

## HIGH DEGREE VERTICES IN THE POWER OF CHOICE MODEL COMBINED WITH PREFERENTIAL ATTACHMENT

**Malyshkin Yu.**

Senior lecturer at Applied Physics department, Tver State University.  
*Russia, 170100, Tver, 33 Zhelyabova str., TSU.*

---

*Received 30.08.2016, revised 17.03.2017.*

---

We find asymptotics for the first  $k$  highest degrees of the degree distribution in an evolving tree model combining the local choice and the preferential attachment. In the considered model, the random graph is constructed in the following way. At each step, a new vertex is introduced. Then, we connect it with one (the vertex with the largest degree is chosen) of  $d$  ( $d > 2$ ) possible neighbors, which are sampled from the set of the existing vertices with the probability proportional to their degrees. It is known that the maximum of the degree distribution in this model has linear behavior. We prove that  $k$ -th highest degree has a sublinear behavior with a power depends on  $d$ . This contrasts sharply with what is seen in the preferential attachment model without choice, where all highest degrees in the degree distribution has the same sublinear order. The proof is based on showing that the considered tree has a persistent hub by comparison with the standard preferential attachment model, along with martingale arguments.

**Keywords:** random graphs, preferential attachment, choice.

*Vestnik TvGU. Seriya: Prikladnaya Matematika [Herald of Tver State University. Series: Applied Mathematics], 2017, no. 1, pp. 31–43.*

### 1. Introduction

In the present work, we further explore how the addition of choice (see, e.g., [2, 7, 11, 12]) affects the standard preferential attachment model (see [1, 8]). The preferential attachment graph model is a time-indexed inductively constructed sequence of graphs, formed in the following way. We start with some initial graph and then on each step we add a new vertex and an edge between it and one of the old vertices, chosen with probability proportional to its degree. Many different properties of this model have been obtained in both the math and physics literature (see [1, 3, 8, 13]).

In the current work, we are interested in the first  $k$  maximums of the degree distribution. For the preferential attachment model, this problem is studied in [4]. It is shown in [4] that the  $k$  highest degrees  $\Delta_i(n)$ ,  $i \in \{1, \dots, k\}$ , at time  $n$  satisfy

$$\frac{n^{1/2}}{g(n)} \leq \Delta_1(n) \leq g(n)n^{1/2} \text{ and } \frac{n^{1/2}}{g(n)} \leq \Delta_i(n) \leq \Delta_{i-1}(n) - \frac{n^{1/2}}{g(n)}, \quad k \geq i \geq 2,$$

with high probability for any function  $g(n)$  with  $g(n) \rightarrow \infty$  as  $n \rightarrow \infty$ .

In [12], the limited choice is introduced into the preferential attachment model. More specifically, at each step we independently (from each other) choose  $d$  existing vertices with a probability proportional to their degrees and connect the new vertex with the vertex with the smallest degree. In [12], it is shown that the maximal degree at time  $n$  in such a model grows as  $\log \log n / \log d$  with high probability. If instead of a vertex with the smallest degree we pick one with the highest degree, we would get the max-choice model that was introduced in [11]. In [11], the exact first-order asymptotics for the maximal degree in this model was obtained and almost sure convergence of the appropriately scaled maximal degree was shown. In the current work, we provide such asymptotics for  $k$  highest degrees.

Let us describe the max-choice model. Fix  $d \in \mathbb{N}$ ,  $d \geq 2$ . Introduce a countable non-random set of vertices  $V = \{v_i, i \in \mathbb{N}\}$ . Define a sequence of random trees  $\{P_n\}$ ,  $n \in \mathbb{N}$ , by the following inductive rule. Let  $P_1$  be the one-edge tree which consists of vertices  $v_1$  and  $v_2$  and an edge between them. Given  $P_n$ , we construct  $P_{n+1}$  by adding one vertex and drawing one edge in the following way.

First, we add a vertex  $v_{n+2}$  to  $P_n$ , hence the vertices set  $V(P_{n+1})$  of  $P_{n+1}$  is  $V(P_{n+1}) = \{v_i, i = 1, \dots, n+2\}$ . Note that the randomness of  $P_n$  caused by its edge set  $\mathcal{E}_n$ . Denote  $\mathcal{F}_n = \sigma\{\mathcal{E}_1, \dots, \mathcal{E}_n\}$ . Let  $X_n^1, \dots, X_n^d$  be i.i.d. vertices of  $V(P_n)$  chosen with the conditional probability

$$\Pr [X_n^1 = v_i | \mathcal{F}_n] = \frac{\deg v_i(n)}{2n}, \quad v_i \in V(P_n),$$

where  $\deg v_i(n)$  is the degree of  $v_i$  in  $P_n$  (note that,  $\sum_{v_i} \deg v_i(n) = 2n$ ).

Second, create a new edge between  $v_{n+2}$  and  $Y_n$ , where  $Y_n$  is whichever of  $X_n^1, \dots, X_n^d$  has the largest degree. In the case of a tie, choose according to an independent fair coin toss (this choice will not affect the degree distribution). This model is called the *max-choice preferential attachment tree model*. For any fixed  $k \in \mathbb{N}$ , let  $M_1(n) \geq M_2(n) \geq \dots \geq M_k(n)$  be the degrees of  $k$  highest degree vertices at time  $n$  (if there are less than  $k$  vertices at time  $n$  put  $M_k(n) = 1$ ).

