Coupling on-line and off-line analyses for random power law graphs

Fan Chung *† Linyuan Lu ‡†

Abstract

We develop a coupling technique for analyzing on-line models by using off-line models. This method is especially effective for a growth-deletion model which generalizes and includes the preferential attachment model for generating large complex networks that simulate numerous realistic networks. By coupling the on-line model with the off-line model for random power law graphs, we derive strong bounds for a number of graph properties including diameter, average distances, connected components and spectral bounds. For example, we prove that a power law graph generated by the growth-deletion model almost surely has diameter $O(\log n)$ and average distance $O(\log \log n)$.

1 Introduction

In the past few years, it has been observed that a variety of information networks including Internet graphs, social networks and biological networks among others [1, 3, 4, 5, 17, 18, 20] have the so-called *power law* degree distribution. A graph is called a power law graph if the fraction of vertices with degree k is proportional to $\frac{1}{k^\beta}$ for some constant $\beta > 0$. There are basically two different models for random power law graphs.

The first model is an "on-line" model that mimics the growth of a network. Starting from a vertex (or some small initial graph), a new node and/or new edge is added at each unit of time following the so-called *preferential attachment* scheme [3, 4, 18]. The endpoint of a new edge is chosen with the probability proportional to their (current) degrees. By using a combination of adding new nodes and new edges with given respective probabilities, one can generate large power law graphs with exponents β between 2 and 3 (see [3, 7] for rigorous proofs). Since realistic networks encounter both growth and deletion of vertices

*University of California, San Diego, fan@ucsd.edu
†Research supported in part by NSF Grants DMS 0100472 and ITR 0205061
‡University of California, San Diego, llu@math.ucsd.edu

and edges, here we consider a growth-deletion line model that generalizes and includes the preferential attachment model. Detailed definitions will be given in Section 3.

The second model is an “off-line” model of random graphs with given expected degrees. For a given sequence $\mathbf{w}$ of weights w_v, a random graph in $G(\mathbf{w})$ is formed by choosing the edge between u and v with probability proportional to the product of w_u and w_v. The Erdős-Rényi model $G(n, p)$ can be viewed as a special case of $G(\mathbf{w})$ with all w_i's equal. Because of the independence in the choices of edges, the model $G(\mathbf{w})$ is amenable to a rigorous analysis of various graph properties and structures. In a series of papers [10, 11, 12, 20], various graph invariants have been examined and sharp bounds have been derived for diameter, average distance, connected components and spectra for random power law graphs and, in general, random graphs with given expected degrees.

The on-line model is obviously much harder to analyze than the off-line model. There has been some recent work on the on-line model beyond showing the generated graph has a power law degree distribution. Bollobás and Riordan [7] have derived a number of graph properties for the on-line model by “coupling” with $G(n, p)$, namely, identifying (almost regular) subgraphs whose behavior can be captured in a similar way as graphs from $G(n, p)$ for some appropriate p.

In this paper, our goal is to couple the on-line model with the off-line model of random graphs with the same power law degree distribution so that we can apply the techniques from the off-line model to the on-line model. The basic idea is similar to the martingale method but with substantially differences. The main difference is that there is a *fixed* probability space Ω for the martingale. Although a martingale involves a sequence of functions with consecutive functions having small bounded differences, each function is defined on Ω. For the on-line model, the probability space for the random graph generated at each time instance is *different* in general. We have a sequence of probability spaces where two consecutive ones have “small” differences. To analyze this, we need to examine the difference and relationship of two distinct random graph models, each of which can be viewed as a probability space. In Section 4, we will define the *dominance* of one random graph model over another. Several key lemmas for controlling the differences are also given there.

The main result of this paper is to show the following results for the random graph G generated by the on-line model $G(p_1, p_2, p_3, p_4, m)$ with $p_1 > p_3, p_2 > p_4$, as defined in Section 5:

1. Almost surely the degree sequence of the random graph generated by growth-deletion model $G(p_1, p_2, p_3, p_4, m)$ follows the power law distribu-

tion with exponent $\beta = 2 + (p_1 + p_2)/(p_1 + 2p_2 - p_3 - 2p_4)$.

2. Suppose $m > \log^{2+\epsilon} n$. For $p_2 < p_3 + 2p_4$, we have $2 < \beta < 3$. Almost surely a random graph in $G(p_1, p_2, p_3, p_4, m)$ has diameter $\Theta(\log n)$ and average distance $O(\frac{\log \log n}{\log(1/(\beta-2)})$. We note that the average distance is defined to be the average over all distances among pairs of vertices in the same connected component.

3. Suppose $m > \log^{2+\epsilon} n$. For $p_2 \geq p_3 + 2p_4$, we have $\beta > 3$. Almost surely a random graph in $G(p_1, p_2, p_3, p_4, m)$ has diameter $\Theta(\log n)$ and average distance $O(\frac{\log n}{\log d})$ where d is the average degree.

4. Suppose $m > \log^{2+\epsilon} n$. Almost surely a random graph in $G(p_1, p_2, p_3, p_4, m)$ has Cheeger constant at least $1/2 + o(1)$.

5. Suppose $m > \log^{2+\epsilon} n$. Almost surely a random graph in $G(p_1, p_2, p_3, p_4, m)$ has spectral gap λ at least $1/8 + o(1)$.

We note that the *Cheeger constant* h_G of a graph G which is sometimes called the *conductance* is defined by

$$h_G = \frac{|E(A, \bar{A})|}{\min\{\mathrm{vol}(A), \mathrm{vol}(\bar{A})\}}$$

where $\mathrm{vol}(A) = \sum_{x \in A} deg(x)$. The Cheeger constant is closely related to the spectral gap λ of the Laplacian of a graph by the Cheeger inequality

$$2h_G \geq \lambda \geq h_G^2/2.$$

Thus both h_G and λ are key invariants for controlling the rate of convergence of random walks on G.

2 Strong properties of off-line random power law graphs

For random graphs with given expected degree sequences satisfying a power law distribution with exponent β, we may assume that the expected degrees are $w_i = ci^{-\frac{1}{\beta-1}}$ for i satisfying $i_0 \leq i < n + i_0$.Here c depends on the average degree and i_0 depends on the maximum degree m, namely, $c = \frac{\beta-2}{\beta-1} dn^{\frac{1}{\beta-1}}, i_0 = n(\frac{d(\beta-2)}{m(\beta-1)})^{\beta-1}$.

Average distance and diameter

Fact 1 ([11]) *For a power law random graph with exponent $\beta > 3$ and average degree d strictly greater than 1, almost surely the average distance is $(1 + o(1))\frac{\log n}{\log \tilde{d}}$ and the diameter is $\Theta(\log n)$.*

Fact 2 ([11]) *Suppose a power law random graph with exponent β has average degree d strictly greater than 1 and maximum degree m satisfying $\log m \gg \log n/\log\log n$. If $2 < \beta < 3$, almost surely the diameter is $\Theta(\log n)$ and the average distance is at most $(2 + o(1))\frac{\log \log n}{\log(1/(\beta-2))}$.*

For the case of $\beta = 3$, the power law random graph has diameter almost surely $\Theta(\log n)$ and has average distance $\Theta(\log n/\log\log n)$.

Connected components

Fact 3 ([10]) *Suppose that G is a random graph in $G(\mathbf{w})$ with given expected degree sequence $\mathbf{w}$. If the expected average degree d is strictly greater than 1, then the following holds:*
(1) Almost surely G has a unique giant component. Furthermore, the volume of the giant component is at least $(1 - \frac{2}{\sqrt{de}} + o(1))\mathrm{Vol}(G)$ if $d \geq \frac{4}{e} = 1.4715\ldots$, and is at least $(1 - \frac{1+\log d}{d} + o(1))\mathrm{Vol}(G)$ if $d < 2$.
(2) The second largest component almost surely has size $O(\frac{\log n}{\log d})$.

Spectra of the adjacency matrix and the Laplacian

The spectra of the adjacency matrix and the Laplacian of a non-regular graph can have quite different distribution. The definition for the Laplacian can be found in [8].

Fact 4 ([12])

1. *The largest eigenvalue of the adjacency matrix of a random graph with a given expected degree sequence is determined by m, the maximum degree, and $\tilde{d}$, the weighted average of the squares of the expected degrees. We show that the largest eigenvalue of the adjacency matrix is almost surely $(1 + o(1)) \max\{\tilde{d}, \sqrt{m}\}$ provided some minor conditions are satisfied. In addition, suppose that the k^{th} largest expected degree m_k is significantly larger than $\tilde{d}^2$. Then the k^{th} largest eigenvalue of the adjacency matrix is almost surely $(1 + o(1))\sqrt{m_k}$.*

2. *For a random power law graph with exponent $\beta > 2.5$, the largest eigenvalue of a random power law graph is almost surely $(1+o(1))\sqrt{m}$ where m is the maximum degree. Moreover, the k largest eigenvalues of a random power law graph with exponent β have power law distribution with exponent $2\beta - 1$ if the maximum degree is sufficiently large and k is bounded above by a function depending on β, m and d, the average degree. When*

$2 < \beta < 2.5$, the largest eigenvalue is heavily concentrated at $cm^{3-\beta}$ for some constant c depending on β and the average degree.

3. *We will show that the eigenvalues of the Laplacian satisfy the semi-circle law under the condition that the minimum expected degree is relatively large ($\gg$ the square root of the expected average degree). This condition contains the basic case when all degrees are equal (the Erdös-Rényi model). If we weaken the condition on the minimum expected degree, we can still have the following strong bound for the eigenvalues of the Laplacian which implies strong expansion rates for rapidly mixing,*

$$\max_{i \neq 0} |1 - \lambda_i| \leq (1 + o(1)) \frac{4}{\sqrt{\bar{w}}} + \frac{g(n) \log^2 n}{w_{\min}}$$

where $\bar{w}$ is the expected average degree, $w_{\min}$ is the minimum expected degree and $g(n)$ is any slow growing function of n.

