Evolving (Preferential Attachment) Networks

Sidney Redner, Boston University, physics.bu.edu/~redner
collaborators: P. Krapivsky, F. Leyvraz

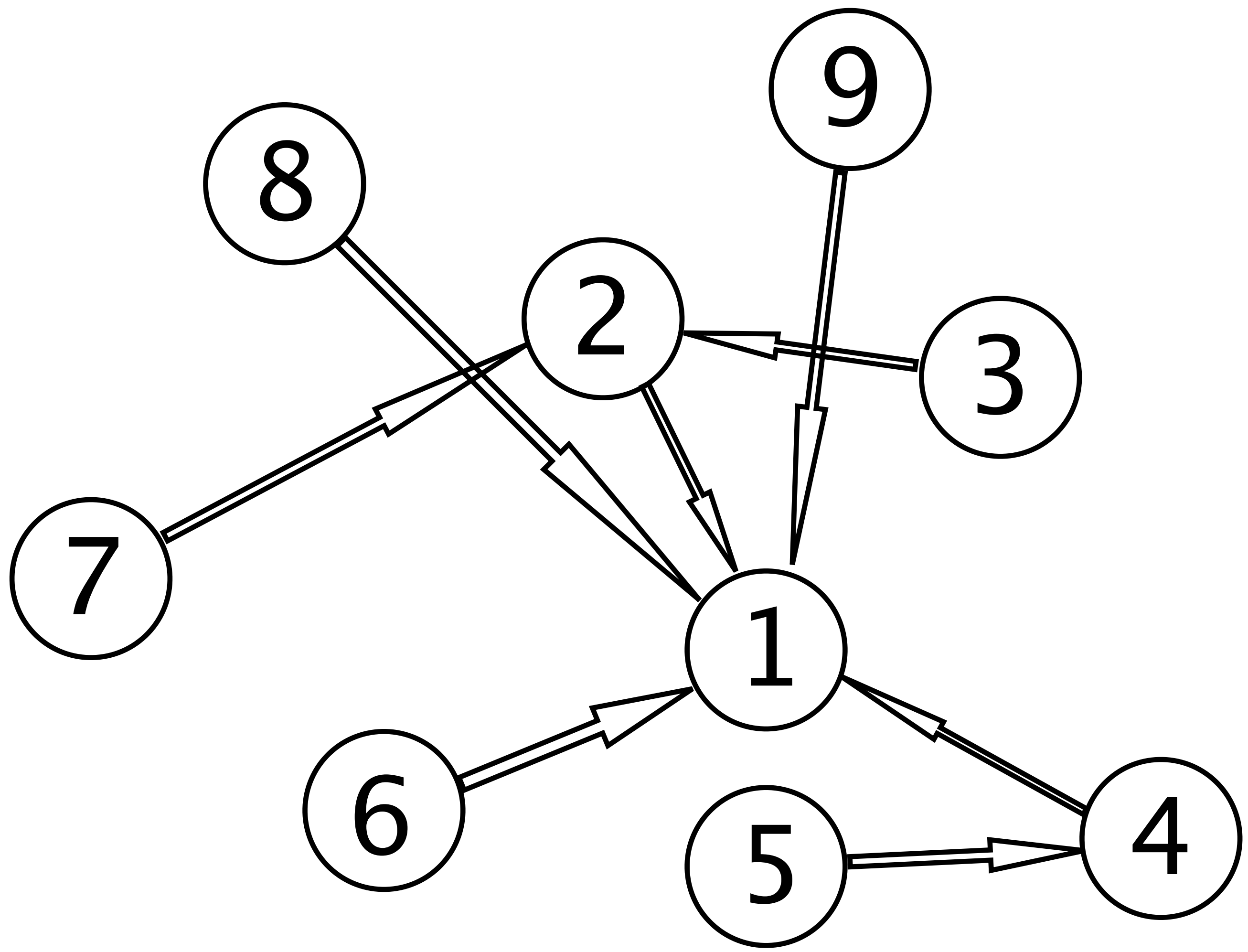
E(uropean) N(etwork) on RA(ndom) GE(ometry), June 2008

Outline:

- master equation: degree distributions
- ubiquitous linear preferential attachment:
  - *redirection & copying*
- a deeper look: correlations, genealogy, fitness
- do the rich really get richer?
- web growth model
- finiteness & fluctuations

Preferential Attachment Network

Yule (1926), Simon (1955)
Barabasi & Albert (1999)



1. Introduce nodes one at a time
2. Attach to one earlier node with k links at rate A_k

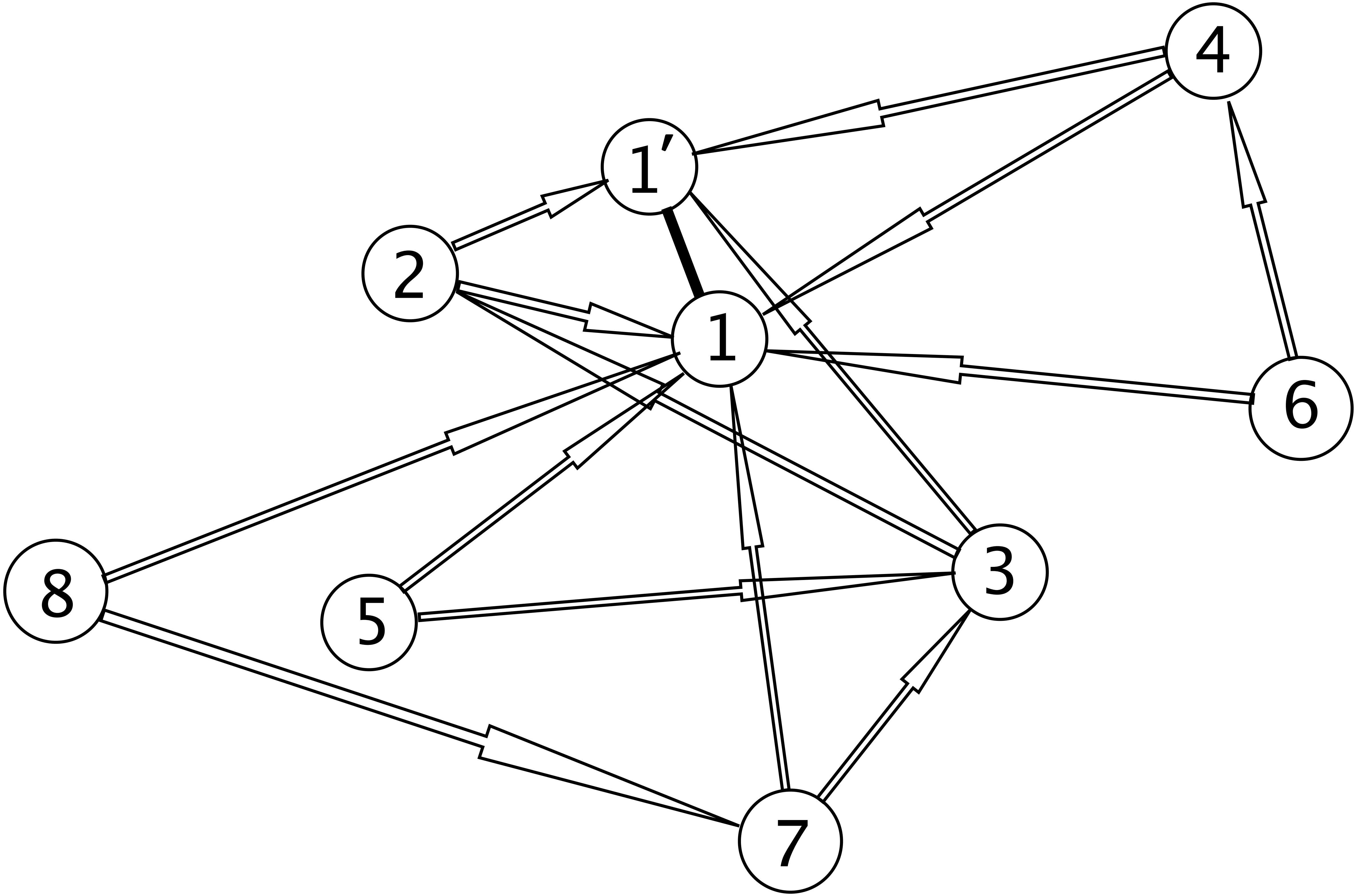
preferential attachment: A_k *increasing in k*

"rich get richer"

Basic observable: N_k, *average* number of nodes with k links

the degree distribution

Larger m=2



Master Equation Approach

KRL, KR (2000)
Dorogovtsev et al. (2000)

Basic observable: N_k, *average* number of nodes with k links

the degree distribution

Master Equation:

attach to node of degree k-1 | attach to node of degree k | create node of degree 1

$$\frac{dN_k}{dN} = \frac{A_{k-1}N_{k-1}}{A} - \frac{A_k N_k}{A} + \delta_{k,1}$$

$A = \text{total rate}$

Attachment Rate: $A_k \sim k^\gamma$

Total Rate:

$$A = A(N) = \sum_{j=1}^{\infty} A_j N_j = \sum_{j=1}^{\infty} j^\gamma N_j \equiv M_\gamma(N)$$

Moment equations:

$$\dot{M}_0 \equiv \sum_j \dot{N}_j = 1; \quad \dot{M}_1 \equiv \sum_j j\dot{N}_j = 2$$

These suggest: $A(N) = \sum_j j^{\gamma} N_j \propto \mu(\gamma) N \quad \text{for } 0 \leq \gamma \leq 1$

$$N_k(N) \equiv N \; n_k$$

Converts rate eqns. to linear recursions

$$\frac{dN_k}{dN} = \frac{A_{k-1} N_{k-1}}{A} - \frac{A_k N_k}{A} + \delta_{k,1}$$

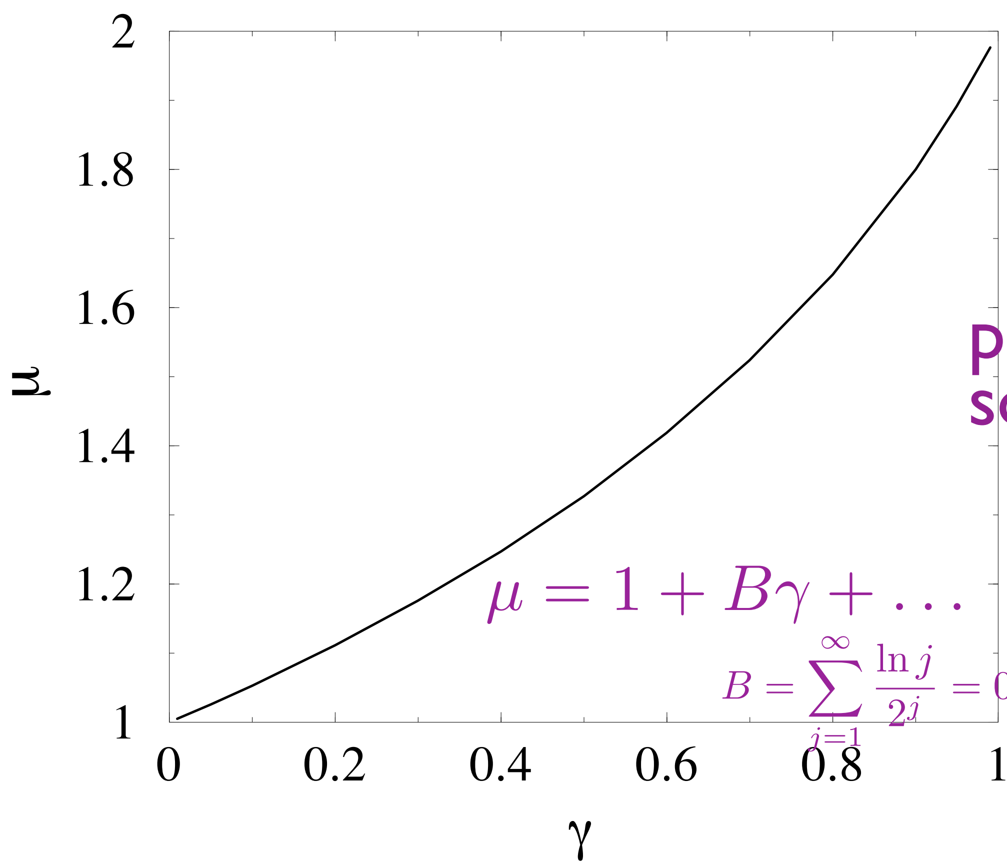
$$\Rightarrow \quad n_k = \frac{A_{k-1} n_{k-1} - A_k n_k}{\mu} + \delta_{k,1}$$

Formal Solution

$$n_k = \frac{A_{k-1}n_{k-1} - A_k n_k}{\mu} + \delta_{k,1} \longrightarrow n_k = \frac{\mu}{A_k}\prod_{j=1}^{k}\left(1+\frac{\mu}{A_k}\right)^{-1}$$

$$\text{determination of } \mu = \sum_{k=1}^{\infty} A_k n_k: \longrightarrow 1 = \sum_{k=1}^{\infty}\prod_{j=1}^{k}\left(1+\frac{\mu}{A_k}\right)^{-1}$$

numerical solution for $A_k = k^\gamma$



$$\mu = 2 - C(1-\gamma) + \dots$$

$$C = 4\sum_{j=1}^{\infty}\frac{\ln j}{(j+1)(j+2)} = 2.407\dots$$

perturbative solutions

$$\mu = 1 + B\gamma + \dots$$

$$B = \sum_{j=1}^{\infty}\frac{\ln j}{2^j} = 0.5078\dots$$

Formal solution: $n_k = \frac{\mu}{A_k} \prod_{j=1}^{k} \left(1 + \frac{\mu}{A_j}\right)^{-1}$ $\quad n_k = \frac{A_{k-1}n_{k-1} - A_k n_k}{\mu} + \delta_{k,1}$

Asymptotics for $A_k \sim k^\gamma$

$$n_k \sim \begin{cases} k^{-\gamma}\, e^{-[\mu k^{1-\gamma}/(1-\gamma)]} & 0 \le \gamma < 1 \end{cases}$$

universal, generic

a little more carefully...

$$n_k = \frac{\mu}{A_k} \prod_{j=1}^{k} \left(1 + \frac{\mu}{A_j}\right)^{-1} \sim \frac{\mu}{k} \exp\left[-\int_1^k \ln\left(1 + \frac{\mu}{j^\gamma}\right)\right] dj$$

$$\sim \frac{\mu}{k} \exp\left[-\int_1^k \left(\frac{\mu}{j^\gamma} - \frac{\mu^2}{2j^{2\gamma}} + \cdots\right)\right] dj$$

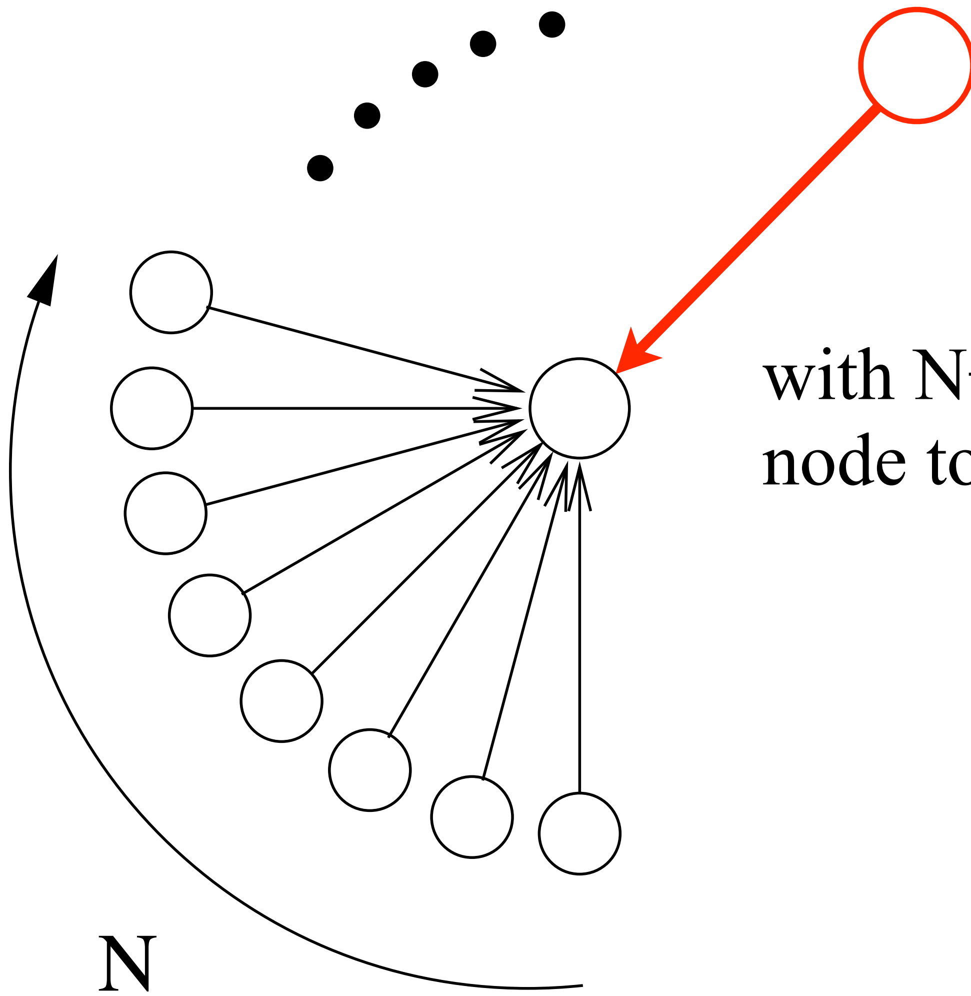
$$n_k \sim \begin{cases} k^{-\gamma} \exp\left[-\mu\left(\frac{k^{1-\gamma}-1}{1-\gamma}\right)\right] & \frac{1}{2} < \gamma < 1 \\ k^{\frac{\mu^2-1}{2}} \exp\left[-2\mu\sqrt{k}\right] & \gamma = \frac{1}{2} \\ k^{-\gamma} \exp\left[-\mu\frac{k^{1-\gamma}}{1-\gamma} + \frac{\mu^2}{2}\frac{k^{1-2\gamma}}{1-2\gamma}\right] & \frac{1}{3} < \gamma < \frac{1}{2} \\ \vdots & \\ \vdots & \end{cases}$$

Formal solution: $n_k = \frac{\mu}{A_k} \prod_{j=1}^{k} \left(1 + \frac{\mu}{A_j}\right)^{-1}$

Asymptotics for $A_k \sim k^\gamma$

$$n_k \sim \begin{cases} k^{-\gamma}\, e^{-[\mu k^{1-\gamma}/(1-\gamma)]} & 0 \le \gamma < 1 \\ \\ \text{“bible”} & \gamma > 2 \end{cases}$$

Creating a "Bible"



with N+1 total nodes: attach new node to bible with probability $\frac{N^\gamma}{N + N^\gamma}$

bible probability: $$\mathcal{P} = \prod_{N=1}^{\infty} \frac{1}{1 + N^{1-\gamma}} = \begin{cases} \text{zero} & \gamma \leq 2 \\ \text{non-zero} & \gamma > 2 \end{cases}$$

Formal solution: $n_k = \frac{\mu}{A_k} \prod_{j=1}^{k} \left(1 + \frac{\mu}{A_j}\right)^{-1}$

