Network theory

Part II

Complexity in Social Systems
AA 2023/2024
Maxime Lucas
Lorenzo Dall'Amico

CENTAI

ISI Foundation

Recap last lecture

<u>Types of networks</u>

Un/directed
Weighted
Bipartite

<u>Concepts</u>

<u>Degree</u>
Weights
<u>Adjacency matrix</u>
<u>Paths/components</u>
<u>Clustering coefficient</u>
Centralities

<u>Properties</u>

<u>Scale-free</u>
<u>Sparseness</u>
<u>Connectedness</u>
<u>Small-worldness</u>
<u>High clustering</u>

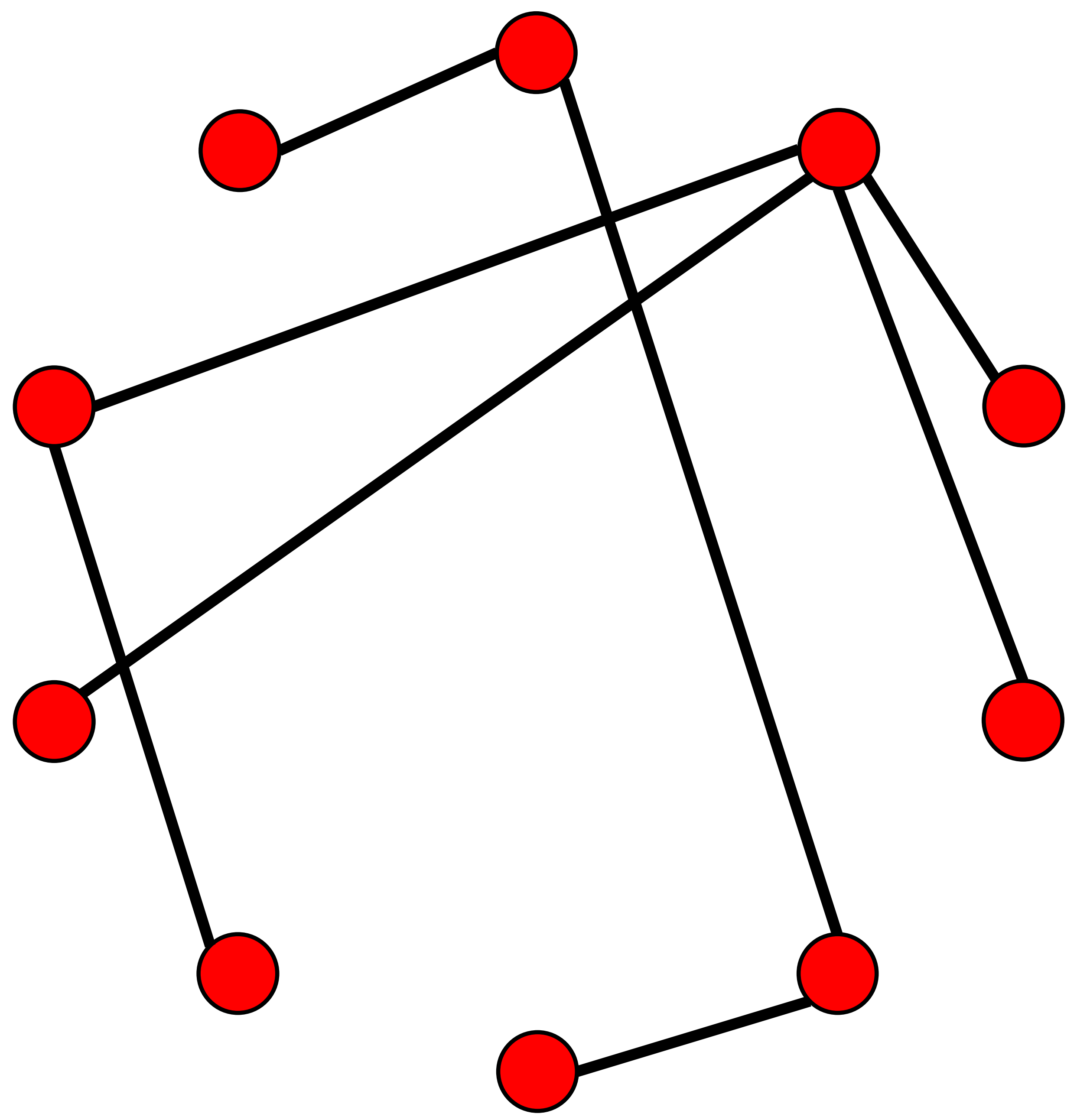
Today's topic: Random models

<u>What?</u>

<u>Ensembles of networks with constraints</u>
but otherwise maximally random

<u>Why?</u>

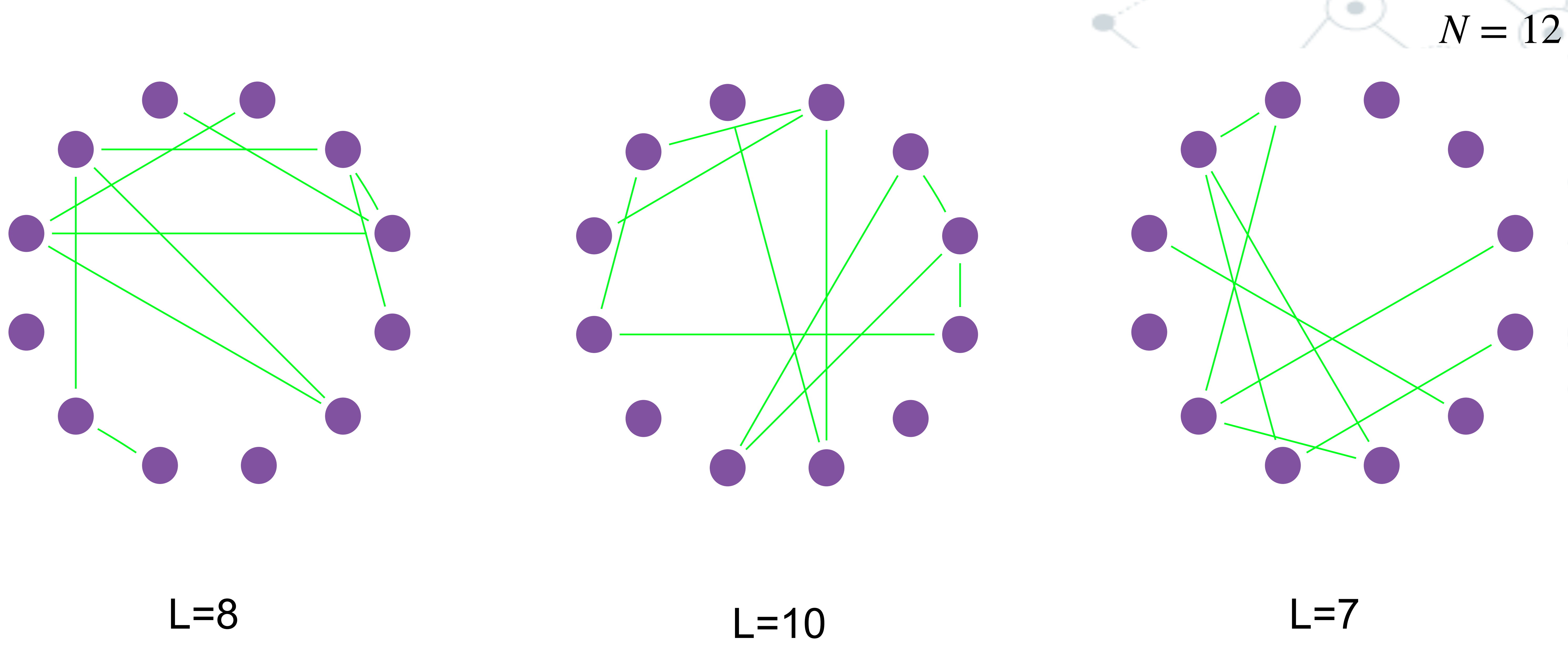
They help us understand
what is **<u>structure</u>**
and what is **<u>random</u>**



Erdos-Renyi random network model

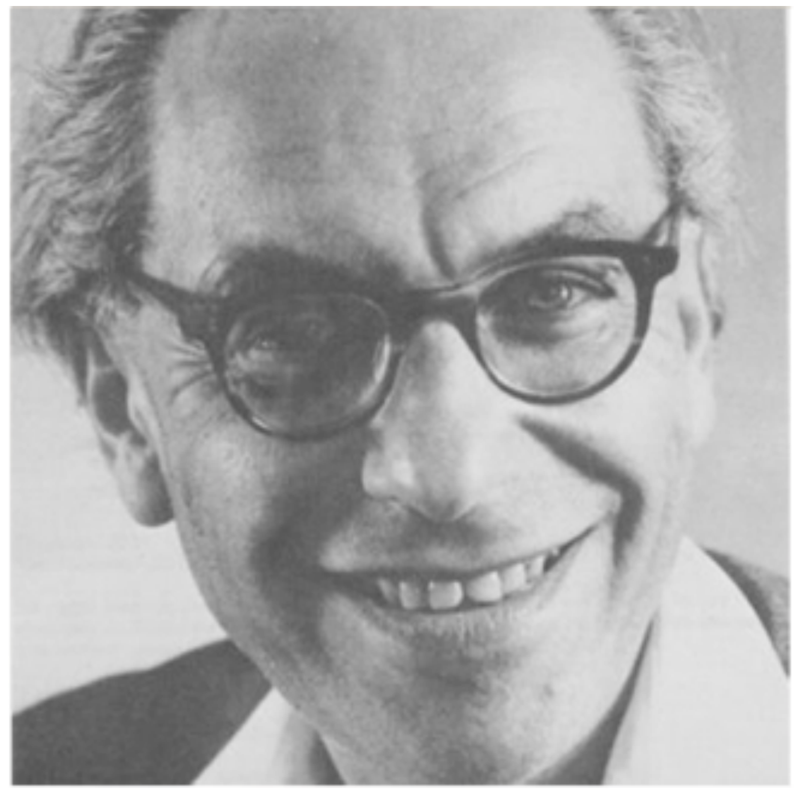
***G*(*N*, *L*) Model**

N labeled nodes are connected with *L* randomly placed links. Erdős and Rényi used this definition in their string of papers on random networks [2-9].



$p = 1/6$

Pál Erdös
(1913-1996)



Alfréd Rényi
(1921-1970)

Erdös-Rényi model (1960)

Erdos-Renyi random network model

Probability of a network in the ensemble

probability to have exactly L links in a network of N nodes and probability p

The maximum number of links
in a network of N nodes.

$$P(L) = \binom{\binom{N}{2}}{L} p^L (1-p)^{\frac{N(N-1)}{2} - L}$$

Number of different ways we can
choose L links among all potential links.

Erdos-Renyi random network model

Average degree

$$P(L) = \binom{\binom{N}{2}}{L} p^L (1-p)^{\frac{N(N-1)}{2}L}$$

Micro-recap

$$P(x) = \binom{T}{x} p^x (1-p)^{T-x}$$

$$< x > = Tp$$

$$< x^2 > = p(1-p)T + p^2T^2$$

$$\sigma_x = [p(1-p)T]^{1/2}$$

Average degree

$$< L > = \sum_{L=0}^{\binom{N}{2}} LP(L) = p\frac{N(N-1)}{2}$$

$$< k > = 2L/N = p(N-1)$$

We are constraining the average degree!
So if we want SPARSENESS, we need small p

Erdos-Renyi random network model

Degree distribution

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

$$<k> = p(N-1) \qquad \sigma_k^2 = p(1-p)(N-1)$$

For large N

$$\frac{\sigma_k}{<k>} = \frac{1-p}{p}\frac{1}{N-1}]^{1/2} \simeq \frac{1}{(N-1)^{1/2}}$$

For large N and small k:

<k> << N

$$\binom{N-1}{k} = \frac{(N-1)!}{k!(N-1-k)!} = \frac{(N-1)(N-1-1)(N-1-2)...(N-1-k+1)(N-1-k)!}{k!(N-1-k)!} = \frac{(N-1)^k}{k!}$$

$$\ln[(1-p)^{(N-1)-k}] = (N-1-k)\ln(1-\frac{<k>}{N-1}) = -(N-1-k)\frac{<k>}{N-1} = -<k>(1-\frac{k}{N-1}) \cong -<k>$$

$$(1-p)^{(N-1)-k} = e^{-<k>}$$

$$\ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n = x - \frac{x^2}{2} + \frac{x^3}{3} - ... \quad \text{for} \quad |x| \leq 1$$

Erdos-Renyi random network model

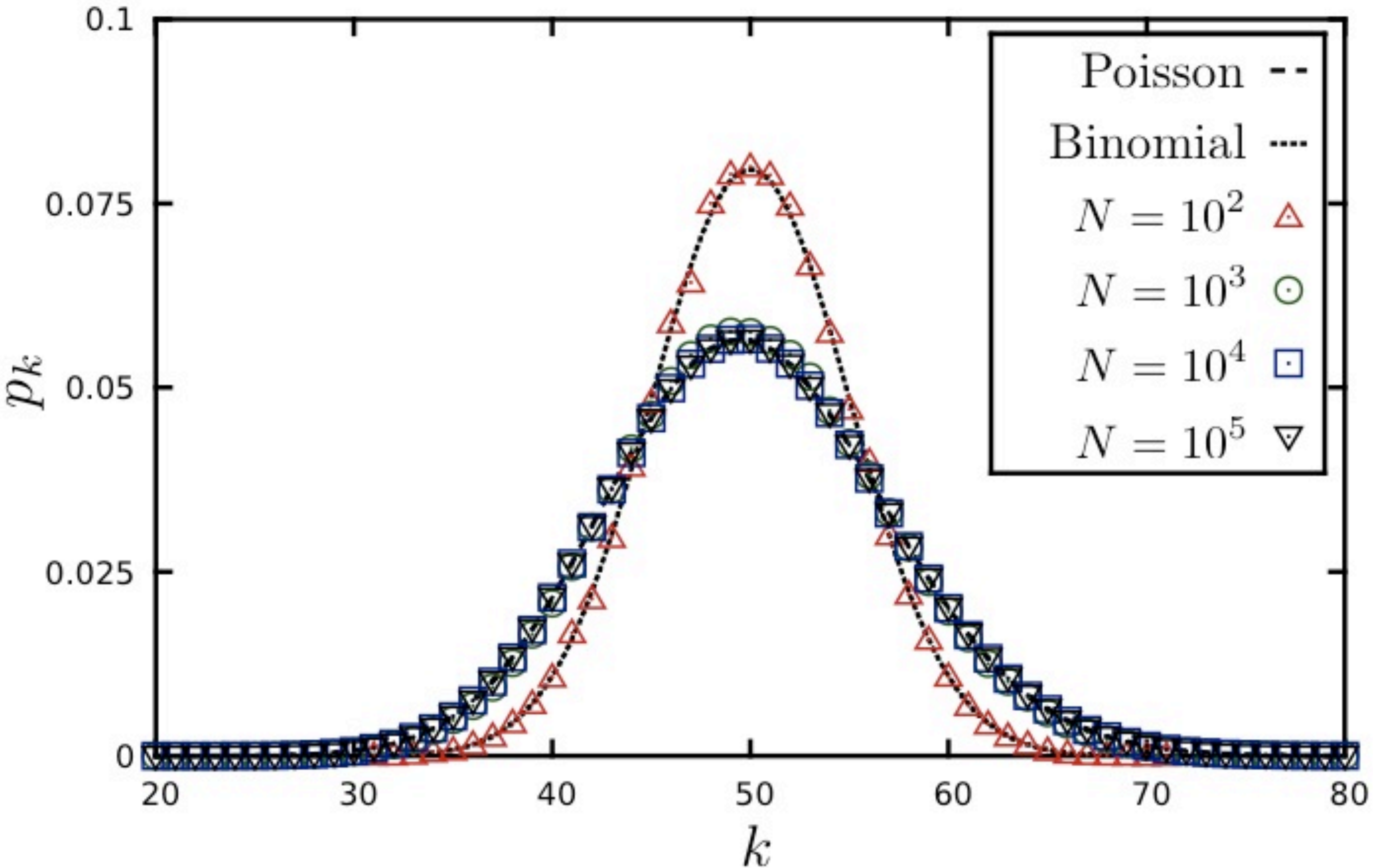
Poisson limit of degree distribution

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

$$p(k) = e^{-<k>} \frac{<k>^k}{k!}$$

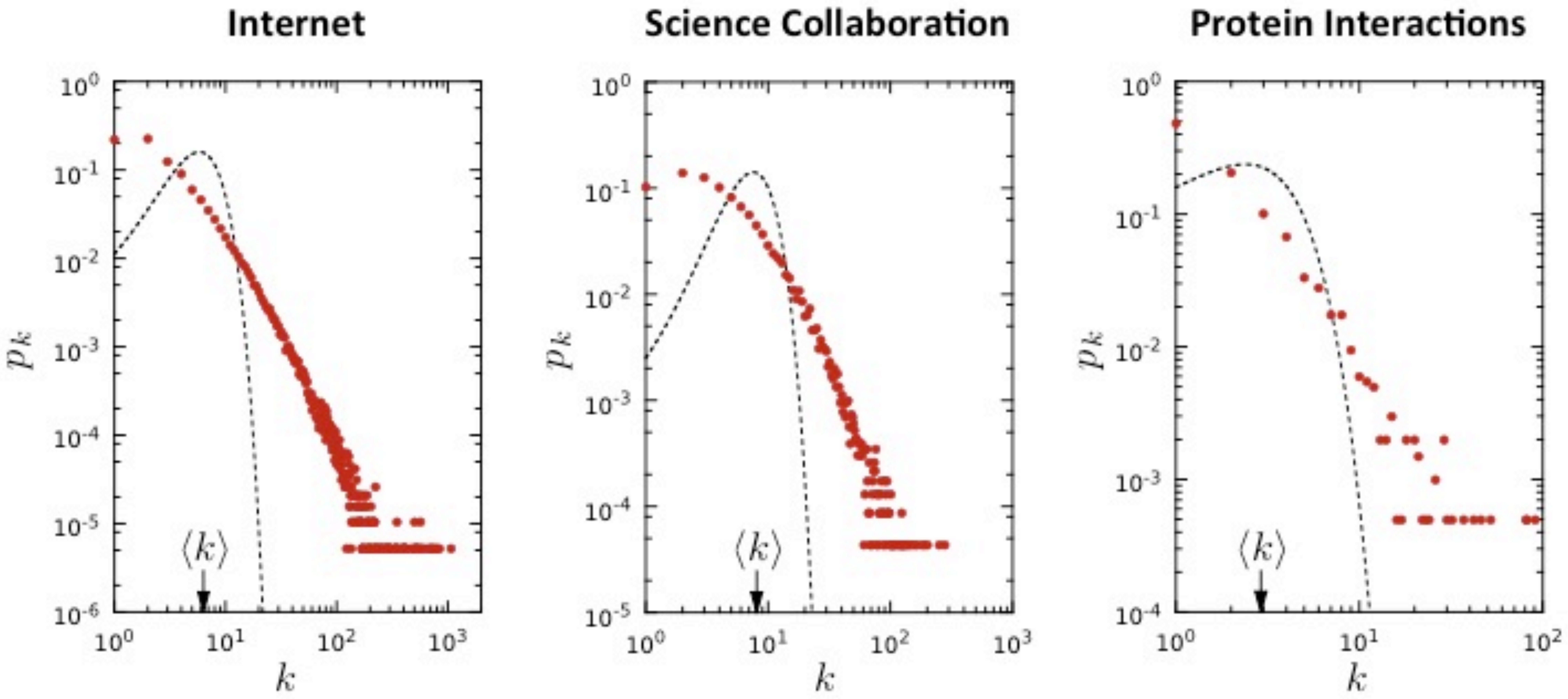
Does not depend on N

Both peak around <k>
Dispersion controlled by <k> or p



Erdos-Renyi random network model

And nope.. ER does not reproduce realistic degree distributions



Dashed line: Poisson with <k> computed from data

Erdos-Renyi random network model

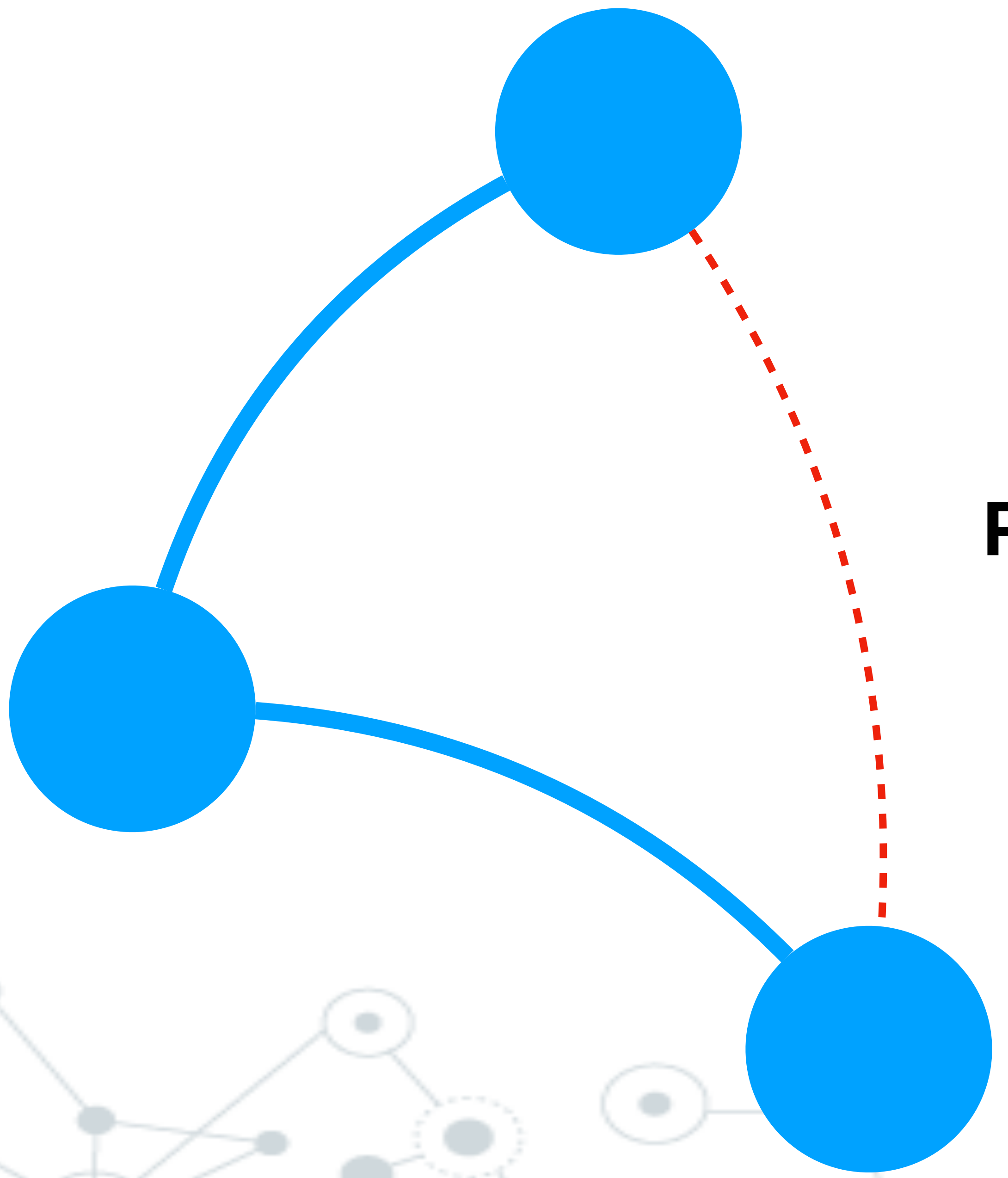
List of results:

- we can reproduce sparseness using N and p
- Degree distribution is binomial/poisson NOT broad/powerlaw

Erdos-Renyi random network model

What about clustering?

$$C_i = \frac{2L_i}{k_i(k_i - 1)}$$



Probability?

$$\langle L_i \rangle = p\frac{k_i(k_i - 1)}{2}$$

$$C_i = p = \frac{<k>}{N}$$

We CAN constrain the clustering (but uniform)!
So if we want high clustering, we need large p!

We are constraining the average degree!
So if we want SPARSENESS, we need small p

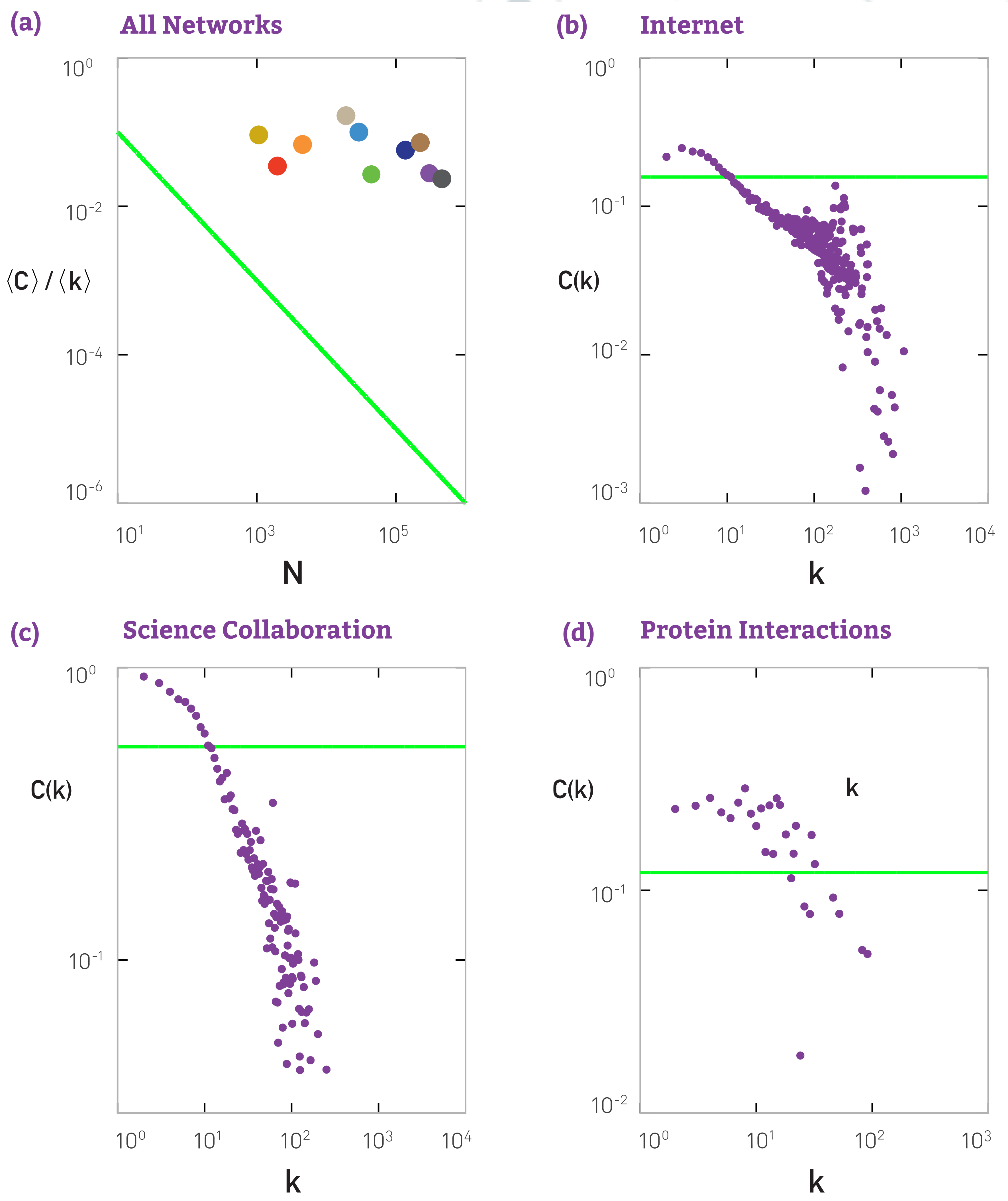
Erdos-Renyi random network model

List of results:

- We can reproduce sparseness using N and p
- Degree distribution is binomial/poisson NOT broad/powerlaw
- We can reproduce high clustering, but not low density (or viceversa)

Network	Size	$\langle k \rangle$	ℓ	ℓ_{rand}	C	C_{rand}	Reference
WWW, site level, undir.	153 127	35.21	3.1	3.35	0.1078	0.00023	Adamic, 1999
Internet, domain level	3015–6209	3.52–4.11	3.7–3.76	6.36–6.18	0.18–0.3	0.001	Yook *et al.*, 2001a, Pastor-Satorras *et al.*, 2001
Movie actors	225 226	61	3.65	2.99	0.79	0.00027	Watts and Strogatz, 1998
LANL co-authorship	52 909	9.7	5.9	4.79	0.43	1.8×10^{-4}	Newman, 2001a, 2001b, 2001c
MEDLINE co-authorship	1 520 251	18.1	4.6	4.91	0.066	1.1×10^{-5}	Newman, 2001a, 2001b, 2001c
SPIRES co-authorship	56 627	173	4.0	2.12	0.726	0.003	Newman, 2001a, 2001b, 2001c
NCSTRL co-authorship	11 994	3.59	9.7	7.34	0.496	3×10^{-4}	Newman, 2001a, 2001b, 2001c
Math. co-authorship	70 975	3.9	9.5	8.2	0.59	5.4×10^{-5}	Barabási *et al.*, 2001
Neurosci. co-authorship	209 293	11.5	6	5.01	0.76	5.5×10^{-5}	Barabási *et al.*, 2001
E. coli, substrate graph	282	7.35	2.9	3.04	0.32	0.026	Wagner and Fell, 2000
E. coli, reaction graph	315	28.3	2.62	1.98	0.59	0.09	Wagner and Fell, 2000
Ythan estuary food web	134	8.7	2.43	2.26	0.22	0.06	Montoya and Solé, 2000
Silwood Park food web	154	4.75	3.40	3.23	0.15	0.03	Montoya and Solé, 2000
Words, co-occurrence	460.902	70.13	2.67	3.03	0.437	0.0001	Ferrer i Cancho and Solé, 2001
Words, synonyms	22 311	13.48	4.5	3.84	0.7	0.0006	Yook *et al.*, 2001b
Power grid	4941	2.67	18.7	12.4	0.08	0.005	Watts and Strogatz, 1998
C. Elegans	282	14	2.65	2.25	0.28	0.05	Watts and Strogatz, 1998

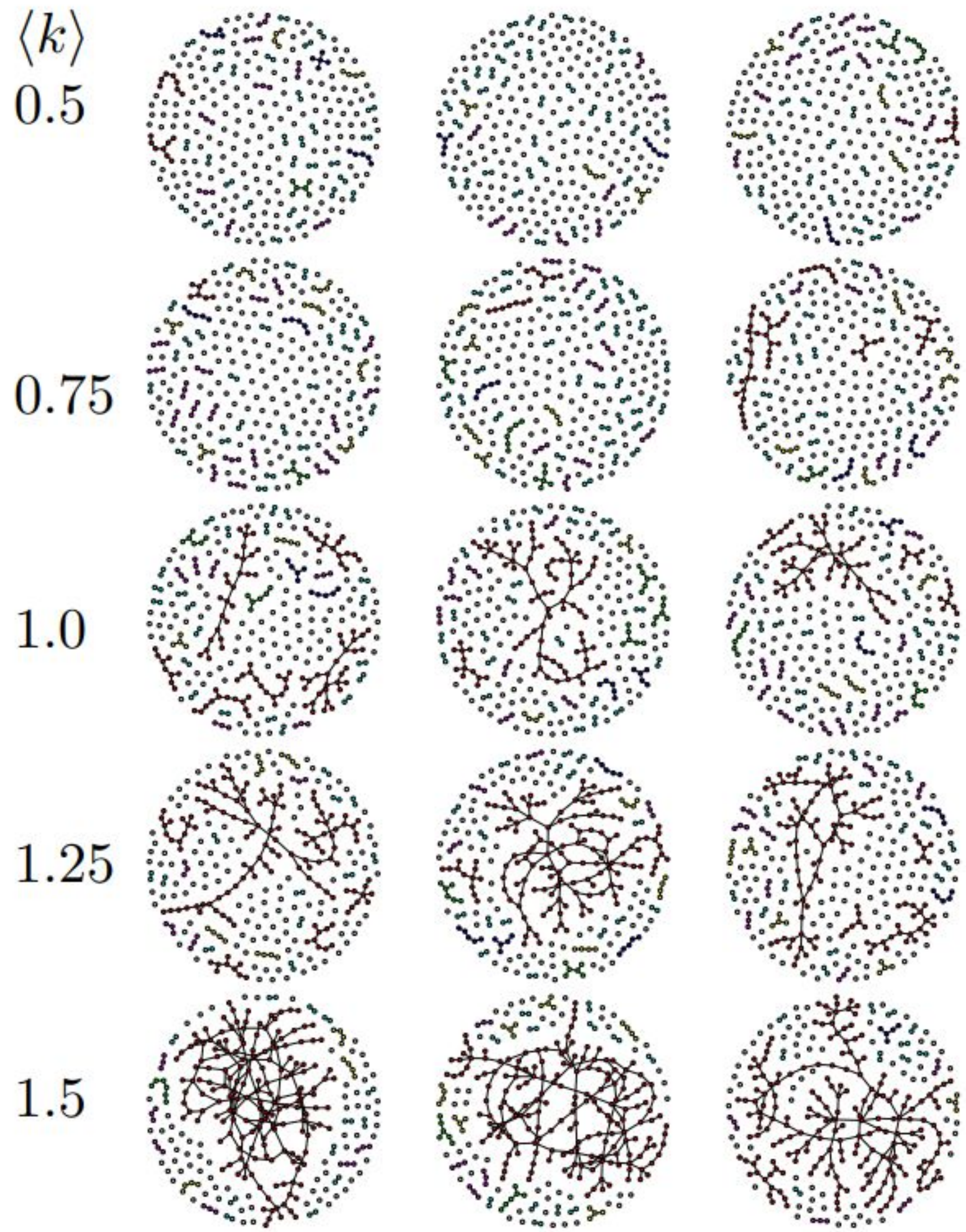
C seems independent of N



Green curve is <C>

Erdos-Renyi random network model

What about connectedness? Let's guess a criterion!



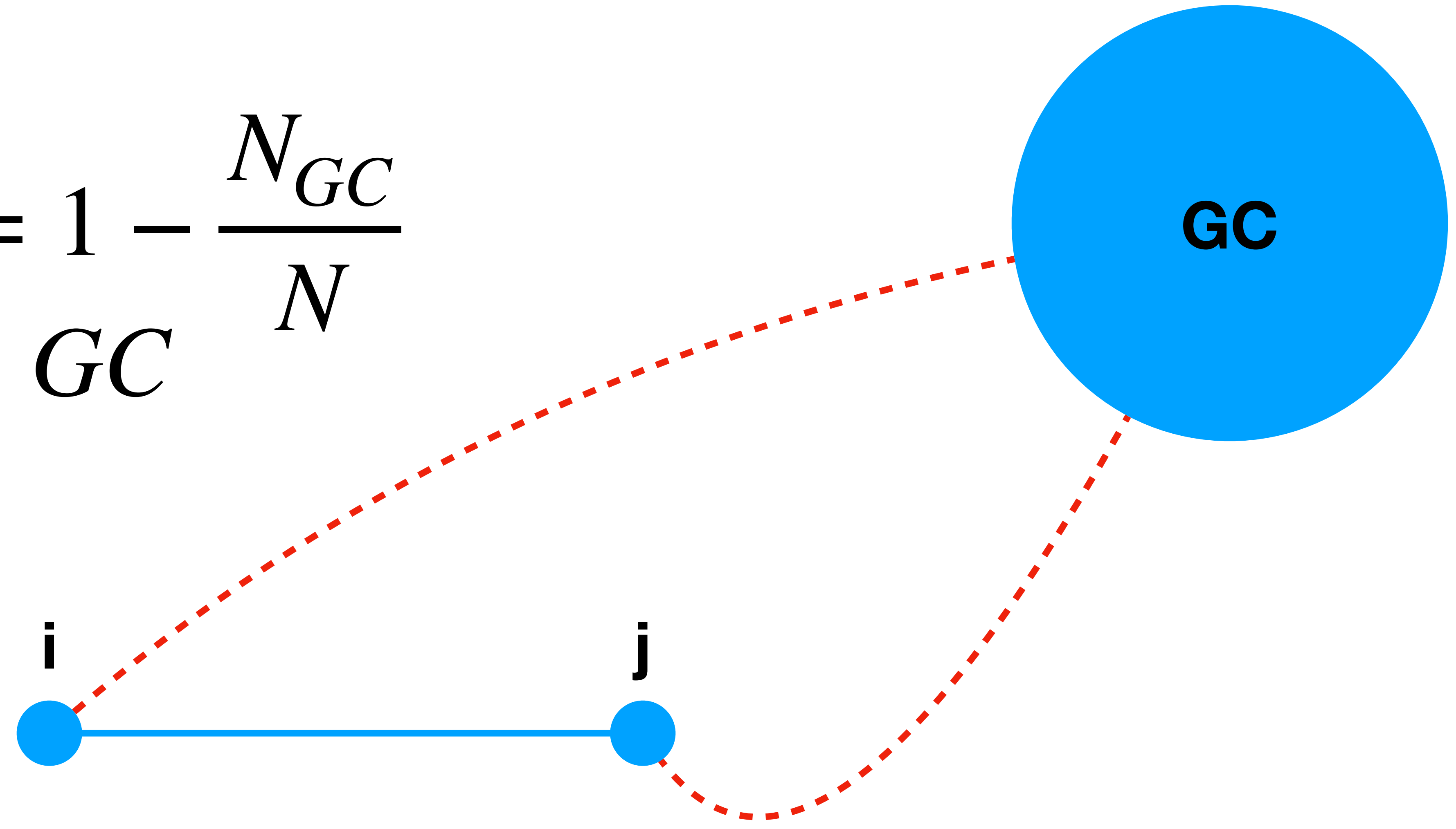
$$< k_c > = 1$$

Necessary! Ok.. but sufficient?

Erdos and Renyi, 1959

$$u = 1 - \frac{N_{GC}}{N}$$

$$i \notin GC$$



Probability that i is not in GC?

1) $i \nsim j \in GC \quad \rightarrow \quad (1-p)$

2) $i \sim j \notin GC \quad \rightarrow \quad (pu)$

Erdos-Renyi random network model

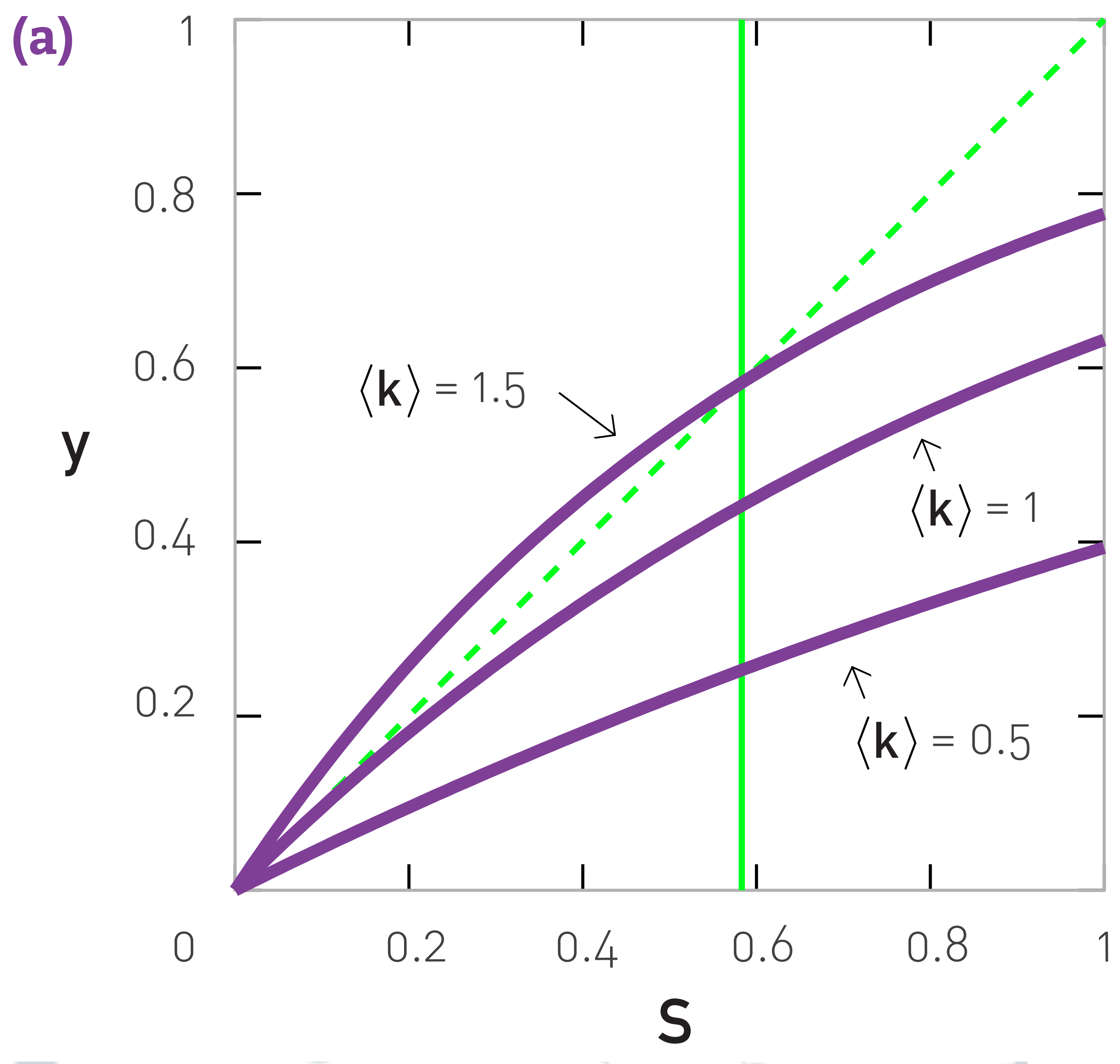
What about connectedness? Let's guess a criterion!