3 A growth-deletion model for generating random power law graphs

One explanation for the ubiquitous occurrence of power laws is the simple growth rules that can result in a power law distribution (see [3, 4]). Nevertheless, realistic networks usually encounter both the growth and deletion of vertices and edges. Here we consider a general on-line model that combine deletion steps with the *preferential attachment model*.

Vertex-growth-step: Add a new vertex v and form a new edge from v to an existing vertex u chosen with probability proportional to d_u.

Edge-growth-step: Add a new edge with endpoints to be chosen among existing vertices with probability proportional to the degrees.

Vertex-deletion-step: Delete a vertex randomly.

Edge-deletion-step: Delete an edge randomly.

For non-negative values p_1, p_2, p_3, p_4 summing to 1, we consider the following growth-deletion model $G(p_1, p_2, p_3, p_4)$:
At each step, with probability p_1, take a vertex-growth step;
With probability p_2, take an edge-growth step;

With probability p_3, take a vertex-deletion step;

With probability $p_4 = 1 - p_1 - p_2 - p_3$, take an edge-deletion step.

Here we assume $p_3 < p_1$ and $p_4 < p_2$ so that the number of vertices and edge grows as t goes to infinity. If $p_3 = p_4 = 0$, the model is just the usual preferential attachment model which generates power law graphs with exponent $\beta = 2 + \frac{p_1}{p_1+2p_2}$. An extensive survey on the preferential attachment model is given in [22] and rigorous proofs can be found in [3, 13].

Previously, Bollobás considered edge deletion after the power law graph is generated [7]. Very recently, Cooper, Frieze and Vera [14] independently consider the growth-deletion model with vertex deletion only. We will show (see Section 5) the following:

Suppose $p_3 < p_1$ and $p_4 < p_2$. Then almost surely the degree sequence of the growth-deletion model $G(p_1, p_2, p_3, p_4)$ follows the power law distribution with the exponent

$$\beta = 2 + \frac{p_1 + p_3}{p_1 + 2p_2 - p_3 - 2p_4}.$$

We note that a random graph in $G(p_1, p_2, p_3, p_4)$ almost surely has expected average degree $(p_1 + p_2 - p_4)/(p_1 + p_3)$. For of p_i's in certain ranges, this value can be below 1 and the random graph is not connected. To simulate graphs with specified degrees, we consider the following modified model $G(p_1, p_2, p_3, p_4, m)$, for some integer m which generates random graphs with the expected degree $m(p_1 + p_2 - p_4)/(p_1 + p_3)$.

At each step, with probability p_1, add a new vertex v and form m new edges from v to existing vertices u chosen with probability proportional to d_u.

With probability p_2, take m edge-growth steps;

With probability p_3, take a vertex-deletion step;

With probability $p_4 = 1 - p_1 - p_2 - p_3$, take m edge-deletion steps.

Suppose $p_3 < p_1$ and $p_4 < p_2$. Then almost surely the degree sequence of the growth-deletion model $G(p_1, p_2, p_3, p_4, m)$ follows the power law distribution with the exponent β the same as the exponent for the model model $G(p_1, p_2, p_3, p_4)$.

$$\beta = 2 + \frac{p_1 + p_3}{p_1 + 2p_2 - p_3 - 2p_4}.$$

Many results for $G(p_1, p_2, p_3, p_4, m)$ can be derived in the same fashion as for $G(p_1, p_2, p_3, p_4)$. Indeed, $G(p_1, p_2, p_3, p_4) = G(p_1, p_2, p_3, p_4, 1)$ is usually the hardest case because of the sparseness of the graphs.

4 Comparing random graphs

In the early work of Erdős and Rényi on random graphs, they first used the model $F(n,m)$ that each graph on n vertices and m edges are chosen randomly with equal probability, where n and m are given fixed numbers. This model is apparently different from the later model $G(n,p)$, for which a random graph is formed by choosing independently each of the $\binom{n}{2}$ pairs of vertices to be an edge with probability p. Because of the simplicity and ease to use, $G(n,p)$ is the model for the seminar work of Erdős and Rényi. Since then, $G(n,p)$ has been widely used and often been referred to as the Erdős-Rényi model. For $m = p\binom{n}{2}$, the two models are apparently correlated in the sense that many graph properties that are satisfied by both random graph models. To precisely define the relationship of two random graph models, we need some definitions.

A graph property P can be viewed as a set of graphs. We say a graph G satisfies property P if G is a member of P. A graph property is said to be *monotone* if whenever a graph H satisfies A, then any graph containing H must also satisfy A. For example, the property A of containing a specified subgraph, say, the Peterson graph, is a monotone property. A random graph G is a probability distribution $Prob(G = \cdot)$. Given two random graphs G_1 and G_2 on n vertices, we say G_1 *dominates* G_2, if for any monotone graph property A, the probability that a random graph from G_1 satisfies A is greater than or equal to the probability that a random graph from G_2 satisfies A, i.e.,

$$Pr(G_1 \text{ satisfies } A) \geq Pr(G_2 \text{ satisfies } A).$$

In this case, we write $G_1 \geq G_2$ and $G_2 \leq G_1$. For example, for any $p_1 \leq p_2$, we have $G(n,p_1) \leq G(n,p_2)$.

For any $\epsilon > 0$, we say G_1 *dominates* G_2 with an error estimate ϵ, if for any monotone graph property A, the probability that a random graph from G_1 satisfies A is greater than or equal to the probability that a random graph from G_2 satisfies A up to an ϵ error term, i.e.,

$$Pr(G_1 \text{ satisfies } A) + \epsilon \geq Pr(G_2 \text{ satisfies } A).$$

If G_1 *dominates* G_2 with an error estimate $\epsilon = \epsilon_n$, which goes to zero as n approaches the infinity, we say G_1 is almost surely dominates G_2. In this case, we write almost surely $G_1 \succeq G_2$ and $G_2 \preceq G_1$.

For example, for any $\delta > 0$, we have almost surely

$$G(n,(1-\delta)\frac{m}{n}) \preceq F(n,m) \preceq G(n,(1+\delta)\frac{m}{n}).$$

We can extend the definition of domination to graphs with different sizes in the following sense. Suppose that random graph G_i has n_i vertices for $i = 1, 2$, and $n_1 < n_2$. By adding $n_2 - n_1$ isolated vertices, the random graph G_1 is extended to the random graph G_1' with the same size as G_2. We say G_2 dominates G_1 if G_2 dominates G_1'.

We consider random graphs that are constructed inductively by *pivoting* at one edge at a time. Here we assume the number of vertices is n.

Edge-pivoting : For an edge $e \in K_n$, a probability q ($0 \leq q \leq 1$), and a random graph G, a new random graph G' can be constructed in the following way. For any graph H, we define

$$
\begin{aligned}
& Pr(G' = H) \\
= & \begin{cases} (1-q)Pr(G = H) & \text{if } e \notin E(H), \\ Pr(G = H) + qPr(G = H \setminus \{e\}) & \text{if } e \in E(H). \end{cases}
\end{aligned}
$$

It is easy to check that $Pr(G' = \cdot)$ is a probability distribution. We say G' is constructed from G by pivoting at the edge e with probability q.

For any graph property A, we define the set A_e to be

$$A_e = \{H \cup \{e\} | H \in A.\}$$

Further, we define the set $A_{\bar{e}}$ to be

$$A_{\bar{e}} = \{H \setminus \{e\} | H \in A.\}$$

In other words, A_e consists of the graphs obtained by adding the edge e to the graphs in A; $A_{\bar{e}}$ consists of the graphs obtained by deleting the edge e from the graphs in A. We have the following useful lemma.

Lemma 1 *Suppose G' is constructed from G by pivoting at the edge e with probability q. Then for any property A, we have*

$$Pr(G' \in A) = Pr(G \in A) + q[Pr((A \cap A_e)_e) - Pr(A \cap A_{\bar{e}})].$$

In particular, if A is a monotone property, we have

$$Pr(G' \in A) \geq Pr(G \in A).$$

Thus, G' dominates G.

Proof: The set associated with a property A can be partitioned into the following subsets. Let $A_1 = A \cap A_e$ be the graphs of A containing the edge e, and

$A_2 = A \cap A_{\bar{e}}$ be the graphs of A not containing the edge e. We have

$$\begin{aligned}
& Pr(G' \in A) \\
= \; & Pr(G' \in A_1) + Pr(G' \in A_2) \\
= \; & \sum_{H \in A_1} Pr(G' = H) + \sum_{H \in A_2} Pr(G' = H) \\
= \; & \sum_{H \in A_1} (Pr(G = H) + qPr(G = H \setminus \{e\})) \\
& + \sum_{H \in A_2} (1-q)Pr(G = H) \\
= \; & Pr(G \in A_1) + Pr(G \in A_2) + qPr(G \in (A_1)_{\bar{e}}) \\
& -qPr(A_2) \\
= \; & Pr(G \in A) + q[Pr((A \cap A_e)_{\bar{e}}) - Pr(A \cap A_{\bar{e}})].
\end{aligned}$$

If A is monotone, we have $A_2 \subset (A_1)_e$. Thus,

$$Pr(G' \in A) \geq Pr(G \in A).$$

Lemma 1 is proved. □

Lemma 2 *Suppose G'_i is constructed from G_i by pivoting the edge e with probability q_i, for $i = 1, 2$. If $q_1 \geq q_2$ and G_1 dominates G_2, then G'_1 dominates G'_2.*

Proof: Following the definitions of A, and letting A_1 and A_2 be as in the proof of Lemma 1, we have

$$\begin{aligned}
& Pr(G'_2 \in A) \\
= \; & Pr(G_2 \in A) + q_2[Pr(G_2 \in (A_1)_{\bar{e}}) - Pr(G_2 \in A_2)] \\
= \; & Pr(G_2 \in A) + q_2 Pr(G_2 \in ((A_1)_{\bar{e}} \setminus A_2)) \\
\geq \; & Pr(G_1 \in A) + q_1 Pr(G_1 \in ((A_1)_{\bar{e}} \setminus A_2)) \\
= \; & Pr(G_1 \in A) + q_1[Pr(G_1 \in (A_1)_{\bar{e}}) - Pr(G_1 \in A_2)] \\
= \; & Pr(G'_1 \in A).
\end{aligned}$$

The proof of Lemma 4.2 is complete. □

Let G_1 and G_2 be the random graphs on n vertices. We define $G_1 \cup G_2$ to be the random graph as follows:

$$Pr(G_1 \cup G_2 = H) = \sum_{H_1 \cup H_2 = H} Pr(G_1 = H_1)Pr(G_2 = H_2)$$

where H_1, H_2 range over all possible pairs of subgraphs that are not necessarily disjoint.

