The Scale-Free Property

Vander Luis de Souza Freitas
vander.freitas@ufop.edu.br

UFOP

Agenda

- So far:
  - Introduction
    - Overview of the course program
  - Graph theory
    - General network characteristics
    - Paths and distances
    - Clustering coefficient

Agenda

- Last class we discussed:
  - Random networks

Agenda

- Today:
  - Discussion about the homework
  - Watts-Strogatz model
  - The Scale-Free Property
  - Generative algorithms

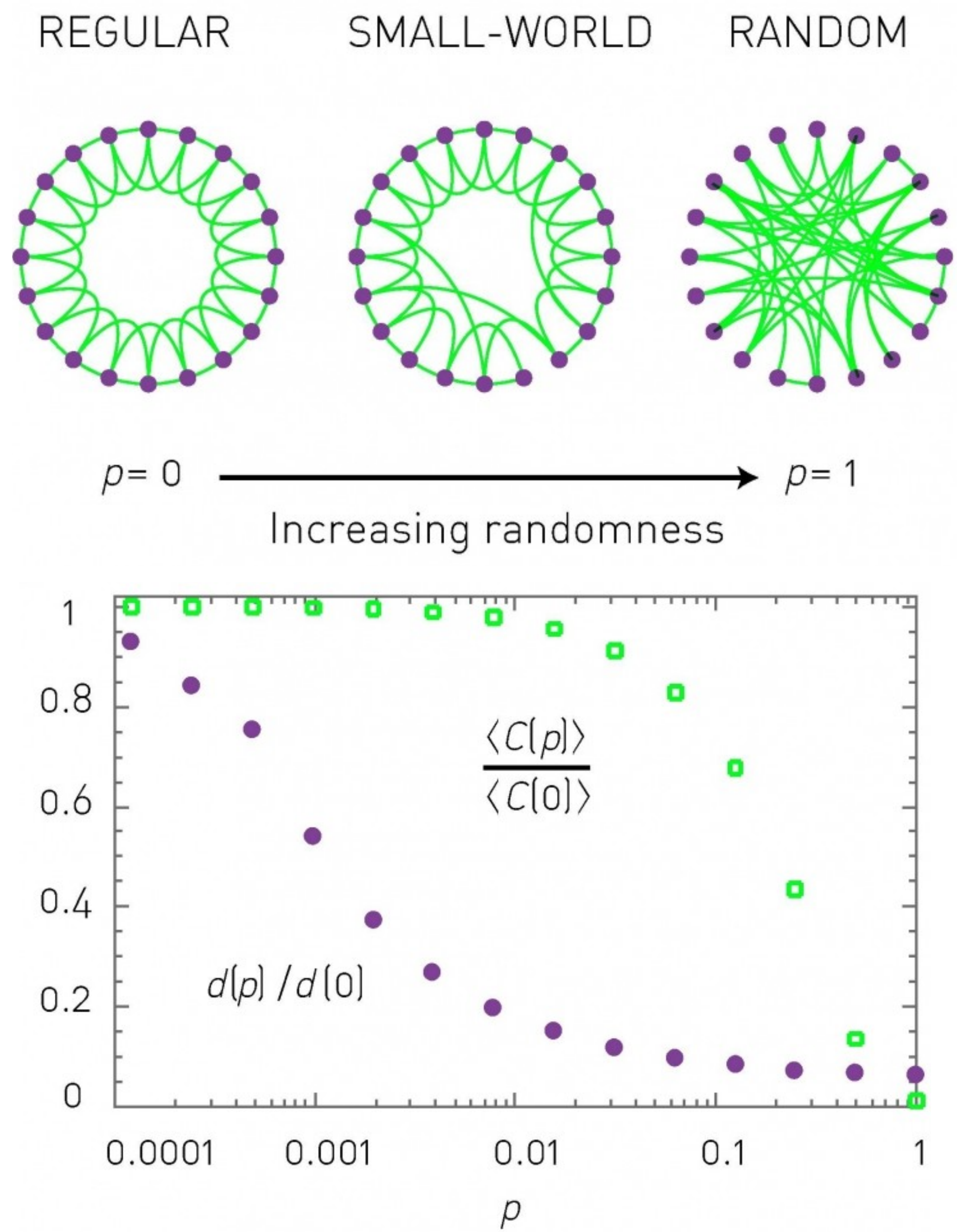
Discussion

Watts-Strogatz model

Small-world network:

an extension of the random network model, with

- High clustering coefficient;
- Small average path length.



http://networksciencebook.com/chapter/3

The Scale-Free Property

Zomming into the first map of the WWW (nd.edu domain):

$N = 300,000, \quad L = 1,500,000$

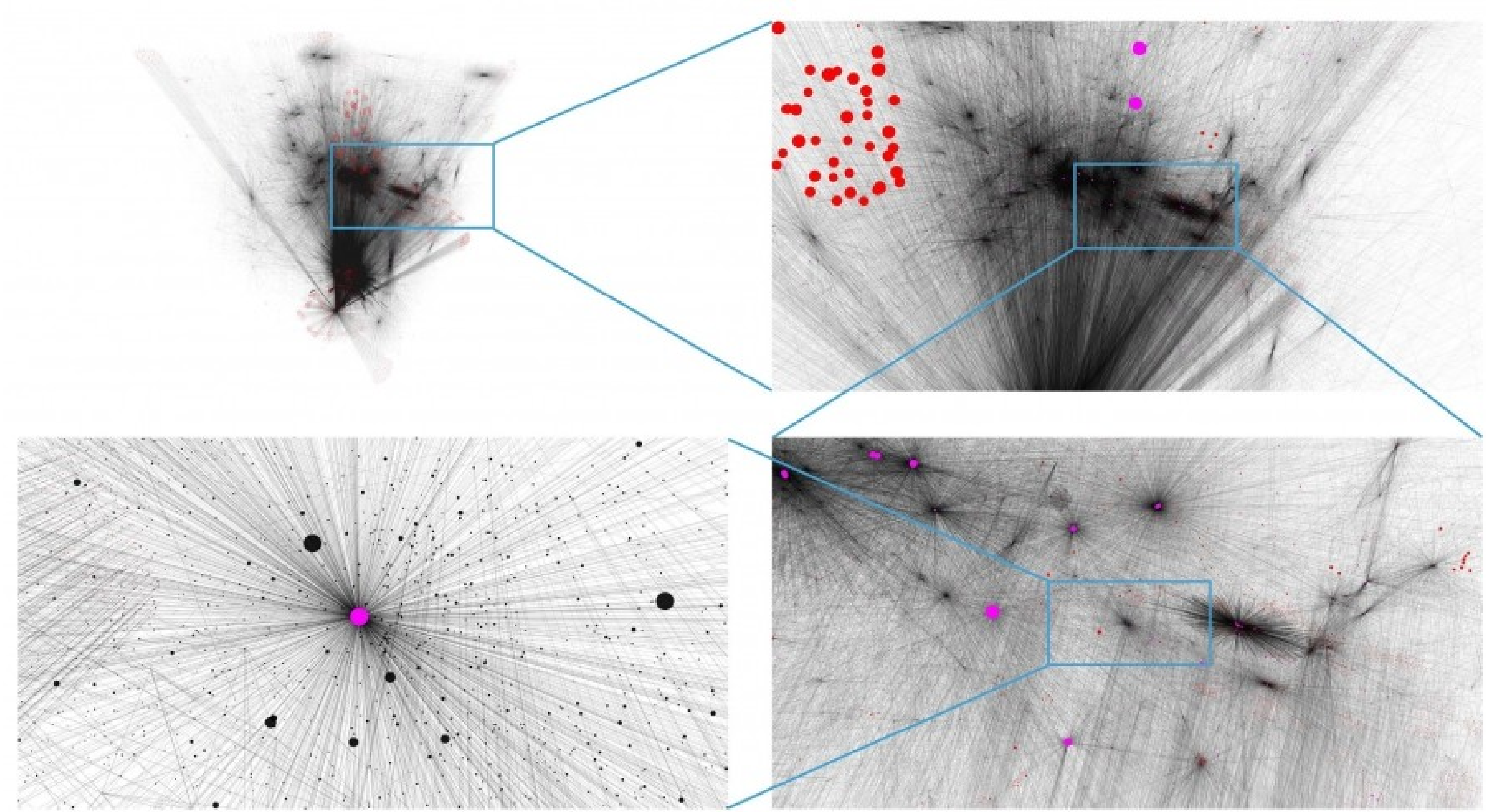
(http://networksciencebook.com/images/ch-04/video-4-1.webm)

