DEI
DEPARTAMENTO
DE ENGENHARIA INFORMÁTICA
TÉCNICO LISBOA

Last class

Let a million of variants bloom!

Models of evolving networks

Today

Going beyond degree distributions

1. Assortativity in complex networks
2. Robustness of complex networks

Network Science, 2025/2026

DEI
DEPARTAMENTO
DE ENGENHARIA INFORMÁTICA
TÉCNICO LISBOA

Last class

Let a million of variants bloom!

Models of evolving networks

Network Science, 2025/2026

Classes of models in network science

- ***Static & generative models.*** ER model, Watts-Strogratz model, Configuration Model, etc.

- ***Evolving network models.*** BA model, Initial attractiveness model, fitness model, internal links model, node deletion model, accelerated model, aging model, costs model, minimal model, ranking model, duplication model, hierarchical networks model, etc.

Do we need degree-based preferential attachment to get to scale-free (or strongly heterogeneous) networks?

Can we get to power-laws following different principles?

Scale-free growth by ranking

Fortunato et al. PRL (2006)

- Growth + preferential attachment model in which attachment probability of a new node to an old vertex *s* given by the form

$$\Pi_s = \frac{R_s^{-\alpha}}{\sum_j R_j^{-\alpha}}$$

where R_s denotes the rank of the node *s* for some specific attribute and where α is a positive parameter.

Scale-free growth by ranking

Fortunato et al. PRL (2006)

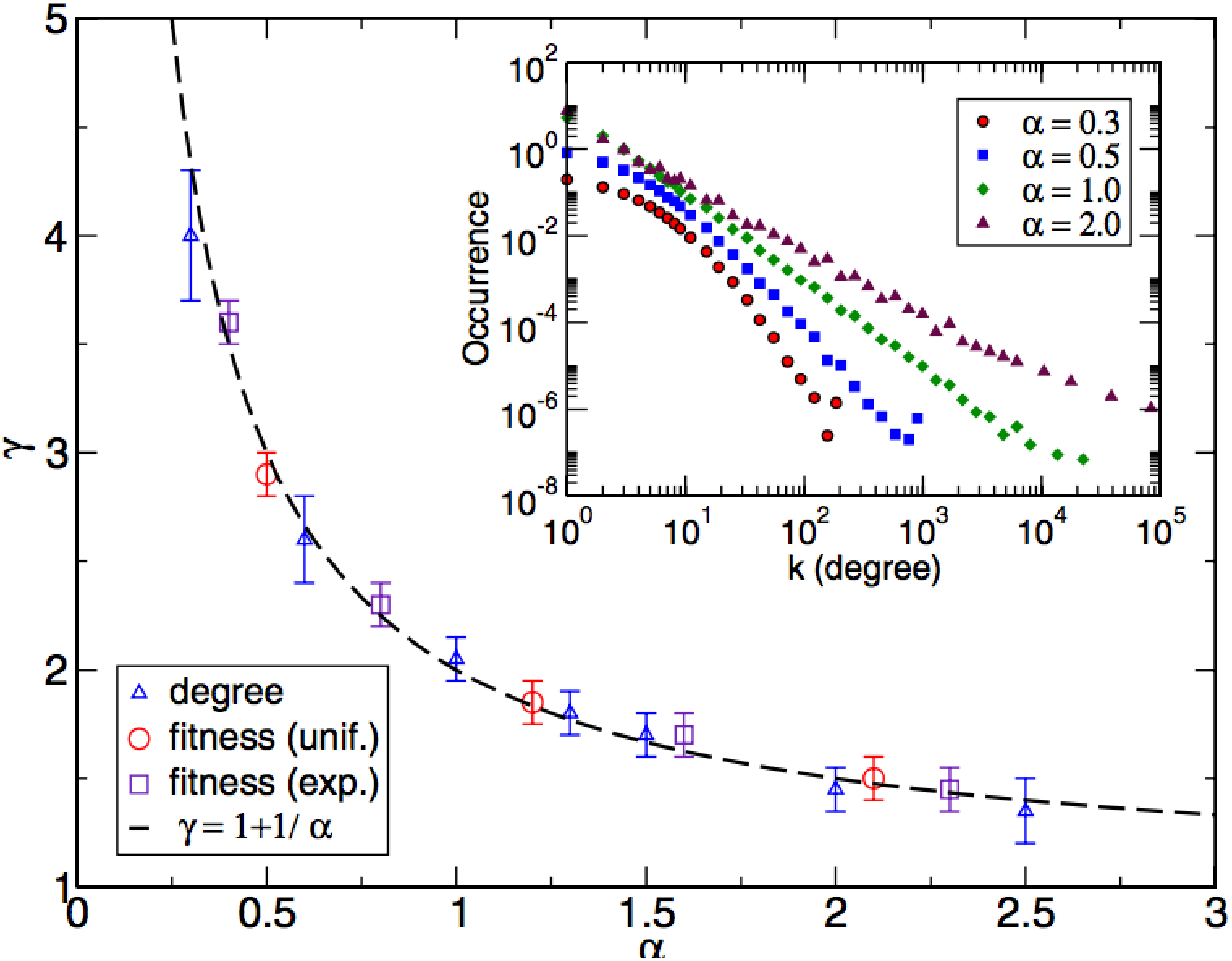
- *Example:* R_s denotes denotes the age-ranking of the nodes of a growing network.

$$P(k) \sim k^{-\left(1+\frac{1}{\alpha}\right)}$$

- Similar behavior will occur if, for instance, we consider the ranking in terms of the in-degree of a node (think for instance on the WWW).
- We get the same type of scaling if we assign a fitness value taken to each node from a given distribution. If we rank the nodes based on this, we get the same behavior.

Scale-free growth by ranking (simulations)

Fortunato et al. PRL (2006)



$$P(k) \sim k^{-\left(1+\frac{1}{\alpha}\right)}$$

Scale-free growth by ranking

Fortunato et al. PRL (2006)

- *Example:* R_s denotes denotes the age-ranking of the nodes of a growing network.

Challenge : Imagine a growth model Page Rank. Would you get to a power-law degree distribution?

- Similar behavior will occur if, for instance, we consider the ranking in terms of the in-degree of a node (think for instance on the WWW).
- A similar pattern will occur if we assign a fitness value taken to each node from a given distribution. If we rank the nodes based on this, we get the same behavior.

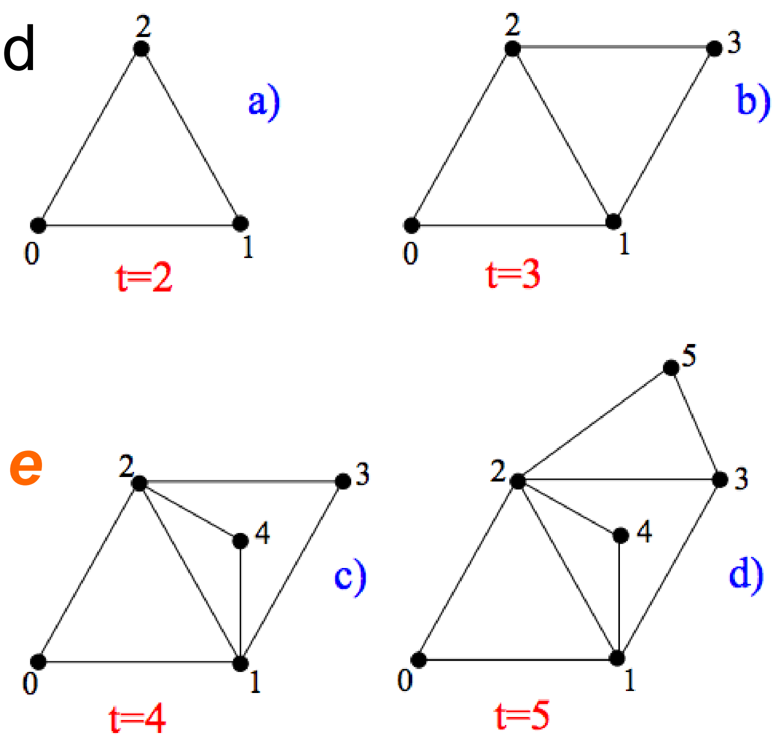
DMS minimal model or link-selection model

Dorogovtsev, ***M***endes, ***S***amukhin Phys. Rev. E 63, 062101 (2001)

- ***Growth***: At each time step we add a new node to the network.

- ***Link selection***: each new node selects a link ***e*** at random, and connects itself to the two ends of ***e***

Note: you may also assume that it connects to only one of the two ends of ***e***.

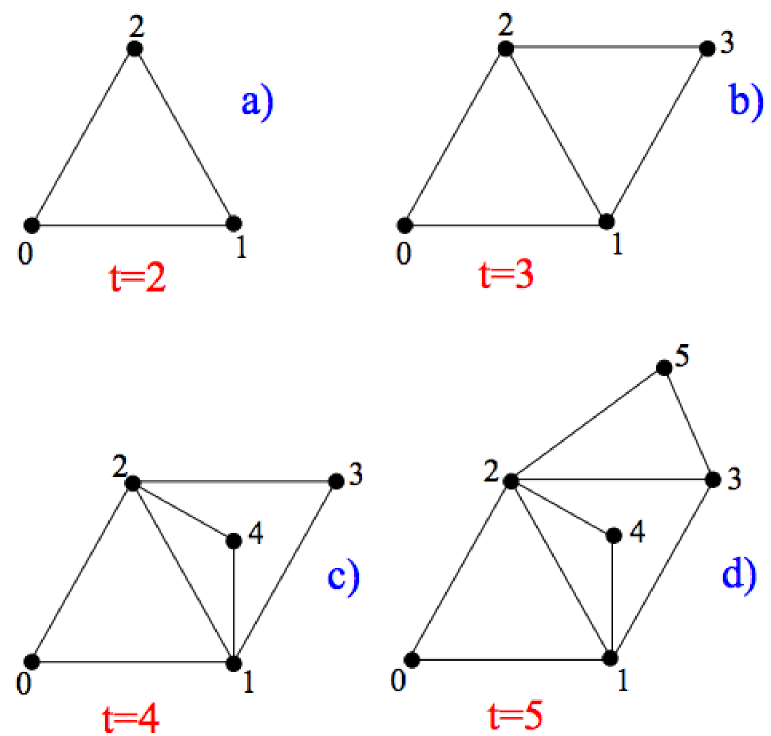


DMS minimal model: degree distribution

Dorogovtsev, **M**endes, **S**amukhin Phys. Rev. E 63, 062101 (2001)

What's the probability that a node i, of degree k_i, gets a link from the new node?

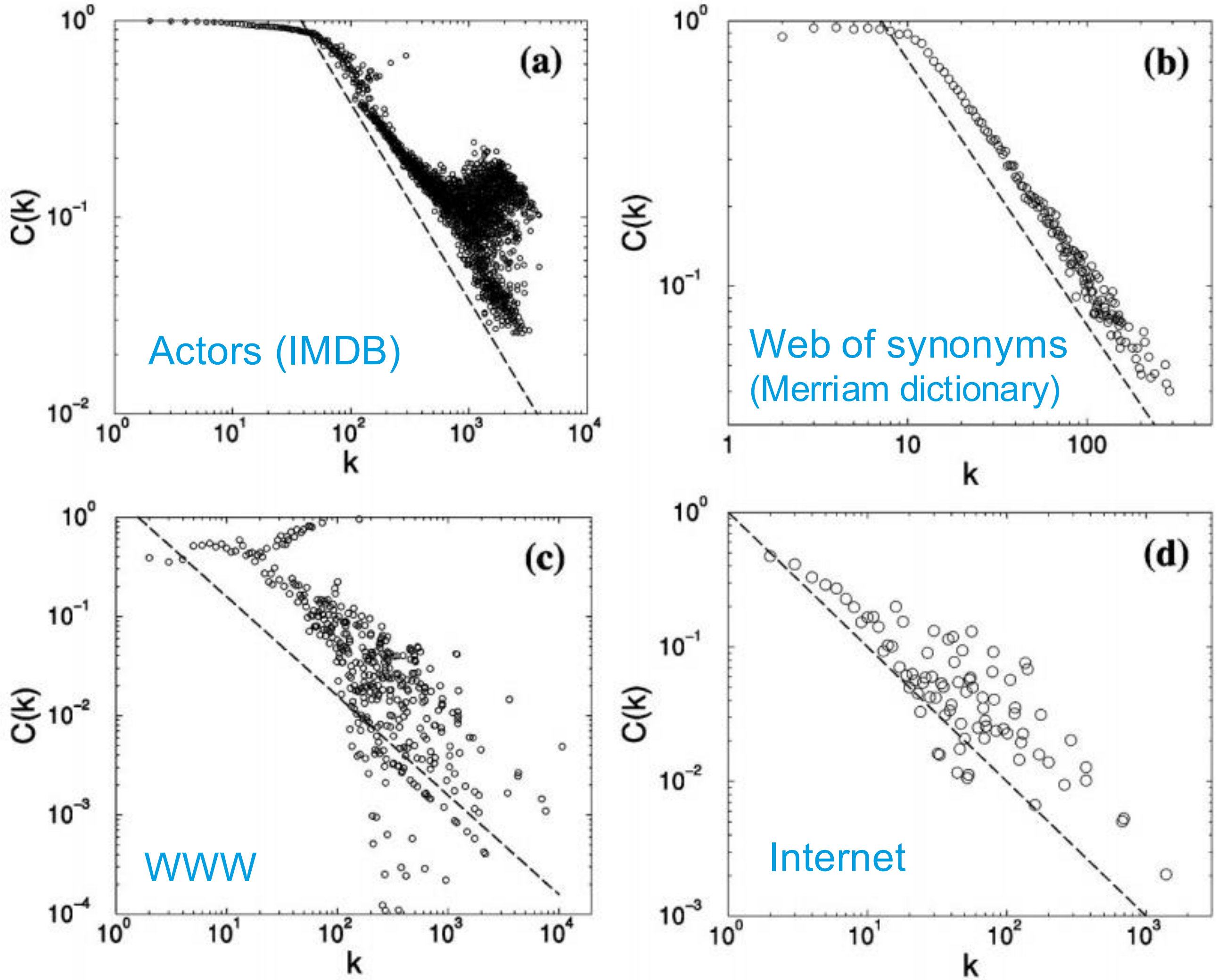
$$\Pi_i = \frac{k_i}{E} = \frac{k_i}{2t-1} \sim \frac{k_i}{\sum_j k_j}$$



i.e., linear preferential attachment, as the Barabási-Albert model ☺

$$P(k) \sim k^{-3}$$

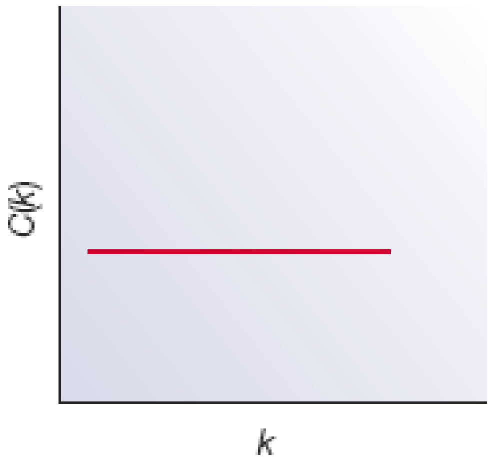
Scaling of clustering coefficient



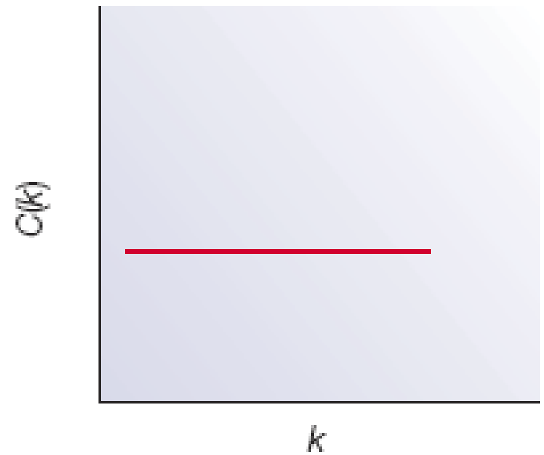
$$C(k) \sim k^{-\beta}$$

DMS Minimal model: Clustering coeff. distribution

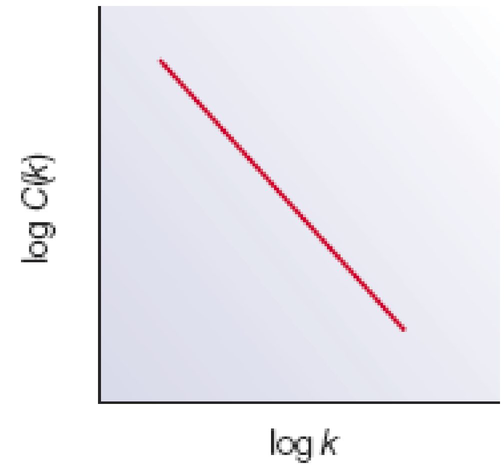
Random Networks

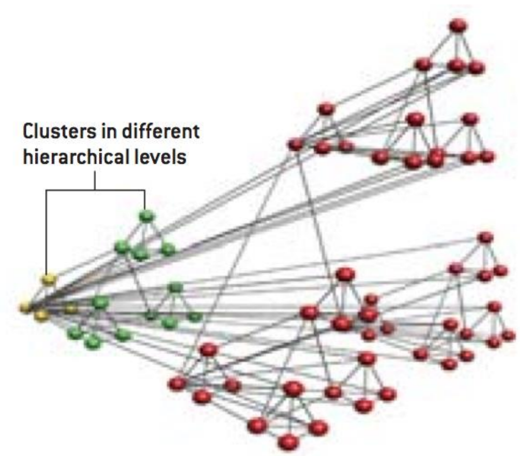


Barabási-Albert model

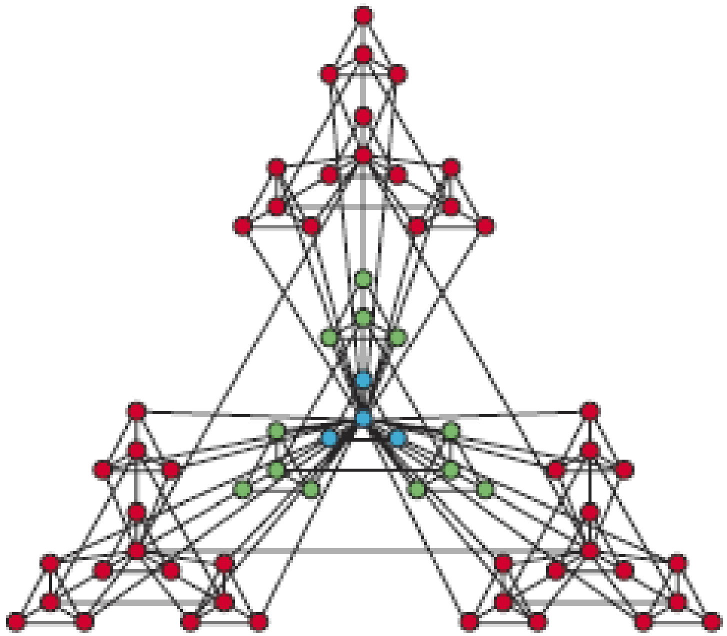


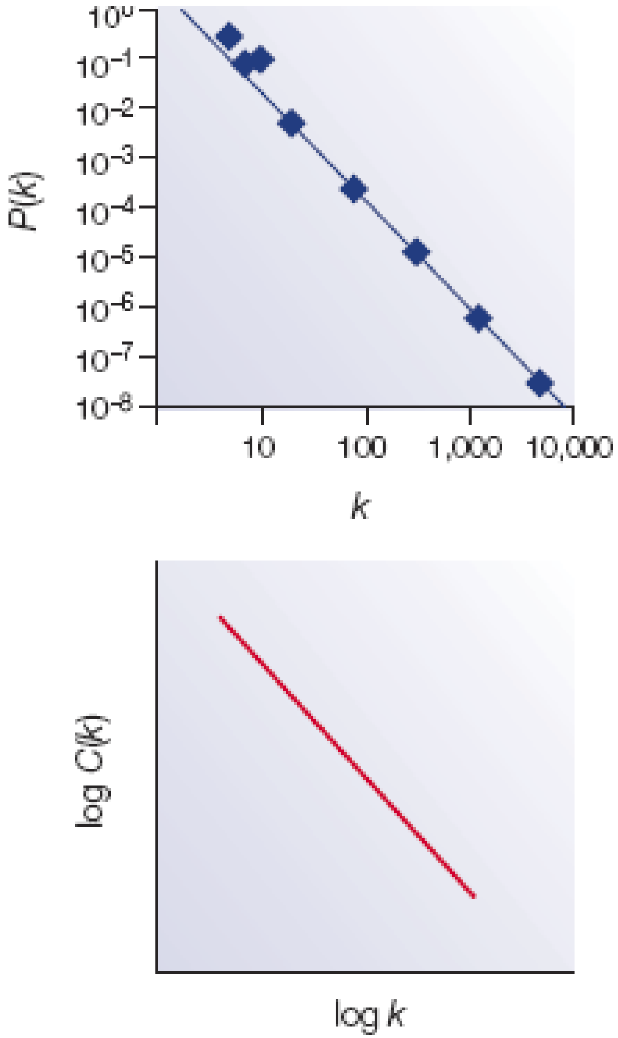
DMS Minimal model



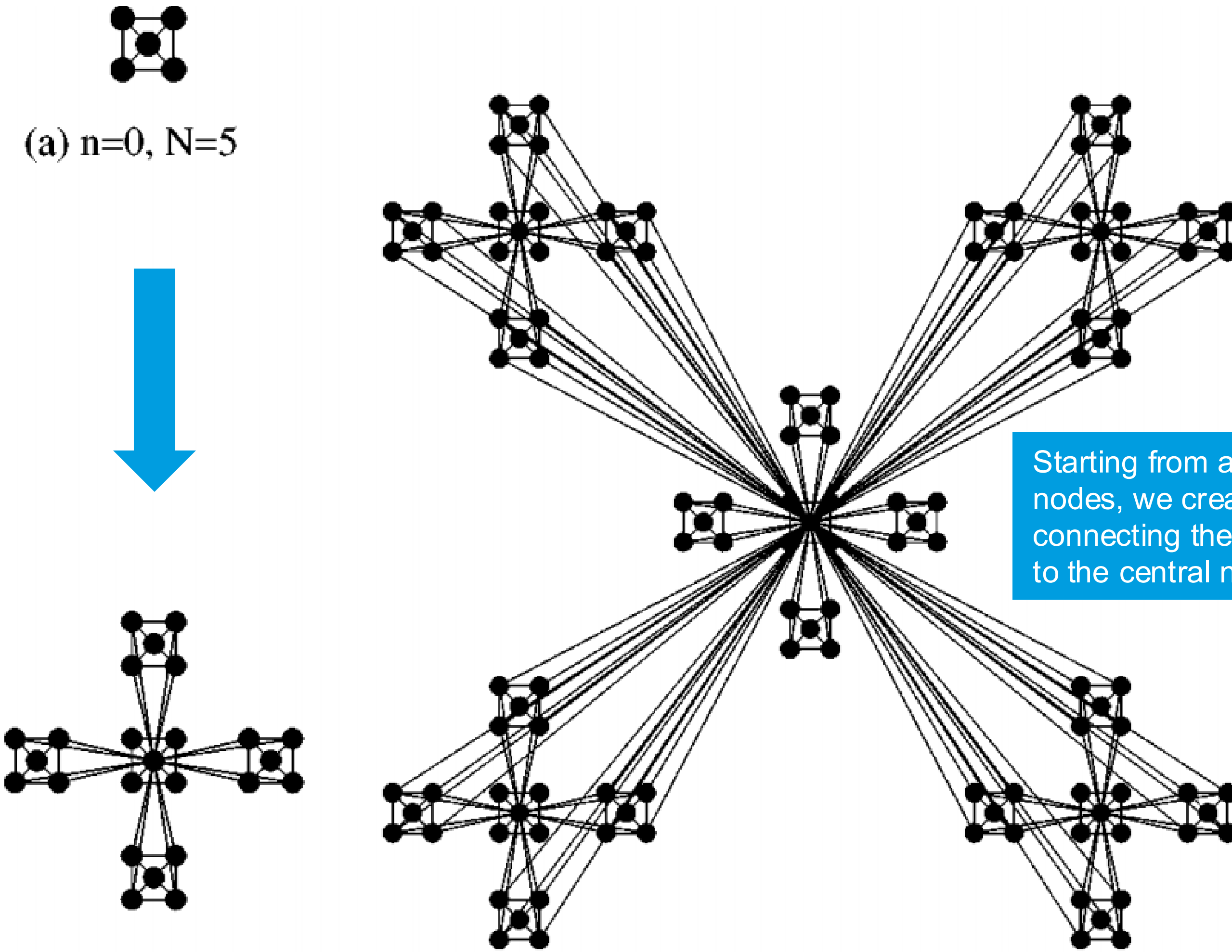


Hierarchical networks





Deterministic Models of Hierarchical Networks



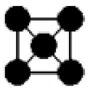
(a) n=0, N=5

(b) n=1, N=25

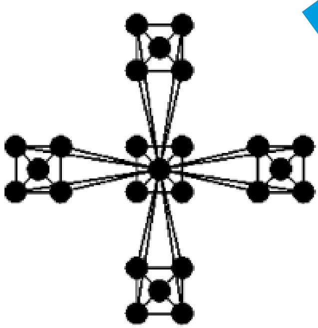
(c) n=2, N=125

Starting from a fully connected cluster of five nodes, we create four identical replicas, connecting the peripheral nodes of each cluster to the central node of the original cluster.

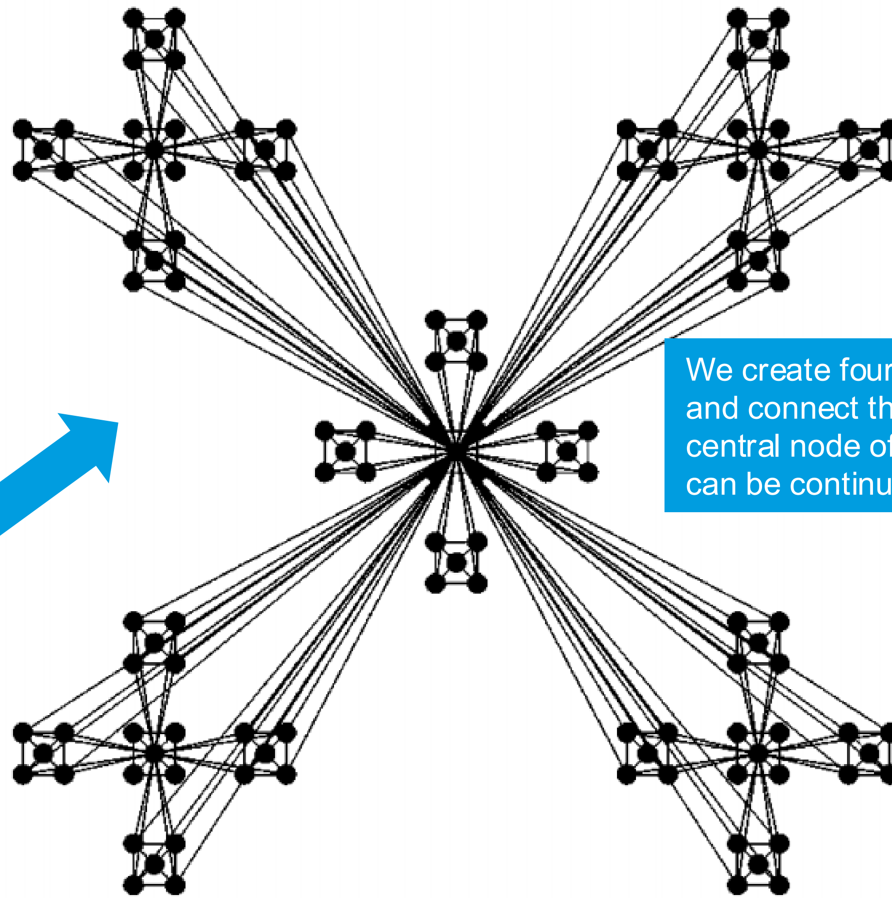
Deterministic Models of Hierarchical Networks



(a) n=0, N=5



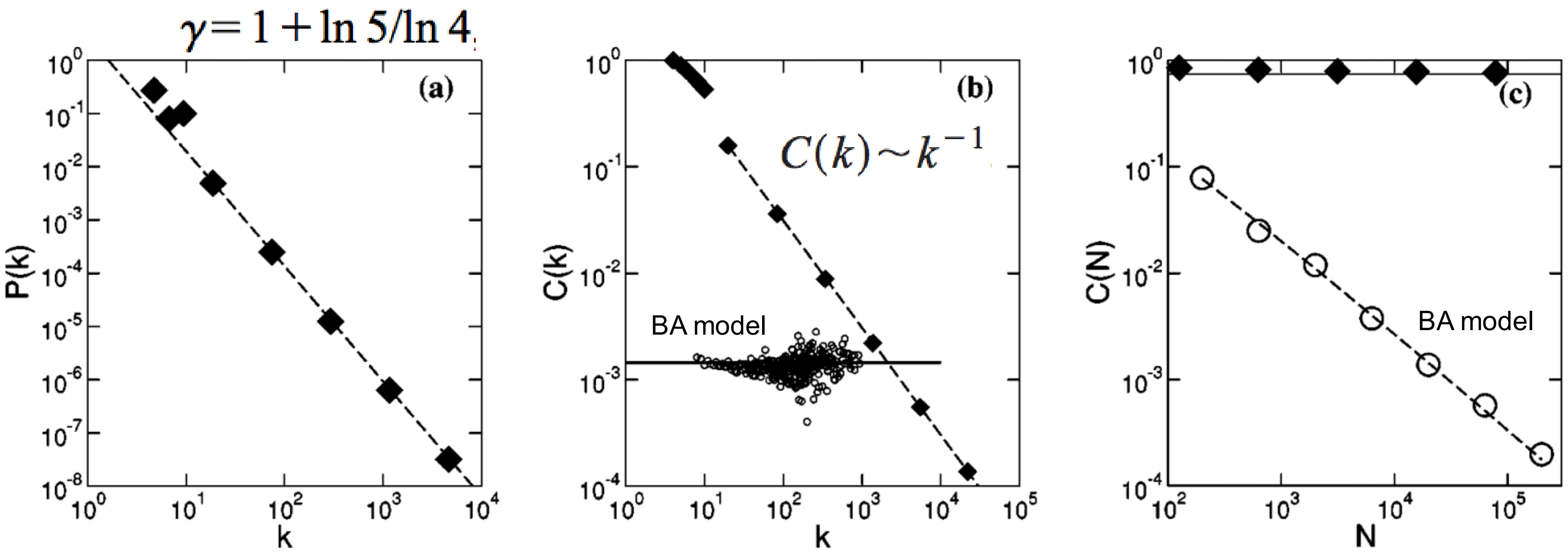
(b) n=1,
N=25



(c) n=2, N=125

We create four replicas of the obtained cluster, and connect the peripheral nodes again, to the central node of the original module. This process can be continued indefinitely

Hierarchical networks



Ravasz et al.,
Phys Rev E 67,
026112 (2003)

Duplication/copying models

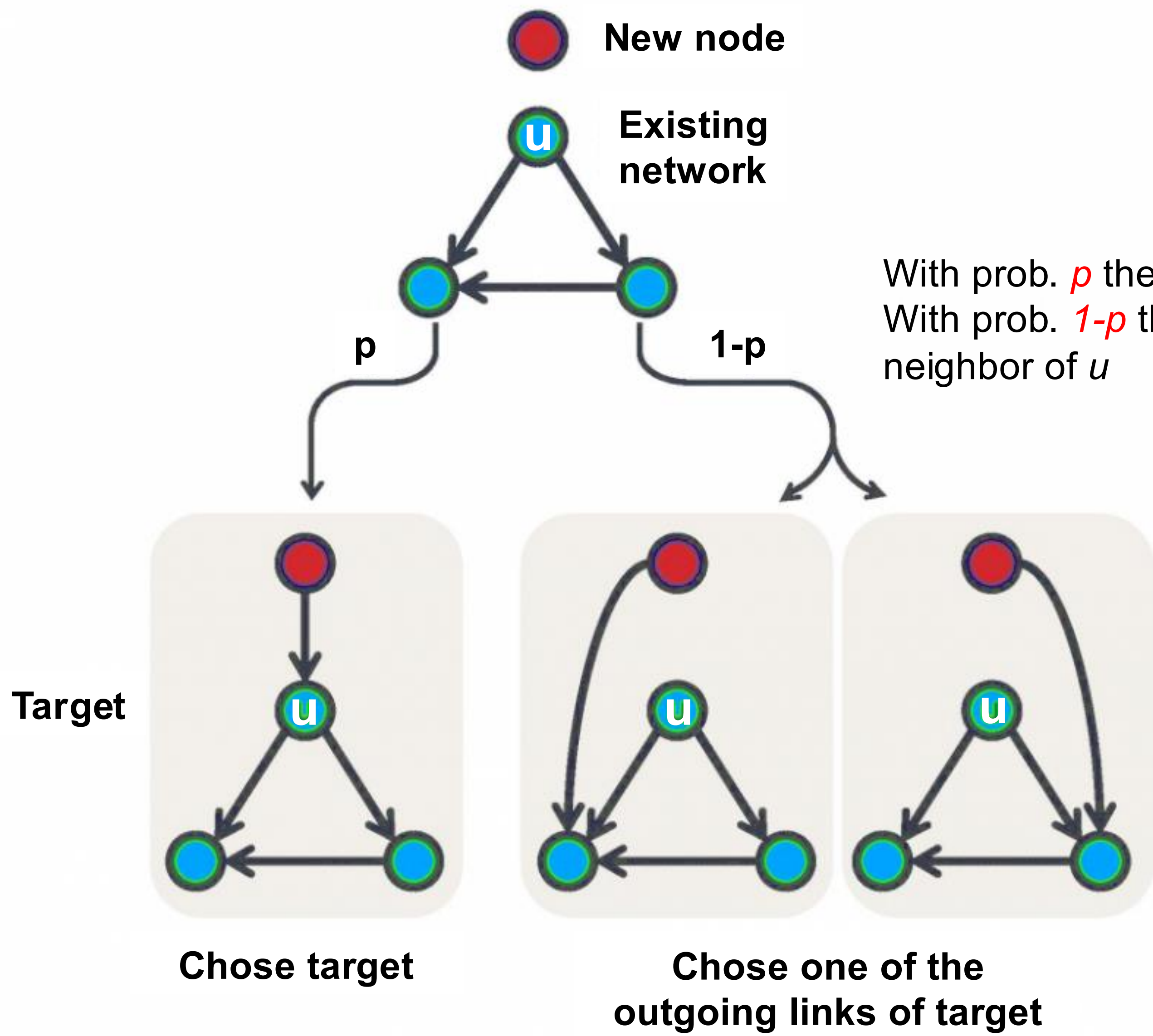
Wagner, Kleinberg, & others

Analytic results for copying models have been given by Chung et al.

