Hierarchical scale-free network is fragile against random failure

Takehisa Hasegawa[*]

Graduate School of Information Science, Tohoku University, 6-3-09, Aramaki-Aza-Aoba, Sendai, Miyagi 980-8579, Japan

Koji Nemoto[†]

Department of Physics, Hokkaido University, Kita 10 Nisi 8, Kita-ku, Sapporo, Hokkaido 060-0810, Japan

We investigate site percolation in a hierarchical scale-free network known as the Dorogovtsev-Goltsev-Mendes network. We use the generating function method to show that the percolation threshold is 1, i.e., the system is not in the percolating phase when the occupation probability is less than 1. The present result is contrasted to bond percolation in the same network of which the percolation threshold is zero. We also show that the percolation threshold of intentional attacks is 1. Our results suggest that this hierarchical scale-free network is very fragile against both random failure and intentional attacks. Such a structural defect is common in many hierarchical network models.

I. INTRODUCTION

The prominent resilience of real networks to random failure and intentional attacks is one of the important issues in network science [1–4]. Many real networks are scale free, i.e., the degree distribution $p(k)$ is a power law denoted by $p(k) \propto k^{-\gamma}$ with $2 \leqslant \gamma \lesssim 3$. Albert *et al.* examined the robustness of networks against two types of attacks: random failure in which nodes are sequentially removed with equal probability and intentional attack, which preferentially removes nodes of large degrees [5]. They showed that scale-free networks with small γ are highly robust against random failure, i.e., the network remains intact until almost all nodes have been removed. On the other hand, such networks are very fragile to intentional attacks because removal of a small fraction of hubs destroys the network.

Random failures and intentional attacks in networks can be interpreted as percolation problems and have been well studied (see Refs. [2,3] and references therein). It is well known that the site (bond) percolation model with the probability p that each site (bond) is occupied (open) has a percolation threshold p_c above which the largest connected component is $O(N)$, N being the number of nodes. Such a network is said to be *robust* (*fragile*) against failure if p_c $(1 − p_c)$ is very small when a random failure is regarded as a node vacancy in site percolation and as a closed link in bond percolation. The local tree approximation for uncorrelated networks [6], which is a standard theory in network science [3], confirms that scale-free networks with heavy-tailed degree distributions are robust against random failures [7,8], i.e., p_c is zero for both bond and site percolations when $\gamma \leqslant 3$. This approximation can be applied to the case of intentional attacks to show that uncorrelated scale-free networks with small γ are fragile against such attacks [7,9]. This theory can be extended so as to treat clustered networks and correlated networks [10–15].

Apart from the local tree approximation, bond percolation in growing and hierarchical networks has been studied extensively (see below), whereas, almost no analytical studies have focused on site percolation. To date, little attention has been paid to the difference between bond percolation and site percolation in complex networks, possibly because the critical properties of the two percolation models are not qualitatively different within the local tree approximation [7,8]. However, what we will demonstrate here is that the opposite can occur. We examine site percolation in a hierarchical scale-free network known as the Dorogovtsev-Goltsev-Mendes (DGM) network [16] or the (1,2) flower [17,18]. Dorogovtsev [19] calculated bond percolation in this network by renormalization group and showed that the percolation threshold p_c is zero. On the other hand, we use generating functions to show that the percolation threshold p_c of site percolation is not zero but 1. That is, this network is fragile even against random failure. Our analytical result is supported perfectly by Monte Carlo simulation.

