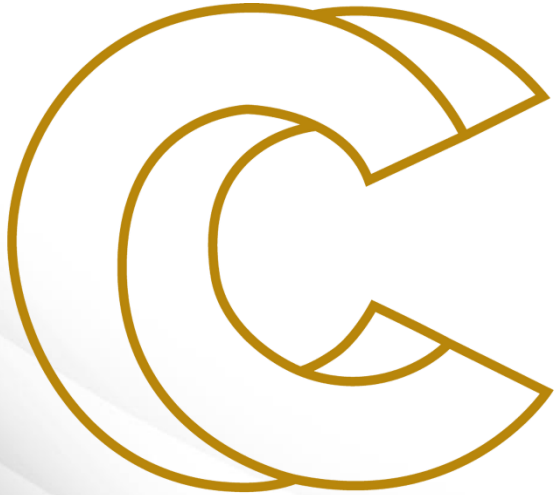
Sabancı Üniversitesi

TÜBİTAK



EURO 4SEE

Residue Networks: Understanding the Time Evolution and Mutations of Proteins from a Graph Theoretical Perspective

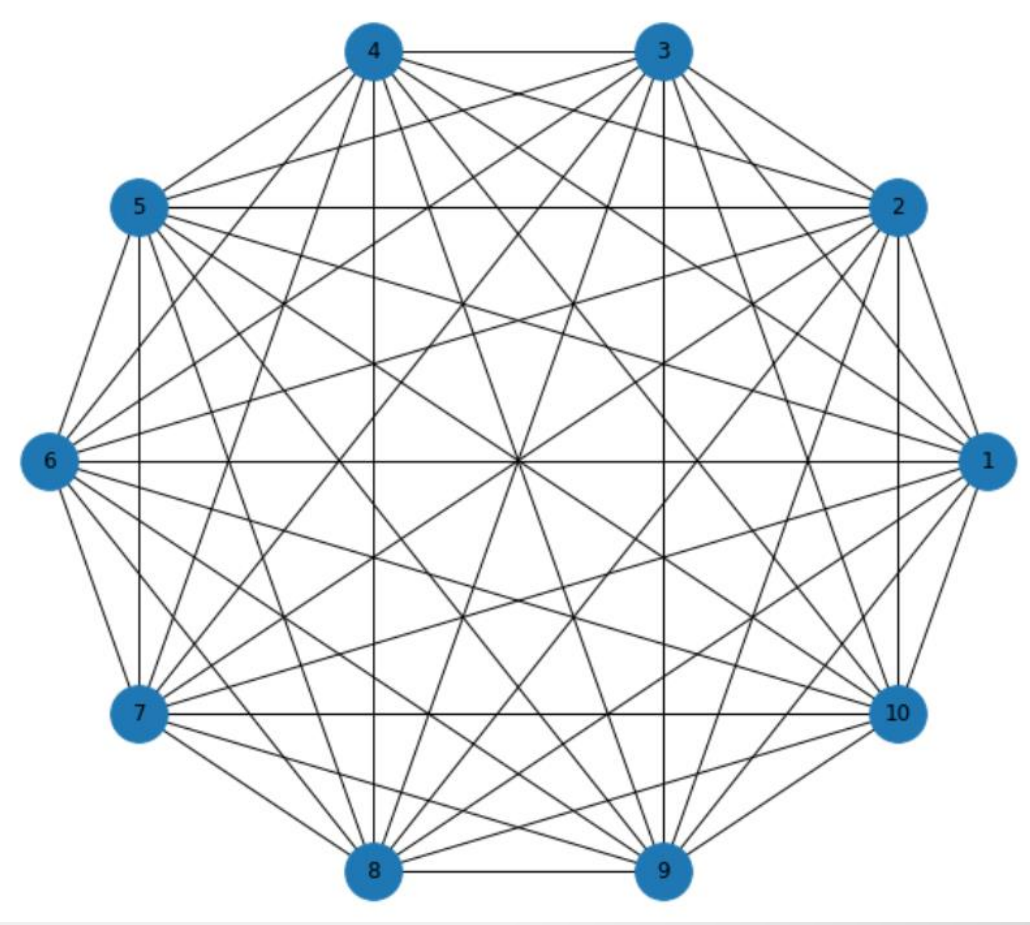
Tandac Furkan Guclu
Sabancı University

ncc@ulakbim.gov.tr

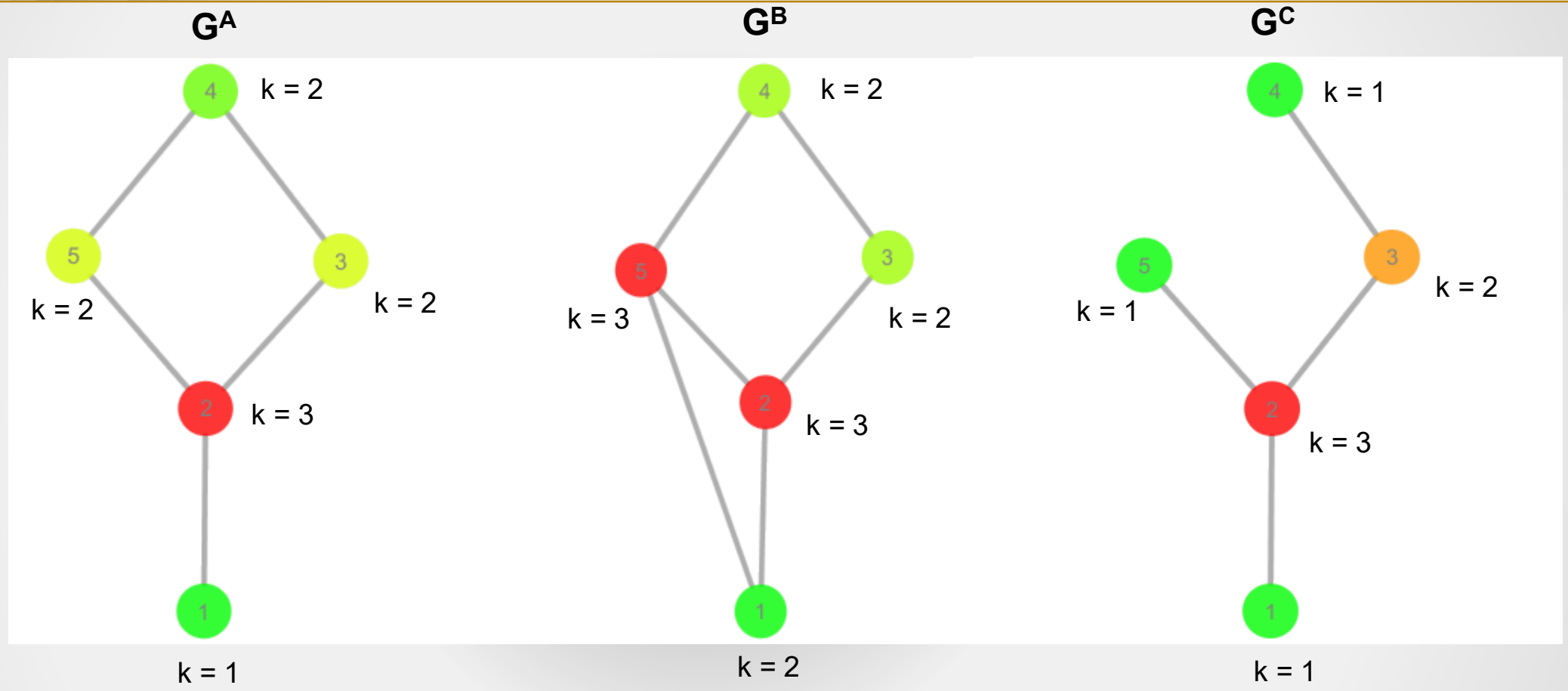
Lesson 1: Introduction to graphs and measures

Graph or network

- Consists of vertices (nodes) and edges (links)
- Nodes and edges may be weighted
- Stored as adjacency matrix or edge/node list
- Here our graphs are
  - Undirected
  - Unweighted
  - No self-loops
  - No parallel edges
- Parameters we use
  - Degree or number of neighbors (*k*)
  - Average shortest path length (<*L*>)
  - Clustering coefficient (*C*)
  - Betweenness centrality (BC)



Degree or number of neighbors (*k*)



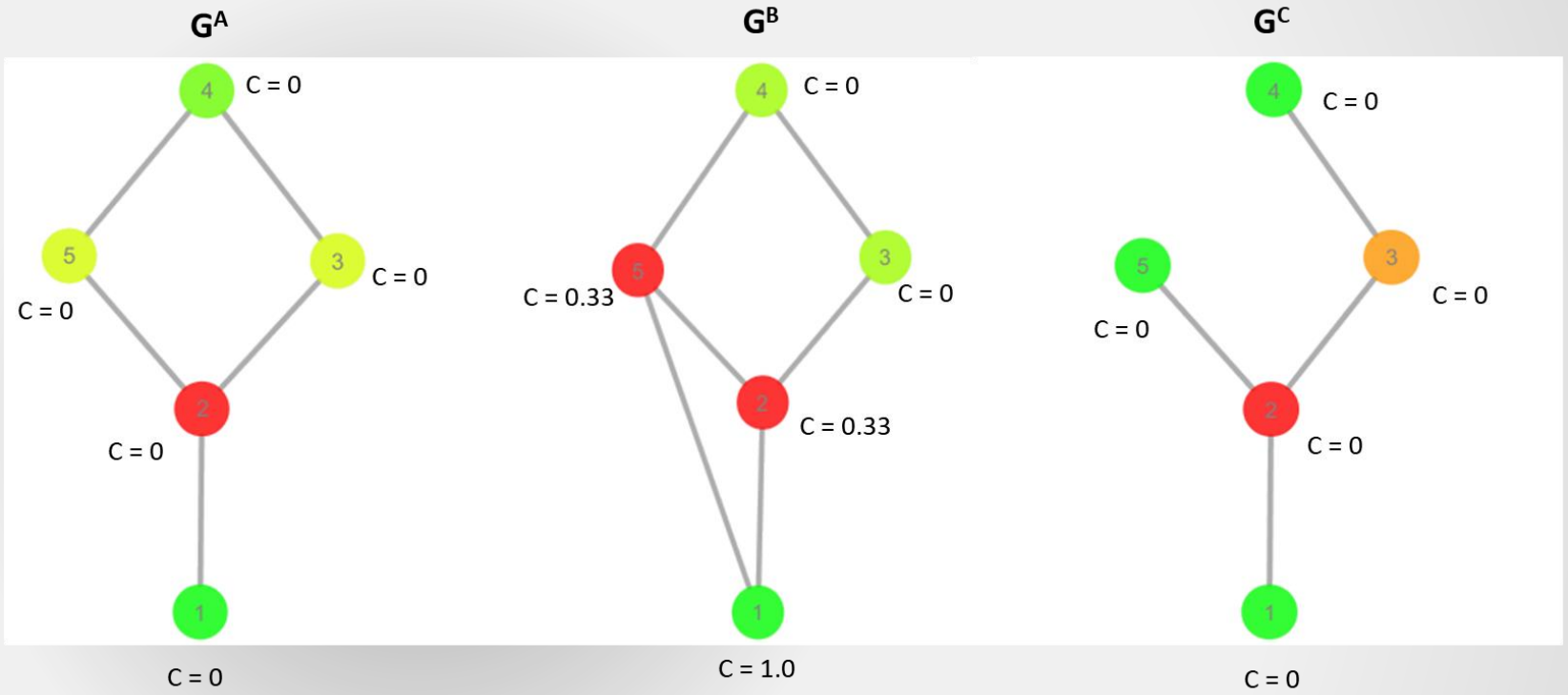
Local information about network.

