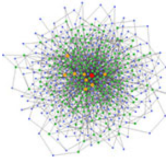
Self-similar hooking networks

Hosam Mahmoud
Department of Statistics
The George Washington University

Talk at AofA 2019



June 24, 2019

Plan

- Introducing self-similar hooking networks
- Models of randomness: uniform and preferential attachment
- Local and global degree profiles
- I rely on intuition and graphics
- Some exact results: Stirling numbers of both kinds
- Gaussian asymptotic distributions

Self-similar hooking networks

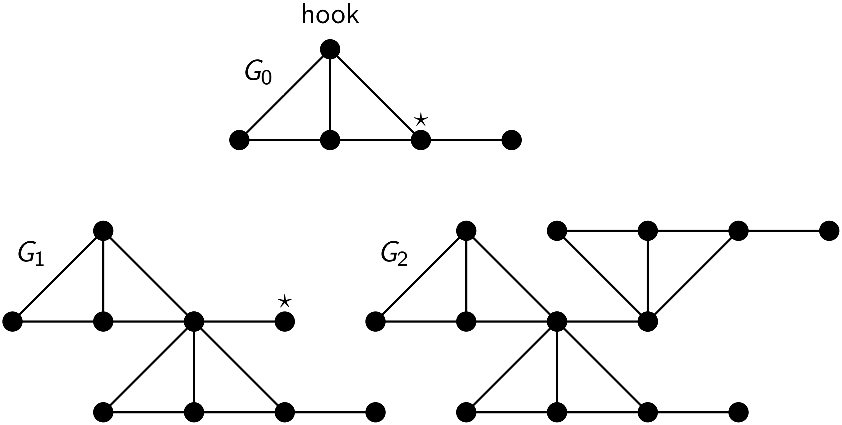


Figure: A seed (top) and two networks grown from it (second row). The nodes chosen for hooking are marked with stars.

PORT: Barabási-Albert network

We call the class of graphs so built *self-similar hooking networks*. The class includes many known families of graphs. For example, certain variations of scale-free trees

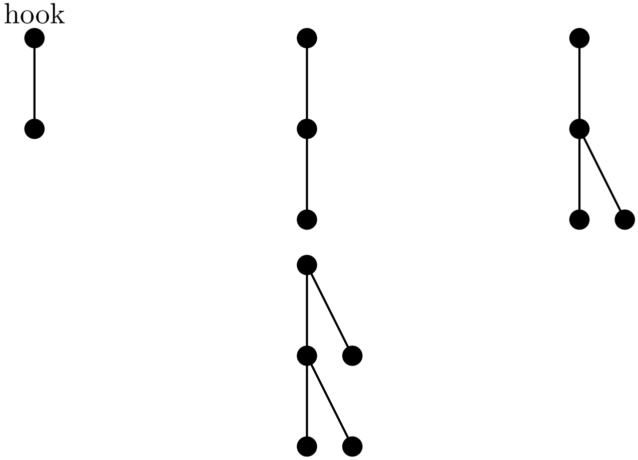
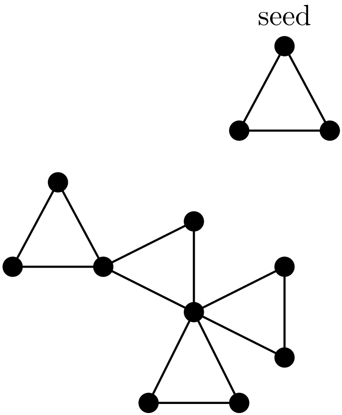


Figure: PORT: a scale-free tree, Barabási-Albert network.

Cactus graphs



Mariam Bahrani and Jérémie Lumbroso are working on this model

Scope of the study

Our intent is the study of two kinds of degree profiles in the self-similar hooking network:

- local: tracks the evolution of a specific vertex in the network since its appearance through later points in time.
  What happens to a node that appears at step j, when the network ages to n?

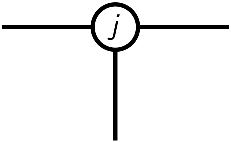


Figure: Degree of node j at time n.

- global: looks at the distribution of degrees in the entire network via raw counts. For example, $X_n^{(1)}$ may be the number of nodes of the degree of the hook by age n.

Scope of the results

- **Local:** We find the exact distribution, which involves Stirling numbers of the first kind. Asymptotically, we observe phase transitions—the degree of *early* nodes have an asymptotic normal distribution (as the notwork size grows), the *intermediate* nodes (linear in age) have an asymptotic Poisson distribution, and the very *late* arrivals have a degenerate distribution. The terms early, intermediate and late will be made precise in the sequel.
- **global:** We study only the nodes of the two smallest degrees to show how in principle the nodes of small degrees interact in an asymptotic multivariate normal distribution. For regular seed graphs, we show that we can push the boundary somewhat to included higher degrees.

