Mechanisms of complex network growth: Synthesis of the preferential attachment and fitness models

Michael Golosovsky*
Racah Institute of Physics, Hebrew University of Jerusalem, 91904 Jerusalem, Israel

(Received 11 December 2017; revised manuscript received 7 May 2018; published 12 June 2018)

We analyze growth mechanisms of complex networks and focus on their validation by measurements. To this end we consider the equation $\Delta K = A(t)(K + K_0)\Delta t$, where K is the node's degree, ΔK is its increment, $A(t)$ is the aging constant, and K_0 is the initial attractivity. This equation has been commonly used to validate the preferential attachment mechanism. We show that this equation is undiscriminating and holds for the fitness model [Caldarelli *et al.*, Phys. Rev. Lett. **89**, 258702 (2002)] as well. In other words, accepted method of the validation of the microscopic mechanism of network growth does not discriminate between "rich-gets-richer" and "good-gets-richer" scenarios. This means that the growth mechanism of many natural complex networks can be based on the fitness model rather than on the preferential attachment, as it was believed so far. The fitness model yields the long-sought explanation for the initial attractivity K_0, an elusive parameter which was left unexplained within the framework of the preferential attachment model. We show that the initial attractivity is determined by the width of the fitness distribution. We also present the network growth model based on recursive search with memory and show that this model contains both the preferential attachment and the fitness models as extreme cases.

DOI: 10.1103/PhysRevE.97.062310

I. INTRODUCTION

Power-law distributions were brought to the attention of scientific community about a century ago [1–3], and they made a sharp contrast with previously known Gaussians. While continuous power-law distributions were usually attributed to the multiplicative random noise, the generative mechanism of discrete power-law distributions remained elusive until de Solla Price suggested his cumulative advantage model which he developed by studying network of citations to scientific papers [4]. This model assumes a network consisting of nodes that appear with constant rate N, each node extending $\sim c$ edges to other nodes. The probability of attachment between a new node i and a target node j is

$$\Pi_{ij} = \frac{K_j + K_0}{\sum_l (K_l + K_0)}, \tag{1}$$

where K_j is the target node's in-degree (the number of incoming edges) and the sum is over all nodes. The initial attractivity K_0 ensures that newly born nodes start to acquire edges immediately.

Equation (1) yields a power-law degree distribution, $p(K) \sim K^{-\gamma}$ with the exponent

$$\gamma = 2 + \frac{K_0}{c}. \tag{2}$$

Price didn't look for experimental verification of Eq. (2), it was sufficient for him that Eq. (1) yields the power-law distribution with $\gamma \geqslant 2$ which is very similar to well-documented Pareto distributions of bibliometric indicators captured by Lotka's, Bradford's, and Zipf's laws. In the absence of any clue, Price postulated $K_0 = 1$.

Price's cumulative advantage model didn't spread beyond the information science community since citation network was the only complex network then known. Following proliferation of digitized information in 1990s, a number of information, biological, and social complex networks came to the forefront of scientific research, most of them exhibiting power-law degree distributions with $\gamma \sim 3$ [5–8]. To account for these distributions Barabasi and Albert suggested the preferential attachment model [9], which is very similar to but not identical with the Price's cumulative advantage. The core assumption of the Barabasi-Albert model is that a newly born node i attaches to a target node j with probability

$$\Pi_{ij} \sim K_j, \tag{3}$$

where K_j is the target node's degree for undirected networks and the sum of in- and out-degrees, $K_j = K_j^{\text{in}} + K_j^{\text{out}}$, for directed networks. Newman [7] showed that for directed networks Eq. (3) can be mapped onto the Price's model. Indeed, since statistical distribution of out-degrees in most complex networks is narrow, Eq. (3) can be written as $\Pi_{ij} \sim (K_j^{\text{in}} + c)$ where $c = \overline{K_j^{\text{out}}}$. On another hand, Eq. (1) can be written as $\Pi_{ij} \sim (K_j^{\text{in}} + K_0)$. These equations are equivalent and c plays the role of initial attractivity K_0. It should be noted, however, that while Price conjectured $K_0 = 1$, the Barabasi-Albert model postulates $K_0 = c$. Both models generate networks with the power-law degree distribution although with different exponents: $\gamma \geqslant 2$ for the former and $\gamma = 3$ for the latter. Equation (2) of the Barabasi-Albert model correctly predicts the exponent of the measured power-law distributions in many complex networks and that is why that model became the paradigm for complex network research (see Appendix A 1 a), much in the same way as the Ising model established itself as a paradigm for studies of magnetism.

*michael.golosovsky@mail.huji.ac.il

After preferential attachment has been generally accepted as the most plausible model of network growth, the next challenge became to validate it by measurements, namely, to check the validity of Eqs. (2) and (1) as well. The best way to do this is to trace evolution of individual nodes and to check whether it conforms to Eq. (1). This requires shifting the perspective from incoming nodes to target nodes. To perform corresponding modification of Eq. (1) we denote by ΔK_j the number of new edges that a node j garners between time t and $t+\Delta t$. One can show (see Appendix A 1 b) that

$$\Delta K_j = \tilde{A}(t_j)(K_j + K_0)\Delta t, \tag{4}$$

where t_j is the node's age at time t, K_j is its current degree, and $\tilde{A}(t_j)$ is the aging function, the same for all nodes. Numerous measurements of growth dynamics of complex networks validated Eq. (4) and yielded a very small initial attractivity, $K_0 \approx 1$ (see Appendix A 1 b). This finding poses a problem for the preferential attachment model. On the one hand, most complex networks are characterized by the power-law distribution with $\gamma \sim 3$ and $c = \overline{K_j^{\text{out}}} \gg 1$. According to Eq. (2), this implies initial attractivity $K_0 \approx c \gg 1$ while the directly measured initial attractivity is much smaller, $K_0 \approx 1$. So far, this inconsistency was swept under the rug and didn't affect wide popularity of the preferential attachment model.

This model was eagerly embraced by the complex network community and overshadowed alternative models, the most important of the latter being the fitness model suggested by Caldarelli *et al.* [10] and further developed in Refs. [11–15]. The fitness model assumes that the propensity of a node to attract edges is determined by some static node's attribute named fitness—a constant number that can include similarity [16–20] (also known as as homophily in social networks [21]) and other factors which shall be determined from measurements. The probability of a new node i to attach to some target node j is

$$\Pi_{ij} \sim \eta_j A(t_i - t_j), \tag{5}$$

where η_j is the target node's fitness, $t_i - t_j$ is the age of node j with respect to node i and $A(t_i - t_j)$ is the aging function. Equation (5) becomes similar to Eq. (3) after aging function is introduced in the latter as well (see Appendix A 1 a). Namely, both equations state that the attachment probability is determined by some target node's attribute: node's degree in Eq. (3) and node's fitness in Eq. (5). The crucial difference between the two is that the node's degree is the dynamic attribute which changes during the growth process while the node's fitness is the static attribute that does not vary with time.

The goal of our study is the reevaluation of the preferential attachment as the plausible growth mechanism of natural complex networks. Conventional evaluation protocol of the preferential attachment model is based on Eq. (4). We show here that this equation is not specific to the preferential attachment model and if the network growth follows the fitness model [Eq. (5)], Eq. (4) also holds. This means that even if the validity of Eq. (4) for some growing network has been established, this is is not a sufficient proof that this network grows according to the preferential attachment model, it can grow according to the fitness model as well.

II. FITNESS MODEL WITH BROAD FITNESS DISTRIBUTION MIMICS PREFERENTIAL ATTACHMENT

Consider a directed acyclic network that grows according to Eq. (5). Every node is endowed with a certain fitness η drawn from some distribution $\rho(\eta)$ where $\int_0^\infty \rho(\eta)\,d\eta = 1$. We assume that the node's fitness remains constant during node's life. We further assume that K, the degree of each node, grows following an inhomogeneous Poisson process, in such a way that ΔK, the number of edges garnered by a node during time window $(t, t+\Delta t)$, is represented by the Poisson distribution, $P(\Delta K;\lambda) = \frac{\lambda^{\Delta K}}{\Delta K!}e^{-\lambda}$. The Poissonian rate,

$$\lambda = \eta A(t)\Delta t, \tag{6}$$

is node-specific and is determined by node's fitness η. $A(t)$ is the aging function which is normalized as follows: $\int_0^\infty A(\tau)\,d\tau = 1$. Under this constraint, the fitness η is the long-time limit of node's degree, namely, $\eta \sim K(t\to\infty)$.

