A Weak Local Limit for Preferential Attachment Graphs

Noam Berger[a]*, Christian Borgs[b], Jennifer T. Chayes[c], Amin Saberi[d]†

[a] *Mathematics Department, UCLA, Los Angeles, CA 90095.*
[b] *Microsoft Research New England, Cambridge, MA 02142.*
[c] *Management Science and Engineering, Stanford University, Palo Alto, CA 94305.*

August 19, 2009

Abstract

We give an explicit construction of the weak local limit of a class of preferential attachment graphs. This limit contains all local information and allows several computations that are otherwise hard, for example joint degree distributions and, more generally, the limiting distribution of subgraphs in balls of any given radius k around a random vertex in the preferential attachment graph.

1 Introduction

In early nineties, it was realized that the Internet has a power-law degree distribution [FFF99, AJB99]. This observation was justified theoretically by the so-called preferential attachment model of Barabási and Albert [BA99], which was later used to explain the observed power-law degree sequence of a host of real-world networks, including social and biological networks, in addition to technological ones. The analysis of [BA99] was non-rigorous; the first rigorous analysis of a preferential attachment model, in particular proving that it has small diameter, was given by Bollobás and Riordan [BR04]. Since these works there has been a tremendous amount of study, both non-rigorous and rigorous, of the random graph models that explain the power-law degree distribution; see [AB02] and [BR02] and references therein for some of the non-rigorous and rigorous work, respectively.

Also motivated by the growing graphs appearing in real-world networks, for the past five years or so, there has been much study in the mathematics community of notions of graph limits. In this context, most of the work has focused on dense graphs. In particular, there

---

*This work was done while the first and last authors were visiting Microsoft Research.

† Email Addresses: berger@math.ucla.edu, borgs@microsoft.com, jchayes@microsoft.com, saberi@stanford.edu.

have been a series of papers on a notion of graph limits defined via graph homomorphisms [BCL$^+$06b, BCL$^+$06a, BCL$^+$07, LS06]; these have been shown to be equivalent to limits defined in many other senses [BCL$^+$06a, BCL$^+$07]. Although most of the results in this work concern dense graphs, the paper [BCL$^+$06b] also introduces a notion of graph limits for sparse graphs with bounded degrees in terms of graph homomorphisms; using expansion methods from mathematical physics, [BCL$^+$] proves some general results on this type of limit for sparse graphs. Another recent work [BR07] concerns limits for graphs which are neither dense nor sparse in the above senses; they have average degrees which tend to infinity.

Earlier, a notion of a weak local limit of a sequence of graphs with bounded degrees was given by Benjamini and Schramm [BS01] (this notion was in fact already implicit in [Ald98]). Interestingly, it is not hard to show that the Benjamini-Schramm limit coincides with the limit defined via graph homomorphisms in the case of sparse graphs of bounded degree; see [Ele07] for yet another equivalent notion of convergent sequences of graphs with bounded degrees.

As observed by Lyons [Lyo05], the notion of graph convergence introduced by Benjamini and Schramm is meaningful even when the degrees are unbounded, provided the *average degree* stays bounded. Since the average degree of the Barabási-Albert graph is bounded by construction, it is therefore natural to ask whether this graph sequence converges in the sense of Benjamini and Schramm.

The analysis of this problem is the subject of this paper. We establish the existence of the Benjamini-Schramm limit for the the Barabási-Albert graph by giving an explicit construction of the limit process, and use it to derive various properties of the limit.

Our proof uses a representation which we first introduced in [BBCS05] to analyze processes which model the spread of viral infections on preferential attachment graphs. Our representation expresses the preferential attachment model process as a combination of several Pólya urn processes. The classic Pólya urn model was of course proposed and analyzed in the beautiful work of Pólya and Eggenberger in the early twentieth century [EP23]; see [Dur96] for a basic reference. Despite the fact that our Pólya urn representation is *a priori* only valid for a limited class of preferential attachment graphs, we give an approximating coupling which proves that the limit constructed here is the limit of a much wider class of preferential attachment graphs.

Our alternative representation contains much more independence than previous representations of preferential attachment and is therefore simpler to analyze. In order to demonstrate this, we also give a few applications of the limit. In particular, we show how to use the limit to calculate the degree distribution and the joint degree distribution of a typical vertex with its parent.

2 Definition of the Model and Statements of Results

2.1 Definition of the Model

The preferential attachment graph we define generalizes the model introduced by Barabási and Albert [BA99] and rigorously constructed in [BR04]. Fix an integer $m \geq 2$ and a real number $0 \leq \alpha < 1$. We will construct a sequence of graphs (G_n) (where G_n has n vertices labeled $1, \ldots, n$) as follows:

G_1 contains one vertex and no edges, and G_2 contains two vertices and m edges connecting them. Given G_{n-1} we create G_n the following way: We add the vertex n to the graph, and choose m vertices $w_1, ..., w_m$, possibly with repetitions, from G_{n-1}. Then we draw edges between n and each of $w_1, ..., w_m$. Repetitions in the sequence $w_1, ..., w_m$ result in multiple edges in the graph G_n.

We suggest three different ways of choosing the vertices $w_1, ..., w_m$. The first two ways, the independent and the conditional, are natural ways which we consider of interest, and are the two most common interpretations of the preferential attachment model. The third way, i.e. the sequential model, is less natural, but is much easier to analyze because it is exchangeable, and therefore by De-Finetti's Theorem (see [Dur96]) has an alternative representation, which contains much more independence. We call this representation the Pólya urn representation because the exchangeable system we use is the Pólya urn scheme.

1. The independent model: $w_1, ..., w_m$ are chosen independently of each other conditioned on the past, where for each $i = 1, \ldots, m$, we choose w_i as follows:

   With probability α, we choose w_i uniformly from the vertices of G_{n-1}, and with probability $1 - \alpha$, we choose w_i according to the preferential attachment rule, i.e., for all $k = 1, \ldots, n-1$,
   $$\mathbf{P}\left(w_i = k\right) = \frac{\deg_{n-1}(k)}{Z}$$
   where Z is the normalizing constant $Z = \sum_{k=1}^{n-1} \deg_{n-1}(k) = 2m(n-2)$.

2. The conditional model: Here we start with some predetermined graph structure for the first m vertices. Then at each step, $w_1, ..., w_m$ are chosen as in the independent case, *conditioned* on them being different from one another.

3. The sequential model: $w_1, ..., w_m$ are chosen inductively as follows: With probability α, w_1 is chosen uniformly, and with probability $1 - \alpha$, w_1 is chosen according to the preferential attachment rule, i.e., for every $k = 1, \ldots, n-1$, we take $w_1 = k$ with probability $(\deg_{n-1}(k))/Z$ where as before $Z = 2m(n-2)$. Then we proceed inductively, applying the same rule, but with two modifications:

   (a) When determining w_i, instead of the degree $\deg_{n-1}(k)$, we use
   $$\deg'_{n-1}(k) = \deg_{n-1}(k) + \#\{1 \leq j \leq i-1 | w_j = k\}$$
   and normalization constant
   $$Z' = \sum_{k=1}^{n-1} (\deg'_{n-1}(k)) = 2m(n-2) + i - 1.$$

(b) The probability of uniform connection will be

$$\tilde{\alpha} = \alpha \frac{2m(n-1)}{2m(n-2) + 2m\alpha + (1-\alpha)(i-1).} = \alpha + O(n^{-1}) \tag{1}$$

rather than α.

We will refer to all three models as versions of the preferential attachment graph, or PA-graph, for short. Even though we consider the graph G_n as undirected, it will often be useful to think of the vertices $w_1, \ldots, w_n$ as vertices which "received an edge" from the vertex n, and of n as a vertex which "sent out m edges" to the vertices $w_1, \ldots, w_n$. Note in particular, that the degree of a general vertex v in G_n can be written as $m + q$, where m is the number of edges sent out by v and q is the (random) number of edges received by v.

2.2 Pólya Urn Representation of the Sequential Model

Our first theorem gives the Pólya urn representation of the sequential model. To state it, we use the standard notation $X \sim \beta(a, b)$ for a random variable $X \in [0, 1]$ whose density is equal to $\frac{1}{Z} x^{a-1}(1-x)^{b-1}$, where $Z = \int_0^1 x^{a-1}(1-x)^{b-1} dx$. We set

$$u = \frac{\alpha}{1-\alpha}.$$

Theorem 2.1. *Fix m, α and n. Let $\psi_1 = 1$, let $\psi_2, \ldots, \psi_n$ be independent random variables with*

$$\psi_j \sim \beta\left(m + 2mu, (2j-3)m + 2mu(j-1)\right), \tag{2}$$

and let

$$\varphi_j = \psi_j \prod_{i=j+1}^{n} (1 - \psi_i), \qquad S_k = \sum_{j=1}^{k} \varphi_j, \qquad \text{and} \qquad I_k = [S_{k-1}, S_k). \tag{3}$$

Conditioned on $\psi_1, \ldots, \psi_n$, choose $\{U_{k,i}\}_{k=1,\ldots n,\, i=1,\ldots,m}$ as a sequence of independent random variables, with $U_{i,k}$ chosen uniformly at random from $[0, S_{k-1}]$. Join two vertices j and k if $j < k$ and $U_{k,i} \in I_j$ for some $i \in \{1, \ldots, m\}$ (with multiple edges between j and k if there are several such i). Denote the resulting random multi-graph by G_n.

Then G_n has the same distribution as the sequential PA-graph.

Figure 1 illustrates this theorem.

It should be noted that the $\alpha = 0$ case of the sequential model defined here differs slightly from the model of Bollobás and Riordan [BR04] in that they allow (self-)loops, while we do not. In fact, a minor alteration of our Pólya urn representation models their graph, and we suspect that a minor alteration of their pairing representation can model our graph.

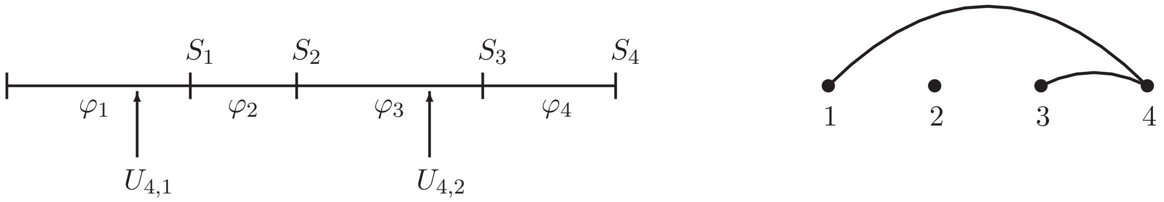

Figure 1: The Pólya-representation of the sequential model for $m = 2$, $n = 4$, and $k = 4$. The variables $U_{4,1}$ and $U_{4,2}$ are chosen uniformly at random from $[0, S_3]$.

2.3 Definition of the Pólya-point graph model.

2.3.1 Motivation

The Benjamini-Schramm notion [BS01] of weak convergence involves the view of the graph G_n from the point of view of a "root" k_0 chosen uniformly at random from all vertices in G_n. More precisely, it involves the limit of the sequence of balls of radius $1, 2, \dots,$ about this root, see Definition 2.1 in Section 2.4 below for the details.

