New Complex Networks for Social Relations

Hans J. Herrmann

Computational Physics, IfB
ETH, Zürich, Switzerland

Marta C. González
José Soares Andrade Jr.
Pedro Lind
Herman Singer

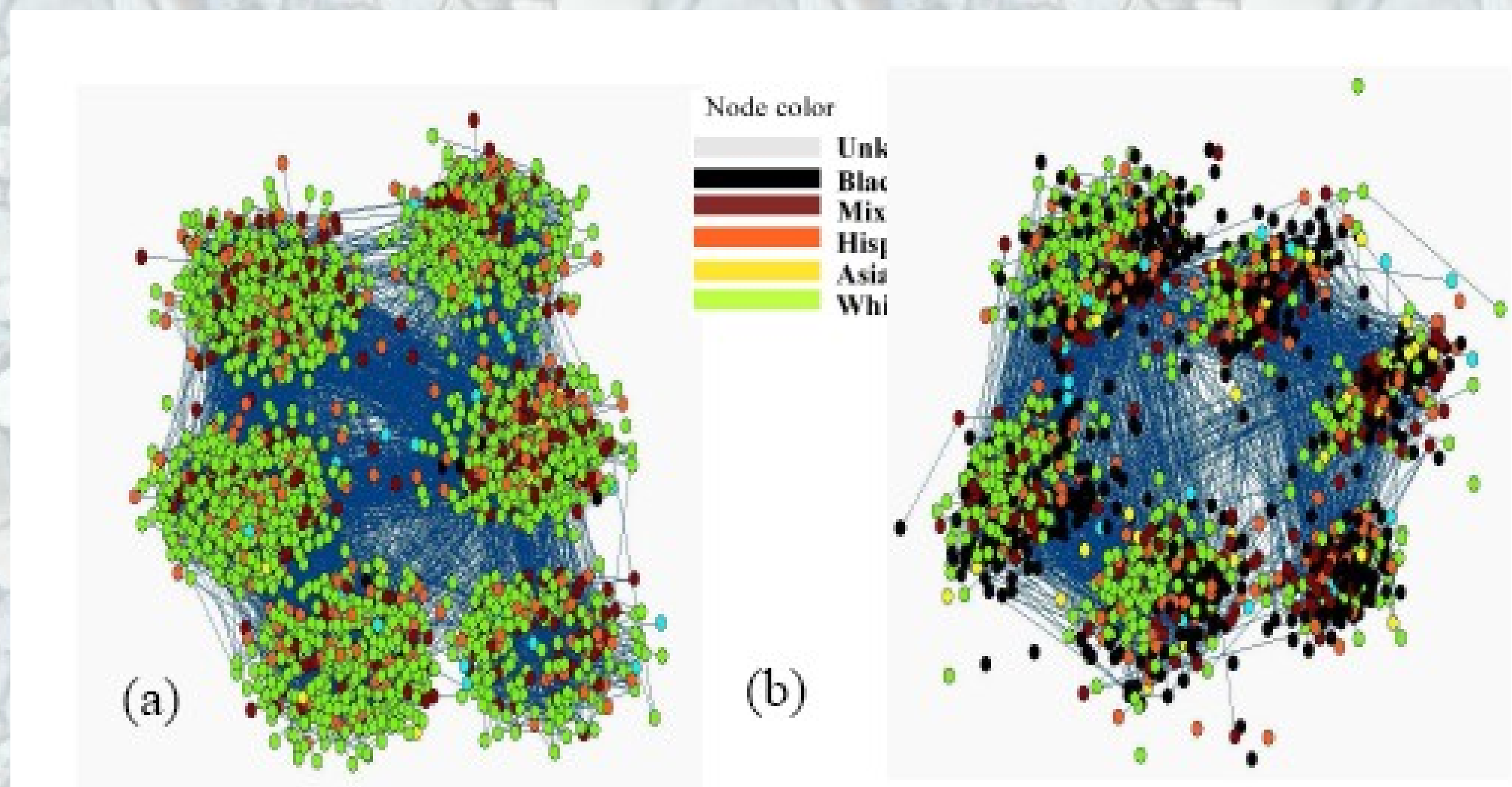
International Workshop on Challenges and Visions in the Social Sciences
ETH Zürich
August 18-23, 2008

Motivation

- The statistical spreading of information or contacts in societies has many practical implications. It involves non-equilibrium phenomena where fluctuations and correlations play an important role.
- Usually these processes are studied by sociologists using surveys including many parameters and obtaining qualitative results. Recently techniques from statistical physics and in particular complex networks and non-linear dynamics have been used to make simplified models and obtain quantitative laws.

Data from schools

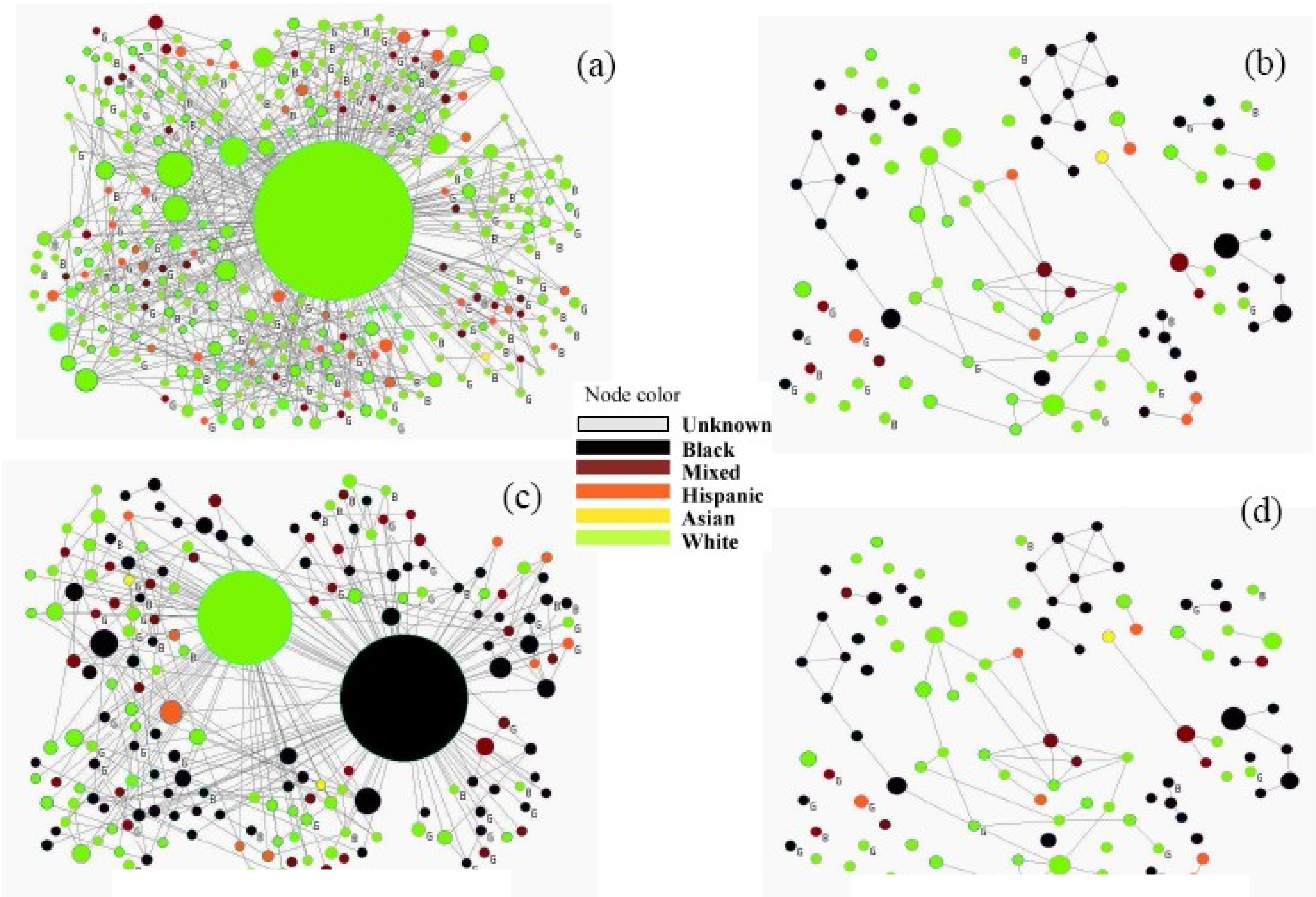
Survey interviewing 90118 student from 84 schools in US
(Add Health Program)



visualization using „pajec“

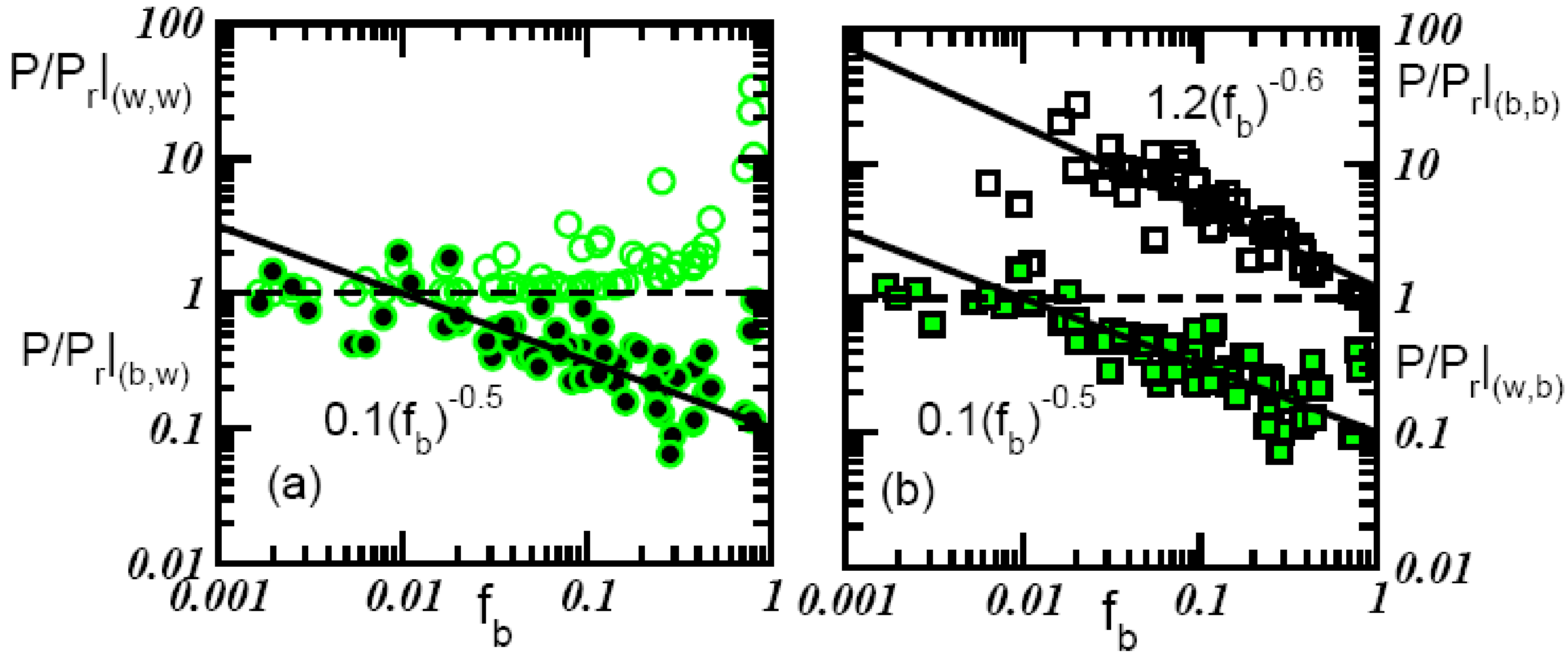
with T. Vicsek and J. Kertész

Schools



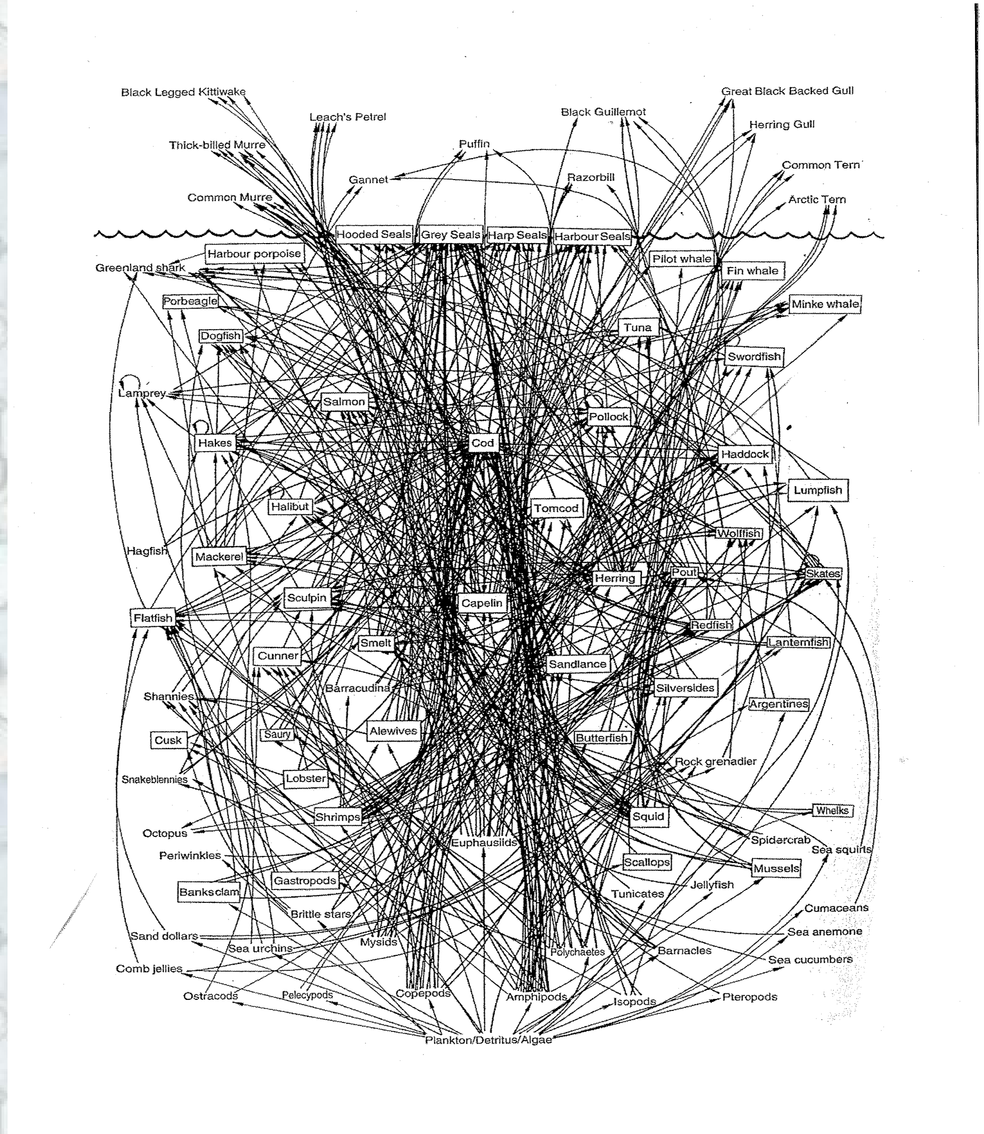
3-cliques **4-cliques**

Schools



affinities between black and white students

Complex Networks



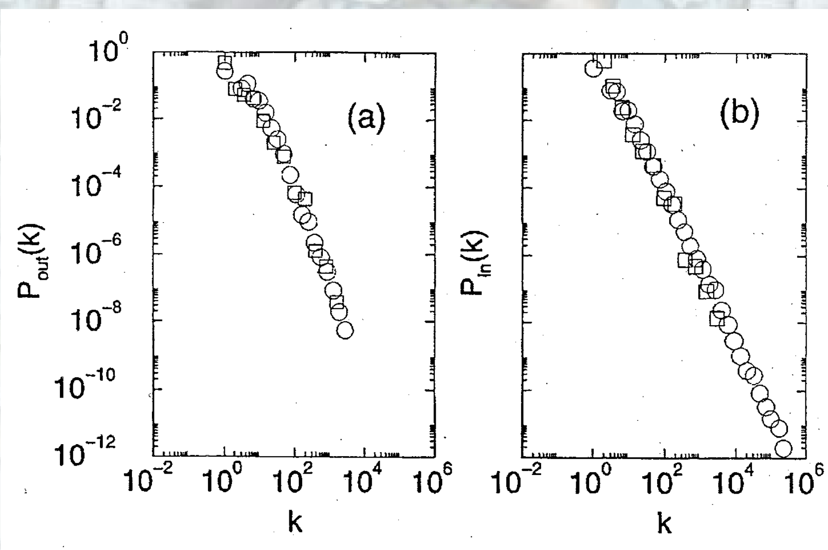
Food web of the North Atlantic Ocean

Scale-free networks

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

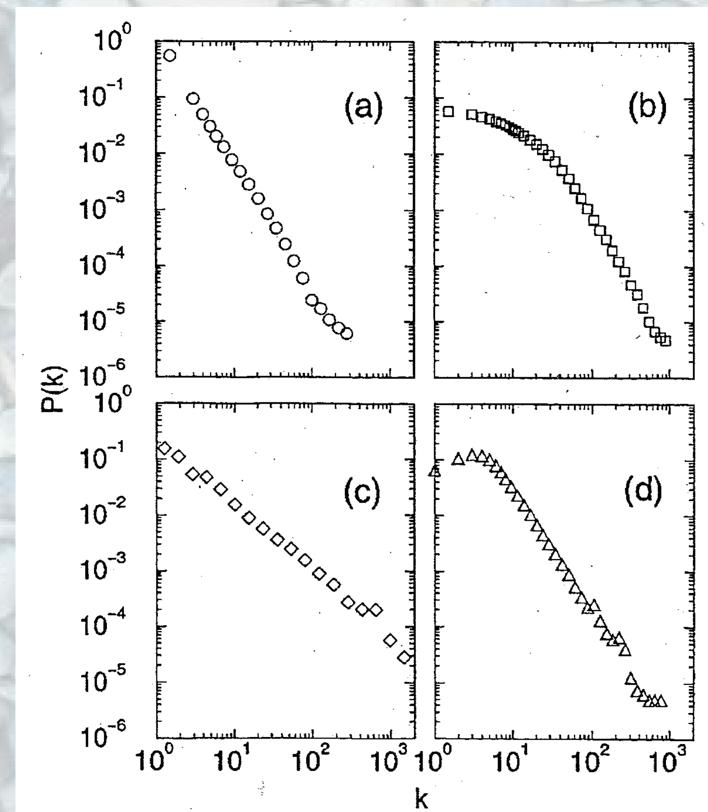
WWW:

$\gamma_{out} = 2.4$ $\gamma_{in} = 2.1$



Internet $\gamma = 2.4$

actors $\gamma = 2.3$



HEP

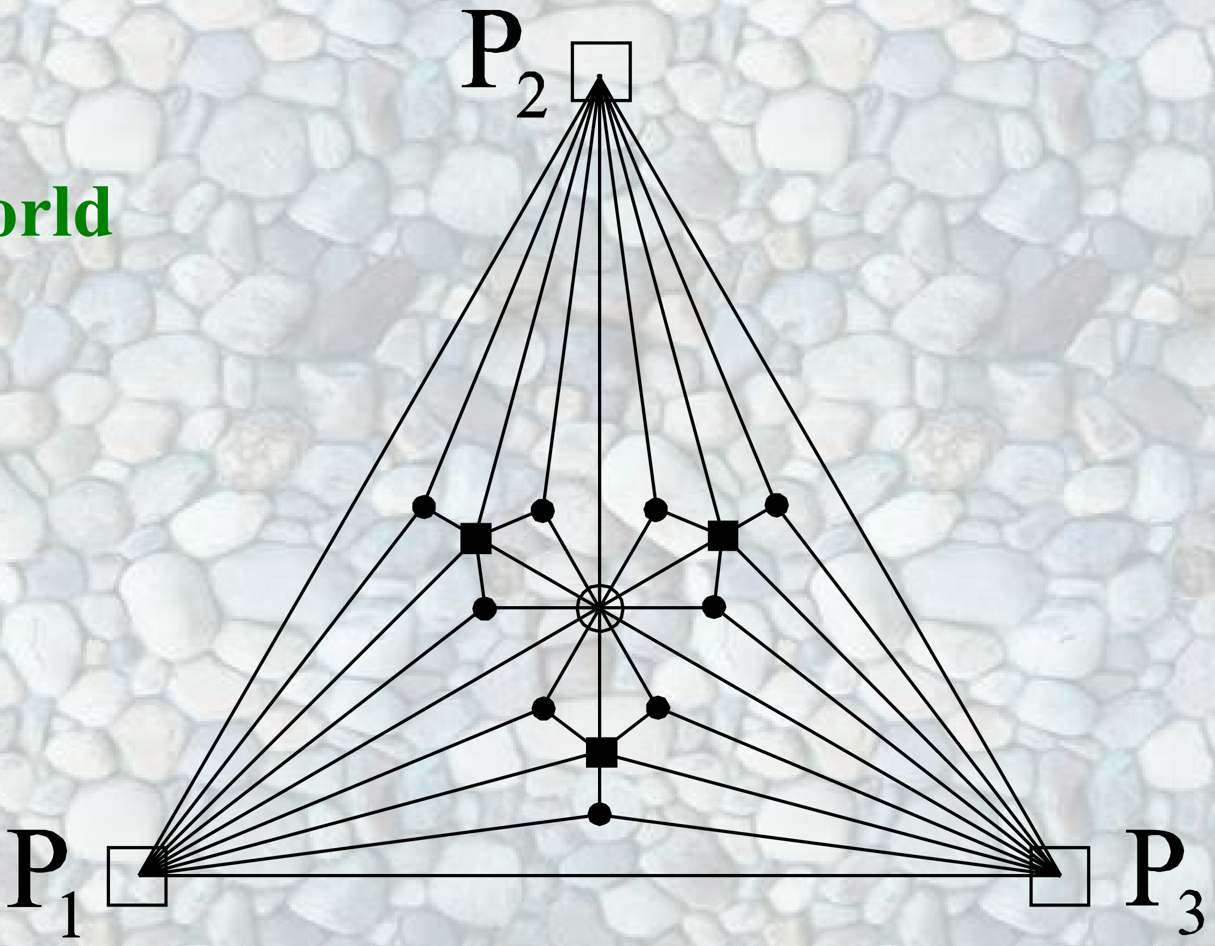
neuroscience $\gamma = 2.1$

scientific collaborations

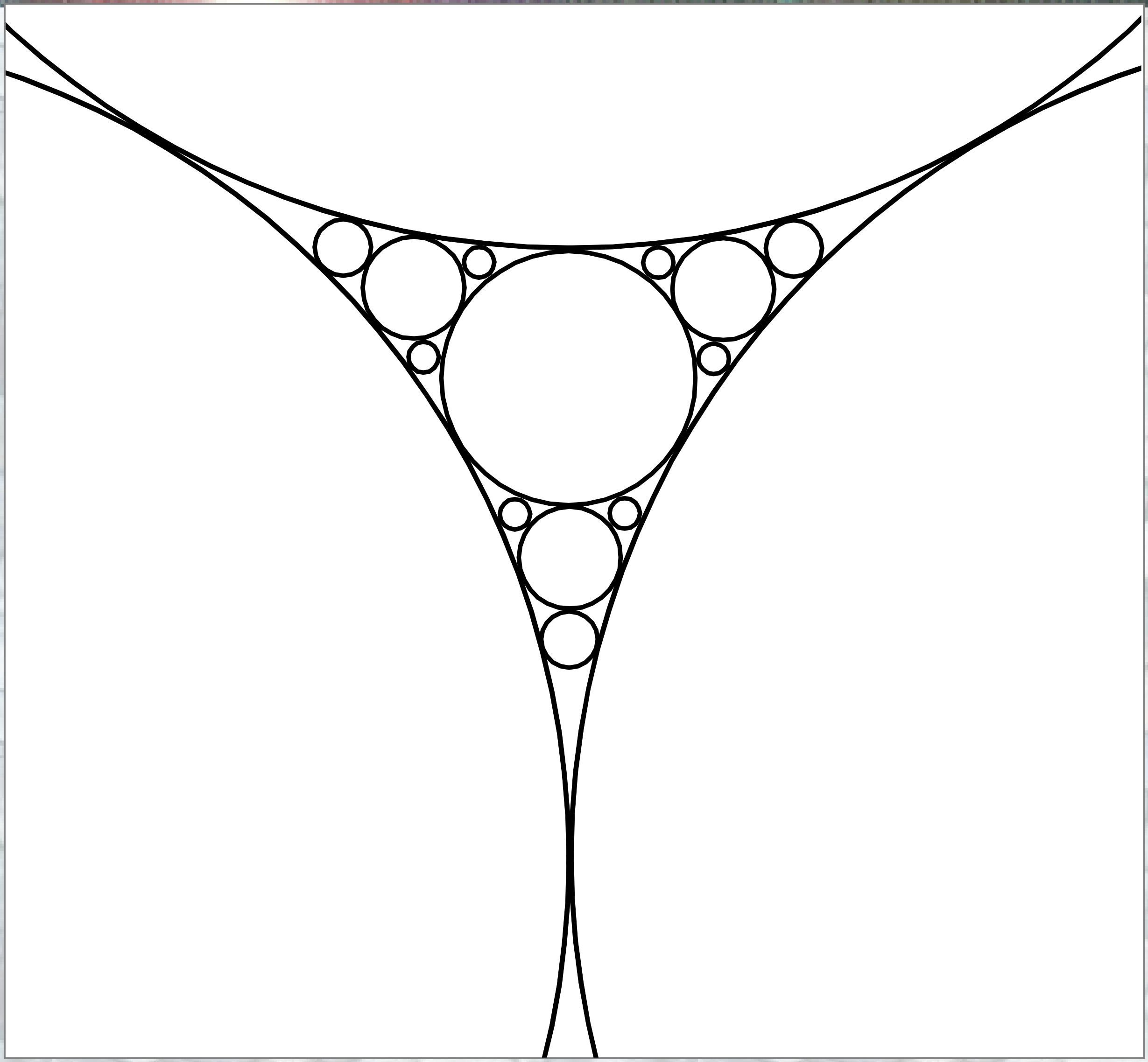
Model: Barabasi-Albert $\gamma = 3$

Apollonian network

- scale-free
- (ultra) small world
- Euclidean
- space-filling
- matching



Apollonian packing



Other applications

- Force networks in polydisperse packings
- Highly fractured porous media
- Networks of roads
- Systems of electrical supply lines