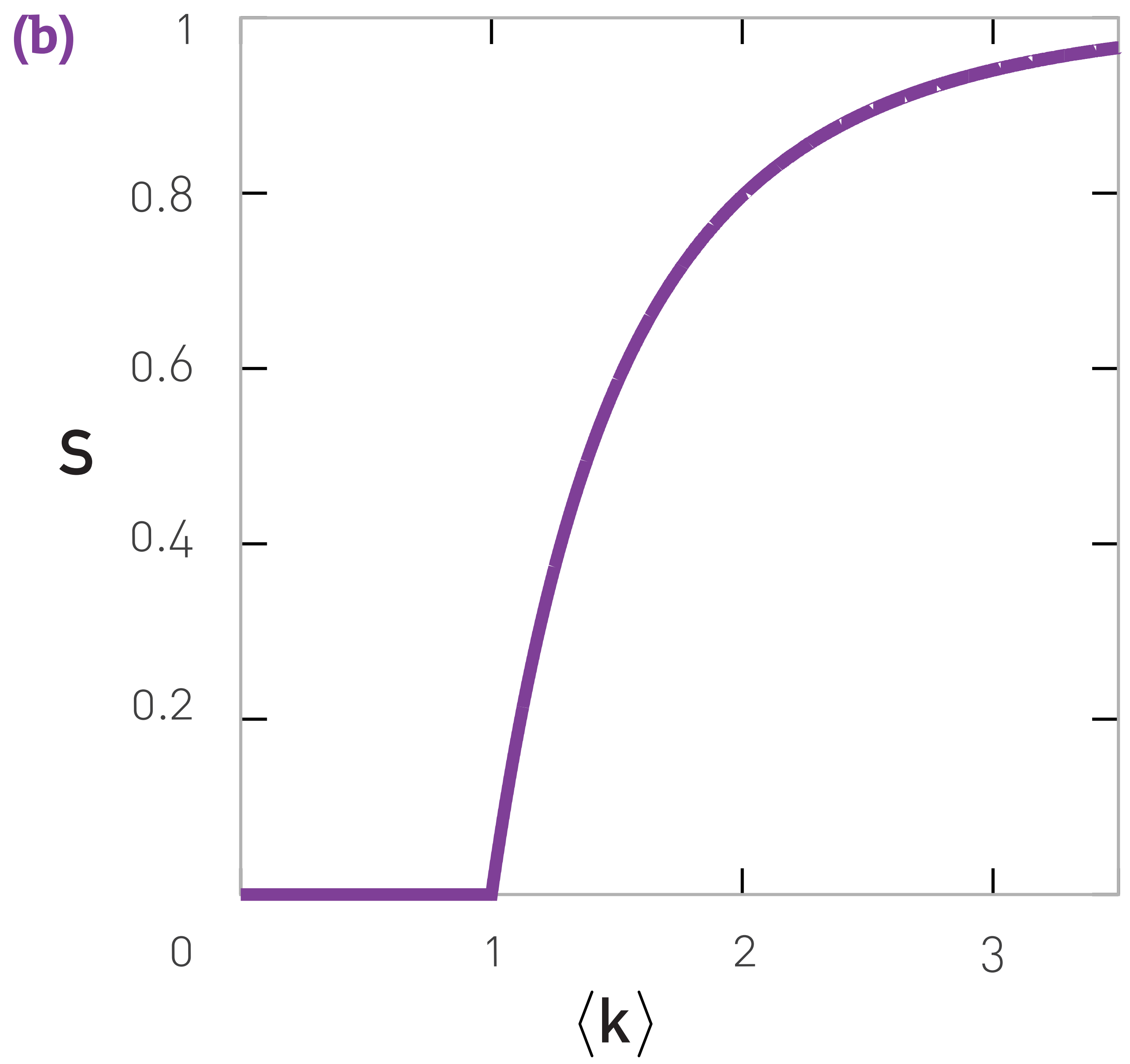
<k> << N

$$(1-p+pu)^{N-1} = u \qquad N_{GC} = N(1-u) \qquad p = \frac{<k>}{N-1} \quad \rightarrow \quad (N-1)\ln\left[1 - \frac{<k>}{N-1}(1-u)\right] = \ln u$$

$$\rightarrow \quad -<k>(1-u) \approx \ln u \quad \rightarrow \quad u \sim e^{-<k>(1-u)}, \quad S = N_{GC}/N = 1-u \quad \rightarrow \quad S = 1 - e^{-<k>S}$$

Does not have a closed solution: let's solve graphically





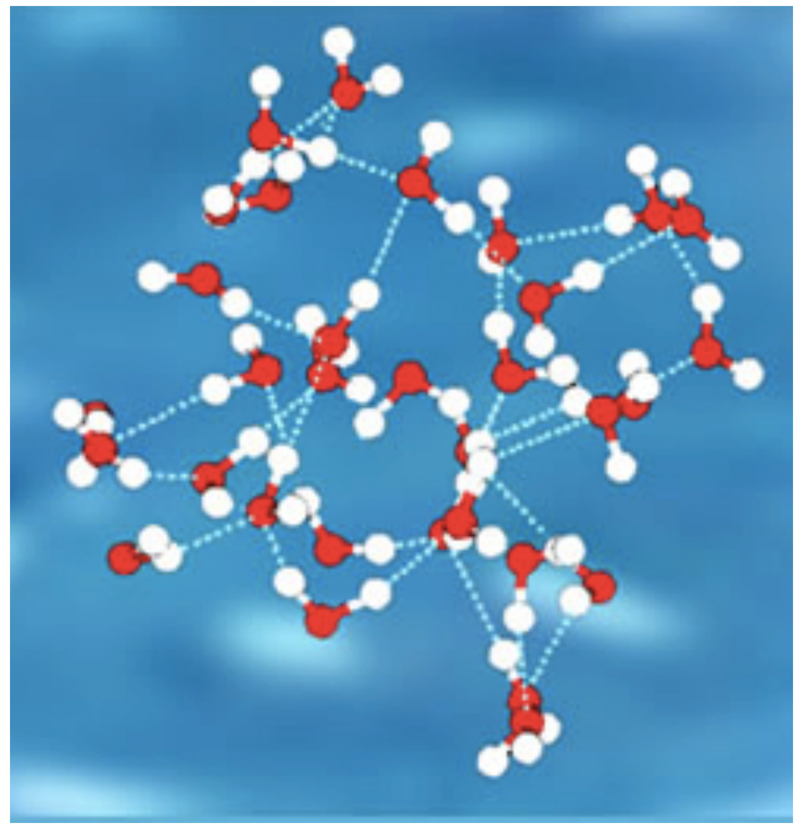
Derive both sides!