It turns out that for the sequential model, this limit is naturally described in terms of the random variables S_{k-1} introduced in Theorem 2.1. To explain this, it is instructive to first consider the ball of radius 1 around the random root k_0. This ball contains the m neighbors of k_0 that were born before k_0 and received an edge from k_0 under the preferential attachment rule described above, as well as a random number q_0 of neighbors that were born after k_0 and send an edge to k_0 at the time they were born. We denote these neighbors by $k_{01}, \dots, k_{0m}$ and $k_{0,m+1}, \dots, k_{0,m+q_0}$, respectively.

Let

$$\chi = \frac{1+2u}{2+2u} \qquad \text{and} \qquad \psi = \frac{1-\chi}{\chi}, \tag{4}$$

and note that $\frac{1}{2} \leq \chi < 1$ and $0 < \psi \leq 1$. As we will see, the random variables S_{k-1} behave asymptotically like $(k/n)^\chi$, implying in particular that the distribution of S_{k_0-1} tends to that of a random variable $x_0 = y_0^\chi$, where y_0 is chosen uniformly at random in $[0, 1]$. The limiting distribution of $S_{k_{01}-1}, \dots, S_{k_{0m}-1}$ turns out to be quite simple as well: in the limit these random variables are i.i.d. random variables x_{0i} chosen uniformly from $[0, x_0]$, a distribution which is more or less directly inherited from the uniform random variables $U_{k,i} \in [0, S_{k_0-1}]$ from Theorem 2.1. The limiting distribution of the random variables $S_{k_{0,m+1}-1}, \dots, S_{k_{0,m+q_0}-1}$ is slightly more complicated to derive, and is given by a Poisson process in $[x_0, 1]$ with intensity

$$\frac{\gamma_0}{2(u+1)\chi} \frac{x^{\psi-1}}{x_0^\psi} dx.$$

Here γ_0 is a "random strength" which is essentially the original β-distributed random variable ψ_{k_0}. We will see that after the normalization and in the limit the distribution of γ_0 will converge to $\Gamma(m+2mu,1)$. As usual, we use $\Gamma(a,b)$ to denote a distribution on $[0,\infty)$ which has density $\frac{1}{Z}x^{a-1}e^{-bx}$, with $Z=\int_0^\infty x^{a-1}e^{-bx}dx$.

Next, consider the branching that results from exploring the neighborhood of a random vertex in G_n in a ball of radius bigger than one. In each step of this exploration, we will find two kinds of children of the current vertex k: those which were born before k, and were attached to k at the birth of k, and those which were born after k, and were connected to k at their own birth. There are always either m or $m-1$ children of the first kind (if k was born after its parent, there will be $m-1$ such children, since one of the m edges sent out by k was sent out to k's parent, otherwise there will be m children of the first type). The number of children of the second kind is a random variable.

In the limit $n\to\infty$, this branching process leads to a random tree whose vertices, $\bar{a}$, carry three labels: a "strength" $\gamma_{\bar{a}}\in(0,\infty)$ inherited from the β-random variables ψ_k, a "position" $x_{\bar{a}}\in[0,1]$ inherited from the random variables S_{k-1}, and a type which can be either E (for "early") or L (for "late"), reflecting whether the vertex k was born before or after its parent. While the strengths of vertices of type L turns out to be again $\Gamma(m+2mu,1)$-distributed, this is not the case for vertices of type E, since a vertex with higher values of ψ_k has a larger probability of receiving an edge from its parent. In the limit, this will be reflected by the fact that the strength of vertices of type E is $\Gamma(m+2mu+1,1)$-distributed.

2.3.2 Formal Definition

The main goal of the previous subsection was to give an intuition of the structure of the neighborhood of a random vertex. We will show that asymptotically, the branching process obtained by exploring the neighborhood of a random vertex k_0 in G_n is given by a random tree with a certain distribution. In order to state our main theorem, we give a formal definition of this tree.

Let F be the Gamma distribution $\Gamma(m+2mu,1)$, and let F' be the Gamma distribution $\Gamma(m+2mu+1,1)$. We define a random, rooted tree $(T,0)$ with vertices labeled by finite sequences

$$\bar{a}=(0,a_1,a_2,\ldots,a_l)$$

inductively as follows:

- The root 0 has a position $x_0=y_0^\chi$, where y_0 is chosen uniformly at random in $[0,1]$.
- In the induction step, we assume that $\bar{a}=(0,a_1,a_2,\ldots,a_l)$ and the corresponding variable $x_{\bar{a}}\in[0,1]$ have been chosen in a previous step. Define $(\bar{a},j)$ as $(0,a_1,a_2,\ldots,a_l,j)$, $j=1,2,\ldots$, and set

  $$m_-(\bar{a})=\begin{cases} m & \text{if } \bar{a} \text{ is the root or of type } E\\ m-1 & \text{if } \bar{a} \text{ is of type } L.\end{cases}$$

  We then take

  $$\gamma_{\bar{a}}\sim\begin{cases} F & \text{if } \bar{a} \text{ is the root or of type } L\\ F' & \text{if } \bar{a} \text{ is of type } E.\end{cases}$$

independently of everything previously sampled, choose $x_{(\bar{a},1)}, \ldots, x_{(\bar{a},m_-(\bar{a}))}$ i.i.d. uniformly at random in $[0, x_{\bar{a}}]$, and $x_{(\bar{a},m_-(\bar{a})+1)} \ldots, x_{(\bar{a},m_-(\bar{a})+q_{\bar{a}})}$ as the points of a Poisson process with intensity

$$\rho_{\bar{a}}(x)dx = \frac{\gamma_{\bar{a}}}{2(u+1)\chi} \frac{x^{\psi-1}}{x_{\bar{a}}^{\psi}} dx \tag{5}$$

on $[x_{\bar{a}}, 1]$. The children of $\bar{a}$ are the vertices $(\bar{a}, 1)$, ..., $(\bar{a}, m_-(\bar{a}) + q_{\bar{a}})$, with $(\bar{a}, 1)$, ..., $(\bar{a}, m_-(\bar{a}))$ called of type E, and the remaining ones called of type L.

We continue this process ad infinitum to obtain an infinite, rooted tree $(T, 0)$. We call this tree the Pólya-point graph, and the point process $\{x_{\bar{a}}\}$ the Pólya-point process.

2.4 Main Result

We are now ready to formulate our main result, which states that in all three versions, the graph G_n converges to the Pólya-point graph in the sense of [BS01].

Let $\mathcal{G}$ be the set of rooted graphs, i.e., the set of all pairs (G, x) consisting of a connected graph G and a designated vertex x in G, called the root. Two rooted graphs $(G, x), (G', x') \in \mathcal{G}$ are called isomorphic if there is an isomorphism from G to G' which maps x to x'. Given a finite integer r, we denote the rooted ball of radius r around x in $(G, x) \in \mathcal{G}$ by $B_r(G, x)$. We then equip $\mathcal{G}$ with the σ-algebra generated by the events that $B_r(G, x)$ is isomorphic to a finite, rooted graph (H, y) (with r running over all finite, positive integers, and (H, y) running over all finite, rooted graphs), and call (G, x) a random, rooted graph if it is a sample from a probability distribution on $\mathcal{G}$.

Definition 2.1. *Given a sequence of random, finite graphs G_n, let $k_0^{(n)}$ be a uniformly random vertex from G_n. Following [BS01], we say that an infinite random, rooted graph (G, x) is the weak local limit of G_n if for all finite rooted graphs (H, y) and all finite r, the probability that $B_r(G_n, k_0^{(n)})$ is isomorphic to (H, y) converges to the probability that $B_r(G, x)$ is isomorphic to (H, y).*

The main result of the paper is the following theorem.

Theorem 2.2. *The weak local limit of the all three variations of the Preferential Attachment model is the Pólya-point graph.*

As alluded to before, the points $x_{\bar{a}}$ of the Pólya-point process represent the random variables S_{k-1} of the vertices in G_n, which in turn behave like $(k/n)^{\chi}$ as $n \to \infty$. The variable $y_{\bar{a}} = x_{\bar{a}}^{1/\chi}$ thus represents the birth-time of the corresponding vertex in G_n. This is made precise in the following corollary to the proof of Theorem 2.2. As the theorem, the corollary holds for all three versions of the Preferential Attachment model.

Corollary 2.3. *Given $r < \infty$ and $\epsilon > 0$ there exists a $n_0 < \infty$ such that for $n \geq n_0$, the Pólya-point process and the preferential attachment graph G_n can be coupled in such a way* *that with probability at least* $1-\epsilon$*, there exists an isomorphism* $\bar{a} \mapsto k_{\bar{a}}$ *from ball of radius* r *about* 0 *in* $(T,0)$ *into the ball of radius* r *about a uniformly random vertex* k_0 *in* G_n*, with the property that*

$$\left| y_{\bar{a}} - \frac{k_{\bar{a}}}{n} \right| \le \epsilon$$

for all $\bar{a}$ *with distance at most* r *from the root in* $(T,0)$*. Here* $y_{\bar{a}}$ *is defined as* $y_{\bar{a}} = x_{\bar{a}}^{1/\chi}$*.*

The numerator $x_{\bar{a}}^{\psi} = y_{\bar{a}}^{1-\chi}$ in (5) thus expresses the fact that in the preferential attachment process, earlier vertices are more likely to attract many neighbors than later vertices.

3 Proof of Weak Distributional Convergence for the Sequential Model

In this section we prove that the sequential model converge to the Pólya-point tree.

3.1 Pólya Urn Representation of the Sequential Model

In early twentieth century, Pólya proposed and analyzed the following model known as the Pólya urn model (see [Dur96]). The model is described as follows. We have a number of urns, each holding a number of balls, and at each step, a new ball is added to one of the urns. The probability that the ball is added to urn i is proportional to $N_i + u$ where N_i is the number of balls in the i-th urn and u is a predetermined parameter of the model.

Pólya showed that this model is equivalent to another process as follows. For every i, choose at random a parameter (which we call "strength" or "attractiveness") p_i, and at each step, *independently* of our decision in previous steps, put the new ball in urn i with probability p_i. Pólya specified the distribution (as a function of u and the initial number of balls in each urn) for which this mimics the urn model. A particularly nice example is the case of two urns, each starting with one ball and $u = 0$. Then p_1 is a uniform $[0,1]$ variable, and $p_2 = 1 - p_1$. Pólya showed that for general values of u and $\{N_i(0)\}$, the values of $\{p_i\}$ are determined by the β-distribution with appropriate parameters.

It is not hard to see that there is a close connection between the the preferential attachment model of Barabási and Albert and the Pólya urn model in the following sense: every new connection that a vertex gains can be represented by a new ball added in the urn corresponding to that vertex. We use this idea to give an equivalent description of the preferential attachment graph which is easier to analyze. We will see throughout the paper the properties of this description that make it useful for understanding the graph.

