Numerical evaluation of the upper critical dimension of percolation in scale-free networks

Zhenhua Wu,[1] Cecilia Lagorio,[2] Lidia A. Braunstein,[1,2] Reuven Cohen,[3] Shlomo Havlin,[3] and H. Eugene Stanley[1]
[1]*Center for Polymer Studies, Boston University, Boston, Massachusetts 02215, USA*
[2]*Departamento de Física, Facultad de Ciencias Exactas y Naturales, Universidad Nacional de Mar del Plata, Funes 3350, 7600 Mar del Plata, Argentina*
[3]*Minerva Center of Department of Physics, Bar-Ilan University, Ramat Gan, Israel*
(Received 14 February 2007; revised manuscript received 1 June 2007; published 27 June 2007)

We propose numerical methods to evaluate the upper critical dimension d_c of random percolation clusters in Erdős-Rényi networks and in scale-free networks with degree distribution $\mathcal{P}(k) \sim k^{-\lambda}$, where k is the degree of a node and λ is the broadness of the degree distribution. Our results support the theoretical prediction, $d_c = 2(\lambda-1)/(\lambda-3)$ for scale-free networks with $3<\lambda<4$ and $d_c=6$ for Erdős-Rényi networks and scale-free networks with $\lambda>4$. When the removal of nodes is not random but targeted on removing the highest degree nodes we obtain $d_c=6$ for all $\lambda>2$. Our method also yields a better numerical evaluation of the critical percolation threshold p_c for scale-free networks. Our results suggest that the finite size effects increases when λ approaches 3 from above.

DOI: 10.1103/PhysRevE.75.066110 PACS number(s): 89.75.Hc, 89.75.Da, 05.10.−a

Recently much attention has been focused on the topic of complex networks, which characterize many natural and man-made systems, such as the Internet, airline transport system, power grid infrastructures, and the world wide web (WWW) [1–4]. Many studies on these systems reveal a common power law degree distribution, $\mathcal{P}(k) \sim k^{-\lambda}$ with $k \geqslant k_{\min}$, where k is the degree of a node, λ is the exponent quantifying the broadness of the degree distribution [5], and $k_{\min}$ is the minimum degree. Networks with power law degree distribution are called scale-free (SF) networks. The power law degree distribution represents topological heterogeneity of the degree in SF networks resulting in the existence of hubs that connect significant fraction of nodes. In this sense, the well studied Erdős-Rényi (ER) networks [6–8] are homogeneous and can be represented by a characteristic degree $\langle k \rangle$, the average degree of a node, while SF networks are heterogeneous and do not have a characteristic degree.

The embedded dimension of ER and SF networks can be regarded as infinite ($d=\infty$) since the number of nodes within a given "distance" increases exponentially with the distance compared to an Euclidean d dimensional lattice network where the number of nodes within a distance L scales as L^d. Percolation theory is a powerful tool to describe a large number of systems in nature such as porous and amorphous materials, random resistor networks, polymerization process, and epidemic spreading and immunization in networks [9,10]. Percolation theory study the topology of a network of N nodes resulting from removal of a fraction $q \equiv 1-p$ of nodes (or links) from the system. It is found that in general there exists a critical phase transition at $p=p_c$, where p_c is the critical percolation threshold. Above p_c, most of the nodes (order N) are connected, while below p_c the network collapses into small clusters of sizes of order $\ln N$. For lattices in $d \geqslant 6$, all percolation exponents remain the same and the system behavior can be described by mean field theory [9,10]. This is because at $d_c=6$ the spatial constraints on the percolation clusters become irrelevant and each shortest path between two nodes in the percolation cluster at criticality can be considered as a random walk.

