Network Science

Class 6: Evolving Networks

Albert-László Barabási
With
Emma K. Towlson, Sebastian Ruf, Michael Danziger, and Louis Shekhtman

www.BarabasiLab.com

Questions

1. Bianconi-Barabasi Model
2. Bose-Einstein Condensation
3. Initial attractiveness
4. Role of internal links.
5. Node deletion.
6. Accelerated growth.

Introduction

Section 1

EVOLVING NETWORK MODELS

The BA model is only a minimal model.

Makes the simplest assumptions:

- linear growth $\langle k \rangle = 2m$
- linear preferential attachment $\Pi(k_i) \propto k_i$

Does not capture

variations in the shape of the degree distribution
variations in the degree exponent
the size-independent clustering coefficient

Hypothesis:

The BA model can be adapted to describe most features of real networks.

We need to incorporate mechanisms that are known to take place in real networks: addition of links without new nodes, link rewiring, link removal; node removal, constraints or optimization

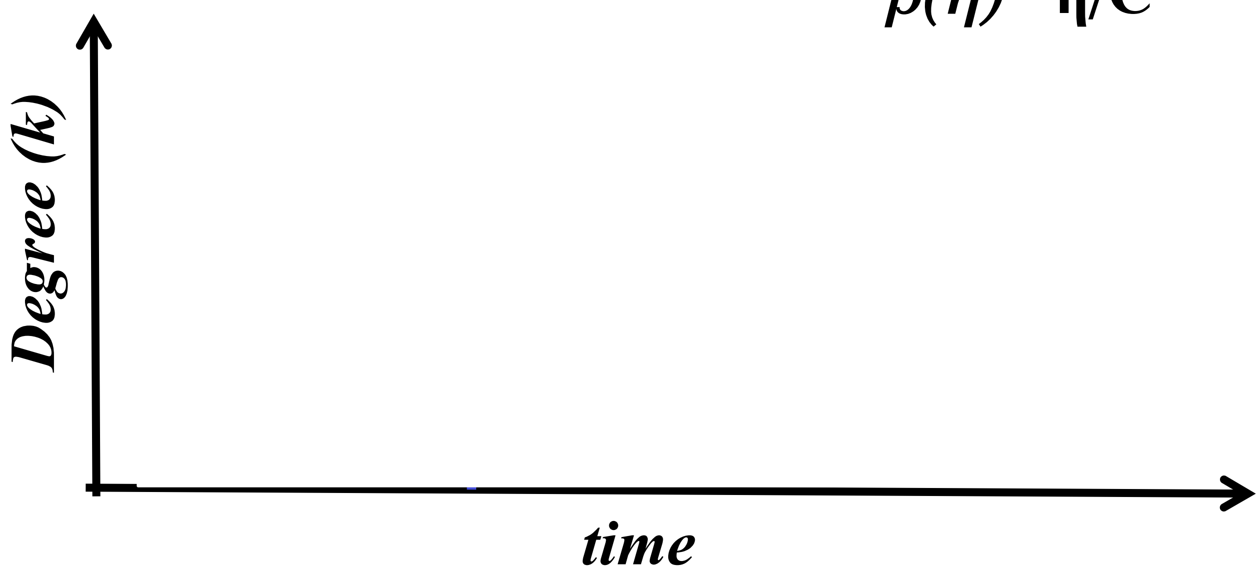
Bianconi-Barabasi model

Can Latecomers Make It?

SF model: $k(t) \sim t^{1/2}$ **(first mover advantage)**

Fitness model: **fitness** (η) $\Pi(k_i) \cong \dfrac{\eta_i k_i}{\sum_j \eta_j k_j}$ $k(\eta,t) \sim t^{\beta(\eta)}$

$\beta(\eta) = \eta / C$



Bianconi & Barabási, *Physical Review Letters* 2001; *Europhys. Lett.* 2001.

