Complex networks: an introduction

Alain Barrat

CPT, Marseille, France

ISI, Turin, Italy

http://www.cpt.univ-mrs.fr/~barrat
http://cxnets.googlepages.com

CENTRE DE PHYSIQUE THÉORIQUE DE MARSEILLE

CNRS CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Cx-Nets

INSTITUTE FOR SCIENTIFIC INTERCHANGE FOUNDATION

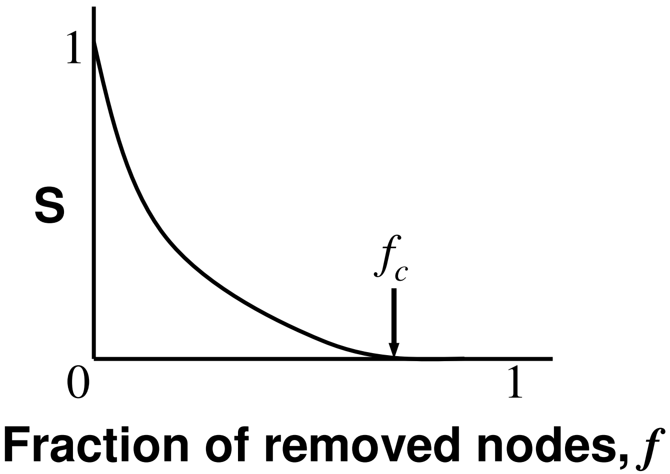
Plan of the lecture

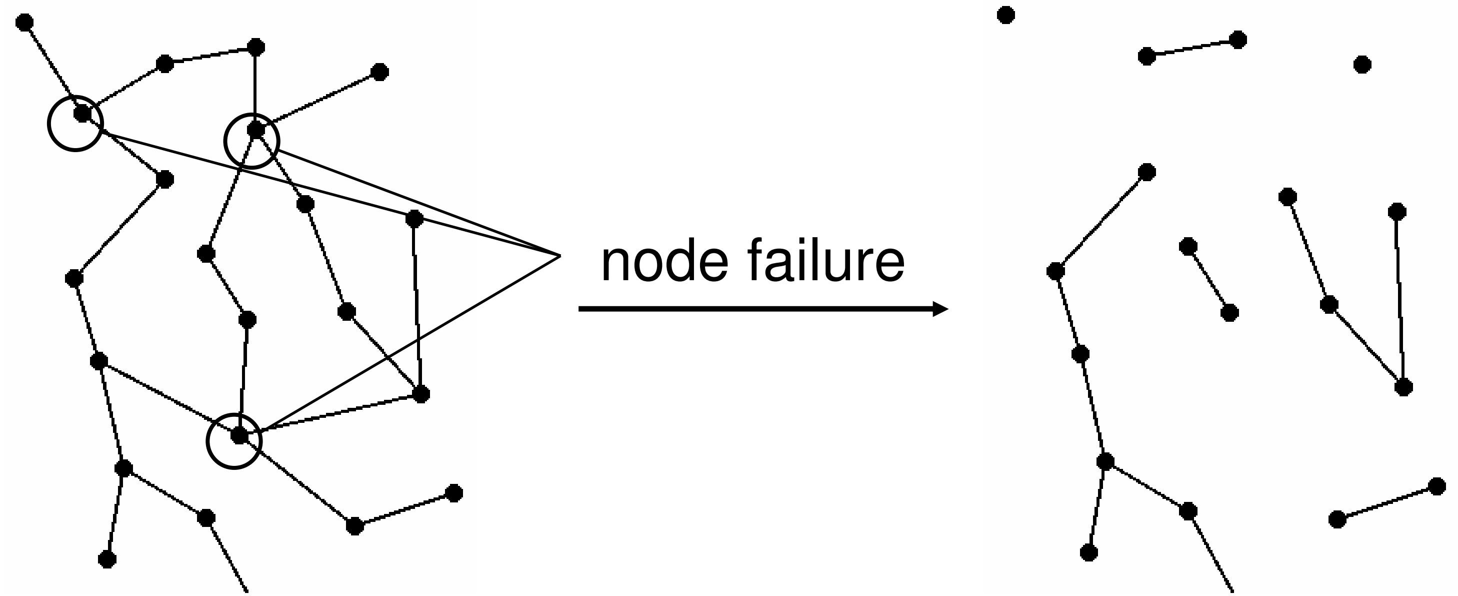
I. INTRODUCTION
   I. Networks: definitions, statistical characterization
   II. Real world networks
II. DYNAMICAL PROCESSES
   I. Resilience, vulnerability
   II. Random walks
   III. Epidemic processes
   IV. (Social phenomena)
   V. Some perspectives

