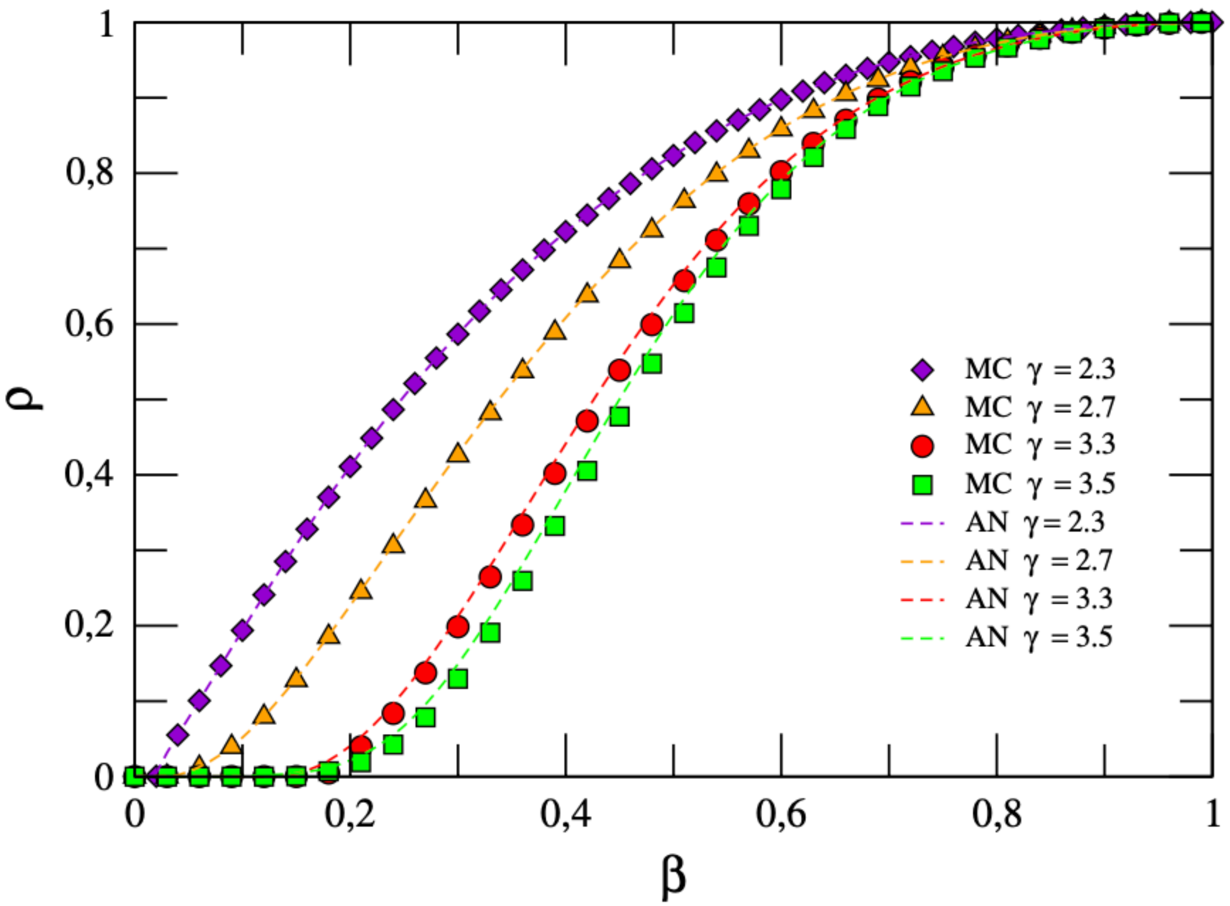
Physics of Life Data Epidemiology

Lect 12: Network epidemiology 2

Chiara Poletto
mail: chiara.poletto@unipd.it
web: chiara-poletto.github.io
bsky: @chpoletto.bsky.social

analytical considerations: epidemic threshold

Comparison between stochastic simulations and Markov chain integration for a random power law network $p(k) \sim k^{-\gamma}$



limit of the approximation:

the approximation may be not so good for network with certain structural properties:

- networks with high clustering
- presence of communities
- assortative/dis-assortative

[S. Gómez, et al. EPL, 89 3 38009 (2010)]

analytical considerations: epidemic threshold

Mapping with statistical mechanics

- SIS belongs to the wide class of processes that present an absorbing state (contact processes)
- the unique stationary solution is the one where no individual is infected
- active epidemic state is a quasi-stationary state
- survival time of an SIS, $T_s(\beta, N)$:
    - $T_s(\beta, N) \sim \log N$ for $\beta \to 0$
    - $T_s(\beta, N)$ is super-polynomial for high β

[Ferreira et al PRE (2012)]

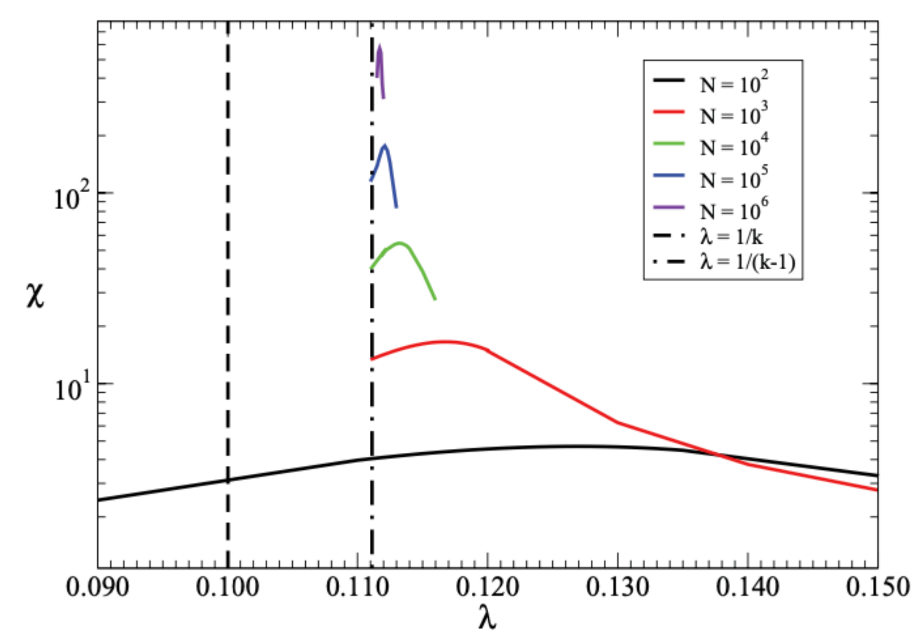
analytical considerations: epidemic threshold

Mapping with statistical mechanics

- SIS belongs to the wide class of processes that present an absorbing state (contact processes)
- the unique stationary solution is the one where no individual is infected
- active epidemic state is a quasi-stationary state
- survival time of an SIS, $T_s(\beta, N)$:
  - $T_s(\beta, N) \sim \log N$ for $\beta \to 0$
  - $T_s(\beta, N)$ is super-polynomial for high β
- around β_c we have a continuous phase transition
- critical behaviour - e.g. large fluctuations, critical exponents $\rho(\beta) \sim (\beta - \beta_c)^{b_c}$
- quasi stationary approach to numerically characterise the phase transition,

  the susceptibility, $\chi = N\dfrac{\langle\rho^2\rangle - \langle\rho\rangle^2}{\langle\rho\rangle}$, peaks at

  the critical point



[Ferreira et al PRE (2012)]

mean field approaches

In physics and probability theory, mean-field theory studies the behavior of high-dimensional stochastic models by studying a simpler model that approximates the original by averaging over degrees of freedom.

The effect of all the other individuals on any given individual is approximated by a single averaged effect, thus reducing a many-body problem to a one-body problem.

I assume that i behaves as an average node, i.e.

- Prob$[\sigma_i(t) = 0, \sigma_j(t) = 1]$ = Prob$[\sigma_i(t) = 0]$ Prob$[\sigma_j(t) = 1]$
- Prob$[\sigma_i(t) = 1] \equiv$ Prob$[\sigma(t) = 1] = \rho(t)$

mean field approaches

In physics and probability theory, mean-field theory studies the behavior of high-dimensional stochastic models by studying a simpler model that approximates the original by averaging over degrees of freedom.

The effect of all the other individuals on any given individual is approximated by a single averaged effect, thus reducing a many-body problem to a one-body problem.

thus

$$\frac{d}{dt}\text{Prob}[\sigma_i(t) = 1] = -\mu\rho(i,t) + \beta\sum_j A_{ij}\text{Prob}[\sigma_i(t) = 0, \sigma_j(t) = 1]$$

