# The degree sequence of the preferential attachment model

Lu Yu

October 24, 2024

**Abstract**

For the preferential attachment model of random networks, Bollobás *et al.* proved that when the number $n$ of vertices is large enough, the proportion $P(d)$ of vertices with degree $d$ obeys a power law $P(d) \propto d^{-3}$ for all $d \leq n^{1/15}$. They asked that how far the result can be extended to degrees $d > n^{1/15}$. We give an answer by extending the range to all $d \leq Cn^{1/2-\mu}$, for any arbitrarily small constant $\mu > 0$ and arbitrarily large constant $C > 0$. The answer is near optimal, since the maximum degree is $O(n^{1/2})$.

***Keywords***— preferential attachment model, degree sequence, power law

# 1 Introduction

Systems as genetic networks and the world wide web are best described as some random "dynamic" process. A common property observed in such large real-world networks is that the degree sequence exhibits a tail that decays polynomially. Barabási and Albert [BA99] showed that this feature is a consequence of two mechanisms:

1. the network expands by adding new vertices;

2. new vertices attach preferentially to already existing nodes.

They suggested that after many steps, the proportion $P(d)$ of vertices with degree $d$ should obey a power law $P(d) \propto d^{-\gamma}$, where $\gamma > 0$ is a constant. A heuristic argument suggesting $\gamma = 3$ was given in [BA99, p.511]. By empirical study, Redner

[Red98] showed that the probability $P(d)$ that a paper is cited $d$ times follows the power law $P(d) \propto d^{-3}$ with exponent $\gamma = 3$. Bollobás et al. [Bol+01] proposed a precise model for this kind of dynamic process and confirmed the conjecture $\gamma = 3$.

Fix an integer $m \geq 1$. The model of Bollobás et al. [Bol+01, p.281] defines a random graph process $(G_m^n)_{n\geq 1}$. For every integer $n \geq 1$, $G_m^n$ is a random directed graph with $n$ vertices, and every vertex has outdegree $m$. Roughly speaking, $G_m^{n+1}$ is constructed from $G_m^n$ by adding one new vertex, together with $m$ new edges pointing from this new vertex to $m$ vertices uniformly chosen from existing ones. Fact 1.1 proves that $P(d) \propto d^{-3}$ for all $d \leq n^{1/15}$. More precisely, let $\#_m^n(d)$ be the number of vertices of $G_m^n$ with indegree $d$, which is a random variable. For every integer $d \geq 0$, set

$$\alpha_{m,d} := \frac{2m(m+1)}{(d+m)(d+m+1)(d+m+2)}. \tag{1}$$

One has $\alpha_{m,d} \sim 2m(m+1)d^{-3}$ when $d \to +\infty$. Here the exponent $\gamma = 3$ is independent of $m$.

**Fact 1.1** ([Bol+01, Theorem 1])**.** *Fix $\epsilon > 0$. Then with probability tending to 1, as $n \to +\infty$, one has*

$$(1-\epsilon)\alpha_{m,d} \leq \frac{\#_m^n(d)}{n} \leq (1+\epsilon)\alpha_{m,d}$$

*for every integer $d \in [0, n^{1/15}]$.*

In [Bol+01, p.287], the authors raised Question 1.2.

*Question* 1.2. How far can Fact 1.1 be extended to degrees $d > n^{1/15}$?

Theorem 1.3 gives an answer to Question 1.2 by extending the bound $n^{1/15}$ to $Cn^{\frac{1}{2}-\mu}$ for any positive numbers $\mu$ and $C$.

**Theorem 1.3** (Corollary 3.3)**.** *For any $\epsilon, \mu, C > 0$, with probability tending to 1 when $n \to +\infty$, we have*

$$\left| \frac{\#_m^n(d)}{n} - \alpha_{m,d} \right| \leq \epsilon$$

*for every integer $d \in [0, Cn^{\frac{1}{2}-\mu}]$.*

Throughout the paper, log denotes the natural logarithm. In Theorem 3.2, we prove that ore generally, for every non-decreasing, divergent sequence $(a_n)_{n\geq 2}$ of positive real numbers, the upper bound $\sqrt{n}/(a_n\sqrt{\log n})$ is also feasible. As proved in [Bol+01, Sec. 3] and [Mór05, Theorem 3.1], with probability tending to 1 when

$n \to +\infty$, the maximum degree of $G_m^n$ is $O(n^{1/2})$. Therefore, the range $[0, Cn^{\frac{1}{2}-\mu}]$ in Theorem 1.3 is almost the largest feasible range. In this sense, Theorem 1.3 is an almost optimal solution to Question 1.2.

Bollobás and Riordan [BR04, Theorems 1 and 13] also determined the asymptotic behavior of the diameter of this graph process. For a directed graph $G$, let $\mathrm{diam}(G)$ be its diameter. If $m \geq 2$, then with probability tending to 1 when $n \to +\infty$, one has

$$\mathrm{diam}(G_m^n) \sim \log n / \log \log n.$$

In particular, the probability that $G_m^n$ is connected tends to 1. By contrast, for $m = 1$, the diameter of every connected component of $G_1^n$ is around $\gamma^{-1} \log n$, where $\gamma \in \mathbb{R}$ is the unique solution to $\gamma e^{\gamma+1} = 1$.