Robustness

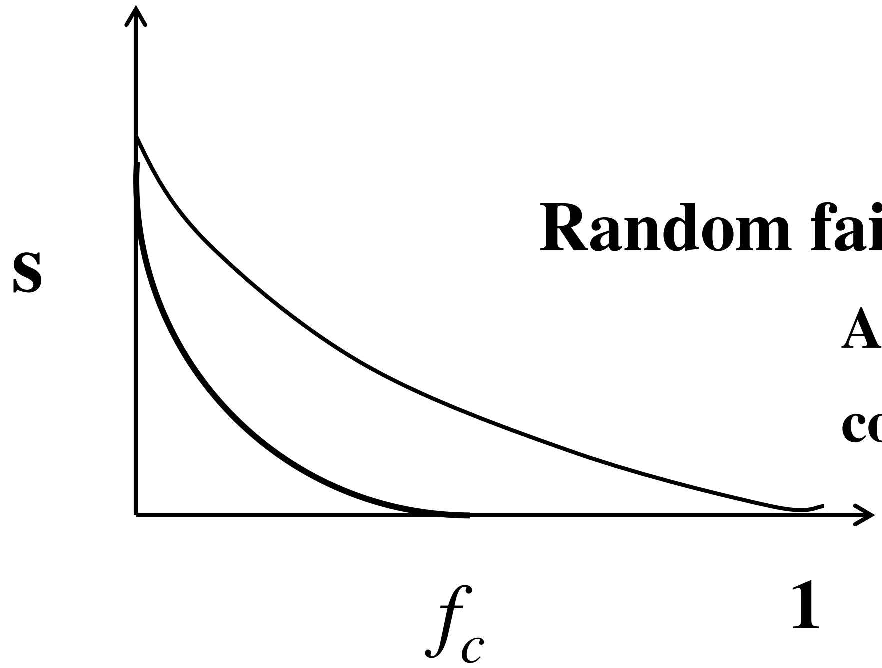
Complex systems maintain their basic functions
even under errors and failures
(cell → mutations; Internet → router breakdowns)