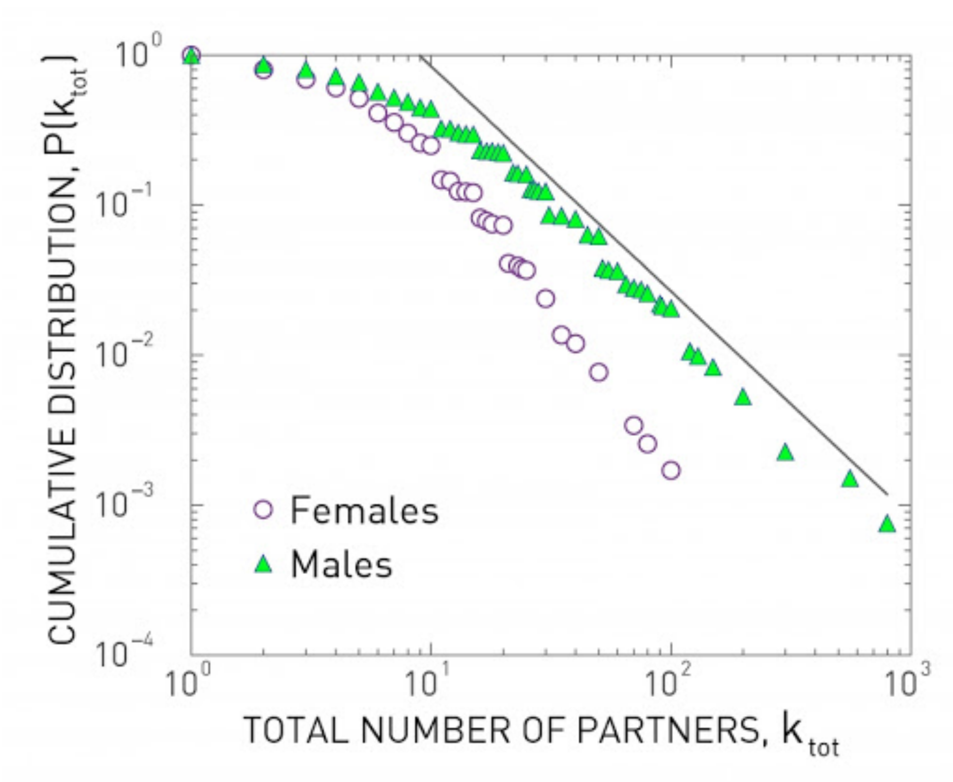
becomes

$$\frac{d}{dt}\text{Prob}[\sigma(t) = 1] = -\mu\rho(t) + \beta\,\text{Prob}[\sigma(t) = 0]\text{Prob}[\sigma(t) = 1]\sum_j A_{ij}$$

population dynamics SIS!

$$\frac{d}{dt}\rho(t) = -\mu\rho(t) + \beta\left(1-\rho(t)\right)\rho(t)k :$$ **homogenous mean field approach**

from homogenous to heterogeneous mean field approach

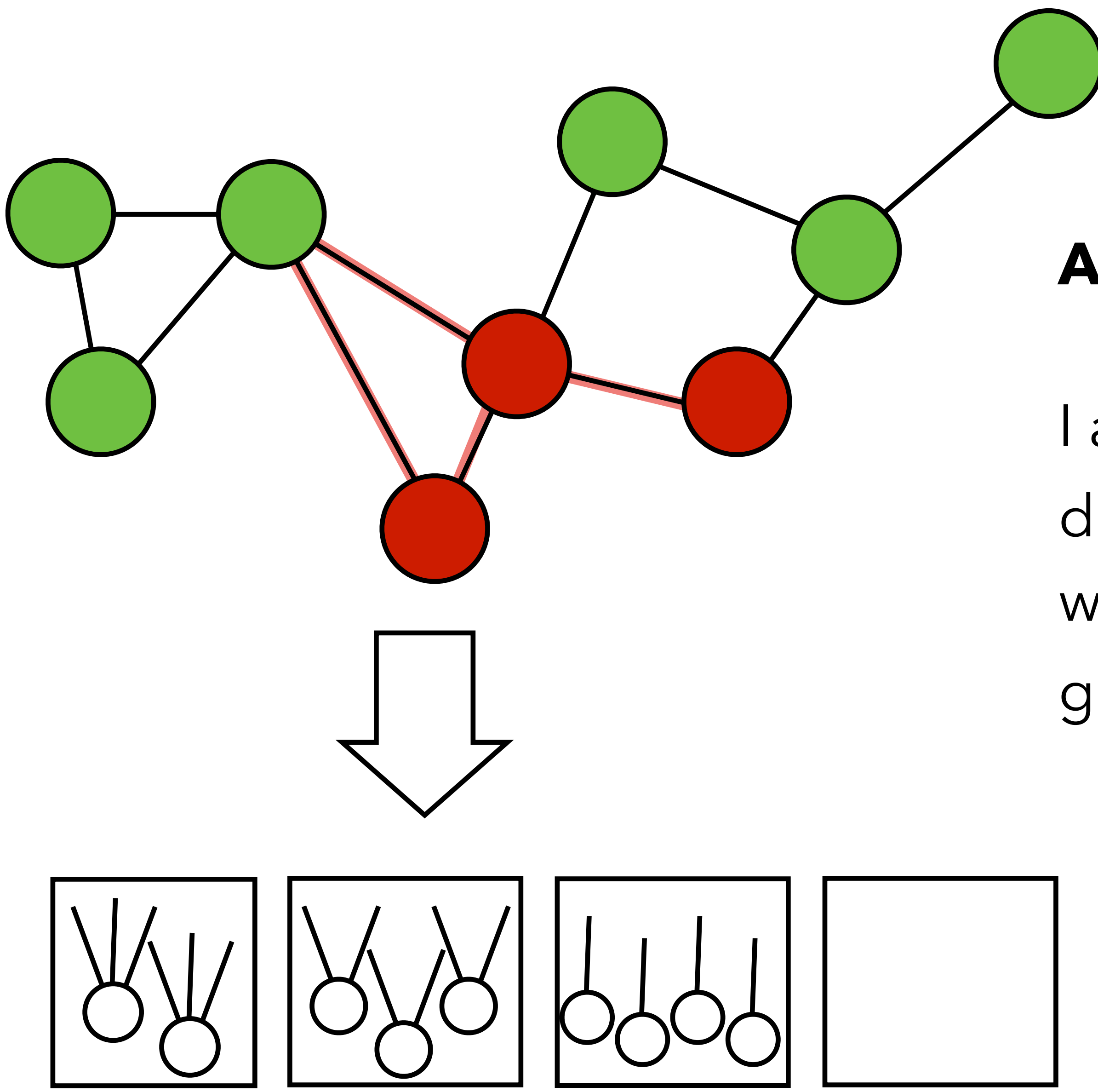


[B. Lewin. (ed.), Sex i Sverige. Om sexuallivet i Sverige 1996 [Sex in Sweden. On the Sexual Life in Sweden 1996]. National Institute of Public Health, Stockholm, 1998.]

marked heterogeneities in the degree : accounting for them with the mean field framework

from i behaving as the average node **to** i behaving as the average node within its degree class k_i

heterogenous mean field approach



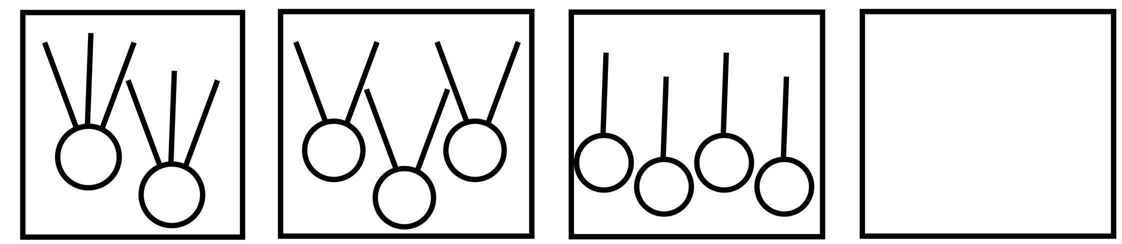
$k = 3$ $k = 2$ $k = 1$ $k = \ldots$

Also degree-based mean field approach

I assume I know only the network degree distribution $P(k)$ → I approximate the network with a random uncorrelated network with that given $P(k)$ **(configurational model)**

[Pastor-Satorras & Vespignani PRL 2001, review :Pastor Satorras et al. Rev Mod Phys 2015]

heterogenous mean field approach



$k = 3 \quad k = 2 \quad k = 1 \quad k = ...$

Degree based compartments

$s_k = \dfrac{S_k}{N_k}$, fraction of susceptible nodes of degree k in the network

$\rho_k = \dfrac{I_k}{N_k}$, fraction of infected nodes of degree k in the network

N_k number of nodes with degree k in the network

Total fraction of ρ and s:

$$\rho = \sum_k P(k)\rho_k$$

$$s = \sum_k P(k)s_k$$

heterogenous mean field approach

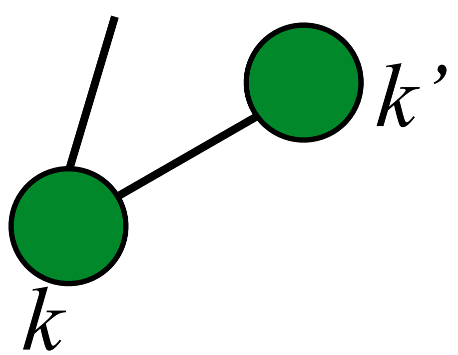
$$\frac{d}{dt}\rho(\mathbf{i},t) = -\mu\rho(\mathbf{i},t) + \beta\sum_j A_{ij}\mathrm{Prob}[\sigma_{\mathbf{i}}(t) = 0, \sigma_{\mathbf{j}}(t) = 1]$$

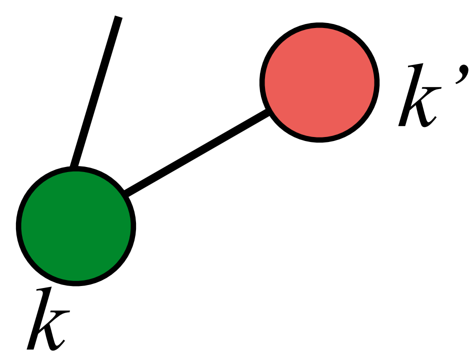
$\Theta_k(t)$ = Density of infectious within the neighbour of a node with degree k

One equation for each degree class: $\frac{d}{dt}\rho_k(t) = -\mu\rho_k(t) + \beta k(1-\rho_k(t))\Theta_k(t)$

heterogenous mean field approach

$\Theta_k(t)$ = Density of infectious within the neighbour of a node with degree k