In standard percolation theory, and in statistical mechanics in general, most of the models are defined on some lattice in a d dimensional space. In this case d_c, the upper critical dimension (UCD), is defined as the lowest dimension d for which the critical exponents take their mean field value. It is well known that the UCD for percolation in d-dimensional lattices is 6. Studies of percolation in ER networks, yield the same critical exponents as in mean-field values of regular percolation in infinite dimensions. This is because in ER networks spatial constraints do not appear and the symmetry is almost the same as in Euclidean lattices, i.e., there is a typical number of links per node. However, SF networks with $2<\lambda<4$ have different critical exponents than ER networks [11,12]. The regular mean-field exponents are recovered only for SF networks with $\lambda>4$. This is due to the fact that for the classical mean field one needs two conditions (a) no spatial constraint (b) translational symmetry, meaning that all nodes have similar neighborhood. The second condition does not apply for SF networks with $\lambda<4$ due to the broad degree distribution and thus we expect a new type of mean field exponents [4]. Indeed, for SF networks with $3<\lambda<4$, the UCD was shown to be [12,13]

$$d_c \equiv \frac{2(\lambda-1)}{\lambda-3}. \qquad (1)$$

Thus, d_c is larger than 6 and for $\lambda \to 3$, $d_c \to \infty$.

The networks discussed in this manuscript are not embedded in any kind of space, and, in fact, cannot be embedded in any finite dimensional space due to the exponential increase in the number of nodes per shell. Therefore, it may seem unnatural to discuss the upper critical dimension, or any form of geometrical (rather than chemical) dimension.

To make sense of the upper critical dimension, two alternate paths can be considered. One possibility is to start with an embedded network model, such as the ones suggested in Refs. [14–16]. Such models describe networks embedded in a finite dimensional space, where the links between nodes are not completely random, but rather nodes are connected to

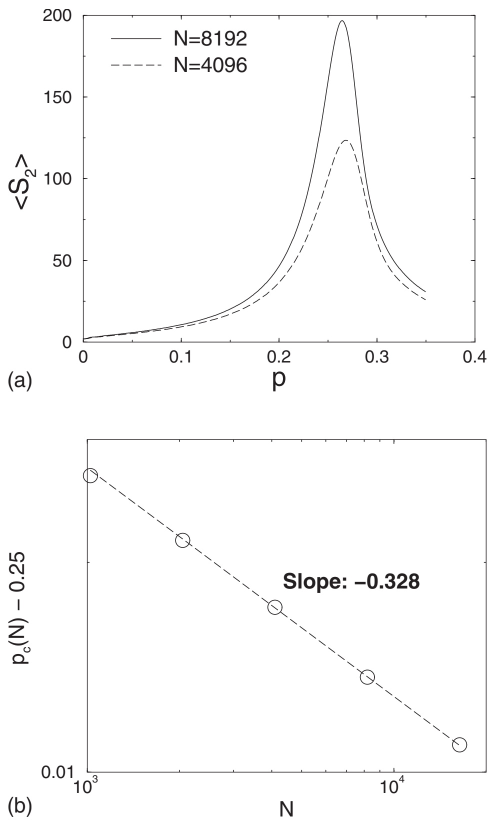

FIG. 1. (a) The average size of the second largest cluster $\langle S_2 \rangle$ as a function of the concentration p of links present in the ER networks. The typical number of realizations for each curve is 10^6. (b) Log-log plot of $p_c(N)-p_c(\infty)$ as a function of N, where $p_c(\infty)=1/\langle k \rangle=0.25$ for ER with $\langle k \rangle=4$.

geometrically close nodes. In this case it is expected that percolation in such networks will become similar to percolation in mean field networks, in terms of the critical exponents when the embedding dimension is at least $d \geqslant d_c$.

An alternative approach to the critical dimensions, requiring no change in the network generation model, starts with a mean field, infinite dimensional network. Then, percolation is performed on this network. When the percolation threshold is reached, one attempts to embed the critical percolation clusters in a finite dimensional space, where neighboring nodes in the cluster are connected by a link of distance one in the lattice. In this case, it is expected that the minimum dimension needed to allow such an embedding is the upper critical dimension, d_c, and that the fractal dimension of the embedded cluster will be d_f. For $d \geqslant d_c$ the relation between the geometrical distance R between two sites on the cluster and the shortest paths is similar to a random walk, i.e., $R^2 \sim \ell$. This is since space limitations for $d \geqslant d_c$ do not appear.