S: fraction of giant component





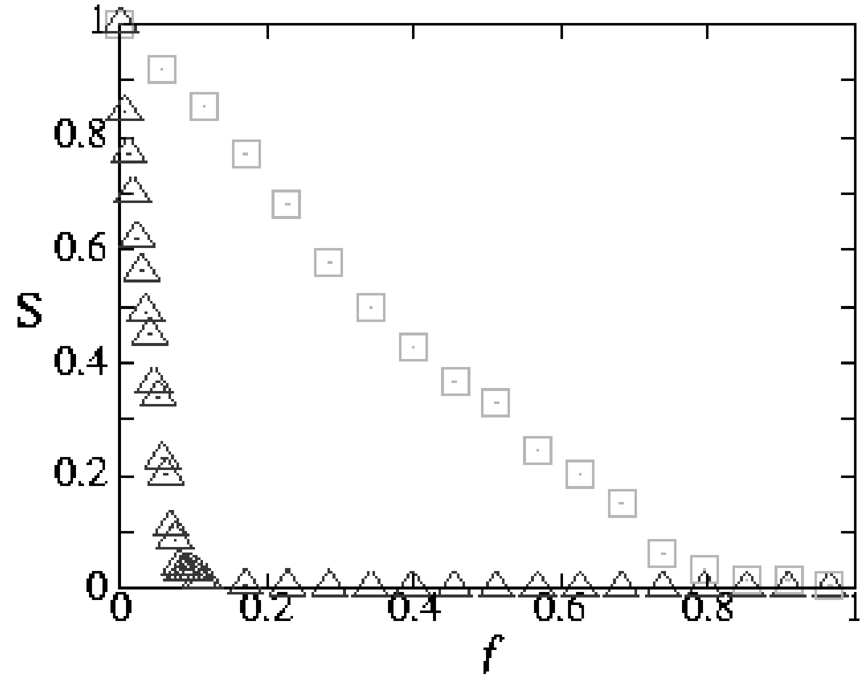
Case of Scale-free Networks



Random failure f_c =1 $(2 < \gamma \leq 3)$

Attack =progressive failure of the most connected nodes f_c <1

Internet maps

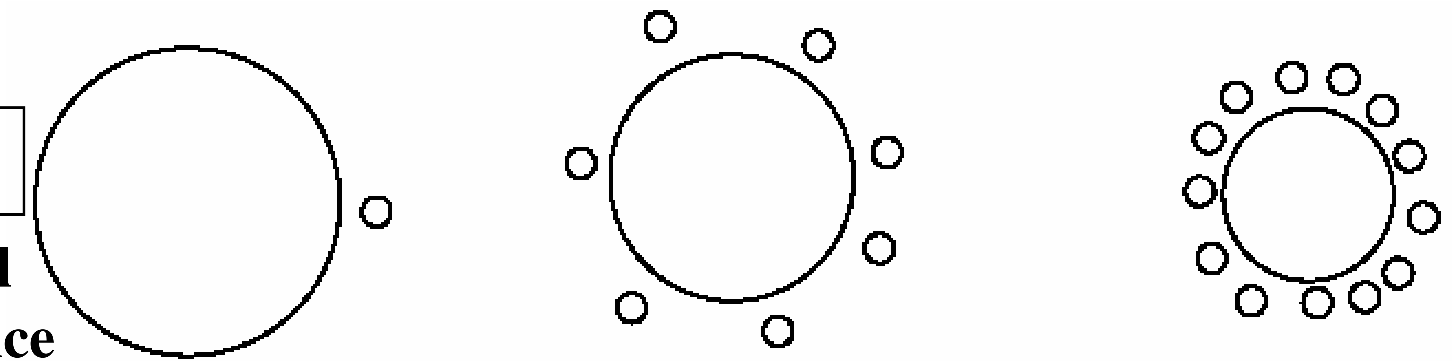


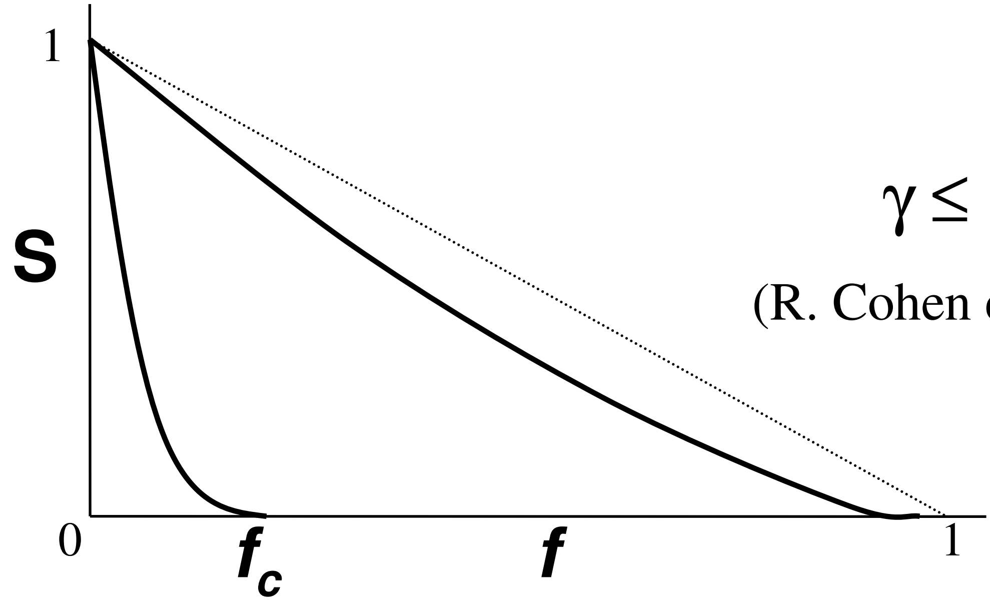
R. Albert, H. Jeong, A.L. Barabasi, Nature **406** 378 (2000)

