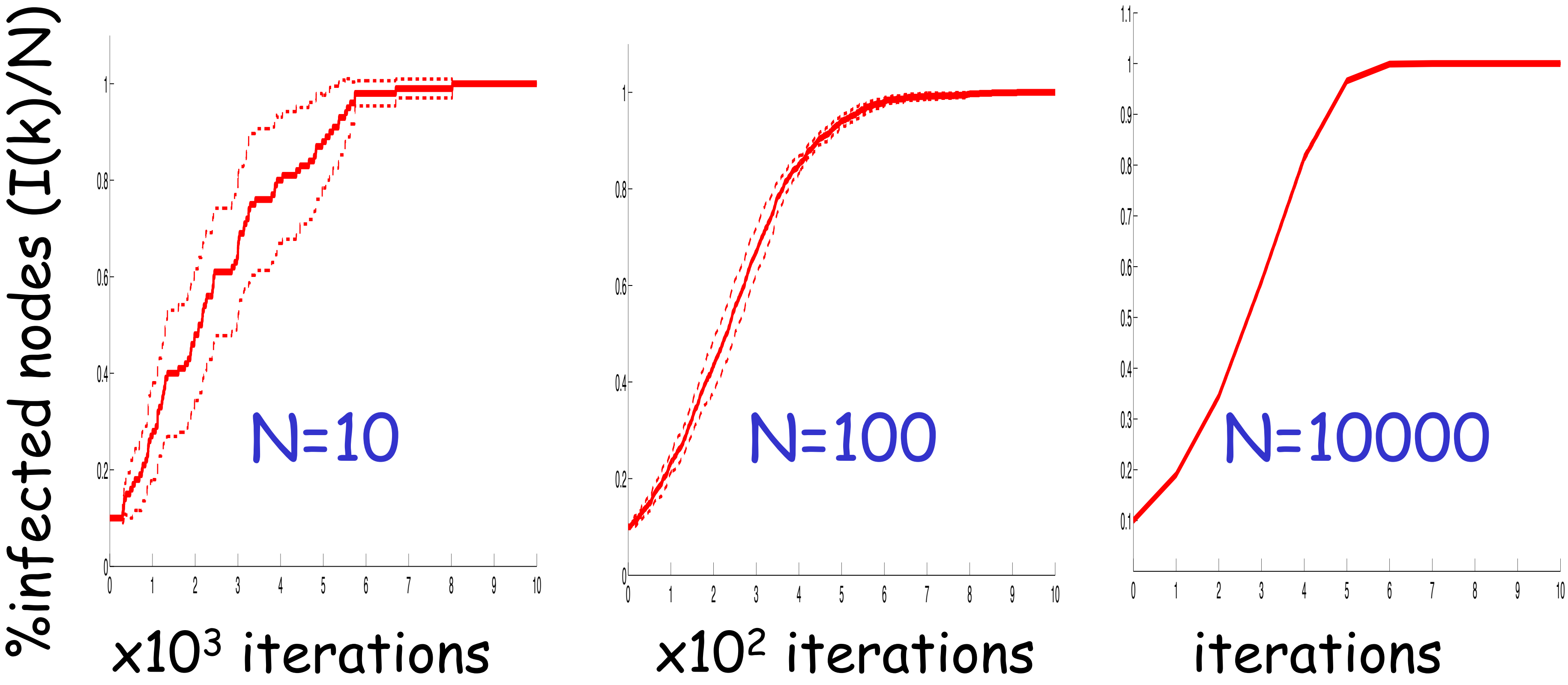
Performance Evaluation

Second Part
Lecture 2

Giovanni Neglia
INRIA – EPI Maestro
16 January 2012

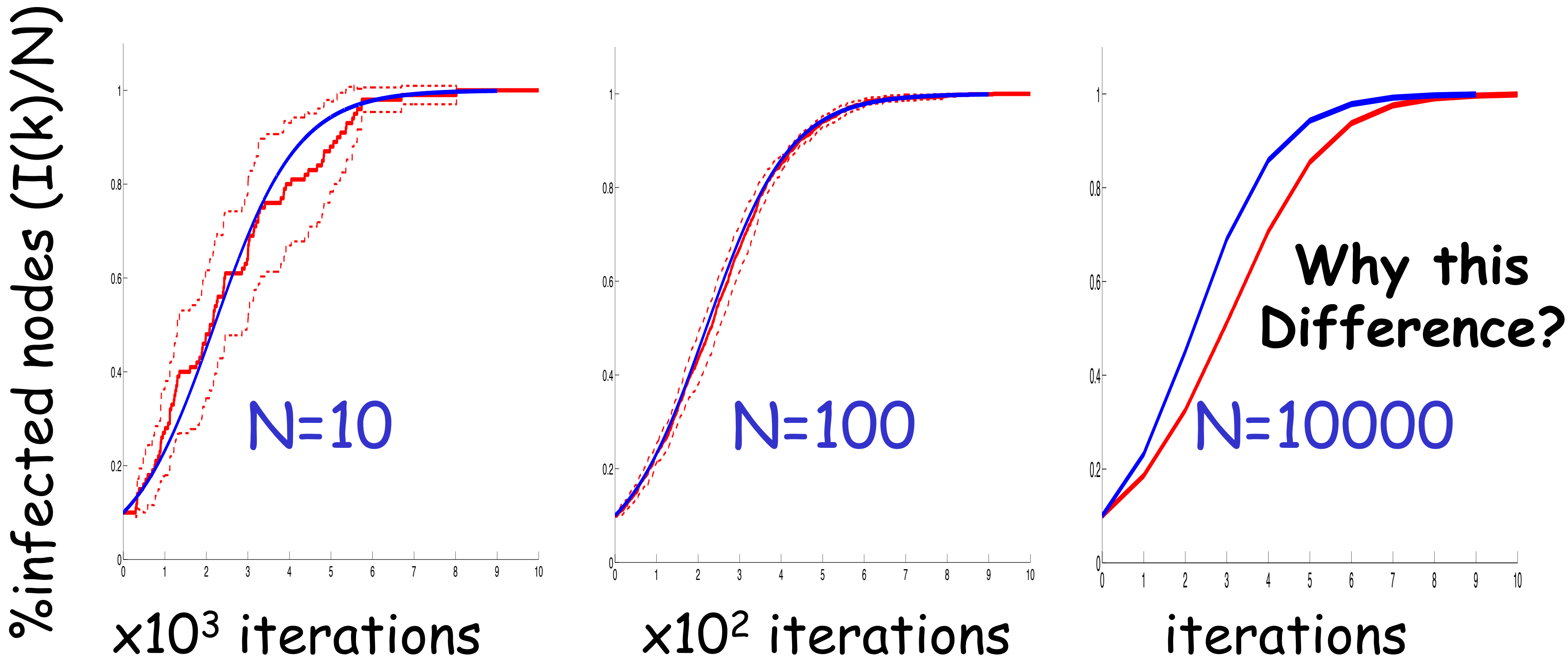
On approximation quality

$p=10^{-4}$, $I(0)=N/10$, 10 runs



On approximation quality
$p=10^{-4}$, $I(0)=N/10$, 10 runs

Model vs Simulations

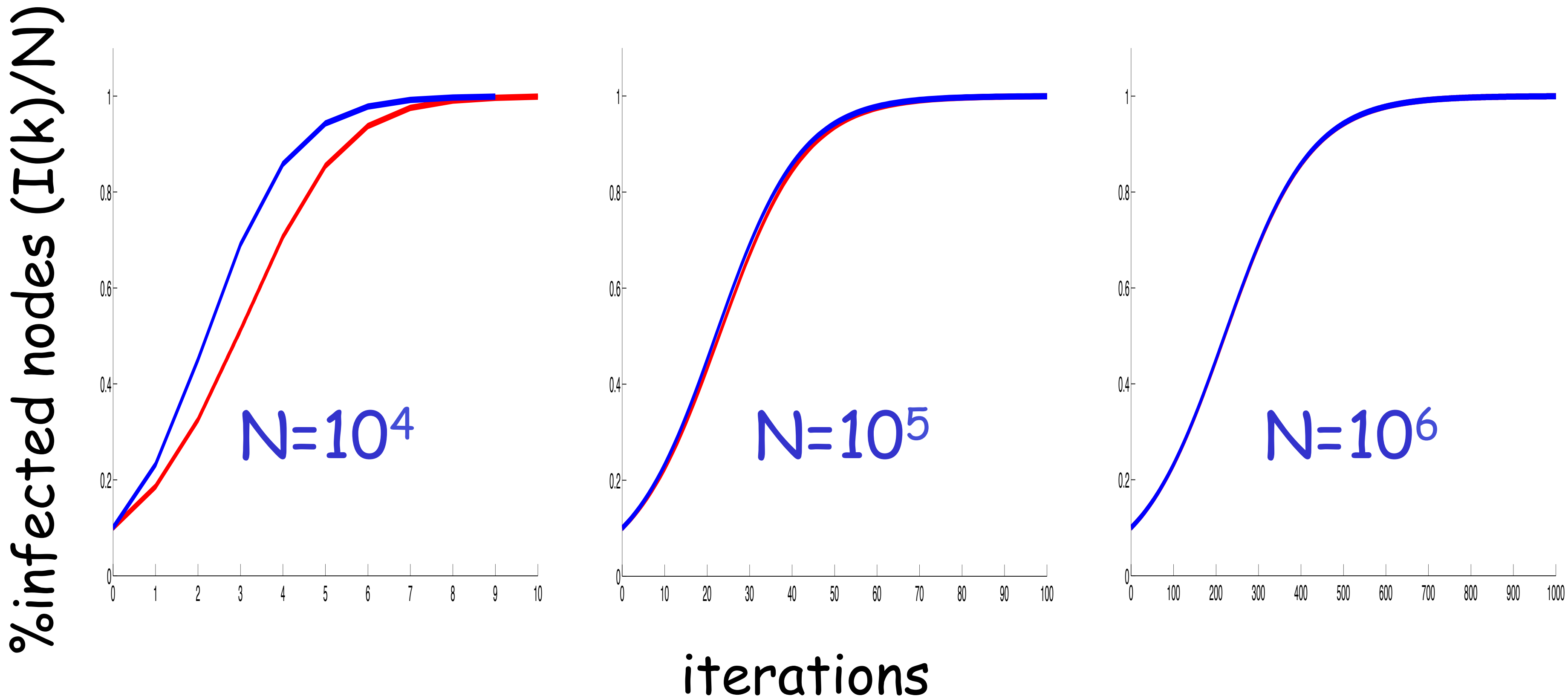


Why the difference?

- N should be large (the larger the better)
