Growth models

How do Power law DDs arise?

One Possible Answer: The Barabási-Albert Model of Preferential Attachment (Rich gets Richer)

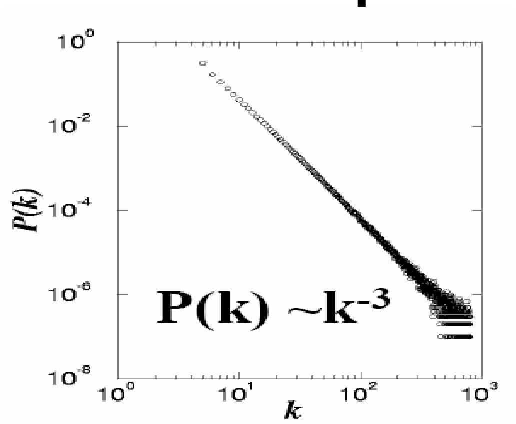
GROWTH :

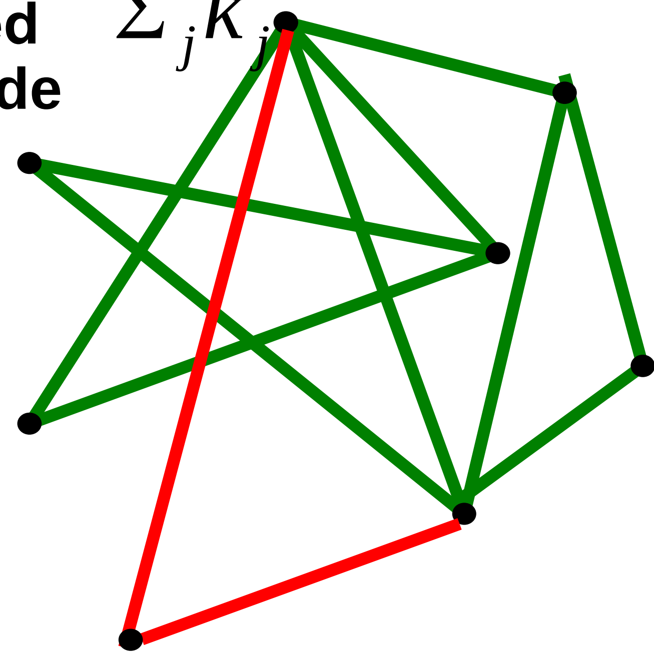
Starting with a small number of nodes (n_0) at every timestep we add a new node with *m* ($<=n_0$) edges (connected to the nodes already present in the system).

(2) PREFERENTIAL ATTACHMENT :

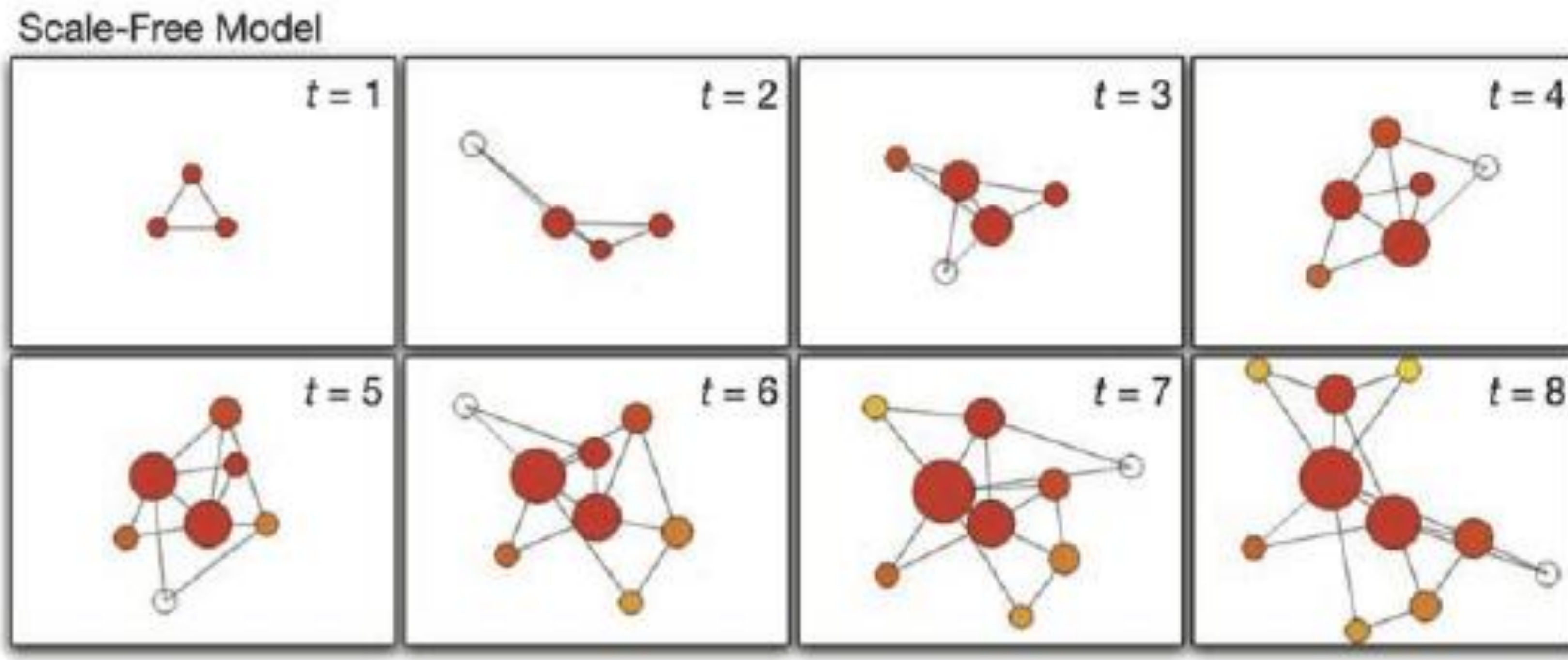
$$\Pi(k_i) = \frac{k_i}{\Sigma_j k_j}$$

The probability Π that a new node will be connected to node *i* depends on the connectivity k_i of that node





Time evolution of degree



Time evolution of degree (Preferential attachment)

