The Degree Sequence of a Scale-Free Random Graph Process with Hard Copying

Gao-Rong Ning[1]*, Xian-Yuan Wu[1]* and Kai-Yuan Cai[2]†

[1]School of Mathematical Sciences, Capital Normal University, Beijing, 100037, China. Email: ninggaorong@163.com; wuxy@mail.cnu.edu.cn

[2]Department of Automatic Control, Beijing University of Aeronautics and Astronautics, Beijing, 100083, China. Email: kycai@buaa.edu.cn

Abstract: We consider a simple random graph process with *hard* copying as following: At any Time-Step t, with probability $0 < \alpha \leq 1$ a new vertex v_t is added and then m edges incident with v_t are added in the manner of *preferential attachment*; or with probability $1-\alpha$ a existing vertex is copied uniformly at random. We prove in the paper that, when α large enough, the model possesses a mean degree sequence as $d_k \sim Ck^{-(1+2\alpha)}$, where d_k be the limit mean proportion of vertices of degree k. Note that in the present model, while a vertex with large degree is copied, the number of added edges is just its degree, so the number of added edges is not upper bounded.

1 Introduction and the statement of the main result

Real-world networks such as economic companies, biological oscillators, social networks, and the World Wild Web (internet) *etc.* can be modeled by random complex graphs [7, 15, 16, 17, 19, 22]. By studying random complex graphs, various topological properties such as degree-distribution [6, 8, 12, 14], diameter [1, 3, 10], clustering [9, 18], stability [4, 5, 11] and spectral gap [2] of these real-world networks have been presented. One of the most basic properties of real-world networks is the power law degree distribution. As indicated in [6], this property should be a consequence of two generic mechanisms:

1. Evolution: new vertices and edges are added continuously, and

AMS classification: 60K35; 05C 80.

Key words and phrases: degree sequence; power law; hard copying; random graph process

*Supported in part by the Natural Science Foundation of China

†Supported in part by the Natural Science Foundation of China and MicroSoft Research Asia under grant 60633010

2. Preferential attachment: new vertices preferentially attach to vertices that already well connected,

note that mathematical model with the above mechanisms is called BA model. Then many new models with the BA mechanisms have been introduced and aimed to explain the underlying causes for the emergence of power law degree distributions. This can be observed in 'LCD model' [10], the generalization of 'LCD model' due to Buckley and Osthus [8], the very general models defined by Copper and Frieze [13], Copper, Frieze and Vera [14] *etc.*

Beyond the power law degree distributions, other degree distributions including the exponential degree distributions of random graph process are also studied. Actually, phase transitions on degree distributions of random graph processes are studied in the recent works [20] and [21] of Wu *et al.*. More precisely, [20] studied a model with edge deletions and showed that, while a relevant parameter varies, the model exhibits power law degree distribution, a special degree distribution lying between power law and exponential, and exponential degree distribution in turn. And [21] studied a mixed model of BA process and 'classical' process and showed that, while the pure 'classical' process possesses an exponential degree sequence, the pure BA process and the mixed ones possess power law degree sequences.

In this paper we will study a 'copying' model and show its power law degree distribution. The basic idea of 'copying' comes from the fact that a new web page is often made by copying an old one. It is well known that Kumar *et al.* [15] had already studied a kind of 'copying' model to explain the observed power laws in the web graphs in a way different to the BA model: The model is parameterized by a *copy factor* $\alpha \in (0, 1)$ and a constant out-degree $d \geq 1$. At each time step, one vertex u is added and u is then given d out-links. To generate the out-links, first, choose a existing vertex p uniformly at random and then with probability $1 - \alpha$ take the i^{th} out-link of p to be the i^{th} out-link of u, with remaining probability, choose a vertex from the existing vertices uniformly at random to be the destination of the i^{th} out-link of u. It is proved in [15] that the above 'copying' model possesses a power law degree sequence as $d_k \sim Ck^{-(2-\alpha)/(1-\alpha)}$.