Examples

- The authors of a new webpage tend to borrow links from other webpages on related topics.
- Similar arguments may be used for social networks.
- Genes that code for proteins duplicate. Since the proteins coded for by each copy are the same, their interactions are also the same, i.e., the new gene copies its edges in the interaction network from the old.
- Similar arguments have been used for Metabolic networks and other.
- Etc.

Copying model (simplest version)

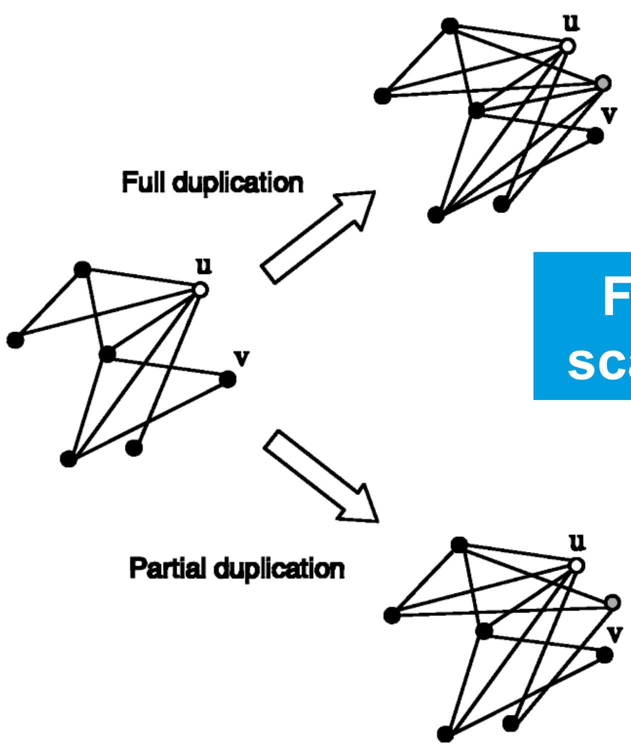


With prob. *p* the new node links to *u*
With prob. *1-p* the new node links to a neighbor of *u*

Other version: Partial duplication model

(Vazquez, Flammini, Maritan and Vespignani, 2003)

- At every time-step a randomly chosen vertex is duplicated at random creating a new node *s*.
- Each of *s* links is either kept with probability ***1−α*** or it is rewired (or removed) with probability ***α*** (equivalent to a mutation).



For low “mutation” probabilities (α), scale-free behavior (<k^2>→) emerges.

Conclusion: Understanding topological variety

Models with Pref. Attachment

Barabási-Albert Model

$\Pi \sim k$ ➔ $\gamma=3$

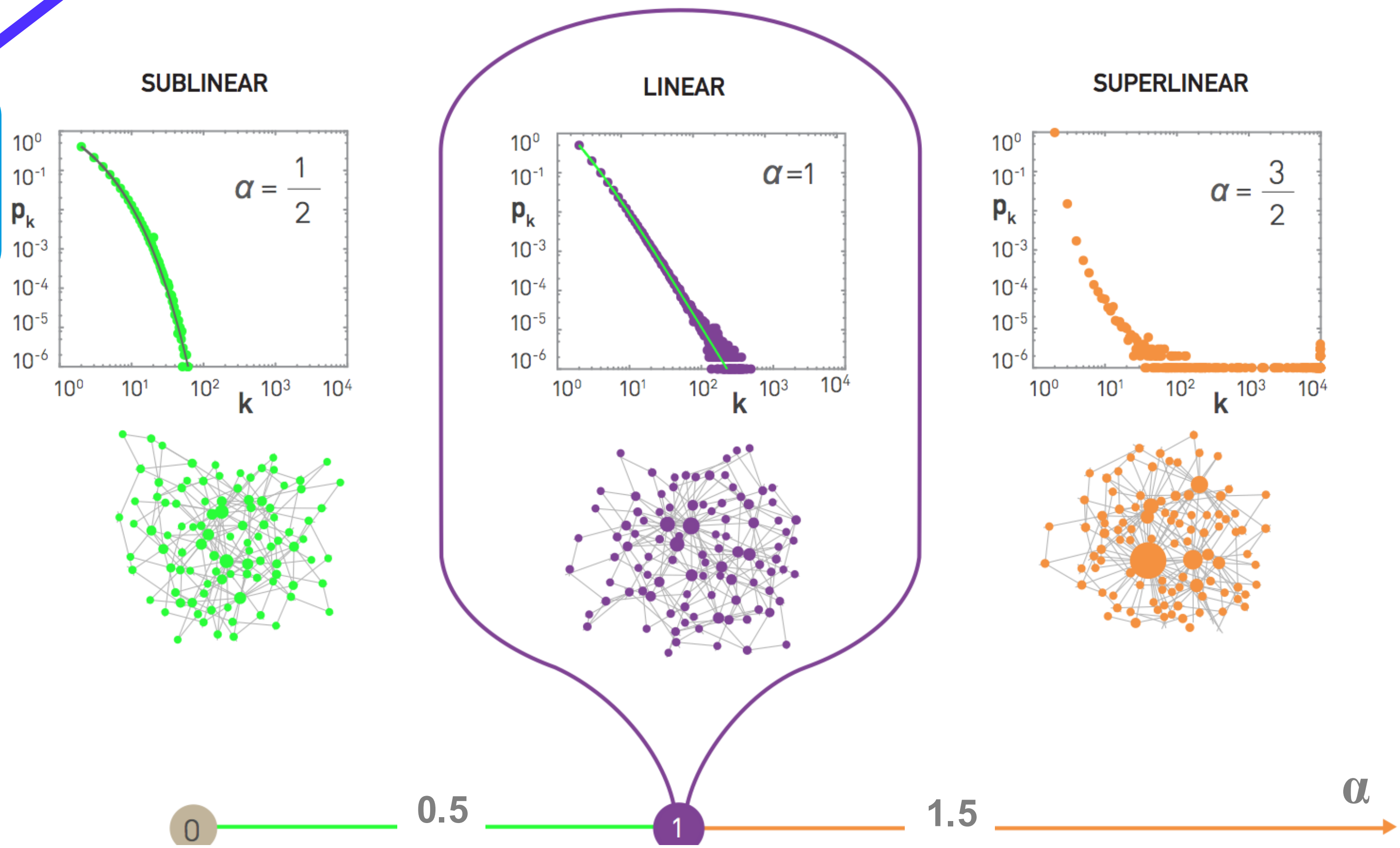
Non-linear Pref. Attachment

$$\Pi_i = \frac{k_i^{\alpha}}{\sum_j k_j^{\alpha}}$$

$\alpha<1$ ➔ exponential dist

$\alpha=1$ ➔ power-law

$\alpha>1$ ➔ winners-take all



Conclusion: Understanding topological variety

Models with Pref. Attachment

Barabási-Albert Model
Π~k ➔ γ=3

Non-linear Pref. Attachment

$$\Pi_i = \frac{k_i^\alpha}{\sum_j k_j^\alpha}$$

α<1 ➔ exponential dist
α=1 ➔ power-law
α>1 ➔ winners-take all

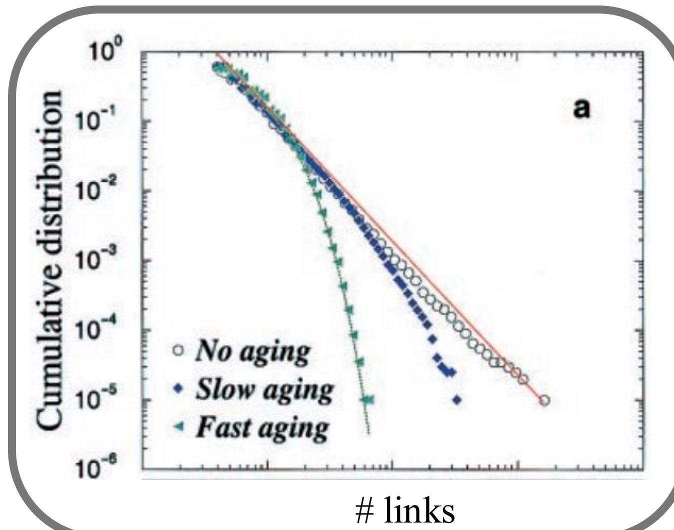
Initial attractiveness
Π~A+k ➔ γ=3+A/m

Fitness models

$$\Pi_i = \frac{\eta_i k_i}{\sum_j \eta_j k_j}$$

Ageing models

$$(t_i - t)^{-\nu}$$



Conclusion: Understanding topological variety

Models with Pref. Attachment

Barabási-Albert Model
Π~k ➔ γ=3

Non-linear Pref. Attachment

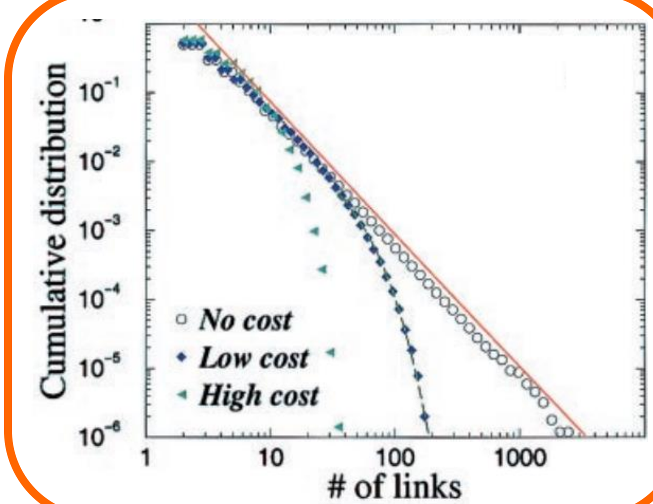
$$\Pi_i = \frac{k_i^{\alpha}}{\sum_j k_j^{\alpha}}$$

α<1 ➔ exponential dist
α=1 ➔ power-law
α>1 ➔ winners-take all

Initial attractiveness
Π~A+k ➔ γ=3+A/m

Growth w/costs

$$k > k_{\max}$$



Fitness models

$$\Pi_i = \frac{\eta_i k_i}{\sum_j \eta_j k_j}$$

Ageing models

$$(t_i - t)^{-\nu}$$

Conclusion: Understanding topological variety

Models with Pref. Attachment

Barabási-Albert Model

Π~k ➔ γ=3

Non-linear Pref. Attachment

$$\Pi_i = \frac{k_i^{\alpha}}{\sum_j k_j^{\alpha}}$$

α<1 ➔ exponential dist

α=1 ➔ power-law

α>1 ➔ winners-take all

Initial attractiveness

Π~A+k ➔ γ=3+A/m

n internal links

i) double pref. attachment

$$\gamma = 2 + \frac{m}{m+2n} < 3$$

ii) random attachment

$$\gamma = 3 + \frac{2n}{2m} > 3$$

Growth w/costs

$$k > k_{\max}$$

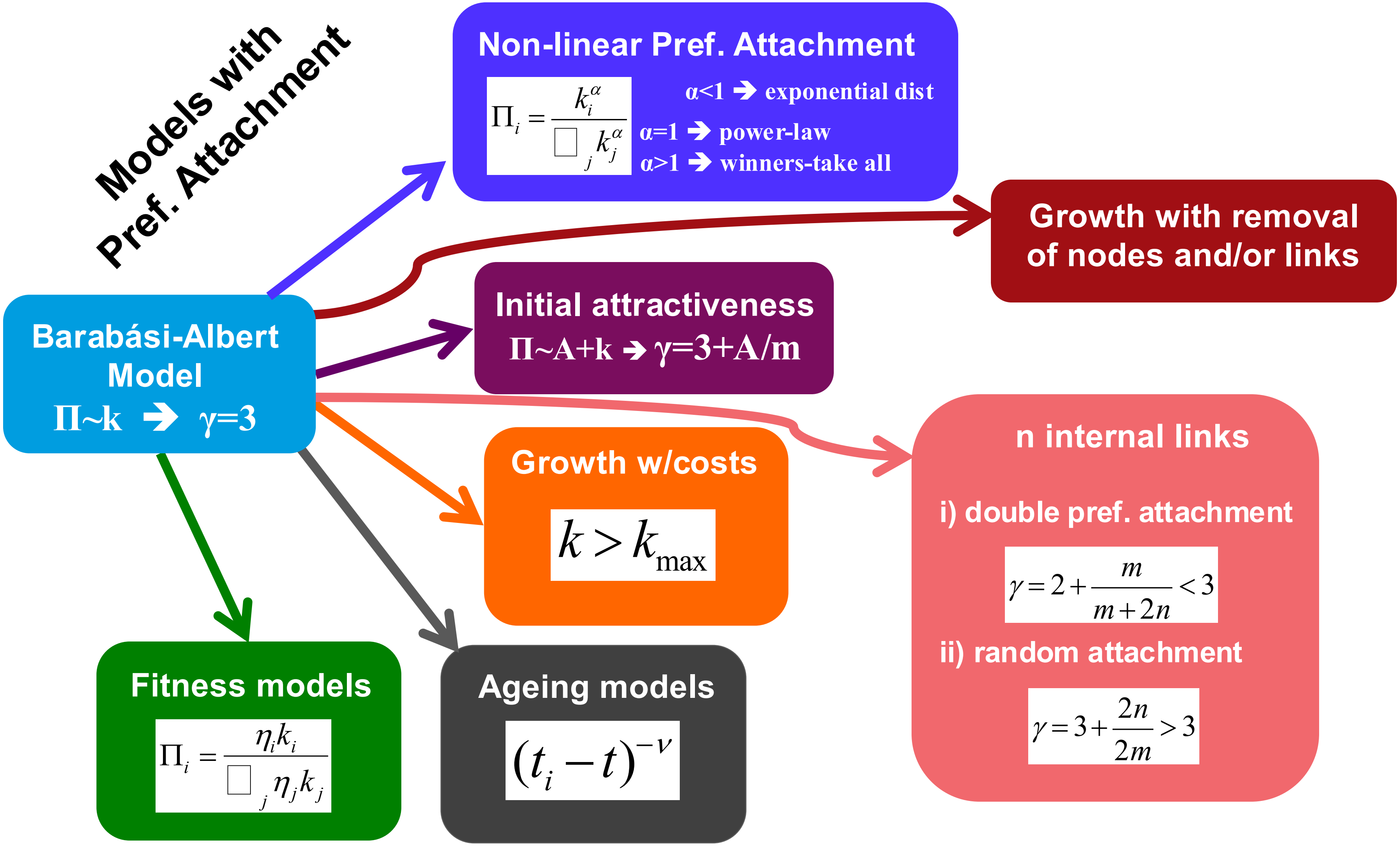
Fitness models

$$\Pi_i = \frac{\eta_i k_i}{\sum_j \eta_j k_j}$$

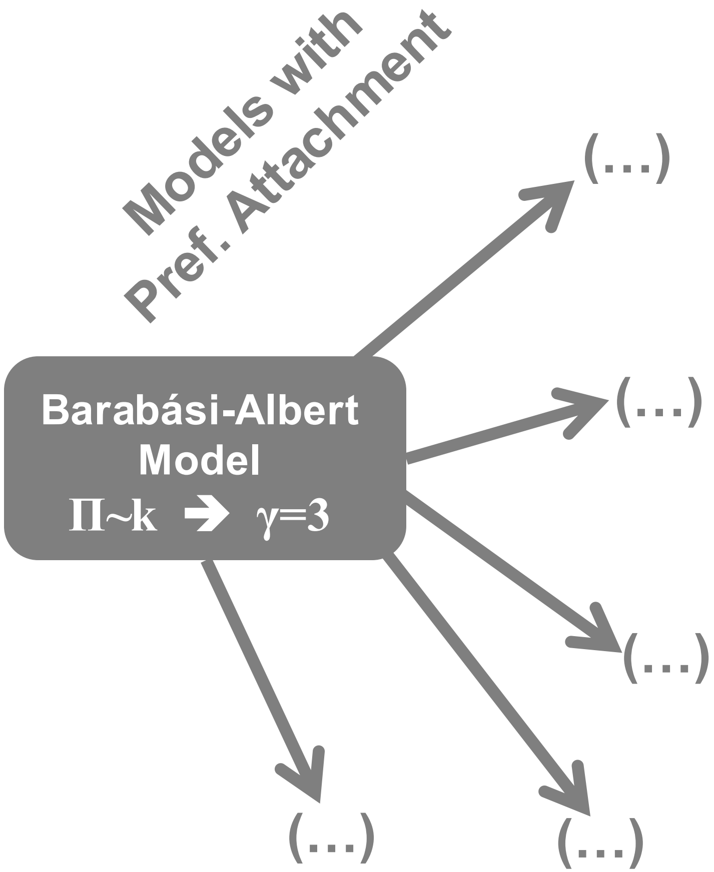
Ageing models

$$(t_i - t)^{-\nu}$$

Conclusion: Understanding topological variety



Conclusion: Understanding topological variety

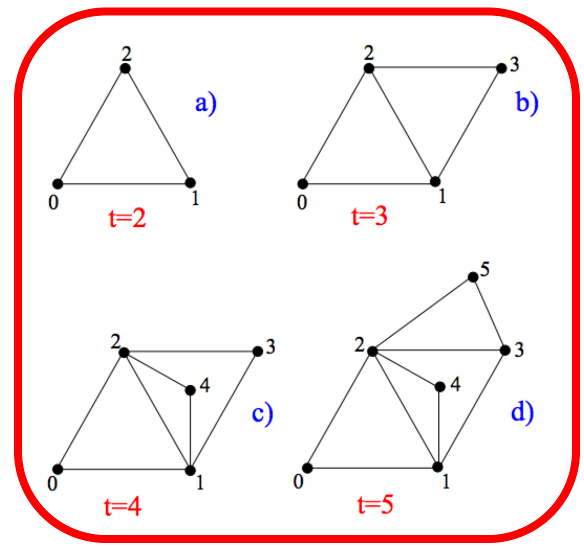


Models without Pref. Attachment?

Conclusion: Understanding topological variety

Models without Pref. Attachment

Minimal or link-selection model
$\Pi \sim k \rightarrow \gamma = 3$



Conclusion: Understanding topological variety

Models with Pref. Attachment

Barabási-Albert Model
$\Pi \sim k \rightarrow \gamma=3$

(…)

(…)

(…)

Models without Pref. Attachment

Ranking models

$$\Pi_s = \frac{R_s^{-\alpha}}{\sum_j R_j^{-\alpha}}$$

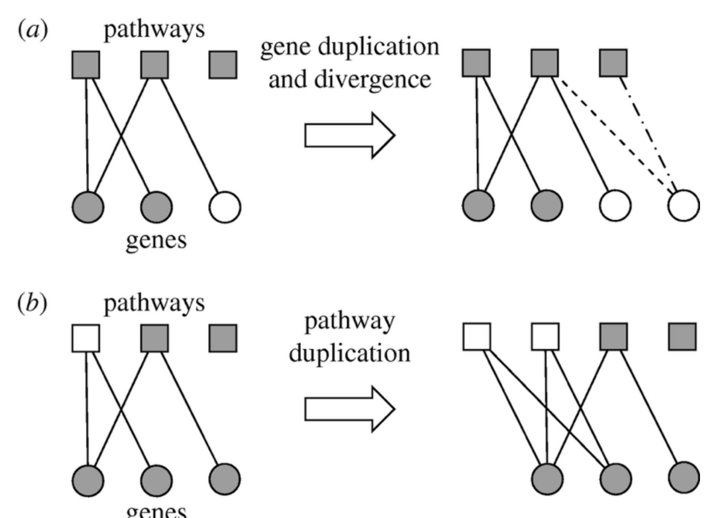
$$P(k) \sim k^{-\left(1+\frac{1}{\alpha}\right)}$$

Minimal or link-selection model
$\Pi \sim k \rightarrow \gamma=3$

Duplication models

$$\Pi_s = \frac{R_s^{-\alpha}}{\sum_j R_j^{-\alpha}}$$

$$P(k) \sim k^{-\left(1+\frac{1}{\alpha}\right)}$$



Conclusion: Understanding topological variety

Models with Pref. Attachment

Barabási-Albert Model
$\Pi \sim k \rightarrow \gamma = 3$

(…)

(…)

(…)

(…)

(…)

Models without Pref. Attachment

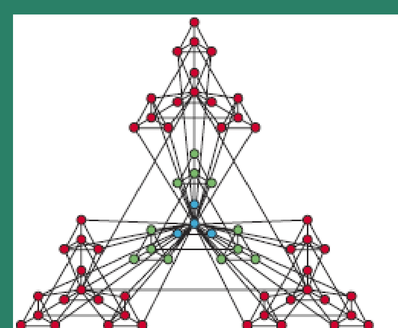
Ranking models

$$\Pi_s = \frac{R_s^{-\alpha}}{\sum_j R_j^{-\alpha}}$$

$$P(k) \sim k^{-\left(1+\frac{1}{\alpha}\right)}$$

Minimal or link-selection model
$\Pi \sim k \rightarrow \gamma = 3$

Hierarchical nets



Duplication models

$$\Pi_s = \frac{R_s^{-\alpha}}{\sum_j R_j^{-\alpha}}$$

$$P(k) \sim k^{-\left(1+\frac{1}{\alpha}\right)}$$

Conclusion: Understanding topological variety

Power-laws: BA-model, minimal model, etc, etc.

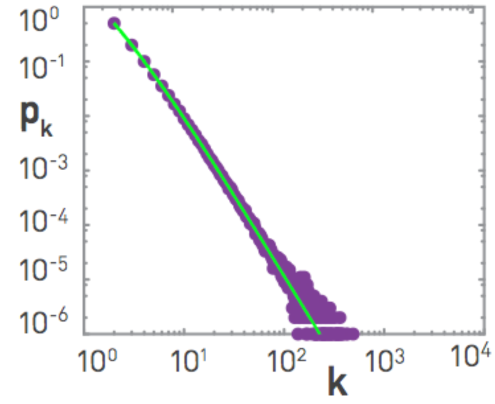
Stretched exponentials: If preferential attachment is sublinear you get a stretched exponential.

Fitness-induced corrections: Ranking models, Fitness models, initial attractiveness model.

Small-degree saturations: Initial attractiveness adds a random component to preferential attachment, particularly for low degrees.

High degree cutoffs: Node and link removal, costs and cutoffs, and node ageing, can induce high-degree cutoffs.

Hierarchical structure & power-law dep. in clustering: Minimal/link-selection model, duplication models and hierarchical networks model.



Conclusion: Understanding topological variety

Power-laws: BA-model, minimal model, etc, etc.

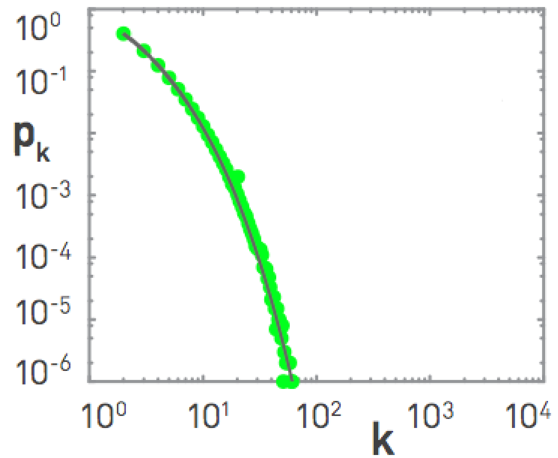
Stretched exponentials: If preferential attachment is sublinear you get a stretched exponential.

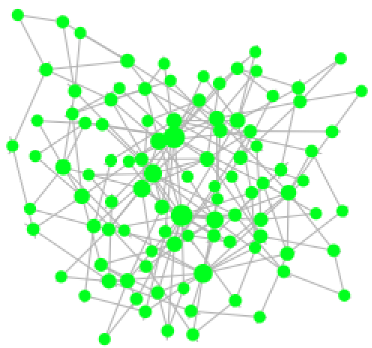
Fitness-induced corrections: Ranking models, Fitness models, initial attractiveness model.

Small-degree saturations: Initial attractiveness adds a random component to preferential attachment, particularly for low degrees.

High degree cutoffs: Node and link removal, costs and cutoffs, and node ageing, can induce high-degree cutoffs.

Hierarchical structure & power-law dep. in clustering: Minimal/link-selection model, duplication models and hierarchical networks model.





Conclusion: Understanding topological variety

Power-laws: BA-model, minimal model, etc, etc.

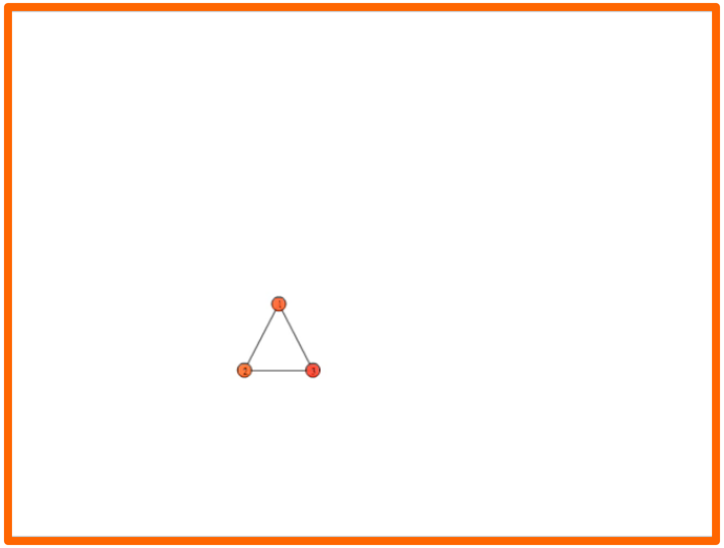
Stretched exponentials: If preferential attachment is sublinear you get a stretched exponential.

Fitness-induced corrections: Ranking models, Fitness models, initial attractiveness model.

Small-degree saturations: Initial attractiveness adds a random component to preferential attachment, particularly for low degrees.

High degree cutoffs: Node and link removal, costs and cutoffs, and node ageing, can induce high-degree cutoffs.

Hierarchical structure & power-law dep. in clustering: Minimal/link-selection model, duplication models and hierarchical networks model.



Conclusion: Understanding topological variety

Power-laws: BA-model, minimal model, etc, etc.