$$1 = \left[\frac{d}{dS}(1 - e^{-<k>S})\right]_{S=0}$$

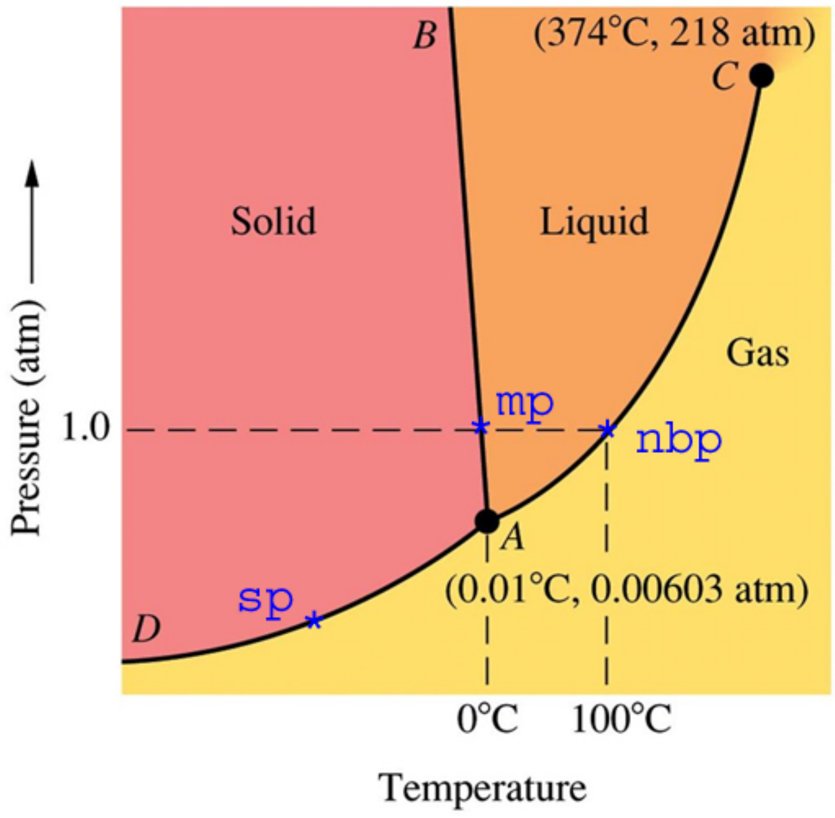
$$<k> = 1$$

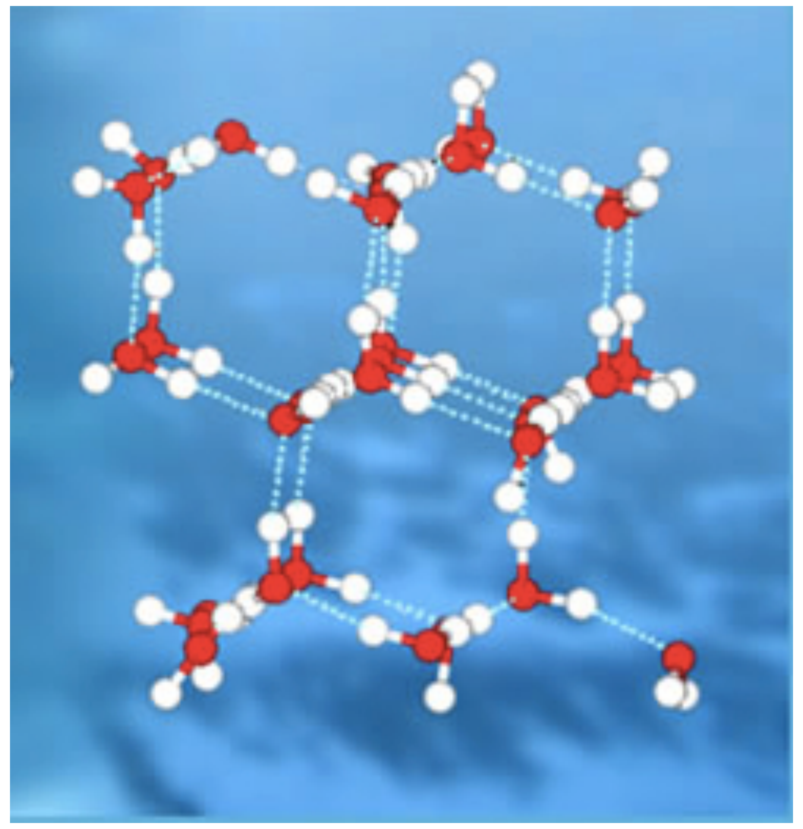
Phase transitions

Water-Ice phase transition

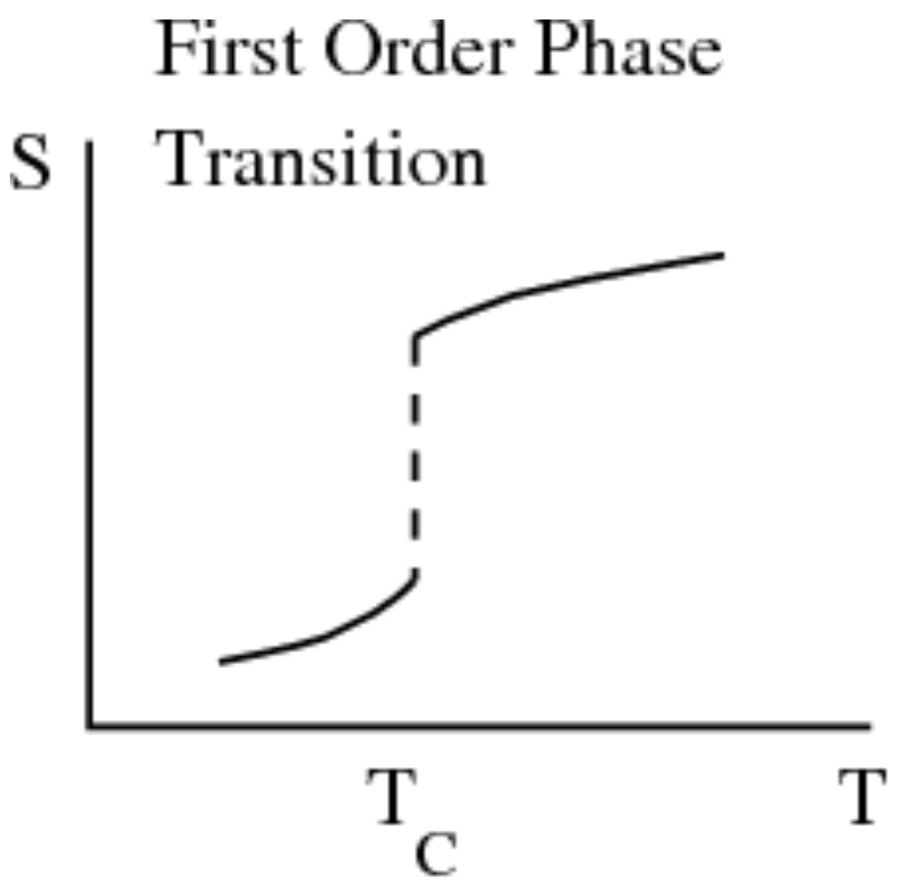


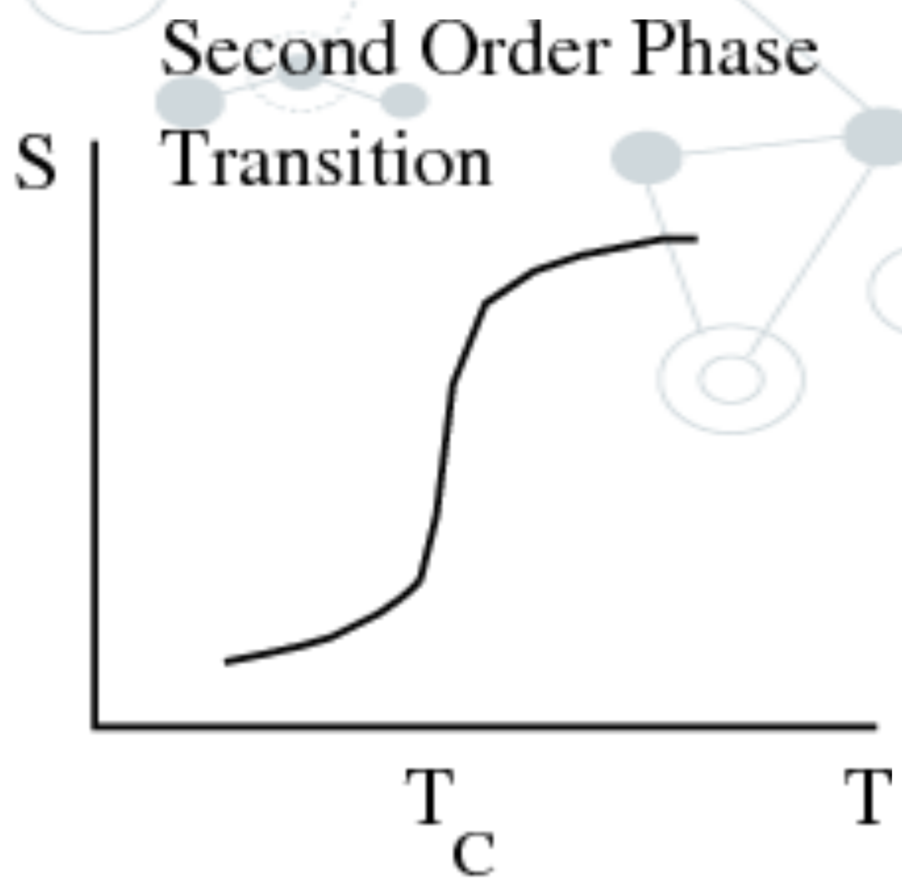
Water

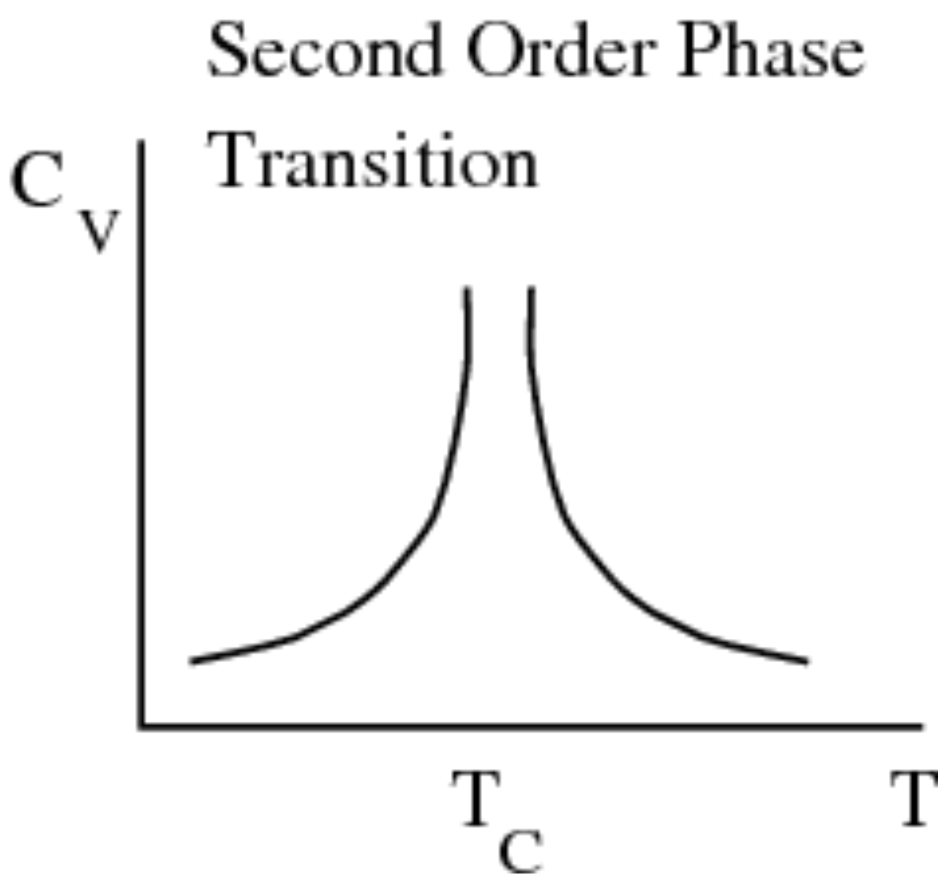




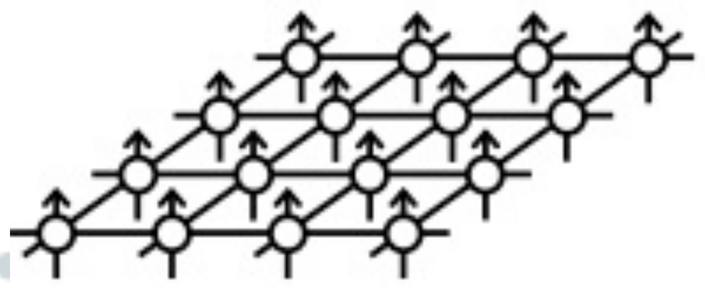
Ice



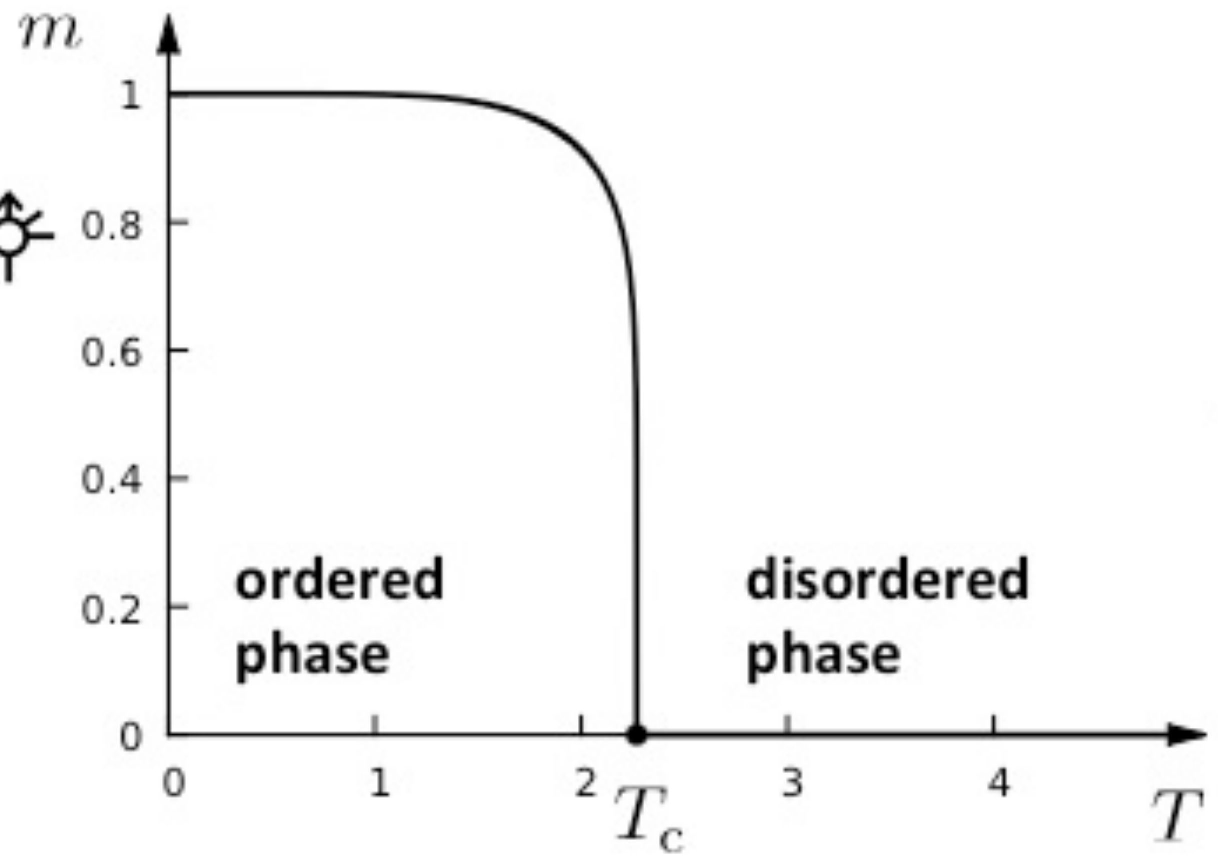


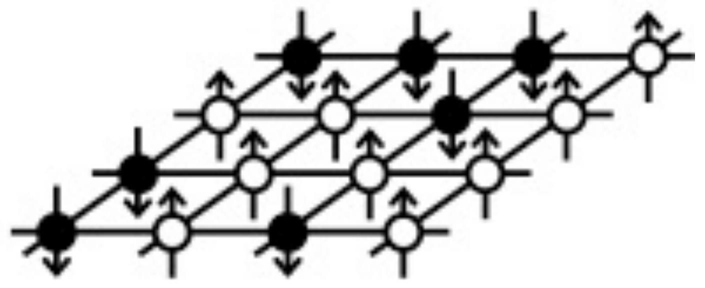


Magnetic phase transition



ordered phase

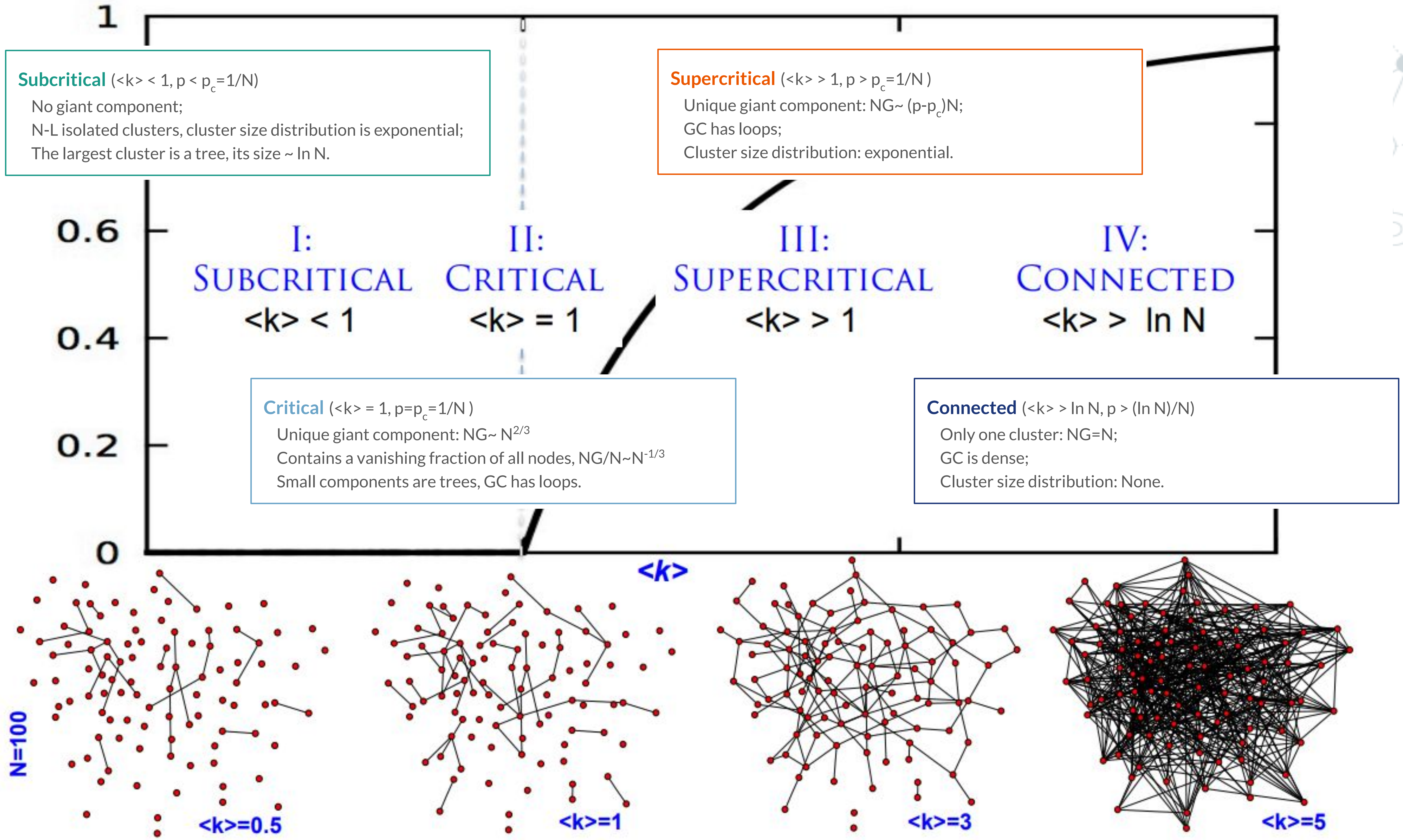




disordered phase

Many properties of a system at a phase transition are universal

Phase transition in connectedness



Subcritical ($\langle k \rangle < 1$, $p < p_c = 1/N$)
- No giant component;
- N-L isolated clusters, cluster size distribution is exponential;
- The largest cluster is a tree, its size ~ ln N.

Supercritical ($\langle k \rangle > 1$, $p > p_c = 1/N$)
- Unique giant component: $N_G \sim (p-p_c)N$;
- GC has loops;
- Cluster size distribution: exponential.

Critical ($\langle k \rangle = 1$, $p = p_c = 1/N$)
- Unique giant component: $N_G \sim N^{2/3}$
- Contains a vanishing fraction of all nodes, $N_G/N \sim N^{-1/3}$
- Small components are trees, GC has loops.

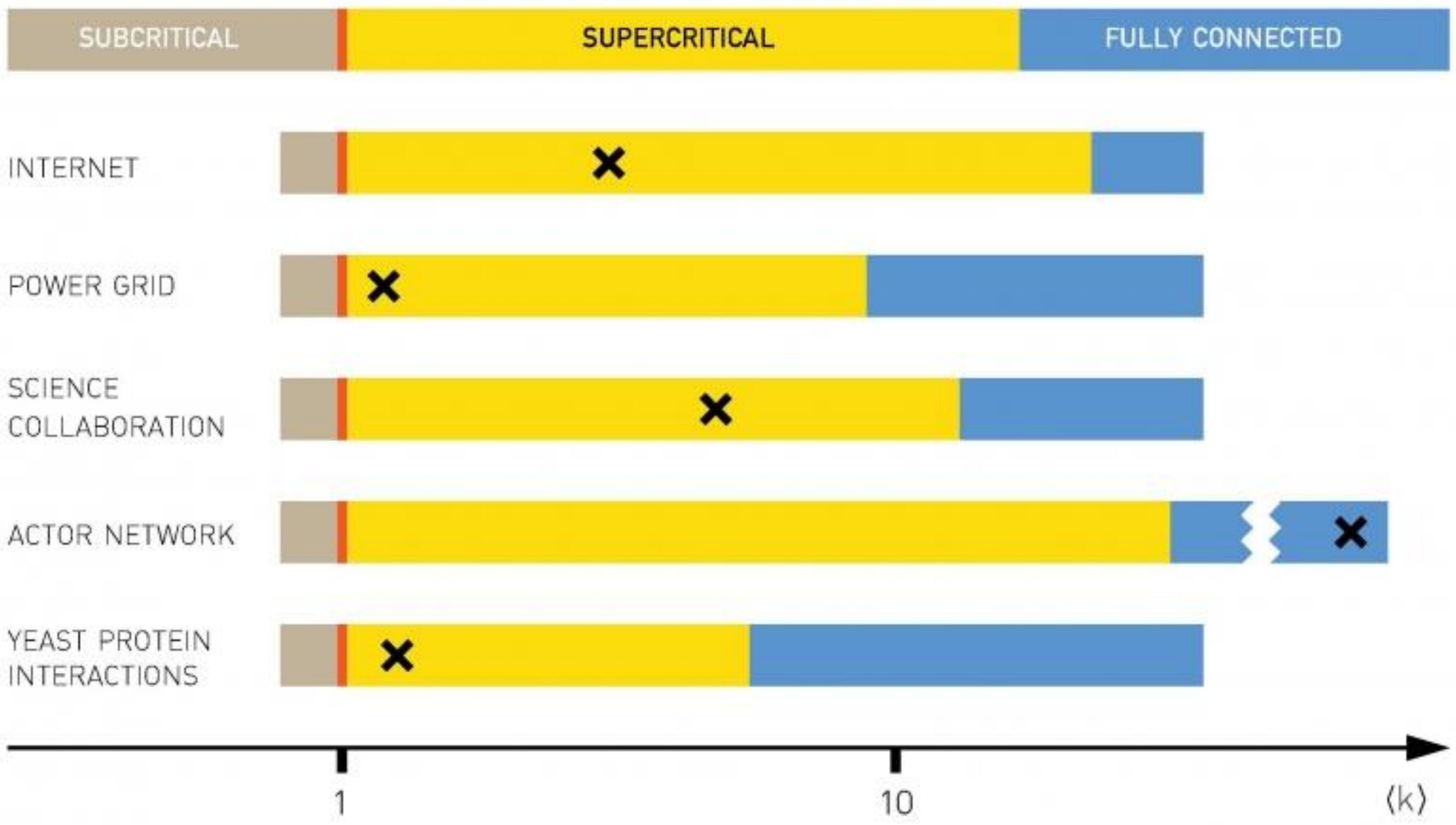
Connected ($\langle k \rangle > \ln N$, $p > (\ln N)/N$)
- Only one cluster: $N_G = N$;
- GC is dense;
- Cluster size distribution: None.

Cluster size distribution

$$p(s) = \frac{e^{-\langle k \rangle s}(\langle k \rangle s)^{s-1}}{s!}$$

$$p(s) \sim s^{-3/2} e^{-(\langle k \rangle - 1)s + (s-1)\ln \langle k \rangle}$$

Erdos-Renyi random network model



Most real networks are in the supercritical regime

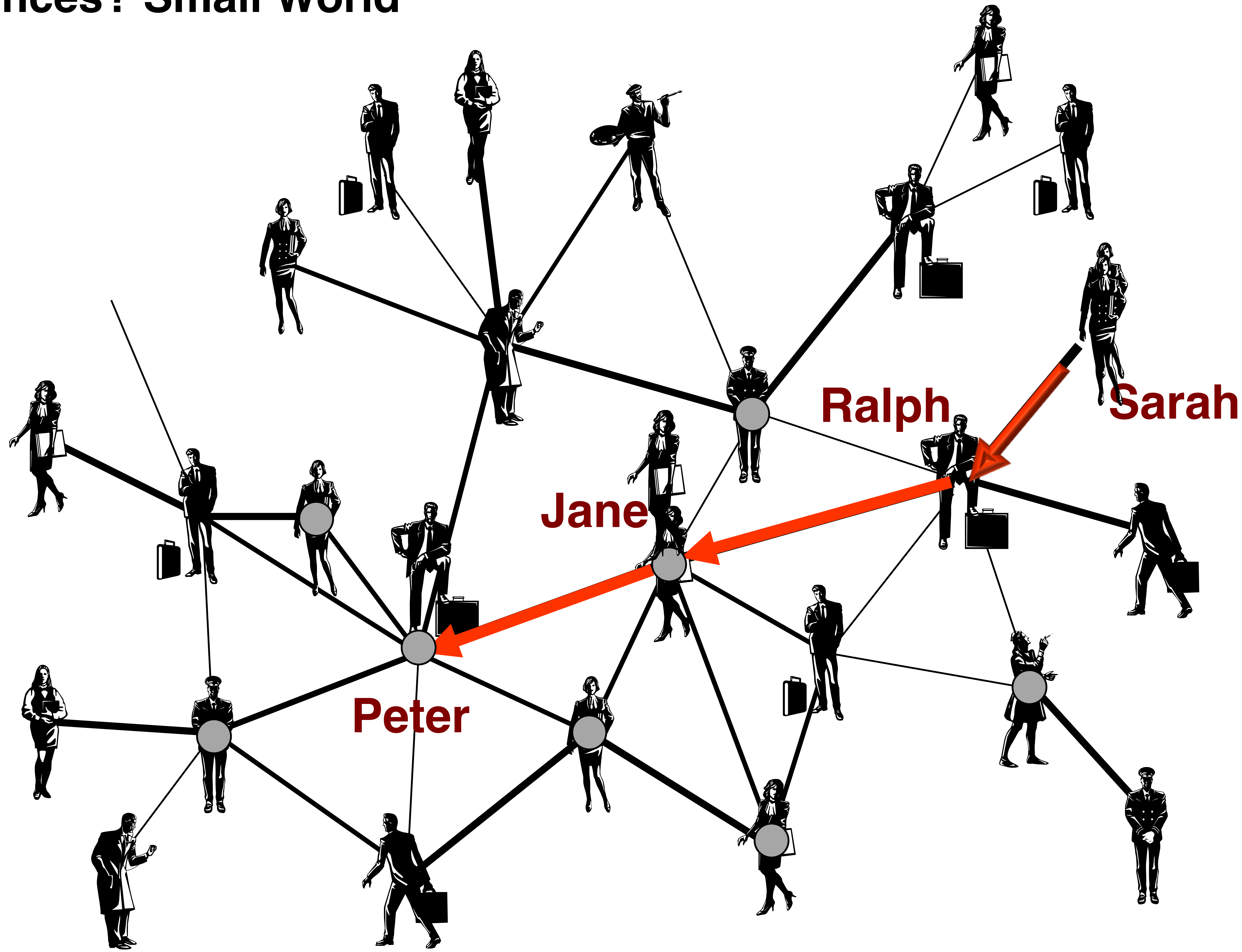
Random network theory then implies that they should have: Giant Component + many disconnected ones
-> but real networks are usually fully connected

List of results:

- We can reproduce sparseness using N and p
- Degree distribution is binomial/poisson NOT broad/powerlaw
- We can reproduce high clustering, but not low density (or viceversa)
- We can reproduce connectedness with p ~ 1/N

Erdos-Renyi random network model

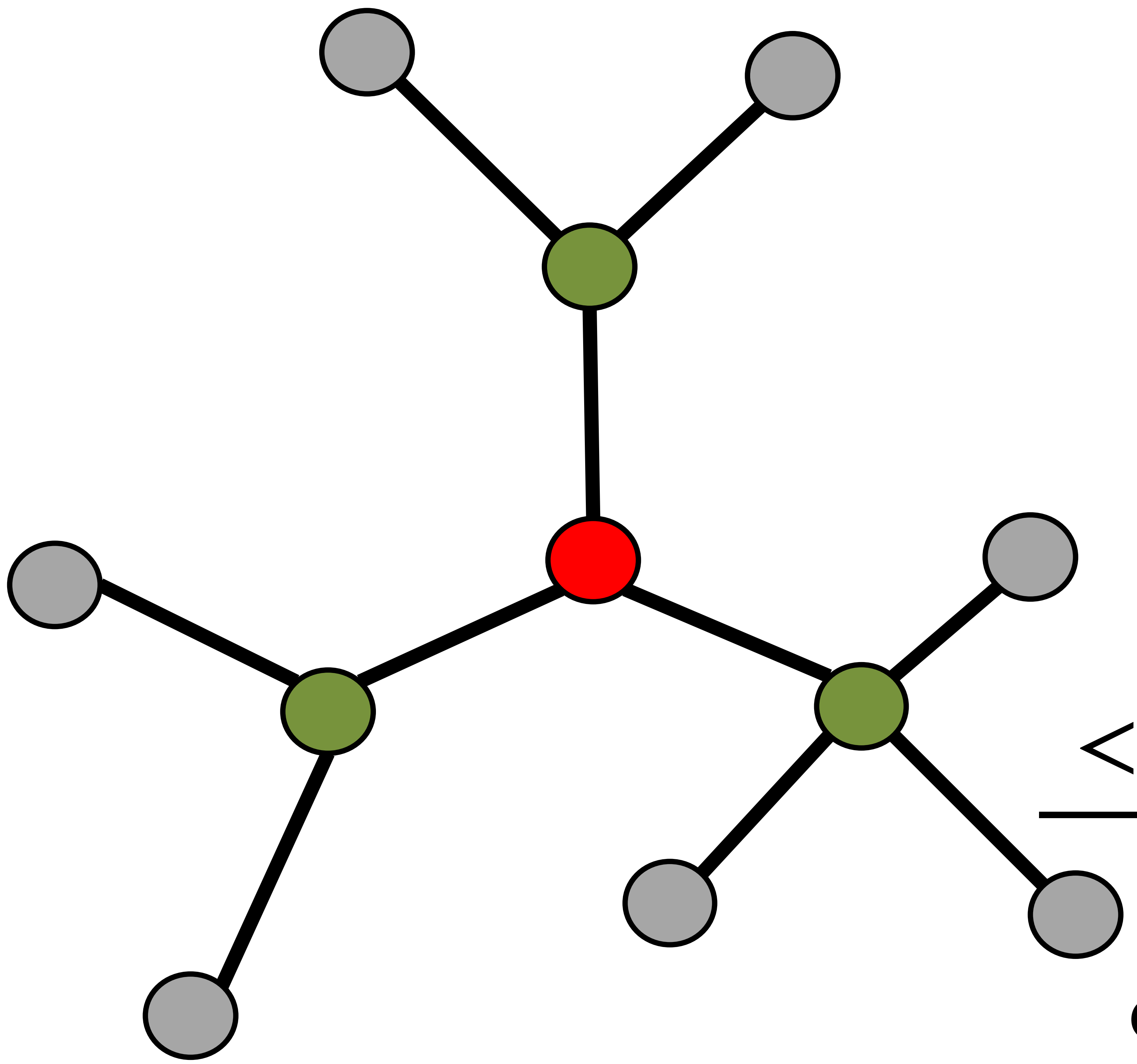
What about distances? Small World



Frigyes Karinthy, 1929
Stanley Milgram, 1967

Erdos-Renyi random network model

Let’s try an easy case



<k> nodes at distance d=1

<k>^2 nodes at distance d=2

<k>^3 nodes at distance d=3

.....