Since Eq. (6) is memoryless, the number of edges that each node garners through the period from $t=0$ to t also follows Poisson distribution with the node-specific rate

$$\Lambda = \eta \int_0^t A(\tau)\,d\tau. \tag{7}$$

We consider the set of N nodes that joined the network at the same moment which we set as $t=0$. Among these, we focus on the subset of nodes that garnered K edges by time t. Their number is

$$n(K,t) = N\int_0^\infty \frac{\Lambda^K}{K!}e^{-\Lambda}\rho(\eta)\,d\eta, \tag{8}$$

where $\rho(\eta)$ is the fitness distribution and $\Lambda(\eta)$ dependence is given by Eq. (7). During time window $(t, t+\Delta t)$ each of these $n(K,t)$ nodes garners $\sim\lambda(\eta)$ edges, in such a way that the average number of new edges garnered by a node from this subset is

$$\overline{\Delta K} = \frac{N\int_0^\infty \lambda \frac{\Lambda^K}{K!}e^{-\Lambda}\rho(\eta)\,d\eta}{n(K,t)}. \tag{9}$$

We substitute Eq. (6) into Eq. (9), note that $\lambda = \Lambda\tilde{A}(t)$, where $\tilde{A}(t) = \frac{A(t)}{\int_0^t A(\tau)\,d\tau}$, use the equality

$$\Lambda P(K;\Lambda) = (K+1)P(K+1;\Lambda) \tag{10}$$

and come to

$$\overline{\Delta K} = \tilde{A}(t)(K+1)\frac{n(K+1,t)}{n(K,t)}\Delta t. \tag{11}$$

Here $n(K+1,t)$ is the number of nodes that garnered $K+1$ edges by time t. For broad fitness distributions and for $K \gg 1$, $n(K+1,t) \approx n(K,t)$ (see Appendix B), in such a way that Eq. (11) reduces to

$$\overline{\Delta K} = \tilde{A}(t)(K+1)\Delta t. \tag{12}$$

This expression is nothing else but Eq. 4 with $K_0 = 1$. A similar result was obtained earlier by Burrell [22] using a different approach. Note that Eq. (12) reduces to $\frac{\overline{\Delta K}}{\Delta t} = \tilde{A}(t)(K+1)$. This has uncanny resemblance to the famous expression describing the photon emission rate for two-level atomic systems, $\frac{dN_{ph}}{dt} = Bn_2(N_{ph}+1)$, where n_2 is the number of atoms in the

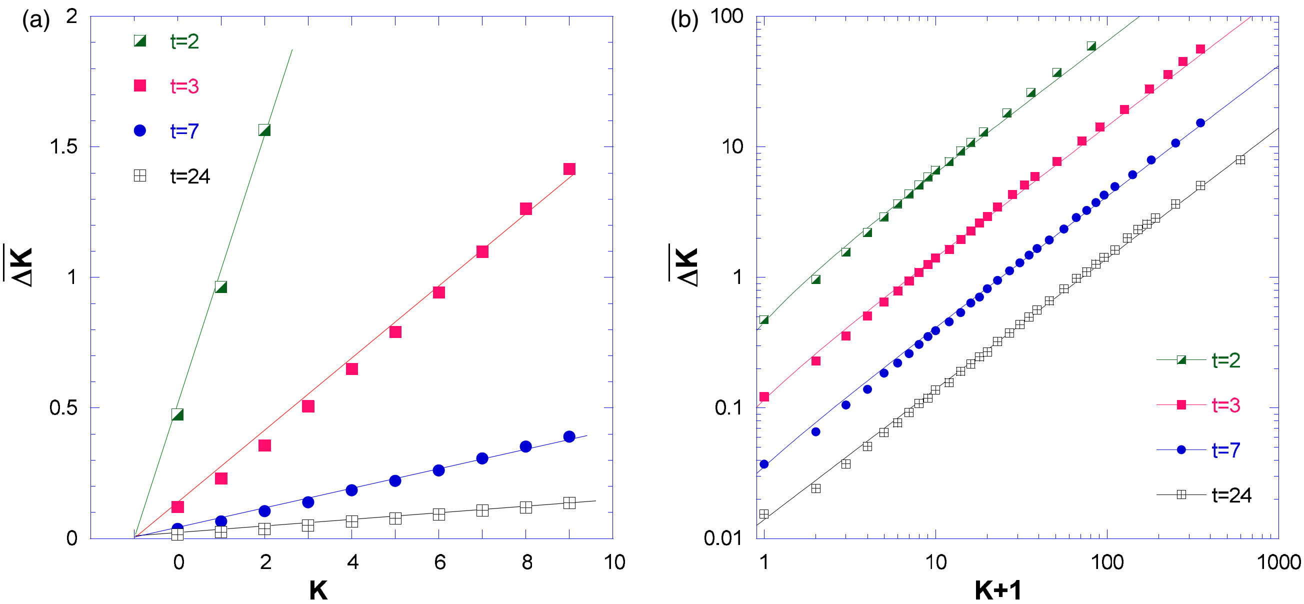

FIG. 1. Numerical simulation of the growth dynamics of 400 000 nodes with the aging function $A(t) = \frac{0.035t}{|t-2.4|^{1.3}}$ and lognormal fitness distribution with $\mu = 1.6$ and $\sigma = 1.1$. $\overline{\Delta K}$ is the mean growth rate, K is the number of accumulated edges, and t is the age. (a) $\overline{\Delta K}$ versus K for small K. The data for all t lie on straight lines with common intercept ~ -1, suggesting $K_0 \approx 1$. (b) $\overline{\Delta K}$ versus K for all K. Continuous lines show fits to Eq. (12) with $K_0 = 0.7$, 0.8, 0.85, and 1 for $t = 2$, 3, 7, and 24, correspondingly.

upper state, N_{ph} is the number of photons, and B is the Einstein coefficient for stimulated emission. Thus, the node degree K is the analog of N_{ph} and K_0 is the analog of spontaneous emission.

To validate Eqs. (11) and (12) through numerical simulation we considered a set of 400 000 nodes with a lognormal fitness distribution $\rho(\eta) = \frac{1}{\sqrt{2\pi}\sigma\eta} e^{-\frac{(\ln\eta-\mu)^2}{2\sigma^2}}$ where $\mu = 1.6$ and $\sigma = 1.1$ We simulated the growth of these nodes using Eq. (6) and the aging function $A(t) = \frac{0.035t}{|t-2.4|^{1.3}}$. (The distribution and the aging function were taken from our recent measurements of citation dynamics of physics papers [23].) The time was run from $t = 0$ to $t = 25$ with steps $\Delta t = 1$, in such a way that $\sum_0^{t=25} A(t) = 1$. For each node j in this set we determined $K_j(t)$, the total number of edges accumulated after time t, and $\Delta K_j(t)$, the number of additional edges gained between t and $t+1$. For every t we grouped all nodes into 40 logarithmically spaced bins, each bin containing the nodes with close values of K. For each bin, we determined ΔK distribution and found its mean, $\overline{\Delta K}$. Figure 1(a) plots $\overline{\Delta K}$ versus K for small K. We observe straight lines with common intercept ~ -1 as suggested by Eq. (12). To fit the whole $\overline{\Delta K}(K)$ dependence we used the following equation:

$$\overline{\Delta K} = \tilde{A}(t)(K + K_0), \tag{13}$$

where K_0 is the fitting parameter. Figure 1(b) shows that this equation fits the data fairly well for all K.

Figure 2 shows $\overline{\Delta K}$ versus $(K + K_0)$ dependences for lognormal fitness distributions with different σ and for time slices $\Delta t = 1$. These dependences are also well fitted by Eq. (13) with $K_0 \sim 1$.

To estimate K_0 from the data more precisely, we turn to Eq. (13). It indicates that at small K, $\overline{\Delta K} \to \tilde{A}(t)K_0$. On another hand, Eq. (11) yields for $\Delta t = 1$

$$\overline{\Delta K}|_{K=0} = \tilde{A}(t)\frac{n(1,t)}{n(0,t)}. \tag{14}$$

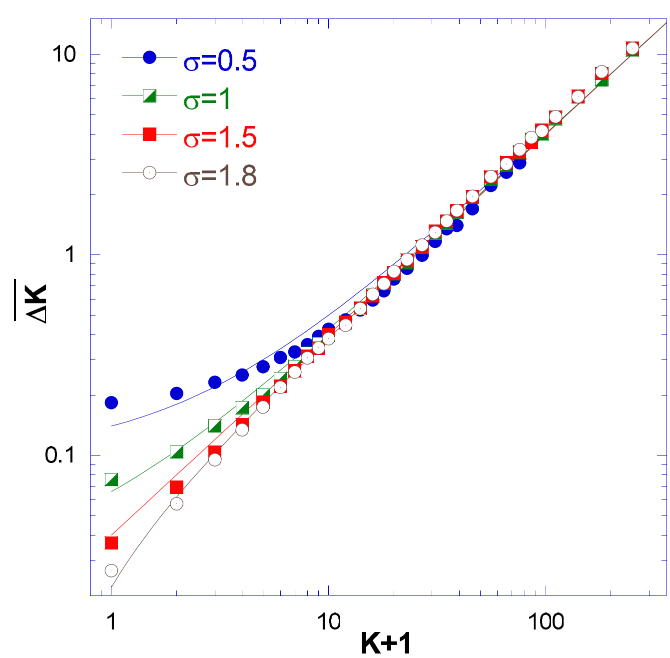

FIG. 2. Numerical simulation of the growth dynamics of the set of 400 000 nodes with different lognormal fitness distributions having the same $\mu = 1.6$ and different σ. The symbols show results of numerical simulation, continuous lines show linear approximation $\overline{\Delta K} = A(K + K_0)$ with $A = 0.04$ and $K_0 = 3.5, 1.65, 1$, and 0.55 for $\sigma = 0.5, 1, 1.5$, and 1.8, correspondingly.

