

INTRODUCTION TO NETWORK SCIENCE

János Kertész

janos.kertesz@gmail.com

9. RANDOM FAILURES, INTENTIONAL ATTACKS AND CASCADING CATASTROPHES

Stability of networks

Complex systems have many units, components, with different kinds of interactions, which all can break down.

First guess: Complexity causes vulnerability

Wrong: There must be mechanisms to make complex systems robust – otherwise we could not observe them!

Robustness should be reflected in the networks deduced from complex systems.

Stability of networks

Topological **robustness against random failures:**

Elements of the network (nodes or links) break down, i.e., are randomly removed – and the networks still functions!

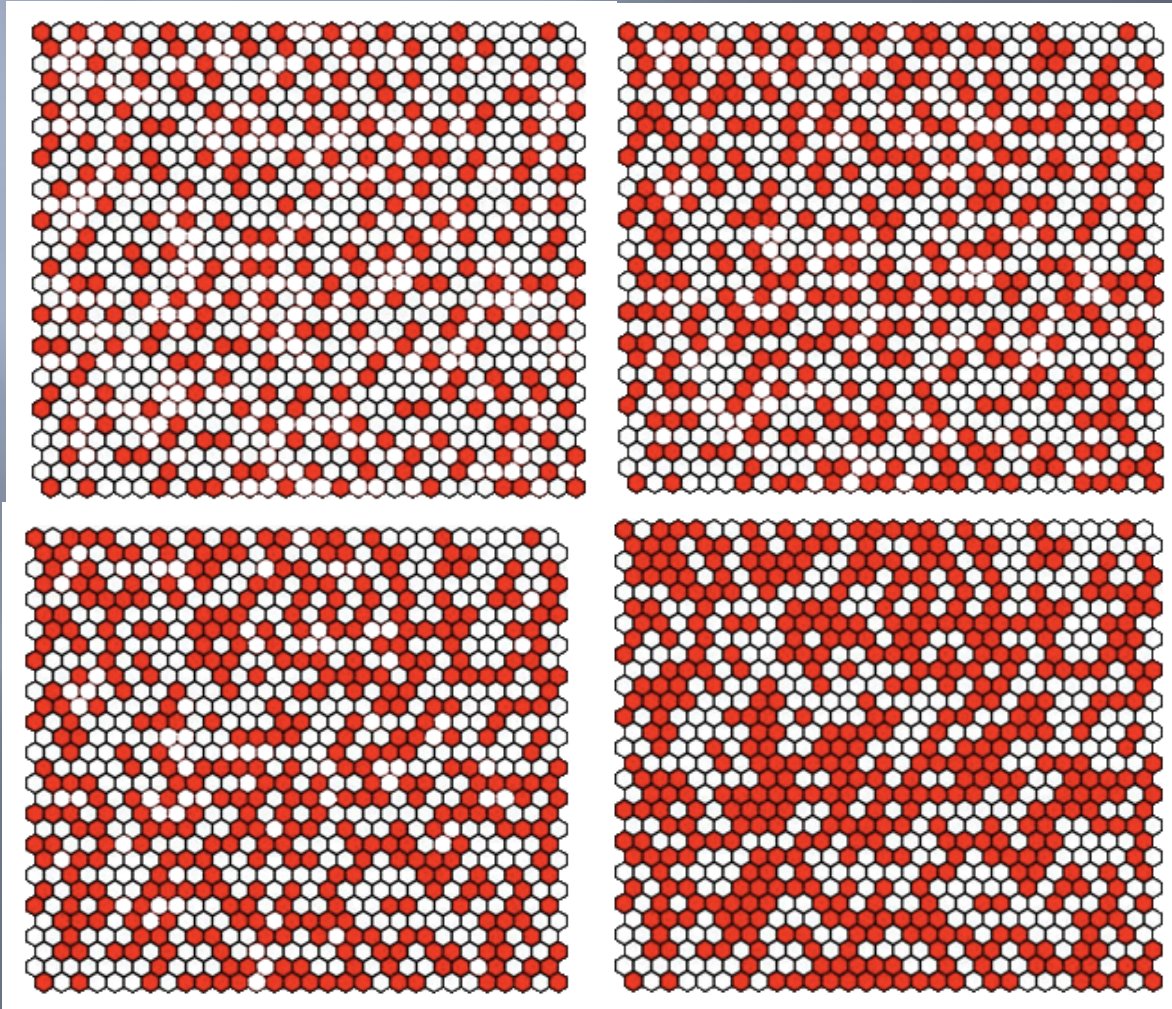
A simple definition of functioning: The survival of the giant component.

Robustness against intentional attack:

Elements are removed in a vicious/efficient manner to cause most harm/effect.

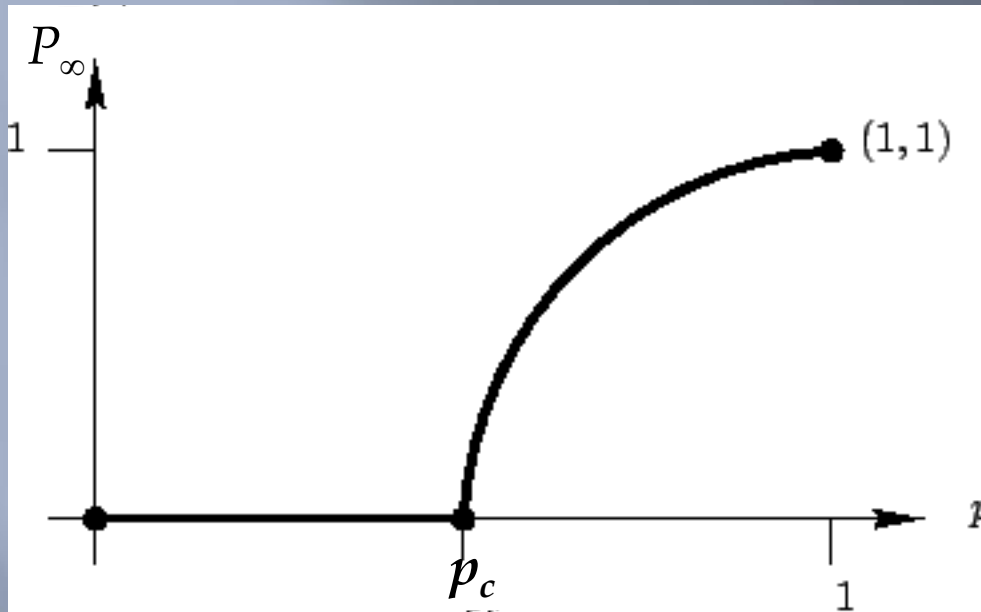
Percolation (recapitulation)

The random failure case is a percolation problem on a complex network (site or bond \leftrightarrow node or link)

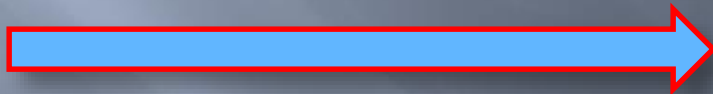


We do not **occupy** sites at random but **remove** them randomly – the same theory can be applied

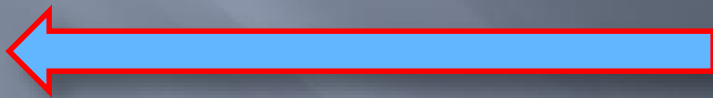
Percolation



P_∞ is the relative weight or density of the infinite cluster or giant component.

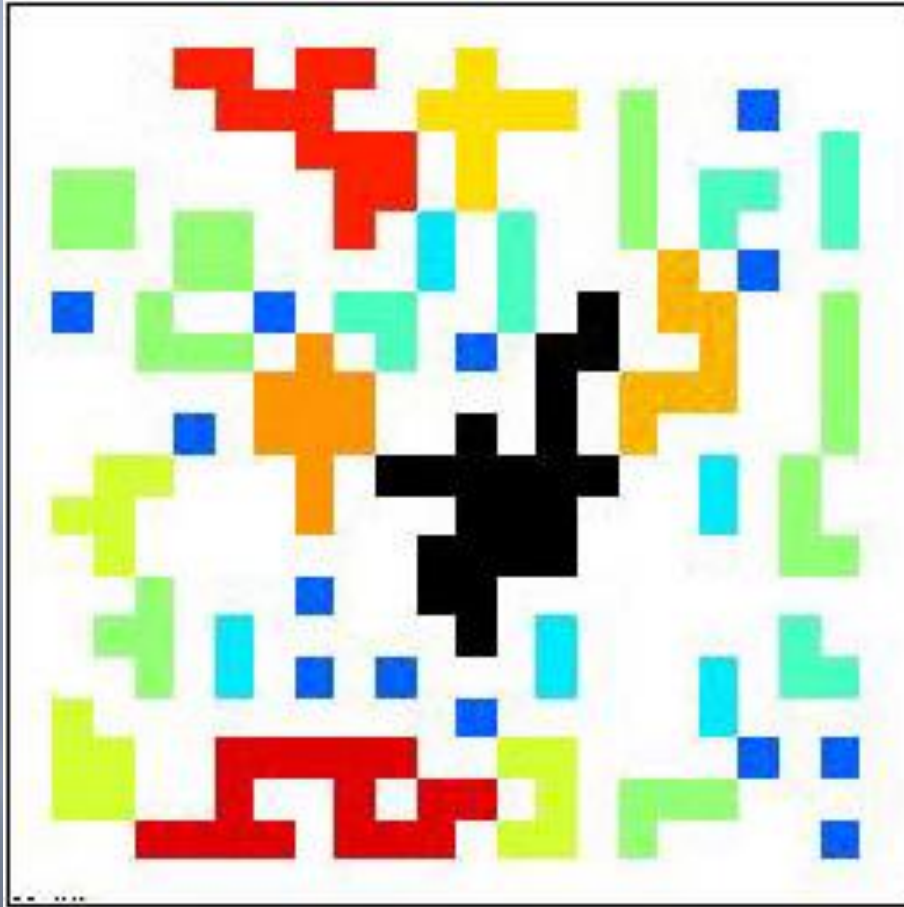


occupation/activation

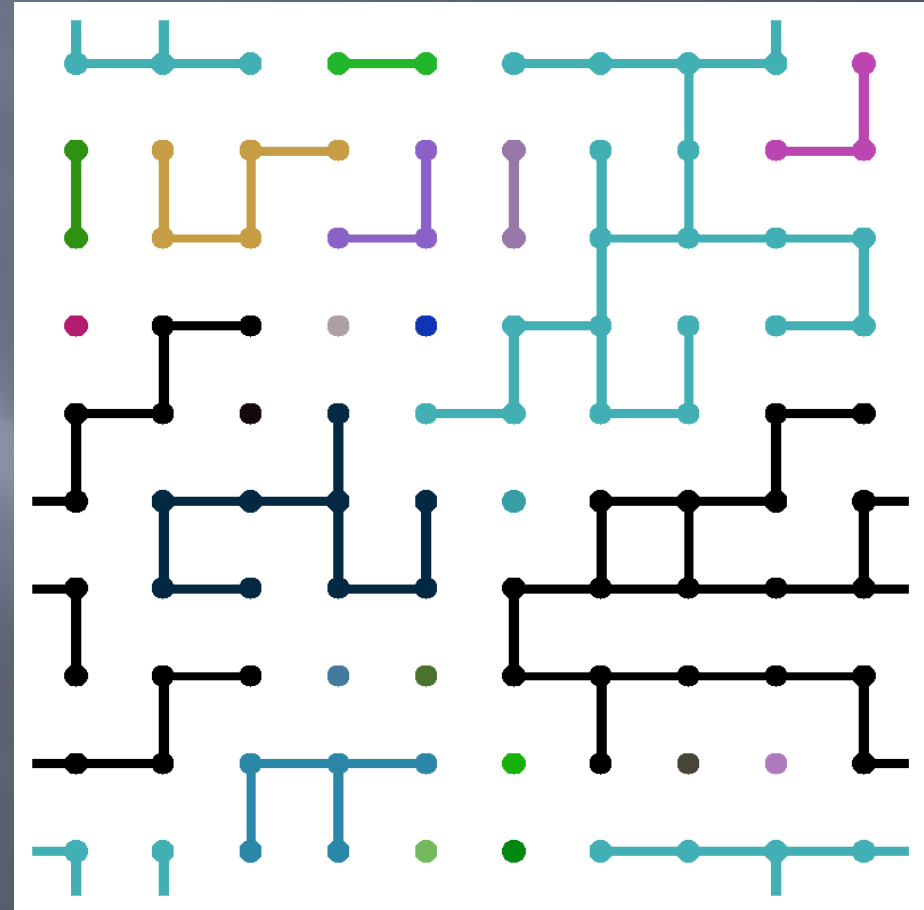


failure/removal

Percolation



Site percolation
Wrong nodes removed

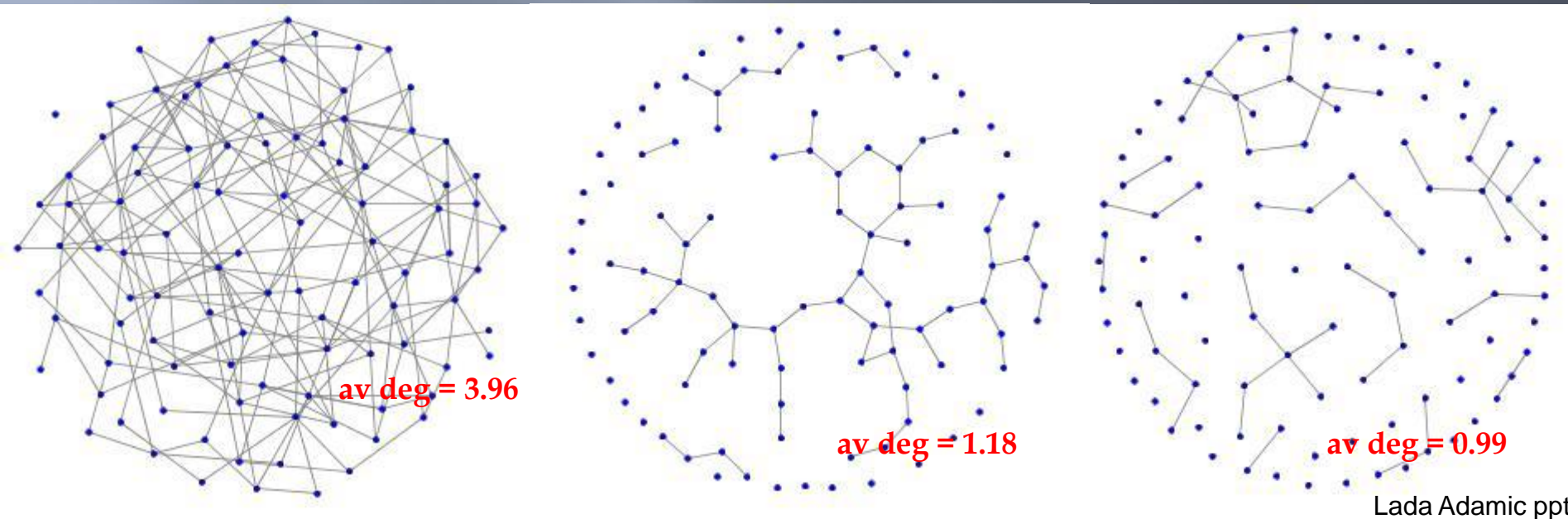


Bond percolation
Broken links removed

Percolation

Already mentioned results:

1. Erdős-Rényi graph has a phase transition at $\langle k \rangle = 1$:
 $\langle k \rangle < 1$ no giant component
 $\langle k \rangle > 1$ there is a giant component



Bond percolation: links are removed until the giant component falls apart.
Note that

$$p_c = \frac{L}{N(N-1)/2} = \frac{N/2}{N(N-1)/2} = \frac{1}{N-1} \rightarrow 0$$

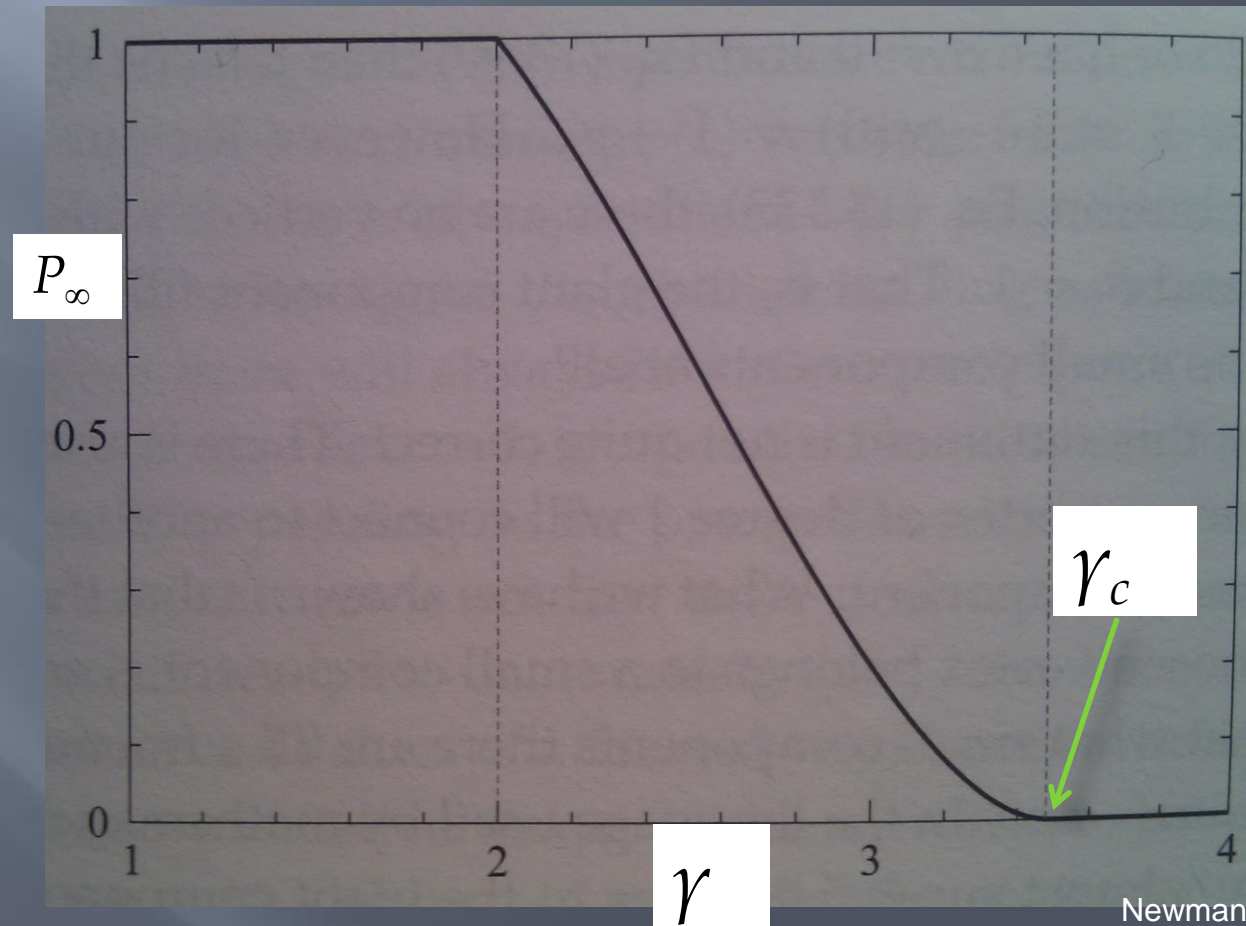
Percolation

2. Configuration model (random network with arbitrary degree distribution):

Molloy-Reed criterion: There is giant component if

$$\langle k^2 \rangle - 2\langle k \rangle > 0$$

This suggests that for small enough γ breaking the giant component will be difficult.



