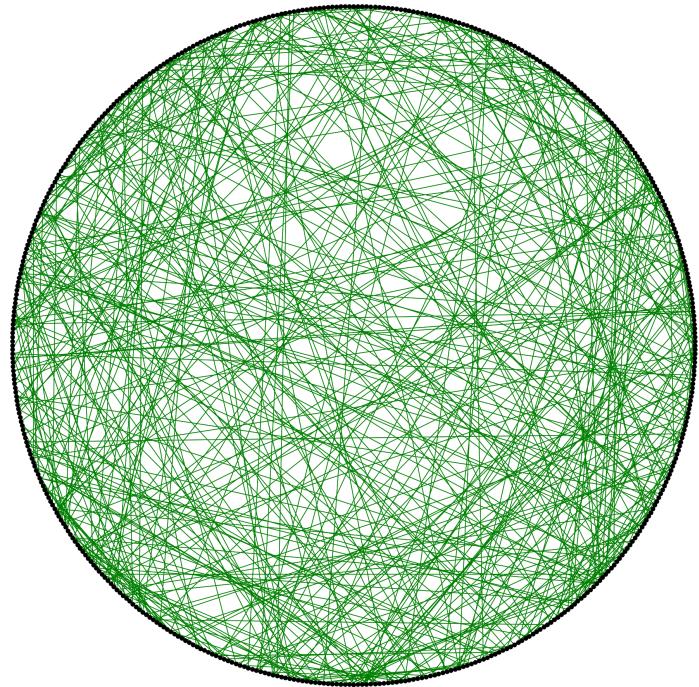


---

A study of some  
**Real networks and random graphs**

---



Second assignment of the course *Introduction to complex systems*

Breebaart Rik  
Cristinelli Giacomo

Haoran Dai  
Vedder Casper

# 1 Introduction

Over the last years there has been a tremendous amount of interest in Complex Networks. This is not without good reason, as in some sense complex networks dictate modern society. We are all part of a complex societal network, or when one hops on the train from Utrecht to Amsterdam you partake in the intricate transport network and when you switch on the light you use electricity coming from the energy network. For this sake studying complex networks is easily justified.

However, as these systems are widely different in many regards, it is thus important to formalize the notion of a *network* to see what unites them.

## 1.1 Networks and Graph Theory

The notion of a network can be formalised as a *graph*. A graph consists out of a set of nodes:

$$V := (v_1, \dots, v_n)$$

In a social network these would then for example be different persons. To make the network complete one also introduces a set of edges:

$$E := \{(v_{i_1}, v_{j_1}) \dots (v_{i_m}, v_{j_m})\}$$

Again, for a social network these would then be the interactions between people. Note that this is only an example, in principle the nodes and edges can be anything. One could in principle make the edges quite complicated. For example one can attach a relative *weight* to the edge, making some connections more important than others. Edges can also be either directed or undirected, undirected graphs are graphs that have the property that if there is a connection  $v_i \rightarrow v_j$ , there is an equally important connection  $v_j \rightarrow v_i$ . In Figure 1 we have shown a unweighted directed network, the directedness can be observed by the arrows on the edges. However, for practical purposes we will mostly be focusing on graphs that are both unweighted and undirected.

To analyse networks it is often convenient to construct an *adjacency matrix*. This is a matrix where the  $i, j$  element gives the weighted value of the edge  $v_i \rightarrow v_j$ . For unweighted undirected networks the adjacency matrix is thus a symmetric matrix containing only zeroes and ones. The diagonal contains the interaction of nodes with themselves.

With this information we can now define some characteristic statistics for networks. To begin, it is useful to define the density of an undirected network with  $n$  nodes:

$$D := \frac{\text{Number of edges}}{\text{Maximal number of edges}} = \frac{2|E|}{n(n-1)} \quad (1)$$

Where  $|E|$  is the number of edges. Heuristically it makes sense that the density will usually be a small quantity. Again taking the example of a social network, we find it intuitive that this is so. When considering the world, maintaining  $\approx 7$  billion friends would probably be quite tiring. If the density of a network is small, the adjacency matrix is a so called *sparse matrix*. This will be useful for computational reasons. We then define the *degree*  $k$  of a node, as the amount of edges it has. In terms of the adjacency matrix this is just summing over the row it lives in i.e.  $k_i = \sum_j a_{ij}$ .

Another interesting statistic is the *mean shortest path*. The distance between nodes is given by  $l(v_i, v_j)$ , which counts the numbers of edges need to be passed from  $v_i$  to  $v_j$ . For networks with  $n$  nodes the average of this number is then:

$$\ell := \frac{2}{n(n-1)} \sum_{i,j} \ell(v_i, v_j) \quad (2)$$

The last statistic we need to introduce is the clustering coefficient. Which measures how the edges are centered around a node. It is defined as:

$$C_i := \frac{\# \text{ of edges between neighbors}}{k_i(k_i-1)/2} = \frac{1}{k_i(k_i-1)} \sum_{j \neq k} a_{ij} a_{ik} a_{jk} \quad (3)$$

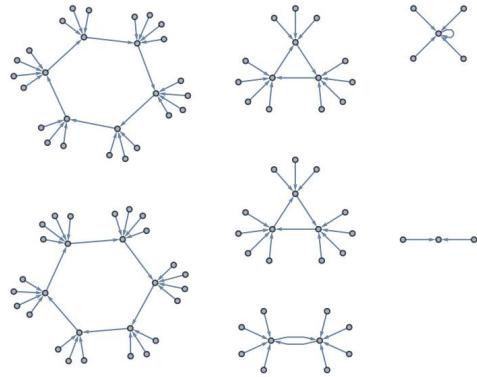


Figure 1: A network produced by having the edges generated by  $i \rightarrow i^2 \bmod 43$  with  $i \in (1, 100)$ .

and indicates the fraction of possible triangles that exist through node  $i$ . An important global statistic of the network is then the average clustering given by:

$$\langle C_i \rangle := \frac{1}{n} \sum_{i=1}^n C_i \quad (4)$$

Now armed with our formalism we can actually consider some models, as we will discuss hereafter.

## 1.2 Models of Complex Networks

As a way of explaining real life models, lots of toy models were constructed. In this subsection we will briefly discuss some of them and the intuition they give. A natural candidate for models is a model with random components, a so called Random Networks. A famous random model is the Erdos-Renyi model [1]. This is a model where between all nodes there is a chance  $p$  an edge is formed. These models then have a binomial distribution as their degree distribution:

$$\mathbb{P}(k = m) = \binom{n-1}{m} p^m (1-p)^{n-1-m} \quad (5)$$

From this then follows that  $\langle k \rangle = p(n - 1)$ . Another feature of this is that for sparse networks the probability distribution will tend towards a Poisson distribution. Random networks are interesting for several reasons, but also because they contain the small world phenomenon.

This is a feature of some networks that was noticed by Milgram in 1967 [2]. Milgram found experimentally that in a certain social network in the US the average smallest path length  $\ell$  was remarkably small, the desired node could usually be reached within 6 steps. This feature is also seen very well in academic networks. A famous example of this is the *Erdos number*, this is a number that shows the collaboration distance between some mathematician and famous mathematician Paul Erdos, which among working mathematician has a median of 5 [3].

Random networks also contain this feature, having a very small average smallest path length, which intuitively makes sense due to random nature connecting nodes all over the network. However, random networks usually have a very small clustering coefficient, as this scales inversely with the amount of nodes. This is not a feature shared by many real life networks.

This motivated Watts and Strogatz in 1998 to develop the *Small World* or *Watts-Strogatz model* [4]. Their model contains random aspects, which can lead to the small world phenomenon. But to accommodate the high clustering seen in many real life networks, it also contain features of a *regular lattice*. The Watts-Strogatz model starts with a ring lattice, a lattice where the nodes are connected to their neighbours on the lattice, then with some random process some edges are rewired to random other nodes. Tweaking the values of the system then determines where it lies on the spectrum between randomness and a regular lattice. This gives very interesting statistics as these models can contain both very small average smallest path lengths, but also have a high degree of clustering. It therefore mimics some aspects of real life networks very well. A notable downfall of this model is, however, that it produces a degree distribution that is not very similar to real life models.

A statistic where both models discussed fail to match up with many real life model is the degree distribution. Many real life models exhibit a degree distribution that looks very much like a power law:

$$p_k \sim k^\gamma$$

As power laws look the same on each scale considered<sup>1</sup> these degree distributions are known as scale free. An actually quantitatively good model would thus also have to contain this feature. A model that has this feature is the Albert-Barabasi model [5], however this model does not have a realistic clustering. It is fair to see that these models cannot be called realistic, however they still give valuable insights in the workings and mechanisms of complex networks.

In this report we will study various models more in depth. First we will dive into the data and study various real life models, after this we switch gears and go on to the modeling discussing some features of the Watts-Strogatz model more in depth.

---

<sup>1</sup>This is actually unique for power laws.

## 2 Real Networks

In this section we will be studying some real networks. We have taken these networks from the Stanford database [6]. For our analysis we have chosen to study two widely different networks: One we expect to be scale free, while we expect the other one not to have this feature.