Thus, it is reasonable that when λ is smaller, the network is more complex (due to bigger hubs) and a higher upper critical dimension is expected. However, Eq. (1), that was shown analytically to be valid for $N \to \infty$ was never verified or tested numerically. It is also interesting to determine the range of N values where the results of Eq. (1) can be observed. Here we propose two numerical methods to measure directly the value of d_c for ER and SF networks with $\lambda > 3$ [17].

TABLE I. The main results for SF and ER networks. The critical percolation threshold $p_c(\infty)$ indicates the numerical value calculated according to Eqs. (6) and (7). Theoretical Θ is the theoretical prediction of Θ [from Eqs. (1) and (3)]. Numerical Θ_1 and Θ_2 are the numerical value we obtained from simulations using the two methods. The SF networks were generated with $k_{\min}=2$.

λ	$p_c(\infty)$	Theoretical Θ	Numerical Θ_1	Numerical Θ_2
3.30	0.1271	0.130		0.140
3.50	0.2039	0.200	0.234	0.192
3.65	0.2574	0.245	0.260	
3.75	0.2911	0.273	0.275	0.246
3.85	0.3234	0.298	0.284	0.263
4.50	0.5009	1/3	0.326	
ER ($\langle k \rangle=4$)	0.25	1/3	0.328	0.335

Method I: Finite-size scaling arguments in d-dimensional lattice networks predict [9,10] that the critical threshold $p_c(L)$ approaches $p_c \equiv p_c(\infty)$ via

$$p_c(L) - p_c(\infty) \sim L^{-1/\nu}, \tag{2}$$

where L is the linear lattice size and ν is the correlation critical exponent. Equation (2) for lattices can be generalized to networks of N nodes via the relation $L^d=N$, i.e., $p_c(N) - p_c(\infty) \sim N^{(-1/d\nu)}$. Since networks can be regarded as embedded in infinite dimension and since above d_c all exponents are the same, we replace d by d_c,

$$p_c(N) - p_c(\infty) \sim N^{-1/d_c\nu} \equiv N^{-\Theta_1}. \tag{3}$$

For ER and SF networks with $\lambda > 4$, we have $d_c=6$ and $\nu=1/2$, thus from Eq. (3) follows

$$p_c(N) - p_c(\infty) \sim N^{-1/3}. \tag{4}$$

For SF networks with $3 < \lambda < 4$, we have $\nu=1/2$ and substituting Eq. (1) in Eq. (3), it yield

$$p_c(N) - p_c(\infty) \sim N^{(3-\lambda)/(\lambda-1)}. \tag{5}$$

We denote $\Theta_1 \equiv 2/d_c$ as the theoretical value of method I to distinguish between method I and method II described below. To measure Θ_1, using the finite size scaling of Eq. (3), we have to compute the dependence of the percolation threshold $p_c(N)$ of ER and SF networks on the system size N. To calculate $p_c(N)$, we apply the second largest cluster method [9,10], which is based on determining $p_c(N)$ by measuring the value of p_c at the maximum value of the average size of the second largest cluster $\langle S_2 \rangle$. It is known that $\langle S_2 \rangle$ has a sharp peak as a function of p at p_c [9,10]. To detect this peak we perform a Gaussian fit around the peak and estimate the peak position which is $p_c(N)$ [18].