- p should be small
  - $p^{(N)}=p_0/N^2$
- For $N=10^4$ $p=10^{-4}$ is not small enough!
- What if we do the correct scaling?

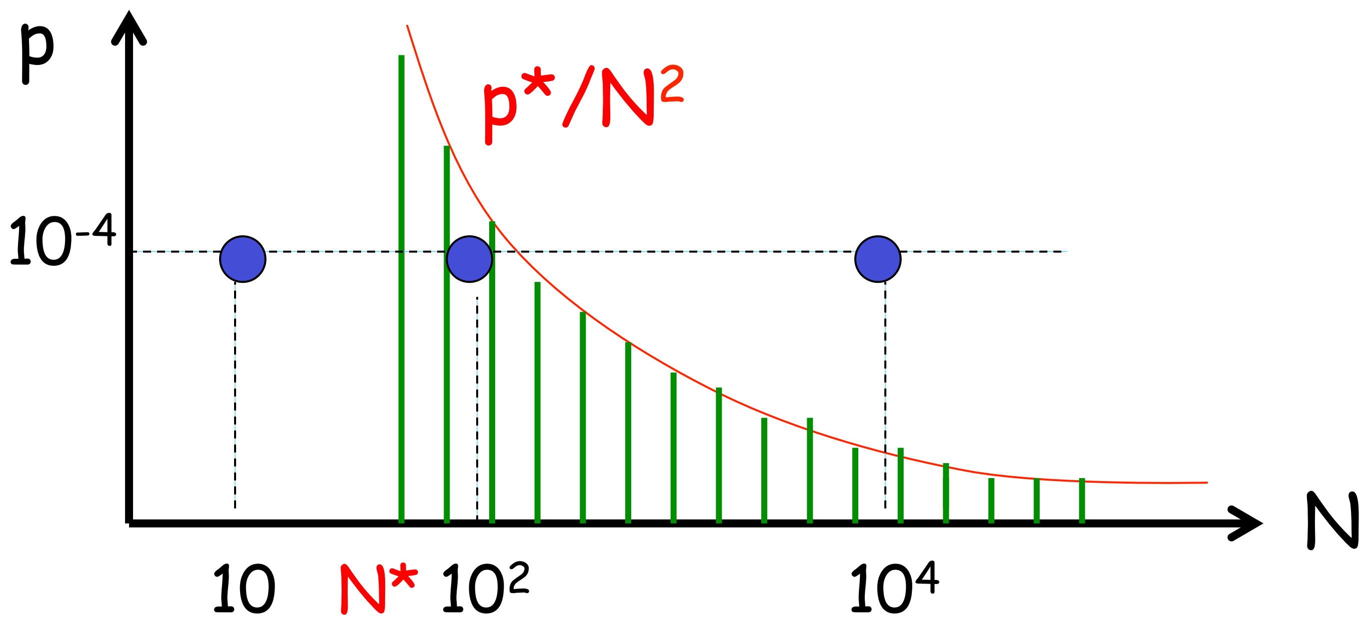
On approximation quality
$p=10^4/N^2$, $I(0)=N/10$, 10 runs

Model vs Simulations

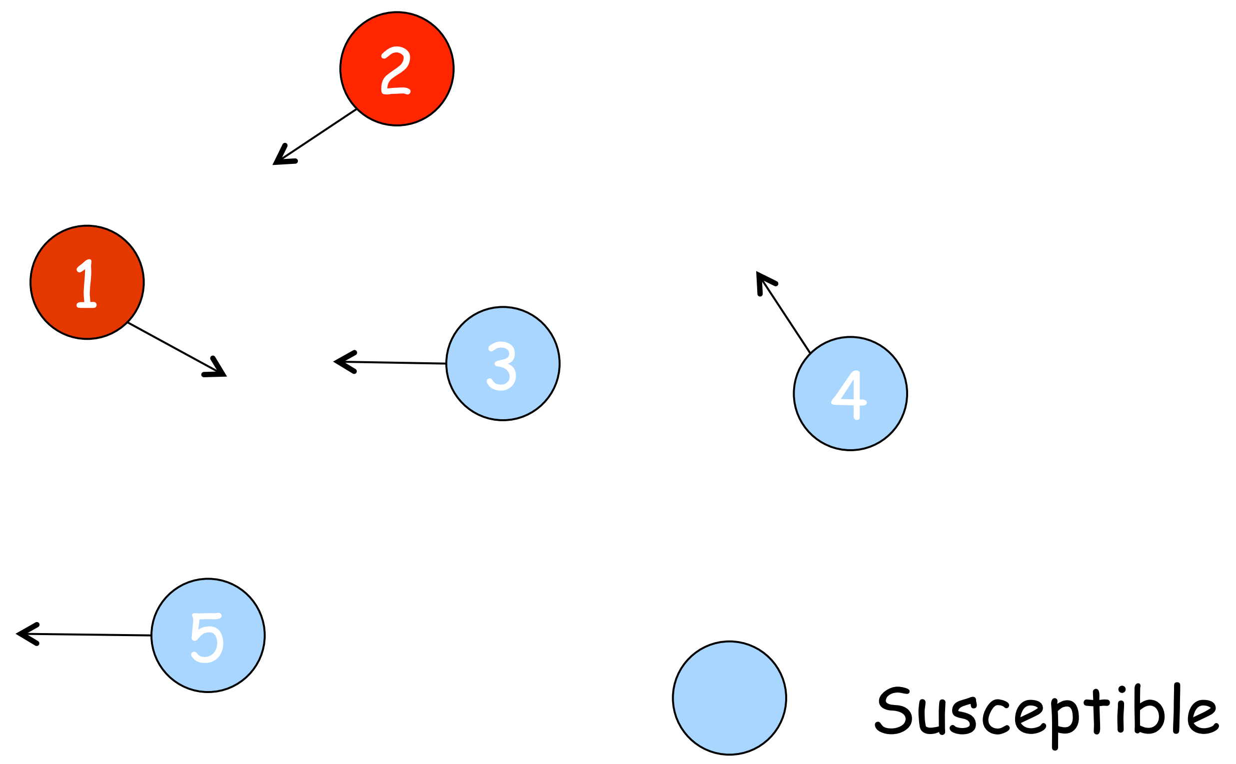


Lesson

- You need to check (usually by simulation) in which parameter region the fluid model is a good approximation.
  - e.g. $N>N^*$ $p<p^*/N^2$

