

## Zero Pearson coefficient for strongly correlated growing trees

S. N. Dorogovtsev,<sup>1,2</sup> A. L. Ferreira,<sup>1</sup> A. V. Goltsev,<sup>1,2</sup> and J. F. F. Mendes<sup>1</sup>

<sup>1</sup>*Departamento de Física, I3N, Universidade de Aveiro, 3810-193 Aveiro, Portugal*

<sup>2</sup>*A. F. Ioffe Physico-Technical Institute, 194021 St. Petersburg, Russia*

(Received 22 November 2009; published 31 March 2010)

We obtained Pearson's coefficient of strongly correlated recursive networks growing by preferential attachment of every new vertex by  $m$  edges. We found that the Pearson coefficient is exactly zero in the infinite network limit for the recursive trees ( $m=1$ ). If the number of connections of new vertices exceeds one ( $m > 1$ ), then the Pearson coefficient in the infinite networks equals zero only when the degree distribution exponent  $\gamma$  does not exceed 4. We calculated the Pearson coefficient for finite networks and observed a slow power-law-like approach to an infinite network limit. Our findings indicate that Pearson's coefficient strongly depends on size and details of networks, which makes this characteristic virtually useless for quantitative comparison of different networks.

DOI: [10.1103/PhysRevE.81.031135](https://doi.org/10.1103/PhysRevE.81.031135)

PACS number(s): 05.40.-a, 05.10.-a, 05.50.+q, 87.18.Sn

### I. INTRODUCTION

The Pearson coefficient  $r$  is used as an integral characteristic of structural correlations in a network. Pearson's coefficient characterizes pairwise correlations between degrees of the nearest-neighboring vertices in networks. Some observable quantities in correlated networks (e.g., the size of a giant connected component near the point of its emergence) are directly expressed in terms of this coefficient [1,2]. The Pearson coefficient is the normalized correlation function of the degrees of the nearest-neighbor vertices [3–5]. The coefficient is the ratio,

$$r = \frac{\langle jk \rangle_e - \langle k \rangle_e^2}{\langle k^2 \rangle_e - \langle k \rangle_e^2} = \frac{\langle (j - \langle k \rangle_e)(k - \langle k \rangle_e) \rangle_e}{\sigma^2}, \quad (1)$$

where  $\langle jk \rangle_e$  is the average product of the degrees  $j$  and  $k$  of the end vertices of an edge,  $\langle k \rangle_e = \langle k^2 \rangle / \langle k \rangle$  is the average degree of an end vertex of an edge,  $\langle k^2 \rangle_e = \langle k^3 \rangle / \langle k \rangle$  is the average square of the degree of an end vertex of an edge, and

$$\sigma^2 \equiv \frac{\langle k^3 \rangle}{\langle k \rangle} - \frac{\langle k^2 \rangle^2}{\langle k \rangle^2} \quad (2)$$

is for normalization. (Note that the equality  $\langle k^n \rangle_e = \langle k^{n+1} \rangle / \langle k \rangle$  is valid for an arbitrary network.) Here  $\langle \dots \rangle$  and  $\langle \dots \rangle_e$  denote averaging over vertices and edges, respectively [see Eqs. (16), (18), and (20) below]. Pearson's coefficient can be positive (in average, assortative mixing of the nearest neighbors' degrees) or negative (disassortative mixing) and takes values in the range from  $-1$  to  $1$ . The Pearson coefficient  $r$  is a convolution of the joint distribution of nearest-neighbor degrees,  $P(j, k) \equiv e_{jk}$ . This joint distribution is the probability that the ends of a randomly chosen edge have degrees  $j$  and  $k$ ,  $\sum_{jk} e_{jk} = 1$ . Being an integral characteristic of degree-degree correlations, the Pearson coefficient misses details of these correlations, compared to  $e_{jk}$  [6–8]. Despite this fact, Pearson's coefficient is widely used for characterization and comparison of real-world networks [4]. Note that the compared real networks have different sizes. In this paper we show that since  $r$  is a markedly size-dependent quantity, Pearson's coefficient may be used for comparison of networks only with a very critical attitude. We calculate Pear-

son's coefficient for the simplest growing complex networks, namely, recursive random networks with preferential attachment of new vertices. We describe the size dependence of  $r$  in these networks (Fig. 1).