Degree distribution

N_n = number of sites at generation n

$m(k,n)$ = number of vertices of degree k

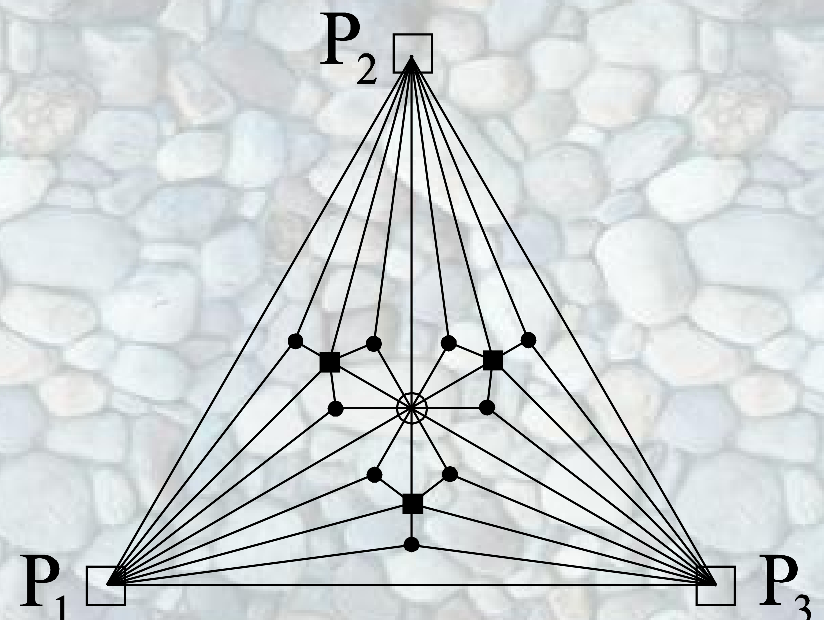
cummulative distribution $W(k) = \sum_{k' > k} m(k',n) / N_n$

scale-free: $P(k) \propto k^{-\gamma}$

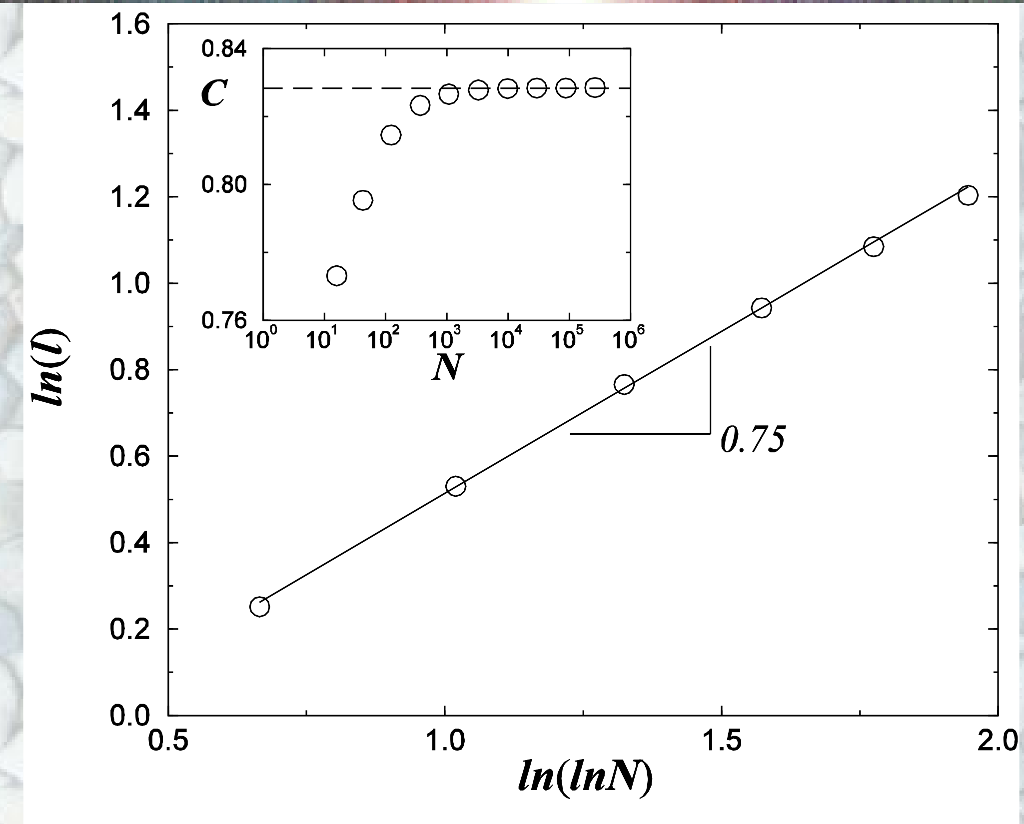
$\Rightarrow \quad W(k) \propto k^{1-\gamma}$

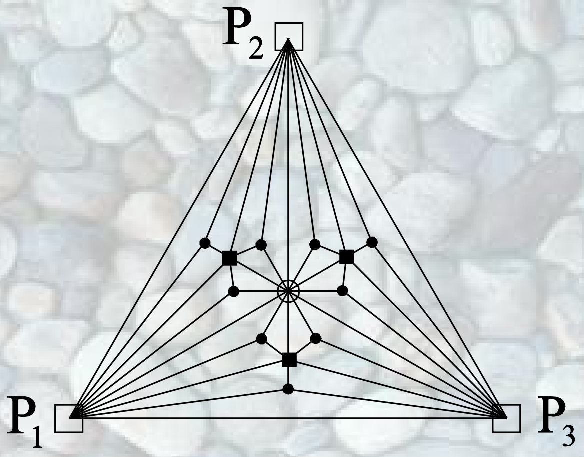
k	$m(k,n)$
3^n	3
3^{n-1}	3×2
3^{n-2}	3×2^2
$\vdots$	$\vdots$
3^2	$3 \times 2^{n-1}$
3	3×2^n
1	2^{n+1}

$$\Rightarrow \quad \gamma = 1 + \frac{\ln 3}{\ln 2} \approx 2.585$$



Small-world properties





clustering coefficient

$$C = \frac{2}{k(k-1)} \times \text{ number of connections between neighbors}$$

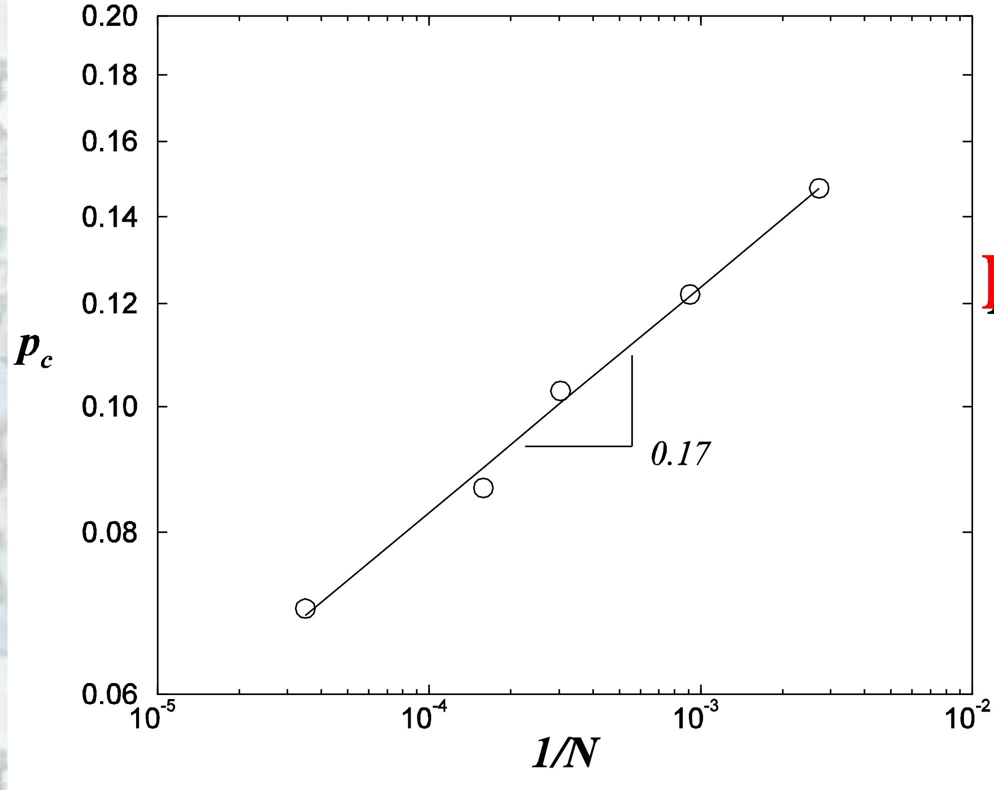
$$C = 0.828$$

shortest path

l = ⟨chemical distance between two sites⟩

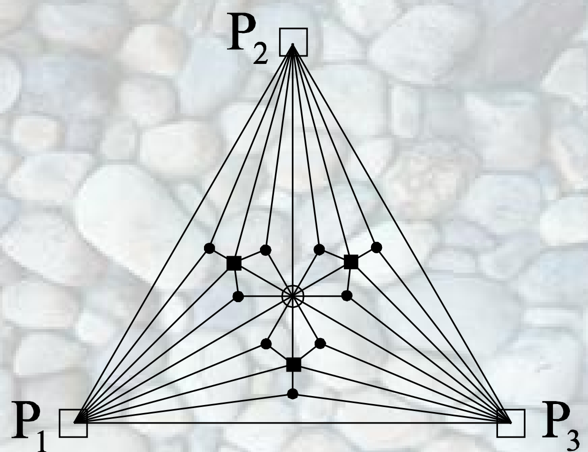
$$l \propto (\ln N)^{3/4}$$

Percolation threshold



epidemics

random failure



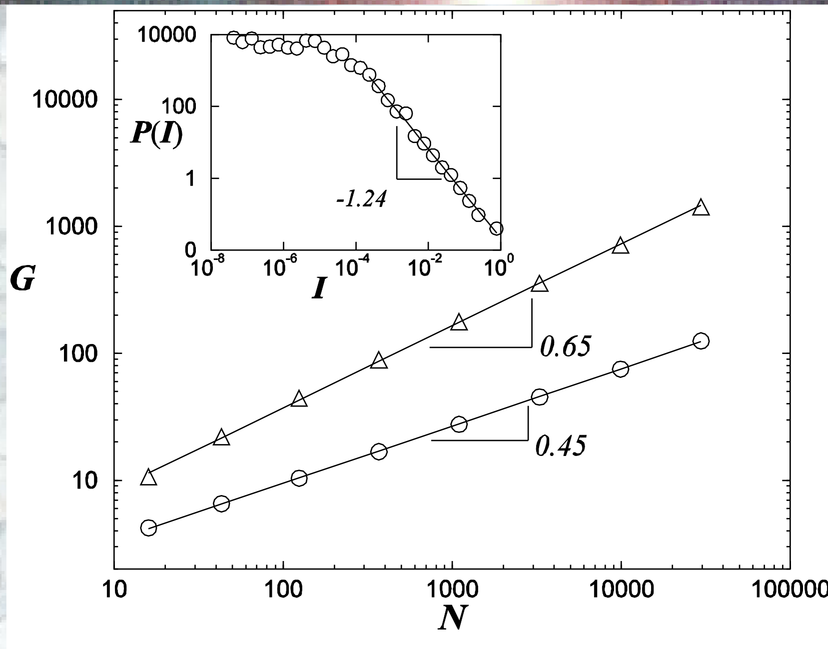
bond percolation

at p_c : P_1, P_2 and P_3 simultaneously connected

$p_c \propto L^{-\frac{1}{\nu}}, L = \sqrt{N}$ → $\nu \approx 3$

porous media → Archie's law

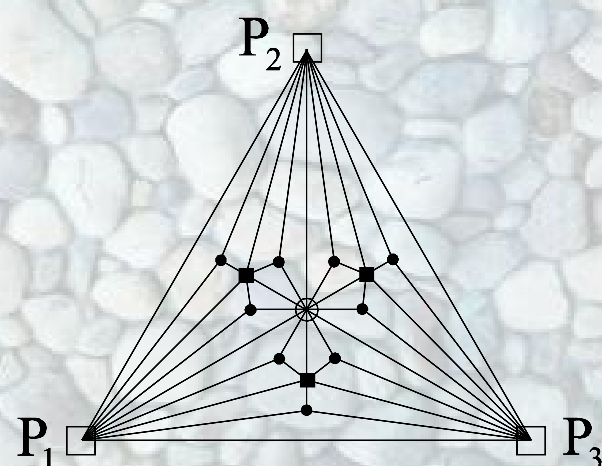
Electrical conductance



flux

fuses =

malicious attack



- from center to all sites of generation n:

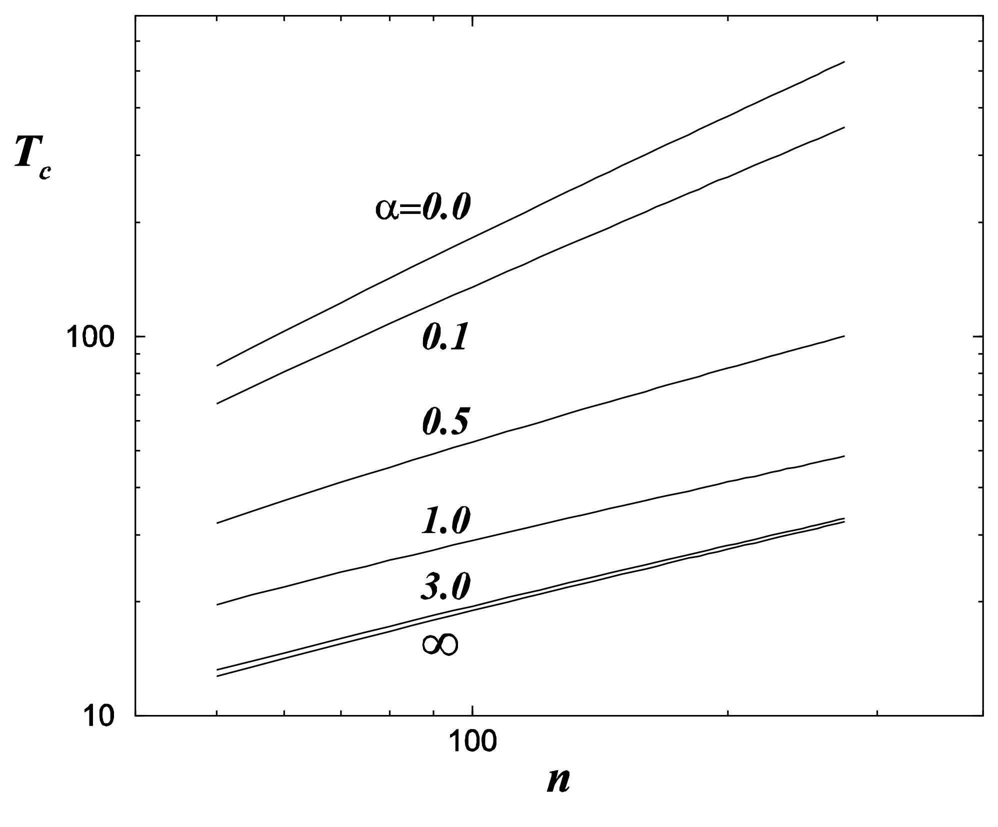
  $\left.\begin{matrix} S_1 \\ S_2 \end{matrix}\right\} = \text{number of outputs of} \left\{\begin{matrix} \text{system} \\ \text{each site} \end{matrix}\right.$

  $n-1 \rightarrow n \Rightarrow S_1 \rightarrow 3 * S_1, S_2 \rightarrow 2 * S_2 \Rightarrow \quad z = 2/3$

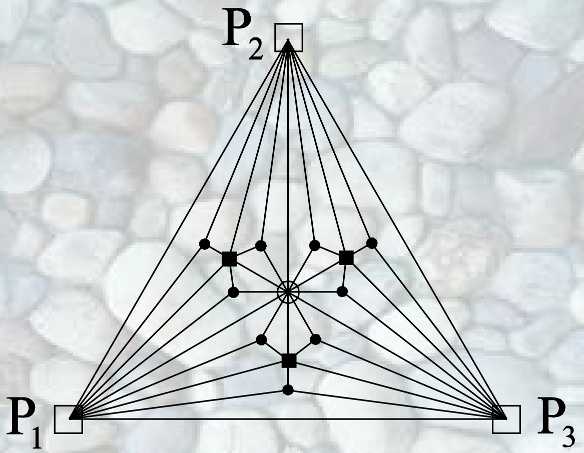
  current distribution $P(I) \propto I^{-1.24}$

- from P_1 to P_2 : $z \approx 0.45$

Critical temperature of Ising model



opinion



coupling constant $J_n \propto n^{-\alpha}$

correlation length J_n diverges at T_c

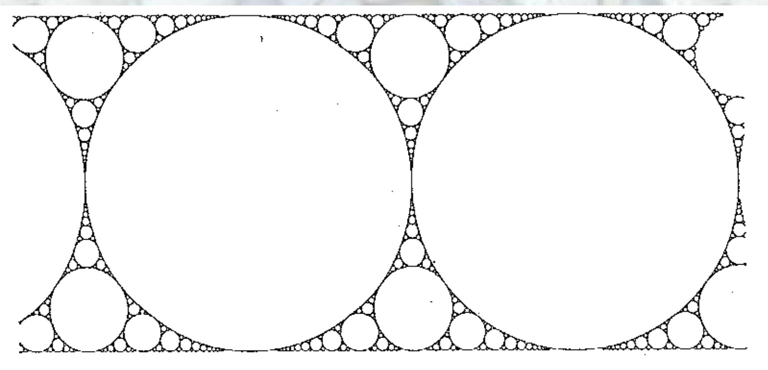
free energy, entropy, specific heat are smooth

magnetization $m : e^{-T^{\lambda}}$ $T \rightarrow$ Ą

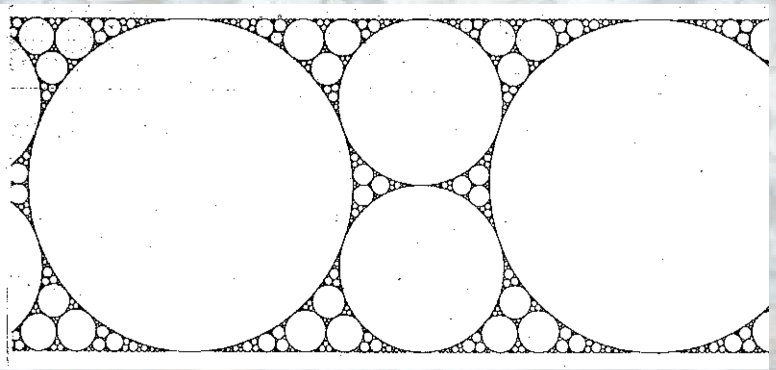
Generalized Apollonians

m=0

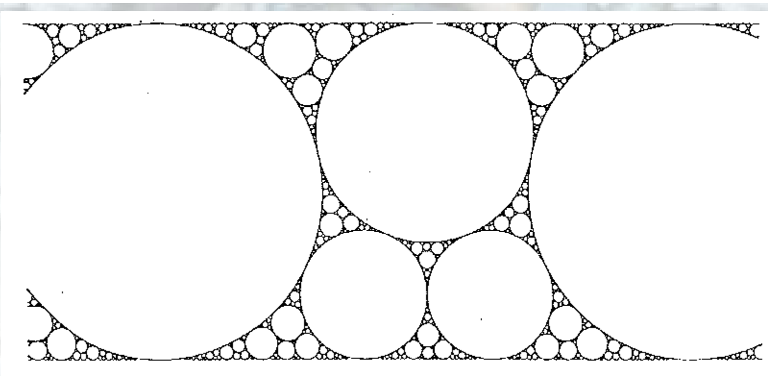
n=∞



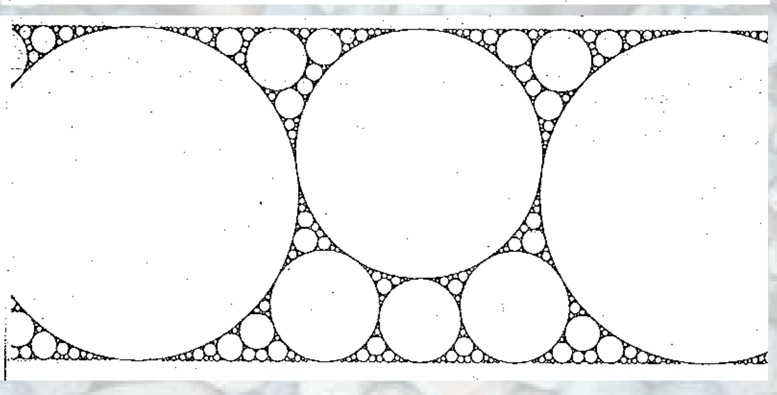
m=1



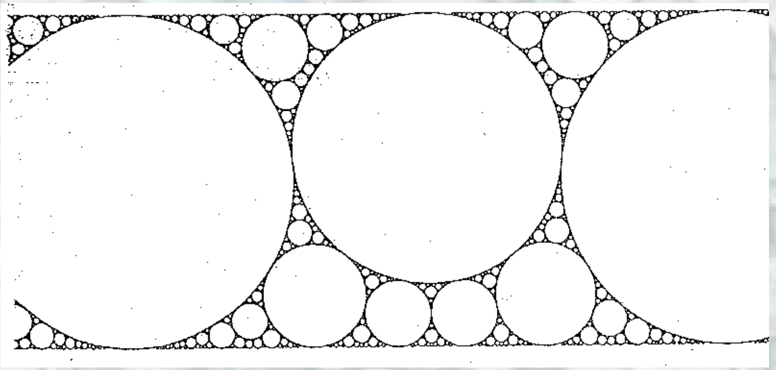
m=2



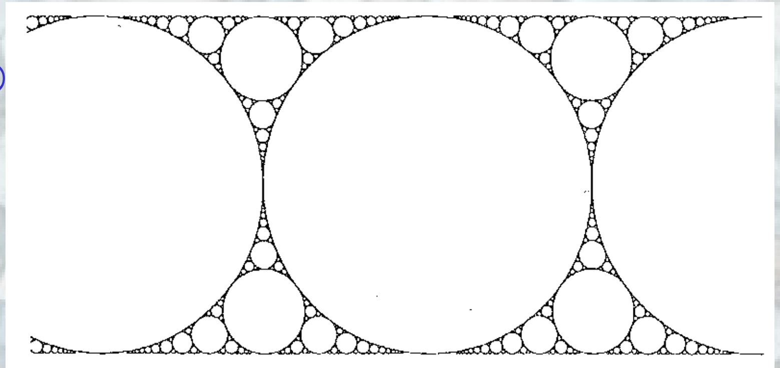
m=3



m=4



m=∞



Osculatory packing

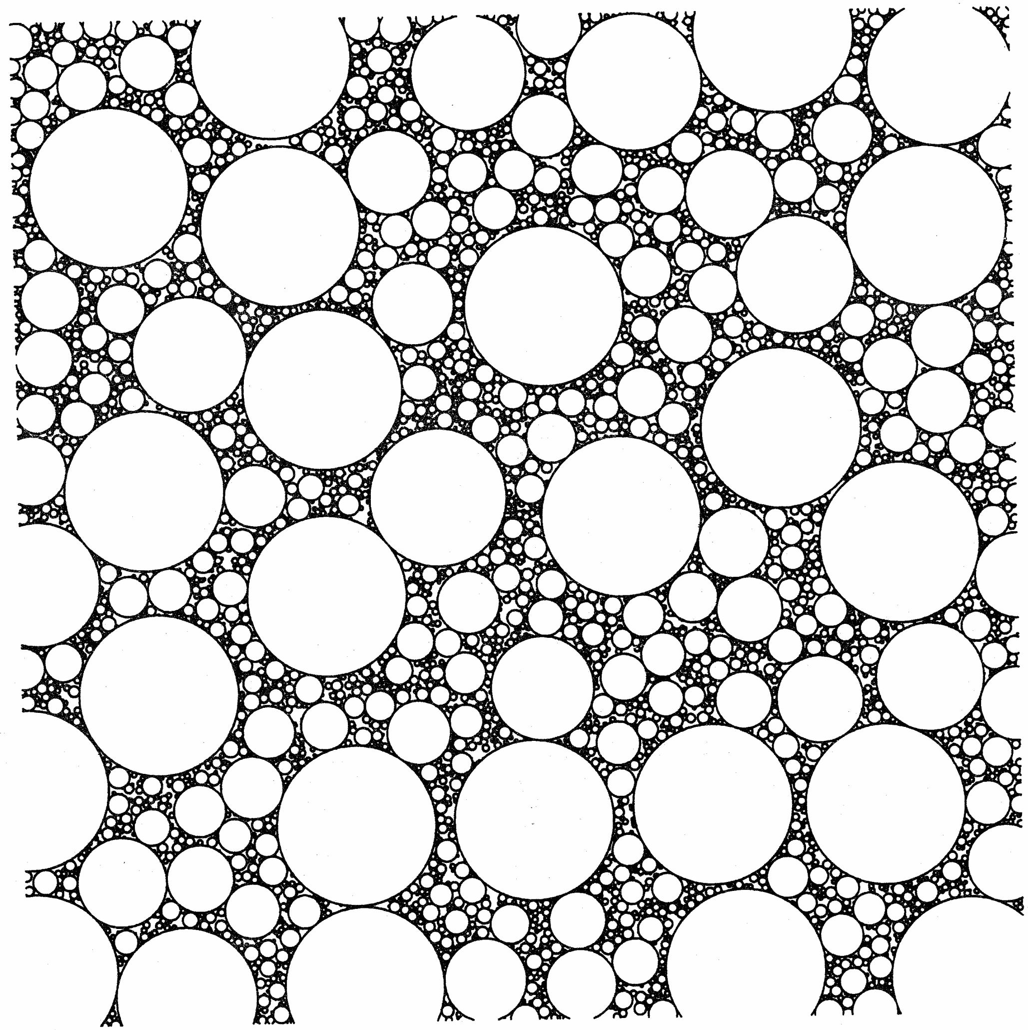
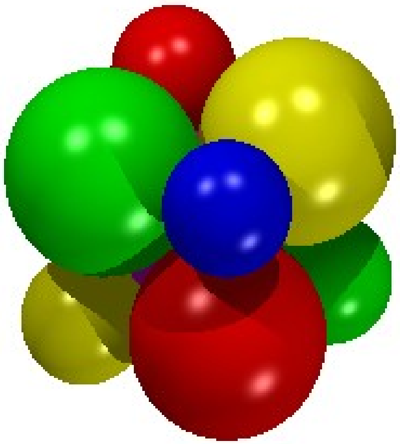
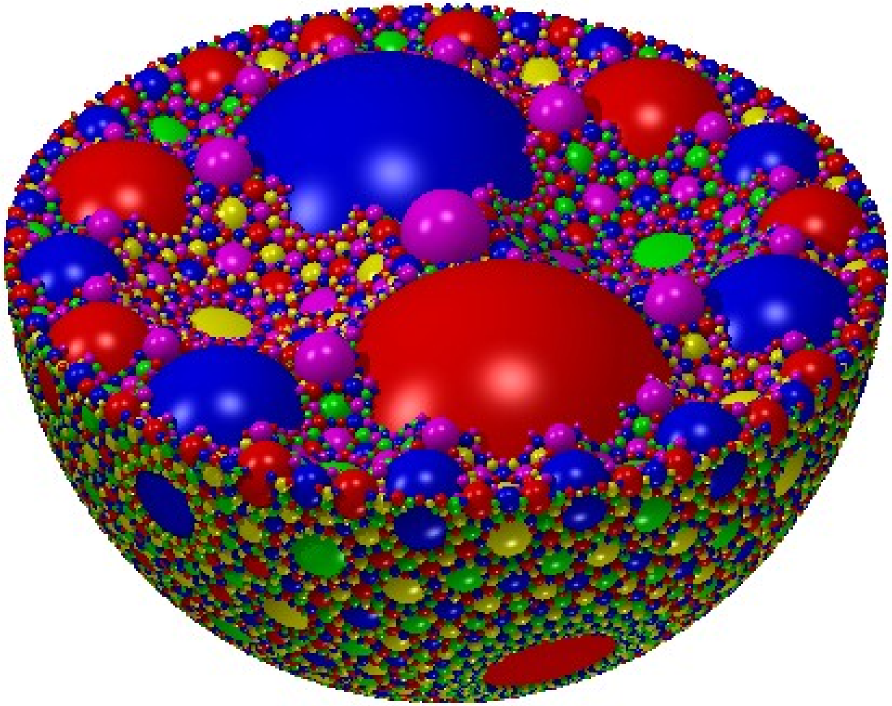


Fig. 3. Osculatory packing for white distribution of circle centers with an upper bound for allowed circle radii.

