

Intrinsic degree-correlations in the static model of scale-free networks

J.-S. Lee, K.-I. Goh, B. Kahng^a, and D. Kim

School of Physics and Center for Theoretical Physics, Seoul National University, Seoul 151-747, Korea

Received 31 October 2005

Published online 17 February 2006 – © EDP Sciences, Società Italiana di Fisica, Springer-Verlag 2006

Abstract. We calculate the mean neighboring degree function $\bar{k}_{\text{nn}}(k)$ and the mean clustering function $C(k)$ of vertices with degree k as a function of k in finite scale-free random networks through the static model. While both are independent of k when the degree exponent $\gamma \geq 3$, they show the crossover behavior for $2 < \gamma < 3$ from k -independent behavior for small k to k -dependent behavior for large k . The k -dependent behavior is analytically derived. Such a behavior arises from the prevention of self-loops and multiple edges between each pair of vertices. The analytic results are confirmed by numerical simulations. We also compare our results with those obtained from a growing network model, finding that they behave differently from each other.

PACS. 89.75.Da Systems obeying scaling laws – 89.75.Fb Structure and organization in complex systems – 05.65.+b Self-organized systems

1 Introduction

Structural properties of complex networks have drawn much attentions recently [1–3]. Degree, the number of edges connected to a given vertex, is a primary quantity to characterize the network structure. In many real-world networks, degrees are inhomogeneous and their distribution follows a power law $P_d(k) \sim k^{-\gamma}$. Such networks are called scale-free (SF) networks [4]. The degree-degree correlation is also important to characterize network structure. The correlation between two degrees of vertices connected via an edge is measured by the mean neighboring degree function $\bar{k}_{\text{nn}}(k)$, which is defined as the mean degree of neighboring vertices of vertices with degree k [5]. The correlation among three vertices centered at a vertex i is measured through the local clustering coefficient C_i , defined as $C_i = 2e_i/k_i(k_i - 1)$, where e_i is the number of connections among the k_i neighbors of a vertex i . k_i is the degree of the vertex i . The clustering function $C(k)$ is the averaged one of C_i over the vertices with degree k [6, 7].

While Barabási and Albert (BA) introduced a model to generate SF networks, the model is applied to growing systems where the number of vertices increases with time [4]. As an extension of the Erdős-Rényi (ER) model of random graph to SF networks [8], where the number of vertices in the system is fixed, Goh *et al.* introduced the so-called static model [9]. The term ‘static’ originates from the fact that the number of vertices N is fixed. The static model was followed by other similar-type models such as

the hidden variable model [10–12]. In the static model, each vertex i ($i = 1, \dots, N$) has a prescribed weight p_i summed to 1, which is given as

$$p_i = \frac{i^{-\mu}}{\sum_{j=1}^N j^{-\mu}} \approx \frac{1-\mu}{N^{1-\mu}} i^{-\mu}, \quad (1)$$

where the Zipf exponent μ is in the range $0 < \mu < 1$. To construct the network, in each time step, two vertices i and j are selected with the probability p_i and p_j , respectively. If $i = j$ or an edge connecting i and j already exists, do nothing, implying that self-loops and multiple edges are not allowed, respectively. This condition is called the fermionic constraint hereafter. Otherwise, an edge is added between i and j . This process is repeated NK times. The resulting network is a scale-free one with the degree exponent given as [9, 13]

$$\gamma = 1 + \frac{1}{\mu}. \quad (2)$$

Since a pair of vertices is selected with rate $2p_i p_j$, where the factor 2 comes from the two cases of (i, j) and (j, i) , one may think that there is no degree correlation, which is the case we can observe when $\gamma > 3$. However, when $2 < \gamma < 3$, due to the fermionic constraint, the degree-degree correlation arises intrinsically. In this case, the degree correlation occurs for vertices with large degree, while it is still absent for vertices with small degree. In this paper, we investigate such degree correlations in terms of the functions $\bar{k}_{\text{nn}}(k)$ and $C(k)$ in the static model and their crossover behavior in terms of system size N .

^a e-mail: kahng@phya.snu.ac.kr

Many SF networks in the real-world and artificial networks include degree correlations within them. For example, the mean neighboring degree function $\bar{k}_{\text{nn}}(k)$ behaves $\sim k^{-\nu}$ with $\nu > 0$ for the Internet [5] and the protein interaction network [14], while $\nu < 0$ for social networks such as the coauthorship network. The case with $\nu > 0$ ($\nu < 0$) is called disassortative (assortative) mixing [15]. When a network contains hierarchical and modular structure within it, it is suggested that the mean clustering function $C(k)$ behaves as $C(k) \sim k^{-\beta}$ for large k as observed in metabolic networks and the hierarchy model [6,16]. Occurrence of such degree correlations in real-world networks may be related to their own functional details. For example, the assortativity of the social network arises from the social relationship between bosses, while the disassortativity of the Internet comes from the network design to allow data packets flow efficiently. The three-degree correlation may be related to the control system in biological network such as the feed-back or feed-forward loop structure [17]. Such degree correlations in real world networks appear in the combination of those due to the fermionic constraint and their functionalities. On the other hand, the static model is frequently used to study various dynamical properties of complex networks. Therefore, the knowledge of the intrinsic degree correlations we study here would be helpful in understanding the degree-correlation a SF network has for functional activity. For the purpose, Catanzaro and Pastor-Satorras [18] studied the degree-correlation function $\bar{k}_{\text{nn}}(k)$ for the static model, but their study relies on numerics in the end. Here we present analytic solutions for $\bar{k}_{\text{nn}}(k)$ as well as clustering function $C(k)$. We mention that $\bar{k}_{\text{nn}}(k)$ for a related model was analyzed by Park and Newman [19] while $\bar{k}_{\text{nn}}(k)$ and $C(k)$ for the BA-type growing network models are studied by Barrat and Pastor-Satorras [20] using the rate equation approach [21]. On the other hand, it was desirable to introduce uncorrelated SF network as a null model to check the correctness of analytic solutions in various problems on SF networks. For the purpose, Boguñá et al. [22] and Catanzaro et al. [23] introduced a way to construct uncorrelated SF network by restricting degree of each vertex to be less than the cutoff value k_c , beyond which the intrinsic correlation arises in $2 < \gamma < 3$. The cutoff value they used scales as $\sim N^{1/2}$, independent of γ , which was based on the configuration model introduced by Molloy and Reed [24]. Such cutoff is also implicit in the model introduced by Chung and Lu [25]. However, we show that while the natural cutoff of the static model is $\sim N^{1/(\gamma-1)}$, the vertex correlations appear for degrees larger than a crossover value, $k_{c1} \sim N^{(\gamma-2)/(\gamma-1)}$, which is smaller than $N^{1/2}$ for $2 < \gamma < 3$. We mention that $k_{c1} \sim \text{const.}$ ($\sim N^{1/2}$) as $\gamma \rightarrow 2$ ($\gamma \rightarrow 3$) so that for $\gamma \rightarrow 2$ all the nodes have nontrivial vertex correlations and for $\gamma \rightarrow 3$, there are no correlations.