Continues time, mean field approximation:

$$k_i(t+\delta t) = k_i(t) + m\Pi(k_i)\delta t$$

$$\frac{dk_i(t)}{dt} = m\Pi(k_i) = m\frac{k_i}{\sum_i k_i} = \frac{mk_i}{2mt}$$

$$\frac{dk_i(t)}{dt} = \frac{k_i(t)}{2t}$$

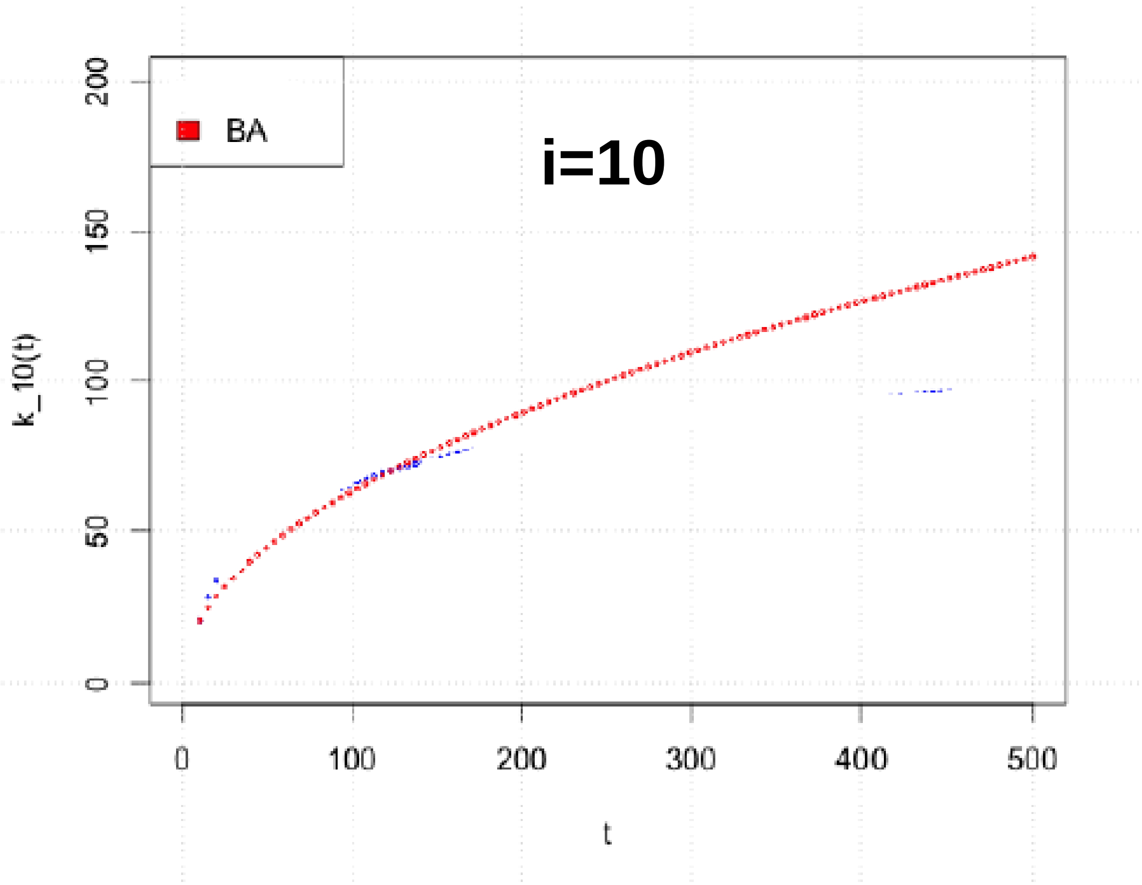
initial conditions: $k_i(i) = m$ t=i, k=m

$$\int_m^{k_i(t)} \frac{dk_i}{k_i} = \int_i^t \frac{dt}{2t}$$

Solution:

$$k_i(t) = m\left(\frac{t}{i}\right)^{1/2}$$

Time evolution of degree



$$BA : k_i(t) = m \left(\frac{t}{i}\right)^{1/2},$$

Random attachment

Continues time approximation $(\Pi(k_i) = \frac{1}{n_0+t-1} \approx \frac{1}{t})$

$$k_i(t+\delta t) = k_i(t) + m\Pi(k_i)\delta t$$

$$k_i(t+\delta t) = k_i(t) + \frac{m}{t}\delta t$$

Differential equation:

$$\frac{dk_i(t)}{dt} = \frac{m}{t}$$

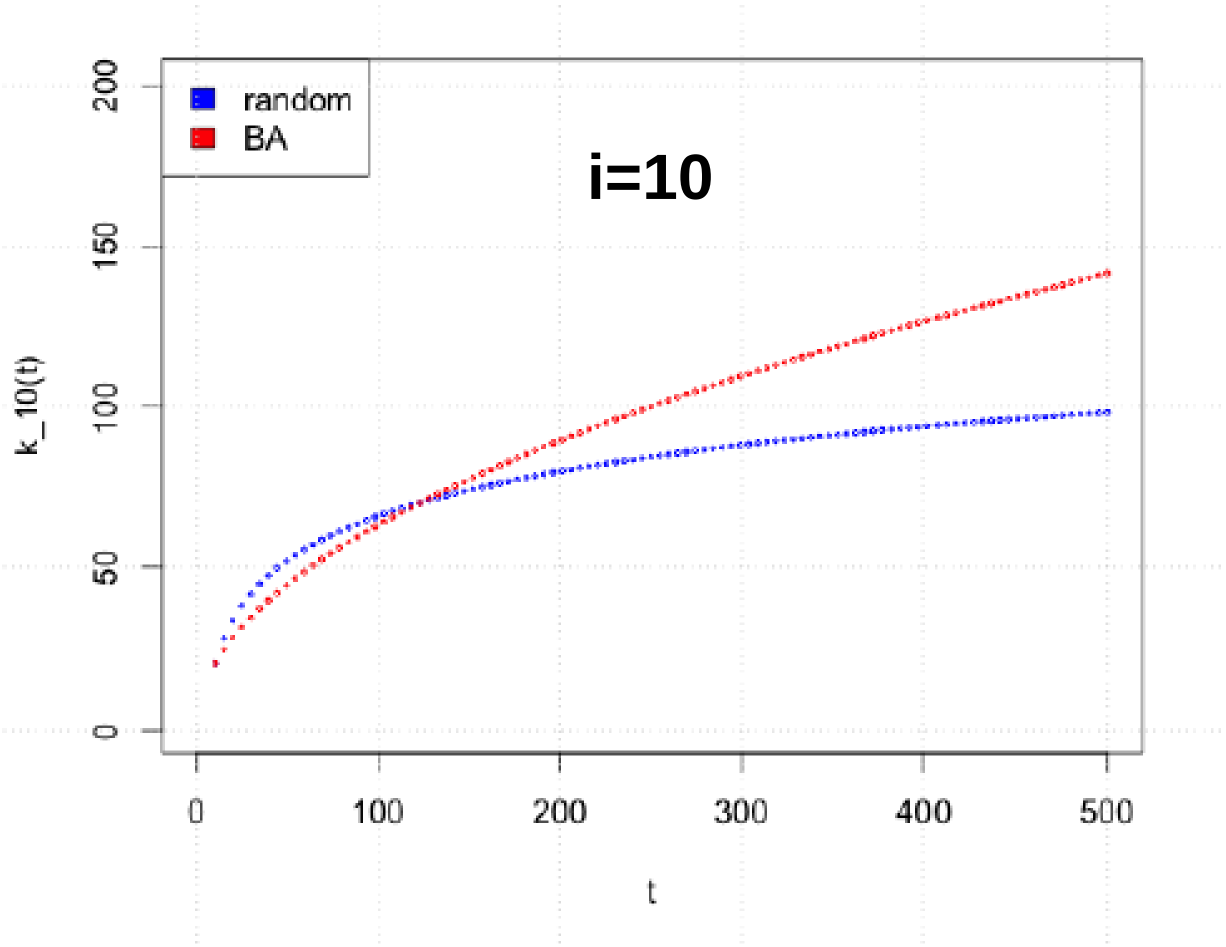
with initial conditions: $k_i(t=i) = m$

$$\int_m^{k_i(t)} \frac{dk_i}{m} = \int_i^t \frac{dt}{t}$$

Solution:

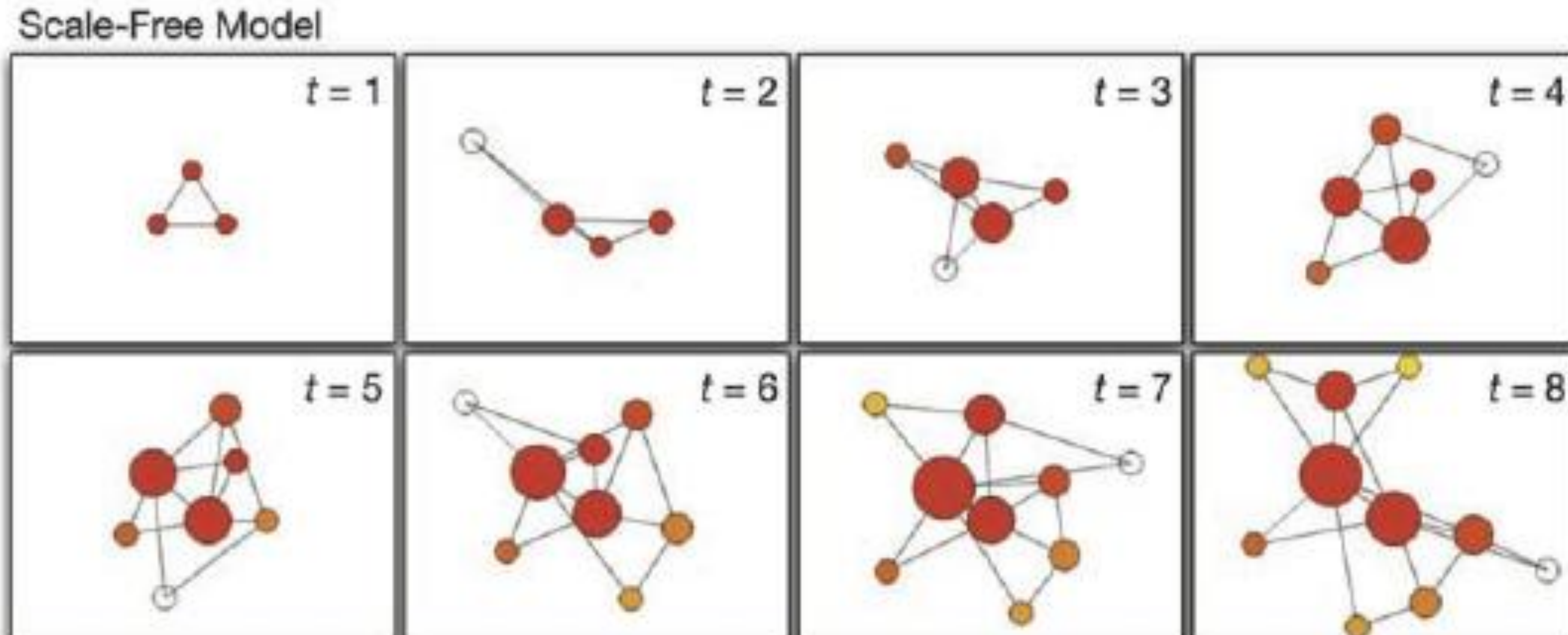
$$k_i(t) = m\left(1 + \log\left(\frac{t}{i}\right)\right)$$