- **Growth**

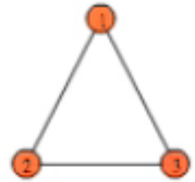
  In each timestep a new node j with m links and fitness η_j is added to the network, where η_j is a random number chosen from a *fitness distribution* $\rho(\eta)$. Once assigned, a node's fitness does not change.

- **Preferential Attachment**

  The probability that a link of a new node connects to node i is proportional to the product of node i's degree k_i and its fitness η_i,

$$\Pi_i = \frac{\eta_i k_i}{\sum_j \eta_j k_j}. \qquad (6.1)$$

Section 2 | Fitness Model



$$\frac{\partial k_i}{\partial t} = m \frac{\eta_i k_i}{\sum_k \eta_j k_j}$$

$$k_{\eta_i}(t, t_i) = m \left(\frac{t}{t_i}\right)^{\beta(\eta_i)}.$$

$$\left\langle \sum_j \eta_j k_j \right\rangle$$

over all possible realizations of the quenched fitnesses η. Since each node is born at a different time t_0, we can write the sum over j as an integral over t_0

$$\left\langle \sum_j \eta_j k_j \right\rangle = \int d\eta \rho(\eta) \eta \int_1^t dt_0 k_\eta(t, t_0). \quad (6.34)$$

By replacing $k_\eta(t, t_0)$ with (6.3) and performing the integral over t_0, we obtain

$$\left\langle \sum_j \eta_j k_j \right\rangle = \int d\eta \rho(\eta) \eta m \frac{t - t^{\beta(\eta)}}{1 - \beta(\eta)}. \quad (6.35)$$

The dynamic exponent $\beta(\eta)$ is bounded, i.e. $0<\beta(\eta)<1$, because a node can only increase its degree with time ($\beta(\eta)>0$) and $k_i(t)$ cannot increase faster than t ($\beta(\eta)<1$). Therefore in the limit $t\to\infty$ in (6.35) the term $t^{\beta(\eta)}$ can be neglected compared to t, obtaining

$$\left\langle \sum_j \eta_j k_j \right\rangle \overset{t\to\infty}{=} Cmt(1 - O(t^{-\varepsilon})), \quad (6.36)$$

where $\varepsilon = (1 - \max_\eta \beta(\eta)) > 0$ and

$$C = \int d\eta \rho(\eta) \frac{\eta}{1 - \beta(\eta)}. \quad (6.37)$$

$$\frac{\partial k_i}{\partial t} = m \frac{\eta_i k_i}{\sum_k \eta_j k_j}$$

$$C = \int d\eta \rho(\eta) \frac{\eta}{1-\beta(\eta)}$$

$$\frac{\partial k_\eta}{\partial t} = \frac{\eta k_\eta}{Ct},$$

$$k_{\eta_i}(t, t_i) = m\left(\frac{t}{t_i}\right)^{\beta(\eta_i)}.$$

$$\beta(\eta) = \frac{\eta}{C},$$

To complete the calculation we need to determine C from (6.37). After substituting $\beta(n)$ with η/C , we obtain

$$1 = \int_0^{\eta_{\max}} d\eta \rho(\eta) \frac{1}{\frac{C}{\eta} - 1}, \tag{6.40}$$

where η_{max} is the maximum possible fitness in the system. The integral (6.40) is singular. However, since $\beta(\eta)=\eta/C< 1$ for any η, we have $C > \eta_{max}$, thus the integration limit never reaches the singularity. Note also that since

$$\sum_j \eta_j k_j \leq \eta_{\max} \sum_j k_j = 2mt\eta_{\max} \tag{6.41}$$

we have $C \leq \eta_{max}$.