Asymptotics for $A_k \sim k^\gamma$

$$n_k \sim \begin{cases} k^{-\gamma}\, e^{-[\mu k^{1-\gamma}/(1-\gamma)]} & 0 \le \gamma < 1 \\ \\ \text{"best seller"} & 1 < \gamma \le 2 \\ \text{"bible"} & \gamma > 2 \end{cases}$$

"Best-Seller" Phase: $1<\gamma\leq 2$

fact: $A(N) = \sum_k k^\gamma N_k \sim N^\gamma$

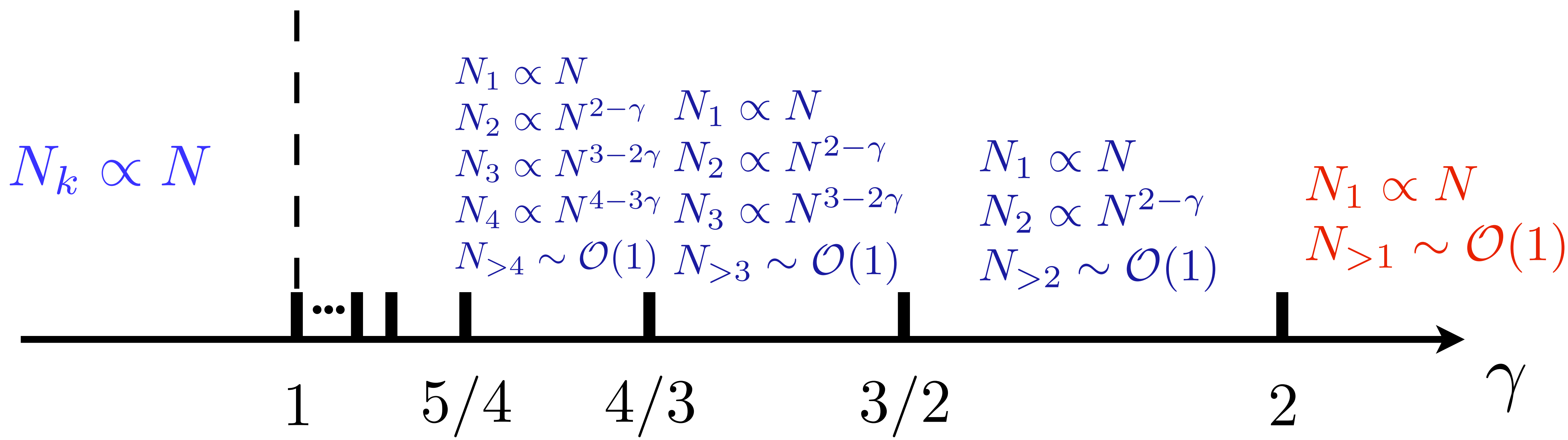
$$\dot{N}_1 = 1 - \frac{N_1}{A} \quad \rightarrow N_1 \propto N$$

$$\dot{N}_2 = \frac{N_1 - 2^\gamma N_2}{N^\gamma} \propto N^{1-\gamma} \rightarrow N_2 \propto N^{2-\gamma}$$

$$\dot{N}_3 = \frac{2^\gamma N_2 - 3^\gamma N_3}{N^\gamma} \propto N^{2-2\gamma} \rightarrow N_3 \propto N^{3-2\gamma}$$

$$\longrightarrow N_k \propto N^{k-(k-1)\gamma}$$

for positive exponent



Formal solution: $n_k = \frac{\mu}{A_k} \prod_{j=1}^{k} \left(1 + \frac{\mu}{A_j}\right)^{-1}$

Asymptotics for $A_k \sim k^\gamma$

$$n_k \sim \begin{cases} k^{-\gamma}\, e^{-[\mu k^{1-\gamma}/(1-\gamma)]} & 0 \le \gamma < 1 \\ k^{-\nu},\ \nu > 2\ \ \textbf{!!!} & \gamma = 1 \\ \text{“best seller”} & 1 < \gamma \le 2 \\ \text{“bible”} & \gamma > 2 \end{cases}$$

Non-Universal Degree Distributions

$$\begin{aligned} n_k &= \frac{\mu}{A_k} \prod_{j=1}^{k} \left(1 + \frac{\mu}{A_j}\right)^{-1} \\ &\sim \frac{\mu}{k} \exp\left[-\int_1^k \ln\left(1 + \frac{\mu}{j}\right)\right] dj \\ &\sim \mu\, k^{-(1+\mu)} \equiv k^{-\nu} \end{aligned}$$

$$\nu > 2 \quad \text{for } \mu = \sum A_j n_j \sim \sum j n_j < \infty$$

Special case, $A_k = k$: $\qquad n_k = \dfrac{4}{k(k+1)(k+2)}$

Non-Universal Degree Distributions

example: $A_k = \begin{cases} \alpha & k = 1 \\ k & k \geq 2 \end{cases}$

substitute into $1 = \sum_{k=1}^{\infty} \prod_{j=1}^{k} \left(1 + \frac{\mu}{A_j}\right)^{-1}$ & split off A_1 term:

$$\mu = A_1 \sum_{k=2}^{\infty} \prod_{j=2}^{k} \left(1 + \frac{\mu}{A_j}\right)^{-1}$$

$$= \alpha \sum_{k=2}^{\infty} \Gamma(2+\mu) \frac{\Gamma(1+k)}{\Gamma(1+\mu+k)}$$

some black magic:

$$\sum_{k=2}^{\infty} \frac{\Gamma(a+k)}{\Gamma(b+k)} = \frac{\Gamma(a+2)}{(b-a-1)\Gamma(b+1)}$$

$$\rightarrow \mu(\mu - 1) = 2\alpha$$

$$\nu = \frac{3 + \sqrt{1 + 8\alpha}}{2} \quad !!$$

full solution: $n_1 = \frac{\mu}{\mu + \alpha}, \quad n_k = \frac{\mu\alpha}{\mu+\alpha} \frac{\Gamma(2+\mu)\,\Gamma(k)}{\Gamma(1+\mu+k)}$

Shifted Linear Attachment $A_k = k + \lambda$

$$
\begin{aligned}
n_k &= \frac{\mu}{A_k} \prod_{j=1}^{k} \left(1 + \frac{\mu}{A_j}\right)^{-1} \\
&= (2+\lambda)\,\frac{\Gamma(3+2\lambda)}{\Gamma(1+\lambda)}\,\frac{\Gamma(k+\lambda)}{\Gamma(k+3+2\lambda)} \\
&\sim k^{-(3+\lambda)} \qquad -1 < \lambda < \infty
\end{aligned}
$$

Role of the Out-Degree m

Master equation:

$$\frac{dN_k}{dt} = \frac{m}{A}\left[(k-1)N_{k-1} - kN_k\right] + \delta_{k,m}$$

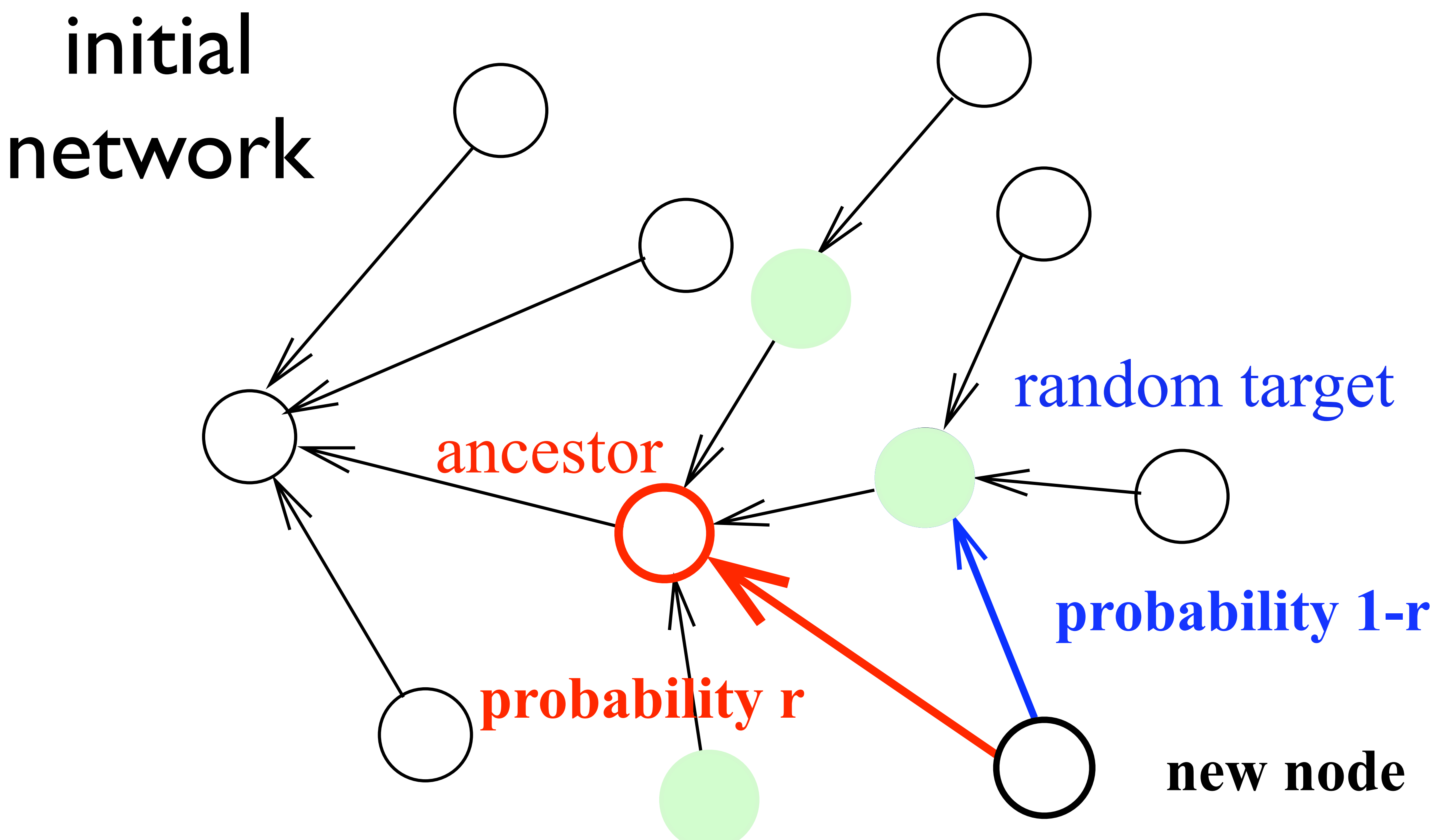
Strictly linear attachment rate:

$$n_k = \frac{2m(m+1)}{k(k+1)(k+2)} \sim k^{-3} \qquad \text{for} \quad k \geq m$$

Shifted linear attachment rate:

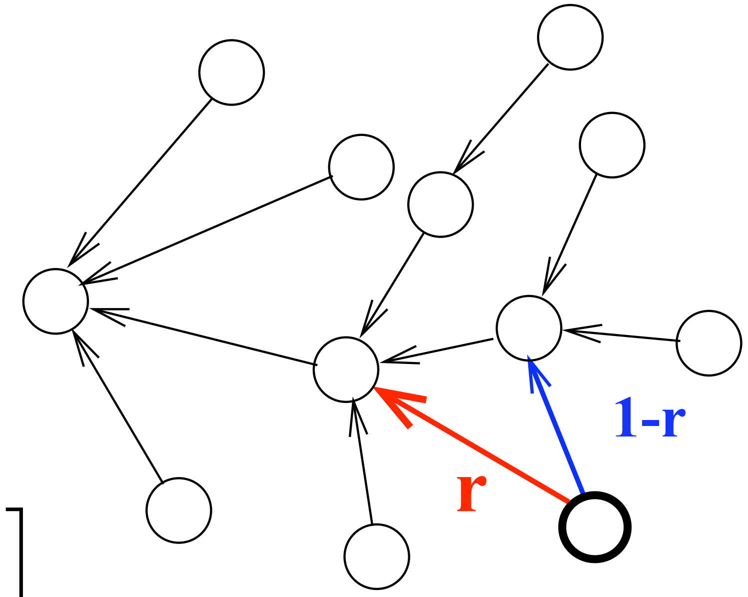
$$n_k \propto \frac{\Gamma(k+\lambda)}{\Gamma(k+3+\lambda+\lambda/m)} \sim k^{-(3+\lambda/m)} \qquad \text{for } k \geq m$$

Shifted Linear Attachment =
Uniform Attachment + Redirection



attachment rate to ancestor node:
∝ number of upstream neighbors

Master Equation:



$$\frac{dN_k}{dN} = \frac{1-r}{M_0}\left[N_{k-1} - N_k\right] + \delta_{k,1} + \frac{r}{M_0}\left[(k-2)N_{k-1} - (k-1)N_k\right]$$

$$= \frac{r}{M_0}\left\{\left[(k-1) + \frac{1}{r} - 2\right]N_{k-1} - \left[k + \frac{1}{r} - 2\right]N_k\right\} + \delta_{k,1}$$

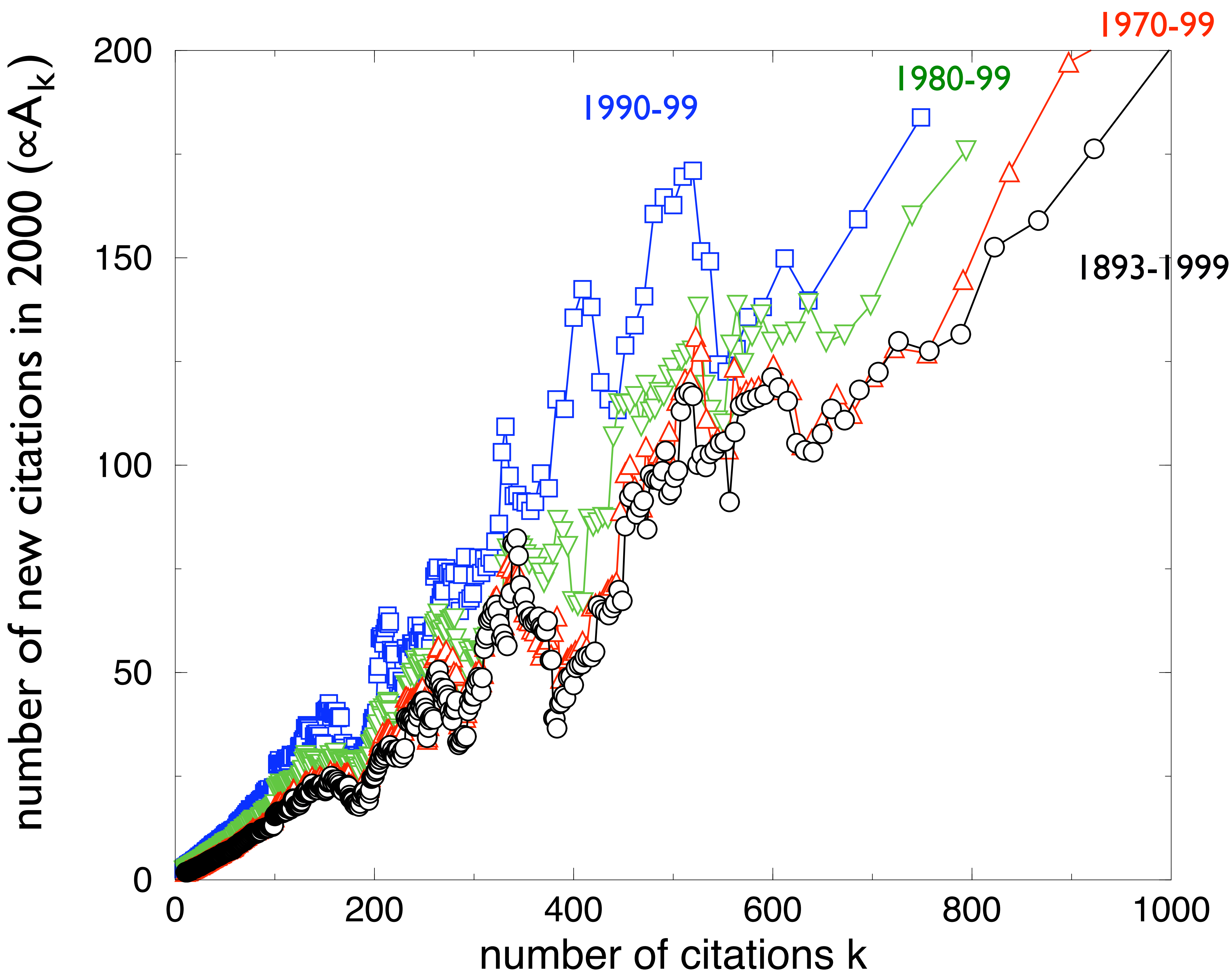
⇨ *shifted* linear attachment rate:

$$A_k = k + \left(\frac{1}{r} - 2\right) \equiv k + \lambda$$

$O(1)$ rule → (shifted) linear preferential attachment!!

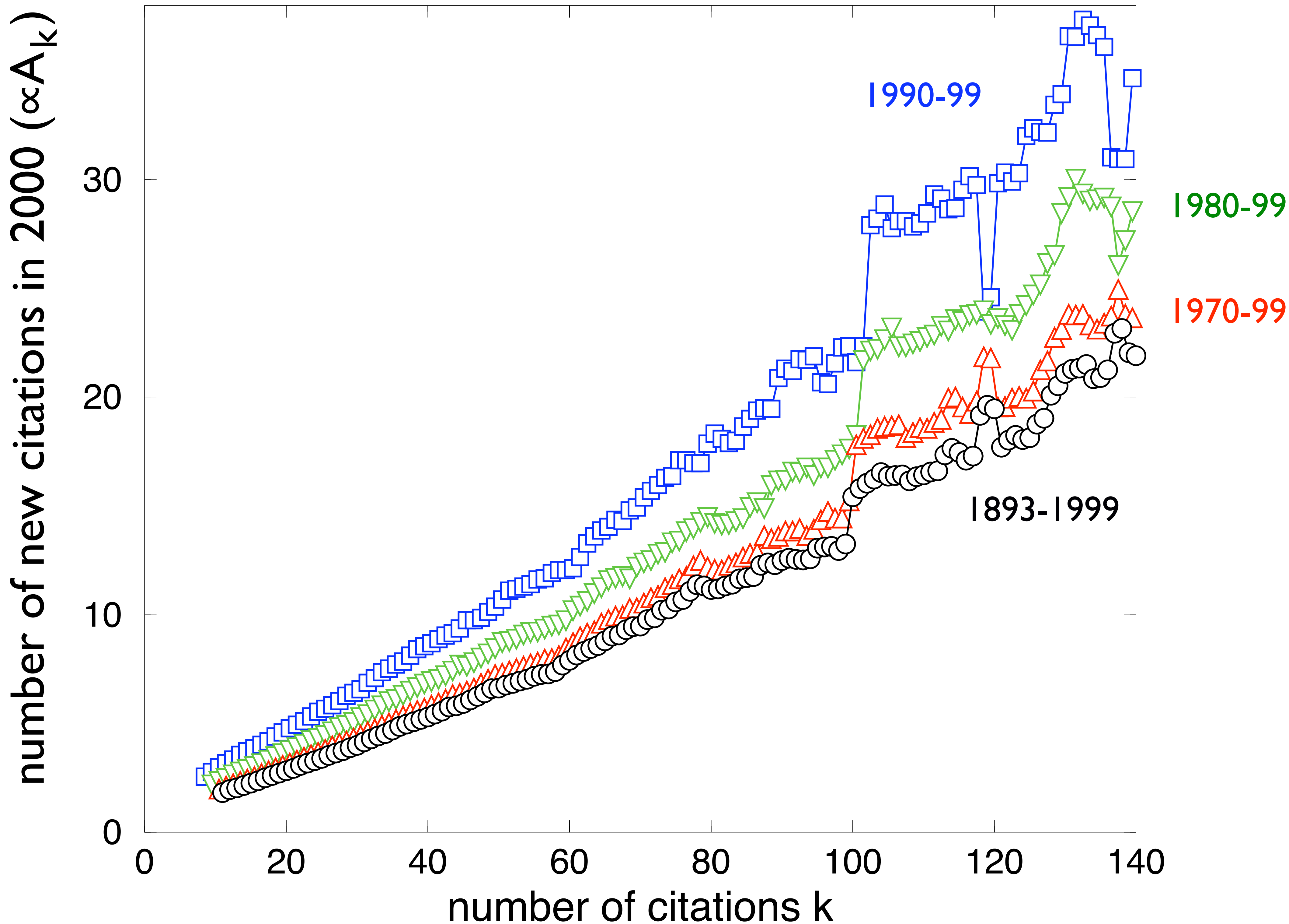
Why Care About *Linear* Preferential Attachment?