Random failure

Based on the MR criterion, we can introduce a parameter, the **inhomogeneity ratio**, indicating on which side of the transition we are:

$$\kappa \equiv \frac{\langle k^2 \rangle}{\langle k \rangle} = 2$$

$\kappa > 2$: a giant cluster exists

$\kappa < 2$: many disconnected clusters

Critical point: $\kappa = 2$

How does the degree distribution change upon node deletion?

First example: ER graph. A randomly diluted ER graph is also an ER graph. Degrees: Poisson-distributed

$$P(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}$$

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

$$\kappa = \frac{\langle k^2 \rangle}{\langle k \rangle} = (1 + \langle k \rangle) = 2 \Rightarrow \langle k \rangle = 1$$

Random failure

What about general degree distributions?

Let us consider the configuration model.

Removing a node changes the degrees of the neighboring nodes and the degree distribution.

Let us call the fraction of nodes removed $f = 1 - p$.

Random failure

The probability that an originally k -degree node becomes a $k' < k$ degree node is a binomial distribution (remove $k - k'$ nodes from k at random) :

$$\binom{k}{k'} f^{k-k'} (1-f)^{k'} \quad k' \leq k$$

The probability that we chose a node with k nodes is $P(k)$
Thus the resulting new degree distribution will be

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

How do the moments change? (We need them for the MR criterion.)

E.g., for $\langle k' \rangle_f$
Similar for $\langle k'^2 \rangle_f$

$$\langle k' \rangle_f = \sum_{k'=0}^{\infty} k' P'(k') = \sum_{k'=0}^{\infty} k' \sum_{k=k'}^{\infty} P(k) \binom{k}{k'} f^{k-k'} (1-f)^{k'}$$

Random failure

$$\sum_{n=0}^m \binom{m}{n} f^{m-n} (1-f)^n = 1$$

$$\langle k' \rangle_f = \sum_{k'=1}^{\infty} \sum_{k=k'}^{\infty} k' P(k) \frac{k!}{(k')!(k-k')!} f^{k-k'} (1-f)^{k'-1} (1-f) =$$

$$\sum_{k=1}^{\infty} (1-f) k P(k) \sum_{k'=1}^k \frac{(k-1)!}{(k'-1)!(k-k')!} f^{k-k'} (1-f)^{k'-1} =$$

$$\sum_{k=1}^{\infty} (1-f) k P(k) = (1-f) \langle k \rangle$$

Trick 1: $\sum_{k'=1}^{\infty} \sum_{k=k'}^{\infty} = \sum_{k=1}^{\infty} \sum_{k'=1}^k$ summation order

Trick 2: $k! = k(k-1)!$ both in denominator and nominator

Trick 3: $\sum_{n=0}^m \binom{m}{n} f^{m-n} (1-f)^n = 1$ (norm of binom. distr. and $k'-1 \rightarrow n, k-1 \rightarrow m$ substitution)

Random failure

$$\langle k' \rangle_f = (1 - f) \langle k \rangle$$

Similarly for the second moment:

$$\langle k'^2 \rangle_f = (1 - f)^2 \langle k^2 \rangle + f(1 - f) \langle k \rangle$$

Can be derived in a similar manner as $\langle k' \rangle_f$ using

Trick 4:
$$\langle k'^2 \rangle_f = \langle k'(k'-1) \rangle_f + \langle k' \rangle_f$$

Now we can apply the RM criterion to the new degree distributions to determine the critical value of $f = (1-p)$

$$\kappa'_c \equiv \frac{\langle k'^2 \rangle_f}{\langle k' \rangle_f} = (1 - f_c) \kappa + f_c = 2$$

with

$$\kappa = \frac{\langle k^2 \rangle}{\langle k \rangle}$$

Random failure

$$\kappa'_c \equiv \frac{\langle k'^2 \rangle_f}{\langle k' \rangle_f} = (1 - f_c)\kappa + f_c = f_c(1 - \kappa) + \kappa = 2$$

$$f_c = 1 - \frac{1}{\kappa - 1} = \frac{\langle k^2 \rangle - 2\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}$$

If a random network (configuration nw) with original degree distribution $P(k)$ suffers from failures, which destroy the nodes randomly with probability f , the network will be resilient until f_c , i.e., up to that point the giant component will exist and the network can (hopefully) fulfill its function.

Random failure

Many networks have power law degree distributions. Let us consider the configuration model with power law distribution of degrees. Only $1 < \gamma < \gamma_c$ is interesting (existence of giant component).

$$\langle k^m \rangle = \sum_k k^m P(k) \sim \sum_k k^{m-\gamma}$$

In a infinite system the summation goes to ∞

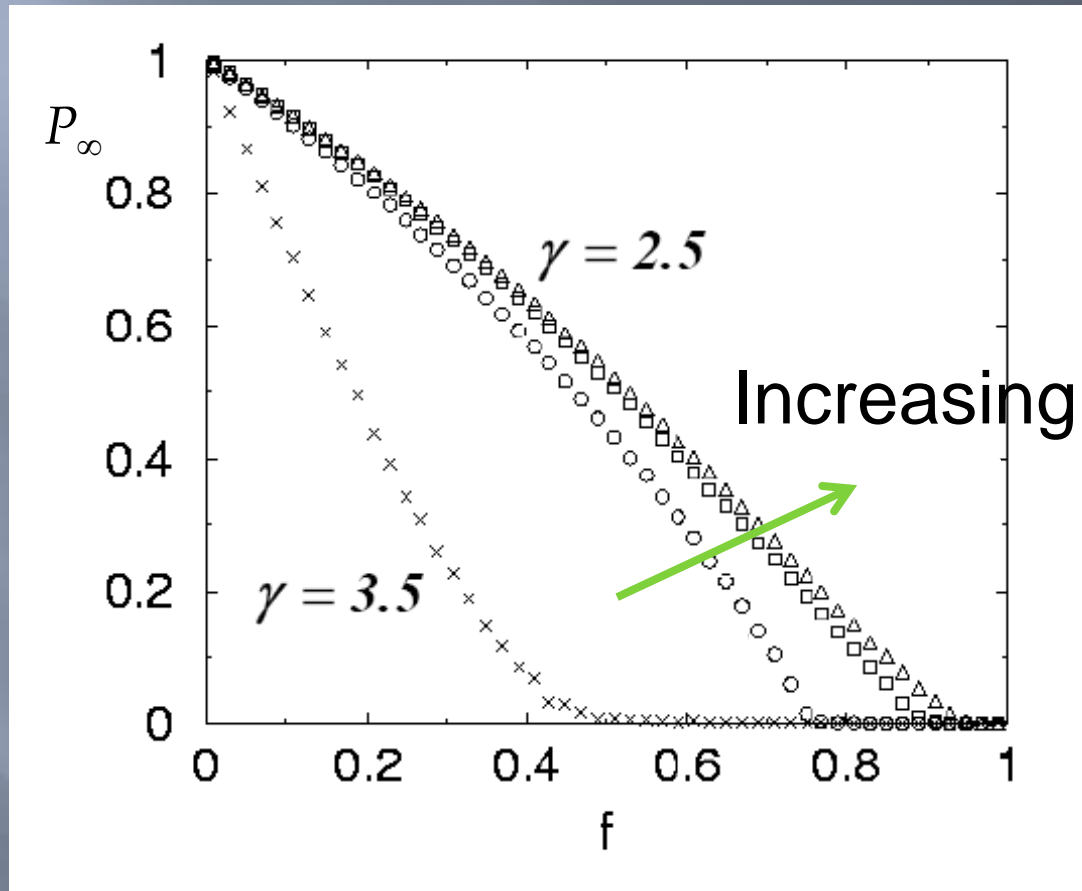
Moments exist only if $m < \gamma - 1$.

If $\gamma \leq 3$, the second moment diverges.

$$f_c = 1 - \frac{1}{\gamma - 1} = \frac{\langle k^2 \rangle - 2\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}$$

For $\gamma \leq 3$ the critical value $f_c = 1$, i.e., in an infinite system all finite fractions of the nodes have to be removed to destroy the giant component! **Resilience against failure**

Random failure



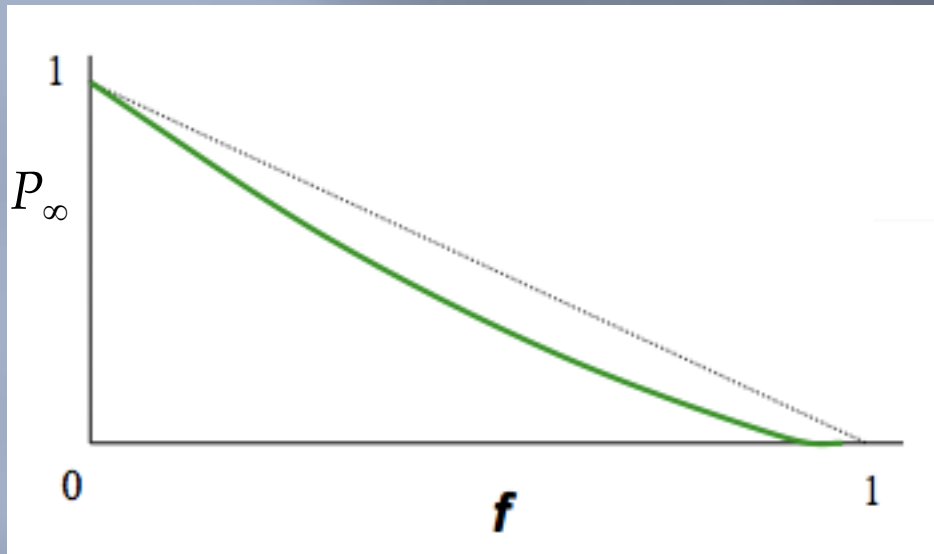
Random failure

While the derivations used massively the random nature of the configuration model, the lesson seems to be valid for empirical networks!

The reason is that the rare but important hubs assure the survival of the giant component and this is present in any power law degree distribution.

Empirical networks are finite, there we cannot expect $f_c = 1$ but values close to it, if the network is large.

Random failure



Empirical study: the Internet

Internet

Router level map, $N=228,263$; $\gamma=2.1\pm 0.1$; $\kappa=28$

→ $f_c=0.962$

AS level map, $N=11,164$; $\gamma=2.1\pm 0.1$; $\kappa=264$

→ $f_c=0.996$

Random failure

Error tolerance of scale free complex networks:
Possible source of their ubiquity.

Complex systems have to develop mechanisms to protect themselves against random failures. The scale free topology is one of them.

This robustness is not always advantageous:
How to stop the spreading of an infection (e.g., computer virus) on a scale free network? Isolating the disease by random removal of nodes (e.g, vaccination) is not a good strategy.

Intentional attacks

The robustness of scale free networks is due to the hubs, which are difficult to hit by chance.

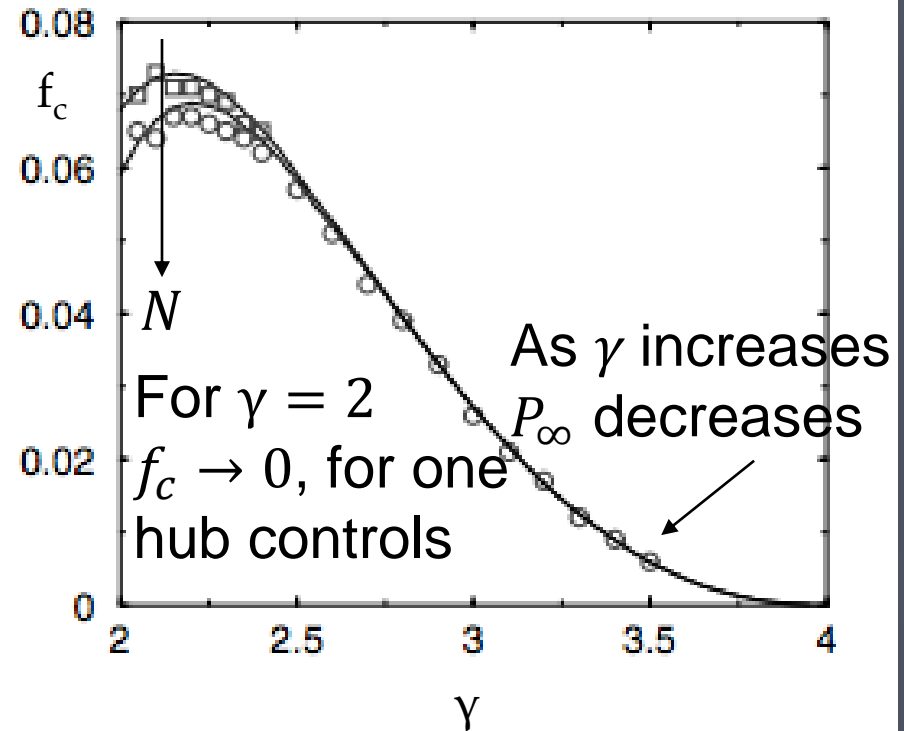
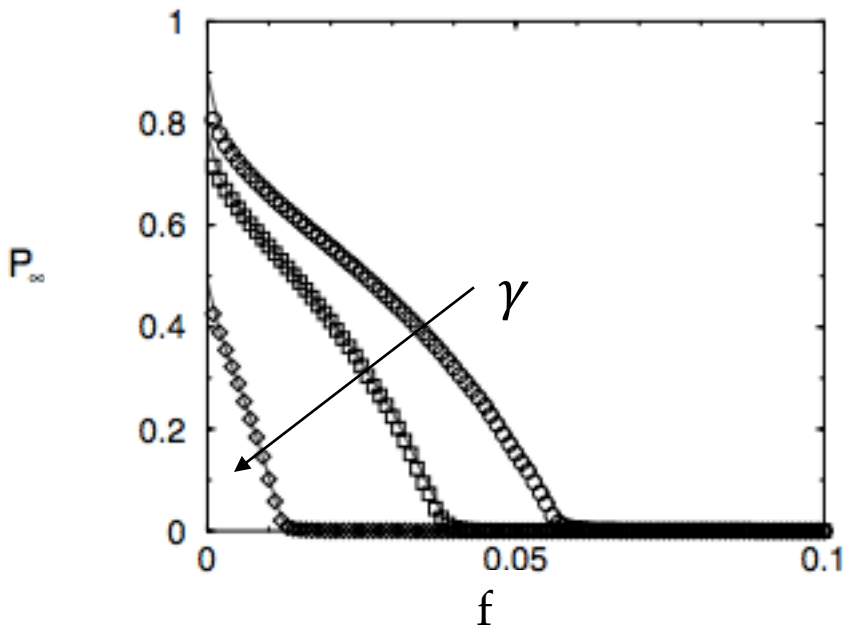
What if the removal is not random but vicious: Somebody (e.g, a terrorist) wants to paralyze the Internet.

The targeted attack should go against the hubs! We assume that the terrorist knows the network and destroys the a fraction of the highest connectivity nodes.

A more peaceful interpretation: How to do efficient vaccination?

Intentional attacks

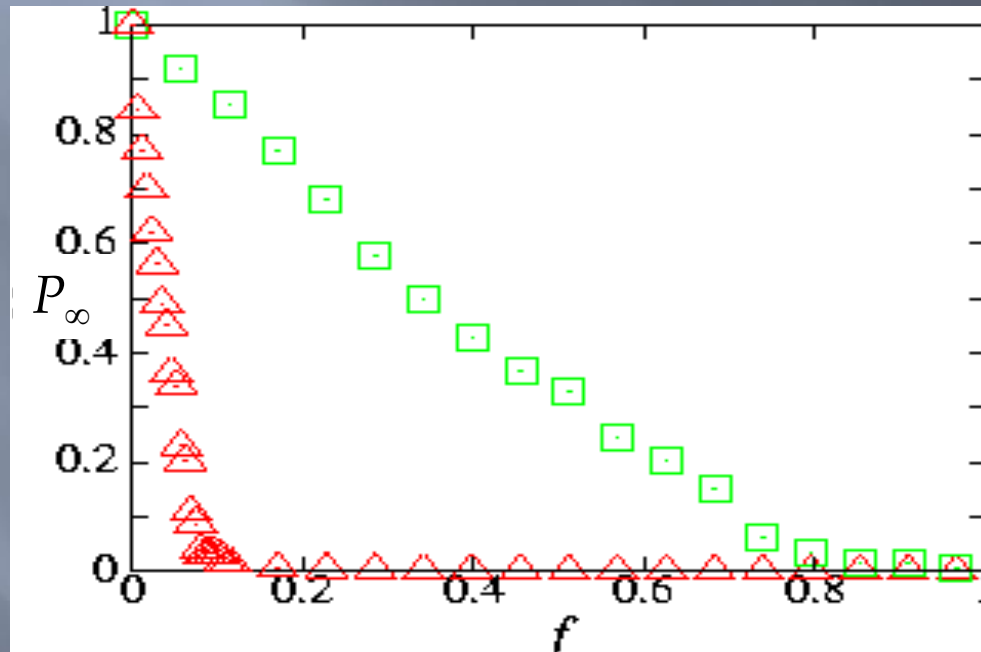
The analytical approach is based on the fact that this time a fraction f of the hubs is removed instead of nodes at random.



