Fitness-based network growth with dynamic feedback.

I. E. Smolyarenko
Department of Mathematical Sciences, Brunel University,
Uxbridge, Middlesex UB8 3PH, United Kingdom
(Dated: December 30, 2013.)

We study a class of network growth models in which the choice of attachment by new nodes is governed by intrinsic attractiveness, or *fitness*, of the existing nodes. The key feature of the models is a feedback mechanism whereby the distribution from which fitnesses of new nodes are drawn derives from the evolving instantaneous node degree distribution. In the case of linear mapping between fitnesses and degrees, the fixed point degree distribution is asymptotically power-law, while in the non-linear case the distributions converge to the stretched exponential form.

Keywords: Complex Networks, Fitness

I. INTRODUCTION.

Models of dynamically evolving complex networks have proved to be powerful tools for describing arrays of interacting agents in various studies of natural and societal phenomena [1]. Network growth models can be broadly separated into two classes. Models belonging to the first class can be generally characterized as having the growth rules governed by the (dynamically evolving) current network topology. The paradigmatic example of this type of model is the preferential attachment mechanism [1–3] where a new node finds a *parent* to which it attaches depending on the parent's number of connections (*degree*) at the time of attachment. Models of this type are well-known to generate the topological characteristics, such as power-law degree distributions, that are frequently observed in empirical network systems.

A significant requirement of the preferential attachment rules is that every new node joining the network must possess complete and updated information about the degrees of every existing node in the network. In a practical setting, such information may not always be readily available. For example, when concluding a business deal or establishing a partnership, information about the overall reputation of a company may be more accessible than the number of their current suppliers and clients. Similarly, research collaborations are typically established on the basis of prospective collaborators' expertise and reputations rather than simply the total number of past (or current, depending on how a link is defined) collaborations.

Such considerations provided one of the motivations for the study of a different class of growth mechanisms, variously known as hidden variable or *fitness*-based models [4–7]. The models of this type are characterized by probabilistic rules for forming connections between nodes based on a static measure of intrinsic node attractiveness, usually termed fitness. Both the distribution of fitnesses and the connection rules are given by *a priori* arbitrary functions, thus allowing a considerable amount of tuning in such models. This feature enables fitness-based models to mimic a variety of network topologies, in particular, subject to some constraints, they can be tuned to reproduce a given type of degree distributions and even degree correlation functions [4, 5, 7]. This tunability makes fitness-based models useful as a modeling tool, but also imparts a degree of arbitrariness which makes them less attractive as a robust explanation for the universality of naturally observed behaviors. Neither of the above two classes of models is likely to be observed empirically in a pure form. However, many realistic models would contain elements of both mechanisms: The degree of a node is indeed a realistic measure of attractiveness, but the relation of the actual proxy used to guide new connections to instantaneous node degrees may be indirect.

As shown in [7], distributions of node degrees in fitness-based network growth models are generically broader than the "input" distributions of fitnesses. For example, if all nodes have equal fitnesses (a delta-peaked distribution) the resulting degree distribution is exponential, while an exponential distribution of fitnesses leads to stretched exponential distribution of degrees. Crucially, however, this broadening saturates at power-law distributions, so that a fitness distribution that asymptotically behaves as a power law generates a degree distribution with a matching asymptotic power law tail [6, 7].

A natural question therefore arises whether power-law behavior can be generated in fitness-based models as a fixed point. Quite generally, ubiquitous power-law degree distributions of observed networks are likely to be enforced by fixed point behaviors insensitive to microscopic details of diverse network growth mechanisms, and it is therefore of interest to identify the possible fixed point scenarios. This consideration motivates the models considered in this paper: The central goal in constructing the models is to retain the concept of node fitness as separate from the node degree, while allowing for feedback from the dynamically evolving network topology to the fitness distribution.

II. THE LINEAR MODEL.

A. The tree model.

The growth models are formulated in discrete time: at every integer time step a new node joins the network, attaching to a parent node chosen according to the probabilistic rule specified below. We first consider the simplest version of the model which results in a tree network: Each new node has a single parent node to which it attaches at the time of joining the network, and there are no rewiring mechanisms. Therefore the total number of nodes at time t is deterministic: $N_t = t + N_0$, where $N_0 \sim O(1)$ is the number of nodes in the seed network. In the following, all consideration is restricted to the asymptotic $t \gg 1$ regime, and sub-leading N_0-dependent terms are consistently dropped.

Denoting the fitness of the existing i-th node as x_i, the probability that the new node joining at time $t + 1$ chooses node i as its parent is linear in fitness:

$$\pi_t(i) = \frac{x_i}{\sum_{j=1}^{t} x_j}. \tag{1}$$

In the asymptotic long-time limit the normalizing denominator simplifies: $\sum_j^t x_j \to t\bar{x}_t$, where $\bar{x}_t$ is the expectation of node fitness at time t.

The key element of the model which implements the feedback feature is the following modeling assumption: the fitness of each new node x_{t+1} is drawn at random from the instantaneous degree distribution $p_t(k)$ at time t. The core theoretical motivation for studying this feedback mechanism is the fact that the "power law generates power law" results in Refs. [6, 7] provide a strong hint that this type of feedback may generate a power-law fixed point. As will be shown below, this conjecture is indeed confirmed by an explicit calculation, with the fixed point characterized by a unique power-law exponent. On the empirical side, fitness of a new node can be thought of as a *credential*. The proposed model corresponds to the situation where the credential of a node is fixed at the time of entry into the network, as it may happen in, *e.g.*, some professional networks where the reputation of a mentor or an *alma mater* may be determinative of a new entrant's credentials.

Analysis of the model is simplified by the fact that $\bar{x}_t = \bar{k}$, where $\bar{k}$ is the average node degree. Since the growth mechanism generates a tree, up to $O(1)$ corrections inherited from the seed value N_0, the total number of edge endings is twice the number of nodes, and hence $\bar{k} = 2$ independent of t.