### 2.1 Astrophysics and Roads

For the scale free network, we consider the *astrophysics* collaboration network [7]. This network is obtained from <https://arxiv.org/archive/astro-ph>, using data between January 1993 and April 2003. We expect this network to be scale free, due to the collaborative nature of astrophysics. Astrophysicist have to work together to analyse their gigantic data sets, therefore there will be nodes having a large amount of edges (hubs). Especially the more successful astrophysicist are likely to have lots of collaborators. Hubs are also a feature of scale free networks, which strengthens our suspicions that this network could be scale free. It also goes without saying that especially in academia it is often lonely at the top. Because of this most astrophysicist are likely to just work together with their few close collaborators, as not everyone can be part of the large collaborations. This would then result in a distribution with most of the spectrum at lower  $k$ , it then makes sense that the spectrum will decay towards higher  $k$ . A power law contains these features. The other network we will be considering is widely different, namely it is the road network of California [8], which we will refer to as Road-CA. In this network the nodes are the endpoints and intersections of the roads. The roads are then given as undirected edges. We expect this to look more like a Poisson distribution, or at least we do not expect it to be scale free. As scale free networks have hubs, for road networks this would be very impractical. A roundabout with about 6 roads connected to it is already quite complicated, so we expect there to be very little nodes with a very large amount of edges. Also intuitively it makes sense that the degree distribution would peak somewhere around 4, as we expect that usually when two roads cross, this would give two roads going in and two roads going out (see for example Figure 2). Thus we expect a peaked distribution that quickly decays with  $k$ , sounding more like a Poisson distribution than a power law. However Poisson networks are usually related to randomness in the system, road networks are constructed with reasons so are probably quite ordered. For this reasons a Poisson distribution might not be suitable to describe this.

### 2.2 Investigating the Networks

Both networks are obtained as a list of edges indicated by their current node  $i$  and the node  $j$  to which they are directed. These connections can be described using an adjacency matrix  $A$  in which the non-zeros values  $a_{i,j}$  of the matrix indicate the edge from  $i$  to  $j$ . In our case the networks we will be investigating are undirected and weightless. Therefore in the case that  $a_{i,j}$  is non-zero then  $a_{i,j} = 1$  and for all entries in the adjacency matrix we have that  $a_{j,i} = a_{i,j}$ . The adjacency matrix corresponding to the networks will thus be symmetric along the diagonal.

In many cases of complex networks, the adjacency matrix of the graph can become incredibly large. This is also the case for the two networks chosen. In Table 1 a selection of attributes of the two networks is shown. Here we can see that the astrophysics network contains  $N = 18872$  nodes and the Road-CA network contains  $N = 1965206$  node. Which in the case everyone would be connected to everyone would result in  $N^2$  edges (also including self-connections). This is however not the case. Not all astrophysicists interact with each other and not all roads are directly connected to all other roads (then we would have intersections with 1965206 roads). Instead each node only has a small number of edges to other nodes. In the case of the astrophysics network there are a total of  $|E| = 198110$  undirected edges which thus results in a total of  $2|E|$  non-zeros in the adjacency matrix. For the Road-CA network there are a total of  $|E| = 2766607$  edges. With this in mind we can look at the sparsity of the adjacency matrices

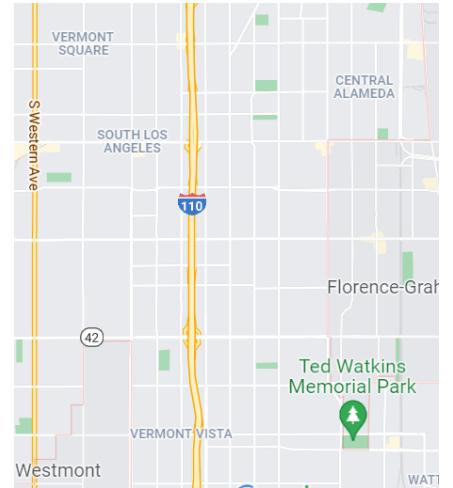


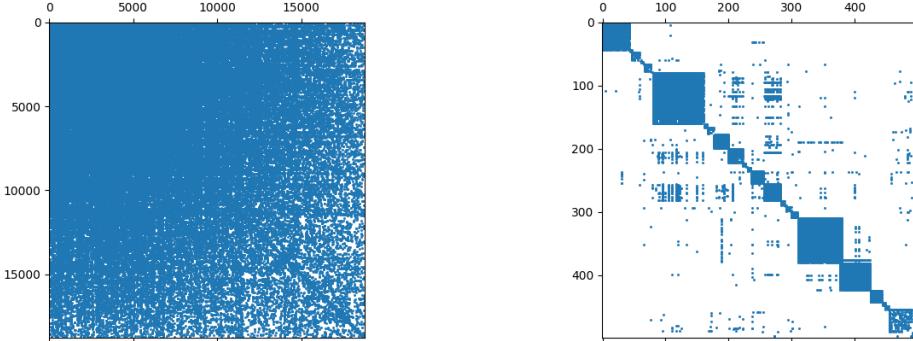
Figure 2: Overview of a road network, in this case of a section of the Californian city of Los Angeles, taken from Google Maps.

	Astrophysics	Road-CA
Number of nodes	18872	1965206
Number of edges	198110	2766607
Density	1.12e-03	1.43 e-06
Average degree	21.10	2.82
Number of self-connections	60	0
Average Clustering	0.63	0.046
Average degree of friends	65.33	3.17

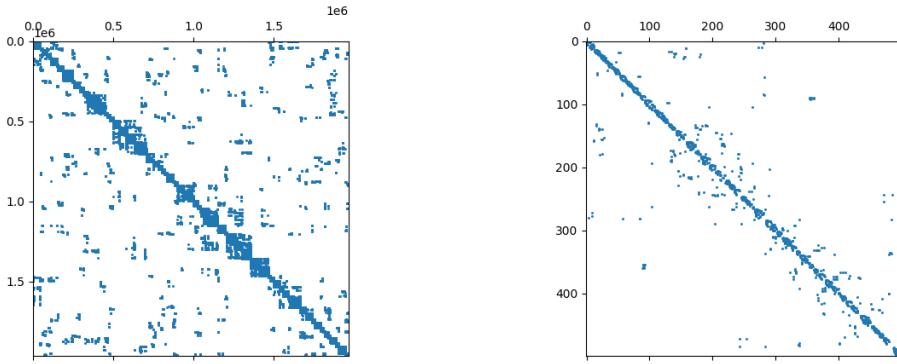
Table 1: Table of attributes of the astrophysics and Road-CA networks.

corresponding to these networks. We can do this by looking at the density of the network which is computed according to eq.1. In table 1 we can see that the density of the two networks is incredibly low and thus the network is very sparse. However, the road map is much more sparse when compared to the astrophysics network. This is also reflected in the average degree of the two networks. On average an astrophysicist collaborates with 21.10 peers while the road connections in California are on average connected to 2.82 roads. These differences are easy to understand as the number of roads which connect at a certain junction are usually around 2 to 4 while the number of collaborators in an astrophysics papers might differ by a large amount. Some important experimental papers might be part of a collaboration containing multiple universities and many researchers while others might only have a few collaborators. In figure 3 the adjacency matrices corresponding to the two networks are shown. Here on first sight the adjacency matrix corresponding to the astrophysics network does not seem sparse (see fig.3a). This first observation does not reflect the results we have obtained for the density and average number of degrees. However, taking a closer look at the matrix (see fig.3b) we can see that the matrix is indeed sparse. The matrix did not seem to be sparse due to the large number of nodes and the number of edges where the blue pixels indicating the non-zeros were too large for the image to show the full details of the matrix. Looking at the first 500 nodes of the matrix we can see that most connections are close to the diagonal. This reflects that collaboration happens more often between closely related peers. This would be the case if the data was obtained by "crawling" through papers. The "crawler" would then start at a certain paper, store the authors of that paper and their collaborators as nodes and then "crawling" to the papers of the collaborators and repeating this process. In that case the nodes of two researchers which collaborate would be close together and thus more non-zeros near the diagonal would form. In this manner we would also expect to see large clusters (regions of high density) due to authors in a certain field collaborating with each other. This can also be seen in fig.3b where cluster regions of higher density can be observed near the diagonal.

As opposed to the astrophysics matrix, the full adjacency matrix of the Road-CA network does not seem dense (see fig.3c). In its adjacency matrix we can clearly see the sparsity of the matrix. We can also see that the connections are all close to the diagonal. This would be expected for a road network as there is a high likelihood that two roads close together (and most likely labeled close together) are connected. While roads which are far apart are usually not connected. We can also see some regions of higher density near the diagonal. This might reflect the road networks of cities in which more roads are closely packed together and connected to each other. For example a city block will result in square network connections. To further investigating the clustering of the two networks we can look at the clustering coefficients. In table 1 the average clustering coefficient is shown. This value is computed using eq.3 for the clustering coefficient and the average using eq.4. To compute this, the sum over the indices of the adjacency matrix can be performed using matrix multiplication. For which we will only look at the diagonal of the matrix  $B^3$  where  $B$  is the adjacency matrix  $A$  with the diagonal set to 0 to remove self-connections. From the results we can see that the clustering of the astrophysics network is higher than that of the Road-CA network. This can be easily understood by the difference between the two networks. The astrophysics network describes collaborations for which it is much more likely to have people working together when working on a similar subject. Resulting in triangles of connections. While for roads it is much less likely that a triangle structure occurs. That is, since this does not allow for convenient junctions. Next to that we are considering the road network of California. In America roads are often ordered in a square checker board sort of structure (see Figure 2), which also explains the data.



