

# Optimal paths in complex networks with correlated weights: The worldwide airport network

Zhenhua Wu,<sup>1</sup> Lidia A. Braunstein,<sup>1,2</sup> Vittoria Colizza,<sup>3</sup> Reuven Cohen,<sup>4</sup> Shlomo Havlin,<sup>4</sup> and H. Eugene Stanley<sup>1</sup>

<sup>1</sup>Center for Polymer Studies, Boston University, Boston, Massachusetts 02215, USA

<sup>2</sup>Departamento de Física, Facultad de Ciencias Exactas y Naturales, Universidad Nacional de Mar del Plata, Funes 3350, 7600 Mar del Plata, Argentina

<sup>3</sup>School of Informatics and Department of Physics, Indiana University, Bloomington, Indiana 47406, USA

<sup>4</sup>Minerva Center and Department of Physics, Bar-Ilan University, Ramat Gan, Israel

(Received 2 August 2006; published 6 November 2006)

We study complex networks with weights  $w_{ij}$  associated with each link connecting node  $i$  and  $j$ . The weights are chosen to be correlated with the network topology in the form found in two real world examples: (a) the worldwide airport network and (b) the *E. Coli* metabolic network. Here  $w_{ij} \sim x_{ij}(k_i k_j)^\alpha$ , where  $k_i$  and  $k_j$  are the degrees of nodes  $i$  and  $j$ ,  $x_{ij}$  is a random number, and  $\alpha$  represents the strength of the correlations. The case  $\alpha > 0$  represents correlation between weights and degree, while  $\alpha < 0$  represents anticorrelation and the case  $\alpha = 0$  reduces to the case of no correlations. We study the scaling of the lengths of the optimal paths,  $\ell_{\text{opt}}$ , with the system size  $N$  in strong disorder for scale-free networks for different  $\alpha$ . We find two different universality classes for  $\ell_{\text{opt}}$  in strong disorder depending on  $\alpha$ : (i) if  $\alpha > 0$ , then for  $\lambda > 2$  the scaling law  $\ell_{\text{opt}} \sim N^{1/\lambda}$ , where  $\lambda$  is the power-law exponent of the degree distribution of scale-free networks, and (ii) if  $\alpha \leq 0$ , then  $\ell_{\text{opt}} \sim N^{\nu_{\text{opt}}}$  with  $\nu_{\text{opt}}$  identical to its value for the uncorrelated case  $\alpha = 0$ . We calculate the robustness of correlated scale-free networks with different  $\alpha$  and find the networks with  $\alpha < 0$  to be the most robust networks when compared to the other values of  $\alpha$ . We propose an analytical method to study percolation phenomena on networks with this kind of correlation, and our numerical results suggest that for scale-free networks with  $\alpha < 0$ , the percolation threshold  $p_c$  is finite for  $\lambda > 3$ , which belongs to the same universality class as  $\alpha = 0$ . We compare our simulation results with the real worldwide airport network, and we find good agreement.

DOI: [10.1103/PhysRevE.74.056104](https://doi.org/10.1103/PhysRevE.74.056104)

PACS number(s): 89.75.Hc

## I. INTRODUCTION

Recently attention has 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 biological and social interaction systems [1–3]. Network structure systems are visualized by nodes representing individuals, organizations, or computers and by links between them representing their interactions. Significant topological features were discovered, such as the clustering and small-world properties [4]. There exists evidence that many real networks possess a scale-free (SF) degree distribution characterized by a power-law tail given by  $P(k) \sim k^{-\lambda}$ , where  $k$  is the degree of a node and  $\lambda$  measures the broadness of the distribution [5]. In most studies, all links or nodes in the network are regarded as identical. Thus the topological structure of the network determines all the other properties of the network, such as robustness, percolation threshold [6,7], the average shortest path length,  $\ell_{\text{min}}$  [8], and transport [9].

In many real world networks the links are not equally weighted. For example, the links between computers in the Internet network have different capacities or bandwidths and the airline network links between different pairs of cities have different numbers of passengers. To better understand real networks, several studies have been carried out on weighted network models [10–19]. For weighted networks, an important quantity that characterizes the information flow is the optimal path, which is the path that minimizes the total weight along the path [15–21]. Optimal paths play important roles in many dynamic systems such as current flow of random resistor networks, where the dynamics of current flow is

strongly controlled by the optimal path [18,19]. The above studies assume that the weights on the links are purely random—i.e., uncorrelated with the topology of the network.

However, recently several studies of networks with weights on the links, such as the worldwide airport network (WAN) and the *E. coli* metabolic network [22–27], show that the weights are correlated with the network topology [22,24]. For the WAN, each link is a direct flight between two airports  $i$  and  $j$  and the weight  $w_{ij}$  is the number of passengers between them during a period of time. The mean traffic on links can be characterized by  $\langle T_{ij} \rangle \sim \langle k_i k_j \rangle^\theta$ , where  $k_i$  and  $k_j$  are the degrees of nodes  $i$  and  $j$  and  $\theta > 0$  [24,27].

In this paper, we study how the correlations between the topology and weights affect the robustness, the percolation threshold, the scaling of the optimal path length, and the minimum spanning tree by studying the weighted SF model. The WAN was found to be a SF network with  $\lambda \approx 2$  [22,26]. We model the dynamic transport process of the WAN as a SF network with correlated weights representing the number of passengers [22]. We study robustness and scaling of the optimal paths and find good agreement with the real WAN results.

## II. NETWORKS WITH TOPOLOGICALLY CORRELATED WEIGHTS

### A. Network model with topologically correlated weights

In studies on weighted networks such as the WAN and *E. Coli* metabolic network, the weights on the links represent different quantities. The weight in the *E. Coli* metabolic network represents the flux of a link, i.e., the relative activity of

the reaction on that link [23]. For both real networks, the weight associated with a link measures the preference level of that link. The higher the weight is, the easier it is to go through that link. In this sense, the inverse of the weight on both networks is actually a plausible evaluation of the “cost” to traverse that link. In the WAN, the “cost” includes the airfare, time, convenience, etc. We assume that the higher is the cost on a link, the less traffic it has.

From this perspective, we apply the optimization problem [17] to the SF network with generalized correlated weights [24,28]. To each link connecting a pair of nodes  $i$  and  $j$  in a SF network we assign a weight  $w_{ij}$ , representing the cost to transverse that link, with the form

$$w_{ij} \equiv x_{ij}(k_i k_j)^\alpha, \quad (1)$$

where  $x_{ij}$  is a random number,  $\alpha$  is a parameter that controls the strength of the correlation between the topology and the weight, and  $k_i$  and  $k_j$  are the degree of node  $i$  and node  $j$ . A random number  $x_{ij}$  could be chosen to mimic the statistical distribution of the real network, and in this paper  $x_{ij}$  is taken from a uniform distribution between 0 and 1 [29]. Here we will be interested in the entire range of  $\alpha$ .