The following Lemma is a generalization of Lemma 2.

Lemma 3 *If G_1 dominates G_3 with an error estimate ϵ_1 and G_2 dominates G_4 with an error estimate ϵ_2, then $G_1 \cup G_2$ dominates $G_3 \cup G_4$ with an error estimate $\epsilon_1 + \epsilon_2$.*

Proof: For any monotone property A and any graph H, we define the set $f(A, H)$ to be

$$f(A,H) = \{G | G \cup H \in A\}.$$

We observe that $f(A, H)$ is also a monotone property. Therefore,

$$\begin{aligned}
& Pr(G_1 \cup G_2 \in A) \\
= & \sum_{H \in A} \sum_{H_1 \cup H_2 = H} Pr(G_1 = H_1) Pr(G_2 = H_2) \\
= & \sum_{H_1} Pr(G_1 = H_1) Pr(G_2 \in f(A, H_1)) \\
\geq & \sum_{H_1} Pr(G_1 = H_1)(Pr(G_4 \in f(A, H_1)) - \epsilon_2) \\
\geq & Pr(G_1 \cup G_4 \in A) - \epsilon_2.
\end{aligned}$$

Similarly, we have

$$Pr_(G_1 \cup G_4 \in A) \geq Pr_(G_3 \cup G_4 \in A) - \epsilon_1.$$

Thus, we get

$$Pr(G_1 \cup G_2 \in A) \geq Pr_(G_3 \cup G_4 \in A) - (\epsilon_1 + \epsilon_2),$$

as desired. $\square$.

Suppose ϕ is a sequence of random graphs $\phi_{G_1}, \phi_{G_2}, \ldots$, where the indices of ϕ range over all graphs on n vertices. Recall that a random graph G is a probability distribution $Prob(G = \cdot)$ over the space of all graphs on n vertices. For any random graph G, we define $\phi(G)$ to be the random graph defined as follows:

$$Pr(\phi(G) = H) = \sum_{H_1 \cup H_2 = H} Pr(G = H_1) Pr(\phi_{H_1} = H_2).$$

We have

Lemma 4 *Let ϕ_1 and ϕ_2 be two sequences of random graphs where the indices of ϕ_1 and ϕ_2 range over all graphs on n vertices. Let G be any random graph. If*

$$Pr(G \in \{H|\phi_1(H) \textit{dominates } \phi_2(H) \textit{with an error estimate } \epsilon_1\}) \geq 1 - \epsilon_2,$$

then $\phi_1(G)$ dominates $\phi_2(G)$ with an error estimate $\epsilon_1 + \epsilon_2$.

Proof: For any monotone property A and any graph H, we have

$$\begin{aligned} & Pr(\phi_1(G) \in A) \\ = & \sum_{H \in A} \sum_{H_1 \cup H_2 = H} Pr(G = H_1) Pr(\phi_1(H_1) = H_2) \\ = & \sum_{H_1} Pr(G = H_1) Pr(\phi_1(H_1) \in f(A, H_1)) \\ \geq & \sum_{H_1} Pr(G = H_1) Pr(\phi_2(H_1) \in f(A, H_1)) - \epsilon_1 - \epsilon_2 \\ \geq & Pr(\phi_2(G) \in A) - (\epsilon_1 + \epsilon_2), \end{aligned}$$

as desired, since $f(A, H) = \{G | G \cup H \in A\}$ is also a monotone property. $\square$.

Let G_1 and G_2 be the random graphs on n vertices. We define $G_1 \setminus G_2$ to be the random graph as follows:

$$Pr(G_1 \setminus G_2 = H) = \sum_{H_1 \setminus H_2 = H} Pr(G_1 = H_1) Pr(G_2 = H_2)$$

where H_1, H_2 range over all pairs of graphs.

Lemma 5 *If G_1 dominates G_3 with an error estimate ϵ_1 and G_2 is dominated by G_4 with an error estimate ϵ_2, then $G_1 \setminus G_2$ dominates $G_3 \setminus G_4$ with an error estimate $\epsilon_1 + \epsilon_2$.*

Proof: For any monotone property A and any graph H, we define the set $\phi(A, H)$ to be

$$\phi(A, H) = \{G | G \setminus H \in A\}.$$

We observe that $\phi(A, H)$ is also a monotone property. Therefore,

$$\begin{aligned} & Pr(G_1 \setminus G_2 \in A) \\ = & \sum_{H \in A} \sum_{H_1 \setminus H_2 = H} Pr(G_1 = H_1) Pr(G_2 = H_2) \\ = & \sum_{H_2} Pr(G_2 = H_2) Pr(G_1 \in \phi(A, H_2)) \\ \geq & \sum_{H_2} Pr(G_2 = H_2)(Pr(G_3 \in \phi(A, H_2)) - \epsilon_1) \\ \geq & Pr(G_3 \setminus G_2 \in A) - \epsilon_1. \end{aligned}$$

Similary, we define the set $\theta(A, H)$ to be

$$\theta(A, H) = \{G | H \setminus G \in A\}.$$

We observe that the complement of the set $\theta(A, H)$ is a monotone property. We have

$$\begin{aligned} & Pr(G_3 \setminus G_2 \in A) \\ = & \sum_{H \in A} \sum_{H_1 \setminus H_2 = H} Pr(G_3 = H_1) Pr(G_2 = H_2) \\ = & \sum_{H_1} Pr(G_3 = H_1) Pr(G_2 \in \theta(A, H_1)) \\ \geq & \sum_{H_1} Pr(G_3 = H_1)(Pr(G_4 \in \theta(A, H_1)) - \epsilon_2) \\ \geq & Pr(G_3 \setminus G_4 \in A) - \epsilon_2. \end{aligned}$$

Thus, we get

$$Pr(G_1 \cup G_2 \in A) \geq Pr_(G_3 \cup G_4 \in A) - (\epsilon_1 + \epsilon_2),$$

as desired.

A random graph G is called *edge-independent* (or independent, for short) if there is an edge-weighted function $p \colon E(K_n) \to [0, 1]$ satisfying

$$Pr(G = H) = \prod_{e \in H} p_e \times \prod_{e \notin H} (1 - p_e).$$

For example, a random graph with a given expected degree sequence is edge-independent. Edge-independent random graphs have many nice properties, several of which we derive here.

Lemma 6 *Suppose that G and G' are independent random graph with edge-weighted functions p and p', then $G \cup G'$ is edge-independent with the edge-weighted function p'' satisfying*

$$p''_e = p_e + p'_e - p_e p'_e.$$

Proof: For any graph H, we have

$$\begin{aligned}
Pr(G \cup G' = H) &= \sum_{H_1 \cup H_2 = H} Pr(G = H_1) Pr(G' = H_2) \\
&= \sum_{H_1 \cup H_2 = H} \prod_{e_1 \in H_1} p_{e_1} \prod_{e_2 \in H_2} p'_{e_2} \prod_{e_3 \notin H_1} (1 - p_{e_3}) \prod_{e_4 \notin H_2} (1 - p'_{e_4}) \\
&= \prod_{e \notin H} (1 - p_e)(1 - p'_e) \prod_{e \in H} (p_e(1 - p'_e) + (1 - p_e)p'_e + p_e p'_e) \\
&= \prod_{e \in H} p''_e \times \prod_{e \notin H} (1 - p''_e).
\end{aligned}$$

□

Lemma 7 *Suppose that G and G' are independent random graph with edge-weighted functions p and p', then $G \setminus G'$ is independent with the edge-weighted function p'' satisfying*

$$p''_e = p_e(1 - p'_e).$$

Proof: For any graph H, we have

$$\begin{aligned}
Pr(G \setminus G' = H) &= \sum_{H_1 \setminus H_2 = H} Pr(G = H_1) Pr(G' = H_2) \\
&= \sum_{H_1 \setminus H_2 = H} \prod_{e_1 \in H_1} p_{e_1} \prod_{e_2 \in H_2} p'_{e_2} \prod_{e_3 \notin H_1} (1 - p_{e_3}) \prod_{e_4 \notin H_2} (1 - p'_{e_4}) \\
&= \prod_{e \in H} (p_e(1 - p'_e)) \prod_{e \notin H} (1 - p_e - p_e p'_e) \\
&= \prod_{e \in H} p''_e \times \prod_{e \notin H} (1 - p''_e).
\end{aligned}$$

□

Let $\{p_e\}_{e \in E(K_n)}$ be a probability distribution over all pairs of vertices. Let G_1 be the random graph of one edge, where a pair e of vertices is chosen with probability p_e. Inductively, we can define the random graph G_m by adding one more random edge to G_{m-1}, where a pair e of vertices is chosen (as the new edge) with probability p_e. (There is a small probability to have the same edges chosen more than once. In such case, we will keep on sampling until we have exactly m different edges.) Hence, G_m has exactly m edges. The probability that G_m has edges $e_1, \ldots, e_m$ is proportional to $p_{e_1} p_{e_2} \cdots p_{e_m}$. The following lemma states that G_m can be sanwitched by two independent random graphs with exponentially small errors if m is large enough.

Lemma 8 *Assume $p_e = o(\frac{1}{m})$ for all $e \in E(K_n)$. Let G' be the independent random graph with edge-weighted function $p'_e = (1-\delta)mp_e$. Let G'' be the independent random graph with edge-weighted function $p''_e = (1+\delta)mp_e$. Then G_m dominates G' with error $e^{-\delta^2 m/4}$, and G_m is also dominated by G'' within an error estimate $e^{-\delta^2 m/4}$.*

Proof: For any Graph H, we define

$$f(H) = \prod_{e \in H} p_e$$

For any graph property B, we define

$$f(B) = \sum_{H \in B} f(H).$$

Let C_k be the set of all graphs with exact k edges.

