Frontiers of Network Science
Fall 2023

Class 11: Evolving Networks II Degree Correlations
(Chapters 5 & 7 in Textbook)

Boleslaw Szymanski

based on slides by
Albert-László Barabási
and Roberta Sinatra

Evolving network models

EVOLVING NETWORK MODELS

The BA model is only a minimal model.

Makes the simplest assumptions:

- linear growth
- linear preferential attachment

$$\langle k \rangle = 2m$$

$$\Pi(k_i) \propto k_i$$

Does not capture

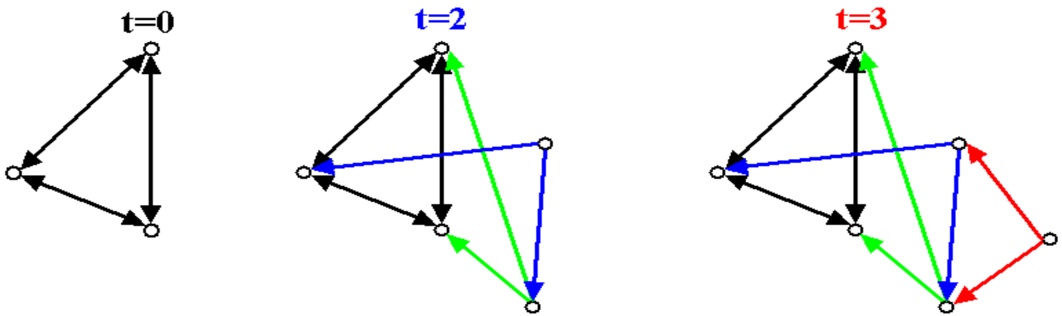
variations in the shape of the degree distribution
variations in the degree exponent
the size-independent clustering coefficient

Hypothesis:
The BA model can be adapted to describe most features of real networks.

We need to incorporate mechanisms that are known to take place in real networks: addition of links without new nodes, link rewiring, link removal; node removal, constraints or optimization

BA ALGORITHM WITH DIRECTED EDGES

(the simplest way to change the degree exponent)



$$\frac{\partial k_i}{\partial t} \propto \Pi(k_i) = A\frac{k_i}{\sum_j k_j} = \boxed{\frac{k_i}{t}}$$

Undirected BA network: $\sum_j k_j = 2t$

Directed BA network: $\sum_j k_j = t$

$$k_i(t) = m\frac{t}{t_i}$$

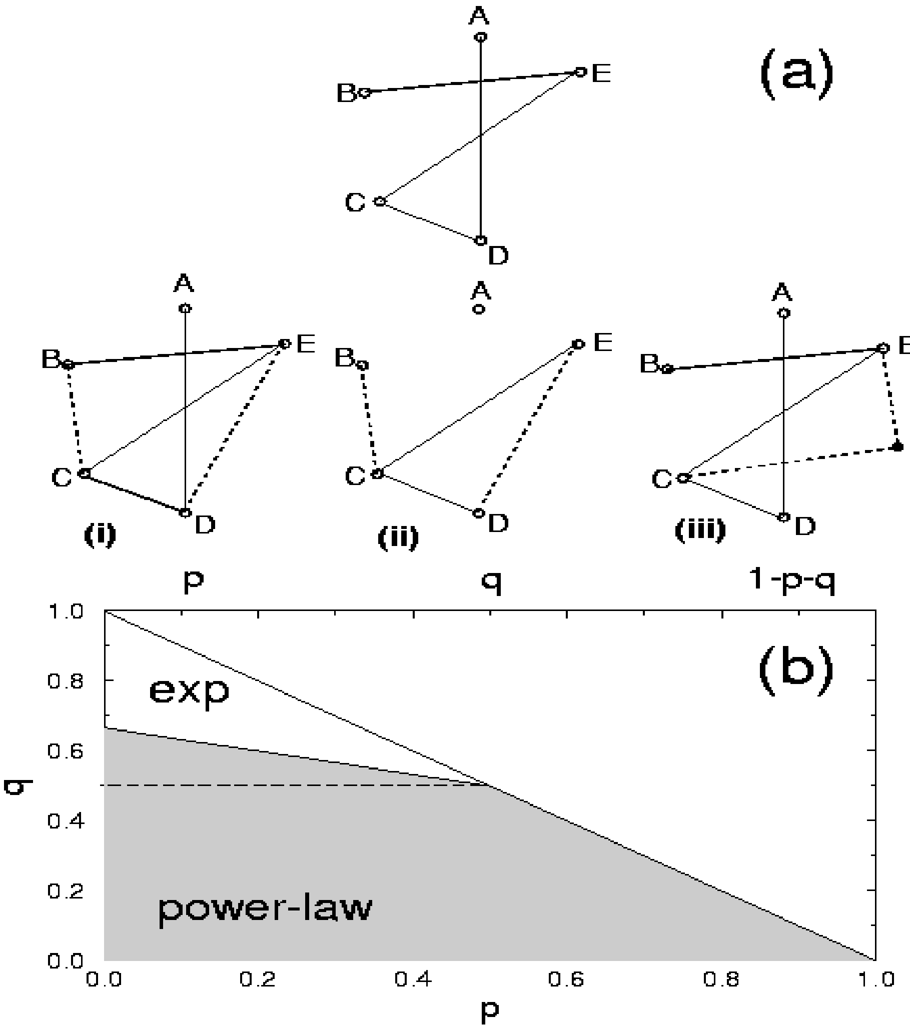
$$P_{in}(k) \sim k^{-2}$$

β=1: dynamical exponent

γ_{in}=2: degree exponent; $P(k_{out})=\delta(k_{out}-m)$

Undirected BA: β=1/2; γ=3

EXTENDED MODEL: Other ways to change the exponent



Extended Model

- prob. p : internal links
- prob. q : link deletion
- prob. 1-p-q : add node

$$P(k) \sim (k+\kappa(p,q,m))^{-\gamma(p,q,m)}$$

$$\gamma \in [1,\infty)$$

EXTENDED MODEL: Small-k cutoff

$$P(k) \sim (k+\kappa(p,q,m))^{-\gamma(p,q,m)} \qquad \gamma \in [1,\infty)$$

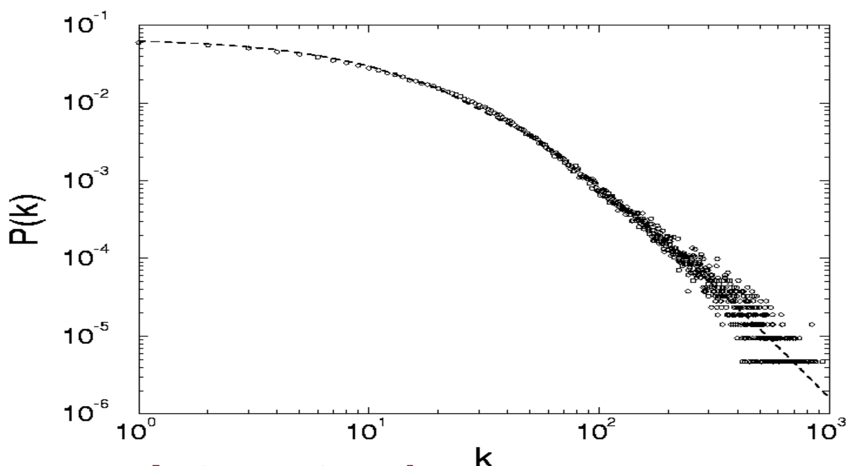
Extended Model

- prob. p : internal links
- prob. q : link deletion
- prob. 1-p-q : add node

→**Predicts a small-k cutoff**

→a correct model should predict all aspects of the degree distribution, not only the degree exponent.

→Degree exponent is a continuous function of *p*,*q*, *m*



Actor network

p=0.937

m=1

κ = 31.68

γ = 3.07

NONLINEAR PREFERENTIAL ATTACHMENT: MORE MODELS

- Non-linear preferential attachment:

$$\Pi(k) = \frac{k^{\alpha}}{\sum_i k_i^{\alpha}}$$

$\rightarrow$ P(k) does not follow a power law for $\alpha \neq 1$

$\Rightarrow$ $\alpha<1$: stretch-exponential $P(k) \approx exp\left(-(k/k_0)^{\beta}\right)$

$\Rightarrow$ $\alpha>1$: no-scaling ($\alpha>2$: “gelation”)

P. Krapivsky, S. Redner, F. Leyvraz, Phys. Rev. Lett. 85, 4629 (2000)

INITIAL ATTRACTIVENESS

BA model: *k=0* nodes cannot aquire links, as *Π(k=0)=0*
(the probability that a new node will attach to it is zero)

$$\Pi(k) \approx A + k^{\alpha},\ \alpha \leq 1$$

A - initial attractiveness

Initial attractiveness shifts the degree exponent:

$$\gamma_{in} = 2 + \frac{A}{m}$$

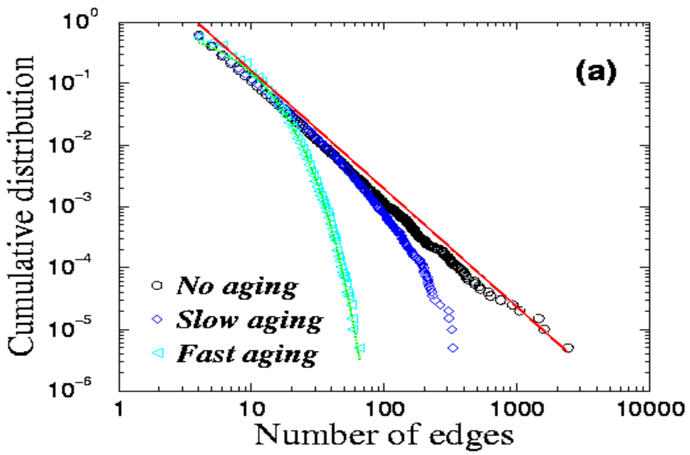
Note: the parameter A can be measured from real data, being the rate at which k=0 nodes acquire links, i.e. Π(k=0)=A

Dorogovtsev, Mendes, Samukhin, Phys. Rev. Lett. 85, 4633 (2000)

GROWTH CONSTRAINTS AND AGING CAUSE CUTOFFS

- Finite lifetime to acquire new edges

L. A. N. Amaral et al., PNAS 97, 11149 (2000)

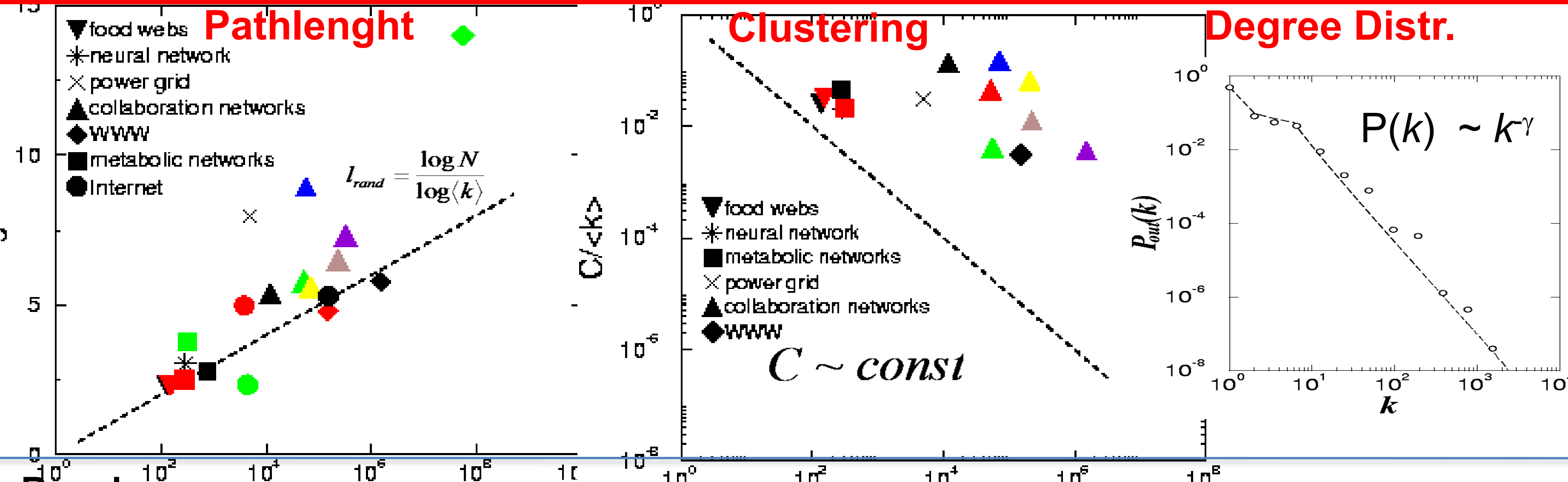


- Gradual aging: $$\Pi(k_i) \propto k_i (t - t_i)^{-\nu}$$

γ increases with ν

S. N. Dorogovtsev and J. F. F. Mendes, Phys. Rev. E 62, 1842 (2000)

THE LAST PROBLEM: HIGH, SYSTEM-SIZE INDEPENDENT C(N)