Clustering coefficient (C)

For G^B:

$$C = \frac{2T(u)}{\mathrm{k}(u)(\mathrm{k}(u)-1)}$$

$$C_{Node2} = \frac{2 \times 1}{3 \times 2} = 0.33$$



Connectivity of the neighbors of a node, semi-local information about connectivity.

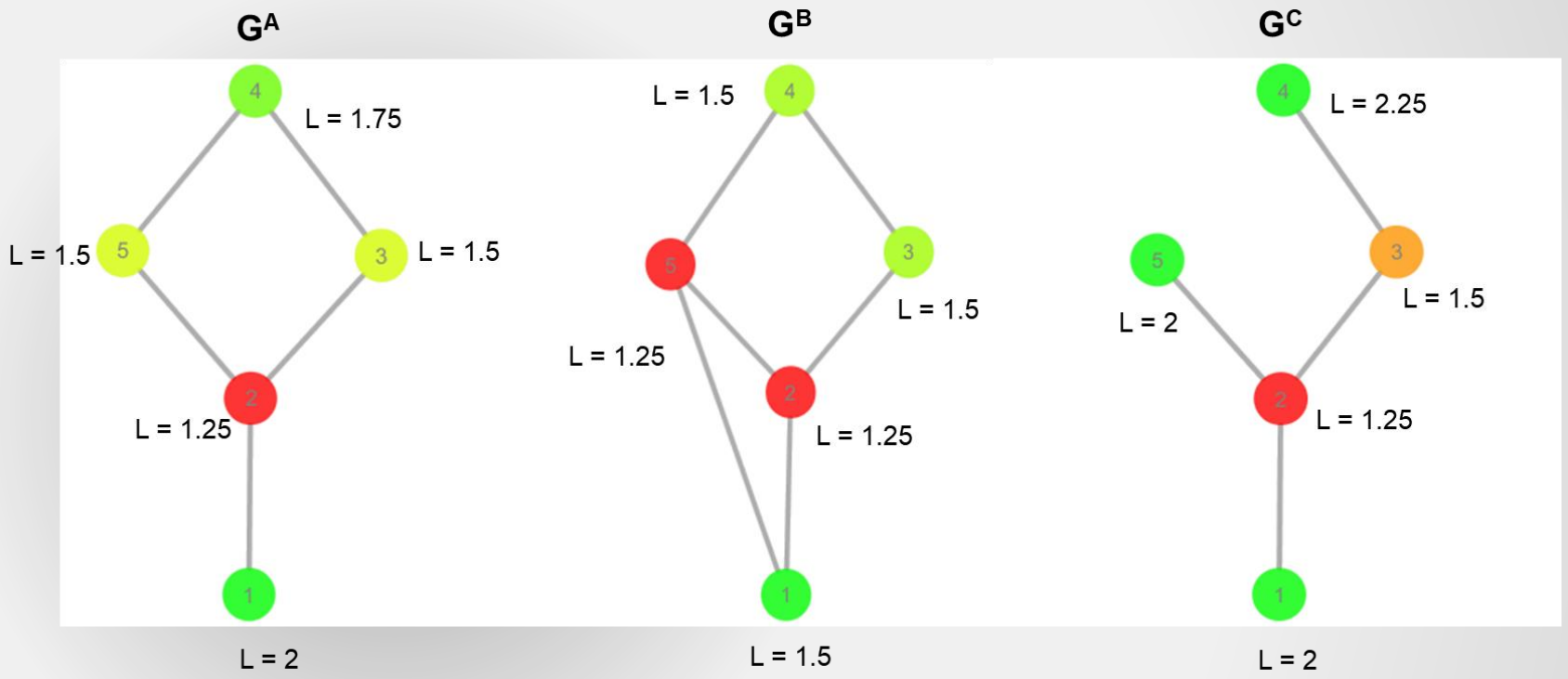
Average Shortest Path Length (<L>)

For G^A:

$$< L_i >= \frac{1}{n-1}\sum_{i}^{n} L_{ij}$$

$$< L >_{Node\ 2} = \frac{L_{21} + L_{25} + L_{23} + L_{24}}{N}$$

$$< L >_{Node\ 2} = \frac{1+1+1+2}{4} = 1.25$$



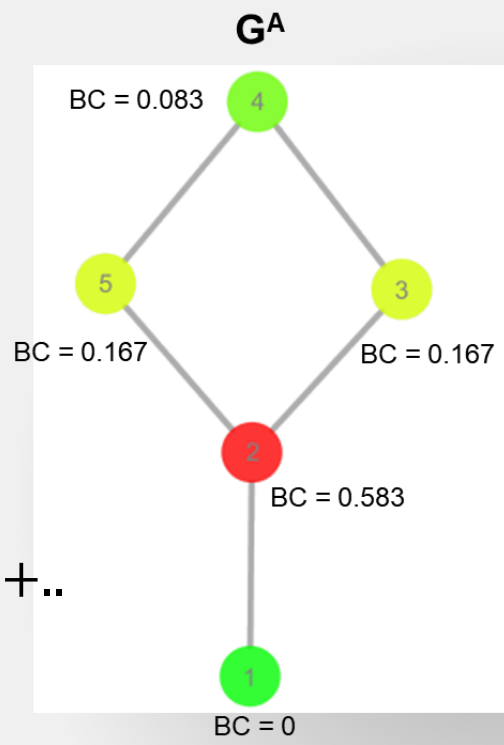
Information about the accessibility of a node in units of edges.

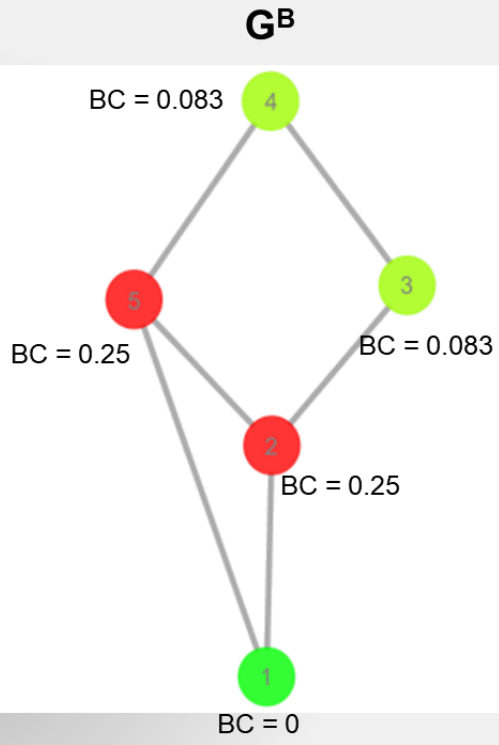
Betweenness centrality (BC)

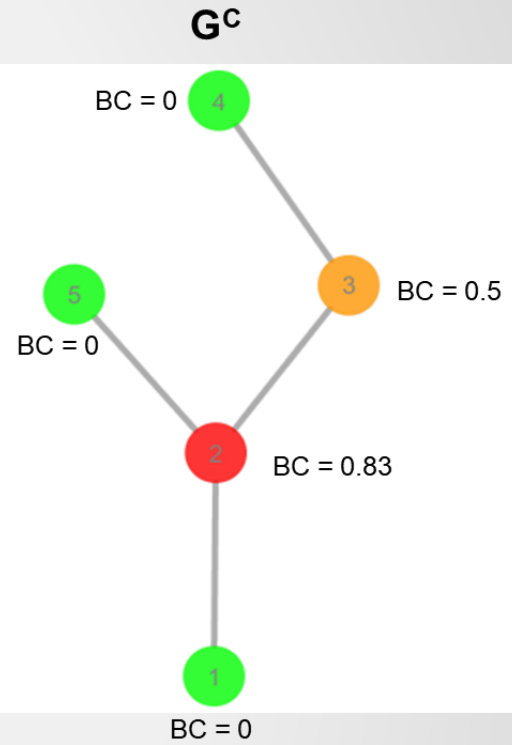
For G^A:

$$BC(X) = \sum \left(\frac{P(a,b|X)}{P(a,b)} \right)$$

$$BC_{Node\ 2} = \frac{P_{13}(2)}{P_{13}} + \frac{P_{15}(2)}{P_{15}} + \frac{P_{14}(2)}{P_{14}} + ..$$





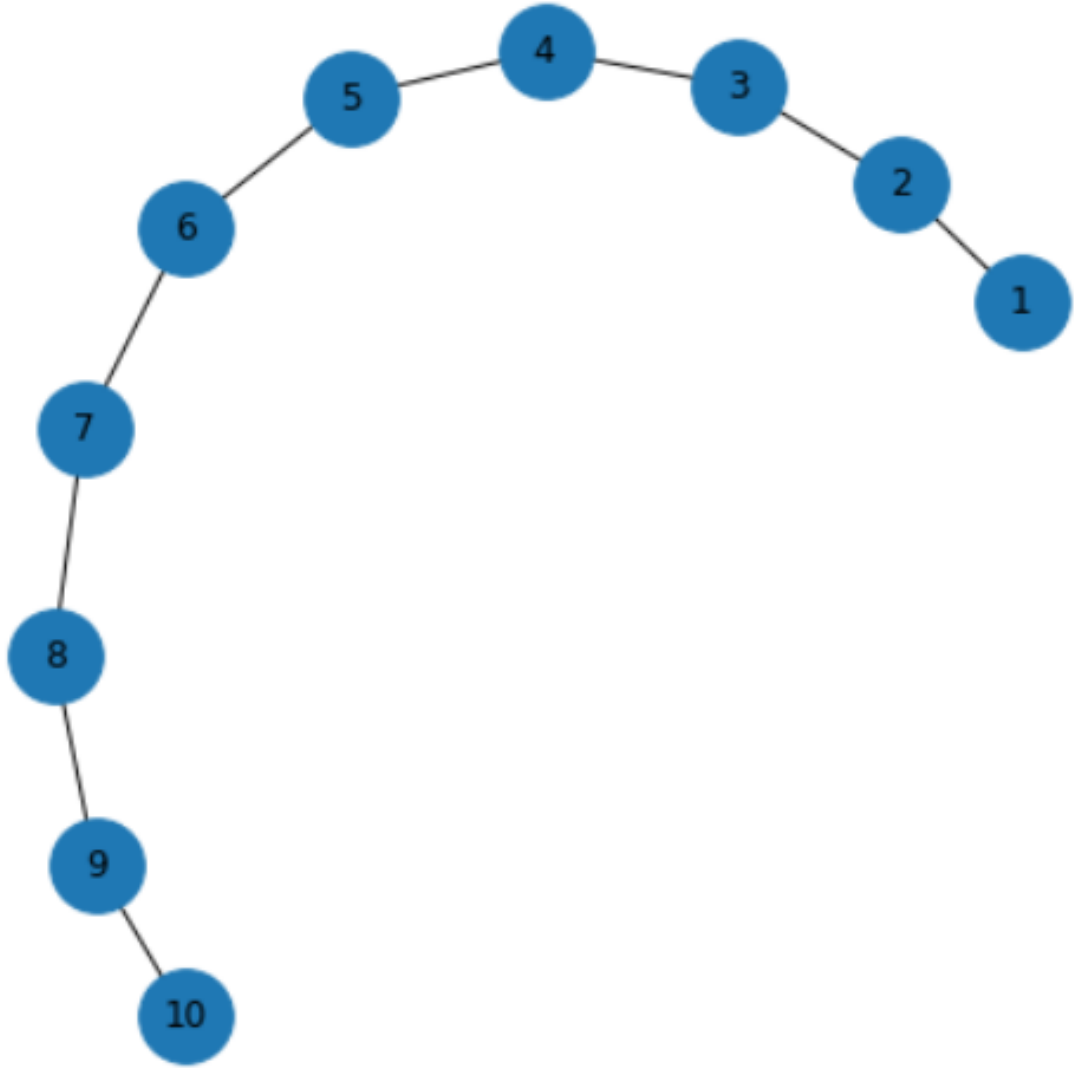


$$BC_{Node\ 2} = \frac{1}{1} + \frac{1}{1} + \frac{2}{2} + 0 + \frac{1}{2} + 0 = 3.5$$