Let $p_t(k|x,\tau)$ be the probability that a node which joined the network at time τ with fitness x has degree k at time $t > \tau$. It follows from Eq. (1) that $p_t(k|x,\tau)$ satisfies the following rate equation in the $t \gg 1$ regime:

$$p_{t+1}(k|x,\tau) = p_t(k-1|x,\tau)\frac{x}{2t}[1-\delta_{k,1}] + p_t(k|x,\tau)\left(1 - \frac{x}{2t}\right), \tag{2}$$

where $\delta_{i,j}$ is the Kronecker symbol. In the continuous time limit, the corresponding generating function $G(t,s|x,\tau) = \sum_{k=1}^{\infty} p_t(k|x,\tau)s^k$ satisfies

$$\frac{\partial}{\partial \ln t} G(t,s|x,\tau) = \frac{x}{2}(s-1)G(t,s|x,\tau). \tag{3}$$

The boundary condition is determined by the fact that a newly-joined node is connected only to its parent, and therefore has degree equal to 1: $\lim_{t\to\tau+} p_t(k|x,\tau) = \delta_{k,1}$, hence $G(\tau+, s|x,\tau) = s$. The equation on $G(t,s|x,\tau)$ is solved by

$$G(t,s|x,\tau) = s\left(\frac{\tau}{t}\right)^{x(1-s)/2}. \tag{4}$$

The global generating function $G(t,s) = \sum_{k=1}^{\infty} p_t(k)s^k$ is obtained by averaging $G(t,s|x,\tau)$ over the fitness x and the time of joining τ. The feedback mechanism in the model dictates that x is distributed according to p_τ, therefore

$$G(t,s) = \int_0^t \frac{d\tau}{t} \sum_{x=1}^{\infty} p_\tau(x) s \left(\frac{\tau}{t}\right)^{x(1-s)/2} = s\int_0^1 dz G\left(zt, z^{\frac{1-s}{2}}\right), \tag{5}$$

where the second equality is obtained by changing the integration variable to $z = \tau/t$, and substituting the definition of the global generating function.

The generating function $G(s) = \sum_{k=1}^{\infty} p(k)s^k$ of the corresponding stationary distribution $p(k)$ therefore satisfies the following integral equation:

$$G(s) = s\int_0^1 dz G\left(z^{\frac{1-s}{2}}\right). \tag{6}$$

It is convenient to transform this equation back to the distribution $p(k)$ itself:

$$p(k) = \sum_{n=1}^{\infty} \frac{2p(n)}{n}\left(\frac{n}{n+2}\right)^k. \tag{7}$$

The matrix operator on the right is positive, and it is easy to see that it possesses an eigenvalue equal to 1 corresponding to a positive eigenvector:

$$\sum_{k=1}^{\infty} \frac{2}{n}\left(\frac{n}{n+2}\right)^k = 1. \tag{8}$$

It is also easy to check that this equation automatically satisfies the $\bar{k} = 2$ property, which is enforced by the growth rules:

$$\bar{k} = \sum_{k=1}^{\infty} kp(k) = \sum_{k=1}^{\infty}\sum_{n=1}^{\infty} \frac{2p(n)}{n} k \left(\frac{n}{n+2}\right)^k = \sum_{n=1}^{\infty} \frac{2p(n)}{n+2} \frac{1}{\left(1 - \frac{n}{n+2}\right)^2} = \sum_{n=1}^{\infty} \frac{1}{2}(n+2)p(n) = 1 + \bar{k}/2, \quad (9)$$

hence $\bar{k} = 2$.

The calculation above can be generalized to calculate the second and the third moments of $p(k)$. For example, expressing the second moment of p_k via the l.h.s. of Eq. (7) gives

$$\overline{k^2} = \sum_{k=1}^{\infty} k^2 p(k) = \sum_{k=1}^{\infty}\sum_{n=1}^{\infty} \frac{2p(n)}{n} k^2 \left(\frac{n}{n+2}\right)^k = \sum_{n=1}^{\infty} \frac{1}{2} p(n)(n+1)(n+2), \quad (10)$$

hence

$$\overline{k^2} = \frac{1}{2}\overline{k^2} + \frac{3}{2}\bar{k} + 1, \quad (11)$$

so that $\overline{k^2} = 8$. Similarly,

$$\overline{k^3} = \sum_{n=1}^{\infty} \frac{1}{4} p(n)(n+2)(3n^2+6n+2) = \frac{3}{4}\overline{k^3} + 3\overline{k^2} + \frac{7}{2}\bar{k} + 1, \quad (12)$$

thus $\overline{k^3} = 128$.

The technique, however, runs into a seeming contradiction if an attempt is made to calculate the 4th moment: the coefficient of $\overline{k^4}$ in the r.h.s. is greater than 1, seemingly implying that the moment is negative. The latter is impossible, however, since Krein-Rutman theorem together with Eq. (8) ensures that the dominant eigenvector of $K_{nk} = \frac{2}{n}\left(\frac{n}{n+2}\right)^k$ corresponding to the eigenvalue 1 is positive. It follows that the fourth moment of the degree distribution does not exist, thus confirming the asymptotic power-law nature of the distribution. The precise exponent of the power-law decay of $p(k)$ can be obtained by asymptotic matching of the coefficient of the fractional moment. Consider the fractional moment of the distribution defined as

$$\mu_\beta = \sum_{k=1}^{\infty} k^\beta p(k). \quad (13)$$

Substituting this definition into Eq. (7), we find

$$\mu_\beta = \sum_{k=1}^{\infty} k^\beta \sum_{n=1}^{\infty} \frac{2p(n)}{n} \left(\frac{n}{n+2}\right)^k = \sum_{n=1}^{\infty} \frac{2p(n)}{n} \mathrm{Li}_{-\beta}\left(\frac{n}{n+2}\right). \quad (14)$$