Let us formulate our main theorem:

**Theorem 1.** For  $d \in \mathbb{N}$ ,  $d > 2$ ,  $k \in \mathbb{N}$ ,  $k > 1$  and any  $\epsilon > 0$ ,

$$\lim_{n \rightarrow \infty} \mathbb{P}(n^{c^{d-1}d/2-\epsilon} < M_k(n) < n^{c^{d-1}d/2+\epsilon}) = 1,$$

where  $c = 1 - x^*/2$ ,  $x^*$  is the unique positive solution of the equation  $1 - (1 - x/2)^d = x$  in the interval  $0 \leq x \leq 1$ .

Our proof is based on the existence of the  $k$ -th *persistent hub*, i.e. a single vertex that in some finite random time becomes the  $k$ -th highest degree vertex for all time after. Using this, instead of analyzing the  $k$ -th highest degree over all vertices we effectively only need to analyze the degree of just one vertex. The existence of the  $k$ -th persistent hub is stated in the following result.

**Proposition 1.** There exist random variables  $N_l$  and  $K_l$ ,  $1 \leq l \leq k$ , that are finite almost surely so that at any time  $n \geq N_l$ ,  $\deg v_{K_l}(n) = M_l(n)$  and  $M_1(n) > M_2(n) > \dots > M_l(n) > \deg v_i(n)$  for any  $i \neq K_1, \dots, K_l$ .

The purpose of this proposition is to simplify analysis of the dynamics of  $M_k(n)$ . Indeed, let  $L_k(n)$  be the number of vertices at time  $n$  that has degree equal to  $M_k(n)$ . The effect of Proposition 1 is that for some random and sufficiently large  $N_k < \infty$ ,  $L_k(n) = 1$  for all  $n \geq N_k$ .

If  $M_{k-1}(n) = M_k(n)$ , then  $M_k(n+1) = M_k(n)$ , cause  $M_{k-1}(n)$  and  $M_k(n)$  could not be increased at the same time and we should increase  $M_{k-1}(n)$  before  $M_k(n)$ . If  $M_{k-1}(n) > M_k(n)$ , to increase  $M_k(n)$  we need to draw an edge to a vertex with the degree  $M_k(n)$ . Therefore the dynamics of  $M_k(n)$  is given by the formula

$$M_k(n+1) - M_k(n) = \mathbf{1}\{\deg Y_{n+1}(n) = M_k(n), M_{k-1}(n) > M_k(n)\}, \text{ with}$$

$$\begin{aligned} \mathbb{E}(M_k(n+1) - M_k(n) | \mathcal{F}_n) &= \\ &= \left( \hat{c}_k^d(n) - \left( \hat{c}_k(n) - \frac{M_k(n)L_k(n)}{2n} \right)^d \right) \mathbf{1}\{M_{k-1}(n) > M_k(n)\}, \end{aligned}$$

where

$$\hat{c}_l(n) = 1 - \frac{1}{2n} \sum_{i=1}^{l-1} M_i(n), \quad 1 \leq l \leq k, \quad n \in \mathbb{N}.$$

Note that  $\hat{c}_l(n) \geq 0$  cause the sum of the degrees is  $2n$ . From here, we will refer to  $\hat{c}_k^d(n) - \left( \hat{c}_k(n) - \frac{M_k(n)L_k(n)}{2n} \right)^d$  as  $p_{n,k}$ . Note that cause  $M_k(n+1) - M_k(n)$  could only take values 0 and 1, if  $M_{k-1}(n) > M_k(n)$  then  $p_{n,k}$  equals to the probability to increase  $k$ -th maximal degree at the  $n$ -th step conditional on  $\mathcal{F}_n$ .

Before starting the proof, let us describe its structure and main ideas. We will prove Proposition 1 and Theorem 1 using an induction over  $k$ . To do so, we consider them as independent theorems for each  $k$ . For  $k = 1$ , the convergence  $\frac{M_1(n)}{n} \rightarrow x^*$  almost surely and the existence of the persistent hub were proven in [11]. We will fix  $k_0 > 1$  and, using statements of Theorem 1 and Proposition 1 for  $k < k_0$  (from here we refer to them as induction hypothesis), prove them for  $k = k_0$ . In Section 2, we prove initial estimates using Theorem 1 for  $k < k_0$ . In Section 3, we use these estimates to prove the existence of the persistent hub and, so, prove Proposition 1 for  $k = k_0$ . In Section 4, we use Proposition 1 along with lemmas from Section 2 to prove Theorem 1 for  $k = k_0$ .

## 2. Initial estimates

We assume that Theorem 1 and Proposition 1 hold for  $k < k_0$ . In this section, we obtain an initial estimate on  $M_{k_0}(n)$  along with some technical lemmas.

Recall that  $c = 1 - x^*/2$ , where  $x_*$  is the solution of the equation  $1 - (1 - x/2)^d = x$  in the interval  $0 \leq x \leq 1$ . Define the function

$$f(x, y) = \frac{1}{2} \sum_{i=0}^{d-1} y^{d-i-1} (y - x/2)^i, \quad x, y \in \mathbb{R}_+.$$

Note that  $f(x, y) = \frac{y^d - (y - x/2)^d}{x}$  for  $x \neq 0$ . We will need the following estimates.