# 2 Preferential attachment model

We review the preferential attachment model introduced in [Bol+01, Section 2]. The new edges are created in such a way that a vertex is linked with probability proportional to its current degree. In particular, large degrees tend to become even larger. Thus, the model captures the two mechanisms of Barabási and Albert [BA99].

The preferential attachment model provides a simplified model of the growth of the world wide web. In the **directed** graph, the vertices represent sites, and the arrows show how a site cites earlier sites. A site may link to some part to its own web page, so loops are allowed in the graph. It may contain several citations of another page, so multiple edges are permitted.

Fix a sequence of vertices $(v_i)_{i \geq 1}$. For every integer $m \geq 1$, we shall define a random graph process $(G_m^n)_{n \geq 1}$. For every integer $n \geq 1$, $G_m^n$ is a random *directed* graph on $(v_i)_{i=1}^n$, where every vertex has outdegree $m$. For every arrow from $v_j$ to $v_k$, one has $k \leq j$. Thus, there exist $mn$ edges in $G_m^n$.

For a directed graph $G$, the total degree of a vertex $v \in G$ is written as $d_G(v)$. For $m = 1$, the process $(G_1^n)_{n \geq 1}$ is defined inductively.

**Definition 2.1.** Let $G_1^1$ be the graph with one vertex $v_1$ and one loop. Given $G_1^{k-1}$, form $G_1^k$ by adding the vertex $v_k$ with a single edge directed from $v_k$ to $v_i$, where $i$ is chosen randomly from $\{1, 2, \ldots, k\}$ with

$$P(i = s) = \begin{cases} d_{G_1^{k-1}}(v_s)/(2k-1) & s < k; \\ 1/(2k-1) & s = k. \end{cases}$$

For an integer $m > 1$, we use the process $G_1$ on vertices $v_1', v_2', \ldots$ to define a process $G_m$ on $v_1, v_2, \ldots$, as follows.

**Definition 2.2.** For every integer $n \geq 1$, let $G_m^n$ be the random graph formed from $G_1^{mn}$, by gluing $m$ vertices $v'_{mk-m+1}, v'_{mk-m+2}, \ldots, v'_{mk}$ as a new vertex $v_k$ for every integer $1 \leq k \leq n$.

For any integers $m, n \geq 1$, let $\Gamma_m^n$ be the set of all possibilities of the random graph $G_m^n$, with the corresponding probability measure $\mathbb{P}$. The subscript $m$ is omitted when $m = 1$. By construction, there exists a canonical surjection

$$\Gamma^{mn} \to \Gamma_m^n.$$

Let $\mathbb{N}$ (resp. $\mathbb{N}^+$) be the set of non-negative (resp. positive) integers. For every integer $i \in [1, n]$, define a function

$$d_i : \Gamma^n \to \mathbb{N}^+, \quad G \mapsto d_G(v_i),$$

and set $D_i = \sum_{j=1}^{i} d_j$. Then $2i \leq D_i \leq n + i$. We recall two facts about these random variables that are used in the proof of Theorem 3.1.

**Fact 2.3** ([Bol+01, (3)]). *For every integer $k \geq 1$, on $\Gamma^n$ one has*

$$\mathbb{P}\left(|D_k - 2\sqrt{kn}| \geq 4\sqrt{n \log n}\right) = o(n^{-1})$$

*when $n \to +\infty$.*

For integers $a \geq b \geq 0$, set $(a)_b := a!/(a-b)!$.

**Fact 2.4** ([Bol+01, (4)]). *For all integers $s, d, k, n \geq 0$ with $n \geq d + k + s$, on $\Gamma^n$ one has*

$$\mathbb{P}\left(d_{k+1} = d + 1 \mid D_k = 2k + s\right) = (s + d)2^d \frac{(n - k - s)_d}{(2n - 2k - s)_{d+1}}.$$

Bollobás and Riordan [BR04, p.8] proposed an alternative description of the distribution of $G_1^n$ in terms of pairings, which is more suitable for calculation.

**Definition 2.5.** For an integer $n \geq 1$, an $n$-pairing is a partition of the set $\{1, 2, \ldots, 2n\}$ into $n$ pairs. In each pair, the smaller (resp. larger) element is called the left (resp. right) endpoint. Let $\Pi_n$ be the set of all $n$-pairings.

Given an $n$-pairing $\mathcal{P} \in \Pi_n$, we order the $n$ pairs in $\mathcal{P}$ by increasing order of their right endpoints. For $1 \leq i \leq n$, let $\ell_i(\mathcal{P})$ (resp. $r_i(\mathcal{P})$) be the right endpoint of the $i$-th pair in this order. Then

$$2 \leq r_1(\mathcal{P}) < r_2(\mathcal{P}) < \cdots < r_n(\mathcal{P}) \leq 2n.$$

By convention, $r_0(\mathcal{P}) = 0$.

Define a map $\phi_n : \Pi_n \to \Gamma^n$ as follows. For every $\mathcal{P} \in \Pi_n$ and every integer $1 \leq k \leq n$, there exists a unique integer $i \in [0, k)$ with $\ell_k(\mathcal{P}) \in (r_i(\mathcal{P}), r_{i+1}(\mathcal{P}))$. We draw an arrow from $v_k$ to $v_{i+1}$, which results in a diagram $\phi_n(\mathcal{P}) \in \Gamma^n$.

