

# Scale-Free Stochastic Networks

## Abstract

*This report will look into the underlying mechanisms of fat tailed degree distributions in real-world networks. A proposed underlying mechanism for this phenomenon is preferential attachment, which is in practice implemented via the Barabasi-Albert model of a growing network. After analysis on the degree distributions of the Barabasi-Albert model, this will be compared to other schemes, using random attachment and random walks. The results show that the fat tailed distributions occur in the Barabasi-Albert model, and also using random walks, for the length of the walk  $L > 1$ . The random attachment model exhibits an exponential cut-off. Performing a chi-square test for both the Barabasi-Albert model and the random attachment model, comparing the data to the theoretical forms, the resulting  $p$ -values gave no reason to doubt the theoretical derivations. For the Barabasi-Albert model, the predicted largest value of the degree did not match the mean average of the largest observed degree, but it matched the modal average, it was hypothesised that this effect could be due to the high skewness and kurtosis in fat-tailed distributions. The random walks appeared to follow fat tailed distributions for  $L > 1$ , but had an even higher skewness and kurtosis for  $L = 1$ . The  $L = 0$  case appeared to follow the same form of the random attachment distribution.*

## I. AIMS & INTRODUCTION

THE Barabasi-Albert, or price model is a network model that uses the concept of preferential attachment. This more closely resembles the citation network (where the degree is the number of citations and the nodes are scientific papers) than the Erdos-Reyni random graphs. Citation networks, have a fat-tailed distribution, so that the degrees are present on all scales. The aims are to investigate the Barabasi-Albert model, in terms of how well theoretical predictions match with the simulated degree distribution, this will also be done for random attachment. Then the report will look into using random walks on the network as the mechanism for choosing which edges are connected.

## II. METHODS

### I. Algorithm

The algorithm used is defined as follows:<sup>2</sup>

1. Set up initial network  $\mathcal{G}_0$
2. Increment time  $t \rightarrow t + 1$
3. Add one new node
4. Add  $m$  edges as follows:
  - Connect one end of the new edge to the new node.
  - Connect the other end to an existing node with probability  $\Pi$
5. Repeat from 2

In the script that was written, there was no need to increment time, since it is proportional to the number of nodes and the number of nodes is automatically known by the “graph” object, creating a time variable and incrementing it would have caused a slight inefficiency. The initial network that was used was a complete network with the number of nodes equal to  $m + 1$ , this was so that all the initial nodes had the minimum degree  $k = m$ . For preferential attachment (Barabasi-Albert model)  $\Pi \propto k$ , for random attachment  $\Pi \propto \text{constant}$ .

In the case of using random walks in order to choose the nodes that are connected to the new node, after setting up an initial network, choose a node at random, and pick any neighbouring node from the currently selected node at random, do this  $L$  times, the last node to be selected is then the vertex to be added. This process is then repeated  $m$  times. The program was checked to be working by querying the graphs for node number and degree at each time step, and seeing that the node number increases by 1 at each time step and edge number increasing by  $m$  at each time step. The parameters required by the program are  $m$ , and the method for which to choose which vertices to connect to.

## II. Derivation of Degree Distribution for Preferential Attachment

The master equation for the Barabasi-Albert model is

$$n(k, t+1) = n(k, t) + m\Pi(k-1, t)n(k-1, t) - m\Pi(k, t)n(k, t) + \delta_{km}$$

where  $n$  is the number of nodes of degree  $n$ . The first term on the right  $n(k, t)$  comes from the starting number at time  $t$ , the second term  $m\Pi(k-1, t)n(k-1, t)$  comes from the average number of nodes with degree  $k-1$  that is increased in degree as a result of one time step. The third term  $m\Pi(k, t)n(k, t)$  comes from the average number of nodes with degree  $k$  that is increased in degree as a result of a time step. The  $\delta_{km}$  term comes from the node that is added with degree  $m$ .

The probability  $\Pi$  is given by:

$$\Pi(k, t) = \frac{k}{2E(t)} = \frac{k}{2mN(t)}$$

where  $E$  is the number of edges at time  $t$ , since the number of edges added is  $m$  each time, it's proportional to the number of nodes  $N$  with proportionality constant  $m$ . In terms of probability distribution  $p(k, t) = n(k, t)/N(t)$ :

$$N(t+1)p(k, t+1) - N(t)p(k, t) = \frac{1}{2} [(k-1)p(k-1, t) - kp(k, t)] + \delta_{km}$$

Since  $N(t+1) = 1 + N(t)$ , taking the limit as  $t \rightarrow \infty$  of  $p$ ,  $p(k, t) \rightarrow p_\infty(k)$ :

$$p_\infty(k) = \frac{1}{2} [(k-1)p_\infty(k-1) - kp_\infty(k)] + \delta_{km}$$

First assuming that  $k \neq m$  rearranging this gives:

$$\frac{p_\infty(k)}{p_\infty(k-1)} = \frac{k-1}{k+2}$$

which has the solution,

$$p_\infty(k) = A \frac{\Gamma(k)}{\Gamma(k+3)} = \frac{A}{k(k+1)(k+2)}$$

for any  $A$ . Considering the case where  $m = k$  and  $p_\infty(m-1) = 0$ , gives,

$$\begin{aligned} p_\infty(m) &= 1 - \frac{mp_\infty(m)}{2} \\ \implies p_\infty(m) &= \frac{2}{m+2} \end{aligned}$$