Notation

$$|V| = v;$$

$$|E| = \eta;$$

$$h = \text{hook degree}$$

Technically speaking, we do not have one seed, but a sequence $S_0, S_1, S_2, \ldots$. Assume the nodes of the seed are numbered $1, \ldots, v$, an arbitrary numbering in which the hook is numbered 1. We take this numbering as canonical to be preserved in every copy of S.

Generalized harmonic numbers

$$H_n^{(s)}(x) = \frac{1}{(1+x)^s} + \frac{1}{(2+x)^s} + \cdots + \frac{1}{(n+x)^s} .$$

Node specification

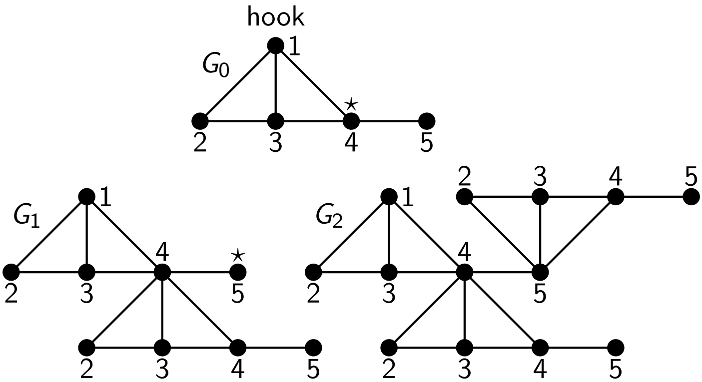


Figure: Canonical numbering.

We can thus speak of node (j,r): *r*th node in the seed copy that joins in the *j*th step.

Degrees

Let $d(r)$ be the degree of node r in the seed, for $1 \leq r \leq v$. We use the notation $\Delta_{j,n}^{(r)}$ to denote the degree at age n of node (j, r). For instance, we have $\Delta_{0,2}^{(4)} = 6$, $\Delta_{1,2}^{(2)} = 2$, and $\Delta_{2,2}^{(3)} = 3$ in the running example.

Models of randomness

Two models of randomness arise naturally.

Uniform attachment: In this model, the hooking vertex in G_{n-1} is selected uniformly at random from among all the existing nodes in G_{n-1}. Here, every node in the existing network is given a fair equal chance to be the hooking vertex.

Preferential attachment: In this model, the hooking vertex in G_{n-1} is selected with a probability proportional to its degree in G_{n-1}. Here, popular vertices are more likely to be successful in attracting and hooking the next copy of the seed, pursuing the Mathew principle according to which "the rich get richer," or "success breeds success."

Colin Desmarais will tell us about models that combine these under a bigger umbrella

Local degree profile under uniform attachment

Proposition: mean

Let $0 \le j \le n$. For $1 \le r \le v$, let $\Delta_{j,n}^{(r)}$ be the degree of the rth vertex in copy S_j of the seed in a uniform attachment self-similar hooking network at age n. For $(j, r) = (0, 1)$ or $2 \le r \le v$ and any $j \ge 0$, we have

$$\mathbb{E}[\Delta_{j,n}^{(r)}] = d(r) + \frac{h}{(v-1)}\Big(H_n\Big(\frac{1}{v-1}\Big) - H_j\Big(\frac{1}{v-1}\Big)\Big)$$

$$= \begin{cases} \frac{h}{(v-1)}\big(\ln n - \psi\big(\frac{1}{v-1}\big) - v + 1\big) + d(r) \\ \qquad - \frac{h}{(v-1)} H_j\big(\frac{1}{v-1}\big) + O\big(\frac{1}{n}\big), & \text{if } \; j \ge 0 \text{ is fixed;} \\ \frac{h}{(v-1)} \ln \frac{n}{j} + d(r) + O\big(\frac{1}{j}\big), & \text{if } \; n \ge j \to \infty. \end{cases}$$

Variance

$$
\begin{aligned}
\mathbb{V}\mathrm{ar}[\Delta_{j,n}^{(r)}] &= \frac{h^2}{(v-1)}\Big(H_n\Big(\frac{1}{v-1}\Big) - H_j\Big(\frac{1}{v-1}\Big)\Big) \\
&\qquad - \frac{h^2}{(v-1)^2}\Big(H_n^{(2)}\Big(\frac{1}{v-1}\Big) - H_j^{(2)}\Big(\frac{h^2}{v-1}\Big)\Big) \\
&= \begin{cases} \frac{h^2}{(v-1)} \ln n + O(1), & \text{if } \ j \geq 0 \text{ is fixed;} \\ \frac{h^2}{(v-1)} \ln \frac{n}{j} + O\big(\frac{1}{j}\big), & \text{if } \ n \geq j \to \infty. \end{cases}
\end{aligned}
$$

