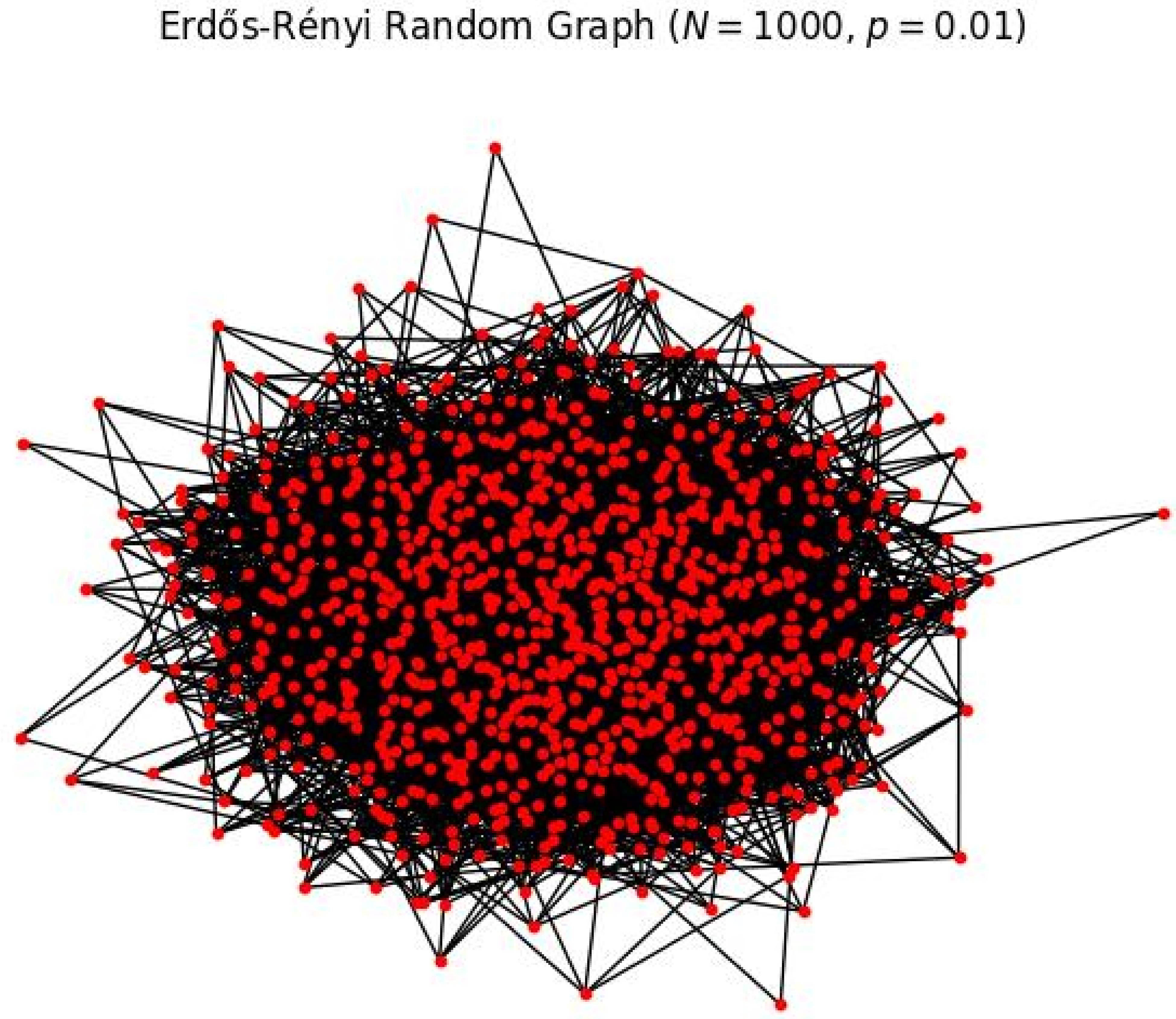
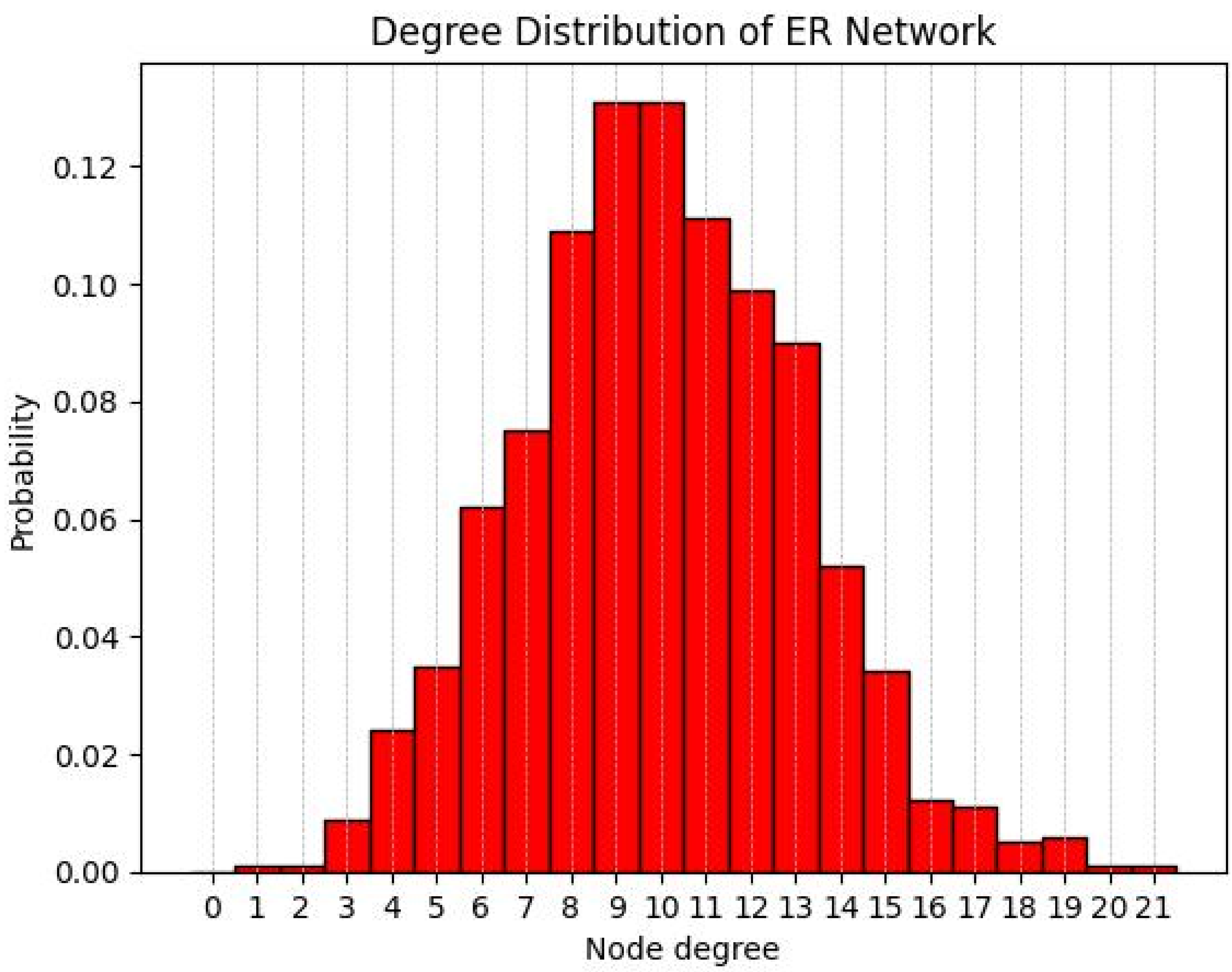
1. Degree distribution of Erdős-Rényi networks (5 pts)

The Erdős-Rényi (ER) model is a model for generating random networks where N nodes are randomly connected such that the probability that a pair of nodes is linked is p. In this exercise, we are going to study the degree distribution of an ER network.

a) (1 pt) Create an instance of the ER network with $N = 1000$ nodes and $p = 0.01$ and visualize it.