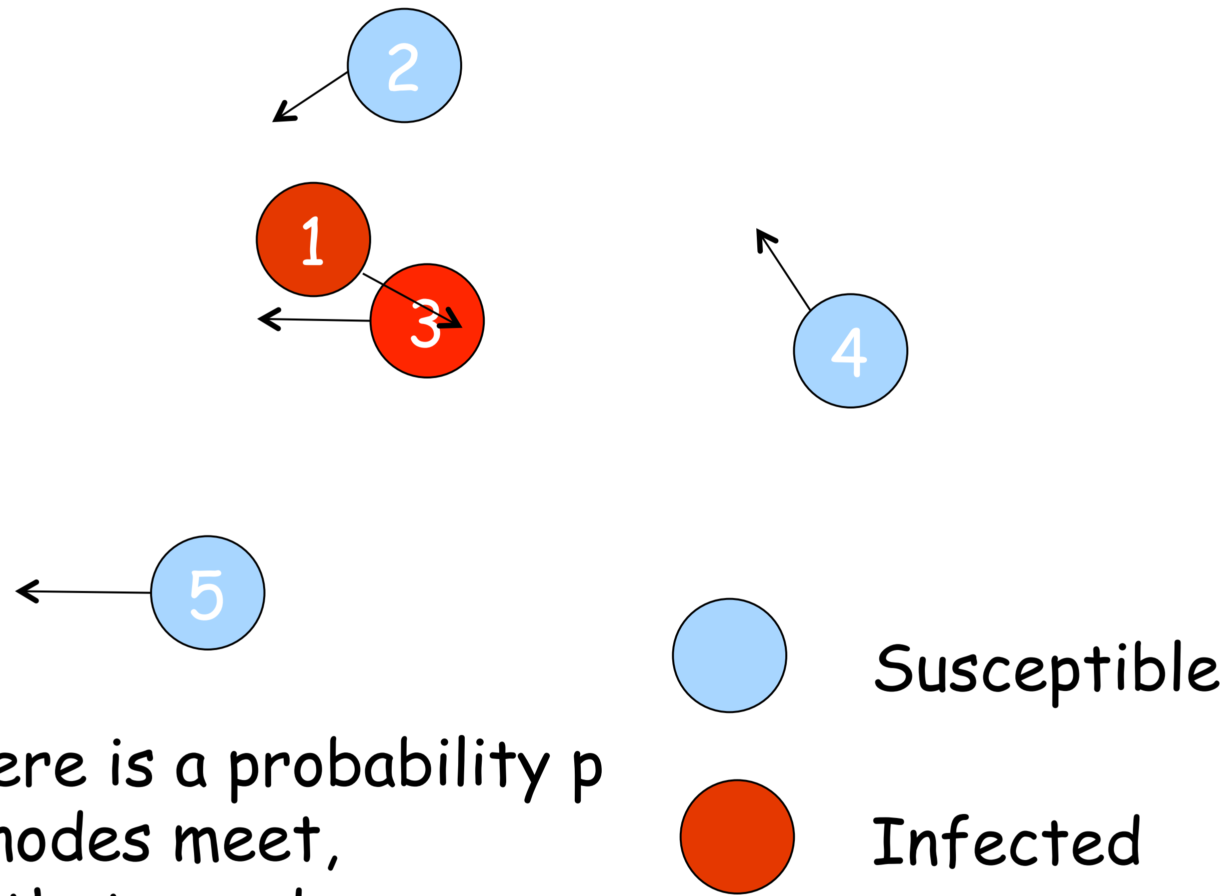


SIS model



At each slot there is a probability p that two given nodes meet, a probability r that a node recover.

SIS model



At each slot there is a probability p
that two given nodes meet,
a probability r that a node recovers.

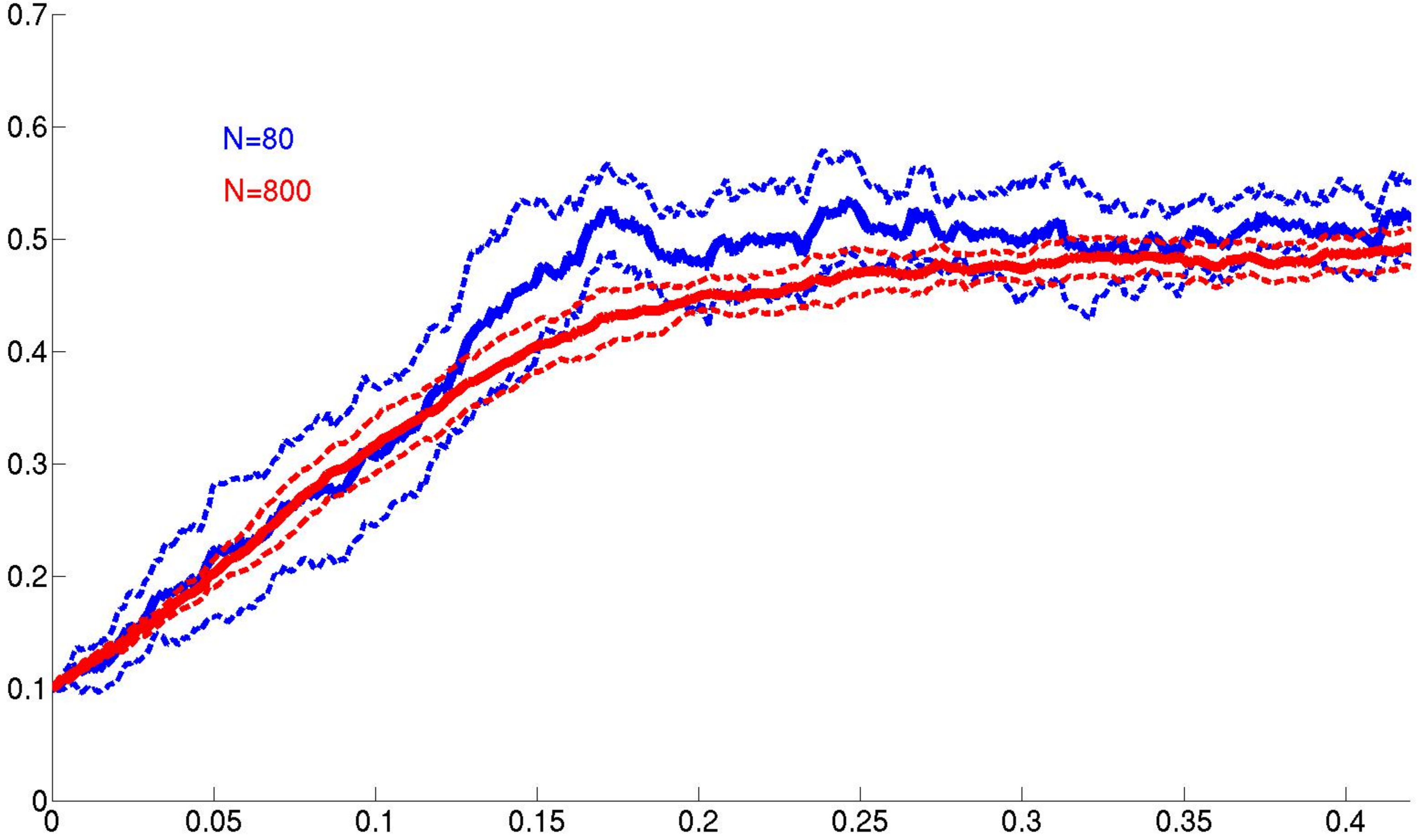
Study of the SIS model

- $d\mu_2(t)/dt = p_0\, \mu_2(t)(1-\mu_2(t)) - r_0\, \mu_2(t)$
- If $p_0 > r_0$ $\mu_2(\infty) = 1 - r_0/p_0$
- $Prob(X_1^{(N)}(k)=1) \approx \mu_2(k\varepsilon(N))$
  - $Prob(X_1^{(N)}(\infty)=1) \approx \mu_2(\infty) = 1 - r_0/p_0$
- What is the steady state distribution of the MC?
  - (0,0,0,…0) is the unique absorbing state and it is reachable from any other state
  - Who is lying here?

