Learning Network Growth Mechanisms from Degree Distributions in 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 index directly to the model parameters, allowing direct inference of the parameters from the degree distribution alone.

Keywords: networks, discrete extremes, power law, preferential attachment

1 Introduction

Networks have become very powerful tools for representing and analysing complex systems, with uses in a large array of fields, from using stochastic block models to detect communities online (Latouche et al, 2011), to using Exponential Random Graph Models (ERGMs) to analyse the global trade network (Setayesh et al, 2022), and even using mechanistic models to investigate the patterns in neural systems (Betzel and Bassett, 2017).

Along with the recent rise of interest in networks, has been a debate amongst those that study them regarding the claims that most real networks are scale-free. This claim is most often equivalent to claiming that the degree distributions of real networks follow a power law, that is the fraction of nodes with degree k is proportional to $k^{-\alpha}$, and therefore has a regularly varying tail with extreme value index $1/\alpha$. Broido and Clauset (2019) provide empirical evidence, using almost one thousand networks, that these scale-free networks do not make up a large majority of real networks. They compare the fits of a power law model against that of several non-scale-free models only finding strong evidence for scale-freeness in four percent and weak evidence in fifty-two percent of the networks. On the other side of this debate is Voitalov et al (2019) who disagrees and claims that these networks are not nearly as rare and only appear to so as a result of an unrealistic expectation of a power law without deviations or noise. Evidence of these deviations from a power law is shown by Lee et al (2024) who demonstrate that a lot of networks are partially scale-free, in that the body of the degree distribution is modelled well by a power law while the tail is lighter than what is implied by the body (although still regularly varying).

Most of the study into the degrees of real networks, including those mentioned here, has been purely descriptive in the sense that no information about the networks growth is revealed (aside from relating a power law to preferential attachment).

The study of network growth began to gain popularity when Barabási and Albert (1999) showed that a network that has grown in such a way that nodes gain edges at a rate proportional to their degree results in a network that is scale-free. This is why the presence of real scale-free networks is such a hot topic, since if it is shown that they are often scale-free then a model such as preferential attachment gains a lot of justification for being used to model network growth. The preferential attachment model they introduced uses a preference function, $b(k) = k + 1$ where k is a node in-degree, that is fairly simple and provided the foundations for many modifications and extensions to the model. Krapivsky and Redner (2001) considered a more general preference function of the form $b(k) = (k+1)^\alpha$ and showed that the degree distribution only has regular variation in the tail when $\alpha = 1$, no regular variation when $0 < \alpha < 1$ (and therefore not power law), and when $\alpha > 1$ a finite number of nodes end up with all edges after a certain point resulting in a degenerate degree distribution. Wang and Resnick (2024) returns to a linear preference function of the form $b(k) = k + \varepsilon$ but adds the possibility for reciprocal edges to be sent, this results in joint in-degree out-degree distribution that is multivariate regular varying and has the property of hidden regular variation. Rudas et al (2007) follows in the footsteps of Krapivsky and

Redner (2001) and considers a preferential attachment tree with a preference function $b(\cdot)$ and uses theory from continuous time branching processes to derive a limiting degree distribution in terms of the preference function $b(\cdot)$. Research in this area tends to only focus on the theoretical asymptotic results of network growth models with little analysis of real networks. This paper aims to address this gap, asking if a network is assumed to come from a preferential attachment model can we use the degree distribution alone to directly infer the model parameters?

It is important to think about how we intend to consider the tail of the degree distributions, because if not done correctly we may end up essentially discounting the effects of the largest degrees (often the most influential nodes) deviating from a power law. For this reason we will use methods from discrete extreme value theory, as they best capture the nature of the data involved, more specifically we will use results from Shimura (2012) to study how the tail of the degree distribution is affected by the preference function since we wish to align what we do with the recent findings that the a lot of real degree distributions are regularly varying.

The remainder of the paper is as follows: Section 2 gives a detailed description of the preferential attachment model alongside the theoretical results for the survival of the limiting degrees with a focus on the tail behaviour in terms of the preference function $b(\cdot)$. It is shown that a regularly varying degree distribution can only be the result of an asymptotically linear preference function. This section also introduces a preference function that can guarantee regular variation in the degrees while remaining flexible via a set of parameters that can provide behaviour similar to that in Krapivsky and Redner (2001) up until a threshold. Section 3 utilises the preference function proposed at the end of Section 2, and illustrates how the extreme value index (EVI) of the degree distribution varies with the model parameters. Section 4.1 consists of a simulation study where networks are simulated from the proposed model for various parameter combinations, demonstrating that the parameters can be recovered from fitting the model to only the degree distribution. Section 4.2 fits the model to some real data and provides posterior estimates for the preference function. Section 5 concludes the article.