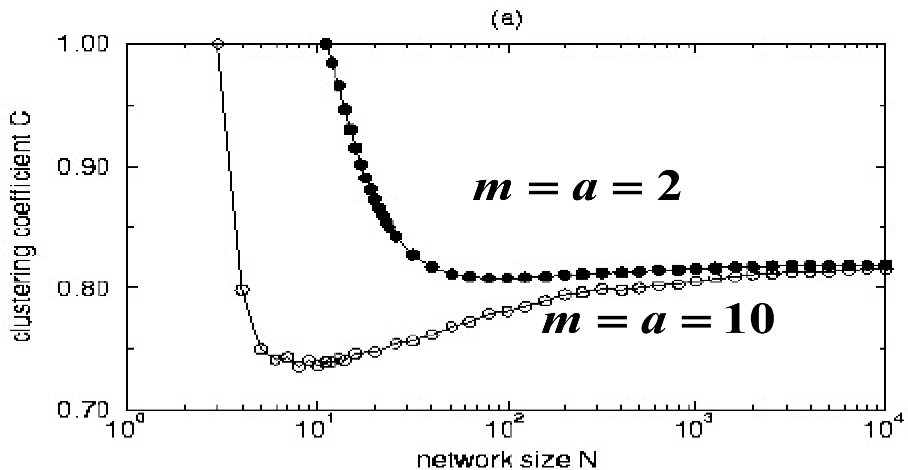


Regular network	$l \approx N^{1/D}$		$C \sim const$		$P(k)=\delta(k-k_d)$	
Erdos-Renyi	$l_{rand} \approx \frac{\log N}{\log\langle k\rangle}$		$C_{rand} = p = \frac{\langle k\rangle}{N}$		$P(k) = e^{-\langle k\rangle}\frac{\langle k\rangle^k}{k!}$	
Watts-Strogatz	$l_{rand} \approx \frac{\log N}{\log\langle k\rangle}$		$C \sim const$		*Exponential*	
Barabasi-Albert	$l \approx \frac{\ln N}{\ln \ln N}$		$C \sim \frac{(\ln N)^2}{N}$		P(k) ~ $k^{-\gamma}$	

A MODEL WITH HIGH CLUSTERING COEFFICIENT

- Each node of the network can be either active or inactive.
- There are *m* active nodes in the network in any moment.

1. Start with *m* active, completely connected nodes.
2. Each timestep add a new node (active) that connects to *m* active nodes.
3. Deactivate one active node with probability: $P_d(k_i) \propto (a + k_j)^{-1}$



$$\Pi(k) \approx a + k$$

$$P(k) \approx k^{-2-a/m}$$

C→ C* when N→∞

K. Klemm and V. Eguiluz, Phys. Rev. E 65, 036123 (2002)

Linear growth, linear pref. attachment	$\gamma=3$	Barabási and Albert, 1999
Nonlinear preferential attachment $\Pi(k_i)\sim k_i^{\alpha}$	no scaling for $\alpha\neq 1$	Krapivsky, Redner, and Leyvraz, 2000
Asymptotically linear pref. attachment $\Pi(k_i)\sim a_\infty k_i$ as $k_i\rightarrow\infty$	$\gamma\rightarrow 2$ if $a_\infty\rightarrow\infty$ $\gamma\rightarrow\infty$ if $a_\infty\rightarrow 0$	Krapivsky, Redner, and Leyvraz, 2000
Initial attractiveness $\Pi(k_i)\sim A+k_i$	$\gamma=2$ if $A=0$ $\gamma\rightarrow\infty$ if $A\rightarrow\infty$	Dorogovtsev, Mendes, and Samukhin, 2000a, 2000b
Accelerating growth $\langle k\rangle\sim t^{\theta}$ constant initial attractiveness	$\gamma=1.5$ if $\theta\rightarrow 1$ $\gamma\rightarrow 2$ if $\theta\rightarrow 0$	Dorogovtsev and Mendes, 2001a
Internal edges with probab. p Rewiring of edges with probab. q	$\gamma=2$ if $q=\dfrac{1-p+m}{1+2m}$ $\gamma\rightarrow\infty$ if $p,q,m\rightarrow 0$	Albert and Barabási, 2000
c internal edges or removal of c edges	$\gamma\rightarrow 2$ if $c\rightarrow\infty$ $\gamma\rightarrow\infty$ if $c\rightarrow -1$	Dorogovtsev and Mendes, 2000c
Gradual aging $\Pi(k_i)\sim k_i(t-t_i)^{-\nu}$	$\gamma\rightarrow 2$ if $\nu\rightarrow -\infty$ $\gamma\rightarrow\infty$ if $\nu\rightarrow 1$	Dorogovtsev and Mendes, 2000b
Multiplicative node fitness $\Pi_i\sim\eta_i k_i$	$P(k)\sim\dfrac{k^{-1-C}}{\ln(k)}$	Bianconi and Barabási, 2001a
Edge inheritance	$P(k_{in})=\dfrac{d}{k_{in}^{\sqrt{2}}}\ln(ak_{in})$	Dorogovtsev, Mendes, and Samukhin, 2000c
Copying with probab. p	$\gamma=(2-p)/(1-p)$	Kumar *et al.*, 2000a, 2000b
Redirection with probab. r	$\gamma=1+1/r$	Krapivsky and Redner, 2001
Walking with probab. p	$\gamma\simeq 2$ for $p>p_c$	Vázquez, 2000
Attaching to edges	$\gamma=3$	Dorogovtsev, Mendes, and Samukhin, 2001a
p directed internal edges $\Pi(k_i,k_j)\propto(k_i^{in}+\lambda)(k_j^{out}+\mu)$	$\gamma_{in}=2+p\lambda$ $\gamma_{out}=1+(1-p)^{-1}+\mu p/(1-p)$	Krapivsky, Rodgers, and Redner, 2001

Section 11: Summary

Number of Nodes

$N = t$

Number of Links

$N = mt$

Average Degree

$\langle k \rangle = 2m$

Degree Dynamics

$k_i(t) = m\ (t/t_i)^{\beta}$

Dynamical Exponent

$\beta = 1/2$

Degree Distribution

$p_k \sim k^{-\gamma}$

Degree Exponent

$\gamma = 3$

Average Distance

$\langle d \rangle \sim \log N / \log \log N$

Clustering Coefficient

$\langle C \rangle \sim (\ln N)^2 / N$

The network grows, but the degree distribution is stationary.

Section 11: Summary

Number of Nodes
$N = t$

Number of Links
$N = mt$

Average Degree
$\langle k \rangle = 2m$

Degree Dynamics
$k_i(t) = m\,(t/t_i)^{\beta}$

Dynamical Exponent
$\beta = 1/2$

Degree Distribution
$p_k \sim k^{-\gamma}$

Degree Exponent
$\gamma = 3$

Average Distance
$\langle d \rangle \sim \log N / \log \log N$

Clustering Coefficient
$\langle C \rangle \sim (\ln N)^2 / N$

Consequently, the modeling philosophy behind the model is simple: *to understand the topology of a complex system, we need to describe how it came into being.*

The network grows, but the degree distribution is stationary.

Section 11: Summary

Number of Nodes
$N = t$

Number of Links
$N = mt$

Average Degree
$\langle k \rangle = 2m$

Degree Dynamics
$k_i(t) = m\ (t/t_i)^\beta$

Dynamical Exponent
$\beta = 1/2$

Degree Distribution
$p_k \sim k^{-\gamma}$

Degree Exponent
$\gamma = 3$

Average Distance
$\langle d \rangle \sim \log N / \log \log N$

Clustering Coefficient
$\langle C \rangle \sim (\ln N)^2 / N$

- The model predicts γ=3 while the degree exponent of real networks varies between 2 and 5 (Table 4.2).

- Many networks, like the WWW or citation networks, are directed, while the model generates undirected networks.

- Many processes observed in networks, from linking to already existing nodes to the disappearance of links and nodes, are absent from the model.

- The model does not allow us to distinguish between nodes based on some intrinsic characteristics, like the novelty of a research paper or the utility of a webpage.

- While the Barabási-Albert model is occasionally used as a model of the Internet or the cell, in reality it is not designed to capture the details of any particular real network. It is a minimal, proof of principle model whose main purpose is to capture the basic mechanisms responsible for the emergence of the scale-free property. Therefore, if we want to understand the evolution of systems like the Internet, the cell or the WWW, we need to incorporate the important details that contribute to the time evolution of these systems, like the directed nature of the WWW, the possibility of internal links and node and link removal.

LESSONS LEARNED: evolving network models

1. There is no universal exponent characterizing all networks.

2. Growth and preferential attachment are responsible for the emergence of the scale-free property.
3. The origins of the preferential attachment is system-dependent.
4. Modeling real networks:
   - identify the microscopic processes that take place in the system
   - measure their frequency from real data
   - develop dynamical models that capture these processes.

5. If the model is correct, it should correctly predict not only the degree exponent, but both small and large k-cutoffs.

LESSONS LEARNED: evolving network models

Philosophical change in network modeling:

ER, WS models are static models – the role of the network modeler it to cleverly place the links between a fixed number of nodes to that the network topology mimic the networks seen in real systems.

BA and evolving network models are dynamical models: they aim to reproduce how the network was built and evolved.

Thus their goal is to capture the network dynamics, not the structure.

→ as a byproduct, you get the topology correctly

LESSONS LEARNED: evolving network models

Philosophical change in network modeling:

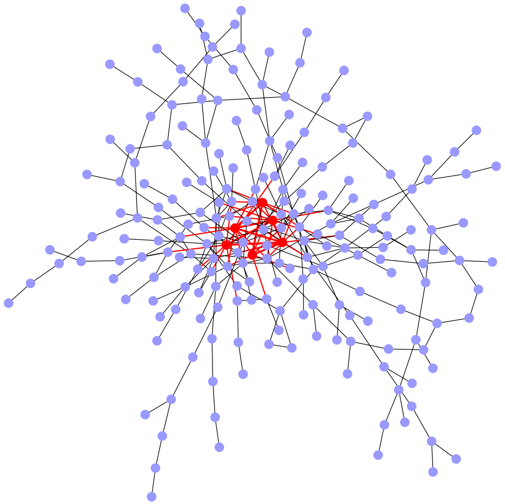
ER, WS models are static models – the role of the network modeler it to cleverly place the links between a fixed number of nodes to that the network topology mimic the networks seen in real systems.

BA and evolving network models are dynamical models: they aim to reproduce how the network was built and evolved.

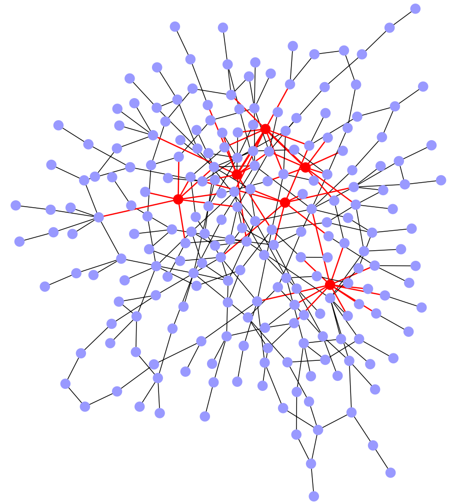
Thus their goal is to capture the network dynamics, not the structure.