**Lemma 1.**  $f(x, y) < 1$  for  $y^{d-1} < 2/d$  and  $0 \leq x \leq 2y$ .

*Proof.*  $f(0, y) = d/2y^{d-1} < 1$ . Since  $f(x, y)$  is the decreasing function over  $x$  for  $0 \leq x \leq 2y$ , we have that  $f(x, y) < 1$  for corresponding  $x$ .  $\square$

**Lemma 2.**  $c^{d-1} < 2/d$ .

*Proof.* Note that  $x^* = 1 - (1 - x^*/2)^d = 1 - c^d$ . Therefore,

$$c^{d-1} = \frac{c^d}{c} = \frac{1 - x^*}{1 - x^*/2} < 2(1 - x^*).$$

Now show that  $x^* > 1 - 1/d$ . Due to convexity of  $1 - (1 - x/2)^d$  on  $[0, 2]$ , it is enough to show that  $1 - (1 - (1 - 1/d)/2)^d - (1 - 1/d) > 0$ :

$$1 - (1 - (1 - 1/d)/2)^d - (1 - 1/d) = 1/d - \left(\frac{d+1}{2d}\right)^d = 1/d - (1/2)^d(1 + 1/d)^d > 0,$$

for  $d > 2$  (could be easily proved by an induction starting with  $d = 3$ ).  $\square$

We will frequently use the following lemma of [5].

**Lemma 3.** *Suppose that a sequence of positive numbers  $r_n$  satisfies*

$$r_{n+1} = r_n \left(1 + \frac{\alpha}{n+x}\right), \quad n \geq k$$

for fixed  $\alpha > 0$ ,  $k > 0$  and  $x$ . Then  $r_n/n^\alpha$  has a positive limit.

Now, we formulate our initial estimate.

**Lemma 4.** *There is  $\gamma > 0$  (which do not depend on  $k_0$ ) such that, with probability 1,  $\inf_n M_{k_0}(n)/n^\gamma > 0$ .*

*Proof.* For fixed  $n_0 \in \mathbb{N}$ , define

$$C_{n+1} = \frac{4n}{4n - (c - \delta)^{d-1}} C_n = \left(1 + \frac{(c - \delta)^{d-1}}{4n - (c - \delta)^{d-1}}\right) C_n, \quad n \geq n_0,$$

with  $C_{n_0} = 1$  and  $0 < \delta < c$ . By Lemma 3, we have that  $C_n/n^{(c-\delta)^{d-1}/4}$  converges to a positive limit.

Introduce events

$$Q_{k_0}(n_0) = \{N_{k_0-1} < n_0, \hat{c}_{k_0}(n) > c - \delta \forall n > n_0\}.$$

By the induction hypothesis ( $N_{k_0-1} < \infty$  almost surely and  $\hat{c}_{k_0}(n) \rightarrow c$  in probability),  $\mathbb{P}(Q_{k_0}(n_0)) \rightarrow 1$  when  $n_0 \rightarrow \infty$ . Introduce Markov moments

$$\lambda_{k_0}(n_0) = \inf\{n > n_0 : L_k(n) > 1 \text{ for some } k < k_0 \text{ or } \hat{c}_{k_0}(n) < c - \delta\}.$$

Note that  $\lambda_{k_0}(n_0) = \infty$  on  $Q_{k_0}(n_0)$ . Put  $A_{k_0}(n) = C_n/M_{k_0}(n)$ . We will prove that  $A_{k_0}(n \wedge \lambda_{k_0}(n_0))$  (where  $x \wedge y = \min(x, y)$ ) is a supermartingale for  $n > n_0$ . Hence by Doob's theorem (Corollary 3, p. 509 of [14]) it converges almost surely to some finite limit. Therefore, there is a random variable  $B_{k_0, n_0}$ , which is positive on  $Q_{k_0}(n_0)$ , such

that  $\frac{M_{k_0}(n)}{n^\gamma} \geq B_{k_0, n_0}$  almost surely for  $\gamma = (c - \delta)^{d-1}/4$  and  $n \geq n_0$ . Consequently, we have

$$\mathbb{P}\left(\inf_{n \in \mathbb{N}} \frac{M_{k_0}(n)}{n^\gamma} > 0\right) \geq \mathbb{P}\left(\inf_{n \in \mathbb{N}} \frac{M_{k_0}(n)}{n^\gamma} > 0, Q_{k_0}(n_0)\right) = \mathbb{P}(Q_{k_0}(n_0)) \rightarrow 1.$$

Now prove that  $A_{k_0}(n \wedge \lambda_{k_0}(n_0))$  is a supermartingale, which concludes our proof.

Recall that if  $M_{k_0-1}(n) > M_{k_0}(n)$  (holds if  $L_{k_0-1} = 1$ , in particular for  $n_0 < n < \lambda_{k_0}(n_0)$ ) then  $p_{n, k_0}$  equals to the probability to increase  $k_0$ -th maximal degree at the  $n$ -th step conditional on  $\mathcal{F}_n$ . Note that

$$\begin{aligned} p_{n, k_0} &= \hat{c}_{k_0}^d(n) - \left(\hat{c}_{k_0}(n) - \frac{M_{k_0}(n)L_{k_0}(n)}{2n}\right)^d \geq \hat{c}_{k_0}^d(n) \left(1 - \left(1 - \frac{M_{k_0}(n)}{2n\hat{c}_{k_0}(n)}\right)^d\right) \geq \\ &\geq \hat{c}_{k_0}^d(n) \left(1 - \left(1 - \frac{M_{k_0}(n)}{2n\hat{c}_{k_0}(n)}\right)^2\right) = \hat{c}_{k_0}^d(n) \left(\frac{M_{k_0}(n)}{n\hat{c}_{k_0}(n)} - \frac{(M_{k_0}(n))^2}{(2n\hat{c}_{k_0}(n))^2}\right) = \\ &= \hat{c}_{k_0}^{d-1}(n) \frac{M_{k_0}(n)}{n} \frac{4n\hat{c}_{k_0}(n) - M_{k_0}(n)}{4n\hat{c}_{k_0}(n)} \geq \hat{c}_{k_0}^{d-1}(n) \frac{2M_{k_0}(n)}{4n} = \hat{c}_{k_0}^{d-1}(n) \frac{M_{k_0}(n)}{2n}. \end{aligned}$$