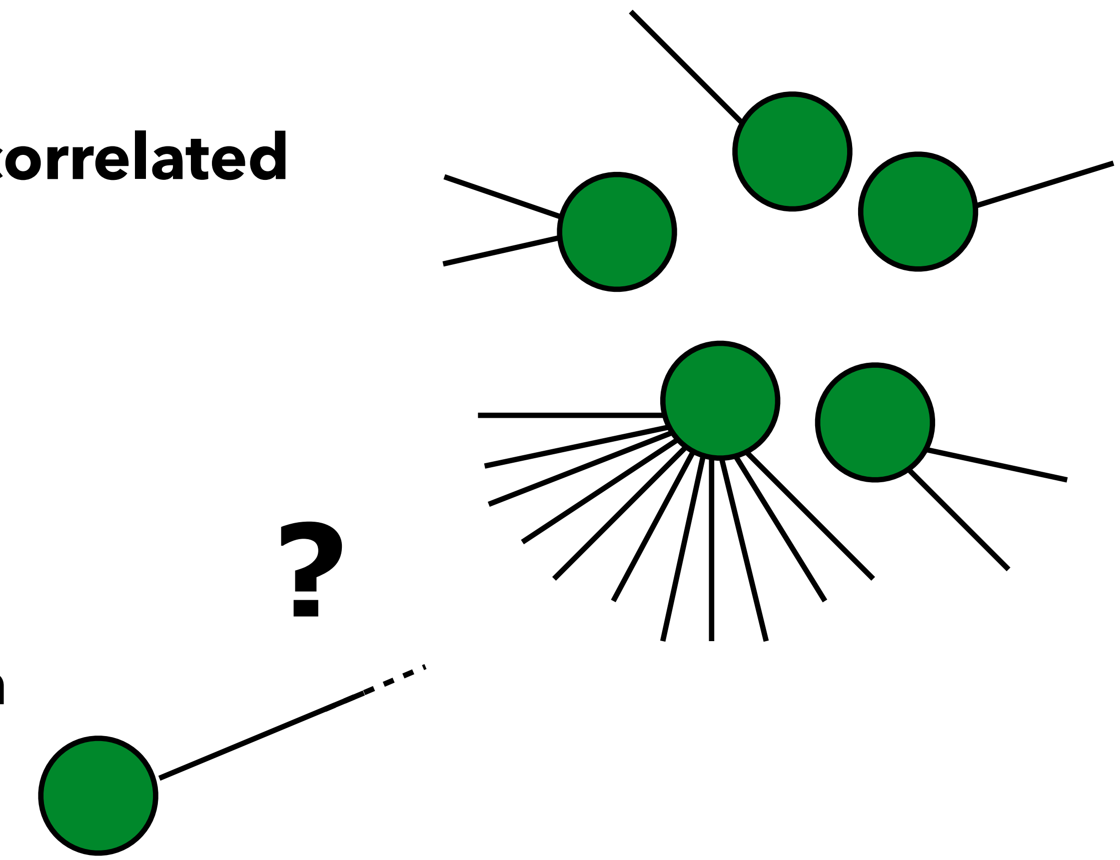
probability of contact with k': $P(k'|k)$ **x** number of infectious within the k'-class: $\rho_{k'}$

$$\Theta_k(t) = \sum_{k'} P(k'|k)\rho_{k'}$$

heterogenous mean field approach

hypothesis: the network is uncorrelated

$$P(k'|k) = \frac{k'P(k')}{\sum_{k'} k'P(k')} = \frac{k'P(k')}{\langle k \rangle}$$



If I make a connection at random

I will do it more likely with

someone that is very social (more stubs)

$\Theta(t)$ does not depend on k (as expected due to the absence of correlation)

heterogenous mean field approach

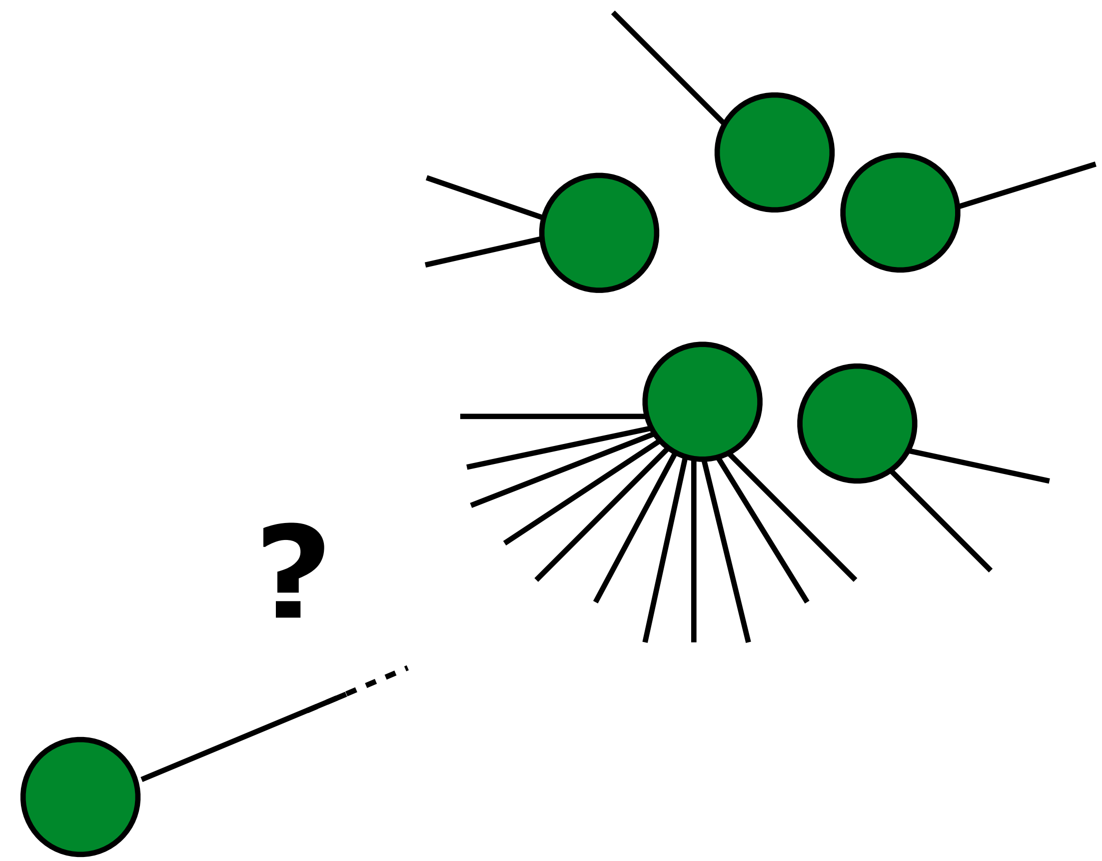
average nearest neighbour degree:

$$k_{\mathrm{nn},i} = \frac{1}{k_i} \sum_{j \in \upsilon(i)} k_j$$

$$P(k'|k) = \frac{k'P(k')}{\langle k \rangle}$$

$$k_{\mathrm{nn},i} = \sum_{k'} k' P(k'|k_i) = \frac{1}{\langle k \rangle} \sum_{k'} k'^2 P(k')$$

$$k_{\mathrm{nn},i} = \frac{\langle k^2 \rangle}{\langle k \rangle}$$



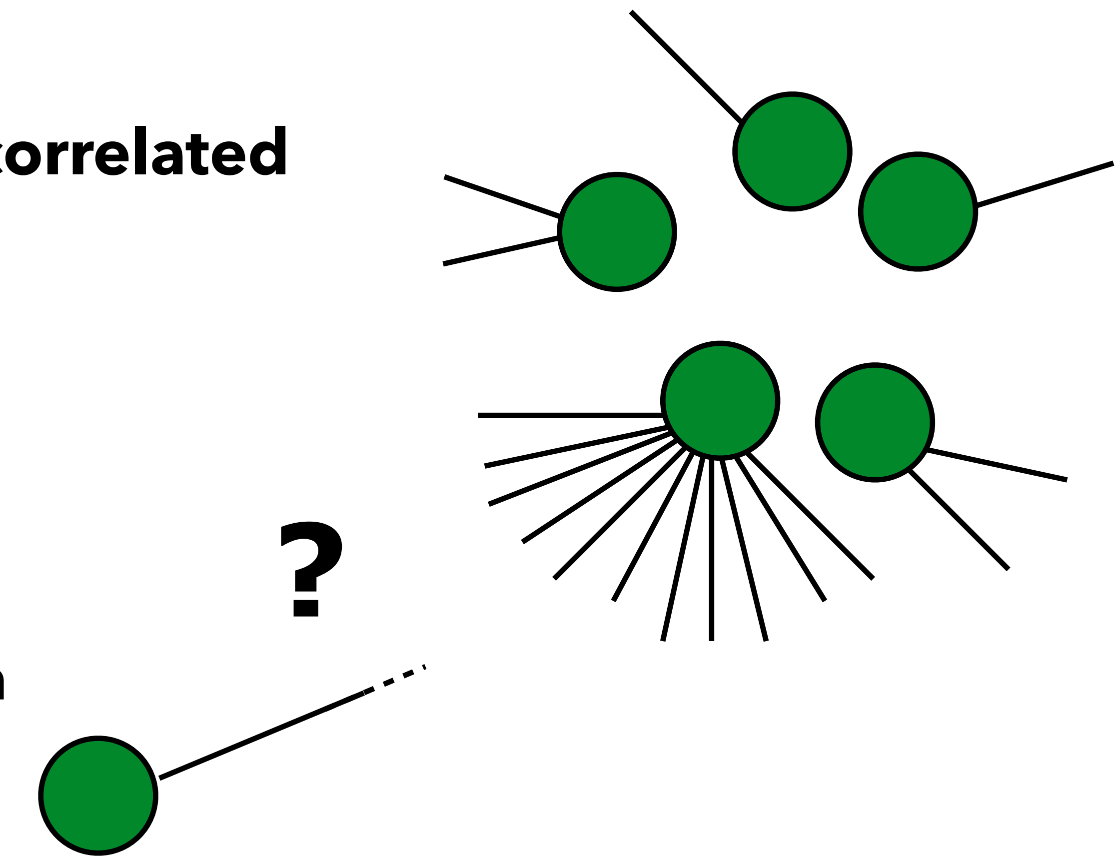
my friend has more friends than me ...