Before discussing the main topic, we emphasize that percolation characteristics on some graphs are delineated by two nontrivial transition points, namely, p_{c1} and p_{c2} [20–22]. According to the value of p, the system shows one of the three phases: (i) nonpercolating phase $(0 \leqslant p \leqslant p_{c1})$ in which all clusters are of finite size, (ii) critical phase $(p_{c1} < p < p_{c2})$ (called the patchy phase in Ref. [23]) in which infinitely many infinite clusters exist, and (iii) percolating phase $(p_{c2} \leqslant p \leqslant 1)$ in which the system has a unique infinite cluster. Here an infinite cluster is a cluster whose size is on the order $O(N^\alpha)(0 < \alpha \leqslant 1)$. Note that p_{c2} is equal to the percolation threshold p_c, $p_c = p_{c2}$. By using the order parameter $m \equiv \lim_{N \to \infty} s_{\max}(N; p)/N$ and the fractal exponent $\psi \equiv \lim_{N \to \infty} \log_N s_{\max}(N)$ [24,25], we represent these phases as: (i) $m = 0$ and $\psi = 0$, (ii) $m = 0$ and $0 < \psi < 1$, and (iii) $m > 0$ and $\psi = 1$, respectively. Here $s_{\max}(N; p)$ is the mean size of the largest cluster in a graph of size N at a given value of p. As known, $p_{c1} = p_{c2}$ on Euclidean lattices, whereas, $p_{c1} < p_{c2}$ on transitive nonamenable graphs [21,22]. Also, in complex networks, some growing network models [25–29] and hierarchical network models [23,30–33] yield $0 = p_{c1} < p_{c2}$ for bond percolation, whereas, $p_{c1} = p_{c2}$ for some static network models, such as uncorrelated networks

---
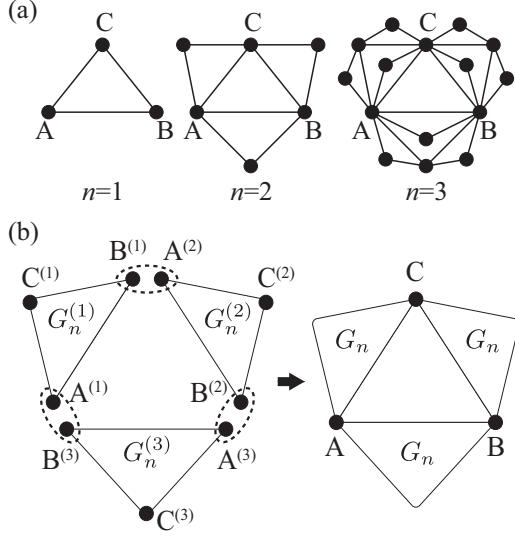
[*]hasegawa@m.tohoku.ac.jp

[†]nemoto@statphys.sci.hokudai.ac.jp

FIG. 1. (a) Realization of the DGM network with generation $n = 1$–3. (b) Construction of G_{n+1} from G_n. Three copies of G_n ($G_n^{(k)}$, $k = 1$–3) are connected by identifying $A^{(1)}$ and $B^{(3)}$ to be the new A, $A^{(3)}$ and $B^{(2)}$ to be the new B, and $A^{(2)}$ and $B^{(1)}$ to be the new C, respectively.

[2,3]. In this paper, we analytically show that $p_{c1} = 0$ and $p_{c2} = 1$ for site percolation, whereas, $p_{c1} = p_{c2} = 0$ for bond percolation [19]. We also demonstrate that $p_{c1} = p_{c2} = 1$ for intentional attacks on the DGM network.

II. MODEL

The DGM network was proposed as a deterministic growing network [16]. Let us denote the DGM network with generation ($=$ time) n by G_n. The model starts from a triangle at $n = 1$. At each time step n, every link in G_n adds a new node, which links to both end nodes of the link, to create G_{n+1}. The realizations of the first three generations G_1, G_2, and G_3 are shown in Fig. 1(a). This model also is regarded as a recursively constructed hierarchical network: G_1 consists of a triangle of nodes A, B, and C. We refer to these nodes as *roots*. Then, G_{n+1} is constructed from three copies of G_n that are joined at the roots as shown in Fig. 1(b).

The structural properties of this network have been described in Refs. [16,17]. The number of links in G_n is 3^n, and the number of nodes is $(3^n + 3)/2$. The number of nodes $N_n(\ell)$ of degree $k = 2^\ell$ ($\ell = 1, 2, \ldots, n$) is $N_n(\ell) = 3^{n-\ell}$ for $\ell < n$ and $N_n(n) = 3$. Thus, the degree distribution $p(k)$ is a power law, $p(k) \propto k^{-\gamma}$ with $\gamma = 1 + \ln 3/ \ln 2$. Furthermore, the DGM network is a small-world network because the diameter of G_n is n and the clustering coefficient is $4/5$ in the limit $n \to \infty$.