In Section 2, we derive the mean neighboring degree function $\bar{k}_{\text{nn}}(k)$ and the mean clustering function $C(k)$ analytically. Comparisons between the results of our analytic derivations and numerical simulations are given in Section 3. Section 4 summarizes our results.

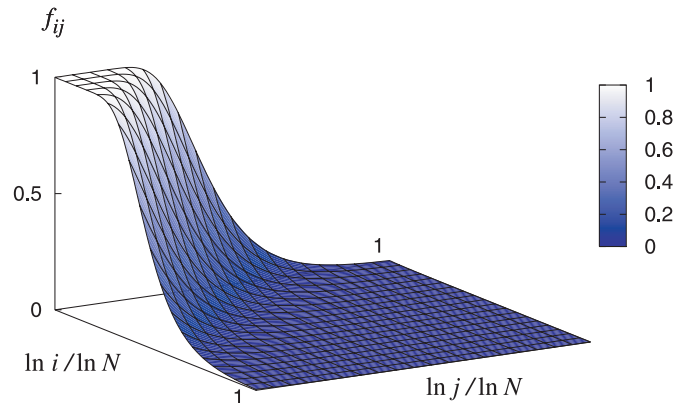


Fig. 1. The connection probability f_{ij} of an edge has two distinct regions where $f_{ij} \approx 1$ or $\approx 2KNp_i p_j$ due to the fermionic constraint when $2 < \gamma < 3$.

2 Analytic results

In the static model, the notion of the grand canonical ensemble is applied [13], where the number of edges is a fluctuating variable while keeping the SF nature of the degree distribution. Each pair of vertices (i, j) is connected independently with the probability f_{ij} , given by

$$f_{ij} = 1 - e^{-2KNp_i p_j}, \quad (3)$$

because the probability that vertices i and j ($i \neq j$) are not connected after NK trials is given by $(1 - 2p_i p_j)^{NK} \simeq e^{-2KNp_i p_j}$. That is, if we denote the adjacency matrix element by a_{ij} ($=0, 1$) then its ensemble average is given by f_{ij} ; i.e., $\langle a_{ij} \rangle = f_{ij}$, $\langle \dots \rangle$ denoting the grand canonical ensemble average. For $i = j$, $f_{ij} = 0$ because of the prevention of self-loops. Since $2KNp_i p_j \sim KN^{2\mu-1}/(ij)^\mu$ for finite K , when $0 < \mu < 1/2$, corresponding to the case $\gamma > 3$, $2KNp_i p_j$ is small in the thermodynamic limit, therefore,

$$f_{ij} \approx 2KNp_i p_j. \quad (4)$$

This is called the bosonic limit. On the other hand, when $1/2 < \mu < 1$, corresponding to the case $2 < \gamma < 3$, $2KNp_i p_j$ may diverge in the thermodynamic limit, therefore, f_{ij} is not necessarily of the form of equation (4), but it reduces to

$$f_{ij} \approx \begin{cases} 1 & \text{when } ij \ll N^{2-\frac{1}{\mu}} \\ 2KNp_i p_j & \text{when } ij \gg N^{2-\frac{1}{\mu}}. \end{cases} \quad (5)$$

This is the manifestation of the fermionic constraint, the prevention of multiple edges. Thus, for $2 < \gamma < 3$, f_{ij} has two distinct regions in the (i, j) plane as shown in Figure 1.

2.1 Degree and degree distribution:

The degree k_i of a vertex i is given in terms of the adjacency matrix as $k_i = \sum_j a_{ij}$. For completeness, we present

here known results for the mean degree $\langle k_i \rangle$ [13]. It is obtained through the formula $\langle k_i \rangle = \sum_{j \neq i} f_{ij}$ which can be evaluated by using its integral form as

$$\begin{aligned} \langle k_i \rangle &= \sum_{j \neq i} f_{ij} \approx \int_1^N dj f_{ij} \\ &= \frac{1}{\mu} a^{\frac{1}{\mu}} N^{1-\frac{1}{2\mu}} \int_{aN^{-\frac{1}{2}}}^{aN^{\mu-\frac{1}{2}}} dy \frac{1-e^{-xy}}{y^\gamma}, \end{aligned} \quad (6)$$

where $\gamma = 1 + \frac{1}{\mu}$, $x = aN^{\mu-\frac{1}{2}}i^{-\mu}$ and $y = aN^{\mu-\frac{1}{2}}j^{-\mu}$ with

$$a = \sqrt{2K(1-\mu)^2}. \quad (7)$$

The integral in equation (6) denoted as $I(x)$ is evaluated as

$$\begin{aligned} I(x) &= \int_{aN^{-\frac{1}{2}}}^{aN^{\mu-\frac{1}{2}}} dy \frac{1-e^{-xy}}{y^\gamma} \\ &\approx \begin{cases} \frac{a^{2-\gamma} N^{\frac{\gamma-2}{2}}}{\gamma-2} (1-N^{\mu-1})x & \text{for } x < 1/aN^{\mu-\frac{1}{2}} \\ \frac{a^{2-\gamma} N^{\frac{\gamma-2}{2}}}{\gamma-2} x + q_0(\gamma)x^{\gamma-1} & \text{for } x > 1/aN^{\mu-\frac{1}{2}}, \end{cases} \end{aligned} \quad (8)$$

with $q_0(\gamma) \equiv \int_0^\infty dr(1-e^{-r}-r)/r^\gamma$, which is a negative constant. Therefore, we obtain that

$$\langle k_i \rangle \approx 2K(1-\mu) \left(\frac{N}{i}\right)^\mu + \mathcal{A}, \quad (9)$$

where \mathcal{A} is the correction, of which the leading term is

$$\mathcal{A} \approx \begin{cases} -2K(1-\mu)N^{2\mu-1}/i^\mu & \text{for } i > a^{\frac{2}{\mu}}N^{2-\frac{1}{\mu}} \\ (2K(1-\mu)^2)^{\frac{1}{\mu}}N^{2-\frac{1}{\mu}}q_0(\gamma)/(i\mu) & \text{for } i < a^{\frac{2}{\mu}}N^{2-\frac{1}{\mu}}. \end{cases} \quad (10)$$