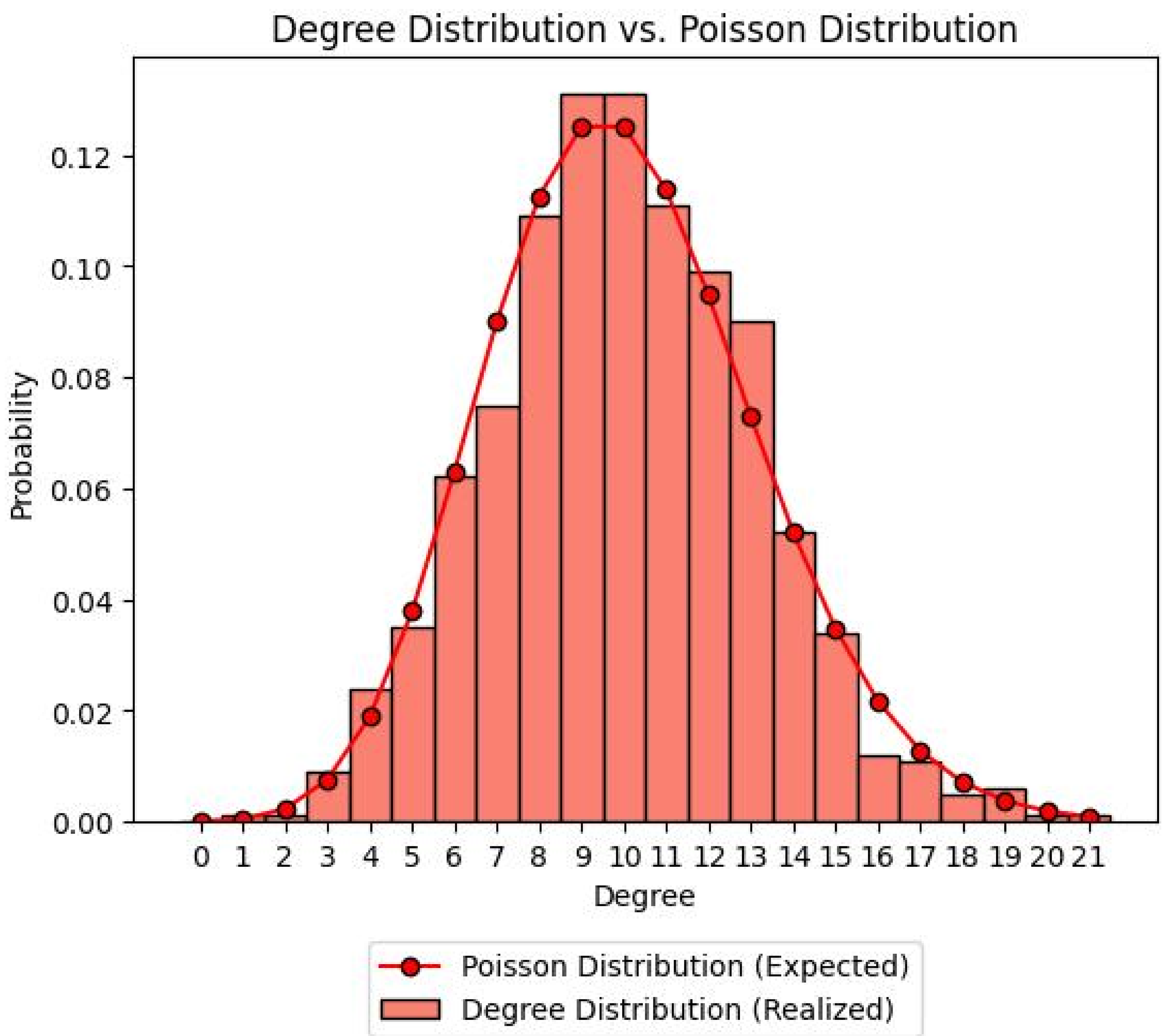


b) (2 pt) Plot the degree distribution of the generated ER network. First, get the histogram of the degree and normalize it to transform it into a probability mass function.



The Erdös- Renyi random graph $G(n,p)$ defines a family of graphs, each of which starts with n isolated nodes, and we place an edge between each distinct node with probability p. In $G(n,p)$ model, the probability of obtaining any one particular random graph with m edges is $p^m(1-p)^{N-m}$, where $N=\binom{n}{2}$.
Erdös- Renyi model has a binomial distribution over degree k. $P(k)$ denotes the fraction of nodes with degree k, and we have $P(k)=\binom{n-1}{k}p^k(1-p)^{n-1-k}$, which is a binomial distribution which has a peak at $\bar{k}=p(n-1)$, and variance of $\sigma^2=p(1-p)(n-1)$

c) (2 pt) Plot the Poisson distribution expected for an ER network and the degree distribution of the realized ER network in one plot.



It is possible to notice that the Poisson distribution with $\lambda = n \cdot p$ is a great approximation for the Erdös-Renyi network degree distribution, and the answer is related to the next question.

Challenge exercise (3 pts) (pen and paper)

Strictly speaking, the degree of an ER network is binomially distributed. Show that the Poisson distribution is the limiting case of the binomial distribution for large N.

As we know, the binomial distribution tends toward the Poisson distribution as $N \to \infty$, $p \to 0$, $\lambda = Np$ stays constant

We need to show that $\binom{N}{x} p^x (1-p)^{N-x} \to \frac{\lambda^x e^{-\lambda}}{x!}$ as $N \to \infty$, $\lambda = Np$ constant

and we have to remember that $e^x = \lim_{N\to\infty} \left(1+\frac{x}{N}\right)^N$

Starting from the binomial formula, we can use $\lambda = Np$ and we obtain

$$\binom{N}{x} p^x (1-p)^{N-x} = \binom{N}{x}\left(\frac{\lambda}{N}\right)^x \left(1-\frac{\lambda}{N}\right)^{N-x} = \frac{N!}{x!\,(N-x)!}\left(\frac{\lambda}{N}\right)^x\left(1-\frac{\lambda}{N}\right)^{N-x} = \frac{\lambda^x}{x!}\cdot\frac{N!}{(N-x)!}\cdot\frac{1}{N^x}\left(1-\frac{\lambda}{N}\right)^{N-x}$$

So we have obtained the factor $\frac{\lambda^x}{x!}$ and we only have to prove that

$$\lim_{N\to\infty} \frac{N!}{(N-x)!}\cdot\frac{1}{N^x}\left(1-\frac{\lambda}{N}\right)^{N-x} = e^{-\lambda}$$

First we look at the term $\frac{N!}{(N-x)!}\cdot\frac{1}{N^x} = \frac{N(N-1)(N-2)\dots(N-x+1)\cancel{(N-x)!}}{\cancel{(N-x)!}\;N^x}$

Considering that in the numerator we have x terms, we can rewrite

$$= \frac{N}{N}\cdot\frac{N-1}{N}\cdot\frac{N-2}{N}\cdots\frac{N-x+1}{N} = 1\left(1-\frac{1}{N}\right)\left(1-\frac{2}{N}\right)\cdots\left(1-\frac{x+1}{N}\right)$$

The binomial formula is now $\frac{\lambda^x}{x!}\left(1-\frac{1}{N}\right)\left(1-\frac{2}{N}\right)\cdots\left(1-\frac{x+1}{N}\right)\left(1-\frac{\lambda}{N}\right)^N\left(1-\frac{\lambda}{N}\right)^{-x}$

where we have decomposed

$$\left(1-\frac{\lambda}{N}\right)^{N-x} = \left(1-\frac{\lambda}{N}\right)^N\left(1-\frac{\lambda}{N}\right)^{-x}$$

