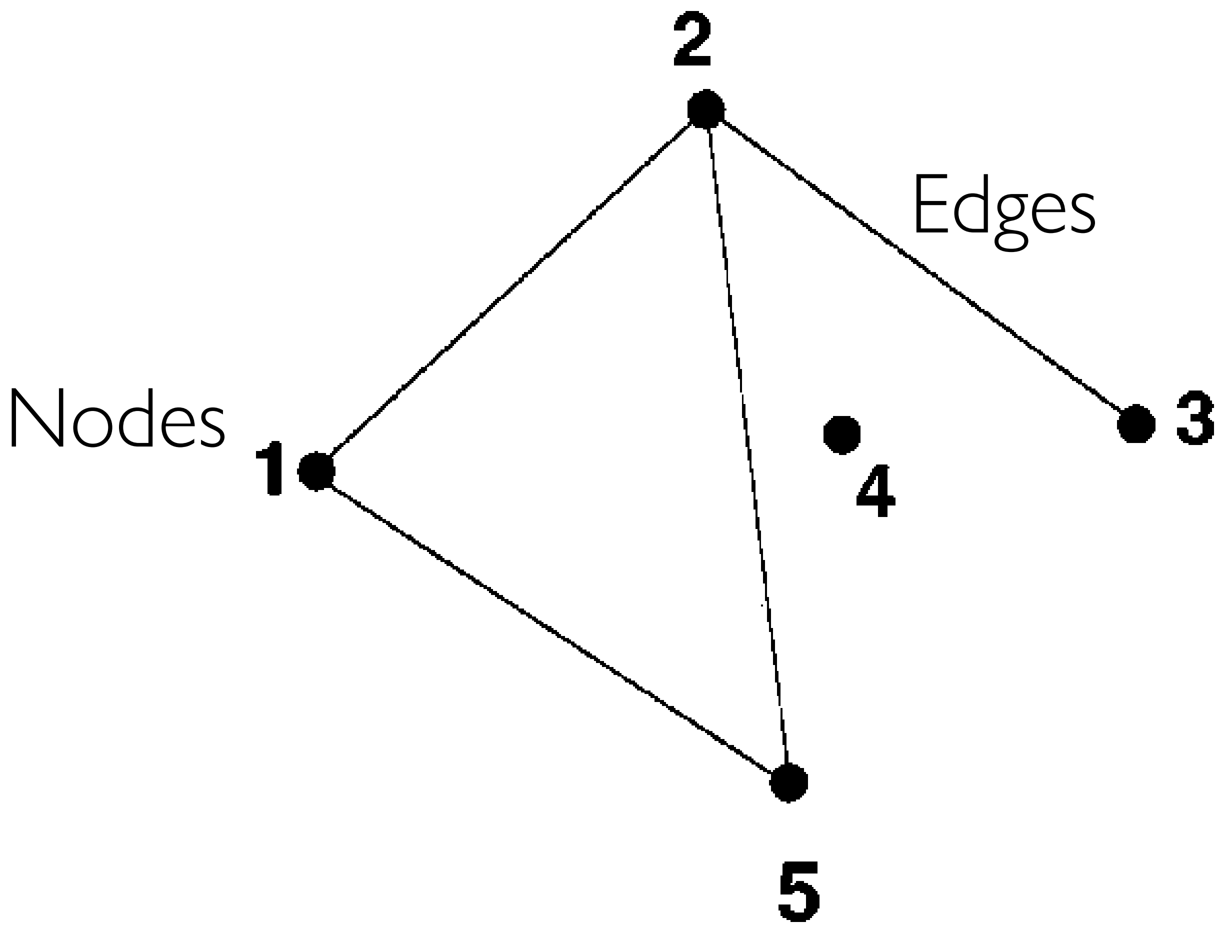
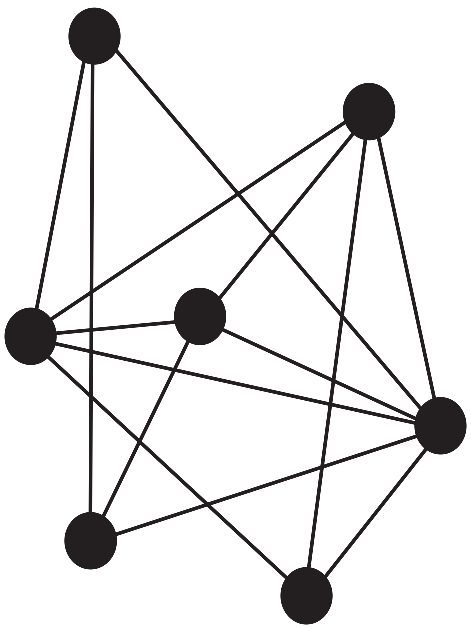
Graph Theory and Connectivity

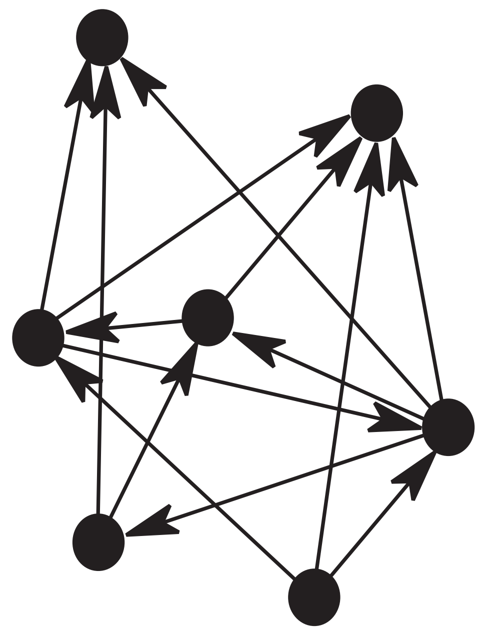
What is a graph?



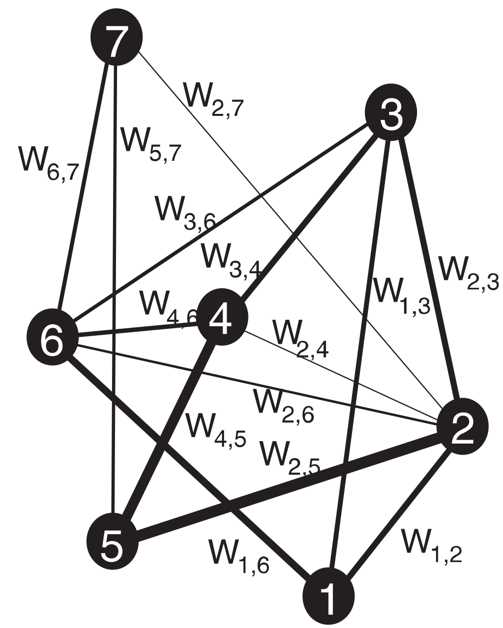
Different types of graphs



Directed

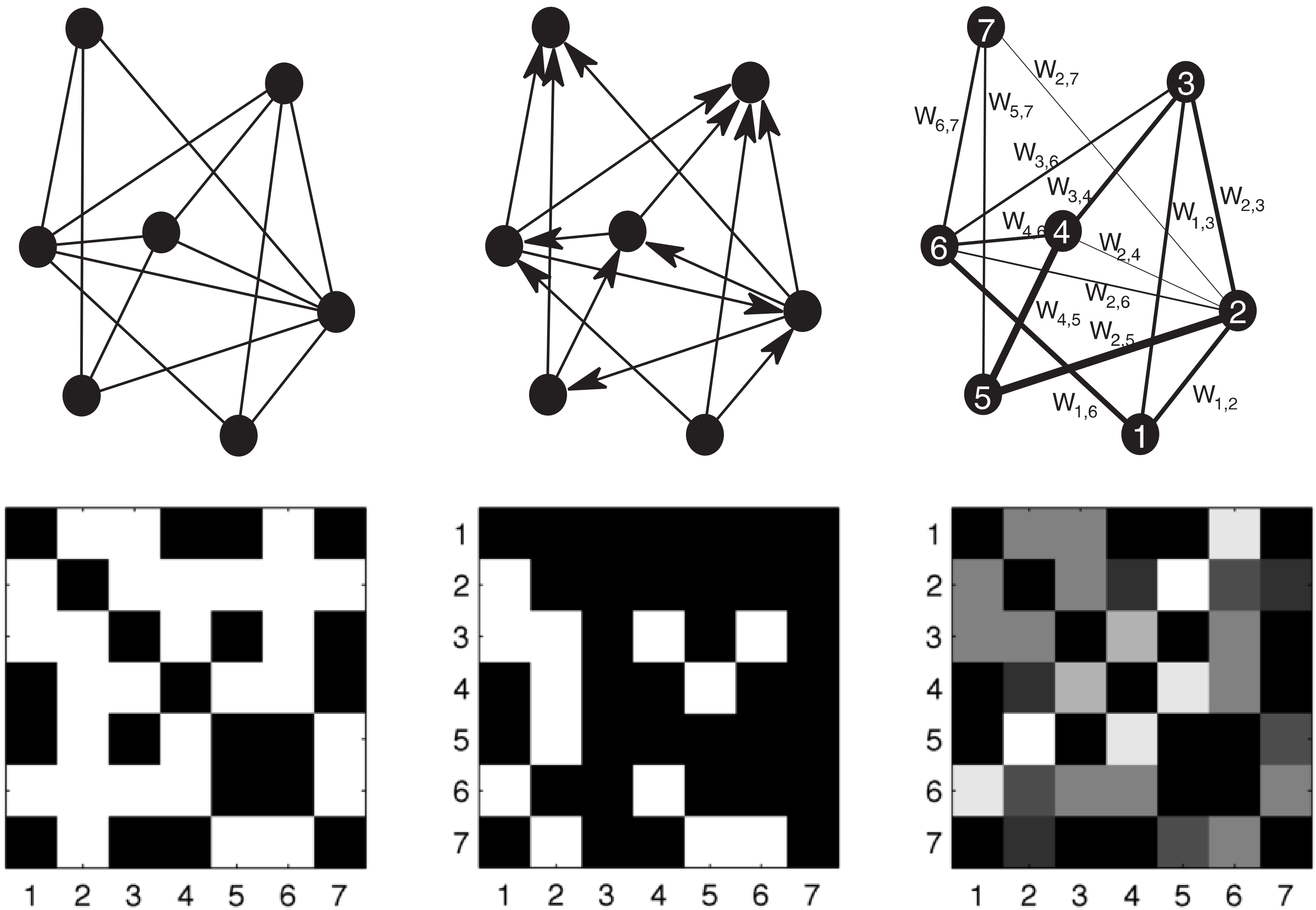


Weighted



Boccaletti et al. 2006

Matrix Form



7
3
6
4
2
5
1
W2,7
W6,7
W5,7
W3,6
W3,4
W1,3
W2,3
W4,6
W2,4
W2,6
W4,5
W2,5
W1,6
W1,2
1 2 3 4 5 6 7
1 2 3 4 5 6 7
1 2 3 4 5 6 7
1 2 3 4 5 6 7
1 2 3 4 5 6 7
1 2 3 4 5 6 7

Why do we care about graphs?