Exact distribution

Theorem

Let $0 \leq j \leq n$. For $1 \leq r \leq v$, let $\Delta_{j,n}^{(r)}$ be the degree of the rth vertex in copy S_j of the seed in a self-similar hooking network at age n. For $(j, r) = (0, 1)$ or $2 \leq r \leq v$ and any $j \geq 0$, we have

$$\mathbb{P}(\Delta_{j,n}^{(r)} = k) = \begin{cases} \frac{1}{\langle j+1+\frac{1}{v-1}\rangle_{n-j}(v-1)^m} \sum_{i=m}^{n-j} \left[{n-j \atop i} \right] \binom{i}{m} (j+1)^{i-m}, \\ \qquad \qquad \text{for} \;\; k = d(r) + mh, m \geq 0; \\ 0, \qquad \qquad \text{otherwise.} \end{cases}$$

Remark

It is evident from Theorem 1 that the admissible degrees of the rth node of any generation are additive translates of the initial degree of r in S_j by multiple of h. Indeed, every time that node is chosen as the hooking vertex, the hook of a new copy of the seed is identified with it, increasing its degree by yet another h.

Phases

The mean and variance of the network node degrees indicate a phase change, as the relation between j and n changes.

- **Early:** Fixed j: the asymptotic mean is $\frac{h}{v-1} \ln n$, as $n \to \infty$, and the role of j provides only negligible lower order asymptotics.
- **Early:** j grows to infinity with n, but remains $o(n)$, such as the case
$$j = j(n) = \lceil 5n^{\frac{3}{4}} + 2\sqrt{n} + 6 \rceil.$$
In this range, j provides essential leading-term asymptotics. In the example, the mean is $\frac{h}{v-1} \ln \frac{n}{j} = \frac{h}{4(v-1)} \ln n$.
- **Intermediate:** $j \sim \alpha n$, for $0 < \alpha < 1$, such as the case $\lceil \frac{3}{7} n \rceil$, the asymptotic mean is $d(r) + \frac{h}{v-1} \ln \frac{1}{\alpha}$.
- **Late:** $j \sim n$ The asymptotic mean of late nodes is just $d(v)$, showing that the late arrivals, such as the case
$$j = \lfloor n - 3 \ln(n+1) + 11 \rfloor,$$
have negligible probability of participating in the hooking events; their degrees remain the same as the initial.

Phases in the asymptotic distribution

Phases in the mean are reflected in the asymptotic distributions.

Theorem

Let $0 \le j \le n$. For $1 \le r \le v$, let $\Delta_{j,n}^{(r)}$ be the degree of the rth vertex in copy S_j of the seed in a uniform attachment self-similar hooking network at age n. For $(j, r) = (0, 1)$ or $2 \le r \le v$ and any $j \ge 0$, we have

$$\frac{\Delta_{j,k}^{(r)} - \frac{h}{v-1} \ln n}{\sqrt{\ln n}} \xrightarrow{\mathcal{D}} \mathcal{N}\Big(0, \frac{h^2}{v-1}\Big), \qquad \text{for } j \text{ fixed};$$

$$\frac{\Delta_{j,k}^{(r)} - \frac{h}{v-1} \ln \frac{n}{j}}{\sqrt{\ln \frac{n}{j}}} \xrightarrow{\mathcal{D}} \mathcal{N}\Big(0, \frac{h^2}{v-1}\Big), \qquad \text{for } j \to \infty \text{ and } j = o(n);$$

$$\Delta_{j,k}^{(r)} \xrightarrow{\mathcal{D}} d(r) + h\,\mathrm{Poisson}\Big(\frac{1}{v-1} \ln \frac{1}{\alpha}\Big), \qquad \text{for } j \sim \alpha n, \quad 0 < \alpha < 1;$$

$$\Delta_{j,k}^{(r)} \xrightarrow{P} d(r), \qquad \text{for } j = n - o(n).$$

A glimpse at the proof

Let $\mathcal{H}_{j,k}^{(r)}$ be the event that node (j, r) is chosen as the hooking vertex in the network G_{k-1} to attach S_k, and let $\mathbb{I}_{\mathcal{H}_{j,k}^{(r)}}$ be its indicator. Note that the events $\mathcal{H}_{j,k}^{(r)}$, for $k = j+1, \dots, n$ are independent. Under uniform attachment, the event $\mathcal{H}_{j,k}^{(r)}$ occurs with probability $1/\tau_{k-1}$. Therefore, the indicator of this event is a Bernoulli random variable with the moment generating function

$$1 - \frac{1}{\tau_{k-1}} + \frac{e^t}{\tau_{k-1}}.$$

At step $k > j$, if the event $\mathcal{H}_{j,k}^{(r)}$ occurs, the degree of node (j, r) goes up by h; otherwise it remains unchanged. Whence, we have a stochastic recurrence:

$$\Delta_{j,n}^{(r)} = d(r) + h\,\mathbb{I}_{\mathcal{H}_{j,j+1}^{(r)}} + h\,\mathbb{I}_{\mathcal{H}_{j,j+2}^{(r)}} + \cdots + h\,\mathbb{I}_{\mathcal{H}_{j,n}^{(r)}}.$$