Section 2 Fitness Model

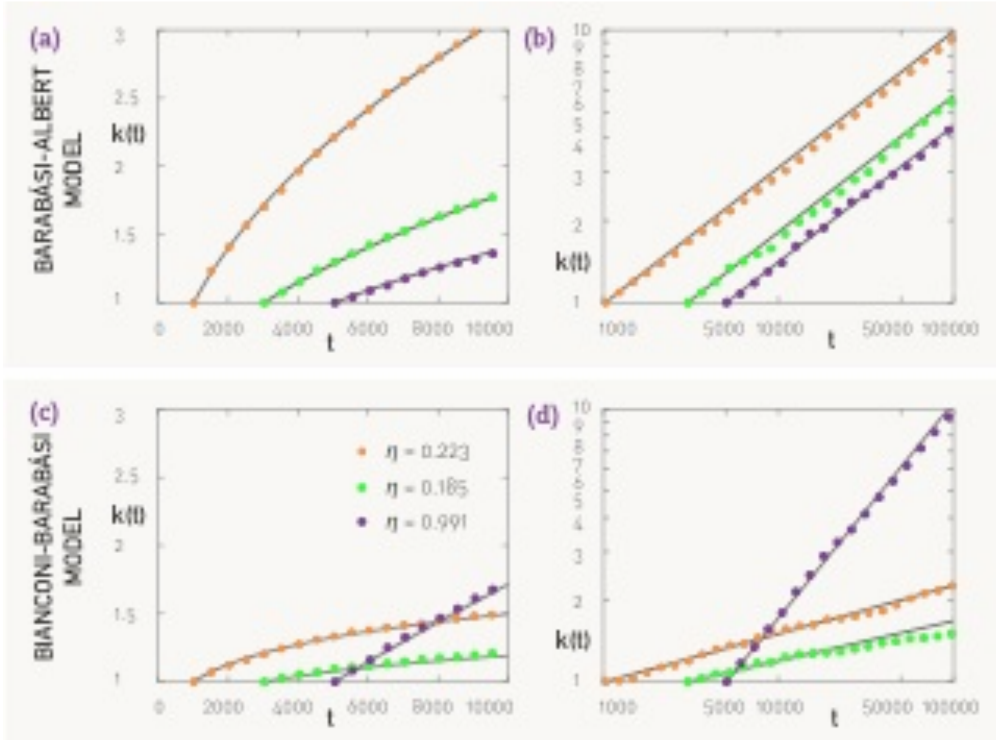
BA model: $k(t) \sim t^{\frac{1}{2}}$

(first mover advantage)

BB model: $k(\eta,t) \sim t^{\beta(\eta)}$

(fit-gets-richer)