**Example 2.6.** We illustrate the construction with the case $n = 2$. One has $\Pi_2 = \{(12)(34), (13)(24), (23)(14)\}$. (Every 2-pairing is written in the order defined above.) For $\mathcal{P} = (12)(34)$, the graph $\phi_2(\mathcal{P})$ is depicted below.

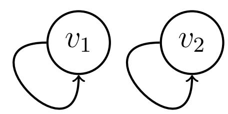


Figure 1: The degrees are $d_1 = d_2 = 2$.

The map $\phi_2$ sends both the 2-pairings $(13)(24)$ and $(23)(14)$ to the following graph.

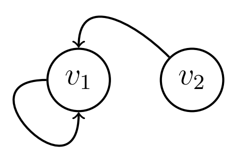


Figure 2: The degree sequence is $d_1 = 3$, $d_2 = 1$.

In the probability space $\Gamma^2$, the first (resp. second) graph has probability 1/3 (resp. 2/3), matting the probabilities from Definition 2.1.

**Fact 2.7.** [BR04, p.8] *For every integer $n \geq 1$, the pushforward of the uniform probability measure on $\Pi_n$ along the map $\phi_n : \Pi_n \to \Gamma^n$ is $\mathbb{P}$.*

**Lemma 2.8.** *For all positive integers $n \geq k$ and all $\mathcal{P} \in \Pi_n$, one has $r_k(\mathcal{P}) = D_k\left(\phi_n(\mathcal{P})\right)$.*

*Proof.* For every integer $0 \leq i \leq k-1$, by construction of $\phi_n(\mathcal{P})$, the indegree of $v_{i+1}$ is

$$d_{i+1}(\phi_n(\mathcal{P})) - 1 = \#\left\{j \in \mathbb{Z} \mid i < j \leq n, \ell_j(\mathcal{P}) \in (r_i(\mathcal{P}), r_{i+1}(\mathcal{P}))\right\}.$$

Every integer in $(r_i(\mathcal{P}), r_{i+1}(\mathcal{P}))$ is the left endpoint of a pair in $\mathcal{P}$, so

$$d_{i+1}(\phi_n(\mathcal{P})) = r_{i+1}(\mathcal{P}) - r_i(\mathcal{P}).$$

Thus, one gets

$$r_k(\mathcal{P}) = \sum_{i=0}^{k-1} (r_{i+1}(\mathcal{P}) - r_i(\mathcal{P})) = \sum_{i=0}^{k-1} d_{i+1}(\phi_n(\mathcal{P})) = D_k(\phi_n(\mathcal{P})).$$

□

# 3 The degree sequence

Fix an integer $m \geq 1$. For integers $n \geq 1$ and $d \geq 0$, let $\#_m^n(d) : \Gamma_m^n \to \mathbb{N}$ be the function giving the number of vertices of $G_m^n$ with indegree $d$. Recall the definition of $\alpha_{m,d}$ from (1). Theorem 3.1 estimates the expectation of this random variable. Section 5 is devoted to its proof.

**Theorem 3.1.** *Let $(a_n)_{n\geq 2}$ be a non-decreasing sequence of positive real numbers with $\lim_n a_n = +\infty$. Then for every $\epsilon > 0$, there exists an integer $N \geq 2$, such that for any integers $n \geq N$ and $0 \leq d \leq \sqrt{n}/(a_n\sqrt{\log n})$, one has*

$$\left| \frac{\mathbb{E}\left(\#_m^n(d)\right)}{n} - \alpha_{m,d} \right| < \epsilon.$$

Theorem 3.2 gives the asymptotic behavior of this random variable.

**Theorem 3.2.** *For any $\epsilon, \delta > 0$ and every non-decreasing sequence $(a_n)_{n\geq 2}$ of positive real numbers with $\lim_n a_n = +\infty$, there exists an integer $n_0 \geq 2$, such that for any integers $n \geq n_0$ and $0 \leq d \leq \sqrt{n}/(a_n\sqrt{\log n})$, one has*

$$\mathbb{P}\left( \left| \frac{\#_m^n(d)}{n} - \alpha_{m,d} \right| > \epsilon \right) \leq \delta.$$

*In particular, the convergence in probability $\#_m^n(d)/n \xrightarrow{p} \alpha_{m,d}$ holds for every integer $d \geq 0$.*

*Proof.* For any integers $n \geq 1$ and $d \geq 0$, define a discrete Doob martingale on the probability space $\Gamma_m^n$ by $X_k = \mathbb{E}[\#_m^n(d) \mid G_m^k]$ for all integers $k \geq 0$. Then $X_0 = \mathbb{E}[\#_m^n(d)]$ and for every $k \geq n$, one has $X_k = \#_m^n(d)$. By [Bol+01, p.287], for every $k \geq 0$, one has

$$|X_{k+1} - X_k| \leq 2. \tag{2}$$

By Theorem 3.1, there exists an integer $N \geq \max\left(2, 32\epsilon^{-2}\log(2/\delta)\right)$ such that for any integers $n \geq N$ and $0 \leq d \leq \sqrt{n}/(a_n\sqrt{\log n})$, one has

$$\left| \frac{\mathbb{E}\left(\#_m^n(d)\right)}{n} - \alpha_{m,d} \right| < \epsilon/2.$$