heterogenous mean field approach

hypothesis: the network is uncorrelated

$$P(k'|k) = \frac{k'P(k')}{\sum_{k'} k'P(k')} = \frac{k'P(k')}{\langle k \rangle}$$

If I make a connection at random
I will do it more likely with
someone that is very social (more stubs)



$$\Theta_k(t) = \sum_{k'} P(k'|k)\rho_{k'} = \frac{\sum_{k'} k'P(k')\rho_{k'}(t)}{\langle k \rangle} \equiv \Theta(t)$$

$\Theta(t)$ **does not depend on** k **(as expected due to the absence of correlation)**

heterogenous mean field approach

$$\frac{d}{dt}\rho(\mathbf{i},t) = -\mu\rho(\mathbf{i},t) + \beta\sum_j A_{ij}\underline{\text{Prob}[\sigma_{\mathbf{i}}(t) = 0, \sigma_{\mathbf{j}}(t) = 1]}$$

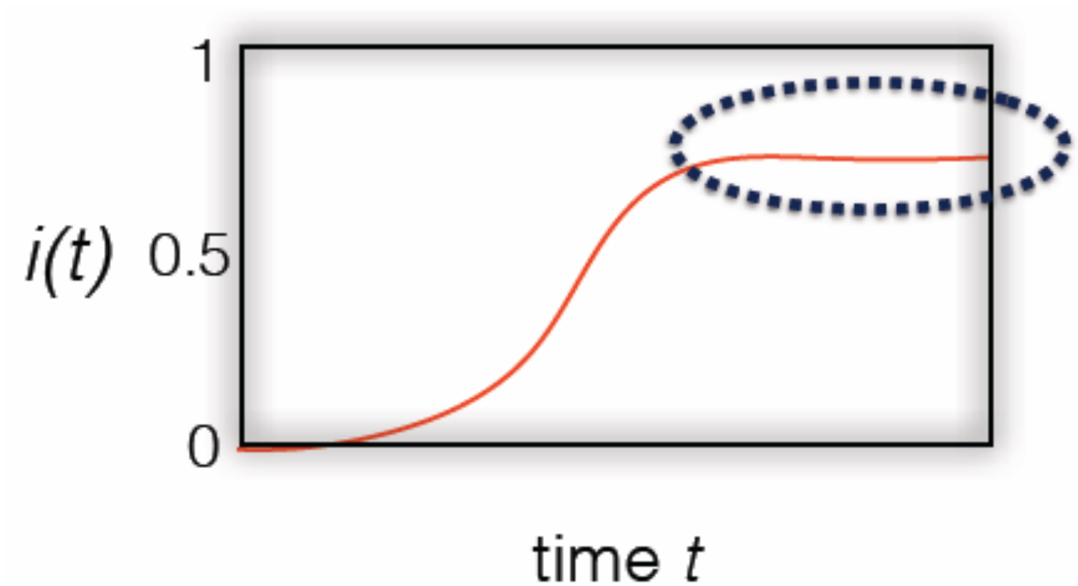
$\Theta_k(t)$ = Density of infectious within the neighbour of a node with degree k

One equation for each degree class: $\frac{d}{dt}\rho_k(t) = -\mu\rho_k(t) + \underline{\beta k(1-\rho_k(t))\Theta_k(t)}$

$$\Theta_k(t) \equiv \Theta(t) = \frac{\sum_{k'} k' P(k')\rho_{k'}(t)}{\langle k \rangle}$$

$$\frac{d}{dt}\rho_k(t) = -\mu\rho_k(t) + \beta k(1-\rho_k(t))\Theta(t)$$

$$\frac{d}{dt}\rho_k(t) = 0 \;\rightarrow\; \rho_k = \frac{\beta k\Theta}{\mu + \beta k\Theta}$$



heterogenous mean field approach

$$\rho_k = \frac{\beta k \Theta}{\mu + \beta k \Theta}, \text{ with } \Theta(t) \equiv \frac{\sum_{k'} k' P(k') \rho_{k'}(t)}{\langle k \rangle}$$

$\frac{\sum_k k P(k)}{\langle k \rangle}$ $\frac{\sum_k k P(k)}{\langle k \rangle}$

$$\Theta = \frac{1}{\langle k \rangle} \sum_k \frac{k^2 P(k) \beta \Theta}{\mu + \beta k \Theta}$$: self consistent equation for Θ

- Trivial solution $\Theta = 0$: disease-free equilibrium
- Non-trivial solution ?

heterogenous mean field approach

$$\Theta = \boxed{\frac{1}{\langle k \rangle} \sum_k \frac{k^2 P(k) \beta \Theta}{\mu + \beta k \Theta}}^{\,f(\Theta)}$$: self consistent equation for Θ

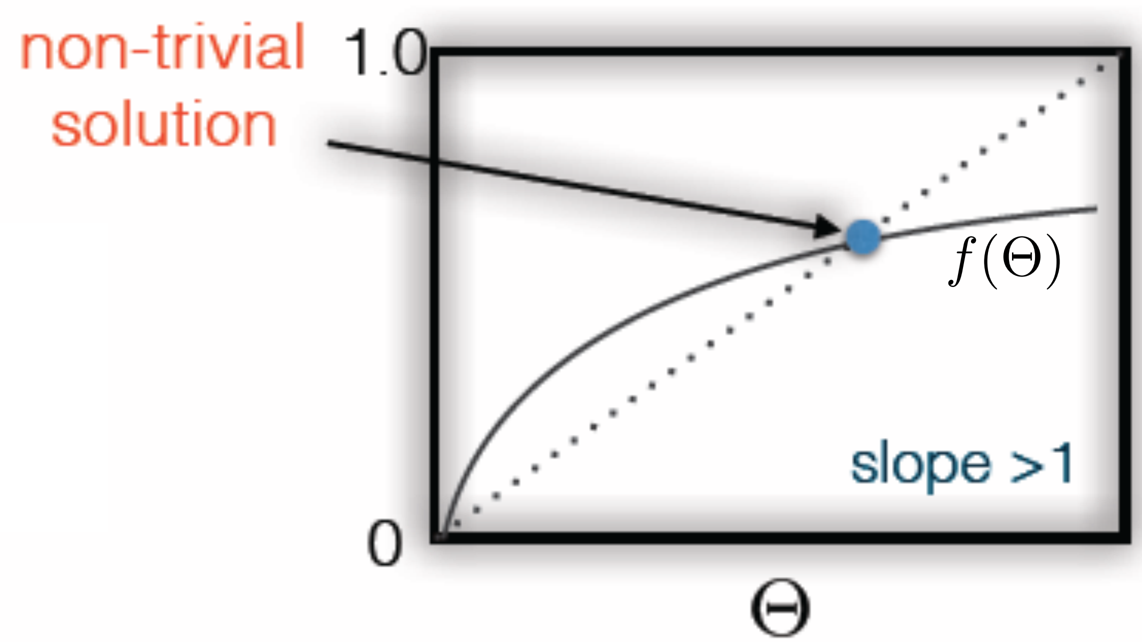
Non-trivial solution?

A solution exists when Θ and $f(\Theta)$ cross in the interval

$0 < \Theta \leq 1$

i.e. $\left.\frac{d}{d\Theta} f(\Theta)\right|_{\Theta=0} \geq 1$

$$\left.\frac{d}{d\Theta}\left(\frac{1}{\langle k \rangle} \sum_k \frac{k^2 P(k) \beta \Theta}{\mu + \beta k \Theta}\right)\right|_{\Theta=0} \geq 1$$



heterogenous mean field approach

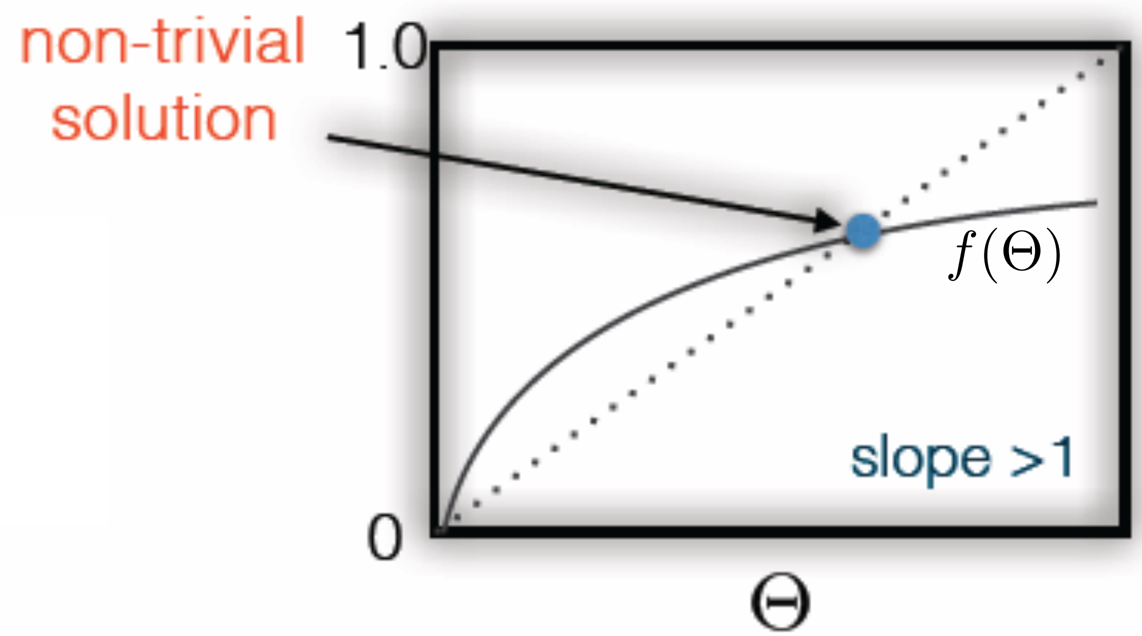
$$\Theta = \underbrace{\boxed{\frac{1}{\langle k \rangle} \sum_k \frac{k^2 P(k) \beta \Theta}{\mu + \beta k \Theta}}}_{}^{f(\Theta)}$$: self consistent equation for Θ