By definition of  $p_{n, k_0}$ , for  $1/M_{k_0}(n+1)$  we get

$$\begin{aligned} \mathbb{E}(1/M_{k_0}(n+1)|\mathcal{F}_n) &= \\ &= \mathbb{E}\left(\frac{\mathbf{1}\{M_{k_0}(n+1) = M_{k_0}(n) + 1\}}{M_{k_0}(n) + 1} + \frac{\mathbf{1}\{M_{k_0}(n+1) = M_{k_0}(n)\}}{M_{k_0}(n)} \middle| \mathcal{F}_n\right) = \\ &= \left(\frac{p_{n, k_0} \mathbf{1}\{M_{k_0-1}(n) > M_{k_0}(n)\}}{M_{k_0}(n) + 1} + \frac{1 - p_{n, k_0} \mathbf{1}\{M_{k_0-1}(n) > M_{k_0}(n)\}}{M_{k_0}(n)}\right) = \\ &= \left(\frac{M_{k_0}(n) + 1 - p_{n, k_0} \mathbf{1}\{M_{k_0-1}(n) > M_{k_0}(n)\}}{(M_{k_0}(n) + 1)M_{k_0}(n)}\right). \end{aligned}$$

Therefore, if  $n_0 < n < \lambda_{k_0}(n_0)$ , then

$$\begin{aligned} \mathbb{E}(1/M_{k_0}(n+1)|\mathcal{F}_n) &= \frac{M_{k_0}(n) + 1 - p_{n, k_0}}{M_{k_0}(n)(M_{k_0}(n) + 1)} = \frac{1}{M_{k_0}(n)} \left(1 - \frac{p_{n, k_0}}{M_{k_0}(n) + 1}\right) \leq \\ &\leq \frac{1}{M_{k_0}(n)} \left(1 - \frac{p_{n, k_0}}{2M_{k_0}(n)}\right) \leq \frac{1}{M_{k_0}(n)} \left(1 - \frac{\hat{c}_{k_0}^{d-1}(n)}{4n}\right) \leq \\ &\leq \frac{1}{M_{k_0}(n)} \left(1 - \frac{(c - \delta)^{d-1}}{4n}\right), \end{aligned}$$

which concludes the proof.  $\square$

### 3. Persistent hub

We assume that Theorem 1 and Proposition 1 holds for  $k < k_0$ . In this section, we prove Proposition 1 for  $k = k_0$  under this assumption. Our method of the proof bases on the comparison of our model with the standart preferential attachment model, and we use the technique of [5] developed for the last one. We divide the proof of Proposition 1 into two parts. First, we prove that degrees of only finite number of vertices could at some time become  $k$ -th maximal. Second, we prove that two vertices could have a  $k$ -th highest degree at the same time only for finite number of time moments.

Let us introduce some notations:

$$\chi_k(n) = \min\{i \geq n : \deg v_n(i) = M_k(i)\},$$

$$U_k = \sum_{n=1}^{\infty} \mathbf{1}\{\chi_k(n) < \infty\},$$

$$\psi_{i,j}(n) = \min_{l \geq n} \{\deg v_i(l) = \deg v_j(l)\}.$$

Here  $U_k$  is the number of vertices (of  $V$ ) whose degrees were  $k$ -th maximal at some moments,  $\chi_k(n)$  is the moment it happens for the vertex  $v_n$ .

**Lemma 5.**  *$U_k$  is finite almost surely.*

To prove the lemma, we first need a result (which is stated below) from [11] on a random walk that describes the evolution of degrees of two vertices in the preferential attachment model without choices.

Let  $T_n = T_n(n_0, A_{n_0}, B_{n_0}) = (A_n, B_n)$  for  $n \geq n_0$  be random walks on  $\mathbb{Z}^2$  started from some point  $(A_{n_0}, B_{n_0})$  that at time  $n$  move one step right or one step up with the conditional probabilities  $\frac{A_n}{A_n+B_n}$  and  $\frac{B_n}{A_n+B_n}$  respectively. Also, introduce the stopping times  $\pi(i, j) = \min\{n \geq n_0 : A_n = B_n | A_{n_0} = i, B_{n_0} = j\}$  and the function  $q(i, j) = \mathbb{P}(\pi(i, j) < \infty)$ . Although, the arguments of  $q$  and  $\pi$  are integers, sometimes in estimates we will write noninterges in arguments meaning the value of the floor function of it.

Lemma 4.2 from [11] stated that

**Lemma 6.** *The following inequality holds for any positive integers  $i$  and  $j$*

$$\mathbb{P}(\psi_{i,j}(n) < \infty | \mathcal{F}_n) \leq q(\deg v_i(n), \deg v_j(n)).$$

Let us prove Lemma 5.