A glimpse at the proof: mean

$$
\begin{aligned}
\mathbb{E}\big[\Delta_{j,n}^{(r)}\big] &= d(r) + h\,\mathbb{P}\big(\mathcal{H}_{j,j+1}^{(r)}\big) + \cdots + h\,\mathbb{P}\big(\mathcal{H}_{j,n}^{(r)}\big) \\
&= d(r) + h \sum_{k=j+1}^{n} \frac{1}{\tau_{k-1}} \\
&= d(r) + \frac{h}{(v-1)} \sum_{k=j+1}^{n} \frac{1}{k + \frac{1}{v-1}} \\
&= d(r) + \frac{h}{(v-1)} \Big(H_n\Big(\frac{1}{v-1}\Big) - H_j\Big(\frac{1}{v-1}\Big)\Big).
\end{aligned}
$$

The asymptotic average follows from the approximation of harmonic numbers by logarithms.

A glimpse at the proof: variance

By the independence of the hooking events, we similarly have

$$\begin{aligned}
\mathbb{V}\mathrm{ar}\big[\Delta_{j,n}^{(r)}\big] &= h^2 \sum_{k=j+1}^{n} \mathbb{V}\mathrm{ar}\big[\mathbb{I}_{\mathcal{H}_{j,k}^{(r)}}\big] \\
&= \frac{h^2}{(v-1)} \sum_{k=j+1}^{n} \frac{1}{k+\frac{1}{v-1}} - \frac{h^2}{(v-1)^2} \sum_{k=j+1}^{n} \frac{1}{(k+\frac{1}{v-1})^2} \\
&= \frac{h^2}{(v-1)}\Big(H_n\Big(\frac{1}{v-1}\Big) - H_j\Big(\frac{1}{v-1}\Big)\Big) \\
&\qquad - \frac{h^2}{(v-1)^2}\Big(H_n^{(2)}\Big(\frac{1}{v-1}\Big) - H_j^{(2)}\Big(\frac{1}{v-1}\Big)\Big).
\end{aligned}$$

The asymptotic average follows from the approximation of harmonic numbers by logarithms.

A glimpse at the proof: exact and symptotic distribution

Define the moment generating function $\phi^{(r)}_{j,n} = \exp(\Delta^{(r)}_{j,n}\, t)$. The stochastic recurrence yields

$$\phi^{(r)}_{j,n} = \mathbb{E}\Big[e^{(d(r)+h\,\mathbb{I}_{\mathcal{H}^{(r)}_{j,j+1}}+h\,\mathbb{I}_{\mathcal{H}^{(r)}_{j,j+2}}+\dots+h\,\mathbb{I}_{\mathcal{H}^{(r)}_{j,n}})t}\Big] = e^{d(r)t}\,\mathbb{E}\Big[\prod_{k=j+1}^{n} e^{h\,\mathbb{I}_{\mathcal{H}^{(r)}_{j,k}}\,t}\Big];$$

decomposition into a product is by independence of the events:

$$\begin{aligned}
\phi^{(r)}_{j,n} &= e^{d(r)t}\,\frac{(\tau_j - 1 + e^{ht})(\tau_{j+1} - 1 + e^{ht})\cdots(\tau_{n-1} - 1 + e^{ht})}{\tau_j\tau_{j+1}\cdots\tau_{n-1}}\\
&= e^{d(r)t}\,\frac{((v-1)(j+1) + e^{ht})\cdots((v-1)n + e^{ht})}{((v-1)(j+1)+1)\cdots((v-1)n+1)}\\
&= e^{d(r)t}\,\frac{(j+1+\frac{e^{ht}}{v-1}))\cdots(n+\frac{e^{ht}}{v-1})}{(j+1+\frac{1}{v-1})\cdots(n+\frac{1}{v-1})}\\
&= e^{d(r)t}\,\frac{\Gamma(n+1+\frac{e^{ht}}{v-1})\,\Gamma(j+1+\frac{1}{v-1})}{\Gamma(n+1+\frac{1}{v-1})\,\Gamma(j+1+\frac{e^{ht}}{v-1})}.
\end{aligned}$$

A glimpse at the proof: exact distribution

To obtain a probability generating function

$$\zeta_{j,n}(u) = \sum_{i=0}^{\infty} \mathbb{P}(\Delta_{j,n}^{(r)} = i) u^i,$$

we substitute $\ln u$ for t in the generating function

$$\zeta_{j,n}^{(r)} = u^{d(r)} \frac{(j+1+\frac{u^h}{v-1})\ldots(n+\frac{u^h}{v-1})}{(j+1+\frac{1}{v-1})\cdots(n+\frac{1}{v-1})}$$

$$= u^{d(r)} \frac{\langle j+1+\frac{u^h}{v-1}\rangle_{n-j}}{\langle j+1+\frac{1}{v-1}\rangle_{n-j}}.$$