Non-trivial solution?

$$\frac{d}{d\Theta} \left(\frac{1}{\langle k \rangle} \sum_k \frac{k^2 P(k) \beta \Theta}{\mu + \beta k \Theta} \right) \Bigg|_{\Theta=0} \geq 1$$

$$\frac{\beta}{\mu \langle k \rangle} \sum_k k^2 P(k) \geq 1$$

$$\frac{\beta \langle k^2 \rangle}{\mu \langle k \rangle} \geq 1$$



heterogenous mean field approach

$$\Theta = \boxed{\frac{1}{\langle k \rangle} \sum_k \frac{k^2 P(k) \beta \Theta}{\mu + \beta k \Theta}}^{\,f(\Theta)}$$: self consistent equation for Θ

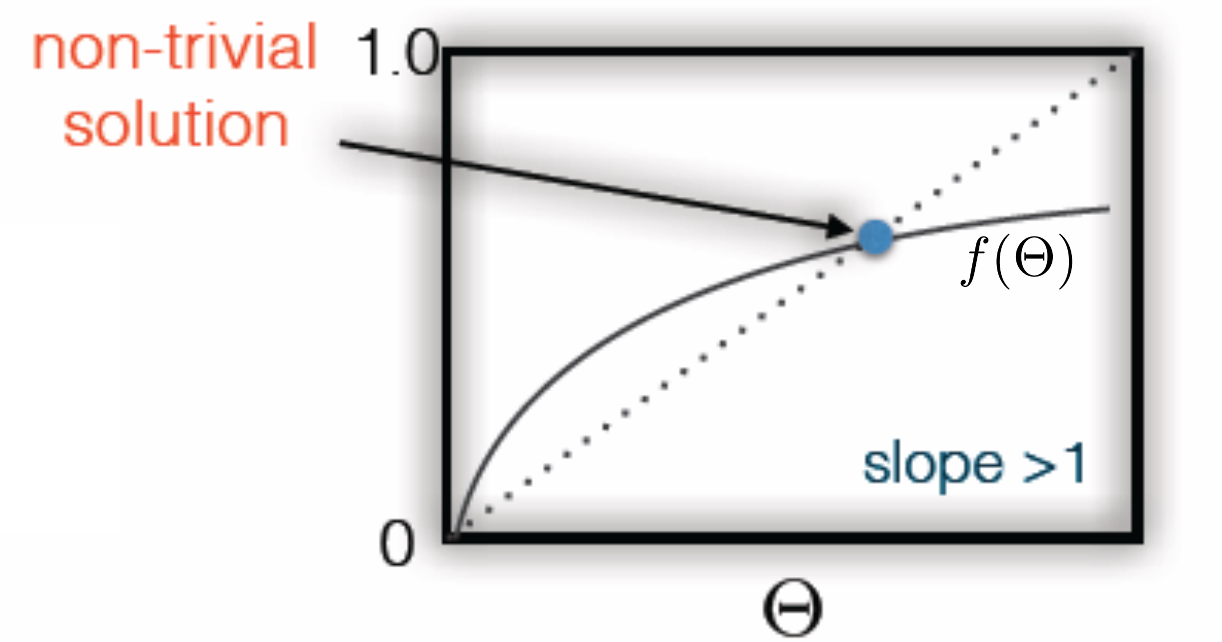
Non-trivial solution ?

$$\frac{d}{d\Theta}\left(\frac{1}{\langle k \rangle} \sum_k \frac{k^2 P(k) \beta \Theta}{\mu + \beta k \Theta} \right)\Bigg|_{\Theta=0} \geq 1$$

$$\frac{\beta}{\mu \langle k \rangle} \sum_k k^2 P(k) \geq 1$$

$$\boxed{\frac{\beta \langle k^2 \rangle}{\mu \langle k \rangle} \geq 1}$$

Condition for an endemic state



heterogenous mean field approach

$\frac{\beta \langle k^2 \rangle}{\mu \langle k \rangle} \geq 1$ → **epidemic threshold** $\beta_c = \frac{\mu \langle k \rangle}{\langle k^2 \rangle}$

Implications:

- For homogenous networks $\langle k^2 \rangle \simeq \langle k \rangle^2$ recovering:

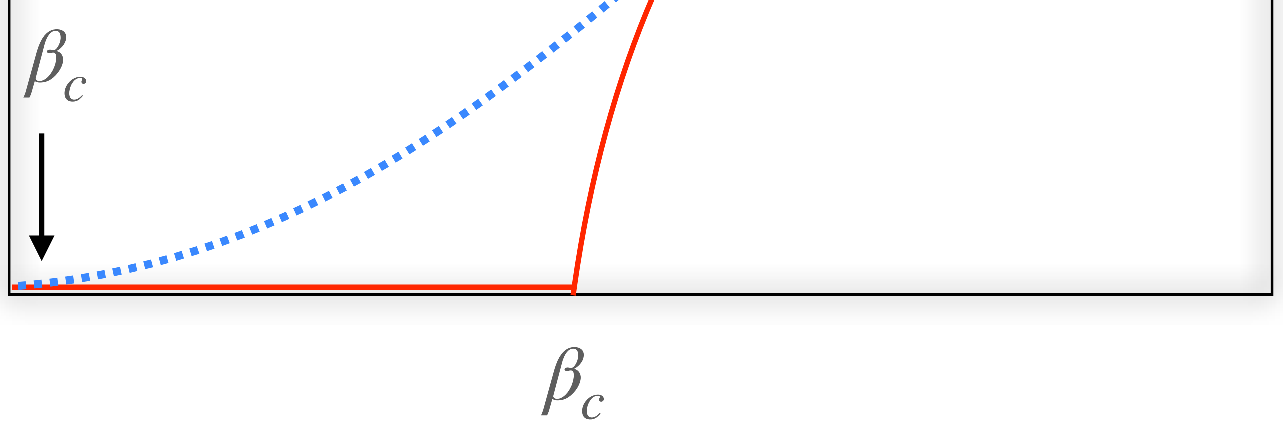
  $$\beta_c = \frac{\mu \langle k \rangle}{\langle k^2 \rangle} = \frac{\mu}{\langle k \rangle}$$

- Recalling that in Scale-Free networks with $2 < \gamma \leq 3$ we have $\langle k \rangle \to c$ and $\langle k^2 \rangle \to \infty$ as $N \to \infty$

  $\beta_c = \frac{\mu \langle k \rangle}{\langle k^2 \rangle} \to 0$ → **The epidemic threshold vanishes for $N \to \infty$**

heterogenous mean field approach

$$\frac{\beta\langle k^2\rangle}{\mu\langle k\rangle} \geq 1 \quad \rightarrow$$ **epidemic threshold** $\beta_c = \frac{\mu\langle k\rangle}{\langle k^2\rangle}$



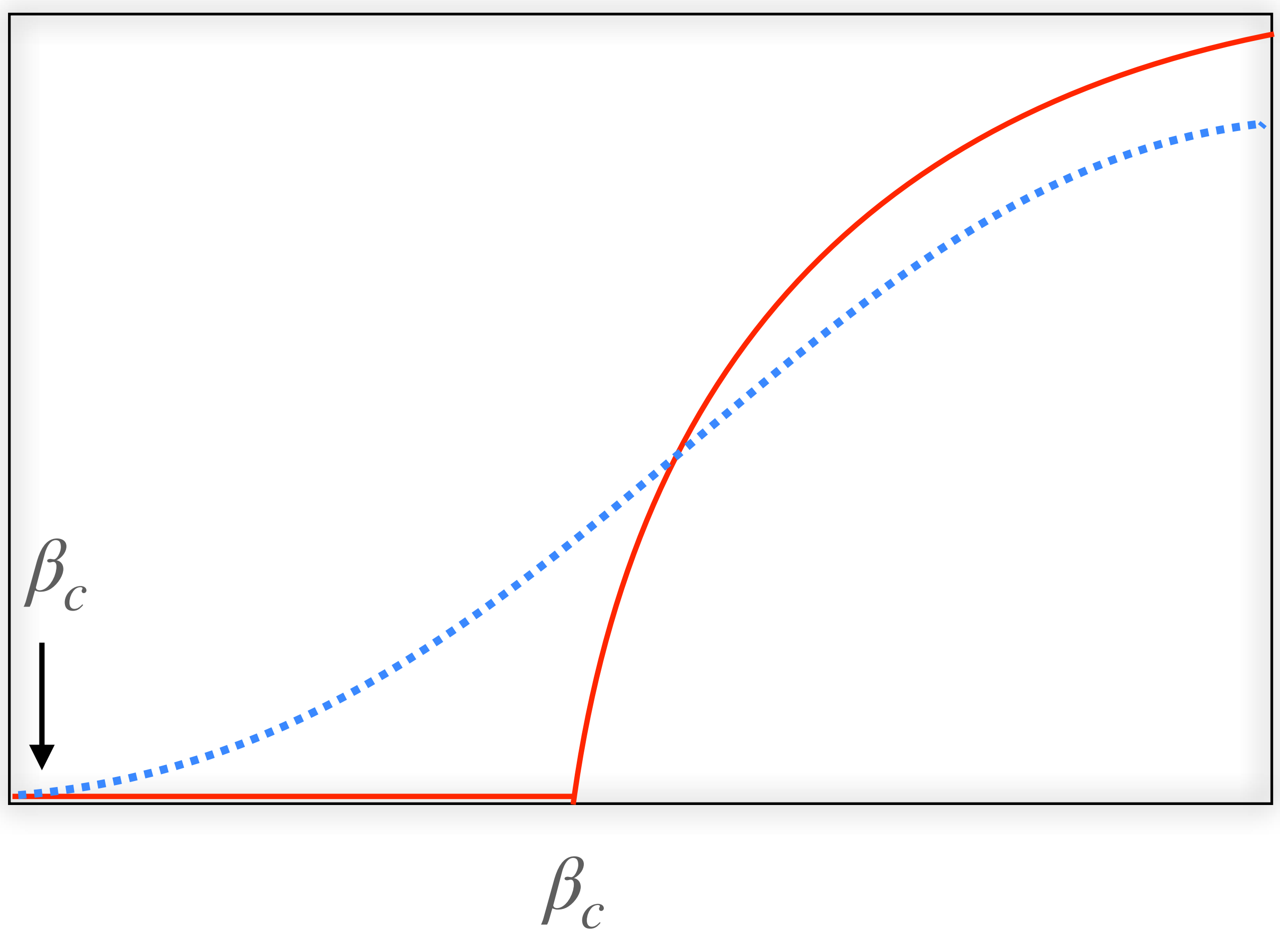
immunisation in heterogeneous networks

imagining of immunising a fraction g of individuals chosen at random

$\beta(1-g) \leq \beta_c$

then $g_c \simeq 1$

random immunisation is totally ineffective



[Pastor-Satorras & Vespignani, PRE 65, 036104 (2002)]
[Dezso & Barabasi cond-mat/0107420; Havlin et al. preprint (2002)]