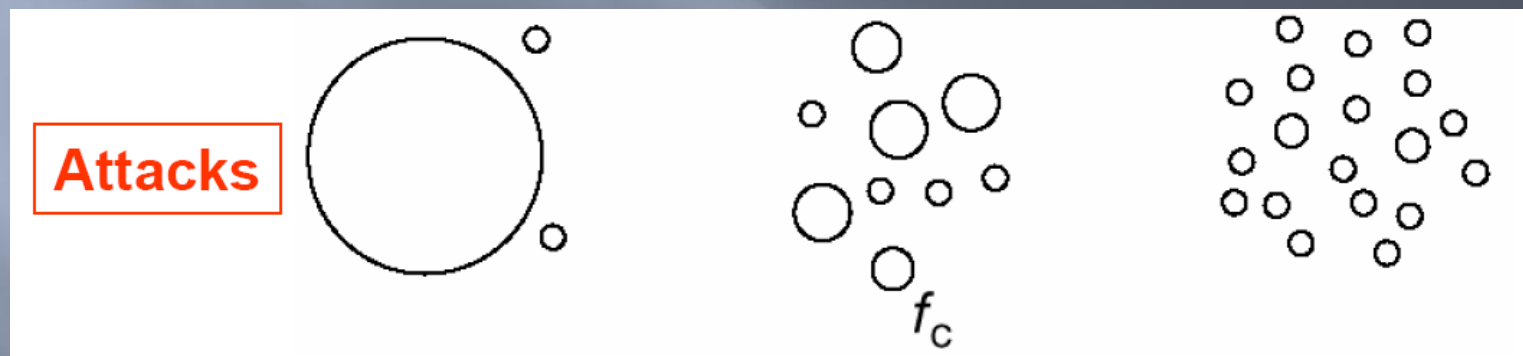
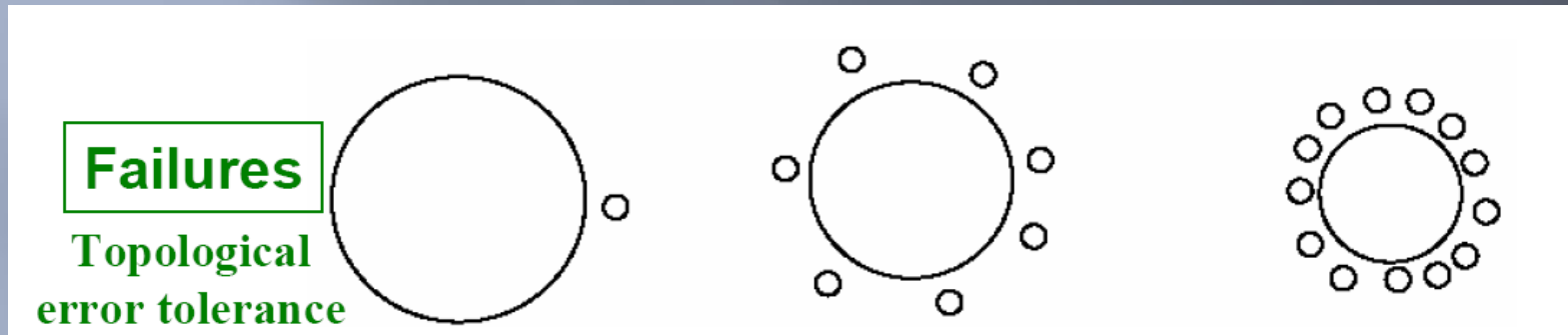
The calculations show that the situation changes dramatically: There will always be a threshold and its value is rather low.

Intentional attacks

Empirical comparison between random failure and intentional attack: Internet



Intentional attacks



$$f_c^{\frac{2-\gamma}{1-\gamma}} = 2 + \frac{2-\gamma}{3-\gamma} K_{min} (f_c^{\frac{3-\gamma}{1-\gamma}} - 1)$$

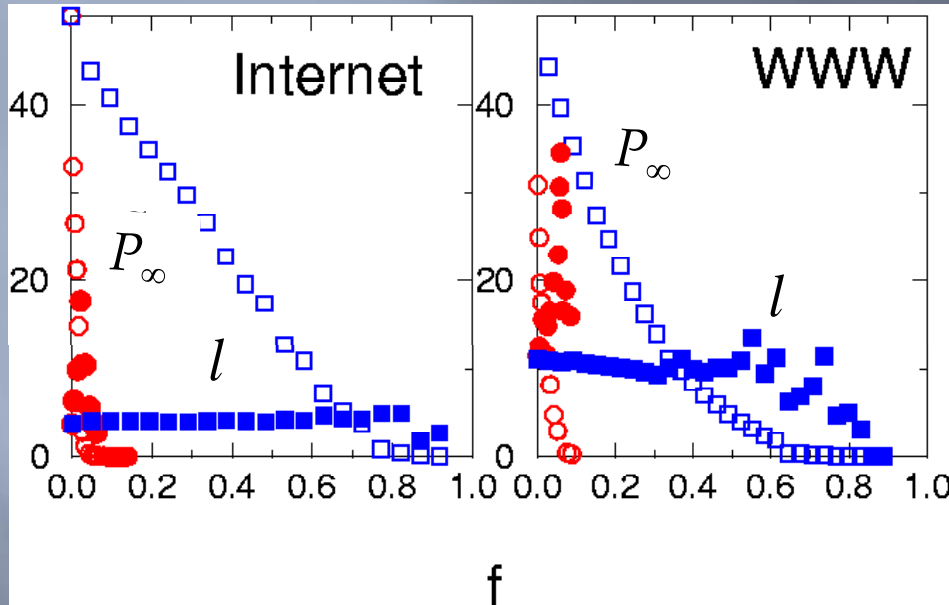
For scale free conf. model (implicit eq)

Cohen et al, 2001

Scale-free networks are more error tolerant, but also more vulnerable to attacks

Intentional attacks

Real scale-free networks show the same dual behavior



- blue squares: random failure
- red circles: targeted attack
- open symbols: P_∞
- filled symbols: l

- break down if 5% of the nodes are eliminated selectively (always the highest degree node)
- resilient to the random failure of 50% of the nodes.

Similar results have been obtained for metabolic networks and food webs.

All this is static!

Cascading breakdown

In spite of the overall tolerance against random failures complex systems show sometimes extreme vulnerability against them.

As these are nonlinear **dynamic** systems some failures may trigger further ones leading to collapse

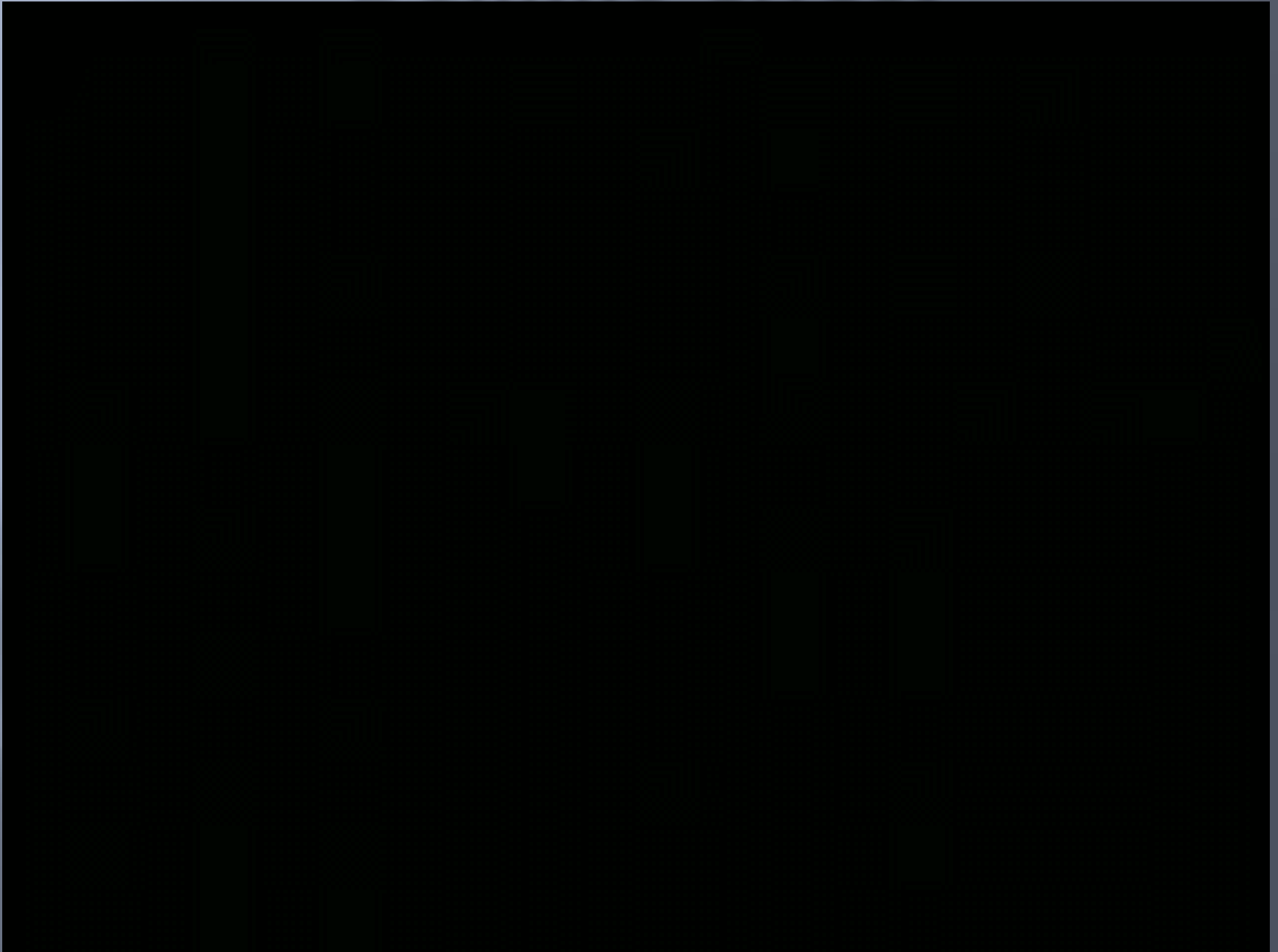
“Avalanches”, “cascading failures”, “domino effect”

Avalanche



teton
GRAVITY RESEARCH
SUBSCRIBE

Domino effect

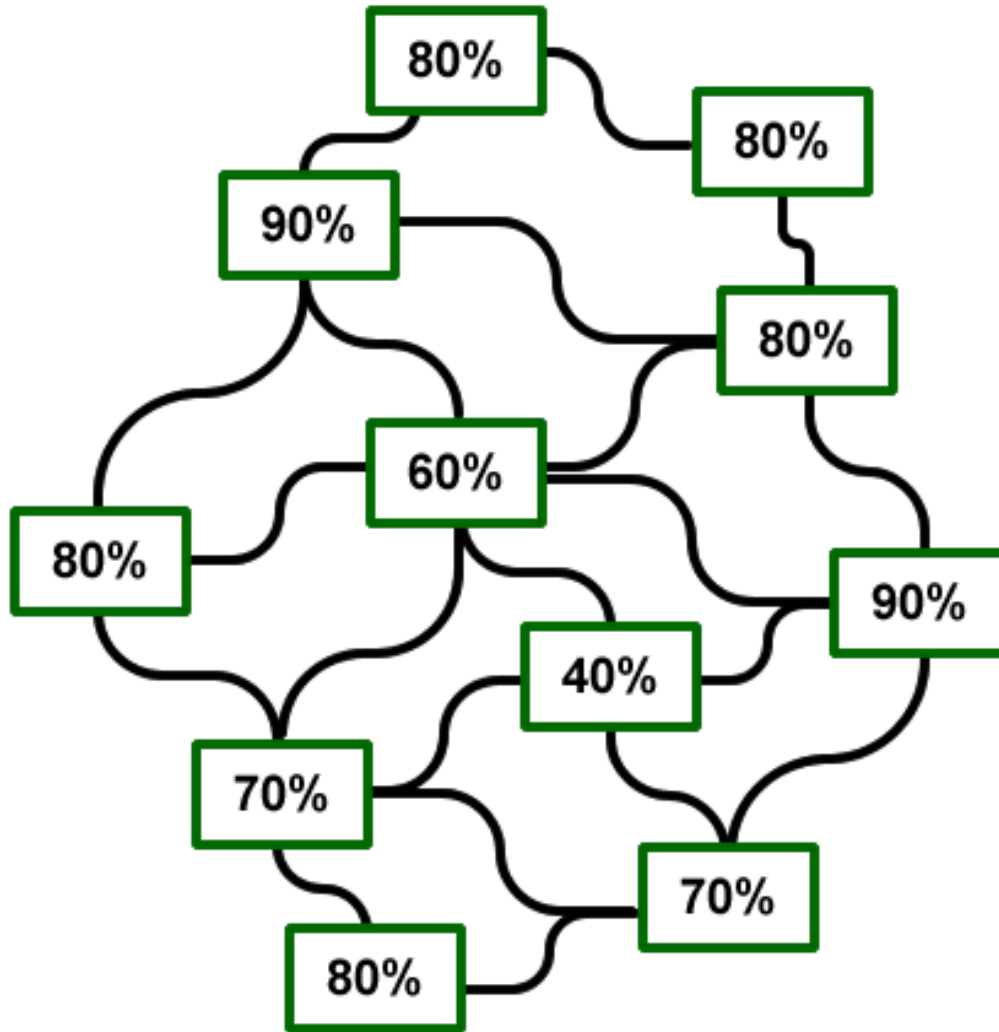


Cascading Failures

Finance:



Cascading Failures



Network running normally

Cascading Failures

Major North-East Blackout

Origin

A 3,500 MW power surge (towards Ontario) affected the transmission grid at 4:10:39 p.m. EDT. (Aug-14-2003)

Before the blackout



After the blackout

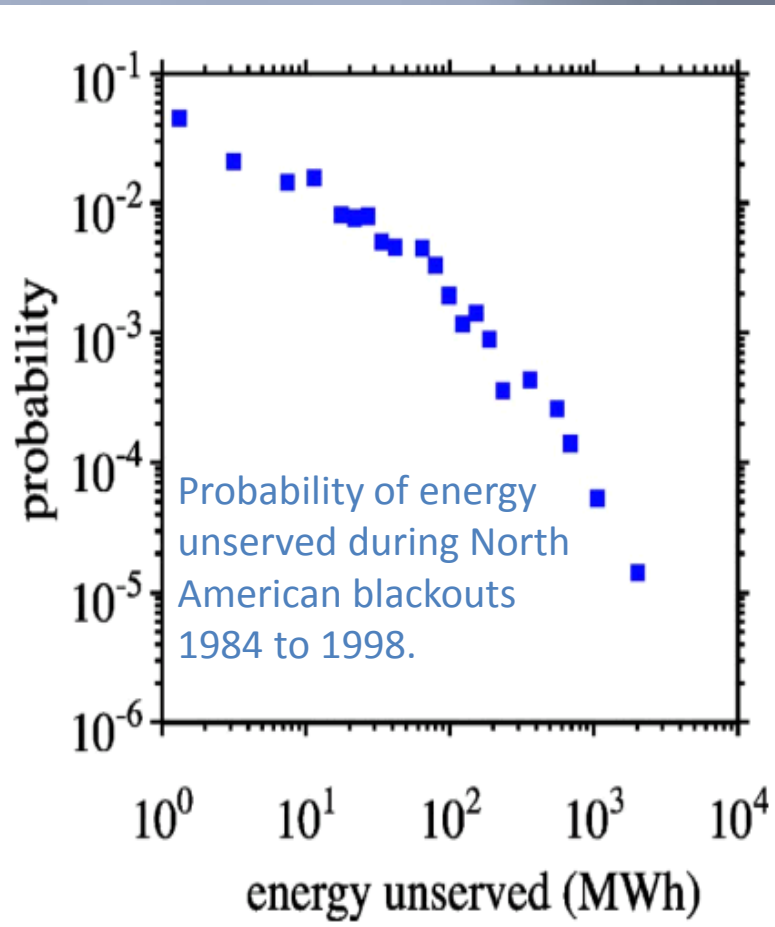


Consequences

More than 508 generating units at 265 power plants shut down during the outage. In the minutes before the event, the NYISO-managed power system was carrying 28,700 MW of load. At the height of the outage, the load had dropped to 5,716 MW, a loss of 80%.

Cascading Failures

Size distribution of blackouts



$$P(S) \sim S^{-\alpha}, 1 < \alpha < 2$$

Source	Exponent	Quantity
North America	2.0	Power
Sweden	1.6	Energy
Norway	1.7	Power
New Zealand	1.6	Energy
China	1.8	Energy

Self-organized criticality

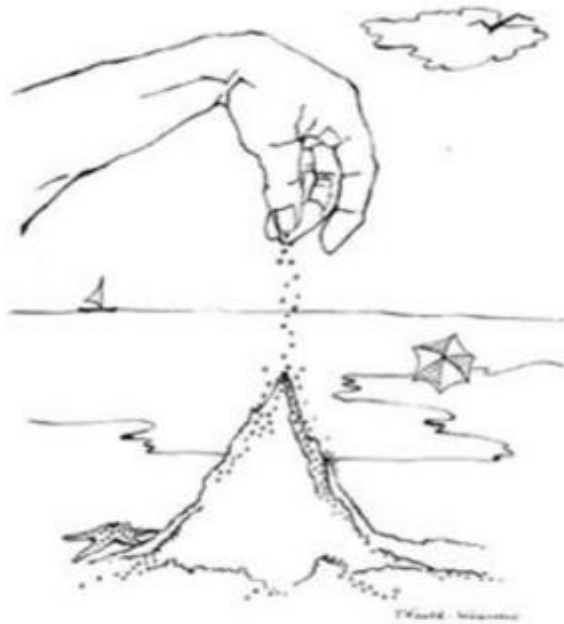
Per Bak (1987): Why is power-law so ubiquitous?
The critical state emerges as a result of self-organization!