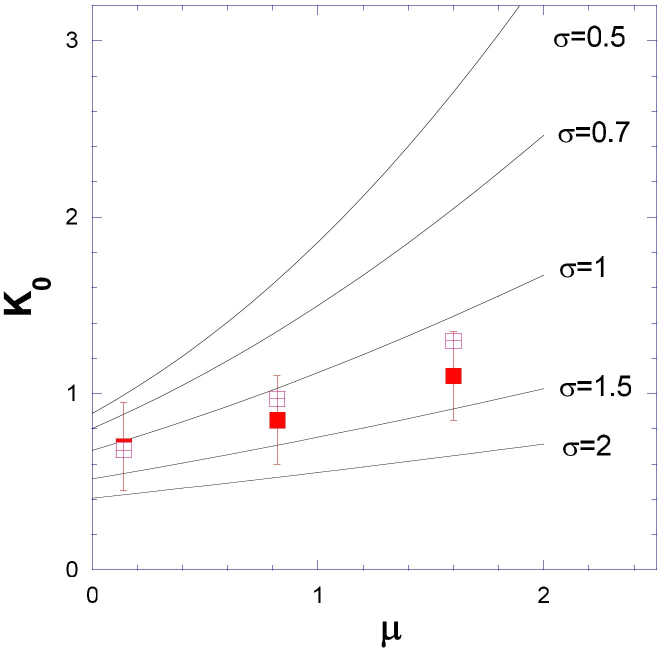

FIG. 3. Initial attractivity K_0 estimated from Eq. 15 in dependence of the parameters of the lognormal fitness distribution, μ and σ. The filled squares show our measurements for physics, economics, and mathematics papers published in 1984 (see Ref. [24]); the open squares show our expectations based on measured μ and σ of the lognormal fitness distribution for these very data sets. The measured values of K_0 are close to those predicted by Eq. (15).

Thus,

$$K_0 \approx \frac{n(1,t)}{n(0,t)} = \frac{\int_0^\infty \Lambda e^{-\Lambda} \rho(\Lambda)\, d\Lambda}{\int_0^\infty e^{-\Lambda} \rho(\Lambda)\, d\Lambda}. \quad (15)$$

We note that $\rho(\Lambda)$ follows the lognormal distribution which is nothing else but the fitness distribution with shifted mean, $\mu' = \mu + \log[\int_0^t A(\tau)\, d\tau]$. Since $\int_0^t A(\tau)\, d\tau \to 1$ in the long time limit, the difference between μ and μ' becomes increasingly small at long t. Figure 3 shows K_0 calculated according to Eq. (15) as a function of μ and σ. We observe that K_0 increases with μ and decreases with σ. These dependences can be captured by the approximate empirical expression

$$K_0 \approx \frac{e^{\frac{\mu}{1+\sigma}}}{(1+\sigma^2)^{0.6}}. \quad (16)$$

For reasonable values of μ from 0 to 2 and σ from 1 to 2, K_0 lies between 0.5 and 1.5. It is determined by σ, and, to a lesser extent, by μ. All this means the following: if K_0 is measured using Eq. (13) using extrapolation from large K, one always gets $K_0 = 1$. On another hand, since most fitness distributions are broad, then the estimates made using Eq. (13) for small K, as it is usually done in most studies, yield $K_0 = 0.5$–1.5. Figure 3 shows that for narrow fitness distribution, $\sigma \ll 1$, K_0 can be higher.

We plot in Fig. 3 the measured values of K_0 which were inferred from our studies of citation dynamics of scientific papers published in 1984. We considered three research fields: physics, economics, and mathematics and found that the fitness distributions for all these fields are lognormals with different μ but the same, $\sigma = 1.1$. The measured and calculated initial attractivities K_0 are in good agreement and are all close to 1.

Thus, our numerical simulation supports Eq. (13) with initial attractivity $K_0 \approx 1$, as it was postulated by de Solla Price [4]. The natural question arises—why is $K_0 \approx 1$ so widespread? Figure 3 shows that $K_0 \approx 1$ corresponds to $\sigma = 1$–1.5 irrespective of μ. Nguyen and Tran [14] used numerical simulation to study complex networks with lognormal fitness distribution that grow according to Eq. (5). They found that the resulting network structure strongly depends on the width of the fitness distribution σ, in particular, the power-law degree distribution appears only for $\sigma \approx 1$ and its exponent γ is close to 3. This observation implies that the initial attractivity is coupled to the exponent of the degree distribution, in such a way that within the framework of the fitness model, the universality of $K_0 \approx 1$ in complex networks is a consequence of the fact that most of them exhibit power-law degree distributions with $\gamma \sim 3$.

Equation (13) is commonly accepted as an evaluation tool of growth mechanism of real complex networks. We demonstrate here that this equation is undiscriminating and holds for the preferential attachment and the fitness model as well. It should be noted that both models are phenomenological, namely, to explain the degree distribution in growing complex networks these models make plausible but unsubstantiated assumptions regarding the microscopic mechanism of attachment. In fact, the preferential attachment model deduces the attachment mechanism from the degree distribution. On another hand, there are models that start from some realistic attachment mechanism and deduce the degree distribution and other network properties basing on this mechanism. In what follows we demonstrate one of such models based on recursive search. We have developed this model to account for citation dynamics of scientific papers [23], and in what follows we generalize it beyond citation networks. This model unifies the fitness and the preferential attachment mechanisms in one framework.

III. RECURSIVE SEARCH MODEL

Consider a growing unweighted directed acyclic network. The nodes enter one by one, every new node is endowed by a certain fitness η, and it extends a certain number of edges to older nodes. Then the time increases by one unit, a new node enters, and the process repeats itself.

A new node i can chose the target node j in two ways, as it is illustrated in Fig. 4. The direct attachment occurs when the choice is based on target node's fitness. We assume that the probability of this process is

$$\pi_{ij} = \eta_j A(t_i - t_j), \quad (17)$$

where η_j is the target node's fitness, t_i, t_j are the moments when the nodes i, j joined the network, $t_i - t_j$ is the age of the target node j with respect to the new node i, and $A(t_i - t_j)$ is the aging function.

Alternatively, the new node i can find the target node j in the network neighborhood of one of the previously chosen nodes k. We denote by Θ_{ikj} the probability of indirect attachment through the node k and assume that after node i connects to the intermediate node k, it can attach to any of its ancestors j with equal probability which does not depend on the attributes

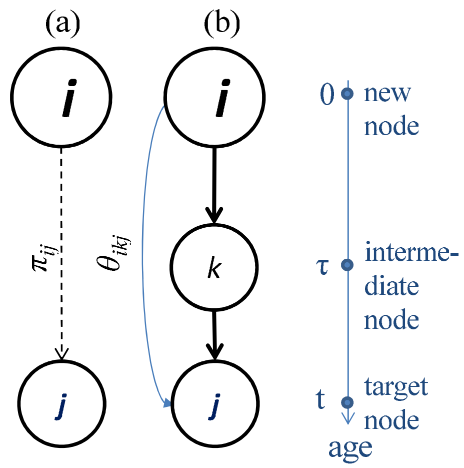

FIG. 4. (a) Direct attachment. A new node i finds the target node j through fitness-based search and attaches to it with probability π_{ij} given by Eq. (17). (b) Indirect attachment. The node i finds node j by exploring the network neighborhood of an already connected node k and attaches to node j with probability Θ_{ikj} given by Eq. (18).

of node j but is determined by the age of node k with respect to node i, namely,

$$\Theta_{ikj} = \frac{T(t_i - t_k)}{K_k^{\rm out}}. \tag{18}$$

Here $K_k^{\rm out}$ is the out-degree of node k and $T(t_i - t_k)$ is the memory function. Θ_{ij}, the total probability of indirect attachment between the nodes i and j, is the sum of Θ_{ikj}-s over all intermediate nodes k which lie on two-hop paths connecting nodes i and j. Using adjacency matrix ($a_{mn} = 1$ if nodes m,n are connected, and $a_{mn} = 0$ otherwise) we write this probability as the sum over all nodes l,

$$\Theta_{ij} = \sum_l a_{il} a_{lj} \Theta_{ilj}. \tag{19}$$

The probability of a new node i to extend an edge to an older node j is the sum of direct and indirect contributions,

$$\Pi_{ij} = \frac{(\pi_{ij} + \Theta_{ij})(1 - a_{ij})}{\sum_l (\pi_{il} + \Theta_{il})(1 - a_{il})}, \tag{20}$$

where the sum in denominator is over all nodes, and the factor $(1 - a_{ij})$ takes into account that if two nodes are already connected, no other connection between them can be established anymore. We cast Eq. (20) in a more compact form, $\Pi_{ij} = \eta_j \tilde{A}(t) + \tilde{\Theta}_{ij}$, where $\tilde{A}_t$ is the normalized aging constant and $\tilde{\Theta}_{ij}$ is the normalized probability of indirect attachment.

