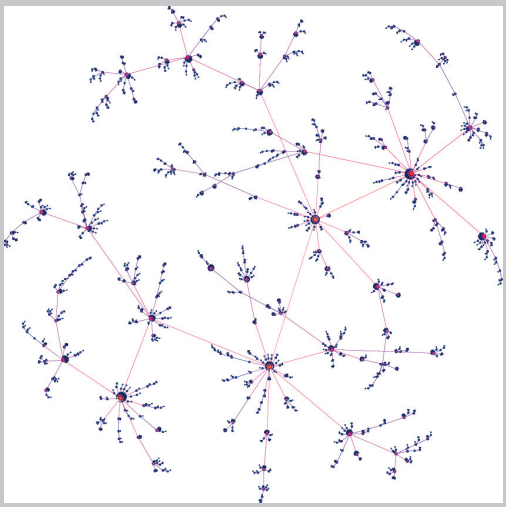
Colección de Estudios de Física 104

Dynamics and collective phenomena of social systems



Carlos Gracia Lázaro

Dynamics and Collective Phenomena of Social Systems

Colección de Estudios de Física
Vol. 104

Esta colección recoge las tesis presentadas en el Departamento de Física de la Materia Condensada de la Universidad de Zaragoza desde su constitución en 1987.

Colección de Estudios de Física **Vol. 104**

Dynamics and Collective Phenomena of Social Systems

Carlos Gracia Lázaro

Prensas de la Universidad
Universidad Zaragoza

GRACIA LÁZARO, Carlos

Dynamics and collective phenomena of social systems / Carlos Gracia Lázaro. — Zaragoza : Prensas de la Universidad de Zaragoza, 2013

XIII, 181 p. ; 25 cm. — (Colección de estudios de física ; 104)

Bibliografía: p. 173-181. — ISBN 978-84-15770-24-4

1. Juegos, Teoría de los–Tesis doctorales. 2. Elección social–Modelos matemáticos–Tesis doctorales

519.83(043.2)

316.6:519.87(043.2)

1.ª edición, 2013

Prensas de la Universidad de Zaragoza. Edificio de Ciencias Geológicas, c/ Pedro Cerbuna, 12, 50009 Zaragoza, España. Tel.: 976 761 330. Fax: 976 761 063

puz@unizar.es http://puz.unizar.es

Impreso en España

Imprime: Servicio de Publicaciones. Universidad de Zaragoza

D.L.: Z-87-2013

Imagine there’s no countries, it isn’t hard to do.
Nothing to kill or die for, and no religion too.

John Lennon.

A María

Publications related to this thesis

Most of the results presented in this thesis have been published in the following papers:

L. M. Floría, C. Gracia-Lázaro, J. Gómez-Gardeñes and Y. Moreno. *Social network reciprocity as a phase transition in evolutionary cooperation*. Physical Review E **79**, 026106 (2009).

C. Gracia-Lázaro, L.F. Lafuerza, L. M. Floría and Y. Moreno. *Residential segregation and cultural dissemination: An Axelrod-Schelling model*. Physical Review E **80**, 046123 (2009).

C. Gracia-Lázaro, L. M. Floría and Y. Moreno. *Selective advantage of tolerant cultural traits in the Axelrod-Schelling model*. Physical Review E **83**, 056103 (2011).

C. Gracia-Lázaro, F. Quijandría, L. Hernández, L. M. Floría and Y. Moreno. *Coevolutionary network approach to cultural dynamics controlled by intolerance*. Physical Review E **84**, 067101 (2011).

C. Gracia-Lázaro, Jose A. Cuesta, Angel Sánchez and Yamir Moreno. *Human behavior in Prisoner's Dilemma experiments suppresses network reciprocity*. Scientific Reports **2**, 325 (2012).

C. Gracia-Lázaro, A. Ferrer, G. Ruíz, A. Tarancón, J. A. Cuesta, A. Sánchez and Y. Moreno. *Heterogeneous networks do not promote cooperation when humans play a Prisoner's Dilemma.* Proceedings of the National Academy of Sciences USA **109**, 12922 (2012).

J. Gómez-Gardeñes, C. Gracia-Lázaro, L. M. Floría and Y. Moreno. *Evolutionary dynamics on interdependent populations*. Physical Review E **109**, 12922. (2012).

Contents

Chapter 1

Introduction.

This thesis focuses on the study of social systems through methods of theoretical physics, in particular proceedings of statistical physics and complex systems, as well as mathematical tools like game theory and complex networks. There already exists predictive and analysis methods to address these problems in sociology, but the contribution of physics provides new perspectives and complementary and powerful tools. This approach is particularly useful in problems involving stochastic aspects and nonlinear dynamics. The contribution of physics to social systems provides not only prediction procedures, but new insights, especially in the study of emergent properties that arise from holistic approaches.

We study social systems by introducing different agent-based models (ABM). When possible, the models are analyzed using mathematical methods of physics, in order to achieve analytical solutions. In addition to a theoretical approach, experimental treatment is performed via computer simulations both through Monte Carlo methods and deterministic or mixed procedures. This working method has proved very fruitful for the study of several open problems.

The book is structured as follows. This introduction presents the mathematical formalisms used in the investigations, which are structured in two parts: in part I we deal with the emergence of cooperation, while in part II we analyze cultural dynamics under the perspective of tolerance.

1.1 Game theory.

In 1944, mathematician John von Neumann and economist Oskar Morgenstern established a definition of game and its components [1]:

"First, one must distinguish between the abstract concept of a game, and the individual plays of that game. The game is simply the totality of the rules which describe it. Every particular instance at which the game is played – in a particular way – from beginning to end, is a play. Second, the corresponding distinction should be made for the moves, which are the component elements of the game. A move is the occasion of a choice between various alternatives, to be made either by one of the players, or by some device subject to chance, under conditions precisely prescribed by the rules of the game. The move is nothing but this abstract 'occasion', with the attendant details of description, – i.e. a component of the game. The specific alternative chosen in a concrete instance – i.e. in a concrete play – is the choice. Thus the moves are related to the choices in the same way as the game is to the play. The game consists of a sequence of moves, and the play of a sequence of choices. Finally, the rules of the game should not be confused with the strategies of the players.[...] Each player selects his strategy – i.e. the general principles governing his choices – freely. [...] The rules of the game, however, are absolute commands. If they are ever infringed, then the whole transaction by definition ceases to be the game described by those rules." **(von Neumann and Morgenstern. *Theory of Games and Economic Behavior* (1944))**.

Game theory (GT) is an area of applied mathematics that uses models to study interactions with formalized incentive structures (*i.e.* games), and is therefore a field closely related to decision theory. The origins of game theory go far back in time: In 1713, James Waldegrave proposed a solution for a two player game [2]. Nevertheless, GT as a specific field did not appear until 1928 through a series of papers published by John von Neumann [3]. Modern game theory was comprehensively formalized in 1944 by John von Neumann and Oskar Morgenstern [1], and experienced a very important step forward with John Nash's contribution of strategic equilibrium: the Nash equilibrium [4]. GT was firstly developed as a tool for understanding economic behavior, but now is applied in many fields such as biology, physics, sociology, traffic control, etcetera; in fact, it applies to a wide variety of agents including humans, microorganism and nonhuman animals. Further refinements to game theory include evolving populations and underlying topologies, among others.

1.1.1 Definitions.

In this section is a brief introduction to game theory related to the focus of this thesis, a deeper study can be found in [5, 6, 7].
A game is a mathematical representation of a conflict situation. The outcome (pay-off) results from mutual interaction between different agents or players. A player is defined as a decision maker: a person, a people group, an animal or whatever kind of element. Based on the number of players $N \geq 2$, games are classified as two-person games, three-person games, and in general as N-person games. The interactions between the players are governed by rules that state the actions each player can take,

the information each player has available and the outcomes of the actions. A strategy defines the actions that a player will follow in every scenario.

There is a distinction between games with *perfect*, *complete* and *incomplete* information. *Perfect* information describes the situation when each player has available the information to determine all of the possible scenarios, strategies, actions and outcomes all game long: so, players have full information about the actions that have already taken place. In games with *complete* information, each player knows the rules of the game and the payoff functions of all the players, but players may not see all of the actions chosen by other players. In *incomplete* information games, players may not know some information about the other players (actions, strategies, payoffs) or about the rules (*e.g.* game's length).

According to the updating, games can be classified as *simultaneous* and *sequential* games. In *simultaneous* games, players choose their actions simultaneously, therefore, players may predict other players' action but don't know it. By extension, a game can be classified as *simultaneous* if decisions are not taken simultaneously but players' actions are in ignorance of others players' actions. On the contrary, in *sequential* games players make decisions in sequential order and later players have some knowledge on actions already taken by earlier players.

Based on the total outcome, there are *zero-sum* games and *non-zero-sum* games. In a *zero-sum* game, a player's gain (or loss) is balanced by the losses (or gains) of the other players(s), *i.e.* the total pay-offs for the players, for every combination of the available actions, sum to zero. Otherwise, in *non-zero-sum* games, total payoff is different from zero. Attending the indiscernibility of players, games can be classified between *symmetric* or *asymmetric* games. In *symmetric* games, payoffs depend only on the actions, not on who is chosen them. Otherwise, the game is *asymmetric*.

1.1.2 Normal form.

A game can be can be represented through different forms, such as extensive and normal form. Normal form is a description of a game by way of a matrix that relates players' actions to payoff functions. In order to have a normal form description of a game, we take in consideration the following data:

i) A set N of players, $N = \{1, 2, \ldots, n\}$.

ii) Each player u has a finite number of actions, represented by the actions set $A_u = \{1, 2, \ldots, m_u\}$.

iii) Each player u has a payoff function associated ($P_u : A_1 \times A_2 \times \ldots \times A_n \to \mathbb{R}$) that provides the payoff of player u.

Definition: A game in normal form is a structure $G = \langle N, \mathbf{A}, \mathbf{F} \rangle$, where $N = \{1, 2, \ldots, n\}$ is a set of players, $\mathbf{A} = \{A_1, A_2, \ldots, A_n\}$ is an n-tuple of actions sets, one for each player, and $\mathbf{F} = \{F_1, F_2, \ldots, F_n\}$ is an n-tuple of payoff functions.

The normal form of a two-person symmetric game is given by a pair (A, P), where A is a nonempty set, the set of actions, and P is a real-valued function defined on $A \times A$, (*i.e.*, $P(ij) \in \mathbb{R},\ \forall\ i, j \in A$.). The *payoff matrix* $\mathbf{P}$, defined as $P_{ij} = P(i, j)$, represents the payoff of player I, given the actions chosen by players I and II are i and j respectively. Note that $\mathbf{P}$ is a $m \times m$ matrix, where m is the number of possible actions. Inasmuch as the game is symmetric, player II's payoff is given by $P(j, i)$. In particular, a two-player m-action symmetric game is defined by the matrix:

$$\mathbf{P} = \begin{pmatrix} P_{11} & P_{12} & \ldots & P_{1m} \\ P_{21} & P_{22} & \ldots & P_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ P_{m1} & P_{m2} & \ldots & P_{mm} \end{pmatrix} = \begin{pmatrix} P(1,1) & P(1,2) & \ldots & P(1,m) \\ P(2,1) & P(2,2) & \ldots & P(2,m) \\ \vdots & \vdots & \ddots & \vdots \\ P(m,1) & P(m,1) & \ldots & P(m,m) \end{pmatrix} \tag{1.1}$$

1.1.3 Pure and mixed strategies.

A player's strategy determine the action the player will take at any stage of the game. A strategy profile or strategy combination is a set of strategies for each player which fully specifies all actions in a game. A strategy profile consists of one and only one strategy for every player. Then, strategy and move are different concepts: A move is an action taken by a player at some point of the game. On the other hand, a strategy is a player's algorithm that relates every scenario to the player's actions [8, 9]. Although the terms action and strategy represent different concepts, sometimes have been used interchangeably, especially for the last twenty years. This is due to the use of repeated one-round games in evolutionary dynamics, where the players action change is sometimes studied according to updating rules. In this thesis we try to recover the original meaning of such terms; nevertheless, in chapter 2 we use the notation *strategy, updating rule*. Note that original notation provides more levels: *move, action, strategy, updating rule*.

A pure strategy R determines the move a player will make for any scenario. A player's strategy set $\{R_1, R_2, \ldots, R_m\}$ is the set of pure strategies available to that player. The convex linear combination set S of pure strategies

$$S \;=\; \{\mathbf{p} = (p_1, \ldots, p_m) \;\in \mathbb{R}^m \;:\; p_i \geq 0,\; \sum_{i=1}^{m} p_i = 1\} \tag{1.2}$$

is the set of mixed strategies. Therefore, a mixed strategy **p** is an assignment of a probability p_i to each pure strategy R_i.

1.1.4 Minimax rule.

Minimax is a decision algorithm for minimize losses under the maximum loss scenario; similarly, maximin rule consists in maximizing the minimum payoff. John von Neumann proved the minimax theorem in 1928 [3]: He stated that in every two-person zero-sum game with finitely many pure strategies, there exists a value V and a mixed strategy $\boldsymbol{p}^j$ for each player j, such that:

i) Given player 2's strategy $\boldsymbol{p}^2$, the best payoff possible for player 1 is V, and
ii) Given player 1's strategy $\boldsymbol{p}^1$, the best payoff possible for player 2 is $-V$.

Subsequently, it has been extended to other types of games. A generalization of Neumann's minimax theorem is the Sion's minimax theorem that states [10]:

Let X be a compact convex subset of a linear topological space and Y a convex subset of a linear topological space. Let f be a real-valued function on $X \times Y$ such that $f(x, \cdot)$ is upper semicontinuous and quasiconcave on Y, $\forall x \in X$, and $f(\cdot, y)$ is lower semicontinuous and quasi-convex on X, $\forall y \in Y$. Then:

$$\min_{x \in X} \sup_{y \in Y} f(x, y) \;=\; \sup_{y \in Y} \min_{x \in X} f(x, y). \tag{1.3}$$

In two-person zero-sum games, the minimax algorithm can be summarized as choosing the best move for yourself (higher payoff) assuming that your opponent will choose the worse for you (lower payoff). In an iterated game, minimax method implies considering all possible moves for all players and rounds.

1.1.5 Dominant strategy. Nash equilibrium. Stability.

A strategy is a strict *(resp. weak)* dominant strategy if it provides greater *(resp. greater or equal)* payoff to a given player than any other strategy, no matter other

players' strategies. Let *B* be the set of strategies for which the function $\mathbf{v} \rightarrow \mathbf{v}P\mathbf{w}$ gets its maximum value; then *B* is called set of best responses to **w**.

A set of strategies is a *Nash equilibrium* if no player can get greater payoff by unilaterally changing its strategy. Let S_i be the strategy set for player i, $S = S_1 \times S_2 \times ... \times S_n$ be the set of strategy profiles and $f_i(x)$ be the *player-i*'s payoff function for $x \in S$. Let x_i be a strategy profile of player i and x_{-i} be a strategy profile of all players except for player i. A strategy profile $x^* \in S$ is a weak Nash equilibrium if:

$$f_i(x_i^*, x_{-i}^*) \geq f_i(x_i, x_{-i}^*),\ \forall i, x_i \in S_i, x_i \neq x_i^* \quad . \tag{1.4}$$

In the same way, a strategy profile $x^* \in S$ is a strict Nash equilibrium if:

$$f_i(x_i^*, x_{-i}^*) > f_i(x_i, x_{-i}^*),\ \forall i, x_i \in S_i, x_i \neq x_i^* \quad . \tag{1.5}$$

A game is finite if the number of players and the number of pure strategies each player has are both finite. Nash proved that, taking into account mixed strategies, at least one Nash equilibrium exists for all finite games [4].

There is a intermediate equilibrium between strict and weak Nash equilibria. A Nash equilibrium is *stable* if an infinitesimal change in probabilities for any player u implies: *i)* Any other player $v \neq u$ do not gets a higher payoff, *ii)* Player u gets a lower payoff.

1.1.6 Evolutionary game theory.

Evolutionary game theory (EGT) is defined as the application of game theory to evolving populations, providing a useful framework to model Darwinian competition. The origins of EGT can be found in John Maynard Smith and George R. Price's study about the way in which animal conflict can be modeled through survival strategies in hostile scenarios [11]. Evolutionary game theory has been successfully used not only in many aspects of biology (*e.g.* the basis of altruistic behaviors, the emergence of multicellular organisms, group selection, sexual selection, parental care, coevolution or ecological dynamics), but also in other sciences such as economics or sociology [12, 13, 14, 15, 16].

Replicator equations.

The usual method for studying evolutionary dynamics in GT is through replicator equations. These replicator equations establish a relation between the growth rate of the proportion of players using a certain strategy (that can be interpreted, for example, as a specie's population growth or a behavior evolution) and the difference between the average payoff of that strategy and the average payoff of the whole population (respectively the specie's fitness or the behavior reward) [12, 13]. Replicator equation assume infinite and well-mixed populations and continuous time. Usually, the interest is not in the transient but in the steady-state solutions: the stable states.

The replicator equation can be obtained from Darwinian arguments. Once postulated that expected offspring of a kind (mutation, race, specie, ...) is proportional to the fitness (that provides food, welfare, safety, ...), the growth ratio $\dot{x_i}/x_i$ of that kind i can be assumed as the difference between kind's fitness and mean population fitness, that is, $\dot{x_i}/x_i = f_i(\mathbf{x}) - \bar{f}(\mathbf{x})$. Assuming continuous time, the evolution of a population distributed in n (finite) kinds is given by:

$$\dot{x_i} = x_i[f_i(x) - \mu(x)], \quad \mu(x) = \sum_{i=1}^{n} x_i f_i(x) \ , \tag{1.6}$$

where x_i ($i = 1, 2, \ldots, n$) is the ratio of kind i in the population, $f_i(x)$ is the fitness of type i, and $\mu(x)$ is the average population fitness. Since $\sum_i x_i = 1$, the population vector $\mathbf{x} = (x_1, \ldots, x_n)$ evolves in the (n-1)-simplex defined by the n vertices $x_i = \delta_{ij},\ j = 1, \ldots, n$. Eq. 1.6 is the most general form of replicator equation.

Under the assumption that kind's fitness is a linear function of population ratio, the replicator equation can be written as:

$$\dot{x_i} = x_i \left((\mathbf{Px})_i - \mathbf{x^T Px} \right), \tag{1.7}$$

where $\mathbf{P}$ is the payoff matrix and contain the fitness information, $(\mathbf{Px})_i$ represents the expected payoff of i-kind and $\mathbf{x^T Px}$ stands for the whole population's mean payoff.

Replicator dynamics.

An alternative way to get the replicator equation, often used in social sciences, is based on the concept of imitation, rather than offspring, most related to biology

[15, 17, 18]. Consider a iterated two-player n-action symmetric game and a population of N players. Let be t the round number, $i = 1, 2, \ldots, n$ the actions, x_i^t the proportion of population choosing action i and $P_i^t = P_i(x^t)$ is the expected payoff of a player choosing i at time (round) t. The evolutionary dynamics can be model in the following way (replicator dynamics): assume that each time lapse dt, a randomly and equiprobably player u (the past action of u was i) chooses a random player v (the past action of v was j) and u changes its action to j, if the payoff P_j^t of v is greater than his payoff P_i^t, with a probability proportional to the payoff difference $P_j^t - P_i^t$. That is, once two players with past actions i and j have been chosen, the first player's change probability is given by:

$$\Pi_{i\to j}^t = \beta(P_j^t - P_i^t)\Theta(P_j^t - P_i^t) \ , \tag{1.8}$$

where $\Theta(y)$ is the Heaviside function ($\Theta(y) = 1$ if $y > 0$ and 0 otherwise).

If the actions are ordered (without loss of generality), such that $P_1^t \leq P_2^t \leq \ldots \leq P_n^t$, the expected ratio of players choosing action i at time $t + dt$ will be given by:

$$\begin{aligned} \bar{x}_i^{t+dt} &= x_i^t - \frac{1}{N}x_i^t(dt)\sum_{j=i+1}^{n} x_j^t\beta(P_j^t - P_i^t) + \sum_{j=1}^{i}\frac{1}{N}x_i^t x_j^t(dt)\beta(P_i^t - P_j^t) \\ &= x_i^t + \frac{x_i^t\beta(dt)}{N}(P_i^t - \bar{P}^t) \ , \end{aligned} \tag{1.9}$$

where $\bar{P}^t = \sum_j P_j^t x_j^t$ is the whole population mean payoff. Under the assumption of large population size, we can replace $\bar{x}_i^{t+dt}$ by x_i^{t+dt}, getting:

$$\dot{x}_i^t = \frac{\beta}{N}x_i^t(P_i^t - \bar{P}^t) \ . \tag{1.10}$$

and making $\beta = N$ (time rescaling), we get the replicator equation:

$$\dot{x}_i = x_i(P_i - \bar{P}) \ . \tag{1.11}$$

The strategy above described, **replicator dynamics**, has been used extensively to describe a large variety of problems. The main advantage is that it can be applied in

finite and non-well-mixed populations but, unfortunately, the mean-field description is not valid anymore and those problems must be numerically solved, usually through extensive numerical simulations. In fact, the first part of this thesis is focused in such kind of problems.

Other update rules.

A necessary ingredient of evolutionary game models is the way in which agents choose their actions. In order to frame the thesis' subject of study, this description is focused on repeated two-player symmetric games. In simultaneous games (*i.e.*, players choose their actions simultaneously), at a given time step, every player plays the game with all her neighbors, usually using the same strategy in all pairings. If there is an underlying topology, the neighbors of each player are given by the network(s) of the model [16]. Once all the games are played each agent collects the total payoff. Subsequently, players decide the action they will take in the next round: This decision constitutes the strategy (update rule). Besides the aforementioned **replicator dynamics**, some of the most used strategies are:

Unconditional imitation: A randomly chosen player i compares its payoff P_i with its neighbor with the largest payoff, say player j. If $P_j > P_j$ player i will imitate in the next round the last action taken by j. Otherwise, player i will repeat action [19].

Moran rule: A randomly chosen player i chooses one of its neighbors j proportionally to its payoff P_j. In the next round agent i will chose the last action taken by j [20].

Fermi rule: A randomly chosen player i compares its payoff P_i with a random neighbor j. If $P_j > P_i$ player i will imitate in the next round the last action taken by j with a probability proportional to:

$$\Pi_{i\to j} = \frac{1}{1+e^{-\beta(P_j-P_i)}} \quad . \tag{1.12}$$

Otherwise, player i will repeat action [21].

Best-response: The best response is the strategy consisting on choosing, for the next round, the action which would have produced the higher payoff in the past round, once each player knows the chosen action by its neighbors. [22].

In some kind of problems, instead of considering that the strategies depend directly on the payoff, it is more realistic to consider the fitness as a function of payoff. For instance, **weak selection** describes situations in which the effects of payoff

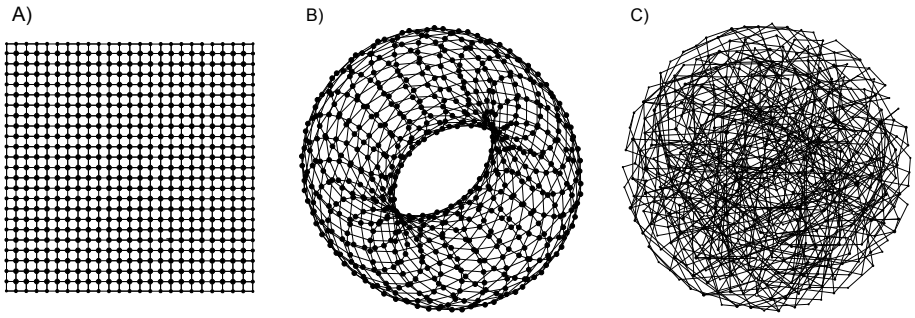


Figure 1.1: *Examples of graphs.* Although the three graphs have a similar number of nodes, only the one on the right can be considered a complex network. A) Regular lattice of 25x25 (625 nodes). B) Regular lattice of 25x25 (625 nodes) with periodical boundary conditions. C) Complex network consisting of 700 nodes and a heterogeneous connectivity according to a binomial distribution. The diameter of each node is proportional to its connectivity.

differences are small. Weak selection has been extensively studied in evolutionary biology, and recently in evolutionary game dynamics [24]. Nevertheless, not all the usual strategies take into consideration the payoffs: **Majority rule** consists on the imitation of most common action in the neighborhood, and **analysis players' strategies**, described in chapters 5 and 6, consists on the imitation of neighborhood actions with a probability proportional to their frequency [23].

1.2 Complex networks.

The study of the relations among elements of different systems unveil underlying networks: Regardless of its origins, many networks of different areas can be characterized through common schemes, showing similar properties [25, 26, 27, 28]. Examples of this can be found in biology (*e.g.*, regulatory, metabolic, signaling or neuronal networks), sociology (*e.g.*, scientific colaborations, coworking relations or information exchange networks) and technology (*e.g.*, internet). The topology of the interaction network may provide the key to understanding many complex systems and, in fact, complex networks have become a new paradigm for complexity [29].

1.2.1 Definitions.

A graph $G = \{N, L\}$ consists of two sets N and L, such that N is not empty and L is a set of couples of elements of N. If L is ordered (unordered), then the graph is directed (undirected). The elements n_i of N are called vertices or nodes and the elements l_i of L are the links or edges. Let N and K be the number of elements in

N and L, respectively. We can refer to a node by its order number i, then a link l consisting of couple (i, j) can be referred by l_{ij} or by l_k, where k is the order number of the link l in set L. The link l_{ij} is said to be incident in nodes i and j, or to join i and j. Two nodes (i,j) are adjacent, connected or neighbors if there exists a link l_{ij} incident in nodes i and j. $G' = \{N', L'\}$ is a subgraph of $G = \{N, L\}$ if $N' \subset N$ and $L' \subset L$. In order to get a matricial representation of graph G, we consider the adjacency matrix A: a NxN square matrix with components a_{ij} such that $a_{ij} = 1$ *if and only if* the link l_{ij} exists, otherwise $a_{ij} = 0$. The components a_{ii} of the diagonal of the adjacency matrix satisfy $a_{ii} = 0$ and it is a symmetric matrix $a_{ij} = a_{ji}$ for undirected graphs. An alternative matricial representation is given by the incidence matrix B: a NxK matrix with components b_{ik} such that $b_{ik} = 1$ *if and only if* $l_k = l_{ij}$ for a node j, otherwise $b_{ik} = 0$.

1.2.2 Properties.

Let $G = \{N, L\}$ be an undirected graph. The connectivity k_i or degree of a node $i \in G$ is defined as:

$$k_i = \sum_{j \in L} a_{ij} \ , \tag{1.13}$$

and represents the number of incident links in i. In the same way, if $G' = \{N', L'\}$ is a directed graph, the out-degree and in-degree of a node $i \in G'$ are defined as:

$$\begin{aligned} k_i^{out} &= \sum_{j \in L'} a_{ij} \ , \\ k_i^{in} &= \sum_{j \in L'} a_{ji} \ , \end{aligned} \tag{1.14}$$

and represent the number of outgoing and ingoing links respectively, and the total degree of the node i in a directed graph is $k_i = k_i^{out} + k_i^{in}$. In order to characterize the topology of an undirected graph, the degree distribution $P(k) = P_k$ is defined for $k = 0, 1, 2, \ldots$ as the fraction of nodes with connectivity k in the graph. Likewise, for directed graphs $P^{out}(k), P^{in}(k)$ are defined in the same way. As well, the n-moments of $P(k)$, defined as $\langle k^n \rangle = \sum_k k^n P(k)$, give us information about networks' topology and therefore about their behavior under the dynamical processes that we study in this work. To deal the degree correlations, $P(k'|k)$ is defined as the likelihood that a given link connecting a degree k node i is connected to a node j of degree k', and satisfies the normalization equation $\sum_{k'} P(k'|k) = 1$ and the detailed balance condition $kP(k'|k)P(k) = k'P(k|k')P(k')$. Nevertheless, in finite size real

networks we can obtain clearer results computing $k_{nn}(k)$, that is, the average degree of the neighbors of degree k nodes:

$$k_{nn}(k) \equiv \sum_{i=1}^{N}\sum_{j=1}^{N} a_{ij}\frac{k_j}{k_i} = \sum_{k'} k' P(k'|k) \quad . \tag{1.15}$$

Depending on $k_{nn}(k)$, the networks can be uncorrelated, assortatives ($k_{nn}(k)$ increases with k) or dissortatives ($k_{nn}(k)$ decreases with k).

The shortest path d_{ij} or geodesic between two nodes (i,j) is the minimum number of links required to connect i and j. In order to characterize the size of a network G, beside the values of N (number of nodes) and L (number of links), the diameter Diam(G) of G is defined as the maximum value of the shortest paths in G. For fully connected graphs, the characteristic path length L is defined as:

$$L \;=\; \frac{1}{N(N-i)}\sum_{i=1}^{N}\sum_{j=1,\ j\neq i}^{N} d_{ij} \quad , \tag{1.16}$$

as it diverges for disconnected graphs, can be useful the efficiency:

$$E \;=\; \frac{1}{N(N-i)}\sum_{i=1}^{N}\sum_{j=1,\ j\neq i}^{N} \frac{1}{d_{ij}} \quad . \tag{1.17}$$

With the purpose of characterizing the connectedness of a node i, together with the degree, the closeness c_i of i is defined as the inverse of the average distance from i to all nodes $c_i = N(\sum_{j=1,j\neq i}^{N} d_{ij})^{-1}$. Besides, the betweenness of i is defined as the average fraction of shortest paths passing through i ($n_{jk}(i)$) over total the shortest paths (n_{jk}) for every pair of nodes:

$$b_i \;=\; \sum_{j=1}^{N}\sum_{k=1,\ k\neq j}^{N} \frac{n_{jk}(i)}{n_{jk}} \quad . \tag{1.18}$$

Another typical property of complex networks is clustering, it characterizes the likelihood that two nodes with a common neighbor are connected. One of the many measures of clustering is transitivity T, which is defined as the normalized ratio of

the number of transitive (fully-connected) triples divided by the amount of connected (fully or simply-connected) triples.

$$T = \frac{3(\#transitive\ triples\ in\ G)}{\#connected\ triples\ in\ G} \quad . \tag{1.19}$$

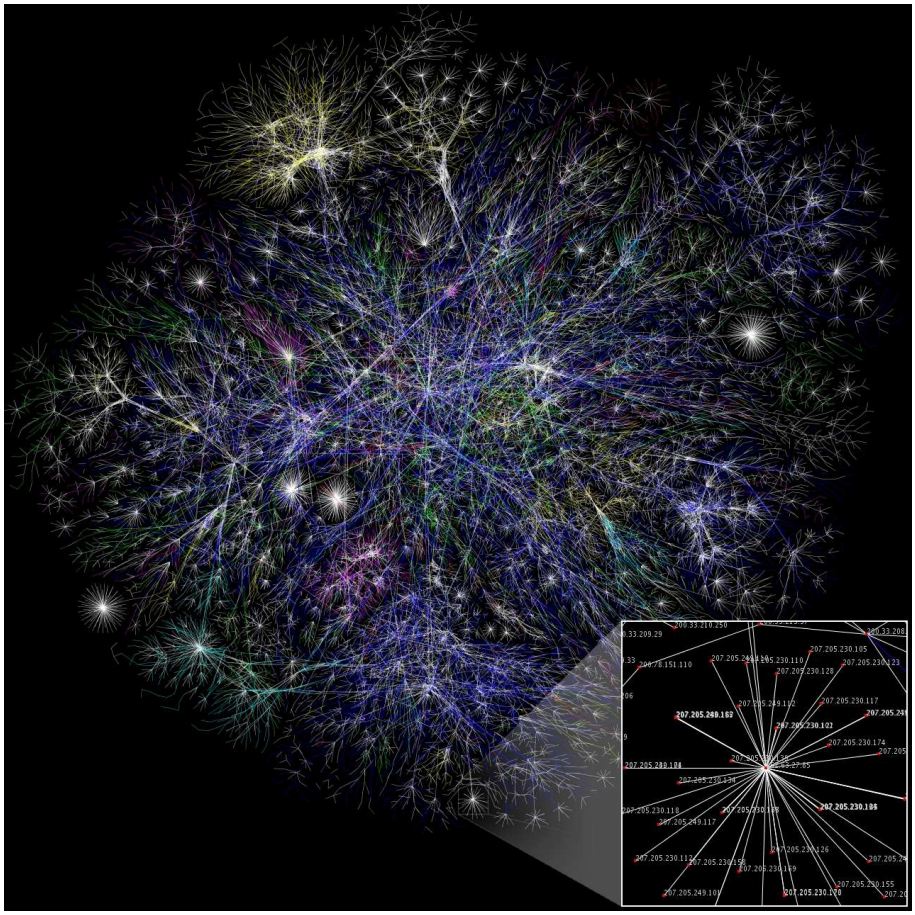
When a subgraph $G' = \{N', L'\}$ of $G = \{N, L\}$ has higher connectivity than G, *i.e.* the nodes of G' are tightly connected, G' is called a cluster. More precise definitions of cluster are the n-clique and the k-plex. A clique or 1-clique is a fully-connected subgraph G' of G. More generally, a n-clique is a subgraph G' such that the largest geodesic d_{ij} between two nodes (i,j) of G' is n. On the other hand, a subgraph $G' = \{N', L'\}$ is a k-plex if it is a maximal subgraph such that $k'_i \geq N' - k, \ \ \forall i \in N'$, that is, each node of G' has at least $N' - k$ neighbors in G' and there is not another graph in G containing G' that satisfy this property.

1.2.3 Weighted networks.

So far, we have considered unweighted networks, which means that the links between nodes are either present or not, without an assigned value. However, many real networks exhibit heterogeneity in the links. This feature has been studied in many fields such as social networks [33, 34, 35, 36, 37], metabolic networks [38, 39], predator-prey interactions [40, 41], neural networks [35, 44], traffic of the passengers in airline networks [42, 43], internet traffic [60], etcetera. Weighted networks (*i.e.* networks in which each link is characterized by a value) provide a very useful tool to describe these systems.

A weighted network $G^W = \{N, L, W\}$ consists of three sets N, L and W, such that $N = \{n_1, n_2, ..., n_N\}$ is a not empty set of N nodes, $L = \{l_1, l_2, ..., l_K\}$ is a set of K couples of elements of N (the links) and $W = \{w_1, w_2, ..., w_K\}$ is a set of real numbers (weights) associated to the links. A Weighted network G^W can be described by a $N \times N$ matrix W, the weights matrix , such that its component w_{ij} is the weight of the link from node i to node j, assumed that $w_{ij} = 0$ if such connection does not exist. The network is symmetric if $w_{ij} = w_{ji}, \ \forall i, j$.

In a weighted network G^W, the weight distribution $Q(w)$ is defined as the probability for a given edge to have weight w. The node weight (or strength, or weighted connectivity) is defined as $s_i = \sum_{j \in N} w_{ij}$. If there are not correlations between weights and connectivity, one obtains $s|k \sim \langle w \rangle k$. In the same way, measure coefficients of unweighted networks (such as shortest paths, clustering coefficient, etcetera) can be generalized to weighted networks [28].



Example of real network. Partial map of the Internet: Each node represents an IP address while each line symbolizes a link between two nodes, the length of the lines indicate the delay. This graph represents less than 30% of the Class C networks reachable by the data collection program. Different colors represent different allocations. *The Opte Project (2005).*

1.2.4 Real networks: distributions and topologies.

Complex networks can represent a huge range of real systems consisting on many highly connected elements that can be found in different fields, such as sociology, biology or technology. Although the use of complex network theory makes up an approximation that implies loss of information, it provides a holistic approach and details about emergent phenomena. Different sorts of systems are characterized by different kinds of networks, with dissimilar properties as degree distributions, path lengths, clustering, degree correlations, etc.

Regular graphs.

In graph theory, a regular graph (RG) is a graph where each vertex has the same degree connectivity. In addition, in a regular directed graph the indegree and outdegree of each vertex are equal to each other. A RG with vertices of degree k is called a k-regular graph. A network is a regular graph *if and only if* the vector $u = (1, \ldots, 1)$ is an eigenvector of its adjacency matrix A_{ij}. In addition, the eigenvalue of u is the constant degree k of the graph. Eigenvectors v corresponding to other eigenvalues are orthogonal to u, for such eigenvectors $v = (v_1, \ldots, v_n)$, we have $\sum_{i=1}^{n} v_i = 0$ (see, *e.g.* [78]).

RG of degrees $k = 0, 1, 2$ are trivial, but higher degree RG can be complex networks. In some parts of this work (*e.g.* chapter 2) we use a kind of RG, called random regular graph (RRG), characterized by a random distribution of links. Starting from a regular lattice, a RRG can be generated by randomization of links through a rewiring process.

Random graphs.

Erdös and Rényi initiated in 1959 the study of graphs that grow through random procedures [76]. The original Erdös-Rényi graph $G_{N,K}^{ER}$ consists of a set of N nodes, firstly disconnected, and later linked by connecting K pairs of nodes at random. In the same way, $G_{N,p}^{ER}$ represents a graph generated though a set of N nodes, firstly disconnected, and later linked by connecting each pair of nodes with probability p (figure 1.2). Although both processes generate different (but similar) kinds of network, for large values of N they provide the same distribution. The graph $G_{N,p}^{ER}$ has on average $\binom{n}{2}p$ edges. The degree distribution of a node i is binomial:

$$P(k_i = k) = \binom{n-1}{k} p^k (1-p)^{n-1-k} \quad . \tag{1.20}$$

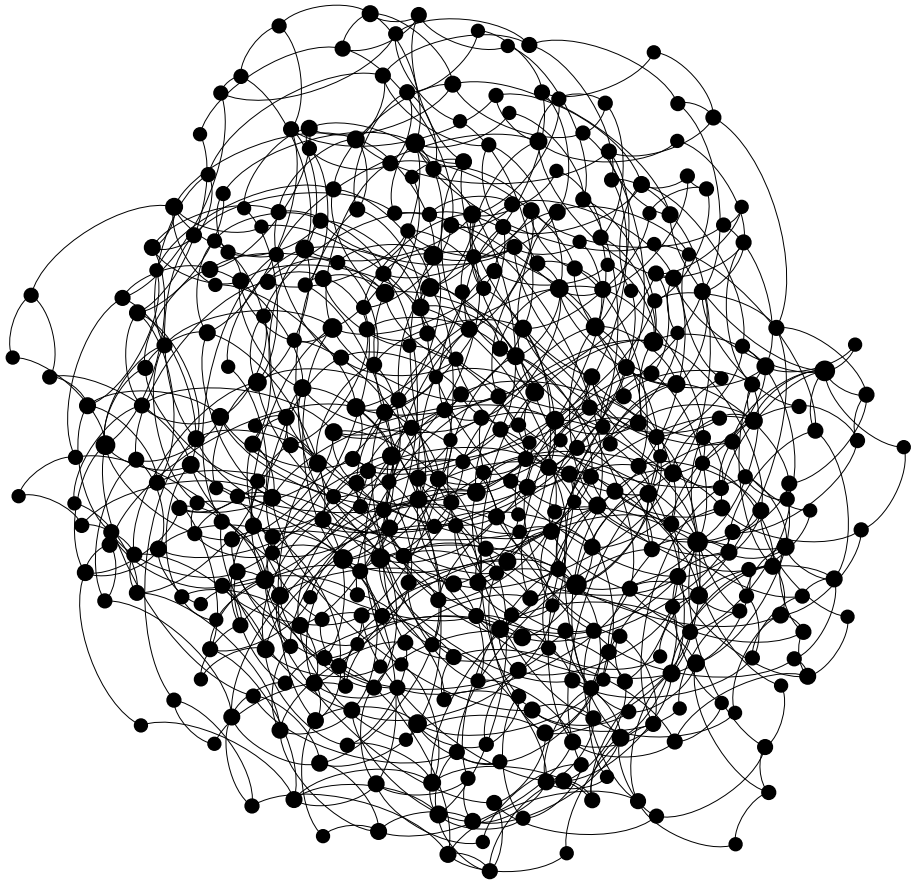


Figure 1.2: Random network according to a Poisson distribution of the connectivity. The diameter of each node is proportional to its connectivity. Network size $N = 800$, averaged degree $\langle k \rangle = 4$

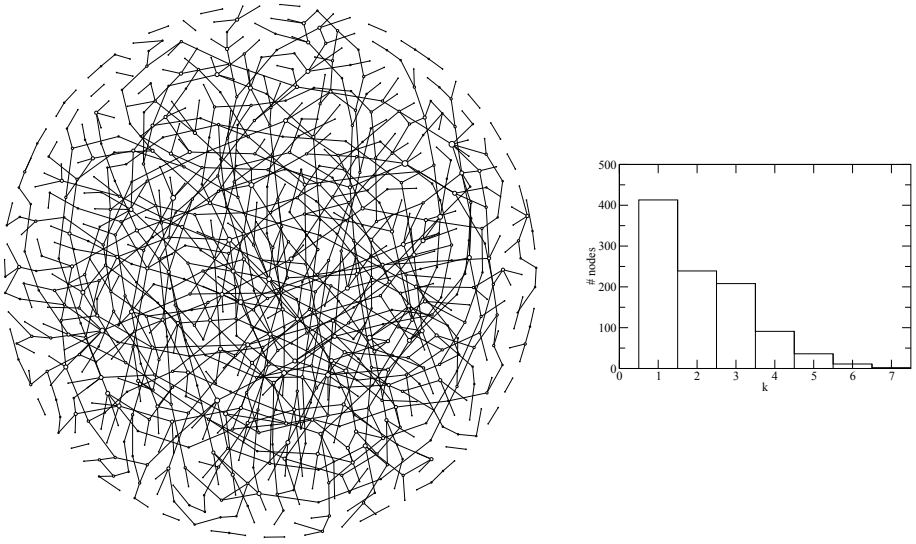


Figure 1.3: Random network generated through Erdös-Rényi mechanism (left) and its connectivity histogram (right). The diameter of each node is proportional to its connectivity. Network size $N = 1000$, averaged degree $\langle k \rangle = 2$

In [77], Erdös and Rényi showed that the topology of $G_{N,p}^{ER}$ depends mainly on p, in fact:

- If $p < 1/N$, then the size S_{max} of the greatest connected component of graph $G_{N,p}^{ER}$ will be $S_{max} \lesssim \ln N$, with a probability increasing with N.
- If $p = 1/N$, then almost surely $S_{max} \simeq N^{2/3}$.
- If $p > 1/N$, then $G_{N,p}^{ER}$ will almost surely have a giant component and no other component will contain more than $O(\ln N)$ nodes.

Small-world networks

Many real networks, such as social networks [62, 63, 64], show the property that, although the mean connectivity is much low that the system size, most nodes can be reached from every other by a small number of steps. A small-world (SW) network is a network where the characteristic path length grows proportionally to the logarithm of the network size:

$$L \propto \log N \quad . \tag{1.21}$$

The Watts and Strogatz model consists of an algorithm to make graphs that provides small-world networks with a high clustering coefficient [25]. Starting from a ring of N nodes (that is, a circle of nodes in which each node is connected to its $2m$ nearest nodes), a rewiring process is performed, so that with probability p each node is disconnected from its clockwise neighbor and connected to a random node. On one extreme, for $p = 0$ the ring remains unchanged and we have a regular lattice; at the other extreme, for $p = 1$ the procedure provides a random graph with minimum connectivity $k_{min} = m$.

According to the SW definition, most pairs of nodes will be connected by at least one short path. Furthermore, from the high clustering coefficient property follows that SW networks contain a high number of cliques and many highly connected subgraphs. Real SW networks usually have high degree nodes that serve as hubs of short paths, and the degree distribution are fat-tailed. Anyway, very different graphs (both real and artificial networks) can be defined as small-world networks as long as they satisfy the property 1.21.

Scale-free networks

The degree distribution of many real networks follows a power-law, at least asymptotically. That is, P(k) goes for large values of k as

$$P(k) \sim k^{-\gamma} \ , \tag{1.22}$$

where the value of the constant γ is usually $2 < \gamma < 3$. Power-laws are the only functions $f(x)$ that presents scale invariance: do not change if independent variable x is multiplied by a common factor, apart from a dilatation. In reference to this property, such networks are called scale-free networks. Many networks belonging to a wide range of subjects appear to be scale-free: Social networks (by example the collaboration networks, as the collaboration of movie actors in films or the coauthors relationships), biological networks as the protein-protein interaction networks, sexual relations in humans (related with the diffusion of sexually transmitted diseases), semantic networks, many technological networks as the World Wide Web, etcetera. Although the average degree $\langle k \rangle$ is bounded, the variance

$$\sigma^2 = < k^2 > - < k >^2 \ , \tag{1.23}$$

diverges as the second moment $< k^2 >$:

$$< k^2 > = \int_{k_{min}}^{k_{max}} k^2 P(k) \sim k_{max}^{3-\gamma} = k_{max}^a \ , \tag{1.24}$$

where the exponent $a > 0$.

The Barabási-Albert (BA) model

In 1999, A. L. Barabási and R. Albert mapped of a portion of the Web, and show that some nodes (hubs), had a connectivity degree very higher than the rest ones, and that the degree distribution of the whole network follows a power-law distribution [45]. They found that other social and biological networks also had similar properties. Barabási and Albert proposed a mechanism, the preferential attachment, to explain the emergence of the power-law distribution. However, this mechanism only produces a specific kind of scale-free networks, and many other mechanisms have been discovered since (and earlier). The preferential attachment mechanism generate a graph $G^{BA}_{N,K}$ according to the next rules: starting from m_0 isolated nodes, a new node j with $m < m_0$ links is added. The likelihood that j will connect to a given node i is proportional to the i-degree:

$$P(j \to i) = \frac{k_i}{\sum_l k_l} \quad . \tag{1.25}$$

The operation of addition a node is repeated $N - m_0$ times. At the end, the graph obtained, known as Barabási-Albert (BA) network, will have N nodes and $K \simeq mN$ links, with $< k >= 2m$ (figure 1.4).

The BA model has been solved in the mean-field approximation [45]: In the thermodynamic limit $t \to \infty$, the degree distribution obtained is $P(k) \sim k^{-\gamma}$, with $\gamma = 3$. For the same value of N and K, BA graphs have smaller average distance than ER graphs, resulting $L \sim logN/log(logN)$ [46]. Furthermore, the clustering coefficient decreases with respect to the size of the system as $C \sim N^{-0.75}$, that is, slower than that observed for ER graphs $C \sim 1/N$. Several variations of the model have been studied, such as directed graphs [47] or alternative mechanisms for preferential attachment [48], among many others.

Configuration model

Starting from a set of N nodes and a given connectivity vector V (that is, V is the connectivities sequence $k_1, k_2, \ldots, k_N$, and k_i the connectivity of link i), the configuration model (CM) provides an algorithm to build up a network [49, 50]. The CM takes the set $G^{conf}_{N,D}$ of all graphs with N nodes and connectivity vector V, and consider all the elements of $G^{conf}_{N,D}$ with equal probability. The algorithm consists of the following steps: firstly assigns k_i half-edges to each node i, after this connects by pairs the half-edges in an equiprobable way. A giant component emerges almost surely when $Q = \sum_k k(k-2)P(k) > 0$ and the maximum degree k_{max} is not too large [49] (figure 1.5).

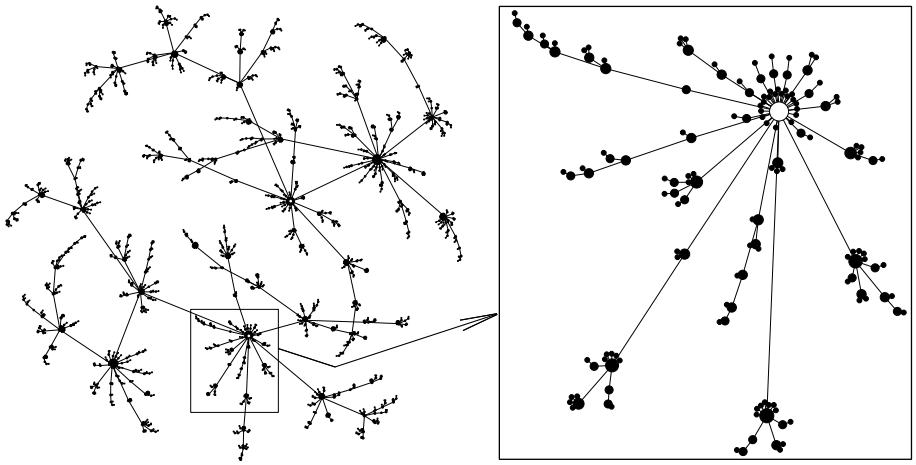


Figure 1.4: SF network generated through Barabási-Albert mechanism. Red (blue) colored circles represent the high (low) connectivity nodes. Network size $N = 2000$, averaged degree $\langle k \rangle = 2$. The subgraph (detail) highlights the scale invariance property of power-law distributions.

The configuration model proposed by Molloy and Reed provides networks with degree correlations, in the sense that the expected degree of the neighbors of a given node i is not independent of k_i. Starting from the CM, *Catanzaro, Boguña, and Pastor-Satorras* proposed the uncorrelated configuration model (UCM), capable to generate random uncorrelated scale-free networks [51]. The model adds a restriction on the maximum possible degree of the vertices, governed by the structural cutoff $k_{max} \simeq \sqrt{kN}$.

