Tail Flexibility in the Degrees of Preferential Attachment Networks

Thomas William Boughen[1], Clement Lee[1],
Vianey Palacios Ramirez[1]

[1]School of Mathematics, Statistics and Physics, Newcastle University.

Abstract

Devising the underlying generating mechanism of a real-life network is difficult as, more often than not, only its snapshots are available, but not its full evolution. One candidate for the generating mechanism is preferential attachment which, in its simplest form, results in a degree distribution that follows the power law. Consequently, the growth of real-life networks that roughly display such power-law behaviour is commonly modelled by preferential attachment. However, the validity of the power law has been challenged by the presence of alternatives with comparable performance, as well as the recent findings that the right tail of the degree distribution is often lighter than implied by the body, whilst still being heavy. In this paper, we study a modified version of the model with a flexible preference function that allows super/sub-linear behaviour whilst also guaranteeing that the limiting degree distribution has a heavy tail. We relate the distributions tail heaviness directly to the model parameters, allowing direct inference of the parameters from the degree distribution alone.

Keywords: networks, extremes

1 Introduction

The degree distribution of a network can be very informative about the networks structure, giving some understanding about the inequality in the influence of the nodes. Often the degree distribution is one of the simplest places to look when trying to understand this, as the full evolution of the network is usually not available. This makes modelling the degree distribution an important topic in research with contributions such as [ref].

When looking at the degree distributions of real networks, a power law is an attractive option to use when modelling as it is seemingly observed in a lot of networks. Fitting the discrete power law to the degree distribution come with the added benefit of suggesting preferential attachment as the generating mechanism since it has been shown to generate networks with a power law degree distribution.

If the power law is an adequate model to use for real degree distributions has become a heated debate, with alternatives such as (Broido and Clauset, 2019) seeming to outperform the power law in some cases. Additionally, [ref] find that for many real degree distributions the body may follow a power law, the right tail does not or at least seems to be lighter than is implied by the body whilst still being heavy. One method for modelling the tail of the degree distributions of real networks is to use a discretised variation of the generalised Pareto distribution (GPD) usually dubbed the Integer GPD (GPD) where $X|X > v \sim \mathrm{IGPD}(\xi, \sigma, v)$ is defined by the survival:

$$\Pr(X > x|X > v) = \left(\frac{\xi(x-v)}{\sigma} + 1\right)^{-1/\xi}, \qquad x = v+1, v+2, \ldots$$

for $v \in \mathbb{Z}^+, \sigma > 0, \xi \in \mathbb{R}$.

Shimura (2012) introduces a quantity that will help in determining what domain of attraction a discrete distribution belongs to, that is:

For a distribution F with survival function $\bar{F}$ and some $n \in \mathbb{Z}^+$ let:

$$\Omega(F, n) = \left(\log \frac{\bar{F}(n+1)}{\bar{F}(n+2)}\right)^{-1} - \left(\log \frac{\bar{F}(n)}{\bar{F}(n+1)}\right)^{-1}$$

(Shimura, 2012) then states that if $\lim_{n\to\infty} \Omega(F, n) = 1/\alpha$ ($\alpha > 0$), then F is heavy tailed with $\bar{F}(n) \sim n^{-\alpha}$. Additionally, if $\lim_{n\to\infty} \Omega(F, n) = 0$ then the distribution is light tailed.

These models that seem to outperform the power law do not however come with the added benefit of suggesting a generating mechanism for the networks in the same way the power law does. Efforts to make the preferential attachment model more flexible have not yet remedied this issue with (Krapivsky and Redner, 2001) finding that using a sub/super-linear preference function yields a Weibull tail and a degenerate network, where one vertex eventually gains all new edges, respectively.

(Rudas et al, 2007) introduces theoretical results for a PA model with a general preference function in the cases that the network being generated is a tree; presenting an opportunity to design a preference function that will capture the behaviours observed in the degree distributions of real networks. This paper studies a flexible class of preference functions that allow for a power law body and a tail that is heavy and flexible, and proposes a strategy for fitting this model to real data. With good fits, this method does not only allow the tail heaviness to be quantified, but it also allows for some inference on the preference function assuming a general preferential attachment model is the underlying generating mechanism.

2 General Preferential Attachment

The focus of this section is the limiting degree distribution of a preferential attachment network with preference function $b(\cdot)$ where each vertex brings with it m edges. This preference function is subject to the following conditions:

$$b : \mathbb{N} \mapsto \mathbb{R}^+ \setminus \{0\},$$

$$\sum_{k=0}^{\infty} \frac{1}{b(k)} = \infty. \tag{1}$$

Given these two conditions an expression for the survival of the limiting degree distribution can be found in the case that $m = 1$; obtained by considering a branching process that is equivalent to the growth of the network, as in (Rudas et al, 2007). Theorem 1 from (Rudas et al, 2007) states that for the tree $\Upsilon(t)$ at time t:

$$\lim_{t\to\infty} \frac{1}{|\Upsilon(t)|} \sum_{x\in\Upsilon(t)} \varphi(\Upsilon(t)_{\downarrow x}) = \lambda^* \int_0^{\infty} e^{-\lambda^* t} \mathbb{E}\left[\varphi(\Upsilon(t))\right] dt$$

where λ^* satisfies $\hat{\rho}(\lambda^*) = 1$.