- Our neuroscientist half:
  - Assess connections among networks of neurons
- Our computational half:
  - Visualize and solve difficult computational problems
- Focus of today will be on the former

Neurons are Connected

- We know how to characterize neural responses in isolation
  - STA
  - Information Theory
  - "Noise Correlations"

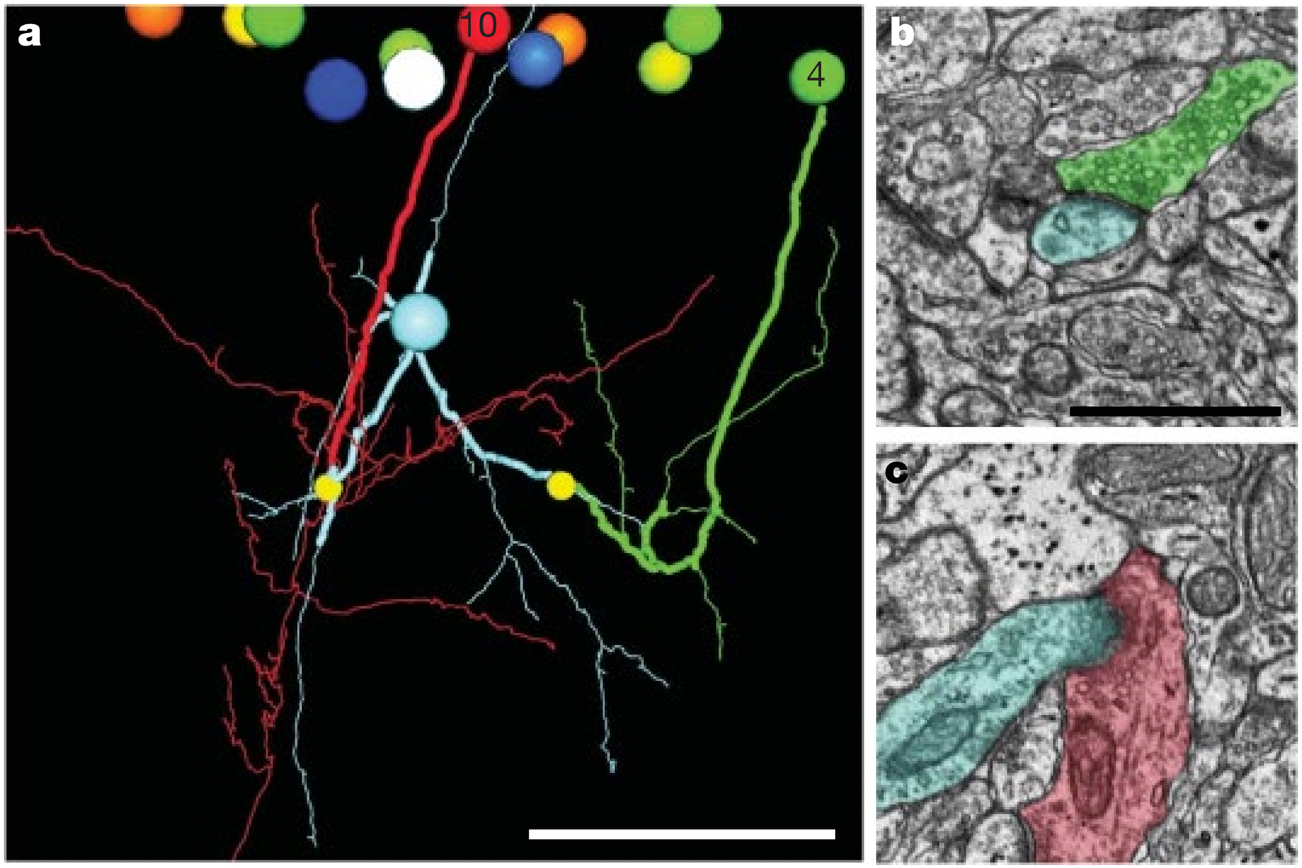
Neurons are Connected

- Neurons are not isolated
- Neurons are connected
- Neural connections are probably functionally important

Neurons are Connected

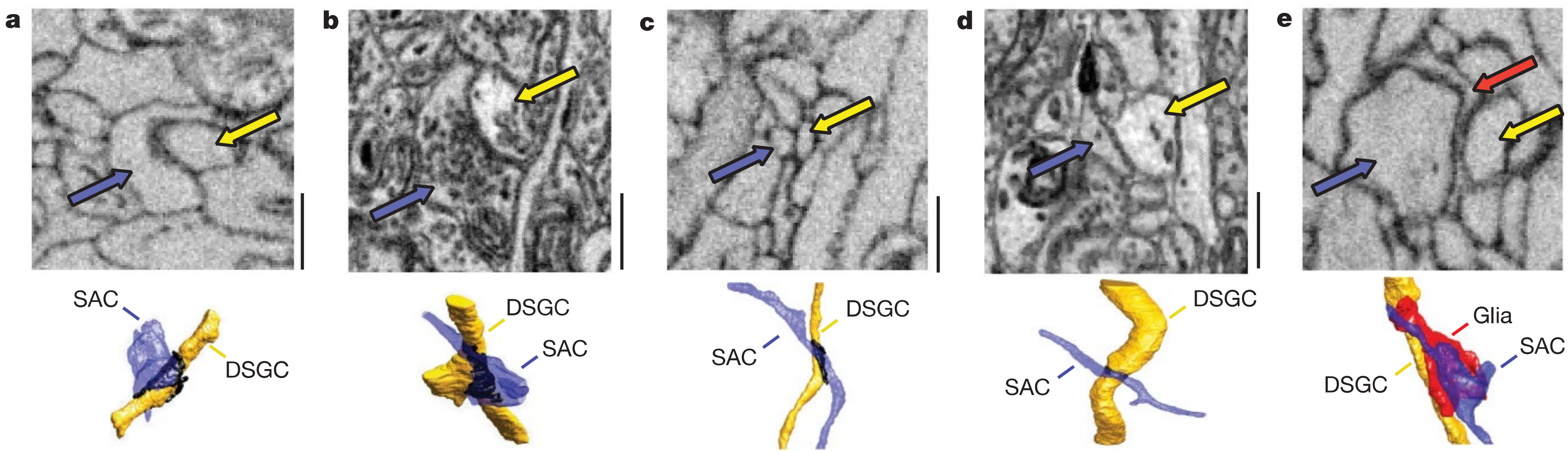
- How to measure neural connectivity?
  - Anatomical Connectivity
  - Effective Connectivity
  - Functional Connectivity