Distribution of Citations of Physical Review



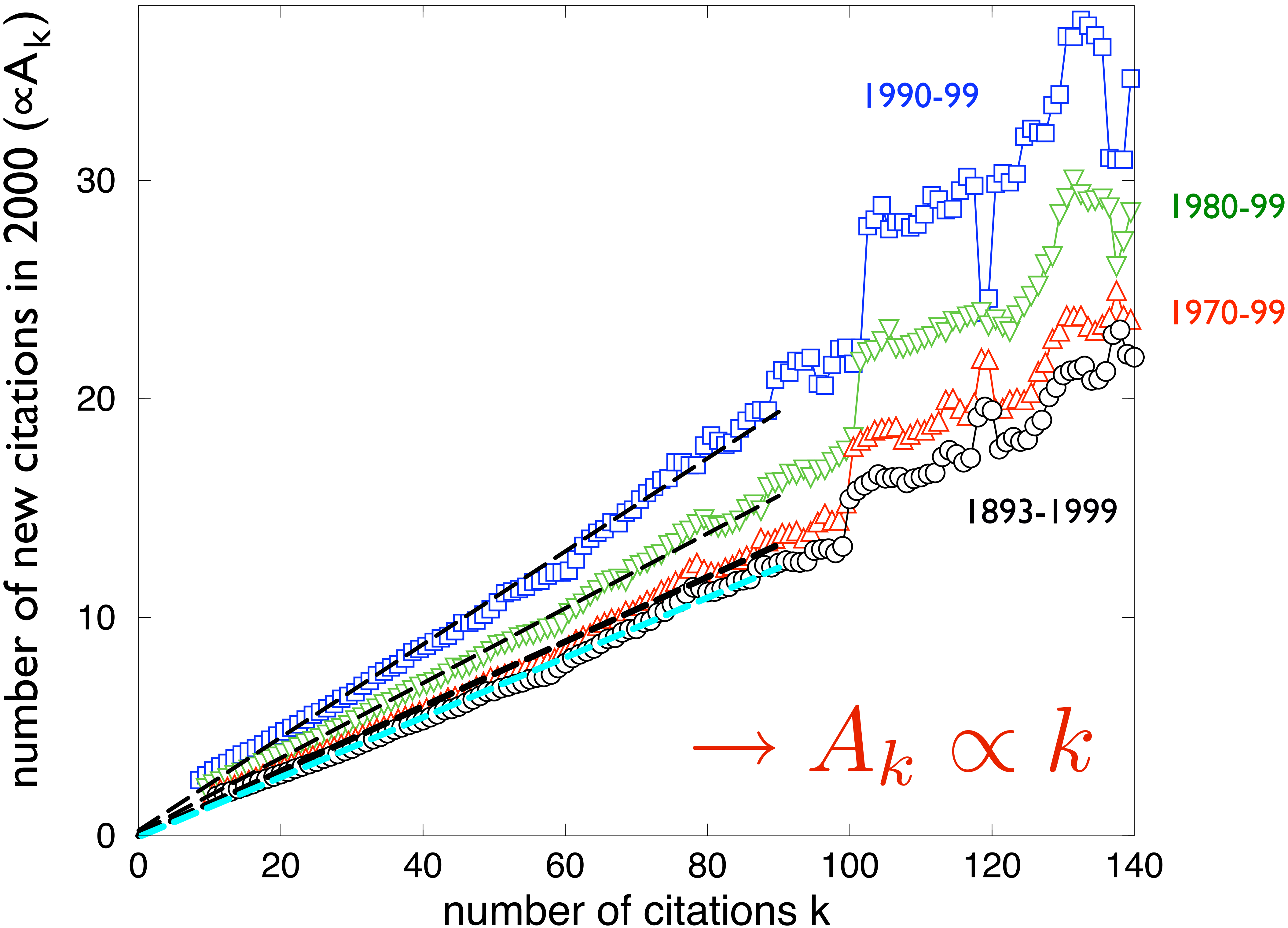
Why Care About *Linear* Preferential Attachment?

Distribution of Citations of Physical Review



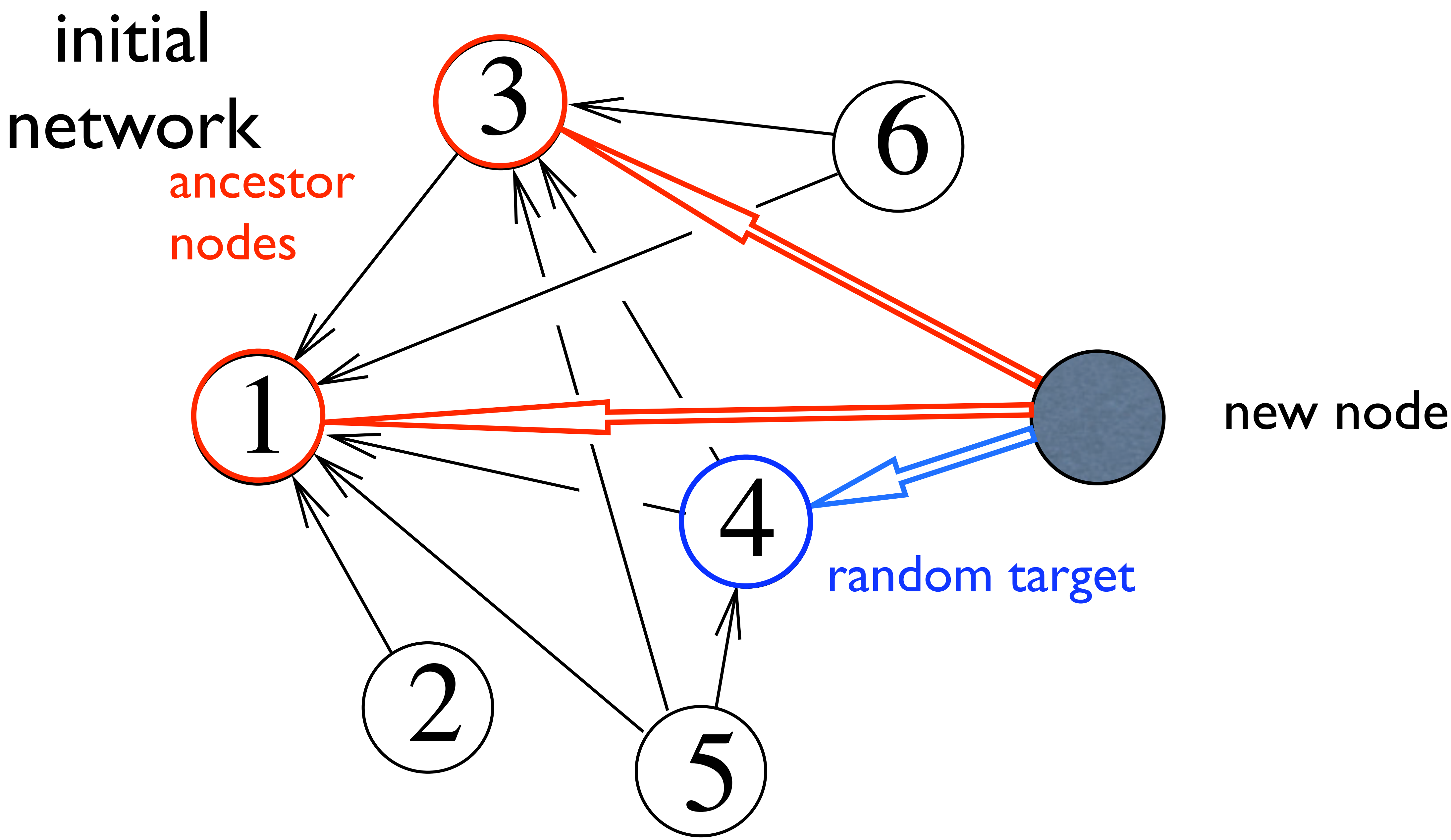
Why Care About *Linear* Preferential Attachment?

Distribution of Citations of Physical Review



Random Attachment + Copying

(a lazy person's approach to references)



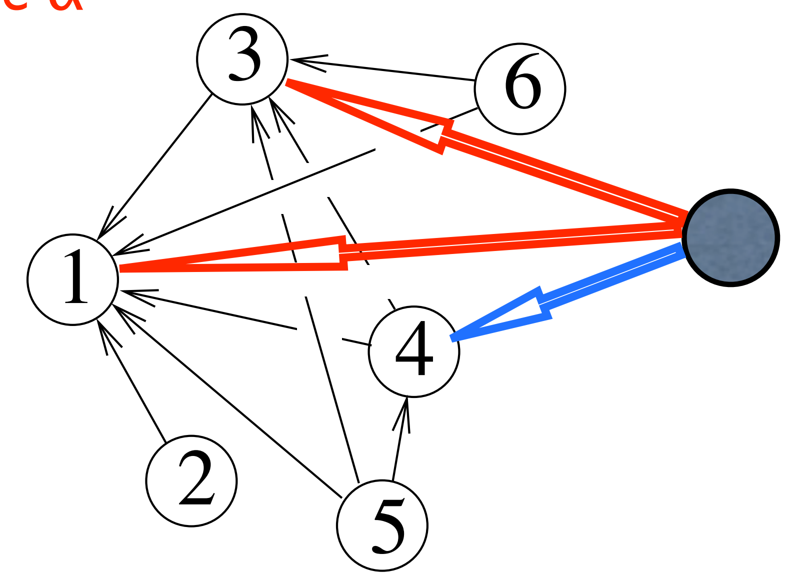
Mean Number of Links L(N)

Evolution equation:

ancestors of node α

$$L(N+1) = L(N) + \frac{1}{N}\sum_{\alpha}(1+j_\alpha)$$

$$= L(N) + 1 + \frac{L(N)}{N}$$

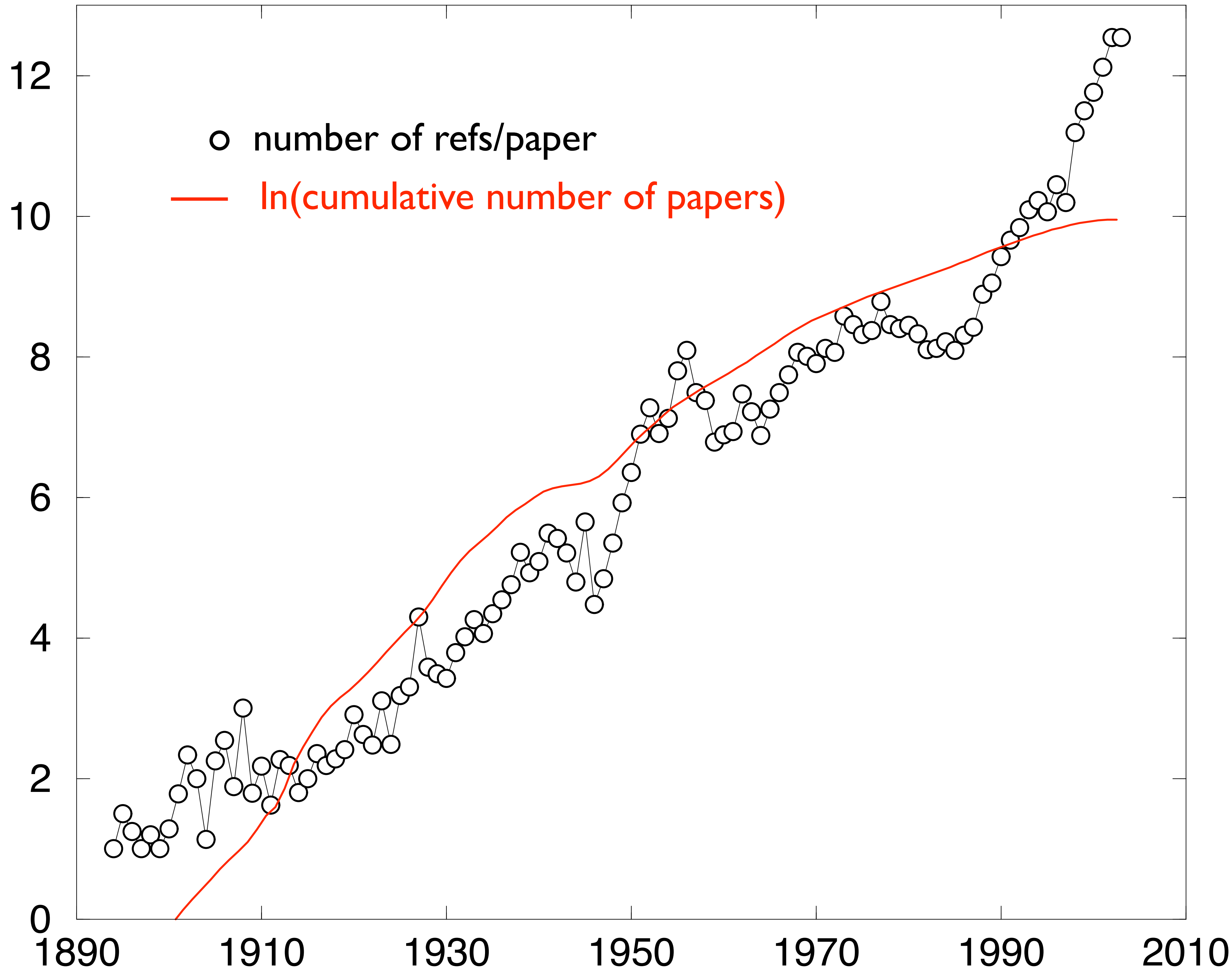


Solution:

$$L(N) = N(H_N - 1)$$

$$= N\ln N - N(1-\gamma) + \frac{1}{2} - \frac{1}{12N} + \dots$$

Comparison with Physical Review Reference Data



Degree Distributions

In-Degree: $P_i(N)$: number of nodes with in-degree i

attach to node of degree i: directly or thru i-1 descendants

attach to node of degree i-1

$$P_i(N+1) - P_i(N) = -\frac{i+1}{N} P_i(N) + \frac{i}{N} P_{i-1}(N) + \delta_{i,0}$$

$$P_i(N) = \frac{N}{(i+1)(i+2)}$$

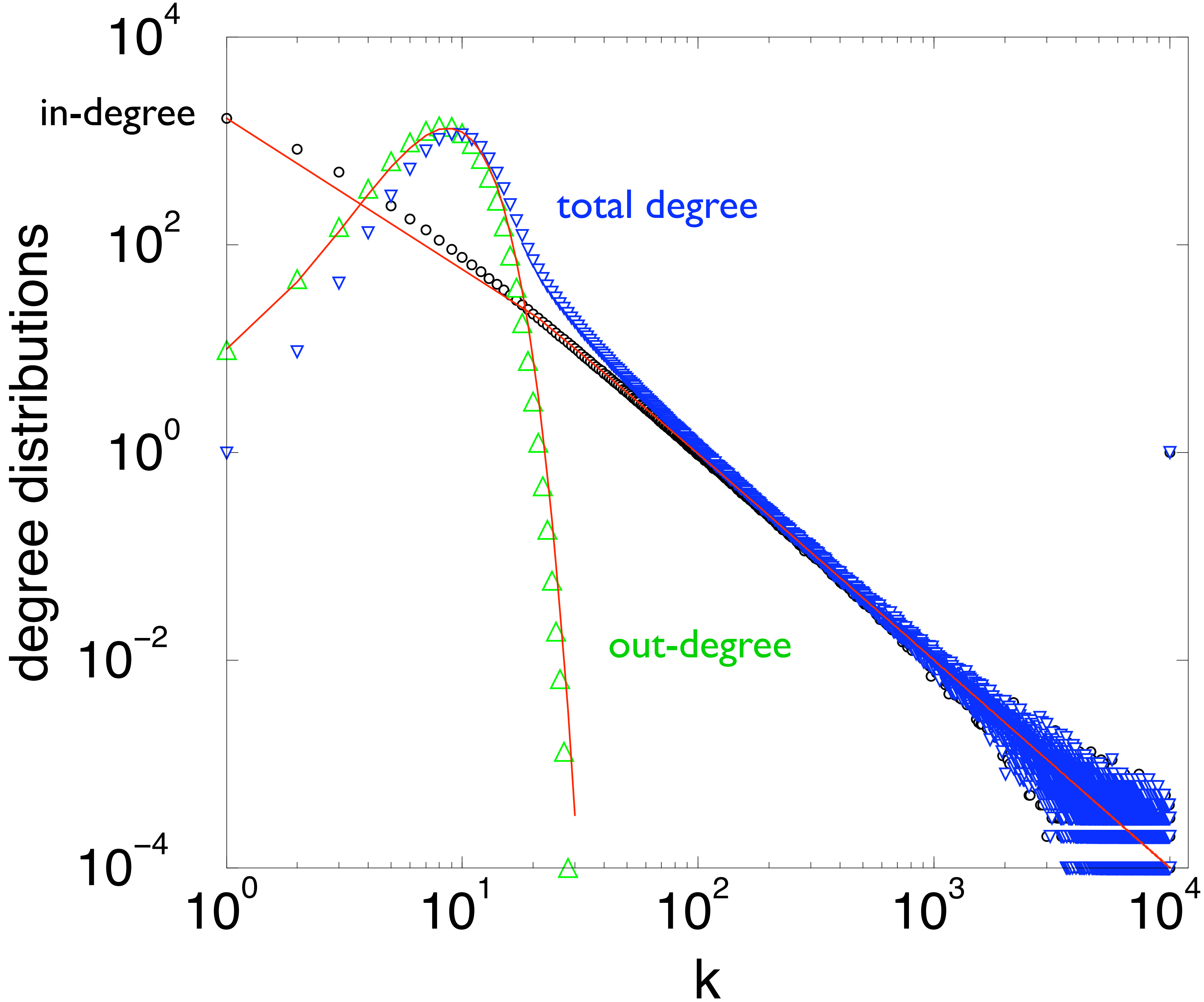
Out-Degree: $Q_j(N)$: number of nodes with out-degree j

$$Q_j(N+1) - Q_j(N) = +\frac{1}{N} Q_{j-1}(N)$$

pick a node of out-degree j-1
→ new node has out-degree i

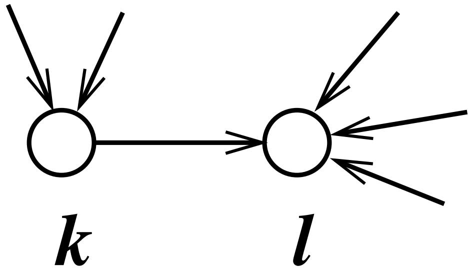
$$Q_j(N) \to \frac{(\ln N)^j}{j!}$$

The Degree Distributions



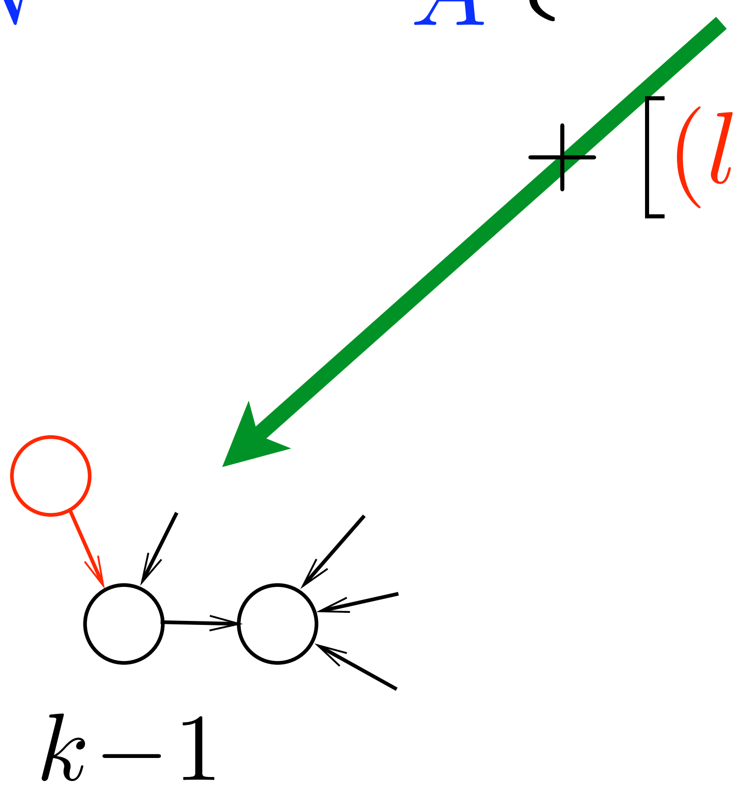
Degree Correlations

$C_{kl}(N) \equiv$ number of nodes of degree k (≥1) that attach to an ancestor of degree l (≥2)