This gives the normalisation constant  $A$ :

$$\begin{aligned} \frac{A}{(m+2)(m+1)m} &\stackrel{!}{=} \frac{2}{m+2} \\ \implies A &= 2m(m+1) \end{aligned}$$

Thus:

$$p_\infty(k) = \frac{2m(m+1)}{k(k+1)(k+2)} \quad \text{for } k \geq m$$

Checking whether this is normalised:

$$\begin{aligned}
 \sum_{k=m}^{\infty} p_{\infty}(k) &= \sum_{k=m}^{\infty} \frac{2m(m+1)}{k(k+1)(k+2)} \\
 &= \sum_{k=m}^{\infty} \frac{m(m+1)}{k} - \frac{2m(m+1)}{k+1} + \frac{m(m+1)}{k+2} \\
 &= (m+1) - 2m + \frac{m(m+1)}{m+2} \\
 &\quad + m - \frac{2m(m+1)}{m+2} + \vdots \\
 &\quad + \frac{m(m+1)}{m+2} - \vdots + \vdots \\
 &= 1
 \end{aligned}$$

As required.

### III. Derivations of Degree Distribution for Random Attachment

The form of the master equation is unchanged, this time the attachment probability is:

$$\Pi(k, t) = \frac{1}{N(t)} = \frac{1}{t + N(0)}$$

substituting into the master equation gives:

$$n(k, t+1) - n(k, t) = \frac{m}{t+t_0}(n(k-1, t) - n(k, t)) + \delta_{km}$$

substituting  $p(k, t) = n(k, t)/N(t)$ :

$$(t+1+N(0))p(k, t+1) - (t+N(0))p(k, t) = m(p(k-1, t) - p(k, t)) + \delta_{km}$$

Taking  $t \rightarrow \infty$ :

$$p_{\infty}(k) = m(p_{\infty}(k-1) - p_{\infty}(k)) + \delta_{km} \quad (1)$$

Starting from this equation there are two methods that is used to derive  $p_{\infty}(k)$ .

#### III.1 Method I

Assume  $k \neq m$ , then rearrange equation 1:

$$p_{\infty}(k) = mp_{\infty}(k-1)/(m+1)$$

for the  $k = m$  case:

$$p_{\infty}(m) = \delta_{mm}/(m+1) = 1/(m+1)$$

Setting  $p_{\infty}(m-1) = 0$  as before. Hence, iterating the previous equation we can conclude that:

$$p_{\infty}(k) = \frac{1}{m+1} \left( \frac{m}{m+1} \right)^{k-m} \quad \text{for } k \geq m$$

Checking the normalisation:

$$\begin{aligned}
 \sum_{k=m}^{\infty} p_{\infty}(k) &= \frac{1}{1+m} \sum_{k=m}^{\infty} \left( \frac{m}{m+1} \right)^{k-m} \\
 &= \frac{1}{1+m} \sum_{k=0}^{\infty} \left( \frac{m}{m+1} \right)^k \\
 &= \frac{1}{(1+m)} \frac{1}{(1 - \frac{m}{m+1})} \\
 &= 1
 \end{aligned}$$

As required.

### III.2 Method II

The probabilities can be transformed using a discrete mellin transform:<sup>1</sup>

$$G_\infty(k) = \sum_k z^k p_\infty(k)$$

Multiplying equation 1 by  $z^k$  and summing over all  $k$  gives:

$$\begin{aligned} G_\infty(k) &= z^m - mG_\infty(k) + \sum_k mz^k p_\infty(k-1) \\ G_\infty(k)(1+m) &= z^m + \sum_k z^{k+1} p_\infty(k) \\ G_\infty(k)(1+m) &= z^m + mzG_\infty(k) \\ G_\infty(k) &= \frac{z^m/m}{\frac{1+m}{m} - z} \end{aligned}$$

Now using the inverse Mellin transform, the probabilities are obtained:

$$\begin{aligned} p_\infty(k) &= \mathcal{M}^{-1}[G_\infty(z)] = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} z^{-k-1} G_\infty(z) dz \\ p_\infty(k) &= \frac{1}{2m\pi i} \int_{c-i\infty}^{c+i\infty} \frac{z^{m-k-1}}{\frac{1+m}{m} - z} dz \end{aligned}$$

Where  $c \in [0, \frac{1+m}{m}]$  (in this particular case), is between the regions on the complex plane in which there are residues in the integrand.