*Proof.* By Lemma 4, we get  $M_k(n) \geq Mn^\gamma$  for some random  $M > 0$  almost surely. Hence, at time  $n$  there are at least  $k$  vertices  $v_{i_1}, \dots, v_{i_k}$  with degrees not less then  $Mn^\gamma$  with probability 1. A degree of the vertex  $v_{n+1}$  could become  $k$ -th maximal only if at some moment  $\tilde{n} > n$  its degree becomes higher than at least one of the degrees  $\deg v_{i_1}(\tilde{n}), \dots, \deg v_{i_k}(\tilde{n})$ . Due to Lemma 6 (as in [11]), we could construct  $k$  versions  $\pi_l(i, j)$ ,  $1 \leq l \leq k$ , of  $\pi(i, j)$ , such that

$$\mathbf{1}\{\chi_k(n+1) < \infty\} \leq \sum_{l=1}^k \mathbf{1}\{\psi_{i_l, n+1}(n) < \infty\} \leq$$

$$\leq \sum_{l=1}^k \mathbf{1}\{\pi_l(\deg v_{i_l}(n), 1) < \infty\} \leq \sum_{l=1}^k \mathbf{1}\{\pi_l(Mn^\gamma, 1) < \infty\} \quad a.s.$$

Fix  $C > 0$ . Then

$$\begin{aligned} U_k \mathbf{1}\{M > C\} &= \sum_{n=1}^{\infty} \mathbf{1}\{\chi_k(n) < \infty\} \mathbf{1}\{M > C\} < \\ &< \sum_{n=1}^{\infty} \sum_{l=1}^k \mathbf{1}\{\pi_l(Mn^\gamma, 1) < \infty\} \mathbf{1}\{M > C\} \leq \sum_{n=1}^{\infty} \sum_{l=1}^k \mathbf{1}\{\pi_l(Cn^\gamma, 1) < \infty\} \mathbf{1}\{M > C\} \leq \\ &\leq \sum_{n=1}^{\infty} \sum_{l=1}^k \mathbf{1}\{\pi_l(Cn^\gamma, 1) < \infty\}. \end{aligned}$$

Corollary 15 of [5] gives us the following estimate:

$$q(i, 1) \leq \frac{Q(i)}{2^i} \text{ for any integer } i$$

for some polynomial function  $Q(x)$ . Therefore, the expectations

$$\mathbb{E} \mathbf{1}\{\pi_l(Cn^\gamma, 1) < \infty\} = q(Cn^\gamma, 1) \leq \frac{Q(Cn^\gamma)}{2^{Cn^\gamma}}$$

forms a convergent series, and the last sum is finite almost surely by Borel-Cantelli Lemma. Since  $M > 0$  with probability 1,

$$\mathbb{P}(U_k < \infty) = \mathbb{P}(\{U_k < \infty\} \bigcup_{n \in \mathbb{N}} \{M > 1/n\}) = 1.$$

□

Now let  $J_k$  denote the set of vertices whose degrees become  $k$ -th maximal at some moment. According to Lemma 5,  $J_k$  is finite almost surely. Introduce random moments

$$\zeta_l(v_i, v_j) = \inf\{n > \zeta_{l-1}(v_i, v_j) :$$

$$\deg v_i(n-1) \neq \deg v_j(n-1) \text{ and } \deg v_i(n) = \deg v_j(n)\}, \quad \zeta_0(v_i, v_j) = 0,$$

$$N(v_i, v_j) = \sup\{l : \zeta_l(v_i, v_j) < \infty\},$$

$$\xi_k = \sup\{\zeta_{N(v_i, v_j)}(v_i, v_j) | v_i \in J_k, v_j \in J_k\}.$$

Note that almost sure finiteness of  $\xi_k$  implies Proposition 1 cause any vertex that become  $k$ -th maximal at any time is in  $J_k$ , and an order of degrees of vertices from  $J_k$  does not change after the moment  $\xi_k$ . Thus, to complete the proof of Proposition 1 we need the following lemma:

**Lemma 7.**  $\xi_k$  is finite almost surely.

*Proof.* Since  $J_k$  is finite almost surely, it is enough to prove that for any  $v_i, v_j \in V$   $N(v_i, v_j)$  is finite almost surely. To do so we will use the random walk  $T_n$  with  $n_0 = \max\{i + 1, j + 1\}$ ,  $A_{n_0} = \deg v_i(n_0)$ ,  $B_{n_0} = \deg v_j(n_0)$ . Let  $R(n_0, i, j)$  be the number of times  $n > n_0$  such that  $A_n = B_n$ , and let  $n_0 \leq \rho_1(i, j) < \rho_2(i, j) < \dots$  be moments when either  $\deg v_i$  or  $\deg v_j$  is changed. Then due to the coupling used in the proof of Lemma 4.2 from [11] there is version of  $T$ , such that  $\min\{\deg v_i(\rho_n), \deg v_j(\rho_n)\}$  is dominated by  $\min\{A_n, B_n\}$  for  $n \geq n_0$ , which implies  $N(v_i, v_j) \leq R$  (since  $A_n + B_n = \deg v_i(\rho_n) + \deg v_j(\rho_n)$ ).

