4

Scale-free networks in complex systems

In the previous chapter we introduced the concept of *localized* information for a multi-agent model that simulates stock market dynamics. This means that the information is shared only within the various groups of agents and not at a global level. In particular, in the CA example, we considered the agents to gather in *clusters*, neglecting, in first approximation, the exact structure of these clusters (each agent in a cluster can talk with all the other agents of the same cluster) but just looking at their size.

In this chapter we further investigate the effect of the topology of the interactions between the elementary parts of complex systems by considering specific kinds of networks. This study is motivated by the peculiar nature of recent empirical findings that took place in different areas of research, from biology to engineering and from physics to the social sciences. The main result is that most of the networks show, despite having a different microscopic dynamics, a tendency to self-organize into structures that share common features. In particular, these networks are characterized by a power law distribution, $P(k) \sim k^{-\alpha}$, in the number of connections per node, k, over several orders of magnitude. Networks that fulfill this propriety of scale-invariance are referred to as "scale-free".

In this chapter we explore the implication of scale-free topologies, and in particular the one generated with the Barabási-Albert algorithm, in the antiferromagnetic (AF) Ising model and in a stochastic model of opinion formation. In the first case we show that the implicit disorder and frustration lead to a spin-glass phase transition not observed for the AF Ising model on standard lattices. We further illustrate that the opinion formation model produces a coherent, turbulent-like dynamics for a certain range of parameters. In this regime, the particular structure of the network plays an important role in sustaining the turbulent state even in the case of randomly undecided agents. This chapter represents the joint results of the two separate works in Refs. [BLT05b] and [BSLW05].

4.1 Introduction: empirical evidence for scale-free networks in Nature

One of the most intriguing issues in modern statistical mechanics concerns the role played by the topology of the interactions between the elementary parts of complex systems. This problem gained relevance after the results of several empirical studies carried out from the end of the 1990s, which shed some light on the basic principles of structural self-organization. A few examples include the food webs [Wea02, CGA02, MS02], power grids and neural networks [WS98, Aea00], cellular networks [J*et al.*00, Jea01], sexual contacts [Lea01], Internet routers [FFF99, PSV01, YJB02], the World Wide Web [AJB99, Kea00], actor collaborations [WS98, AJB99, Aea00, BA99], the citation networks of scientists [Red98, Vaz01] and the stock market [BCLM03, OCK$^+$03].

Although different in the underlying interaction dynamics or *micro-physics*, all these networks have shown a tendency to self-organize in structures that share common features. In particular, the number of connections, k, for each element, or node, of the network follow a power law distribution, $P(k) \sim k^{-\alpha}$. Networks that fulfill this property are referred to as *scale-free* networks (SFNs). In addition many of these networks are characterized by a high clustering coefficient, C, in comparison with random graphs [Bol85]. The clustering coefficient, C, is computed as the average of local clustering, C_i, for the ith node, defined as

$$C_i = \frac{2y_i}{z_i(z_i - 1)}, \tag{4.1}$$

where z_i is the total number of nodes linked to the site i and y_i is the total number of links between those nodes. As a consequence both C_i and C lie in the interval [0,1]. The high level of clustering found supports the idea that a *herding* phenomenon is a common feature in social and biological communities.

Numerical studies on SF networks have demonstrated how the topology plays a fundamental role in infection spreading [PSV01] and tolerance against random and preferential node removal [AJB00, CEbAH00, CNSW00]. A detailed description of the progress in this emerging field of statistical mechanics can be found in the recent reviews of Refs. [AB02, DM02].

Motivated by these findings, we investigate the influence of scale-free structures, and in particular the ones generated by using the Barabási-Albert algorithm described in the next section, in two different numerical models. The first of them is the well known AF Ising model, extensively studied on regular lattices, Sec. 4.3, while the second, Sec 4.4, is a model of stochastic opinion formation similar, in the decision making process, to the model presented in the previous

chapter.

4.2 The Barabási-Albert model

In order to study the effects of a scale-free topology on the previously mentioned models, we need an algorithm that is able to generate such a kind of structure.

In the present study we make use of the popular Barabási-Albert model [AJB99] (BA). This is based on two main assumption: (i) linear growth and (ii) preferential attachment. In practice the network is initialized with m_0 disconnected nodes. At each step a new node with m edges is added to the pre-existing network. The probability that an edge of the new node is linked with the ith node is expressed by

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j}. \tag{4.2}$$

The iteration of this preferential growing process yields a SFN, $P(k) \sim k^{-\alpha}$ where $\alpha = 3$. Note that, due to the probabilistic nature of the algorithm, this result is exactly true only if we average the exponent over several realizations of the network, especially if the number of nodes that we are considering is relatively small (for large networks the self-averaging effect becomes relevant). As a consequence, the physical quantities computed for a single realization are not characteristic of the system: they need to be averaged over an ensemble of networks to gain physical significance. As we show in the following sections, this implicit randomness leads to important physical consequences.

It is also worth noting that the Barabási-Albert model cannot reproduce a high clustering coefficient. In fact, the value of this coefficient depends on the total number of nodes in the network [AB02] and $C \to 0$ in the thermodynamic limit, $N \to \infty$. In order to account for this, we introduce a further step in the growth process, namely the triad formation proposed by Holme and Kim [HK02]. In this case, if the new added node is linked with an older node, i, having other links, then with a certain probability, θ, the next link of the new node, if any remain, will be added to a randomly selected neighbour of node i. This method of introducing friends to friends, while preserving the scale-free nature of the networks, generates high clustering coefficients that *do not depend* on the number of nodes in the network. The only tunable parameter that changes the value of the clustering coefficient is the clustering probability θ. An example of SFN generated with this algorithm is shown in Fig. 4.1 for 500 nodes [1].

[1] Another model for acquaintance networks, showing properties similar to the one presented in this work, has been proposed by Davidsen *et al.* [DEB02].

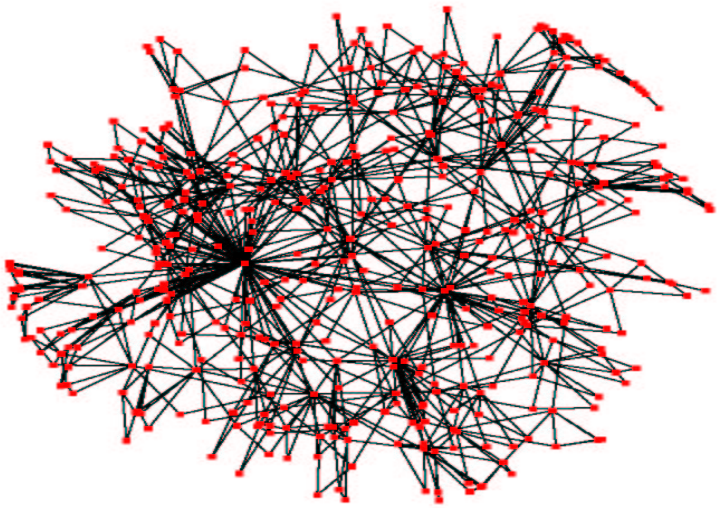

Figure 4.1: Example of a scale-free network. The number of nodes is 500 with clustering probability $\theta = 0.9$ and $m_0 = m = 2$, so that each new node is linked twice. The number of nodes has been kept small in order to preserve the clarity of the plot. Note that, for such small networks, a large scale invariant range is obtained only if one considers the ensemble average over several realizations. This plot has been realized with the Pajek software [sof05].

4.3 Spin-glass behaviour of the antiferromagnetic Ising model on a scale-free network

Inspired by he ubiquity of SFNs in natural systems, physicists have started to investigate the dynamics of standard models, such as the Ising model, in the case where the interactions between elements are described by this complex structure. An intriguing issue concerns how the unusual topology acts to influence the cooperative behaviour of the spins. Studies of the ferromagnetic (FM) Ising model on a SFN, using several theoretical techniques [AHS02, DGM02, IT02, Her04] including the Monte Carlo (MC) method [Her04], have found considerable robustness of ferromagnetic ordering against thermal fluctuations for the degree distribution exponent $\alpha \leq 3$.

The robustness feature is naturally expected as SFNs have large connectivities. This is analogous to the FM Ising model on a regular lattice above the lower critical spatial dimension, $d_l = 2$. There the ordered phase is very robust against thermal fluctuations. However, for the antiferromagnetic (AF) case with

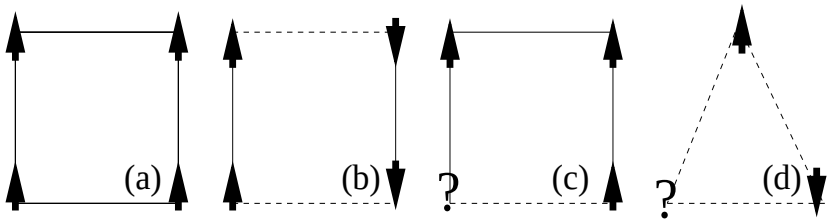

Figure 4.2: If we consider the continuous line to be a ferromagnetic bond, $J = +1$, and the dotted line to be an antiferromagnetic bond, $J = -1$, then it is clear that the configurations in (a) and (b) reach unique ground state, apart from the Z_2 symmetry, which minimize the Hamiltonian of Eq. (4.3). For the configuration (c) the situation is different. The spin denoted by the question mark (?) cannot decide a preferential orientation: the system is said to be *frustrated*. The presence of only antiferromagnetic bonds on a triangular lattice produce the same result, (d). In this case one of the three spins in the elementary cell is necessarily unsatisfied.

a SFN, the situation is different.

Two factors come to play a central role in the dynamics of the AF-SFN model; namely the competition induced by the AF interaction in the elementary triangles of the network and the randomness related to the non-regular connections. The abundance of elementary triangles in the network leads to "frustration". For example, only two of the three spins can be anti-aligned. More generally, frustration refers to the inability of the system to remain in a single lowest energy state (ground state). These ingredients lead the AF-SFN to belong to a class of randomly frustrated systems commonly referred to as spin glasses (SGs). Examples of non-frustrated and frustrated configurations are given in Fig. 4.2.

Most studies of SGs have been performed on regular lattices. These studies have shown that frustration and randomness are the key ingredients for SG behavior, characterized by a frozen random spin orientation at low temperatures [BY86, KR04]. Spin glasses on a SFN with mixed AF and FM bonds have been investigated recently by Kim *et al.* [KRKK05]. They found, for $\gamma \leq 3$ and even distributions of the two kinds of interaction, that the system is always in a SG state for any finite temperature. A study of the pure AF Ising model on a SFN is of great theoretical interest since, in fact, it does possess all the characteristics of a SG. Reviews on SG can be found in Refs. [BY86, KR04].

In order to shed some light on this issue we consider the AF Ising model on a SFN and, more precisely, the Barabási-Albert network with tunable clustering [HK02] introduced in the previous section. The thermodynamics quantities

of interest, along with the order parameters for the spin glass behaviour are calculated by using the replica exchange algorithm [HN96], that we introduce in the next subsection. An accurate determination of the transition temperature between paramagnetic phase and spin glass phase is archive by the evaluation of the Binder parameter, Subs. 4.3.3.

4.3.1 Model and algorithm: the replica exchange method

The AF Ising is quite simple to describe. Once the SFN has been constructed at the beginning of the simulation, we assign to each vertex an Ising spin, and to each link an AF interaction. The Hamiltonian can be written as follows

$$H = -\sum_{\langle ij \rangle} J_{ij}\, \sigma_i\, \sigma_j\,, \tag{4.3}$$

where the summation is performed over the connected spins σ_i and σ_j occupying sites i and j, respectively. The coupling interaction $J_{ij} = J = -1$ is AF.

For each size of the network under consideration we create an ensemble of configurations in order to have a reliable estimate of the physical quantities in which we are interested. All the simulations have been carried out fixing $\theta = 0.9$, corresponding to an average clustering coefficient of $C \sim 0.39$, close to the value found in many real systems [AB02]. It is worth recalling that this coefficient, due to the modified BA algorithm that we use, is *independent* of the network size. As previously mentioned, each vertex with the local cluster coefficient $C_i > 0$ together with its neighbours, compose elementary triangles, Fig. 4.3. Due to the AF interactions the system is locally frustrated, that is the Hamiltonian cannot be minimized for all the nodes at the same time: some of them must remain unsatisfied.

It follows that C is related to the degree of frustration of each network. Due to the probabilistic algorithm used for their construction, the value of C fluctuates around a mean value from one network to the next and, therefore, provides a source of randomness that, as we will see, gives rise to the spin glass properties of the model. This probabilistic growth is not shared by other algorithms which use recursion formulas to generate scale-free structures, such as, for example, the Apollonian networks [AH05]. In this case, once one fixes the number of iterations of the algorithm, which is proportional to the number of nodes of the final network, one also fixes its topology. The element of randomness is therefore missing in the Apollonian procedure.

As a random system, each realization of a network of size N will differ in the "structure" of connectivities. Therefore, in order to have reliable statistics, we average over many realizations of the SFN for each specified size. The system

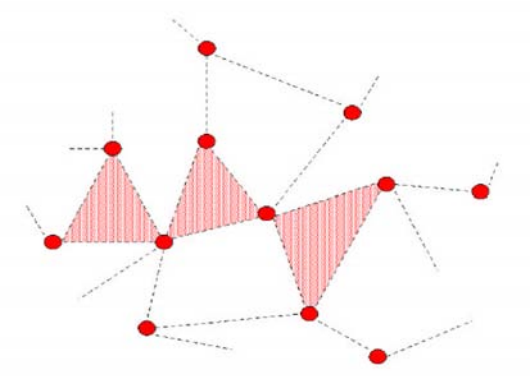

Figure 4.3: Representation of a part of the network. The shaded triangles are frustrated.

sizes that we simulate are $N = 1024$, 2048, 4096, and 8192. In general, one takes into account more realizations for small system sizes and less for large system sizes as the latter tend to self-average. However, since the self-averaging of physical quantities for larger system sizes is interfered with by the increase of ground state degeneracy, we do not take less realizations. Instead all physical quantities of interest for each system size are averaged over 1000 network realizations. Moreover, for each realization of the network, we fix $m_0 = m = 5$, corresponding to a coordination number on a regular lattice of approximately 10. In the thermodynamic limit, the average connectivity for the BA network is $\langle k \rangle = 2m = 10$, emphasizing the fact that we are implicitly dealing with a high dimensional system.

Another peculiarity of SFNs is the existence of a broad distribution of "hubs", that is nodes with a large number of connections, k. The energy difference in a spin flip actually depends on the number of connections of the spin itself, $\Delta E_i = -2\sigma_i \sum_{j=1}^{k_i} \sigma_j$. Thus in the AF case for the i^{th} spin with k_i connections, the hubs are more likely to "freeze" into a particular configuration compared to the nodes with just a few links. This property resembles the spin glass behaviour of particular alloys where some elements freeze into a particular orientation at a higher temperature than others.

The calculation of the thermal averages of the physical quantities of interest is performed using the *replica exchange* MC method [HN96]. In this method the evolution of M replicas, each in equilibrium with a heat bath of inverse temperature β_m for the m^{th} replica, is simulated in parallel. Given a set of inverse temperatures, $\{\beta\}$, the probability distribution of finding the whole system in a

state $\{X\} = \{X_1, X_2, \ldots, X_M\}$ is

$$P(\{X, \beta\}) = \prod_{m=1}^{M} \tilde{P}(X_m, \beta_m), \tag{4.4}$$

with

$$\tilde{P}(X_m, \beta_m) = Z(\beta_m)^{-1} \exp(-\beta_m H(X_m)), \tag{4.5}$$

and $Z(\beta_m)$ is the partition function at the m^{th} temperature. We can then define an exchange matrix between the replicas in our Markov chain, $W(X_m, \beta_m | X_n, \beta_n)$, that is the probability to switch the configuration X_m at the temperature β_m with the configuration X_n at β_n. By using the detailed balance condition, required to keep the entire system at equilibrium, on the transition matrix

$$\begin{aligned} &P(\ldots, \{X_m, \beta_m\}, \ldots, \{X_n, \beta_n\}, \ldots) \cdot W(X_m, \beta_m | X_n, \beta_n) \\ &= P(\ldots, \{X_n, \beta_m\}, \ldots, \{X_m, \beta_n\}, \ldots) \cdot W(X_n, \beta_m | X_m, \beta_n), \end{aligned} \tag{4.6}$$

along with Eq. (4.5), we have that

$$\frac{W(X_m, \beta_m | X_n, \beta_n)}{W(X_n, \beta_m | X_m, \beta_n)} = \exp(-\Delta), \tag{4.7}$$

where $\Delta = (\beta_n - \beta_m)(H(X_m) - H(X_n))$. With the above constrains we can choose the matrix coefficients according to the standard Metropolis method and, therefore, we have

$$W(X_m, \beta_m | X_n, \beta_n) = \begin{cases} 1 & \text{if} \quad \Delta < 0, \\ \exp(-\Delta) & \text{if} \quad \Delta > 0. \end{cases} \tag{4.8}$$

In our simulation we restrict the exchange to temperatures next to each other; that is, we consider only the terms $W(X_m, \beta_m | X_{m+1}, \beta_{m+1})$. This choice is motivated by the fact that the acceptance ratio decays exponentially with $(\beta_n - \beta_m)$.