This is negligible compared with the first term on the right hand side of equation (9) in the thermodynamic limit, $N \rightarrow \infty$. The average degree is then

$$\bar{k} \equiv \frac{2\langle L \rangle}{N} = \frac{1}{N} \sum_i \langle k_i \rangle = 2K, \quad (11)$$

where $\langle L \rangle$ is the mean number of edges in the grand canonical ensemble. From equation (9), one can easily obtain that the degree exponent is related to the Zipf exponent μ as $\gamma = 1 + 1/\mu$ given in equation (2)

2.2 Mean neighboring degree function $\bar{k}_{\text{nn}}(\mathbf{k})$:

We pay attention to the case $2 < \gamma < 3$. To evaluate the mean neighboring degree function $\bar{k}_{\text{nn}}(k)$, we first calculate the mean neighboring degree in terms of i , i.e., $\bar{k}_{\text{nn}}(i)$ and convert it to $\bar{k}_{\text{nn}}(k)$ by using the relation of $\langle k_i \rangle$ versus i . To proceed, we use the expression,

$$\begin{aligned} \bar{k}_{\text{nn}}(i) &= \left\langle \frac{\sum_{j \in \text{nn of } i} k_j}{k_i} \right\rangle = \left\langle \frac{\sum_{j,k} a_{ij} a_{jk}}{\sum_j a_{ij}} \right\rangle \approx \frac{\langle \sum_{j,k} a_{ij} a_{jk} \rangle}{\langle k_i \rangle}, \end{aligned} \quad (12)$$

where the ensemble average is applied to the numerator and the denominator separately. Its validity is checked numerically, which is shown in Section 3. The denominator was already derived, and the numerator is evaluated as follows:

$$\begin{aligned} \left\langle \sum_{j,k} a_{ij} a_{jk} \right\rangle &= \sum_{j,k \neq i} f_{ij} f_{jk} + \sum_j f_{ij} \approx \int_1^\infty dj \int_1^\infty dk f_{ij} f_{jk} + \langle k_i \rangle, \end{aligned} \quad (13)$$

where $a_{ij}^2 = a_{ij}$ is used and the double sum is approximated by the double integral. The validity of the transformation from the discrete double sum to the double integration is discussed in Appendix A where $2 < \gamma < 3$. Such an approximation introduces at most an $O(1)$ factor on the amplitude of the leading order terms for large N , as will be mentioned below. Applying the similar method used in equation (6), we evaluate the integration in equation (13) as

$$\begin{aligned} \int_1^\infty dj \int_1^\infty dk f_{ij} f_{jk} &= \frac{a^{\frac{2}{\mu}} N^{(2-\frac{1}{\mu})}}{\mu^2} \\ &\times \int_{aN^{-\frac{1}{2}}}^{aN^{\mu-\frac{1}{2}}} dy \frac{1-e^{-xy}}{y^\gamma} \int_{aN^{-\frac{1}{2}}}^{aN^{\mu-\frac{1}{2}}} dz \frac{1-e^{-yz}}{z^\gamma}, \end{aligned} \quad (14)$$

where $x \equiv aN^{\mu-1/2}i^{-\mu}$ is in the range $aN^{-1/2} < x < aN^{\mu-1/2}$ (see the definition of a in Eq. (7)). The last part of the integral in equation (14) is $I(y)$ defined in equation (8). Therefore, we substitute the leading term of equations (8) into (14) and obtain

$$\begin{aligned} \int_1^\infty dj \int_1^\infty dk f_{ij} f_{jk} &\approx \frac{a^{1+\frac{1}{\mu}} N^{\frac{3}{2}-\frac{1}{2\mu}}}{\mu(1-\mu)} x^{\gamma-2} \int_{aN^{-\frac{1}{2}}}^{aN^{\mu-\frac{1}{2}}} dq \frac{1-e^{-q}}{q^{\gamma-1}}, \end{aligned} \quad (15)$$

in which we change the variable of integration as $q = xy$. The integral in the right hand side of equation (15) is evaluated in three parts as

$$\begin{aligned} \int_{aN^{-\frac{1}{2}}}^{aN^{\mu-\frac{1}{2}}} dq \frac{1-e^{-q}}{q^{\gamma-1}} &= \int_0^\infty dq \frac{1-e^{-q}}{q^{\gamma-1}} \\ &- \int_0^{aN^{-\frac{1}{2}}} dq \frac{1-e^{-q}}{q^{\gamma-1}} - \int_{aN^{\mu-\frac{1}{2}}}^\infty dq \frac{1-e^{-q}}{q^{\gamma-1}}. \end{aligned} \quad (16)$$

The first term is denoted as

$$q_1(\gamma) \equiv \int_0^\infty dq \frac{1-e^{-q}}{q^{\gamma-1}}, \quad (17)$$

which is a positive constant for $2 < \gamma < 3$. The second term is, since $aN^{-1/2} \ll 1$,

$$\int_0^{aN^{-\frac{1}{2}}} dq \frac{1-e^{-q}}{q^{\gamma-1}} \approx \frac{a^{3-\gamma} N^{-(3-\gamma)/2}}{3-\gamma} x^{3-\gamma}. \quad (18)$$

The last term is calculated as, when $x \ll 1/aN^{\mu-\frac{1}{2}}$ ($i \gg a^{\frac{2}{\mu}}N^{2-\frac{1}{\mu}}$),

$$\int_{axN^{\mu-\frac{1}{2}}}^{\infty} dq \frac{1-e^{-q}}{q^{\gamma-1}} = \int_0^{\infty} dq \frac{1-e^{-q}}{q^{\gamma-1}} - \int_0^{axN^{\mu-\frac{1}{2}}} dq \frac{1-e^{-q}}{q^{\gamma-1}} \\ \approx q_1(\gamma) - \frac{a^{3-\gamma}N^{(3-\gamma)(\mu-\frac{1}{2})}}{3-\gamma}x^{3-\gamma}, \quad (19)$$

and, when $x \gg 1/aN^{\mu-\frac{1}{2}}$ ($i \ll a^{\frac{2}{\mu}}N^{2-\frac{1}{\mu}}$),

$$\int_{axN^{\mu-\frac{1}{2}}}^{\infty} dq \frac{1-e^{-q}}{q^{\gamma-1}} \approx \frac{a^{1-\frac{1}{\mu}}N^{(1-\frac{1}{\mu})(\mu-\frac{1}{2})}}{\frac{1}{\mu}-1}x^{2-\gamma}. \quad (20)$$

Combining all the contributions, when $x \ll 1/aN^{\mu-\frac{1}{2}}$, the second term on the right hand side of equation (16) becomes the leading order term, while when $x \gg 1/aN^{\mu-\frac{1}{2}}$, the first one does. Thus,

$$\int_1^{\infty} dj \int_1^{\infty} dk f_{ij}f_{jk} \approx \begin{cases} \frac{a^4N^{3\mu-1}}{(1-\mu)(2\mu-1)}i^{-\mu} & \text{for } i > a^{\frac{2}{\mu}}N^{2-\frac{1}{\mu}} \\ q_1(\gamma)\frac{a^{\frac{2}{\mu}}N^{3-\mu-\frac{1}{\mu}}}{\mu(1-\mu)}i^{-1+\mu} & \text{for } i < a^{\frac{2}{\mu}}N^{2-\frac{1}{\mu}}. \end{cases} \quad (21)$$

