

EXISTENCE OF A PERSISTENT HUB IN THE CONVEX PREFERENTIAL ATTACHMENT MODEL

PAVEL GALASHIN

ABSTRACT. A vertex of a randomly growing graph is called a persistent hub if at all but finitely many moments of time it has the maximal degree in the graph. We establish the existence of a persistent hub in the Barabási–Albert random graph model with probability one. We also extend this result to the class of convex preferential attachment graphs, where a vertex of degree k gets a new edge with probability proportional to some convex function of k .

1. INTRODUCTION

The preferential attachment model was introduced by R. Albert and A. L. Barabási in [3] in order to create a natural model for a dynamically growing random network with a scale-free power-law distribution of degrees of vertices. This distribution appears in many large real random graphs such as internet, social networks, etc.

Since then the model became very popular and has been investigated mathematically and empirically in many works, for example [4, 9, 11, 13, 7]. Many generalizations have been suggested: [1, 5, 12] etc.

The Barabási–Albert (BA) preferential attachment model is defined as follows:

- (1) Before the first step we have a natural number m_0 and a tree which contains one vertex v_1 and zero edges.
- (2) At the k -th step ($k \geq 1$) we attach one new vertex v_k and m_0 new edges to the graph. These edges are attached one-by-one. Every edge connects v_k to some old vertex v_i of the graph. This vertex v_i is chosen randomly, with probability proportional to

DEPARTMENT OF MATHEMATICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MA, 02139

E-mail address: galashin@mit.edu.

Date: May 16, 2015.

Key words and phrases. Preferential attachment, persistent hub, Barabási–Albert model.

its degree $\deg(v_i)$. The degrees are being refreshed after every edge attachment.

The convex preferential attachment model involves the same algorithm, but the old vertex v_i is chosen with probability proportional to $\mathcal{W}(\deg(v_i))$, where $\mathcal{W} : \mathbb{Z}_+ \rightarrow \mathbb{R}_+$ is a fixed positive, convex and unbounded function defined on the set of non-negative integers. Formal definitions of these models are given below.

1.1. The main result. Investigation of vertices of maximal degree became one of the most popular research directions in preferential attachment, because the presence of vertices with large degrees is one of the features of preferential attachment model as opposed to the classical Erdős–Rényi model. While the graph grows, different vertices can have maximal degrees at different steps.

If a vertex has the maximal degree at some step, then it is called a *hub*. If some vertex is a hub for all but finitely many steps, then it is called a *persistent hub*. This notion has appeared in the literature, see, for example, [5].

Now we can ask the following question: does the hub change infinitely many times, or does there appear a persistent hub after some number of steps?

The following theorem answers this question:

Theorem 1. *In the BA and convex preferential attachment models with probability 1 there exist numbers n and k such that at any step after the n -th step the vertex v_k has the highest degree among all vertices. In other words, the persistent hub appears with probability one.*

1.2. Previous research. S. Dereich and P. Mörters consider in [5] a model very similar to the BA preferential attachment model with $m_0 = 1$: one starts with a single vertex and a fixed concave function $f : \mathbb{Z}_+ \rightarrow \mathbb{R}_+$, such that $f(n) \leq n + 1$ for all $n \in \mathbb{Z}_+$. At the k -th step:

- one new vertex v_k is added
- for every old vertex v_i the edge e_{ik} connecting v_i to v_k is added with probability

$$\frac{f(\deg(v_i))}{k}.$$

The main difference is that here for $i \neq j$ the decisions of adding the edges e_{ik} and e_{jk} are made **independently**, while in the BA model exactly one (if $m_0 = 1$) old vertex becomes connected to the new vertex.

The question of existence of a persistent hub for this model has been answered completely in [5], but the difference between the models makes it too hard to apply those results to the BA case.

Theorem 1.7 from [5] states that a persistent hub appears if and only if

$$\sum_{k=1}^{\infty} \frac{1}{f(k)^2} < \infty.$$

This condition is of course satisfied if f is convex and unbounded. But, even though the models are different, it remains an interesting open question if the same result also holds in our situation after we drop the convexity assumption. A weaker conjecture is that a persistent hub in the BA model appears if the weight function is superlinear.

1.3. Outline of the paper. The proof of Theorem 1 assumes $m_0 = 1$ and deals with random trees. We extend it to the case of an arbitrary m_0 in the last subsection.

- (1) In Subsection 1.4 we give precise specifications of all considered models.
- (2) In Section 2 we investigate the joint behavior of two fixed vertices. The main result of this section is Proposition 3, which states that if at the n -th step there is some vertex v_k with a high degree, then with high probability the degree of v_{n+1} will be lower than the degree of v_k at every step.
- (3) In Section 3 we prove the main result:
 - (a) In Subsection 3.1 we prepare some tools to make a comparison between the convex model and the BA model.
 - (b) In Subsection 3.2 we show that the maximal degree tends to infinity fast enough, even though the degree of any fixed vertex may be bounded.
 - (c) In Subsection 3.3 we use Borel–Cantelli lemma to show that with probability 1 all but finite number of vertices never have the highest degree.
 - (d) For every pair of the remaining vertices we prove in Subsection 3.4 that there is only a finite number of steps at which their degrees coincide. This completes the proof for the case $m_0 = 1$.
 - (e) In Subsection 3.5 we modify the steps of the proof so that they work for an arbitrary m_0 .

