Eur. Phys. J. B **26**, 521–529 (2002)
DOI: 10.1140/epjb/e20020122

THE EUROPEAN PHYSICAL JOURNAL B
EDP Sciences

Epidemic outbreaks in complex heterogeneous networks

Y. Moreno[1], R. Pastor-Satorras[2,a], and A. Vespignani[1]

[1] The Abdus Salam International Centre for Theoretical Physics, PO Box 586, 34100 Trieste, Italy
[2] Departament de Física i Enginyeria Nuclear, Universitat Politècnica de Catalunya Campus Nord, Mòdul B4, 08034 Barcelona, Spain

Received 20 September 2001 and Received in final form 4 February 2002

Abstract. We present a detailed analytical and numerical study for the spreading of infections with acquired immunity in complex population networks. We show that the large connectivity fluctuations usually found in these networks strengthen considerably the incidence of epidemic outbreaks. Scale-free networks, which are characterized by diverging connectivity fluctuations in the limit of a very large number of nodes, exhibit the lack of an epidemic threshold and always show a finite fraction of infected individuals. This particular weakness, observed also in models without immunity, defines a new epidemiological framework characterized by a highly heterogeneous response of the system to the introduction of infected individuals with different connectivity. The understanding of epidemics in complex networks might deliver new insights in the spread of information and diseases in biological and technological networks that often appear to be characterized by complex heterogeneous architectures.

PACS. 89.75.-k Complex systems – 87.23.Ge Dynamics of social systems – 05.70.Ln Nonequilibrium and irreversible thermodynamics

1 Introduction

The epidemiology of heterogeneous networks has largely benefited from the need of understanding the spreading of human sexual diseases in the complex web of sexual partnership [1–3]. Epidemic modeling considered that population groups can be characterized in classes having different sexual activity or number of sexual contacts. This fact leads to models dealing with heterogeneous populations which are known to enhance the spread of infections as well as make them harder to eradicate (for a review see [4]). In this perspective, a limiting case is represented by the newly identified classes of complex networks (for a review see [5]). The highly heterogeneous topology of these networks is mainly reflected in the small average path lengths among any two nodes (small-world property) [6,7], and in a power law distribution (scale-free property), $P(k) \sim k^{-2-\gamma}$, for the probability that any node has k connections to other nodes [8]. While regular networks present finite connectivity fluctuations ($\langle k \rangle \simeq \langle k^2 \rangle$), scale-free (SF) networks are a limiting case of heterogeneity where connectivity fluctuations are diverging if $0 < \gamma \leq 1$. In other words, the network nodes possess a statistically significant probability of having a virtually unbounded number of connections compared to the average value. SF networks find real examples in several technological systems such as the Internet [9,10] and the world-wide-web (WWW) [11], as well as in natural systems such as food-webs, and metabolic or protein networks [5]. The need to understand the dynamics of information transmission, the error tolerance [12–14], and other properties of complex networks has therefore called for the study of epidemic modeling in complex networks [15–19].

A surprising result, originated by the inspection of the susceptible-infected-susceptible (SIS) model, has shown that the spread of infections is tremendously strengthened on SF networks [18,19]. Opposite to standard models, epidemic processes in these networks do not possess, in the limit of an infinite network, any epidemic threshold below which the infection cannot produce a major epidemic outbreak or an endemic state. In principle, SF networks are prone to the persistence of diseases whatever infective strength they may have. This feature reverberates also in the choice of immunization strategies [20–22] and changes radically many standard conclusions on epidemic spreading. This study appears particularly relevant in the case of technological networks, for instance for the spreading of digital viruses in the Internet [18], and it has soon been generalized by showing that also the susceptible-infected-removed (SIR) model shows the same absence of epidemic threshold [23]. These results highlight the study of epidemic models in complex networks as potentially relevant also in human and animal epidemiology [23], as confirmed

[a] e-mail: romu@complex.upc.es

recently by the experimental observation that the web architecture of sexual contacts is best described by a scale-free topology in which individuals have widely different connectivities [24].

In this paper we provide a detailed analytical and numerical study of the SIR model on two prototype complex networks: the Watts-Strogatz (WS) model and the Barabási-Albert (BA) model. The first model is a small-world network with bounded connectivity fluctuations, while the second one is the prototype example of SF network. The analytical approach allows us to recover the total size of the epidemics in an infinite population, in agreement with earlier estimates [23]. We are able to find the analytic expression for the critical threshold as a function of the moments of the connectivity distribution and we confirm the absence of any finite threshold for connectivity distributions $P(k) \sim k^{-2-\gamma}$ with $0 < \gamma \leq 1$, in the infinite size limit. We obtain the general analytic expression for the total density of infected individuals and the epidemic threshold at arbitrary γ values. Finite size network effects can be easily evaluated from the analytic expressions. Time evolution and other effects of heterogeneity such as the relative infection incidence in different connectivity classes can be predicted. In order to confirm the analytical findings we perform large scale numerical simulations on the WS and BA networks. Numerical results are in perfect agreement with the analytical predictions and confirm that the interplay of complex networks topology and epidemic modeling leads to a new and interesting theoretical framework, whose predictions and implications need to be exhaustively explored.

During the completion of this paper we became aware of a work by May and Lloyd [25] which reports a comprehensive study of the SIR model in scale free networks. This work extends the preliminary account provided in reference [23].

