Inefficient epidemic spreading in scale-free networks

Carlo Piccardi* and Renato Casagrandi
Dipartimento di Elettronica e Informazione, Politecnico di Milano, Piazza Leonardo da Vinci, 32, I-20133 Milano, Italy
(Received 14 September 2007; published 26 February 2008)

Highly heterogeneous degree distributions yield efficient spreading of simple epidemics through networks, but can be inefficient with more complex epidemiological processes. We study diseases with nonlinear force of infection whose prevalences can abruptly collapse to zero while decreasing the transmission parameters. We find that scale-free networks can be unable to support diseases that, on the contrary, are able to persist at high endemic levels in homogeneous networks with the same average degree.

DOI: 10.1103/PhysRevE.77.026113 PACS number(s): 89.75.Hc, 05.45.−a, 87.19.X-, 87.23.Ge

In a series of influential papers, Pastor-Satorras and Vespignani (PV) [1–3] analyzed how the network structure can affect the key outcomes of the classical susceptible-infective-susceptible (SIS) epidemiological model [4]. As originally noticed in 1927 by Kermack and McKendrick for the susceptible-infective-recovered (SIR) model, "In general a threshold density of population is found to exist, which depends upon the infectivity, recovery and death rates peculiar to epidemic. No epidemic can occur if the population density is below this threshold value" [5]. The existence of such a nonzero epidemic threshold value in the SIR (and SIS) models is a milestone of epidemiological thinking and any evidence that could contradict it would receive a great deal of attention. This is exactly what happened to PV's works. They have in fact shown that when implemented over realistic, scale-free networks (SFNs) the SIS process has unexpected dynamical properties, because the nonzero epidemic threshold does no longer exist as the network size tends to infinity. In other words, the strong heterogeneous nature of SFNs crucially enhances the spreading ability of the epidemic agent in comparison to what happens in homogeneous networks (HNs). As sharply summarized by PV, "This implies that SFNs are prone to the spreading and the persistence of infections whatever spreading rate the epidemic agents might possess" [2]. A striking conclusion that has shattering consequences on the way we usually study the invasion and persistence of pathogens in a variety of contexts, ranging from bacterial infections in humans to computer viruses in the Internet. The result is so simple and clear to claim for very general statements such as "SFNs [are] the ideal media for the propagation of infections, bugs, or unsolicited information. [...] SFNs have the peculiar property of being prone to the spreading of infections" [6]. The risk of such generalizations is to leave the reader with the impression that "SFNs are completely prone to epidemic spreading allowing the onset of large epidemics whatever the spreading rate of infection" [7]. Recent studies have shown that, when more realistic aspects of the infection process are accounted for, an epidemic threshold reemerges in SFNs. In addition to the finite size of the network [8], such features include, e.g., connectivity-dependent transmission rates [9] or degree-dependent deactivation of links [10]. Here we want to make the caveat that in a number of significant epidemic processes different from the standard SIS model, not only an epidemic threshold reappears even in theoretical SFNs of infinite size but, more important, SFNs can be much less efficient than HNs in favoring the disease spread. In particular, we show that this happens in cases where the mean-field epidemiological model exhibits a saddle-node bifurcation, namely, when the disease can survive and establish at high endemic levels in populations that it will be unable to invade from zero [11]. Saddle-node bifurcations emerge in models of important infectious diseases such as hepatitis B [12], tuberculosis [13], or HIV [14]. The prototypical mechanisms that lead to saddle-node bifurcations are nonconstant transmission rates [15], nonconstant recovery rates [16], or both of the above [17]. Departures from the traditional SIS model are relevant not only to the transmission of human diseases but also when the "infectious material" [18] being exchanged is of intellectual [19], commercial [20,21], or social [22] nature.

We consider the population as a network composed of N individuals (nodes). The ith node has degree k_i, i.e., it is connected by $0<k_{\min}\leq k_i\leq k_{\max}$ edges to other nodes. At any time instant t, every node i is either susceptible or infective. During the time interval Δ, an infected node can return susceptible with probability $\gamma\Delta$, while a susceptible node can become infected with probability $\beta G_i\Delta$. The quantity β is the disease transmission rate, whereas the function G_i (detailed below) accounts for the number of infected neighbors of node i. The state of all nodes is updated synchronously with time step Δ. There are two main generalizations of the traditional SIS model in the family of contact processes studied here. The force of infection [23] is nonlinear, as the transmission rate takes the form

$$\beta=\beta(t)=\beta_0+\beta_1 y(t),\quad \beta_0>0,\beta_1\geq 0,$$

where $0\leq y(t)\leq 1$ represents the current disease prevalence, i.e., the fraction of infected individuals in the network at time t. By varying the values attributed to the parameters characterizing the state-independent (β_0) and the state-dependent ($\beta_1 y$) component of the transmission rate β, a wide spectrum of epidemic processes can be described, from the standard SIS model ($\beta_1=0$) to some marketing models [20] for product diffusion ($\beta_0=0$). The function G_i is assumed to