Consider:

$$\frac{\text{Res}(f(z), z_0)}{m} = \frac{1}{2m\pi i} \lim_{R \rightarrow \infty} \left( \int_{c-iR}^{c+iR} \frac{z^{m-k-1}}{\frac{1+m}{m} - z} dz + \int_{\mathcal{C} \in \mathbb{C}} \frac{z^{m-k-1}}{\frac{1+m}{m} - z} dz \right)$$

where  $\mathcal{C}$  encloses a semicircle from  $-iR$  to  $iR$ , in either direction/chirality,  $f(z)$  is the integrand  $\frac{z^{m-k-1}}{\frac{1+m}{m} - z}$ , and  $z_0$  is the location of either of its residues.

Using the substitution  $z = Re^{i\theta} - c$ ,  $dz = izd\theta$

$$\begin{aligned} \left| \int_{\mathcal{C} \in \mathbb{C}} \frac{z^{m-k}}{\frac{1+m}{m} - z} id\theta \right| &< \int_{\mathcal{C} \in \mathbb{C}} \left| \frac{z^{m-k}}{\frac{1+m}{m} - z} \right| d\theta \\ &\approx \int_{\mathcal{C} \in \mathbb{C}} |R^{m-k-1}| d\theta \end{aligned}$$

Which goes to zero as  $R \rightarrow \infty$  for the constraint on the network  $k \geq m$ . Hence the integral  $f(z)$  in the region  $\Re(z) > c$  or  $\Re(z) < c$  is entirely due to the vertical line in the complex plane.

$$p_\infty(k) = \frac{\text{Res}(f(z), z_0)}{m}$$

Taking the Laurent expansion of  $f(z)$  about  $z = 0$  gives, for  $b := m - k$ ,  $a := \frac{1+m}{m}$

$$f(z) = z^b \left( \frac{1}{az} + \frac{1}{a^2} + \frac{z}{a^3} + \dots \right)$$

Taking the positive residues since on the region  $\Re(z) < c$ , the integral goes around the residue in a right handed chirality:

$$p_\infty(k) = \frac{1}{m} a^{b-1} = \frac{1}{m} \left( \frac{1+m}{m} \right)^{m-k-1} = \frac{1}{m+1} \left( \frac{1+m}{m} \right)^{m-k}$$

Now checking the solution, by integrating over  $\Re(z) > c$ , the Laurent expansion about  $z = a$  is:

$$-\frac{a^{b-1}}{z-a} - (b-1)a^{b-2} - \dots$$

This time taking the negative of the residue since the integral goes round the residue in a left handed chirality,

$$p_\infty(k) = \frac{1}{m}a^{b-1}$$

as before. Both methods give the same solution.

## IV. Largest Expected Degree for a Given Network Size

### IV.1 Preferential Attachment Case

network of size  $N$  for large  $N$  can be approximated with  $p_\infty(k)$ . For at least 1 to have the degree  $k_{max}$ , the probability,

$$P(k > k_{max}) \stackrel{!}{=} \frac{1}{N}$$

Hence

$$\begin{aligned} 1/N &= \sum_{k=k_{max}}^{\infty} \frac{2m(m+1)}{k(k+1)(k+2)} \\ &\approx m(m+1) \int_{k=k_{max}}^{\infty} \frac{1}{k} - \frac{2}{k+1} + \frac{1}{k+2} \\ &= m(m+1) \ln \left( \frac{k(k+2)}{(k+1)^2} \right) \Big|_{k=k_{max}}^{\infty} \end{aligned}$$

Assessing the convergence of the top limit:

$$\lim_{k \rightarrow \infty} \ln \left( \frac{k(k+2)}{(k+1)^2} \right) = \ln(1) = 0$$

Hence,

$$\begin{aligned} \frac{k_1(k_1+2)}{(k_1+1)^2} &= \exp \left( \frac{-1}{Nm(m+1)} \right) \\ \implies (1-A)k_1^2 + 2(1-A)k_1 - A &= 0 \\ \implies k_1 &= -1 + \sqrt{1 + \frac{A}{1-A}} \end{aligned}$$

where  $A := \exp \left( \frac{-1}{Nm(m+1)} \right) < 1$  since the argument in the exponential is always negative.