$\beta(\eta) = \eta/C$

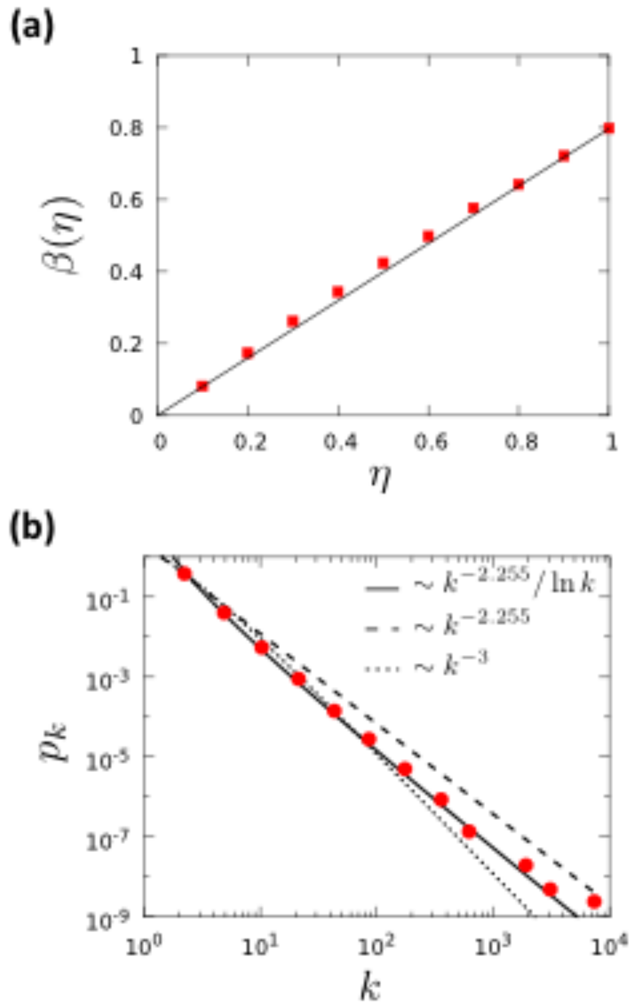


$$p_k \sim C \int dy \frac{\rho(\eta)}{\eta} \left(\frac{m}{k} \right)^{\frac{C}{\eta}+1} .$$

Uniform fitness distribution:
fitness uniformly distributed in the [0,1] interval.

$$p_k \sim \int_1^0 d\eta \frac{C^*}{\eta} \frac{1}{k^{1+C^*/\eta}} \sim \frac{k^{-(1+C^*)}}{\ln k},$$

C = 1.255* $\gamma = 2.255$



$$k_{\eta_i}(t,t_i) = m\left(\frac{t}{t_i}\right)^{\beta(\eta_i)}.$$

$$\beta(\eta) = \frac{\eta}{C},$$

$$p_k \sim C\int d\eta \frac{\rho(\eta)}{\eta}\left(\frac{m}{k}\right)^{\frac{C}{\eta}+1}.$$

If there is a single dynamic exponent β, the degree distribution follows the power law $p_k \sim k^{-\gamma}$ with degree exponent γ=1/β+1. In the Bianconi-Barabási model we have a spectrum of dynamic exponents $\beta(\eta)$, thus p_k is a weighted sum over different power-laws. To calculate p_k we need to determine the cumulative probability that a randomly chosen node's degree satisfies $k_\eta(t)>k$. This cumulative probability is

$$P(k_\eta(t) > k) = P\left(t_0 < t\left(\frac{m}{k}\right)^{C/\eta}\right) = t\left(\frac{m}{k}\right)^{C/\eta}. \tag{6.42}$$

Thus, the degree distribution is given by the integral

$$p_k = \int_{\eta_{max}}^{0} d\eta \frac{\partial P(k_\eta(t) > k)}{\partial t} \propto \int d\eta \rho(\eta)\frac{C}{\eta}\left(\frac{m}{k}\right)^{\frac{C}{\eta}+1}, \tag{6.43}$$

- **Equal Fitnesses**
When all fitnesses are equal, the Bianconi-Barabási model reduces to the Barabási-Albert model. Indeed, let us use $\rho(\eta) = \delta(\eta - 1)$, capturing the fact that each node has the same fitness $\eta = 1$. In this case (6.5) yields $C = 2$. Using (6.4) we obtain $\beta = 1/2$ and (6.6) predicts $p_k \sim k^{-3}$, the known scaling of the degree distribution in the Barabási-Albert model.

$$C = \int d\eta \rho(\eta) \frac{\eta}{1-\beta(\eta)}$$

$$\beta(\eta) = \frac{\eta}{C},$$

$$p_k \sim C \int d\eta \frac{\rho(\eta)}{\eta} \left(\frac{m}{k}\right)^{\frac{C}{\eta}+1}.$$

Section 6.2 Uniform Fitnesses

- **Uniform Fitness Distribution**

  The model's behavior is more interesting when nodes have different fitnesses. Let us choose η to be uniformly distributed in the [0,1] interval. In this case C is the solution of the transcendental equation (6.5)

$$C = \int d\eta \rho(\eta) \frac{\eta}{1-\beta(\eta)}$$

$$\exp(-2/C) = 1 - 1/C\,, \tag{6.7}$$

whose numerical solution is $C^* = 1.255$. Consequently, (6.4) predicts that each node i has a different dynamic exponent, $\beta(\eta_i) = \eta_i / C^*$.