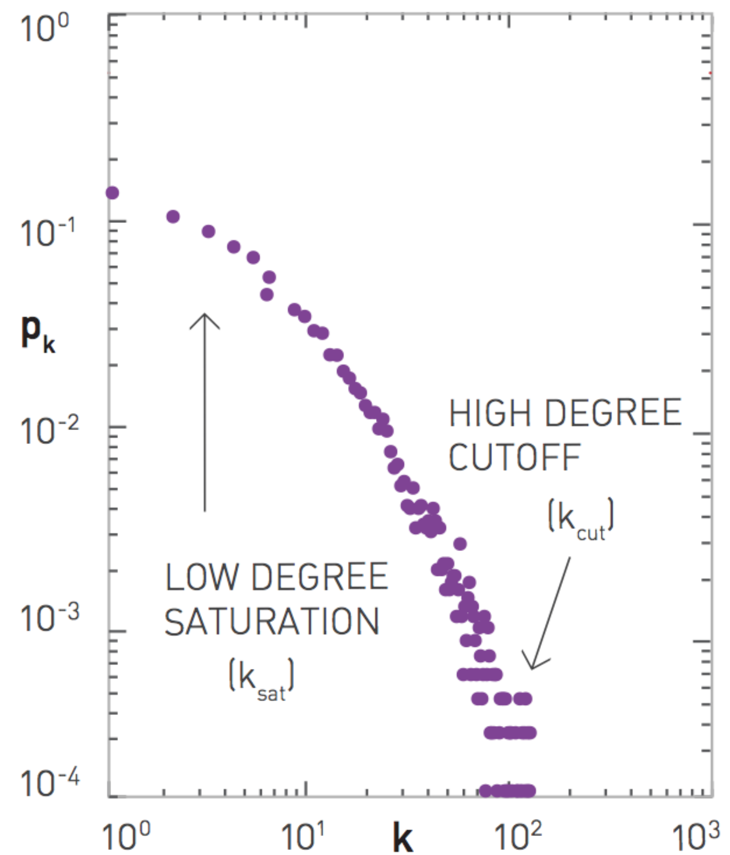
Stretched exponentials: If preferential attachment is sublinear you get a stretched exponential.

Fitness-induced corrections: Ranking models, Fitness models, initial attractiveness model.

Small-degree saturations: Initial attractiveness adds a random component to preferential attachment, particularly for low degrees.

High degree cutoffs: Node and link removal, costs and cutoffs, and node ageing, can induce high-degree cutoffs.

Hierarchical structure & power-law dep. in clustering: Minimal/link-selection model, duplication models and hierarchical networks model.



Conclusion: Understanding topological variety

Power-laws: BA-model, minimal model, etc, etc.

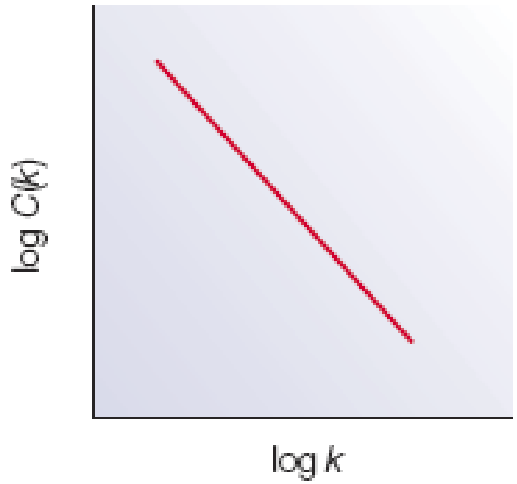
Stretched exponentials: If preferential attachment is sublinear you get a stretched exponential.

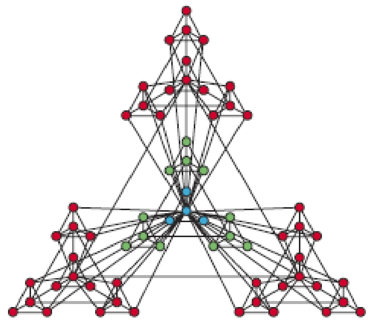
Fitness-induced corrections: Ranking models, Fitness models, initial attractiveness model.

Small-degree saturations: Initial attractiveness adds a random component to preferential attachment, particularly for low degrees.

High degree cutoffs: Node and link removal, costs and cutoffs, and node ageing, can induce high-degree cutoffs.

Hierarchical structure & power-law dep. in clustering: Minimal/link-selection model, duplication models and hierarchical networks model.





Classes of models in network science

- ***Static & generative models.*** ER model, Watts-Strogratz model, Configuration Model, etc.

- ***Evolving network models.*** BA model, Initial attractiveness model, fitness model, internal links model, node deletion model, accelerated model, aging model, costs model, minimal model, ranking model, duplication model, hierarchical networks model, etc.

Which models or principles should I consider to justify each case:

1) Network with a power-law degree distribution with $\gamma=3$.
2) Network with a power-law degree distribution with $\gamma=3$ with a significant exponential cutoff for large $k \sim k_{max}$.
3) Network with a power-law degree distribution with $\gamma=3$ and large clustering coefficient.
4) Power-law degree distribution with $\gamma<3.0$.
5) Power-law degree distribution with $\gamma>3.0$.
6) A network with exponential degree distribution.
7) A power-law degree distribution with saturation for low k.

Solutions:

1) BA model
2) BA model with costs or cutoffs
3) DMS minimal model
4) BA model with internal links chosen through preferential attachment
5) BA model with internal links chosen randomly (preferential attachment with initial attractiveness will also work)

DEI
DEPARTAMENTO
DE ENGENHARIA INFORMÁTICA
TÉCNICO LISBOA

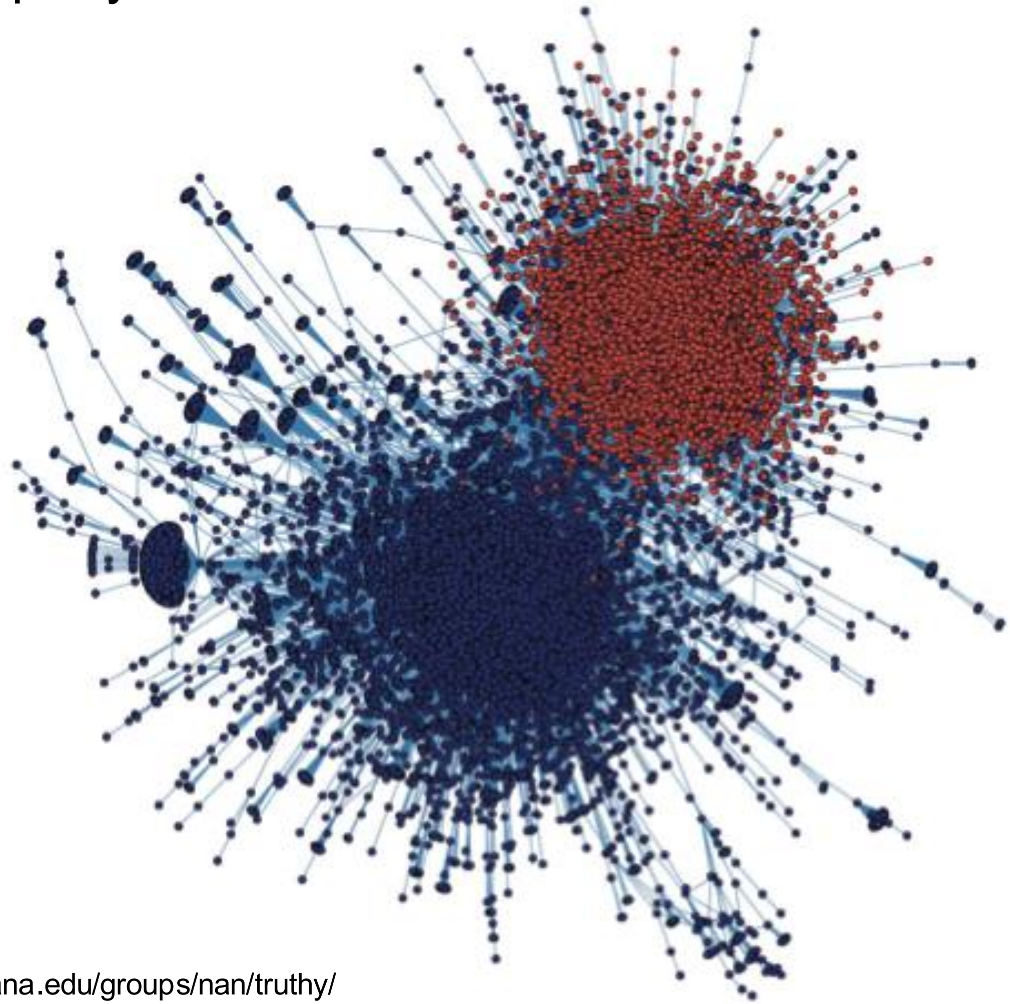
Going beyond degree distributions

Assortativity in complex networks

Network Science, 2025/2026

Political Homophily in Twitter

Assortativity = homophiliy



Network of Retweets
Political retweet network:
red: right
blue: left

Conover, Ratkiewicz, et al. 2011
Data available from: http://cnets.indiana.edu/groups/nan/truthy/

Mixing patterns

- ***Assortative mixing***: “likes link with likes”
- ***Disassortative mixing***: “likes link with dislikes”

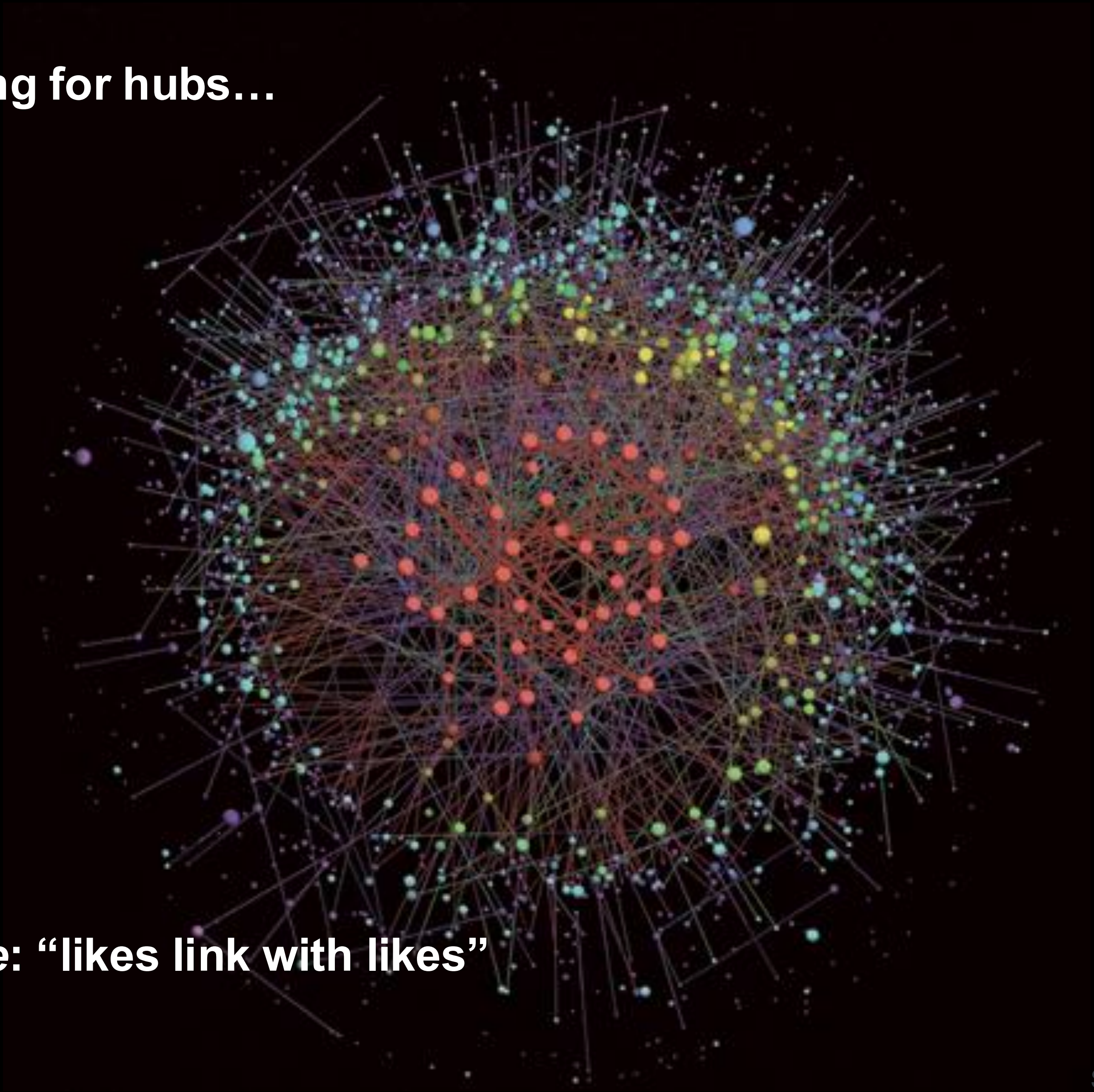
Nodes can be (dis)assortative on various attributes (age, sex, geography), topological attributes (degree, clustering, etc.).

Examples:

Assortative mixing (in social nets): political beliefs, race, obesity, etc.

Disassortative mixing: food webs (predator/prey), dating networks (gender), economic networks (producers / consumers)

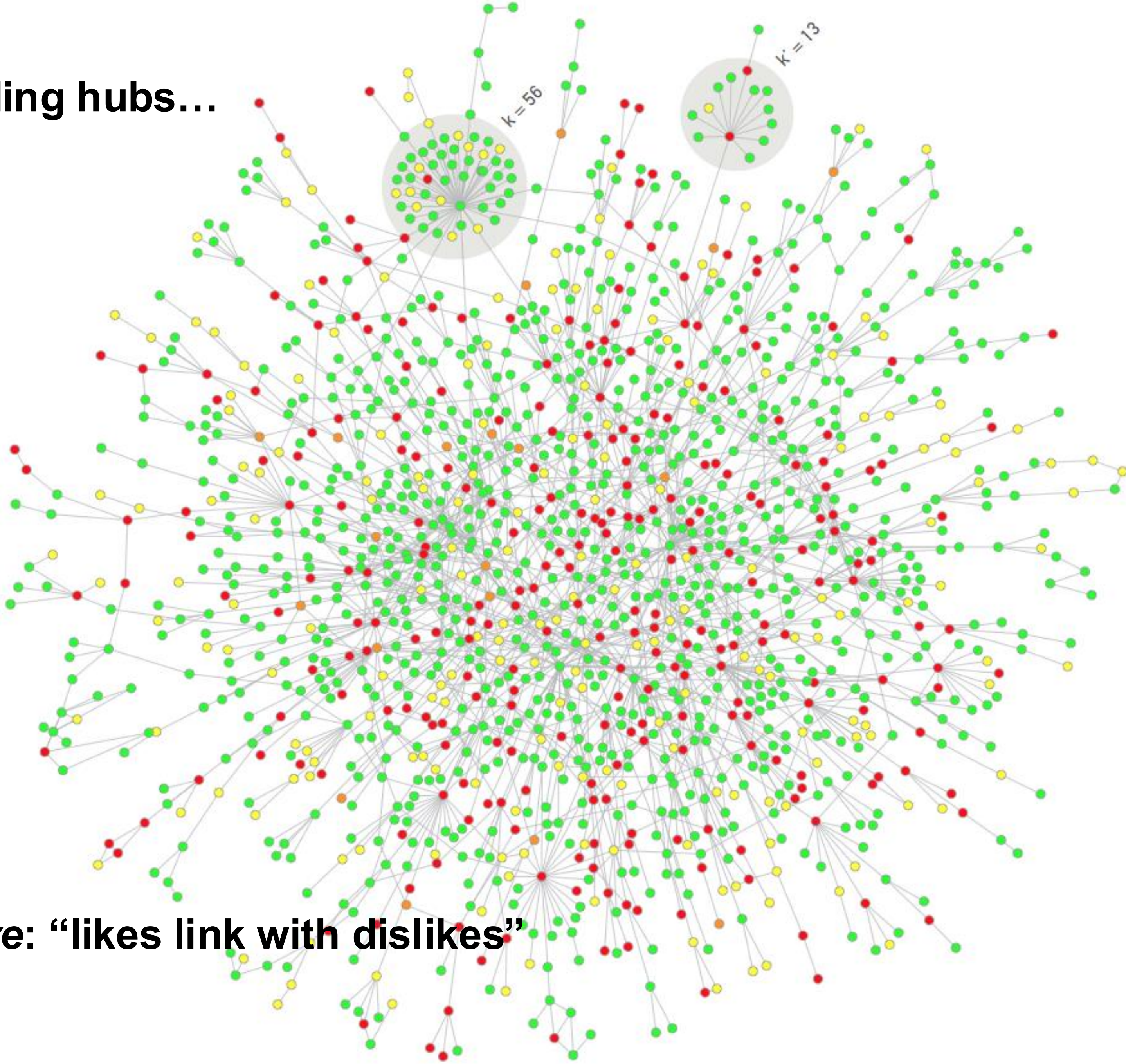
Hubs looking for hubs…



***Assortative*: "likes link with likes"**

Science, 2025/26

Hubs avoiding hubs…



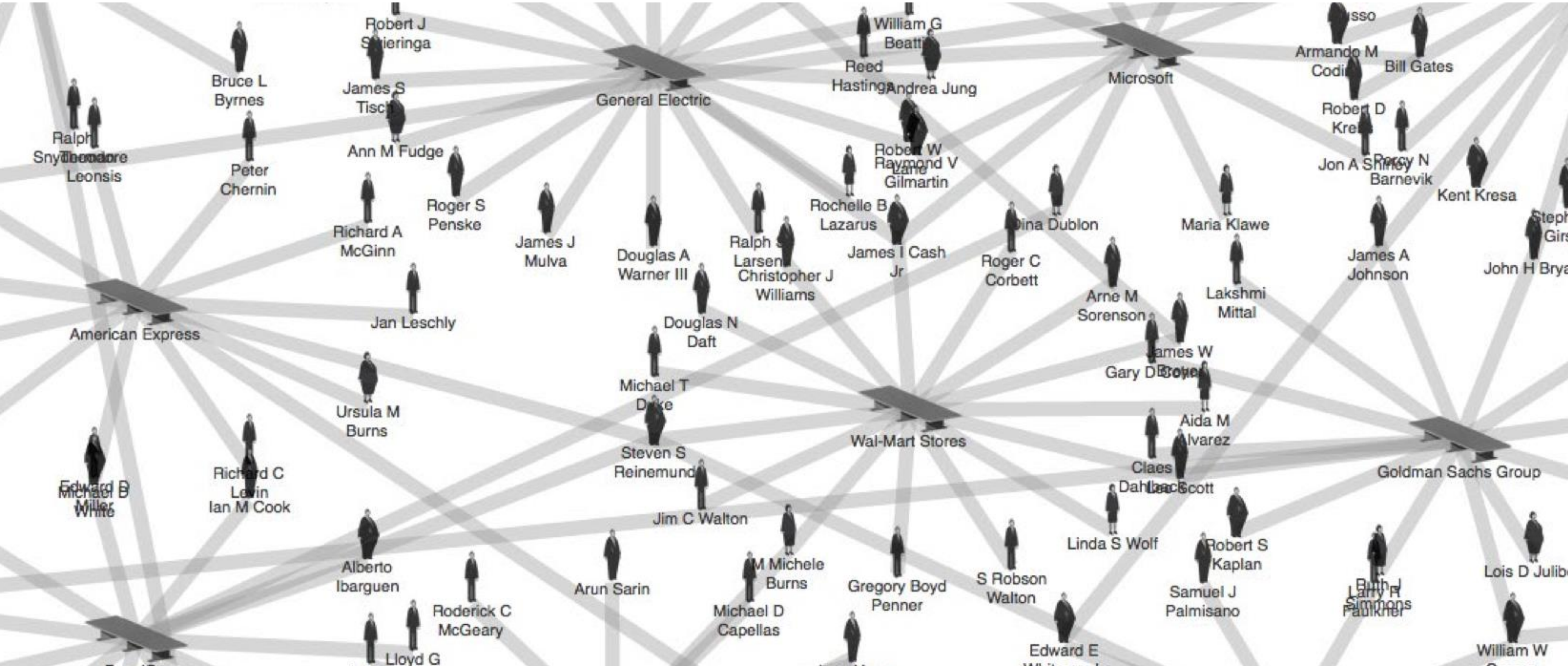
Disassortative: **“likes link with dislikes”**

Degree assortativity

- ***Social networks***: Hubs tend to date each other. If popularity is attractive, this is even clearer among the most popular.

Degree assortativity

- ***Social networks***: Hubs tend to date each other. If popularity is attractive, this is even clearer among the most popular.



Degree assortativity

- ***Social networks***: Hubs tend to date each other. If popularity is attractive, this is even clearer among the most popular.
- ***Biological & technological networks*** (e.g., PPI nets): hubs avoid linking to other hubs, connecting instead to many small degree nodes.

***Assortativity* is a preference for nodes to attach to others that are similar in some way.**
Though the specific measure of similarity may vary, network theorists often examine assortativity in terms of a node's degree

We will discuss 3 key measures

- ***Social networks***: Hubs tend to date each other. If po… m…
- ***Bi…*** hu… m…

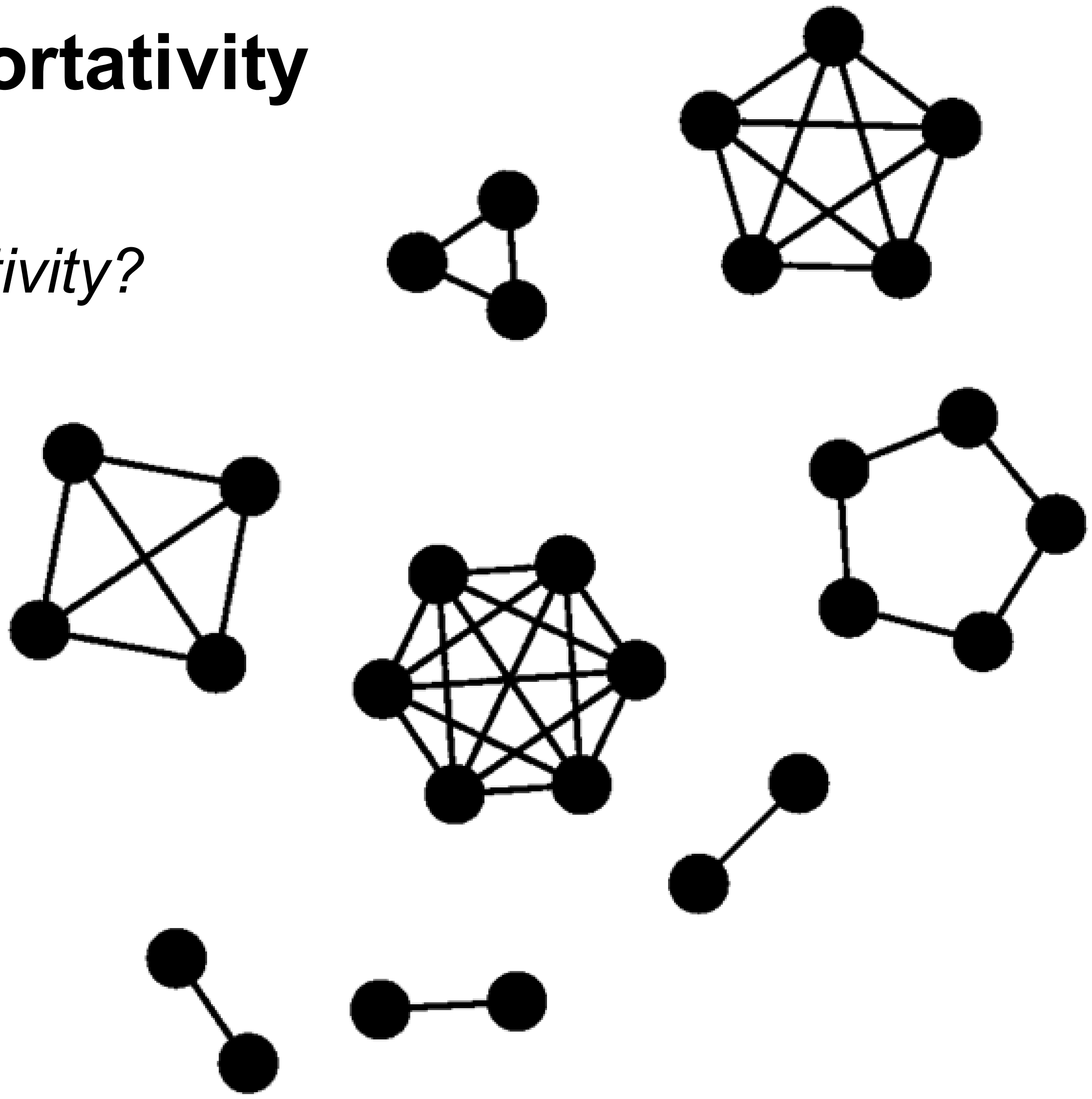
Four assortativity measures:

1. Pearson coefficient
2. Degree correlation matrix
3. Degree correlation function
4. Rich-club coefficient

Assortativity **is a preference for nodes to attach to others that are similar in some way. Though the specific measure of similarity may vary, network theorists often examine assortativity in terms of a node's degree**

Degree assortativity

Perfect assortativity?



In the absence of degree correlations…

…one should get

$$p(k_1, k_2) = \frac{k_1 \cdot k_2}{2E}$$

which, in general, fails for real-world networks.

However, it offers a neat reference point.

A general picture: measuring mixing patterns
(of any kind)

Assortativity of <u>discrete attributes</u> c_i (e.g., color, gender, modularity);

The key question for any assortativity measure:
How much more often do attributes match across edges than expected at random?

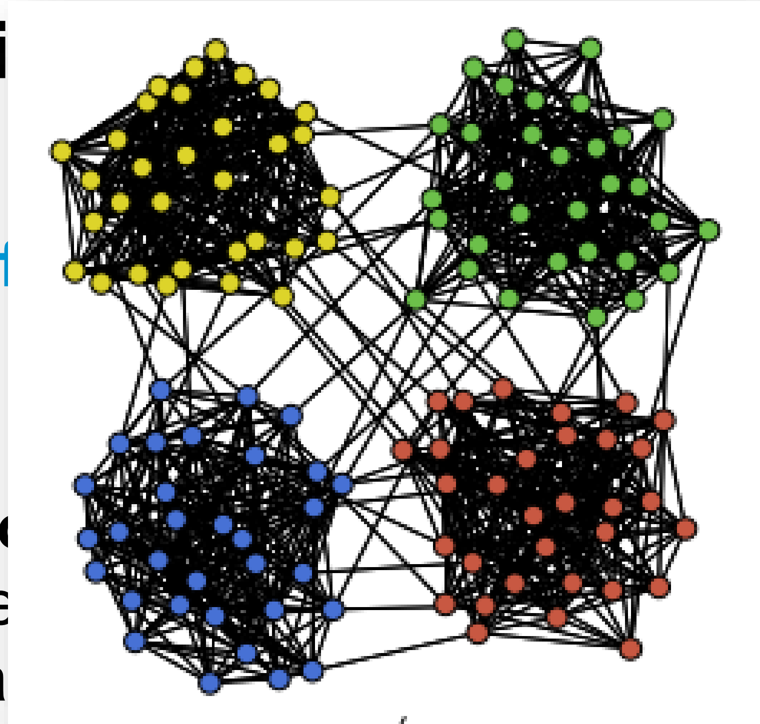
$$M = \frac{E_{same-label} - \langle E_{same-label} \rangle_{rand}}{E} = \frac{1}{2E} \sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2E} \right) \delta(c_i, c_j)$$

$E_{same-label}$=number of edges between vertices of the same label
$\langle E_{same-label} \rangle$=expected number of edges between vertices of the same label in a randomized network
A_{ij} = adj. matrix
c_i, c_j = labels of the class, $\delta(c_i, c_j)=\delta_{ij}=1$ iff $c_i=c_j$; 0, otherwise.

A general pi… …ixing patterns
(of any kind)

Assortativity o… …g., color, gender, modularity);

The key questi… …easure:
How much more … …across edges than expected a…



$$M=\frac{E_{same-label}-\left\langle E_{same-label}\right\rangle_{rand}}{E}=\frac{1}{2E}\sum_{ij}\left(A_{ij}-\frac{k_ik_j}{2E}\right)\delta(c_i,c_j)$$

$E_{same-label}$=number of edges between vertices of the same label
$<E_{same-label}>$=expected number of edges between vertices of the same label in a randomized network
A_{ij} = adj. matrix
c_i, c_j = labels of the class, $\delta(c_i, c_j)=\delta_{ij}=1$ iff $c_i=c_j$; 0, otherwise.

A general picture: measuring mixing patterns
(of any kind)

Assortativity of discrete attributes c_i (e.g., color, gender, modularity);

The key question for any assortativity measure:
How much more often do attributes match across edges than expected at random?

$$M = \frac{Q}{Q_{\max}} = \frac{\sum_{ij}\left(A_{ij} - \frac{k_i k_j}{2E}\right)\delta(c_i, c_j)}{2E - \sum_{ij}\frac{k_i k_j}{2E}\delta(c_i, c_j)}$$

A_{ij} = adj. matrix
c_i, c_j = labels of the class, $\delta(c_i, c_j)=\delta_{ij}=1$ iff $c_i=c_j$; 0, otherwise.

A general picture: measuring mixing patterns
(of any kind)

Assortativity of scalar attributes x_i (e.g., age, income, degree, clustering, etc.);

Again, how much more often do attributes match across edges than expected at random? Correlation of values across edges, gives the so called Assortativity coeff., also known as the Pearson Coeff.:

covariance is a measure of the joint variability of two random variables.

$$r = \frac{\text{covariance}}{\text{variance}}$$

If the greater values of one variable mainly correspond with the greater values of the other variable, and the same holds for the lesser values, (i.e., the variables tend to show similar behavior), the covariance is positive

A general picture: measuring mixing patterns

(of any kind) across edges

Assortativity of scalar attributes x_i (e.g., age, income, degree, clustering, etc.);

Again, how much more often do attributes match across edges than expected at random? Correlation of values across edges, gives the so called Assortativity coeff., also known as the Pearson Coeff.:

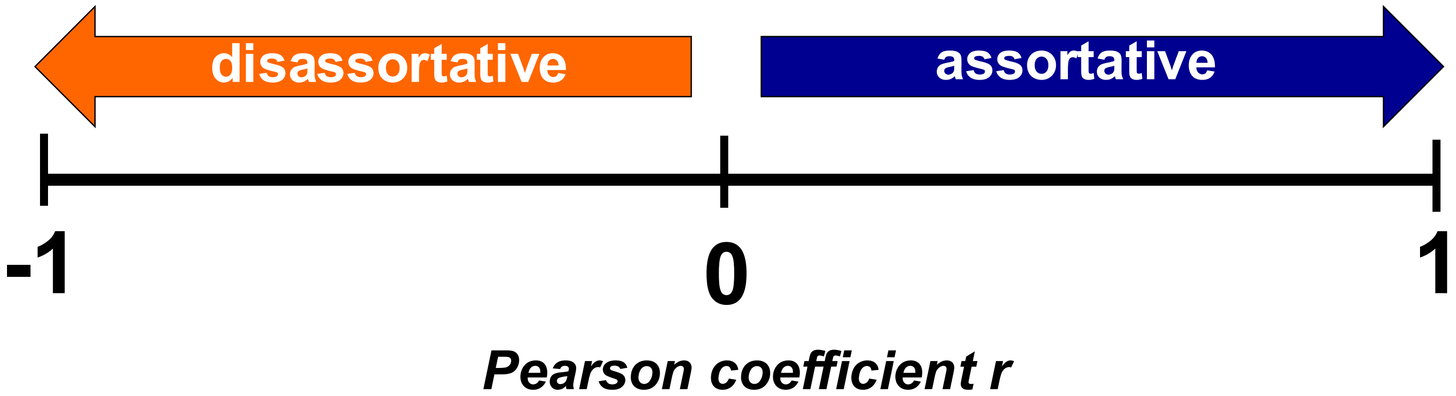
covariance is a measure of the joint variability of two random variables.