To analyze Eq. (19) we adopt the mean field approximation and transform the sum to an integral. Since the memory function $T(t_i - t_k)$ depends only on the age of the target node [Eq. (18)], we replace the time by age. We focus on some new node i. Its age is zero, the age of the target node j is t, and the age of the intermediate node k is τ (see Fig. 4). We assume that the nodes join the network at constant rate N and consider $N\delta\tau$ old nodes that joined the network in the narrow time window $\tau,\tau + \delta\tau$. A new node i has $k_i^{\rm out}(\tau)\delta\tau$ edges ending up in this network slice, while the target node j has $k_j^{\rm in}(t - \tau)\delta\tau$ edges originating in this slice. [The functions $k_l^{\rm out}(\tau)$, $k_l^{\rm in}(\tau)$ are defined for every node l through the following relations: $\int_0^\infty k_l^{\rm out}(\tau)\,d\tau = K_l^{\rm out}$, $\int_0^t k_l^{\rm in}(\tau)\,d\tau = K_l^{\rm in}(t)$, where $K_l^{\rm out}$ is the total number of outgoing edges and $K_l^{\rm in}(t)$ is the total number of incoming edges accumulated by the node l by time t.] Without loss of generality we assume that all nodes extend the same number of outgoing edges, $K_l^{\rm out} = c$. For sufficiently small $\delta\tau$, the number of intermediate nodes k connecting nodes i and j that belong to the slice $(\tau,\tau + \delta\tau)$ is $\sim \frac{k_i^{\rm out}(\tau) k_j^{\rm in}(t-\tau)}{N}\delta\tau$. Each of these nodes can induce indirect attachment of the node i to node j with probability $T(\tau)/c$. The probability of attachment between the nodes i and j is the probability of direct attachment plus contributions of all such slices,

$$\Pi_{ij} = \eta_j \tilde{A}(t) + \int_0^t \frac{k_i^{\rm out}(\tau) T(\tau)}{Nc} k_j^{\rm in}(t - \tau)\,d\tau. \tag{21}$$

To analyze Eq. (21) we replace the kernel $\frac{k_i^{\rm out}(\tau)T(\tau)}{Nc}$ by the exponential $qe^{-\Gamma\tau}$ (such replacement has been justified in our recent measurements of citation dynamics [23]). Then Eq. (21) reduces to

$$\Pi_{ij} = \eta_j \tilde{A}(t) + \int_0^t q e^{-\Gamma(t-\tau)} k_j(\tau)\,d\tau, \tag{22}$$

where the right-hand side depends only on the properties of the target node j. We also swapped τ and $t - \tau$ under the integral using the properties of convolution. If Γ is very small, then Eq. (22) reduces to

$$\Pi_{ij} \approx \eta_j \tilde{A}(t) + q K_j(t), \tag{23}$$

where $K_j(t)$ is the total degree of the node j at time t. Equation (23) has been considered by Refs. [16,25–27], and it is nothing else but the preferential attachment with additive fitness (see Appendix A 2 b). This equation is also at the core of the Bass model for diffusion of innovations in the infinite market.

In the opposite case of large Γ, the main contribution to the integral in Eq. (22) comes from recent edges garnered between t and $t - 1/\Gamma$. Thus, we can retain in the integral only those τ that fall in the time window $(t, t - 1/\Gamma)$, namely, $t - \tau \ll t$. In view of this relation we approximate $k_j(\tau)$ by $k_j(t) - (t - \tau)\frac{dk_j}{d\tau}|_t$, perform integration, and after some algebra arrive at

$$\Pi_{ij} \approx \eta_j \tilde{A}(t) + \frac{q}{\Gamma} k_j(t - 1/\Gamma). \tag{24}$$

Equation (24) indicates that the attachment probability is determined by the recent growth rate- the dominant feature of the self-exciting (Hawkes) process. Similar equation was suggested by Refs. [12,28,29] under the name of preferential attachment with gradually vanishing memory.

To validate Eq. (21) by measurements we need to shift the perspective from the incoming node to the target node. We focus on one such target node j and consider N new nodes i that joined the network during time window $(0,\Delta t)$. From this batch of new nodes the node j garners $\Delta K_j \approx \overline{\Pi_{ij}} c N \Delta t$ edges where Π_{ij} is given by Eq. (21). The averaging is performed over the batch of new nodes i. We also note that $\overline{k_i^{\rm out}(\tau)} = \overline{k_k^{\rm in}(\tau)}$. This relation stems from the fact that the outgoing edge from the node i is the incoming node for the node k, as it is shown in Fig. 4. With respect to the node j, the above function is nothing else but the age-resolved nearest-neighbor connectivity, $\overline{k_k^{\rm in}(\tau)} = k_j^{nn}(\tau)$. In view of this relation we transform Eq. (21) into

$$\Delta K_j = \left[\eta_j \tilde{A}(t)Nc + \int_0^t k_j^{nn}(t-\tau)T(t-\tau)k_j^{\rm in}(\tau)\,d\tau\right]\Delta t, \tag{25}$$

where we swapped τ and $t-\tau$ in the integral. The right-hand side of Eq. (25) contains only attributes of the node j making this equation similar to Eq. (13). For small t the fitness term dominates, in such a way that Eq. (25) reduces to Eq. (5). Hence, this is no surprise that for broad fitness distribution it reduces to Eq. (13) with initial attractivity $K_0 \approx 1$. Indeed, our measurements of citation dynamics validated Eq. (25) and the data shown on Fig. 3 were obtained using this very equation.

IV. DISCUSSION

Our analysis shows that if network growth is considered from the perspective of a target node and is studied using the mean-field approximation, namely, by averaging over many similar nodes, one cannot distinguish between the preferential attachment and the fitness models; both of them yield Eq. (4). Thus, in all that concerns the mean-field network dynamics, preferential attachment model is equivalent to fitness model, in other words, the rich-gets-richer reduces to the fit-gets-richer (good-gets-richer) [10,30]. This is surprising since these two models are based on different premises. The preferential attachment model assumes that all nodes are born equal, the inequality in their degree coming by chance. After this inequality has been established, it is amplified by the autocatalytic process represented by Eq. (1). In contrast, the fitness model and the fitness-based recursive search model assume that the nodes are born unequal, each newly born node is endowed with a certain fitness. The latent inequality in fitnesses becomes evident when the nodes have been developing for some time. Surprisingly, the two opposing assumptions underlying network growth—all nodes are born equal or different—result in the same Eq. (4).

This does not mean that the two models are equivalent. While the preferential attachment model does not specify the initial attractivity, the fitness model with aging explains it perfectly well: it is determined by the shape of the fitness distribution. With respect to the power-law degree distribution in complex networks: the preferential attachment relates its to the strategy by which the new node attaches to old nodes, while the fitness model implies that this distribution is inherited from the fitness distribution. The fitness model successfully explains the first-mover advantage, degree distribution for the nodes of the same age, different trajectories of the nodes of the same age, etc. (see Appendix A 1 c). However, this model does not account for the nonlinear growth observed in some networks.

Although it could seem that the fitness model is a more appropriate framework to conceptualize network growth, Eqs. (1) and (4) can still be valid since the preferential attachment is a structural rather than explanatory model. Indeed, the relation $\Pi_{ij} \sim K_j$ does not imply that a new node i crawls through the whole network in order to gain information about degrees of all other nodes j. What occurs in reality is that the network grows following some local rule and this rule becomes imprinted in the network topology. When the network growth is analyzed, the changes in topology are visible while the underlying microscopic growth rule is not. This feeds the illusion that the growth dynamics is determined by network topology while in reality the reverse is true.

The challenge is to uncover the microscopic rules of network growth that produce the given network topology. We showed that the recursive search is one of the plausible microscopic mechanisms of network growth. What is the relation of this mechanism to the genuine preferential attachment, namely, the algorithm whereby a new node finds well-connected older nodes and attaches to them? It has been generally believed (see Appendix A 4) that the recursive search is one of realizations of this algorithm, since if a new node makes a random choice among the neighbors of already chosen nodes, it has high probability of picking up highly connected nodes. We demonstrate here [Eq. (23)] that this strategy works in a straightforward way only if the recursive search does not have memory. In reality, recursive search has rather short memory [23], and it is not clear whether highly connected nodes can be found by this simple strategy: random choice among the neighbors of already chosen nodes. In our studies of citation dynamics we found [23] that the recursive search there follows a more clever strategy: the search in the network neighborhood of the previously chosen nodes is not random but has preference for those neighbors that are connected to several already chosen nodes. The cartoon picture of this strategy is as follows. Simple recursive search: if Alice is linked to Bob, and Bob is linked to Frank, there is a chance that Alice will link to Frank. Clever recursive search: if Alice is linked to Bob and Charlie, and both of them are linked to Frank, then Alice will link to Frank almost for sure. Thus, if a new node identifies a target node in the network vicinity of two or more previously chosen nodes, the probability of attachment to such node exceeds the sum of probabilities per each path, namely, multiple paths interfere constructively, reinforcing one another. The synergetic interaction between the paths to the next-nearest neighbors ensures that a new node finds highly connected nodes. This strategy of exploring next-nearest neighbors can still be considered as a local strategy, but in fact, it is one step towards global search and this is one of the ways how the genuine preferential attachment emerges within the framework of the recursive search model.

ACKNOWLEDGMENTS

I am grateful to Lev Muchnik for valuable discussions.

APPENDIX A: ANALYSIS OF THE NETWORK GROWTH MODELS

1. Preferential attachment

a. Generalization

In what follows we do not make distinction between the Price's and Barabasi-Albert approaches and relate to Eq. (1) with unspecified K_0 as the preferential attachment model. Its success in explaining the seemingly universal power-law degree distribution in complex networks prompted several theoretical generalizations:

(1) *Initial attractivity*. References [31,32] analyzed Eq. (3), in which the attachment rule was modified to $\Pi_{ij} \sim (K_j + K_0)$, where K_j is the total degree and K_0 is an arbitrary number. It was found that the power-law degree distribution is retained but its exponent is modified in accordance with Eq. (2).

(2) *Accelerated network growth*. While the original Barabasi-Albert model assumed that new nodes and edges appear at the same rate, Leskovec *et al.* [33] found that the number of edges often grows faster than the number of nodes, in such a way that these networks shrink with time. Assuming that the average number of outgoing edges per node increases exponentially with time, $c = c_0 t^{\Theta}$, Ref. [8] found that the power-law degree distribution in these shrinking networks is retained but its exponent increases by $\Delta\gamma = \frac{2\Theta}{1-\Theta}$.

(3) *Link editing.* In the original Barabasi-Albert model the existing links can't be deleted, the links between old nodes can't be added, and the old nodes can't be removed. Although this is true for citation networks, other networks such as Wikipedia edits, allow link or node editing. Ghoshal *et al.* [32] analyzed networks for which link editing and node removal are allowed and showed that these processes act disruptively on network topology, although there is a wide range of parameters for which the power-law degree distribution is conserved.