To improve the speed of the simulations, we implement the fast Monte Carlo algorithm for percolation proposed by Newman and Ziff [19]. Basically, for each realization, we prepare one instance of N nodes network with the desired structure as the reference network. Then we prepare another set of N nodes with no links as our target network. To keep track of the size of the second largest cluster instead of the

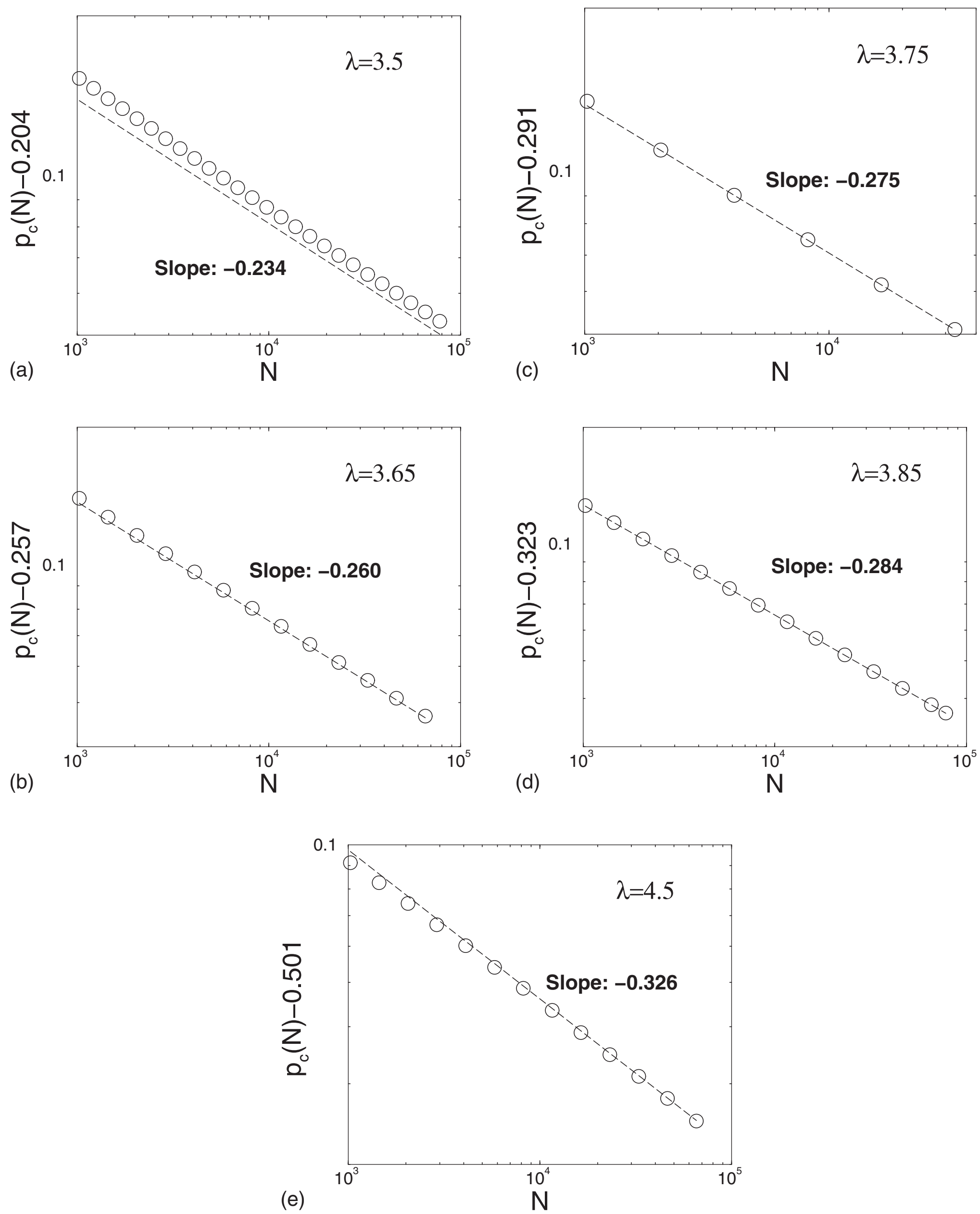

FIG. 2. Log-log plots of $p_c(N)-p_c(\infty)$ as a function of N for SF networks with $k_{\min}=2$ and different value of λ. The dashed line is the reference line with indicated slope.

largest one, we use a sorted list of all the clusters in descending order according to their sizes. In the beginning is a list of N clusters of size one. As we choose the links in random order from the reference network and make the connection in the target network, we update the list of the cluster size but always keep them in descending order. The concentration value, p, of each newly connected link is calculated by the number of links after adding this link in the target network divided by the total number of links in the reference network. When each link is connected, we record S_2 at the concentration value p of this newly connected link. We divide the range 0 to 1 into 1000 bins. After many realizations, we take the average of S_2 for each bin.