The limiting survival can be viewed as the limit of the empirical proportion of vertices with degree over a threshold k, that is:

$$\bar{F}(k) = \lim_{t\to\infty} \frac{\sum_{x\in\Upsilon(t)} \mathbb{I}\left\{\deg(x, \Upsilon(t)_{\downarrow x}) > k\right\}}{\sum_{x\in\Upsilon(t)} 1}$$

which by the previously stated theorem can also be written as:

$$\bar{F}(k) = \frac{\int_0^{\infty} e^{-\lambda^* t} \mathbb{E}\left[\mathbb{I}\left\{\deg(x, \Upsilon(t)) > k\right\}\right] dt}{\int_0^{\infty} e^{-\lambda^* t} dt} = \prod_{i=0}^{k} \frac{b(i)}{\lambda^* + b(i)}$$

Additionally, using the fact that $f(k) = \bar{F}(k-1) - \bar{F}(k)$ the probability mass function (p.m.f.) can be shown to be:

$$f(k) = \frac{\lambda^*}{\lambda^* + b(k)} \prod_{i=0}^{k-1} \frac{b(i)}{\lambda^* + b(i)}, \qquad k \in \mathbb{N}.$$

which aligns with the result from (Rudas et al, 2007), where the expression is derived using the same branching process results.

Using results from (Shimura, 2012) the domain of attraction of this limiting degree distribution can be found.

Proposition 2.1. *If* $b(k) \to \infty$ *as* $k \to \infty$ *then*

$$\bar{F}(k) = \prod_{i=0}^{k} \frac{b(i)}{\lambda^* + b(i)}$$

and

$$\lim_{k\to\infty} \Omega(F, n) = \lim_{k\to\infty} \frac{b(k+1) - b(k)}{\lambda^*}.$$

Proposition 2.1 implies that the only way to obtain a heavy tailed degree distribution using a non-decreasing preference function is to have that preference function be asymptotically linear. The next subsection focuses on a particular class of these preference functions that provide flexible behaviour but guarantee a heavy tail.

2.1 Preferential Attachment with flexible heavy tail

Previously, in (Krapivsky and Redner, 2001), a preference function of the form $k^\alpha + \varepsilon$ has been considered. However, it has been noted that super-linear preferential attachment ($\alpha > 1$) leads to a degenerate limiting degree distribution where at some point one vertex gains the connection from every vertex that joins the network, additionally sub-linear preferential attachment ($\alpha < 1$) has been shown to lead to a light tailed limiting degree distribution. Consider a preference function of the form:

$$b(k) = \begin{cases} k^\alpha + \varepsilon, & k < k_0 \\ k_0^\alpha + \varepsilon + \beta(k - k_0), & k \geq k_0 \end{cases}$$

for $\alpha, \beta, \varepsilon > 0$ and $k_0 \in \mathbb{N}$.

Using a preference function with guaranteed linear behaviour in the limit, allows for the inclusion of sub/super linear behaviour without losing the heavy tails or ending up with a degenerate degree distribution.

The limiting degree distribution resulting from using a preference function of this form can be found to have survival:

$$\bar{F}(k) = \begin{cases} \prod_{i=0}^{k} \frac{i^\alpha+\varepsilon}{\lambda^*+i^\alpha+\varepsilon}, & k < k_0 \\ \left(\prod_{i=0}^{k_0-1} \frac{i^\alpha+\varepsilon}{\lambda^*+i^\alpha+\varepsilon}\right) \frac{\Gamma(\lambda^*+k_0^\alpha+\varepsilon)/\beta)}{\Gamma((k_0^\alpha+\varepsilon)/\beta)} \frac{\Gamma\left(k-k_0+1+\frac{k_0^\alpha+\varepsilon}{\beta}\right)}{\Gamma\left(k-k_0+1+\frac{\lambda^*+k_0^\alpha+\varepsilon}{\beta}\right)}, & k \geq k_0. \end{cases} \quad (2)$$

with λ^* satisfying $\widehat{\rho}(\lambda^*) = 1$ where:

$$\hat{\rho}(\lambda) = \sum_{n=0}^{k_0} \prod_{i=0}^{n-1} \frac{i^\alpha + \varepsilon}{\lambda + i^\alpha + \varepsilon} + \left(\frac{k_0^\alpha + \varepsilon}{\lambda - \beta}\right) \prod_{i=0}^{k_0-1} \frac{i^\alpha + \varepsilon}{\lambda + i^\alpha + \varepsilon}$$

which must be solved numerically for most parameter choices. Also, note that $\lambda > \beta$ since when $\lambda \leq \beta$ the infinite sum no longer converges and instead goes to infinity.

Some examples of what the degree distribution looks like are shown below in Figure 1:

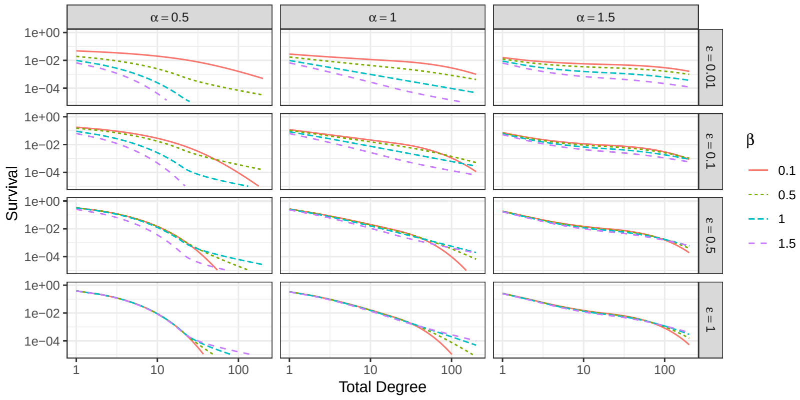

Figure 1: Theoretical survival distributions of the limiting degree distributions, according to various combinations of $(\alpha, \beta, \varepsilon)$ and $k_0 = 20$ of the proposed preferential attachment model.