The Scale-Free Property

Zomming into the first map of the WWW (nd.edu domain):

$N = 300,000, \quad L = 1,500,000$

Nodes with $k > 50$ are in red and those with $k > 500$ are in purple.



http://networksciencebook.com/chapter/4

The Scale-Free Property

Power Laws and Scale-Free Networks:

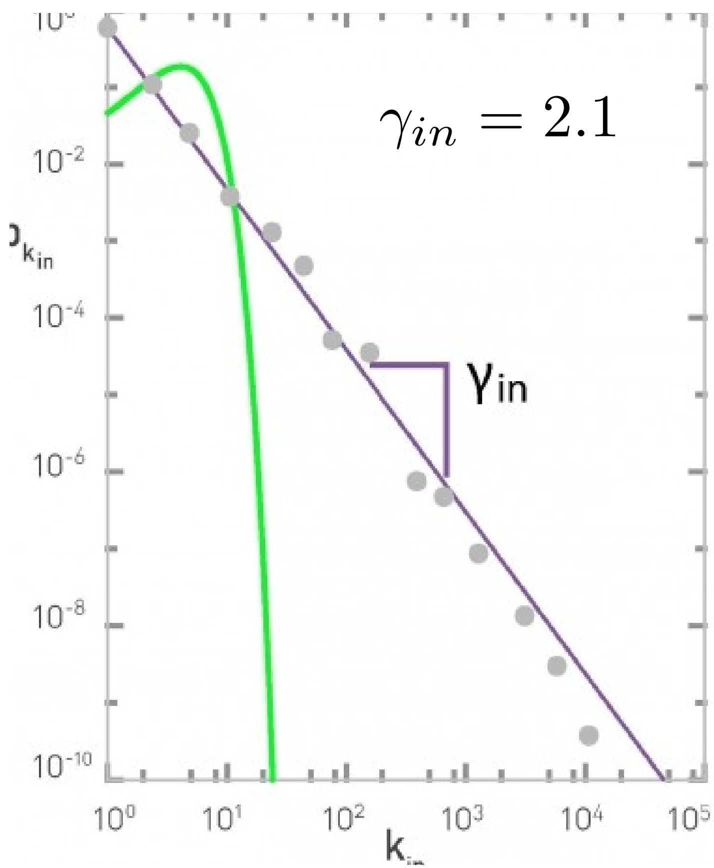
$$p_k \sim k^{-\gamma} \tag{1}$$

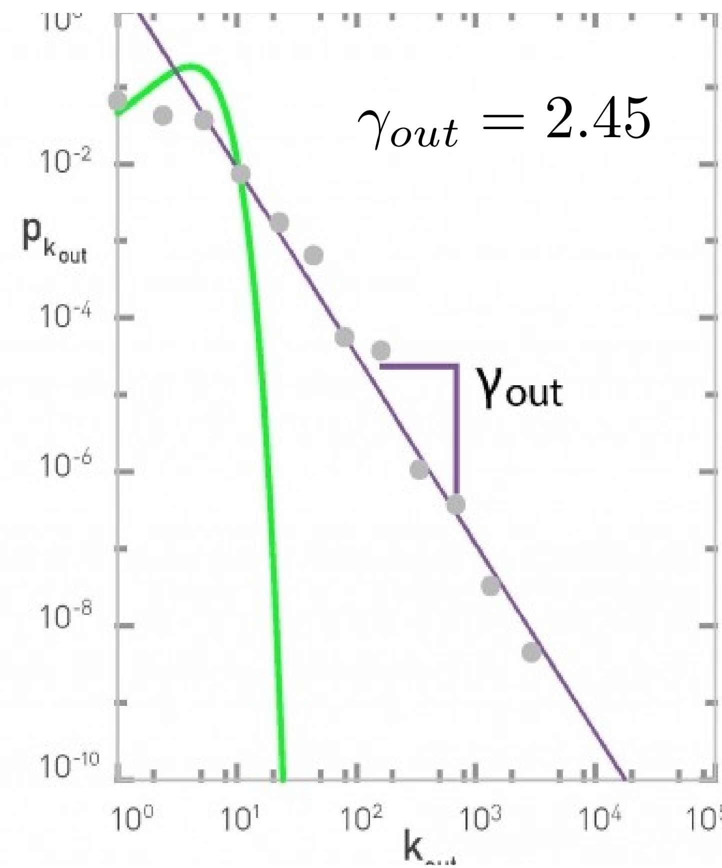
$$\log p_k \sim -\gamma \log k \tag{2}$$

$\log p_k$ is expected do depend linearly on $\log k$, with degree exponent γ.

The degree distribution of the WWW: $\langle k_{in} \rangle = \langle k_{out} \rangle = 4.6$

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





http://networksciencebook.com/chapter/4

Power-law distribution - discrete formalism

The probability that a node has exactly *k* links:

$$p_k = Ck^{-\gamma} \tag{3}$$

The constant C is determined by the normalization condition:

$$\sum_{k=1}^{\infty} p_k = 1 \tag{4}$$

Using (3) we obtain

$$C \sum_{k=1}^{\infty} k^{-\gamma} = 1$$

Hence

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

where $\zeta(\gamma)$ is the Riemann-zeta function. Thus

$$p_k = \frac{k^{-\gamma}}{\zeta(\gamma)} \tag{5}$$

Diverges at k=0

http://networksciencebook.com/chapter/4

Power-law distribution - continuum formalism

Power-law degree distribution:

$$p(k) = Ck^{-\gamma} \tag{6}$$

Using the normalization condition:

$$\int_{k_{min}}^{\infty} p(k)dk = 1 \tag{7}$$

We obtain

$$C = \frac{1}{\int_{k_{min}}^{\infty} k^{-\gamma} dk} = (\gamma - 1)k_{min}^{\gamma-1} \tag{8}$$

In the continuum formalism the degree distribution has the form

$$p(k) = (\gamma - 1)k_{min}^{\gamma-1} k^{-\gamma} \tag{9}$$

where k_{min} is the smallest degree for which (5) holds.

http://networksciencebook.com/chapter/4

Power-law distribution - continuum formalism

$p(k)$ is the probability that a randomly selected node has degree *k*.