Throughout the paper every weight function is assumed to be convex.

1.4. Definitions of the models. First we define all models for $m_0 = 1$, because the major part of the proof is concentrated around this case.

1.4.1. *Basic model for $m_0 = 1$.* Let X_n^k be the degree of the vertex v_k before the n -th step. Note that before the n -th step there are n vertices and $n - 1$ edges in the tree, and the total degree of all vertices has a very simple form:

$$\sum_{k=1}^n X_n^k = 2(n - 1).$$

By p_n^k denote the probability that the new edge at the n -th step is attached to the vertex v_k , $k \leq n$. Then

$$p_n^k := \begin{cases} 1, & k = n = 1, \\ \frac{X_n^k}{2(n-1)}, & n > 1, 1 \leq k \leq n. \end{cases}$$

Since in this paper we are not interested in the topological structure of the tree, we can just consider the Markov chain of vectors $X_n := (X_n^k)_{1 \leq k \leq n}$.

1.4.2. *Generalized model for $m_0 = 1$.* Let $\mathcal{W} : \mathbb{Z}_+ \rightarrow \mathbb{R}_+$ be a strictly positive function.

In this model, a vertex v_k of degree X_n^k has weight $\mathcal{W}(X_n^k)$, and the probability p_n^k that the new edge is attached to the vertex v_k at the n -th step is defined as a ratio of the weight of v_k to the total weight of all vertices:

$$p_n^k := \begin{cases} 1, & k = n = 1, \\ \frac{\mathcal{W}(X_n^k)}{w_n}, & n > 1, 1 \leq k \leq n. \end{cases}$$

where

$$w_n := \sum_{k=1}^n \mathcal{W}(X_n^k).$$

(Here, unlike the basic model, w_n is a random variable).

This model is also common, for example, in [5], [14] the cases of superlinear ($\mathcal{W}(n) \gg n$) and sublinear ($\mathcal{W}(n) \ll n$) preferential attachment are considered, in [15] the asymptotical degree distribution for a wide range of weight functions is given, and in [16] the Hausdorff dimension of some natural measure on the leaves of the limiting tree is evaluated.

1.4.3. *Convex model for $m_0 = 1$.* The convex model is a special case of the generalized model. Here $\mathcal{W}(n)$ must be convex and unbounded. Note that $\mathcal{W}(n)$ is not assumed to be increasing.

The convex model includes several popular special cases. We have already discussed the basic model. In [14] the case $\mathcal{W}(n) = n^p$, $p > 1$, is considered. In [12] and [11] the case $\mathcal{W}(n) = n + \beta$, $\beta > -1$ is

considered. We call this case *the linear model*. Some similar models have been considered earlier, for example, in [6] the concave preferential attachment rule is investigated.

The convexity condition is pretty mild, but, on the other hand, it is very convenient and simplifies proofs and calculations.

1.4.4. Generalized model for an arbitrary m_0 .

- We start with no vertices and no edges.
- Let $n \geq 0$ be the number of the current step. Put $N := \lfloor \frac{n}{m_0} \rfloor$.
- If $n \equiv 0 \pmod{m_0}$, then add one new vertex v_N .
- Connect the vertex v_N to exactly one of the vertices v_0, \dots, v_{N-1} , if this set is not empty.
- A vertex v_i , $0 \leq i \leq N - 1$ is chosen with probability

$$\frac{\mathcal{W}(\deg(v_i))}{\sum_{j=0}^{N-1} \mathcal{W}(\deg(v_j))}.$$

The basic, linear and convex models are just special cases of the generalized model, so it is enough to specify the generalized model for an arbitrary m_0 .

2. PAIRWISE VERTEX DEGREE ANALYSIS

In this section we investigate a random walk on the two-dimensional integer lattice. In terms of preferential attachment, we consider two fixed vertices, and we are interested only in steps at which the degree of one of these vertices increases. We obtain a random walk by putting a point on \mathbb{N}^2 , whose coordinates are equal to the degrees of these two vertices.

Consider the following random walk R_k on \mathbb{N}^2 . From the point (A, B) it moves either to the point $(A + 1, B)$ with probability $\frac{\mathcal{W}(A)}{\mathcal{W}(A) + \mathcal{W}(B)}$ or to the point $(A, B + 1)$ with probability $\frac{\mathcal{W}(B)}{\mathcal{W}(A) + \mathcal{W}(B)}$. Note that the sum of the coordinates of R_k increases by 1 at every step.

2.1. The number of paths. We are interested in the probability that R_k moves from some fixed point to the diagonal $\{(m, m)\}_{m \in \mathbb{N}}$. It means that the degrees of the two considered vertices become equal.

The event $\{R_k \text{ crosses the diagonal}\}$ can be partitioned into events $\{R_k \text{ moves to the point } (m, m), \text{ and this is the first time it crosses the diagonal}\}_{m \in \mathbb{N}}$. We evaluate the probabilities of these events. To do it, we first need to count all admissible paths connecting the initial point and the point (m, m) , where by *admissible* we mean that only the endpoints of this path may belong to the diagonal.