The minimum spanning tree (MST) of SF networks with correlation of Eq. (1) represents the structure carrying the maximum total traffic in a real WAN and the skeleton of the most used paths in the *E. Coli* metabolic network. The MST is the tree spanning all nodes of the network with the minimum total weight. With weights according to Eq. (1) and  $\alpha = -0.5$ , the MST is the same as the tree that maximizes the total traffic out of all possible spanning trees of the WAN because the structure of the MST is only determined by the relative order of the links according to the weight, which is preserved under the inverse transformation. For this reason, in our simulation, we only show results for  $\alpha = 1, 0, -1$  representing  $\alpha > 0$ ,  $\alpha = 0$ , and  $\alpha < 0$  [24,30]. As an optimized tree, the MST playing the role of the network skeleton is widely used in different fields, such as the design and operation of communication networks, the traveling salesman problem, the protein interaction problem, optimal traffic flow, and economic networks [24,31–36]. Thus studying the effect of correlations of the type of Eq. (1) on the structure of the MST may explain the transport on such weighted networks and possibly will lead to a better understanding of the origin of such correlations in real networks. Moreover, the MST is the union of all optimal paths in the strong disorder (SD) limit [20], where one link on a path dominates the total weight of the path. Thus, the scaling of the optimal paths also reveals an important aspect of the structure of the MST.

### B. Scaling of the length of the optimal path

The case  $\alpha = 0$  represents the uncorrelated case, which is studied in Ref. [17]. In the SD regime, the total weight of a path is controlled by a single link with the highest cost on that path [15–21]. For uncorrelated SF networks, the length of the optimal path in SD,  $\ell_{\text{opt}}$ , scales with  $N$  as [17]

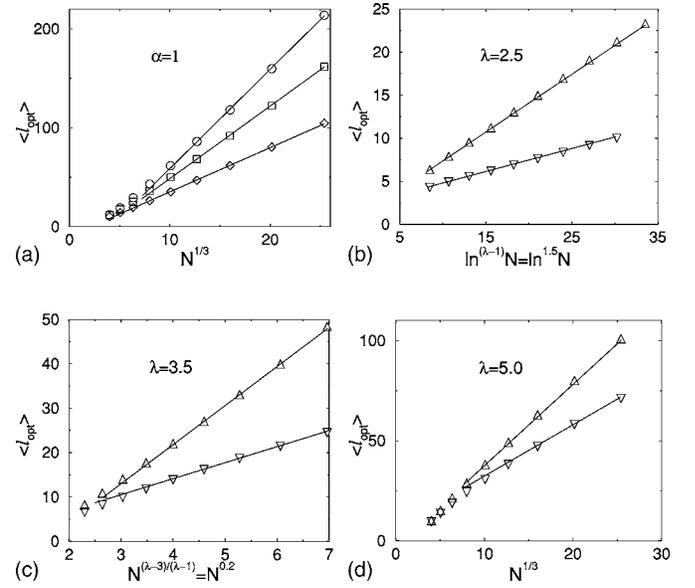


FIG. 1.  $\langle \ell_{\text{opt}} \rangle$  as a function of (a)  $N^{1/3}$ , (b)  $\ln^{(\lambda-1)} N$ , (c)  $N^{(\lambda-3)/(\lambda-1)}$ , and (d)  $N^{1/3}$  in SD for (a)  $\alpha = 1$  with  $\lambda = 5$  ( $\circ$ ),  $\lambda = 3.5$  ( $\square$ ), and  $\lambda = 2.5$  ( $\diamond$ ); (b)  $\lambda = 2.5$  with  $\alpha = 0$  ( $\triangle$ ) and  $\alpha = -1$  ( $\nabla$ ); (c)  $\lambda = 3.5$  with  $\alpha = 0$  ( $\triangle$ ) and  $\alpha = -1$  ( $\nabla$ ); and (d)  $\lambda = 5.0$  with  $\alpha = 0$  ( $\triangle$ ) and  $\alpha = -1$  ( $\nabla$ ).

$$\ell_{\text{opt}} \sim \begin{cases} N^{\nu_{\text{opt}}}, & \lambda > 3, \\ \ln^{\lambda-1} N, & 2 < \lambda \leq 3, \end{cases} \quad (2)$$

with  $\nu_{\text{opt}} = 1/3$  for  $\lambda \geq 4$  and  $\nu_{\text{opt}} = (\lambda - 3)/(\lambda - 1)$  for  $3 < \lambda < 4$ .

From the definition of the weights [see Eq. (1)], in the case  $\alpha \neq 0$ , we expect that correlations will affect the links connected to the high degree nodes (hubs). The optimal paths behave either as *hub-phobic* or as *hub-philic* depending on whether  $\alpha$  is positive or negative. Hub-phobic ( $\alpha > 0$ ) means that the optimal paths dislike to go through hubs because the cost is higher. Hub-philic ( $\alpha < 0$ ) means that the optimal paths like to go through hubs because they cost less. Thus  $\alpha$  is a parameter controlling the importance level of hubs in the optimized transport process. Due to the importance of the hubs in SF networks, we expect that the scaling of  $\ell_{\text{opt}}$  will be affected by such correlations. We expect in the hub-phobic regime that the optimal paths should be topologically similar to an Erdős-Rényi (ER) graph [17], where there are almost no hubs. Thus  $\ell_{\text{opt}}$  will scale with  $N$  with exponent  $1/3$  as in ER networks [17] for all the values of  $\lambda > 2$ . On the other hand, in the hub-philic regime, the hubs are preferred and the optimal paths will avoid small or intermediate degrees, which are less important for the SF networks and have minor influence on the optimal paths. Thus in this case we expect that  $\ell_{\text{opt}}$  will scale similarly to the uncorrelated case.