(4) *Aging.* Equation (1) assumes that the attachment probability Π_{ij} does not depend on node's age. However, the common sense tells us that Π_{ij} should decrease with increasing age of the target node, in such a way that recent nodes become more popular. References [28,34–37] considered a general case of the preferential attachment with aging

$$\Pi_{ij} = \frac{A(t_j)(K_j + K_0)}{\sum_l A(t_l)[K_l + K_0]}, \tag{A1}$$

where t_j is the age of node j with respect to node i and $A(t)$ is the aging function which is usually assumed to follow exponential or power-law dependence, $A(t) \propto 1/t^{\nu}$. In the latter case Dorogovtsev and Mendes [34] showed that the power-law degree distribution is retained only for $\nu < 1$ while for $\nu > 1$ the aging effect overcomes the preferential attachment and the degree distribution does not follow the power-law dependence.

(5) *Memory*. Although Eq. (A1) accounts for aging, it attributes equal weight to all edges, as if the attachment were a Markov process. However, the recent edges are usually more important than the older ones, namely, attachment process can have memory. To account for memory, Refs. [23,28,29,37–40] replaced Eq. (A1) by

$$\Pi_{ij}(t_j) \propto \int_0^{t_j} A(t_j - \tau) k_j(\tau)\, d\tau, \tag{A2}$$

where $\Delta K_j(\tau) = k_j \delta\tau$ is the number of edges garnered by the node j in the time window $(\tau, \tau + \delta\tau)$, t_j is the node's age, and $A(t_j - \tau)$ is the memory kernel.

(6) *Nonlinear preferential attachment.* While Eq. (1) assumes a linear relation between the attachment probability Π_{ij} and the node's degree K_j, Krapivsky and Redner [41] considered a general case of nonlinear preferential attachment

$$\Pi_{ij} = \frac{(K_j + K_0)^{1+\delta}}{\sum_l (K_l + K_0)^{1+\delta}}, \tag{A3}$$

where $\delta \neq 0$. It was shown that the power-law degree distribution is associated with the linear case, $\delta = 0$. For superlinear attachment, $\delta > 0$, the network becomes the hub-and-spoke or "winner-takes-all," while for sublinear attachment, $\delta < 0$, the network becomes like a gel where every node is connected to all other nodes and degree distribution is the stretched exponential rather than power-law. Subsequently, Krapivsky and Krioukov [42] showed that for weak superlinear attachment, $0 < \delta \ll 1$, there is a vast asymptotic regime for which the network retains its power-law degree distribution.

In summary, theoretical studies indicate that the preferential attachment mechanism is plausibly robust. Namely, it generates complex networks with the power-law degree distribution and the exponent $2 < \gamma \leqslant 3$ under the following conditions: the attachment probability is linear or weakly nonlinear, aging is weak, and initial attractivity is positive and small. These conditions are quite reasonable and that is why the preferential attachment model has been accepted as the most plausible generative mechanism of growing complex networks with the power-law degree distribution.

b. Model validation by measurements

The microscopic mechanisms of the evolution of growing networks and the protocols of their validation were discussed in Refs. [20,43–46]. In what follows we present a more specific overview with the focus on preferential attachment.

Straightforward verification of the preferential attachment model requires analysis of decisions made by incoming nodes. The measurements aimed at quantitative analysis of such decisions are widespread in psychology but they are rare in physics, biology, and computer science, the fields where majority of complex networks appear. A more conventional way to uncover the growth mechanism of complex networks is to trace evolution of individual nodes. To this end, the perspective shall be shifted from the incoming node to the target node. To perform such shift we go back to Eq. (A3) and assume that new nodes appear at a constant rate N. Consider N new nodes that joined the network during time window $(t, t + \Delta t)$. Each new node extends in average c edges to existing nodes. Consider one such target node j. From the batch of new nodes, it garners approximately $\Delta K_j = \Pi_{ij} c N \Delta t$ edges where Π_{ij} is the attachment probability. We substitute there the general expression for Π_{ij}, which includes nonlinearity, aging, and initial attractivity, but not memory, and find

$$\Delta K_j = \tilde{A}(t_j)(K_j + K_0)^{1+\delta} \Delta t, \tag{A4}$$

where t_j is the target node's age at time t, K_j is its current degree, K_0 is the initial attractivity, and the aging function is $\tilde{A}(t) = \frac{A(t)}{\sum_l A(t_l)[K_l + K_0]^{1+\delta}} cN$. [Note the difference between $A(t)$ and $\tilde{A}(t)$: for the Barabasi-Albert model $A(t) = 1$ while $\tilde{A}(t) = \frac{2N}{t}$.] Equation (A4) is the basis for comparison of the preferential attachment model to measurements.

To validate Eq. (A4) one usually considers a set of all nodes of the same age and measures each one's degree at time t and at $t + \Delta t$. Then one calculates $\Delta K_j = K_j(t + \Delta t) - K_j(t)$, the number of additional edges that each node garnered during the time window $(t, t + \Delta t)$, plots ΔK_j versus K_j, and makes one's best to fit this scatter plot using Eq. (A4) [20,47,48]. This fit is by no means trivial. The catch here is that ΔK_j is a discrete stochastic variable and Eq. (A4) predicts its mean value but says nothing about the variance. Our measurements for citation networks [24] indicate that ΔK_j distribution (for fixed K_j) follows negative binomial distribution with high variance-to-mean ratio >2, in such a way that the variance of ΔK_j is considerably greater than that for the Poisson distribution. In other words, $\Delta K_j(K_j)$ dependence is so noisy that direct fitting of ΔK_j versus K_j using Eq. (A4) is not very informative.

To circumvent the problem of noise one can use logarithmic binning of K_j, plot a histogram of ΔK_j, and find the trend. This method was originated by Newman [18] and since then it has been adopted by many others [20,44,49–53].

Another way to counter the noise problem is to plot cumulative function $\int_0^K \Delta K(K)\,dK$ versus K. In the context of complex networks this procedure was first applied by Jeong *et al.* [54] and subsequently by Refs. [55,56]. However, there is a pitfall here. If ΔK_j were continuous variable with symmetric distribution, this cumulative procedure should certainly work. However, since ΔK_j is a non-negative discrete variable with highly skewed distribution, the cumulative procedure can distort the results. In particular, when applied to validation of Eq. (A4), this procedure overestimates the initial attractivity [51].

Yet another strategy is to use the raw ΔK_j versus K_j plots and to apply sophisticated numerical fitting procedure to find parameters of Eq. (A4) [44,57].

In summary, the measurements aimed at validation of Eq. (A4) showed the following:

(1) *Preferential attachment*. The growth of many complex networks does follow Eq. (A4) [20,43–46]. However, some of these networks exhibit preferential attachment [namely, linear or quasilinear $\Delta K(K)$ dependence] only for nodes with low and moderate degrees while the nodes with high degree exhibit antipreferential attachment (namely, decreasing $\Delta K(K)$ dependence) [44,49].

(2) *Linear or nonlinear PA?* Early measurements claimed linear or close-to-linear preferential attachment [18,54]. Later measurements using large data sets (citations to scientific papers [24,53] and patent citations [52,58]) revealed superlinear attachment with the exponent $1+\delta \sim 1.25$. Social networks (scientific collaboration [18], movie actors [54,55]) exhibit sublinear preferential attachment with the exponent $1+\delta = 0.8$–0.9.

(3) *Aging function.* Our measurements of citations to scientific papers [51] yielded $\tilde{A}(t) \sim (t-\Delta)^{-\nu}$, namely, a power-law decay with small delay $\Delta \sim 1-2$ yr and the exponent $\nu = 2$. Patent citations yield a similar power-law aging function with $\nu = 1.3$–1.6 [52,58]. Zeng *et al.* [59] provide an overview of aging effects in citation networks.

(4) *Initial attractivity.* Early measurements of citations to scientific papers were not statistically representative to make reliable estimate of K_0 [47,54]. Subsequent studies of patent citations yielded small $K_0 \sim 1$ [58]. Our high statistics measurements of citations to scientific papers [51] also yielded small $K_0 \sim 1$. Eom and Fortunato [55] analyzed network of citations between the American Physical Society (APS) journals and found a bigger number, $K_0 \sim 7$ for younger papers and $K_0 \sim 1$–2 for the papers that are at least five years old. (Note, however, that Ref. [55] used cumulative procedure which is known to overestimate K_0 [51]). Recent studies of Higham *et al.* [52,53] yielded $K_0 = 1$–1.8 for patent citations and $K_0 = 1$ for the *Physical Review* citations. Thus, all measured initial attractivities are small, $K_0 \ll c$, and better conform to Price's conjecture, $K_0 = 1$, than to the Barabasi-Albert conjecture, $K_0 = c$.

Thus, Eq. (A4) has been qualitatively validated for many complex networks. The attachment in most of them turned out to be linear or close to linear, although deviations from the linearity are well documented. However, Eq. (2) is not supported by measurements. Indeed, the measurements indicate small initial attractivity K_0. In this case Eq. (2) yields the power-law degree distributions with $\gamma \geqslant 2$ and this is in contrast to the power-law degree distribution with $\gamma \sim 3$ observed in majority of complex networks [51,54,55,58]. This inconsistency notwithstanding, the preferential attachment model became a paradigm of complex network growth and a platform for network characterization.

c. Specific predictions of the preferential attachment model and their validation

After scientific community became persuaded that the growth of complex networks is accounted for by the preferential attachment model, the research shifted from the model validation to analysis of its predictions. Indeed, besides the power-law degree distribution, the complex networks generated by Eq. (1) should acquire a very special structure [7,8]:

(1) *First mover advantage*. The preferential attachment model predicts strong positive correlation between the node's age and degree, namely, the degree of the old nodes should be substantially higher than that of the recent nodes. The measurements reveal such correlation but it is not strong and most new edges do not necessarily go to old nodes [60,61].