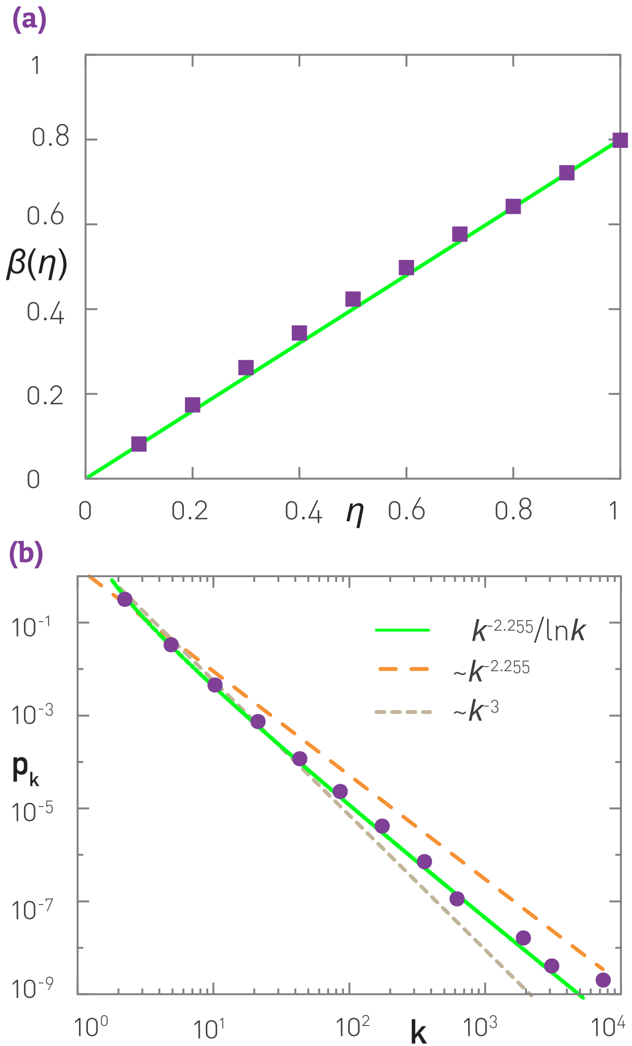
$$\beta(\eta) = \frac{\eta}{C}\,,$$

Using (6.6) we obtain

$$p_k \sim C \int dy \frac{\rho(\eta)}{\eta} \left(\frac{m}{k}\right)^{\frac{C}{\eta}+1}.$$

$$p_k \sim \int_0^1 d\eta \frac{C^*}{\eta} \frac{1}{k^{1+C^*/\eta}} \sim \frac{k^{-(1+C^*)}}{\ln k}\,, \tag{6.8}$$

predicting that the degree distribution follows a power law with degree exponent $\gamma = 2.255$. Yet, we do not expect a perfect power law, but the scaling is affected by an inverse logarithmic correction $1/\ln k$.