The replica exchange method is extremely efficient for simulating systems such as spin glasses, that can otherwise become frozen in some particular configuration at low temperatures when using a standard Metropolis algorithm for the configuration update. In this case, as we lower the temperature, the system can become trapped into a local minimum of the free-energy where the barriers are so high that the time required for the system to move to another allowed region of the configuration space diverges to infinity as a function of the system size. If the system is trapped in a local minimum then the ergodicity condition is no longer fulfilled and the measure that one makes becomes biased by the particular region of the configuration space that is being sampled. By using the exchange replica method, instead, we keep switching the temperatures between the M

copies of the system and, as long as the higher temperature is in a hot phase (where, the system can easily explore all the configuration space), then we are in principle able to explore all the configuration space for the lower temperatures as well. Another advantage of this method is that the replica exchange reduces drastically the temporal correlation in the system dynamics at each temperature. This enables one to collect more independent measures for the thermal averages of the physical quantities and, therefore, reduces the uncertainty.

It is important to stress that, before starting the actual simulations, some care is required in selecting the set of inverse temperatures, $\{\beta\}$. In fact, the method is efficient only when a fairly large transition probability is maintained in the range of interest. From Eq. (4.8), we can see that, in the hot phase, temperatures can be more coarsely spaced while in the cold phase the temperatures need to be closer to each other. An optimal set of temperatures can be obtained by iterating, in preliminary runs, the following map [HN96]:

$$\begin{aligned} \tilde{\beta}_1 &= \beta_1, \\ \tilde{\beta}_m &= \tilde{\beta}_{m-1} + (\beta_m - \beta_{m-1}) \cdot p_m / c, \end{aligned} \tag{4.9}$$

where p_m is the acceptance ratio for the switch between two configurations at the m^{th} temperature and $c = \sum_{m=1}^{M} p_m/(M-1)$ is a normalization factor. The initial value for the set $\{\beta\}$ is uniform in the interval of interest and we ensure that β_1 belongs to the hot phase. For each iteration of the map, a run of a few thousand MC steps is carried out to calculate the acceptance ratios, p_m, which are then plugged into Eq. (4.9) in order to obtain a new set of inverse temperatures. After a few iterations, the map of Eq. (4.9) converges to a fixed point, $\{\beta^\star\}$, which sets the values of the temperatures to be used in our simulations.

In using this method, we define a "local" MC (LMC) update as a MC update for each spin of each replica, either consecutively through all elements of the network or randomly. Given that we can group the inverse temperatures in even and odd pairs, (β_m, β_{m+1}), after each LMC update we alternate attempts to switch configurations from one temperature to the next. According to this procedure, we define a Monte Carlo step (MCS) as a LMC plus a half (m odd or even) exchange trial.

For each realization of the network we start from a random configuration of the spins and then perform 10^3 LMC updates in order to reach thermal equilibrium. After this transient period, we run the simulation for 3×10^5 MCSs while taking a total of 6×10^4 measures for the thermal averages, that is one every 5 MCSs (temporal correlations are lost very quickly by using the replica exchange method). We consider low temperatures in a search for the possible existence of a phase transition. The thermal averages obtained for each network are then averaged over the ensemble of networks. In the following, we indicate

$\langle ... \rangle$ as the thermal average and $[...]_{\rm av}$ as the ensemble average. The statistical errors in the plots, where reported, are calculated via the bootstrap method.

4.3.2 Spatial correlations and specific heat

As a first step we investigate the extent of spatial correlation of the spins in the SFN by making use of the spatial autocorrelation function which is defined on a regular lattice as

$$\xi(\delta) = \left[\frac{1}{L_d} \langle \sigma_i \sigma_{i+\delta} \rangle \right]_{\rm av}, \tag{4.10}$$

where L_d is the total number of pairs at distance δ and depends just on the dimension considered. In a SFN the situation is more complicated since there may be several paths leading from a certain node to another. We then define δ as the *minimum* path between two nodes and the denominator of Eq. (4.10) becomes dependent on δ, $L_d \equiv L_d(\delta)$. The results, averaged over 50 configurations, between the temperatures $T = 5.0$ and $T = 2.1$, are shown in Fig. 4.4 for $N = 1024$. All the temperatures in the present paper are expressed in units of J/k_B, where J is the coupling strength between spins and k_B is the Boltzmann constant.

In order to give a better interpretation of the plot in Fig. 4.4 we remind the reader about an important propriety of SFNs; that is their "small world structure". The "hubs", in fact, play a fundamental role in linking sites otherwise very distant. Moreover, the average path length increases just logarithmically with the size of the network [AB02]. In the plot of Fig. 4.4, for $N = 1024$ nodes, an upper limit of $\delta = 6$ is encountered. While all the 50 configurations reach $\delta = 6$, only a few networks exceed this limit.

The plot emphasizes how neighboring spins, on average, tend to be anti-correlated, as expected in the AF case. The autocorrelation decreases with the distance from the node under consideration. The temperature dependence is also in accord with the expectations. The absolute value of the correlation decreases with increasing temperature and vice versa. Indeed, the highest and lowest temperatures form a perfect boundary for all the curves. This is an expected result, since thermal effects always tend to reduce the correlation between the spin interactions.

We also study the behaviour of the specific heat, C_ν, defined as follows

$$C_\nu(T) = \left[\frac{1}{N k_B T^2} (\langle E^2 \rangle - \langle E \rangle^2) \right]_{\rm av}. \tag{4.11}$$

Although no singularity is expected for this quantity in the spin-glass transition, it is interesting to compare its behaviour with other studies. The dependence of

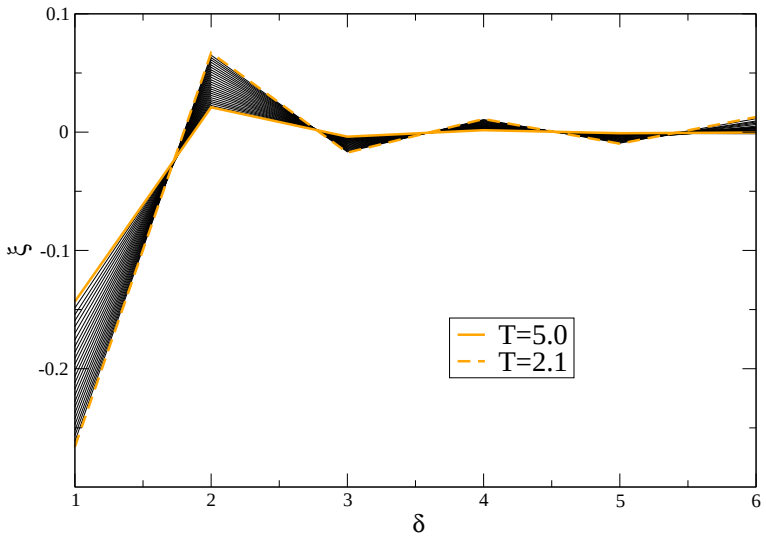

Figure 4.4: Spatial autocorrelation, $\xi(\delta)$, for $N = 1024$ averaged over 50 network configurations for temperatures between $T = 5.0$ and $T = 2.1$. The plot shows that next neighbour spins tend to be anti-parallel as in the standard AF Ising model. The AF interaction in the triangular units of the system results in high frustration. Note that the number of nodes at large distances is much smaller than the ones at smaller distances and so the average calculated for $\delta = 5$ and $\delta = 6$ includes just a few samples. This is a consequence of the "small-world" effects in SFNs.

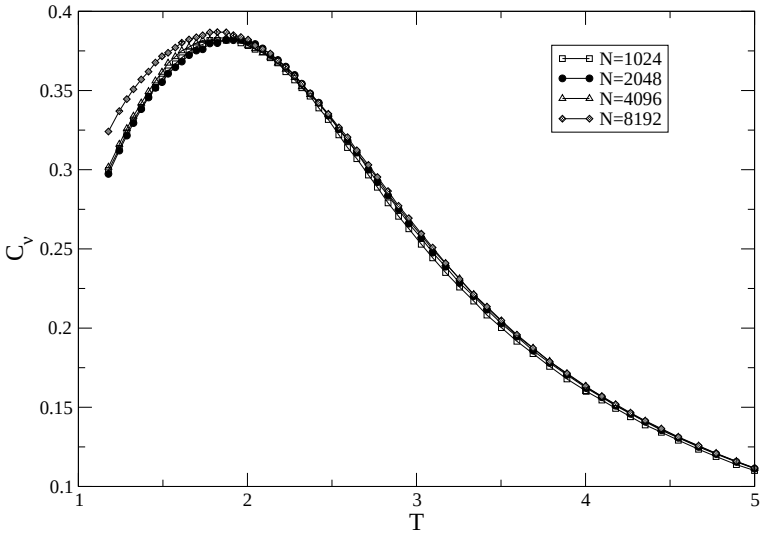

Figure 4.5: Specific heat, C_ν, as a function of the temperature and system size. The plot has been obtained by averaging over 50 network configurations for each N. Note that the specific heat does not scale with the size of the system.

the specific heat on temperature is reported in Fig. 4.5. The statistical errors, in this case, are smaller than the size of the symbols and therefore are not reported. A common Schottky peak of the specific heat for a finite system is observed at the temperature $T \sim 2.0$, independent of the system size. Below this point, we found that C_ν decreases and goes to zero as $T \to 0$.

This behaviour follows from simple entropy considerations. In fact, since we are dealing with a finite Ising system, the entropy is bounded at each finite temperature as well,

$$S(T) = \int_0^T \frac{C_\nu(T)}{T} dT < 2^N, \tag{4.12}$$

and, necessarily, $C_\nu \to 0$ for $T \to 0$.

The next subsection is dedicated to study of the SG behaviour and the phase transition of the system. In order to achieve this task, we evaluate the corresponding order parameters: the overlap parameter and the Binder parameter.

4.3.3 Observing spin glass behaviour

With the presence of frustration and randomness in the AF-SFN model, we expect to observe a spin glass transition, i.e., a transition from a temporal disordered to a temporal ordered phase at low temperatures.

This feature is not shared by the so-called fully frustrated systems [SOT04]. This type of transition might be characterized by the order parameter such as that suggested by Edward and Anderson [EA75], defined as follows

$$q_{EA} = \left[\frac{1}{N} \sum_i \langle \sigma_i \rangle^2 \right]_{\text{av}} . \tag{4.13}$$

However, an ergodic Markov chain of a system having Z_2 symmetry will ensure the thermal average of the i^{th} spin vanishes. Therefore a finite value of this measure simply reflects the non-ergodicity in the MC update.

A more appropriate quantity that is often used to characterize the SG state is the overlap parameter, q, defined as [Par83, BY88]

$$q = \frac{1}{N} \sum_i \sigma_i^{(\alpha)} \sigma_i^{(\beta)}, \tag{4.14}$$

where the superscripts α and β denote two copies of the same configuration of connectivity at the same temperature. The actual value of q is extracted from both the thermal and disorder average, $[\langle ... \rangle]_{\text{av}}$.

Using the replica exchange MC simulation, the two copies, α and β, are allocated at each temperature of the parallel tempering. This means, if the measurement is performed on M temperatures, there are M pairs of identical networks. The Metropolis spin update is performed on each replica for every MC step. As a part of the equilibration steps of the algorithm described in the previous section, we exchange two α (and β) replicas of neighboring temperatures, according to a certain probability. Then, for each temperature, the α and β replicas are superimposed every 5 MCSs in order to measure the overlap parameters, as defined in Eq. (4.14).

In particular, for the Ising system, due to the Z_2 symmetry, it is important to evaluate the absolute value of the order parameter,

$$|q| \equiv \left[\langle | \frac{1}{N} \sum_i \sigma_i^{(\alpha)} \sigma_i^{(\beta)} | \rangle \right]_{\text{av}} , \tag{4.15}$$

to overcome the implication of the Z_2 symmetry of the Hamiltonian, that is the configurations σ_i and $-\sigma_i$ have equal Boltzmann weights. That is, if the system is at thermal equilibrium and if we take quite long MCS then the usual q should average to zero. The existence of a spin glass phase is indicated by the convergence of $|q|$ to a finite value as we increase the network size. At the same time, a convergence of $|q|$ to zero at high temperatures is anticipated. In the latter case the system is in the paramagnetic phase.

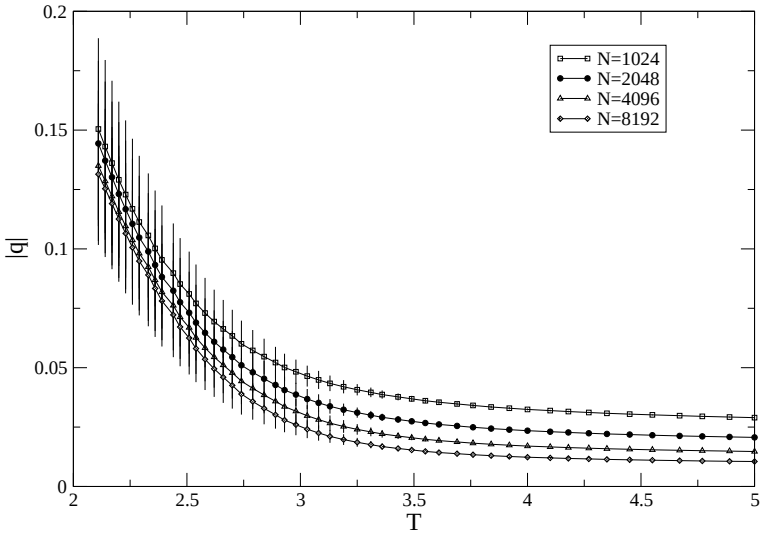

Figure 4.6: Temperature dependence of the overlap parameter, q, for different system sizes N. The increasing value of q at low temperatures indicates a SG phase. For a given network size, 1000 realizations of the SFN are averaged over.

The temperature dependence of $|q|$, resulting from the simulations, is shown in Fig. 4.6. The existence of a SG phase is indicated by the finite value of $|q|$ in the low temperature region, and the approach of $|q|$ to zero at higher temperatures associated with the paramagnetic phase. For high temperatures and large networks, $|q|$ approaches zero in accord with the thermodynamic limit, where $|q| = 0$ [Ogi85].

The existence of these two different phases can also be observed from the distribution of q, as shown in Fig. 4.7. For higher temperatures we observe simple Brownian fluctuations of the values of q, leading to a singly peaked Gaussian distribution characteristic of a paramagnetic state. By decreasing the temperature, the distribution spreads out, reflecting the increasing number of metastable disordered states associated with a substantial frustration. At lower temperatures the distribution develops double peaks reflecting the Z_2 symmetry and a finite value of $|q|$, representative of the SG phase. We note that the shape of the observed distribution at low temperatures is different from that of the conventional Ising system where the double peaks approach delta-like double peaks reflecting a simple doubly degenerate ground state [Dot01].

