Degree distributions in preferential attachment graphs
Part I: Multivariate Approximations

Nathan Ross (University of Melbourne)

Joint work with Erol Peköz (Boston U) and Adrian Röllin (NU Singapore)

Preferential attachment random graphs:

- Popularized by Barabasi and Albert in 1999 to explain so-called "power law" behavior of the degree distribution in some real world networks, for example
  - vertices are html pages on the internet with edges the hyperlinks between webpages,
  - vertices are movie actors with an edge between two actors if they have appeared in a movie together.
- General idea: graph evolves sequentially by adding vertices one at a time. Each new vertex connects to some number of existing vertices in a random way so that connections to vertices with high degree are favored.

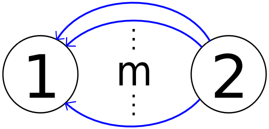
Outline

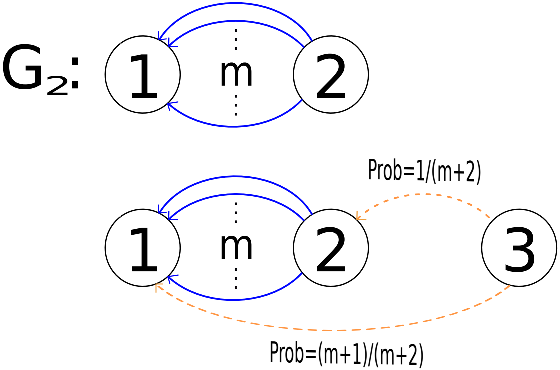
- Precisely define the model we study.
- State results.
- Main idea of the proof.

- Vertex $n+1$ sequentially attaches m outgoing edges to vertices $\{1, \dots, n\}$.
- The chance that an outgoing edge attaches to vertex j is proportional to

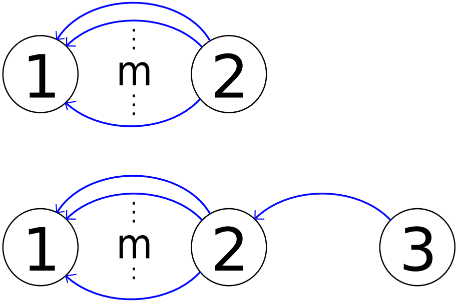
$$1 + \text{in-degree of vertex } j \text{ at that moment.}$$

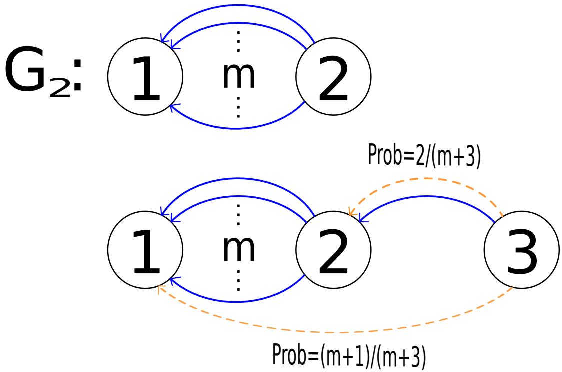
G_2:



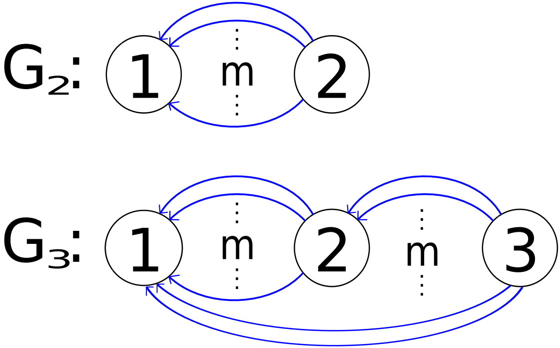

G₂:
1
m
2
Prob=1/(m+2)
1
m
2
3
Prob=(m+1)/(m+2)

G_2:





G₂:
1
m
2
Prob=2/(m+3)
1
m
2
3
Prob=(m+1)/(m+3)