m=20

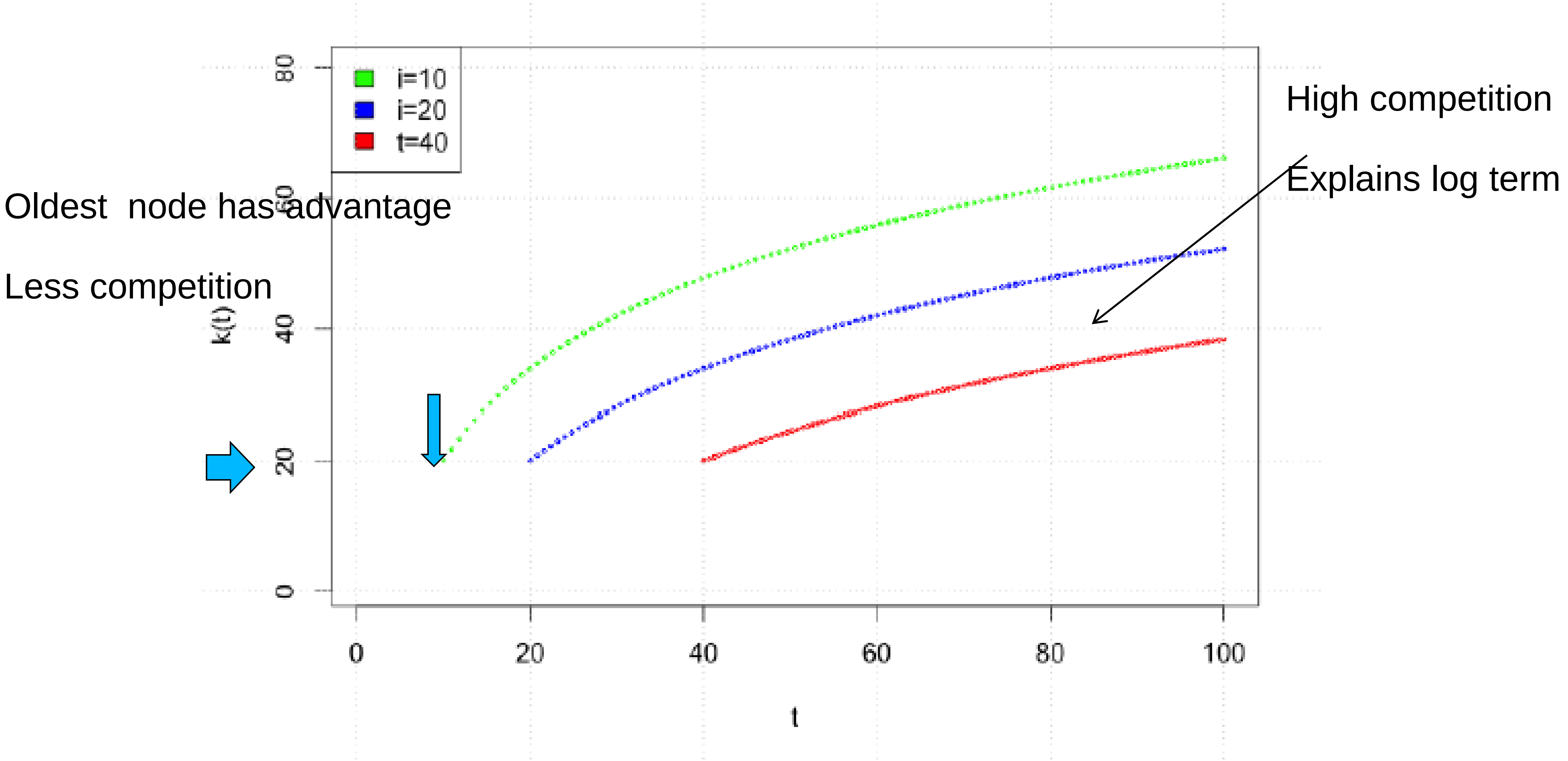


$$BA : k_i(t) = m\left(\frac{t}{i}\right)^{1/2}, \quad RG : k_i(t) = m\left(1 + \log\left(\frac{t}{i}\right)\right)$$

Time evolution of degree

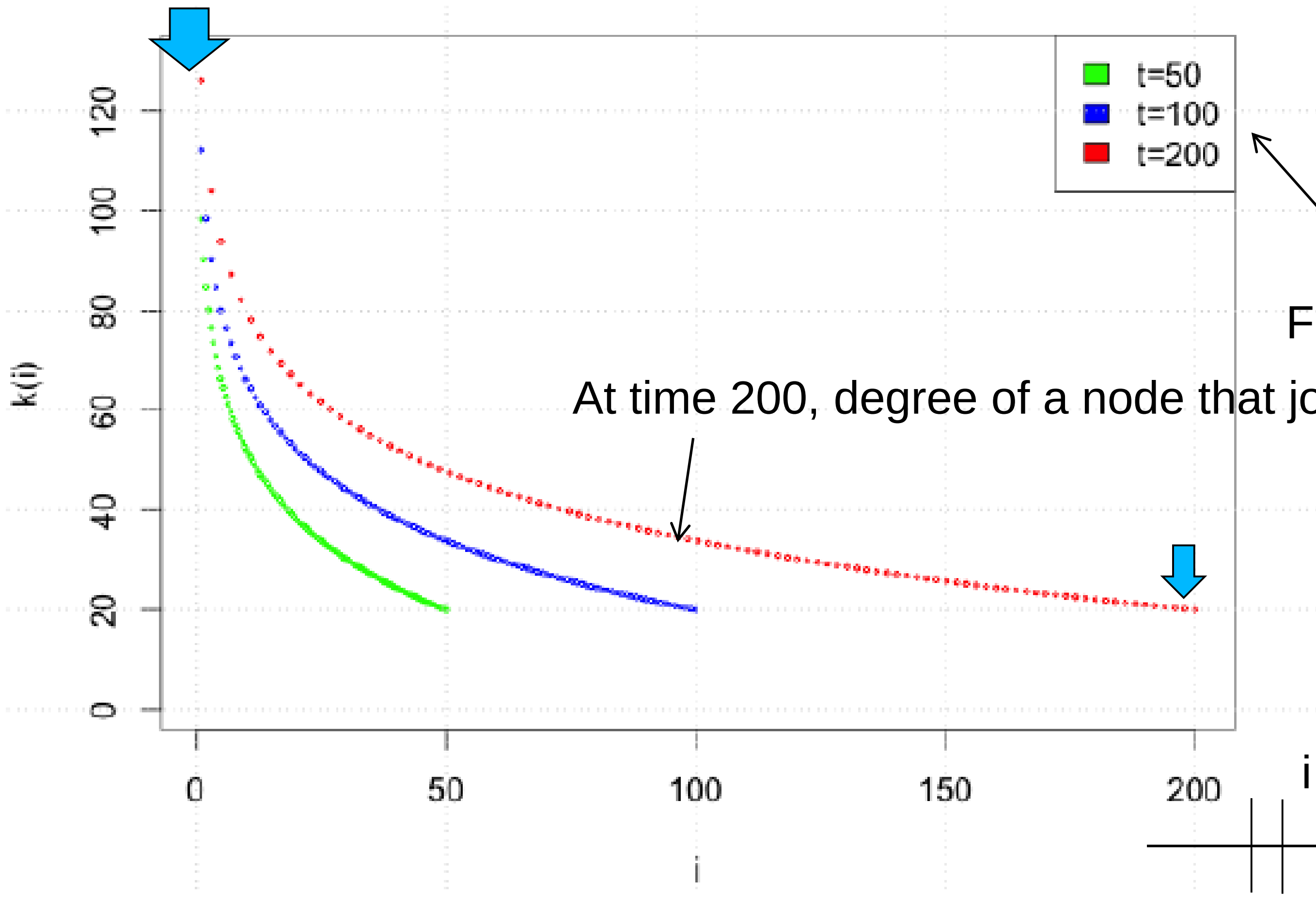


Time evolution of degree (random)