$$1 + <k> + <k>^2 + <k>^3 + \ldots . = N(d)$$

$$\frac{<k>^{d_{max}+1} - 1}{<k> - 1} = N(d_{max}) = N \quad \rightarrow \quad d_{max} \simeq \frac{\log N}{\log <k>}$$

Geometric series

Assume <k> >> 1

Wrong! This is actually closer to the average distance!

$$<d> \simeq \frac{\log N}{\log <k>}$$

This is small world: <d> << N for large N

<d>: avg shortest path

Small world property

Erdos-Renyi random network model

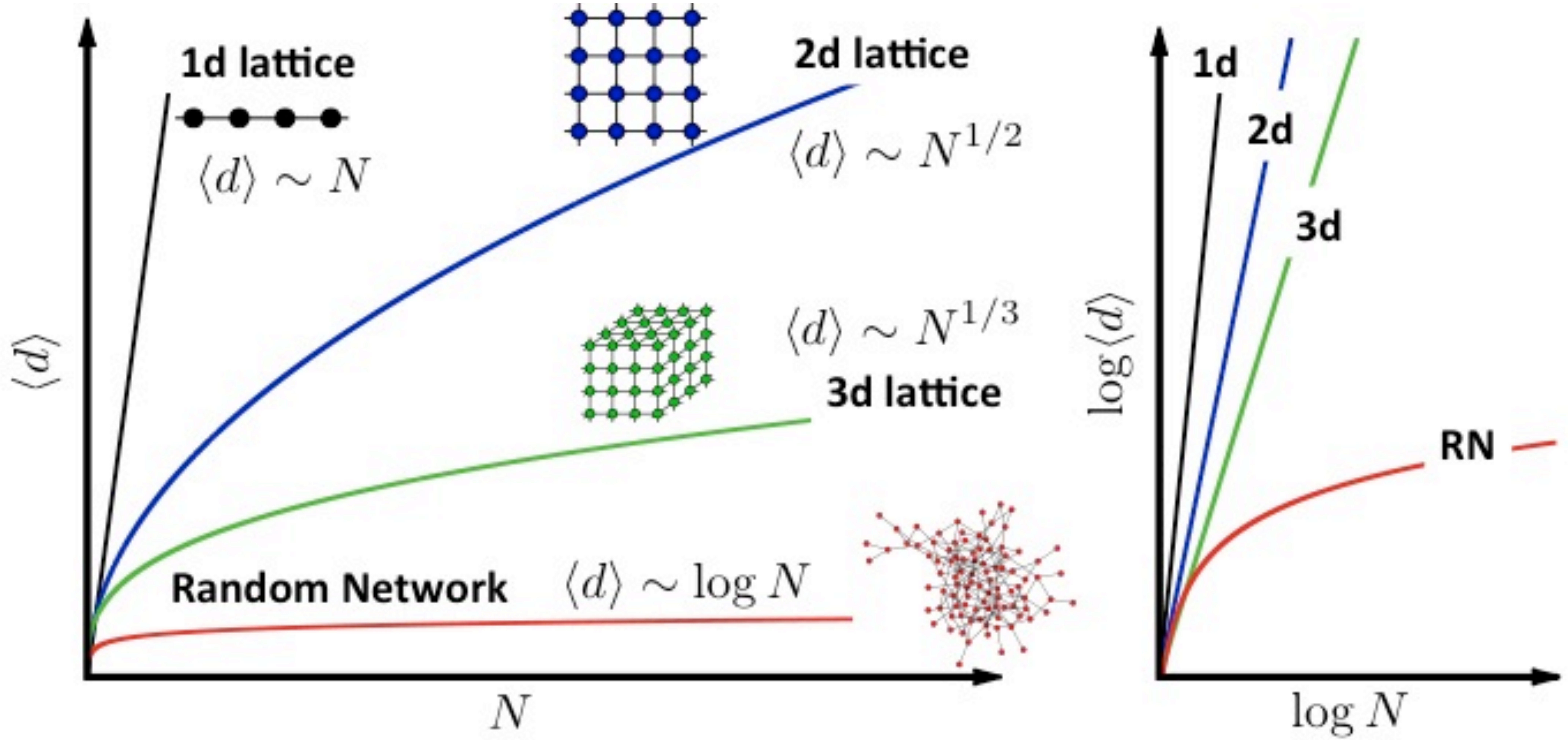
List of results:

- We can reproduce sparseness using N and p
- Degree distribution is binomial/poisson NOT broad/powerlaw
- We can reproduce high clustering, but not low density (or viceversa)
- We can reproduce connectedness with p ~ 1/N
- Small worldness

NETWORK	N	L	$\langle k \rangle$	$\langle d \rangle$	d_{max}	$\frac{\ln N}{\ln \langle k \rangle}$
Internet	192,244	609,066	6.34	6.98	26	6.58
WWW	325,729	1,497,134	4.60	11.27	93	8.31
Power Grid	4,941	6,594	2.67	18.99	46	8.66
Mobile Phone Calls	36,595	91,826	2.51	11.72	39	11.42
Email	57,194	103,731	1.81	5.88	18	18.4
Science Collaboration	23,133	93,439	8.08	5.35	15	4.81
Actor Network	702,388	29,397,908	83,71	3,91	14	3,04
Citation Network	449,673	4,707,958	10.43	11,21	42	5.55
E. Coli Metabolism	1,039	5,802	5.58	2.98	8	4.04
Protein Interactions	2,018	2,930	2.90	5.61	14	7.14

Is small-world surprising?

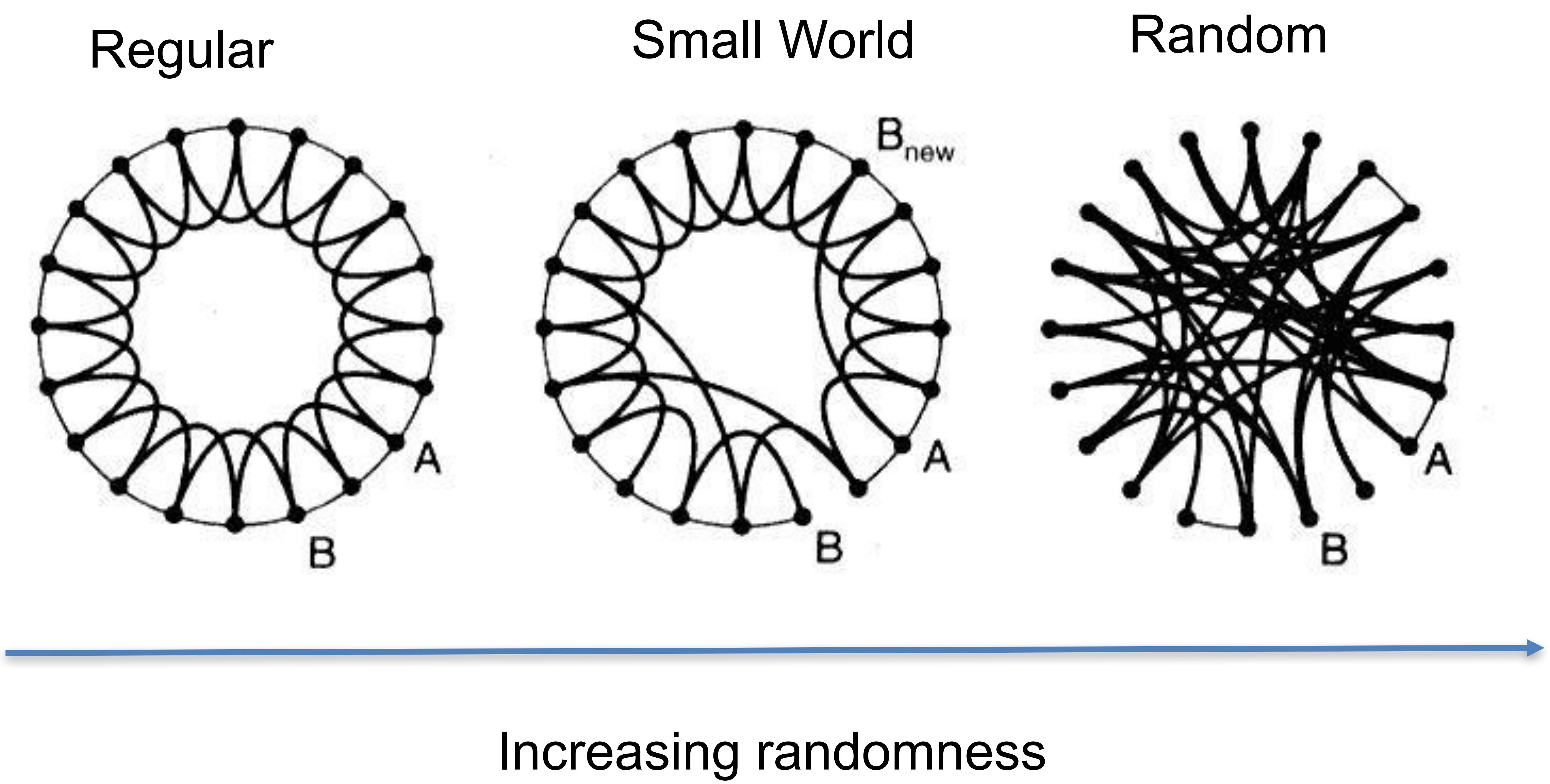
Compared to lattices (for which we have more intuition), yes

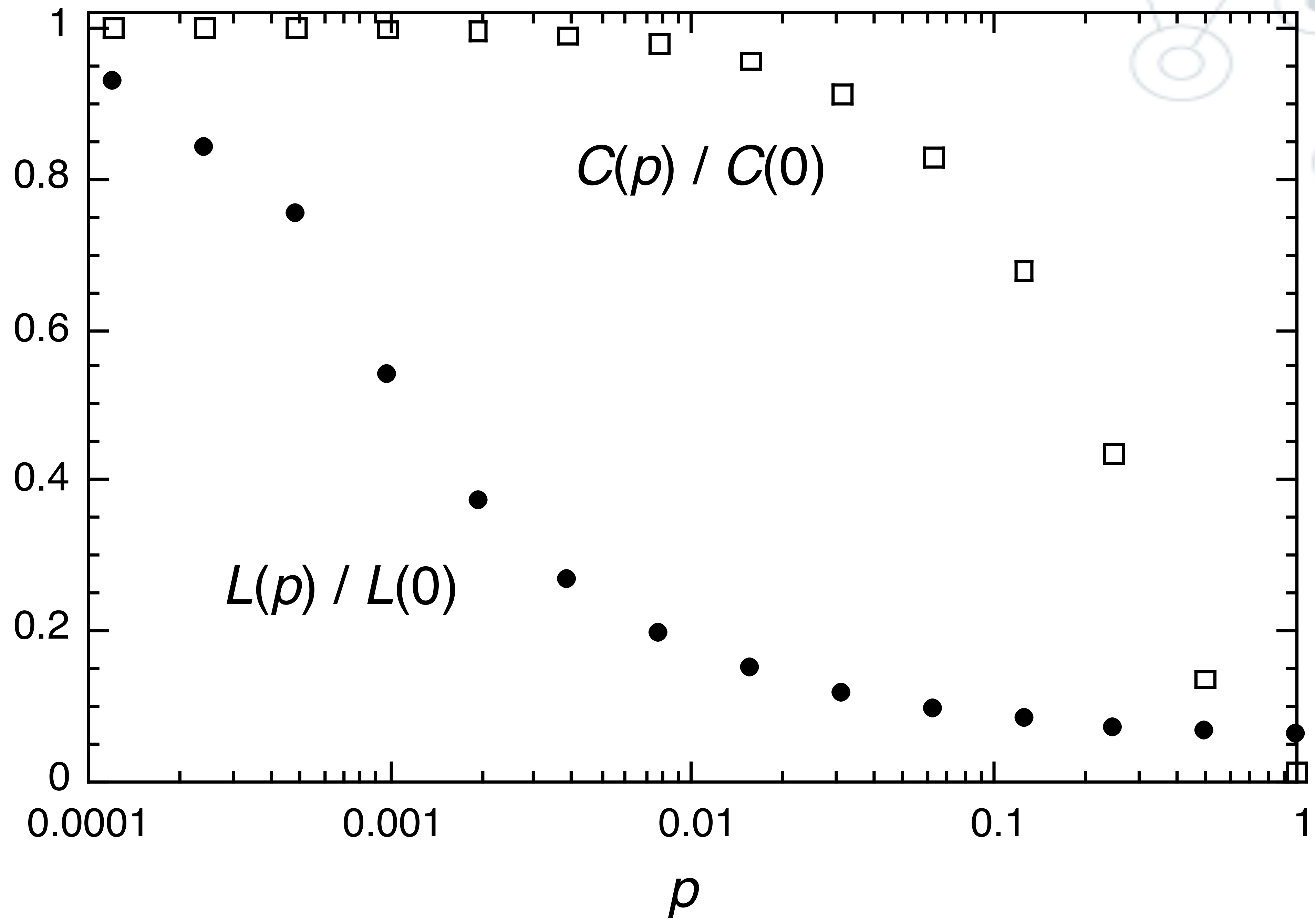


Can we reconcile SW and high C? Watts-Strogatz model

Regular lattices: not small world, and high clustering
Random network: small world, but low clustering

C(p): avg clustering coeff as a function of p
L(p): average shortest path length as a function of p





Watts, Duncan J., and Steven H. Strogatz. "Collective dynamics of ‘small-world’networks." *nature* 393.6684 (1998): 440-442.

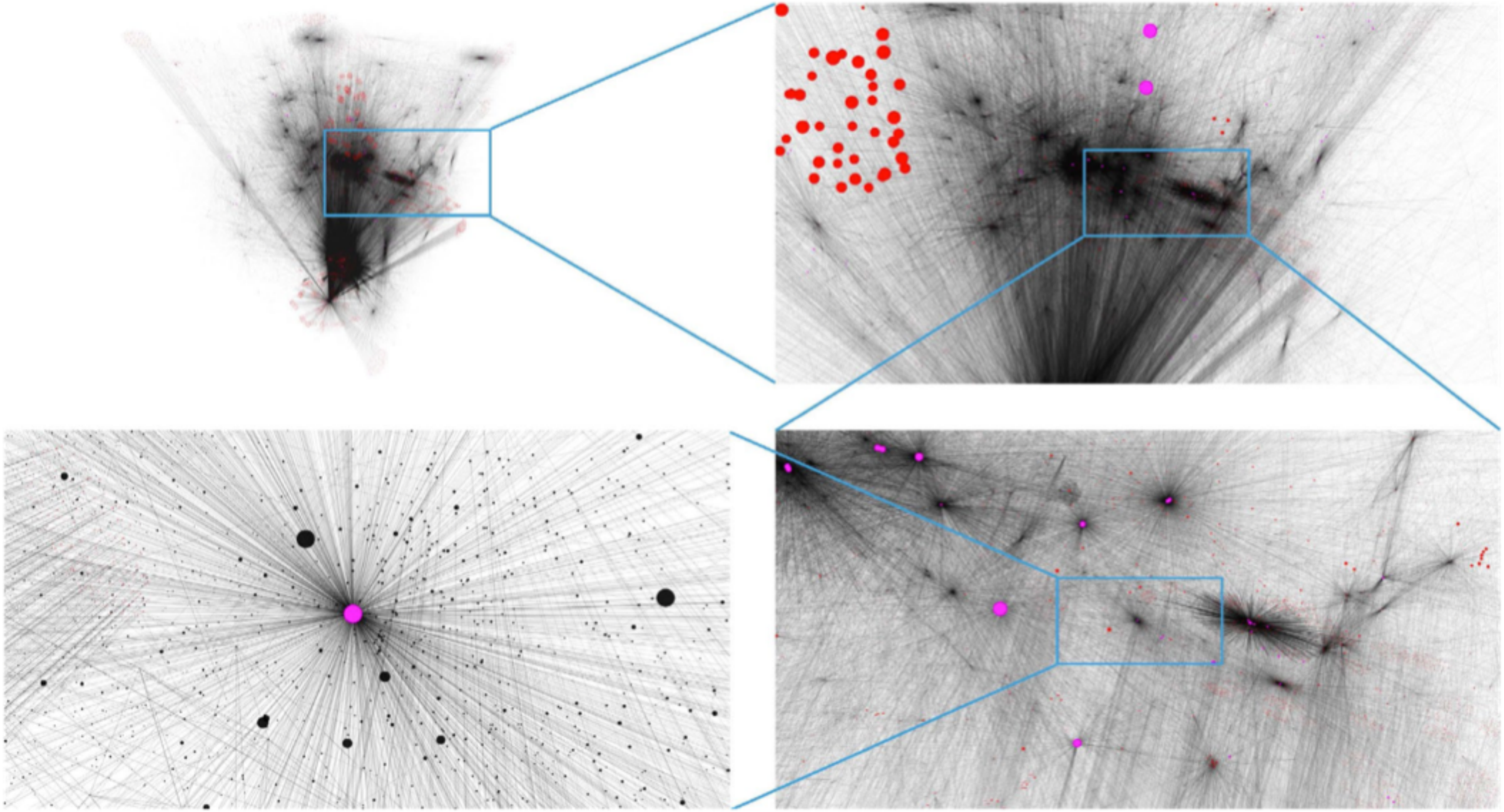
Erdos-Renyi random network model

List of results:

- We can reproduce sparseness using N and p
- Degree distribution is binomial/poisson NOT broad/powerlaw
- We can reproduce high clustering, but not low density (or viceversa)
- We can reproduce connectedness with p ~ 1/N
- Small worldness emerges naturally.

Network	Degree Distribution	Path Length	Clustering Coefficient
Real-world networks	Broad	Short	Large
ER graphs	Poissonian	Short	Small

What does scale-freeness mean? Hubs



What does scale-freeness mean?

A scale-free network is a network whose degree distribution follows a power law.