To derive this representation, let us consider first a two urn model, with the number of balls in one urn representing the degree of a particular vertex k, and the number of balls in the other representing the sum of the degrees of the vertices $1, \dots, k-1$. We will start this process at the point when $n = k$ and k has connected to precisely m vertices in $\{1, \dots, k-1\}$. Note that at this point, the urn representing the degree of k has m balls, while the other one has $(2k-3)m$ balls.

Consider a time in the evolution of the preferential attachment model when we have $n-1 \geq k$ old vertices, and $i-1$ edges between the new vertex n and $\{1, \ldots, k-1\}$. Assume that at this point the degree of k is d_k, and the sum of the degrees of $1 \ldots, k-1$ is $d_{<k}$. At this point, the probability that the i^{th} edge from n to $\{1, \ldots, n-1\}$ is attached to k is

$$\tilde{\alpha}\frac{1}{n-1} + (1-\tilde{\alpha})\frac{d_k}{2m(n-2)+(i-1)}$$

while the probability that it is attached to one of the nodes $1, \ldots, k-1$ is

$$\tilde{\alpha}\frac{k-1}{n-1} + (1-\tilde{\alpha})\frac{d_{<k}}{2m(n-2)+(i-1)}.$$

Thus, conditioned on connecting to $\{1, \ldots, k\}$, the probability that the i^{th} edge from n to $\{1, \ldots, n-1\}$ is attached to k is

$$\frac{1}{Z}\big(2mu + d_k\big)$$

while the conditional probability that it is attached to one of the nodes $1, \ldots, k-1$ is

$$\frac{1}{Z}\big(2mu(k-1) + d_{<k}\big)$$

where

$$u = \frac{\tilde{\alpha}}{1-\tilde{\alpha}} \frac{(n-2)+(i-1)/2m}{n-1}$$

and Z is an appropriate normalization constant. Note that our definition of $\tilde{\alpha}$ was chosen in such a way that the above definition of u agrees with that from (4). Taking into account that the two urns start with m and $(2k-3)m$ balls, respectively, we see that the evolution of the two bins is a Pólya urn with strengths ψ_k and $1-\psi_k$, where $\psi_k \sim B_k = \beta\left(m + 2mu, (2k-3)m + 2mu(k-1)\right)$.

Proof:[Proof of Theorem 2.1] Using the two urn process as an inductive input, we can now easily construct the Pólya graph defined in Theorem 3.5. Indeed, let $X_t \in \{1, 2, \ldots, \lceil \frac{t}{m} \rceil\}$ be the vertex receiving the t^{th} edge in the sequential model (the other endpoint of this edge being the vertex $\lceil \frac{t}{m} \rceil + 1$). For $t \leq m$, X_t is deterministic (and equal to 1), but starting at $t = m+1$, we have a two urn model, starting with m balls in each urn. As shown above, the two urns can be described as Pólya-urns with strengths $1-\psi_2$ and ψ_2. Once $t > 2m$, X_t can take three values, but conditioned on $X_t \leq 2$, the process continues to be a two urn model with strengths $1-\psi_2$ and ψ_2. To determine the probability of the event that $X_t \leq 2$, we now use the above two urn model with $k=3$, which gives that the probability of the event $X_t \leq 2$ is $1-\psi_3$, at least as long as $t \leq 3m$. Combining these two urn models, we get a three urn model with strengths $(1-\psi_2)(1-\psi_3)$, $\psi_2(1-\psi_3)$, and ψ_3. Again, this model remains valid for $t > 3m$, as long as we condition on $X_t \leq 3$.

Continuing inductively, we see that the sequence X_t evolves in stages:

- For $t = 1, \dots, m$, the variable X_t is deterministic: $X_t = 1$.
- For $t = m+1, \dots, 2m$, the distribution of $X_t \in \{1, 2\}$ is described by a two urn model with strengths $1 - \psi_2$ and ψ_2, where $\psi_2 \sim B_2$.
- In general, for $t = m(k-1)+1, \dots, km$, the distribution of $X_t \in \{1, \dots, k\}$ is described by a k urn model with strengths

  $$\varphi_j^{(k)} = \psi_j \prod_{i=j+1}^{k} (1 - \psi_i), \qquad j = 1, \dots k. \tag{6}$$

  Here $\psi_k \sim B_k$ is chosen at the beginning of the k^{th} stage, independently of the previously chosen strengths $\psi_1, \dots, \psi_{k-1}$ (for convenience, we set $\psi_1 = 1$).

Note that the random variables $\varphi_j^{(k)}$ can be expressed in terms of the random variables introduce in Theorem 2.1 as follows: by induction on k, it is easy to show that

$$S_k = \prod_{j=k+1}^{n} (1 - \psi_k). \tag{7}$$

This implies that

$$\varphi_j^{(k)} = \frac{\psi_j}{S_k},$$

which relates the strengths $\varphi_j^{(k)}$ to the random variables defined in Theorem 2.1, and shows that the process derived above is indeed the process given in the theorem. $\Box$

In order to apply Theorem 2.1, we will use two technical lemmas, whose proofs will be deferred to a later section. The first lemma states a law of large numbers for the random variables S_k.

Lemma 3.1. *For every ϵ there exist $K < \infty$ such that for $n \geq K$, we have that with probability at least $1 - \epsilon$,*

$$\max_{k \in \{1, \dots, n\}} \left| S_k - \left(\frac{k}{n}\right)^{\chi} \right| \leq \epsilon$$

and

$$\max_{k \in \{K, \dots, n\}} \left| S_k - \left(\frac{k}{n}\right)^{\chi} \right| \leq \epsilon \left(\frac{k}{n}\right)^{\chi}.$$

The second lemma concerns a coupling of the sequence $\{\psi_k\}_{k \geq 1}$ and an i.i.d. sequence of Γ-random variables $\{\chi_k\}_{k \geq 1}$, where $\chi_k \sim \Gamma(m + 2mu, 1)$. To describe the coupling, we define a sequence of functions $f_k : [0, \infty) \to [0, 1)$ by

$$\mathbf{P}(\psi_k \leq f_k(x)) = \mathbf{P}(\chi_k \leq x). \tag{8}$$

Then $f_k(\chi_k)$ has the same distribution as ψ_k, implying that $(\{\chi_k\}_{k \geq 1}, \{f_k(\chi_k)\}_{k \geq 1})$ defines a coupling between $\{\chi_k\}_{k \geq 1}$ and $\{\psi_k\}_{k \geq 1}$.

Lemma 3.2. *Let* f_k *be as in* (8), *and let* $\{\chi_k\}_{k\geq 1}$ *i.i.d. random variables with distribution* $\Gamma(m+2mu,1)$. *Given* $\epsilon>0$ *there exist a* $K<\infty$ *so that the following holds:*
(i) With probability at least $1-\epsilon$,

$$\chi_k \leq \log^2 k \quad \textit{for all} \quad k \geq K; \tag{9}$$

(ii) For $k \geq K$ *and* $x \leq \log^2 k$,

$$\frac{1-\epsilon}{2mk(1+u)} x \leq f_k(x) \leq \frac{1+\epsilon}{2mk(1+u)} x. \tag{10}$$

We defer the proof of Lemmas 3.1 and 3.2 to Section 3.6.

3.2 The Exploration Tree of G_n

Let $K_r = K_r(G_n, k_0)$ be the set of vertices in G_n which have distance at most r from the random root k_0, and let $\hat{B}_r(G_n, k_0)$ be the graph on K_r that contains all edges in G_n for which at most one endpoint has distance $\leq r$ from k_0. When proving that the preferential model converges to the Pólya-point graph, we will use the notion of convergence given in Definition 2.1, but instead of the standard ball of radius r, we will use the ball modified ball $\hat{B}_r(G_n, k_0)$. (It is obvious that this definition is equivalent).

We will prove our results by induction on r, using the exploration procedure outlined in Section 2.3.1 in the inductive step. To this end, it will be convenient to endow the rooted graph (G_n, k_0) with a structure which is similar to the one defined for the Pólya-point graph. More precisely, we will inductively define a rooted tree $(T_r^{(n)}, 0)$ on sequence of integers $\bar{a} = (0, a_1, a_2, \ldots, a_l)$, and a homomorphism

$$\mathbf{k}^{(r)} : T_r^{(n)} \to \hat{B}_r(G_n, k_0) : \bar{a} \mapsto k_{\bar{a}}$$

as follows.

We start our inductive definition by mapping 0 into a vertex k_0 chosen uniformly at random from the vertex set $\{1, \ldots, n\}$ of G_n. Given a vertex $\bar{a} = (0, a_1, a_2, \ldots, a_l) \in T_r^{(n)}$ and its image $k_{\bar{a}}$ in G_n, let $d_{\bar{a}}$ be the degree of $k_{\bar{a}}$ in G_n, and let $k_{\bar{a}_-}, k_1, \ldots, k_{d_{\bar{a}}-1}$ be the neighbors of $k_{\bar{a}}$ in G_n, where $\bar{a}_- = (0, a_1, a_2, \ldots, a_{l-1})$. Recalling that edges were created one by one during the sequential preferential attachment process, we order $k_1, \ldots, k_{d_{\bar{a}}-1}$ in such a way that for all $i = 1, \ldots, d_{\bar{a}} - 2$, the edge $(k_{\bar{a}}, k_i)$ was born before the edge $(k_{\bar{a}}, k_{i+1})$. We then define the children of $\bar{a}$ to be the points $(\bar{a}, 1), \ldots, (\bar{a}, d_{\bar{a}} - 1)$. This defines $T_{r+1}^{(n)}$. The map $\mathbf{k}^{(r+1)}$ is the extension of $\mathbf{k}^{(r)}$ which maps $(\bar{a}, 1), \ldots, (\bar{a}, d_{\bar{a}} - 1)$ to the vertices $k_1, \ldots, k_{d_{\bar{a}}-1}$, respectively. We call a vertex $(\bar{a}, i)$ early or of type E if $k_i < k_{\bar{a}_-}$ and late or of type L otherwise. Note that the root and vertices of type E have m children of type E, while vertices of type L have $m-1$ children of type E.

To make the dependence on G_n explicit, we often use the notation $T_r(G_n)$ for the tree $T_r^{(n)}$, and the notation $\mathbf{k}^{(r)}(G_n)$ for the map $\mathbf{k}^{(r)}$. Note that $\mathbf{k}^{(r)}$ does not, in general, give a graph isomorphism between $T_r^{(n)}$ and $\hat{B}_r(G_n, k_0)$. But if the map is injective when restricted to $T_r^{(n)}$, it is a graph isomorphism. To prove Theorem 2.2, it is therefore enough to show that for all r, the map $\mathbf{k}^{(r)}$ is injective and the tree $T_r^{(n)}$ converges in distribution to T_r, the ball of radius r in the Pólya-point graph $(T, 0)$.

3.3 Regularity Properties of the Pólya-Point Process

In order to prove Theorem 2.2, we will use some simple regularity properties of the Pólya-point process.

Recall the definition of the Pólya-point graph $(T, 0)$ and the Pólya-point process $\{x_{\bar{a}}\}$ from Section 2.3.2, as well as the notation $\rho_{\bar{a}}(x)dx$ for the intensity defined in (5). As usual, we define the height of a vertex $\bar{a} = (0, a_1, a_2, \ldots, a_l)$ in T as it's distance l from the root. We denote the rooted subtree of height r in $(T, 0)$ by $(T_r, 0)$.