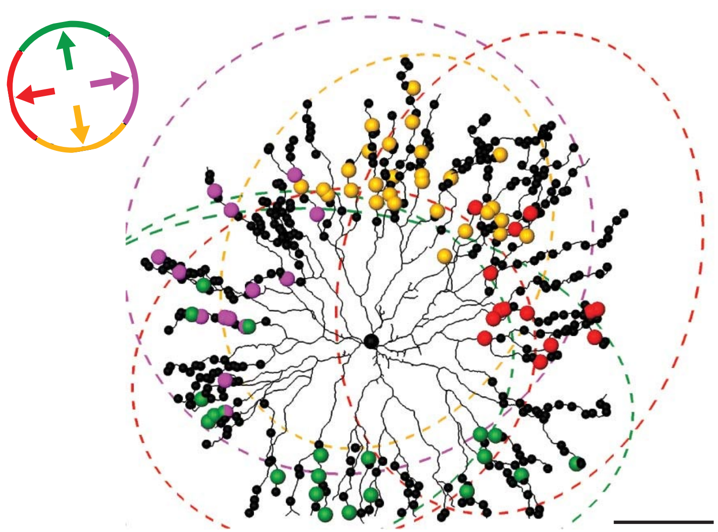
Anatomical Connectivity

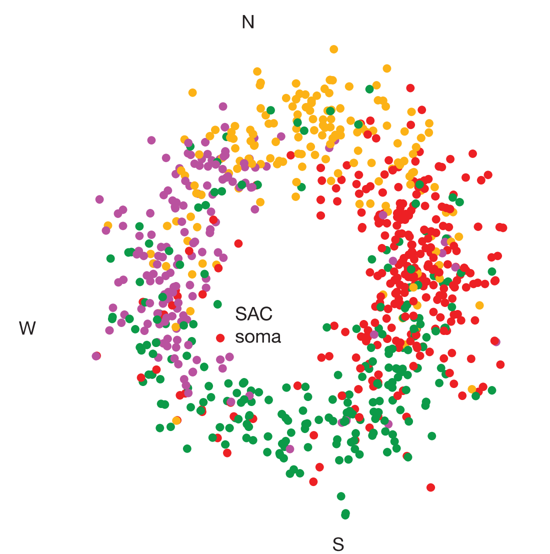


Bock et al. 2011

Anatomical Connectivity

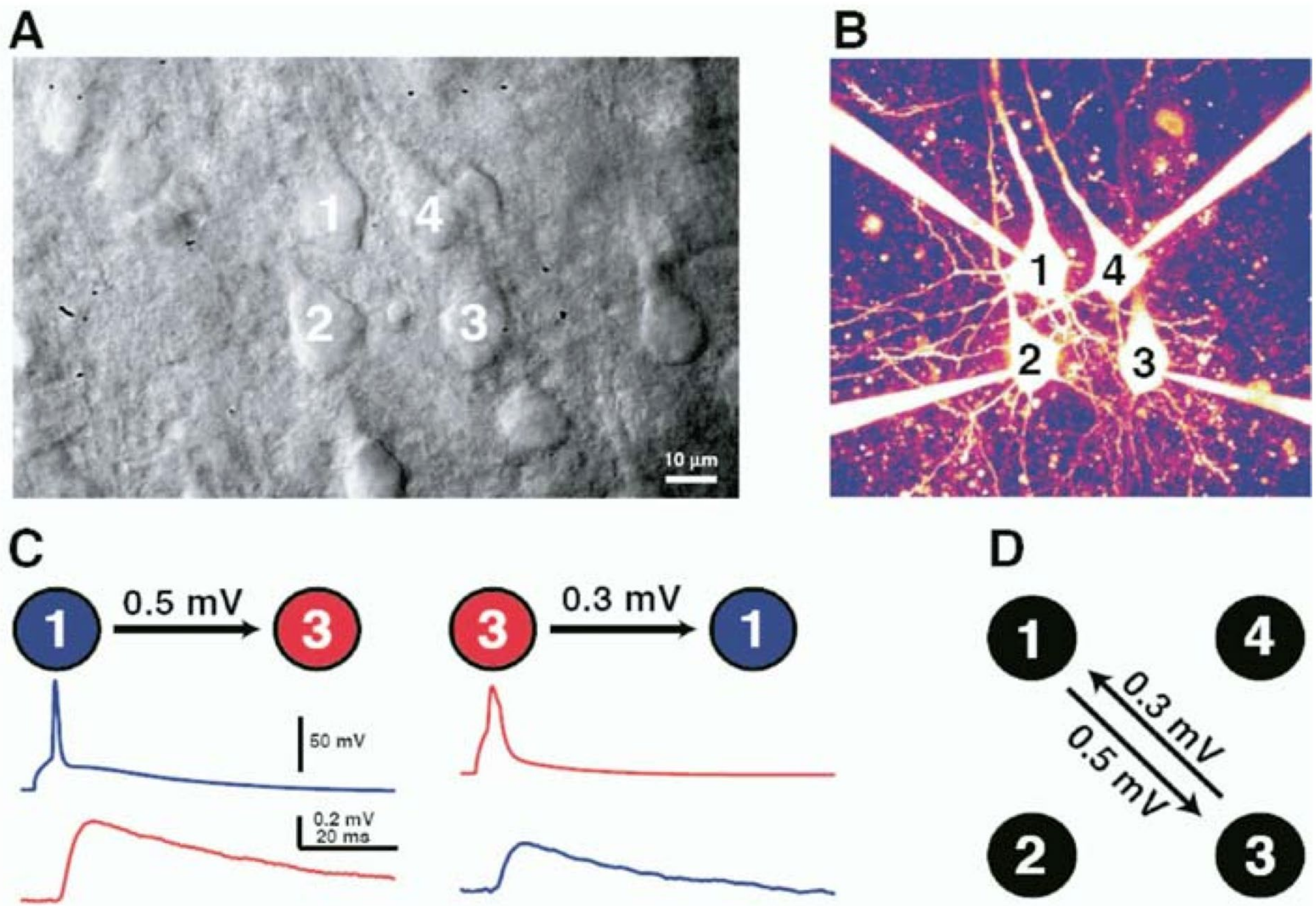






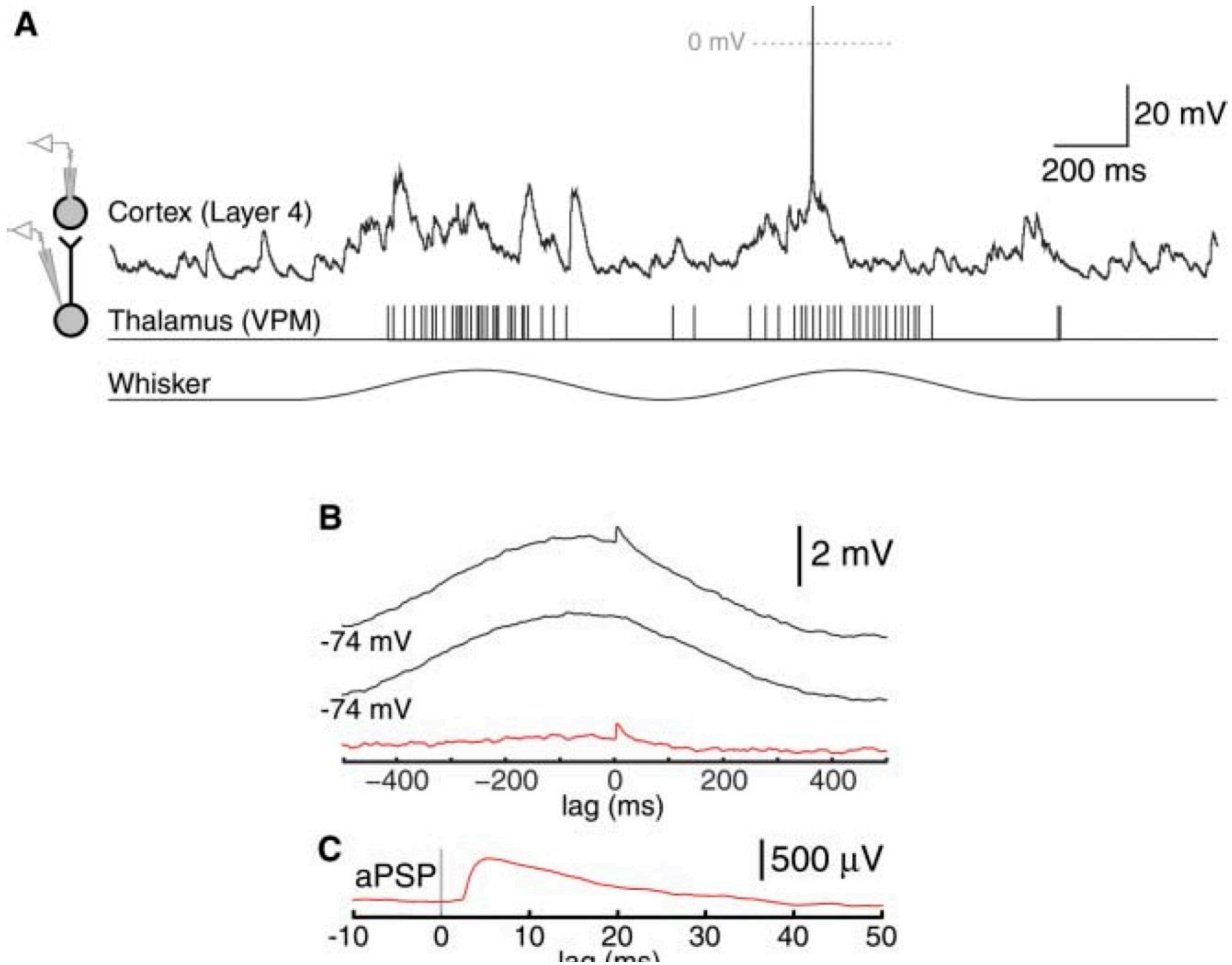
Briggman et al. 2011

Effective Connectivity



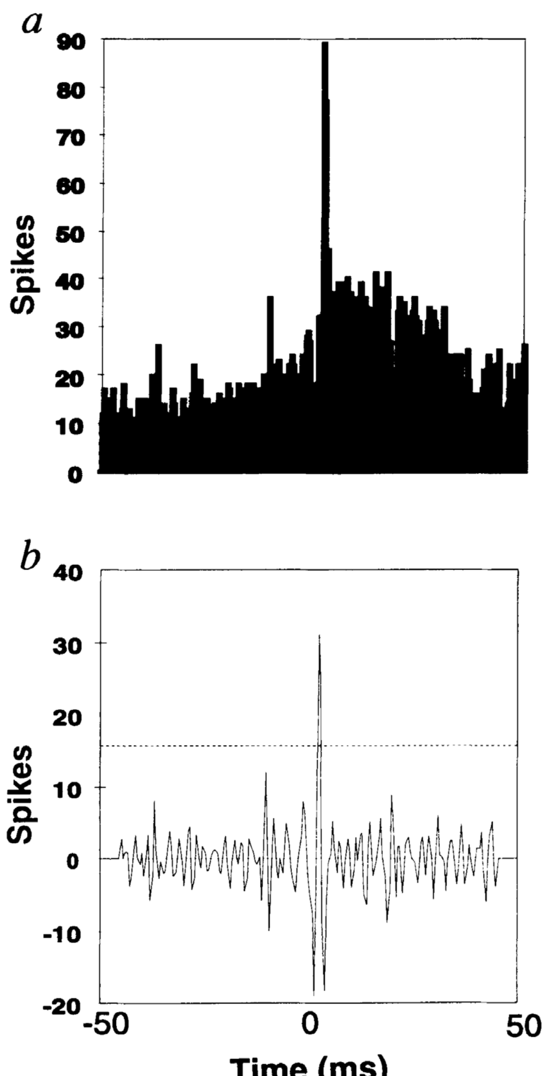
Song et al 2005

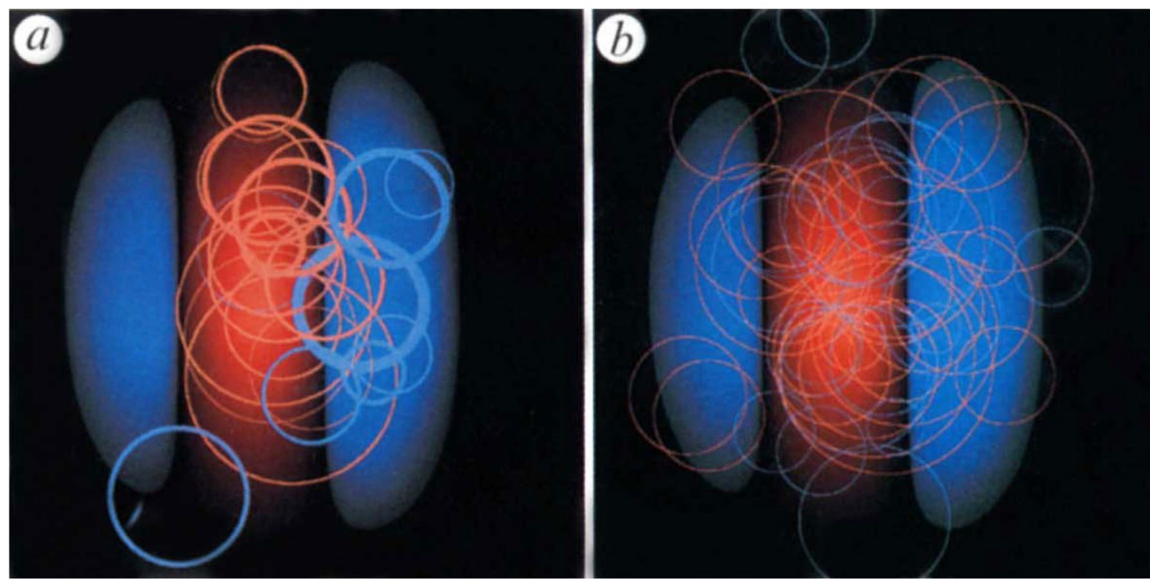
Effective/Functional Connectivity



Bruno and Sakmann 2006

Functional Connectivity



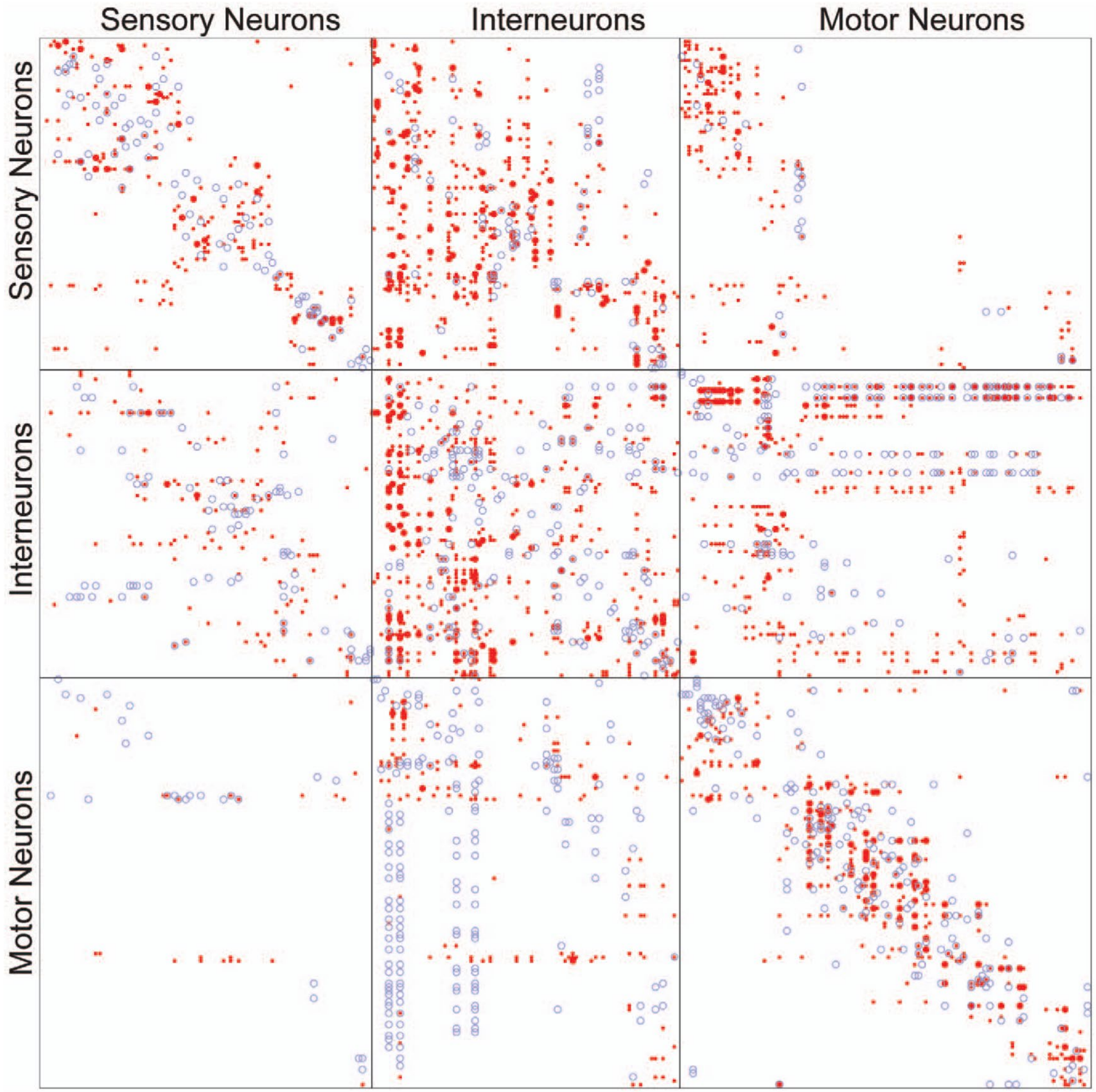


Reid and Alonso1995

Now what?