### IV.2 Random Attachment Case

Again use the requirement that

$$P(k > k_{max}) \stackrel{!}{=} \frac{1}{N}$$

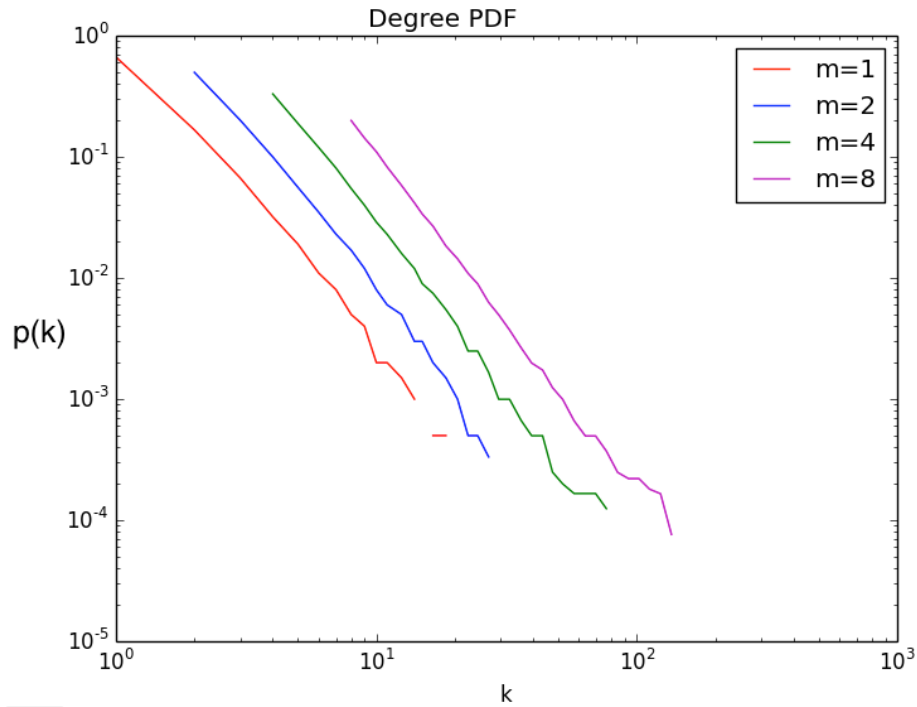
So that

$$\begin{aligned}
 1/N &= \sum_{k=k_1}^{\infty} p_{\infty}(k) \\
 &= \frac{1}{m+1} \sum_{k=k_1}^{\infty} \left( \frac{1+m}{m} \right)^{m-k} \\
 &= \left( \frac{1+m}{m} \right)^{m-k_1} \frac{1}{m+1} \sum_{k=0}^{\infty} \left( \frac{1+m}{m} \right)^{-k} \\
 &= \left( \frac{1+m}{m} \right)^{m-k_1} \\
 \Rightarrow k_1 &= \frac{\ln \left( N \left( \frac{m+1}{m} \right)^m \right)}{\ln \frac{m+1}{m}}
 \end{aligned}$$

### III. RESULTS

#### I. Preferential Attachment Case

##### I.1 Degree Distribution

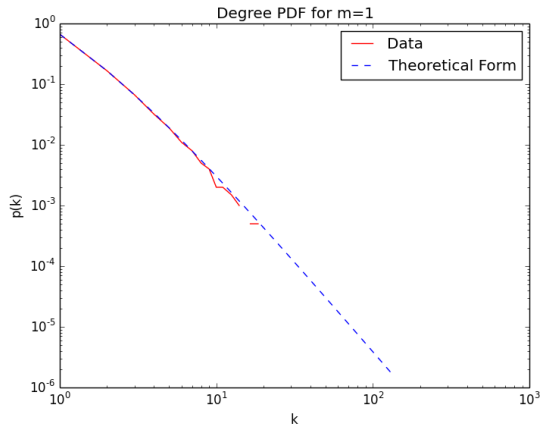


**Figure 1:** Degree distribution for preferential attachment, for different values of  $m$ , log-binned to reduce the bad statistics caused by the fat tail.

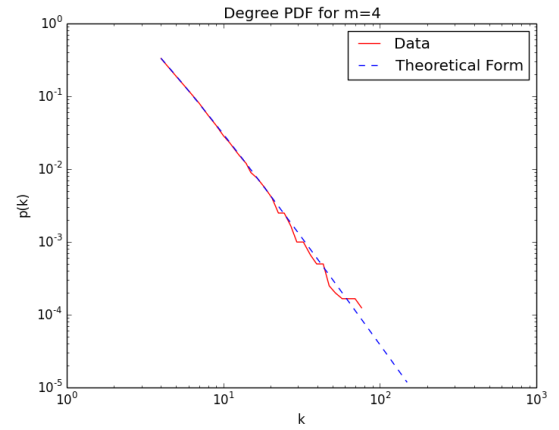
1

The results were obtained choosing the parameters,  $m = 1, 2, 4, 8$ , so that the program samples a reasonable scale in  $m$ , whilst not taking too long.

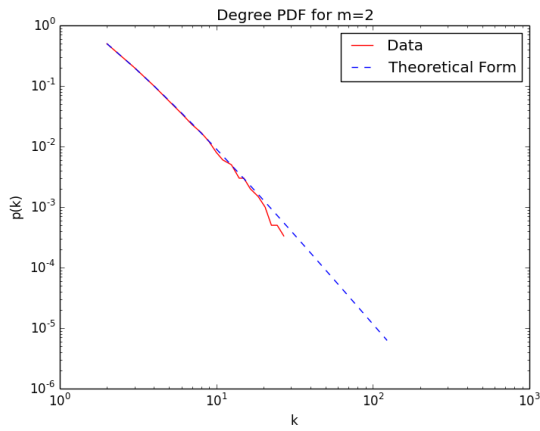
<sup>1</sup>see appendix



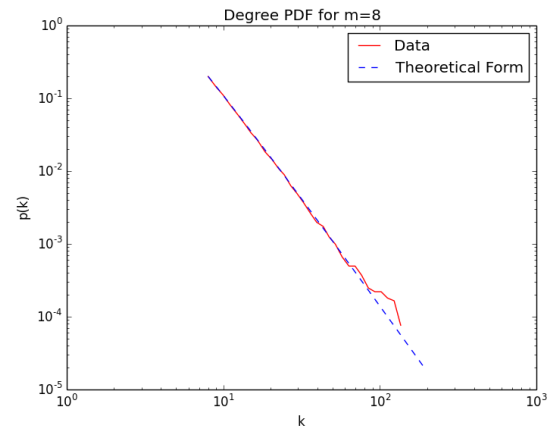
**Figure 2:** Degree distribution for preferential attachment, with theoretical values also plotted for  $m = 1$ , log-binned to reduce the bad statistics caused by the fat tail.