Then

$$\mathbb{P}\left(\left|\frac{\#_m^n(d)}{n} - \alpha_{m,d}\right| > \epsilon\right) \leq \mathbb{P}\left(\left|\frac{\#_m^n(d)}{n} - \frac{\mathbb{E}\left(\#_m^n(d)\right)}{n}\right| > \frac{\epsilon}{2}\right)$$
$$= \mathbb{P}\left(\left|\mathbb{E}\left(\#_m^n(d)\right) - \#_m^n(d)\right| > \frac{n\epsilon}{2}\right)$$
$$\overset{(a)}{\leq} 2\exp\left(-\frac{n\epsilon^2}{32}\right) \leq \delta,$$

where (a) uses (2) and the Azuma-Hoeffding inequality (see, e.g., [RS13, Theorem 2.2.2]). ☐

**Corollary 3.3.** *For any $\epsilon, \mu, C > 0$, with probability tending to 1 when $n \to +\infty$, we have*

$$\left|\frac{\#_m^n(d)}{n} - \alpha_{m,d}\right| \leq \epsilon$$

*for every integer $d \in [0, Cn^{\frac{1}{2}-\mu}]$.*

*Proof.* Consider the function

$$h : (1, +\infty) \to (0, +\infty), \quad x \mapsto \frac{x^\mu}{\sqrt{\log x}}.$$

Its derivative is

$$h'(x) = \frac{x^{\mu-1}(2\mu \log x - 1)}{2(\log x)^{3/2}},$$

so $h$ is increasing on the interval $[e^{\frac{1}{2\mu}}, +\infty)$. For every integer $n \geq 2$, let

$$a_n = \begin{cases} C^{-1}h(e^{\frac{1}{2\mu}}), & n < e^{\frac{1}{2\mu}}; \\ C^{-1}h(n), & \text{otherwise.} \end{cases}$$

Then $(a_n)_{n\geq 2}$ is a non-decreasing, divergent sequence of positive real numbers. The result follows from Theorem 3.2. ☐

*Remark* 3.4. Since [Bol+01] formulated the model rigorously, many variants have been developed to describe the growth process obeying preferential attachment. They all lead to the same asymptotic behavior. In particular, the occurrence of power laws and the power law exponent do not depend sensitively on the respective choices ([Dur09, p.90] and [VDH16, p.166]).

# 4 Conditional distribution

Lemma 4.1 basically shows that to predict the degree $d_{k+1}$ in $G_1^n$, knowing the sum $D_k$ of the previous degrees is as good as knowing all previous degrees $d_1, \ldots, d_k$.

**Lemma 4.1.** *Let $n > k$ be two non-negative integers. Let $a_1, \ldots, a_k$ be positive integers such that their sum $A_k := \sum_{i=1}^k a_i$ lies in $[2k, n+k]$. Then for every integer $d \in [0, n+k-A_k]$, one has*

$$\mathbb{P}(d_{k+1} = d+1 \mid d_1 = a_1, \ldots, d_k = a_k) = \mathbb{P}(d_{k+1} = d+1 \mid D_k = A_k).$$

*Proof.* Let $s = A_k - 2k$. Let $S := \{\mathcal{P} \in \Pi_n \mid r_k(\mathcal{P}) = 2k+s\}$. Let $L$ be the set of partial pairings on $\{1, 2, \ldots, 2k+s\}$ with $k$ pairs and $s$ unpaired elements, where the $k$-th pair has right endpoint $2k+s$. Let $R$ be the set of partial pairings on $\{2k+s+1, \ldots, 2n\}$ with $s$ unpaired elements and $n-k-s$ pairs. Let $u : S \to L$ and $v : S \to R$ be maps restricting $n$-pairings to partial pairings on $\{1, 2 \ldots, 2k+s\}$ and on $\{2k+s+1, \ldots, 2n\}$ respectively. Let

$$S' := \{\mathcal{P} \in S \mid r_{k+1}(\mathcal{P}) = 2k+s+d+1\}.$$

Let $u' : S' \to L$ be the restriction of $u : S \to L$. Let $R'$ be the partial pairing on $\{2k+s+d+2, \ldots, 2n-1, 2n\}$ with $s+d-1$ unpaired elements and $n-k-s-d$ pairs. Let $v' : S' \to R'$ be the map be the map restricting $n$-pairings to partial pairings on $\{2k+s+d+2, \ldots, 2n-1, 2n\}$. Let $L_0 \subset L$ be the subset of partial pairings on $\{1, \ldots, 2k+s\}$, where for every $1 \leq i \leq k$, the $i$-th pair has right endpoint $A_i := \sum_{j=1}^i a_j$. Let $S_0 = u^{-1}(L_0)$ and $S_0' := S' \cap S_0$.

By Lemma 2.8, one has

$$\begin{aligned}
&\{\mathcal{P} \in \Pi_n \mid D_k(\phi_n(\mathcal{P})) = 2k+s\} = S, \\
&\{\mathcal{P} \in \Pi_n \mid D_k(\phi_n(\mathcal{P})) = 2k+s,\ d_{k+1}(\phi_n(\mathcal{P})) = d+1\} = S', \\
&\{\mathcal{P} \in \Pi_n \mid d_1(\phi_n(\mathcal{P})) = a_1, \ldots, d_k(\phi_n(\mathcal{P})) = a_k,\ d_{k+1}(\phi_n(\mathcal{P})) = d+1\} = S_0', \\
&\{\mathcal{P} \in \Pi_n \mid d_1(\phi_n(\mathcal{P})) = a_1, \ldots, d_k(\phi_n(\mathcal{P})) = a_k\} = S_0.
\end{aligned} \tag{3}$$

By the size of a set, we mean the cardinality. Every fiber of $(u, v) : S \to L \times R$ has the same size. Indeed, for every $\mathcal{L} \in L$ and every $\mathcal{R} \in R$, the fiber $(u, v)^{-1}(\mathcal{L}, \mathcal{R})$ is in bijection with the ways to associate each of the $s$ unpaired elements of $\mathcal{L}$ with a distinct unpaired element of $\mathcal{R}$. Therefore, the size of this fiber is $s!$.