(2) *Trajectory of the nodes of the same age*. The basic preferential attachment model predicts that the node's degree grows with time according to the rule $K_j(t_j) = K_0[(1+\frac{t_j}{t})^{\frac{c}{c+K_0}} - 1]$ where t is the age of the network at the moment when the node was born and t_j is the node's age [8]. Thus, the node's degree grows with time with deceleration and the trajectories of the nodes of the same age should be very similar. However, the measurements show that these trajectories strongly diverge [24,62] and do not necessarily decelerate with time. In particular, citation networks demonstrate "sleeping beauties" [63] whose trajectories accelerate with time.

(3) *Degree distribution for the nodes of the same age*. According to the preferential attachment model, this distribution is narrow and close to exponential, $p(K) \sim K^{K_0-1}(1-t_j^{\frac{c}{c+K_0}})^K$ [7,8]. Measurements on Wikipedia and citation networks showed that degree distribution for the nodes of the same age is much wider than exponential and is better described by the power-law or lognormal function [23,47,62,64,65].

(4) *Degree-degree correlation*. Within the framework of the preferential attachment model, the assortativity of the resulting network is determined by the initial attractivity K_0 [8,18]. In particular, for $K_0 = c$ (the Barabasi-Albert model), the network shall be neutral, for $K_0 > c$ it shall be assortative, and for $K_0 < c$ (Price's model) it shall be disassortative [8]. Direct measurements of the initial attractivity yield small $K_0 \sim 1$ implying that most networks should be disassortative. While social networks are indeed disassortative, citation networks are

not [8,23,66–68]. Thus, contrary to model prediction, there is no straightforward relation between the initial attractivity K_0 and the network assortativity.

(5) *Clustering coefficient*. The preferential attachment mechanism predicts that in large networks the clustering coefficient shall be vanishingly small [69]. However, many real networks have high clustering coefficient [43].

Thus, several specific predictions of the preferential attachment model are inconsistent with measurements. This is unsurprising since these predictions were made assuming linear preferential attachment in the absence of aging [Eq. (1)]. Since most studied networks exhibit weakly nonlinear attachment and strong aging, the proper account of these two factors could modify some of the above predictions and make them consistent with observations. However, the problem of the wide degree distribution of the nodes of the same age and the paradox of the first-mover advantage can be hardly solved in such a way. To address these problems one needs to go beyond the framework of the preferential attachment model [Eqs. (A1)–(A4)].

2. Fitness-based preferential attachment

The difficulties associated with the application of the preferential attachment model and its derivatives for the quantitative account of complex network growth call for alternative approaches. One alternative is the attachment probability which is proportional not to node's degree but to some other node's attribute such as local clustering coefficient [70], node's rank [71], or PageRank coefficient [72]. The most popular alternative is the Bianconi-Barabasi model [73] that introduced fitness, an empirical parameter that characterizes the propensity of nodes to attract edges. The core assumption of the model is that the node's fitness is a constant number and does not change with time.

a. Multiplicative fitness

How fitness can be incorporated into dynamic equation of network growth? The Bianconi-Barabasi model [73] introduces fitness on top of the preferential attachment, namely, it postulates that the attachment probability is the product of node's fitness η_j and degree,

$$\Pi_{ij} = \eta_j(K_j + K_0). \tag{A5}$$

[To be consistent with Eq. (1) we introduced here initial attractivity K_0.] Solution of Eq. (A5) yields the node's trajectory $K_j(t)$ which strongly depends on fitness: a high-fitness latecomer can outperform a low-fitness old node. Thus, fitness solves the problem of the first-mover advantage and the problem of degree distribution for the nodes of the same age which is now determined by fitness distribution rather than by the acquired degree.

The obvious way to measure the Bianconi-Barabasi's fitness is through Eq. (A5). Thus, Kong *et al.* [62] studied the network of WWW internet pages, analyzed trajectories of the pages of the same age, and successfully fitted them using Eq. (A5). The fitness turned out to be constant, as expected, and the fitness distribution turned out to be wide.

The most striking prediction of the Bianconi-Barabasi model is that for wide fitness distributions there are supercritical nodes that eventually take a lion share of edges. Such supercritical nodes were indeed observed [74,75], and this successful prediction brought a wide popularity to the Bianconi-Barabasi model [6,13,76–78].

While Eq. (A5) explains several features of complex networks, such as degree distribution of the nodes of the same age, it does not account for aging. To convert the Bianconi-Barabasi model into a quantitative tool that can be be compared to measured node's trajectories, Wang *et al.* [76] extended Eq. (A5) to

$$\Pi_{ij} = \eta_j A_j(t_j)(K_j + K_0), \tag{A6}$$

where $A_j(t_j)$ is the aging function, specific for each node, and t_j is the node's age. [Reference [76] denoted the aging function by $P_j(t)$ while we denote it by $A_j(t)$ to be consistent with Eq. (A1)]. The Wang *et al.* model [Eq. (A6)] builds upon the earlier approach of Ref. [79], which introduced the node's relevance, $X_j(t) \sim \eta_j A_j(t)$.

Equation (A6) was validated using citation network of physics papers covered by the APS database [76]. The aging function was approximated by the lognormal dependence $A_j(t_j) = \frac{1}{\sqrt{2\pi}\sigma_j t_j} e^{-(\frac{(\ln t_j - \mu_j)^2}{2\sigma_j^2})}$ where μ_j and σ_j are specific parameters for each node. Pham *et al.* [30,46] developed a software package based on Eq. (A6) and converted it into a practical and useful platform for quantitative description of the complex network growth.

This success notwithstanding, Eq. (A6) has several problems. First, there are too many parameters: in addition to dynamic node attributes (degree K_j and age t_j), Eq. (A6) adds three static attributes: fitness η_j and two parameters μ_j, and σ_j that characterize the aging function for each node. Second, the Wang-Song-Barabasi and its parent Bianconi-Barabasi model assume *linear* preferential attachment. This is an unlucky coincidence that both these models were validated using citation networks which exhibit *nonlinear* preferential attachment [24,51,53]. While the Wang *et al.* model can be extended to account for nonlinearity, this extension requires an additional fitting parameter, an attachment exponent, in such a way that the resulting model becomes too sophisticated.

b. Additive fitness

The multiplicative fitness of the Bianconi-Barabasi model is not the only way the fitness can be introduced into dynamic equation of network growth. Ref. [80] introduced fitness through optimization procedure, while Refs. [11,16,55,81] introduced fitness additively, as follows:

$$\Pi_{ij} \propto (K_j + \eta_j). \tag{A7}$$

Equation (A7) is nothing else but Eq. (1) where fitness η_j replaces the initial attractivity K_0.

The growth dynamics described by Eqs. (A5) and (A7) are not that different as it could seem. In fact, the combination of nonlinear preferential attachment [Eq. (A4)] with additive fitness [Eq. (A7)] mimics Eq. (A5), in particular, it yields supercritical nodes. To demonstrate this we adopt continuous approximation of Ref. [8] and replace ΔK_j in Eq. (A4) by

$\frac{dK_j}{dt}\Delta t$. In view of Eq. (A7), Eq. (A4) can be recast as follows:

$$\frac{dK}{dt} = \tilde{A}(t)(K+\eta)^{1+\delta}, \tag{A8}$$

where we replaced K_0 by η and dropped the index j for brevity. We solve Eq. (A8) for $\delta > 0$ and find

$$K(t) = \frac{\eta}{\left[1 - \delta\eta^{\delta}\int_0^t \tilde{A}(\tau)\,d\tau\right]^{\frac{1}{\delta}}} - \eta. \tag{A9}$$

To analyze Eq. (A9) we assume for simplicity that the integral $\int_0^t \tilde{A}(\tau)\,d\tau$ converges as $t \to \infty$. We introduce $\eta_{\text{crit}} = [\delta\int_0^\infty \tilde{A}(\tau)\,d\tau]^{-\frac{1}{\delta}}$, in such a way that Eq. (A9) reduces to

$$K(t) = \frac{\eta}{\left[1 - \left(\frac{\eta}{\eta_{\text{crit}}}\right)^{\delta}\frac{\int_0^t \tilde{A}(\tau)\,d\tau}{\int_0^\infty \tilde{A}(\tau)\,d\tau}\right]^{\frac{1}{\delta}}} - \eta, \tag{A10}$$

where t is the node's age. For $\eta < \eta_{\text{crit}}$ Eq. (A10) yields $K(t)$ that increases with time and eventually achieves saturation, $K(\infty) = \frac{\eta}{[1-(\frac{\eta}{\eta_{\text{crit}}})^{\delta}]^{\frac{1}{\delta}}} - \eta$. However, for $\eta \geqslant \eta_{\text{crit}}$, $K(t)$ does not achieve saturation, namely, the node's trajectory becomes supercritical. [In fact, it undergoes a finite-time singularity at certain t_0, in such a way that Eqs. (A9) and (A10) hold only for $t < t_0$.] Thus, for the superlinear preferential attachment, $\delta > 0$, Eq. (A7) predicts the supercritical nodes, exactly as Eq. (A5) does.