(if one increases the other increases as well, or the opposite)

$$r = \frac{\text{cov}}{\text{var}} = \frac{\sum_{ij}\left(A_{ij} - \frac{k_i k_j}{2E}\right)x_i x_j}{\sum_{ij}\left(k_i \delta_{ij} - \frac{k_i k_j}{2E}\right)x_i x_j}$$

= Traditional correlation between numbers but we only look at the correlations between numbers associated with nodes that are connected, and compare with what you would obtain in the case of random graph in the absence of correlations.

Pearson correlation coefficient r



Degree assortativity w/ Pearson correlation coefficient

$$x_i \square \; k_i \square$$

$$r = \frac{\sum_{ij}\left(A_{ij} - \frac{k_i k_j}{2E}\right) k_i k_j}{\sum_{ij}\left(k_i \delta_{ij} - \frac{k_i k_j}{2E}\right) k_i k_j} = \frac{S_A S_1 - S_2^2}{S_3 S_1 - S_2^2}$$

$$S_1 = \sum_i k_i = 2E$$

$$S_2 = \sum_i k_i^2$$

$$S_3 = \sum_i k_i^3$$

$$S_A = \sum_{ij} A_{ij} k_i k_j$$

- Positive values of ***r*** ([-1,1]) indicate a correlation between nodes of the same degree, while negative values indicate relationships between nodes of different degree.

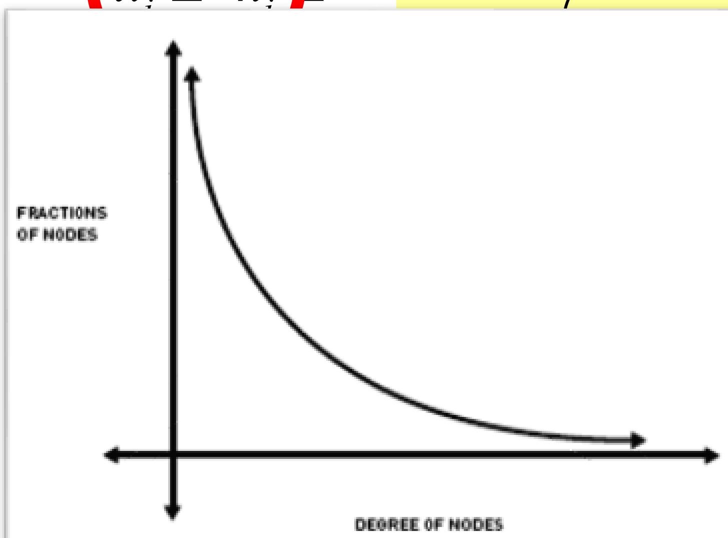
M. E. J. Newman. Assortative mixing in networks. Phys. Rev. Lett., 89: 208701, 2002.
M. E. J. Newman. Mixing patterns in networks. Phys. Rev. E, 67: 026126, 2003.

Degree assortativity w/ Pearson correlation coefficient

x_i □ k_i □

$$r = \frac{\sum_{ij}\left(A_{ij} - \frac{k_i k_j}{2E}\right)k_i k_j}{\ldots \frac{k_i k_j}{2E}\Big) k_i k_j} = \frac{S_A S_1 - S_2^2}{S_3 S_1 - S_2^2}$$

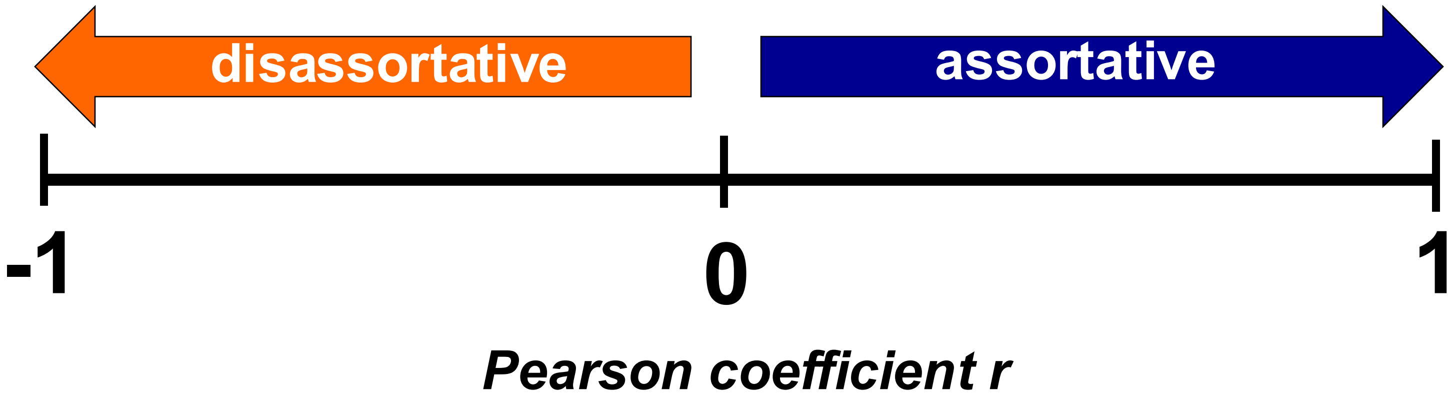
$$S_1 = \sum_i k_i = 2E$$
$$S_2 = \sum_i k_i^2$$
$$S_3 = \sum_i k_i^3$$
$$S_A = \sum_{ij} A_{ij} k_i k_j$$



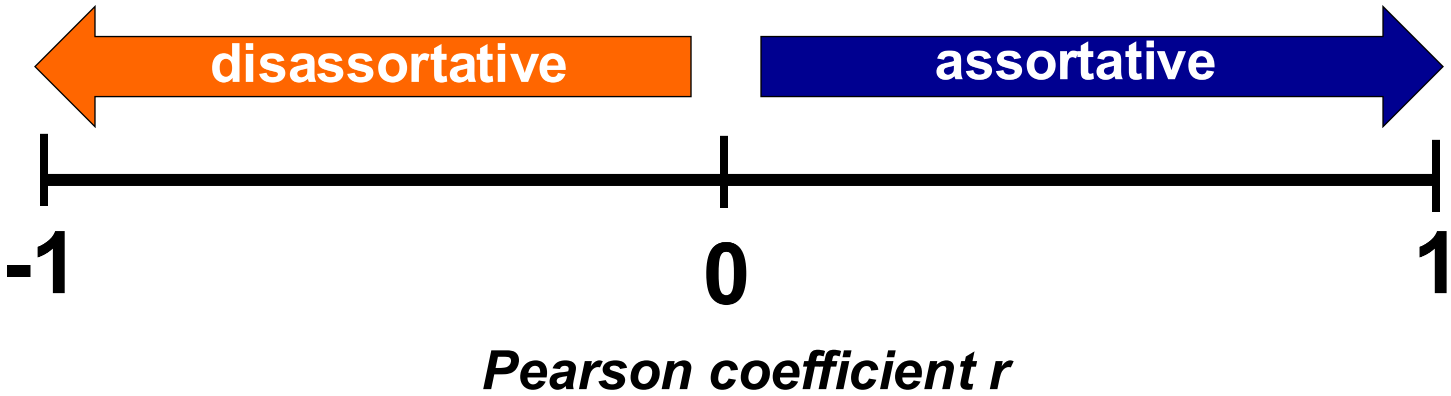
-1,1]) indicate a correlation
same degree, while negative
onships between nodes of different
degree.

- Gives a larger weight to the more abundant degree classes, which in many cases might not express the variations of the correlation function behavior.

Pearson correlation coefficient r



Pearson correlation coefficient r



R with igraph:

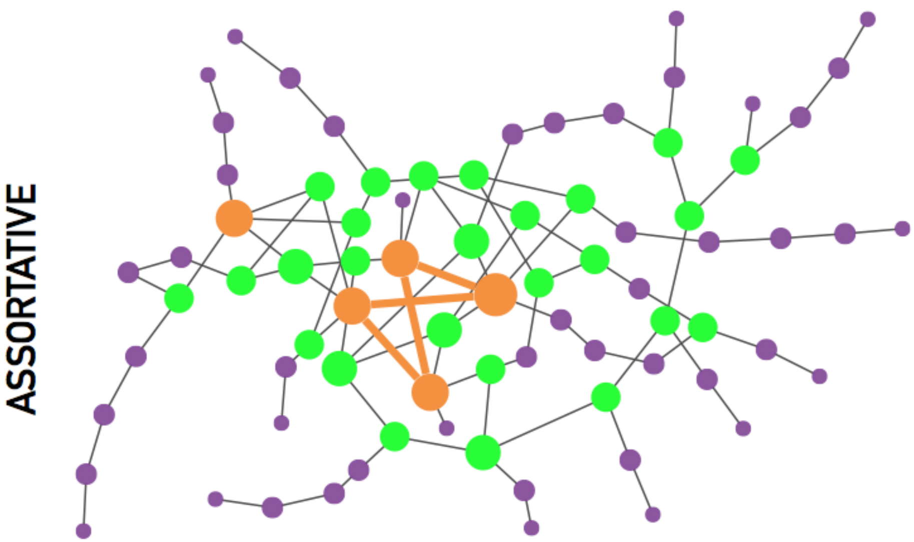
```
assortativity.degree()
```

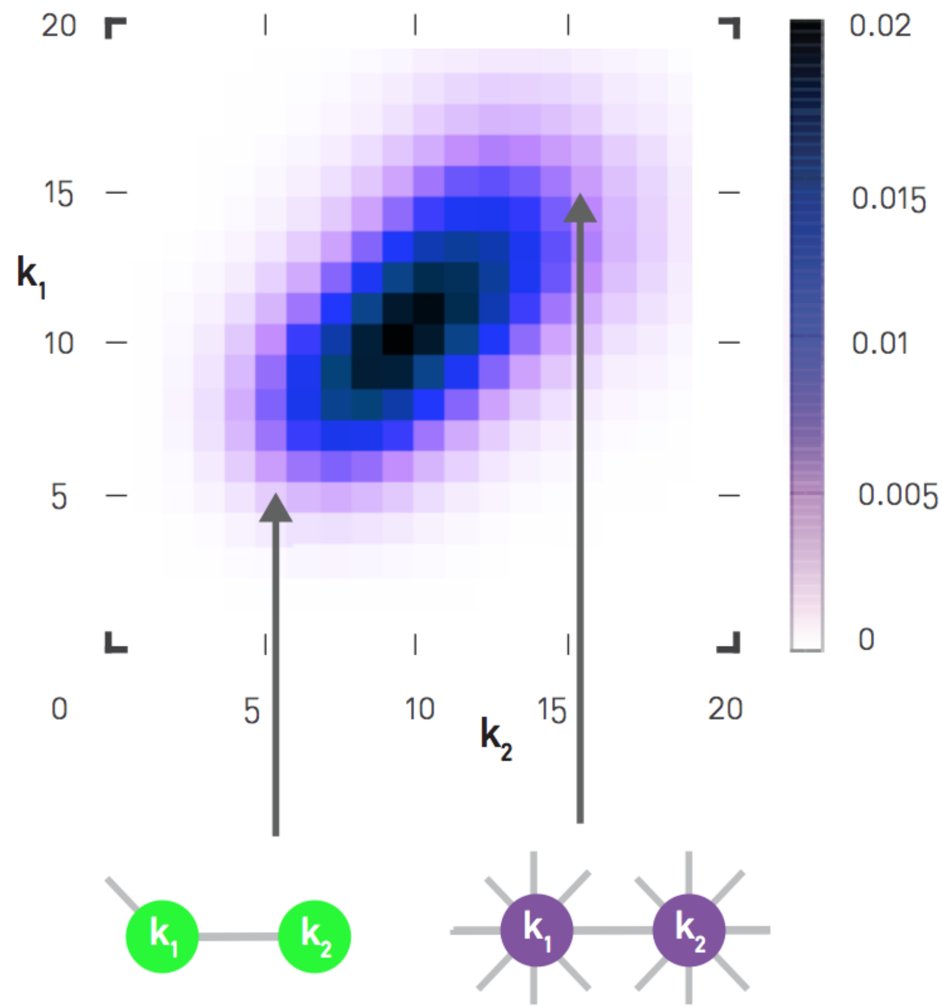
Assortativity coefficient or the Pearson correlation coefficient

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
Barabási and Albert (w)		0

Degree correlation matrix e_{ij}

Probability of finding a node with degrees i and j
at the two ends of a randomly selected link.



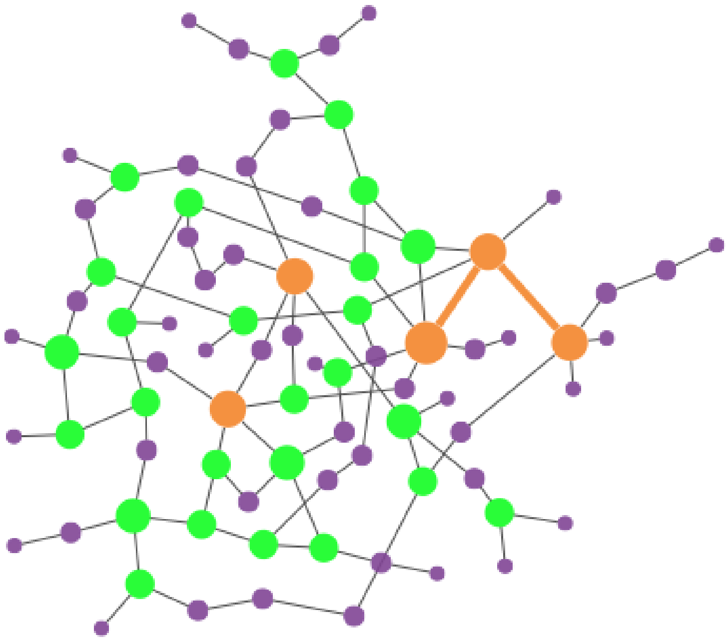


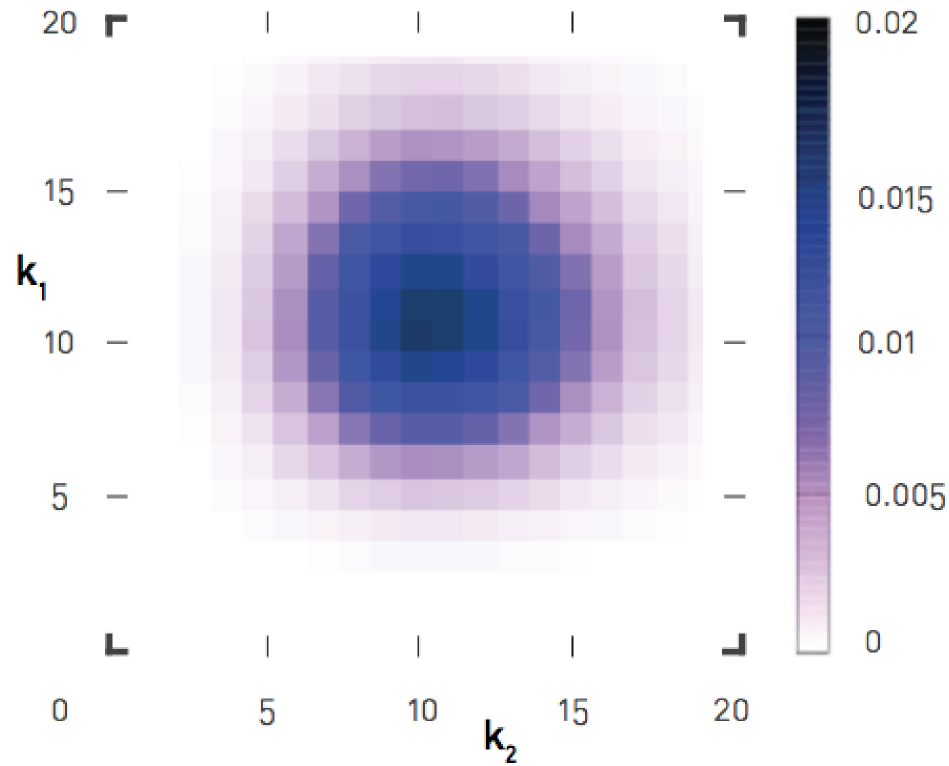
Nodes with comparable degrees tend to link to each other.

Degree correlation matrix e_{ij}

Probability of finding a node with degrees i and j
at the two ends of a randomly selected link.

NEUTRAL



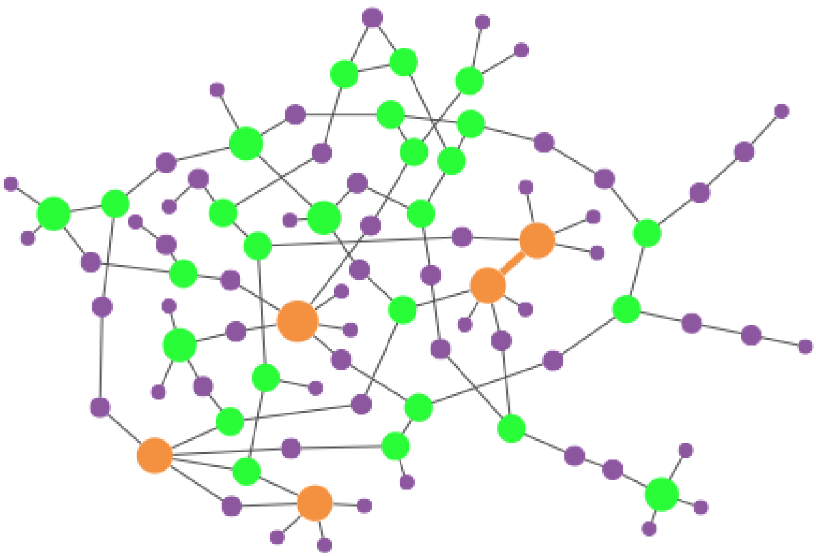


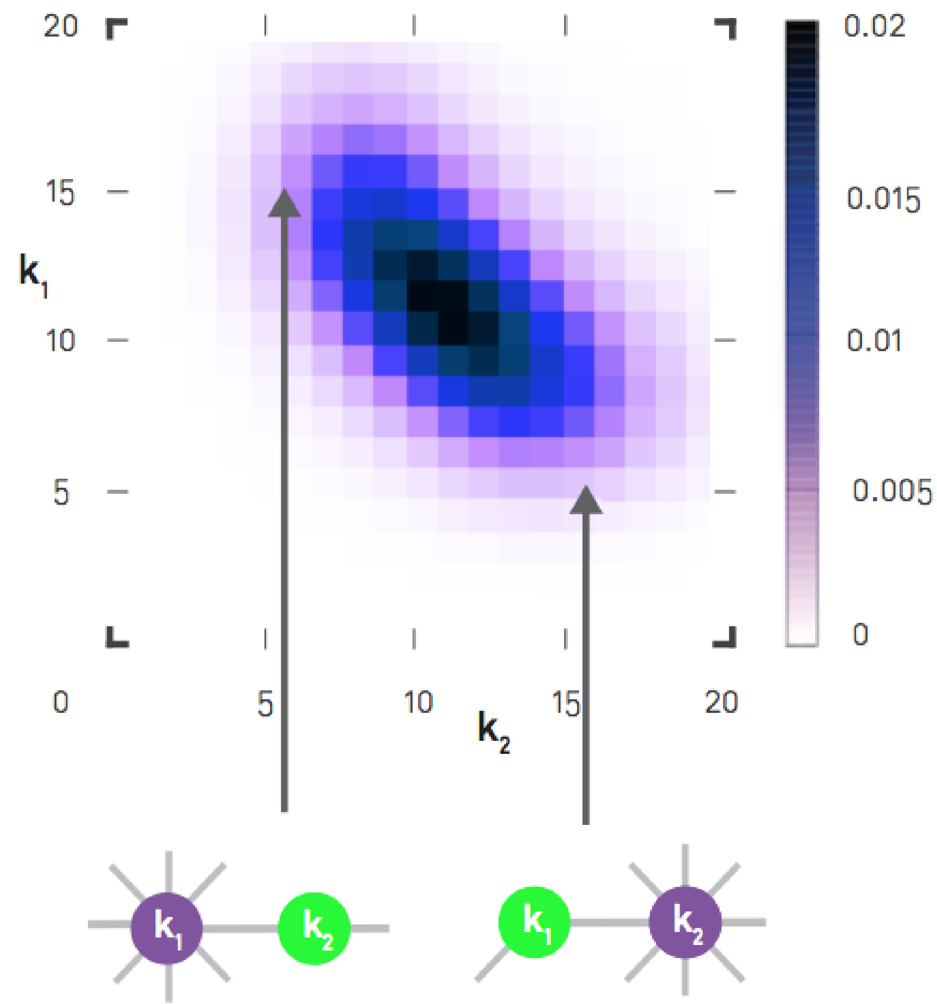
Density of links is symmetric around the average degree, indicating a lack of correlations.

Degree correlation matrix e_{ij}

Probability of finding a node with degrees i and j
at the two ends of a randomly selected link.

DISASSORTATIVE

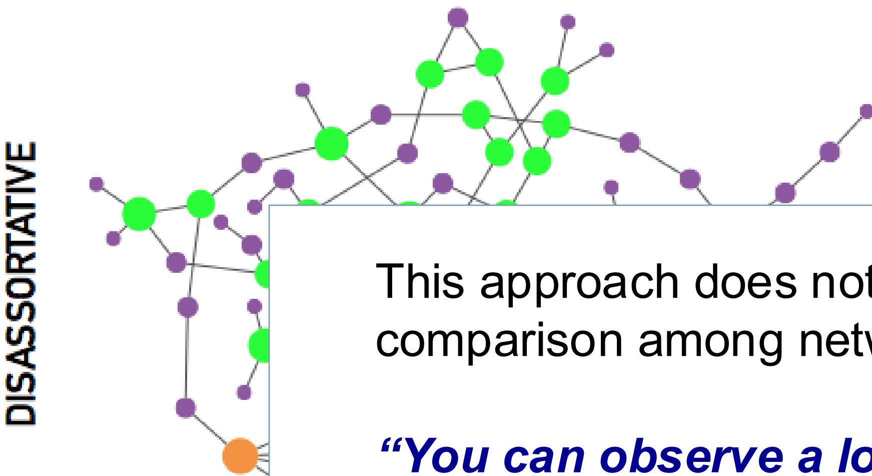


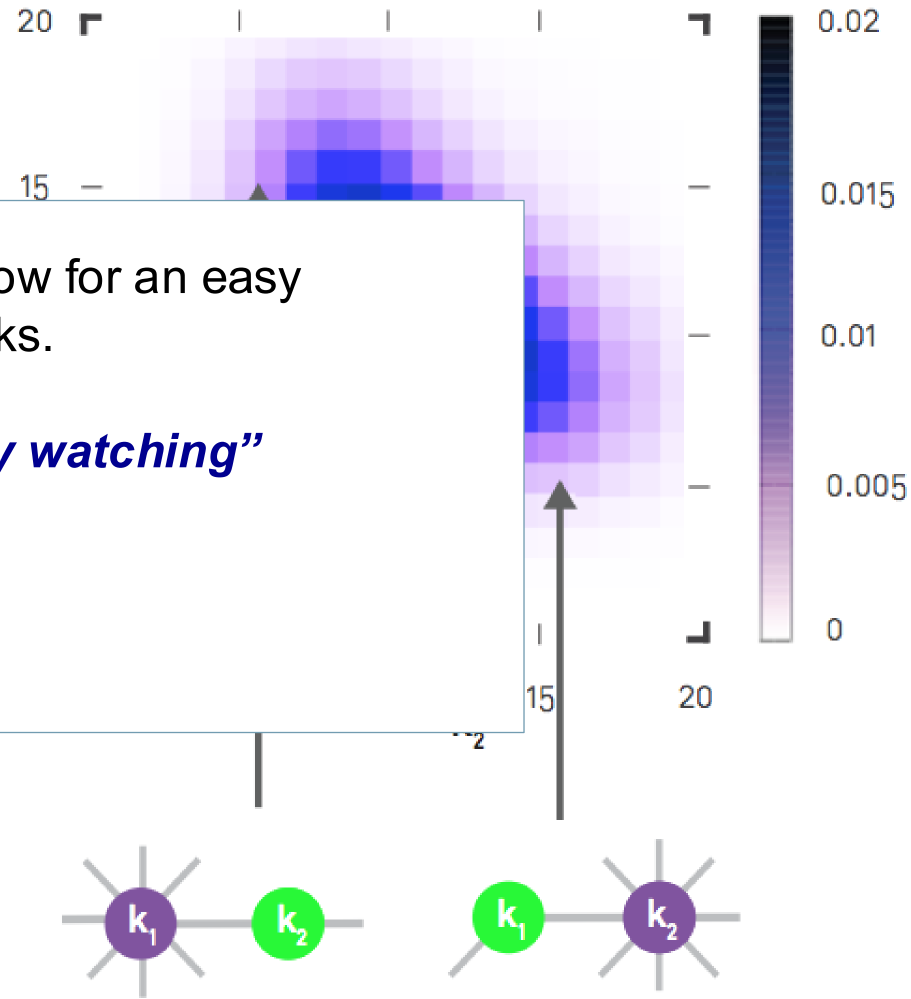


Hubs tend to connect to small-degree nodes and small-degree nodes to hubs.

Degree correlation matrix e_{ij}

Probability of finding a node with degrees i and j
at the two ends of a randomly selected link.





This approach does not allow for an easy comparison among networks.

***"You can observe a lot by watching"* (Yogi Berra)**,

but not always…

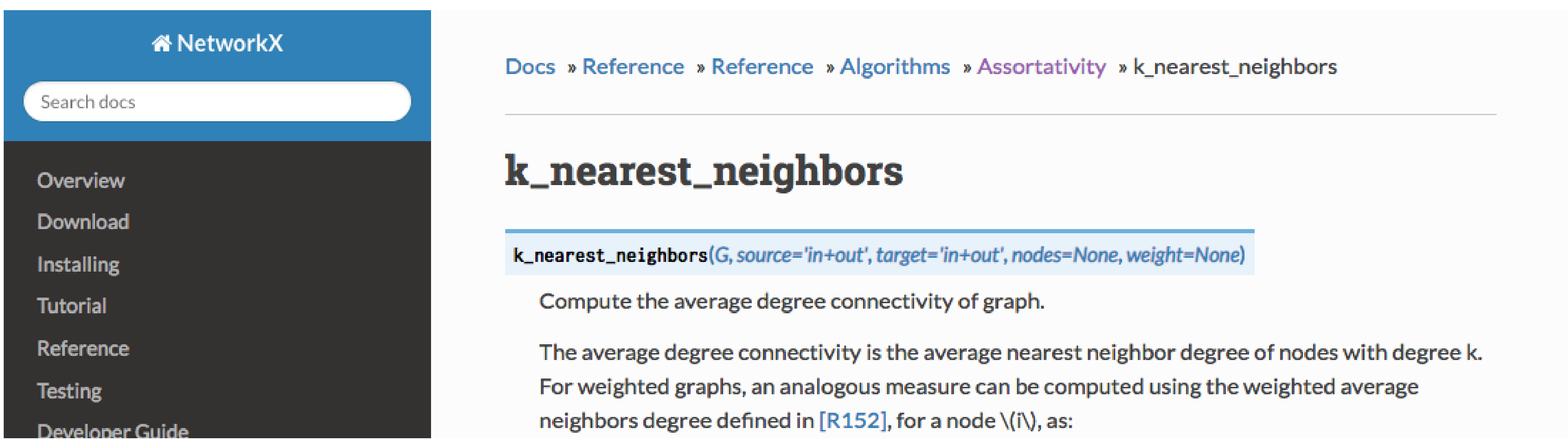
https://en.wikiquote.org/wiki/Yogi_Berra

What's the average degree of your friends?

- Degree correlation function is the average degree of the neighbors of all degree-***k*** nodes

$$k_{nn}(k) = \sum_{k'} k' P(k' | k)$$

conditional probability that, starting from a k-degree node, we reach a neighbor with degree k'



What's the average degree of your friends?

- Degree correlation function is the average degree of the neighbors of all degree-***k*** nodes

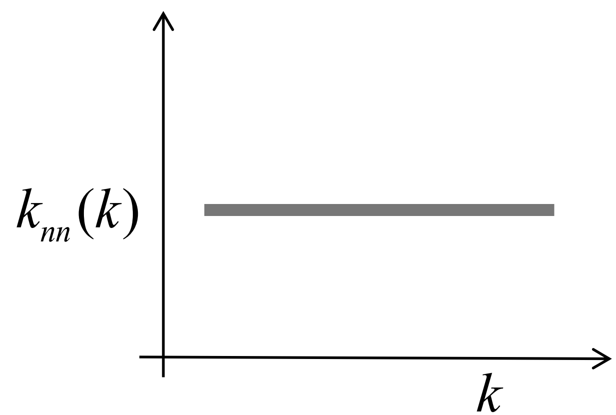
$$k_{nn}(k) = \sum_{k'} k' P(k' \mid k)$$

conditional probability that starting from a k-degree node, we reach a neighbor with degree k'

- Neutral network: $k_{nn}(k) = \dfrac{\langle k^2 \rangle}{\langle k \rangle}$

The average degree of a node's *i* neighbors is independent of the degree of *i*…

…it does not depend solely on the average degree, but also on the variance of the degrees!



What's the average degree of your friends?

- Degree correlation function is the average degree of the neighbors of all degree-***k*** nodes

$$k_{nn}(k) = \sum_{k'} k' P(k' | k)$$

conditional probability that, following a link of a k-degree node we reach a degree-k'

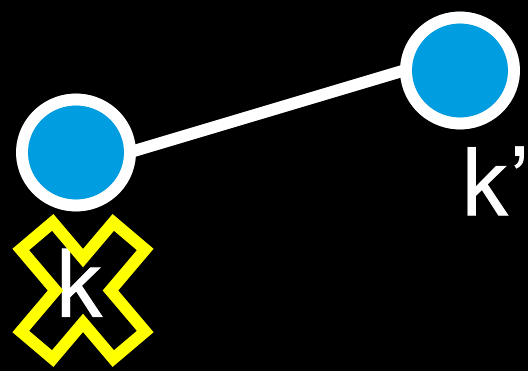
- Neutral network: $k_{nn}(k) = \frac{\langle k^2 \rangle}{\langle k \rangle}$

?

Back to our original challenge:
<u>What is the average degree</u> of the neighbors of degree-*k* nodes?

$$k_{nn}(k) = \sum_{k'} k' P(k' \mid k)$$

conditional probability that starting from a
k-degree node, we reach a neighbor with degree k'



In the absence of correlations, k does not play any role. Thus, P(k'|k) is given by the ***probability*** $q_{k'}$ ***of finding a node with degree k' at the end of a randomly chosen link***:

$$P(k' \mid k) = q_{k'}$$

What is the probability q_k that the node at the end of a randomly chosen link has degree k ?

$$q_k = Const.k.p_k$$

Ensures that we have a probability

The higher is the degree of a node, the higher is the chance that it is located at the end of the chosen link.