Apollonian packing in 3D



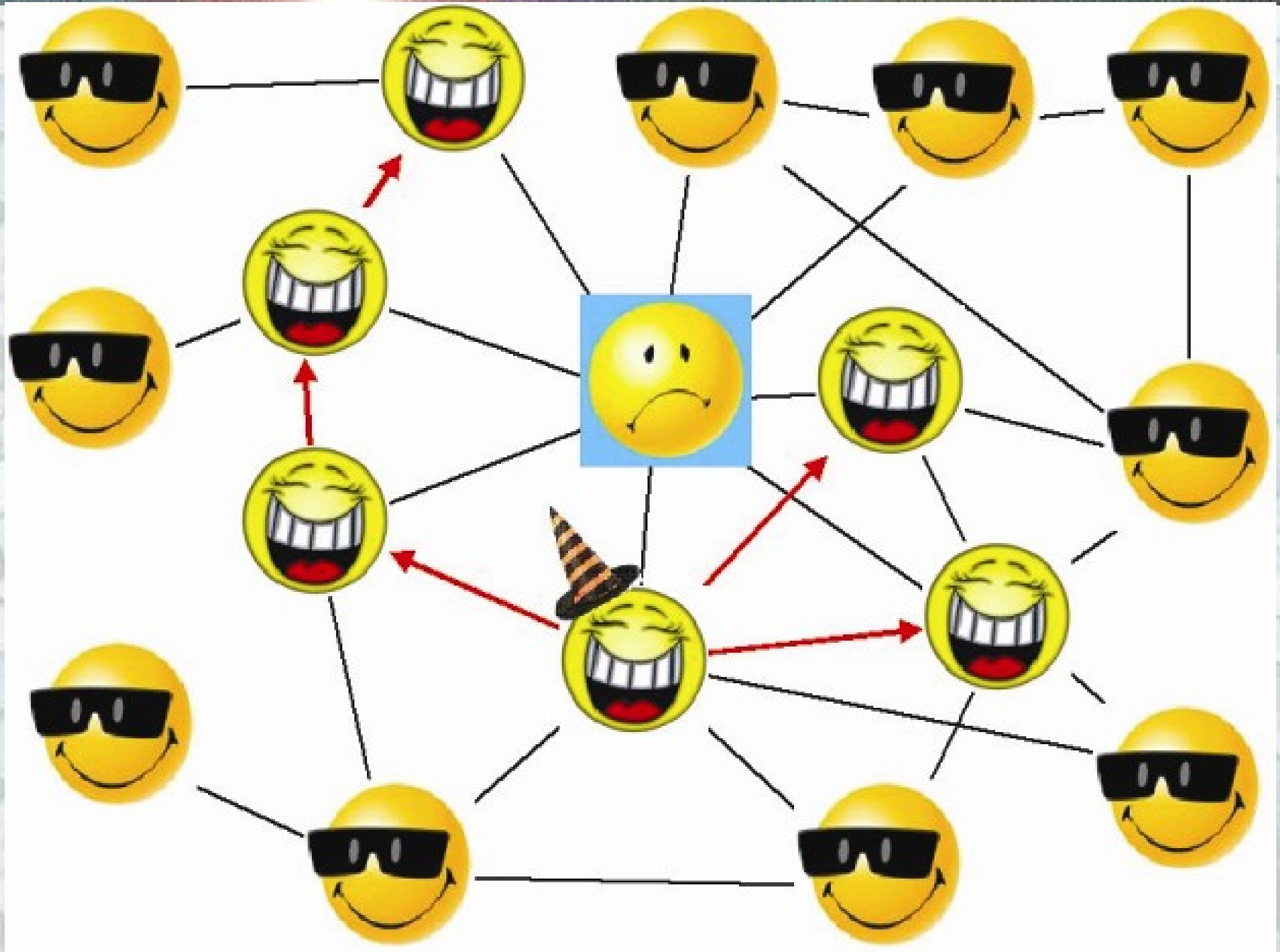


Tetrahedron

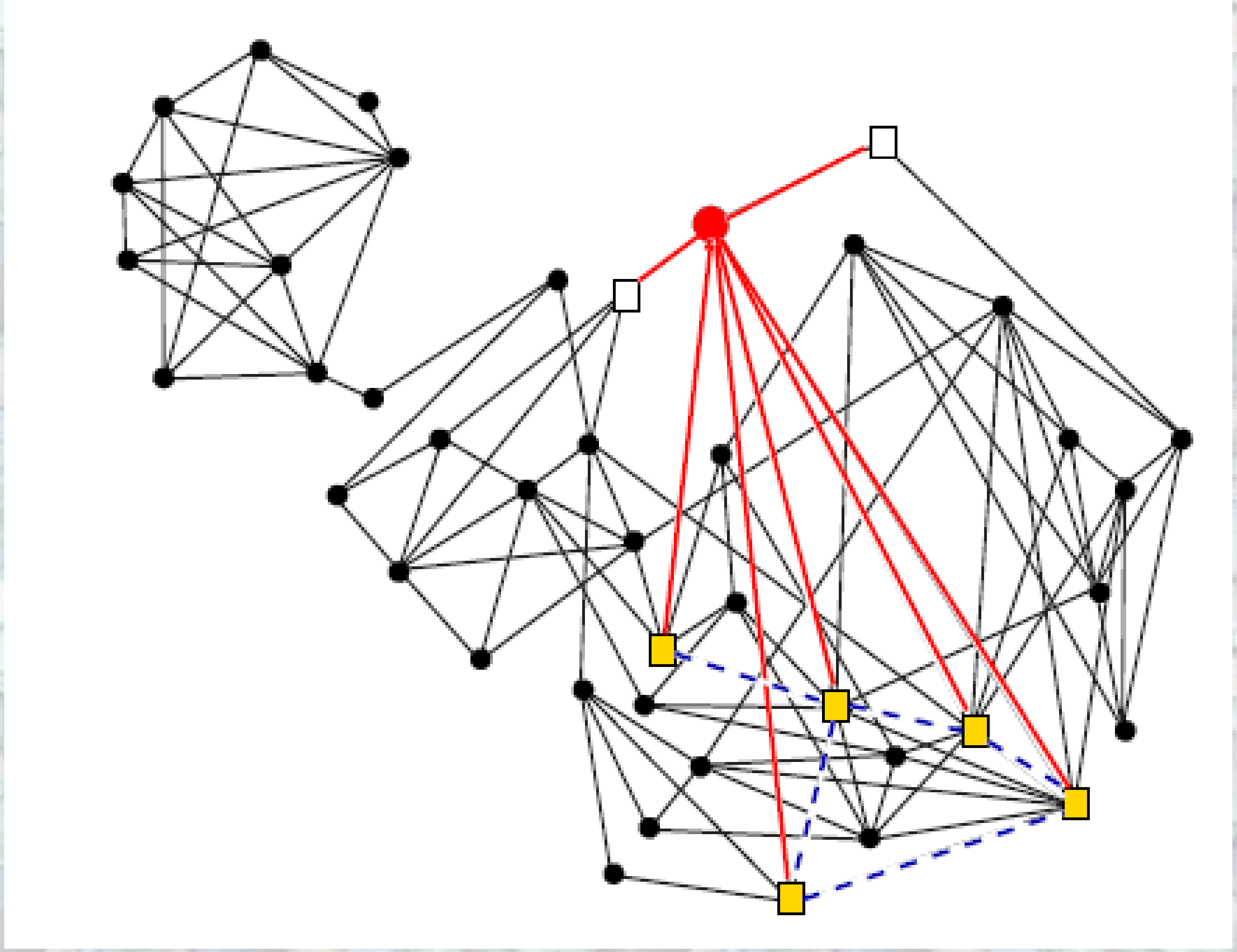
Conclusions on Apollonian networks

- Stable against random failure.
- Opinions stabilize at any interaction strength.
- Supply current distribution follows power-law while point-to-point distribution is log-normal.
- Apollonian networks are between small world $(l \propto \ln N)$ and ultrasmall world $(l \propto \ln \ln N)$
- Force networks of polydisperse packings have few contacts between grains and their rigidity increases with density like a power-law.
- Porous medium follows Archie's law. $\Phi_{\chi} = 0$.

Gossip



Gossip on network



Gossip propagation

Spreading time τ is the minimum time it takes for a gossip to reach all accessible persons.

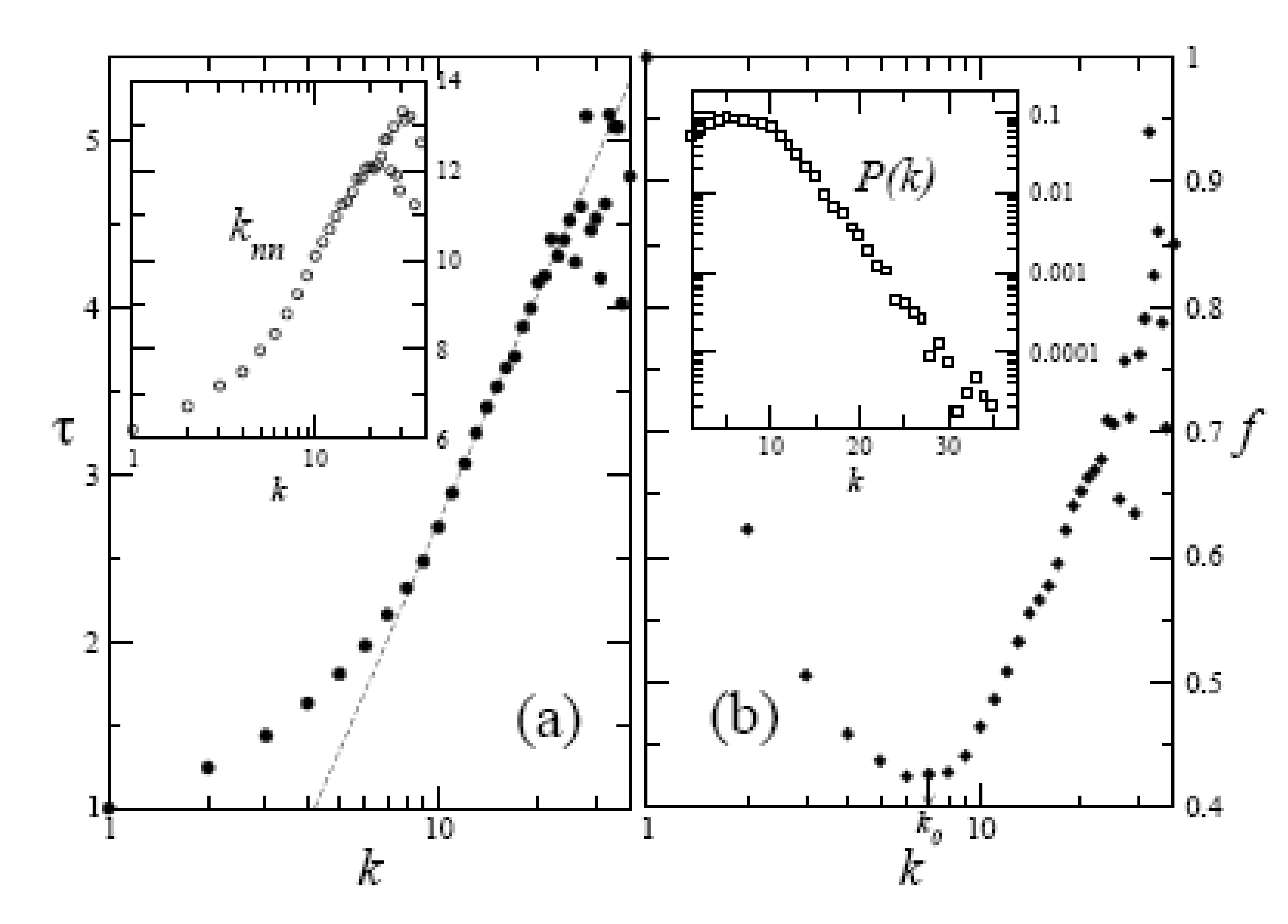
f

n is the total number of persons that eventually get the gossip.

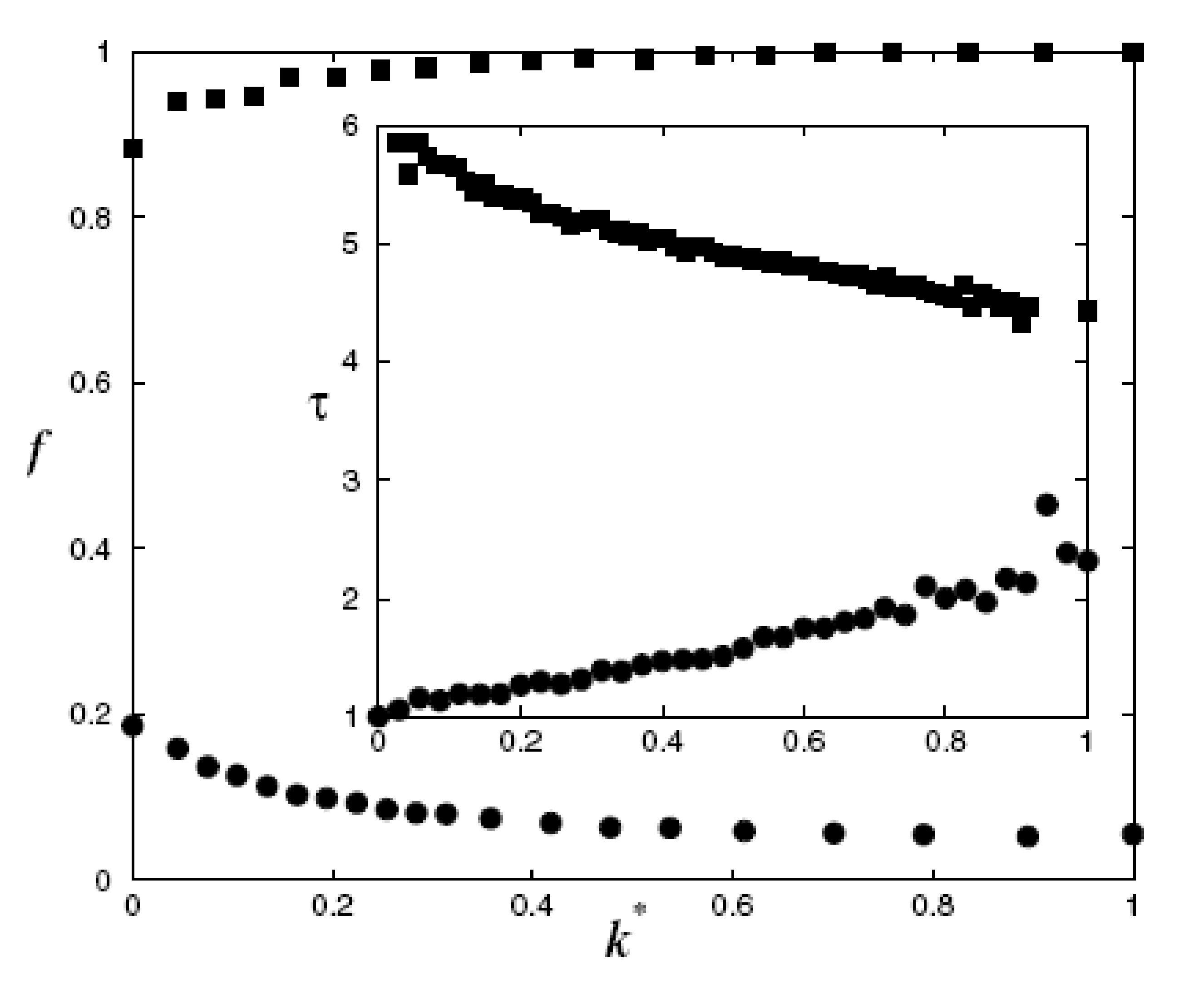
We also define the „spread factor" f:

$$f = \frac{n_f}{k}$$ where k is the degree of the victim.