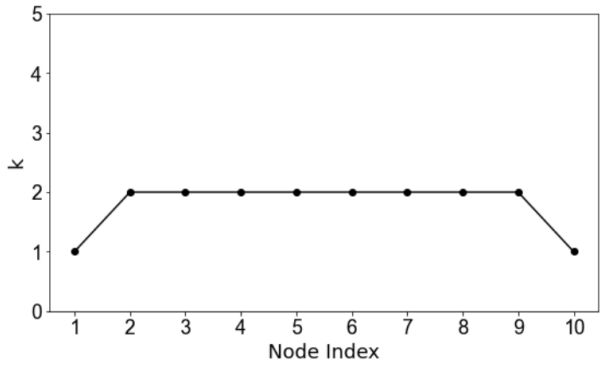
Further normalized by (n-1)(n-2)/2

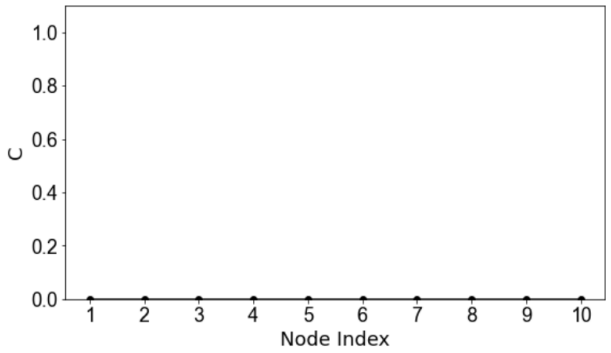
$$BC_{Node\ 2} = \frac{3.5}{6} = 0.583$$

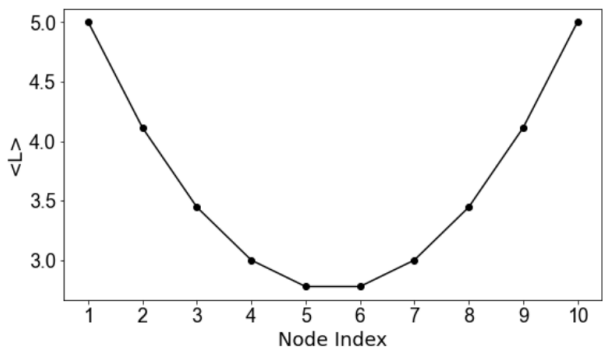
More toy graphs…

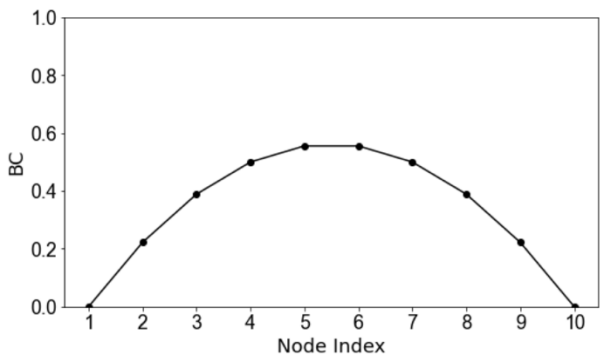


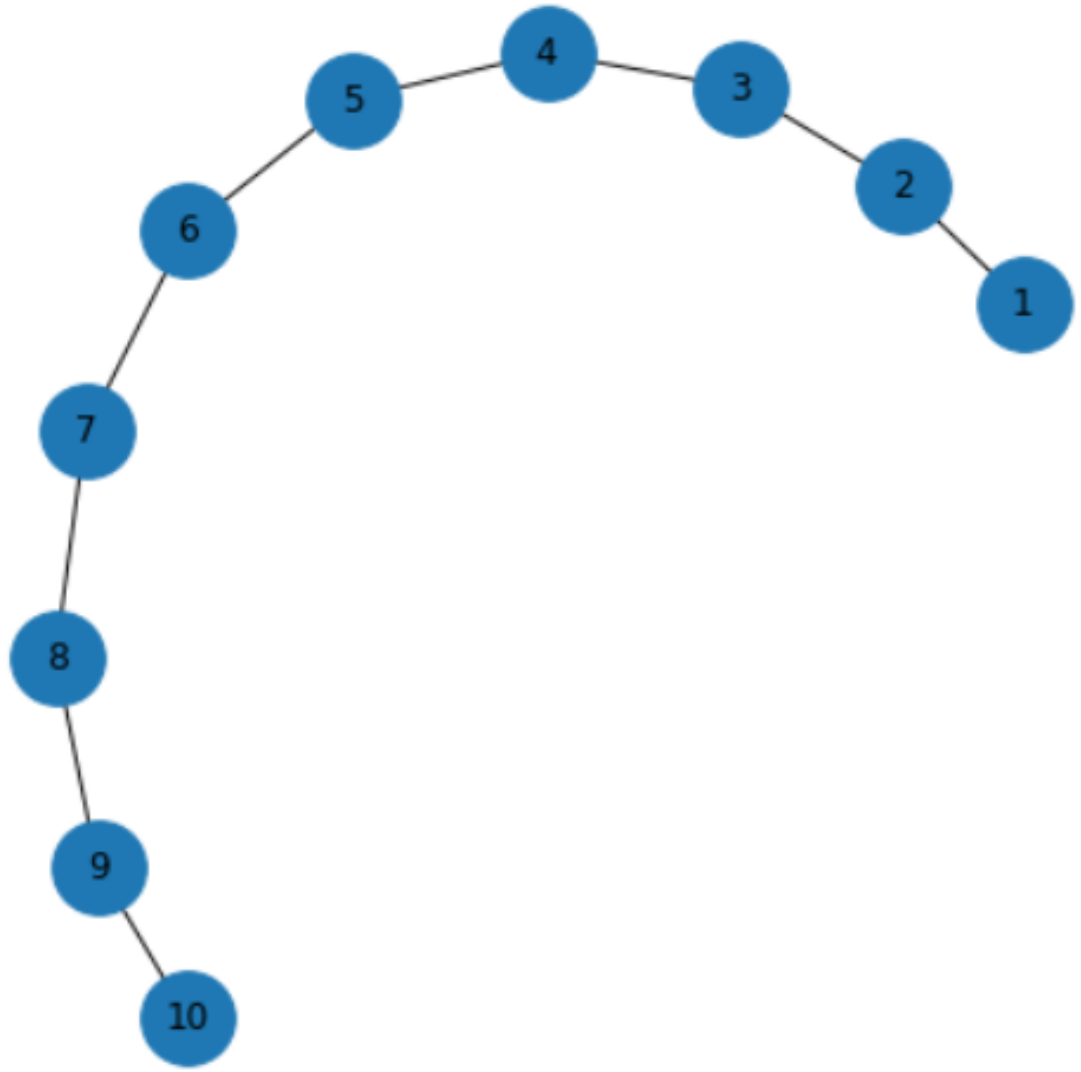
Number of nodes: 10
Number of edges: 9
Average degree: 1.8