$$\beta(\eta_i) = \eta_i / C^*.$$

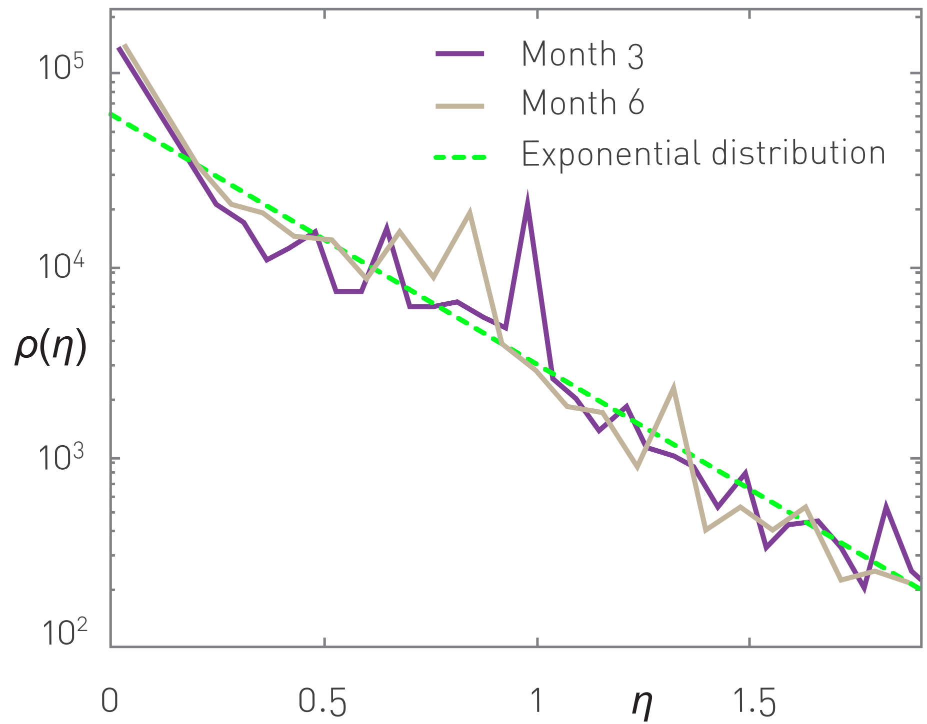
$C^* = 1.255.$

$$p_k \sim \int_0^1 d\eta \frac{C^*}{\eta} \frac{1}{k^{1+C^*/\eta}} \sim \frac{k^{-(1+C^*)}}{\ln k},$$

Measuring Fitness

$$k_{\eta_i}(t,t_i) = m\left(\frac{t}{t_i}\right)^{\beta(\eta_i)}.$$

$$\log k_{\eta_i}(t,t_i) = \beta(\eta_i)\log t + B_i,$$

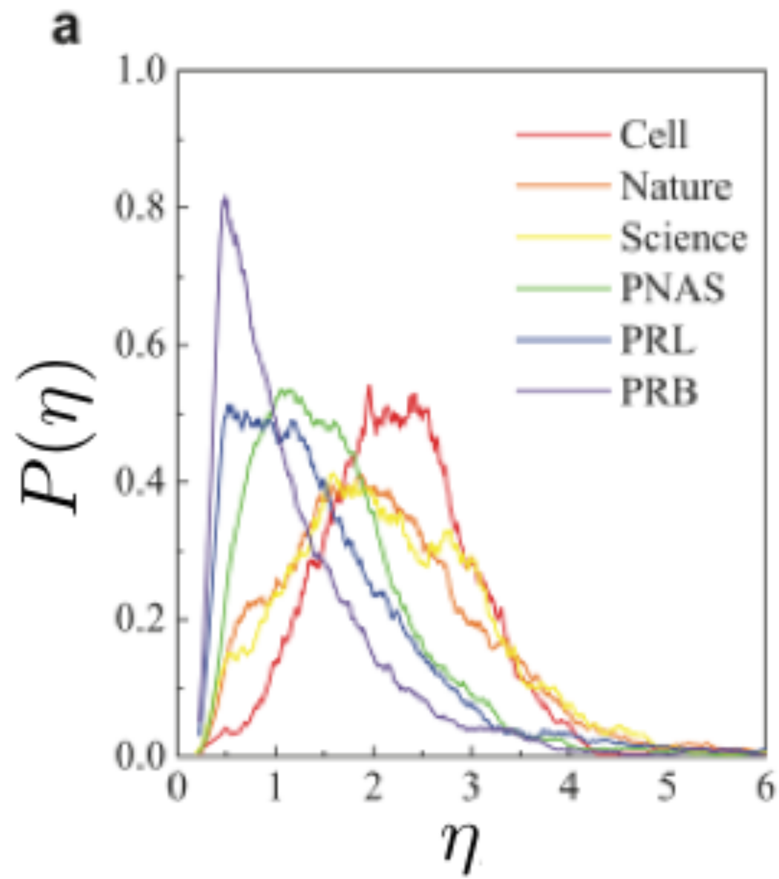


$$\Pi_i \sim \eta_i c_i^t P_i(t),$$

$$P_i(t) = \frac{1}{\sqrt{2\pi}\sigma_i t} e^{-\frac{(\ln t - \mu_i)^2}{2\sigma_i^2}}.$$

$$c_i^t = m\left(e^{\frac{\beta\eta_i}{A}\Phi\left(\frac{\ln t - \mu_i}{\sigma_i}\right)} \right),$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2} dy$$