Gossip in schools

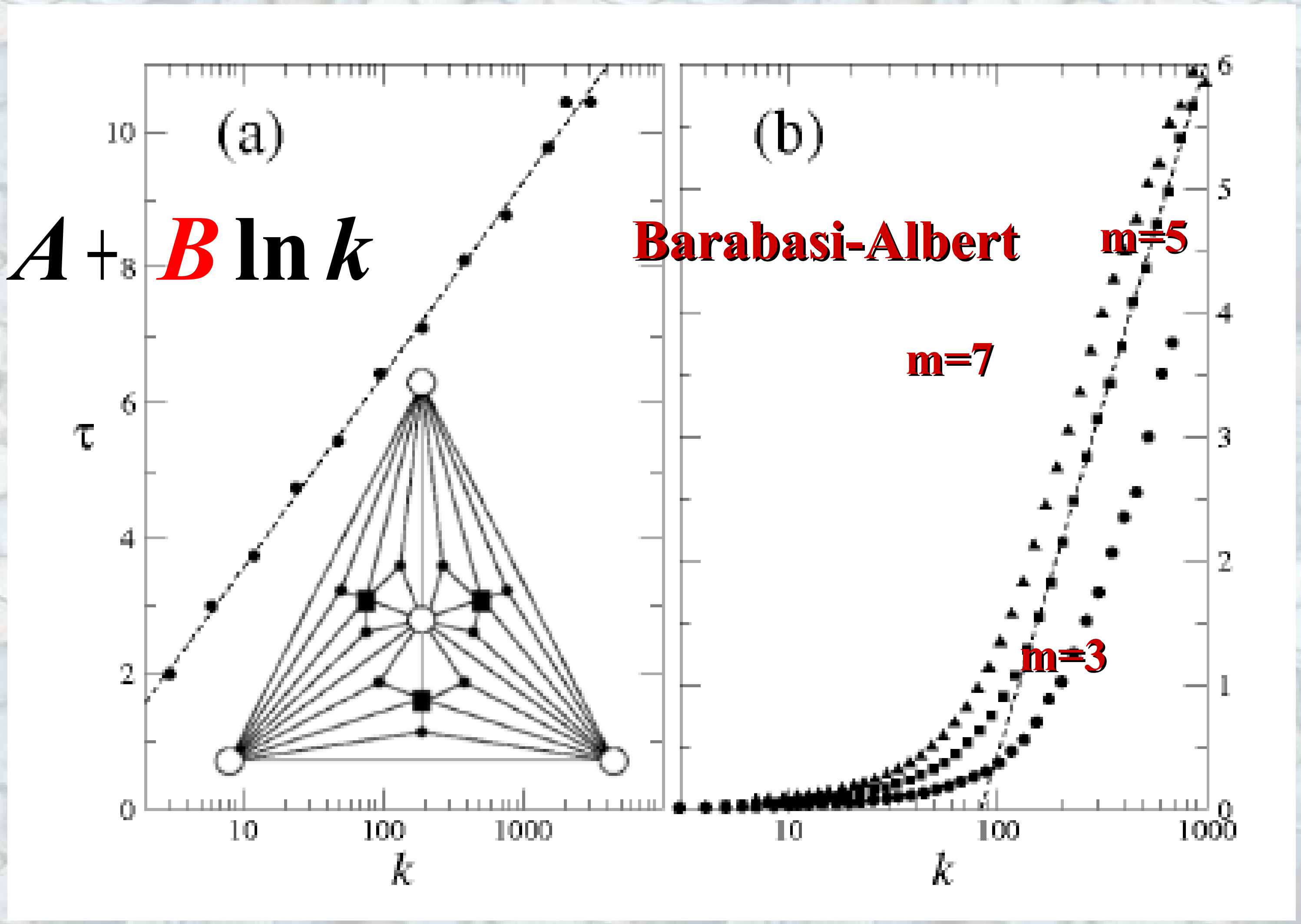


Gossip on random graph

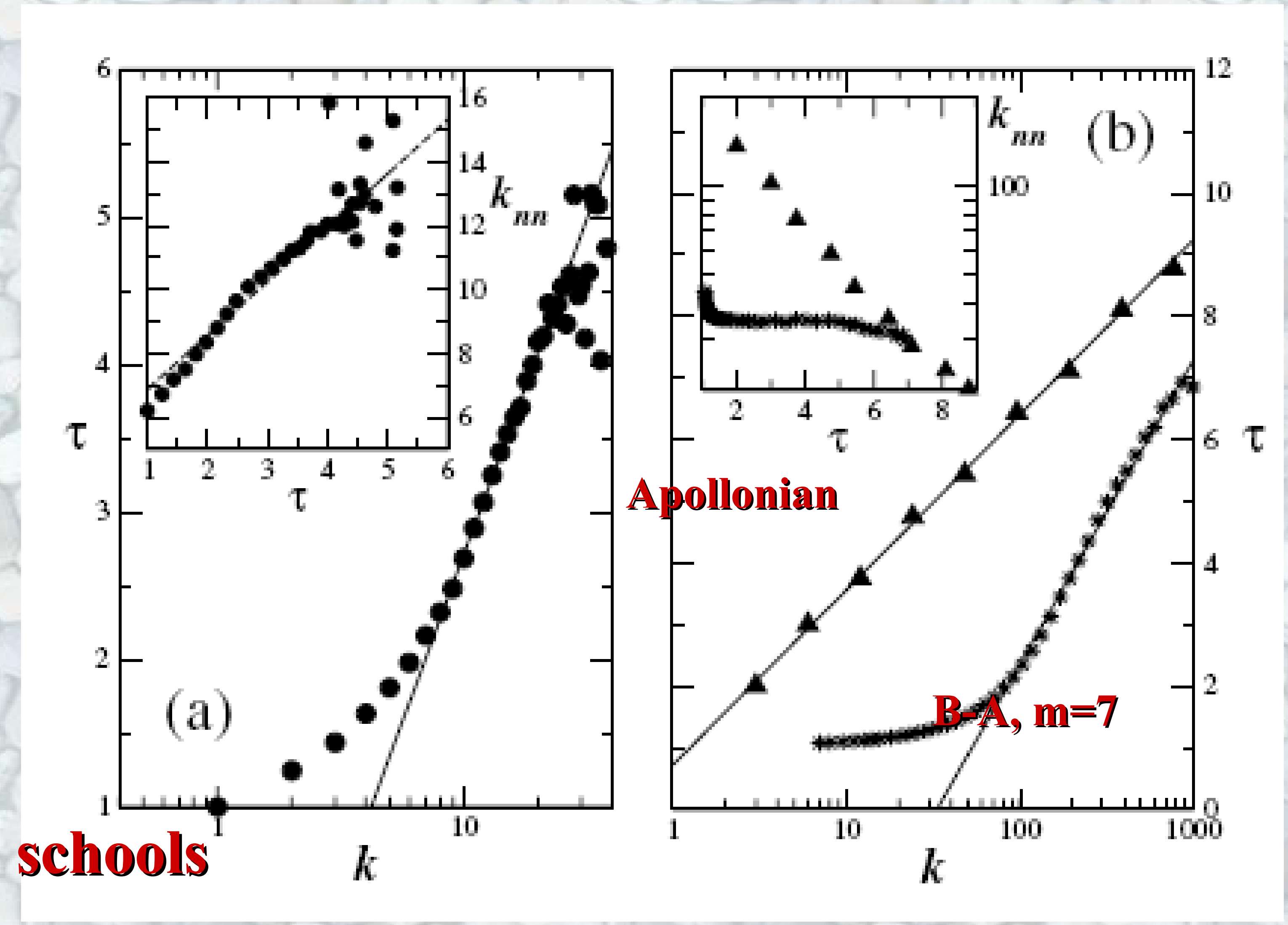


Spreading time

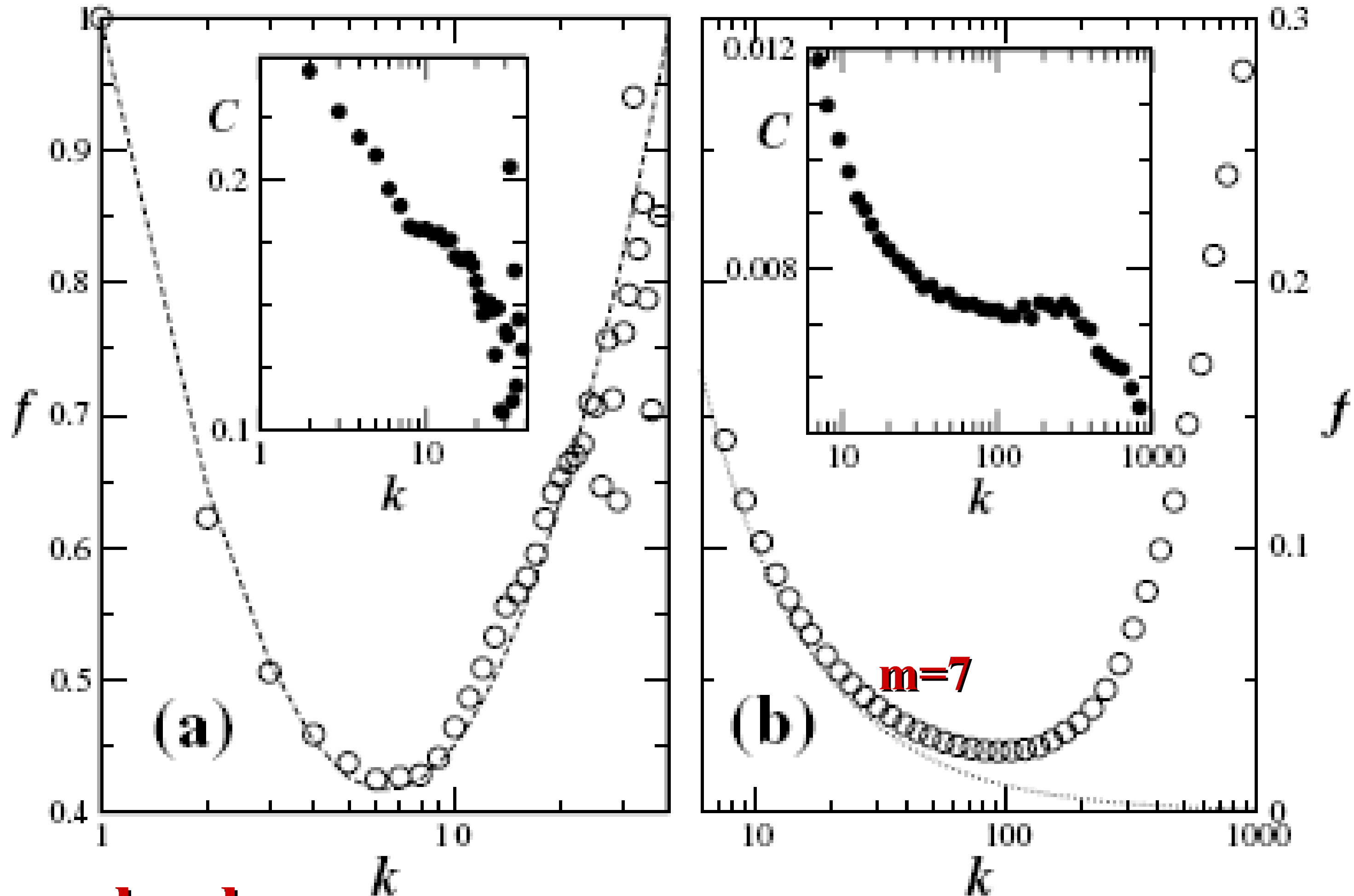
$$\tau = A + B \ln k$$



Spreading time



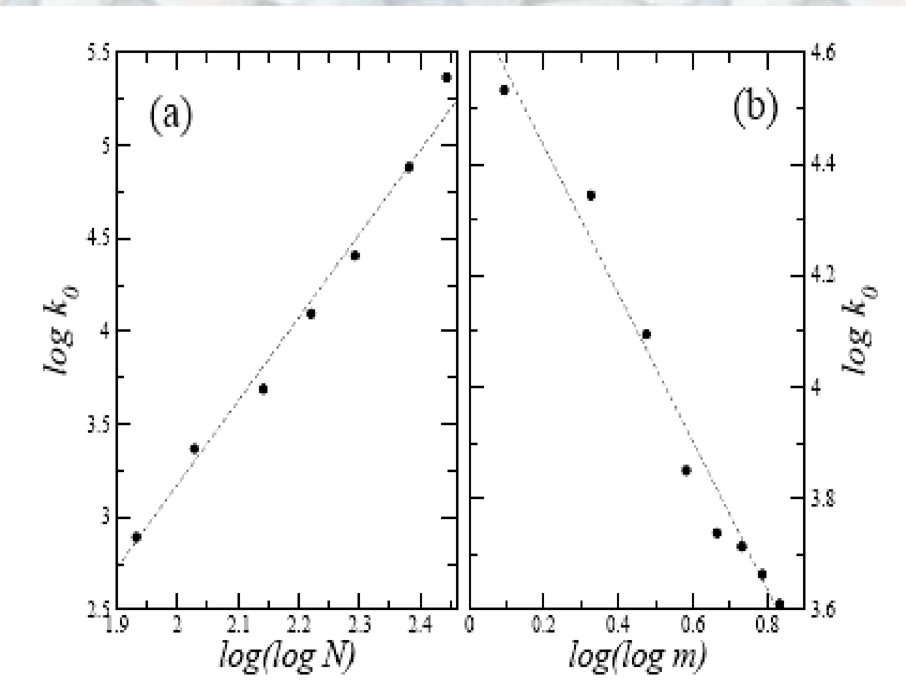
Spreading factor

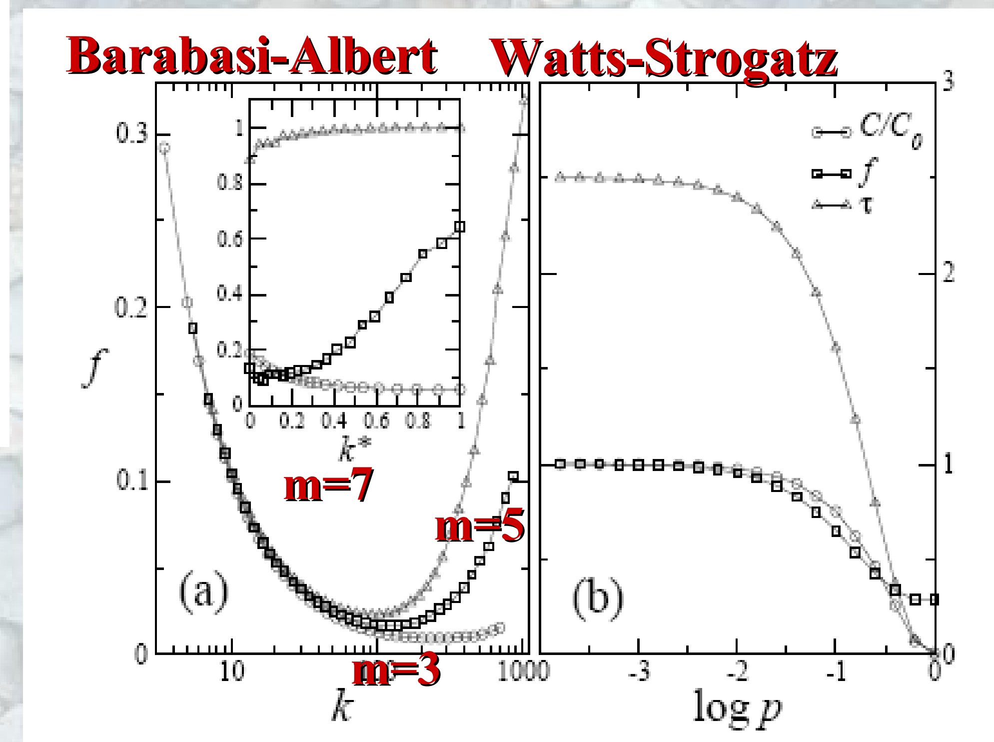


schools

Barabasi-Albert network

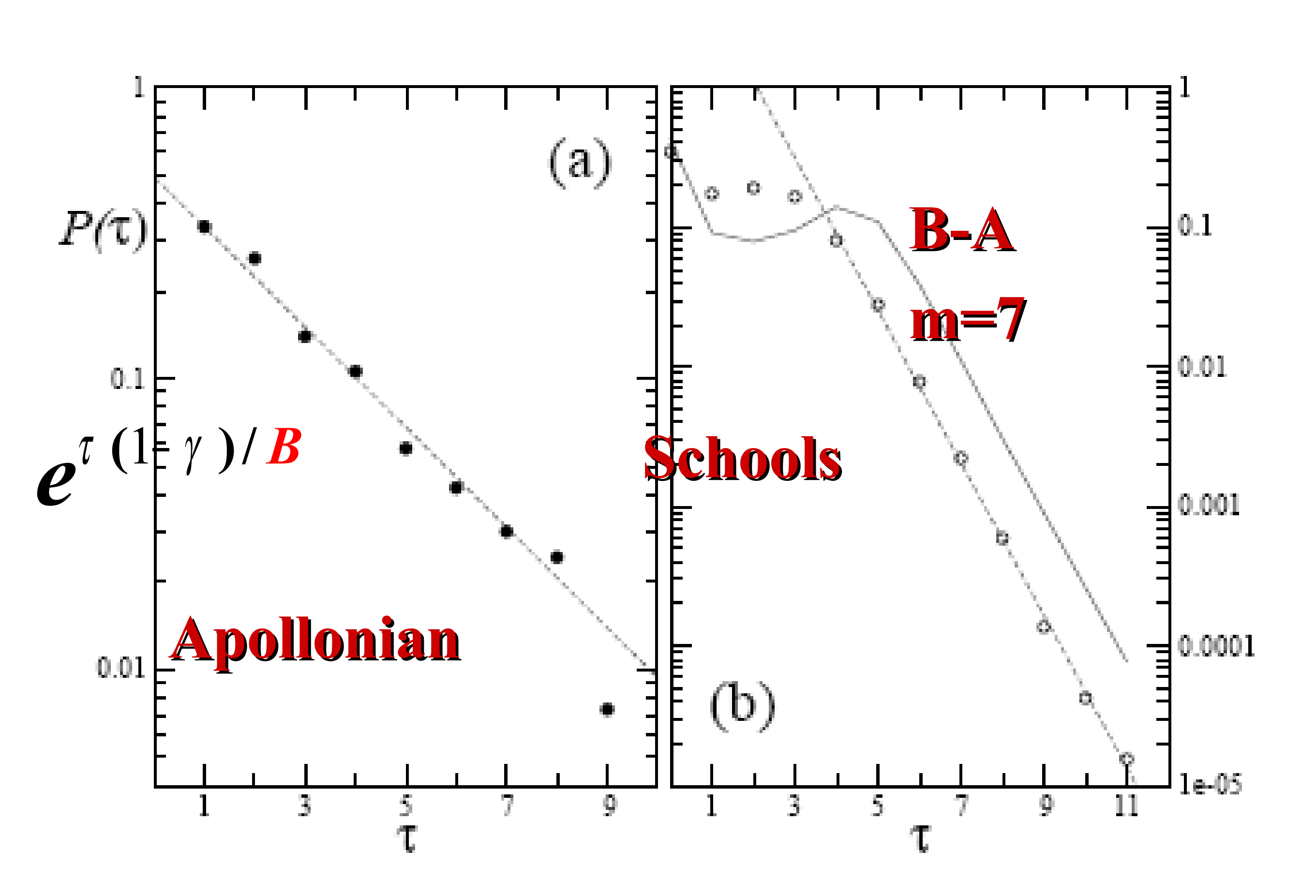
Spreading factor



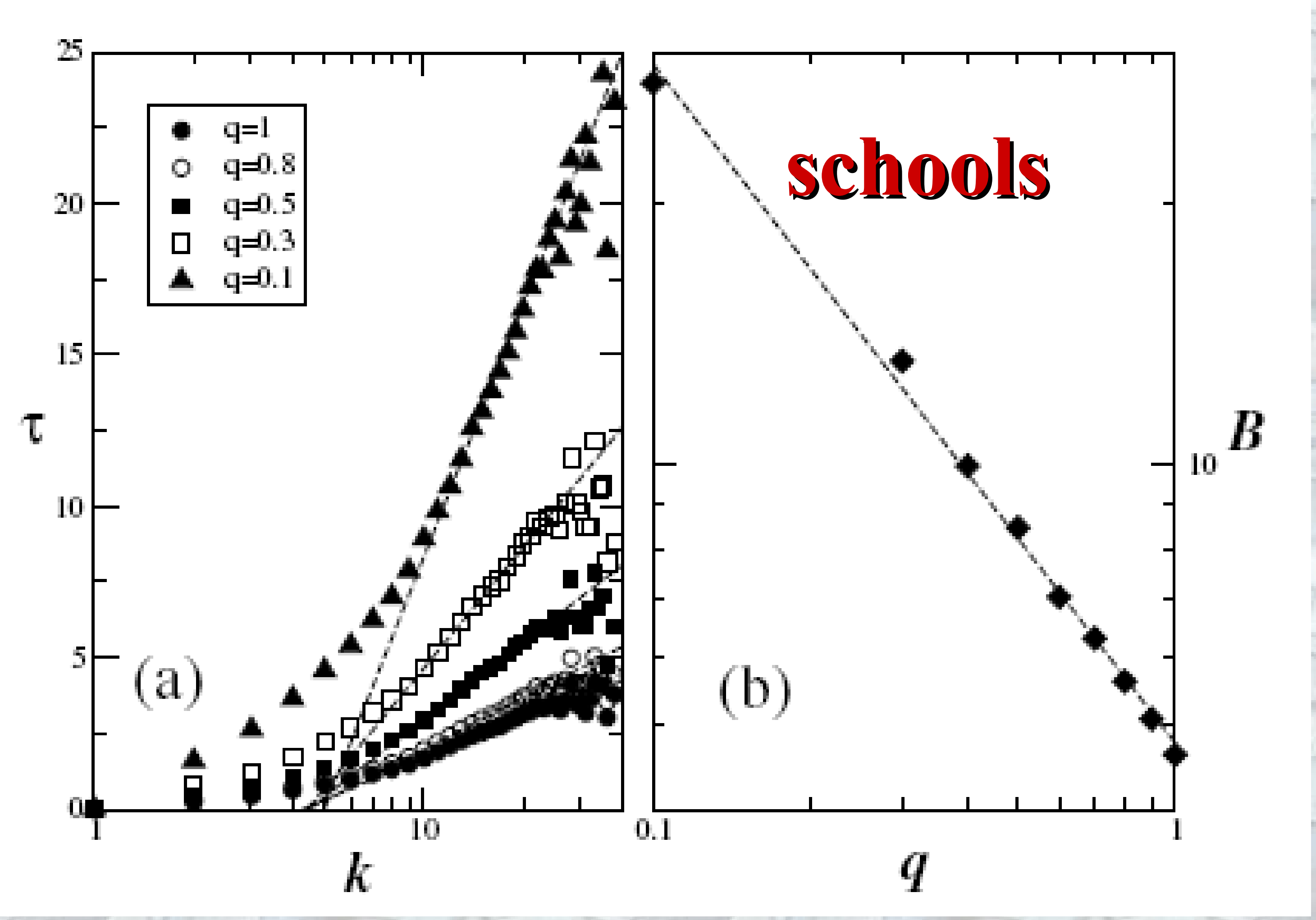


Distribution of τ

$$P(\tau) = e^{\tau(1-\gamma)/B}$$

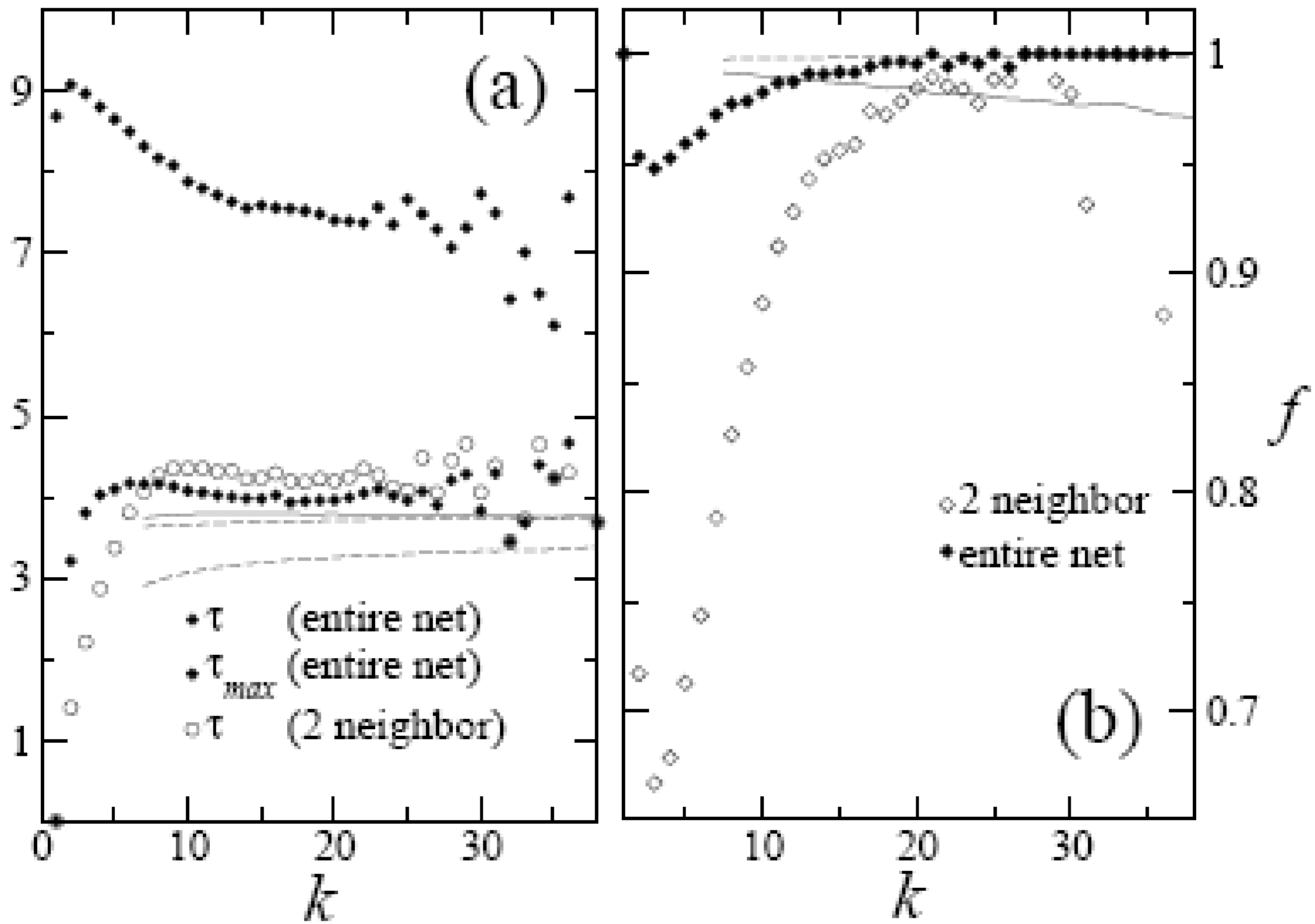


Gossip with probability q

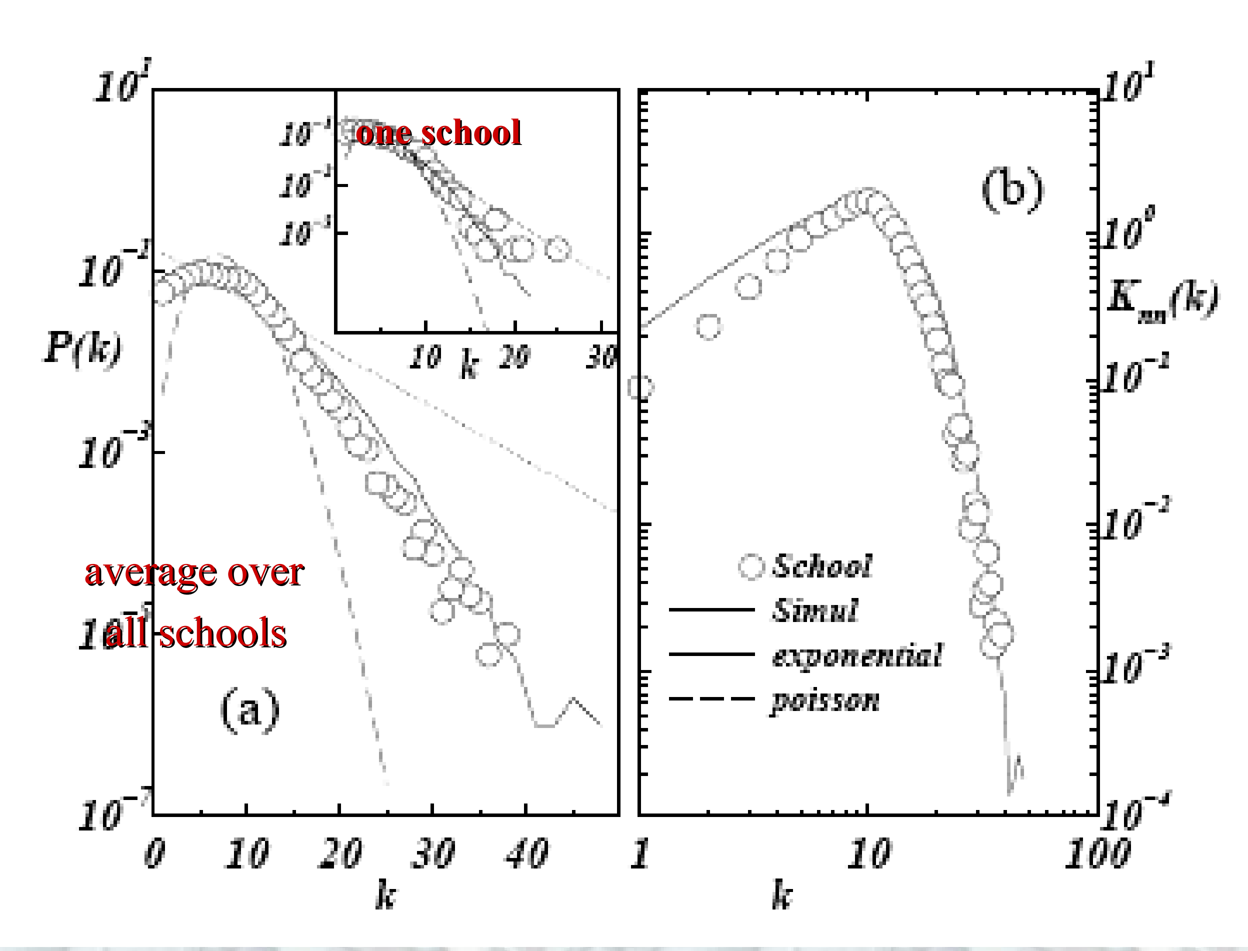


Gossip about famous people

symbols = schools   lines = Barabasi-Albert

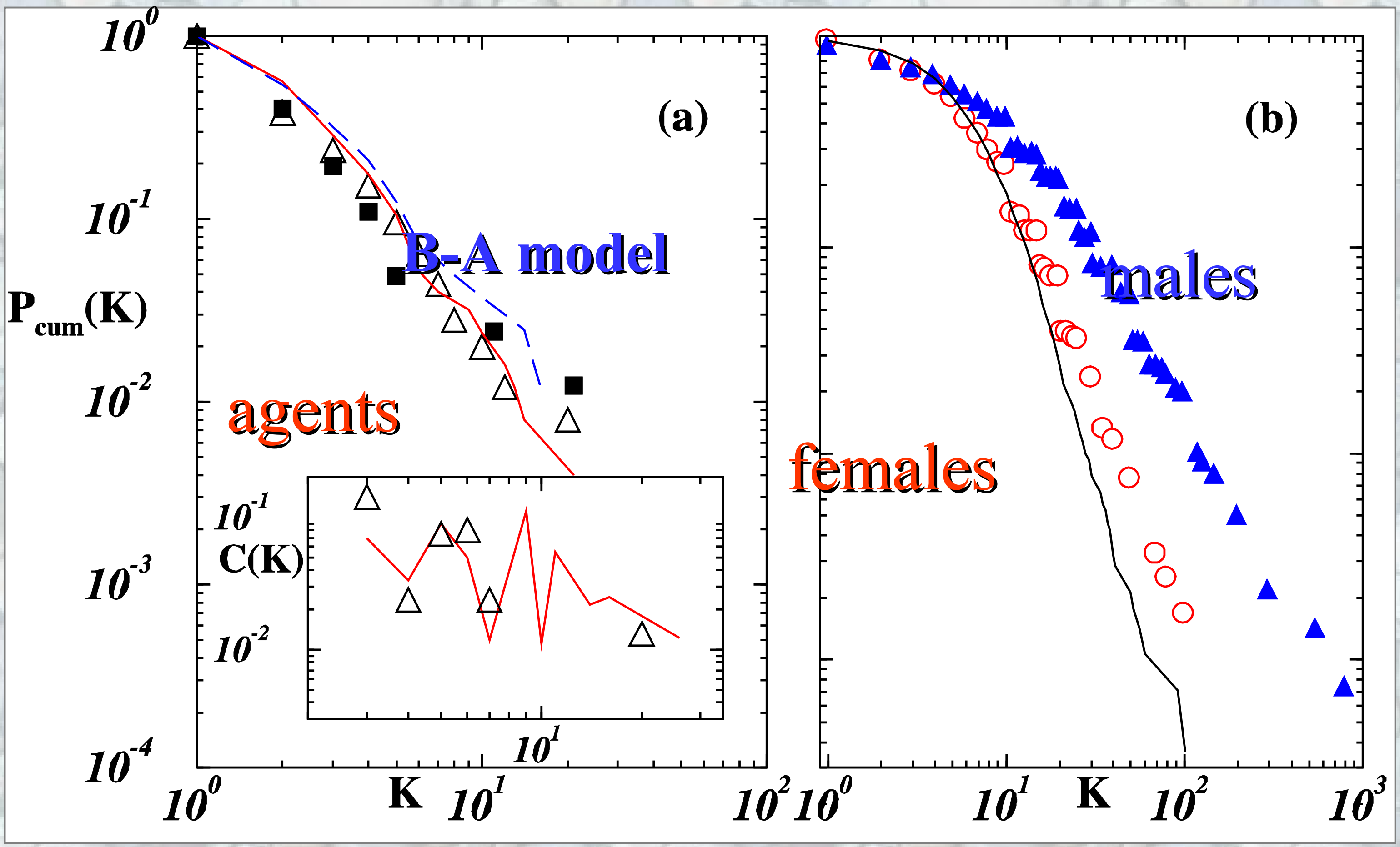


Degree distribution



Sexual contact networks

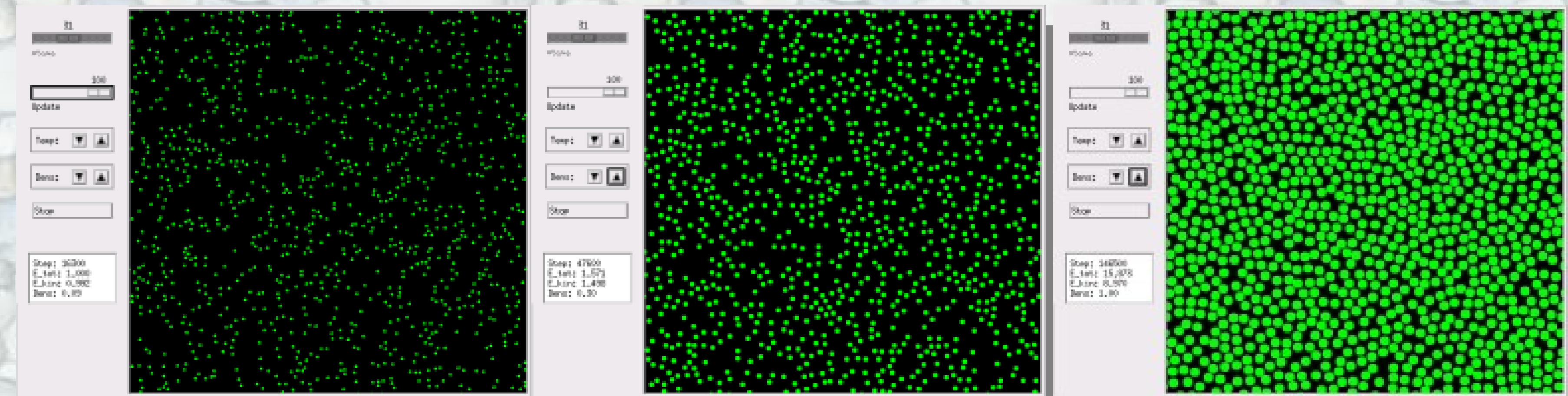
homosexual (Colorado Springs) **heterosexual** (Sweden)