Lemma 3.3. *Fix $0 \leq r < \infty$ and $\epsilon > 0$. Then there are constants $\delta > 0$, $C < \infty$, $K < \infty$ and $N < \infty$ such that with probability at least $1 - \epsilon$, we have that*

- $x_{\bar{a}} \geq \delta$ *for all vertices* $\bar{a}$ *in* T_r
- $\gamma_{\bar{a}} \leq C$
- $\rho_{\bar{a}}(\cdot) \leq K$
- $|T_r| \leq N$

Proof: The proof of the Lemma is easily obtained by induction on r. We leave it to the reader. □

Corollary 3.4. *For all $\epsilon > 0$ and all $r < \infty$ there is a constant $\delta > 0$ such that with probability at least $1 - \epsilon$, we have*

$$\min_{\substack{\bar{a}, \bar{b} \in T_r \\ \bar{a} \neq \bar{b}}} |x_{\bar{b}} - x_{\bar{a}}| \geq \delta.$$

Proof: This is an immediate consequence of the continuous nature of the random variables $x_{\bar{a}}$ and the statements of Lemma 3.3. □

3.4 The Neighborhood of Radius One

Before proving our main theorem, Theorem 2.2, for the sequential model, we establish the following lemma, which will serve as the base in an inductive proof of our main theorem.

Lemma 3.5. *Let G_n be the sequential preferential attachment graph, let k_0 be chosen uniformly at random in $\{1, \ldots, n\}$, and let $k_{01}, \ldots, k_{0,m+q_0}$ be the neighbors of k_0, ordered as in Section 3.2 by the birth times of the edges $\{k_0, k_{0i}\}$. Then (G_n, k_0) and the Pólya-point process $\{x_{\bar{a}}\}$ can be coupled in such a way that for all $\epsilon > 0$ there are constants $C, N < \infty$, $\delta > 0$ and $n_0 < \infty$ such that for $n \geq n_0$, with probability at least $1 - \epsilon$, we have that*

1. $T_1 = T_1(G_n)$ *and* $|T_1(G_n)| \le N$*;*

2. $|x_{\bar a} - S_{k_{\bar a}-1}| \le \epsilon$ *for all* $\bar a \in T_1$*,*

3. $k_0, k_{01}, \dots, k_{0m+q_0}$ *are pairwise distinct and* $k_{\bar a} \ge \delta' n$ *for all* $\bar a \in T_1$*;*

4. $\chi_{k_{\bar a}} = \gamma_{\bar a} \le C \quad$ *for all* $\quad \bar a \in T_1$.

Proof:(i-ii): We start by proving the first two statements. Choose y_0 uniformly at random in $[0,1]$, let $x_0 = y_0^\chi$, and let $x_{0,1}, \dots, x_{0,m+q_0'}$ be the positions of the children of 0 in $(T, 0)$. Define $k_0 = \lceil n y_0 \rceil$, so that k_0 is distributed uniformly in $\{1, \dots, n\}$, and for $i = 1, \dots, m$, define k_{0i} by

$$S_{k_{0i}-1} \le \frac{x_{0,i}}{x_0} S_{k_0-1} < S_{k_{0i}}$$

By Theorem 2.1 and the observation that $U_{k_0,1} = \frac{x_{0,1}}{x_0}, \dots, U_{k_0,m} = \frac{x_{0,m}}{x_0}$ are i.i.d. random variables chosen uniformly at random from $[0,1]$, we have that $k_{01}, \dots, k_{0m}$ have the required distribution.

Given $\epsilon > 0$ choose δ, C, K and N in such a way that the statements of Lemma 3.3 and Corollary 3.4 hold for $r = 1$, and let $\epsilon' = \min\{\epsilon, \delta/4\}$. By Lemma 3.1 there exists a constant $n_0 < \infty$ such that for $n \ge n_0$, we have that

$$|S_{k_0-1} - x_0| \le \epsilon' \quad \text{and} \quad |S_{k_{0i}-1} - x_{0i}| \le \epsilon' \quad \text{for all} \quad i = 1, \dots, m \tag{11}$$

with probability at least $1 - 2\epsilon$.

To understand the limiting distribution of the remaining neighbors, $k_{0,m+1}, \dots, k_{0,m+q_0}$, of k_0, we observe that conditioned on the random variables $\psi_1, \dots, \psi_n$, each vertex $k > k_0$ has m independent chances of being connected to k_0, corresponding to the m independent events $\{X_{k,i} = k_0\}$, $i = 1, \dots, m$, where we used the shorthand $X_{k,i}$ for the endpoint of the i^{th} edge sent out from k (it is related to the random variables X_t introduced in the proof of Theorem 2.1 via $X_{k,i} = X_{(k-2)m+i}$). Let

$$P_{k \to k_0} = \varphi_{k_0} \frac{1}{S_{k-1}} = \frac{S_{k_0}}{S_{k-1}} \psi_{k_0} \tag{12}$$

be the probability of the event $\{X_{k,i} = k_0\}$, and let $N_{y_0}(y) = \sum_{i=1}^m \sum_{k=k_0}^{\lceil ny \rceil} \mathbb{I}(X_{k,i} = k_0)$ where $\mathbb{I}(A)$ is the indicator function of the event A. We want to show that $N_{y_0}(\cdot)$ converges to a Poisson process on $[y_0, 1]$.

By Lemma 3.3, we have that $k_0 \ge n x_0 \ge n\delta$ with probability at least $1 - \epsilon$, which allows us to apply Lemmas 3.1 and 3.2 to show that for n large enough, with probability at least $1 - 2\epsilon$, we have

$$\hat P_{k \to k_0}(1 - \epsilon) \le P_{k \to k_0} \le (1 + \epsilon) \hat P_{k \to k_0} \quad \text{where} \quad \hat P_{k \to k_0} = \frac{1}{nm} \frac{\chi_{k_0}}{2(1+u)} \frac{n}{k_0} \Big(\frac{k_0}{k}\Big)^\chi.$$

For $y > y_0$, let $\hat N_{y_0}(y) = \sum_{i=1}^m \sum_{k=k_0}^{\lceil ny \rceil} \hat Y_{k \to k_0}^{(i)}$ where $\{\hat Y_{k \to k_0}^{(i)}\}$ are independent random variables such that $\hat Y_{k \to k_0}^{(i)} = 1$ with probability $\hat P_{k \to k_0}$ and $\hat Y_{k \to k_0}^{(i)} = 0$ with probability $1 - \hat P_{k \to k_0}$. It follows from standard results on convergence to Poisson processes (and the fact that γ_0 has the same distribution as χ_{k_0}) that $\hat{N}_{y_0}(\cdot)$ converges weakly to a Poisson process with density $\frac{\gamma_0}{2(u+1)y_0}\left(\frac{y_0}{y}\right)^{\chi}$ on $[y_0, 1]$. A change of variables from y to $x = y^{\chi}$ now leads to the Poisson process with density

$$\frac{\gamma_0}{2(1+u)x_0^{\psi}}\frac{x^{\psi-1}}{\chi}$$

on $[x_0, 1]$. Combined with a last application of Lemma 3.1 to bound the difference between $S_{k_{0i}-1}$ and $(k_{0i}/n)^{\chi}$, this proves that $x_{0,m+1}, \dots, x_{m+q_0'} \in [x_0, 1]$ and $k_{0,m+1}, \dots, k_{0,m+q_0}$ can be coupled in such a way that for n large enough, with probability at least $1 - 3\epsilon$, we have that $q_0 = q_0' \leq Q = N - m - 1$, $\chi_{k_0} = \gamma_0 \leq C$ and

$$|x_{0,i} - S_{k_{0,i}-1}| \leq \epsilon' \qquad \text{for} \qquad i = m+1, \dots, m+q_0. \tag{13}$$

Since $\epsilon > 0$ was arbitrary, this completes the proof of the first two statements of the lemma.

(iii) To prove the third statement, we use the bounds (11) and (13), and a final application of Lemma 3.1, to establish the existence of two constants $\delta' > 0$ and $n_0' < \infty$ such that for $n \geq n_0'$, with probability at least $1 - 4\epsilon$,

$$k_{\bar{a}} \geq \delta' n \quad \text{for all} \quad \bar{a} \in T_1(G_n) \tag{14}$$

and

$$|k_{\bar{a}} - k_{\bar{b}}| \geq \delta' n \quad \text{for all} \quad \bar{a}, \bar{b} \in T_1(G_n) \quad \text{with} \quad \bar{a} \neq \bar{b},$$

implying in particular that $k_0, k_{01}, \dots, k_{0m+q_0}$ are pairwise distinct.

(iv) To prove the last statement, let us assume that that $\gamma_0 \leq C$, and that $k_{01}, \dots, k_{0,m+q}$ are pairwise distinct, with $k_{0i} < k_0$ for $i \leq m$, $k_{0i} > k_0$ for $i > m$, $\min k_{0i} \geq n\delta'$ and $q \leq Q$. Let A be the event that we have chosen k_0 as the uniformly random vertex and that the neighbors of k_0 are the vertices $k_{01}, \dots, k_{0,m+q}$. Let χ^{A,γ_0} be the collection of random variables $\{\chi_k\}_{k \neq k_0}$ conditioned on $\chi_{k_0} = \gamma_0$ and A. We will want to show that χ^{A,γ_0} can be coupled to a collection of independent random variables $\{\hat{\chi}_k\}_{k \neq k_0}$ such that $\chi^{A,\gamma_0} = \{\hat{\chi}_k\}_{k \neq k_0}$ with probability at least $1 - \epsilon$, and

$$\hat{\chi}_k \sim \begin{cases} F_k' & \text{if } k \in \{k_{01}, \dots, k_{0m}\} \\ F_k & \text{otherwise} \end{cases} \tag{15}$$

Let $\rho(\cdot \mid A, \chi_{k_0})$ be the density of the random variable χ^{A,γ_0}, and let $\mathbf{P}(\cdot)$ be the joint distribution of G_n and the random variables $\chi_1, \dots, \chi_n$. By Bayes' theorem,

$$\rho(\cdot \mid A, \chi_{k_0} = \gamma_0) = \frac{\mathbf{P}(A \mid \cdot, \chi_0)}{\mathbf{P}(A \mid \chi_{k_0})}\rho_0(\cdot), \tag{16}$$

where ρ_0 is the original density of the random variables $\{\chi_k\}_{k \neq k_0}$ (we denote the corresponding probability distribution and expectations by P_0 and E_0, respectively).