Then the size of every fiber of $u$ is $s! \cdot \#R$. We claim that every fiber of $(u', v') : S' \to L \times R'$ has the same size $(s+d)!$. To see this, note that for every $\mathcal{L} \in L$ and every $\mathcal{R}' \in R'$, an element of the fiber $(u', v')^{-1}(\mathcal{L}, \mathcal{R}')$ can be described in two steps. For $\mathcal{L} \in L$, let $U_{\mathcal{L}}$ be the union of $\{2k+s+1, \ldots, 2k+s+d\}$ with the $s$ unpaired elements of $\mathcal{L}$. Firstly, we choose one element of $U_{\mathcal{L}}$ as the $(k+1)$-th left endpoint, i.e., the element paired with $2k+s+d+1$. There exist $\#U_{\mathcal{L}} = s+d$ ways. Then, we associate each of the $s+d-1$ remaining elements of $U_{\mathcal{L}}$ with a distinct unpaired element of $\mathcal{R}'$. There exist $(s+d-1)!$ ways. Hence this fiber has size $(s+d)!$, completing the proof of the claim.

Then every fiber of $u' : S' \to L$ has size $(s+d)! \cdot \#R'$. As every fiber of $u : S \to L$ (resp. $u' : S' \to L$) has the same size, we have

$$\frac{\#S_0}{\#S} = \frac{\#L_0}{\#L} = \frac{\#S_0'}{\#S'}.$$

Therefore, one has

$$\begin{aligned}
&\mathbb{P}(d_{k+1} = d+1 \mid d_1 = a_1, \dots, d_k = a_k) \\
\overset{\text{(a)}}{=}&\frac{\mathbb{P}(\mathcal{P} \in S_0')}{\mathbb{P}(\mathcal{P} \in S_0)} \\
=&\frac{\#S_0'}{\#S_0} = \frac{\#S'}{\#S} \\
=&\frac{\mathbb{P}(D_k = 2k+s, d_{k+1} = d+1)}{\mathbb{P}(D_k = 2k+s)} \\
=&\mathbb{P}(d_{k+1} = d+1 \mid D_k = A_k),
\end{aligned}$$

where (a) uses (3). □

We need Lemma 4.2 for the proof of Lemma 4.3.

**Lemma 4.2.** *Let $(A_n)_{n\geq 1}, (B_n)_{n\geq 1}, (x_n)_{n\geq 1}$ be sequences of positive real numbers. Let $(\delta_n)_{n\geq 1}$ be a sequence of non-negative integers. If $\delta_n x_n = o\left(\frac{A_n B_n}{A_n+B_n}\right)$, then $\left(\frac{A_n+O(x_n)}{B_n+O(x_n)}\right)^{\delta_n} = (1+o(1))\left(\frac{A_n}{B_n}\right)^{\delta_n}$.*

*Proof.* One may assume that $\delta_n \geq 1$ for every $n \geq 1$. Then

$$x_n \leq \delta_n x_n = o\left(\frac{A_n B_n}{A_n + B_n}\right) = o(B_n).$$

Then for every constant $C > 0$, one has

$$(A_n + B_n)x_n = o\left(\frac{A_n B_n}{\delta_n}\right) = o\left(\frac{A_n(B_n - Cx_n)}{\delta_n}\right).$$

So

$$\frac{A_n + Cx_n}{B_n - Cx_n} - \frac{A_n}{B_n} = \frac{C(A_n + B_n)x_n}{(B_n - Cx_n)B_n} = o\left(\frac{A_n}{\delta_n B_n}\right).$$

One obtains

$$\frac{A_n + O(x_n)}{B_n + O(x_n)} \le \left(1 + o(\delta_n^{-1})\right) \frac{A_n}{B_n},$$

which implies

$$\left(\frac{A_n + O(x_n)}{B_n + O(x_n)}\right)^{\delta_n} \le (1 + o(1)) \left(\frac{A_n}{B_n}\right)^{\delta_n}.$$

Similarly, one has

$$\left(\frac{A_n + O(x_n)}{B_n + O(x_n)}\right)^{\delta_n} \ge (1 + o(1)) \left(\frac{A_n}{B_n}\right)^{\delta_n}.$$

□

In Lemma 4.3, we estimate the distribution of $d_{k+1}$ when the sum of the previous degrees $d_1, \dots, d_k$ are known. This is used in the proof of Theorem 3.1. Set

$$M_1(n) := \sqrt{n}, \quad M_2(n) := \frac{n}{\sqrt[4]{a_n}}. \tag{4}$$

**Lemma 4.3.** *Let $(a_n)_{n\ge1}$ be a sequence of positive real numbers with $a_n \to +\infty$ and $a_n = o(n/\log n)$. Then as $n \to +\infty$, for all integers $k \in [M_1(n), n - M_2(n)]$, $d \in [0, \sqrt{n}/(a_n\sqrt{\log n})]$ and*

$$A \in [2\sqrt{kn} - 5\sqrt{n\log n}, 2\sqrt{kn} + 5\sqrt{n\log n}] \cap [2k, k+n],$$

*on the probability space $\Gamma^n$ one has*

$$\mathbb{P}(d_{k+1} = d + 1 \mid D_k = A) = (1 + o(1)) \sqrt{\frac{k}{n}} \left(1 - \sqrt{\frac{k}{n}}\right)^d.$$

*The implicit function represented by the $o(1)$ term is independent of $A$.*

*Proof.* Let $M_i = M_i(n)$ $(i = 1, 2)$. One has $M_1 = o(n - M_2)$ and $M_2 = o(n)$.