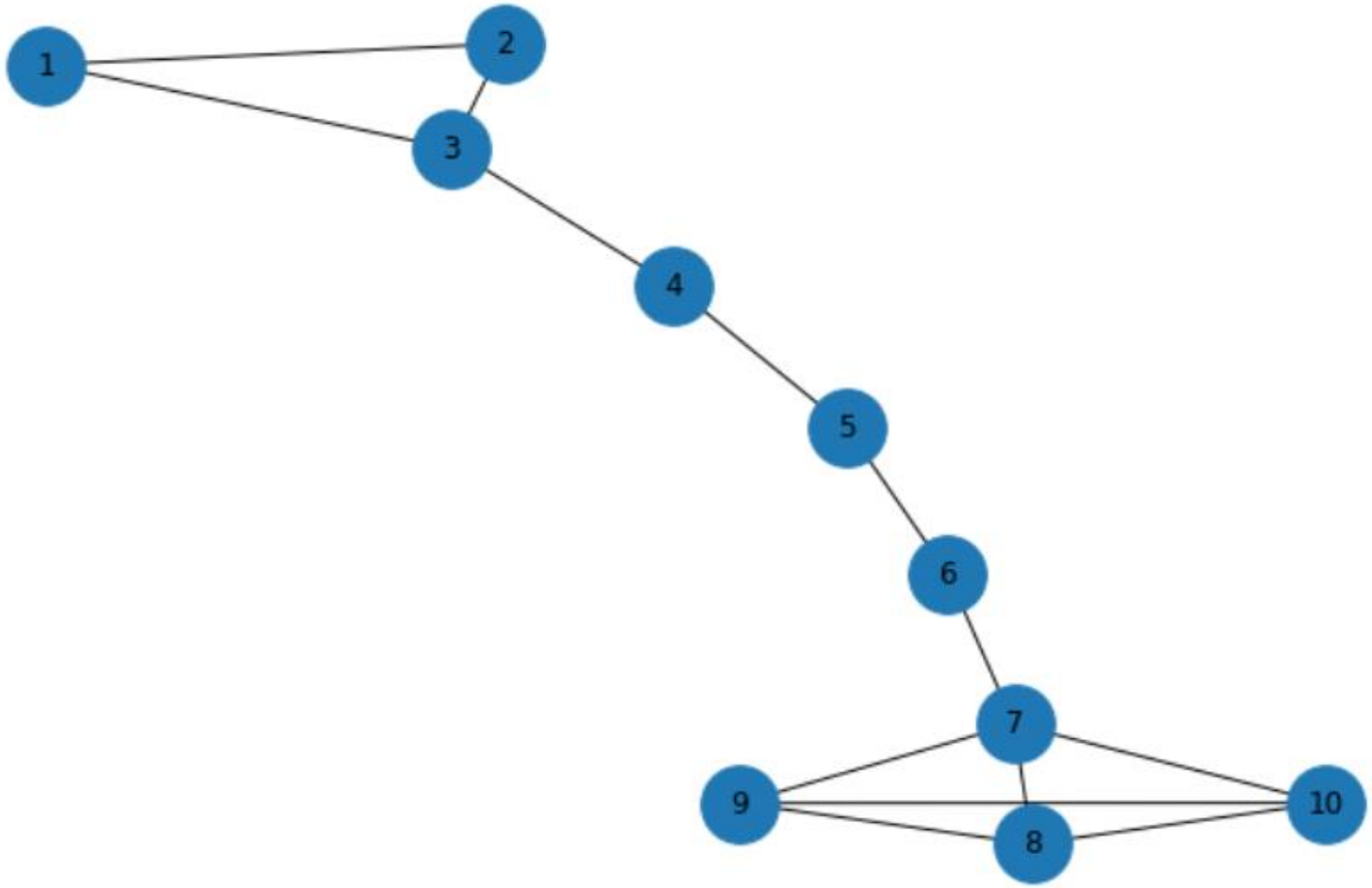






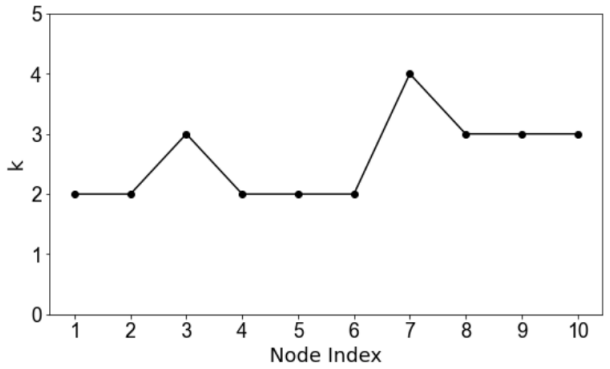


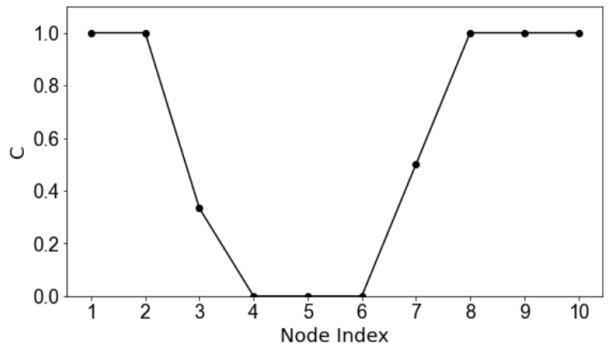
Number of nodes: 10
Number of edges: 9
Average degree: 1.8

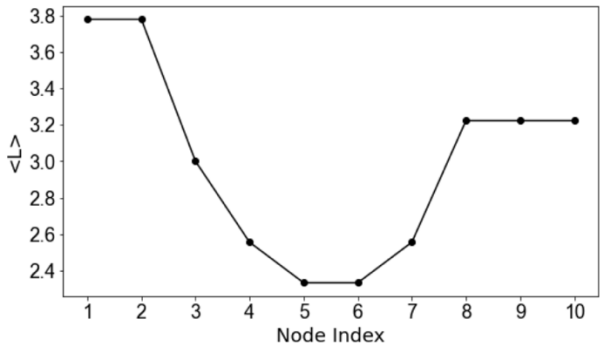


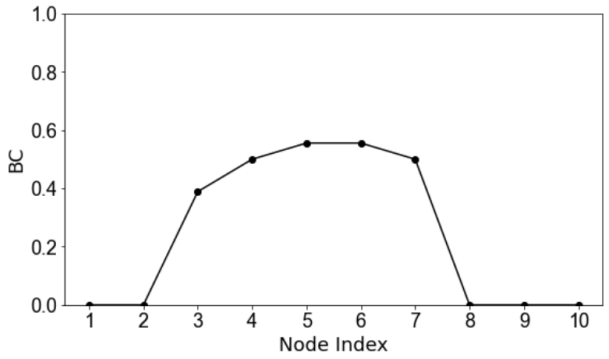
Number of nodes: 10
Number of edges: 13
Average degree: 2.6

[1,3]
[8,10]
[7,10]
[7,9]



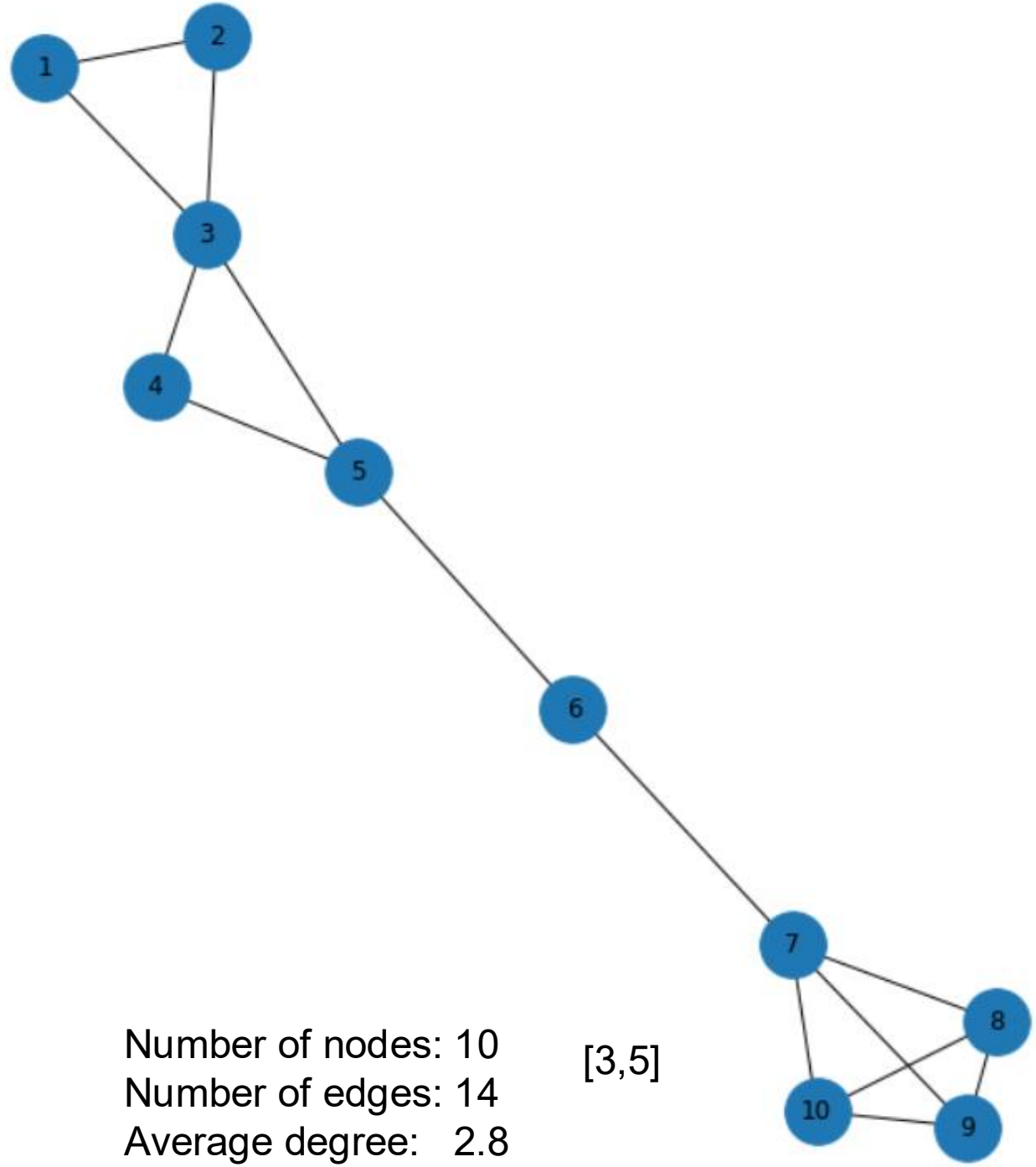






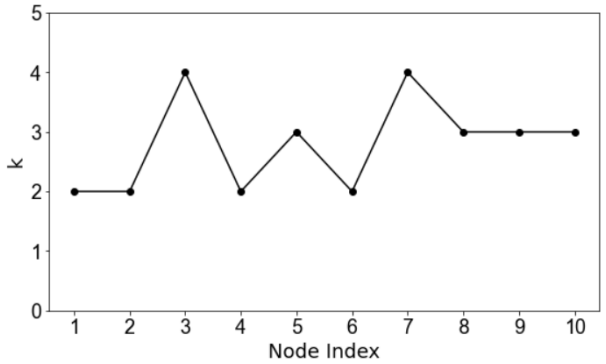
Number of nodes: 10
Number of edges: 13
Average degree: 2.6

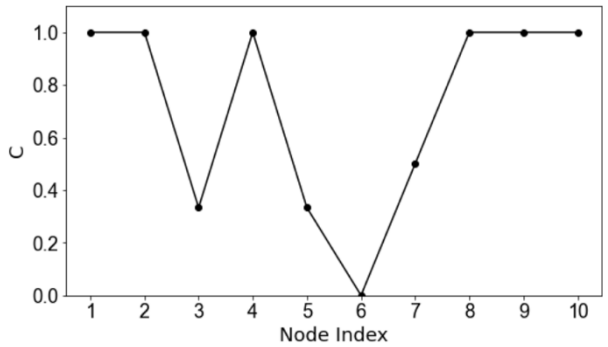
[1,3]
[8,10]
[7,10]
[7,9]

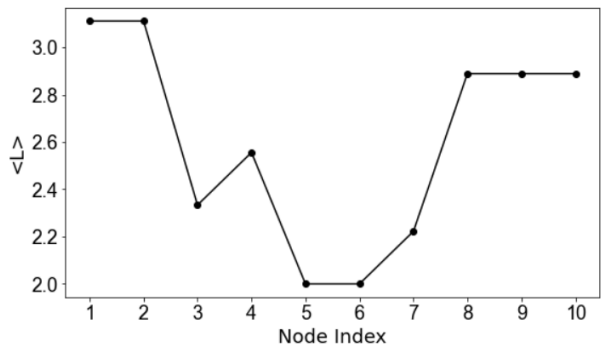


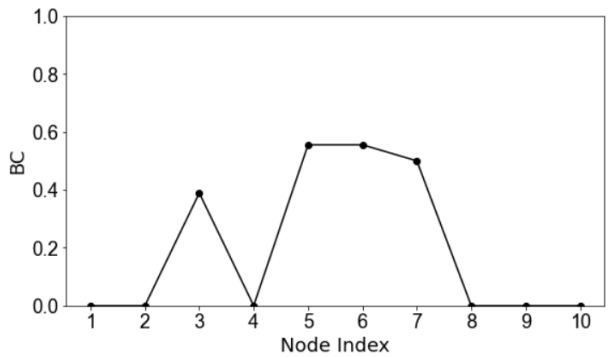
Number of nodes: 10
Number of edges: 14
Average degree: 2.8