2 The SIR model

Our theoretical understanding of epidemic spreading is based on compartmental models, in which the individuals in the population are divided in a discrete set of states [4,26]. In this framework, diseases which result in the immunization or death of infected individuals can be characterized by the classical susceptible-infected-removed (SIR) model [4,26]. In this model individuals can only exist in three different states: susceptible (healthy), infected, or removed (immunized or dead). In a homogeneous system, the SIR model can be described in terms of the densities of susceptible, infected, and removed individuals, $S(t)$, $\rho(t)$, and $R(t)$, respectively, as a function of time. These three magnitudes are linked through the normalization condition

$$S(t) + \rho(t) + R(t) = 1, \tag{1}$$

and they obey the following system of differential equations:

$$\begin{aligned} \frac{\mathrm{d}S}{\mathrm{d}t} &= -\lambda \overline{k} \rho S, \\ \frac{\mathrm{d}\rho}{\mathrm{d}t} &= -\mu\rho + \lambda \overline{k} \rho S, \\ \frac{\mathrm{d}R}{\mathrm{d}t} &= \mu\rho. \end{aligned} \tag{2}$$

These equations can be interpreted as follows: infected individuals decay into the removed class at a rate μ, while susceptible individuals become infected at a rate proportional to both the densities of infected and susceptible individuals. Here, λ is the microscopic spreading (infection) rate, and $\overline{k}$ is the number of contacts per unit time that is supposed to be constant for the whole population. In writing this last term of the equations we are assuming the *homogeneous mixing* hypothesis [4], which asserts that the force of the infection (the per capita rate of acquisition of the disease by the susceptible individuals) is proportional to the density of infectious individuals. The homogeneous mixing hypothesis is indeed equivalent to a mean-field treatment of the model, in which one assumes that the rate of contacts between infectious and susceptibles is constant, and independent of any possible source of heterogeneity present in the system. Another implicit assumption of this model is that the time scale of the disease is much smaller than the lifespan of individuals; therefore we do not include in the equations terms accounting for the birth or natural death of individuals.

The most significant prediction of this model is the presence of a nonzero epidemic threshold λ_c [26]. If the value of λ is above λ_c, $\lambda > \lambda_c$, the disease spreads and infects a finite fraction of the population. On the other hand, when λ is below the threshold, $\lambda < \lambda_c$, the total number of infected individuals (the epidemic prevalence), $R_\infty = \lim_{t\to\infty} R(t)$, is infinitesimally small in the limit of very large populations (the thermodynamic limit). In order to see this point, let us consider the set of equations (2), in which, without lack of generality, we set $\mu = 1$. Integrating the equation for $S(t)$ with the initial conditions $R(0) = 0$ and $S(0) \simeq 1$ (*i.e.*, assuming $\rho(0) \simeq 0$, a very small initial concentration of infected individuals), we obtain

$$S(t) = \mathrm{e}^{-\lambda \overline{k} R(t)}. \tag{3}$$

Combining this result with the normalization condition (1), we observe that the total number of infected individuals R_∞ fulfills the following self-consistent equation:

$$R_\infty = 1 - \mathrm{e}^{-\lambda \overline{k} R_\infty}. \tag{4}$$

While $R_\infty = 0$ is always a solution of this equation, in order to have a nonzero solution the following condition must be fulfilled:

$$\frac{\mathrm{d}}{\mathrm{d}R_\infty} \left(1 - \mathrm{e}^{-\lambda \overline{k} R_\infty}\right)\Big|_{R_\infty = 0} > 1. \tag{5}$$

This condition is equivalent to the constraint $\lambda > \lambda_c$, where the epidemic threshold λ_c takes the value $\lambda_c = \overline{k}^{-1}$ in this particular case. Performing a Taylor expansion at $\lambda = \lambda_c$ it is then possible to obtain the epidemic prevalence behavior $R_\infty \sim (\lambda - \lambda_c)$ (valid above the epidemic threshold). In the language of the physics of nonequilibrium phase transitions [27], the epidemic threshold is completely equivalent to a critical point. In analogy with critical phenomena, we can consider R_∞ as the order parameter of a phase transition and λ as the tuning parameter. In particular, it is easy to recognize that the SIR model is a generalization of the dynamical percolation model, that has been extensively studied in the context of absorbing-state phase transitions [27].

3 The SIR model in complex networks

In order to address the effects of contact heterogeneity in epidemic spreading, let us consider the SIR model defined on a network with general connectivity distribution $P(k)$ and a finite average connectivity $\langle k \rangle = \sum_k kP(k)$. Each node of the network represents an individual in its corresponding state (susceptible, infected, or removed), and each link is a connection along which the infection can spread. The disease transmission on the network is described in an effective way: At each time step, each susceptible node is infected with probability λ, if it is connected to one or more infected nodes. At the same time, each infected individual becomes removed with probability μ, that, without lack of generality, we set equal to unity.

In order to take into account the heterogeneity induced by the presence of nodes with different connectivity, we consider the time evolution of the magnitudes $\rho_k(t)$, $S_k(t)$, and $R_k(t)$, which are the density of infected, susceptible, and removed nodes of connectivity k at time t, respectively. These variables are connected by means of the normalization condition

$$\rho_k(t) + S_k(t) + R_k(t) = 1. \tag{6}$$

Global quantities such as the epidemic prevalence are therefore expressed by an average over the various connectivity classes; *i.e.*, $R(t) = \sum_k P(k)R_k(t)$. At the mean-field level, these densities satisfy the following set of coupled differential equations:

$$\frac{\mathrm{d}\rho_k(t)}{\mathrm{d}t} = -\rho_k(t) + \lambda k S_k(t)\Theta(t), \tag{7}$$

$$\frac{\mathrm{d}S_k(t)}{\mathrm{d}t} = -\lambda k S_k(t)\Theta(t), \tag{8}$$

$$\frac{\mathrm{d}R_k(t)}{\mathrm{d}t} = \rho_k(t). \tag{9}$$

The factor $\Theta(t)$ represents the probability that any given link points to an infected site. This quantity can be computed in a self-consistent way [18]: The probability that a link points to a node with s links is proportional to $sP(s)$. Thus, the probability that a randomly chosen link points to an infected node is given by

$$\Theta(t) = \frac{\sum_k kP(k)\rho_k(t)}{\sum_s sP(s)} = \frac{\sum_k kP(k)\rho_k(t)}{\langle k \rangle}\,. \tag{10}$$

In this approximation we are neglecting the connectivity correlations in the network, *i.e.*, the probability that a link points to an infected node is considered independent of the connectivity of the node from which the link is emanating. A more refined approximation would consider the network correlations as given by the conditional probability $P(k\,|\,k')$ that a node with given connectivity k' is connected to a node with connectivity k [28]. Nevertheless, as we will see in the next sections, this rather crude approximation is quite able to give account of many of the properties shown by computer simulations of the model.