Probability that a randomly chosen node has degree between k_1 and k_2:

$$\int_{k_1}^{k_2} p(k)dk \tag{10}$$

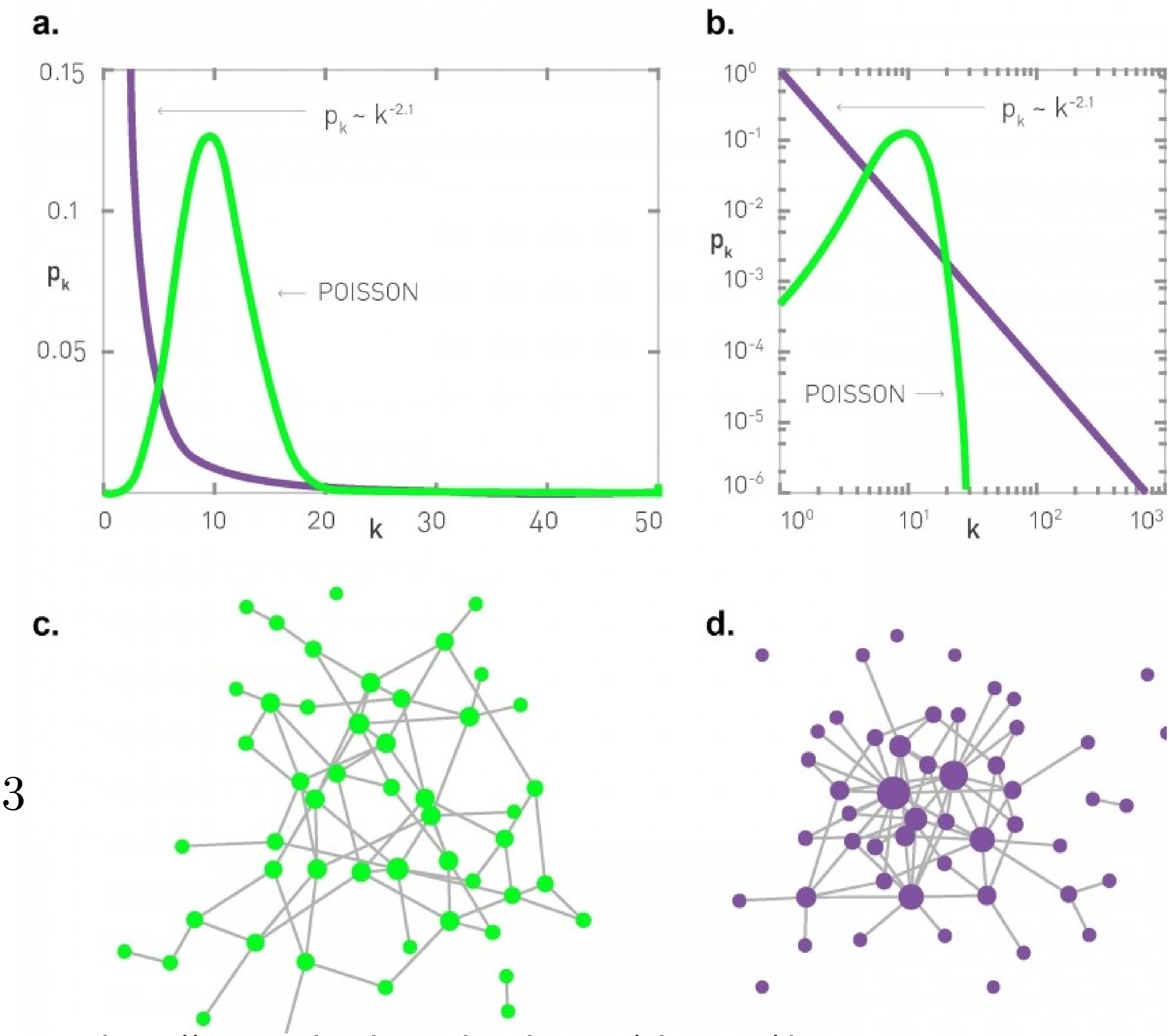
The continuum formalism provides physical interpretation.

The Scale-Free Property

Poisson vs Power-law distributions:

$\langle k \rangle = 11$

$N = 50,\ \langle k \rangle = 3$



http://networksciencebook.com/chapter/4

The Scale-Free Property

Example: the WWW, with $\langle k \rangle = 4.6, \quad N \approx 10^{12}$

In a Poisson distribution, the number of nodes with at least 100 links is pretty much none:

$$N_{k \geq 100} = 10^{12} \sum_{k=100}^{\infty} \frac{(4.6)^k}{k!} e^{-4.6} \simeq 10^{-82}$$

However, in a power-law distribution:

$$N_{k \geq 100} = 4 \times 10^9$$

i.e. more that four billion nodes, considering $\gamma_{in} = 2.1$.