The second term $\langle k_i \rangle$ on the right-hand-side of equation (13) can be neglected compared with equation (21) for all range of $2 < \gamma < 3$ and i . Therefore,

$$\bar{k}_{\text{nn}}(i) \approx \begin{cases} \frac{a^2}{2\mu-1}N^{2\mu-1} & \text{when } i > a^{\frac{2}{\mu}}N^{2-\frac{1}{\mu}} \\ q_1(\gamma)a^{\frac{2}{\mu}-2}N^{3-2\mu-\frac{1}{\mu}}i^{2\mu-1}/\mu & \text{when } i < a^{\frac{2}{\mu}}N^{2-\frac{1}{\mu}}, \end{cases} \quad (22)$$

and using equation (9) for $k = \langle k_i \rangle$,

$$\bar{k}_{\text{nn}}(k) \approx \begin{cases} \frac{a^2}{2\mu-1}N^{2\mu-1} & \text{when } k < N^{1-\mu} \\ q_1(\gamma)2K(1-\mu)^{\frac{1}{\mu}}N^{2-\frac{1}{\mu}}k^{-2+\frac{1}{\mu}}/\mu & \text{when } k > N^{1-\mu}. \end{cases} \quad (23)$$

Here we note that the coefficient of $N^{2\mu-1}$ when $i > a^{\frac{2}{\mu}}N^{2-\frac{1}{\mu}}$ (or when $k < N^{1-\mu}$) is not exact but is in between $a^2/(2\mu-1)$ and $2\mu a^2/(2\mu-1)$ as explained in Appendix A (see Eq. (44)). In terms of the degree exponent γ we rewrite equation (23) as

$$\bar{k}_{\text{nn}}(k) \sim \begin{cases} N^{(3-\gamma)/(\gamma-1)} & \text{when } k > k_{c1} \\ N^{3-\gamma}k^{-(3-\gamma)} & \text{when } k < k_{c1}, \end{cases} \quad (24)$$

where the crossover degree k_{c1} scales as $k_{c1} \sim N^{(\gamma-2)/(\gamma-1)}$.

2.3 Clustering function $C(\mathbf{k})$:

The clustering function $C(k)$ is the mean of the local clustering coefficient C_i over the vertices with degree k . To calculate $C(k)$, we first calculate C_i and convert it to $C(k)$ by

using the relation equation (9). As we introduced before, C_i is defined as $C_i = 2e_i/k_i(k_i-1)$, where e_i is the number of connections among the k_i neighbors. In the grand canonical ensemble, C_i is calculated as

$$C_i = \left\langle \frac{e_i}{k_i(k_i-1)/2} \right\rangle. \quad (25)$$

However, we approximate it as

$$C_i \approx \frac{\langle e_i \rangle}{\langle k_i(k_i-1)/2 \rangle}, \quad (26)$$

which enables us to calculate it analytically. The validity of this approximation is checked numerically in Section 3. We evaluate the denominator and numerator separately.

The denominator is evaluated as

$$\left\langle \frac{k_i(k_i-1)}{2} \right\rangle = \frac{1}{2} \sum_{j \neq k (\neq i)} f_{ij}f_{ik} \\ \approx \frac{1}{2} \sum_{j,k (\neq i)} f_{ij}f_{ik} \approx 2K^2(1-\mu)^2N^{2\mu}/i^{2\mu}. \quad (27)$$

The numerator is evaluated as

$$\langle e_i \rangle = \frac{1}{2} \sum_{j \neq k (\neq i)} f_{ij}f_{jk}f_{ki} \\ \approx \frac{1}{2} \sum_{j,k} f_{ij}f_{jk}f_{ki} \approx \frac{1}{2} \int_1^N dj \int_1^N dk f_{ij}f_{jk}f_{ki} = \frac{a^{\frac{2}{\mu}}N^{2-\frac{1}{\mu}}}{2\mu^2} \\ \times \int_{aN^{-\frac{1}{2}}}^{aN^{\mu-\frac{1}{2}}} dy \int_{aN^{-\frac{1}{2}}}^{aN^{\mu-\frac{1}{2}}} dz \frac{(1-e^{-xy})(1-e^{-yz})(1-e^{-zx})}{y^\gamma z^\gamma}. \quad (28)$$

Possible errors involved in using the integral form for the double sum is discussed in Appendix A and will be mentioned below. The evaluation of the integrals of equation (28) is carried out depending on the magnitude of x . When $x \gg 1/aN^{\mu-\frac{1}{2}}$, we obtain

$$\int_{aN^{-\frac{1}{2}}}^{aN^{\mu-\frac{1}{2}}} dz \frac{(1-e^{-yz})(1-e^{-zx})}{z^\gamma} \\ \approx \begin{cases} q_0(\gamma) (x^{\gamma-1} - (x+y)^{\gamma-1}) & \text{when } y < 1/aN^{\mu-\frac{1}{2}} \\ q_0(\gamma) (x^{\gamma-1} + y^{\gamma-1} - (x+y)^{\gamma-1}) & \text{when } y > 1/aN^{\mu-\frac{1}{2}}. \end{cases} \quad (29)$$

Thus $\langle e_i \rangle$ is written as

$$\langle e_i \rangle = -\frac{a^{\frac{2}{\mu}}N^{2-\frac{1}{\mu}}}{2\mu^2}q_0(\gamma)(\mathcal{B} + \mathcal{C}), \quad (30)$$

where \mathcal{B} and \mathcal{C} are expressed in the integral forms as

$$\begin{aligned} \mathcal{B} &= \int_{aN^{-\frac{1}{2}}}^{aN^{\mu-\frac{1}{2}}} dy \frac{1-e^{-xy}}{y^\gamma} [(x+y)^{\gamma-1} - x^{\gamma-1} - y^{\gamma-1}] \\ &= \int_{aN^{-\frac{1}{2}/x}^1 dq \frac{(\gamma-1)(1-e^{-x^2q})}{q^{\gamma-1}} \\ &\quad - \int_{aN^{-\frac{1}{2}/x}^1 dq \frac{1-e^{-x^2q}}{q} \\ &\quad + \int_1^{aN^{\mu-\frac{1}{2}}/x} dq \left(\frac{\gamma-1}{q^2} - \frac{1}{q^\gamma} \right) (1-e^{-x^2q}) \end{aligned} \quad (31)$$

and

$$\mathcal{C} = -q_0(\gamma) \int_{aN^{-\frac{1}{2}}}^{1/aN^{\mu-\frac{1}{2}}} dy \frac{1-e^{-xy}}{y^\gamma} y^{\gamma-1}. \quad (32)$$

