AMSC808N/CMSC828V

Processes on complex networks

Keywords:
network growth,
preferential attachment,
power-law degree distribution,
random failure vs attack,
percolation

Maria Cameron

References

- M. Newman, *The structure and function of complex networks*, SIAM Review, 45/2, 167—256, 2003
- A.-L. Barabasi and R. Albert, *Emergence of Scaling in Random Networks*, Science, 286, 509—512, 1999
- A. L. Barabasi, Network Science, 2017
- R. Albert, H. Jeong, and A.-L. Barabasi, *Error and attack tolerance of complex networks*, Nature, 406, 378—382, 2000
- R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, *Resilience of the Internet to Random Breakdowns*, Physical Review Letters, 85/21, 4626—4628
- D. Callaway, M. Newman, S. Strogatz, and D. Watts, *Network Robustness and Fragility: Percolation on Random Graphs,* Physical Review Letters, 85/25, 5468—5471
- M. Newman, Spread of epidemic disease on networks, Physical review E, 66, 016128 (2002)

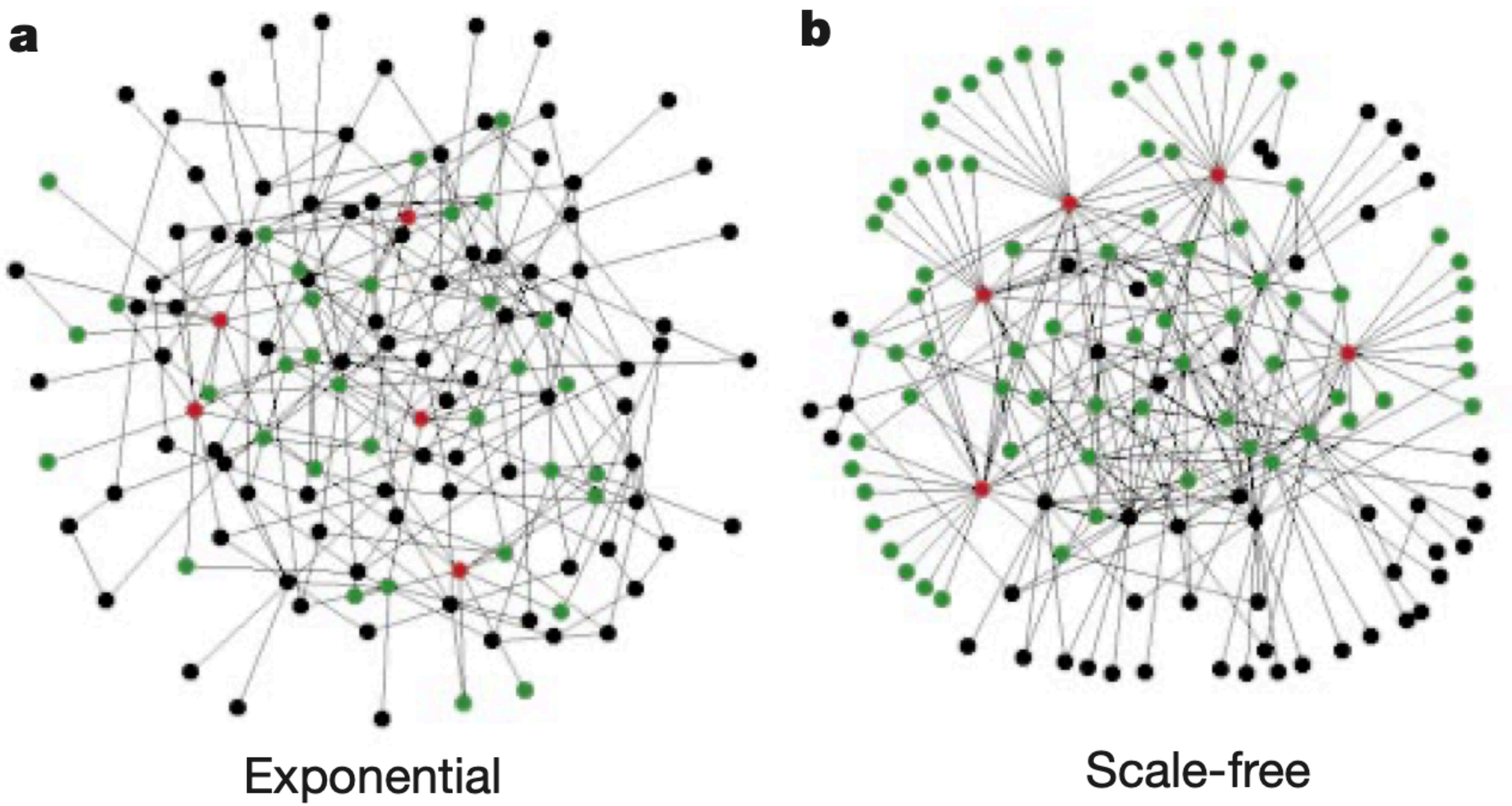
Degree fluctuations in real networks

A. L. Barabasi, Network Science

Network	# verts N	# edges L	Mean degree $\langle k \rangle$	$\langle k_{in}^2 \rangle$	$\langle k_{out}^2 \rangle$	Mean 2nd moment $\langle k^2 \rangle$	γ_{in}	γ_{out}	$p_k \sim k^{-\gamma}$ 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*-

Exponential vs power-law networks

Figure is from R. Albert, H. Jeong, and A.-L. Barabasi



Red: 5 nodes with highest degree, green: their first neighbors.

Growth and preferential attachment lead to power-law degree distribution

A.-L. Barabasi and R. Albert (1999)

- Observed that numerous real-world networks exhibit power-law degree distribution $p_k \sim k^{-\gamma}$
- Argued that this is the result of two factors: (1) growth and (2) preferential attachment
- Proposed a simple growth model leading to $p_k \sim k^{-3}$

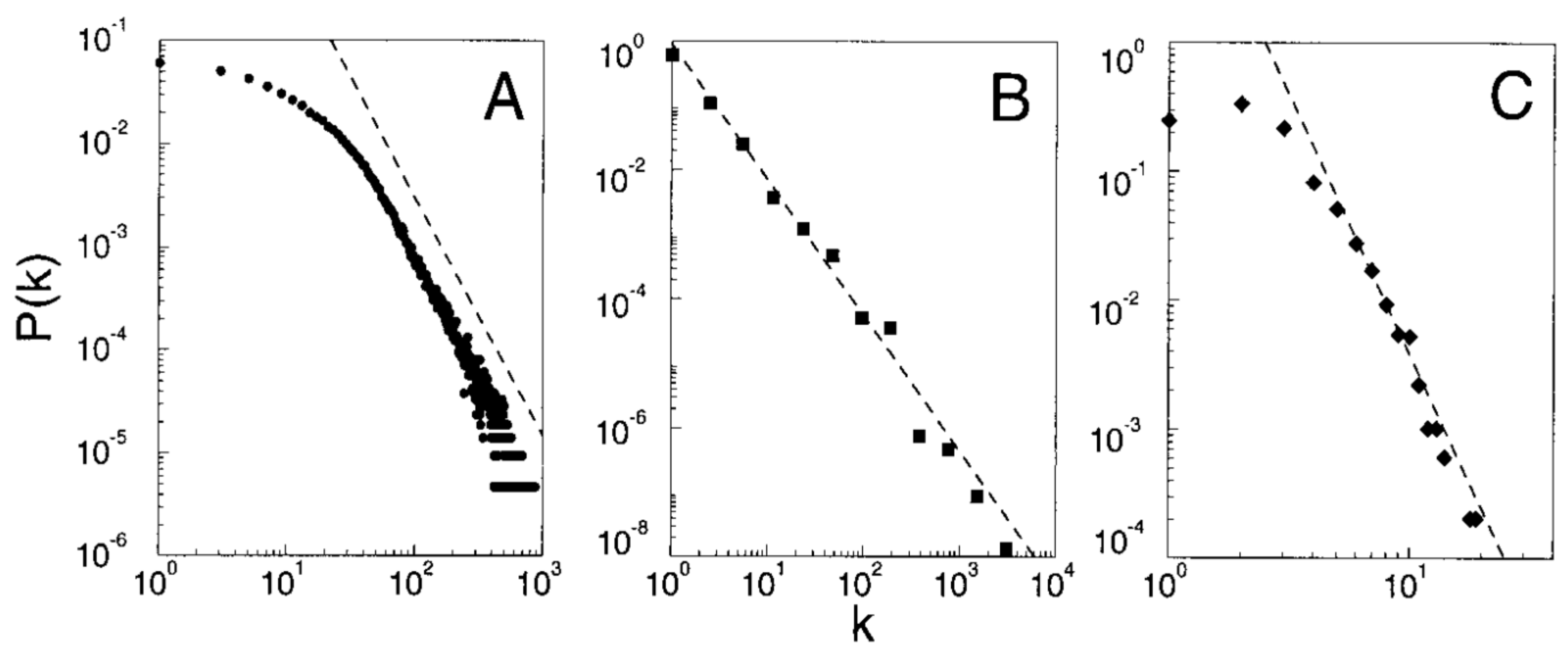
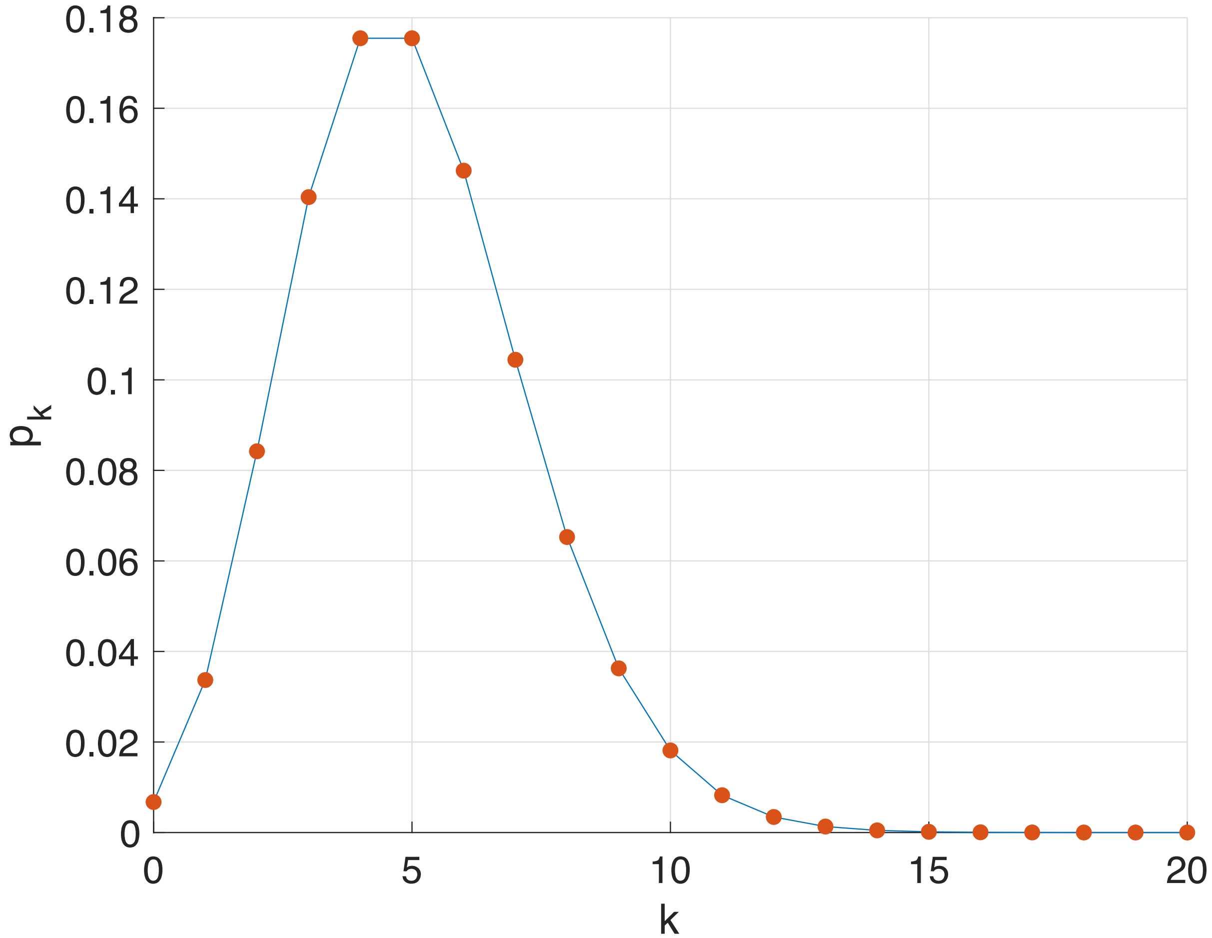
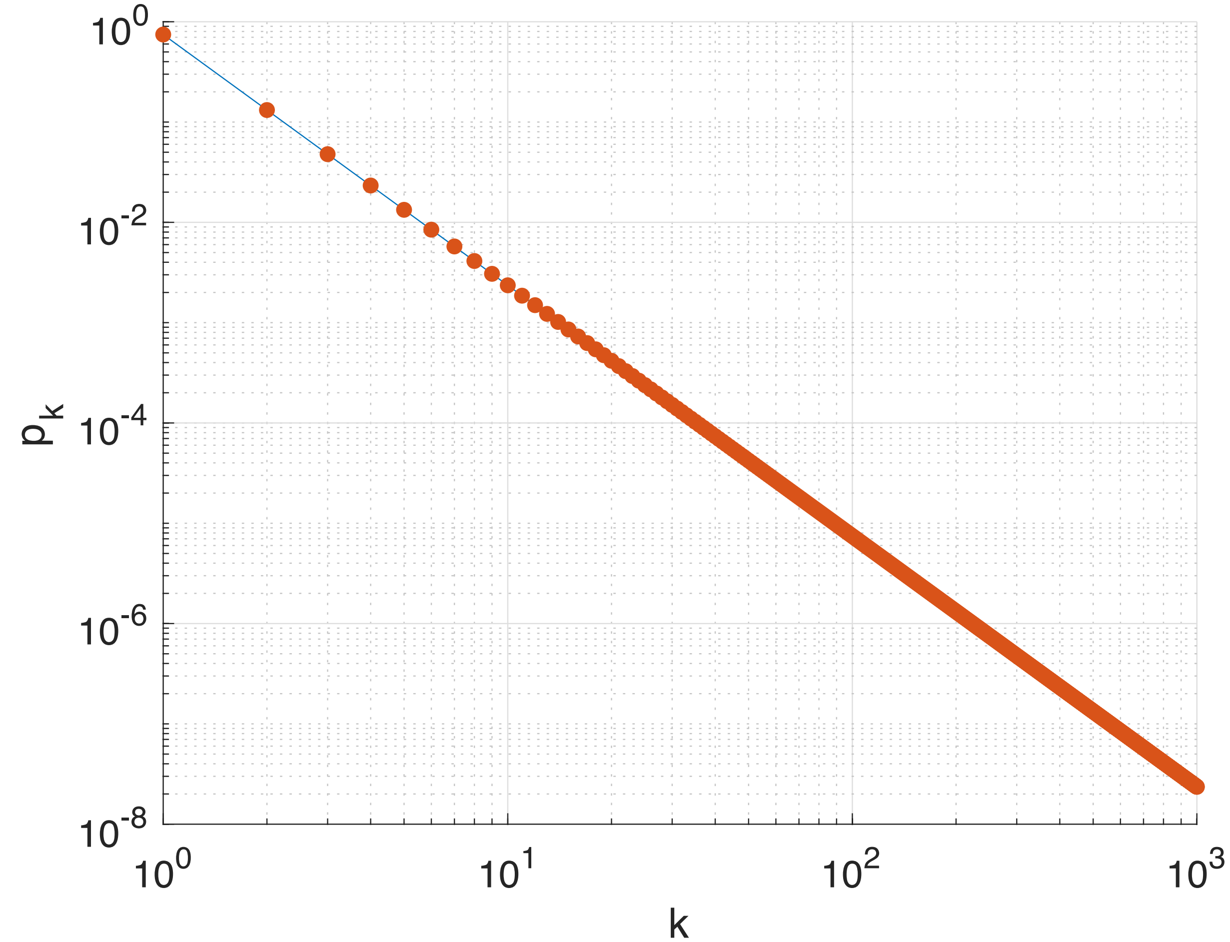


