ALEA, Lat. Am. J. Probab. Math. Stat. 12 (2), 903–915 (2015)

The power of choice over preferential attachment

Yury Malyshkin and Elliot Paquette

Tver State University
Sadovij per. 35,
Tver, 170002, Russia
E-mail address: yury.malyshkin@mail.ru

Department of Mathematics, Weizmann Institute of Science
Herzl St 234,
Rehovot, 7610001, Israel
E-mail address: elliot.paquette@gmail.com

Abstract. We introduce a new type of preferential attachment tree that includes choices in its evolution, like with Achlioptas processes. At each step in the growth of the graph, a new vertex is introduced. A fixed number d of possible neighbor vertices are selected independently and with probability proportional to degree. Between them, the vertex with smallest degree is chosen, and a new edge is created. We determine with high probability the largest degree of this graph up to some additive error term.

1. Introduction

In the present work we consider an alteration of the preferential attachment model, in the spirit of the Achlioptas processes (see Achlioptas et al. (2009); Riordan and Warnke (2012)). The preferential attachment graph is a time-indexed sequence of graphs constructed the following way. We start with a single edge, and at each time step we add a new vertex. We then select an old vertex with probability proportional to the degree of the vertex, and we add a new edge between the new vertex and the selected vertex. This model is widely studied and many of its properties are known, such as the maximum degree, the limiting degree distribution, and the diameter of the graph (for instance see Barabási and Albert (1999); Flaxman et al. (2005); Dommers et al. (2010); Móri (2005)). In particular, in Flaxman et al. (2005) it was shown that at time t, for any function f with $f(t) \to \infty$ as $t \to \infty$, $\frac{t^{1/2}}{f(t)} \leq \Delta(t) \leq t^{1/2} f(t)$ with high probability, where $\Delta(t)$ is the highest

Received by the editors March 7, 2015; accepted September 7, 2015.
2010 *Mathematics Subject Classification.* 05C80.
Key words and phrases. Preferential Attachment, Random Graphs, Maximum Degree.
YM gratefully acknowledges the support of the Weizmann Institute of Science, where this work was performed. EP gratefully acknowledges the support of NSF Postdoctoral Fellowship DMS-1304057.

degree of the preferential attachment graph at time t. In Móri (2005), this was strengthened to say that over the course of all time, $\Delta(t)t^{-1/2}$ converges almost surely to a non-degenerate positive random variable. We say that some event $\mathcal{E}_n$ occurs with high probability as $n \to \infty$ if $\mathbb{P}(\mathcal{E}_n) \to 1$ as $n \to \infty$. When it is clear which parameter is turning to infinity we omit it.

We will consider an alteration of this model that allows limited choice into its evolution. Fix an integer $d > 0$ and define a sequence of trees $\{P_m\}$ given by the following rule. Let P_1 be the one-edge tree. Given P_m, define P_{m+1} by first adding one new vertex v_{m+1}. Let $X_m^1, X_m^2, \ldots X_m^d$ be i.i.d. vertices from $V(P_m)$ (here $V(P)$ is the set of vertices of P) chosen with probability

$$\mathbb{P}\left[X^1 = w\right] = \frac{\deg w}{2m}.$$

Note that as the graph has m edges, $\sum_w \deg w = 2m$. Finally, create a new edge between v_{m+1} and Y_m, where Y_m is whichever of $\left\{X_m^1, X_m^2, \ldots X_m^d\right\}$ has smallest degree. In the case of a tie, choose amongst these smallest degree vertices with equal probability. We call this the *min-choice preferential attachment tree*.

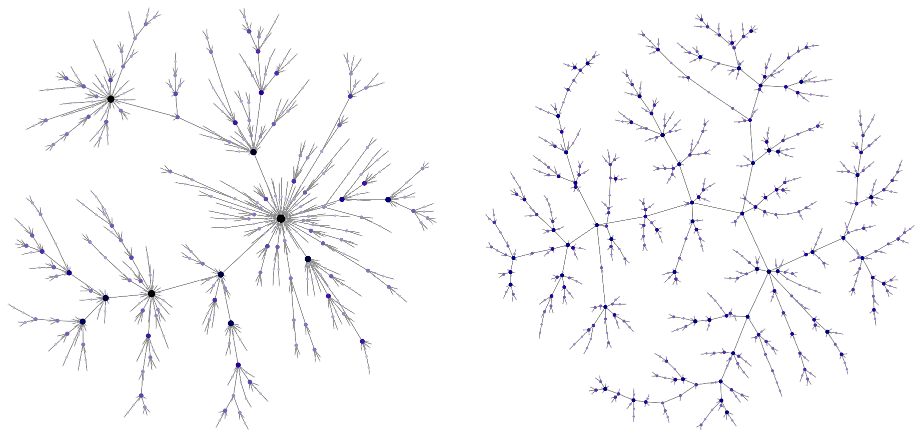
In D'Souza et al. (2007), similar models of randomly evolving networks were introduced. Among others, they study a model in which they sample $\left\{X_m^1, X_m^2, \ldots, X_m^d\right\}$ uniformly at random, and then they draw an edge to the minimal degree vertex.

This is in turn strongly related to the original model of Azar et al. (1999), in which this type of choice was introduced to study load balancing. In its simplest form, this amounts to studying balls thrown randomly into bins. Suppose we have n bins and n balls, and at each step we put a new ball into one of the bins, choosing the bin randomly and uniformly. In this model the number of balls in the most loaded bin is about $\log n/\log\log n$, as $n \to \infty$. Adding d choices to this model significantly reduces this number. More precisely, we alter the model so that at each step we independently select d bins and put the ball in the bin that contains fewer balls. In the case that multiple bins hold a minimal number of balls, break the tie uniformly at random. As a result the number of balls in the most loaded bin is $\log\log n/\log d + \Theta(1)$.