Even in the region of $x > 1/aN^{\mu-\frac{1}{2}}$, the leading term is determined depending on the magnitude of x . When $x > 1 > 1/aN^{\mu-\frac{1}{2}}$, the first term of the integral \mathcal{B} is the most dominant one compared with the other terms of \mathcal{B} and \mathcal{C} , which is evaluated as $\approx(\gamma-1)q_1(\gamma)x^{2(\gamma-2)}$. When $1 > x > 1/aN^{\mu-\frac{1}{2}}$, however, the third term is most dominant and evaluated as $\approx 2(\gamma-1)\ln(1/x)x^2$. Therefore, the numerator is evaluated as

$$\langle e_i \rangle \approx \begin{cases} \frac{a^\mu N^{2-\frac{1}{\mu}}}{2\mu^2} [-q_0(\gamma)q_1(\gamma)(\gamma-1)x^{2(\gamma-2)}], & \text{when } x > 1 \\ \frac{a^\mu N^{2-\frac{1}{\mu}}}{2\mu^2} 2(\gamma-1)\ln(1/x)x^2, & \text{when } 1 > x > 1/aN^{\mu-\frac{1}{2}}. \end{cases} \quad (33)$$

Thus, we get C_i in the region of $x > 1/aN^{\mu-\frac{1}{2}}$ to the leading order as

$$C_i \approx \begin{cases} N^{1-\frac{1}{\mu}} \ln(i^\mu/N^{\mu-\frac{1}{2}}), & \text{when } N^{1-\frac{1}{2\mu}} < i < N^{2-\frac{1}{\mu}} \\ N^{5-4\mu-\frac{2}{\mu}} i^{2(2\mu-1)}, & \text{when } i < N^{1-\frac{1}{2\mu}}. \end{cases} \quad (34)$$

Equivalently,

$$C(k) \approx \begin{cases} N^{1-\frac{1}{\mu}} \ln(N^{\frac{1}{2}}/k), & \text{when } N^{1-\mu} < k < N^{\frac{1}{2}} \\ N^{3-\frac{2}{\mu}} k^{-2(3-\gamma)}, & \text{when } N^{\frac{1}{2}} < k. \end{cases} \quad (35)$$

Let us consider the case of $x \ll 1/aN^{\mu-\frac{1}{2}}$ ($i \gg N^{2-\frac{1}{\mu}}$). In this case,

$$\int_{aN^{-\frac{1}{2}}}^{aN^{\mu-\frac{1}{2}}} dz \frac{(1-e^{-yz})(1-e^{-zx})}{z^\gamma} \approx -q_0(\gamma)((x+y)^{\gamma-1} - y^{\gamma-1}), \quad (36)$$

for $y > 1/aN^{\mu-\frac{1}{2}}$ and is almost negligible for $y < 1/aN^{\mu-\frac{1}{2}}$. Thus

$$\begin{aligned} \langle e_i \rangle &= -\frac{q_0(\gamma)a^\mu N^{2-\frac{1}{\mu}}}{2\mu^2} \int_{1/aN^{\mu-\frac{1}{2}}}^{aN^{\mu-\frac{1}{2}}} [(x+y)^{\gamma-1} - y^{\gamma-1}] \frac{1-e^{-xy}}{y^\gamma} \\ &\approx -\frac{q_0(\gamma)a^\mu(\gamma-1)N^{2-\frac{1}{\mu}}}{2\mu^2} x^2 \ln(a^2 N^{2\mu-1}). \end{aligned} \quad (37)$$

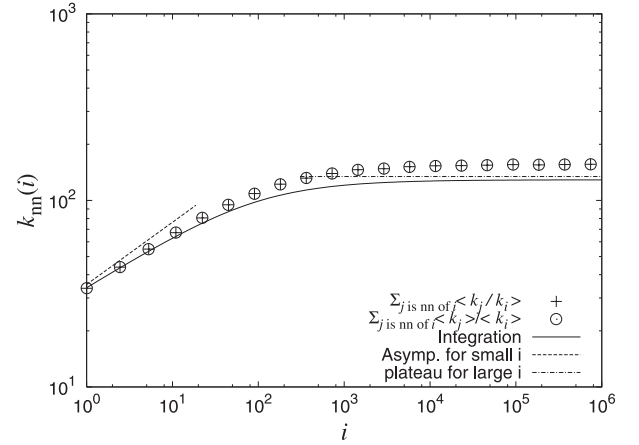


Fig. 2. Plot of $\bar{k}_{\text{nn}}(i)$ versus i . To check the validity of the approximation equation (12), we plot $\sum_{j \in \text{nn of } i} \langle \frac{k_j}{k_i} \rangle$ (+) and $\frac{\sum_{j \in \text{nn of } i} \langle k_j \rangle}{\langle k_i \rangle} = \frac{\sum_{j,k} f_{ij} f_{jk}}{\sum_j f_{ij}} + 1$ (o) for $N = 10^6$. We can see that the approximation is valid. We compare them with $\frac{\int \int dk f_{ij} f_{jk}}{\int dj f_{ij}} + 1$ (solid line). They agree with each other for small i , however, it is in disagreement in the plateau region as expected in Appendix A. We also plot with the first leading term presented in the text in the asymptotic regions with dot-dashed line and dashed line, respectively.

Therefore we get when $x \ll 1/aN^{\mu-\frac{1}{2}}$ (i.e., $i \gg a^\mu N^{2-\frac{1}{\mu}}$, or $k \ll N^{1-\mu}$),

$$C_i = C(k) \sim AN^{1-\frac{1}{\mu}} \ln N, \quad (38)$$

with $A = -q_0(\gamma)a^{\frac{2}{\mu}-2}(\gamma-2)^2(3-\gamma)$, when $k < N^{1-\mu}$. In Appendix A, the error introduced in equation (28) is estimated and is found not to change equation (35). For equation (38), however, we find $C(k) \sim AN^{1-\frac{1}{\mu}}(\ln N + D)$ with an undetermined constant D . Equations (35) and (38) are the main results of this subsection.