Let us assume that β is close to the critical value α at which the moment becomes divergent. In this regime the sum is dominated by large values of n, and their contribution is obtained by using the leading asymptotics of the polylogarithm $\mathrm{Li}_{-\beta}(z) \sim \Gamma(1+\beta)(-\ln z)^{-\beta-1} + O(1)$ [8]. Equation (14) takes the form

$$\mu_\beta = \sum_{n=1}^{\infty} \frac{\Gamma(1+\beta)}{2^\beta} n^\beta p(n) + r_\beta = \frac{\Gamma(1+\beta)}{2^\beta}\mu_\beta + r_\beta, \quad (15)$$

where r_β is $\sim O(1)$. For any value β such that $\frac{\Gamma(1+\beta)}{2^\beta} < 1$, the above expression relates μ_β and r_β: $\mu_\beta = r_\beta/(1 - \Gamma(1+\beta)/2^\beta)$. Since μ_β is positive by definition for all values β such that μ_β exists, this implies that $r_\beta > 0$ for $\beta < \alpha$, where α is defined by

$$1 = \Gamma(1+\alpha)/2^\alpha, \quad (16)$$

or $\alpha \approx 3.45987$. On the other hand, r_β is given by a convergent series, hence it is regular at $\beta = \alpha$, and therefore μ_β exhibits a $1/(\alpha - \beta)$ divergence as $\beta \to \alpha$, and does not exist for $\beta \geq \alpha$. It follows that $p(k)$ asymptotically behaves as $k^{-\alpha-1}$. The asymptotic power-law behavior of the degree distribution confirms the core conjecture stated earlier: if the fitness distribution of newly joined nodes dynamically tracks the degree distribution, the latter converges to a fixed point characterized by a power-law decay with a unique value of the exponent. The $1/(\alpha-\beta)$ decay is also consistent with absence of any logarithmic corrections to $k^{-\alpha-1}$, since it matches the first-order pole of the Riemann zeta-function $\zeta(s)$ at $s\to 1$. This will also be demonstrated below using an explicit calculation.

The full shape of the distribution is reasonably well approximated by the discrete analog of the power-law function [9],

$$p_{\rm fit}(k) \approx C_\alpha \Gamma(k+a)/\Gamma(k+a+\alpha+1), \tag{17}$$

where $C_\alpha = \alpha\Gamma(a+\alpha+1)/\Gamma(a+1)$, and numerical fitting gives $a \approx 1.0731$. Fig. 1 shows the degree distribution obtained after 4×10^9 time steps of direct simulation of this model, together with the fitting function Eq. (17), and its the asymptotic power law tail $C_\alpha k^{-\alpha-1}$.

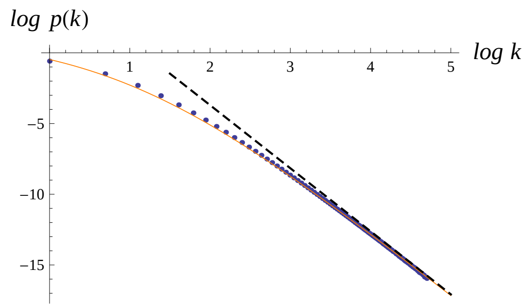

FIG. 1: Log-log plots of the degree distribution function: simulation (dots), the fitting function $p_{fit}(k)$ (solid line), and the asymptotic power law (dashed line) in the linear model.

Although Eq. (6) does not appear to possess a closed form analytic solution, the kernel in Eq. (6) is a compact operator with spectral radius equal to 1, hence the solution can be obtained via a convergent sequence of iterations. Changing variables in Eq. (6) from z to $y = z^{(1-s)/2}$, we obtain the following iterative relation:

$$G_{n+1}(s) = \frac{2s}{1-s}\int_0^1 y^{\frac{1+s}{1-s}} G_n(y)dy. \tag{18}$$

Choosing $G_0(s) = (s\alpha/(1+\alpha+a))\,{}_2F_1(1, 1+a; 2+a+\alpha; s)$ (corresponding to the fitted distribution (17)) as the zeroth order approximation, the first iteration can be performed analytically, giving

$$G_1(s) = \frac{2s\alpha}{(3-s)(1+a+\alpha)}\,{}_3F_2\left(1, 1+a, \frac{3-s}{1-s}; 2+a+\alpha, 1+\frac{3-s}{1-s}; 1\right). \tag{19}$$

The results of subsequent iterations do not have closed form expressions in terms of standard special functions, and have to be performed numerically. Figure 2 shows $G_0(s)$ and the (numerical approximation to) the fixed point solution $G_\infty(s)$.

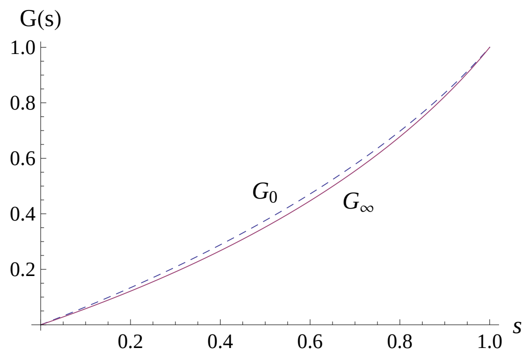

FIG. 2: The zeroth order (dashed), and the fixed-point (solid line) generating functions in the linear model.