The equations (7–9), combined with the initial conditions $R_k(0) = 0$, $\rho_k(0) = \rho_k^0$, and $S_k(0) = 1 - \rho_k^0$ completely define the SIR model on any complex network with connectivity distribution $P(k)$. We will consider in particular the case of a homogeneous initial distribution of infected nodes, $\rho_k^0 = \rho^0$. In this case, in the limit $\rho^0 \to 0$, we can substitute $\rho_k(0) \simeq 0$ and $S_k(0) \simeq 1$. Under this approximation, equation (8) can be directly integrated, yielding

$$S_k(t) = \mathrm{e}^{-\lambda k \phi(t)} \tag{11}$$

where we have defined the auxiliary function

$$\phi(t) = \int_0^t \Theta(t')\mathrm{d}t' = \frac{1}{\langle k \rangle}\sum_k kP(k)R_k(t), \tag{12}$$

and in the last equality we have made use of the definition (10).

In order to get a closed relation for the total density of infected individuals, it results more convenient to focus on the time evolution of the averaged magnitude ϕ. To this purpose, let us compute its time derivative:

$$\frac{\mathrm{d}\phi(t)}{\mathrm{d}t} = \frac{1}{\langle k \rangle}\sum_k kP(k)\rho_k(t) \tag{13}$$

$$= \frac{1}{\langle k \rangle}\sum_k kP(k)(1 - R_k(t) - S_k(t)) \tag{14}$$

$$= 1 - \phi(t) - \frac{1}{\langle k \rangle}\sum_k kP(k)S_k(t). \tag{15}$$

Introducing the obtained time dependence of $S_k(t)$ we are led to the differential equation for $\phi(t)$

$$\frac{\mathrm{d}\phi(t)}{\mathrm{d}t} = 1 - \phi(t) - \frac{1}{\langle k \rangle}\sum_k kP(k)\mathrm{e}^{-\lambda k\phi(t)}. \tag{16}$$

Once solved equation (16), we can obtain the total epidemic prevalence R_∞ as a function of $\phi_\infty = \lim_{t\to\infty}\phi(t)$. Since $R_k(\infty) = 1 - S_k(\infty)$, we have

$$R_\infty = \sum_k P(k)\left(1 - \mathrm{e}^{-\lambda k \phi_\infty}\right). \tag{17}$$

Equations (16, 17) constitute thus an alternative representation of the model, with respect to equations (7–9). It is worth remarking that identical equations can be obtained in the framework developed for the HIV transmission in heterogeneous populations by May and Anderson [3].

For a general $P(k)$ distribution, equation (16) cannot be solved in a closed form. However, we can still get useful information on the infinite time limit; *i.e.* at the end of the epidemics. Since we have that $\rho_k(\infty) = 0$ and consequently $\lim_{t\to\infty} \mathrm{d}\phi(t)/\mathrm{d}t = 0$, we obtain from equation (16) the following self-consistent equation for ϕ_∞

$$\phi_\infty = 1 - \frac{1}{\langle k\rangle}\sum_k kP(k)\mathrm{e}^{-\lambda k\phi_\infty}. \tag{18}$$

The value $\phi_\infty = 0$ is always a solution. In order to have a non-zero solution, the condition

$$\frac{\mathrm{d}}{\mathrm{d}\phi_\infty}\left(1 - \frac{1}{\langle k\rangle}\sum_k kP(k)\mathrm{e}^{-\lambda k\phi_\infty}\right)\Bigg|_{\phi_\infty=0} > 1 \tag{19}$$

must be fulfilled. This relation implies

$$\frac{1}{\langle k\rangle}\sum_k kP(k)(\lambda k) = \lambda\frac{\langle k^2\rangle}{\langle k\rangle} > 1. \tag{20}$$

This condition defines the epidemic threshold

$$\lambda_c = \frac{\langle k\rangle}{\langle k^2\rangle} \tag{21}$$

below which the epidemic prevalence is null, and above which it attains a finite value. That is, the threshold is inversely proportional to the connectivity fluctuations $\langle k^2\rangle$. For regular networks, in which $\langle k^2\rangle < \infty$, the threshold has a finite value and we are in the presence of a standard phase transition. On the other hand, networks with strongly fluctuating connectivity distribution, show a *vanishing* epidemic threshold for increasing network sizes; *i.e.* $\langle k^2\rangle \to \infty$ for $N \to \infty$. It is worth remarking that real networks have always a finite size N and thus an effective threshold, depending on the magnitude of $\langle k\rangle$ and $\langle k^2\rangle$, that can be easily calculated as a function of N [25]. This apparent threshold, however is not an intrinsic quantity and it is extremely small for systems with large enough N. A more detailed discussion of finite-size effects can be found in reference [29].

4 Exponentially distributed networks: The Watts-Strogatz model

The class of exponential networks refers to random graph models which produce a connectivity distribution $P(k)$ peaked at an average value $\langle k\rangle$ and decaying exponentially fast for $k \gg \langle k\rangle$ and $k \ll \langle k\rangle$. Typical examples of such a network are the Erdös and Rényi model [30] and the small-world model of Watts and Strogatz (WS) [31]. The latter has recently been the object of several studies as a good candidate for the modeling of many realistic situations in the context of social and natural networks. In particular, the WS model shows the "small-world" property common in random graphs [6]; *i.e.*, the diameter of the graph – the shortest chain of links connecting any two vertices – increases very slowly, in general logarithmically, with the network size [32]. On the other hand, the WS model has also a local structure (clustering property) that is not found in random graphs with finite connectivity [31,32]. The WS graph is defined as follows [31,32]: The starting point is a ring with N nodes, in which each node is symmetrically connected with its $2m$ nearest neighbors. Then, for every node each link connected to a clockwise neighbor is rewired to a randomly chosen node with probability p, and preserved with probability $1 - p$. This procedure generates a random graph with a connectivity distributed exponentially for large k [31,32], and an average connectivity $\langle k\rangle = 2m$. The graph has small-world properties and a non-trivial "clustering coefficient"; *i.e.*, neighboring nodes share many common neighbors [31,32]. The richness of this model has stimulated an intense activity aimed at understanding the network's properties upon changing p and the network size N [6,15,31–34]. At the same time, the behavior of physical models on WS graphs has been investigated, including epidemiological percolation models [13,15,16] and models with epidemic cycles [17].