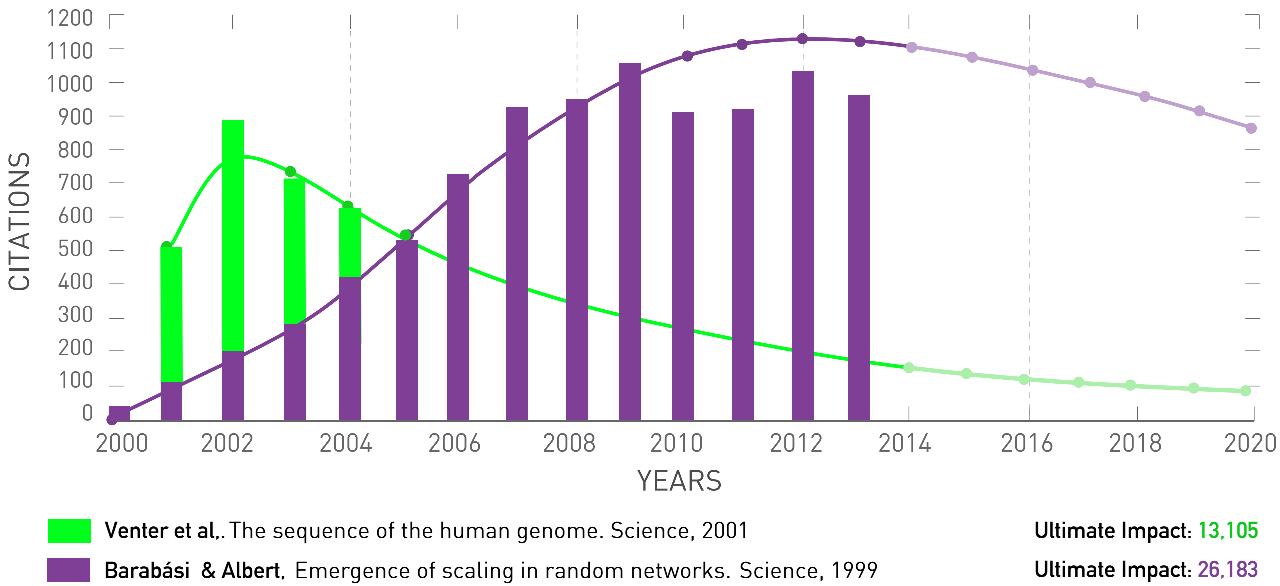
Section 3 — The Fitness of a scientific publication

$$c_i^t = m\left(e^{\frac{\beta\eta_i}{A}\Phi\left(\frac{\ln t-\mu_i}{\sigma_i}\right)}\right),$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-y^2/2}dy$$

Ultimate Impact: t → ∞

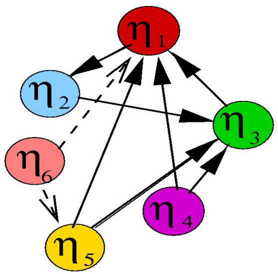
$$c_i^{\infty} = m(e^{\eta_i}-1)$$



Bose-Einstein condensation

MAPPING TO A QUANTUM GAS

Network



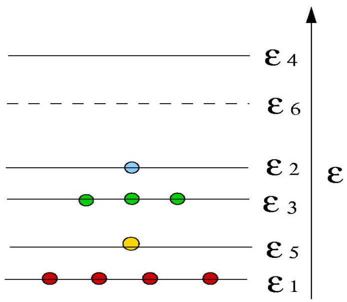
$$\Pi_i = \frac{\eta_i k_i}{\Sigma_j \eta_j k_j}$$