1.3 Sociophysics.

The use of methods of probability theory and statistics for dealing with large populations in solving physical problems constitutes Statistical Physics as a branch of physics, and its procedures are successfully used in a wide variety of fields of physics involving many interacting entities. In the light of its usefulness, this successful framework has recently been extended to other sciences including chemistry, biology, neurology, and even some social sciences, such as economics and sociology.

Sociophysics is a multidisciplinary research field that applies theories and methods originally developed by physicists in order to study social topics, usually those including uncertainty or stochastic processes and nonlinear dynamics [142]. In Statistical Physics, the elementary components of studied systems usually are simple objects whose behavior is conditioned by some well-know laws: the statistical study

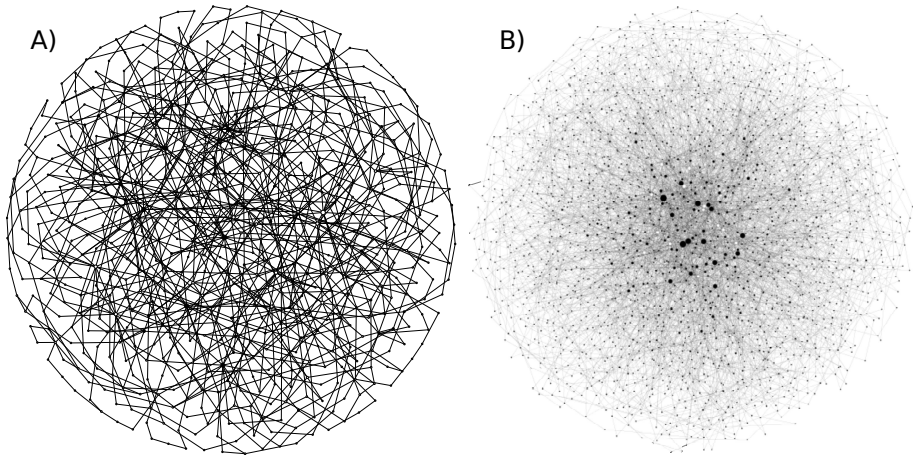


Figure 1.5: *Examples of configuration model graphs:* The C.M. algorithm allows to make a network according to a given connectivity vector. Both graphs above consist of N nodes and a heterogeneous connectivity conforming to a truncated power-law distribution $P(k) = k^{-2.7}$. A)$N = 600,\ k_{min} = 2,\ k_{max} = 15$. B)$N = 2000,\ k_{min} = 2,\ k_{max} = 44$. The diameter of each node is proportional to its connectivity.

focuses on collective effects due to the interactions of a large number of elements. In contrast, the basic constituents of social phenomena are humans, and each individual interacts with a variable number of individuals that, for most individuals and problems, is negligible compared to the system size. People are not single objects following simple behavior rules, and modeling of social topics involves a great simplification of reality. Clearly, this approach does not try to model the complexity of individuals, but focus on the nature of interactions, seeking an holistic approach and drawing conclusions about the overall system. These conclusions include transitions from order to disorder phase, transient and stationary states, scaling properties, and so on. Obviously, the limitations in the modeling of agents (people) condition the validity of the results, and should be taken with caution. However, in most collective systems, global properties do not depend on microscopic configuration but on global features as symmetries, conservation laws, temperature, external fields, noise, etc. Following this holistic view, the modeling of social systems includes only the most important characteristics of individuals, nevertheless criterion common to all scientific modeling process.

A high number of scientists and multidisciplinary work teams are involved in this branch of knowledge and the topics have been dealt in recent years with this approach are so varied as opinion dynamics, ethnic segregation, cultural dynamics, crowd behavior, social hierarchies, language dynamics and spreading phenomena among others.

1.3.1 Agent based models.

An agent-based model (ABM) is a class of numerical models for simulating the interactions of autonomous agents with the purpose of studying their effects on the system as a whole. Usually ABMs rely Monte Carlo methods to introduce randomness and the high numbers of interactions is performed trough specific computer programs. The origin of agent based (AB) modeling dates back to the late 1940s when the mathematician John von Neumann (starting from the Alan Turing idea) designed the von Neumann machine, a theoretical system based on self replication. The concept was then improved by Stanislaw Ulam: Ulam suggested to build the machine as a set of cells on a grid [30]. The concept was taken up by von Neumann, who created the first of the devices later termed cellular automata [31]. The development of computers led to AB modeling widespread since the 1990s. ABMs have been used to deal with a wide range of problems in several fields as biology (as *e.g.*, spread of epidemics, population dynamics), biomedical applications, economics, dynamics of ancient civilizations, logistics, traffic control, workforce management, distributed computing, people's migrations, language dynamics and social network effects.

One of the firsts ABM designed to explore a social issue was developed by Thomas Schelling in 1971 [52, 53]. The Schelling's residential segregation model studies the effects of a preference for people to be in a similar neighborhood and consists of a regular lattice, with a density of empty sites, whose nodes mimic agents of different ethnicities. After an initial distribution of the agents, at each elementary dynamical step an agent is randomly chosen and it moves to a empty site chosen at random if its ratio of other ethnic neighbors is higher than a tolerance threshold T. After a long enough transition time, he found that the agents remain in a mixed distribution only for very high values of the tolerance threshold, but agents form segregated neighborhoods for other values of tolerance (Figure 1.6).

Opinion Dynamics is a social topic very dealt with AB modeling, it studies the chance for a social group to reach agreement or disagreement about a question. Outside this framework, in the early 70's Clifford an Sudbury defined a model for population dynamics [54], the latter named Voter Model has been used in fields so varied such as social dynamics, population genetics, chemistry and probability theory. The description of Voter Model is not at all complex: each agent is provided of a variable that take two possible values, and at each time step a randomly chosen agent selects at random a neighbor and imitates its variable's value. Although its extremely simple design and the fact that can be solved exactly in any finite dimension [55], the model has been and is still studied in many sciences and according to different variants (as e.g., different topologies, noise and external fields). Other opinion dynamics model, the majority rule model [56], explores a similar topic through a different procedure: Starting from a complete graph whose nodes (agents) are provided of a variable (opin-

ion) that can take two possible values (-1, +1), at each time step a random group of r agents is selected and they share the majority opinion inside the group. The group size r is taken at each interaction from a given distribution. The model includes an asymmetry: when r is even, a value of opinion (e.g. +1) can be promoted in case of a tie. Under the mean field assumption, Krapivsky and Redner [57] solved the model for a fixed value of r. They found three fixed points: one unstable fixed point that corresponds to a situation in which the population is evenly distributed in both values of opinion variable, and two stable fixed points that correspond to one-opinion frozen states. The majority rule model has been studied under different network topologies, multi-state opinion and plurality rule [58], as well as modifications that include mobility, external fields, variable connectivity, etcetera ...

Other context of social research corresponds to the cases in which opinion is modeled as a vector of variables. These models are usually grouped under the name of cultural dynamics. The most representative and studied model in cultural dynamics was introduced by Robert Axelrod in 1997 [59]. The well-known Axelrod's model for culture dissemination explores the principle of homophily, developing the idea that a social agent will convince similar people easily than dissimilar ones, and therefore similar people tend to become even more alike. It consists of a regular lattice, whose nodes mimic cultural agents. Each one of these nodes is provided with a set of F cultural features that can assume q possible integer values. The parameter q represents the possible traits that each feature can assume. Thus a culture is modeled as a vector of F integer variables. After assigning the traits at random, the system evolves as follow: at each elementary dynamical step, an individual i randomly chosen imitates a feature's trait of a random neighbor j with a probability equal to the cultural overlap $\omega_{i,j}$ between both agents, defined as $\omega_{ij} = (\sum_{f=1}^{F} \delta_{\sigma_f(i),\sigma_f(j)})/F$, where $\delta_{x,y}$ is the Kronecker's delta which is 1 if $x = y$ and 0 otherwise. After a long enough transient, for a low value of the initial cultural diversity q, the system reaches a frozen monocultural state, in which all agents share every trait. On the other hand, for high values of the initial cultural diversity, the system can't arrive to cultural convergence but remains at a multicultural state, characterized by agents who hardly share features with their neighbors (Figure 1.7). The usual order parameter is the relative size of main culture $\langle S_{max} \rangle / N$, *i.e.* the maximum number of agents sharing every trait divided by the total population. Several others order parameters can be used as the number g of different cultural domains in the asymptotic state over the total population $g = \langle N_g \rangle / N$. Then, the final states above mentioned are characterized by $\langle S_{max} \rangle / N \sim 1,0$ and $g \sim 0,1$ respectively. Both states are bounded by a phase transition at a critical value q_c depending on F: this transition is continuous for $F = 2$, but discontinuous for $F > 2$. Axelrod model has been studied with many variants as random noise [147], mass media effects [150] and different network topologies [141, 145] among others.

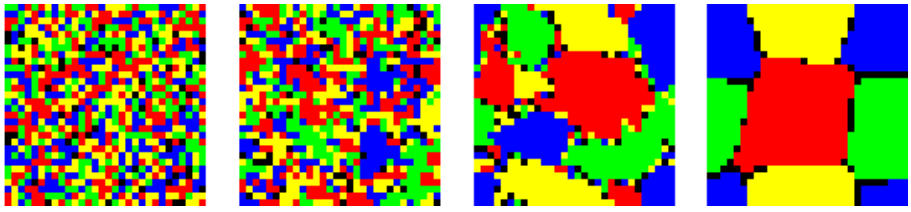


Figure 1.6: *Dynamics of the Schelling model.* The Schelling's model of residential segregation studies the effects of intolerance in the population distribution. People move if their neighborhood is unfriendly according to a tolerance parameter. The model shows segregation in neighborhoods for intermediate and high values of intolerance [52, 53]. Colors represent ethnicities and opposite sides are connected.

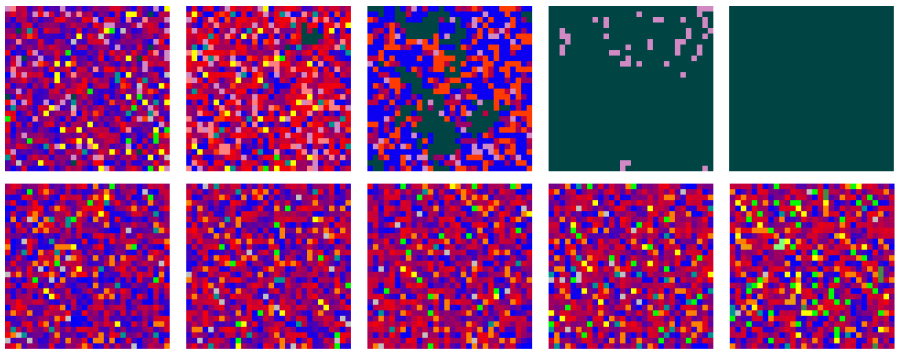


Figure 1.7: *Dynamics of the Axelrod model.* Top panels: For low values of the initial cultural diversity (q), the system converges to a frozen monocultural state, characterized by a cultural group shared by all the agents. Bottom panels: Otherwise, for high values of q the system remains in a multicultural state in which agents do not share many traits with their neighborhood [59]. Colors represent different cultures and periodic boundary conditions are applied.

1.3.2 Topology.

The description of interacting relations (who interacts with whom, how often, in which way do agents relate), is a key of the modeling process in social dynamics. The usual tool to describe the relationship between agents is a network in which agents are represented by the nodes and links represent the interaction chances. ABMs rely on different network structures, and outcomes usually strongly depend on the topology chosen. In a society, people are not usually connected everyone to everyone but in several ways. This fact can be modeled through different network kinds (directed or indirected, weighted or unweighted, etc) and topologies, including regular and complex networks (*e.g.* lattice, small-world, RRN, ER or SF). Although real networks, and particularly social networks, often differ from artificial networks in many features, the dependence of model behavior with topology provides a tool to understand in what way individual relationship influence on society.

A social network is a graph in which the nodes represent individuals or social groups and the links symbolize the relationships among them [60, 61]. In fact, many of the concepts used in the analysis of complex networks are based in social features, such as small-world property [62, 63, 64], as well as theoretical tools such as node centrality or clustering index [65]. Graph theory has provided a very useful tool for measurement of different social topics as collaboration networks, friendship or social interdependence. Additionally, new technologies provide new kinds of social interactions characterized by lower dependence on the physical location and higher connectivity of individuals [66]. Complex network theory offers a tool to study the properties of the contacts structure and the dynamics involved in the forming process. Examples of social networks are scientific colaborations (such as coauthorship networks), actors coworking relations, information exchange, online social networks friendship, etcetera.

The influence of network topology on the behavior of social ABMs mentioned in section 1.3.1 has been studied in the last decades, including several areas such as language dynamics, social behavior, rumors spreading, opinion models, cultural dynamics, etcetera (see, *e.g.* [67]). For example, in opinion models, consensus critical values show a strong dependence on the underlying topology [57]. Similarly, Axelrod model for cultural dissemination displays dependence of the phase transition on the network structure [145, 146]; beyond this dependence, the dynamics of the model can be used to build a network through a rewiring process, which in turn provides new system behaviors [166, 167].

In strategic games [12, 17], the role of network topology is usually introduced through the condition that, at every time step, one or more individuals interact with one or more agents chosen among their network neighbors. The specific model def-

inition depends, among other parameters, on the kind of synchronization (*e.g.* synchronous, serial) and on the strategies and updating rules considered. In particular, the dependence of social observables on the underlying topology has been widely analyzed through EGT models. For example, prisoner dilemma in complex networks has been object of several studies for the last twenty years, focusing in the influence of the structure of interactions among people on the level of cooperation [16, 94, 18, 109, 95]. The studies showed a strong dependence of the cooperation level on the topology, that is, heterogeneity enhances cooperation. Nevertheless, this dependence is very sensitive to the type of strategies and updating rules considered [79].

Part I

The emergence of cooperation.

Presentation of Part I.

"*First, let it be remembered that we have innumerable instances, both in our domestic productions and in those in a state of nature, of all sorts of differences of inherited structure which are correlated with certain ages and with either sex. We have differences correlated not only with one sex, but with that short period when the reproductive system is active, as in the nuptial plumage of many birds, and in the hooked jaws of the male salmon. We have even slight differences in the horns of different breeds of cattle in relation to an artificially imperfect state of the male sex; for oxen of certain breeds have longer horns than the oxen of other breeds, relatively to the length of the horns in both the bulls and cows of these same breeds. Hence, I can see no great difficulty in any character becoming correlated with the sterile condition of certain members of insect communities; the difficulty lies in understanding how such correlated modifications of structure could have been slowly accumulated by natural selection. This difficulty, though appearing insuperable, is lessened, or, as I believe, disappears, when it is remembered that selection may be applied to the family, as well as to the individual, and may thus gain the desired end. Breeders of cattle wish the flesh and fat to be well marbled together. An animal thus characterized has been slaughtered, but the breeder has gone with confidence to the same stock and has succeeded. Such faith may be placed in the power of selection that a breed of cattle, always yielding oxen with extraordinarily long horns, could, it is probable, be formed by carefully watching which individual bulls and cows, when matched, produced oxen with the longest horns; and yet no one ox would ever have propagated its kind. Here is a better and real illustration: According to M. Verlot, some varieties of the double annual stock, from having been long and carefully selected to the right degree, always produce a large proportion of seedlings bearing double and quite sterile flowers, but they likewise yield some single and fertile plants. These latter, by which alone the variety can be propagated, may be compared with the fertile male and female ants, and the double sterile plants with the neuters of the same community. As with the varieties of the stock, so with social insects, selection has been applied to the family, and not to the individual, for the sake of gaining a serviceable end. Hence, we may conclude that slight modifications of structure or of instinct, correlated with the sterile condition of certain members of the community, have proved advantageous; consequently the fertile males and females have flourished, and transmitted to their fertile offspring a tendency to produce sterile members with the same modifications. This process must have been repeated many times, until that prodigious amount of difference between the fertile and sterile females of the same species has been produced which we see in many social insects.*" (**Charles Darwin. *The Origin of Species by means of Natural Selection.* (1859)**).

Understanding how cooperative behavior emerges in different contexts is one of the major questions of modern science. The presence of cooperation in hostile environments, that is, when selfish behavior provides higher individual fitness as does cooperation, has been studied in many areas including microbiology, species evolution, population dynamics, economy and sociology.

Theory of evolution is based on natural selection, which in turn is based on the *survival-of-the-fittest* rule. The limited resources available in a habitat promotes competition between organisms of the same or different species that have to struggle to survive; the final purpose of this competition is to provide offspring. In such competition, cooperation is, at least in the first instance, faced to individual interest of survival and reproductive success. Nevertheless, cooperation and even altruism have evolved and persist, and evolutionists have studied this question extensively for the last 150 years.

Cooperation between biological entities pursuing their own ends is key to understanding biological issues such as the emergence of multicellular organisms or gregarious behavior, but also to analyze human societies: people form families, tribes, cities, nations, coworker networks, companies, research teams, associations, etcetera.

Altruism is defined as a form of cooperation in which there is no direct benefit to the organism carrying out the behavior, while mutually beneficial relationship refers to cooperative behavior in which there is a direct benefit to the actor as well as the recipient [68]. There are several proposed mechanisms which help to explain cooperative behavior; they are not necessarily mutually exclusive, so that a combination of some of them may be applied to explaining a particular case of cooperative behavior. The most widely accepted and studied ones are:

Kin selection refers to evolutionary strategies that favor the persistence and reproductive success of an agent's relatives, even at a cost to the survival or reproduction of the agent. Kin selection implies cross-generational genetic changes driven by interactions between relatives and may be applied when relatives influence one another on survival and offspring. William D. Hamilton established, in 1964, a mathematic condition to explain altruistic behavior through kin selection [69]. According to Hamilton's rule, an altruistic action from agent i to j may be justified when:

$$rB > C \ , \tag{1.26}$$

where r is the genetic relatedness of i to j, defined as the probability that two genes taken at random from the same locus in both individuals are identical by descent, B is the reproductive benefit obtained by j and C is the reproductive cost paid by i.

Group selection mechanism in biology takes the assumption that genes can spread into a population because of the benefits they provide to the community [70]. Although group selection is not widely accepted by evolutionists in biology for several reasons (*e.g.*, the different time scales between groups dynamics and reproduction), it is often applied to other areas such as human behavioral sciences [71].

Reciprocity refers to situations where cooperation and/or altruism is enhanced by the probability of future mutual interactions. There are three types of reciprocity that have been deeply studied:

i) **Direct reciprocity** mechanism was proposed by Robert Trivers in 1971 [72]. If there is a probability of repeated interactions between the same two players with a *cooperate* action available, then a strategy of mutual cooperation may be favored even when *non-cooperate* actions brings larger short-term benefits. Direct reciprocity can enhance cooperation only if the probability w of another encounter between the same two individuals is higher than the cost-to-benefit ratio c/b of the cooperative action $w > c/b$.

ii) **Indirect reciprocity** mechanism do not requires that the same two individuals interact again. In a mutual interaction, actions are observed by third party individuals who might inform others. Thus, social approval promotes cooperation by indirect reciprocity [73]. Indirect reciprocity can enhance cooperation only if the probability p, of knowing a random agent's reputation is higher than the cost-to-benefit ratio c/b of the cooperative action $p > c/b$.

iii) **Network reciprocity**. Real populations have spatial structures or underlying networks which imply that some individuals interact more often than others. According to the so-called *lattice reciprocity* mechanism, the cooperative action can take advantage of the topology of the network, so that cooperators clusters are often resilient to invasion by the defective action [19].

In this part of the thesis, we focus on the emergence of cooperation in complex networks. In the framework of evolutionary game theory, among other games that provides a satisfactory description of a wide range of situations (*e.g.*, the Public Goods Game), the Prisoner's Dilemma (PD) has become a standard for studying the cooperation. First, in chapters 2-3-4 we investigate in detail the dynamics of PD in different artificial networks under the assumption of a widely accepted strategy: the replicator dynamics. Later, in chapter 5 we take into consideration observed strategies in human behavior and study the consequences of such rules. Finally, in chapter 6 we test the predictions by a large-scale experiment.

Chapter 2

The dipole model: Thermodynamic study of a social system.

In Evolutionary Dynamics the understanding of cooperative phenomena in natural and social systems has been the subject of intense research during decades. We focus attention here on the so-called "Lattice Reciprocity" mechanisms that enhance evolutionary survival of the cooperative phenotype in the Prisoner's Dilemma game when the population of darwinian replicators interact through a fixed network of social contacts. Exact results on a "Dipole Model" are presented, along with a mean-field analysis as well as results from extensive numerical Monte Carlo simulations. The theoretical framework used is that of standard Statistical Mechanics of macroscopic systems, but with no energy considerations. We illustrate the power of this perspective on social modelling, by consistently interpreting the onset of lattice reciprocity as a thermodynamical phase transition that, moreover, cannot be captured by a purely mean-field approach.

2.1 Introduction.

Is the term "social temperature" just a rhetoric figure (suggestive metaphor), or on the contrary, could it be given a precise meaning? By working out in detail the evolutionary dynamics of the most studied social dilemma (the Prisoner's Dilemma) on a simple kind of artificial social networks we will show here that the formal framework of equilibrium statistical mechanics is, to a large extent, applicable to the rigorous

description of the asymptotic behavior of strategic evolution, thus providing the key for a formal quantitative meaning of the term social "temperature" in these contexts.

Evolutionary game theory, in contrast with classical game theory that focusses on the decision making process of (rational) agents, is concerned with entire populations of agents programmed to use some strategy in their interactions with other agents. The agents are replicators, *i.e.* entities which have the means of making copies of themselves (by inheritance, learning, infection, imitation, etc...), whose reproductive success depends on the payoff obtained during interaction. As the payoff depends on the current composition of strategies among the interacting agents, this yields a feedback loop that drives the evolution of the strategic state of the population[12][17][79][80].

This darwinian feedback (frequency-dependent fitness) dynamics depends strongly not only on the particular game, and on the specifics of the way strategies spread, but also on the (social) structure of connections describing the interactions. Under the assumption of a well-mixed population (*social panmixia* assumption), the temporal evolution of the proportion of strategies among the population is governed by a differential equation named *replicator equation* (see below). Well-known celebrated folk's theorems (see, e.g. [79]) establish a connection between the asymptotic behavior of this equation and the powerful concepts of classical game theory based on the notion of best reply (Nash). However, if the social panmixia assumption is abandoned, and individuals only interact with their neighbors in a social network, the asymptotic of evolutionary dynamics generically differ in a substantial way from this "well-mixed population" description. The social structure of strategic interactions turns out to be of importance regarding the evolutionary outcome of the strategic competition.

We will consider here the Prisoner's Dilemma (PD), a two-players-two-strategies game, where each player chooses one of the two available strategies, cooperation or defection: A cooperator receives R when playing with a cooperator, and S when playing with a defector, while a defector earns P when playing with a defector, and T (temptation) against a cooperator. When $T > R > P > S$, the game is a PD (while if $T > R > S > P$ it is called Snowdrift game, also "Chicken" or "Hawks and Doves"). Given the payoff's ordering, whatever the value of the prior assign of probability to the co-player's strategy is, the expected payoff is higher for defection, and that is what a rational agent should choose. In the PD game only the defective strategy is a strict best response to itself and to cooperation, thus it is an easy example of game with an unbeatable [80] strategy. Still, though there is no difficulty in the making of the strategic decision from Nash analysis, two cooperators are better off than two defectors, hence the social dilemma.

In graph-structured populations, a large body of research [19, 81, 82, 83, 84, 85, 86, 87, 16, 90, 101, 102, 103, 104] (and references therein) on evolutionary dynamics

of the PD game has convincingly show the so-called *lattice reciprocity* effects: The cooperative phenotype can take advantage of the topology of the social net, so that clusters of cooperators are often resilient to invasion by the (continuum-unbeatable) defective phenotype. This enhancement of asymptotic macroscopic levels of cooperation due to the structure and topology of strategic interactions includes, but it is far more general than, the so-called space reciprocity mechanisms, where social nets are discretizations (solid state lattices) of the euclidian space, and diffusion approximations are often useful [105]. In this regard, one should stress the accumulated evidence that (*i*) many interesting social nets [27, 28, 45] are far away from being regular lattices, and (*ii*) freedom of connectivity scales (scale-free complex networks) enhances [18, 106, 107, 109, 111, 112, 113] the lattice reciprocity mechanisms up to unexpectedly high values of the temptation parameter T of the dilemma, where cooperation is very expensive (but affordable in an evolutionary sense).

In this chapter we investigate in detail the lattice reciprocity mechanisms in an artificial network (Dipole Model) that models the competition for influence on a population of social PD-imitators of two antagonist Big Brothers (nodes connected to the whole population, but with no direct connection between them).

2.2 Natural strategic selection on graphs.

We specify here the evolutionary game dynamics scenario, meaning the game parametrization, the microscopic strategic dynamics (replication mechanism or strategic updating rule), and the social structure of contacts that we will consider along the chapter.

We normalize the PD payoffs to the reward for cooperating, $R = 1$, and fix the null payoff at punishment $P = 0$. Note that provided the (differential or relative) selective advantage among two individuals depends on their payoff's difference (see below), one can arbitrarily fix the zero payoff level. Then only two parameters $T = b > 1$ and $S = \epsilon \leq 0$ are tuned. Note that the range $\epsilon > 0$ defines a game named Hawks and Doves (also Chicken and Snowdrift) where punishment and sucker's payoff have the reverse order. We will occasionally comment on this range of parameters.

Moreover, we do not restrict our computations to $2R > T + S$. This restriction means that the total payoff for the two players is higher if both cooperate ($2R$) than if one cooperates and the other defects ($T + S$), and is usually incorporated in iterated games studies of the PD to prevent agents taking turns at defection and then sharing the payoffs. For the specifics of the replicator dynamics (memory-less, markovian) in the next paragraph, one should not expect that this restriction qualitatively matters.

Regarding the replication mechanism, we implement the finite population (size $N \gg 1$) analogue of replicator dynamics [15, 18]. At each time step t, which represents one generation of the discrete evolutionary time, each agent i plays once with each one of the agents in its neighborhood and accumulates the obtained payoffs, P_i. Then, the individuals, i, update synchronously their strategies by picking up at random a neighbor, j, and comparing their respective payoffs P_i and P_j. If $P_i > P_j$, nothing happens and i keeps the same strategy for the next generation. On the contrary, if $P_j > P_i$, with probability $\Pi_{i\to j} = \eta(P_j - P_i)$, i adopts the strategy of its neighbor j for the next round robin with its neighbors, before which all payoffs are reset to zero. Here η is a number small enough to make $\Pi_{i\to j}$ an acceptable probability; its physical meaning is related to the characteristic inverse time scale: the larger it is, the faster evolution takes place.

From a theoretical point of view, this specific choice of the dynamics has the virtue of leading directly (see, *e.g.* [17]), under the hypothesis of a well-mixed population and very large population size, to the celebrated replicator equation for the frequencies p_α of strategies α(= C or D) in the population:

$$\dot{p}_\alpha = p_\alpha(f_\alpha - \bar{f}) \tag{2.1}$$

where f_α is the payoff of an α-strategist and $\bar{f}$ is the average payoff for the whole population. Note that time unit in equation (2.1) is scaled to η^{-1}.

For the payoffs of the Prisoner's Dilemma the asymptotic frequency of cooperators, from the replicator equation, is driven to extinction, $p_c = 0$, while for the Hawks and Doves game, its asymptotic value is $\epsilon/(b - 1 + \epsilon)$. As stated in the introductory section, we will be concerned here mainly with populations that are not well-mixed, where predictions based on this nonlinear differential equation are often of little use.

Regarding the structure of connections between interacting agents, we will consider here that it is given by a fixed graph (*i.e.* connections between players do not change by rewiring) where agents are represented by nodes, and a link between nodes indicates that they interact (play). If k_i is the number of neighbors of agent i (connectivity or degree), and Δ is the maximal possible one-shot-payoff difference ($\Delta = \max\{b, b-\epsilon\}$), we will assume $\eta = (\max\{k_i, k_j\}\Delta)^{-1}$ for the specification of the probability $\Pi_{i\to j}$ of invasion of node i by the strategy of neighbor j. This simple choice, introduced in [18], assures that $\Pi_{i\to j} < 1$; in heterogeneous networks it has also the effect of slowing down the invasion processes from or to highly connected nodes, with respect to the rate of invasion processes between poorly connected nodes, a feature not without consequences [114].

We now introduce some notation, which is familiar to statistical physicists: The configuration (*strategic microstate* l) of a population of N agents at time t is specified

by the sequence $l = \{s_i(t)\}$ ($i = 1, ..., N$), where $s_i(t) = 1$ (or 0) denotes that node i is at this time a cooperator (resp. defector). The set of all possible 2^N configurations is called the phase space. Stationary probability densities of microstates $\mathcal{P}(l)$ ($l = 1, ...2^N$) are then representatives of *strategic macro-states*. The average cooperation c_l of microstate l is defined as

$$c_l = \frac{1}{N}\sum_i^N s_i \tag{2.2}$$

We denote by $\Pi_{l'l}$ the probability that the strategic microstate of the population at time $t+1$ is l', provided that it is l at time t. Note that $\sum_{l'} \Pi_{l'l} = 1$. A microstate $\hat{l}$ is a *frozen* equilibrium configuration if the probability that it changes in one time step is null, and then $\Pi_{\hat{l}\hat{l}} = 1$ and $\Pi_{l'\hat{l}} = 0$ if $l' \neq \hat{l}$. We will assume generic real values (irrational) of the payoff parameters, so that if a configuration contains a C-D link it cannot be a frozen configuration. The only possible frozen equilibrium configurations are *all*-C and *all*-D. However, for a very wide class of graphs, and a wide range of model parameters they are not the only possible stationary probability measures.

We now illustrate by means of easy examples the evolution of PD on graphs. Our first and simplest example is a star-shaped graph consisting of a central node connected to $N-1$ peripheral nodes. It is straightforward to check that any initial condition with cooperators at the central node and (at least) at $[(b-\epsilon(N-1))/(1-\epsilon)]+1$ peripheral nodes has a positive probability of evolving in one time step to a configuration with a higher number of cooperators, and a null probability of evolving towards less cooperators. Thus, all those configurations evolve asymptotically to the *all*-C equilibrium. The rest of configurations evolve towards the *all*-D equilibrium. Therefore, if $N > (b-\epsilon+2)$ both equilibria are attractors (absorbing states), in the sense that some configurations different from themselves evolve to them; the phase space is partitioned into two basins of attraction. If $N < b-\epsilon+2$, only the *all*-D frozen equilibrium is attractor. The stationary probability densities $\mathcal{P}^*(l)$ of the star are pure point measures (two- or one- Dirac delta peaks) in the thermodynamic limit $N \to \infty$.

Now take a star and add some arbitrary number of links between its peripheral nodes. We call this network a crown, whose head is the central node. If the head is occupied at t_0 by a defector, it will remain so forever, because the payoff of a peripheral cooperator is strictly lower than head's payoff. Sooner or later the head (center) of the crown will be imitated by the whole crown, and the evolution will stop when everybody be defecting. But, what happens to a cooperator on the head? The answer is dependent on both, the net topology of the crown periphery and the cooperators disposition there: To ensure fixation of cooperation at the head node, it

suffices that a subset C of peripheral nodes occupied by cooperators, and with no direct links to the rest of the periphery, have a size $n_C > bk_{max} - \epsilon(N - n_C - 1)$, where k_{max} is the maximal degree in the rest of the periphery. Under this proviso all-C is the unique absorbing microstate of all corresponding initial conditions.

Finally consider the graph schematized in Fig. 2.1, composed of the following:

(*a*) A component $\mathcal{F}$ of n_F nodes with arbitrary connections among them.

(*b*) A node, say node 1, that is connected to all the nodes in $\mathcal{F}$ and has no other links.

(*c*) A component $\mathcal{C}$ of n_C nodes with arbitrary connections among them.

(*d*) A node, say node 2, that is connected to all the nodes in $\mathcal{F}$ and $\mathcal{C}$, but not to node 1.

This is what we will call a Dipole Model network. It is a two-headed (nodes 1 and 2) crown (with periphery $\mathcal{F}$) plus a tail $\mathcal{C}$ hanging on head 2. To strength the special status of the head nodes, let us nickname them as "Big Brothers". They certainly enjoy a sort of omnipresence that fits well with the character of Orwell's famous social sci-fiction novel 1984. In the following section we prove that for this simple network there exists a non-trivial stationary probability density of microstates $\mathcal{P}^*(l)$ for the strategic evolution of the PD game.

2.3 The Dipole Model.

The analysis of evolutionary dynamics of the PD on the Dipole network shows that there is a non-trivial invariant measure in phase space. Let us consider the set $\mathcal{I}$ of initial conditions defined by: (*i*) Big Brother 1 is a defector, (*ii*) Big Brother 2 is a cooperator, and (*iii*) all nodes in component $\mathcal{C}$ are cooperators. Note that this set contains 2^{n_F} different configurations. We now prove that, provided some sufficient conditions, this is a minimally invariant set of the evolutionary dynamics.

First, one realizes that Big Brother 1 cannot be invaded by the cooperative strategy: The payoff of a cooperator node i in $\mathcal{F}$ is $P_i^c = k_i^c + 1 + \epsilon(k_i - k_i^c + 1)$, where k_i is the number of its neighbors in $\mathcal{F}$ and $k_i^c \leq k_i$ is the number of those that are cooperators. The payoff of Big Brother 1 (BB1) is then $P_1 \geq (k_i^c + 1)b$. For the PD game, where $\epsilon \leq 0$, the inequality $P_1 > P_i^c$ always holds, so that BB1 will always be a defector. (Note also that for the Hawks and Doves game, a sufficient condition for $P_1 > P_i^c$ is $b > 1 + \epsilon(k_F + 1)$, where k_F ($< n_F$) is the maximal degree in component $\mathcal{F}$, *i.e.* the maximal number of links that a node in $\mathcal{F}$ shares within $\mathcal{F}$.) We thus conclude that defection is fixed at BB1.

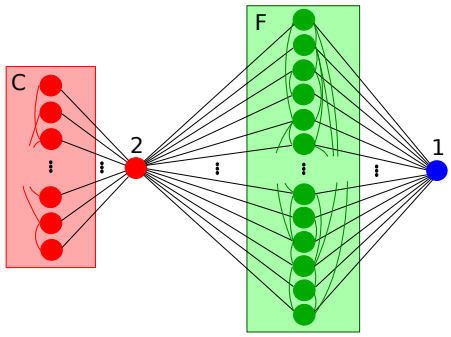


Figure 2.1: Structure of the Dipole Network. Two nodes (1 and 2) are connected to all nodes in $\mathcal{F}$, whose elements can be arbitrarily linked to each other. Moreover, node 2 is also linked to a set $\mathcal{C}$ (with arbitrary internal connections as well). Initial conditions are indicated by colors: red for cooperators (node 2 and set $\mathcal{C}$), blue for defectors (node 1), and green means arbitrary (set $\mathcal{F}$). See the text for further details.

Second, thanks to its interaction with set $\mathcal{C}$, Big Brother 2 resists invasion, provided its size n_C is above a threshold: The payoff of a defector node i in $\mathcal{F}$ is $P_i^d = (k_i^c + 1)b$, where k_i^c is the number of its cooperator neighbors in $\mathcal{F}$, while the payoff of Big Brother 2 (BB2) is $P_2 = n_C + n_F\epsilon + n_F^c(1-\epsilon)$, where $n_F^c \leq n_F$ is the number of cooperators in $\mathcal{F}$. Thus, a sufficient condition for $P_2 > P_i^d$ is $n_C > b(k_F + 1) - n_F\epsilon$. With this proviso, BB2 will always be a cooperator, which in turn implies that all the nodes in the component $\mathcal{C}$ will remain always cooperators. Note that for $\epsilon \leq 0$ and $b > 1$, the absence of the component $\mathcal{C}$ could imply invasion of node 2, that would lead to fixation of the defective strategy on the whole network.

The previous argument proves that provided the sufficient conditions $n_C > b(k_F + 1) - \epsilon n_F$ and $b > 1 + \epsilon(k_F + 1)$ hold, the subset $\mathcal{I}$ of phase space defined by (*i*), (*ii*), and (*iii*) is an invariant set. As this set does not contain equilibria, no stochastic trajectory evolves from it to a frozen equilibrium configuration.

Finally, one realizes that $\mathcal{I}$ is indeed minimal, because at any time, a defector in $\mathcal{F}$ has a positive probability to be invaded by the cooperation strategy (at least from BB2), and a cooperator in $\mathcal{F}$ has a positive probability of being invaded by the defection strategy (at least from BB1). Therefore, any strategic configuration of the set $\mathcal{I}$ is reachable in one time step from any other, *i.e.* for all pairs (l, l') of microstates in $\mathcal{I}$, the transition probability $\Pi_{l'l} > 0$. Consequently, $\mathcal{I}$ does not contain proper invariant subsets: it is minimally invariant. Moreover, following Perron-Frobenius theorem, there exists a unique stationary macro-state $\mathcal{P}^*(l)$. This provides a rigorous framework for the interpretation of results from numerical Monte Carlo simulation studies in evolutionary dynamics on dipole models, provided the sufficient conditions above.

While nodes in $\mathcal{C}$ and Big Brother 2 are permanent cooperators, and Big Brother 1 is a permanent defector, nodes in $\mathcal{F}$ are forced to fluctuate. This partition of the network into sets of nodes where each particular strategy is fixed forever, and a set of fluctuating nodes, turns out to be a generic feature of the discrete replicator dynamics (neighbor imitation proportional to payoffs difference) on many network settings [109, 112]. The simplicity of the Dipole Network model allows on it an easy formal proof of existence of this partition, so providing an illustration of both, its origins and generic character. It also shows the formal applicability of equilibrium statistical physics formalism to characterize the asymptotic behavior of evolutionary dynamics on these graphs. This will be made in the next section for specific choices of structural traits for the subgraph $\mathcal{F}$.

Let us note that if a direct link between BB1 and BB2 is added, then (see appendix 2.4.3) asymptotic fixation in the whole network of either cooperation or defection will occur, depending on the relative size n_C/n_F of components $\mathcal{C}$ and $\mathcal{F}$.

The name dipole for this structure of connections is suggested by the strategic polar ($\mathcal{C}-\mathcal{F}-\mathcal{D}$) aspect of the whole graph. Note also that the number of $\mathcal{C}-\mathcal{F}$ and $\mathcal{F}-\mathcal{D}$ connections scales linearly with the size n_F of the fluctuating interior, that is to say that the poles (C and D) act as an externally imposed (AC) field on $\mathcal{F}$, whose strength is proportional to the internal levels of cooperation. As the cooperation (and then the fitness) levels are self-sustained (as proved by the previous theorem), this is a closed macroscopic system with a non-trivial self-sustained social activity of cooperation at evolutionary equilibrium.

The interest of the Dipole Model is by no means restricted to a mere academic illustration: First of all, we can make a technical use of it in macroscopic stability analysis studies of PD-evolution on highly heterogeneous complex networks. Indeed, the fluctuations inside the subset $\mathcal{F}$ are the effect of the competition for invasion among two non-neighboring hubs (hugely connected nodes), where opposite pure strategies have reached fixation, in their common neighborhood. This is a local strategic configuration that mimics those that are often observed in stochastic simulations of evolutionary dynamics in highly heterogeneous (scale-free) networks [109, 112]. Simple multipolar network models can easily be constructed (e.g. by establishing direct links from $\mathcal{C}$ to $\mathcal{F}$ in a way that simple sufficient conditions guarantee that the theorem still holds), that are indeed indistinguishable from typical strategic patterns found in the numerical simulations on scale-free networks. This makes the Dipole net a very useful technical device to analyze the stability mechanisms of the cooperator clusters [109, 112] in scale-free structured populations, as well as the kind of temporal fluctuations of cooperation that one should expect in the fluctuating set of nodes.

Regarding potentialities for econo-socio-physics applications of the Dipole model, it could be viewed as a sort of schematic (then simplistic, cartoon-like) model for

the competition for influence of two powerful superstructural institutions (*e.g.* like "mass media", political parties, or lobbies) on a target population, in strongly polarized strategic contexts. The analysis rigorously provides sufficient conditions for the parameter values where fixation of strategic traits is proved impossible, so that temporal fluctuations dominate forever the target population of social imitators $\mathcal{F}$. The influence on each individual of the two competing institutions is simulated here through the omnipresent ("Big Brother" nodes 1 and 2) neighbors, whose own high appeal for imitation (the strength of Big Brother's influence) is in turn conditioned by the strategic composition of the target population. Here the interest could well be the study of the influence that metric and topological network characteristics of the social structure have on the strategic macro-state, and thus on the quantitative values of *social indicators*. We address some aspects of this issue in the next section.

At a more general level, the design of experiments in social sciences as well as theoretical studies of artificial societies could greatly benefit from having at hand simple but non-trivial "exactly soluble statistical-mechanical models" that may provide safe guides to develop further intuitions on social phenomena that demands more comprehension.

2.4 The role of social structure in Big Brothers competition.

In this section we present some analytical and numerical results on the evolutionary dynamics of games in the Dipole Model for different choices of topologies of the fluctuating set $\mathcal{F}$. The sufficient conditions stated in the previous section are assumed hereafter. We are interested in the situation where $n_F \gg 1$, *i.e.* large size of the fluctuating population.

First we will briefly comment on the straightforward limiting case when the macroscopic set $\mathcal{F}$ is a fully connected set, so that $k_F = n_F - 1$. This is the well-mixed population limit, for which it is easy to show that the replicator equation (2.1) is an exact description. The payoffs of polar nodes BB1 and BB2 are given by $P_1 = bcn_F$ and $P_2 = n_C + cn_F + \epsilon(1-c)n_F$, while the payoffs of a cooperator node and a defector node in $\mathcal{F}$ are $P_c = cn_F + \epsilon(n_F - cn_F + 1)$ and $P_d = (cn_F + 1)b$. One easily realizes that $P_c < P_d$, provided the sufficient condition ($b > 1 + \epsilon n_F$) for fixation of defection at node 1. Thus the (one time step) probabilities Q_{DC} (invasion of a cooperator node in $\mathcal{F}$) and Q_{CD} (invasion of a defector node in $\mathcal{F}$) are

$$\begin{aligned} Q_{DC} &= \frac{1}{(n_F+1)}\frac{P_1-P_c}{\Delta(n_F+1)} + \frac{(1-c)n_F}{(n_F+1)}\frac{P_d-P_c}{\Delta(n_F+1)} \\ Q_{CD} &= \frac{1}{(n_F+1)}\frac{P_2-P_d}{\Delta(n_F+n_C)} \quad . \end{aligned} \tag{2.3}$$

Assuming that the size of $\mathcal{F}$ is macroscopic, $n_F \gg 1$, the fraction of cooperators c in $\mathcal{F}$ evolves according to

$$\dot{c} = (1-c)Q_{CD} - cQ_{DC} \quad . \tag{2.4}$$

Now, if $n_F \gg 1$, and $n_C/(n_F)^2 \to 0$, then both Q_{CD} and the first term in the right-hand side of Q_{DC} vanish, and we arrive to the differential equation

$$\dot{c} = \frac{c(1-c)}{\Delta}(\epsilon(1-c) - (b-1)c) \quad . \tag{2.5}$$

This is, with a simple re-scaling of time, the replicator equation (2.1): note that in the limit $n_F \gg 1$ that we have considered, the probability that a node in $\mathcal{F}$ picks up a Big Brother when updating its strategy is negligible, and then the evolution inside the complete graph $\mathcal{F}$ is overwhelmingly determined by the internal connections, and thus by the replicator equation. In other words, in this limit of maximal possible connectivity, BB1 and BB2 are no longer bigger than the nodes in $\mathcal{F}$ and their influence on the fluctuating set is negligibly small in the thermodynamic limit.

We now turn attention to situations where $k_F \ll n_F$, far from the social panmixia. In subsection 2.4.1 we will explicitly solve the opposite trivial case of disconnected $\mathcal{F}$ set ($k_F = 0$), which turns out to reduce to the standard textbook ideal two-states model of Statistical Physics. After that, in subsection 2.4.2, the "random regular" network structure for $\mathcal{F}$ is seen to be amenable to a plausible mean-field approach, but insufficient to explain the phenomenology shown by Monte Carlo numerical results. These show beyond any doubt a critical behavior, a transition point separating two qualitatively different types of social macro-states. This transition is sensibly interpreted as the onset of lattice reciprocity. In other words, *lattice reciprocity is a true critical social phenomenon*.

2.4.1 $\mathcal{F}$ is a disconnected graph (ideal-gas).

Let us now obtain some explicit results for one of the simplest choices for the topology of connections inside the fluctuating set, namely $k_F = 0$. In this case each node

in $\mathcal{F}$ is only connected to Big Brothers. This is in fact an effective single node problem, where homogeneity (*i.e.*, mean field assumption) in $\mathcal{F}$ is exact; in other words, the absence of internal interactions in the set $\mathcal{F}$ is a sort of ideal-gas condition easy to exactly deal with in the large size limit.

Note that the sufficient conditions for fixation of defection at BB1 and of cooperation at BB2 are respectively, $b > 1 + \epsilon$, and $n_C > b - \epsilon n_F$. Denoting by $c(t)$ the instantaneous fraction of cooperators in $\mathcal{F}$, one finds for the (one time step) probability Q_{DC} of invasion of a cooperator node in $\mathcal{F}$

$$Q_{DC} = \frac{cb - (1+\epsilon)/n_F}{2\Delta} \quad , \tag{2.6}$$

and using the notation $A = \epsilon + (n_C - b)/n_F$ and $B = 1 + n_C/n_F$

$$Q_{CD} = \frac{A + c(1-\epsilon)}{2\Delta B} \quad , \tag{2.7}$$

for the probability of invasion of a defector node in $\mathcal{F}$. Note that $A > 0$ due to the non-invasion of BB2 (sufficient) condition.

Provided $n_F \gg 1$, the fraction of cooperators c in $\mathcal{F}$ evolves according to the differential equation (2.4), which after insertion of expressions (2.6) and (2.7), and re-scaling of time, becomes

$$\dot{c} = f(c) \equiv A_0 + A_1 c + A_2 c^2 \quad , \tag{2.8}$$

where the coefficients are

$$A_0 = A \quad , \tag{2.9}$$

$$A_1 = 1 - \epsilon - A + B(1+\epsilon)/n_F \quad , \tag{2.10}$$

$$A_2 = -(1 - \epsilon + bB) \quad , \tag{2.11}$$

One can easily check ($A_0 > 0$ and $A_2 < 0$) that there is always one positive root c^* of $f(c)$, which is the asymptotic value for any initial condition $0 \leq c(0) \leq 1$ of equation (2.8).

For $\epsilon = 0$, in the so called weak PD game (*i.e.* at the border between the PD and the Hawks and Doves game), if one further assumes that the relative size $\mu(F)$ of the component F is large enough, *i.e.* $\mu(F) \to 1$, and $\mu(C) \to 0$, one easily obtains

that the stationary solution of equation (2.8) behaves as $c^* \simeq (b+1)^{-1}$ near the limit $\mu(F) \to 1$.

From the point of view of the set $\mathcal{F}$, when $n_F \gg 1$, the model corresponds to a non-interacting (ideal) set of independent phenotypic strategists that fluctuate due to a polar field (Big Brothers influence) whose strength is self-consistently determined by the average cooperation c. This problem is equivalent to the equilibrium of an ideal paramagnetic salt in a noisy (telegraphic) magnetic AC field of intensity proportional to the average magnetization.

A typical and correct statistical-physicists approach "from scratch" to this two-states model is the familiar micro-canonical setting: At (dynamical) macroscopic equilibrium, the probability of each strategic micro-state $l = \{s_i\}$ of fixed value of $c_l = c$ is uniform

$$P_l = \Omega^{-1} \ , \tag{2.12}$$

where $\Omega = n_F!/((cn_F)!(n_F - cn_F)!))$ is their number. The lack of information $S = \ln \Omega$ of the macro-state as a function of global cooperation $n_F c$, *i.e.* the relation $S(n_F c)$, can be regarded as the analogue of the micro-canonical fundamental "thermodynamical" relation, and its first derivative is the intensive parameter β (thus the analogue of the inverse thermodynamical temperature), that after using Stirling's approximation is easily obtained as

$$\beta = \ln\left(\frac{1-c}{c}\right) \ . \tag{2.13}$$

This relation is the analogue of a thermodynamical equation of state, which simply expresses the connection of the equilibrium value of the macroscopic cooperation level c to the "entropic" intensive parameter β. Note that c is determined by the balance condition ($\dot{c} = 0$):

$$\frac{1-c}{c} = \frac{Q_{DC}}{Q_{CD}} \ , \tag{2.14}$$

from where the equation of state (2.13) determines β as a function of model parameters (*i.e.* b, ϵ, and n_C/n_F). For example, when $\epsilon = 0$, $\beta = \ln b > 0$, indicating that the disorder of the activity increases with increasing cooperation. The maximal value of $\beta \to \infty$ corresponds to zero disorder ($b \to \infty$), while its minimal zero value corresponds to highest possible value (at $b = 1$) of cooperation ($c = (1/2)$). Note that values of $b < 1$ correspond to negative β values, where entropy decreases with increasing values of cooperation, outside the PD domain. [1]

[1] Note: The Stag Hunt game corresponds to $b < 1$ and $\epsilon \leq 0$, and it is the archetype of coordination games. A clear case for the deep importance of this game in Social Studies can be found in [116].