Remarkably, for all infinite recursive trees of this kind, we find that the Pearson coefficient is exactly zero at any value of the degree distribution exponent  $\gamma$ , although all these networks are strongly correlated. The statement that Pearson's coefficient is zero in the range  $\gamma > 4$  is essentially nontrivial. The point is that the zero value of the Pearson coefficient in the range  $\gamma \leq 4$  where the third moment of the degree distribution diverges is clear. Indeed, in this region, the denominator in definition (1) diverges in the infinite network limit. [If  $\gamma \leq 3$ , the numerator in definition (1) also diverges. This divergence, however, is slower than that of the denominator, see Eq. (33), which leads to zero Pearson's coefficient in this range.] In contrast, the denominator approaches a finite value if  $\gamma > 4$  and, for zero  $r$ , the numerator in Eq. (1) must be zero. In this range, the zero value of Pearson's coefficient

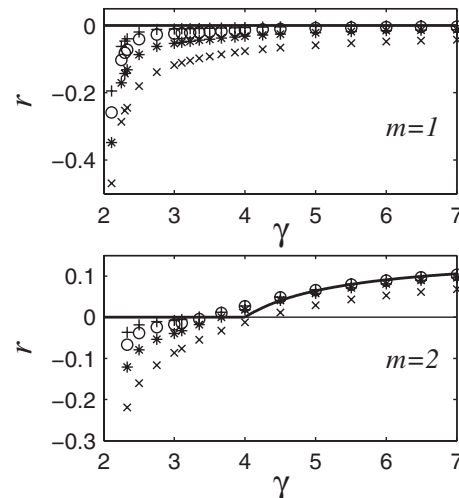


FIG. 1. Pearson's coefficient  $r$  vs degree distribution exponent in the recursive networks with  $m=1$  and  $2$  for the networks of  $10^3$  ( $\times$ ),  $10^4$  ( $*$ ),  $10^5$  ( $\circ$ ), and  $10^6$  ( $+$ ) vertices. The thick lines show the Pearson coefficient for the infinite networks.

means a surprising exact mutual compensation of different degree-degree correlations in these trees. There are two opposing kinds of degree-degree correlations in complex networks: assortative and disassortative. Here, assortativity is a tendency of high degree vertices to have high degree neighbors and low degree vertices to have low degree neighbors. In contrast, disassortative mixing means neighborhood of vertices with contrasting (low and high) degrees. These tendencies may be opposing in different ranges of degrees; see discussion in Sec. IV (Fig. 4). That is, assortative and disassortative mixing may coexist. In particular, this is the case for random recursive trees. Our results show that, in these growing networks, the different kinds of correlations completely compensate each other in the infinite network limit.

**II. MAIN RESULTS**

We study recursive networks in which each new vertex is attached to  $m$  existing ones chosen with probability proportional to a linear function of vertex degree,  $k+A \equiv k+am$ . This rule generates scale-free correlated networks with a degree distribution exponent  $\gamma=3+A/m=3+a$ . In our simulations, we used a complete graph of  $m$  vertices as an initial configuration. We studied sufficiently large networks of  $10^3-10^6$  vertices, and our observations did not depend on initial conditions. Note that at  $m > 1$ , the generated recursive networks contain multiple connections. These multiple edges are essential if  $\gamma < 3$ . In this range of exponent  $\gamma$ , the exclusion of multiple connections would change degree-degree correlations in a network. Nonetheless, one can show that the removal of the multiple edges will not change Pearson's coefficient for the infinite networks (see the Appendix).

In the infinite network limit, Pearson's coefficient  $r_\infty(\gamma \leq 4) = 0$  since the third moment of the degree distribution [see the denominator of Pearson's coefficient definition (1)] diverges. Here, we define  $r_\infty \equiv r(t \rightarrow \infty)$ , where  $t$  is the number of vertices in a network. For  $\gamma > 4$ , we find

$$r_\infty = \frac{(m-1)(a-1)[2(1+m)+a(1+3m)]}{(1+m)[2(1+m)+a(5+7m)+a^2(1+7m)]}, \quad (3)$$

which shows that for  $m=1$  (i.e., for random recursive trees), Pearson's coefficient  $r_\infty=0$  for any value of  $\gamma$ . One can also see that  $r_\infty(\gamma=4)=0$  for any  $m$ . In particular, for uniformly random attachment, i.e.,  $\gamma \rightarrow \infty$ , we have

$$r_\infty(\gamma \rightarrow \infty) = \frac{(m-1)(1+3m)}{(1+m)(1+7m)}. \quad (4)$$