→ as a byproduct, you get the topology correctly

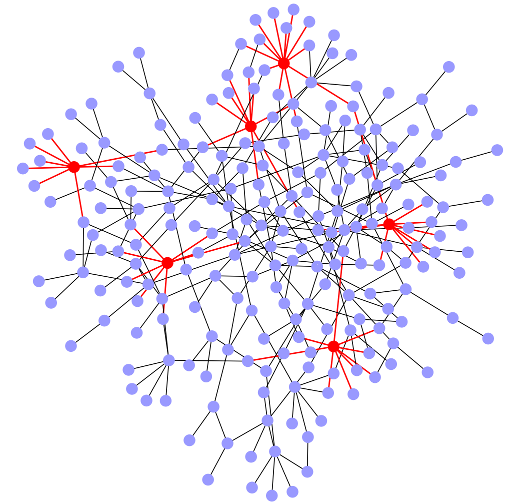
DEGREE CORRELATIONS IN NETWORKS



Assortative:
hubs show a tendency to link to each other.



Neutral:
nodes connect to each other with the expected random probabilities.



Disassortative:
Hubs tend to avoid linking to each other.

Quantifying degree correlations (three approaches):

→ full statistical description (Maslov and Sneppen, Science 2001)
→ degree correlation function (Pastor Satorras and Vespignani, PRL 2001)
→ correlation coefficient (Newman, PRL 2002)

STATISTICAL DESCRIPTION

e_{jk}: probability to find a node with degree *j* and degree *k* at the two ends of a randomly selected edge

$$\sum_{j,k} e_{jk} = 1 \qquad \sum_{j} e_{jk} = q_k$$

q_k: the probability to have a degree *k* node at the end of a link.

Where: $$q_k = \frac{kp_k}{\langle k \rangle}$$

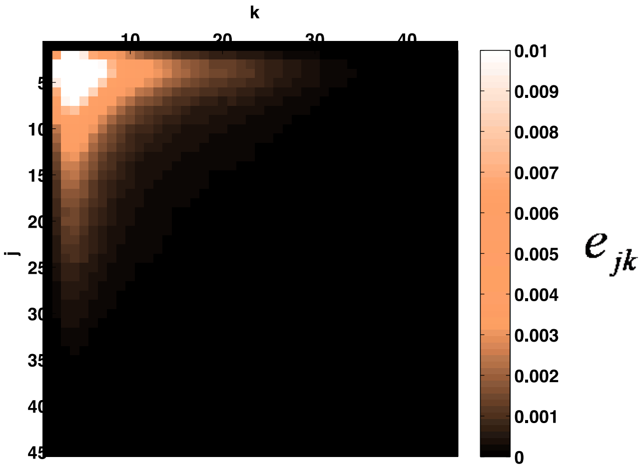
Probability to find a node at the end of a link is biased towards the more connected nodes, i.e. $q_k = Ckp_k$, where C is a normalization constant . After normalization we find $C = 1/\langle k \rangle$, or $q_k = kp_k/\langle k \rangle$

If the network has no degree correlations: $$e_{jk} = q_j q_k$$

Deviations from this prediction are a signature of *degree correlation.*

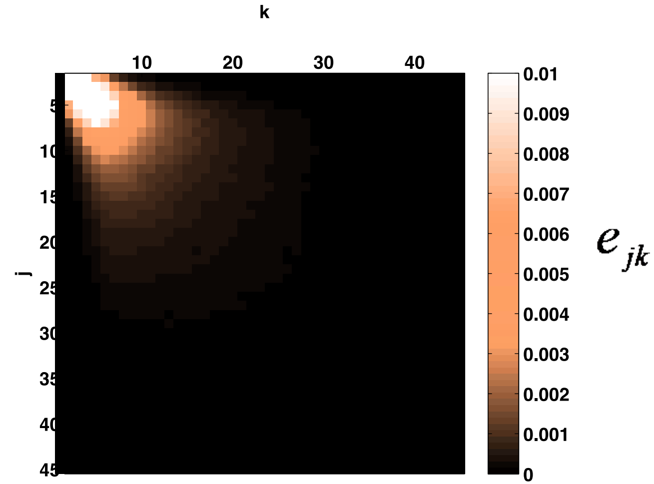
M. E. J. Newman, Phys. Rev. Lett. **89,** 208701 (2002)

EXAMPLE: e_{jk} FOR A SCALE-FREE NETWORK

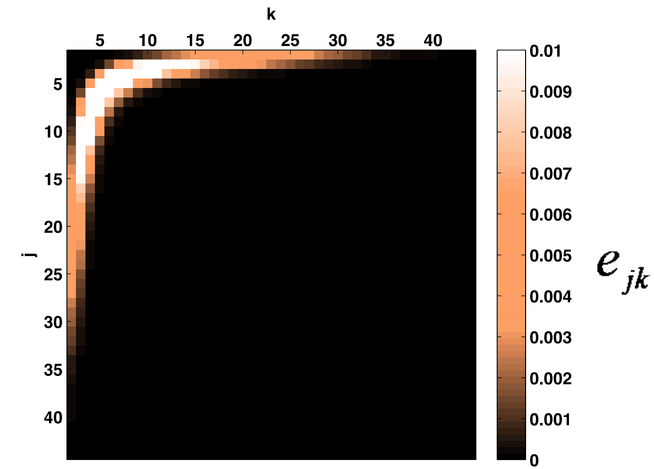


Neutral

Assortative: More strength in the diagonal, hubs tend to link to each other.



Disassortative: Hubs tend to connect to small nodes.



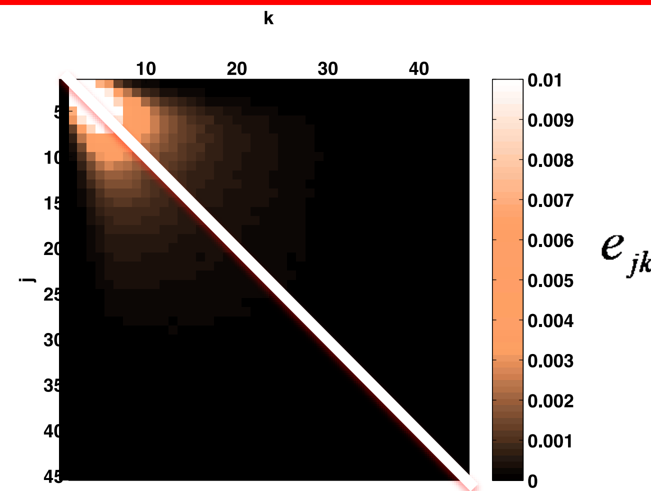
Each matrix is the average of a 100 independent scale-free networks, generated using the static model with $N=10^4$, $\gamma=2.5$ and $<k>=3$.

EXAMPLE: e_{jk} FOR A SCALE-FREE NETWORK

Perfectly assortative network:

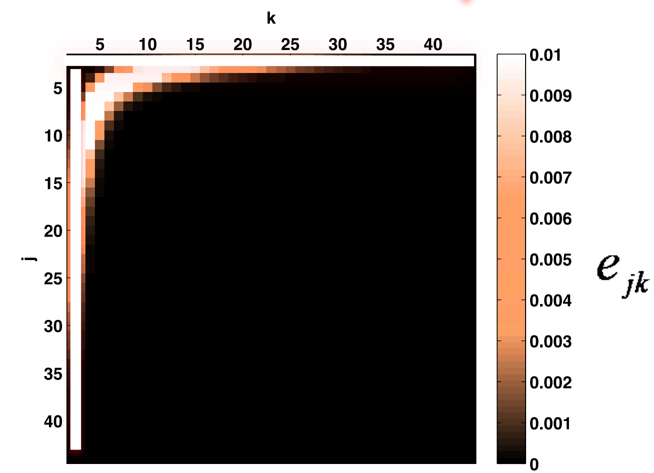
$$e_{jk}=q_k\delta_{jk}$$

Assortative:
More strength in the diagonal, hubs tend to link to each other.



Perfectly disassortative network:

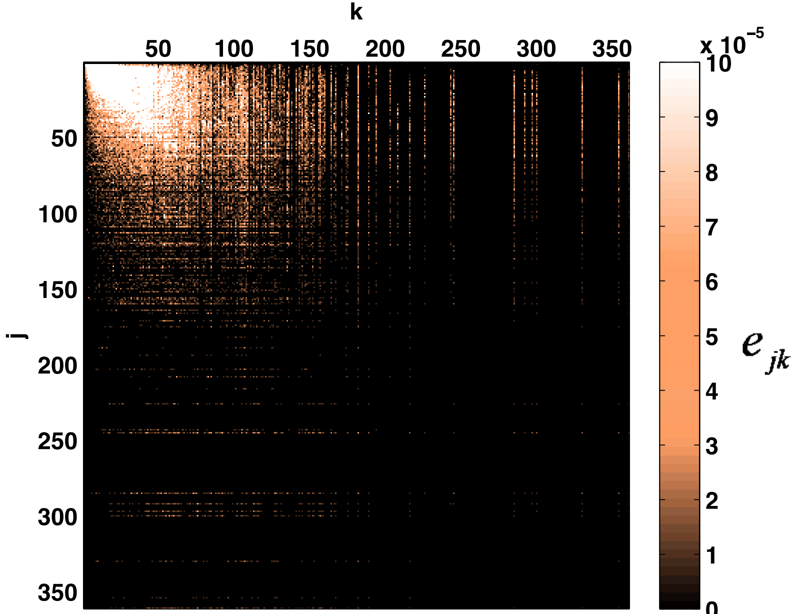
Disassortative:
Hubs tend to connect to small nodes.



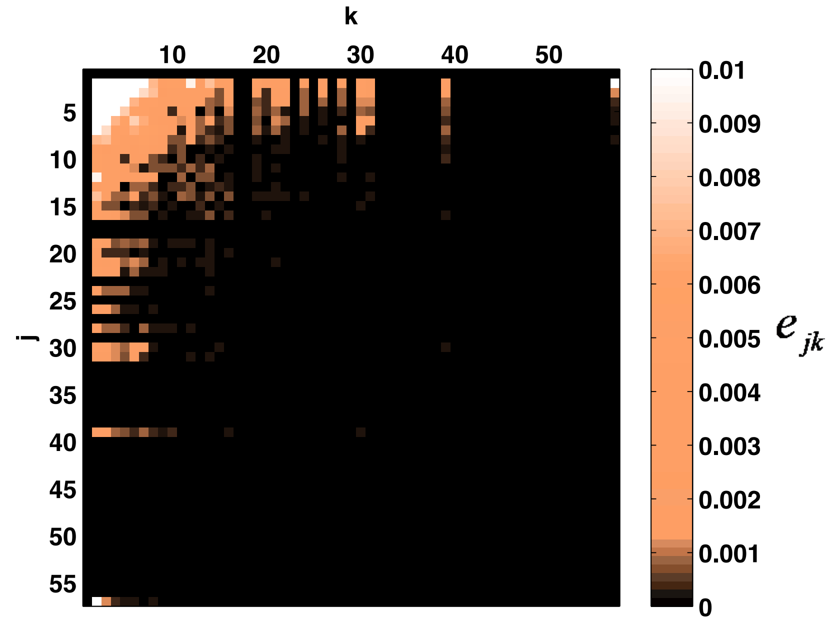
Each matrix is the average of a 100 independent scale-free networks, generated using the static model with $N=10^4$, $\gamma=2.5$ and $\langle k \rangle=3$.