Claim : For a graph monotone property A and an integer k, we have

$$\frac{f(A \cap C_k)}{f(C_k)} \leq \frac{f(A \cap C_{k+1})}{f(C_{k+1})}.$$

Proof: Both $f(A \cap C_k) f(C_{k+1})$ and $f(A \cap C_{k+1}) f(C_k)$ are homogenous polynimals on $\{p_e\}$ of degree $2k+1$. We compare the coefficients of a general monomial

$$p_{e_1}^2 \cdots p_{e_r}^2 p_{e_{r+1}} \cdots p_{e_{2k-r+1}}$$

in $f(A \cap C_k) f(C_{k+1})$ and $f(A \cap C_{k+1}) f(C_k)$. The coefficient c_1 of the monomial in $f(A \cap C_k) f(C_{k+1})$ is the number of $(k-r)$-subset $\{e_{i_1}, e_{i_2}, \cdots, e_{i_{k-r}}\}$ of $e_{r+1}, \ldots, e_{2k-r+1}$ satisfying that the graph with edges $\{e_1, \ldots, e_r, e_{i_1}, e_{i_2}, \cdots, e_{i_{k-r}}\}$ belongs to A_k. The coefficient c_2 of the monomial in $f(A \cap C_k) f(C_{k+1})$ is the number of $(k-r+1)$-subset $\{e_{i_1}, e_{i_2}, \cdots, e_{i_{k-r+1}}\}$ of $e_{r+1}, \ldots, e_{2k-r+1}$ satisfying that the graph with edges $\{e_1, \ldots, e_r, e_{i_1}, e_{i_2}, \cdots, e_{i_{k-r+1}}\}$ belongs to A_{k+1}. Since A is monotone, if the graph with edges $\{e_1, \ldots, e_r, e_{i_1}, e_{i_2}, \cdots, e_{i_{k-r}}\}$ belongs to A_k then the graph with edges $\{e_1, \ldots, e_r, e_{i_1}, e_{i_2}, \cdots, e_{i_{k-r+1}}\}$ must belong to A_{k+1}. Hence c_1 is always less than or equal to c_2. Thus, we have

$$f(A \cap C_k) f(C_{k+1}) \leq f(A \cap C_{k+1}) f(C_k).$$

The claim is proved.

Now let $p'_e = \frac{(1-\delta)mp_e}{1+(1-\delta)mp_e} = (1+o(1))(1-\delta)mp_e$. In other words, $\frac{p'_e}{1-p'_e} = (1-\delta)mp_e$.

$$\begin{aligned}
Pr(G' \in A) &= \sum_{k=0}^{n} Pr(G' \in A \cap C_k) \\
&\leq \sum_{k=0}^{m} Pr(G' \in A \cap C_k) + \sum_{k=m+1}^{n} Pr(G' \in C_k) \\
&= \prod_e (1-p'_e) \sum_{k=0}^{m} ((1-\delta)m)^k f(A \cap C_k) + Pr(G' \text{ has more than } m \text{ edges}) \\
&\leq \prod_e (1-p'_e) \sum_{k=0}^{m} ((1-\delta)m)^k f(C_k) \frac{f(A \cap C_m)}{f(C_m)} + Pr(G' \text{ has more than } m \text{ edges}) \\
&\leq \frac{f(A \cap C_m)}{f(C_m)} \prod_e (1-p'_e) \sum_{k=0}^{m} ((1-\delta)m)^k f(C_k) + Pr(G' \text{ has more than } m \text{ edges}) \\
&= \frac{f(A \cap C_m)}{f(C_m)} \sum_{k=0}^{m} Pr(G' \in C_k) + Pr(G' \text{ has more than } m \text{ edges}) \\
&\leq \frac{f(A \cap C_m)}{f(C_m)} + Pr(G' \text{ has more than } m \text{ edges}) \\
&= Pr(G_m \in A) + Pr(G' \text{ has more than } m \text{ edges})
\end{aligned}$$

Now we estimate the probability that G' has more than m edges. Let X_e be the 0-1 random variable with $Pr(X_e = 1) = p'_e$. Let $X = \sum_e X_e$. Then $E(X) = (1+o(1))m(1-\delta)$. Now we apply the following large deviation inequality.

$$Pr(X - E(X) > a) \leq e^{-\frac{a^2}{2(E(X)+a/3)}}.$$

We have

$$\begin{aligned}
Pr(X > m) &= Pr(X - E(X) > (1+o(1))\delta m) \\
&\leq e^{-(1+o(1))\frac{\delta^2 m^2}{2(1-\delta)m+\delta m/3}} \\
&\leq e^{-\delta^2 m^2/2}.
\end{aligned}$$

For the other direction, let $p''_e = \frac{(1+\delta)mp_e}{1+(1+\delta)mp_e} = (1+o(1))(1+\delta)mp_e$, which implies $\frac{p''_e}{1-p''_e} = (1+\delta)mp_e$.

$$\begin{aligned}
Pr(G'' \in A) &= \sum_{k=0}^{n} Pr(G'' \in A \cap C_k) \\
&\geq \sum_{k=m}^{n} Pr(G' \in A \cap C_k) \\
&= \prod_e (1-p''_e) \sum_{k=m}^{n} ((1+\delta)m)^k f(A \cap C_k) \\
&\geq \prod_e (1-p'_e) \sum_{k=m}^{n} ((1+\delta)m)^k f(C_k) \frac{f(A \cap C_m)}{f(C_m)} \\
&\geq \frac{f(A \cap C_m)}{f(C_m)} \prod_e (1-p'_e) \sum_{k=m}^{n} ((1+\delta)m)^k f(C_k) \\
&= \frac{f(A \cap C_m)}{f(C_m)} (1 - \sum_{k=0}^{m-1} Pr(G' \in C_k)) \\
&\geq \frac{f(A \cap C_m)}{f(C_m)} - Pr(G'' \text{ has less than } m \text{ edges}) \\
&= Pr(G_m \in A) - Pr(G'' \text{ has less than } m \text{ edges})
\end{aligned}$$

Now we estimate the probability that G'' has less than m edges. Let X_e be the 0-1 random variable with $Pr(X_e = 1) = p''_e$. Let $X = \sum_e X_e$. Then $E(X) = (1+o(1))m(1+\delta)$. Now we apply the following large deviation inequality.

$$Pr(X - E(X) < a) \leq e^{-\frac{a^2}{2E(X)}},$$

We have

$$\begin{aligned}
Pr(X < m) &= Pr(X - E(X) < (1+o(1))\delta m) \\
&\leq e^{-(1+o(1))\frac{\delta^2 m^2}{2(1+\delta)m}} \\
&\leq e^{-\delta^2 m^2/3}.
\end{aligned}$$

The proof of Lemma is completed. □

5 The coupling of the growth deletion model

We will prove the following:

Theorem 1 *Suppose $p_3 < p_1$ and $p_4 < p_2$. Then*

1. *Almost surely the degree sequence of the growth-deletion model* $G(p_1, p_2, p_3, p_4, m)$ *follows the power law distribution with the exponent*

$$\beta = 2 + \frac{p_1 + p_3}{p_1 + 2p_2 - p_3 - 2p_4}.$$

2. *Suppose* $m > \log^{2+\epsilon} n$. *Let* S_1 *be the set of vertices* i *satisfying* $t^{1/2} \leq i \leq t$. *Almost surely the induced graph of* $G(p_1, p_2, p_3, p_4, m)$ *on* S_1 *dominates an (off-line) random graph* G_1 *(also on* S_1*) with expected degree sequence* $w_i = (1 - o(1))m(\frac{t}{i})^{p_1/(p_1-p_3)(\beta-1)}$ *(for* $i \in S_1$*).*

3. *Suppose* $m > \log^{2+\epsilon} n$. *Let* S_2 *be the set of vertices* i *satisfying* $t^{1/2} \leq i \leq t$. *Almost surely the induced graph of* $G(p_1, p_2, p_3, p_4, m)$ *on* S_2 *is dominated by an (off-line) random graph* G_2 *(also on* S_2*)) with expected degree sequence* $w_i = (1 + o(1))m(\frac{t}{i})^{p_1/(p_1-p_3)(\beta-1)}$ *(for* $i \in S_2$*).*

Let n_t (or τ_t) be the number of vertices (or edges) at time t. We first establish the following lemmas on the number of vertices and the number of edges.

Lemma 9 *For any* k *and* t, *in* $G(p_1, p_2, p_3, p_4)$, *the number* n_t *of vertices at time* t *satisfies*

$$(p_1 - p_3)t - \sqrt{2kt \log t} \leq n_t \leq (p_1 - p_3)t + \sqrt{2kt \log t}. \tag{1}$$

with probability at least $1 - \frac{2}{t^k}$.

Proof: The expected number of vertices n_t satisfies the following recurrence relation:

$$E(n_{t+1}) = E(n_t) + p_1 - p_3$$

Hence, $E(n_{t+1}) = (p_1 - p_3)t$. Since we assume $p_3 < p_1$, the graph grows as time t increases. By Azuma's martingale inequality, we have

$$Pr(|n_t - E(n_t)| > a) \leq 2e^{\frac{a^2}{2t}}.$$

By choosing $a = \sqrt{2kt \log t}$, with probability at least $1 - \frac{2}{t^{k+1}}$, we have

$$(p_1 - p_3)t - \sqrt{2kt \log t} \leq n_t \leq (p_1 - p_3)t + \sqrt{2kt \log t}. \tag{2}$$

□

Lemma 10 *For any* ϵ *and* k, *in* $G(p_1, p_2, p_3, p_4)$, *the number* τ_t *of edges at time* t *satisfies*

$$|E(\tau_t) - m\frac{(p_1 + p_2 - p_4)(p_1 - p_3)}{p_1 + p_3}t| \leq \frac{4}{\sqrt{\epsilon}}t^{\frac{1+\epsilon}{2}},$$

with probability at least $1 - \frac{2}{k}t^{-k\epsilon}$, *if* $t \geq (k \log k)^{\frac{1}{\epsilon}}(\frac{2p_3(p_1+p_2-p_4)}{p_1^2-p_3^2})^{2/\epsilon}$

Proof:

The expected number of edges satisfies

$$E(\tau_{t+1}) = E(\tau_t) + mp_1 + mp_2 - p_3 E(\frac{2\tau_t}{n_t}) - mp_4.$$

Let $\tau = m\frac{(p_1+p_2-p_4)(p_1-p_3)}{p_1+p_3}$.