Discrete formalism

$$p_k = Ck^{-\gamma}$$

$$\sum_{k=1}^{\infty} p_k = 1$$

$$C \sum_{k=1}^{\infty} k^{-\gamma} = 1 \qquad C = \frac{1}{\sum_{k=1}^{\infty} k^{-\gamma}} = \frac{1}{\zeta(\gamma)}$$

$$p_k = \frac{k^{-\gamma}}{\zeta(\gamma)}$$

Continuous formalism

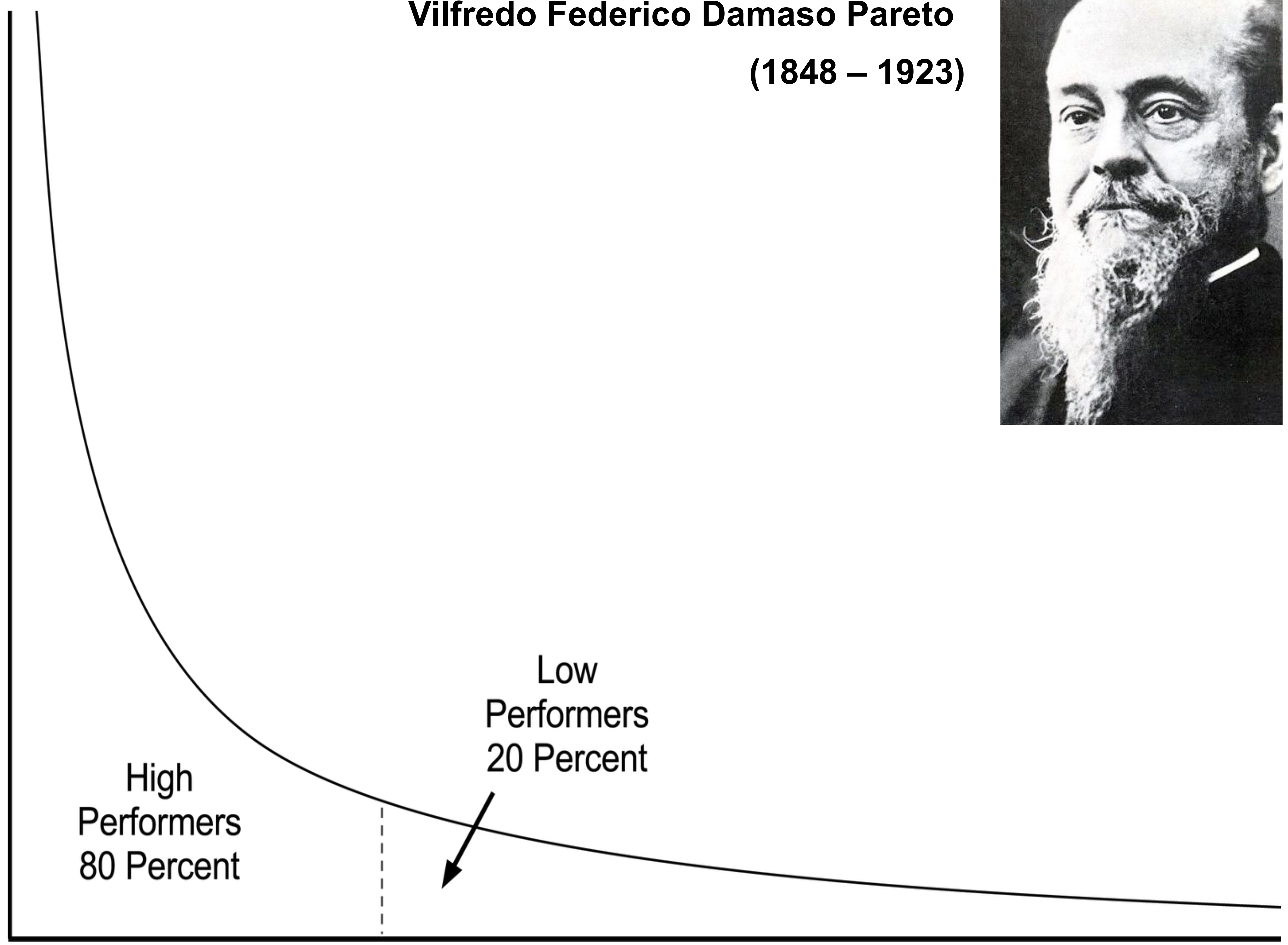
$$p(k) = Ck^{-\gamma}$$

$$\int_{k_{min}}^{\infty} p(k)dk = 1$$

$$C = \frac{1}{\int_{k_{min}}^{\infty} p(k)dk} = (\gamma - 1)k_{min}^{\gamma-1}$$

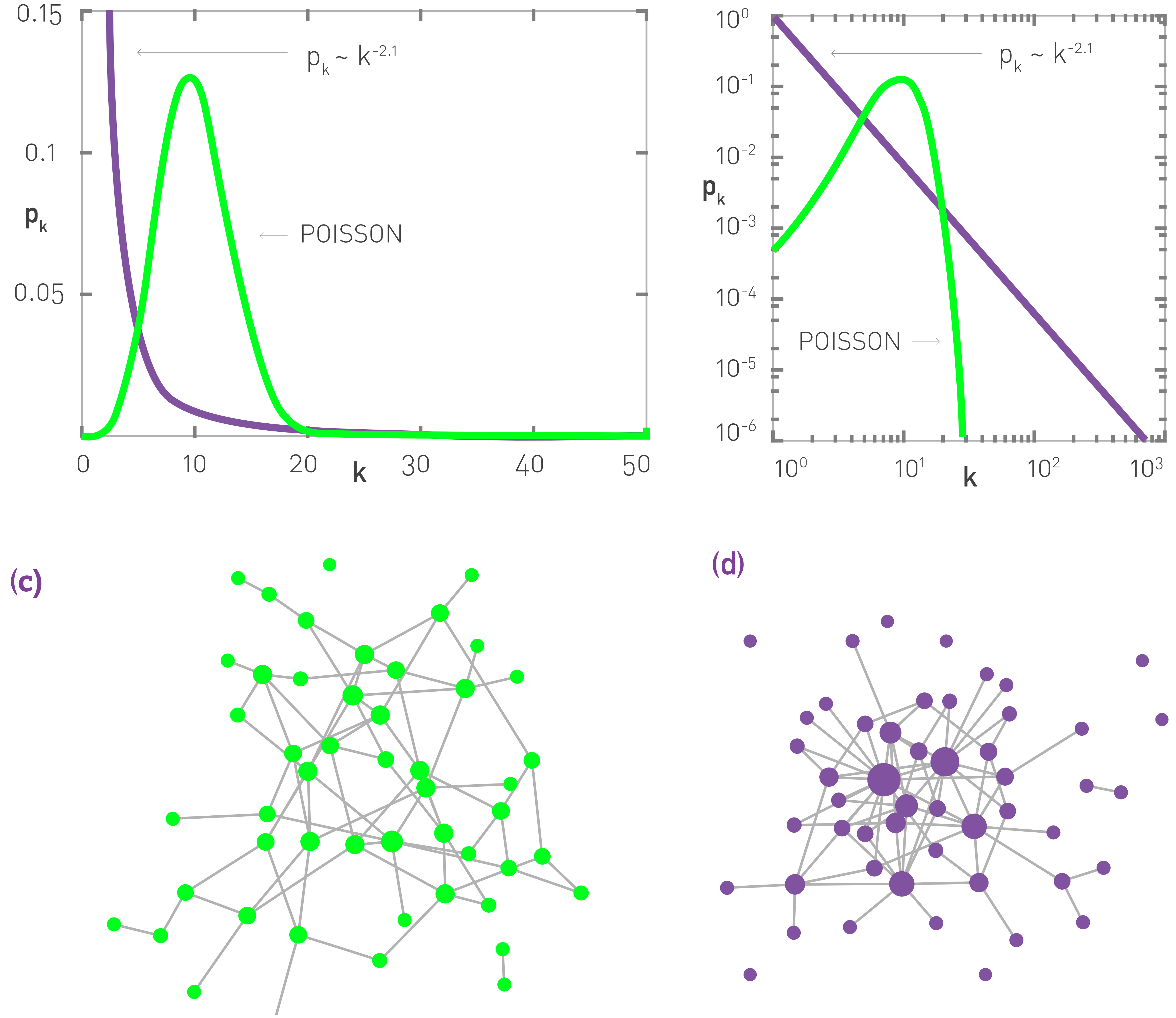
$$p(k) = (\gamma - 1)k_{min}^{\gamma-1}k^{-\gamma}$$

Why is scale-freeness important?



Vilfredo Federico Damaso Pareto

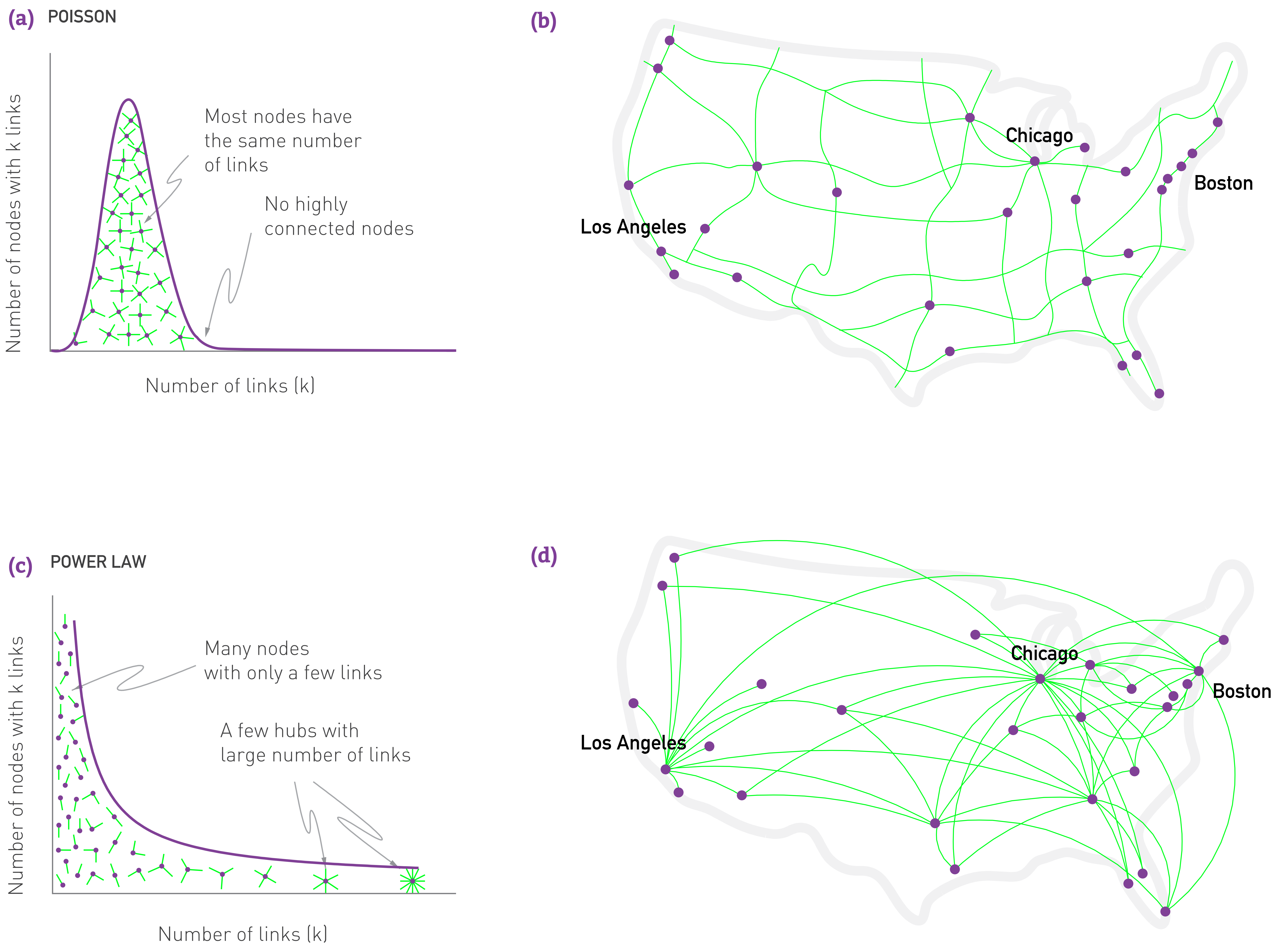
(1848 – 1923)



Hubs!

Why is scale-freeness important?

Implies heterogeneity
Changes network "topology"
Affects dynamical processes



Why is scale-freeness important?

One hub to rule them all. **How does the network size affect the size of the largest hub?**

Power laws "diverge" often, but networks are finite, hence max degree exists

Assume only one node

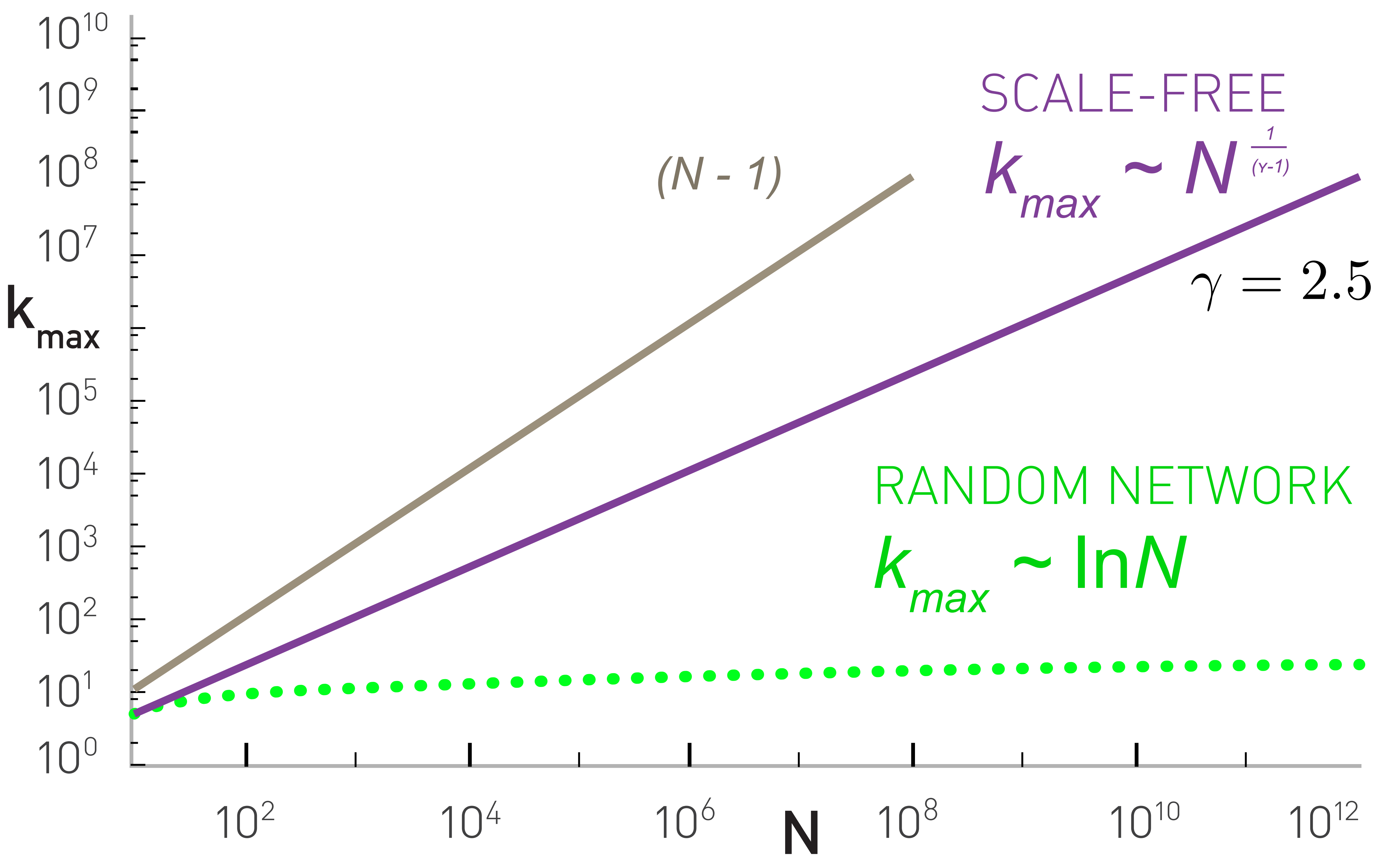
$$\int_{k_{max}}^{\infty} p(k)dk \simeq \frac{1}{N}$$

By def, for scale free

$$\int_{k_{max}}^{\infty} p(k)dk = (\gamma-1)k_{min}^{\gamma-1}\int_{k_{max}}^{\infty} k^{-\gamma}dk = \frac{\gamma-1}{-\gamma+1}k_{min}^{\gamma-1}[k^{-\gamma+1}]_{k_{max}}^{\infty} = \frac{k_{min}{}^{\gamma-1}}{k_{max}{}^{\gamma-1}} \simeq \frac{1}{N}$$

$$k_{max} = k_{min}N^{\frac{1}{\gamma-1}}$$

- k_{max}, increases with the size of the network ==> bigger system, bigger hub
- For γ>2, k_{max} increases slower than N ==> decreasing fraction of links as N increases.
- For γ=2 k_{max}~N ==> The size of the biggest hub is O(N)
- For γ<2 k_{max} increases faster than N: condensation phenomena ==> the largest hub will grab an increasing fraction of links. Anomaly!



Why is scale-freeness important?

More divergences!

$$< k^m >= \int_{k_{min}}^{\infty} k^m p(k) dk \qquad p(k) = (\gamma - 1) k_{min}^{\gamma-1} k^{-\gamma}$$

$$< k^m >= (\gamma - 1) k_{min}^{\gamma-1} \int_{k_{min}}^{\infty} k^{m-\gamma} dk = \frac{\gamma - 1}{m - \gamma + 1} k_{min}^{\gamma-1} \left[k^{m-\gamma+1} \right]_{k_{min}}^{\infty}$$

if $m - \gamma + 1 < 0$: $\qquad < k^m >= \frac{\gamma - 1}{m - \gamma + 1} k_{min}^{m}$

if $m - \gamma + 1 > 0$: $\qquad < k^m > \rightarrow \infty$

This implies:

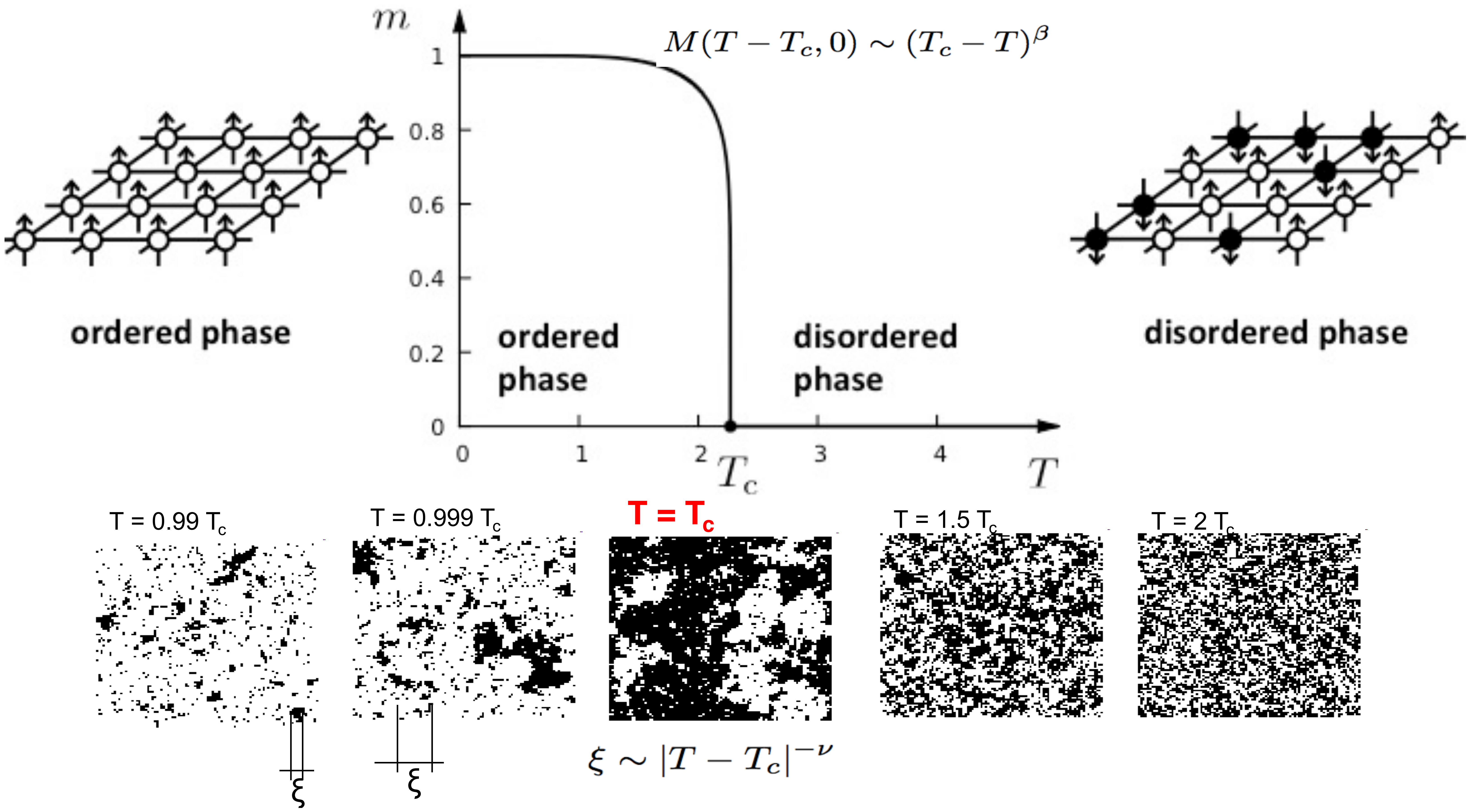
$$\text{For } \gamma < 3, \quad < k^2 > \rightarrow \infty$$

As N goes to infinity: this means there is no single scale

Network	Size	$\langle k \rangle$	κ	γ_{out}	γ_{in}
WWW	325 729	4.51	900	2.45	2.1
WWW	4×10^7	7		2.38	2.1
WWW	2×10^8	7.5	4000	2.72	2.1
WWW, site	260 000				1.94
Internet, domain*	3015–4389	3.42–3.76	30–40	2.1–2.2	2.1–2.2
Internet, router*	3888	2.57	30	2.48	2.48
Internet, router*	150 000	2.66	60	2.4	2.4
Movie actors*	212 250	28.78	900	2.3	2.3
Co-authors, SPIRES*	56 627	173	1100	1.2	1.2
Co-authors, neuro.*	209 293	11.54	400	2.1	2.1
Co-authors, math.*	70 975	3.9	120	2.5	2.5
Sexual contacts*	2810			3.4	3.4
Metabolic, *E. coli*	778	7.4	110	2.2	2.2
Protein, *S. cerev.**	1870	2.39		2.4	2.4
Ythan estuary*	134	8.7	35	1.05	1.05
Silwood Park*	154	4.75	27	1.13	1.13
Citation	783 339	8.57			3
Phone call	53×10^6	3.16		2.1	2.1
Words, co-occurrence*	460 902	70.13		2.7	2.7
Words, synonyms*	22 311	13.48		2.8	2.8

Why is scale-freeness important?