There are a few differences between our model and the bin and ball model with d choices. First, the d-choice preferential attachment model tends to select higher degree vertices because of the size biasing. Second, the ball and bin model tends to select empty bins frequently at the beginning of the process, while adding a new vertex to the d-choice preferential attachment model always increases the degree of an existing vertex (this is also true in the model of D'Souza et al. (2007), but it alone does not greatly increase the maximum degree). Both influences tend to create higher degree vertices and more loaded bins. Note that the combined influences of these effects have a large impact in the models without two choices. The degree distribution in the preferential attachment model follows a power law Barabási and Albert (1999), while the load distribution in the bin and ball model is asymptotically Poisson.

Our main theorem shows that these differences are in some sense less powerful than the power of d choices.

Theorem 1.1. *With high probability, the maximum degree of P_m is $\frac{\log\log m}{\log d} + \Theta(1)$.*



(A) The preferential attachment tree after 1000 vertices have been added.

(B) The min-choice preferential attachment tree with $d = 2$ after 1000 vertices have been added.

Before going deep into the proof, we will outline the approach. Define $F_m(k)$ to be the weight under the size bias distribution given to vertices of the graph P_m of degree greater than k, i.e.

$$F_m(k) = \sum_{i=1}^{m} (\deg v_i) \mathbf{1} \{\deg v_i \geq k\}.$$

Note that $F_m(1) = \sum_{i=1}^m \deg v_i = 2m$, as there are always m edges in the graph. If it holds that $F_m(k) > 0$ for some $k > 0$, there is a vertex of P_m with degree at least k, while if $F_m(k) < k$ then all vertices of P_m have degrees less than k.

Now $F_m(k)$ as a function of k is a Markov chain in m which evolves according to the following rule, valid for $k > 1$,

$$F_{m+1}(k) - F_m(k) = \begin{cases} 1, & \mathbb{P} = \left(\frac{F_m(k)}{2m}\right)^d \\ k, & \mathbb{P} = \left(\frac{F_m(k-1)}{2m}\right)^d - \left(\frac{F_m(k)}{2m}\right)^d \\ 0, & \text{otherwise,} \end{cases} \tag{1.1}$$

where $\mathbb{P}$ is the probability to connect the new vertex to a vertex whose degree exceeds or equals k, conditional on the process up to time m. We begin by showing that for fixed k, as $m \to \infty$, we have that $F_m(k)/(2m)$ converges to some deterministic constant $\alpha_{k,d}$ (Lemma 3.1). This sequence $\{\alpha_{k,d}\}_{k \in \mathbb{N}}$ is defined inductively by the rules that $\alpha_{1,d} = 1$, $0 < \alpha_{i+1,d} \leq \alpha_{i,d}$ and

$$2\alpha_{k,d} + (k-1)\alpha_{k,d}^d = k\alpha_{k-1,d}^d. \tag{1.2}$$

This equation has unique solution on $(0, 1)$ since the left hand side is monotone on it.

These $\alpha_{k,d}$ decay doubly exponentially, but only after a long enough, d-dependent burn-in time. For these initial steps, very careful analysis is required to ensure that they even decrease. For $d = 2$, we solve equation (1.2) directly. Applying the quadratic formula, the following rational upper bounds can be verified inductively using (1.2) and monotonicity. After these first steps, we may use an approximate

Exact Value	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}
Bound	1	$\frac{3}{4}$	$\frac{9}{16}$	$\frac{2}{5}$	$\frac{3}{10}$	$\frac{1}{5}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{60}$	$\frac{1}{600}$

FIGURE 1.1. Rational upper bounds for $\alpha_{k,2}$, $1 \leq k \leq 10$.

recurrence relation to estimate the decay for larger k (see Lemma 3.2). To obtain the initial estimates for $\alpha_{k,d}$, $d > 2$ we will use a comparison with $\alpha_{k,2}$. By multiplying equation (1.2) by $\alpha_{k,d}^{d-2}$ we get

$$2\alpha_{k,d}^{d-1} + (k-1)\alpha_{k,d}^{2d-2} = k\alpha_{k-1,d}^{d}\alpha_{k,d}^{d-2} \leq k\alpha_{k-1,d}^{2d-2}.$$

Therefore if we put $u_{k,d} = \alpha_{k,d}^{d-1}$ we have that

$$2u_{k,d} + (k-1)u_{k,d}^2 < ku_{k-1,d}^2$$

and hence $\alpha_{k,d} \leq \alpha_{k,2}^{\frac{1}{d-1}}$. Note that as a by-product of showing the convergence of $F_j(k)/(2j)$ and of the decay of $\alpha_{k,d}$, we will have shown the in-probability convergence of the empirical degree distribution.

We then proceed by a bootstrapping method known as *layered induction*, whereby good control on $(F_j(k))_{j=1}^m$ can be used to get better control on $(F_j(k+1))_{j=1}^m$. However, there are some complications in this procedure that make this more technical than textbook layered induction.

First, due to having some initial control on $F_j(k)/(2j)$, which is afforded by the convergence to $\alpha_{k,d}$, we can find a decaying solution $f(k)$ to the recurrence $f(k) = kf(k-1)^d$ where $f(k_0) > \alpha_{k_0,d}$ for some fixed k_0. For slowly growing $j_0(m)$ we have with high probability that $F_j(k_0)/(2j) < f(k_0)$ for all $j \geq j_0(m)$.