In order to restore the coefficients $p(k)$ from $G_\infty(s)$, Eq. (18) is interpreted as an integral representation of $G(s) = G_\infty(s)$, which therefore allows to perform its analytical continuation from $[0, 1]$ to the unit circle. The degree

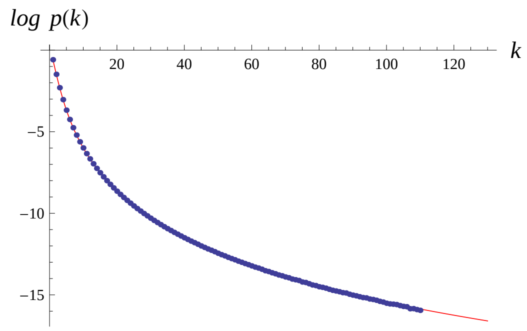

FIG. 3: The logarithm of the node degree distribution in the linear model: analytical results complemented by numerical evaluation of the fixed point generating function (solid line) *vs.* direct numerical simulation of the network growth process.

distribution $p(k)$ now straightforwardly follows from the application of the residue theorem. Figure 3 shows the degree distribution obtained using this method, plotted together with the outcome of the direct simulation of the model.

We now demonstrate explicitly that the power-law asymptotics of $p(k)$ is not augmented by any logarithmic corrections. Figure 4 shows the the plot of $p(k)k^{\alpha+1}/C_\alpha$. It is worth remarking that, although the ratio $p(k)k^{\alpha+1}/C_\alpha$ is seen numerically to approach saturation, the approach is sufficiently slow that it is not feasible to unambiguously demonstrate the absence of logarithmic corrections using numerical results. However, this can be achieved analytically by analyzing the asymptotic behavior of truncated divergent moments of $p(k)$.

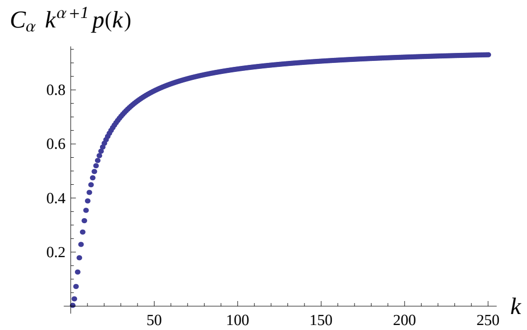

FIG. 4: Node degree distribution in the linear model scaled by the inverse asymptotic power-law ansatz.

Consider truncated fractional moments of degree $\beta > \alpha$:

$$M_\beta(K) = \sum_{k=1}^{K} k^\alpha p(k). \tag{20}$$

Substituting this into the self-consistency condition on $p(n)$ given by Eq. (7), we obtain

$$M_\beta(K) = \sum_{k=1}^{K}\sum_{n=1}^{\infty} k^\beta \left(\frac{n}{n+2}\right)^k \frac{2p(n)}{n} = \sum_{n=1}^{\infty} \frac{2p(n)}{n}\left[\mathrm{Li}_{-\beta}\left(\frac{n}{n+2}\right) - \left(\frac{n}{n+2}\right)^{K+1} \Phi\left(\frac{n}{n+2}, -\beta, K+1\right)\right], \tag{21}$$

where Φ is the Lerch zeta-function. In the asymptotic $K \to \infty$ regime, the behavior of this truncated moment is controlled by the large-k asymptotics of $p(k)$. More precisely, for a fixed β, we require $(\beta-\alpha)K \gg 1$. We now substitute for $p(k)$ an asymptotic ansatz $p(k) \sim \widetilde{C} k^{-\alpha-1} f(\ln k)$, where $\widetilde{C}$ is a proportionality constant, and $f(x)$ is slower than exponential so that $f(\ln k)$ is slower than any power of k. Consistently utilizing the large-K approximation, we find the dominant behavior of $M_\beta(k)$ to be given by

$$M_\beta(K) \sim \widetilde{C} \int^{K} n^{\beta-\alpha-1} f(\ln n) dn \sim K^{\beta-\alpha} \int_0^{\infty} e^{-t(\beta-\alpha)} f(\ln K - t) dt. \tag{22}$$

In order to extract the dominant large-K behavior of the r.h.s. in Eq. (21), we approximate the double sum in r.h.s. in Eq. (21) by integrals over variables scaled by K. This gives the following leading asymptotics:

$$\tilde{C}K^{\beta-\alpha}2^{-\beta}\int_0^\infty d\zeta\zeta^\beta e^{-\zeta}\int_{\zeta/2}^\infty \frac{dy}{y}y^{\alpha-\beta}f(\ln K-\ln y), \tag{23}$$

where ζ corresponds to $2k/n$, and y corresponds to K/n. The same procedure gives the following representation for the moment in the l.h.s. of (21):

$$M_\beta(K)\sim\tilde{C}K^{\beta-\alpha}\int_0^\infty \frac{dy}{y}y^{\alpha-\beta}f(\ln K-\ln y) \tag{24}$$

To check consistency, we first substitute the simplest ansatz $f(\ln K) = 1$, corresponding to a pure power-law asymptotic tail in $p(k)$. This immediately gives in the l.h.s. of Eq. (20) $M\beta(K) \sim \tilde{C}K^{\beta-\alpha}/(\beta-\alpha)$. The inner integral in the r.h.s. is $(\zeta/2)^{\alpha-\beta}/(\beta-\alpha)$, and therefore the outer integral over ζ takes the form

$$\tilde{C}\frac{K^{\beta-\alpha}}{\beta-\alpha}2^{-\alpha}\int_0^\infty d\zeta\zeta^\alpha e^{-\zeta}=\tilde{C}\frac{K^{\beta-\alpha}}{\beta-\alpha}\frac{\Gamma(\alpha+1)}{2^\alpha}, \tag{25}$$

the last factor being equal to 1 by virtue of the definition of α, Eq. (16). This result shows that a pure asymptotic power-law form of $p(n)$, which was originally conjectured by identifying the critical index α separating convergent and divergent moments is indeed consistent with the dominant asymptotic behavior of the truncated divergent moments.