2 Tail Behaviour of Preferential Attachment Model

The model that we will focus on in this paper is the General Preferential Attachment model in Rudas et al (2007) and is defined as follows:

Starting at time $t = 0$ with an initial network of m vertices that each have no edges, at times $t = 1, 2, ...$ a new vertex is added to the network bringing with it m directed edges (with the new vertex as the source); the target for each of these edges are selected from the vertices already in the network with weights proportional to some preference function $b(\cdot)$ of their degree, where $b : N \mapsto \mathbb{R}^+ \setminus \{0\}$ is such that:

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

Special cases of this model include the Barabási-Albert (BA) model when $b(k) = k+\varepsilon$, which leads to a power-law degree distribution with EVI $\xi = 1/2$, and the Uniform Attachment (UA) model where $b(k) = c$ leading to a degree distribution that is not regularly varying.

The survival function of the limiting degree distribution, called the limiting survival hereafter, under condition 1 can be analytically derived in the case where $m = 1$.

Consider a continuous time branching process $\zeta(t)$ driven by a Markovian pure birth process, with $\zeta(0) = 0$ abd birth rates depending on a non-negative function $b(\cdot)$:

$$\Pr(\zeta(t + dt) = k + 1|\zeta(t) = k) = b(k)dt + o(dt).$$

Now let $\Upsilon(t)$ be the tree determined by $\zeta(t)$ as follows: $\Upsilon(t) = \{\emptyset\}$ and $\Upsilon(t) = G$ where each existing node x in $\Upsilon(t)$ gives birth to a child with rate $b(\deg(x, \Upsilon(t))$ independently of the other nodes where $\deg(x, \Upsilon(t))$ is the degree of node x in the tree $\Upsilon(t)$ at time t.

Theorem 1 from Rudas et al (2007) states that for the tree $\Upsilon(t)$ at time t:

$$\lim_{t\to\infty} \frac{1}{|\Upsilon(t)|} \sum_{x\in\Upsilon(t)} \varphi(\Upsilon(t)_{\downarrow x}) = \lambda^* \int_0^{\infty} e^{-\lambda^* t} \mathbb{E}\left[\varphi(\Upsilon(t))\right] dt \tag{2}$$

where λ^* satisfies $\hat{\rho}(\lambda^*) = 1$ and $\hat{\rho}$ is the Laplace transform of the density of the point process associated with the pure birth process that corresponds to a nodes individual growth, that is:

$$\hat{\rho}(\lambda) := \int_0^{\infty} e^{-\lambda t} \rho(t) \mathrm{d}t.$$

The limiting survival can be viewed as the limit of the empirical proportion of vertices with degree over a threshold $k \in \mathbb{N}$, 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 can also be written using Equation 2 as:

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

Therefore, the corresponding probability mass function of the degree distribution $f(k) = \bar{F}(k-1) - \bar{F}(k)$ is

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

We are interested in how the tail behaviour of the discrete limiting degree distribution is affected by the preference function b.

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

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

Shimura (2012) states that if $\lim_{k\to\infty} \Omega(F, k) = 1/\alpha$ ($\alpha > 0$), then F is regularly varying with $\bar{F}(k) \sim k^{-\alpha}$. On the other hand, if $\lim_{k\to\infty} \Omega(F, k) = 0$ then we will refer to the distribution as light-tailed. This allows us to show the following:

Proposition 2.1. *If $\bar{F}(k) = \prod_{i=0}^{k} \frac{b(i)}{\lambda^*+b(i)}$ and $b(k) \to \infty$ as $k \to \infty$, then*

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

See Appendix C for the details of the proof.

Proposition 2.1 aligns with the result from Krapivsky and Redner (2001) demonstrating that a sub-linear preference function will lead to a light-tailed distribution, as $\lim_{k\to\infty} b(k+1) - b(k) = 0$ if $b(k) = k^\alpha$ where $\alpha < 1$. Proposition 2.1 also aligns with the fact that BA model produces a regularly varying degree distribution with EVI $\xi = 0.5$ by considering the preference function $b(k) = k + \varepsilon$, as $\lim_{k\to\infty} b(k+1) - b(k) = 1$ leaving the tail index to be $1/\lambda^*$ which using $\hat{\rho}$ can be found to be $1/2$. So, in order for the degree distribution to be regularly varying we need that the limit $\lim_{k\to\infty} b(k+1) - b(k)$ exists and is positive. To determine the class of functions that will result in regularly varying limiting degree distributions, we use the following result:

Proposition 2.2. *Consider a GPA model with preference function $b(\cdot)$ satisfying Proposition 2.1. Then the limiting degree distribution is regularly varying with EVI c/λ^* if and only if $\lim_{k\to\infty} b(k)/k = c > 0$.*

Proof

From Proposition 2.1, we have that:

$$\lim_{k\to\infty}[b(k+1)-b(k)] = c > 0.$$

Now, setting $b_k = b(k)$ and $a_k = k$:

$$\lim_{k\to\infty}[b(k+1)-b(k)] = \lim_{k\to\infty}\frac{b_{k+1}-b_k}{a_{k+1}-a_k} = c > 0,$$

meaning by Theorem A.1 in Appendix A:

$$\lim_{k\to\infty}\frac{b_k}{a_k} = \lim_{k\to\infty}\frac{b(k)}{k} = c > 0 \qquad \square.$$