Fig. 1. The distribution function of connectivities for various large networks. (**A**) Actor collaboration graph with $N = 212{,}250$ vertices and average connectivity $\langle k \rangle = 28.78$. (**B**) WWW, $N = 325{,}729$, $\langle k \rangle = 5.46$ (6). (**C**) Power grid data, $N = 4941$, $\langle k \rangle = 2.67$. The dashed lines have slopes (A) $\gamma_{actor} = 2.3$, (B) $\gamma_{www} = 2.1$ and (C) $\gamma_{power} = 4$.

Poisson vs power-law



Poisson distribution is sharply peaked at $z = \langle k \rangle$, indicating that there is a characteristic scale for *k*.

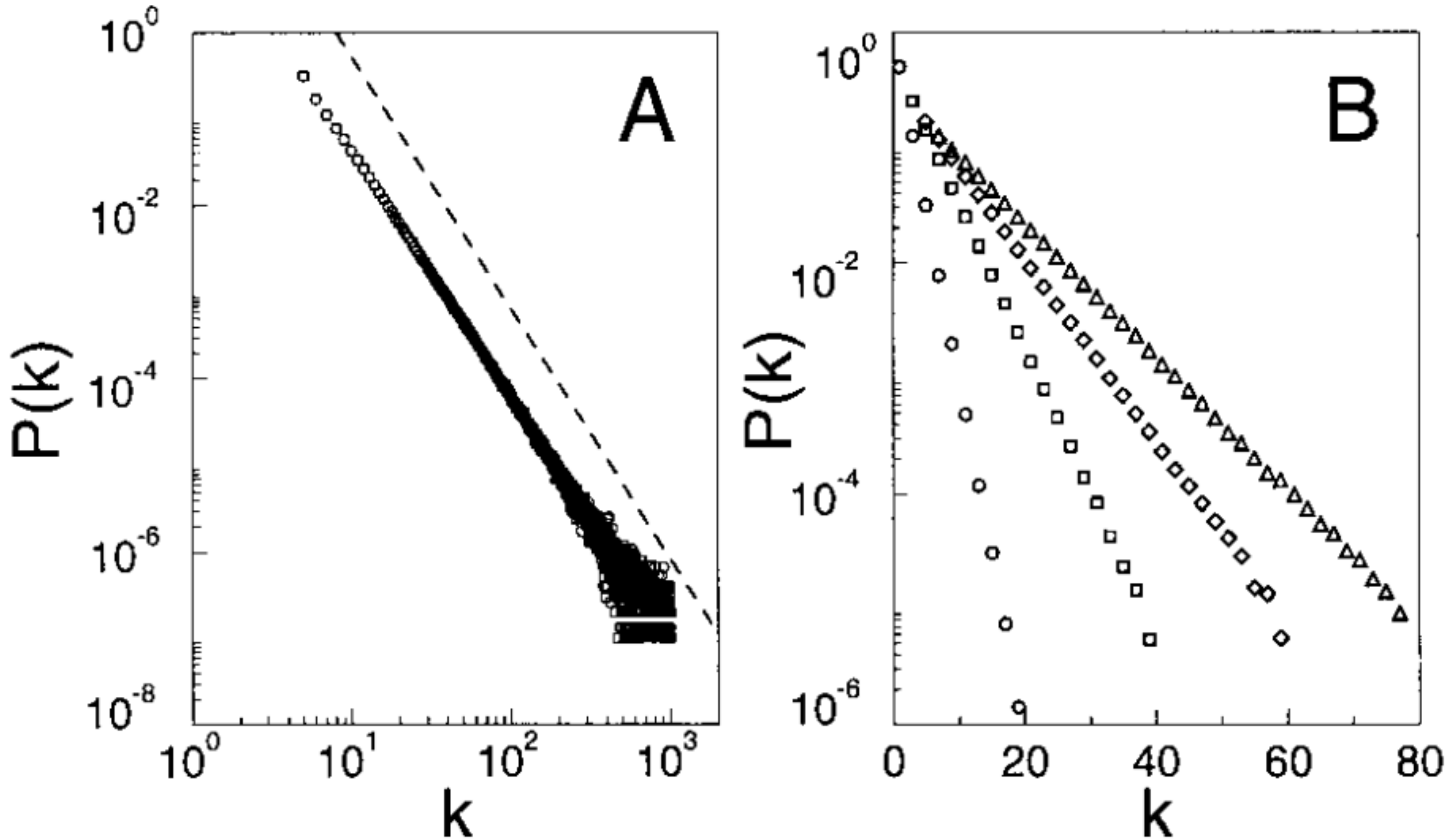


Power-law distribution does not have a characteristic scale.

Barabasi-Albert growth model

Preferential attachment

- **Start** with m vertices and no edges
- **Step 1**: add a vertex and link it to all vertices.
- **Step 2, 3, 4, …**: add a vertex with m edges and link it to m different vertices. The probability that at step t the new vertex will be linked to vertex i is $P(k_i) = k_i/\Sigma_j k_j$, where k_i is the degree of vertex i.
- After t steps, there will be $m + t$ vertices and mt edges.



(A) The power-law connectivity distribution at t ? 150,000 (circles) and t ? 200,000 (squares) as obtained from the model, using $m_0 = m =$?5. The slope of the dashed line is ? ? 2.9.

(B) The exponential connectivity distribution for model A, in the case of $m_0 = m =$?1 (circles), $m_0 = m =$?3 (squares), $m_0 = m =$?5 (diamonds) , and $m_0 = m =$?7 (triangles)?.

Emergence of a power law

A.-L. Barabasi and R. Albert, (1999)

Make time continuous to facilitate calculations

The rate at which a vertex acquires edges is $\frac{dk_i}{dt} = \frac{k_i}{2t}$.

Justification: the rate must be proportional to k_i and all rates must sum up to m.

$$\sum_i \frac{dk_i}{dt} = \frac{1}{2t} \sum_i k_i = \frac{2mt}{2t} = m$$

Initially, $k_i(t_i) = m$. Here, t_i is the time at which vertex i is added.

Emergence of a power law

A.-L. Barabasi and R. Albert, (1999)

Make time continuous to facilitate calculations

The rate at which a vertex acquires edges is $\frac{dk_i}{dt} = \frac{k_i}{2t}$. Initially, $k_i(t_i) = m$.

Justification: the rate must be proportional to k_i and all rates must sum up to m.

$$\sum_i \frac{dk_i}{dt} = \frac{1}{2t} \sum_i k_i = \frac{2mt}{2t} = m$$

Hence, $k_i(t) = m \left(\frac{t}{t_i} \right)^{1/2}$, where t_i is the time when vertex i was added.

Emergence of a power law

A.-L. Barabasi and R. Albert, (1999)

Make time continuous to facilitate calculations

The rate at which a vertex acquires edges is $\frac{dk_i}{dt} = \frac{k_i}{2t}$. Initially, $k_i(t_i) = m$.

Justification: the rate must be proportional to k_i and all rates must sum up to m.

$$\sum_i \frac{dk_i}{dt} = \frac{1}{2t} \sum_i k_i = \frac{2mt}{2t} = m$$

Hence, $k_i(t) = m \left(\frac{t}{t_i} \right)^{1/2}$, where t_i is the time when vertex i was added.

The probability that at time t a vertex i has $< k$ edges is $P[k_i(t) < k] = P\left[t_i > \frac{m^2 t}{k^2} \right]$.

Emergence of a power law

A.-L. Barabasi and R. Albert, (1999)

Make time continuous to facilitate calculations

The rate at which a vertex acquires edges is $\frac{dk_i}{dt} = \frac{k_i}{2t}$. Initially, $k_i(t_i) = m$.

Justification: the rate must be proportional to k_i and all rates must sum up to m.

$$\sum_i \frac{dk_i}{dt} = \frac{1}{2t} \sum_i k_i = \frac{2mt}{2t} = m$$

Hence, $k_i(t) = m \left(\frac{t}{t_i} \right)^{1/2}$, where t_i is the time when vertex i was added.

The probability that at time t a vertex i has $< k$ edges is $P[k_i(t) < k] = P\left[t_i > \frac{m^2 t}{k^2}\right]$.

$$P\left[t_i > \frac{m^2 t}{k^2}\right] = 1 - P\left[t_i \leq \frac{m^2 t}{k^2}\right] = 1 - \frac{m^2 t/k^2 + m}{t + m}$$

Emergence of a power law

A.-L. Barabasi and R. Albert, (1999)

Make time continuous to facilitate calculations

The rate at which a vertex acquires edges is $\frac{dk_i}{dt} = \frac{k_i}{2t}$. Initially, $k_i(t_i) = m$.