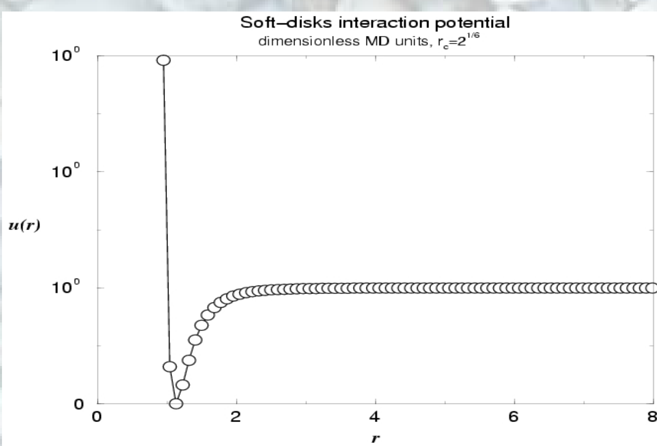


F. Liljeros et al, Nature **411**, 907 (2001)

Agent model

soft-disks in 2d





interaction potential

$$u(r_{ij}) = U_0 \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right] + U_0 \quad , \quad r_{ij} \le r_c = 2^{1/6}\sigma$$

collision time

$$\tau_{coll} = \frac{1}{\rho\, 2r_c} \sqrt{\frac{m}{\pi T k_B}}$$

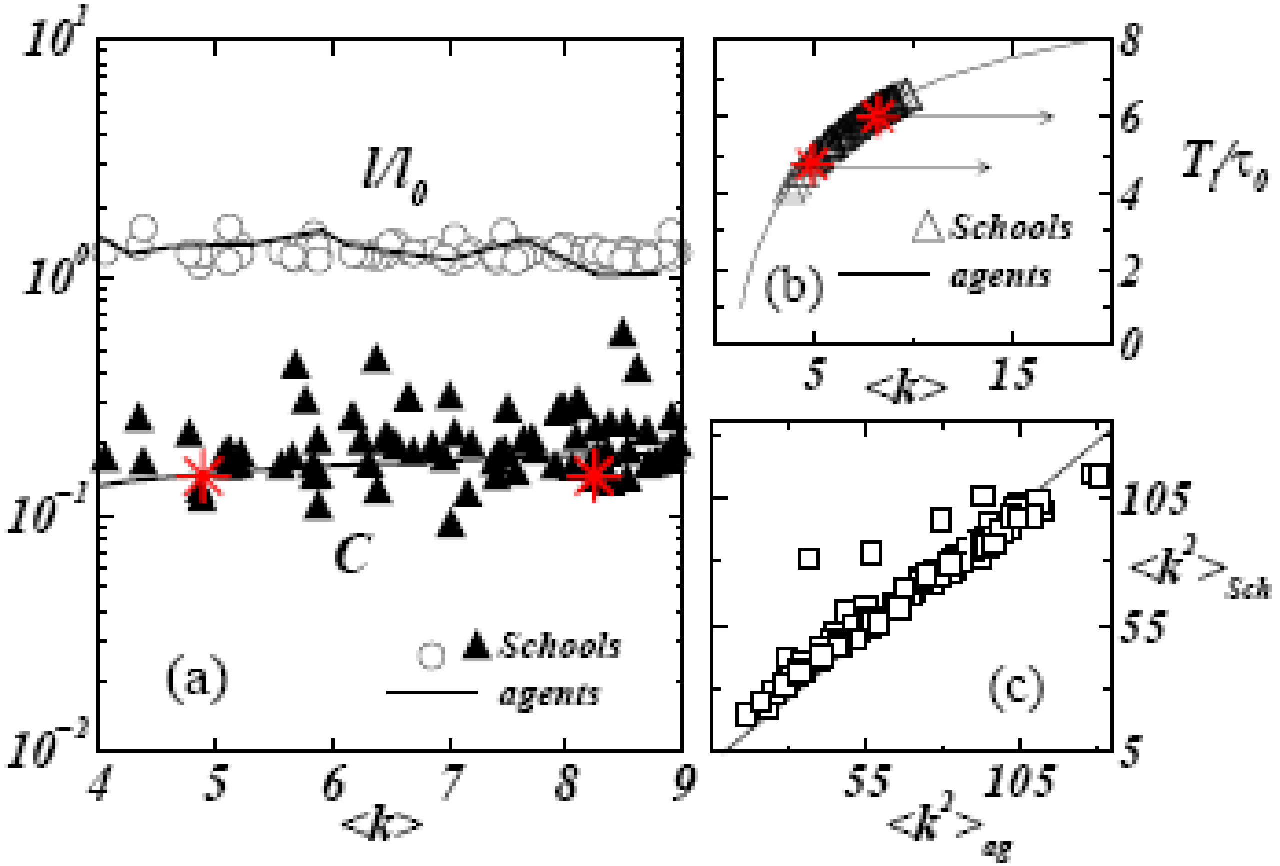
Marta González

Evolution of network

Evolution of network with life-times

Comparison to schools

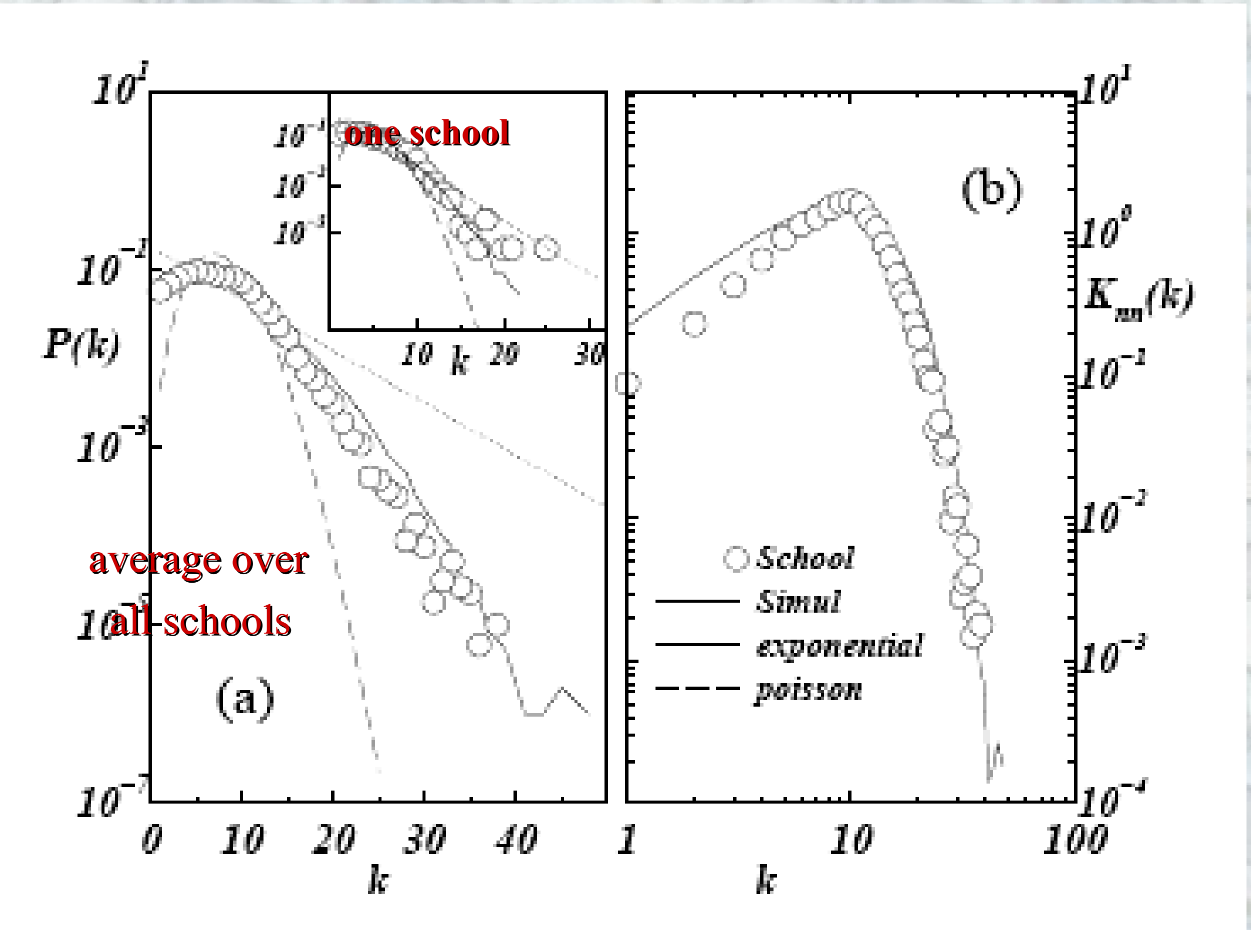
Shortest path and clustering coefficient



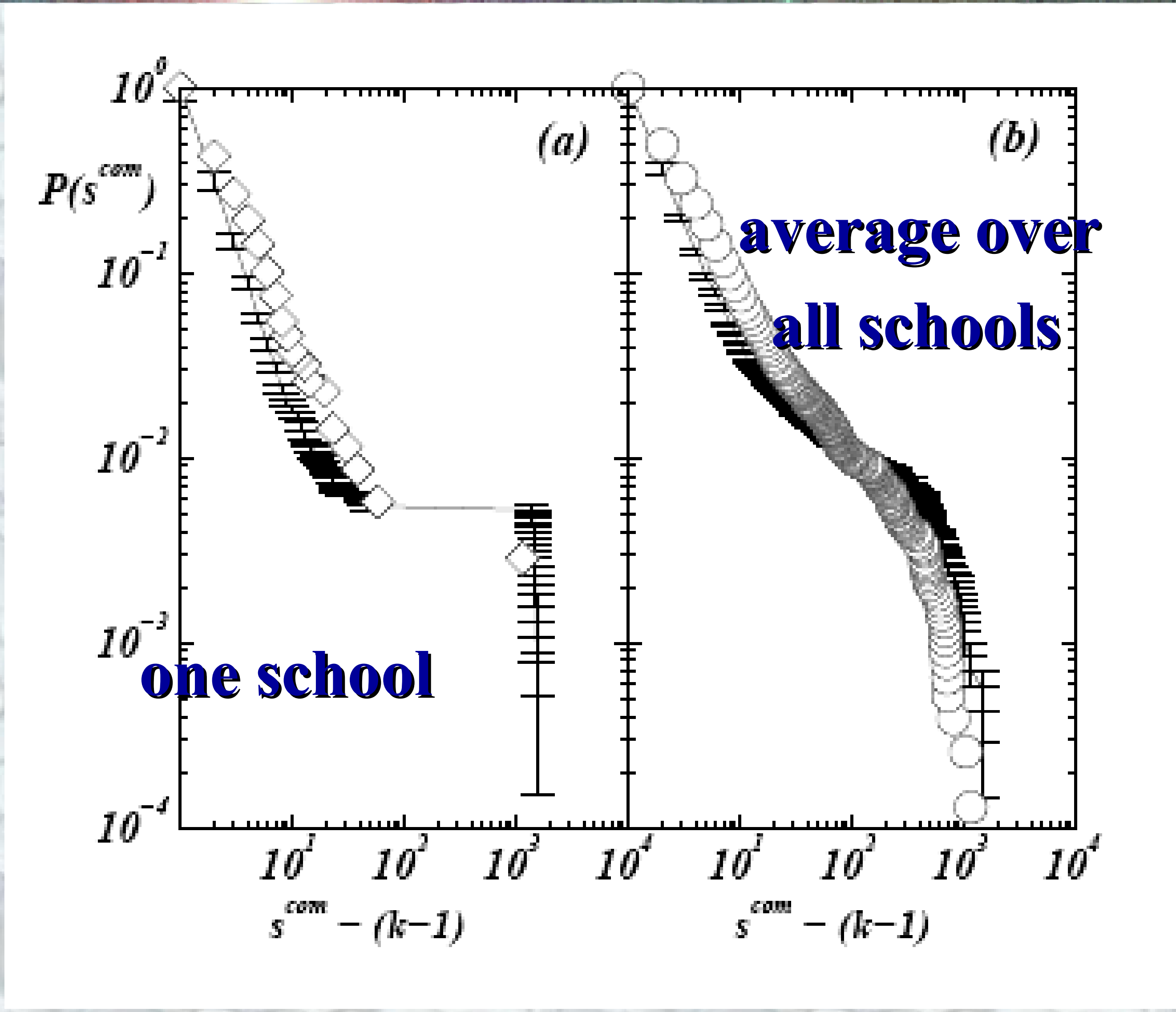
González, Lind and Herrmann, Physical Review Letters 96, 088702 (2006)

Comparison to schools

Degree distribution



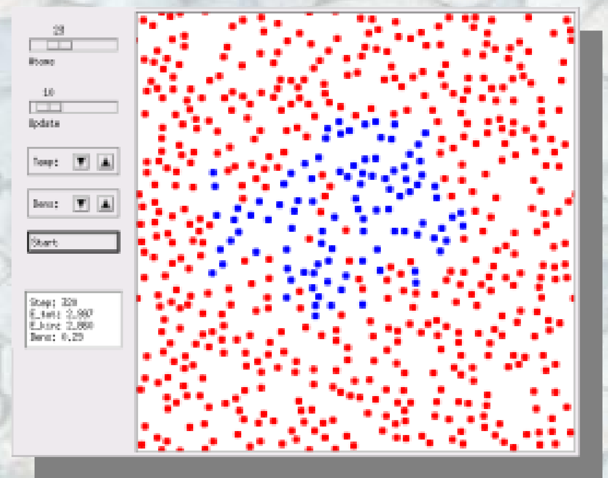
3-clique communities

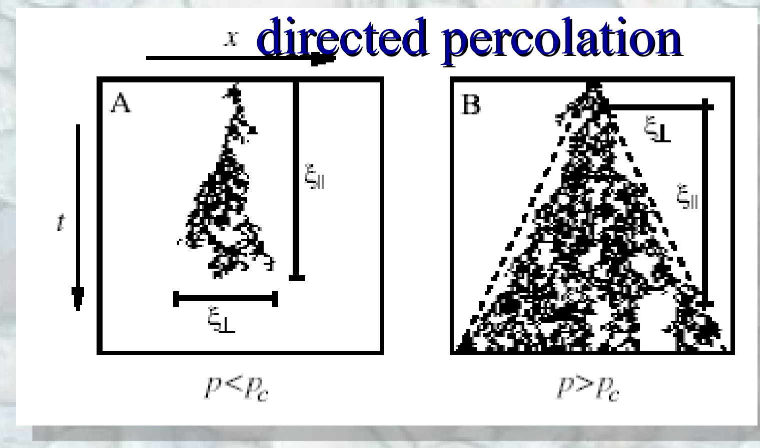


mbols:
hools

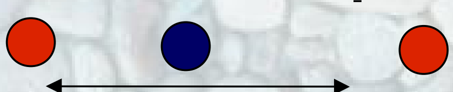
ll lines:
gent model

SIS model of epidemic





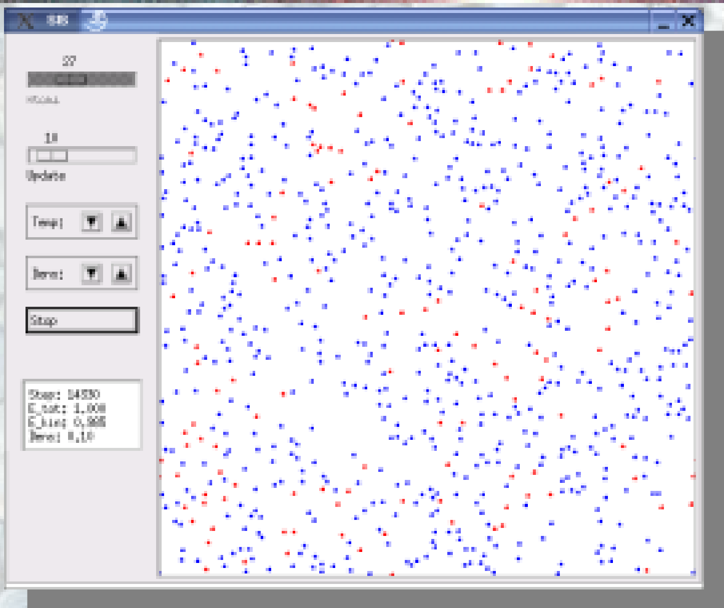
infection rate λ



Δt_{inf} = (MD time-steps)Δt

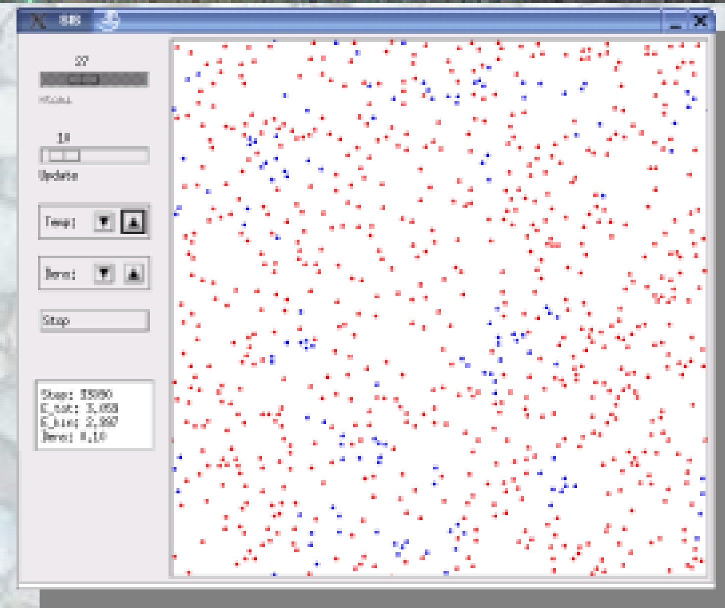
$$\lambda \approx \frac{\Delta t_{inf}}{\tau_{coll}} \propto \Delta t_{inf}\, \rho\, T^{1/2}$$

SIS model for epidemics

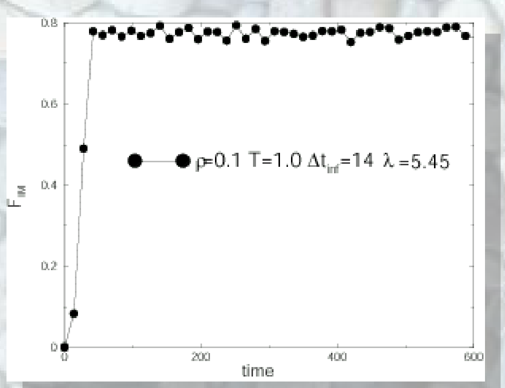


Susceptible ●

Infected ●

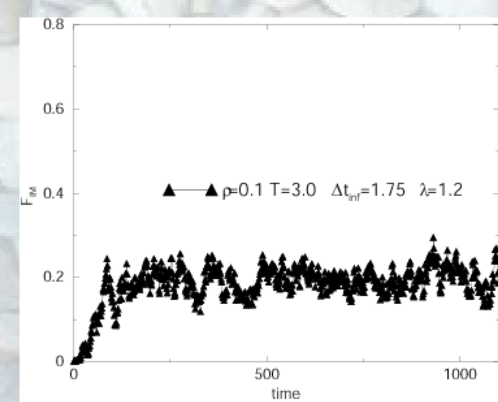


λ = 5.45



Steady State

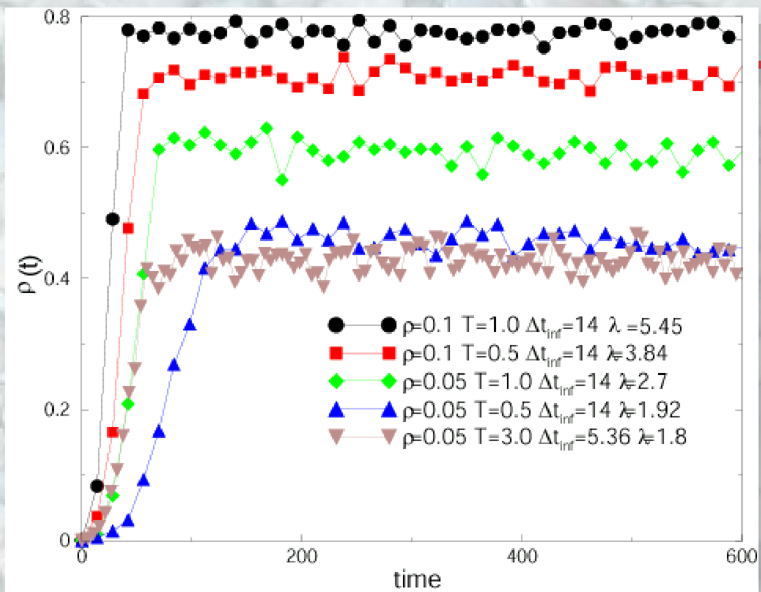
λ =1.2

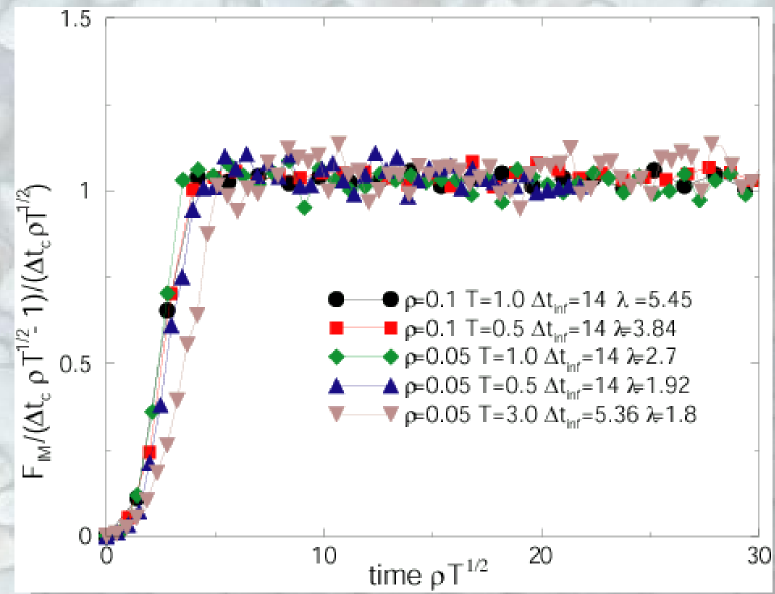


Data collapse

order parameter

$$F_{IM}(\lambda) \equiv \rho(t)_{\inf}\big|_{t\to\infty,\lambda>\lambda_c} = \begin{cases} 0 & \textbf{for } \lambda \le \lambda_c \\ 1 & \textbf{for } \lambda > \lambda_c \end{cases}$$

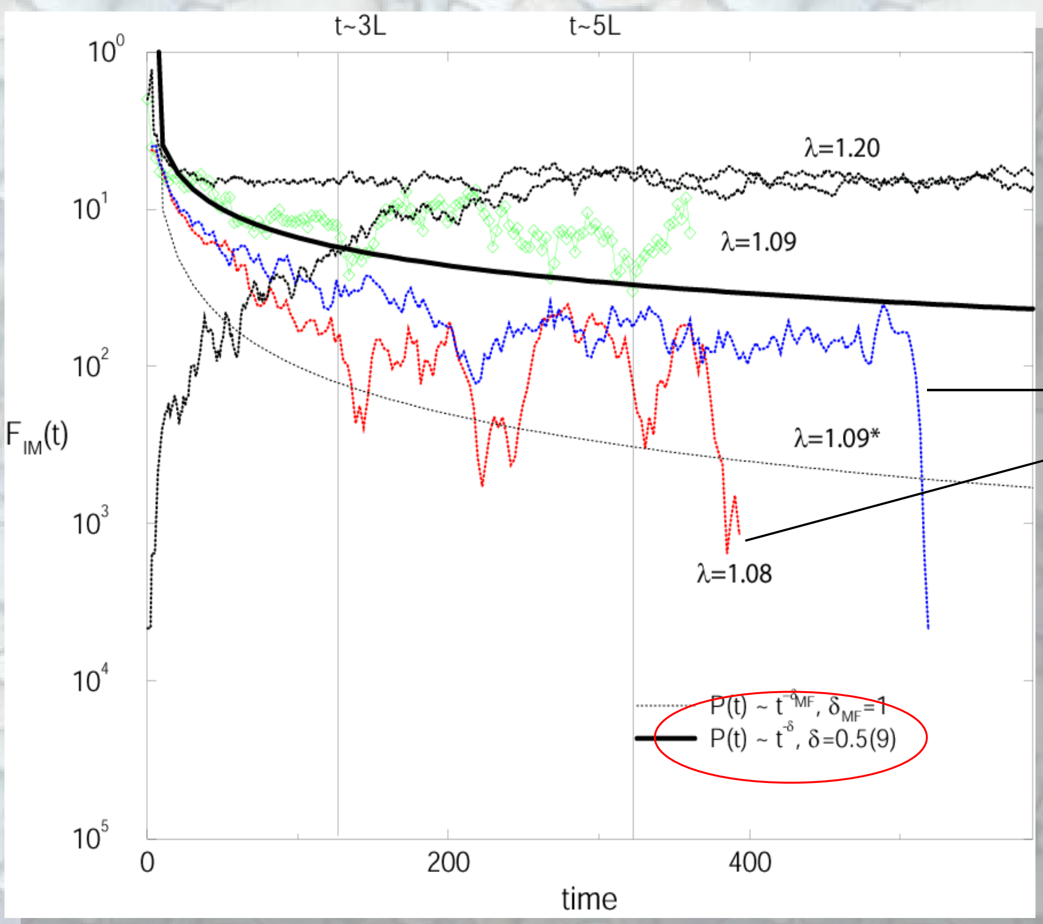




scaling law for $\lambda > \lambda_c$:

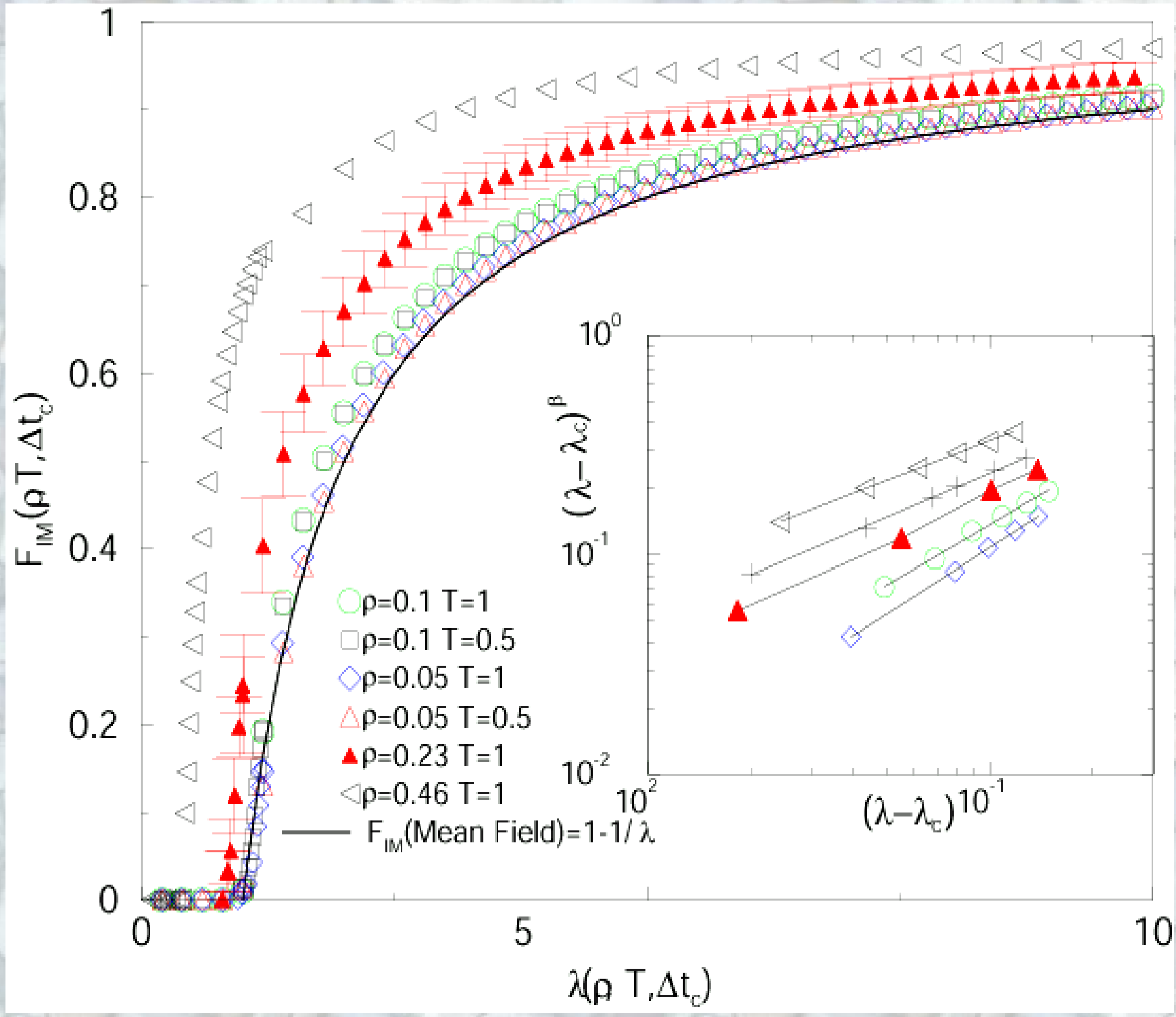
$$F_{IM}(\lambda) \approx 1 - \frac{1}{\lambda} \quad \textbf{with} \quad \lambda \equiv \frac{\Delta t_{\inf}}{\tau_{coll}} \propto \Delta t_{\inf}\, \rho\, T^{1/2}$$