The SD limit can be reached by approaching absolute zero temperature if passing through a link is regarded as an activation process with a random activation energy  $\varepsilon_{ij}$  and  $w_{ij} \equiv \exp(\beta \varepsilon_{ij})$ , where  $\beta$  is the inverse temperature, since for  $\beta \rightarrow \infty$  a single link in the path dominates the sum  $\sum w_{ij}$  along the path. For a general form of weight, which is not exponential but rather such as Eq. (1), the SD limit can be reached

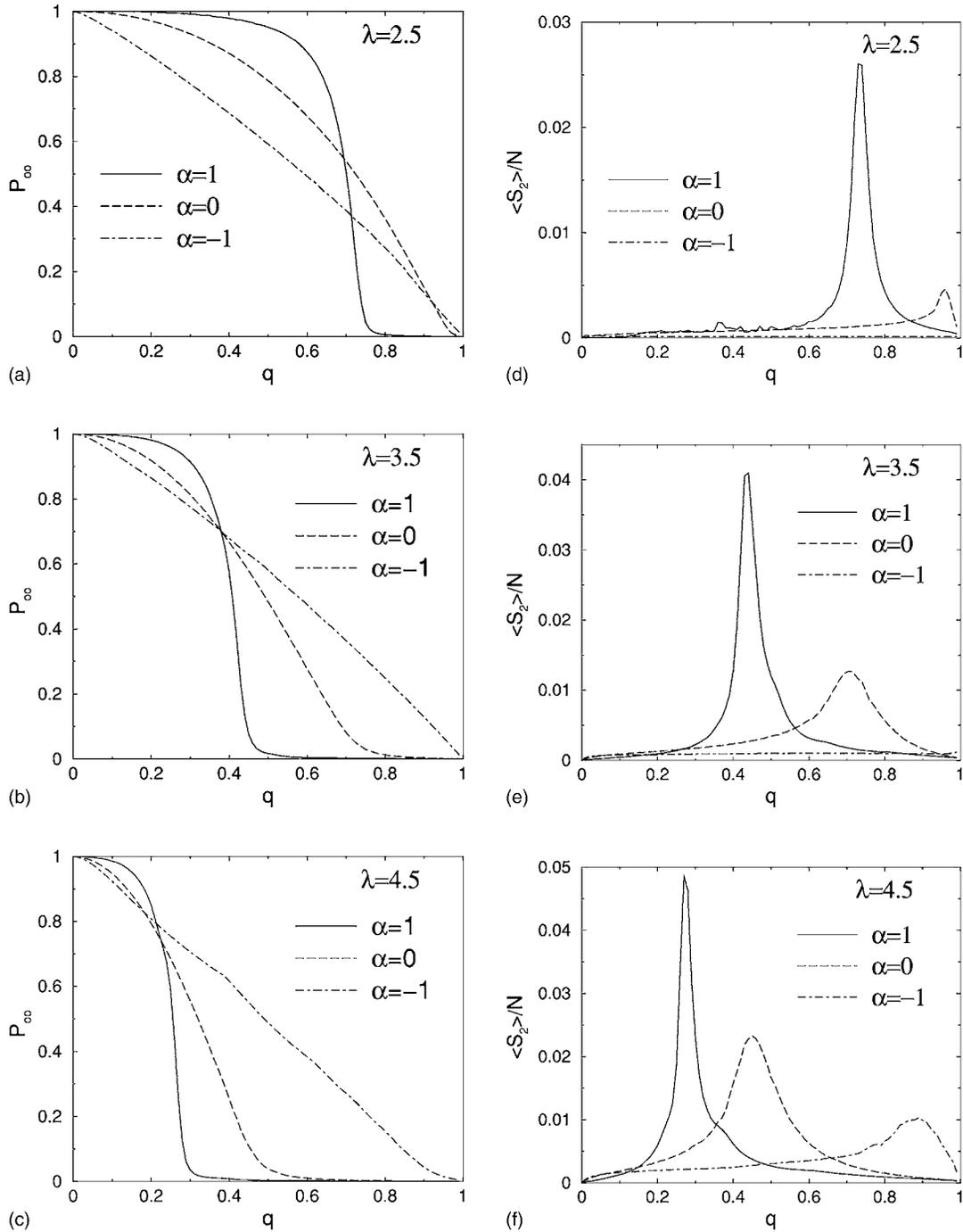


FIG. 2. Left column (a)–(c):  $P_\infty$  as a function of  $q$ , the concentration of removed links with (a)  $\lambda=2.5$ , (b)  $\lambda=3.5$ , and (c)  $\lambda=4.5$ . Right column (d)–(f):  $\langle S_2 \rangle / N$  as a function of  $q$  with (d)  $\lambda=2.5$ , (e)  $\lambda=3.5$ , and (f)  $\lambda=4.5$ .

by generating the MST because it provides the optimal path between any two nodes of a network in the SD limit [20]. In this paper, we use the MST to compute the optimal paths in the SD limit.

To model the WAN, we generate the SF networks using the Molloy-Reed algorithm with the constraint of disallowing parallel links (two or more links connecting the same two nodes) and self-loops (node connecting with itself) [37,38]. Then for each link  $ij$  in the network, we assign a weight according to Eq. (1).

We build the MST on the largest connected component of a graph using Prim's algorithm [39]. The tree starts from any node in the largest connected component and grows to the nearest neighbor of the tree with the minimum cost. This process ends when the MST includes all the nodes of the largest connected component. This process is analogous to invasion percolation used in physics [40]. As the MST is a tree, there is only a single path between any pair of nodes, which is the optimal path in the SD limit [20].

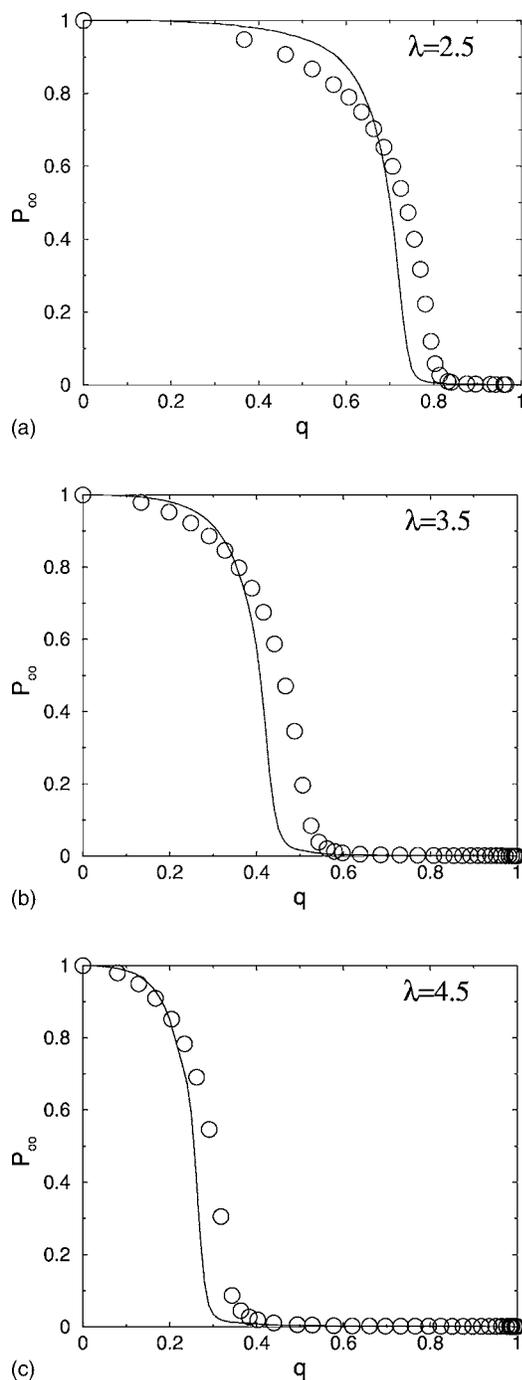


FIG. 3.  $P_\infty$  as a function of  $q$  with (a)  $\lambda=2.5$ , (b)  $\lambda=3.5$ , and (c)  $\lambda=4.5$  for strategy I ( $\circ$ ), attacking nodes according to degree of nodes, and strategy II (solid line), attacking links according to its weight for the SF networks in the hub-phobic regime ( $\alpha > 0$ ).