We thus have to determine the probability of A conditioned on $\chi_1, \ldots, \chi_n$. With the help of Theorem 2.1, this probability is easily calculated, and is equal to

$$\begin{aligned}\mathbf{P}(A \mid \{\chi_k\}) &= \frac{1}{n}\prod_{i=1}^{m} P_{k_0\to k_{0i}} \prod_{j=1}^{q} m P_{k_{0,m+j}\to k_0}\Big(1 - P_{k_{0,m+j}\to k_0}\Big)^{m-1} \\ &\quad\times \prod_{k>k_0 : k\notin\{k_{0,m+1},\ldots,k_{0,m+q}\}} \Big(1 - P_{k\to k_0}\Big)^m \\ &= \frac{1}{n}\prod_{i=1}^{m} P_{k_0\to k_{0i}} \prod_{j=1}^{q} \frac{m P_{k_{0,m+j}\to k_0}}{1 - P_{k_{0,m+j}\to k_0}} \prod_{k>k_0}\Big(1 - P_{k\to k_0}\Big)^m\end{aligned}$$

where $P_{k\to k'}$ is defined in (12). By Lemma 3.1, this implies that given any $\epsilon' > 0$, we can find $n_0 < \infty$ such that for $n \geq n_0$, we have that with probability at least $1-\epsilon'$ with respect to P_0,

$$\begin{aligned}(1-\epsilon')\mathbf{P}(A \mid \{\chi_k\}) &\leq \frac{1}{n}\left(\prod_{i=1}^{m}\psi_{k_{0i}}\left(\frac{k_{0i}}{k_0}\right)^{\chi}\prod_{j=m+1}^{m+q} m\psi_{k_0}\left(\frac{k_0}{k_{0j}}\right)^{\chi}\right)\exp\Big(-m\psi_{k_0}\sum_{k>k_0}\Big(\frac{k_0}{k}\Big)^{\chi}\Big) \\ &\leq (1+\epsilon')\mathbf{P}(A \mid \{\chi_k\})\end{aligned}$$

To estimate $\mathbf{P}(A \mid \chi_{k_0}) = E_0[\mathbf{P}(A \mid \{\chi_k\})]$, we combined this bound with the deterministic upper bound

$$\begin{aligned}\mathbf{P}(A \mid \{\chi_k\}) &\leq \frac{1}{n}\prod_{i=1}^{m} P_{k_0\to k_{0i}}\prod_{j=1}^{q} m P_{k_{0,m+j}\to k_0} \leq \frac{1}{n}(m\psi_{k_0})^q\prod_{i=1}^{m}\psi_{k_{0i}} \\ &\leq \frac{C'}{n}\left(\prod_{i=1}^{m}\psi_{k_{0i}}\left(\frac{k_{0i}}{k_0}\right)^{\chi}\prod_{j=m+1}^{m+q} m\psi_{k_0}\left(\frac{k_0}{k_{0j}}\right)^{\chi}\right)\exp\Big(-m\psi_{k_0}\sum_{k>k_0}\Big(\frac{k_0}{k}\Big)^{\chi}\Big)\end{aligned}$$

where $C' = (\delta')^{-(m+Q)}\sup_{n\geq 1} e^{mnf_{\delta' n}(C)}$.

These bounds imply that given any $\epsilon' > 0$, we can find an $n_0 < \infty$ such that for $n \geq n_0$, with probability at least $1-\epsilon'/2$ with respect to P_0, we have

$$\sqrt{1-\epsilon'}\prod_{i=1}^{m}\frac{\psi_{k_{0i}}}{E_0(\psi_{k_{0i}})} \leq \frac{\mathbf{P}(A \mid \{\chi_k\})}{\mathbf{P}(A \mid \chi_{k_0})} \leq \sqrt{1+\epsilon'}\prod_{i=1}^{m}\frac{\psi_{k_{0i}}}{E_0(\psi_{k_{0i}})}.$$

With the help of Lemma 3.2, this shows that for n large enough, with probability at least $1-\epsilon'$, we have

$$(1-\epsilon')\prod_{i=1}^{m}\frac{\chi_{k_{0i}}}{E_0(\chi_{k_{0i}})} \leq \frac{\mathbf{P}(A \mid \{\chi_k\})}{\mathbf{P}(A \mid \chi_{k_0})} \leq (1+\epsilon')\prod_{i=1}^{m}\frac{\chi_{k_{0i}}}{E_0(\chi_{k_{0i}})}.$$

Recalling (16) and the definition of the random variables $\{\hat\chi_k\}_{k\neq k_0}$, we therefore have shown that with probability at least $1-\epsilon'$ with respect to P_0,

$$(1-\epsilon')\hat\rho(\cdot) \leq \rho(\cdot \mid A, \chi_{k_0} = \gamma_0) \leq (1+\epsilon')\hat\rho(\cdot) \tag{17}$$

where $\hat{\rho}$ is the density of the random variables $\{\hat{\chi}_k\}_{k\neq k_0}$. (We denote the corresponding product measure by $\hat{P}$).

To continue, we need to transform statement which happen with high probability with respect to P_0 into statement which happen with high probability with respect to $\hat{P}$. To this end, we consider the general case of two probability measures μ and ν such that ν is absolutely continuous with respect to μ, $\nu = f\mu$ for some non-negative function $f \in L_2(\mu)$. Let Ω_0 be an event which happens with probability $1-\epsilon'$ with respect to μ. Then

$$\nu(\Omega_0^c) = \int f 1_{\Omega_0^c} \leq \sqrt{E_\mu(f^2)\mu(\Omega_0^c)} = \sqrt{\epsilon' E_\mu(f^2)}, \tag{18}$$

implying that Ω_0 happens with probability at least $1-\sqrt{\epsilon' E_\mu(f^2)}$ with respect to ν.

Applying this bound to the probability measures P_0 and $\hat{P}$, we see that the bound (17) holds with probability at least $1-\sqrt{2\epsilon'}$ with respect to $\hat{P}$, provided n (and hence $k_{01}, \dots, k_{0m}$) is large enough. Using this fact, one than easily shows that

$$\|\hat{\rho} - \rho(\cdot \mid A, \chi_{k_0} = \gamma_0)\|_1 \leq 2\epsilon' + 2\sqrt{2\epsilon'}.$$

Choosing ϵ' sufficiently small ($\epsilon' = \epsilon^2/32$ is small enough), we see that the right hand side can be bounded by ϵ, which proves that χ^{A,γ_0} and $\{\hat{\chi}_k\}_{k\neq k_0}$ can be coupled in such a way that they are equal with probability at least $1-\epsilon$, as required. $\square$

3.5 Proof of Convergence for the sequential model

In this section we show that the sequential model converges to the Pólya-point graph. Indeed, we prove slightly more, namely the following proposition.

Proposition 3.6. *Given $\epsilon > 0$ and $r < \infty$, there are constants $C, N < \infty$, $\delta > 0$ and $n_0 < \infty$ such that for $n \geq n_0$, the rooted sequential attachment graph (G_n, k_0) and the Pólya-point process $\{x_{\bar{a}}\}$ can be coupled in such a way that with probability at least $1-\epsilon$, the following holds*

1. *$T_r(G_n) = T_r$ and $|T_r(G_n)| \leq N$;*
2. *$|x_{\bar{a}} - S_{k_{\bar{a}}-1}| \leq \epsilon$ for all $\bar{a} \in T_r$;*
3. *$\mathbf{k}^{(r)}(G_n)$ is injective, and $k_{\bar{a}} \geq \delta n$ for all $\bar{a} \in T_r$;*
4. *$\gamma_{\bar{a}} = \chi_{k_{\bar{a}}} \leq C$ for all $\bar{a} \in T_r$.*

Proof: For $r = 1$, this follows from Lemma 3.5 and Lemma 3.3.

Assume by induction that the lemma holds for $r < \infty$, and fix T_r, $\mathbf{k}^{(r)}(G_n)$, $\{x_{\bar{a}}\}_{\bar{a}\in T_r}$, $\{\gamma_{\bar{a}}\}_{\bar{a}\in T_r}$ and $\{\chi_{k_{\bar{a}}}\}_{\bar{a}\in T_r}$ in such a way that 1–4 hold (an event which has probability at least $1-\epsilon$ by our inductive assumption).

Consider a vertex $\bar{a} \in \partial T_r = T_r \setminus T_{r-1}$. We want to explore the neighborhood of $k_{\bar{a}}$ in G_n. To this end, we note that for all $\bar{b} \in T_{r-1}$, the neighborhood of $k_{\bar{b}}$ is already determined by our conditioning on $\mathbf{k}^{(r)}(G_n)$, implying in particular that none of the edges sent out from $k_{\bar{a}}$ can hit a vertex $k \in K_{r-1}$, unless, of course, $\bar{a}$ is of type L and k happens to be the parent of $k_{\bar{a}}$ — in which case the edge between k and $k_{\bar{a}}$ is already present. To determine the children of type E of the vertex $k_{\bar{a}}$, we therefore have to condition on not hitting the set K_{r-1}. But apart from this, the process of determining the children of $k_{\bar{a}}$ is exactly the same as that of determining the children of the root k_0. Since $|K_r| \leq N$, $k \geq \delta n$ for all $k \in K_r$, and $\chi_k \leq C$ for all $k \in K_r$, we have that $\sum_{k \in K_r} \varphi_k \leq C'/n$ for some $C' < \infty$, implying that conditioning on $k \notin K_{r-1} \subset K_r$ has only a negligible influence on the distribution of the children of $k_{\bar{a}}$. We may therefore proceed as in the proof of Lemma 3.5 to obtain a coupling between a sequence of i.i.d. random variables $x_{\bar{a},i}$ distributed uniformly in $[0, x_{\bar{a}}]$ and the children $k_{\bar{a},i}$ of $k_{\bar{a}}$ that are of type E. As before, we obtain that for n large enough, with probability at least $1-\epsilon$, we have $|S_{k_{\bar{a},i}-1} - x_{\bar{a},i}| \leq \epsilon$.

Repeating this process for all $k_{\bar{a}} \in \partial K_r = K_r \setminus K_{r-1}$, we obtain a set of vertices E_{r+1} consisting of all children of type E with parents in ∂K_r. It is easy to see that with probability tending to one as $n \to \infty$, the set E_{r+1} has no intersection with K_r, so we will assume this for the rest of this proof.

Next we continue with the vertices of type L. Assume that we have already determined all children of type L for a certain subset $U_r \subset \partial K_r$, and denote the set children obtained so far by L_{r+1}. We decompose this set as $L_{r+1} = \bigcup_{i=1}^{m} L_{r+1}^{(i)}$, where $L_{r+1}^{(i)} = \{k \in L_{r+1} \colon X_{i,k} \in U_r\}$. Consider a vertex $\bar{a} \in \partial K_r \setminus U_r$. Conditioning on the graph explored so far is again not difficult, and now amounts to two conditions:

1. $X_{k,i} \neq k_{\bar{a}}$ if $k \in K_r \cup L_{r+1}^{(i)}$, since all the edges sent out from this set have already been determined.