*Author to whom correspondence should be addressed. carlo.piccardi@polimi.it

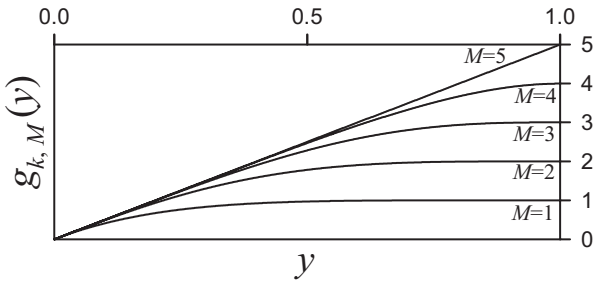

FIG. 1. Examples of the functions $g_{k,M}(y)$ for $k=5$.

increase, yet saturate, with the number I_i of infectives among the neighbors of node i, i.e., $G_i=\min\{I_i,M\}$ where $M\in\{1,2,\ldots,k_{\max}\}$ is the saturation level. A value $M=1$ is adopted in all the network studies implementing the classical SIS model we are aware of. Practically, setting M to unity is equivalent to assuming that the chance for a susceptible node of being infected does not further increases if more than one of its neighbors are carrying the virus.This is quite a typical situation in the word-of-mouth spread of information and rumors or in highly pathogenic diseases.

The spread of every epidemic process depends very much on the network structure. Consistently with [1–3], we compare the spread of the above process in random HNs and SFNs. We constructed our HN by linking pairs of randomly selected nodes, until all of them have exactly the same number $\bar{k}$ of connections, thus the resulting degree distribution p_k is spiky ($p_{\bar{k}}=1$, and $p_k=0$ for all $k\neq\bar{k}$). Other random networks that have been used in the literature can fairly be considered as homogeneous in their connectivity properties. They include the Erdős-Rényi network [24] and the Watts-Strogatz "small-world" network [25], in particular, with rewiring probability equal to 1 [1,2]. We anticipate that all the results presented below are qualitatively identical for all the above mentioned types of random homogeneous networks and, moreover, largely independent of the numerical value assigned to the average degree $\bar{k}$. Letting $\Delta\to 0$, the dynamics of the disease prevalence y in HNs can be described by the following mean-field model:

$$\dot{y}=-\gamma y+(\beta_0+\beta_1 y)(1-y)g^{-}_{k,M}(y), \tag{1}$$

where

$$g_{k,M}(y)=\sum_{I=0}^{k}\min\{I,M\}\binom{k}{I}y^I(1-y)^{k-I} \tag{2}$$

(see Fig. 1) is the expected number of infectives (saturated to M) among the neighbors of any node, because $\binom{k}{I}y^I(1-y)^{k-I}=P(I)$ is the probability that a degree k node has exactly I infected neighbors. Note that $g_{k,M}(0)=0$ and $g'_{k,M}(0)=k$, i.e., the functions $g_{k,M}(y)$ are equivalent up to the first derivative for all values of M if $y\to 0$. In the absence of saturation ($M=k$), $g_{k,k}(y)=ky$ is the standard contagion term of the SIS model. Here we focus on the effects of low saturation values. In the limit case of $M=1$, Eq. (2) reads $g_{k,1}(y)=1-(1-y)^k$ and represents the probability that a node has at least one infected neighbor.