For every real number $c > 0$, define a function

$$f_c : (0, +\infty) \to \mathbb{R}, \quad x \mapsto x(c - x)/(c + x).$$

It is concave, since $f_c''(x) = -4c^2/(c + x)^3 < 0$ for all $x \in (0, +\infty)$. Define a function

$$f : [M_1, n - M_2] \to (0, +\infty), \quad x \mapsto \sqrt{n} f_{\sqrt{n}}(\sqrt{x}) = \frac{\sqrt{xn}(\sqrt{n} - \sqrt{x})}{\sqrt{n} + \sqrt{x}}.$$

Then $f$ attains its minimum at one or both of the endpoints. One has

$$f(M_1) = \sqrt{nM_1}\frac{1-\sqrt{\frac{M_1}{n}}}{1+\sqrt{\frac{M_1}{n}}} \sim \sqrt{nM_1};$$

$$f(n-M_2) = \sqrt{n(n-M_2)}\frac{1-\sqrt{1-\frac{M_2}{n}}}{1+\sqrt{1-\frac{M_2}{n}}} \sim \sqrt{n(n-M_2)}\frac{M_2}{4n} \sim \frac{M_2}{4}.$$

Set $x_n := 4\sqrt{n \log n}$. Since $a_n \to +\infty$, one has $d = o(x_n)$. As $x_n^2/n = o(M_1)$, we find $x_n = o(\sqrt{nM_1}) = o\left(f(M_1)\right)$. By assumption, one has $x_n = o(n/\sqrt{a_n}) = o(M_2) = o(f(n-M_2))$. Thus, one obtains $x_n = o(f(k)) = o\left(A_nB_n/(A_n+B_n)\right)$, where $A_n = \sqrt{kn} - k$ and $B_n = n - \sqrt{kn}$. By Lemma 4.2, one has

$$\frac{\sqrt{kn} - k + O(x_n)}{n - \sqrt{kn} + O(x_n)} = (1+o(1))\sqrt{\frac{k}{n}}. \tag{5}$$

For every real number $c > 0$, define a function

$$g_c : (0, c) \to (0, +\infty), \quad x \mapsto \frac{(c-x)^2}{2c-x}.$$

It is decreasing, since $g_c'(x) = -(c-x)(3c-x)/(2c-x)^2 < 0$ for all $x \in (0, c)$. Define a function

$$g : [M_1, n-M_2] \to (0, +\infty), \quad x \mapsto \sqrt{n}g_{\sqrt{n}}(\sqrt{x}) = \frac{(\sqrt{n}-\sqrt{x})^2\sqrt{n}}{2\sqrt{n}-\sqrt{x}}.$$

Then $g$ is decreasing, so

$$\min g = g(n-M_2) = \frac{n(1-\sqrt{1-M_2/n})^2}{2-\sqrt{1-M_2/n}} \sim n(\frac{M_2}{2n})^2 = \frac{M_2^2}{4n}.$$

As $0 \leq d \leq \sqrt{n}/(a_n\sqrt{\log n})$, one has

$$d \cdot x_n = o(\frac{M_2^2}{n}) = o(g(k)) = o\left(\frac{A_n'B_n}{A_n'+B_n}\right),$$

where $A_n' = (\sqrt{n} - \sqrt{k})^2$. By Lemma 4.2, one has

$$\left(\frac{n+k-2\sqrt{kn}+O(x_n)}{n-\sqrt{kn}+O(x_n)}\right)^d = (1+o(1))\left(1-\sqrt{\frac{k}{n}}\right)^d. \tag{6}$$

By assumption on $A$, one has $s := A - 2k = 2\sqrt{kn} - 2k + O(x_n)$. Then

$$
\begin{aligned}
&\mathbb{P}(d_{k+1} = d+1 \mid D_k = A) \\
&\overset{\text{(a)}}{=} \left(2\sqrt{kn} - 2k + O(x_n)\right) 2^d \frac{\left(n + k - 2\sqrt{kn} + O(x_n)\right)^d}{\left(2n - 2\sqrt{kn} + O(x_n)\right)^{d+1}} \\
&= \frac{\sqrt{kn} - k + O(x_n)}{n - \sqrt{kn} + O(x_n)} \left(\frac{n + k - 2\sqrt{kn} + O(x_n)}{n - \sqrt{kn} + O(x_n)}\right)^d \\
&\overset{\text{(b)}}{=} (1 + o(1)) \sqrt{\frac{k}{n}} \left(1 - \sqrt{\frac{k}{n}}\right)^d,
\end{aligned}
$$

where (a) uses Fact 2.4 and $d = o(x_n)$, while (b) relies on (5) and (6). □

# 5 Proof of Theorem 3.1

There exists a real number $c > 1$ such that the function

$$f : [c, +\infty) \to \mathbb{R}, \quad x \mapsto x(\log x)^{-1-\epsilon}$$

is increasing. For every integer $n \geq c$, replace $a_n$ by $\min(a_n, n(\log n)^{-1-\epsilon})$. Thus, one may assume that $a_n = o(n/\log n)$ and keep the sequence $(a_n)_{n\geq 2}$ non-decreasing.