It is a standard fact about Pólya urn model that if  $T_n = (A_n, B_n)$  starts from a point  $(a, b)$ , then the fraction  $A_n/(A_n + B_n)$  tends in law to a random variable  $H(a, b)$  as  $n$  tends to infinity, where  $H(a, b)$  has beta probability distribution:

$$H(a, b) \sim \text{Beta}(a, b).$$

(See, e.g., Theorem 3.2 in [10] or Section 4.2 in [6]). Thus, the limit of  $A_n/(A_n + B_n)$  exists almost surely, and it takes the value  $1/2$  with probability 0 for any starting point of the process  $T$ . Hence, this fraction can be equal to  $1/2$  only finitely many times almost surely, and so  $R$  is finite almost surely, which completes the proof.  $\square$

#### 4. Final result

Fix  $0 < \delta < 2/d - c^{d-1}$  (by Lemma 2,  $2/d > c^{d-1}$ ). For any fixed  $n_0 > 0$ , we introduce the events

$$D_{k_0}(n_0, \delta) = \{L_l(n) = 1, c - \delta < \hat{c}_{k_0}(n) < c + \delta, M_{k_0}(n) > n^\gamma/n_0 \forall n \geq n_0 \forall k \leq k_0\},$$

and the Markov moment

$$\begin{aligned} \eta_{k_0}(n_0, \delta) &= \inf\{n \geq n_0 : L_k(n) > 1 \text{ for some } k \leq k_0, \text{ or} \\ &\hat{c}_{k_0}(n) > c + \delta, \text{ or } \hat{c}_{k_0}(n) < c - \delta, \text{ or } M_{k_0}(n) \leq n^\gamma/n_0\}. \end{aligned}$$

Note that by the induction assumption for  $k < k_0$ , Proposition 1 and Lemma 4 (both for  $k = k_0$ ) we have that

$$\mathbb{P}(D_{k_0}(n_0, \delta)) = \mathbb{P}(\eta_{k_0}(n_0, \delta) = \infty) \rightarrow 1 \text{ as } n_0 \rightarrow \infty.$$

Now, let prove Theorem 1 for  $k = k_0$

**Lemma 8.** *With probability 1,  $M_k(n)/n \rightarrow 0$ .*

*Proof.* Recall that if  $M_{k-1}(n) > M_k(n)$  (in particular, on  $D_k(n_0, \delta)$ ) then  $p_{n,k}$  equals to the conditional probability to increase  $M_k(n)$  conditional on  $\mathcal{F}_n$ . Note that for  $n$  such that  $n_0 \leq n \leq \eta_{k,C}$ ,

$$p_{n,k} = \hat{c}_k^d(n) - \left( \hat{c}_k(n) - \frac{M_k(n)}{2n} \right)^d = \frac{M_k(n)}{2n} \left( \sum_{i=0}^{d-1} \hat{c}_k^{d-i-1}(n) \left( \hat{c}_k(n) - \frac{M_k(n)}{2n} \right)^i \right).$$

Hence,  $\frac{p_{n,k}}{M_k(n)} = \frac{1}{n} f\left(\frac{M_k(n)}{n}, \hat{c}_k(n)\right)$ . From Lemmas 1, 2, it follows that for any small enough  $\delta > 0$  there is  $\beta > 0$  so that  $f(x, y) < 1 - \beta$  if  $y < c + \delta$  for  $0 \leq x \leq 2y$ .



Consider the expectation:

$$\mathbb{E} \left( \frac{M_k(n+1)}{M_k(n)} \middle| \mathcal{F}_n \right) = \frac{p_{n,k}(M_k(n)+1)}{M_k(n)} + 1 - p_{n,k} = 1 + \frac{p_{n,k}}{M_k(n)}.$$

Therefore, for small enough  $\delta > 0$  there is  $\beta > 0$  such that  $\mathbb{E}(M_k(n+1)|\mathcal{F}_n) < (1 + (1-\beta)/n)M_k(n)$  for  $n_0 \leq n < \eta_k(n_0, \delta)$ . Set  $A_k(n) = M_k(n)/C_{n,k}$ , where  $C_{n+1,k} = (1 + (1-\beta)/n)C_{n,k}$ ,  $n \geq n_0$ ,  $C_{n_0,k} = 1$ . We have that

$$\begin{aligned} & \mathbb{E} \left( \frac{A_k(n+1 \wedge \eta_k(n_0, \delta))}{A_k(n \wedge \eta_k(n_0, \delta))} \middle| \mathcal{F}_n \right) = \\ & \mathbb{E} \left( \frac{A_k(n+1)}{A_k(n)} \mathbf{1}\{n+1 \leq \eta_k(n_0, \delta)\} + \mathbf{1}\{n+1 > \eta_k(n_0, \delta)\} \middle| \mathcal{F}_n \right) = \\ & \mathbf{1}\{\eta_k(n_0, \delta) > n\} \mathbb{E} \left( \frac{A_k(n+1)}{A_k(n)} \middle| \mathcal{F}_n \right) + \mathbf{1}\{\eta_k(n_0, \delta) \leq n\} = \\ & \mathbf{1}\{\eta_k(n_0, \delta) > n\} \frac{C_{n,k}}{C_{n+1,k}} \mathbb{E} \left( \frac{M_k(n+1)}{M_k(n)} \middle| \mathcal{F}_n \right) + \mathbf{1}\{\eta_k(n_0, \delta) \leq n\} \leq \\ & \mathbf{1}\{\eta_k(n_0, \delta) > n\} \frac{1 + (1-\beta)/n}{1 + (1-\beta)/n} + \mathbf{1}\{\eta_k(n_0, \delta) \leq n\} = 1. \end{aligned}$$