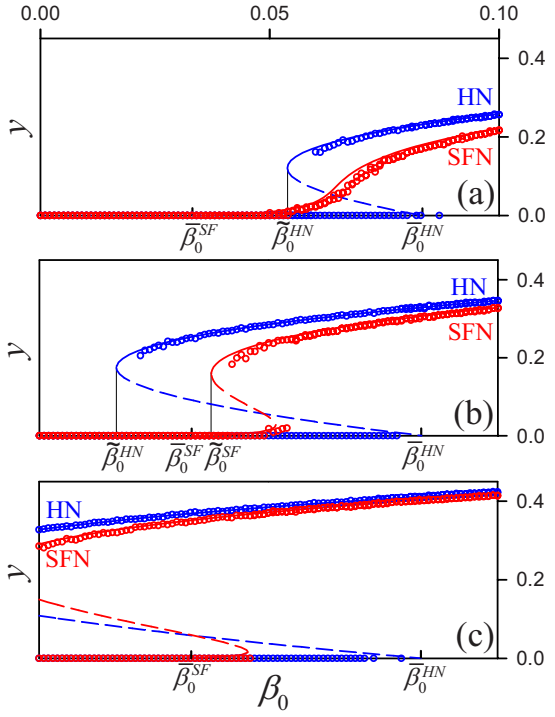

FIG. 2. (Color online). The nontrivial equilibria $y=Y(\beta_0)$ of the HN and SFN mean-field models (1) and (3), for $M=1$, $\bar{k}=12$, and β_1 set to 1 (a), 1.25 (b), and 1.5 (c). Solid (dashed) curves denote stable (unstable) equilibria and have been numerically obtained by continuation [26]. For the SFN model we used a number $n=k_{\max}-k_{\min}+1=600$ of node degrees, which is consistent with a finite-size network of $N\cong(n/k_{\min})^2=10^4$ nodes [3]. As for the stochastic simulations, at each value of β_0 we simulated the epidemiological process over a time horizon of 3×10^4 steps ($\Delta=10^{-2}$) starting from two different network states, i.e., by randomly infecting 2 and 40 % of nodes. In circles are the averages over the last 10^4 steps.

The analysis of the mean-field model (1) reveals that a saddle-node bifurcation is possible in the family of epidemic processes under study. The trivial equilibrium $y=0$ exists for all parameter values, but it is stable if and only if $\beta_0<\bar{\beta}_0^{\mathrm{HN}}=\gamma/\bar{k}$ (easy to prove via linearization). In contrast, the nontrivial equilibria branch $Y=Y(\beta_0)$ is implicitly defined by the condition

$$\beta_0=\frac{\gamma Y-\beta_1 Y(1-Y)g^{-}_{k,M}(Y)}{(1-Y)g^{-}_{k,M}(Y)}$$

and it is such that (a) $Y\to 0$ as $\beta_0\to\bar{\beta}_0^{\mathrm{HN}}$ and (b) $Y\to 1$ as $\beta_0\to\infty$. Property (a) shows that the nontrivial equilibria branch intersects the trivial one at the bifurcation point $\beta_0=\bar{\beta}_0^{\mathrm{HN}}$, exactly as in the standard SIS model. However, as shown in Fig. 2, the slope of $Y(\beta_0)$ at the intersection can be negative for sufficiently large β_1's, because

$$\left.\frac{d\beta_0}{dY}\right|_{Y\to 0}=\gamma\left(\frac{1}{\bar{k}}-\frac{g''_{k,M}(0)}{2\bar{k}^2}\right)-\beta_1$$

and $g''_{\bar{k},M}(0) \leq 0$ for all $\bar{k}$, M (see again Fig. 1). The negative slope at the intersection and property (b) imply the existence of a limit point occurring at another parameter threshold $\beta_0 = \tilde{\beta}_0^{\mathrm{HN}} < \bar{\beta}_0^{\mathrm{HN}}$, where a saddle-node bifurcation occurs. To summarize the general behavior of the epidemiological process over a HN, we find that if $\beta_0 < \tilde{\beta}_0^{\mathrm{HN}}$ the disease cannot persist; if $\beta_0 > \bar{\beta}_0^{\mathrm{HN}}$ there is a unique stable high endemic equilibrium, and if $\tilde{\beta}_0^{\mathrm{HN}} < \beta_0 < \bar{\beta}_0^{\mathrm{HN}}$ the disease can persist or not depending upon initial conditions.