Back to the Convergence Result

- Define $\underline{\mathbf{M}}^{(N)}(t)$ with t real, such that
  - $\underline{\mathbf{M}}^{(N)}(k\varepsilon(N))=\mathbf{M}^{(N)}(k)$ for k integer
  - $\underline{\mathbf{M}}^{(N)}(t)$ is affine on $[k\varepsilon(N),(k+1)\varepsilon(N)]$
- Consider the Differential Equation
  - $d\boldsymbol{\mu}(t)/dt=\mathbf{f}(\boldsymbol{\mu})$, with $\boldsymbol{\mu}(0)=\mathbf{m}_0$
- Theorem
  - For all T>0, if $\mathbf{M}^{(N)}(0) \to \mathbf{m}_0$ in probability (/mean square) as $N \to \infty$, then $\sup_{0 \le t \le T} ||\underline{\mathbf{M}}^{(N)}(t)-\boldsymbol{\mu}(t)|| \to 0$ in probability (/ mean square)

Some examples



Nothing to do with $t=\infty$?

- Theorem 3: The limits when N diverges of the stationary distributions of $\mathbf{M}^{(N)}$ are included in the Birkhoff center of the ODE
  - Birkhoff center: the closure of all the recurrent points of the ODE (independently from the initial conditions)
  - What is the Birkhoff center of

    $di(t)/dt=p_0\, i(t)(1-i(t)) - r_0\, i(t)$?

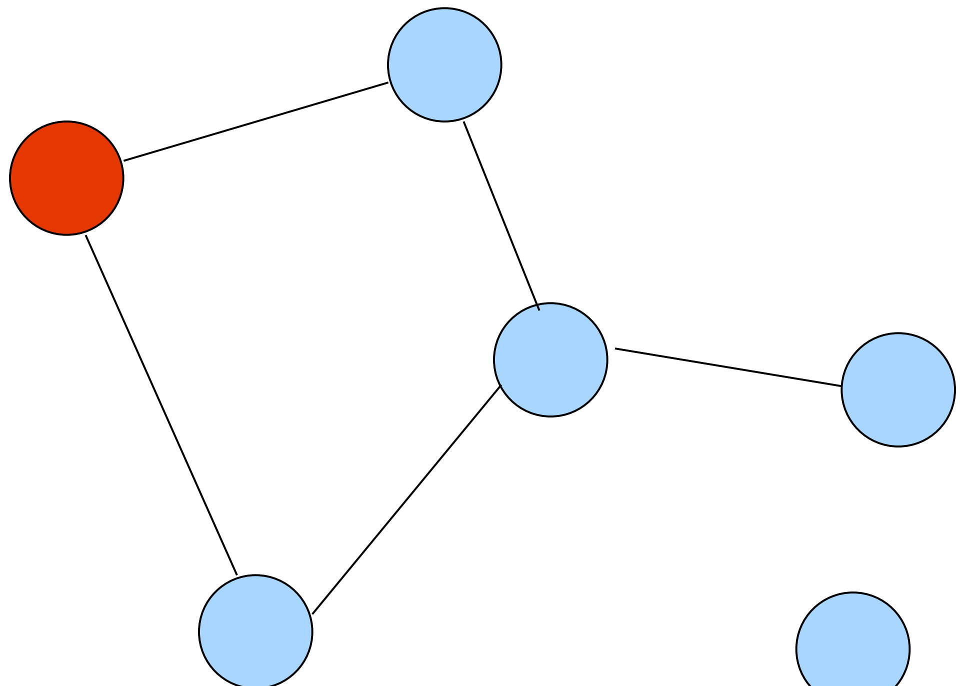
Nothing to do with $t=\infty$?

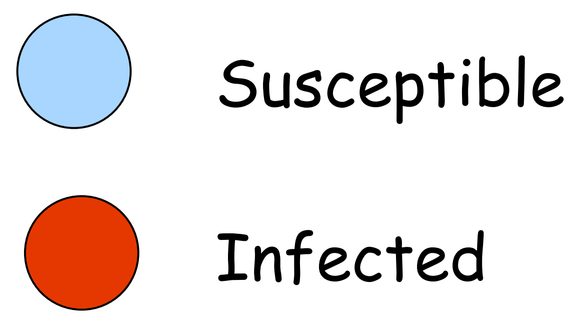
- Theorem 3: The limits when N diverges of the stationary distributions of $\mathbf{M}^{(N)}$ are included in the Birkhoff center of the ODE
- Corollary: If the ODE has a unique stationary point $\mathbf{m}^*$, the sequence of stationary distributions $\mathbf{M}^{(N)}$ converges to $\mathbf{m}^*$

Outline

- Limit of Markovian models
- Mean Field (or Fluid) models
  - exact results
  - Extensions
    - Epidemics on graphs
      - Reference: ch. 9 of Barrat, Barthélemy, Vespignani "Dynamical Processes on Complex Networks", Cambridge press
  - Applications to networks

SI on a graph

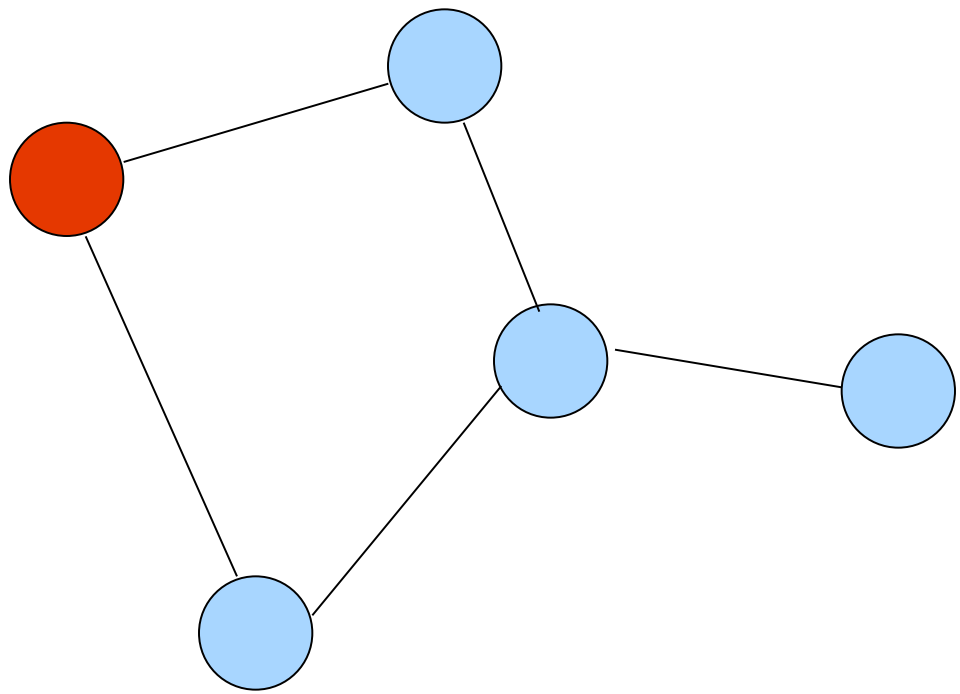




At each time slot, each link outgoing from an infected node spreads the disease with probability p_g