Example: sandpile (more language than concrete physical example)



Avalanche statistics:
power-law
(not for sand...)

Self-organized criticality



Per Bak's sand pile



*Power-law distributed
avalanches in a rice pile*

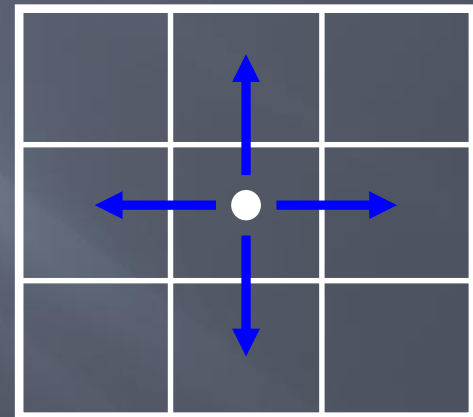
SOC Model

- Starting with a flat surface $Z(x,y) = 0$ for all x and y .
- Add a grain of sand: $Z(x,y) = Z(x,y) + 1$.
- Start avalanche if $Z(x,y) > Z_c$:

$$Z(x, y) = Z(x, y) - 4$$

$$Z(x \pm 1, y) = Z(x \pm 1, y) + 1$$

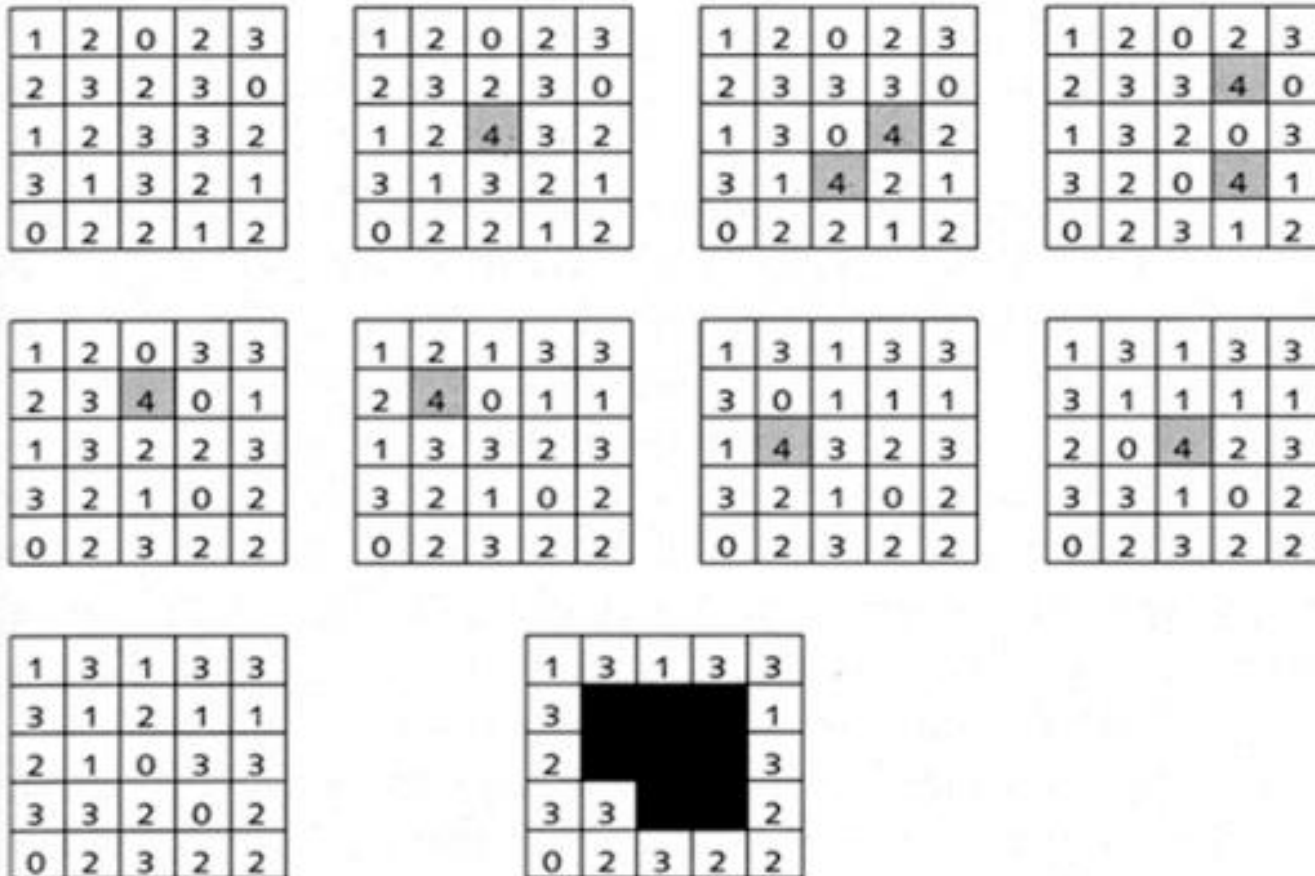
$$Z(x, y \pm 1) = Z(x, y \pm 1) + 1$$



At the boundary “grains” leave the system assuring constant density on the average

SOC-model

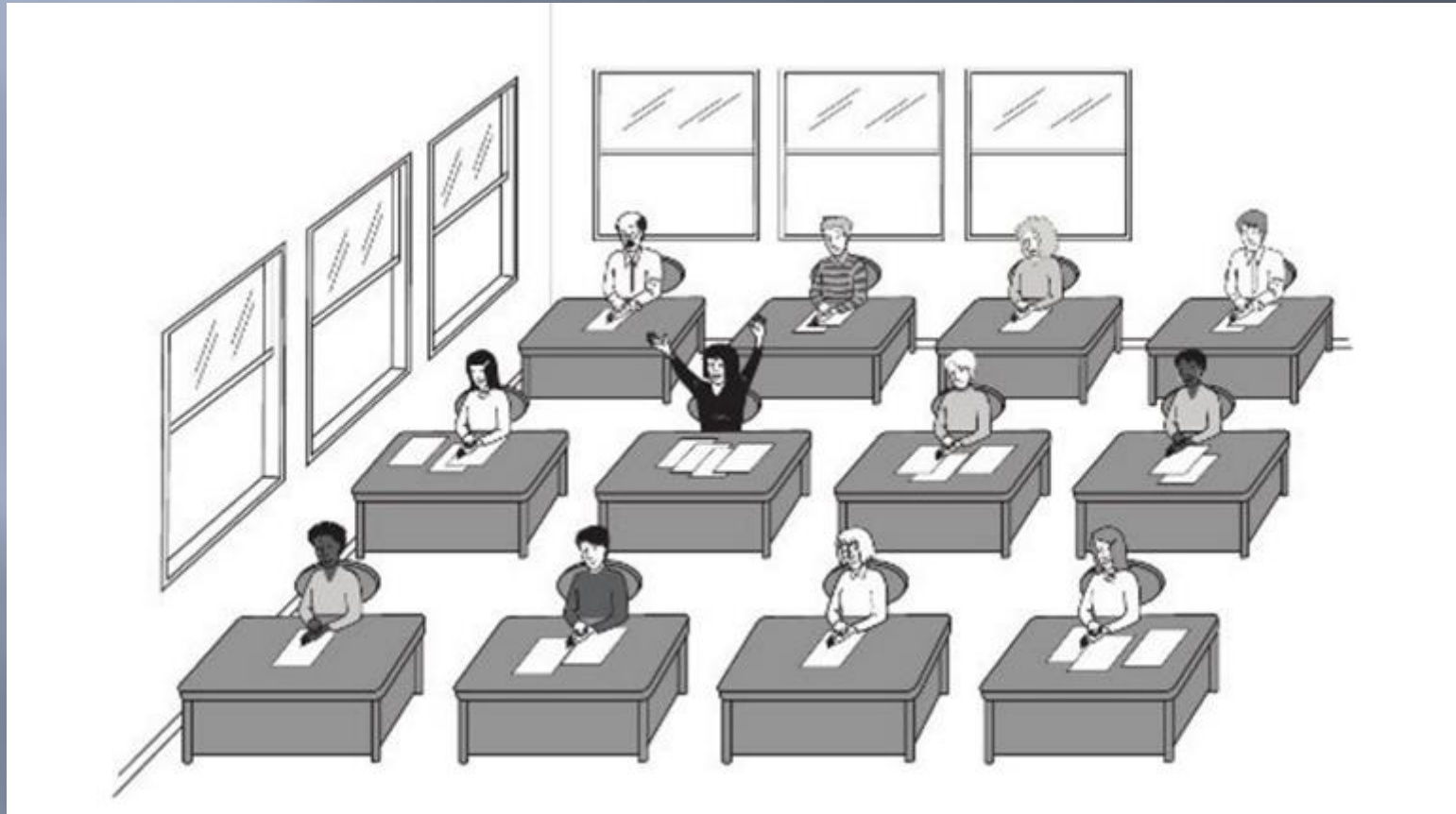
Self-organized criticality

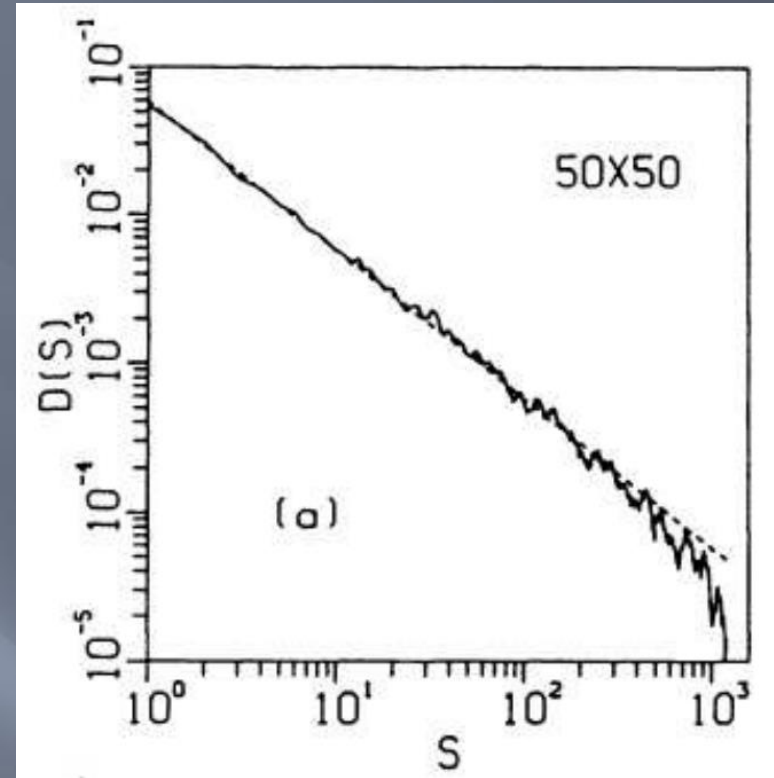
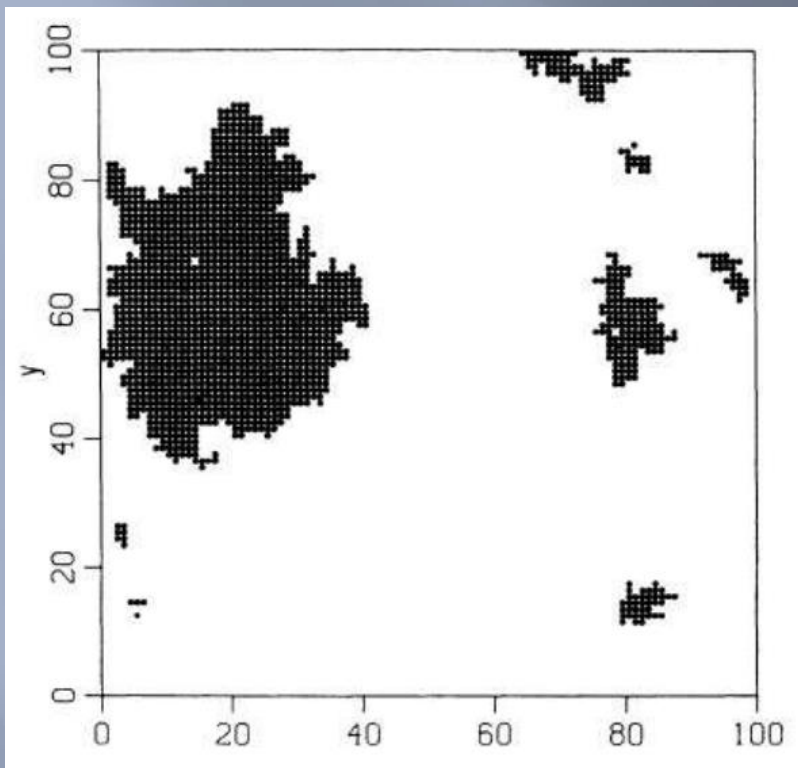


Not failure
but
avalanche
statistics
power law.

Non-
linearity
due to
feedback
causes
power law

SOC-model





Shapes of effected regions

Size distribution of
avalanches

Similar results for the Manna model, where 2 “grains”
are toppled irrespective of the dimension of the lattice

SOC

Simple mechanism leading to SOC: Let us play the Manna model in 1D with periodic BC (a ring of size L) and constant density ρ . Grains are conserved.

If $\rho \ll 1$ the system goes to an absorbing state: at stationarity nothing moves.

If $\rho > 1$ there is no absorbing state: There will always be sites with $z \geq z_c = 2$

It can be shown that there is a non-trivial $\rho_c < 1$ such that even for $\rho_c < \rho$ there is no absorbing stationary state in the TDL.

This is a continuous phase transition with the density of active sites (i.e., sites with $z \geq 2$) as order parameter.

What if we use open BC, i.e., grains can leave the system and we add grains if avalanches stop?

(Manna model)

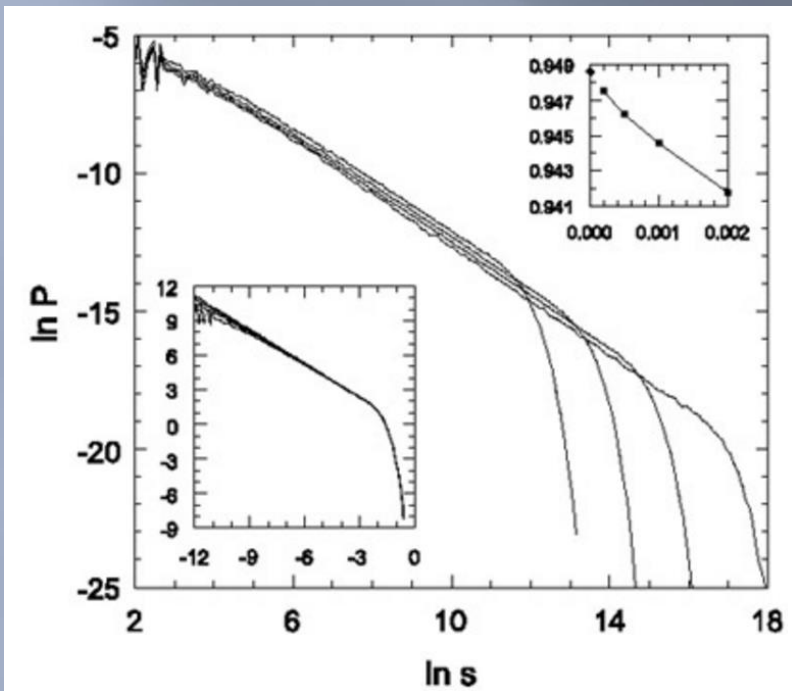
If $\rho < \rho_c$ the system is gets frozen. Then new grains are added.

If $\rho > \rho_c$ the activity persists until grains flow out of the system.

As a consequence, the system drives itself to the critical state.