An equivalent algorithm to find the MST is the “bombing algorithm” [15,17], where we start with the full network and remove links in descending order of the weight only when the removal of a link does not disconnect the graph. The algorithm ends, and the MST is obtained, when no more links can be removed without disconnecting the graph. Then we calculate the mean value of  $\ell_{\text{opt}}$  between a pair of randomly chosen nodes and average over many pairs and many realizations.

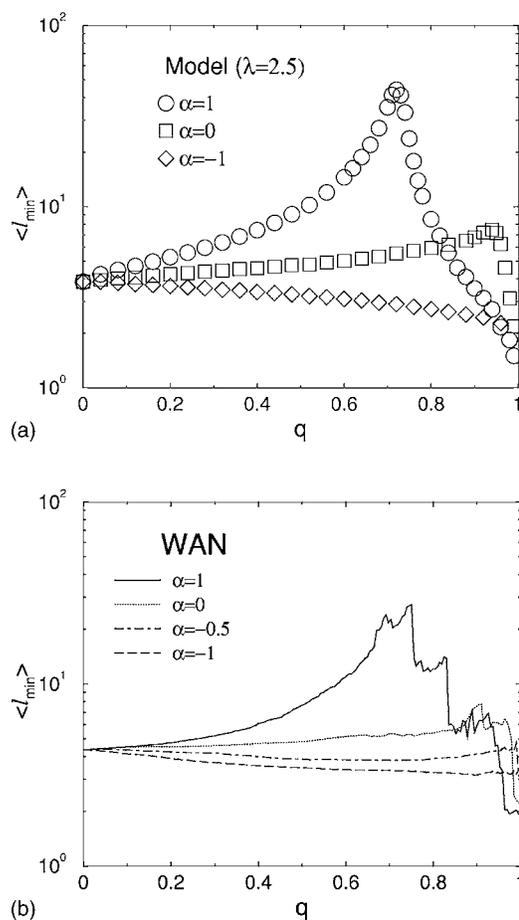


FIG. 4. (a)  $\langle \ell_{\text{min}} \rangle$ , the average length of the shortest path as a function of  $q$ , the concentration of removed links in descending order of the weight for simulated SF networks with  $\alpha=1$  ( $\circ$ ),  $\alpha=0$  ( $\square$ ), and  $\alpha=-1$  ( $\diamond$ ). (b) The same calculation for the real WAN with  $\alpha=1$  (solid line),  $\alpha=0$  (dotted line),  $\alpha=-0.5$  (dot-dashed line), and  $\alpha=-1$  (dashed line).

Figure 1(a) shows  $\ell_{\text{opt}}$  as a function of  $N^{1/3}$  for several values of  $\lambda > 2$  with  $\alpha=1$ . The linear behavior supports the expected scaling for the SF networks in the hub-phobic regime ( $\alpha > 0$ ):

$$\ell_{\text{opt}} \sim N^{1/3} \quad [\lambda > 2, \alpha > 0]. \quad (3)$$

Equation (3) supports our hypothesis that if the hubs are dynamically avoided, the SF networks behave the same as an ER networks.

In the hub-philic regime ( $\alpha < 0$ ), we find that  $\ell_{\text{opt}}$  scales the same as for the uncorrelated case ( $\alpha=0$ ). In Figs. 1(b)–1(d), we plot the scaling behavior of  $\ell_{\text{opt}}$  as a function of  $N$  according to the scaling found for the uncorrelated case [17]. Our numerical results suggest that the scaling behavior of  $\ell_{\text{opt}}$  in the hub-philic regime has the same scaling form of Eq. (2) as the uncorrelated case. However, in the hub-philic regime,  $\ell_{\text{opt}}$  is much smaller than that of the uncorrelated case, which shows the advantage for the optimal paths to pass through the hubs.

### C. Robustness

Percolation properties of random networks with given degree distributions have been extensively studied [6,7,41–43]. One of the most striking results [6], due to its important implications, is the absence of a percolation threshold in the uncorrelated random SF networks with  $\lambda < 3$ . In other words, in this type of network, one has to remove almost all nodes before the network collapses into disconnected components because the few hubs always keep the remaining network connected [6]. Translated into epidemic context, it means that an epidemic threshold below which the epidemics cannot propagate approaches zero as  $N \rightarrow \infty$  [44].

Robustness of a network can be characterized by the fraction of the links or nodes one has to remove in order to disconnect the whole network. Nevertheless, almost all the analytical results on robustness obtained up to now implicitly refer to uncorrelated networks and little is known about the effects of topologically correlated weights on the percolation properties of networks. To test the robustness in weighted correlated networks, we calculate the ratio  $P_\infty \equiv \langle S \rangle / N$  as a function of  $q$ , where  $\langle S \rangle$  is the average mass of the largest remaining cluster,  $N$  is the mass of the whole network, and  $q$  is the fraction of removed links.  $P_\infty$  basically is the probability to find a node belonging to the giant component. In order to compute  $P_\infty$  after building the network, we assign weights on links according to Eq. (1). Then we remove a fraction  $q$  of links in descending order of weights and calculate  $P_\infty$  through  $\langle S \rangle / N$ . In order to identify the percolation threshold  $p_c = 1 - q_c$  we compute also the average mass of the second largest component  $\langle S_2 \rangle$  and estimate  $q_c$  from its maximum value [40].

In Fig. 2 we show the results for networks with  $N=8192$  and different values of  $\lambda$  and  $\alpha$ . The peak of  $\langle S_2 \rangle$  indicates the position of  $p_c(N)$ . We can see that for uncorrelated case ( $\alpha=0$ ), as expected  $p_c(N)$  is close to zero for  $\lambda=2.5$ . In the hub-phobic ( $\alpha>0$ ) regime,  $p_c(N)$  is always finite as expected due to its topological similarity with ER networks. However, in the hub-philic regime ( $\alpha<0$ ), first we observe that the  $p_c(N)$  is always smaller than the uncorrelated case, which means that networks with  $\alpha<0$  are more robust than uncorrelated ones. Second,  $p_c(N=8192)$  in the hub-philic regime ( $\alpha<0$ ) is very close to zero even for  $\lambda=3.5$ , which suggests that  $p_c(\infty)$  might be zero. This is very unusual because it is known that in the uncorrelated case,  $p_c(\infty)$  is finite for  $\lambda > 3$  [6]. A similar behavior  $p_c(N)$  close to zero is obtained for  $\alpha<0$  and  $3 < \lambda < 4$ . Thus, for these values of  $\lambda$ , the case of the hub-philic regime might correspond to a new universality class. This question will be discussed in Sec. II D.