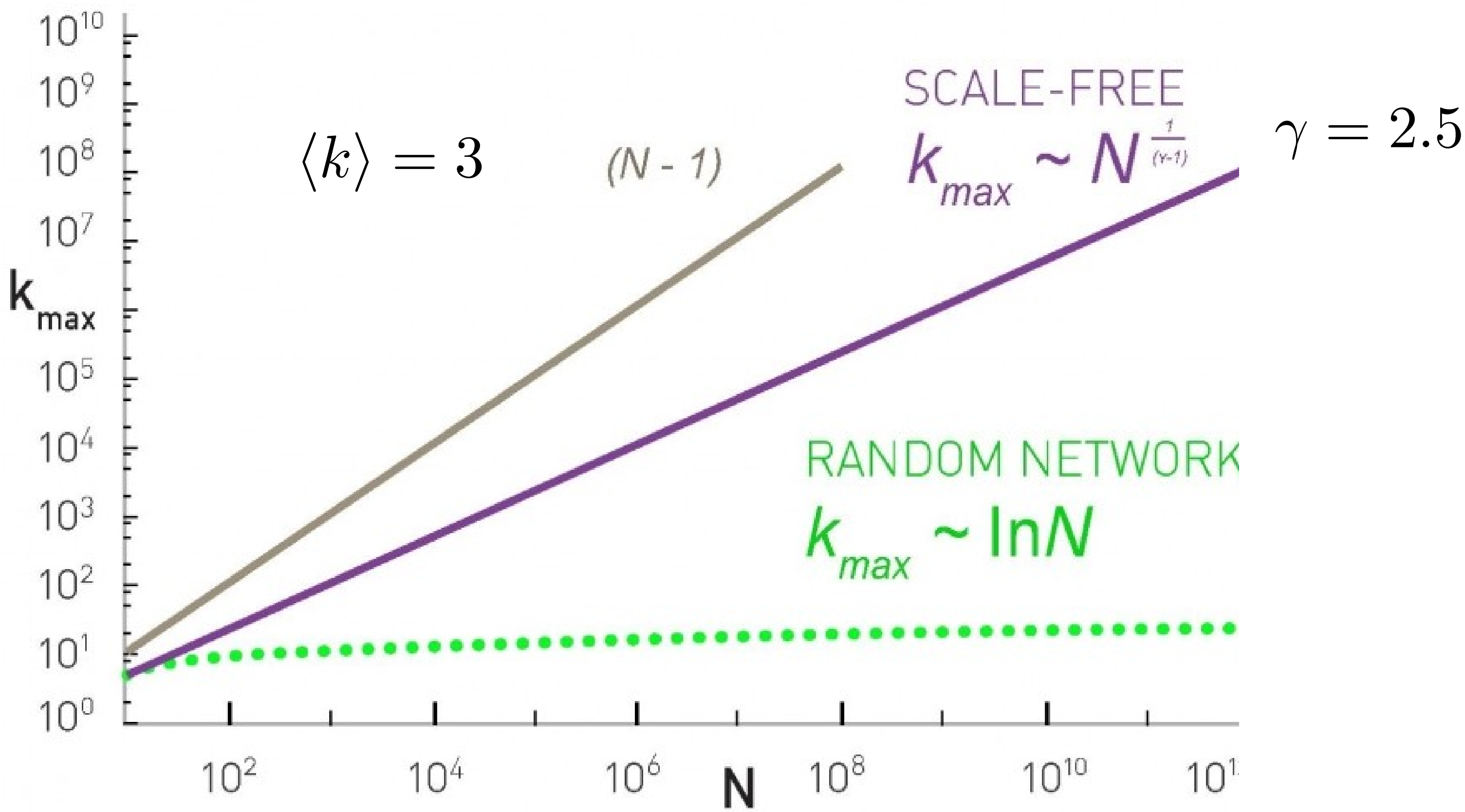
http://networksciencebook.com/chapter/4

The largest hub

We expect that there is at maximum one single node with degree k_{max}, due to the fact that networks are finite. From Eqs. (6) to (9) we obtain a natural cutoff, that follows:

$$k_{max} = k_{min} N^{\frac{1}{\gamma - 1}} \tag{11}$$

Hence, considering that the degree distribution is a power-law: the larger the network, the larger is the degree of its biggest hub.



http://networksciencebook.com/chapter/4

The Meaning of Scale-Free

The nth moment of the degree distribution is defined as

$$\langle k^n \rangle = \sum_{k_{min}}^{\infty} k^n p_k \approx \int_{k_{min}}^{\infty} k^n p(k) dk$$

Important interpretation:

- *n=1*: The first moment is the average degree;
- *n=2*: The second moment helps us to calculate the variance

  $$\sigma^2 = \langle k^2 \rangle - \langle k \rangle^2$$

  that measures the spread in degrees;
- *n=3*: The third moment determines the *skewness* of a distribution, telling us how symmetric is p_k around the average $\langle k \rangle$.

http://networksciencebook.com/chapter/4

The Meaning of Scale-Free

The n[th] moment of the degree distribution of a scale-free network is:

$$\langle k^n \rangle = \int_{k_{min}}^{k_{max}} k^n p(k) dk = C \frac{k_{max}^{n-\gamma+1} - k_{min}^{n-\gamma+1}}{n-\gamma+1}$$

While typically k_{min} is fixed, k_{max} increases with the system size, following (11). For very large networks, with $k_{max} \to \infty$, the n[th] moment depends on:

- If $n-\gamma+1 \leq 0$ then $k_{max}^{n-\gamma+1}$ goes to zero as k_{max} increases. Therefore all moments that satisfy $n \leq \gamma - 1$ are finite.
- If $n-\gamma+1 > 0$ then $\langle k^n \rangle$ goes to infinity as $k_{max} \to \infty$. Therefore all moments larger than $\gamma - 1$ diverge.

This means that in the limit $N \to \infty$, the scale-free networks have average degree but the other moments go to infinity. Most scale-free networks have $2 < \gamma < 3$

http://networksciencebook.com/chapter/4

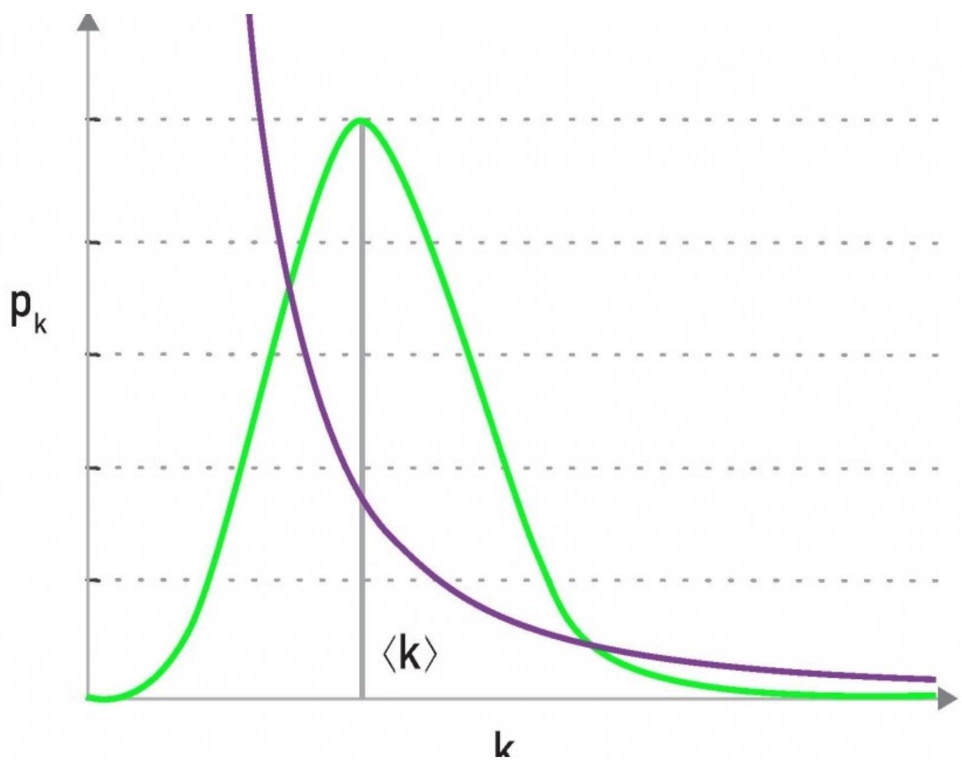
The Meaning of Scale-Free

Random networks have a scale:

$\sigma_k = \langle k \rangle^{1/2}$ is always smaller than $\langle k \rangle$ and means that nodes have degree in the range $k \in [\langle k \rangle - \langle k \rangle^{1/2}, \langle k \rangle + \langle k \rangle^{1/2}]$. The average degree serves as a scale.

Scale-free networks lack a scale:

The n^{th} moments diverge. There is no meaningful internal scale when $\gamma < 3$. They are "scale-free".