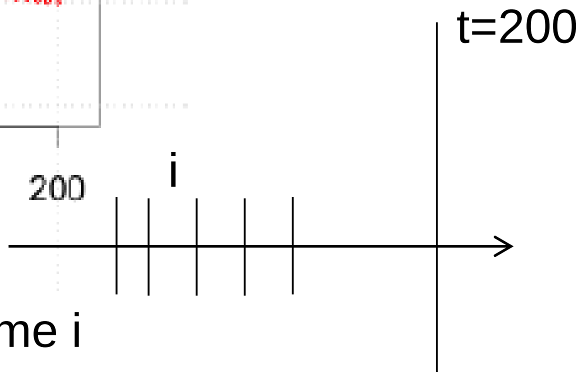
$$k_i(t) = m\left(1 + \log\left(\frac{t}{i}\right)\right), \quad m = 20, \; i = 10, 20, 40, \; t \geq i$$

Snapshot



Fix the timestep

At time 200, degree of a node that joins at i= 100



Degree of a node at t, which joined at time i

$$k_i(t) = m\left(1 + \log\left(\frac{t}{i}\right)\right), \quad m = 20, \;\; t = 50, 100, 200, \;\; i \leq t$$

Preferential attachment

Time evolution of a node degree

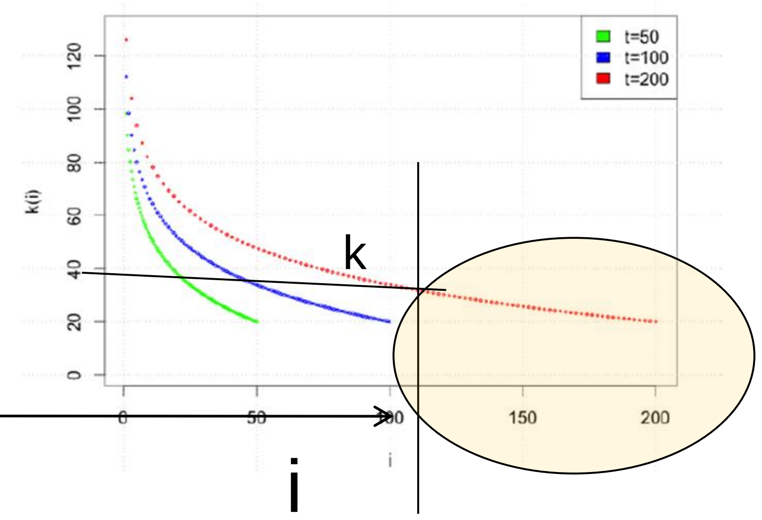
Find all the nodes at time t,
whose degree is <= k (say 40)

$$k_i(t) = m\left(\frac{t}{i}\right)^{1/2}$$

Nodes with $k_i(t) \le k$:

$$m\left(\frac{t}{i}\right)^{1/2} \le k$$

$$i \ge \frac{m^2}{k^2} t$$



Fraction of nodes with $k_i(t) \le k$ (CDF):

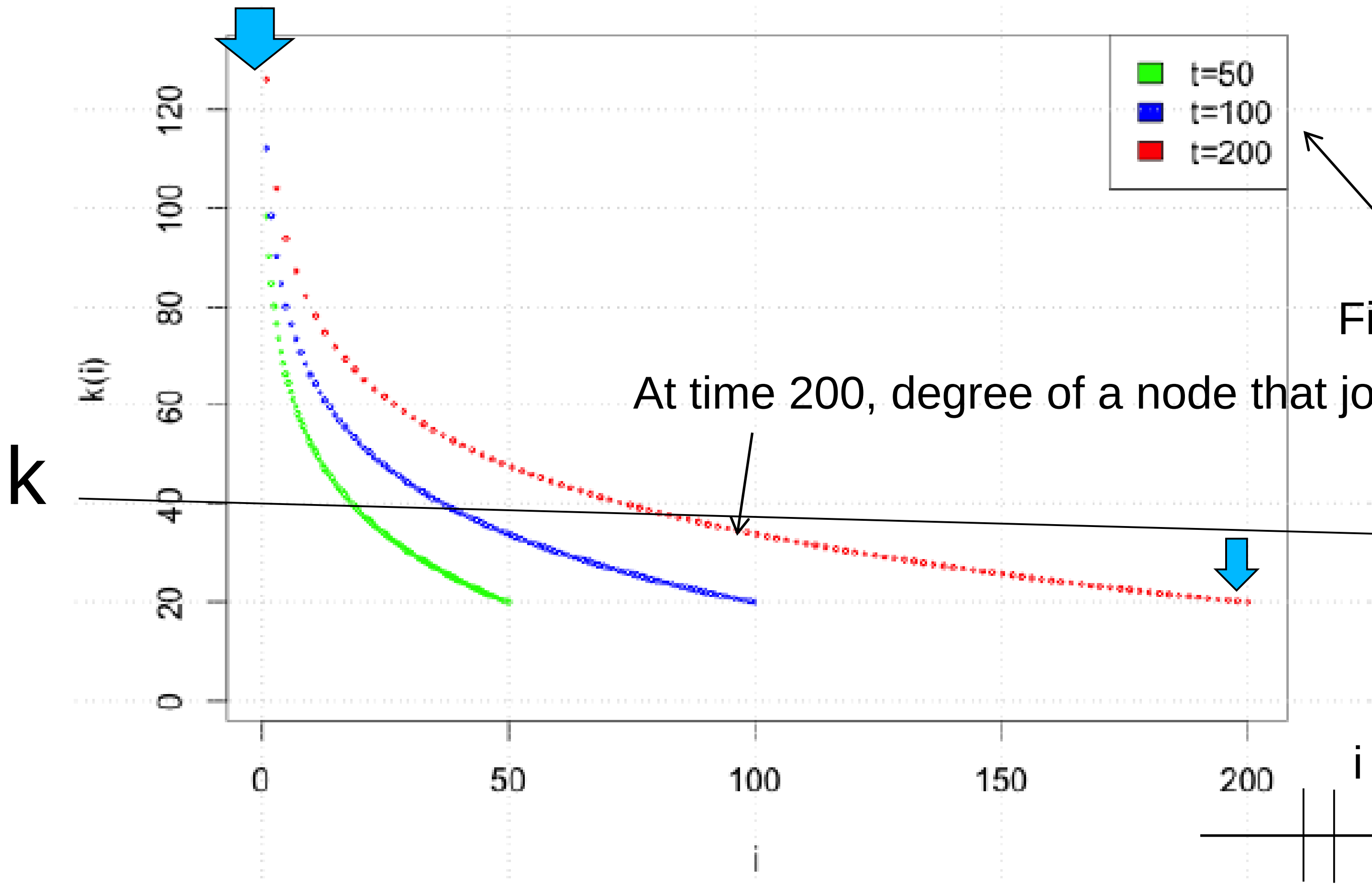
$$F(k) = P(k_i(t) \le k) = \frac{n_0 + t - i}{n_0 + t} = \frac{n_0 + t - m^2 t/k^2}{n_0 + t} \approx 1 - \frac{m^2}{k^2}$$