The survival (Equation 2) can be connected to the IGPD mentioned in Section 1 by using Sterling's approximation of the gamma function to obtain:

$$\bar{F}(k) \begin{cases} = \prod_{i=0}^{k} \frac{i^\alpha+\varepsilon}{\lambda^*+i^\alpha+\varepsilon}, & k < k_0, \\ \approx \left(\prod_{i=0}^{k_0-1} \frac{i^\alpha+\varepsilon}{\lambda^*+i^\alpha+\varepsilon}\right) \left(\frac{\beta(k+1-k_0)}{k_0^\alpha+\varepsilon} + 1\right)^{-\lambda^*/\beta}, & k \geq k_0, \end{cases}$$

meaning that for $k \geq k_0$ the limiting degree distribution is similar to IGPD $\left(\frac{\beta}{\lambda^*}, \frac{k_0^\alpha+\varepsilon}{\lambda^*}, k_0 - 1\right)$.

To assess how close of an approximation this is the theoretical conditional survivals are shown in Figure 2 in colour and their IGPD approximations are shown in grey. The approximation seems to hold up fairly well even for large degrees.

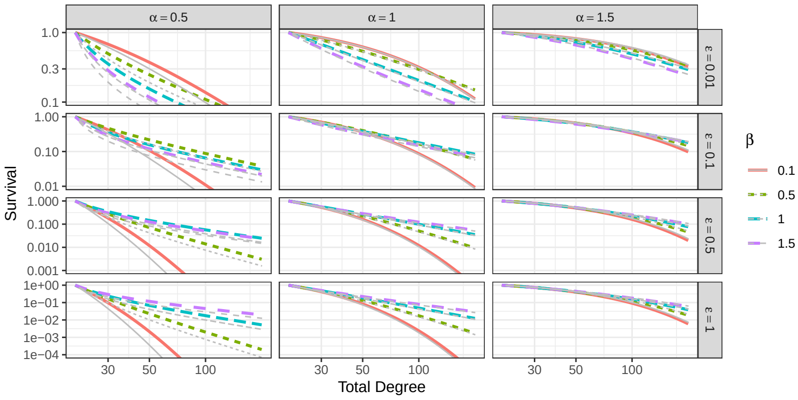

Figure 2: Theoretical conditional survivals (grey) alongside their IGPD approximations (coloured).

Since $\beta > 0$, the shape parameter of the IGPD is positive and thus the distribution is heavy tailed. Additionally the value of the shape parameter ξ is shown in Figure 3 for various parameter choices:

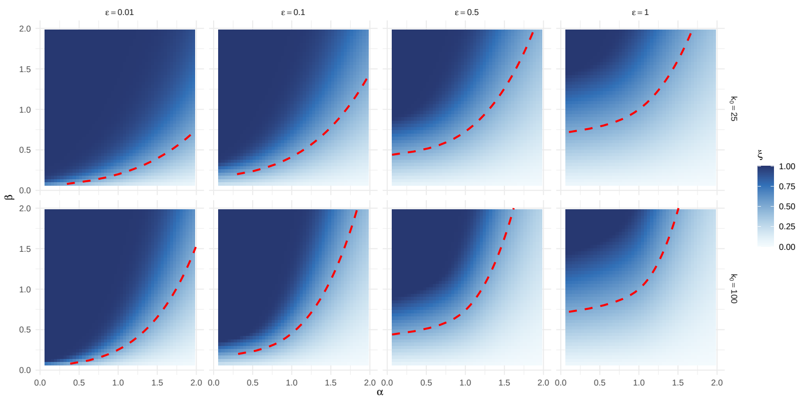

Figure 3: Heat maps of ξ for various combinations of the parameters of the proposed model.

The darker regions on the heat maps correspond to a heavier tail and the lighter to a lighter tail, the red dashed line shows combinations of α and β that produce a limiting degree distribution with the same tail heaviness as the Barabási-Albert model, $\xi = 0.5$.

2.1.1 Piecewise Linear Model

When $\alpha = 1$ the preference function simplifies to a piecewise linear function with β representing the ratio of the gradients of the two components:

$$b(k) = \min(k, k_0) + \varepsilon + \mathbb{I}\{k \geq k_0\}\beta(k - k_0)$$

for some $\beta, \varepsilon, k_0 > 0$.

Following Equation 2 the limiting degree distribution is:

$$\bar{F}(k) = \begin{cases} \frac{\Gamma(\lambda^*+\varepsilon)\Gamma(k+1+\varepsilon)}{\Gamma(\varepsilon)\Gamma(k+1+\lambda^*+\epsilon)}, & k < k_0 \\ \frac{\Gamma(\lambda^*+\varepsilon)\Gamma(k_0+\varepsilon)}{\Gamma(\varepsilon)\Gamma(k_0+\lambda^*+\varepsilon)} \times \frac{\Gamma\left(\frac{k_0+\varepsilon+\lambda^*}{\beta}\right)}{\Gamma\left(\frac{k_0+\varepsilon}{\beta}\right)} \times \frac{\Gamma\left(k-k_0+\frac{k_0+\varepsilon}{\beta}+1\right)}{\Gamma\left(k-k_0+\frac{k_0+\varepsilon+\lambda^*}{\beta}+1\right)}, & k \geq k_0 \end{cases} \tag{3}$$

Just like the previous model, the survival can be approximated by:

$$\bar{F}(k) \approx \begin{cases} \frac{\Gamma(\lambda^*+\varepsilon)}{\Gamma(\varepsilon)}(1+\varepsilon)^{\lambda^*}\left(\frac{k}{1+\varepsilon}+1\right)^{-\lambda^*}, & k < k_0 \\ \frac{\Gamma(\lambda^*+\varepsilon)\Gamma(k_0+\varepsilon)}{\Gamma(\varepsilon)\Gamma(k_0+\lambda^*+\varepsilon)} \times \left(\frac{\beta(k+1-k_0)}{k_0+\varepsilon}+1\right)^{-\lambda^*/\beta}, & k \geq k_0 \end{cases}$$

Some examples of the possible degree distributions are shown in the central row of Figure 1.