Non-equilibrium phase transitions

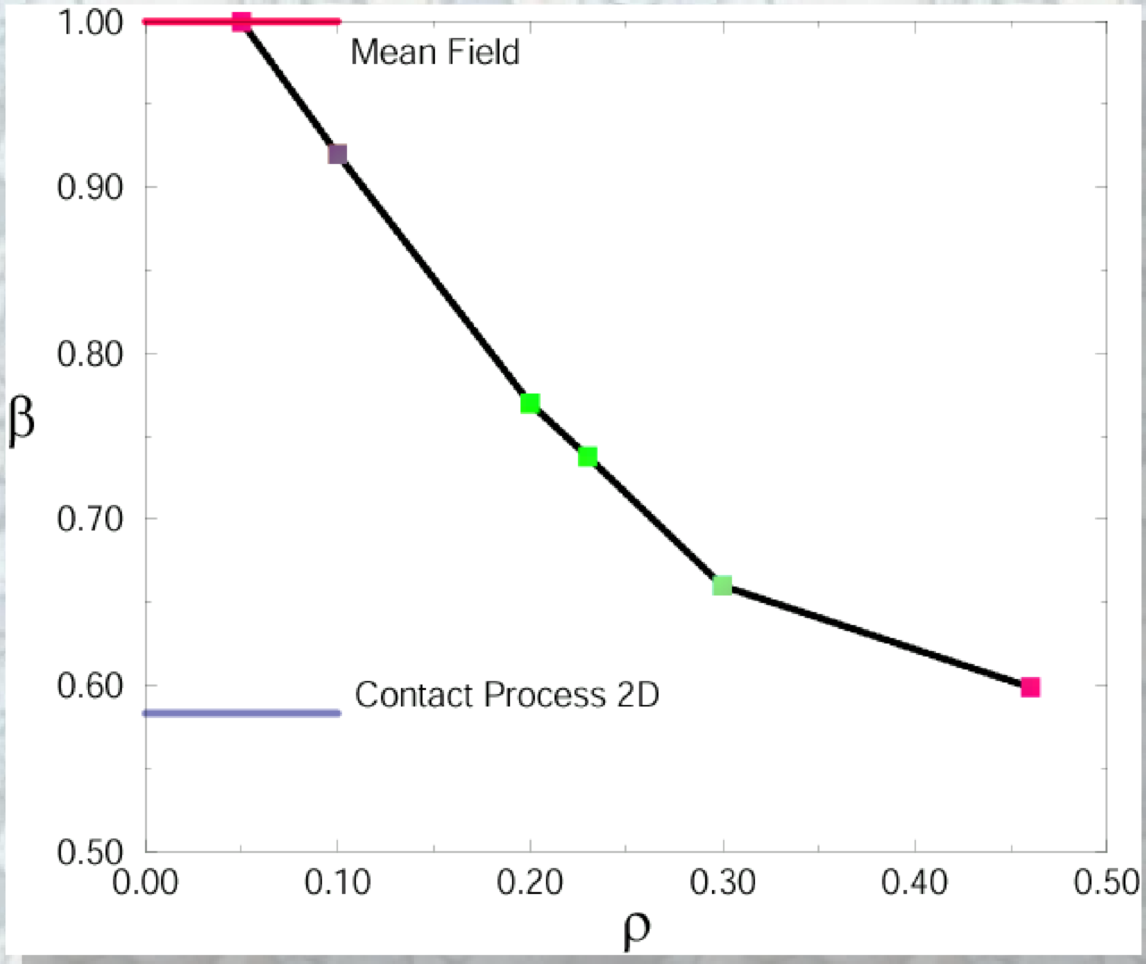


System goes into the absorbing state

Influence of the density



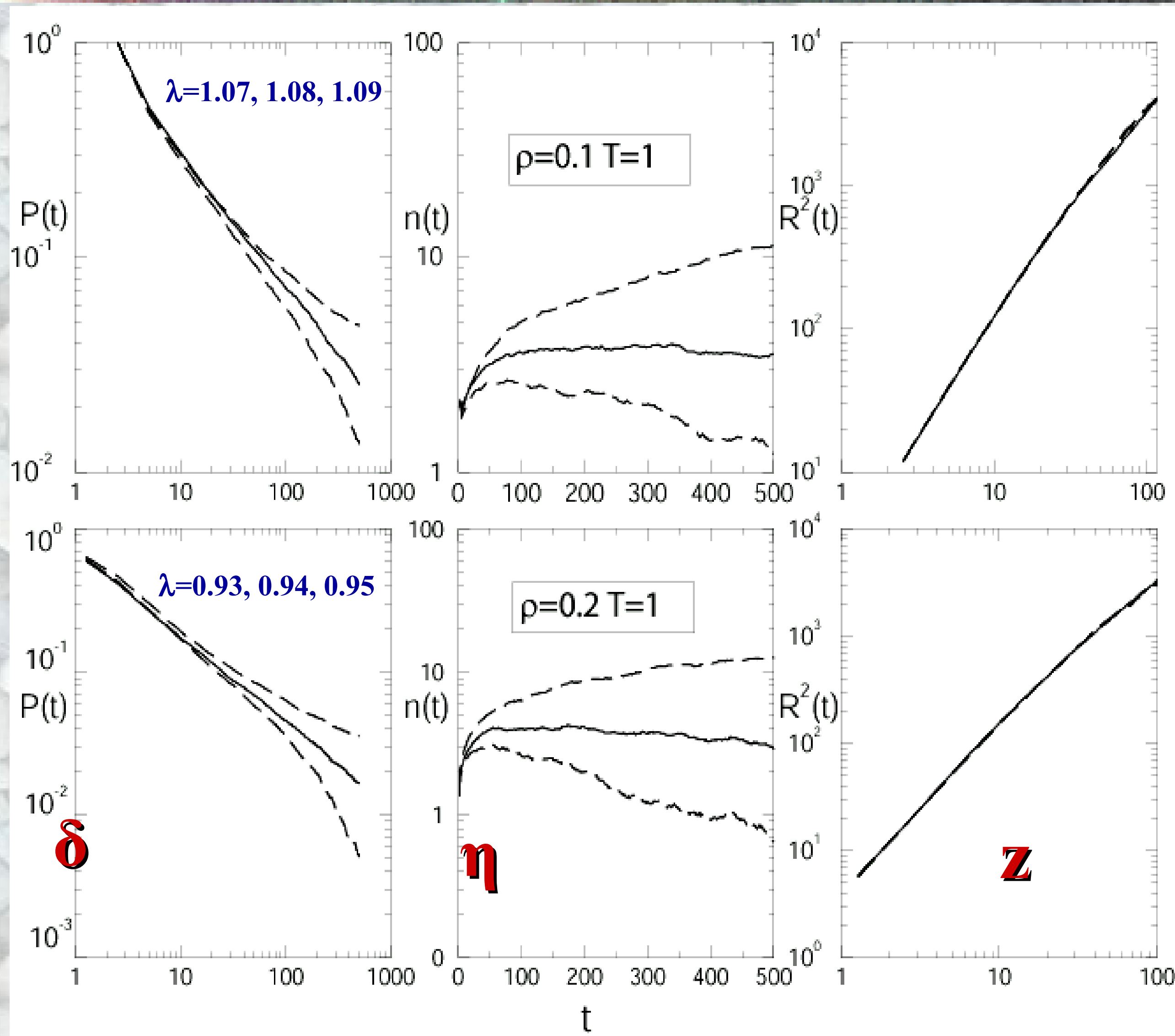
Critical exponent β



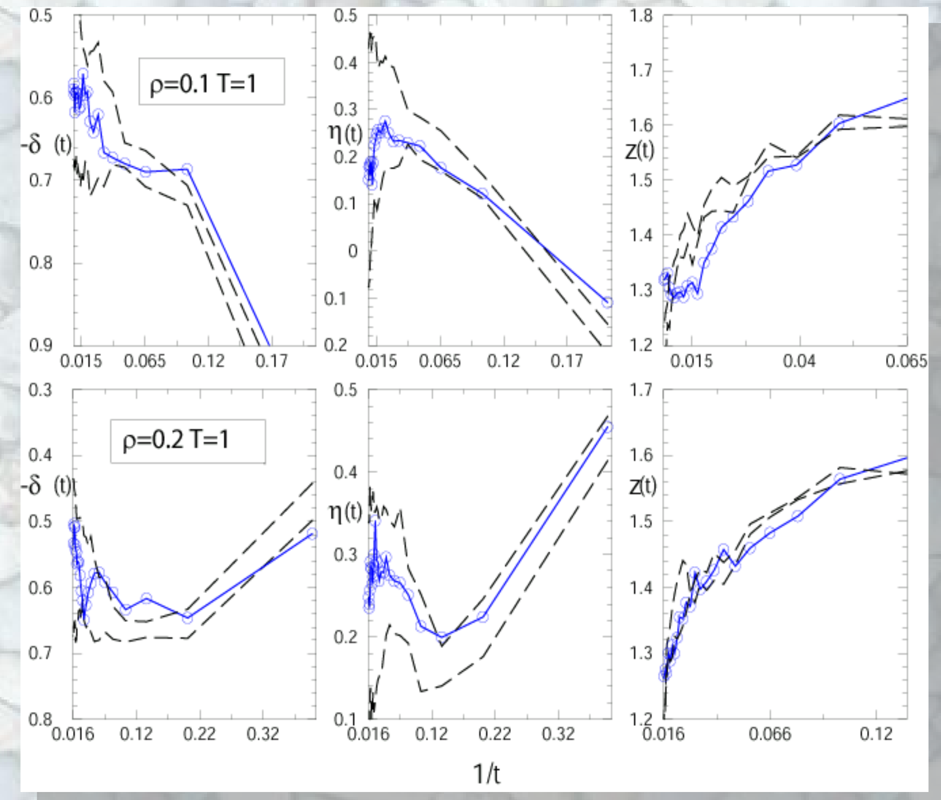
Other critical properties

(t) = probability
hat infection stays

(t) = average number
f infected agents



Critical exponents



$$\delta_t = \frac{\log[P(t)/P(t/\Delta)]}{\log \Delta}$$

$$P(t) \propto t^{-\delta}\left(1 + \frac{a}{t} + \frac{b}{t^{\delta'}} + \ldots\right)$$

$$\delta_t \propto \delta + \frac{a}{t} + \frac{\delta' b}{t^{\delta'}} + \ldots$$

- *P. Grassberger, J. Phys. A., (1989) 3673-3679*
- *J. Marro and R. Dickman, "Non-equilibrium Phase Transitions…", (1999) Cambridge University Press.*

Critical exponents

	Mean Field	Moving Agents $\rho=0.1$	Moving Agents $\rho=0.2$	Contact Process 2D
λ_c	1	1.0(8)	0.9(4)	1.6488(1)
β	1	0.9(2)	0.7(7)	0.583(4)
δ	1	0.5(9)	0.5(3)	0.4505(10)
η	0	0.1(5)	0.2(5)	0.2295(10)
z	1	1.3(0)	1.2(7)	1.1325(10)

$4\delta+2\eta=dz$ holds.

Networks properties

Labeling of Cluster sizes

Susceptible ○ RGB scale for cluster size

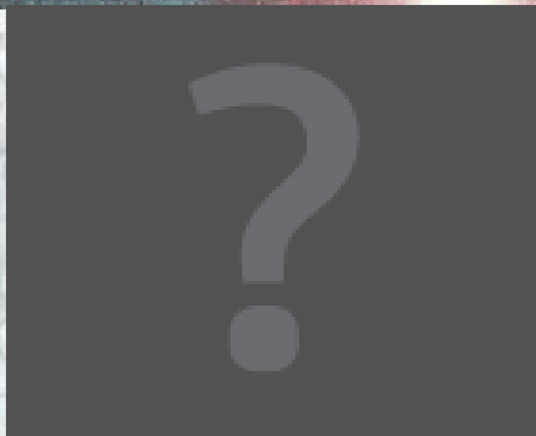
density = 0.1 T = 1, λ = 2

density = 0.23 T = 1, λ = 2

Rules for generating the network

- Each time one agent infects other a bond between the two is created.
- The infection lasts (Δtinf time steps), when one of the agents is recovered the bond disappears and the agent becomes susceptible to infection again.

Cluster size distribution



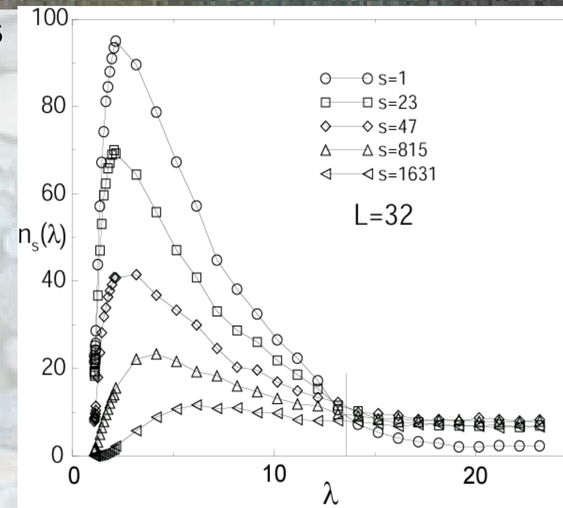
Susceptible ◯

Clusters size

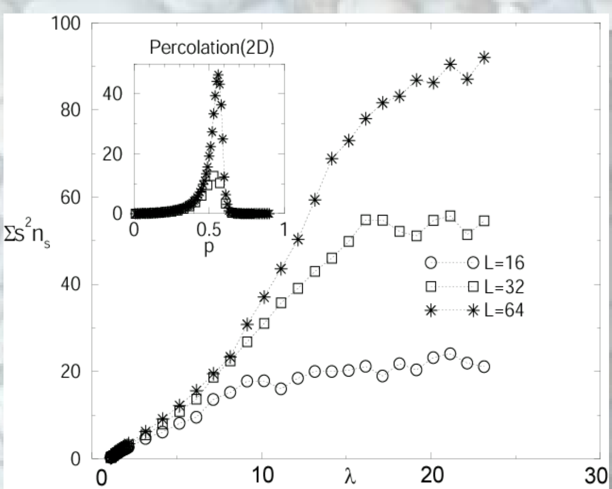
N =16, $\rho = 0.1$, $\lambda = 10$

same scaling as in percolation

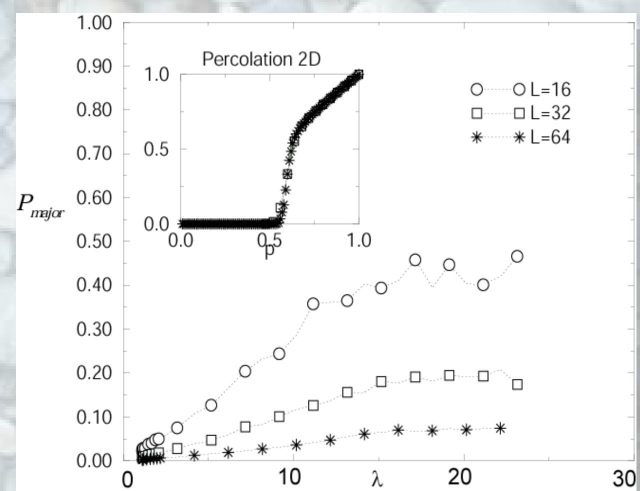
Cluster numbers



Second moment of the cluster size distribution

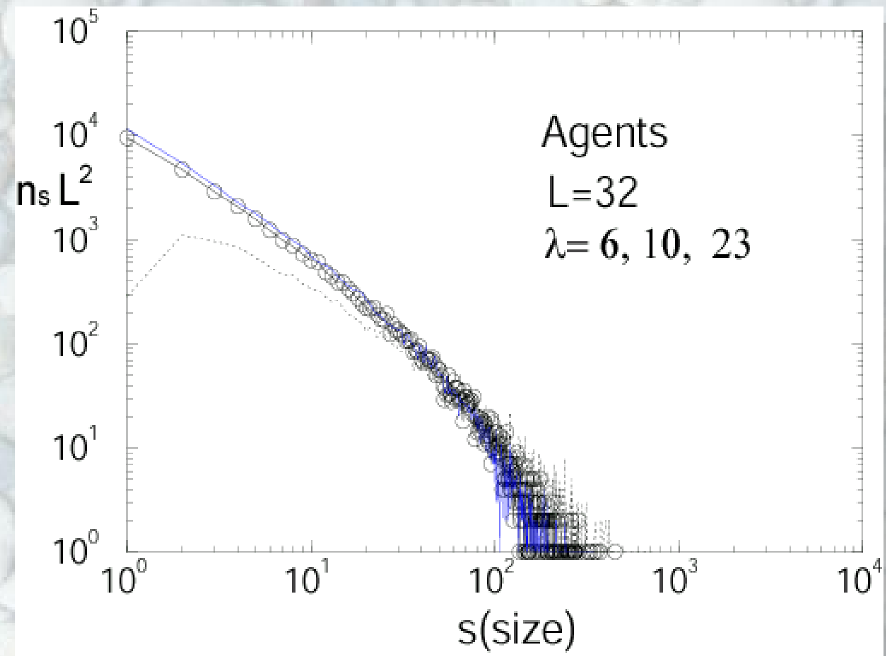


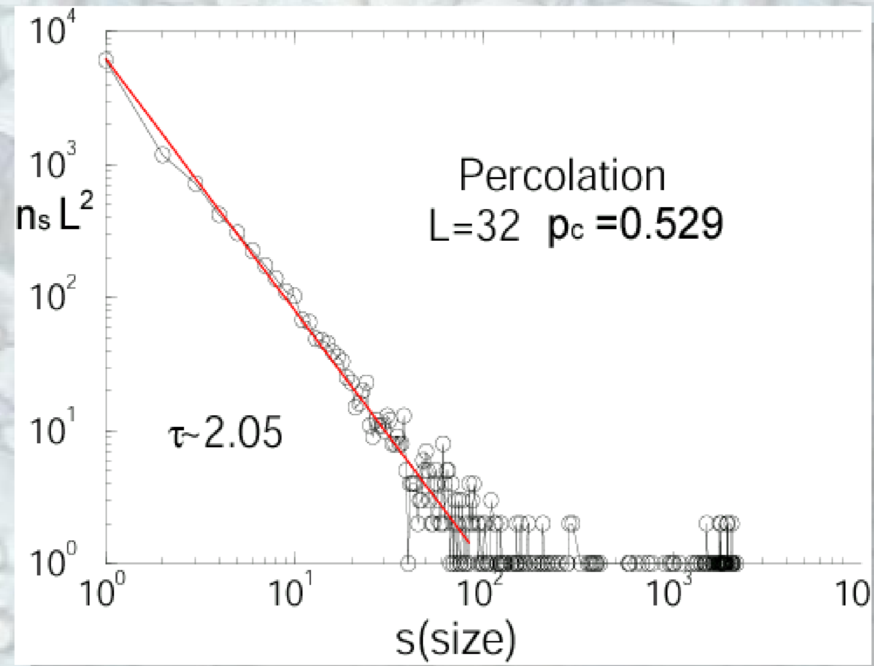
Probability that an agent belongs to the biggest cluster



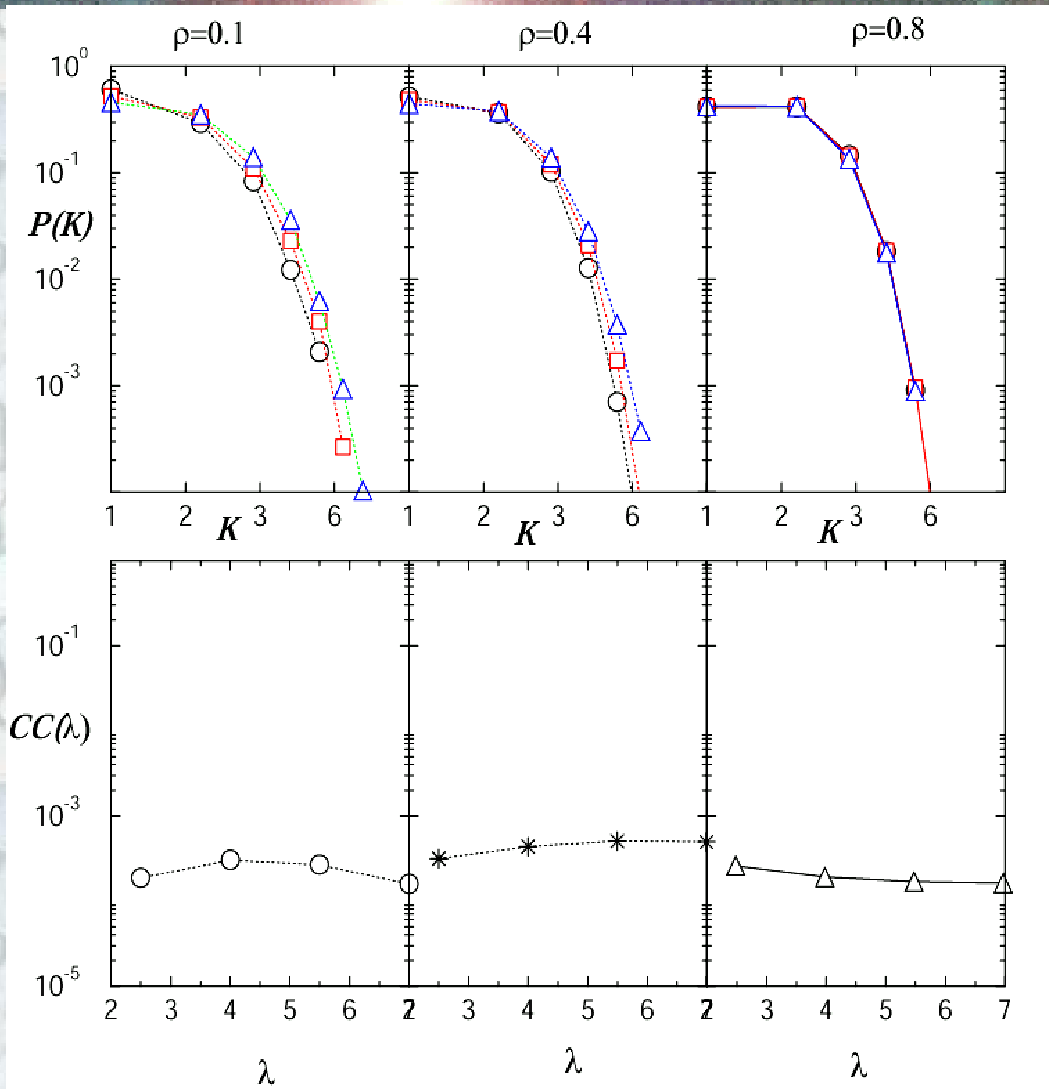
In Collab. with:
A.D. Araújo,
UFC, Brazil

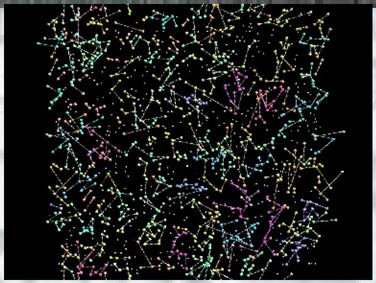
Cluster numbers



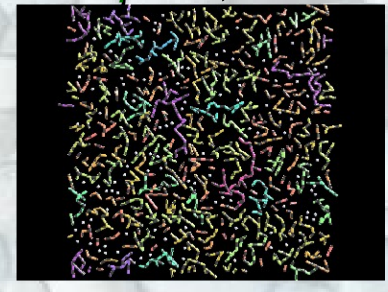


Clustering coefficient and degree distribution

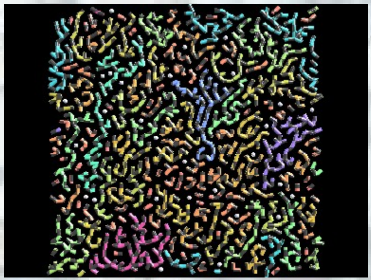




ρ=0.1, λ=3



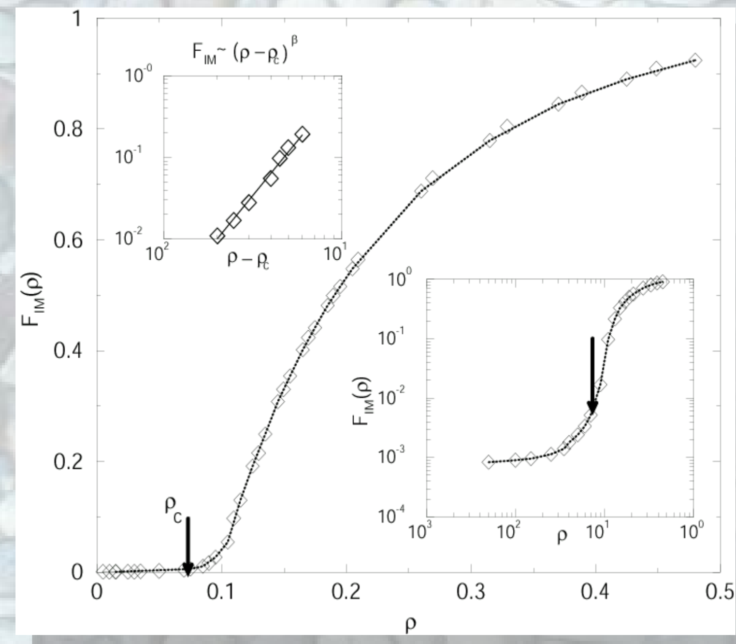
ρ=0.4 λ=3



ρ=0.8, λ=3

Power-law distribution of infection time

$\rho = 0.1, \gamma = 2.4$



$$P(\Delta t_{\inf}) = (\gamma - 1)\Delta t_{\inf}^{-\gamma} \qquad 2 < \gamma \leq 3 \qquad \Delta t_{\inf} > 1$$

β=2.6(5)

$$\lambda_c \equiv \frac{\langle t_{\inf} \rangle}{\tau_{coll}} = \frac{\gamma - 1}{\gamma - 2} \rho \, 2 r_o \sqrt{\frac{\pi T k_B}{m}} > 1$$