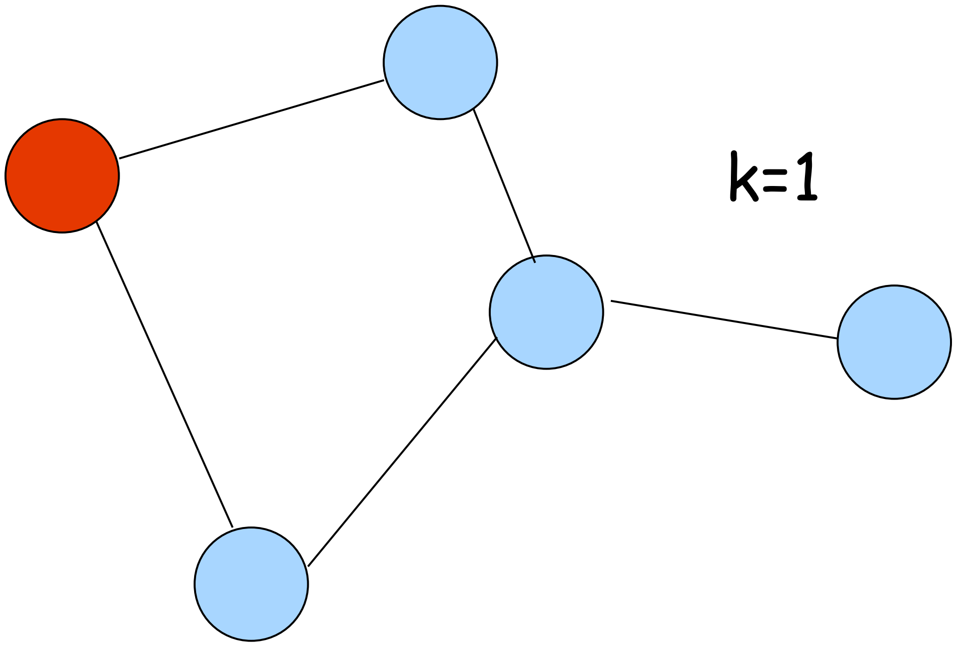
Can we apply Mean Field theory?

- Formally not, because in a graph the different nodes are not equivalent…
- …but we are stubborn



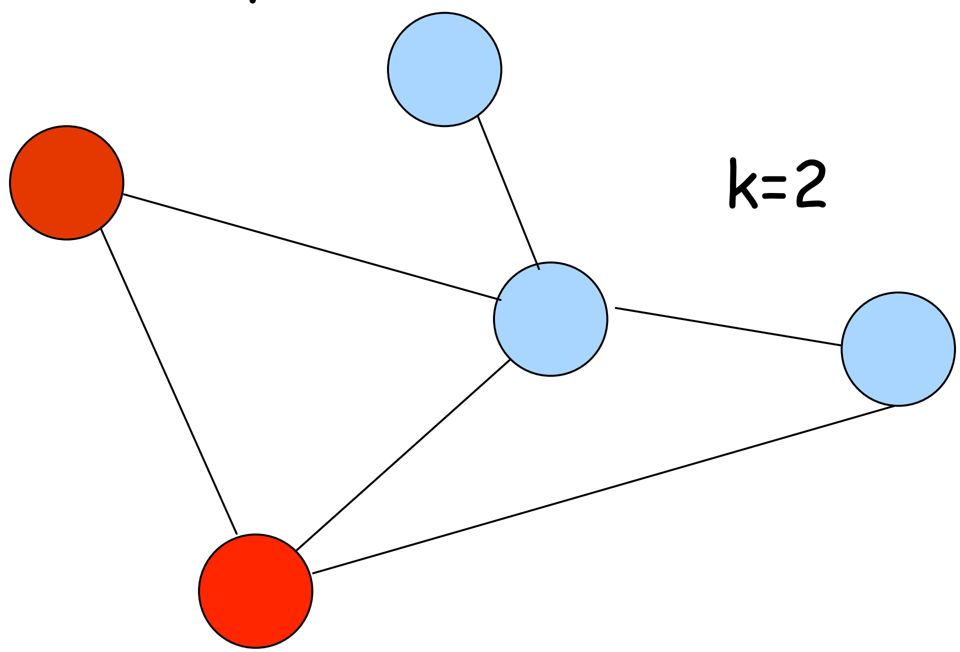
Derive a Mean Field model

- Consider all the nodes equivalent
- e.g. assume that at each slot the graph changes, while keeping the average degree <d>
  - Starting from an empty network we add a link with probability <d>/(N-1)



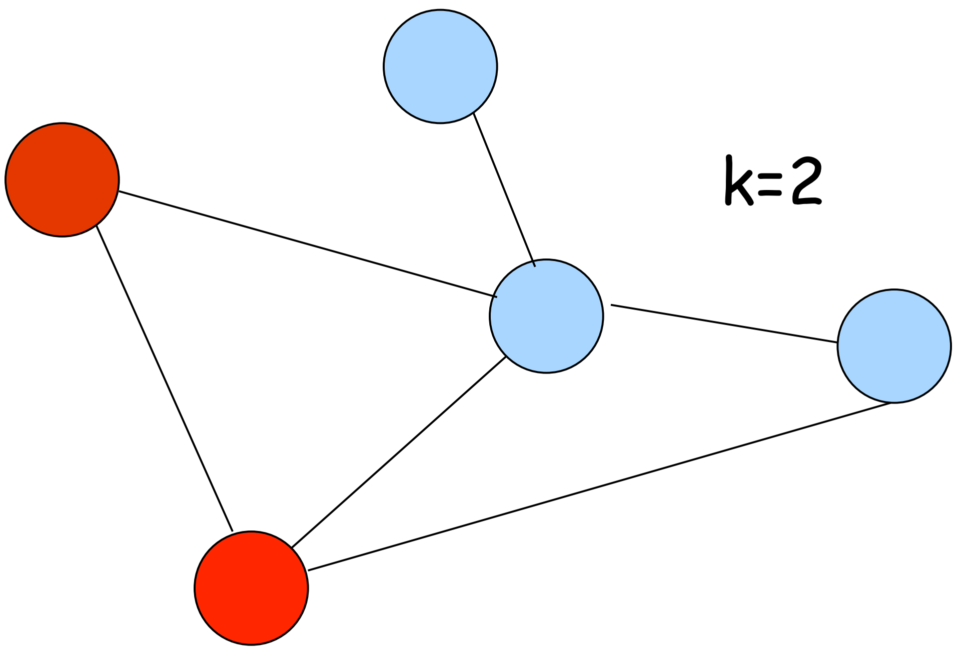
Derive a Mean Field model

- Consider all the nodes equivalent
- e.g. assume that at each slot the graph changes, while keeping the average degree <d>
  - Starting from an empty network we add a link with probability <d>/(N-1)



Derive a Mean Field model

- If $I(k)=I$, the prob. that a given susceptible node is infected is $q_I=1-(1-\langle d\rangle/(N-1)\ p_g)^I$
- and $(I(k+1)-I(k)|I(k)=I) =_d \mathrm{Bin}(N-I,\ q_I)$



Derive a Mean Field model

- If $I(k)=I$, the prob. that a given susceptible node is infected is $q_I=1-(1-\langle d\rangle/(N-1)\ p_g)^I$
- and $(I(k+1)-I(k)|I(k)=I) =_d \mathrm{Bin}(N-I,\ q_I)$
  - Equivalent to first SI model where $p=\langle d\rangle/(N-1)\ p_g$
  - We know that we need $p^{(N)}=p_0/N^2$
- $i^{(N)}(k) \approx \mu_2\,(k\varepsilon(N))=1/((1/i_0-1)\exp(-k\ p_0/N)+1)=$
  $= 1/((1/i_0-1)\exp(-k\ \langle d\rangle\ p_g)+1)$
  - The percentage of infected nodes becomes significant after the **outbreak time** $1/(\langle d\rangle p_g)$
- How good is the approximation practically?
  - It depends on the graph!