- Consider neurons as nodes and synapses as edges
- Connectivity measures dictate edge
  - Locations
  - Directedness
  - Weights

C elegans adjacency matrix



Varshney et al. 2011

Now what?

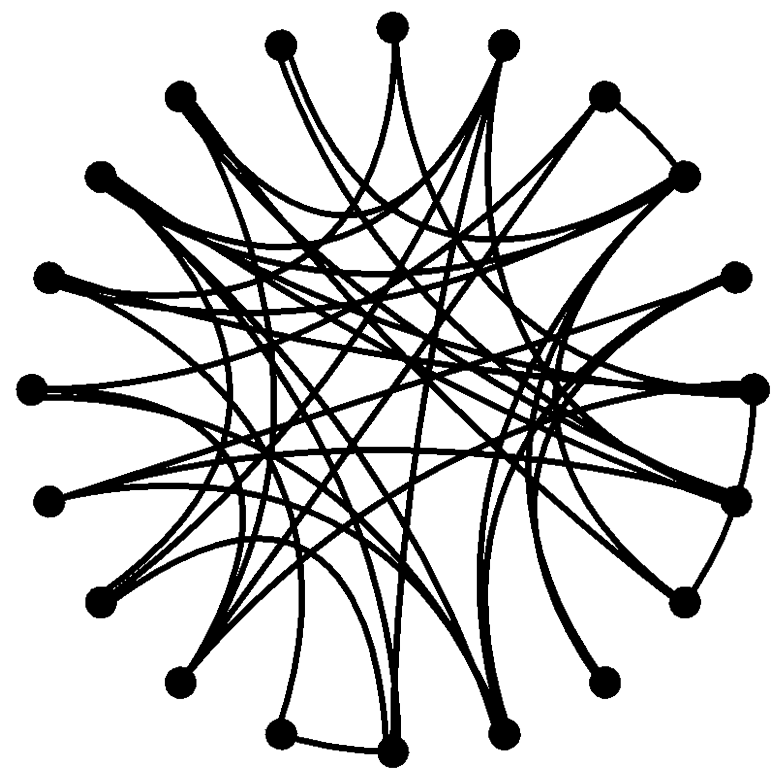
- We need priors to make interpretations of our graphs meaningful
- We need summary measures to describe big networks in the first place

Graph Priors

- Random (Erdős–Rényi) Graphs
- Regular Graphs
- Small-World Graphs

Random Graphs

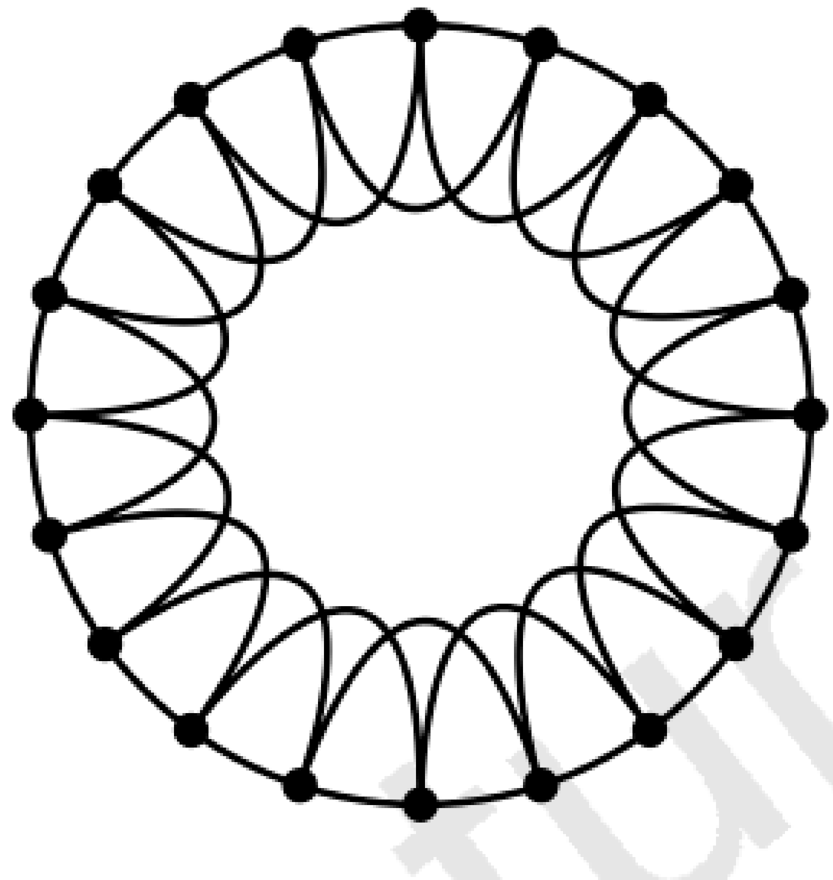
- Defined by a uniform, independent connection probability between any two nodes



Watts and Strogatz 1996

Regular Graphs

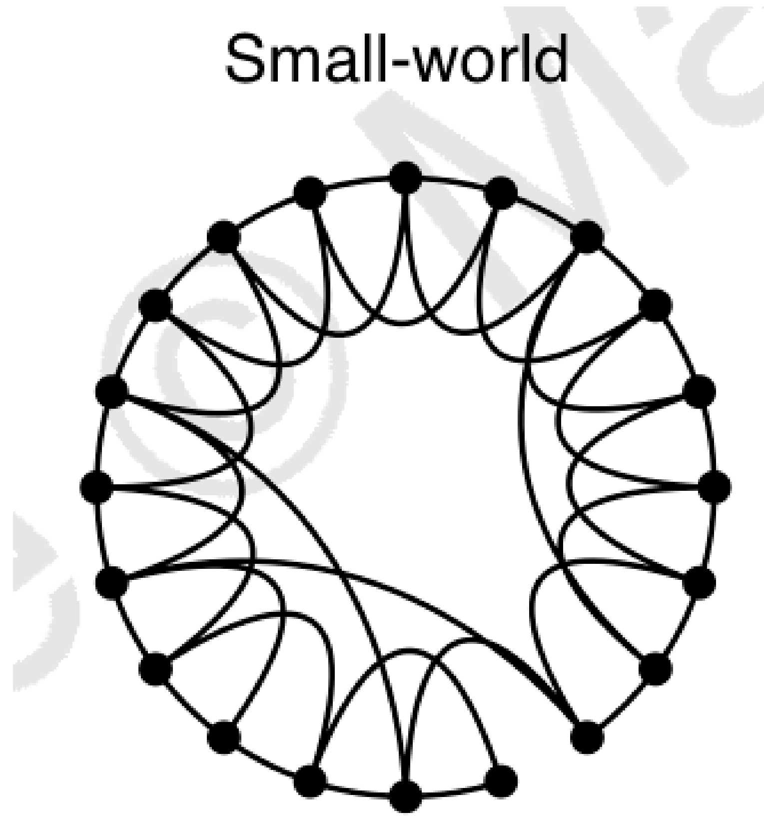
- Deterministic edge distributions
- Often determined by Euclidian distance



Watts and Strogatz 1996

Small-World Graphs

- Generate regular graph
- Randomly shuffle edge connections from a subset of nodes
- Determined by uniform shuffling probability



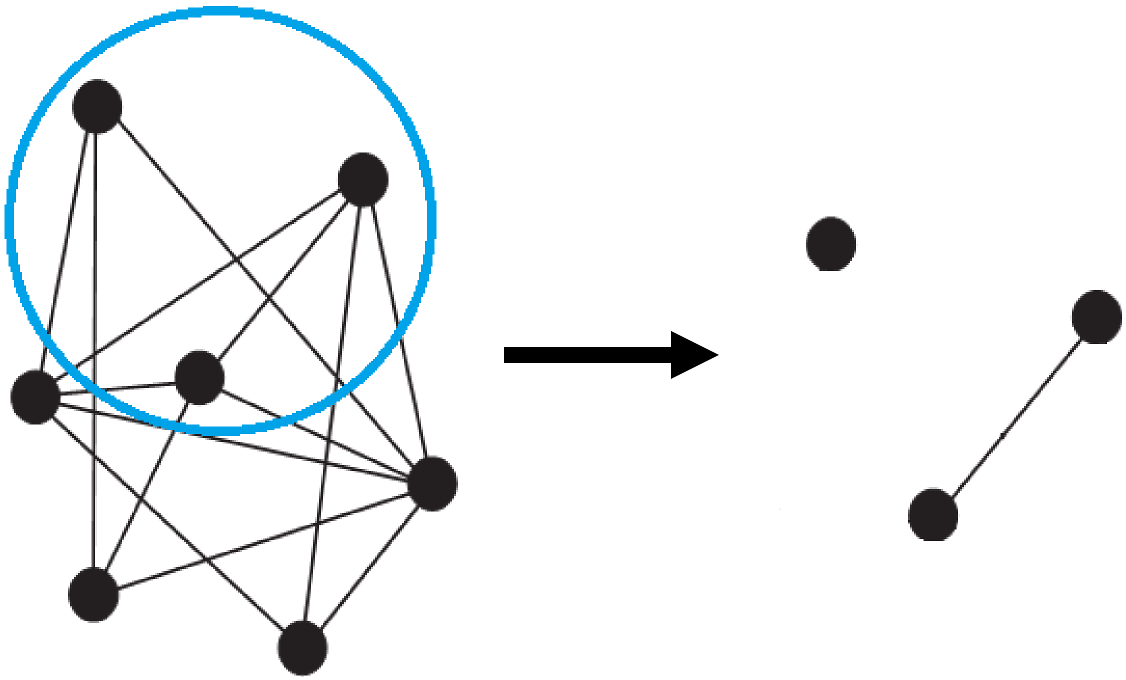
Watts and Strogatz 1996

Graph measures

- Motif frequencies
- Clustering coefficient
- Characteristic path length
- Degree distribution

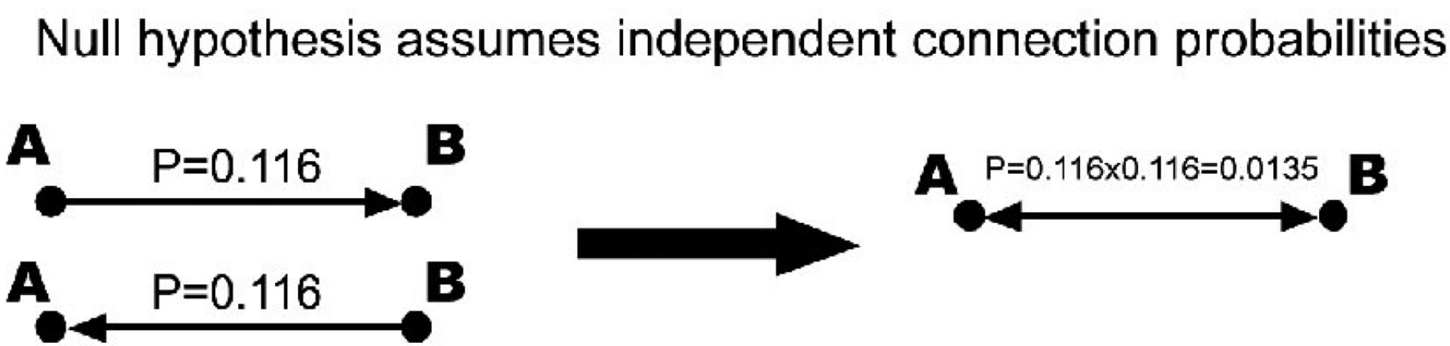
Motif Frequencies

- *Subgraph*: A subset of nodes & their connected edges lifted from a larger graph