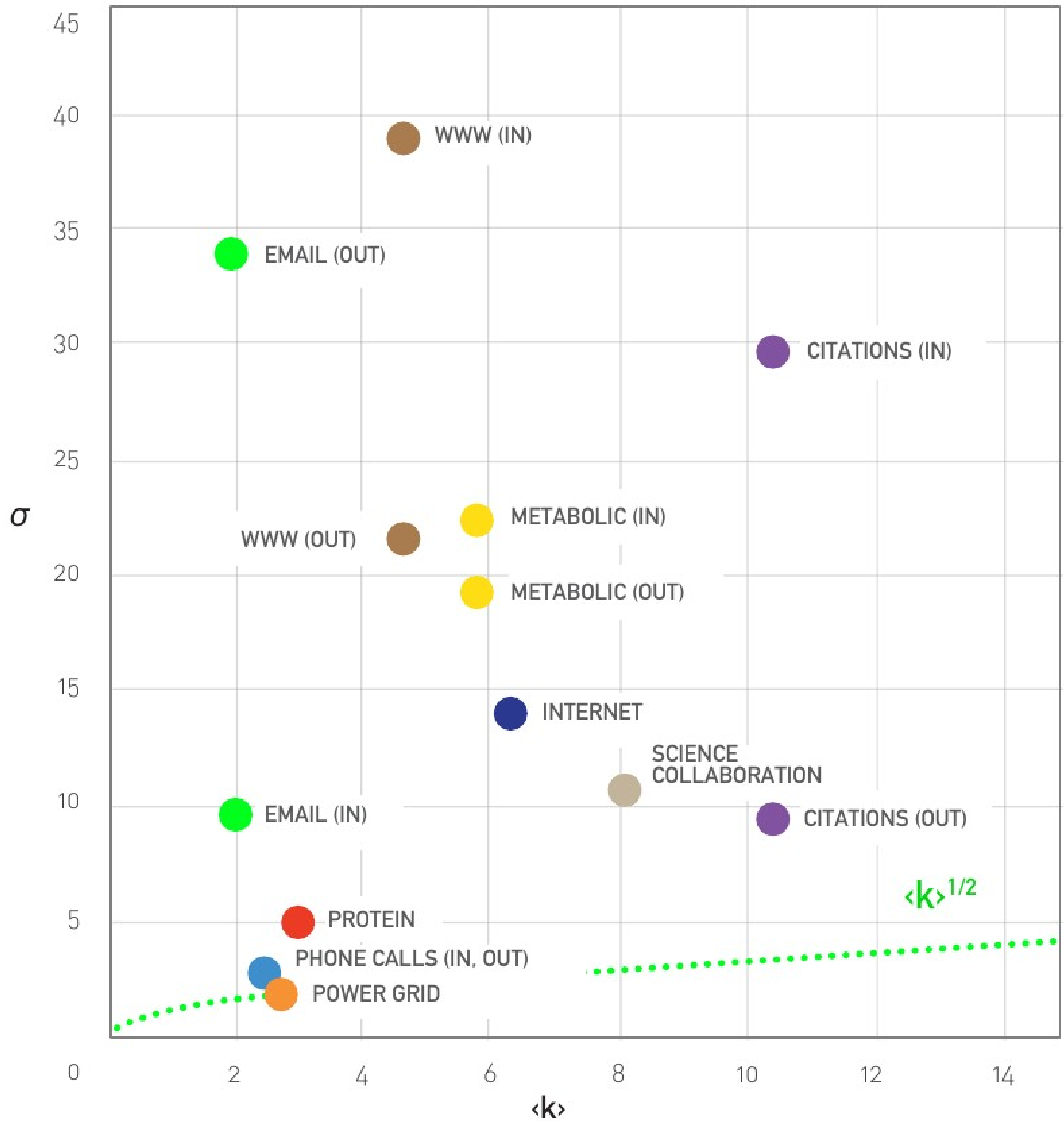
http://networksciencebook.com/chapter/4

The Meaning of Scale-Free

Network	N	L	$\langle k \rangle$	$\langle k_{in}^2 \rangle$	$\langle k_{out}^2 \rangle$	$\langle k^2 \rangle$	γ_{in}	γ_{out}	γ
Internet	192,244	609,066	6.34	-	-	240.1	-	-	3.42*
WWW	325,729	1,497,134	4.60	1546.0	482.4	-	2.00	2.31	-
Power Grid	4,941	6,594	2.67	-	-	10.3	-	-	Exp.
Mobile-Phone Calls	36,595	91,826	2.51	12.0	11.7	-	4.69*	5.01*	-
Email	57,194	103,731	1.81	94.7	1163.9	-	3.43*	2.03*	-
Science Collaboration	23,133	93,437	8.08	-	-	178.2	-	-	3.35*
Actor Network	702,388	29,397,908	83.71	-	-	47,353.7	-	-	2.12*
Citation Network	449,673	4,689,479	10.43	971.5	198.8	-	3.03*	4.00*	-
E. Coli Metabolism	1,039	5,802	5.58	535.7	396.7	-	2.43*	2.90*	-
Protein Interactions	2,018	2,930	2.90	-	-	32.3	-	-	2.89*-