Then for $k > k_0$, we can nearly dominate $(F_j(k))_{j=j_0}^{2m}$ by a random walk with increments distributed as $k\,\mathrm{Bernoulli}(f(k-1))$. Hence, as k grows, the increments of the random walk become progressively more heavy tailed. For this reason, to get control with high probability, we require the random walk to evolve for some amount of time, depending on k, before we can expect to get control on the ratio $F_j(k)/(2j)$. Hence, in Lemma 3.3, we define

$$\phi(m,k) = \lceil (\log\log m)^{1/3} \rceil C^{d^{k+1}}$$

for some sufficiently large C and show that with high probability $F_j(k) \leq 2jf(k)$ for all (j,k) with $\phi(m,k) \leq j \leq m$ and $k_0 \leq k \leq k_*(m)$, where $k_*(m) = \log\log m / \log d + \Theta(1)$.

As a result of this procedure, for $k > k_*(m)$ have strong control on $F_j(k)/(2j)$ for $j \geq \phi(m,k)$ but relatively weak control (which just follows from monotonicity in k) for small $j < \phi(m,k)$. Hence, we have effectively shown a bound of the form $F_j(k)/(2j) < j^{-\beta_0}$ that holds for all $j \geq j_0(m)$. We then show that this bound can be bootstrapped finitely many times to produce polynomial decay of order $F_j(k)/(2j) < j^{-\beta}$ where $\beta > \frac{1}{d}$ (Lemma 3.5). At this point, the probability of $F_j(k)$ increasing for any $j > j_1(m)$, where $j_1(m)$ is some slowly growing function, tends to 0, and we conclude the proof.

To summarize, we go through 4 steps.

(1) We get starting estimates for k less than some fixed $k_0 > 0$ and $\omega(m) \leq j \leq m$, $\omega(m) \to \infty$ as $m \to \infty$ (see Lemma 3.1).

(2) We get improved estimates for $k_0 \leq k \leq k_*(m)$ that decrease doubly exponentially in k but are only valid for $\phi(m,k) \leq j \leq m$ (see Lemma 3.3).
(3) We then get estimates of the form $F_j(k) \leq 2j^{1-\beta}$ for some $0 < \beta < 1$ that hold for $k > k_*(m)$ and $(\log\log m)^M \leq j \leq m$. By increasing k finitely many times, we can make β very close to 1 (see Lemma 3.5).
(4) Once β is sufficiently large for some $k = k_*(m) + r$, we show that in fact $F_j(k+1)$ must be 0 (see Lemma 3.6).

2. Discussion

Theorem 1.1 answers a question about the degree sequence of the tree, which uses no topological features of the graph. In the case of the standard preferential attachment model, the diameter is known to be logarithmic Pittel (1994); Dommers et al. (2010). It would be interesting to know if this remains the case in the min-choice preferential attachment tree or if the diameter is larger. In Rudas et al. (2007), the authors derive the limiting law of the preferential attachment tree viewed from a random vertex; a deeper, narrower tree should be expected in the case of the min-choice tree.

The *max-choice* preferential attachment model also presents an interesting model. This corresponds to choosing the vertex of larger degree instead of smaller degree. There one sees largest degree which is linear or nearly linear in m (see Malyshkin and Paquette (2014)).

The preferential attachment model fits naturally inside a larger class of processes where the new vertex chooses a neighbor in the old graph with probability proportional to some power α of the degree, which was first studied in Krapivsky et al. (2000). In the case that $\alpha > 1$, the tree has a single dominant vertex Oliveira and Spencer (2005). This "persistent hub" (using terminology of Dereich and Mörters (2009)) has degree of order m, while all other vertices have bounded degree. The min-choice adaptation can be made to these models as well, first sampling d vertices with probability proportional to the power α of the degree and then choosing the vertex with minimal degree. Simulations suggest that for α large enough (around 1.8) a single vertex dominates the others, while for α up to 1.5 the tree remains more diffuse. This leaves open the possibility of a sharp transition in behavior for some critical value of α.

3. Proofs

For the first step we prove the following.

Lemma 3.1. *For any $\epsilon > 0$, any $\omega(m) \to \infty$, $\omega < m$ and any $k \geq 1$ fixed, we have that*

$$\mathbb{P}\left[\exists j,\ m \geq j \geq \omega(m)\ :\ |F_j(k) - 2\alpha_{k,d} j| > \epsilon j\right] \to 0.$$

Proof: We prove this lemma using induction over k. The base case, $k = 1$, is immediate as $F_j(1) = 2j$ for all j. Define the event $\mathcal{A}$ by

$$\mathcal{A}(\omega_0(m), k-1, \delta) = \{\ \forall\ j, \omega_0(m) \leq j < m, |F_j(k-1) - 2j\alpha_{k-1,d}| < \delta j\}.$$

From the induction hypothesis, we have that $\mathcal{A}$ holds with high probability for any $\delta > 0$ fixed and any $\omega_0(m) \to \infty$. Therefore, it suffices to show that there is a $\delta > 0$

and a $\omega_0(m) \to \infty$ so that

$$\mathbb{P}\left[\mathcal{A}(\omega(m), k, \epsilon)^c \cap \mathcal{A}(\omega_0(m), k-1, \delta)\right] \to 0.$$

We will only prove the upper bound, i.e. that $F_j(k) - 2\alpha_{k,d} j \leq \epsilon j$ with high probability. The lower bound follows from an identical argument.