Using the generating function, we write

$$\zeta_{j,n}^{(r)} = u^{d(r)} \frac{u^{d(r)}}{\langle j+1+\frac{1}{v-1}\rangle_{n-j}} \sum_{i=0}^{n-j} \left[{n-j \atop i} \right] \Big(j+1+\frac{u^h}{v-1}\Big)^i$$

$$= \frac{u^{d(r)}}{\langle j+1+\frac{1}{v-1}\rangle_{n-j}} \sum_{i=0}^{n-j} \left[{n-j \atop i} \right] \sum_{m=0}^{i} (j+1)^{i-m} \Big(\frac{u^h}{v-1}\Big)^m \binom{i}{m}$$

A glimpse at the proof: early phase

We start with the phase in which j is fixed. For this case, by the Stirling approximation in (see handout), we write

$$\phi_{j,n}^{(r)}\Big(\frac{t}{\sqrt{\ln n}}\Big) = e^{d(r)t/\sqrt{\ln n}}\,\frac{\Gamma\big(n+1+\frac{e^{ht/\sqrt{\ln n}}}{v-1}\big)\,\Gamma\big(j+1+\frac{1}{v-1}\big)}{\Gamma\big(n+1+\frac{1}{v-1}\big)\,\Gamma\big(j+1+\frac{e^{ht/\sqrt{\ln n}}}{v-1}\big)}$$

$$\sim n^{e^{ht/\sqrt{\ln n}}/(v-1)-1/(v-1)}.$$

Going through a local expansion of the exponential, we write

$$\phi_{j,n}^{(r)}\Big(\frac{t}{\sqrt{\ln n}}\Big) \sim e^{\Big(\Big(\frac{1}{v-1}\Big(1+\frac{ht}{\sqrt{\ln n}}+\frac{h^2t^2}{2\ln n}+O\Big(\frac{t^3}{\ln n}\Big)\Big)\Big)-\frac{1}{v-1}\Big)\ln n}.$$

Equivalently, we can write this relation as

$$\phi_{j,n}^{(r)}\Big(\frac{t}{\sqrt{\ln n}}\Big)e^{-\frac{h}{v-1}\,t\sqrt{\ln n}} \sim e^{\Big(\frac{h^2t^2}{2(v-1)}+O\Big(\frac{t^3}{\sqrt{\ln n}}\Big)\Big)}.$$

So, at any fixed $t\in\mathbb{R}$, we can finally the convergence

$$\mathbb{E}\Big[e^{\frac{\Delta_{j,k}^{(r)}-\frac{h}{v-1}\ln n}{\sqrt{\ln n}}\,t}\Big] \to e^{\frac{h^2t^2}{2(v-1)}}.$$

A glimpse at the proof: Poisson phase

In the intermediate and late phases $j \sim \alpha n$, for $\alpha \in (0, 1]$, no scaling is required to get convergence in distribution. Instead, we have

$$\begin{aligned}\phi_{j,n}^{(r)}(t) &= e^{d(r)t}\,\frac{n^{e^{ht}/(v-1)-1/(v-1)}(1+O(1/n))}{j^{e^{ht}/(v-1))-1/(v-1)}(1+O(1/j))}\\ &\to e^{d(r)t}\left(\frac{1}{\alpha}\right)^{\frac{1}{v-1}(e^{ht}-1)}\\ &= e^{d(r)t+\frac{\ln\frac{1}{\alpha}}{v-1}(e^{ht}-1)};\end{aligned}$$

The Poisson moment generating function. Technically speaking

$$\Delta_{j,k}^{(r)} \xrightarrow{\mathcal{D}} d(r) + \mathrm{Poi}\left(\frac{1}{\alpha}\right).$$

Local degree profile under preferential attachment

Suppose the affinity of a vertex in the network to attract the next hook is proportional to its degree. At any step, if a vertex is chosen for hooking, its degree goes up by h, otherwise it stays the same. The total affinity (which we call ν_n) in all the vertices of the network G_n is the sum of all the degrees of the nodes in G_n, which is twice the number of edges in that network:

$$\nu_n = 2\eta(n+1).$$

Letting $\mathbb{F}_n$ be the sigma field generated by the first n steps, preferential attachment conditional distribution is

$$\Delta_{j,n}^{(r)} \,|\, \mathbb{F}_{n-1} = \begin{cases} \Delta_{j,n-1}^{(r)} + h, & \text{with probability } \frac{\Delta_{j,n-1}^{(r)}}{\nu_{n-1}}; \\ \Delta_{j,n-1}^{(r)}, & \text{with probability } 1 - \frac{\Delta_{j,n-1}^{(r)}}{\nu_{n-1}}. \end{cases}$$

colors in a Pólya urn

I always use white and blue.
Today I make an exception: **red** and **blue**.