In the following we focus on the WS model with $p = 1$; it is worth noticing that even in this extreme case the network retains some memory of the generating procedure. The network, in fact, is not locally equivalent to a random graph in that each node has at least m neighbors. In the limit $p \to 1$, the connectivity distribution of the WS network, as defined previously, takes the form [35]

$$P(k) = \frac{m^{k-m}}{(k-m)!}\mathrm{e}^{-m} \qquad \text{for } k \geq m. \tag{22}$$

For general regular networks, for which $\langle k^n\rangle < \infty$ for all values of n, equations (16, 17) can be approximately solved in the limit $\phi(t) \to 0$, by expanding the exponentials under the summation signs. Thus, for the case of the total epidemic prevalence R_∞ in equation (17),

$$R_\infty \simeq \sum_k P(k)\lambda k\phi_\infty = \langle k\rangle\,\lambda\phi_\infty. \tag{23}$$

That is, R_∞ is linearly proportional to ϕ_∞.

On the other hand, by expanding the exponential in equation (16) and keeping the most relevant terms, we obtain:

$$\frac{\mathrm{d}\phi}{\mathrm{d}t} \simeq 1 - \phi - \frac{\sum_k kP(k)(1 - \lambda k\phi + \lambda^2k^2\phi^2/2)}{\langle k\rangle} \tag{24}$$

$$= \phi\left(-1 + \lambda\frac{\langle k^2\rangle}{\langle k\rangle} - \lambda^2\phi\frac{\langle k^3\rangle}{2\,\langle k\rangle}\right). \tag{25}$$

The resulting previous equation can be integrated, to yield

$$\phi(t) = \frac{2(\lambda - \lambda_c)\,\langle k\rangle}{\langle k^3\rangle\,\lambda^2 + A\mathrm{e}^{-(\lambda-\lambda_c)t/\lambda_c}}, \tag{26}$$

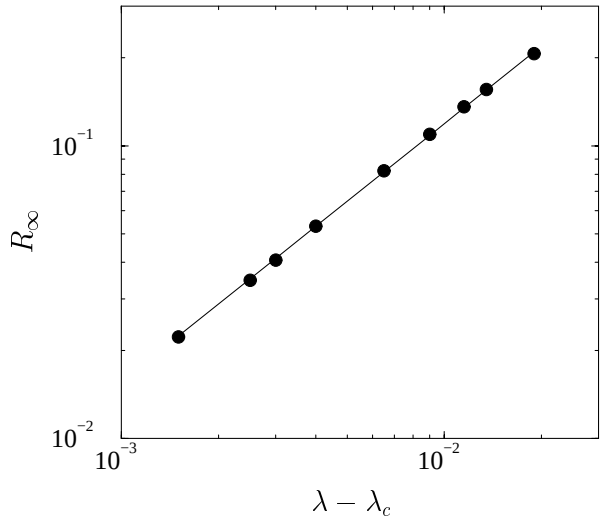

Fig. 1. Total density of infected individuals R_∞ as a function of $\lambda - \lambda_c$ for the SIR model in WS networks of size $N = 10^6$. The value of $\lambda_c = 0.184(5)$ is in good agreement with the analytical prediction. The full line is a fit to the form $R_\infty \sim (\lambda - \lambda_c)^\beta$ with $\beta = 0.9(1)$.

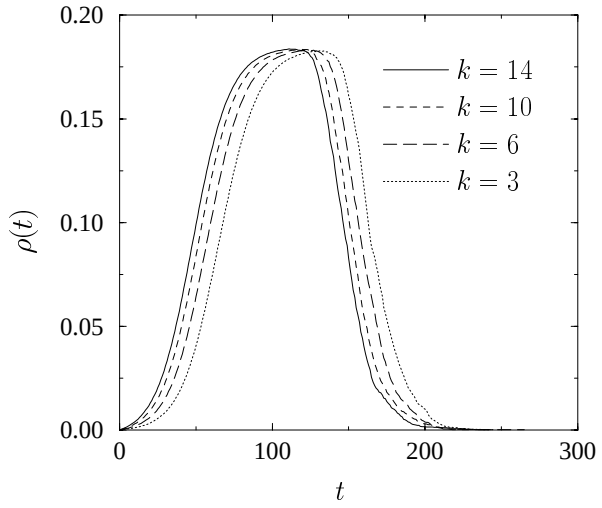

Fig. 2. Total density of infected nodes as a function of time for the SIR model in WS networks, starting from initial conditions peaked in nodes of connectivity k (initial infected individuals randomly distributed only among the nodes of connectivity k). The spreading rate is fixed to $\lambda = 0.20$. The network size is $N = 10^6$.

where λ_c is defined as in equation (21) and A is an integration constant. That is, for $\lambda < \lambda_c$, $\phi_\infty \to 0$, while from $\lambda > \lambda_c$, we recover the well-known mean-field behavior $\phi_\infty \sim (\lambda - \lambda_c)$, which translated to the total epidemic prevalence R_∞ yields

$$R_\infty \sim (\lambda - \lambda_c). \tag{27}$$

In the particular case of the WS networks, we expect to observe the behavior dictated by equation (27), with an epidemic threshold that, using the connectivity distribution equation (22), is given by

$$\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle} = \frac{2}{1 + 4m} \cdot \tag{28}$$