(a) Adjacency matrix of the astrophysics network. (b) Adjacency matrix corresponding to first 500 nodes of the astrophysics network.



(c) Adjacency matrix of the Road-CA network. Full size. (d) Adjacency matrix corresponding to first 500 nodes of the Road-CA network.

Figure 3: Representation of the adjacency matrix of the astrophysics and Road-CA network. Where non-zeros are indicated in blue. Here for each of the networks two figures are shown. Figure (a) and (c) are the full matrix while figure (b) and (d) are the adjacency matrix corresponding to the first 500 nodes.

## 2.3 Degree distribution

The degree distribution contains a lot of information about the network. For example if there are any hubs present in it. The normalised degree distributions of the two networks are shown in Figure 4. We see that as we expected the Road-CA network is indeed very strongly peaked around a value of 4. It then decays rather quickly, also agreeing with our expectation. When considering the astrophysics network, we see that up until  $k \approx 3$  the values increases. Meaning that astrophysicist's usually do not work alone. After this peak it looks approximately scale free, which also agrees with our previously discussed expectations. This is seen most clearly in the log-log plot, where a power law would be seen as a straight line. However, a power law does not work for the whole distribution. As for very small  $k$ , the prediction of a power law fails badly, as there the distribution increases with  $k$ . Hubs are present in the astrophysics network. As from considering the distribution, we see that there are quite some nodes with a degree that is higher than 100, meaning more than 100 collaborators. This is also characteristic for a power law tail.

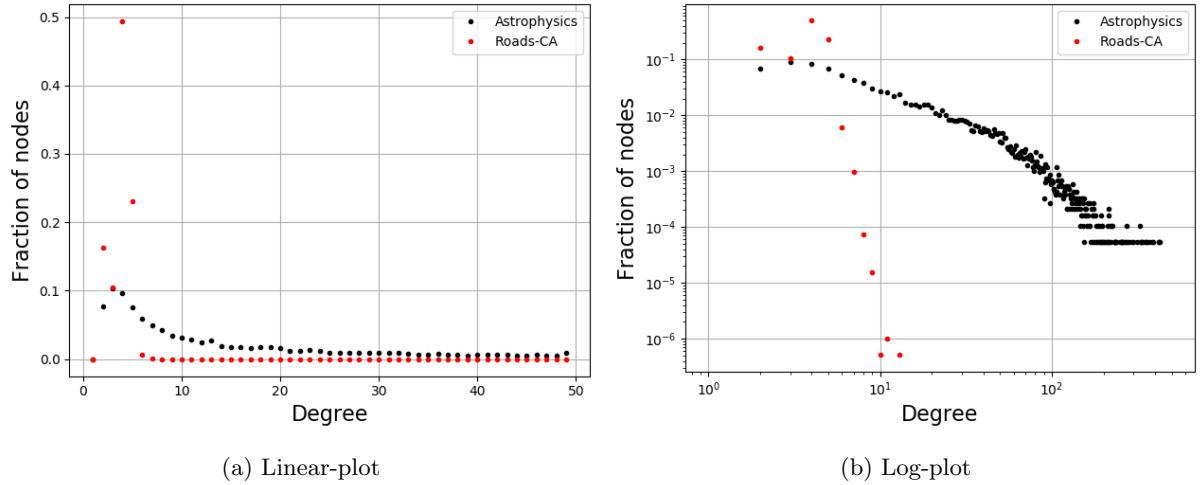


Figure 4: The degree distributions of the astrophysics network compared to the degree distribution of the Road-CA network. For a correct comparison both degree distributions are normalised to one. The linear plot ranges from  $k = 0$  to 50 as on this scale most interesting features appear on small  $k$ .

We reserve a proper analysis of the scale free or Poisson like behaviour for the later sections where we will fit these distribution to the data.

## 2.4 The Friendship Paradox

We will now consider the following paradox, known as the *friendship paradox*. It states that on average your friends will have more friends than you have. Which sounds like a very harsh truth, so we will study this more in depth before accepting this. In terms of networks this means the following: The nodes that a certain node  $v_i$  is connected with have on average more connections than node  $v_i$ .

We will investigate this in the following way using our two networks: First we will pick at random a node in our system, then we look to which nodes this node is connected and calculate the degree of these nodes. We will repeat this process  $10^4$  time and compute the average of the previous iterations for each step. We repeat this procedure 50 times and take the average from that. This is done to showcase that the procedure with an increasing number of iterations (or random nodes chosen) always converges to the same number. The evolution of this average during the iterations is shown in 5. Here we can clearly see due to the decrease of the standard deviation that the system converges to the same value for each repetition of the procedure when enough iterations (randomly sampled nodes) are taken.

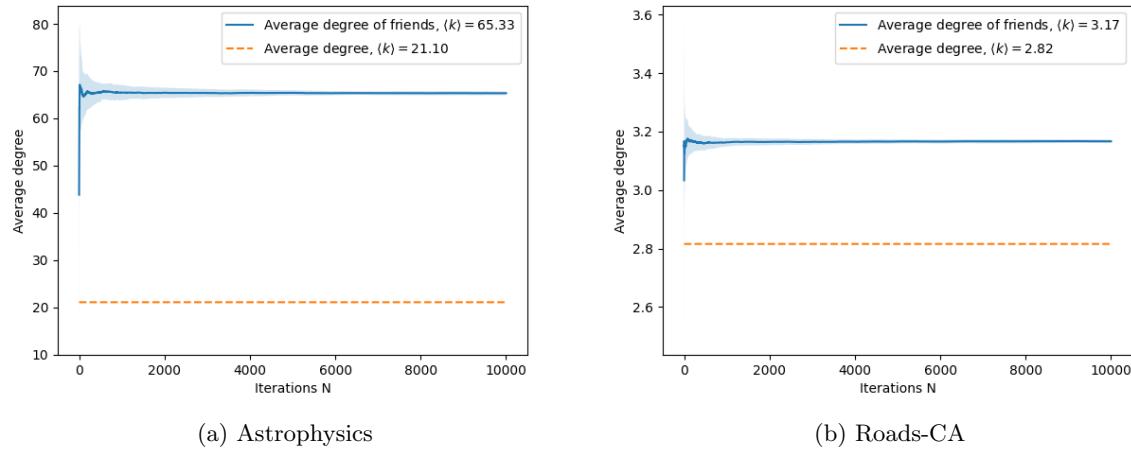


Figure 5: The average degree of friends compared to the average degree of the full matrix as a function of iterations. The analysis is done by adding the degree of a random nodes neighbours each iteration step and computing the average of all neighbours degree up to that point. This is done for  $10^4$  steps. Each point is thus an addition of the neighbours of a new randomly selected node. This process is repeated 50 times and the average is taken and a standard deviation is given. This is done to see the dependency on the initial random coordinate and to investigate the convergence.

In Figure 5 we clearly see the friendship paradox taking shape. For the astrophysics case the friends have 3 times more friends on average. While for the Road-CA network the friends degree is only slightly higher (friends being somewhat of a misnomer for road intersections).

The Friendship paradox can be explained though, it is in fact a simple sampling bias. Nodes with lots of edges will be counted more often, as they have more friends. In this calculation nodes with lots of edges are thus effectively weighted with the number of their friends. This means that very connected nodes are weighted very heavily, raising the average degree significantly (as is observed in Figure 5). Heuristically this means you are more likely to be friends with people having a lot of friends. There might be a lot of people having only a few friends, however the chance is small that you are one of these friends, as they only have a few. Another example of this sampling bias is the gym: When going to the gym, people at the gym tend to look fitter than you, as these fit people go to the gym more often, making you see them more often.

The difference between these two numbers also contains information, it is expected that for networks with lots of high degree nodes (hubs) this number will be higher, due to the friends with a high degree pushing up the number. As scale free networks usually have hubs in them, scale free networks likely have a large difference between these two numbers.

We can also investigate an alternative to the *friendship paradox* namely the statement that:

For a general graph with no isolated vertices. Where we uniformly pick a random vertex  $v_0 \in V$  and from that vertex a random vertex  $v_1 \in V$  which is a neighbour of  $v_0$ . Then for a sufficiently large sample set we will find that:

$$\mathbb{E}(k_{v1}) > \mathbb{E}(k_{v0}) \quad (6)$$

in which  $k_v$  denotes the degree of vertex  $v \in V$ .