There are 2 important elements:

- Conservation (grains do not disappear in the bulk)
- Separation of time scales (new grain comes if system is at rest)



$$\mathcal{P}(s; L) = as^{-\tau} \mathcal{G}\left(\frac{s}{bL^D}\right)$$

Finite size scaling

lattice	D	τ
simple chain	2.27(2)	1.117(8)
rope ladder	2.24(2)	1.108(9)
nnn chain	2.33(11)	1.14(4)
Futatsubishi	2.24(3)	1.105(14)
square	2.748(13)	1.272(3)
jagged	2.764(15)	1.276(4)
Archimedes	2.76(2)	1.275(6)
nc diagonal square	2.750(14)	1.273(4)
triangular	2.76(2)	1.275(5)
Kagomé	2.741(13)	1.270(4)
honeycomb	2.73(2)	1.268(6)
Mitsubishi	2.75(2)	1.273(6)
SC	3.38(2)	1.408(3)
BCC	3.36(2)	1.404(4)
BCCN	3.38(3)	1.408(4)
FCC	3.35(4)	1.402(8)
FCCN	3.38(4)	1.408(7)

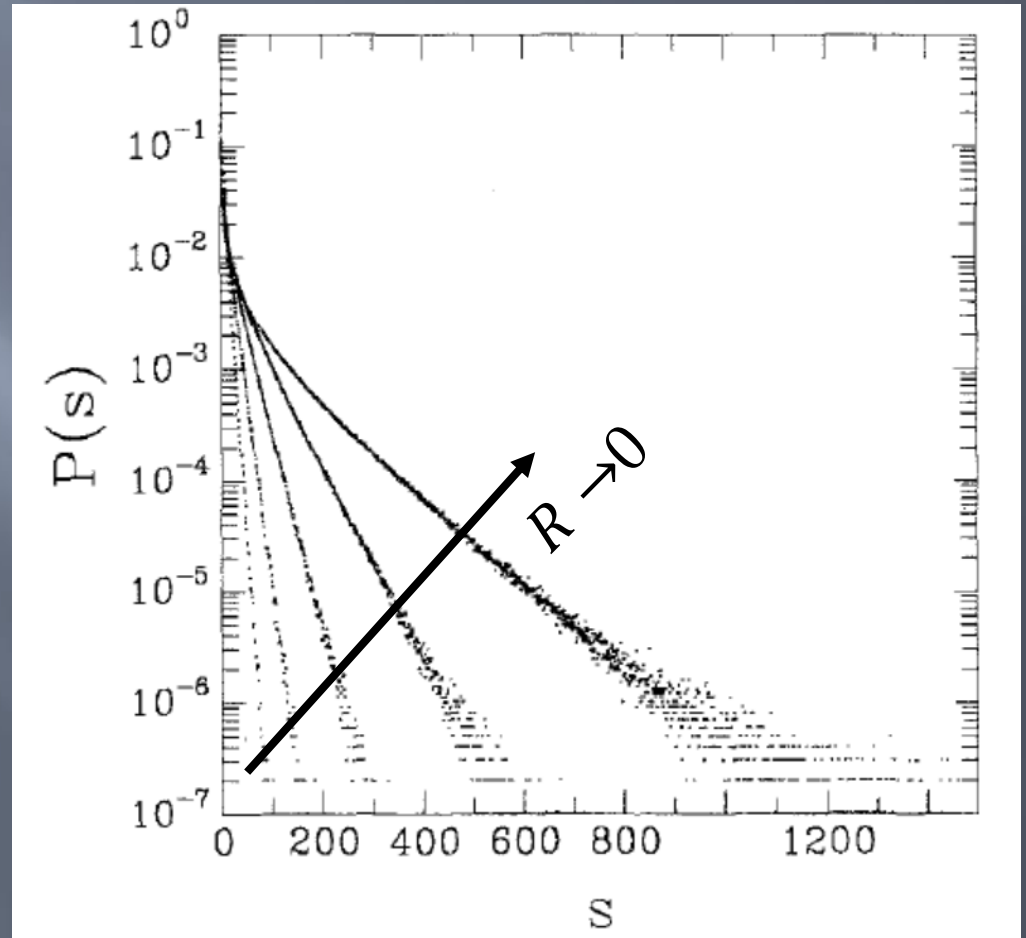
Role of conservation:

Let R be the prob. that during toppling a particle vanishes.

$$P(s) \sim s^{-\tau} \exp(-s/s^*)$$

With

$$s^* \sim R^{-\psi}$$



Cumulative distribution

SOC on networks

Generalization of BTW to uncorrelated scale free networks: node-dependent thresholds (= degree k)

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

Using theory of branching processes the SNU group could calculate the exponents

Mean field, independent of γ

γ -dependent exponents

	simul	theor
γ	τ_m	τ_t
∞	1.52(1)	1.50
5.0	1.52(3)	1.50
3.0*	1.66(2)	1.50
2.8	1.69(3)	1.56
2.6	1.75(4)	1.63
2.4	1.89(3)	1.71
2.2	1.95(9)	1.83
2.01	2.09(8)	2.0

Cascading Failures

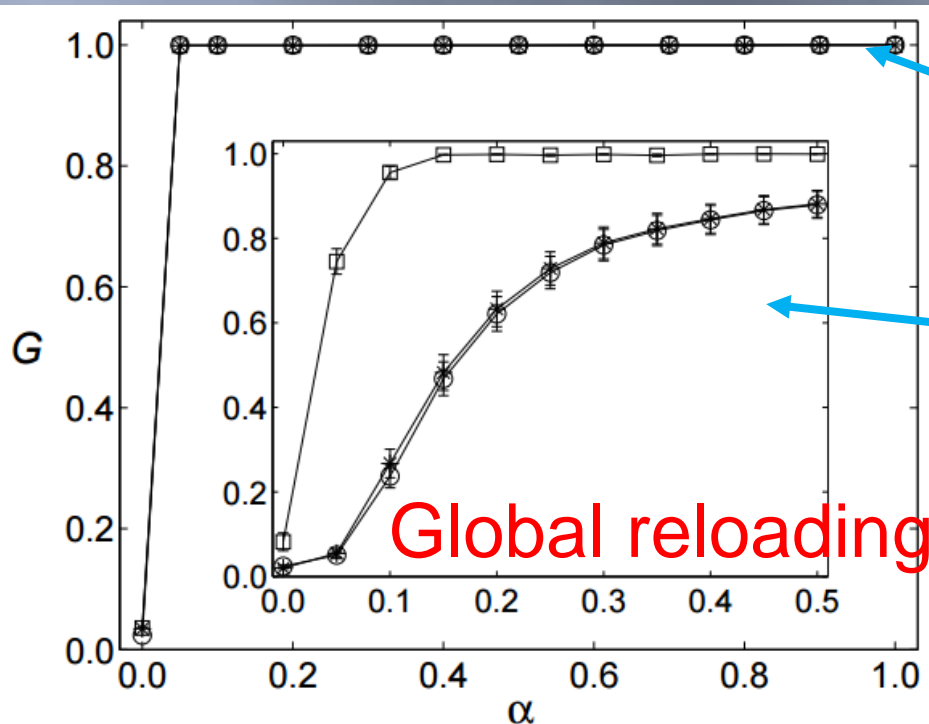
Start: loads $S = \text{betweenness} = \# \text{ shortest paths}$

Capacities: $(1+\alpha)S$

Initially the system works fine

After a failure, load has to be redistributed among the neighbors \rightarrow avalanche

Simulation of the model networks. G (not damaged part) vs α and the initiator. * max load, o: hub, \square : random



Regular network
 $k = 3$

Scale free nw
 $\langle k \rangle = 3, \gamma = 3$

Global reloading

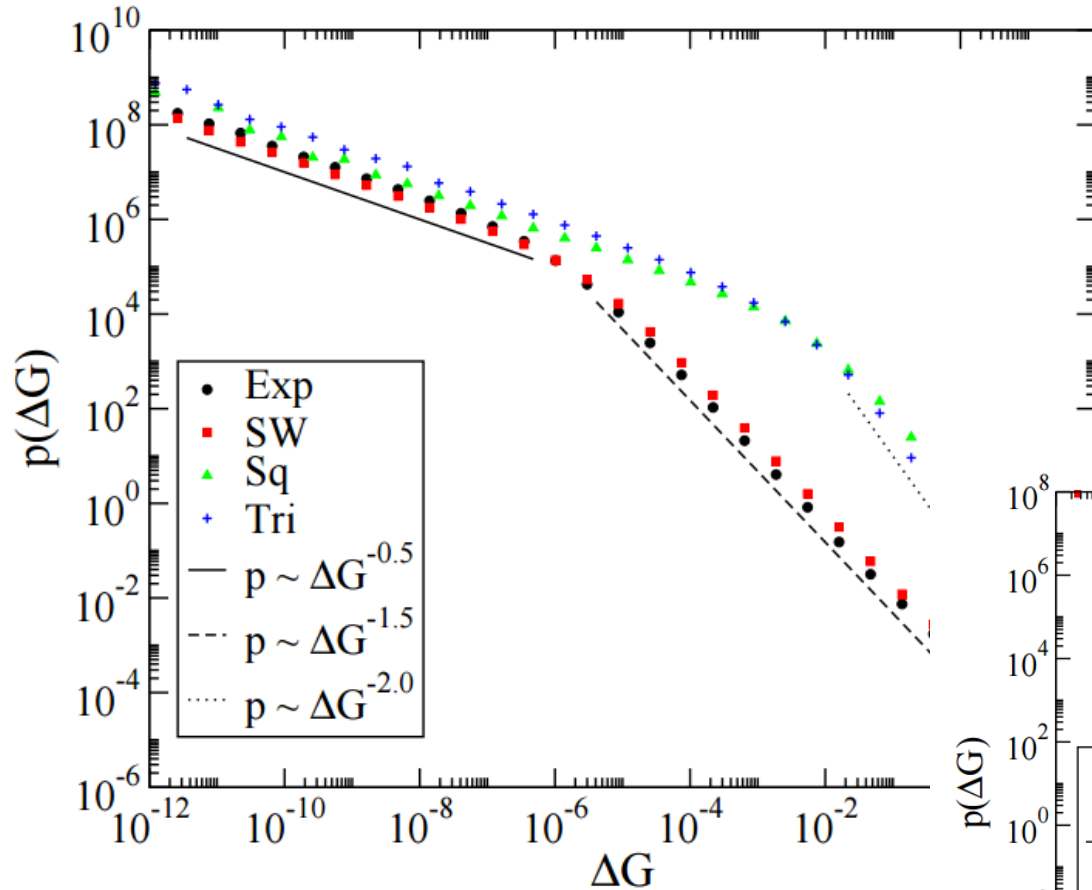
Blackouts

Load: el. current (random fuse model)

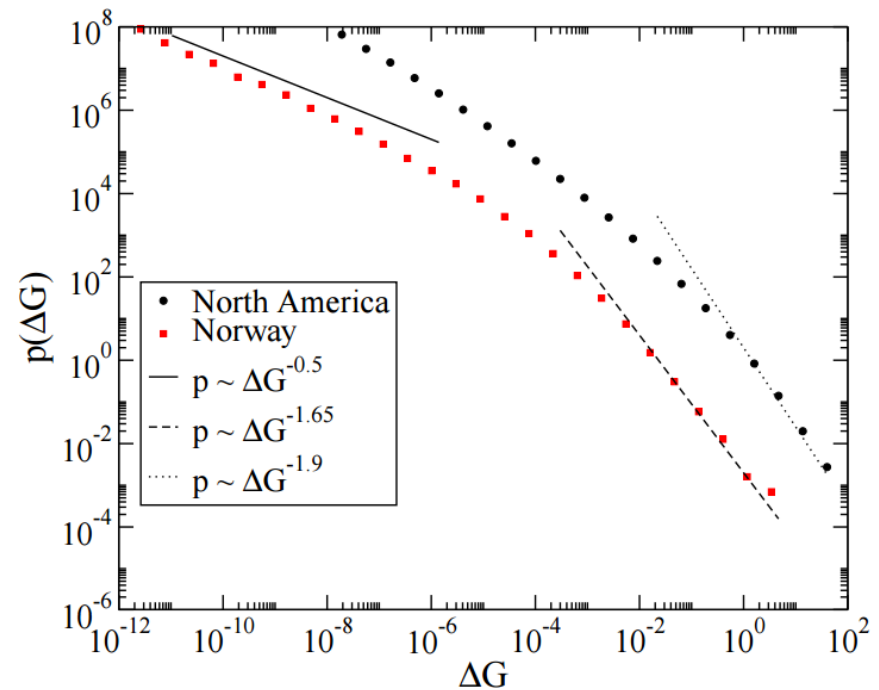
Consider a network (power grid) made up of resistors
Introduce current at a random point and a drain at another. Fix voltage.

1. Solve Kirchoff's eqs,
2. Set thresholds at $(1 + \alpha)i$
3. Remove a link at random,
4. Solve Kirchoff
5. Remove links where current exceeds threshold
6. Go to 3 until no more bad links found

Blackouts



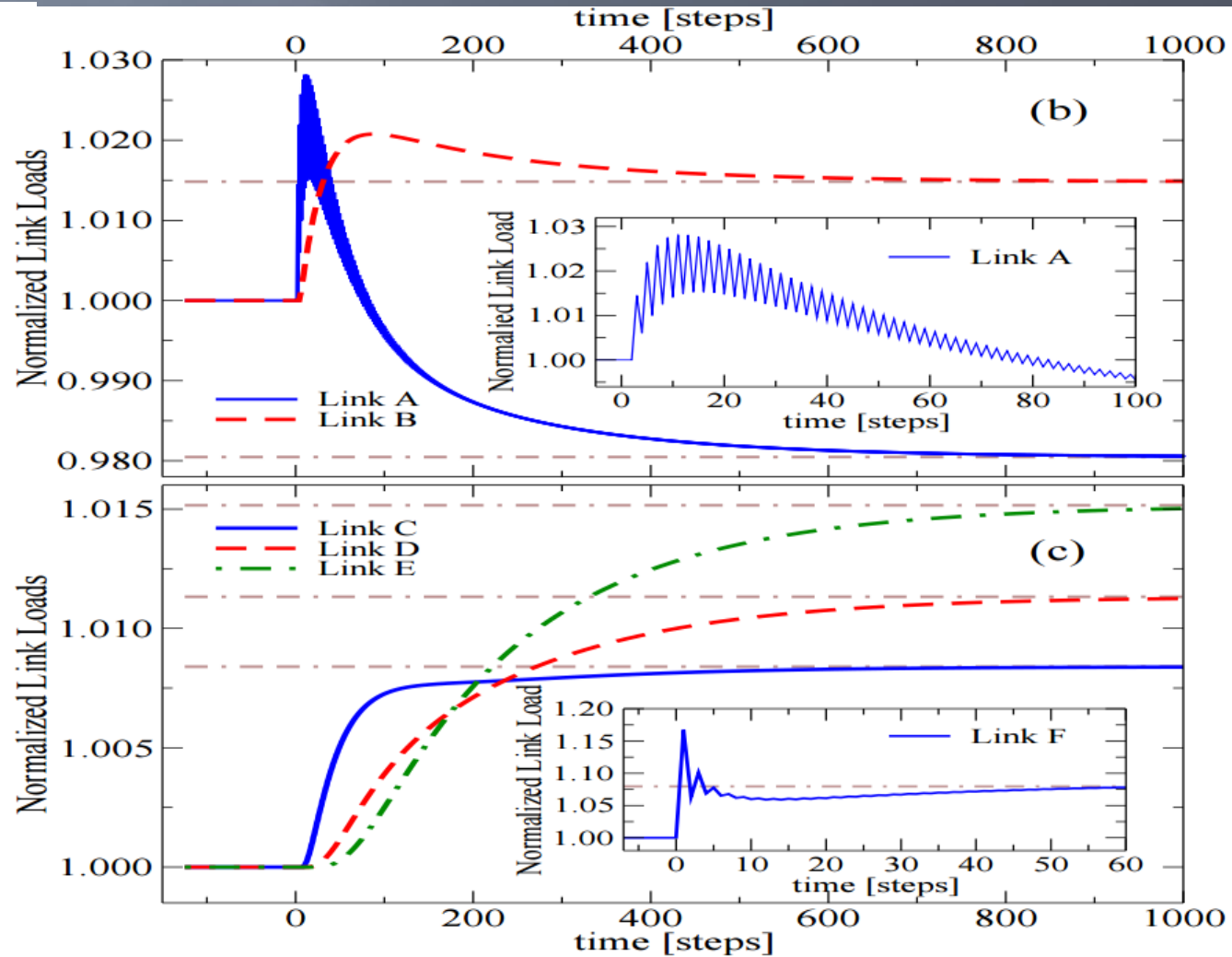
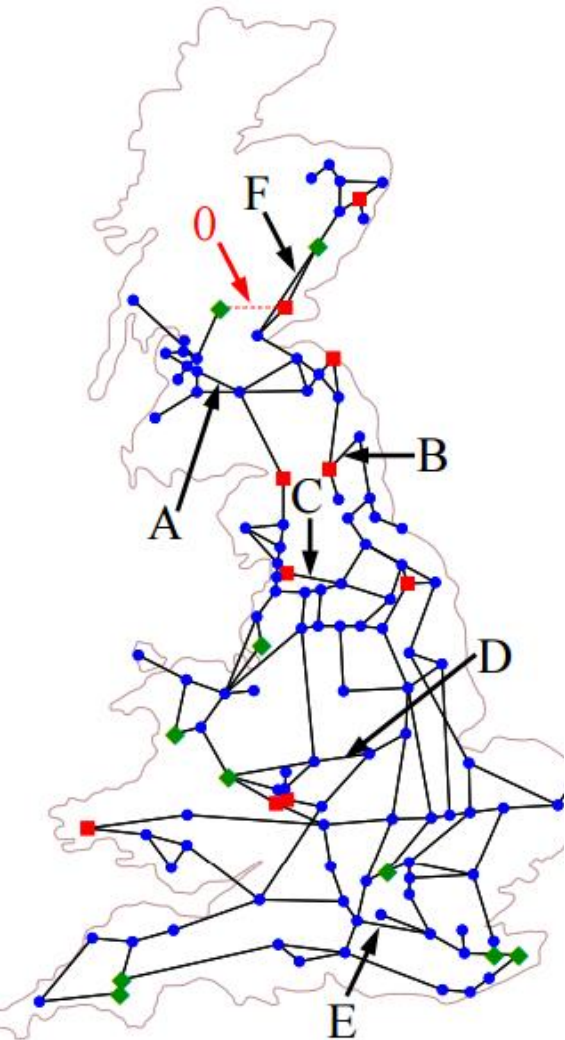
Pdf of $\Delta G =$ cond. loss