The above 'copying' model had provided another mechanism which leads to power law degree sequence of random graphs. But, in fact, at any time step, the probability that a given existing vertex is chosen to be the destination of some out-link of the new vertex is proportional to the number of its neighbors (degree), in point of view of degree growth, this is coincident to *preferential attachment*.

In this paper we will introduce and study another 'copying' model created by *lazy* copiers. Our copiers are so lazy that the only thing they want to do is 'copying'. Clearly, the copiers corresponding to the copying action discussed in [15] should be more clever and diligent: for the chosen vertex p, they have to distinguish which link be a original out-link of p first and then decide whether or not to copy it.

Let's consider the following random process G_t, $t = 2, 3, \cdots$. Assume that graph $G_t = (V_t, E_t)$ and $t = |V_t|$, $e_t = |E_t|$. In order to simplify the statement and the proof of our main result, technically, we start our process at Time-Step 2.

Time-Step 2: To begin the process, we start with G_2 consisting of vertices v_1, v_2 and $2m$ multi-edges between them.

Time-Step $t \geq 3$:

- With probability $\alpha > 0$ we add a new vertex v_t to G_{t-1} and then add m random edges incident with v_t. The m random neighbors $w_1, w_2, \ldots, w_m$ are chosen independently and for any $1 \leq i \leq m$, $w \in V_{t-1}$,
$$\mathbb{P}(w_i = w) = \frac{d_w(t-1)}{2e_{t-1}}, \tag{1.1}$$
where $d_w(t-1)$ denotes the degree of vertex w in G_{t-1}. Thus neighbors are chosen by *preferential attachment.*

- With probability $1-\alpha$ we generate vertex v_t by copying a existing vertex v_i, $1 \leq i \leq t-1$ from V_{t-1} uniformly at random. Note that in this case, all neighbors of v_t are those of the copied vertex v_i.

As defined above, our copying is executed in a direct and simple way, and we call it *hard* copying. With hard copying, e_t may increase nonlinearly, this makes bounding e_t a rather hard problem.

Now, Let $D_k(t)$ be the number of vertices with degree $k \geq 0$ in G_t and let $\overline{D}_k(t)$ be the expectation of $D_k(t)$. The main result of this paper follow as:

Theorem 1.1 *Assume that $2m(1-\alpha) < \alpha$. Then, for all $k \geq 0$, the limit $d_k = \lim_{t\to\infty} \frac{\overline{D}_k(t)}{t}$ exists and satisfies*
$$d_k = 0,\ 0 \leq k < m;\ d_m = \frac{2\alpha}{m+2\alpha};\ d_k = \prod_{i=m+1}^{k} \left(1 + \frac{1+2\alpha}{i+2\alpha}\right) d_m,\ \forall\ k > m.$$
Obviously, $d_k \sim Ck^{-(1+2\alpha)}$ for some constant C.

We follow the basic procedures in [13] and [14] to prove our main theorem. The rest of the paper is organized as follows. In Section 2, we bound the maximum degree and then bound e_t, the number of edges in G_t. In Section 3, using the estimations given in Section 2, we establish the recurrence for $\overline{D}_k(t)$. Finally, in section 4, we derive the approximation of $\overline{D}_k(t)$ by a recurrence with respect to k and then solve the recurrence in k to finish the proof of Theorem 1.1.

2 Bounding the degree and the number of edges

In this section, we first bound the maximum degree in G_t and then bound e_t. Actually, we will give four kinds of estimations to e_t, as will be seen in section 3, the four estimations are all necessary for establishing the recurrence of $\overline{D}_k(t)$.

For $t \geq 2$, let V_t^o be set of *original* vertices in V_t, namely

$$V_t^o := \{v \in V_t : v = v_1,\ v_2 \text{ or } v \text{ is added as a new vertex at some Time-Step } 3 \leq s \leq t\}.$$

For any times s and t with $3 \leq s \leq t$, if $v_s \in V_t^o$, then,

$$d_{v_s}(s) = \frac{1}{2} d_{v_1}(2) = \frac{1}{2} d_{v_2}(2) = m. \tag{2.1}$$

We say an event happens *quite surely* (qs) if the probability of the complimentary set of the event is $O(t^{-K})$ for any $K > 0$.