To investigate this we performed the same procedure as with the *friendship paradox*. However, we first removed the isolated vertices of the network. Such that we obtained a graph with no isolated vertices. For each iteration we then uniformly selected a node  $v_0$  from the graph and then randomly select a single neighbor vertex  $v_1$  of this node. This process is then iterated  $10^4$  times and for each step the average degree of both  $v_0$  and  $v_1$  of the previous iterations is computed. This procedure is then repeated 50 times to properly showcase the convergence at increasing number of iterations. In figure 6 a plot showcasing the result for both the astrophysics and Road-CA network is shown. This is again computed using  $10^4$  iterations for each run and these are repeated 50 times to produce an average. Here we can see that for enough iterations the statement holds and that initially much more fluctuations occur. We can see that in the case of the astrophysics network the average degree of friend  $v_1$  is roughly twice that of the

average of  $v_0$ . While the Road-CA network again only has a slightly larger average for the degree of  $v_1$  compared to that of  $v_0$ . Although the isolated connections are removed from the network we still see that the connections of a friend are on average larger than that of a node itself. This is again since nodes with a large number of connections are selected more often as  $v_1$  since  $v_1$  has more friends who can select him. Resulting in a higher average for  $k_{v1}$ . Which can again be understood as you are more likely being friends with people already having a lot of friends. We can also see that comparing this *single friend* result to that of the *friendship paradox* that the average of the single friend  $v_1$  is lower than the average of all friends of  $v_0$ . This is since in the case a single friend is randomly chosen there is a lower chance that the hub (node with many friends) is chosen for the average. While when computing the average of all friends it is always added if the hub is a friend of  $v_0$ . We can clearly see this when comparing the astrophysics results however only a slight difference is found between the Road-CA results. Which again indicates that there are no large hubs in the Road-CA network while there are in the astrophysics network.

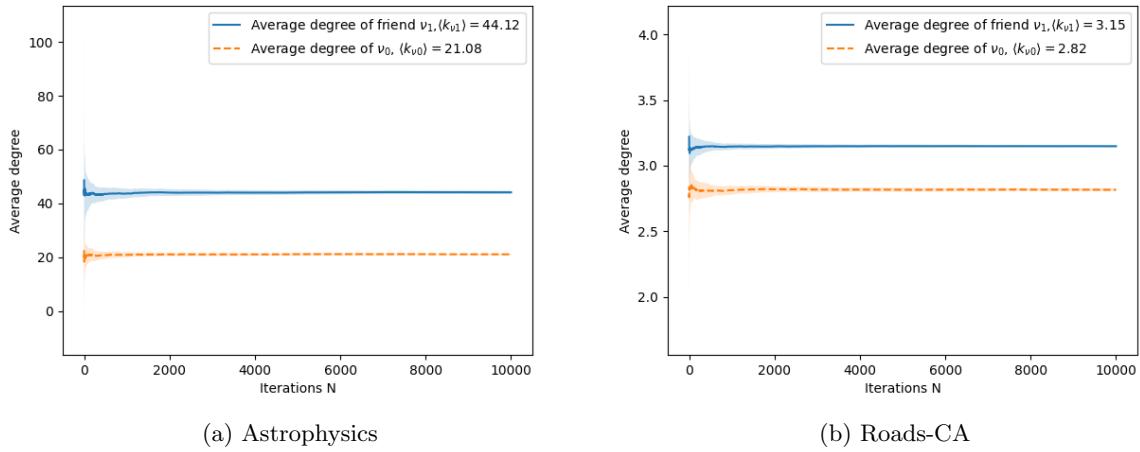


Figure 6: The average degree of a randomly selected friend  $v_1$  of  $v_0$  and the average degree of the randomly selected node  $v_0$ . The analysis is done by adding the degree  $k_{v1}$  of  $v_1$  which is a neighbour of the randomly selected node  $v_0$  for each iteration step and computing the average of all degrees up to that point. This is done for  $10^4$  steps. Each point is thus the addition of a neighbour of a new randomly selected node. This process is repeated 50 times and the average is taken. The standard deviation of the averages is shown in the regions around the average plot.

## 2.5 Poisson or Scale Free distribution

In the previous sections we have only qualitatively discussed the features of the distributions. However, to make this more concrete we will actually fit a power law or Poisson distribution to it. We do this by using the SciPy optimize function, we then use SciPy ChiSquare to calculate  $\chi^2$ . From this we then calculate the  $\chi^2_{red}$  which is our measure of how good the fit is. In this context  $\chi^2_{red} \approx 1$  means it is a good fit. We use  $\chi^2_{red}$  to calculate the  $p$  value, if this is smaller than 0.05 we reject the hypothesis. Concretely we will fit two functions to the data, with two free parameters. For the power law we use:

$$P(k) = Ak^\gamma \quad (7)$$

And for the Poisson distribution we use:

$$P(k) = A \frac{\lambda^k}{k!} e^{-\lambda} \quad (8)$$

In these functions  $A$ ,  $\lambda$  and  $\gamma$  are fitted for. The values of  $\lambda$  and  $\gamma$  are shown in the Figures 7,8 and 9. We will now go into the various concrete possibilities:

### Astrophysics and a Poisson distribution:

First we study the distribution of the astrophysics network. We start by taking the Poisson distribution as our null hypothesis. If we fit this distribution to the full network, this fit is very bad as it has a  $\chi^2_{red}$

of order  $10^{95}$ . This corresponds to a  $p$  value of about 0 and we thus reject this hypothesis. We can also take as our hypothesis that it fits the start of the distribution. This fit is shown in 7b. This results in a  $\chi^2_{red}$  of 8024, which is slightly better. However this also corresponds to a  $p$  value of about 0, thus this hypothesis is also rejected.

### Astrophysics and a power law distribution:

Now we consider as null hypothesis the power law. If we fit this to the full network the  $\chi^2_{red}$  is 45 (the fit is shown in 7a). This is significantly better than the Poisson fit, but still amounts to a  $p$  value of about zero so we reject this hypothesis as well. We also notice that the fit intersects the data. We estimate where it intersects it by minimizing the difference between the data points and the fit. However, as this data set has scatter, this is not a unique way of calculating this. Another way would be calculating where the difference between the data points and the fit switches sign. This, however, yields multiple points in the  $k$  range (25, 30), thus giving a slightly lower estimate. As this way is ambiguous we have chosen the other way of calculating it.

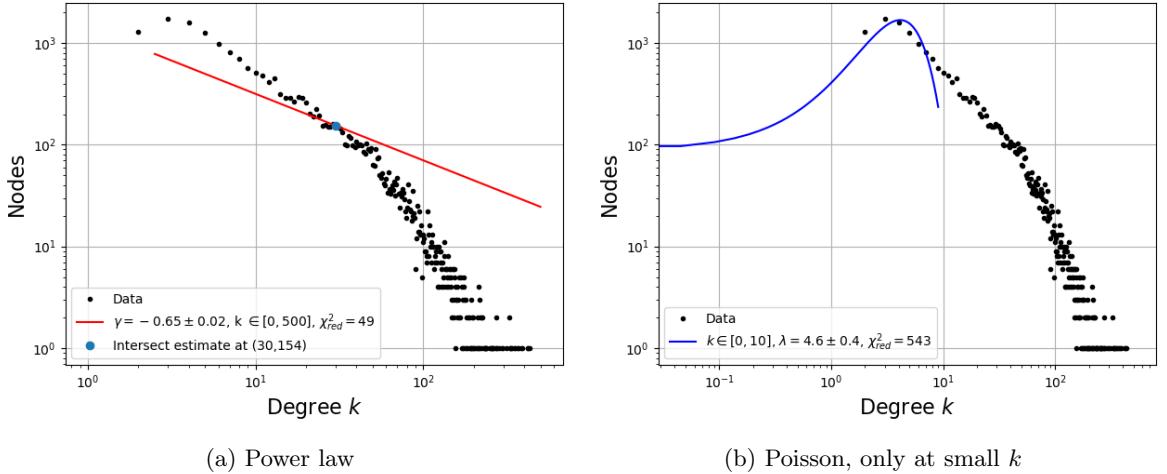


Figure 7: Both the power law as the Poisson distribution fitted to the full distribution of the astrophysics network, for the Poisson distribution only small  $k$  is considered. The legends show the  $\chi^2_{red}$  as well as the fitted parameters  $\alpha$  and  $\lambda$  with their uncertainties. For the power law we also give the coordinate where the fit intersects the data.

Instead of considering the full distribution, we can also consider the tail. This is motivated by the appearance of hubs, which is characteristic for scale free networks. We have shown this for various starting values of  $k$ , some of which are shown in Figure 8. These fits actually become quite good. For example, if we take as null hypothesis that the distribution ranging from 50 to 500 is a power law we get a  $\chi^2_{red}$  of 0.84, which has a corresponding  $p$  value of 0.97, indicating a good fit. Studying this more in depth we find that the distribution with starting values above  $k = 38$  pass our test. We thus do not reject the null hypothesis that the distributions with starting values above  $k = 38$  satisfy a power law distribution.