Figure 1, obtained by simulations of the networks of  $10^3, 10^4, 10^5$ , and  $10^6$  vertices (the number of runs for each point was between 50 and 500 for  $\gamma < 3$  and between  $5 \times 10^3$  and  $10^4$  for  $\gamma > 3$ ), demonstrates the size dependence of Pearson's coefficient in these networks. This figure shows that, even for large networks, the deviations from the limiting infinite network values are significant. We find the following asymptotic size dependences  $\delta r(t) = r(t) - r_\infty$ :

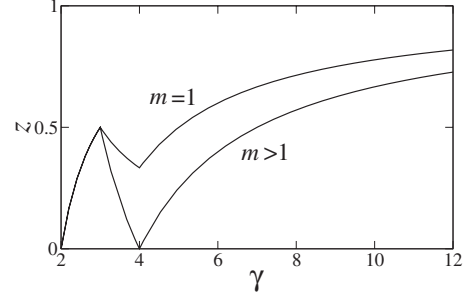


FIG. 2. Variation of exponent  $z[\delta r(t) = r(t) - r_\infty \propto t^{-z}]$  with  $\gamma$  exponent of a degree distribution.

$$\delta r(t) \propto \begin{cases} -t^{-(\gamma-2)/(\gamma-1)}, & \text{any } m, & \gamma < 3 \\ -t^{-1/(\gamma-1)}, & m = 1 \\ +t^{-(4-\gamma)/(\gamma-1)}, & m > 1 \end{cases} \quad \left. \begin{matrix} \\ \\ \end{matrix} \right\} \begin{matrix} 3 < \gamma < 4 \\ \gamma > 4. \end{matrix} \quad (5)$$

Here we show the signs of the asymptotes but ignore their factors. The positive sign of the asymptotes at  $m > 1, 3 < \gamma < 4$ , and  $\gamma > 4$  means that  $r(t)$  approaches the infinite-size limit  $r_\infty$  from above. In this situation ( $m > 1, \gamma > 3$ ), Pearson's coefficient varies with size nonmonotonously: first increases and then diminishes to  $r_\infty$  (see the second panel in Fig. 1). Introducing the exponent  $z$ ,  $\delta r(t) \propto t^{-z}$  for large network sizes  $t$ , we arrive at the dependences  $z(\gamma)$  shown in Fig. 2. Note that  $z < 1$ , so the infinite network limit is approached slowly. The relaxation to the infinite network values,  $r_\infty$ , is especially slow if the exponent  $\gamma$  is close to 2 (at any  $m$ ) or to 4 (only if  $m > 1$ ). In the specific case of  $m > 1, \gamma = 4$ , we obtain the logarithmic relaxation,

$$\delta r(t) \propto \frac{1}{\ln t}. \quad (6)$$

We measured the dependence of the exponent  $z$  on  $\gamma$  in the simulated networks. The result, shown in Fig. 3, demonstrates an agreement with above analytical predictions.

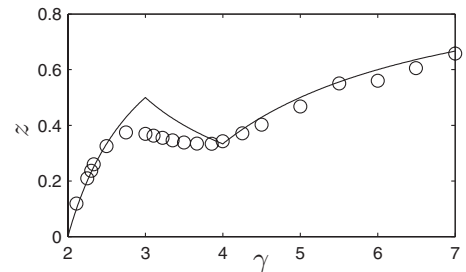


FIG. 3. The measured dependences of exponent  $z$  on  $\gamma$  for recursive trees ( $m=1$ ). The solid curve is the theoretical dependence  $z(\gamma, m=1)$ . The data points are obtained by power-law fitting of the Pearson coefficient of simulated trees with sizes close to the largest simulated size ( $10^6$  vertices).

### III. PEARSON COEFFICIENT OF INFINITE NETWORKS

Let us derive the Pearson coefficient (5) of the infinite recursive networks. The derivation is based on rate equations for the average number  $N_k(t)$  of vertices of degree  $k$  in a recursive network at time  $t$  and the average number  $E_{jk}(t)$  of edges connecting vertices of degrees  $j$  and  $k$ . We have

$$\sum_{k=m} N_k(t) = t \quad (7)$$

for vertices, and so

$$N_k(t) = tp_k(t), \quad (8)$$

where  $p_k(t) \equiv P(k, t)$  is a degree distribution of the network of size  $t$ .