Lemma 2. *Let $m \geq A > B$ be some natural numbers. By $\mathcal{G}(A, B, A', B')$ denote the number of admissible paths connecting (A, B) to (A', B') .*

Then

$$\mathcal{G}(A, B, m, m) = \frac{(2m - 1 - A - B)!(A - B)}{(m - A)!(m - B)!}.$$

Proof. By $\mathcal{A}(A, B, A', B')$ denote the number of different up-right paths connecting the point (A, B) with the point (A', B') . Then

$$\mathcal{A}(A, B, A', B') = \binom{A' + B' - A - B}{A' - A},$$

because this is the number of ways to choose at which of $A' + B' - A - B$ steps the path goes up, and at the remaining steps the path goes to the right.

By

$$\mathcal{B}(A, B, A', B') := \mathcal{A}(A, B, A', B') - \mathcal{G}(A, B, A', B')$$

denote the number of non-admissible paths between these two points.

To evaluate $\mathcal{G}(A, B, m, m)$ we use André's reflection principle. Let us show that there is a one-to-one correspondence between all paths from (A, B) to $(m - 1, m)$ and all non-admissible paths from (A, B) to $(m, m - 1)$. Consider an arbitrary path between (A, B) and $(m - 1, m)$. It crosses the diagonal, because $A > B$ but $m - 1 < m$. Now we perform the following operation: all steps before the intersection with the diagonal remain the same while all steps after the intersection are inverted (right \leftrightarrow up). The part of the path after the intersection connected the point (k, k) and the point $(m - 1, m)$ for some k . Therefore, after the inversion it connects the point (k, k) and the point $(m, m - 1)$. Hence, now we have a non-admissible path from (A, B) to $(m, m - 1)$. This process can be reversed, because the first intersection point with the diagonal remains the same, hence the required bijection is constructed.

We get a formula

$$\mathcal{B}(A, B, m, m - 1) = \mathcal{A}(A, B, m - 1, m).$$

Since all admissible paths from (A, B) to (m, m) must have an inner point $(m, m - 1)$, we get the following chain of equalities, which concludes the proof:

$$\begin{aligned}
 \mathcal{G}(A, B, m, m) &= \mathcal{G}(A, B, m, m-1) \\
 &= \mathcal{A}(A, B, m, m-1) - \mathcal{B}(A, B, m, m-1) \\
 &= \mathcal{A}(A, B, m, m-1) - \mathcal{A}(A, B, m-1, m) \\
 &= \binom{2m-1-A-B}{m-A} - \binom{2m-1-A-B}{m-A-1} \\
 &= \frac{(2m-1-A-B)!}{(m-A)!(m-B-1)!} - \frac{(2m-1-A-B)!}{(m-A-1)!(m-B)!} \\
 &= \frac{(2m-1-A-B)!}{(m-A-1)!(m-B-1)!} \left(\frac{1}{m-A} - \frac{1}{m-B} \right) \\
 &= \frac{(2m-1-A-B)!(A-B)}{(m-A)!(m-B)!} .
 \end{aligned}$$

□

2.2. The upper bound for the diagonal intersection probability. By $q(A, m)$ denote the probability that R_k moves from the point $(A, 1)$ to the point (m, m) following an admissible path.

Proposition 3. *There exists a polynomial (with coefficients depending only on the weight function \mathcal{W}) $P(\cdot)$ such that for sufficiently large A and for any $m \geq A$ it is true that*

$$q(A, m) < \frac{P(A)}{2^A m^{3/2}} .$$

Proof. We evaluate upper bounds for the number of paths $\mathcal{G}(A, 1, m, m)$ and for the probability of every fixed path from $(A, 1)$ to (m, m) separately.

Lemma 4. *There exists a polynomial $P_1(\cdot)$ such that*

$$\mathcal{G}(A, 1, m, m) \leq \frac{P_1(A) 2^{2m}}{2^A m^{3/2}} \quad \forall A, m \geq A .$$

Proof. By Lemma 2,

$$\begin{aligned}
 \mathcal{G}(A, 1, m, m) &= \frac{(2m-2-A)!(A-1)}{(m-A)!(m-1)!} \\
 &= \frac{(2m-2)!}{(m-1)!(m-1)!} \cdot \frac{A-1}{2m-1-A} \cdot \frac{(m-A+1) \cdots (m-1)}{(2m-A) \cdots (2m-2)} .
 \end{aligned}$$

In the last expression, the first fraction is a binomial coefficient. Note that the numerator and denominator of the last fraction have the same

number of factors $(A - 1)$, and every factor of the numerator is at most the half of the corresponding factor of the denominator. Therefore

$$\mathcal{G}(A, 1, m, m) \leq \frac{2^{2m}}{\sqrt{m}} \cdot \frac{P_1(A)}{m} \cdot \frac{1}{2^A}$$

(all appearing constants are already included in the polynomial). The lemma is proved. \square

Lemma 5. *There exist a polynomial $P_2(\cdot)$ and a number A_1 such that if $m \geq A > A_1$ then for every path S from $(A, 1)$ to (m, m) it is true that*

$$p(S) \leq \frac{P_2(A)}{2^{2m}}.$$

Proof. Consider a composite path consisting of two simple paths:

$$S^* = S_1, S_2$$

where

$$S_1 = (A, 1), (A, 2), \dots, (A, A),$$

$$S_2 = (A, A), (A+1, A), (A+1, A+1), (A+2, A+1), (A+2, A+2) \dots, (m, m).$$

Proposition 6. *Among all paths with the same endpoints S^* has the largest probability.*

Proof. The probabilities of any two paths with the same endpoints are two fractions with same numerators but with different denominators. Therefore it is sufficient to find the path with a minimal denominator. Every denominator is a product of several expressions of the form $\mathcal{W}(A_k) + \mathcal{W}(B_k)$ where $A_k + B_k$ is fixed. Due to the convexity of \mathcal{W} , the smaller $|A_k - B_k|$ is the smaller $\mathcal{W}(A_k) + \mathcal{W}(B_k)$ is. Clearly, the path S^* minimizes $|A_k - B_k|$ at each step. \square

By $\mathbb{P}(S)$ denote the probability of the path S .