3. Fitness-only models

The fitness model suggested by Caladarelli *et al.* [10] and further developed in Refs. [11–15] assumes that the probability of attachment between a new node i and the target node j depends only on their fitnesses, η_i and η_j, and does not depend on their degrees. Reference [10] assumed that $\Pi_{ij} = f(\eta_i,\eta_j)$ where η_i and η_j are node fitnesses, and $f(\eta_i,\eta_j)$ is the symmetric function of its arguments (linking function). Reference [10] considered additive linking function but the later publication of the same group [82] introduced multiplicative linking function, $f(\eta_i,\eta_j) \sim \eta_i\eta_j$. The latter assumption became more popular and it allows the following generalization. Consider a target node j. If fitness is determined by similarity and all nodes belong to the same community, then the node j will garner edges with the rate $\Delta K_j \propto \overline{\eta_i}\eta_j$ where $\overline{\eta_i}$ is the average fitness of incoming nodes. This average fitness can be absorbed into the aging function, in such a way that the probability of a new node i to attach to existing node j is

$$\Pi_{ij} \sim \eta_j A(t_j). \tag{A11}$$

What is fitness? It should be noted that fitness in the sense of Caldarelli *et al.* [10] [see Eq. (A11)] is very different from the Bianconi-Barabasi's fitness [73] since the latter has been defined in the context of preferential attachment through Eq. (A5). On the one hand, the fitness includes the notion of similarity known as homophily in social networks [21]. Indeed, complex networks are rarely uniform, they consist of communities and subcommunities. New nodes tend to attach to similar nodes, namely, to those belonging to the same community. To measure similarity one can use overlap of contents or bibliographies for citation networks and WWW pages [16,17] or overlap of common neighbors in the general case [18–20]. Another ingredient of fitness is associated with quality or talent. This component is not easy to estimate when the node first appears, it can be measured only after it has garnered some edges.

4. Explanatory models

The preferential attachment mechanism of network growth presents a major conceptual difficulty because it is global and not local. Indeed, Eqs. (A1) and (A4) imply that each incoming node shall know degrees of all other nodes. Although this can be true for collaboration and some other social networks [83,84], in general, a new node has bounded knowledge—it is familiar only with a limited set of nodes [85].

While we showed that the fitness model explains the initial attractivity, it can hardly serve as an explanatory platform of network growth since it is too phenomenological and devoid of specific details characterizing real networks.

The most popular explanatory mechanism of network growth is the recursive search [86] also known as link copying or redirection [87], random walk or local search [88,89], triple(triangle) formation [90], triadic closure [91], or forest fire model [92,93]. This mechanism assumes that a new node attaches to a randomly found node, explorers the network neighborhood of the latter, and with some probability attaches to one [88] or all [87,94] of its ancestors. This is a one-level recursive search while Refs. [92,93,95] considered a multilevel recursive search, whereby a new node explores network vicinity of all previously chosen nodes. Vazquez [86] showed that one-level recursive search mechanism reduces to the following probability of a new node i to attach to a target node j,

$$\Pi_{ij} = \lambda + qK_j. \tag{A12}$$

Here λ is the probability of random search, qK_j is the probability of recursive search, and K_j is the target node's degree. Thus, Eq. (A12) reduces to Eq. (1). Note, however, that the derivation of Eq. (A12) by Ref. [86] was based on very simplifying assumptions: all nodes have the same probability of being randomly chosen, only one ancestor of the randomly-found node is chosen by incoming node, no aging, no memory, etc. These assumptions are too restrictive. In Sec. III we demonstrate a more realistic recursive search model which emerged from our recent studies of citation networks [23]. This model embeds fitness-based search into recursive search with memory.

APPENDIX B: INITIAL ATTRACTIVITY ANALYSIS

To explore the limits of approximation $K \gg 1$ for which

$$n(K+1,t) \approx n(K,t), \tag{B1}$$

we consider the ratio $\frac{n(K+1,t)}{n(K,t)} = \frac{\int_0^\infty \frac{\Lambda^{K+1}}{(K+1)!}e^{-\Lambda}\rho(\Lambda)\,d\Lambda}{\int_0^\infty \frac{\Lambda^K}{K!}e^{-\Lambda}\rho(\Lambda)\,d\Lambda}$ and note that for the uniform distribution, $\rho(\Lambda) = \text{const}$, Eq. (B1) is satisfied exactly for every K. In order to study to what extent this relation holds for nonuniform distributions, we assume a log-normal distribution, $\rho(\Lambda) = \frac{1}{\sqrt{2\pi}\sigma\Lambda}e^{-\frac{(\ln\Lambda-\mu)^2}{2\sigma^2}}$. Then $\frac{n(K+1,t)}{n(K,t)} =$

$\frac{\int_0^\infty \frac{\Lambda^{K+1}}{(K+1)!} e^{-\Lambda} e^{-\frac{(\ln\Lambda-\mu)^2}{2\sigma^2}} d\Lambda}{\int_0^\infty \frac{\Lambda^{K}}{(K)!} e^{-\Lambda} e^{-\frac{(\ln\Lambda-\mu)^2}{2\sigma^2}} d\Lambda}$. The expression $\Lambda^K e^{-\Lambda}$, when considered as function of Λ, is a bell-shaped function with a peak at $\Lambda_{\max} = K$. The width of this peak is $\Delta\Lambda \sim K^{\frac{1}{2}}$, and the relative width is $\frac{\Delta\Lambda}{\Lambda_{\max}} \approx \frac{1}{K^{\frac{1}{2}}}$. For $K \gg 1$ this peak is much narrower than any lognormal distribution with $\sigma > 1$. In this case, the lognormal function is almost constant across the peak of the function $\Lambda^K e^{-\Lambda}$ and we can replace it by its value at the peak, $\Lambda = \Lambda_{\max}$. Since $\int_0^\infty \frac{\Lambda^K}{(K)!} e^{-\Lambda} = 1$ then, for $K \gg 1$, $\frac{n(K+1,t)}{n(K,t)} \approx e^{-\frac{(\ln(K+1)-\mu)^2-(\ln K-\mu)^2}{2\sigma^2}} = e^{-\ln(1+\frac{1}{K})\frac{\ln K-\mu}{\sigma^2}} \approx e^{-\frac{\ln K-\mu}{K\sigma^2}}$. Since $\frac{\ln K}{K} \ll 1$ for $K \gg 1$, then $\frac{n(K+1,t)}{n(K,t)} \approx 1$. Thus, the latter relation holds for $K \gg 1$ and $\sigma > 1$.