An alternative (and equivalent in the thermodynamic limit) setting is to consider the whole space of 2^{n_F} configurations $l = \{s_i\}_{i=1}^{n_F}$, of unrestricted c_l, but under the condition that the average value $c = \sum_l \mathcal{P}_l c_l$ is fixed. This is the analogue of the canonical setting. The normalization factor $Z = \sum_l \exp(-\beta c_l)$ is the analogue of the familiar canonical partition function (Boltzmann's Zustandsumme), that due to the agents independence ($k = 0$) is easily factorized as $Z = (1 + \exp(-\beta))^{n_F}$.

In the canonical setting a most informative macroscopic quantity is the "heat capacity" analogue: The fluctuations of c_l along representative (typical) stochastic trajectories at equilibrium under the evolutionary dynamics of the game are, following the standard thermodynamical formalism, given by $\partial c/\partial(\beta^{-1})$, so that this quantitative social indicator detects very precisely sudden variations of the macroscopic cooperation with payoff's parameters. In this ideal-gas kind of case there are no critical points and fluctuations do not diverge. For example, for $\epsilon = 0$ they are given by the (Bernouillian) binomial variance $n_F c(1-c) = n_F b/(b+1)^2$.

2.4.2 $\mathcal{F}$ is a random regular graph.

Random regular networks are random networks of fixed degree k. All nodes being thus equivalent, a sensible approach is to assume (mean-field like, see *e.g.* [117]) that the fraction of instantaneous cooperators in the neighborhood of a node is the fraction c of the whole set $\mathcal{F}$. In other words, one neglects local fluctuations of c. The contribution of the internal interactions to the variation of c is then of the "replicator equation" type, as discussed above for the complete graph case. The difference here is that if $k_F \ll n_F$ the contribution of the interactions with Big Brothers cannot be longer neglected.

Mean-field approximation.

The payoffs of Big Brothers BB1 and BB2 are given by $P_1 = bcn_F$ and $P_2 = n_C + cn_F + \epsilon(1-c)n_F$, while the payoffs of a cooperator node and a defector node at $\mathcal{F}$ under the mean-field assumption are:

$$P_c = ck + 1 + \epsilon(k(1-c)+1) \quad , \qquad P_d = (ck+1)b \ . \tag{2.15}$$

The differential equation for c is then

$$\begin{aligned} \dot{c} \ &= \ \frac{(1-c)(P_2 - P_d)}{(k+2)Bn_F\Delta} - \frac{c(P_1 - P_c)}{(k+2)n_F\Delta} \\ &\quad + \frac{(1-c)ck(P_c - P_d)}{(k+2)^2\Delta} \ , \end{aligned} \tag{2.16}$$

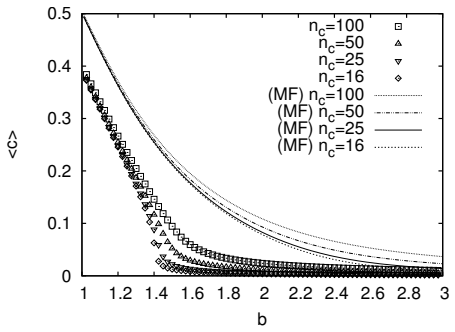


Figure 2.2: Macroscopic cooperation in a random regular graph structure for the set $\mathcal{F}$, with $k = 4$, and $n_F = 4000$, and $\epsilon = 0$. A decreasing sequence of n_C/n_F, as indicated in figure, has been used. Symbols represent numerical Monte Carlo results, and the different lines represent the mean-field predictions as given by the solution ($\dot{c} = 0$) of Eq. (2.17).

which under the assumption $kb \ll n_F$, takes the form

$$\dot{c} = f(c) \equiv \frac{1}{(k+2)^2 B\Delta}(A'_0 + A'_1 c + A'_2 c^2 + A'_3 c^3) \ , \tag{2.17}$$

where the coefficients are

$$\begin{aligned} A'_0 &= (k+2)(B-1+\epsilon) \ , & (2.18)\\ A'_1 &= 2(2(1-\epsilon)-B) + k(2(1-\epsilon) - B(b-\epsilon)) \\ &\quad +k^2 B\epsilon \ , & (2.19)\\ A'_2 &= 2(\epsilon - 1 - Bb) + k(\epsilon - 1 - B(1+\epsilon)) \\ &\quad +k^2 B(1-b-2\epsilon) \ , & (2.20)\\ A'_3 &= k^2 B(b-1+\epsilon) \ , & (2.21) \end{aligned}$$

Note that the assumption $n_C > b - n_F\epsilon$ (*i.e.* the condition for Big Brother 2 to be a permanent cooperator) implies that $A'_0 > 0$, so that $\dot{c}(0) > 0$ and one positive root, say c^*, of $f(c)$ is then ensured, in agreement with the theorem of section 2.3. In Fig. 2.2 we show the asymptotic value of the average cooperation c versus the temptation parameter b, as obtained from (2.17), for several different values of n_C/n_F, $\epsilon = 0$, and $k = 4$.

Within the mean field approximation, it is possible to obtain explicitly the equilibrium macro-state, *i.e.* the stationary probability distribution density $\mathcal{P}^*_l$, which as expected from section 2.3 turns out to be of the Boltzmann type. Let us consider two

different (arbitrary) strategic microstates $l = \{s_i\}$ ($i = 1, ..., n_F$), and $l' = \{s'_i\}$, of the fluctuating set. For any pair of microstates (l, l') we define the following numbers:

$$n_{11} = \sum_i \delta_{s_i,s'_i}\delta_{s'_i,1} \ , \tag{2.22}$$

$$n_{10} = \sum_i (1-\delta_{s_i,s'_i})\delta_{s'_i,0} \ , \tag{2.23}$$

$$n_{00} = \sum_i \delta_{s_i,s'_i}\delta_{s'_i,0} \ , \tag{2.24}$$

$$n_{01} = \sum_i (1-\delta_{s_i,s'_i})\delta_{s'_i,1} \ , \tag{2.25}$$

i.e., n_{11} is the number of nodes that are cooperators in both microstates, n_{10} that of the nodes that are cooperators in l but defectors in l', etc... Using equation (2.2) it is straightforward to obtain

$$c_l - c_{l'} = \frac{1}{n_F}(n_{10} - n_{01}) \ . \tag{2.26}$$

Now, let us assume that the probabilities that a node i changes strategy are independent of node i (homogeneity assumption, mean-field), and denote them by Q_{CD} (transition from defector to cooperator) and Q_{DC} (for the transition from cooperator to defector). Then we can easily see that the transition probabilities between the microstates l and l' are given by

$$\Pi_{l,l'} = (1-Q_{DC})^{n_{11}}(1-Q_{CD})^{n_{00}}Q_{DC}^{n_{01}}Q_{CD}^{n_{10}}, \tag{2.27}$$

$$\Pi_{l',l} = (1-Q_{DC})^{n_{11}}(1-Q_{CD})^{n_{00}}Q_{DC}^{n_{10}}Q_{CD}^{n_{01}}, \tag{2.28}$$

Henceforth, denoting $\exp(-\beta) = Q_{CD}/Q_{DC}$, one easily obtains the expression:

$$\Pi_{l,l'}\exp(-\beta c_{l'} n_F) = \Pi_{l',l}\exp(-\beta c_l n_F) \ , \tag{2.29}$$

from where the unique solution to the fixed point equation

$$\Pi_{l,l'}\mathcal{P}^*_{l'} = \mathcal{P}^*_l \ , \tag{2.30}$$

is easily found to be:

$$\mathcal{P}^*_l = Z^{-1}\exp(-\beta c_l n_F) \ , \tag{2.31}$$

where Z is the analogue of the canonical partition function

$$Z = \left[\frac{Q_{CD} + Q_{DC}}{Q_{DC}} \right]^{n_F} \quad . \tag{2.32}$$

Note that Eq. (2.29) expresses the "detailed balance" condition, which is thus proved to be satisfied. As it is well-known [118], the canonical probability distribution density (2.31) is the unique density that maximizes the lack of information (entropy), $S = -\sum_l \mathcal{P}_l \ln \mathcal{P}_l$, among those (compatible) densities that share a common value for the macroscopic average of cooperation $c = \sum_l \mathcal{P}_l c_l$. This provides a "generalized thermodynamic" meaning to the parameter β: it is no other than the intensive entropic parameter associated to cooperation, that is, the Lagrange multiplier [119, 120] associated to the restriction $c = \sum_l \mathcal{P}_l c_l$ on the compatible measures (canonical restricted maximization of entropy), that is:

$$\beta = \frac{\partial S}{n_F \partial c} \quad . \tag{2.33}$$

The parameter β simply measures how fast the entropy of the equilibrium macro-state increases versus global cooperation variations. Its formal role is that of an analogue of inverse thermodynamical temperature. Let us note that, at variance with many works in evolutionary game dynamics (see [16] and references therein) where an analogue of temperature is introduced *ad hoc* as a parameter entering into the definition of the (stochastic) strategic updating rules, the parameter β (2.33) is a kind of emergent property that characterizes the equilibrium macro-state, and thus is a function of the model parameters (not a model parameter itself).

The fluctuations of the micro-states cooperation c_l, namely $(n_F)^2(\sum_l(\mathcal{P}_l c_l^2) - (\sum_l \mathcal{P}_l c_l)^2)$ are given by $n_F c(1-c)$. This is the analogue of the heat capacity. The dependence on the game and network parameters $b, \epsilon, n_C/n_F, k$ of the fluctuations of cooperation is obtained by solving for the cooperation equilibrium value $\dot{c} = 0$ in (2.17), and plotted in Fig. 2.3(panel b) for $k = 4$, $\epsilon = 0$, and decreasing values of the ratio n_C/n_F.

Numerical results, and the mean-field failure.

In this subsection we compare the mean-field results with those obtained from Monte Carlo simulations implementing the updating rules on the dipole model with a random regular network structure for the fluctuating set $\mathcal{F}$.

In order to illustrate the Boltzmannian character of the stationary probability density $\mathcal{P}^*(l)$, we plot in Fig. 2.4 the numerical estimates of $\ln\left(\frac{P(c)}{g(c)}\right)$, where $P(c)$ is

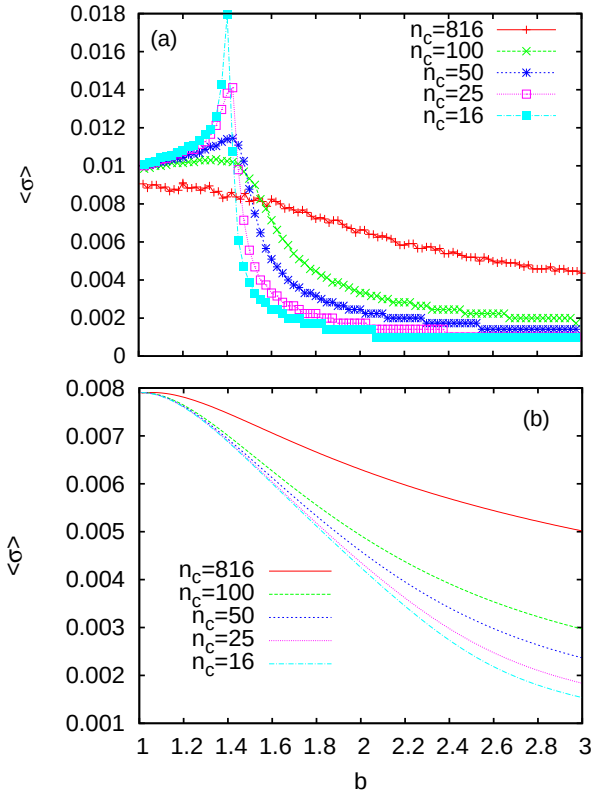


Figure 2.3: Fluctuations of cooperation in a random regular graph structure for the set $\mathcal{F}$. The upper panel (a) shows, for $k = 4$, $\epsilon = 0$, $n_F = 4000$, and a decreasing sequence of n_C/n_F values as indicated, the fluctuations of cooperation observed in Monte Carlo simulations. The lower panel (b) shows the mean-field predictions. The mean-field approach is shown in text to be unable to predict the observed phase transition. This qualifies network reciprocity as a true "critical" social phenomenon.

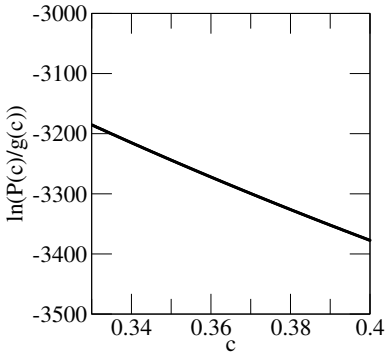


Figure 2.4: Plot of $\ln\left(\frac{P(c)}{g(c)}\right)$ versus cooperation c, showing the Boltzmannian character of the stationary probability density of microstates, for a random regular network structure for the set $\mathcal{F}$. The parameter values are $b = 1.1$, $n_F = 5000$, $k = 4$, $n_C = 500$, and $\epsilon = 0$. The results shown here correspond to 5×10^4 Monte Carlo steps (after a long enough transient), for each one of the 1.5×10^3 different network realizations and/or initial conditions.

the probability that a microstate has an average cooperation c (2.2), as inferred from the simulation results, and $g(c) = n_F!/((cn_F)!(n_F - cn_F)!)$ is the degeneracy of c (i.e., the number of microstates l such that $c_l = c$). The data correspond to a random regular network structure for the component $\mathcal{F}$ with degree $k = 4$, and parameter values $b = 1.1$, $n_F = 5000$, $n_C = 500$, and $\epsilon = 0$. As one can see from the perfect straight line shape of the plot, the data are fully consistent with the Boltzmann's density (2.31).

Though the system evolution is governed by dynamical rules (strategic updating) which "a priori" could be thought to lead to non-equilibrium behaviors, one finds that the asymptotic regime of the PD evolutionary dynamics in the dipole model is a true macroscopic equilibrium regime, where the formalism of generalized thermodynamics [120] applies.

The results of the asymptotic value of the average cooperation c versus the temptation to defect b are presented in Fig. 2.2 for (relatively small) values of n_C/n_F ranging from 0.025 down to 4×10^{-3}, but still satisfying the sufficient condition for the fixation of cooperation at BB2. The comparison with the mean-field predictions show that the mean-field approximation overestimates the cooperation value. Most notably, for very small values of n_C/n_F, the numerical results show, at about $b \simeq 1.4$, a fast decay of cooperation to values close to zero (thus suggesting the existence of a phase transition), while the corresponding decay for the mean-field prediction is smooth in the whole range.

To which extent the mean-field approximation fails for low values of the parameter n_C/n_F, can be appreciated by confronting its prediction, $n_F c(1 - c)$, for

the fluctuations of cooperation with the results from Monte Carlo simulations. In Fig. 2.3 (panel a) we see how a peak in cooperation fluctuations is revealed, when $n_C/n_F \to 0$, signaling the occurrence of a phase transition between two qualitatively different equilibrium macroscopic behaviors, that correspond to low and high temptation regimes. The mean-field assumption is thus qualitatively wrong if the payoff received from $\mathcal{C}$ by Big Brother 2 becomes negligible versus the size n_F.

The reasons for this qualitative failure of the mean-field approximation rely on the lattice reciprocity of internal interactions, which is totally absent in the mean field approximation. Let us remind here our remark above on the replicator-equation-type of effect of internal interactions in equation (2.8) because of the mean-field assumption. The transition signaled by the divergence of fluctuations at b^* reveals the onset of internal lattice reciprocity, a conclusion that we now substantiate (see also appendix 2.4.4 below).

For $b > b^*$, say in the low-temperature (high temptation) phase, the macro-state is dominated by fast defection invasions on the relatively few nodes that are instantaneous cooperators due to sporadic interactions with Big Brother 2. In the appendix 2.4.4 we show that, in the low c and low n_C/n_F regime, the BB-imitation events in a given node are typically separated by intervals of time of about c^{-1} time units large. In those large intervals when Big Brother 2's influence is null, the very few and mostly isolated instantaneous cooperators are quickly invaded by defector internal neighbors. In this regime lattice reciprocity has no chance to develop, and cooperation is only weakly sustained by the sporadic influence of BB2.

On the contrary, for $b < b^*$ (high temperature, or low temptation phase) the local fluctuations of the neighbors strategic field favor the building up of clusters of cooperators that resist invasions during time intervals that are comparable to the characteristic time intervals between BB-imitation events. Under these circumstances the "extra payoff" that BB2 receives from $\mathcal{C}$ does not anymore need to be high in order to sustain high levels of cooperation. Internal lattice reciprocity enhances the probability of highly cooperative micro-states, so that the macro-states below transition differs substantially from those of the high-temptation phase. This was not captured by the mean-field approximation, for these effects require a sizable likelihood of occurrence for the local fluctuations of the strategic field, and the neglect of them is all a mean-field approach is based upon.

To summarize the discussion of the results shown in figure, a random regular structure of interactions inside $\mathcal{F}$ is enough to support lattice reciprocity mechanisms that cannot be captured by a simple mean-field approach. The onset of lattice reciprocity in the dipole model is furthermore interpreted as a "thermodynamical" phase transition, in a rigorous formal sense (divergence of the fluctuations of an equilibrium extensive parameter, the cooperation c). One is then lead to a sensible and precise for-

mal framework where such a term as "social temperature" is not a vague metaphor, but it denotes a truly quantitative parameter, a legitimate (measurable, observable) social indicator.

2.4.3 What if BB1 and BB2 are directly connected?

If a direct connection between Big Brothers is added (for the set of initial conditions specified in section 2.3, and the conditions on parameters given ibidem), one must compare their respective payoffs to see who can invade the other. One easily finds that the payoff of the defector BB1 is higher than that of the cooperator BB2 provided the following condition holds:

$$\frac{n_C}{n_F} < c(b+\epsilon-1) - \epsilon + \frac{b-\epsilon}{n_F} \quad , \tag{2.34}$$

where c is the (instantaneous) average cooperation in $\mathcal{F}$. In this case, BB2 will be invaded with a non-zero probability. Once this eventuality occurs, no cooperator (in $\mathcal{F}$ or in $\mathcal{C}$) can later invade BB2 because all of them have lower payoffs, and fixation of defection in the whole network will occur. Note that as the average cooperation in $\mathcal{F}$ fluctuates, the condition above must be satisfied at the precise time when BB2 has chosen (by chance) to compare its payoff with BB1, and that due to the high connectivity of BB2 (which is now $n_C + n_F + 1$) the later event occurs with a very low probability for macroscopic values of n_F. In other words, the eventual invasion of BB2 from BB1 and the subsequent fixation of defection in the whole network can take on a very long time.

If the opposite condition holds, say if

$$\frac{n_C}{n_F} > c(b+\epsilon-1) - \epsilon + \frac{b-\epsilon}{n_F} \quad , \tag{2.35}$$

when BB1 has chosen to compare its payoff with BB2, then invasion of BB1 will occur with a non-zero probability. After this has occurred, BB1 becomes a fluctuating node (for it could be eventually invaded by an instantaneous defector in $\mathcal{F}$), but in the long term fixation of cooperation in the whole network will occur.

The introduction of a direct connection between Big Brothers in the Dipole Model makes fixation of opposite strategies on them impossible, and then asymptotic fixation on the whole network of either defection or cooperation will occur, depending on the relative size n_C/n_F of components $\mathcal{C}$ and $\mathcal{F}$.

2.4.4 Low c approximation.

In order to simplify expressions we assume hereafter $\epsilon = 0$ and $k = 4$, and denote $\delta = n_C/n_F$. For the case of a random regular graph structure of the fluctuating set $\mathcal{F}$, the probability $\Pi^{BB}_{C\leftarrow D}$ that an instantaneous defector node chooses to imitate Big Brother 2 (invasion event from BB2) is, to first order in n_F^{-1},

$$\Pi^{BB}_{C\leftarrow D} = \frac{1}{(k+2)}\frac{c+\delta}{(1+\delta)b} \tag{2.36}$$

while the probability $\Pi^{BB}_{D\leftarrow C}$ of an invasion event from BB1 to an instantaneous cooperator node in $\mathcal{F}$ is, to first order in n_F^{-1},

$$\Pi^{BB}_{D\leftarrow C} = \frac{c}{(k+2)} \tag{2.37}$$

Thus, for $\delta \leq c$, typical intervals between invasion events from Big Brothers in a node are (of the order of) c^{-1} time units large. For large values of the temptation, where the value of c is expected to be very small, the dynamics is consequently dominated, for typically very large intervals of time, by internal strategic interactions. Let us analyze them.

The internal neighbors of a cooperator i are overwhelmingly likely instantaneous defectors in this "low c" regime, so that i will be quickly invaded by them. The only chance for it to resist invasion would be that its instantaneous neighborhood microstate had at least two cooperator neighbors and that $b < (3/2)$ (note that in this strategic configuration, the payoff of i is $P_i = 3$ and that of its typical defector neighbors is $2b$). These neighborhood microstates (cooperative clusters) are so rare fluctuations that low values of the temptation b are necessary for their non-negligible occurrence. Provided b is below the transition value, the resilience to invasion (lattice reciprocity) of cooperative clusters enhances the likelihood of these fluctuations, which in turn reinforces the clusters resilience, and so on. This positive feedback mechanism of cooperative fluctuations enhancement is thus what triggers the transition to highly cooperative macro-states, and qualifies lattice reciprocity as a critical social phenomenon.

2.4.5 General case: $\mathcal{F}$ is a general random graph.

Let us pay attention to the behavior of the model when $\mathcal{F}$ is a graph characterized by a given degree distribution $P(k)$. While the random regular network (RRN) considered in the previous section corresponds to the distribution $P(k) = \delta(k - k_F)$, now we

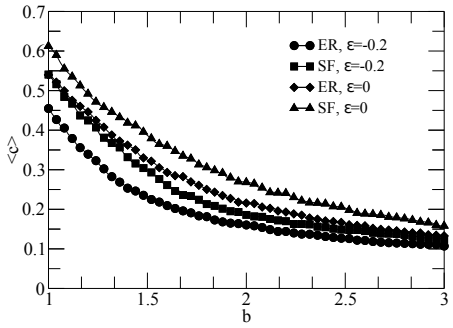


Figure 2.5: Fraction of cooperators as a function of the temptation to defect b in a ER (BA) topology for the set $\mathcal{F}$, with $n_F = 4000$, $n_C = 3804$ and $\epsilon = 0,\ -0.2$. Each point is averaged over 1600 realizations (40 networks, 40 initial conditions).

consider two cases: when $\mathcal{F}$ is a Erdös-Rényi (ER) or Barabási-Albert (BA) network. While in homogeneous ER graphs, $P(k)$ follows a Poisson distribution centered at k_μ, in BA networks the degree distribution follows a power-law, and the different connectivity distribution determines the system behavior; as will be detailed below, the hubs of the set $\mathcal{F}$ in the BA case are a deciding factor.

With regard to the roots of cooperation, highly connected nodes are critical elements in the intensity of lattice reciprocity mechanism. This is mainly due to the effect that the network topology has on the distribution of strategies: the formation of clusters of cooperators prevents the invasion of the strategy D in them. If the connectivity distribution is such that there are very connected vertices, the hubs having an initial strategy C are likely to form cooperative clusters, and the change of strategy of these hubs is very unlikely. According to this argument, BA networks will promote cooperation more than ER ones.

We have numerically studied the system by performing Monte Carlo simulations after implementing in the fluctuating set $\mathcal{F}$ a network topology generated by a routine taken from [140]. This algorithm provides a random network such that $P(k)$ depends on a continuous parameter α: $\alpha = 0$ (*resp.*, 1) generates a Scale-Free (*resp.*, Poisson distribution) graph. We have scanned the parameter space $\{\alpha,\ \epsilon,\ b\}$, with $0 \leq \alpha \leq 1$ (SF to ER). The condition $b > 1 + \epsilon(k_F + 1)$, in practice $\epsilon \geq 0$, involves PD. In the same way, the constraint $n_C > b(k_F + 1) - \epsilon n_F$ implies a large number of BB2 stabilizer agents (set $\mathcal{C}$) for a high connectivity, as in the BA case.

Regarding macroscopic cooperation, the numerical results obtained do not show any abrupt transition when the parameter α varies: the gradual change from $\alpha = 0$ (SF) to 1 (ER) implies a gradual change in c. The results of the asymptotic value of the average cooperation c can be seen in figure 2.5. The figure shows greater

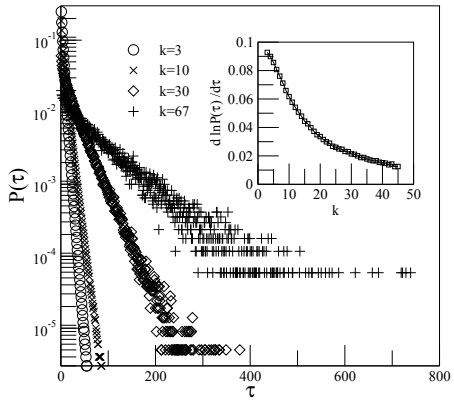


Figure 2.6: The main panel shows the probability for a fluctuating node to remain as a cooperator for a period τ, for $\epsilon = -0.2$ (PD), $b = 2$ and a BA structure for the set $\mathcal{F}$. The permanence probability follows a power law. The inside panel shows the slopes of the main panel, for all connectivity classes. Nodes with higher connectivity are more resilient to strategy change, which provides a microscopic basis for lattice reciprocity. See the text for further details.

cooperation c in BA graphs than in ER ones, as we argued above. Moreover, the dependence of c on ϵ (always for $\epsilon < 0$, *i.e.*, PD game) turns out to be approximately linear for all values of b, depending very little on the type of network:

$$c(\epsilon) = c(\epsilon = 0)(1 + \kappa\epsilon), \ (\epsilon < 0), \ \kappa \approx 1.$$

This relation is valid for all values of b, as verified by low variance of

$\phi_\epsilon(b) = \frac{c(b)|_{\epsilon'}}{c(b)|_{\epsilon=0}}$, (e.g. $\sigma^2(\phi_\epsilon) = 0.004$ for $\epsilon' = -0.2$).

Fluctuations.

Given that the elements of the set $\mathcal{F}$ always have a non-zero probability of strategy change, we have studied the distribution of the *characteristic cooperation time* τ_c, which is defined as the average period that a node remains as cooperator. The simulations showed that τ_c is independent of α, that is, the characteristic cooperation time for a given connectivity is independent of the network topology. Figure 2.6 shows the probability $P(\tau)$that a node of the set $\mathcal{F}$ keeps strategy C during a period τ, in a BA graph, for $\epsilon = 0$, $b = 2$ and different connectivity classes. The probability of permanence decreases exponentially over time and, as expected, nodes with higher connectivity have higher characteristic cooperation time, which **constitutes a manifestation of the microscopics roots of lattice reciprocity**, and explains why BA

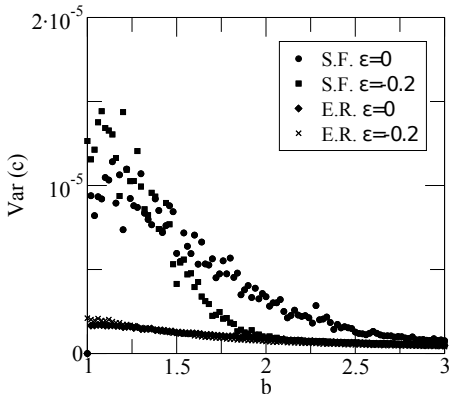


Figure 2.7: Fluctuations of cooperation in a ER (BA) topology for the set $\mathcal{F}$, with $n_F = 4000$, $n_C = 3804$ and $\epsilon = -0.2,\ 0.2$.

networks show higher cooperation rates than ER ones. The inset of figure 2.6 shows the slopes of main figure, and represents the coefficient λ of the relation:

$$P(\tau) \propto e^{-\lambda\tau} \tag{2.38}$$

As shown, $\lambda(k)$ is monotonically decreasing, which highlights the relationship between connectivity and network reciprocity.

As in the previous sections, we study the fluctuations of cooperation along stochastic trajectories. Nevertheless, for a ER or BA structure for the set $\mathcal{F}$, the analogy between the cooperation rate variance and the heat capacity is not as justified as in the random regular structure and, therefore, the study is restricted to Monte Carlo simulations. As shown in figure 2.7, BA structures presents more fluctuations than ER ones.

2.5 Heat transfer: Two dipoles in contact.

The system shown so far constitutes a Markov process. We have seen in section (2.4.2) that, in the scope of applicability of mean field approximation (as when $\mathcal{F}$ is a RRN), the model satisfies detailed balance condition (2.29). Therefore, in the stationary state we deal with a reversible Markov process described by the Boltzmann distribution (2.30). However, the analogy of the model with a real particle system has a drawback: the equivalent of energy is the total number of cooperators cn_F, which means that the model is not conservative. Despite that, it's hard not to wonder how two dipoles behave when they come into contact. The interest of this extension of the

original model is wide: On one hand, we can study the validity of the social temperature concept and evaluate its relation to the physical temperature. Furthermore, the expanded model allows to study the interaction between two evolutionary dynamics, both with two pure strategies.

2.5.1 The double dipole model.

In order to simulate the heat transfer, we consider two dipoles D_1 and D_2, composed by respective fluctuating sets F_i, the hubs $BB1_i$ and $BB2_i$ and the cooperating sets C_i. The subscript i indicates the dipole subsystem to which it belongs, $i = 1, 2$. Both dipole subsystems have the same size: $n_{C_1} = n_{C_2} = n_C$, $n_{F_1} = n_{F_2} = n_F$. Therefore, from now on, these sizes will be denoted by n_C and n_F. The parameter ϵ is common for both dipoles, and the difference in payoffs is determined by the respective parameters b_i. The definition of temperature, $\beta^{-1} = -(ln\frac{Q_{CD}}{Q_{CD}})^{-1} = f(b)$, leading to the Boltzmann distribution (2.30), takes us to use b as an independent variable. This procedure allows to control the temperature of each of the two dipoles when they reach their respective stationary states.

Let I' be the set of initial conditions: the Big Brothers $BB1_i$ are defectors, Big Brothers $BB2_i$ are cooperators, and all nodes in sets C_i are also cooperators. There is not constraint to initial strategies for elements in sets F_i. To ensure that I' is a invariant set of the evolutionary dynamics, we maintain the restrictions of section 2.3, now take the form: $n_C > b_{max}(k_F + 1) - \epsilon n_F$ and $b_i > 1 + \epsilon(k_F + 1)$, where $b_{max} = sup\{b_i\}$.

2.5.2 Dynamics.

The subsystems D_1 and D_2, after setting their parameters, are left to evolve according to the usual dynamic to reach equilibrium. Once reached their respective stationary states $(t = 0)$, we connect both fluctuating sets as follow: We choose at random a node i_1 of F_1 and a neighbor j_1, in the same way, we choose another node i_2 of F_2 and his neighbor j_2. Subsequently, we remove links connecting those nodes, and connect i_1 to i_2, and j_1 to j_2. We repeat this process $k\kappa n_F/2$ times, where κ is a coupling parameter. After that, the system will evolve again according to the habitual dynamics. When a node i of dipole n choose a neighbor j of dipole m to play, both agents take n's pay-off matrix, i.e. the temptation to defect is b_n. We have studied the evolution of the system, the observables and stationary states.

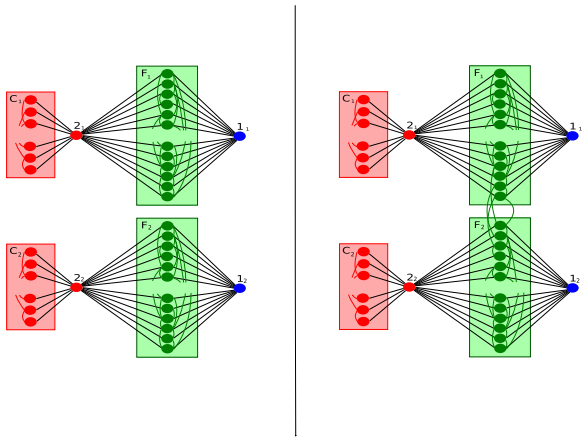


Figure 2.8: Structure of the Double Dipole Network. On the left side is represented the model before thermal contact: Without any link between the sets F_1 and F_2. On the right, after the thermal contact: With links connecting nodes of F_1 to nodes of F_2. Nodes 1_i and 2_i (Big Brothers) are linked to all nodes in F_i. Furthermore, nodes 2_i are also connected to all nodes in C_i. Nodes in sets F_i, C_i are internally connected, forming a defined network. Initial conditions (on left side) are indicated by colors: red for cooperators (nodes 2_i and set C_i), blue for defectors (nodes 1_i), and green means arbitrary (sets F_i). See the text for further details.

2.5.3 Effective temperature.

The intrinsic observables of the extended model are c_i. However, it is possible to study the final temperature of each subsystem by introducing the concept of effective temperature.

By $\exp(-\beta) = Q_{CD}/Q_{DC}$, we know that stationary state temperature in an isolated dipole, when $\mathcal{F}$ is a RRN, is a function $f(b, \epsilon, n_C, n_F)$. Besides, if all other parameters keep fixed, the cooperation is given by a monotonically decreasing function $c(b)$. Therefore, if we keep constant all other parameters, each value of c corresponds to a value of b.

We define the effective temperature of a subsystem $D_i(b_i, \epsilon, n_C, n_F)$ as the value of β^{-1} that would fall to his level c_i of cooperation in an isolated dipole $D(b_i, \epsilon, n_C, n_F)$ in stationary state. The effective temperature is applicable to each subsystem, once put in thermal contact and reached a new stationary state.

We can study now the temperatures of the subsystems before and after putting them in contact. The problem is that, knowing the asymptotic value of c_i, we can not infer β for any topology in set $\mathcal{F}$, unless we use an approximation as MF. Therefore, we use the observed value of c as relatives of the temperature. The Big Brothers'

influence on fluctuating sets F_i is determined by its payoffs, which in turn are a function of c_i, b_i, n_C, n_F and ϵ. An interpretation of the influence of BBs on F_i is that the sets F_i are thermodynamic systems in contact with respective heat baths. The influence of the heat baths remains after putting the dipoles in contact, therefore the effective temperature of both subsystems need not be equal once it has reached the new stationary state, unless the coupling parameter is $\kappa \geq 0.5$.

In the model, there are only two strategies or accessible levels by element, and there exists a extern source of cooperation: The sets C_i. It is therefore possible to have configurations in which there are more elements adopting cooperating strategy than defect one and the system can be characterized by a negative effective temperature. This is only possible if the external field (heat bath) has enough influence, that is, for high values of n_C. Negative temperatures are possible in both dipole, before and after heat contact, for high enough values of n_C.

2.5.4 Mean-field approximation.

For an isolated dipole, in subsection 2.4.2 we assumed that the fraction of instantaneous cooperators $c[i]$ in the neighborhood of a node i, is the fraction c of the whole set $\mathcal{F}$. In order to generalize MF approximation, now we make $c[i]$ to correspond to the weighted average of the cooperation: We assume that for a fluctuant node i_1 in the dipole D_1, in thermal contact κ with another dipole D_2, the fraction of cooperators in the neighborhood of i_1 is $c[i_1] = (1-\kappa)c_1 + \kappa c_2$. In the same way, we take $c[i_2] = \kappa c_1 + (1-\kappa)c_2$. In the model, according to 2.5.2, when a node of dipole n chooses a neighbor to play, they take b_n, in this approximation we assume the average temptation to defect $\bar{b} = (b_1 + b_2)/2$ for heat interactions (between two fluctuating sets).

Under these assumption, the payoffs of Big Brothers $BB1_i$ and $BB2_i$ are given by $P_{1i} = b_i c_i n_F$ and $P_{2i} = n_C + c_i n_F + \epsilon(1-c_i)n_F$, while the payoffs of a cooperator node and a defector node at F_i are:

$$\begin{aligned} P_{c1} &= k(\lambda c_1 + \kappa c_2) + 1 + \epsilon(k(\lambda(1-c_1) + \kappa(1-c_2)) + 1) \ , \\ P_{c2} &= k(\kappa c_1 + \lambda c_2) + 1 + \epsilon(k(\kappa(1-c_1) + \lambda(1-c_2)) + 1) \ , \\ P_{d1} &= (k\lambda c_1 + 1)b_1 + k\kappa c_2 \bar{b} \ , \\ P_{d2} &= k\kappa c_1 \bar{b} + (k\lambda c_2 + 1)b_2 \ , \end{aligned} \tag{2.39}$$

where $\lambda = 1 - \kappa$.

$b_i > 1$ implies $P_{di} > P_{ci}$. Without loss of generality, we assume that $b_2 > b_1$. If there is not constraint to κ, can not ensure that $P_{ci} > P_{cj}$ nor $P_{di} > P_{dj}$ for any $i \neq j$. For nodes in F_i, each time step, the probabilities Q_{DCi} (a cooperator changes its strategy to defector) and Q_{CDi} (a defector changes its strategy to cooperator) are:

$$\begin{aligned}
Q_{DC1} &= \frac{P_{11} - P_{c1}}{(k+2)n_F\Delta} \\
&\quad + k\frac{\lambda(1-c_1)(P_{d1} - P_{c1}) + \kappa(1-c_2)(P_{d2} - P_{c1})H(P_{d2} - P_{c1})}{(k+2)^2\Delta}, \\
Q_{CD1} &= \frac{P_{21} - P_{d1}}{(k+2)Bn_F\Delta} + k\frac{\kappa c_2(P_{c2} - P_{d1})H(P_{c2} - P_{d1})}{(k+2)^2\Delta}, \\
Q_{DC2} &= \frac{P_{12} - P_{c2}}{(k+2)n_F\Delta} \\
&\quad + k\frac{\kappa(1-c_1)(P_{d1} - P_{c2})H(P_{d1} - P_{c2}) + \lambda(1-c_2)(P_{d2} - P_{c2})}{(k+2)^2\Delta}, \\
Q_{CD2} &= \frac{P_{22} - P_{d2}}{(k+2)Bn_F\Delta} + k\frac{\kappa c_1(P_{c1} - P_{d2})H(P_{c1} - P_{d2})}{(k+2)^2\Delta},
\end{aligned} \tag{2.40}$$

where $H(x)$ represents Heaviside's step function, that takes value 1 if $x > 0$, 0 otherwise.

After a time step, the fraction of cooperators in set F_i is given by:

$$c_i(t+1) = c_i(t) + (1 - c_i(t))Q_{CDi} - c_i(t)Q_{DCi} \ . \tag{2.41}$$

Now, we can replace (2.39) and (2.40) in (2.41) to achieve two coupled finite difference equations for c_i. These equations provide the evolution of the system and the stationary state, according to MF approximation.

2.5.5 Numerical results.

We compare the mean-field results, obtained by evaluating the expression (2.41) iteratively, with experimental ones obtained from Monte Carlo simulations. The values of c_i, once stationary state is reached, are related to the effective temperatures according to 2.5.3.

In order to study transition phenomena, we measure the cooperation evolution from the stationary states for the isolated dipoles D_i to the stationary state after thermal contact. Fig. 2.9 shows the evolution of cooperation, by comparing simulations

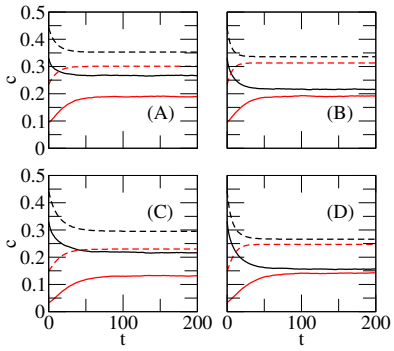


Figure 2.9: Evolution of macroscopic cooperation after putting in contact two dipoles ($t = 0$), with $k = 4$, $n_F = 4000$, $\epsilon = 0$, $n_C = 100$, $b_1 = 1.1$, and for different values of $b_2 = 1.5$ (A,B), $b_2 = 1.8$ (C,D) and coupling parameter: $\kappa = 0.25$ (A,C), $\kappa = 0.5$ (B,D). Solid lines represent the results of the simulations, while dotted lines represent the MF approximation. Black lines represent to dipole D_1 and red lines to dipole D_2. Time unit corresponds to a Monte Carlo step. See the text for further details.

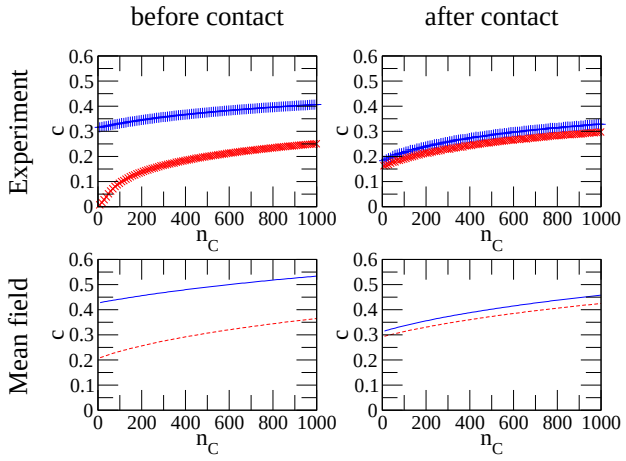


Figure 2.10: Macroscopic cooperation levels c for both subsystems, versus the size n_C of pure cooperators sets, with $k = 4$, $n_F = 4000$, $\epsilon = 0$, $\kappa = 0.5$ and RRN structures in fluctuating sets F_i. **Upper panels** show the results of Monte Carlo simulations, after a long enough transient, averaged over 5×10^4 steps and 5×10^3 different network realizations. The temptation parameters of both sybsystems are fixed at $b_1 = 1.1$ (blue, +) and $b_2 = 1.5$ (red, x). The upper-left panel corresponds to the stationary state before putting in contact the dipoles, and the upper-right panel corresponds to the stationary state after exchanging $1.5n_F$ links. **Lower panels**: Mean field estimate for the same values of parameters.

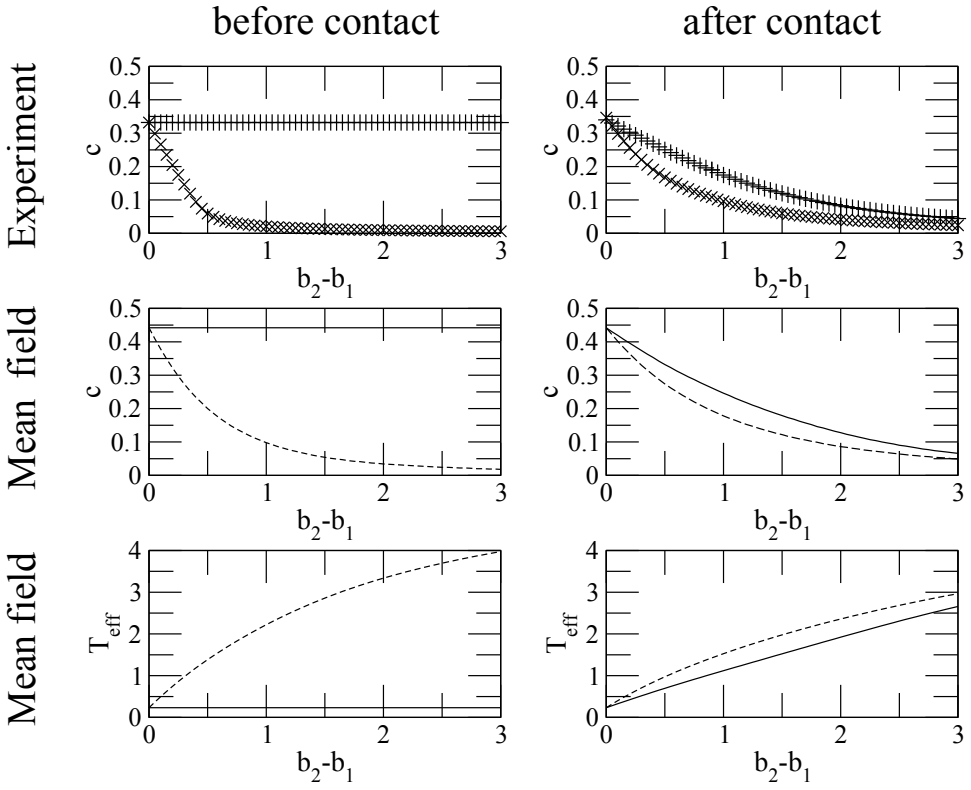


Figure 2.11: **Upper panels**: Cooperation levels in dipoles D_1 (symbols +) and D_2 (x) versus the difference $b_2 - b_1$ of temptation parameters. The rest of parameters have been fixed to $b_1 = 1.1$, $k = 4$, $n_F = 4000$, $n_C = 100$, $\epsilon = 0$ and $\kappa = 0.5$. Sets F_i are endowed with a random regular graph structure. Symbols correspond to the value of c once reached stationary state, averaged over 5×10^4 Monte Carlo steps and for 5×10^3 different networks, before (left) and after (right) putting in contact the dipoles. **Middle panels**: Mean-field estimation of cooperation levels in dipoles D_1 (solid lines) and D_2 (dashed lines), before (left) and after (right) contact, for the same values of parameters. **Lower panels**: Mean-field estimation of effective temperatures T_{eff} of dipoles D_1 (solid lines) and D_2 (dashed lines), before (left) and after (right) contact, for the same values of parameters.
In all panels, the temptation parameter b_1 of dipole D_1 remains constant (+, solid lines), while varying the respective parameter b_2 of dipole D_2 (x, dashed lines).

results with MF estimate, for different values of coupling parameter κ and $b_2 - b_1$ (in monotone bijection to initial temperature difference). One can observe that the mean-field predictions provide higher cooperation values. This failure of MF approximation, as in the section 2.4.2, highlights the importance of lattice reciprocity, which is ignored in MF approximation. As expected, this MF estimate, provided by the finite difference equations (2.41), is coherent at $t = 0$ with the MF estimate of the section 2.4.2, provided by the fixed point of the differential equation (2.16).

The results for the asymptotic values of the average cooperation, that is, the values of c_i at the stationary states before and after the thermal contact, versus the size of fluctuating sets C_i are represented in figure (2.10). As shown, for low values of n_C the difference in cooperation between the dipoles is smaller than for high values of n_C. This is a consequence of the BB's influence, which grows as do the size of the set of cooperating nodes. Put another way, the influence of heat bath increases with n_C/n_F. This has a direct explanation because the size of n_C does not affect the payoff of fluctuating nodes, but it does increase the payoff of cooperator hubs $BB2_i$.

The upper panel of figure 2.11 shows the average cooperation c versus the difference in the temptation to defect $b_2 - b_1$, once fixed $n_C = 100$, $n_F = 4000$, $\epsilon = 0$, $\kappa = 0.5$, $b_1 = 1.1$ and taking b_2 as independent variable. For high values of b_2, the proportion of cooperators of both subsystems after heat transfer and once reached the new stationary state tend to equate: The heat bath effect decreases with b_{max}, and cooperation final ratios tend to be equal when b_2 increases. When we use the MF approximation, the same results can be analyzed from the perspective of effective temperature, as shown in the lower panel of figure 2.11. Effective temperature decreases as c_i increases, and then it increases when b_i increases. The final temperature difference depends on the value of the coupling parameter κ, a higher κ implies lower difference. However, $\kappa > 0.5$ means that in average, nodes have more links to the other fluctuating set than to his. Therefore, for small enough n_c size and $\kappa > 0.5$, the final temperature of initially hotter dipole will be lower than the opposite dipole final temperature.

2.6 Prospective remarks

The plausibility of a thermodynamical perspective on evolutionary game dynamics studies is not a new issue, for it is somehow implicit (or at least connatural) to a body of research literature on statistical mechanics of strategic interactions [16, 121]. What our simple analysis here shows is that it can sometimes be strengthened up to a formal interpretation of quantitative macroscopic social indicators as thermodynamic quantities. In the extent that it helps to understand and to quantitatively characterize the phenomenology of social and economical models, it should be recognized as

a powerful theoretical perspective. What is even more important, this perspective emphasizes the central role of quantitative (experimental, observational) studies in social sciences, and could provide, in those contexts, alternate valuable meanings to quantitative social indicators and even suggestions for new and better ones.

Any "general-physics" trained scientist recognizes that entropy reasoning is an extraordinary powerful tool for the analysis of macroscopic behavior in (material) traditional-physics systems. It turns out that some of the models (at least a bunch of interesting ones) of social phenomena are to a large extent amenable to a macroscopic description where thermodynamical concepts have proved to be essential. Of course, some notions like *e.g.* "First Law of Thermodynamics" could be often absent in these new contexts. However we emphasize that the absence of energy as a variable in social models is not a shortcoming for the applicability *mutatis mutandi* of many aspects of the thermodynamical formalism to these models. A word of caution is nevertheless worth here regarding typical system sizes in controlled social experiments, where finite size effects could be hugely determinant. Also, one should not expect always social processes to be amenable to equilibrium descriptions, what makes them even more interesting from the physicists point of view.