Robustness is directly related to the attacking and immunization strategies [45]. The calculations in Fig. 2 show the robustness of networks against attacks according to the link weights. There are other types of nonrandom attacking strategies such as the following: strategy I, intentionally attacking node according to its degree [42,46], and strategy II, intentionally attacking links according to its weight in the SF network at the hub-phobic regime ( $\alpha>0$ ). It was found that strategy I is an efficient way of attacking a SF network

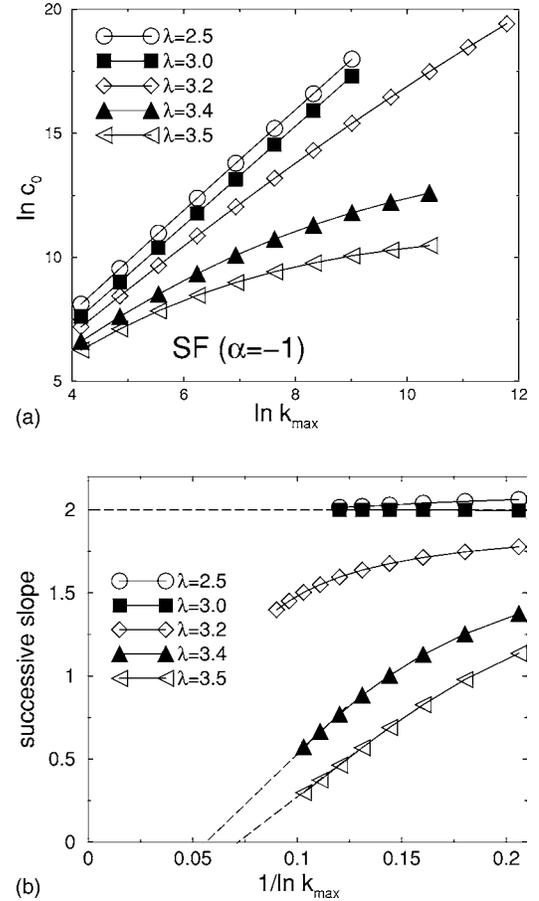


FIG. 5. (a)  $\ln c_0$  as a function of  $\ln k_{\max}$ . (b) Plot of the successive slopes of (a) as a function of  $1/\ln k_{\max}$ .

[42,46]. For strategy II, in the hub-phobic regime, the links connecting to hubs will be attacked first. Next we compare the efficiency of these two types of intentional attack strategies. To fairly compare the two strategies, we calculate  $P_\infty$  of both intentional attack strategies as a function of  $q$ , the fraction of removed links. For strategy I, attacking a node is equivalent to attack all the links of that node. In Fig. 3, we see that these two strategies are similar to each other. Comparing the position of  $q_c$  for these two strategies, we can conclude that strategy II is actually slightly more efficient than strategy I.

From Fig. 2(a) we can see that

$$p_c(\alpha = -1) < p_c(\alpha = 0) < p_c(\alpha = 1), \quad (4)$$

a result which is not surprising because as  $\alpha$  increases more central links connecting high degree nodes are removed. A surprising result is that

$$P_\infty(\alpha = 1) > P_\infty(\alpha = 0) > P_\infty(\alpha = -1) \quad (5)$$

for a wide range of  $q$  values—e.g., for  $\lambda=2.5$  below  $q_c \approx 0.7$ . This is probably due to the fact that when removing central links connecting high degree nodes the remaining network does not become disconnected as shown in the study of the  $k$  core of networks [47]. The effect of removing those central links is mostly to increase the distance.