http://networksciencebook.com/chapter/4

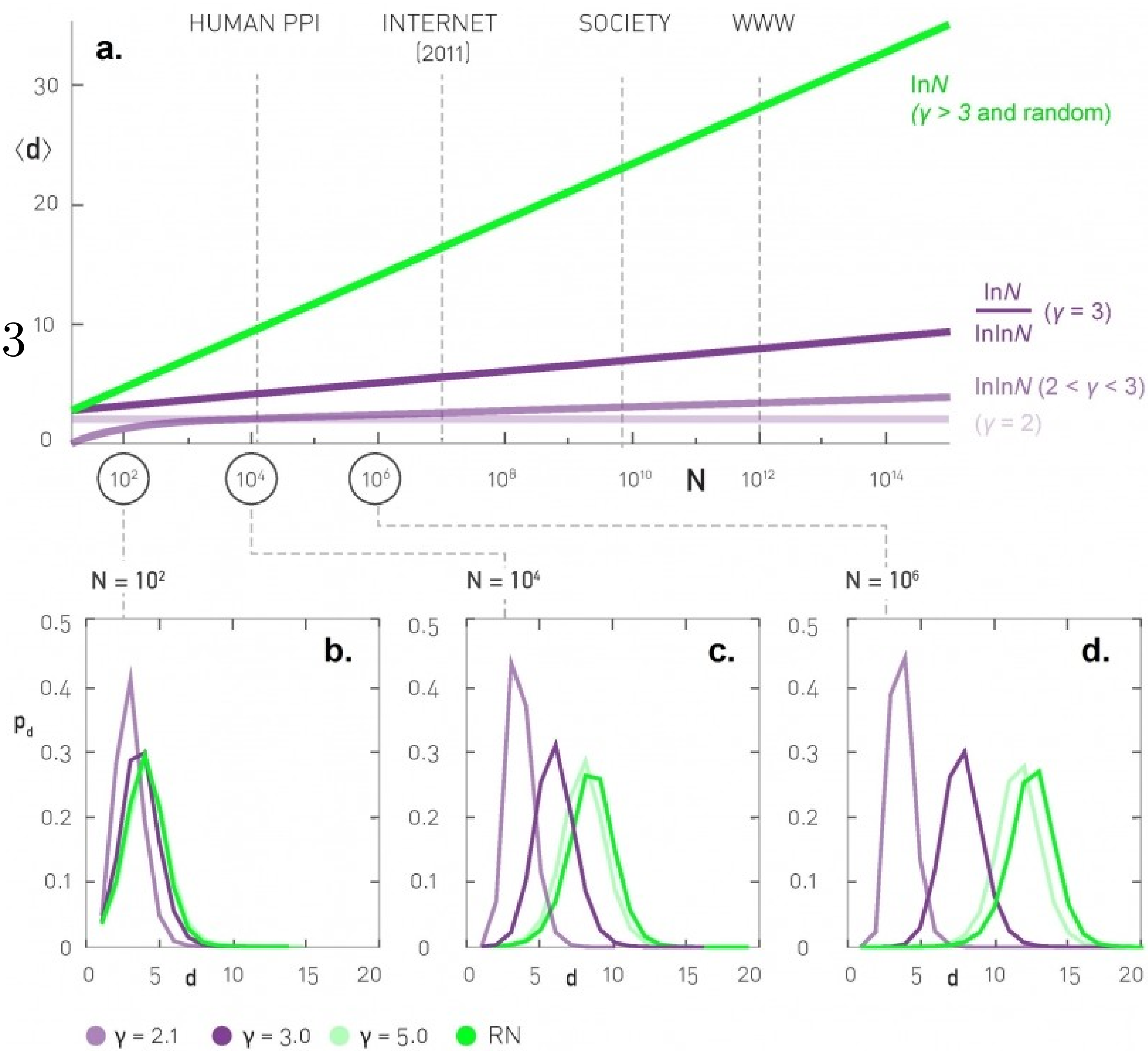
The Meaning of Scale-Free



http://networksciencebook.com/chapter/4

Ultra-small property

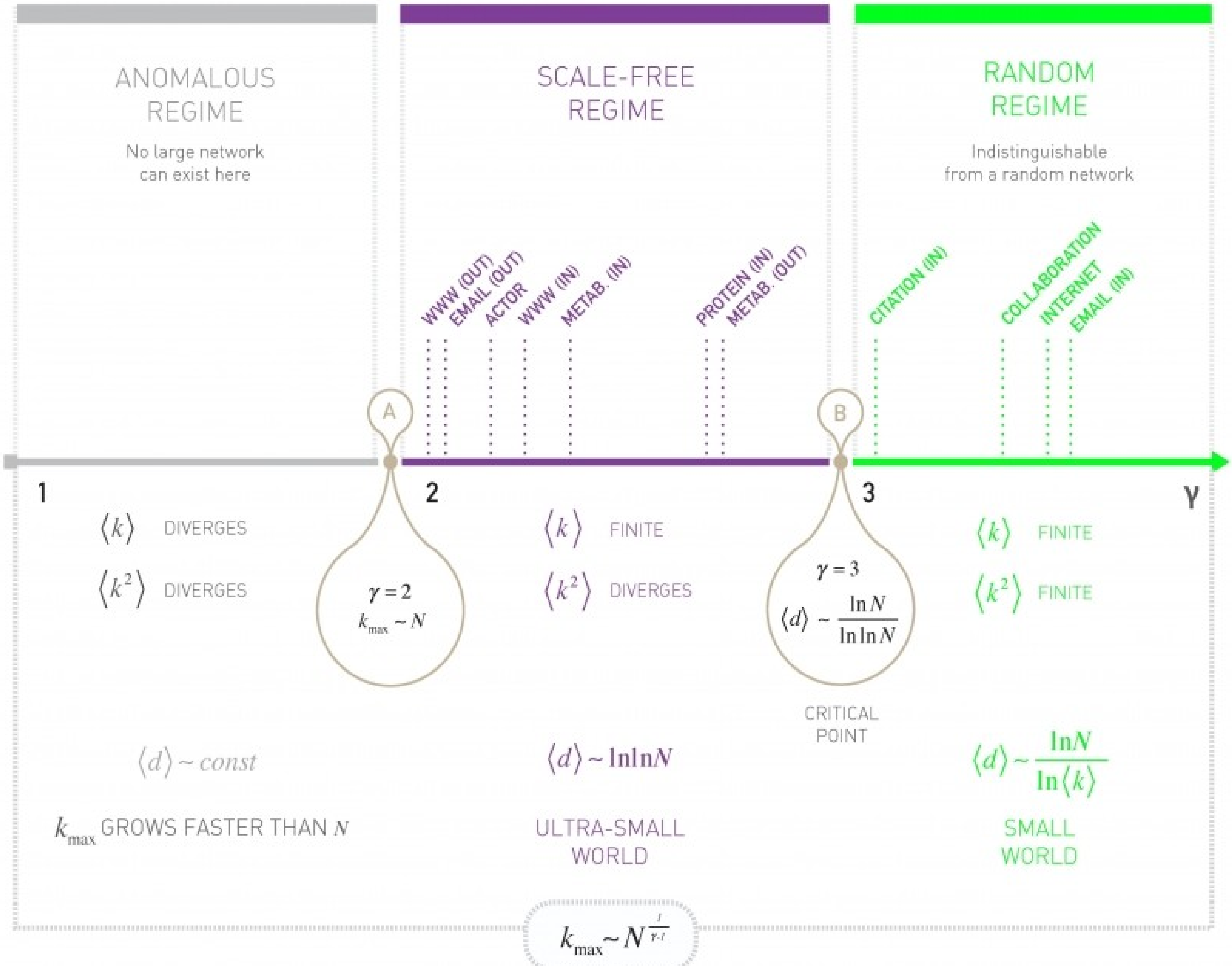
$$\langle d \rangle \sim \begin{cases} const. & \gamma = 2 \\ \ln \ln N & 2 < \gamma < 3 \\ \dfrac{\ln N}{\ln \ln N} & \gamma = 3 \\ \ln N & \gamma > 3 \end{cases}$$



Most scale-free networks of practical interest are not only “small world”, but are “ultra-small world”.