The more degree-k nodes are in the network (i.e., the higher is p_k), the more likely that a degree k node is at the end of the link.

$$\sum_k q_k = 1 = Const.\sum_k kp_k$$

$$\Rightarrow \quad Const = \frac{1}{\langle k \rangle}$$

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

$$k_{nn}(k) = \sum_{k'} k' \frac{k' p_{k'}}{\langle k \rangle} = \frac{\langle k^2 \rangle}{\langle k \rangle}$$

What's the average degree of your friends?

- Degree correlation function is the average degree of the neighbors of all degree-***k*** nodes

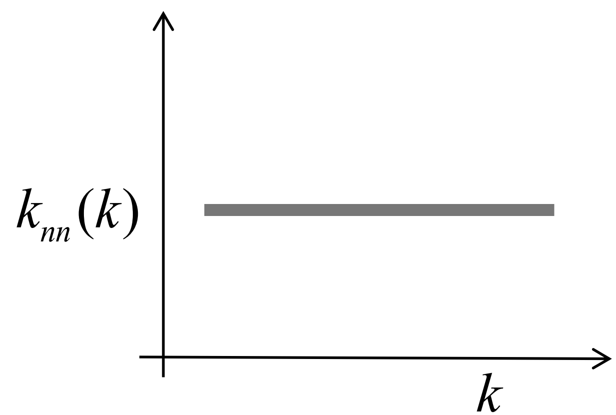
$$k_{nn}(k) = \sum_{k'} k' P(k' \mid k)$$

conditional probability that starting from a
k-degree node, we reach a neighbor with degree k'

- Neutral network: $k_{nn}(k) = \frac{\langle k^2 \rangle}{\langle k \rangle}$

The average degree of a node's *i* neighbors
is independent of the degree of *i*…

…it does not depend solely on the average
degree, but also on the variance of the degrees!



Just to recall:

The 2nd moment of a scale-free network diverges for large N

- For Scale-free networks we have

$$\langle k^n \rangle \, \square \, \square_{k_{min}}^{k_{max}} k^n P_k \, dk = C \frac{k_{max}^{n-\gamma+1} - k_{min}^{n-\gamma+1}}{n-\gamma+1}$$

$$\langle k^2 \rangle \, \square \, N^{\frac{3-\gamma}{\gamma-1}}$$

Diverges for γ<3

- n=0 sums to one.
- n=1 gives the **average** degree
- n=2 helps us to calculate the **variance**
- n=3 determines the **skewness**

$$k = \langle k \rangle \pm \times$$

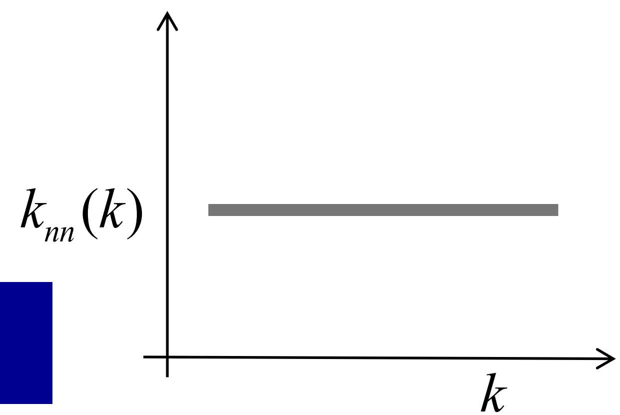
What's the average degree of your friends?

- De 

of the ne

k-degree node, we reach a neighbor with degree k'

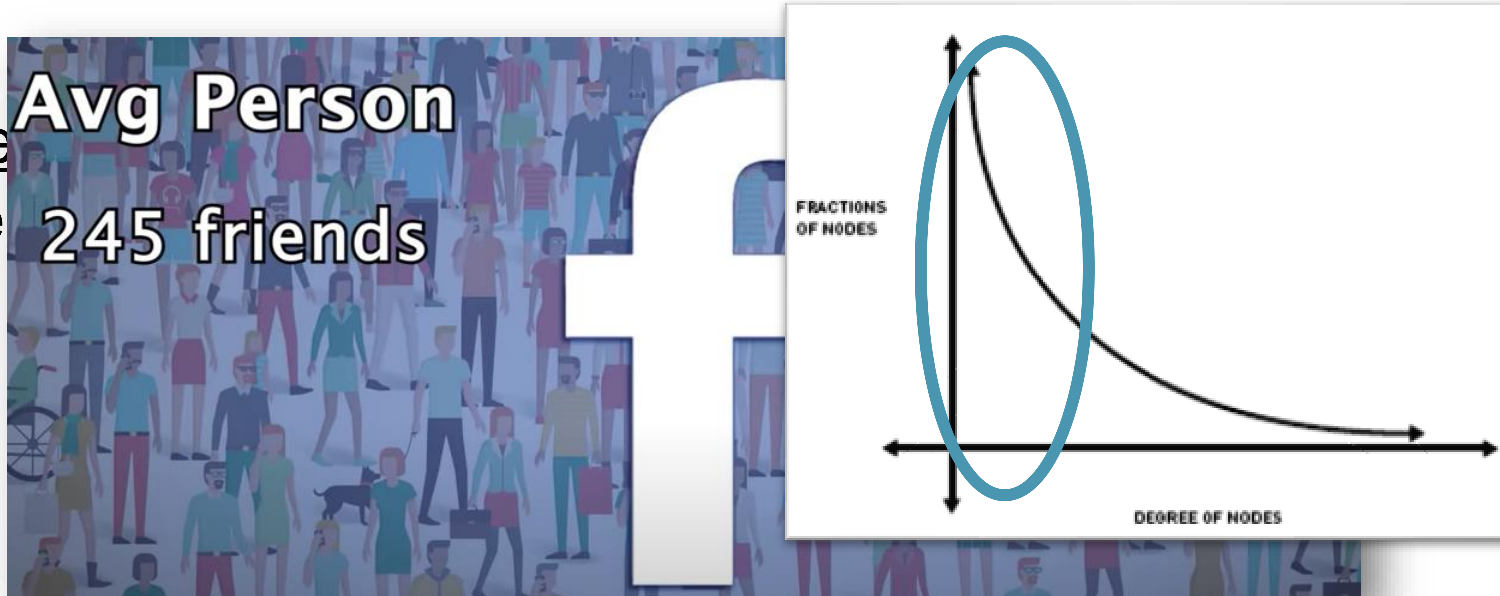
- Neutral network: $$k_{nn}(k) = \frac{\langle k^2 \rangle}{\langle k \rangle}$$



Friendship paradox:
On average my friends are more popular than I am

What’s the average degree of your friends?

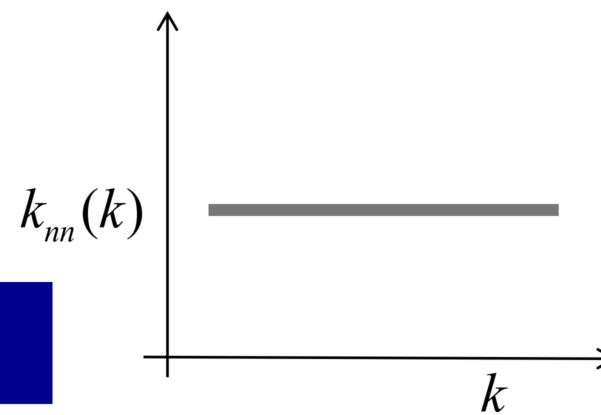
- De
  ne



k-degree node, we reach a neighbor with degree k'

- Neutral network: $$k_{nn}(k) = \frac{\langle k^2 \rangle}{\langle k \rangle}$$

Friendship paradox:
On average my friends are more popular than I am



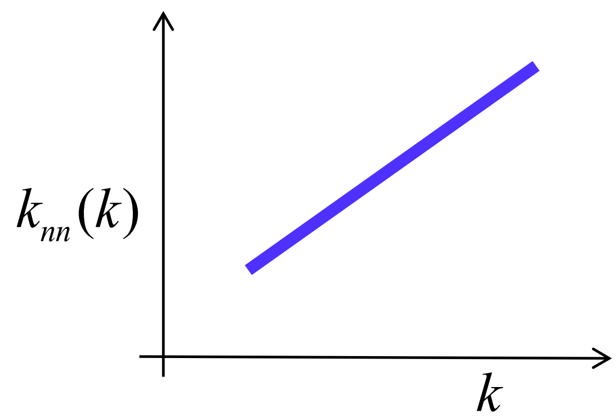
What's the average degree of your friends?

- Degree correlation function is the average degree of the neighbors of all degree-***k*** nodes

$$k_{nn}(k) = \sum_{k'} k' P(k' | k)$$

conditional probability that following a link of a k-degree node we reach a degree-k'

- Assortative network:



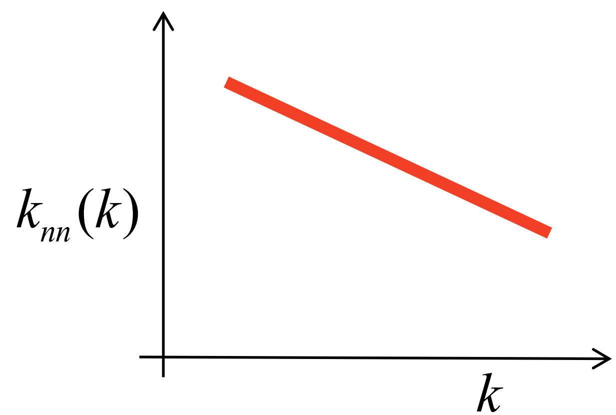
What’s the average degree of your friends?

- Degree correlation function is the average degree of the neighbors of all degree-***k*** nodes

$$k_{nn}(k) = \sum_{k'} k' P(k' \mid k)$$

conditional probability that following a link of a k-degree node we reach a degree-k'

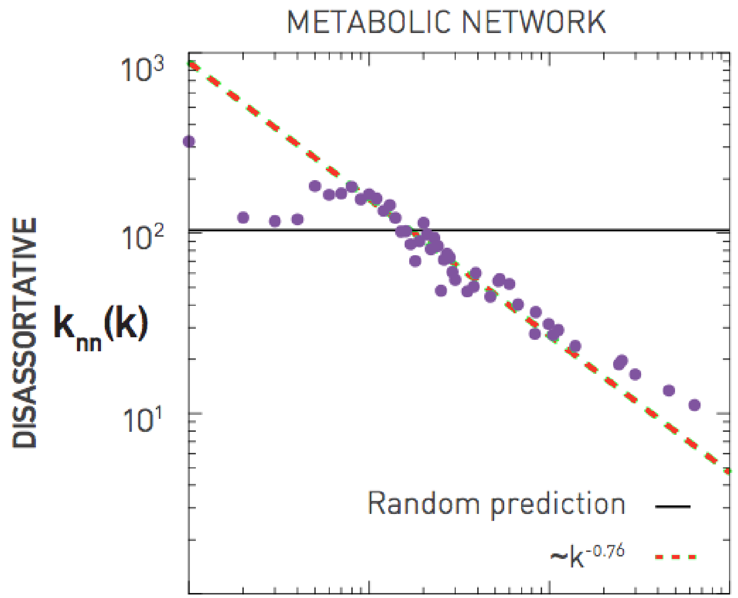
- Disassortative network:

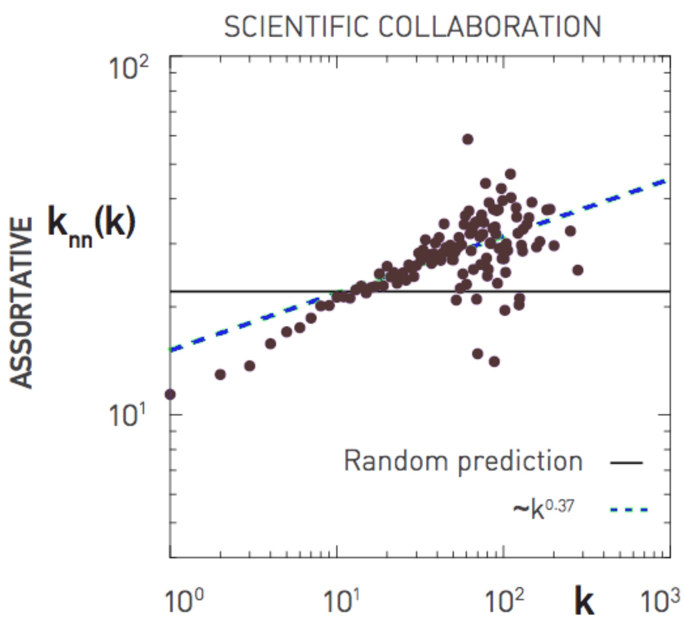


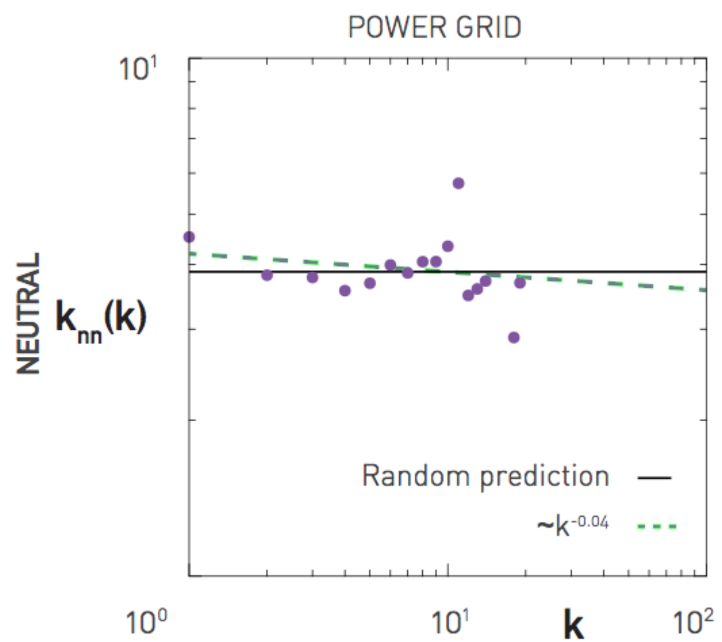
Examples

$$k_{nn}(k) \sim k^{\mu}$$

Disassortative (μ<0)
Neutral (μ=0)
Assortative (μ>0)





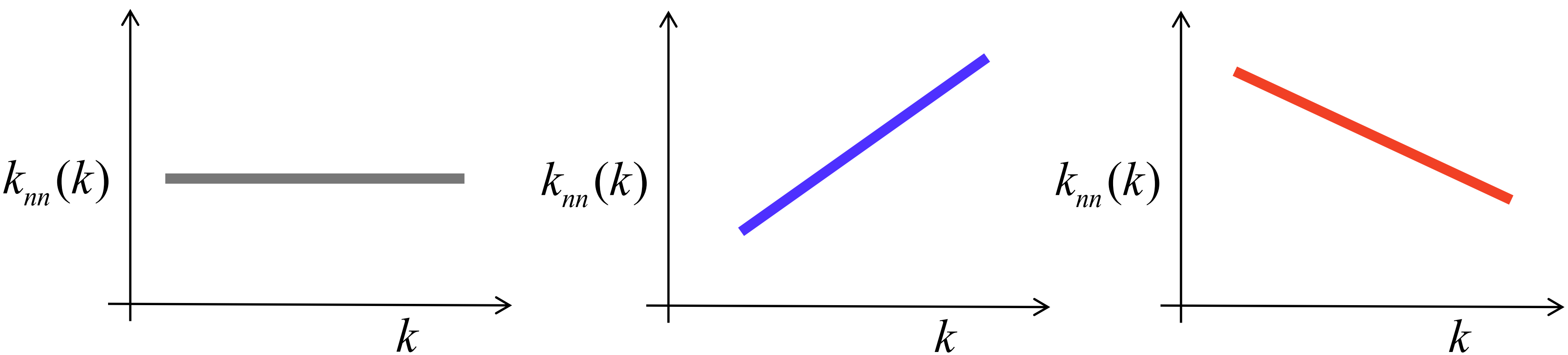


Examples

$$k_{nn}(k) \sim k^{\mu}$$

Disassortative (μ<0)
Neutral (μ=0)
Assortative (μ>0)

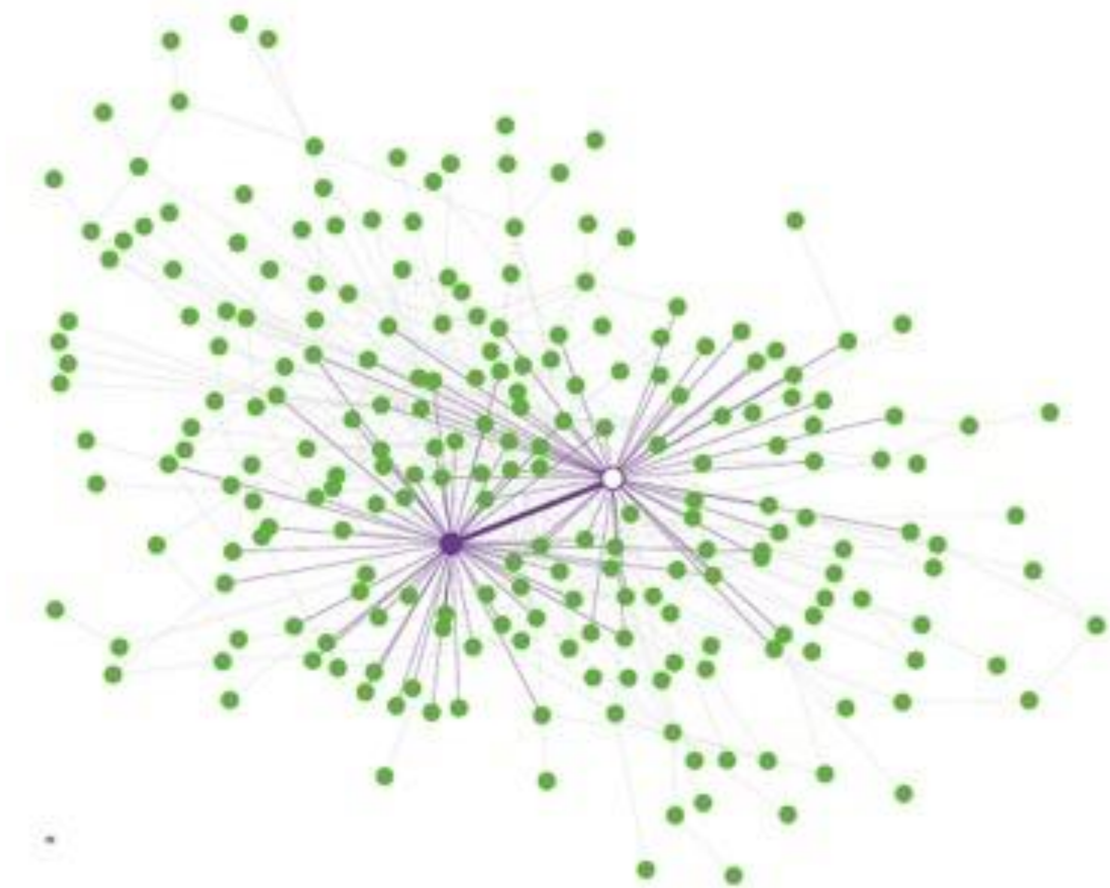
Neutral (μ=0) Assortative (μ>0) Disassortative (μ<0)



Challenge:
Does assortativity have an impact on the significance of the friendship paradox?

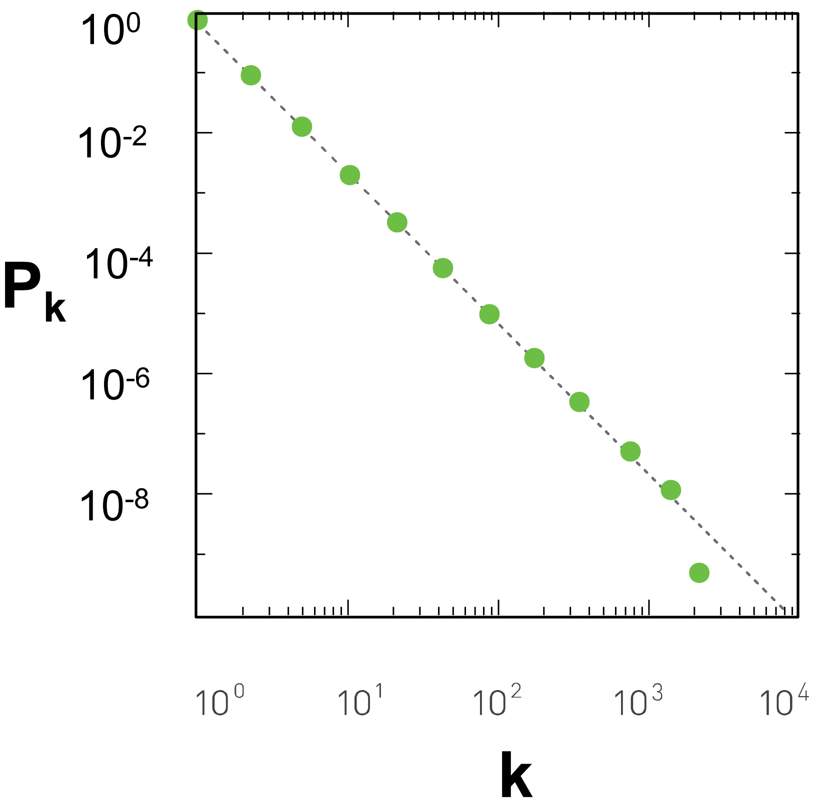
Structural cut-offs of power-laws & degree correlations

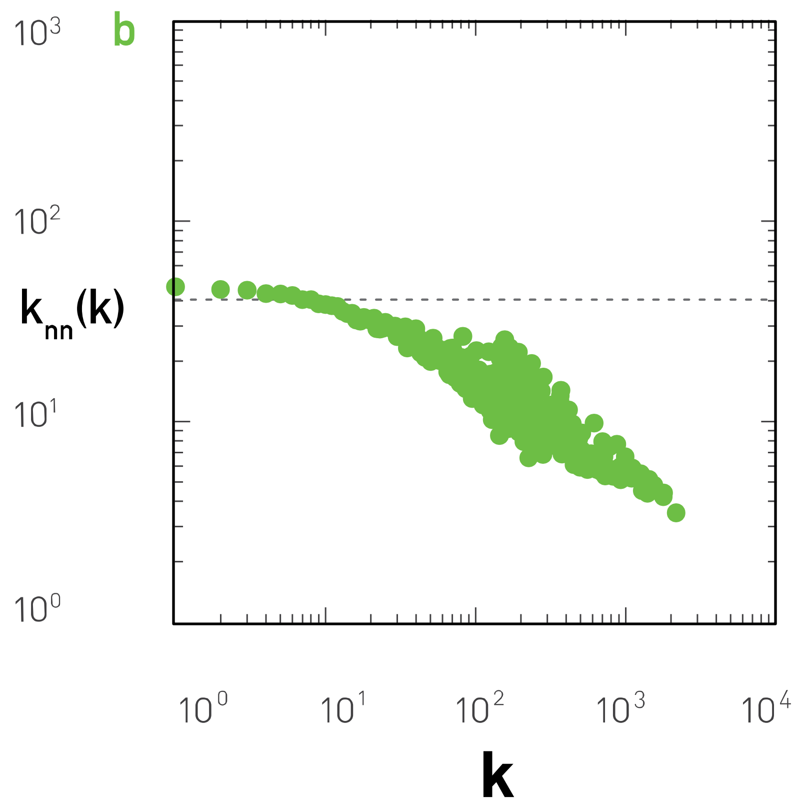
- For $\gamma \leq 3$ we may have some problems. Example: the BA-model… You have few hubs with a very large degree.



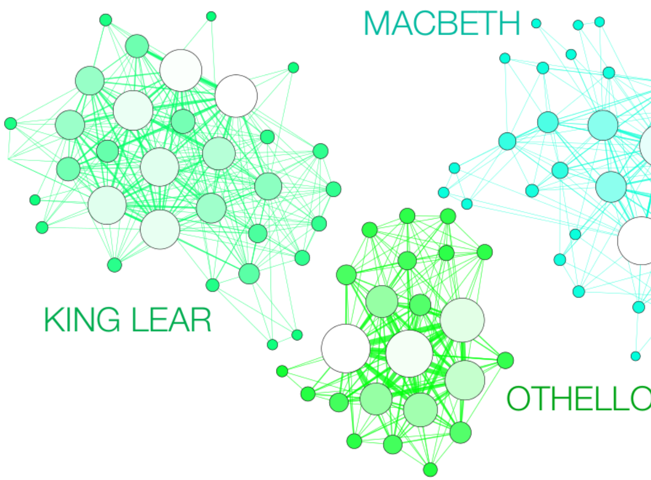
Structural cut-offs of power-laws & degree correlations

- For γ≤3 we may have some problems. Example: the BA-model… You have few hubs with a very large degree.





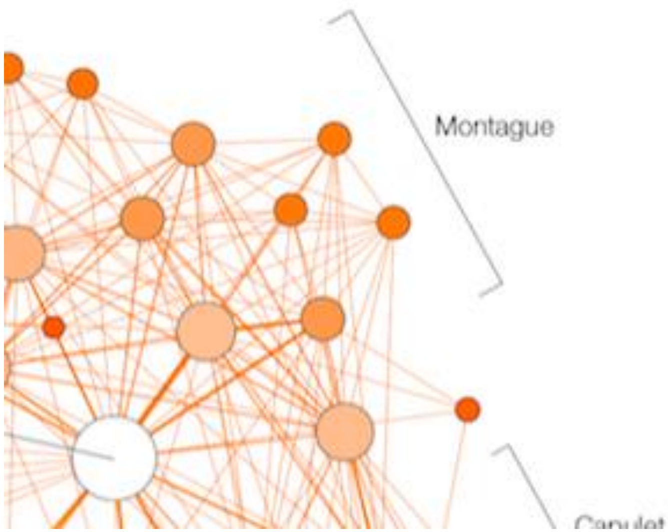
Analyzing real networks

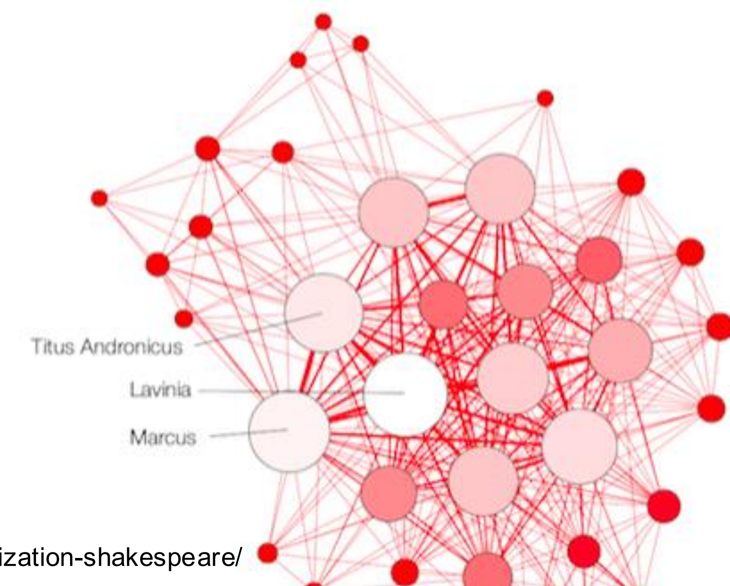


1. Limitations you should be aware of:

Finite size effects lead to “Structural cutoffs”

2. Ways out of these problems…

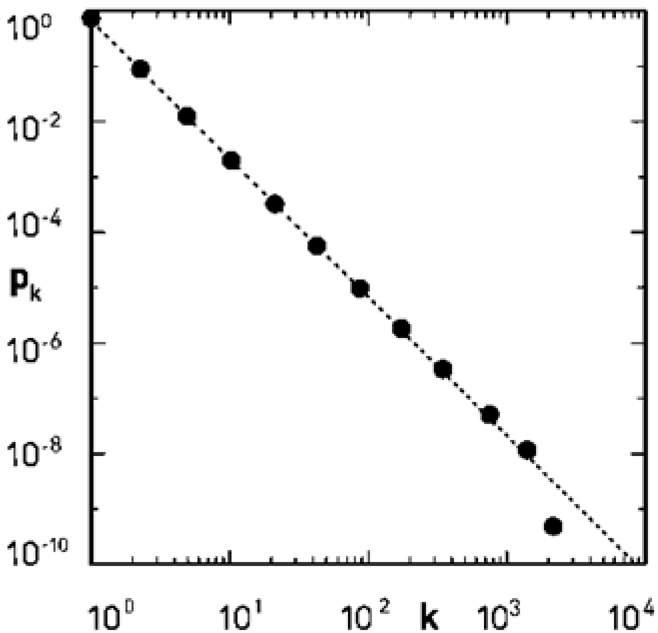


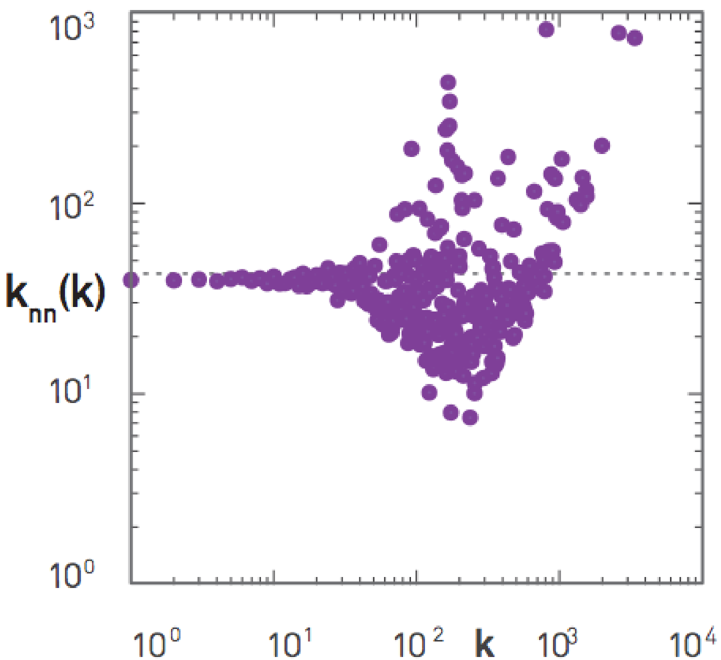


Images from: http://www.martingrandjean.ch/network-visualization-shakespeare/

Structural cut-offs of power-laws & degree correlations

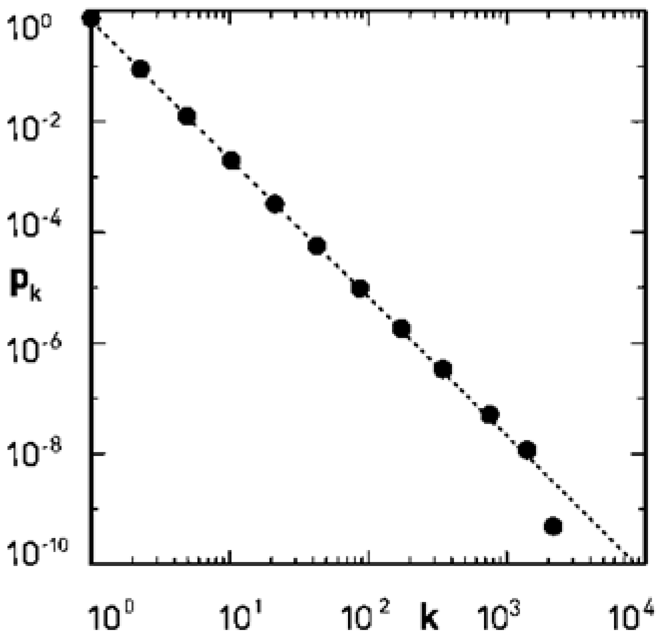
- For γ≤3 we may have some problems. Example: the BA-model… If we relax the original model and allowing multiple links we get a neutral network as expected.