REAL-WORLD EXAMPLES

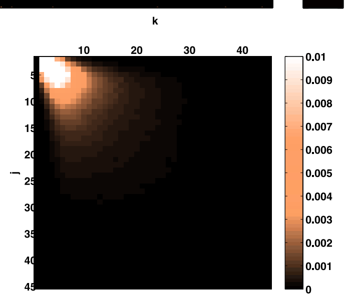
Astrophysics co-authorship network



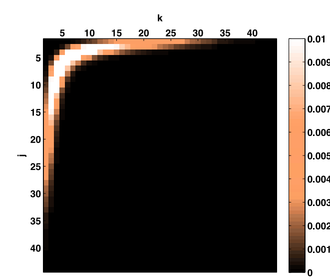
Yeast PPI



Assortative: More strength in the diagonal, hubs tend to link to each other.

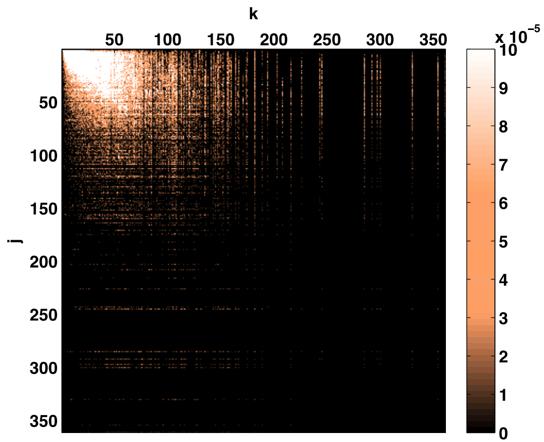


Disassortative: Hubs tend to connect to small nodes.



PROBLEM WITH THE FULL STATISTICAL DESCRIPTION

(1) Difficult to extract information from a visual inspection of a matrix.



(2) Based on e_{jk} and hence requires a large number of elements to inspect:

$$\frac{k_{\max}(k_{\max}-1)}{2} - 1 - k_{\max}$$

Nr. of independent elements

Undirected network: k_{max} x k_{max} matrix

$$\sum_{j,k} e_{jk} = 1 \qquad \sum_{j=1,k_{\max}} e_{jk} = q_k$$

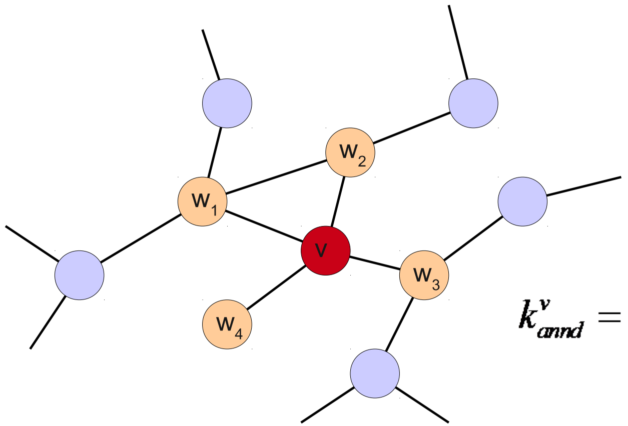
Constraints

We need to find a way to reduce the information contained in e_{jk}

M. E. J. Newman, Phys. Rev. Lett. **89,** 208701 (2002)

Average next neighbor degree

$k_{annd}(k)$: average degree of the first neighbors of nodes with degree k.



$$k^{v}_{annd} = \frac{4+3+3+1}{4}$$

$$k_{annd}(k) = \sum_{k'} k' P(k' \mid k) = \frac{\sum_{k'} k' e_{kk'}}{\sum_{k'} e_{kk'}}$$

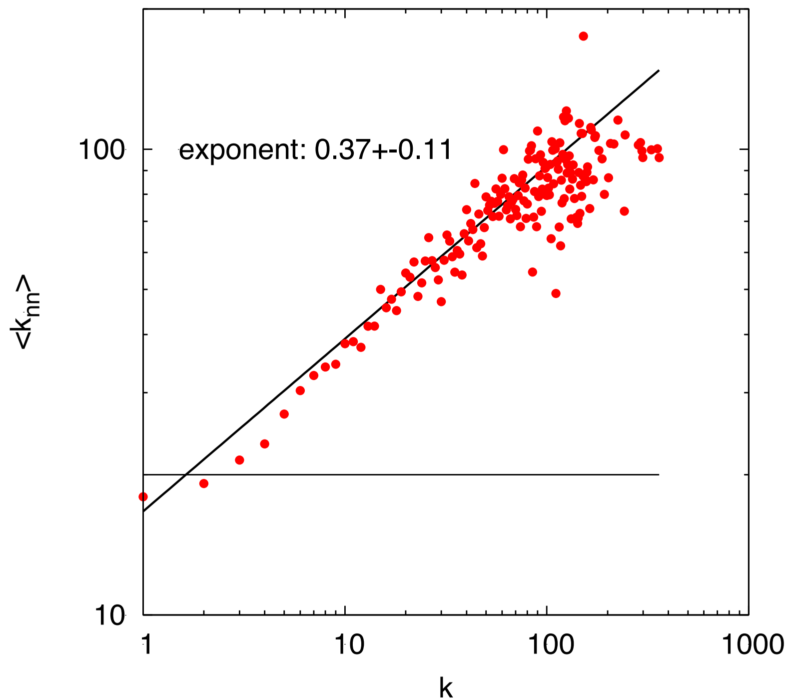
No degree correlations:

$$k_{annd}(k) = \frac{\sum_{k'} k' e_{kk'}}{\sum_{k'} e_{kk'}} = \frac{\sum_{k'} k' q_k q_{k'}}{q_k} = \sum_{k'} k' q_{k'} = \sum_{k'} k' \frac{k' p(k')}{< k >} = \frac{\langle k^2 \rangle}{\langle k \rangle}$$

If there are no degree correlations, $k_{annd}(k)$ is independent of k.

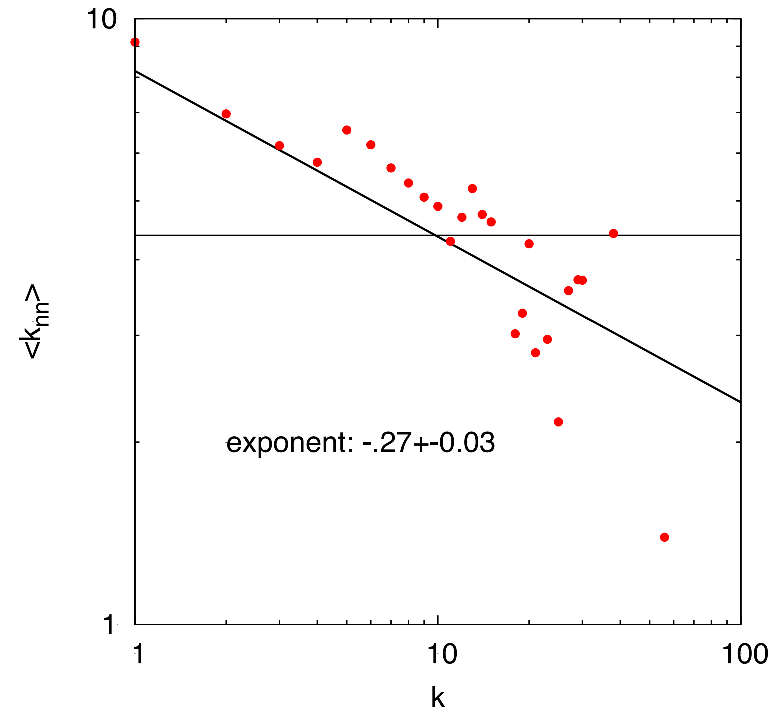
R. Pastor-Satorras, A. Vázquez, A. Vespignani, Phys. Rev. E **65**, 066130 (2001)

$k_{annd}(k)$ FOR REAL NETWORKS



Astrophysics co-authorship network

Assortative



Yeast PPI

Disassortative

Average next neighbor degree

$k_{annd}(k)$: average degree of the first neighbors of nodes with degree k.

constraint:

$$\sum_k k_{annd}(k) \cdot kNp_k = \sum_k k^2 \cdot Np_k$$

$$\langle k_{annd}(k)k \rangle = \langle k^2 \rangle$$

$\longrightarrow$ k_{max}-1 independent elements

$k_{annd}(k)$ is a k-dependent function, hence it has much fewer parameters,

and it is easier to interpret/read.

R. Pastor-Satorras, A. Vázquez, A. Vespignani, Phys. Rev. E **65**, 066130 (2001)

PEARSON CORRELATION

If there are degree correlations, e_{jk} will differ from $q_j q_k$. The magnitude of the correlation is captured by <jk>-<j><k> difference, which is:

$$\sum_{jk} jk(e_{jk} - q_j q_k)$$

<jk>-<j><k> is expected to be:

positive for *assortative* networks,
zero for *neutral* networks,
negative for *dissasortative* networks

To compare different networks, we should normalize it with its maximum value; the maximum is reached for a *perfectly assortative network*, i.e. $e_{jk}=q_k\delta_{jk}$

normalization:

$$\sigma_r^2 = \max \sum_{jk} jk(e_{jk} - q_j q_k) = \sum_{jk} jk(q_k \delta_{jk} - q_j q_k)$$

$$r = \frac{\sum_{jk} jk(e_{jk} - q_j q_k)}{\sigma_r^2} \qquad -1 \le r \le 1$$

$r \le 0$ *disassortative*
$r = 0$ *neutral*
$r \ge 0$ *assortative*

M. E. J. Newman, Phys. Rev. Lett. **89,** 208701 (2002)

REAL NETWORKS

Social networks are *assortative*

Network	n	r
Physics coauthorship (a)	52 909	0.363
Biology coauthorship (a)	1 520 251	0.127
Mathematics coauthorship (b)	253 339	0.120
Film actor collaborations (c)	449 913	0.208
Company directors (d)	7 673	0.276
Internet (e)	10 697	−0.189
World-Wide Web (f)	269 504	−0.065
Protein interactions (g)	2 115	−0.156
Neural network (h)	307	−0.163
Marine food web (i)	134	−0.247
Freshwater food web (j)	92	−0.276
Random graph (u)		0
Callaway *et al.* (v)		$\delta/(1+2\delta)$
Barabási and Albert (w)		0

Biological, technological networks are *disassortative*

r>0: assortative network:
Hubs tend to connect to other hubs.

r<0: disassortative network:
Hubs tend to connect to small nodes.

RELATIONSHIP BETWEEN r AND k_{annd}

$$r = \frac{\sum_{kj} kj(e_{kj} - q_k q_j)}{\sigma_r^2} = \frac{\sum_k k q_k \sum_j \frac{j e_{kj}}{q_k} - \left(\sum_k k q_k\right)^2}{\sigma_r^2} = \frac{\sum_k k k_{annd}(k) q_k - \frac{\langle k^2 \rangle^2}{\langle k \rangle^2}}{\sigma_r^2}$$

$$k_{annd}(k) = \sum_{k'} k' P(k' \mid k) = \frac{\sum_{k'} k' e_{kk'}}{\sum_{k'} e_{kk'}} = \frac{\sum_{k'} k' e_{kk'}}{q_k}$$

In general case we need to know q_k and $k_{annd}(k)$ to calculate r.