In contrast to HNs, the SFNs are characterized by a heterogeneous degree distribution which takes the form $p_k \approx k^{-\gamma}$ when N (thus $k_{\max}$) tends to infinity and can be created with the Barabási-Albert algorithm of preferential attachment [27] yielding for large N's to $\gamma = 3$, $\bar{k} = 2k_{\min}$, and a degree distribution $p_k = 2k_{\min}(k_{\min}+1)/[k(k+1)(k+2)]$ (e.g., Ref. [28]). Analogously to Refs. [1,2], we describe the mean-field behavior of the epidemic process in a SFN by the following set of $n = k_{\max} - k_{\min} + 1$ equations:

$$\dot{y}_k = -\gamma y_k + (\beta_0 + \beta_1 y)(1 - y_k) g_{k,M}(\tilde{y}), \tag{3}$$

where $0 \leq y_k \leq 1$ is the fraction of infected nodes at time t among those with degree k, $y = \Sigma_{k=k_{\min}}^{k_{\max}} p_k y_k$ is the current disease prevalence, and $\tilde{y} = \Sigma_{k=k_{\min}}^{k_{\max}} k p_k y_k / \bar{k}$ is the expected proportion of infectives at time t among the neighbors of a node, because $k p_k / \bar{k} = q_k$ is the degree distribution of the neighbors (e.g., Ref. [29]). The epidemiological scenario emerging from the analysis of the SFN model (3) has some significant differences with respect to that seen in HNs. The major similarity between the two is in that, despite model (3) is a high-dimensional system, its qualitative asymptotic behavior cannot be other than stationary (i.e., convergence to equilibria). In fact, the off-diagonal entries of the Jacobian matrix are given by

$$\frac{\partial \dot{y}_k}{\partial y_h} = (1 - y_k)\left(\beta_1 p_h g_{k,M}(\tilde{y}) + \frac{(\beta_0 + \beta_1 y) g'_{k,M}(\tilde{y}) h p_h}{\bar{k}} \right)$$

and are strictly positive for all $0 < y_k < 1$. Thus, system (3) is monotone and irreducible. Given that the domain $(0,1)^n$ is invariant, this implies that (almost) all trajectories converge to an equilibrium state [30]. The equilibria Y_k of model (3) are the solutions of the n equations

$$Y_k = \frac{(\beta_0 + \beta_1 Y) g_{k,M}(\tilde{Y})}{\gamma + (\beta_0 + \beta_1 Y) g_{k,M}(\tilde{Y})}, \quad k = k_{\min}, \ldots, k_{\max}.$$

Using the above definitions of y and $\tilde{y}$, we obtain the following pair of algebraic equations that must be nullified with respect to the unknowns $Y, \tilde{Y}$:

$$F_1(Y, \tilde{Y}) = Y - \sum_{k=k_{\min}}^{k_{\max}} p_k \frac{(\beta_0 + \beta_1 Y) g_{k,M}(\tilde{Y})}{\gamma + (\beta_0 + \beta_1 Y) g_{k,M}(\tilde{Y})},$$

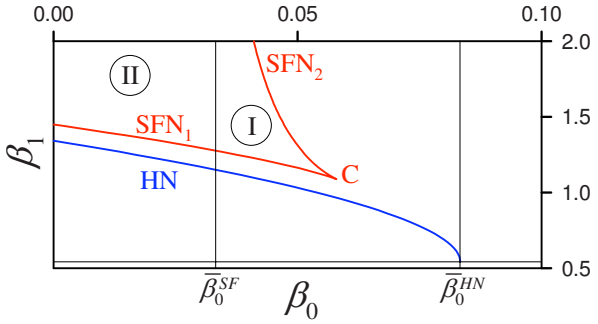

FIG. 3. (Color online) The saddle-node bifurcation curves of the mean-field models (1) and (3). For transmission parameter values in regions I and II, SFNs support the coexistence of two stable equilibria: the high endemic state and either a low endemic state (region I) or a totally susceptible population (region II). Above the HN curve, model (1) has also two attractors, namely, the high endemic state and the totally susceptible population. Unspecified parameters as in Fig. 2.

$$F_2(Y, \tilde{Y}) = \tilde{Y} - \frac{1}{\bar{k}} \sum_{k=k_{\min}}^{k_{\max}} k p_k \frac{(\beta_0 + \beta_1 Y) g_{k,M}(\tilde{Y})}{\gamma + (\beta_0 + \beta_1 Y) g_{k,M}(\tilde{Y})}. \tag{4}$$