**Figure 4:** Degree distribution for preferential attachment, with theoretical values also plotted for  $m = 4$ , log-binned to reduce the bad statistics caused by the fat tail.



**Figure 3:** Degree distribution for preferential attachment, with theoretical values also plotted for  $m = 2$ , log-binned to reduce the bad statistics caused by the fat tail.



**Figure 5:** Degree distribution for preferential attachment, with theoretical values also plotted for  $m = 8$ , log-binned to reduce the bad statistics caused by the fat tail.

Compared to the theoretical result, the data matches, with some fluctuations at the tail end of the distribution. This shows that there is indeed no typical scale for the degree. A  $\chi^2$ -test was performed and it was found that all the p-values were 1.0 up to the accuracy of the variable type, hence there is no reasonable doubt that the theoretical model matches the simulation.

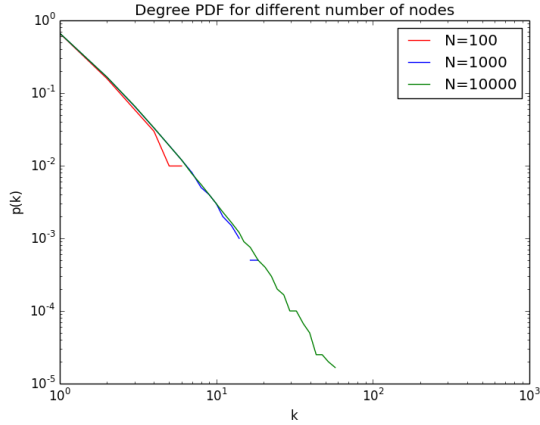
$m$	$\chi^2$	Degrees of Freedom	p-value
1	3.40391064383	35	1.0
2	2.30599926631	34	1.0
4	0.8779577978	34	1.0
8	0.757281441045	33	1.0

## I.2 Finite-Size effects

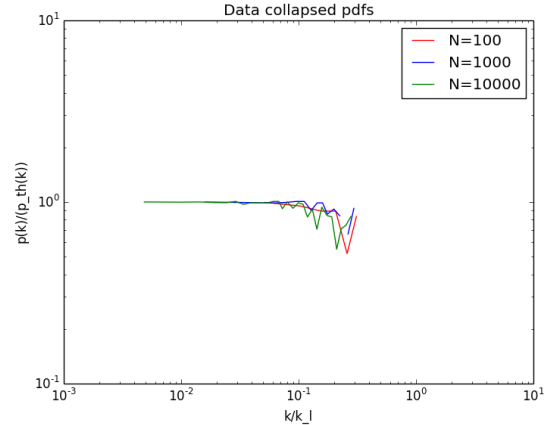
Looking at the distribution for varying numbers of nodes, it's found that the mean-averages of the maximum degree do not correspond to the theoretical values, however, the modal-average is close to the predicted values:

$N$	$\mu(\hat{k}_l)$	$\sigma(\hat{k}_l)$	mode	median	predicted value
101	19.1174	0.000484533360565	16	18	13.23
1001	63.28	0.0168545389732	43	59	43.75
10001	221.9	4.94349066956	144	190	140.43

The differences between the predicted values and the observed values could be due to the high kurtosis and skewness of fat-tailed distributions, as exhibited by the data. There is no typical scale for degree distribution in the limit that the network size goes to infinity.

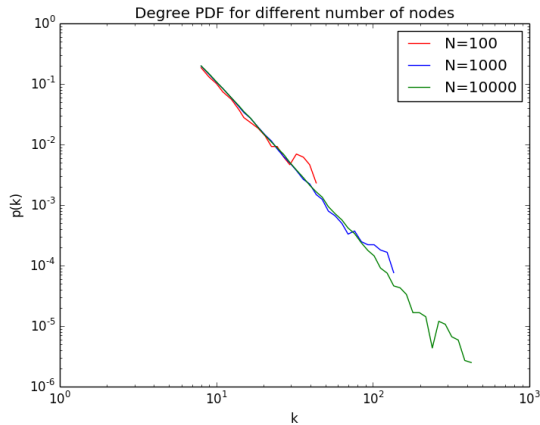


**Figure 6:** Degree distribution for different values of  $N$ , for  $m = 1$  log-binned to reduce the bad statistics caused by the fat tail. Averaged over multiple iterations.