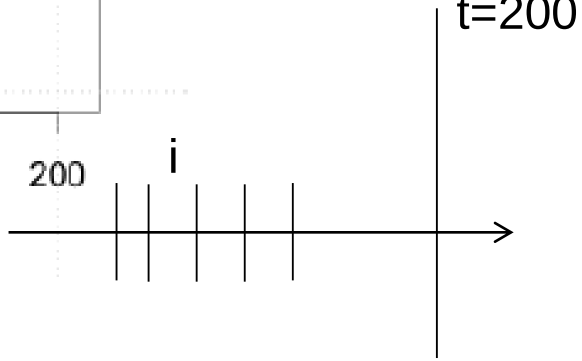
Distribution function:

$$P(k) = \frac{d}{dk} F(k) = \frac{2m^2}{k^3}$$

Power law

Snapshot





Degree of all the nodes joining before i

$$k_i(t) = m\left(1+\log\left(\frac{t}{i}\right)\right), \quad m=20,\ t=50,100,200,\ i\le t$$

Random attachment

Find all nodes that at time t has degree less than k, $k_i(t) \leq k$?
(for example $k_i(t) \leq 40$)

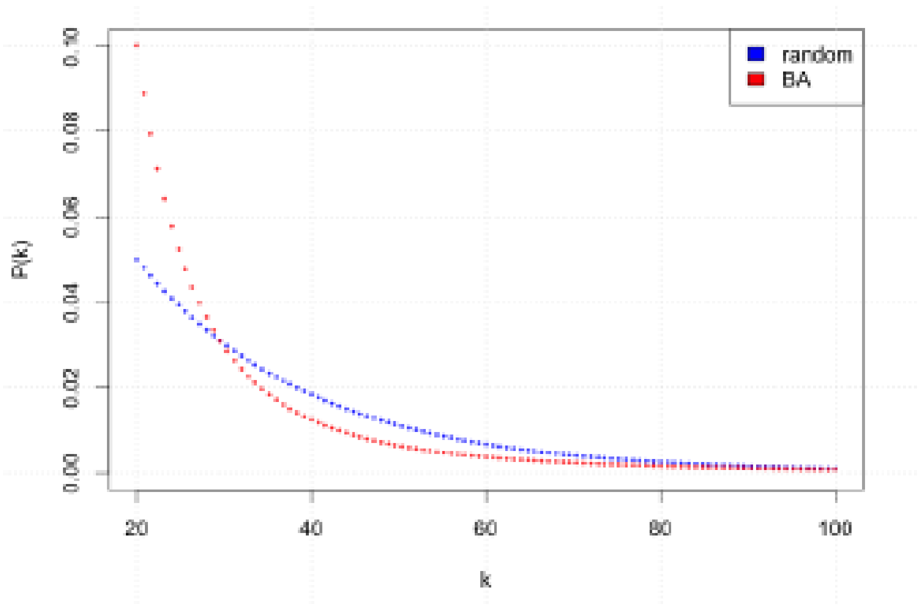
$$\begin{aligned} k_i(t) = m\left(1 + \log\left(\frac{t}{i}\right)\right) &\leq k \\ \log\left(\frac{t}{i}\right) &\leq \frac{k}{m} - 1 \\ \frac{t}{i} &\leq e^{\frac{k-m}{m}} \\ i &\geq te^{\frac{m-k}{m}} \end{aligned}$$

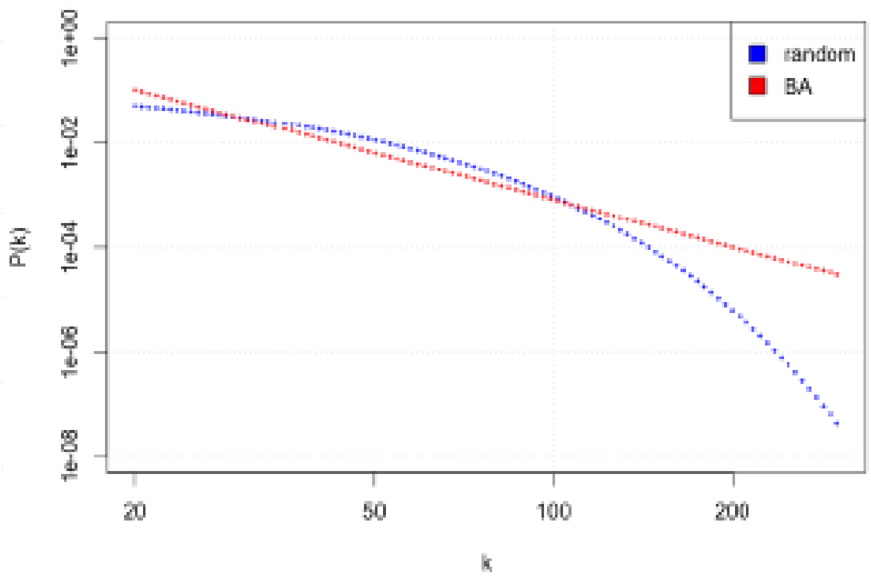
Fraction of nodes with degrees $k_i(t) \leq k$ (CDF):

$$F(k) = P(k_i(t) \leq k) = \frac{n_0 + t - i}{n_0 + t} = \frac{n_0 + t - te^{\frac{m-k}{m}}}{n_0 + t} \approx 1 - e^{\frac{m-k}{m}}$$

$$P(k) = \frac{d}{dk}F(k) = \frac{1}{m}e^{-\frac{k-m}{m}} = \frac{e}{m}e^{-\frac{k}{m}}, \quad k \geq m$$