For edges, we have

$$\frac{1}{2} \sum_{j,k} E_{jk}(t) \equiv mt, \quad (9)$$

and so

$$E_{jk}(t) \equiv 2mte_{jk}(t), \quad (10)$$

where  $e_{jk}$  is the degree-degree distribution for edges. Using the standard rate or master-equation approaches [9–11] to this kind of networks (specifically, to the recursive networks growing due the preferential attachment mechanism), we write the following rate equations for  $N_k(t)$  and  $E_{jk}(t)$ :

$$N_m(t+1) = N_m(t) + 1 - m \frac{m+ma}{m(2+a)t} N_m(t),$$

$$N_{k>m}(t+1) = N_k(t) + m \frac{k-1+ma}{m(2+a)t} N_{k-1}(t) - m \frac{k+ma}{m(2+a)t} N_k(t) \quad (11)$$

(note that the mean vertex degree in these networks is  $\langle k \rangle = 2m$ ) and

$$\begin{aligned} E_{mk}(t+1) &= E_{mk}(t) + m \frac{k-1+ma}{m(2+a)t} N_{k-1}(t) \\ &\quad - m \left[ \frac{m+ma}{m(2+a)t} E_{mk}(t) + \frac{k+ma}{m(2+a)t} E_{mk}(t) \right], \\ E_{jk>m}(t+1) &= E_{jk}(t) + m \left[ \frac{k-1+ma}{m(2+a)t} E_{j,k-1}(t) \right. \\ &\quad \left. + \frac{j-1+ma}{m(2+a)t} E_{j-1,k}(t) \right] - m \left[ \frac{j+ma}{m(2+a)t} E_{jk}(t) \right. \\ &\quad \left. + \frac{k+ma}{m(2+a)t} E_{jk}(t) \right]. \quad (12) \end{aligned}$$

We separately write out equations for the case of  $k=m$ . Here, we used the fact that the probability to attach a new vertex to a vertex  $i$  of degree  $k_i$  in this model is

$$\frac{k_i+ma}{\sum_i (k_i+ma)} = \frac{k_i+a}{(\langle k \rangle + ma)t} = \frac{k_i+a}{m(2+a)t}. \quad (13)$$

Assuming a stationary regime at the limit  $t \rightarrow \infty$  [ $N_k(t) \equiv tp_k$ ,  $E_{jk}(t) \equiv 2mte_{jk}$ ], we reduce these equations to

$$p_m = 1 - m \frac{m+ma}{m(2+a)} p_m,$$

$$p_{k>m} = m \frac{k-1+ma}{m(2+a)} p_{k-1} - m \frac{k+ma}{m(2+a)} p_k, \quad (14)$$

$$2m(k+m+2+2ma+a)e_{mk} = (k-1+ma)p_{k-1},$$

$$\begin{aligned} m(j+k+2+2ma+a)e_{jk>m} &= m(k-1+ma)e_{j,k-1} \\ &\quad - m(j-1+ma)e_{j-1,k}. \quad (15) \end{aligned}$$

To obtain the Pearson coefficient [see definition (1)], we must find  $\langle k^2 \rangle$ ,  $\langle k^3 \rangle$ , and  $\langle jk \rangle_e$ , where

$$\langle jk \rangle_e = \sum_{jk} jke_{jk}. \quad (16)$$

Multiplying both sides of the second equation of system (14) by  $k^2$  and  $k^3$  and summing over  $k$ , and taking into account that

$$p_m = \frac{2+a}{2+a+m+ma}, \quad (17)$$

which follows from the first equation of system (14), we obtain

$$\langle k^2 \rangle = \sum_k k^2 p_k = \frac{m}{a} (2+5ma+a+2m), \quad (18)$$

$$\begin{aligned} \langle k^3 \rangle &= \frac{m}{a(a-1)} [6(1+m) + a(5+21m+8m^2) \\ &\quad + a^2(1+9m+16m^2)]. \quad (19) \end{aligned}$$

respectively.