Nowadays, it is somewhat generally accepted that physics in general, and statistical physics in particular, offers a powerful tool-box for problem solving in social sciences and many other areas. Recent trends in cognitive science [122] have correctly emphasized the power of the "diversity of perspectives" in problem solving, so it does not come as a surprise that adding physical perspectives to social models may sometimes pave the way to the needed breakthrough. Perhaps one should also wonder about the possibility of reverse flow in these interdisciplinary approaches to social sciences. After all, the proper use of a tool helps to its reshaping, and one could perhaps expect some kind of feedback. In other words, is there any new physics that we can learn from the study of social and economic complex systems? Only the recourse to empirical and quantitative methods in the study of social phenomena may likely give clues for sensible answers to this question.

Chapter 3

Cooperation in changing environments: Irreversibility in the transition to cooperation in complex networks.

In this chapter, we study the evolutionary dynamics of the prisoner's dilemma game in differents complex networks, focusing on its reversibility under adiabatic variations of the payoff matrix parameter b (temptation to defect). We find that, for the networks considered, the process is reversible provided it is kept away from the absorbing states. Nevertheless, irreversibility appears when the level of cooperation reaches a tipping point, emerging a hysteresis cycle whose shape depends on the underlying topology.

3.1 Introduction.

Evolutionary dynamics has been widely used to describe the evolution of biological, economic and social systems [12]. The replicator dynamics of *evolutionary game theory* (EGT) provides a powerful tool to study the progress of strategies through the lens of evolution [79, 17]. In this respect, one of the hot topics of the evolutionary game dynamics is the understanding of the observed evolutionary survival of cooperative behavior among individuals despite selfish actions provide higher fitness (reproductive success). Possibly, the most used EGT model to formalize the emergence of cooperation is iterated *Prisoner's Dilemma* (PD), a symmetric two-player two-actions game where each player choose one of the two available actions: coop-

eration or defection: A cooperator receives R from another cooperator and S from a defector; a defector receives T from a cooperator and P from another defector; payoffs satisfy $T > R > P > S$. Several studies on the iterated PD on complex networks [19, 81, 82, 83, 84, 85, 86, 87, 16, 90, 101, 102, 103, 104] show that the cooperation level depends strongly on the topology of the network. The existence of cooperation enhancing mechanisms based on the interaction structure now is widely accepted: The clustering of cooperators could provide high enough payoff to the cooperator nodes to resist invasion of defectors, even when defection is favored by the one-round two-players game analysis. For small values of $P - S$ (*i.e.*, $P - S \ll T - R$), cooperation decreases slowly when $T - R$ increases from zero, and becomes zero at a value of $T/R > 1$ that depends on the network considered.

Recent studies of replicator dynamics [109] on graphs show that fixation of cooperation on certain nodes occurs after transients, in which the trajectories are characterized by a partition of the network into three sets: the set C of pure cooperators (nodes where cooperation is fixed), the set D of pure defectors (nodes where defection is fixed), and the set F of fluctuating nodes (nodes that never reach an unchanging action). Furthermore, robustness of cooperation in the evolutionary PD on complex networks has been recently studied [112], showing that the level of cooperation under different network structures is robust against variation of initial conditions. The aim of the present study is to investigate evolutionary PD on complex networks in changing environments, in particular its reversibility under variations of temptation to defect T, and to determine how topology affects reversibility.

3.2 The model.

We consider a two-players two-actions game, where each player chooses one of the two available actions, cooperation or defection: A cooperator earns R when playing with a cooperator, and S when playing with a defector, while a defector earns P when playing with a defector, and T (temptation to defect) against a cooperator. When $T > R > P > S$, the game is called Prisoner's Dilemma (PD), while if $T > R > S > P$ it is called Snowdrift Game (SG). In this work we study a variant of PD called weak Prisoner's Dilemma, placed in its boundary respect to SG, that is $T > R > P = S$. In PD (including weak variant), whatever the opponent's action, the payoff is never higher for cooperation, and a rational agent should choose defection. Still, two cooperator agents receive higher payoff ($2R$) than two defector ones ($2P$), which leads to the dilemma. Provided the relative selective advantage among two individuals depends on their payoff's difference (see below), we can normalize without loss of generality the pay-off matrix taking $R = 1$ and fix the punishment $P = 0$. Then only a parameter $T = b > 1$ is a system variable.

In this study we implement the following replication mechanism: At each time step, each agent i plays once with each one of its neighbors (*i.e.* agents connected to i) and accumulates the obtained payoffs, P_i. After that, the individuals, i, update synchronously their actions choosing a neighbor j at random, and comparing their respective payoffs P_i and P_j. If $P_i \geq P_j$, nothing happens and i preserves its action. Otherwise, if $P_j > P_i$, i adopts the action of its neighbor j with probability $\Pi_{ji} = \eta(P_j - P_i)$. Next, all payoffs are reset to zero. Here, η is a positive real number, related to the characteristic inverse time scale: the larger it is, the faster evolution takes place. We consider that players and connections between them are given by a fixed graph where agents are represented by nodes, and a link between nodes indicates that they interact. We choose here the maximum value of η that preserves the probabilistic character of Π_{ji}, that is, $\eta = (\max\{k_i, k_j\}b)^{-1}$, where k_i is the number of neighbors of agent i (connectivity or degree). This choice, introduced in [18], slows down the invasion processes from or to highly connected nodes (hubs), with respect to the rate of invasion processes between poorly connected nodes.

Our aim is the study of the reversible (or irreversible) character of cooperation level c under the variation of the temptation to defect parameter b, where c is defined as the number of cooperator nodes divided by the total population $c = N_c/N$. In order to study the system's behavior, we choose an initial value of $b = b_0$ such that the asymptotic cooperation value c is close to a half: $c(b_0) \simeq 0.5$. Once the system has reached stationary state, we decrease b in a quasi-static way, that is, in steps $\Delta b < 0$ small enough to ensure that the system remains very close to equilibrium. Along this process, we compute the stationary value of cooperation $c(b)$ for each value of b. To avoid getting stuck in the absorbing states we deal with large enough networks sizes ($N > 10^5$), considering that fluctuations decrease according to the square root of the system size. Once the system has almost reached the absorbing state $c = 0$, we reverse the sign of the increase in b, *i.e.* $\Delta b > 0$, to almost reach the other absorbing state $c = 1$, and then again decrease b to complete the cycle.

3.3 Results.

To study the influence of network topology in the reversibility of the process, we consider three different network models: Random Regular Graphs (RRG), Erdös-Rényi and Scale-free networks. In the case of RRG (*i.e.*, random networks of fixed degree k, which means that every node has the same number of neighbors), adiabatic cycles are identical; that is, **the behavior observed in the numerical simulations with RRG corresponds to a reversible process**.

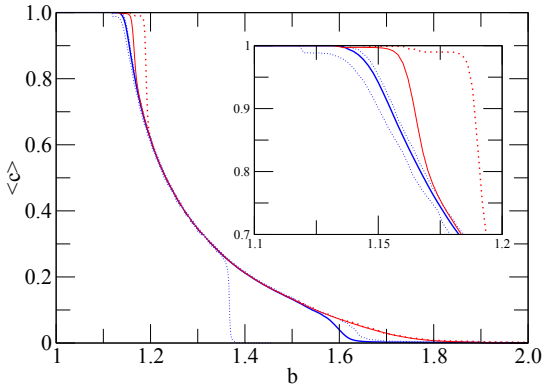


Figure 3.1: Cooperation level $\langle c \rangle$ versus the temptation to defect b averaged over 10^3 ER networks (solid lines) and envelopes (pointed lines). Red lines represent semicycles with increasing b and blue lines represent semicycles with decreasing b. The network size is $N = 1.2 \times 10^5$. See the text for further details.

3.3.1 Erdös-Rényi networks.

Erdös-Rényi (ER) networks are random graphs characterized by a binomial degree distribution of any particular node, this distribution is Poisson for large number N of nodes. To study the processes reversibility, we have performed numerical simulations in 10^3 independent networks of size $N = 1.2 \times 10^5$ generated through Erdös-Rényi algorithm. As outcome, for reduced cycles, that is, when the return points are far from absorbing states ($1 - N_c(b_{min}) \gg 1$, $N_c(b_{max}) \gg 1$) the processes are reversible and the level of cooperation is independent of the sign of the increase in b. Nevertheless, when return points are close enough to the absorbing states ($c(b_{min}) \approx 1, c(b_{max}) \approx 0$), ER networks show a dramatic irreversibility. In fact, once the level of cooperation reaches a tipping point, all processes are irreversible. In particular, there is a strong resilience of cooperation (defection) when increasing (decreasing) the value of b. However, the backward and forward transition curves are identical for intermediate values of cooperation. The proximity ϵ of the tipping points $c(b_{min}), c(b_{max})$ to the absorbent states in both ends of cycle turns out to be similar: $1 - c(b_{min}) = \epsilon \approx c(b_{max})$ and, for the networks size used, it takes on the value $\epsilon \approx 2 \times 10^{-3}$.

As a result, once the population has reached a cooperation level above (below) a tipping point, the system shows a reticence to retrieve the past level of cooperation when the parameter b increases (decreases). This phenomenon is independent of the particular ER network, being observed in all network realizations. Figure 3.1 shows the level of cooperation $\langle c \rangle$ versus the temptation to defect b, averaged over 10^3 realizations in distinct ER networks. Different realizations show different b-

increasing and b-decreasing curves, whose envelopes are depicted as dotted lines in Figure 3.1. Remarkably, the dispersion of the different curves is much larger for the b-decreasing direction.

3.3.2 Scale-free networks.

Scale-free (SF) networks are random graphs whose degree distribution $P(k)$ follows a power law, that is, $P(k) \sim ck^{-\gamma}$. We ran simulations in 5×10^3 independent networks of size $N = 1.2 \times 10^5$ generated through the Barabási-Albert algorithm. Although most of the SF networks show nearly reversible behavior, around 5% of networks show a strong hysteresis. Nevertheless, irreversibility in SF networks should not be considered as a rare event: Increasing the network size increases the proportion of networks that show irreversible behavior. The explanation for this fact is that the use of larger networks allows to approach closer the absorbing states $c = 0, 1$ without getting stuck in them. Based on this argument, we have separated realizations showing a reversible behavior from irreversible ones. In these latter cases, hysteresis shows up only for low values of b; in other words, when cooperation is very small, backward and forward $c(b)$ curves are almost identical. Moreover, the behavior of the system in b-increasing semicycles is always similar, the cooperation level $c(b)$ taking approximately the same value in all realizations, regardless they are reversible or irreversible. On the contrary, $c(b)$ curves are different for different (irreversible) realizations in b-decreasing semicycles, and show a substantially larger dispersion that those of ER networks.

The results of the average cooperation level $\langle c \rangle$ as a function of the temptation to defect b, for SF networks showing irreversible behavior, are presented in figure 3.2. The return points b_{min}, b_{max} were chosen such that $c(b_{max}) = 1 - c(b_{min}) = \epsilon$, for a value of $\epsilon = 10^{-3}$. Note that, despite the small value of ϵ, the network size N is large enough to asses that we are not dealing with pathological cases, since a value $c = 0.001$ involves a number of cooperators $N_c = 120$. In the same way, $c = 0.999$ implies 120 defector nodes. As shown in envelopes (dotted lines), the degree of irreversibility varies greatly from each realization. Specifically, irreversibility depends on the particular network, since for a given network repeated cycles share approximately the same $c(b)$ curves for a given (forward or backward) direction. A most remarkable feature of the irreversibility in SF networks is that, for irreversible network realizations, the value of the temptation to defect needed to reach a cooperation level of $c = 10^{-3}$ is $b_{min} < 1$, that is to say, outside the PD game range.

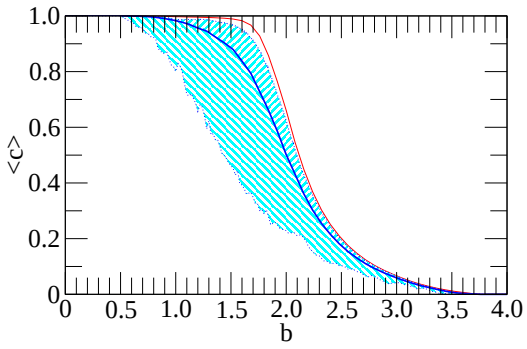


Figure 3.2: Cooperation level $\langle c \rangle$ versus the temptation to defect b averaged over 100 SF networks (solid lines) and envelopes (dotted lines). Red lines represent semicycles with increasing b and blue lines represent semicycles with decreasing b. Only irreversible realizations are shown. The network size is $N = 1.2 \times 10^5$. See the text for further details.

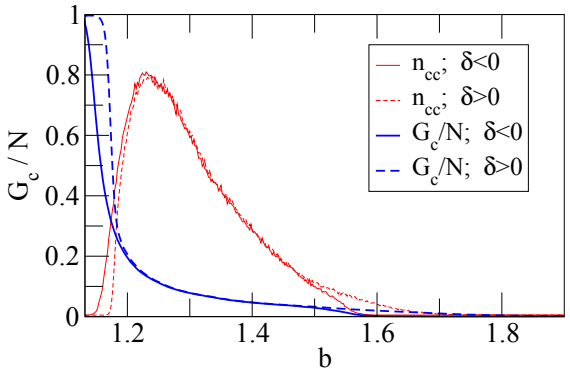


Figure 3.3: Number n_{cc} of cooperator clusters (blue, thick lines) and relative size of main cooperator cluster G_c/N (red, thin lines) in ER networks. Solid lines represent b-decreasing half-cycles and dashed lines represent b-increasing half-cycles. The system size is $N = 1.2 \times 10^5$. We have averaged over 50 simulations.

3.4 Microscopic roots.

Previous studies [109, 112, 113] have shown that, in the asymptotic states of the evolutionary dynamics of the PD game, under the updating of the actions explained above, the network is generically partitioned into three sets of nodes: Pure cooperators (nodes where cooperation has reached fixation), pure defectors, and fluctuating strategists (nodes where fixation is impossible so that defection and cooperation alternate forever). Pure cooperators resist invasion by grouping together in cooperator clusters, each of these connected subgraphs keeping around it a cloud of fluctuating strategists. The basis for an understanding of the irreversible behavior in ER networks is found by looking along both (b-increasing and b-decreasing) branches at the details of this microscopic organization of cooperation. In particular, in what follows we pay attention to the number and size of pure cooperator clusters as a function of b. Figure 3.3 shows the averaged relative size $\langle G_c/N\rangle$ of the largest cooperator cluster, and the average $\langle n_{cc}\rangle$ of the number of cooperator clusters versus the temptation to defect b, in both semicycles for ER networks.

Let us first analyze the b-increasing semicycle. In typical configurations near the absorbent state $c = 1$, the pure cooperators percolate the network conforming a giant cooperator cluster whose averaged relative size $\langle G_c/N\rangle \simeq 1$. As the temptation to defect b increases, starting from such configurations, the existence of a single very large cluster of pure cooperators allows initially for a very efficient resilience to invasion by defectors until a value of $b \simeq 1.16$ is reached. From there on, invasion processes are dramatically enhanced so inducing the fragmentation of the large cluster: $\langle G_c/N\rangle$ decreases quickly, the large cluster giving birth to an increasing number n_{cc} of small clusters of pure cooperators, that at $b \simeq 1.23$ reaches its maximum value $n_{cc} \simeq 160$ when the large cluster size has been reduced to $\langle G_c/N\rangle \simeq 0.15$. Further increase of b reduces both the number of pure cooperator clusters and the size of the largest one: At $b \simeq 1.8$ basically only the largest cluster remains with a very small size which keeps decreasing further beyond the tipping point (typically found at $b \geq 2$).

Now we analyze the b-decreasing semicycle. Back from the typical configuration reached past the tipping point near the absorbing state $c = 0$, when decreasing the temptation value b the very small size of the remaining pure cooperator cluster cannot benefit (*i.e.*, enlarge its size) enough from cooperative fluctuations nearby; correspondingly the level of cooperation $\langle c\rangle$ remains well below the values observed for the b-increasing branch. It is not until a value of $b \simeq 1.6$ is reached, that $\langle G_c/N\rangle$ starts a significant increase. Simultaneously, some cooperative fluctuations in the cloud of fluctuating agents form separated small cooperator clusters, so that n_{cc} also starts to significantly detach from zero. At around $b \simeq 1.5$ both $\langle G_c/N\rangle$ and n_{cc} (as well as the average level of cooperation $\langle c\rangle$) show already values that are very

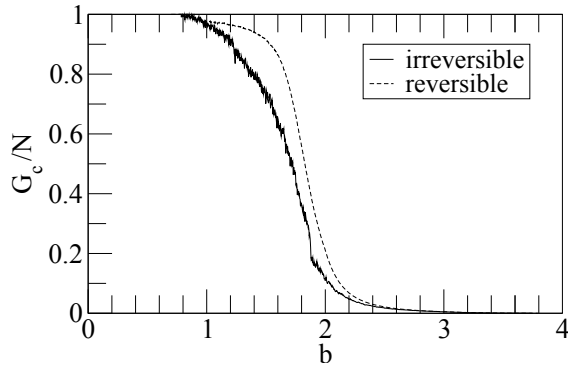


Figure 3.4: Relative size of the main cooperator cluster G_c/N for reversible processes (dashed line) and irreversible ones (solid line) in the b-decreasing semicycle ($\Delta b < 0$) on SF networks. Averaged over the 100 different networks studied that show irreversible behavior. The system size is $N = 1.2 \times 10^5$.

close to those exhibited by the b-increasing branch. However, once reached the value $b \simeq 1.23$, where n_{cc} has its maximum value (and, as explained in previous paragraph, the fragmentation of the large cluster of pure cooperators reached an end in the b-increasing branch), though further decrease in b leads to an increase of $\langle G_c/N \rangle$, and a concomitant decrease of n_{cc} due to the connection of small cooperator clusters to the largest one, these processes take place at a slower pace than the corresponding fragmentation occurring for the b-increasing branch. The consequence is that the cooperation level values in this range of b values for the b-decreasing branch are significantly lower than those for the b-increasing semicycle. Note that though the values of $\langle G_c/N \rangle$, n_{cc}, and $\langle c \rangle$ in the range of intermediate $1.23 \leq b \leq 1.5$ values are very similar in both branches, the system keeps memory of the path followed, demonstrating the importance of the particular topological details of the organization of cooperator clusters.

A significant difference, regarding the microscopic organization of cooperation, between ER and SF networks, is the observation first reported in [109] that for SF networks pure cooperators group together in a single cluster, while in ER networks they are disaggregated into several cooperator clusters for generic values of b. In our simulations here we are using network sizes that are larger than those used in [109] by a factor of 30, and for SF networks we have observed nodes that, though being isolated from the main cooperator cluster, remain cooperators during observational time scales. Strictly speaking they are not pure cooperators, for the probability of invasion by the defective strategy is not strictly zero (in all the cases analyzed), though it turns out to be exceedingly small, due to the large connectivity (degree) of these nodes. These quasi-pure cooperators appear in both, reversible and irreversible network re-

alizations. For a network size of $N = 1.2 \times 10^5$ its number is never larger than 8 for reversible realizations and b-increasing branches of irreversible ones, and not larger than 14 for b-decreasing branches of the latter. Their contribution both direct and indirect (through the cloud of fluctuating strategists each one keeps nearby) to the level c of cooperation can be considered as negligible. Still one cannot discard *a priori* an eventual role they might play in the reshaping of the main cooperator cluster during the hysteresis cycle of particular irreversible realizations.

In figure 3.4 we plot the relative size of the cooperator cluster $\langle G_c/N \rangle$ averaged over 100 irreversible realizations for both forward and backward branches of the cycle. Contrary to what happens for ER networks at high values of the temptation to defect, when starting to decrease it from b_{min}, the size of the cooperator cluster in SF networks initially follows very closely the values of the forward branch until $b \simeq 2.5$. However, significant differences in the average cooperation value $\langle c \rangle$ (see figure 3.2) are already noticeable from $b \simeq 3$, indicating that the contribution from the cloud of fluctuating strategies is lower for the backward branch. When further decreasing b down from $b \simeq 2.5$, the averaged size of the cooperator cluster takes on values progressively lower than in the b-increasing branch. This agrees nicely with the observation just made in the previous sentence on the cloud of fluctuating strategies, for the growth of the cooperator cluster originates from the cooperative fluctuations in its frontier, and thus the strength of these fluctuations determines the pace of the cluster size growth. The difference between forward and backward branches persists down to the tipping point, which somewhat surprisingly occurs for values of b outside the PD game range.

Chapter 4

Evolutionary dynamics on interdependent populations.

Although several mechanisms can promote cooperative behavior, there is no general consensus about why cooperation survives when the most profitable action for an individual is to defect, specially when the population is well mixed. Here we show that when a replicator like evolutionary game dynamics takes place on interdependent networks, cooperative behavior is fixed on the system. Remarkably, we analytically and numerically show that this is even the case for well mixed populations. Our results open the path to new mechanisms able to sustain cooperation and can provide hints for controlling its raise and fall in a variety of biological and social systems.

4.1 Introduction

The onset of global cooperation in large populations of unrelated agents when defective actions provide the largest short-term benefits at the individual level constitutes one of the most amazing puzzles for evolutionary dynamics [14, 171, 172, 124]. During the last decade, the structure of the interactions among individuals seems to have provided a way out for cooperation to survive in those scenarios, such as the Prisoner's Dilemma (PD) game, in which defective behaviors are evolutionary favored under the well-mixed assumption [16, 88]. Although recent results have shown that network reciprocity is not always a viable mechanism to explain cooperation among humans (see, *e.g.*, chapter 6), larger cooperative levels are achieved if an evolutionary game dynamics takes place on top of structured populations and networks, in which nodes account for players and links represent the existence of game interactions. Moreover, further including real structural patterns of large systems [28] (scale-

free distribution for the number of contacts a player has [108, 109], the small-world properties [81], nonzero density of triads [173], etc) provides also high cooperative outputs.

On the other hand, in most cases, a real population – be it a biological or a social system – is not isolated and interactions take place at and between different levels (or layers) following different rules [174, 175]. Think of for instance in an economical system, where different levels account for different competitive markets and their interdependencies (developers, manufactures, providers). The rules governing the interactions at one layer are not necessarily the same that those driving the dynamics at another layer – admittedly, within each layer competition should exist while this is not necessarily the case for inter-layer interactions. Thus, a natural question arises as to whether the observed degree of interdependency in real systems is a relevant factor for the emergence and survival of cooperative behavior.

The previous interdependency, which is also referred to as multiplexity, can be easily incorporated into the framework of any dynamical process by coupling two or more networked populations in which links between individuals of the same population involve a different dynamical relationship to those stablished between members of different populations [176, 177, 178]. In this chapter, we focus on the case in which an evolutionary PD game drives the interactions between agents of the same population. On its turn, the existence of links between agents of different populations allow the two networks to interact. We will assume that the latter interactions are ruled by the Snowdrift (SD) game. In this way, defection is punished when facing other defectors outside the original population, thus balancing the evolutionary advantage that defectors find by exploiting cooperators in their respective populations.

We henceforth analyze what new emergent behavior results from the multilevel nature of a system made up by two populations that interact through a number of links connecting nodes located at each subsystem. Exact analytical calculations can be carried out for the case in which the population of each layer is well mixed, through the nonlinear analysis of the two-coupled-variable replicator equation for the strategic densities in both layers. Our results show the emergence of a new *polarized* state in which all the individuals in one of the populations cooperate while all in the other population defect. In addition we find *quasi-polarized* states, so that all the agents in one population are defectors, while most of the other one cooperate. Moreover, we also numerically show that the previous results hold for the case of networked populations. As we will discuss later on, our findings provide new mechanisms for the rise and survival of cooperation and for its control.

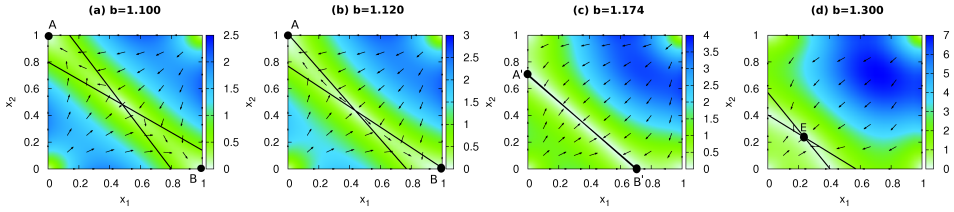


Figure 4.1: Phase portrait of replicator equation (4.1) for the symmetric case ($\beta = 1$) and weak PD ($r = 0$) intra-population game, for different values of the temptation b, with $p = 0.3$, and $\epsilon = -0.4$. The direction of velocity field is indicated by the arrows, and its modulus by the colors. We also plot the interior nullclines. For low values of b **(a)**, the polarized states A and B are attractors. They lose stability at $b = b^{up}$ **(b)**, in favor of the quasi-polarized states A' and B'. These in turn destabilize at $b = b^{c}$ **(c)** when the nullclines coincide in a line of marginally stable equilibria. From there on, the interior equilibrium E becomes the global attractor **(d)**.

4.2 The model: Evolutionary dynamics on two interacting populations.

Let us first describe the evolutionary dynamics of two interacting populations of size N_1 and N_2. Two agents belonging to the same population α ($= 1, 2$) play a PD game so that a cooperator facing a cooperator (defector) in population α obtains a payoff $R = 1$ ($S = 0$). On the other hand, a defector facing a cooperator (defector) obtains a benefit of $T = b > 1$ ($P = r \geq 0$). The games played between agents of different populations follow the same parameterization except for the situation in which two defectors meet. In this case, the associated punishment is negative, $P = \epsilon < 0 -$ thus, inter-populations games follow the SD formulation. Importantly, the strategists' competition for replication only occurs among own-population players. That is to say that there is no "interbreeding" (as it happens for different species, in biological contexts) or "strategic diffusion" (as for functionally heterogeneous layers in social or economical contexts) among the individuals of different populations. In terms of imperfect (or/and irrelevant) knowledge, the strategists from a population are unaware of the replicating success of strategies in the other population (or/and this information is irrelevant for its replication).

4.3 Well-mixed populations: Analytical formulation.

To start with, consider the case in which agents of the same population (layer) are well-mixed. Let us also assume that both N_1 and N_2 are large enough, *i.e.*, $N_1, N_2 \gg 1$. Under these simple assumptions, an exact analytical description via the analysis

of the phase portrait of the two-dimensional replicator equation for two-by-two matrix games is possible. In our well-mixed population approximation an individual in population α has $N_\alpha - 1$ neighbors inside this population. Moreover, for interactions between the two layers, we suppose that any pair of nodes (each one of a different population) is present with probability p. Thus, the number of inter-population links is equal to $p \cdot N_1 \cdot N_2$.

Let us call x_α the fraction of cooperators in the population α. The replicator equations for the evolutionary game dynamics are

$$\begin{aligned} \dot{x}_1 &= x_1(1-x_1)[(N_1-1)(x_1(1-b+r)-r)+ \\ &\quad N_2 p(x_2(1-b+\epsilon)-\epsilon)] \\ \dot{x}_2 &= x_2(1-x_2)[(N_2-1)(x_2(1-b+r)-r)+ \\ &\quad N_1 p(x_1(1-b+\epsilon)-\epsilon)] \ . \end{aligned} \tag{4.1}$$

The results of the theoretical analysis (see section 4.4 for details) of these coupled deterministic equations are illustrated in Fig. 4.1 for the symmetric (thus non-generic) case $N_1 = N_2$, and the simple weak ($r = 0$) PD game for those intra-population encounters. Below we will comment on the main qualitative changes for the generic case, *i.e.*, whenever both the size proportion $\beta = N_1/N_2 \neq 1$ and general PD ($r > 0$) game for intra-population interactions apply.

The analysis of Fig. 4.1 shows a rather natural non-linear resolution of the conflict introduced by fitness-punishment (ϵ) to inter-populations defective encounters. Briefly said, even-symmetric ($x_1 = x_2$) states D (both populations are fully defective) and C (fully cooperative populations) are both, for any $b > 1^+$, unstable against perturbations in all directions, and stability resides instead on odd-symmetric polarized states [A (all-D in population 1 and all-C in population 2) and its symmetric transformed B (all-C in population 1 and all-D in population 2)] for strictly positive temptation b less than a bound $b^{up}(\epsilon; p) = 1 - p\epsilon$ (see Fig. 4.1.a). At this critical (bifurcation) value of b the interior nullclines $\dot{x}_1 = 0$ and $\dot{x}_2 = 0$ (see Fig. 4.1.b) touch states A and B respectively. Increasing the value of the temptation b above b^{up} the polarized states lose their stability in favor of the quasi-polarized states [A' (all-D in 1 and mostly C in 2) and its symmetric B'], which detach from A and B and become attractors. At $b = b^c = 1 - \frac{p\epsilon}{1-p}$ the interior nullclines coincide (see Fig. 4.1.c) becoming a line (A'B') of marginally stable equilibria. Finally, for $b > b^c$ (see Fig. 4.1.d) the global attractor is the interior even-symmetric state E, the intersection of the interior nullclines, which keeps approaching, as b increases, the neighborhood of the high b limit attractor, say the state D of fully defective populations.

This scenario remains qualitatively unchanged for strictly positive values of the parameter r, provided $0 < r < -p\epsilon$, the only change being that the bifurcation value b^c, where the quasi-polarized states loose stability, becomes:

$$b^c = 1 + \frac{r - p\epsilon}{1 - p} \, . \tag{4.2}$$

In other words, the weak PD limit ($r = 0$) for the intra-population game is structurally stable respect to (small enough) positive parametric variations of the game parameter r. For $r > -p\epsilon$, the scenario changes drastically: D is now a stable equilibrium, but still, for $b < b^{up}$ (which doesn't depend on r), the polarized states are also stable equilibria. Only for larger $b > b^{up}$ values of the temptation, D becomes the unique global attractor. Summarizing the results for the symmetric case, the attractor states for increasing values of b from $b = 1^+$ follow the sequence:

$$\text{A, B} \overset{b^{up}}{\rightarrow} \text{A', B'} \overset{b^c}{\rightarrow} \text{E} \, . \tag{4.3}$$

when $0 \leq r < -p\epsilon$ while, when $r > -p\epsilon$, the sequence is:

$$\text{D, A, B} \overset{b^{up}}{\rightarrow} \text{D} \, . \tag{4.4}$$

For the general case $N_1 \neq N_2$, the lack of the population interchange symmetry modifies some of the features seen in the symmetric case. Without loss of generality,

(a)	D, A, B $\overset{b^{up}_B}{\rightarrow}$ D, A $\overset{b^{up}_A}{\rightarrow}$ D
(b)	A, B $\overset{b^{up}_B}{\rightarrow}$ A $\overset{b^{up}_A}{\rightarrow}$ A' $\overset{b^{c}_A}{\rightarrow}$ E
(c)	A, B $\overset{b^{up}_B}{\rightarrow}$ A $\overset{b^{up}_A}{\rightarrow}$ A'
(d)	A, B $\overset{b^{up}_B}{\rightarrow}$ A, B' $\overset{b^{c}_B}{\rightarrow}$ A $\overset{b^{up}_A}{\rightarrow}$ A'
(e)	A, B $\overset{b^{up}_B}{\rightarrow}$ A, B' $\overset{b^{c}_B}{\rightarrow}$ A $\overset{b^{up}_A}{\rightarrow}$ A' $\overset{b^{c}_A}{\rightarrow}$ E
(f)	A, B $\overset{b^{up}_B}{\rightarrow}$ A, B' $\overset{b^{up}_A}{\rightarrow}$ A', B' $\overset{b^{c}_B}{\rightarrow}$ A'
(g)	A, B $\overset{b^{up}_B}{\rightarrow}$ A, B' $\overset{b^{up}_A}{\rightarrow}$ A', B' $\overset{b^{c}_B}{\rightarrow}$ A' $\overset{b^{c}_A}{\rightarrow}$ E

Table 4.1: Sequence of attractors in phase space for equation (4.1), as b increases from $b = 1^+$. The arrow indicates a bifurcation at the b value that appears over the arrow. The scenarios (a)-(g) correspond to different ranges of values of the parameters r, β, p and ϵ, that are made explicit in section 4.4. Note that except for the scenario (a), that corresponds to $r > -\beta p\epsilon$, polarized and quasi-polarized states dominate the asymptotic behavior.

we assume that $\beta = N_1/N_2 > 1$. On one hand, the lower bound of r for the stability of the fully defective state D becomes now $r = -\beta p\epsilon$. On the other hand, the bifurcation values at which the polarized states lose their stability are now different,

$$b_B^{up} = 1 - \frac{p\epsilon}{\beta} < b_A^{up} = 1 - \beta p\epsilon \ , \tag{4.5}$$

as well as the bifurcation values (provided they exist) at which quasi-polarized states destabilize, $b_B^c < b_A^c$, where

$$b_B^c = 1 + \frac{r^2 - (p\epsilon)^2}{(r + \beta p\epsilon) - p(\beta r + p\epsilon)} \tag{4.6}$$

$$b_A^c = 1 + \frac{\beta(r^2 - (p\epsilon)^2)}{(\beta r + p\epsilon) - p(r + \beta p\epsilon)} \ . \tag{4.7}$$

Let us note that the polarized state A, where the defective population is of larger size, turns out to have a wider range of stability, as well as a larger basin of attraction, than the state B. The results of the complete analysis of the replicator equation (4.1) are summarized in Table 4.1, where we show the sequences of attractors coexisting in phase space. The seven scenarios (a)-(g) correspond to different ranges of values of the parameters r, β, p and ϵ (see section 4.4 for further details).

From the previous analysis of well mixed populations, one sees that polarized and quasi-polarized states appear as generic attractors of the evolutionary dynamics for wide ranges of model parameters, which in turn has the effect of enhancing in a remarkable way the asymptotic levels of cooperation in the two-populations system.

4.4 Phase portrait analysis of the two-variable replicator equation

The replicator equation that describes the continuum time evolution of the cooperator fractions $x_1(t)$, $x_2(t)$ in subpopulations 1 and 2 can be written as:

$$\dot{x}_1 = \mathcal{F}_1(x_1, x_2) \ , \tag{4.8}$$

$$\dot{x}_2 = \mathcal{F}_2(x_1, x_2) \ , \tag{4.9}$$

where the velocities $\mathcal{F}_{1,2}$, after time rescaling, are explicitly given as:

$$\mathcal{F}_1(x_1, x_2) = x_1(1 - x_1)[\beta(x_1(1 - b + r) - r) + p(x_2(1 - b + \epsilon) - \epsilon)] \ , \tag{4.10}$$

$$\mathcal{F}_2(x_1,x_2) = x_2(1-x_2)[(x_2(1-b+r)-r)+\beta p(x_1(1-b+\epsilon)-\epsilon)] \ . \quad (4.11)$$

The unit square $0 \leq x_1, x_2 \leq 1$ is the invariant set of interest here. To follow the phase portrait variation of a two-degrees of freedom nonlinear system like equation (4.9) is pretty straightforward for one-parameter variations. We are dealing with a model where $b,\ r,\ \epsilon,\ \beta$, and p are free model parameters, each one inside their natural range, *i.e.*, $b > 1^+$, $0 \leq r \leq 1$, $\epsilon < 0^-$, $\beta \geq 1$, and $0 \leq p \leq 1$. In our systematics below, we will consider continuum variation of b, from $b = 1^+$ up to infinity, at fixed values of the other parameters and so we will obtain the "critical" (bifurcation) points $b^*(\epsilon, r; \beta, p)$, where the phase portrait of the evolution experiences *qualitative changes*: Somewhat, the direction of increasing temptation b is often most considered in recent literature on PD games. But we will pay due attention also to variations of the parameter r, and find two important critical values that do not depend on the value of the temptation b, so that different scenarios of phase transitions (inside the well-mixed population approximation to the thermodynamical limit $N_1, N_2 \to \infty$) as b varies do appear. Finally, we choose also β as an interesting (*e.g.*, for control applications) parameter to vary, and find also two critical values that are temptation independent, that in turns, increase the number of those scenarios.

The best visualization of the velocity field is a phase portrait where fixed (equilibrium) points and nullclines are also plotted, as in Fig. 1 in the main text. A nullcline is the locus of points defined by $\mathcal{F}_i(x_1,x_2) = 0$ for some i. The nullclines that correspond to $\mathcal{F}_1(x_1,x_2) = 0$ are the straight lines

$$x_1 = 0 \ , \quad (4.12)$$

$$x_1 = 1 \ , \quad (4.13)$$

$$x_2 = \frac{-x_1\beta(b-1-r)-(\beta r+p\epsilon)}{p(b-1-\epsilon)} \ , \quad (4.14)$$

while those that correspond to $\mathcal{F}_2(x_1,x_2) = 0$ are

$$x_2 = 0 \ , \quad (4.15)$$

$$x_2 = 1 \ , \quad (4.16)$$

$$x_2 = \frac{-x_1\beta p(b-1-\epsilon)-(r+\beta p\epsilon)}{(b-1-r)} \ . \quad (4.17)$$

The possible equilibria are the crossing points of any line from the first group with any other line from the second one, so there are nine candidates. Moreover, only

solutions in the unit square, $0 \leq x_1, x_2 \leq 1$, interest us, and this exclude two of the crossing points (see below), leaving the following seven possibilities, namely the four corners of the unit square:

- $A = (0, 1)$,
- $B = (1, 0)$,
- $C = (1, 1)$,
- $D = (0, 0)$,

and those whose location depends on parameter values:

- We call A' the crossing point of nullclines (4.12) and (4.17), whose coordinates are $x_1(A') = 0$ and

$$x_2(A') = \frac{-(r + \beta p\epsilon)}{(b - 1 - r)} \ . \tag{4.18}$$

- We call B' the crossing point of nullclines (4.14) and (4.15), so that $x_2(B') = 0$ and

$$x_1(B') = \frac{-(\beta r + p\epsilon)}{\beta(b - 1 - r)} \ . \tag{4.19}$$

- Finally, we call E the crossing of (4.14) and (4.17). Its coordinates are obtained as:

$$x_1(E) = \frac{(b - 1 - r)(\beta r + p\epsilon) - p(b - 1 - \epsilon)(\beta r + p\epsilon)}{\beta[(p(b - 1 - \epsilon)^2 - (b - 1 - r)^2]} \ , \tag{4.20}$$

$$x_2(E) = \frac{(b - 1 - r)(r + \beta p\epsilon) - p(b - 1 - \epsilon)(r + \beta p\epsilon)}{(p(b - 1 - \epsilon)^2 - (b - 1 - r)^2} \ . \tag{4.21}$$

The (missing in the list) crossings of (4.13) - (4.17), and of (4.14) - (4.16), are easily seen to be always outside the unit square for the range of parameters considered. Also inside this range, the non-generic event of *nullclines' coincidence* only could happen provided $\beta = 1$ and $r > -p\epsilon$, at a value $b^c(\epsilon, r; \beta = 1, p) = 1 + \frac{r - p\epsilon}{1 - p}$. Only then, the exotic (forced by symmetry) situation in which there is a segment of marginally stable equilibria occurs.

To determine the bifurcation points, one uses the spectral analysis of tangent space perturbations around equilibria. The linearized evolution of small perturbations around the fixed point $\mathbf{x}^*$ is given by the matrix:

$$\left(\begin{array}{cc} \frac{\partial \mathcal{F}_1}{\partial x_1} & \frac{\partial \mathcal{F}_1}{\partial x_2} \\ \\ \frac{\partial \mathcal{F}_2}{\partial x_1} & \frac{\partial \mathcal{F}_2}{\partial x_2} \end{array} \right)_{\mathbf{x}=\mathbf{x}^*} \tag{4.22}$$

In what follows, the presentation of the results from the phase portrait analysis of the nonlinear coupled ODE (4.11) tries to rationalize them in terms of evolutionary game theoretic concepts, within a thermodynamical limit (statistical physics) perspective.

4.4.1 Symmetric case: $N_1 = N_2 (= N)$.

For simplicity, as well as for illustrate neatly the systematics that we follow, we analyze first the case of equal population sizes. For this case, where populations are identical (though distinguishable), the population interchange symmetry imposes that phase portrait is invariant under permutation of coordinates ($x_1 \leftrightarrow x_2$), a non-generic property that limits severely the possible scenarios. The stability analysis of the equilibria shows that there are two generic scenarios for the sequence of bifurcations that appear when b increases from 1^+ up to infinity:

(s_1) If $r > r_c = -p\epsilon$ there is only one bifurcation at $b^{up}(r, \epsilon, \beta = 1, p) = 1 - p\epsilon$. For $b < b^{up}$, the phase portrait has three stable equilibria with their own basins of attraction: D, A, and B. The equilibria C, A' and B' are unstable, and E is outside the unit square. At $b = b^{up}$, A and B destabilize (through collision with A' and B' that exit the unit square) becoming saddle equilibria, and D becomes the unique global attractor for $b > b^{up}$. This translates into the following sequence of attractors when temptation increases from 1^+:

$$\text{D, A, B} \overset{b^{up}}{\rightarrow} \text{D} \, . \tag{4.23}$$

(s_2) If $r < r_c(p, \epsilon)$, however, D is always unstable, and there are two bifurcations at b^{up} and b^c (and note that $b^{up} < b^c$). For $b < b^{up}$ the equilibria C, D are sources, E is a saddle, and A and B are attractors, becoming saddle equilibria at b^{up} where A' and B' enter into the unit square. For $b^{up} < b < b^c$ A' and B' are the only attractors. At b^c the segment A'B' of marginally stable equilibria is the limit set for all trajectories (nullcline's coincidence). For $b > b^c$ E becomes the unique (and even-symmetric) global attractor. This last bifurcation restores

the symmetry of the asymptotic evolution that was spontaneously broken at lower b values. The sequence of stationary limiting (point) densities is:

$$\text{A, B} \overset{b^{up}}{\rightarrow} \text{A', B'} \overset{b^{c}}{\rightarrow} \text{E} \ . \tag{4.24}$$

Note that the condition $r = r_c(p, \epsilon)$ that separates the regimes where the equilibrium D is unstable ($r < -p\epsilon$) or attractor ($r > -p\epsilon$), corresponds to the exact compensation of the surplus rN of defective intra-population interactions of a defector and the punishment $p\epsilon N$ it receives from inter-population interactions. Below this critical value, full defection is unstable to cooperative fluctuations. But, as we have just seen, even in case the punishment from coupling is weaker than surplus, polarized states have their own basins of attraction, away from whole defection, at low values of $b > 1^+$. This can be rationalized from the role that punishment plays in our -no interbreeding, punishing defective coupling- setting. Populations' strategic polarization emerges as stable generic asymptotic state of evolution, even when defectors can afford external punishment (D being then fully stable): The duplex (two coupled populations) has always the option to become polarized or quasi-polarized provided the initial conditions belong to its basin of attraction.

4.4.2 General case: $N_1 \neq N_2$.

The parameter p determines the fraction of inter-to-intra-population interactions any agent plays per unit time in the symmetric ($N_1 = N_2$) case. This fraction changes to βp and p/β ($\beta > 1$) for small and large populations respectively, when symmetry of population interchange is absent. This combination of parameters regulates how important to the replicating power (fitness) of an individual the inter-population coupling is, and we then see that for the largest population the effective coupling p/β is smaller. This makes the polarized state A (where population 1 is defective) more robust than the polarized state B, and provided both are attractors, the basin of attraction of A is correspondingly larger. This is a major qualitative change in the phase portrait of the velocity field of evolution in the absence of symmetry. The concomitant change is the shift, and in more extreme cases the disappearance, of the bifurcations associated to the quasi-polarized equilibria A' and B', *i.e.*, $b^{up}_{A,B}$ and $b^{c}_{A,B}$:

$$b^{up}_{B}(r, \epsilon; \beta, p) = 1 - (p/\beta)\epsilon \ , \tag{4.25}$$

$$b^{up}_{A}(r, \epsilon; \beta, p) = 1 - \beta p\epsilon \ , \tag{4.26}$$

$$b^{c}_{B}(r, \epsilon; \beta, p) = 1 + \frac{r^2 - (p\epsilon)^2}{(r + \beta p\epsilon) - p(\beta r + p\epsilon)} \ , \tag{4.27}$$

$$b_A^c(r,\epsilon;\beta,p) = 1 + \frac{\beta(r^2-(p\epsilon)^2)}{(\beta r + p\epsilon) - p(r+\beta p\epsilon)} \ . \tag{4.28}$$

Note that the minimum of this set of values is b_B^{up}, its maximum is b_A^c, and that the relative order of the other two values is parameter dependent. Several new generic scenarios of phase portrait variations naturally follows from these major effects, when the "population interchange" symmetry is absent. Still, let us remark that the evolutionary attractiveness of the odd-symmetric polarized (A and B) and quasi-polarized (A' and B') asymptotic densities still dominates ample regions of parameter space.

A first scenario, similar to the first one seen above for the symmetric case, is found when $r > r_c^A(\epsilon;\,\beta,p) = -\beta p\epsilon$. In this scenario, the fully defective state D is stable for all $b > 1$ values. For very low values of b, A and B are also stable. Due to asymmetry, the instabilities of A and B occur at different bifurcation values, $b_B^{up} < b_A^{up}$, so that state B destabilizes first when b increases from $b = 1^+$, as expected, *i.e.*,

(*i*) If $r_c^A < r$ there are only two bifurcations at $b_B^{up} < b_A^{up}$. For all $b > 1^+$, C is unstable and E is outside the unit square. For $b < b_B^{up}$, the states D, A and B are attractors. At b_B^{up}, B collides with the unstable B' that exits the unit square, then becoming a saddle with unstable direction corresponding to defective fluctuations in cooperative population 1. The same happens *mutatis mutandi* ($1 \leftrightarrow 2$ interchange) to A at b_A^{up}, leaving finally D (for $b > b_A^{up}$) as the global attractor.

$$\text{D, A, B} \overset{b_B^{up}}{\rightarrow} \text{D, A} \overset{b_A^{up}}{\rightarrow} \text{D} \tag{4.29}$$

At $r = r_c^A$, for a defective individual in population 2, and state D, the "internal surplus - coupling punishment" balance exactly compensates. This means that changing to cooperator makes no difference to its replicating power, and thus a zero eigenvalue appears in the spectrum of the Jacobian (linear stability) matrix of the fully defective state D. Inside the range $r < r_c^A$, D is always unstable face to cooperative fluctuations in the smaller population. Further down in surplus (r) values, at $r = r_c^B = -(p/\beta)\epsilon$, D becomes also unstable face to cooperative fluctuations in the large population. In other words, when decreasing r from large (compared to r_c^A) positive values of intra-population surplus, to 0^+ (*weak PD* limit), there are two critical values, where qualitative changes of the phase portrait occur, that coincide with the change of stability of D from stable ($r > r_c^A$) to saddle ($r_c^B < r < r_c^A$), to source ($r < r_c^B$).

Provided $r < r_c^A$, if one consider the high b ($\rightarrow \infty$) limit, one easily finds that it can be either "mixed type" (state E, interior to the unit square) or "quasi-polarized"

(state A', on the vertical $x_1 = 0$) regarding its convergence to virtually full defection. The transition between these two qualitatively different "high temptation limit behaviors", for given values of ϵ, p, and r, is controlled by the value of the population ratio β and it occurs at the critical value:

$$\beta_c^A(\epsilon, r;\, p) = \frac{p(r-\epsilon)}{r - p^2\epsilon}\ . \tag{4.30}$$

At this value of the population ratio, the bifurcation value b_A^c (where A' collides with state E, this one entering into the unit square) formally diverges, so that the collision occurs (or doesn't), depending on the value of the population ratio β, for fixed value of p, r, and ϵ.

On the other side, the bifurcation value at b_B^c only occurs provided $r < r_c^B$, but its relative order with respect to b_A^{up} depends also on the value of β with a critical value at:

$$\beta_c^B(\epsilon, r;\, p) = \frac{-p\epsilon(p^2\epsilon - r) - \sqrt{p^2\epsilon^2(p^2\epsilon - r)^2 - 4p^2\epsilon(r-\epsilon)(p^2\epsilon^2 - r^2)}}{2p^2\epsilon(r-\epsilon)}\ . \tag{4.31}$$

The different possible combinations of all the previous possibilities give the following scenarios:

(*ii*) If $r_c^B < r < r_c^A$, then the stable linear manifold of the saddle point D ($x_2 = 0$) does not allow B' to be a stable equilibrium, while its unstable direction ($x_1 = 0$) pushes evolution to polarized A or quasi-polarized A' states; C is a always a source for all $b > 1$. Two different scenarios are realized depending on the inter-population ratio value, β:

(ii_1) If $\beta > \beta_c^A$ (see Eq. 4.30), bifurcations only occur at $b_B^{up} < b_A^{up}$. At b_B^{up}, the collision of B and the unstable exiting B' occurs, while at b_A^{up}, it takes place the collision of A with the entering state A'. The corresponding sequence of attracting equilibria is given by:

$$\text{A, B} \overset{b_B^{up}}{\rightarrow} \text{A} \overset{b_A^{up}}{\rightarrow} \text{A'} \tag{4.32}$$

(ii_2) If $\beta_c^A > \beta$, besides the bifurcations described in (ii_1), there is an additional bifurcation at b_A^c, where A' collides with state E that enters into the unit square. The corresponding sequence of attracting equilibria is given by:

$$\text{A, B} \overset{b_B^{up}}{\rightarrow} \text{A} \overset{b_A^{up}}{\rightarrow} \text{A'} \overset{b_A^c}{\rightarrow} \text{E} \tag{4.33}$$

The presence or absence of the bifurcation b_A^c determines whether the approach to the high temptation limit is via "mixed interior type" E state, or "edge quasi-polarized type" A' state, so that for values of β below critical (β_c^A), virtually full defection $(1^-, 1^-)$ is approached with non-zero cooperation levels in both populations as b diverges.

