The number of descendants in preferential attachment graphs

Tiffany Lo
Joint work with Svante Janson

Stockholm University

September 9, 2026

Preferential attachment models (PAM)

- A popular class of random graphs for modelling real-world networks since it was proposed by Barabási and Albert (1999) to explain the *power-law* degree sequences observed in real-world networks.
- We study the 'sequential' version of *directed* PAM, with parameters $m \in \{1, 2, \dots\}$ and $\rho > -m$.
- No self-loops, but contains multiple edges.

The sequential PAM

1. G_1: One vtx (labelled 1) with no edges.
2. G_2: Vertices 1 and 2, joined by m outgoing edges, all directed from vtx 2 to 1.
3. Given G_{n-1}, build G_n by adding vtx n to G_{n-1}, and add m outgoing edges of n one by one: Choose vtx $k \in [n-1]$ as the recipient of the i-th edge of n w.p. $\propto$

   $$\rho + \text{degree of } k \text{ in } G_{n-1} + \#\text{first } i-1 \text{ edges attached to } k.$$

4. We obtain G_n after attaching all m edges of vtx n.
   $\mathrm{PA}(n, m, \rho)$: the law of G_n.

An example ($m = 2$, $\rho > -2$)

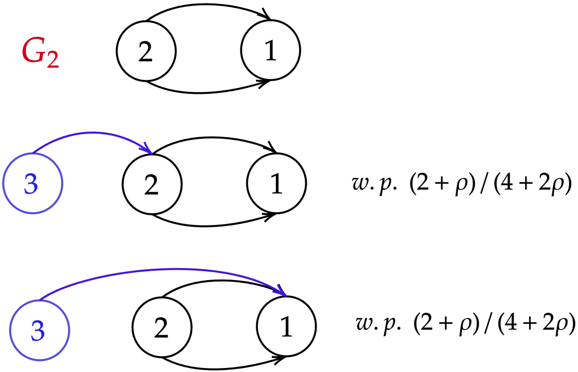
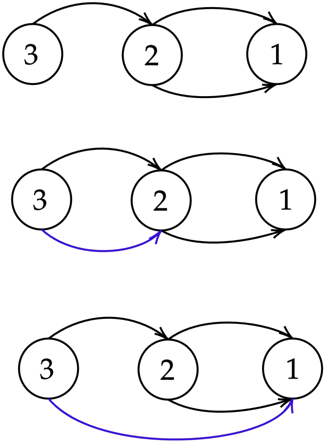


Figure: 2 ways of adding the 1st edge of vtx. 3.

Example ctd.



$w.p.\ (3+\rho)/(5+2\rho)$

$w.p.\ (2+\rho)/(5+2\rho)$

Figure: Two ways of adding the 2nd edge of vtx. 3.

The number of descendants

- # descendants: # vertices (including n) in G_n that can be reached from n via a directed path.

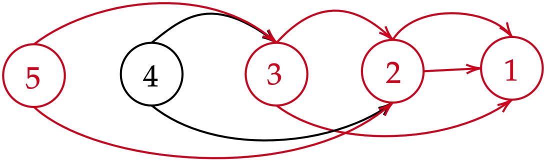


Figure: $n = 5$ and $m = 2$. The descendants are marked in red

Related work: $m = 1$

- $X^{(n)}$: # descendants in G_n.
- When $m = 1$, $X^{(n)} = 1+$ depth of vtx n.

$$\frac{X^{(n)}}{\log n} \xrightarrow{p} \frac{1+\rho}{2+\rho}; \tag{1}$$

  other results include exact distribution, CLT and Poisson approximation. See Kuba & Wagner (2010), Dobrow & Smythe (1996) and Panholzer & Prodinger (2007).

- We thus restrict ourselves to $m \geq 2$.

Related work: uniform attachment

G_1 – a single vtx (labelled 1) with no edges. To build G_n from G_{n-1}, add vtx n and its $m \geq 1$ outgoing edges to G_{n-1}; choose the endpoint of each edge uniformly among the existing $[n-1]$ vertices.