III. GENERATING FUNCTION

Let us consider site percolation in G_n with an occupation probability p. We calculate the mean size $s_{\text{root}}(N_n; p)$ of the root cluster, which includes at least one of the roots A, B, and C, rather than the mean largest cluster size $s_{\text{max}}(N_n; p)$. Here, A, B, and C have the largest degree $k = 2^n$ in G_n and can, therefore, be regarded as *hubs*. Also, the root cluster is always unique because these roots are connected directly in the
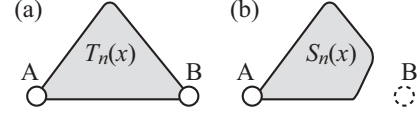

FIG. 2. Schematic of (a) $T_n(x)$ and (b) $S_n(x)$.

cluster. Since this root cluster is expected to become the largest cluster, we assume that $s_{\text{max}}(N_n; p)$ is well approximated by $s_{\text{root}}(N_n; p)$. This assumption is verified numerically below.

To evaluate $s_{\text{root}}(N_n; p)$, we consider the following two quantities in G_n: The probability that the size of the root cluster is k provided that both A and B are occupied [denoted as $t_k^{(n)}(p)$, we call such clusters doubly occupied] and the probability that the size of the root cluster is k provided that A is occupied and B is unoccupied [denoted as $s_k^{(n)}(p)$, we call such clusters singly occupied]. For convenience, A and B are not included in counting the cluster size k for $t_k^{(n)}(p)$ and $s_k^{(n)}(p)$, but C is included. We now introduce the generating functions $T_n(x)$ and $S_n(x)$ for $t_k^{(n)}(p)$ and $s_k^{(n)}(p)$ (Fig. 2), which are defined as

$$T_n(x) \equiv \sum_{k=0}^{\infty} t_k^{(n)}(p) x^k, \tag{1a}$$

$$S_n(x) \equiv \sum_{k=0}^{\infty} s_k^{(n)}(p) x^k. \tag{1b}$$

Here, $T_n(1) = S_n(1) = 1$ for all n. Given the self-similar structure, the recursion relations for these generating functions are readily obtained as

$$T_{n+1}(x) = px T_n^3(x) + q T_n(x) S_n^2(x), \tag{2a}$$

$$S_{n+1}(x) = px T_n(x) S_n^2(x) + q S_n^2(x), \tag{2b}$$

where $q \equiv 1 - p$ [Figs. 3(a) and 3(b)]. The initial condition is $T_1(x) = S_1(x) = q + px$. For example, the first term of the right-hand side in Eq. (2a) represents the contribution of the first graph of Fig. 3(a) in which root C is occupied. It consists of the factor px for the occupied root C and the factor $T_n^3(x)$ for three doubly occupied root clusters. Here, x accounts for the occupied root C, which is not counted in $T_n(x)$ or $S_n(x)$, but is counted in $T_{n+1}(x)$ or $S_{n+1}(x)$. The second term of the right-hand side in Eq. (2a) represents the contribution of the second graph of Fig. 3(a). The factor q means the probability of root C being unoccupied, and $T_n(x)S_n^2(x)$ is given from one doubly occupied root cluster and two singly occupied root clusters.

Now, we consider the mean size of the root cluster. By $f_k^{(n)}(p)$, we denote the probability that the size of the root cluster in G_n is k. For evaluating the generating function $F_n(x) \equiv \sum_{k=0}^{\infty} f_k^{(n)}(p) x^k$, we only need to count possible contributing diagrams as shown in Fig. 3(c). Noting that all roots A, B, and C in G_{n+1} are not counted as occupied in $T_n(x)$ and $S_n(x)$, we easily find that $F_{n+1}(x)$ is evaluated as