Let $\omega_0(m) \to \infty$ and $\delta > 0$ be considered fixed, with appropriate values to be determined later. For j such that $\omega_0(m) \leq j \leq m$,

$$F_j(k) = F_{\omega_0(m)}(k) + \sum_{i=\omega_0}^{j} \chi_i \leq 2\omega_0(m) + \sum_{i=\omega_0}^{j} \chi_i,$$

almost surely, where

$$\chi_i = F_i(k) - F_{i-1}(k) = \begin{cases} 1, & \mathbb{P} = \left(\frac{F_{i-1}(k)}{2(i-1)}\right)^d \\ k, & \mathbb{P} = \left(\frac{F_{i-1}(k-1)}{2(i-1)}\right)^d - \left(\frac{F_{i-1}(k)}{2(i-1)}\right)^d \\ 0, & \text{otherwise.} \end{cases}$$

Note that $\frac{F_{i-1}(k-1)}{2(i-1)} \leq \alpha_{k-1,d} + \delta/2$. We may construct variables η_i whose law given $\sigma(F_{i-1}(k))$ is

$$\eta_i = \begin{cases} 1, & \mathbb{P} = \left(\frac{F_{i-1}(k)}{2(i-1)}\right)^d \\ k, & \mathbb{P} = (\alpha_{k-1,d} + \delta/2)^d - \left(\frac{F_{i-1}(k)}{2(i-1)}\right)^d \\ 0, & \text{otherwise} \end{cases}$$

so that on the event $\mathcal{A}(\omega_0(m), k-1, \delta)$, we have $\chi_i \leq \eta_i$.

Then it follows that

$$F_j(k) \leq 2\omega_0(m) + \sum_{i=\omega_0}^{j} \eta_i.$$

Let π be the first $j \geq \omega_0(m)$ so that $F_j(k) \leq 2(\alpha_{k,d} + \epsilon/2)j$. We will estimate the probability that $\pi \leq \omega(m)$. Set $g_i = F_i(k)/(2i) - \alpha_{k,d}$. If $\omega_0(m) \leq i < \pi$, then $g_i > \epsilon/2$.

We can expand the law of η_i as

$$\eta_i = \begin{cases} 1, & \mathbb{P} = (\alpha_{k,d} + g_{i-1})^d \\ k, & \mathbb{P} = \alpha_{k-1,d}^d - \delta/2(\sum_{l=0}^{d-1} \alpha_{k-1,d}^l (\delta/2)^{d-1-l}) + (\alpha_{k,d} + g_{i-1})^d \\ 0, & \text{otherwise.} \end{cases}$$

Choose δ such that $\delta/2(\sum_{l=0}^{d-1} \alpha_{k-1,d}^l (\delta/2)^{d-1-l}) = \epsilon/(4k)$. We may construct i.i.d. variables

$$\eta_i^{'} = \begin{cases} 1, & \mathbb{P} = \alpha_{k,d}^d \\ k, & \mathbb{P} = \alpha_{k-1,d}^d - \alpha_{k,d}^d + \epsilon/(4k) \\ 0, & \text{otherwise} \end{cases}$$

so that $\eta_i \leq \eta_i'$ on $\mathcal{A}(\omega_0(m), k-1, \delta)$ for $i > \omega_0(m)$ such that $F_i(k)/(2i) \geq \alpha_{k,d}$. Set ρ to be the first time after π that $F_i(k)/(2i) \leq \alpha_{k,d}$.

Note that from the definition of $\alpha_{k,d}$, it follows that $\mathbb{E}\eta_i^{'} = 2\alpha_{k,d} + \epsilon/4$. Now we obtain the estimate

$$\begin{aligned}\mathbb{P}(\pi > \omega(m)) &\leq \mathbb{P}(\sum_{i=\omega_0}^{\omega(m)} \eta_i + 2\omega_0(m) > 2\alpha_{k,d}\omega(m) + \epsilon\omega(m)/2) \\ &\leq \mathbb{P}(\sum_{i=\omega_0}^{\omega(m)} \eta_i^{'} > 2\alpha_{k,d}(\omega(m) - \omega_0(m)) + \epsilon\omega(m)/2 - \omega_0(m)(2 - 2\alpha_{k,d})).\end{aligned}$$

Choose $\omega_0(m) = \frac{\epsilon}{6(2-2\alpha_{k,d}-\epsilon/3)}\omega(m)$, so that $\epsilon\omega(m)/2 - \omega_0(m)(2-2\alpha_{k,d}) = (\omega(m) - \omega_0(m))\epsilon/3$. Then we have

$$\mathbb{P}(\pi > \omega(m)) \leq \mathbb{P}(\sum_{i=\omega_0}^{\omega(m)} \eta_i^{'} > (2\alpha_{k,d} + \epsilon/3)(\omega(m) - \omega_0(m))) \leq C_1 e^{-C_2\omega(m)},$$

where C_1, C_2 are some positive constants (which still depend on k and ϵ).

Now we estimate the probability that $F_j(k)$ reaches the line $2\alpha_{k,d}j + \epsilon j$ when started from time $\pi > \omega_0$. From monotonicity, we may assume that $F_\pi(k) = \lfloor 2\alpha_{k,d}\pi + \epsilon\pi/2 \rfloor$. Let $M_a(j)$ denote the random walk with increments distributed as η_1', started from level a and stopped when the process crosses the line $2\alpha_{k,d}j$.

Define the following function

$$p(m, r_1, r_2) = \sup_{t \geq \omega_0(m)} \mathbb{P}\left[\exists\ j \geq t\ :\ M_{\lfloor 2\alpha_{k,d}t + r_1 t \rfloor}(j - t) \geq \lfloor (2\alpha_{k,d} + r_2)j \rfloor\right].$$

We claim that for all fixed $\epsilon/4 < r_1 < r_2$, we have $p(m, r_1, r_2) \to 0$. This follows from a simple tail bound estimate, and we will delay the proof until the end.