Justification: the rate must be proportional to k_i and all rates must sum up to m.

$$\sum_i \frac{dk_i}{dt} = \frac{1}{2t} \sum_i k_i = \frac{2mt}{2t} = m$$

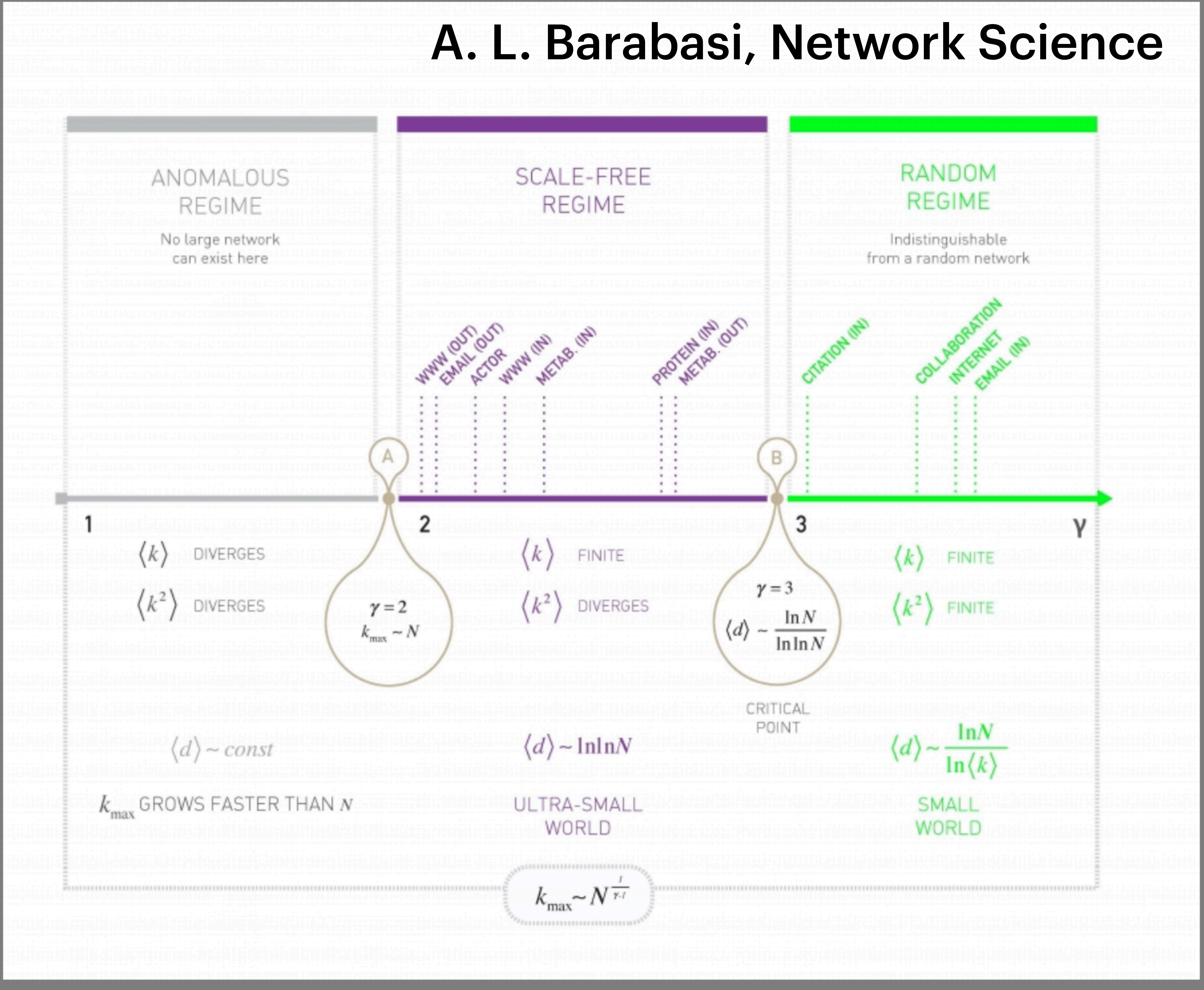
Hence, $k_i(t) = m \left(\frac{t}{t_i} \right)^{1/2}$, where t_i is the time when vertex i was added.

The probability that at time t a vertex i has $< k$ edges is $P[k_i(t) < k] = P\left[t_i > \frac{m^2 t}{k^2}\right]$.

$$P\left[t_i > \frac{m^2 t}{k^2}\right] = 1 - P\left[t_i \leq \frac{m^2 t}{k^2}\right] = 1 - \frac{m^2 t / k^2 + m}{t + m}$$

Now, find the pdf: $p(k) = \frac{\partial P[k_i(t) < k]}{\partial k} = \frac{2m^2 t}{k^3 (t + m)} \to \frac{2m^2}{k^3}$.

The γ Dependent Properties of Scale-Free Networks



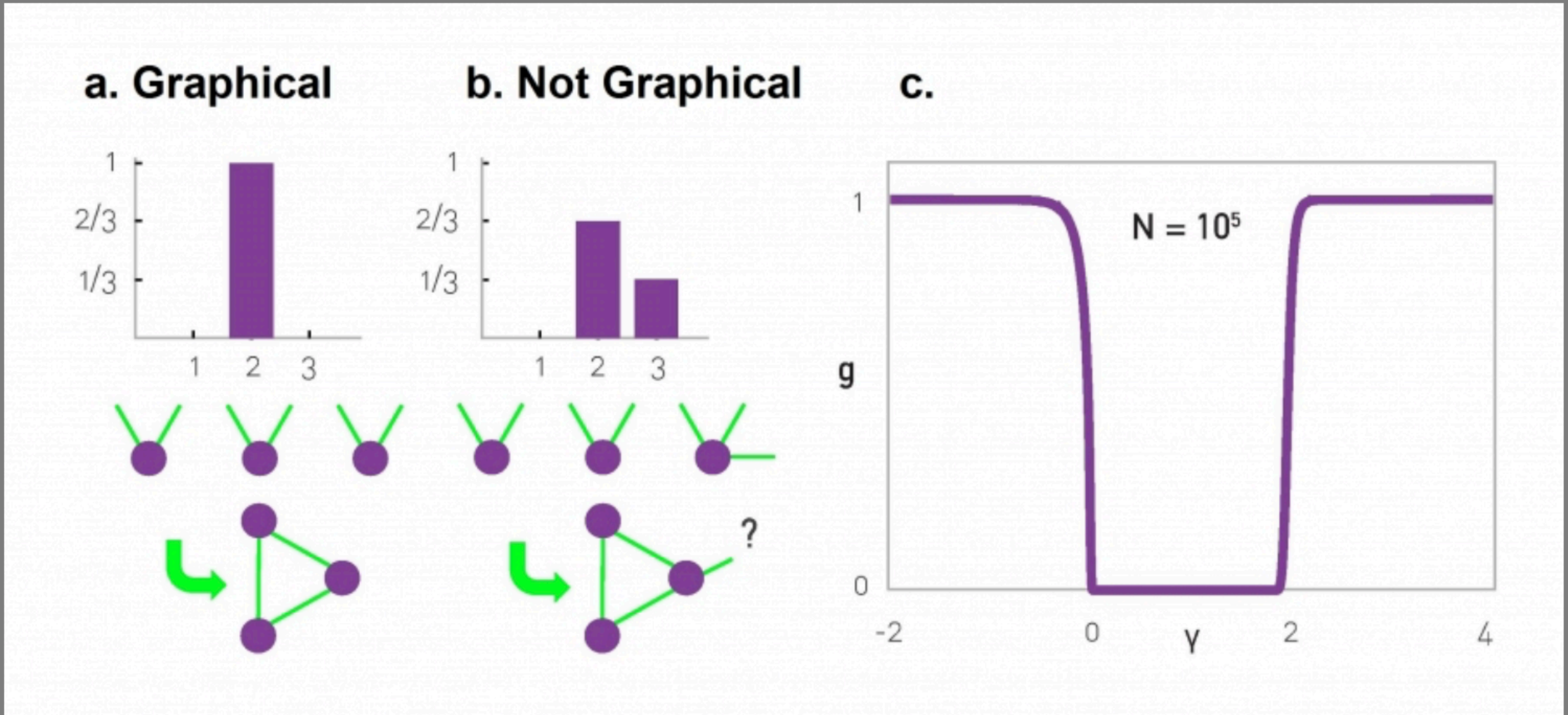


Image 4.14

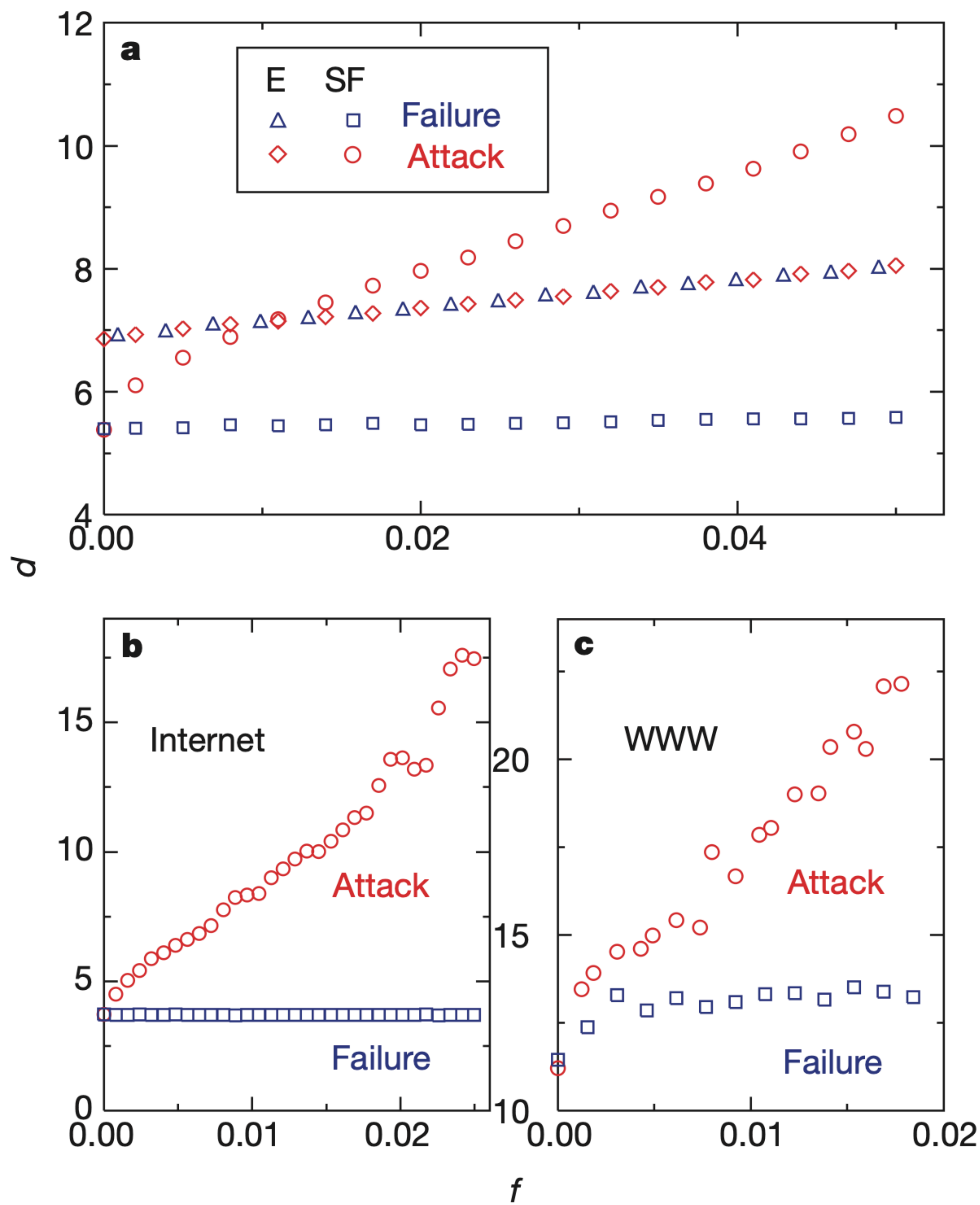
Networks With γ ‹ 2 are Not Graphical

- Degree distributions and the corresponding degree sequences for two small networks. The difference between them is in the degree of a single node. While we can build a simple network using the degree distribution (a), it is impossible to build one using (b), as one stub always remains unmatched. Hence (a) is *graphical*, while (b) is not.
- Fraction of networks, *g*, for a given *γ* that are graphical. A large number of degree sequences with degree exponent *γ* and $N = 10^5$ were generated, testing the graphicality of each network. The figure indicates that while virtually all networks with $\gamma > 2$ are graphical, it is impossible to find graphical networks in the $0 < \gamma < 2$ range. After [39].

Error and Attack tolerance

R. Albert, H. Jeong, and A.-L. Barabasi

- Two types of random networks: **Poisson** and **scale-free**
- Two types of disturbances: **random failures** and **targeted attacks**.
- **Poisson random graphs** are ***equally tolerant*** to **random failures** and **targeted attacks**.
- **Scale-free random graphs** are ***highly tolerant*** to **random failures** but ***extremely vulnerable*** to **targeted attacks**.