Obviously, we have an upper bound for $\mathbb{P}(S_2)$:

$$\mathbb{P}(S_2) \leq \frac{1}{2^{2(m-A)}} = \frac{2^{2A}}{2^{2m}}.$$

Now to conclude the lemma proof it suffices to show that

$$\mathbb{P}(S_1) \leq \frac{P_2(A)}{2^{2A}}.$$

for some polynomial $P_2(A)$ and sufficiently large A .

The explicit formula for $\mathbb{P}(S_1)$ looks as follows:

$$\mathbb{P}(S_1) = \frac{\mathcal{W}(1)}{\mathcal{W}(1) + \mathcal{W}(A)} \frac{\mathcal{W}(2)}{\mathcal{W}(2) + \mathcal{W}(A)} \cdots \frac{\mathcal{W}(A-1)}{\mathcal{W}(A-1) + \mathcal{W}(A)}.$$

We introduce some notations. By $\widetilde{\mathcal{W}}_A(\cdot)$ denote a function that interpolates \mathcal{W} at the points 1 and A in a linear way.

$\widetilde{\mathcal{W}}_A(n)$ has a form $\widetilde{\mathcal{W}}_A(n) = k_A n + b_A$ for some real numbers k_A and b_A . Let $\beta(A) := b_A/k_A \in [-\infty, +\infty]$.

The function $\mathcal{W}(\cdot)$ is required to be unbounded, so let $A_0 \in \mathbb{N}$ be such that $\mathcal{W}(A_0) > \mathcal{W}(1)$ and put $\beta_0 := \beta(A_0)$. Then for any $A > A_0$ it is obvious that

$$(1) \quad -1 < \beta(A) < \beta_0 .$$

Remark 7. *The function $\beta(A) + 1$ is not necessarily separated from zero, unlike the linear model. On the contrary, if $A = o(\mathcal{W}(A))$ then $\lim_{A \rightarrow \infty} \beta(A) = -1$.*

Every fraction here will increase if we replace $\mathcal{W}(k)$ by $\widetilde{\mathcal{W}}_A(k)$, because all fractions are less than 1, and we add the non-negative number $\widetilde{\mathcal{W}}_A(k) - \mathcal{W}(k)$ to both numerator and denominator. Therefore,

$$\mathbb{P}(S_1) \leq \frac{\widetilde{\mathcal{W}}_A(1)}{\widetilde{\mathcal{W}}_A(1) + \widetilde{\mathcal{W}}_A(A)} \frac{\widetilde{\mathcal{W}}_A(2)}{\widetilde{\mathcal{W}}_A(2) + \widetilde{\mathcal{W}}_A(A)} \cdots \frac{\widetilde{\mathcal{W}}_A(A-1)}{\widetilde{\mathcal{W}}_A(A-1) + \widetilde{\mathcal{W}}_A(A)} .$$

We know that $\widetilde{\mathcal{W}}_A(n) = k_A n + b_A$. After substituting it and reducing all fractions by k_A we get

$$\mathbb{P}(S_1) \leq \frac{1 + \beta(A)}{1 + A + 2\beta(A)} \frac{2 + \beta(A)}{2 + A + 2\beta(A)} \cdots \frac{A - 1 + \beta(A)}{2A - 1 + 2\beta(A)} .$$

Now we want to replace $\beta(A)$ by a larger number β_0 . Note that if $\beta(A) < \beta_0$, then for any $B, C, D, E \in \mathbb{N}$ such that $D + E\beta(A) > 0$ the following holds:

$$(2) \quad \frac{B + C\beta(A)}{D + E\beta(A)} < \frac{B + C\beta_0}{D + E\beta_0} \Leftrightarrow BE - CD < 0 .$$

The condition on the right-hand side is satisfied, thus after replacing $\beta(A)$ by β_0 we get the following inequality:

$$\begin{aligned} \mathbb{P}(S_1) &\leq \frac{(1 + \beta_0)(2 + \beta_0) \cdots (A + \beta_0 - 1)}{(A + 1 + 2\beta_0) \cdots (2A - 1 + 2\beta_0)} \\ &= \frac{\Gamma(A + \beta_0)\Gamma(A + 2\beta_0 + 1)}{\Gamma(\beta_0 + 1)\Gamma(2A + 2\beta_0)} . \end{aligned}$$