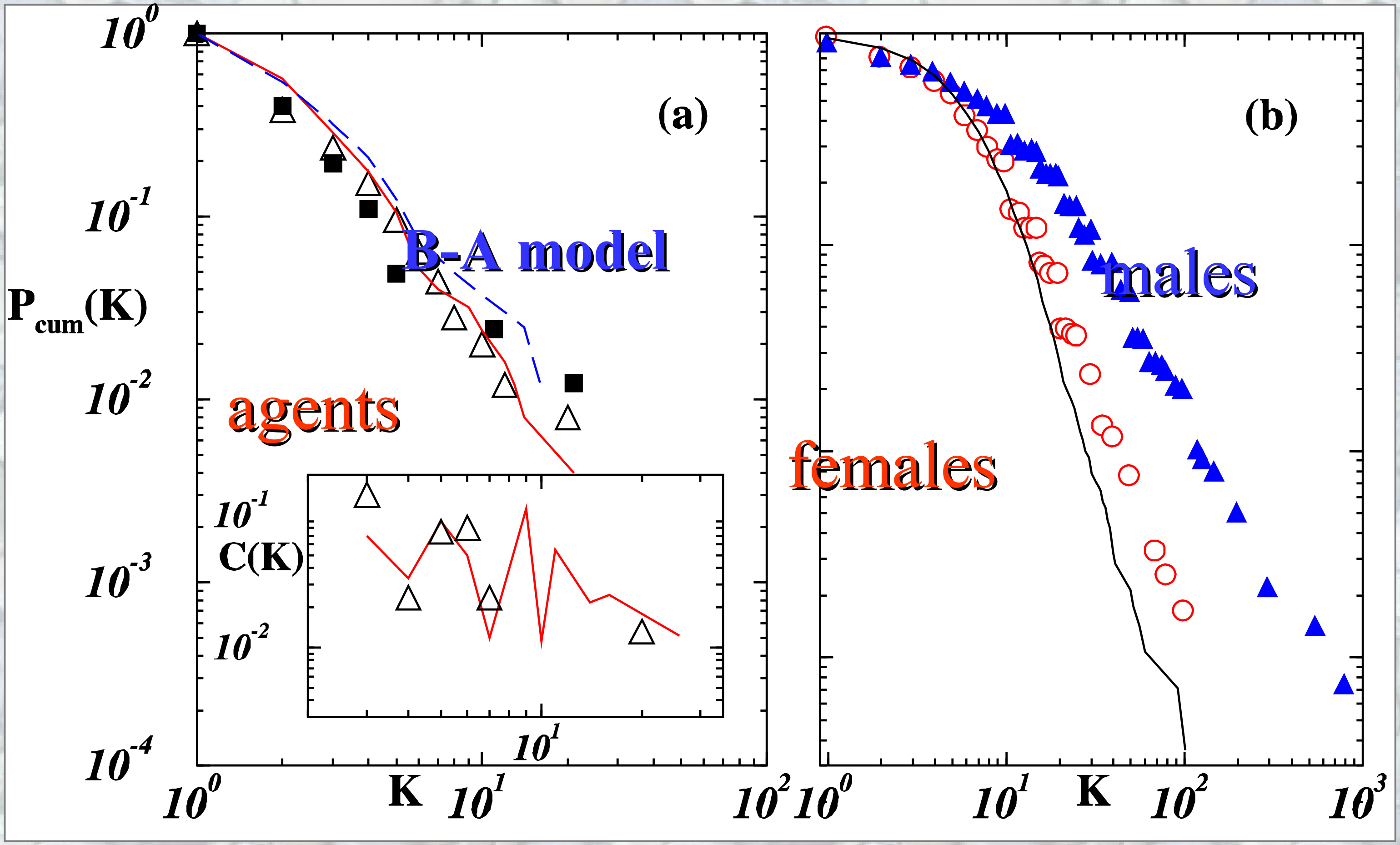
$$\rho_c\big|_{calc} = 0.00718 \qquad \rho_c\big|_{num} = 0.006(5)$$

Conclusions on epidemics

- ✓ Novel effects are observed studying the SIS model of infection on a system of moving agents. A continuous range of critical exponents is found as a function of the density number of agents.
- ✓ Geometrical properties like the degree distribution and the clustering coefficient of the network of infected agents depend on the rate of infection and not on the density of agents. The distribution of cluster sizes is not a power law at the transition to spreading.
- ✓ Introducing a power law distribution of infection times the epidemic threshold becomes zero, but there is still a critical rate of infection that depends on the exponents of the distribution and the mean interaction time among the agents .

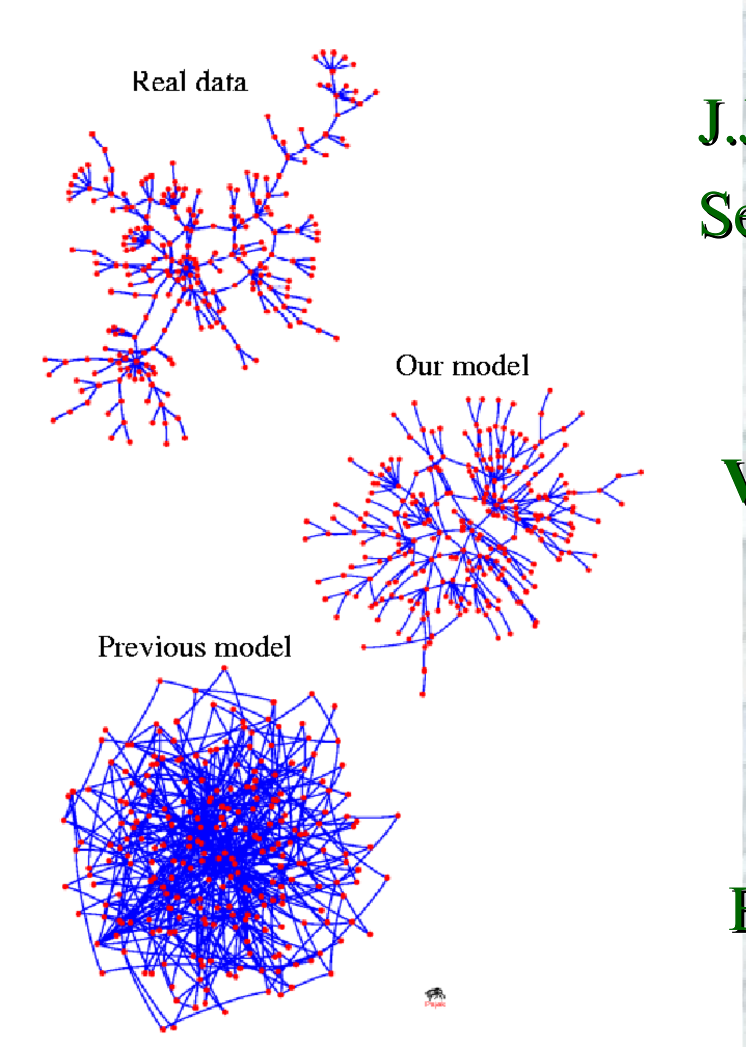
Sexual contact networks

homosexual (Colorado Springs) **heterosexual** (Sweden)



F. Liljeros et al, Nature **411**, 907 (2001)

Comparison with real data



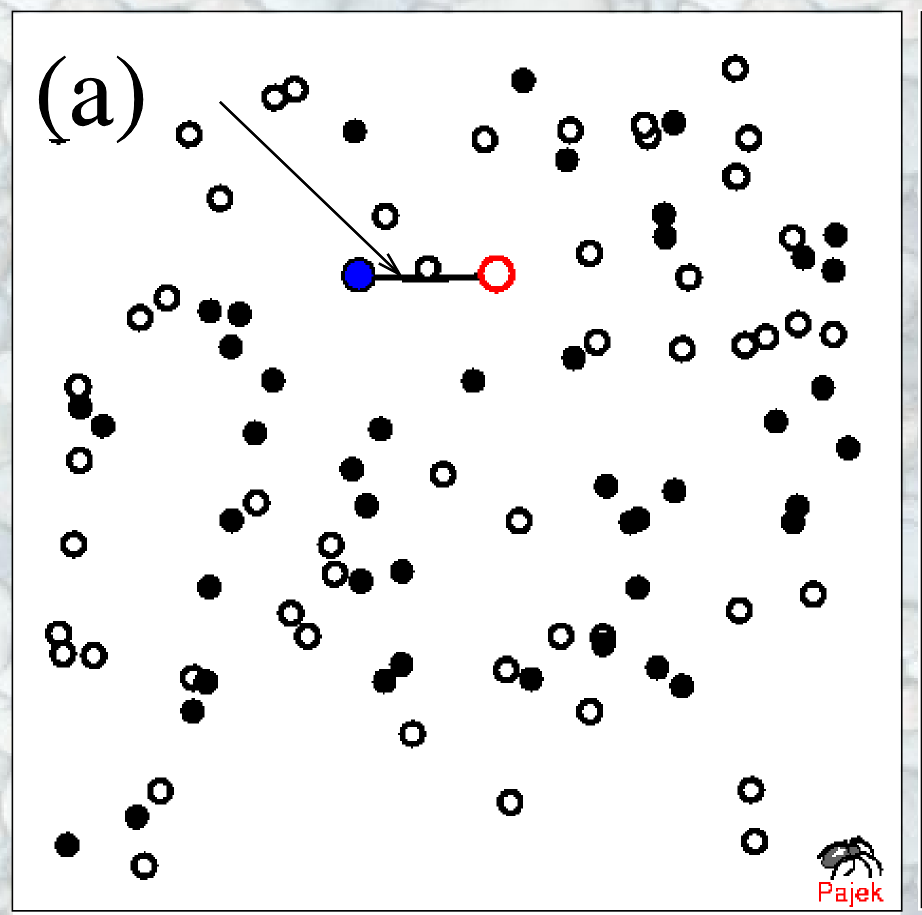
J.J. Potterat et al,

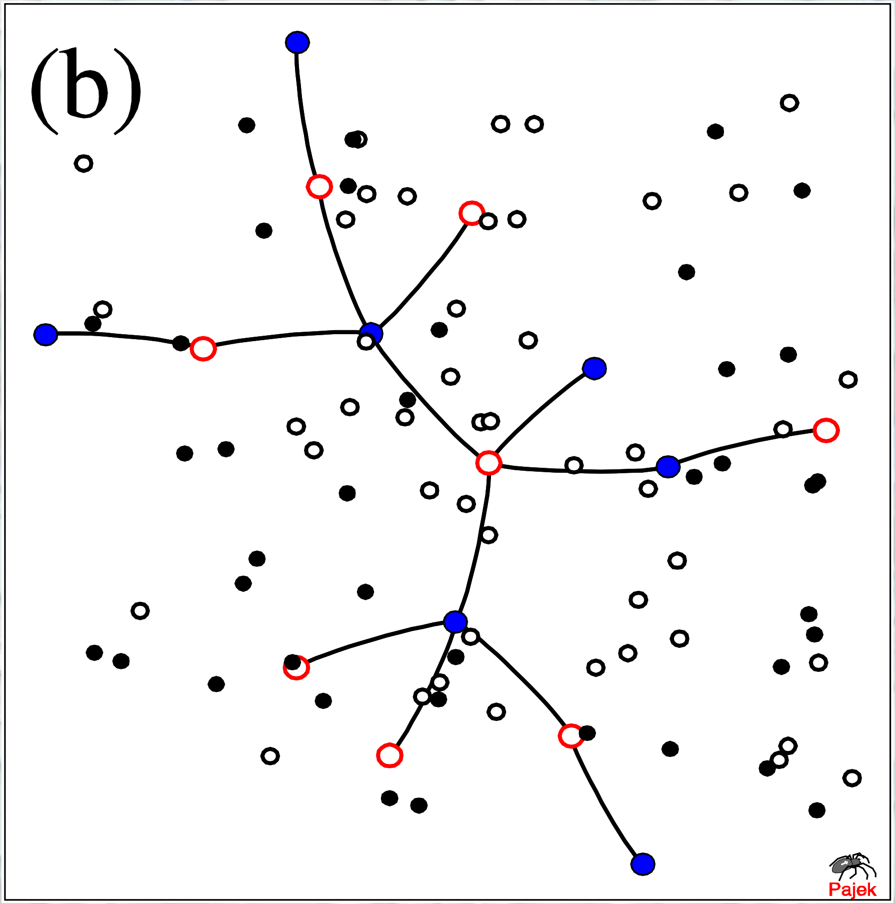
Sex.Transm.Infect. **78**, 59 (2002)

Velocity of agent proportional

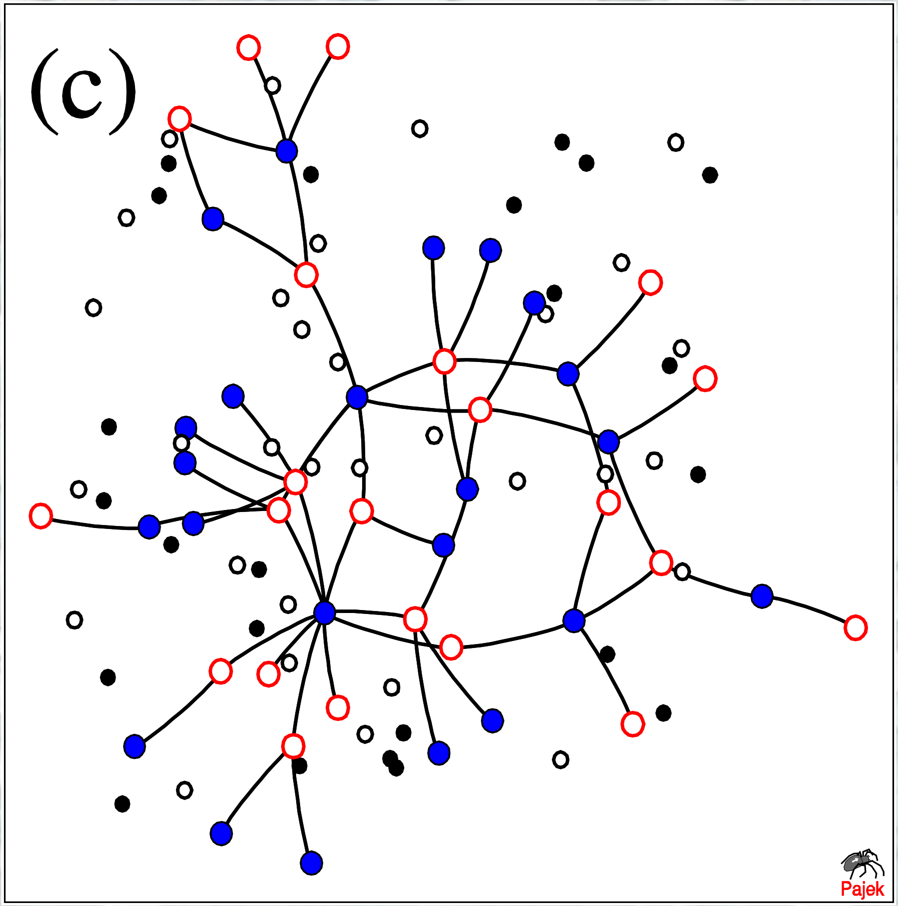
Barabasi-Albert scale-free network

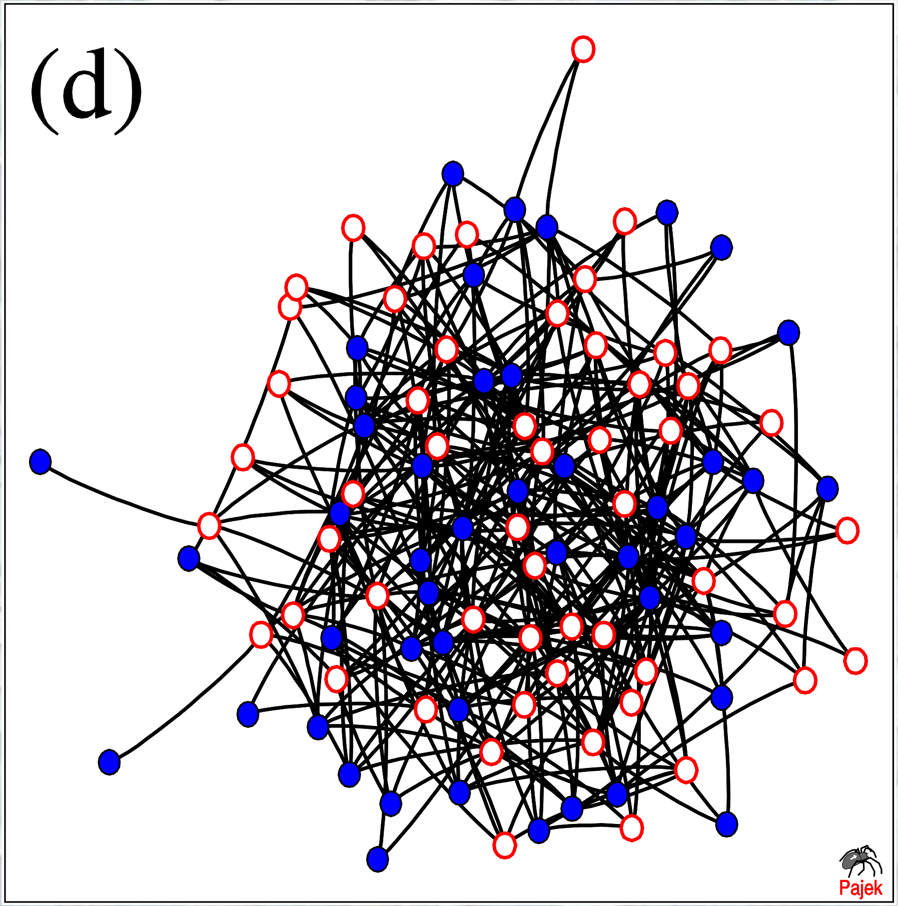
Bipartite growing network





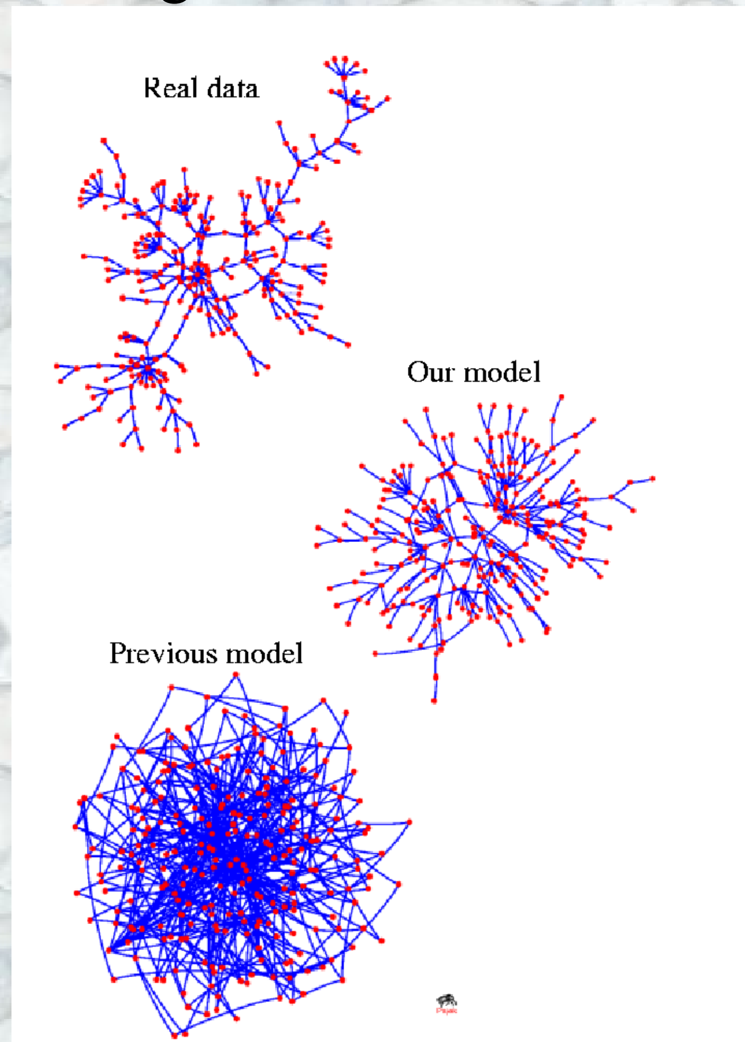
Bipartite growing network



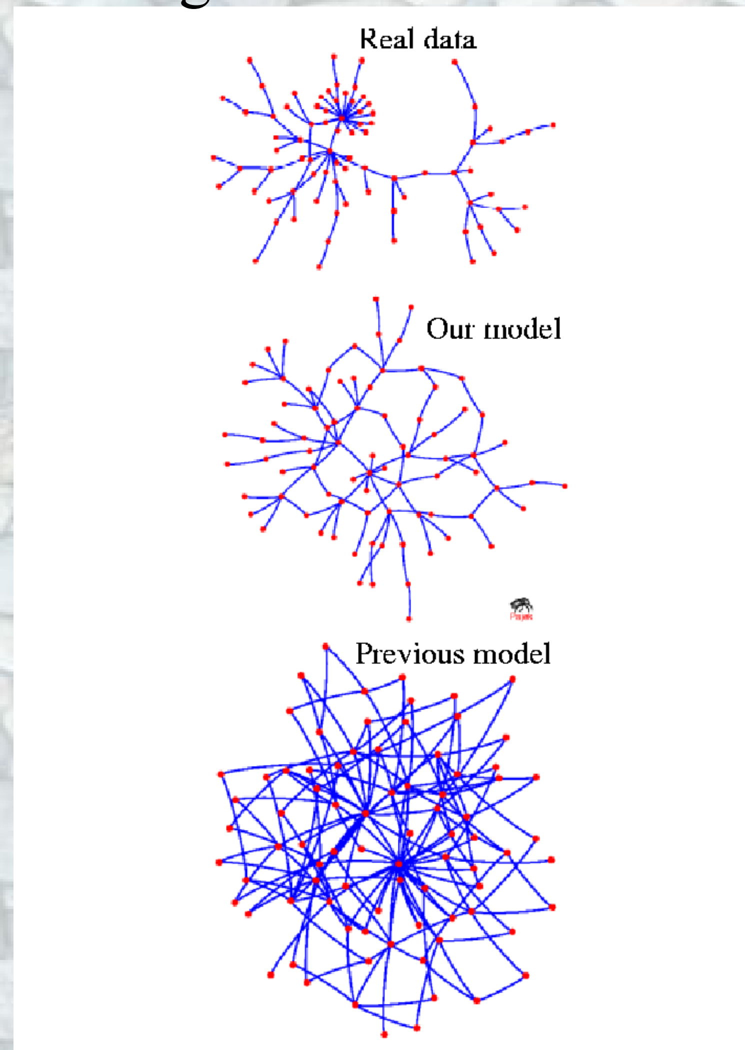


Comparison with real data

homogeneous:

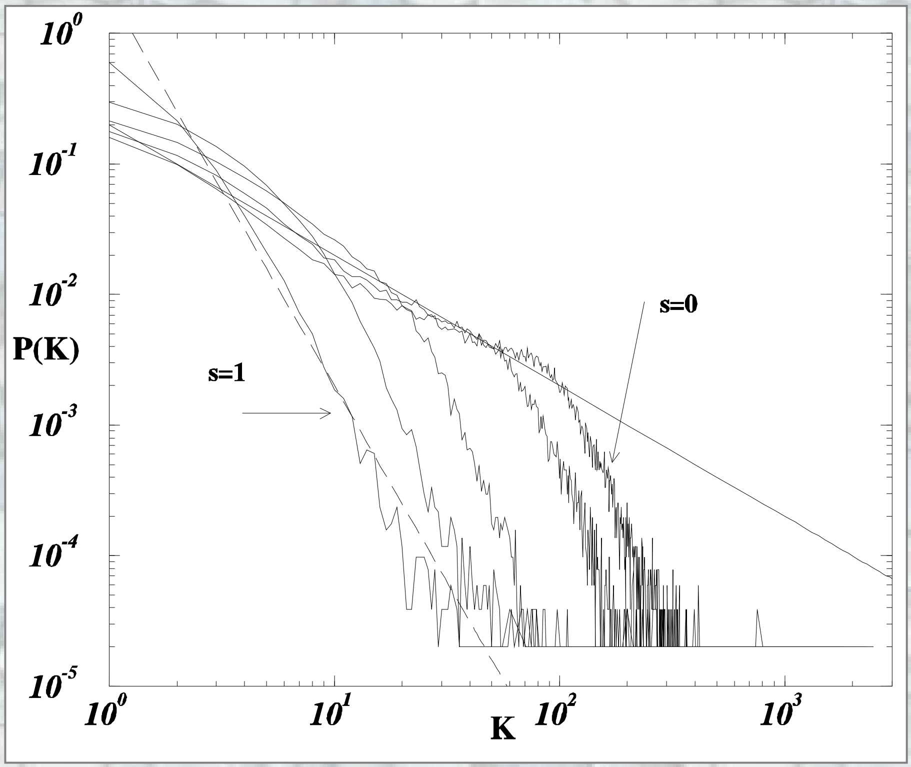


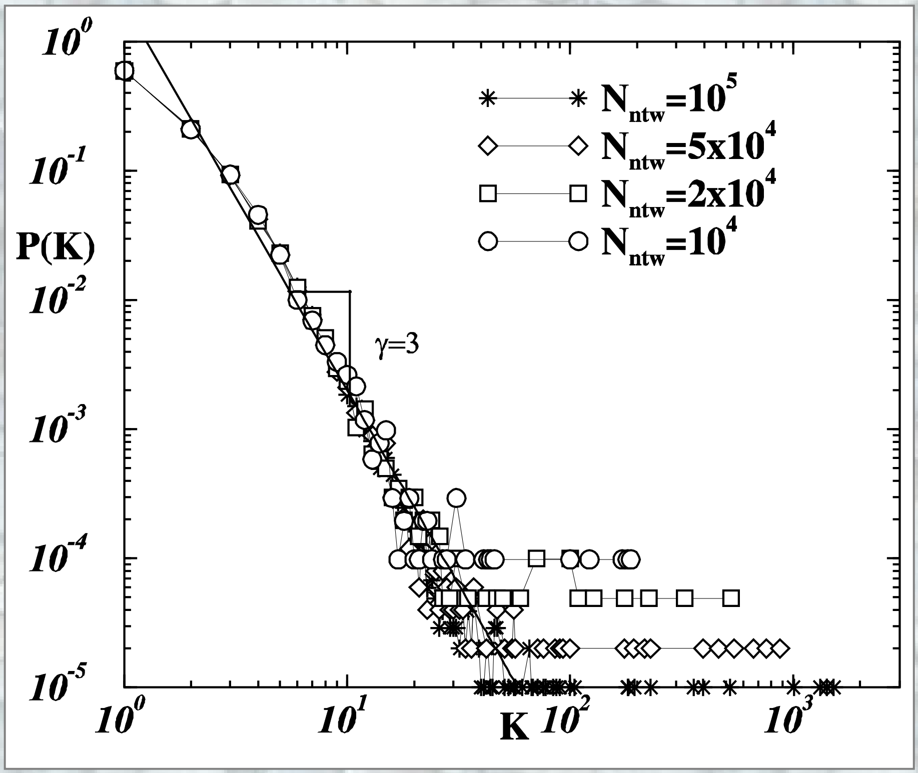
heterogeneous:



Sexual contact networks

Degree distribution



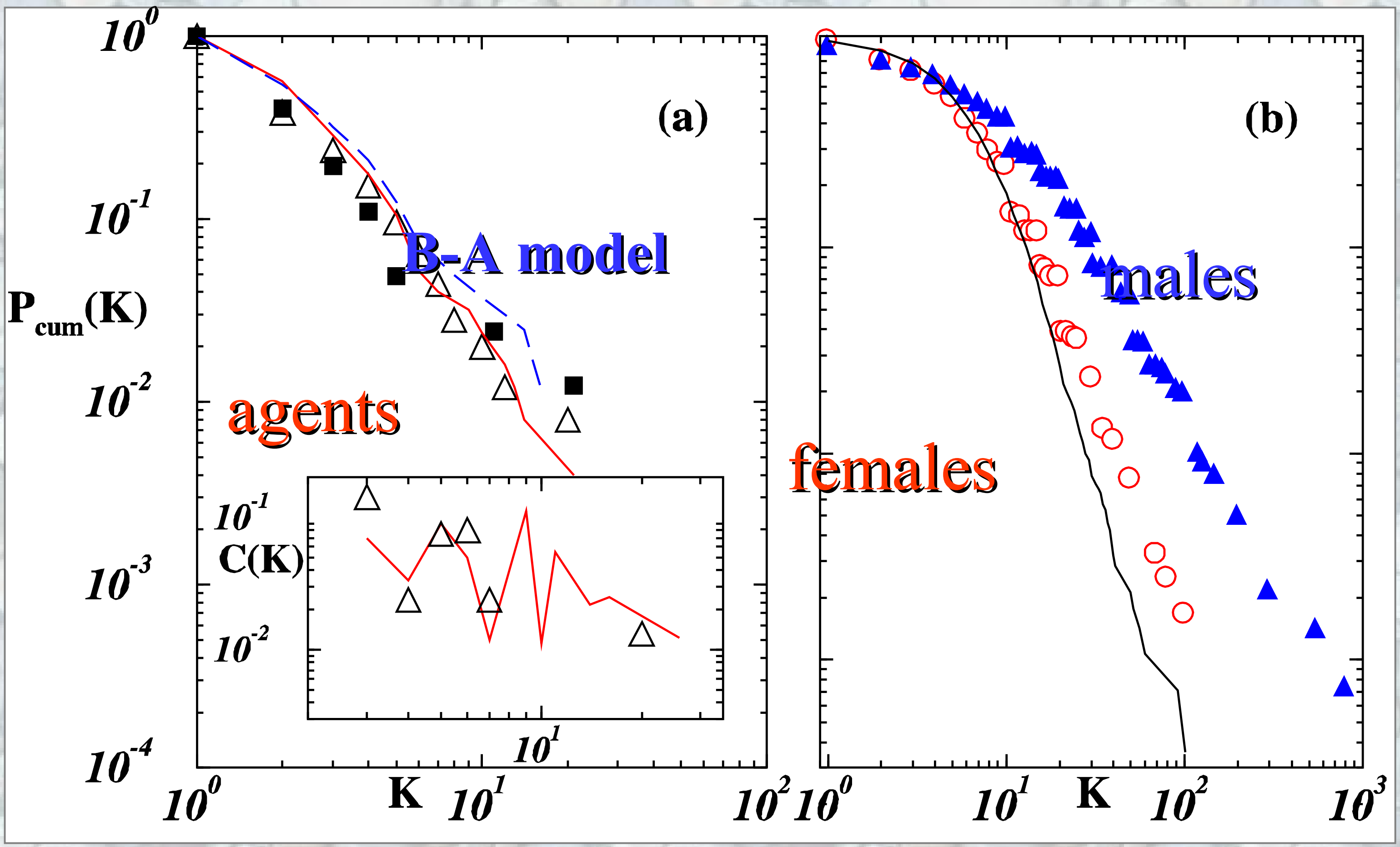


Sexual contacts with any agent.