Let ρ_1 be the first time after π that the process drops below the line $2\alpha_{k,d}j$ and returns to level greater than $\lfloor 2\alpha_{k,d}j + \epsilon j/2 \rfloor - k$ without crossing $2\alpha_{k,d}j + \epsilon j$. Likewise, let $\rho_i \geq \rho_{i-1}$ be the i^{th} time that this happens. Given that $\rho_i < \infty$, for ρ_i to occur, it must be that the process crosses from level $\lfloor 2\alpha_{k,d}j + \epsilon j/3 \rfloor$ to level $\lfloor 2\alpha_{k,d}j + 3\epsilon j/8 \rfloor$, provided m is sufficiently large, and hence

$$\mathbb{P}\left[\rho_i < \infty \mid \rho_{i-1} < \infty\right] \leq p(m, \epsilon/3, 3\epsilon/8).$$

We now decompose the probability of $F_j(k)$ exceeding $2\alpha_{k,d}j + \epsilon j$ according to the renewal times ρ_j.

$$\begin{aligned}&\mathbb{P}\left[\exists\ j \geq \pi\ :\ F_j(k) > 2\alpha_{k,d}j + \epsilon j\right] \\ &\quad\leq \sum_{i=0}^{\infty} \mathbb{P}\left[\exists\ j, \rho_{i+1} \geq j \geq \rho_i\ :\ F_j(k) > 2\alpha_{k,d}j + \epsilon j \mid \rho_i < \infty\right] \mathbb{P}\left[\rho_i < \infty\right] \\ &\quad\leq \sum_{i=0}^{\infty} p(m, \epsilon/2, \epsilon) p(m, \epsilon/3, 3\epsilon/8)^i = o(1).\end{aligned}$$

It remains to show that for all fixed $\epsilon/4 < r_1 < r_2$, we have $p(m, r_1, r_2) \to 0$. Set $S_j = M_a(j) - \mathbb{E}M_a(j)$. The event that

$$\mathcal{E} = \{\exists\ j \geq t\ :\ M_{\lfloor 2\alpha_{k,d}t + r_1 t \rfloor}(j - t) \geq \lfloor (2\alpha_{k,d} + r_2)j \rfloor\}$$

has

$$\mathcal{E} \subseteq \{\exists\ n \geq 0\ :\ S_n \geq (r_2 - \epsilon/4)n + (r_2 - r_1)t - 1\}.$$

From Hoeffding's inequality, we have that for fixed n, there is a constant $c = c(k,\epsilon) > 0$ so that

$$\mathbb{P}\left[S_n \geq t\right] \leq \exp(-ct^2/n).$$

Summing over all n, we get that

$$\begin{aligned}
&\mathbb{P}\left[\exists\ n \geq 0\ :\ S_n \geq (r_2-\epsilon/4)n + (r_2-r_1)t - 1\right] \\
&\qquad \leq \sum_{n=0}^{\infty} \exp(-c((r_2-\epsilon/4)n + (r_2-r_1)t)^2/n) \\
&\qquad \leq \exp(-2c(r_2-\epsilon/4)(r_2-r_1)t) \sum_{n=1}^{\infty} \exp(-c(r_2-\epsilon/4)n) \\
&\qquad \leq \frac{\exp(-2c(r_2-\epsilon/4)(r_2-r_1)t)}{1-\exp(-c(r_2-\epsilon/4))}.
\end{aligned}$$

This goes to 0 uniformly in $t \geq w_0(m)$, and hence the proof is complete.

□

Lemma 3.2. *Let* $0 < f(k_0) < \frac{1}{e^2 k_0}$ *and inductively define* $f(k+1) = f(k)^d(k+1)$ *for* $k \geq k_0$. *Then there are constants* $c_1 > 0$ *and* $c_2 > 0$ *so that for all* $j \geq 0$,

$$\exp(-c_1 d^j) \leq f(k_0+j) \leq \exp(-c_2 d^j).$$

Proof: It is easily verified by induction that $f(k)$ can be expressed using the following rule for $k > k_0$,

$$\log f(k) = d^{k-k_0} \sum_{i=1}^{k-k_0} d^{-i} \log(k_0+i) + d^{k-k_0} \log f(k_0). \tag{3.1}$$

Thus, from the positivity of the $\log(k_0+i)$ term, it follows immediately that

$$\log f(k) > d^{k-k_0} \log f(k_0),$$

so that the lower bound holds with $c_1 = -\log f(k_0)$. For the upper bound, we note that $\log(k_0+i) \leq \log(k_0) + i$ and hence

$$\sum_{i=1}^{k-k_0} d^{-i} \log(k_0+i) \leq \sum_{i=1}^{\infty} d^{-i}(\log(k_0)+i) = \log(k_0) + d/(d-1)^2 \leq \log(k_0) + 2.$$

Thus from (3.1), we have that

$$\log f(k) \leq d^{k-k_0}\left(\log(k_0) + 2 + \log(f(k_0))\right),$$

As we have $e^2 k_0 f(k_0) < 1$, we may take $c_2 = -\log(e^2 k_0 f(k_0))$ to complete the proof.

□

Now set $\rho(m) = \lceil (\log\log m)^{1/3} \rceil$ and define $\phi(m,k)$ to be $\rho(m)C^{d^{k+1}}$ where C is an integer sufficiently large that

$$\log C > c_1 \vee (\log 4 + c_1 d^{-k_0}). \tag{3.2}$$

Let $k_* = k_*(m)$ be the smallest integer so that

$$C^{d^{k_*+1}} \geq m^{1/2}.$$

Note that this makes $k_* = \frac{\log\log m}{\log d} + \Theta(1)$.

Lemma 3.3. *If $\alpha_{k_0,d} < \frac{1}{e^2 k_0}$, then with high probability, for all $k_0 \le k \le k_*$ and for all j with $m \ge j \ge \phi(m,k)$,*

$$\frac{F_j(k)}{2j} \le f(k),$$

where $f(k)$ defined as in Lemma 3.2 with $f(k_0) = (\alpha_{k_0,d} + \frac{1}{e^2 k_0})/2$.