G₂:
1
m
2
G₃:
1
m
2
m
3

- Vertex $n+1$ sequentially attaches m outgoing edges to vertices $\{1,\ldots,n\}$.
- The chance an outgoing edge attaches to vertex j is proportional to

$$1 + \text{in-degree of vertex } j \text{ at that moment.}$$

- We're interested in the joint distributional behavior of

$$W_j(n) = 1+ \text{ in-degree of vertex } j \text{ in } G_n.$$

- Actually we study $W_j(n)$ through $S_k(n)= \sum_{j=1}^{k} W_j(n)$.

Main result

- $S_k(n)$ is the sum of "weights" of the first k vertices in G_n.
- $X_1, \ldots, X_r$ are independent rate one exponential variables,

$$Z_k := (X_1 + \cdots + X_k)^{1/(m+1)}, \qquad 1 \le k \le r,$$

- $\mathbf{Z} = (Z_1, Z_2, \ldots, Z_r)$ and $\mathbf{S}(n) = \frac{(S_1(n), S_2(n), \ldots, S_r(n))}{(m+1)n^{m/(m+1)}}$.

Then:

$$\sup_K |\mathbb{P}[\mathbf{S}(n) \in K] - \mathbb{P}[\mathbf{Z} \in K]| \le \frac{C(r)}{n^{m/(m+1)}},$$

for some constant $C(r)$, where the supremum ranges over all convex subsets $K \subset \mathbb{R}^r$.

Immediate corollaries:

- Same rate of convergence of scaled joint degree counts $(W_1(n), \ldots, W_r(n))$ to limit $(Z_1, Z_2 - Z_1, \ldots, Z_r - Z_{r-1})$.
- Same rate of convergence of scaled maximum degree $\max_{1 \leq j \leq r} W_j(n)$ to limit $\max_{1 \leq j \leq r}(Z_j - Z_{j-1})$.

Generalizations:

- Different initial "seed" graphs.
- Different rule for defining the m edge PA graph.

Related results for the case $m = 1$:

- Our previous work (2013) showed rates of convergence of marginal distributions (though limits described differently).
- Flaxman, Frieze, Fenner (2005) showed the rate of growth of the maximum degree is $\sqrt{n}$.
- Móri (2005) showed a.s. convergence of the scaled joint degrees and the maximum using martingale arguments (no rates and limits not identified).

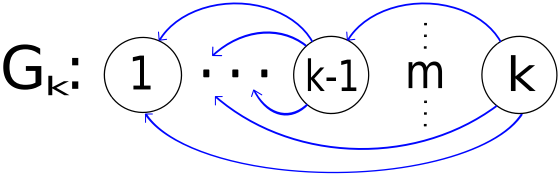
Applications:

- Bubeck, Mossel, Rácz (2014) use our results in a statistical inference problem.
- Curien, Duquesne, Kortchemski, Manolescu (2014) use our results to show the PA graph with $m = 1$ "converges" to an object related to Aldous's Brownian CRT.

Key proof idea

1. Let $\mathrm{Polya}(b, w; n)$ denote the law of the number of white balls in n draws and replacements of of a classical Pólya urn started with b black and w white balls. Then for $k \geq 2$:

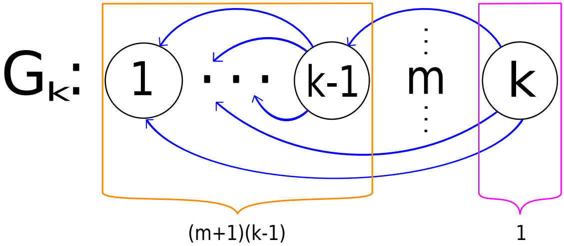
$$S_{k-1}(n)|S_k(n) \stackrel{d}{=} \mathrm{Polya}(1, (k-1)(m+1); S_k(n)-(k-1)m-k).$$