3 Numerical simulations

We now discuss numerical check of the analytic results derived in Section 2. For the purpose, the static model network is generated with $K = 2$ and $\mu = 2/3$ ($\gamma = 2.5$) and with varying system size N . All data below are averaged over 10^4 network configurations. For the case of $\bar{k}_{\text{nn}}(k)$, we first check the approximation, equation (12). To proceed, we measure $\sum_{j \in \text{nn of } i} \langle k_j/k_i \rangle$ and $\sum_{j \in \text{nn of } i} \langle k_j \rangle / \langle k_i \rangle$ separately in Figure 2, finding that the data overlap and the approximation is valid. Next we directly enumerate the function, $\int_1^N dj \int_1^N dk f_{ij} f_{jk} / \int_1^N dj f_{ij} + 1$ (solid line) and compare it with the evaluation (dashed line) within leading order, equation (23). The extra term of ‘1’ comes from the 2nd term $\langle k_i \rangle$ of equation (13). For small i , the two lines seem to be consistent, however, for large i , they somewhat deviate in the intermediate region of i . However, we confirm that our analytic solution is valid within leading order by the finite size scaling plot. In Figure 3, we

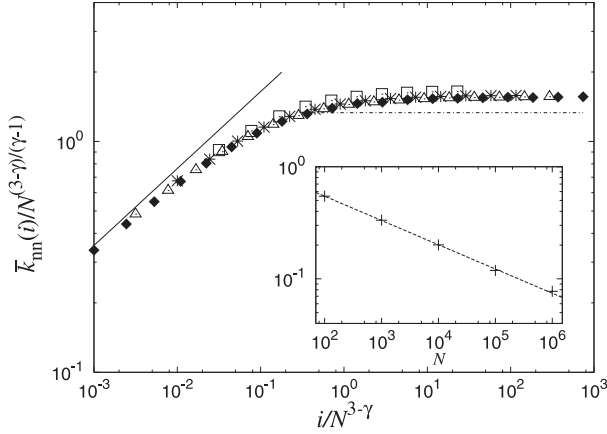


Fig. 3. Plot of size-dependent behavior of $\bar{k}_{nn}(i)$. Data of different network sizes $N = 10^3$ (\square), 10^4 ($*$), 10^5 (\triangle) and 10^6 (\blacklozenge) collapse into a single curve in the scaling plot. Inset: plot of the difference between the leading order analytic expression and the simulation value of $i = 1$, divided by the simulation value. As N increases, the relative difference decreases, showing that the analytic solution converges to the numerical result.

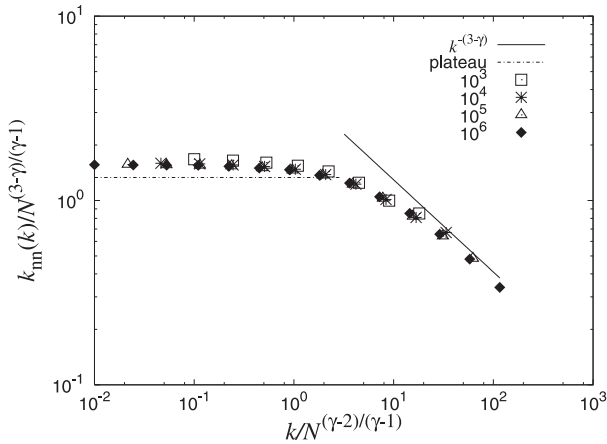


Fig. 4. Plot of $\bar{k}_{nn}(k)$ versus k for different system sizes $N = 10^3$ (\square), 10^4 ($*$), 10^5 (\triangle) and 10^6 (\blacklozenge). Data for different system sizes collapse in the scaling plot. Solid and dot-dashed lines indicate the analytic results of leading order for large and small k , respectively.

plot $\bar{k}_{nn}(i)$ for different $N = 10^3, 10^4, 10^5$ and 10^6 finding that the data collapse into a single curve by the rescalings of $i \rightarrow i/N^{2-1/\mu}$ and $\bar{k}_{nn}(i) \rightarrow \bar{k}_{nn}(i)/N^{2\mu-1}$. Moreover, we find as we increase the system size that the numerical simulation data approaches our analytic solution for small i (inset of Fig. 3). We also check the behavior of $\bar{k}_{nn}(k)$ numerically. Under the rescaling of $k \rightarrow k/N^{1-\mu}$ and $\bar{k}_{nn}(k) \rightarrow \bar{k}_{nn}(k)/N^{2\mu-1}$, the data for different system sizes collapse well, confirming the validity of our analytic result.

Next, the local clustering coefficient function C_i is measured. We first check the approximation introduced in equation (26) in Figure 5, finding that they overlap each other except for large i . This discrepancy originates from the fact that the vertices with large i are mostly those

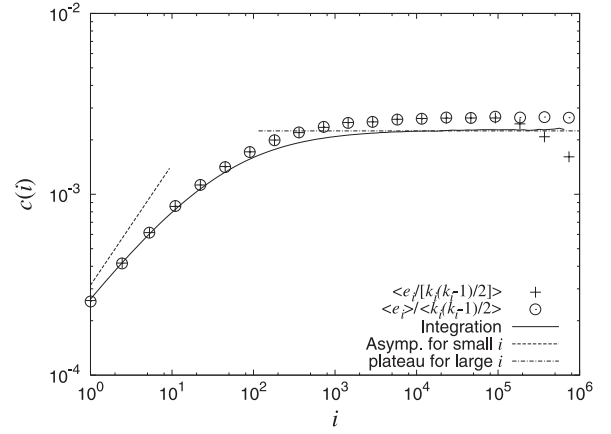


Fig. 5. Plot of C_i . We can see that the approximation, $\langle \frac{2e_i}{k_i(k_i-1)} \rangle \approx \frac{2\langle e_i \rangle}{\langle k_i(k_i-1) \rangle}$ is valid for the large i limit. The dot-dashed line indicates the analytic result. In the plateau region, the discrepancy between the analytic and the numerical results decrease as system size N increases. The dashed line indicates the analytic results, equations (34) and (38).

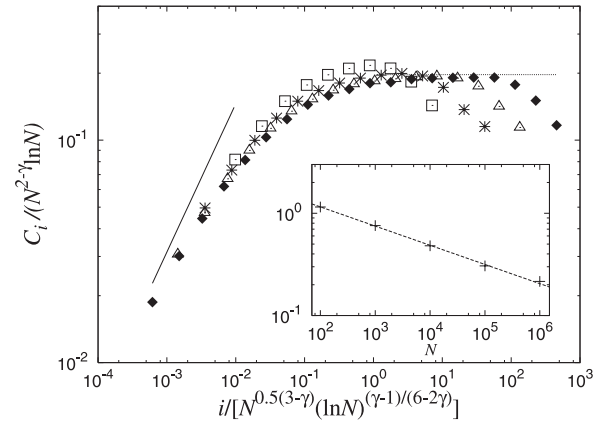
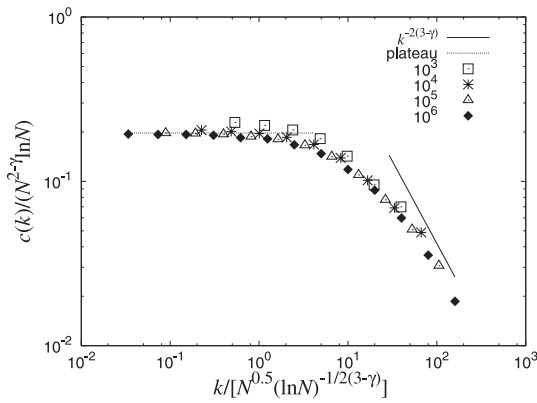