Failures vs. attacks

Failures

Topological error tolerance

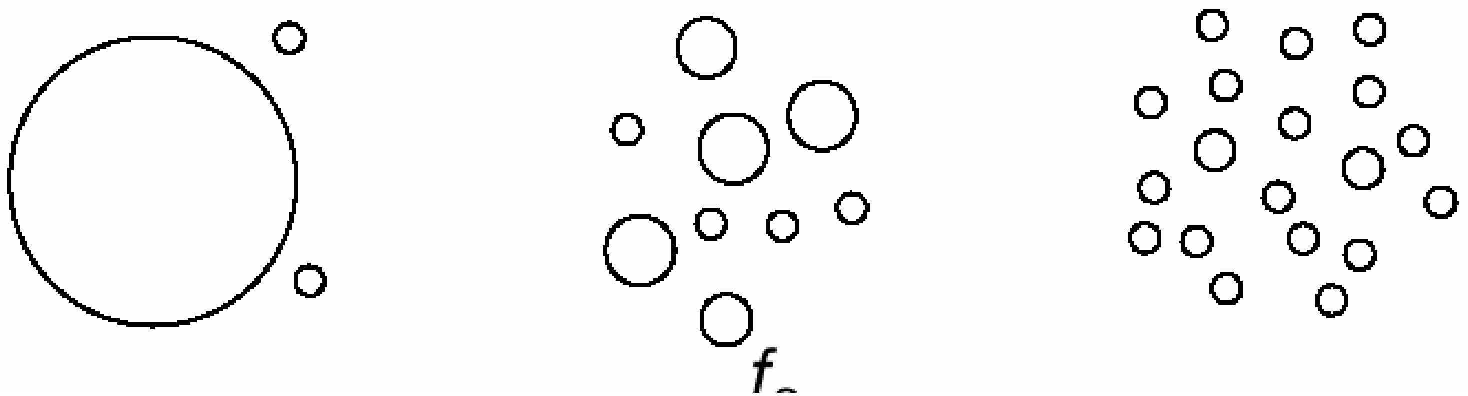




$\gamma \leq 3$: f_c=1

(R. Cohen et al PRL, 2000)

Attacks



Failures = percolation

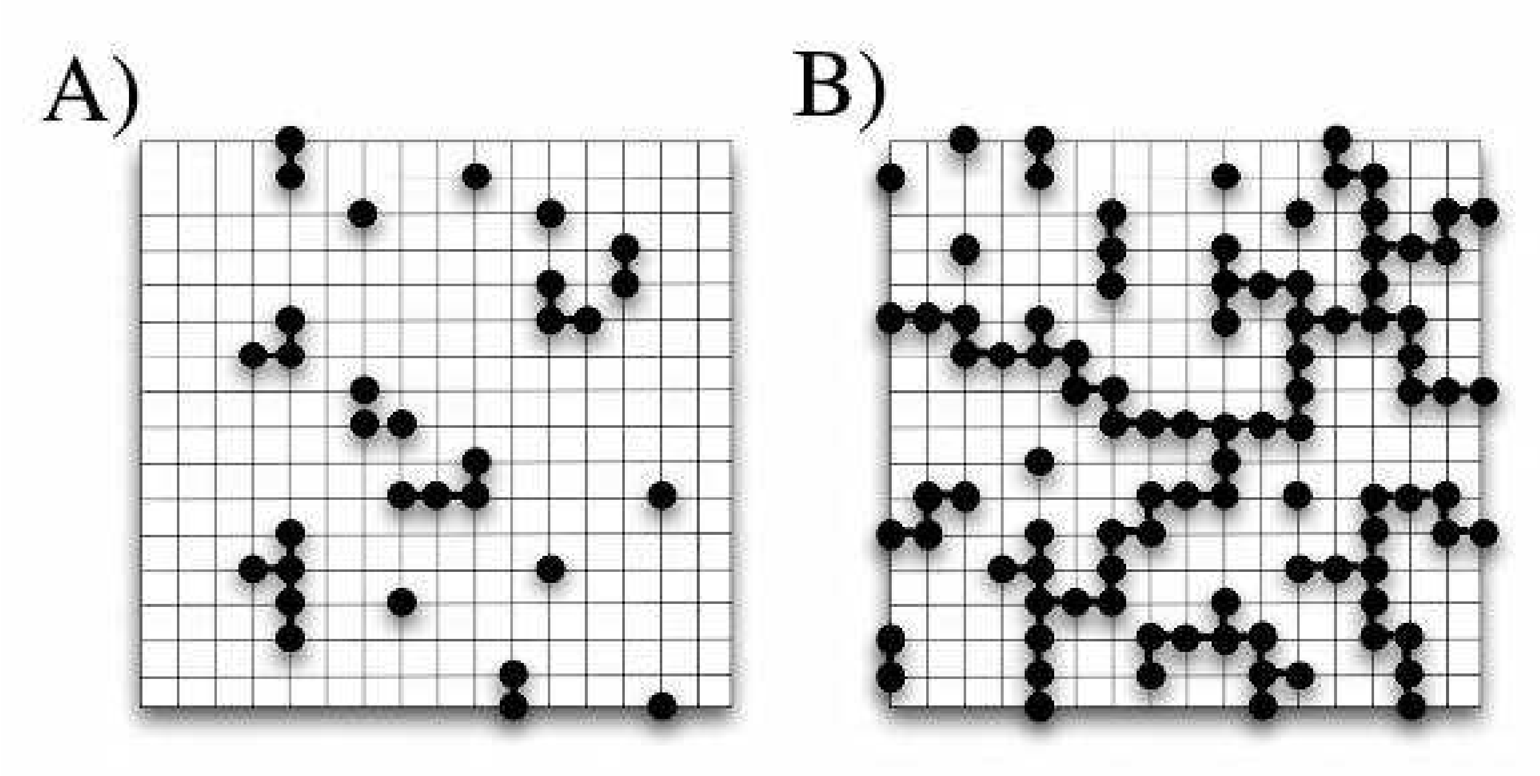
f=fraction of nodes removed because of failure