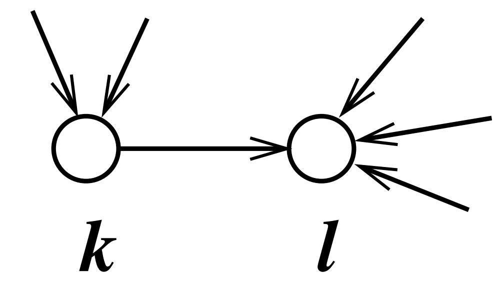
Master equation (for linear rate A_k=k):

$$\frac{dC_{kl}}{dN} = \frac{1}{A}\Big\{\big[(k-1)C_{k-1,l} - kC_{kl}\big] + \big[(l-1)C_{k,l-1} - lC_{kl}\big]\Big\} + (l-1)N_{l-1}\delta_{k1}$$



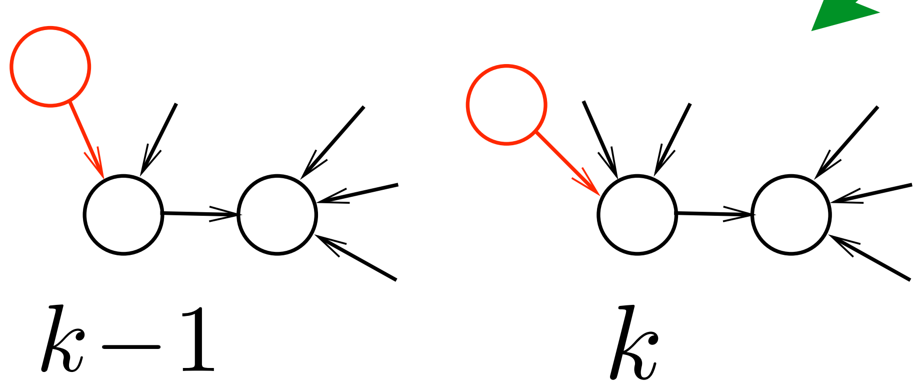
Degree Correlations

$C_{kl}(N) \equiv$ number of nodes of degree k (≥1) that attach to an ancestor of degree l (≥2)



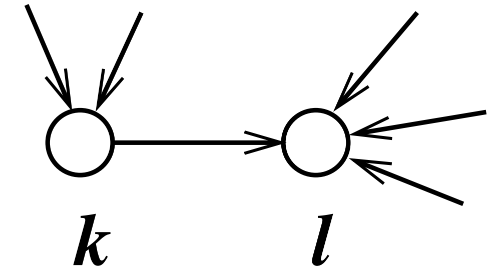
Master equation (for linear rate A_k=k):

$$\frac{dC_{kl}}{dN} = \frac{1}{A}\Big\{\big[(k-1)C_{k-1,l} - kC_{kl}\big] + \big[(l-1)C_{k,l-1} - lC_{kl}\big]\Big\} + (l-1)N_{l-1}\delta_{k1}$$



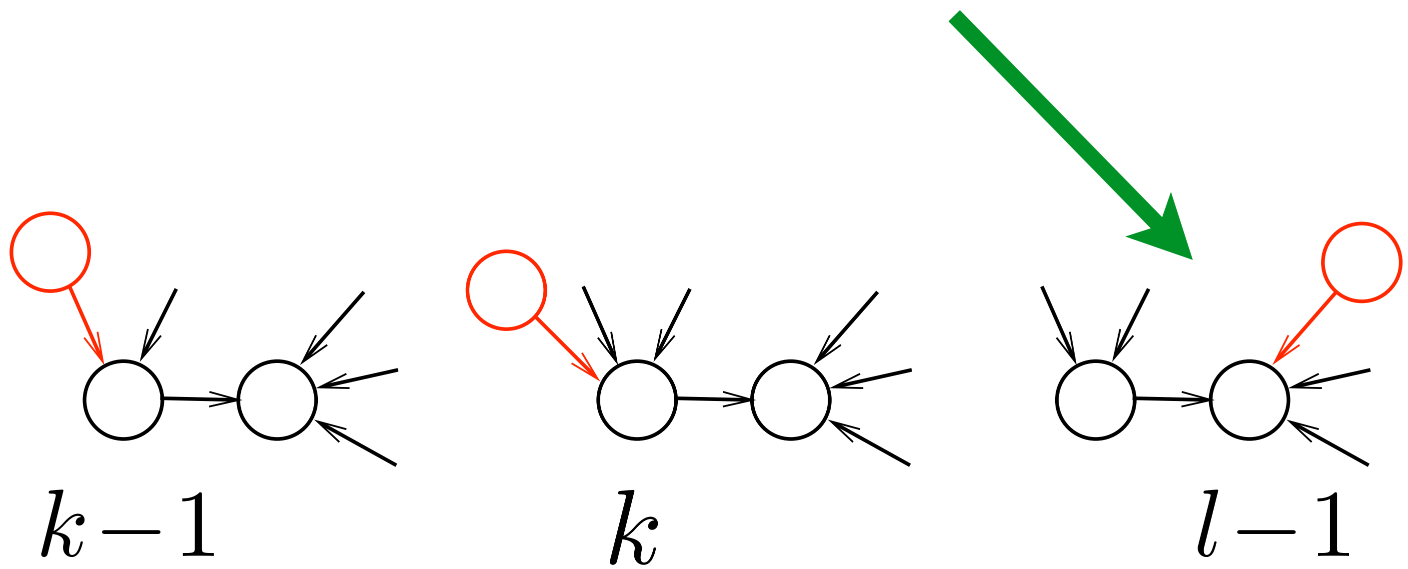
Degree Correlations

$C_{kl}(N) \equiv$ number of nodes of degree k (≥1) that attach to an ancestor of degree l (≥2)



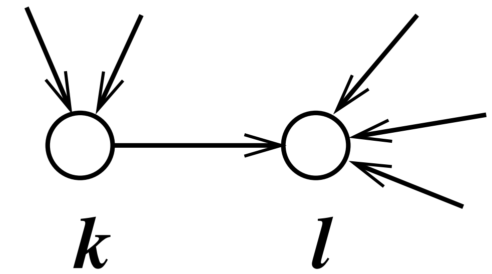
Master equation (for linear rate A_k=k):

$$\frac{dC_{kl}}{dN} = \frac{1}{A}\Big\{\big[(k-1)C_{k-1,l} - kC_{kl}\big] + \big[(l-1)C_{k,l-1} - lC_{kl}\big]\Big\} + (l-1)N_{l-1}\delta_{k1}$$



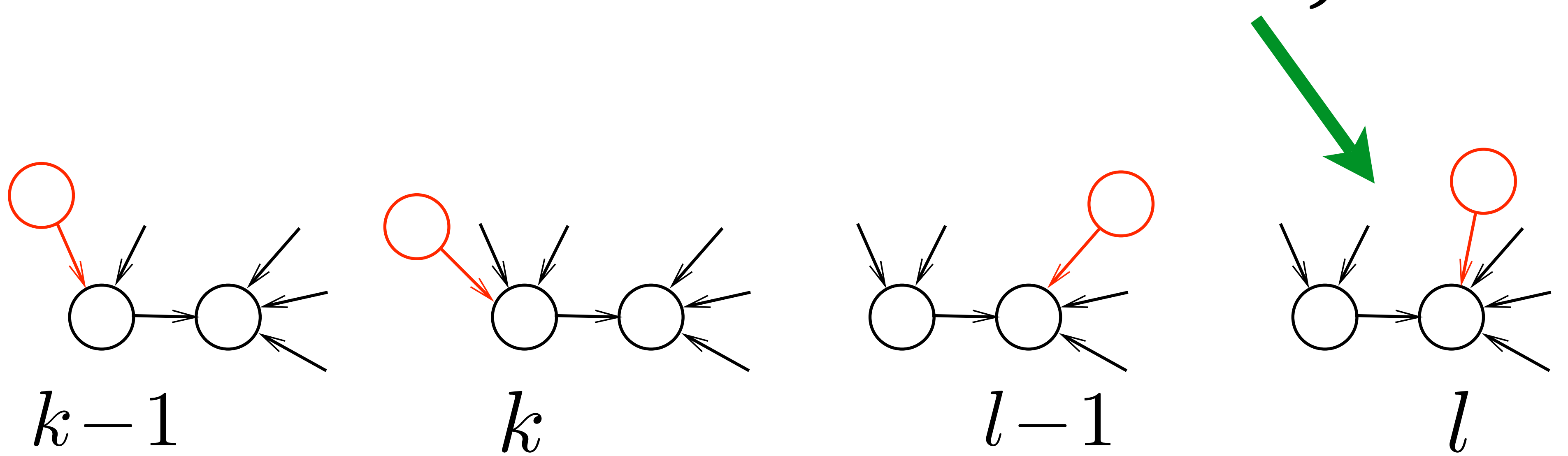
Degree Correlations

$C_{kl}(N) \equiv$ number of nodes of degree k (≥1) that attach to an ancestor of degree l (≥2)



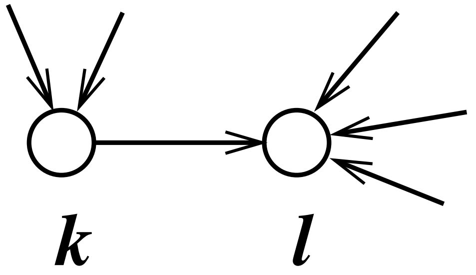
Master equation (for linear rate A_k=k):

$$\frac{dC_{kl}}{dN} = \frac{1}{A}\Big\{\big[(k-1)C_{k-1,l} - kC_{kl}\big] + \big[(l-1)C_{k,l-1} - lC_{kl}\big]\Big\} + (l-1)N_{l-1}\delta_{k1}$$



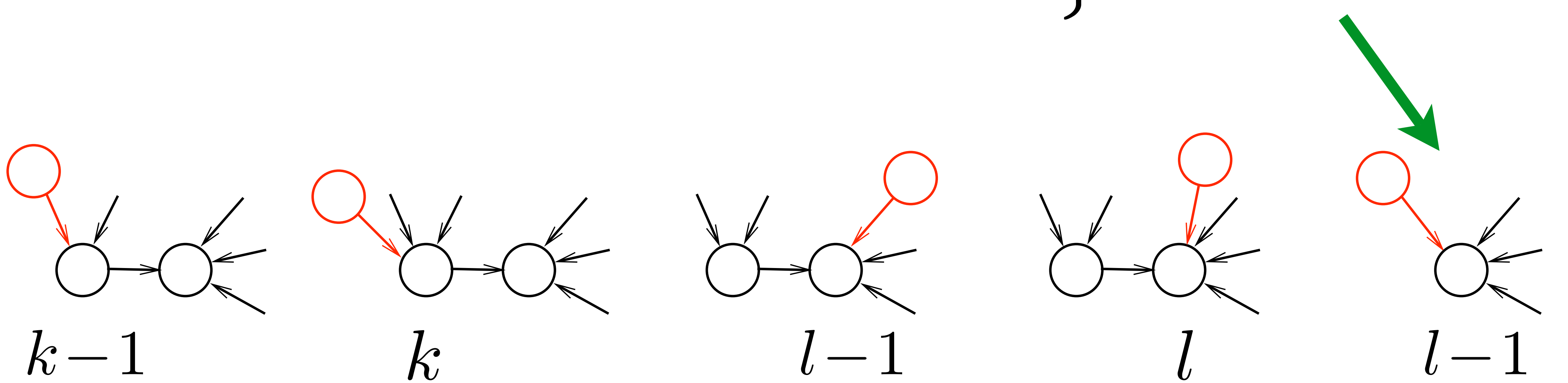
Degree Correlations

$C_{kl}(N) \equiv$ number of nodes of degree k (≥1) that attach to an ancestor of degree l (≥2)



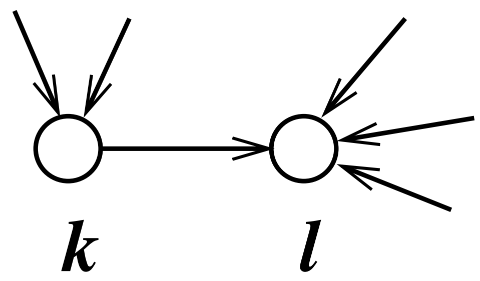
Master equation (for linear rate A_k=k):

$$\frac{dC_{kl}}{dN} = \frac{1}{A}\Big\{\big[(k-1)C_{k-1,l} - kC_{kl}\big] + \big[(l-1)C_{k,l-1} - lC_{kl}\big]\Big\} + (l-1)N_{l-1}\delta_{k1}$$



Degree Correlations

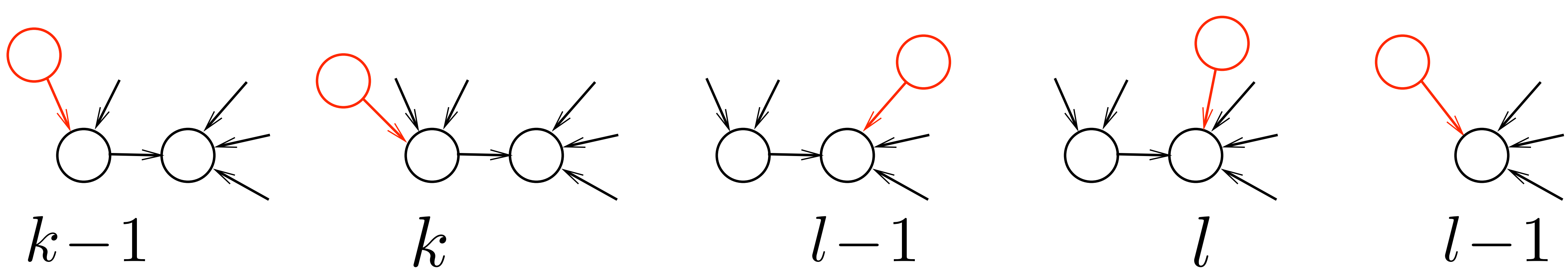
$C_{kl}(N) \equiv$ number of nodes of degree k (≥1) that attach to an ancestor of degree l (≥2)



Master equation (for linear rate A_k=k):

closed equation (no hierarchy)

$$\frac{dC_{kl}}{dN} = \frac{1}{A}\Big\{\big[(k-1)C_{k-1,l} - kC_{kl}\big] + \big[(l-1)C_{k,l-1} - lC_{kl}\big]\Big\} + (l-1)N_{l-1}\delta_{k1}$$



Asymptotic solution for $A_k = k$

For $N \to \infty$, $A \to 2N$, $C_{kl} \to N n_{lk}$
the master equation becomes:

$$(k+l+2)n_{kl} = (k-1)n_{k-1,l} + (l-1)n_{k,l-1} + (l-1)n_{l-1}\,\delta_{k,1}$$

Small trick: let $n_{kl} = \dfrac{\Gamma(k)\,\Gamma(l)}{\Gamma(k+l+3)}\, m_{kl}$

Reduces the master equation to:

$$m_{kl} = m_{k-1,l} + m_{k,l-1} + 4(l+2)\delta_{k1}$$

Define the generating function: $\mathcal{M}(x,y) = \sum_{k=1}^{\infty}\sum_{l=2}^{\infty} m_{kl}\, x^k y^l$

Generating function solution: $\mathcal{M}(x,y) = \dfrac{4xy^2(4-3y)}{(1-x-y)(1-y)^2}$

Master equation solution:

$$n_{kl} = \frac{4(l-1)}{k(k+1)(k+l)(k+l+1)(k+l+2)} + \frac{12(l-1)}{k(k+l-1)(k+l)(k+l+1)(k+l+2)}$$

Asymptotics:
$(k, l \to \infty)$

$$n_{kl} \sim \begin{cases} 16\dfrac{l}{k^5} & l \ll k \\ \dfrac{4}{k^4}\dfrac{4y(y+4)}{(1+y)^4} & l \sim k \\ \dfrac{4}{k^2 l^2} & l \gg k \end{cases}$$

$$n_{kl} \neq n_k n_l \propto (kl)^{-3}$$

Age-Degree Distribution

Basic question: *when* do connections occur?