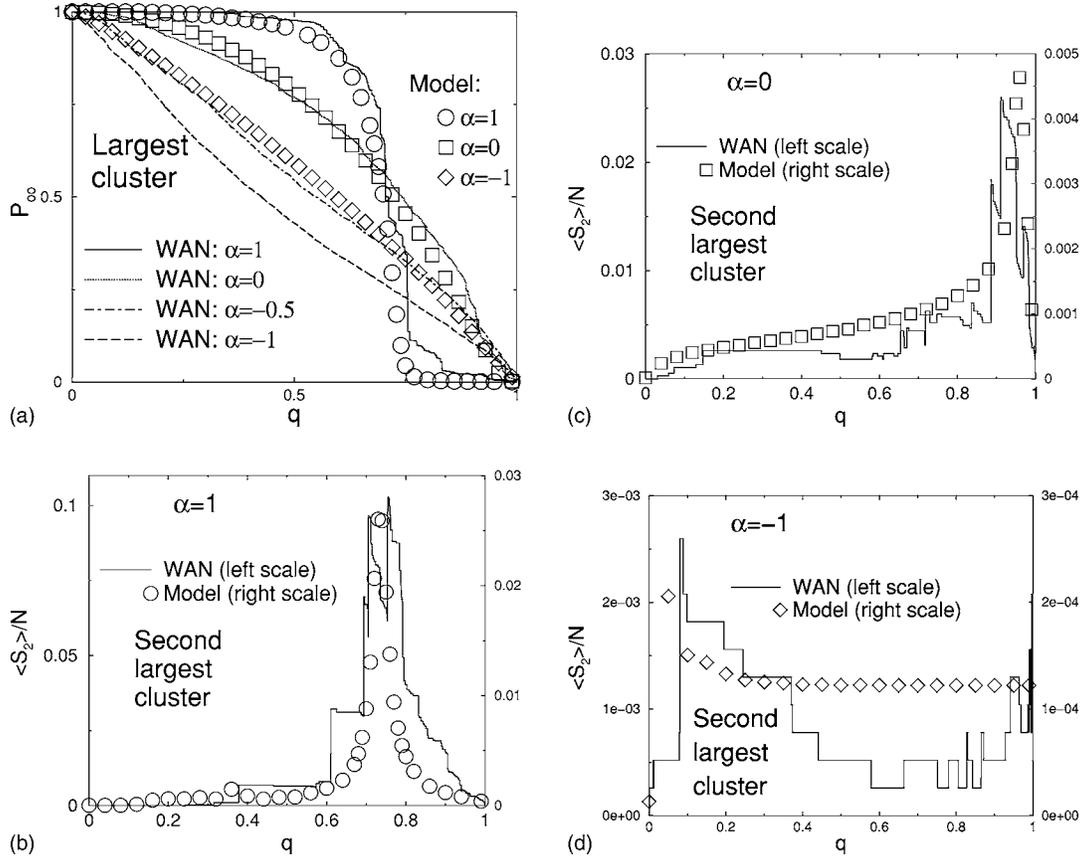


FIG. 6. (a)  $P_{\infty}$  as a function of  $q$ , the concentration of removed links in descending order of the weight for our model with  $\lambda=2.5$  (symbols) and the real WAN (lines). For the model:  $\alpha=1$  ( $\circ$ ),  $\alpha=0$  ( $\square$ ), and  $\alpha=-1$  ( $\diamond$ ). For the WAN:  $\alpha=1$  (solid line),  $\alpha=0$  (dotted line),  $\alpha=-0.5$  (dot-dashed line), and  $\alpha=-1$  (dashed line). (b)–(d)  $\langle S_2 \rangle / N$  as a function of  $q$  for our model with  $\lambda=2.5$  (symbols) and the real WAN (solid lines) with (b)  $\alpha=1$ , (c)  $\alpha=0$ , and (d)  $\alpha=-1$ .

To test our hypothesis we calculate the average length of the shortest path  $\ell_{\min}$  of the largest cluster as a function of  $q$  for the three different correlation cases. In Fig. 4(a) we plot  $\langle \ell_{\min} \rangle$  as function of  $q$ . We can see that

$$\langle \ell_{\min} \rangle_{\alpha=1} > \langle \ell_{\min} \rangle_{\alpha=0} > \langle \ell_{\min} \rangle_{\alpha=-1} \quad (6)$$

for a wide range of  $q$  as predicted. The peaks for  $\alpha=1$  and  $\alpha=0$  are at the same position as the corresponding  $q_c$  of Fig. 2(d) because when  $q > q_c$  the whole network is fragmented. We compute  $\langle \ell_{\min} \rangle$  also for the real WAN and obtain similar results [see Fig. 4(b) and Sec. III].

#### D. Percolation analysis of correlated scale-free networks

An important question is whether, for  $\alpha < 0$  and  $3 < \lambda < 4$ ,  $p_c(\infty)$  is zero. In simulations we observe that  $p_c(\infty)$  is incredibly close to zero for  $3 < \lambda < 4$ . However, it might be a result of finite-size effects and  $p_c$  is indeed very close to zero but not zero. To further test this possibility, we propose the following analytical method.

Near  $p_c$ , the structure of the giant component is like a tree. The giant component at the percolation threshold can be viewed as a growing process of a tree. Assuming that at

some layer the number of nodes with degree  $k$  is  $n(k)$ , then the number of nodes of degree  $k'$  at the next level  $n(k')$  satisfies

$$n(k') = \sum_k n(k)(k-1) \frac{k' \mathcal{P}(k')}{\langle k' \rangle} \phi(k', k), \quad (7)$$

where  $\phi(k', k)$  is the probability that a node with degree  $k$  is connected to a node of degree  $k'$ . Here we are interested in the hub-philic regime ( $\alpha=-1$ ). Since we remove the links in descending order, we can assume that any link with weight above  $1/c$  are cut. The links left are only links with weights below  $1/c$ , which is represented by the condition  $kk' > c$ . Equation (7) can be therefore simplified to

$$n(k') = \sum_{k=c/k'}^{\infty} n(k)(k-1) \frac{k' \mathcal{P}(k')}{\langle k' \rangle}. \quad (8)$$

The dimension of the vector  $n(k)$  is actually the maximum degree  $k_{\max}$ . For a SF network with system size  $N$ ,  $k_{\max} \sim N^{1/(\lambda-1)}$  [6,48]. Thus controlling  $k_{\max}$ , the dimension of the vector  $n(k)$  is actually equivalent to controlling the system size  $N$ . For a fixed  $k_{\max}$ , if Eq. (8) has at least one eigenvalue that is above 1, the giant component will grow to infinity, which means it is above  $p_c$ ; on the other hand, if all of the

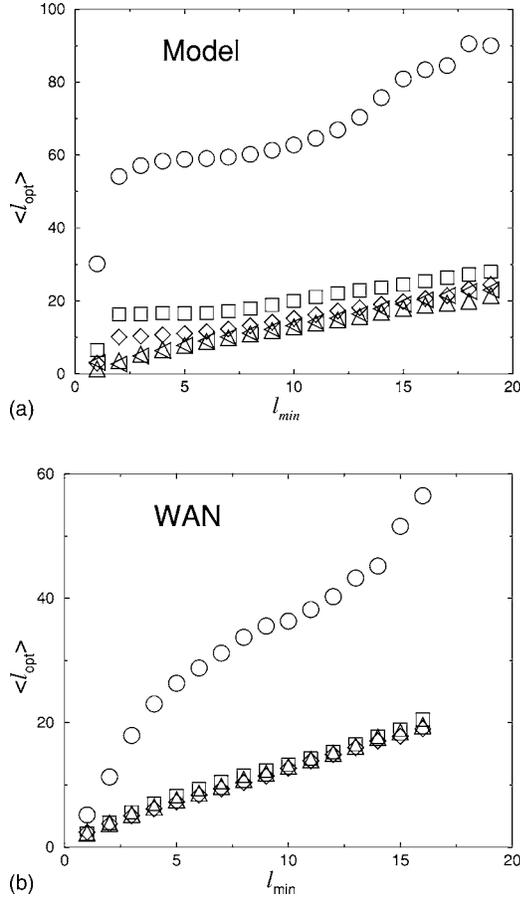


FIG. 7. (a) Plot of  $\langle \ell_{\text{opt}} \rangle$  as a function of  $\ell_{\text{min}}$  for our model of correlated weighted SF network with  $\lambda=2.5$ , in the SD limit ( $\circ$ ),  $\alpha=2$  ( $\square$ ),  $\alpha=1$  ( $\diamond$ ),  $\alpha=0$  ( $\triangle$ ), and  $\alpha=-1$  ( $\triangleleft$ ). (b) Same plot as in (a) for the real WAN with  $\alpha=1$  ( $\circ$ ),  $\alpha=0$  ( $\square$ ),  $\alpha=-1$  ( $\diamond$ ), and  $\alpha=-0.5$  ( $\triangle$ ).

eigenvalues are below 1, the system is below  $p_c$ . We basically change the  $c$  value and compute the normal of vector  $n(k)$  recursively and find the critical  $c$  value  $c^*$  at which the normal of vector  $n(k)$  changes from converging at  $c^*$  to diverging at  $c^*+1$ . Thus, from the definition of percolation,  $p_c(N)$  can be numerically calculated using the relation

$$p_c(N) = \int_{kk' > c^*} \frac{kk' \mathcal{P}(k) \mathcal{P}(k')}{\langle k \rangle^2} dk dk'. \quad (9)$$

If  $c^*(k_{\text{max}})$  diverges as  $k_{\text{max}} \rightarrow \infty$ , from Eq. (9) it follows that  $p_c(\infty)$  is zero. Otherwise,  $p_c(\infty)$  is finite. Thus, we convert the question of whether  $p_c(\infty)$  is zero for  $3 < \lambda < 4$  into the question of whether  $c^*(k_{\text{max}})$  diverges with  $k_{\text{max}}$ .