Using this result we can understand how the preference function is directly connected with the EVI and regular variation of the degree distribution. We use this result to create a preference function that guarantees regular variation in the tail of the degree distribution, aligning with analysis of real networks, whilst allowing for the tail to deviate from the shape implied by the body. This gives the model the capability to produce realistic behaviour in the degrees like what was in Lee et al (2024) by using a piecewise function inspired by the observed deviation from the power law after a certain threshold.

$$b(k) = \begin{cases} k^\alpha + \varepsilon, & k < k_0, \\ k_0^\alpha + \varepsilon + \beta(k-k_0), & k \geq k_0 \end{cases} \tag{3}$$

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

This preference function, as per Proposition 2.1, will produce a degree distribution with EVI $\xi = \beta/\lambda^*$ guaranteeing regular variation. We study this preference function further in the next section.

3 A Preferential Attachment Model with Flexible Regular Variation

In the previous section, we found that using a preference function with linearity in the limit allows for the inclusion of sub/super-linear behaviour below the threshold, while simultaneously guaranteeing regular varation of the degrees. In this section, we study how the limiting degree distribution varies with the preference function according to Equation 3. Its survival is

$$\bar{F}(k) = \begin{cases} \prod_{i=0}^{k} \frac{i^\alpha+\varepsilon}{\lambda^*+i^\alpha+\varepsilon}, & k < k_0, \\ \left(\prod_{i=0}^{k_0-1} \frac{i^\alpha+\varepsilon}{\lambda^*+i^\alpha+\varepsilon}\right) \frac{\Gamma(\lambda^*+k_0^\alpha+\varepsilon)/\beta)}{\Gamma((k_0^\alpha+\varepsilon)/\beta)} \frac{\Gamma\left(k-k_0+1+\frac{k_0^\alpha+\varepsilon}{\beta}\right)}{\Gamma\left(k-k_0+1+\frac{\lambda^*+k_0^\alpha+\varepsilon}{\beta}\right)}, & k \geq k_0, \end{cases} \tag{4}$$

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

$$\hat{\rho}(\lambda) = \sum_{n=0}^{k_0} \prod_{i=0}^{n-1} \frac{i^\alpha+\varepsilon}{\lambda+i^\alpha+\varepsilon} + \left(\frac{k_0^\alpha+\varepsilon}{\lambda-\beta}\right) \prod_{i=0}^{k_0-1} \frac{i^\alpha+\varepsilon}{\lambda+i^\alpha+\varepsilon} \tag{5}$$

which must be solved numerically for most parameter choices. Also, note that $\lambda > \beta$.

Some examples of the limiting degree distribution for various parameter combinations are shown below on log-log scale in Figure 1:

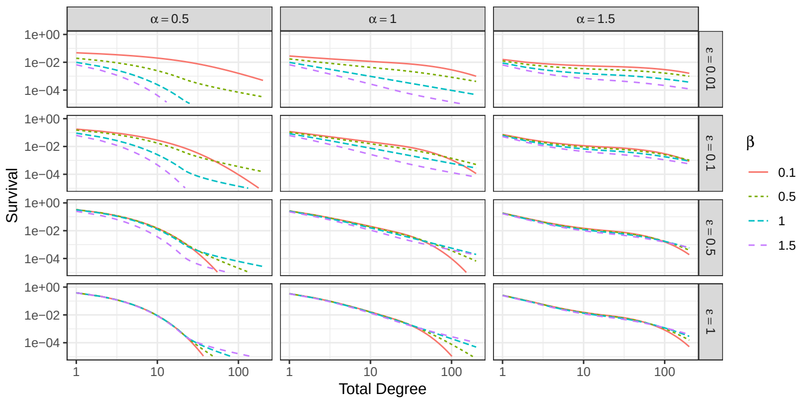

Figure 1: Theoretical survival distributions of the limiting degree distributions, according to various combinations of $(\alpha, \beta, \varepsilon)$ and $k_0 = 20$ of the proposed preferential attachment model.

Figure 1 demonstrates that this model can capture a wide variety of shapes for the survival function, including a large range of possible tail indices ranging from 0.035 ($\alpha = 1.5, \beta = 0.1, \varepsilon = 1$) to 0.999 ($\alpha = 0.5, \beta = 1.5, \varepsilon = 0.01$).