For every integer $n > 1$, let $M_1(n), M_2(n)$ be as in (4). For every integer $k \in [1, n]$, let $d'_k : \Gamma_m^n \to \mathbb{N}^+$ be the function giving the total degree of vertex $v_k$, and set $D'_k := \sum_{i=1}^k d'_i$. By construction of $G_m^n$, one has $d'_k = \sum_{i=1}^m d_{mk-k+i}$ and $D'_k = D_{mk}$ as functions $\Gamma^{mn} \to \mathbb{N}^+$.

Assume $k \in [M_1(n), n - M_2(n) - 1]$. One has

$$
\begin{aligned}
&M_1(mn) = \sqrt{mn} \leq m\sqrt{n} = mM_1(n) \leq mk, \\
&mk + m \leq m(n - M_2(n)) = m(n - \frac{n}{\sqrt[4]{a_n}}) \overset{\text{(a)}}{\leq} mn - \frac{mn}{\sqrt[4]{a_{mn}}} = mn - M_2(mn),
\end{aligned}
\tag{7}
$$

where (a) uses the non-decreasingness of the sequence $(a_n)$. Set

$$x'_{n,k} := 5\sqrt{mn\log(mn)} + 2m\sqrt{kn} - 2\sqrt{(mk+m)mn} - \frac{\sqrt{n}}{a_n\sqrt{\log n}} - m.$$

When $n$ is large enough, one has

$$\frac{\sqrt{n}}{a_n\sqrt{\log n}} + m + \frac{2m\sqrt{n}}{\sqrt{k+1} + \sqrt{k}} < \sqrt{mn\log(mn)},$$

so

$$x'_{n,k} > 4\sqrt{mn\log(mn)}. \tag{8}$$

Let $S_{n,k}$ denote the set

$$\{A \in \mathbb{Z} \mid 2mk \le A \le mn + mk,\ |A - 2m\sqrt{nk}| \le x'_{n,k}\}.$$

Consider $A \in S_{n,k}$. If $d'_k = d + m$ (i.e., $v_k \in G^n_m$ has indegree $d$) and $D'_k = A$, then for every integer $j \in [mk, mk + m - 1]$, one has an inequality

$$\begin{aligned}
&|D_j - 2\sqrt{jmn}| \le |D_j - D_{mk}| + |D_{mk} - 2m\sqrt{nk}| + |2m\sqrt{nk} - 2\sqrt{jmn}| \\
\le&(d + m) + |A - 2m\sqrt{nk}| + (2\sqrt{jmn} - 2m\sqrt{nk}) \\
\le&5\sqrt{mn\log(mn)}
\end{aligned} \tag{9}$$

for functions on $\Gamma^{mn}$. Let $\mathcal{B} := \{(b_1, \dots, b_m) \in \mathbb{N}^m \mid \sum_{i=1}^{m} b_i = d\}$. Recall that $\#\mathcal{B} = \binom{d+m-1}{m-1}$. Then one has

$$\begin{aligned}
&\mathbb{P}(d'_{k+1} = d + m \mid D'_k = A) = \mathbb{P}(d_{mk+1} + \cdots + d_{mk+m} = d + m \mid D_{mk} = A) \\
\overset{\text{(a)}}{=}& \sum_{(b_1,\dots,b_m)\in\mathcal{B}} \mathbb{P}(d_{mk+i} = 1 + b_i, \forall 1 \le i \le m \mid D_{mk} = A) \\
=& \sum_{(b_1,\dots,b_m)\in\mathcal{B}} \prod_{v=0}^{m-1} \mathbb{P}(d_{mk+v+1} = 1 + b_{v+1} \mid D_{mk} = A, d_{mk+i} = 1 + b_i, \forall 1 \le i \le v) \\
\overset{\text{(b)}}{=}& \sum_{(b_1,\dots,b_m)\in\mathcal{B}} \prod_{v=0}^{m-1} \mathbb{P}(d_{mk+v+1} = 1 + b_{v+1} \mid D_{mk+v} = A + v + \sum_{u=1}^{v} b_u) \\
\overset{\text{(c)}}{=}& \sum_{(b_1,\dots,b_m)\in\mathcal{B}} \prod_{v=0}^{m-1} \left( (1 + o(1)) \sqrt{\frac{mk + v}{mn}} \left(1 - \sqrt{\frac{mk + v}{mn}}\right)^{b_{v+1}} \right) \\
=&(1 + o(1)) \sum_{(b_1,\dots,b_m)\in\mathcal{B}} \prod_{j=1}^{m} \sqrt{\frac{k}{n}} \left(1 - \sqrt{\frac{k}{n}}\right)^{b_j} \\
=&(1 + o(1)) \sum_{(b_1,\dots,b_m)\in\mathcal{B}} \left(\frac{k}{n}\right)^{m/2} \left(1 - \sqrt{\frac{k}{n}}\right)^{d} \\
=&(1 + o(1)) \binom{d + m - 1}{m - 1} \left(\frac{k}{n}\right)^{m/2} \left(1 - \sqrt{\frac{k}{n}}\right)^{d}.
\end{aligned} \tag{10}$$

Here (a) is from

$$\{(d_{mk+1},\ldots,d_{mk+m})\in(\mathbb{N}^+)^m \mid \sum_{i=1}^{m} d_{mk+i}=d{+}m\}=\{(b_1{+}1,\ldots,b_m{+}1)\mid(b_1,\ldots,b_m)\in\mathcal{B}\}.$$

Next, (b) uses Lemma 4.1, while (c) follows from (9), (7) and Lemma 4.3.