Fig. 6. Size-dependence of local clustering coefficient C_i . Data of various network sizes $N = 10^3$ (\square), 10^4 ($*$), 10^5 (\triangle) and 10^6 (\blacklozenge) are collapsed in the rescaling plot. Inset: plot of the difference between the analytic solution within the leading order and the simulation value for $i = 1$, divided by the simulation value as a function of N . The decreasing behavior with increasing N indicates that the analytic solution is asymptotically valid.

located at dangling ends with degree 1. Thus, the formation of triangles or wedge shapes is rare and their numbers fluctuate highly. Next, we also check the validity of the approximation from the discrete summation and the continuous integration, equation (28). For small i , the approximation is reasonably valid as shown in Figure 5, which can be expected in Appendix A. However, for large i in the flat region, the approximation shows some discrepancy, but it is likely that the discrepancy decreases as system size N increases. To check the size-dependent behavior of C_i , we plot C_i versus i with rescalings of $C_i \rightarrow C_i/N^{1-\frac{1}{\mu}} \ln N$ and $i \rightarrow i/[N^{(4\mu-4+\frac{1}{\mu})/(4\mu-2)} \ln^{1/(4\mu-2)} N]$ for different system sizes $N = 10^3, 10^4, 10^5$ and 10^6 . We find that the data collapse reasonably well as shown in Figure 6.

Table 1. Degree and system-size dependence of $\bar{k}_{\text{nn}}(i)$, $\bar{k}_{\text{nn}}(k)$, C_i and $C(k)$.

Range	$\bar{k}_{\text{nn}}(i)$	$\bar{k}_{\text{nn}}(k)$	C_i	$C(k)$
$2 < \gamma < 3$				
$i > N^{3-\gamma}$ $k < N^{\frac{\gamma-2}{\gamma-1}}$		$\sim N^{\frac{3-\gamma}{\gamma-1}}$		$\sim N^{-(\gamma-2)} \ln N$
$N^{3-\gamma} > i > N^{\frac{1}{2}(3-\gamma)}$ $N^{\frac{\gamma-2}{\gamma-1}} < k < N^{\frac{1}{2}}$	$\sim N^{\frac{(3-\gamma)(\gamma-2)}{\gamma-1}} i^{\frac{3-\gamma}{\gamma-1}}$	$\sim N^{3-\gamma} k^{-(3-\gamma)}$	$\sim N^{-(\gamma-2)} \ln \left(\frac{i^{1/(\gamma-1)}}{N^{(3-\gamma)/(2\gamma-2)}} \right)$	$\sim N^{-(\gamma-2)} \ln \left(\frac{N^{1/2}}{k} \right)$
$N^{\frac{1}{2}(3-\gamma)} > i$ $N^{\frac{1}{2}} < k$			$\sim N^{-\frac{2\gamma^2+9\gamma+11}{\gamma-1}} i^{\frac{2(3-\gamma)}{\gamma-1}}$	$\sim N^{5-2\gamma} k^{-2(3-\gamma)}$
$\gamma = 3$				
whole range		$\sim \ln N$		$\sim (\ln N)^2 / N$
$\gamma > 3$				
whole range		$\sim O(1)$		$\sim 1/N$


Fig. 7. Plot of $C(k)$ for different system size $N = 10^3$ (\square), 10^4 ($*$), 10^5 (\triangle) and 10^6 (\blacklozenge). The data are well collapsed in the rescaling plot.

And we also check the behavior of $C(k)$. By rescaling of $C(k) \rightarrow C(k)/N^{1-\frac{1}{\mu}} \ln N$ and $k \rightarrow k \ln^{1/2(2-\frac{1}{\mu})}/N^{1/2}$, the data of $C(k)$ for different system sizes also collapse into a single curve reasonably well as shown in Figure 7. Thus, our numerical simulation results show that, although several approximations are involved in deriving the analytic results of Section 2, they are valid to the leading orders in N as $N \rightarrow \infty$.

4 Conclusions and discussion

We have studied analytically the mean neighboring degree function $\bar{k}_{\text{nn}}(k)$ and the clustering function $C(k)$ in the static model for the case of $2 < \gamma < 3$ and checked the results by numerical simulations. Due to the prevention of self-loop and multiple edges, there occur intrinsic degree correlations, which appear for $2 < \gamma < 3$ in the k -dependent form of $k_{\text{nn}}(k)$ and $C(k)$ for large k . Our results are summarized in Table 1 together with those for the case of $\gamma \geq 3$. It would be interesting to compare our results with those obtained in the generalized BA-type growth model [20]. In this model, $\bar{k}_{\text{nn}}(k) \sim N^{(3-\gamma)/(\gamma-1)} k^{-(3-\gamma)}$ when $\gamma < 3$, $\sim \ln N$ when $\gamma = 3$, and $\sim \ln k$ when $\gamma > 3$. On the other hand, $C(k) \sim N^{(4-2\gamma)/(\gamma-1)} k^{-(3-\gamma)}$ for

$k > (\ln N)^{1/(3-\gamma)}$ and $\sim (\ln N) N^{(4-2\gamma)/(\gamma-1)} k^{-2(3-\gamma)}$ for $k < (\ln N)^{1/(3-\gamma)}$ when $\gamma < 3$, $\sim (\ln N)^2 / N$ when $\gamma = 3$ and $\sim N^{-1} k^{\gamma-3}$ in the range $k < N^{1/(\gamma-3)}$ when $\gamma > 3$. Therefore, it appears that the degree correlation functions $\bar{k}_{\text{nn}}(k)$ and $C(k)$ behave differently for the cases of the static model and the BA-type growth model.

This work is supported by the KRF Grant No. R14-2002-059-010000-0 in the ABRL program funded by the Korean government MOEHRD.