The survival function (4) can be connected to the discrete version of the generalised Pareto distribution (GP), called the Integer GP (IGP) (Rohrbeck et al, 2018) with conditional survival:

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

for $v \in \mathbb{Z}^+, \sigma > 0, \xi \in \mathbb{R}$, denoted as $X|X > u \sim \mathrm{IGP}(\xi, \sigma, u)$ where ξ is the shape parameter controlling the tail index

By Equation 4 and using Stirling's approximation:

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

Therefore,

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

meaning that for $k \geq k_0$ the limiting degree distribution (for large k_0^α) is approximated by $\mathrm{IGP}\left(\frac{\beta}{\lambda^*}, \frac{k_0^\alpha + \varepsilon}{\lambda^*}, k_0 - 1\right)$.

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

In agreement with Proposition 2.2, $\beta > 0$ ensures that the shape parameter of the IGP is positive and thus the distribution is regularly varying. Additionally the value of the shape parameter ξ is shown in Figure 3 for various parameter choices. The darker and lighter regions on the heat maps correspond to a heavier and a lighter tail, respectively, the red dashed line shows combinations of α and β that produce a limiting degree distribution with the same tail index as the Barabási-Albert model, that is, $\xi = 0.5$.

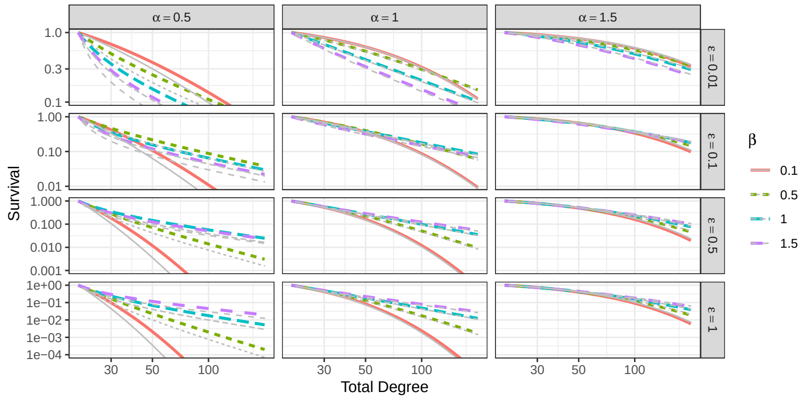

Figure 2: Theoretical conditional survivals (grey) alongside their IGP approximations (coloured).

Through the connection to the IGP, fitting this model is almost equivalent to fitting the IGP to the degree and estimating the parameters but instead of only describing the shape of the degree distribution we would also gain estimates for the shape of the preference function while can help understand the mechanisms underlying the growth of the network.

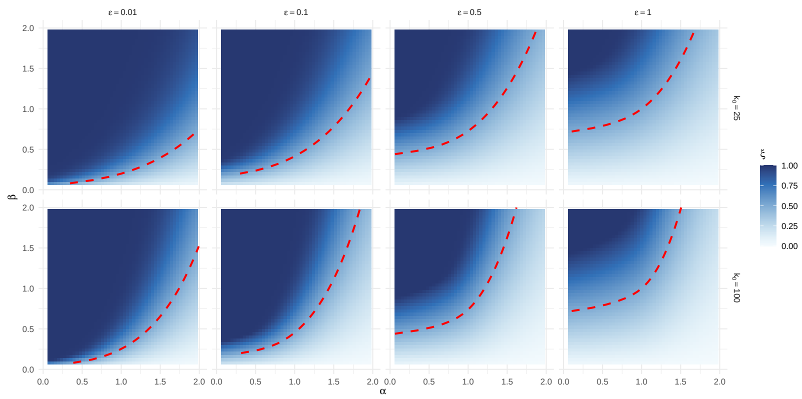

Figure 3: Heat maps of ξ for various combinations of the parameters of the proposed model.

To perform inference of the model parameters, we consider a network with degree count vector $\boldsymbol{n} = (n_0, n_1, \dots, n_M)$, where M is the maximum degree. We can then write the likelihood as:

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

where $B(y, z)$ is the the beta function, and $l \geq 0$ is a variable that allows truncating the data such that the minimum degree is l. This will allow the model to be fitted whilst ignoring the influence of the lower degrees (those less than l) as the model does not capture the behaviour well at the lower degrees since Rudas et al (2007) only provides results for the case of a preferential attachment tree.

4 Applications

4.1 Simulated Data

This first subsection aims to show that the parameters of the model (and therefore the preference function) in Section 3 can be recovered from simulating a network from the model, and fitting it to the observed degree distribution, using the likelihood in (3).

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 and using the likelihood from the previous section alongisde the priors:

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

where Gamma(a, b) is the Gamma distribution with shape a and rate b, and U(a, b) is uniform distribution with lower and upper bounds a and b, to obtain a posterior distribution, up to the proportionality constant. Posterior samples can then be obtained by an adaptive Metropolis-Hastings Markov chain Monte Carlo (MCMC) algorithm. For these simulated networks $l = 0$.