2. For $k \notin K_r \cup L_{r+1}^{(i)}$, the probability that $k_{\bar{a}}$ receives the i^{th} edge from k is different from the probability given in (12), since the random variables $X_{k,i}$ has been probed before: we know that $X_{k,i} \notin K_{r-1}$ since otherwise k had sent out an edge to a vertex in K_{r-1}, which means that k would have been a child of type L in K_r. We also know that $X_{k,i} \notin U_r$, since otherwise $k \in L_{r+1}^{(i)}$. Instead of (12), we therefore have to use the modified probability
$$P_{k \to k_{\bar{a}}} = \varphi_{k_{\bar{a}}} \frac{1}{\widetilde{S}_{k-1}}$$
where
$$\widetilde{S}_{k-1} = \sum_{\substack{k' > k_{\bar{a}}: \\ k' \notin K_{r-1} \cup U_r}} \varphi_{k'}$$

Since $\tilde{S}_{k-1} \leq S_{k-1} \leq \tilde{S}_{k-1} + C'/n$ by our inductive assumption, we can again refer to

Lemma 3.1 to approximate $P_{k\to k_{\bar{a}}}$ by

$$\hat{P}_{k\to k_{\bar{a}}} = \frac{1}{nm}\frac{\chi k_{\bar{a}}}{2(1+u)}\frac{n}{k_{\bar{a}}}\Big(\frac{k_{\bar{a}}}{k}\Big)^{\chi}.$$

From here on the proof of our inductive claim is completely analog to the proof of Lemma 3.5. We leave it to the reader to fill in the (straightforward but slightly tedious) details. ☐

3.6 Estimates for the Pólya Urn Representation

In this section we complete the work started in Section 3.1 by proving Lemmas 3.1 and 3.2.

Proof:[Proof of Lemma 3.1] Fix ϵ, and recall that

$$\chi = \frac{1+2u}{2+2u}.$$

For every k,

$$\begin{aligned} E(\psi_k) &= \frac{m+2mu}{(2k-2)m+2kmu} = \frac{\chi}{k} + ck^{-2} \\ &= 1 - \left(\frac{k-1}{k}\right)^{\chi} + ck^{-2}, \end{aligned} \tag{19}$$

where $0 < c < 1$. And

$$\mathrm{var}(\psi_k) \le E\left[\psi_k^2\right] = \frac{m+2mu+1}{(2k-2)m+2kmu+1}\,\frac{m+2mu}{(2k-2)m+2kmu} \le \frac{1}{k^2}. \tag{20}$$

Take

$$\psi_k' = \left\{ \begin{array}{ll} \psi_k & \text{if } \psi_k < 1/2 \\ 0 & \text{otherwise} \end{array} \right. .$$

Then

$$E(\log(1-\psi_k')) = \chi \log\left(\frac{k-1}{k}\right) + O(k^{-2}),$$

$$\mathrm{var}(\log(1-\psi_k')) = O(k^{-2}),$$

and

$$\mathbf{P}(\psi_k \ne \psi_k') = \mathbf{P}(\psi_k \ge 1/2) \le 4E(\psi_k^2) = O(k^{-2}).$$

Note that

$$S_k = \prod_{j=k+1}^{n}(1-\psi_j) = \exp\left(\sum_{j=k+1}^{n}\log(1-\psi_j)\right)$$

and thus, with probability $1 - O(K^{-2})$,

$$S_k = \exp\left(\sum_{j=k+1}^{n} \log(1 - \psi'_j)\right)$$

for all $k \geq K$. Finally, by Doob's inequality, we have that

$$\mathbf{P}\left(\max_{K \leq k \leq n}\left|\sum_{j=k+1}^{n}\left(\log(1-\psi'_j) - E[\log(1-\psi'_j)]\right)\right| \geq \epsilon\right) = O\Big(\frac{1}{\epsilon^2 K^2}\Big).$$

Putting everything together, we get that there exists a constant $K(\epsilon)$ not depending on n such that with probability at least $1 - \epsilon$, we have that

$$\left(\frac{k}{n}\right)^{\chi} e^{-\epsilon} < S_k < \left(\frac{k}{n}\right)^{\chi} e^{\epsilon} \quad \text{for all} \quad K(\epsilon) \leq k \leq n. \tag{21}$$

For $k < K(\epsilon)$, we bound $S_k \leq S_K$ to conclude that with probability at least $1 - \epsilon$,

$$\left|S_k - \Big(\frac{k}{n}\Big)^{\chi}\right| = O\Big(\Big(\frac{K}{n}\Big)^{\chi}\Big).$$

The lemma now follows. $\Box$

Proof:[Proof of Lemma 3.2] (i) Let $a = m + 2mu$, so that $\chi_k \sim \Gamma(a, 1)$. Then

$$\mathbf{P}(\chi_k \geq \log^2 k) \leq E[e^{\frac{1}{2}\chi_k}]e^{-\frac{1}{2}\log^2 k} = 2^a k^{-\frac{1}{2}\log k}.$$

Since the right hand side is sumable, this implies the first statement of the lemma.

(ii) Let $b_k = (2k-3)m + 2mu(k-1) - 1$, and let $\chi'_k = \chi_k / b_k$. Then f_k can be defined by

$$\mathbf{P}(\psi_k \leq f_k(x)) = \mathbf{P}(\chi'_k \leq x/b_k)$$

In order to prove the second statement of the lemma, it is clearly enough to prove that for all sufficiently large k, we have

$$(1-\epsilon)\frac{x}{b_k} \leq f_k(x) \leq \frac{x}{b_k} \quad \text{for} \quad x \leq \log^2 k,$$

which in turn is equivalent to showing that

$$\mathbf{P}\Big(\psi_k \leq (1-\epsilon)x\Big) \leq \mathbf{P}\Big(\chi'_k \leq x\Big) \leq \mathbf{P}\Big(\psi_k \leq x\Big) \quad \text{for} \quad x \leq \frac{\log^2 k}{b_k} \tag{22}$$

provided k is large enough.

We start by proving that

$$\Delta(x) := \mathbf{P}(\psi_k \leq x) - \mathbf{P}(\chi'_k \leq x) \geq 0.$$

To this end, we rewrite

$$\mathbf{P}(\psi_k \le x) = \frac{1}{Z_\beta} \int_0^x y^{a-1}(1-y)^b dy,$$

and

$$\mathbf{P}(\chi'_k \le \lambda) = \frac{1}{Z_\gamma} \int_0^\lambda y^{a-1} e^{-by} dy,$$

where $a = m + 2mu$, $b = b_k$, and $Z_\gamma = \int_0^\infty y^{a-1}e^{-by}dy$ and $Z_\beta = \int_0^1 y^{a-1}(1-y)^b dy$ are the appropriate normalization factors. For $x \le 1$, we express $\Delta(x)$ as

$$\Delta(x) = \frac{1}{Z_\gamma} \int_0^x dy y^{a-1} e^{-by} \left(e^\delta \exp\left(-b \sum_{k=2}^\infty \frac{y^k}{k} \right) - 1 \right),$$

where $e^\delta = Z_\gamma / Z_\beta$. Note that $\delta > 0$ by the fact that $(1-x) \le e^{-x}$. It is also easy to see that $\delta \to 0$ as $k \to \infty$; indeed, we have $\delta = O(b^{-1}) = O(k^{-1})$.

Consider the derivative

$$\frac{d\Delta(x)}{dx} = \frac{x^{a-1}e^{-bx}}{Z_\gamma} \left(e^\delta \exp\left(-b \sum_{k=2}^\infty \frac{x^k}{k} \right) - 1 \right),$$

and let x_0 be the unique root, i.e., let $x_0 \in (0,1)$ be the solution of the equation

$$\delta = b \sum_{k=2}^\infty \frac{x_0^k}{k}.$$

Then $\Delta(x)$ is monotone increasing for $0 < x < x_0$ and monotone decreasing for all $x > x_0$. Together with the observation that $\Delta(x) > 0$ for all sufficiently small x, and $\Delta(x) \to 0$ as $x \to \infty$, we conclude that $\Delta(x) \ge 0$ for $0 \le x < \infty$. This proves that $\mathbf{P}(\chi'_k \le x) \le \mathbf{P}(\psi_k \le x))$ for all $x \ge 0$.

To prove the lower bound in (22), we will prove that

$$\tilde{\Delta}(x) = \mathbf{P}\Big(\chi'_k \le x\Big) - \mathbf{P}\Big(\psi_k \le (1-\epsilon)x\Big) \ge 0 \quad \text{if} \quad x \le \frac{\epsilon}{4} \le \frac{1}{8}.$$

We decompose the range of x into two regions, depending on whether $x \ge \frac{4a}{b\epsilon}$ or $x \le \frac{4a}{\epsilon b}$.

In the first region, we express $\tilde{\Delta}(x)$ as

$$\begin{aligned} \tilde{\Delta}(x) &= \mathbf{P}\Big(\psi_k \ge (1-\epsilon)x\Big) - \mathbf{P}\Big(\chi'_k \ge x\Big) \\ &= \frac{e^\delta}{Z_\gamma} \int_{x(1-\epsilon)}^1 dy y^{a-1}(1-y)^b - \frac{1}{Z_\gamma} \int_x^\infty dy y^{a-1} e^{-by}. \end{aligned}$$

We then bound

$$\begin{aligned} \int_x^\infty dy (2y)^{a-1} e^{-by} &\le \int_x^{2x} dy y^{a-1} e^{-by} \int_{2x}^\infty dy y^{a-1} e^{-by} \\ &\le \int_x^{2x} dy y^{a-1} e^{-by} + 2^{a-1} e^{-bx} \int_x^\infty dy y^{a-1} e^{-by} \end{aligned}$$

proving that

$$\int_x^\infty dy(2y)^{a-1}e^{-by} \le (1-2^{a-1}e^{-bx})^{-1}\int_x^{2x} dyy^{a-1}e^{-by} \le 2\int_x^{2x} dyy^{a-1}e^{-by}, \tag{23}$$

where we have used $bx \ge a\log 2$ in the last step.
On the other hand, using that $(1-y)^b \ge e^{-by(1+x)}$ if $y \le 2x \le 1/2$, we have that

$$\begin{aligned} e^\delta \int_{x(1-\epsilon)}^1 dyy^{a-1}(1-y)^b &\ge \int_{x(1-\epsilon)}^{2x(1-\epsilon)} dyy^{a-1}e^{-by(1+x)} \\ &= \int_x^{2x} dyy^{a-1}(1-\epsilon)^a e^{-by(1+x)(1-\epsilon)} \\ &\ge (1-\epsilon)^a e^{-2bx^2}e^{\epsilon bx}\int_x^{2x} dyy^{a-1}e^{-by} \\ &\ge 2\int_x^{2x} dyy^{a-1}e^{-by}. \end{aligned}$$

Combined with (23), this proves that $\tilde\Delta(x) \ge 0$ if $\epsilon bx \ge 4a$.
For $\epsilon bx \le 4a$, we bound

$$\begin{aligned} \tilde\Delta(x) &= \frac{1}{Z_\gamma}\left(\int_0^x dyy^{a-1}e^{-by} - e^\delta \int_0^{x(1-\epsilon)} dyy^{a-1}(1-y)^b\right) \\ &\ge \frac{1}{Z_\gamma}\left(\int_0^x dyy^{a-1}e^{-by} - e^\delta \int_0^{x(1-\epsilon)} dyy^{a-1}e^{-by}\right) \\ &= \frac{1}{Z_\gamma}\left(\int_{(1-\epsilon)x}^x dyy^{a-1}e^{-by} - (e^\delta - 1)\int_0^{x(1-\epsilon)} dyy^{a-1}e^{-by}\right) \\ &\ge \frac{1}{Z_\gamma}\left(\epsilon x[(1-\epsilon)x]^{a-1}e^{-bx} - (e^\delta-1)x^a\right) \\ &\ge \frac{x^a}{Z_\gamma}\left(\epsilon 2^{1-a}e^{-4a/\epsilon} - (e^\delta - 1)\right). \end{aligned}$$

Since $\delta \to 0$ as $b \to \infty$, we see that the right hand side becomes positive if $k \ge K$ for some $K < \infty$ that depends on a and ϵ (it grows exponentially in a/ϵ). □

4 Approximating Coupling for the Independent and the Conditional Models.

We only give full details for the independent model. The approximating coupling for the conditional model is very similar, and the proof that it works is identical.