Origin of the name



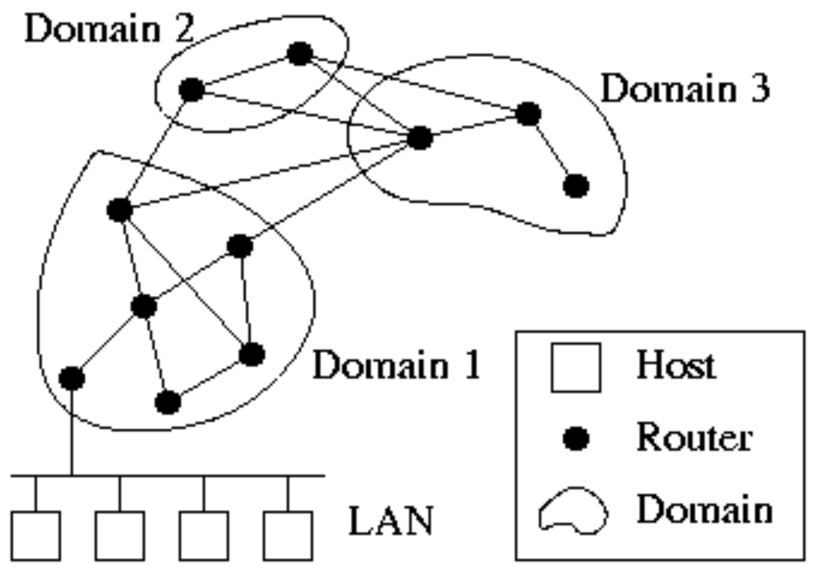
Correlation length diverges at the critical point: the whole system is correlated!

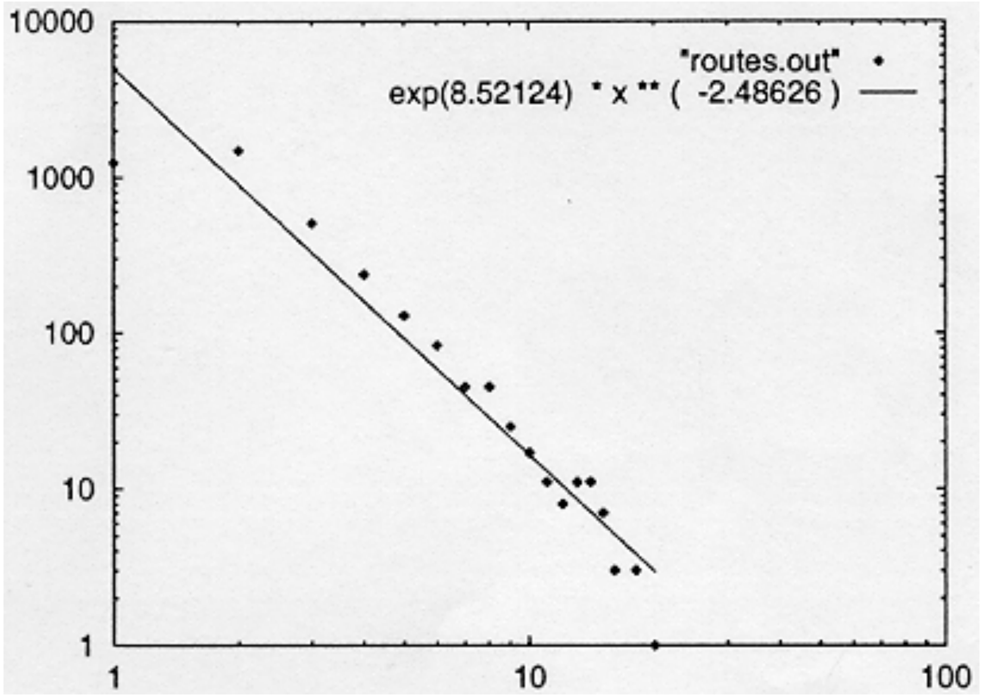
Scale invariance: there is no characteristic scale for the fluctuation **(scale-free behavior).**

Universality: exponents are independent of the system's details.

Why is scale-freeness important?

Universality?



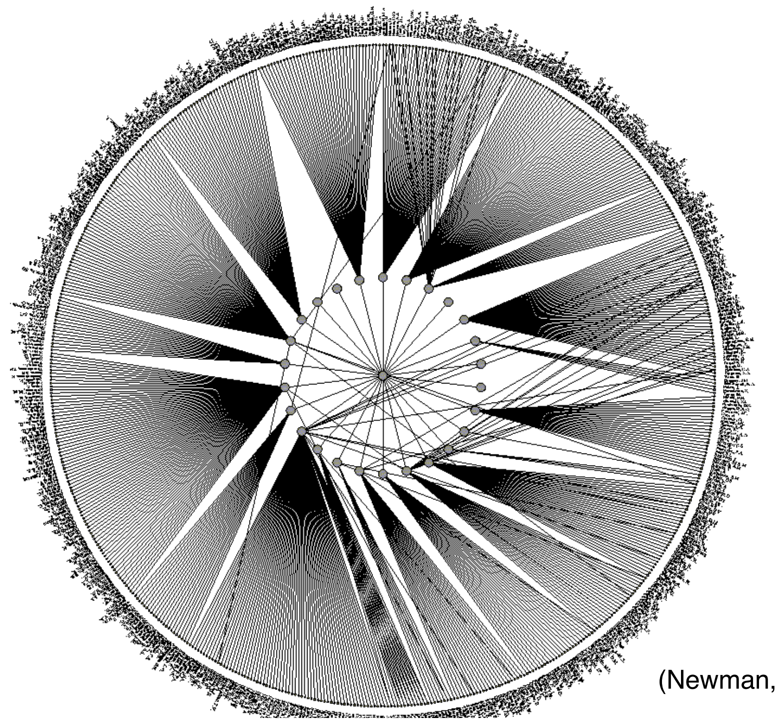


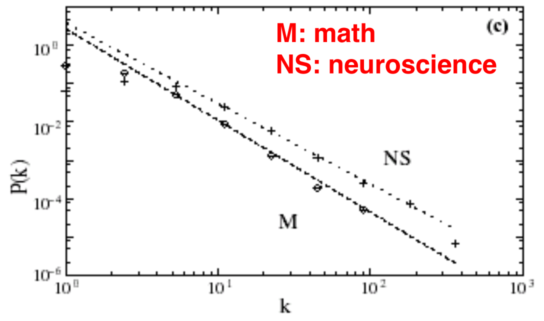
(Faloutsos, Faloutsos and Faloutsos, 1999)

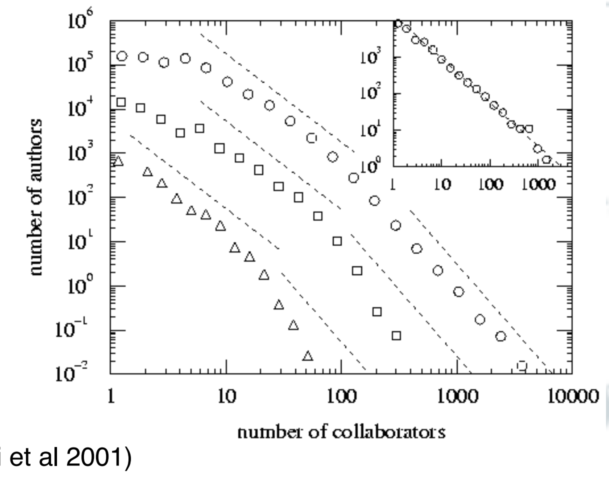
Scale free networks are rare:
https://www.nature.com/articles/s41467-019-08746-5

Scalefree networks well done:
https://arxiv.org/abs/1811.02071

Nodes: scientist (authors)
Links: joint publication

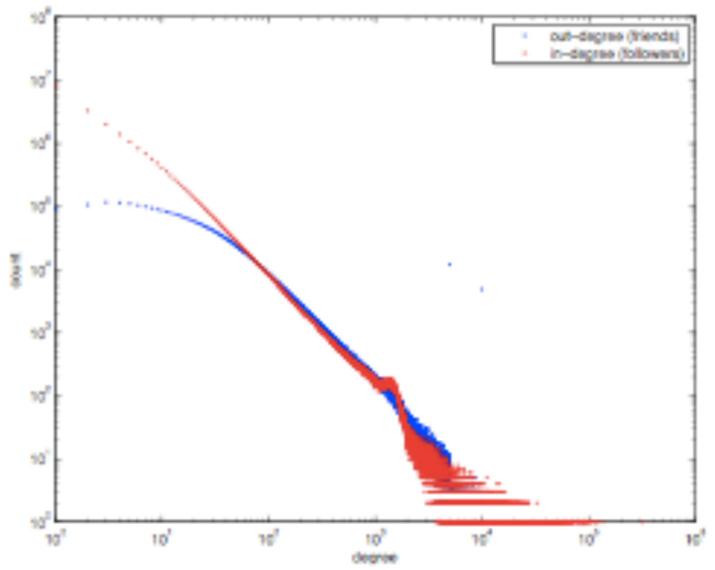




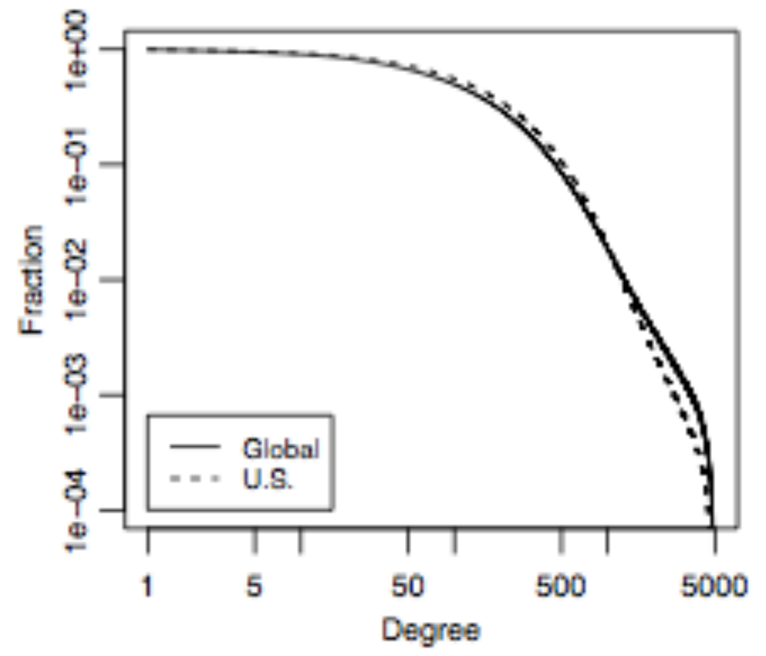


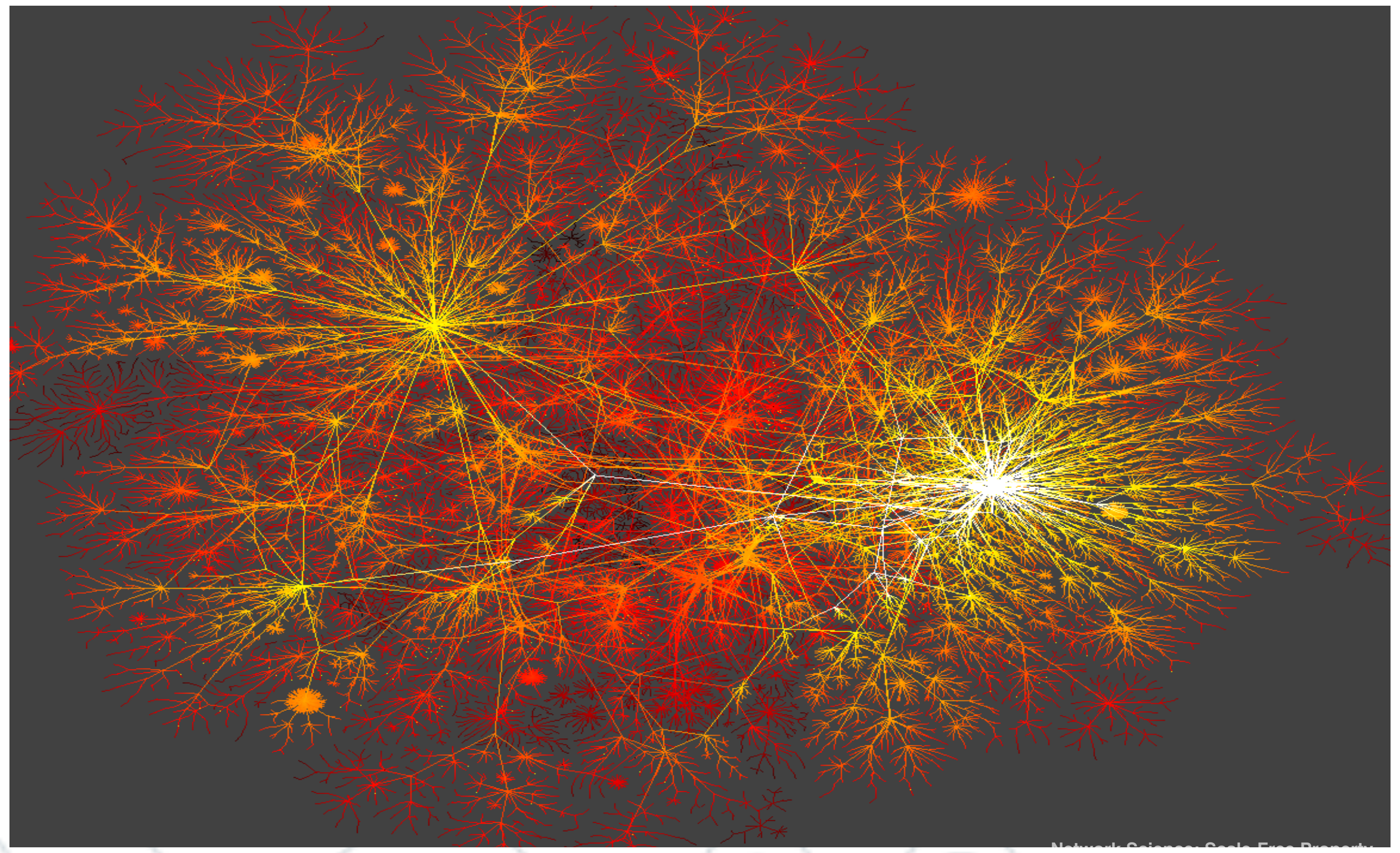
(Newman, 2000, Barabasi et al 2001)

Twitter:

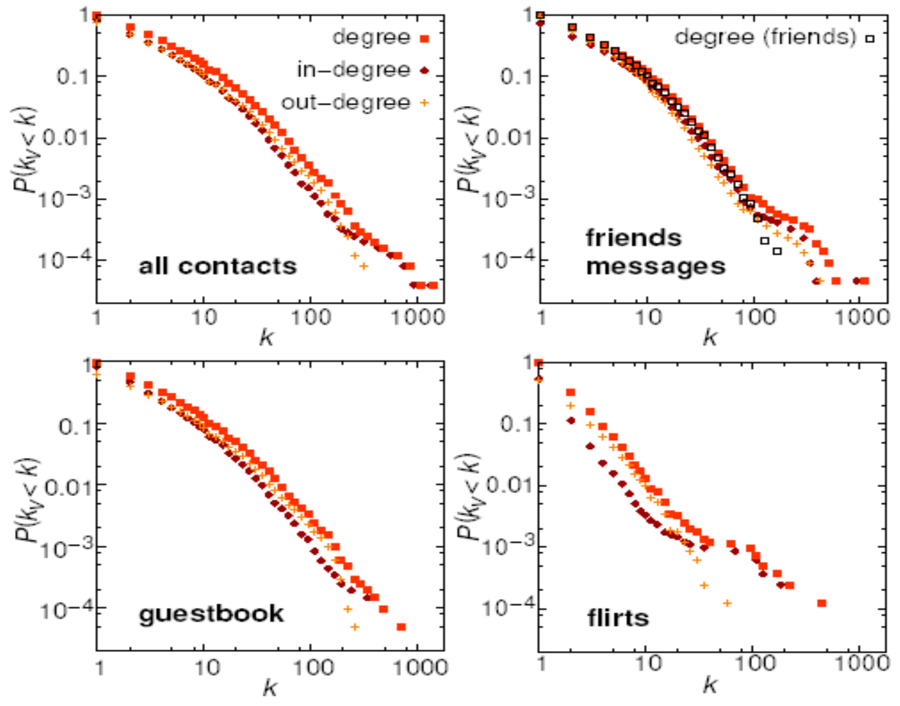


Facebook



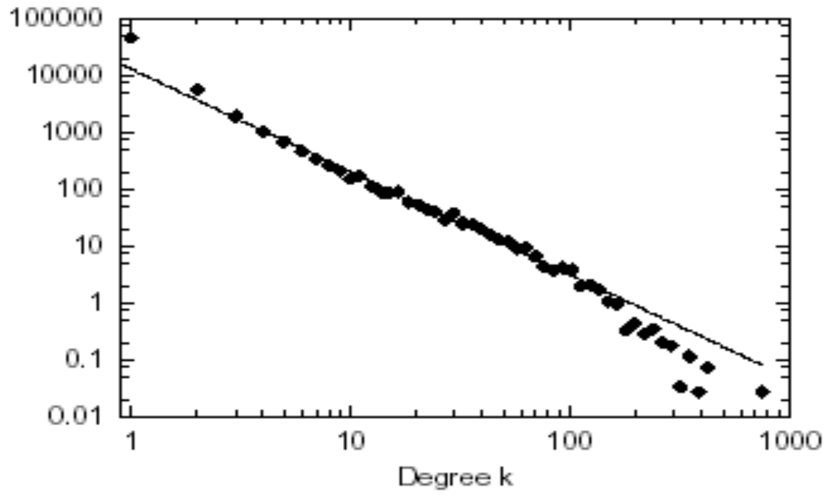


Pussokram.com online community;
512 days, 25,000 users.



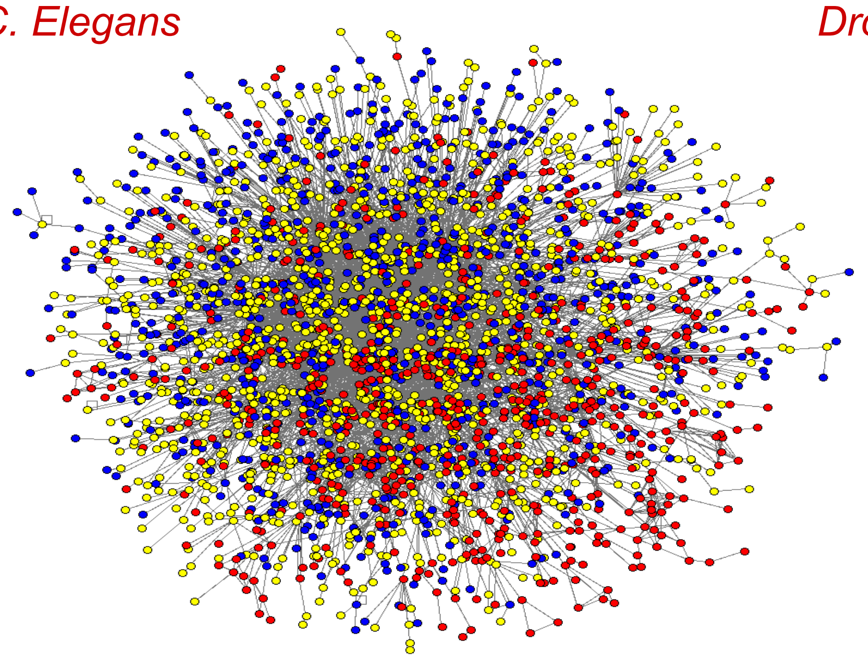
Holme, Edling, Liljeros, 2002.