To get  $\langle jk \rangle_e$ , we multiply both sides of the second equation of system (15) by  $jk$  and sum them over  $j$  and  $k$ . We also take into account the following general equality:

$$\langle k \rangle_e = \sum_{jk} ke_{jk} = \frac{\langle k^2 \rangle}{\langle k \rangle}, \quad (20)$$

which gives

$$\langle k \rangle_e = \frac{1}{2a} (2+5ma+a+2m). \quad (21)$$

As a result we find

$$\langle jk \rangle_e = \frac{m}{a^2} [2+2m+a(5+7m)+a^2(2+7m)], \quad (22)$$

and so at  $a > 0$ , i.e.,  $\gamma > 3$ , the numerator of definition (1) is

$$\langle jk \rangle_e - \langle k \rangle_e^2 = \frac{m-1}{4a^2} [4(1+m) + 4a(1+2m) + a^2(1+3m)]. \quad (23)$$

Importantly, this numerator is zero if  $m=1$ . So we see that  $r_\infty(m=1)=0$  at least if  $\gamma > 3$ . On the other hand, if  $m > 1$ , the numerator is finite in the range  $\gamma > 3$ . This shows that  $r_\infty(m > 1)$  must be zero at  $3 < \gamma \leq 4$  since the denominator diverges if  $\gamma \leq 4$ . Finally, at  $\gamma \leq 3$ ,  $r_\infty=0$  at any  $m$  [see the next section; Eq. (33)]. Substituting relations (18), (19), and (23) into definitions (1) and (2) readily gives the resulting expression (3) for the stationary value  $r_\infty$  of Pearson's coefficient.

#### IV. SIZE DEPENDENCE

Equations (11) and (12) allow one to derive the full size dependence of the Pearson coefficient. Instead of these cumbersome straightforward calculations, we obtain the asymptotic behavior of  $r(t)$  in an easier way, using known results for the asymptotics of degree-degree correlations in these networks [9,12–14]. The derivation is based on the following expression of the Pearson coefficient in terms of the average vertex degree  $\bar{k}_{nn}(k)$  of the nearest neighbors of the vertices of degree  $k$ :

$$r = \frac{\langle k \rangle \sum_k k(k-1) p_k [\bar{k}_{nn}(k) - \langle k^2 \rangle / \langle k \rangle]}{\langle k \rangle \langle k^3 \rangle - \langle k^2 \rangle^2} \quad (24)$$

$$= \frac{\langle k \rangle \sum_k k^2 p_k \bar{k}_{nn}(k) - \langle k^2 \rangle^2}{\langle k \rangle \langle k^3 \rangle - \langle k^2 \rangle^2}. \quad (25)$$

This expression directly follows from definition (1). The leading asymptotics of  $\bar{k}_{nn}(k)$  can be obtained by using the known exact asymptotics of  $n_{kl}$  for these recursive networks [9]. Here,  $n_{kl}$  is the probability that a descendant vertex of degree  $k$  is connected to an ascendant vertex of degree  $l$ . This quantity satisfies the following relations:

$$\begin{aligned} \sum_l n_{kl} &= p_k, \\ \sum_l (n_{kl} + n_{lk}) &= k p_k, \\ \sum_l l(n_{kl} + n_{lk}) &= k p_k \bar{k}_{nn}(k). \end{aligned} \quad (26)$$

At  $m=1$ , according to Ref. [9],

$$\begin{aligned} n_{kl} &\sim l^{\gamma-2} k^{1-2\gamma} \quad \text{if } 1 \ll l \ll k, \\ n_{kl} &\sim l^{1-\gamma} k^{-2} \quad \text{if } 1 \ll k \ll l. \end{aligned} \quad (27)$$

For  $\gamma > 3$  and any  $m$ ,  $\bar{k}_{nn}(k, t)$  approaches a stationary limit as the network grows. Figure 4 demonstrates this relaxation for two values of the degree distribution exponent  $\gamma$ , namely, for  $\gamma=3.1$  and 4 in the case of  $m=1$ . We can approximate the deviation from the infinite network limit by