We bound the degree in G_t from top as follows

Lemma 2.1 *Assume that* $2m(1-\alpha) < 1$ *and* $v_s \in V_t^o$. *Then*

$$d_{v_s}(t) \leq (t/s)^{\alpha/2 + m(1-\alpha)} (\log t)^3 \quad qs. \tag{2.2}$$

Proof: Let Y be the $\{0,1\}$-valued random variable with $\mathbb{P}(Y=1) = \alpha = 1 - \mathbb{P}(Y=0)$. Then using the fact that $e_t \geq mt$, we have

$$\mathbb{E}(d_{v_s}(t+1) \mid G_t) \leq d_{v_s}(t) + YB\left(m, \frac{d_{v_s}(t)}{2mt}\right) + (1-Y)mB\left(1, \frac{d_{v_s}(t)}{t}\right), \tag{2.3}$$

where $B(\cdot,\cdot)$ be the general Binomial random variable.

Using the fact (2.1) and the relation (2.3), Lemma 2.1 follows from the same argument as used in [13], [14] and [20]. □

For any $v \in V_t$, if v is copied at Time-Step s from some vertex v_r, $1 \leq r \leq s-1$, we call v the *daughter* vertex of v_r and call v_r the *mother* vertex of v. Denote by $D(v, G_t)$ the set of all descendants of v in G_t. By the definition of the model, we know that, for any $v_s \in V_t^o$ and $v \in D(v_s, G_t)$, $d_v(t)$ is same distributed as $d_{v_s}(t)$. Now, denote by Δ_t the maximum degree in G_t, then, by Lemma 2.1 and the above analysis, we have

$$\Delta_t \leq t^{\alpha/2 + m(1-\alpha)} (\log t)^3, \quad qs. \tag{2.4}$$

For any $v_s \in V_t^o$, let $f_{v_s}(t) = |D(v_s, G_t)|$ be the number of all descendants of v_s, then, we have

Lemma 2.2 *For any* $s \geq 1$, *if* v_s *is a original vertex, i.e., for some* $t \geq 2$, $v_s \in V_t^o$, *then*

$$f_{v_s}(t) \leq (t/s)^{1-\alpha} (\log t)^3, \quad qs. \tag{2.5}$$

Proof: Let Y be the random variable used in the proof of Lemma 2.1, then,

$$\mathbb{E}(f_{v_s}(t+1) \mid G_t) = f_{v_s}(t) + (1-Y)B\left(1, \frac{f_{v_s}(t)}{t}\right). \tag{2.6}$$

The Lemma follows from the relation (2.6) and the same argument as used in Lemma 2.1. □

Now we begin to bound e_t, the number of edges in G_t. Let a_t be the number of edges added at Time-Step $t+1$, i.e., $e_{t+1} = a_t + e_t$. By the definition of the model, we have $a_t \leq \max\{\Delta_t, m\} = \Delta_t$, $\forall\ t \geq 2$; on the other hand, noticing that the number of multi-edges between any given vertices pair is fewer than $2m$, we have

$$\Delta_2 = 2m, \quad \Delta_{t+1} \leq \Delta_t + 2m,\ \forall\ t \geq 2.$$

This gives the following determined upper bound on e_t,

$$e_t = 2m + \sum_{s=2}^{t-1} a_s \leq 2m + \sum_{s=2}^{t-1} 2m(s-1) = O(t^2). \tag{2.7}$$

For random upper bounds on e_t, firstly, we prove a crude one as

$$e_t \leq O\left(t(\log t)^6\right), \quad qs. \tag{2.8}$$

Indeed, we have

$$2e_t = \sum_{s=1}^{t} d_{v_s}(t) = \sum_{v_s \in V_t^o} \sum_{v \in D(v_s, G_t)} d_v(t).$$

By Lemma 2.1 and Lemma 2.2,

$$\sum_{v_s \in V_t^o} \sum_{v \in D(v_s, G_t)} d_v(t) \leq \sum_{s=1}^{t} \left[(t/s)^{\alpha/2 + (m+1)(1-\alpha)} (\log t)^6 \right] = O\left(t(\log t)^6\right), \quad qs.$$

Note that for the last equality we have used the condition $2m(1-\alpha) < \alpha$, which is given in the statement of Theorem 1.1.