Kiel University log files
112 days, N=59,912 nodes

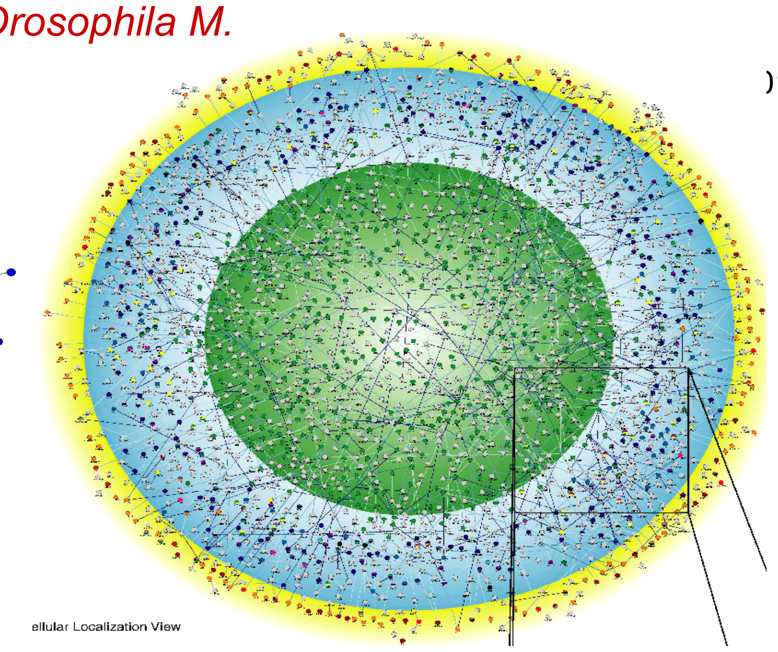


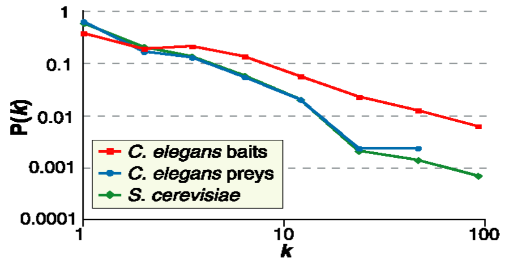
Ebel, Mielsch, Bornholdtz, PRE 2002.

C. Elegans

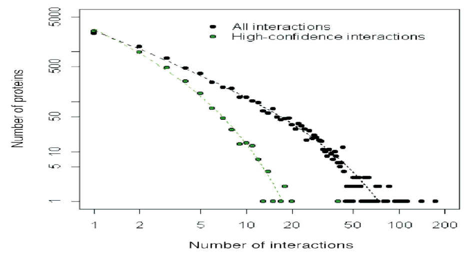


Drosophila M.





Li *et al.* Science 2004



Giot *et al.* Science 2003

Why is scale-freeness important?

Effects on the distances (smaller than in random)

$$
\langle l \rangle \sim \begin{cases} const. & \gamma = 2 \\ \dfrac{\ln \ln N}{\ln(\gamma - 1)} & 2 < \gamma < 3 \\ \dfrac{\ln N}{\ln \ln N} & \gamma = 3 \\ \ln N & \gamma > 3 \end{cases}
$$

Ultra Small World

Small World

Size of the biggest hub is of order O(N). Most nodes can be connected within two layers of it, thus the average path length will be independent of the system size.

The average path length increases slower than logarithmically. In a random network all nodes have comparable degree, thus most paths will have comparable length. In a scale-free network the vast majority of the path go through the few high degree hubs, reducing the distances between nodes.

Some key models produce γ=3, so the result is of particular importance for them. This was first derived by Bollobas and collaborators for the network diameter in the context of a dynamical model, but it holds for the average path length as well.

The second moment of the distribution is finite, thus in many ways the network behaves as a random network. Hence the average path length follows the result that we derived for the random network model earlier.

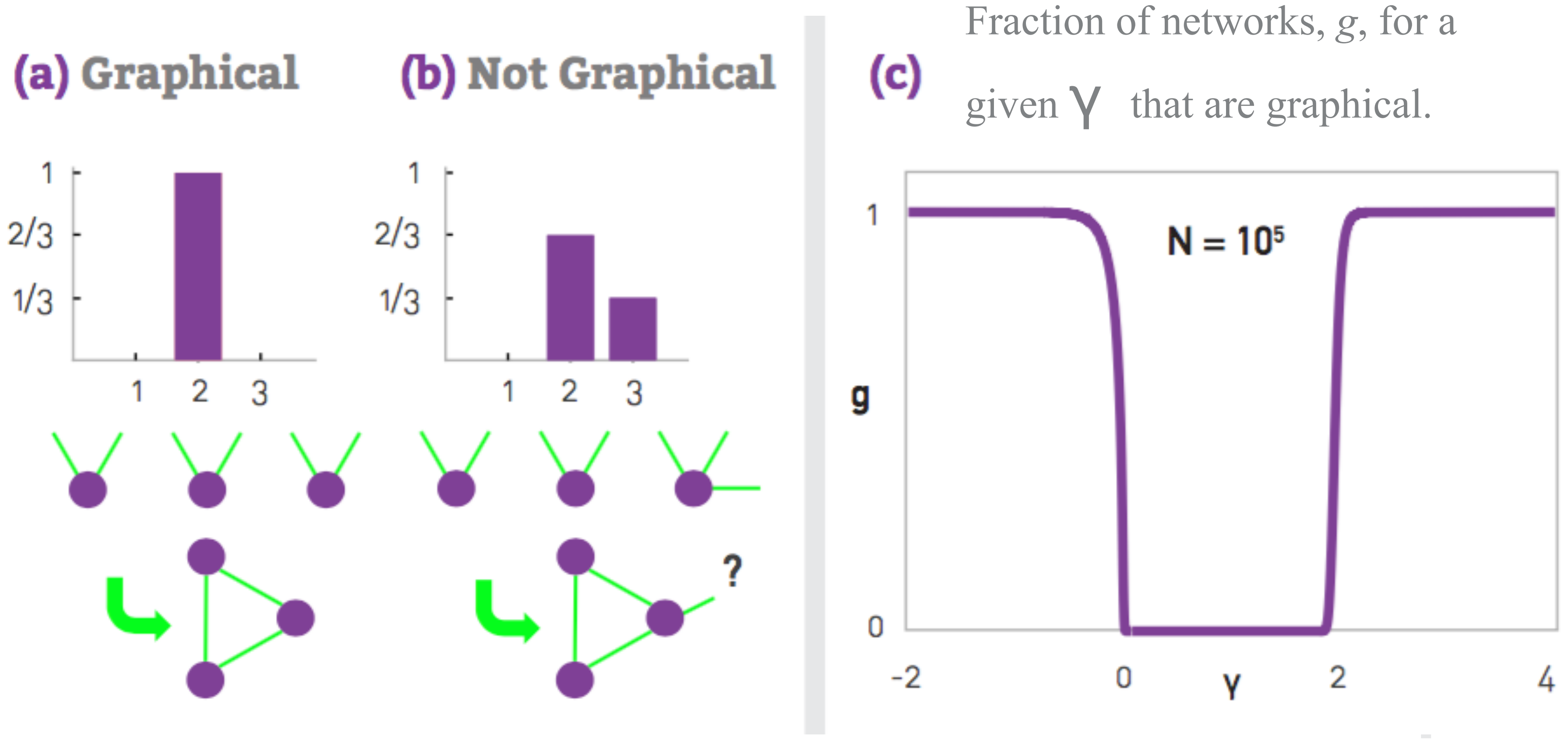
Cohen, Havlin Phys. Rev. Lett. 90, 58701(2003); Cohen, Havlin and ben-Avraham, in Handbook of Graphs and Networks, Eds. Bornholdt and Shuster (Willy-VCH, NY, 2002) Chap. 4; Confirmed also by: Dorogovtsev et al (2002), Chung and Lu (2002); (Bollobas, Riordan, 2002; Bollobas, 1985; Newman, 2001

Why is scale-freeness important?

Recap

Graphical: a degree sequence that can be turned into a graph

Small gamma? Graphicality



P. Erdős and T. Gallai. Graphs with given degrees of vertices. Matematikai Lapok, 11:264-274, 1960.

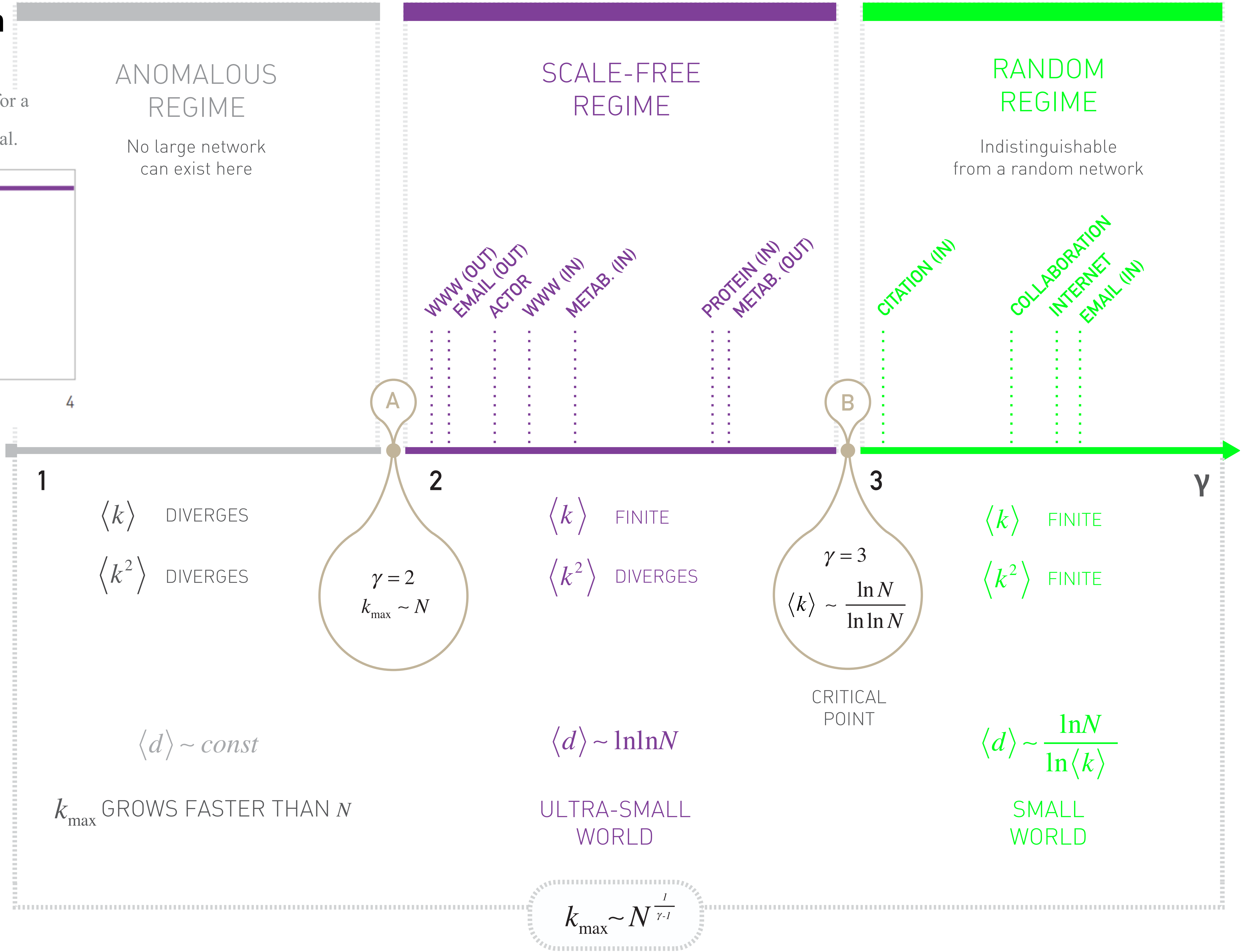
C.I. Del Genio, H. Kim, Z. Toroczkai, and K.E. Bassler. Efficient and exact sampling of simple graphs with given arbitrary degree sequence. PLoS ONE, 5: e10012, 04 2010.

V. Havel. A remark on the existence of finite graphs. Casopis Pest. Mat., 80:477-480, 1955.

Large gamma?

$$k_{max} = 10^3 \rightarrow k_{max} = k_{min} N^{\frac{1}{\gamma-1}}$$

$$N = \left(\frac{k_{max}}{k_{min}}\right)^{\gamma-1} \simeq 10^8$$

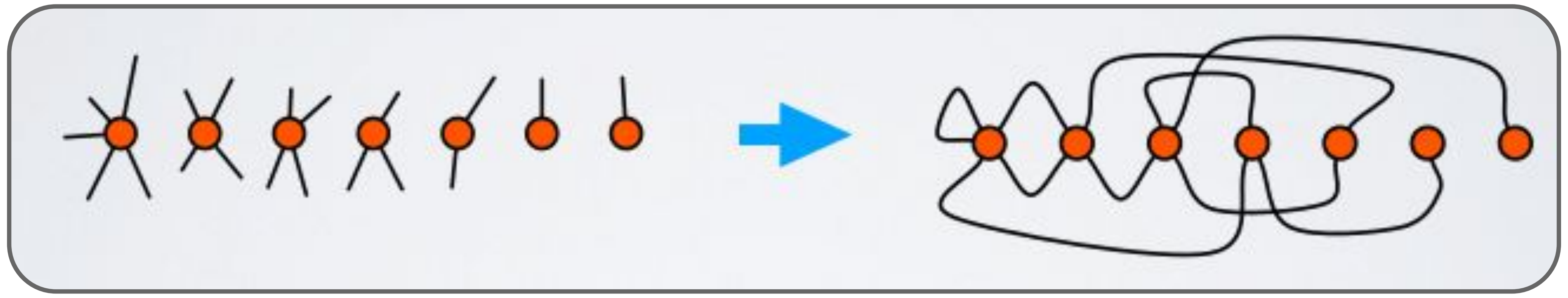


Can we constrain random models to be scale-free?

(Growing models next time) Now configuration model: fix the degree sequence, shuffle rest

Original idea:

1. Given a degree sequence $\vec{k} = \{k_1, k_2, \ldots, k_n\}$
2. Assign to each node $i \in V$ k_i bs
3. Select random pairs of unmatched stubs and connect them
4. Repeat 3 while there are unmatched stubs



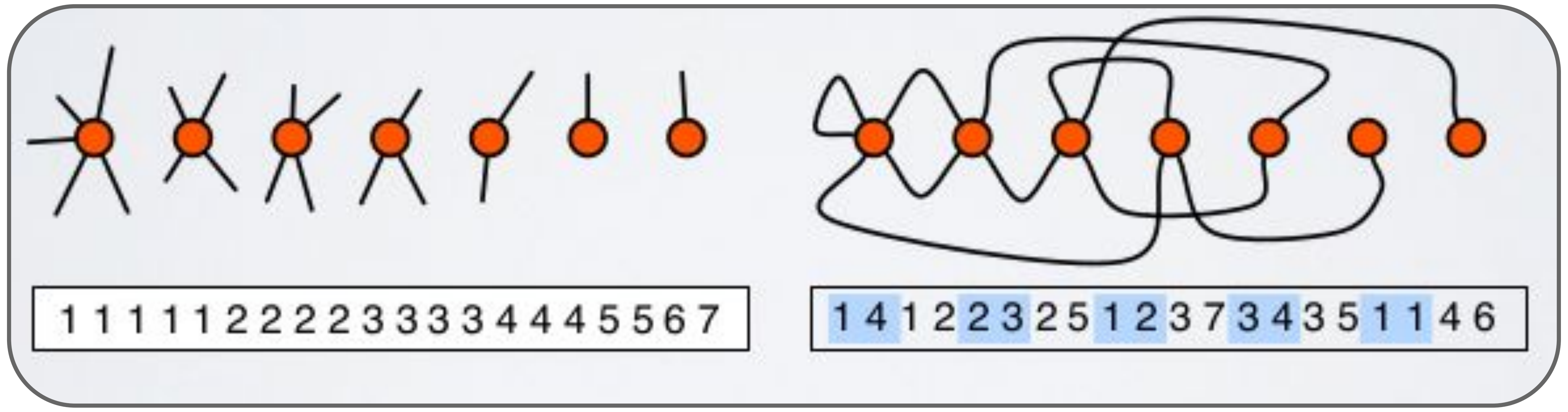
Such process produces a configuration model that preserves the input degree sequence, allowing:

- multi-links,
- self-links

expected number of self-loops and multi-links goes to zero in the $N \to \infty$ limit.

An effective algorithm

1. Take an array $\vec{v}$ ith length 2m and fill it with ki indices of each node $i \in V$
2. Make a random permutation of the array $\vec{v}$
3. Read the content of the array as ordered pairs
4. Each pair of consecutive node indices create a links in the configuration network

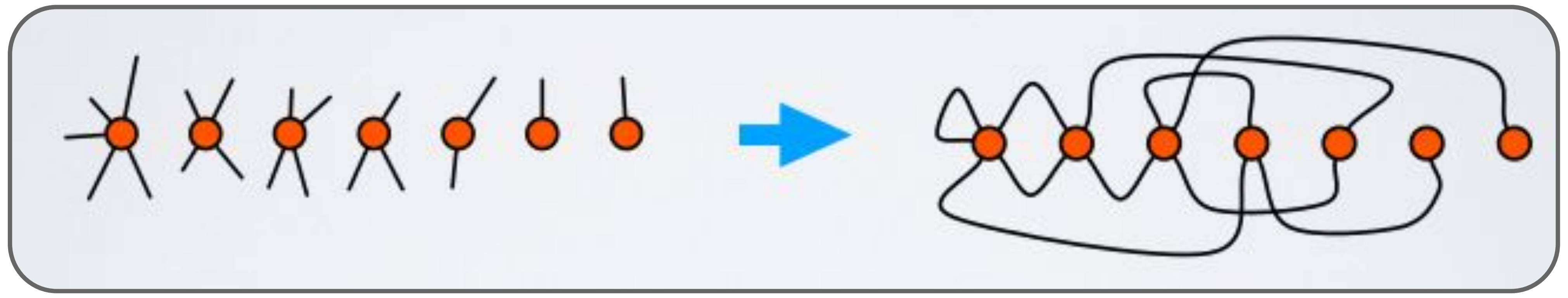


Can we constrain random models to be scale-free?

Clustering

$$C_g = \sum_{k_i, k_j = 1}^{\infty} q_{k_i} q_{k_j} \frac{(k_i - 1)(k_j - 1)}{2m} = \frac{1}{2m}\left[\sum_{k=0}^{\infty}(k-1)q_k\right]^2$$

where q_k denotes the probability that a random edge reaches a degree-k vertex



Excess degree

$$q_k = \frac{kp(k)}{<k>} \qquad 2m = N<k>$$

probability that a random edge reaches a degree- vertex

Probability of edge (i, j)

$$p(k_i, k_j) = \frac{k_i k_j}{2m}$$

$$C_g = \frac{1}{N<k>^3}\left[\sum_{k=1}^{\infty}(k-1)kp(k)\right]^2 = \frac{(<k^2> - <k>)^2}{N<k>^3} \sim \frac{const}{N}$$

Average degree of neighbours

$$<k_{nn}> = \sum_k k q_k = \frac{<k^2>}{<k>}$$

Micro-canonical model!
Canonical version: **Chung-Lu model**
(https://arxiv.org/pdf/1910.11341.pdf)

Molloy-Reed criterion (**homework!**)

Network	Degree Distribution	Path Length	Clustering Coefficient
Real-world networks	Broad	Short	Large
ER graphs	Poissonian	Short	Small
Configuration model	Custom, can be broad	Short	Small

Can we constrain random models to be scale-free?

Elements of Molloy-Reed criterion

Definition 4. *A node is in the giant component of the network if, at least one of its links reach a node that is also in the giant component of the network.*
A node reached by following a link of a network is in the giant component if at least one of the nodes reached by following one of the other links of the node is also in the giant component.

Probability following link to node in GC satisfies

$$S' = 1 - \sum_k \frac{k}{\langle k \rangle} P(k)(1 - S')^{k-1}.$$

Prob reach k node

$$q_k = \frac{kp(k)}{< k >}$$

Prob other links link to GC

$$1 - (1 - S')^{k-1}$$

$$S' = \sum_k \frac{k}{\langle k \rangle} P(k) \left[1 - (1 - S')^{k-1}\right]$$

$$S' = 1 - \sum_k \frac{k}{\langle k \rangle} P(k)(1 - S')^{k-1}.$$

Probability node is not in GC == prob all its edges link to non GC

$$1 - S = \sum_k P(k)(1 - S')^k. \qquad S = 1 - \sum_k P(k)(1 - S')^k.$$

Again we need a graphical solution:

$$f(S') = S'$$

$$g(S') = 1 - \sum_k \frac{k}{\langle k \rangle} P(k)(1 - S')^{k-1}$$

$$\left.\frac{dS'}{dS'}\right|_{S'=0} = \left.\frac{d(1 - \sum_k \frac{k}{\langle k \rangle} P(k)(1 - S')^{k-1})}{dS'}\right|_{S'=0},$$

$$1 = \left.\sum_k \frac{k(k-1)}{\langle k \rangle} P(k)\right|_{S'=0},$$

$$1 = \frac{\langle k(k-1) \rangle}{\langle k \rangle}$$

Molloy-Reed criterion

$$\frac{\langle k(k-1) \rangle}{\langle k \rangle} > 1,$$

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

(Check proposition 8.4.1-2-3 here)