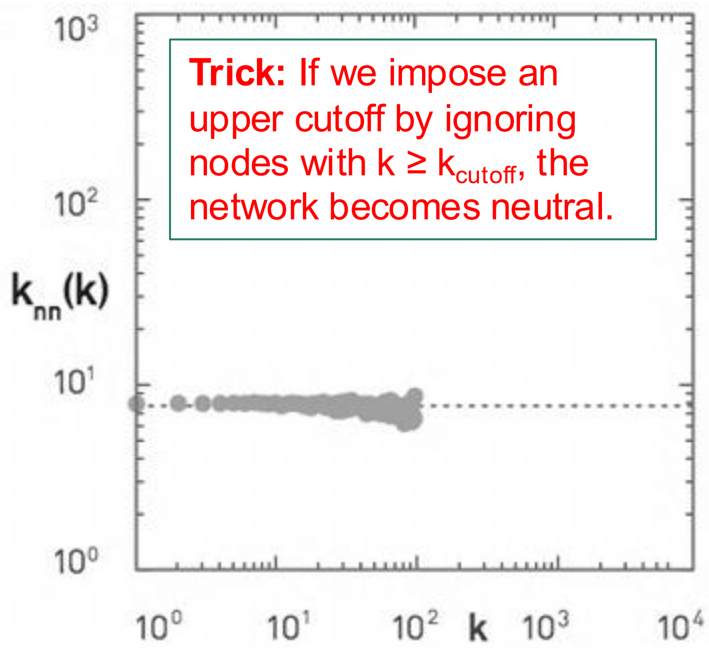




Structural cut-offs of power-laws & degree correlations

- For γ≤3 we may have some problems. Example: the BA-model… Same happens if we impose a structural cutoff.

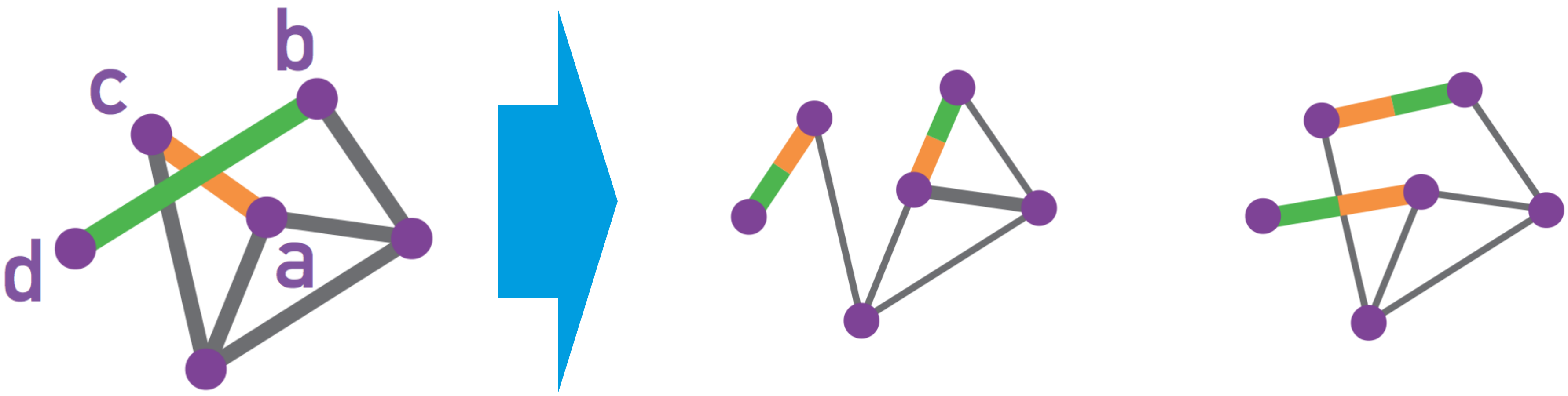




Trick: If we impose an upper cutoff by ignoring nodes with $k \geq k_{cutoff}$, the network becomes neutral.

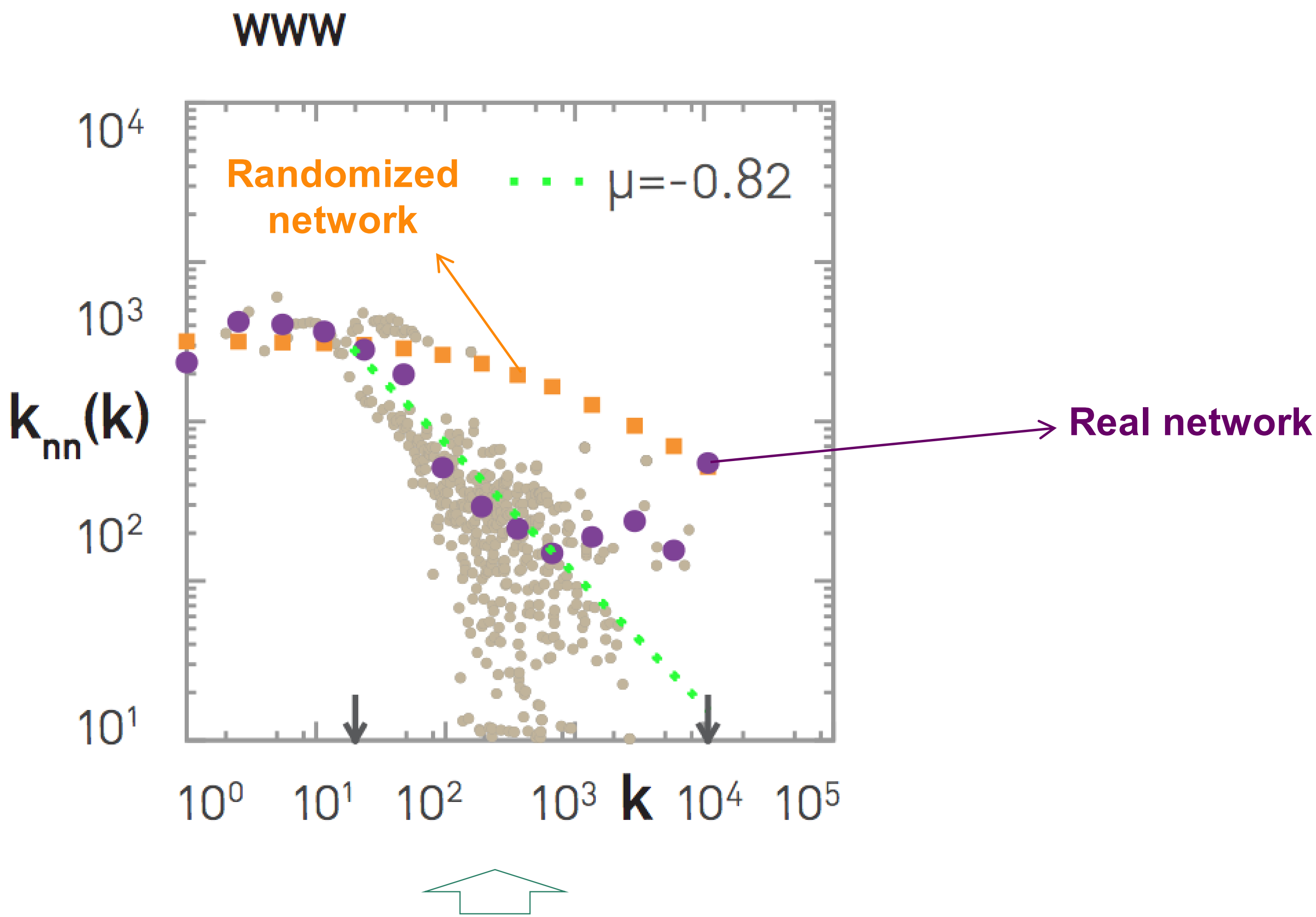
Analyzing real networks

- To be sure, instead of simply computing μ, we should compare our networks with a null (neutral) model of the same network.
- **Trick:** Randomize your network (without changing the degree dist.) and compare the original one with its shuffled version.



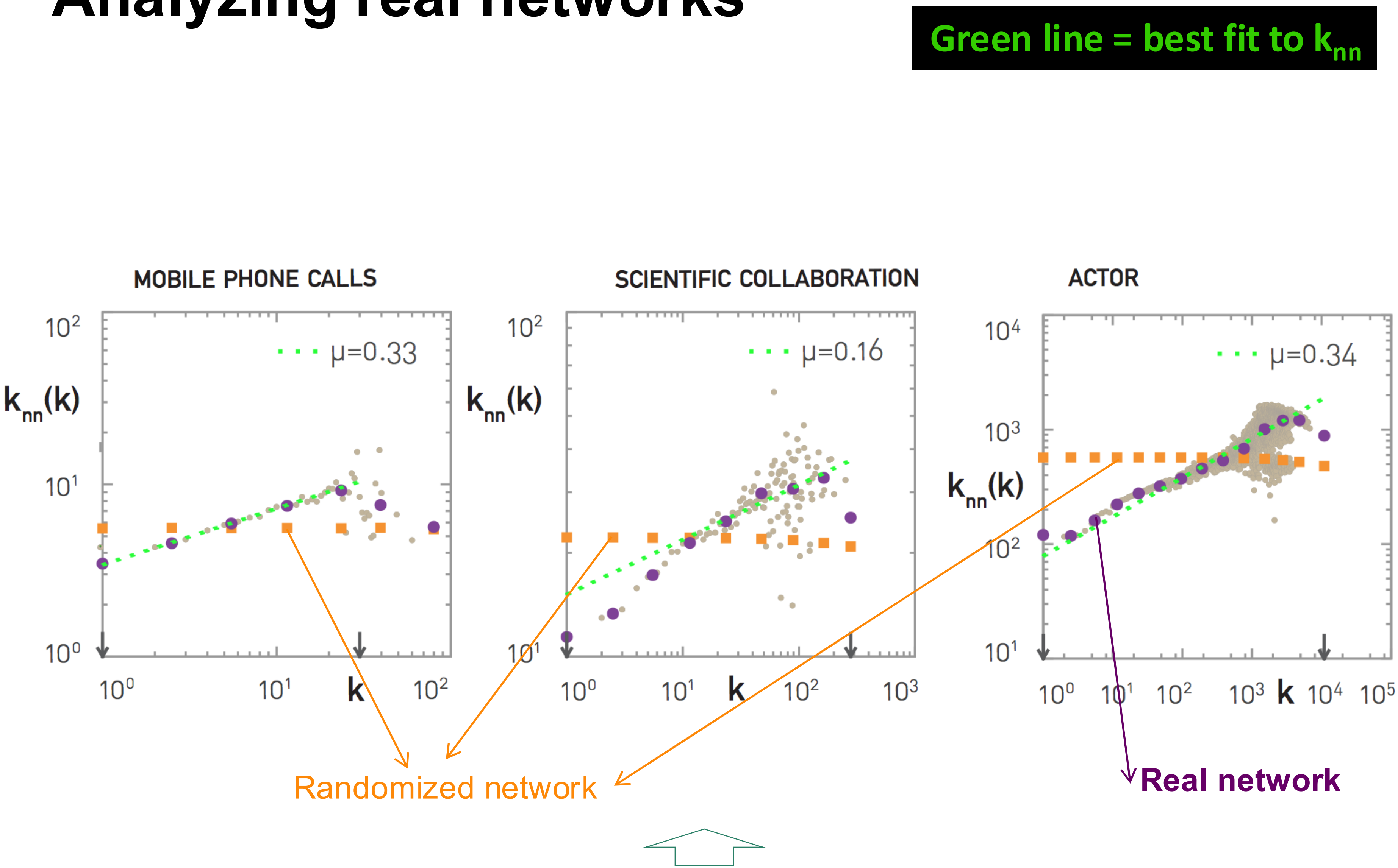
Analyzing real networks

Green line = best fit to k_{nn}



Disassortative

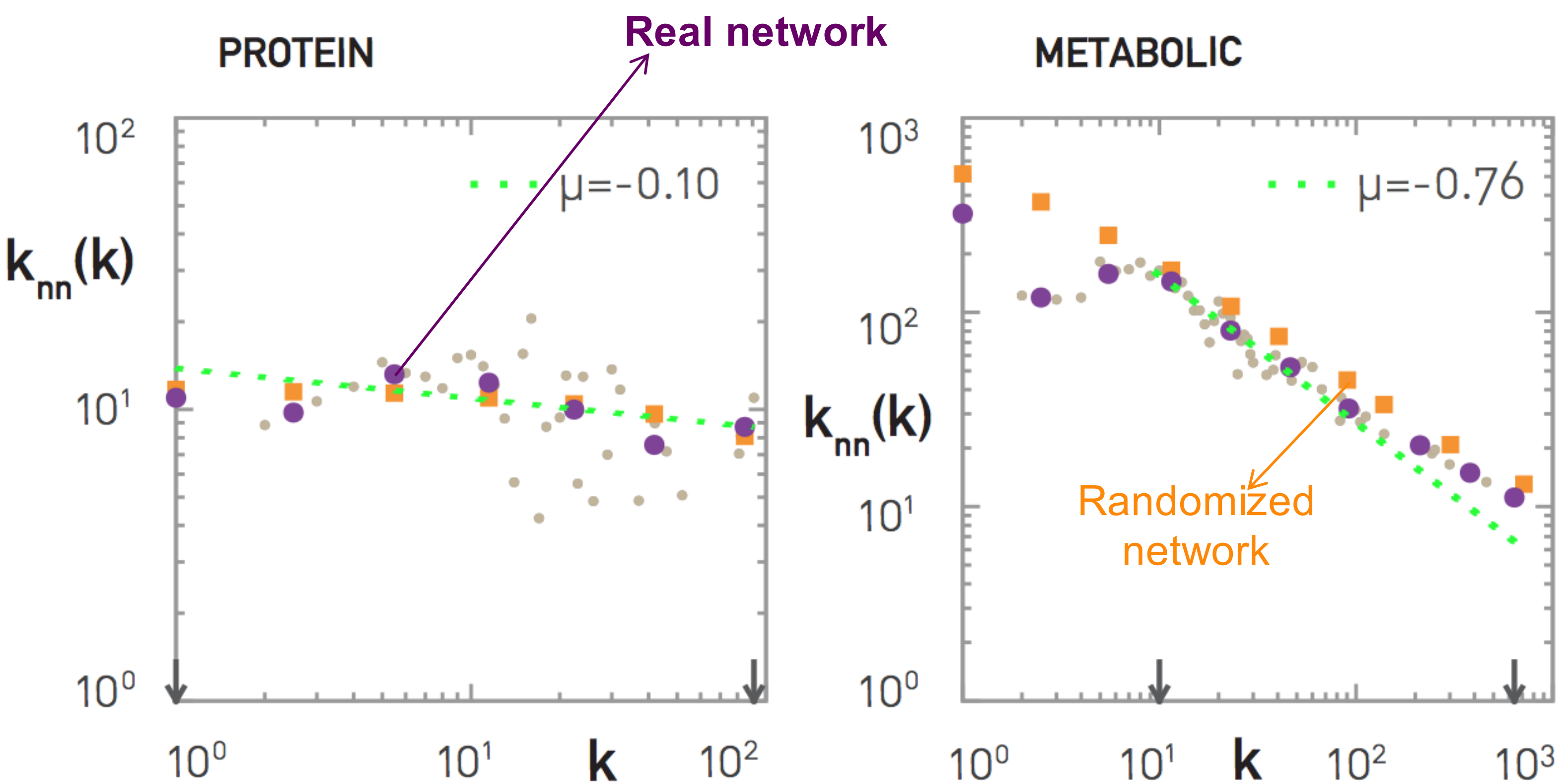
Analyzing real networks



Assortative

Analyzing real networks

Green line = best fit to k_{nn}



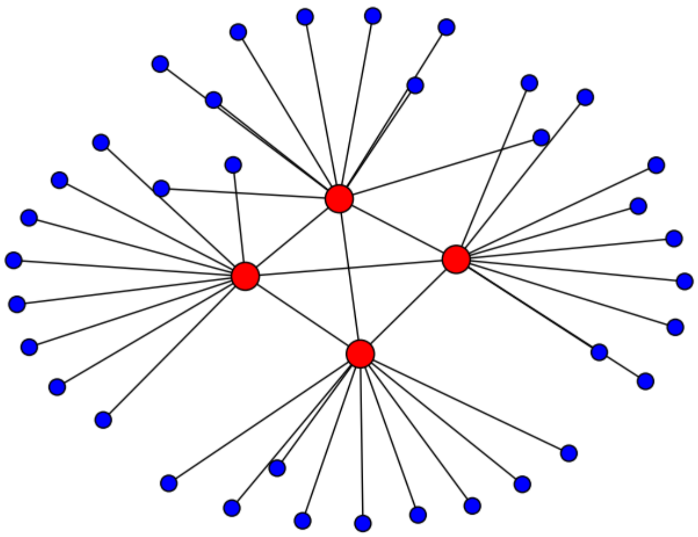
Disassortative (yet, just structural... They're neutral!)

Alternative measures: Rich-club coefficient

Colizza et al. Nature Physics (2006) & Opsahl et al. Phys Rev Lett (2008)

Sometimes we want to measure extent to which well-connected nodes also connect to each other

Often highly connected nodes create a league of highly connect nodes, even in disassortative networks.



Alternative measures: Rich-club coefficient

Colizza et al. Nature Physics (2006) & Opsahl et al. Phys Rev Lett (2008)

Designed to measure the extent to which well-connected nodes also connect to each other.

$$\phi(k) = \frac{E_{>k}}{N_{>k}(N_{>k} - 1)/2}$$

where N_k is the number of nodes with degree larger than k, and E_k is the number of edges among those nodes.

Alternative measures: Rich-club coefficient

Colizza et al. Nature Physics (2006) & Opsahl et al. Phys Rev Lett (2008)

Designed to measure the extent to which well-connected nodes also connect to each other.

$$\phi(k) = \frac{E_{>k}}{N_{>k}(N_{>k}-1)/2}$$

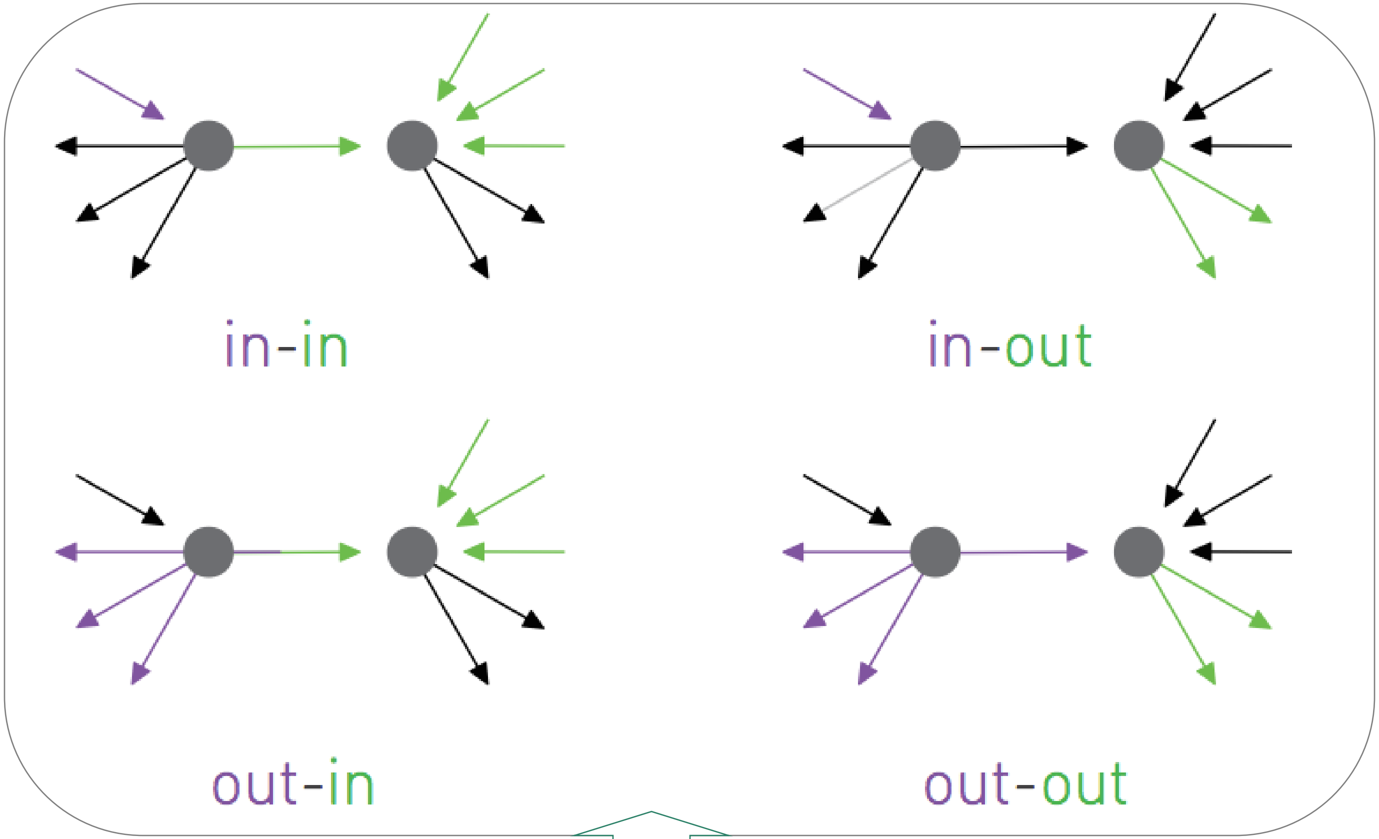
Alternatively, we can randomize the network (Xulvi-Brunet method) and compute a normalize version of it:

$$\phi_{NORM}(k) = \frac{\phi(k)}{\phi_{Rand}(k)}$$

NetworkX
Network Analysis in Python
2.5

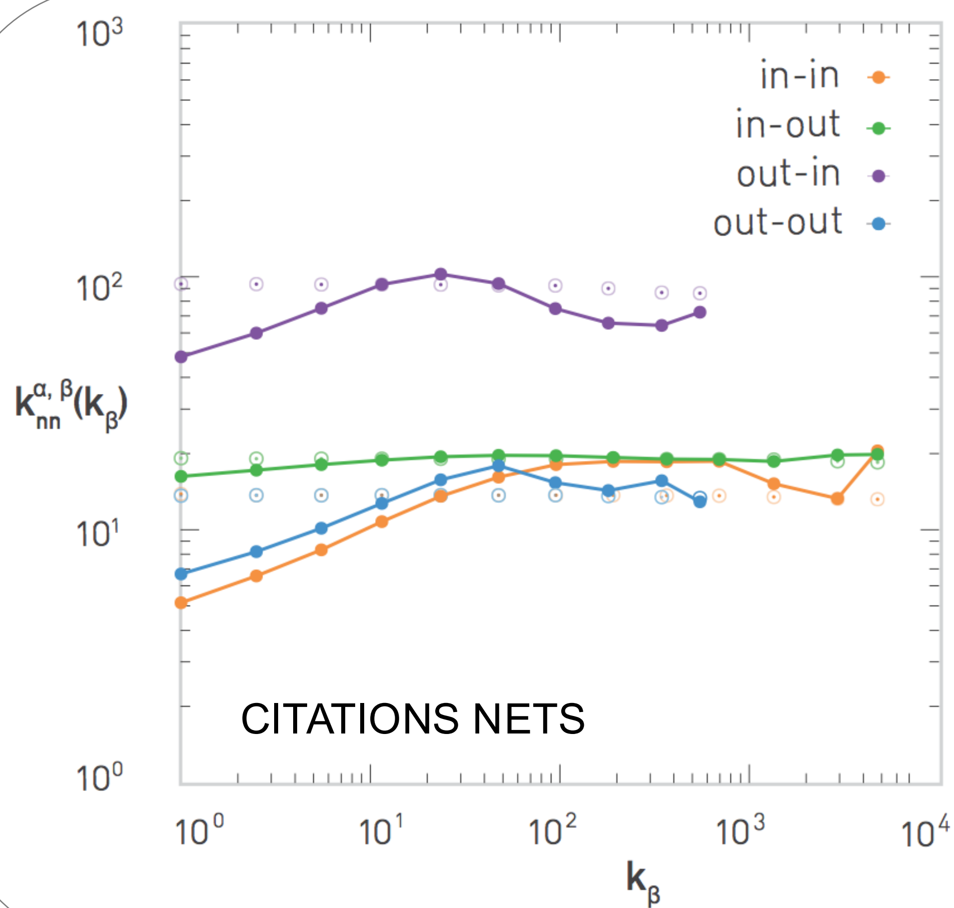
networkx.algorithms.richclub.rich_club_coefficient

Side note: Analyzing *directed* networks



We get 4 possible k_{nn} values

Side note: Analyzing *directed* networks



Models (examples)

- **Erdős-Rényi Model.** Neutral by definition.
- **Configuration model.** It is also neutral by definition. Yet, if we force it to be simple (absence of multiple links and self-loops) then it will become disassortative.

- **BA model.** Neutral ➔ $k_{nn}(k) \sim \frac{m}{2} \ln N$
- **BA model** with initial attractiveness ($-m<A<0$). Disassortative ➔ $k_{nn}(k) \sim k^{-|A|/m}$
- **BA model** with initial attractiveness ($A>0$, $\gamma>3$). Assortative (but weak) ➔ $k_{nn}(k) \sim \ln\left(\frac{k}{m+A}\right)$

Let’s create a model…

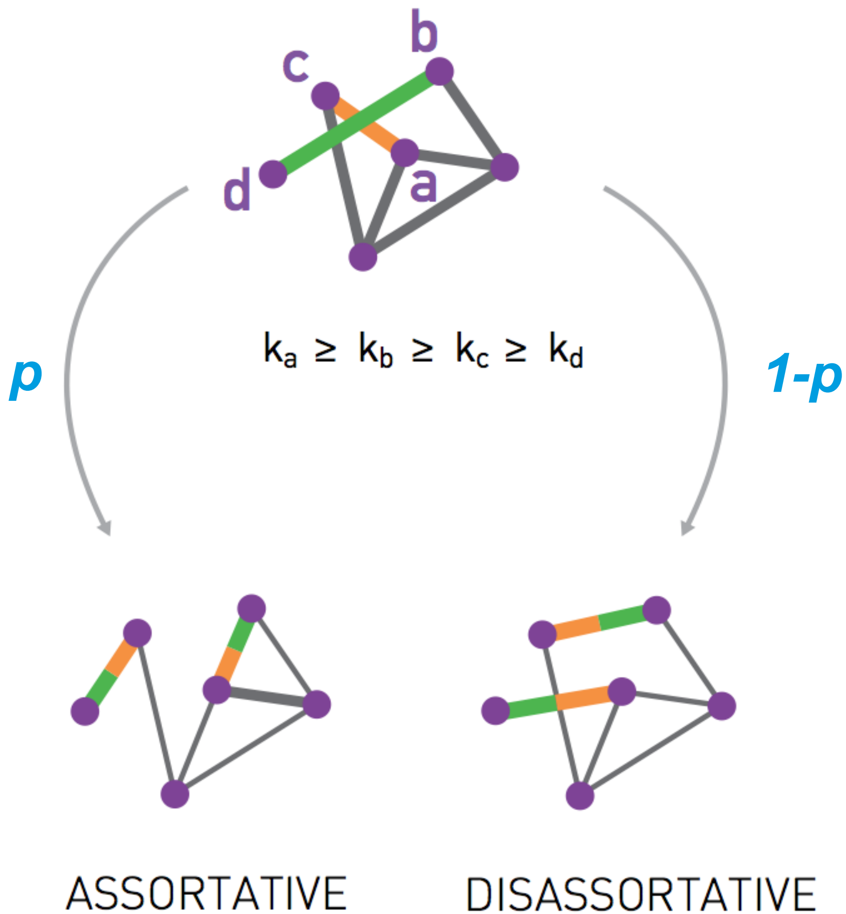
Propose a simple model to answer the following question:

What’s the impact of assortativity on the average path length & the diameter of a network?

Take a ***Random graph*** as an example

Tuning degree correlations
Xulvi-Brunet et al. algorithm

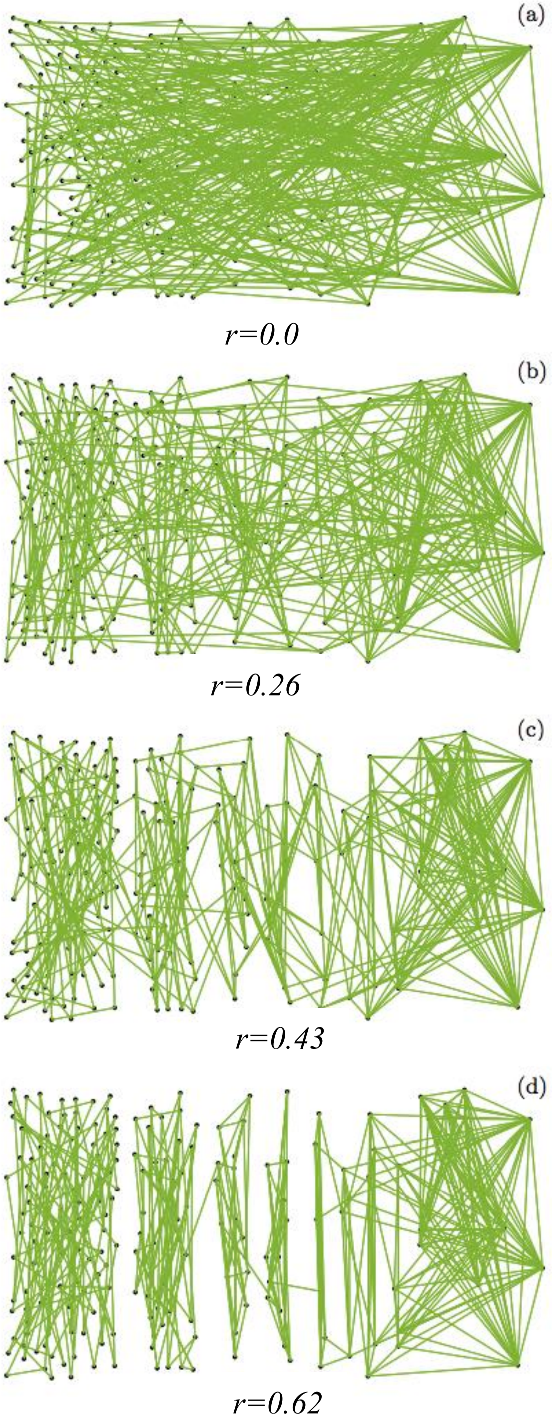
- Algorithm to generate networks with a given degree correlations.