Secondly, we try to give a estimation to $\mathbb{E}(e_t)$, the expectation of the number of edges in G_t. By the definition of the model, we have

$$\mathbb{E}(e_{t+1}|G_t) = e_t + \alpha m + (1-\alpha)\frac{2e_t}{t}, \tag{2.9}$$

so

$$\mathbb{E}(e_{t+1}) = \mathbb{E}(e_t)\left(1 + \frac{2(1-\alpha)}{t}\right) + \alpha m. \tag{2.10}$$

Let

$$\eta_t := e_t - \mu t,$$

where $\mu = \dfrac{\alpha m}{1 - 2(1-\alpha)}$. Then, (2.10) implies that

$$\mathbb{E}(\eta_{t+1}) = \mathbb{E}(\eta_t)\left(1 + \frac{2(1-\alpha)}{t}\right).$$

Thus, $\mathbb{E}(\eta_t) = O(t^{2(1-\alpha)})$ and we have

$$\mathbb{E}(e_t) = \mu t + O(t^{2(1-\alpha)}). \tag{2.11}$$

Finally, we have the following probability estimation on e_t as

Lemma 2.3 *Assume that $2m(1-\alpha) < 1$. Take $\varepsilon_0 > 0$ such that $1 + 2\varepsilon_0 + 2m(1-\alpha) < 2$, then*

$$\mathbb{P}\left(|e_t - \mu t| > t^{\frac{1}{2}+\varepsilon_0+m(1-\alpha)}\right) = O(t^{-\varepsilon_0}). \tag{2.12}$$

Proof: To get the estimation (2.12), we have to bound $\mathrm{Var}(e_t)$, the variance of e_t. First of all, we have

$$\mathrm{Var}(e_{t+1}) = \mathrm{Var}(a_t + e_t) = \mathrm{Var}(e_t) + \mathrm{Var}(a_t) + 2\left(\mathbb{E}(a_t e_t) - \mathbb{E}(a_t)\mathbb{E}(e_t)\right). \tag{2.13}$$

By definition, we have

$$\mathbb{E}(a_t^2 \mid G_t) = \alpha m^2 + (1-\alpha)\sum_{s=1}^{t} \frac{d_{v_s}^2(t)}{t}.$$

Then, by Lemma 2.1 and Lemma 2.2,

$$\begin{aligned}
\mathbb{E}(a_t^2) &= \alpha m^2 + \frac{(1-\alpha)}{t}\mathbb{E}\left(\sum_{v_s\in V_t^o}\sum_{v\in D(v_s,G_t)} d_v^2(t)\right) \\
&\le \alpha m^2 + \frac{(1-\alpha)}{t}\sum_{s=1}^{t}\left[(t/s)^{\alpha+2m(1-\alpha)}(\log t)^6\right]\left[(t/s)^{1-\alpha}(\log t)^3\right] + O(t^{-10}) \\
&= O\left(t^{2m(1-\alpha)}(\log t)^9\right).
\end{aligned} \tag{2.14}$$

In addition, by (2.9) and (2.11), we have

$$\mathbb{E}(a_t) = \alpha m + 2(1-\alpha)\mu + O(t^{2(1-\alpha)-1}). \tag{2.15}$$

Thus

$$\mathrm{Var}(a_t) = O\left(t^{2m(1-\alpha)}(\log t)^9\right). \tag{2.16}$$

For the term $\mathbb{E}(a_t e_t)$, using (2.9), it is clear that

$$\mathbb{E}(a_t e_t | G_t) = e_t \mathbb{E}(a_t | G_t) = e_t\left(m\alpha + 2(1-\alpha)\frac{e_t}{t}\right),$$

then

$$\mathbb{E}(a_t e_t) = m\alpha\mathbb{E}(e_t) + \frac{2(1-\alpha)}{t}\mathbb{E}(e_t^2). \tag{2.17}$$