A result due to Janson (RSA, 2024):

Thm (# descendants in UAM)

For $m \geq 2$,

$$\frac{X^{(n)}}{n^\alpha} \xrightarrow{d} \frac{\pi(m-1)^{1/m}}{m \sin(\pi/m)} \xi^{1-\alpha} \quad \text{as } n \to \infty, \tag{2}$$

where $\alpha = (m-1)/m$ and $\xi \sim \mathrm{Gamma}(m/(m-1), 1)$.
Furthermore, all moments in (2) converge as $n \to \infty$.

Main result - preferential attachment

In analogue to the result of Janson (RSA, 2024):

Thm (# descendants in sequential PAM)

For $m \geq 2$,

$$\frac{X^{(n)}}{n^{\nu}} \xrightarrow{d} C_{m,\rho}\xi^{1-\nu} \quad \text{as } n \to \infty, \tag{3}$$

where $\nu = \frac{(m-1)(m+\rho)}{m(m+\rho+1)}$ and $\xi \sim \mathrm{Gamma}(m/(m-1), 1)$.
Furthermore, all moments in (3) converge as $n \to \infty$.

Remarks: the constant $C_{m,p}$ is explicit.
Similar results also hold for the sequential PAM that allows possible self-loops.

Proof ideas

Pólya urns and sequential PAMs

An alternative construction of sequential PAM due to Berger, Borgs, Chayes & Saberi (2014):

Pólya urn representation (assuming $\rho = 0$)

Let $(B_k)_1^\infty$ be independent variables with $B_1 = 1$ and

$$B_k \sim \mathrm{Beta}(m, (2k-3)m), \quad k \geq 2.$$

Given $(B_k)_1^\infty$, construct for each $n \geq 1$ a directed graph on n vertices – each vtx $2 \leq k \leq n$ has m outgoing edges, an edge of vtx k attaches to $i \in [k-1]$ w.p. $B_i \prod_{j=i+1}^{k-1}(1-B_j)$.

Let $\mathrm{PU}(n, m)$ be the law of the n-vertex graph above:

Thm

For every $n \geq 1$ and $m \geq 1$, $\mathrm{PU}(n, m) = \mathrm{PA}(n, m, 0)$.

Why an urn – heuristics

- We attach an edge of vtx n by considering vtx $n-1, n-2, \ldots$ and so on.
- Once the edge reaches vtx k: it attaches to k w.p. $\propto$ degree of k, else lands in $[k-1]$ w.p. $\propto$ total degree of $[k-1]$.
- That is a draw from a *classical* Pólya urn – red $=$ vtx k, black $=$ a vertex in $[k-1]$ – started at

$$r = m, \qquad b = (2k-3)m,$$

whose a.s. limit of the fraction of reds is $\mathrm{Beta}(m, (2k-3)m)$.

A stochastic recursion

- Assume $m = 2$.

- D_n: The subgraph consisting of the descendants and the edges joining them.
  Y_k: # edges in D_n joining $[k]$ and $\{k+1, \dots, n\}$.
  Note that $Y_{n-1} = 2$.
  Z_k: # edges in Y_k that hit vtx k.

- Expose D_n backwards, from vtx n down to 1:

A stochastic recursion

First sample $(B_k)_{k=1}^n$; start with two 'incomplete' edges from n.
When at vtx $k \in \{n-1, \dots, 2\}$, try to attach each of the incomplete edges in Y_k to k w.p. B_k.
Add vertex k and two incomplete edges (from k) to D_n if at least one edge in Y_k hits k.

An example

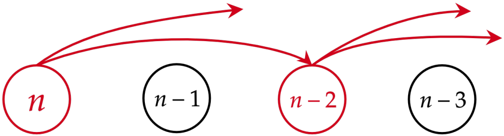


Figure: $Y_{n-1} = Y_{n-2} = 2,\ Y_{n-3} = 3$ and $Z_{n-2} = 1$.