E = “Exponential” = “Poisson” = “Erdos-Renyi”

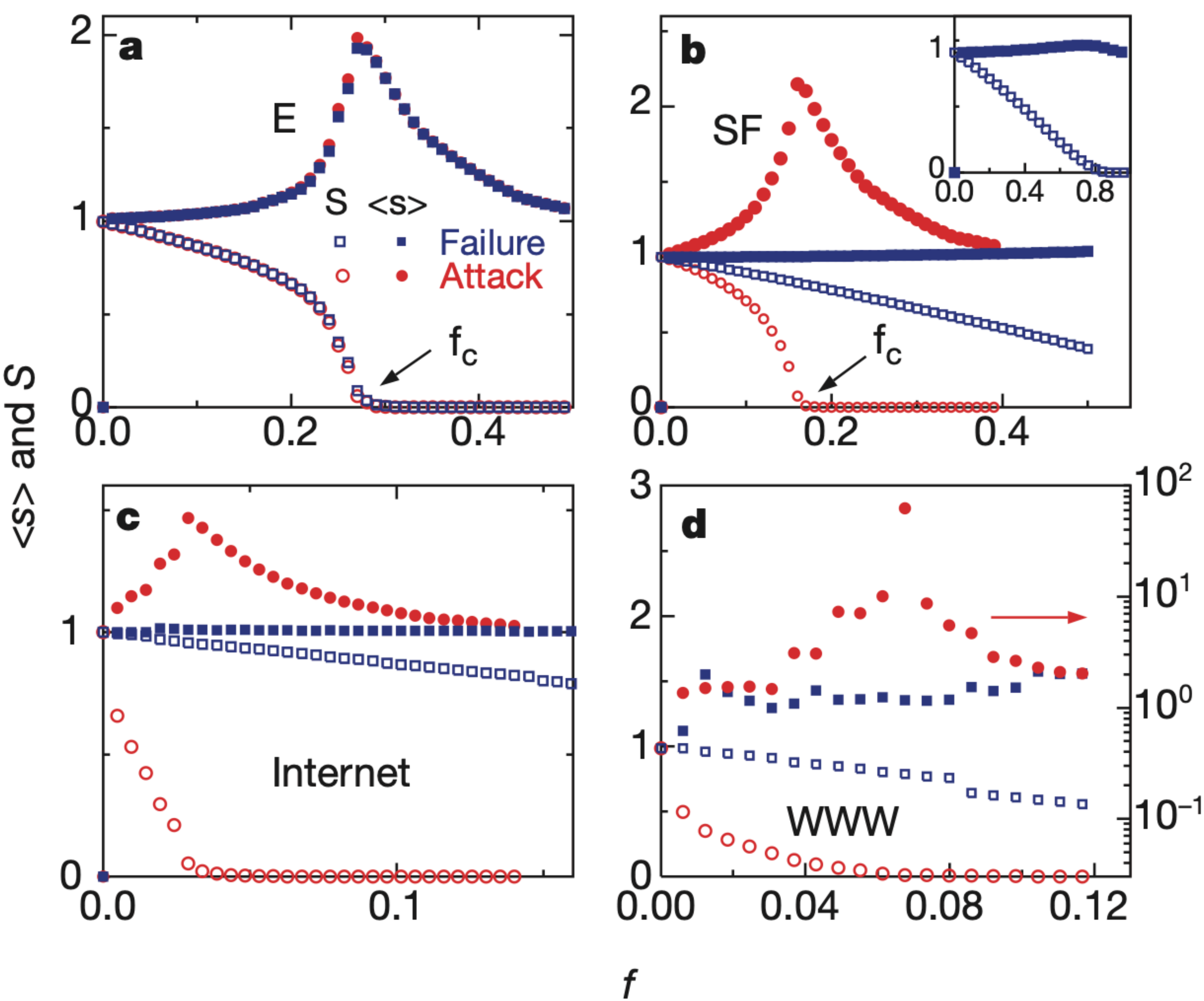
SF = “Scale-free” = “Power law”

d = the average length of the shortest path

f = fraction of removed nodes

Failure = removal of randomly picked nodes

Attack = removal of nodes of highest degree



a
E
S <s>
Failure
Attack
f_c
b
SF
f_c
c
Internet
d
WWW
<s> and S
f

A debate about scale-free networks

- **Scale-free networks are rare**. A. Broido, A. Clauset.
    - Nature Communications 10, 1017 (2019)
    - a supplement
    - ArXiv preprint (2018)(contains more details)
    - An article in Quanta Magazine
- **A. L. Barabasi's response: Love is All You Need. (2018)**
    - *Conceptual problem.* Power law is an idealized model. Real networks formed as a result of more complex processes.
    - *Methodological problem.* The criterion for a power-law networks set up by B&C is highly artificial. Even some truly scale-free networks fail to satisfy it.

Resilience to random breakdowns

Cohen, Erez, ben Avraham, Havlin, 2000

- Recall the criterion for the phase transition from no giant component to its existence

$$0 = z_2 - z_1 = \langle k^2 \rangle - 2\langle k \rangle, \quad \text{or} \quad \kappa := \frac{\langle k^2 \rangle}{\langle k \rangle} = 2$$

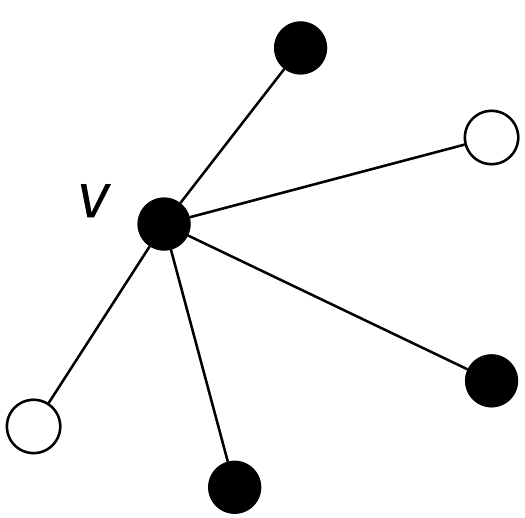
Resilience to random breakdowns

Cohen, Erez, ben Avraham, Havlin, 2000

- Recall the criterion for the phase transition from no giant component to its existence

$$0 = z_2 - z_1 = \langle k^2 \rangle - 2\langle k \rangle, \quad \text{or} \quad \kappa := \frac{\langle k^2 \rangle}{\langle k \rangle} = 2$$

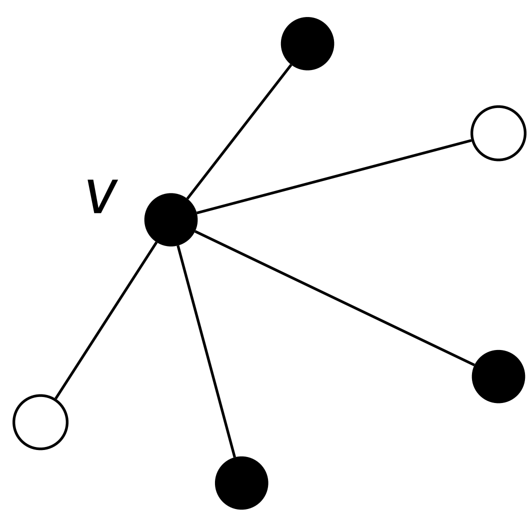
- Imagine that each node is destroyed with probability *p*.



k_0 = the original number of first neighbors of *v*

k = the number of first neighbors of *v* that are not destroyed

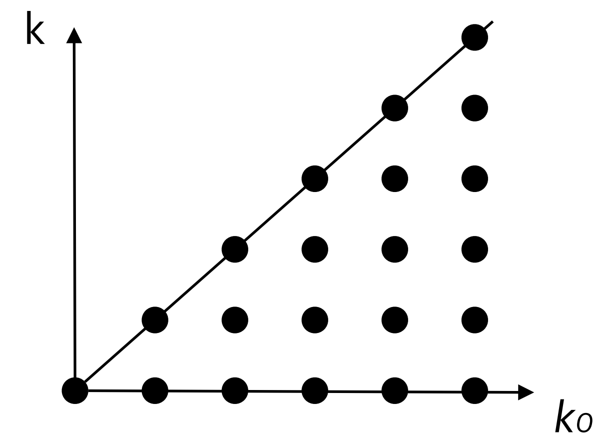
$$P'(k) = \sum_{k_0=k}^{\infty} P(k_0) \binom{k_0}{k} (1-p)^k p^{k_0-k}$$



k_0 = the original number of first neighbors of v

k = the number of first neighbors of v that are not destroyed

$$P'(k) = \sum_{k_0=k}^{\infty} P(k_0) \binom{k_0}{k} (1-p)^k p^{k_0-k}$$

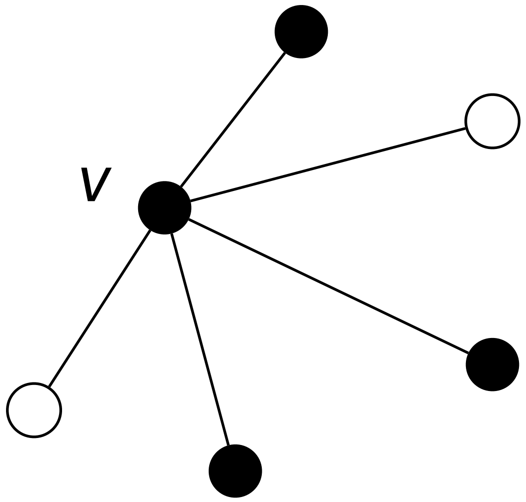


$$\langle k \rangle = \sum_{k=0}^{\infty} k \sum_{k_0=k}^{\infty} P(k_0) \binom{k_0}{k} (1-p)^k p^{k_0-k}$$

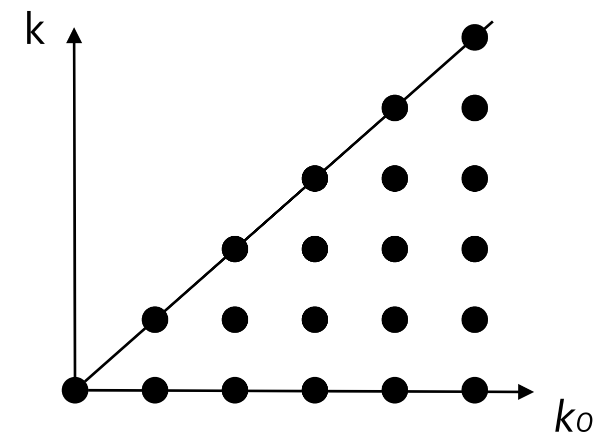
$$= \sum_{k_0=0}^{\infty} P(k_0) \sum_{k=0}^{k_0} k \binom{k_0}{k} (1-p)^k p^{k_0-k}$$

k_0 = the original number of first neighbors of v

k = the number of first neighbors of v that are not destroyed



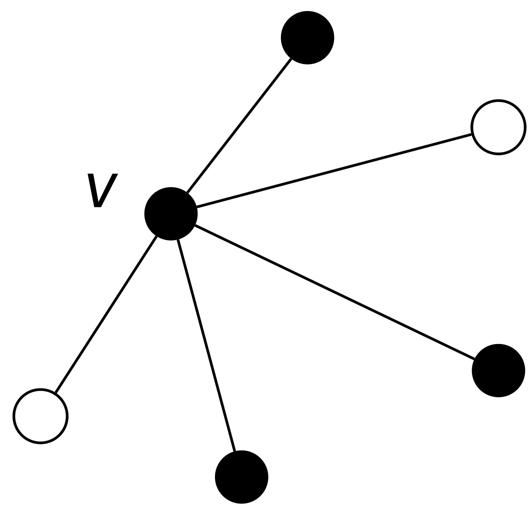
$$P'(k) = \sum_{k_0=k}^{\infty} P(k_0) \binom{k_0}{k} (1-p)^k p^{k_0-k}$$



$$\langle k \rangle = \sum_{k=0}^{\infty} k \sum_{k_0=k}^{\infty} P(k_0) \binom{k_0}{k} (1-p)^k p^{k_0-k}$$

$$= \sum_{k_0=0}^{\infty} P(k_0) \sum_{k=0}^{k_0} k \binom{k_0}{k} (1-p)^k p^{k_0-k}$$

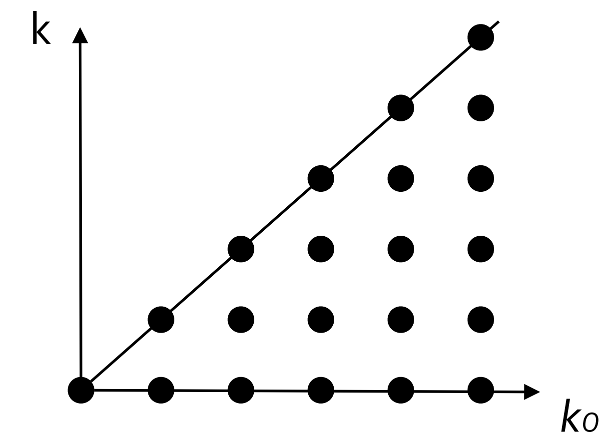
$$(x+y)^{k_0} = \sum_{k=0}^{k_0} \binom{k_0}{k} x^k y^{k_0-k}$$



k_0 = the original number of first neighbors of v

k = the number of first neighbors of v that are not destroyed

$$P'(k) = \sum_{k_0=k}^{\infty} P(k_0) \binom{k_0}{k} (1-p)^k p^{k_0-k}$$

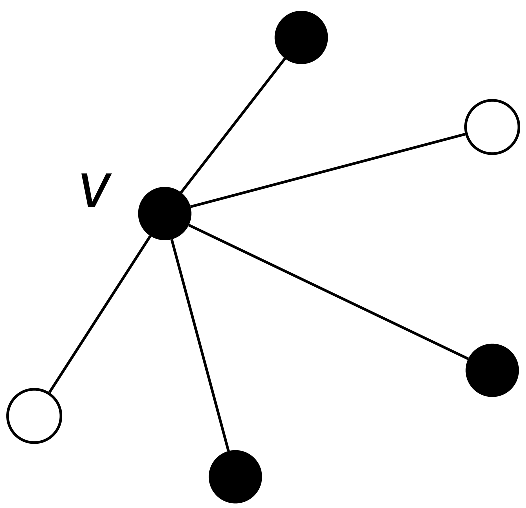


$$\langle k \rangle = \sum_{k=0}^{\infty} k \sum_{k_0=k}^{\infty} P(k_0) \binom{k_0}{k} (1-p)^k p^{k_0-k}$$

$$= \sum_{k_0=0}^{\infty} P(k_0) \boxed{\sum_{k=0}^{k_0} k \binom{k_0}{k} (1-p)^k p^{k_0-k}}$$

$$(x+y)^{k_0} = \sum_{k=0}^{k_0} \binom{k_0}{k} x^k y^{k_0-k}$$

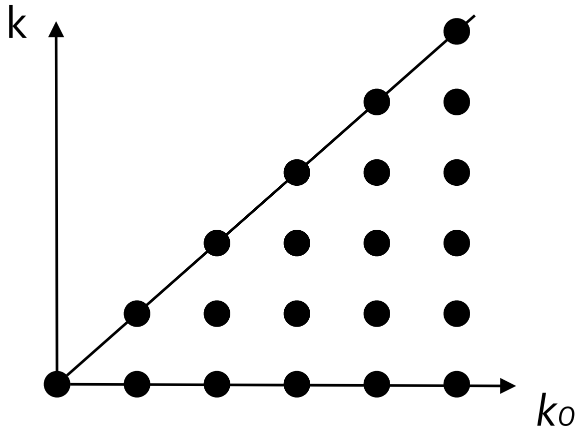
$$x\frac{d}{dx}(x+y)^{k_0} = xk_0(x+y)^{k_0-1} = \sum_{k=0}^{k_0} k \binom{k_0}{k} x^k y^{k_0-k}$$



k_0 = the original number of first neighbors of v

k = the number of first neighbors of v that are not destroyed

$$P'(k) = \sum_{k_0=k}^{\infty} P(k_0) \binom{k_0}{k} (1-p)^k p^{k_0-k}$$



$$\langle k \rangle = \sum_{k=0}^{\infty} k \sum_{k_0=k}^{\infty} P(k_0) \binom{k_0}{k} (1-p)^k p^{k_0-k}$$