Assuming: $k_{annd}(k) = a \cdot k + b$

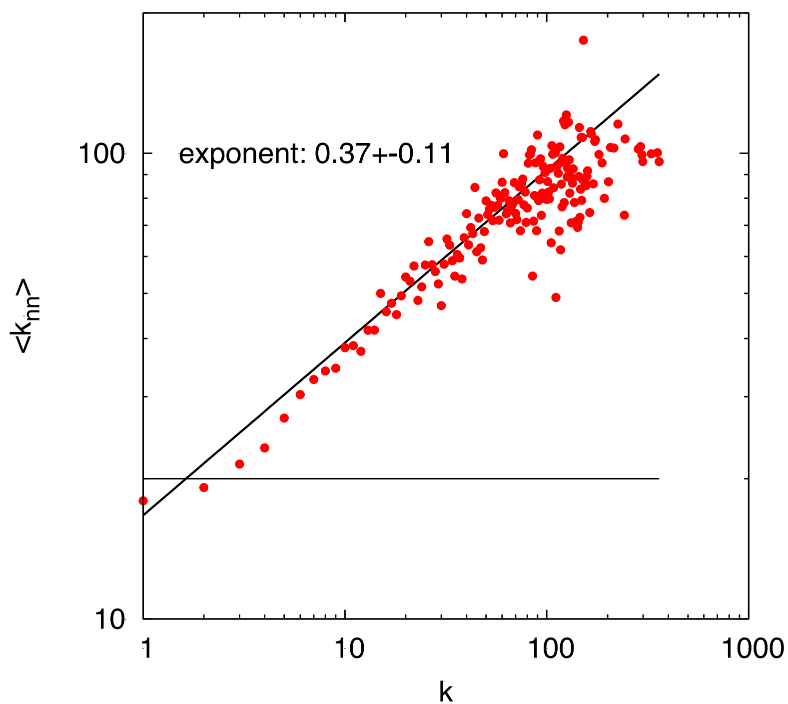
Using the constraint for ANND:

$$\langle k^2 \rangle = \langle k_{annd}(k) k \rangle = \sum_{k'} a \cdot k^2 p_k + b \cdot k p_k = a \langle k^2 \rangle + b \langle k \rangle \longrightarrow b = \frac{(1-a)\langle k^2 \rangle}{\langle k \rangle}$$

$$r = \frac{\sum_k k \cdot (a \cdot k + b) q_k - \frac{\langle k^2 \rangle^2}{\langle k \rangle^2}}{\sigma_r^2} = \frac{\sum_k k \cdot \left(a \cdot k + \frac{(1-a)\langle k^2 \rangle}{\langle k \rangle}\right) \frac{k p_k}{\langle k \rangle} - \frac{\langle k^2 \rangle^2}{\langle k \rangle^2}}{\sigma_r^2} =$$

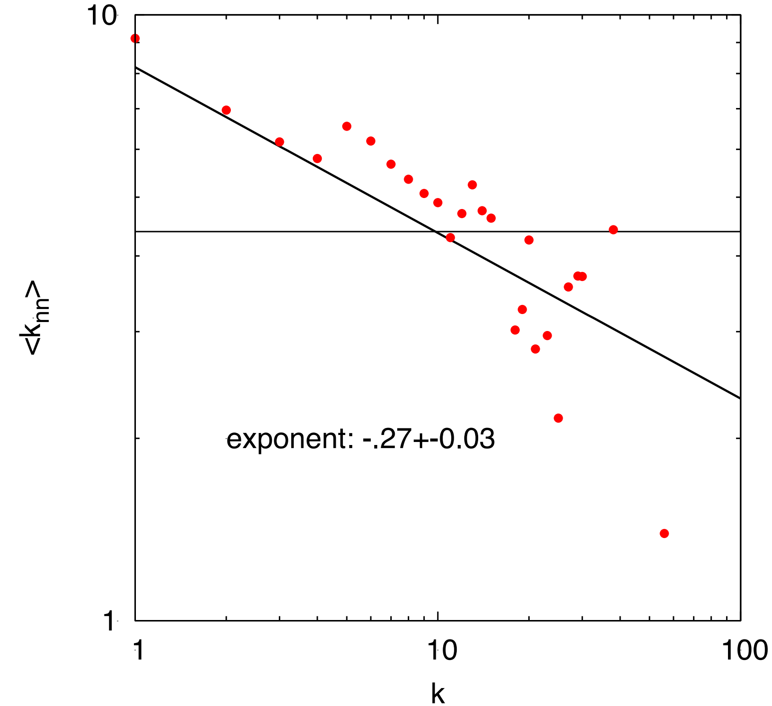
$$= \frac{a\left(\sum_k k^3 \frac{p_k}{\langle k \rangle} - \frac{\langle k^2 \rangle^2}{\langle k \rangle^2}\right) + \frac{\langle k^2 \rangle^2}{\langle k \rangle^2} - \frac{\langle k^2 \rangle^2}{\langle k \rangle^2}}{\sigma_r^2} = a$$

PROBLEM WITH THE PREVIOUS DEVIATION: $k_{annd}(k) \sim k^{\beta}$



Astrophysics co-authorship network

Assortative



Yeast PPI

Disassortative

CONNECTION WITH *ANND*

Assuming: $k_{annd}(k) = a \cdot k^{\beta}$

Using the constraint for ANND: $\langle k^2 \rangle = \langle k_{annd}(k) k \rangle = \sum_{k'} a \cdot k^{\beta+1} p_k = a \langle k^{\beta+1} \rangle \longrightarrow a = \frac{\langle k^2 \rangle}{\langle k^{\beta+1} \rangle}$

$$r = \frac{\sum_k k \cdot a k^{\beta} \cdot q_k - \frac{\langle k^2 \rangle^2}{\langle k \rangle^2}}{\sigma_r^2} = \frac{\sum_k a \cdot k^{\beta+2} \frac{p_k}{\langle k \rangle} \cdot - \frac{\langle k^2 \rangle^2}{\langle k \rangle^2}}{\sigma_r^2} = \frac{\frac{\langle k^2 \rangle}{\langle k^{\beta+1} \rangle} \frac{\langle k^{\beta+2} \rangle}{\langle k \rangle} - \frac{\langle k^2 \rangle^2}{\langle k \rangle^2}}{\sigma_r^2} =$$

$$= \frac{1}{\sigma_r^2} \frac{\langle k^2 \rangle}{\langle k \rangle} \left(\frac{\langle k^{\beta+2} \rangle}{\langle k^{\beta+1} \rangle} - \frac{\langle k^2 \rangle}{\langle k \rangle} \right)$$

$$\sigma_r^2 = \sum_{jk} jk (q_k \delta_{jk} - q_j q_k) = \frac{\langle k^3 \rangle}{\langle k \rangle} - \frac{\langle k^2 \rangle^2}{\langle k \rangle^2}$$

$$\beta < 0 \quad \rightarrow \quad r < 0$$

$$\beta = 0 \quad \rightarrow \quad r = 0$$

$$\beta > 0 \quad \rightarrow \quad r > 0$$

CONNECTION BETWEEN R AND k_{ANND}

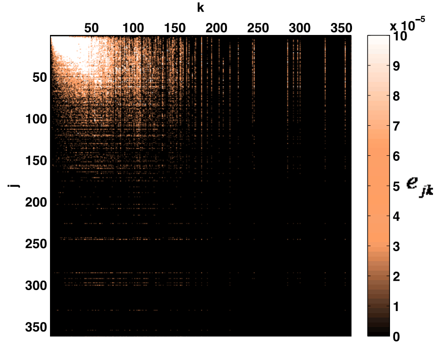
$$\beta = 0: \quad \frac{\langle k^{\beta+2}\rangle}{\langle k^{\beta+1}\rangle} - \frac{\langle k^2\rangle}{\langle k\rangle} = \frac{\langle k^2\rangle}{\langle k\rangle} - \frac{\langle k^2\rangle}{\langle k\rangle} = 0 \quad \Rightarrow \quad r = 0$$

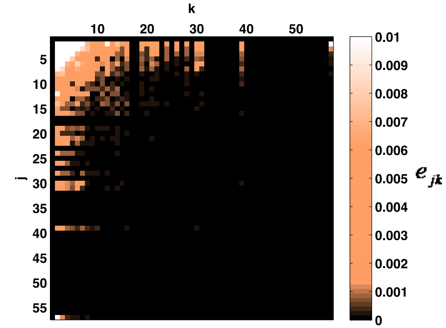
$$\langle k^{\alpha+\beta}\rangle = \sum_{k_{\min}}^{k_{\max}} k^{\alpha+\beta} p_k \begin{array}{l} < k_{\max}^{\beta} \sum_{k_{\min}}^{k_{\max}} k^{\alpha} p_k = k_{\max}^{\beta} \langle k^{\alpha}\rangle \\ > k_{\min}^{\beta} \sum_{k_{\min}}^{k_{\max}} k^{\alpha} p_k = k_{\min}^{\beta} \langle k^{\alpha}\rangle \end{array}$$

$$\longrightarrow \quad \begin{array}{l} 0 > \beta > -1: \quad \dfrac{\langle k^2\rangle}{\langle k\rangle} > \left(\dfrac{k_{\min}}{k_{\max}}\right)^{-\beta} \dfrac{\langle k^{\beta+2}\rangle}{\langle k^{\beta+1}\rangle} > \dfrac{\langle k^{\beta+2}\rangle}{\langle k^{\beta+1}\rangle} \quad \Rightarrow \quad r < 0 \\ +1 > \beta > 0: \quad \dfrac{\langle k^{\beta+2}\rangle}{\langle k^{\beta+1}\rangle} > \left(\dfrac{k_{\min}}{k_{\max}}\right)^{\beta} \dfrac{\langle k^2\rangle}{\langle k\rangle} > \dfrac{\langle k^2\rangle}{\langle k\rangle} \quad \Rightarrow \quad r > 0 \end{array}$$

DEGREE CORRELATION IN NETWORKS

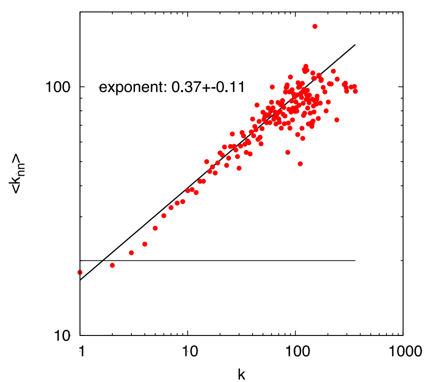
e_{jk}

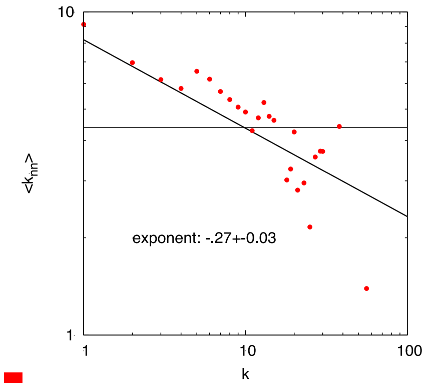




$$\frac{k_{max}\left(k_{max}-1\right)}{2}-k_{max}-1$$

$k_{annd}(k)$





$k_{max}-1$

r 0.31 -0.16 1

GENERATING NETWORK WITH GIVEN ASSORTATIVITY

We have a desired e_{jk} distribution, which also specifies p_k.

1. Generate a network with the desired degree distribution using the configuration model.
2. Choose two links at random from the network: (v_1,w_1) and (v_2,w_2).
3. Measure the degrees j_1, k_1, j_2, k_2 of nodes v_1, w_1, v_2, w_2. Replace the two selected links with two new ones (v_1,v_2) and (w_1,w_2) with probability

$$P = \begin{cases} \dfrac{e_{j_1 j_2} e_{k_1 k_2}}{e_{j_1 k_1} e_{k_2 j_2}} & \text{if } e_{j_1 j_2} e_{k_1 k_2} < e_{j_1 k_1} e_{k_2 j_2} \\ 1 & \text{otherwise} \end{cases}$$

1. Repeat from step 2.

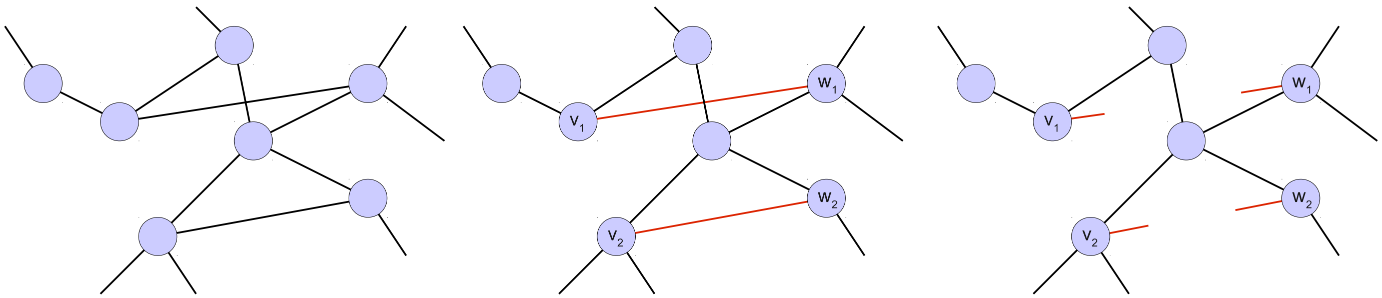
The algorithm is ergodic and satisfies detailed balance, therefore in the long time limit it samples the desired network ensemble correctly.