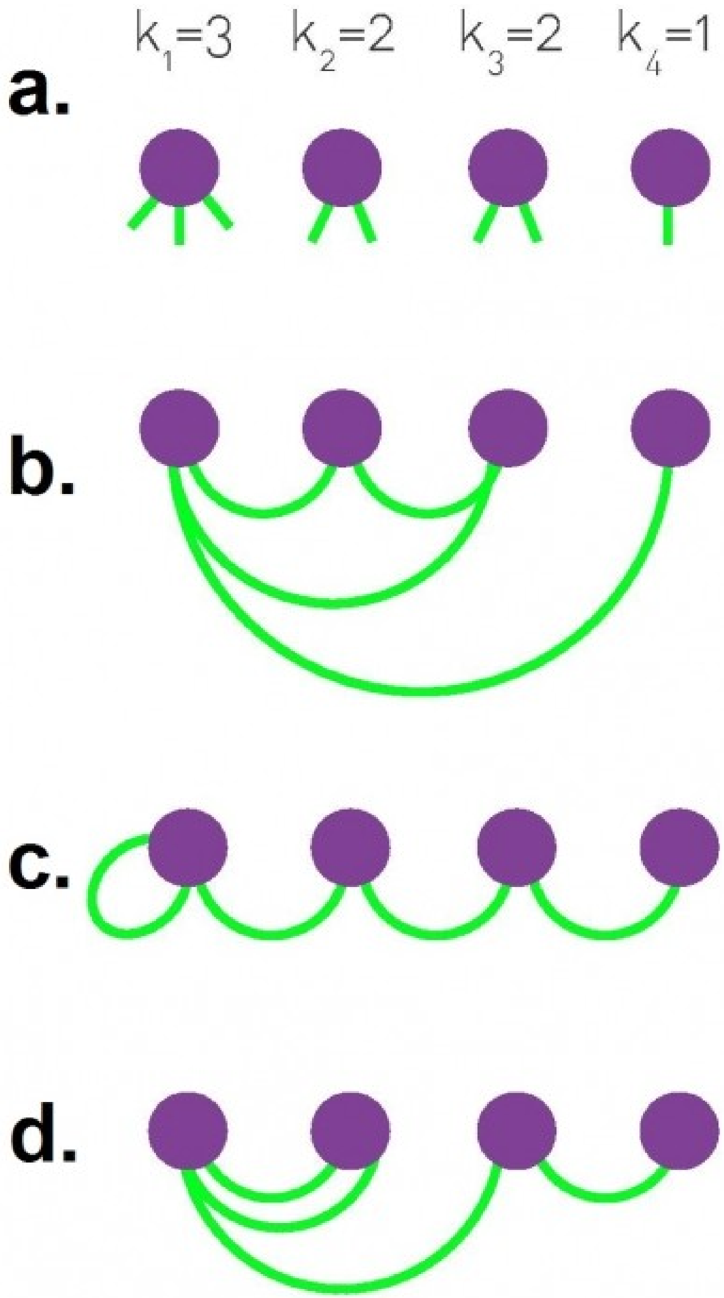
http://networksciencebook.com/chapter/4

The role of the degree exponent



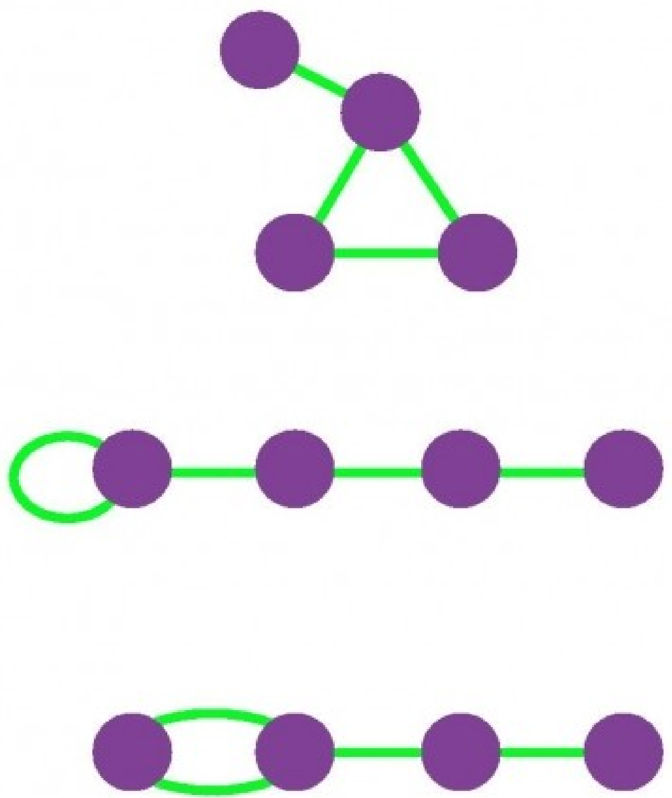
http://networksciencebook.com/chapter/4

Generating Networks with Arbitrary Degree Distribution



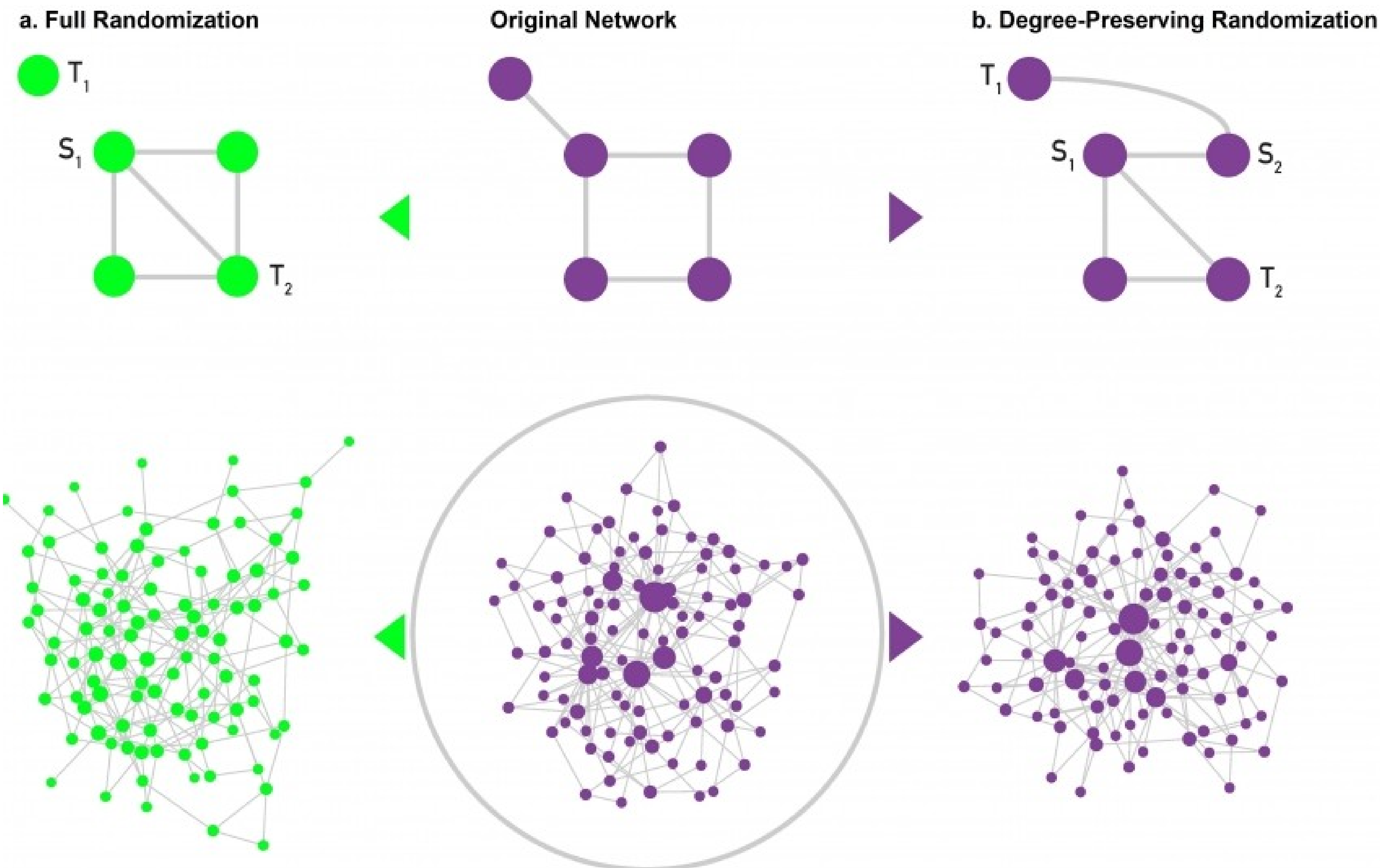
The Configuration Model builds a network whose nodes have predefined degrees.