An accurate evaluation of critical temperature of the phase transition is

achieved via the Binder parameter defined as follows

$$g_L = \frac{1}{2}\left(3 - \frac{[\langle q^4\rangle]_{\rm av}}{[\langle q^2\rangle]^2_{\rm av}}\right). \tag{4.16}$$

Here $\langle q^2\rangle$ and $\langle q^4\rangle$ are respectively the second and the fourth cumulant moment of the overlap[2]. The Binder parameter is constrained in the range $0 \leq g_L \leq 1$. At high temperature, where thermal fluctuations overcome all cooperative interaction, the system is expected to exist in the paramagnetic phase where there is no spatial autocorrelation. As a result, the distribution of q should be Gaussian centered at $q = 0$. In this case the ratio of the cumulants, $\langle q^4\rangle/\langle q^2\rangle^2 \to 3$, resulting in $g_L \to 0$. At low temperatures, the cooperative interaction becomes dominant and the ratio of the cumulants approaches unity so that $g_L \to 1$.

Fig. 4.8 (inset) displays the temperature dependence of the Binder parameter for a variety of network sizes. A spin glass state is observed for lower temperatures where the Binder parameter deviates from zero, and increases with the system size while approaching to 1. In the thermodynamic limit, we expect $g_L \to 1$ just below the critical temperature. From the same plot, Fig. 4.8, we also observe a crossing point in the size dependence of g_L at $T \sim 4.0$, which indicates the critical temperature for the SG phase transition. For temperatures above $T \sim 4.0$ the Binder parameter, while remaining always above zero, does indeed order in an opposite manner indicative of a genuine crossing of the curves and in accord with a pure spin glass transition at finite temperature. This feature which is not observed for uniformly distributed AF and FM bonds, as $T_c = \infty$ in the thermodynamic limit [KRKK05]. However, the value of the transition temperature is not determined with high accuracy by the crossing of the Binder parameter. In fact, finite size effects seem to slightly distort the tendency for very small networks, as in the case of $N = 1024$. At the same time, the statistical errors in the paramagnetic phase for large networks, see $N = 8192$, appear to be significant and some points are scattered.

A more accurate estimate of the critical temperature, T_c, for finite size systems can be obtained using scaling arguments. For a SG system on a regular lattice, the Binder parameter depends on the system size L as

$$g_L = \tilde{g}_L[(T - T_c)L^{1/\nu}], \tag{4.17}$$

with $\nu > 0$. At T_c the Binder cumulant does not depend on L. For the SFN, the system size scales logarithmically with the number of nodes N [AB02], and

[2]To avoid systematic correlation errors that could bias the results if we were evaluating this average over g_L directly [KY96], the second and fourth order cumulants are averaged prior to taking their ratio.

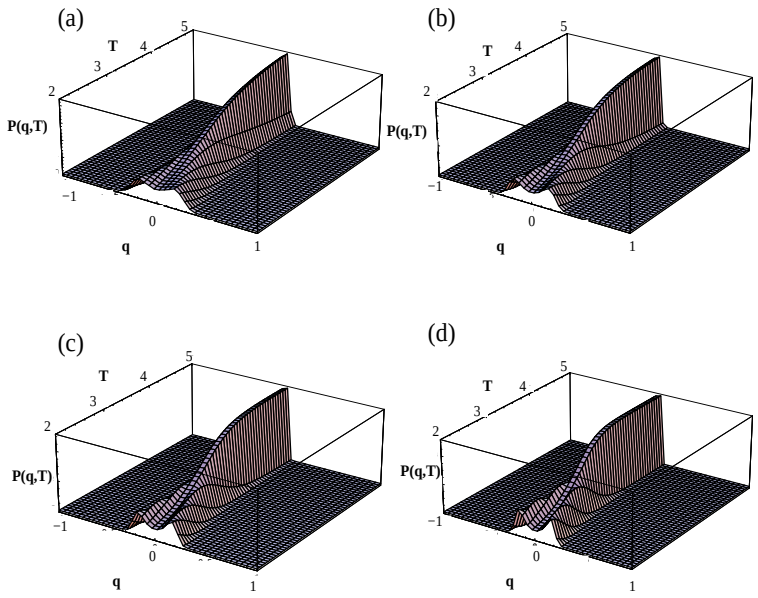

Figure 4.7: The distribution of q at various temperatures for different system sizes, including (a) $N = 1024$, (b) $N = 2048$, (c) $N = 4096$ and (d) $N = 8192$.

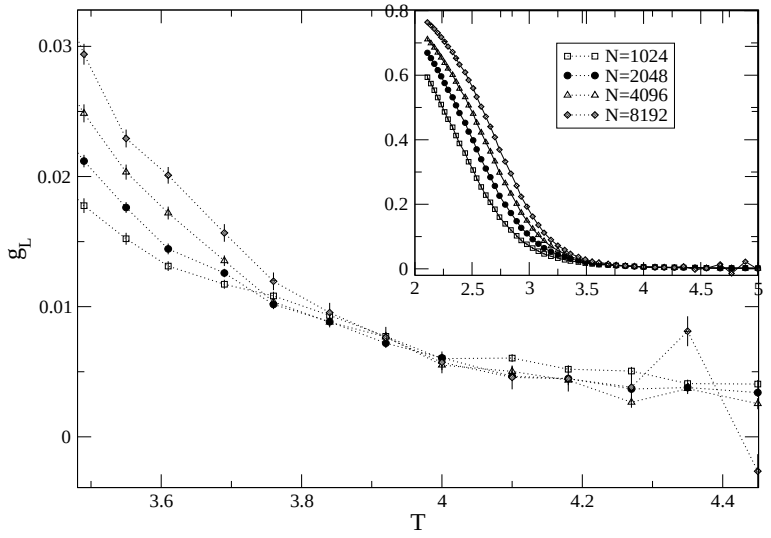

Figure 4.8: Scaling behaviour of the Binder cumulant, g_L, for different system sizes. Each system size is averaged over 1000 realizations of the network configuration.

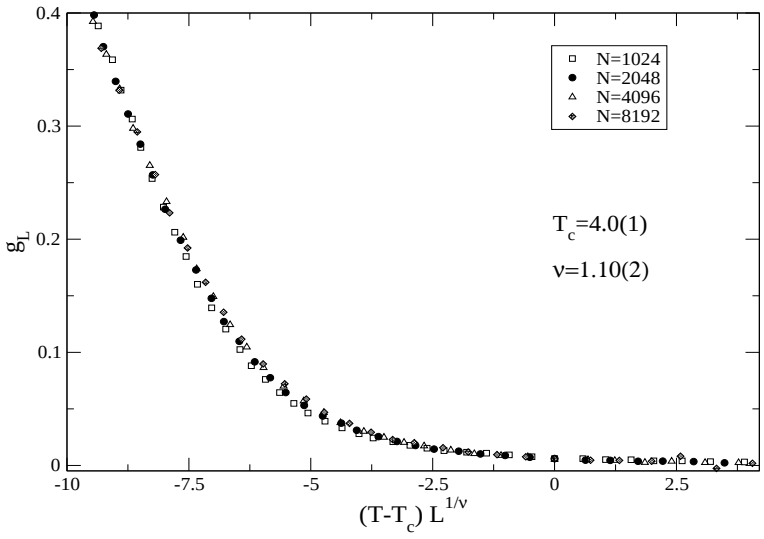

Figure 4.9: Scaling plot of the data illustrated in Fig. (4.8), fitted to Eq. 4.17.

therefore we take $L = \log(N)$. This slow increase in the diameter of the system, as well as the average path length, is a manifestation of the "small-world" property of this network, induced by the presence of a large number of highly connected hubs which create shortcuts between the nodes. An important implication of this feature is that we cannot embed the network in any finite dimensional lattice: we are implicitly dealing with a high dimensional system. The correlation length, in this case, is still well defined although its value gets close to the densely-connected, mean field limit as we increase the average connectivity of the nodes, $\langle k \rangle = 2m$.

The parameters T_c and ν are determined by constraining the temperature dependence of the Binder parameter for each network size to lie on a single curve. The curves following the scaling bahaviour of Eq. (4.17) are shown in Fig. 4.9. From this fit we estimate the critical temperature $T_c \sim 4.0(1)$ and the exponent of the SG correlation length $\nu \sim 1.10(2)$. It is important to underline that this kind of behaviour is not observed for an AF system on a regular triangular lattice, where frustration alone is present.

4.4 Stochastic model of opinion formation on a scale-free network

In the previous sections we analyzed the bahaviour of the antiferromagnetic Ising model on the scale-free topology generated by the BA algorithm with tunable clustering. Here the network, describing the interactions among the spins, displays a power law bahaviour in the connectivity of the nodes, $P(k) \sim k^{-\alpha}$, with $\alpha \sim 3$, along with a clustering coefficient that is independent on the number of nodes.

We now consider the effect of the scale-free topology generated by the same algorithm on a model of stochastic opinion formation. In this example each of the N nodes of the network, representing a person acting in a determined social group, is characterized by a spin value, $\sigma = \pm 1$, which mimics one of the two possible opinions that an agent can take. We neglect, in the first approximation, the network dynamics. This is equivalent to assuming that the time scale for evolving the network is much longer that the time needed for people to make a decision.

The dynamics of the spins, randomly assigned at the beginning of the simulation, follows a stochastic process that mimics the human uncertainty in decision making [KHH02, BT04], similarly to the spin dynamics in the CA model studied in chapter 3. The values are updated synchronously according to a local probabilistic rule: $\sigma_i(t+1) = +1$ with probability p_i and $\sigma_i(t+1) = -1$ with probability $1 - p_i$. The probability p_i is determined, by analogy with heat bath dynamics with formal temperature $k_B T = 1$, by

$$p_i(t) = \frac{1}{1 + e^{-2I_i(t)}}, \tag{4.18}$$

where the local field, $I_i(t)$, is

$$I_i(t) = a\xi(t)\tilde{N}_i^{-1} \sum_{j=1}^{\tilde{N}_i} \sigma_j(t) + h_i \eta_i(t) x(t). \tag{4.19}$$

The first term on the right-hand side of Eq. (4.19) represents the time dependent interaction strengths between the node i and his/her $\tilde{N}_i$ information sources, which are the first neighbours in the network. The second term instead reflects the personal reaction to the system feedback, that is the average opinion,

$$x(t) = \frac{1}{N} \sum_{j=1}^{N} \sigma_j(t), \tag{4.20}$$

resulting from the previous time step. The terms $\xi(t)$ and $\eta_i(t)$ are random variables uniformly distributed in the interval (-1,1) with no correlation in time nor in the network. They represent the conviction, at time t, with which agent i responds to his/her group (common for all the agents) and the global opinion of the network respectively. The strength term, a, is constant and common for the whole network, while h_i is specifically chosen for every individual from a uniform distribution in $(0,\kappa)$ and are both constant in the dynamics of the system. By varying the parameter κ we can give more or less weight to the role of feedback in the model. The strength coefficients a and h_i in the local field, I_i, characterizing the attributes of the agents, play a key role in the dynamics of the model. They represent the relative importance that each agent of the network gives, respectively, to his/her group and to the variation of the average opinion itself.

4.4.1 Numerical simulations

At first we investigate the importance of the group strength a by fixing $\kappa = a$. In this case the dynamical behaviour is similar to that found in the stock market context in Refs. [Kai00, KHH02, BT04]. For $a \lesssim 1$ the resulting time series of average opinion is largely uncorrelated Gaussian noise with no particularly interesting features, as illustrated in Fig. 4.10(i).

As soon as we exceed the value of $a \approx 1$ a turbulent-like regime sets in, characterized by large intermittent fluctuations, as illustrated in Fig. 4.10(ii $\rightarrow$ iv). These large fluctuations, or *coherent events*, can be interpreted in terms of a multiplicative stochastic process with a weak additive noise background [Nak98, KHH02]. For $a \gtrsim 2.7$ we observe that the bursts of the time series begin to saturate the bounds $-1 \leq r \leq 1$.

In Fig. 4.11 we plot the probability distribution functions (PDFs), P, associated with the time series of Fig. 4.10. The large fluctuations, for a greater than ≈ 1, are reflected in the fat tails of the relative PDFs. Decreasing the value of a, and so the number of coherent events, the PDF converges to a Gaussian distribution generated by a random Poisson process.

The personal response to the change in the average opinion also plays an important role in the turbulent-like regime of the simulation. In order to study the impact of this term on the dynamics we change the parameter κ while keeping a fixed at 1.8. The results are summarized by the PDF plots in Fig. 4.12. For $\kappa = 0$ the behaviour of the time series is still turbulent-like, which makes it clear how the network group interaction is, in reality, the crucial factor for the appearance of coherent events. As expected, incrementing the value of κ leads to a progressive crossover toward a noise regime. It is important to notice how

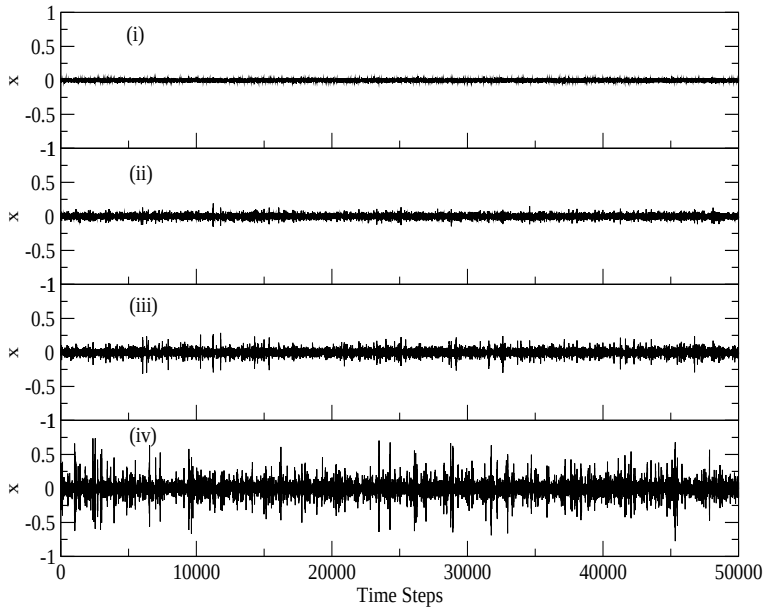

Figure 4.10: Time series of the average opinion, x, for different values of the group interaction strength parameter a: (i) $a = 0.8$, (ii) $a = 1.5$, (iii) $a = 1.8$ and (iv) $a = 2.3$. The parameters used for the simulations are: $N = 10^4$ nodes, clustering probability $\theta = 0.9$, initial nodes and links per new node $m_0 = m = 5$ and we take the upper bound of the distribution of personal response strengths equal to the group interaction strength, that is $\kappa = a$. The results involve 10 realizations of the scale free network each displayed for 5000 time steps. For values of a greater than 1 a turbulent-like state, characterized by large fluctuations, starts to appear in the process of opinion formation. The clustering probability $\theta = 0.9$, related to the triad formation in the network, fixes the clustering coefficient to $C \approx 0.39$. This value is similar to that found for many real systems [AB02, DM02].

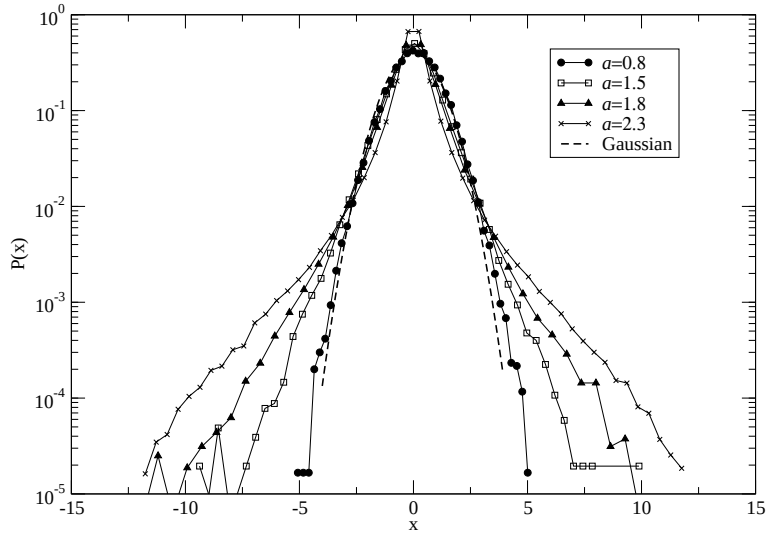

Figure 4.11: PDFs of the time series relative to Fig. 4.10. The shapes of the distributions converge to a Gaussian for small values of the group interaction strength $a = \kappa$. A Gaussian distribution is also plotted for comparison. All the PDFs in this paper are obtained by averaging over 50 realizations of the SFN. In order to compare the fluctuations at different scales, the time series in the plot have been normalized in the usual way according to Eq. (2.2).