p=1-f

p=probability of a node to be present in a percolation problem

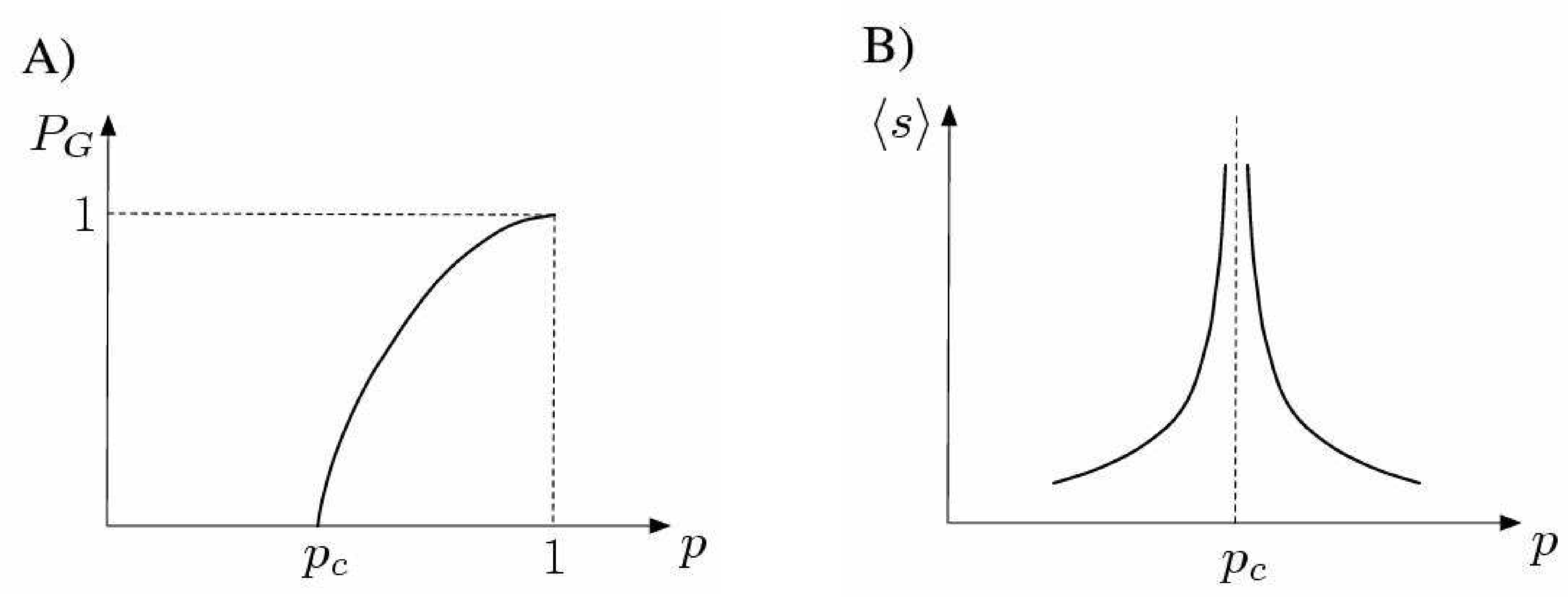
Question: existence or not of a giant/percolating cluster, i.e. of a connected cluster of nodes of size O(N)

Percolation



Question: existence or not of a giant/percolating cluster, i.e. of a connected cluster of nodes of size O(N)

Percolation



Question: existence or not of a giant/percolating cluster, i.e. of a connected cluster of nodes of size O(N)