In order to compare with the analytical predictions we have carried out large scale simulations of the SIR model in the WS network with $p = 1$. In our simulations we consider the WS network with parameter $m = 3$, which corresponds to an average connectivity $\langle k \rangle = 6$. Simulations were implemented on graphs with number of nodes ranging from $N = 10^3$ to $N = 3 \times 10^6$, averaging over at least 10^4 different epidemic outbreaks, performed on at least 10 different realization of the random network. In Figure 1, we show the total density of removed nodes at the end of the epidemic outbreak as a function of the parameter λ. The graph exhibits an epidemic threshold at $\lambda_c = 0.184(5)$ that is approached with a roughly linear behavior by R_∞. A linear fit to the form $R_\infty \sim (\lambda - \lambda_c)^\beta$ provides an exponent $\beta = 0.9(1)$, in reasonable agreement with the analytical findings. This confirms that the SIR model in exponentially bounded complex networks has a behavior similar to that obtained with the mean-field hypothesis. Actually, since the connectivity fluctuations are very small in the WS graph ($\langle k^2 \rangle \sim \langle k \rangle$), as a first approximation we can consider the WS model as a homogeneous one in which each node has the same number of links, $k \simeq \langle k \rangle$. In order to provide further evidence to this effective homogeneity, we show in Figure 2 the time evolution of the density of infected nodes for epidemic outbreaks starting only on nodes with a given connectivity k. The total epidemic prevalence is almost constant for all connectivity k, with a slight shift of the peak time of the outbreak. The figure clearly shows that the system reacts almost identically to this heterogeneous initial conditions, confirming that the mean-field assumption is correctly depicting the system's behavior. We shall see in the next section that this is not the case for SF networks.

5 Power-law distributed networks: The Barabási-Albert model

The Barabási-Albert (BA) graph was introduced as a model of growing network (such as the Internet or the world-wide-web) in which the successively added nodes establish links with higher probability pointing to already highly connected nodes [8]. This is a rather intuitive phenomenon on the Internet and other social networks, in which new individuals tend to develop more easily connections with individuals which are already well-known and widely connected. The BA graph is constructed using the following algorithm [8]: We start from a small number m_0 of nodes; every time step a new vertex is added, with m links that are connected to an old node i with probability

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j}, \tag{29}$$

where k_i is the connectivity of the ith node. After iterating this scheme a sufficient number of times, we obtain a

network composed by N nodes with connectivity distribution $P(k) \sim k^{-3}$ and average connectivity $\langle k \rangle = 2m$ (in this work we will consider the parameters $m_0 = 5$ and $m = 3$). Despite the well-defined average connectivity, the scale invariant properties turn out to play a major role on the physical properties of these networks (for instance, the resilience to attack [13,14]).

In the continuous k approximation, that substitutes the discrete variable k for a continuous variable in the range $[m, \infty[$, the connectivity distribution of the BA model takes the form

$$P(k) = \frac{2m^2}{k^3} \qquad \text{for } k \geq m. \tag{30}$$

For this distribution, the first moment is finite, $\langle k \rangle = 2m$, but the second moment diverges with the network size, $\langle k^2 \rangle \sim m^2 \log N$ [36]. In view of the general result equation (21), we observe that a finite network composed by N nodes, should exhibit an *effective* epidemic threshold

$$\lambda_c(N) \sim \frac{1}{\log N}, \tag{31}$$

which appears as a consequence of finite size effects, as costumarily encountered in nonequilibrium statistical systems [27]. For very large networks, however, $\lambda_c(N)$ will tend to zero and we will observe a null threshold in the thermodynamic limit [25,29].

The equation for R_∞, with the connectivity distribution (30) is

$$R_\infty = 1 - 2m^2 \int_m^\infty k^{-3} \mathrm{e}^{-\lambda k \phi_\infty} \mathrm{d}k. \tag{32}$$

This integral can be performed and expressed in terms of the incomplete Gamma function [37]. Expanding the obtained result for small ϕ_∞ yields

$$R_\infty \simeq 2\lambda m \phi_\infty. \tag{33}$$

On its turn, the equation for $\phi(t)$, with the connectivity distribution (30), is

$$\frac{\mathrm{d}\phi(t)}{\mathrm{d}t} = 1 - \phi(t) - m \int_m^\infty k^{-2} \mathrm{e}^{-\lambda k \phi} \mathrm{d}k. \tag{34}$$

Expressing the previous integral in terms of incomplete Gamma functions and expanding for small $\phi(t)$ we are led to the equation

$$\frac{1}{\lambda m} \frac{\mathrm{d}\phi(t)}{\mathrm{d}t} \simeq \phi \left[1 - \gamma_E - \frac{1}{\lambda m} - \ln(\lambda m \phi) \right]. \tag{35}$$

This equation can be integrated, to yield

$$\phi(t) \simeq \frac{1}{\lambda m} \exp\left(1 - \gamma_E - \frac{1}{\lambda m} + A \mathrm{e}^{-\lambda m t} \right), \tag{36}$$

where A is an integration constant. The stationary regime for long times is

$$\phi_\infty \simeq \frac{\mathrm{e}^{1-\gamma_E}}{\lambda m} \mathrm{e}^{-1/\lambda m}, \tag{37}$$

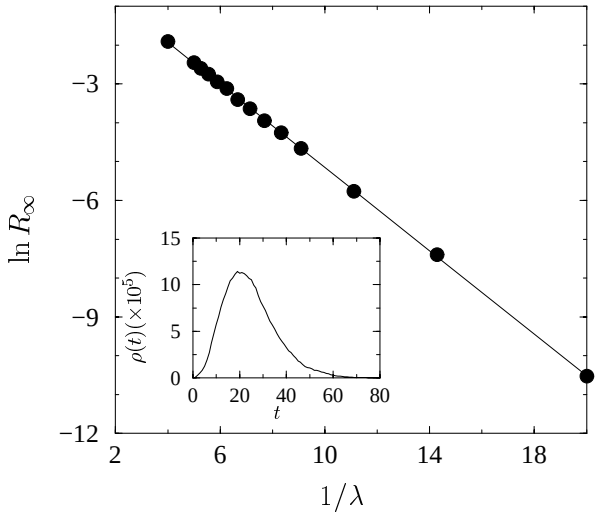

Fig. 3. Total density of infected individuals R_∞ as a function of $1/\lambda$ for the SIR model in BA networks of size $N = 10^6$. The linear behavior on the semi-logarithmic scale proves the stretched exponential behavior predicted by equation (38). The inset show the time profile of the average density of infected individuals at the spreading rate $\lambda = 0.09$.