By Stirling's formula for any $z \geq 1$ it is true that $\Gamma(z + 1) \asymp \sqrt{z} \left(\frac{z}{e}\right)^z$. After applying this and hiding all the constants into the polynomial we

get

$$\begin{aligned}
\mathbb{P}(S_1) &\leq \frac{P_4(A)e^{2A+2\beta_0}}{e^{A+2\beta_0}e^{A+\beta_0}} \frac{(A+\beta_0-1)^{A+\beta_0-1}(A+2\beta_0)^{A+2\beta_0}}{(2A+2\beta_0-1)^{2A+2\beta_0-1}} \\
&\leq P_3(A) \cdot \left(\frac{A+\beta_0-1}{2A+2\beta_0-1}\right)^{A+\beta_0-1} \cdot \left(\frac{A+2\beta_0}{2A+2\beta_0-1}\right)^{A+2\beta_0} \\
&= P_3(A) \cdot \frac{1}{2^{2A+3\beta_0-1}} \cdot \left(\frac{A+\beta_0-1}{A+\beta_0-1+1/2}\right)^{A+\beta_0-1} \cdot \left(\frac{A+2\beta_0}{A+2\beta_0-(1/2+\beta_0)}\right)^{A+2\beta_0} \\
&\leq P_2(A) \cdot \frac{1}{2^{2A}}.
\end{aligned}$$

The last inequality is not as obvious as the other ones. Note that

$$\left(\frac{x}{x+a}\right)^x = \left(1 - \frac{a}{x+a}\right)^x \leq \exp(-ax/(a+x))$$

and that for large x and bounded a this expression is also bounded by some constant, which has also been already included into the polynomial. \square

The conclusion of Proposition 3 follows from our lemmas by multiplication of the corresponding inequalities. \square

Corollary 8. *By $q(A)$ denote the probability that our random walk moves from the point $(A, 1)$ to the diagonal. Then for sufficiently large values of A and for some polynomial $P(\cdot)$ it is true that*

$$q(A) < \frac{P(A)}{2^A}.$$

Proof. By Proposition 3 we get that

$$q(A) \leq \sum_{m=A}^{\infty} q(A, m) \leq \frac{P(A)}{2^A} \sum_{m=A}^{\infty} \frac{1}{m^{3/2}}.$$

It remains to note that the series $\sum \frac{1}{m^{3/2}}$ is convergent. \square

2.3. Limit distribution of the random walk in the linear case.

Suppose $\mathcal{W}(n) = n + \beta$, $\beta > -1$. In this case, the following proposition provides an explicit asymptotic form of the random walk distribution.

Proposition 9. *If R_k starts at the point $(A, 1)$ then the quantity $A_k/(A_k + B_k)$ tends with probability 1 to some random variable $H(A)$ as k tends to infinity. Moreover, $H(A)$ has a beta probability distribution:*

$$H(A) \sim \text{Beta}(1 + \beta, A + \beta).$$

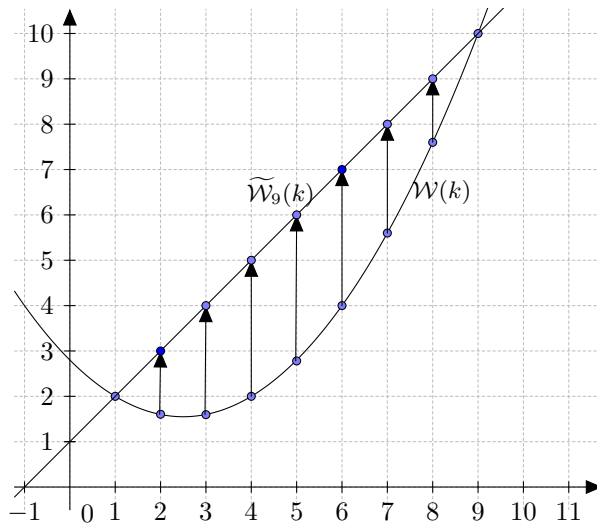


FIGURE 1. Replacing the weight function increases the relative weight of all vertices except for the current hub.

Proof. As noted in [2], this random walk is a special case of Pólya urn model with initial parameters $(1 + \beta, A + \beta)$. Recall that for the urn model the limit distribution of that fraction is well known, see, for example, [10] or [8]. \square

3. THE PROOF OF THE MAIN RESULT

3.1. Comparison between the convex model and the linear model. Motivated by (1), for the convex model with weight function \mathcal{W} we introduce the linear model with $\widetilde{\mathcal{W}}(n) = n + \beta_0$ and call it *the linear comparison model*.

Lemma 10 (The comparison lemma). *Suppose in the convex model before the n -th step the vertex v_t has the maximal degree m , and the degrees of all other vertices are fixed. Let $m > A_0$. By p denote the probability that the next edge is attached to the vertex v_t . Now consider exactly the same situation (all the degrees remain the same), but in the linear comparison model. By \tilde{p} denote the probability that the next edge is attached to the vertex v_t in the linear case.*

Then $p \geq \tilde{p}$.