Figure 1(a) shows $\langle S_2\rangle$ as a function of p, for two different system sizes of ER networks with $\langle k\rangle=4$. The position of the peak, obtained by fitting the peak with a Gaussian function, yields $p_c(N)$. Figure 1(b) shows $p_c(N)$ as a function N. Using $p_c(\infty)\equiv 1/\langle k\rangle=0.25$ [6,7], the fitting of Eq. (3) gives the exponent $\Theta_1=0.328\pm 0.003$, very close to the theoretical prediction for ER, $\Theta=1/3$, Eq. (4). We performed the same simulations for ER with other average degrees, $\langle k\rangle=5$ and 6, and obtained similar results for Θ_1.

To determine $p_c(\infty)$ for random SF networks, we use the exact analytical results [20]

$$p_c(\infty)\equiv\frac{1}{\kappa_0-1}. \qquad (6)$$

Here $\kappa_0\equiv\langle k_0^2\rangle/\langle k_0\rangle$ is computed from the original degree distribution $(\mathcal{P}(k_0))$ for which the network is constructed. How-

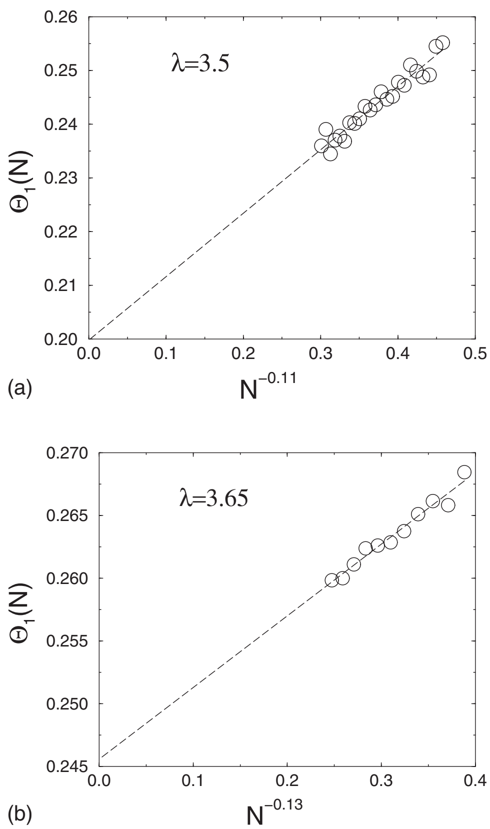

FIG. 3. The exponent $\Theta_1(N)$ as a function of N^{-x} for SF networks with $k_{\min}=2$ and different value of λ: (a) λ=3.5, where $x \approx 0.11$ and (b) λ=3.65, where $x \approx 0.13$. The theoretical values $\Theta=0.2$ (λ=3.5) and $\Theta=0.245$ (λ=3.65), are consistent with the asymptotic values of $\Theta_1(N)$ obtained for $N \to \infty$.

ever, the way to compute the value of κ_0 is strongly affected by the algorithm of generating the SF network as explained below.

To generate SF networks with power law exponent λ, we use the configuration model algorithm [21–23]. We first generate a series of random real number u satisfying the distribution $\mathcal{P}(u)=cu^{-\lambda}$, where $c=(\lambda-1)/k_{\min}^{1-\lambda}$ is the normalization factor. Next we truncate the real number u to be an integer number k, which we assume to be the degree of a node. We make k copies of each node according to its degree and randomly choose two nodes and connect them by a link. Notice that the process of truncating the real number u to be an integer number k which is the degree of a node actually slightly changes the degree distribution because any real number $n \leq u < n+1$, where n is an integer number, will be truncated to be equal n. Thus, the actual degree distribution we obtain using this algorithm is

$$\mathcal{P}(k) = \int_k^{k+1} cu^{-\lambda}du = \frac{1}{k_{\min}^{1-\lambda}}[k^{1-\lambda} - (k+1)^{1-\lambda}]. \qquad (7)$$

We use Eq. (7) to compute κ_0 and $p_c(\infty)$ defined in Eq. (6). Table I shows the calculated results of $p_c(\infty)$ for several values of λ.