Mapping onto a Pólya urn

To keep track of the degree of node $y = (j, r)$, we map the edges onto a Pólya urn via a color code. We think of an edge (x, y) joining vertices x and y in the network as two "half edges," one half edge accounting for the incidence with x, the other half edge accounting for the incidence with y.
Let us color each half edge incident with $y = (j, r)$ with **red** and color all the rest of the half edges with **blue** .

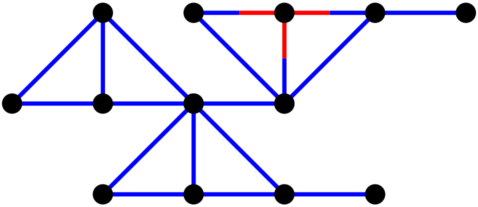


Figure: Coloring scheme as a Pólya urn.

Pick a red

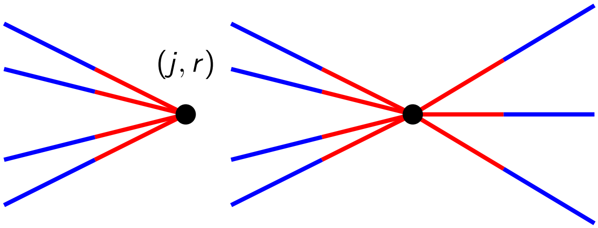


Figure: pick a red $\implies$ add h red $2\eta - h$ blue balls.

Pick a blue

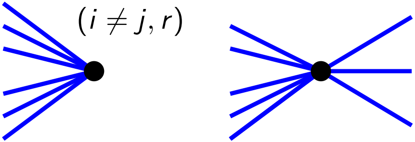


Figure: pick a blue $\implies$ add 2η blue balls.

Exact moments

Let W_m be the number of white balls after m draws from the urn. We take the initial state to correspond to G_j. Thus, initially we have $W_0 = d(y)$ white balls and $2\eta(j+1) - d(y)$ blue balls in the urn. Our interest is in the count W_{n-j}, the degree of y, when the network is at age n.

These dynamics are captured by the Pólya replacement matrix

$$\begin{pmatrix} h & 2\eta - h \\ 0 & 2\eta \end{pmatrix}.$$

In view of the 0 entry, this Pólya urn scheme falls in the class of triangular urn schemes.

Recently investigated, Chen Chen, Panpan Zhang, Mahmoud (2015) gives the exact moments:

$$\mathbb{E}\left[\left(\Delta_{j,n}^{(r)}\right)^k\right] = \mathbb{E}[W_{n-j}^k] = \frac{h^k}{(n-j)!}\sum_{i=1}^{k}(-1)^{k-i}\left\{ {k \atop i} \right\}\left\langle \frac{d(r)}{h} \right\rangle_i \left\langle 1+\frac{ih}{2\eta} \right\rangle_{n-j}.$$

Gamma-type limit

In particular, we have the average

$$\mathbb{E}\big[\Delta_{j,n}^{(r)}\big] = \frac{d(r)\,\Gamma\big(n-j+1+\frac{h}{2\eta}\big)}{\Gamma(n-j+1)\,\Gamma\big(1+\frac{h}{2\eta}\big)}.$$

With the proper scaling, we have convergence of moments

$$\mathbb{E}\Big[\Big(\frac{\Delta_{j,n}^{(r)}}{(n-j)^{h/(2\eta)}}\Big)^k\Big] \to \frac{h^k\,\Gamma\big(k+\frac{d(r)}{h}\big)}{\Gamma\big(\frac{d(r)}{h}\big)\,\Gamma\big(1+\frac{hk}{2\eta}\big)}, \qquad \text{provided } n-j \to \infty.$$

These moments uniquely determine a limiting distribution of the gamma type, as **Svante Janson** calls them.

Global profile: uniform attachment

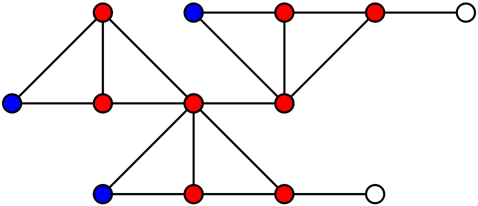
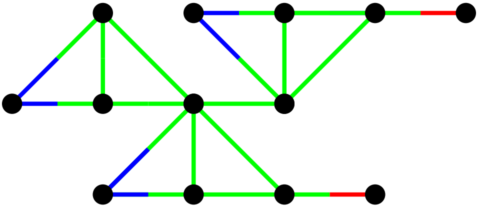


Figure: Coloring scheme for uniform attachment.