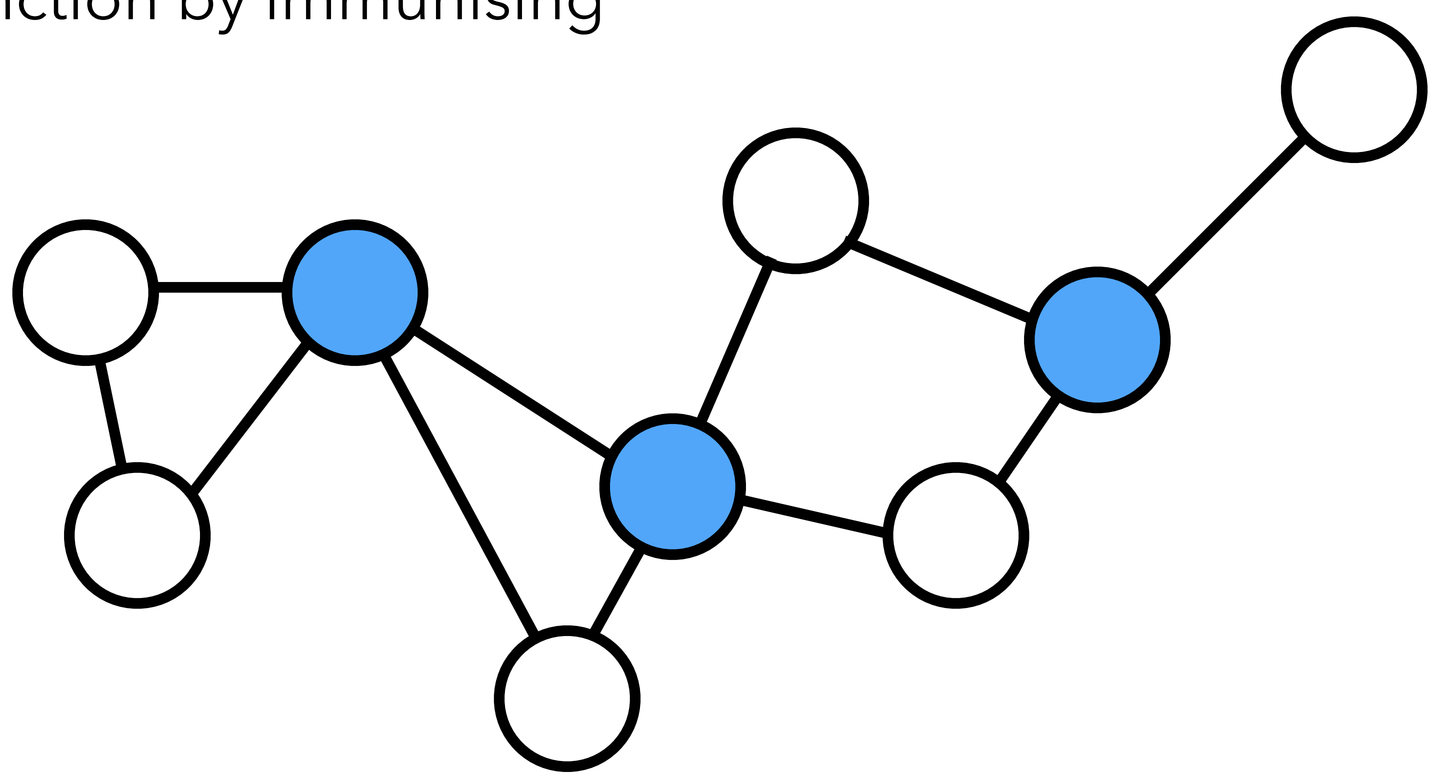
immunisation in heterogeneous networks

hubs act as super-spreaders

⇒ targeted immunisation is extremely effective

I can drive the epidemic to extinction by immunising only the hubs

(which are the hubs?)



[Pastor-Satorras & Vespignani, PRE 65, 036104 (2002)]
[Dezso & Barabasi cond-mat/0107420; Havlin et al. preprint (2002)]

immunisation in heterogeneous networks

average nearest neighbour degree:

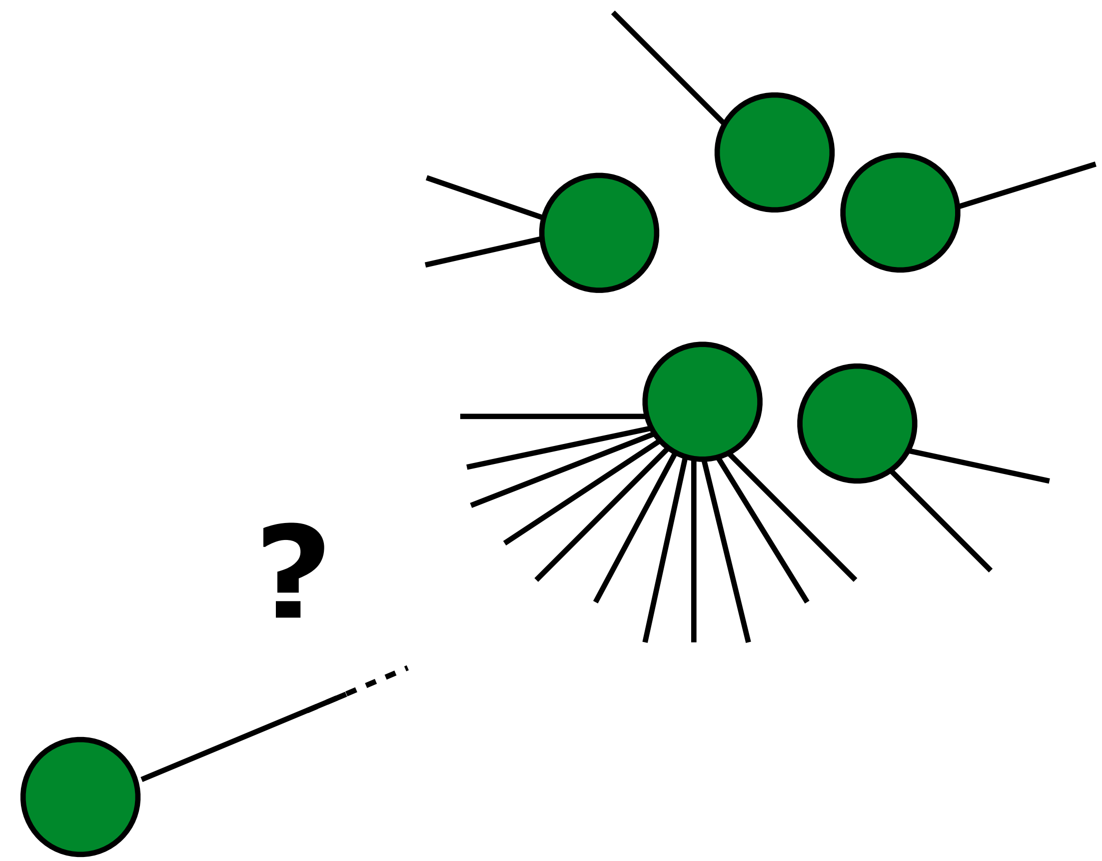
$$k_{\mathrm{nn},i} = \frac{1}{k_i} \sum_{j \in v(i)} k_j$$

$$P(k'|k) = \frac{k'P(k')}{\langle k \rangle}$$

$$k_{\mathrm{nn},i} = \sum_{k'} k' P(k'|k_i) = \frac{1}{\langle k \rangle} \sum_{k'} k'^2 P(k')$$

$$\boxed{k_{\mathrm{nn},i} = \frac{\langle k^2 \rangle}{\langle k \rangle}}$$

my friend has more friends than me ...