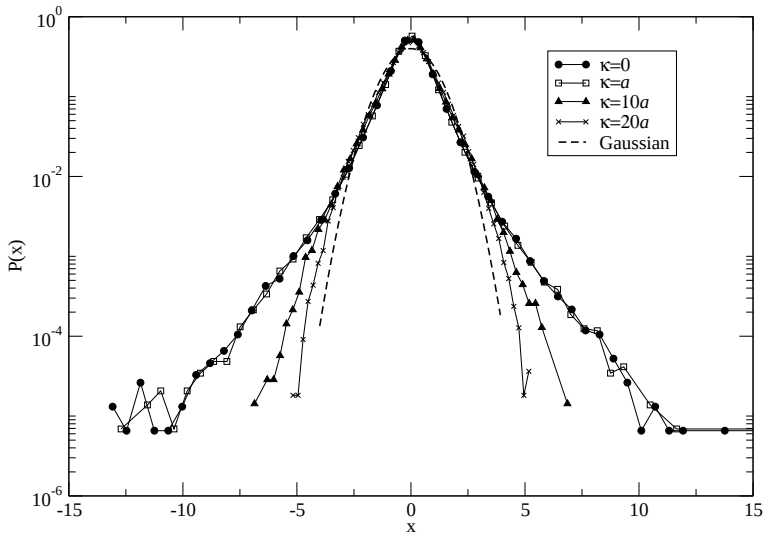

Figure 4.12: The importance of the personal response, related to the global opinion strength parameter κ is shown by the change of shapes of the PDFs for group interaction strength $a = 1.8$. For large values of κ the time series of global opinion approaches Gaussian noise. The time series of x has been normalized – see the caption of Fig. 4.11.

this regime is reached for $\kappa > 10a$. The group interactions continue to play an important role even when the average value of h_i is large compared to a.

In order to test the relevance of the network structure on the process of opinion formation, the previous simulations have been repeated, with a large number of nodes, N, and $\kappa = a$, for different values of the clustering parameter, θ, and the node-edge parameter, m. While varying θ, does not lead to any substantial difference in the dynamics of the model, the increase of the average number of links per node, $\bar{k} = 2m$, has a dramatic effect in the turbulent-like phase, as shown in Fig. 4.13. Here the kurtosis, ($K_x = \langle x^4 \rangle / \langle x^2 \rangle^2$, where $\langle \ldots \rangle$ denotes the temporal average), of the time series of the average opinion, used to quantify the deviation from a noise regime, is plotted against m. It is evident that an increase in the average number of links per node gives rise to more turbulence characterized by larger fluctuations and broader tails in the PDF. Large scale synchronizations are more likely to occur for large m. This behaviour is intrinsically related to the model of Eqs. (4.18) and (4.19). In fact, the turbulent-like regime is a consequence of the random fluctuations of the interaction strengths between agents around a bifurcation value separating the ordered and disordered phase.

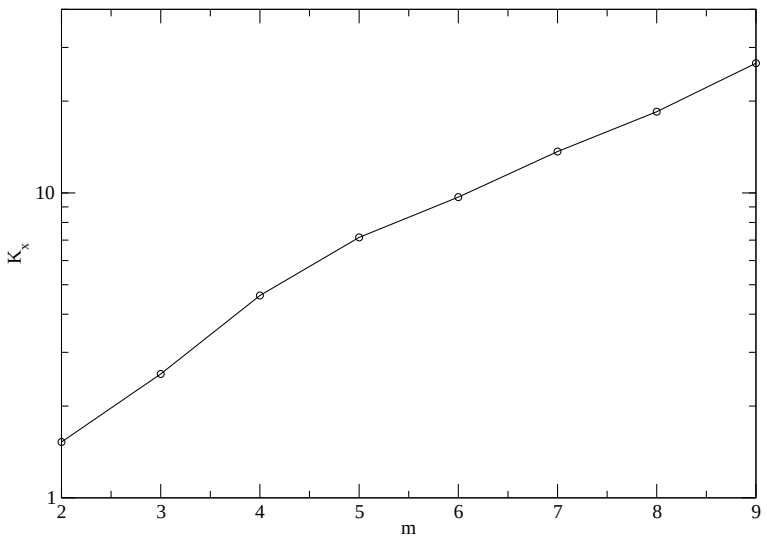

Figure 4.13: Dependence of the kurtosis, defined as $K_x = \langle x^4 \rangle / \langle x^2 \rangle^2$, where $\langle \dots \rangle$ denotes the temporal average, as a function of the node-edge parameter m. For a Gaussian noise process $K_x = 3$ while for $K_x > 3$ large deviations from the average start to appear. The final value for each m has been obtain after the average over 50 configurations of the network. The calculations show an exponential increase for K_x.

If we consider the case where $m \to \infty$ along with the thermodynamic limit, $N \to \infty$, then the coupling strengths between agents can be approximated well by the average strength over all the network and a mean field approach becomes appropriate to describe the dynamics of the model. Krawiecki *et al.* [KHH02] proposed the following map

$$x(t+1) = A\xi(t)x(t) + h\eta(t), \tag{4.21}$$

as a mean field approximation of a stochastic dynamical system similar to the one used in the present work. Here A and h are coupling coefficients and $\xi(t)$ and $\eta(t)$ random numbers in the interval (-1,1). The map of Eq. (4.21) is a generic model for *on-off intermittency* and *attractor bubbling* extensively studied in chaos theory [PST93, ABS94, PHH94, OS94, AS97].

It is also worth pointing out that an increase of $\bar{k}$ is related to a decrease in the average path length between nodes; that is, the network "shrinks" and becomes more compact. In relation to our previous discussion, the more compact the network is the more the dynamics of our system corresponds to that of the mean field approximation. It becomes easier for the agents to synchronize. This characteristic of compactness, referred to as the *small world effect* [Bol85, AB02, DM02], is actually very common in both real and artificial networks.

4.4.2 Comparison with random networks

We further investigate the importance of the SFN topology and the the small world effect in our model by performing a numerical simulation of the same system but using a *random network* (RN) or *random graph* as the underlying topology. Given a fixed a number of nodes, N, a RN is defined by the probability p that two nodes are linked together [Bol85, AB02, DM02]. In this case $\bar{k} = pN$ and, moreover, there exists a critical value, $p \equiv p_c \simeq 1/N$, for which the the network undergoes a topological phase transition where it moves from a phase where it is composed of a collection of small, disjoint, sub-networks to a phase where a giant cluster emerges [3]. Random networks, while preserving small world properties, have a Poisson degree distribution [AB02], $P(k)$, and small clustering coefficients. As previously mentioned, we make use of the RN to test the robustness of our model with respect to the topology used and to learn about the most important properties relevant to the dynamics. In order to do so, we fix the number of agents and the average number of links for the

[3] Note the analogy between the random network theory and the standard percolation theory on a lattice [Sta85] where the structural properties of the system are studied as a function of the percolation probability.

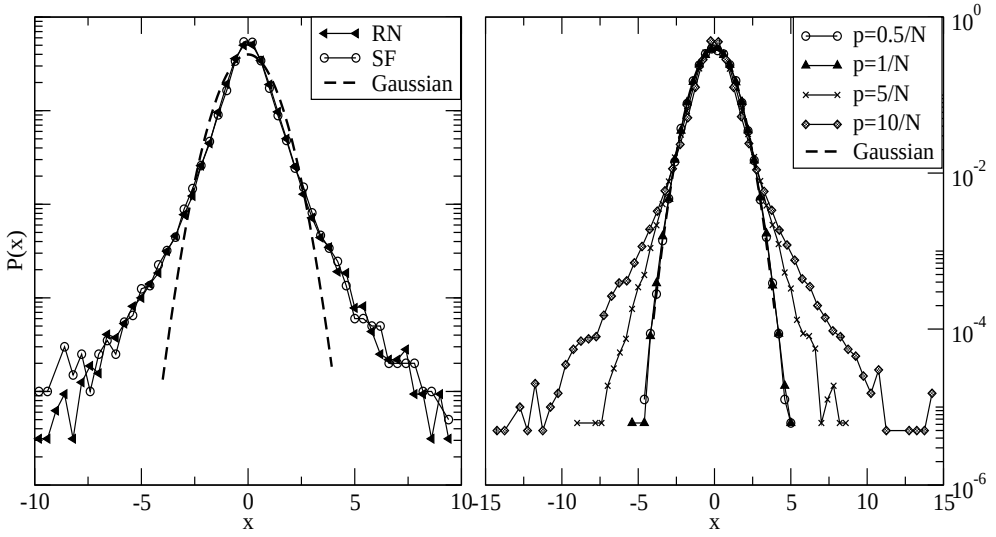

Figure 4.14: Left: comparison between the PDFs of our model obtained on a SFN and on a RN with number of nodes $N = 10^4$ and average links per node $\bar{k} = 10$. For the SFN the parameters used are $m = m_0 = 5$ for the links of each new node and $\theta = 0.9$ for the clustering probability while for the RN $p = 10/N$. From a statistical point of view the characteristic features of the PDFs have their origin in the model dynamics as opposed to the fine features of the network. Right: Dependence of the opinion fluctuations on the parameter p on a RN. The parameters used for the dynamics are $a = 1.8$ and $\kappa = a$ for the group and global opinion response respectively.

SFN and RN, namely $N = 10^4$ and $\bar{k} = 10$. Then we perform independent numerical simulations on the two topologies. Note that for the RN, $\bar{k} = 10$ requires $p = 10/N$, that is ten times greater than the percolation threshold. The results, shown in Fig. 4.14 (left) demonstrate how the dynamics of the two systems are largely equivalent under the adopted constraints. In Fig. 4.14 (right) we also show the dependence of the dynamics on the parameter p for the RN. At the critical threshold, that is the value of p for which a giant cluster appears, there is still no trace of turbulent-like activity giving rise to fat tails. Yet, in this case each agent has, on average, just one link and there cannot be any small world properties.

These results confirm that the critical topological characteristic leading to herding behaviour in the framework of stochastic opinion formation is the presence of mean field effects enhanced by small-world structure. The more infor-

mation (links) that an agent has, the more likely it is for him/her to have an opinion in accord with other agents.

In the following we extend our model in order to include indecision in the process of opinion formation.

4.4.3 The influence of indecision

We now extend our model in order to include the concept of indecision. In practice a certain agent i, at a time step t, may take neither of the two possible decisions, $\sigma_i = \pm 1$, but remain in a neutral state. Keeping faith to the spirit of the model, we address this problem introducing an *indecision probability*, ϵ: that is the probability to find, at each time step, a certain agent undecided. This is equivalent to introducing time dependent failures in the structure of the network by setting $\sigma = 0$.

Focusing on the turbulent-like regime, the shape of the PDF in the opinion fluctuations changes according to different concentrations of undecided persons. The results of the simulations, in Fig. 4.15, show how the dynamics of the model move from an intermittent state for $\epsilon = 0$ toward a noise state for $\epsilon \approx 0.6$. The convergence to a Gaussian distribution is obtained only for quite high concentrations of undecided agents at about 60%. The robustness of the turbulent-like behaviour is related to the intrinsic robustness of SFN against random failures [AJB00, CEbAH00, CNSW00]. In fact, because there is a large absolute number of poorly connected nodes, related to the power law shape of $P(k)$, the probability of setting one of them to inactive is much higher compared to the "hubs" that are relatively rare.

We can claim that, in large social networks governed by stochastic reactions in their elements, large fluctuations in the average opinion can appear even in the case in which a large part of the network is actually "inactive" provided that the structure is scale-free and the indecision is randomly distributed. The existence of large hubs provides for the survival of extended sub-networks in which synchronization can give rise to coherent events. The structure of the network itself supplies the random indecision.

Now we address the question of how the dynamics may change if we do not choose randomly the inactive nodes but we target the nodes having the most links. What we do in practice is to sort the nodes according to their number of links and then deactivate the ones having the largest number of links in decreasing order. Fig. 4.16 illustrates how the fragmentation process is much faster and the noise regime is reached already when only the 10% of the hubs are deactivated. As emphasized in Ref. [AJB00, CEbAH00, CNSW00], the hubs have a great importance in the structural properties of SFNs and specifically

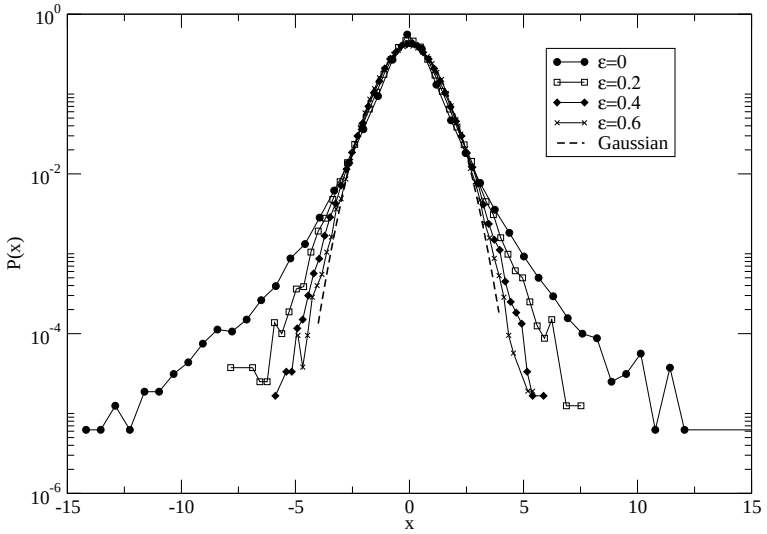

Figure 4.15: Transition from coherent bahaviour, indecision probability $\epsilon = 0$, to noise using a random selection for the inactive agents. For $\epsilon \approx 0.6$ we reach a noise-like behaviour. The parameters used in the simulation are: $N = 10^4$ nodes, $\theta = 0.9$ for the clustering probability, $m = m_0 = 5$ for the links of each new node, $a = 1.8$ and $\kappa = a$ for the group and global opinion response respectively.

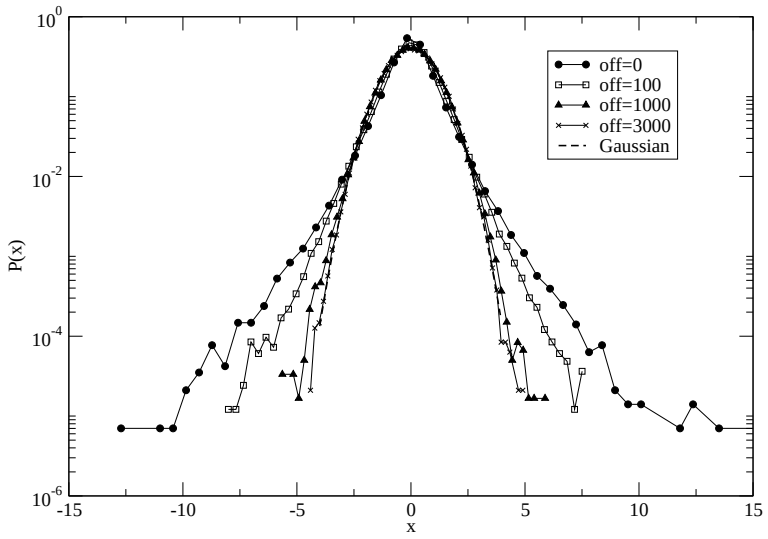

Figure 4.16: In this simulation we progressively turn off the largest hubs in the network. Once we have turned off about the 10% of agents, $N = 10^4$, the coherence in opinion formation disappears. The parameters used in the simulations are the same as in Fig. 4.15.

targeting these vertices can lead to sudden isolation of a large fraction of the nodes of the network.