Claim 1: For any $s > \frac{2k \log k p_3^2 (p_1+p_2-p_4)^2}{(p_1^2-p_3^2)^2}$, with probability at least $1 - \sum_{i=s+1}^{t} \frac{2}{t^{k+1}}$, we have

$$|E(\tau_t) - \tau t| < 4m\sqrt{\frac{s}{\log s}}\sqrt{t \log t}. \tag{3}$$

We will prove Claim 1 by induction on t. Let $C_s = 4\sqrt{\frac{s}{\log s}}$. For $t \le s$, the total number of edges is at most $2t$. We have

$$\begin{aligned} |E(\tau_t) - \tau t| &\le 2mt \\ &\le 4m\sqrt{\frac{s}{\log s}}\sqrt{t\log t} \\ &\le C_s\sqrt{t \log t} \end{aligned}$$

Now we assume that $\tau_t - \tau t \le C_s \sqrt{t \log t}$ holds with probability at least $1 - \sum_{i=s+1}^{t} \frac{2}{i^{k+1}}$. For $t+1$, we have

$$\begin{aligned} & |E(\tau_{t+1}) - \tau(t+1)| \\ = & |E(\tau_t) + mp_1 + mp_2 - p_3 E(\frac{2\tau_t}{n_t}) - mp_4 - m(t+1)| \\ = & |E(\tau_t) - \tau t - 2p_3(E(\frac{\tau_t}{n_t}) - \frac{\tau}{p_1 - p_3})| \\ = & |(1 - \frac{2p_3}{(p_1-p_3)t})(E(\tau_t) - mt) - 2p_3(E(\frac{\tau_t}{n_t}) - \frac{\tau_t}{(p_1-p_3)t})| \\ \le & |E(\tau_t) - \tau t| + 2p_3|E(\tau_t)||\frac{1}{(p_1-p_3)t - \sqrt{2kt\log t}} - \frac{1}{(p_1-p_3)t}| \\ \le & C_s\sqrt{t\log t} + (2p_3\tau + C_s\sqrt{\frac{\log t}{t}})(\frac{\sqrt{2k}}{(p_1-p_3)^2}\sqrt{\frac{\log t}{t}} + O(\frac{\log t}{t})) \\ \le & C_s\sqrt{t\log t} + \frac{2m\sqrt{2k}p_3(p_1+p_2-p_4)}{p_1^2 - p_3^2}\sqrt{\frac{\log t}{t}} + O(\frac{\log t}{t}) \\ \le & C_s\sqrt{(t+1)\log(t+1)} + O(\frac{\log t}{t}). \end{aligned}$$

Here we apply the inequality (1). The inequality fails with probability at most

$$\sum_{i=s+1}^{t} \frac{2}{i^{k+1}} + \frac{2}{(t+1)^{k+1}} = \sum_{i=s+1}^{t+1} \frac{2}{i^{k+1}}.$$

We have proved Claim 1. The proof of Lemma 10 follows by choosing $s = t^{\epsilon}$. $\square$

Lemma 10 can be further strengthened as follows:

Lemma 11 *In $G(p_1, p_2, p_3, p_4, m)$, with probability at least $1 - O(t^{-k})$, the total number of edges is*

$$\tau_t = \frac{(p_1 + p_2 - p_4)(p_1 - p_3)}{p_1 + p_3} mt + O(kmt^{1 - \frac{p_1}{4(p_1+p_2)}} \log^2 t).$$

Proof: We claim that

$$d_i(t) = O(km(\frac{t}{i})^{1 - \frac{p_1}{2(p_1+p_2)}} \log^2 t) \tag{4}$$

holds with probability at least $1 - O(\frac{1}{t^k})$. (This will be proved later.)

For any $s \leq t$, let $D_s(t) = \max_{s \leq i \leq t} d_i(t)$ and $\tau_s(t) = \{ij \in E(G_t) | s \leq i, j \leq t\}$. Inequality (7) implies with probability at least $1 - O(\frac{1}{t^{k-1}})$, we have

$$\begin{aligned} \tau_t - \tau_s(t) &\leq \sum_{i \leq s} D_i(t) \\ &\leq \sum_{i \leq s} C \left(\frac{t}{i}\right)^{1 - \frac{p_1}{2(p_1+p_2)}} mk \log^2 t \\ &\leq \frac{C}{\frac{p_1}{2(p_1+p_2)}} mkt \log^2 t \left(\frac{t}{s}\right)^{-\frac{p_1}{2(p_1+p_2)}}. \end{aligned}$$

By choosing $s = \sqrt{t}$, we have

$$\tau_t = \tau_s(t) + O(kmt^{1 - \frac{p_1}{4(p_1+p_2)}} \log^2 t). \tag{5}$$

Since $\tau_t - \tau_s(t) = o(t)$, it suffices to estimate $\tau_s(t)$ instead of $m(t)$.

We define a different random process G'_t, which is exactly the same as $G_t(p_1, p_2, p_3, p_4)$ upto $t \leq s$. For $t \geq s$, G'_t differs from $G_t(p_1, p_2, p_3, p_4)$ only at the vertex-deletion step in the following way. Suppose a vertex i is chosen to be deleted at $G_t(p_1, p_2, p_3, p_4)$. We check the degree $d_i(t)$. If $i \leq s$ or $d_i(t) < Ckm(\frac{t}{i})^{1 - \frac{p_1}{2(p_1+p_2)}} \log^2 t$, (where C is the hidden constant in equation (4), vertex i is deleted in the same way as $G_t(p_1, p_2, p_3, p_4)$. Otherwise the vertex i is kept. Let $\tau'_s(t)$ be $\{ij \in E(G'_t) | s \leq i, j \leq t\}$. We have

$$Pr(\tau_s(t) \neq \tau'_s(t)) \leq \sum_{l=s}^{t} O(l^{-k}) = O(s^{-k+1}). \tag{6}$$

We note that $\tau'_s(t+1) - \tau'_s(t)$ is always bounded by $Ckm(\frac{t}{s})^{1 - \frac{p_1}{2(p_1+p_2)}} \log^2 t$ (denoted by C_t, for short). We apply the martingale inequality to $\tau'_s(t)$. We have

$$
\begin{aligned}
& Pr(|\tau_s'(t) - E(\tau_s'(t))| > a) \\
\leq\ & 2e^{-\frac{a^2}{\sum_{l=s}^{t} C_l^2}} \\
\leq\ & 2e^{-\frac{a^2}{\sum_{l=s}^{t}(Ckm(\frac{l}{s})^{1-\frac{p_1}{2(p_1+p_2)}}\log^2 l)^2}} \\
\leq\ & 2e^{-\frac{a^2}{C'C^2 t^{3-\frac{p_1}{p_1+p_2}} s^{-2+\frac{p_1}{p_1+p_2}} m^2 k^2 \log^4 t}}.
\end{aligned}
$$

We choose $s = \sqrt{t}$ and $a = \sqrt{C'}Ct^{1-\frac{p_1}{4(p_1+p_2)}} mk^{3/2}\log^{5/2} t$. With probability at least $1 - O(t^{-k/2+1})$, we have

$$
|\tau_s'(t) - E(\tau_s'(t))| = O(t^{1-\frac{p_1}{4(p_1+p_2)}} mk^{3/2}\log^{5/2} t).
$$

We note that $\tau_s(t)$ is always less than $2mt$. By equation (6), we have

$$
E(\tau_s(t)) - E(\tau_s'(t)) \leq 4mtO(t^{-k/2+1/2}) = O(mt^{-k/2+3/2}).
$$

Hence with probability at least $1-O(t^{-k/2+1})-O(t^{-k/2+1/2}) = 1-O(t^{-k/2+1})$, we have

$$
\begin{aligned}
|\tau_s(t) - E(\tau_s(t))| \leq & \\
& |\tau_s'(t) - E(\tau_s'(t))| + E(\tau_s(t)) - E(\tau_s'(t)) \\
= \ & O(t^{1-\frac{p_1}{4(p_1+p_2)}} mk^{3/2}\log^{5/2} t) + O(mt^{-k/2+3/2}) \\
= \ & O(t^{1-\frac{p_1}{4(p_1+p_2)}} mk^{3/2}\log^{5/2} t)
\end{aligned}
$$

Combining with inequality (5), with probability $1 - O(t^{-k/2+1})$,

$$
\begin{aligned}
& |\tau_t - E(\tau_t)| \\
\leq\ & |\tau_t - \tau_s(t)| + |E(\tau_t) - E(\tau_s(t))| + |\tau_s(t) - E(\tau_s(t))| \\
\leq\ & O(mkt^{1-\frac{p_1}{4(p_1+p_4)}}\log^2 t) + O(mkt^{1-\frac{p_1}{4(p_1+p_4)}}\log^2 t) \\
& + O(mkt^{1-\frac{p_1}{4(p_1+p_4)}}\log^2 t) \\
=\ & O(mkt^{1-\frac{p_1}{4(p_1+p_4)}}\log^2 t)
\end{aligned}
$$

It remains to prove inequality (4).

We compare $G(p_1, p_2, p_3, p_4, m)$ with the following preferential model $G(p_1, p_2, m)$ without deletion. At each step, with probability p_1, take a vertex-growth step and add m edges from the new vertex to the current graph;

With probability p_2, take an edge-growth step and m edges are added into the current graph;

With probability $1 - p_1 - p_2$, do nothing.