Tuning degree correlations
Xulvi-Brunet et al. algorithm

- Example:

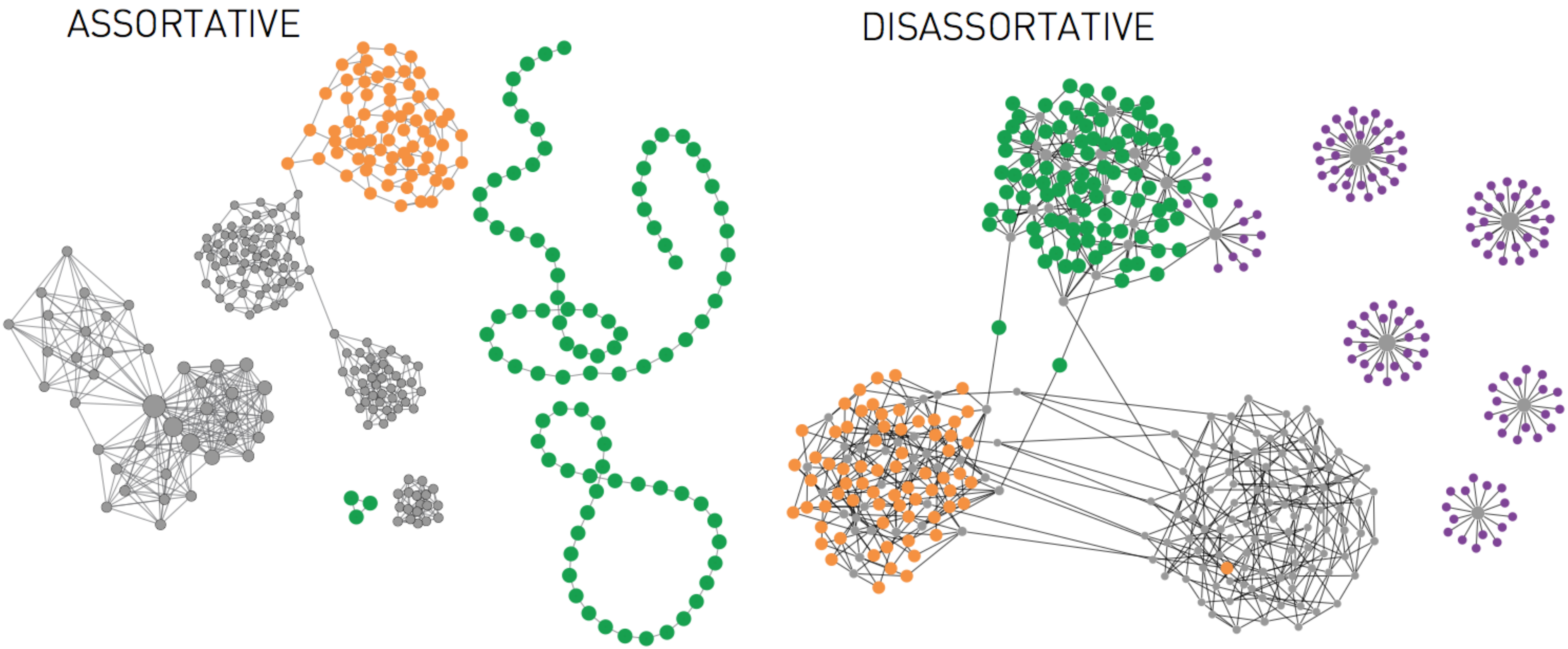
scale-free BA model, linear pref. attach.



Tuning degree correlations
Xulvi-Brunet et al. algorithm

- Example:

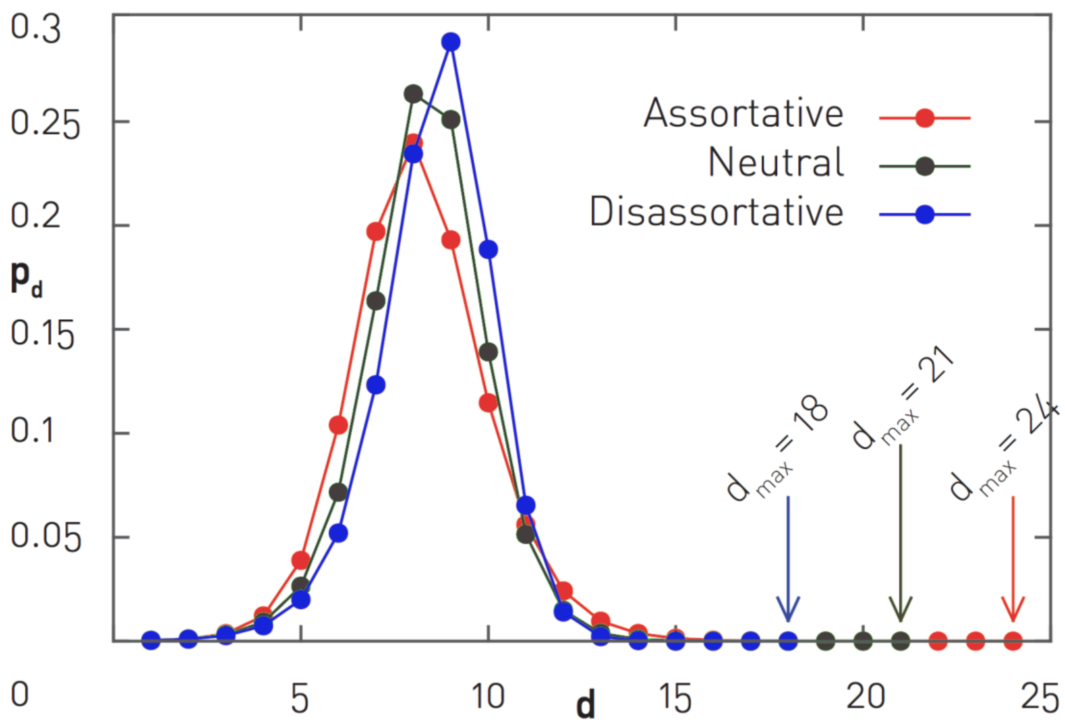
scale-free BA model, linear pref. attach.



Impact of assortativity on <L> & diameter?

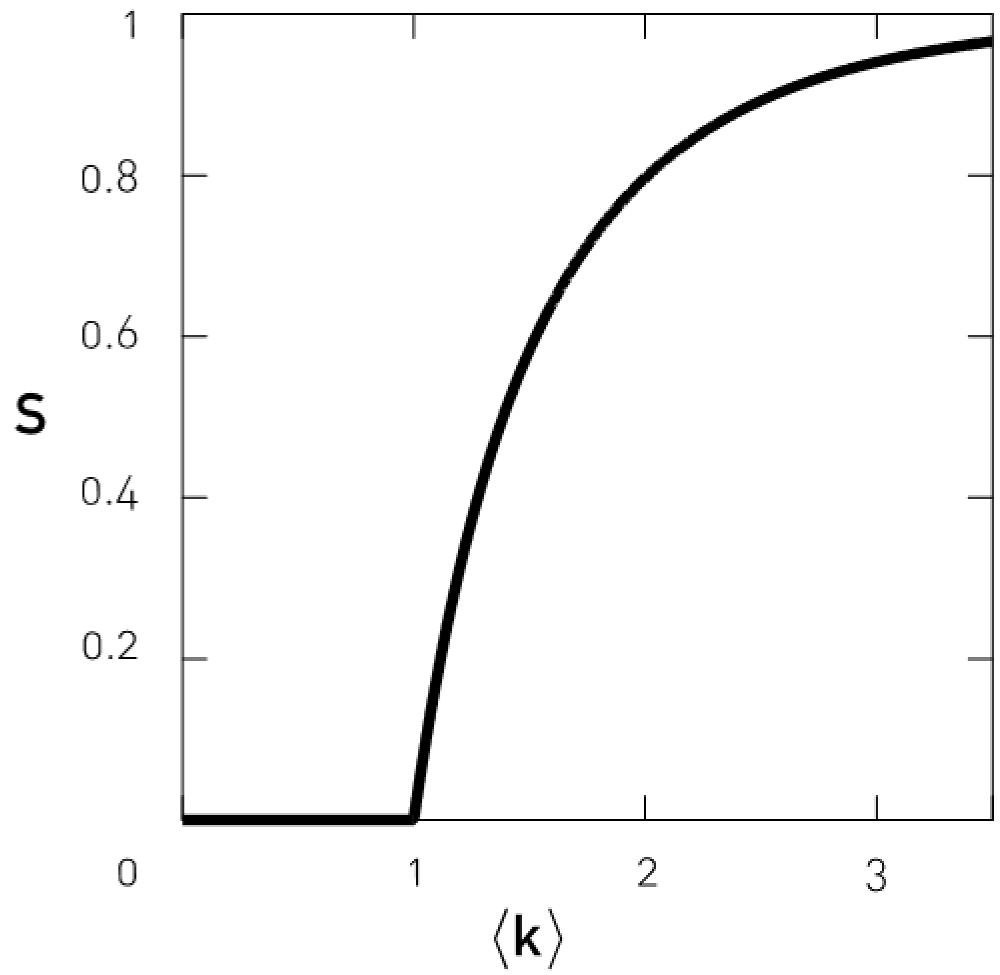
Example: ***Random graphs***

- Assortativity
  - reduces the APL (<L>, hubs get closer)
  - increases the diameter of the network (foster chains of low degree nodes).



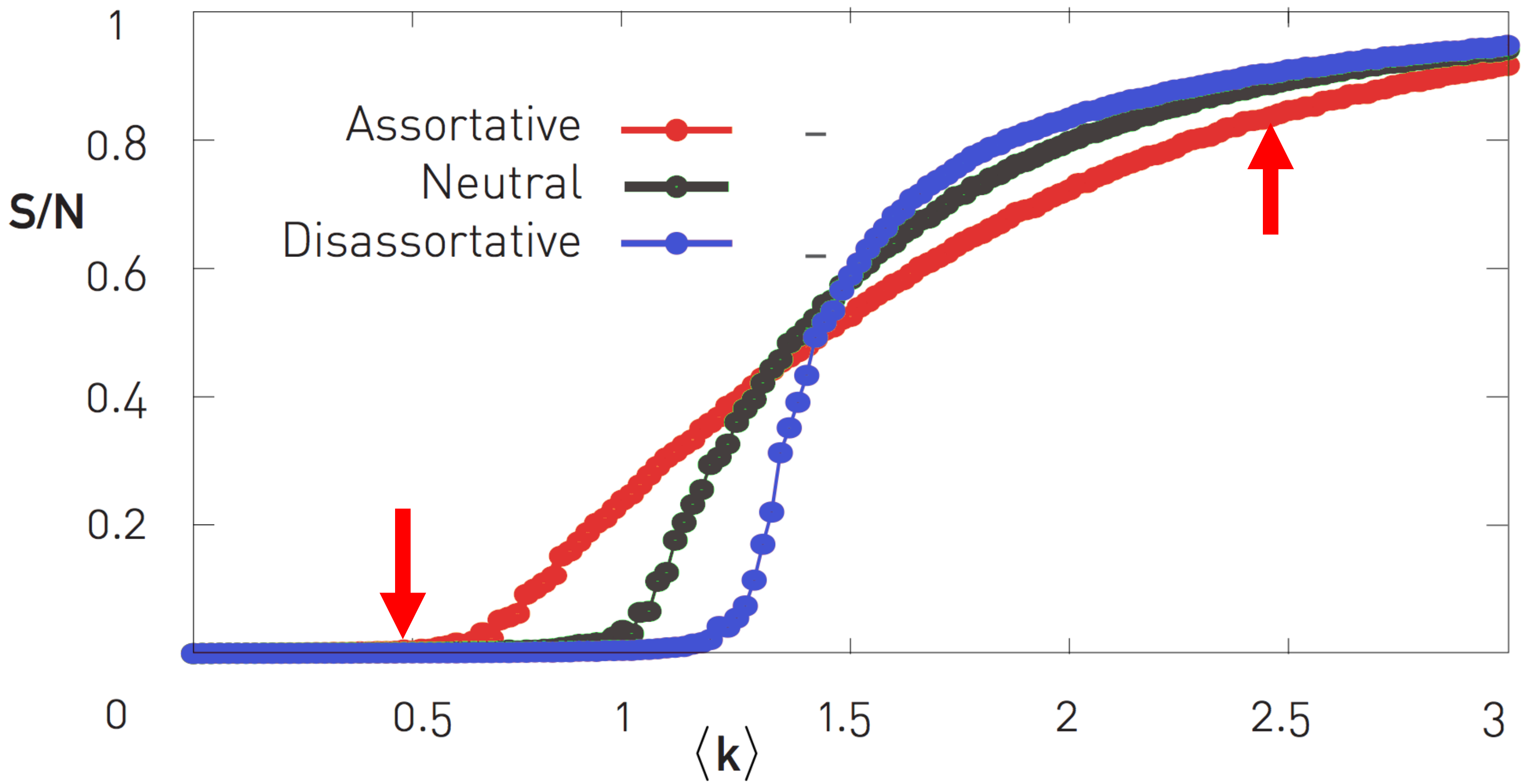
Impact: assortativity & giant components?

Example: ***Random graphs***



Impact: assortativity & giant components

Example: ***Random graphs***



- It is easier to start a giant component if high-degree nodes are connected.
- The giant component is smaller for large k in assortative nets: since hubs are forced to be linked, they miss low-degree nodes.

Conclusion

- Degree correlations are present in most real networks
- Once present, degree correlations change a network’s behavior.
- Degree correlations show that there’s much more beyond the degree distributions, allowing to quantify patterns that govern the way nodes link to each other that are not captured by the degree distributions.
- Our knowledge of the impact of degree correlations in many dynamical systems grounded on complex networks is still largely incomplete…

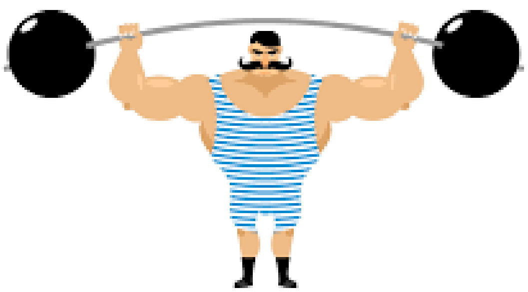
What's next?

Resilience of complex networks and cascading effects

Does it really matter?

- Robustness in ***biology and medicine***: there are countless protein misfolding errors, missed cell reactions, and mutations which are neutral and others that lead to diseases.
- Stability of ***human societies and institutions***: social, economical and political networks are constantly being perturbed by wars, political and economical cycles, etc.
- ***Ecology and sustainability***: analyze the disruptive effect of climate & human activity in ecological networks.
- Ultimate goal in ***engineering***: design communication systems, cars or airplanes which cope with occasional component failures.

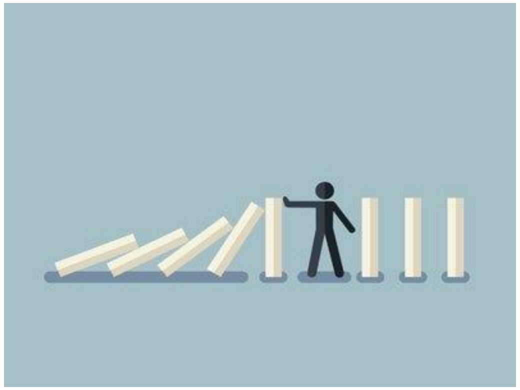
Synopsis



robustness

Random failures & attacks

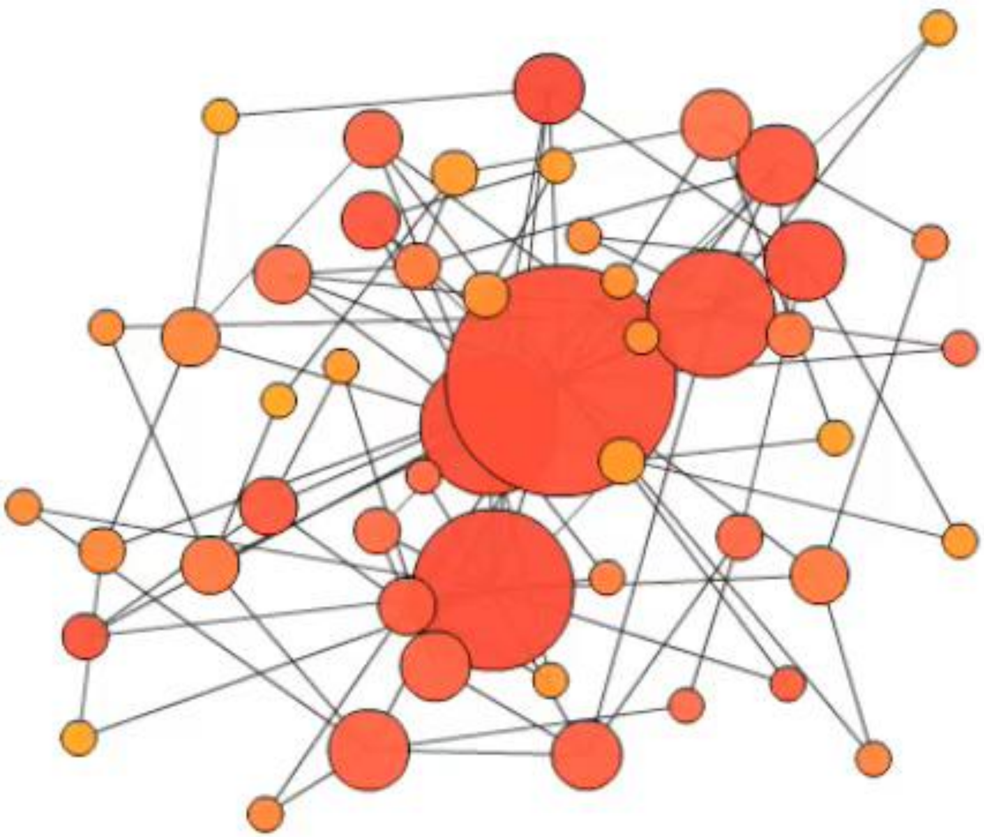
Cascading effects



Building robustness

Modeling cascading failures

Next step: Network robustness



How many nodes do we have to delete to fragment the network into isolated components?

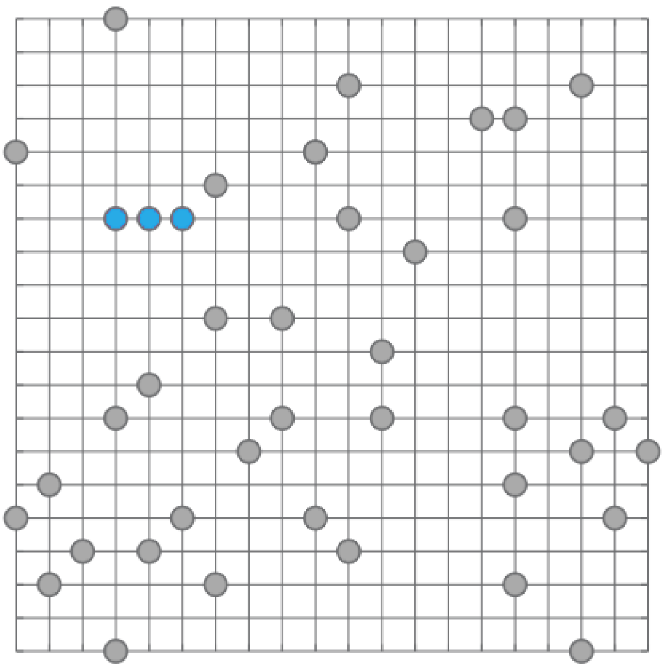
Scale-free nets show an usual behavior: we must remove almost all of its nodes to destroy its giant component.

What's the origin of this result?

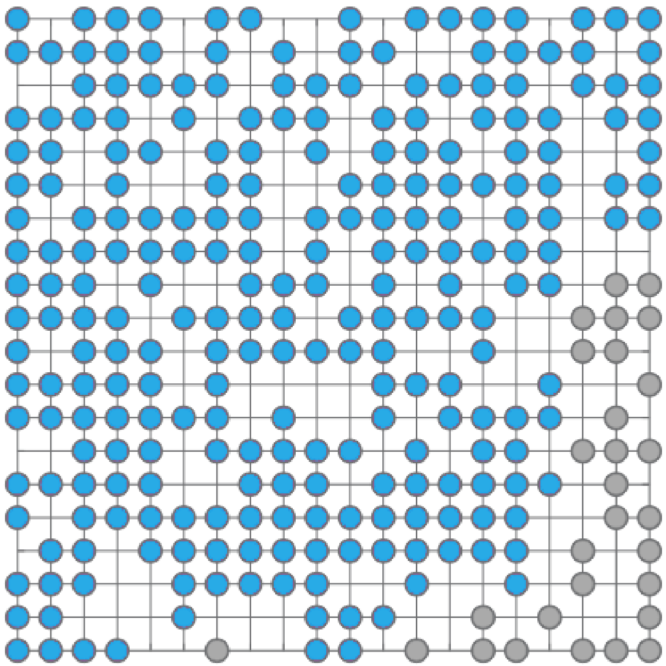
Robustness toolkit: percolation theory

Example: site percolation in 2D

- Let's place pebbles with probability p at each intersection



p=0.1

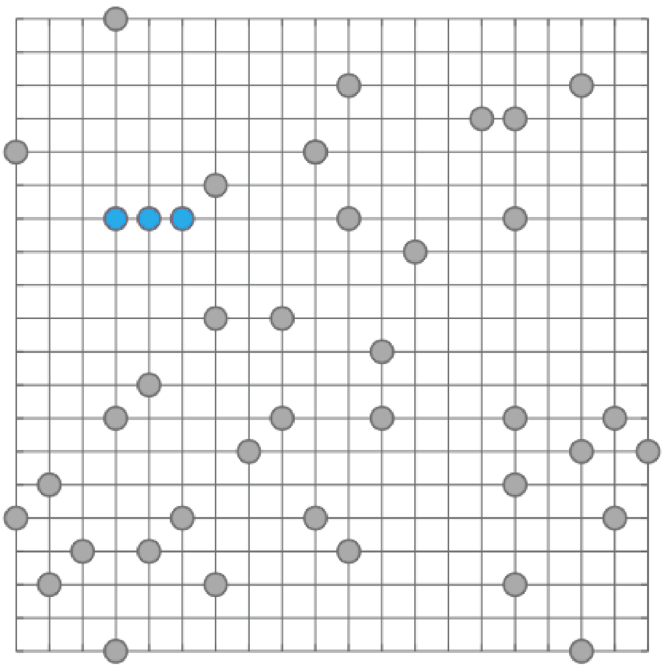


p=0.7

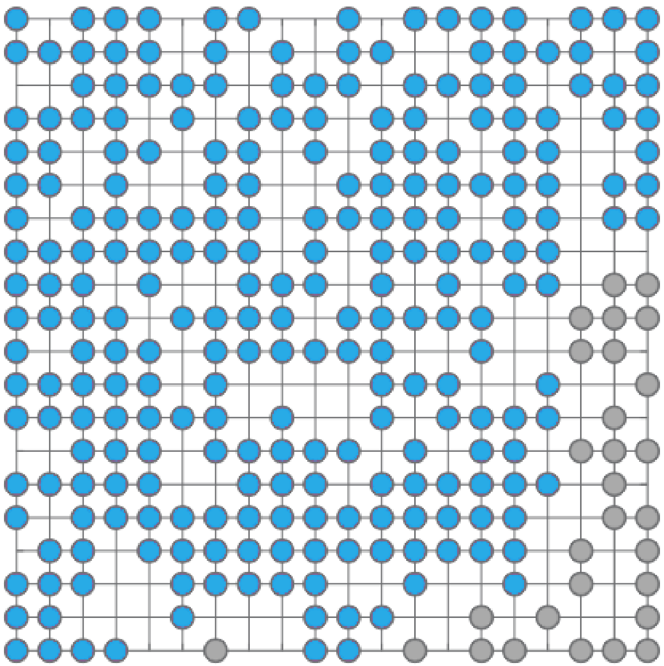
Robustness toolkit: percolation theory

Example: site percolation in 2D

- What's the expected size of the largest cluster?
- What's the average cluster size?



p=0.1

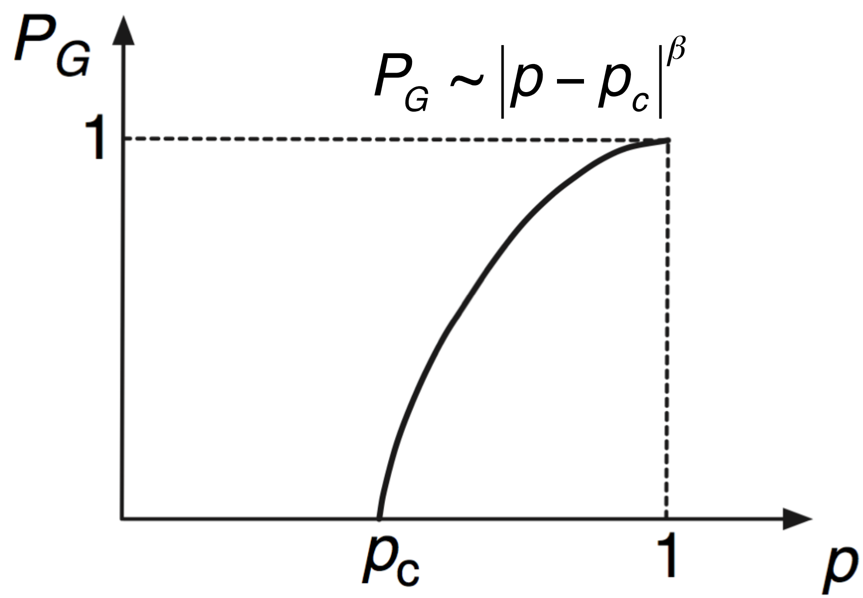


p=0.7

Robustness toolkit: percolation theory

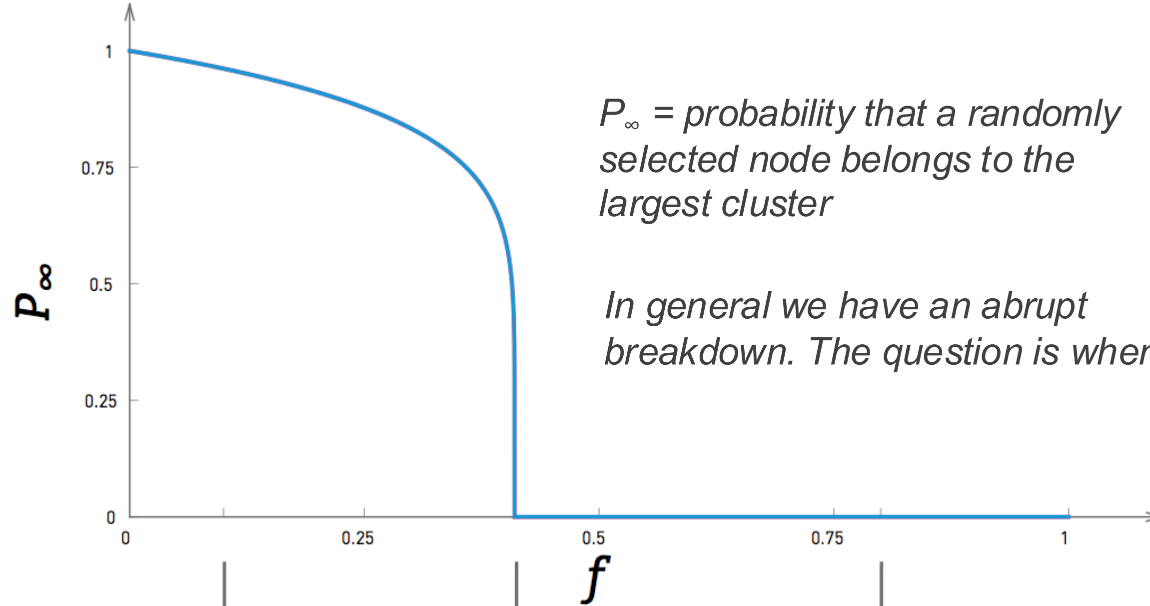
Example: site percolation in 2D

- What's the expected size of the largest cluster?



Probability for a node to belong to the largest cluster

Robustness as an inverse percolation problem



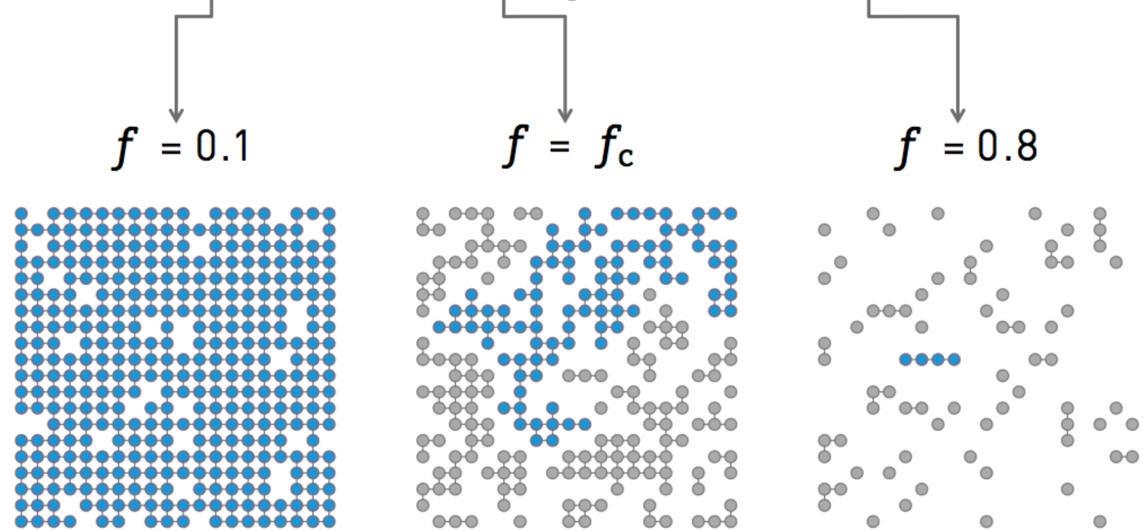
P_∞ = probability that a randomly selected node belongs to the largest cluster

In general we have an abrupt breakdown. The question is when.