and by inserting this result into the expression for the total epidemic prevalence we find

$$R_\infty \sim \mathrm{e}^{-1/\lambda m}. \tag{38}$$

That is, in an infinite network the function R_∞ is non-zero for any non-zero value of λ, which is in agreement with the predicted threshold $\lambda_c = 0$. This result recovers the findings of Lloyd and May [23] as well as the behavior obtained by considering a diverging connectivity variance in the results reported by May and Anderson for HIV spreading in heterogeneous populations [3]. By following this framework is also possible to relate the absence of the epidemic threshold to the divergence of the basic reproductive number, customarily defined in traditional epidemiological modeling [23,25].

The numerical simulations performed on the BA network confirm the picture extracted from the analytic treatment. We consider the SIR model on BA networks of size ranging from $N = 10^3$ to $N = 10^6$, with $m = 3$ and thus $\langle k \rangle = 6$. As predicted by the analytic calculations, Figure 3 shows that R_∞ decays with λ as $R_\infty \sim \exp(-C/\lambda)$, where C is a constant. In order to rule out the presence of finite size effects hiding an abrupt transition (the so-called smoothing out of critical points [27]), we have inspected the behavior of the stationary persistence for network sizes varying over three orders of magnitude. The total absence of scaling of R_∞ and the perfect agreement for any size with the analytically predicted exponential behavior allows us to definitely confirm the absence of any finite epidemic threshold. A closer look at R_∞ is given in Figure 4. While Figure 3 reports the average over 10^4–10^5 epidemic outbreaks, Figure 4 reports an illustration of the behavior of the cumulative probability $P(R_\infty > R)$ of having an outbreak which affects more

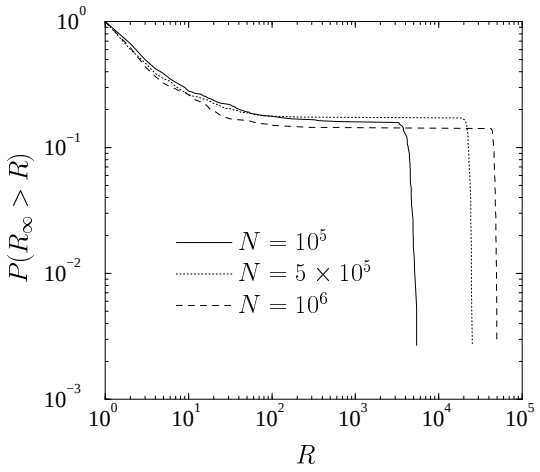

Fig. 4. Cumulated outbreak epidemic distribution for the SIR model in BA networks. The spreading rate is fixed to $\lambda = 0.09$.

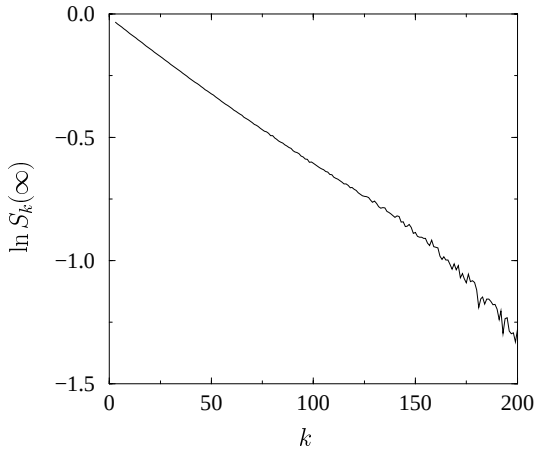

Fig. 5. Density $S_k(\infty)$ of susceptible nodes as a function of k for the SIR model in BA networks. Epidemics start from random initial conditions (initial infected individuals randomly distributed among all nodes). The spreading rate is fixed to $\lambda = 0.09$. The network size is $N = 10^6$. The linear-log plot recovers the exponential form predicted in equation (11).

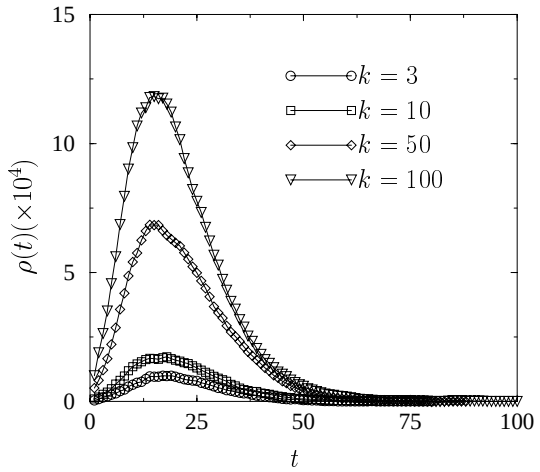

Fig. 6. Total density of infected nodes as a function of time for the SIR model in BA networks, starting from initial conditions concentrated in nodes of connectivity k (initial infected individuals randomly distributed among the nodes of connectivity k). The spreading rate is fixed to $\lambda = 0.09$. The network size is $N = 10^6$.

than R individuals in a single realization at $\lambda = 0.09$. The figure shows a finite probability of having outbreaks involving a number of individuals of the order of the network size N. The large plateau corresponds to a gap between large events and small outbreaks that give rise to a zero density of infected individuals in the $N \to \infty$ limit. Accordingly to the predictions, the plateau extends proportionally to N for increasing network sizes.