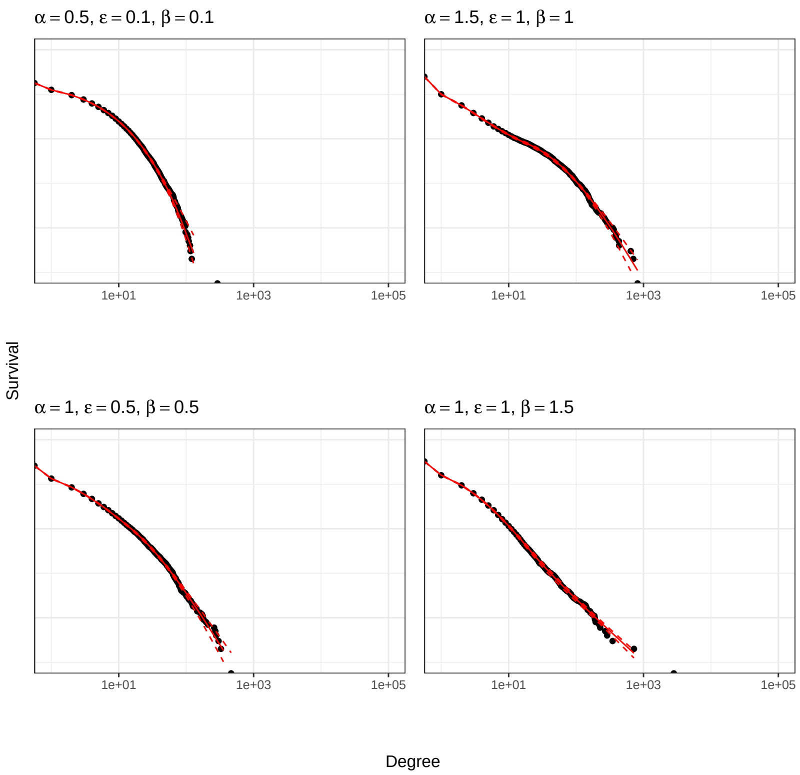

Figure 4: Posterior estimates of survival function for data simulated from the proposed model with various combinations of $(\alpha,\beta,\varepsilon)$ and $k_0 = 20$.

Figure 4 and Figure 5 shows the estimations demonstrates the usefulness of the model, as we can recover the model parameters well from only the final degree distribution of the simulated network. This indicates that the method may also be applied to the degree distributions of real networks, estimating the model parameters assuming they evolved according to the GPA scheme.

4.2 Real Data

In this section, we fit the proposed model to the degree distributions of various real networks and learn about the mechanics of their growth. While we also compare the fit to that of the mixture distribution by Lee et al (2024) we note that the proposed method has the additional benefit of learning directly about the growth of a network from the inference results. The data consists of 12 networks sourced from KONECT and the Network Data Repository(Rossi and Ahmed, 2015):

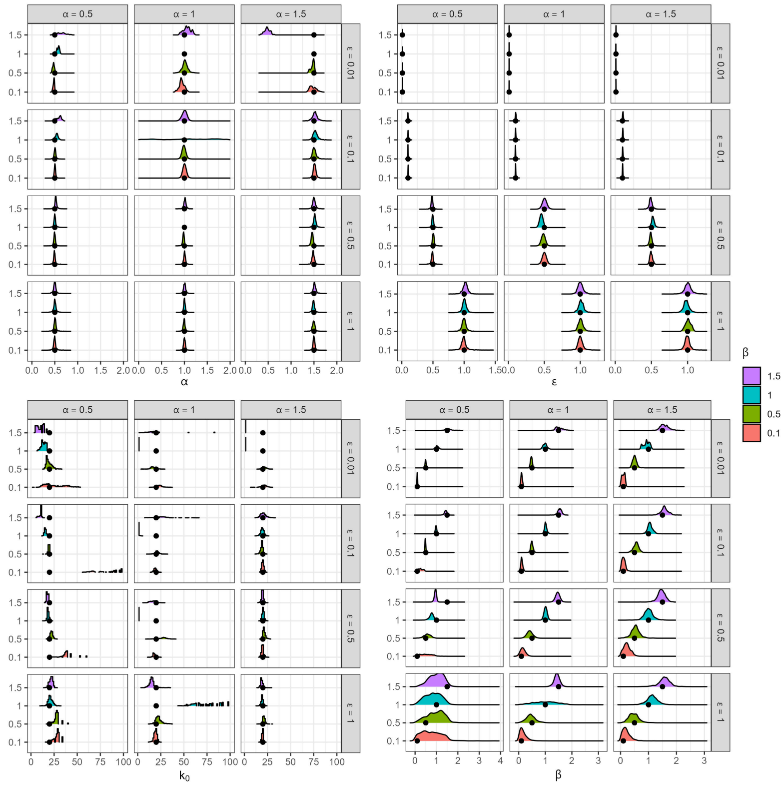

Figure 5: : Posterior estimates of paramters for data simulated from the proposed model with various combinations of (α,β,ε) and $k_0 = 20$.