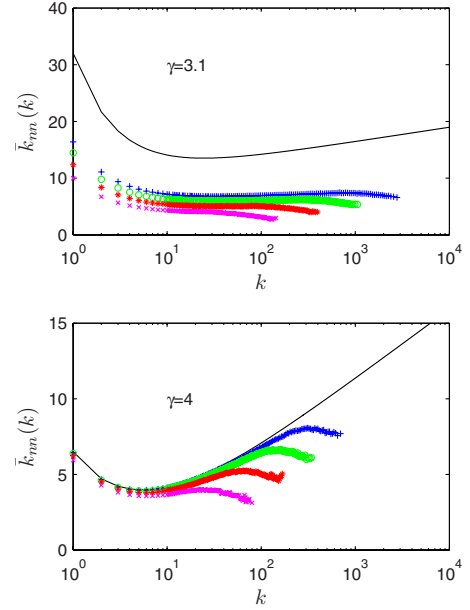


FIG. 4. (Color online) Average degree of the nearest neighbors of a vertex versus the degree of this vertex for the scale-free recursive networks of various sizes with the degree distribution exponent values of 3.1 and 4, and  $m=1$ . The solid lines show the stationary dependences  $\bar{k}_{nn}(k)$  in the infinite networks. These curves were found by numerical solving the stationary limits of the rate equations. Both these dependences, substituted into expression (24) or (25) give exactly zero Pearson coefficient. The positive slope of  $\bar{k}_{nn}(k)$  means assortative mixing, while the negative slope means disassortative mixing. The nonmonotonous dependencies indicate the combination of this kind of degree-degree correlations, which—in this case—exactly compensate each other. The points show the results of simulation of the recursive networks of  $10^3$  ( $\times$ ),  $10^4$  ( $*$ ),  $10^5$  ( $\circ$ ), and  $10^6$  ( $+$ ) vertices. Note a very slow convergence of the results for finite networks to the infinite network limit at  $\gamma$  close to 3.

$$[\bar{k}_{nn}(k) - \bar{k}_{nn}(k, t)] k p_k \sim \int_{k_{\text{cut}}}^{\infty} dl l(n_{kl} + n_{lk}), \quad (28)$$

where  $k_{\text{cut}} = k_{\text{cut}}(t)$  is the time-dependent cutoff of the degree distribution. In these networks,  $k_{\text{cut}}(t) \sim t^{1/(\gamma-1)}$ . Substituting Eq. (27) into Eq. (28) results in

$$\bar{k}_{nn}(k) - \bar{k}_{nn}(k, t) \sim c_1 k^{\gamma-3} t^{(3-\gamma)/(\gamma-1)} + c_2 k^2 \gamma^{-3} t^{-(2\gamma-3)/(\gamma-1)} \quad (29)$$

at large degrees  $k < k_{\text{cut}}$  for the networks with  $\gamma > 3$ . Here,  $c_1$  and  $c_2$  are constants depending on  $\gamma$ , which we do not calculate.

Similarly, for  $2 < \gamma < 3$ , where  $\bar{k}_{nn}(k, t)$  diverges as  $t \rightarrow \infty$  at any  $m$ , we can write an asymptotic estimate,

$$\bar{k}_{nn}(k, t) k p_k \sim \int_m^{k_{\text{cut}}} dl l(n_{kl} + n_{lk}). \quad (30)$$

Substituting Eq. (27) into Eq. (30) we obtain

$$\bar{k}_{nn}(k, t) \sim c_3 k^{\gamma-3} t^{(3-\gamma)/(\gamma-1)} + c_4 k^{2\gamma-3} t^{(3-2\gamma)/(\gamma-1)} \quad (31)$$

at large degrees  $k < k_{\text{cut}}$  for the networks with  $2 < \gamma < 3$ . Here,  $c_3$  and  $c_4$  are some constants. Note that there is some difference between asymptotics (31) and the corresponding result in Ref. [13]. The difference is apparently due to the fact that we derived these expressions by using the exact asymptotics taken from Ref. [9], while the result of Ref. [13] was obtained by using the continuous approximation for the rate equations. Nonetheless, one can easily check that, for  $2 < \gamma < 3$ , the expression for  $\bar{k}_{nn}(k, t)$  from Ref. [13] leads to the same asymptotic behavior of the Pearson coefficient as in the present paper.

We substitute asymptotics (29) and (31) into expression (25) for the Pearson coefficient and take into account the leading terms in the numerator and the denominator. The combinations of these leading terms are different in different areas of  $\gamma$  and  $m$  since the quantities  $\bar{k}_{nn}(k, t)$ ,  $\langle k^2 \rangle$ , and  $\langle k^3 \rangle$  in expression (25) change their asymptotic behavior at two special points, namely,  $\gamma=3$  and 4. In particular, at these points ( $\gamma=3$  and 4) the second and, respectively, the third moments of the degree distribution become divergent. For example, if  $2 < \gamma < 3$  at any  $m$ , relation (25) with substituted asymptotics (29) takes the form

$$r(t) \sim \left[ \left( c_1' t^{(3-\gamma)/(\gamma-1)} \int t^{1/(\gamma-1)} dk k^2 k^{-3+\gamma} k^{-\gamma} \right. \right. \\ \left. \left. + c_2' t^{(3-2\gamma)/(\gamma-1)} \int t^{1/(\gamma-1)} dk k^2 k^{2\gamma-3} k^{-\gamma} \right) \right. \\ \left. - c_3' \left( \int t^{1/(\gamma-1)} dk k^{2-\gamma} \right)^2 \right] / \left( 2m \int t^{1/(\gamma-1)} dk k^{3-\gamma} \right), \quad (32)$$