Figure 5 shows numerical results of  $c^*(k_{\text{max}})$  as  $k_{\text{max}}$  increases. It shows that for  $\lambda < 3$ ,  $c^*$  grows with  $k_{\text{max}}$  as a power law, which confirms that when  $\lambda < 3$ ,  $p_c(\infty)$  is zero [6]. However, for  $3 < \lambda < 4$ , Fig. 5(b) suggests that the successive slopes of  $\ln c^*$  versus  $\ln k_{\text{max}}$  approach zero for  $3 < \lambda < 4$ , which indicates a finite  $p_c(\infty)$  when  $3 < \lambda < 4$ . We can see the strong finite-size effect from Fig. 5, where the convergence of  $c^*$  for  $\lambda=3.5$  happens only for  $k_{\text{max}} > 10^4$

corresponding to  $N \sim k_{\text{max}}^{\lambda-1} > 10^{10}$ , which we are not able to reach in simulations with present computing power.

### III. IMPACT OF THE HUB-PHILIC CORRELATION ON REAL WORLD NETWORKS

Real world networks such as the WAN and the *E. Coli* metabolic network are found to have a hub-philic type correlation ( $\alpha=-0.5$ ) [22–24]. Using the WAN as an example, the passengers tend to go through the large airports, which actually shortens and optimizes the transport performance of the entire WAN network. Our simulations show the advantage of networks in the hub-philic regime over uncorrelated networks from the perspective of the optimal paths in the SD regime [see Figs. 1(b)–1(d)]. Also they show that networks in the hub-philic regime are also more robust than uncorrelated networks (see Fig. 2). Associated to the WAN there is only one value of  $\alpha=-0.5$ , so to see the effect of different  $\alpha$  on the real WAN, we use the following method. Assigning the weights according to Eq. (1), we first compute  $x_{ij}^{\text{real}}$ , the value of  $x_{ij}$  from the real weights of the WAN. Thus we leave  $\alpha$  as a parameter and we can change its value to get either the hub-philic or hub-phobic regime. In this case, when  $\alpha=-0.5$ , the weight is the same as in the original WAN.

Figure 6 compares the robustness calculations [see also Figs. 2(a) and 2(d)] of our numerical results with the real WAN. From the calculation of both the largest and second largest clusters, we can see that real WAN has behavior similar to the network model with generalized correlated weights. Notice that Fig. 6(a) shows that  $\alpha=-0.5$  is more robust than  $\alpha=-1$ , which shows the advantage of  $\alpha=-0.5$  in the real WAN [30]. The second largest cluster calculation [see also Figs. 6(b)–6(d) and Fig. 2(d)] indicates where the  $p_c$  is; we see that the WAN and our model are surprisingly similar.

To compare the optimal paths between the real WAN and the SF network with generalized correlated weights, we calculate  $\langle \ell_{\text{opt}} \rangle$  as a function of  $\ell_{\text{min}}$ . Figure 7 shows the simulation results for the SF model and for the real WAN [Fig. 7(b)]. We observe that for  $\alpha \leq 0$ ,  $\langle \ell_{\text{opt}} \rangle$  is almost the same as  $\ell_{\text{min}}$  because in the hub-philic regime the optimal paths tend to go through the hubs, which shortens the length of the optimal paths. In the hub-phobic regime ( $\alpha > 0$ ), the optimal paths tends to avoid the hubs, which naturally generate large length of the optimal paths. We also see that in both the real WAN and the model for  $\alpha > 0$ , and for short  $\ell_{\text{min}}$ ,  $\langle \ell_{\text{opt}} \rangle$  increases sharply, while for large  $\ell_{\text{min}}$ ,  $\langle \ell_{\text{opt}} \rangle$  increases weakly with  $\ell_{\text{min}}$ . This is probably related to the crossover from strong disorder in short length scales to weak disorder in large length scales that was observed in several earlier studies [16,19,20].

### IV. CONCLUSIONS

Motivated by studies of the WAN and *E. Coli* metabolic network [22,24], we analyze the three different regimes of a network model with generalized correlated weights [22,28]—Eq. (1)—as defined by the value of the parameter  $\alpha$ : the hub-philic regime ( $\alpha < 0$ ), the hub-phobic regime

( $\alpha > 0$ ), and the uncorrelated regime ( $\alpha = 0$ ). We study the properties of the optimal path and find two universality classes for the length of the optimal path  $\ell_{\text{opt}}$ : (i) For networks in the hub-phobic regime ( $\alpha > 0$ ), the optimal path on SF networks with any  $\lambda > 2$  behaves the same as the optimal path in ER networks. (ii) For  $\alpha \leq 0$ , the optimal paths in SF networks belongs to the same universality class as the optimal path in SF networks with uncorrelated weights.

We also calculate the robustness of the SF networks with correlated weights and find that SF networks in the hub-philic regime are more robust compared to SF networks in the other two regimes. We propose an analytical method to study the percolation transition of networks with weights of Eq. (1) and the numerical results suggest that the  $p_c(\infty)$  of networks in the hub-philic regime, although close to zero, is not zero for  $3 < \lambda < 4$ . We also observe strong finite-size effects for SF networks in the hub-philic regime.

In the last section, we compare the simulation results of the weighted network model defined by Eq. (1) with the

actual WAN. We calculate the  $\langle \ell_{\text{opt}} \rangle$  for two nodes separated with fixed distance  $\ell_{\text{min}}$  and observe similar behaviors and the crossover from strong disorder to weak disorder for both the model and the real WAN. We also compare the numerical results of the robustness calculation with those obtained in the real WAN and find good agreement.