The large heterogeneity of these networks can be pictorially characterized by inspecting the epidemic evolution in each class of connectivity k. We know from equation (11), that the susceptibles densities $S_k(t)$ decay much faster in the highly connected classes. In particular, we have that $S_k(\infty) \sim \exp(-\lambda k \phi_\infty)$. In Figure 5, we report $S_k(\infty)$ as a function of k on a semi-logarithmic scale. The plot shows the expected linear relation in k. Since $R_k(\infty) = 1 - S_k(\infty)$, the curves clearly show that the higher is the nodes' connectivity, the higher is the relative prevalence of the epidemic outbreak. Classes of nodes with few connections have a small density of removed individuals (total number of infected individuals), while highly connected classes ($k \gg 100$) are almost totally affected by the infection. A further striking evidence of the peculiar behavior of the SF networks is obtained by inspecting epidemic outbreaks starting on nodes with different connectivity k. While an analytical solution for this case is very troublesome, the numerical investigation presents clear-cut results. In Figure 6 we present the infection incidence profile for epidemic outbreaks started on sites with different connectivities k. The population results much weaker (higher number of infected individuals) to epidemics starting on highly connected individuals. This weakness points out that the best protection of these networks can be achieved by targeted immunization programs [4,22].

6 Generalized scale-free networks

Recently there has been a burst of activity in the modeling of SF complex networks. The recipe of Barabási and Albert [8] has been followed by several variations and generalizations [38–41] and the revamping of previous mathematical works [42,43]. All these studies propose methods to generate SF networks with variable exponent γ. The analytical treatment presented in the previous section for the SIR model can be easily generalized to SF networks with connectivity distribution with $\gamma > 0$. Let us consider a generalized SF network with a normalized connectivity distribution given by

$$P(k) = (1+\gamma) m^{1+\gamma} k^{-2-\gamma}, \tag{39}$$

where we are approximating the connectivity k as a continuous variable and assuming m the minimum connectivity of any node. According to the general result equation (21), the epidemic threshold for infinite networks is given in this case, as a function of γ, as

$$\lambda_c = \frac{\gamma - 1}{\gamma m} \qquad \text{if } \gamma > 1, \tag{40}$$

$$\lambda_c = 0 \qquad \text{if } \gamma \leq 1. \tag{41}$$

When the network size is finite, finite size effects come at play and we recover a non-zero effective threshold $\lambda_c(N)$ for $\gamma \leq 1$. It can be easily checked, however, that the effective threshold tends to zero as a power law of N [29].

To obtain the explicit expression for ϕ_∞ and R_∞ we must solve the equations (16, 17) for the general connectivity distribution (39). While the differential equation (16) cannot be solved in a closed form for general γ, we can obtain an approximation to the steady state value at long times, ϕ_∞, solving the algebraic equation

$$\phi_\infty = 1 - \frac{1}{\langle k \rangle} \sum_k k P(k) \mathrm{e}^{-\lambda k \phi_\infty}. \tag{42}$$

Integrating equation (17) in the continuous k approximation, the total epidemic prevalence R_∞ takes the form

$$R_\infty = 1 - (1+\gamma)(\lambda m \phi_\infty)^{1+\gamma} \Gamma(-1-\gamma, \lambda m \phi_\infty), \tag{43}$$

where $\Gamma(a, z)$ is the incomplete Gamma function [37]. In the limit $\phi_\infty \to 0$, we can perform a Taylor expansion of the right hand side of equation (43), which at lowest order yields:

$$R_\infty \sim \frac{\gamma + 1}{\gamma} \lambda m \phi_\infty. \tag{44}$$

In order to find the infinite time limit value ϕ_∞, we must solve the equation (42). Substituting the form of the generalized connectivity distribution (39), we obtain

$$\phi_\infty = 1 - \gamma (\lambda m \phi_\infty)^\gamma \Gamma(-\gamma, \lambda m \phi_\infty). \tag{45}$$

The leading behavior of the right hand side of this equation depends on the particular value of γ considered. Thus we can consider the following cases:

(a) $0 < \gamma < 1$: In this case, we have

$$\phi_\infty \sim (\lambda m)^{\gamma/(1-\gamma)}. \tag{46}$$

Combining this result with equation (44), we obtain

$$R_\infty \sim \lambda^{1/(1-\gamma)}. \tag{47}$$

In this range of values of γ we recover the absence of the epidemic threshold, and the associated critical behavior, as we have already shown in Section 5.

(b) $1 < \gamma < 2$: The nontrivial solution for ϕ_∞ is now

$$\phi_\infty \sim \left(\lambda - \frac{\gamma - 1}{\gamma m} \right)^{1/(\gamma - 1)}, \tag{48}$$

which yields

$$R_\infty \sim (\lambda - \lambda_c)^{1/(\gamma-1)}, \tag{49}$$

with an epidemic threshold λ_c given by equation (40).

(c) $\gamma > 2$: The most relevant terms in the expansion of ϕ_∞ are now

$$\phi_\infty \sim \lambda - \frac{\gamma - 1}{\gamma m}, \tag{50}$$

that for the epidemic prevalence yields the behavior

$$R_\infty \sim (\lambda - \lambda_c). \tag{51}$$

The threshold λ_c is again given by the general expression (40). In other words, we recover the usual epidemic framework in networks with connectivity distribution that decays faster than k to the fourth power. Obviously, an exponentially bounded network is included in this last case. It is worth remarking that similar results have been obtained for static percolation in the analysis of the resilience to damage of scale-free networks [13,14].

7 Conclusions

The presented results for the SIR model in complex networks confirm the epidemiological picture proposed in previous works. The topology of the network has a great influence in the overall behavior of epidemic spreading. The connectivity fluctuations of the network play a major role by strongly enhancing the infection's incidence. This issue assumes a particular relevance in the case of SF networks that exhibit connectivity fluctuations diverging with the increasing size N of the web. SF networks are therefore very weak in face of infections presenting an effective epidemic threshold that is vanishing in the limit $N \to \infty$. In the case of the SIR model in an infinite population this corresponds to the absence of any epidemic threshold below which major epidemic outbreaks are impossible. These results strengthen the epidemiological framework for complex networks reported for the SIS model [18,19] and proposed as well for the SIR model [23]. In both models the heterogeneous connectivity allows certain regions of the network to be always in the active phase, thus shifting to zero the epidemic threshold and inducing the same functional form of the order parameter. The emerging picture is likely going to stimulate the re-analysis of several concepts of standard epidemiology such as the "core group" or the characteristic number of contacts that appears to be ill-defined in SF networks.