If we now take the limit

$$\lim_{N\to\infty}\binom{N}{x}p^x(1-p)^{N-x} = \lim_{N\to\infty}\frac{\lambda^x}{x!}\left(1-\frac{1}{N}\right)\left(1-\frac{2}{N}\right)\cdots\left(1-\frac{x+1}{N}\right)\left(1-\frac{\lambda}{N}\right)^N\left(1-\frac{\lambda}{N}\right)^{-x}$$

we can notice that

- $\lim_{N\to\infty}\left(1-\frac{1}{N}\right) = \lim_{N\to\infty}\left(1-\frac{2}{N}\right) = \dots = \lim_{N\to\infty}\left(1-\frac{x+1}{N}\right) = 1$
- $\lim_{N\to\infty}\left(1-\frac{\lambda}{N}\right)^N = e^{-\lambda}$ (for the property $e^x = \lim_{N\to\infty}\left(1+\frac{x}{N}\right)^N$)
- $\lim_{N\to\infty}\left(1-\frac{\lambda}{N}\right)^{-x} = 1$

So the final result is $\lim_{N\to\infty}\binom{N}{x}p^x(1-p)^{N-x} = \frac{\lambda^x}{x!}e^{-\lambda}$, Poisson formula

proving that the Poisson distribution is the limiting form of the Binomial distribution ∎

2. Erdős-Rényi graph ensemble (5 pts)

The Erdős-Rényi (ER) model is a random graph model, meaning that it generates a network from a set of networks, rather than a single deterministic network. The set of all graphs a model may generate is called a *graph ensemble*. More precisely, a graph ensemble is a distribution of graphs where each graph G_i has a certain probability π_i. We let $G(N, p)$ denote the ensemble of graphs generated by the Erdős-Rényi model with N nodes and edge probability p.

Let's say we have a variable X whose value is a function of network G. We can compute X for each network, but we may be more interested in the expected value of X across the ensemble generated by a model. The expected value, or *ensemble average*, of quantity X is defined as $\langle X \rangle = \sum_i \pi_i X(G_i)$.

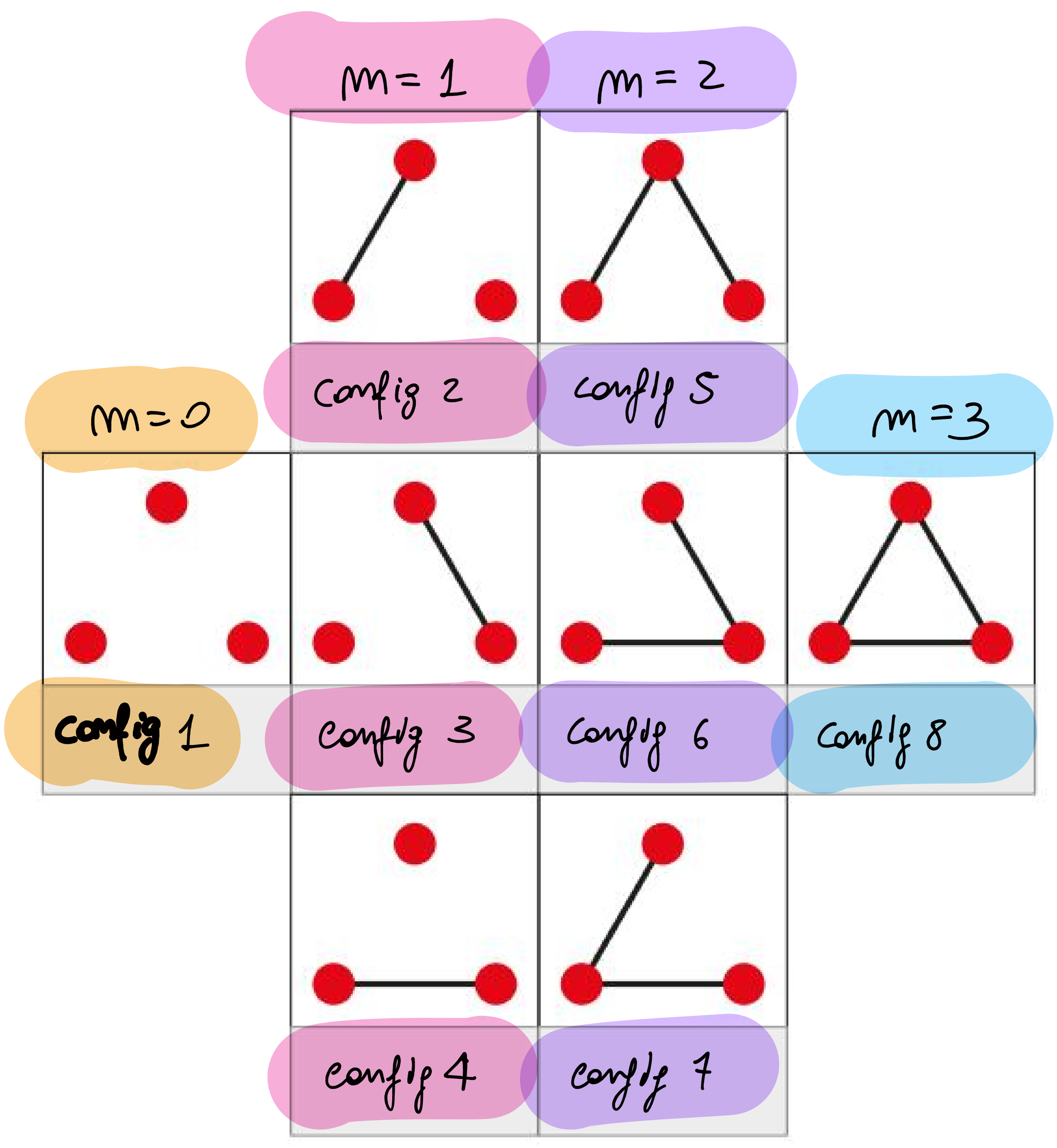
Let us define the following quantities for graph G: the average degree $k(G)$, the diameter of the largest connected component $d^*(G)$, and the average clustering coefficient $c(G)$ (assuming that the clustering coefficient equals to 0 for nodes of degree 0 and 1).

a) (2 pt) Calculate, using pen and paper, the formulas for expected average degree $\langle k \rangle$ and expected diameter $\langle d^* \rangle$ for the ER model $G(N = 3, p)$. Remember to simplify the formulas you get as results.

- EXPECTED AVERAGE DEGREE $\langle k \rangle$

The general formula can be written as $\langle k \rangle = \sum_{i=1}^{2^{\frac{N(N-1)}{2}}} \pi_i \cdot \langle k_i \rangle$, where the probability of each graph is
$\pi_i = p^m (1-p)^{3-m}$, where m is the number of edges.