Thus,  $A_k(n \wedge \eta_k(n_0, \delta))$  is a supermartingale. By Lemma 3, we have that  $C_{n,k}n^{-1+\beta}$  converges to a positive limit. Therefore, by Doob's theorem we have that  $A_k(n \wedge \eta_k(n_0, \delta))$  tends to a finite limit with probability 1, and, in particular, there is a random constant  $B_k > 0$  so that  $M_k(n \wedge \eta_k(n_0, \delta)) \leq B_k n^{1-\beta}$  almost surely. Thus,  $M_k(n \wedge \eta_k(n_0, \delta))/n \rightarrow 0$  almost surely as  $n \rightarrow \infty$ , and, since  $\mathbb{P}(\eta_k(n_0, \delta) = \infty) \rightarrow 1$  as  $n_0 \rightarrow \infty$ ,  $M_k(n)/n \rightarrow 0$  almost surely as  $n \rightarrow \infty$ .  $\square$

Now, consider the expectation for  $n_0 \leq n \leq \eta_k(n_0, \delta)$  and some  $0 < \alpha < 1$

$$\begin{aligned} \mathbb{E} \left( \frac{M_k(n+1)/(n+1)^\alpha}{M_k(n)/n^\alpha} \middle| \mathcal{F}_n \right) &= \frac{M_k(n) + \hat{c}_k^d(n) - \left( \hat{c}_k(n) - \frac{M_k(n)}{2n} \right)^d}{M_k(n)} \frac{n^\alpha}{(n+1)^\alpha} = \\ &= \left( 1 + \frac{1}{2n} \left( d\hat{c}_k^{d-1}(n) + \sum_{i=1}^{d-1} (-1)^i \binom{i}{d} \hat{c}_k^{d-1-i}(n) \left( \frac{M_k(n)}{2n} \right)^i \right) \right) \frac{1}{(1+1/n)^\alpha}. \end{aligned}$$

By the induction assumption and Lemma 8,

$$\Delta_k(n) = d\hat{c}_k^{d-1}(n) + \sum_{i=1}^{d-1} (-1)^i \binom{i}{d} \hat{c}_k^{d-1-i}(n) \left( \frac{M_k(n)}{2n} \right)^i \rightarrow c^{d-1}d \text{ a.s. as } n \rightarrow \infty.$$

In particular, for any  $\epsilon > 0$

$$\mathbb{P}(\Delta_k(n) < c^{d-1}d + \epsilon \text{ for } n > n_0) \rightarrow 1 \text{ and}$$

$$\mathbb{P}(\Delta_k(n) > c^{d-1}d - \epsilon \text{ for } n > n_0) \rightarrow 1 \text{ as } n_0 \rightarrow \infty.$$

Therefore,

$$\begin{aligned}
& \mathbb{P} \left( \mathbb{E} \left( \frac{M_k(n+1)/(n+1)^{c^{d-1}d/2-\epsilon}}{M_k(n)/n^{c^{d-1}d/2-\epsilon}} \middle| \mathcal{F}_n \right) > 1, \text{ for } n > n_0 \right) \geq \\
& \geq \mathbb{P} \left( \mathbb{E} \left( \frac{M_k(n+1)/(n+1)^{c^{d-1}d/2-\epsilon}}{M_k(n)/n^{c^{d-1}d/2-\epsilon}} \middle| \mathcal{F}_n \right) > 1, \right. \\
& \left. \Delta_k(n) > c^{d-1}d - \epsilon \text{ for } n > n_0, \eta_k(n_0, \delta) = \infty \right) \geq \\
& \geq \mathbb{P} \left( \frac{1 + (c^{d-1}d/2 - \epsilon/2)/n}{(1 + 1/n)^{c^{d-1}d/2-\epsilon}} > 1, \right. \\
& \left. \Delta_k(n) > c^{d-1}d - \epsilon, \text{ for } n > n_0, \eta_k(n_0, \delta) = \infty \right) \rightarrow 1 \text{ as } n_0 \rightarrow \infty \text{ and} \\
& \mathbb{P} \left( \mathbb{E} \left( \frac{M_k(n+1)/(n+1)^{c^{d-1}d/2+\epsilon}}{M_k(n)/n^{c^{d-1}d/2+\epsilon}} \middle| \mathcal{F}_n \right) < 1, \text{ for } n > n_0 \right) \geq \\
& \geq \mathbb{P} \left( \mathbb{E} \left( \frac{M_k(n+1)/(n+1)^{c^{d-1}d/2+\epsilon}}{M_k(n)/n^{c^{d-1}d/2+\epsilon}} \middle| \mathcal{F}_n \right) < 1, \right. \\
& \left. \Delta_k(n) < c^{d-1}d + \epsilon \text{ for } n > n_0, \eta_k(n_0, \delta) = \infty \right) \geq \\
& \geq \mathbb{P} \left( \frac{1 + (c^{d-1}d/2 + \epsilon/2)/n}{(1 + 1/n)^{c^{d-1}d/2+\epsilon}} < 1, \right. \\
& \left. \Delta_k(n) < c^{d-1}d + \epsilon \text{ for } n > n_0, \eta_k(n_0, \delta) = \infty \right) \rightarrow 1
\end{aligned}$$

as  $n_0 \rightarrow \infty$ . Introduce Markov moments

$$\begin{aligned}
\nu_{k,n_0,\epsilon} = \inf \left\{ n > n_0 : \mathbb{E} \left( \frac{M_k(n+1)/(n+1)^{c^{d-1}d/2-\epsilon}}{M_k(n)/n^{c^{d-1}d/2-\epsilon}} \middle| \mathcal{F}_n \right) \leq 1, \text{ or} \right. \\
\left. \mathbb{E} \left( \frac{M_k(n+1)/(n+1)^{c^{d-1}d/2+\epsilon}}{M_k(n)/n^{c^{d-1}d/2+\epsilon}} \middle| \mathcal{F}_n \right) \geq 1 \right\}.
\end{aligned}$$