Proof. Note that for all $1 \leq k \leq m$, $\mathcal{W}(k) \leq \widetilde{\mathcal{W}}_m(k)$ due to the convexity of \mathcal{W} . Then

$$\begin{aligned} p &= \frac{\mathcal{W}(m)}{\sum_v(\mathcal{W}(\deg(v)))} \geq \frac{\widetilde{\mathcal{W}}_m(m)}{\sum_v(\widetilde{\mathcal{W}}_m(\deg(v)))} \\ &= \frac{m + \beta(m)}{2(n-1) + n\beta(m)} \geq \frac{m + \beta_0}{2(n-1) + n\beta_0} = \tilde{p}. \end{aligned}$$

First we increase the denominator (see Figure 1), then we reduce the fraction, and then we use (2). \square

Remark 11. *Unlike the expressions on the left hand side, \tilde{p} depends only on m and on the total degree of vertices.*

3.2. The maximal degree grows fast enough. In the linear and basic models the degree of any fixed vertex grows fast enough to provide the convergence of the series $\sum q(A)$ with probability 1, see Remark 14 below. Unfortunately, this is not always the case in the convex model, for example, if $\mathcal{W}(n) = 2^{2^n}$ then with positive probability the degree of the first vertex will be bounded, because the second one will be connected to almost all vertices. So, any fixed vertex degree can be bounded. However, the maximal degree (the degree of random vertex), as we will see, grows fast enough with probability 1.

By \mathcal{M}_n denote the maximal degree before the n -th step.

Proposition 12. *There exists a sequence C_n of positive real numbers satisfying the following conditions:*

- (1) C_n grows fast enough: the expression $C_n n^{-1/(4+2\beta_0)}$ converges to a positive finite limit,
- (2) C_n/\mathcal{M}_n is a supermartingale with respect to the filtration $\sigma_n = \sigma(\mathcal{M}_1, \dots, \mathcal{M}_n)$.

Corollary 13. *There exists a positive random variable M such that for all $n \geq 2$,*

$$(3) \quad \mathcal{M}_n \geq M n^{1/(4+2\beta_0)}.$$

Proof. C_n/\mathcal{M}_n is a positive supermartingale, hence by Doob's theorem it tends to a finite limit with probability 1, therefore this sequence with probability 1 is bounded by some positive random variable C . But this implies $\mathcal{M}_n \geq C_n/C$ with probability 1, i.e. with probability 1 for all $n \geq 2$ we get (3). \square

Proof of Proposition 12. Let n be sufficiently large. By V_n denote the set of vertices before the n -th step, and by p_n denote the probability

that the maximal degree increases at the n -th step. We can bound it from below:

$$p_n \geq \frac{\mathcal{W}(\mathcal{M}_n)}{\sum_{v \in V_n} \mathcal{W}(\deg(v))} \geq \frac{\widetilde{\mathcal{W}}_{\mathcal{M}_n}(\mathcal{M}_n)}{\sum_{v \in V_n} \widetilde{\mathcal{W}}_{\mathcal{M}_n}(\deg(v))} \geq \frac{\mathcal{M}_n + \beta_0}{\widetilde{w}_n} =: \widetilde{p}_n .$$

Here $\widetilde{w}_n = 2(n-1) + n\beta_0$.

Denote $\alpha := \frac{1}{4+2\beta_0}$.

For the sequence $Y_n := C_n/\mathcal{M}_n$ to be a supermartingale it is necessary to show that

$$\mathbb{E}(Y_{n+1}|\mathcal{F}_n) \leq Y_n .$$

Note that

$$Y_{n+1}/C_{n+1} = \begin{cases} \frac{1}{\mathcal{M}_{n+1}} & \text{with probability } p_n, \\ \frac{1}{\mathcal{M}_n} & \text{with probability } 1 - p_n. \end{cases}$$

It follows that

$$\begin{aligned} \mathbb{E}(Y_{n+1}/C_{n+1}|\mathcal{F}_n) &= \frac{p_n}{\mathcal{M}_{n+1}} + \frac{1-p_n}{\mathcal{M}_n} \\ &= \frac{p_n\mathcal{M}_n + \mathcal{M}_n + 1 - p_n\mathcal{M}_n - p_n}{\mathcal{M}_n(\mathcal{M}_n + 1)} \\ &= \frac{\mathcal{M}_n + 1 - p_n}{\mathcal{M}_n(\mathcal{M}_n + 1)} = \frac{1}{\mathcal{M}_n} - \frac{p_n}{\mathcal{M}_n(\mathcal{M}_n + 1)} \\ &\leq \frac{1}{\mathcal{M}_n} - \frac{\widetilde{p}_n}{\mathcal{M}_n(\mathcal{M}_n + 1)} \leq \frac{1}{\mathcal{M}_n} - \frac{\widetilde{p}_n}{2\mathcal{M}_n^2} \\ &= \frac{1}{\mathcal{M}_n} - \frac{1 + \beta_0/\mathcal{M}_n}{2\mathcal{M}_n\widetilde{w}_n} \leq \frac{1}{\mathcal{M}_n} - \frac{1}{2\mathcal{M}_n\widetilde{w}_n} \\ &= \frac{1}{\mathcal{M}_n} \left(1 - \frac{1}{2(2(n-1) + n\beta_0)} \right) \\ &= \frac{1}{\mathcal{M}_n} \left(1 - \frac{1/(4+2\beta_0)}{n - 4/(4+2\beta_0)} \right) \\ &= \frac{1}{\mathcal{M}_n} \left(1 - \frac{\alpha}{n - 4\alpha} \right) = \frac{1}{\mathcal{M}_n} \left(\frac{n - 5\alpha}{n - 4\alpha} \right) . \end{aligned}$$

Now it is clear that the following inequality is sufficient for Y_n to be a supermartingale:

$$\frac{C_{n+1}}{\mathcal{M}_n} \left(\frac{n - 5\alpha}{n - 4\alpha} \right) \leq \frac{C_n}{\mathcal{M}_n} .$$

To make this inequality true put, for example, $C_{n+1} = C_n(1 + \frac{\alpha}{n-5\alpha})$

A simple computation shows that the sequence C_n/n^α has a positive and finite limit, therefore C_n satisfies both conditions from the statement of the proposition. This completes the proof. \square

3.3. Finite number of hubs. In this subsection we prove that the set of vertices that have been hubs at some step is finite with probability 1.