Remark 3.4. By using this and Lemma 3.1 for $d = 2$ along with the estimate $\alpha_{k,d} \le \alpha_{k,2}^{\frac{1}{d-1}}$ we obtain that there is a $k_0 = k_0(d)$ such that $\alpha_{k_0,d} < \frac{1}{e^2 k_0(d)}$.

Proof: The case $k = k_0$ follows from Lemma 3.1 with $\omega(m) = \phi(m, k_0)$. We now show how the proof follows by layered induction. Let $\mathcal{G}_k$ be the event

$$\mathcal{G}_k = \{F_j(k) \le 2jf(k)\ ,\ \forall\ j\ :\ m \ge j \ge \phi(m,k)\}.$$

For any $j \ge \phi(m, k+1)$,

$$\frac{F_j(k+1)}{2j} = \frac{F_{\phi(m,k)}(k+1)}{2j} + \frac{1}{2j} \sum_{i=\phi(m,k)}^{j} \xi_i(k+1),$$

where $\xi_i(k+1) = F_{i+1}(k+1) - F_i(k+1)$ follows the rule in (1.1). Let $X_{j,k}$ be distributed as

$$X_{j,k} \sim (k+1)\,\mathrm{Binom}(j - \phi(m,k), f(k)^d).$$

On $\mathcal{G}_k$, the sum $\sum_{i=\phi(m,k)}^{j} \xi_i(k+1)$ is stochastically dominated by $X_{j,k}$.

Consider the event

$$\mathcal{E}(k+1) = \{\exists j \ge \phi(m,k+1) : \sum_{i=\phi(m,k)}^{j} \xi_i(k+1) > 3/2\mathbb{E}X_{j,k}\}.$$

On complement of $\mathcal{E}(k+1)$ we obtain, setting $l = j - \phi(m, k+1)$,

$$\begin{aligned}
\frac{F_j(k+1)}{2j} &= \frac{F_{\phi(m,k)}(k+1)}{2\phi(m,k+1)+l} + \frac{1}{2(\phi(m,k+1)+l)} \sum_{i=\phi(m,k)}^{\phi(m,k+1)+l} \xi_i(k+1) \\
&\le \frac{\phi(m,k)}{\phi(m,k+1)+l} + \frac{3(k+1)(\phi(m,k+1)+l-\phi(m,k))f^d(k)}{4(\phi(m,k+1)+l)} \\
&\le \frac{\phi(m,k)}{\phi(m,k+1)} + \frac{3}{4}(k+1)f^d(k) \\
&\le C^{d^{k+1}-d^{k+2}} + \frac{3}{4}f(k+1) \\
&\le \frac{1}{4}e^{-c_1 d^{k+1-k_0}} + \frac{3}{4}f(k+1) \\
&\le f(k+1),
\end{aligned}$$

where we have applied (3.2) in the fifth line. Hence we obtain that $\mathcal{E}(k+1)^c \subseteq \mathcal{G}_{k+1}$, and thus we may bound

$$\begin{aligned}
\mathbb{P}\left[\,\exists\ k\ , k_* \ge k > k_0\ :\ \mathcal{G}_k \text{ fails}\right] &\le \sum_{k=k_0}^{k_*-1} \mathbb{P}\left[\mathcal{G}_{k+1}^c \cap \mathcal{G}_k\right] \\
&\le \sum_{k=k_0}^{k_*-1} \mathbb{P}\left[\mathcal{E}_{k+1}\ \cap\ \mathcal{G}_k\right].
\end{aligned}$$

We estimate the probability of this event using standard Chernoff bounds. In the following $c > 0$ is an absolute constant.

$$\begin{aligned}\mathbb{P}\left[\mathcal{E}(k+1) \cap \mathcal{G}_k\right] &\leq \mathbb{P}\left[\ \exists\ j \geq \phi(m, k+1)\ :\ X_{j,k} > \tfrac{3}{2}\mathbb{E}X_{j,k}\right] \\ &\leq \sum_{l=0}^{m} e^{-c(\phi(m,k+1)-\phi(m,k))f^d(k)-clf^d(k)} \\ &= e^{-c(\phi(m,k+1)-\phi(m,k))f^d(k)} \sum_{l=0}^{m} e^{-clf^d(k)} \\ &\leq \frac{1}{1-e^{-cf^d(k)}} e^{-c(\phi(m,k+1)-\phi(m,k))f^d(k)}\end{aligned}$$

Here we use that $f(k) \leq f(k_0)$ and hence there is an absolute constant C' so that $C'(1 - e^{-cf^d(k)}) \geq f^d(k)$. Applying Lemma 3.2,

$$\begin{aligned}\mathbb{P}\left[\mathcal{E}(k+1) \cap \mathcal{G}_k\right] &\leq C' \frac{e^{-\rho(m)(C^{d^{k+2}}-C^{d^{k+1}})\exp(-c_1 d^{k-k_0})}}{\exp(-c_1 d^{k-k_0})} \\ &\leq C' \exp(-\rho(m)(Ce^{-c_1})^{d^k} + c_1 d^{k-k_0}).\end{aligned}$$

Therefore we may conclude that

$$\mathbb{P}\left[\ \exists\ k\ , k_* \geq k > k_0\ :\ \mathcal{G}(k) \text{ fails}\right]\ \leq\ \sum_{k=k_0}^{k_*} C' \exp(-\rho(m)(Ce^{-c_1})^{d^k} + c_1 d^{k-k_0}).$$

It can be checked that for m sufficiently large, this bound is monotone decreasing in k, and hence we have that

$$\mathbb{P}\left[\ \exists\ k\ , k_* \geq k > k_0\ :\ \mathcal{G}(k) \text{ fails}\right] \leq k_* \exp(-A\rho(m))$$

for some absolute constant A. As $k_* = O(\log\log m)$, this tends to 0 with m, which completes the proof of Lemma 3.3.