4.4.4 Agent induced indecision: the three state model

In the previous section we introduced random and targeted failures in order to study the response of the system to changes in the network topology. In a real social network the reason behind the indecision of a person follows much more complex rules and can depend on different factors as, for example, unsatisfactory information obtained by his/her sources. As seen from Eq. (4.19), the opinion of each agent depends on the poll of his/her network links. Suppose now that the agent i has $\tilde{N}_i$ neighbours where $\tilde{N}_i/2$ of these share the opinion $+1$ while the remaining $\tilde{N}_i/2$ share the opposite opinion. In this case, unless we give specific weights to each node, the agent i will not have an easy task in choosing one of the two possible positions because of a lack of popular consensus. Based on this idea derived from common sense, we can extend our two state model by introducing an *induced indecision probability*, μ, dependent on the information available to the agents at each time step. In particular we define the global opinion of the neighbours of the i^{th} node as $s_i(t) = \sum_{j=1}^{\tilde{N}_i} \sigma_j(t)$ and the indecision probability

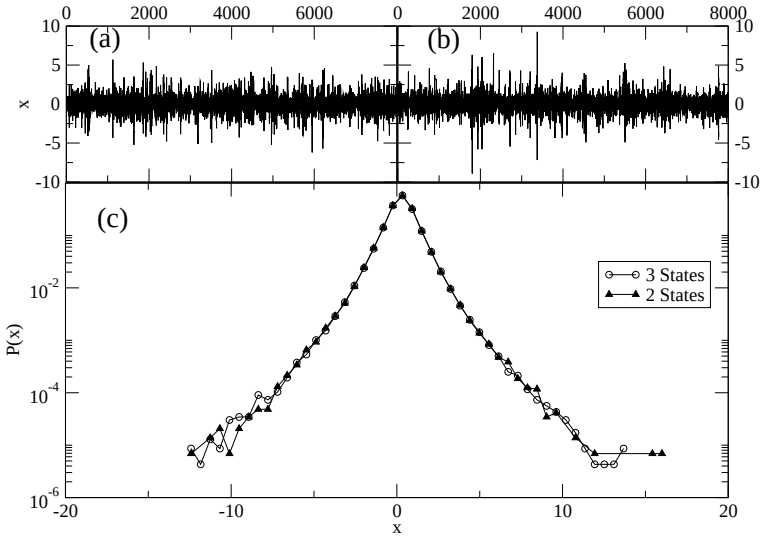

Figure 4.17: (a) A window of the normalized time series generated by the two-state model with parameters $N = 10^4$ nodes, $\theta = 0.9$ for the clustering probability, $m = m_0 = 5$ for the links of each new node, $a = 1.8$ and $\kappa = a$ for the group and global opinion response respectively. (b) Window of the normalized time series generated by the three states model with the same parameters as (a) and indecision probability width $\varsigma = 1$. (c) Comparison between the PDFs generated by the two and three-state models with the aforementioned parameters obtained over 50 realizations of the SFN. No relevant differences can be observed.

for the i^{th} node at time t

$$\mu_i(s,t) = c_i \; e^{-s_i^2(t)/2\varsigma}, \tag{4.22}$$

where the indecision probability width, ς, is a parameter of the model and c_i a normalization constant that depends just on the structure of the network. The latter is calculated at the beginning of the simulation by imposing $\sum_{s=-\tilde{N}_i}^{\tilde{N}_i} \mu_i(s,0) = 1$, i.e. the sum of the indecision probabilities over all possible global opinions is one. The model of Eq. (4.22) assumes a Gaussian probability, centered on $\sigma_i = 0$, for the distribution of indecision of the i^{th} agent. That is, the probability of having this agent in a state with $\sigma_i = 0$ is greater when there is not a large agreement in the opinion of the his/her sources.

The analysis of the time series generated by the three state model does not present any relevant difference if compared with the two state model with the same parameters, Fig. 4.17.

We also plot the PDF for the number of inactive agents, $N_s(t)$, during the simulation, Fig. 4.18. It is interesting to notice how this distribution is not

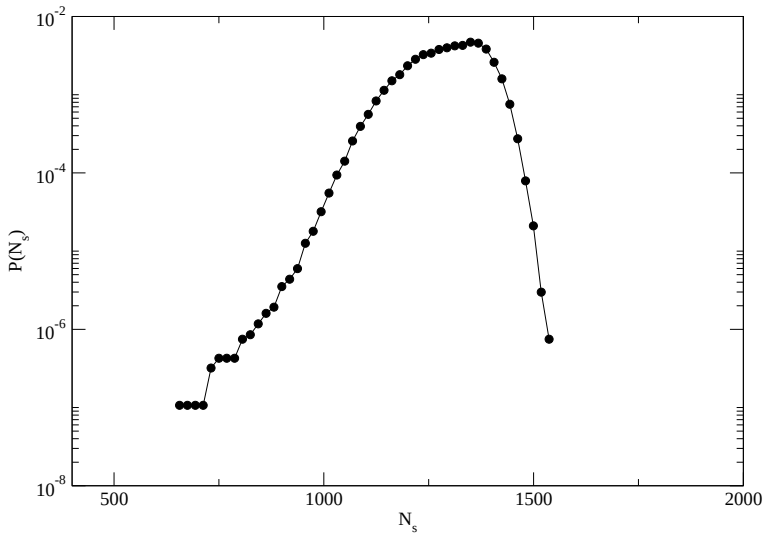

Figure 4.18: PDF of the number of inactive agents, $\sigma_i(t) = 0$, during the simulation of the three state model. The parameters used are the same as used in Fig. 4.17.

Gaussian distributed around the average but it is skewed on one side. Moreover, only a small fraction of agents is undecided, of the order of 10/15 %. This is consistent with the observation that in opinion polls most of the participants actually indicate an opinion.

4.4.5 Possible application: opinion formation and the stock market

The model for opinion formation discussed thus far can be tested against the best known real social network: the stock market. In this specific case the spin value corresponds to the will of an agent to by or sell a stock. The main idea is to compare our results with some *stylized facts* concerning the price time series, $P_r(t)$ and, in particular, with the properties of the logarithm of the price fluctuations, or *returns*, which we recall to be defined as $r(t) = \ln P_r(t+1) - \ln P_r(t)$. As we already observed in the previous chapters, the returns show features that are independent of the particular market and can be considered as universal [MS99] and their intermittent behaviour resembles, at least in some aspects, the hydrodynamic turbulence [MS99, MS95, GBP+96, MS97], also characterized by power law tails in the PDF. In this case the large coherent events are related to crashes

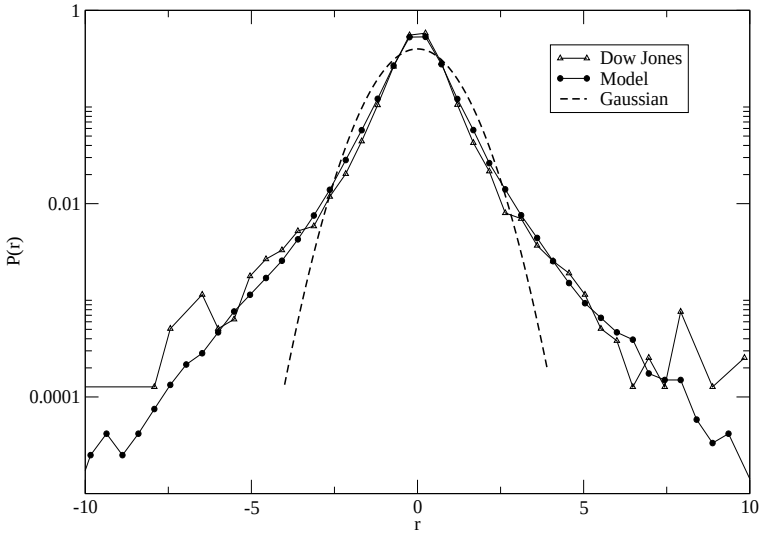

Figure 4.19: Comparison between the PDF of our model and the time series of the Dow Jones index from 13/1/1930 to 13/4/2004. The parameters of the model used to reproduce the PDF in the plot are: $N = 10^4$ nodes, $\theta = 0.9$ for the clustering probability, $m = m_0 = 5$ for links of each new node, $a = 1.8$ and $\kappa = a$ for the group and global opinion response respectively. A Gaussian is also superimposed in order to emphasize the fat tails.

or other anomalous variations of price.

Considering the assumption of Eq. (3.3) still valid, the returns are proportional to the average opinion $r(t) \approx x(t)$ and we compare the time series of average opinion generated by the two state model against the time series of daily closures of the Dow Jones index. The data set spans the range 13/1/1930 to 13/4/2004 for a total of 18645 samples. In Fig. 4.19, a comparison between the two PDFs is shown. The similarities between the model and the Dow Jones is remarkable. Both distributions have a *leptokurtic* shape and, in particular, they are described by power law tails, expressing the turbulent-like dynamics of the time series[4]. Note that, in contrast to the self-organized model for stock market dynamics proposed by Bak *et al.* [BPS97], here the price feedback is not an essential ingredient for the reproduction of the correct shape of the distribution. Rather it is the herding behaviour that plays the main role, as observed from Fig. 4.12.

[4]The problem of the actual shape of the PDF for the stock market returns is still a matter of debate in the econophysics community [MS99, RRN+01, GC02, MJ03]. A solution to this problem would be of a great interest, especially for the practical application of option pricing.

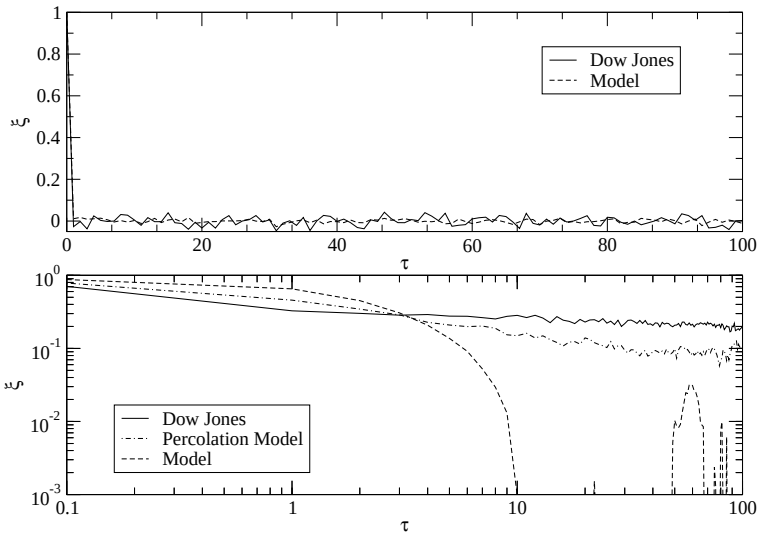

Figure 4.20: Autocorrelation functions for the fluctuations $x(t)$ (top) and the volatility $v(t)$ (bottom). The parameters used to produce the analyzed set are: $N = 10^4$ nodes, $\theta = 0.9$ for the clustering probability, $m = m_0 = 5$ for the links of each new node, $a = 1.8$ and $\kappa = a$ for the group and global opinion response respectively.

The similarities between the artificial time series generated by the virtual social network and the stock market extend beyond the fat tails in the PDF of the fluctuations to temporal correlations. As already done in Chapter 3, we calculate the autocorrelation functions, Eq. (3.5), both for the returns and for the volatility, Eq. (2.7). While the time series of returns generated by the model and the Dow Jones index have an equivalent behaviour, Fig. 4.20 (top), the same similarities do not hold for the volatility, Fig. 4.20 (bottom). We observe a qualitatively different correlation: while for the market we observe a power law behaviour[5], the memory in the time series generated by the model decays exponentially like a short-range correlated random processes [MS99]. This second point illustrates how non-trivial memory effects in the stock market cannot be taken into account by a simple heath bath dynamics.

In Fig. 4.20 (bottom) we also reproduce the autocorrelation function for the CA model presented in Chapter 3. In this model a heat bath dynamics, similar to the one used in the present simulations, is applied to dynamical percolation clusters, used as a paradigm for agent aggregation. The temporal evolution of

[5]In Chapter 3 we already came across the well known phenomenon of *volatility clustering*, which relates to the persisting correlation for the volatility of daily data [MS99].

the clusters, the size of which follows a power law distribution, is related to a forest-fire dynamics in which some potential traders are attracted in the market by other already active traders while, at the same time, some of them may temporarily quit the trading. Large fluctuations in the price changes are due to the synchronization in the of the larger clusters in the market at a particular time. The main qualitative difference between this model and the one presented so far is that the former presents a decay rate much closer to that of the real market. At this point it is important to underline that the main difference between the two models is related to the network dynamics. While in the present simulation the network is fixed, in the CA the interaction between agents is time dependent and localized in separate clusters. We can argue that the dynamics of the networks and, in particular, the clustering of agents in different sub-networks can play an important role in the correlation properties of the stock market volatility. In reality, this fact appears quite natural if we use the autocorrelation function, defined in Eq. (3.5), in order to estimate the degree of memory in a process. If, for example, the variable under investigation is the sum over many independent Markovian processes, as in the CA case, then the resulting autocorrelation is given the convolution of the common exponential decay, $\propto e^{-\beta\tau}$, with the distribution of the decay rates, $g(\beta)$,

$$\xi(\tau) \propto \int_0^\infty g(\beta) e^{-\beta\tau} d\beta. \tag{4.23}$$

According to the shape of this distribution, the observed macroscopic variable can show a behaviour characteristic of a long memory processes, like the $1/f$ Fourier spectrum [VdZ50]. Power law tails in the autocorrelation, $\xi(\tau) \propto \tau^{-\gamma}$, are produced from the distribution $g(\beta) = \Gamma(\gamma)^{-1}\beta^{\gamma-1}$, where Γ is the gamma function and γ a generic real exponent [Sor04]. This fact strengthens the idea that the stock market is organized in a hierarchy of sub-networks where each of them can be considered, from a physical point of view, at local equilibrium. For time periods shorter than the typical time scale necessary for the networks to evolve, the only link between the sub-systems composing the market is the feedback coming from the price history. This idea is closely related to the concept of *subordination* used in probability theory [Fel68]. The superposition of distributions, as a possible explanation of fat-tailed processes, has been proposed recently by Beck [Bec01, Bec03] in the context of hydrodynamic turbulence and then extended also to other systems [Bec05] including the stock market [AI03, KF03].

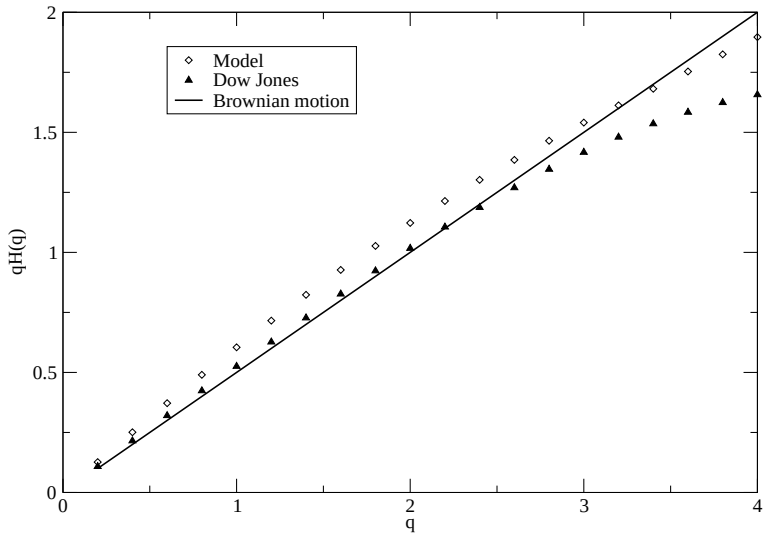

Figure 4.21: Structure function exponents for the Dow Jones index and our model. A deviation from a linear behaviour is evident. The hypothetical spectrum of a 1D Brownian motion is also shown for comparison.

4.4.6 Multifractal analysis

In order to further test the model, we calculate the multifractal characteristics of the time series generated by our model, along with the Dow Jones index, by calculating the spectrum of Hurst exponents, Eq. (3.6), in the same fashion as in Sec. 3.4, where we refer the reader for a broader discussion on this quantity.

The results, shown in Fig. 4.21, underline a multifractal structure for both the time series, which cannot be associated with a simple random walk as imposed by the classical *efficient market hypothesis* [Bac00].