- `as-caida20071105`: network of autonomous systems of the Internet connected with each other from the CAIDA project
- `dimacs10-astro-ph` : co-authorship network from the "astrophysics" section (astro-ph) of arXiv
- `ego-twitter`: network of twitter followers
- `facebook-wosn-wall`: subset of network of Facebook wall posts
- `maayan-faa`: USA FAA (Federal Aviation Administration) preferred routes as recommended by the NFDC (National Flight Data Center)
- `maayan-Stelzl`: network representing interacting pairs of proteins in humans
- `moreno-blogs-blogs`: network of URLs found on the first pages of individual blogs
- `opsahl-openflights`: network containing flights between airports of the world.
- `pajek-erdos`: co-authorship network around Paul Erdø"s
- `reactome`: network of protein–protein interactions in humans
- `sx-mathoverflow`: interactions from the StackExchange site MathOverflow
- `topology`: network of connections between autonomous systems of the Internet

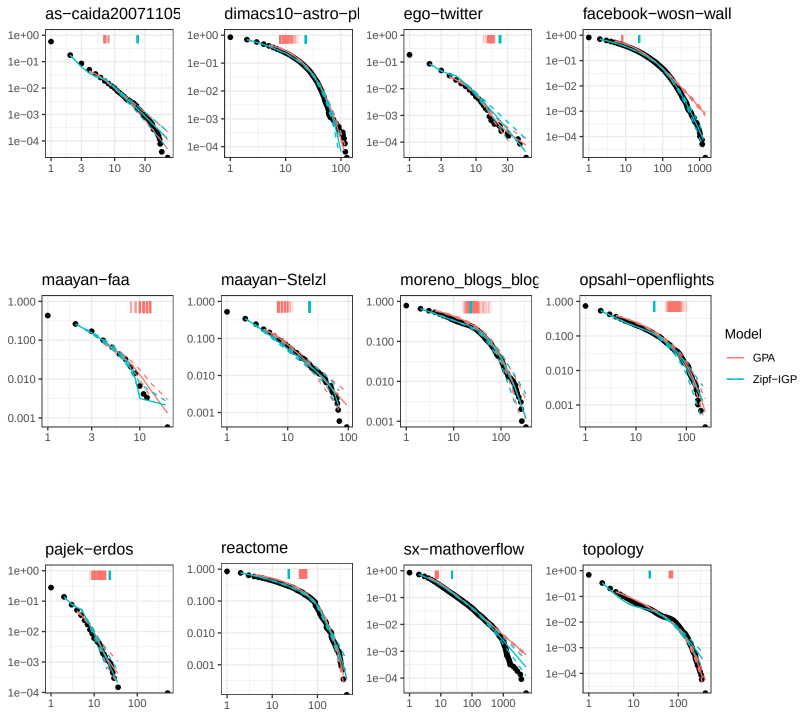

Figure 6: Posterior estimates (solid red) of survival for several real data sets and their 95% credible intervals (dotted red).

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 from Lee et al (2024). In most cases, the GPA model gives a similar fit 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. Generally, the GPA model performs similarly to the Zipf-IGP when estimating the tail behaviour of the degree distribution. In the cases of substantial discrepancies, it is either because the GPA model fits the tail better than the Zipf-IGP model does, or because of the threshold being estimated as too low forcing almost all of the data to be modelled by the linear part of the GPA. This again shows the effects that small degrees have on this model, which is somewhat expected as the theory used for this model is for trees and none of these real networks (nor many real networks) are.

Figure 8 shows the estimated preference function $b(k)$ alongside the 95% credible interval on a log-log plot. Although the credible interval becomes very large for the largest degrees, this is expected as not all of these networks had data in that region, and for those that do the credible interval is much narrower, as is the case for