In our case, $N=3$, so we have $2^3 = 8$ possible configurations, defined in the image below



So, we have that $\langle k \rangle = \sum_{i=1}^{8} \pi_i \cdot \langle k_i \rangle = (1-p)^3 \cdot 0 + 3 \cdot \frac{2}{3} \cdot p(1-p)^2 + 3 \cdot \frac{4}{3} \cdot p^2(1-p) + 2p^3 =$

$= 2p(1-p)^2 + 4p^2(1-p) + 2p^3 =$

$= 2p(1-p)\left[(1-p) + 2p\right] + 2p^3 =$

$= 2p\left[(1-p)\left[(1-p)+2p\right] + p^2\right] =$

$= 2p\left[1 + \cancel{p^2} - \cancel{2p} + \cancel{2p} - \cancel{2p^2} + \cancel{p^2}\right] = 2p$

So the final result is

$\langle k \rangle = 2p$

- **EXPECTED DIAMETER $\langle d^* \rangle$**

Following the same procedure, we have that

$$\langle d^* \rangle = \sum_{i=1}^{8} \pi_i \langle d^*_i \rangle = (1-p)^3 \cdot 0 + 3(1-p)^2 p + 6(1-p)p^2 + 1p^3 =$$

$$= 3p(1-p)^2 + 6p^2(1-p) + p^3 = 3p(1-p)[(1-p)+2p] + p^3 =$$

$$= 3p(1-p)(1+p) + p^3 = 3p(1-p^2) + p^3 = 3p - 3p^3 + p^3 =$$

$$= 3p - 2p^3 = p(3-2p^2)$$

So the final result is $\langle d^* \rangle = p(3-2p^2)$

b) (1 pt) For ER models with sufficiently large p, the expected value of the average clustering coefficient $\langle c \rangle$ equals p. Explain why this is the case.

As we know, the clustering coefficient of a node measures the local density of edges in the neighborhood of that node.
From a mathematical point of view, it is the ratio of the actual number of links in the neighborhood of a node to the maximum number links in its neighborhood.
If we consider a node i, of degree k_i, in one of the $2^{\binom{N}{2}}$ graphs of the ensemble defined by the Erdös-Renyi graph $G(N,p)$, we have that the cluster coefficient of this node is:

$$c_i = \frac{p \cdot \frac{k_i(k_i-1)}{2}}{\frac{k_i(k_i-1)}{2}} = p.$$

Notice that at the denominator we have the maximum number of links in the neighborhood of a node with degree k_i, while at the numerator there is the possible number of the links, due to the "probabilistic" view of the ER graph.

This ratio only depends from p,
so we can say that the average cluster coefficient of a graph j from the ensemble defined by $G(N,p)$ is $\langle c \rangle_j = \frac{\sum_{i=1}^{N} c_i}{N} = \frac{p \cdot N}{N} = p$.

If we average this result over the $2^{\binom{N}{2}}$ possible configurations, we can say that $\langle c \rangle_{G(N,p)} = \frac{\sum_{j=1}^{2^{\binom{N}{2}}} \langle c \rangle_j}{2^{\binom{N}{2}}} = \frac{p \cdot 2^{\binom{N}{2}}}{2^{\binom{N}{2}}} = p.$

So we can confirm that the average clustering coefficient of a ER graph is equal to p.

c) (2 pt) Explain what happens to $\langle c \rangle$, if $N \to \infty$ with $\langle k \rangle$ bounded.

From a logical point of view, if we consider a random graph with N nodes, where N is going to infinity, and the average degree $\langle k \rangle$ is bounded, this means that there is some constant M such that $k < M \ \forall N$. As N grows, the number of nodes increases rapidly, but the average number of connections per node ($\langle k \rangle$) stays below the fixed bound M.
Remembering that the clustering coefficient c measures the connectedness of a node's neighbors, as N increases with k bounded, each node will have a small, fixed number of connections on average.
With a sparse, limited number of connections, it becomes increasingly unlikely that two neighbors of a node will also be connected to each other.
Therefore, as $N \to \infty$, with $\langle k \rangle$ bounded, the average clustering coefficient $\langle c \rangle \to 0$.

From a mathematical point of view, in an ER random graph with N nodes, there are $\binom{N}{2}$ possible edges (pairs of nodes), and considering that $\langle k \rangle$ is fixed, in total there should be $N \cdot \langle k \rangle$ edges in the graph.
If we now consider the probability p of any particular edge existing in the graph, given that there are $N \cdot \langle k \rangle$ edges in total and $\binom{N}{2}$ possible edges, p can be expressed as:

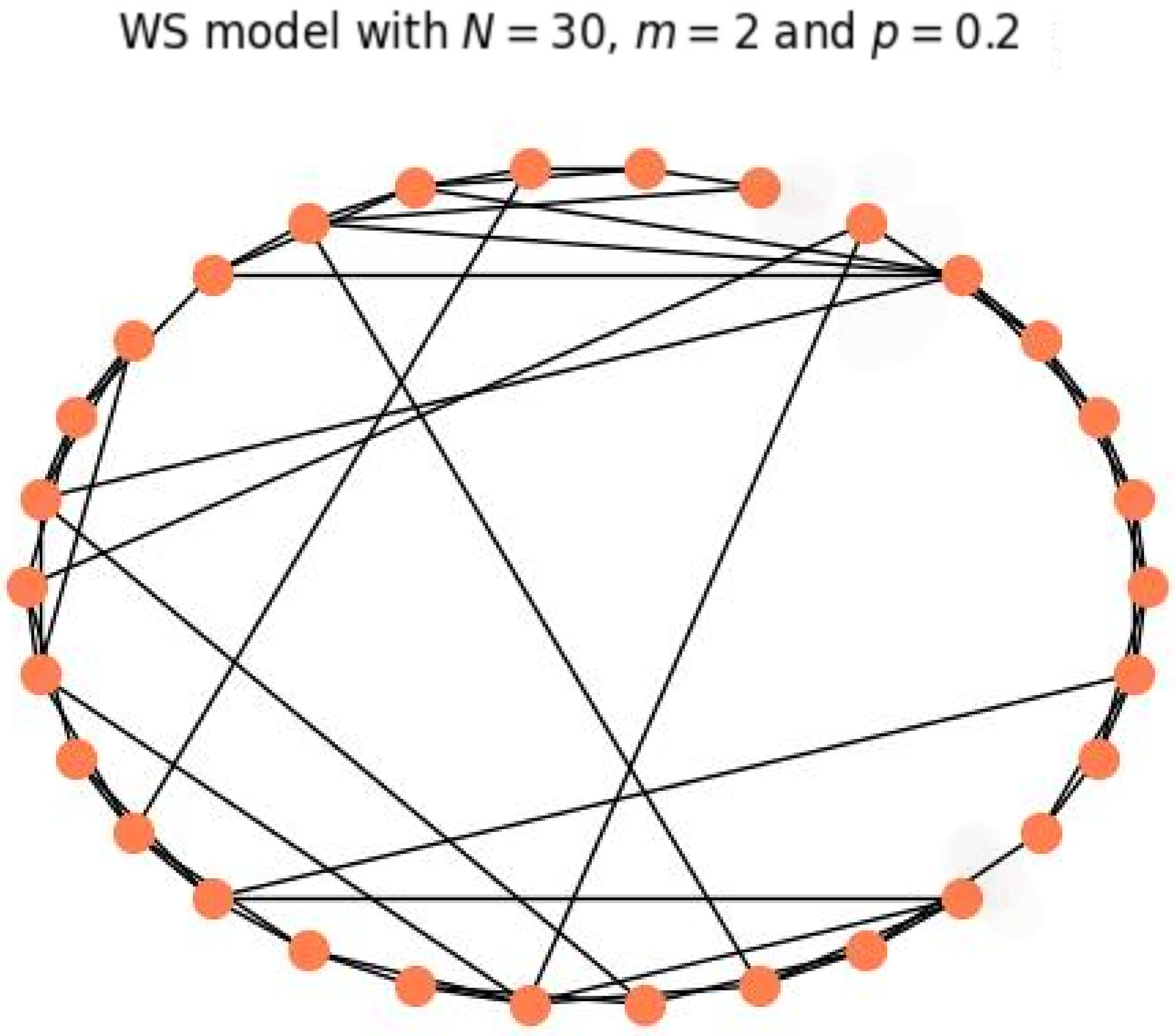
$$p = \frac{N \cdot \langle k \rangle}{\binom{N}{2}} = \frac{2\cancel{N}\langle k \rangle}{\cancel{N}(N-1)} \cong \frac{2\langle k \rangle}{N}.$$ Finally, $$\lim_{N \to \infty} p = \lim_{N \to \infty} \frac{2\langle k \rangle}{N} = 0.$$

This demonstrates that if $N \to \infty$ with $\langle k \rangle$ bounded, the probability p of any particular edge tends to zero, and of course also $\langle c \rangle \to 0$.