$$F_{n+1}(x) = p^3 x^3 T_n^3(x) + 3p^2 q x^2 T_n(x) S_n^2(x) + 3pq^2 x S_n^2(x) + q^3. \tag{3}$$
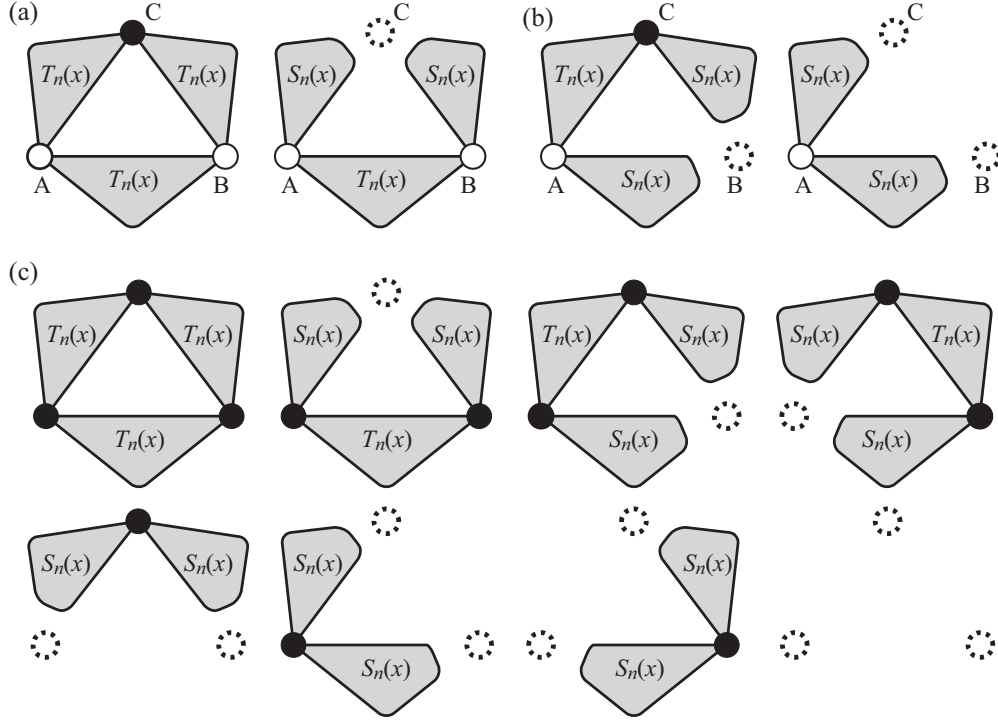
FIG. 3. Possible contributions to (a) $T_{n+1}(x)$, (b) $S_{n+1}(x)$, and (c) $F_{n+1}(x)$. Each solid (dashed) circle represents the root node being occupied (unoccupied). The black circles represent occupied root nodes, which should be taken into account by multiplying x. For example, the first diagram of (a) represents $px T_n^3(x)$, the second diagram of (b) represents $q S_n^2(x)$, and the second one in the second line of (c) represents $pq^2 x S_n^2(x)$.

For $n = 1$, we have $F_1(x) = p^3 x^3 + 3p^2 q x^2 + 3pq^2 x + q^3$. Then, $s_{\text{root}}(N_{n+1}; p) = F_{n+1}'(1)$ is given by

$$s_{\text{root}}(N_{n+1}; p) = 3p^2 T_n'(1) + 6pq S_n'(1) + 3p, \qquad (4)$$

and $s_{\text{root}}(N_1; p) = 3p$. Here, the prime denotes the first derivative with respect to x. By Eq. (2), we have $T_n'(1)$ and $S_n'(1)$ recursively as

$$T_{n+1}'(1) = (2p + 1)T_n'(1) + 2q S_n'(1) + p, \qquad (5a)$$

$$S_{n+1}'(1) = p T_n'(1) + 2 S_n'(1) + p, \qquad (5b)$$

with $T_1'(1) = S_1'(1) = p$.

Now, we consider the fractal exponent of the root cluster ψ_{root}, where $s_{\text{root}}(N_n; p) \propto N_n^{\psi_{\text{root}}}$. For $n \gg 1$, the recursion relations (5) are approximated as $\mathbf{x_{n+1}} = \mathbf{A} \cdot \mathbf{x_n}$, where

$$\mathbf{x_n} = \begin{pmatrix} T_n'(1) \\ S_n'(1) \end{pmatrix}, \qquad \mathbf{A} = \begin{pmatrix} 2p+1 & 2q \\ p & 2 \end{pmatrix}. \qquad (6)$$