M. E. J. Newman, Phys. Rev. E **67,** 026126 (2003)

GENERATING NETWORK WITH GIVEN ASSORTATIVITY

2. Choose two edges random from the network: (v_1,w_1) and (v_2,w_2).
3. Measure the degrees j_1, k_1, j_2, k_2 of vertices v_1, w_1, v_2, w_2. Replace the two selected edges with two new ones (v_1,v_2) and (w_1,w_2) with probability

$$P = \begin{cases} \dfrac{e_{j_1 j_2} e_{k_1 k_2}}{e_{j_1 k_1} e_{k_2 j_2}} & \text{if } e_{j_1 j_2} e_{k_1 k_2} < e_{j_1 k_1} e_{k_2 j_2} \\ 1 & \text{otherwise} \end{cases}$$



1 2 3 4

GENERATING NETWORK WITH GIVEN ASSORTATIVITY

If we only specify *r* we have great degree of freedom in choosing e_{jk}.

Possible choice for disassortative case:

$$e_{jk}^{(d)} = q_j x_k + x_j q_k - x_j x_k$$

Where x_k is any normalized distribution.

This form satisfies the constraints on e_{jk}:

$$\sum_{jk} e_{jk} = \sum_{jk} q_k x_j + x_k q_j - x_k x_j = 1 + 1 - 1 = 1 \qquad \sum_{j} e_{jk} = \sum_{j} q_k x_j + x_k q_j - x_k x_j = q_k + x_k - x_k = q_k$$

The *r* value can be easily calculated:

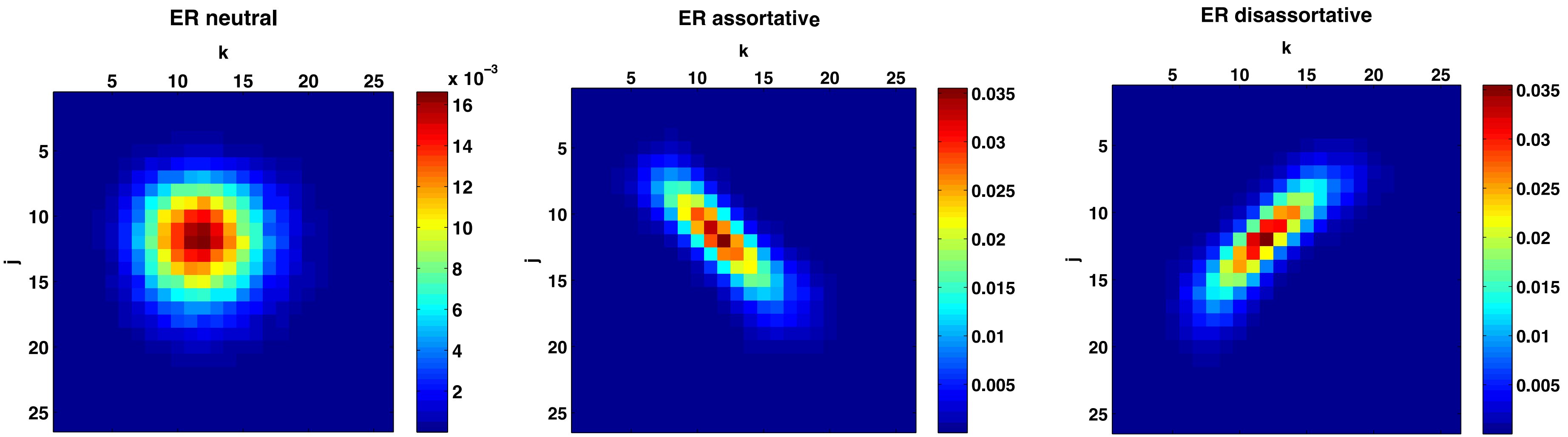
$$r_d = \frac{\sum_{jk} jk\left(q_k x_j + x_k q_j - x_k x_j - q_k q_j\right)}{\sigma_r^2} = \frac{2\langle k\rangle_q \langle k\rangle_x - \langle k\rangle_x^2 - \langle k\rangle_q^2}{\sigma_r^2} = -\frac{\left(\langle k\rangle_x - \langle k\rangle_q\right)^2}{\sigma_r^2}$$

Assortative case:

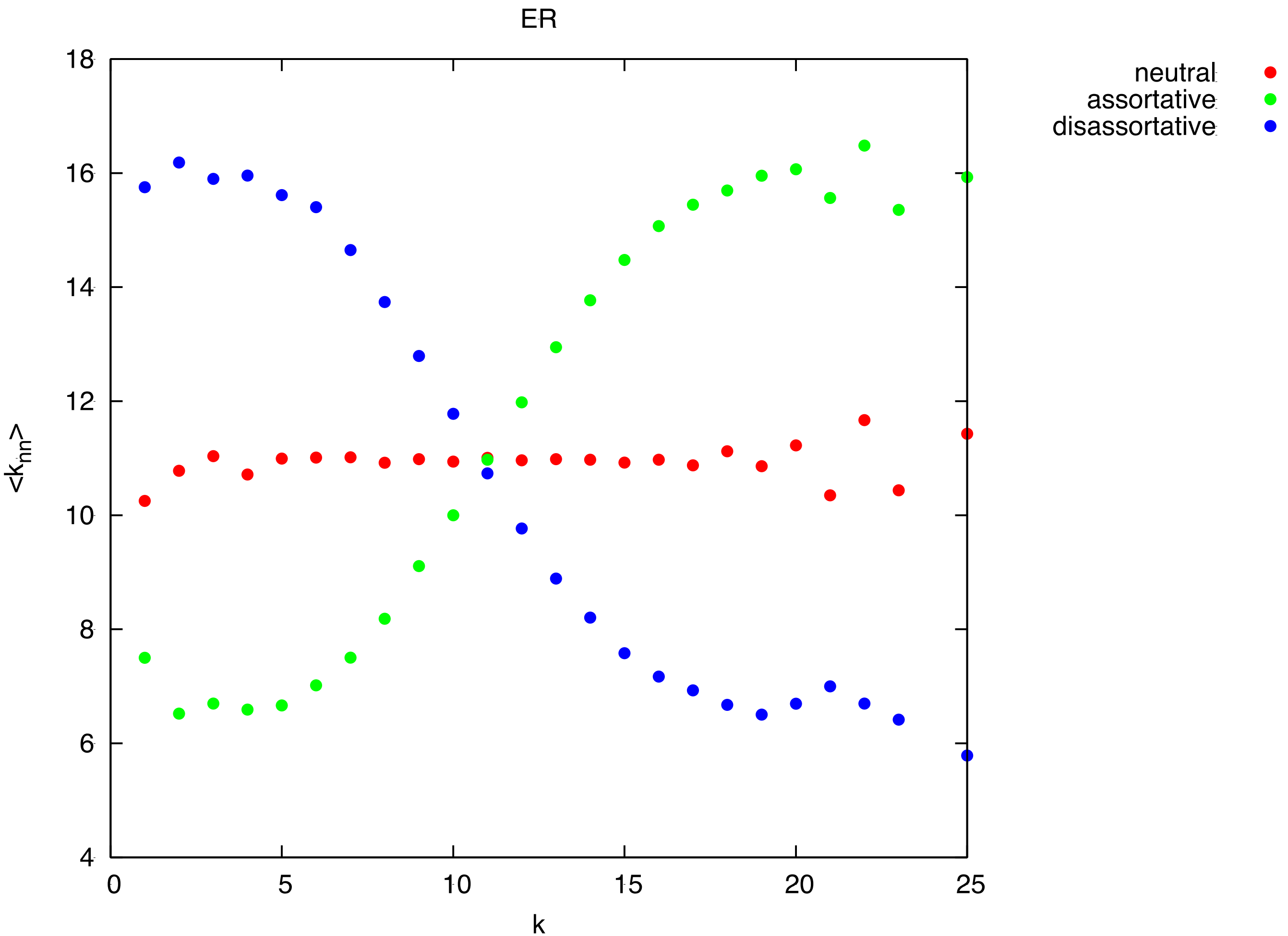
$$e_{jk}^{(a)} = q_j q_k - e_{jk}^{(d)} \longrightarrow r_a = -r_d$$

M. E. J. Newman, Phys. Rev. E **67,** 026126 (2003)

EXAMPLE: Erdős-Rényi



EXAMPLE: Erdős-Rényi



Structural cut-off

High assortativity→ high number of links between the hubs.

If we allow only one link between two nodes, we can simply run out of hubs to connect to each other to satisfy the assortativity criteria.

Number of edges between the set of nodes with degree *k* and degree *k'*:

$$E_{kk'} = e_{kk'} \langle k \rangle N$$

Maximum number of edges between the two groups:

$$m_{kk'} = \min\{kN_k, k'N_{k'}, N_k N_{k'}\}$$

There cannot be more links between the two groups, than the overall number of edges joining the nodes with degree k.

If we only have **simple edges**, we cannot have more links between the two groups, than if we connect every node with degree k to every node with degree k' **once**.

This is true even if we allow multiple edges.

M. Boguñá, R. Pastor-Satorras, A. Vespignani, EPJ B **38,** 205 (2004)

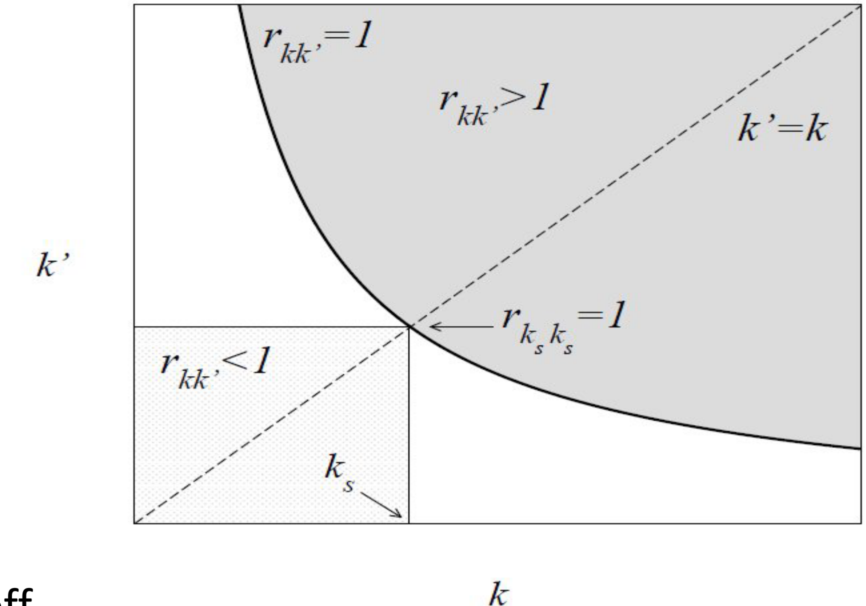
Structural cut-off

$$E_{kk'} = e_{kk'}\langle k\rangle N$$

$$m_{kk'} = \min\{kN_k, k'N_{k'}, N_kN_{k'}\}$$

The ratio of $E_{kk'}$ and $m_{kk'}$ has to be ≤ 1 in the physical region!