Consider a decreasing sequence of events

$$B_M = \{\forall n \mathcal{M}_n > Mn^{1/(4+2\beta_0)}\} .$$

for any real $M > 0$.

Let $v_{l(n)}$ be some vertex which has the maximal degree before the n -th step (in general, there can be several such vertices). Consider the following event

$$H_n = \{\text{the vertex } v_{n+1} \text{ has the same degree as } v_{l(n)} \text{ at some future step}\}.$$

We recall that the joint behaviour of vertices $v_{l(n)}$ and v_{n+1} is described by the random walk from Section 2, starting from the point $(\mathcal{M}_n, 1)$. Using Corollary 8, we get that for any $M > 0$ and for sufficiently large n the following is true:

$$\mathbb{P}(H_n \cap B_M) \leq \max_{A \geq Mn^{1/(4+2\beta_0)}} \frac{P(A)}{2^A} \leq \frac{P_1(Mn^{1/(4+2\beta_0)})}{2^{Mn^{1/(4+2\beta_0)}}} .$$

where P, P_1 are some polynomials. The expressions at the right hand side form a convergent series, therefore, using Borel–Cantelli lemma one can show that the event $H_n \cap B_M$ occurs for only finitely many indices n with probability 1. Moreover, because of (3), we see that $P(B_M) \rightarrow 1$ as $M \rightarrow 0$. Therefore, the event H_n also occurs for only finitely many indices n with probability 1.

Hence only finitely many vertices have been hubs.

Remark 14. *In the linear and basic models the proof can be simplified using any fixed vertex for comparison instead of the leader $l(n)$, because in these models even the degree of any fixed vertex grows fast enough, i.e. like some power of n .*

3.4. Finite number of leader changes between any two fixed vertices. It remains to prove the following result:

Theorem 15. *For any two vertices the set of all steps at which their degrees coincide is finite.*

Proof. Consider any two vertices and the corresponding two-dimensional random walk. Suppose the random walk starts from the point (A_k, B_k) , which means that the degrees of these two vertices were at first equal to A_k and B_k respectively. We can always consider k to be sufficiently large so that $A_k + B_k > A_0$. Consider the linear two-dimensional comparison random walk (according to Lemma 10) starting from the same point, but with the other weight function $\widetilde{\mathcal{W}}(A) = A + \beta_0$.

First we introduce some notation. By $\Delta_n := |A_n - B_n|$ denote the difference between A_n and B_n , and by $\widetilde{\Delta}_n := |\widetilde{A}_n - \widetilde{B}_n|$ denote the corresponding difference in the linear comparison model.

Proposition 16. *There exists a coupling between Δ_n and $\widetilde{\Delta}_n$, such that Δ_n stochastically dominates $\widetilde{\Delta}_n$ for any $n \geq k$.*

Proof. We construct both functions Δ_n and $\widetilde{\Delta}_n$ on the same probability space preserving every independency relation each of them must satisfy.

Using induction on n , we now show that for every n , $\Delta_n \geq \widetilde{\Delta}_n$ with probability 1. For $n = k$ it is true.

Now consider a (maximal by inclusion) set $L \subset \Omega$ of positive measure p such that the functions Δ_n and $\widetilde{\Delta}_n$ are constants on L , and, by induction, $\Delta_n \geq \widetilde{\Delta}_n$ on L .

By q denote the probability that Δ_n increases by 1 (therefore, it decreases by 1 with probability $1 - q$), and by \widetilde{q} denote the probability that $\widetilde{\Delta}_n$ increases by 1. Let $\widetilde{\Delta}_n$ be positive on L . Then, by Lemma 10, $q > \widetilde{q}$. Let L' be a subset of L on which $\Delta_{n+1} = \Delta_n + 1$, and \widetilde{L}' be a subset of L on which $\widetilde{\Delta}_{n+1} = \widetilde{\Delta}_n + 1$. Clearly, the probability of the set L' is greater than the probability of the set \widetilde{L}' , therefore we can choose them in such a way that $\widetilde{L}' \subset L'$. So on L the induction inequality $\Delta_{n+1} \geq \widetilde{\Delta}_{n+1}$ now holds.

The only remaining set is the set where $\widetilde{\Delta}_n = 0$. On its subset where $\Delta_n \neq 1$ the required inequality $\Delta_{n+1} \geq \widetilde{\Delta}_{n+1}$ holds automatically, and now all we need is to note that Δ_n and $\widetilde{\Delta}_n$ are of the same parity (because their parities both change at every step), so the set where $\widetilde{\Delta}_n = 0$ and $\Delta_n = 1$ is empty. This concludes the construction of the functions Δ_n and $\widetilde{\Delta}_n$. □

Now we show that with probability 1 the sequence $\widetilde{\Delta}_n$ is equal to zero only finitely many times. Then it is also true for Δ_n , because $\Delta_n \geq \widetilde{\Delta}_n$.