4.5 Discussion and conclusion

In relation with the availability of larger databases and faster computers, more and more studies have been devoted to the empirical analysis of the characteristic topological structures of complex systems. Despite the different Nature of the subjects under examination, it has been found that the networks describing the interactions been the various elementary constituents, show a remarkable homogeneity in their statistical features. In particular a robust power law distribution in the connectivity of the nodes, along with a high clustering, has been pointed out. These empirical findings, apart from bringing structural self-

organization problems under a different prospective, open an entirely new area of investigation in computational physics where most of the numerical work in literature has been developed on regular lattices. Not much is known regarding the behaviour of even the well-established models on complex topologies, such as SFNs.

In the present chapter we have investigated the role played by these networks, and in particular the one generated via the Barabási-Albert algorithm, by studying their impact on the dynamics of two different models, namely the AF Ising and a model for stochastic opinion formation.

In the AF Ising model the calculation of the overlap parameter reveals a spin glass behaviour at low temperatures, produced by the random frustration of scale-free network. The critical temperature separating the SG and the the paramagnetic phases is found to be $T_c = 4.0(2)$. Such behaviour is not observed for the same model on regular triangular lattices. Hence the topology of the interactions plays a critical role in the dynamics of this system.

In the model for opinion formation the choice of a SFN is further motivated by a series of recent studies on social aggregation [AB02, DM02]. In this case the results of the numerical simulations show that for a certain range in the parameter space the fluctuations of opinion have a non-trivial turbulent-like dynamics determined by the synchronization of large parts of the network.

Also in this case the SFN topology plays a key part in the dynamics of the model. In fact, introducing inactive agents and spreading the undecided agents randomly on the network, does not spoil the turbulent-like state even for high concentrations of "gaps", up to approximately 60% of agents. This is a consequence of the implicit robustness of SFNs against random failures. If instead of selecting randomly the undecided individuals we aim directly to the "hubs" of the network the situation changes. In this case the network is disaggregate, composed of very small sub-networks and isolated nodes. Synchronization cannot significantly effect the resulting global opinion and the time series approximates Gaussian noise.

The evolution of opinion formation generated by the model in the turbulent-like regime has also been tested against a time series of daily closures for the Dow Jones index. The stock market, in fact, can be considered as the most studied network of social interactions. The results show a very good agreement with some stylized facts of the financial market like the broad tails in the PDFs, temporal correlations and a multifractal spectrum. We also notice an interesting discrepancy in the autocorrelation function for the volatility. Comparing the present results with those obtained in Chapter 3, we conjecture that the persistence in the volatility memory can be explained by considering the market as constituted by sub-systems at local equilibrium and weakly interacting with

each other.

In conclusion, we have seen in this chapter how complex systems tend to self-organize in networks with specific connectivities and how the structures influence the dynamics of theoretical models. Although regular lattices present numerous practical advantages and must be considered as the starting point for every model, we must keep in mind that the world around us is not as simple as we would like it to be: sometimes the structure of the elementary interactions can be very complex and cannot be ignored.

5

Symbiosis in the Bak-Sneppen model for biological evolution

In the introduction of this thesis, Sec. 1.3.3, we have seen that changes in ecosystems are everything but a smooth phenomenon. Geological eras in which almost nothing happens are suddenly interrupted by periods of high activity where the ecology goes through very rapid modifications which can be related with the disappearance of some species or with a mutation in their genetic code. Despite the fact that the mass extinctions observed in the fossil records can be related to exogenous events, the entangled evolution of species in an ecology is probably one of the most fascinating examples of complex dynamics.

In this chapter we extend, after a brief review, one of the most popular physical models on the subject: the Bak-Sneppen [BS93] model. In the new version that we propose, the Local Interactions Bak-Sneppen (LIBS) model, we introduce explicit local interactions between species of the ecology. This "environmental" perturbation modifies the intrinsic fitness of each element of the ecology, leading to higher survival probability, even for the less fit. While the system still self-organizes toward a critical state, the distribution of fitness broadens, losing the classical step-function shape. The LIBS model finds application also to the economic world where the distribution of global-fitness can be related to the size distribution of firms in the most developed markets. In this respect the evolution of firms is seen as a punctuated equilibrium process in which the convolution of mutual interest can justify the spreading in size of the firms themselves. This chapter is based on the work published in Ref. [BLT06].

5.1 Introduction: the Bak-Sneppen model

Throughout the thesis we have seen that the presence of power laws is ubiquitous in Nature. In the late eighties, the theory of self-organized criticality (SOC) gave an attempt to explain such phenomenon via a mechanism of charge-discharge, triggered once a certain threshold is exceeded. In this framework, systems move from one metastable state to another by going through a series of avalanche-like events. As a consequence, in SOC at least two characteristic temporal scale

are involved in the dynamics: the characteristic time of the avalanches and the characteristic needed to reach the activation threshold [Jen98]. For references and a broader discussion on the subject we refer the reader to Sec. 2.2.

Among the various SOC models developed so far, the one proposed in 1993 by Bak and Sneppen [BS93] for biological evolution still enjoys a great popularity. The main idea behind this model is that each species can be uniquely characterized by a single parameter called *fitness*. The fitness of a species represents its degree of adaptation with respect to the external environment. Highly adapted species will hardly undergo any successful, spontaneous mutations. At the opposite end of the scale, if a species has a very low degree of fitness it needs to mutate in order to survive and its mutation automatically influences the other species belonging to the same environment. These concepts can be easily formulated as a simple 1D model. Suppose that the ecology can be represented by a periodic array of N cells and to each cell, i, is assigned a fitness, $B(i)$, taken from a uniform distribution between 0 and 1. Once we have fixed the initial condition, for each discrete time-step, the dynamical evolution of the system works as follows:

a) locate the species with minimum fitness – that is, the one most likely to mutate, i_{min},

b) change the fitness of i_{min} and that of its neighbours (species related) according to

$$\begin{aligned} B(i_{min}-1) &\to u_1, \\ B(i_{min}) &\to u_2, \\ B(i_{min}+1) &\to u_3, \end{aligned} \tag{5.1}$$

where the new fitness value, u_i, is a random number taken from a uniform distribution bounded between 0 and 1.

From numerical [BS93] and analytical [FBS93] studies it has been shown that the values of the fitness evolve to a step function, in the thermodynamic limit ($N \to \infty$), characterized by a single value, B_c. For $B < B_c$ the distribution of fitness, $P(B)$, is uniformly equal to zero while for $B > B_c$ we have $P(B) = 1/(1-B_c)$, determined by the normalization condition. An example is shown in Fig. 5.1 (a) and (b).

In this model it is also possible to define an intermittent dynamics, Fig. 5.2, that resembles closely the "punctuated equilibrium" observed in the fossil records [EG72], Fig. 1.3. The size of the avalanches is defined by fixing a threshold for the fitness, B_0 and by considering $B_m(t)$ as the minimal fitness at time step t. If at a certain time step, t_1, it happens that $B_m(t_1) < B_0$ then we can measure the interval of time, T_a, needed for having again $B_m(t_1 + T_a) > B_0$. In this case an

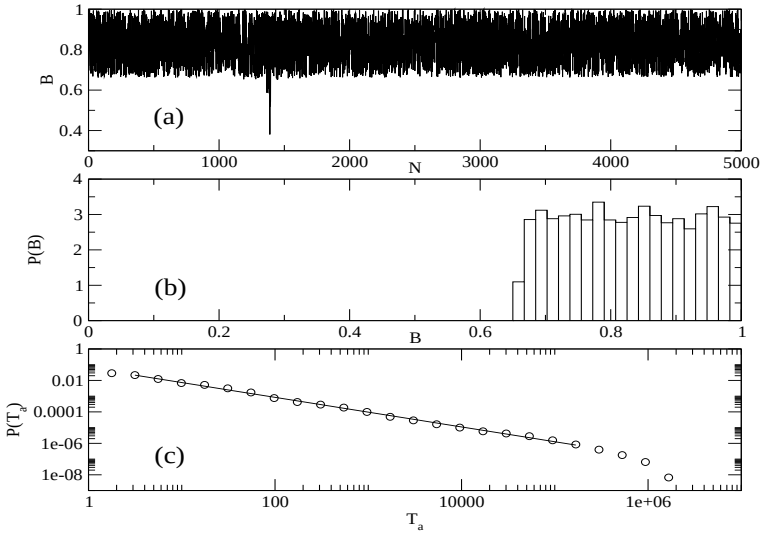

Figure 5.1: (a): Snapshot of the fitness, B, after $8 \cdot 10^7$ time steps in a stable configuration. All the values are above a critical threshold, except for those around 1500, where an avalanche is clearly visible. (b): Probability distribution of the fitness, $P(B)$, shown in (a). While the distribution is equal to zero on the left hand side, a plateau is evident for $B > B_c \sim 0.667$. In the thermodynamic limit the plateau will become exactly a constant. In this case the simulation has been carried out for $N = 5000$ species. (c): Probability distribution, $P(T_a)$, of avalanche duration, T_a, in the stationary regime for the Bak-Sneppen model. The avalanche time series has been recorded in the stationary state of the system, after $\sim 10^7$ iterations for $N = 2000$ species. The power law exponent is $\gamma \sim 1$.

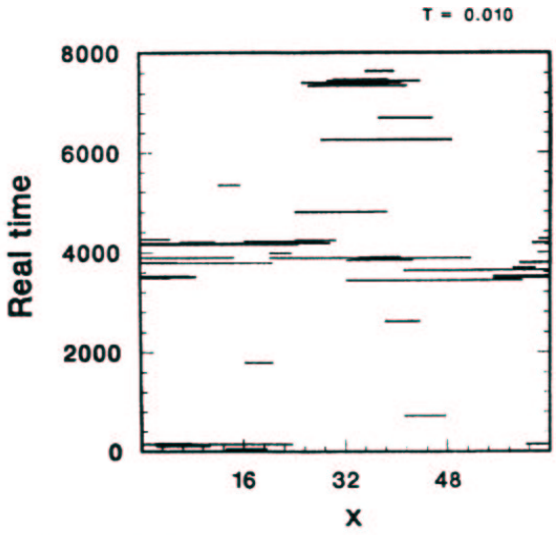

Figure 5.2: Mutation activity against *real time* for an ecology of 64 species (here denoted by x) taken by the original paper of Bak and Sneppen [BS93]. The real time is calculated by assuming that the time to mutation of a species with minimum barrier B_{min} is $\propto \exp(B_{min}/T_{car})$, where T_{car} ($T_{car} = 0.010$ in the plot) is a characteristic temporal scale. It is evident that avalanches of mutations are localized in particular periods of time, similarly to the "punctuated equilibrium" observed by Eldredge and Gould [EG72].

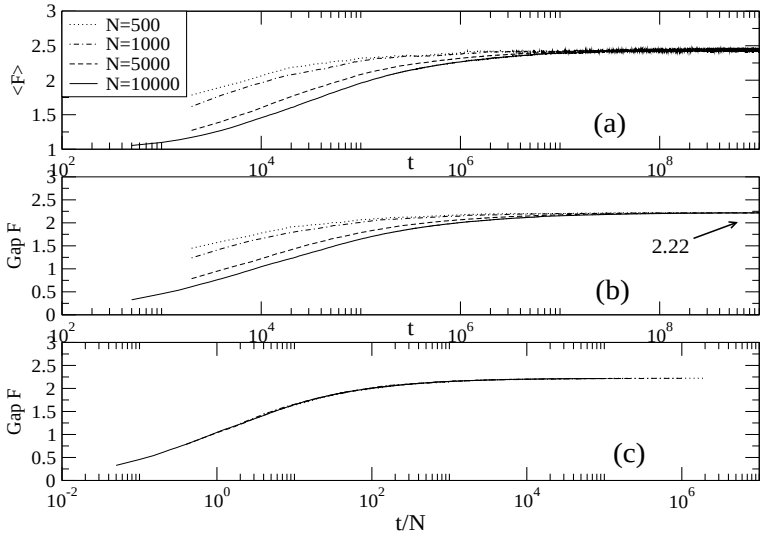

Figure 5.3: (a) Average value ($\langle \ldots \rangle$) and (b) gap (Gap) function for the global fitness F and different system sizes for simulations up to 10^9 mutations, t. Note that the gap function converges approximately to the value $F_c \sim 2.22$ that corresponds to the critical threshold for this model. A simple rescaling of the time, $t \to t/N$, collapses the curves onto a single curve as shown in (c).

avalanche of duration, or size, T_a has taken place in order to restore a minimal fitness in the system. If $B_0 = B_c$ then we have $P(T_a) \sim T_a^{-\gamma}$: the system is critical, see Fig. 5.1(c). Suggestions of a power law behaviour in the distribution of extinction size is supported also by empirical observations [NP99]. In this case the index of the power law is $\gamma \sim 2$.

In conclusion, according to the BS model the great mass extinctions of species, like dinosaurs for example, can be explained in terms of burst-like dynamics. A small perturbation in a critical self-connected system can trigger a chain reaction that may influence a great part of the species in the ecosystem. Moreover, time series of fossils samples seem to be in agreement with this avalanche dynamics in the extinction/evolution of species, as previously pointed out. A more detailed discussion of the BS model goes beyond the scope of this short introduction. For a general review see Ref. [PMB96].

5.2 Species living in symbiosis: the LIBS model

In the previous section we have seen that, despite its simplicity, the BS model evolves according to a complex dynamics and it is able to explain some empirical features of the biological evolution [BS93]. An implicit assumption in the model is that every species is deeply connected to its environment. A mutation on a single element automatically triggers a mutation in its neighbors.

But is this approximation always appropriate? Consider for example three species in a one dimensional array and suppose that $B_i = B_{min}$ while $B_{i-1} \cong B_{i+1} \cong 1$. In the standard BS model the i^{th} cell undergoes to mutation that also triggers a change in B_{i-1} and B_{i+1}. From a biological point of view it means that two extremely well adapted species have to mutate in order to cope with the mutation of the i^{th} species. This can be interpreted as a very particular (pessimistic) case – such as, for example, the case where the i^{th} species is the main source of food for both the other species.

In order to stress this idea we use some examples from different areas in which a similar evolutionary dynamics can be applied. Suppose that a new unfit or unskilled player joins a strong team. Will this player trigger a regression in the team performance or will the team compensate for this lack of skills? This is a small perturbation after all.

Another example comes from economics. In this case, it has been shown [VBT01, BCLM03, SFA04], that the dynamics of different firms is correlated. In fact, it is not unusual for a company to own large amounts of stock of other companies and so on. The result is an entangled environment, where the evolution of a firm is, in a way, linked to the evolution of its network of interaction. Is it then possible, in this case, for an wealthy environment to sustain an unfit element, or will its lack of "fitness" bring to the brink of the financial collapse all the other partners, as the BS model would suggest?

We provide an answer to these questions using a modified version of the BS model that takes into account the feedback of the environment on the single element. We refer to this model incorporating Local Interactions in the BS model as the LIBS model. For the sake of simplicity we do not consider the topology of the interaction, that may be very complex; rather, we use a simple 1D array. The influence of the network structure on the dynamics of the model will be briefly discussed in Sec. 5.5 although an exhaustive investigation is left for future work.

As a first approximation we consider our species to be arranged on a one dimensional array with nearest neighbor interactions. This means that the micro-environment is composed of three cells. The value of the fitness, B, for each cell

is taken, according to the BS model, from a uniform distribution between zero and one. The fitness parameter, B_i, of the i^{th} cell represents the *self-fitness* of the species. Motivated by the aforementioned examples, we add an environmental contribution to the self-fitness that leads to a *global-fitness*, F_i, according to

$$F_i = B_i + \Lambda_{i,i-1} B_{i-1} + \Lambda_{i,i+1} B_{i+1}, \tag{5.2}$$

where $\Lambda_{i,i-1}$ and $\Lambda_{i,i+1}$ are the fractions of fitness that the i^{th} cell shares with its neighbours. The matrix of Λs is not symmetric, reflecting the fact that the contribution in one direction can be very different that the contribution in the other. This is equivalent to considering a directed *weighted graph* with a trivial necklace topology.

In the sport example, the global fitness corresponds to the fit players that contribute to sustaining the unskilled team-mate. From the economic point of view it represents the capability of a firm to gain benefits from its partnerships with other firms. In this particular case, B_i represents the wealth generated by the firm itself, while the other two terms represent the contribution, in different forms, from the linked firms. In general, we can consider the new terms in the definition of F_i as short ranged random forces acting on the i^{th} cell.