Using (2.9) again, we have

$$\mathbb{E}(a_t)\mathbb{E}(e_t) = m\alpha E(e_t) + \frac{2(1-\alpha)}{t}\mathbb{E}(e_t)^2. \tag{2.18}$$

Substituting (2.16), (2.17) and (2.18) into (2.13), we get

$$\begin{aligned}
\mathrm{Var}(e_{t+1}) &= \left(1 + \frac{4(1-\alpha)}{t}\right)\mathrm{Var}(e_t) + O\left(t^{2m(1-\alpha)}(\log t)^9\right) \\
&= \left(1 + \frac{4(1-\alpha)}{t}\right)\mathrm{Var}(e_t) + O\left(t^{2m(1-\alpha)+\varepsilon_0}\right),
\end{aligned} \tag{2.19}$$

where $\varepsilon_0 > 0$ is given in the statement of the Lemma. The recurrence (2.19) can be solved directly to get

$$\mathrm{Var}(e_t) = \prod_{s=3}^{t-1}\left(1+\frac{4(1-\alpha)}{s}\right)\left(\mathrm{Var}(e_3) + O\left(\sum_{s=3}^{t-1}\frac{s^{2m(1-\alpha)+\varepsilon_0}}{\prod_{j=3}^{s}(1+4(1-\alpha)/j)}\right)\right)$$

for large t, this implies that

$$\mathrm{Var}(e_t) = O\left(t^{1+2m(1-\alpha)+\varepsilon_0}\right). \tag{2.20}$$

The Lemma follows immediately from (2.11), (2.20) and the Chebychev's inequality. □

3 Establishing The Recurrence for $\overline{D}_k(t)$

Before we establish the recurrence for $\overline{D}_k(t)$, we have to bound the multi-edges first. For $t \geq 2$, let

$$Z_t = \{v \in V_t : \exists\ u \in V_t \text{ s.t. there are multi-edges between } u \text{ and } v\}$$

and $X_t = |Z_t|$, the cardinality of random set Z_t. Clearly, the number of multi-edges in G_t is less than $2mX_t$.

Lemma 3.1 *For any $\epsilon > 0$, we have*

$$\mathbb{E}(X_t) = O\left(t^{\alpha/2+m(1-\alpha)+\epsilon}\right). \tag{3.1}$$

Proof: By the definition of the model, we have

$$\mathbb{E}(X_{t+1} \mid G_t) \leq X_t + (1-\alpha)\frac{X_t}{t} + \alpha\binom{m}{2}\frac{\Delta_t}{e_t}.$$

Taking expectation and then using (2.4) and the fact that $e_t \geq mt$, we have

$$\begin{aligned}\mathbb{E}(X_{t+1}) &\leq \left(1+\frac{1-\alpha}{t}\right)\mathbb{E}(X_t) + O\left(t^{\alpha/2+m(1-\alpha)-1}(\log t)^3\right)\\ &= \left(1+\frac{1-\alpha}{t}\right)\mathbb{E}(X_t) + O\left(t^{\alpha/2+m(1-\alpha)-1+\epsilon}\right). \end{aligned}\tag{3.2}$$

Using the argument between (2.19) and (2.20), the Lemma follows immediately from (3.2). □

Now, we try to establish the recurrence for $\overline{D}_k(t)$. Put $D_k(t) = 0, 0 \leq k < m$, for all $t \geq 2$. For $k \geq m$, we have

$$\begin{aligned}\overline{D}_k(t+1) = \overline{D}_k(t) &+ \alpha m\mathbb{E}\left(-\frac{kD_k(t)}{2e_t} + \frac{(k-1)D_{k-1}(t)}{2e_t} - O\left(\frac{\Delta_t}{e_t}\right)\right)\\ &+(1-\alpha)(k-1)\mathbb{E}\left(-\frac{D_k(t)}{t} + \frac{D_{k-1}(t)}{t} - O\left(\frac{X_t}{t}\right)\right) + \alpha I_{k=m}.\end{aligned}\tag{3.3}$$

The terms $O\left(\frac{\Delta_t}{e_t}\right)$ and $O\left(\frac{X_t}{t}\right)$ account for the probabilities that we create more than one degree changes due to new vertex addition and vertex copying from Z_t respectively.