We calculate $\langle S_2 \rangle$ for SF networks for different values of

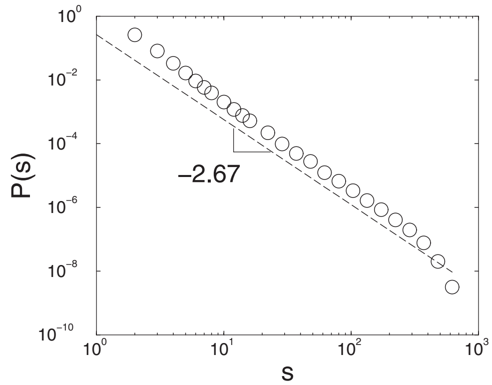

FIG. 5. The probability distribution of the cluster sizes at $p_c(N)$ for $N=2048$ (○) and $N=16\,384$ (□). The dashed line is the reference line with slope −2.67.

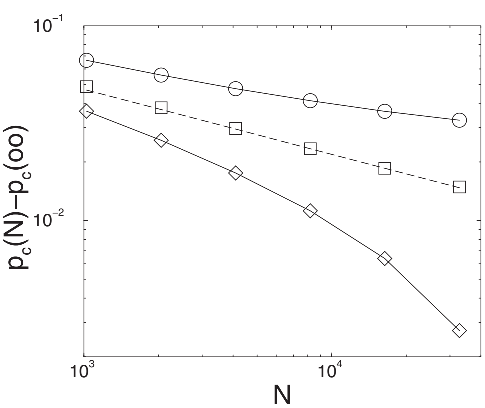

FIG. 4. Log-log plot of $p_c(N)-p_c(\infty)$ as a function of N for SF networks with λ=2.5, $k_{\min}=2$ for a targeted attack. The dashed line is the best fit with slope −0.33. Since we do not have a good estimation for $p_c(\infty)$, we modified $p_c(\infty)$ to get the best straight line in log-log plot, $p_c(\infty)=0.23$ (○), $p_c(\infty)=0.25$ (□), and $p_c(\infty)=0.26$ (◇). When $p_c(N)-p_c(\infty)$ is linear (dashed line) in the log-log plot, the slope yields the exponent $\Theta \approx 0.33$, i.e., $d_c=6$.

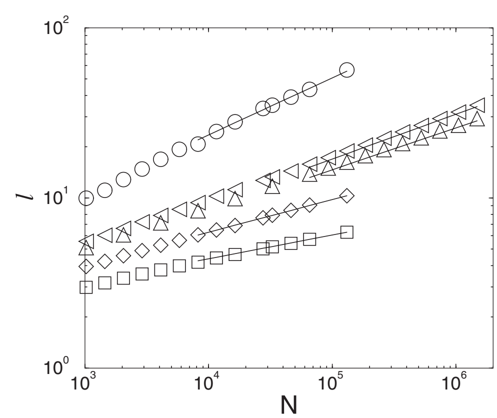

FIG. 6. Log-log plot of the average distance of the IIC, ℓ as a function of the system size N. Results are for ER with $\langle k \rangle=4$ (○); SF with λ=3.3 (□), λ=3.5 (◇), λ=3.75 △), λ=3.85 (◁). The best fit for large N is shown as the solid lines and the exponents obtained from the slope of the best fit are 0.335 (1/3, ○), 0.14 (0.13, □), 0.192 (0.2, ◇), 0.246 (0.273, △), and 0.263 (0.298, ◁). The first value in the parentheses is the expected theoretical value.