3. Implementing the Watts-Strogatz model (6 pts)

In this exercise, you will implement the Watts-Strogatz (WS) small-world model, which is a simple network model that yields a small diameter as well as a high level of clustering. In practice, the WS model is a one-dimensional ring lattice where some of the links have been randomly rewired. The model has three parameters: network size N, m (each node on the ring connects to m nearest neighbors to the left and m nearest neighbors to the right), and p, the probability of rewiring one end of each link to a random endpoint node.

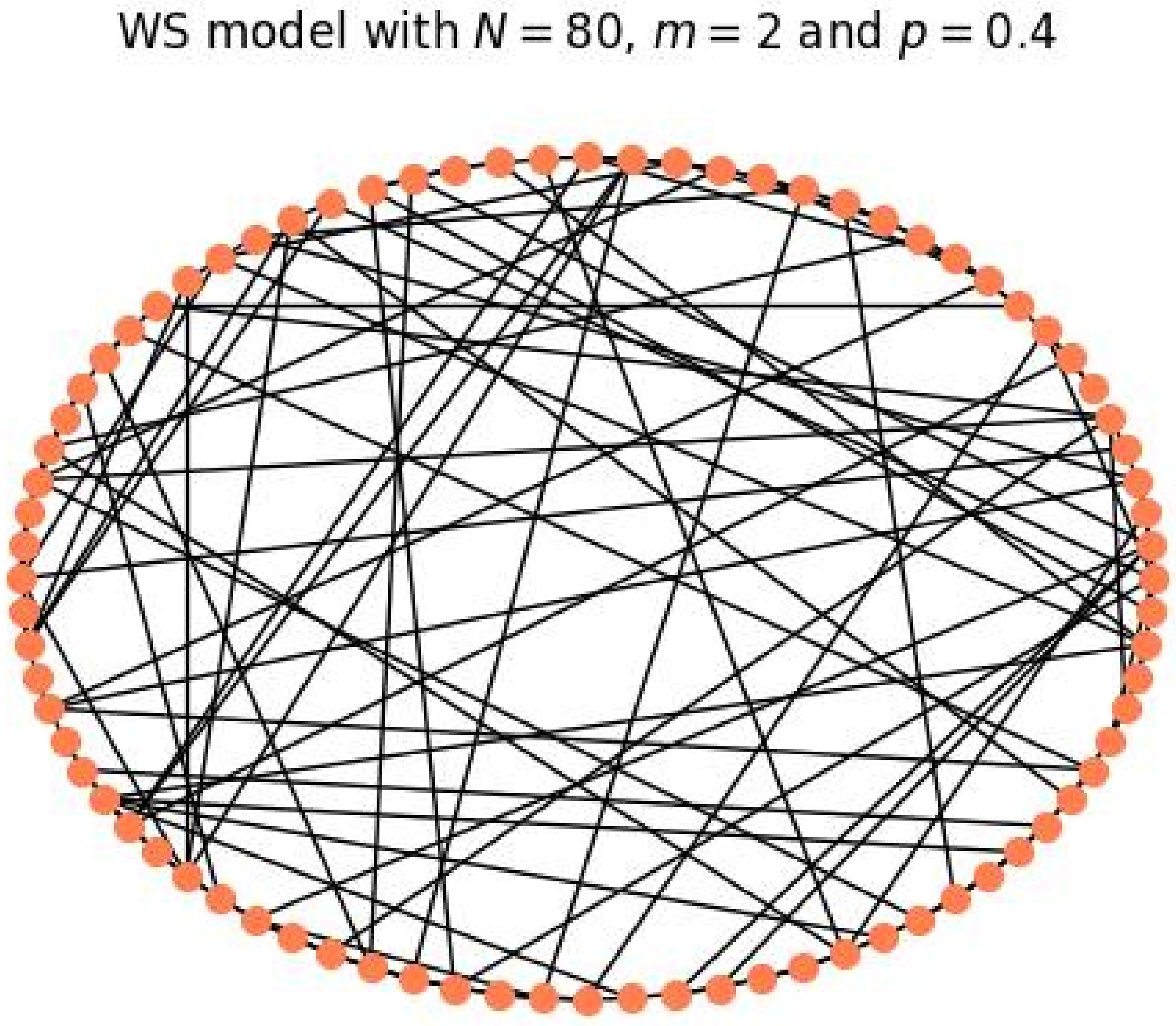
a) (3 pt) Implement the WS model and generate a network using two different sets of parameters: $N = 30$, $m = 2$, $p = 0.2$, and $N = 80$, $m = 2$, $p = 0.4$. Visualize the networks using a circular layout algorithm (`nx.draw_circular(G)`), and check that the networks look right. For each network, report the total number of links and the number of rewired links. Note that NetworkX has a ready-made function for the WS model. However, the task is to program your own function, so do not use it (except for checking results, if in doubt).



TOTAL NUMBER OF LINKS : 60

TOTAL NUMBER OF REWIRED LINKS: 15

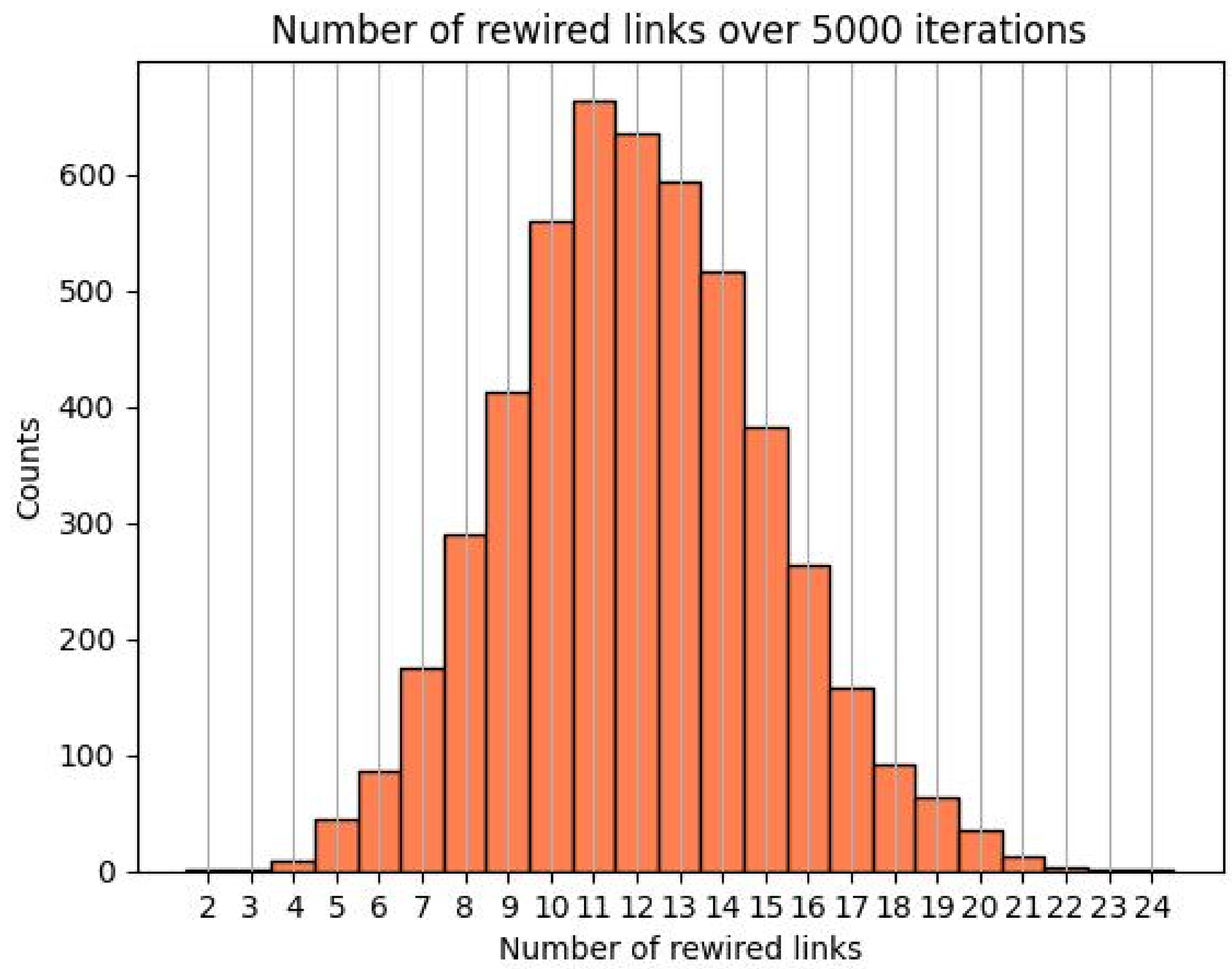
EXPECTED TOTAL NUMBER OF REWIRED LINKS: 12