By Lemma 2.3, the term $\mathbb{E}\left(\frac{kD_k(t)}{2e_t}\right)$ can be expressed as

$$\mathbb{E}\left(\frac{kD_k(t)}{2e_t}\middle| |e_t-\mu t|\le t^{1/2+\varepsilon_0+m(1-\alpha)}\right)\mathbb{P}\left(|e_t-\mu t|\le t^{1/2+\varepsilon_0+m(1-\alpha)}\right)$$

$$+\mathbb{E}\left(\frac{kD_k(t)}{2e_t}\middle| |e_t-\mu t|> t^{1/2+\varepsilon_0+m(1-\alpha)}\right)\mathbb{P}\left(|e_t-\mu t|> t^{1/2+\varepsilon_0+m(1-\alpha)}\right)$$

$$=\frac{\mathbb{E}\left(kD_k(t)|\,|e_t-\mu t|\le t^{1/2+\varepsilon_0+m(1-\alpha)}\right)\mathbb{P}\left(|e_t-\mu t|\le t^{1/2+\varepsilon_0+m(1-\alpha)}\right)}{2\mu t}$$

$$\times\left(1+O\left(t^{-1/2+\varepsilon_0+m(1-\alpha)}\right)\right)+O(t^{-\varepsilon_0}), \tag{3.4}$$

where we used the fact that $kD_k(t)\le 2e_t$ to hand the second term. In addition, we have

$$\mathbb{E}\left(kD_k(t)|\,|e_t-\mu t|\le t^{1/2+\varepsilon_0+m(1-\alpha)}\right)\mathbb{P}\left(|e_t-\mu t|\le t^{1/2+\varepsilon_0+m(1-\alpha)}\right)$$

$$=k\overline{D}_k(t)-\mathbb{E}(kD_k(t);|e_t-\mu t|>t^{1/2+\varepsilon_0+m(1-\alpha)}), \tag{3.5}$$

and

$$\mathbb{E}(kD_k(t);|e_t-\mu t|>t^{1/2+\varepsilon_0+m(1-\alpha)})$$

$$=\mathbb{E}(kD_k(t);|e_t-\mu t|>t^{1/2+\varepsilon_0+m(1-\alpha)},e_t\le O(t(\log t)^6))$$

$$+\mathbb{E}(kD_k(t);|e_t-\mu t|>t^{1/2+\varepsilon_0+m(1-\alpha)},e_t> O(t(\log t)^6))$$

$$\le O(t(\log t)^6)\mathbb{P}(|e_t-\mu t|>t^{1/2+\varepsilon_0+m(1-\alpha)})$$

$$+O(t^2)\mathbb{P}(e_t>O(t(\log t)^6))$$

$$\le O(t^{1-\varepsilon_0}(\log t)^6)+O(t^{-10})=O(t^{1-\varepsilon_0}(\log t)^6). \tag{3.6}$$

Note that to get (3.6), we used the fact that $kD_k(t)\le 2e_t$ and the bounds on e_t given in (2.7) and (2.8).