Motif Frequencies

- Analyze the likelihood of all possible N-sized subgraphs
- Usually compared against random priors

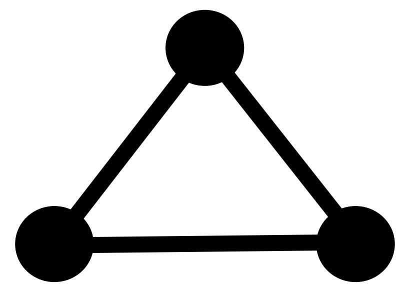


Song et al. 2005

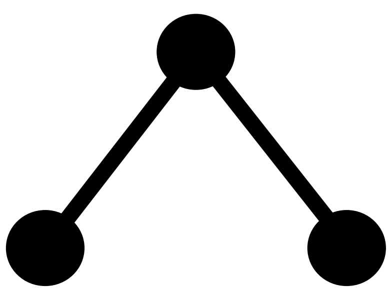
Motifs: Random Prior

For simple, undirected, unweighted random graphs with connection probability $\boldsymbol{p}$:

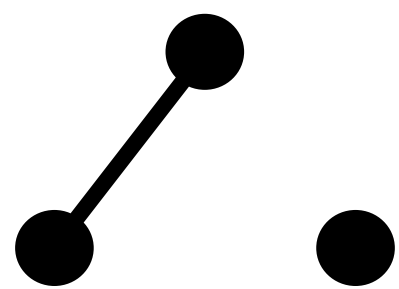
$P(edges|subgraph) = p$



$P(edges|subgraph) = p^3$

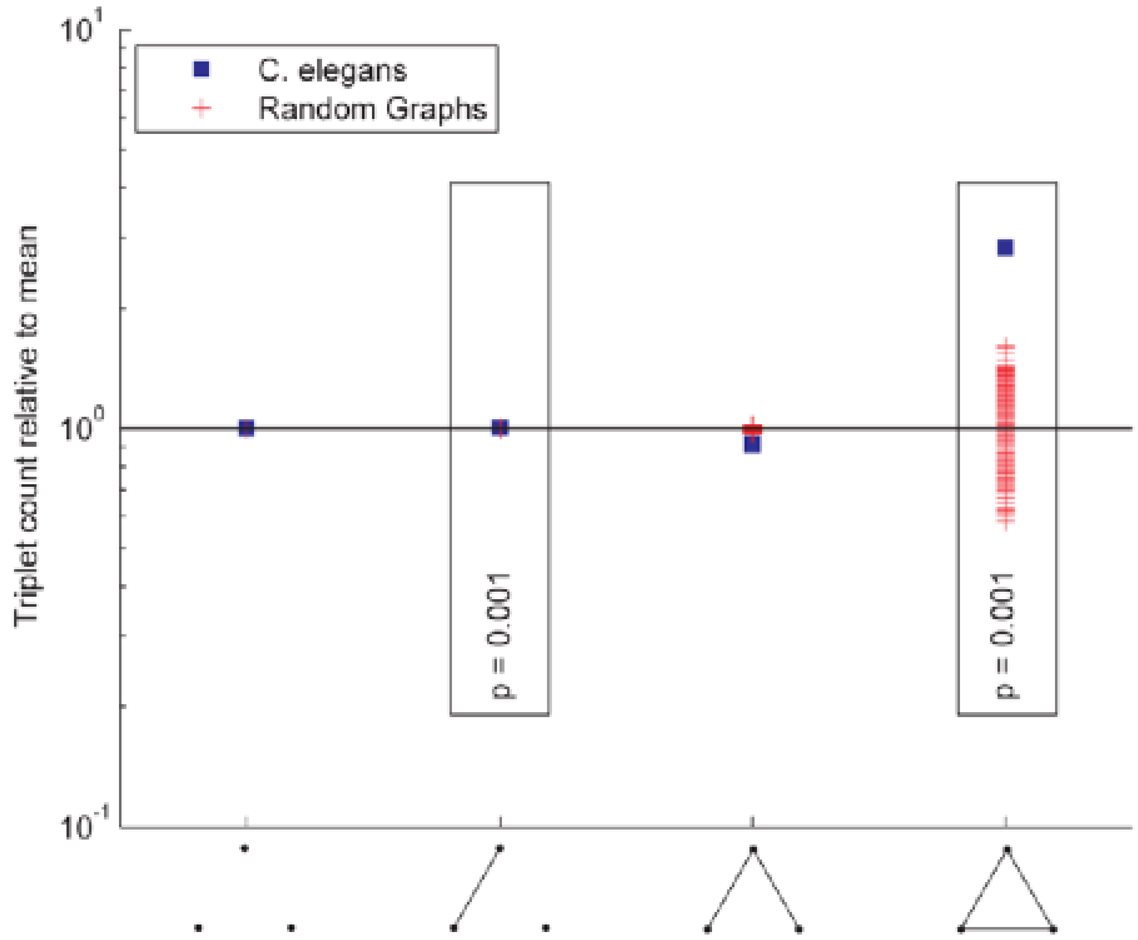


$P(edges|subgraph) = C_2^3 p^2 (1-p)$



$P(edges|subgraph) = C_1^3 p (1-p)^2$

Motifs: *C elegans* vs. Random



Varshney et al. 2011

Clustering Coefficient

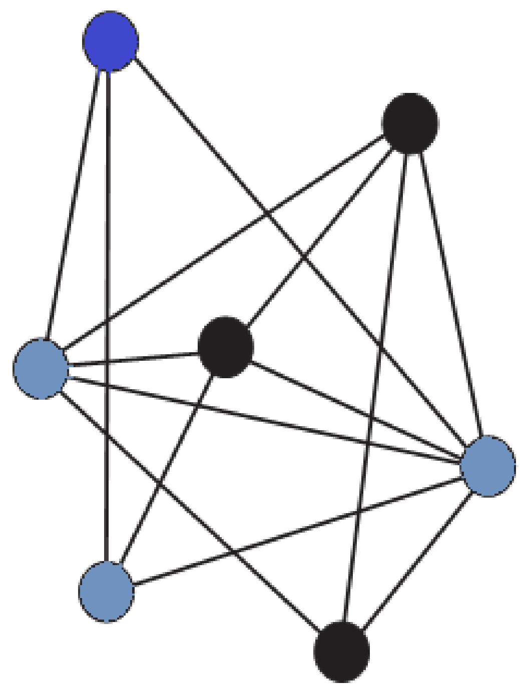
- *Complete Graph*: A graph of N nodes in which each node is connected to all other nodes in the network
- Complete subgraphs known as "cliques"
- For a simple, unweighted, undirected network:

$$E \equiv \textbf{number of edges}$$

$$E = \frac{1}{2}N(N-1)$$

Clustering Coefficient

- *Neighborhood*: For some node, the subgraph of all nodes connected to it.



Clustering Coefficient

- Two types: local and average
  - Local: Completeness (clique-ness) of the neighborhood of node i
  - For a neighborhood with n_i nodes and adjacency matrix with binary elements of the type c_{jk}:

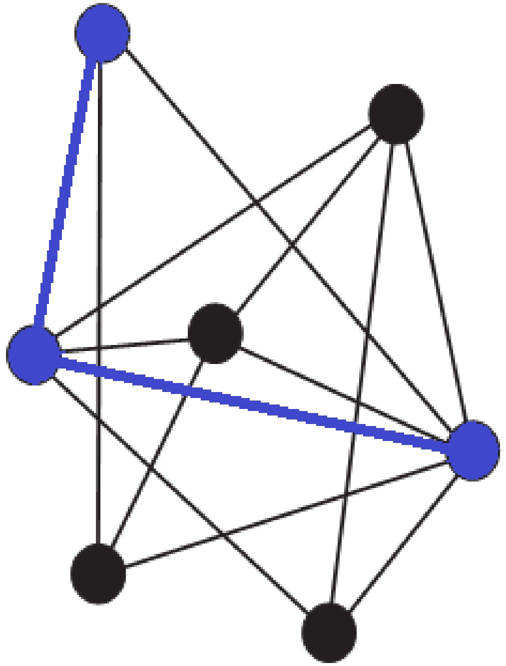
$$C_i = \frac{\sum_{j<k} c_{jk}}{\frac{1}{2} n_i (n_i - 1)}$$

  - Average: Mean clustering across all N nodes of the full graph:

$$\langle C \rangle = \frac{1}{N} \sum_{i=1}^{N} C_i$$

Characteristic Path Length

- *Path*: Alternating sequence of edges and nodes, beginning and terminating with a node
- *Path Length:* Sum of edge weights in a path



Characteristic Path Length

- *Minimum Path Length:* For a given pair of nodes, the minimum edge count among all possible paths
- Solved using Dijkstra's algorithm

Minimum Path Length Solution

```
function Li = Dijkstra(A,i)
% Takes adjacency matrix A, starting node index i
% mark non-existent edges as having weights of inf
% Li gives the minimum distance to each node
% for directed graphs, columns dictate the "from" node

n              = size(A,1);      % node count
Li             = inf(n,1);       % distance functions initialized to inf
Li(i)          = 0;              % starting point w/0 dist by definition
uv             = 1:n;            % indices of unvisited nodes

while any(uv)
    [~,ci]     = min(Li(uv));    % find terminal index of shortest path
    current    = uv(ci);         % greedily mark as "current"
    uv(ci)     = [];             % ...and, in turn, as "visited"

    Li(uv)     = min(Li(uv),Li(current) + A(uv,current));
                                 % minimum of previous distance & current
end
```

Characteristic Path Length

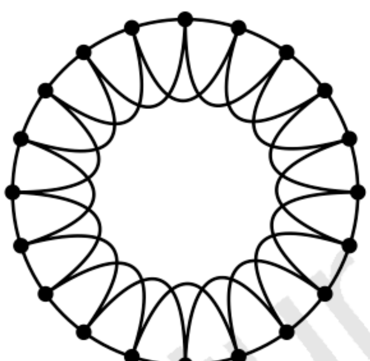
- *Characteristic Path Length:* The mean minimum path length across all pairs of (different) nodes:

```
>> n = size(A,1);
>> dists = [];
>> for i = 1:n
dtemp = Dijkstra(A,i);
dtemp(i) = [];
dists = vertcat(dists,dtemp);
end
>> mean(dists)
```