3 Recovery

The goal of was to provide a possible preference function that is able to explain the growth of real networks with varied tail behaviour in their degree distributions. This section aims to show that the parameters of the model in Section 2.1 can be recovered from the degree distribution of a network simulated from it.

The procedure for recovering the parameters begins with simulating a network from the model with $N = 100,000$ vertices and $m = 1$ given some set of parameters $\boldsymbol{\theta} = (\alpha, \beta, \varepsilon, k_0)$, obtaining the degree counts in the form of a vector of degrees $\boldsymbol{x} = (1, 2, \dots, M)$ and the number of vertices with those degrees $\boldsymbol{n}$ where M is the maximum degree. However, when it comes to fitting this model to real data in the next section, small degrees are extremely influential and will often cause the model to have a bad fit even though the rest of the data looks similar to that of the simulations; truncating the data above some value, say $l < k_0$, should allow the model to be more effectively fitted. This accounts in some way for the fact that this model only applies to trees, and could not possibly capture the behaviour for nodes with small degrees. Using the p.m.f derived in (Rudas et al, 2007), the likelihood given that $x_i \geq l$ for all $i = 1, 2, \dots, M$ is:

$$L(\boldsymbol{x},\boldsymbol{n}|\boldsymbol{\theta}) = \left(\frac{\lambda^*}{\lambda^*+\varepsilon}\right)^{n_0} \left(\prod_{j=l}^{k_0-1} \frac{j^\alpha+\varepsilon}{\lambda^*+j^\alpha+\varepsilon}\right)^{\left(\sum_{x_i\geq k_0} n_{x_i}\right)} \prod_{l\leq x_i<k_0} \left(\frac{\lambda^*}{\lambda^*+x_i^\alpha+\varepsilon}\prod_{j=l}^{k_0-1}\frac{j^\alpha+\varepsilon}{\lambda^*+j^\alpha+\varepsilon}\right)^{n_i}$$
$$\times \prod_{x_i\geq k_0}\left(\frac{\mathrm{B}(x_i-k_0+(k_0^\alpha+\varepsilon)/\beta, 1+\lambda^*/\beta)}{\mathrm{B}((k_0^\alpha+\varepsilon)/\beta, \lambda^*/\beta)}\right)^{n_i}$$

where $B(y,z)$ is the the beta function.

This likelihood allows inference to be made about the parameters using a Bayesian approach using the priors:

$$\begin{aligned}
\alpha &\sim \mathrm{Ga}(1, 0.01),\\
\beta &\sim \mathrm{Ga}(1, 0.01),\\
k_0 &\sim \mathrm{U}(1, 10,000),\\
\varepsilon &\sim \text{Ga(1,0.01)},
\end{aligned}$$

to obtain a posterior distribution that can then be used in an adaptive Metropolis-Hastings Markov chain Monte Carlo (MCMC) algorithm to obtain posterior samples. The results of this inference are shown in **?@fig-rec1** and Figure 5. For these simulated networks $l = 0$.

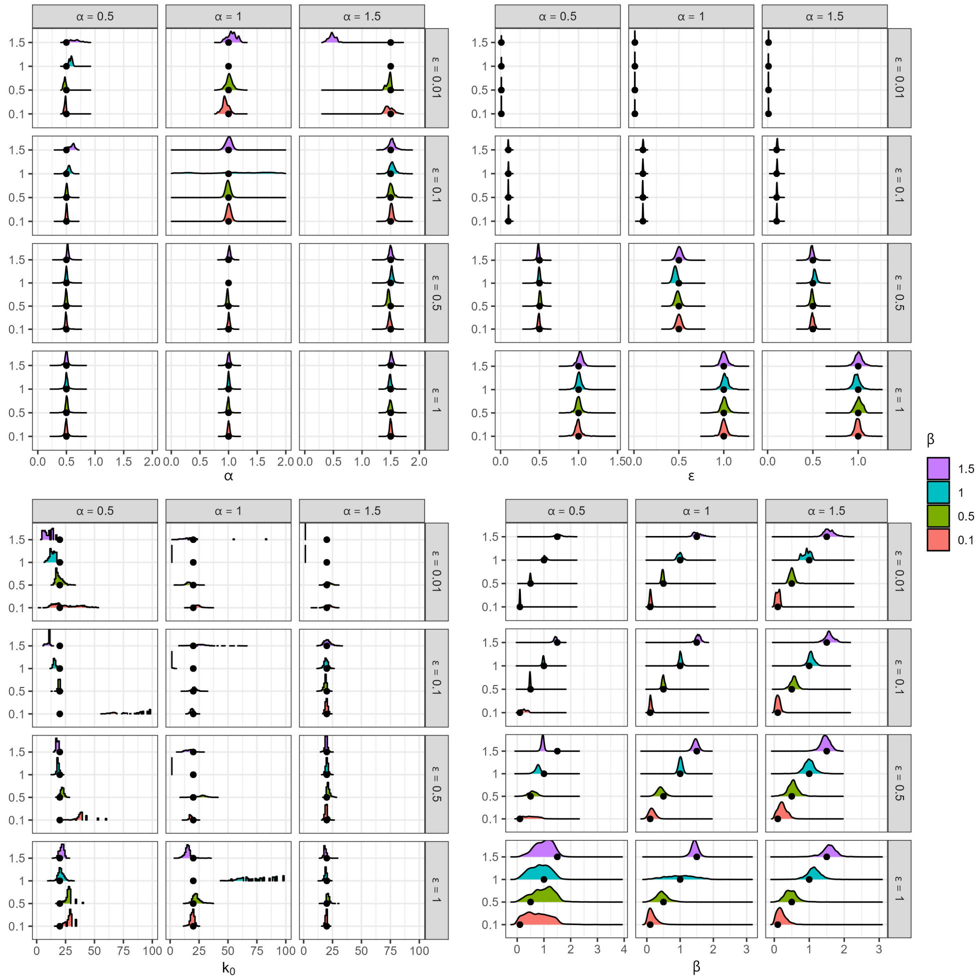

Figure 4: Posterior estimates of paramters for data simulated from the proposed model with various combinations of (α,β,ε) and $k_0 = 20$.