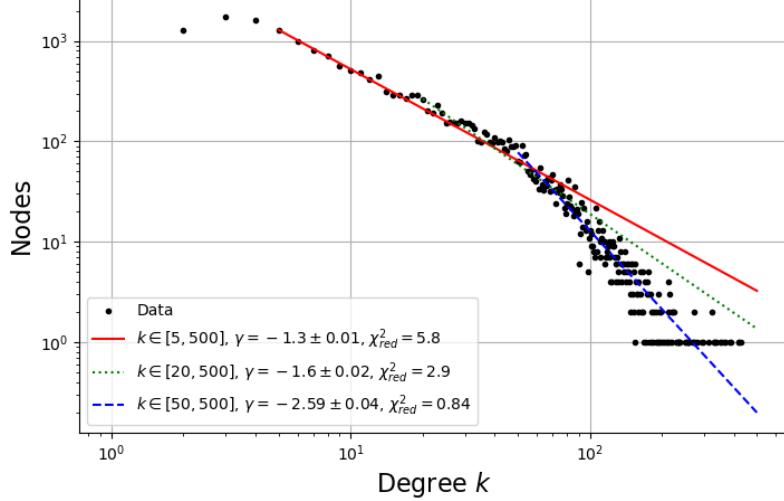


Figure 8: Power law fitted to the tail of the distribution, given on a log-log scale.

#### Roads-CA and a Poisson distribution:

Now we will consider the Road-CA network. We consider  $k$  ranging from 0 to 10, as above this there are no nodes. This network is very different, as we already discussed earlier. For example there are no nodes with a very large amount of edges. First we take as our null hypothesis the Poisson distribution. This fit is shown in Figure 9b and yields a  $\chi^2_{red}$  of order  $10^5$ , which means zero as  $p$  value. We thus reject this hypothesis. Varying this for several final or starting values of  $k$  does not change this.

#### Roads-CA and a power law distribution:

We can also take the power law as null hypothesis, this fit is shown in Figure 9a. We also reject this hypothesis as its  $\chi^2_{red}$  is also of order  $10^5$ , resulting in a  $p$  value of 0. Again we have investigated only considering the tail. For starting  $k = 5$  a fit is possible with the lower (although still not acceptable)  $\chi^2_{red}$  of 86, however this is calculated with only a few data points so we do not give any merit to this result. We have also estimated the intersection point between the global power law fit and the data again. We estimate this point as (3,228323), however when using the method of looking where it switches sign one would find 2, 4 or 6 as intersecting points (note that the log plot does not show the point at 0,0).

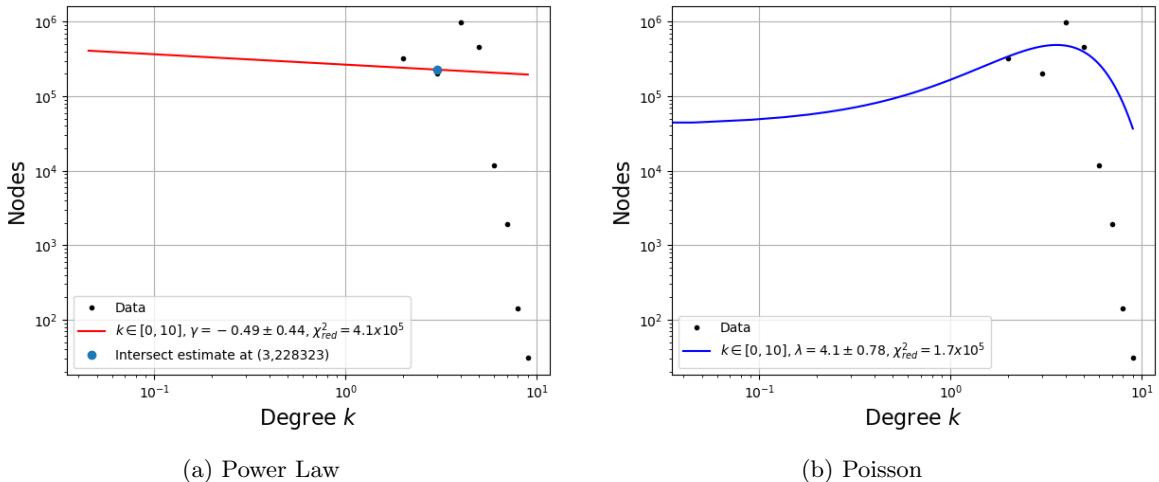


Figure 9: Both the power law as the Poisson distribution fitted to the Road-CA network.

We can thus conclude that both networks do not contain features of a Poisson distribution. This could be related to them not being randomly generated, as completely random models (i.e. the Erdos-Renyi

model) exhibit this distribution. The networks not being random makes sense as astrophysicist's have reasons to collaborate with some people rather than others and roads are made to go somewhere. The power law finds some use in describing the astrophysics network. As we cannot reject the hypothesis that the tail is described by a power law distribution. We think this makes sense due to the high  $k$  region having lots of hubs, that appears because of large collaborations of astrophysicist's.

We have also noticed that fitting empirical data can be quite tricky and one has to make some choices which fraction of the degree distribution one considers. To make this less arbitrary one could for example use the estimators from [9].

## 2.6 Conclusion

In this section we have studied two very different networks. Namely, a scientific astrophysics collaboration network (*astrophysics*) and the road network of road intersections in California (*Road-CA*). First we have investigated the basic statistics of the networks, which are shown in table 1. In the section we also study the adjacency matrix of these models in this section, we see that both networks are described by sparse matrices. We mainly notice that the astrophysics is much more clustered than the Road-CA network, which we briefly discuss. Hereafter we study the degree distributions of the networks. We observe that the astrophysics network has a lot more nodes containing a large number of connections. The Road-CA network quickly peaks at  $k = 4$  and then also decays very fast to zero. We also discuss the friendship paradox, stating that on average your friends have more friends than you have. We show that this feature also appear in our networks. We then discuss why it is not really a paradox, but rather some sort of sampling bias. In the last section we perform a more quantitative study of the form of the degree distributions. Here we find that in general both networks are not well described by either a power law or a Poisson distribution. The exception is the tail of the astrophysics network, which can be described by a power law.

### 3 Random graph: the Watts-Strogatz and the Erdos-Renyi model

Random network, first used by Erdos [3] to give a probabilistic construction of a graph with large girth and large chromatic number, wasn't become objects of interest until later that Erdos and Renyi began a systematic study of random graphs and defined several key measurement on their properties. Among those, the Erdos-Renyi model,  $G(N, p)$ , was the precedent, which is constructed by connecting  $N$  labeled nodes randomly, and each edge is included in the graph with probability  $p$ , independently from every other edge. Equivalently, the probability of generating a graph that has  $N$  nodes and  $M$  edges is  $p^M(1-p)^{\binom{N}{2}-M}$ . Due to its randomness, the adjacency matrix is always formulated in a disorder way, even for small  $p$ , as the Figure 10 shows.

On the base of that, Watts and Strogatz published their 1998 paper “Collective dynamics of ‘small-world’ networks” [4], which had a phenomenal influence on the field of complex systems and was one of the defining studies for the following success of network science to emerge as an interdisciplinary field. It was not only the first of a series of studies trying to explain the small-world effect as based on “six degrees of separation” experiment, it introduced a simple and intuitive network model which had, at its core, the defining properties to obtain a “complex” system. Specifically, a Watts-Strogatz model  $WS(N, 2r, p)$  is constructed first by arranging  $N$  vertices in a circle, where each node is linked to its  $2r$  neighbors,  $r$  clockwise and  $r$  anti-clockwise. Then, with probability  $p$ , it rewires each edge going clockwise randomly to another available node. For small  $p$ , it is clear that most nodes are still connected to its neighbor and therefore, the adjacency matrix  $A$  for a Watts-Strogatz model looks like a strip lying on the diagonal with width  $2r$ , just as in Figure 11. Due to the fact that there are no self-connected nodes under this setting, the element lying exactly on the diagonal will equals to 0, which means that  $A[i, i] = 0$  for  $\forall i \in \{0, \dots, N-1\}$ . Even though the width of the “strip” depends only on  $r$ , unlike Erdos-Renyi model, the rewiring probability  $p$  determines its existence and thus, the level of randomized state; when  $p \rightarrow 1$ , the adjacency matrix of Watts-Strogatz model will also approaching a randomized state, but in a different way comparing to the Erdos-Renyi model.[10] That means there is a discrepancy between Watts-Strogatz model and Erdos-Renyi model even for  $p \rightarrow 1$ , and this difference will be made evident in the next section. As the plot (b) and (c) in Figure 11 indicates, a Watts-Strogatz model contains both a disorder structure, just like Erdos-Renyi model, and a crystal structure, which inherited from nonrandom network.

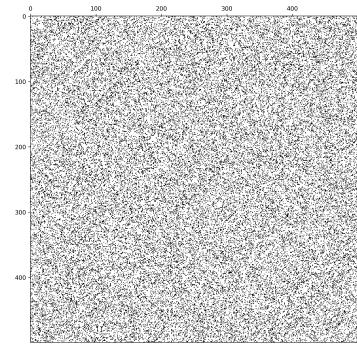


Figure 10: Example of adjacency matrices of the Erdos-Renyi model for  $p = 0.2$ .