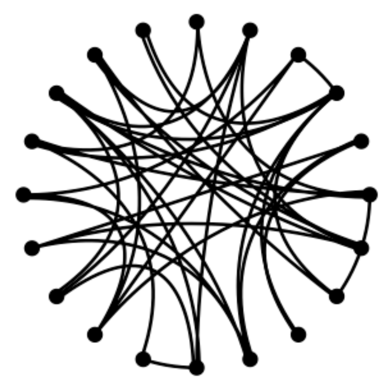
L and C of prior graphs

- **Regular:** L ~ N C ~ 1
- **Random:** L ~ log N C ~ 1/N

Regular



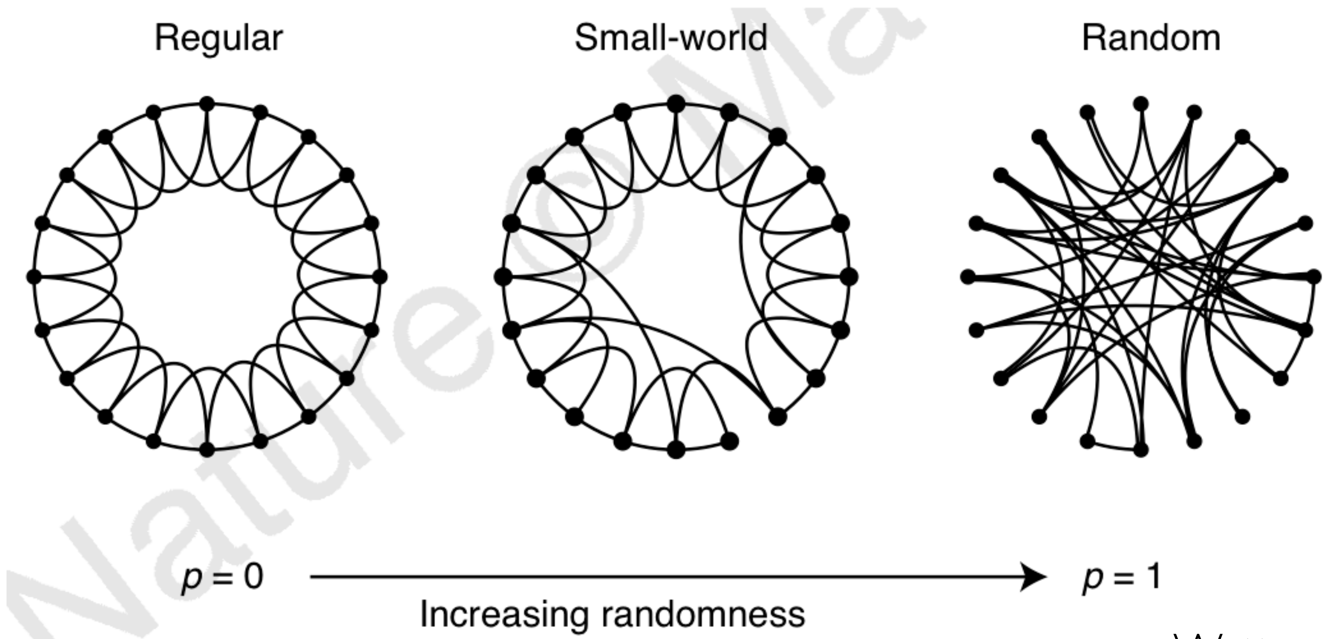
Random



Watts and Strogatz 1996

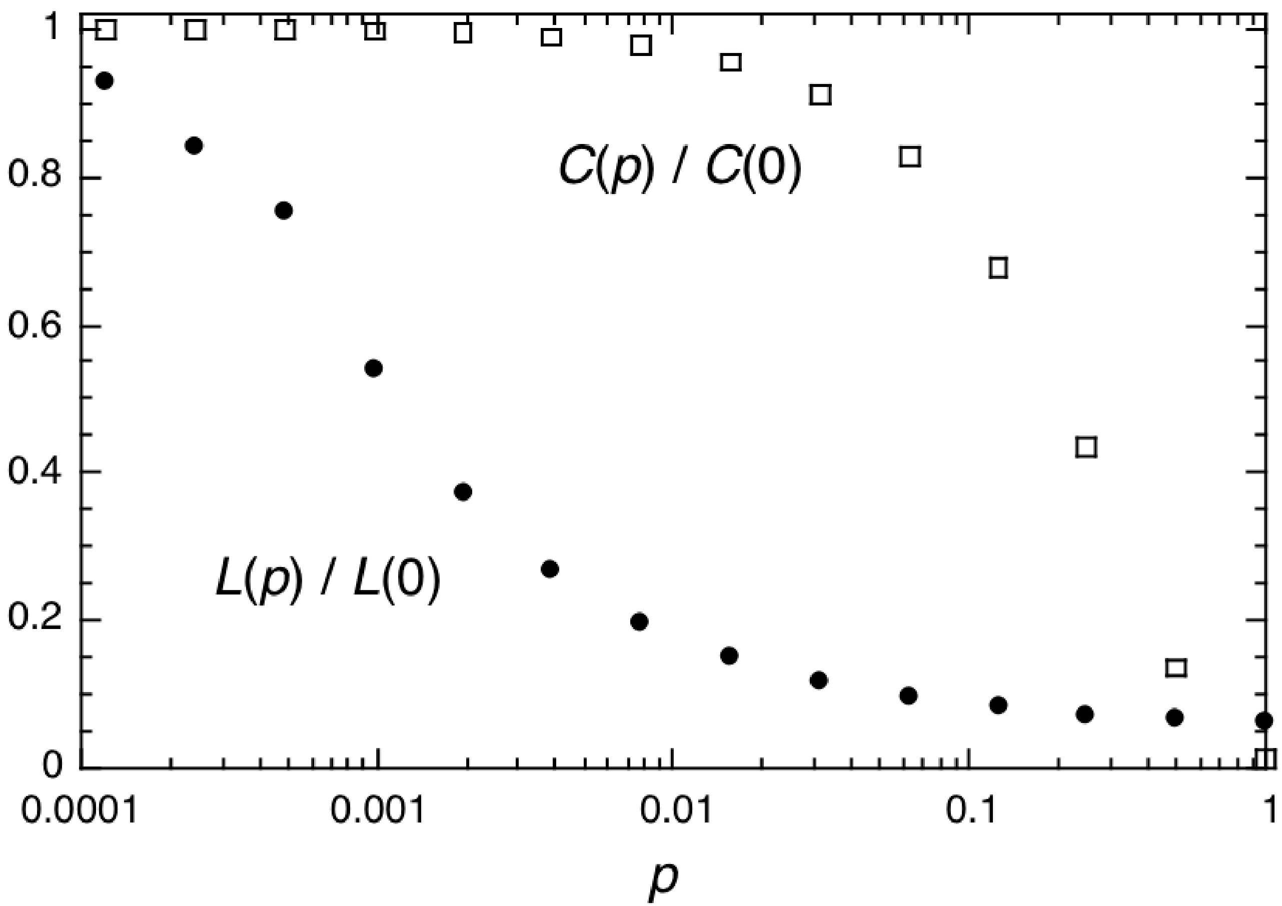
L and C of prior graphs

- **Regular:** L ~ N C ~ 1
- **Random:** L ~ log N C ~ 1/N
- **Small World:** L ~ log N C ~ 1



Watts and Strogatz 1996

"Sweet Spot"



Real Networks in the "Sweet Spot"

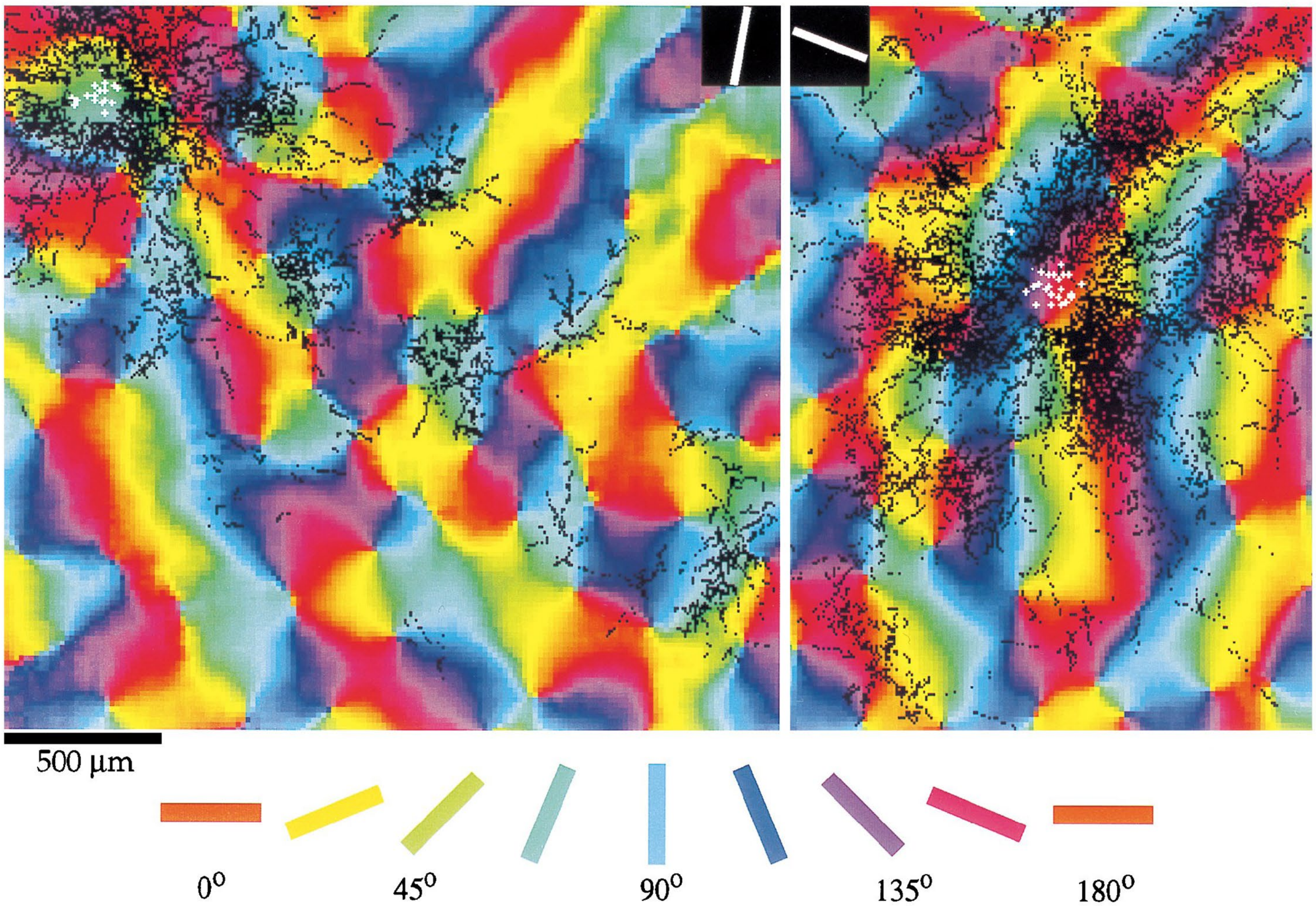
Table 1 Empirical examples of small-world networks

	L_{actual}	L_{random}	C_{actual}	C_{random}
Film actors	3.65	2.99	0.79	0.00027
Power grid	18.7	12.4	0.080	0.005
C. elegans	2.65	2.25	0.28	0.05