(*iii*) If $r < r_c^B$, both quasi-polarized states A' and B' enter into the unit square at b_B^{up} and b_A^{up}, respectively. B' always destabilizes at b_B^c ($> b_B^{up}$ always) to become a saddle through collision with the exiting unstable interior equilibrium E. This may happens before [as in (iii_1) and (iii_2) below] or after [as in (iii_3) and (iii_4)] the entrance of A' at b_A^{up} depending on β value (relative to β_c^B). And finally note that the bifurcation at b_A^c only occurs for $\beta < \beta_c^A$, as analyzed above, to arrive to the following possible four scenarios:

(iii_1) If $\max(\beta_c^A, \beta_c^B) < \beta$, then $b_B^c < b_A^{up}$, and b_A^c is absent:

$$\text{A, B} \overset{b_B^{up}}{\rightarrow} \text{A, B'} \overset{b_B^c}{\rightarrow} \text{A} \overset{b_A^{up}}{\rightarrow} \text{A'} \tag{4.34}$$

(iii_2) If $\beta_c^B < \beta < \beta_c^A$, then $b_B^c < b_A^{up}$, and b_A^c occurs:

$$\text{A, B} \overset{b_B^{up}}{\rightarrow} \text{A, B'} \overset{b_B^c}{\rightarrow} \text{A} \overset{b_A^{up}}{\rightarrow} \text{A'} \overset{b_A^c}{\rightarrow} \text{E} \tag{4.35}$$

(iii_3) If $\beta_c^A < \beta < \beta_c^B$, then $b_A^{up} < b_B^c$, and b_A^c is absent:

$$\text{A, B} \overset{b_B^{up}}{\rightarrow} \text{A, B'} \overset{b_A^{up}}{\rightarrow} \text{A', B'} \overset{b_B^c}{\rightarrow} \text{A'} \tag{4.36}$$

(iii_4) If $\beta < \min(\beta_c^A, \beta_c^B)$, then $b_A^{up} < b_B^c$, and b_A^c occurs:

$$\text{A, B} \overset{b_B^{up}}{\rightarrow} \text{A, B'} \overset{b_A^{up}}{\rightarrow} \text{A', B'} \overset{b_B^c}{\rightarrow} \text{A'} \overset{b_A^c}{\rightarrow} \text{E} \tag{4.37}$$

This analysis provides the three-dimensional phase diagram (r, β, b) for fixed, though arbitrary, ϵ and p. It exhibits a wealthy of different macroscopic phases separated by critical lines and surfaces. It shows that polarized and quasi-polarized phases dominate wide regions in parameter space. This illustrates the effects of inter-population trade of fitness (even under the simplest possible structure of inter and intra population contacts) on the evolution of PD replicators.

$$\text{A, B} \overset{b^{up}}{\rightarrow} \text{A', B'} \overset{b^c}{\rightarrow} \text{E} \,. \tag{4.38}$$

when $0 \leq r < -p\epsilon$ while, when $r > -p\epsilon$, the sequence is:

$$\text{D, A, B} \overset{b^{up}}{\rightarrow} \text{D} \,. \tag{4.39}$$

4.5 Random networks.

On the other hand, for structured populations, where individuals interact with their neighbors as dictated by a given network of contacts, it is known that under some assumptions cooperation is enhanced, a phenomenon called network reciprocity.

While for well mixed populations, the stability of polarized states extends down to $b = 1^+$, one should expect that at small $b > 1$ values, the enhancement of cooperative fluctuations due to network reciprocity in the defective population 1 destabilizes the polarized states below some critical value b^{low}. Moreover, one should also expect b^{low} to decrease with the parameter p, because higher values of p increase the payoff that a (defector) individual in population 1 obtains from encounters with (cooperator) individuals of population 2, thus decreasing the resilience of cooperative fluctuations ("network reciprocity") in population 1. In other words, for low values of b, the interaction between populations acts against network reciprocity. These expectations are fully confirmed by the results from simulations of the evolutionary dynamics in populations with a random network structure of intra-population contacts, using the discrete version of replicator dynamics.

In Fig. 4.2 we show the average cooperation $\langle c \rangle$ level (over a sample of 200 different realizations) on the two-population system as a function of b for different values of p, and parameters as indicated. The two populations have a random (Erdös-Rényi [28]) network of contacts with average degree $\langle k \rangle = 6$. In the initial conditions, the individuals of both populations were chosen cooperators with probability $1/2$. The plateau at $\langle c \rangle = 1/2$ points out the asymptotic polarized state. Moreover, the states with $\langle c \rangle < 1/2$ correspond to quasi-polarized regimes where all the individuals in one population are defectors, while those with $\langle c \rangle > 1/2$, at values of $b < b^{low}$, results from states where all the individuals in one population are cooperators. This represent a new type of quasi-polarized states that were not found to be attractors of the dynamics for well mixed populations. The comparison with the average cooperation level for non-interacting populations ($p = 0$ in Fig. 4.2) confirms that for low values of b the inter-population interaction acts against network reciprocity.

From a complementary perspective, the networked populations show new attractors, impossible to be such for coupled well-mixed populations, because they are the effect of network reciprocity. On the other hand, for larger values of b, the populations' coupling favors the achievement of substantial levels of cooperation, well beyond the typical values of b for which network reciprocity ceases to be effective, being an effect already present in the well-mixed case. This clarify farther the confluent effects of these two different mechanisms of cooperation enhancement.

Finally, the robustness of polarized and quasi-polarized states suggests to use the coupling to a defective population as an engineered (control) procedure to induce

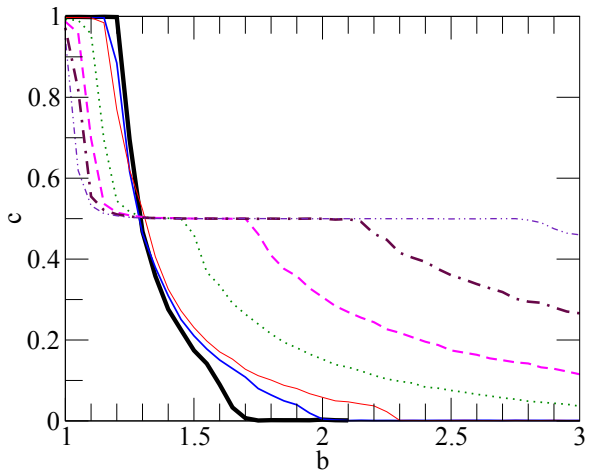


Figure 4.2: Average level of cooperation in the two-population system as a function of b, for different values of the fraction p of inter-population contacts. Other parameters are $r = 0$, $\epsilon = -0.4$, $N_1 = N_2 = 10^3$. The two populations have a random (Erdös-Rényi) network of contacts with average degree $\langle k \rangle = 6$. See the text for further details.

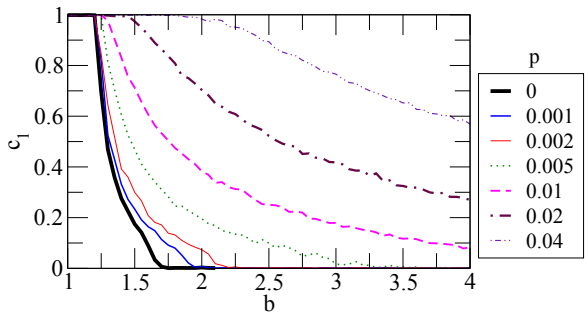


Figure 4.3: Average level of cooperation in the population 1 as a function of b, for different values of the fraction p of inter-population contacts. The population 1 (of size $N_1 = 10^3$) has been coupled to a smaller population 2 ($N_2 = 10^2$). While initial strategies in population 1 are equiprobables (random initial conditions), the population 2 starts from the absorbent state of fully defection. Other parameters are $r = 0$, $\epsilon = -0.4$. Both populations have a random (Erdös-Rényi) network of contacts with average degree $\langle k \rangle = 6$.

high levels of cooperation in a target population. To check for this possibility, we have coupled a large population 1 with random (equiprobable in strategies) initial conditions to a smaller defective population 2. In Fig. 4.3 we show the asymptotic average level of cooperation in a target population of size $N_1 = 10^3$ for different values of the average number, $N_2 \cdot p$, of inter-population contacts per individual of the target population. The results suggest that such arrangements can provide new mechanisms to control and/or sustain cooperation in different kind of systems.

Summarizing, two PD populations SD-coupled in conditions of strict inbreeding (no inter-population strategic diffusion) evolve easily to polarized and quasi-polarized strategic probability densities in the well-mixed thermodynamical limit of the evolutionary replicator dynamics. This happens also when population structure is a complex network of contacts, where other mechanisms (known as network reciprocity) of enhanced cooperation also operate. The confluence of both mechanisms has been analyzed in depth showing that polarization opposes network reciprocity at small values of the temptation parameter, while both act (synergy) together enhancing cooperation in one of the layers for higher temptation values. This phenomenon, that could be rationalized as the effect of incorporating a punishment to defective inter-population encounters, illustrate the remarkable effects that structural multiplexity introduces in evolutionary dynamics.

Chapter 5

Human behavior in Prisoner's Dilemma experiments suppresses network reciprocity.

During the last few years, much research has been devoted to strategic interactions on complex networks. In this context, the Prisoner's Dilemma has become a paradigmatic model, and it has been established that imitative evolutionary dynamics lead to very different outcomes depending on the details of the network. We here report that when one takes into account the real behavior of people observed in the experiments, both at the mean-field level and on utterly different networks the observed level of cooperation is the same. We thus show that when human subjects interact in an heterogeneous mix including cooperators, defectors and moody conditional cooperators, the structure of the population does not promote or inhibit cooperation with respect to a well mixed population.

5.1 Introduction

In recent years, the physics of complex systems has widened its scope by considering interacting many-particle models where the interaction goes beyond the usual concept of force. One such line of research that has proven particularly interesting is evolutionary game theory on graphs [16, 88], in which interaction between agents is given by a game while their own state is described by a strategy subject to an evolutionary process [12, 91]. A game that has attracted a lot of attention in this respect is the Prisoner's Dilemma (PD) [92, 93], a model of a situation in which cooperative actions lead to the best outcome in social terms, but where free riders or non-cooperative in-

dividuals can benefit the most individually. In mathematical terms, this is described by a payoff matrix (entries correspond to the row player's payoffs and C and D are respectively the cooperative and non-cooperative actions)

	C	D
C	1	S
D	T	0

(5.1)

with $T > 1$ (temptation to free-ride) and $S < 0$ (detriment in cooperating when the other does not).

In a pioneering work, Nowak and May [19] showed that the behavior observed in a repeated Prisoner's Dilemma was dramatically different on a lattice than in a mean-field approach: Indeed, on a lattice the cooperative strategy was able to prevail by forming clusters of alike agents who outcompeted defection. Subsequently, the problem was considered in literally hundreds of papers [16, 94, 18, 109, 95], and very many differences between structured and well-mixed (mean-field) populations were identified, although by no means they were always in favor of cooperation [96, 97]. In fact, it has been recently realized that this problem is very sensitive to the details of the system [88, 115], in particular to the type of evolutionary dynamics [79] considered. For this reason experimental input is needed in order to reach a sound conclusion about what has been referred to as 'network reciprocity'.

Here, we show that using the outcome from the experimental evidence to inform theoretical models, the behavior of agents playing a PD is the same at the mean field level and in very different networks. To this end, instead of considering some *ad hoc* imitative dynamics [19, 98, 21], our players will play according to the strategy recently uncovered by Grujić *et al.* [23] in the largest experiment reported to date about the repeated spatial PD, carried out on a lattice as in Nowak and May's paper [19] with parameters $T = 1.43$ and $S = 0$.

The results of the experiment were novel in several respects. First, the population of players exhibited a rather low level of cooperation (fraction of cooperative actions in every round of the game in the steady state), hereafter denoted by $\langle c \rangle$. Most important, however, was the unraveling of the structure of the strategies. The analysis of the actions taken by the players showed a heterogeneous population consisting of "mostly defectors" (defected with probability larger than 0.8), a few "mostly cooperators" (cooperated with probability larger than 0.8), and a majority of so-called moody conditional cooperators. This last group consisted of players that switched from cooperation to defection with probability $P_i^{DC} = 1 - d - \gamma c_i = 1 - P_i^{CC}$ and from defection to cooperation with probability $P_i^{CD} = a + \beta c_i = 1 - P_i^{DD}$, c_i being the fraction of cooperative actions in player i's neighborhood in the previous iteration. Conditional cooperation, i.e., the dependency of the chosen strategy on the

amount of cooperation received, had been reported earlier in related experiments [99] and observed also for the spatial repeated PD at a smaller scale [100]. The new ingredient revealed in Grujić *et al.*'s experiment [23] was the dependence of the behavior on the own player's previous action, hence the reason to call them "moody".

5.2 Results

To study how the newly unveiled rules influence the emergence of cooperation in an structured population of individuals, we first report results from numerical simulations of a system made up of $N = 10^4$ individuals who play a repeated PD game according to the experimental observations. To this end, we explored the average level of cooperation in four different network configurations: a well-mixed population in which the probability that a player interacts with any other one is the same for all players, a square lattice, an Erdös-Rényi (ER) graph and a Barabási-Albert (BA) scale-free (SF) network. It is worth mentioning that the dependence on the payoff matrix only enters through the parameters describing the players' behavior (d, γ, a, β and the fractions of the three types of players). Once these parameters are fixed the payoffs do not enter anywhere in the evolution, as this is only determined by the variables c_i, the local fractions of cooperative actions within each player's neighborhood. Thus there is no possibility to explore the dependence on the payoffs because we lack a connection between them and the behavioral parameters.

In Figure 5.1 we present our most striking result. The figure represents, in a color-coded scale, the average level of cooperation as a function of the fraction of mostly cooperators, ρ_C, and mostly defectors, ρ_D, for a BA network of contacts. The same plots but for the rest of topologies explored (lattice and ER graphs) produce indistinguishable results with respect to those shown in the figure. We therefore conclude that the average level of cooperation in the system *does not* depend on the underlying structure. This means that, under the assumption that the players follow the behavior of the experiment in [23], *there is no network reciprocity*, i.e., no matter what the network of contacts looks like, the observed level of cooperation is the same. This latter finding is in stark contrast to most previous results coming out from numerical simulations of models in which many different updating rules —all of them based upon the relative payoffs obtained by the players— have been explored.

Mean-field Approach. The previous numerical findings can be recovered using a simple mean-field approach to the problem. Let the fractions of the three types of players be ρ_C, ρ_D and ρ_X, for mostly cooperators, mostly defectors, and moody conditional cooperators, respectively, with the obvious constraint $\rho_X = 1 - \rho_D - \rho_C$.

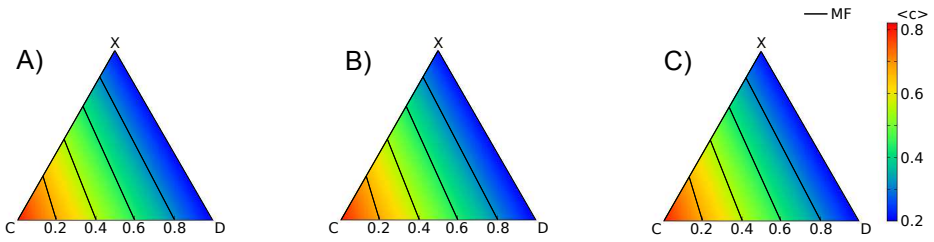


Figure 5.1: **Dependency of the average level of cooperation on the density of strategists.** Density plot of $\langle c \rangle$, as a function of the fractions of the three strategies -mostly cooperators, C, mostly defectors, D, and moody conditional cooperators, X-. Pannels A), B) and C) correspond to a regular lattice ($k = 8$), Erdös-Rényi($\langle k \rangle = 6$) and Barabási-Albert ($\langle k \rangle = 6$) network of contacts respectively, but independence of cooperation level on the topology make them indistinguishable. The system is made up of $N = 10^4$ players and the rest of parameters, taken from [23], are: $d = 0.38$, $a = 0.15$, $\gamma = 0.62$, $\beta = -0.1$. The thin lines represent the mean-field estimations [c.f. Eq. (5.5)] for $\langle c \rangle = 0.32, 0.44, 0.56, 0.68$. They very accurately match the contour lines of the density plot corresponding to those values of $\langle c \rangle$, thus proving that the same outcome is obtained in a complete graph (mean-field). Simulation results have been averaged over 200 realizations.

Denoting by $P_t(A)$ the cooperation probability at time t for strategy $A(= C, D, X)$ of the repeated PD we have

$$\langle c \rangle_t = \rho_C P(C) + \rho_D P(D) + \rho_X P_t(X), \tag{5.2}$$

where $P_t(C) = P(C)$ and $P_t(D) = P(D)$ are known constants [in our case $P(C) = 0.8$, $P(D) = 0.2$]. The probability of cooperation for conditional players in the next time step can be obtained as

$$P_{t+1}(X) = (d + \gamma \langle c \rangle_t) P_t(X) + (a + \beta \langle c \rangle_t)[1 - P_t(X)], \tag{5.3}$$

where the first term in the right hand side considers the probability that a conditional cooperator keeps playing as a cooperator, whereas the second terms stands for the situation in which a moody conditional cooperator switched from defection to cooperation. Asymptotically

$$\lim_{t \to \infty} P_t(X) = P(X), \qquad \lim_{t \to \infty} \langle c \rangle_t = \langle c \rangle.$$

From Eq. (5.3),

$$P(X) = \frac{a + \beta \langle c \rangle}{1 + a - d + (\beta - \gamma) \langle c \rangle}, \tag{5.4}$$

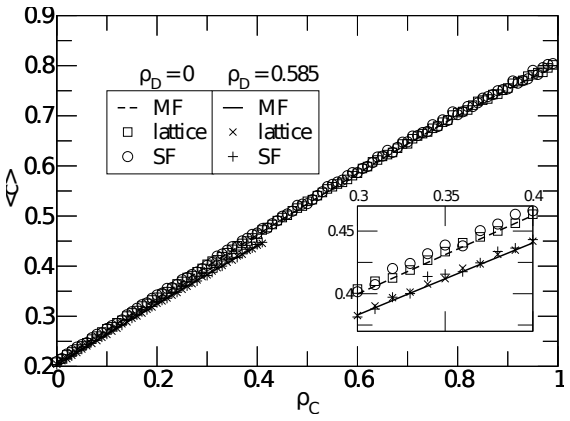


Figure 5.2: **Absence of Network Reciprocity.** Average cooperation level in the stationary state, $\langle c \rangle$, as a function of the density ρ_C of mostly cooperators and two different values of the density ρ_D of mostly defectors, for two different kinds of networks: regular lattice ($k = 8$), and Barabási-Albert network ($\langle k \rangle = 8$). The network size is $N = 10^4$ and the rest of parameters are as in Figure 5.1. Lines represent the mean-field estimations. Results are averages over 200 realizations. The inset is a zoom that highlights how the different curves compare.

thus (5.2) implies (with the replacement $\rho_X = 1 - \rho_C - \rho_D$)

$$A\rho_C + B\rho_D = 1, \tag{5.5}$$

where

$$A \equiv \frac{P(C) - P(X)}{\langle c \rangle - P(X)}, \qquad B \equiv \frac{P(D) - P(X)}{\langle c \rangle - P(X)}, \tag{5.6}$$

are functions of $\langle c \rangle$. From Eq. (5.5) it follows that the curves of constant $\langle c \rangle$ are straight lines in the simplex. Figure 5.1 clearly demonstrates this fact: The straight lines are plots of Eq. (5.5) for different values of $\langle c \rangle$. It can be seen that they are parallel to the color stripes, and that the values of $\langle c \rangle$ they correspond to accurately fit those of the simulations.

Figure 5.2 depicts the curve $\langle c \rangle$ vs. ρ_C for two different values of ρ_D, as obtained from Eq. (5.5) and compared to simulations. This figure illustrates the excellent quantitative agreement between the mean-field result and the simulation results. The match between the analytical and numerical results is remarkable, as it is the fact that the result does not depend on the underlying topology. This is the ultimate consequence of the lack of network reciprocity: the cooperation level on any network can be accurately modeled as if individuals were playing in a well-mixed population.

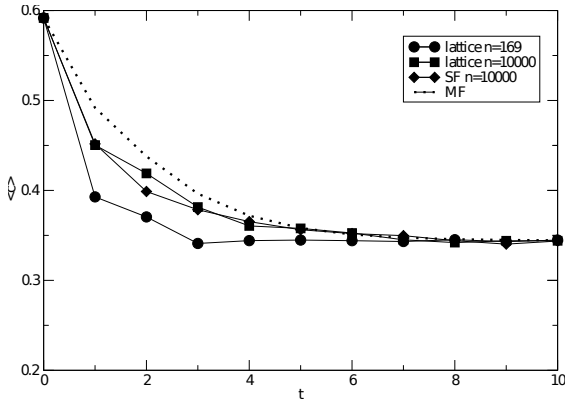


Figure 5.3: **Asymptotic level of Cooperation**. Time evolution of the cooperation level until the stationary state is reached. The results have been obtained from numerical simulations on different networks with different sizes. The Mean-Field curve is the solution of Eq. (5.3). $P(C) = 2/3$, $P(D) = 1/3$, $P(X; t = 0) = 1$, $\langle k \rangle = 8$, $\rho_D = 0.586$, $\rho_C = 0.053$, $d = 0.345$, $a = 0.224$, $\gamma = 0.64$, $\beta = -0.072$. Averages have been taken over 10^3 realizations.

The steady state is reached after a rather short transient, as illustrated in Figure 5.3. This figure compares the approach of the cooperation level to its stationary state as obtained iterating Eq. (5.3) and from numerical simulations on different networks with different sizes. The initial cooperation level has been set to $\langle c \rangle_0 = 0.592$, close to the value observed in the experiment of Ref. [23]. The transient does exhibit a weak dependence on the underlying topology and specially on the network size, but for the largest simulated size ($N = 10^4$) the curves are all very close to the mean-field prediction.

Distribution of Payoffs. The only observable on which the topology does have a strong effect is the payoff distribution among players. Figure 5.4 shows these distributions for the three studied topologies, and at two different times —short and long. Smooth at short times, this distribution peaks around certain values at long times. This reflects the fact that payoffs depend on the number of neighbors of different types around a given player, which yields a finite set of values for the payoffs (the centers of the peaks). These numbers occur with different probabilities (determining the height of the peaks), according to the distribution

$$Q(\mathbf{k}) = \sum_{k \geq 1} \binom{k}{k_C \; k_D} \rho_C^{k_C} \rho_D^{k_D} \rho_X^{k_X} p(k), \tag{5.7}$$

where $p(k)$ is the degree distribution of the network and $\mathbf{k} = (k_C, k_D, k_X)$, but it is understood that $k_X = k - k_C - k_D$. The standard convention is assumed that the multinomial coefficient $\binom{k}{k_C\ k_D} = 0$ whenever $k_C < 0$, $k_D < 0$ or $k_X < 0$.

The approach to a stationary distribution of payoffs exhibits a much longer transient. This is due to the fluctuations in the payoffs arising from the specific actions (cooperate or defect) taken by the players. These fluctuations damp out as the accumulated payoffs approach their asymptotic values. Thus, the peak widths shrink proportionally to $t^{-1/2}$. In fact, one can show that the probability density for the distribution of payoffs Π for strategy Z can be approximated as

$$W_Z(\Pi) = \sum_{\mathbf{k}\geq 1} G\big(\Pi - a_k(Z)\mu(\mathbf{k}), \sqrt{t}a_k(Z)\sigma(\mathbf{k})\big)Q(\mathbf{k}), \tag{5.8}$$

where $G(x,\gamma) \equiv (2\pi\gamma^2)^{-1/2}e^{-x^2/2\gamma^2}$, the mean payoff per neighbor received by a Z strategist against a cooperator is

$$a_k(Z) \equiv \frac{1}{k}\{P(Z) + T[1 - P(Z)]\},$$

with $k = k_C + k_D + k_X$, and the average cooperation level in the neighborhood of the focal player and its variance are

$$\begin{aligned} \mu(\mathbf{k}) &\equiv k_C P(C) + k_D P(D) + k_X P(X), \\ \sigma(\mathbf{k})^2 &\equiv k_C P(C)[1 - P(C)] + k_D P(D)[1 - P(D)] \\ &\quad + k_X P(X)[1 - P(X)]. \end{aligned}$$

The approximate total payoff distribution, $W(\Pi) = \rho_C W_C(\Pi) + \rho_D W_D(\Pi) + \rho_X W_X(\Pi)$, is compared in Figure 5.4 with the results of the simulations for the longest time.

5.3 Discussion

In this work we have shown both analytically and through numerical simulations that if we take into account the way in which humans are experimentally found to behave when facing social dilemmas on lattices, no evidence of network reciprocity is obtained. In particular, we have argued that if the players of a Prisoners' Dilemma adopt an update rule that only depends on what they see from their neighborhood, then cooperation drops to a low level —albeit nonzero— irrespective of the underlying network. Moreover, we have shown that the average level of cooperation obtained from simulations is very well predicted by a mean-field model, and it is found to depend only on the fractions of different strategists. Additionally, we have also shown

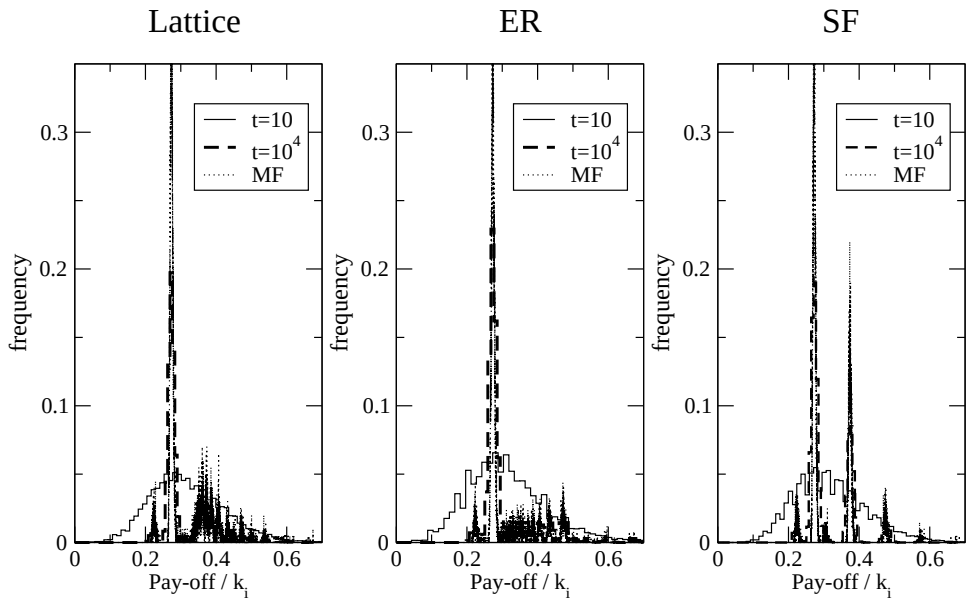


Figure 5.4: **Payoffs Distribution**. Distribution of the pay-off per neighbor in the stationary state for different network topologies: regular lattice ($k = 8$), Erdös Rényi ($\langle k \rangle = 8$) and Barabási-Albert network ($\langle k \rangle = 8$). Solid and dashed lines represent the results of numerical simulations for two values of time: $t = 10$ (solid lines) and $t = 10^4$ (dashed lines) while pointed curves represent the theoretical estimations for the density probabilities at $t = 10^4$, as obtained from Eq. (5.8). $N = 10^4$, $\rho_D = 0.586$, $\rho_C = 0.053$, and other parameters are as in Figure 5.1. The simulation results are averages over 10^3 realizations.

that the underlying network of contacts does manifest itself in the distribution of payoffs obtained by the players, and has a slight influence on the transient behavior.

To conclude, it is worth mentioning that our results only make sense when applied to evolutionary game models aimed at mimicking human behavior in social dilemmas. The independence on the topology seems to reflect the fact that humans update their actions according to a rule that ignores relative payoffs. Interestingly, absence of network reciprocity has also been observed in numerical simulations using best response dynamics [89], an update rule widely used in economics that does not take into account the neighbors's payoffs. This suggests that the result that networks do not play any role in the repeated PD may be general for any dynamics that does not take neighbors' payoffs into account. We want to stress that the same kind of models thought of in a strict biological context are ruled by completely different mechanisms which do take into account payoff (fitness) differences. Therefore, in such contexts lattice reciprocity does play its role. In any case, our results call for further experiments that uncover what rules are actually governing the behavior of players engaged in this and other social dilemmas.

Chapter 6

Heterogeneous networks do not promote cooperation when humans play a Prisoner's Dilemma.

It is not fully understood yet why we cooperate with strangers on a daily basis. In an increasingly global world, where interaction networks and relationships between individuals are becoming more complex, different hypotheses have been put forward to explain the foundations of human cooperation on a large scale and to account for the true motivations that are behind this phenomenon. In this context, population structure has been suggested to foster cooperation in social dilemmas, but theoretical studies of this mechanism have yielded contradictory results so far, and the issue lacks a proper experimental test in large enough systems. We have performed the largest experiments to date with humans playing a spatial Prisoner's Dilemma on a lattice and on a scale-free network (1229 subjects). We observed that the level of cooperation reached in both networks is the same, comparable to that of smaller networks or unstructured populations. We have also found that subjects respond to the cooperation they observe in a reciprocal manner, being more likely to cooperate if in the previous round many of their neighbors and themselves did so. This implies that humans do not consider neighbors' payoffs when making their decisions in this dilemma, but only their actions. Our results, that are in agreement with recent theoretical predictions based on this behavioral rule, suggest that population structure has little relevance as a cooperation-promoter or inhibitor among humans.

6.1 Introduction

The strong cooperative attitude of humans defies the paradigm of *homo economicus* and poses an evolutionary conundrum [123, 124]. This is so because many of our interactions can be framed as Prisoner's Dilemmas [93, 125, 126] or Public Goods Games [127], famous for bringing about a "tragedy of the commons" [128]. Several mechanisms have been suggested as putative explanations of cooperative behavior [129], among which the existence of an underlying network of contacts constraining who one can interact with has received very much attention. This mechanism was first proposed by Nowak and May [19], whose simulations on a square lattice with agents that imitate the behavior of their neighbor with the highest payoff showed high levels of cooperation in the Prisoner's Dilemma. The ensuing two decades have witnessed a wealth of theoretical studies that have concluded that this so-called "network reciprocity" [129] is indeed possible under a variety of circumstances, but in many other contexts networks do not promote −or even inhibit− cooperation [16, 88]. The effect of regular and homogeneous networks on cooperation is very sensitive to the details of the model (e.g., dynamics, clustering), while theoretical results and simulations indicate that heterogeneous networks should be particularly efficient in fostering cooperation in social dilemmas [108, 109, 88]. A natural way to shed some light on these partially contradictory results would be to test experimentally the predictions of the different models. Such tests are currently lacking [130], as the few available experimental works only deal—with some exception [23]— with very small networks [131, 132, 100]. Interestingly, the only theoretical result [133] that takes into account the behavioral information extracted from experiments predicts that neither homogeneous nor heterogeneous networks would influence the cooperative behavior in the Prisoner's Dilemma, i.e., the observed cooperation level should be the same as if every player interacted with every other one.

6.2 The experiments.

Here, we close the cycle by testing the above theoretical predictions [133] and contributing to the current debate on the existence and effects of network reciprocity by performing experiments on large samples of structured populations of individuals who interact through a Prisoner's Dilemma (PD) game. Specifically, we have designed a setup in which 1229 human subjects were placed either in a square lattice or in a scale-free network, and for more than 50 rounds they played a 2×2 multiplayer PD game with each of their k neighbors, taking only one action, either to cooperate (C) or to defect (D)—the action being the same against all opponents. The experiment was simultaneously carried out on two different virtual networks: a 25×25

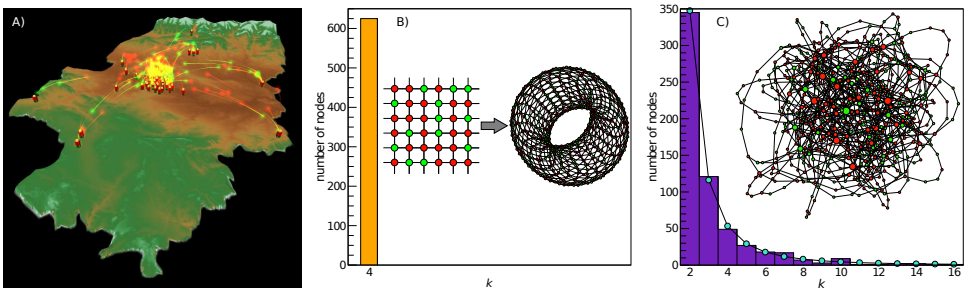


Figure 6.1: **Players in the experiment were sitting in different physical locations, but played in two virtual networks.** Panel A is a snapshot at round 10 of a graphic animation illustrating the activity during the experiment. On a map of Aragón the image displays small buildings representing the schools. Arrows (green for cooperate and red for defect) represent actual actions taken by players. They travel towards the school where their randomly assigned neighbors were sitting. Buildings are colored green and red, proportional to the respective number of cooperative and defective actions taken by the subjects in that school. The height of the yellow column on top of each building is proportional to the school's accumulated payoffs. Panels B and C show snapshots of the two networks at that same round, along with their degree distributions (in the case of the heterogeneous network, both the theoretical distribution and the actual realization corresponding to the network of the experiment are represented). Colors indicate the corresponding player's action (green for cooperate, red for defect). The size of a node is proportional to its degree.

regular lattice with $k = 4$ and periodic boundary conditions (625 subjects), and a heterogeneous network with a fat-tailed degree distribution (604 subjects, the number of neighbors varied between $k = 2$ and $k = 16$). Figure 6.1 depicts a snapshot of a visual representation of the experiment as well as of the two networks. Subjects played a repeated (weak) Prisoner's Dilemma (PD) with all their neighbors for an initially undetermined number of rounds. Payoffs of the PD were set to be 7 ECUs for mutual cooperation, 10 ECUs for a defector facing a cooperator, and 0 ECUs for any player facing a defector (weak PD [19]). We note that this choice of payoffs is as in Grujić et al.'s experiment on a smaller regular lattice [23] (see Figure 6.1) and such that cooperation should reach a high level according to the available simulations [19, 88, 108, 109]. The size of each network was large enough so that clusters of cooperators could form (the underlying mechanism by which cooperators may thrive [134, 115]).

On this general setup, we carried out two treatments, which we will refer to as experiment and control. In the experiment, subjects remained at the same positions in the network with the same neighbors throughout all the rounds played. In the control treatment we removed the effect of the network by shuffling the neighbors of each subject in every round. Therefore, in this phase, the players were always connected to

the same number of neighbors, but these neighbors changed from round to round. On the screen, subjects saw the actions and normalized payoffs of their neighbors from the previous round, who in the control treatment were different from their current neighbors with high probability. All treatments of the experiment were carried out in sequence with the same subjects. Players were also fully informed of the different setups they were going to run through. The number of rounds in each treatment was randomly chosen between 50 and 70 in order to avoid subjects knowing in advance when it was going to finish, resulting in 51 and 59 rounds for the experimental and control treatments, respectively.

6.3 Results and Discussions.

Figures 6.2A and 6.2B show the fraction of cooperative actions, c, in each round for the two networks and for both treatments. The first feature worth noticing in this figure is that, in the experiment phase, the level of cooperation in either network quickly drops from initial values around 60% to values around 40% and finally settles at a slower pace around 30%, much lower than theoretical models predict [19, 16, 88]. This is especially remarkable for the heterogeneous network, on which no previous results are available, and is in stark contrast with the predictions that this kind of networks should be particularly efficient in promoting cooperation [108, 109, 88]. In the control, the initial level of cooperation is already at these low values. This behavior is consistent with previous findings in experiments with smaller lattices [100, 23] as well as with unstructured populations [135, 136]. Regarding the slow decay undergone by these curves after the first quick drop in the level of cooperation, we believe that this is associated to a process of learning (see below). However, the most remarkable result that this figure provides is that, quite unexpectedly, the network does not have any influence in the evolution of the level of cooperation. In fact, both curves are nearly identical—the slightly lower values obtained for the lattice are likely to arise from the small difference in the initial level of cooperation—despite the very different nature of the networks of contacts between the players.

The experimental result we have just reported is in very good agreement with the theoretical prediction in [133]. This prompts us to investigate in detail what is the players' behavior, as the reason why this prediction was different from earlier ones is the use of the update rule observed in [23]. The distributions of subjects by their individual cooperation levels (averaged over the whole experiment) depicted in Figures 6.2C and 6.2D show quite some heterogeneity of behavior: a few subjects have a high level of cooperation (above 70%), a sizable fraction cooperated less than 20% of the rounds, whereas the bulk of subjects have intermediate levels of cooperation. Importantly, the comparison of these distributions of actions, which turn out to be

statistically indistinguishable (see Kolmogorov-Smirnov test data on Table 6.1 of the appendix 6.5), provides additional evidence that the behavior observed in the two networks is the same. This finding, along with the identical behavior of the cooperation level, suggests that subjects use the same strategies in the lattice and in the heterogeneous network, regardless of the fact that in the latter the number of neighbors of each individual is heterogeneously distributed.

Figure 6.4 provides further evidence of the significance of the moody conditional cooperation by means of a nonparametric bootstrap check. The series of actions taken by every individual are randomly reassigned to other positions in the lattice or the network and the probability of cooperation is recomputed. This is done 10^6 times and the results show that the two probabilities become independent of the context. Of course, such a reshuffling will not change the dependence on the player's own previous action, as the order of the actions is not altered, and consequently there are still two distinct lines corresponding to the probability of cooperation following cooperation or defection, but the dependence on the number of cooperators in the previous round is fully removed.

The existence of (almost pure) cooperators and defectors aside from moody conditional cooperators can be further supported through a comparison with the same histograms but for the control condition (see Figure 6.6 of the appendix), since for the latter a larger number of subjects are in the region that would correspond to defectors. This can be interpreted as an indication that a fraction of—probably—moody conditional cooperators changed to a defective strategy, given that retaliation is ineffective in the control condition. Furthermore, performing running averages of the levels of cooperation during the experiment condition (see Figures 6.7 and 6.8 of the appendix) shows that the number of subjects whose level of cooperation is below a given threshold increases with time—irrespective of the precise value of the threshold. Not only this gives support to the existence of this kind of players, but it is consistent with a continuous (albeit small) flow of players who change from moody conditional cooperation to defection—a behavior that could be understood as a generalized form of a grim strategy. Notice that this flow can account for the slow decay observed all along the run of the experiment and control observed in Figures 6.2A and 6.2B.

Finally, another important point that our experiment allows to address to some extent is the dependence of the actions on the connectivity of the participants for the heterogeneous network. The results are displayed in Figure 6.5, where we represent the average cooperation level c as a function of the connectivity of the players, k, for both treatments: experiment and control. As can be seen from the plots, there might be some trend towards lower levels of cooperation with increasing degree for small connectivities, particularly in the control (the levels for the first three values

of the degree in the experiment are not statistically different). However, looking at the figure as a whole it becomes clear that there does not seem to be any statistically significant trend. It has to be borne in mind that in this type of networks the number of hubs or large-degree nodes is intrinsically small, and therefore the statistics for them is not very accurate (notice the size of the error bars). Going beyond this results would require much larger networks (which would still have the same problem for their higher degree nodes). Additionally, the bottom panels of Figure 6.5 show the frequency of cooperative actions of nodes with degree k after playing as C or D with respect to the fraction of their neighbors that cooperated in the previous round. The results are a clear evidence that moody conditional cooperation is indeed the general behavior even if one disaggregates the data in terms of their degree. As we have already stated above for the total level of cooperation, for higher degrees the statistics is poorer and the analysis does not lead to such clear-cut results.

6.4 Conclusions

To sum up, we have performed a large-scale experimental test of the hypothesis of network reciprocity, i.e., that the existence of a structure in the population may affect cooperation in social dilemmas. Our experiment shows that, when it comes to human behavior, the existence of an underlying network of contacts does not seem to have any influence in the global outcome. We want to stress that this conclusion applies only to human cooperation, and network reciprocity may still be relevant in other contexts, e.g., in microbiology [137]. Players seem to act by responding to the level of cooperation in their neighborhood, and this renders the network irrelevant. In addition, players behave in a 'moody' manner, being significantly less likely to cooperate following a defection of their own. The levels of cooperation attained in a regular lattice and in a highly heterogeneous network (hitherto thought to be a cooperation enhancer) are indistinguishable, and the responsive behavior of subjects appears to be independent of the number of neighbors they have or on the payoff differences they observe. The results are in full agreement with the theoretical prediction in [133]; the fact that the key hypothesis in that model is that people behave in the way we have just described, provides further support to our finding of moody conditional cooperation in networked Prisoner's Dilemmas.

Our results have implications for policy making when cooperation is a desired behavior. Although further experiments with other social dilemmas still need to assess the range of applicability of our conclusions, the present study suggests that imposing a network structure might be a sterile effort. It should be noted, however, that this caveat does not imply that networking is useless to achieve cooperation—results would probably be very different if the network is allowed to be formed by

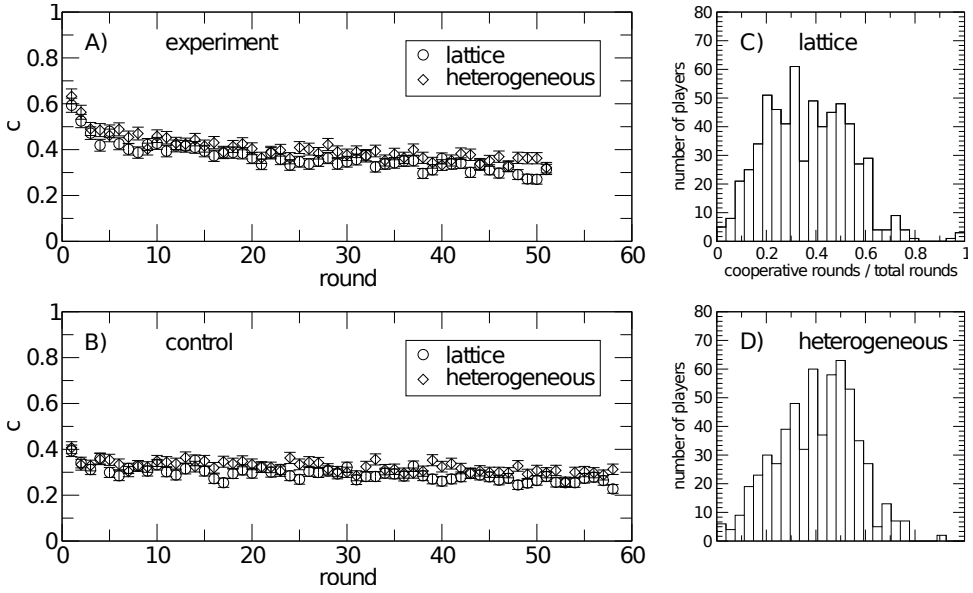


Figure 6.2: **The level of cooperation declines and is independent of the network of contacts.** Fraction of cooperative actions (level of cooperation) per round during the experiment (panel A) and the control (panel B) for both networks, and histograms of cooperative actions in the lattice (panel C) and in the heterogeneous network (panel D). The histograms (panels C and D) show the number of subjects ranked according to the fraction of cooperative actions they perform along the experiment in the two networks. A Kolmogorov-Smirnov test shows that the distributions are statistically indistinguishable (see appendix 6.5). They illustrate the high heterogeneity in subjects' behavior, their levels of cooperation ranging from nearly zero to almost one in a practically continuous distribution. The corresponding histograms for the control (Figureexp.figS4 of the appendix) show that a sizable group of subjects lowered their levels of cooperation hence becoming mostly defectors. Actually, the decline in the level of cooperation observed in the experiment (panels **A** and B) can be explained as a constant flow of subjects to more defective strategies (for evidence supporting this hypothesis see Figures 6.7 and 6.8 of the appendix).

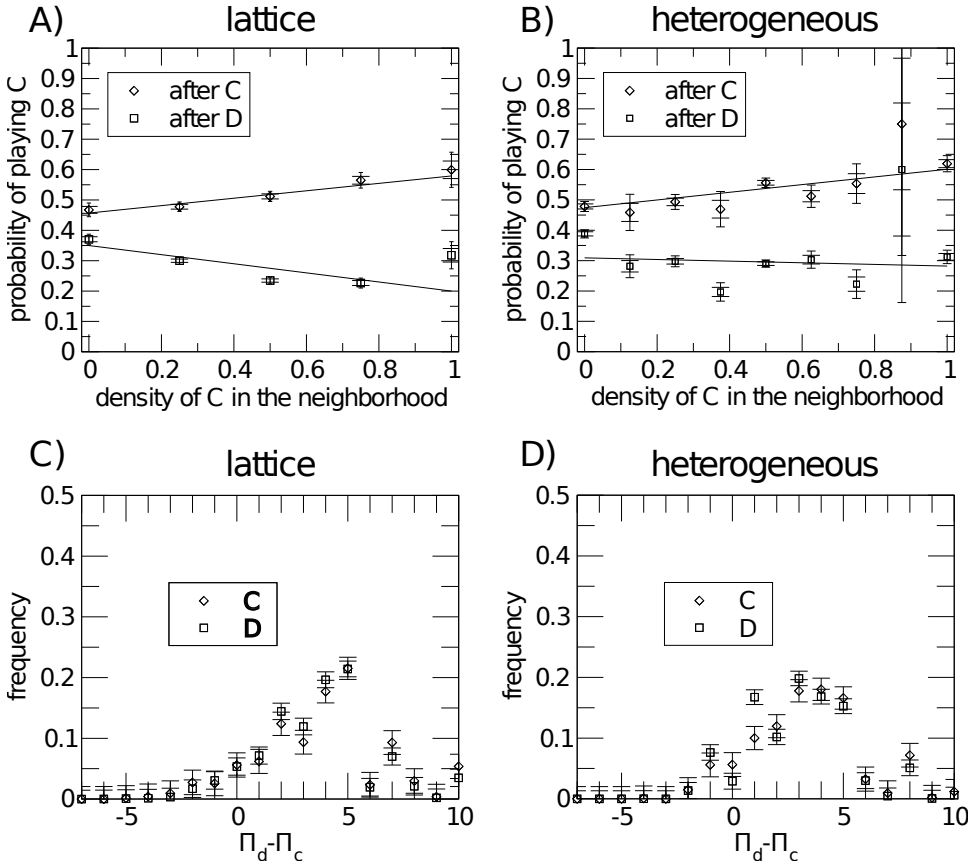


Figure 6.3: **Players' behaviour depends both on the level of cooperation in the neighborhood and on their previous action.** Frequency of cooperative actions after a cooperative/defective action, conditioned to the context (fraction of cooperative actions in the neighborhood in the previous round) observed in the lattice (A) and in the heterogeneous network (B). Details of the linear fits and comparison with randomizations to prove statistical significance can be found in the appendix. The plots demonstrate that there is a relevant dependence on the context for subjects that cooperated in the previous round (i.e., were in a "cooperative mood"), the cooperation probability increasing with the fraction of cooperative neighbors much as for the conditional cooperators found by Fischbacher *et al.*[99]. However, after having defected, this dependence is less clear, and if anything, it suggest an exploiting behavior—subjects who defected are less prone to cooperate the more cooperation they find around. Panels C and D show how subjects who cooperated or defected are distributed according to the largest payoff-per-link difference in their neighborhoods between the two actions. These plots reveal that a player's decision to cooperate or defect was independent on the payoffs-per-link they observed (an information that was explicitly provided during the experiment).