Returning now to the full structure involving the conjectured logarithmic corrections given by f, we note that if the condition $K(\beta-\alpha) \gg 1$ is satisfied, the integrals are dominated by values of y such that $\ln y \ll \ln K$. Therefore an asymptotic expansion can be obtained by expanding f in powers of $\ln y$ near $\ln K$. Since $f(x)$ is assumed slower than exponential to ensure that it cannot affect the overall exponent of the power law, each subsequent derivative of f is parametrically (in $\ln K$) smaller than the preceding one. (This argument, of course, pre-supposes the existence of the derivatives, however, $f(x)$ can be assumed to be obtained by analytic continuation from discrete points $\ln n$.) The moment in the l.h.s. now takes the form

$$M_\beta(K)\sim\tilde{C}K^{\beta-\alpha}\left[\frac{f(\ln K)}{\beta-\alpha}+\frac{f'(\ln K)}{(\beta-\alpha)^2}+\frac{f''(\ln K)}{(\beta-\alpha)^3}+\dots\right].$$

The expansion on the right, on the other hand, has a more complicated structure. E.g., in the first order, the inner integral gives $f'(\ln K)(\zeta/2)^{\alpha-\beta}\left[1/(\beta-\alpha)^2+\ln(\zeta/2)/(\beta-\alpha)\right]$. After the second integration, the first term exactly matches the corresponding term in the expansion of $M_\beta(K)$, however, the second term gives an additional contribution proportional to $f'(\ln K)$ with a positive coefficient $\int_0^\infty \exp(-\zeta)\left(\frac{\zeta}{2}\right)^\alpha \ln\left(\frac{\zeta}{2}\right)d\zeta \approx .6857$. Therefore, $f'(\ln K) = 0$.

Similarly, at all higher orders $m \geq 1$, $\int_{\zeta/2}^\infty \frac{dy}{y}y^{\alpha-\beta}\ln^m y$ gives $(\zeta/2)^{\alpha-\beta}/(\beta-\alpha)^{m+1}$ times a polynomial in powers of $(\beta-\alpha)\ln(\zeta/2)$ with positive coefficients, and each resulting integral $\int_0^\infty \exp(-\zeta)\left(\frac{\zeta}{2}\right)^\alpha \ln^l\left(\frac{\zeta}{2}\right)d\zeta$ with integer l's is also positive for the given value of α. Hence, all derivatives of f must vanish, and its constant value is absorbed in $\tilde{C}$. We have thus shown that the asymptotic decay of $p(k)$ at large k is a pure power law with the exponent $\alpha+1$.

B. Generalization to models with re-wiring.

We now generalize the model considered in the previous subsection by allowing for addition of new edges connecting existing nodes. At each discrete time $t+1$ either (with probability q) a new node joins the network, acquiring a fitness value as described above, and connecting to an existing node according to Eq. (1), or (with probability $1-q$) a new edge is added connecting existing nodes i and j with probability

$$\Pi_t(i,j)=\frac{x_ix_j}{\frac{1}{2}\sum_{n=1}^{N_t}\sum_{m=1}^{N_t}x_nx_m}. \tag{26}$$

The total number of nodes N_t is now a stochastic variable, however, in the $t \gg 1/q$ limit it can be replaced by its expectation qt, corrections being sub-leading in $1/t$. It is also immediately evident that the expected node degree is $\bar{k}_q = 2t/qt = 2/q$ (again neglecting sub-leading terms in $1/t$). The relation $\bar{x}_t = \bar{k}_q$ is still true, and therefore the denominator in Eq. (26) asymptotically converges to $\frac{1}{2}N_t^2\bar{x}_t^2 \to 2t^2$, while the denominator in Eq. (1) still converges to $N_t\bar{x}_t \to 2t$. Consequently, the total probability that a node with fitness x acquires a link at time $t+1$ is

$$q\frac{x}{2t} + (1-q)\frac{x\sum_j x_j}{2t^2} \to \frac{x}{2t}(2-q).$$

Denoting $\nu = 2-q$ for brevity, equation (2) now takes the form

$$p_{t+1}(k|x,\tau) = p_t(k-1|x,\tau)\frac{x\nu}{2t}[1-\delta_{k,1}] + p_t(k|x,\tau)\left(1-\frac{x\nu}{2t}\right). \tag{27}$$

Repeating the steps leading to Eq. (6) leads to

$$G(s) = s\int_0^1 dzG\left(z^{\nu\frac{1-s}{2}}\right), \tag{28}$$

and the analog of Eq. (7) is

$$p(k) = \sum_{n=1}^{\infty}\frac{2p(n)}{n\nu}\left(\frac{n\nu}{n\nu+2}\right)^k. \tag{29}$$

As an example, Fig. 5 and Fig. 6 show the generating function and the corresponding degree distribution for the case $q=1/2$ obtained by employing the iteration procedure outlined in the previous subsection.

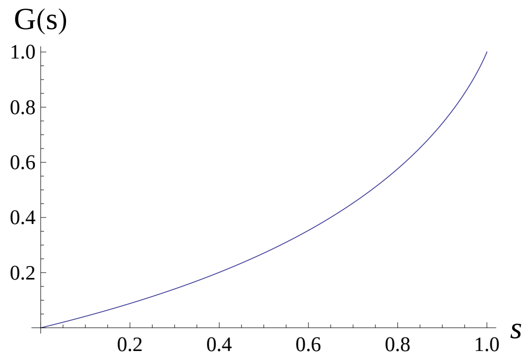

FIG. 5: The fixed-point generating function in the case $q = 1/2$.