N=279

Watts and Strogatz 1996

Tree Shrew Small World

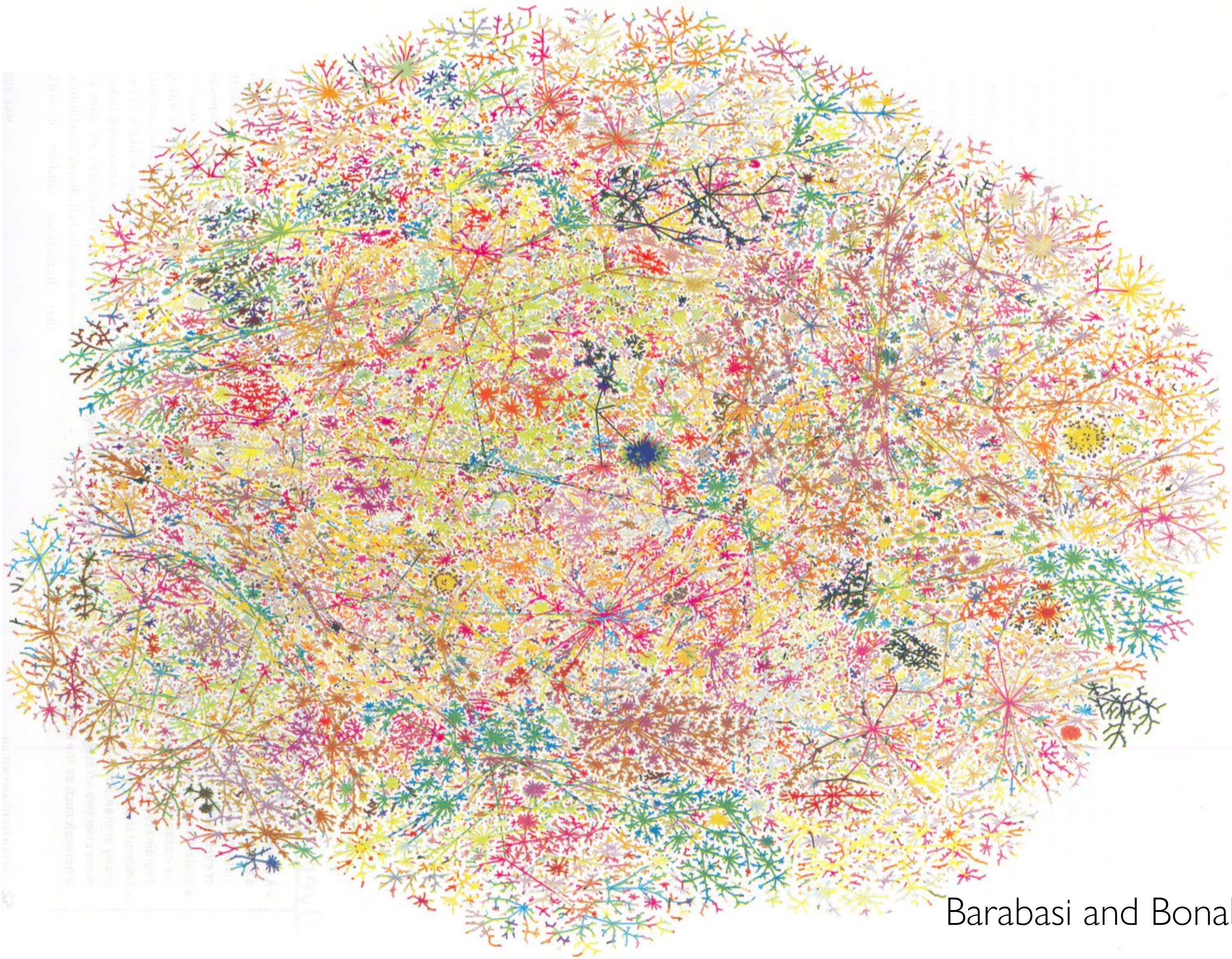


Bosking et al. 1997

Degree Distributions

- The **degree** of a node k_i is simply the number of edges connected to that node
- The **degree distribution** $P(k)$ is the probability across a network of a node having degree k.

Hubs: High-Degree Nodes

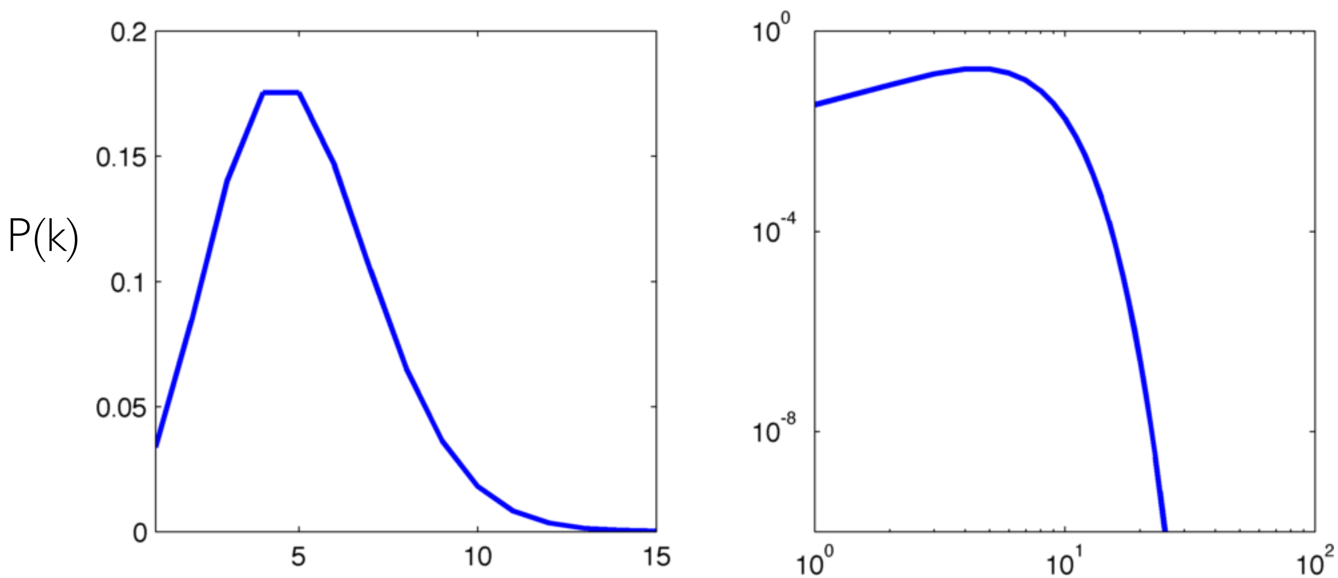


Barabasi and Bonabeau 2003

Random Networks

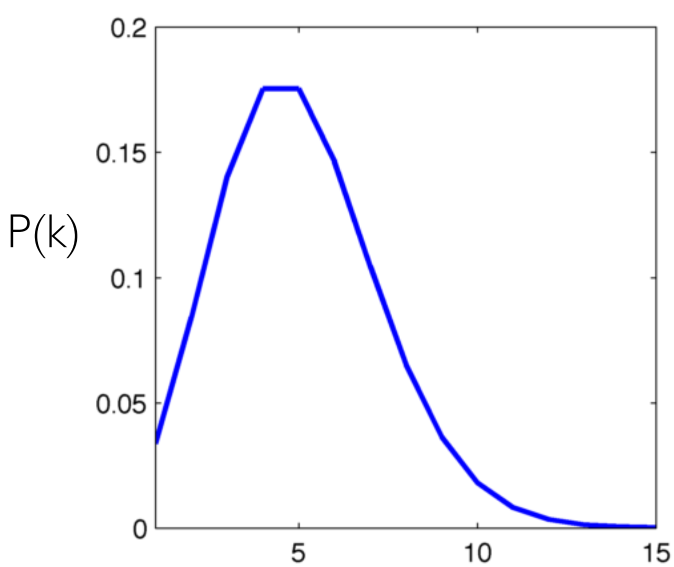
- The degree distribution of a random network is:

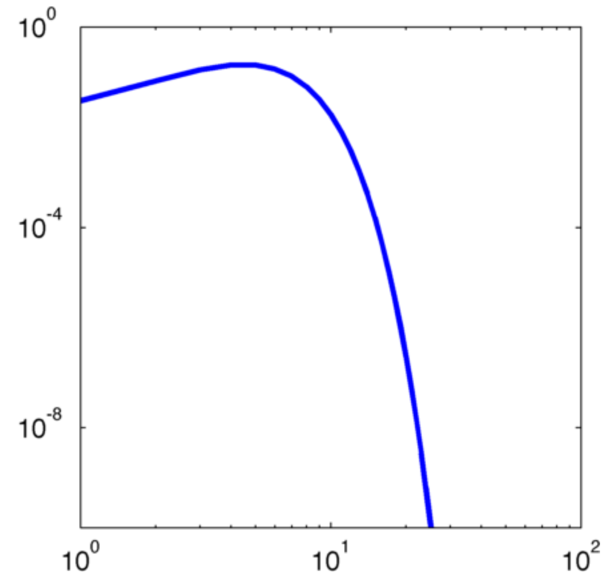
$$P(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k}$$