$$r_{kk'} = \frac{E_{kk'}}{m_{kk'}} \leq 1$$



$\longrightarrow$ $r_{k_s k_s} = 1$ defines the structural cut-off

M. Boguñá, R. Pastor-Satorras, A. Vespignani, EPJ B **38,** 205 (2004)

Structural cut-off for uncorrelated networks

Uncorrelated networks:

$$m_{kk'} = \min\{kN_k, k'N_{k'}, N_k N_{k'}\}$$

$$m_{k_s k_s} = k_s N_{k_s} = k_s N p_{k_s}$$

$$m_{k_s k_s} = N_{k_s}^2 = N^2 p_{k_s}^2$$

$$e_{kk'} = q_k q_{k'} = \frac{kk' p_k p_{k'}}{\langle k \rangle^2} \longrightarrow r_{kk'} = \frac{E_{kk'}}{m_{kk'}} = \frac{\langle k \rangle N e_{kk'}}{m_{kk'}}$$

$$r_{k_s k_s} = \frac{\langle k \rangle N \cdot k_s^2 \cdot p_{k_s}^2}{\langle k \rangle^2 k_s p_{k_s} N} = \frac{k_s p_{k_s}}{\langle k \rangle} = q_{k_s} < 1 \quad \forall k_s$$

$$r_{k_s k_s} = \frac{\langle k \rangle N \cdot k_s^2 \cdot p_{k_s}^2}{\langle k \rangle^2 N^2 \cdot p_{k_s}^2} = \frac{k_s^2}{\langle k \rangle N} \longrightarrow \boxed{k_s(N) = (\langle k \rangle N)^{1/2}}$$

***$k_s(N)$* represents a *structural cutoff*:**
one cannot have nodes with degree larger than *$k_s(N)$* ,

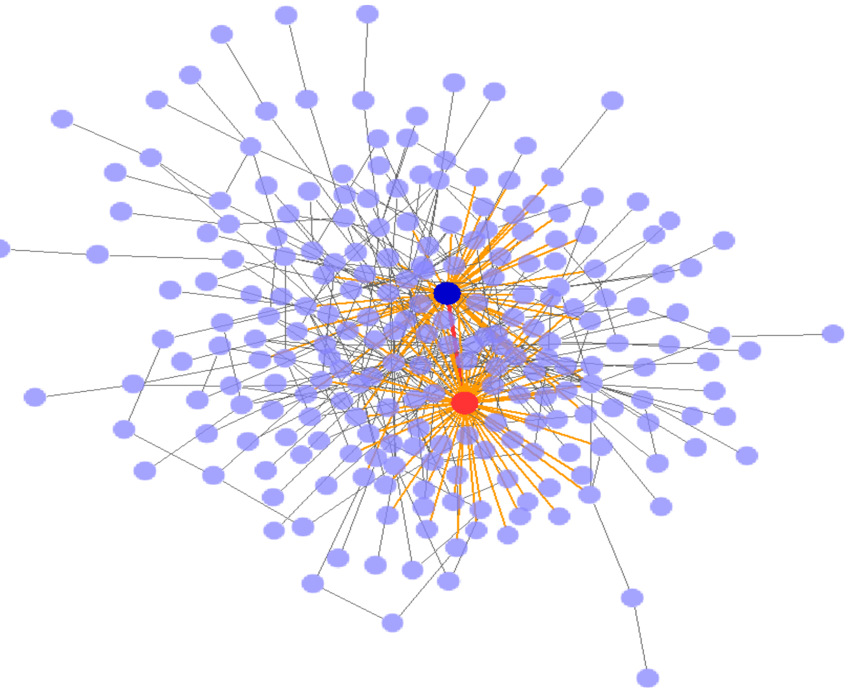
→*if there are nodes with $k > k_s(N)$ we cannot find sufficient links between the highly connected nodes to maintain the neutral nature of the network.*

Solution:

(a) Introduce a structural cutoff (i.e. do not allow nodes with $k > k_s(N)$

(b) Let the network become more dissasortative, having fewer links between hubs.

Example: Degree sequence introduces disassortativity



Scale-free network generated with the configuration model ($N=300$, $L=450$, $\gamma=2.2$).

The measured r=-0.19! → Dissasortative!

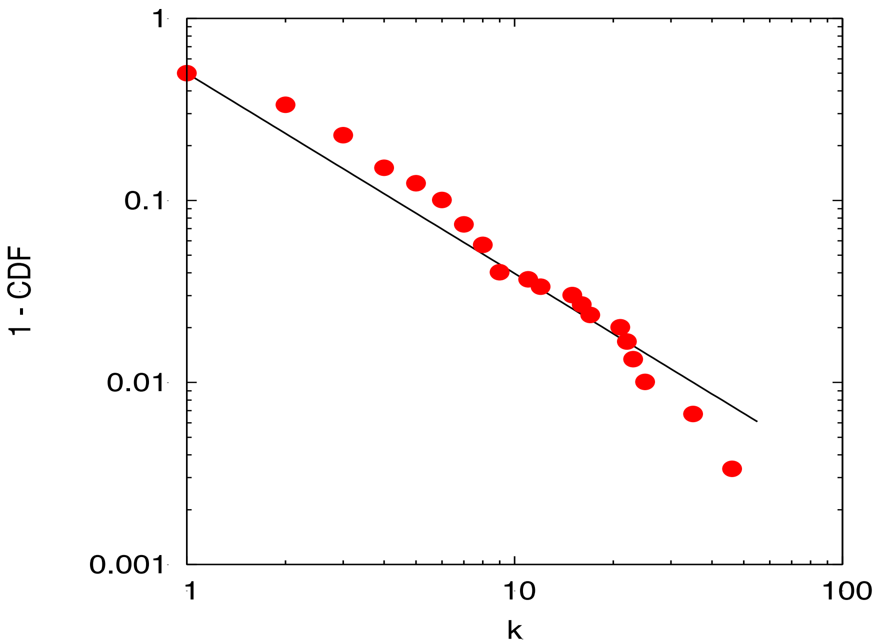
Red hub: 55 neighbors.
Blue hub: 46 neighbors.

Let's calculate the expectation number of links between red node (k=55) and blue node (k=46) for uncorrelated networks!

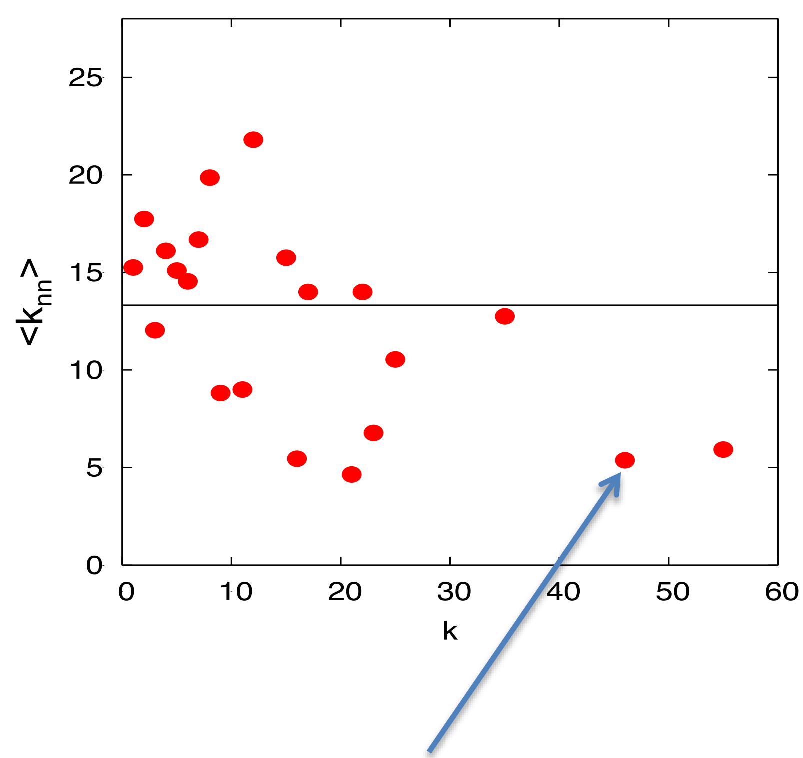
Here $N_{55}=N_{46}=1$, hence $m_{55,46}=1$ so $r_{55,46}=E_{55,46}$

$$E_{55,46} = \langle k\rangle N \cdot e_{55,46} = 900 \cdot \frac{\underbrace{55}_{k}\,\overbrace{\frac{1}{300}}^{p_k} \cdot \overbrace{46}^{k'}\,\overbrace{\frac{1}{300}}^{p_{k'}}}{\underbrace{3^2}_{\langle k\rangle}} \approx 2.8 > 1$$

In order for the network to be neutral, we need 2.8 links between these two hubs.

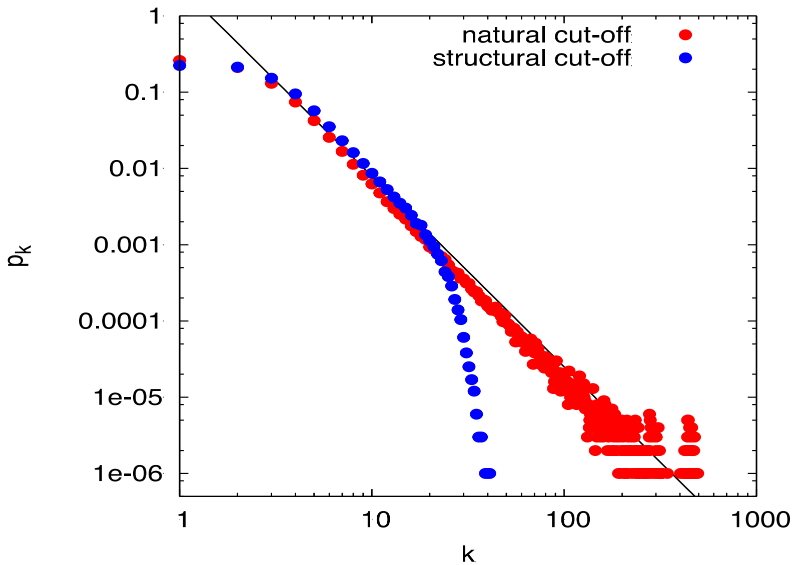


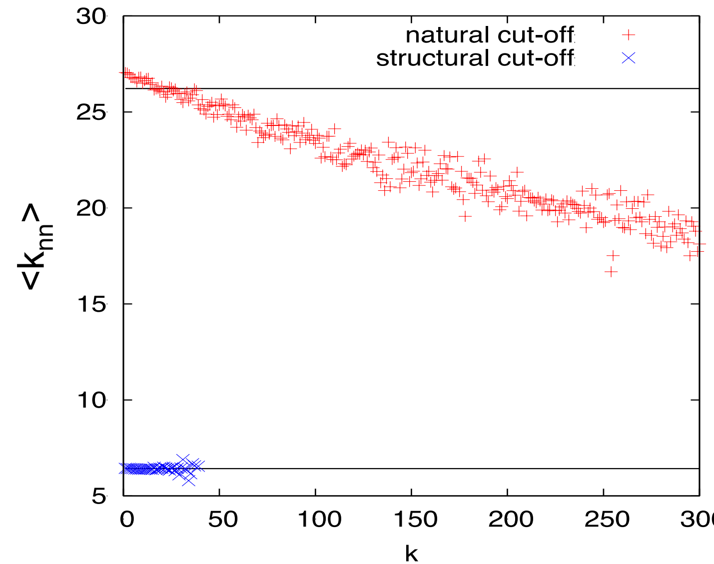
$$1 - CDF = P(k' > k) = 1 - \sum_{k'}^{k} p_{k'}$$



The largest nodes have $k_{nn} < \langle k_{nn} \rangle$

The effect is particularly clear for N=10,000:





The red curves are those of interest to us: one can see that a clear dissasortativity property is visible in this case.