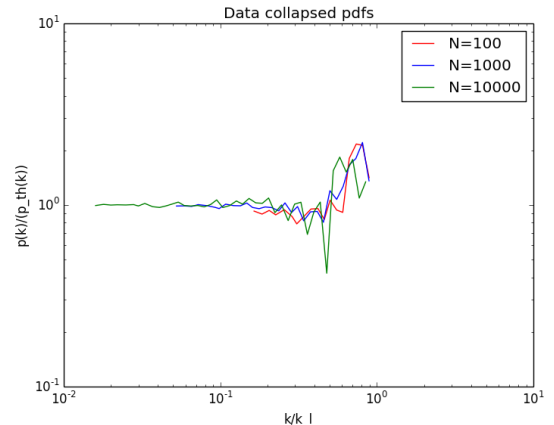


**Figure 7:** Collapsed degree distributions, for  $m = 1$

The degree distributions can be collapsed by plotting  $p_{data}/p_{theoretical}$  vs  $k/k_1$ , since the tail cut off of the distribution scales by the largest value.



**Figure 8:** Degree distribution for different values of  $N$ , for  $m = 8$  log-binned to reduce the bad statistics caused by the fat tail. Averaged over multiple iterations.



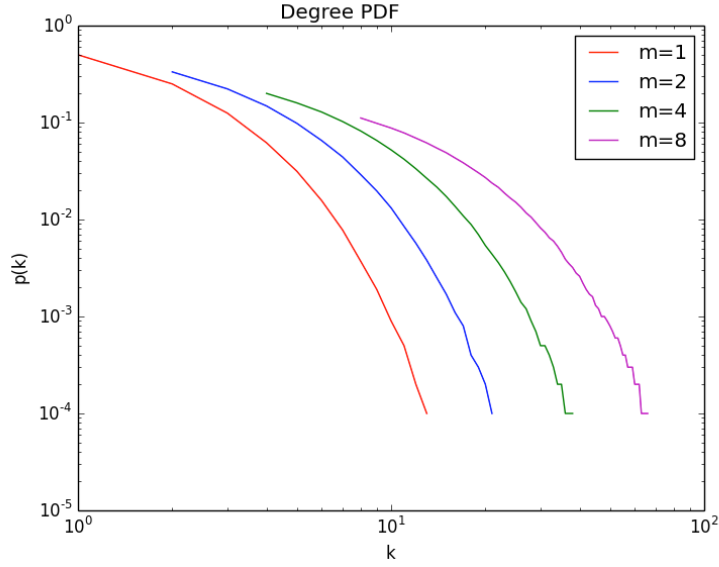
**Figure 9:** Collapsed degree distributions, for  $m = 8$

For larger values of  $m$ , it becomes apparent that the distribution rises for the the tail ends of the data. This occurs with all 3 profiles and thus it cannot simply be ignored and explained due to the poor statistics at the end of the fat-tailed distributions. These are caused by nodes which would otherwise have had a larger degree, if the system size was not limited. This effect is only noticeable if the network is growing fast enough, so for  $m = 8$  the effect is visible. On the profile of the collapsed pdf for  $m = 8$ , all the bumps are increasing by the same order of magnitude with respect to the scaling behaviour of the largest degree, and the magnitude of the probability.



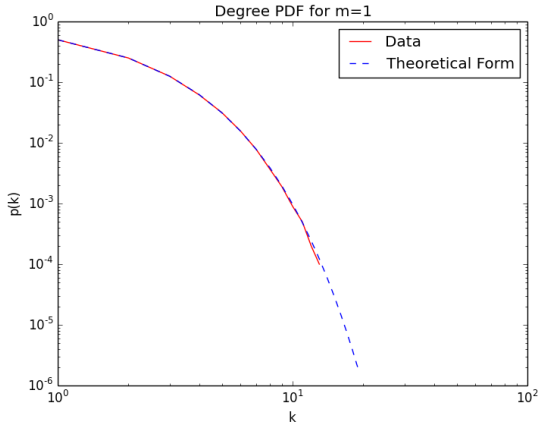
## II. Random Attachment Case

### II.1 Degree Distribution

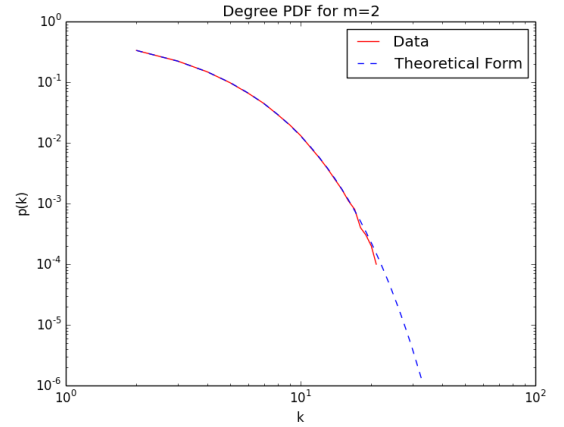


**Figure 10:** Degree distribution for random attachment, for different values of  $m$ , log-binned to reduce the bad statistics caused by the fat tail.