At the beginning of the simulation the self-fitness is drawn from a uniform distribution between zero and one. The same is done for the link weights, Λ_{ij}. As explained above, in general, for two cells i and j, $\Lambda_{ij} \neq \Lambda_{ji}$.

Assuming that the neighbours can cooperate in defining the fitness of a species (optimistic view), the extremal dynamics is moved from B_{min} to F_{min}. Once the site with minimum global fitness, i_{min}, is located, then the self-fitness and the interactions of this species are redrawn according to the following rules:

$$\begin{aligned}
\Lambda_{i_{min},i_{min}-1} &\to u_1, \\
\Lambda_{i_{min}-1,i_{min}} &\to u_2, \\
B(i_{min}) &\to u_3, \\
\Lambda_{i_{min},i_{min}+1} &\to u_4, \\
\Lambda_{i_{min}+1,i_{min}} &\to u_5,
\end{aligned} \tag{5.3}$$

where the new values for the changed quantities are taken from a uniform distribution between zero and one, as in the BS model. However, in contrast to the BS model, a change in the fitness of the i^{th}_{min} species does not automatically trigger a change in the neighbours. Only the interactions are changed.

5.3 Numerical simulations

In order to test the stability of the model we monitor the *average fitness* and the *gap function*, $G(t)$, for both B and F. The gap function is nothing but the tracking function of the minimum of $B_{min}(t)$ (or $F_{min}(t)$). At $t = 0$ we have $G(0) = B_{min}(0)$ (or $G(0) = F_{min}(0)$). As the evolution proceeds eventually for a certain t_1 we will have $B_{min}(t_1) > B_{min}(0)$ (or $F_{min}(t_1) > F_{min}(0)$) as the minimum values are converging toward the critical value. The gap function is then updated as $G(t_1) = B_{min}(t_1)$ (or $G(t_1) = F_{min}(t_1)$) and so on. It is easy to see that in the stationary state the gap function converges toward the critical value[1].

In Fig. 5.3 the time series of average values and the gap function of F are plotted for different number of species in the ecology. The time to reach the stable state depends strongly on the size: for $N = 10^4$, the largest system in our simulations, we need approximately $t \sim 10^8$ mutations to achieve the equilibrium. Note also that a simple rescaling, $t \to t/N$, leads to a collapse of these curves. The relaxation times in the BS model are, approximately, one order of magnitude lower compared to the LIBS model of the same size (or in rescaled time).

A snapshot of the grid in the stable configuration is shown in Fig. 5.4. We notice immediately that that the local fitness (B) is no longer distributed like a step function (as for BS). Rather a long, exponential, tail of low fitness is evident, as shown in Fig. 5.5 (Left). The cells with a higher local fitness still have a greater probability to survive but the global-fitness (F), or the presence of environmental partnerships, widens the possibility of survival, even for some species with a lower degree of self-fitness.

If we examine the global-fitness, a single avalanche is present – as in the classical BS model. Moreover the probability distribution function for the avalanche duration, shown in Fig. 5.6 and computed with respect to F, is power law dis-

[1]For simulations on finite BS systems, a perfectly stationary state can never be achieved during a finite number of mutations. This drawback, discussed in Refs. [Hea00, Tab01, Hea02], is due to spurious correlations in the dynamics of the avalanches induced by the finiteness of the lattice as $G(t)$ gets closer to the critical value and, therefore, their average duration is supposed to diverge. As soon as we get very close to this point, an artifact regime sets in and the gap function starts to saturate toward $B = 1$. The phase in which $G(t) \sim B_c$ can be regarded as a transition point for the physically meaningful state: the larger the system is, that is the closer it is to the thermodynamic limit, the slower is the drift from this point and the system can be regarded, to a good approximation, as stable. An accurate study of this phenomenon in relation to the LIBS model, although very interesting, is not of fundamental importance in the context of the present work. Therefore we will consider the system to be stable as soon as the gap function and the average reach a plateau.

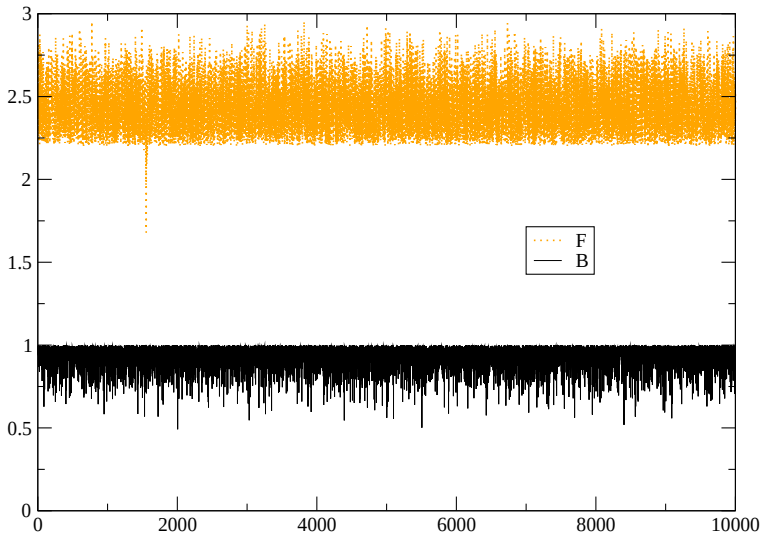

Figure 5.4: Snapshot of the grid, $N = 10^4$, in a stable configuration for fitness B and global fitness F.

tributed, in relation with the criticality of the model. The index of the distribution turns out to be different from that of BS: the change in the dynamics has also led to a change in the universality class of the model.

The distribution of global-fitness, shown in Fig. 5.5 (Right) differs significantly from the step-function of the BS model. It displays a polynomial decay (4^{th} order fit in the plot) above a critical threshold, as a result of the convolution of stochastic variables. Remarkably, a similar nontrivial distribution can also be found in the size distribution of firms, suggesting a possible practical application of the LIBS model as we are going to see in the next Sec. 5.6.

5.4 Second order neighbours in the LIBS model

In order to further investigate the importance of the environment in the LIBS model we now include the second nearest-neighbour interactions in the simple one dimensional topology. The global-fitness of Eq. (5.2) becomes

$$F_i = B_i + \Lambda_{i,i-1}B_{i-1} + \Lambda_{i,i+1}B_{i+1} + \Lambda_{i,i-2}B_{i-2} + \Lambda_{i,i+2}B_{i+2}, \qquad (5.4)$$

where the second order coefficients are not independent random numbers but are rather related via $\Lambda_{i,i\pm2} = \Lambda_{i,i\pm1}\cdot\Lambda_{i\pm1,i\pm2}$. The reason behind this choice, that can be easily extended to the nth order neighbours, is motivated by the assumption

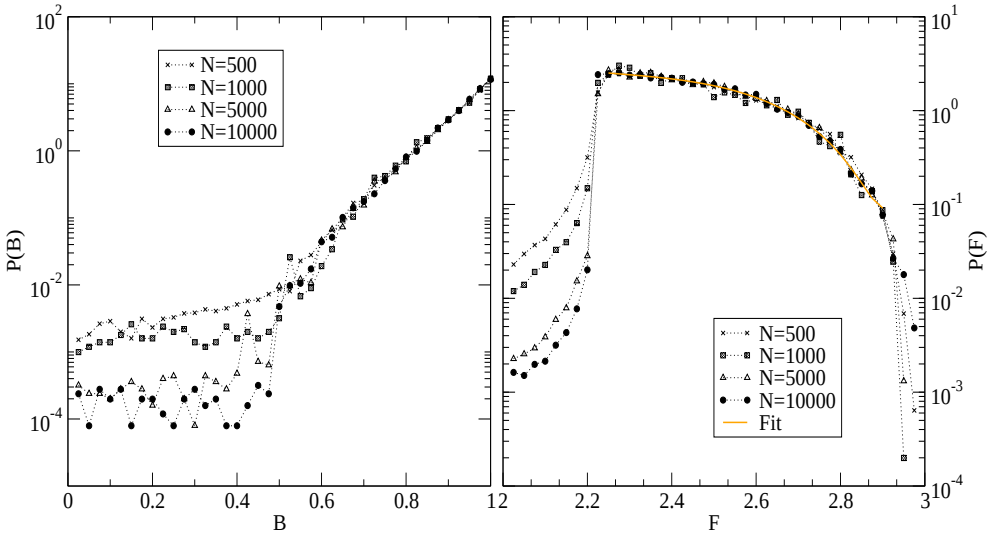

Figure 5.5: Left: probability distribution function, $P(B)$, for fitness B. The step distribution of the standard BS model has been replaced with an exponential distribution with a cut-off at $B \sim 0.55$. Right: Probability distribution, $P(F)$, for the global-fitness, F, and fit with a fourth order polynomial. In this case a sharp threshold is visible, as in the standard BS model, indicating that poor fitness environments undergo mutation. As we consider larger N the transition, at $F_c \sim 2.22$, gets sharper and sharper as expected by a finite size analysis. The values of the $P(F)$ below this threshold are related to the recorded avalanches. The distributions shown in these plots are the results of an average over 50 different configurations in the stable regime.

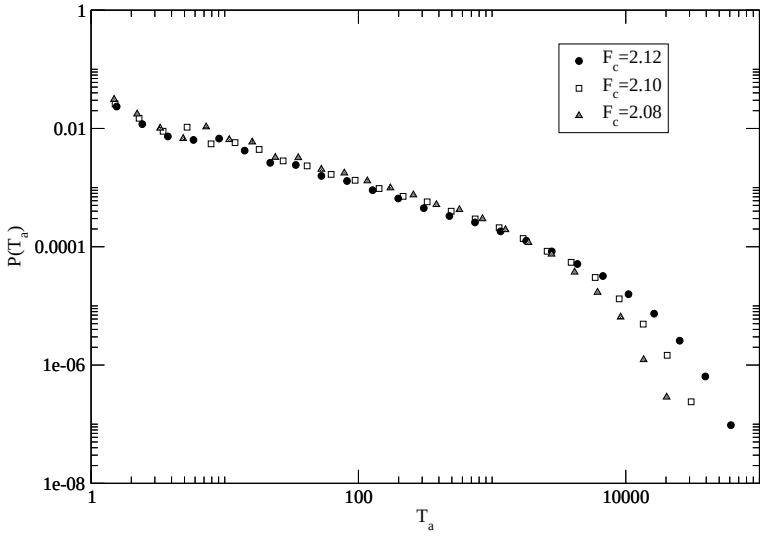

Figure 5.6: Probability distribution, $P(T_a)$, of the avalanche duration, T_a, in the LIBS model. Three different thresholds during the stable regime are considered. The critical exponent is $\gamma \sim 0.83$. For $F_c = 2.20$ we are very close to the critical point and very large avalanches are present.

that the higher order interactions are damped by the "distance" between the two species and therefore $\Lambda_{i,i\pm2} \leq \Lambda_{i,i\pm1}$. By using this formulation, we attempt to mimic a hierarchical dependence of the global fitness in the ecology: species become explicitly related to their second nearest neighbours via the mediation of their first neighbours and so on. Using these constraints the mutation rules remain the same as in Eq. (5.3) since a change in the first order coefficients triggers automatically a change in the higher order ones.

The dynamics that results from the numerical simulations is similar to the first-neighbour LIBS model. After an extensive transient we reach a critical stationary state characterized by avalanches of mutations with size, T_a, which are power law distributed. The distributions for B and F, in the stable regime, are shown in Fig. 5.7 along with the distribution obtained by considering just the first neighbour interaction.

In the case of B we notice that by enlarging the neighborhood, the distributions show a slower decaying rate and they appear to be smoother. In this case we have an exponential decay all the way down to zero, without any clear cut-off for low B. Despite their fitness, all the species have a chance of survival if sustained by a healthy environment. Regarding the global fitness, instead, a polynomial decay is still evident, although the order is higher compared to the

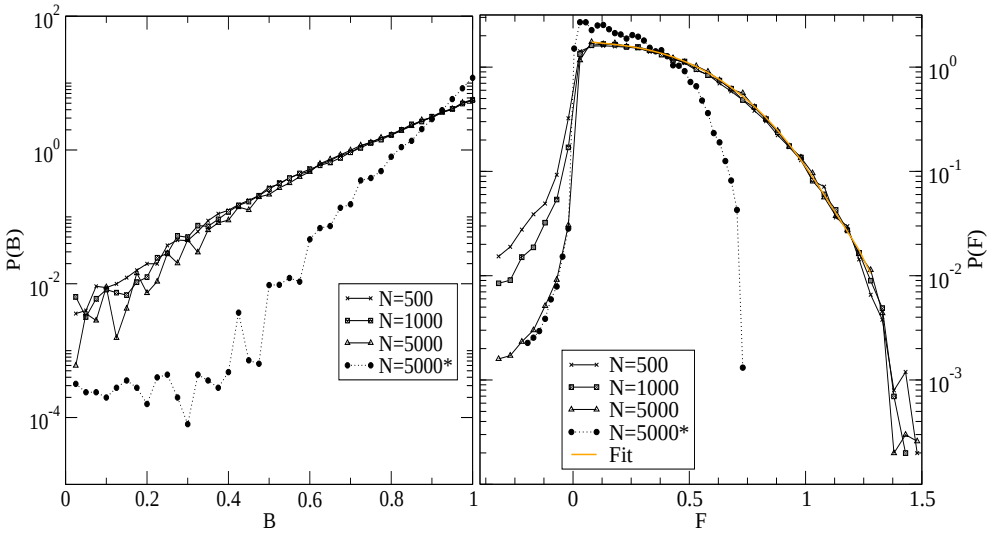

Figure 5.7: Left: distribution of self-fitness, B, for the second order neighbours for various number of cells, N. The distribution for the first order neighbours is also plotted (*). Note that an extension of the neighborhood leads to a slower decay rate. Moreover, a clear cut-off for low B is no longer evident. Right: distribution of global fitness F. In this case the distributions have been rescaled according to $F \to F - F_c$. For the second order neighbours we have $F_c \sim 2.37$. The fit with a sixth order polynomial, also shown, produces a smaller χ^2 compared to the fourth order used for the first nearest neighbour case.

first neighbours case. It is also important to note that a relatively large change in the theoretical range for F, for which the bounds are now $0 \leq F_i \leq 5$, does not lead to a consequent rise in the threshold value. It just moves slightly from $F_c \sim 2.22$ to $F_c \sim 2.37$. Nevertheless, this is equivalent to saying that in the previous case, in order to be considered "fit", a species had to exceed roughly 74% of the possible range for F. Now just 48% is sufficient! In conclusion, a hierarchical extension of the cooperation between species in the LIBS model leads to an easier adaptation and survival probability: the more compact the ecosystem is the higher will be the chances of survival of each single species as long as they cooperate for their mutual interest.

5.5 LIBS model on complex topologies: beyond "democracy"

In the previous section we considered the behaviour of the LIBS model on regular topologies, such as a 1D grid. However, as we have seen in Chapter 4, most of the time the interactions between species, and complex systems in general, are actually governed by complex scale-free networks. If we consider the LIBS model on a topology in which the number of connections per species is not homogeneous then the different kinds of convolution between the B values can generate a different shapes in the distribution of the global-fitness. This can be easily deduced by writing Eq. (5.2) in a general form as

$$F_i = B_i + \sum_{j=1}^{k_i} \Lambda_{i,j} B_j, \tag{5.5}$$

where the sum over j is extended to all the k_i neighbours of the i^{th} species and no particular topology is specified.

For an isotropic model on a D-dimensional lattice, k_i is equal for all the species and depends only on D and the definition of neighborhood: the theoretical boundaries for F are the same for each element and we can talk about a "democratic" model. From Eq. (5.5), we can immediately see that, by adopting a complex network as the underlying structure for the interactions between species, as most likely for a complex system [AB02, DM02], we move to a model in which each species may have a different boundary value for the global-fitness, since $0 \leq F_i \leq k_i$. This inequality has a straightforward interpretation: species with a large number of connections will have a higher barrier against environmental changes because they can rely on numerous resources.

Moreover, the fact that in Eq. (5.5) the global fitness depends directly on the number of links per species can lead to a loss of criticality in the model. In fact, hubs will hardly mutate and the avalanches will be limited to the poorly connected nodes[2]: the mutations are localized in a few, possibly distant, nodes of the network and the correlations that give rise to power laws are, therefore, lost.