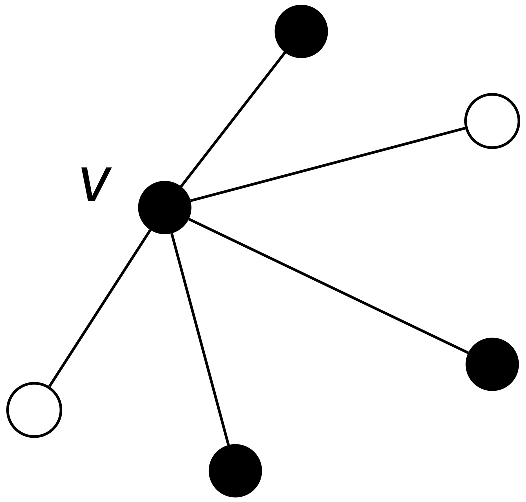
$$= \sum_{k_0=0}^{\infty} P(k_0) \sum_{k=0}^{k_0} k \binom{k_0}{k} (1-p)^k p^{k_0-k}$$

$$(x+y)^{k_0} = \sum_{k=0}^{k_0} \binom{k_0}{k} x^k y^{k_0-k}$$

$$x\frac{d}{dx}(x+y)^{k_0} = xk_0(x+y)^{k_0-1} = \sum_{k=0}^{k_0} k \binom{k_0}{k} x^k y^{k_0-k}$$

$$x \mapsto 1-p, \quad y \mapsto p$$

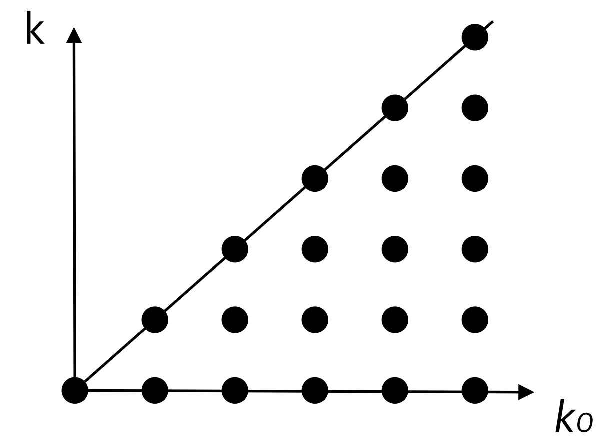
$$\sum_{k=0}^{k_0} k \binom{k_0}{k} (1-p)^k p^{k_0-k} = k_0(1-p)$$



k_0 = the original number of first neighbors of v

k = the number of first neighbors of v that are not destroyed

$$P'(k) = \sum_{k_0=k}^{\infty} P(k_0) \binom{k_0}{k} (1-p)^k p^{k_0-k}$$

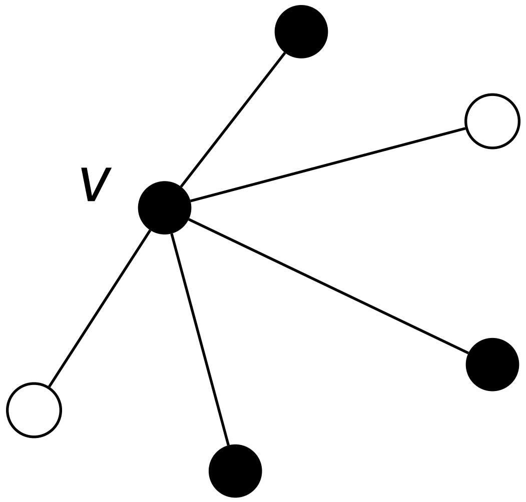


$$\langle k \rangle = \sum_{k=0}^{\infty} k \sum_{k_0=k}^{\infty} P(k_0) \binom{k_0}{k} (1-p)^k p^{k_0-k}$$

$$= \sum_{k_0=0}^{\infty} P(k_0) \sum_{k=0}^{k_0} k \binom{k_0}{k} (1-p)^k p^{k_0-k}$$

$$= \sum_{k_0=0}^{\infty} P(k_0) k_0 (1-p) = \langle k_0 \rangle (1-p)$$

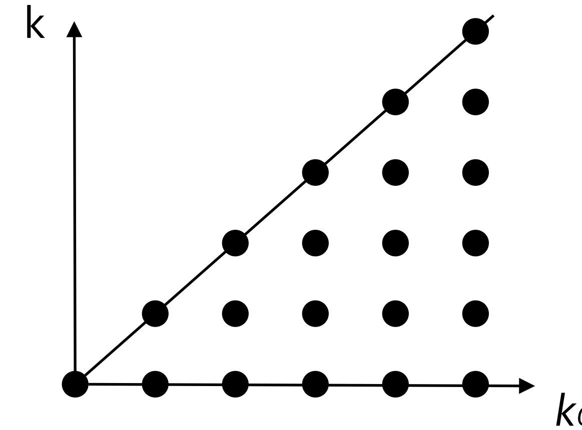
$$\langle k \rangle = \langle k_0 \rangle (1-p)$$



k_0 = the original number of first neighbors of v

k = the number of first neighbors of v that are not destroyed

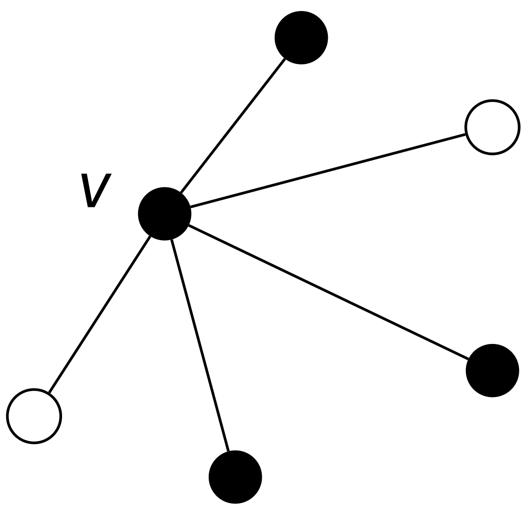
$$P'(k) = \sum_{k_0=k}^{\infty} P(k_0) \binom{k_0}{k} (1-p)^k p^{k_0-k}$$



$$\langle k^2 \rangle = \sum_{k=0}^{\infty} k^2 \sum_{k_0=k}^{\infty} P(k_0) \binom{k_0}{k} (1-p)^k p^{k_0-k}$$

$$= \sum_{k_0}^{\infty} P(k_0) \sum_{k=0}^{k_0} k^2 \binom{k_0}{k} (1-p)^k p^{k_0-k}$$

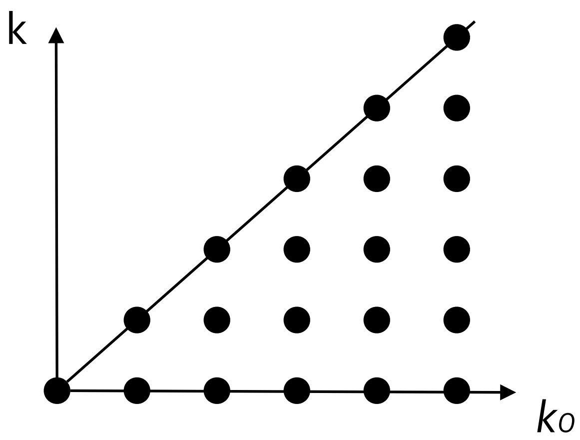
$$\langle k \rangle = \langle k_0 \rangle (1-p)$$



k_0 = the original number of first neighbors of v

k = the number of first neighbors of v that are not destroyed

$$P'(k) = \sum_{k_0=k}^{\infty} P(k_0) \binom{k_0}{k} (1-p)^k p^{k_0-k}$$



$$\langle k^2 \rangle = \sum_{k=0}^{\infty} k^2 \sum_{k_0=k}^{\infty} P(k_0) \binom{k_0}{k} (1-p)^k p^{k_0-k}$$

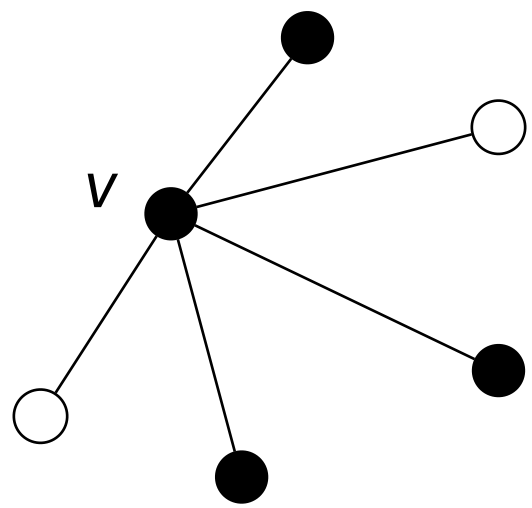
$$= \sum_{k_0}^{\infty} P(k_0) \boxed{\sum_{k=0}^{k_0} k^2 \binom{k_0}{k} (1-p)^k p^{k_0-k}}$$

$$\langle k \rangle = \langle k_0 \rangle (1-p)$$

$$\left(x \frac{d}{dx} \right)^2 (x+y)^{k_0} = x \frac{d}{dx} \left[x k_0 (x+y)^{k_0-1} \right] = x k_0 (x+y)^{k_0-1} + x^2 k_0 (k_0-1)(x+y)^{k_0-2}$$

$$= \sum_{k=0}^{k_0} k^2 \binom{k_0}{k} x^k y^{k_0-k}$$

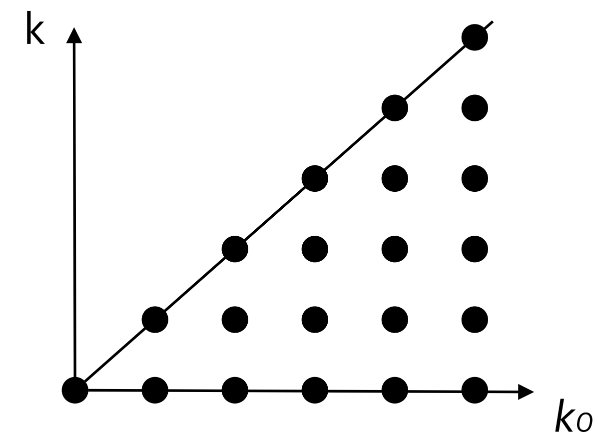
$$x \mapsto 1-p, \quad y \mapsto p$$



k_0 = the original number of first neighbors of v

k = the number of first neighbors of v that are not destroyed

$$P'(k) = \sum_{k_0=k}^{\infty} P(k_0) \binom{k_0}{k} (1-p)^k p^{k_0-k}$$



$$\langle k^2 \rangle = \sum_{k=0}^{\infty} k^2 \sum_{k_0=k}^{\infty} P(k_0) \binom{k_0}{k} (1-p)^k p^{k_0-k}$$

$$= \sum_{k_0}^{\infty} P(k_0) \boxed{\sum_{k=0}^{k_0} k^2 \binom{k_0}{k} (1-p)^k p^{k_0-k}}$$

$$\langle k \rangle = \langle k_0 \rangle (1-p)$$

$$\left(x\frac{d}{dx}\right)^2 (x+y)^{k_0} = x\frac{d}{dx}\left[x k_0 (x+y)^{k_0-1}\right] = x k_0 (x+y)^{k_0-1} + x^2 k_0 (k_0-1)(x+y)^{k_0-2}$$