1

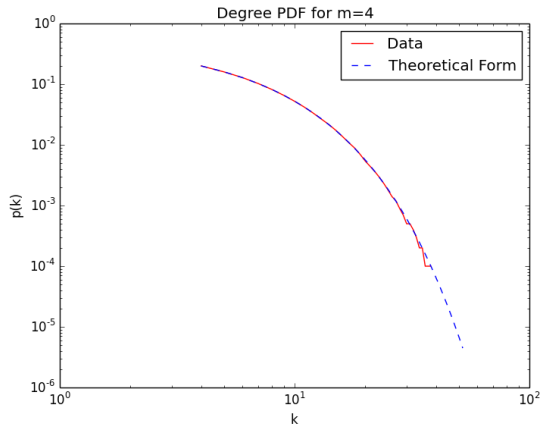


**Figure 11:** Degree distribution for random attachment, with theoretical values also plotted for  $m = 1$ , log-binned to reduce the bad statistics caused by the fat tail.

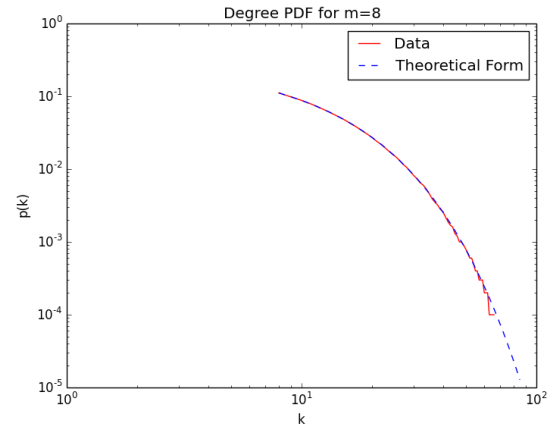


**Figure 12:** Degree distribution for random attachment, with theoretical values also plotted for  $m = 2$ , log-binned to reduce the bad statistics caused by the fat tail.

<sup>1</sup>see appendix



**Figure 13:** Degree distribution for random attachment, with theoretical values also plotted for  $m = 4$ , log-binned to reduce the bad statistics caused by the fat tail.



**Figure 14:** Degree distribution for random attachment, with theoretical values also plotted for  $m = 8$ , log-binned to reduce the bad statistics caused by the fat tail.

$m$	$\chi^2$	Degrees of Freedom	p-value
1	1.59266343244	18	0.99999826223
2	2.30599926631	29	0.99999996191
4	0.8779577978	47	1.0
8	0.757281441045	77	1.0

Visually it can be seen from Figures 9 – 12 that the simulation behaves exactly as expected. Doing a  $\chi^2$ -test gives  $p$ -values close to 1, beyond any reasonable doubt that the theoretical forms are incorrect, since the p-values are above any of the standard thresholds of significance.

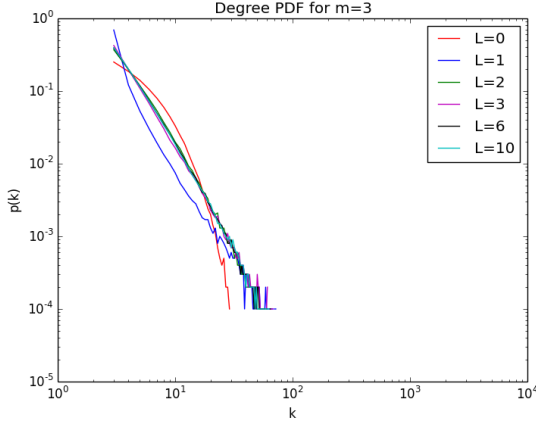
## II.2 Largest Degree

For  $N = 10000 + m$ , the predicted values versus the obtained values:

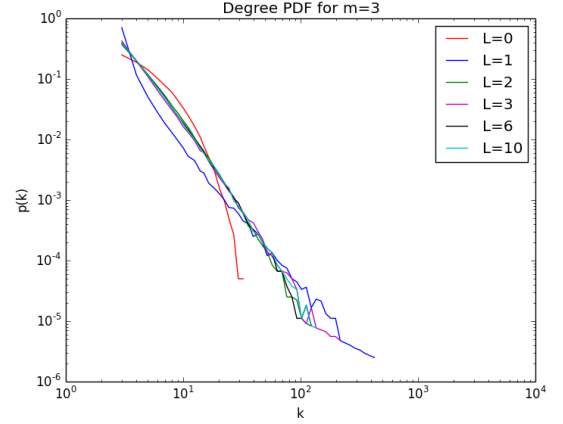
$m$	$\mu(\hat{k}_l)$	predicted value
1	14.11	14.29
2	24.51	24.72
4	42.36	45.28
8	73.77	86.20

The first three values match the theory, the last value does not, however the standard deviation increases with  $m$ , and the larger  $m$  was averaged over a number iterations since it takes the longest to run. These values correspond to theory better than the mean average values for the preferential attachment case, because this time the data is not highly skewed, and there is some typical degree that can be defined by the mean.