We construct a coupling between the sequential model and the independent model such that almost surely a subset of density 1 of the vertices has the exact same connections in both graphs. From that will imply that both models have the same weak distributional limit.

Throughout our analysis, we condition the events on the the σ-algebra generated by $\{\psi_h\}_{h=1}^{\infty}$. As before, we use $\varphi_k^{(n)}$ for

$$\psi_k \prod_{i=k+1}^{n} (1-\psi_i).$$

We start with a somewhat technical claim about the evolution of degrees in the sequential model. Let $d_n(k)$ be the degree of vertex k when the graph contains n vertices. By construction,

$$d_n(k) = m + \sum_{i=(k-1)m+1}^{nm} \mathcal{U}_i$$

where the variables $\{\mathcal{U}_i\}$ are defined as follows: Let $\{\hat{U}_i\}_{i=1}^{\infty}$ be i.i.d. $U[0,1]$ variables, independent of the φ_i-s. Then $\mathcal{U}_i = \mathbf{1}_{\hat{U}_i < \varphi_k^{([i/m])}}$. Note that conditioned on $\{\varphi_k^{([i/m])}\}_{i\geq k}$, $\{\mathcal{U}_i\}$-s are independent, each being Bernoulli $\varphi_k^{([i/m])}$.

Claim 4.1. *For the sequential preferential attachment model, for every j and k such that $j > k$, let $d_n(k)$ be the degree of vertex k when the graph contains n vertices. Then almost surely for every k,*

$$U_k := \sup_n \frac{d_n(k)}{n^{1-\chi}} < \infty. \tag{24}$$

In addition, as k goes to infinity, U_k converges to 0 in probability.

Note that $\frac{1}{2} \leq \chi < 1$. The proof of this Claim is presented in Section 4.1.

Now, We can construct the coupling: Let $V = 1, 2, \ldots$ be the vertices of the preferential attachment graph. For $1 \neq n \in V$ and $i = 1, \ldots, m$ let e_n^i and f_n^i be the i-th vertex that n is connected to in, respectively, the sequential and the independent models.

By construction, $e_2^i = f_2^i = 1$ for all i. Once we know e_l^i and f_l^i for every $l < n$ and $i = 1, \ldots, m$, we determine e_n^i and f_n^i for $1 \leq i \leq m$ as follows: Let D_1 be the distribution of $\{e_n^i\}_{i=1}^m$ based on the sequential rule and conditioned on $\{e_l^i\}_{l<n\ ;1\leq i\leq m}$, and let D_2 be the distribution of $\{f_n^i\}_{i=1}^m$ based on the independent rule and conditioned on $\{e_l^i\}_{l<n\ ;1\leq i\leq m}$. Let D be a coupling of D_1 and D_2 that minimizes the total variation distance. Then we choose $\{e_n^i\}_{i=1}^m$ and $\{f_n^i\}_{i=1}^m$ according to D.

Let $k \in V$. We now estimate the probability that there exists $n > k$ and i such that $e_n^i = k \neq e_l^i$ or $e_n^i \neq k = f_n^i$. For given n, we call the event that such i exists A_n.

We are interested in

$$E_k = \sum_{n=k+1}^{\infty} \mathbf{P}\left(A_n \,\middle|\, \bigcap_{i=k+1}^{n-1} A_i^c \right).$$

Claim 4.2.

$$\mathbf{P}\left(A_n \,\middle|\, \bigcap_{h=k+1}^{n-1} A_h^c, d_{n-1}^k\right) = O\left(\max\left\{\frac{d_{n-1}^k}{n^2}, \frac{(d_{n-1}^k)^3}{n^3}\right\}\right) \tag{25}$$

Note that the proof is very simple for $m = 2$ (as the reader can easily verify). The proof below, for general m, contains no further ideas. However, the formulas become more complicated, as the reader can see below.

Proof: Let us denote d_{n-1}^k with d. The probability in (25) is the difference between the probabilities of the new edge connecting to k according to the sequential and the independent model, which is

$$O\left(\max_{r=0,\ldots,m}\left|\left(\frac{d}{n}\right)^r\left(\frac{n-d}{n}\right)^h - \prod_{l=0}^{r-1}\frac{d+l}{n+l}\prod_{u=0}^{h-1}\frac{n-d+u}{n+r+u}\right|\right) \tag{26}$$

where $h = m - r$. Equation (26) follows from the fact that the probability of having r connections to k and h connections to other vertices is $\binom{m}{r}\left(\frac{d}{n}\right)^r\left(\frac{n-d}{n}\right)^h$ in the independent model and $\binom{m}{r}\prod_{l=0}^{r-1}\frac{d+l}{n+l}\prod_{u=0}^{h-1}\frac{n-d+u}{n+r+u}$ in the sequential model. Therefore the probability in (25) is of the order of magnitude of the largest of these differences.

In order to bound equation (26), note that the second product in the absolute value is always bigger than or equal to the first term. Also it is at most

$$\left(\frac{d+m}{n+m}\right)^r\left(\frac{n-d+m}{n+m}\right)^h - \left(\frac{d}{n}\right)^r\left(\frac{n-d}{n}\right)^h$$

For $r \geq 3$, the second term at most $f(m)(\frac{d}{n})^3$ which is of $O(\frac{d^3}{n^3})$. Note that m is constant. For $r = 1$, the above equation becomes

$$\begin{aligned}\frac{(n-d+m)^m n^m - (n-d)^m(n+m)^m}{(n+m)^m n^m} &= \frac{f(m)\left[n^m(n-d)^{m-1} - n^{m-1}(n-d)^m\right] + O(n^{m-2})}{(n+m)^m n^m}\\ &= O(\frac{d}{n^2})\end{aligned}$$

The calculation for $r = 1, 2$ is very similar and it gives the same bound of $O(\frac{d}{n^2})$. □

From Claims 4.1 and 4.2, we get that $E_k \leq u_k k^{-\frac{1}{2}}$. Our last step is to show that for every r, almost surely a fraction 1 of the vertices have the same neighborhood of radius r in the sequential and the independent models, and therefore both graphs have the same weak distributional limit.

Fix r, and fix some $\epsilon > 0$. By Lemma 3.3, there exist $\alpha > 0$ and $M < \infty$ such that in the sequential graph, all but fraction ϵ of the vertices k, (1) The ball of radius r around k, denoted by $B_r(k)$, contains no more than M vertices. (2) The oldest vertex (the vertex with the smallest index) in this ball is no older than αk. Let A be the set of vertices satisfying (1) and (2). Call $k \in A$ bad if there is a difference in the neighborhood of radius r of vertex k in the two models. For every k, let

$$H_k = \sum_{j: j \in B_r(k)} E_j.$$

H_k bounds from above the probability that that a vertex in $k \in A$ is bad. On the other hand, for all vertices $k \in A$, $H_k \le M U_k (\alpha k)^{-\frac{1}{2}}$ and in particular for k large enough H_k is bounded by an inverse polynomial.

Fix some $\gamma > 1$. Consider the intervals $I_n = [\gamma^n, \gamma^{n+1} - 1]$. The expected fraction of bad k's in I_n is bounded by $\max_{j \in I_n \cap A} E(H_j)$. This quantity decays exponentially with n. Therefore, for a given δ, by Markov's inequality, the probability that the fraction of bad k's in I_n exceeds δ, decays exponentially with n. Therefore, by Borel-Cantelli lemma, there exists an n_0 such that for any $n \ge n_0$, the fraction of bad vertices in each I_n is less than δ.

The above argument implies that for any integer r, and $\delta, \epsilon > 0$, there is an n_0 such that the fraction of vertices $n \ge n_0$ whose neighborhood of radius r is different in two models is bounded by $\epsilon + \delta$. That implies that almost surely, a fraction 1 of the vertices have the same neighbors in balls of radius r in the sequential and the independent models. Therefore both processes have the same weak distributional limit.

4.1 Proof of Claim 4.1

We use the Pólya urn representation the prove the claim.

As before, we use $\varphi_k^{(n)}$ for

$$\psi_k \prod_{i=k+1}^{n} (1 - \psi_i).$$

By construction,

$$d_n(k) = m + \sum_{i=(k-1)m+1}^{nm} \mathcal{U}_i$$

where the variables $\{\mathcal{U}_i\}$ are defined as follows: Let $\{\hat{U}_i\}_{i=1}^{\infty}$ be i.i.d. $U[0,1]$ variables, independent of the φ_i-s. Then $\mathcal{U}_i = \mathbf{1}_{\hat{U}_i < \varphi_k^{([i/m])}}$. Note that conditioned on $\{\varphi_k^{([i/m])}\}_{i \ge k}$, the $\{\mathcal{U}_i\}$-s are independent, each being Bernoulli $\varphi_k^{([i/m])}$.

First, we want to understand the expectation of $d_n(k))$. Let $\mathcal{F}$ be the σ-algebra generated by $\{\psi_h\}_{h=1}^{\infty}$. Let $\hat{d}_n(k) = E(d_n(k)|\mathcal{F})$.

$$U_k = \sup_{n \ge k} \frac{d_n(k)}{n^{1-\chi}} = \left(\sup_{n \ge k} \frac{d_n(k)}{\hat{d}_n(k)} \right) \left(\frac{\hat{d}_n(k)}{n^{1-\chi}} \right) \tag{27}$$

We will show that the first term of the product is less than infinity with probability 1 and uniformly in k, and the σ-algebra $\mathcal{F}$. Furthermore, the second term is less than infinity with probability 1 and goes to zero as k goes to infinity in probability.

The bound on the first term follows immediately from the following lemma which is a simple application of law of large numbers. We include the proof for the sake of completeness.

Lemma 4.3. *Suppose $X_1, X_2, \cdots$ are independent Bernoulli random variables and $E(X_i) = a_i$. Then it is possible to find a coupling with a rate one Poisson random variable M such that with probability 1, for any l,*

$$S_l = \sum_{i=1}^{l} X_i \le 2(1+Z)\sum_{i=1}^{l} a_i$$

where $Z = sup_t \frac{M_t}{t} < \infty$.