It follows from Proposition 9 and from the absolute continuity of beta-distribution that the probability of every particular value equals

to zero. Therefore with probability 1 $A_n/(A_n + B_n)$ converges to some $y \neq \frac{1}{2}$. Hence this fraction can be equal to $\frac{1}{2}$ only finitely many times, and it means that $\tilde{\Delta}_n$ equals to zero only finitely many times with probability 1, q.e.d. □

Now the result of Theorem 1 for $m_0 = 1$ obviously follows from those of Subsections 3.3 and 3.4.

From Theorem 1 we can easily deduce an important known result about the behaviour of maximal degrees in the linear model from [12]:

Corollary 17. *In the linear model the variable \mathcal{M}_n satisfies the following:*

$$\mathcal{M}_n n^{-1/(2+\beta)} \rightarrow \mu ,$$

where μ is an almost surely positive and finite random variable.

Proof. We know that \mathcal{M}_n behaves like the degree of some fixed vertex. Moreover, it is known that in the linear model the degree of every vertex is asymptotically equivalent to $n^{-1/(2+\beta)}$ multiplied by some random constant. □

3.5. Generalization to the case of an arbitrary m_0 . The case of $m_0 > 1$ is often considered to be much more complicated than the case of $m_0 = 1$, because the graph is not a tree. But it turns out that all steps of the presented proof (pairwise vertex degree analysis, the sufficiently fast growth of the maximal degree, finite number of hubs and leader changes) remain literally the same for $m_0 > 1$, except for just one change: in the pairwise analysis part, the random walk related to the degrees of the vertices starts not from the point $(A, 1)$, but from the point (A, m_0) . This obstacle can be easily avoided by introducing a new (convex and unbounded) weight function $\mathcal{W}'(n) := \mathcal{W}(n + m_0 - 1)$. Then the random walk with the weight function \mathcal{W} starting from the point (A, m_0) is isomorphic to the random walk with the weight function \mathcal{W}' starting from the point $(A - m_0 + 1, 1)$, and for this case we have already provided all necessary bounds.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Professor M. A. Lifshits for his guidance, patience, time and valuable remarks. I am also grateful to Andrey Alpeev, who suggested to consider the convex model, and to Vladimir Zolotov for many fruitful conversations. This research is supported by JSC "Gazprom Neft" and by Chebyshev Laboratory

(Department of Mathematics and Mechanics, St. Petersburg State University) under RF Government grant 11.G34.31.0026.

REFERENCES

- [1] William Aiello, Fan Chung, and Linyuan Lu. Random evolution in massive graphs. In *Handbook of massive data sets*, volume 4 of *Massive Comput.*, pages 97–122. Kluwer Acad. Publ., Dordrecht, 2002.
- [2] Ágnes Backhausz. Limit distribution of degrees in random family trees. *Electron. Commun. Probab.*, 16:29–37, 2011.
- [3] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.
- [4] Béla Bollobás, Oliver Riordan, Joel Spencer, and Gábor Tusnády. The degree sequence of a scale-free random graph process. *Random Structures Algorithms*, 18(3):279–290, 2001.
- [5] Steffen Dereich and Peter Mörters. Random networks with sublinear preferential attachment: degree evolutions. *Electron. J. Probab.*, 14:no. 43, 1222–1267, 2009.
- [6] Steffen Dereich and Peter Mörters. Random networks with concave preferential attachment rule. *Jahresber. Dtsch. Math.-Ver.*, 113(1):21–40, 2011.
- [7] C. Godrèche, H. Grandclaude, and J. M. Luck. Statistics of leaders and lead changes in growing networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2010(02):P02001, 2010.
- [8] Norman L. Johnson and Samuel Kotz. *Urn models and their application*. John Wiley & Sons, New York-London-Sydney, 1977. An approach to modern discrete probability theory, Wiley Series in Probability and Mathematical Statistics.
- [9] Paul L Krapivsky and Sidney Redner. Organization of growing random networks. *Physical Review E*, 63(6):066123, 2001.
- [10] Hosam M. Mahmoud. *Pólya urn models*. Texts in Statistical Science Series. CRC Press, Boca Raton, FL, 2009.
- [11] T. F. Móri. On random trees. *Studia Sci. Math. Hungar.*, 39(1-2):143–155, 2002.
- [12] Tamás F. Móri. The maximum degree of the Barabási-Albert random tree. *Combin. Probab. Comput.*, 14(3):339–348, 2005.
- [13] M. E. J. Newman. The structure and function of complex networks. *SIAM Rev.*, 45(2):167–256 (electronic), 2003.
- [14] Roberto Oliveira and Joel Spencer. Connectivity transitions in networks with super-linear preferential attachment. *Internet Math.*, 2(2):121–163, 2005.
- [15] Anna Rudas, Bálint Tóth, and Benedek Valkó. Random trees and general branching processes. *Random Structures Algorithms*, 31(2):186–202, 2007.
- [16] Anna Rudas and Imre Péter Tóth. Entropy and Hausdorff dimension in random growing trees. *Stoch. Dyn.*, 13(1):1250010, 30, 2013.