We have

$$\begin{aligned}
&\mathbb{P}(d'_{k+1}=d+m)\\
=&\mathbb{P}(d'_{k+1}=d+m,\,|D'_k-2m\sqrt{kn}|>x'_{n,k})\\
+&\sum_{A\in S_{n,k}}\mathbb{P}(D'_k=A)\mathbb{P}(d'_{k+1}=d+m\mid D'_k=A)\\
\overset{(a)}{=}&o(\frac{1}{mn})+\sum_{A\in S_{n,k}}\mathbb{P}(D'_k=A)\,(1+o(1))\binom{d+m-1}{m-1}\left(\frac{k}{n}\right)^{m/2}\left(1-\sqrt{\frac{k}{n}}\right)^d\\
=&o(\frac{1}{mn})+(1+o(1))\binom{d+m-1}{m-1}\left(\frac{k}{n}\right)^{m/2}\left(1-\sqrt{\frac{k}{n}}\right)^d\Big(1-\mathbb{P}(|D'_k-2m\sqrt{nk}|>x'_{n,k})\Big)\\
\overset{(b)}{=}&o(\frac{1}{mn})+(1+o(1))\binom{d+m-1}{m-1}\left(\frac{k}{n}\right)^{m/2}\left(1-\sqrt{\frac{k}{n}}\right)^d(1-o(\frac{1}{mn}))\\
=&o(\frac{1}{n})+(1+o(1))\binom{d+m-1}{m-1}\left(\frac{k}{n}\right)^{m/2}\left(1-\sqrt{\frac{k}{n}}\right)^d.
\end{aligned}\tag{11}$$

Here, we apply Fact 2.3 to $D_{km}$ on $\Gamma^{mn}$ and make use of (8) in both (a) and (b). In (a), we also use (10) to transform the summation. Thus, we get

$$\begin{aligned}
\mathbb{E}(\#^n_m(d))&=\sum_{i=0}^{n-1}\mathbb{P}(d'_{i+1}=d+m)\\
&=O(M_1(n))+O(M_2(n))+\sum_{k=M_1(n)}^{n-M_2(n)-1}\mathbb{P}(d'_{k+1}=d+m)\\
&\overset{(a)}{=}o(n)+(1+o(1))\,n\binom{d+m-1}{m-1}\sum_{k=M_1(n)}^{n-M_2(n)-1}\left(\frac{k}{n}\right)^{m/2}\left(1-\sqrt{\frac{k}{n}}\right)^d\frac{1}{n}\\
&\overset{(b)}{=}o(n)+(1+o(1))\,n\binom{d+m-1}{m-1}\int_0^1\kappa^{m/2}(1-\sqrt{\kappa})^d d\kappa\\
&=o(n)+(1+o(1))\,n\alpha_{m,d},
\end{aligned}$$

where (a) uses (11). In (b), the Riemann sum with rectangles of width $1/n$ is approximated by the Riemann integral when $n \to \infty$.

We obtain $\left|\frac{\mathbb{E}(\#_m^n(d))}{n} - \alpha_{m,d}\right| = o(1)$, which completes the proof.


## Acknowledgments

I thank my supervisor Philippe Bich for his support. I am grateful to the referees for their suggestions, which have greatly improved the exposition.


# References


[BA99] Albert-László Barabási and Réka Albert. "Emergence of scaling in random networks". In: *Science* 286.5439 (1999), pp. 509–512.

[Bol+01] Béla Bollobás, Oliver Riordan, Joel Spencer, and Gábor Tusnády. "The degree sequence of a scale-free random graph process". In: *Random Structures & Algorithms* 18.3 (2001), pp. 279–290.

[BR04] Béla Bollobás and Oliver Riordan. "The diameter of a scale-free random graph". In: *Combinatorica* 24.1 (2004), pp. 5–34.

[Dur09] Richard Durrett. *Random Graph Dynamics*. Cambridge University Press, 2009. DOI: 10.1017/CBO9780511546594.

[Mór05] Tamás F. Móri. "The maximum degree of the Barabási-Albert random tree". In: *Combinatorics, Probability and Computing* 14.3 (2005), pp. 339–348.

[Red98] Sidney Redner. "How popular is your paper? An empirical study of the citation distribution". In: *The European Physical Journal B-Condensed Matter and Complex Systems* 4.2 (1998), pp. 131–134.

[RS13] Maxim Raginsky and Igal Sason. "Concentration of Measure Inequalities in Information Theory, Communications, and Coding". In: *Foundations and Trends® in Communications and Information Theory* 10.1–2 (2013), pp. 1–246. DOI: 10.1561/0100000064.

[VDH16] Remco Van Der Hofstad. *Random Graphs and Complex Networks*. Cambridge University Press, 2016. DOI: 10.1017/9781316779422.