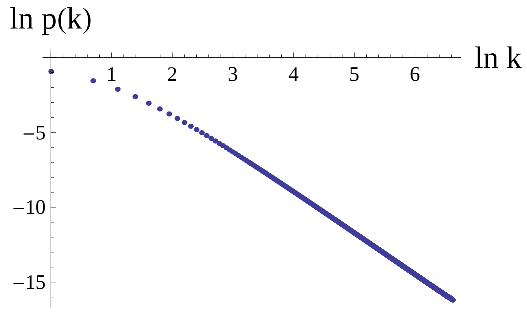

FIG. 6: Degree distribution in the case $q = 1/2$.

Calculation of the second moment of $p(k)$ gives

$$\overline{k^2}_q = \frac{1+\nu}{(1-\nu^2/2)(1-\nu/2)}, \tag{30}$$

so that it exists only if $2-\sqrt{2} < q \le 1$. Generalizing the derivation that led to Eq. (16) one now finds that the exponent α_q of the asymptotic decay $p_k \sim k^{-\alpha_q-1}$ is given by the solution of the equation

$$1 = \Gamma(1+\alpha_q)(\nu/2)^{\alpha_q}. \tag{31}$$

The absence of logarithmic corrections is established straightforwardly using the results of the previous subsection. Since $1 \leq \nu < 2$, solutions of Eq. (31) lie in the interval $1 < \alpha_q \leq \alpha$. This is the range where the estimated exponents of most of the empirically observed networks can be found. Figure 7 shows the graph of the exponent $\alpha_q + 1$ plotted against the expected node degree $\bar{k}_q = 2/q$. The dots on the graph correspond to the empirically observed values for a number of networks that have been compiled by KONECT (the Koblenz Network Collection) [10]. Only the data corresponding to undirected networks with simple edges are shown [11].

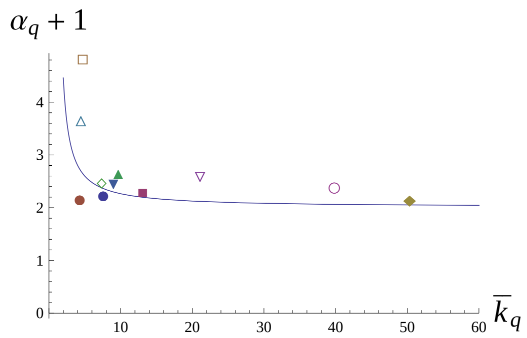

FIG. 7: Exponent of the power-law decay against the expected node degree, together with empirically observed values.

III. THE NONLINEAR MODELS.

A natural generalization of the models considered in the previous Section is to allow for non-linear mapping between fitnesses and degrees. For technical reasons, it is convenient to transfer the non-linearity into the definition of the attachment probability:

$$\pi_t(i) = \frac{f(x_i)}{\sum_{j=1}^{t} f(x_j)}, \tag{32}$$

where $f(x)$ is the linking function that implements the mapping. The fitnesses x_i, as before, are assigned to each new node probabilistically from the instantaneous distribution of degrees. To avoid cluttering the calculation, only the tree version of the model is considered here. Repeating the steps leading to Eq. (6), we obtain the following equation on the average stationary generating function of the degree distribution:

$$G(s) = s \int_0^1 dz \sum_n p(n) z^{(f(n)/\bar{f})(1-s)}, \tag{33}$$

where a crucial assumption has been made that in the long-time limit the average linking function

$$\bar{f} = \lim_{t\to\infty} (1/t) \sum_{j=1}^{t} f(x_j)$$

is finite. The limits of validity of this assumption will be discussed below.

Expanding in the powers of s, we find the analog of Eq. (7):

$$p(k) = \sum_{n=1}^{\infty} \frac{\bar{f} p(n)}{f(n)} \left(\frac{f(n)}{\bar{f} + f(n)} \right)^k. \tag{34}$$

It is straightforward to verify that the sum rule $\bar{k} = 2$, enforced by the growth rules, is satisfied. The assumption that $\bar{f}$ is finite is automatically true for any $f(x)$ that grows asymptotically at large x no faster than linear. This follows from the fact that $\bar{k} = \sum_k k p_t(k)$ is finite and equal to 2 for any distribution, whether stationary or not. Therefore, if $f(x) < Ax$ for some finite constant A, then $\mathbb{E}[(1/t) \sum_{j=1}^{t} f(x_j)] = \int_0^t (d\tau/t) \sum_{k=1}^{\infty} f(k) p_\tau(k) \leq A \int_0^\tau (d\tau/t) \sum_{k=1}^{\infty} k p_\tau(k) = 2A$, and it can be similarly shown that the variance vanishes as $t \to \infty$ [12].

Let us first consider the case where $f(x) \sim x$ as $x \to \infty$. The coefficient of proportionality can be set to one since it is scaled out of Eq. (32). Using again the fractional moment method, we find the analog of Eq. (16):

$$1 = \Gamma(\alpha_f + 1)/\bar{f}^{\alpha_f} \tag{35}$$

which determines the critical index α_f separating convergent and divergent moments and therefore determines the power-law asymptotics of the node degree distribution. Unlike the closed form structure of Eq. (16), the equation on α_f involves $\bar{f}$ whose value has to be determined self-consistently from the full stationary distribution $p(k)$. Qualitatively, $\bar{f}$ decreases if the the weight of the linking function is moved away from lower values of k. It is easy to see from Eq. (35) that $\bar{f} = (\alpha_f/e)(2\pi\alpha_f)^{1/2\alpha_f}$ asymptotically at large α_f, and numerically the relation is approximately linear at all $\bar{f} \geq 1$, as seen in Fig. 8.