Percolation in complex networks

q=probability that a randomly chosen link does **not** lead to a giant percolating cluster

$$q = \sum_k \frac{kP(k)}{\langle k \rangle} q^{k-1}$$

Average over degrees

Proba that the link leads to a node of degree k

Proba that none of the outgoing k-1 links leads to a giant cluster

NB: uncorrelated random networks

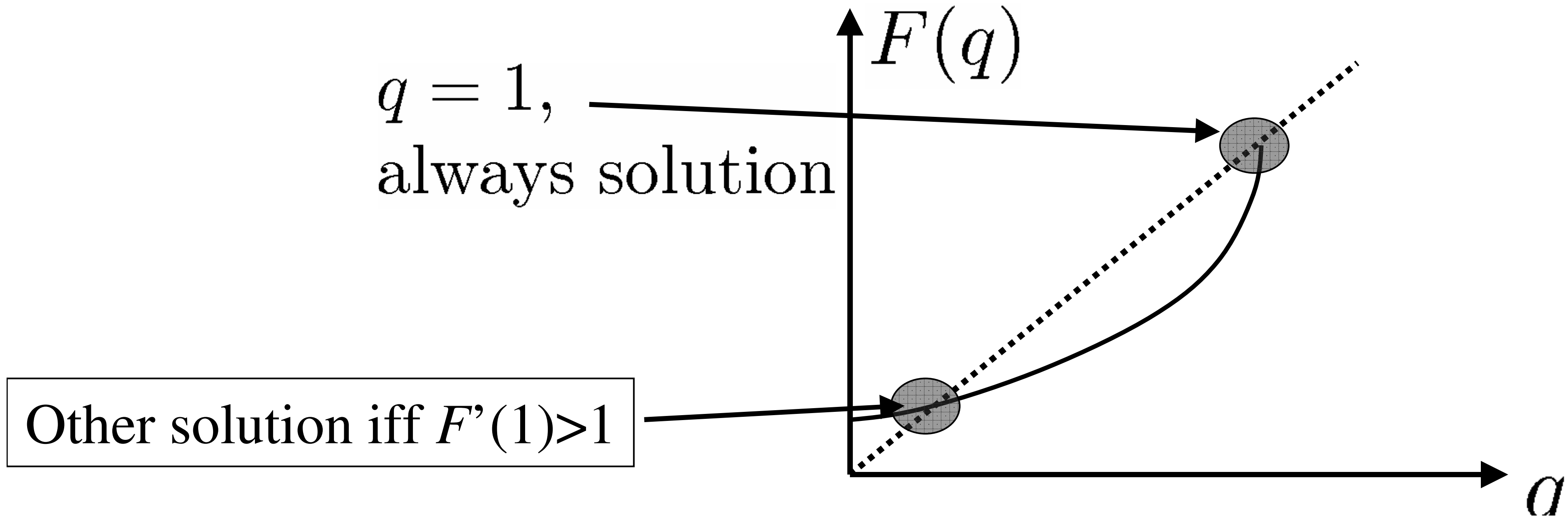
Percolation in complex networks

q=probability that a randomly chosen link does **not** lead to a giant percolating cluster

$$q = F(q), \text{with} \quad F(q) = \frac{1}{\langle k \rangle} \sum_k k P(k) q^{k-1}$$

$$F(0) > 0, F(1) = 1$$

$$F'(q), F''(q) > 0$$



Percolation in complex networks

q=probability that a randomly chosen link does **not** lead to a giant percolating cluster

$$q = F(q), \text{with} \qquad F(q) = \frac{1}{\langle k \rangle} \sum_k k P(k) q^{k-1}$$

$$F'(1) \geq 1 \quad \Longleftrightarrow \quad \boxed{\langle k^2 \rangle \geq 2\langle k \rangle}$$

“Molloy-Reed” criterion for the existence of a giant cluster in a random uncorrelated network

Back to random failures

Initial network: $P_0(k)$, $\langle k \rangle_0$, $\langle k^2 \rangle_0$

After removal of fraction f of nodes: $P_f(k)$, $\langle k \rangle_f$, $\langle k^2 \rangle_f$

Node of degree k_0 becomes of degree k with proba

$$C_{k_0}^{k}(1-f)^k f^{k_0-k}$$

$$\Rightarrow P_f(k) = \sum_{k_0} P_0(k_0) C_{k_0}^{k}(1-f)^k f^{k_0-k}$$

Back to random failures

Initial network: $P_0(k)$, $\langle k \rangle_0$, $\langle k^2 \rangle_0$

After removal of fraction f of nodes: $P_f(k)$, $\langle k \rangle_f$, $\langle k^2 \rangle_f$

$$\begin{cases} \langle k \rangle_f = (1-f)\langle k \rangle_0 \\ \langle k^2 \rangle_f = (1-f)^2 \langle k^2 \rangle_0 + f(1-f)\langle k \rangle_0 \end{cases}$$