Connection to the number of descendants

Observe that

$$\mathbb{1}[Z_k \geq 1] = \mathbb{1}[\text{vtx } k \text{ is a descendant}].$$

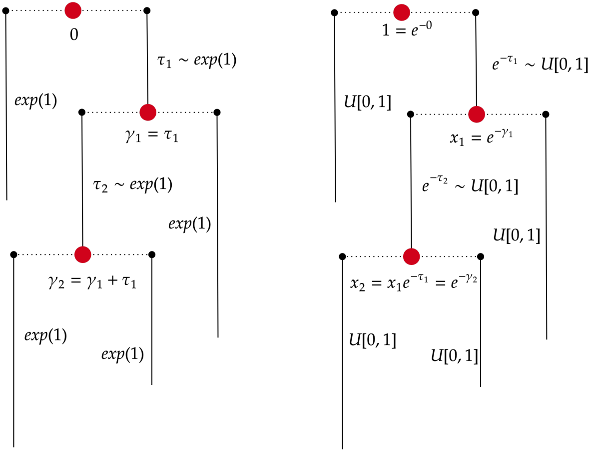
Thus, counting vtx n separately,

$$X^{(n)} := \#\text{descendants of vtx } n = 1 + \sum_{k=1}^{n-1} \mathbb{1}[Z_k \geq 1].$$

We study Y_k in three phases: ‘early’, ‘middle’ and ‘late’.

The early part of D_n and a Yule process

We couple the early phase to a Yule process. Assume still $m = 2$.
Start with a Yule process that has 2 particles at time 0.
Apply the time change $t \to e^{-t}$:



The early part of D_n and a Yule process ctd.

- 'Early' vtx in D_n: those in D_n with labels $k \geq n_1 = n/\log n$.
- Y_k: $\#$ edges in D_n joining the vtx sets $[k]$ and $\{k+1, \ldots, n\}$.
- $\widehat{D}_n$: D_n with original labels $k \in [n]$ replaced with $(k/n)^\chi$.
  $(\chi = m/(2m+\rho)$, with $\chi = 1/2$ when $\rho = 0)$
- $\widehat{\mathcal{Y}}$: The time-changed Yule tree, taking splitting times as vtx labels.
- $\widehat{\mathcal{Y}}_t$: $\#$ edges in $\widehat{\mathcal{Y}}$ alive at time t.

The early part of D_n and the Yule process

(Informal) Thm: Yule process approximation

We can couple $\widehat{D}_n$ and $\widehat{\mathcal{Y}}$ such that when restricted to vtx with labels in $[(n_1/n)^\chi, 1]$ and the edges starting from them, w.h.p.

1. they have the same shape;
2. the perturbation in vertex labels are 'small' enough;
3. $Y_{n_1} = \widehat{\mathcal{Y}}_{(n_1/n)^\chi}$.

Idea of proof:

The early part of D_n is w.h.p. tree-like.

The *Pólya urn representation* tells us how we should rescale and couple the vtx labels of D_n and $\widehat{\mathcal{Y}}$.

The gamma distribution

- $\mathcal{Y}_t$: # living particles in the original Yule process at time t.
- By standard properties of the Yule process, for $x = e^{-t} \in [0, 1]$,

$$x\widehat{\mathcal{Y}}_x = e^{-t}\mathcal{Y}_t \xrightarrow{a.s.} \xi, \quad x \to 0,$$

  where $\xi \sim \mathrm{Gamma}(2, 1)$.
- Thus, by the previous theorem,

$$\left(\frac{n_1}{n}\right)^{\chi} Y_{n_1} \xrightarrow{p} \xi.$$

Middle and late phases: a supermartingale

Recall the stochastic recursion where at each vtx k, we attach each of the incomplete edges in Y_k in D_n to k w.p. B_k. Then,

$$Y_{k-1} = Y_k - Z_k + 2 \cdot \mathbb{1}[Z_k \geq 1], \quad 1 \leq k \leq n-1.$$