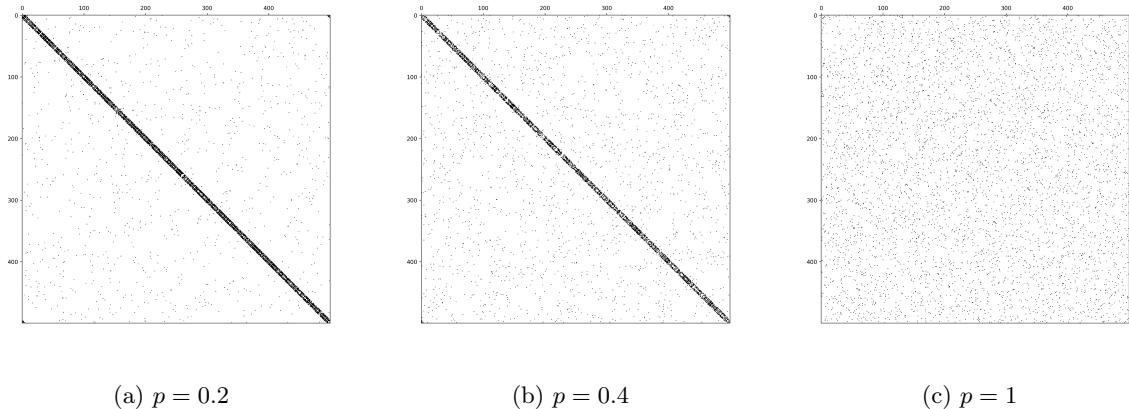


Figure 11: Example of adjacency matrices for the Watts-Strogatz model.

In general terms, the Watts-Strogatz model starts with a regular, locally connected structure, and its

rewiring process introduces long-range edges between the nodes. By varying the probability  $p$ , it interpolating between two well-studied physical systems: a crystal and disorder, namely, capturing the characteristics of both ‘high network clustering’ and ‘short characteristic path length’ without losing the randomized framework. For even small amounts of rewired contacts, the probability that two neighbors are connected (typically large in social networks) hardly changes, while almost immediately, short paths between individual nodes appear, explaining how social networks can be both: highly clustered but with a small amount of necessary steps to reach one node from another.[10]

The Watts-Strogatz model is widely used in the network literature, often to explore the influence of the small-world effect on the outcome of dynamic processes taking their course on the network. One important feature of the model is that the mean degree, i.e. the average number of edges per node, is constant, which is a first-order control parameter for a variety of dynamic systems based on it, e.g. random walks or epidemic spreading[11].

In this section, we will first analyze the two models from a theoretical perspective. We will compute the degree distribution, the clustering coefficient and the mean shortest path distance for the Watts-Strogatz model, elaborating the comparison with more random Erdos-Renyi model. After that, we will provide some simulation results to provide external validity to the theoretical prediction.

### 3.1 Theoretical approach

In the usual notation, we will refer to the Watts-Strogatz model with  $N$  nodes, neighborhood coefficient  $r$  and probability  $p$  as  $WS(N, 2r, p)$  and to the Erdos-Renyi model as either  $G(N, p)$  or  $G(N, M)$  where  $M$  is the total number of edges, related to  $p$  by the obvious formula  $p = M \binom{N}{2}^{-1}$ .

#### The degree distribution of the Watts-Strogatz model (including Bonus part)

By construction, no edges are added or removed from the graph, so the average degree remains the same. Singularly, for each node  $i$ , there are at least  $r$  edges, clockwise in the initial configuration, that can either be left untouched or rewired. In both cases, they keep being edges of the vertex  $i$ . This sets a lower bound for the degree  $k_i$  that we can write as

$$k_i = r + n_i \quad (9)$$

for some  $n_i \geq 0$ . Moreover, we can split  $n_i$  into the values  $A_i + B_i$  where  $A_i$  is the number of anticlockwise untouched edges and  $B_i$  is the number of edges that were rewired to  $i$ . To better understand this notation, consider the simple example in Figure 12. Here, the rewired edges are highlighted in red and the untouched

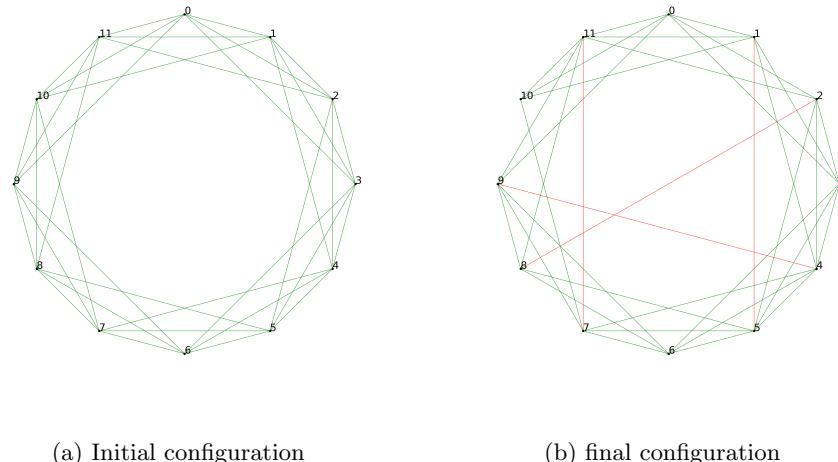


Figure 12: Example of Watts-Strogatz model with  $n = 12, r = 3$  and  $p = 0.1$

ones in green. We see that the perturbed edges are  $(1, 2) \xrightarrow{\text{rewired to}} (1, 5), (7, 8) \rightarrow (7, 11), (8, 10) \rightarrow (8, 2)$

and  $(9, 10) \rightarrow (9, 4)$ . This implies

$$A_2 = A_8 = 2 \quad \& \quad A_{10} = 1 \quad \& \quad B_2 = B_4 = B_5 = B_{11} = 1.$$

For the other values of  $i$ , we have that  $A_i = 3$  and  $B_i = 0$ .

By construction, every vertex  $i$  has  $r$  clockwise edges that can be left untouched with probability  $(1-p)$ . This reflects to the distribution of the variable  $A_i$ . In particular, we see that, for  $0 \leq n \leq r$ ,

$$P(A_i = n) = \binom{r}{n} (1-p)^n p^{r-n}. \quad (10)$$

On the other hand, every vertex  $i$  has  $(rN)$  edges that can be rewired towards  $i$ , each with probability  $p/N$ . As before, we can say that, for  $0 \leq n \leq rN$ ,

$$P(B_i = n) = \binom{rN}{n} \left(1 - \frac{p}{N}\right)^n \left(\frac{p}{N}\right)^{rN-n}.$$

By the Poisson limit Theorem, we get that

$$\lim_{N \rightarrow \infty} \binom{rN}{n} \left(1 - \frac{p}{N}\right)^n \left(\frac{p}{N}\right)^{rN-n} = \frac{(rp)^n}{n!} e^{-rp}.$$

Therefore, for  $N$  sufficiently large, this binomial distribution can be approximated by the Poisson of parameter  $rp$

$$P(B_i = n) \simeq \frac{(rp)^n}{n!} e^{-rp}. \quad (11)$$

All together we can write, for  $m \geq r$ ,

$$\begin{aligned} p_m := P_p(k_i = m) &= \sum_{n=0}^r P_p(r + A_i + B_i = m | A_i = n) P_p(A_i = n) \\ &= \sum_{n=0}^{\min\{m-r, r\}} P_p(r + A_i + B_i = m | A_i = n) P_p(A_i = n) \end{aligned}$$

where we took the  $\min\{m-r, r\}$  as the maximum index because  $n$  cannot be greater than this value. To see this, keep in mind that  $A_i$  is the number of untouched anticlockwise edges of the node  $i$ . Clearly, there is no chance for it to be greater than  $r$ . Moreover, if  $m \leq 2r$ , there is also no chance of having  $A_i > m-r$  and degree equal to  $m$  because it would mean

$$m = r + B_i + A_i > r + B_i + m - r = B_i + m \implies B_i < 0,$$

which is obviously not possible. Then, we can further manipulate the previous equation by using the results in Equation (10) and (11) and obtain

$$\begin{aligned} p_m &= \sum_{n=0}^{\min\{m-r, r\}} P_p(B_i = m - r - n) P_p(A_i = n) \\ &\simeq \sum_{n=0}^{\min\{m-r, r\}} \binom{r}{n} (1-p)^n p^{r-n} \frac{(rp)^{m-r-n}}{(m-r-n)!} e^{-pr} \end{aligned} \quad (12)$$

We will see in the next section how this distribution compares with the simulation results.