Solving the characteristic equation of $\mathbf{A}$ yields the largest eigenvalue $\lambda(\mathbf{A})$,

$$\lambda(\mathbf{A}) = \tfrac{1}{2}(2p + 3 + \sqrt{1 + 4p - 4p^2}). \qquad (7)$$

By noting $s_{\text{root}}(N_n; p) \propto \lambda(\mathbf{A})^n$, we have the fractal exponent of the root cluster as

$$\psi_{\text{root}}(p) = \frac{\ln\left[\tfrac{1}{2}(2p + 3 + \sqrt{1 + 4p - 4p^2})\right]}{\ln 3} \qquad (8)$$

for $0 < p \leqslant 1$. At $p = 0$, $\psi_{\text{root}}(p) = 0$.

IV. RESULT

From Eq. (8), it is apparent that $\ln 2 / \ln 3 < \psi_{\text{root}} < 1$ for $0 < p < 1$. This indicates that $p_{c1} = 0$ and $p_{c2} = 1$ and the order parameter is zero except at $p = 1$. In Fig. 4(a), $m_{\text{root}}(N_n; p) = s_{\text{root}}(N_n; p)/N_n$ is plotted over several generations using Eqs. (4) and (5). We also performed the Monte Carlo simulation of site percolation in the DGM network for different generations (n ranging from 8 to 13). The number of percolation trials at each p is 200 000. The order parameters $m(N_n; p) = s_{\text{max}}(N_n; p)/N_n$ obtained by Monte Carlo simulation for $n = 9$ and 12 are shown in Fig. 4(a). The numerical results of $m(N_n; p)$ lie precisely on the analytical curves of $m_{\text{root}}(N_n; p)$. From Fig. 4(a), we find that, across the range of p, the order parameter decays to zero as n increases, although the convergence is very slow. The giant component of $O(N)$ disappears at $p < 1$ in the thermodynamic limit, implying that this network is essentially fragile against random failures.

In Fig. 4(b), we plot the fractal exponent $\psi_{\text{root}}(p)$ given by Eq. (8) (solid line) and $\psi(N_n; p)$ obtained from the Monte Carlo simulations (symbols). Here, the fractal exponent $\psi(N; p) = d \ln s_{\text{max}}(N; p)/d \ln N$ of a finite graph is evaluated as the difference $\psi(N_n; p) \approx [\ln s_{\text{max}}(N_{n+1}; p) - \ln s_{\text{max}}(N_{n-1}; p)]/(\ln N_{n+1} - \ln N_{n-1})$. When $p \gtrsim 0.2$, the Monte Carlo results lie in the theoretical curve. When $p \lesssim 0.2$, the data points deviate from this curve, but this deviation can be diminished by increasing n.

The DGM network is also fragile against intentional attacks. Note that, in G_n, the degree of the node added at generation
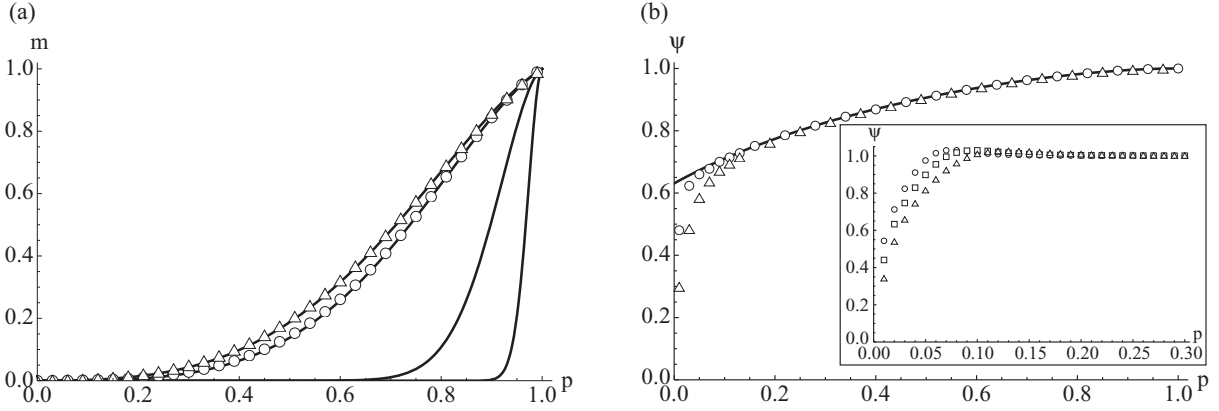
FIG. 4. (a) Order parameter and (b) fractal exponent profiles for different generations. The solid lines in (a) are the numerical evaluations of $m_{root}(N_n; p)$ given by Eqs. (4) and (5) for generations $n = 9$, 12, 100, and 1000 (from left to right). The solid line in (b) represents $\psi_{root}(p)$ given by Eq. (8). The open circles and triangles denote the results obtained from the Monte Carlo simulation for (a) $m(N_n; p)$ and (b) $\psi(N_n; p)$ with $n = 12$ and 9, respectively. The inset of (b) shows the plot of the fractal exponent $\psi(N_n; p)$ of the configuration model having the same degree distribution as that of the DGM network. The number of nodes is $N_n = 265\,722$ (circles), 88 575 (squares), and 29 526 (triangles). The number of trials over each realization is 1000, and the number of graph realizations is 100.