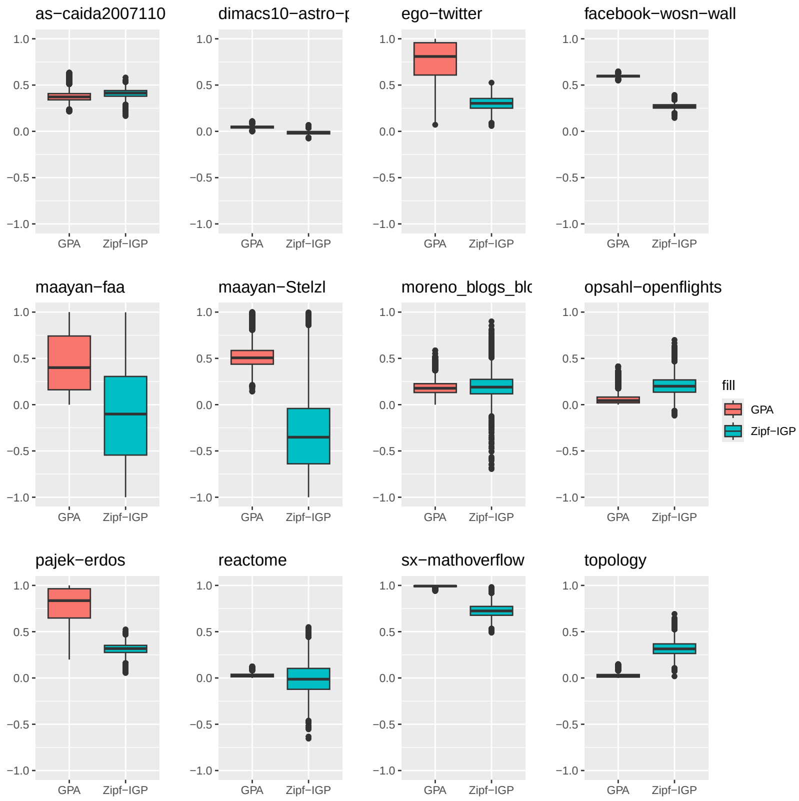

Figure 7: Posterior estimates (solid red) of survival for several real data sets and their 95% credible intervals (dotted red).

`sx-mathoverflow`. Looking at the shape of the preference function, there appears to be two distinct shapes of preference function. The first appears mostly flat (similar to uniform attachment) for the smallest degrees and then after a threshold preferential attachment kicks in, some with this shape are `pajek-erdos` and `sx-mathoverflow`. The second distinct shape appears to provide some clear preferential attachment behaviour that then slows down after a certain point, examples of this are seen in the two infrastructure networks `opsahl-openflights` and `topology`. This slowing down could be viewed as a kind of diminishing returns on the degree of a vertex i.e. as a vertex gets larger gaining more connections has less of an effect than it did before some threshold k_0.

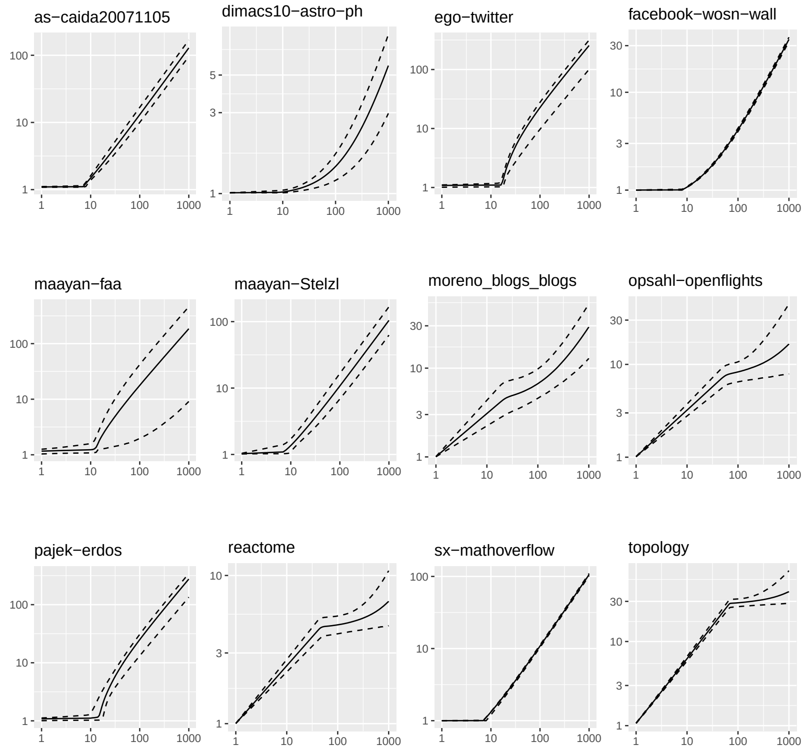

Figure 8: Posterior estimate for preference function (solid) with 95% credible interval (dashed) on log-log scale.

5 Conclusion and Discussion

In this paper we introduced a class of preference functions that, under the GPA scheme, generate a network with a flexible yet heavy-tailed degree distribution. From the simulation study we showed that the parameters can be recovered from fitting the model to the degrees alone. We also applied this method to the degree distributions of real networks, estimating their model parameters assuming they evolved in the same way. Not only did this yield fairly good fits for the degree distribution, similar to that of the Zipf-IGP, it also came with the added benefit of giving a posterior estimate for a preference function.

One limitation of this method is that the lowest degrees needed to be truncated as they had a very large effect on the fit of the model as a result of using theory developed for trees and applying it to general networks. Future work could apply theory developed for general networks using a similar method to this, allowing us to compare the results here something that is more accurate. This could include fixing the out-degree of new nodes at a constant greater than one, or allowing the out-degree of new nodes to vary.