[Cohen et la PRL 2003; Christakis & Fowler PLoS ONE 2010]

immunisation in heterogeneous networks

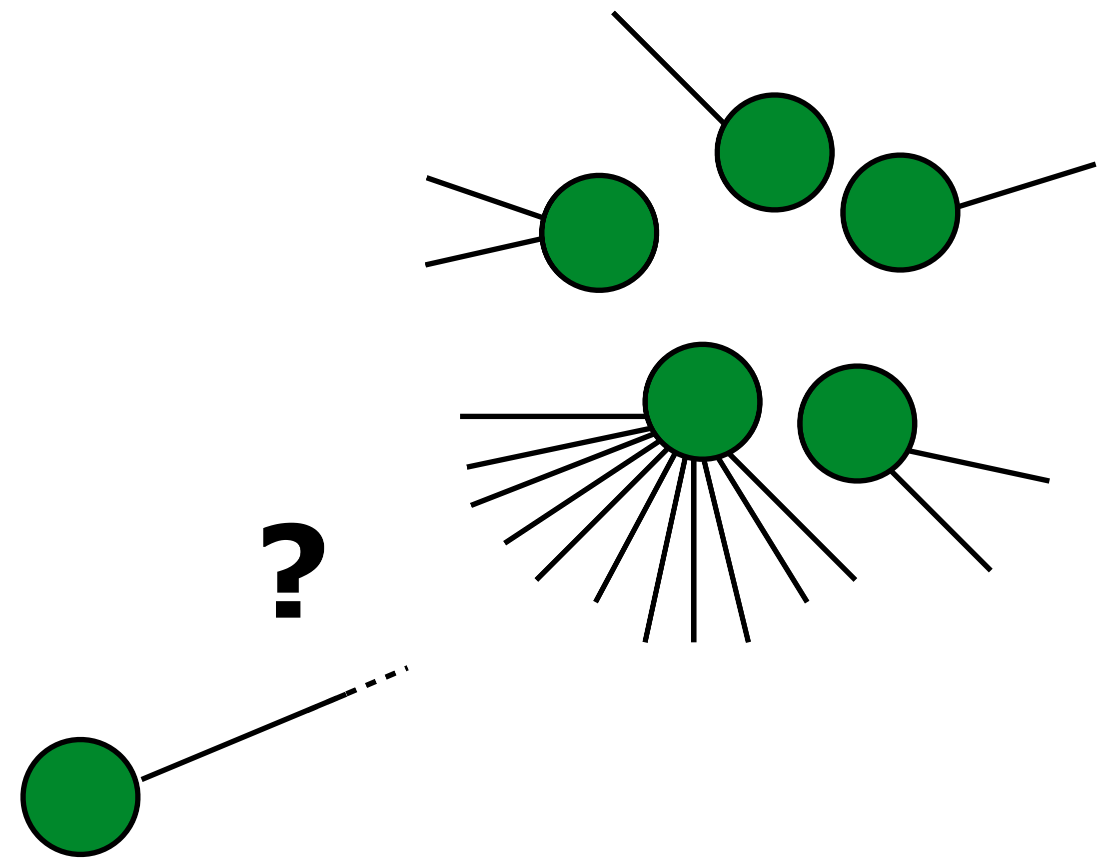
average nearest neighbour degree:

$$k_{\mathrm{nn},i} = \frac{1}{k_i} \sum_{j \in \upsilon(i)} k_j$$

$$P(k'|k) = \frac{k' P(k')}{\langle k \rangle}$$

$$k_{\mathrm{nn},i} = \sum_{k'} k' P(k'|k_i) = \frac{1}{\langle k \rangle} \sum_{k'} k'^2 P(k')$$

$$\boxed{k_{\mathrm{nn},i} = \frac{\langle k^2 \rangle}{\langle k \rangle}}$$



my friend has more friends than me …

Acquaintance immunisation : I select a random node and I vaccinate one of its random friends

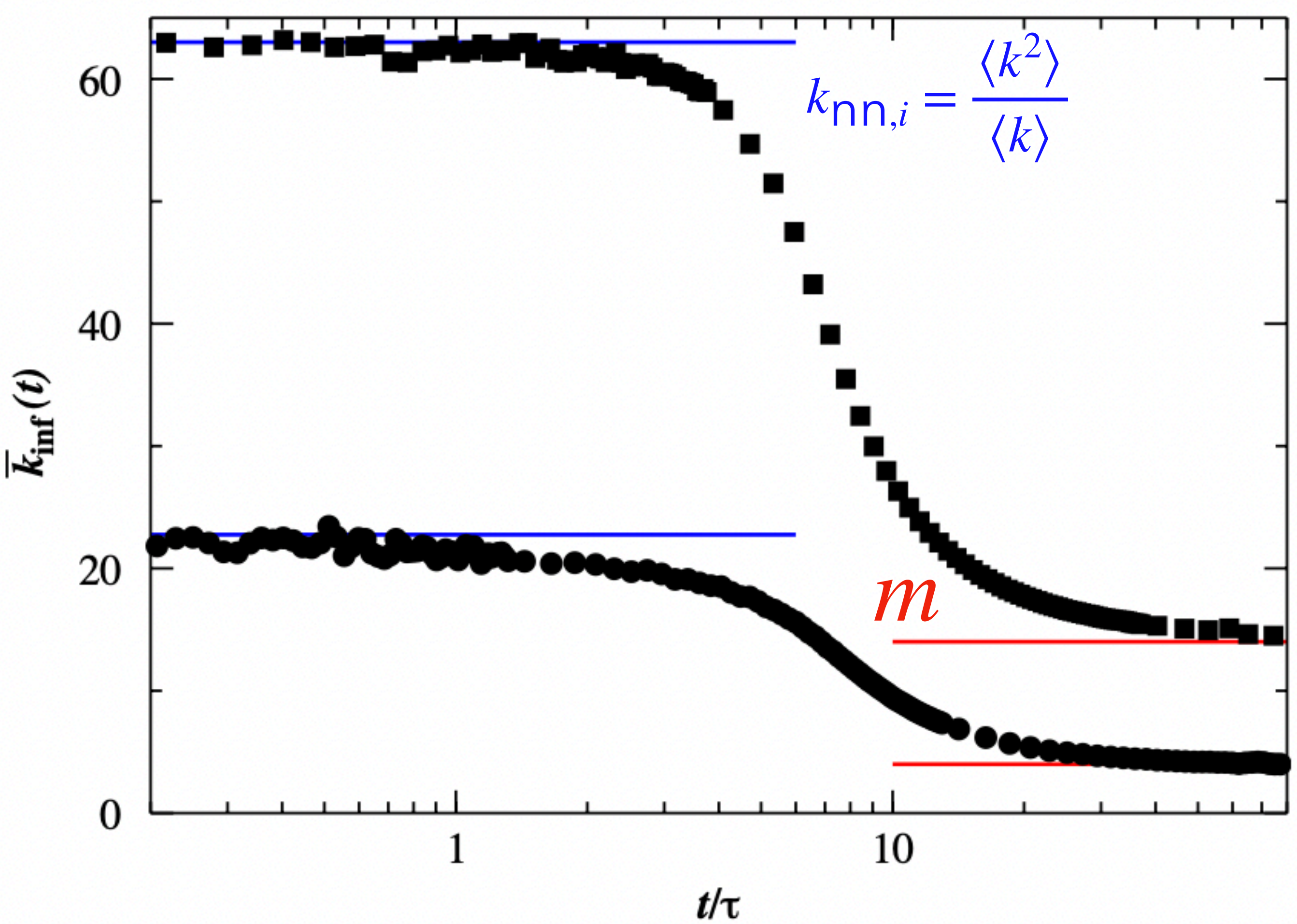
[Cohen et la PRL 2003; Christakis & Fowler PLoS ONE 2010]

more on the role of hubs

cascade phenomenon

average degree of newly infected nodes

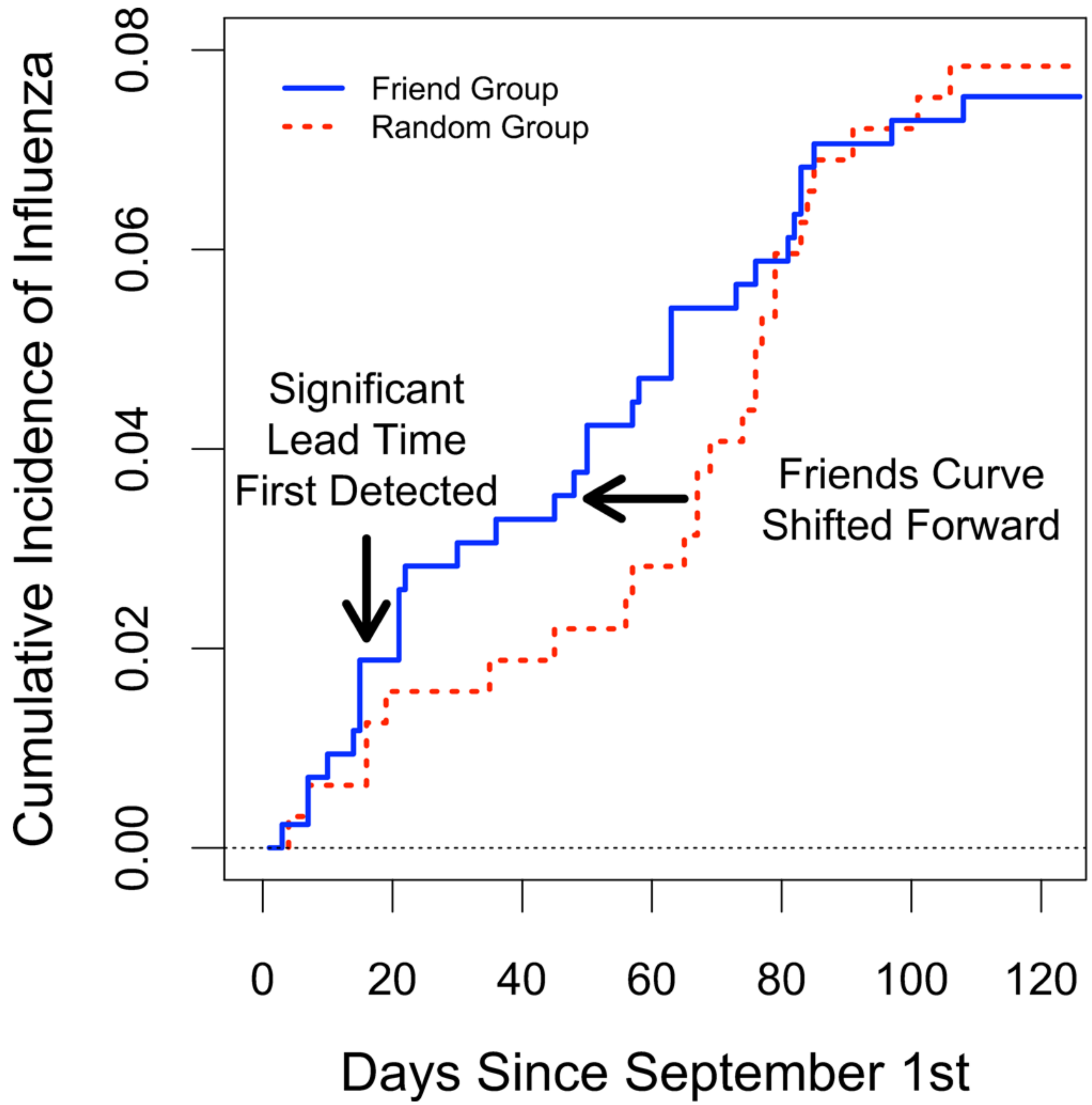
$$\bar{k}_{\text{inf}}(t) = \frac{\sum_k k\left[I_k(t) - I_k(t-1)\right]}{I(t) - I(t-1)}$$