We claim that the degree $d_u(t)$ in the model $G(p_1, p_2, p_3, p_4, m)$ (with deletion) is stochastically dominated by the degree sequence $d_u(t)$ in the model $G(p_1, p_2, m)$ (without deletion). This can be shown by the following balls-and-bins argument. The number of balls in the first bin (denoted by a_1) represents the degree of u while the number of balls in the other bin (denoted by a_1) represents the sum of degrees of the vertices other than u. When an edge incident to u is added to the graph $G(p_1, p_2, p_3, p_4, m)$, it increases both a_1 and a_2 by 1. When an edge not incident to u is added into the graph, a_2 increases by 2 while a_1 remains the same. Without loss of generality, we can assume a_1 is less than a_2 in the initial graph. If an edge uv, which is incident to u, is deleted later, we delay adding this edge until the very moment the edge is to be deleted. At the moment of adding the edge uv, two bins have a_1 and a_2 balls respectively. When we delay adding the edge uv, the number of balls in two bins are still a_1 and a_2 comparing with $a_1 + 1$ and $a_2 + 1$ in the original random process. Since $a_1 < a_2$, the random process with delay dominates the original random process. If an edge vw which is not incident to u is deleted, we also delay adding this edge until the very moment the edge is to be deleted. Equivalently, we compare the process of a_1 and a_2 balls in bins to the process with a_1 and $a_2 + 2$ balls. The random process without delay dominates the one with delay. Therefore, for any u, the degrees of u in the model without deletion dominates the degrees in the model with deletion.

It remains to obtain an appropriate upper bound of $d_u(t)$ for model $G(p_1, p_2, m)$. If a vertex u is added at time i, we label it by i. Let us remove the idle steps and re-parameterize the time. Let $\alpha = \frac{p_1}{p_1+p_2}$. We observe that the model $G(p_1, p_2, m) = G(\alpha, 1-\alpha, 0, 0, m)$. The claim is true by using the upper bound for the degrees of $G(\alpha, 1-\alpha, 0, 0, m)$ as proved in Lemma 13 in the appendix.

The proof of Lemma 13 is complete. □.

We have the following result on the degrees of the model $G(p_1, p_2, p_3, p_4, m)$.

Lemma 12 *If the vertex i ($i \geq \sqrt{t}$) survives up to time t, then, with probability at least $1 - O(t^{-k})$, the degree $d_i(t)$ of the model $G(p_1, p_2, p_3, p_4)$ satisfies*

$$d_i(t) \leq (\frac{t}{i})^{\frac{p_1(p_1+2p_2-p_3-2p_4)}{2(p_1+p_2-p_3-p_4)(p_1-p_3)}}(m + C\log^2 t).$$

If $m > \log^{2+\epsilon} n$, then with probability at least $1 - t^{-k}$ (any constant k), we have

$$d_i(t) \geq (1 - o(1))m(\frac{t}{i})^{\frac{p_1(p_1+2p_2-p_3-2p_4)}{2(p_1+p_2-p_3-p_4)(p_1-p_3)}}.$$

Proof: Let $\eta = p_1 - p_3$ and $\tau = \frac{(p_1+p_2-p_4)(p_1-p_3)}{p_1+p_3}m$. By lemma 11, with probability at least $1 - O(t^{-k})$, the total number of edges is

$$\tau_t = \tau t + O(kmt^{1-\frac{p_1}{4(p_1+p_2)}}\log^2 t).$$

By lemma 9, the number n_t of vertices at time t satisfies

$$n_t = \eta t + O(\sqrt{2kt\log t}),$$

with probability at least $1 - \frac{2}{t^k}$.

Let C_l be the event that either

$$|\tau_l - \frac{(p_1+p_2-p_4)(p_1-p_3)}{p_1+p_3}mt| > Ckml^{1-\frac{p_1}{4(p_1+p_2)}}\log^2 l,$$

or $|n_t - \eta t| > C\sqrt{2kt\log t}$.

The probability that one of C_l $(i \le l \le t)$ occurs is at most

$$\sum_{l=i}^{t} O(t^{-k}) = O(i^{-k+1}) = O(t^{-k/2+1}).$$

Let X_t be the truncated degree of i at time t as follows. $X_t = d_i(t)$ if none of events C_l $(i \le l \le t)$ occurs, 0 otherwise. At $t+1$, with probability of $1 - \frac{1}{n_t}$, the vertex i survives at time $t+1$. Condition on this probability, we have

$$\begin{aligned}
E(e^{\lambda X_{t+1}}|G_t; X_t = x) \quad \le \quad & e^{\lambda x}\frac{1}{1-\frac{1}{n_t}}(p_1(1 - \frac{x}{2\tau_t} + \frac{x}{2\tau_t}e^{\lambda})^m \\
& +p_2(1 - \frac{x}{\tau_t} + \frac{x}{\tau_t}e^{\lambda})^m \\
& +p_3(1 - \frac{x}{n_t} + \frac{x}{n_t}e^{-\lambda}) \\
& +p_4(1 - \frac{x}{\tau_t} + \frac{x}{\tau_t}e^{-\lambda})^m) \\
& +O(i^{-k+1})
\end{aligned}$$

For any $\delta = o(1)$, we can choose λ small enough so that $e^{\lambda} \le 1 + \lambda(1+\delta)$ and $e^{-\lambda} \le 1 - \lambda(1-\delta)$ hold. We have

$$\begin{aligned}
E(e^{\lambda X_{t+1}}|G_t; X_t = x) \quad \le \quad & e^{\lambda x}\frac{1}{1-\frac{1}{n_t}}(p_1 e^{\frac{x}{2\tau_t}\lambda m(1+\delta)} + p_2 e^{\frac{x}{\tau_t}\lambda m(1+\delta)} \\
& +p_3 e^{\frac{x}{n_t}\lambda m(-1+\delta)} + p_4 e^{\frac{x}{2\tau_t}\lambda(-1+\delta)}) + O(i^{-k+1}) \\
\le \quad & e^{\lambda x(1+\frac{p_1 m(1+\delta)}{2\tau_t}+\frac{p_2 m(1+\delta)}{\tau_t}+\frac{p_3 m(-1+\delta)}{n_t}+\frac{p_4 m(-1+\delta)}{\tau_t})}e^{1/n_t} + O(i^{-k+1}) \\
\le \quad & e^{\lambda x(1+\frac{p_1(p_1+2p_2-p_3-2p_4+O(\delta))}{2(p_1+p_2-p_4)(p_1-p_3)t}+o(t^{-1-\epsilon}))}e^{-1/n_t} + O(i^{-k+1})
\end{aligned}$$

Let $\lambda_{t+1} = (1 + \frac{p_1+2p_2-p_3-2p_4+O(\delta)}{2(p_1+p_2-p_4)t} + o(t^{-1-\epsilon}))\lambda_t$, for $t \geq i+1$. We have

$$\begin{aligned}
\lambda_t &= \lambda_{l+1} \prod_{l=i+1}^{t-1} (1 + \frac{p_1(p_1+2p_2-p_3-2p_4+O(\delta))}{2(p_1+p_2-p_4)(p_1-p_3)l} + o(l^{-1-\epsilon}))\lambda_l \\
&\approx \lambda_{i+1} e^{\sum_{l=i+1}^{t-1} \frac{p_1+2p_2-p_3-2p_4+O(\delta)}{2(p_1+p_2-p_4)l} + o(l^{-1-\epsilon})} \\
&\approx \lambda_{l+1} e^{\frac{p_1(p_1+2p_2-p_3-2p_4+O(\delta))}{2(p_1+p_2-p_4)(p_1-p_3)l} \log \frac{t}{i} + o(1)} \\
&\approx \lambda_{l+1} \left(\frac{t}{i}\right)^{\frac{p_1(p_1+2p_2-p_3-2p_4)+O(\delta)}{2(p_1+p_2-p_4)(p_1-p_3)}}
\end{aligned}$$

By choosing $\delta = O(\frac{1}{\log t})$ and $\lambda_t = O(\delta)$, we have

$$\lambda_{i+1} = \frac{1}{\log t} \left(\frac{i}{t}\right)^{\frac{p_1(p_1+2p_2-p_3-2p_4)}{2(p_1+p_2-p_4)(p_1-p_3)}}.$$

Hence

$$E(e^{\lambda_l X_l}) \leq E(e^{\lambda_{l-1} X_{l-1}}) e^{1/(\eta l + O(\sqrt{2kl \log l}))} + O(i^{-k+1}).$$

$$\begin{aligned}
E(e^{\lambda_t X_t}) &\leq E(e^{\lambda_{t-1} X_{t-1}}) e^{1/(\eta t + O(\sqrt{2kt \log t}))} + O(i^{-k+1}) \\
&\leq \cdots \\
&\leq (1+o(1)) E(e^{\lambda_{i+1} X_{i+1}}) \prod_{l=i+1}^{t-1} e^{1/(\eta l + O(\sqrt{2kl \log l}))} \\
&\approx (1+o(1)) \left(\frac{t}{i}\right)^{p_1-p_3}.
\end{aligned}$$

Let $a = (m + Ck \log^2 t) \left(\frac{i}{t}\right)^{\frac{p_1+2p_2-p_3-2p_4}{2(p_1+p_2-p_4)}}$ for some absolute constant C. We have

$$Pr(X_t > a) \leq e^{-\lambda_t a} E(e^{\lambda_t X_t}) \leq e^{-k \log t}.$$

Thus, we have

$$Pr(d_i(t) > a) \leq Pr(X_t > a) + O(i^{-k+1}) \leq t^{-k} + t^{-k/2+1}.$$

as desired.