Fraction of removed nodes:

$$f = 1\text{-}p$$

e.g., the fraction of nodes that fail



$0 < f < f_c$:

There is a giant component.

$P_\infty \sim |f - f_c|^\beta$

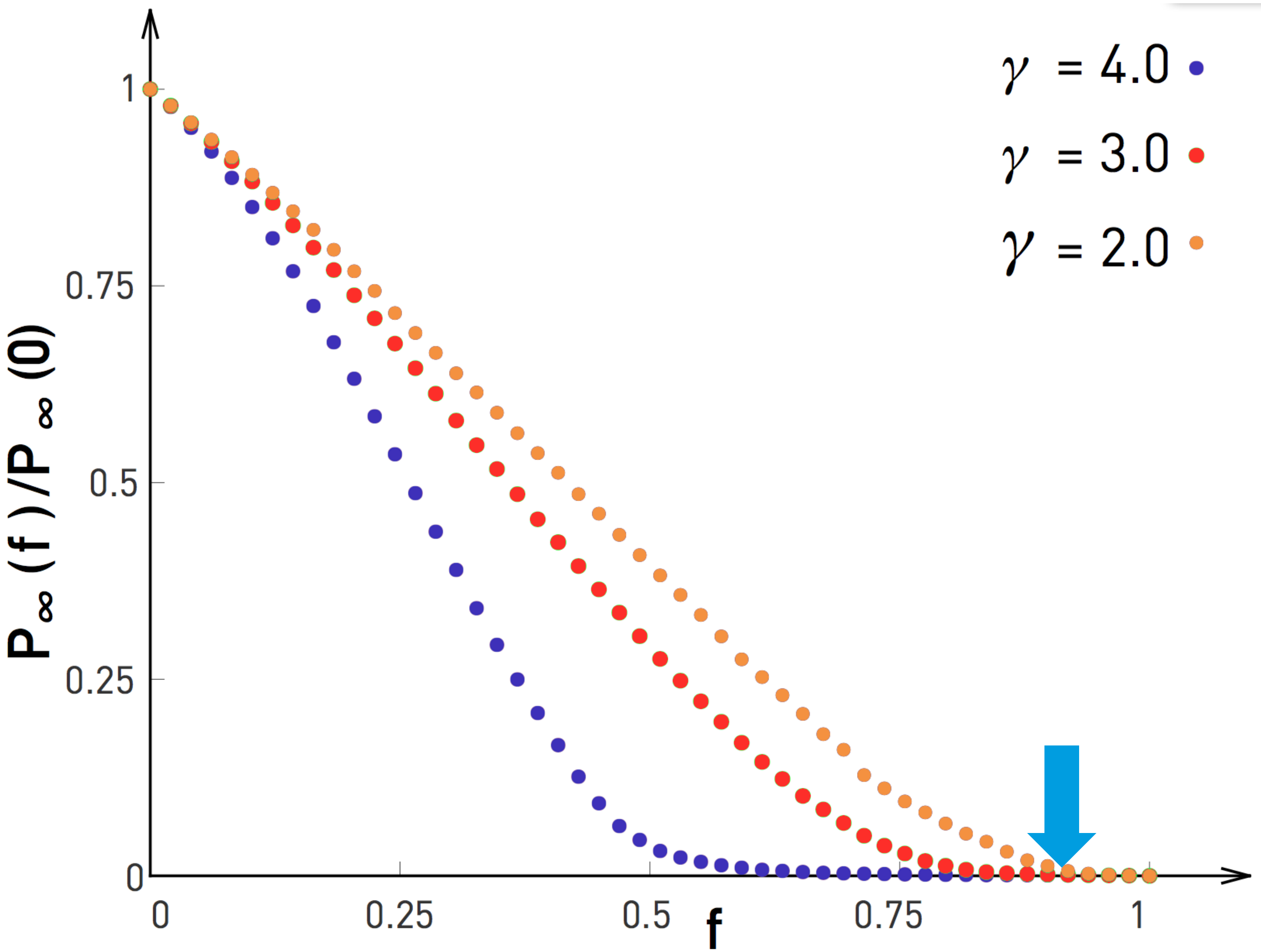
$f = f_c$:

The giant component vanishes.

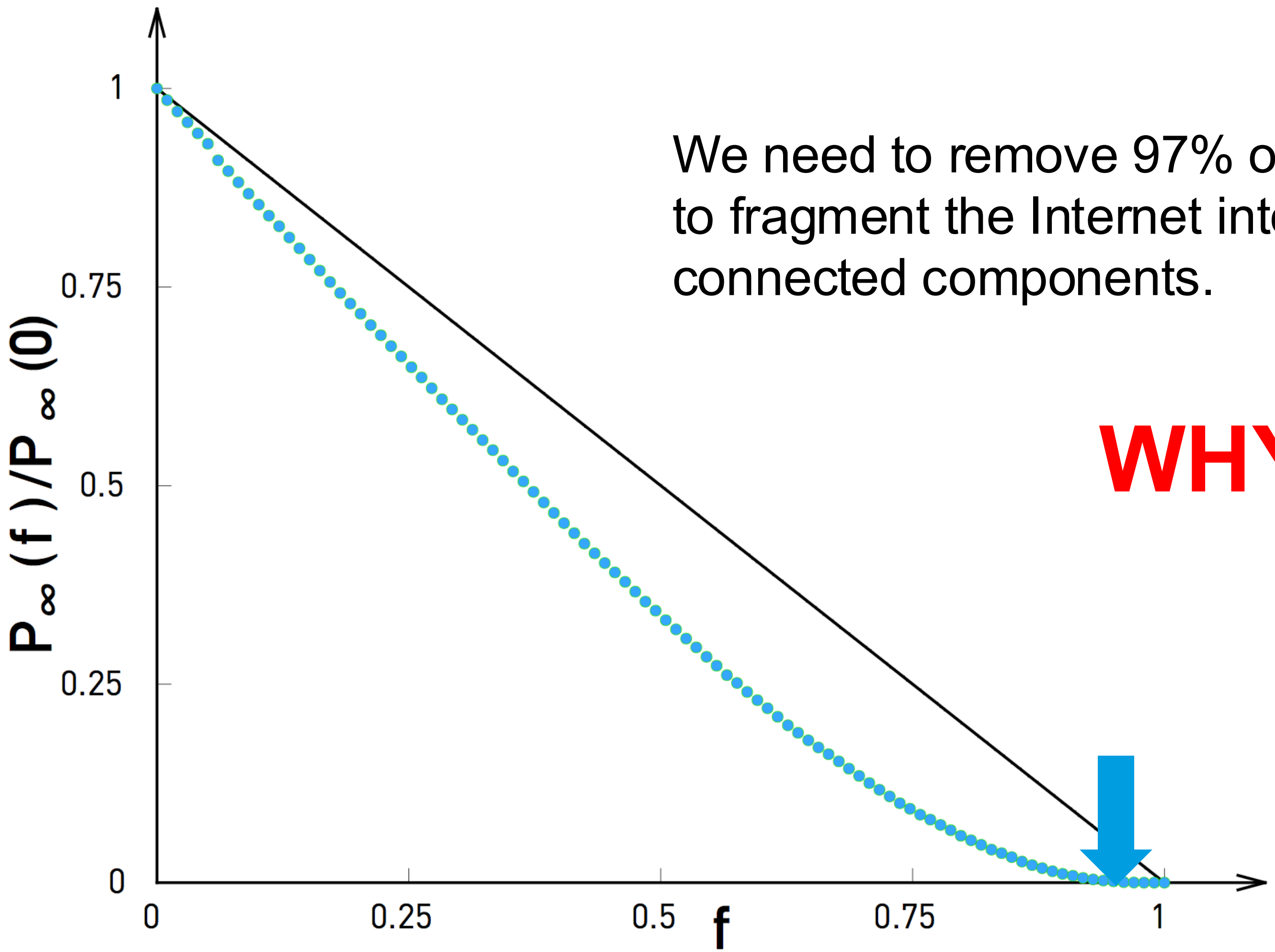
$f > f_c$:

The lattice breaks into many tiny components.

Scale-free networks (simulations using models)



Scale-free networks (Internet / Simulations)



We need to remove 97% of the routers to fragment the Internet into dis-connected components.

WHY?

Analytical insights: Molloy-Reed criterion (1995)

For details see for instance, Barrat et al. Dynamical processes in complex networks, 2008, section 6.4.

- In general the breakdown of any complex network is not gradual: it occurs abruptly in a phase transition at some critical fraction of nodes removed.
- A (infinite) randomly wired complex network (without loops!) w/ arbitrary degree distribution shows a giant component if

$$\langle k^2 \rangle - \langle k \rangle > \langle k \rangle$$

or

$$\frac{\langle k^2 \rangle}{\langle k \rangle} > 2$$

nearest neighbors average degree

Analytical insights: Molloy-Reed criterion (1995)

For details see for instance, Barrat et al. Dynamical processes in complex networks, 2008, section 6.4.

- In general the breakdown of any complex network is not gradual: it occurs abruptly in a phase transition at some critical fraction of nodes removed.
- Intuition: To form a chain each individual must hold the hand of two other individuals. This means that average degree of our neighbors should be > 2.

$$\langle k^2 \rangle - \langle k \rangle > \langle k \rangle$$

or

$$\frac{\langle k^2 \rangle}{\langle k \rangle} > 2$$

nearest neighbors average degree

(Cohen et al., 2000)

Molloy-Reed criterion

For details see for instance, Barrat et al. Dynamical processes in complex networks, 2008, section 6.4.

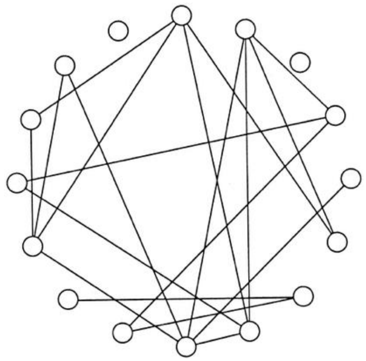
- From this, it can be shown that the giant component is present if the fraction f of removed nodes is

$$f < f_C = 1 - \frac{1}{\frac{\langle k^2 \rangle}{\langle k \rangle} - 1}$$

For instance for a random network we have $\langle k^2 \rangle = \langle k \rangle (1 + \langle k \rangle)$

getting

$$f_c^{ER} = 1 - \frac{1}{\langle k \rangle}$$



(Cohen et al., 2000)

Molloy-Reed criterion

- From this, it can be shown that the giant component is present if the fraction f of removed nodes is

$$f < f_C = 1 - \frac{1}{\frac{\langle k^2 \rangle}{\langle k \rangle} - 1}$$

- As you may remember, in ***scale-free nets*** with $2 \leq \gamma < 3$ the second moment $\langle k^2 \rangle \square \times$
- Thus, when $2 \leq \gamma < 3$, we get a giant component whenever ***p=1-f>0*** (or ***f<1***), i.e., **always**!!

(Cohen et al., 2000)

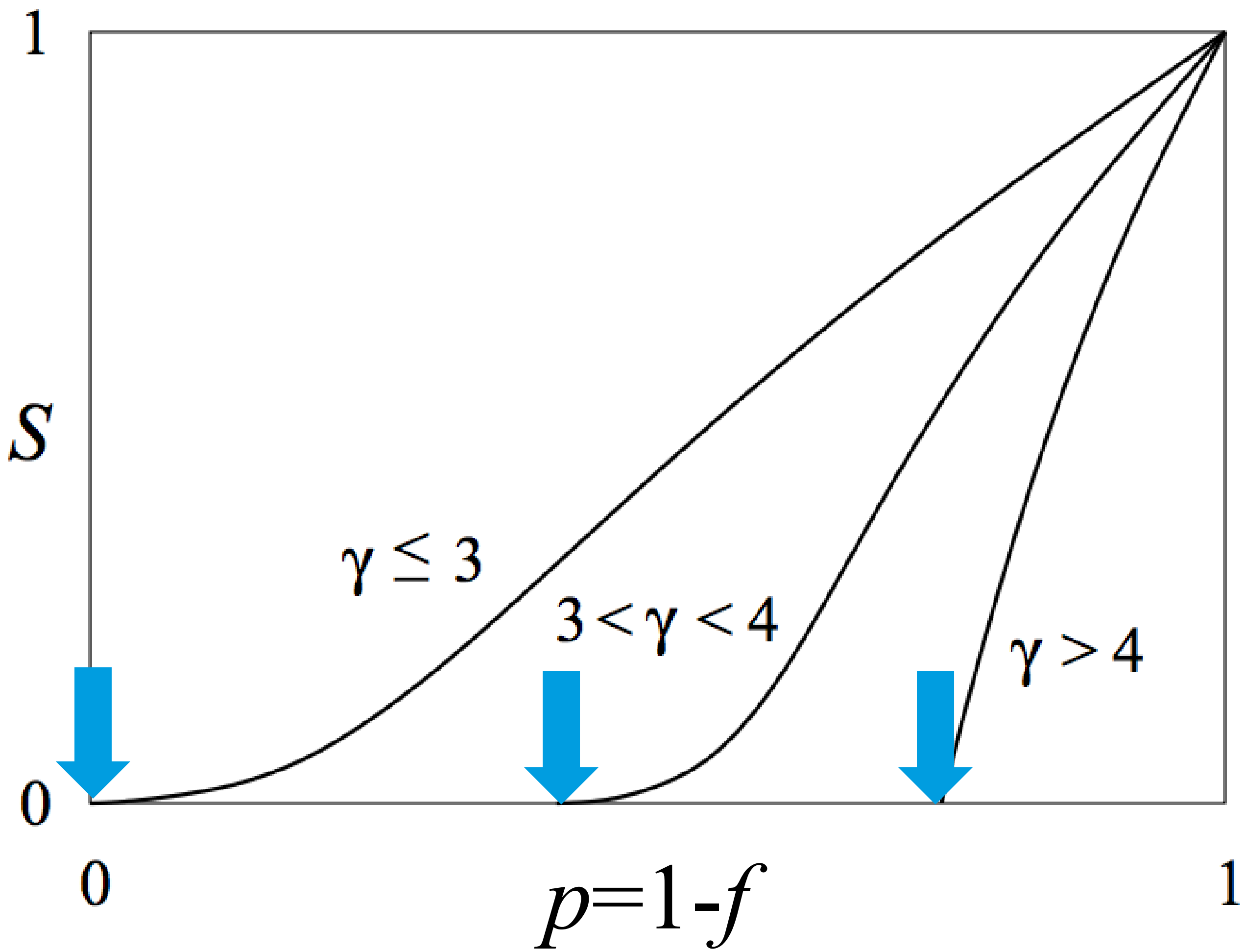
Molloy-Reed criterion

- From this, it can be shown that the giant component is present if the fraction f of removed nodes is

This means that
to fragment a scale-free network
we must remove all its nodes!!

- As you may remember, in ***scale-free nets*** with $2 \le \gamma < 3$ the second moment $\langle k^2 \rangle \to \infty$
- Thus, when $2 \le \gamma < 3$, we get a giant component whenever ***p=1-f>0*** (or ***f<1***), i.e., **always**!!

Scale-free networks (analytics)



Finite size effects

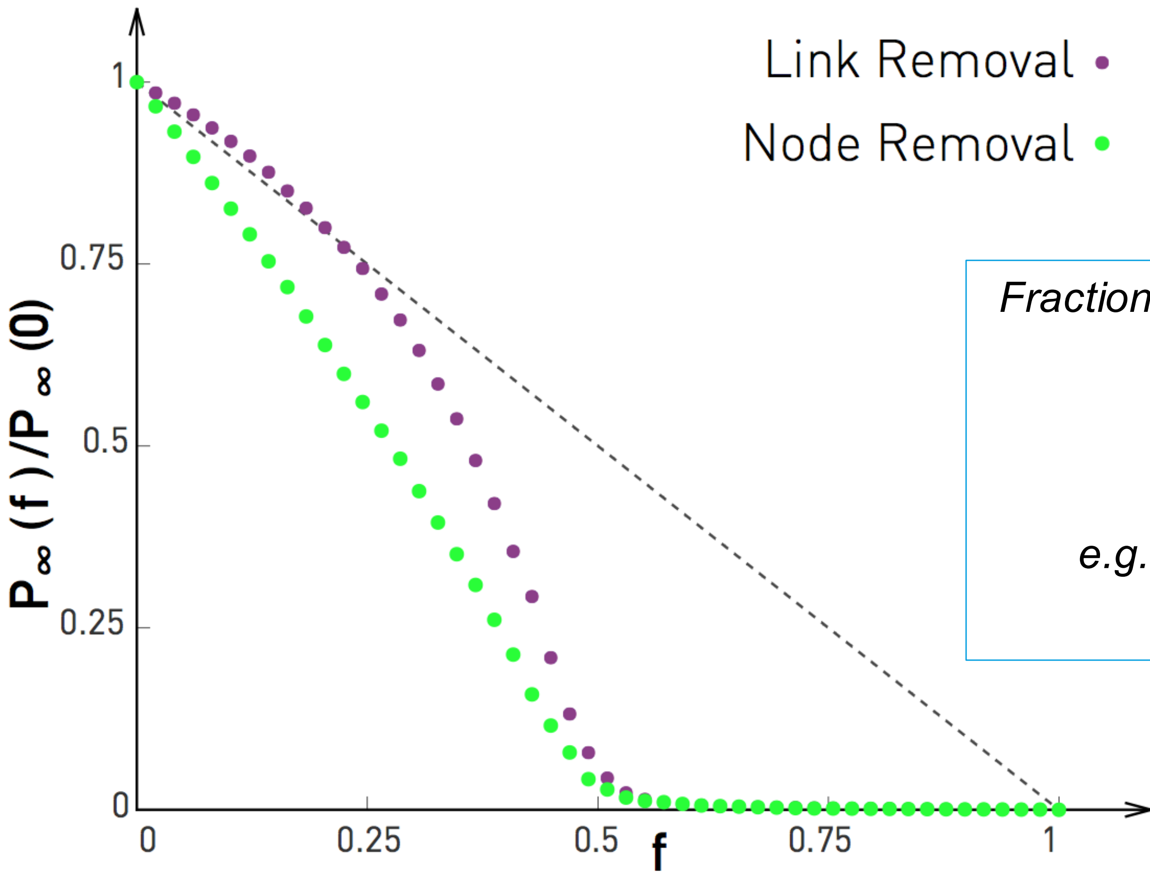
- For finite networks, naturally, $<k^2>$ does not diverge. Also the "abrupt" transition becomes smoother. If you consider a finite N and a power-law with $2 \leq \gamma < 3$ one gets

$$f_c \approx 1 - \frac{3-\gamma}{\gamma-2} k_{\min}^{2-\gamma} k_{\max}^{\gamma-3}$$

- Example: for N=10^3, minimum degree (k_{min}) = 1, γ=2.5, we get a maximum degree (k_{max}) ~ $N^{1/(1-\gamma)}$ ~100 and a critical f_c = 0.9<1.

What happens if we randomly remove links instead of nodes?

ex: Random network with <k>=2



Fraction of removed nodes / links:

$$f = 1\text{-}p$$

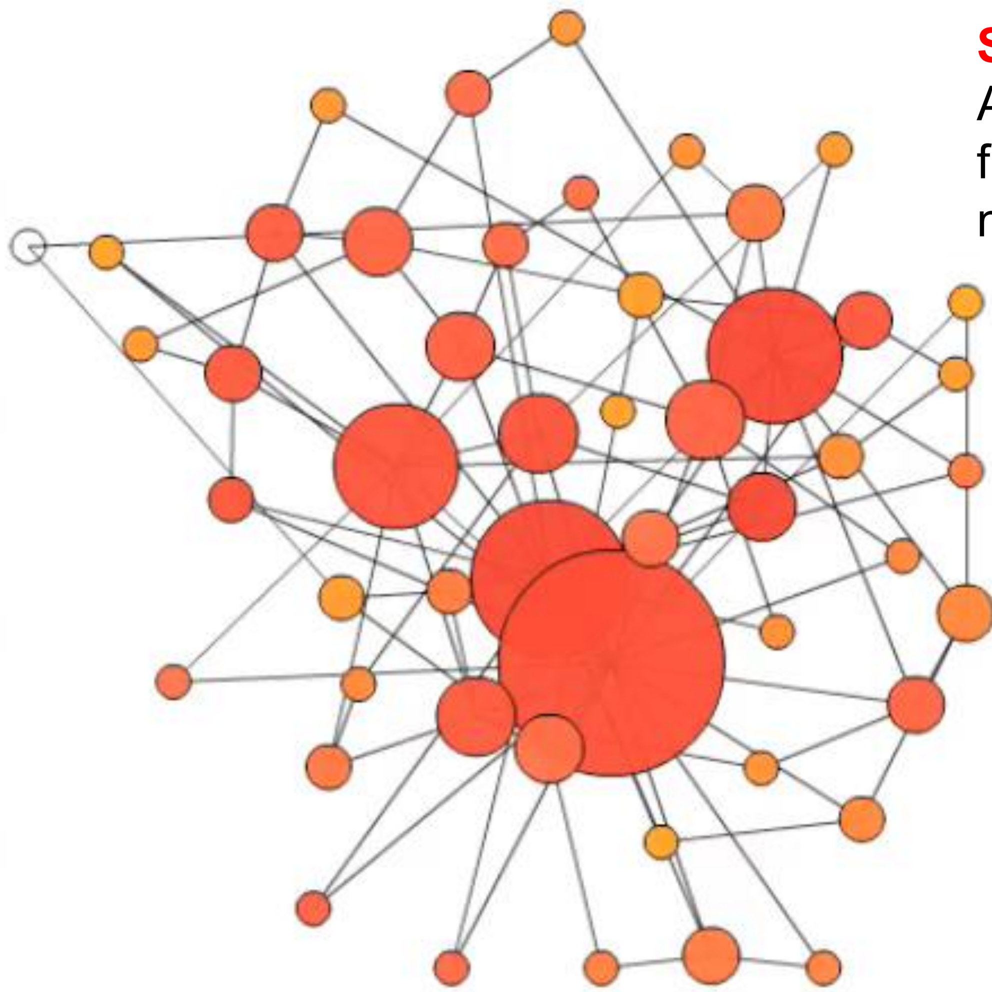
e.g., the fraction of nodes or links that fail

On average each node removes <k> links. Hence the removal of a fraction f of nodes is equivalent to a removal of a fraction f<k> links

Conclusion

- We discussed a fundamental property of real world networks: robustness to random failures
- The breakdown threshold of a network depends of $<k>$ and on $<k^2>$, which are uniquely defined by the degree distribution.
- For $\gamma<3$ the breakdown threshold rapidely converges to one, which means that we have to remove almost all nodes such that the network falls apart.

Achilles' heel of scale-free networks

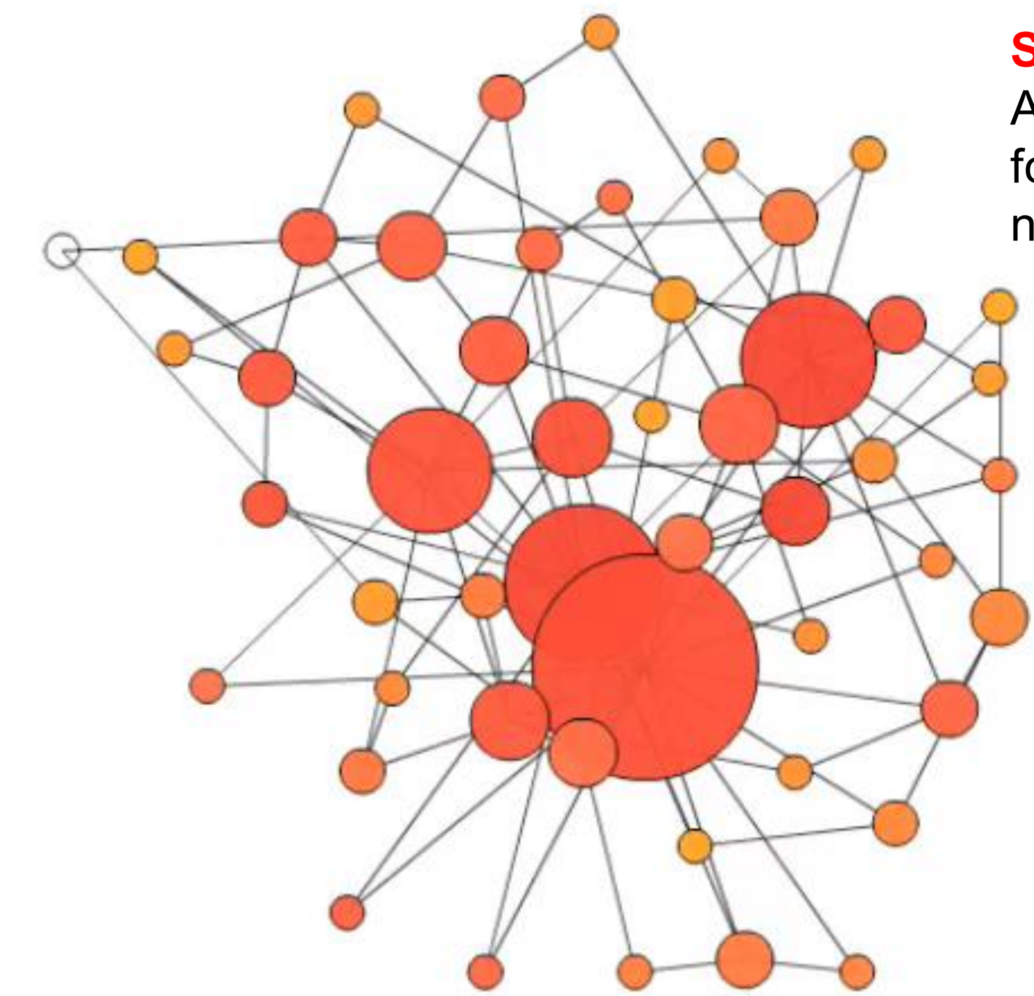


Scale-free networks under attack

Attack first the highest-degree node, followed by the next highest degree node, and so on.

Video by Dashun Wang

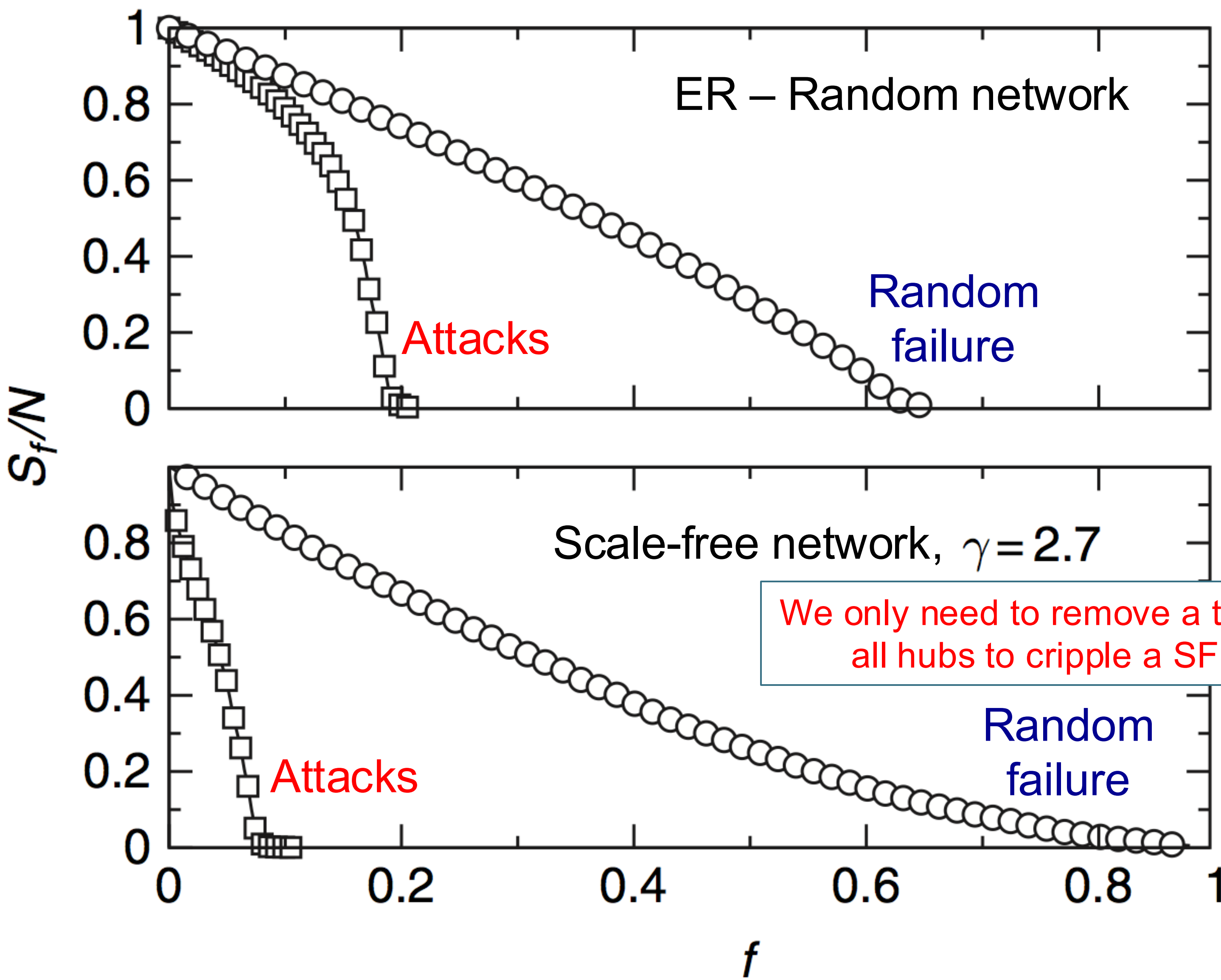
Achilles' heel of scale-free networks



Scale-free networks under attack
Attack first the highest-degree node, followed by the next highest degree node, and so on.

Video by Dashun Wang

Robustness against targeted attacks



We only need to remove a tiny fraction of all hubs to cripple a SF network.

Robustness beyond the degree distribution

NETWORK	RANDOM FAILURES (REAL NETWORK)	RANDOM FAILURES (RANDOMIZED NETWORK)	ATTACK (REAL NETWORK)
Internet	0.92	0.84	0.16
WWW	0.88	0.85	0.12
Power Grid	0.61	0.63	0.20
Mobile-Phone Call	0.78	0.68	0.20
Email	0.92	0.69	0.04
Science Collaboration	0.92	0.88	0.27
Actor Network	0.98	0.99	0.55
Citation Network	0.96	0.95	0.76
E. Coli Metabolism	0.96	0.90	0.49
Yeast Protein Interactions	0.88	0.66	0.06

Robustness beyond the degree distribution

- What's the impact of clustering on the robustness of a network?

  Holme et al. (2002)

- What's the role of degree-degree correlations?

- Resilience and robustness of weighted networks?

  See Dall'Asta et al., 2005, 2006

- Beyond topology: Failure of a single node leads to a redistribution of traffic on the network which may trigger subsequent overloads and failure of the next most-loaded node…