The trivial solution of Eq. (4) $Y = \tilde{Y} = 0$ has a branching point at

$$\det \begin{pmatrix} \partial F_1/\partial Y & \partial F_1/\partial \tilde{Y} \\ \partial F_2/\partial Y & \partial F_2/\partial \tilde{Y} \end{pmatrix}_{Y=\tilde{Y}=0} = 1 - \frac{\beta_0}{\gamma} \frac{\langle k^2 \rangle}{\bar{k}} = 0,$$

namely, at $\beta_0 = \bar{\beta}_0^{\mathrm{SF}} = \gamma \bar{k} / \langle k^2 \rangle < \bar{\beta}_0^{\mathrm{HN}}$. This threshold is the same lower bound value that marks the existence of a (stable) nontrivial solution in the standard SIS process over a finite SFN [3]. Differently from the outcome of the SIS process, however, we notice that model (3) can have nontrivial solutions also for $\beta_0 < \bar{\beta}_0^{\mathrm{SF}}$. This can be numerically verified by analyzing how the solutions of Eq. (4) vary with respect to the transmission coefficients β_0 and β_1, as shown in Fig. 2. Sufficiently large values of β_0, no matter the value β_1, guarantee the disease persistence at high endemic levels in both HNs and SFNs. The most interesting result arises at intermediate values of β_1. Figure 2(b) shows the existence of a saddle-node bifurcation in both the HN (at $\beta_0 = \tilde{\beta}_0^{\mathrm{HN}}$) and the SFN (at $\beta_0 = \tilde{\beta}_0^{\mathrm{SF}}$) models. Decreasing the value attributed to parameter β_0 causes an abrupt transition from a high endemic state to extremely low or null prevalences. Notably, in all the numerical analyses we have performed, it turns out that $\tilde{\beta}_0^{\mathrm{HN}} < \tilde{\beta}_0^{\mathrm{SF}}$. In words, this is similar to saying that HNs can support diseases that would be unable to circulate over SFNs. The fact that for large values of β_1 [see Fig. 2(c)] the thresholds $\tilde{\beta}_0^{\mathrm{HN}}$, $\tilde{\beta}_0^{\mathrm{SF}}$ become negative does not change the overall picture.

In order to gain deeper insight on the results of Fig. 2, we performed a bifurcation analysis of models (1) and (3) in the parameter plane (β_0, β_1). For any fixed value of β_1 above the cusp point C in Fig. 3, the epidemiological process over a SFN has two saddle-node bifurcations when β_0 is varied.

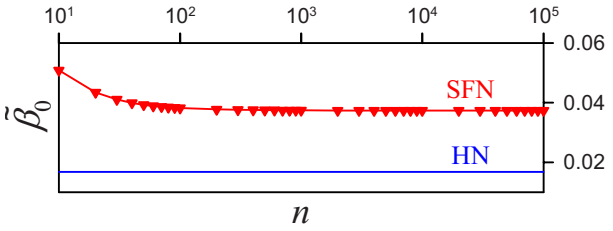

FIG. 4. (Color online) The value of β_0 at the saddle-node bifurcation in both HN and SFN models as a function of $n=k_{\max}-k_{\min}+1$. The value $n=10^5$ corresponds to a finite-size SFN of $N\cong(n/k_{\min})^2\cong 2.8\times 10^8$ nodes [3]. Other parameters set to $\bar{k}=2k_{\min}=12$, $M=1$, and $\beta_1=1.25$.

They can be recognized in Fig. 2(b), where the unstable equilibrium collides respectively with the high ($\tilde{\beta}_0^{\mathrm{SF}}$, branch SFN_1) or the low endemic equilibrium (branch SFN_2). Since the HN curve of Fig. 3 is systematically placed to the left of curve SFN_1, we find that $\tilde{\beta}_0^{\mathrm{HN}}<\tilde{\beta}_0^{\mathrm{SF}}$ for all fixed values of β_1. One could conjecture that the gap between the two thresholds $\tilde{\beta}_0^{\mathrm{HN}}$ and $\tilde{\beta}_0^{\mathrm{SF}}$ is due to the finite-size approximation of the SFN, and that it can vanish as $k_{\max}$ (thus n) becomes increasingly large. This is not true, though, as it can be checked by computing $\tilde{\beta}_0^{\mathrm{SF}}$ as a function of n over many orders of magnitude. Figure 4 makes evident that $\tilde{\beta}_0^{\mathrm{SF}}$ settles to a constant positive limit, which is well above $\tilde{\beta}_0^{\mathrm{HN}}$.