λ and N and compute $p_c(N)$ by fitting with a Gaussian function near the peak of $\langle S_2 \rangle$ as for ER networks. Using the values of $p_c(\infty)$ for SF networks displayed in Table I, we obtain Θ_1 by a power law fitting with Eq. (3) as shown in Fig. 2. As we can see for $\lambda=4.5$, 3.85 and 3.75 we obtain quite good agreement with the theoretical values. However for $\lambda=3.65$ and 3.5, the values of Θ_1 become better when fitting only the last several points (largest N) and still have large deviations from their theoretical values. This strong finite size effect is probably since for $\lambda \to 3$ the largest percolation cluster at the criticality becomes smaller [24]. Thus, we expect that as N increase, the exponent $\Theta_1(N)$ obtained by simulations should approach the theoretical value of Θ_1 of Eq. (5). To better estimate Θ_1 we assume finite size corrections to scaling for Eq. (5), i.e.,

$$p_c(N) - p_c(\infty) \sim N^{-\Theta_1}(1 + N^{-x}). \tag{8}$$

Thus, the actual $\Theta_1(N)$ obtained from simulation is the successive slopes

$$\Theta_1(N) \equiv -\partial \ln[p_c(N) - p_c(\infty)]/\partial(\ln N), \tag{9}$$

from which we can see that $\Theta_1(N)$ approaches Θ_1 as a power law

$$\Theta_1(N) - \Theta_1 \sim N^{-x}. \tag{10}$$

Indeed, Fig. 3 shows the exponent $\Theta_1(N)$ as a function of N^{-x} for $\lambda=3.5$ and 3.65. Figure 3(a) shows that for $\lambda=3.5$ and $x=0.11$, we obtain a straight line and $\Theta_1(N)$ approaches 0.2 as $N \to \infty$, consistent with the theoretical value of Θ (Table I). Figure 3(b) shows, for $\lambda=3.65$ and $x=0.13$, $\Theta_1(N)$ is again a straight line that approaches 0.245 for $N \to \infty$, consistent with the theory.

Next we estimate the value of d_c for SF network under targeted attack on the largest degree nodes [25–28]. For this case since the hubs are removed we expect that for all $\lambda > 2$, d_c will be the same as for ER, i.e., $d_c=6$. In Fig. 4, we plot $p_c(N) - p_c(\infty)$ for SF with $\lambda=2.5$ under targeted attack. Indeed from Eq. (3) by changing $p_c(\infty)$ and fitting the best straight line in log-log plot, we obtain $\Theta_1 \approx 0.33$, i.e., $d_c \approx 6$, as expected.

Further supports of the analytical approach, we evaluate by simulations $\mathcal{P}(s)$, the probability distribution of the cluster sizes at $p_c(N)$, which should follow a power law for SF networks [11],

$$\mathcal{P}(s) \sim s^{-\tau} = s^{-[2+1/(\lambda-2)]}, \quad 2 < \lambda < 4. \tag{11}$$

Figure 5 shows the simulations results for SF networks $\lambda = 3.5$. The dashed line is the reference line with slope -2.67, which is the theoretical value of τ from Eq. (5), showing good agreement between theory and simulations.

Method II: We calculate the upper critical dimension using an alternative approach. Since at the upper critical dimension $N \sim R^{d_c}$ and $\ell \sim R^2$, where R is the geometrical distance, N is the system size and ℓ is the average distance among all the pairs of nodes in the infinite incipient cluster at criticality p_c, we obtain

$$l \sim N^{2/d_c} \sim N^{\Theta_2}. \tag{12}$$

Thus measuring ℓ versus N yields $\Theta_2 \equiv 2/d_c$ from which d_c is directly obtained. Using $p_c(\infty)$ from Table I, we bomb $1 - p_c$ fraction of links randomly and calculate the average distance of all the pairs from the largest cluster in the remained network, which is the infinite incipient cluster. For each system size, we uses 10 000 realizations. Figure 6 shows the results and the exponents obtained using method II are also shown in Table I. Comparing the results from these two methods in Table I, we can see that Θ_1 is closer to the theoretical values when λ approaches 3 and Θ_2 is closer when λ approaches 4, which suggests different behaviors of finite size effects.