The high heterogeneity of SF networks finds signatures also in the peculiar susceptibility to infections starting on the most connected individuals and the different relative incidence within populations of varying connectivity k. It is reasonable to expect that these features can point at better protection methods for these networks which appear to have practical realization in many technological and biological systems. In this perspective, the introductions of many elements of realism and a better knowledge

of the networks temporal pattern are fundamental ingredients towards a better understanding of the spreading of information and epidemics in a wide range of complex interacting systems.

This work has been partially supported by the European Network Contract No. ERBFMRXCT980183. R.P.-S. acknowledges financial support from the Ministerio de Ciencia y Tecnología (Spain). We thank M.-C. Miguel, R.V. Solé, S. Visintin, and R. Zecchina for helpful comments and discussions. We are grateful to A.L. Lloyd and R.M. May for enlightening suggestions and for pointing out to us some fundamental references.

References

1. H.W. Hethcote, J.A. Yorke, Lect. Notes Biomath. **56**, 1 (1984).
2. R.M. May, R.M. Anderson, Nature **326**, 137 (1987).
3. R.M. May, R.M. Anderson, Phil. Trans. R. Soc. Lond. B **321**, 565 (1988).
4. R.M. Anderson, R.M. May, *Infectious diseases in humans* (Oxford University Press, Oxford, 1992).
5. S.H. Strogatz, Nature **410**, 268 (2001).
6. D.J. Watts, *Small worlds: The dynamics of networks between order and randomness* (Princeton University Press, New Jersey, 1999).
7. L.A.N. Amaral, A. Scala, M. Barthélémy, H.E. Stanley, Proc. Natl. Acad. Sci. USA **97**, 11149 (2000).
8. A.-L. Barabási, R. Albert, Science **286**, 509 (1999).
9. M. Faloutsos, P. Faloutsos, C. Faloutsos, Comput. Commun. Rev. **29**, 251 (1999).
10. G. Caldarelli, R. Marchetti, L. Pietronero, Europhys. Lett. **52**, 386 (2000).
11. R. Albert, H. Jeong, A.-L. Barabási, Nature **401**, 130 (1999).
12. R.A. Albert, H. Jeong, A.-L. Barabási, Nature **409**, 542 (2000).
13. D.S. Callaway, M.E.J. Newman, S.H. Strogatz, D.J. Watts, Phys. Rev. Lett. **85**, 5468 (2000).
14. R. Cohen, K. Erez, D. ben-Avraham, S. Havlin, Phys. Rev. Lett. **86**, 3682 (2001).
15. M.E.J. Newman, D.J. Watts, Phys. Rev. E **60**, 5678 (1999).
16. C. Moore, M.E.J. Newman, Phys. Rev. E **61**, 5678 (2000).
17. G. Abramson, M. Kuperman, Phys. Rev. Lett. **86**, 2909 (2001).
18. R. Pastor-Satorras, A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001).
19. R. Pastor-Satorras, A. Vespignani, Phys. Rev. E **63**, 066117 (2001).
20. H.W. Hethcote, Theor. Pop. Biol. **14**, 338 (1978).
21. R.M. May, R. M. Anderson, Math. Biosci. **72**, 83 (1984).
22. R. Pastor-Satorras, A. Vespignani, Phys. Rev. E **65**, 036104 (2002).
23. A.L. Lloyd, R.M. May, Science **292**, 1316 (2001).
24. F. Liljeros *et al.*, Nature **411**, 907 (2001).
25. R.M. May, A.L. Lloyd, Phys. Rev. E **64**, 066112 (2001).
26. J.D. Murray, *Mathematical Biology* (Springer Verlag, Berlin, 1993).
27. J. Marro, R. Dickman, *Nonequilibrium phase transitions in lattice models* (Cambridge University Press, Cambridge, 1999).
28. R. Pastor-Satorras, A. Vázquez, A. Vespignani, Phys. Rev. Lett. **87**, 258701 (2001).
29. R. Pastor-Satorras, A. Vespignani, Phys. Rev. E **65**, 035108 (2002).
30. P. Erdös, P. Rényi, Publ. Math. Inst. Hung. Acad. Sci. **5**, 17 (1960).
31. D.J. Watts, S.H. Strogatz, Nature **393**, 440 (1998).
32. A. Barrat, M. Weigt, Eur. Phys. J. B **13**, 547 (2000).
33. M. Barthélémy, L.A.N. Amaral, Phys. Rev. Lett. **82**, 3180 (1999).
34. M. Argollo de Menezes, C. Moukarzel, T.J.P. Penna, Europhys. Lett. **50**, 574 (2000).
35. A. Barrat, M. Weigt, Eur. Phys. J. B **13**, 547 (2000).
36. S.N. Dorogovtsev, J.F.F. Mendes, *Evolution of networks, 2001*, e-print `cond-mat/0106144`.
37. M. Abramowitz, I.A. Stegun, *Handbook of mathematical functions.* (Dover, New York, 1972).
38. R. Albert, A.-L. Barabási, Phys. Rev. Lett. **85**, 5234 (2000).
39. S.N. Dorogovtsev, J.J.F.F. Mendes, A.N. Samukhin, Phys. Rev. Lett. **85**, 4633 (2000).
40. P.L. Krapivsky, S. Redner, F. Leyvraz, Phys. Rev. Lett. **85**, 4629 (2000).
41. B. Tadic, Physica A **293**, 273 (2001).
42. S. Bornholdt, H. Ebel, Phys. Rev. E **64**, 035104 (2001).
43. H.A. Simon, Biometrika **42**, 425 (1955).