Same game on real data

Dynamic effect due to transients

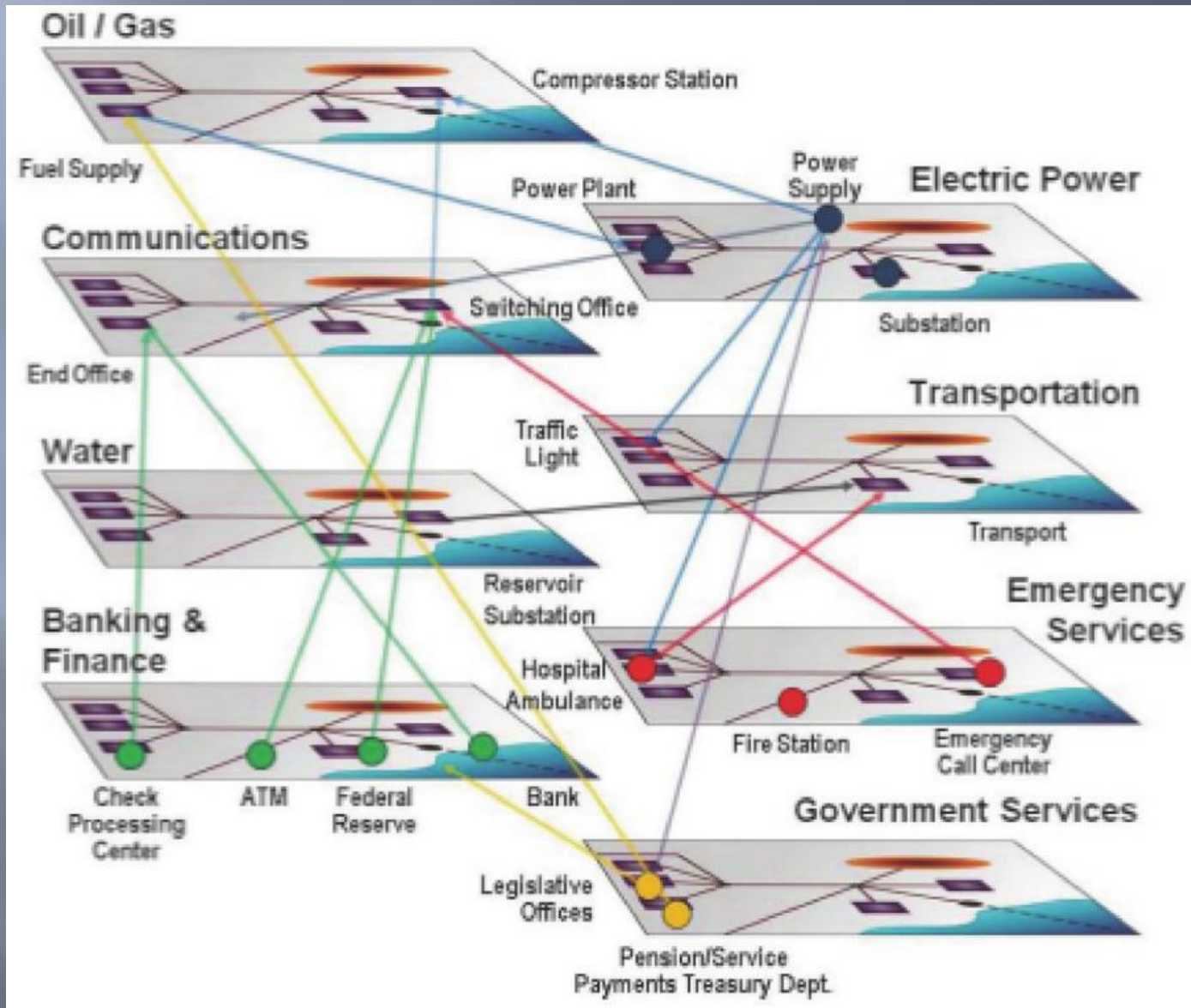
Britain: high-voltage power transmission grid



- Generator
- Sink

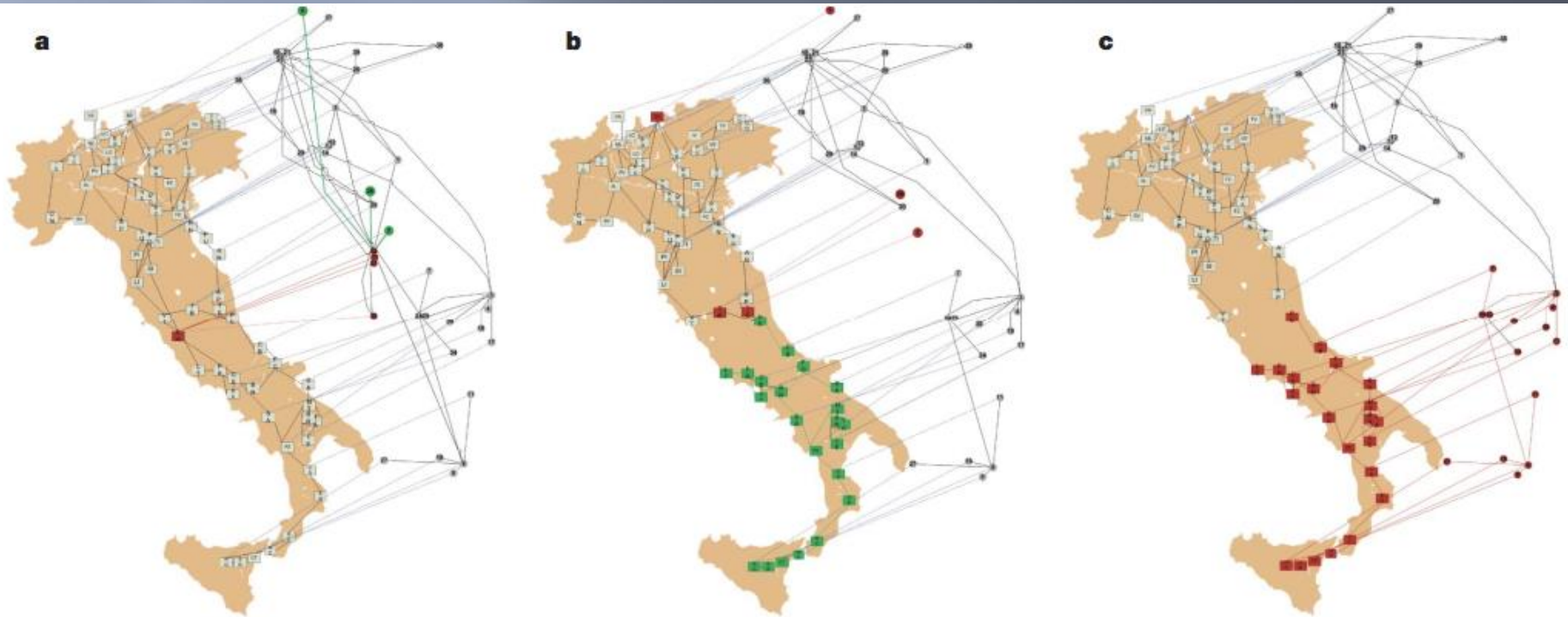
Cut at 0

Networks of Networks



Cascading Failures: Coupled Networks

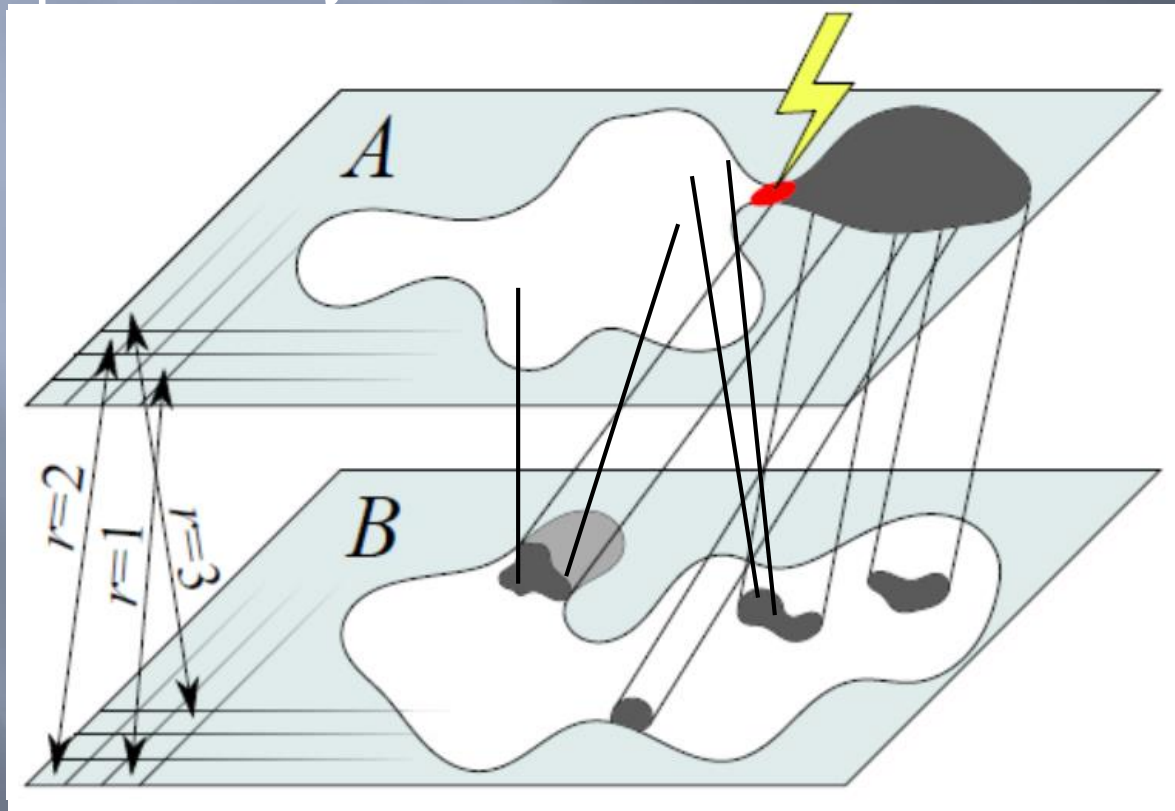
Networks are coupled: Energy supply, ICT, health service etc.



Modeling blackout in Italy: Interplay btw. Power grid and computer network needed for maintaining.

Cascading Failures

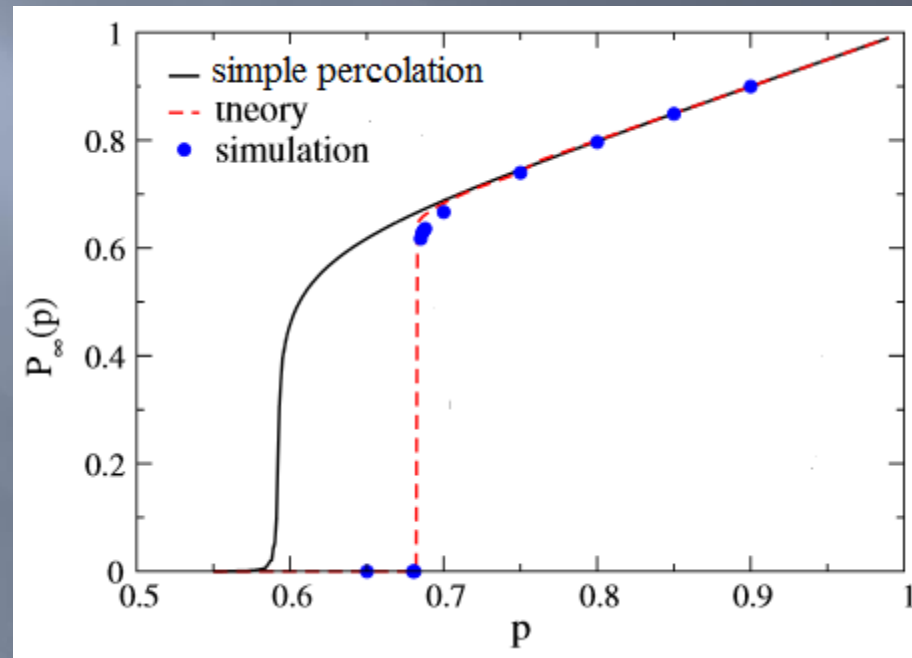
Coupled interdependent networks (Buldyrev et al. Nature 2010). Two kinds of links: Connection, dependency. The model:



Dependency link

Nodes not belonging to the giant component are sequentially removed. Analytical results for one-to-one dependency link.

Interdependent networks



Cascading failures on two interdependent square lattices. P_∞ is the relative size of the largest avalanche. For comparison, P_∞ for simple square lattice site percolation is also shown. First order transition: Cascading collapse

The threshold

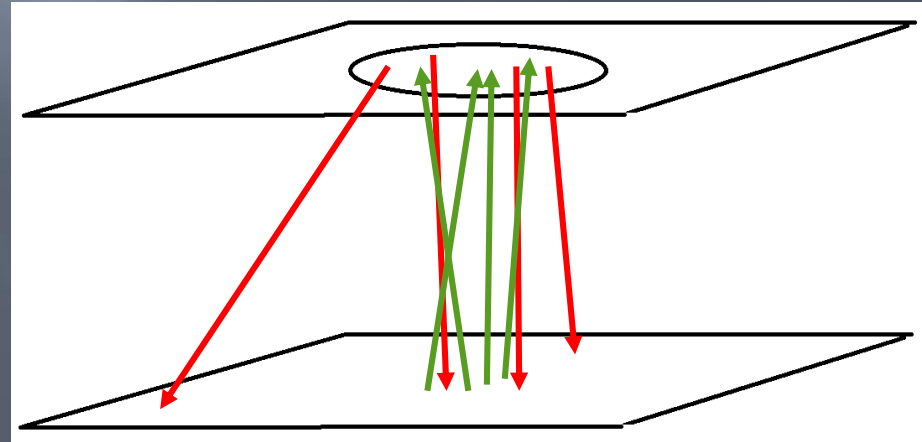
Let us consider a pair of square lattices of size N with random dependency links resulting in a one-to-one relationship. Let the initial dilution be $q = 1 - p$. Since we ignore everything but the giant component, the remaining number of nodes is $NP_\infty(p)$. Thus the other lattice will be diluted by $q_1 = 1 - p_1 = 1 - P_\infty(p)$. This is projected back to the first network.

The remaining density in the first step is $p_1 = P_\infty(p)$

In the second step we have to correct for the original dilution:

$p_2 = pP_\infty(p_1)$. Introducing $p_0 = p$ we have for the i -th

$$\text{step: } p_i = \frac{p_0}{p_1} \frac{p_1}{p_2} \frac{p_2}{p_3} \dots \frac{p_{i-2}}{p_{i-1}} P_\infty(p_{i-1}) = \frac{p_0}{p_{i-1}} P_\infty(p_{i-1})$$



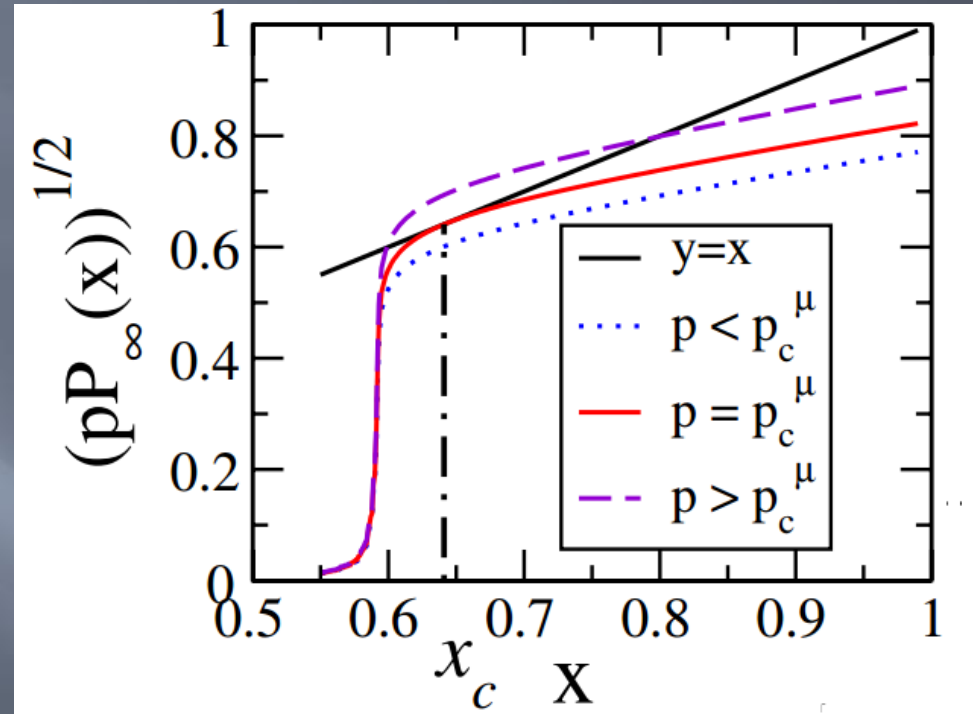
The threshold

$p_i = \frac{p}{p_{i-1}} P_\infty(p_{i-1})$ leading asymptotically to the equation $x = \sqrt{p P_\infty(x)}$. We know $P_\infty(x)$ only numerically but to high accuracy. Graphical solution:

$p_c(\text{normal}) = 0.59274$
 $p_c(\text{interdep}) = 0.6827$
 $P_\infty(p_c(\text{interdep})) = 0.602$

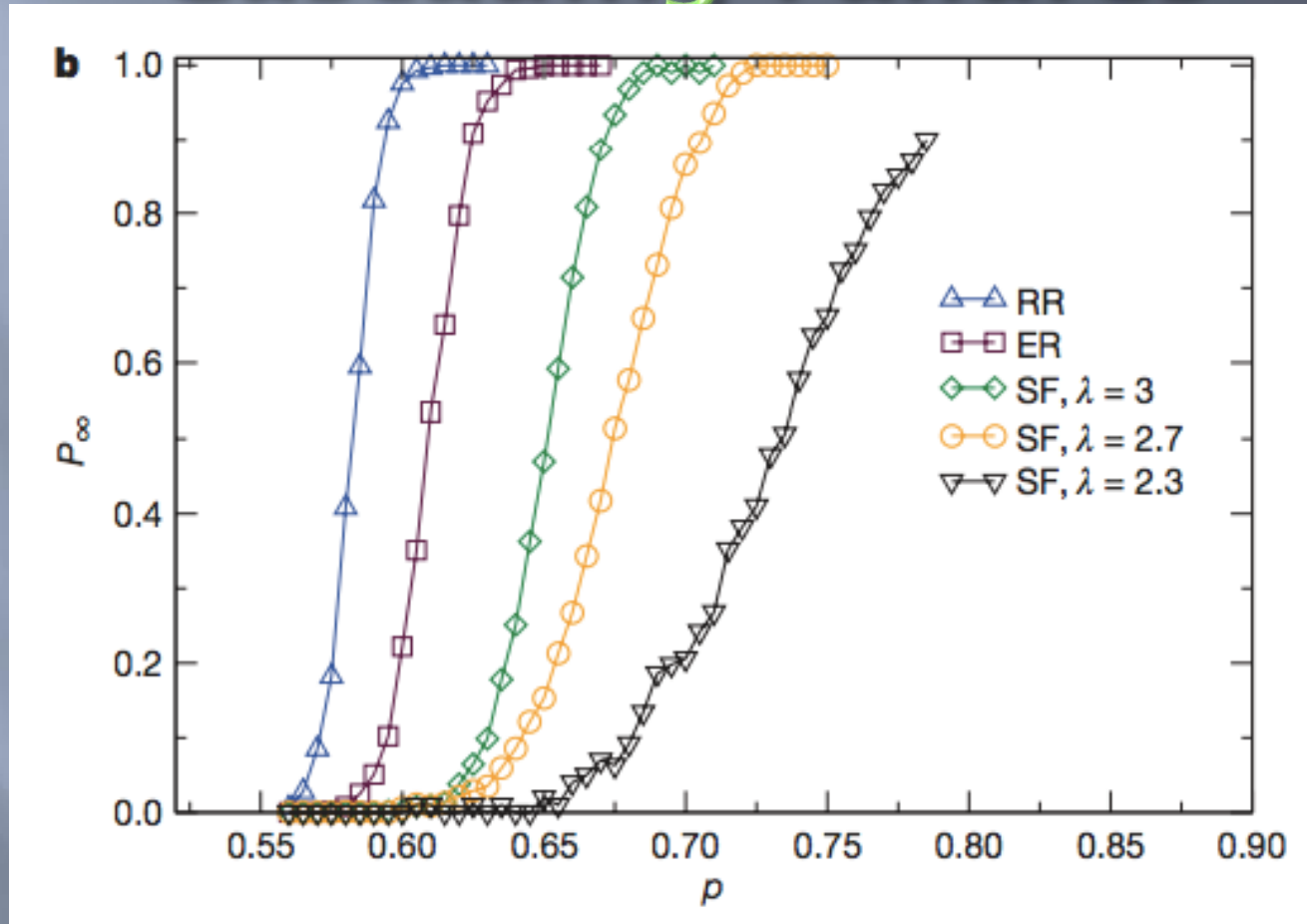
For ER network exact: $f = \exp[(f-1)/2f]$

$$p_c = [2\langle k \rangle f(1-f)]^{-1} = 2.4554/\langle k \rangle$$



The last point with solution.