We can account for this inconvenience by assuming that the dependency of F from the neighbours species is of a non-linear kind,

$$F_i = B_i + \frac{1}{k_i^{\vartheta}} \sum_{j=1}^{k_i} \Lambda_{i,j} B_j. \tag{5.6}$$

where the extra parameter $\vartheta \in [0,1]$ is a measure of the influence of the environment on the i^{th} species. For $\vartheta \to 0$ we recover the model of Eq. (5.5) where just few species can easily mutate, while for $\vartheta \to 1$ we recover a completely democratic model that behaves in a way similar to the 1D LIBS. We argue that, in the real world, this parameter should be something in between.

A snapshot of the ecology fitness resulting from simulations carried out on a Barabási-Albert (BA) scale-free network with tunable clustering (see Sec. 4.2) and $\vartheta = 0.5$ is shown in Fig. 5.8 for a stable configuration. Note how the nodes on the left had side, that correspond to the older ones in the BA algorithm, that is the ones with the larger number of connections, are also the fittest: by relying on a vast environment they can afford to have a smaller value of B. The avalanches are present on the right hand side where nodes with less connections are concentrated.

Moreover, the distribution of F in this case matches the distribution of connectivity of the network, giving rise to a power law shape Fig. 5.9 reports results from an ensemble average over 10 network configurations.

Despite the preliminary nature of the results for the LIBS model on a complex topology, we have seen how its extremal dynamics can lead to a power law shape in the distribution of the fitness parameter which characterizes the status of a species. Note that in this example we neglected the reason for the formation of the SF topology. It would be of a great interest to develop a model which can evolve naturally to this configuration while retaining, at the same time, the main features of the LIBS. This issue will be addressed in our future studies [3].

[2]This situation of "freezing" of the large "hubs" has some analogies with spin glass theory where some species freeze in a random configuration leading to a rough landscape of energy levels at small temperatures [MPV87].

[3]It is also worth pointing out that another parameter related to $P(F)$ is the domain of values for B itself. In the present case we take B to be uniform in the interval (0,1). In fact, a change in this distribution, while preserving the dynamics of the model, would lead to a

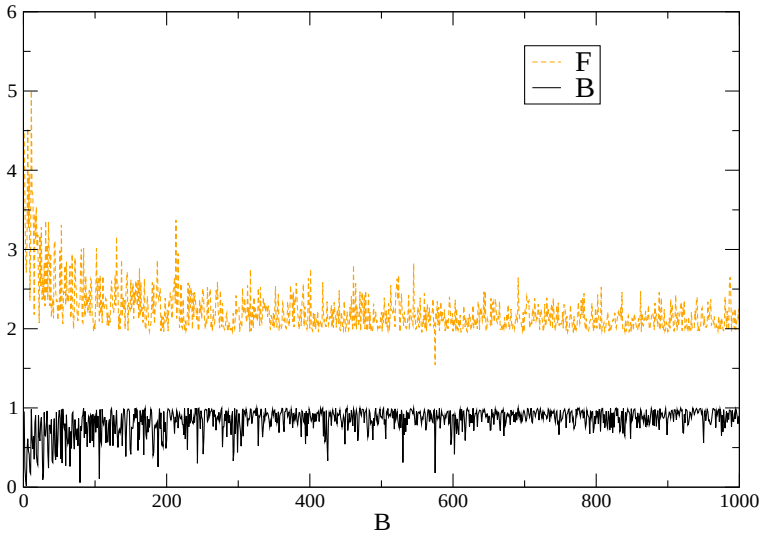

Figure 5.8: Snapshot of the grid, $N = 10^3$, in a stable configuration for fitness B and global fitness F. In this case the clustering of the network is $C \sim 0.39$. The nodes on the left-hand side, the ones with more connections, are characterized by a high value of F. The self-fitness, B, in this case can be even very small without inducing any mutation. Avalanches are present for the poorly connected nodes on the right-hand side.

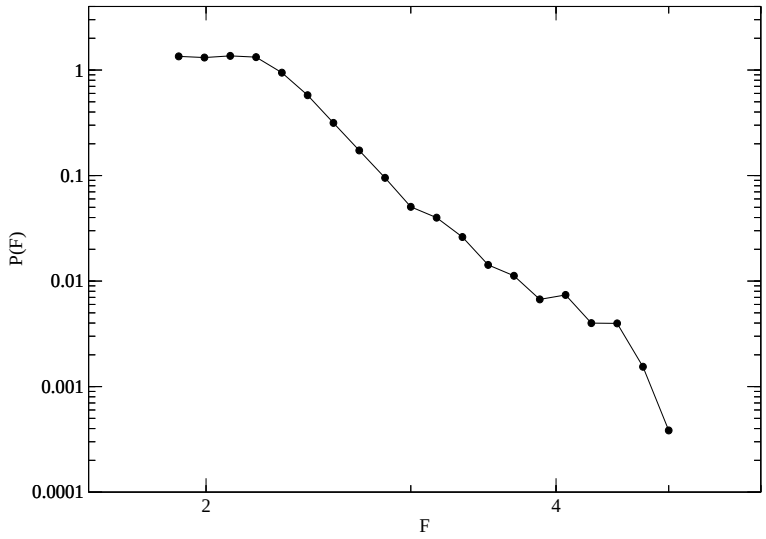

Figure 5.9: Probability distribution function of F on a BA network with $C \sim 0.39$ and averaged over 10 networks. The power law shape is a consequence of the structure of the network and depends both on ϑ and the distribution of B.

5.6 LIBS model and evolutionary economy: a possible application to evolution of firms

We have seen in the previous sections that the LIBS model produces a broad distribution (polynomial) of F in the 1D case and a power law in the in the same quantity when the underlying topology is represented by a scale-free network. These findings can suggest a possible application for the LIBS model in the context of the *evolutionary theory of economics*. This theory, which takes inspiration from the Darwinian principles of evolution, has become popular among many economists and the literature on the subject is rapidly growing. In a nutshell, the main idea behind this theory is that the development of economic entities, such as firms, is ruled by a complex network of interdependencies and competition. Therefore, their evolution is similar to that of species in an ecosystem which, in order to survive, have to "adapt" to the environment [4].

Recently, some empirical studies have tried to shed some light on this debated issue. In particular, Axtell [Axt01] analyzed the size distribution of U.S. companies, as defined by the total number of employees, during 1997. He found

different shape in the final distribution of F.

[4] For a recent review on the subject see Ref. [Day04].

that it could be well represented by a Zipf distribution, $P(s) \sim s^{-\alpha}$, with $\alpha \sim 1$ and s being the size of the firm. Further investigation of this issue has been carried out by Gaffeo *et al.* [GGP03], who analyzed a database of companies for the G7 countries from 1987 to 2000. To some extent, this analysis confirmed the findings of Axtell, namely a power law distribution with $\alpha \sim 1$. However, $\alpha \sim 1$ was obtained only for a particular definition of the firm size, and particular business periods. More generally, they found a robust power law behaviour. However, the index was seen to change with the time window analyzed and the definition of firm size used. Moreover, it has been shown empirically [VBT01, BCLM03, SFA04], that there exists a complex network of interactions between the various firms composing the stock markets.

In this framework, we can assume that the global fitness, F in the LIBS model, represents the actual size of a firm. The numerical results discussed in the previous sections then show that its distribution, in equilibrium condition, come to be very similar to the power law observed empirically.

5.7 Discussion and conclusion

The results obtained with the LIBS model confirm the relevance of self-organized criticality in complex systems and, in particular, economics. The concept of mutual cooperation, introduced via the global-fitness, can explain the ubiquity of broad tails in the distribution of characteristic quantities of physical and social systems in terms of a convolution of variables between elements of the network of interaction. In the economic context, this asymptotic behaviour can be related to the empirical findings concerning the distribution of the size of firms. The possible relevance of self-organized criticality in economics has already been suggested by recent theoretical and empirical studies [BCSW93, BPS97, PA00, BT04], while possible applications of the BS model in this field can be found in Refs. [CVV01, Yam01, ACP04]. The application of the SOC concept to social sciences can, in general, be motivated by empirical observations of the "intermittent" activity in human dynamics at every level, from wars to revolutions and, in particular, intellectual production where moments of frenetic activity can alternate with long breaks, with lengths which cannot be predicted.

This process is, in a way, similar to the discharge, via avalanches, needed in the classical sandpile model, to restore the critical slope. In a real economic world a wide series of changes, similar to avalanches, can be triggered by exogenous or endogenous shocks related to structural changes at macroeconomic level, for example the creation and successive enlargement of the European community or the fall of the Soviet empire, or at microeconomic level, the invention of a

new technology [Day04]. Since the shocks leading to avalanches are of different nature, we also expect the existence of different time scales involved in the self-organization process. In SOC systems, in general, the existence of a sharp separation between time scales, energy storage and relaxation, appears to be a strict prerequisite. In the BS model, as in the LIBS model, by mutating one unit at a time, we implicitly assume that the times to extinction, t_e, of species are well separated. If we consider, for example, an exponential dependence on the global fitness, that is $t_e(i) \propto e^{F_i}$, we can reproduce the "punctuation" phenomenon empirically observed in the fossil record, see Fig 5.2. In economic terms we can still assume this behaviour: changes of unfit firms can be simply related to small microeconomic fluctuations that can happen in time scales of the order of weeks or months while much longer times are needed to change the fitness of a highly adapted company. In the latter case radical changes are needed, as for example a switch from one political regime to another, an event that may take centuries to happen.

In conclusion, we have extended the Bak-Sneppen model for biological evolution by introducing explicitly local interactions between elements of the ecology. Numerical simulations have shown how the dynamics of the model, while still leading to a self-organized critical state, can be dramatically effected by environmental forces, leading to smoother distributions in both the intrinsic fitness, B, and the global fitness, F. As already pointed out by Grassberger [Gra95] the BS model cannot be taken too seriously for describing the punctuated equilibrium of biological evolution. Nevertheless, because of its simplicity, it can easily be used as paradigm for other complex systems. In the present work we suggest a possible application of our extension of the BS model to the economic world. In particular the distribution of global-fitness can be related to the size distribution of firms in the most developed markets. In this respect the evolution of firms is seen as a punctuated equilibrium process in which the convolution of mutual interest can justify the spreading in size of the firms themselves. We have emphasized that the actual shape of the distribution of global-fitness is related to the topology of the interaction. Future work will be devoted to a more exhaustive study of the LIBS model on complex topologies.

6

Conclusion

The interdisciplinary field of "complexity" has gained more and more popularity among the physics community in the past years. Different areas of research, traditionally precluded, such as biology or economics, are now systematically studied with methods borrowed from statistical mechanics, chaos theory, stochastic processes, critical phenomena etc...

On occasion, this fervent enthusiasm has led also to paths that have not much to do with science, especially when the aims of the research are not well defined. In fact, the term "complexity", in its broad etymological meaning, can be misleading and blindly related to all sort of problems which are just complicated to solve: in this case the boundary between physics and philosophy become thin.

In order to avoid these ambiguities, that unfortunately can be found also in literature, the present Thesis has focused on the study of particular, and formally well defined, complex behaviour of some specific systems such as the stock market, spin glasses and ecologies of species.

In the first part of this thesis, which globally covers Chapters 2 and 3, we investigated, via data analysis and numerical simulations, the phenomena of self-organization and herding in stock market dynamics.

Chapter 2, in particular, has been dedicated to the research of empirical evidence for *self-organized criticality* and *phase transitions*. The results have shown that after an appropriate wavelet filtering, the temporal evolution of the fluctuations in the price dynamics is characterized by an avalanche-like behaviour similar to the one found in sand pile models. A statistical analysis over the avalanches reveals a power law distribution in size, duration and laminar times. These empirical findings, although not exhaustive, indicate that a charge-discharge process, typical of self-organized criticality, can play an important role in stock market dynamics. The results of the analysis are robust against various markets and time scales. If further confirmed, this behaviour should be considered as a novel stylized fact of the financial markets.

Further indications of self-organized behaviour in the market has been revealed in the evolution of the price index. In fact, during specific periods, it follows a power law with superimposed embedded log-periodic oscillations. This phenomenon, with a time scale of a few years to a few months for the short em-

bedded structures, has been found in various western markets from 2000 until the beginning 2003. Its presence can be due to a self-induced critical state, similar to phase transitions in physics. Moreover, we argued that these log-periodic oscillations which characterize these periods result from an underlying discrete scale invariance of the system with a preferred scaling ratio of $\lambda \sim 2$. These results are confirmed also by an independent Lomb analysis over the frequencies of the oscillations.

In Chapter 3 a stochastic cellular automata is used in order to capture the herding behaviour believed to be responsible for the large fluctuations observed in the stock market. By using simple percolation rules, distinct dynamical clusters of trading agents emerge on a 2D grid representing the *open market*. The size of the groups of traders, inside which the information is shared, defines a hierarchy in the trading power. Although the process of decision making is stochastic, large fluctuations, like crashes, can take place as a result of the synchronization, or herding, in the decision to buy or sell a stock by most of the agents composing the largest clusters. The dynamical bahaviour of the cellular automata, tested against a time series of daily closures of the S&P500 index, reproduces most of the stylized facts of the real data such as the power law wings in the probability distribution function of the returns as well as their correlation properties. Also the multifractal analysis proves the good agreement between model and empirical data. These results point out the importance of the hierarchical behaviour in the trading dynamics. The bigger networks have also a bigger weight and, alone, can deeply influence the market behaviour.

In Chapter 4 we investigated another important property that results from the self-organization of a complex system; that is the network of the interactions between the elements composing the system itself. During the past few years, this area of research had a burst in popularity mainly due to a series of empirical findings which brought to light a preferential structural organization that is independent of the specific nature of the systems under consideration. In particular, it has been found that most networks of interaction present a power law distribution in the connectivity of each element (*scale-free*) along with a high clustering of elements.

Motivated by these findings, we studied the implications of this particular topology in two different models. The first of them is the antiferromagnetic Ising model, widely studied in statistical mechanics, while the second is a Boolean model of opinion formation. The results of the numerical simulations have shown how both models are deeply influenced by this structure.

The antiferromagnetic Ising model displays a spin glass transition at low temperatures induced by the random frustration generated by the network itself. This behaviour is not observed in regular lattices.

The the model for opinion formation, instead, shows a dynamical turbulent-like regime for a certain range of the parameters. Thanks to the peculiar nature of the scale-free network, where a relative large number of hubs are present, this behaviour is found to be relatively robust against the random indecision of the agents. This model can be considered also as a paradigm for financial trading where one of the two opinions represent the will to buy or sell a stock. The results are in agreement with different stylized facts such as the distribution of returns and the multifractal spectrum. However, discrepancies in the behaviour of the temporal correlation in the volatility have been found. We argued that this fact may be due to the oversimplified nature of the simulation where just one single network of a fixed size has been considered. In comparing these results with those obtained with the cellular automata, we find further proof of the existence of a hierarchical organization in market activity where almost independent networks of different sizes (or trading power) are involved at the same time.

The issue of self-organization, and in particular self-organized criticality, has been addressed also in a model of species evolution/extinction, in Chapter 5. In this case we have investigated the role played by the environment in an extremal dynamics model, named Local Interactions Bak-Sneppen (LIBS). Here the interaction between the species are explicitly expressed by random strength coefficients. The results of the simulations show that the probability of extinction/mutation is deeply related to the environment that the species live. Cooperation can broaden the possibility of survival even for the weaker links. Moreover, the broad shape of the distribution of the global fitness suggests a parallel with the data observed for the size distribution of firms in the G7 countries and a possible application of the model in evolutionary economics.

In conclusion, we studied and characterized different manifestations of the emerging behaviour of complex systems. We do not certainly pretend that this work could be exhaustive for the understanding of this broad subject where so much work is still left to do, but, at the same time, we hope to have shed some light on some important issues concerning, in particular, self-organization and the role played by complex topologies.

The best way to conclude this work is with an humble citation from Sir I. Newton:

> *I do not know what I may appear to the world, but to myself I seem to have been only like a boy playing on the seashore and diverting myself in now and then finding a smoother pebble or a prettier shell than ordinary whilst the great ocean of truth lay all undiscovered before me.*