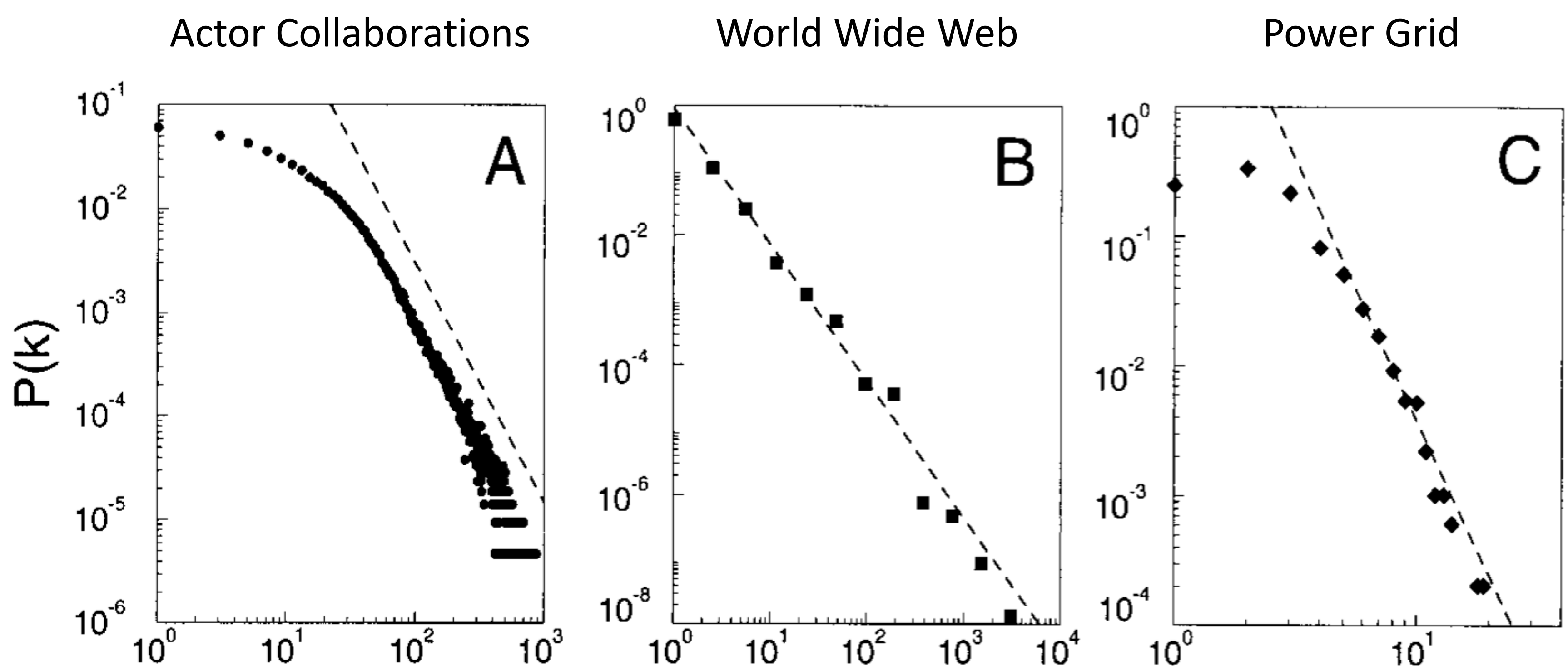
Random Networks

- For random networks:
  - $P(k) \propto a^{-k}$





Real Networks



- Empirically, in many real networks
  - $P(k) \propto k^{-\gamma}$

Barabasi and Albert 1999

Scale Invariance

- A constant scaling of the inputs leads to a constant scaling of the outputs

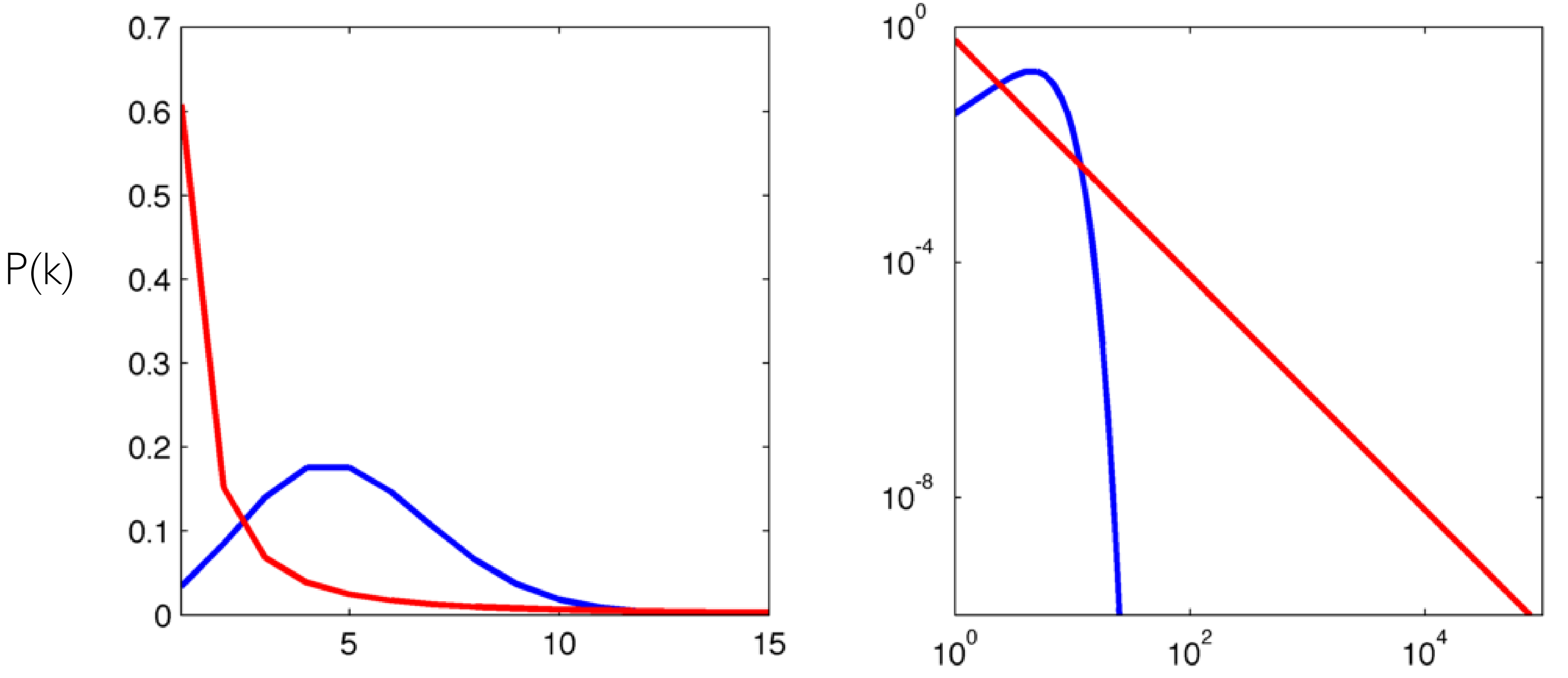
$$f(cx) = c^{\gamma} f(x)$$

$$f(x) = ax^{\gamma}$$

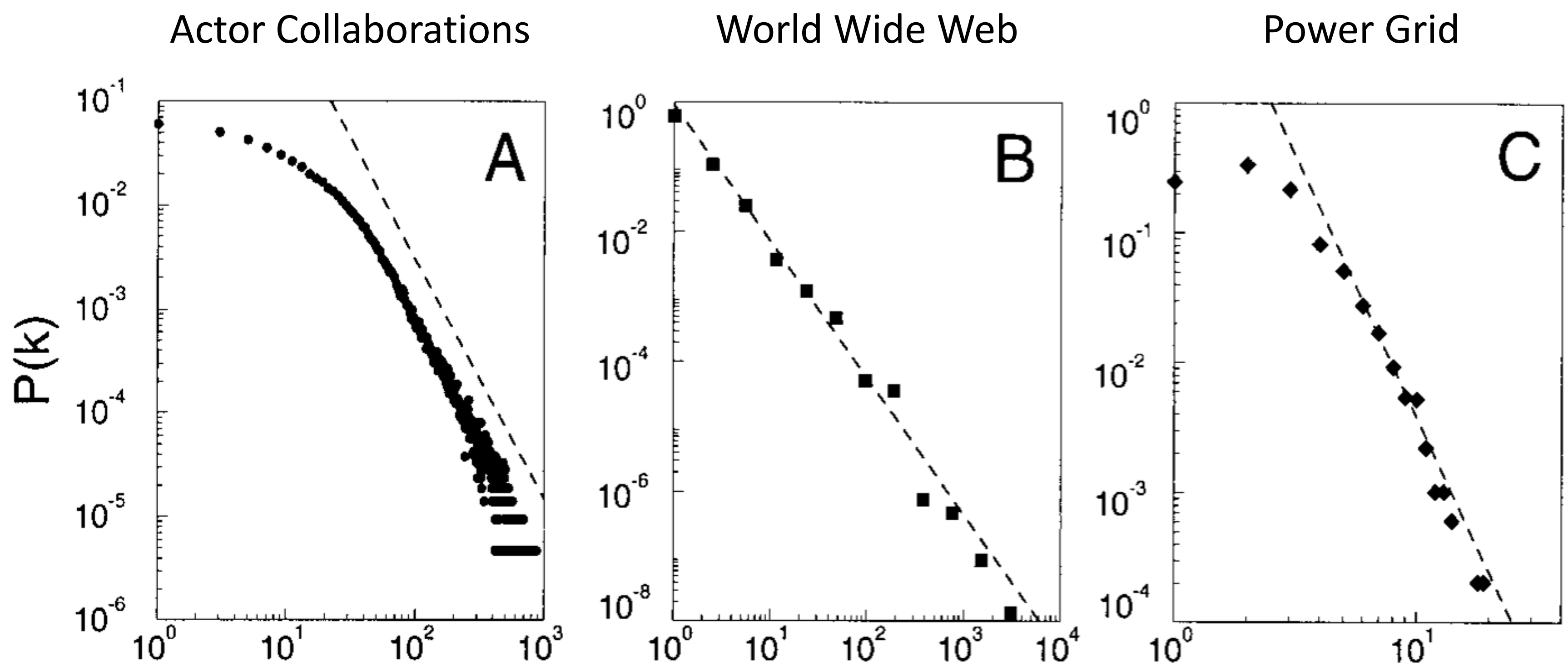
- Power laws lead to scale invariance
- A power law degree distribution defines a *'scale free' network*

Scale Free Networks

- Scale free networks have a "heavy tail"
- Thus, scale free networks have hubs



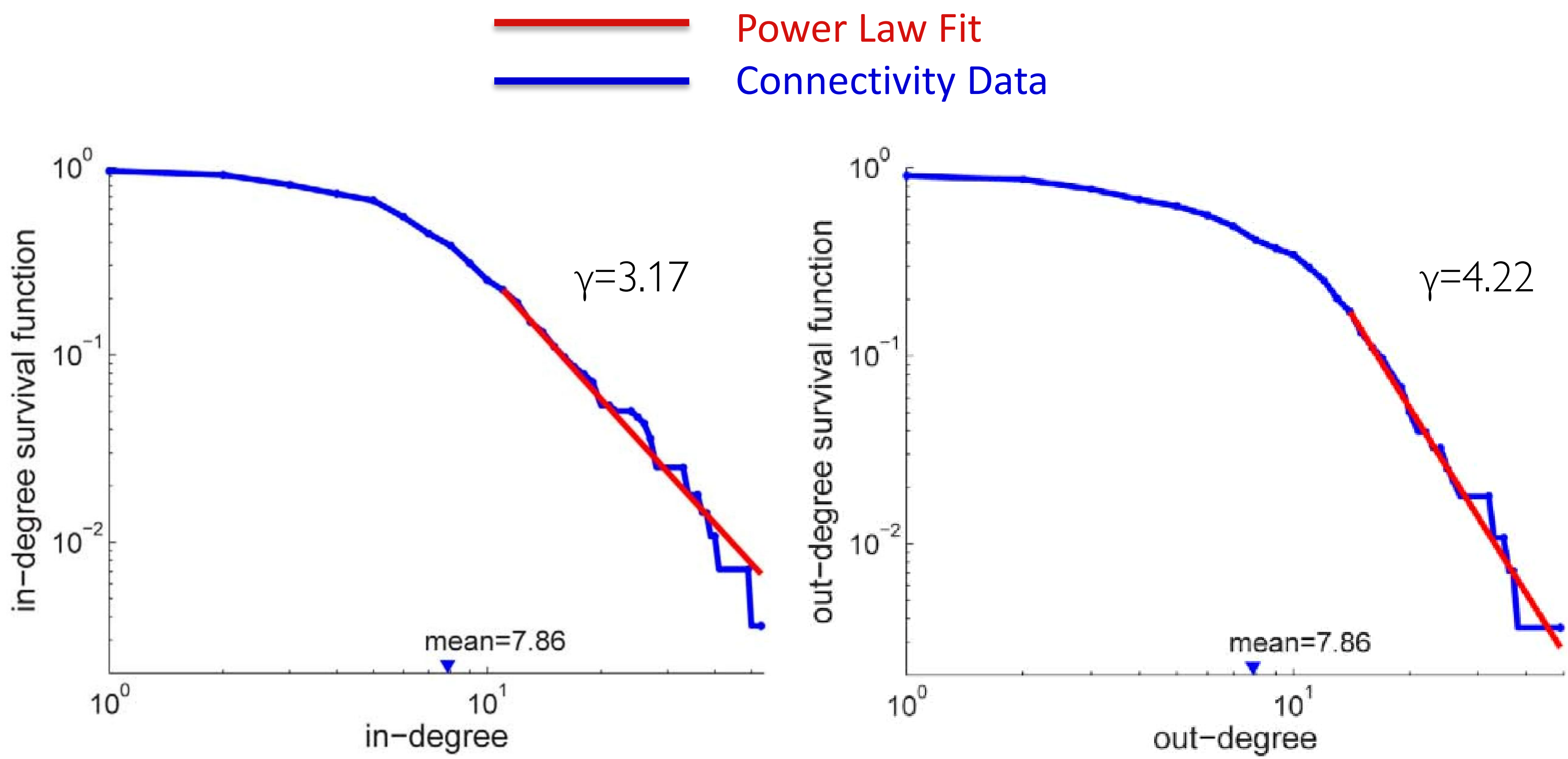
Real Networks: *Low k behavior*



- Power law is only good for asymptotic k
- Low k show binomial behavior

Barabasi and Albert 1999

C. Elegans



Varshney et al. 2011

Recap

- Graphs useful for two things from our perspective
  - Quantifying network connectivity
  - Formulating problems in easily-analyzable format
- Neural networks are
  - Clustered and connected
  - Have highly likely hubs
  - Best approximated by small-world priors: a mix of random and regular