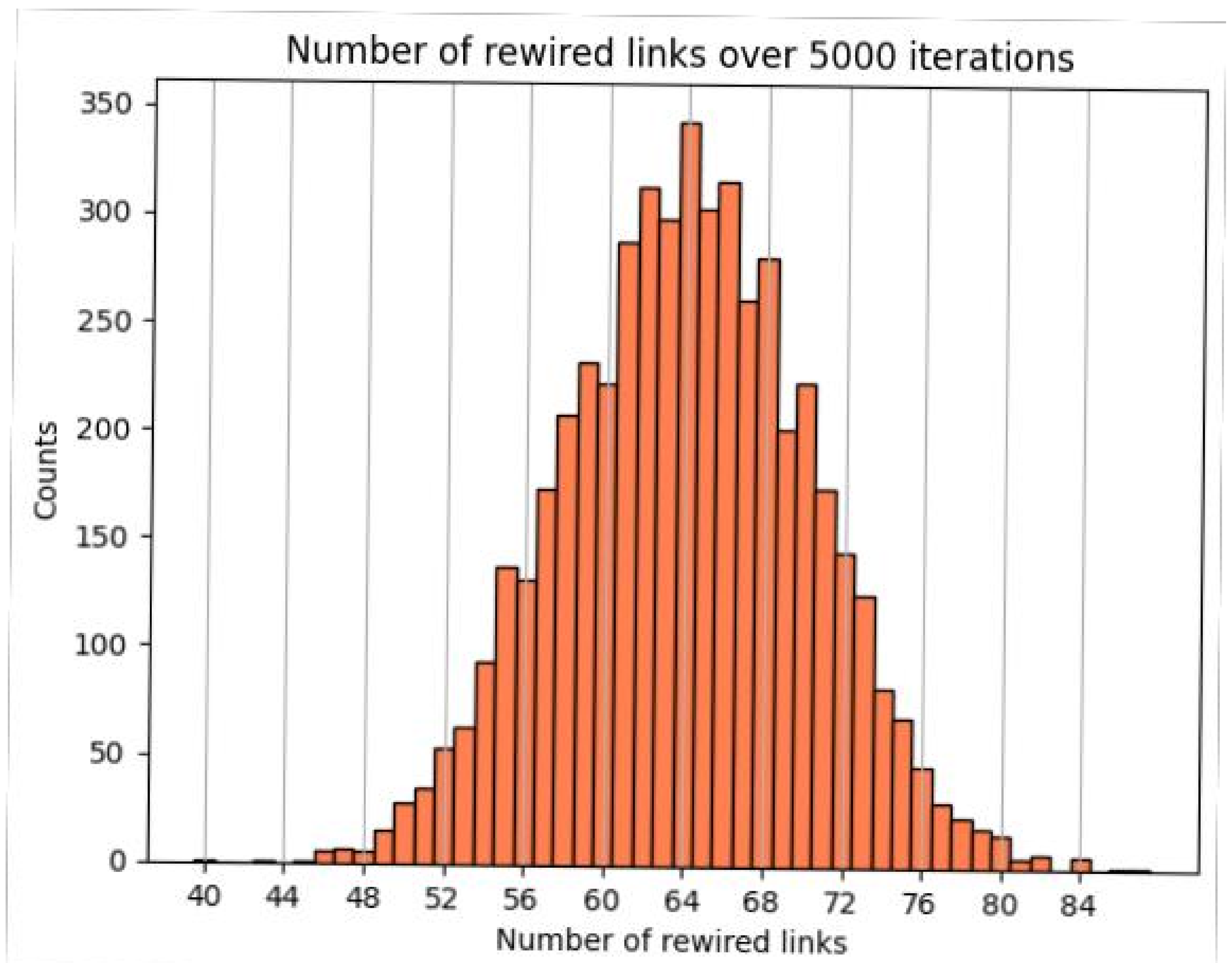


TOTAL NUMBER OF LINKS: 160

TOTAL NUMBER OF REWIRED LINKS: 58

EXPECTED TOTAL NUMBER OF REWIRED LINKS: 64



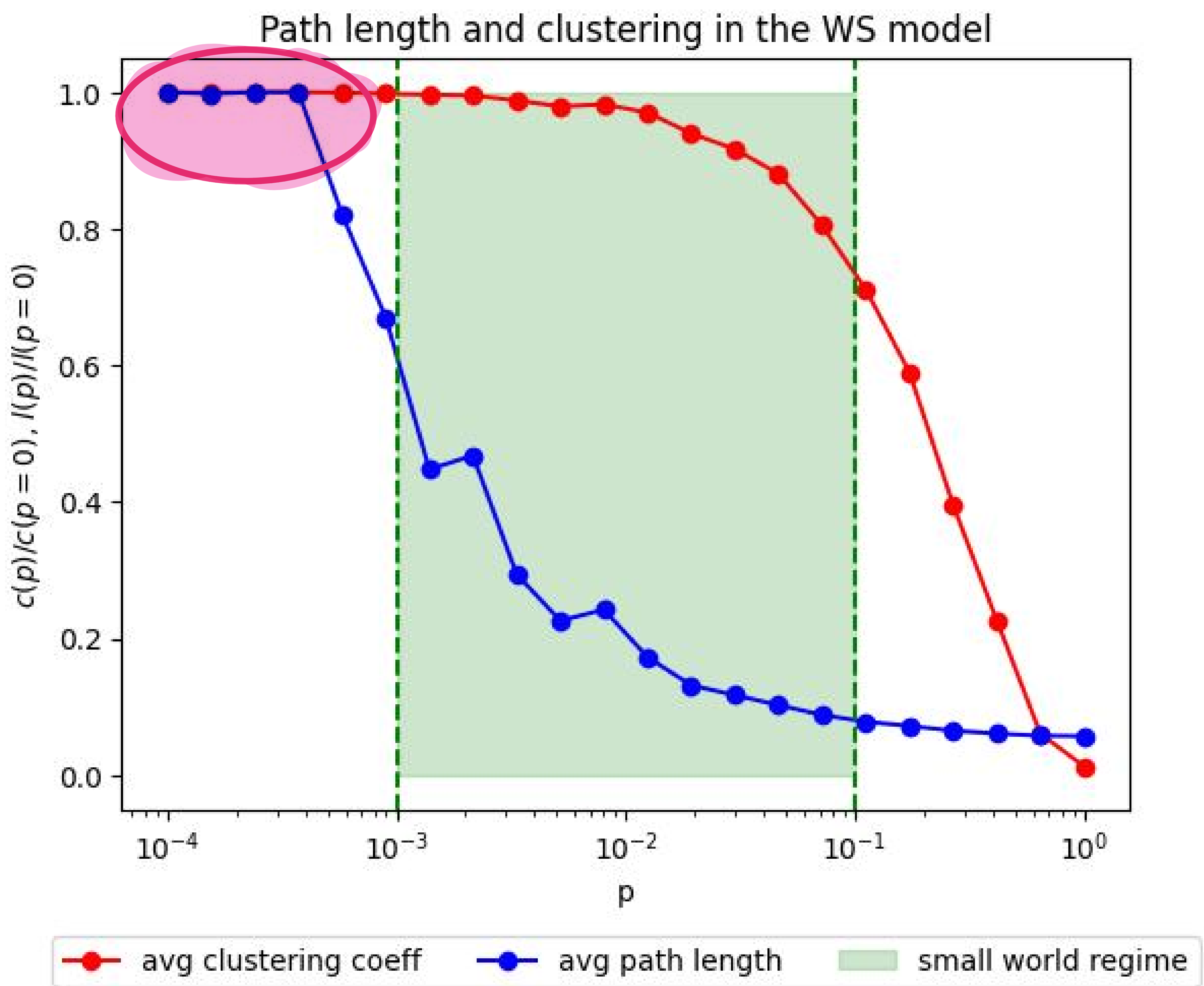


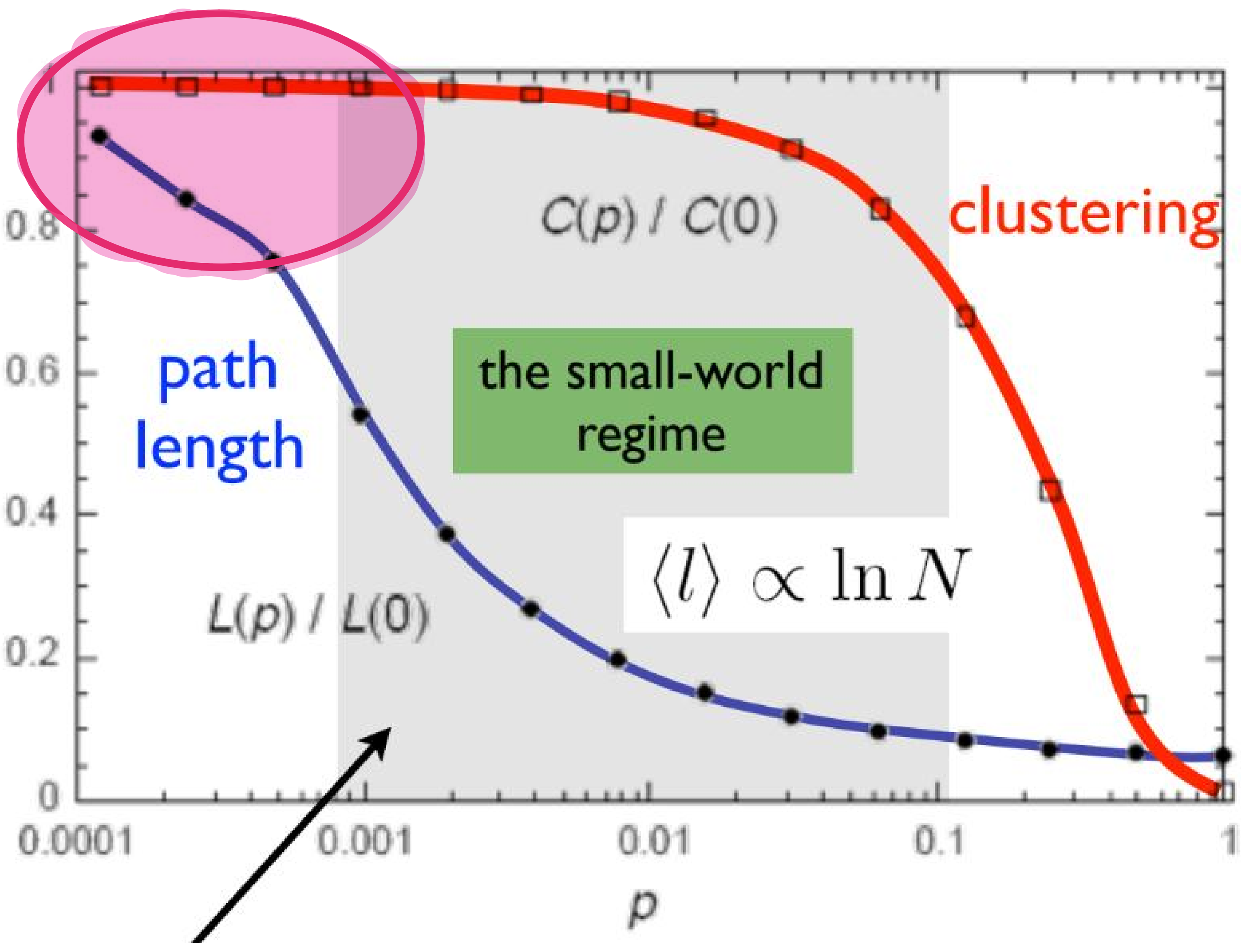
From the graph above, it is possible to notice that the predicted number of links is indeed correct, while in a single run, the number of rewired links is not always the expected one. For this reason, I performed 5000 runs for each of the two models, it is possible to see that the distributions of rewired links are effectively centered around the values we expected.

b) (3 pt) Compute, for WS model networks with $N = 1000$ and $m = 4$, *relative* clustering coefficient $c(p)/c(p = 0)$ and *relative* average path length $l(p)/l(p = 0)$. Plot them as a function of p for $p = 0.001, \ldots, 1$ in one figure. Use a logarithmic x-axis in your plot.

Then answer the following questions:

- Are your results in line with the plots in the lecture slides?
- Why does the clustering coefficient decrease as the probability increases?
- What happens to the average path length? Why?