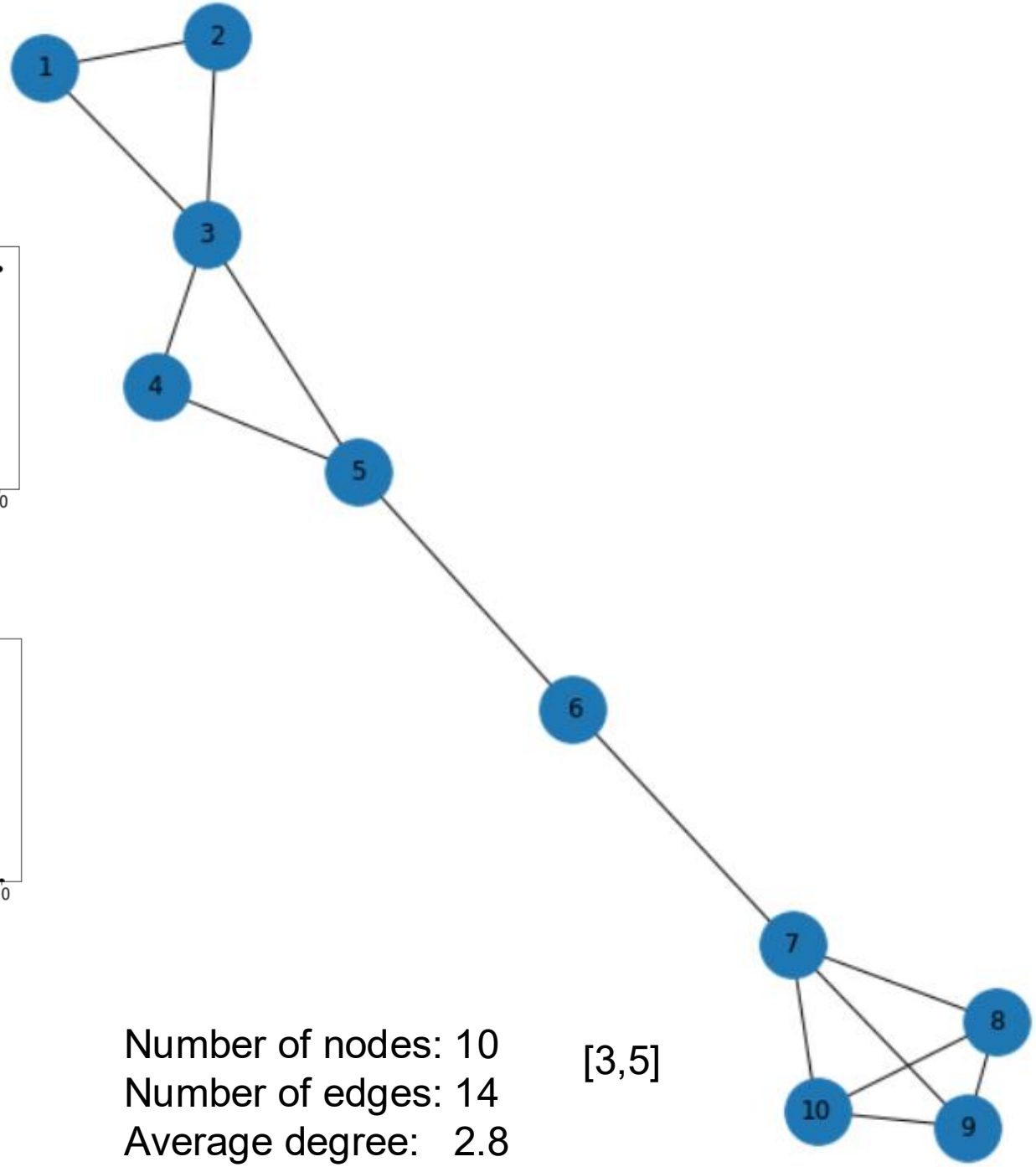
[3,5]





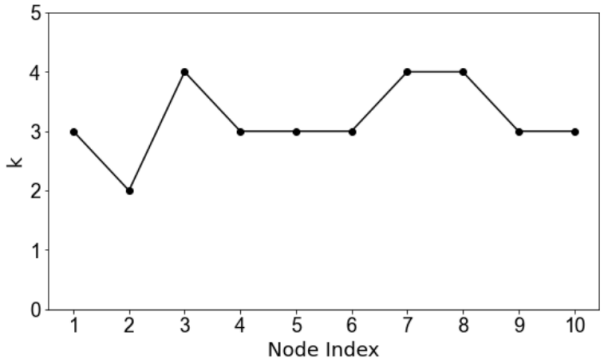


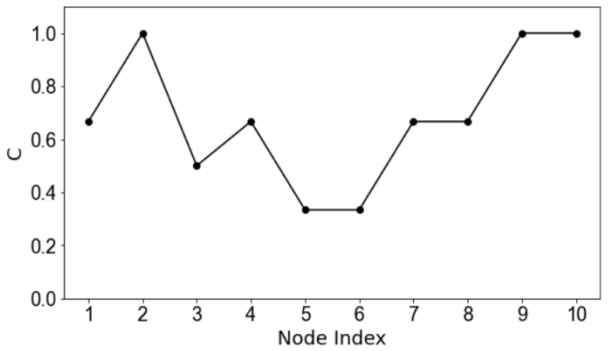


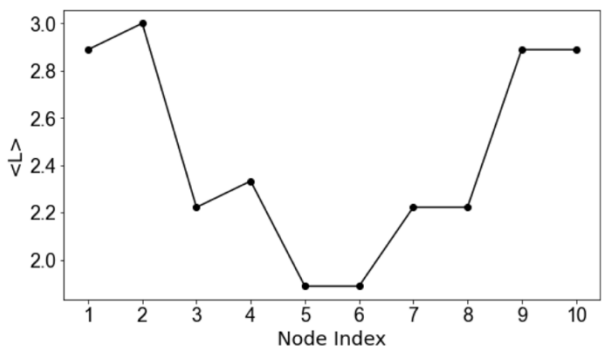


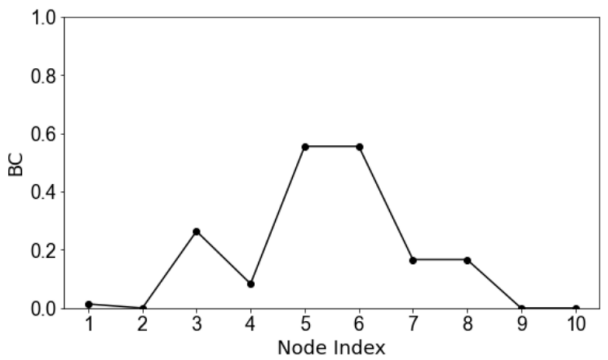
Number of nodes: 10
Number of edges: 14
Average degree: 2.8

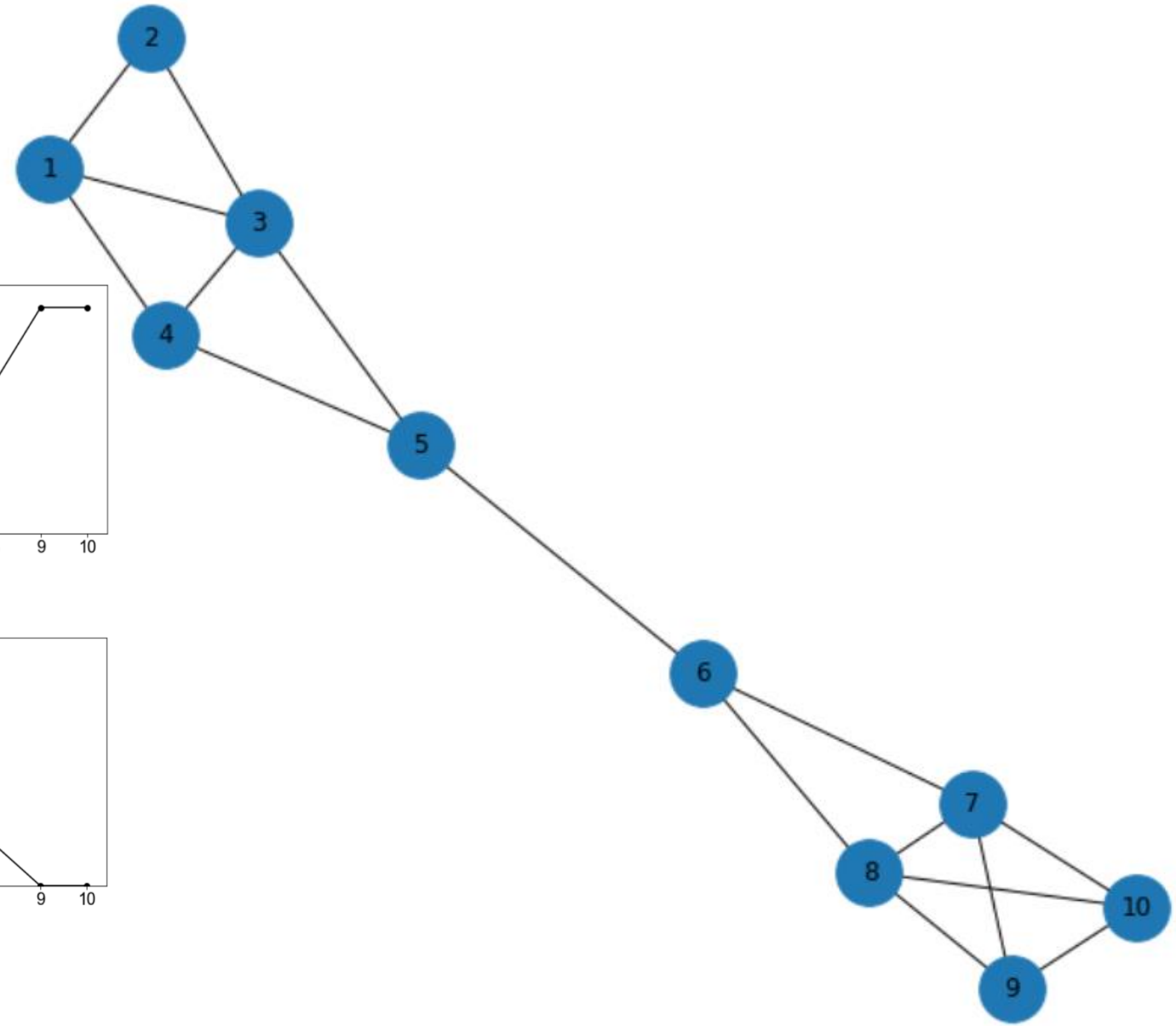
[3,5]