□

As a consequence, we have that for $m \geq j \geq \phi(m, k)$, and all $k_0 \leq k \leq k_*$

$$F_j(k_*) \leq F_j(k) \leq 2jf(k) \leq 2j\exp(-c_2 d^{k-k_0})$$

with high probability. We can therefore find some constant $\beta > 0$ so that

$$\frac{F_j(k_*)}{2j} \leq \left(\frac{\rho(m)}{\phi(m, k+1)}\right)^{\beta}$$

for $m \geq j \geq \phi(m, k)$, and all $k_0 \leq k \leq k_*$. For each $\log\log m \leq j \leq m$ we could find k such that $\phi(m, k+1) \geq j \geq \phi(m, k)$, which implies that there is a $\beta_0 > 0$ constant so that with high probability

$$\frac{F_j(k_*)}{2j} \leq \frac{1}{j^{\beta_0}} \tag{3.3}$$

for all $\log\log m/\log d \leq j \leq m$.

A large enough value of β_0 would complete the proof. If $\beta_0 > \frac{1}{d}$, then $F_j(k_* + 1)$ would be identically 0 for all j with high probability. Regardless of the value of β_0, it is possible to supercharge this result by letting the recurrence run a little longer.

Lemma 3.5. *If there is an absolute constant $M_1 > 0$ so that with high probability for some $k \leq \log\log m$,*

$$F_j(k) \leq 2(j)^{1-\beta},\ \forall\ j:\ m \geq j \geq (2\log\log m)^{M_1},$$

for some $\beta < \frac{1}{d}$, then there is an absolute constant $M_2 > 0$ so that with high probability

$$F_j(k+1) \leq 2(j)^{1-(d-0.5)\beta},\ \forall\ j:\ m \geq j \geq (2\log\log m)^{M_2}.$$

Proof: We let $\mathcal{C}$ be the event used as the hypothesis of the lemma. Set $j_0 = (2\log\log m)^{M_1}$. Then for $j \geq j_0$, we have that

$$\Delta_j = F_j(k+1) - F_{j_0}(k+1)$$

is dominated on $\mathcal{C}$ by $(k+1)$ times a sum of independent Bernoulli variables with means at most $j_0^{-d\beta}$. Thus, we may find an absolute constant $c > 0$ so that $\Delta_j/(k+1)$ is stochastically dominated by a Poisson variable X_j with mean

$$\mathbb{E}X_j = c\sum_{i=j_0}^{j} i^{-2\beta} \leq \frac{c}{1-2\beta}j^{1-2\beta},$$

with the inequality following by comparison with a Riemann sum. From standard tail bounds for Poisson variables, we may find a constant C' so that

$$\mathbb{P}\left[\exists\ j \geq j_0\ :\ \Delta_j \geq C'\mathbb{E}X_j\right] \leq C'\sum_{j=j_0}^{\infty} \exp(-j^{1-d\beta}/C'),$$

which is $o(1)$ using the hypothesis that $1 - d\beta > 0$. Thus it follows that with high probability

$$F_j(k+1) \leq F_{j_0}(k+1) + (k+1)C'j^{1-d\beta}.$$

As $k \leq \log\log m$, we have that $(k+1) \leq 2\log\log m$ for large enough m. Choose M_2 sufficiently large that both of $j_0(2\log\log m) \leq (2\log\log m)^{M_2(1-(d-0.5)\beta)}$ and $C'(2\log\log m) \leq (2\log\log m)^{0.5M_2\beta}$ for all m sufficiently large. Then we conclude for all $j \geq (2\log\log m)^{M_2}$,

$$F_j(k+1) \leq j^{1-(d-0.5)\beta} + j^{0.5\beta}j^{1-d\beta} = 2(j)^{1-(d-0.5)\beta},$$

as desired.

□

Lemma 3.6. *There is a $M = M(\beta_0) > 0$ and an integer $r = r(\beta_0) > 0$ so that setting $j_0 = (2\log\log m)^M$, then with high probability*

$$F_j(k_* + r) = F_{j_0}(k_* + r)$$

for all $m \geq j \geq j_0$.

Proof: We may apply Lemma 3.5 some $r'(\beta_0)$ many times to conclude that there is an $M = M(\beta_0)$ so that with high probability $F_j(k_* + r') \leq 2j^{1-\beta}$ for all $j \geq j_0$ and for some $\beta > \frac{1}{d}$.

Let $\mathcal{C}$ be the event

$$\mathcal{C} = \{F_j(k_* + r') \leq 2j^{1-\beta},\ \forall\ j \geq j_0\}.$$

It now follows from the usual recurrence argument that

$$\mathbb{P}\left[\exists\, j;\ j_0 \le j \le m\ :\ F_j(k_*+r'+1) > F_{j_0}(k_*+r'+1) \middle|\ \mathcal{C}\right] = O\left(\sum_{i=j_0}^{m} i^{-2\beta}\right) = o(1),$$

as $i^{-2\beta}$ is summable. Thus taking $r = r'+1$, we have shown the desired claim. □

We now prove the final theorem.

Proof of Theorem 1.1: From Lemma 3.6, it follows that with high probability,

$$F_m(k_*+r) = F_{j_0}(k_*+r).$$

As $F_{j_0}(k_*+r)$ is almost surely at most $O((\log\log m)^{M+1})$, it follows that the maximum degree of the P_m graph after m steps is $(\log\log m)^{M+1}$ with high probability. Note that $k_*+r = \Theta(j_0^{1/M_1})$, and hence with high probability, P_{j_0} has no vertices of degree k_*+r. Thus in fact, it follows that with high probability $F_{j_0}(k_*+r) = 0$, so that with high probability $F_m(k_*+r) = 0$ and the maximum degree of the graph is at most $k_*+r = \log\log m/\log d + \Theta(1)$.