[1] M. Mitzenmacher, Internet Math. **1**, 226 (2004).
[2] M. Newman, Contemp. Phys. **46**, 323 (2005).
[3] A. Clauset, C. R. Shalizi, and M. E. J. Newman, SIAM Rev. **51**, 661 (2009).
[4] D. D. S. Price, J. Am. Soc. Inf. Sci. **27**, 292 (1976).
[5] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, Phys. Rep. **424**, 175 (2006).
[6] G. Caldarelli, *Scale-Free Networks: Complex Webs in Nature and Technology* (Oxford University Press, New York, 2007).
[7] M. Newman, *Networks* (Oxford University Press, New York, 2010).
[8] A.-L. Barabasi, *Network Science* (Cambridge University Press, Cambridge, 2015).
[9] R. Albert and A.-L. Barabasi, Rev. Mod. Phys. **74**, 47 (2002).
[10] G. Caldarelli, A. Capocci, P. DeLosRios, and M. A. Muñoz, Phys. Rev. Lett. **89**, 258702 (2002).
[11] C. Bedogne' and G. J. Rodgers, Phys. Rev. E **74**, 046115 (2006).
[12] M. V. Simkin and V. P. Roychowdhury, J. Am. Soc. Inf. Sci. Technol. **58**, 1661 (2007).
[13] S. Ghadge, T. Killingback, B. Sundaram, and D. A. Tran, Intl. J. Parallel Emergent Distributed Syst. **25**, 223 (2010).
[14] K. Nguyen and D. A. Tran, in *Handbook of Optimization in Complex Networks*, edited by M. T. Thai and P. M. Pardalos (Springer, New York, 2012), pp. 39–53.
[15] J. M. Luck and A. Mehta, Phys. Rev. E **95**, 062306 (2017).
[16] F. Menczer, Proc. Natl. Acad. Sci. U. S. A. **101**, 5261 (2004).
[17] V. Ciotti, M. Bonaventura, V. Nicosia, P. Panzarasa, and V. Latora, EPJ Data Sci. **5**, 7 (2016).
[18] M. E. J. Newman, Phys. Rev. E **64**, 025102(R) (2001).
[19] D. Liben-Nowell and J. Kleinberg, in *Proceedings of the 12th International Conference on Information and Knowledge Management—CIKM* (Association for Computing Machinery, New York, 2003), pp. 556–559.
[20] A. Mislove, H. S. Koppula, K. P. Gummadi, P. Druschel, and B. Bhattacharjee, in *Dynamics on and of Complex Networks*, edited by A. Mukherjee *et al.* (Birkhäuser, New York, 2013), Vol. 2, pp. 19–40.
[21] Y. Bramoullé, S. Currarini, M. O. Jackson, P. Pin, and B. W. Rogers, J. Econ. Theory **147**, 1754 (2012).
[22] Q. L. Burrell, J. Am. Soc. Inf. Sci. **54**, 372 (2003).
[23] M. Golosovsky and S. Solomon, Phys. Rev. E **95**, 012324 (2017).
[24] M. Golosovsky and S. Solomon, Phys. Rev. Lett. **109**, 098701 (2012).
[25] D. M. Pennock, G. W. Flake, S. Lawrence, E. J. Glover, and C. L. Giles, Proc. Natl. Acad. Sci. U. S. A. **99**, 5207 (2002).
[26] Z.-G. Shao, X.-W. Zou, Z.-J. Tan, and Z.-Z. Jin, J. Phys. A **39**, 2035 (2006).
[27] G. J. Peterson, S. Presse, and K. A. Dill, Proc. Natl. Acad. Sci. U. S. A. **107**, 16023 (2010).
[28] M. Wang, G. Yu, and D. Yu, Physica A **387**, 4692 (2008).
[29] J. P. Gleeson, D. Cellai, J.-P. Onnela, M. A. Porter, and F. Reed-Tsochas, Proc. Natl. Acad. Sci. U. S. A. **111**, 10411 (2014).
[30] T. Pham, P. Sheridan, and H. Shimodaira, Sci. Rep. **6**, 32558 (2016).
[31] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, Phys. Rev. Lett. **85**, 4633 (2000).
[32] G. Ghoshal, L. Chi, and A.-L. Barabási, Sci. Rep. **3**, 2920 (2013).
[33] J. Leskovec, J. Kleinberg, and C. Faloutsos, ACM Trans. Knowledge Disc. Data (TKDD) **1**, 2 (2007).
[34] S. N. Dorogovtsev and J. F. F. Mendes, Phys. Rev. E **62**, 1842 (2000).
[35] K. B. Hajra and P. Sen, Physica A **368**, 575 (2006).
[36] Y. Wu, T. Z. Fu, and D. M. Chiu, J. Informetrics **8**, 650 (2014).
[37] L. Ostroumova Prokhorenkova and E. Samosvat, J. Complex Netw. **4**, 475 (2016).
[38] B. A. Miller and N. T. Bliss, IEEE Signal Process. Lett. **19**, 356 (2012).
[39] M. Rosvall, A. V. Esquivel, A. Lancichinetti, J. D. West, and R. Lambiotte, Nat. Commun. **5**, 4630 (2014).
[40] O. Mokryn, A. Wagner, M. Blattner, E. Ruppin, and Y. Shavitt, PLoS ONE **11**, e0156505 (2016).
[41] P. L. Krapivsky and S. Redner, Phys. Rev. E **63**, 066123 (2001).
[42] P. Krapivsky and D. Krioukov, Phys. Rev. E **78**, 026114 (2008).
[43] J. Leskovec, L. Backstrom, R. Kumar, and A. Tomkins, in *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining—KDD '08* (Association for Computing Machinery, New York, 2008), pp. 462–470.
[44] J. Kunegis, M. Blattner, and C. Moser, in *Proceedings of the 5th Annual ACM Web Science Conference, WebSci '13* (Association for Computing Machinery, New York, 2013), pp. 205–214.
[45] M. Perc, J. R. Soc. Interface **11**, 20140378 (2014).
[46] T. Pham, P. Sheridan, and H. Shimodaira, PLoS ONE **10**, e0137796 (2015).
[47] S. Redner, Phys. Today **58**, 49 (2005).
[48] A. Garavaglia, R. Hofstad, and G. Woeginger, J. Stat. Phys. **168**, 1137 (2017).
[49] A. Capocci, V. D. P. Servedio, F. Colaiori, L. S. Buriol, D. Donato, S. Leonardi, and G. Caldarelli, Phys. Rev. E **74**, 036116 (2006).
[50] C. P. Massen and J. P. Doye, Physica A **377**, 351 (2007).
[51] M. Golosovsky and S. Solomon, J. Stat. Phys. **151**, 340 (2013).
[52] K. W. Higham, M. Governale, A. B. Jaffe, and U. Zulicke, Phys. Rev. E **95**, 042309 (2017).
[53] K. Higham, M. Governale, A. Jaffe, and U. Zülicke, J. Informetrics **11**, 1190 (2017).

[54] H. Jeong, Z. Néda, and A.-L. Barabási, EPL (Europhys. Lett.) **61**, 567 (2003).

[55] Y.-H. Eom and S. Fortunato, PLoS ONE **6**, e24926 (2011).

[56] M. Medo, Phys. Rev. E **89**, 032801 (2014).

[57] Reference [44] analyzed 49 networks, claimed preferential attachment for each of them, and reported the corresponding parameters found from the sophisticated analysis of ΔK versus K plots. Reference [44] also showed raw data for six networks, and five of them clearly do not support the preferential attachment mechanism. These networks demonstrate preferential attachment at low degree and anti-preferential attachment for high degree, similar to what was observed in Ref. [49] for Wikipedia edits. Thus, although the fitting procedure of Ref. [44] can be valid, the interpretation of the results without inspection of raw data is incomplete.

[58] G. Csárdi, K. J. Strandburg, L. Zalányi, J. Tobochnik, and P. Érdi, Physica A **374**, 783 (2007).

[59] A. Zeng, Z. Shen, J. Zhou, J. Wu, Y. Fan, Y. Wang, and H. E. Stanley, Phys. Rep. **714**, 1 (2017).

[60] M. E. J. Newman, EPL (Europhys. Lett.) **86**, 68001 (2009).

[61] M. E. J. Newman, EPL (Europhys. Lett.) **105**, 28002 (2014).

[62] J. S. Kong, N. Sarshar, and V. P. Roychowdhury, Proc. Natl. Acad. Sci. U. S. A. **105**, 13724 (2008).

[63] Q. Ke, E. Ferrara, F. Radicchi, and A. Flammini, Proc. Natl. Acad. Sci. U. S. A. **112**, 7426 (2015).

[64] B. A. Huberman and L. A. Adamic, arXiv:cond-mat/9901071.

[65] L. A. Adamic, B. A. Huberman, A.-L. Barabási, R. Albert, H. Jeong, and G. Bianconi, Science **287**, 2115 (2000).

[66] X. Geng and Y. Wang, EPL (Europhys. Lett.) **88**, 38002 (2009).

[67] Z. Xie, Z. Ouyang, Q. Liu, and J. Li, Physica A **456**, 167 (2016).

[68] I. Sendiña-Nadal, M. M. Danziger, Z. Wang, S. Havlin, and S. Boccaletti, Sci. Rep. **6**, 21297 (2016).

[69] L. O. Prokhorenkova, Optimization Lett. **11**, 279 (2017).

[70] J. P. Bagrow and D. Brockmann, Phys. Rev. X **3**, 021016 (2013).

[71] S. Fortunato, A. Flammini, and F. Menczer, Phys. Rev. Lett. **96**, 218701 (2006).

[72] J. Zhou, A. Zeng, Y. Fan, and Z. Di, Scientometrics **106**, 805 (2016).

[73] G. Bianconi and A.-L. Barabasi, Phys. Rev. Lett. **86**, 5632 (2001).

[74] A.-L. Barabasi, C. Song, and D. Wang, Nature (London) **491**, 40 (2012).

[75] M. Golosovsky, Phys. Rev. E **96**, 032306 (2017).

[76] D. Wang, C. Song, and A.-L. Barabasi, Science **342**, 127 (2013).

[77] T. Carletti, F. Gargiulo, and R. Lambiotte, Eur. Phys. J. B **88**, 18 (2015).

[78] M. Bell, S. Perera, M. Piraveenan, M. Bliemer, T. Latty, and C. Reid, Sci. Rep. **7**, 42431 (2017).

[79] M. Medo, G. Cimini, and S. Gualdi, Phys. Rev. Lett. **107**, 238701 (2011).

[80] F. Papadopoulos, M. Kitsak, M. Á. Serrano, M. Boguñá, and D. Krioukov, Nature (London) **489**, 537 (2012).

[81] G. Ergün and G. Rodgers, Physica A **303**, 261 (2002).

[82] V. D. P. Servedio, G. Caldarelli, and P. Buttà, Phys. Rev. E **70**, 056126 (2004).

[83] D. Centola, V. M. Eguíluz, and M. W. Macy, Physica A **374**, 449 (2007).

[84] D. Centola, Science **329**, 1194 (2010).

[85] With the proliferation of informational databases such as Google Scholar, Scopus, ISI Web of Science, etc., global information on many complex networks became easily accessible. In particular, for citation networks, the incentive to cite a certain paper may indeed come from the number of previous citations. Thus, the preferential attachment model is becoming a self-fulfilling prophecy.

[86] A. Vazquez, EPL (Europhys. Lett.) **54**, 430 (2001).

[87] P. L. Krapivsky and S. Redner, Phys. Rev. E **71**, 036118 (2005).

[88] M. O. Jackson and B. W. Rogers, Am. Econ. Rev. **97**, 890 (2007).

[89] S. R. Goldberg, H. Anthony, and T. S. Evans, Scientometrics **105**, 1577 (2015).

[90] Z.-X. Wu and P. Holme, Phys. Rev. E **80**, 037101 (2009).

[91] T. Martin, B. Ball, B. Karrer, and M. E. J. Newman, Phys. Rev. E **88**, 012814 (2013).

[92] J. Leskovec, J. Kleinberg, and C. Faloutsos, in *Proceedings of the ACM Transactions on Knowledge Discovery from Data (TKDD), 2007* (Association for Computing Machinery, New York, 2005).

[93] L. Šubelj and M. Bajec, in *Proceedings of the 22nd International Conference on World Wide Web, WWW '13 Companion, 2013* (Association for Computing Machinery, New York, 2013), pp. 527–530.

[94] R. Lambiotte, P. L. Krapivsky, U. Bhat, and S. Redner, Phys. Rev. Lett. **117**, 218301 (2016).

[95] R. Lambiotte and M. Ausloos, EPL (Europhys. Lett.) **77**, 58002 (2007).