Natural cutoffs in scale-free networks

All real networks are finite → let us explore its consequences.
→ We have an expected maximum degree, K_{max}

Estimating K_{max}

$$\int_{K_{\max}}^{\infty} P(k)dk \approx \frac{1}{N}$$

Why: the probability to have a node larger than K_{max} should not exceed the prob. to have one node, i.e. *1/N* fraction of all nodes

$$\int_{K_{\max}}^{\infty} P(k)dk = (\gamma-1)K_{\min}^{\gamma-1}\int_{K_{\max}}^{\infty} k^{-\gamma}dk = \frac{(\gamma-1)}{(-\gamma+1)}K_{\min}^{\gamma-1}\left[k^{-\gamma+1}\right]_{K_{\max}}^{\infty} = \frac{K_{\min}^{\gamma-1}}{K_{\max}^{\gamma-1}} \approx \frac{1}{N}$$

Natural cutoff: $K_{\max} = K_{\min} N^{\frac{1}{\gamma-1}}$

Structural cut-off for uncorrelated networks

Structural cutoff: $\boxed{k_s(N) \sim (\langle k \rangle N)^{1/2}}$

$$e_{kk'} = q_k q_{k'} = \frac{kk' p_k p_{k'}}{\langle k \rangle^2}$$

Natural cut-off: $k_{\max}(N) \sim N^{\frac{1}{\gamma-1}}$

γ=3: $k_s(N)$ and $k_{max}(N)$ scale the same way, i.e. $\sim N^{1/2}$.

γ<3: $k_{\max} > k_s$ $\longrightarrow$ The size of the largest hub is above the structural cutoff, which means that it cannot have enough links to the other hubs to maintain its neutral status.

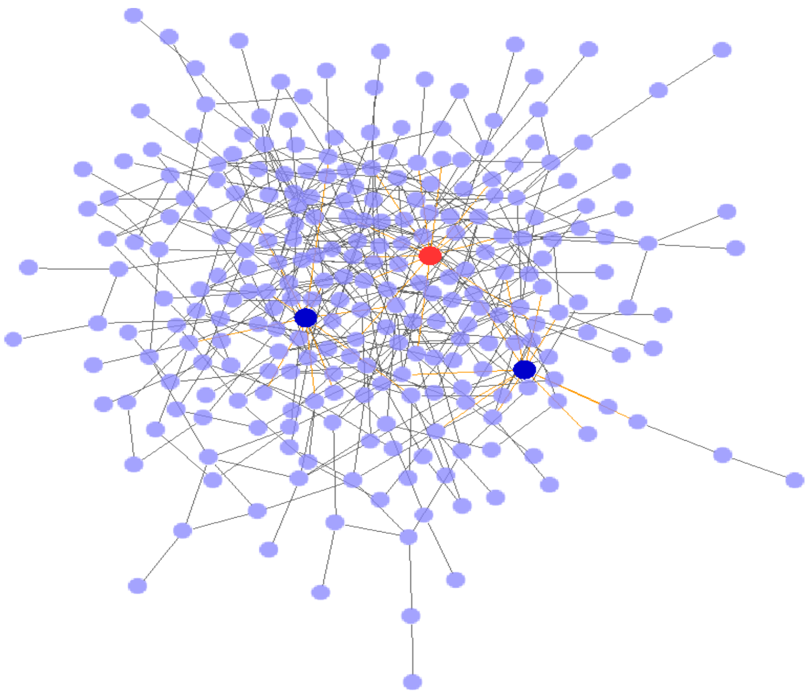
→ ***disassortative mixing***

→a randomly wired network with γ<3 will be

(a) dissasortative

(b) Or will have to have a cutoff at $k_s(N) < k_{max}(N)$

Example: introducing a structural cut-off



Scale-free network generated with the configuration model (N=300, L=450, γ=2.2) with structural cut-off ~ N½.

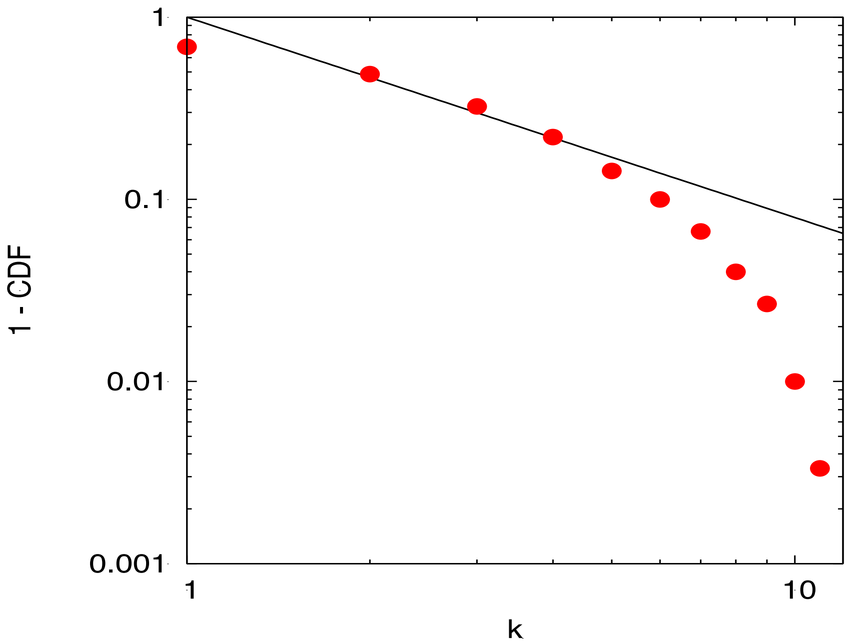
r=0.005 → neutral

Red hub: 12 neighbors.

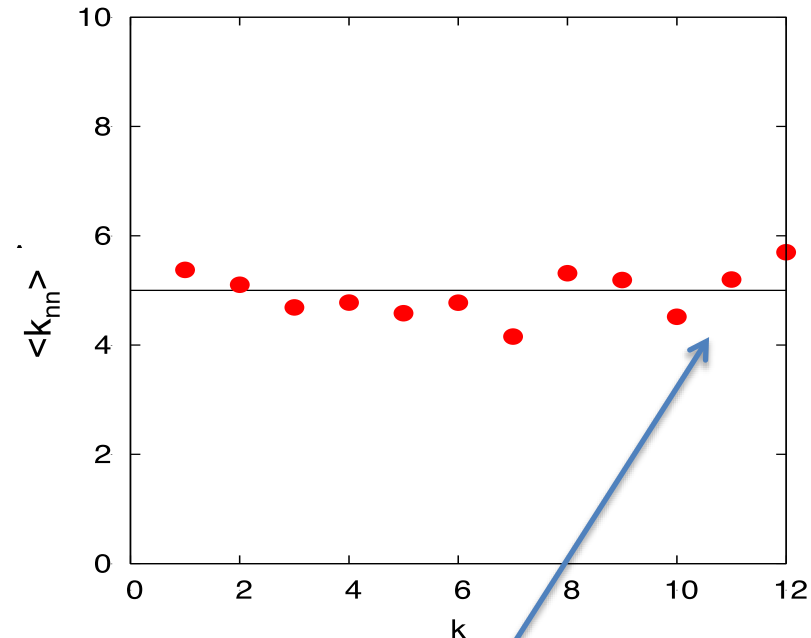
Blue hubs: 11 neighbors.

Again we can calculate the expectation number of edges between the hubs.

$$E_{11,12} = \langle k \rangle N \cdot e_{11,12} = 900 \cdot \frac{\underbrace{12}_{k} \underbrace{\frac{1}{300}}_{p_k} \cdot \underbrace{11}_{k'} \underbrace{\frac{2}{300}}_{p_{k'}}}{\underbrace{3^2}_{\langle k \rangle}} \approx 0.3 < 1$$

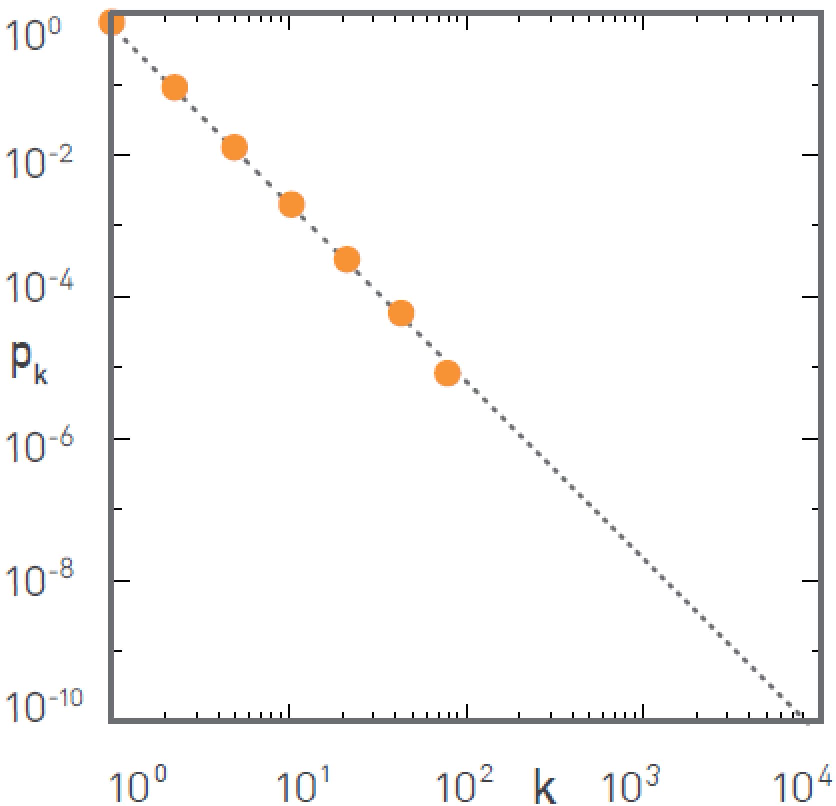


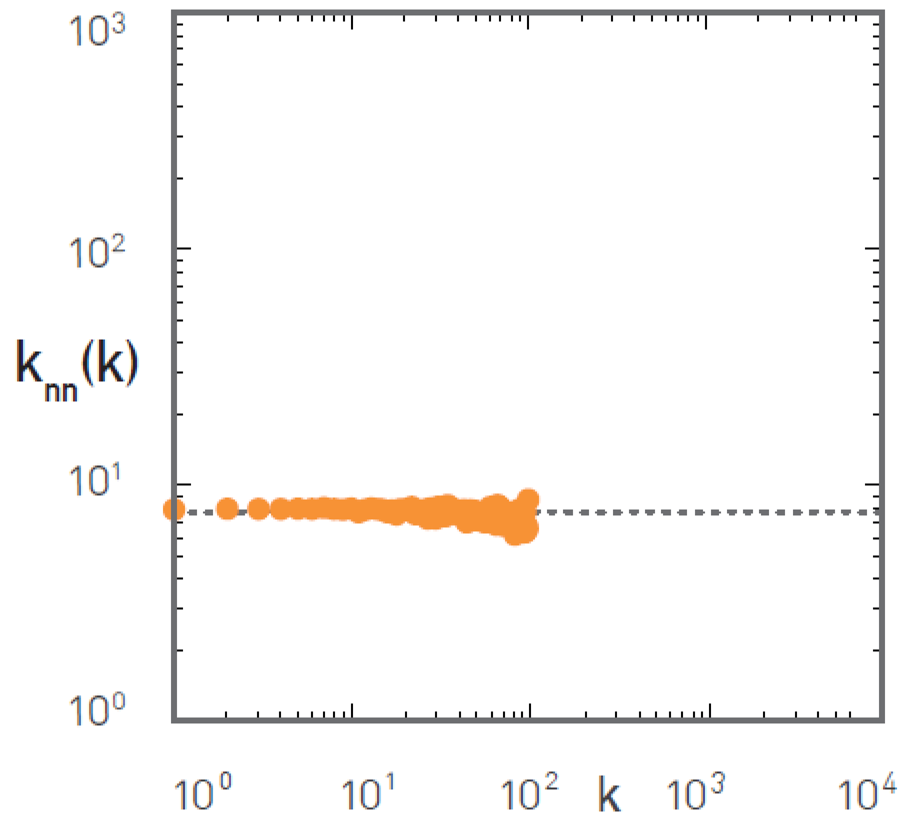
$$1 - CDF = P(k' > k) = 1 - \sum_{k'}^{k} p_{k'}$$



The largest nodes have $k_{nn} \sim \langle k_{nn} \rangle$

The effect is particularly clear for N=10,000:





A clear case of neutral assortativity property is visible in this case thanks to imposing structural cut-off.