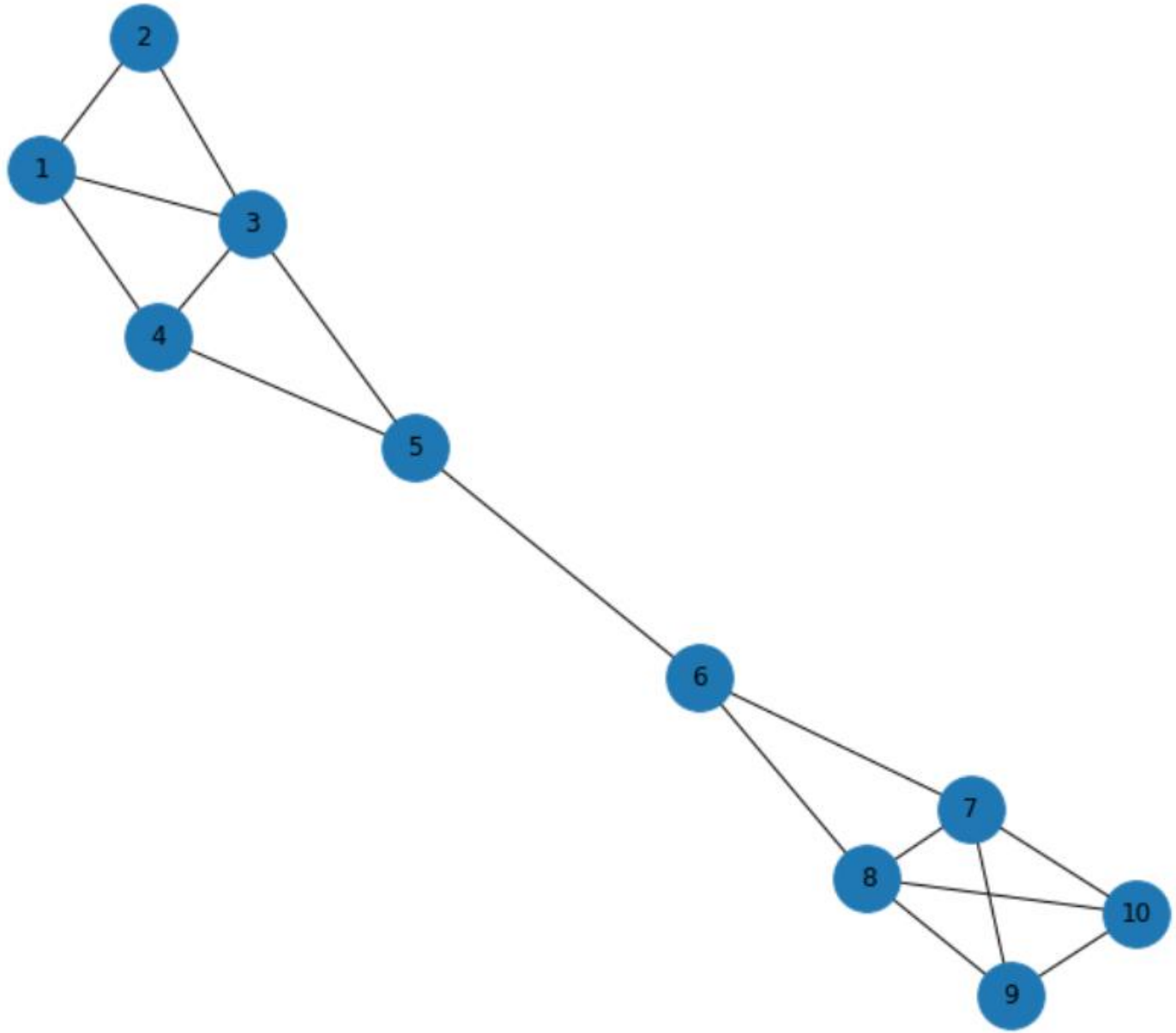






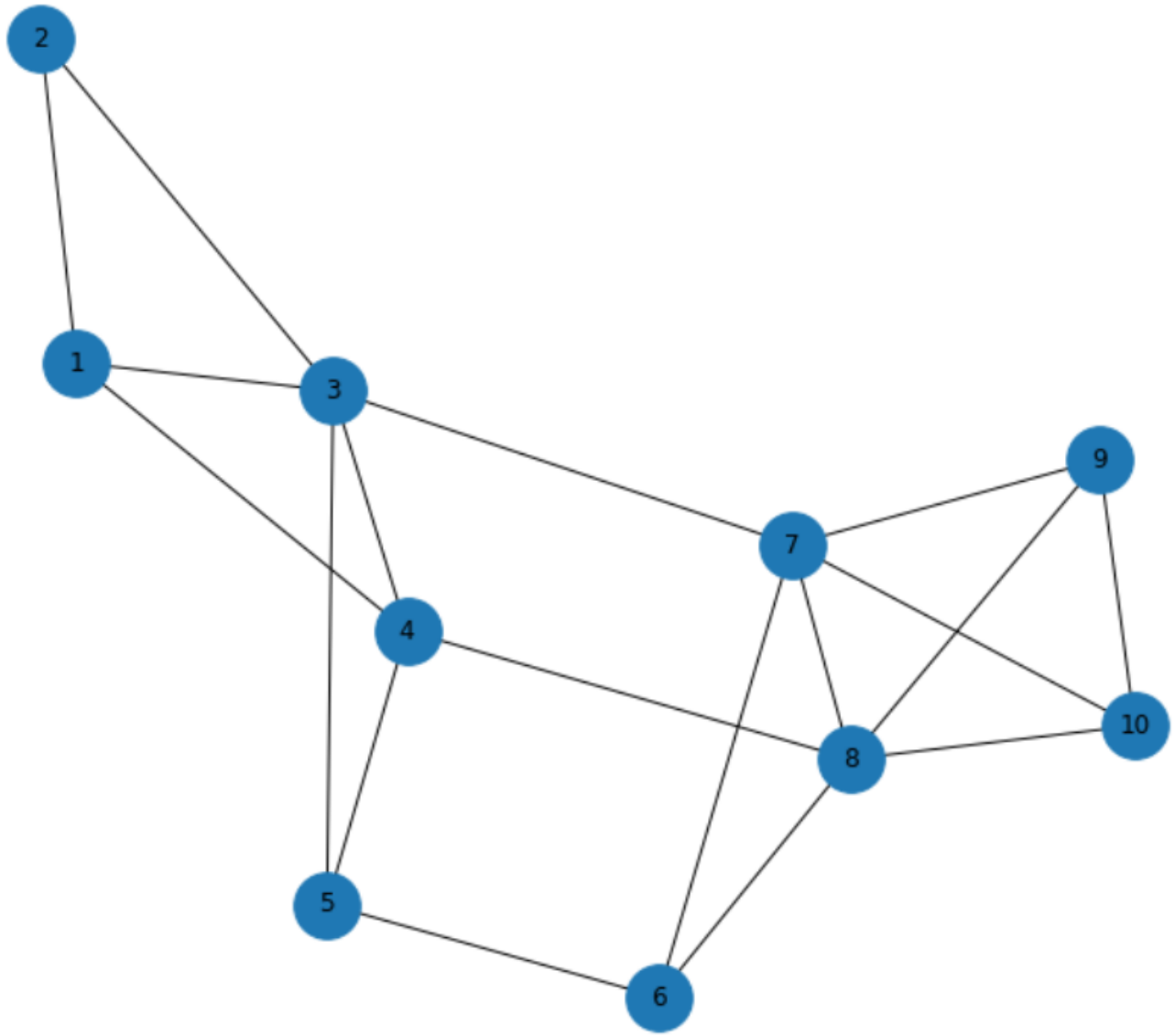
Number of nodes: 10
Number of edges: 16
Average degree: 3.2

[1,4]
[6,8]

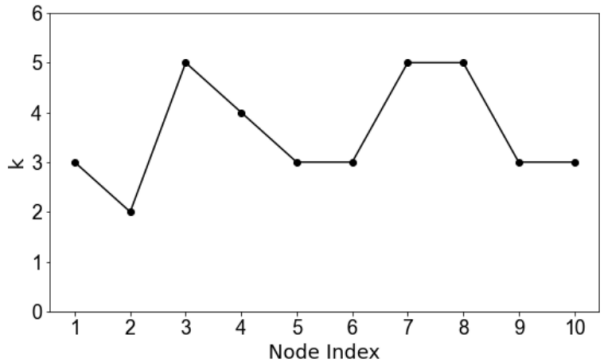


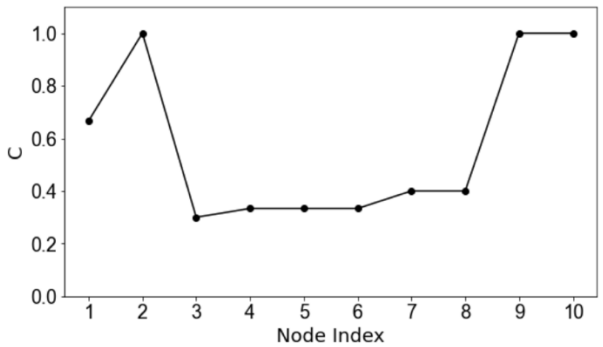
Number of nodes: 10
Number of edges: 16
Average degree: 3.2

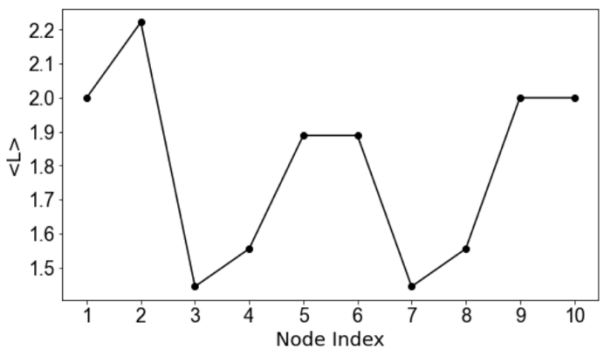
[1,4]
[6,8]

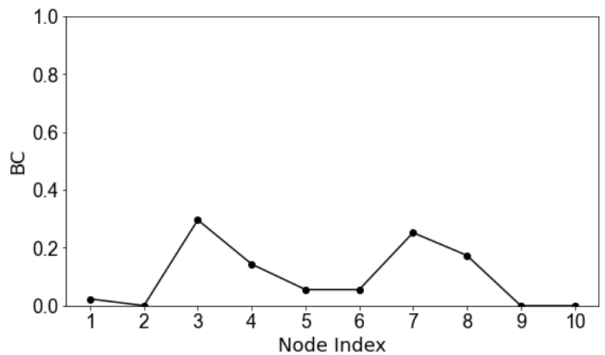


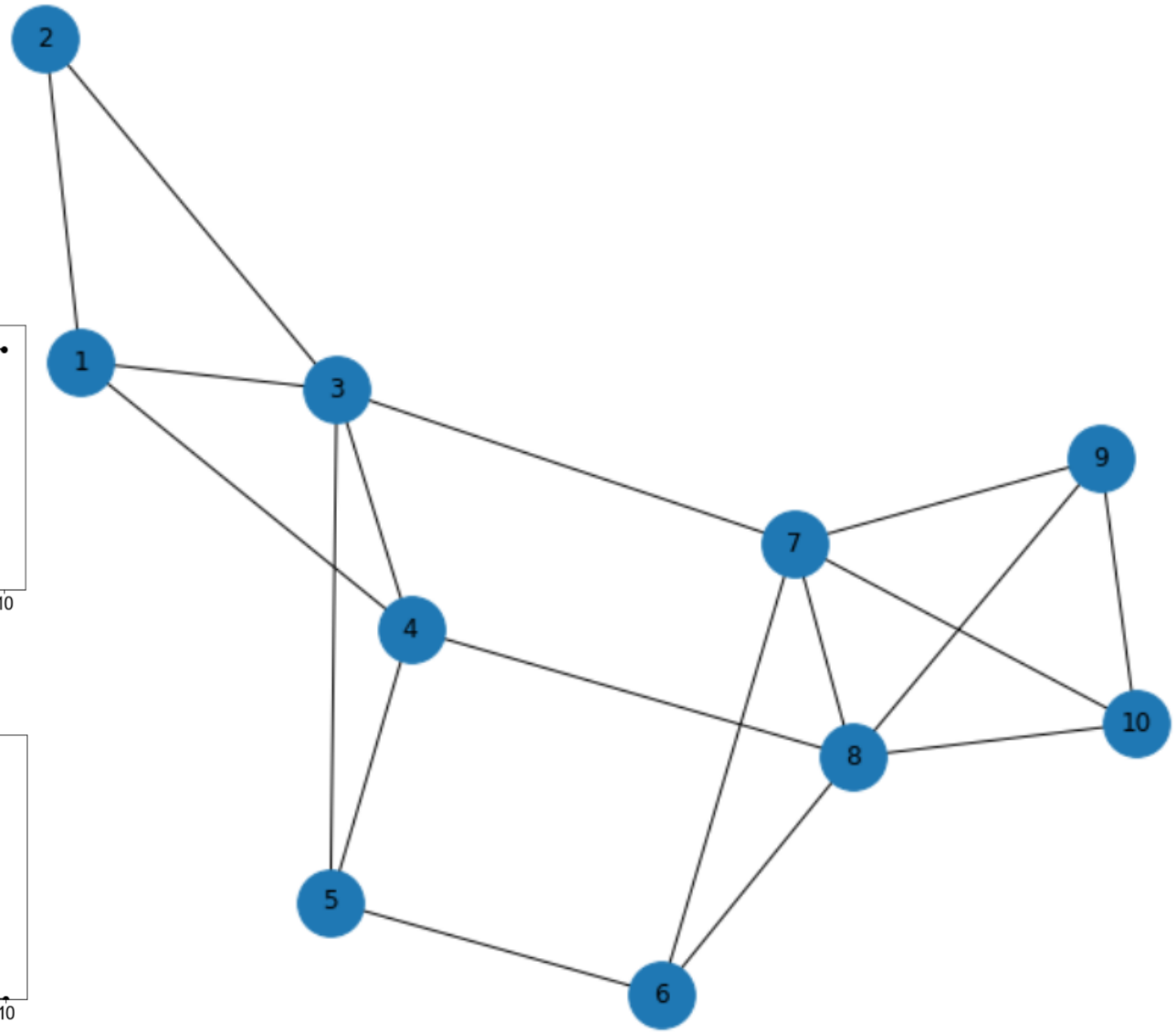
Number of nodes: 10 [3,7]
Number of edges: 18 [4,8]
Average degree: 3.6