ℓ is $2^{n-\ell}$, i.e., the older the node is, the larger the degree is. Let us consider removing the nodes added at generations less than ℓ from G_n. Then, the three clusters, including the nodes added at generation ℓ, can be considered as the largest clusters. Simple reasoning gives the size of these clusters as $3^{n-\ell}$ in G_n, i.e., $\psi(N_n; p) = 1 - \ell/n$. To prevent the fraction of removed nodes $\tilde{p} = 1 - p = 3^{\ell-1}/3^n$ from disappearing in the limit $n \to \infty$, ℓ needs to increase as $\ell = n - c$, where c is some constant. Then, we conclude that $\psi = c/n \to 0$ for $p < 1$, implying that $p_{c1} = 1$. Thus, we have $p_{c1} = p_{c2} = 1$ for intentional attacks.

V. SUMMARY

To summarize, we have examined site percolation in the DGM network. We have shown that $p_{c1} = 0$ and $p_{c2} = 1$ for site percolation (random failure), whereas, $p_{c1} = p_{c2} = 0$ for bond percolation [19]. We also have demonstrated that $p_{c1} = p_{c2} = 1$ for intentional attacks. We conclude that this hierarchical network is fragile against both random failure and intentional attacks.

How universal are the behaviors observed here among complex networks? The origin of the observed fragility should be assigned to the hierarchical structure of the DGM network. We have numerically observed site percolation in the configuration model having the same degree distribution as that of the DGM network [see the inset of Fig. 4(b)]. The numerically obtained $\psi(N; p)$ approaches 1 even for small $p(>0)$ when N increases. This means $p_{c2} = 0$, which is consistent with the result by the local tree approximation [7,8] (note that degree exponent γ of this network is less than 3). Furthermore, for site-bond percolation in another hierarchical scale-free network, called the decorated (2,2) flower, the critical open bond probability

is 1 except for the case of no dilution [34]. An important point here is that the difference in fragility between bond and site percolations in the DGM network may be related to a geometrical property, i.e., the number of *ends*. Consider an infinite graph G. The number of ends of G, $e(G)$, is given as the supremum of the number of infinitely connected components in $G \backslash S$, where $G \backslash S$ is the graph obtained from G by removing an arbitrary finite subset S of nodes or edges. If G is locally finite and transitive, $p_c = 1$ when $e(G) = \infty$ [35] and $p_c < 1$ when $e(G) = 1$ [37]. In the infinite DGM network G_∞, which is neither locally finite nor transitive, the number of ends for the deletion of *nodes* and that for the deletion of *edges* can be considerably different: In the case of node and edge deletions, $e(G_\infty)$ is infinite and unity, respectively. In other words, G_∞ can disintegrate if a finite number of nodes is removed, which occurs with a nonzero probability when $p < 1$, but is robust against edge removals, thus, reflecting the qualitative difference in percolation threshold between site and bond percolations. Such a structural defect appears to be common among previously identified hierarchical scale-free networks, such as the (decorated) (u, v) flower, the Ravasz-Barabási hierarchical network [38], and a hierarchical network proposed by Barabási *et al.* [39]. We expect that fragility against random failures is also common in hierarchical network models. The validity of this conjecture will be investigated in future studies.