Cascading Failures



Result: Enormous sensitivity to the ratio p of initially removed nodes: “first order” transition, jump in P_∞ for $N \rightarrow \infty$. Note that transition is there for SF! It is getting first order for $N \rightarrow \infty$.

Hybrid phase transition

Really first order?

Indeed, the order parameter (P_∞) is discontinuous.

But: There is scaling at p_c !

$$P_\infty = \begin{cases} 0 & \text{if } p < p_c \\ m_0 + r(p - p_c)^\beta & \text{if } p \geq p_c \end{cases}$$

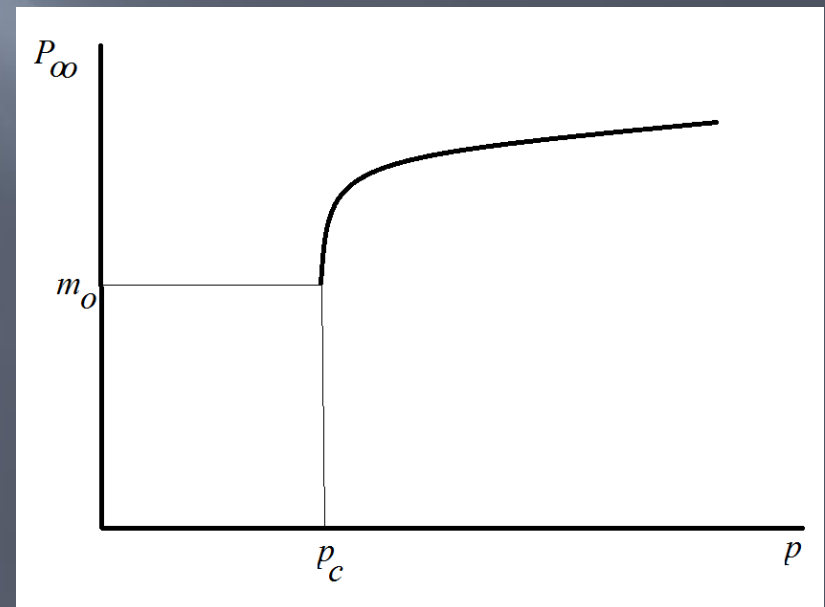
There is “nothing” on the other side of the transition:

Absorbing phase

Two (coupled) critical phenomena

- Order parameter
- Avalanches

Accordingly: Two sets of exponents



Hybrid phase transition

One is (as usual) related to the order parameter:

β describes how it changes at p_c , γ is the exponent of the fluctuations of P_∞ and ν can be defined through finite size scaling.

In ordinary percolation an equivalent formulation can be given through the scaling of clusters.

Here not! All finite clusters are of size 1 or 2.

But the avalanches scale near p_c . Two divergent scales!

Simulations can be carried out using the efficient method by the Kahng group (SNU).

		β	τ	σ	γ	$\bar{\nu}$
ordinary ER		1	1.5	0.5	1	3
interdependent ER	m	0.5 ± 0.01	-	-	1.05 ± 0.05	2.1 ± 0.02
	a	-	1.5 ± 0.01	1.0 ± 0.01	0.5 ± 0.01	1.85 ± 0.02
ordinary 2D		0.139	1.055	0.286	2.389	2.667
interdependent 2D	m	0.53 ± 0.02	-	-	1.35 ± 0.10	2.2 ± 0.20
	a	-	1.59 ± 0.02	0.70 ± 0.05	0.5 ± 0.05	2.1 ± 0.2

Hybrid phase transition

Usual relationships like $\beta = \frac{\tau-1}{\sigma}$ or $\gamma = \frac{2-\tau}{\sigma}$ do not hold.

However, exponents are not unrelated!

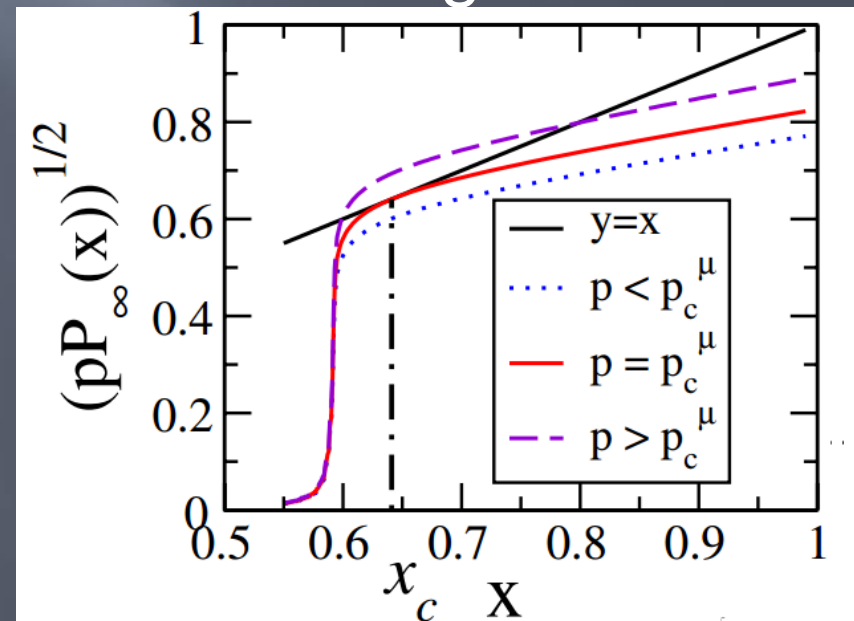
Sum rule: A site can either belong to the giant mutually connected component or it has been eliminated by one of the avalanches:

$1 = P_\infty(p) + \int_p^1 \sum_s s p_s(p') dp'$ where

$p_s(p)$ is the number of avalanche of size s generated between $(p, p + dp)$.

This leads to $\frac{2-\tau}{\sigma} = 1 - \beta$

Furthermore, $\beta = 1/2$ as it is clear from the determination of the critical point.



Hybrid percolation transition

Many examples: k-core percolation, extended epidemic process, threshold model, etc.

Mechanism of hybrid percolation transition (ER):

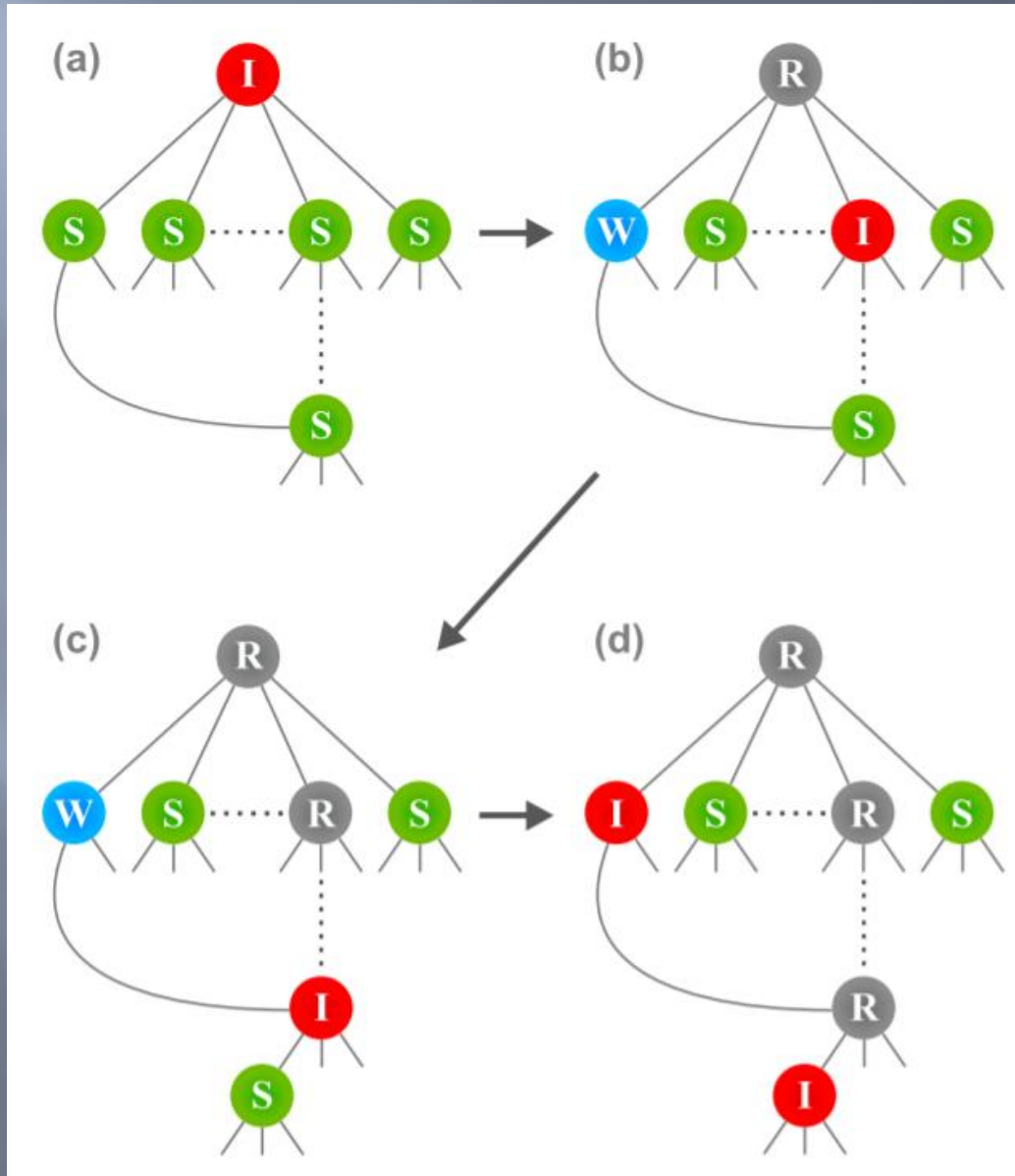
Ordinary percolation at criticality: Critical branching process.

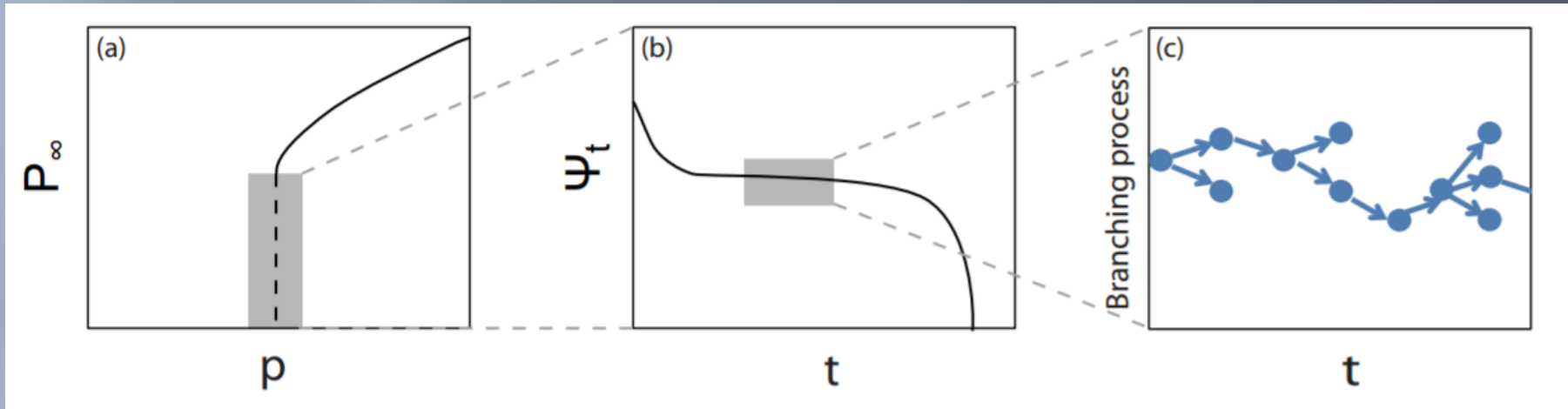
Here 2 states after contact with “infected” node: Infected or weakened.

The infected constitute a critical branching tree.

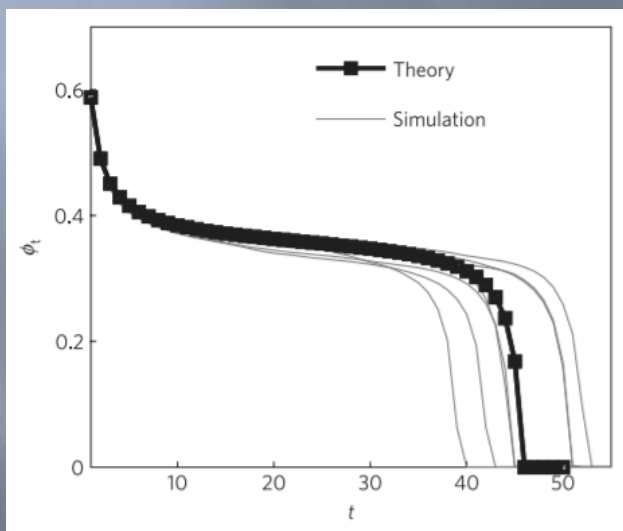
Due to finiteness of the sample, there is a characteristic time ($O(N^{1/3})$) that a weakened node meets an infected \rightarrow this pushes the process into the supercritical regime \rightarrow breakdown

Hybrid percolation transition



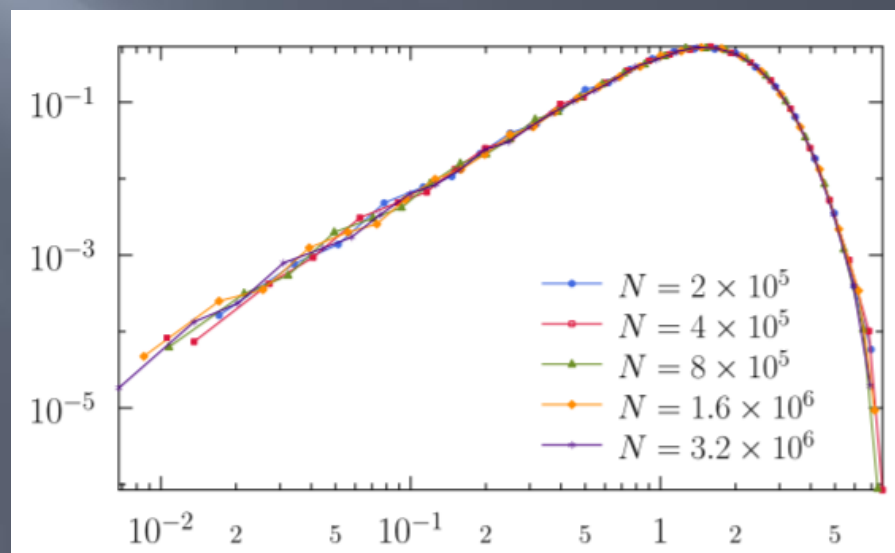


Zhou et al 2014



Average and sample to sample var. of the evolution of the OP

$$P_t N^{1/3}$$



Prob. that large loop at t

Lee et al. 2017

Stopping spreading

Spreading can be good (innovations) or bad (disease, cascading failures). If bad, we want to hinder.

Simple epidemic models: spreading depends largely on the underlying network. E.g., SIR model epidemic threshold $\frac{\mu}{\beta} < \kappa - 1$, with $\kappa = \frac{\langle k^2 \rangle}{\langle k \rangle}$ inhomogeneity ratio

For scale free networks with $\gamma \leq 3$ always spreading!
Hubs are good spreaders.

Immunization: Random vaccination does not work.

Vaccinate the hubs!

How to know? Needs global info.

As prob. of having a neighbor with degree k is $k p_k / z$ (my friends are more popular than I am): Vaccinate the contact persons of randomly selected ones.

Stopping cascades

Idea: Cut some links before failure escalates.