Figure 5

4 Application to Real Data

Turning now to real data, the goal is to fit the model to the degree distributions of real networks from various sources. [describe networks being used]. Alongside fitting this model to the degree distributions, we then compare the fit to that of an mixture distribution that was used in [clement] and is similar to others that have been used to model degree distributions [others].

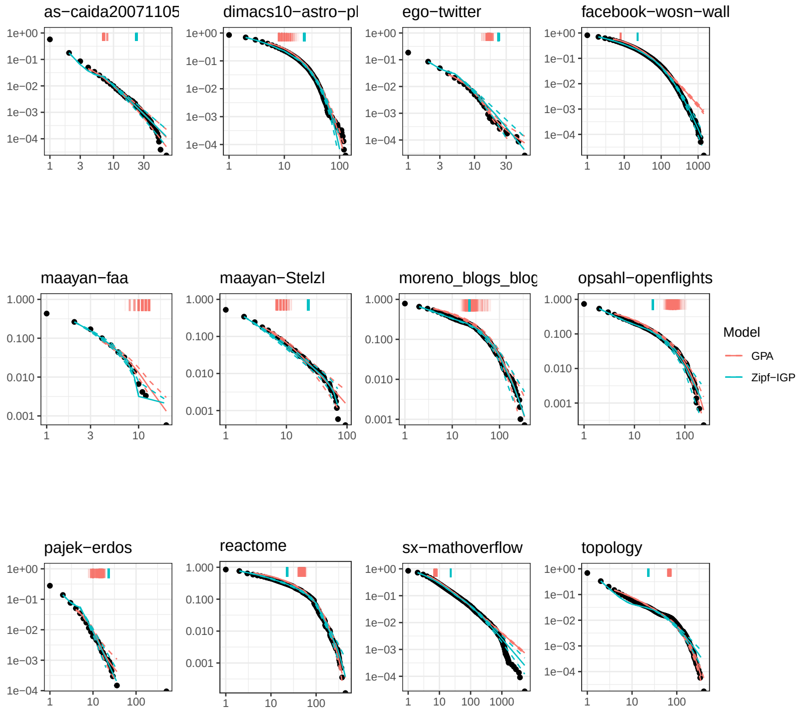

Figure 6: Posterior estimates (solid red) of survival for several real data sets and their 95% credible intervals (dotted red)

Figure 6 displays the posterior estimates of the survival function for various data sets, obtained from fitting the GPA model and the Zipf-IGP mixture model. In most cases, the GPA model does not necessarily provide an improvement in fit when compared to the Zipf-IGP model but where the GPA model fits well we gain additional information about the preference function assuming that the network evolved according the the GPA scheme. Figure 7 shows the posterior of the shape parameter ξ obtained from the Zipf-IGP model alongside the posterior of the equivalent shape parameter β/λ^* obtained from fitting the GPA model.

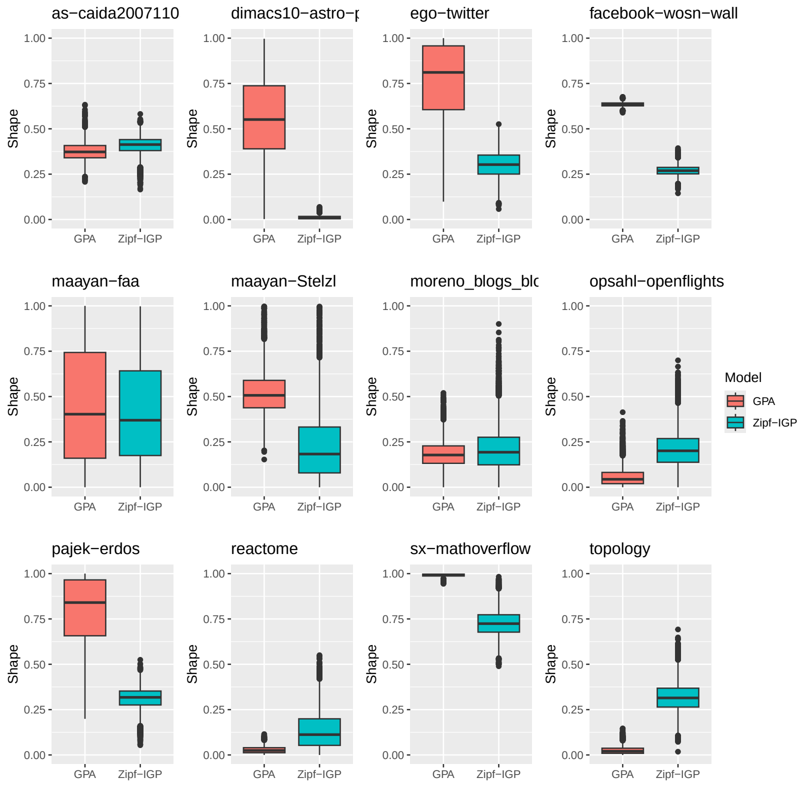

Figure 7: Posterior estimates (solid red) of survival for several real data sets and their 95% credible intervals (dotted red)