[1] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).

[2] M. E. J. Newman, SIAM Rev. **45**, 167 (2003).

[3] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Rev. Mod. Phys. **80**, 1275 (2008).

[4] A. Barrat, M. Barthélemy, and A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge University Press, Cambridge, U.K., 2008).

[5] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **406**, 378 (2000).

[6] M. Molloy and B. Reed, Random Struct. Algorithms **6**, 161 (1995).

[7] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. Lett. **85**, 5468 (2000).

[8] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, Phys. Rev. Lett. **85**, 4626 (2000).

[9] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, Phys. Rev. Lett. **86**, 3682 (2001).

[10] A. V. Goltsev, S. N. Dorogovtsev, and J. F. F. Mendes, Phys. Rev. E **78**, 051105 (2008).

[11] T. Tanizawa, S. Havlin, and H. E. Stanley, Phys. Rev. E **85**, 046109 (2012).

[12] M. E. J. Newman, Phys. Rev. E **68**, 026121 (2003).

[13] J. P. Gleeson, Phys. Rev. E **80**, 036107 (2009).

[14] J. P. Gleeson and S. Melnik, Phys. Rev. E **80**, 046121 (2009).

[15] J. P. Gleeson, S. Melnik, and A. Hackett, Phys. Rev. E **81**, 066114 (2010).

[16] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Phys. Rev. E **65**, 066122 (2002).

[17] H. D. Rozenfeld, S. Havlin, and D. ben-Avraham, New J. Phys. **9**, 175 (2007).

[18] H. D. Rozenfeld and D. ben-Avraham, Phys. Rev. E **75**, 061102 (2007).

[19] S. N. Dorogovtsev, Phys. Rev. E **67**, 045102 (2003).

[20] I. Benjamini and O. Schramm, Electron. Commun. Probab. **1**, 71 (1996).

[21] R. Lyons, J. Math. Phys. **41**, 1099 (2000).

[22] R. Schonmann, Commun. Math. Phys. **219**, 271 (2001).

[23] S. Boettcher, J. L. Cook, and R. M. Ziff, Phys. Rev. E **80**, 041115 (2009).

[24] T. Nogawa and T. Hasegawa, J. Phys. A: Math. Theor. **42**, 145001 (2009).

[25] T. Hasegawa, T. Nogawa, and K. Nemoto, Europhys. Lett. **104**, 16006 (2013).

[26] D. S. Callaway, J. E. Hopcroft, J. M. Kleinberg, M. E. J. Newman, and S. H. Strogatz, Phys. Rev. E **64**, 041902 (2001).

[27] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, Phys. Rev. E **64**, 066110 (2001).

[28] L. Zalányi, G. Csárdi, T. Kiss, M. Lengyel, R. Warner, J. Tobochnik, and P. Érdi, Phys. Rev. E **68**, 066104 (2003).

[29] T. Hasegawa and K. Nemoto, Phys. Rev. E **81**, 051105 (2010).

[30] A. N. Berker, M. Hinczewski, and R. R. Netz, Phys. Rev. E **80**, 041118 (2009).

[31] S. Boettcher, V. Singh, and R. M. Ziff, Nat. Commun. **3**, 787 (2012).

[32] T. Hasegawa, M. Sato, and K. Nemoto, Phys. Rev. E **82**, 046101 (2010).

[33] T. Hasegawa and T. Nogawa, Phys. Rev. E **87**, 032810 (2013).

[34] T. Hasegawa, M. Sato, and K. Nemoto, Phys. Rev. E **85**, 017101 (2012).

[35] The number of ends of an infinite locally finite transitive graph is either 1, 2, or ∞ [36].

[36] B. Mohar, Discrete Math. **95**, 193 (1991).

[37] There is a widely believed conjecture that $p_c < 1$ if $e(G)$ of an infinite locally finite transitive graph G is 1 [21].

[38] E. Ravasz and A.-L. Barabási, Phys. Rev. E **67**, 026112 (2003).

[39] A.-L. Barabási, E. Ravasz, and T. Vicsek, Physica A **299**, 559 (2001).