Global profile: preferential attachment

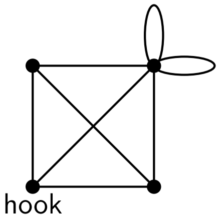


Admissible degrees

Let the distinct degrees in the seed be $d_1 < d_2 < \cdots < d_k$, and let $d_1^* < d_2^* < d_3^* < \cdots$ be the distinct admissible degrees in the network. Necessarily, we have $d_1^* = d_1$, and $d_2^* \le d_1 + h$.

In the seed
n_1 nodes of degree d_1^*
n_2 nodes of degree d_2^*.

Note that n_2 may be 0, as in the seed in the figure, where we have $d_2^* = 6$, in a self-similar network built from this seed, but no vertex in the seed has degree 6.



Regimes

(i) The hook has the smallest degree in the admissible sequence, $h = d_1^*$, and the second admissible degree in the network is $d_2^* = d_1^* + h = 2h$.

(ii) The hook has the smallest degree in the admissible sequence, and the second admissible degree in the network is $d_2^* < d_1^* + h$.

(iii) The hook has the second smallest degree in the admissible sequence: $h = d_2^*$.

(iv) The hook degree is larger than the second smallest in the admissible sequence: $h > d_2^*$.

Pólya urn: uniform

So, we color nodes of degree d_1^* with color 1, nodes of degree d_2^* with color 2, and nodes of higher degrees with color 3. The replacement matrix is different for each regime.

Lemma

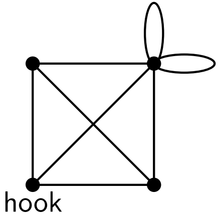
Let $N = v - n_1 - n_2$. The four regimes (i)–(iv) correspond to Pólya urn schemes with the respective replacement matrices

$$(i)\ \begin{pmatrix} n_1-2 & n_2+1 & N \\ n_1-1 & n_2-1 & N+1 \\ n_1-1 & n_2 & N \end{pmatrix}, \qquad (ii)\ \begin{pmatrix} n_1-2 & n_2 & N+1 \\ n_1-1 & n_2-1 & N+1 \\ n_1-1 & n_2 & N \end{pmatrix},$$

$$(iii)\ \begin{pmatrix} n_1-1 & n_2-1 & N+1 \\ n_1 & n_2-2 & N+1 \\ n_1 & n_2-1 & N \end{pmatrix}, \qquad (iii)\ \begin{pmatrix} n_1-1 & n_2 & N \\ n_1 & n_2-1 & N \\ n_1 & n_2 & N-1 \end{pmatrix}.$$

The initial urn contains n_1 balls of color 1, n_2 balls of color 2, and N balls of color 3.

A concrete example



Admissible degrees: $3, 6, 7, 9, 10, 12, 13, 15, \ldots$.

$$d_1^* = 3, \qquad d_2^* = 6$$

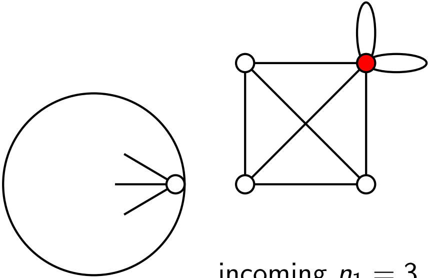
$$n_1 = 3, \qquad n_2 = 0, \qquad N = 1.$$

Dynamics of growth (Regime (i))

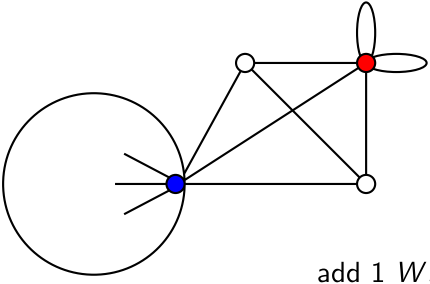
$$n_1 = 3, \qquad n_2 = 0, \qquad N = 1.$$

$$\begin{pmatrix} n_1 - 2 & n_2 + 1 & N \\ n_1 - 1 & n_2 - 1 & N + 1 \\ n_1 - 1 & n_2 & N \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & 2 \\ 2 & 0 & 1 \end{pmatrix}.$$