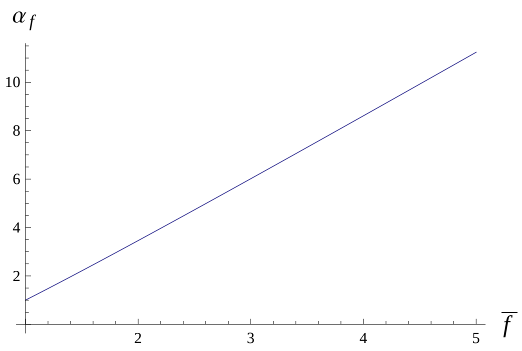

FIG. 8: Solution of Eq. (35) as a function of $\bar{f}$.

For a given choice of $f(k)$ the degree distribution $p(k)$ can be found numerically using iterations of the matrix kernel in Eq. (34) which, as before, has the spectral radius equal to 1. The two examples shown in Fig. 9 and Fig. 10 below correspond to the choices

$$f^{(1)}(x) = \begin{cases} 1/4, & x = 1 \\ x, & x \geq 2 \end{cases} \tag{36}$$

and

$$f^{(2)} = \begin{cases} 1/4, & x = 1 \\ 1/2, & x = 2 \\ x, & x \geq 3. \end{cases} \tag{37}$$

The exponents of the corresponding asymptotic power laws are $\alpha_f^{(1)} = 2.152$ and $\alpha_f^{(2)} = 1.488$, which are found by numerically evaluating $\bar{f}$ and substituting the values $\bar{f}^{(1)} = 1.478$ and $\bar{f}^{(2)} = 1.204$ into Eq. (35).

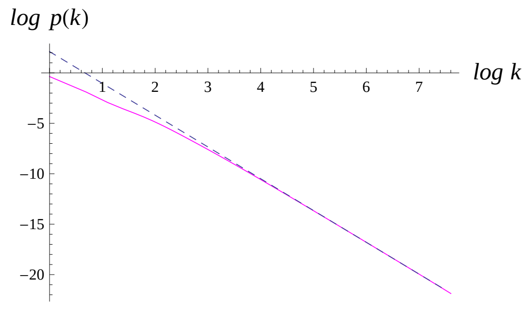

FIG. 9: Logarithm of the node degree distribution in the nonlinear model with the linking function $f^{(1)}$ together with the asymptotic power law with the exponent $\alpha_f^{(1)} = 2.152$. The constant offset is found by numerical fitting.

Let us now consider the cases when $f(x)$ is asymptotically nonlinear. If $f(x)$ grows faster than x, the degree distribution does not possess a stationary fixed point, as can be seen from the following argument. Suppose $f(x) \sim$

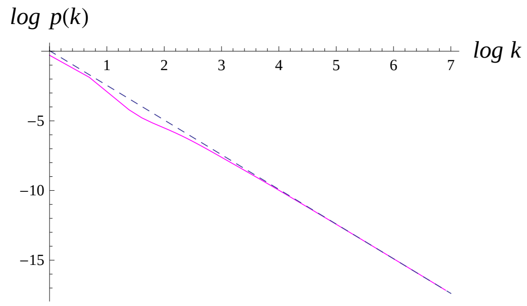

FIG. 10: Logarithm of the node degree distribution in the nonlinear model with the linking function $f^{(2)}$ together with the asymptotic power law with the exponent $\alpha_f^{(2)} = 1.488$. The constant offset is found by numerical fitting.

$x^{1+\epsilon}$, $\epsilon > 0$, and $p(k)$ converges to a stationary distribution with the asymptotic behavior $k^{-2-\gamma}$ (the sum rule $\bar{k} = 2$ requires $\gamma > 0$). Consider Eq. (34) in the limit $k \gg 1$. The factor $[f(n)/(\bar{f} + f(n))]^k$ suppresses all contributions to the sum over n below n_k such that $f(n_k) \sim k\bar{f}$. Therefore the sum in the r.h.s. of Eq. (34) can be estimated as

$$\bar{f} \sum_{n \sim n_k}^{\infty} p(n)/f(n) \sim \bar{f} \sum_{n \sim n_k} n^{-3-\gamma-\epsilon} \propto k^{-(2+\epsilon+\gamma)/(1+\epsilon)}. \tag{38}$$

Since the sum is equal to p_k, we have $(2 + \epsilon + \gamma)/(1 + \epsilon) = 2 + \gamma$, or $0 = \epsilon(1 + \gamma)$. This equation cannot be solved for $\epsilon > 0$ and $\gamma > 0$, thus the assumption that a stationary distribution exists leads to a contradiction.

Finally, if $f(x)$ is asymptotically sub-linear, $f(x) \sim x^{1-\epsilon}$ with $\epsilon > 0$, we assume the following ansatz of $p(k)$ at large k:

$$p(k) \sim \exp\{-\gamma k^{\delta}\} \tag{39}$$

with some positive constants γ and $\delta < 1$. The sum in the r.h.s. of Eq. (34) in the large-k regime can be approximated by the corresponding integral,

$$\sum_n \frac{\bar{f}}{f(n)} \left(\frac{f(n)}{\bar{f} + f(n)}\right)^k \sim \int dn \frac{\bar{f}}{n^{1-\epsilon}} \exp\{-k\bar{f}/n^{1-\epsilon} - \gamma n^{\delta}\}, \tag{40}$$

and evaluated in the saddle-point approximation. The saddle-point equation is solved by

$$n_0 = \left[\frac{(1-\epsilon)k\bar{f}}{\gamma\delta}\right]^{1/(1-\epsilon+\delta)}. \tag{41}$$

Equation (34) requires $\gamma k^{\delta} = k\bar{f}/n_0^{1-\epsilon} + \gamma n_0^{\delta}$, from where it follows that $\delta = \epsilon$, and $\gamma = \frac{\bar{f}}{\epsilon}(1-\epsilon)^{1-1/\epsilon}$. Therefore sub-linear growth of $f(x)$ leads to the stationary distribution acquiring the form of stretched exponential. The graphs in Figs. 11 and 12 show the linear behavior of $\ln^2 p(k)$, and $\ln^4 p(k)$, respectively corresponding to the choices $f(x) \sim \sqrt{x}$ and $f(x) \sim x^{3/4}$.