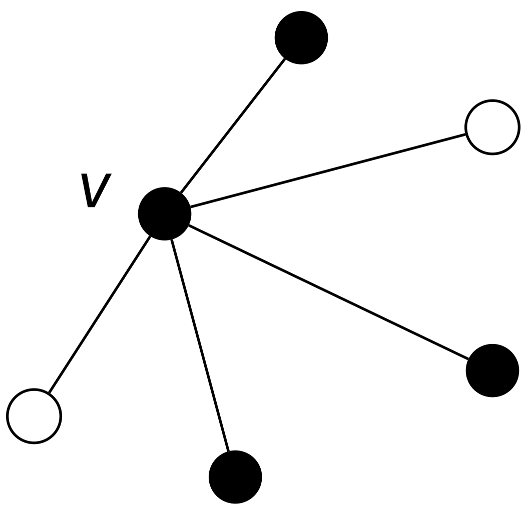
$$= \sum_{k=0}^{k_0} k^2 \binom{k_0}{k} x^k y^{k_0-k}$$

$$x \mapsto 1-p, \quad y \mapsto p$$

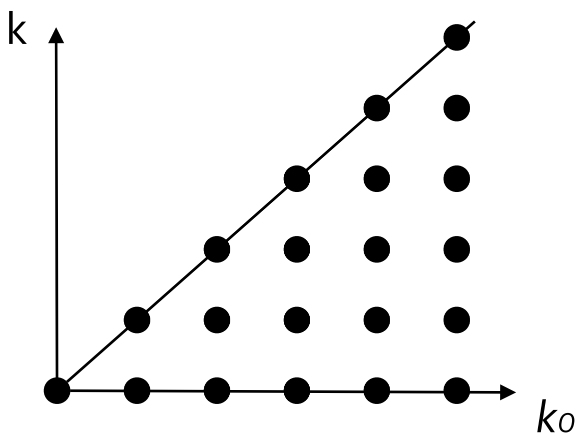
$$\sum_{k=0}^{k_0} k^2 \binom{k_0}{k} (1-p)^k p^{k_0-k} = k_0(1-p) + k_0(k_0-1)(1-p)^2 = k_0^2(1-p)^2 + k_0 p(1-p)$$

k_0 = the original number of first neighbors of v

k = the number of first neighbors of v that are not destroyed



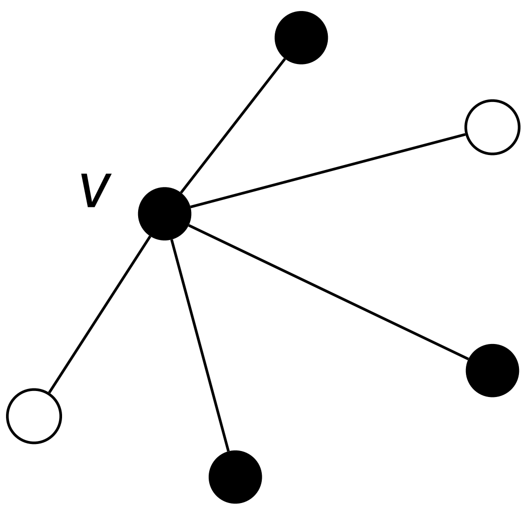
$$P'(k) = \sum_{k_0=k}^{\infty} P(k_0) \begin{pmatrix} k_0 \\ k \end{pmatrix} (1-p)^k p^{k_0-k}$$



$$\langle k^2 \rangle = \sum_{k_0=0}^{\infty} P(k_0) \left[k_0^2 (1-p)^2 + k_0 p (1-p) \right]$$

$$= \langle k_0^2 \rangle (1-p)^2 + \langle k_0 \rangle p (1-p)$$

$$\langle k \rangle = \langle k_0 \rangle (1-p)$$



k_0 = the original number of first neighbors of v

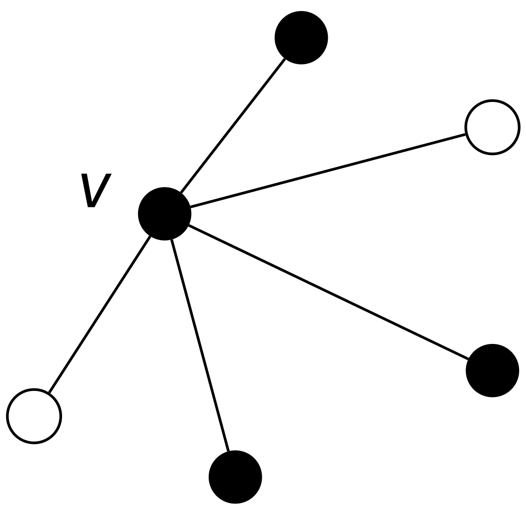
k = the number of first neighbors of v that are not destroyed

$$P'(k) = \sum_{k_0=k}^{\infty} P(k_0) \binom{k_0}{k} (1-p)^k p^{k_0-k}$$

$$\langle k \rangle = \langle k_0 \rangle (1-p)$$

$$\langle k^2 \rangle = \langle k_0^2 \rangle (1-p)^2 + \langle k_0 \rangle p(1-p)$$

$$\begin{aligned} \kappa &= \frac{\langle k^2 \rangle}{\langle k \rangle} \\ &= \frac{\langle k_0^2 \rangle (1-p) + \langle k_0 \rangle p}{\langle k_0 \rangle} \\ &= \frac{\langle k_0^2 \rangle}{\langle k_0 \rangle}(1-p) + p \end{aligned}$$



k_0 = the original number of first neighbors of v

k = the number of first neighbors of v that are not destroyed

$$P'(k) = \sum_{k_0=k}^{\infty} P(k_0) \binom{k_0}{k} (1-p)^k p^{k_0-k}$$

Critical probability of failure

$$\langle k \rangle = \langle k_0 \rangle (1-p)$$

$$\langle k^2 \rangle = \langle k_0^2 \rangle (1-p)^2 + \langle k_0 \rangle p(1-p)$$

$$\kappa = \kappa_0(1-p_c) + p_c = 2$$

$$p_c = 1 - \frac{1}{\kappa_0 - 1}$$

$$\begin{aligned} \kappa &= \frac{\langle k^2 \rangle}{\langle k \rangle} \\ &= \frac{\langle k_0^2 \rangle (1-p) + \langle k_0 \rangle p}{\langle k_0 \rangle} \\ &= \frac{\langle k_0^2 \rangle}{\langle k_0 \rangle}(1-p) + p \end{aligned}$$

Random failures in random graphs with power-law degree distribution

$$P(k_0) = ck_0^{-\alpha}, \quad k_0 = m, m+1, \ldots, K$$

Random failures in random graphs with power-law degree distribution

$$P(k_0) = ck_0^{-\alpha}, \quad k_0 = m, m+1, \ldots, K$$

$$\int_m^K ck_0^{-\alpha} dk_0 = \left[c(1-\alpha)k_0^{1-\alpha}\right]_m^K = c(1-\alpha)[K^{1-\alpha} - m^{1-\alpha}] = 1$$

Hence $\quad c \approx \dfrac{m^{\alpha-1}}{\alpha - 1}$

Random failures in random graphs with power-law degree distribution

$$P(k_0) = ck_0^{-\alpha}, \quad k_0 = m, m+1, \ldots, K \qquad c \approx \frac{m^{\alpha-1}}{\alpha - 1}$$

Estimate maximal vertex degree for a finite network with *N* nodes

$$\int_K^\infty P(k_0) dk_0 = \frac{1}{N}$$

I.e., the probability that a node has at least K first neighbors is 1/N, i.e., we expect to have at most one such a node.

Random failures in random graphs with power-law degree distribution

$$P(k_0) = c k_0^{-\alpha}, \quad k_0 = m, m+1, \ldots, K \qquad c \approx \frac{m^{\alpha-1}}{\alpha - 1}$$

Estimate maximal vertex degree for a finite network with *N* nodes

$$\int_K^\infty P(k_0) dk_0 = \frac{1}{N}$$

I.e., the probability that a node has at least K first neighbors is 1/N, i.e., we expect to have at most one such a node.

$$\int_K^\infty P(k_0) dk_0 = c(\alpha - 1) K^{1-\alpha} = \left(\frac{m}{K}\right)^{\alpha-1} = \frac{1}{N}$$

$$K = m N^{1/(\alpha-1)}$$

Hence $K \to \infty$ as $N \to \infty$.

Random failures in random graphs with power-law degree distribution

$$P(k_0) = ck_0^{-\alpha}, \quad k_0 = m, m+1, \ldots, K$$

$$p_c = 1 - \frac{1}{\kappa_0 - 1}$$

$$\kappa_0 = \frac{\langle k_0^2 \rangle}{\langle k_0 \rangle} = \frac{(3-\alpha)}{(2-\alpha)} \frac{[K^{3-\alpha} - m^{3-\alpha}]}{[K^{2-\alpha} - m^{2-\alpha}]}$$

Random failures in random graphs with power-law degree distribution

$$P(k_0) = ck_0^{-\alpha}, \quad k_0 = m, m+1, \ldots, K$$

$$p_c = 1 - \frac{1}{\kappa_0 - 1}$$

$$\kappa_0 = \frac{\langle k_0^2 \rangle}{\langle k_0 \rangle} = \frac{(3-\alpha)}{(2-\alpha)} \frac{[K^{3-\alpha} - m^{3-\alpha}]}{[K^{2-\alpha} - m^{2-\alpha}]}$$

$$K \to \infty$$

$$\kappa_0 \approx \left| \frac{3-\alpha}{2-\alpha} \right| \begin{cases} m, & \alpha > 3 \\ m^{\alpha-2} K^{3-\alpha}, & 2 < \alpha < 3 \\ K, & 1 < \alpha < 2. \end{cases}$$

Random failures in random graphs with power-law degree distribution

$$P(k_0) = c k_0^{-\alpha}, \quad k_0 = m, m+1, \ldots, K$$

$$p_c = 1 - \frac{1}{\kappa_0 - 1}$$

$$\kappa_0 = \frac{\langle k_0^2 \rangle}{\langle k_0 \rangle} = \frac{(3-\alpha)}{(2-\alpha)} \frac{[K^{3-\alpha} - m^{3-\alpha}]}{[K^{2-\alpha} - m^{2-\alpha}]}$$

$$K \to \infty$$

$$\kappa_0 \approx \left| \frac{3-\alpha}{2-\alpha} \right| \begin{cases} m, & \alpha > 3 \\ m^{\alpha-2} K^{3-\alpha}, & 2 < \alpha < 3 \\ K, & 1 < \alpha < 2. \end{cases}$$

$$p_c = 1 - \frac{\alpha - 2}{(\alpha - 3)m - (\alpha - 2)} < 1, \quad \alpha > 3$$

$$p_c = 1 - \frac{\alpha - 2}{(3-\alpha) m^{\alpha-2} K^{3-\alpha} - (\alpha - 2)} \to 1 \quad \text{as} \quad K \to \infty, \quad 2 < \alpha < 3$$

Random failures in random graphs with power-law degree distribution

$$P(k_0) = ck_0^{-\alpha}, \quad k_0 = m, m+1, \ldots, K$$

$$p_c = 1 - \frac{1}{\kappa_0 - 1}$$

$$\kappa_0 = \frac{\langle k_0^2 \rangle}{\langle k_0 \rangle} = \frac{(3-\alpha)}{(2-\alpha)} \frac{[K^{3-\alpha} - m^{3-\alpha}]}{[K^{2-\alpha} - m^{2-\alpha}]}$$

Most real-world networks:

$$2 < \alpha < 3$$

$$K \to \infty$$

$$\kappa_0 \approx \left| \frac{3-\alpha}{2-\alpha} \right| \begin{cases} m, & \alpha > 3 \\ m^{\alpha-2} K^{3-\alpha}, & 2 < \alpha < 3 \\ K, & 1 < \alpha < 2. \end{cases}$$

Hence $\kappa_0 \to \infty$ as $K \to \infty$

$$p_c = 1 - \frac{\alpha - 2}{(\alpha - 3)m - (\alpha - 2)} < 1, \quad \alpha > 3$$

$$p_c = 1 - \frac{\alpha - 2}{(3-\alpha) m^{\alpha-2} K^{3-\alpha} - (\alpha - 2)} \to 1 \quad \text{as} \quad K \to \infty, \quad 2 < \alpha < 3$$

Random failures in random graphs with power-law degree distribution

$$P(k_0) = ck_0^{-\alpha}, \quad k_0 = m, m+1, \ldots, K$$

$$p_c = 1 - \frac{1}{\kappa_0 - 1}$$

Most real-world networks: $2 < \alpha < 3$

Hence $\kappa_0 \to \infty$ as $K \to \infty$ which is caused by $N \to \infty$.