Erdös-Rényi graph

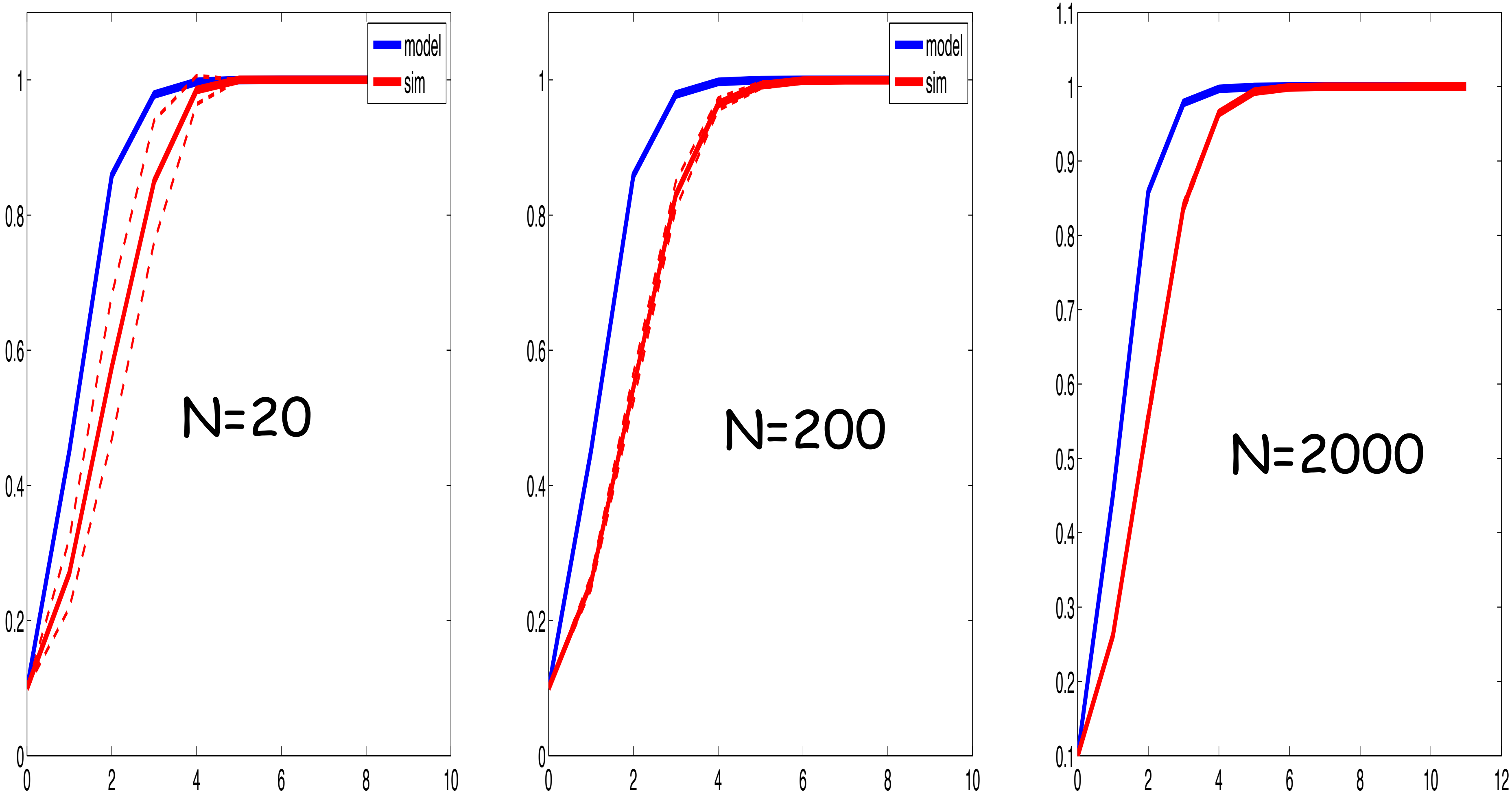
- A ER graph G(N,q) is a stochastic process
  - N nodes and edges are selected with prob. q
- Purpose: abstract from the details of a given graph and being able to reach conclusions depending on its average features

Erdös-Rényi graph

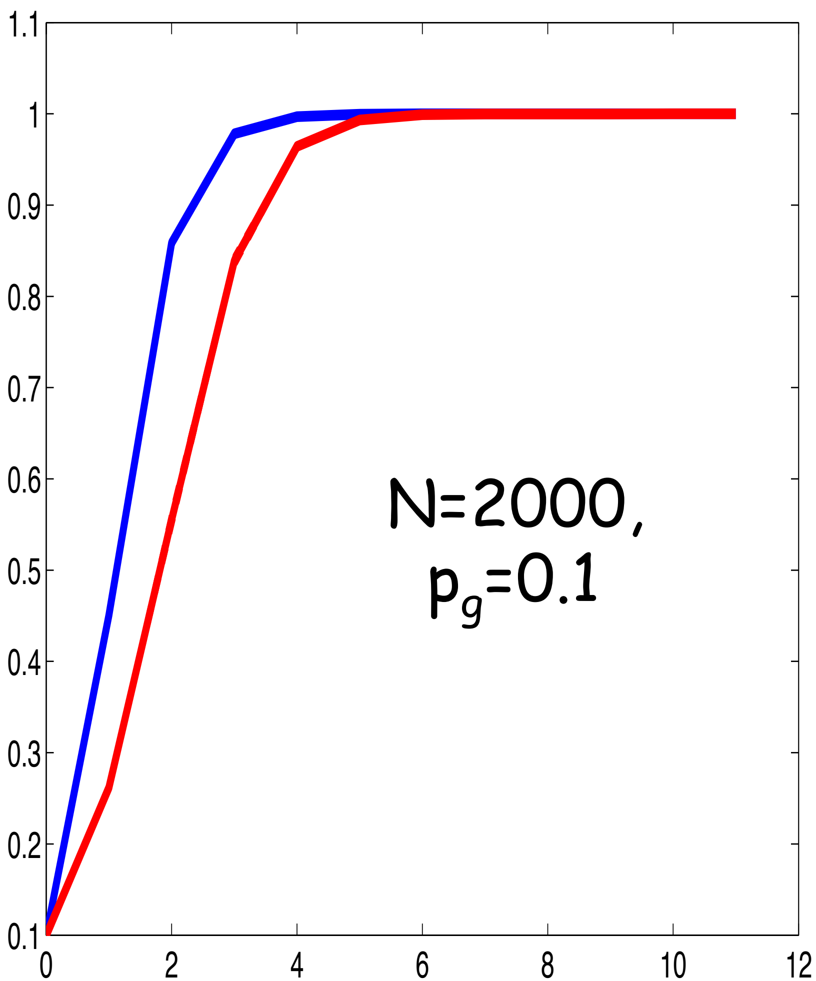
- A ER graph $G(N,q)$ is a stochastic process
  - N nodes and edges are selected with prob. q
  - Degree distribution: $P(d)=C^d_{N-1}\, q^d(1-q)^{N-1-d}$
    - For $N\to\infty$ and Nq constant: $P(d)=e^{-\langle d\rangle}\langle d\rangle^d/d!$
    - $\langle d^2\rangle=\langle d\rangle(1+\langle d\rangle)$
    - Average degree: $\langle d\rangle=q\,(N-1)$
  - Average distance: $\langle l\rangle\approx \log N/\log\langle d\rangle$
    - Small world
- Remark: in the calculations above we had a different sample of an ER graph at each slot, in what follows we consider a single sample