Let $\mathcal{F}_k$ be the σ-algebra generated by the betas $(B_i)_2^n$ and the coin tosses at vtx $k+1, \ldots, n-1$. Also,

$$Z_k \mid \mathcal{F}_k \sim \mathrm{Bin}(Y_k, B_k).$$

So by Markov inequality,

$$\mathbb{E}[Y_{k-1} \mid \mathcal{F}_k] \leq (1 + B_k) Y_k.$$

So we may define the reverse supermartingale

$$W_k = Y_k \prod_{i=1}^{k} (1 + B_i), \quad 0 \leq k \leq n-1.$$

Three phases of the descent

We study W_k in three phases (assuming $m = 2$ and $\rho = 0$):

- Early $k \geq n_1 = n/\log n$:
  $n^{-1/2} W_{n_1} \xrightarrow{d} \tilde{\beta} \cdot \xi$, where $\tilde{\beta}$ is independent of $\xi \sim \mathrm{Gamma}(2, 1)$;
- Flat middle $n_2 \leq k \leq n_1$ $(n_2 \ll n^{1/3})$:
  here $n^{-1/2} W_k$ *barely moves*.
- Late $k \sim n^{1/3}$: where *most* descendants live. For the interpolated processes to real arguments $t \in [0, \infty)$,

$$n^{-1/2} W_{tn^{1/3}} \xrightarrow{d} 4\tilde{\beta}((t^{9/2} + \tfrac{3}{4}\xi t^3)^{1/3} - t^{3/2}) \quad \text{in } C[0, \infty),$$

$$n^{-1/3} Y_{tn^{1/3}} \xrightarrow{d} 4((t^3 + \tfrac{3}{4}\xi t^{3/2})^{1/3} - t) \quad \text{in } C(0, \infty).$$

Returning to the number of descendants

By Doob's decomposition,

$$X^{(n)} = 1 + \sum_{k=1}^{n-1} \mathbb{1}[Z_k \geq 1] = 1 + L_0 + P_0,$$

where for $0 \leq k \leq n-1$,

$$L_k = \sum_{i=k+1}^{n-1} (\mathbb{1}[Z_i \geq 1] - \mathbb{P}(Z_i \geq 1 \mid \mathcal{F}_i))$$

is a reverse martingale and

$$P_k = \sum_{i=k+1}^{n-1} \mathbb{P}(Z_i \geq 1 \mid \mathcal{F}_i) = \sum_{i=k+1}^{n-1} \left(1 - (1 - B_i)^{Y_i}\right).$$

The number of descendants

By bounding $\mathbb{E}L_0^2$ via estimating $\mathbb{P}(Z_k \geq 1)$, one can show

$$n^{-1/3} L_0 \xrightarrow{p} 0.$$

By interpolating the process P_k and applying a similar argument as for $W_{tn^{1/3}}$ and the result for $Y_{tn^{1/3}}$ we reach

$$n^{-1/3} P_0 \xrightarrow{d} 2^{-4/3} 3^{-1/3} \frac{\Gamma(\frac{1}{3})^2}{\Gamma(\frac{2}{3})} \xi^{2/3}.$$

Plugging the above back to $X^{(n)} = 1 + L_0 + P_0$ gives

$$\frac{X^{(n)}}{n^{1/3}} \xrightarrow{d} 2^{-4/3} 3^{-1/3} \frac{\Gamma(\frac{1}{3})^2}{\Gamma(\frac{2}{3})} \xi^{2/3}$$

for $m = 2$ and $\rho = 0$.

Open problems

- Same problem but in different models? E.g. Bernoulli PA graph studied in Dereich and Mörters (2009).
- Other properties of D_n and $G_n \setminus D_n$?
- $|D_n \cap D_{n+1}|$? Recent progress by Fabian Burghart (AofA 2026) for UAM, completely open for PAM.
- $\max\{X^{(n+1)}, ..., X^{(n+i)}\}$ for a fixed $i > 2$ or $i = i(n)$ growing with n at some rate, in both UAM and PAM.

Thank you!