Note that  $\mathbb{P}(\nu_{k,n_0,\epsilon} = \infty) \rightarrow 1$  as  $n_0 \rightarrow \infty$ . Let

$$A_k(n) = \frac{M_k(n)}{n^{c^{d-1}d/2+\epsilon/2}} \text{ and } B_k(n) = \frac{n^{c^{d-1}d/2-\epsilon/2}}{M_k(n)}.$$

Then  $A_k(n \wedge \nu_{k,n_0,\epsilon})$  and  $B_k(n \wedge \nu_{k,n_0,\epsilon})$  are supermartingales, and from Doob's theorem,

$$\frac{M_k(n)}{n^{c^{d-1}d/2-\epsilon}} \rightarrow \infty \text{ and } \frac{M_k(n)}{n^{c^{d-1}d/2+\epsilon}} \rightarrow 0 \text{ almost surely,}$$

which implies our theorem.

### Acknowledgements

The author is grateful to Professor Itai Benjamini for proposing the model and to Maksim Zhukovskii for helpful discussions.

### References

- [1] Barabási A., Albert R. Emergence of scaling in random networks. *Science*, 1999, vol. 286(5439), pp. 509–512.
- [2] D’Souza R.M., Krapivsky P.L., Moor C. The power of choice in growing trees. *The European Physical Journal B*, 2007, vol. 59(4), pp. 535–543.
- [3] Dommers S., Hofstad R., Hooghiemstra G. Diameters in preferential attachment models. *Journal of Statistical Physics*, 2010, vol. 139(1), pp. 72–107.
- [4] Fenner T., Flaxman A., Frieze A. High degree vertices and eigenvalues in the preferential attachment graph. *Internet Mathematics*, 2005, vol. 2(1), pp. 1–19.
- [5] Galashin P.A. *Existence of a persistent hub in the convex preferential attachment model*. ArXiv e-prints. October 2013. URL: <https://arxiv.org/abs/1310.7513>
- [6] Johnson N.L., Kotz S. *Urn models and their application*. John Wiley and Sons, New York, 1977.
- [7] Krapivsky P.L., Redner S. Choice-driven phase transition in complex networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2014.
- [8] Leyvraz F., Krapivsky P.L., Redner S. Connectivity of growing random networks. *Physical Review Letters*, 2000, vol. 85, pp. 4629–4632.
- [9] Kuba M., Mahmoud H., Panholzer A. Analysis of a generalized friedman’s urn with multiple drawings. *Discrete Applied Mathematics*, 2013, vol. 161(18), pp. 2968–2984.
- [10] Mahmoud H. *Polya urn models*. Chapman and Hall/CRC, 2009.
- [11] Malyshkin Y., Paquette E. The power of choice combine with preferential attachment. *Electronic Communication in Probability*, 2014, vol. 19(44), pp. 1–13.
- [12] Malyshkin Y., Paquette E. The power of choice over preferential attachment. *Latin American Journal of Probability and Mathematical Statistics*, 2015, vol. 12(2), pp. 903–915.
- [13] Tamás F. Móri. The maximum degree of the Barabási-Albert random tree. *Combinatorics, Probability and Computing*, 2005, vol. 14(3), pp. 339–348.
- [14] Shiryaev A.N. *Probability*. Second edition. Springer, 1996. 621 p.

**Bibliographic citation**

Malyshkin Yu. High degree vertices in the power of choice model combined with preferential attachment. *Vestnik TvgU. Seriya: Prikladnaya Matematika* [Herald of Tver State University. Series: Applied Mathematics], 2017, no. 1, pp. 31–43.

## ВЕРШИНЫ С НАИБОЛЬШИМИ СТЕПЕНЯМИ В МОДЕЛИ ПРЕДПОЧТИТЕЛЬНОГО ПРИСОЕДИНЕНИЯ С ВЫБОРОМ

Малышкин Ю.А.

Тверской государственной университет, г. Тверь

---

*Поступила в редакцию 30.08.2016, после переработки 17.03.2017.*

---

В работе найдена асимптотика первых  $k$  максимумов распределения степеней в графе для модели предпочтительного присоединения с выбором. В данной модели на каждом шаге добавляется одна вершина. Затем мы случайным образом выбираем  $d$  ( $d > 2$ ) вершин, и проводим ребро из новой вершины в вершину с наибольшей (из выбранных вершин) степенью. Известно, что в данной модели максимальная степень вершины в графе растет линейно относительно общего числа вершин, в то время как в моделях предпочтительного присоединения без выбора первые  $k$  максимумов распределения степеней вершин растут сублинейно с одинаковым показателем. Доказано, что степени  $k$ -ых по величине степеней вершин растут сублинейно относительно размера графа. Доказательство использует существование в графе выделенных вершин и мартингальную технику.

**Ключевые слова:** случайные графы, предпочтительное присоединение, выбор.

### Библиографическая ссылка

Malyshkin Yu. High degree vertices in the power of choice model combined with preferential attachment. *Vestnik TvGU. Seriya: Prikladnaya Matematika* [Herald of Tver State University. Series: Applied Mathematics], 2017, no. 1, pp. 31–43.

### Сведения об авторах

1. **Малышкин Юрий Адреевич**

старший преподаватель кафедры прикладной физики Тверского государственного университета.

*Россия, 170100, г. Тверь, ул. Желябова, д. 33, ТвГУ.*