#### ACKNOWLEDGMENTS

The authors would like to thank Lazaros K. Gallos, Eduardo López, Gerald Paul, and Sameet Sreenivasan for useful discussions and suggestions. We thank Alessandro Vespignani for useful discussions and the analysis of the anonymized WAN data that has been carried out in his laboratory at the School of Informatics. We thank ONR, ONR-Global, UNMdp, NEST Project No. DYSONET012911, and the Israel Science Foundation for support. L.A.B. thanks FONCYT (PICT-O 2004/370) 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, England, 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] D. J. Watts and S. H. Strogatz, *Nature (London)* **393**, 440 (1998).
- [5] A.-L. Barabási and R. Albert, *Science* **286**, 509 (1999).
- [6] R. Cohen, K. Erez, Daniel ben-Avraham, and S. Havlin, *Phys. Rev. Lett.* **85**, 4626 (2000).
- [7] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts, *Phys. Rev. Lett.* **85**, 5468 (2000).
- [8] R. Cohen and S. Havlin, *Phys. Rev. Lett.* **90**, 058701 (2003).
- [9] E. López, S. V. Buldyrev, S. Havlin, and H. E. Stanley, *Phys. Rev. Lett.* **94**, 248701 (2005).
- [10] K. Park, Y.-C. Lai, and N. Ye, *Phys. Rev. E* **70**, 026109 (2004).
- [11] A. Barrat, M. Barthélemy, and A. Vespignani, *Phys. Rev. Lett.* **92**, 228701 (2004).
- [12] S. H. Yook, H. Jeong, A.-L. Barabási, and Y. Tu, *Phys. Rev. Lett.* **86**, 5835 (2001).
- [13] D. Zheng, S. Trimper, B. Zheng, and P. M. Hui, *Phys. Rev. E* **67**, 040102(R) (2003).
- [14] W.-X. Wang, B.-H. Wang, B. Hu, G. Yan, and Q. Ou, *Phys. Rev. Lett.* **94**, 188702 (2005).
- [15] M. Cieplak *et al.*, *Phys. Rev. Lett.* **72**, 2320 (1994).
- [16] M. Porto *et al.*, *Phys. Rev. E* **60**, R2448 (1999).
- [17] L. A. Braunstein *et al.*, *Phys. Rev. Lett.* **91**, 168701 (2003); S. V. Buldyrev, L. A. Braunstein, R. Cohen, S. Havlin, and H. E. Stanley, *Physica A* **330**, 246 (2003).
- [18] Y. M. Strelniker, R. Berkovits, A. Frydman, and S. Havlin, *Phys. Rev. E* **69**, 065105(R) (2004).
- [19] Z. Wu *et al.*, *Phys. Rev. E* **71**, 045101(R) (2005).
- [20] S. Sreenivasan *et al.*, *Phys. Rev. E* **70**, 046133 (2004).
- [21] T. Kalisky, S. Sreenivasan, L. A. Braunstein, S. V. Buldyrev, S. Havlin, and H. E. Stanley, *Phys. Rev. E* **73**, 025103(R) (2006).
- [22] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 3747 (2004).
- [23] E. Almaas, B. Kovács, T. Vicsek, Z. N. Oltvai, and A.-L. Barabási, *Nature (London)* **427**, 839 (2004).
- [24] P. J. Macdonald, E. Almaas, and A.-L. Barabási, *Europhys. Lett.* **72**, 308 (2005).
- [25] R. Guimerá and L. A. N. Amaral, *Eur. Phys. J. B* **38**, 381 (2004).
- [26] R. Guimerá, S. Mossa, A. Turttschi, and L. A. N. Amaral, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 7794 (2005).
- [27] W. Li and X. Cai, *Phys. Rev. E* **69**, 046106 (2004).
- [28] E. Almaas, P. L. Krapivsky, and S. Redner, *Phys. Rev. E* **71**, 036124 (2005).
- [29] Correlated weight with Eq. (1) could be assigned to any network other than the SF networks, but here we focus on SF networks. References [22,24] use  $\theta$  to represent the correlation parameter.  $\alpha$  in Eq. (1) is analogous to  $-\theta$ . So the corresponding  $\alpha$  values for real networks such as the WAN and the *E. coli* are negative, with  $\alpha = -0.5$  for both.
- [30] From the definition and the algorithm for the MST, only the relative order of the links according to the weight determines the structure of the MST. Thus, from Eq. (1), for all values of  $\alpha > 0$ , if  $x_{ij}$  is fixed to be 1, the same MST will be found for a specific structure of network (see also [24]). In our simulation as we do many realizations with different series of random number  $x_{ij}$ , statistically results related the MST will be the same for all value of  $\alpha$  with the same sign. For this reason, we only show the simulation results for  $\alpha = 1, 0$  and  $-1$ . However, for a specific series of prefactor  $x_{ij}$ , different value of  $\alpha$  with the same sign will give different MST. Thus while identifying the MST in the real WAN data, the value of  $\alpha$  does matter.
- [31] M. Khan, G. Pandurangan, and B. Bhargava (unpublished).
- [32] S. Skiena, *Implementing Discrete Mathematics: Combinatorics and Graph Theory With Mathematica* (Addison-Wesley, New

- York, 1990).
- [33] M. L. Fredman and R. E. Tarjan, *J. ACM* **34**, 596 (1987).
- [34] J. B. Kruskal, *Proc. Am. Math. Soc.* **7**, 48 (1956).
- [35] G. Bonanno, G. Caldarelli, F. Lillo, and R. N. Mantegna, *Phys. Rev. E* **68**, 046130 (2003).
- [36] J.-P. Onnela *et al.*, *Phys. Rev. E* **68**, 056110 (2003).
- [37] M. Molloy and B. A. Reed, *Combinatorics, Probab. Comput.* **7**, 295 (1998).
- [38] M. Molloy and B. A. Reed, *Random Struct. Algorithms* **6**, 161 (1995).
- [39] R. K. Ahuja, T. L. Magnanti, J. B. Orlin, *Network Flows: Theory, Algorithms and Applications* (Prentice-Hall, Englewood Cliffs, NJ, 1993).
- [40] *Fractals and Disordered Systems*, edited by A. Bunde and S. Havlin (Springer, New York, 1996).
- [41] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, *Phys. Rev. E* **64**, 026118 (2001).
- [42] R. Cohen, K. Erez, Daniel ben-Avraham, and S. Havlin, *Phys. Rev. Lett.* **86**, 3682 (2001).
- [43] L. Dall'Asta, *J. Stat. Mech.: Theory Exp.* P08011 (2005).
- [44] R. Pastor-Satorras and A. Vespignani, *Phys. Rev. Lett.* **86**, 3200 (2001).
- [45] L. Dall'Asta, A. Barrat, M. Barthelemy, and A. Vespignani, *J. Stat. Mech.: Theory Exp.* P04006 (2006).
- [46] L. K. Gallos, R. Cohen, P. Argyrakis, A. Bunde, and S. Havlin, *Phys. Rev. Lett.* **94**, 188701 (2005).
- [47] S. Carmi, S. Havlin, S. Kirkpatrick, Y. Shavitt, and E. Shir, e-print cond-mat/0601240.
- [48] M. E. J. Newman, *SIAM Rev.* **45**, 167 (2003).