We thank ONR, ONR-Global, UNMdP, NEST Project No. DYSONET012911, and the Israel Science Foundation for support. L.A.B. thanks FONCyT (Grant No. PICT-O 2004/370) for financial support. C.L. thanks Conicet for financial support.

[1] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[2] R. Pastor-Satorras and A. Vespignani, *Evolution and Structure of the Internet: A Statistical Physics Approach* (Cambridge University Press, Cambridge, 2004).
[3] S. N. Dorogovtsev and J. F. F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford University Press, Oxford, 2003).
[4] R. Cohen and S. Havlin (unpublished).
[5] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[6] P. Erdős and A. Rényi, Publ. Math. (Debrecen) **6**, 290 (1959).
[7] P. Erdős and A. Rényi, Publ. Math. (Debrecen) **5**, 17 (1960).
[8] B. Bollobas, *Random Graphs* (Cambridge University Press, Cambridge, 2001).
[9] *Fractals and Disordered Systems*, edited by A. Bunde and S. Havlin (Springer, New York, 1996).
[10] D. Stauffer and A. Aharony, *Introduction to Percolation Theory* (Taylor & Francis, New York, 1994).
[11] R. Cohen, Daniel ben-Avraham, and S. Havlin, Phys. Rev. E **66**, 036113 (2002).
[12] R. Cohen, S. Havlin, and Daniel ben-Avraham, in *Handbook of Graphs and Networks*, edited by S. Bornholdt and H. G. Schuster (Wiley-VCH, Berlin, 2002), Chap. 4.
[13] R. Cohen and S. Havlin, Physica A **336**, 6 (2004).
[14] S. S. Manna, G. Mukherjee, and Parongama Sen, Phys. Rev. E **69**, 017102 (2004).
[15] A. F. Rozenfeld, R. Cohen, D. ben-Avraham, and S. Havlin, Phys. Rev. Lett. **89**, 218701 (2002).
[16] C. P. Warren, L. M. Sander, and I. M. Sokolov, Phys. Rev. E **66**, 056105 (2002).
[17] The case of $2 < \lambda < 3$ is more complex since there is no finite percolation threshold ($p_c \to 0$). However, since for $\lambda \to 3$, d_c

$\to\infty$, Eq. (1) suggests that for $2<\lambda\leqslant 3$ there is no finite upper critical dimension but $d_c=\infty$ for all $\lambda\leqslant 3$.
[18] The finite size dependency of Eq. (3) cannot be obtained from the condition $\kappa\equiv\langle k_0^2\rangle/\langle k_0\rangle=2$, since this formula is valid only in the limit $N\to\infty$.
[19] M. E. J. Newman and R. M. Ziff, Phys. Rev. E **64**, 016706 (2001).
[20] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, Phys. Rev. Lett. **85**, 4626 (2000).
[21] B. Bollobas, Eur. J. Comb. **1**, 311 (1980).
[22] M. Molloy and B. A. Reed, Combinatorics, Probab. Comput. **7**, 295 (1998).
[23] M. Molloy and B. A. Reed, Combinatorics, Probab. Comput. **6**, 161 (1995).
[24] The largest percolation cluster at the criticality was analytically predicted as $\langle S\rangle\sim N^{(\lambda-3)/(\lambda-1)}$ [11].
[25] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. Lett. **85**, 5468 (2000).
[26] R. Cohen, K. Erez, Daniel ben-Avraham, and S. Havlin, Phys. Rev. Lett. **86**, 3682 (2001).
[27] Z. Wu, L. A. Braunstein, V. Colizza, R. Cohen, S. Havlin, and H. E. Stanley, Phys. Rev. E **74**, 056104 (2006).
[28] L. K. Gallos, R. Cohen, P. Argyrakis, A. Bunde, and S. Havlin, Phys. Rev. Lett. **94**, 188701 (2005).