A Supplementary Results

Theorem A.1 (Stolz-Cesàro Theorem (Cesàro, 1888))**.** *Let* $(a_k)_{k\geq 1}$ *and* $(b_k)_{k\geq 1}$ *be two sequence of real numbers. Assume* $(a_k)_{k\geq 1}$ *is a strictly monotone and divergent sequence and the following limit exists:*

$$\lim_{k\to\infty} \frac{b_{k+1} - b_k}{a_{k+1} - a_k} = l.$$

Then,

$$\lim_{k\to\infty} \frac{b_k}{a_k} = l.$$

B Additional Plots

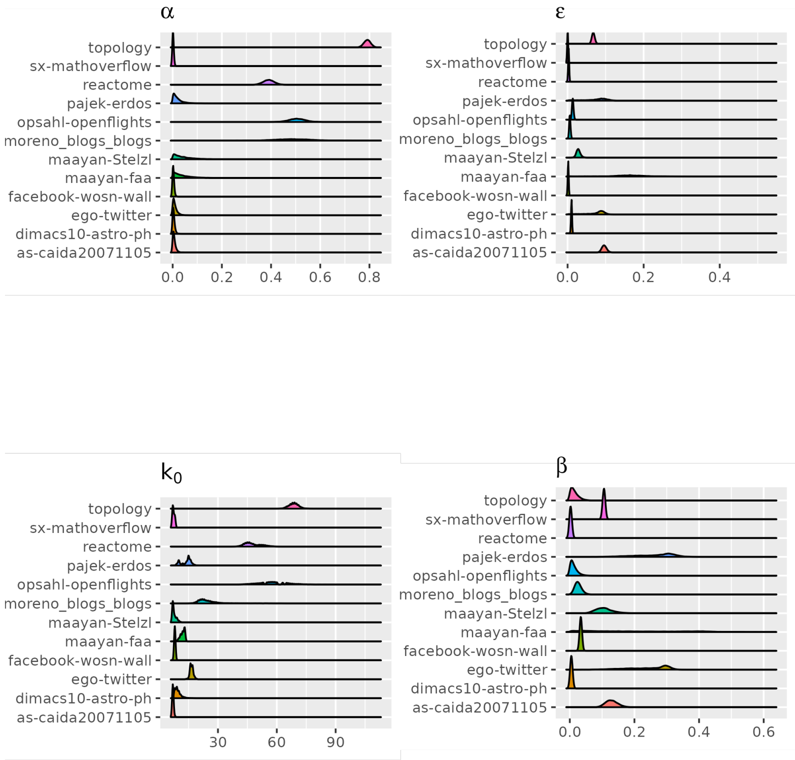

Figure A1: : Posterior estimates of parameters for real data.

C Proofs and Derivations

C.1 Proof of Proposition 2.1

Taking the form of the GPA degree survival:

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

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

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

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

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

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

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

C.2 Derivation of Equation 5

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:

$$\sum_{n=0}^{\infty} \frac{\Gamma(n + x)}{\Gamma(n + x + y)} = \frac{1}{\Gamma(y)} \sum_{n=0}^{\infty} \mathrm{B}(n + x, y)$$

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

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

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

References

Barabási AL, Albert R (1999) Emergence of scaling in random networks. Science 286(5439):509–512. https://doi.org/10.1126/science.286.5439.509

Betzel RF, Bassett DS (2017) Generative models for network neuroscience: prospects and promise. Journal of The Royal Society Interface 14(136):20170623. https://doi.org/10.1098/rsif.2017.0623

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

Cesàro E (1888) Sur la convergence des séries. Nouvelles annales de mathématiques : journal des candidats aux écoles polytechnique et normale 3e série, 7:49–59. URL https://www.numdam.org/item/NAM_1888_3_7___49_1/

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

Latouche P, Birmelé E, Ambroise C (2011) Overlapping stochastic block models with application to the French political blogosphere. The Annals of Applied Statistics 5(1):309 – 336. https://doi.org/10.1214/10-AOAS382

Lee C, Eastoe EF, Farrell A (2024) Degree distributions in networks: Beyond the power law. Statistica Neerlandica 78(4):702–718. https://doi.org/https://doi.org/10.1111/stan.12355

Rohrbeck C, Eastoe EF, Frigessi A, et al (2018) Extreme value modelling of water-related insurance claims. The Annals of Applied Statistics 12(1):246 – 282. https://doi.org/10.1214/17-AOAS1081

Rossi RA, Ahmed NK (2015) The network data repository with interactive graph analytics and visualization. In: AAAI, URL https://networkrepository.com

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

Setayesh A, Sourati Hassan Zadeh Z, Bahrak B (2022) Analysis of the global trade network using exponential random graph models. Applied Network Science 7. https://doi.org/10.1007/s41109-022-00479-7

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

Voitalov I, van der Hoorn P, Hofstad R, et al (2019) Scale-free networks well done. Physical Review Research 1. https://doi.org/10.1103/PhysRevResearch.1.033034

Wang T, Resnick S (2024) Random networks with heterogeneous reciprocity. Extremes 27(1):123–161. https://doi.org/10.1007/s10687-023-00478-6