If we look at the limit for  $p \rightarrow 1$  we see that

$$\begin{aligned} \lim_{p \rightarrow 1} P_p(k_i = m) &\simeq \lim_{p \rightarrow 1} \sum_{n=0}^{\min\{m-r, r\}} \binom{r}{n} (1-p)^n p^{r-n} \frac{(rp)^{m-r-n}}{(m-r-n)!} e^{-pr} \\ &\simeq \lim_{p \rightarrow 1} \binom{r}{0} p^r \frac{(rp)^{m-r}}{(m-r)!} e^{-pr} = \frac{r^{m-r}}{(m-r)!} e^{-r} \end{aligned}$$

where the terms  $(1-p)^n$  for  $n \neq 0$  tends to 0 faster than any other term for  $p \rightarrow 1$ ; thus the only contribution of the sum to the limit is when  $n = 0$ .

## The Erdos-Renyi model

We want to compare the Watts-Strogatz model  $WS(500,r, 1)$  with the Erdos-Renyi model  $G(500,p)$  that has the same average degree of the previous model. We know that  $N = 500$  is sufficiently large such that the Erdos-Renyi model follows a Poisson distribution of parameter  $\lambda = p(N - 1)$ , which is the expected number of edges per node. Precisely

$$p(m) = \frac{\lambda^m}{m!} e^{-\lambda}$$

We know that the average degree of  $WS(500, r = 5, 1)$  is  $2r$ . Therefore,  $\lambda = 10$  and

$$p = \frac{2r}{N - 1} \sim 0.02.$$

It is clear from Figure 13 that these two distributions are slightly different. The reason for this difference is twofold: firstly, the Watts-Strogatz model has a lower bound for the number of edges connecting a node. Equation (9) showed that minimal degree is precisely  $r$ . This is not the case for the Erdos-Renyi model. In fact, without any initial configuration, some nodes of  $G(N, p)$  may remain, in principle, completely without edges. Secondly, while the total number of edges of the Watts-Strogatz model is fixed, the total amount of edges in the Erdos-Renyi may vary around 2500. This results in an higher probability of having a big number of edges connecting a node. All things considered, Figure 13 shows also that the degree of the Watts-Strogatz model tends to concentrate on the expected value  $2r$ , more than the other model. This is due to the stricter construction of the  $WS$ : the degree of each node is  $2r$  in the initial configuration and it is very likely to stay around that number by the rewiring process. On the contrary, the Erdos-Renyi is completely random and it creates edges from scratch, thus its degree behaves in a more diversified way. In section 3.2, we will see how these two distributions relates to the actual empirical ones obtained from the simulations.

## The clustering coefficient of the Watts-Strogatz model

To get an insight of the Watts-Strogatz model, we would start by looking at the  $WS(N, 2r, 0)$ , the model with probability  $p = 0$  to rewire, and try to prove that the clustering coefficient  $C_{0,i}$  is

$$C_{0,i} = \frac{3r - 3}{4r - 2}.$$

The clustering coefficient is defined as a measure structuring of the edges around vertex  $v_i$  as the formula (3) shows. Specifically, the  $k_i$  represents the degree of  $v_i$ , which equals  $2r$  under the settings. After assuming a zero probability of rewiring, each vertex is connected to  $r$  nodes clockwise and  $r$  nodes anti-clockwise, and every vertex shares the homogeneous property. So for all  $v_i$  where  $i \in \{0, 1, 2, \dots, N - 1\}$ , we can separate its neighbors into two groups, one for the nodes lying in the clockwise direction,  $G_C(i)$ , while the other group,  $G_A(i)$ , includes all the nodes at the anticlockwise direction. Noticed that the Watts-Strogatz model is an undirected graph, we would only focus on the link formed clockwise to eliminate the overlap. Therefore, for all  $v_j \in G_A(i)$ , that is  $j \in \{i - r, \dots, i - 1\}$ , each of these vertices links clockwise to other  $r - 1$  nodes either in  $G_A(i)$  or  $G_C(i)$ . For instance, if we take point 6 in the initial configuration of Figure 12 as  $v_i$  and suppose  $r = 3$ . Then for node 3,4,5, each of them has 2 direct clockwise links. Specifically, for node 3, it links to node 4 and 5; for node 4, it links to node 5 and 7; while for node 5, it links to node 7 and 8. As the total number of  $v_j$  is  $r$ , the total amount of links formed between the neighbors of  $v_i$  (both  $G_A(i)$  and  $G_C(i)$ ) and the  $v_j \in G_A(i)$  is:

$$(r - 1)r. \tag{13}$$

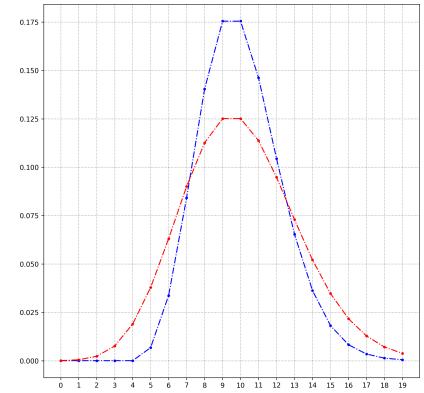


Figure 13: Two Poisson distributions compared:  $G(500,0.02)$  in red and  $WS(500,10,1)$  in blue

Analogously, we can calculate the amount of links between neighbors for  $v_j \in G_C(i)$ . Since we now only concentrate on the clockwise links formed between clockwise neighbors, while the vertex  $v_{i+1}$  has one link beyond the reach of  $v_i$ , that is  $v_{i+r+1}$ , the amount of links the vertex can provide is  $r - 1$ . The amount of those links will decrease as the  $v_j$  moves clockwise towards  $v_{i+r}$ , which provides 0 links between neighbors without overlap. For example, if we again take point 6 in Figure 12 as  $v_i$  and suppose  $r = 3$ , then for node 7, it has 2 direct links to node 8 and 9; for node 8 it has 1 link to node 9; while node 9 have 0 link without overlap. Thus, the sum of links for  $v_j \in G_C(i)$  is:

$$(r - 1) + (r - 2) + \cdots + 0 = \frac{(r - 1)r}{2}.$$

Combining the result in formula (13), we can deduce the  $C_{0,i}$  as follows:

$$C_{0,i} = \frac{(r - 1)r + \frac{(r - 1)r}{2}}{\frac{2r(2r - 1)}{2}} = \frac{3(r - 1)r}{2(2r - 1)r} = \frac{3r - 3}{4r - 2}$$

Also notice that the number of links here can be treated as the amount of triangle founded between two neighbor nodes  $v_j, v_k$  and  $v_i$ , where  $j, k \neq i$ , i.e. the number of triangles with a vertex in  $v_i$ .

On the base of that, we intend to generalize the clustering coefficient of  $WS(N, 2r, p)$  for any other  $p \in [0, 1]$ . To do so, we rely on the geometric description of  $C_{0,i}$  in terms of the number of triangles with a vertex in  $v_i$ . Such a triangle requires a link between two neighbors of  $v_i$ , namely  $v_j$  and  $v_k$  and obviously, the two edges  $(v_i, v_j)$  and  $(v_i, v_k)$ . With  $p > 0$ , the rewiring process may remove any of those links and in turn, may disintegrate the triangle. The probability that a triangle remains untouched is exactly the product of the probability that all three edges exists, which is:

$$P(\text{untouched edge})^3 = (1 - p)^3.$$

However, the probability that a triangle of the initial configuration still exists after the rewiring is not just the probability that it remains untouched: it is possible that, even if one of the edges of the triangle is rewired, for example  $(v_i, v_j)$  is rewired to  $(v_i, v_l)$  for some  $l \neq j, k$ , one of the initial edges of  $v_j$  might be rewired back precisely to  $(v_j, v_i)$ , thus reconstructing the triangle. The probability of this event to happen is very low when  $N$  is large because it is in the scale of  $1/N$ . Therefore, if we wanted to approximate the clustering coefficient for general  $p$ , we would say that it corresponds to the number of unperturbed triangles of the initial configuration, that is

$$C_{p,i} \approx C_{0,i}(1 - p)^3.$$

### Mean shortest path distance for the Watts-Strogatz model

As previously introduced, the *mean shortest path distance*  $\ell$  of a random graph with  $N$  nodes corresponds to the average distance between two nodes. As in equation (2), this is explicitly given by

$$\ell = \binom{N}{2}^{-1} \sum_{i,j} \ell(v_i, v_j)$$

where  $\ell(v_i, v_j)$  is the minimal distance, or lowest number of edges, between the nodes  $v_i, v_j$ . In the case of a Watts-Strogatz model  $WS(N, 2r, p)$ , the lenght of the “geodesic”  $\ell(v_i, v_j)$  is exactly the distance  $d_N^2(i, j)$  defined at the beginning of this section. Therefore,

$$\ell = \frac{2}{N(N - 1)} \sum_{i,j} d_N^2(i, j).$$

In this setting, a *mean-field treatment* of the model gives the following approximation:

$$\ell \approx \frac{N}{r} f(Nrp), \tag{14}$$

where

$$f(x) = \frac{1}{2\sqrt{x^2 + 2x}} \tanh^{-1} \sqrt{\frac{x}{x + 2}}.$$

Furthermore, by the identity  $\tanh^{-1}(x) = \frac{1}{2} \log\left(\frac{1+x}{1-x}\right)$ , we can rewrite the approximation in Equation (14) in the following compact form:

$$\ell \approx \frac{N}{4r} \frac{1}{\sqrt{(Nrp)^2 + 2(Nrp)}} \log \frac{1 + \sqrt{\frac{Nrp}{Nrp+2}}}{1 - \sqrt{\frac{Nrp}{Nrp+2}}}.$$