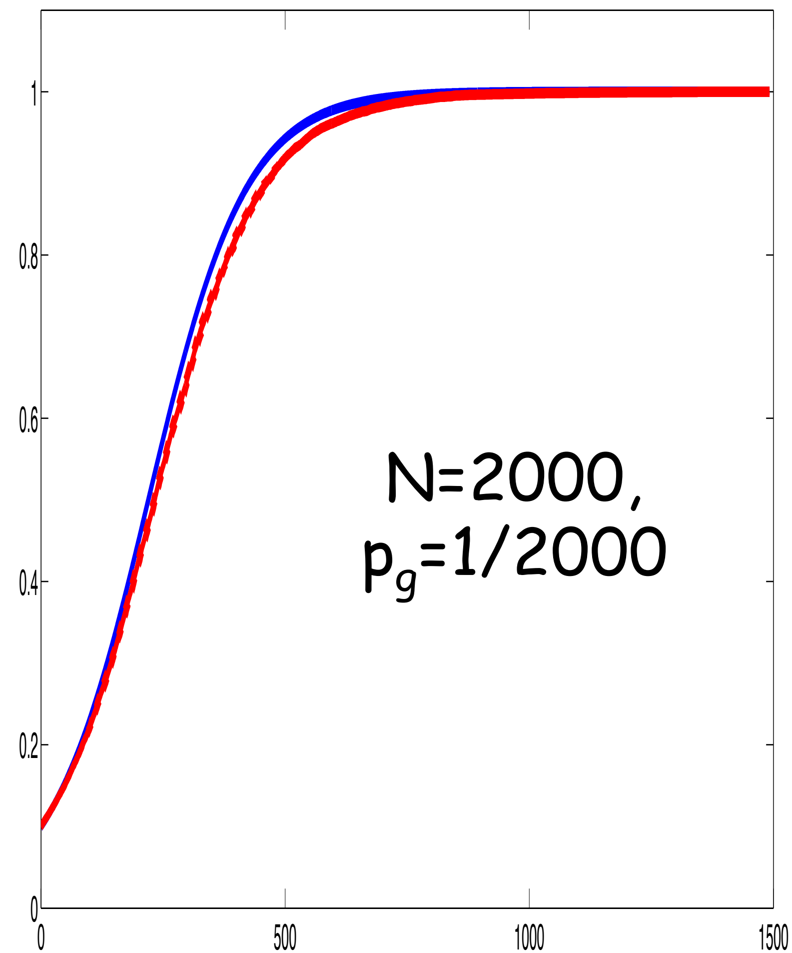
ER $\langle d \rangle$=20, p_g=0.1, 10 runs

$$i^{(N)}(k) \approx 1/((1/i_o-1)\ \exp(-k\ \langle d \rangle\ p_g)+1)$$



ER <d>=20, 10 runs

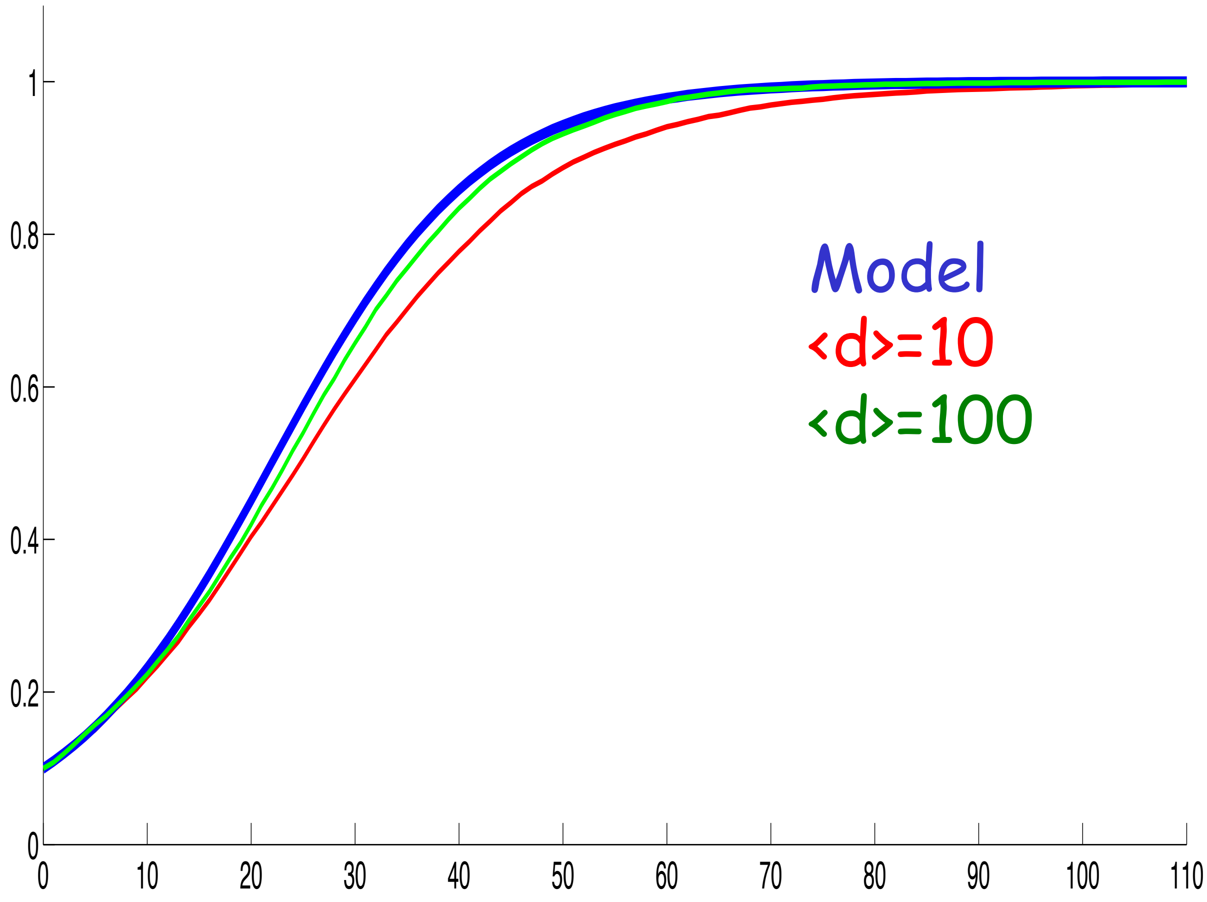




Lesson 1

- System dynamics is more *deterministic* the larger the network is
  - Why?
- For given <d>, the MF solution shows the same relative error