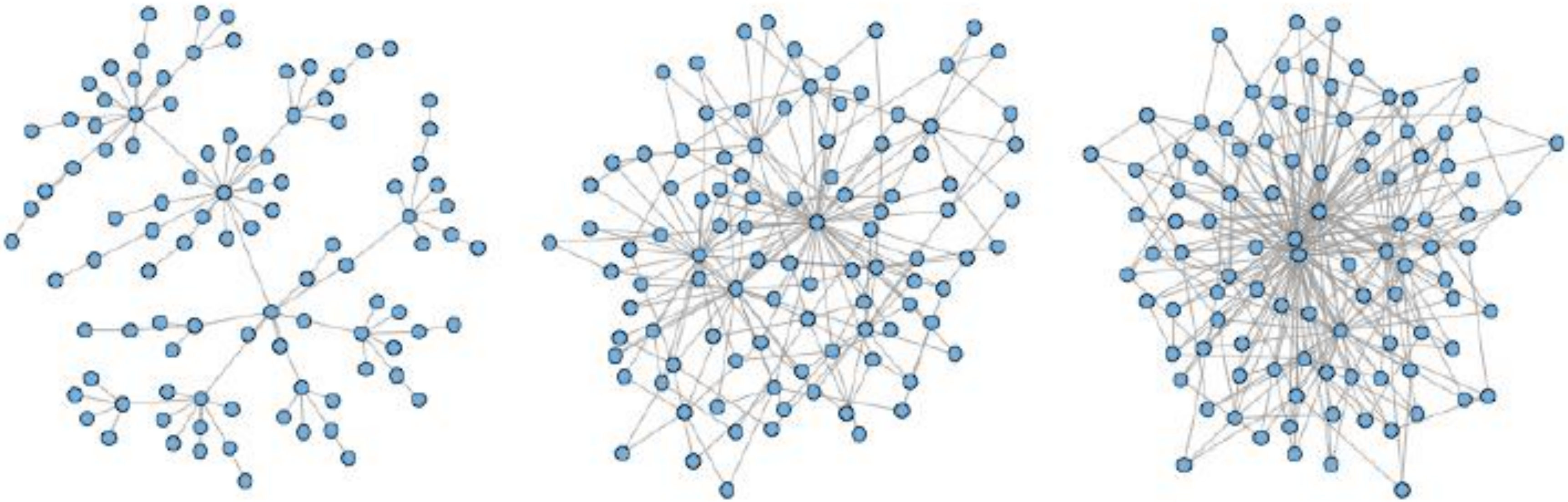
Degree distribution





$$BA : P(k) = \frac{2m^2}{k^3}, \qquad RG : P(k) = \frac{e}{m} e^{-\frac{k}{m}}$$

Preferential attachment



$m = 1$ $m = 2$ $m = 3$

Discrete time solution
Rate equation

How do Power law DDs arise?

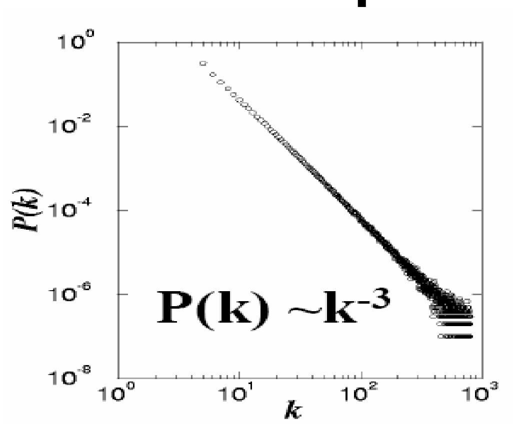
One Possible Answer: The Barabási-Albert Model of Preferential Attachment (Rich gets Richer)

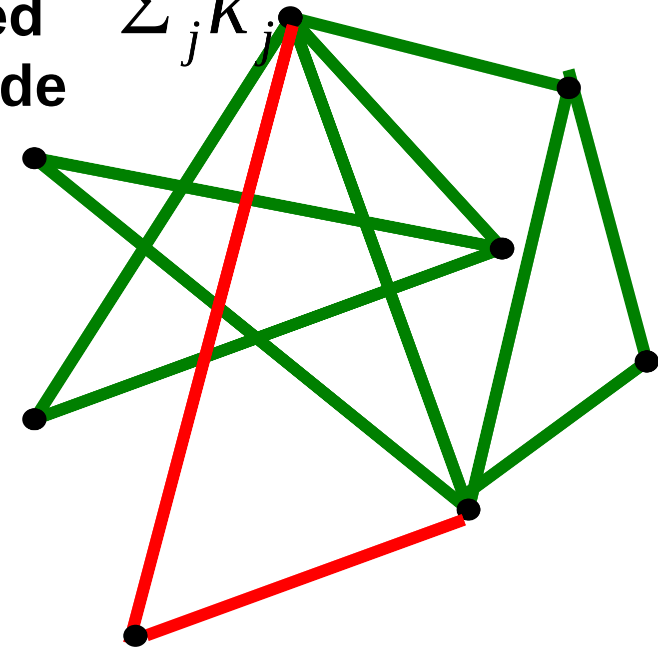
GROWTH :

Starting with a small number of nodes (m_0) at every timestep we add a new node with *m* (<=m_0) edges (connected to the nodes already present in the system).

(2) PREFERENTIAL ATTACHMENT : $$\Pi(k_i) = \frac{k_i}{\Sigma_j k_j}$$

The probability Π that a new node will be connected to node *i* depends on the connectivity k_i of that node