However, it allows self-loops and multi-links.



http://networksciencebook.com/chapter/4

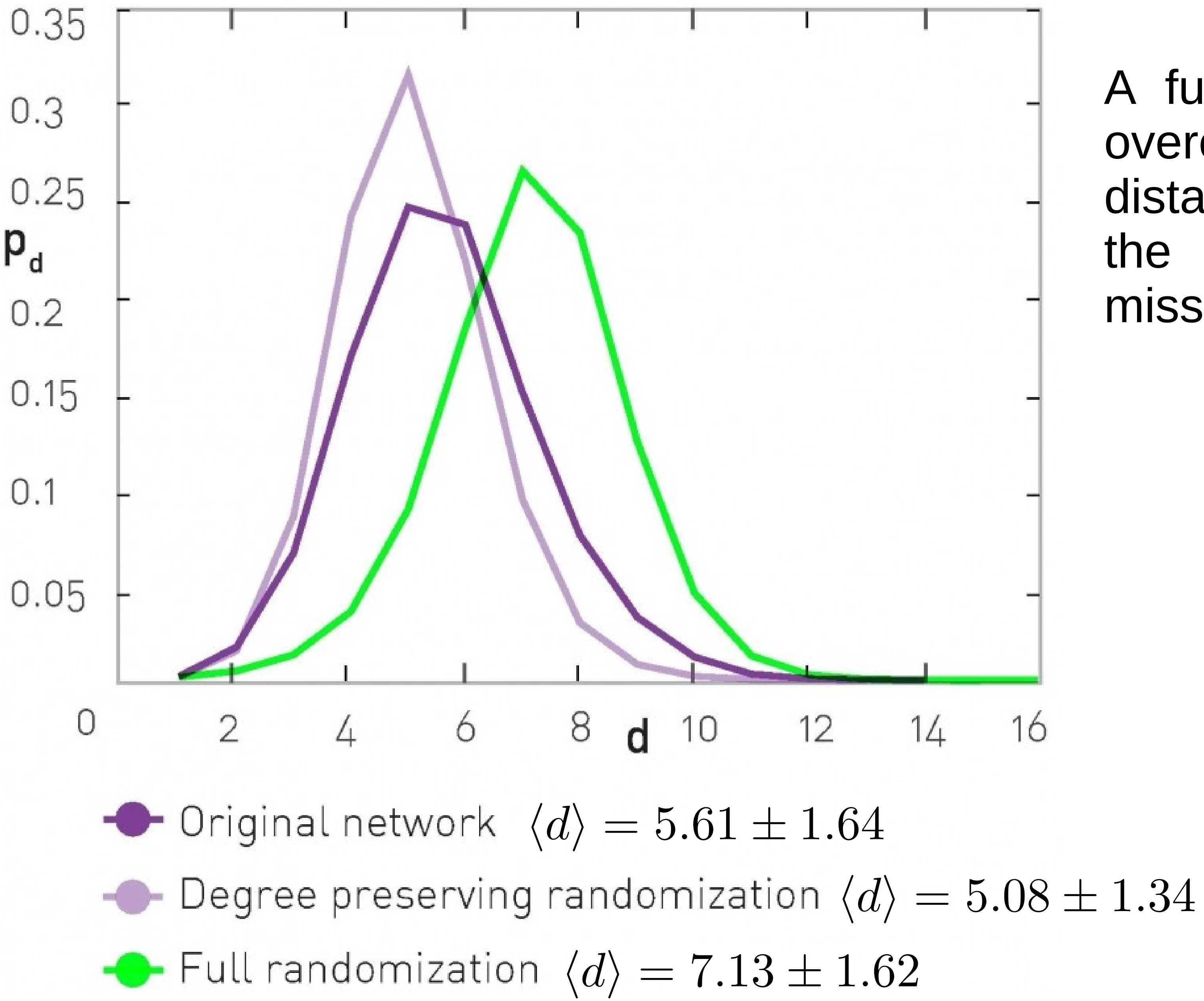
Degree-Preserving Randomization



http://networksciencebook.com/chapter/4

Randomizing Real Networks

Distance distribution:



A full randomization overestimates the distances between the nodes, as it is missing the hubs.

http://networksciencebook.com/chapter/4

Final considerations

- Exponentially Bounded Networks;
- Fat Tailed Networks;

PUREen Access | Published: 04 March 2019

Scale-free networks are rare

Anna D. Broido & Aaron Clauset

Nature Communications **10**, Article number: 1017 (2019) | Cite this article

48k Accesses | **194** Citations | **615** Altmetric | Metrics

At a Glance: Scale-Free Networks

Degree Distribution

Discrete form:

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

Continuous form:

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

Size of the Largest Hub

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

Moments of p_k for $N \to \infty$

$2 < \gamma \leq 3$: $\langle k \rangle$ finite, $\langle k^2 \rangle$ diverges.

$\gamma > 3$: $\langle k \rangle$ and $\langle k^2 \rangle$ finite.

Distance

$$\langle d \rangle \sim \begin{cases} const. & \gamma = 2 \\ \ln \ln N & 2 < \gamma < 3 \\ \frac{\ln N}{\ln \ln N} & \gamma = 3 \\ \ln N & \gamma > 3 \end{cases}$$

http://networksciencebook.com/chapter/4

Centrality

How important a node is?

One may compute node centrality measures such as:

- Degree
- Betweenness
- Closeness
- Eigenvalue
- PageRank

There are many others and there is nothing like the best metric. It depends on the problem.

Centrality

Degree k: Number of nodes a node is connected to.

	A	B	C	D	E	F	G
Ex:	[2,	4,	2,	4,	5,	3,	2]

$$k_i = \sum_{j=1}^{N} A_{ij}$$

Strength s: Accumulated weights from incident edges.

	A	B	C	D	E	F	G
Ex:	[12,	31,	13,	35,	44,	25,	20]

$$s_i = \sum_{j=1}^{N} W_{ij}$$

Betweenness b: Rate of shortest paths that pass through a node.

	A	B	C	D	E	F	G
Ex:	[0,	2.33,	0,	2.67,	5.17,	0.83,	0]

$$b_i = \sum_{v \neq w \neq i} \frac{\sigma_{vw}(i)}{\sigma_{vw}}$$

Diameter D: Length of the longest shortest path. Ex: [3]