Proof: First, note that

$$\sum_i X_i \le 2\sum_{i:a_i \ge 1/2} a_i + \sum_{i:a_i<1/2} X_i.$$

We will couple the second term with a rate 1 Poisson process on $\mathbb{R}_+$. Let $N_l = \sum_{i:a_i<1/2, i\le l} X_i$. For any i such that $a_i < 1/2$, let $b_i = -\log(1-a_i)$. Let M be a rate one Poisson process and $e_l = \sum_{i:a_i<1/2, i\le l} b_i$. It is possible to define a coupling such that for every l, M_{e_l} dominates N_l.

$$S_l \le 2\sum_{i:a_i\ge 1/2, i\le l} a_i + \sup_{t\ge 0}\frac{M_t}{t} e_l \le 2(1+\sup_{t\ge 0}\frac{M_t}{t})\sum_{i=1}^{l} a_i$$

The last inequality is because for any i such that $a_i > 1/2$ we have $a_i \le b_i < 2a_i$ and therefore $e_l < 2\sum_{i:a_i<1/2, i\le l} a_i$.

Furthermore, by the law of large numbers,

$$\lim_{t\to\infty}\frac{M_t}{t} = 1.$$

almost surely, and this is sufficient to show that $Z = \sup_{t\ge 0}\frac{M_t}{t} < \infty$ almost surely.

□

For bounding the second term in equation (27), note that

$$\hat{d_n}(k) = E(d_n(k)|\mathcal{F}) = m + \sum_{i=k+1}^{n} m\varphi_k^{(i)} = m + \psi_k \sum_{i=k+1}^{n} m\frac{\varphi_k^{(i)}}{\psi_k}. \tag{28}$$

Also,

$$\frac{\varphi_k^{(i)}}{\psi_k} = \prod_{h=k+1}^{i}(1-\psi_h) = \exp\left(\sum_{h=k+1}^{i}\log(1-\psi_h)\right) \le \exp\left(-\sum_{h=k+1}^{i}\psi_h\right).$$

By equation (19), $\frac{\chi}{h} \le E(\psi_h) \le \frac{\chi}{h} + \frac{1}{h^2}$. Therefore,

$$\frac{\varphi_k^{(i)}}{\psi_k} \le \exp\left(- \sum_{h=k+1}^{n} E(\psi_h) + R_k \right) \le e^{R_k} \left(\frac{k}{i}\right)^{\chi} \tag{29}$$

where $R_k = \sup_{i \ge k} \sum_{h=k+1}^{i} (\psi_h - E(\psi_h))$. Equations (28) and (29) imply

$$\begin{aligned} \hat{d}(k,n) &\le m + m\psi_k e^{R_k} \sum_{i=k+1}^{n} \left(\frac{k}{i}\right)^{\chi} \\ &\le C\psi_k e^{R_k} k^{\chi} n^{1-\chi}. \end{aligned}$$

For some constant C. For bounding R_k, note that by equation (20), we have $\mathrm{var}(\psi_h) < 1/h^2$. Therefore, Doob's inequality implies that for every $\alpha > 0$

$$\mathbf{P}(R_k \ge \alpha) \le \frac{1}{\alpha^2} \sum_{h=k+1}^{\infty} E((\psi_h - E(\psi_h))^2) \le \frac{1}{\alpha^2 k}. \tag{30}$$

The claim now follows if we notice that e^{R_k} is tight and that $\psi_k k^{\chi}$ converges in probability to zero.

5 Applications

5.1 Finite Ball Distribution

Here we show how to use the ideas above in order to determine the distribution of the graph structure of a ball of radius k around a uniform point of the preferential attachment graph. By Theorem 2.2 this is the same as the distribution of the graph structure of a ball of radius k around the root of the Pólya-point graph. While the calculation below may get quite complicated for large graphs, this gives a complete calculation for every finite graph.

In the sections below we show how to explicitly calculate certain quantities based on the scheme provided here.

Let $G = (V = \{v_0, v_1, v_2, \dots, v_m\}, E)$ be a tree rooted at v_0 such that all path from v_0 are increasing, let $n_0, \dots, n_m$ be non-negative integers and let $L : E \to \{0, 1\}$. We want to find the probability that there exits an injective homomorphism f from G to $\widetilde{P}$ such that:

1. $f(v_0) = \rho$.
2. For $v_j, v_k \in V$ with $e = (v_j, v_k) \in E$ and $j < k$, if $L(e) = 0$ then v_k is a parent of v_j and if $L(e) = 1$ then v_k is a child of v_j.

3. The number of children of $f(v_l)$ in $\tilde{P}$ is n_l plus the number of children of v_l in G, where a child of v_l in G is a vertex v_j s.t. $(v_l, v_j) \in E$ and
   (a) $l > j$ and $L(v_l, v_j) = 0$ or
   (b) $j > l$ and $L(v_l, v_j) = 1$.

We want to calculate $P(G)$, the probability that such homomorphism exists. Obviously, $P(G) = 0$ unless G is a tree.

For every $x_0, \ldots, x_m \in [0,1]$ and $\gamma_0, \ldots, \gamma_m \in (0, \infty)$, the probability density that a homomorphism f exists with $f(v_i) = x_i$ conditioned on $\gamma(x_i) = \gamma_i$ is

$$U(G, x_0, \ldots, x_m, \gamma_0, \ldots, \gamma_m)$$

$$= \frac{x_0^{\psi}}{\psi+1} \prod_{\substack{(v_j, v_k) \in E \\ j < k \\ L(v_j, v_k) = 0}} \mathbf{1}_{\{x_k < x_j\}} x_j^{-1} \cdot \prod_{\substack{(v_j, v_k) \in E \\ j < k \\ L(v_j, v_k) = 1}} \mathbf{1}_{\{x_k > x_j\}} \frac{\gamma_j x_k^{\psi-1}}{x_j^{\psi}}$$

$$\cdot \prod_{v_k \in V} \frac{\exp(-H_k) H_k^{n_k}}{n_k!}$$

with

$$H_k = \frac{\alpha_k 1 - x_k^{\varphi}}{x_k^{\varphi}}.$$

In addition, for $i = 1, \ldots, m$ let $J(i)$ be the (unique) $j < i$ such that $(v_j, v_i) \in E$ and let $L(i) = L((v_j, v_i))$. Let

$$\alpha(i) = \begin{cases} m + 2mu + 1 & \text{if } L(i) = 0 \\ m + 2mu & \text{if } L(i) = 1 \end{cases}.$$

We let $\alpha_0 = m + 2m$.

Then

$$P(G) = \frac{1}{Z \cdot CL(G)} \int_{[0,1]^{m+1} \times (0,\infty)^{m+1}} U(G, x_0, \ldots, x_m, \gamma_0, \ldots, \gamma_m) \prod_{k=0}^{m} e^{\gamma_k} \gamma_k^{\alpha_k - 1} dx_0 \cdots dx_m d\gamma_0 \cdots d\gamma_m$$

with $CL(G)$ being the number of isomorphisms of G and

$$Z = \prod_{i=0}^{m} \Gamma(\alpha_i).$$

5.2 Degree Distribution

The limiting degree distribution of the preferential attachment graph is exactly the degree distribution of the root of the Pólya-point graph. We will now estimate the probability that the root is of degree $k+m$.

Conditioned on the value of $\gamma = \gamma(x_0)$ and on x_0, the degree of the root, denoted by D, is m plus a Poisson variable with parameter $\frac{\gamma}{x_0^\psi} \int_{x_0}^1 x^{\psi-1} dx = \frac{\gamma(1-x_0^\psi)}{\psi x_0^\psi}$

Therefore, conditioned on γ,

$$\mathbf{P}(D = m+k|\gamma) = \frac{\psi+1}{k!} \int_0^1 e^{-\gamma x} (\gamma x)^k x^\psi dx$$

References

[AB02] R. Albert and A. Barabasi. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74:47–98, 2002.

[AJB99] R. Albert, H. Jeong, and A. Barabási. Diameter of the world wide web. *Nature*, 401:130–131, 1999.

[Ald98] David Aldous. Tree-valued Markov chains and Poisson-Galton-Watson distributions. In *Microsurveys in discrete probability (Princeton, NJ, 1997)*, volume 41 of *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, pages 1–20. Amer. Math. Soc., Providence, RI, 1998.

[BA99] A. Barabasi and R. Albert. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.

[BBCS05] N. Berger, C. Borgs, J. Chayes, and A. Saberi. On the spread of viruses on the internet. *Proceedings of the 16th ACM-SIAM Symposium on Discrete Algorithm (SODA)*, pages 301–310, 2005.

[BCL$^+$] C. Borgs, J. Chayes, L. Lovász, V.T. Sós, and J. Kahn. Convergence of sparse graph sequences. In preparation.

[BCL$^+$06a] C. Borgs, J. Chayes, L. Lovász, V.T. Sós, and K. Vesztergombi. Convergent sequences of dense graphs I. subgraph frequencies, metric properties and testing. *Advances in Mathematics, to appear*, 2006. Preprint available at http://research.microsoft.com/ borgs/Papers/ConvMetric.pdf.

[BCL$^+$06b] Christian Borgs, Jennifer Chayes, László Lovász, Vera T. Sós, and Katalin Vesztergombi. Counting graph homomorphisms. In *Topics in discrete mathematics*, volume 26 of *Algorithms Combin.*, pages 315–371. Springer, Berlin, 2006.

[BCL$^+$07] C. Borgs, J. Chayes, L. Lovász, V.T. Sós, and K. Vesztergombi. Convergent graph sequences II. H-colorings, statistical physics and quotients. 2007. Preprint available at http://research.microsoft.com/ borgs/Papers/ConRight.pdf.

[BR02] B. Bollobás and O. Riordan. Mathematical results on scale-free random graphs. In S. Bornholdt and H.G. Schuster, editors, *Handbook of Graphs and Network: From the Genome to the Internet*, pages 1–34. Wiley-VCH, 2002.

[BR04] B. Bollobas and O. Riordan. The diameter of a scale-free random graph. *Combinatorica*, 24:5–34, 2004.

[BR07] Bela Bolloás and Oliver Riordan. Sparse graphs: metrics and random models. 2007.

[BS01] I. Benjamini and O. Schramm. Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.*, 6 (paper 23):1–13, 2001.

[Dur96] R. Durrett. *Probability: Theory and Examples.* Duxbury Press, 1996.

[Ele07] Gábor Elek. On limits of finite graphs. *Combinatorica*, 27(4):503–507, 2007.

[EP23] F. Eggenberger and G. Polya. Uber die statistik verketteter vorgange. *Zeitschrift fur Angewandte Mathematik und Mechanik*, 1923.

[FFF99] Michalis Faloutsos, Petros Faloutsos, and Christos Faloutsos. On power-law relationships of the internet topology. In *SIGCOMM '99: Proceedings of the conference on Applications, technologies, architectures, and protocols for computer communication*, pages 251–262, New York, NY, USA, 1999. ACM Press.

[LS06] László Lovász and Balázs Szegedy. Limits of dense graph sequences. *J. Combin. Theory Ser. B*, 96:933–957, 2006.

[Lyo05] Russell Lyons. Asymptotic enumeration of spanning trees. *Combin. Probab. Comput.*, 14(4):491–522, 2005.