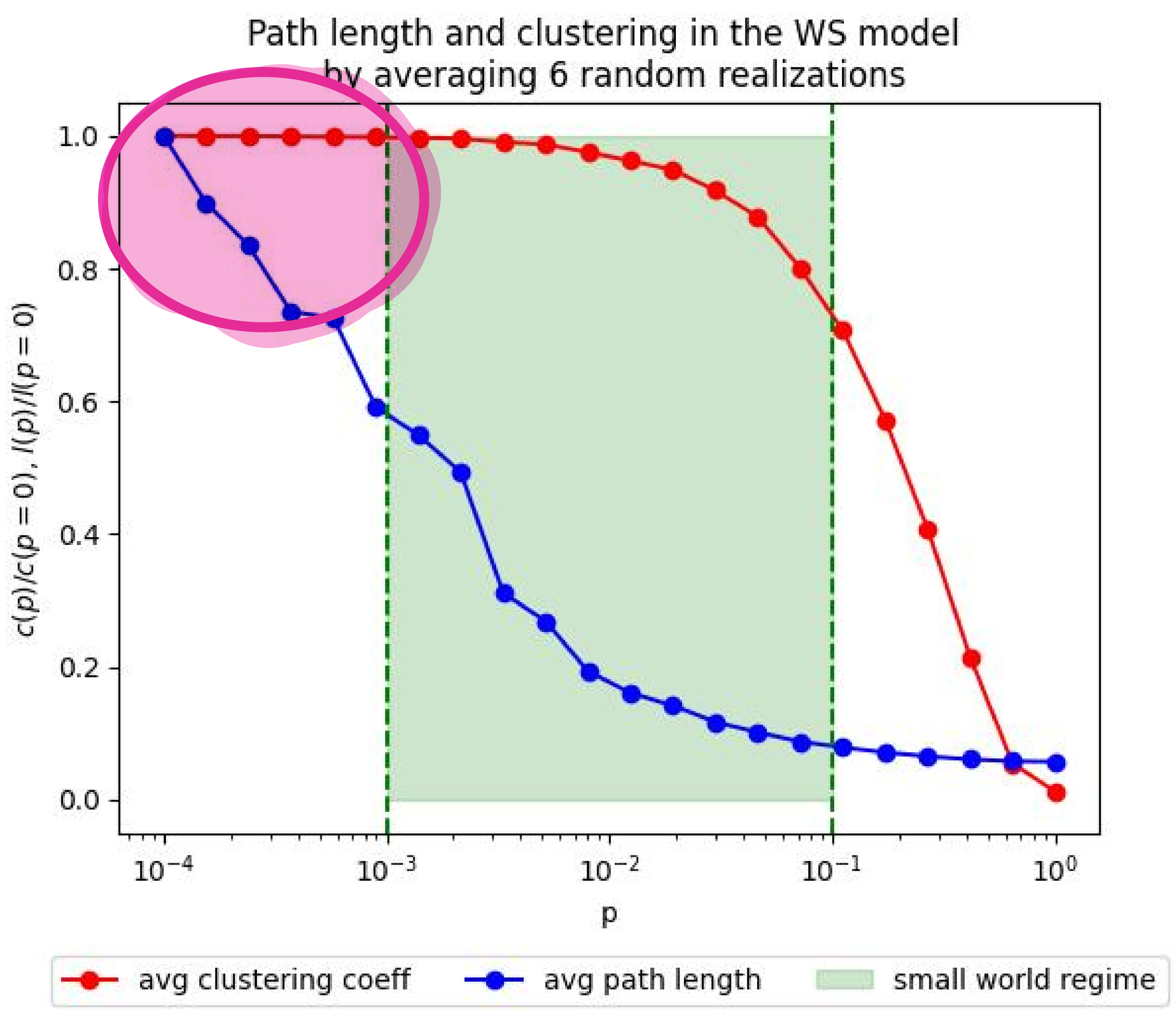
- On the left you can observe the plot obtained experimentally, while on the right there is the plot from the lecture slides. It is possible to observe how in the small-world regime, the trends are extremely similar. What differs between the two graphs is the trend for values of p smaller than those in which the small-world regime occurs (pink highlighted zone). In fact, in the obtained results, there is an overlap between the trend of path length and that of clustering. The explanation for this is provided in the answer to the next question.

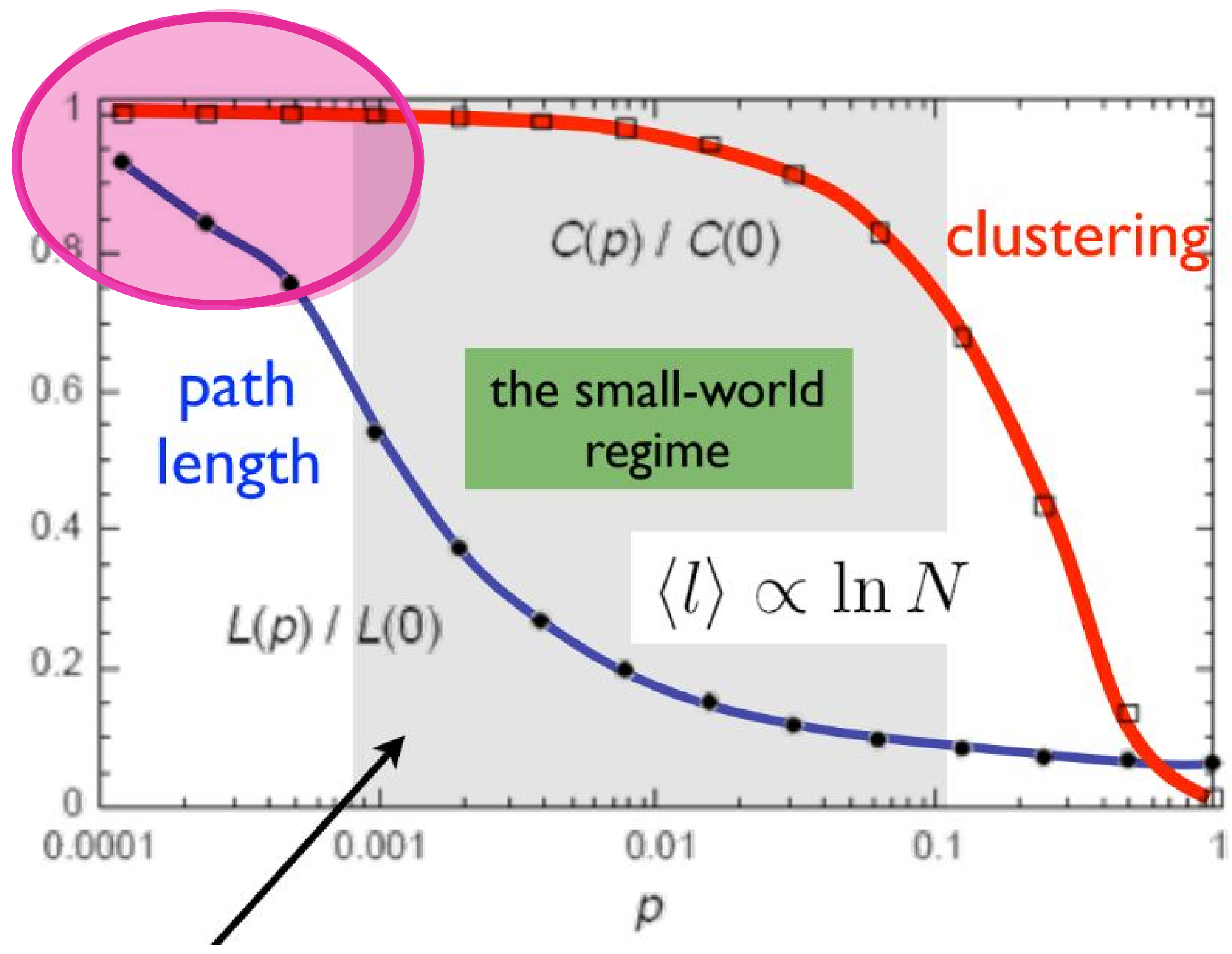
- When we rewire edges in a small world network, we are essentially destroying triangles. This is because the rewiring process is random, and it is unlikely that two nodes that are connected to the same node will also be connected to each other. As the rewiring probability increases, more and more triangles are destroyed, and this is why cluster coefficient of a small-world network decreases as the rewiring probability increases.

- The average path lenght in a WS model decreases as the rewiring probability p increases. This is because the rewiring process creates shortcuts in the network, which reduces the average distance between nodes.

Challenge exercise (2 pts)

You probably observe that your plots in b) are not exactly the same as the one in the lecture slides. Explain why this is the case. Based on your argument, improve your code and plot again $c(p)/c(p=0)$ and $l(p)/l(p=0)$ as a function of p.





On the left it is possible to observe the improved experimental result, while on the right the results presented in class.
As mentioned before, the main difference concerned the overlap between the average clustering coefficient and the average path length, for small values of p. The main reason is due to the fact that we were previously observing a single realization of the WS model. In this case instead, 6 runs were performed, on which the mean of the values obtained was calculated.
Even in the original paper, the data have been averaged over 20 random realizations of the rewiring process, suggesting that, as the number of iterations increases, the two graphs should become increasingly similar. (In this case, 6 were made due to the time required for each iteration to complete.)