Assuming  $p \neq 0$ , we can write

$$\frac{Nrp}{2 + Nrp} = \frac{1}{\frac{2+Nrp}{Nrp}} = \frac{1}{1 + \frac{2}{Nrp}}.$$

Moreover, for  $N$  sufficiently large, (e.g.  $N \gg \frac{1}{p}$ ), we can approximate

$$\frac{1}{1 + \frac{2}{Nrp}} \approx 1 - \frac{2}{Nrp} \quad \& \quad \frac{1}{\sqrt{(Nrp)^2 + 2Nrp}} \approx \frac{1}{Nrp}$$

All together, we obtain that

$$\ell \approx \frac{1}{4pr^2} \log \frac{1 + \sqrt{1 - \frac{2}{Nrp}}}{1 - \sqrt{1 - \frac{2}{Nrp}}} \approx \frac{1}{4pr^2} \log \frac{2 + \frac{2}{Nrp}}{\frac{2}{Nrp}} = \frac{1}{4pr^2} \log(Nrp + 1)$$

This shows that, for sufficiently large  $N$  (e.g.  $N \gg \frac{1}{p}$ ), the mean shortest path  $\ell$  can be roughly approximated by  $\log(N)$ .

### 3.2 Simulation studies

In this section, we will present the simulation results for the degree distribution of the Watts-Strogatz model  $WS(500, 10, p)$  where  $p \in \{0.2, 0.4, 1\}$ . Then, we will compare them with the theoretical predictions we made in previous sections. Note that for each  $p$ , we averaged the results from 50 runs to obtain a standard deviation of at most 0.02 at each point.

#### Empirical and theoretical degree distribution of the Watts-Strogatz model

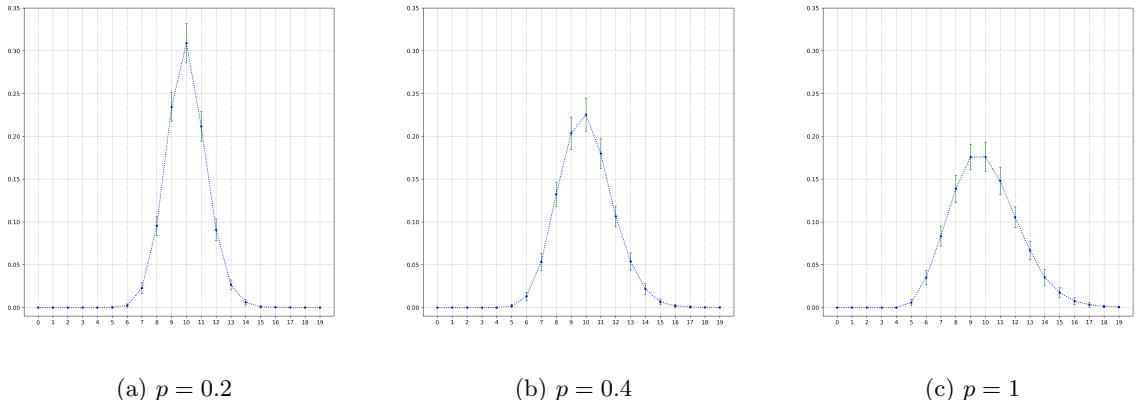


Figure 14: Empirical degree distribution of  $WS(500, 10, p)$ . Average from 50 runs.

The results are shown in Figure 14. Those plots indicates a nearly symmetric distribution with a reducing kurtosis as  $p$  increases. By construction, it was already clear that the minimum number of edges per node is exactly  $r$ . Therefore, we know a priori that  $P(k_i \in \{0, 1, 2, 3, 4\}) = 0$ . This property is satisfied by the empirical results. Moreover, if we plot the theoretical prediction of Equation (12) above the empirical result, just as in Figure 15, it appears that the two results coincide. In fact, the red line, which corresponds to the theoretical prediction, is approximately the same as the blue line, which represents the empirical result. These implies, among many things, that  $N = 500$  is sufficiently large to apply the Poisson limit Theorem and the approximation used in the theoretical section.

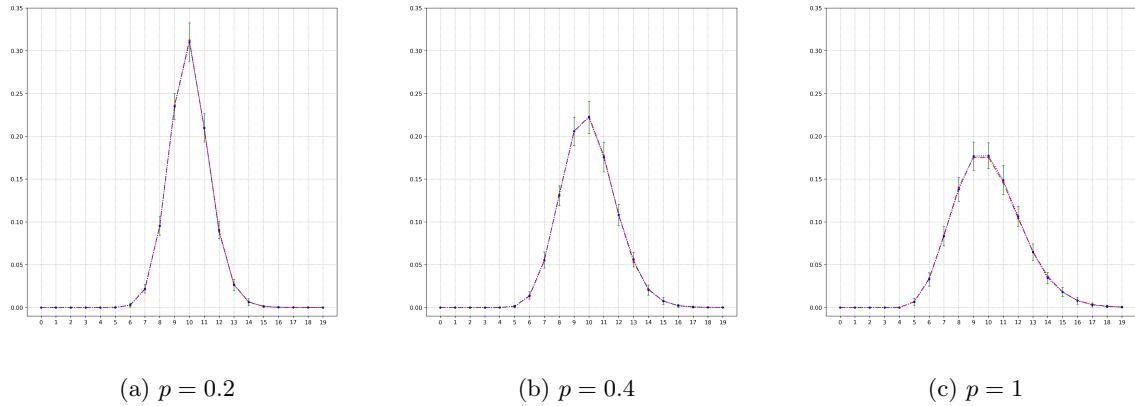


Figure 15: degree distributions compared: in blue the empirical one and in red the one deriving from Equation (12). Average from 50 runs.

### 3.3 Conclusion

In this section, we introduced a couple of example of random graphs: the Erdos-Renyi and the Watts-Strogatz model. The first has no initial configuration and it creates links between nodes from scratch. On the contrary, the second one starts with a locally connected structure and, during a rewiring process, it introduces long-range edges between the nodes. Both of these models have some interesting properties but the more sophisticated construction of the latter makes it funnier to study from a theoretical perspective. Indeed, in the first section we proved that the Watts-Strogatz model has a lower bound for the degree given by the neighborhood coefficient  $r$ . Thanks to this, it became straightforward to compute its distribution, as shown in Equation (12). When compared to the Erdos-Renyi model, whose degree follows a Poisson distribution for a sufficiently large number of nodes, we saw two main differences: no fixed bounds and a lower peak on the expected value as shown in Figure 13.

In the same section, we estimated the clustering coefficient of the Watts-Strogatz model and we proved it to be highly related to the number of triangles of the initial configuration. Then, we proved that the mean shortest path of  $WS(N, 2r, p)$  is of scale  $\log(N)$ , meaning that the model has *small world properties*. Finally, we compared the theoretical results with a simulation of  $WS(500, 10, p)$  for  $p = 0.2, 0.4, 1$  and we saw a clear accordance between them.

## References

- [1] Paul Erdős and Alfréd Rényi. On the evolution of random graphs. In *The structure and dynamics of networks*, pages 38–82. Princeton University Press, 2011.
  - [2] Stanley Milgram. The small world problem. *Psychology today*, 2(1):60–67, 1967.
  - [3] Erdős number and small world phenomenon : Networks course blog for info 2040/cs 2850/econ 2040/soc 2090, <https://blogs.cornell.edu/info2040/2014/09/14/erdos-number-and-small-world-phenomenon/>.
  - [4] Duncan J Watts and Steven H Strogatz. Collective dynamics of ‘small-world’networks. *nature*, 393(6684):440–442, 1998.
  - [5] Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1):47, 2002.
  - [6] Jure Leskovec and Andrej Krevl. SNAP Datasets: Stanford large network dataset collection. <http://snap.stanford.edu/data>, June 2014.
  - [7] Jure Leskovec, Jon Kleinberg, and Christos Faloutsos. Graph evolution: Densification and shrinking diameters. *ACM transactions on Knowledge Discovery from Data (TKDD)*, 1(1):2–es, 2007.

- [8] Jure Leskovec, Kevin J. Lang, Anirban Dasgupta, and Michael W. Mahoney. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters, 2008.
- [9] Ivan Voitalov, Pim van der Hoorn, Remco van der Hofstad, and Dmitri Krioukov. Scale-free networks well done. *Physical Review Research*, 1(3), Oct 2019.
- [10] Benjamin F Maier. Generalization of the small-world effect on a model approaching the erdős–rényi random graph. *Scientific reports*, 9(1):1–9, 2019.
- [11] Albert-László Barabási, Réka Albert, and Hawoong Jeong. Mean-field theory for scale-free random networks. *Physica A: Statistical Mechanics and its Applications*, 272(1-2):173–187, 1999.