Thus, combining (3.4), (3.5) and (3.6),

$$\mathbb{E}\left(\frac{kD_k(t)}{2e_t}\right) = \frac{k\overline{D}_k(t)}{2\mu t}\left(1+O\left(t^{-1/2+\varepsilon_0+m(1-\alpha)}\right)\right)+O(t^{-\varepsilon_0}(\log t)^6)$$

$$\le \frac{k\overline{D}_k(t)}{2\mu t}+\frac{\mathbb{E}(2e_t)}{2\mu t}O\left(t^{-1/2+\varepsilon_0+m(1-\alpha)}\right)+O(t^{-\varepsilon_0}(\log t)^6),$$

using (2.11), we have for $k\ge m$

$$\mathbb{E}\left(\frac{kD_k(t)}{2e_t}\right)=\frac{k\overline{D}_k(t)}{2\mu t}+O\left(t^{-1/2+\varepsilon_0+m(1-\alpha)}\right)+O(t^{-\varepsilon_0}(\log t)^6). \tag{3.7}$$

On the other hand, by inequality (2.4) and Lemma 3.1, for any fixed $\epsilon \in (0, 1-\alpha/2-m(1-\alpha))$, we have

$$\mathbb{E}\left(\frac{\Delta_t}{e_t}\right),\ \mathbb{E}\left(\frac{X_t}{t}\right) = O(t^{-1+\alpha/2+m(1-\alpha)+\epsilon}). \tag{3.8}$$

Let

$$\varepsilon_1 = 1/2 \min\left\{\varepsilon_0, 1-\alpha/2-m(1-\alpha), 1/2-\varepsilon_0-m(1-\alpha)\right\}. \tag{3.9}$$

Now, substitute (3.7) and (3.8) into (3.3), we get the recurrence for $\overline{D}_k(t)$ as

$$\begin{aligned}\overline{D}_k(t+1) \quad &= \overline{D}_k(t) - \left(\frac{k}{2} - (1-\alpha)\right)\frac{\overline{D}_k(t)}{t} + \frac{(k-1)}{2}\frac{\overline{D}_{k-1}(t)}{t} \\ &+ O(t^{-\varepsilon_1}) + \alpha I_{k=m}, \quad \forall \ \ k \geq m.\end{aligned} \tag{3.10}$$

Note that the hidden constant, denote by L, in term $O(t^{-\varepsilon_1})$ is independent of k.

4 Solving (3.10) and The Proof Theorem 1.1

In recurrence (3.10), if we heuristically put $\bar{d}_k = \dfrac{\overline{D}_k(t)}{t}$ and assume it is a constant, we get

$$\frac{(k+2\alpha)}{2}\bar{d}_k \ = \frac{(k-1)}{2}\bar{d}_{k-1} + O(t^{-\varepsilon_1}) + \alpha I_{k=m}.$$

This leads to the consideration of the following recurrence in k:

$$\begin{cases} \dfrac{(k+2\alpha)}{2}d_k = \dfrac{(k-1)}{2}d_{k-1} + \alpha I_{k=m}, \quad k \geq m; \\ \\ d_k = 0, \quad 0 \leq k < m. \end{cases} \tag{4.1}$$

The following Lemma shows that (4.1) is a good approximation to (3.10).

Lemma 4.1 *Suppose that $\{d_k : k \geq 0\}$ be the solution of (4.1), then there exists a constant $M > 0$ such that*

$$|\overline{D}_k(t) - td_k| \leq Mt^{1-\varepsilon_1}, \tag{4.2}$$

for all $t \geq 1$ and $k \geq 0$, where ε_1 is given in (3.9).

Proof: The recurrence can be solved directly as: $d_k = 0$, $0 \leq k < m$; $d_m = \dfrac{2\alpha}{m+2\alpha}$ and

$$d_k = \prod_{i=m+1}^{k}\left(1 + \frac{1+2\alpha}{i+2\alpha}\right)d_m, \quad \forall \ k > m. \tag{4.3}$$

Obviously, d_k decay as $k^{-(1+2\alpha)}$, consequently, for some constant C,

$$d_k \leq C/k \text{ for all } k \geq 1. \tag{4.4}$$

Using (4.4) and the degree estimation given in Lemma 2.1, the Lemma follows from a standard argument which can be found in [14] (see Lemma 5.1) and [20] (see Lemma 3.1). □

Proof of Theorem 1.1: Theorem 1.1 follows immediately from (4.2) and (4.3). □

Acknowledgements

The authors thank Prof. Zhao Dong and Prof. Ke Liu for useful advice and discussion.