Appendix A: Transformation from discrete summation to continuous integration

In several parts of this paper, we use the transformation from the discrete summation to the continuous integration such as

$$\sum_{j,k=1}^N F(i,j,k) \approx \int_1^N dj \int_1^N dk F(i,j,k). \quad (39)$$

Here we discuss its validity. For a monotone decreasing function $F(x)$, one has the well known relation:

$$\int_1^N dx F(x) + F(N) \leq \sum_{n=1}^N F(n) \leq \int_1^N dx F(x) + F(1). \quad (40)$$

When $F(i,j,k)$ is positive, monotonously decreasing and bounded in both j and k , we can apply equation (40) twice to obtain the error in equation (39) as

$$\begin{aligned} \sum_k F(i,N,k) + \sum_j F(i,j,N) - F(i,N,N) &\leq \sum_{j,k} F(i,j,k) \\ &- \int_1^N dj \int_1^N dk F(i,j,k) \leq \sum_k F(i,1,k) \\ &+ \sum_j F(i,j,1) - F(i,1,1). \end{aligned} \quad (41)$$

Thus equation (39) is valid when the ‘‘surface terms’’ in equation (41) are negligible compared with the ‘‘bulk

term", $\int_1^N dj \int_1^N dk F(i, j, k)$. When we consider $\bar{k}_{nn}(i)$, $F(i, j, k)$ is given as $f_{ij}f_{jk}$ and one of the surface terms that require special attention is

$$\sum_k F(i, 1, k) = \sum_k f_{i1}f_{1k} \approx \frac{a^4 N^{3\mu-1}}{1-\mu} i^{-\mu}, \quad (42)$$

with $a = \sqrt{2K(1-\mu)^2}$. It turns out that this surface term is of the same order as the bulk term when $i > a^{\frac{2}{\mu}} N^{2-\frac{1}{\mu}}$. Other surface terms are, however, negligible. Thus, the contribution of the surface term to $\langle k_{nn} \rangle(i)$ is $\sim a^2 N^{2\mu-1}$. Then equation (22) for the case $i > a^{\frac{2}{\mu}} N^{2-\frac{1}{\mu}}$ has to be changed as

$$\frac{\int_1^N dj \int_1^N dk f_{ij}f_{jk}}{\langle k_i \rangle} \leq \bar{k}_{nn}(i) \leq \frac{\int_1^N dj \int_1^N dk f_{ij}f_{jk}}{\langle k_i \rangle} + \frac{\sum_k f(i, 1, k)}{\langle k_i \rangle}. \quad (43)$$

This leads to

$$\frac{a^2 N^{2\mu-1}}{2\mu-1} \leq \bar{k}_{nn}(i) \leq \frac{a^2 N^{2\mu-1}}{2\mu-1} \cdot 2\mu \quad (44)$$

to the leading order in N . Thus, the leading order of $\langle k_{nn} \rangle(i)$ is given only in the form of the bounds when $i > a^{\frac{2}{\mu}} N^{2-\frac{1}{\mu}}$.

When C_i is considered, $F(i, j, k) = f_{ij}f_{jk}f_{ki}$. The most relevant terms are $\sum_k F(i, 1, k)$ and $\sum_j F(i, j, 1)$:

$$\begin{aligned} \sum_k F(i, 1, k) &= \sum_j F(i, j, 1) \\ &= \sum_k f_{i1}f_{1k}f_{ki} \approx -q_0(\gamma)(\gamma-1)a^{\frac{2}{\mu}} N^{2-\frac{1}{\mu}} x^2/\mu, \end{aligned} \quad (45)$$

when $i > a^{\frac{2}{\mu}} N^{2-\frac{1}{\mu}}$. These are of the same order of magnitude as the bulk term in equation (33) up to the $\ln N$ factor. Other boundary terms are smaller in order of magnitude compared with these terms. Thus when $i > a^{\frac{2}{\mu}} N^{2-\frac{1}{\mu}}$, C_i is bounded as

$$\begin{aligned} -q_0(\gamma)a^{\frac{2}{\mu}-2}(\gamma-1)(\gamma-2)^2 N^{1-\frac{1}{\mu}} \ln(a^2 N^{2\mu-1}) &\leq C_i \\ \leq -q_0(\gamma)a^{\frac{2}{\mu}-2}(\gamma-1)(\gamma-2)^2 N^{1-\frac{1}{\mu}} [\ln(a^2 N^{2\mu-1}) + 4\mu]. \end{aligned} \quad (46)$$

The boundary term is important when $\ln N$ is not large enough compared with 4μ .

References

1. R. Albert, A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002)
2. S.N. Dorogovtsev, J.F.F. Mendes, *Evolution of Networks* (Oxford University Press, Oxford, 2003)
3. M.E.J. Newman, SIAM Rev. **45**, 167 (2003)
4. A.-L. Barabási, R. Albert, Science **286**, 509 (1999); A.-L. Barabási, R. Albert, H. Jeong, Physica A **272**, 173 (1999)
5. R. Pastor-Satorras, A. Vazquez, A. Vespignani, Phys. Rev. Lett. **87**, 258701 (2001)
6. E. Ravasz, A.L. Somera, D.A. Mongru, Z.N. Oltvai, A.-L. Barabási, Science **297**, 1551 (2002)
7. A. Vazquez, R. Pastor-Satorras, A. Vespignani, Phys. Rev. E **65**, 066130 (2002)
8. P. Erdős, A. Rényi, Publicationes Mathematicae **6**, 290 (1959); Publications of the Mathematical Inst. of the Hungarian Acad. of Sciences **5**, 17 (1960)
9. K.-I. Goh, B. Kahng, D. Kim, Phys. Rev. Lett. **87**, 278701 (2001)
10. G. Caldarelli, A. Capocci, P. De Los Rios, M.A. Muñoz, Phys. Rev. Lett. **89**, 258702 (2002)
11. B. Söderberg, Phys. Rev. E **66**, 066121 (2002)
12. M. Boguñá, R. Pastor-Satorras, Phys. Rev. E **68**, 036112 (2003)
13. D.-S. Lee, K.-I. Goh, B. Kahng, D. Kim, Nucl. Phys. B **696**, 351 (2004)
14. S. Maslov, K. Sneppen, Science **296**, 910 (2002)
15. M.E.J. Newman, Phys. Rev. Lett. **80**, 208701 (2002)
16. E. Ravasz, A.-L. Barabási, Phys. Rev. E **67**, 026112 (2003)
17. R. Milo et al., Science **298**, 824 (2002)
18. M. Catanzaro, R. Pastor-Satorras, Eur. Phys. J. B **44**, 241 (2005)
19. J. Park, M.E.J. Newman, Phys. Rev. E **68**, 026112 (2003)
20. A. Barrat, R. Pastor-Satorras, Phys. Rev. E **71**, 036127 (2005)
21. G. Szabó, M. Alava, J. Kertész, Phys. Rev. E **67**, 056102 (2003)
22. M. Boguñá, R. Pastor-Satorras, A. Vespignani, Eur. Phys. J. B **38**, 205 (2004)
23. M. Catanzaro, M. Boguñá, R. Pastor-Satorras, Phys. Rev. E **71**, 027103 (2005)
24. M. Molloy, B. Reed, Random Structure and Algorithms **6**, 161 (1995)
25. F. Chung, L. Lu, Annals of Combinatorics **6**, 125 (2002)