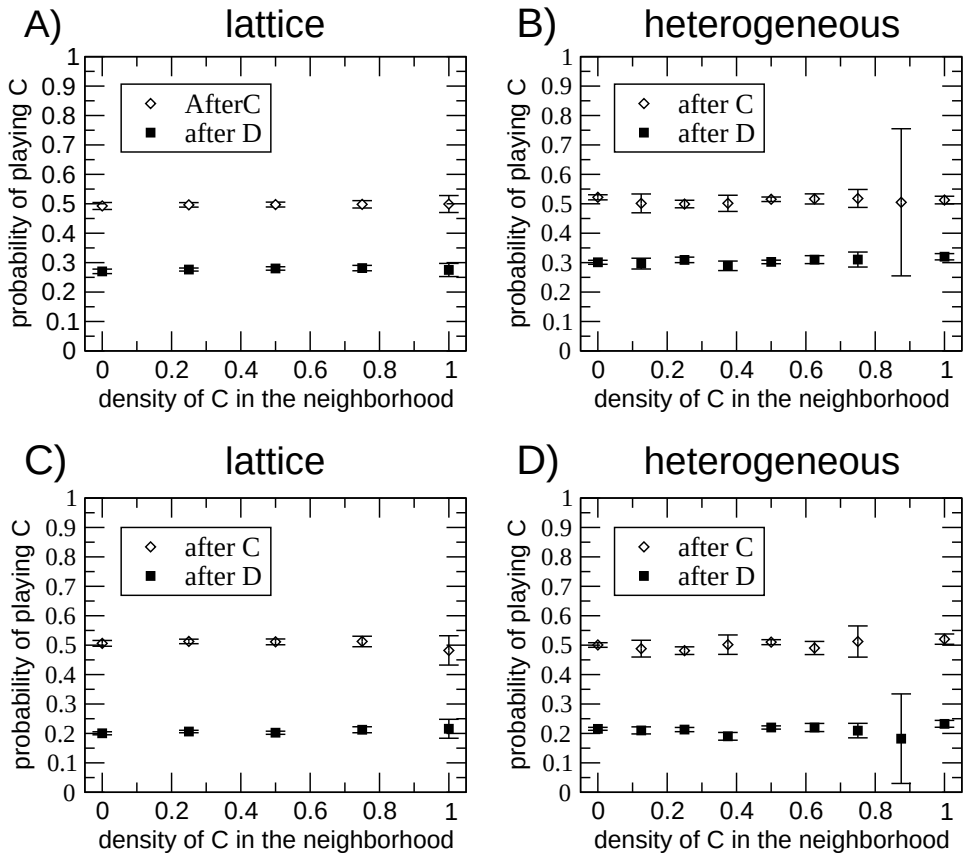


Figure 6.4: **Null hypothesis statistical significance test.** Probability of cooperating after playing C or D, conditioned to the context (fraction of cooperative actions in the neighborhood in the previous round), averaged over 10^6 random shuffling of players. Panel A) corresponds to the experimental treatment in the lattice, panel B) to the same treatment but for the heterogeneous network, panel C) to the control phase in the lattice and panel D) to the same control treatment for the heterogeneous network. The results show that there is no dependence on the context and hence that the results of panels A and B of Figure 3 are statistically relevant. The anomalous variance (or even absence of data) observed at a fraction of C's in the neighborhood close to 0.9 is not a relevant feature of the experimental results but a consequence of the very low probability of having events contributing to that bin of the histogram in the heterogeneous network. This anomaly can also be noticed in Figure 3.

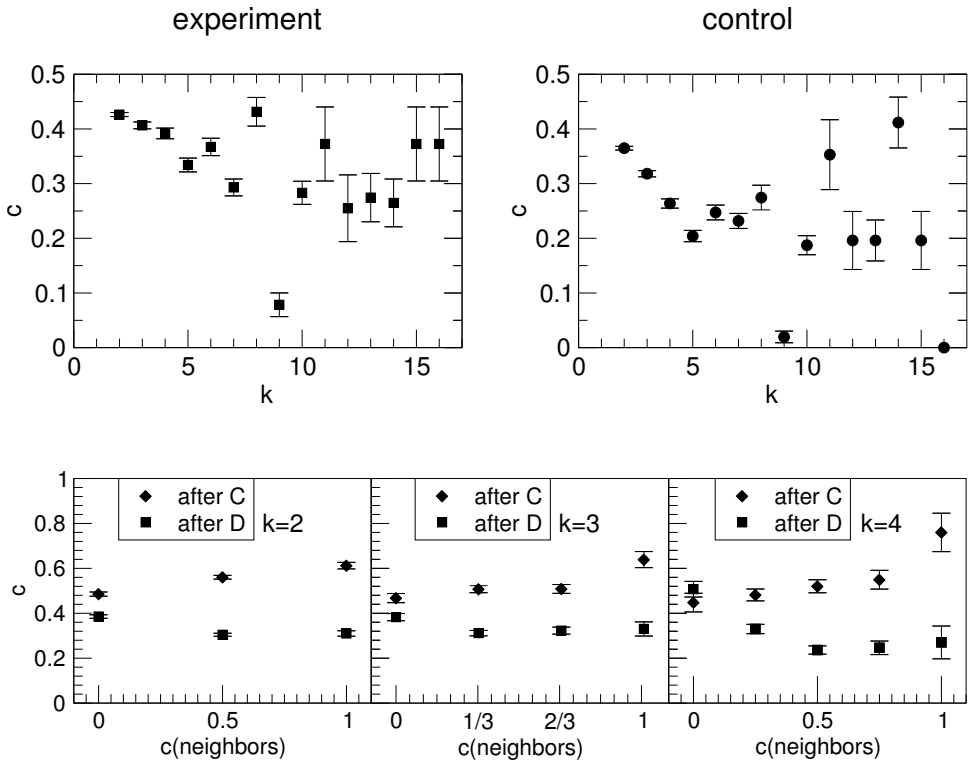


Figure 6.5: **Dependence of the strategies on the connectivity.** The upper panels show the cooperation level c as a function of the connectivity k_i in the heterogeneous network, averaged over all rounds of the experiment (upper left panel) and the control (upper right) of the experiment. In the lower panels, we plot the frequency of cooperative actions of players with degree as indicated, after they have cooperated or defected, as a function of the fraction of cooperative actions in their neighborhood during the previous round, along the experiment treatment in the heterogeneous network. Statistics is restricted to nodes of connectivity $k = 2$ (lower left panel), $k = 3$ (lower center) and $k = 4$ (lower right).

the subjects as part of the game. Recent experiments on groups of up to 20 people [138, 139] strongly suggest this, but to the best of our knowledge no large-scale experiments have been carried out to test this issue. On the other hand, the theoretical work in [133] does not predict the slow decay of the cooperation level observed in the experiments, which we have conjectured that arises from moody conditional cooperators becoming defectors in a generalized grim behavior. Such a change in the percentage of players using different strategies is not included in the theoretical model, and therefore a next step would require to account for such changes and, if possible, to justify them within an evolutionary framework. Finally, given that in our setup players have to play the same action with all their neighbors, it is clear that our results should be related to those of public goods experiments. In fact, conditional cooperation was first observed in that context [99]. Our findings suggest that the "moody" version we have found can also arise in public goods games. If that is the case, it is likely that network reciprocity does not apply to public goods games on networks. Hopefully our experiment will encourage further research in all these directions.

6.5 Appendix 1: Additional material about the experimental results

Here we present further results aimed at supporting the findings shown in the previous sections. As there, we will refer to the basic types of individuals found in the experiment as mostly cooperators (players who cooperate with a high probability regardless of the context), mostly defectors (players who defect with a high probability regardless of the context) and "moody" conditional cooperators (players whose action depends on their previous action as well as the level of cooperation in their neighbourhood, see Fig. 6.3 A and B).

Figure 6.6 shows the histograms of the number of players ranked according to the fraction of cooperative actions they performed along the control phase, in the lattice (panel **A**) and in the heterogeneous network (panel **B**). The same results but for the experimental phase can be found in panels **C** and **D** of Figureexp.fig:2. The comparison between the plots shows a large increase in the fraction of individuals that never or almost never cooperated in the control with respect to the experiment. This is likely to be a consequence of the fact that in the experiment there is an initial amount of cooperation well above 50%, which is not the case in the control. At the other extreme of the plots, the (small) amount of highly cooperative players remains approximately the same, indicating that their motivation has nothing to do with having or not a fixed environment for their interactions. The general picture thus arising

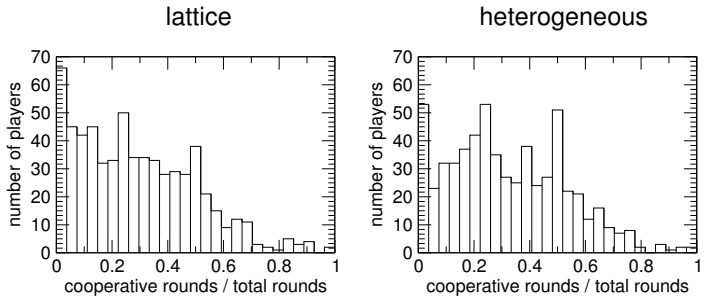


Figure 6.6: **Distribution of cooperative actions in the control.** We represent the number of players that cooperated during the given number of rounds (normalized by the total number of rounds played). The results correspond to the control phase. Similar results were presented Fig.6.2.

from the control part is that there is not much cooperation, and the majority of players do not cooperate other than occasionally.

On the other hand, Figure 6.7 displays the time evolution of the distribution of cooperative actions in the experimental part. The histograms show the players' frequency as a function of the fraction of cooperative actions along successive 10-round periods corresponding to the experimental phase in the lattice (left column) and in the heterogeneous network (right column). The results show evidence of some degree of learning as the experiment progresses: Indeed, the number of people who cooperate never or rarely increases with time. This would be consistent with the decay of cooperation shown in Fig. 6.2 A; while the first, quick drop in cooperation would be explicable within a computer model with a fixed proportion of defectors, cooperators and moody conditional cooperators, the second part of the evolution, a much slower decay, is inconsistent with such a model and must then come from changes in the proportion of the different types of players.

The phenomenon we have just described can also be shown in a different manner, namely by monitoring the evolution of mostly defectors both during the experimental and control parts of the experiment. Figure 6.8 represents the fraction of agents whose probability to cooperate is below a given threshold (indicated in the rightmost legend) at every round (time t). To calculate this quantity, we have taken into account the actions of the players during the previous 10 rounds. The results obtained show an increasing trend (more evident for the experimental phase, top panels) for both the square lattice and the heterogeneous network, which confirms the tendency of the players to learn that they should defect as time goes on.

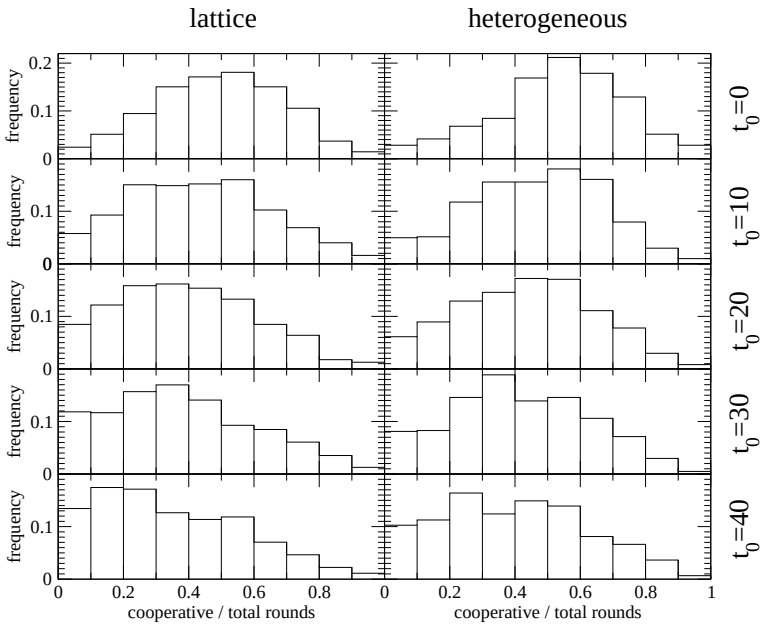


Figure 6.7: **Time evolution of the distribution of cooperative actions.** The different panels show how frequently players cooperated in different time periods. The results correspond to the first treatment (experiment). Rows represent periods 1-10($t_0 = 0$, top), 11-20($t_0 = 10$), 21-30($t_0 = 20$), 31-40($t_0 = 30$) and 41-50($t_0 = 40$, bottom) as indicated.

We also report on the statistical analysis we carried out about the experimental data. First, in order to determine whether or not the likelihood to cooperate differs significantly in the two studied networks, we use the Kolmogorov-Smirnov (KS) test for the two data sets. We take as a first sample the distribution of the probability to cooperate in the lattice, cumulated over all rounds of the experimental phase. The second sample used as input for the KS test corresponds to the same distribution but for the heterogeneous network. These are the distributions represented on Figure 6.2. The maximum difference between the cumulative distributions for the experimental phase is 0.1071 with a corresponding value for $P_{KS} = 0.995$. The statistics of both samples, together with the ones corresponding to the control phase Fig. 6.6 are summarized in Table 6.1.

Finally, Table 6.2 summarizes the statistical fits (obtained from a weighted least squares regression) of the conditional probability P to cooperate, conditioned on the player's action in the previous round (X=after C, after D) and on the density ρ of cooperators in the players' neighborhoods during the previous round. Fits are defined by $P(C|X, \rho) = a + b\rho$. The data fitted correspond to the results shown in Figure 6.3 **A,B**.

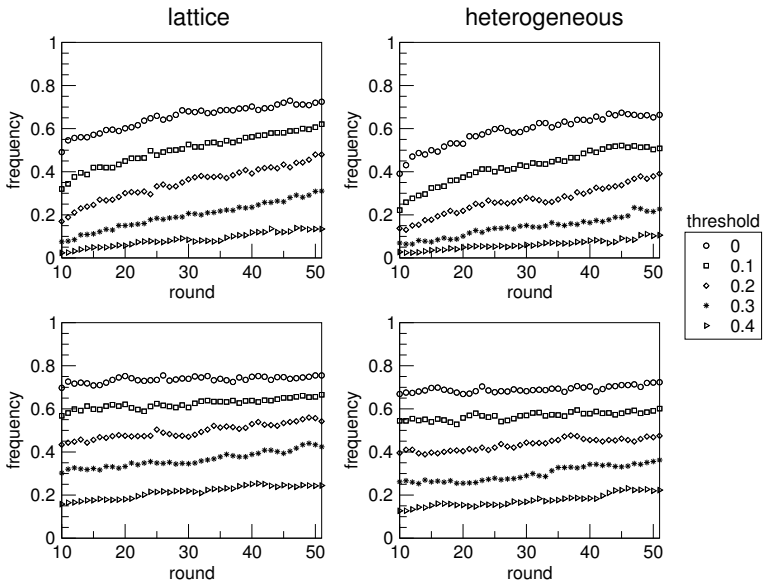


Figure 6.8: **Evolution of the fraction of mostly defectors.** Fraction of agents with a cooperation probability lower than a given $threshold$ as a function of t(=round), according to their cooperative actions through the previous 10 rounds, for different values of the $threshold = 0, 0.1, 0.2, 0.3, 0.4$. Columns represent results for the lattice (left) and the heterogeneous network (right), while rows correspond to the two treatments: experiment (top) and control (bottom).

6.6 Appendix 2: Experimental setup.

6.6.1 Volunteer recruitment and treatment

The experiment was carried out with 1229 volunteers chosen among last year high-school students (17-18 years old) of 42 different High Schools located throughout the geography of the Autonomous Region of Aragón, Spain, whose capital is Zaragoza, where the University of Zaragoza is. 34 High Schools were in the province of Zaragoza, 5 in the province Huesca, and 3 in the province of Teruel. For the recruitment of the students, we contacted the coordinators of a program (Ciencia Viva, "Living Science") of the local government that supports the dissemination of Science among public high schools of Aragón. Moreover, we also contacted many of the private schools of Zaragoza City also offering them the possibility of taking part in the experiment. In all cases, the experiment was referred to as "a social experiment" and nobody (including the high-school teachers in charge of the coordination) knew in advance what the experiment was about (see below).

	experiment		control	
	lattice	heterogeneous	lattice	heterogeneous
mean	0.03703	0.03703	0.03226	0.03226
95% confidence	(0.02434,0.04974)	(0.02335,0.05072)	(0.02549,0.04858)	(0.02607,0.04800)
standard deviation	0.03210	0.03459	0.02918	0.02772
high	0.0976	0.104	0.106	0.0878
low	0	0	0	0
third quartile	0.06560	0.06126	0.05440	0.05795
first quartile	0.006400	0.006623	0.006400	0.01159
median	0.04000	0.03146	0.0448	0.03808
Median absolute deviation	0.02844	0.02937	0.02495	0.02275

Table 6.1: Statistics of the distribution of the probability to cooperate cumulated over all rounds of the experimental and control phases in both networks. See the text for further details.

	lattice		heterogeneous	
	a	b	a	b
after C	0.457 ± 0.015	0.122 ± 0.034	0.475 ± 0.016	0.126 ± 0.039
after D	0.350 ± 0.021	-0.149 ± 0.050	0.309 ± 0.069	-0.0269 ± 0.035

Table 6.2: Values of the fitting parameters for the results shown in Figure 6.3 **A,B**. Fits are defined by $P(C|X,\rho) = a + b\rho$, being X=after C, or after D. See the text for more details.

Following the call for participation, we selected 1300 volunteers. In order to satisfy ethical procedures, all personal data about the participants were anonymized and treated as confidential. Moreover, the Ethical Committee of the University of Zaragoza approved all procedures. On the day of the experiment, the aforementioned 1229 volunteers showed up, with 541 males and 688 females representing the 44.02% and 55.98% of the total number of players, respectively. Out of the 1229 students, 625 played the game on a square lattice (274 males and 351 females keeping the male to female ratio) and 604 on an heterogeneous network. In the first topology, every player had $k = 4$ neighbors while in the second, the connectivity varied between 2 and 16 following a distribution $\frac{N(k)}{N} = P(k) = Ak^{-2.7}$, with $A = (\sum_k P(k))^{-1}$.

All the students played via a web interface specifically created for the experiment (see below) that was accessible through the computers available in the computer rooms of their respective schools. At least one teacher supervised the experiment in each computer room (which at most had a maximum capacity of 20 students), preventing any interaction among the students. To further guarantee that potential interactions among students seating next to each other in the class do not influence the results of the experiment, the assignment of players to the different topologies was completely random. Hence, the odds of having two participants geographically close (i.e., of the same school and seating next to each other) who were also neighbors in the virtual topology was quite small. In addition, as described below, the colors used

to code the two available actions of the game were also selected randomly, further decreasing the likelihood that neighboring participants could influence each other.

We describe in the following section the steps followed by each participant during the experiments. In short, all participants went through a tutorial on the screen, including questions to check their understanding of the game. When everybody had gone through the tutorial, the experiment began, lasting for approximately an hour. At the end of the experiments volunteers were presented a small questionnaire to fill in. Immediately after, all participants received their earnings and their show-up fee. Total earnings in the experiment ranged from 2.49 to 40.48 euros.

6.6.2 Experimental platform and interface

The experiment was run using a fairly sophisticated web application specifically developed to this purpose. The application was made entirely using free software. It was developed in Ruby On Rails, a technology used by other popular websites like Twitter, and has a MySQL database that stores all data needed to carry out the experiment and the subsequent analysis. MySQL is a freely available open source relational database management system based on Structured Query Language (SQL), the most popular language for adding, accessing and managing content in a database.

The application was designed to be used by three different user profiles. First of all, we have the players, who were shown at the beginning a detailed tutorial for a better understanding of the interface and basis of the experiment. Secondly, there are teachers who were responsible for supervising students through their dedicated web monitors, facilitating the work of the central administrator work and contributing to the success of the experiment. Finally, the administrators were responsible for controlling the game and everything that was happening in real time. The application, which was designed using technologies compatible with all platforms, was managed from a standard web browser. There was a last participant, a daemon or process running in the background whose function was to update the results and play instead of players who do not play within the specified time frame for each action.

Considering that the experiment required that around 1300 students could play online simultaneously, we used a server with enough power, and many optimizations were performed in terms of connections to the server, access to database, client-server data exchange, lightness of the interface, control logic, etc. The experiment started on December 20, 2011 at 10:00 CET. The steps followed during the development of the experiment were:

1. Administrators opened the registration process.
2. Players (students) gradually registered.

3. Once all students had registered, teachers informed the administrators via their screen.

4. As soon as the required number of participants have registered (this took around 20 minutes), administrators blocked further registrations and initiated the reading of the tutorial.

5. Students and teachers read the tutorial.

6. Teachers informed (also via their screens) administrators that the reading was completed.

7. The experiment treatment began, which lasted 51 rounds.

8. Students played according to some predefined times (a maximum of 20 seconds per round to choose an action). During these steps, teachers controlled for any potential problem using a chat channel that connected them to the administrators. As mentioned above, if one student did not play within the 20 seconds given for each action, the daemon played automatically (see below). The administrators were able to identify who was not playing and to contact the teachers if the situation persisted. However, the experiment went smoothly and no feedback to the professors for misbehavior was needed.

9. The experiment treatment finished and a brief tutorial on the second one (control) was shown.

10. Once teachers and students had read the tutorial, the former notified the administrators.

11. Administrators started the control treatment, which lasted 59 rounds.

12. Students played as in the previous treatment.

13. Once the control treatment finished, volunteers were presented a short questionnaire to fill in.

14. All participants checked their earnings and were given their show-up fee.

6.6.3 Synchronous play and automatic actions

The experiment assumes synchronous play, thus we had to make sure that every round ended in a certain amount of time. This playing time was set to 20 seconds, which was checked during the testing phase of the programs to be enough to make a decision, while at the same time not too long to make the experiment boring to fast players. If a player did not choose an action within these 20 seconds, the computer made the

decision instead. This automatic decision was randomly chosen to be the player's previous action 90% of the times and the opposite action 10% of the times. We chose this protocol following previous testings performed by the authors of a similar experiment (see [23]). Volunteers were informed that the computer would play for them if their decision took more than the prescribed time-out. However, they were not informed of the precise strategy used by the computer in order to avoid any bias in their own choices of strategy. In any case, for the reliability of the experiment it is important that a huge majority of actions were actually played by humans, not by the computer. This quantity, when averaged over all rounds, yields that the 90% of the actions were chosen by humans, regardless of the underlying network of contacts.

6.6.4 Questionnaires

At the end of the experiments volunteers were presented a small questionnaire to fill in. The list of questions (translated into English) was the following:

1. Describe briefly how you made your decisions in part I (Experiment).
2. Describe briefly how you made your decisions in part II (Control).
3. Did you take into account your neighbors' actions?
4. Is something in the experiment familiar to you? (yes/no).
5. If so, please point out what it reminds you of.
6. If you want to make any comment, please do so below.

The first three questions have a clear motivation, namely to see whether (possibly some) players did have a strategy to decide on their actions. Question 3 was intended to check whether players decided on their own or did look at their environment, because only in this last case imitative or conditionally cooperative strategies make any sense. Questions 4 and 5 focused on the possibility that some of the players recognized the game as a Prisoner's Dilemma because they had a prior knowledge of the basics of game theory. The final question just allowed them to enter any additional comment they would like to make. We did not carry out a more detailed questionnaire to avoid the risk of many players' leaving it blank (the whole experiment was already very long).

Part II

An evolutionary dynamics approach to tolerance.

Presentation of Part II.

In this part of the thesis, we address the second research theme: the tolerance under the framework of social dynamics. The concepts of tolerance and intolerance have been approached from many perspectives including biology, sociology and philosophy [74]. For the purposes of this thesis, we consider tolerance as "*a permissive and indulgent attitude in relation to those subjects whose characters, opinions or behaviors differ from one's own*" and, complementarily, intolerance or bigotry as "*the refusal to accept subjects with different characters, opinions or behaviors from one's own*".

Despite antecedents in classical greek (as well as from the hellenistic, and roman period) philosophers (*e.g.* Socrates, Epictetus, Marc Aurellius), toleration does not become a serious subject of philosophical and political concern in Europe until the 16th and 17th Centuries. Motivated by the Religion's Wars which followed Reformation and Counter-Reformation, thinkers as Milton, Bayle, Spinoza and Locke defended religious tolerance. Among the Enlightenment's philosophers, perhaps Voltaire was the one that most vividly expressed his views in defense of religious tolerance, and surely Kant was the most rigorous one.

At the end of the 18th Century one can see tolerant ideas embodied in practice in the USA Constitution's Bill of Rights. In the 19th Century at the formulation of political liberalism, J.S. Mill argues that the only proper limit of liberty is harm (to others), and that political power should have no authority to regulate those activities and interests of individuals that are purely private and have no secondary effects on others.

Already in the 20th Century, toleration became an important component of what is known as liberal theory. It has been defended by liberal philosophers and political theorists such as Dewey, Berlin, Popper, Dworkin and Rawls, but also criticized by Marcuse and other modern marxist thinkers who worry that toleration and its ideal of state neutrality is merely another hegemonic Western ideology, a useful "superstructural mask". After all, some politically neo-liberal practices in Europe politics are indeed quite far from being tolerant in any ample sense of the term. Nowadays, a concern for racial equality, gender neutrality, an end of prejudice, respect for cultural and ethnic difference, and a general commitment to multiculturalism has fueled ongoing debates about the nature of toleration in our age of globalization and homogeneization.

To a 21st Century social sciences theorist that might worry about these issues, a basic methodological question is wether or not social tolerance can be measured, or at least semi-quantitatively inferred from observations. What about consider it

as a parameter in an ABM investigation? The accepted meaning of the "tolerance" parameter of the Schelling model is but a possible operational use of the concept of toleration. So we see there is already a half-century old tradition of using tolerance as a social parameter in social modeling of some phenomena. Moreover, given that intolerance is the main cause of conflicts at all levels, from two-person disputes to multipart struggle and wars, and considering that, unfortunately, social rejection and self-exclusion based on real or perceived characteristics is inherent in the human condition and has always been present in every culture and time period [74, 75], the cultural diffusion enhanced by development of mass media and new technologies leads to an increasing need to address tolerance under the perspectives of social and cultural dynamics.

Tolerance and intolerance are issues that can be properly addressed through ABMs, as shown, for example, the fact that one of the first social ABM was the residential segregation model developed by Thomas C. Schelling in 1971 [52, 53]. The Schelling's model shows how a preference to have similar neighbors can lead to segregation for relatively small values of intolerance (see Introduction 1.3.1). Nevertheless, ABM have not paid much attention to the study of tolerance, in contrast to other related topics such as homophily [59], opinion formation [153] or rumor spreading [155].

In the Axelrod model (see Introduction 1.3.1 for a basic presentation) the *social influence* on the "cultural" individual characteristics defining the cultural state, appears itself as *homophile satisfaction*, the driving force of cultural change (or cultural evolution). One might say that the "incentive" to modify a cultural trait (*e.g.* conversion to a new, non-inborn faith, changing of musical taste, or going into bio-organic food consumption) is the aspiration to a greater share of cultural features with nearby agents. Simple and of general appealing, homophile satisfaction is the "benefit function" in an economic framework formulation of the model. Economic language is well-fitted to permeate Political Sciences modeling, and human social behavior, though not always, can sometimes be understood in terms of economy of "moral feelings" and/or other categories.

Importantly, the Axelrod's dynamics assumption, namely "the more similar two cultural agents are, the more similar they'll likely become in the future", seems also be rooted on a sensible theorist's inference from social experiences and observations. Note that this assumption leads naturally, for a dimer of cultural automata, to a self-sustained increase of cultural similarity. Though this may suggest that the dynamics is just a trivial accelerated tendency to cultural consensus, when passing from the cultural dimer analysis to the neighborhood's state analysis of a focal agent, one realizes that the (ensemble averaged) mean similarity of the focal isn't forcefully an increasing function of time dynamics: it could decrease in time, so dynamics is not that

trivial. Despite that, the tension between the two extreme macroscopic states, globalization and multiculturalism, is regulated by the (initial) degree q of cultural diversity, as in the dimer analysis: Provided q is so large as to render negligible the probability of sharing some cultural feature, no option other that multiculturalism prevails for the dimer, as well as for the whole macroscopic population, where frozen multicultural patterns dominate the asymptotic states for large initial cultural diversity.

Another basic feature of this modeling framework of cultural dynamics is its highly *non-biased* setting respect to trait values: There is no advantage of particular traits regarding "socio-cultural ineffectiveness" power. A physicists' term for this basic feature is *trait symmetry*. The complete invariance under interchangeability of traits imposes that the homogeneously cultural macroscopic state contains q^F microscopic states, that are equally likely provided traits are uniformly distributed in the initial conditions. Along any particular stochastic trajectory, the macroscopic consensus reaches fixation (thus irreversibly breaking trait's symmetry) through an unbiased random walk, as it occurs in Evolutionary Genetics where neutral characters are fixed in some populations. The same occurs regarding traits frequencies in multicultural macroscopic states.

Our daily experience would easily sanction as too simplistic both (homophily and exact symmetry) basic features of this cultural dynamics: to put it crudely, during the early eighties of last century, when I became more a reggae than heavy metal rocker tasted young Spaniard, is a different socio-economic situation from the coetaneous conversion of a young Mexican "catholic"-born to the "Jehovah's witnesses" faith, that are treated as indistinguishable processes inside this framework. The model, however, is not aimed to address individuals' cultural issues. Also, for a field anthropologist interested, say, in the cultural decline of Patagonia populations, such a bareness of details in the description of cultural dynamics could seem useless, likely a mere kidding exercise. However, no particular cases of important cultural processes (as culture extinction) motivate this model.

Our kind of theorists' social modeling is (no more and no less than) an educated (scientific) attempt to gain insight into basic fundamental mechanisms that operate, in some sense universally, in the emergence of collective social behaviors. We know that the Ising model is invaluably useful, far away its strict applicability to *e.g.* magnetic materials experiments. This important message is well-understood inside our physicists' culture, and, hopefully, will be increasingly so across other scientific cultures. It is amply heard, at least in some branches of social and economical sciences research, in connection to the Schelling model, where residential segregation based on ethnic (racial, unchangeable agent features) differentiation of individuals, is modeled through a very simple homophile satisfaction driven mobility dynamics (see Introduction for a basic presentation).

Although intolerance and homophily are close ideas, the relation between both concepts is not trivial: while homophily refers to the preference to not interact with dissimilar people, this preference does not implies social rejection. Nevertheless, the Axelrod and the Schelling model share the same *social force*, the homophile satisfaction. In the first one it fuels the cultural change (under conditions given by a cultural diversity parameter), while in the second it determines moving decisions (conditioned by a tolerance parameter) in a "geography" of residential neighborhoods. Each model addresses a different specific social issue, and inside own domain, each one is a basic archetype model ultimately based on homophile satisfaction as a social force.

Our goal in this part of the thesis is to incorporate intolerance into the cultural dynamics through different mechanisms. Starting from the Axelrod's model, in chapter 7 we introduce intolerance allowing agents to move from a culturally dissimilar environment to other available places according to a intolerance threshold T. This is possible thanks to the introduction of a density of empty sites h in the lattice of the original model. We show that, when the density h of empty sites is low enough and the agents percolates the lattice, mobility enhances the convergence to monocultural state. Moreover, the transition value q_c depends linearly with the system size. On the other hand, for large enough values of h, when $1 - h$ is below the site percolation threshold, a new multicultural fragmented phase appears at low values of the initial cultural diversity q; however, the monocultural phase of the original Axelrod's model is recovered for intermediate values of q, triggered by mobility, as well as the disordered (multicultural) phase for large values of q. In chapter 8, we extend the previously described model by considering intolerance T as an individual cultural feature susceptible of imitation through the cultural dynamics. This asymmetry introduced in the traits of Axelrod dynamics allows to study the preference of tolerant traits to be present or not in dominant cultures. We consider two options in order to introduce individual intolerance: *social rejection* (*i.e.*, the agents move according to their neighbors' intolerance) and *self-exclusion* (agents move incited by its own intolerance to their neighbors). In both cases we show that tolerant traits are more likely to be present in dominant cultures. Moreover, the advantage of tolerance increases with the density h of empty sites, being higher in the *social rejection* scheme. In order to obtain a more realistic model, given that social networks are heterogeneous, in Chapter 9 we introduce tolerance into cultural dynamics through *network plasticity*, allowing agents to remove links to its dissimilar neighbors and reconnecting them to other individuals chosen at random. This method allows to consider heterogeneous and dynamic networks, with a network dynamics driven by the cultural dynamics. Starting from the dynamics designed by Vazquez *et al.* [166], we introduced tolerance through a parameter Z that modulates the intensity of rewiring mechanism. We show that rewiring mechanism induces opposite effects. As expected, for large values of tolerance Z, rewiring promotes the convergence to globalization. Neverthe-

less, for intermediate values of Z, the rewiring mechanism enhances the formation of disconnected cultural clusters for values of the initial cultural diversity q which present globalization in non-evolving networks. Further, for low values of Z, although clusters are present in transitory states, rewiring promotes cultural spreading between clusters for large enough values of q, and monocultural phase is recovered.

Chapter 7

Residential segregation and cultural dissemination: An Axelrod-Schelling model.

In the Axelrod's model of cultural dissemination, we consider mobility of cultural agents through the introduction of a density of empty sites and the possibility that agents in a dissimilar neighborhood can move to them if their mean cultural similarity with the neighborhood is below some threshold. While for low values of the density of empty sites the mobility enhances the convergence to a global culture, for high enough values of it the dynamics can lead to the coexistence of disconnected domains of different cultures. In this regime, the increase of initial cultural diversity paradoxically increases the convergence to a dominant culture. Further increase of diversity leads to fragmentation of the dominant culture into domains, forever changing in shape and number, as an effect of the never ending eroding activity of cultural minorities.

7.1 Introduction

The use of agent-based models (ABM) [156] in the study of social phenomena provides useful insights about the fundamental causal mechanisms at work in social systems. The large-scale (macroscopic) effects of simple forms of (microscopic) social interaction are very often surprising and generally hard to anticipate, as vividly demonstrated by one of the earliest examples of ABM, the Schelling [52, 53] model of urban segregation, that shows how residential segregation can emerge from indi-

vidual choices, even if people have fairly tolerant preferences regarding the share of like persons in a residential neighborhood.

To gain insights on the question of why cultural differences between individuals and groups persist despite tendencies to become more alike as a consequence of social interactions, Axelrod [59] proposed an ABM for the dissemination of culture, that has subsequently played a prominent role in the investigation of cultural dynamics. Questions concerning the establishment, spread and sustainability of cultures, as well as on the "pros and cons" of cultural globalization versus the preservation and coexistence of cultural diversity, are of central importance both from a fundamental and practical point of view in today's world.

The Axelrod model implements the idea that social influence is "homophilic", *i.e. the likelihood that a cultural feature will spread from an individual to another depends on how many other features they may have already in common* [59]. The resulting dynamics converges to a global monocultural macroscopic state when the initial cultural diversity is below a critical value, while above it homophilic social influence is unable to inforce cultural homogeneity, and multicultural patterns persist asymptotically. This change of macroscopic behavior has been characterized [141, 142, 143, 144] as a non-equilibrium phase transition. Subsequent studies have analyzed the effects on this transition of different lattice or network structures [145, 146], the presence of different types of noise ("cultural drift") [147, 148], as well as the consideration of external fields (influential media, or information feedback) [149, 150] and global or local non-uniform couplings [151]. Along with other models of social dynamics (as *e.g.*, models of opinion formation [153, 154], rumor spreading [155], etc) cultural dynamics are of interest in the field of non-equilibrium phase transitions in lattice models, as other stochastic spatial models motivated by population dynamics or evolutionary biology [80]. Up to now, no investigation of the effects of agent mobility on cultural transmission has been carried out, with the exception of [157], where individuals move following the gradient of a "sugar" landscape (that they consume) and interact culturally with agents in their neighborhood, i.e., mobility is not culturally driven.

In this chapter we incorporate into the Axelrod dynamics of cultural transmission the possibility that agents living in a culturally dissimilar environment can move to other available places, much in the spirit of the Schelling model of residential segregation. This requires the introduction of a density of empty sites h in the discrete space (lattice) where agents live. As anticipated by [157] the expectations are that the agents mobility should enhance the convergence to cultural globalization, in the extent that it acts as a sort of global coupling between agents. It turns out that these expectations are clearly confirmed when the density h of empty sites is low enough so that the set of occupied sites percolates the lattice: The transition value depends

linearly with the number of agents, so that in an infinite system (thermodynamical limit) only global cultural states are possible. However, for large enough values of h, new phenomena appear associated to this mixed Axelrod-Schelling social dynamics, including a new multicultural fragmented phase at very low values of the initial cultural diversity, a (seemingly first order) transition to cultural globalization that is triggered by mobility, and the fragmentation of the dominant culture into separated domains that change continuously as the result of erosive processes caused by the mobility of cultural minorities.

7.2 The model

In the Axelrod model of cultural dissemination, a culture is modelled as a vector of F integer variables $\{\sigma_f\}$ ($f = 1, ..., F$), called cultural *features*, that can assume q values, $\sigma_f = 0, 1, ...q-1$, the possible *traits* allowed per feature. At each elementary dynamical step, the culture $\{\sigma_f(i)\}$ of an individual i randomly chosen is allowed to change (social influence) by imitation of an uncommon feature's trait of a randomly chosen neighbor j, with a probability proportional to the cultural overlap ω_{ij} between both agents, defined as the proportion of shared cultural features,

$$\omega_{ij} = \frac{1}{F}\sum_{f=1}^{F}\delta_{\sigma_f(i),\sigma_f(j)}, \tag{7.1}$$

where $\delta_{x,y}$ stands for the Kronecker's delta which is 1 if $x = y$ and 0 otherwise. Note that in the Axelrod dynamics the mean cultural overlap $\bar{\omega}_i$ of an agent i with its k_i neighbors, defined as

$$\bar{\omega}_i = \frac{1}{k_i}\sum_{j=1}^{k_i}\omega_{ij} \quad , \tag{7.2}$$

not always increases after an interaction takes place with a neighboring agent: indeed, it will decrease if the feature whose trait has been changed was previously shared with at least two other neighbors.

To incorporate the mobility of cultural agents into the Axelrod model, two new parameters are introduced, say the density of empty sites h, and a threshold T ($0 \leq T \leq 1$), that can be called *intolerance*. After each elementary step of the Axelrod dynamics, we perform the following action: If imitation has not occurred and $\omega_{ij} \neq 1$, we compute the mean overlap (7.2) and if $\bar{\omega}_i < T$, then the agent i moves to an empty site that is randomly chosen. Finally, in the event that the agent i randomly chosen is isolated (only empty sites in its neighborhood), then it moves directly to an

empty site. Note, additionally, that in the presence of a density of empty sites, the sum in equation (7.2) runs over neighboring agents, and not on neighboring sites, so that k_i can take on the values 0, 1, $\ldots 4$ for a square lattice geometry.

We define the mobility m_i of an agent i as the probability that it moves in one elementary dynamical step (provided it has been chosen):

$$m_i = (1 - \bar{\omega}_i)\ \Theta(T - \bar{\omega}_i) \ \ , \tag{7.3}$$

where $\Theta(x)$ is the Heaviside step function, that takes the value 1 if $x > 0$, and 0 if $x \leq 0$. For an isolated agent, that moves with certainty, one may convene that its mean cultural overlap is zero, so that expression (7.3) applies as well. The average mobility m of a configuration is the average of the mobility of the agents:

$$m = \frac{1}{N} \sum_{i=1}^{N} m_i \ \ , \tag{7.4}$$

where N is the total number of cultural agents. We will consider below two-dimensional square lattices of linear size L, so that $N = (1-h)L^2$, periodic boundary conditions, and von Neumann neighborhoods, so that the number k_i of neighbors of an agent i is $0 \leq k_i \leq 4$. We fix the number of cultural features to $F = 10$, and vary the parameters q, h and T, as well as the linear size L of the lattice. As it happens also for the genuine (*i.e.*, without mobility) Axelrod model, no qualitative differences appear for different values of $F \geq 3$, the only difference being that larger values of F make it easier the convergence to cultural globalization. One can easily realize that the probability that the overlap between two randomly chosen cultures, say i and j, is positive $P(\omega_{ij} > 0) = 1 - ((q-1)/q)^F$, is an increasing function of the parameter F.

Note that for $F = 1$, no matter how large $q \geq 2$ is, the overlap ω_{ij} is either 0 or 1 so that there is no chance for cultural interaction (imitation). In this limit case each agent keeps forever its own initial culture, and the size of each culture is fixed by the initial conditions (no cultural evolution). In this case the model effectively reduces to a version (one among the many possible variants) of the Schelling model of urban segregation. Specifically, it becomes a Schelling model with myopic long range move. Some recent papers in the physics literature on the Schelling model are [158, 159, 160, 161]. See also [162] for some critical comments on the physical perspective of the Schelling model.

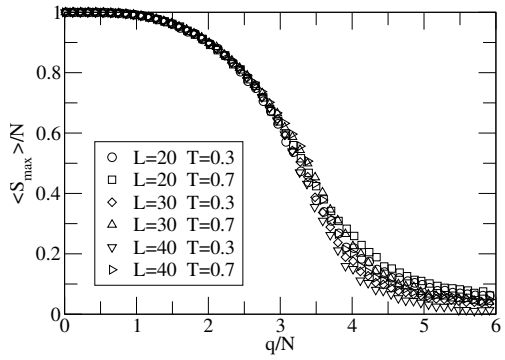


Figure 7.1: Order parameter $\langle S_{max}\rangle/N$ versus scaled initial cultural diversity q/N for a very small density of empty sites $h = 0.05$ and different values of the intolerance $T = 0.3,\ 0.7$, and of the lattice linear size $L = 20,\ 30,\ 40$, as indicated in the inset.

7.3 Results and Discussion

For the initial conditions for the cultural dynamics, N cultural agents are randomly distributed in the $L \times L$ sites of the square lattice, and randomly assigned a culture. The simulation is stopped when the number n_a of active links (*i.e.*, links such that $0 < \omega_{ij} < 1$) vanishes. The results shown below are obtained by averaging over a large number (typically $5 \cdot 10^2\ -\ 10^4$) of different initial conditions.

The usual order parameter for the Axelrod model is $\langle S_{\max}\rangle/N$, where $\langle S_{\max}\rangle$ is the average number of agents of the dominant (most abundant) culture. Large values (close to unity) of the order parameter are the signature of cultural globalization. In Fig. 7.1, we plot the order parameter versus the initial cultural diversity scaled to the population size, q/N, for a small value of the density of empty sites $h = 0.05$, and different values of the intolerance T and of the linear size L. We observe the collapse in a single curve of the graphs corresponding to different lattice sizes and, moreover, that the results are rather insensitive to the intolerance values. Figure 7.2 represents the cultural distribution in both states: ordered phase for low values of q/N, and disordered phase for high values of q/N.

For a fixed value of the initial cultural diversity q, the larger the size N of the population is, the more likely an agent can share a cultural feature with someone else in the population. Hence, as mobility allows contacts with virtually anybody, the increase of the population size enhances the tendency towards cultural globalization, and the monocultural (ordered) phase extends up to higher values of the parameter q. The critical value q_c of the transition between consensus and a disordered multicultural phase diverges with the system size $q_c \sim N$, so that in the thermodynamical limit only global cultural states are possible for a small density h of empty sites.

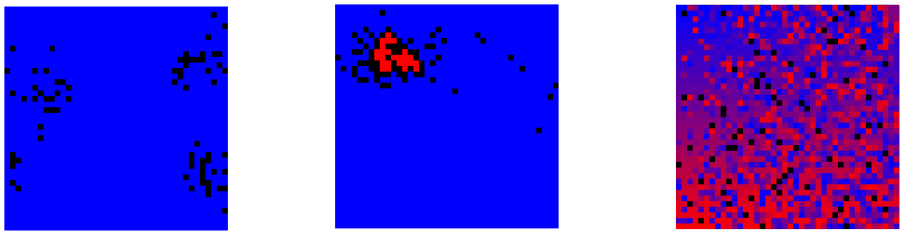


Figure 7.2: Cultural distribution when the density of empty sites is below their percolation threshold: The system is characterized by two phases as in the original Axelrod Model: A monocultural phase for low values of initial cultural diversity q/N (left chart, q/N=0.1) and a multicultural phase for high values of q/N (right chart, q/N=5). The center chart (q/N=1) represents an anomalous state that is present only in some realizations. Each color represents the cultural group that owns the node. Empty cells are represented in black. Here has been taken L=30, h=0.05 and T=0.8.

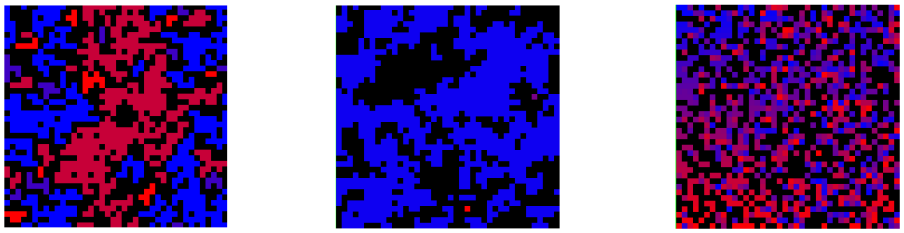


Figure 7.3: Cultural distribution for a empty sites density above their percolation threshold. A new multicultural fragmented phase appears for very low values of the initial cultural diversity (left chart, q/N=0.1), in adition to the two phases of original Axelrod Model (the ordered phase of center chart for q/N=1 and the disordered phase of right chart for q/N=5). Here has been taken L=30, h=0.45 and T=0.8. Colors represent the cultural groups and black sites the empty cells.

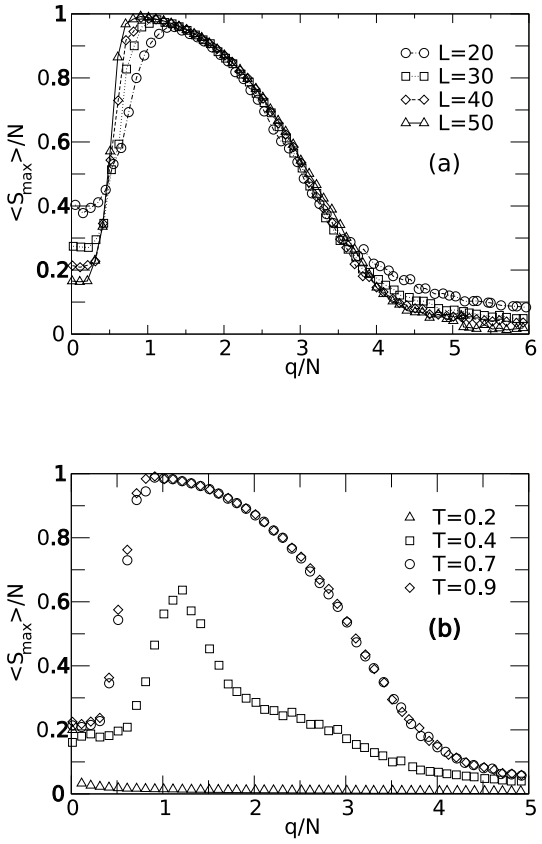


Figure 7.4: Order parameter $\langle S_{max} \rangle / N$ versus scaled initial cultural diversity q/N for an intermediate value of the density of empty sites $h = 0.5$. Panel (a) corresponds to a high value of the intolerance $T = 0.7$, and different lattice linear sizes $L = 20,\ 30,\ 40,\ 50$, while in panel (b) $L = 40$, and different values of the intolerance $T = 0.2,\ 0.4,\ 0.7,\ 0.9$ are used. See the text for further details.

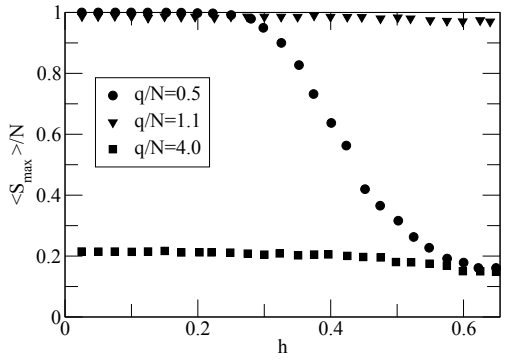


Figure 7.5: Order parameter $\langle S_{max}\rangle/N$ versus density h of empty sites, for three different values of the scaled initial cultural diversity $q/N = 0.5,\ 1.1,\ 4.0$, $T = 0.7$, and linear lattice size $L = 30$.

We will focuss hereafter on larger values of the density h of empty sites, a regime where the cultural dynamics shows strikingly different features. At very low values of the initial cultural diversity q (so that cultural convergence is strongly favored), the asymptotic states are characterized by low values of the order parameter $\langle S_{\max}\rangle/N$. The reason for the absence of cultural globalization in this regime is the existence of disconnected monocultural domains, a fact that requires values of the density $1-h$ of cultural agents at least close to (or below) the site percolation threshold value for the square lattice (0.593). In Fig. 7.5 we plot the order parameter versus the density h of empty sites, for three different values of q/N, intolerance $T = 0.7$, and linear lattice size $L = 30$. For the largest value of $q/N = 4.0$ corresponding to the culturally disordered regime, the order parameter is rather insensitive to the h values. This is also the case for $q/N = 1.1$, a value representative of the cultural globalization regime. However, for the lowest value of $q/N = 0.5$, we observe the decrease of the order parameter when $1-h$ takes on values close to the site percolation threshold, signaling the appearance of the fragmented multicultural regime. This new kind of macroscopic multicultural state is thus of a very different nature from the "genuine" multicultural phase of the original Axelrod model ($h = 0$). Though cultural convergence is locally achieved inside each geometrical cluster, the absence of contacts between clusters makes impossible the existence of globalization. The values of the order parameter in this *fragmented* phase, represented in Fig. 7.4a as a function of q/N with $h = 0.5$ and $T = 0.7$ and for several values of L, decrease with increasing lattice size, and the expectation is that the order parameter vanishes in the thermodynamical limit, because the largest cluster size below percolation should be independent of the lattice size. Left chart of figure 7.3 shows the cultural distribution of this new multicultural fragmented phase, next to the phases of original Axelrod Model.