Sexual contacts only with uninitiated agents.

Sexual contact networks

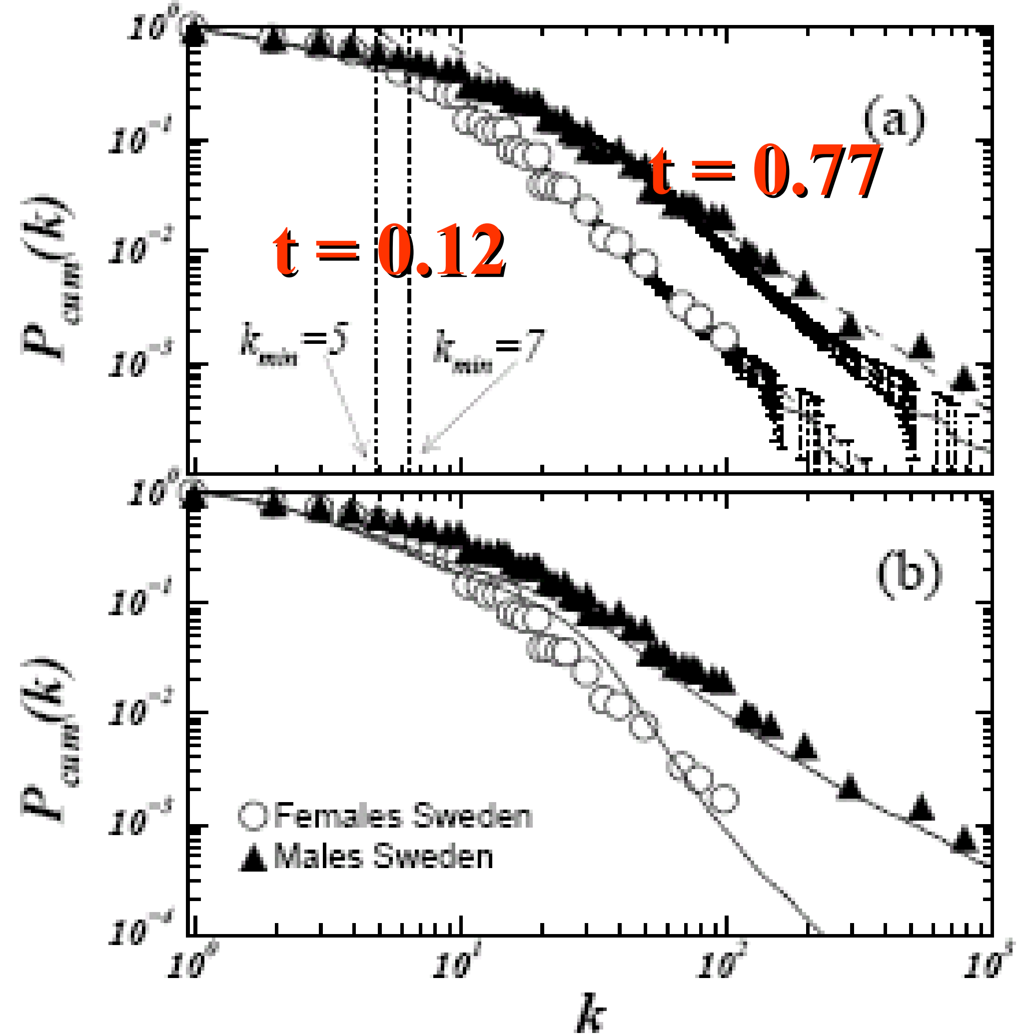
homosexual (Colorado Springs) **heterosexual** (Sweden)



F. Liljeros et al, Nature **411**, 907 (2001)

Sexual contact networks

heterosexual

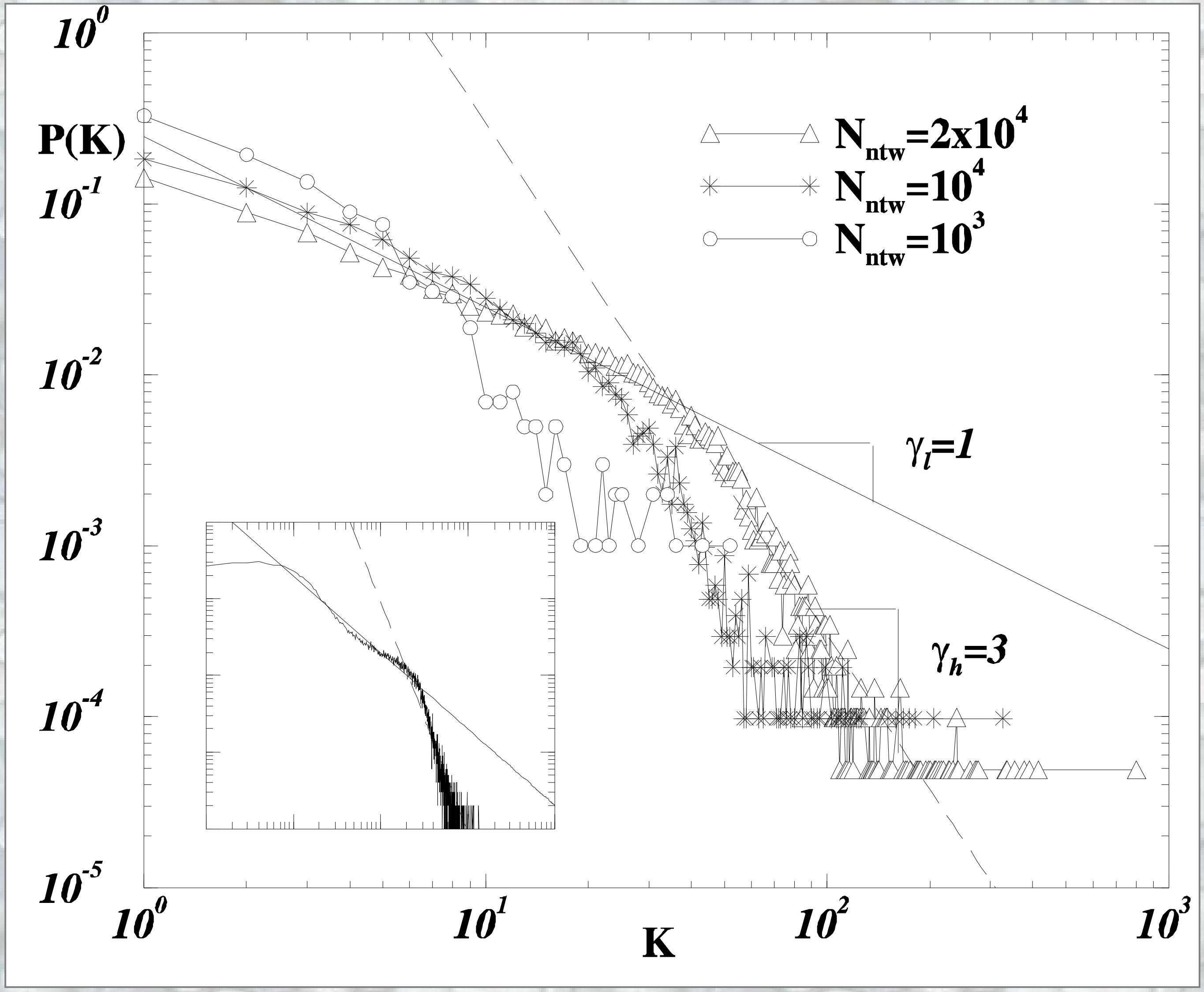


F. Liljeros et al, Nature **411**, 907 (2001)

58% females

42% males

Sexual contact networks



Another friendship model

- **Goal**
  - friendship network model (spatially independent) for a fixed setting (e.g. enrolling at a university)
  - reproduce experimentally measurable quantities
  - natural emergence of community structures

Herman Singer, ETH

Properties of the agent

- agent properties
  - affinity $a_i \in [0,1]$
  - list of past contacts, length n_i
  - maximal acquaintance parameter λ_i
  - absolute importance

$$\prod_i = \frac{k_i}{\sum_{j=1}^{N} k_j}$$

  - relative importance

$$p_{ij} = \frac{f_{ij}}{\sum_{k=1}^{n_i} f_{ik}}$$

- behavior
  - agent can choose between meeting new contacts and meeting again already established contacts
  - similar to aging but important difference:
    - agent can accept new contacts throughout the simulations
    - at the expense of dropping old ones

→ every agent optimizes its interest

→ local, self organized community structure emerges

Friendship in formulas

- friendship is written as the weighted sum of the contributions
  - number of times n_{ij}
  - interest/affinity a_i, a_j

$$f_{ij} = \gamma\, f_1(|a_j - a_i|) + (1-\gamma)\, f_2(n_{ij})$$

  - with $\gamma > 0.8$

friendship functions:

- affinity f_1

$$f_1(|a_j - a_i|) = \left(1 - |a_j - a_i|\right)^{\kappa}$$

  - with $\kappa \geq 2$, here 2,3,5,10

- frequency f_2
  - exponential

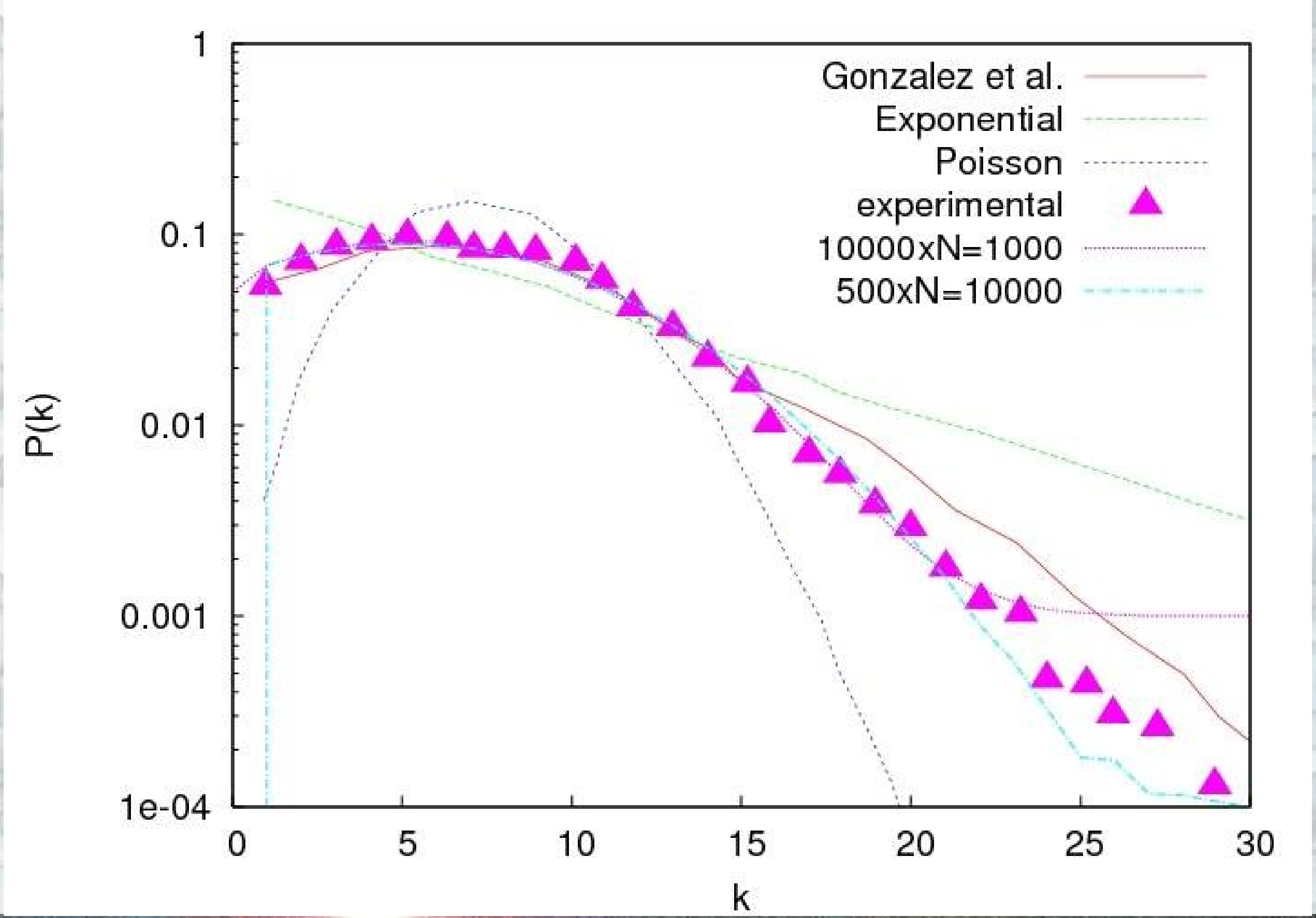
$$f_2(n_{ij}) = 1 - e^{-\lambda_f n_{ij}}$$

  - with $\lambda_f = 0.2 \leftrightarrow$ saturation on ~25 contacts

Algorithm

- $t=0$: initialize N agents, no connections
  with parameters a_i, $n_i=0$, λ_i
- $t \rightarrow t+1$:
  - choose an agent randomly: i
  - decide on contact mode with probability $p = 1 - e^{-\lambda_i n_i}$
    - select an agent from pool with probability Π_j
      - add to contact list
    - otherwise select an agent in contact list with propability p_{ij}
- prevent social isolation and take into account random encounters:
  - introduce threshold $p \geq \theta_p$

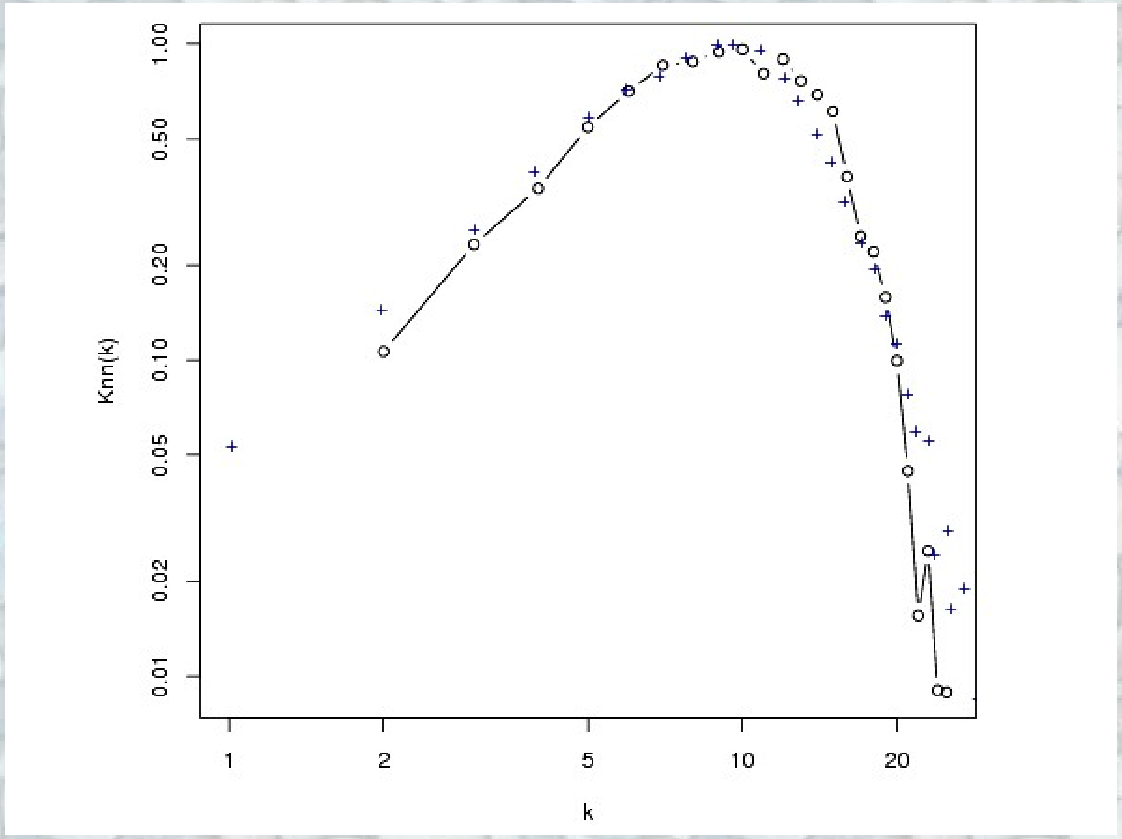
Degree distribution



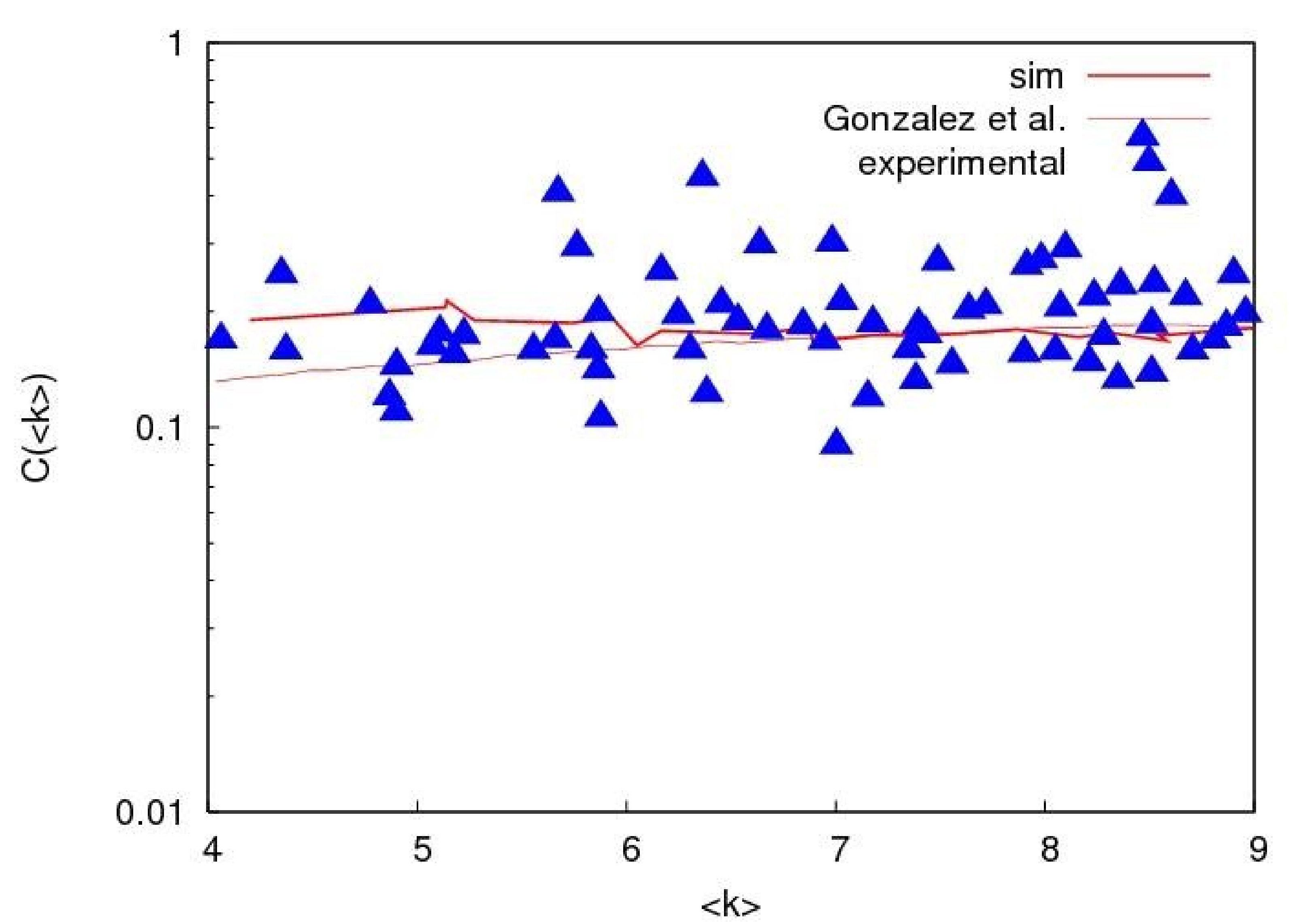
Degree correlation

- P(k|k'): probability at a node with k to find a neighbor with k'

$$\rightarrow \quad K_{nn}(k) = \sum_{k'=1}^{N} P(k \mid k')k'$$

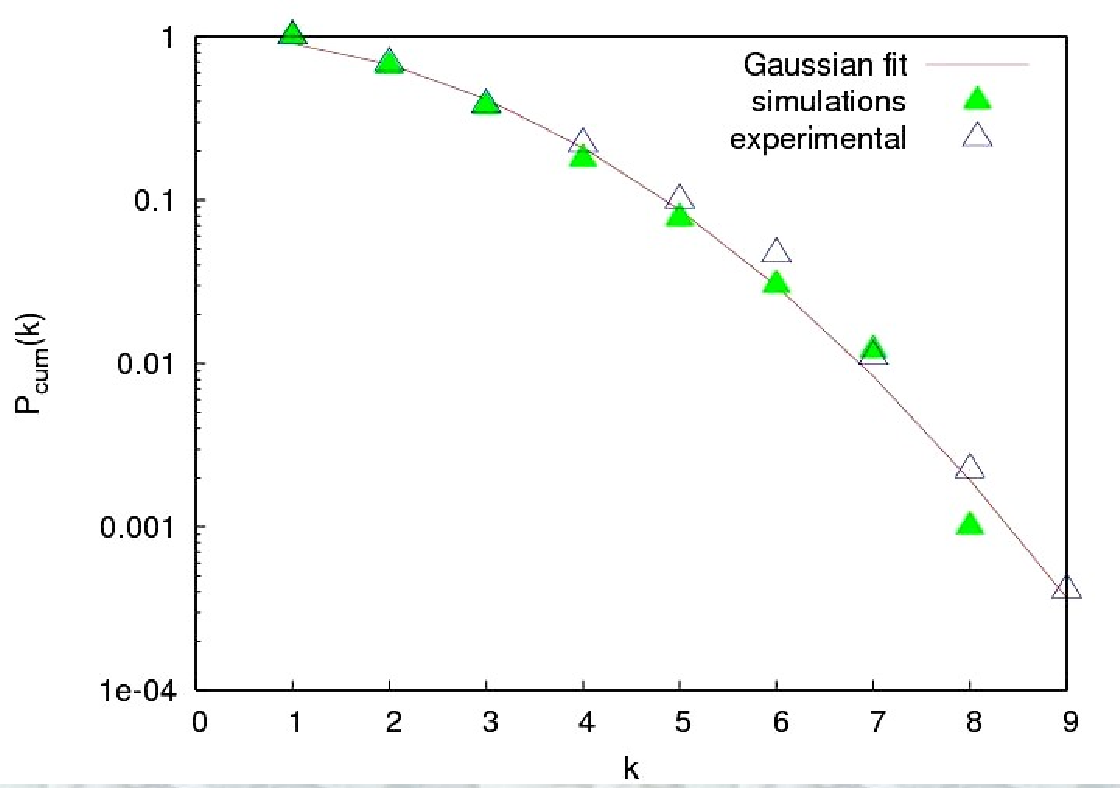


Clustering coefficient



Friendship distribution

- Experimental results
  - 417 high school students
  - "who are your best friends?"
  - probability that a person is mentioned n times



Conclusion

- Social relation is difficult to quantify.
- Many simplified models are possible.
- Parameters can be tuned to agree with some data.
- One can find general conclusions.
- But does anybody believe these generalities?
- Does anybody care about this research?
- Do we need socio-physics?
- What does one learn from modeling?