$N_k(N,a) \equiv$ number of nodes of degree k and age a in a network of N nodes

master equation:

$$\left(\frac{\partial}{\partial N}+\frac{\partial}{\partial a}\right)N_k = \frac{A_{k-1}N_{k-1} - A_k N_k}{A} + \delta_{k,1}\delta_{a,0}$$

node aging

wave equation scaling: $N_k(N,a) = f_k(x);\qquad x = 1-\frac{a}{N}$

$$\longrightarrow\quad -2x\frac{df_k}{dx} = (k-1)f_{k-1} - kf_k$$

(for A_k=k):

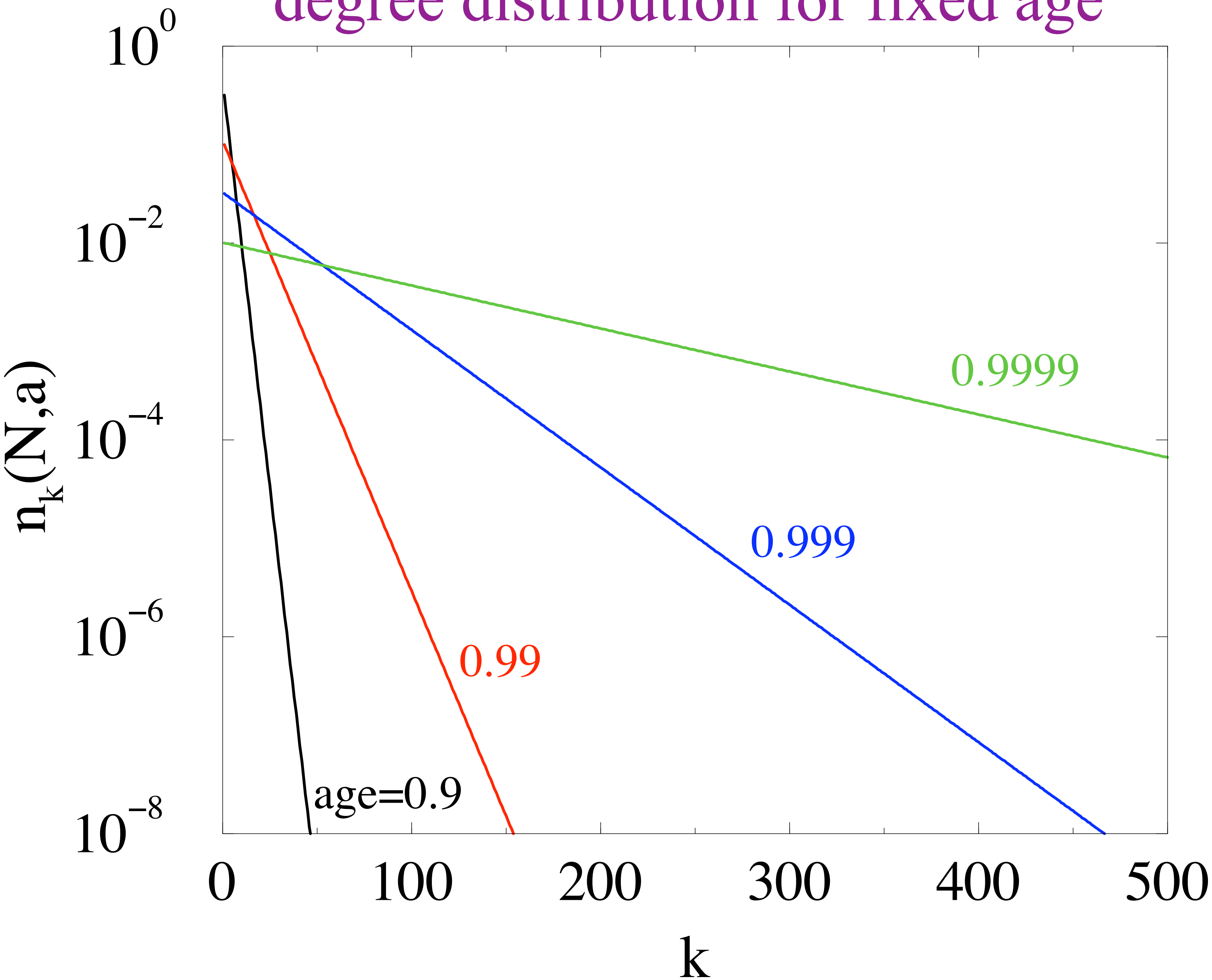
with the ansatz $f_k = \Phi\varphi^{k-1}$ the solution is:

$$\varphi(x) = 1-\sqrt{x}\qquad \Phi(x) = \sqrt{x}$$

Age-Degree Distribution

solution (for A_k=k): $$N_k(N,a) = \sqrt{1-\frac{a}{N}}\left[1-\sqrt{1-\frac{a}{N}}\right]^{k-1}$$
$$\sim (k^*)^{-1}\, e^{-k/k^*}$$

degree distribution for fixed age



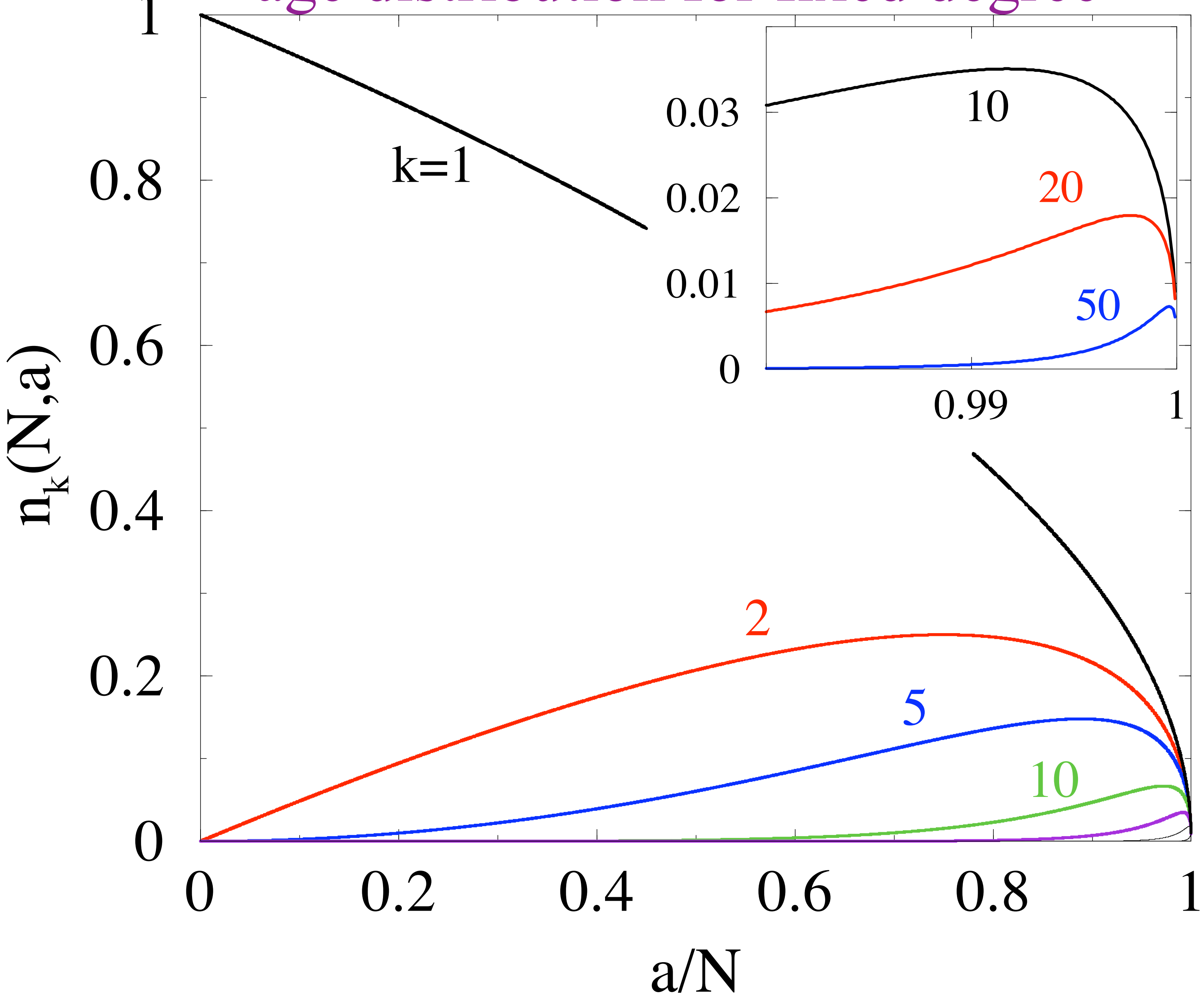
$$k^* \propto \left(1-\frac{a}{N}\right)^{-1/2}$$
$$\to \begin{cases} 1 & a \to 0 \\ \sqrt{N} & a \to N \end{cases}$$

Age-Degree Distribution

solution (for $A_k=k$): $$N_k(N,a) = \sqrt{1-\frac{a}{N}}\left[1-\sqrt{1-\frac{a}{N}}\right]^{k-1}$$

$$\sim (k^*)^{-1} e^{-k/k^*}$$

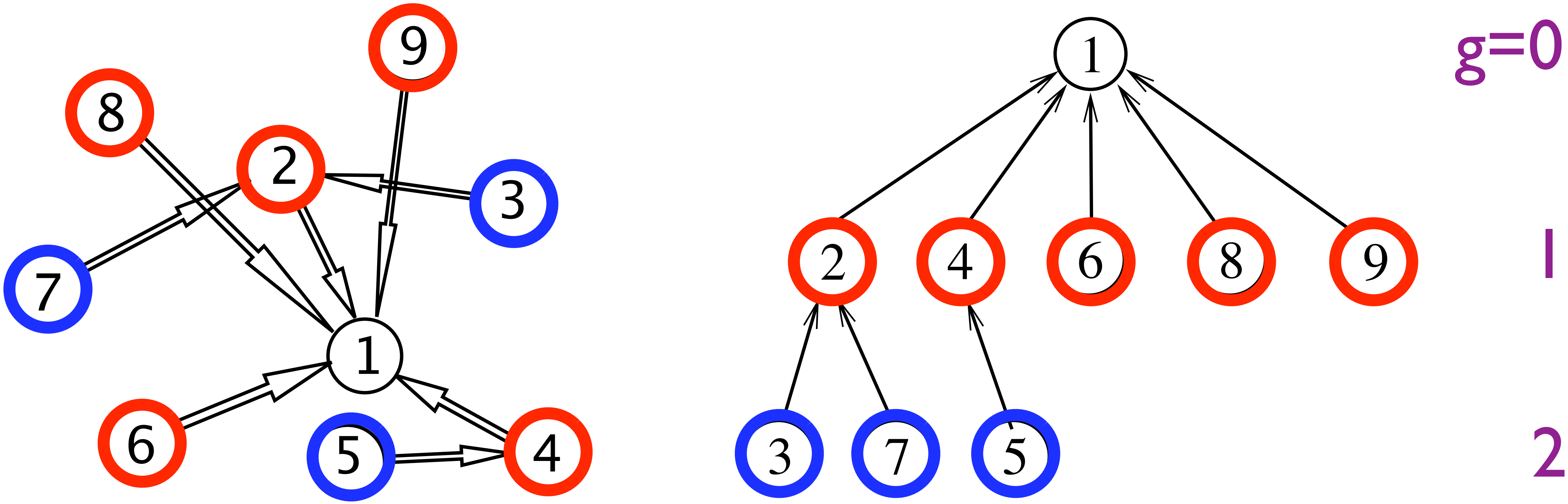
age distribution for fixed degree



$$\langle a_k \rangle = N_k^{-1} \int_0^N a\, N_k(N,a)\, da$$

$$\sim 1 - \frac{4}{k^2}$$

Genealogy (for uniform attachment)



Size of generation g ≡ L_g

Uniform attachment rate equation:

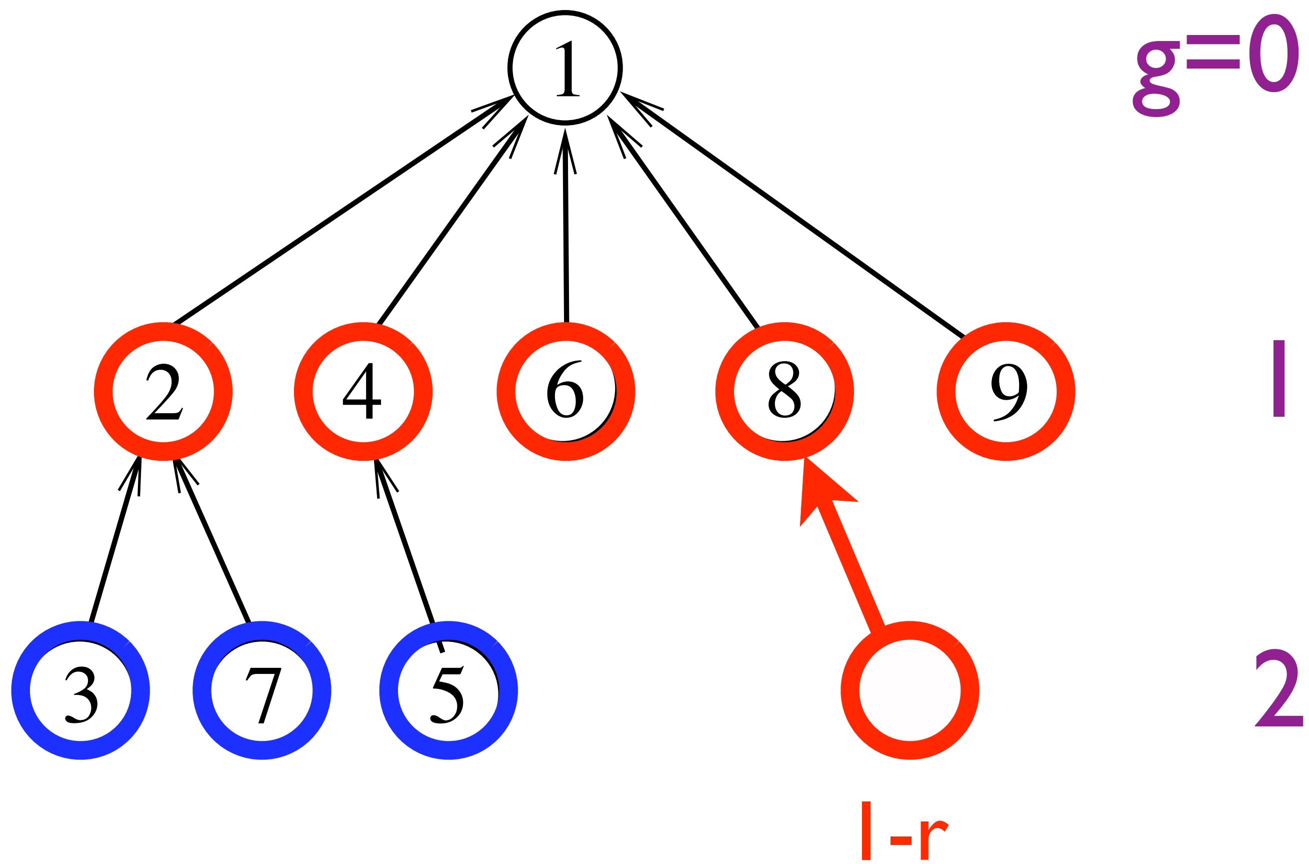
$$\frac{dL_g}{dN} = \frac{L_{g-1}}{N}$$

Solution: $$L_g = \frac{(\ln N)^g}{g!}$$

$L_g = 1$ defines last generation $\longrightarrow g_{\max} \sim e \ln N$

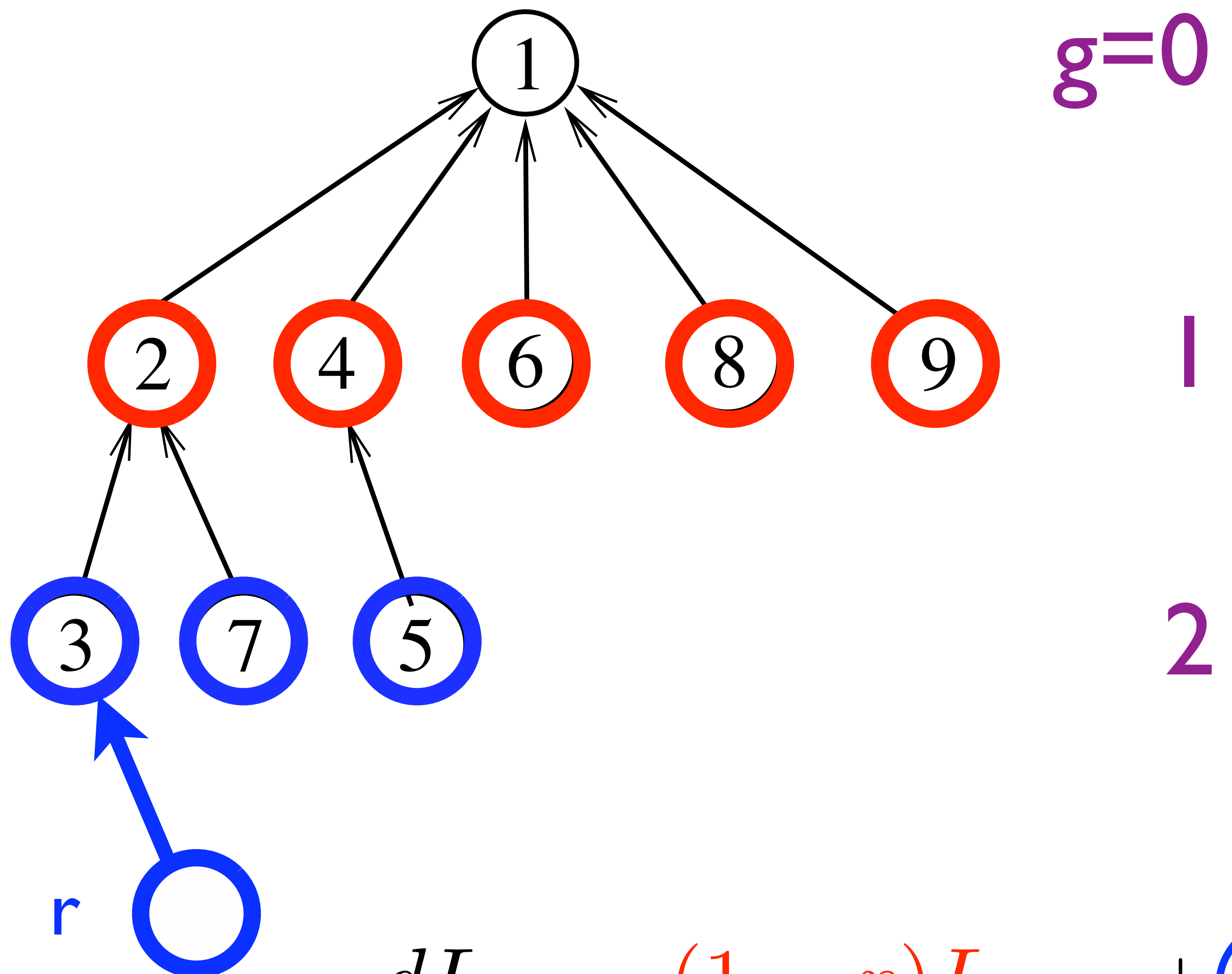
$$\text{diameter} \sim 2e \ln N$$

Genealogy (for A_k=k) *use redirection!*



Rate equation: $$\frac{dL_g}{dN} = \frac{(1-r)L_{g-1} + rL_g}{N}$$

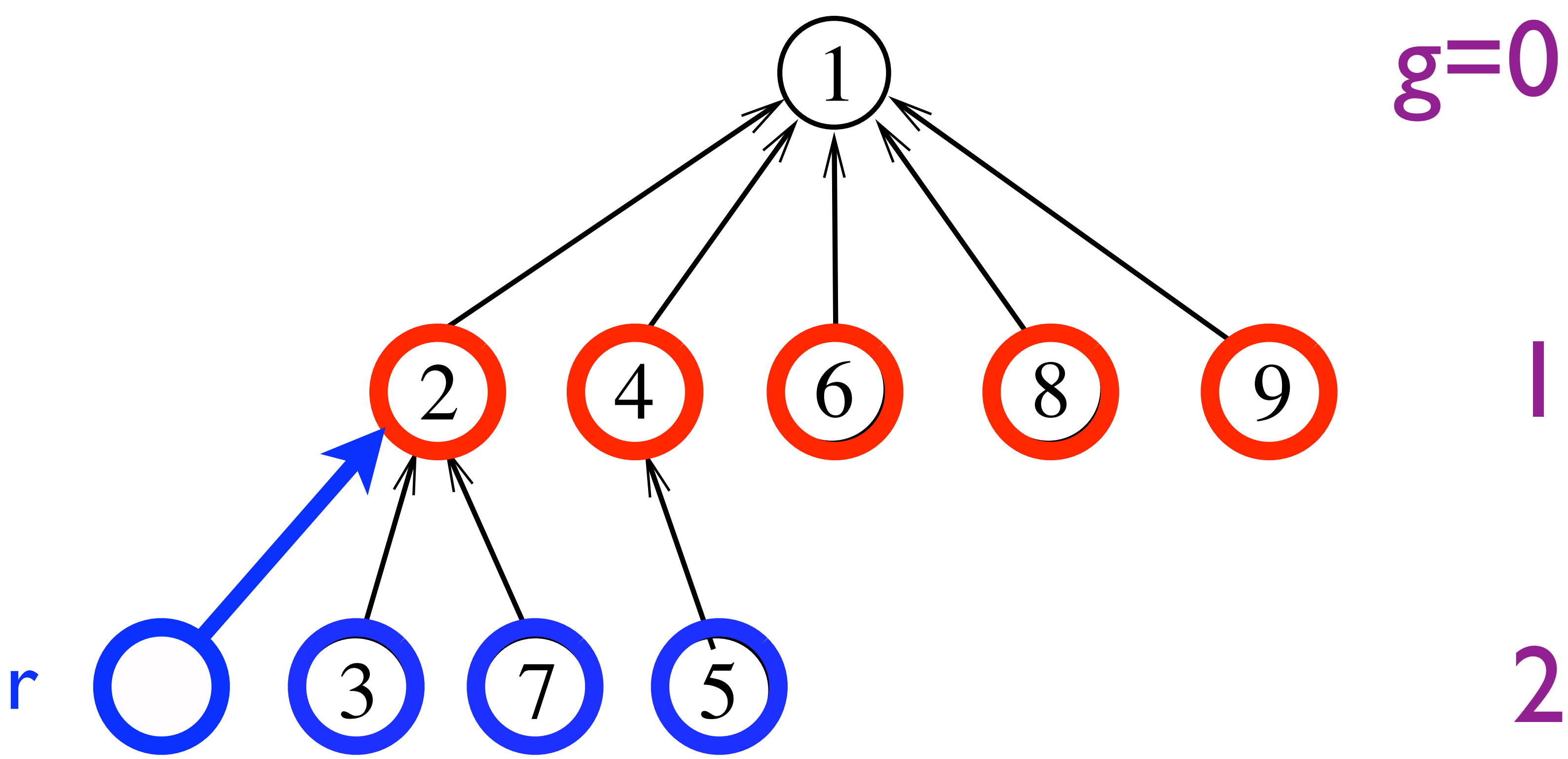
Genealogy (for A_k=k) *use redirection!*



Rate equation:

$$\frac{dL_g}{dN} = \frac{(1-r)L_{g-1} + rL_g}{N}$$

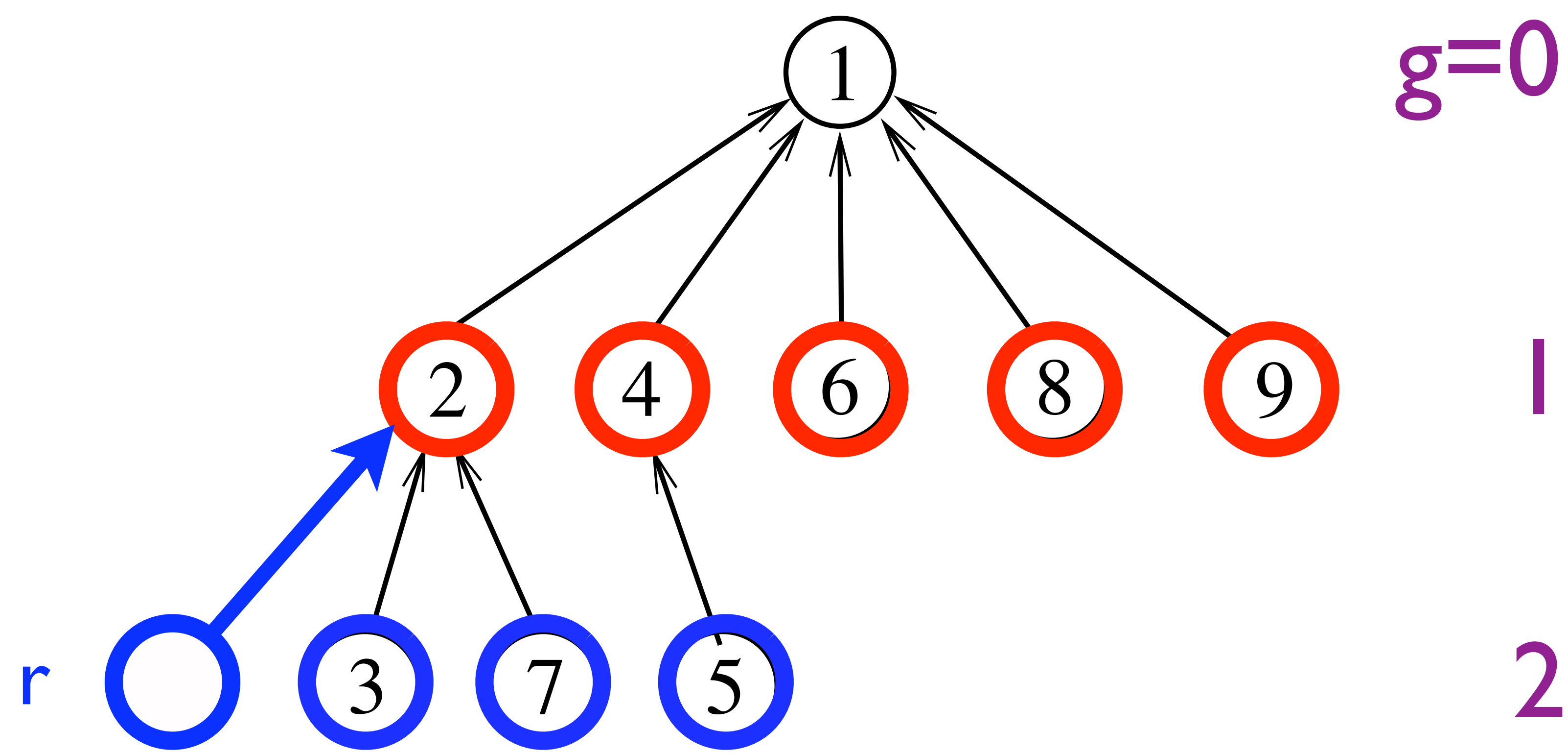
Genealogy (for A_k=k) *use redirection!*



Rate equation: $$\frac{dL_g}{dN} = \frac{(1-r)L_{g-1} + rL_g}{N}$$

For A_k=k, take r=1/2

Genealogy (for A_k=k) *use redirection!*



Rate equation: $$\frac{dL_g}{dN} = \frac{(1-r)L_{g-1} + rL_g}{N}$$

For A_k=k, take r=1/2 $\longrightarrow g_{\max} \sim b \ln N$

morally similar solution as for r=0

Intrinsic Node Fitness

Bianconi & Barabasi (2000)

- each node has fitness η chosen from $p_0(\eta)$

Master equation:

$$\frac{dN_k(\eta)}{dN} = \frac{A_{k-1}(\eta)N_{k-1}(\eta) - A_k(\eta)N_k(\eta)}{A} + p_0(\eta)\delta_{k1}$$

$$A = \int d\eta \sum_k A_k(\eta)N_k(\eta)$$

Solution:

$$n_k(\eta) = p_0(\eta)\frac{\mu}{A_k(\eta)}\prod_{j=1}^{k}\left(1 + \frac{\mu}{A_j(\eta)}\right)^{-1}$$

For $A_k(\eta) = \eta k$:

$$n_k(\eta) = \frac{\mu\, p_0(\eta)}{\eta}\frac{\Gamma(k)\,\Gamma\left(1+\frac{\mu}{\eta}\right)}{\Gamma\left(k+1+\frac{\mu}{\eta}\right)} \to k^{-(1+\mu/\eta)}$$

Determine μ:

$$A = \mu N = \int d\eta \sum_{k\geq 1} A_k(\eta)\,N_k(\eta) \quad \to 1 = \int d\eta\, p_0(\eta)\left(\frac{\mu}{\eta} - 1\right)^{-1}$$

arbitrary charm: BE condensation

bounded charm: $n_k \sim k^{-(1+\mu/\eta_{\rm max})}(\ln k)^{\omega}$

Evolving (Preferential Attachment) Networks

Sidney Redner, Boston University, physics.bu.edu/~redner
collaborators: P. Krapivsky, F. Leyvraz

E(uropean) N(etwork) on RA(ndom) GE(ometry), June 2008

Outline:

- master equation: degree distributions
- ubiquitous linear preferential attachment:
  - *redirection & copying*
- a deeper look: correlations, genealogy, fitness
- do the rich really get richer?
- web growth model
- finiteness & fluctuations

Master Equation: $\frac{dN_k}{dN} = \frac{A_{k-1}N_{k-1}}{A} - \frac{A_k N_k}{A} + \delta_{k,1}$

Formal solution: $n_k = \frac{\mu}{A_k} \prod_{j=1}^{k} \left(1 + \frac{\mu}{A_j}\right)^{-1}$

For $A_k \sim k^\gamma$:

$$n_k \sim \begin{cases} k^{-\gamma}\, e^{-[\mu k^{1-\gamma}/(1-\gamma)]} & 0 \leq k < 1 \\ k^{-\nu},\ \nu > 2 & \gamma = 1 \\ \text{“best seller”} & 1 < \gamma \leq 2 \\ \text{“bible”} & \gamma > 2 \end{cases}$$

Redirection ⇔ *shifted* linear attachment → non-universal ν

Copying → dense network: $\langle k \rangle \propto \ln N$ $n_k = \frac{4}{k(k+1)(k+2)}$, $\langle k \rangle = 2$

Preferential attachment → correlations: $n_{kl} \neq n_k n_l$

Genealogy: older nodes → larger degree; diameter ∝ ln N

Evolving (Preferential Attachment) Networks

Sidney Redner, Boston University, physics.bu.edu/~redner
collaborators: P. Krapivsky, F. Leyvraz

E(uropean) N(etwork) on RA(ndom) GE(ometry), June 2008

Outline:

- master equation: degree distributions
- ubiquitous linear preferential attachment:
  - *redirection & copying*
- a deeper look: correlations, genealogy, fitness
- do the rich really get richer?
- web growth model
- finiteness & fluctuations

Do The Rich Get Richer?

P(k,N)≡ prob. 1st node has degree k in a network of N links

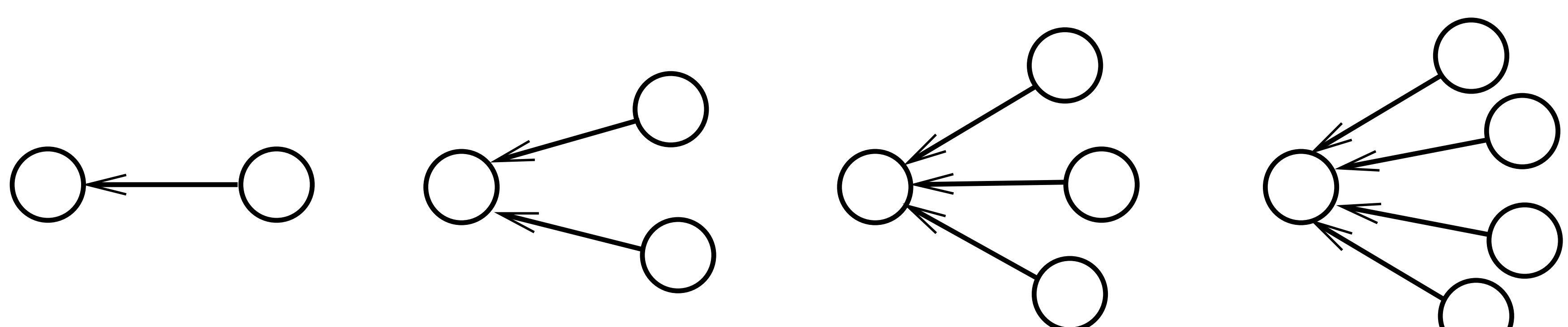
For A_k=k:

attach to 1st node — attach to later node

$$P(k,N+1) = \frac{k-1}{2N}\,P(k-1,N) + \frac{2N-k}{2N}\,P(k,N)$$

Mean degree of 1st node: $\langle k \rangle_{N+1} = \langle k \rangle_N \left(1 + \frac{1}{2N}\right)$ *multiplicative*

Solution: $\langle k \rangle_N = \Lambda \frac{\Gamma\left(N+\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma(N)} \sim \frac{\Lambda}{\sqrt{\pi}} N^{1/2}$



Λ=2 Λ=8/3 Λ=16/5 Λ=128/35

initial "wealth" matters!

Do The Rich Get Richer?

For A_k=1:

attach to 1st node — attach to later node

$$P(k, N+1) = \frac{1}{N+1}\, P(k-1, N) + \frac{N}{N+1}\, P(k, N)$$

Mean degree of 1st node: $\langle k \rangle_{N+1} = \langle k \rangle_N + \dfrac{1}{N+1}$ *additive*

Solution: $\langle k \rangle_N = H_N + \text{const.} \sim \ln N + \text{const.}$

initial "wealth" irrelevant!

A more subtle feature: *leadership*

Does the first node maintain its lead?

For A_k=k:

Largest degree: $N \sum_{k=k_{max}}^{\infty} 4k^{-3} \approx 1 \;\rightarrow\; k_{max} \sim \sqrt{2N}$

1st node degree: $\langle k \rangle_N \sim \Lambda \sqrt{N/\pi}$

can be a winner

For A_k=1:

Largest degree: $N \sum_{k=k_{max}}^{\infty} 2^{-k} = 1 \;\rightarrow\; k_{max} \sim 1 + \frac{\ln N}{\ln 2}$

1st node degree: $\langle k \rangle_N \sim \ln N + \text{const.}$

likely loser

For A_k=1, how does the 1st node become a loser?

Master equation for degree of 1st node:

$$P(k, N+1) = \frac{1}{N+1}\,P(k-1,N) + \frac{N}{N+1}\,P(k,N)$$

Continuum limit & solution:

$$\left(\frac{\partial}{\partial \ln N} + \frac{\partial}{\partial k}\right) P = \frac{1}{2}\,\frac{\partial^2 P}{\partial k^2} \qquad\qquad P(k,N) = \frac{1}{\sqrt{2\pi \ln N}}\,e^{-\frac{(k-\ln N)^2}{2\ln N}}$$

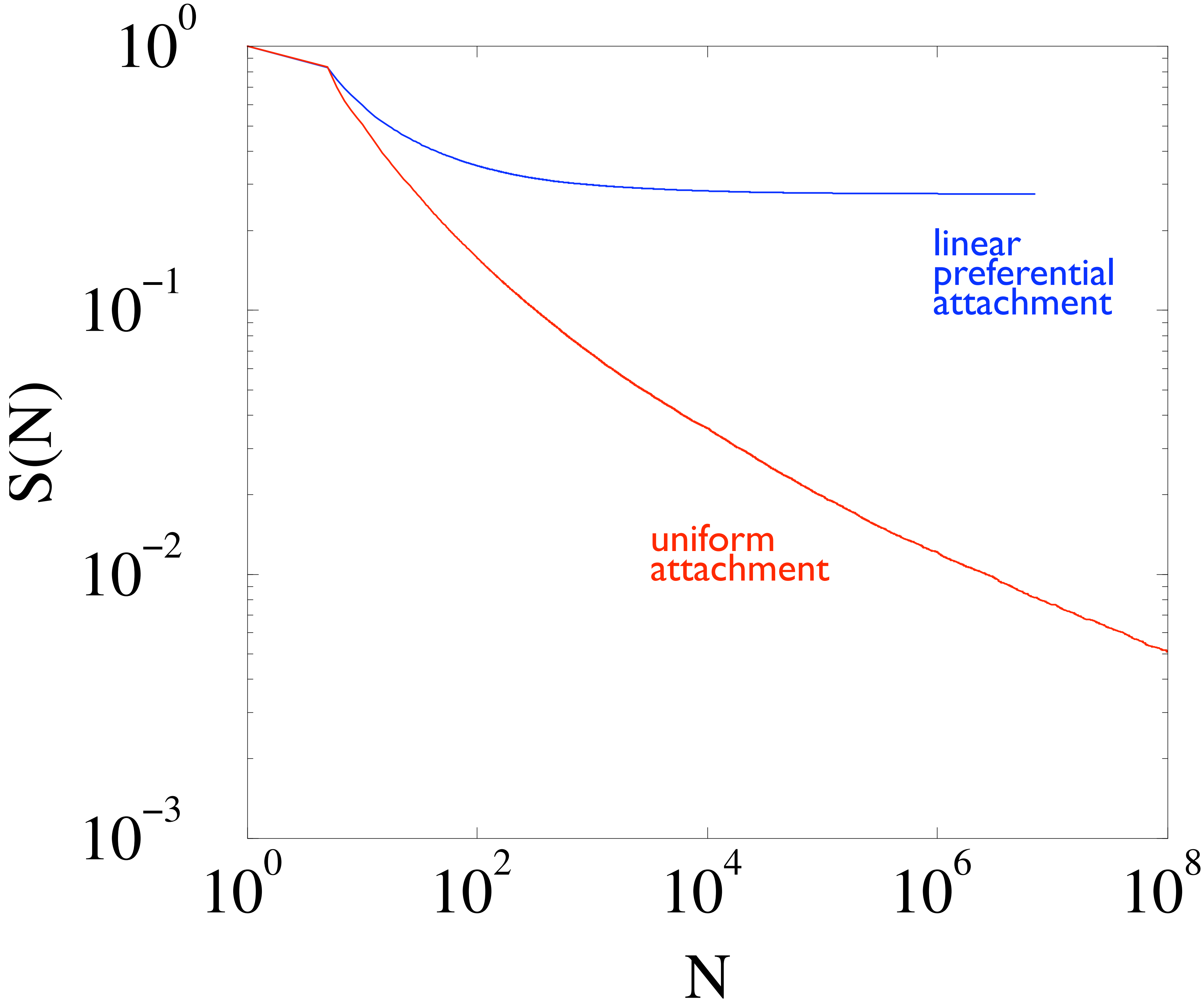
Probability that first node maintains lead:

$$S(N) \quad \sim \quad \int_{\ln N/\ln 2}^{\infty} P(k,N)\,dk$$

$$\propto \quad N^{-\phi}\,(\ln N)^{-1/2}$$

$$\text{with } \phi = 0.086071\ldots \qquad [0.097989\ldots]$$

Lead Probability of the Initial Node

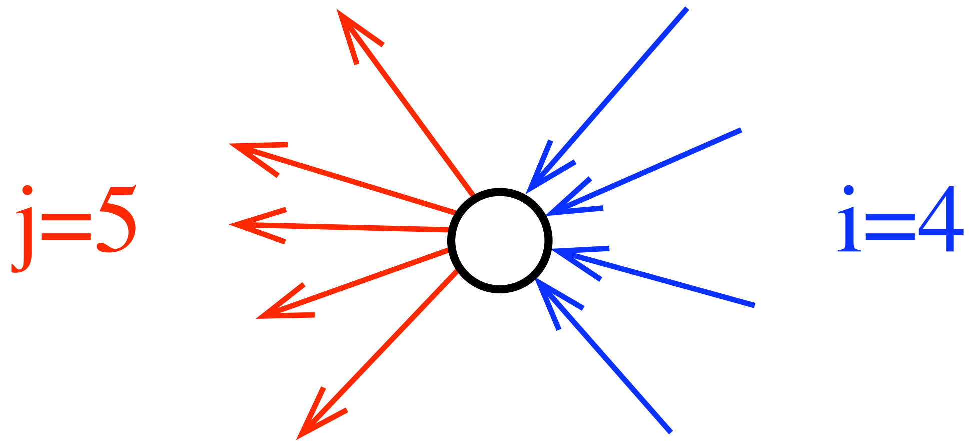


Web-Based Growth Models

Dorogovtsev & Mendes (2000), KRR (2001)

What new features should be included to describe the growth of the www?