Molloy-Reed criterion:
existence of a giant cluster iff $\langle k^2 \rangle_f \geq 2\langle k \rangle_f$

$$f \geq f_c, \text{with } \boxed{f_c = 1 - \frac{\langle k \rangle_0}{\langle k^2 \rangle_0 - \langle k \rangle_0}}$$

$\langle k^2 \rangle_0 \to \infty$ ⟹ $f_c \to 1$ ⟺ **Robustness!!!**

Finite-size effects

Finite number of nodes *N*

⇒ *Finite cut-off for P(k)*

⇒ *Finite* $\kappa = \langle k^2 \rangle / \langle k \rangle$

Example: scale-free network, min. degree m, $P(k) = ck^{-\gamma}$ $k = m, m+1, ..., k_c(N)$

Cut-off $k_c(N)$ defined by

$$N \int_{k_c(N)}^{\infty} P(k) dk = 1$$

$$k_c(N) = mN^{1/(\gamma-1)}$$

$$\kappa = \frac{\langle k^2 \rangle}{\langle k \rangle} = \frac{2-\gamma}{3-\gamma} \cdot \frac{k_c(N)^{3-\gamma} - m^{3-\gamma}}{k_c(N)^{2-\gamma} - m^{2-\gamma}}$$

Finite-size effects

Example: scale-free network, min. degree m, $P(k) = ck^{-\gamma}$ $k = m, m+1, ..., k_c(N)$

$$k_c(N) = mN^{1/(\gamma-1)}$$

$$\kappa = \frac{\langle k^2 \rangle}{\langle k \rangle} = \frac{2-\gamma}{3-\gamma} \cdot \frac{k_c(N)^{3-\gamma} - m^{3-\gamma}}{k_c(N)^{2-\gamma} - m^{2-\gamma}}$$

$N \to \infty$
$k_c \to \infty$

$\gamma > 3$: $\kappa \approx \frac{\gamma-2}{\gamma-3} m$ finite => f_c finite

$2 < \gamma < 3$: $$f_c \approx 1 - \frac{3-\gamma}{\gamma-2} m^{2-\gamma} k_c(N)^{\gamma-3} \longrightarrow 1$$

Ex: N=1000, m=1, γ=2.5
=> k_c = 100, $f_c \approx 0.9$

Attacks: various strategies

- Most connected nodes
- Nodes with largest betweenness
- Removal of links linked to nodes with large k
- Removal of links with largest betweenness
- Cascades
- ...

Attacks: most connected nodes

Removal of a fraction f of nodes, such that these nodes are the most connected ones:

Implicit equation defining the largest degree after removal::

$$f = \sum_{k=k_c(f)+1}^{\infty} P(k)$$

=> Modification of the degree distribution of the remaining nodes

Attacks: most connected nodes

Removal of a fraction f of nodes, such that these nodes are the most connected ones

Modification of the degree distribution of the remaining nodes:

Probability that a neighbor of a given node has been removed=

probability that the neighbor has degree > $k_c(f)$ =

$$r(f) = \sum_{k=k_c(f)+1}^{\infty} \frac{kP(k)}{\langle k \rangle}$$

(in a random uncorrelated network)

Attacks: most connected nodes

Removal of a fraction f of nodes, such that these nodes are the most connected ones

Remaining network=

- ***Cut-off $k_c(f)$***
- ***Random removal with proba $r(f)$***

Molloy-Reed criterion => threshold f_c at which the giant component disappears

$$r(f_c) = 1 - \frac{1}{\kappa(f_c) - 1}$$

$$\kappa(f_c) = \frac{\sum_{k=1}^{k_c(f_c)} k^2 P(k)}{\sum_{k=1}^{k_c(f_c)} k P(k)}$$

Attacks: most connected nodes

Example: scale-free network, min. degree m $P(k) = ck^{-\gamma}$

$$f = \sum_{k=k_c(f)+1}^{\infty} P(k) \quad \Rightarrow \quad k_c(f) = m f^{1/(1-\gamma)}$$

$$r(f) = \sum_{k=k_c(f)+1}^{\infty} \frac{kP(k)}{\langle k \rangle} \approx f^{(2-\gamma)/(1-\gamma)}$$

$$\kappa(f) = \frac{2-\gamma}{3-\gamma} \cdot \frac{k_c(f)^{3-\gamma} - m^{3-\gamma}}{k_c(f)^{2-\gamma} - m^{2-\gamma}}$$