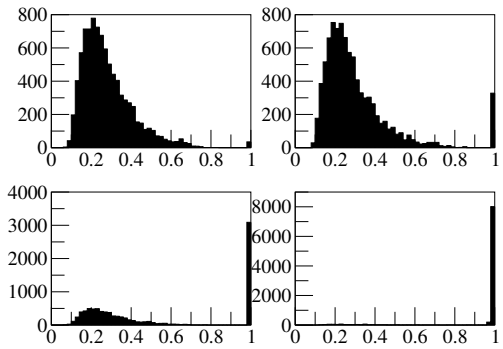


Figure 7.6: Histograms of the values of S_{max}/N, nearby the transition from the fragmented multicultural phase to globalization, for $9. \times 10^3$ realizations at (from left to right and top to bottom) $q = 100,\ 150,\ 250,\ 400$, for $L = 30$, $h = 0.5$, and $T = 0.7$. The histograms display the characteristic behavior of a discontinuous (first order) phase transition.

The increase in q from the very small values that correspond to the fragmented multicultural phase has the seemingly paradoxical effect of increasing the order parameter $\langle S_{\max}\rangle/N$ values, *i.e.*, the increase of the initial cultural disorder promotes cultural globalization. To understand this peculiar behavior, one must consider the effect of the increase of q in the initial mobility of the agents. One expects that the higher the value of q is, the lower the initial values of the cultural overlap ω_{ij} among agents are, and then the higher the initial mobility of agents should be. Under conditions of high mobility, the processes of local cultural convergence are slower than the typical time scales for mobility, so that the agents can easily move before full local consensus can be achieved, propagating their common features, and enhancing the social influence among different clusters. In other words, the attainment of different local consensus in disconnected domains is much less likely to occur, and one should expect the coarsening of a dominant culture domain that reaches a higher size.

A straightforward prediction of this argument is that one should observe higher values of $\langle S_{\max}\rangle/N$ for higher values of the intolerance T, because agents mobility is an increasing function of this parameter (see eq. (7.3)). The numerical results shown in Fig. 7.4b for different values of T and $h = 0.5$ nicely confirm this prediction, in support of the consistency of the previous argument. Interestingly, for very low values of T when mobility is not enhanced, multiculturalism prevails for the whole range of q values. On the contrary, for high values of the intolerance T, an almost full degree of cultural globalization is reached, as indicated by the values $\langle S_{\max}\rangle/N \simeq 1$ of the order parameter. In those final states almost all agents belong to a single connected monocultural cluster. One should also note that, for fixed values of the intolerance T and the density h of empty sites, the previous argument indicates that the relevant variable for this transition is the initial cultural diversity q, and not q/N,

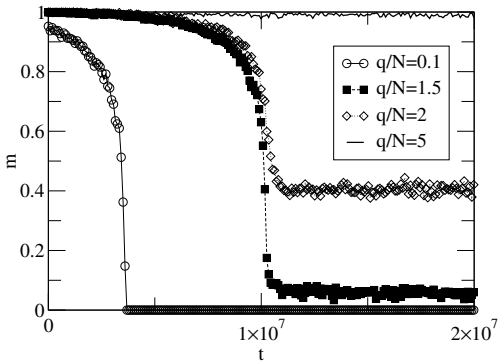


Figure 7.7: Average mobility m versus time t for $h = 0.5$, $L = 30$, $T = 0.7$ and different values of the scaled initial culture diversity q/N as indicated. Unlike the other figures, in this case each curve represents the results of a single realization. See the text for further details.

so that the interval of values of q/N that corresponds to the multicultural fragmented phase shrinks for increasing N values.

To characterize the passage from the multicultural fragmented phase to global consensus with increasing initial cultural diversity, we have computed the histograms of the values of $S_{\max}/N$ at values of q where the order parameter increases, see Fig. 7.6. The histograms display the bimodal characteristics of a first-order transition. In a fraction of realizations, the transient mobility is able to spread social influence among the clusters so that global consensus is finally reached. This fraction increases with q, to the expense of the fraction of realizations where fragmented multiculturality is reached. Note that no significant change of shape and position of the corresponding part of the histogram is noticeable, apart from its progressive reduction to lower volumes, when q increases.

Further increase of the initial cultural diversity q enhances the likelihood of agents sharing no cultural feature with anybody else in the finite population. The presence of these culturally "alien" agents decreases the value of the order parameter and the increase of their number with q is concomitant with the transition to multiculturality in the original Axelrod model (as well as here, for finite populations). We see in Fig. 7.4b that the increase of the intolerance parameter T shifts this transition to higher values of q/N, in agreement with the enhancement of the convergence to globalization that T produces via mobility, as discussed above. Each alien agent has, at all times, a mobility $m_i = 1$, and the average mobility cannot decrease in time to zero value when they appear. In other words, the asymptotic states of the cultural dynamics are no longer characterized by $m = 0$. The time evolution of the average mobility m for particular realizations at $h = 0.5$, $T = 0.7$, $L = 30$ and different values of q/N is shown in Fig 7.7. The value of q/N beyond which the stationary

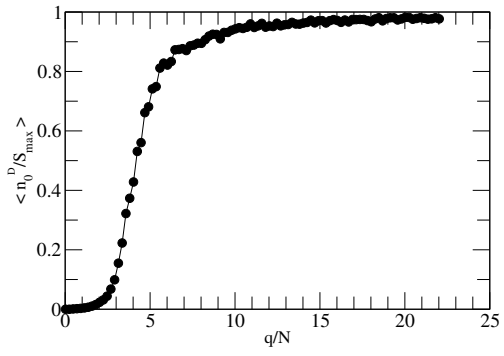


Figure 7.8: Cultural minorities continuously erode the dominant culture domain, that breaks into separate domains and isolated individuals. As a quantitative measure of this erosion phenomenon we plot here the stationary value of the averaged fraction n_0^D/S_{max} of isolated individuals of the dominant culture versus q/N, for $h = 0.5$, $T = 0.7$, and $L = 30$. The inset shows an illustrative configuration where erosion can take place.

average mobility is larger than zero signals the appearance of these alien cultural agents.

In addition, the restless character of the alien agents has an important effect on the geometry of the dominant culture, namely its *erosion*. As an illustrative example, let us consider the situation represented in the inset of Fig. 7.8, in which an agent i of the dominant culture is placed at the frontier of a cluster, having a single neighbor of his kind, and assume that an alien agent j has moved recently to one of the empty neighboring sites of i. When agent i is chosen for an elementary dynamical step, there is a probability $1/2$ of choosing agent j for an imitation trial. As $\omega_{ij} = 0$, and then $\bar{\omega}_i = 1/2$, the agent i will move from there to a randomly chosen empty site whenever the intolerance parameter is $T > 1/2$. We see that, for this particular situation, the erosion of the dominant culture cluster will occur with probability one half.

Note that the erosion of the dominant culture cluster does not change the size $S_{\max}$ of the dominant culture. It simply breaks it up into separate domains, some of them consisting of single (isolated) individuals. These isolated members of the dominant culture will eventually adhere to domains, to be at a later time again exposed to erosion, and so on. Therefore the shape and number of domains of the dominant culture (as well as that of the other ones), fluctuate forever. The number n_0^D of isolated dominant culture agents reaches a stationary value that results from the balance between erosive and adhesive processes. To quantify the strength of the eroding activity of cultural minorities we show in Fig. 7.8 the stationary value of the averaged fraction $\langle \frac{n_0^D}{S_{max}} \rangle$ of isolated individuals of the dominant culture versus the scaled ini-

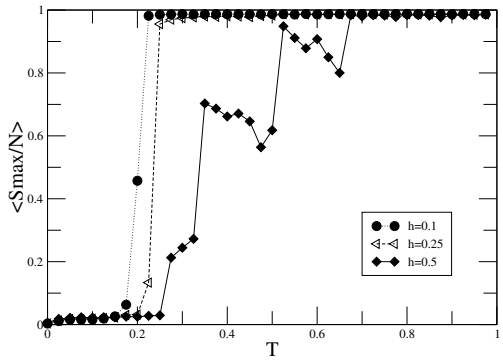


Figure 7.9: Order parameter S_{max}/N versus intolerance T, for differents values of h, q/N=1.1 and L=30. The shape corresponds to discrete and continuous aspects of the algorithm. Results have been averaged for 10^3 different initial conditions. See the text for further details.

tial cultural diversity, for $h = 0.5$, $T = 0.7$, and $L = 30$. Soon after the transition from the fragmented multicultural phase to globalization occurs, erosion increases dramatically, largely contributing to the large values of the stationary mobility m that characterize the multicultural states in the model here introduced.

Figure 7.9 shows the order parameter S_{max}/N versus intolerance T, for a scaled diversity initial $q/N = 1.1$ and different densities of empty sites h. This value of q/N is high enough to avoid cultural globalization in the Axelrod limit, but low enough to allow the overlap. There exists a threshold of intolerance T_c, below which the order parameter is $S_{max}/N \approx 0$. If $T < T_c$ the mobility is low and can not promote cultural convergence. Although empty sites enhance mobility, paradoxically it is found that T_c increases with h. The explanation for this phenomenon is that the influence of h on the mobility is small, but the increase of h implies a decrease of N, therefore also a decrease of q if q/N had been fixed. This in turn implies a increase of mean overlap ω_i, a decrease of mobility m_i and finally the incerase of T_c as observed. For low values of h, if $T > T_c$ the order parameter is found $S_{max}/N \approx 1$, that is, above a critical value of mobility the system reaches monocultural state. By contrast, when h is high enough to allow the formation of site clusters, it can be seen a discontinuous behavior: Now a node can have $m = 0, 1, 2, 3\ or\ 4$ neighbors, and $S_{max}(T)$ presents steps for $T = n/m$.

7.4 Theorical analysis

Let i be a node of dominant culture D, and κ its number of links to his cultural domain. If i has an allien neighbor of culture O, every time step the probability P_{su} for i to stay unchanged, ie, to preserve his features and rest in D-group is:

$$P_{su} = \frac{k + (1 - \omega_{DO})H(\frac{k+\omega_{DO}}{k+1} - T)}{k+1} \ , \tag{7.5}$$

where H represents the Heaviside function. The probability P_{sc} for i to change, that is, to leave the group D is:

$$P_{sc} = \frac{\omega_{DO}}{k+1} \ , \tag{7.6}$$

and the probability P_{su} for i to move is:

$$P_m = \frac{1 - \omega_{DO}}{k+1} H(T - \frac{k - \omega_{DO}}{k+1}) \ , \tag{7.7}$$

From now on, we deal with stationary state, and therefore we take $\omega_{DO} = 0$.

In order to calculate an estimate of n_0^D, we consider:
n^u: number of D-agents exposed to real erosion, ie, having an allien-neighbour they belong to a D-cluster with at least three elements. n^u is the sum of nodes like that, with κ links to D-agents:
$n^u = \sum_{\kappa=1}^{3} n^u(\kappa)$.
n_2^u: D-agents exposed to erosion that belong to a 2-agents domain.
n^E: number of empty sites adjacent to a D-domain.
H_0^D: number of empty sites adjacent to isolated D-agents.

We take into account the useful time step, ie, we only consider a step when a event occurs. That only implies a time translation and does not affect fixed points. We study the most likely events of erosion and adhesion:

- a_1: Simple adhesion event: $n_0^D \to n_0^D - 1$, $\quad \Delta N^u \geq 0$,
- e_1: Simple erosion event: $n_0^D \to n_0^D + 1$, $\quad \Delta N^u = -1$,
- a_2: Double adhesion event: $n_0^D \to n_0^D - 2$,
- e_2: Double erosion event: $n_0^D \to n_0^D + 2$,

and, their respectives probabilities are:

$$\begin{aligned}
P(a_1) &= \frac{n_0^D n^E}{hL^2(n^0+n_0^D+n^u)} \ , \\
P(e_1) &= \Big(1-\frac{n^E}{hL^2}\Big)\frac{1}{n^0+n_0^D+n^u}\sum_{\kappa=1}^{3} n^U(\kappa)\frac{1}{\kappa+1}H\big(T-\frac{\kappa}{\kappa+1}\big) \ , \\
P(a_2) &\simeq \frac{n_0^D}{n^0+n_0^D+n^u}\frac{H_0^D}{hL^2} \ , \\
P(e_1) &= \Big(1-\frac{n^E}{hL^2}\Big)\frac{n_2^U\frac{1}{2}H(T-\frac{1}{2})}{n^0+n_0^D+n^u} \ .
\end{aligned} \tag{7.8}$$

A necessary condition for equilibrium is:

$$P(a_1)+2P(a_2)=P(e_1)+2P(e_2) \tag{7.9}$$

Therefore, under this assumptions (we are not taking into account all possible events), we can replace 7.8 in 7.9 to estimate n_0^D:

$$\begin{aligned}
T<1/2: \ n_0^D &= 0 \ , \\
1/2\le T<2/3: \ n_0^D &= \frac{(hL^2-n^e-H_0^D)(n^u(1)+n_2^u)}{2(n^E+H_0^D)} \ , \\
2/3\le T<3/4: \ n_0^D &= \frac{(hL^2-n^e-H_0^D)(\frac{1}{2}n^u(1)+\frac{1}{3}n^u(2)+\frac{1}{2}n_2^u}{n^E+H_0^D} \ , \\
T\ge 3/4: \ n_0^D &= \frac{(hL^2-n^e-H_0^D)(\frac{1}{2}n^u(1)+\frac{1}{3}n^u(2)+\frac{1}{4}n^u(3)+\frac{1}{2}n_2^u}{n^E+H_0^D}
\end{aligned} \tag{7.10}$$

Figure 7.10 shows the comparison between theorical estimate for the isolated dominant culture agents n_0^D and experimental results. As one can see, although theory underestimates the value of isolated agents, adjustment is quite good. The explanation for this deviation is that we have ignored some kinds of events. However, n_0^D estimate is rooted not only in system parameters, but also in other observables related to the spatial distribution (perimeter, location of aliens), so that rather than a prediction is a check of the analytical method.

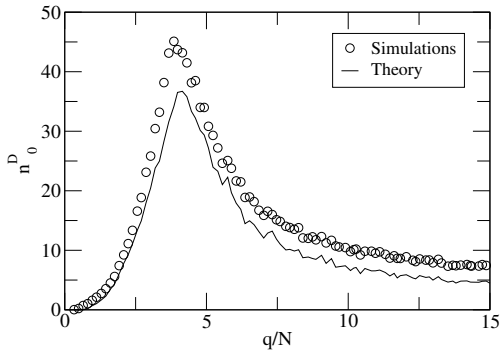


Figure 7.10: Isolated dominant culture agents n_0^D versus the scaled initial cultural diversity q/N, for T=0.8, h=0.3 and L=30. Circles represent simulation results for 10^4 realizations, and lines theorical estimation.

7.5 Conclusions

We have introduced a model of cultural dynamics in which agents can move driven by cultural dissimilarities with their environments, at the style of the Schelling model of urban segregation. The introduction of agents mobility through this segregation mechanism into the Axelrod cultural dynamics leads to an enhancement of the convergence to cultural globalization for small densities of empty sites, so that the behavior of the order parameter (i.e., the relative size of the dominant culture) scales with the number N of cultural agents. That is, the transition to multiculturalism only occurs for finite populations.

Furthermore, for larger densities of empty sites, when cultural agents cannot percolate the lattice, a new type of multicultural fragmented phase appears at low values of the initial cultural diversity q. Though the initial cultural overlap is enough to trigger the local cultural convergence inside each geometrical cluster of agents, cultural globalization is no longer possible due to the lack of cultural transmission between monocultural isolated domains. Provided the values of the intolerance T are high enough, this regime is followed by a new transition to globalization for increasing values of q, that is triggered by the increase in the initial mobility. Moreover, in the genuine Axelrod transition from global consensus to polarization, the shape and number of cultural domains are here dynamically fluctuating by the competitive balance of erosive and adhesive processes associated to the agents mobility.

Chapter 8

Selective advantage of tolerant cultural traits in the Axelrod-Schelling model.

In the previous chapter 7 we introduced the Axelrod-Schelling model, that incorporates into the original Axelrod's model of cultural dissemination the possibility that cultural agents placed in culturally dissimilar environments move to other places, the strength of this mobility being controlled by an intolerance parameter. By allowing heterogeneity in the intolerance of cultural agents, and considering it as a cultural feature, *i.e.*, susceptible of cultural transmission (thus breaking the original symmetry of Axelrod-Schelling dynamics), we address here the question of whether tolerant or intolerant traits are more likely to become dominant in the long term cultural dynamics. Our results show that tolerant traits possess a clear selective advantage in the framework of the Axelrod-Schelling model. We show that the reason for this selective advantage is the development, as time evolves, of a positive correlation between the number of neighbors that an agent has in its environment and its tolerant character.

8.1 Introduction

In the Axelrod-Schelling model introduced in chapter 7, we focused on mobility, considering the driving force of mobility is the agents' cultural dissimilarity with their environment, *i.e.*, homophile (dis)satisfaction, the same that drives cultural transmission. Starting from the Axelrod model for culture dissemination and the residential segregation model of Schelling, two new parameters where introduced, namely the

density h of empty lattice sites (places that are available to moving agents), and an intolerance parameter T that controls the strength of the mobility: If an attempt to cultural interaction (imitation) fails, then the agent i moves to a randomly chosen empty lattice site if its mean cultural similarity $\bar{\omega}_i < T$. T here is a threshold for tolerance, in such a way that high values of T characterize intolerant societies.

In this chapter, we extend the Axelrod-Schelling model by considering intolerance T as a cultural feature, and then it is no longer a parameter (a property of the whole population) but an individual property of agents subjected to cultural transmission. Due to its influence on the dynamics through the rule of mobility, the question of whether or not certain traits of this feature are more likely to be present in the dominant culture makes sense, contrary to what occurs with the rest of cultural features, whose particular traits do not influence the dynamics, and are thus selectively neutral.

We have performed extensive numerical simulations that implement different rules for the mobility of agents, whose results show unambiguously that tolerant traits possess a selective advantage over intolerant ones, *i.e.*, they are better adapted for survival in the long term dynamics. Furthermore, by a stochastic analysis we present arguments showing that the reason of this cultural evolutionary success of tolerant traits is the establishment in the population of a negative correlation between the number k_i of neighboring agents, and the value T_i of the agent intolerance. This is presented in section 8.3. Before, in section 8.2, we reconsider the transition between fragmented multiculturalism and globalization, first analyzed in chapter 7, by using an alternative scheme for mobility with homogeneous intolerance. This new scheme corresponds to the homogeneous version of one of the rules of mobility used in section 8.3 (mobility by social rejection), so that this helps in the interpretation of some of these results, and at the same time, it throws a new light on the understanding of the mechanisms triggering this transition. Finally, we summarize our results in the concluding chapter 10.

8.2 The transition from fragmented multiculturalism to globalization revisited

One of the new phenomena that appear associated to the mixed Axelrod-Schelling social dynamics is the existence, for values of the density $(1-h)$ of agents below the lattice percolation threshold, of a multicultural macroscopic phase at very low values of the initial cultural diversity q. In this regime, the processes of local cultural convergence are faster that the typical time scales at which mobility is able to induce global convergence to a monocultural state. In this multicultural state agents are ag-

gregated into disconnected (monocultural) clusters where different cultural consensus have been achieved. Hence the name *fragmented* for this multicultural phase.

If the value of q is increased (see figure 1a), the behavior for the order parameter $\langle S_{max}\rangle/N$, becomes rather sensitive to the value of the intolerance parameter T: For very low values of T multiculturalism persists, while for very high values, a first order transition to complete globalization is observed. At intermediate values of T, the order parameter increases versus q but complete globalization is not reached. The observation that the increase of the initial cultural diversity promotes cultural globalization may seem paradoxical at a first sight, but it is not difficult to rationalize it by noting that an increase in q has also the effect of enhancing mobility, which is in turn an important driving force towards globalization. Moreover, insofar as higher values of T enhance agents' mobility, the different behaviors that are observed for different values of the intolerance are consistent with this interpretation.

To deepen further our current understanding of the complex competing effects of different parameter variations that lead to the transition *fragmented multiculturalism-globalization*, we study here this transition in a different scheme for the mobility of cultural agents. We remind here that in the original scheme of chapter 7, after an elementary step of the Axelrod dynamics, if imitation has not occurred and $\omega_{ij} \neq 1$, the agent i moves to a randomly chosen empty site whenever $\bar{\omega}_i < T$. If the agent i turns out to be isolated, then it moves with certainty. We refer hereafter to this scheme as A. The mobility m_i of an agent i is defined as the probability that it moves in one elementary dynamical step (provided it has been chosen). Thus in the scheme A:

$$m_i^A = (1 - \bar{\omega}_i)\ \Theta(T - \bar{\omega}_i) \ \ , \tag{8.1}$$

where $\Theta(x)$ is the Heaviside step function, that takes the value 1 if $x > 0$, and 0 if $x \leq 0$.

In the new scheme, hereafter referred to as B, after an elementary step of the Axelrod dynamics, if imitation has not occurred and $\omega_{ij} \neq 1$, the agent i moves to a randomly chosen empty site with probability $(1 - \bar{\omega}_i)\ T$. In the case that agent i is isolated, then it moves with certainty, as in the previous scheme. The mobility of agent i in the scheme B is thus given by

$$m_i^B = (1 - \bar{\omega}_i)^2\ T \ \ . \tag{8.2}$$

As shown in the figure 8.2, in both schemes the mobility is a decreasing function of $\bar{\omega}$. However in the scheme A the mobility vanishes in the interval $\bar{\omega} > T$ (being independent on T for $\bar{\omega} < T$), while it does not vanish in the scheme B, provided $\bar{\omega} \neq 1$ (and $T > 0$), though it takes lower values than in the scheme A for $\bar{\omega} < T$ where it depends linearly on T.

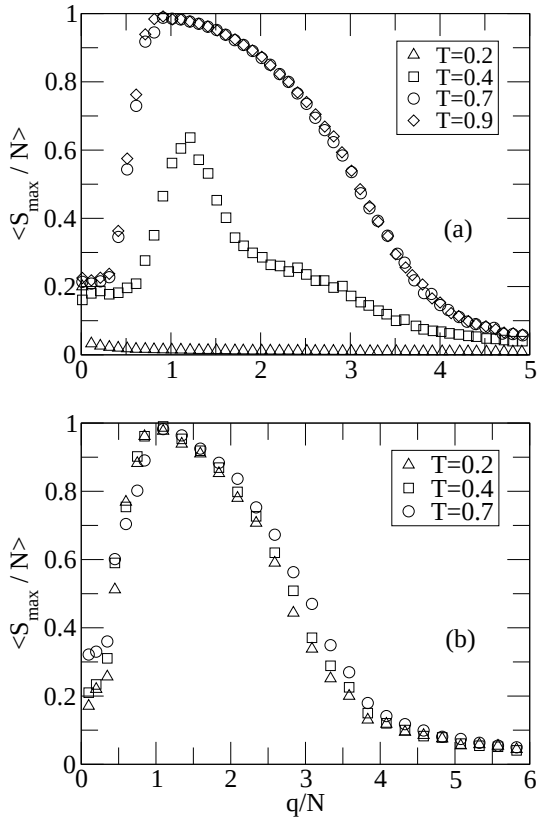


Figure 8.1: Order parameter $\langle S_{max} \rangle / N$ versus scaled initial cultural diversity q/N for a density of empty sites $h = 0.5$ and lattice linear size $L = 40$. Panel (a) corresponds to scheme A for different values of the intolerance parameter. Panel (b) corresponds to scheme B. See the text for further details.

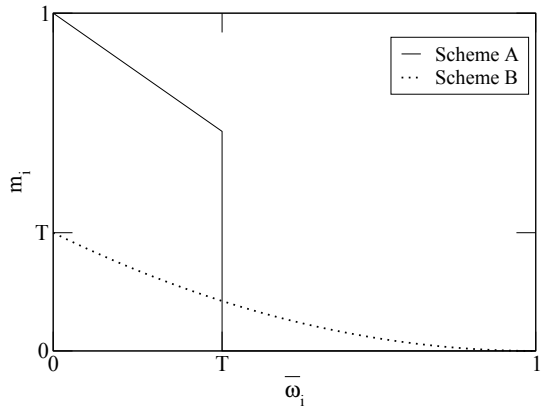


Figure 8.2: Mobility m_i for a node i as a function of his mean overlap $\bar{\omega}_i$ in schemes A and B. In the scheme A the mobility vanishes for $\bar{\omega} > T$. Otherwise, whenever $\bar{\omega} \neq 1$ and $T > 0$, it does not vanish in the scheme B.

In figure 8.1(b) we plot the order parameter versus the scaled initial cultural diversity q/N for $h = 0.5$ and different values of the intolerance T, for the scheme B and a two-dimensional square lattice geometry. In contrast with the results for the scheme A (shown in figure 8.1(a)), the behavior of the order parameter turns out to be rather insensitive to the values of the intolerance T, and the transition from the fragmented multicultural phase to globalization takes place for all the values of T that we have used. How to fit these observations into the interpretation framework given in chapter 7 (succintly reproduced above in a previous paragraph) for the transition?

To have a better picture of the speed at which the processes of cultural convergence take place and what parameters are more influential on them, we have inspected the time evolution of the histograms of $\bar{\omega}$, namely $P(\bar{\omega}, t)$, at values of the initial cultural diversity close (below and above) to the transition. In all cases and for both schemes, this probability density evolves always from being sharply concentrated near $\bar{\omega} = 0$ at $t = 0$, to become later widespread, the centroid shifting to progressively higher values of $\bar{\omega}$ as time goes by, until it concentrates near $\bar{\omega} = 1$, finally becoming a Dirac delta function $\delta(\bar{\omega} - 1)$. The time scale at which this evolution occurs seems not to be influenced by the scheme (A or B) adopted and the influence of the value of T is also minor. The important parameter that mainly determines the time scale of local cultural convergence is the initial cultural diversity q: The lower its value the faster this process takes place. Then, what makes a truly meaningful difference between, on one side, both scheme A at high T values and scheme B at all T values and, on the other side, scheme A at low T values (where the transition to globalization is absent), is that agents with high cultural overlap do not move in the latter.

These results throw a new light over the mechanisms that trigger the transition from the fragmented multicultural phase to cultural globalization. The increase of the initial cultural diversity slows down the local cultural convergence, giving then a chance to mobility to induce global cultural consensus. But it is the mobility of agents with a significant high local cultural overlap (however small its mobility could be, as it is the case for the scheme B at low T values), and not just the amount of overall mobility, what allows the effective cultural transmission among the disconnected clusters of the fragmented states so making possible the coalescence of the giant monocultural cluster characteristic of the globalization state. If mobility is strictly limited to culturally marginal agents, its power of cultural transmission is unable to overcome the fragmentation into disconnected cultural clusters.

8.3 Heterogeneous intolerance.

As we have already mentioned in the introductory section, the mobility of cultural agents in the Axelrod-Schelling model is driven by the same utility (or social driving force) that underlies the cultural dynamics of the Axelrod model (as well as the dynamics of the Schelling model), namely "homophile satisfaction". In the model, those agents that are placed inside fully homogeneous cultural environments don't move. Cultural dissimilarities are the only source of mobility, and the parameter T, that controls the strength of mobility, quantifies the degree of (in)tolerance to cultural dissimilarities. Being a model parameter, tolerance is a quantity characteristic of the whole (artificial) society. In other words, in this context one can speak of tolerant (low value of T) or intolerant societies. However, it seems to us rather natural to consider (artificial) societies where different agents have different degrees of tolerance to cultural dissimilarities. This certainly opens the possibility of new interesting questions to be investigated inside the model.

In what follows, we consider that each cultural agent i has assigned a real number $0 \leq T_i \leq 1$, called intolerance. Moreover, we are going to consider the intolerance of agents as a quantity associated to a *cultural feature*, *i.e.* a component of the cultural vector, and then subjected to temporal changes as a result of cultural interactions. Without loss of generality, one can associate the agents' intolerance to the first component σ_1 of the cultural vector $\{\sigma_f\}$. As this variable takes on integer $(0, 1, ..., q-1)$ values, one has to choose some function $f(x)$ that takes values in the interval $[0, 1]$, and define the intolerance T_i of agent i to be

$$T_i = f(\sigma_1(i)) \ . \tag{8.3}$$

Next we have to specify the particular way in which the agents' intolerances enter into the dynamical rules. Many alternatives can indeed be considered for it, and our

first choice will be the following: After an elementary step of the Axelrod dynamics, if imitation has not occurred and $\omega_{ij} \neq 1$, the agent i moves to a randomly chosen empty site with probability

$$\frac{1}{k_i}\sum_{j=1}^{k_i}(1-\omega_{ij})T_j \quad , \tag{8.4}$$

where the sum extends to the k_i neighbors of i, and if the agent i is isolated ($k_i = 0$) it moves with certainty. In this choice, the intolerance $theta_j$ of a cultural agent j is seen as its degree of hostility towards a culturally dissimilar neighbor i, and is weighted by the cultural dissimilarity $(1-\omega_{ij})$. The mobility of an agent i is here the result of the *social rejection* of its neighbors, due to cultural dissimilarities.

The Axelrod-Schelling model with homogeneous tolerance, as the original Axelrod's model does, assumes an unbiased scenario in the sense that the traits of a cultural feature are completely interchangeable: Nothing in the dynamical rules distinguishes among different traits, and then the likelihood that each particular trait is present in the dominant culture of a realization is the same for all of them, provided they are uniformly distributed in the initial conditions for the dynamics. The particular traits that survive in the dominant culture of a given realization reach fixation by neutral selection, so that averaging over many independent realizations, one obtains a uniform distribution of traits in a large enough sample of dominant cultures.

However this symmetry of the model is broken in our current case of heterogeneous intolerance regarding the cultural feature σ_1, for its particular values do influence the local dynamics through the dynamical rule of mobility. Then, the question of how likely are different traits to prevail and be present in the dominant culture makes now sense in this new symmetry-breaking scenario. Do tolerant traits possess a cultural selective advantage? or, on the contrary, are intolerant traits better adapted to survive? Moreover, by which dynamical mechanisms the "natural" selection of particular T values is built up in the time evolution of the populations of cultural agents?

Note that if one takes for $f(x)$ in equation (8.3) a constant function, so that $T_i = T$ independent of i, one recovers the scheme B introduced in the previous section 8.2. In the extent that the behavior of the order parameter $\langle S_{max}\rangle/N$ (for a density of empty sites $h = 0.5$) in scheme B was seen to be rather insensitive to the value of T, one should expect in the present case of heterogeneous intolerance, that the order parameter for a density of empty sites $h = 0.5$ will be as shown in figure 8.1(b). Thus the choice made above in equation (8.4) is technically convenient for the purpose of investigating the question on the selective advantage of tolerant traits, just because it is expected that it leads to states of cultural globalization in some ranges of the initial cultural diversity, when the very term "dominant culture" is most meaningful.

We consider two-dimensional square lattices of linear size L, with periodic boundary conditions. The number F of cultural features is fixed to $F = 10$, and we have used two values of the density of empty sites, namely $h = 0.05$, representative of the situation in which agents percolate the lattice, and $h = 0.5$ as representative of the opposite case. For $f(x)$ we will consider a simple linear function:

$$T_i = q^{-1}\sigma_1(i) \ . \tag{8.5}$$

For the initial conditions, $N = (1-h)L^2$ agents are randomly distributed on the $L \times L$ lattice sites and randomly assigned a culture. The simulation of the cultural dynamics is stopped when the number of links for which $0 < \omega_{ij} < 1$, commonly called active links, vanishes. Besides the order parameter, we compute the intolerance T_D of the dominant culture, the average intolerance $\langle T \rangle$, and sometimes, the histogram of intolerance values of the final state. The results that we show below are obtained by averaging over a large number (typically $10^3 - 10^4$) of different initial conditions.

In the two panels of Figure 8.3 we show our numerical results for $h = 0.05$ (panel a) and $h = 0.5$ (panel b). First, we confirm the expectations on the behavior of the order parameter discussed above: Given the insensitive character of the order parameter in the scheme B to the value of the intolerance parameter T for both values of h, no effect on $\langle S_{max} \rangle / N$ due to the heterogeneity of agents' intolerance is observed.

The numerical results for the intolerance values T_D of the dominant culture for both values of the density of empty sites clearly show that very tolerant traits are better adapted to survive and become a part of the dominant culture. This occurs in the whole range of values of the initial cultural diversity that leads to values of the order parameter much larger than N^{-1} (so as the term dominant possess a meaning). By comparing the graphs of T_D shown in Figs. 8.3a and 8.3b, we observe that the T_D values are significantly lower for $h = 0.5$ than for $h = 0.05$, so that the strength of the selective advantage of tolerant traits increases when the density h of empty sites is higher. The fact that the average intolerance $\langle T \rangle$ of the final configurations is higher than T_D, provided the order parameter $N^{-1} \ll \langle S_{max} \rangle / N < 1$, indicates that the non-dominant surviving values of the intolerance are typically larger than the dominant one. We further show in Fig. 8.4 that the results regarding the behavior of T_D and $\langle T \rangle$ for $L = 40$, are essentially unchanged for lattice of size 100×100.

In Fig. 8.5 we show the histogram of T_D values, obtained from 2×10^3 realizations, at fixed value of $q/N = 1.1$, for a density of empty sites $h = 0.05$. One should note that though the mean value of the dominant intolerance is at $T_D = 0.07$, the probability density is sharply peaked at $T_D = 0$ and quickly decays to negligible values as T_D increases. In other words, the lower the value of T_D, the more probable, so that the mean value is only indicative of the dispersion scale of the density.

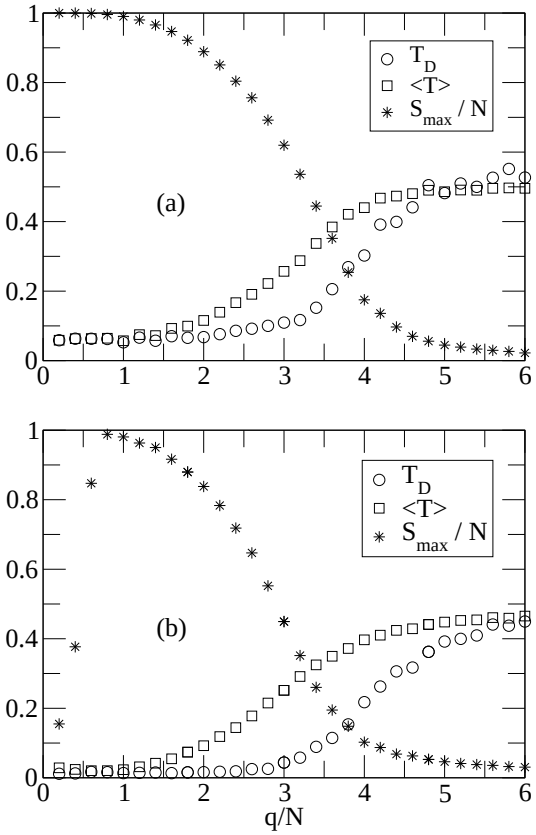


Figure 8.3: Order parameter $\langle S_{max} \rangle / N$ (stars), intolerance T_D of the dominant culture (circles), and average intolerance $\langle T \rangle$ (squares) versus scaled initial cultural diversity q/N for a lattice linear size $L = 40$, for the scheme of mobility corresponding to equation (8.4). Panel (a) corresponds to a density of empty sites $h = 0.05$. Panel (b) corresponds to $h = 0.5$. See the text for further details.

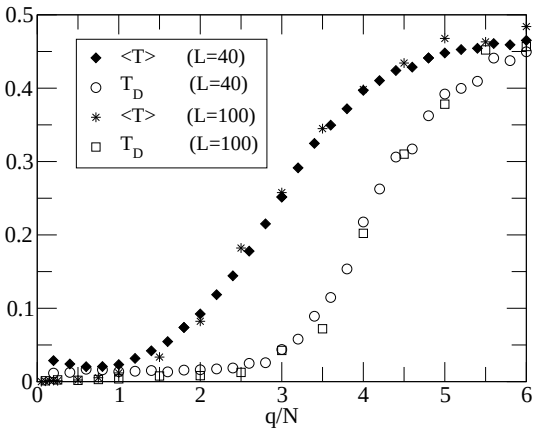


Figure 8.4: Comparison of T_D and $\langle T \rangle$ for $L = 40$ and $L = 100$, $h = 0.5$ and the same mobility rule used in Fig 8.3.

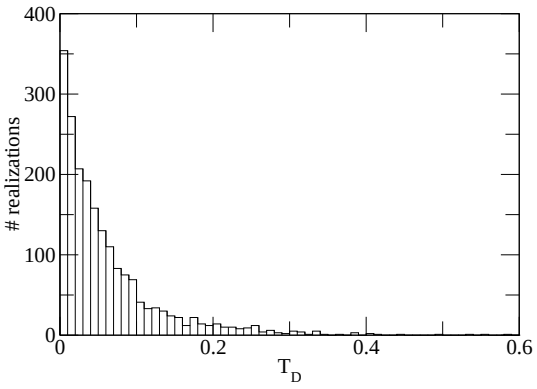


Figure 8.5: Histogram of the values of the intolerance T_D of the dominant culture for 2×10^3 realizations, at scaled initial cultural diversity $q/N = 1.1$, and a density $h = 0.05$ of empty sites. See the text for further details.

In order to explain why tolerant traits are better adapted to prevail in the long term of the dynamics, let us consider the subset $A(T,t)$ of those cultural agents i for which, at time t, $T_i \leq T$, where T is an arbitrarily chosen value of the intolerance (*e.g.*, $T = 0.3$ or more, or less). Let us denote by $n(T,t)$ the cardinal of $A(T,t)$, and call $\mathcal{L}(T,t)$ the set of lattice links (i,j), such that the agent i belongs to $A(T,t)$ and the agent j is not in this set (so $T_j > T$). If time is measured in elementary step units, the difference

$$\Delta n(T,t) = n(T,t+1) - n(T,t) \tag{8.6}$$

can only take on the values $0, \pm 1$. To compute the probability P_+ that $\Delta n(T,t)$ takes on the value $+1$, one has to sum over all links $(i,j) \in \mathcal{L}(T,t)$ the product of the following factors:

a) the probability (N^{-1}) of choosing agent j for a cultural imitation trial,

b) the probability (k_j^{-1}) that its neighbor i is chosen,

c) the probability (ω_{ij}) that agent j imitates an uncommon feature's trait of i, and

d) the probability $\left(\frac{1}{(1-\omega_{ij})\,F}\right)$ that the chosen uncommon feature is σ_1.

Note that for a link (i,j) in the set $\mathcal{L}(T,t)$, the strict inequality $\omega_{ij} < 1$ holds. Then we obtain

$$P_+ = \frac{1}{NF} \sum_{(i,j)\in\mathcal{L}(T,t)} \frac{1}{k_j} \frac{\omega_{ij}}{(1-\omega_{ij})} \quad . \tag{8.7}$$

In a similar way, the probability P_- that $\Delta n(T,t)$ takes on the value -1 is

$$P_- = \frac{1}{NF} \sum_{(i,j)\in\mathcal{L}(T,t)} \frac{1}{k_i} \frac{\omega_{ij}}{(1-\omega_{ij})} \quad . \tag{8.8}$$

We see that the number of agents in the set $A(T,t)$ performs a complicated random walk with left- and right-step probabilities changing in time as dictated by the model dynamics. The expected value of $\Delta n(T,t)$ is given by the difference $(P_+ - P_-)$, then

$$E[\Delta n(T,t)] = \frac{1}{NF} \sum_{(i,j)\in\mathcal{L}(T,t)} \frac{(k_i - k_j)}{k_i k_j} \frac{\omega_{ij}}{(1-\omega_{ij})} \quad . \tag{8.9}$$

This equation is the basis for an understanding of the selective advantage of tolerant traits. Indeed, following equation (8.4), agents with high T_i values promote the

mobility of their neighbors (leaving empty sites in their neighborhoods) more than tolerant agents do, so that one should expect that a negative correlation between values of k_i and T_i may be easily developed in the population, and tolerant agents may likely have larger values of k_i than those of intolerant agents. If this is the case, then equation (8.9) indicates that the random walk performed by $n(T,t)$ will be biased to the right, and the number of tolerant agents will likely increase as time evolves. The cultural selective advantage of tolerant traits has its origin on the bias produced by the negative correlation degree-intolerance ($k_i,\ T_i$) that is directly induced by the dynamical rule of social rejection.

The equation (8.9) allows also to rationalize the observation that the selective advantage of tolerant traits is strengthened by higher values of the density h of empty sites, because higher h values easily allow for higher values of the degree differences ($k_i - k_j$) for $(i,j) \in \mathcal{L}(T,t)$, and so the bias favoring the increase of $n(T,t)$ can be stronger.

We have also considered a second way in which agents' intolerance enter into the mobility rule of the dynamics: After an elementary step of the Axelrod dynamics, if imitation has not occurred and $\omega_{ij} \neq 1$, the agent i moves to a randomly chosen empty site provided

$$\bar{\omega}_i < T_i \ . \tag{8.10}$$

Note that if one takes for $f(x)$ in equation (8.3) a constant function, so that $T_i = T$ independent of i, one recovers the scheme A for homogeneous intolerance, that was used in chapter 7: Intolerance value is a threshold for the cultural overlap. But there is also here an important difference with respect to equation (8.4) regarding the interpretation, or meaning, of the intolerance. In (8.10) what determines whether an agent i moves or not, is its own intolerance value T_i, instead of that of its neighbors, as in the previous case. Though both dynamical rules are based on homophile dissatisfaction, they in fact implement different plausible mechanisms for mobility. Whether the average social rejection (hostility) of my neighbors is more important than my own degree of tolerance with a dissimilar environment or not, in the decision of moving, may be a question with widely different (as well as context-dependent) individual answers, and it is certainly not inside the scope of this work to enter into such a discussion. We regard here both as alternative plausible mechanisms for mobility, which may lead to differences regarding the selective advantage of tolerant traits in the Axelrod-Schelling model with heterogeneous intolerance.

We show in figure 8.6 the results obtained for the dynamical rule associated to equation (8.10). Though the values of T_D are in this scheme higher than those characteristic of the scheme analyzed before, certain degree of selective advantage of tolerant traits is unambiguously observed. Also, the selective advantage is stronger

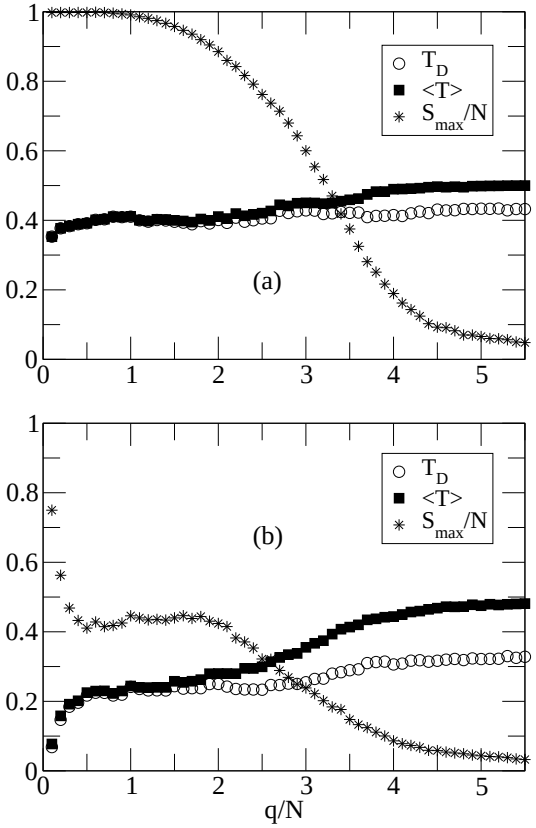


Figure 8.6: Order parameter $\langle S_{max} \rangle / N$ (stars), intolerance T_D of the dominant culture (circles), and average intolerance $\langle T \rangle$ (squares) versus scaled initial cultural diversity q/N for a lattice linear size $L = 40$, for the scheme of mobility corresponding to equation (8.10). Panel (a) corresponds to a density of empty sites $h = 0.05$. Panel (b) corresponds to $h = 0.5$. See the text for further details.

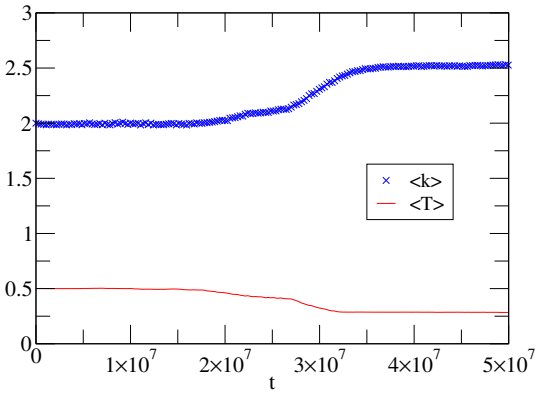


Figure 8.7: Time evolution of the average number of neighbors per agent $\langle k \rangle$ and average intolerance $\langle T \rangle$, for $q/N = 1$, $L = 40$ and $h = 0.5$ as obtained from 200 realizations in the scheme of equation (8.10).

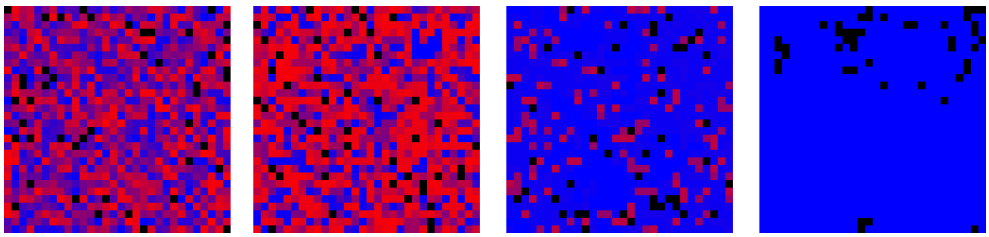


Figure 8.8: Time evolution of the tolerance distribution when the density of empty sites is below their percolation threshold in a representative realization. Each cell represents a node. The color code is a quasi-continuum, from blue for tolerant nodes to red for intolerant ones. Left chart represents the initial conditions, center charts the intermediate states and the right one the stationary state. Here has been taken L=30, h=0.05.

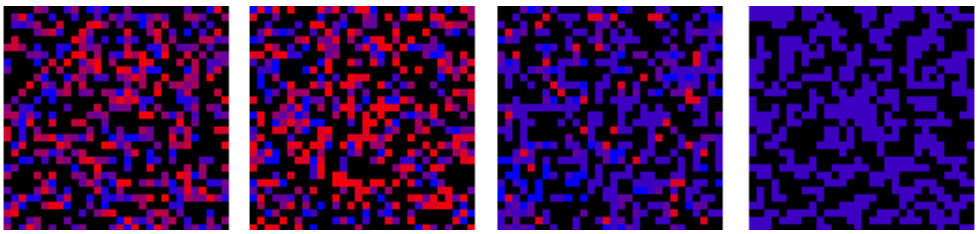


Figure 8.9: Time evolution of the tolerance distribution when the density of empty sites is above their percolation threshold in a representative realization. Codes are the same that in figure 8.8. L=30, h=0.5.

for high density h of empty sites, as before. Now, however, agents move depending on their own intolerance values, and then it is not (at least) as clear as before that a negative correlation degree-intolerance could be established, which would in turn explain the selective advantage of tolerant traits.

A possibility for this comes from the fact that intolerant agents move to empty sites more easily than tolerant agents do, so that a negative ($k_i,\ T_i$) correlation could appear provided the lattice sites occupied by agents are more likely to have agents in their neighborhood than empty sites are. To check for this, we have computed the time evolution of the average number of neighbors $\langle k \rangle$ of agents. Figure 8.7 shows that, after some (long) transient, the average degree of agents increases above its initial value (that is $\langle k \rangle = 4(1-h)$, for a square lattice and von Neumann neighborhood). This increase of $\langle k \rangle$ corresponds to the coalescence of clusters that will become monocultural in due (short) time. Interestingly, we also see in figure 8.6 the decrease of the average intolerance $\langle T \rangle$ as soon as the average degree increases, so giving further support to the argument.

Consequently, also in the case that the agents' mobility is the result of their own intolerance to cultural dissimilarity, the tolerant traits possess selective advantage due to the establishment of a negative ($k_i,\ T_i$) correlation which in this case has its origin in the agents' aggregation processes concomitant to the increase of local cultural overlaps. The observed fact that the selective advantage of tolerant traits is now weaker than in the case when mobility is induced by social rejection, may likely be the effect of two confluent factors; on one hand, the development of a negative degree-intolerance correlation is not now a direct consequence of the dynamical rule, and on the other, as analyzed in previous section 8.2, agents' aggregation processes are much less effective when intolerance enter as a threshold for mobility.