Changing the degree
ER N=1000, <d>p_g=0.1, 10 runs

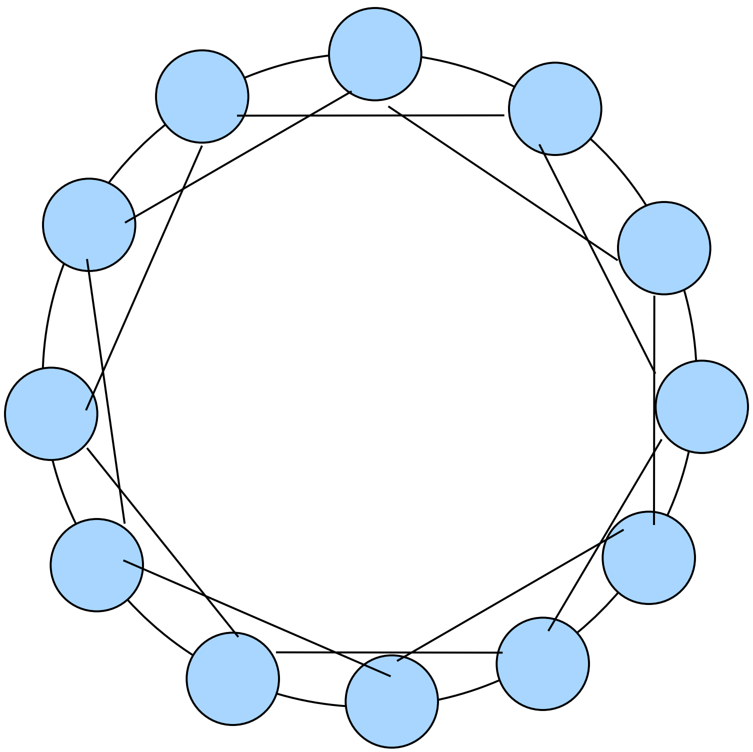


$$i^{(N)}(k) \approx 1/((1/i_0-1)\exp(-k \langle d\rangle p_g)+1)$$

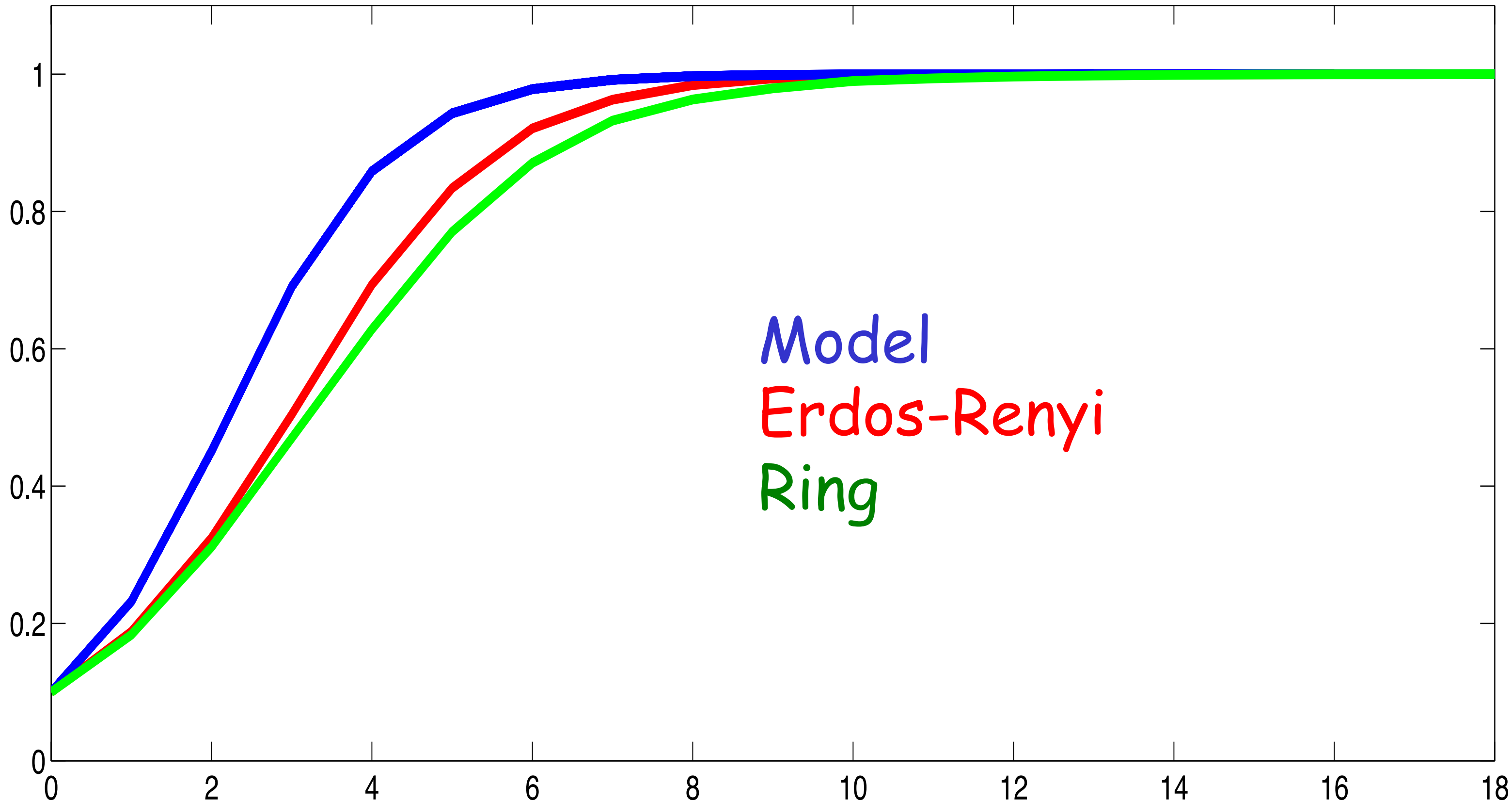
Lesson 2

- The more the graph is connected, the better the MF approximation
  - Why?

A different graph Ring(N,k)



Ring vs ER, N=2000, <k>=10



Lesson 3

- The smaller the clustering coefficient, the better the MF approximation
  - Why?

Outline

- Limit of Markovian models
- Mean Field (or Fluid) models
  - exact results
  - extensions
  - Applications
    - Bianchi's model
    - Epidemic routing

Decoupling assumption in Bianchi's model

- Assuming that retransmission processes at different nodes are independent
  - Not true: if node i has a large backoff window, it is likely that also other nodes have large backoff windows
- We will provide hints about why it is possible to derive a Mean Field model…
- then the decoupling assumption is guaranteed asymptotically

References

- Benaïm, Le Boudec, "A Class of Mean Field Interaction Models for Computer and Communication Systems", LCA-Report-2008-010
- Sharma, Ganesh, Key, "Performance Analysis of Contention Based Medium Access Control Protocols", IEEE Trans. Info. Theory, 2009
- Bordenave, McDonarl, Proutière, "Performance of random medium access control, an asymptotic approach", Proc. ACM Sigmetrics 2008, 1-12, 2008

Bianchi's model

- N nodes,
- K possible stages for each node, in stage i (i=1,…V) the node transmit with probability $q^{(N)}_i$ (e.g. $q^{(N)}_i = 1/W^{(N)}_i$)
- If a node in stage i experiences a collision, it moves to stage i+1
- If a node transmits successfully, it moves to stage 0

Mean Field model

- We need to scale the transmission probability: $q^{(N)}_i = q_i/N$
- $\mathbf{f}^{(N)}(\mathbf{m})=E[\mathbf{M}^{(N)}(k+1)-\mathbf{M}^{(N)}(k)|\mathbf{M}^{(N)}(k)=\mathbf{m}]$
- $f_1^{(N)}(\mathbf{m})=E[M_1^{(N)}(k+1)-M_1^{(N)}(k)|M_1^{(N)}(k)=\mathbf{m}]$
- $P_{idle}=\Pi_{i=1,\dots V}(1-q_i^{(N)})^{m_i N}$
- The number of nodes in stage 1
  - increases by one if there is one successful transmission by a node in stage i<>1
  - Decreases if a node in stage 1 experiences a collision

Mean field model

- $P_{idle}=\Pi_{i=1,\dots V}(1-q_i^{(N)})^{m_i N} \text{->} \exp(-\Sigma_i q_i\, m_i)$
  - Define $\tau(m)= \Sigma_i q_i\, m_i$
- The number of nodes in stage 1
  - increases by one if there is one successful transmission by a node in stage i<>1
    - with prob. $\Sigma_{i>1}\, m_i\, N\, q_i^{(N)}\, P_{idle}/(1-q_i^{(N)})$
  - Decreases if a node in stage 1 experiences a collision
    - with prob. $m_1\, N\, q_1^{(N)}\,(1-P_{idle}/(1-q_1^{(N)})$
- $f_1^{(N)}\,(\mathbf{m})=E[M_1^{(N)}(k+1)-M_1^{(N)}(k)|M_1^{(N)}(k)=\mathbf{m}]=$

  $= \Sigma_{i>1} m_i q_i^{(N)} P_{idle}/(1-q_i^{(N)})$

  $- m_1 q_1^{(N)}(1-P_{idle}/(1-q_1^{(N)}))$

Mean field model

- $P_{idle}=\Pi_{i=1,\dots V}(1-q_i^{(N)})^{m_i N} \to \exp(-\Sigma_i q_i\, m_i)$
  - Define $\tau(\mathbf{m})= \Sigma_i q_i\, m_i$
- $f_1^{(N)}(\mathbf{m})=\Sigma_{i>1} m_i q_i^{(N)} P_{idle}/(1-q_i^{(N)})$
  $- m_1 q_1^{(N)}(1-P_{idle}/(1-q_1^{(N)}))$
- $f_1^{(N)}(\mathbf{m}) \sim 1/N \left(\Sigma_{i>1} m_i q_i\, e^{-\tau(\mathbf{m})} - m_1 q_1 (1-e^{-\tau(\mathbf{m})})\right)$
- $f_1^{(N)}(\mathbf{m})$ vanishes and $\varepsilon(N)=1/N$, continuously differentiable in $\mathbf{m}$ and in $1/N$
- This holds also for the other components
- Number of transitions bounded

=> We can apply the Theorem