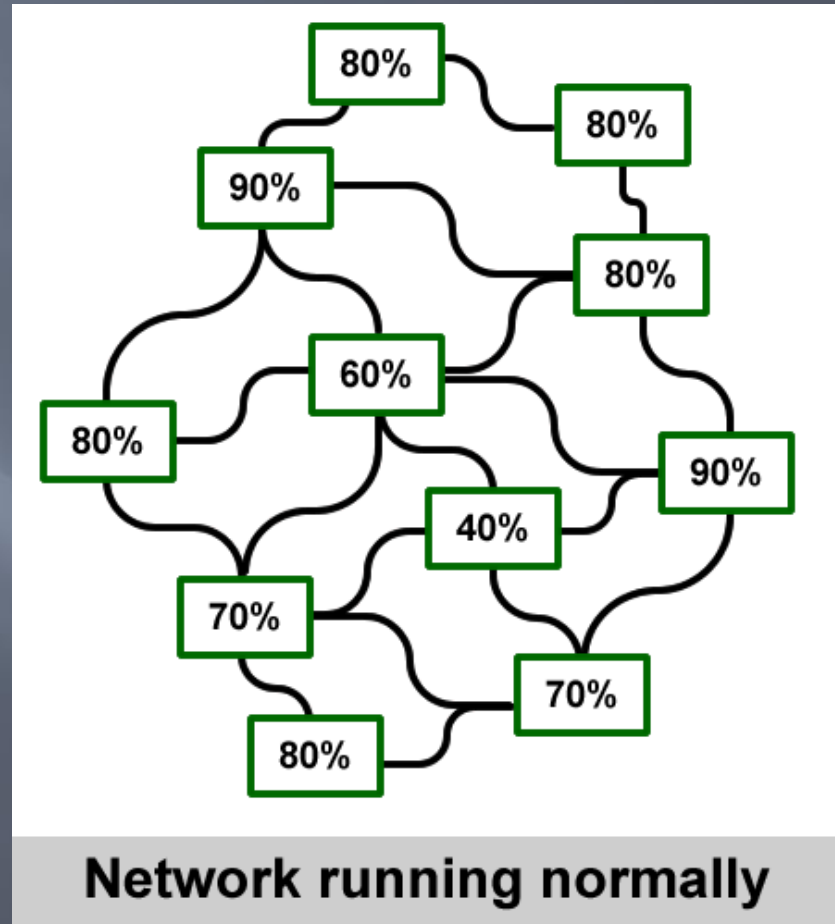
Model: Load=betweenness centrality

Network: Scale free

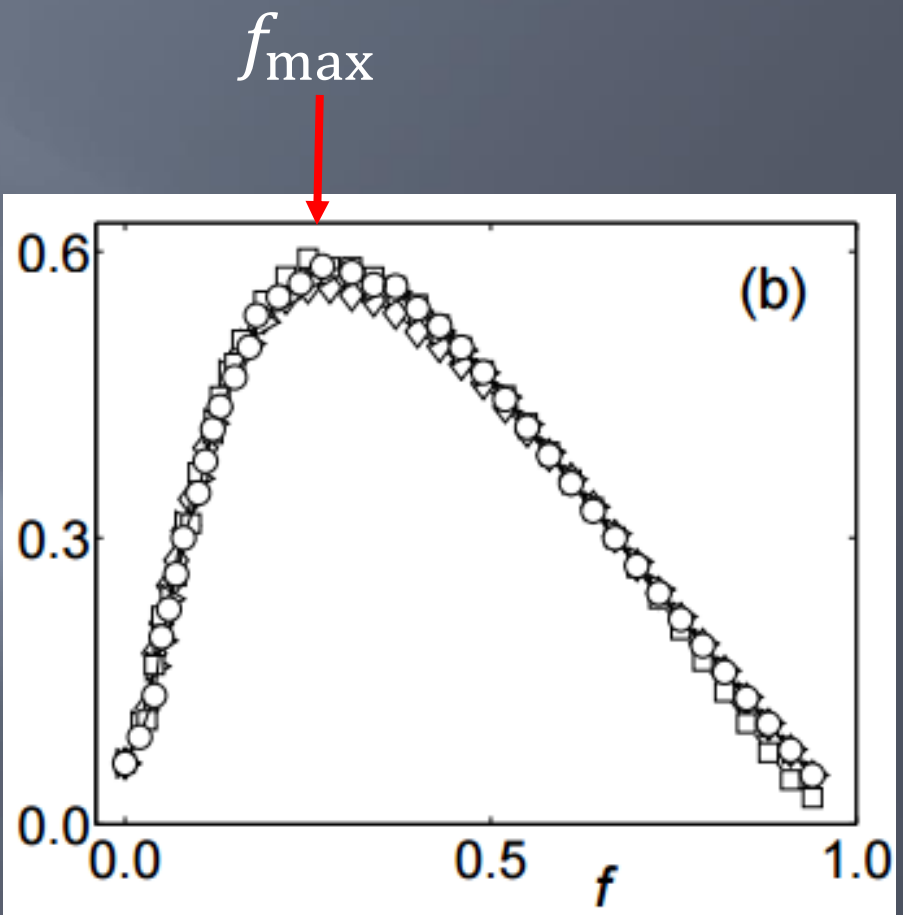
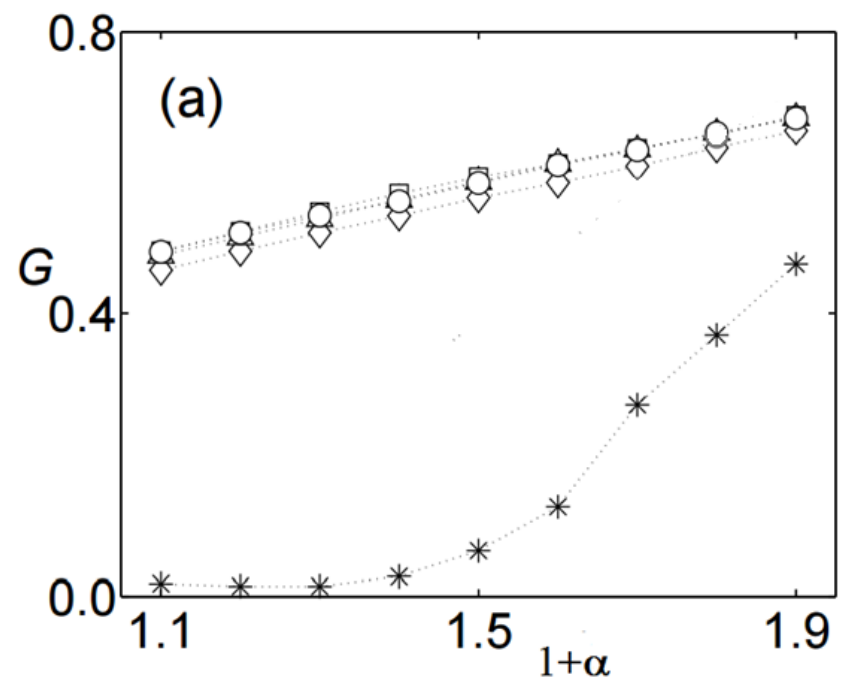
4 suggestions:

After first step in the cascade eliminate ratio f of nodes, where

1. Difference between carried and generated load is minimum
2. Closeness centrality is min. (periphery nodes)
3. Load is smallest
4. Degree is smallest



G : Remaining fraction

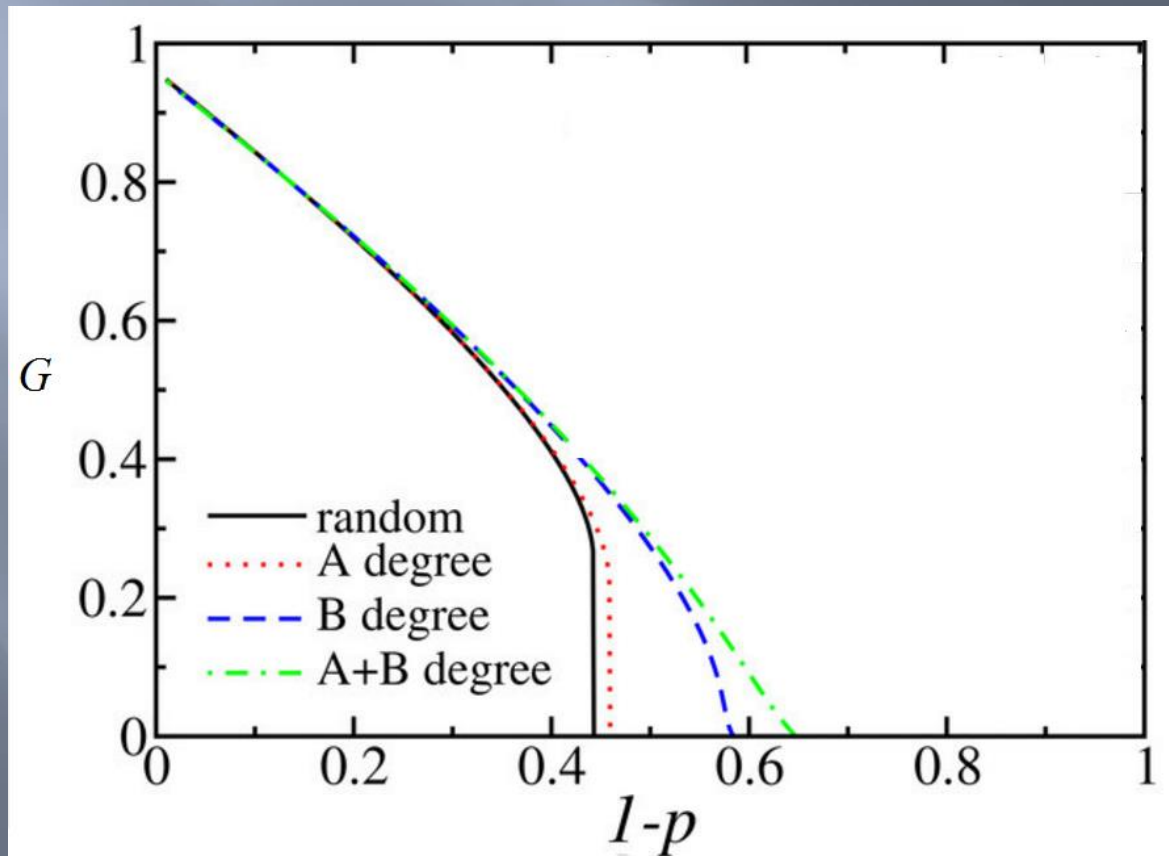


a) * no defense; open circle, square, triangle, diamond correspond to strategies 1,2,3,4, respectively at f_{\max}
b) Dependence on f

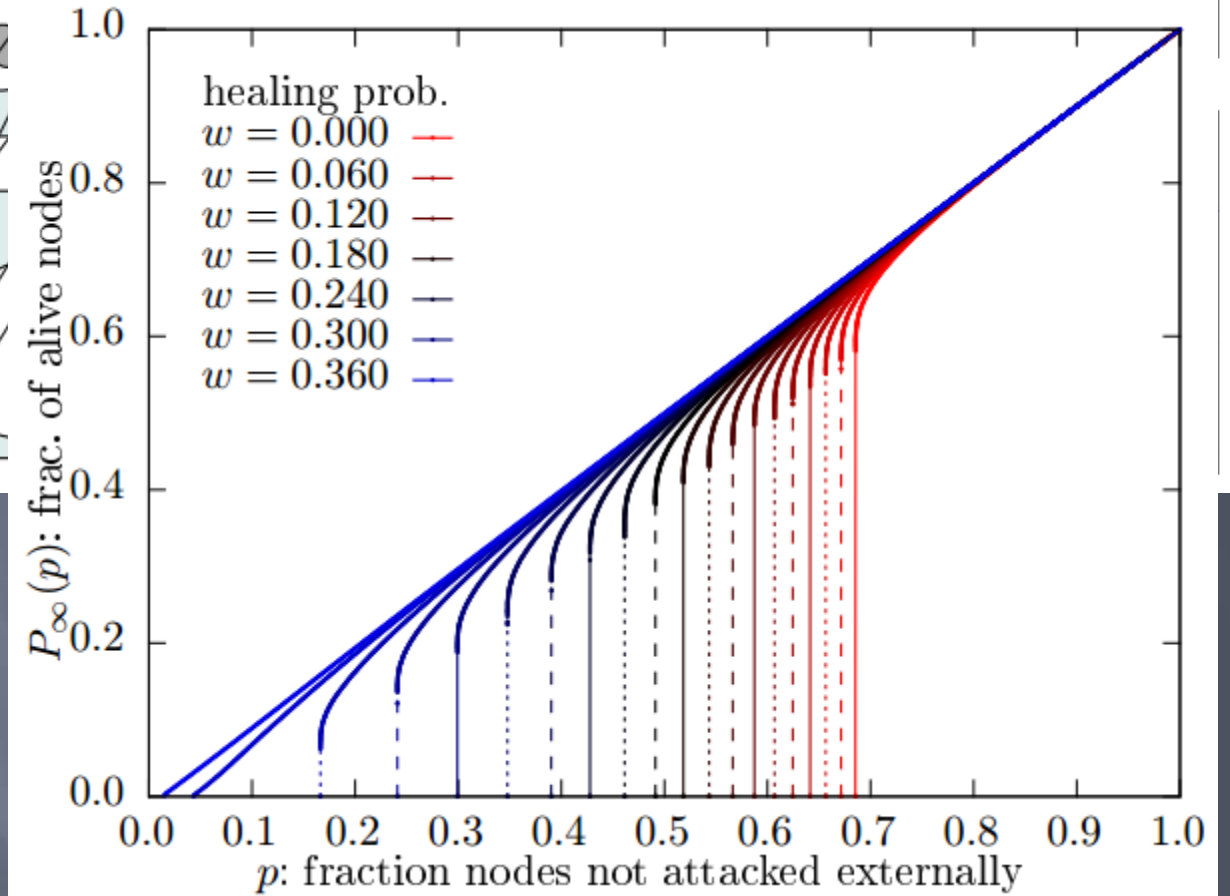
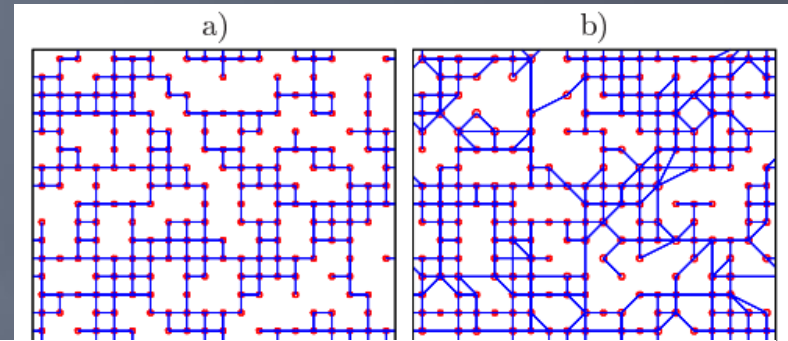
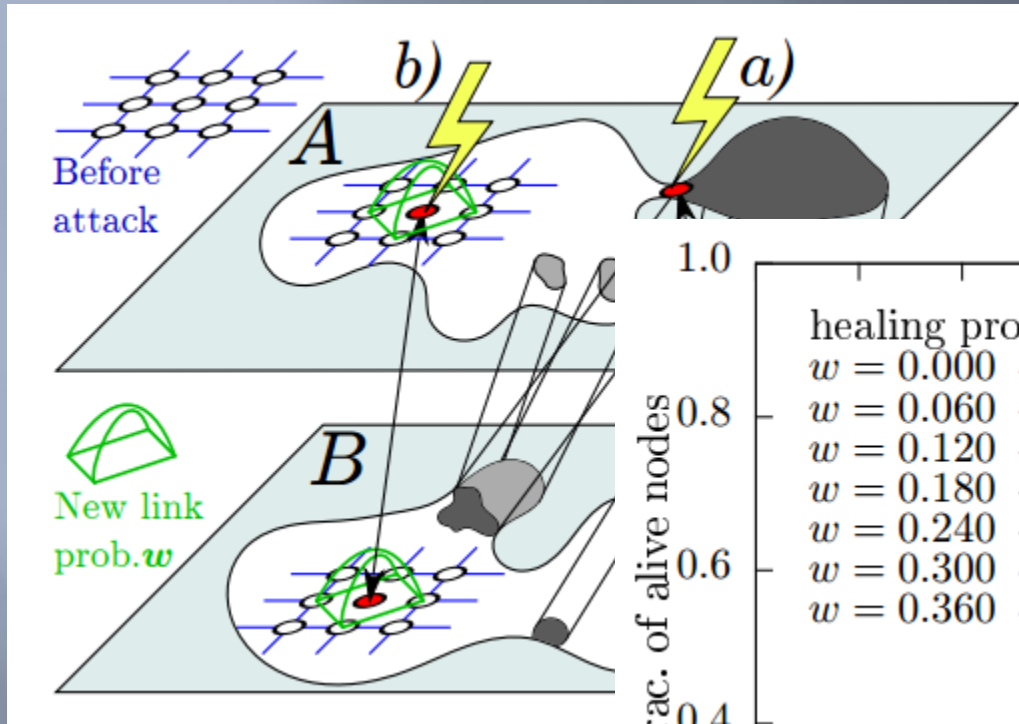
Reducing the effect for interdependent networks

Different strategies:

1. Strengthen high degree nodes.



2. Healing



There is a critical healing prob. above which second order transition

Summary

- Random failures can be treated within percolation theory. Scale free networks are extremely resilient against random failures
- If hubs are attacked intentionally, scale free networks become vulnerable
- Cascading failures may lead to catastrophic breakdown even in scale free networks

Homework

Take the configuration with $N = 10000$ nodes and degree distribution

$$P(k) = \frac{k^{-\gamma}}{\zeta(\gamma)}, \text{ where } \gamma = 2.2 \text{ and } \gamma = 2.9,$$

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

Make a statistics about the systems' threshold of the collapse of the giant component with

a) random failures (random node removal)

c) intentional attacks (removal according to the degree sequence, starting with highest)

Compare the systems with $\gamma = 2.2$ and $\gamma = 2.9$