References

[1] R. Albert, A. Barabási and H. Jeong (1999) *Diameter of the World Wide Web.* Nature, **401**, pp. 103-131.

[2] W. Aiello, F. R. K. Chung and L. Lu (2002) *Random Evolution in Massive Graphs* In *Handbook on Massive Data Sets*, edited by James Abello et al., pp. 510-519. Norwood, MA: Kluwer Academic Publishers

[3] L. A. N. Amaral, A. Scala, M. Barthélémy and H. E. Stanley (2000) *Classes of Small-World Networks,* Proc Natl Acad Sci U S A. 2000 October 10; **97**: pp. 11149-11152.

[4] B. Bollobás (1998) *Modern Graph Theory* Springer-Verlag New York

[5] B. Bollobás (2001) *Random Graph (second edition),* Cambridge University Press

[6] A.-L. Barabási and R. Albert (1999) *Emergence of Scaling in Random Networks,* Science **286**, pp. 509-512

[7] H. R. Bernard, P. D. Killworth, M. J. Evans, C. McCarty and G. A. Shelley (1988) *Studying Social Relations Cross-Culturally,* Ethnology **27**, pp. 155-179

[8] P. G. Buckley and D. Osthus (2004) *Popularity Based Random Graph Model Leading to a Scale-Free Degree Sequence,* Discrete Mathematics, **282**, pp. 53-68.

[9] B. Bollobás and O. Riordan (2002) *Mathematical Results on Scale-Free Random Graphs.* In *Handbook of Graphs and Networks,* pp. 1-34. Berlin: Wiley-VCH.

[10] B. Bollobás and O. Riordan (2004) *The Diameter of a Scale-Free Random Graph,* Combinatorica **4**, pp. 5-34.

[11] B. Bollobás and O. Riordan (2003) *Robustness and Vulnerability of Scale-Free Random Graph,* Internet Mathematics **1**, pp.1-35

[12] B. Bollobás, O. Riordan, J. Spencer and G. Tusnády (2001) *The Degree Sequence of a Scale-Free Random Graph Process* Random Structure and Algorithms, **18**, pp. 279-290.

[13] C. Cooper and A. Frieze (2003) *A General Model of Undireted Web Graphs.* Random Structures and Algorithms, **22**, pp. 311-335.

[14] C. Cooper, A. Frieze and J. Vera (2004) *Random Deletion in a Scale-Free Random Graph Process.* Internet Mathematics **1**, pp. 463-483

[15] R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tomkins and E. Upfal (2000) *Stochastic Models for the Web Graph,* In *41st FOCS*, pp. 57-65.

[16] S. Lehmann, B. Lautrup and A. D. Jackson (2003) *Citation Networks in High Energy Physics,* Phys. Rev. E (Statistical, Nonlinear, and Soft Matter Physics), **68**: 026113

[17] M. Mitzenmacher (2003) *A Brief History of Generative Models for Power Law and Lognormal Distributions,*Internet Mathematics

[18] M. E. J. Newman (2003) *The Structure and Function of the Complex Networks,* SIAM Review, **45**, pp. 167-256.

[19] S. H. Strogatz (2001) *Exploring Complex Networks,* Nature **410**, pp. 268-276

[20] Xuan-Yuan Wu, Zhao Dong, Ke Liu and Kai-Yuan Cai (2008) *On the Degree Sequence and its Critical Phenomenon of an Evolving Random Graph Process,* to appear arXiv:0806.4684v1[math.PR]

[21] Xuan-Yuan Wu, Zhao Dong, Ke Liu and Kai-Yuan Cai (2008) *Phase Transition on The Degree Sequence of a Mixed Random Graph Process,* to appear arXiv:0807.2811v1[math.PR]

[22] D. J. Watts and S. H. Strogatz (1998) *Collective Dynamics of 'Small-World' Networks,* Nature **393**, pp. 440-442.