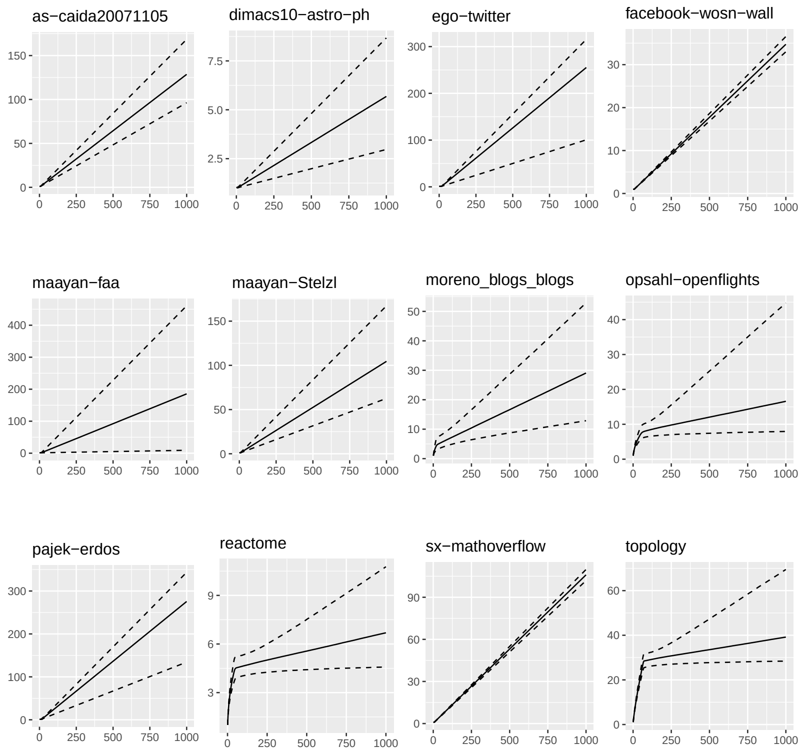
as−caida20071105
dimacs10−astro−ph
ego−twitter
facebook−wosn−wall
maayan−faa
maayan−Stelzl
moreno_blogs_blogs
opsahl−openflights
pajek−erdos
reactome
sx−mathoverflow
topology

Figure 8: df

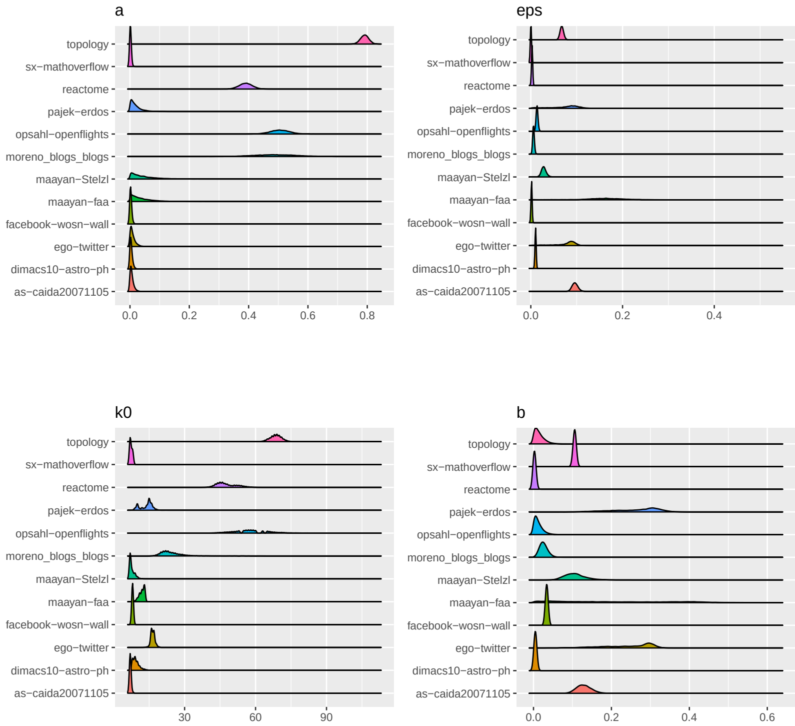

Figure 9: df

5 Conclusion and Discussion

6 Additional Results

6.1 Tail heaviness of general preferential attachment model

Recall the limiting survival function:

$$\bar{F}(k) = \prod_{i=0}^{k} \frac{b(i)}{\lambda^* + b(i)},$$

The aim is to determine how this distribution behaves for different choices of b, specifically what maximum domain of attraction does this distribution belong to and is it affected by the choice of preference function b. Shimura (2012) introduces a quantity that will help in determining what domain of attraction a discrete distribution belongs to, that is:

For a distribution F with survival function $\bar{F}$ and some $n \in \mathbb{Z}^+$ let:

$$\Omega(F,n) = \left(\log \frac{\bar{F}(n+1)}{\bar{F}(n+2)}\right)^{-1} - \left(\log \frac{\bar{F}(n)}{\bar{F}(n+1)}\right)^{-1}$$

(Shimura, 2012) then states that if $\lim_{n\to\infty} \Omega(F,n) = 1/\alpha$ ($\alpha > 0$), then F is heavy tailed with $\bar{F}(n) \sim n^{-\alpha}$. Additionally, if $\lim_{n\to\infty} \Omega(F,n) = 0$ then the distribution is light tailed.

Consider a preference function $b(\cdot)$ that fulfills:

$$\lim_{n\to\infty} b(n) = \infty \tag{4}$$

Substituting in the form of $\bar{F}(n)$ from **?@eq-surv** and taking the limit, subject to Equation 4:

$$\lim_{n\to\infty} \Omega(F,n) = \lim_{n\to\infty} \frac{b(n+2) - b(n+1)}{\lambda^*}. \tag{5}$$

(see appendix for full proof).

Whilst there are many classes of functions that satisfy Equation 4 and Equation 1, the vast majority of these result in $\Omega(F,n) \to 0$, meaning that the limiting degree distribution is light tailed. In fact, in order for the limiting degree distribution to be heavy tailed, b must be asymptotically linear ($b(k) \sim k$).