where  $c_1'$ ,  $c_2'$  and  $c_3'$  are constants. This leads to the result

$$r(t) \sim \frac{[c_1'' \ln t t^{(3-\gamma)/(\gamma-1)} + c_2'' t^{(3-\gamma)/(\gamma-1)}] - c_3'' t^{2(3-\gamma)/(\gamma-1)}}{t^{(4-\gamma)/(\gamma-1)}} \\ \sim -t^{-(\gamma-2)/(\gamma-1)}, \quad (33)$$

where  $c_1''$ ,  $c_2''$ , and  $c_3''$  are constants. In a similar way, we derive the other asymptotics listed in Eqs. (5) and (6).

Interestingly, both the terms in Eqs. (29) and (31) give contributions of the same order of magnitude to  $r(t)$ ; see the numerator of Eq. (33). Note that we suppose that for  $m > 1$  the form of the asymptotics of  $\bar{k}_{nn}(k, t)$  is the same as in Eqs. (29) and (31). To verify our assumption, we inspected the corresponding results for  $\bar{k}_{nn}(k, t)$  in Ref. [13] and found that the asymptotic behavior should be similar at different  $m$  (apart of numerical coefficients) if the exponent  $\gamma$  is fixed. In addition, we checked that results (5) also can be derived by using the asymptotics of  $\bar{k}_{nn}(k, t)$  from Ref. [13]. This confirms our conclusions about the asymptotics of  $r(t)$ .

## V. DISCUSSION AND CONCLUSIONS

Our result, namely, a zero Pearson coefficient of random recursive trees at any  $\gamma > 2$ , naturally leads to the following

questions. What is the class of trees that have zero Pearson coefficient? Is Pearson's coefficient zero for any infinite recursive tree? At present, we cannot answer the first question. As for the second question, the answer is negative. Indeed, as a counterexample, we present the simplest infinite recursive tree with a nonzero Pearson coefficient. This is a star, which is a tree of  $t > 2$  vertices, including  $t-1$  leafs and the hub of degree  $t-1$ . For star of any size, clearly,  $r=-1$ .

Thus, we have studied the Pearson coefficient in strongly correlated growing networks. They form a representative class of networks with strong structural correlations including pairwise correlations between degrees of the nearest-neighbor vertices. Despite these correlations, we have found that in a wide range of infinite correlated networks the Pearson coefficient approaches zero. For any infinite random recursive tree whose growth is driven by arbitrary linear preferential attachment of new vertices, we observed a zero Pearson coefficient. These networks include random recursive trees with rapidly decaying and even exponential degree distribution, where the third moment of the degree distribution is finite. So here the zero value of Pearson's coefficient demands zero correlation function in the numerator of definition (1). This surprising equality to zero indicates an exact mutual compensation of assortative and disassortative degree-degree correlations. In this respect, the recursive trees is a very special case of random recursive networks. Note that in equilibrium networks, dead ends (vertices of degree 1), abundant in random trees, also play a special role for degree-degree correlations (see Ref. [15]).

We have investigated the size dependence of Pearson's coefficient in the growing networks. We have found that the size effect is significant even for very large networks. We have shown that a growing network during its evolution may demonstrate an essential and even nonmonotonous variation of Pearson's coefficient. Due to this marked size dependence, it is hardly feasible to use this integral characteristic of correlations for quantitative comparison of different real-world networks. Instead, for this purpose, one has to use more informative characteristics, for example,  $\bar{k}_{nn}(k)$ . Of course, this quantity is also size dependent [16,17], but at least it provides a more comprehensive view of degree-degree correlations in networks.

## ACKNOWLEDGMENTS

This work was partially supported by PTDC projects No. FIS/71551/2006, No. FIS/108476/2008, and No. SAU-NEU/103904/2008, and also by SOCIALNETS EU project. The authors thank M. Barroso for his help with administering computing facilities used for numerical simulations.