### III. Random Walks



**Figure 15:** Degree distribution for using random walks as the selection method, for different values of  $L$



**Figure 16:** Degree distribution for using random walks as the selection method, for different values of  $L$ , log-binned to reduce the bad statistics caused by the fat tail.

For the case  $L = 0$ , the degree distribution has an exponential cut off, as opposed to a fat tail. This is expected since the selection is essentially random, so it should look like the random attachment case. For  $L = 1$ , the distribution curves towards the fat-tailed distributions, and appears to be even more skewed and with a higher kurtosis than the fat tailed distributions, this could be because for higher  $L$ , the random walker can also leave the nodes with high degree, whereas for the  $L = 1$  case, the random walker only steps towards any node attached to it, which is statistically likely to be a node with high degree. For higher  $L$ , the distribution is equally fat tailed, suggesting the distribution is of the same form as with preferential attachment, where the node at which the random walker ends up must be proportional to  $k$ .

### IV. DISCUSSION

The data was sufficient to test the theoretical forms of the distributions. However for comparing the largest degree, it would be useful to compare the predicted values against a larger number of data points.

None of the results behaved in such a way in which it could not be reconciled with the theoretical predictions of the Barabassi-Albert model or using pure-random attachment.

### V. CONCLUSION

In conclusion, the Barabassi-Albert model and the pure random attachment model were tested against theoretical forms of the degree distribution, high p-values as a result of the  $\chi^2$  test suggests that neither of these contravened the theoretical predictions. the Barabassi-Albert model does indeed give rise to fat tailed distributions for which the degree is scale free. The random-attachment model exponentially cuts off for high degrees. The predicted values of the highest degree matched the observed values, taking the modal average for the Barabassi-Albert model and the mean average for random attachment.

Data collapse has revealed that there is a finite-size effect that becomes noticeable if the network is growing fast enough. These are values of the degree probability which are higher than predicted in the  $t \rightarrow \infty$  limit. These are caused by nodes for which there would have been a higher degree if the system size was larger.

Using random walks as the selection for attaching to a new edge, gives rise to fat tailed distributions for the length of the walk greater than 1. For walks of length 0, it is equivalent to random attachment, since the first node of the walk is randomly selected. For walks of length 1, the data shows that the skewness is higher than in the fat-tailed case, this may be because for walks greater than length 1, the random walker also has a chance to leave the first selected node.

## VI. APPENDICES

### I. Data Binning

A probability mass function of the degree  $k$  can be estimated by:

$$P(k) = \lim_{N \rightarrow \infty} \frac{N_k}{N}$$

Where  $N_k$  is the frequency in which the degree  $k$  occurs. Because  $N$  is always finite, there will be clusters of points where  $P_{\text{estimated}} = 1/N$ , this is the minimum that is possible, but the theoretical probability  $P(k)$  could have a lower value.

Data binning can be used to map the data onto the actual values. The s-axis is divided into bins, where each bin spans:

$$\mathcal{B}_j = [a^j, a^{j+1}[, \quad j \in \mathbb{Z}^+ \cup \{0\} \quad a > 1$$

If the data  $\hat{k} \in \mathcal{B}_j$  occurs, it's recorded as contributing towards the frequency of  $\mathcal{B}_j$ . The estimated probabilities are then transformed thusly:

$$\begin{aligned} \tilde{P}(k_j) &= \frac{N_j}{N\Delta k} \\ \Delta k &:= k_{\max}^j - k_{\min}^j + 1 \\ k_j &:= \sqrt{k_{\max}^j k_{\min}^j} \end{aligned}$$

Where  $k_{\min}^j$  is the minimum integer value of  $\mathcal{B}_j$ ,  $k_{\max}^j$  is the maximum integer value of  $\mathcal{B}_j$ .

In practice  $k_{\min}^j, k_{\max}^j$  was calculated via:

$$\begin{aligned} k_{\min}^j &= \lfloor a^j \rfloor + \delta_{\lfloor a^j \rfloor \lfloor a^{j+1} \rfloor} = \begin{cases} \lfloor a^j \rfloor & \lfloor a^j \rfloor = \lfloor a^{j+1} \rfloor \\ \lceil a^j \rceil & \lfloor a^j \rfloor \neq \lfloor a^{j+1} \rfloor \end{cases} \\ k_{\max}^j &= \lfloor a^{j+1} \rfloor \end{aligned}$$

And the bin number to which a data point should belong is calculated as:

$$j = \lfloor \log_a \hat{k} \rfloor$$

Using a dictionary to record the frequencies, the primary keys being the bin numbers. Bins for which there are no integer values (e.g.  $[1.1, 1.21[$ ) do not exist, in other words they are never included in the data set, otherwise there would be zeros mixed in with the higher values. This has the effect of the apparent loss of detail for low values of  $s_j$  and causes the oscillations near the beginning of the plots in the graphs.

## REFERENCES

- [1] Evans, Tim. *lecture Notes, Networks*
- [2] Evans, Tim, Christensen, Kim *Complexity and Networks Project Notes*  
Imperial College London 2014