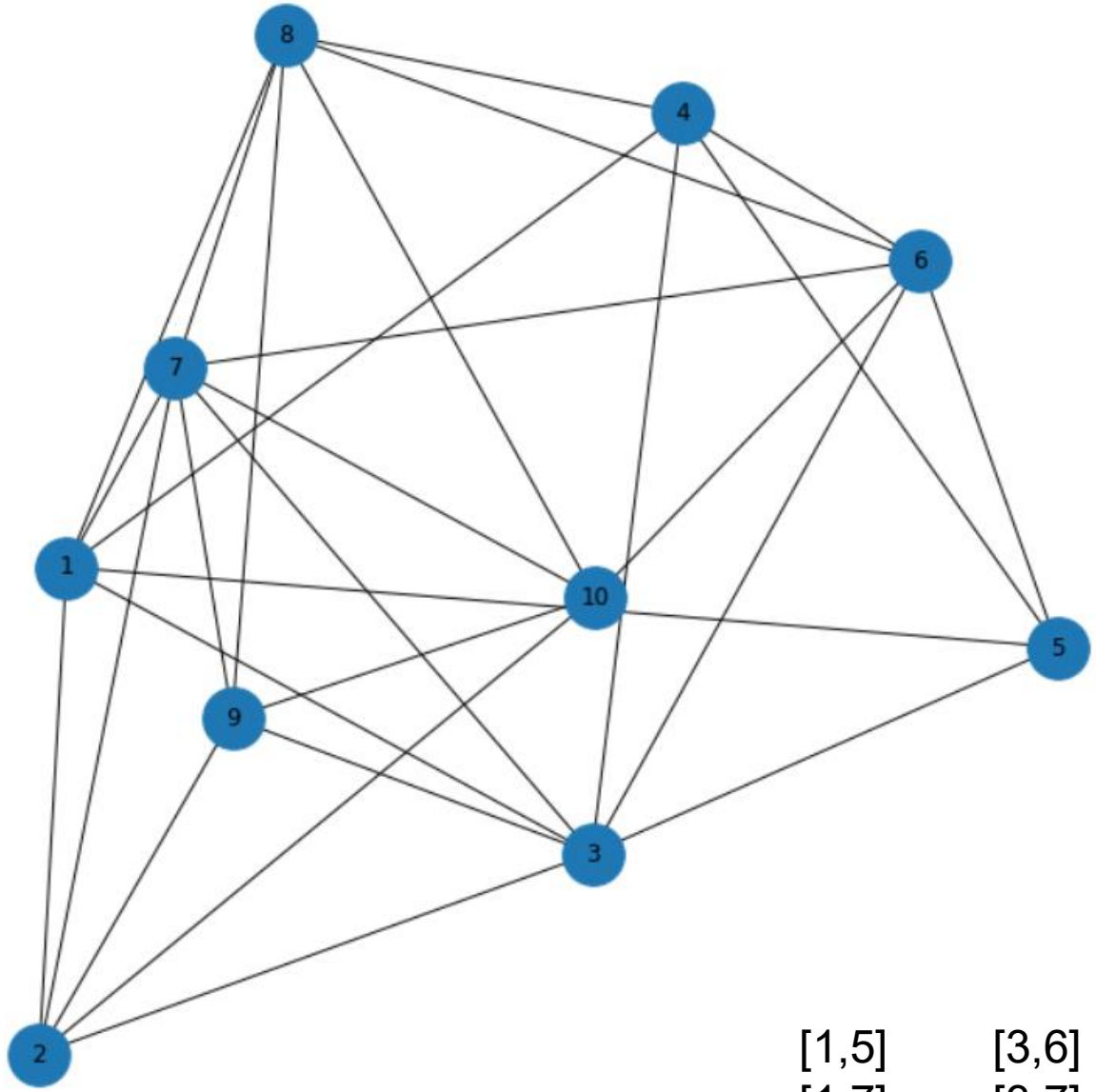






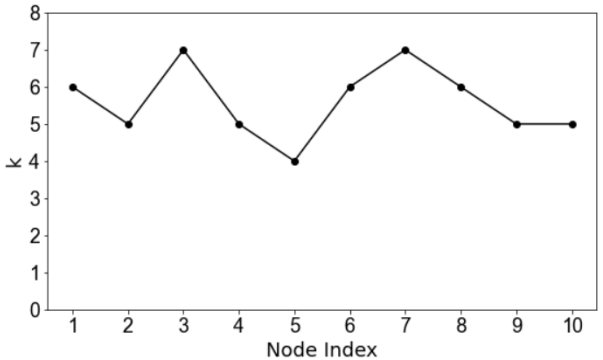


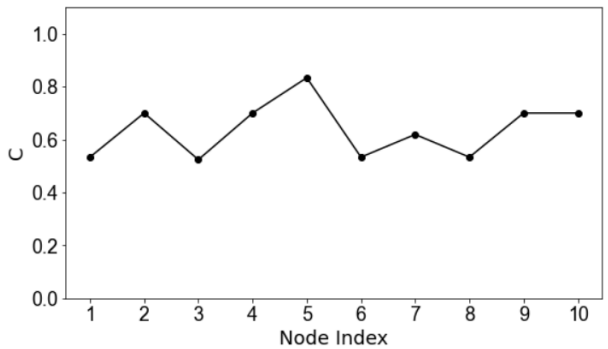
Number of nodes: 10 [3,7]
Number of edges: 18 [4,8]
Average degree: 3.6

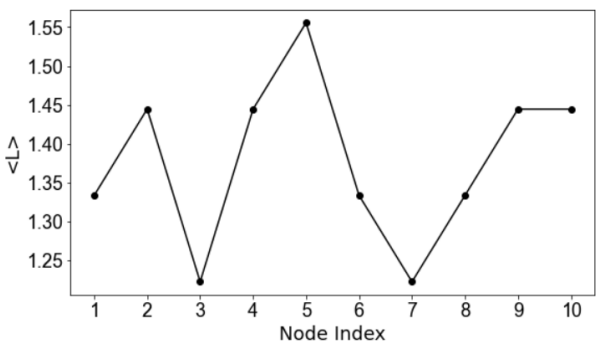


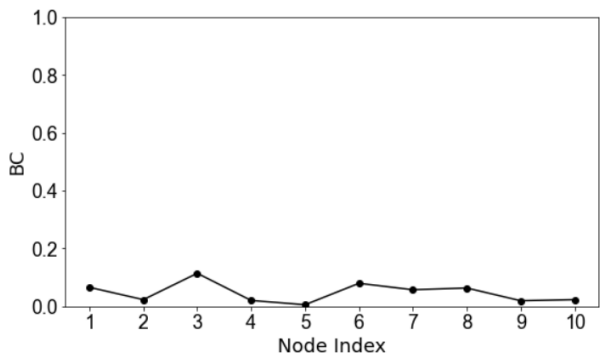
Number of nodes: 10
Number of edges: 28
Average degree: 5.6

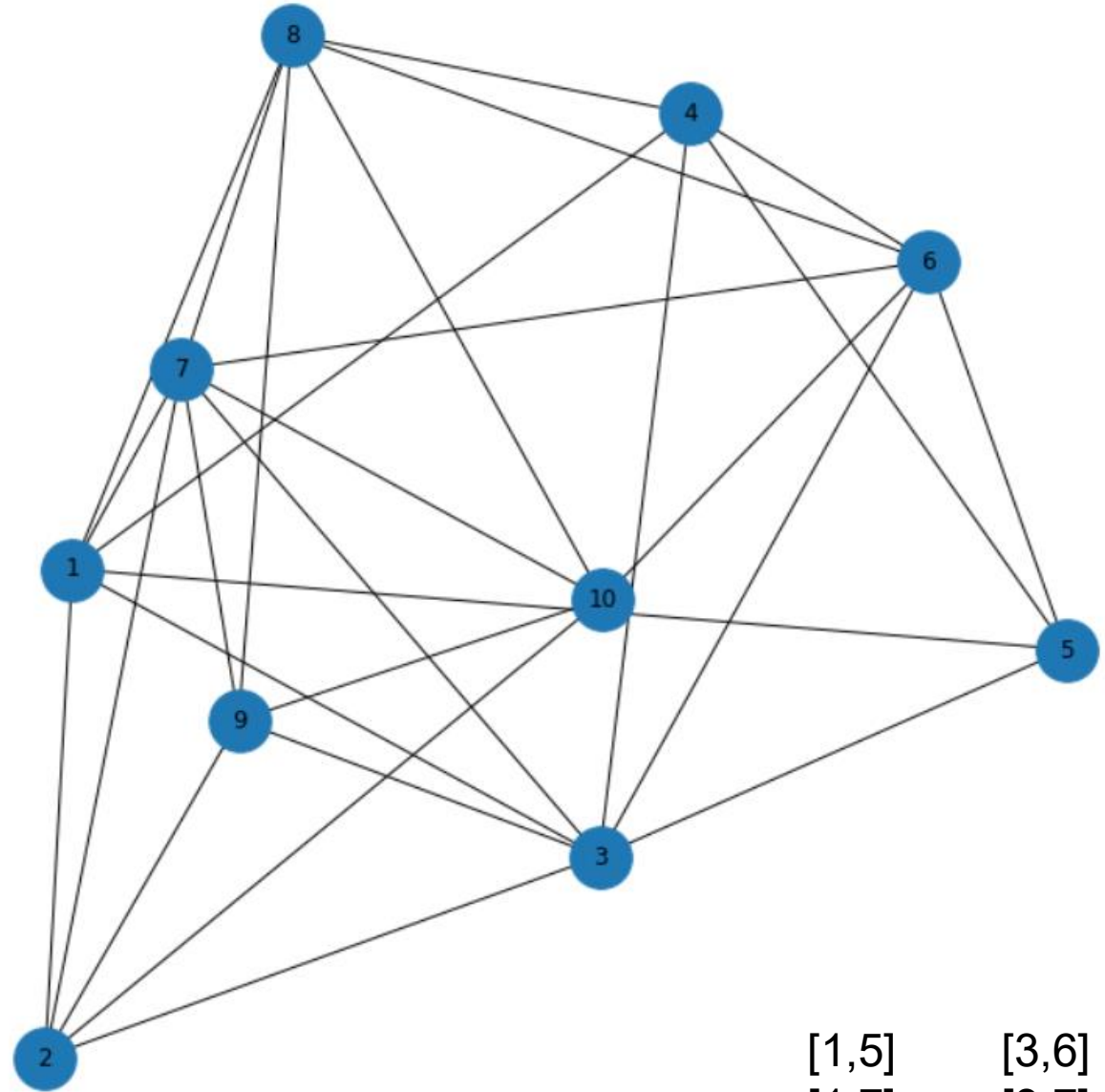
[1,5]
[1,7]
[1,8]
[2,7]
[2,9]
[2, 10]
[3,6]
[3,7]
[3,9]
[4,6]
[6,10]