Dynamics: row 1

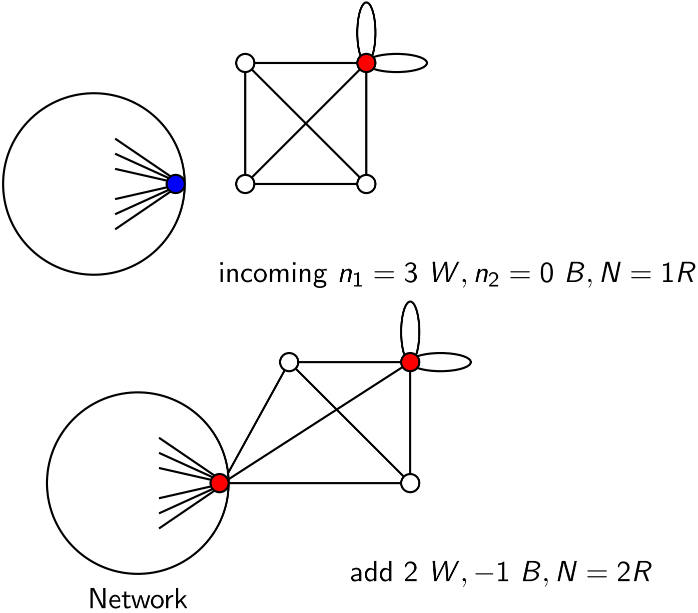


incoming $n_1 = 3\ W, n_2 = 0\ B, N = 1R$

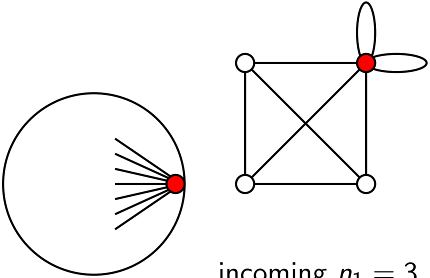


add $1\ W, 1\ B, N = 1R$

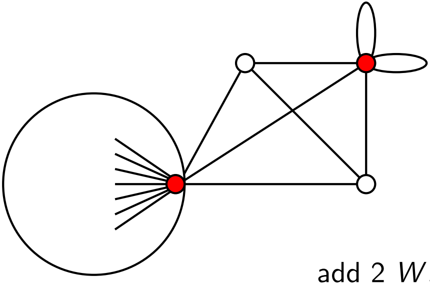
Dynamics: row 2



Dynamics: row 3



incoming $n_1 = 3\ W, n_2 = 0\ B, N = 1R$



Network

add $2\ W, 0\ B, N = 1R$

strong laws

Theorem

Let $D_n^{(i)}$ be the number of nodes of admissible outdegree d_i^*, for $i = 1, 2$ in a uniform attachment self-similar hooking network of age n. Let $\mathbf{D}_n = \begin{pmatrix} D_n^{(1)} \\ D_n^{(2)} \end{pmatrix}$. We then have in Regime

(i) $\frac{1}{n}\,\mathbf{D}_n \xrightarrow{a.s.} \frac{v-1}{v} \begin{pmatrix} n_1 - 1 \\ n_2 + \frac{n_1 - 1}{v} \end{pmatrix}.$

(ii) $\frac{1}{n}\,\mathbf{D}_n \xrightarrow{a.s.} \frac{v-1}{v} \begin{pmatrix} n_1 - 1 \\ n_2 \end{pmatrix}.$

(iii) $\frac{1}{n}\,\mathbf{D}_n \xrightarrow{a.s.} \frac{v-1}{v} \begin{pmatrix} n_1 \\ n_2 - 1 \end{pmatrix}.$

(iv) $\frac{1}{n}\,\mathbf{D}_n \xrightarrow{a.s.} \frac{v-1}{v} \begin{pmatrix} n_1 \\ n_2 \end{pmatrix}.$

Furthermore, for nondegenerate networks, we have the central limit theorem

$$\frac{\mathbf{D}_n - (v-1)\mathbf{v}_1^{(\text{regime})}\, n}{\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}_3(\mathbf{0}, \mathbf{\Sigma}^{(\text{regime})}),$$

for some computable covariance matrix $\mathbf{\Sigma}^{(\text{regime})}$.

Regular graphs

Take seed to be a triangle. This seed is in Regime (i). In the uniform attachment case, the Pólya urn scheme underlying the network growth has replacement matrix

$$\begin{pmatrix} 1 & 1 & 0 \\ 2 & -1 & 1 \\ 2 & 0 & 0 \end{pmatrix}.$$

We can write an explicit central limit theorem:

$$\frac{\begin{pmatrix} D_n^{(1)} \\ D_n^{(2)} \end{pmatrix} - \begin{pmatrix} \frac{4}{3} \\ \frac{4}{9} \end{pmatrix} n}{\sqrt{n}} \xrightarrow{\mathcal{D}} \mathcal{N}_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{9} & -\frac{17}{108} \\ -\frac{17}{108} & \frac{173}{648} \end{pmatrix}\right).$$

Global profile: preferential attachment

To put the uniform attachment strategy in perspective, we write the counterpart central limit theorem for preferential attachment (details omitted):

$$\frac{\begin{pmatrix} \widetilde{D}_n^{(1)} \\ \widetilde{D}_n^{(2)} \end{pmatrix} - \begin{pmatrix} 3 \\ \frac{6}{5} \end{pmatrix} n}{\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{3}{5} & -\frac{6}{5} \\ -\frac{6}{5} & \frac{612}{175} \end{pmatrix} \right).$$

Thank you.