7 Proofs

7.1 Tail heaviness of GPA

Taking the form of the GPA degree survival:

$$\bar{F}(n) = \prod_{i=0}^{n} \frac{b(i)}{\lambda + b(i)}$$

and substituting into the formula for $\Omega(F,n)$:

$$\begin{aligned}
\Omega(F,n) &= \left(\log \frac{\prod_{i=0}^{n+1} \frac{b(i)}{\lambda+b(i)}}{\prod_{i=0}^{n+2} \frac{b(i)}{\lambda+b(i)}}\right)^{-1} - \left(\log \frac{\prod_{i=0}^{n} \frac{b(i)}{\lambda+b(i)}}{\prod_{i=0}^{n+1} \frac{b(i)}{\lambda+b(i)}}\right)^{-1} \\
&= \left(\log \frac{\lambda + b(n+2)}{b(n+2)}\right)^{-1} - \left(\log \frac{\lambda + b(n+1)}{b(n+1)}\right)^{-1} \\
&= \left(\log \left[1 + \frac{\lambda}{b(n+2)}\right]\right)^{-1} - \left(\log \left[1 + \frac{\lambda}{b(n+1)}\right]\right)^{-1}
\end{aligned}$$

Clearly if $b(n) = c$ or $\lim_{n\to\infty} b(n) = c$ for some $c > 0$ then $\Omega(F, n) = 0$. Now consider a non-constant $b(n)$ and re-write $\Omega(F, n)$ as:

$$
\begin{aligned}
\Omega(F,n) &= \left(\log\left[1+\frac{\lambda}{b(n+2)}\right]\right)^{-1} - \frac{b(n+2)}{\lambda} + \frac{b(n+2)}{\lambda} - \left(\log\left[1+\frac{\lambda}{b(n+1)}\right]\right)^{-1} + \frac{b(n+1)}{\lambda} - \frac{b(n+1)}{\lambda} \\
&= \left\{\left(\log\left[1+\frac{\lambda}{b(n+2)}\right]\right)^{-1} - \frac{b(n+2)}{\lambda}\right\} - \left\{\left(\log\left[1+\frac{\lambda}{b(n+1)}\right]\right)^{-1} - \frac{b(n+1)}{\lambda}\right\} + \frac{b(n+2)}{\lambda} - \frac{b(n+}{\lambda}
\end{aligned}
$$

Then if $\lim_{n\to\infty} b(n) = \infty$ it follows that:

$$
\begin{aligned}
\lim_{n\to\infty} \Omega(F,n) &= \lim_{n\to\infty} \left\{\left(\log\left[1+\frac{\lambda}{b(n+2)}\right]\right)^{-1} - \frac{b(n+2)}{\lambda}\right\} - \lim_{n\to\infty} \left\{\left(\log\left[1+\frac{\lambda}{b(n+1)}\right]\right)^{-1} - \frac{b(n+1)}{\lambda}\right\} \\
&\quad + \lim_{n\to\infty}\left(\frac{b(n+2)}{\lambda} - \frac{b(n+1)}{\lambda}\right) \\
&= \frac{1}{2} - \frac{1}{2} + \lim_{n\to\infty}\left(\frac{b(n+2)}{\lambda} - \frac{b(n+1)}{\lambda}\right) \\
&= \frac{1}{\lambda}\lim_{n\to\infty}\left[b(n+2) - b(n+1)\right] \qquad \square
\end{aligned}
$$

7.2 Removing the infinite sum

For a preference function of the form:

$$b(k) = \begin{cases} g(k), & k < k_0 \\ g(k_0) + \beta(k - k_0), & k \geq k_0 \end{cases}$$

for $\beta > 0, k_0 \in \mathbb{N}$ we have that

$$\begin{aligned}
\hat{\rho}(\lambda) = \sum_{n=0}^{\infty} \prod_{i=0}^{n-1} \frac{b(i)}{\lambda + b(i)} &= \sum_{n=0}^{k_0} \prod_{i=0}^{n-1} \frac{g(i)}{\lambda + g(i)} + \sum_{n=k_0+1}^{\infty} \left(\prod_{i=0}^{k_0-1} \frac{g(i)}{\lambda + g(i)} \prod_{i=k_0}^{n-1} \frac{g(k_0) + \beta(i - k_0)}{\lambda + g(k_0) + \beta(i - k_0)} \right) \\
&= \sum_{n=0}^{k_0} \prod_{i=0}^{n-1} \frac{g(i)}{\lambda + g(i)} + \left(\prod_{i=0}^{k_0-1} \frac{g(i)}{\lambda + g(i)} \right) \sum_{n=k_0+1}^{\infty} \prod_{i=k_0}^{n-1} \frac{g(k_0) + \beta(i - k_0)}{\lambda + g(k_0) + \beta(i - k_0)}
\end{aligned}$$

Now using the fact that:

$$\prod_{i=0}^{n} (x + yi) = x^{n+1} \frac{\Gamma(\frac{x}{y} + n + 1)}{\Gamma(\frac{x}{y})}$$