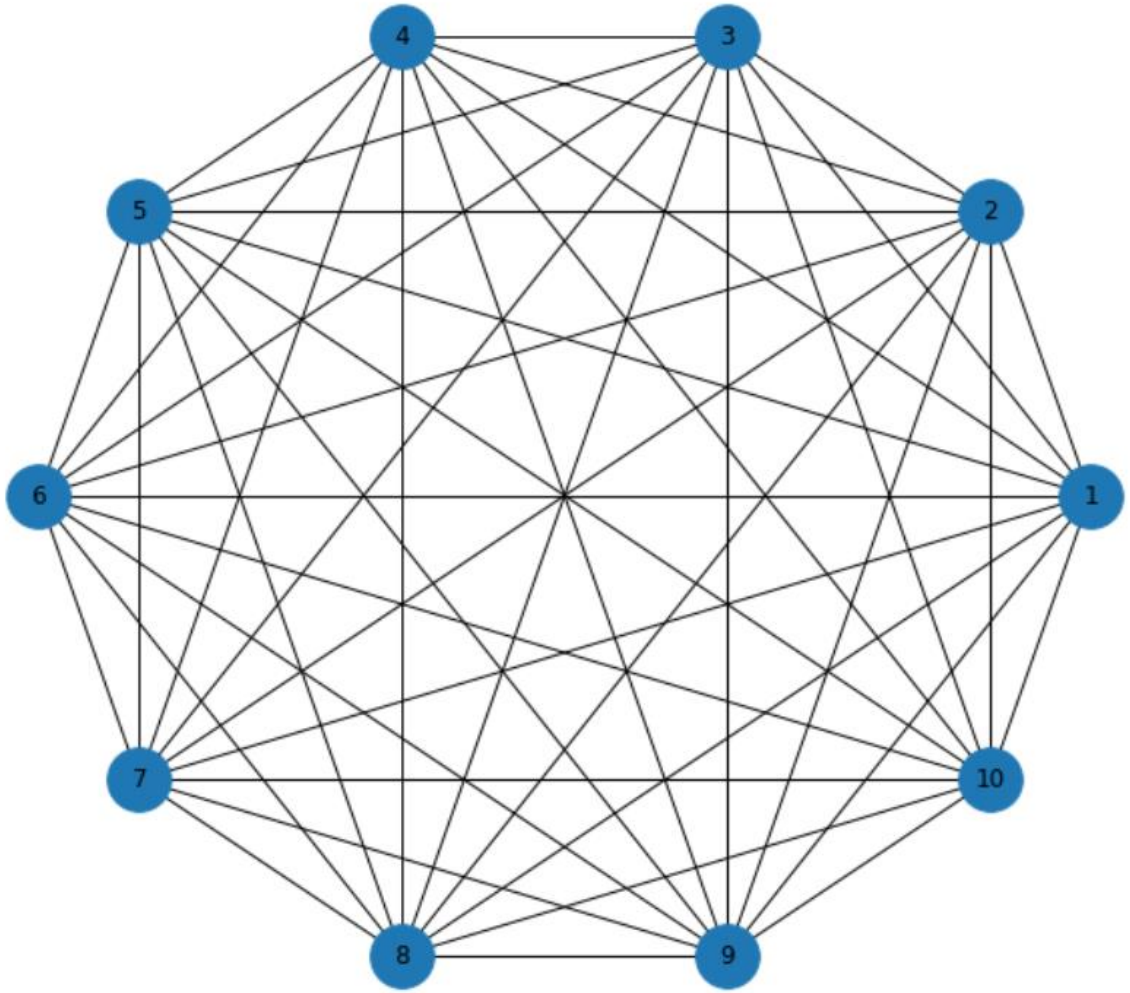






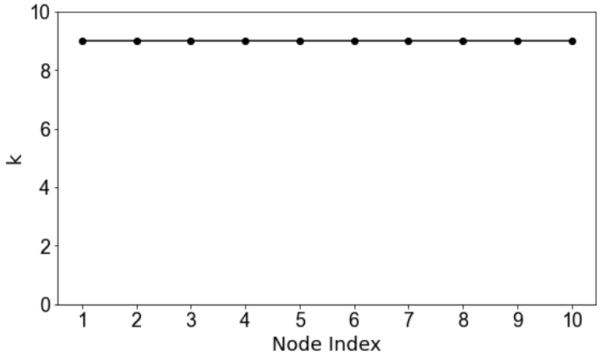
Number of nodes: 10
Number of edges: 28
Average degree: 5.6

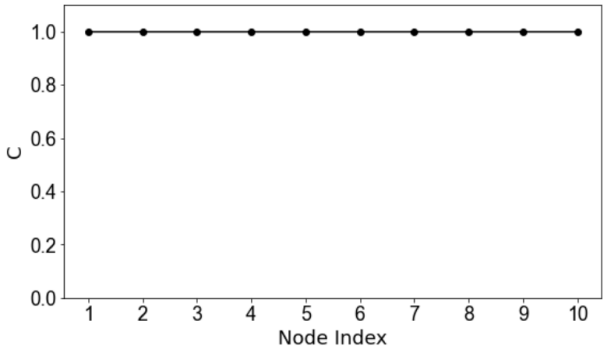
[1,5]
[1,7]
[1,8]
[2,7]
[2,9]
[2, 10]
[3,6]
[3,7]
[3,9]
[4,6]
[6,10]

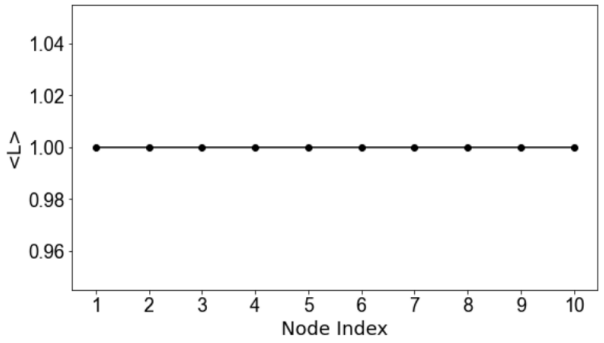


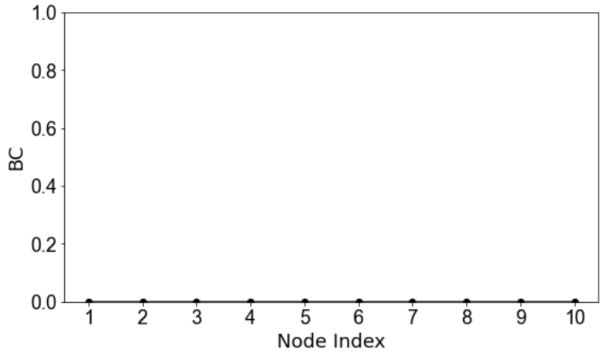
Number of nodes: 10
Number of edges: 45
Average degree: 9

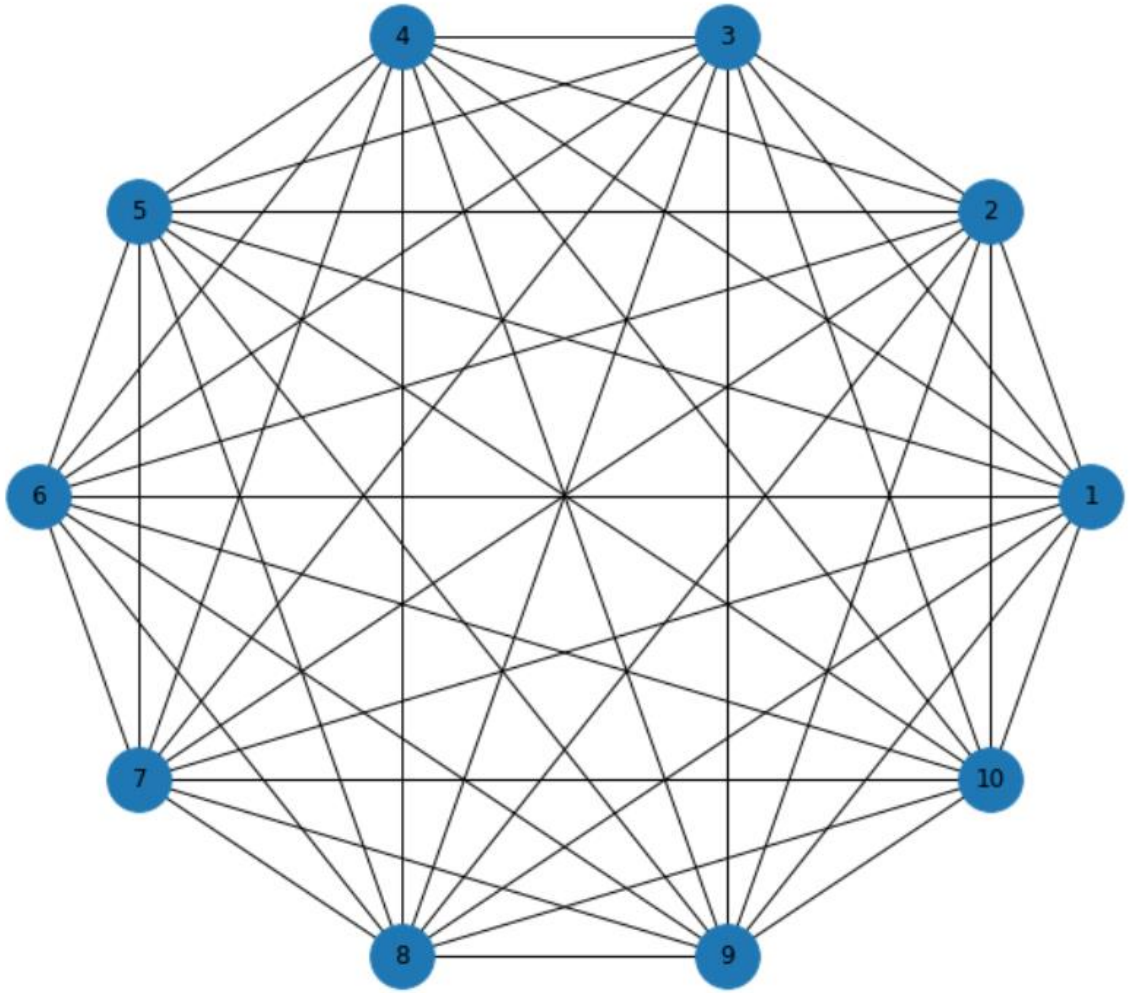
Complete network











Number of nodes: 10
Number of edges: 45
Average degree: 9

Complete network

Next lecture

Code for these toy graphs…

EURO 4SEE

Thanks!

Co-funded by
the European Union

EuroHPC
Joint Undertaking

This project has received funding from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No 101191697. The JU receives support from the Digital Europe Programme and Germany, Türkiye, Republic of North Macedonia, Montenegro, Serbia, Bosnia and Herzegovina.