Stochastic simulations on a Barabási-Albert network (smallest degree m)

[Barthélemy et alJTB 2005]

more on the role of hubs



[Cohen et la PRL 2003; Christakis & Fowler PLoS ONE 2010]

hubs as sentinels:

744 undergraduate students from Harvard College:
random group: 319 individual chosen at random
friends group: 425 individuals who were named as a friend at least once by a member of this random sample

tracked whether they had the flu beginning on September 1, 2009 to December 31, 2009.

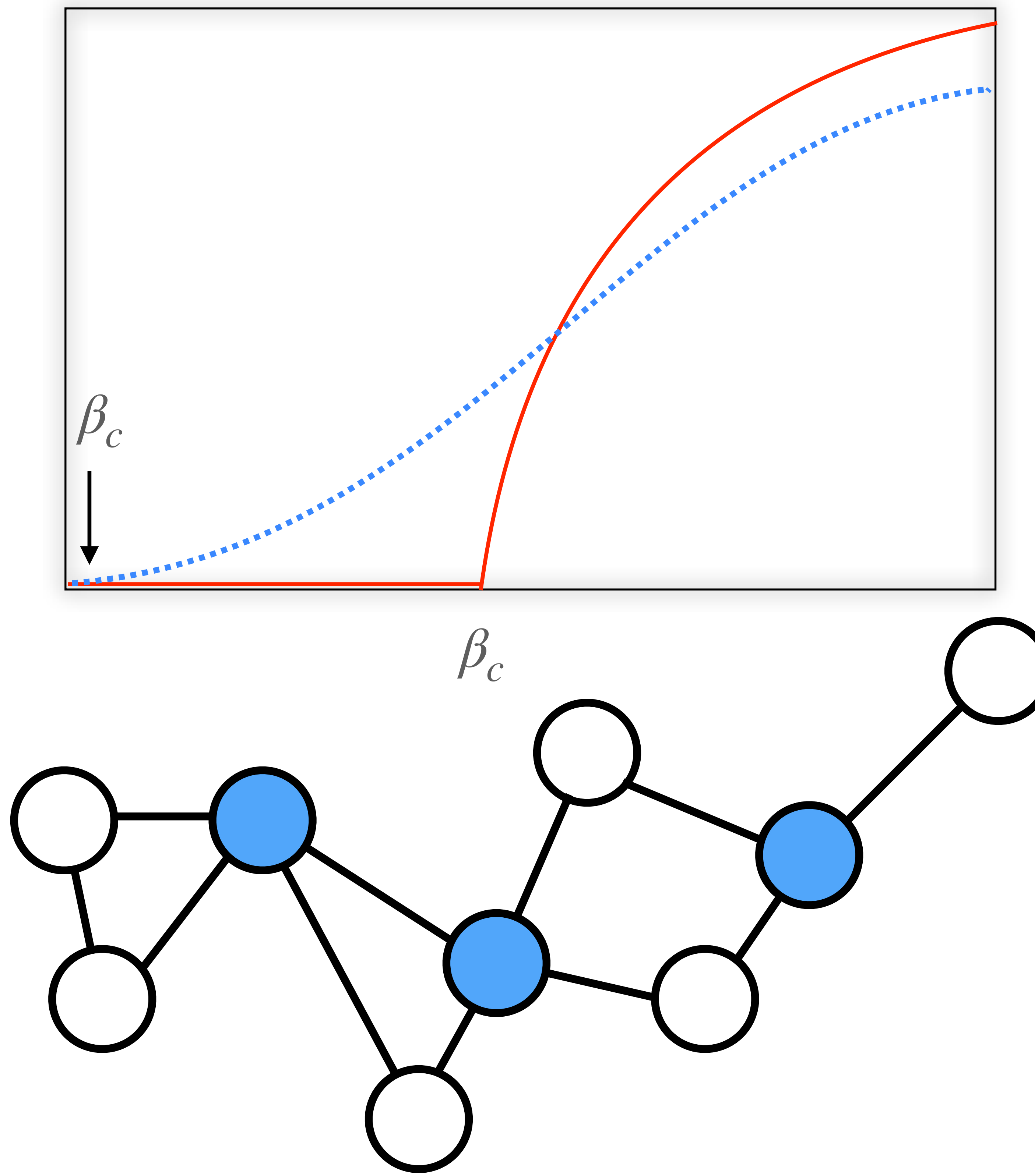
epidemic in the friend group 13.9 days (95% C.I. 9.9–16.6) in advance of the random group

more on the role of hubs

hubs are :

- super-spreader
- sentinels
- super-blokers

(consequences for the heard immunity threshold and for the emergence of new strains)



individual-based vs. heterogeneous

Individual-based mean-field approach

heterogeneous mean-field approach

[Castellano, Pastor-Satorras PRL 2010]

exact structure of the network

-fully account for the network information. Reasonably good for any type of network

-test possible interventions

[Wang et al. SRDS 2003; Gómez et al. 2010]

coarse graining - statistical properties

-simple transparent formula

-deep analytical understanding of the impact of contact heterogeneities

-full network information not available in many cases

[Pastor-Satorras & Vespignani PRL 2001, Pastor Satorras et al. Rev Mod Phys 2015]

HMF vs. IBMF

- In the HMF:

$$\beta_c^{DBMF} = \frac{\mu\langle k\rangle}{\langle k^2\rangle}$$

- In the IBMF:

$$\beta_c^{IBMF} = \frac{\mu}{\Lambda_{max}}$$

- What is the relationship between them?

$$\beta_c^{IBMF} \neq \beta_c^{DBMF}$$

For power law Networks $P(k) \sim k^{-\gamma}$ we have that:

$$\Lambda_{max} \simeq \max\left(\sqrt{k_{max}}, \langle k^2\rangle/\langle k\rangle\right)$$

Specifically:

$$\beta_c \simeq \begin{cases} \mu/\sqrt{k_{max}} & \gamma > 5/2 \\ \mu\langle k\rangle/\langle k^2\rangle & 2 < \gamma < 5/2 \end{cases}$$

[Castellano, Pastor-Satorras PRL 2010]

HMF vs. IBMF

Deriving the HMF from IBMF (continuous time):

$$\dot{\rho}_i = -\mu\rho_i + (1-\rho_i)q_i\,,\;\; q_i = 1 - \prod_{j=1}^{N}[1-\beta A_{ij}\rho_j]$$

Average over the "ensemble" of networks given by $P(k)$ vs. a single realization of the network

The Adjacency Matrix is replaced by an Annealed Adjacency Matrix AAM:

$$\bar{A}_{ij} = \frac{k_j P(k_i\,|\,k_j)}{NP(k_i)}$$

For random networks becomes:

$$\bar{A}_{ij} = \frac{k_i k_j}{N\langle k\rangle}$$

HMF vs. IBMF

The expression for q_i reads:

$$q_i = 1 - \prod_{j=1}^{N} \left[1 - \beta \frac{k_i k_j}{N\langle k \rangle} \rho_j\right]$$

$$q_k = 1 - \prod_{k'} \left[1 - \beta \frac{kk'}{N\langle k \rangle} \rho_{k'}\right]^{N_{k'}}$$

Assuming $\beta\rho_k \ll 1$, remembering that $N_{k'} = NP(k')$

$$q_k \simeq \beta \sum_{k'} NP(k') \frac{kk'}{N\langle k \rangle} \rho_{k'} = \beta k \sum_{k'} \frac{k' P(k')}{\langle k \rangle} \rho_{k'}$$

From individual nodes to degree classes

$$\dot{\rho}_k = -\mu\rho_k + (1 - \rho_k) q_k$$

HMF vs. IBMF

The expression for q_i reads:

$$q_i = 1 - \prod_{j=1}^{N}\left[1 - \beta\frac{k_i k_j}{N\langle k\rangle}\rho_j\right]$$

$$q_k = 1 - \prod_{k'}\left[1 - \beta\frac{kk'}{N\langle k\rangle}\rho_{k'}\right]^{N_{k'}}$$

Assuming $\beta\rho_k \ll 1$, remembering that $N_{k'} = NP(k')$

$$q_k \simeq \beta\sum_{k'} NP(k')\frac{kk'}{N\langle k\rangle}\rho_{k'} = \beta k\sum_{k'}\frac{k'P(k')}{\langle k\rangle}\rho_{k'}$$

From individual nodes to degree classes

$$\dot{\rho}_k = -\mu\rho_k + (1-\rho_k)q_k$$

HMF

$$\frac{d}{dt}\rho_k(t) = -\mu\rho_k(t) + \beta k(1-\rho_k(t))\Theta(t)$$

$$\Theta_k(t) \equiv \Theta(t) = \frac{\sum_{k'} k'P(k')\rho_{k'}(t)}{\langle k\rangle}$$

In the DBMF we are implicitly assuming that $\beta\rho_k \ll 1$

Reason why DBMF is accurate only around the epidemic threshold