8.4 Summary and concluding remarks.

In the Axelrod-Schelling model for cultural dissemination among mobile agents, we have considered the intolerance, that was originally (chapter 7) a model parameter controlling the strength of agents' mobility, as a variable associated to a cultural feature, and thus subjected to cultural transmission. We have performed extensive numerical simulations for two different dynamical rules for mobility, whose respective homogeneous versions are analyzed with respect to the transition from topologically fragmented local consensus to global cultural consensus that occurs at very low values of the initial cultural diversity. In the first of these dynamical rules (mobility by social rejection) agents move due to the intolerance of their neighbors, weighted by their cultural dissimilarity, while in the second one the mobility depends on the agent's own intolerance to the cultural dissimilarity with its environment. In both

cases our results indicate that tolerant traits are selectively advantageous, so that the intolerance values present in the dominant culture are preferentially low. One then sees how the breaking of the original symmetry (indifference of the dynamics respect to particular feature's trait values, that leads to purely neutral selection of dominant characters in cultural evolution) effectively allows for the appearance of natural selection of advantageous traits.

The selective advantage of tolerant traits increases with the density h of empty lattice sites, and is also higher for the first scheme, where mobility is the result of the social rejection from the neighborhood. A stochastic analysis allows the rationalization of all these numerical observations, and points to the dynamical development of a negative correlation between the number of neighbors of an agent and its intolerance value as the origin of the selective advantage of tolerant traits. We should emphasize here that regarding the rule of cultural imitation, nothing privileges tolerant traits over intolerant ones, *i.e.*, Axelrod's cultural interactions are completely unbiased, so the bias towards tolerant traits can only come from the influence of the tolerance cultural feature on the mobility of agents, that shapes the instantaneous network of interactions among cultural agents. One should expect analogous findings for other network updating dynamics as the one considered (in the symmetric context) by [166, 167], also showing topologically fragmented phases, provided the trait symmetry is broken at the network updating rule level.

In this regard, the term tolerance -in the context of the Axelrod-Schelling model- has a very precise and narrow meaning, much more limited than its usual meaning in social science and political philosophy, where it certainly means much more than just a conditioning factor of the mobility of individuals and groups. However, inside the limitations of a simple agent based model like this one, our findings on the "adaptive to survival" character of tolerant traits in cultural dynamics, point to basic mechanisms that can be highly influential in cultural evolution.

Chapter 9

Co-evolutionnary network approach to cultural dynamics controlled by intolerance.

Starting from Axelrod's model of cultural dissemination, we introduce a rewiring probability, enabling agents to cut the links with their unfriendly neighbors if their cultural similarity is below a tolerance parameter. For low values of tolerance, rewiring promotes the convergence to a frozen monocultural state. However, intermediate tolerance values prevent rewiring once the network is fragmented, resulting in a multicultural society even for values of initial cultural diversity in which the original Axelrod model reaches globalization.

9.1 Introduction

The growing interest in the interdisciplinary physics of complex systems, has focussed physicists' attention on agent-based modeling [156, 32] of social dynamics, as a very attractive methodological framework for social sciences where concepts and tools from statistical physics turn out to be very appropriate [142] for the analysis of the collective behaviors emerging from the social interactions of the agents. The dynamical social phenomena of interest include residential segregation [52, 53], cultural globalization [59, 141], opinion formation [153, 168], rumor spreading [155, 169] and others.

The question that motivates the formulation of Axelrod's model for cultural dissemination [59] is how cultural diversity among groups and individuals could survive despite the tendencies to become more and more alike as a result of social interac-

tions. The model assumes a highly non-biased scenario, where the culture of an agent is defined as a set of equally important cultural features, whose particular values (traits) can be transmitted (by imitation) among interacting agents. It also assumes that the driving force of cultural dynamics is the "homophile satisfaction", the agents' commitment to become more similar to their neighbors. Moreover, the more cultural features an agent shares with a neighbor, the more likely the agent will imitate an uncommon feature's trait of the neighbor agent. In other words, the higher the cultural similarity, the higher the social influence.

The simulations of the model dynamics show that for low initial cultural diversity, measured by the number q of different traits for each cultural feature (see below), the system converges to a global cultural state, while for q above a critical value q_c the system freezes in an absorbing state where different cultures persist. The (non-equilibrium) phase transition [110] between globalization and multiculturalism was first studied for a square planar geometry [141, 143, 144], but soon other network structures of social links [145, 146, 152] were considered, as well as the effects of different types of noise ("cultural drift") [147, 148], external fields (modeling *e.g.* influential media, or information feedback) [149, 150, 163, 164], and global or local non-uniform couplings [165, 151].

In all those extensions of Axelrod's model mentioned in the above paragraph, the cultural dynamics occurs on a network of social contacts that is fixed from the outset. However, very often social networks are dynamical structures that continuously reshape. A simple mechanism of network reshaping is agents' mobility, and a scenario (named the Axelrod-Schelling model) where cultural agents placed in culturally dissimilar environments are allowed to move has been analyzed in chapters 7 and 8. In this model, new interesting features of cultural evolution appear depending on the values of a parameter, the (in-)tolerance, that controls the strength of agents' mobility.

A different mechanism of network reshaping has been considered in [166, 167], where a cultural agent breaks its link to a completely dissimilar neighbor, redirecting it to a randomly chosen agent. At variance with the mobility scenario of the Axelrod-Schelling model, that limits the scope of network structures to clusters' configurations on the starting structure (square planar lattice, or others), the rewiring mechanism allows for a wider set of network structures to emerge in the co-evolution of culture and social ties [170].

In this chapter we introduce in the scenario of network rewiring a tolerance parameter Z controlling the likelihood of links rewiring, in such a way that the limit $Z = 1^-$ recovers the case analyzed in [166, 167], where only links with an associated null cultural overlap are broken. Lower values of Z correspond to less tolerant attitudes where social links with progressively higher values of the cultural overlap

may be broken with some probability that depends on these values. The results show a counterintuitive dependence of the tolerance Z on the critical value q_c. On one hand, as expected from [166, 167], rewiring promotes globalization for high values of the tolerance, but on the other hand, very low values of Z (which enhance the rewiring probability) show the higher values of q_c. Indeed, a non monotonous behavior is observed in $q_c(Z)$: Our results unambiguously show that for some intermediate values of the tolerance Z, cultural globalization is disfavored with respect to the original Axelrod's model where no rewiring of links is allowed. In other words, rewiring does not always promote globalization. On the other hand, the resulting network topology depends on q, changing from a Poisson connectivity distribution $P(k)$ to a fat tailed distribution for $q \sim q_c$.

9.2 The model

As in Axelrod's model, the culture of an agent i is a vector of F integer variables $\{\sigma_f(i)\}$ ($f = 1, ..., F$), called cultural *features*, that can take on q values, $\sigma_f(i) = 0, 1, ..., q-1$, the cultural *traits* that the feature f can assume. The N cultural agents occupy the nodes of a network of average degree $\langle k \rangle$ whose links define the social contacts among them. The dynamics is defined, at each time step, as follows:

- Each agent i imitates an uncommon feature's trait of a randomly chosen neighbor j with a probability equal to their *cultural overlap* ω_{ij}, defined as the proportion of common cultural features,

$$\omega_{ij} = \frac{1}{F} \sum_{f=1}^{F} \delta_{\sigma_f(i),\sigma_f(j)}, \tag{9.1}$$

  where $\delta_{x,y}$ denotes the Kronecker's delta which is 1 if $x = y$ and 0 otherwise. The whole set of N agents perform this step in parallel.

- Each agent i disconnects its link with a randomly chosen neighbor agent j with probability equal to its *dissimilarity* $1 - \omega_{ij}$, provided the dissimilarity $1 - \omega_{ij}$ exceeds a threshold (*tolerance*) Z,

$$1 - \omega_{ij} > Z \ , \tag{9.2}$$

  and rewires it randomly to other non-neighbor agent. The tolerance $0 \leq Z \leq 1$ is a model parameter.

First we note that the initial total number of links in the network is preserved in the rewiring process, so the average degree $\langle k \rangle$ remains constant. However, the rewiring process allows for substantial modifications of the network topological features, *e.g.* connectedness, degree distribution, etc. In that respect, except for the limiting situation of very low initial cultural diversity q and a very high tolerance Z (where the likelihood of rewiring could be very low), one should expect that the choices for the initial network of social ties have no influence in the asymptotic behavior of the dynamics.

When the threshold tolerance Z satisfies $\frac{F-1}{F} \leq Z < 1$, only those links among agents with zero cultural overlap are rewired, so the model becomes the one studied in [166, 167]. On the other hand, when the tolerance takes the value $Z = 1$, there is not rewiring likelihood and the original Axelrod's model is recovered. When $Z = 0$ rewiring is always possible provided the cultural similarity is not complete, *i.e.*, $\omega_{ij} \neq 1$, so that it corresponds to the highest intolerance.

The usual order parameter for Axelrod's model is $\langle S_{max} \rangle / N$, where $\langle S_{max} \rangle$ is the average (over a large number of different random initial conditions) of the number of agents sharing the most abundant (dominant) culture, and N is the number of agents in the population. Large values of the order parameter characterize the globalization (cultural consensus) regime. We also compute the normalized size $\langle S_{top} \rangle / N$ of the largest network component (*i.e.*, the largest connected subgraph of the network).

9.3 Results and discussion

We have studied networks of sizes $N = 900, 1600$; averaging over 50 - 2000 replicas. Checks for robustness of main results with larger size of $N = 2500$ were also made. The considered cultural vectors have $F = 10$ cultural features, each one with a variability $q = 5$ - 10000. We studied different values of the tolerance threshold $Z \in (0, 1)$ and different values of the average connectivity $\langle k \rangle = 4, 10, 20, 40$. Each simulation is performed for N, F, $\langle k \rangle$, Z, and q fixed. For the sake of comparison with previous results [166, 167], we will present results for $\langle k \rangle = 4$.

The behavior of the order parameter for different values of Z is seen in Fig. 9.1. Like in [166], three different macroscopic phases are observed with increasing values of q, namely a monocultural phase, with a giant cultural cluster, a multicultural one with disconnected monocultural domains, and finally a multicultural phase with continuous rewiring. The nature of the latter phase has been successfully explained in [166]: At very large values of the initial cultural diversity q, the expected number of pairs of agents sharing at least one cultural trait becomes smaller than the total

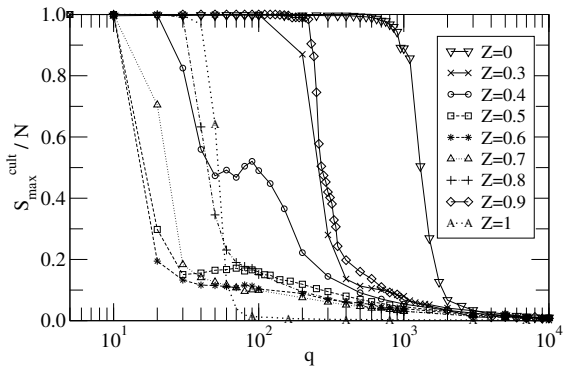


Figure 9.1: Order parameter as a function of the variability q for different values of the tolerance threshold Z. $N = 900$, $\langle k \rangle = 4$, average over 1000 replicas.

number of links in the network, so that rewiring cannot stop. Here we will focus attention on the first two phases and the transition between them.

In figure 9.2 we show the size distribution of the dominant culture over different realizations, measured for different values of q, at a particular fixed value of the tolerance $Z = 0.5$. In the region of q values near the transition from globalization to multiculturalism, the distribution is double peaked, indicating that the transition is first order, as in the original Axelrod's model. The transition value, q_c, may be roughly estimated as the q value where the areas below the peaks of the size distribution are equal. The estimates of the transition points for different values of the tolerance Z are shown in Fig. 9.3. The non monotonous character of the graph $q_c(Z)$ seen in this figure reveals a highly non trivial influence of the tolerance parameter on the co-evolution of cultural dynamics and the network of social ties.

Let us first consider the (most tolerant) case $Z = 0.9$ that, except for the system size N, coincides exactly with the situation considered in [167], *i.e.*, only links with zero cultural overlap are rewired. As discussed in [167], for q values larger than the critical value for a fixed network ($q_c(Z = 1) \simeq 60$), rewiring allows redirecting links with zero overlap to agents with some common cultural trait (compatible agents), so reinforcing the power of social influence to reach cultural globalization. Once all links connect compatible agents, rewiring stops (note: the decrease to zero of a positive cultural overlap cannot be strictly excluded, though it may be considered as a non typical event). From there on, the network structure will remain fixed, and globalization will be reached with the proviso that the network has so far remained connected. This is the case for most realizations (for $N = 900$) up to values of $q \sim 240$. Increasing further the cultural diversity q, increases the frequency of rewiring events and slows down the finding of compatible agents, favoring the topological fragmen-

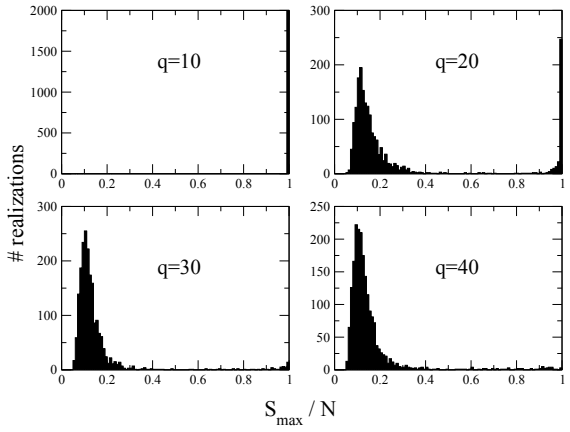


Figure 9.2: Histograms of S_{max}/N for different values of q, and for a fixed tolerance $Z = 0.5$, $N = 900$, $\langle k \rangle = 4$. From this figure one gets $q_c \sim 20$.

tation into network components before rewiring stops. Under these conditions, the asymptotic state will consist of disconnected monocultural components.

On one hand, network plasticity allows to connect compatible agents, so promoting globalization; but on the other hand it may produce network fragmentation, so favoring multiculturalism. What we have seen in the previous paragraph is that for $Z = 0.9$ the first effect prevails over the second one up to $q_c(Z = 0.9) \simeq 240$. Going from there to less tolerant situations (decreasing Z), increases the likelihood of rewiring, making easier that network fragmentation occurs before rewiring stops. This has the effect of decreasing the critical value q_c. In fact, from Fig. 9.3 we see that for $Z = 0.7, 0.6,$ and 0.5 multiculturalism prevails for cultural diversities where the original Axelrod's model shows cultural globalization. In these cases network plasticity promotes multiculturalism in a very efficient way: Agents segregate from neighbors with low cultural similarity and form disconnected social groups where full local cultural consensus is easily achieved, for q values low enough to allow a global culture in fixed connected networks.

For very low values of the tolerance parameter, though network fragmentation occurs easily during the evolution, Fig. 9.3 shows that globalization persists up to very high values of the initial cultural diversity q. To explain this seemingly paradoxical observation, one must realize that network fragmentation is not an irreversible process, provided links connecting agents with high cultural overlap have a positive rewiring probability. Under these circumstances, transient connections among different components occur so frequently so as to make it possible a progressive cultural homogenization between components that otherwise would have separately reached different local consensuses. Fig. 9.4 illustrates the time evolution for $q = 100$ and

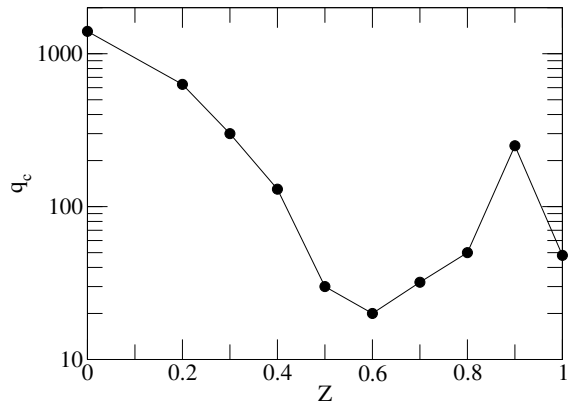


Figure 9.3: Critical value of the diversity q_c versus the tolerance threshold Z, obtained from the distribution of sizes of the dominant culture. $N = 900$, $\langle k \rangle = 4$. See the text for further details.

different values of Z. Panel (a) shows an example of cultural evolution where network fragmentation reverts to a connected monocultural network for $Z = 0.2$. Panel (b), that corresponds to $Z = 0.6$, shows that social fragmentation persists during the whole evolution, while in panel (c), which corresponds to the most tolerant situation ($Z = 0.9$), the network remains connected all the time.

The degree distribution of the network is Poissonian centered about $\langle k \rangle$ for all q values, except for $q \gtrsim q_c$ where it becomes fat tailed, with several lowly connected (and disconnected) sites. For very high q values, in the dynamical phase, the network rewiring is esentially random, so $P_q(k)$ is again Poisson like, centered around $\langle k \rangle$.

9.4 Summary

In this chapter we have generalized the scenario for co-evolution of Axelrod's cultural dynamics and network of social ties that was considered in [166, 167], by introducing a tolerance parameter Z that controls the strength of network plasticity. Specifically, Z fixes the fraction of uncommon cultural features above which an agent breaks its tie with a neighbor (with probability equal to the cultural dissimilarity), so that, the lower the Z value, the higher the social network plasticity.

Our results show that the network plasticity, when controlled by the tolerance parameter, has competing effects on the formation of a global culture. When tolerance is highest, network plasticity promotes cultural globalization for values of the initial cultural diversity where multiculturalism would have been the outcome for fixed networks. On the contrary, for intermediate values of the tolerance, the network

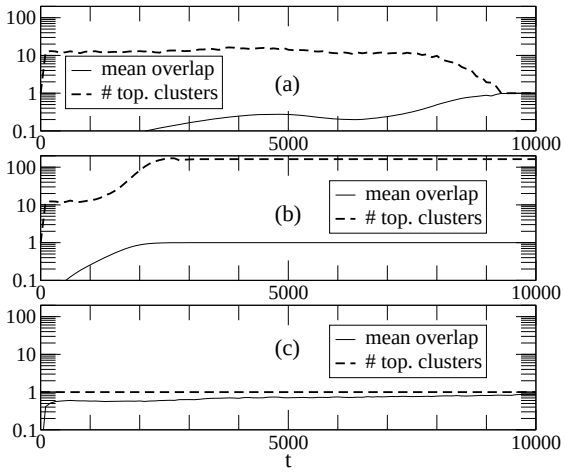


Figure 9.4: Time evolution of mean overlap and number of topological clusters for diferent values of tolerance $Z = 0.2$(a), $Z = 0.6$(b), $Z = 0.9$(c). $N = 900$, $q = 100$. See the text for further details.

plasticity produces the fragmentation of the (artificial) society into disconnected cultural groups for values of the initial cultural diversity where global cultural consensus would have occurred in fixed networks. For very low values of the tolerance, social fragmentation occurs during the system evolution, but the network plasticity is so high that it allows the final cultural homogenization of the transient groups for very high values of the cultural diversity. Intermediate tolerances promote multiculturalism, while both extreme intolerance and extreme tolerance favor the formation of a global culture, being the former more efficient than the latter.

Chapter 10

Conclusions.

Starting from the idea of many interacting entities, we have addressed different social and economic issues using procedures and theoretical tools from complex systems physics, in addition to other fields, such as complex networks and game theory. Although social sciences obviously have their own methods to deal with such kind of problems, this methodology yields new approaches, especially in problems that involve stochastic and/or nonlinear dynamics aspects, and enhances the study of emergent properties arising from aggregating approaches.

In the first part of the thesis, we address a issue related to different sciences such as biology, economics or sociology: the evolution of cooperation in hostile environments, that is, when in the first instance the selfish behavior is more advantageous for the individual that cooperative action. This problem has been dealt with in a variety of ways. In this context, the Prisoner's Dilemma (PD) has become a paradigm for studying the emergence of cooperative behavior. Besides, the thermodynamical perspective on evolutionary game dynamics studies (used, *e.g.*, in chapter 2) is not a new issue, as can be found in research literature on game theory [16, 121], and allows us to interpret the social indicators as physical observables and later to infer analytical results.

In chapter 2 we investigate in detail the dynamics of PD in an artificial network (Dipole Model) that models the influence on a population of two antagonist hubs connected to the whole population, but with no direct connection between them. Based on previous studies [109, 112, 113] that have shown that the asymptotic states of evolutionary PD in complex networks are characterized by three kinds of agents (pure cooperators, pure defectors and fluctuating agents), we designed the model so that the hubs remain as pure strategist; more specifically, we constrained the initial conditions in order to cancel the probability of strategy change for the hubs. The analytical formulation enables a thermodynamic approach of the system, which provides a de-

scription with a range of validity limited by the effects of network topology. The results, besides providing mathematical meaning to the concept of social temperature, help to understand the behavior of a population under the influence of two opposite influential elements (*e.g.*, mass media).

In chapter 3, we study the reversibility of the evolutionary dynamics of PD in different complex networks under adiabatic variations of the temptation to defect. The results show that, for the topologies analyzed, the process is reversible provided it is kept away from absorbing states, but when the cooperation reaches a tipping point the system becomes irreversible showing a hysteresis cycle which is a function of the considered network. The causes of irreversibility vary from one topology to another: the centralization of cooperator clusters around cooperator hubs in scale-free (SF) networks prevents the onset of irreversibility in most SF networks. However, the multiple clusterization of cooperators in Erdös-Rényi (ER) networks determines that, once the tipping-point is reached, irreversible transitions always occur, and irreversibility is more evident around the absorbing states.

In chapter 4 we take into consideration a topic that has been deeply studied in the last years: *multiplex networks*. Usually, real populations (regardless of their nature) are not isolated, but interlinked by interactions between different layers [174, 175]. In addition, the interactions that take place within a given layer may be governed by different rules that the interactions between elements of different layers [176, 177, 178]. In this context, we study the influence of interdependency between different layers on the degree of cooperation in stratified systems. In particular, we focus on the case in which relations within layers are governed by an evolutionary PD, while elements of different layers interact through the Snowdrift (SD) game. This scheme models a situation in which defection is punished in interactions with *outsider defectors*. Our model consists of two populations, provided with an internal structure of contacts, that interact through interpopulation links. When the populations are well-mixed, we carry out analytical calculations that populations that show, in a region of the parameter space, a *polarized* state consisting of a *full-cooperation* community connected to a *full-defection* population. Other regions of the parameter space show *quasi-polarized* states, characterized by a population where every agent defects linked to another where most agents cooperate. In order to deal with networked populations, we solved the system numerically and found that previous states appear when population structure is a complex network of contacts, where network reciprocity promote cooperation. The results show that, while for small values of the temptation to defect parameter polarization opposes network reciprocity, for higher temptation values both mechanisms have the same sign promoting cooperation in a layer. As an application, we find that the cooperation level in a target population can be controlled through a coupled defective population.

However all the above, the statement that underlying network structure enhances cooperation in human interactions is based on some assumptions, namely that personal strategies depends on neighbor's pay-offs. Althoug network reciprocity mechanism in humans has been deeply studied in the last twenty years (see, *e.g.*, [16, 94, 18, 109, 95, 96, 97]), the conclusions are in general contradictory, because the strategies are usually a hypothesis of the models without experimental support [79]. In chapter 5, instead of assuming that people choose following one of the usal strategies, we analyzed the problem taken as starting point the results of recent experiments [100, 23] on the behavior of small human populations in iterated PD games. These works shown that people do not take into account the neighbors' payoffs, but, instead, they consider the cooperation level in their neighborhood. We have studied mathematically the implications of such strategies in heterogeneously-connected large populations. Specifically, we solved analytically the mean-field case and compared the theoretical results with data obtained from numerical simulations made in three network topologies: regular lattice, ER and SF. This comparison show that cooperation level is exactly the same, regardless of the network structure. The consequences of this prediction are very important, because, if eventually confirmed by experiments, they will allow to discard the network reciprocity mechanism in human prisoner's dilemma-like situations. Experimental confirmation implies experiments in heterogeneous networks, and therefore, large scale experiments.

In chapter 6 we show the results of the large-scale experimental test we have performed to test the conclusions above mentioned. Our experiment shows that, regarding human behavior, the underlying topology does not have influence in the observed cooperation level. In particular, the cooperation levels observed in a regular lattice and in a heterogeneous network are indistinguishable, moreover, the behavior of subjects appears to be independent of their connectivity. This conclusion applies only to human cooperation in static networks, therefore network reciprocity may still be relevant in other contexts (*e.g.*, in microbiology [137] or evolving social networks [179]). Our experiment confirm that most people follow the strategy shown in previous works [23], consisting of the imitation of neighbor's actions with a probability that depends on their frequency. Accordingly, the results confirm the theoretical prediction made in chapter 5. These results may be applied to promoting cooperation in real systems, the study suggests that improving network structure might be an innefective policy but invites to incentive individual behavior.

In the second part of the thesis, we study some aspects of social dynamics, focusing our attention in particular on a issue that has been approached from biological, sociological and philosophical perspectives: the intolerance [74]. Intolerance, defined as "the refusal to accept subjects with different characters, opinions or behaviors from one's own", is amenable to be dealt with agent based models (ABM), in fact, one of the first ABM designed to explore a social issue was the segregation

model of Schelling [52]. Despite this, ABM have not paid much attention to the intolerance in itself, while related issues of social dynamics -such as homophily [59] , opinion formation [153] or rumor spreading [155]- have been deeply and widely studied.

In chapter 7, we have introduced a model of cultural dynamics allowing agents to move according to their degree of cultural disagreement with their neighborhood. For small density of empty sites, the introduction of mobility into the Axelrod cultural dynamics promotes the convergence to cultural globalization, with the consequence that the order parameter scales with the system size. Therefore, the transition to multicultural population only occurs for finite populations (*i.e.*, in the thermodynamic limit there exists only monocultural phase). Furthermore, for larger densities of empty sites (*i.e.*, when population density is below the percolation threshold), a new phase (that we call *multicultural fragmented phase*) appears at low values of the initial cultural diversity q. The causes of this new phase can be founded in the early stages of the dynamics: for low values of q, the initial cultural overlap is enough to promote local cultural convergence enhancing the formation of isolated clusters of agents, wich isolation prevents cultural diffusion between different domains. For high enough values of the intolerance threshold T, the increase of q has the effect of increasing the initial mobility, and the monocultural phase of the original Axelrod model is recovered, followed by the multicultural phase for high enough values of q and finite populations. Moreover, in the last transition from order to disordered phase, the dynamics show *erosion-adhesion* processes associated to the agents mobility; in fact, the increase in q enhances the probability of agents without a common cultural feature with anyone else (*aliens*).

In chapter 8, we refined the Axelrod-Schelling model by considering the intolerance T as a variable associated to a cultural feature, and thus subjected to cultural transmission. There are (at least) two natural ways of relating mobility and individual intolerance: agents can move due to the intolerance of their neighbors (*social rejection*) or agents can move motivated by its own intolerance to the cultural dissimilarity with its environment (*self-exclusion*). In both cases our results indicate that tolerant traits are more likely to spread, so that the dominant culture tends to have low values of intolerance. In addition, the selective advantage of tolerant traits increases with the density h of empty sites, and is also higher for the *social rejection* scheme.

While in chapters 7 and 8 we introduced tolerance into the original Axelrod's model through mobility, in Chapter 9 we consider network plasticity, allowing agents to remove links to its dissimilar neighbors. Rewiring mechanism into Axelrod dynamics was considered by Vazquez, González-Avella, Eguíluz and San Miguel [166]; in their model, an agent was able to break its link to an antithetical neighbor (*i.e.*, they do not share any trait), redirecting it to another agent at random. In our research, we

have generalized the model proposed in Refs. [166, 167], by introducing a tolerance parameter Z which modulates the intensity of plasticity. We show that rewiring mechanism can produce opposite effects, depending on the tolerance value Z: while for large values of Z, rewiring enhances monoculturalism compared to non-evolving networks, for intermediate values of the Z, rewiring mechanism promotes clusterization into disconnected cultural groups for values of the initial cultural diversity q which would show monoculturalism in fixed networks. Finally, for low values of Z, transient states show clusterization phenomena, but rewiring (encouraged by low tolerance) enhances cultural transmision between groups for very high values of q, yielding monocultural states. In conclusion, intermediate tolerance values enhance diversity, high values of tolerance greatly promote globalization while low values of tolerance weakly promote it. Note that, unlike the Axelrod-Schelling described in chapters 7-8, the networks used are heterogeneous and dynamic, with a network dynamics given by the cultural dynamics.

Bibliography

[1] Neumann, J.V. Morgenstern, O. *Theory of Games and Economic Behavior.* (Princeton Univ. Press, Princeton, 1944).

[2] Kuhn, H. W. *Preface to Waldegrave's Comments: Excerpt from Montmort's Letter to Nicholas Bernoulli* (London School of Economics and Political Science, London, 1968).

[3] von Neumann, J. *Mathematische Annalen.* **100**, 295, (1928). Translated: *Contributions to the Theory of Games, Volume IV (Annals of Mathematics Studies, 40)*, (A. W. Tucker and R. D. Luce, eds.), Princeton Univ. Press, Princeton, 1959).

[4] Nash, J. *Proc. Natl. Acad. Sci. USA.* **36**, 48, (1950).

[5] Ferguson, T.S. *Game Theory Course.* Mathematics Department, UCLA. http://www.math.ucla.edu/ tom/Game_Theory/Contents.html

[6] Shubik, M. *Game Theory in the Social Sciences: Concepts and Solutions.* (MIT Press, Cambridge, 1984).

[7] Myerson, R. *Game Theory: Analysis of Conflict.* (Harvard Univ. Press, Cambridge, 1991).

[8] Schelling, T.C. *The Strategy of Conflict.* (Harvard Univ. Press, Cambridge, 1960).

[9] Hirshleifer, J. *UCLA Economics Working Papers.* **320**, (1984).

[10] Sion, M. *Pac. J. Math.* **8**, 171, (1958).

[11] Maynard-Smith J. and Price, G. R. *Nature.* **246**, 5427, (1973).

[12] Hofbauer, J. and Sigmund, K., *Evolutionary Games and Population dynamics*. (Cambrige Univ. Press, Cambridge, UK, 1998).

[13] Samuelson, L. , *JEP.* **16(2)**, 16, (2002).

[14] Maynard-Smith, J. *Evolution and the Theory of Games* (Cambrige Univ. Press, Cambridge, UK, 1982)

[15] Weibull, J. W. *Evolutionary game theory.* (MIT Press, Cambridge, 1995).

[16] Szabó, G. and Fath, G. *Phys. Rep.* **446**, 97 (2007).

[17] Gintis,H. *Game Theory Evolving*. (Princeton Univ. Press, Princeton, NJ, 2000).

[18] Santos, F.C. and Pacheco, J.M. *J. Evol. Biol.* **19**, 726 (2006).

[19] Nowak, MA. and May, R.M. *Nature* **359**, 826 (1992).

[20] Moran, P.A.P. *The Statistical Processes of Evolutionary Theory* (Clarendon Press, Oxford, 1962).

[21] Szabó, G. and Töke, C. *Phys. Rev. E* **58**, 69 (1998).

[22] Fudenberg, D. and Tirole, J. *Game Theory.* (MIT Press, Cambridge, 1991).

[23] Grujić, J. Fosco, C. Araujo, L. Cuesta, J.A. and Sánchez, A. *PLoS ONE* **5**(11), e13749 (2010).

[24] Nowak, M.A. Sasaki, A. Taylor, C. and Fudenberg, D. *Nature.* **428**, 646, (2004).

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

[26] Barabási, A.L. *Linked: The New Science of Networks.* (Perseus, Cambridge, 2002).

[27] Newman, M.E.J. *SIAM Review* **45**, 167 (2003).

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

[29] Albert, R. and Barabási, A.L. *Reviews of Modern Physics.* **74**, 47, (2002).

[30] Ulam, S. *A Collection of Mathematical Problems.* (Interscience, New York. 1960).

[31] Neumann, J. V. *Theory of Self-Reproducing Automata.* (University of Illinois Press, Champaign, IL. 1966).

[32] Epstein, J. M. and Axtell, R. *Growing Artificial Societies: Social Science from the Bottom Up* (The MIT Press, Cambridge, MA, 1996).

[33] Granovetter, M. *American J. Sociol.* **78**, 136, (1973).

[34] Marchiori, M. and Latora, V. *Physica A.* **285**, 539, (2000).

[35] Latora, V. and Marchiori, M. *Phys. Rev. Lett.* **87**, 198701, (2001).

[36] Newman, M.E.J. *Phys. Rev. E.* **64**, 016131, (2001).

[37] Newman, M.E.J. *Proc. Natl. Acad. Sci. USA.* **98**, 404, (2001).

[38] Barabasi, A. *et al.*, *Physica A.* **311**, 590, (2002).

[39] Csemerly, P. *Trends Biochem. Sci.* **29**, 331, (2004).

[40] Polis, G.A. *Nature.* **395**, 744, (1998).

[41] McCann, K. Hastings, A. and Huxel, G.R. *Nature.* **395**, 794, (1998).

[42] Guimera, R. Mossa, S. Turtschi and Amaral, L.A.N. *Proc. Natl. Acad. Sci. USA.* **102**, 7794, (2005).

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

[44] Latora, V. and Marchiori, M. *Eur. Phys. J. B* . **32**, 249, (2003).

[45] Barabási, A.L. Albert R. *Science* **286**, 509 (1999).

[46] Bollobás, B. and Riordan O. *Random Structures and Algorithms.* **18**, 279, (2001).

[47] Krapivsky, P.L. and Redner, S. *Phys. Rev. E.* **63**, 66123, (2001).

[48] Dorogovtsev S.N. and Mendes, J.F.F. *Europhys. Lett.* **52**, 2000, (33).

[49] Molloy, M. and Reed, B. *Random Struct. Algorithm.* **6**, 161, (1995).

[50] Molloy, M. and Reed, B. *Combin. Probab. Comput.* **7**, 295, (1998).

[51] Catanzaro, M. Boguña, M. and Pastor-Satorras R. *Phys. Rev. E* . **71**, 27103, (2005).

[52] Schelling, T.C. *J. Math. Sociol.* **1** 143, (1971).

[53] Schelling, T.C. *Micromotives and Macrobehavior* (Norton, New York, 1978).

[54] Clifford, P. and Sudbury, A. *Biometrika* **60 (3)**, 581–588 (1973).

[55] Redner, S. *A Guide to First-passage Processes* (Hardcover, 2001).

[56] Galam, S.*Eur. Physics J.B.* **25,** 403, (2002).

[57] Krapivsky, P.L. and Redner, S. *Phys. Rev. Lett.* **90,** 238701, (2003).

[58] Chen, P. and Redner, S. *J.Phys. A.* **38,** 7239, (2005).

[59] Axelrod, R. *J. Conflict. Res.* **41**, 203 (1997).

[60] Pastor-Satorras, R. and Vespignani, A. *Statistical Mechanics of Complex Networks.* (Cambridge Univ. Press, Cambridge, 2004).

[61] Pastor-Satorras, R. Rubí, J.M. and Díaz-Guilera, A. (Eds.). *Statistical Mechanics of Complex Networks.* (Springer, Berlin, 2003).

[62] Wasserman, S. and Faust, K. *Social Network Analysis: Methods and Applications.* (Cambridge, ENG and New York: Cambridge Univ. Press. 1994).

[63] Milgram, S. *Psychology Today* **2**, 60, (1967).

[64] Korte, C. and Milgram, S. *J. Personal. Social Psychol.* **15**, 101, (1970).

[65] Holland, P. and Leinhardt, S. *Amer. J. Sociology* **76**, 492, (1970).

[66] Wellman, B. *Science.* **293**, 2031, (2001).

[67] Castellano, C. Vilone, D. and Vespignani, A. *Europhys. Lett.* **63**, 153, (2003).

[68] West, S.A. Griffin, A.S. and Gardner, A. *J. Evol. Biol.* **20 (2)**, 415, (2007).

[69] Hamilton, W.D. *J. Theor. Biol.* **7 (1)**, 1, (1964).

[70] Wynne-Edwards, V.C. *Animal dispersion in relation to social behaviour.* (Oliver & Boyd, Edinburgh, 1962).

[71] Wilson, D.S. and Sober, E. *Behavioral and Brain Sciences.* **17 (4)**, 585, (1994).

[72] Trivers, R.*Q. Rev. Biol.* **46**, 35, (1971).

[73] Fowler, J.H. *Nature.* **437**, 7058, (2005).

[74] Razavi, M.A. and Ambuel, D. (eds.) *Philosophy, Religion, and the Question of Intolerance.* (State University of New York Press, Albany 1997).

[75] Nöel, L. *Intolerance, A General Survey.* (McGill-Queen's Univ. Press, NY, 1994).

[76] Erdös, P. and Rényi, A. *Publ. Math.*, **Debrecen 6**, 290 (1959).

[77] Erdös, P. and Rényi, A. *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **5**, 17–61. (1960).

[78] Cvetkovic, D.M. Doob, M. and Sachs, H. *Spectra of Graphs: Theory and Applications* (ed. New York, Wiley, 1998).

[79] Hofbauer, J. and Sigmund, K. *Bull. Am. Math. Soc.* **40**, 479 (2003).

[80] Nowak, M.A. *Evolutionary Dynamics. Exploring the Equations of Life*. (Harvard Univ. Press, Harvard, MA, 2006).

[81] Abramson, G. and Kuperman, M. *Phys. Rev. E.*, **63**, 030901(R) (2001).

[82] Eguíluz, V.M. Zimmermann, M.G. Cela-Conde, C.J. and San Miguel, M. *Am. J. Soc.* **110**, 977 (2005).

[83] Durán, O. and Mulet, R. *Physica D* **208**, 257 (2005).

[84] Hruschka, D.J. and Henrich, J. *J. Theor. Biol.* **239**, 1 (2006).

[85] Wu, Z. Xu, X. Huang, Z. Wang, S. and Wang,Y. *Phys. Rev. E* **74**, 021107 (2006).

[86] Ohtsuki, H. Hauert, C. Lieberman, E. and Nowak, M.A. *Nature* **441**, 502 (2006).

[87] Ohtsuki, H. and Nowak, M.A. *J. Theor. Biol.* **243**, 86 (2006).

[88] Roca, C.P. Cuesta, J.A. and Sánchez, A. *Phys. Life Rev.* **6**, 208 (2009).

[89] Roca, C.P. Cuesta, J.A. and Sánchez, A. *Eur. Phys. J. B* **71**, 587 (2009).

[90] Traulsen, A. Pacheco, J.M. and Nowak, M. A. *J. Theor. Biol.* **246**, 522 (2007).

[91] Gintis, H. *Game Theory Evolving* (2nd Ed, Princeton Univ. Press, Princeton, 2009).

[92] Rapoport, A. and Guyer, M.A. *General Systems* **11**, 203 (1966).

[93] Axelrod, R. *The Evolution of Cooperation* (Basic Books, New York, 1984).

[94] Perc, M. and Szolnoki, A. *BioSystems* **99**, 109 (2010).

[95] Szolnoki, A. and Perc, M. *Phys. Rev. E* **85**, 026104 (2012).

[96] Hauert, C. and Doebeli, M. *Nature* **428**, 643, (2004).

[97] Sysi-Aho, M. Saramäki, J. Kertész, J. and Kaski, K. *Eur. Phys. J. B* **44**, 129 (2005).

[98] Helbing, D. *Physica A* **181**, 29 (1992).

[99] Fischbacher, U. Gächter, S. and Fehr, E. *Econ Lett* **71**, 397 (2001).

[100] Traulsen, A. Semmann, D. Sommerfeld, R.D. Krambeck, H.J. and Milinski, M. *Proc. Natl. Acad. Sci. USA.* **107**, 2962 (2010).

[101] Luthi, L. Pestelacci, E. and Tomassini, M. *Physica A* **387**, 955 (2008).

[102] Ohtsuki, H. and Nowak, M.A. *J. Theor. Biol.* **251**, 698 (2008).

[103] Lozano, S. Arenas, A. and Sánchez, A. *PLoS ONE* **3(4)** e1892 (2008).

[104] Szolnocki, A. and Perc, M. *New J. Phys.* **10**, 043036 (2008).

[105] Hauert, C. *Int. J. Bif. and Chaos* **12**, 1531 (2002)

[106] Santos, F.C. Rodrigues, J,F. and Pacheco, J.M. *Proc. Natl. Acad. Sci. USA.* **273**, 51 (2006).

[107] Santos, F.C. Pacheco, J.M. and Lenaerts, T. *Proc. Natl. Acad. Sci. USA.* **103**, 3490 (2006).

[108] Santos, F.C. and Pacheco, J.M. ,*Phys. Rev. Lett.* **95**, 98104 (2005)

[109] Gómez-Gardeñes, J. Campillo, M. Floría, L. M. and Moreno, Y. *Phys. Rev. Lett.* **98**, 108103 (2007).

[110] Marro, J. and Dickman, R. *Nonequilibrium Phase Transitions in Lattice Models* (Cambridge Univ. Press, Cambridge, UK, 1999).

[111] Rong, Z. Li, X. and Wang, X. *Phys. Rev. E* **76**, 027101 (2007).

[112] Poncela, J. Gómez-Gardeñes, J. Floría, L. M. and Moreno, Y. *New J. Phys.* **9**, 184 (2007).

[113] Gómez-Gardeñes, J. Poncela, J. Floría, L. M. and Moreno, Y. *J. Theor. Biol.* **253**, 296 (2008).

[114] Poncela, J. Gómez-Gardeñes, J. Floría, L. M. Sánchez, A. and Moreno, Y. *PLoS ONE* **3(6)**, e2449 (2008).

[115] Roca, C. P. Cuesta, J. A. and Sánchez, A. *Phys. Rev. E* **80**, 046106 (2009).

[116] B. Skyrms *The Stag Hunt and the Evolution of Social Structure.* (Cambridge Univ. Press, Cambridge 2004).

[117] Edwards, M. Huet, S. Goreaud, F. and Deffuant, G. *J. Art. Soc. & Soc. Sim.* **6**, 4 (2003).

[118] Jaynes, E.T. *Phys. Rev.* **106**, 620 (1957).

[119] Courant, R. and Hilbert, D. *Methods of Mathematical Physics*, vol. I (Interscience, New York, NY, 1953).

[120] Tisza, L. *Generalized Thermodynamics* (MIT Press, Cambridge, MA, 1966).

[121] Blume, L. *Games and Econ. Behav.* **5**, 387 (1993).

[122] Page S.E. *The Difference. How the power of diversity creates better groups, firms, schools,and societies.* (Princeton Univ. Press, Princeton, NJ, 2007).

[123] Fehr, E. and Fischbacher, U. *Nature* **425**, 785 (2003).

[124] Pennisi, E. *Science* **325,** 1196 (2009).

[125] Rapoport, A. and Chammah, A.M. *PrisonerÕs Dilemma*. (University of Michigan Press, Ann Arbor, 1965).

[126] Axelrod, R. and Hamilton, W.D. *Science* **211**, 1390 (1981).

[127] Groves, T. and Ledyard, J. *Econometrica* **45**, 783 (1977).

[128] Hardin, G. *Science* **162**, 1243(1968).

[129] Nowak, MA. *Science* **314**, 1560 (2006).

[130] Helbing, D. and Yu, W. *Proc. Natl. Acad. Sci. USA.* **107**, 5265 (2010).

[131] Cassar, A. *Games Econ. Behav.* **58**, 209 (2007).

[132] Kirchkamp, O. and Nagel, R. *Games Econ. Behav.* **58**, 269 (2007).

[133] Gracia-Lázaro, C. Cuesta, J.A. Sánchez and A. Moreno, Y. *Sci. Rep.* **2**, 325 (2012).

[134] Langer, P. Nowak, M.A. and Hauert, C. *J. Theor. Biol*. **250**, 634 (2008).

[135] Ledyard, J.O. *Public goods: A survey of experimental research, Handbook of experimental economics* 111-251 Nagel, J.H. and Roth A.E. (eds) (Princeton Univ. Press 1995).

[136] Camerer, C.F. *Behavioral Game Theory* (Princeton Univ. Press, Princeton 2003).

[137] Velicer, G.J. *Trends Microbiol.* **11**, 330 (2003).

[138] Fehl, K. van der Post, DJ. Semmann, D. *Ecol. Lett.* **14**, 546 (2011).

[139] Rand, D.G. Arbesman and S. Christakis, N.A. *Proc. Natl. Acad. Sci. USA.* **108**, 19193 (2011).

[140] Gómez-Gardeñes, J. and Moreno, Y. *Phys. Rev. E* **73**, 056124 (2007).

[141] Castellano, C. Marsili, M. Vespignani, A. *Phys. Rev. Lett.* **85**, 3536 (2000).

[142] Castellano, C. Fortunato, S. Loreto, V. *Rev. Mod. Phys.* **81**, 591 (2009).

[143] Vilone, D. Vespignani, A. Castellano, C. *Eur. Phys. J. B* **30**, 399 (2002).

[144] Vázquez, F. Redner, S. *Europhys. Lett.* **78**, 18002 (2007).

[145] Klemm, K. *et al.*, *Phys. Rev. E* **67**, 026120 (2003).

[146] Klemm, K. *et al.*, *Physica A* **327**, 1 (2003).

[147] Klemm, K. *et al.*, *Phys. Rev. E* **67**, 045101(R) (2003).

[148] Klemm, K. *et al.*, *J. Econ. Dyn. Control* **29**, 321 (2005).

[149] Shibanai, Y. Yasuno, S. Ishiguro, I. *J. Conflict Res.* **45**, 80 (2001)

[150] González-Avella, J.C. Cosenza, M. G. Tucci, K. *Phys. Rev. E* **72**, 065102 (2005).

[151] González-Avella, J.C. *et al.*, *Phys. Rev. E* **73**, 046119 (2006).

[152] Guerra, B. Poncela, J. Gómez-Gardenes, J. Latora, V. and Moreno, Y. *Phys. Rev. E* **81**, 056105 (2010).

[153] Galam, S. *J. Stat. Phys.* **61**, 943 (1990).

[154] Galam, S. Chopard, B. Masselor, A. and Droz, M. *Eur. Phys. J. B* **4**, 529 (1998).

[155] Daley, D.J. and Kendall, D.J. *Nature* **204**, 1118 (1964).

[156] Axelrod R. and Tesfatsion, L. *Handbook of Computational Economics, Vol. 2: Agent-Based Computational Economics*, (Tesfatsion, L. and Judd, K. L. (eds), North Holland, Amsterdam, 2006).

[157] Axtell, R. Axelrod, R. Epstein, J. M. and Cohen, M. D. *Comput. Math. Organ. Theory* **1**, 123 (1996).

[158] Vinkovic, D. and Kirman, A. *Proc. Natl. Acad. Sci. (USA)* **103**, 19261 (2006).

[159] Stauffer, D. and Solomon, S. *Eur. Phys. J. B* **57**, 473 (2007)

[160] Dall'Asta, L. Castellano, C. and Marsili, M. *J. Stat. Mech.* L07002 (2008).

[161] Gauvin, L. Vannimenus, J. and Nadal, J.-P. *Eur. Phys. J. B* **70** 293 (2009).

[162] Clark, W.A.V. and Fossett, M. *Proc. Natl. Acad. Sci. (USA)* **105** 4109 (2009).

[163] González-Avella, J. C. Cosenza, M.G. Eguíluz, V.M. and San Miguel, M. *New J. Phys.* **12**, 013010 (2010).

[164] Rodríguez, A.H. and Moreno, Y. *Phys. Rev. E* **82**, 016111 (2010).

[165] González-Avella, Cosenza, M.G. Klemm, K. Eguíluz, V.M. and San Miguel, M. *J. Art. Soc. Soc. Simul.* **10**, (2007)

[166] Vázquez, F. González-Avella, J.C. Eguíluz, V.M. and San Miguel, M. *Phys. Rev. E* **76**, 046120 (2007).

[167] Centola, D. González-Avella, J.C. Eguíluz, V.M. and San Miguel, M. *J. Conflict Resol.* **51**, 905 (2007).

[168] Borge-Holthoefer J. *et al.*, *PLoS ONE* **6(8)**, e23883 (2011).

[169] Moreno, Y. Nekovee, M. Vespignani, A. *Phys. Rev. E* **69**, 055101 (2004).

[170] Holme P. and Newman, M.E.J. *Phys. Rev. E* **74**, 056108 (2006).

[171] Kollock, P. *Annu. Rev. Sociol.* **24**, 183174 (1998).

[172] Pennisi, E. *Science* **309**, 93 (2005).

[173] Assenza, S. Gómez-Gardeñes, J. and Latora, V. *Phys. Rev. E* **78**, 017101 (2008).

[174] Szell, M. Lambiotte, R. and Thurner S. *Proc. Natl. Acad. Sci. USA.* **107**, 13636 (2010).

[175] Mucha, P.J. Richardson, T. Macon, K. Porter, M.A. and Onnela, J.-P. *Science* **328**, 876 (2010).

[176] Buldyrev, S.V. *et al.*, *Nature* **464**, 1025 (2010).

[177] Parshani, R. Buldyrev, S. and Havlin, S. *Proc. Natl. Acad. Sci. USA.* **108**, 1007 (2011).

[178] Wang, Z. Szolnoki, A. and Perc, M. *EPL* **97**, 48001 (2012).

[179] Hanaki, N. Peterhansl, A. Dodds P.S. and Watts D.J. *Management Science* **53(7)** 1036 (2012)

ISBN 978-84-15770-24-4

Prensas de la Universidad
Universidad Zaragoza