We will now prove the lower bound. To do so we provide a coupling between the bin and ball model with d choices and our model. We will use Theorem 6 of Mitzenmacher et al. (2001) for the lower bound estimate on the maximum degree. Let us recall the ball and bin model. Suppose that n balls are sequentially placed into n bins (denote them by $v_1,\ldots,v_n$). Each ball is placed in the least full bin at the time of the placement, among d bins, chosen independently and uniformly at random. Theorem 6 of Mitzenmacher et al. (2001) provides that in this case after all the balls are placed the number of balls in the fullest bin is at least $\log\log n/\log d - \Theta(1)$ with high probability. With a slight change in the proof of this theorem it could be extended to n bins and ϵn balls with the same statement, where $0<\epsilon<1$ is some constant. From here we consider the model with $2m$ bins and m balls and we will use extension of Theorem 6 of Mitzenmacher et al. (2001) for $n = 2m$ and $\epsilon = 1/2$.

Let $N_j^0(k)$ be the number of bins that contain at least k balls at time j. We will need the following lemma.

Lemma 3.7. *There is a coupling such that for all $k \ge 1$ and $1 \le j \le m$*

$$N_j^0(k) \le F_j(k).$$

Note that with this lemma, the proof is now complete, as there is a $k'(m) = \log\log m/\log d - \Theta(1)$ so that with high probability $N_m^0(k') > 0$. And so we turn to proving the lemma by induction over j.

When $j = 1$, the lemma is trivial, as

$$N_1^0(k) = \mathbf{1}\{k=1\} \le 2\mathbf{1}\{k=1\} = F_1(k),\ k \ge 2.$$

Suppose the statement is true for $j \le j_0$. We will show the construction can be extended to $j_0+1 \le m$. The difference $N_{j_0+1}^0(k) - N_{j_0}^0(k)$ takes value 1 with probability

$$\frac{N_{j_0}(k-1)^d - N_{j_0}(k)^d}{n^d} \le \frac{N_{j_0}(k-1)^d}{n^d} \le \frac{F_{j_0}(k-1)^d}{n^d}.$$

If $j_0 \le n/2 = m$ this probability does not exceed $\frac{F_{j_0}(k-1)^d}{(2j)^d}$, and hence the difference $N^0_{j_0+1}(k) - N^0_{j_0}(k)$ is stochastically dominated by $F_{j_0+1}(k) - F_{j_0}(k)$. Therefore there is a coupling such that $N^0_{j_0+1}(k) \le F_{j_0+1}(k)$.

□

Acknowledgements.

The authors are grateful to Professor Itai Benjamini for suggesting the problem and for helpful discussions.

References

D. Achlioptas, R.M. D'Souza and J. Spencer. Explosive percolation in random networks. *Science* **323** (5920), 1453–1455 (2009). MR2502843.

Y. Azar, A.Z. Broder, A.R. Karlin and E. Upfal. Balanced allocations. *SIAM J. Comput.* **29** (1), 180–200 (1999). MR1710347.

A. Barabási and R. Albert. Emergence of scaling in random networks. *Science* **286** (5439), 509–512 (1999). MR2091634.

S. Dereich and P. Mörters. Random networks with sublinear preferential attachment: degree evolutions. *Electron. J. Probab.* **14**, no. 43, 1222–1267 (2009). MR2511283.

S. Dommers, R. van der Hofstad and G. Hooghiemstra. Diameters in preferential attachment models. *J. Stat. Phys.* **139** (1), 72–107 (2010). MR2602984.

R.M. D'Souza, P.L. Krapivsky and C. Moore. The power of choice in growing trees. *Eur. Phys. J. B* **59** (4), 535–543 (2007). MR2358054.

A. Flaxman, A. Frieze and T. Fenner. High degree vertices and eigenvalues in the preferential attachment graph. *Internet Math.* **2** (1), 1–19 (2005). MR2166274.

P.L. Krapivsky, S. Redner and F. Leyvraz. Connectivity of growing random networks. *Phys. Rev. Lett.* **85**, 4629–4632 (2000). DOI: 10.1103/PhysRevLett.85.4629.

Y. Malyshkin and E. Paquette. The power of choice combined with preferential attachment. *Electron. Commun. Probab.* **19**, no. 44, 13 (2014). MR3233206.

M. Mitzenmacher, A.W. Richa and R. Sitaraman. The power of two random choices: a survey of techniques and results. In *Handbook of randomized computing, Vol. I, II*, volume 9 of *Comb. Optim.*, pages 255–312. Kluwer Acad. Publ., Dordrecht (2001). MR1966907.

T.F. Móri. The maximum degree of the Barabási-Albert random tree. *Combin. Probab. Comput.* **14** (3), 339–348 (2005). MR2138118.

R. Oliveira and J. Spencer. Connectivity transitions in networks with super-linear preferential attachment. *Internet Math.* **2** (2), 121–163 (2005). MR2193157.

B. Pittel. Note on the heights of random recursive trees and random m-ary search trees. *Random Structures Algorithms* **5** (2), 337–347 (1994). MR1262983.

O. Riordan and L. Warnke. Achlioptas process phase transitions are continuous. *Ann. Appl. Probab.* **22** (4), 1450–1464 (2012). MR2985166.

A. Rudas, B. Tóth and B. Valkó. Random trees and general branching processes. *Random Structures Algorithms* **31** (2), 186–202 (2007). MR2343718.