Preferential attachment

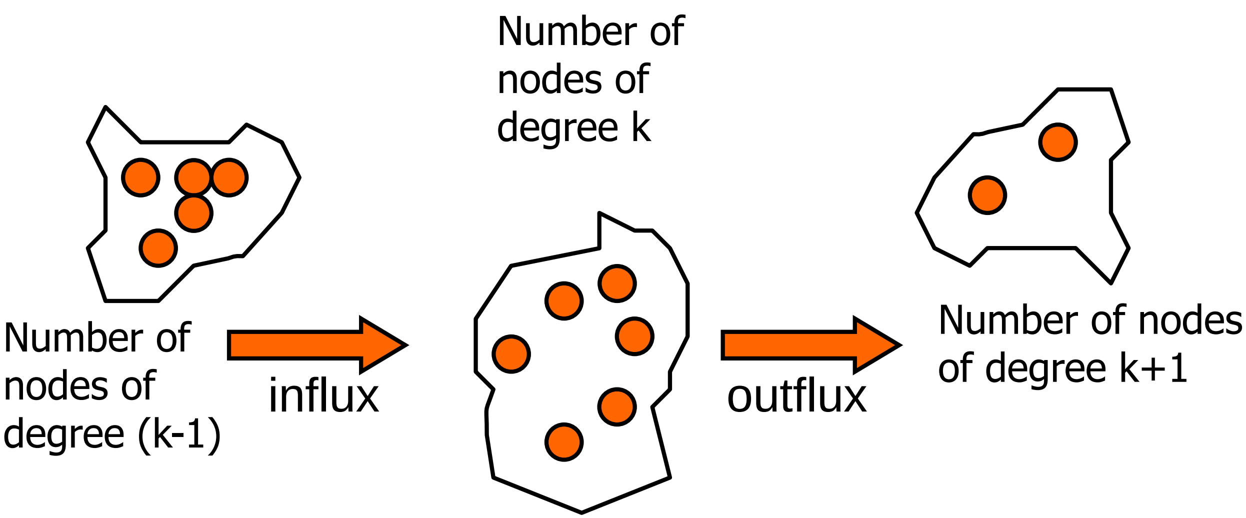
The probability that a new edge attaches to a vertex of degree k—the equivalent of Eq. (58)—is

$$\frac{kp_k}{\sum_k kp_k} = \frac{kp_k}{2m}. \tag{65}$$

Now the mean number of vertices of degree k that gain an edge when a single new vertex with m edges is added is $m \times kp_k/2m = \frac{1}{2}kp_k$, independent of m. The num-

Barabási-Albert Model

- Joining of a node results
  - shift in the k degree nodes to (k+1)
  - shift in the (k-1) degree nodes to k



Barabási-Albert Model

ber np_k of vertices with degree k thus decreases by this same amount, since the vertices that get new edges become vertices of degree $k+1$. The number of vertices of degree k also increases because of influx from vertices previously of degree $k-1$ that have also just acquired a new edge, except for vertices of degree m, which have an influx of exactly 1. If we denote by $p_{k,n}$ the value of p_k when the graph has n vertices, then the net change in np_k per vertex added is

Rate Equation or Master equation:
Discrete time framework

$$(n+1)p_{k,n+1} - np_{k,n} = \tfrac{1}{2}(k-1)p_{k-1,n} - \tfrac{1}{2}kp_{k,n}, \quad (66)$$

for $k > m$, or

$$(n+1)p_{m,n+1} - np_{m,n} = 1 - \tfrac{1}{2}mp_{m,n}, \quad (67)$$

for $k = m$, and there are no vertices with $k < m$.

Looking for stationary solutions $p_{k,n+1} = p_{k,n} = p_k$ as before, the equations equivalent to Eq. (61) for the model are

$$p_k = \begin{cases} \tfrac{1}{2}(k-1)p_{k-1} - \tfrac{1}{2}kp_k & \text{for } k > m, \\ 1 - \tfrac{1}{2}mp_m & \text{for } k = m. \end{cases} \quad (68)$$

Rearranging for p_k once again, we find $p_m = 2/(m+2)$ and $p_k = p_{k-1}(k-1)/(k+2)$, or [123, 249]

$$p_k = \frac{(k-1)(k-2)\ldots m}{(k+2)(k+1)\ldots(m+3)}\, p_m = \frac{2m(m+1)}{(k+2)(k+1)k}.$$

Barabási-Albert Model

In the limit of large k this gives a power law degree distribution $p_k \sim k^{-3}$, with only the single fixed exponent $\alpha = 3$. A more rigorous derivation of this result has been given by Bollobás *et al.* [65].

Price's model

- Directed graph
- Citation network

Consider a directed graph of n vertices, such as a citation network. Let p_k be the fraction of vertices in the network with in-degree k, so that $\sum_k p_k = 1$. New vertices are continually added to the network, though not necessarily at a constant rate. Each added vertex has a certain out-degree—the number of papers that it cites—and this out-degree is fixed permanently at the creation of the vertex. The out-degree may vary from one vertex

In the simplest form of cumulative advantage process the probability of attachment of one of our new edges to an old vertex—i.e., the probability that a newly appearing paper cites a previous paper—is simply proportional to the in-degree k of the old vertex. This however immediately gives us a problem, since each vertex starts with in-degree zero, and hence would forever have zero probability of gaining new edges. To circumvent this problem, Price suggests that the probability of attachment to a vertex should be proportional to $k + k_0$, where k_0 is a constant. Although he discusses the case of general k_0, all his mathematical developments are for $k_0 = 1$, which he justifies for the citation network by saying that one can consider the initial publication of a paper to be its first citation (of itself by itself). Thus the probability of a new citation is proportional to $k + 1$.

The probability that a new edge attaches to *any* of the vertices with degree k is thus

$$\frac{(k+1)p_k}{\sum_k (k+1)p_k} = \frac{(k+1)p_k}{m+1}. \tag{58}$$

Rate equations

The mean number of new citations per vertex added is simply m, and hence the mean number of new citations to vertices with current in-degree k is $(k+1)p_k m/(m+1)$.

$$(n+1)p_{k,n+1} - np_{k,n} = [kp_{k-1,n} - (k+1)p_{k,n}]\frac{m}{m+1}, \quad (59)$$

for $k \geq 1$, or

$$(n+1)p_{0,n+1} - np_{0,n} = 1 - p_{0,n}\frac{m}{m+1}, \quad (60)$$

for $k = 0$.

$$p_k = \begin{cases} [kp_{k-1} - (k+1)p_k]m/(m+1) & \text{for } k \geq 1, \\ 1 - p_0 m/(m+1) & \text{for } k = 0. \end{cases} \tag{61}$$

Rearranging, we find $p_0 = (m+1)/(2m+1)$ and $p_k = p_{k-1}k/(k+2+1/m)$ or

$$\mathrm{B}(x,y) = \frac{(x-1)!\,(y-1)!}{(x+y-1)!}$$

$$\begin{aligned} p_k &= \frac{k(k-1)\ldots 1}{(k+2+1/m)\ldots(3+1/m)}\, p_0 \\ &= (1+1/m)\mathrm{B}(k+1, 2+1/m), \end{aligned} \tag{62}$$

where $\mathrm{B}(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ is Legendre's beta-function, which goes asymptotically as a^{-b} for large a and fixed b, and hence

$$p_k \sim k^{-(2+1/m)}. \tag{63}$$

In other words, in the limit of large n, the degree distribution has a power-law tail with exponent $\alpha = 2 + 1/m$.