Ratio of fractions in giant component

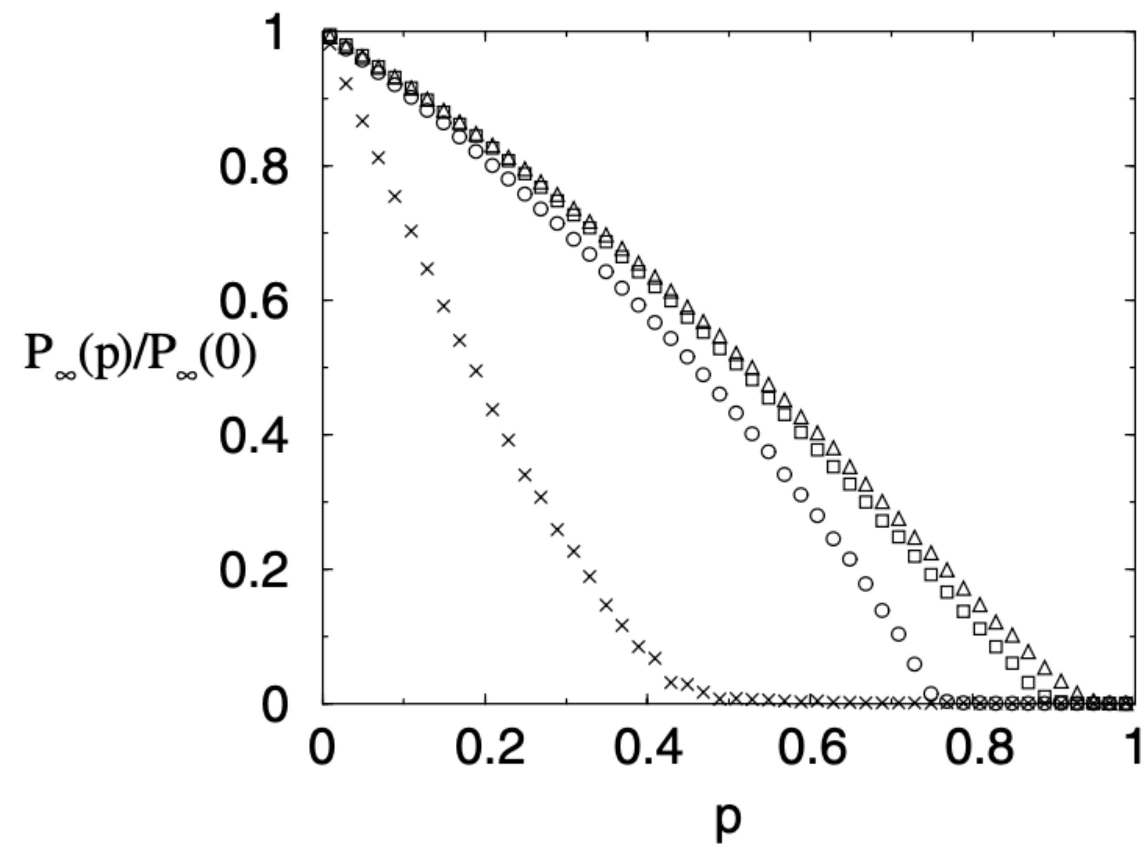


FIG. 1. Percolation transition for networks with power-law connectivity distribution. Plotted is the fraction of nodes that remain in the spanning cluster after breakdown of a fraction p of all nodes, $P_\infty(p)/P_\infty(0)$, as a function of p, for $\alpha = 3.5$ (crosses) and $\alpha = 2.5$ (other symbols), as obtained from computer simulations of up to $N = 10^6$. In the former case, it can be seen that for $p > p_c \approx 0.5$ the spanning cluster disintegrates and the network becomes fragmented. However, for $\alpha = 2.5$ (the case of the Internet), the spanning cluster persists up to nearly 100% breakdown. The different curves for $K = 25$ (circles), 100 (squares), and 400 (triangles) illustrate the finite size effect: the transition exists only for finite networks, while the critical threshold p_c approaches 100% as the networks grow in size.

Percolation on random graphs

Callaway, Newman, Strogatz, Watts (2000)

- The failure probability is allowed to depend on degree: the probability that a vertex of degree k survives (is occupied) is q_k.

Percolation on random graphs

Callaway, Newman, Strogatz, Watts (2000)

- The failure probability is allowed to depend on degree: the probability that a vertex of degree k survives (is occupied) is q_k.
- Note that if $q_k = q$ for all k, then $q = 1 - p$ from Cohen et al.

Percolation on random graphs

Callaway, Newman, Strogatz, Watts (2000)

- The failure probability is allowed to depend on degree: the probability that a vertex of degree k survives (is occupied) is q_k.
- Note that if $q_k = q$ for all k, then $q = 1 - p$ from Cohen et al.
- Method of generating functions is used. Result from Cohen et al. is rederived and refined.

Percolation on random graphs

Callaway, Newman, Strogatz, Watts (2000)

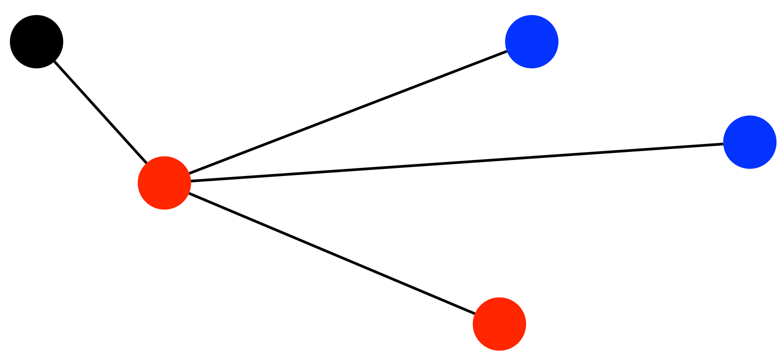
- The failure probability is allowed to depend on degree: the probability that a vertex of degree k survives (is occupied) is q_k.
- Note that if $q_k = q$ for all k, then $q = 1 - p$ from Cohen et al.
- Method of generating functions is used. Result from Cohen et al. is rederived and refined.
- Disappearance of the giant component is shown for targeted attack removing highest degree nodes.

Spread of epidemic disease on network

M. Newman (2002)

SIR model: Susceptible → Infecting → Removed
(L. Reed, W. H. Frost, 1920s, unpublished)

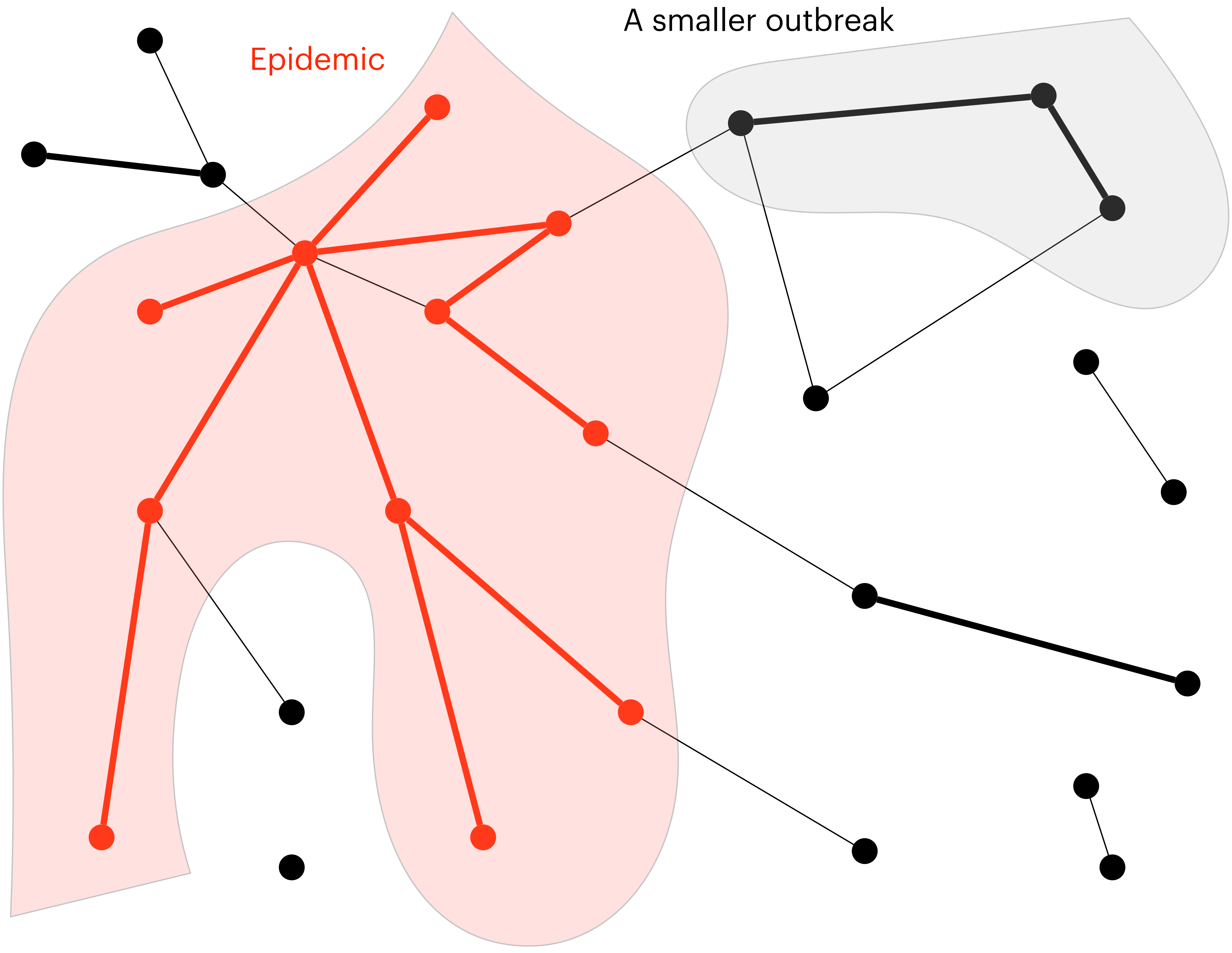
$$\frac{ds}{dt} = -\beta is, \quad \frac{di}{dt} = \beta is - \gamma i, \quad \frac{dr}{dt} = \gamma i \qquad s + i + r = 1$$



r = rate of disease-causing contacts
τ = duration of being infecting

$$T = 1 - e^{-r\tau} = \text{transmission rate}$$

Grassberger (1983): Mapping on the bond percolation problem:
each edge is transmitting with probability *T*.


A smaller outbreak
Epidemic

$$G_0(x) = \sum_{k=0}^{\infty} p_k x^k = \text{generating function for degree distribution}$$

$$G_1(x) = \sum_{k=0}^{\infty} q_k x^k = \sum_{k=0}^{\infty} \frac{(k+1)p_{k+1}}{\sum_{j=0}^{\infty} j p_j} x_k = \frac{G_0'(x)}{z}$$

$$= \text{generating function for the excess degree distribution}$$

$$G_0(x;T) = \sum_{m=0}^{\infty} \sum_{k=m}^{\infty} p_k \binom{k}{m} T^m (1-T)^{k-m} x^m$$

$$= \sum_{k=0}^{\infty} p_k \sum_{m=0}^{k} \binom{k}{m} (xT)^m (1-T)^{k-m}$$

$$= \sum_{k=0}^{\infty} p_k (1 - T + xT)^k = G_0(1 + (x-1)T)$$

$$= \text{generating function for distribution of transmitting edges adjacent to a node}$$

$$G_1(x;T) = G_1(1 + (x-1)T)$$

$$= \text{generating function for distribution of transmitting edges adjacent to a node}$$

arrived at by a randomly chosen edge

$$H_1(x;T) = xG_1(H_1(x;T);T) = \text{generating function for the size of transmitting cluster reached from a randomly chosen edge}$$

$$H_0(x;T) = xG_0(H_1(x;T);T) = \text{generating function for the size of transmitting cluster reached from a randomly chosen vertex}$$

$$P_s(T) = \frac{1}{s!} \left. \frac{d^s H_0}{dx^s} \right|_{x=0} = \frac{1}{2\pi i} \oint \frac{H_0(\zeta;T)}{\zeta^{s+1}} d\zeta$$

$$= \text{probability that transmitting cluster has size } s$$

Recipe for finding the distribution of cluster sizes numerically

$$\langle s \rangle = H_0'(1;T) = 1 + G_0'(1;T) H_1'(1;T)$$

$$= \text{average outbreak size}$$

$$H_1'(1;T) = 1 + G_1'(1;T) H_1'(1;T) = \frac{1}{1 - G_1'(1;T)}$$

$$\langle s \rangle = H_0'(1;T) = 1 + \frac{G_0'(1;T)}{1 - G_1'(1;T)} = 1 + \frac{T G_0'(1)}{1 - T G_1'(1)}$$

$$= \text{average outbreak size}$$

If T is below the epidemic threshold

$$T_c = \frac{1}{G_1'(1)} = \frac{G_0'(1)}{G_0''(1)} = \frac{\sum_{k=1}^{\infty} k p_k}{\sum_{k=1}^{\infty} k(k-1) p_k}$$

Critical transmission:

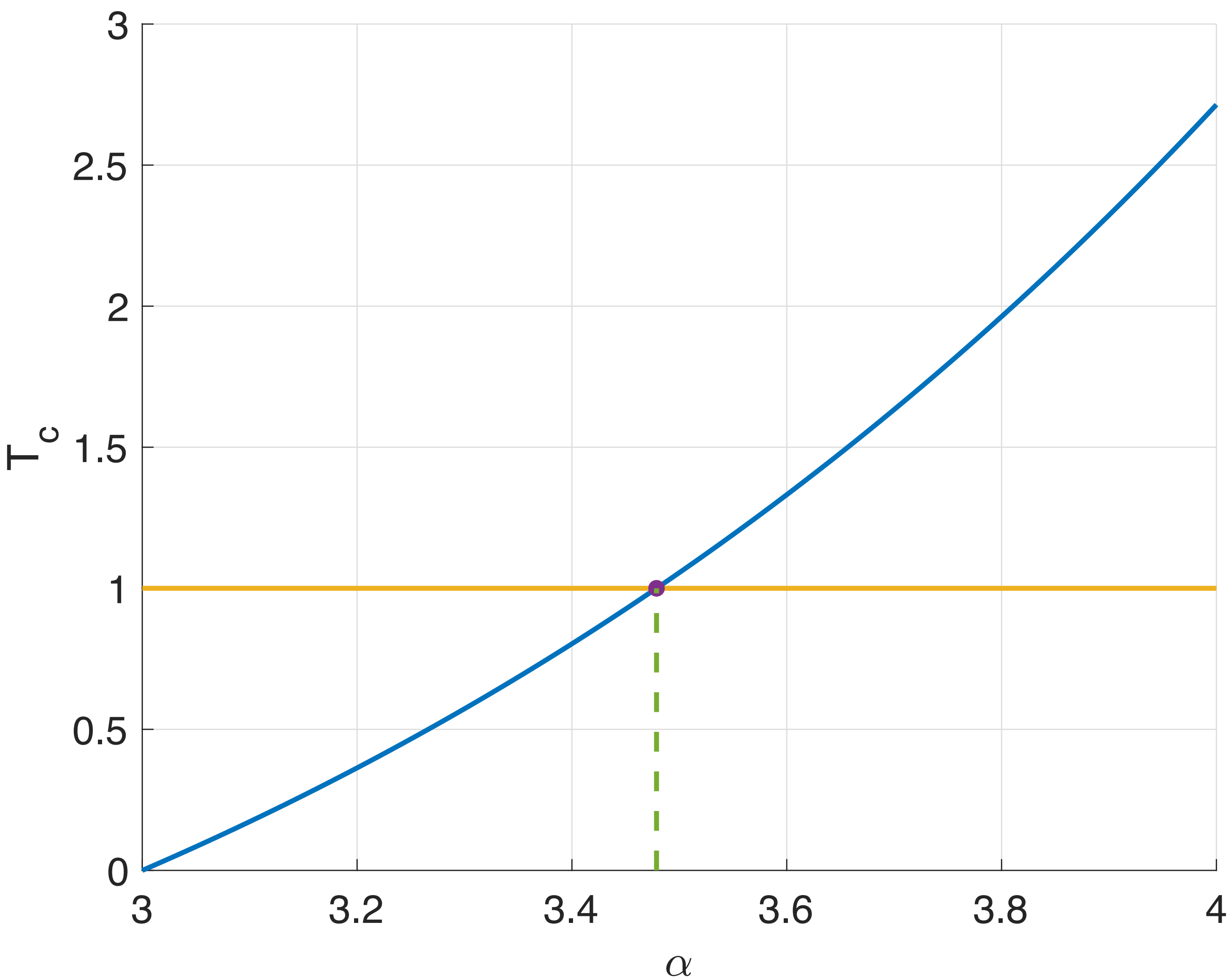
- for $T > T_c$ we have a giant component connected by transmitting edges (an epidemic);
- for $T < T_c$ all components are small (no epidemic).

Critical transmission probability for power law degree distribution

$$p_k = \frac{k^{-\alpha}}{\zeta(\alpha)}, \quad \zeta(\alpha) = \sum_{k=1}^{\infty} k^{-\alpha} = \text{Riemann zeta function}$$

$$T_c = \frac{\sum_{k=1}^{\infty} k p_k}{\sum_{k=1}^{\infty} k(k-1) p_k} = \frac{\zeta(\alpha-1)}{\zeta(\alpha-2)-\zeta(\alpha-1)}$$

- If $\alpha \leq 3$ then $T_c = 0$, hence, there is always an epidemic.
- If $3 < \alpha < \alpha_c \approx 3.4788$, then $0 < T_c < 1$, hence, there is epidemic threshold.
- If $\alpha \geq \alpha_c \approx 3.4788$, no epidemic can occur unless $T = 1$.



- For $T>T_c$, we redefine H_0 as the generating function for outbreaks other than the giant component.
- Note: we cannot use H_0 for the giant cluster as the “no loop” assumption no longer holds.

$$H_0(1;T) = \sum_{s=1}^{\infty} P_s(T) = 1 - S(T), \quad S(T) = \text{fraction in the giant component}$$

$$H_0(1;T) = G_0(u;T), \quad \text{where} \quad u = H_1(1;T)$$

$$H_1(1;T) = G_1(H_1(1;T);T), \quad \text{hence we get an equation for } u: \quad u = G_1(u;T)$$

The quantity *u* is the probability that the vertex at the end of a randomly chosen edge remains uninfected during an epidemic ⍰i.e., that it belongs to one of the finite components⍰.

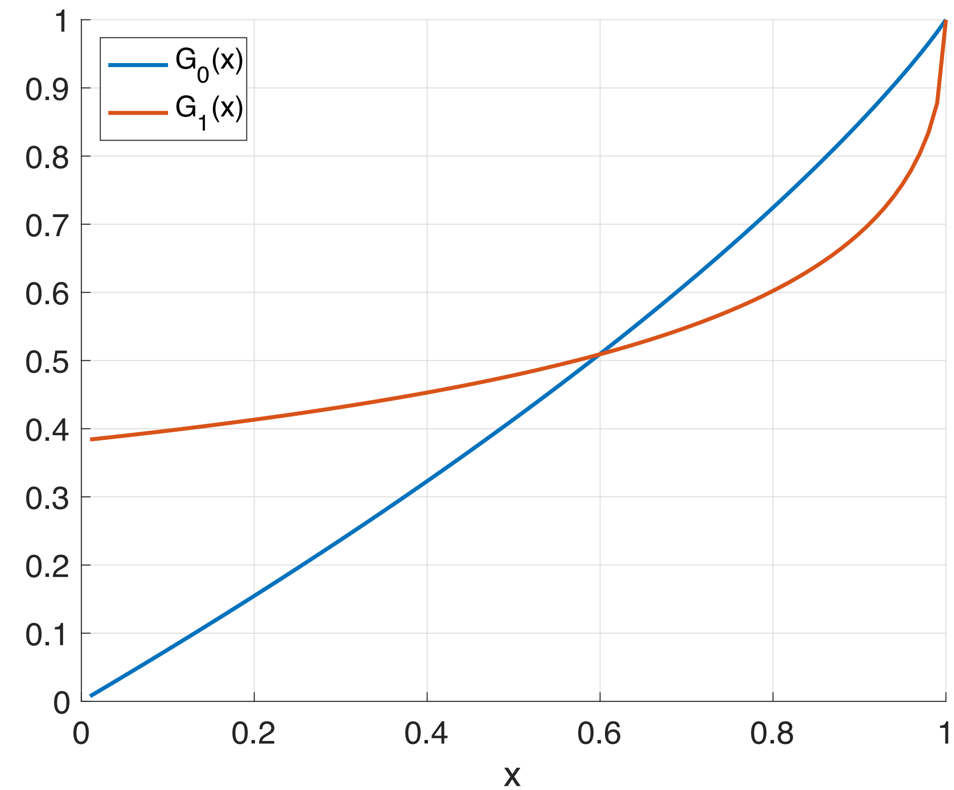
G_0, G_1, u, and S for power law degree distribution

$$Li_\alpha(x) = \sum_{k=1}^{\infty} \frac{x^k}{k^\alpha} = \text{polylogarithm}$$

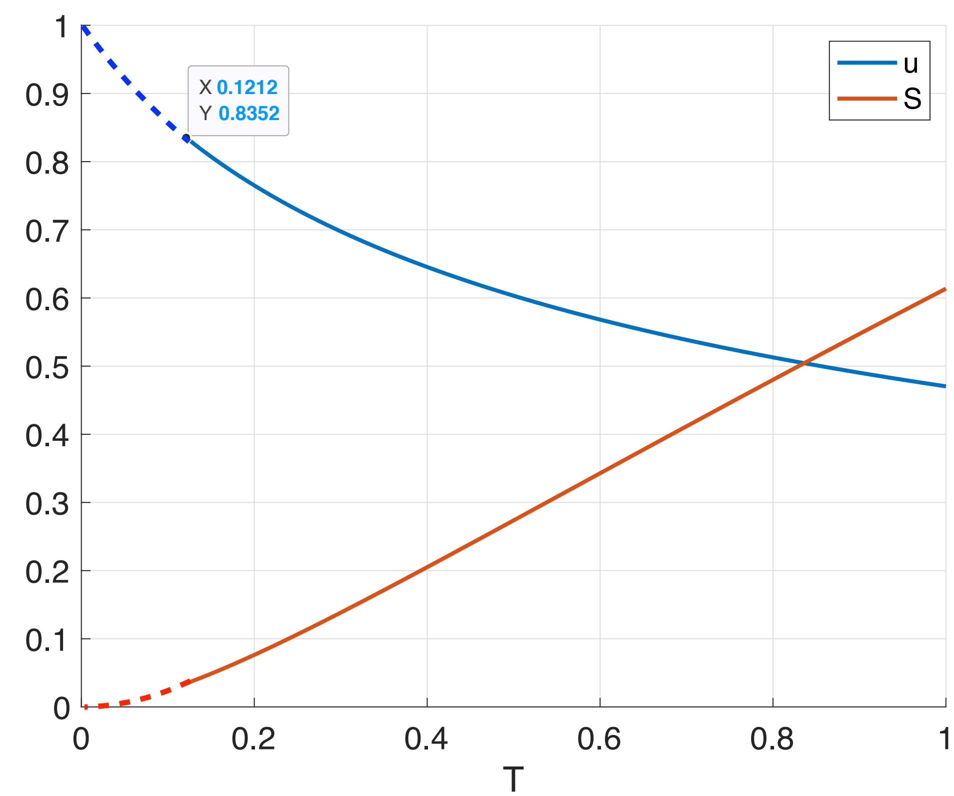
```
function SIR()
close all
fsz = 16;
% power law degree distribution p_k = k^(-a)/zeta(a)
a = 2.5;
G0 = @(x)polylog(a,x)/polylog(a,1);
G1 = @(x)polylog(a-1,x)./(x*polylog(a-1,1));
x=linspace(0,1,100);
%
figure(1);
hold on;
grid;
plot(x,G0(x),'Linewidth',2)
plot(x,G1(x),'Linewidth',2)
legend('G_0(x)','G_1(x)');
xlabel('x','Fontsize',fsz);
set(gca,'Fontsize',fsz)
%
% critical transissibility = 0, hence, there is always an epidemic
nt = 100;
t = linspace(0,1,nt); % transimissibility
u = zeros(nt,1);
S = zeros(nt,1);
for i = 1 : nt
    T = t(i);
    u(i) = fzero(@(x)G1(1-T+T*x)-x,0.3);
    S(i) = 1 - G0(1-T+T*u(i));
end
figure(2);
hold on;
grid;
plot(t,u,'Linewidth',2)
plot(t,S,'Linewidth',2)
legend('u','S');
xlabel('T','Fontsize',fsz);
set(gca,'Fontsize',fsz)
end
```

$$G_0 = \frac{Li_\alpha(x)}{Li_\alpha(1)}$$

$$G_1 = \frac{G_0'(x)}{G_0'(1)} = \frac{Li_{\alpha-1}(x)}{xLi_{\alpha-1}(1)}$$



u = probability that a vertex at the end of a random edge stays uninfected duding the epidemic;
S = fraction in the giant component.



The quantity *u* is the probability that the vertex at the end of a randomly chosen edge remains uninfected during an epidemic ?i.e., that it belongs to one of the finite components?.

The probability that a vertex does not become infected via one of its edges is

v? = 1 - ?*T* +?*Tu,*

which is the sum of the probability (1-?*T)* that the edge is non-transmitting, and the probability *Tu* that it is transmitting but connects to an uninfected vertex. The total probability of being uninfected if a vertex has degree *k* is v^k, and the probability of having degree *k* given that a vertex is uninfected is

$$\frac{p_p v^k}{\sum_{k=0}^{\infty} p_k v^k}. \quad \text{This distribution is generated by} \quad \frac{G_0(vx)}{G_0(v)}.$$

The average vertex degree outside the giant component:

$$z_{\notin Giant} = \frac{d}{dx}\frac{G_0(vx)}{G_0(v)}\bigg|_{x=1} = \frac{vG_0'(v)}{G_0(v)} = \frac{vzG_1(v)}{G_0(v)}$$

Recall that $G_1(x;T) = G_1(1-T+xT)$. Hence $G_1(v) = G_1(u;T) = u$.

Also recall that $G_0(v) = G_0(1-T+Tu) = G_0(u;T) = 1-S(T)$.

Hence $z_{\notin Giant} = \dfrac{vzG_1(v)}{G_0(v)} = \dfrac{(1-T+Tu)u}{1-S(T)}z$

The average vertex degree inside the giant component:

$$z_{\in Giant} = \frac{d}{dx}\frac{G_0(x)-G_0(vx)}{G_0(1)-G_0(v)}\bigg|_{x=1} = \frac{vG_0'(v)}{G_0(v)} = \frac{1-vG_1(v)}{1-G_0(v)}z = \frac{1-u(1-T+Tu)}{S}z$$

Mean degrees for the power law degree distribution

$$p_k = \frac{k^{-\alpha}}{\zeta(\alpha)}, \quad \zeta(\alpha) = \sum_{k=1}^{\infty} k^{-\alpha} = \text{Riemann zeta function}$$

$$z = \frac{Li_{\alpha-1}(1)}{Li_{\alpha}(1)} = \text{the mean degree}$$

$$\alpha = 2.5$$