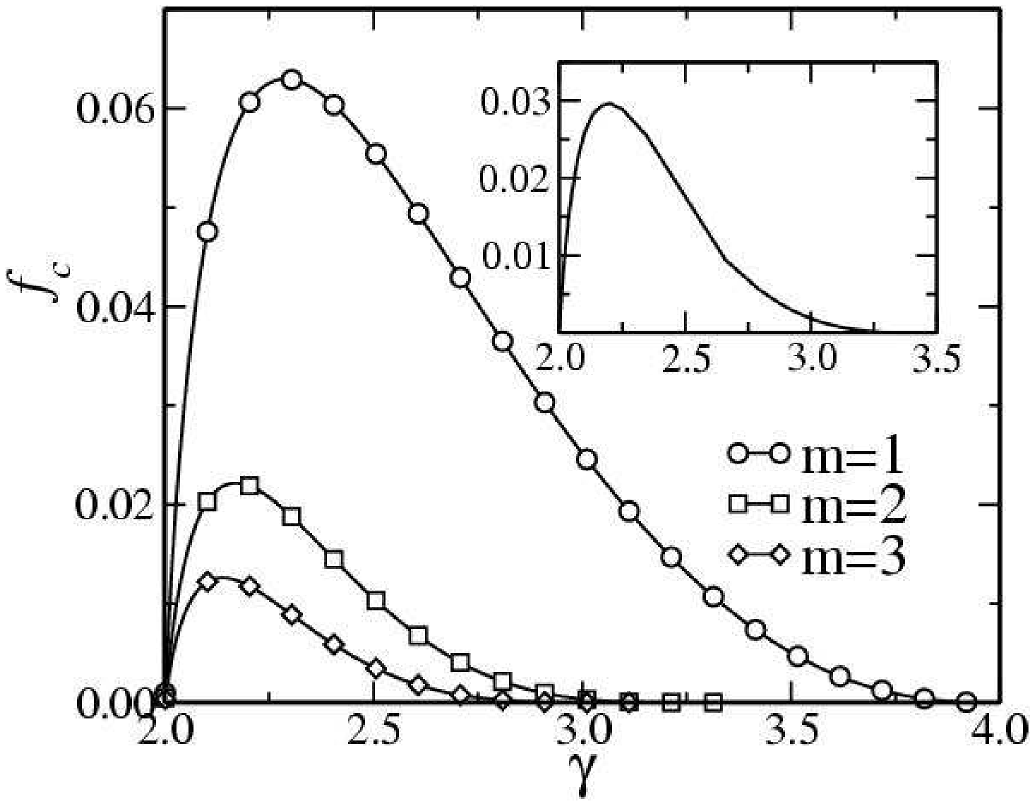
$$r(f_c) = 1 - \frac{1}{\kappa(f_c) - 1}$$

$$f_c^{(2-\gamma)/(1-\gamma)} \approx 2 + \frac{2-\gamma}{3-\gamma} m \left(f_c^{(3-\gamma)/(1-\gamma)} - 1\right)$$

Attacks: most connected nodes

Example: scale-free network, min. degree m $P(k) = ck^{-\gamma}$

$$f_c^{(2-\gamma)/(1-\gamma)} \approx 2 + \frac{2-\gamma}{3-\gamma} m \left(f_c^{(3-\gamma)/(1-\gamma)} - 1 \right)$$



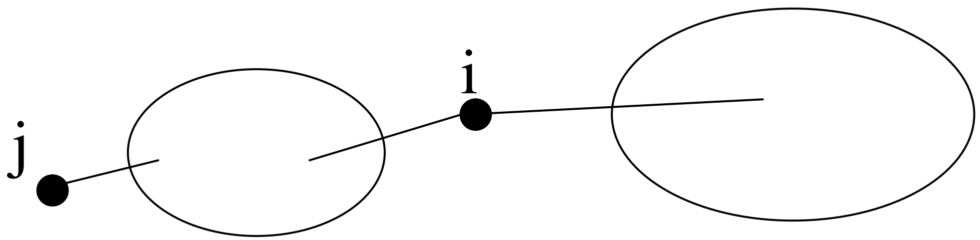
Betweenness

⇒measures the "centrality" of a node i:

for each pair of nodes (l,m) in the graph, there are

- σ^{lm} shortest paths between l and m
- σ_i^{lm} shortest paths going through i

b_i is the sum of $\sigma_i^{lm} / \sigma^{lm}$ over all pairs (l,m)



b_i is large
b_j is small

Attacks: other strategies

- Nodes with largest betweenness
- Removal of links linked to nodes with large k
- Removal of links with largest betweenness
- Cascades
- … **Problem of reinforcement ?**

P. Holme et al (2002); A. Motter et al. (2002);
D. Watts, PNAS (2002); Dall'Asta et al. (2006)…