(i) link directionality

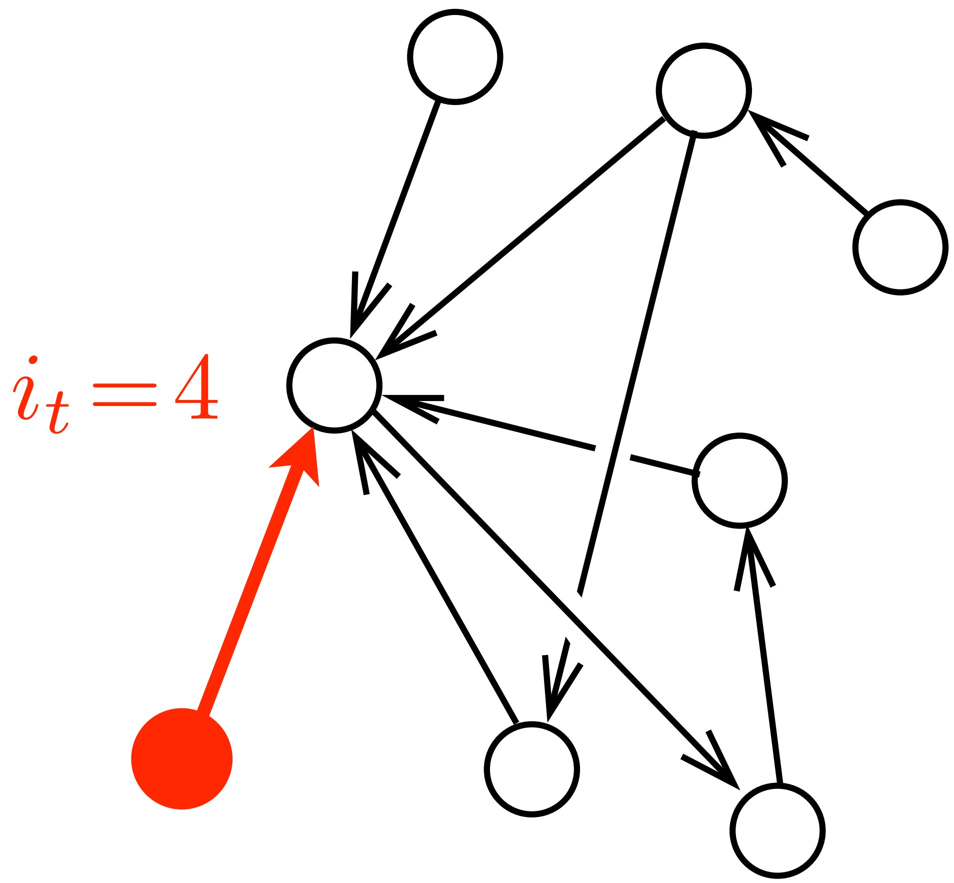


$N_{ij}(N) \equiv$ average number of nodes with in-degree i and out-degree j

(ii) links & nodes should not necessarily be coupled

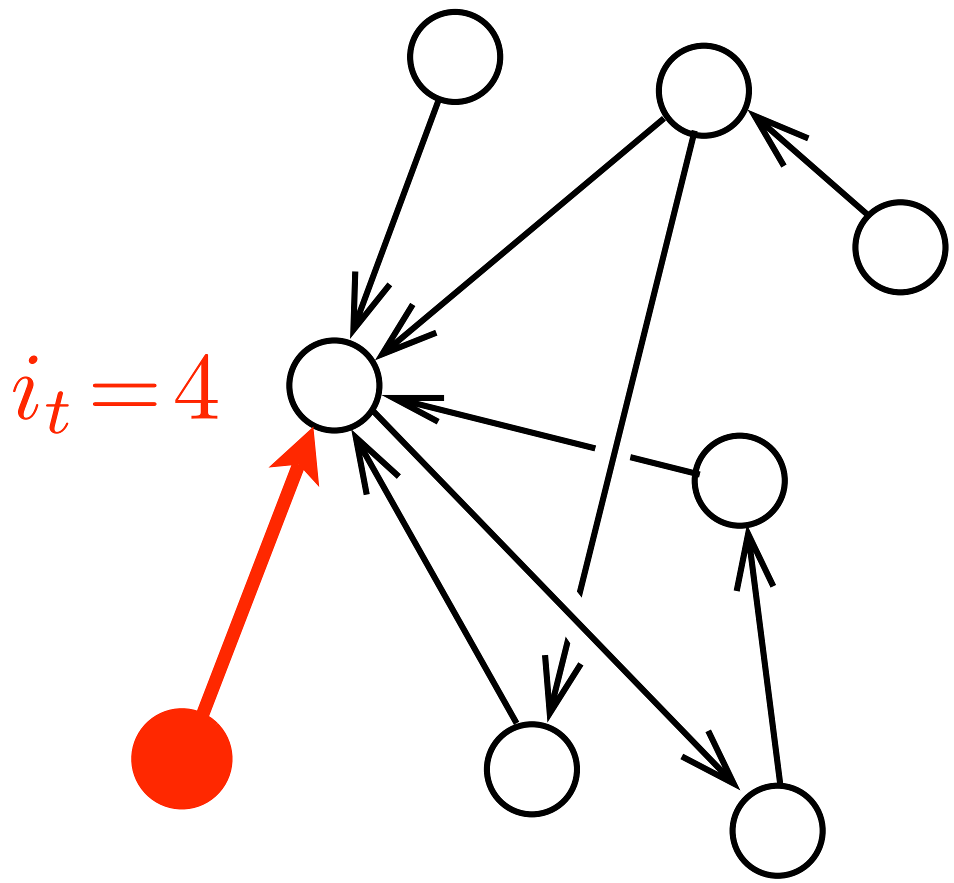
Goal: generate independent in-degree & out-degree distributions *dynamically*

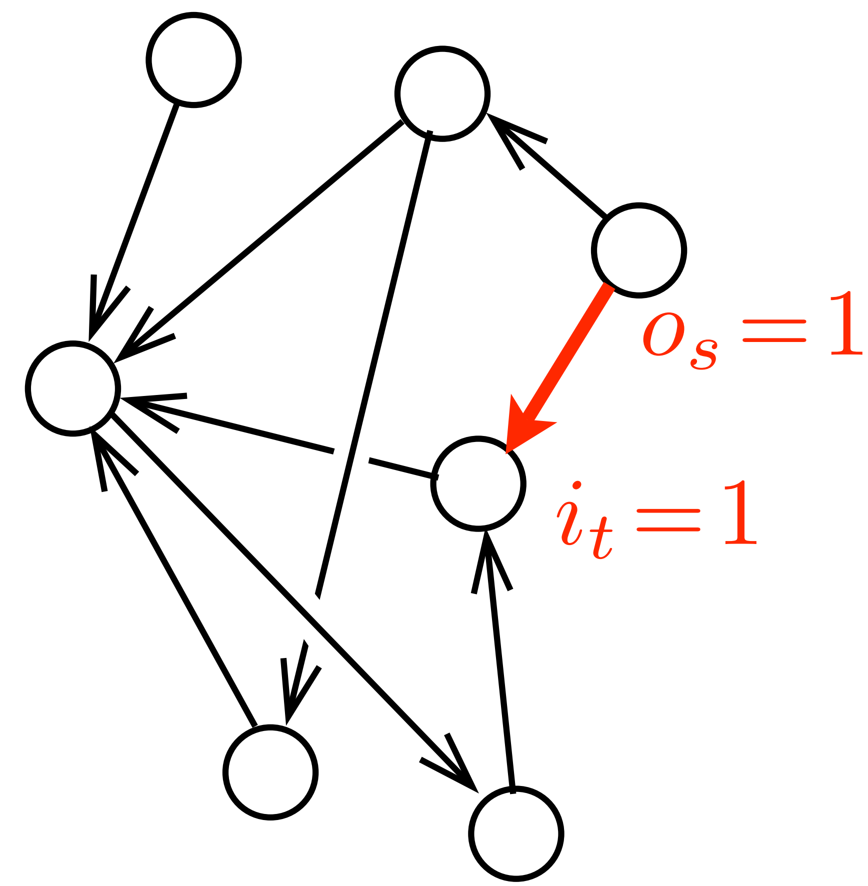
A Web Growth Model



probability p: new node attaches with rate that depends on in-degree of target

A Web Growth Model





probability p: new node attaches with rate that depends on in-degree of target

$$A_i = i + \lambda$$

probability q=1-p: new link whose attachment rate depends on out-degree of source & in degree of target

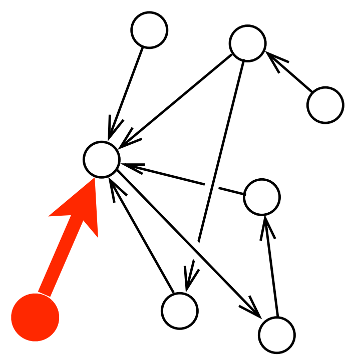
$$C(j, i) = (j + \mu)(i + \lambda)$$

$$\lambda > 0, \mu > -1 \text{ for } i \geq 0 \ \& \ j \geq 1$$

Macroscopic Properties

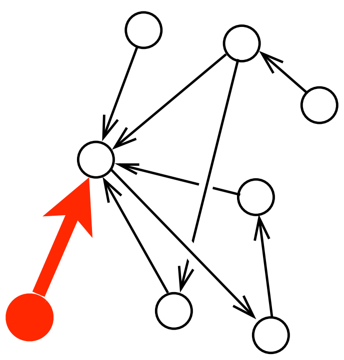
evolution of in/out degrees & number of nodes:

$$(N, I, J) \rightarrow \quad (N+1, I+1, J+1) \quad \text{prob. } p$$

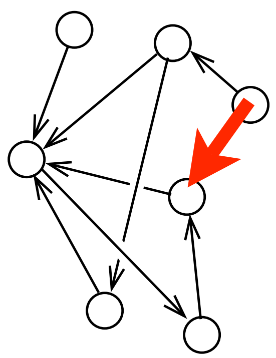


Macroscopic Properties

evolution of in/out degrees & number of nodes:



$$(N, I, J) \to \begin{cases} (N+1, I+1, J+1) & \text{prob. } p \\ (N, I+1, J+1) & \text{prob. } q \end{cases}$$



$$\to N(t) = pt, \qquad I(t) = J(t) = t$$

$$\mathcal{D}_{\text{in}} \equiv \frac{I(t)}{N(t)} = \frac{1}{p} \qquad \mathcal{D}_{\text{out}} \equiv \frac{J(t)}{N(t)} = \frac{1}{p}$$

Master Equation for Degree Distribution

node/link creation — link creation — change in the in-degree

$$\frac{dN_{ij}}{dt} = (p+q)\left[\frac{(i-1+\lambda)N_{i-1,j} - (i+\lambda)N_{ij}}{I+\lambda N}\right] + q\left[\frac{(j-1+\mu)N_{i,j-1} - (j+\mu)N_{ij}}{J+\mu N}\right] + p\,\delta_{i0}\delta_{j1}$$

link creation — change in the out-degree

sum rules:

$$\dot{N} = \sum_{i,j} \dot{N}_{ij} = p \quad \dot{I} = \sum_{i,j} i\dot{N}_{ij} = 1 \quad \dot{J} = \sum_{i,j} j\dot{N}_{ij} = 1$$

$$\textstyle\sum_{i,j}(i+\lambda)N_{ij} = I + \lambda N$$

Degree Densities and Their Master Equations

use: $N_{ij}(t) = t\, n_{ij} \quad N = pt \quad I = J = t$

master eqn → $[i+a(j+\mu)+b]n_{ij} = (i-1+\lambda)n_{i-1,j} + a(j-1+\mu)n_{i,j-1} + p(1+p\lambda)\delta_{i,0}\delta_{j,1}$

$$a = q\,\frac{1+p\lambda}{1+p\mu} \qquad b = 1+(1+p)\lambda$$

In- and out-degree distributions & densities:

$$\mathcal{I}_i(t) = \sum_j N_{ij}(t) = t\, I_i \qquad \mathcal{O}_j(t) = \sum_i N_{ij}(t) = t\, O_j$$

master eqns → $(i+b)I_i = (i-1+\lambda)I_{i-1} + p(1+p\lambda)\delta_{i,0}$

$$\left(j+\frac{1}{q}+\frac{\mu}{q}\right) O_j = (j-1+\mu)O_{j-1} + p\,\frac{1+p\mu}{q}\,\delta_{j,1}$$

Asymptotic Solution

$$I_i \;=\; I_0\,\frac{\Gamma(i+\lambda)\,\Gamma(b+1)}{\Gamma(i+b+1)\,\Gamma(\lambda)} \qquad \sim i^{-\nu_{\rm in}}$$

$$O_j \;=\; O_1\,\frac{\Gamma(j+\mu)\;\Gamma(2+q^{-1}+\mu q^{-1})}{\Gamma(j+1+q^{-1}+\mu q^{-1})\,\Gamma(1+\mu)} \quad \sim j^{-\nu_{\rm out}}$$

$$I_0 = p\frac{1+p\lambda}{b} \qquad O_1 = p\frac{1+p\mu}{1+q+\mu}$$

$$\nu_{\rm in} \;=\; 2+p\lambda$$
$$\nu_{\rm out} \;=\; 1+q^{-1}+\mu p q^{-1}$$

Finiteness and Fluctuations

Questions:

- *How to characterize fluctuations among realizations?*
- *Is a typical graph close to the "average" graph?*

More generally, describe a network by:

$$P(\{N_1, N_2, N_3, \ldots\}) \quad N_k = \#\ \text{nodes of degree}\ k$$

Network evolves by:

$$(N_1, N_2) \rightarrow (N_1, N_2+1),$$ attach to node of degree 1

$$(N_1, N_k, N_{k+1}) \rightarrow (N_1+1, N_k-1, N_{k+1}+1)$$ attach to node of degree >1

Program: write master equation for $P(\{N_1, N_2, N_3, \ldots\})$

& compute $\langle N_k \rangle \quad \langle N_j N_k \rangle \quad \langle N_i N_j N_k \rangle \quad \langle N_k^2 \rangle$

etc.

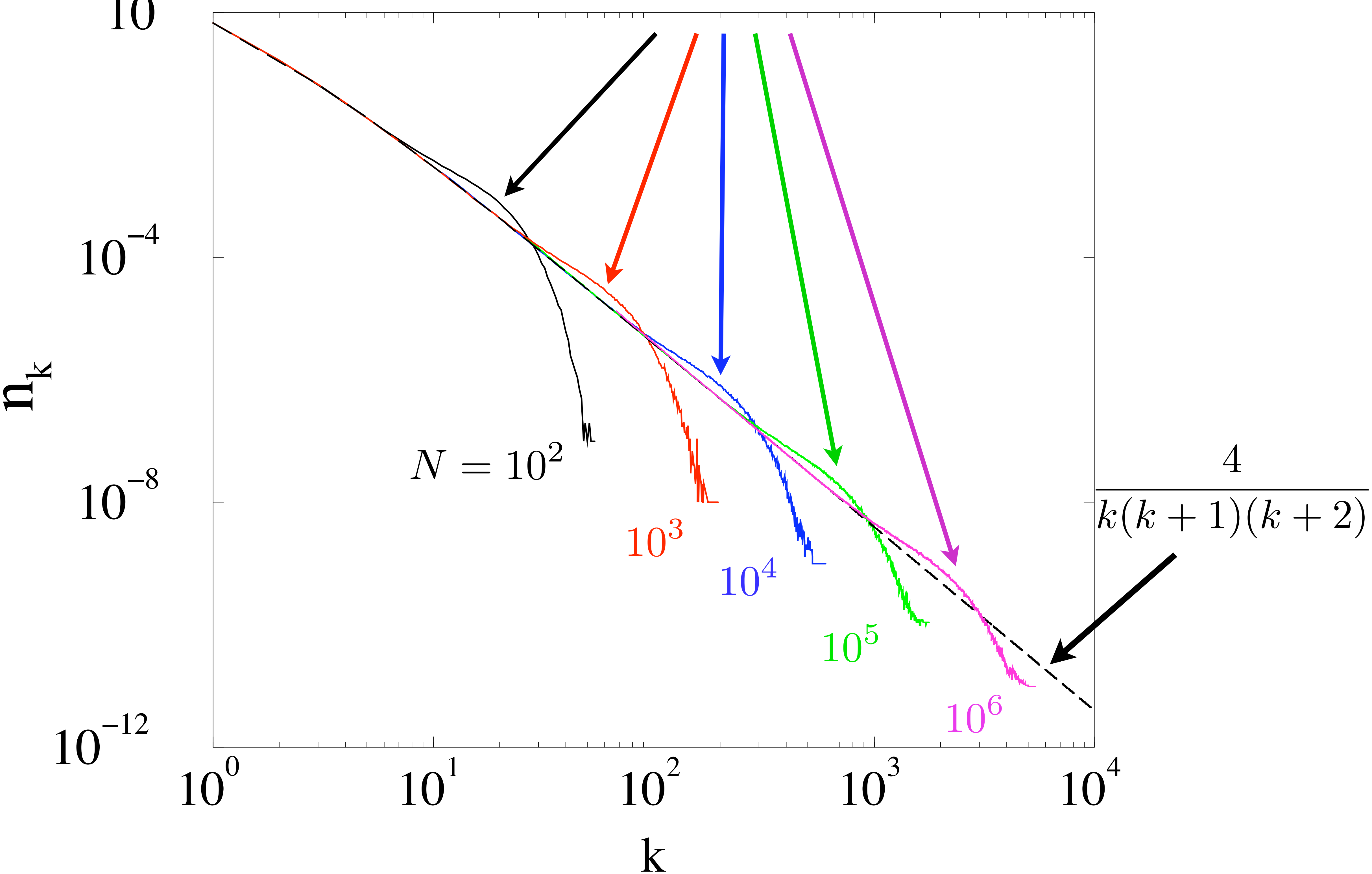
Degree Distribution of a *Finite* Network