IV. CONCLUSION

Universal features of diverse empirically observed large networks imply the existence of some universal drivers of their behavior that are to some extent insensitive to the microscopic details of the network evolution mechanisms. The most likely mathematical expression of such a mechanism is convergence to a fixed point, which necessitates a feedback process whereby the current (or, more generally, the history of the) network topology affects the current network evolution rules. The classical preferential attachment model [2] achieves this by identifying the instantaneous value of the node degree itself as a proxy for node attractiveness. Once the assumption is made that a separate measure of attractiveness, fitness, plays a role in the network evolution, the feedback mechanism needs to be introduced explicitly, as in its absence most fitness-based mechanisms possess a high degree of arbitrariness [7].

The goal of the present study was to investigate a simple class of such models which combine a fitness-based growth mechanism with input from dynamic information about network topology. Of course, such models do not fully catch

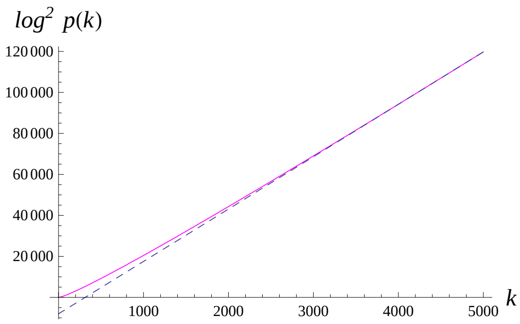

FIG. 11: Square of the logarithm of the node degree distribution in the nonlinear model with the linking function $f(x) = \sqrt{x}$ together with the linear fit.

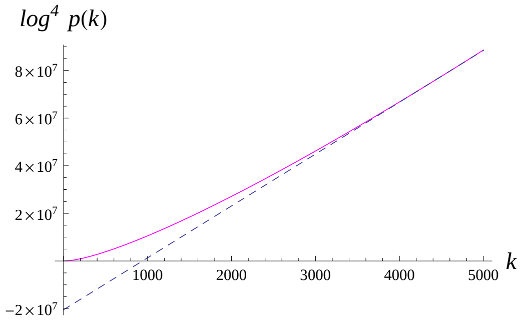

FIG. 12: Square of the logarithm of the node degree distribution in the nonlinear model with the linking function $f(x) = x^{3/4}$ together with the linear fit.

the complexities of realistic growing networks. Most importantly, the fitness of an existing node is taken to be fixed at the time of its creation, disallowing dynamic updating of whatever proxy measure of attractiveness is in operation. (In a sense, the classical preferential attachment model can be viewed as a limiting case of dynamic fitness being instantaneously updated to be equal to the current node degree.) Within these limitations, it was demonstrated that models characterized by dynamic updating of the *distribution* of the incoming node fitnesses exhibit convergence to power-law asymptotics provided that mapping between fitness and attachment probability is linear, and more general stretched exponential behavior in the non-linear case.

The sensitivity of the power-law exponent to the details of attachment rules at low k is similar to the one seen in the preferential attachment model. One could view this feature as an indication that pure fitness-based growth models, even in the presence of the dynamic feedback mechanism studied here, lack some essential stabilization through feedback from network topology. Further study of more realistic fitness-based models with feedback that allow for some fitness dynamics and for edge and node deaths may point towards more robust stabilization mechanisms.

[1] S. N. Dorogovtsev and J. F. F. Mendes, Adv. Phys. **51**, 1079 (2002).
[2] P. Krapivsky and S. Redner, Phys. Rev. E **63**, 066123 (2001).
[3] R. Albert and A.-L. Barabási, Phys. Rev. Lett. **85**, 5234 (2000).
[4] V. D. P. Servedio, G. Caldarelli, and P. Butta, Phys. Rev. E **70**, 056126 (2004).
[5] C. Bedogne and G. J. Rodgers, Phys. Rev. E **74**, 046115 (2006).
[6] X.-J. Xu, L.-M. Zhang, and L.-J. Zhang, Int. J. of Mod. Phys. C **21**, 129 (2010).
[7] I. E. Smolyarenko, K. Hoppe, and G. J. Rodgers, Phys. Rev. E **88**, 012805 (2013).
[8] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, eds., *NIST Handbook of Mathematical Functions* (Cambridge University Press, 2010).
[9] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics: A Foundation for Computer Science* (Addison-Wesley, Reading, MA, 1989).
[10] J. Kunegis, in *Proc. Int. Web Observatory Workshop* (2013) pp. 1343–1350.
[11] Several undirected networks were excluded from the graph in Fig. 7: networks with the degree exponent less than 2 clearly cannot correspond to the type of models considered here, since they do not possess a finite expected degree. Also excluded

were several networks where degree distribution is manifestly not showing power law behavior. The values for the exponents are the ones available at the KONECT web site, and no independent analysis of the raw data to rule power-law behavior in or out (e.g., using the methodology of Ref. [13]) has been performed.

[12] Variance and higher central moments of the linking function that grows at most linearly may formally stay finite at large times in the extreme case of a star-like network, *i.e.* a network that possess $O(1)$ nodes with degrees $k \sim O(t)$. However, existence of finite moments is sufficient to show that the degree distribution converges to a stationary limit in which the sum rule $\bar{k} = 2$ is saturated at $O(1)$ in the $1/t$ expansion, thus excluding the existence of nodes with $O(t)$ degrees, hence this scenario is self-contradictory.

[13] A. Clauset, C. R. Shalizi, and M. E. J. Newman, SIAM Review **51(4)**, 661 (2009).