$$
\begin{aligned}
&E(e^{-\lambda X_{t+1}}|G_t; X_t = x) \\
&\leq \quad e^{-\lambda x}\frac{1}{1-\frac{1}{n_t}}(p_1 e^{-\frac{x}{2\tau_t}\lambda m(1-\delta)} + p_2 e^{-\frac{x}{\tau_t}\lambda m(1-\delta)} \\
&\quad +p_3 e^{\frac{x}{n_t}\lambda m(1+\delta)} + p_4 e^{\frac{x}{2\tau_t}\lambda(1+\delta)}) + O(i^{-k+1}) \\
&\leq \quad e^{-\lambda x(1-\frac{p_1 m(1-\delta)}{2\tau_t}+\frac{p_2 m(1-\delta)}{\tau_t}+\frac{p_3 m(1+\delta)}{n_t}+\frac{p_4 m(1+\delta)}{\tau_t})} + O(i^{-k+1}) \\
&\leq \quad e^{-\lambda x(1-\frac{p_1(p_1+2p_2-p_3-2p_4)-O(\delta)}{2(p_1+p_2-p_4)(p_1-p_3)t}+o(t^{-1-\epsilon}))+\frac{1}{n_t}} + O(i^{-k+1}).
\end{aligned}
$$

Let $\lambda_{t+1} = (1 - \frac{p_1(p_1+2p_2-p_3-2p_4)-O(\delta)}{2(p_1+p_2-p_4)(p_1-p_3)t} + o(t^{-1-\epsilon}))\lambda_t$, for $t \geq i+1$. We have

$$
\begin{aligned}
\lambda_t &= \lambda_{l+1}\prod_{l=i+1}^{t-1}(1 - \frac{p_1(p_1+2p_2-p_3-2p_4)-O(\delta)}{2(p_1+p_2-p_4)(p_1-p_3)l} + o(l^{-1-\epsilon}))\lambda_l \\
&\approx \lambda_{l+1} e^{-\sum_{l=i+1}^{t-1}\frac{p_1(p_1+2p_2-p_3-2p_4)-O(\delta)}{2(p_1+p_2-p_4)(p_1-p_3)l}+o(l^{-1-\epsilon})} \\
&\approx \lambda_{l+1} e^{-\frac{p_1+2p_2-p_3-2p_4+O(\delta)}{2(p_1+p_2-p_4)l}\log\frac{t}{i}+o(1)} \\
&\approx \lambda_{l+1}\left(\frac{t}{i}\right)^{-\frac{p_1(p_1+2p_2-p_3-2p_4)+O(\delta)}{2(p_1+p_2-p_4)(p_1-p_3)}}
\end{aligned}
$$

We choose $\delta = \frac{1}{\log^(1+\epsilon/2)t}$ and $\lambda_{i+1} = \frac{k\log t}{4m} = O(\delta)$. Then,

$$
\lambda_{i+1} = (1+o(1))\lambda_t\left(\frac{i}{t}\right)^{\frac{p_1(p_1+2p_2-p_3-2p_4)}{2(p_1+p_2-p_4)(p_1-p_3)}}.
$$

Let $a = (1-o(1))m\left(\frac{t}{i}\right)^{\frac{p_1(p_1+2p_2-p_3-2p_4)}{2(p_1+p_2-p_4)(p_1-p_3)}}$. We have

$$
\begin{aligned}
Pr(X_t < a) &\leq e^{\lambda_t a}E(e^{-\lambda_t X_t}) \\
&\leq e^{\lambda_t a}E(e^{-\lambda_{i+1}X_{i+1}})e^{\sum_{l=i+1}^{t-1}1/n_l} + te^{\lambda_t a}e^{\sum_{l=i+1}^{t-1}1/n_l}i^{-k+1} \\
&\leq e^{-\lambda_{i+1}m+\lambda_t a+(p_1-p_3)\log\frac{t}{i}} + te^{\lambda_t a}(\frac{t}{i})^{(p_1-p_3)}i^{-k+1} \\
&\leq e^{-\lambda_{i+1}(m-a(1+o(1))\left(\frac{i}{t}\right)^{\frac{p_1(p_1+2p_2-p_3-2p_4)}{2(p_1+p_2-p_4)(p_1-p_3)}})} + O(t^{-k/4+3}) \\
&\leq O(t^{-k/4+3}).
\end{aligned}
$$

Thus, we have

$$
Pr(d_i(t) < a) \leq Pr(X_t < a) \leq O(t^{-k/4+3}).
$$

as desired.

□

Proof of main theorem: The probability that a vertex i survives up to the time t is

$$\prod_{l=i+1}^{t}(1-\frac{p_3}{n_l}) \approx e^{\sum_{l=i+1}^{t} -\frac{p_3}{(p_1-p_3)t}} \approx (\frac{i}{t})^{\frac{p_3}{p_1-p_3}}.$$

Suppose i survives at the time t. Let $\alpha = \frac{p_1(p_1+2p_2-p_3-2p_4)}{2(p_1+p_2-p_4)(p_1-p_3)}$. By Lemma 12, with high probability, we have

$$d_i(t) = (1+o(1))m(\frac{i}{t})^{\alpha}.$$

The number of vertices with degree between x_1 and x_2 is given by

$$\begin{aligned}\sum_{(1+o(1))(\frac{x_2}{m})^{-1/\alpha}t \le i \le (1+o(1))(\frac{x_1}{m})^{-1/\alpha}t} & (\frac{i}{t})^{\frac{p_3}{p_1-p_3}} \\ &\approx ((\frac{x_1}{m})^{\frac{-p_1}{\alpha(p_1-p_3)}} - (\frac{x_1}{m})^{\frac{-p_1}{\alpha(p_1-p_3)}})t \\ &\approx ((\frac{x_1}{m})^{\beta+1} - \frac{x_2}{m})^{-\beta+1}).\end{aligned}$$

Here we apply the following equality:

$$\frac{-p_1}{\alpha(p_1-p_3)} = -\frac{2(p_1+p_2-p_4)}{p_1+2p_2-p_3-2p_4} = -\beta+1.$$

The number of vertices with degree between x and $x+\Delta x$ is

$$(1+o(1))((\frac{x}{m})^{-\beta+1} - \frac{x+\Delta x}{m})^{-\beta+1}) \approx \frac{\beta m^{\beta-1}}{x^{\beta}}\Delta x,$$

Hence, $G(p_1, p_2, p_3, p_4, m)$ is a power law graph with exponent $\beta = 2+\frac{p_1+p_3}{p_1+2p_2-p_3-2p_4}$.

To item (2), we consider $w_i^l = (1-o(1))m(\frac{l}{i})^{\alpha}$ for $t \ge l \ge i \ge \sqrt{t}$, and $\tau_l = (1+o(1))m\frac{(p_1+p_2-p_4)(p_1-p_3)}{p_1+p_3}l$. For $l = \lfloor\sqrt{t}\rfloor, \ldots, t$, we will construct an edge-independent random graph G^l as follows. At $l = \lfloor\sqrt{t}\rfloor$, G^l is an empty graph initially. Inductively, we assume an edge-independent random graph G_j has been constructed, for $j \le l$.

If step $l+1$ is a vertex-growth-step in $G^{l+1}(p_1, p_2, p_3, p_4, m)$, we add a new vertex labelled by $l+1$ to G_l. Let H_v^l be the edge-independent random graph with

$$p_{i,l+1} = m(1-o(1))\frac{w_i^l}{2\tau^l}.$$

We define $G^{l+1} = G^l \cup H_v^l$.

If step $l+1$ is an edge-growth-step in $G^{l+1}(p_1, p_2, p_3, p_4, m)$, Let H_e^l be the edge-independent random graph with

$$p_{i,j} = m(1-o(1))\frac{w_i^l w_j^l}{4\tau_l^2},$$

for all pairs of vertices (i,j) in G^l. We define $G^{l+1} = G^l \cup H_e^l$.

If step $l+1$ is a vertex-deletion-step in $G^{l+1}(p_1,p_2,p_3,p_4,m)$, We delete the same vertex from G^l and call the resulted graph G^{l+1}.

If step $l+1$ is a edge-deletion-step in $G^{l+1}(p_1,p_2,p_3,p_4,m)$, Let H_d^l be the random graph with uniform probability $p = \frac{m}{\tau_l}$. We define $G^{l+1} = G^l \setminus H_d^l$.

Clearly, G^{l+1} is also edge-independent if G^l is edge-independent.

For any two vertices i and j $(i<j)$ in G^l, the edge probability p_{ij}^l satisfies the following recurrence formula.

$$p_{ij}^l = \begin{cases} m(1-o(1))\frac{w_i^j}{2\tau_j} & \text{if } l = j \\ p_{ij}^{l-1} & \text{with probability } p_1+p_3. \\ p_{ij}^{l-1}(1-m(1-o(1))\frac{w_i^l w_j^l}{4\tau_l^2}) + m(1-\delta)\frac{w_i^l w_j^l}{4\tau_l^2} & \text{with probability } p_2 \\ p_{ij}^{l-1}(1-\frac{m}{2\tau_l}) & \text{with probability } p_4 \\ 0 & \text{if either } i \text{ and } j \text{ is deleted, or } l<j \end{cases}$$

We have

$$E(p_{ij}^l) \approx E(p_{ij}^{l-1}) + p_2 m(1-\delta) - p_4\frac{m}{2\tau_l}E(p_{ij}^{l-1}).$$

By solving this recurrence formula, we have

$$E(p_{ij}^l) = (1-o(1))\frac{w_i^l w_j^l}{2\tau_l} = (1-o(1))\frac{(p_1+p_3)}{2m(p_1+p_2-p_4)(p_1-p_3)}\frac{l^{2\alpha-1}}{i^\alpha j^\alpha}.$$

When $l \gg j$, p_{ij}^l concentrates on its expected value. In particular, if $i \le j \le t-\sqrt{t}$, we have $p_{ij}^l \approx (1-o(1))\frac{(p_1+p_3)}{2m(p_1+p_2-p_4)(p_1-p_3)}\frac{l^{2\alpha-1}}{i^\alpha j^\alpha}$. When $j > t-\sqrt{t}$, we have

$$p_{ij}^j = m(1-o(1))\frac{w_i^j}{2\tau_j} = (1-o(1))\frac{(p_1+p_3)}{2m(p_1+p_2-p_4)(p_1-p_3)}\frac{l^{2\alpha-1}}{i^\alpha j^\alpha}.$$

If $j > t-\sqrt{t}$, we have

$$p_{ij}^t \ge p^j ij(1-\frac{m}{2\tau_j})^{t-j} = (1-o(1))p^j ij.$$

Hence, we have

$$p_{ij}^t = (1-o(1))\frac{(p_1+p_3)}{2m(p_1+p_2-p_4)(p_1-p_3)}\frac{l^{2\alpha-1}}{i^\alpha j^\alpha} = (1-o(1))w_i^t w_j^t\frac{1}{2\tau_t},$$

for all $\sqrt{t} \le i < j \le t$. Thus, G^t is the random graph with expected degree sequnence $w_i = w_i^t$.