## APPENDIX: EXCLUSION OF MULTIPLE CONNECTIONS

Let us explain that the exclusion of multiple connections will not change Pearson's coefficient in the infinite network limit. For brevity, let  $m$  be 2. First, we estimate the total number of multiple (double) edges in the recursive net of  $t$  vertices. The probability that a new vertex attaches two times to the same vertex  $i$  is

$$\left[ \frac{k_i + A}{t(4 + A)} \right]^2 \sim \frac{k_i^2}{t^2}. \quad (\text{A1})$$

Therefore, the full probability that this new vertex has two connections is on the order of

$$t \int^{t^{1/(\gamma-1)}} dk k^{-\gamma} \frac{k^2}{t^2} \sim \frac{t^{(3-\gamma)/(\gamma-1)}}{t}. \quad (\text{A2})$$

(Here, we used the cutoff  $k_{\text{cut}} \sim t^{1/(\gamma-1)}$ .) Integrating this number over  $t$ , we obtain the total number of multiple connections in the network,  $\sim t^{(3-\gamma)/(\gamma-1)}$ , which shows that multiple edges are noticeable only if  $\gamma < 3$ . We also estimate the average number of double connections  $\langle d \rangle(k)$  of a vertex of degree  $k$ . For a vertex born at time  $i$ , we can write

$$\frac{\partial \langle d_i \rangle}{\partial t} \sim \frac{\langle k_i \rangle^2}{t^2} \sim i^{-2/(\gamma-1)} t^{2/(\gamma-1)-2}. \quad (\text{A3})$$

(In a recursive network of  $t$  vertices, the average degree of a vertex born at time  $i$  is  $\langle k_i \rangle(t) \sim (i/t)^{-1/(\gamma-1)}$ ; see Refs. [10,14].) Consequently, on average, a vertex born at time  $i$  has

$$\langle d_i \rangle \sim i^{-2/(\gamma-1)} t^{2/(\gamma-1)-1} \quad (\text{A4})$$

double connections, and so

$$\langle d \rangle(k) \sim \frac{k^2}{t}. \quad (\text{A5})$$

Setting the cutoff at  $\sim t^{1/2}$  will ensure the absence of multiple connections in the range  $\gamma < 3$ . Repeating calculations in Sec. IV with this cutoff, we arrive at the asymptotics  $r \sim -t^{-(\gamma-2)/2}$  for  $m > 1$  and  $\gamma < 3$ , and so  $r(t \rightarrow \infty) = 0$  as with multiple edges.

- 
- [1] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, *Rev. Mod. Phys.* **80**, 1275 (2008).  
 [2] A. V. Goltsev, S. N. Dorogovtsev, and J. F. F. Mendes, *Phys. Rev. E* **78**, 051105 (2008).  
 [3] M. E. J. Newman, *Phys. Rev. Lett.* **89**, 208701 (2002).  
 [4] M. E. J. Newman, *SIAM Rev.* **45**, 167 (2003).  
 [5] M. Á. Serrano, M. Boguñá, R. Pastor-Satorras, and A. Vespignani, in *Large Scale Structure and Dynamics of Complex Networks: From Information Technology to Finance and Natural Science*, edited by G. Caldarelli and A. Vespignani (World Scientific, Singapore, 2007), p. 35.  
 [6] R. Pastor-Satorras, A. Vázquez, and A. Vespignani, *Phys. Rev. Lett.* **87**, 258701 (2001).  
 [7] A. Vázquez, R. Pastor-Satorras, and A. Vespignani, *Phys. Rev. E* **65**, 066130 (2002).  
 [8] S. Maslov and K. Sneppen, *Science* **296**, 910 (2002).  
 [9] P. L. Krapivsky and S. Redner, *Phys. Rev. E* **63**, 066123 (2001).  
 [10] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, *Phys. Rev. Lett.* **85**, 4633 (2000).  
 [11] P. L. Krapivsky, S. Redner, and F. Leyvraz, *Phys. Rev. Lett.* **85**, 4629 (2000).  
 [12] M. Boguñá, R. Pastor-Satorras, and A. Vespignani, *Phys. Rev. Lett.* **90**, 028701 (2003).  
 [13] A. Barrat and R. Pastor-Satorras, *Phys. Rev. E* **71**, 036127 (2005).  
 [14] S. N. Dorogovtsev and J. F. F. Mendes, *Phys. Rev. E* **63**, 056125 (2001).  
 [15] P. Bialas and A. K. Oleś, *Phys. Rev. E* **77**, 036124 (2008).  
 [16] S. Weber and M. Porto, *Phys. Rev. E* **76**, 046111 (2007).  
 [17] S. Weber, M.-T. Hütt, and M. Porto, *EPL* **82**, 28003 (2008).