and reindexing the product in the second sum

$$\begin{aligned}
\hat{\rho}(\lambda) &= \sum_{n=0}^{k_0} \prod_{i=0}^{n-1} \frac{g(i)}{\lambda + g(i)} + \left(\prod_{i=0}^{k_0-1} \frac{g(i)}{\lambda + g(i)} \right) \sum_{n=k_0+1}^{\infty} \frac{\Gamma\left(\frac{g(k_0)}{\beta} + n - k_0\right) \Gamma\left(\frac{\lambda + g(k_0)}{\beta}\right)}{\Gamma\left(\frac{\lambda + g(k_0)}{\beta} + n - k_0\right) \Gamma\left(\frac{g(k_0)}{\beta}\right)} \\
&= \sum_{n=0}^{k_0} \prod_{i=0}^{n-1} \frac{g(i)}{\lambda + g(i)} + \frac{\Gamma\left(\frac{\lambda + g(k_0)}{\beta}\right)}{\Gamma\left(\frac{g(k_0)}{\beta}\right)} \left(\prod_{i=0}^{k_0-1} \frac{g(i)}{\lambda + g(i)} \right) \sum_{n=k_0+1}^{\infty} \frac{\Gamma\left(\frac{g(k_0)}{\beta} + n - k_0\right)}{\Gamma\left(\frac{\lambda + g(k_0)}{\beta} + n - k_0\right)} \\
&= \sum_{n=0}^{k_0} \prod_{i=0}^{n-1} \frac{g(i)}{\lambda + g(i)} + \frac{\Gamma\left(\frac{\lambda + g(k_0)}{\beta}\right)}{\Gamma\left(\frac{g(k_0)}{\beta}\right)} \left(\prod_{i=0}^{k_0-1} \frac{g(i)}{\lambda + g(i)} \right) \sum_{n=1}^{\infty} \frac{\Gamma\left(\frac{g(k_0)}{\beta} + n\right)}{\Gamma\left(\frac{\lambda + g(k_0)}{\beta} + n\right)}
\end{aligned}$$

In order to simplify the infinite sum, consider:

$$\begin{aligned}
\sum_{n=0}^{\infty} \frac{\Gamma(n + x)}{\Gamma(n + x + y)} &= \frac{1}{\Gamma(y)} \sum_{n=0}^{\infty} \mathrm{y}(n + x, y) \\
&= \frac{1}{\Gamma(y)} \sum_{n=0}^{\infty} \int_0^1 t^{n+x-1} (1 - t)^{y-1} at
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{\Gamma(y)} \int_0^1 t^{x-1}(1-t)^{y-1} \sum_{n=0}^{\infty} t^n \, at \\
&= \frac{1}{\Gamma(y)} \int_0^1 t^{x-1}(1-t)^{y-1} \frac{1}{1-t} at \\
&= \frac{1}{\Gamma(y)} \int_0^1 t^{x-1}(1-t)^{y-2} at \\
&= \frac{1}{\Gamma(y)} \mathrm{y}(x, y-1) \\
&= \frac{\Gamma(x)}{(y-1)\Gamma(x+y-1)}
\end{aligned}
$$

this infinite sum does not converge when $x \leq 1$ as each term is $O(n^{-x})$. We can now use this in $\hat{\rho}(\lambda)$:

$$
\begin{aligned}
\hat{\rho}(\lambda) &= \sum_{n=0}^{k_0} \prod_{i=0}^{n-1} \frac{g(i)}{\lambda + g(i)} + \frac{\Gamma\left(\frac{\lambda + g(k_0)}{\beta}\right)}{\Gamma\left(\frac{g(k_0)}{\beta}\right)} \left(\prod_{i=0}^{k_0-1} \frac{g(i)}{\lambda + g(i)} \right) \left(\frac{\Gamma\left(\frac{g(k_0)}{\beta}\right)}{\left(\frac{\lambda}{\beta} - 1\right) \Gamma\left(\frac{g(k_0)+\lambda}{\beta} - 1\right)} - \frac{\Gamma\left(\frac{g(k_0)}{\beta}\right)}{\Gamma\left(\frac{g(k_0)+\lambda}{\beta}\right)} \right) \\
&= \sum_{n=0}^{k_0} \prod_{i=0}^{n-1} \frac{g(i)}{\lambda + g(i)} + \left(\prod_{i=0}^{k_0-1} \frac{g(i)}{\lambda + g(i)} \right) \left(\frac{\Gamma\left(\frac{g(k_0)+\lambda}{\beta}\right)}{\left(\frac{\lambda}{\beta} - 1\right) \Gamma\left(\frac{g(k_0)+\lambda}{\beta} - 1\right)} - 1 \right) \\
&= \sum_{n=0}^{k_0} \prod_{i=0}^{n-1} \frac{g(i)}{\lambda + g(i)} + \left(\prod_{i=0}^{k_0-1} \frac{g(i)}{\lambda + g(i)} \right) \left(\frac{\frac{g(k_0)+\lambda}{\beta} - 1}{\frac{\lambda}{\beta} - 1} - 1 \right) \\
&= \sum_{n=0}^{k_0} \prod_{i=0}^{n-1} \frac{g(i)}{\lambda + g(i)} + \left(\prod_{i=0}^{k_0-1} \frac{g(i)}{\lambda + g(i)} \right) \left(\frac{g(k_0) + \lambda - \beta}{\lambda - \beta} - 1 \right) \\
&= \sum_{n=0}^{k_0} \prod_{i=0}^{n-1} \frac{g(i)}{\lambda + g(i)} + \frac{g(k_0)}{\lambda - \beta} \prod_{i=0}^{k_0-1} \frac{g(i)}{\lambda + g(i)} \qquad \square
\end{aligned}
$$

8 References

Broido AD, Clauset A (2019) Scale-free networks are rare. Nature Communications 10(1). https://doi.org/10.1038/s41467-019-08746-5, URL http://dx.doi.org/10.1038/s41467-019-08746-5

Krapivsky PL, Redner S (2001) Organization of growing random networks. Phys Rev E 63:066123. https://doi.org/10.1103/PhysRevE.63.066123, URL https://link.aps.org/doi/10.1103/PhysRevE.63.066123

Rudas A, Tóth B, Valkó B (2007) Random trees and general branching processes. Random Structures & Algorithms 31(2):186–202. https://doi.org/https://doi.org/10.

1002/rsa.20137, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/rsa.20137, https://onlinelibrary.wiley.com/doi/pdf/10.1002/rsa.20137

Shimura T (2012) Discretization of distributions in the maximum domain of attraction. Extremes 15:1–19. https://doi.org/10.1007/s10687-011-0137-7