trimer IC

$$N_k(N) \sim N\, n_k\, F(k/k_{\max})$$

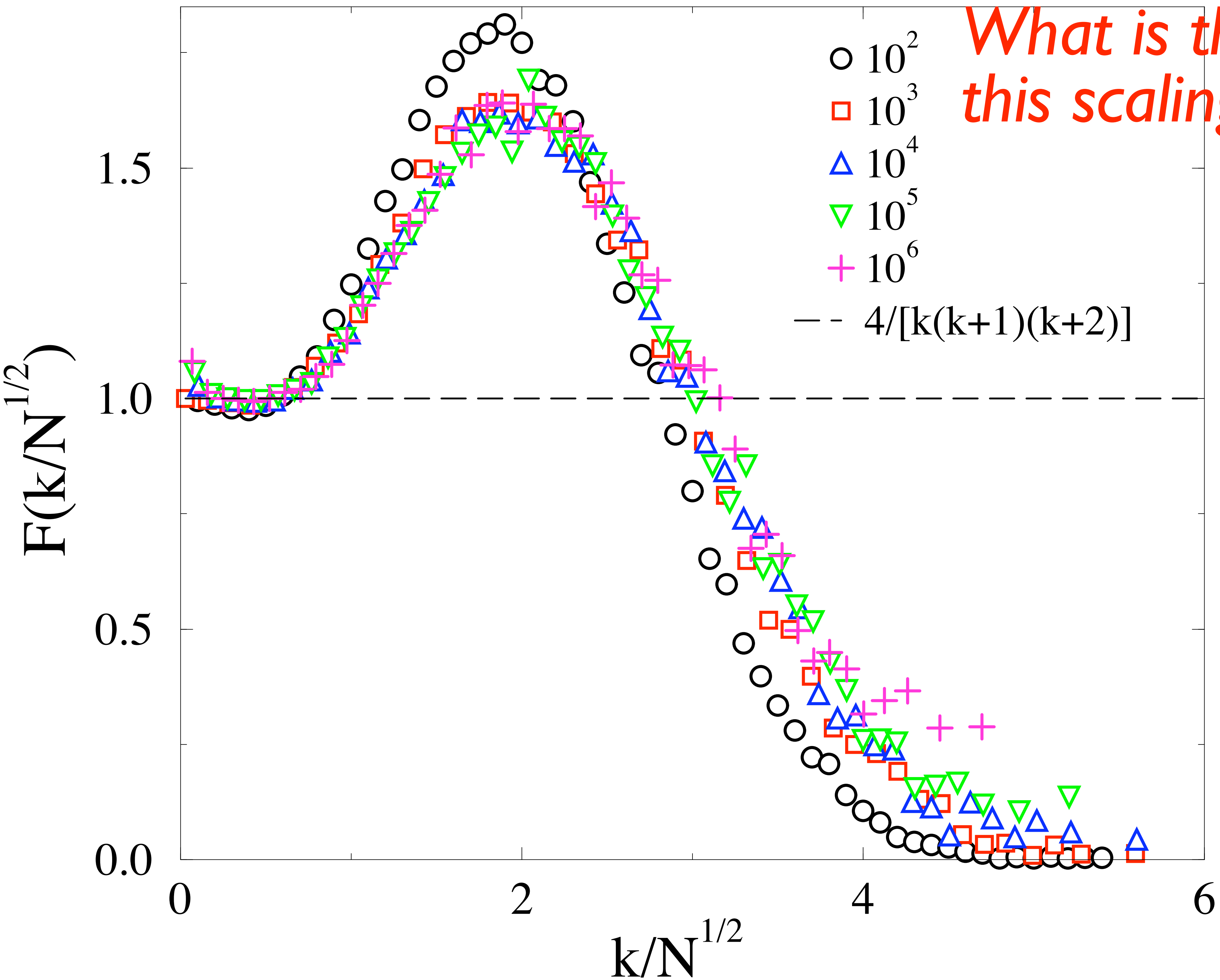
$$k_{\max} \propto N^{1/2}$$

(for $N_k \propto k^{-3}$)



The Scaling Function $F(\xi) = \frac{N_k(N)}{N} \frac{k(k+1)(k+2)}{4}$

What is the form of this scaling function?



Evolution of Fixed-Degree Nodes

k=1:
$$N_1(N+1) = \begin{cases} N_1(N) & \text{prob. } \dfrac{N_1}{2N} \\ N_1(N)+1 & \text{prob. } 1-\dfrac{N_1}{2N} \end{cases}$$

master eqn:
$$\langle N_1(N+1)\rangle = \left\langle \frac{N_1^2(N)}{2N} \right\rangle + \left\langle [N_1(N)+1]\left[1-\frac{N_1(N)}{2N}\right]\right\rangle$$
$$= 1 + \left(1-\frac{1}{2N}\right)\langle N_1(N)\rangle$$

Introduce $\mathcal{G}_1(w) = \sum_{N\geq 1}^{\infty} \langle N_1(N)\rangle\, w^{N-1} \to \dfrac{d\mathcal{G}_1}{dw} = \dfrac{1}{(1-w)^2} + \dfrac{1}{2}\,\mathcal{G}_1 + w\,\dfrac{d\mathcal{G}_1}{dw}$

Solution:
$$\mathcal{G}_1(w) = \frac{2}{3}\,\frac{1}{(1-w)^2} + \frac{4}{3}\,\frac{1}{(1-w)^{1/2}}$$

Invert:
$$\langle N_1(N)\rangle = \frac{2}{3}\,N + \frac{4}{\sqrt{9\pi}}\,\frac{\Gamma\left(N-\frac{1}{2}\right)}{\Gamma(N)} \sim \frac{2}{3}\,N + \frac{4}{\sqrt{9\pi N}}$$

Discrete vs. Continuum

Master equations:

$$\langle N_k(N+1)\rangle - \langle N_k(N)\rangle = \left\langle \frac{(k-1)N_{k-1}(N) - kN_k(N)}{2N} \right\rangle + \delta_{k,1}$$

$$\frac{d\,\langle N_k(N)\rangle}{dN} = \left\langle \frac{(k-1)N_{k-1}(N) - kN_k(N)}{2N} \right\rangle + \delta_{k,1}$$

Solutions for N_1:

discrete (exact): $\langle N_1(N)\rangle = \frac{2}{3}\,N + \frac{4}{3\sqrt{\pi}}\,\frac{\Gamma\left(N - \frac{1}{2}\right)}{\Gamma(N)}$

continuum: $\langle \dot{N}_1\rangle = -\frac{\langle N_1\rangle}{2N} + 1 \qquad \langle N_1(N)\rangle = \frac{2}{3}\,N + \frac{4}{3\sqrt{N}}$

similarly for N_2, N_3, N_4, etc.

mean-square degree:

The processes:

$$N_1(N+1) = \begin{cases} N_1(N) & \text{prob. } \dfrac{N_1}{2N} \\ N_1(N)+1 & \text{prob. } 1-\dfrac{N_1}{2N} \end{cases}$$

master eqn:

$$\langle N_1^2(N+1)\rangle = 1 + \left(1-\frac{1}{N}\right)\langle N_1^2(N)\rangle + \left(2-\frac{1}{2N}\right)\langle N_1(N)\rangle$$

Solution:

$$\begin{aligned}\langle N_1^2(N)\rangle &= \frac{4}{9}N(N+1) - \frac{1}{3}N + \frac{16}{9\sqrt{\pi}}\,\frac{\Gamma\left(N+\frac{1}{2}\right)}{\Gamma(N)} \\ &- \frac{4}{3\sqrt{\pi}}\,\frac{\Gamma\left(N-\frac{1}{2}\right)}{\Gamma(N)} + \frac{35}{9}\,\delta_{N,1}\end{aligned}$$

Dispersion:

$$\sigma_1^2 = \langle N_1^2\rangle - \langle N_1\rangle^2 = \frac{N}{9} - \frac{20}{9\sqrt{\pi}}\,\frac{1}{N^{1/2}} - \frac{16}{9\sqrt{\pi}}\,\frac{1}{N} + \ldots$$

master eqn for pdf:

$$P(N_1, N+1) = \frac{N_1}{2N}\,P(N_1,N) + \left(1-\frac{N_1-1}{2N}\right)P(N_1-1,N)$$

$\rightarrow$ Gaussian, with mean $\frac{2}{3}N$, width $\frac{1}{3}\sqrt{N}$

Evolution of Fixed-Degree Nodes (k>1)

The processes:

$$N_k(N+1) = \begin{cases} N_k - 1 & \text{prob. } \dfrac{kN_k}{2N} \\ N_k + 1 & \text{prob. } \dfrac{(k-1)N_{k-1}}{2N} \\ N_k & \text{prob. } 1 - \dfrac{(k-1)N_{k-1} + kN_k}{2N} \end{cases}$$

master eqn:

$$\langle N_k(N+1)\rangle = \langle N_k(N)\rangle + \left\langle \frac{(k-1)N_{k-1}(N) - kN_k(N)}{2N} \right\rangle$$

soln:

$$\langle N_1(N)\rangle = \frac{2}{3}\,N + \frac{4}{3\sqrt{\pi}}\,\frac{\Gamma\left(N-\frac{1}{2}\right)}{\Gamma(N)}$$

$$\langle N_2(N)\rangle = \frac{1}{6}\,N + \frac{4}{3\sqrt{\pi}}\,\frac{\Gamma\left(N-\frac{1}{2}\right)}{\Gamma(N)} - \frac{3}{2}\,\delta_{N,1}$$

$$\langle N_3(N)\rangle = \frac{1}{15}\,N + \frac{4}{3\sqrt{\pi}}\,\frac{\Gamma\left(N-\frac{1}{2}\right)}{\Gamma(N)} - 3\,\delta_{N,1} - \frac{4}{5\sqrt{\pi}}\,\frac{\Gamma\left(N-\frac{3}{2}\right)}{\Gamma(N)}$$

etc.

Finite-Size Scaling of the Degree Distribution

master eqn:
$$\langle N_k(N+1)\rangle = \langle N_k(N)\rangle + \left\langle \frac{(k-1)N_{k-1}(N) - kN_k(N)}{2N} \right\rangle$$

generating fn:
$$\mathcal{N}(w,z) = \sum_{N=1}^{\infty}\sum_{k=1}^{\infty}\langle N_k(N)\rangle\, w^{N-1}\, z^k$$

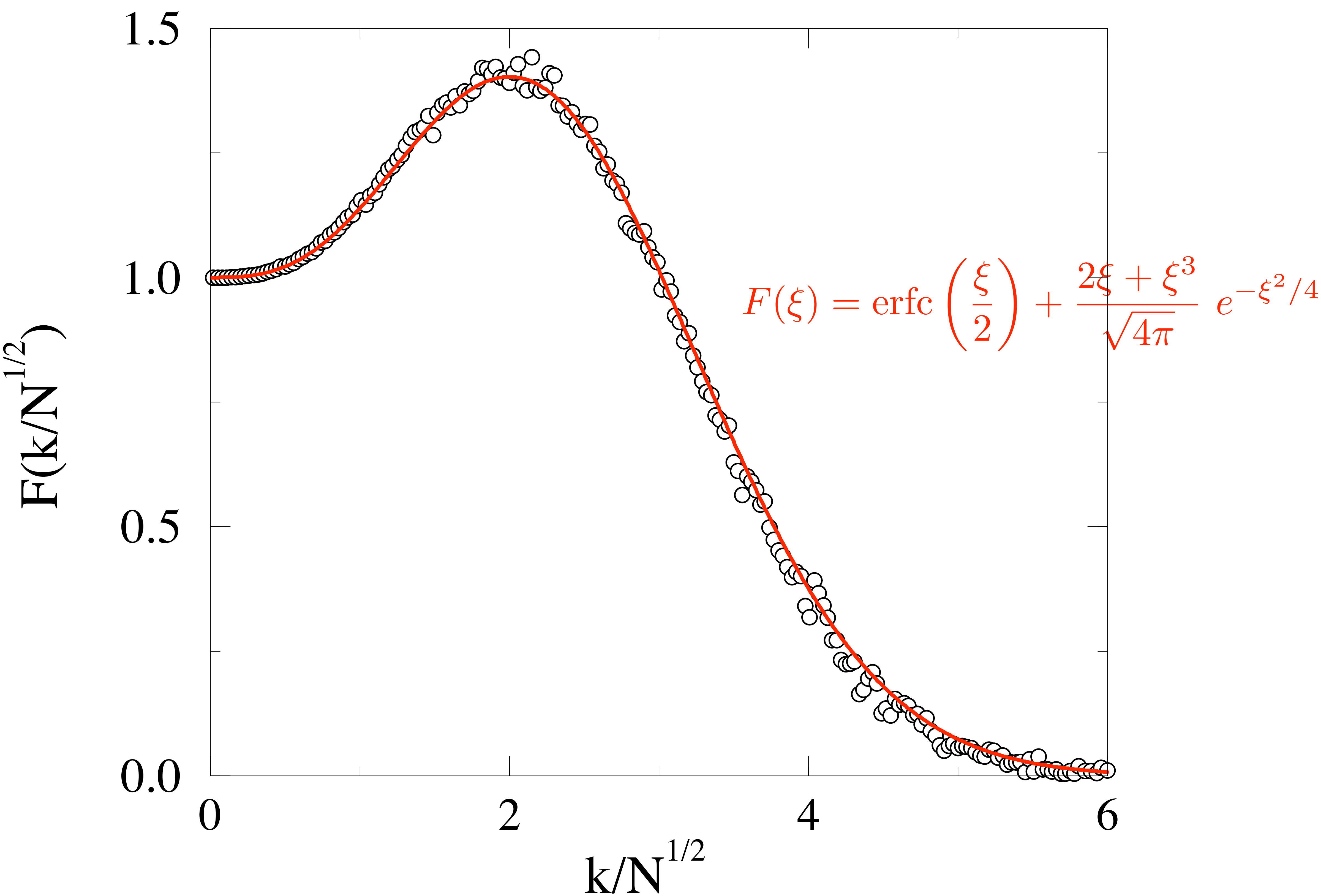
$$\rightarrow \quad \left[2(1-w)\frac{\partial}{\partial w} + z(1-z)\frac{\partial}{\partial z} - 2\right]\mathcal{N} = \frac{2z}{(1-w)^2}$$

soln:
$$\mathcal{N}(w,z) = \frac{(3-2z^{-1})}{(1-w)^2} - \frac{1}{1-w} + \frac{2(z^{-1}-1)}{(1-w)^{3/2}}$$
$$+ \frac{2(1-w)^{-1/2}}{(z^{-1}-1)+(1-w)^{1/2}} - \frac{2(z^{-1}-1)^2}{(1-w)^2}\ln\left[1-z+z(1-w)^{1/2}\right]$$

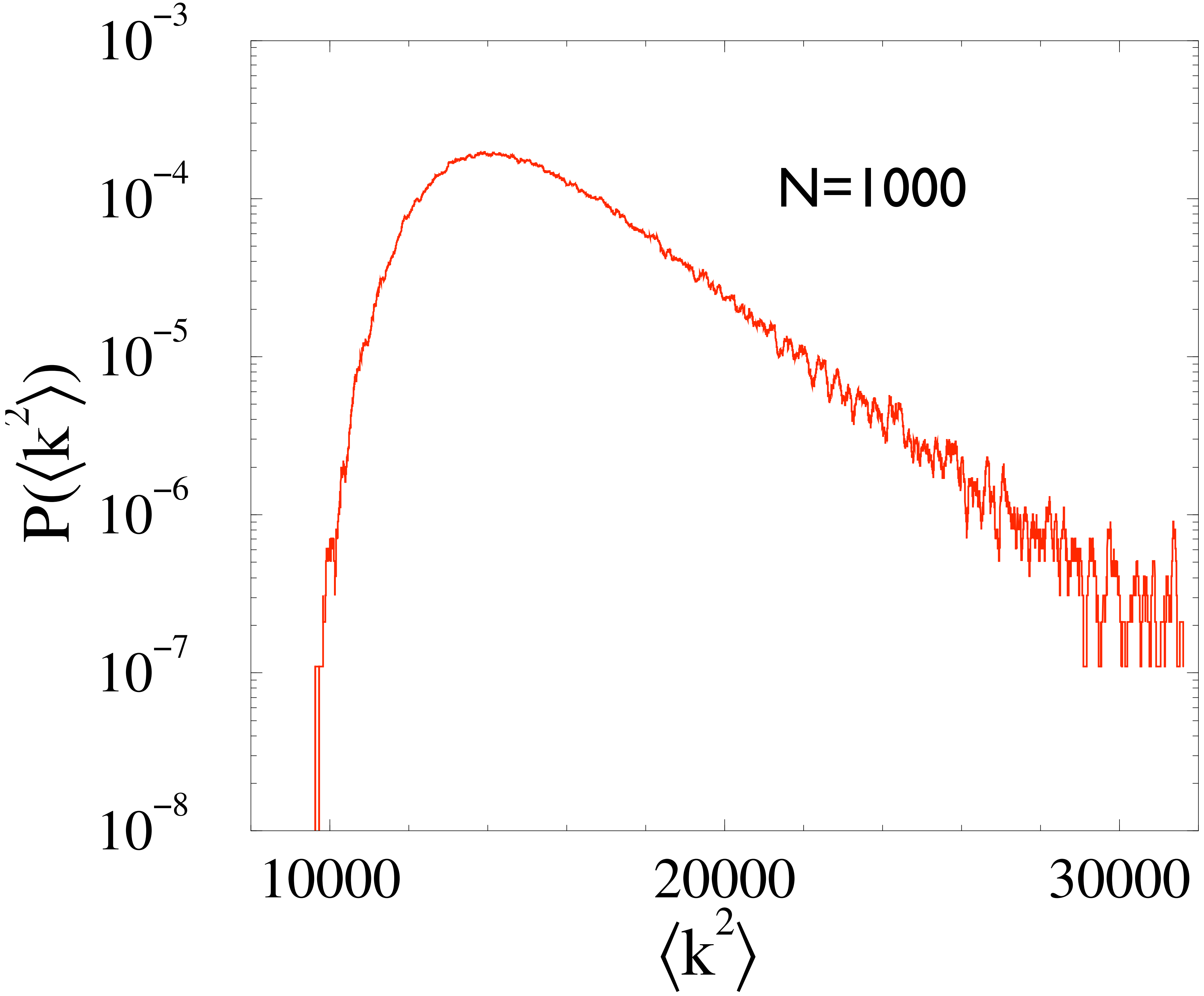
dimer IC

Invert:
$$F(\xi) = \operatorname{erfc}\left(\frac{\xi}{2}\right) + \frac{2\xi+\xi^3}{\sqrt{4\pi}}\, e^{-\xi^2/4}$$

The Scaling Function (dimer initial state)



Diversity:
Distribution of Mean-Square Degree



Outlook

The master equation is a powerful tool to analyze incrementally growing, complex networks

Wide range of degree distributions arise by preferential attachment

Various topological features and more general models are treatable by the ME approach

Some open questions/future possibilities:

Why is *linear* preferential attachment so generic?

Incorporation of spatial structure.

Formulation & analysis of realistic social network models and their dynamics (*communities, frustration, etc*.)

Davidsen et al., PRL (2002)

Networks inspired by biology:

metabolism, epidemics, protein interactions, diseases