The present study contrasts the different abilities of SFNs and HNs in promoting the spread of epidemics with nonlinear force of infection. At vanishing prevalences, in SFNs the susceptible nodes with very high degree are able to sustain the disease by contacting the extremely rare infectives. This is at the basis of PV's results [1,2] and is confirmed by our analysis, since models (1) and (3) reduce to a standard SIS process when $y\to 0$. In contrast, when the prevalence y is large, the high-degree susceptibles have no disproportionate advantage in contacting many infectives, because of the saturation effect in the contact process. Also, the many susceptible nodes with $k_i<\bar{k}$ have a lower probability of contacting an infective node than those of a HN with the same average degree. What is discussed above is the result of two crucial ingredients. First is the ability of the epidemic to establish high prevalences even at low values of β. In our model, this is obtained by accounting for its dependence on y. Second is the existence of a saturation effect in the contact process. As we pointed out, these key features are deeply rooted in a number of important spreading processes. We also cannot exclude that, in principle, different epidemiological features could give rise to the same qualitative picture presented here. Therefore, we can conclude that it is incorrect to state that SFNs (or, more in general, highly heterogeneous networks) are the most efficient media for the propagation of whatever infection. On the contrary, when analyzing nonelementary spreading processes, the interplay between the network topology and the infection peculiarities can give rise to unexpected outcomes.

This work was supported by the Italian Ministry of the University and Research under Grant No. PRIN 2005098133.

[1] R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001).
[2] R. Pastor-Satorras and A. Vespignani, Phys. Rev. E **63**, 066117 (2001).
[3] R. Pastor-Satorras and A. Vespignani, Phys. Rev. E **65**, 036104 (2002).
[4] N. Bailey, *The Mathematical Theory of Infectious Diseases and Its Application* (Griffin, London, UK, 1975).
[5] W. Kermack and A. McKendrick, Proc. R. Soc. London, Ser. A **115**, 700 (1927).
[6] M. Boguñá, R. Pastor-Satorras, and A. Vespignani, Phys. Rev. Lett. **90**, 028701 (2003).
[7] Y. Moreno and A. Vázquez, Eur. Phys. J. B **31**, 265 (2003).
[8] D. Hwang, S. Boccaletti, Y. Moreno, and R. Lopez-Ruiz, Math. Biosci. Eng. **2**, 317 (2005).
[9] R. Olinky and L. Stone, Phys. Rev. E **70**, 030902(R) (2004).
[10] V. M. Eguíluz and K. Klemm, Phys. Rev. Lett. **89** 108701 (2002).
[11] J. Dushoff, W. Huang, and C. Castillo-Chavez, J. Math. Biol. **36**, 227 (1998).
[12] G. Medley, N. Lindop, W. Edmunds, and D. Nokes, Nat. Med. **7**, 619 (2001).
[13] Z. Feng, C. Castillo-Chavez, and A. Capurro, Theor Popul. Biol. **57**, 235 (2000).
[14] W. Huang, K. Cooke, and C. Castillo-Chavez, SIAM J. Appl. Math. **52**, 835 (1992).
[15] P. van den Driessche and J. Watmough, J. Math. Biol. **40**, 525 (2000).
[16] C. Kribs-Zaleta and J. Velasco-Hernández, Math. Biosci. **164**, 183 (2000).
[17] S. Moghadas and M. Alexander, IMA J. Math. Appl. Med. Biol. **23**, 231 (2006).
[18] W. Goffman and V. Newill, Nature (London) **204**, 225 (1964).
[19] L. Bettencourt, A. Cintrón-Arias, D. Kaiser, and C. Castillo-Chavez, Physica A **364**, 513 (2006).
[20] G. Feichtinger, J. Econ. Behav. Organ. **17**, 401 (1992).
[21] N. Meade and T. Islam, Int. J. Forecast. **22**, 519 (2006).
[22] M. Nakamaru and S. Levin, J. Theor. Biol. **230**, 57 (2004).
[23] M. Keeling and K. Eames, J. R. Soc., Interface **2**, 295 (2005).
[24] P. Erdős and A. Rényi, Publ. Math. (Debrecen) **6**, 290 (1959).
[25] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).
[26] W. Govaerts and Y. Kuznetsov, http://www.matcont.ugent.be/
[27] A. Barabási and R. Albert, Science **286**, 509 (1999).
[28] M. Newman, SIAM Rev. **45**, 167 (2003).
[29] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. Hwang, Phys. Rep. **424**, 175 (2006).
[30] H. Smith, SIAM Rev. **30**, 87 (1988).