Key proof idea

1. Let $\mathrm{Polya}(b, w; n)$ denote the law of the number of white balls in n draws and replacements of of a classical Pólya urn started with b black and w white balls. Then for $k \geq 2$:

$$S_{k-1}(n) | S_k(n) \stackrel{d}{=} \mathrm{Polya}(1, (k-1)(m+1); S_k(n) - (k-1)m - k).$$



Key proof idea

1. Let $\mathrm{Polya}(b, w; n)$ denote the law of the number of white balls in n draws and replacements of of a classical Pólya urn started with b black and w white balls. Then for $k \geq 2$:

$$S_{k-1}(n)|S_k(n) \stackrel{d}{=} \mathrm{Polya}(1, (k-1)(m+1); S_k(n)-(k-1)m-k).$$

2. If $B[a, b]$ denotes a beta distributed random variable,

$$\mathrm{Polya}(b, w; n) \stackrel{d}{\approx} nB[w, b].$$

These two points imply the key identity

$$S_{k-1}(n)|S_k(n) \stackrel{d}{\approx} S_k(n)B[(k-1)(m+1), 1]. \qquad (*)$$

Key proof idea

Iterating the key identity

$$S_{k-1}(n) | S_k(n) \overset{d}{\approx} S_k(n) B[(k-1)(m+1), 1], \qquad (*)$$

leads to

$$(S_1(n), \ldots, S_r(n)) \overset{d}{\approx} \left(\prod_{k=1}^{r-1} B[k(m+1), 1], \prod_{k=2}^{r-1} B[k(m+1), 1], \ldots, 1 \right) S_r(n).$$

Beta-Gamma Algebra

Using the basic Beta-Gamma identity,

$$B[a, b] = \frac{G[a]}{G[a] + G[b]},$$

where $G[a]$ and $G[b]$ are independent gamma variables and the LHS is *independent* of the denominator of the RHS, and recalling

$$Z_k := (X_1 + \cdots + X_k)^{1/(m+1)}, \quad 1 \leq k \leq r,$$

we have the matching identity

$$(Z_1(n), \ldots, Z_r(n)) \stackrel{d}{=}$$

$$\left(\prod_{k=1}^{r-1} B[k(m+1), 1], \prod_{k=2}^{r-1} B[k(m+1), 1], \ldots, 1 \right) Z_r(n).$$

$$(S_1(n), \ldots, S_r(n)) \overset{d}{\approx}$$

$$\left(\prod_{k=1}^{r-1} B[k(m+1), 1], \prod_{k=2}^{r-1} B[k(m+1), 1], \ldots, 1 \right) S_r(n).$$

$$(Z_1(n), \ldots, Z_r(n)) \overset{d}{=}$$

$$\left(\prod_{k=1}^{r-1} B[k(m+1), 1], \prod_{k=2}^{r-1} B[k(m+1), 1], \ldots, 1 \right) Z_r(n).$$

So the problem is reduced to quantifying the difference between the marginal distributions of scaled $S_r(n)$ and Z_r.

Bounding

$$d_{\mathrm{Kol}}\left(\frac{S_r(n)}{(m+1)n^{m+1}}, Z_r\right)$$

uses

- Stein's method and
- Z_r as the unique fixed point of a distributional transformation related to the beta-gamma algebra.

E. Peköz, A. Röllin, and N. Ross. Joint degree distributions of preferential attachment random graphs (2014). http://arxiv.org/abs/1402.4686.

E. Peköz, A. Röllin, and N. Ross. Generalized gamma approximation with rates for urns, walks and trees (2013). http://arxiv.org/abs/1309.4183

Related work:

E. Peköz, A. Röllin, and N. Ross. Degree asymptotics with rates for preferential attachment random graphs (2013). *Ann. Appl. Probab.*


STEIN
STEIN
STEIN
STEIN
STEIN
STEIN


STEIN
STEIN
STEIN
STEIN
STEIN


STEIN
STEIN
STEIN
STEIN
STEIN
STEIN
STEIN
STEIN

Thank You!