Now we prove inductively that $G^l(p_1, p_2, p_3, p_4, m)$ dominates G^l within error estimate $o(t^{-K})$ (for any constant K.)

For $l = \sqrt{t}$, the statement is trivial since G^l is an empty graph. We now assume that $G^l(p_1, p_2, p_3, p_4, m)$ dominates G^l within error estimate $o(t^{-K})$ (for any constant K.)

If step $l + 1$ takes a vertex-growth-step, we define the random graph $\phi(H)$ to be the graph consisting m random edges from the new vertex. The other end point of those edges are chosen with probability proportional to their degrees of H. We note that $G^{l+1}(p_1, p_2, p_3, p_4) = \phi(G^l(p_1, p_2, p_3, p_4))$. Since $G^l(p_1, p_2, p_3, p_4, m)$ dominates G^l within error estimate $o(t^{-K})$. Hence, $G^{l+1}(p_1, p_2, p_3, p_4)$ dominates $\phi(G^l)$, which dominates $G^l \cup H_v^l = G^{l+1}$ with an assorciated error term.

If step $l + 1$ takes a edge-growth-step, we define the random graph $\phi(H)$ to be the graph consisting m random edges on the vertices of H. The end points of those edges are chosen with probability proportional to their degrees of H. We note that $G^{l+1}(p_1, p_2, p_3, p_4) = \phi(G^l(p_1, p_2, p_3, p_4))$. Since $G^l(p_1, p_2, p_3, p_4, m)$ dominates G^l within error estimate $o(t^{-K})$. Hence, $G^{l+1}(p_1, p_2, p_3, p_4)$ dominates $\phi(G^l)$, which dominates $G^l \cup H_e^l = G^{l+1}$ with an assorciated error term.

If step $l + 1$ takes a vertex-deletion-step. It is clear $G^{l+1}(p_1, p_2, p_3, p_4)$ dominates G^{l+1} within the same error estimate as in step l.

If step $l + 1$ takes a edge-deletion-step, we note that $G^{l+1}(p_1, p_2, p_3, p_4) = G^l(p_1, p_2, p_3, p_4) \setminus H_d^l$. Since $G^l(p_1, p_2, p_3, p_4, m)$ dominates G^l with an error estimate $o(t^{-K})$. Hence, $G^{l+1}(p_1, p_2, p_3, p_4)$ dominates $G^l \setminus H_e^l = G^{l+1}$.

The total error bound is less that t times the maximum error within each step. Hence the error is $o(t^{-K})$ for any constant K. The proof of item (2) is finished. The proof of item (3) is very similar except that it use the opposite dirction of the domination, and will be omitted here. The theorem is proved. □

References

[1] W. Aiello, F. Chung and L. Lu, A random graph model for massive graphs, *Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing*, (2000) 171-180.

[2] W. Aiello, F. Chung and L. Lu, A random graph model for power law graphs, *Experimental Math.*, **10**, (2001), 53-66.

[3] W. Aiello, F. Chung and L. Lu, Random evolution in massive graphs, *Handbook of Massive Data Sets*, Volume 2, (Eds. J. Abello et al.), Kluwer Academic Publishers, (2002), 97-122.

[4] Albert-László Barabási and Réka Albert, Emergence of scaling in random networks, Science **286** (1999) 509-512.

[5] A. Barabási, R. Albert, and H. Jeong, Scale-free characteristics of random networks: the topology of the world wide web, *Physica* a **281** (2000), 69-77.

[6] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tompkins, and J. Wiener, "Graph Structure in the Web," *proceedings of the WWW9 Conference*, May, 2000, Amsterdam.

[7] B. Bollabás and O. Riordan, Robustness and Vulnerability of Scale-free Random Graphs. *Internet Mathematics*, **1** (2003) no. 1, 1–35.

[8] F. Chung, *Spectral Graph Theory*, AMS Publications, 1997.

[9] Fan Chung and Linyuan Lu, The diameter of random sparse graphs, *Advances in Applied Math.* **26** (2001), 257-279.

[10] Fan Chung and Linyuan Lu, Connected components in a random graph with given degree sequences, *Annals of Combinatorics* **6**, (2002), 125–145.

[11] Fan Chung and Linyuan Lu, The average distance in random graphs with given expected degrees, *Proceeding of National Academy of Science* **99**, (2002), 15879–15882.

[12] Fan Chung, Linyuan Lu and Van Vu, The spectra of random graphs with given expected degrees, *Proceedings of National Academy of Sciences*, **100**, no. 11, (2003), 63 13-6318.

[13] Colin Cooper and Alan Frieze, A general model of undirected Web graphs, *Random Structures and Algorithms*, **22**, (2003), 311-335.

[14] Colin Cooper, Alan Frieze and Juan Vera, Random vertex deletion in a scale free random graph, preprint.

[15] P. Erdős and T. Gallai, Gráfok előírt fokú pontokkal (Graphs with points of prescribed degrees, in Hungarian), *Mat. Lapok* **11** (1961), 264-274.

[16] P. Erdős and A. Rényi, On random graphs. I, *Publ. Math. Debrecen* **6** (1959), 290-291.

[17] H. Jeong, B. Tomber, R. Albert, Z. Oltvai and A. L. Babárasi, The large-scale organization of metabolic networks, *Nature*, **407** (2000), 378-382.

[18] J. Kleinberg, S. R. Kumar, P. Raphavan, S. Rajagopalan and A. Tomkins, The web as a graph: Measurements, models and methods, *Proceedings of the International Conference on Combinatorics and Computing*, 1999.

[19] A. J. Lotka, The frequency distribution of scientific productivity, *The Journal of the Washington Academy of the Sciences*, **16** (1926), 317.

[20] Linyuan Lu, The Diameter of Random Massive Graphs, *Proceedings of the Twelfth ACM-SIAM Symposium on Discrete Algorithms*, (2001) 912-921.

[21] S. Milgram, The small world problem, *Psychology Today*, **2** (1967), 60-67.

[22] Michael Mitzenmacher, A brief history of generative models for power law and lognormal distributions, *Internet Mathematics*, to appear.

Appendix

We consider the preferential attachment model $G(\alpha) = G(\alpha, 1-\alpha, 0, 0)$. We assume that $G(\alpha)$ starts at time $t = 1$ with an initial graph with only one edge. Since the number of edge increases by 1 at a time. The total number of edges at time t is just t.

We label the vertex u by i if u is generated at time i. Let $d_i(t)$ denote the degree of the vertex i at time t. We have the following lemma.

Lemma 13 *There exists a constant C satisfying*

$$Pr(d_i(t) > 2Ck\left(\frac{t}{i}\right)^{1-\frac{\alpha}{2}} \log^2 t) = O(\frac{1}{t^k}) \tag{7}$$

for any k, i and $t \geq t_0$.

Proof: Let $X_t = d_i(t)$. We have $X_i = 0$ and $X_{i+1} = 1$.

For $t \geq i+1$, we have

$$\begin{aligned} & E(e^{\lambda X_{t+1}} | X_t = x) \\ = \;& e^{\lambda x}(1 - \frac{(2-\alpha)x}{2t} + (\frac{(2-\alpha)x}{2t} - \frac{x^2}{4t^2})e^{\lambda} + \frac{x^2}{4t^2}e^{2\lambda}) \\ \leq \;& e^{\lambda x}(1 - \frac{(2-\alpha)x}{2t} + \frac{(2-\alpha)x}{2t}e^{\lambda}. \end{aligned}$$

We choose $\lambda = O(\delta)$ satisfying $e^{\lambda} \leq 1 + \lambda(1+\delta)$. We have

$$\begin{aligned} E(e^{\lambda X_{t+1}} | X_t = x) &\leq e^{\lambda x}(1 - \frac{(2-\alpha)x}{2t} + \frac{(2-\alpha)x}{2t}e^{\lambda}) \\ &\leq e^{\lambda x - \frac{(2-\alpha)x}{2t} + \frac{(2-\alpha)x}{2t}e^{\lambda}} \\ &\leq e^{(1+\frac{(2-\alpha)(1+\delta)}{2t})\lambda x} \end{aligned}$$

Let $\lambda_{t+1} = (1 + \frac{(2-\alpha)(1+\delta)}{2t})\lambda_t$, for $t \geq i+1$. We have

$$\begin{aligned} \lambda_t &= \lambda_{l+1} \prod_{l=i+1}^{t-1} (1 + \frac{(2-\alpha)(1+\delta)}{2l}) \\ &\approx \lambda_{l+1} e^{\sum_{l=i+1}^{t-1} \frac{(2-\alpha)(1+\delta)}{2l}} \\ &\approx \lambda_{l+1} e^{(1-\alpha/2)(1+\delta)\log\frac{t}{i}} \\ &\approx \lambda_{l+1} \left(\frac{t}{i}\right)^{(1-\alpha/2)(1+\delta)} \end{aligned}$$

We choose $\delta = O(\frac{1}{\log t})$ and $\lambda_t = O(\delta)$. We have

$$\lambda_{i+1} = O(\delta \left(\frac{t}{i}\right)^{-(1-\alpha/2)(1+\delta)}) = O(\frac{1}{\log t}\left(\frac{i}{t}\right)^{1-\alpha/2}.$$

Hence

$$E(e^{\lambda_t X_t} \leq \cdots \leq E(e^{\lambda_{i+1} X_{i+1}}) = 1 - o(1).$$

Let $a = Ck\left(\frac{i}{t}\right)^{1-\alpha/2} \log^2 t$ for some absolute constant C. We have

$$Pr(X_t > a) \leq e^{-\lambda_t a} E(e^{\lambda_t X_t}) = e^{-k \log t}$$

The lemma is proved.