$$\eta \longrightarrow e^{-\beta\varepsilon}$$

$$k_{in}(\eta) \longrightarrow n(\varepsilon)$$

$$\rho(\eta) \longrightarrow g(\varepsilon)$$

Bose gas



Fitness η → Energy level ε

New node with fitness η → New energy level ε

Link pointing to node η → Particle at level ε

Network → quantum gas

G. Bianconi and A.-L. Barabási, Physical Review Letters 2001; cond-mat/0011029

BOSE-EINSTEIN CONDENSATION

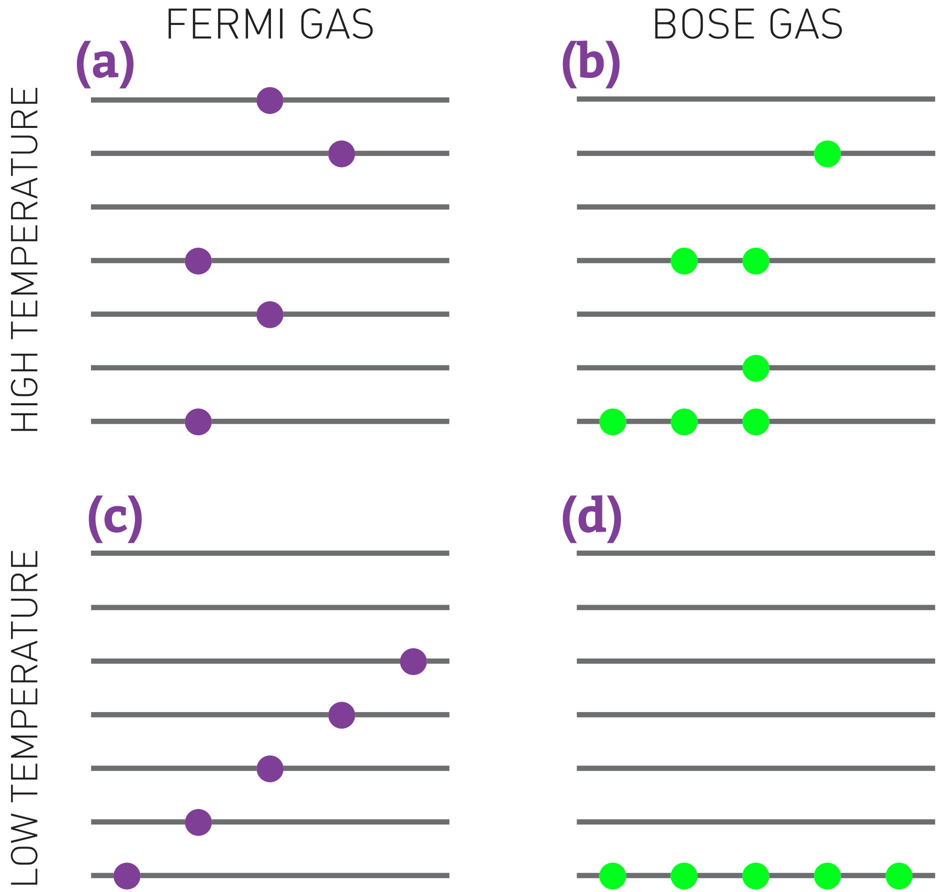
$$\frac{\partial k_i(t, t_i, \varepsilon_i)}{\partial t} = m \frac{e^{-\beta \varepsilon_i} k_i(t, t_i, \varepsilon_i)}{\sum_j e^{-\beta \varepsilon_j} k_j(t, t_j, \varepsilon_j)}.$$

$$k(t, t_i, \varepsilon_i) = m \left(\frac{t}{t_i} \right)^{f(\varepsilon_i)} \quad f(\varepsilon) = e^{-\beta(\varepsilon - \mu)}.$$

The dynamic exponent *f(e)* depends on **m,** determined by the self-consistent equation:

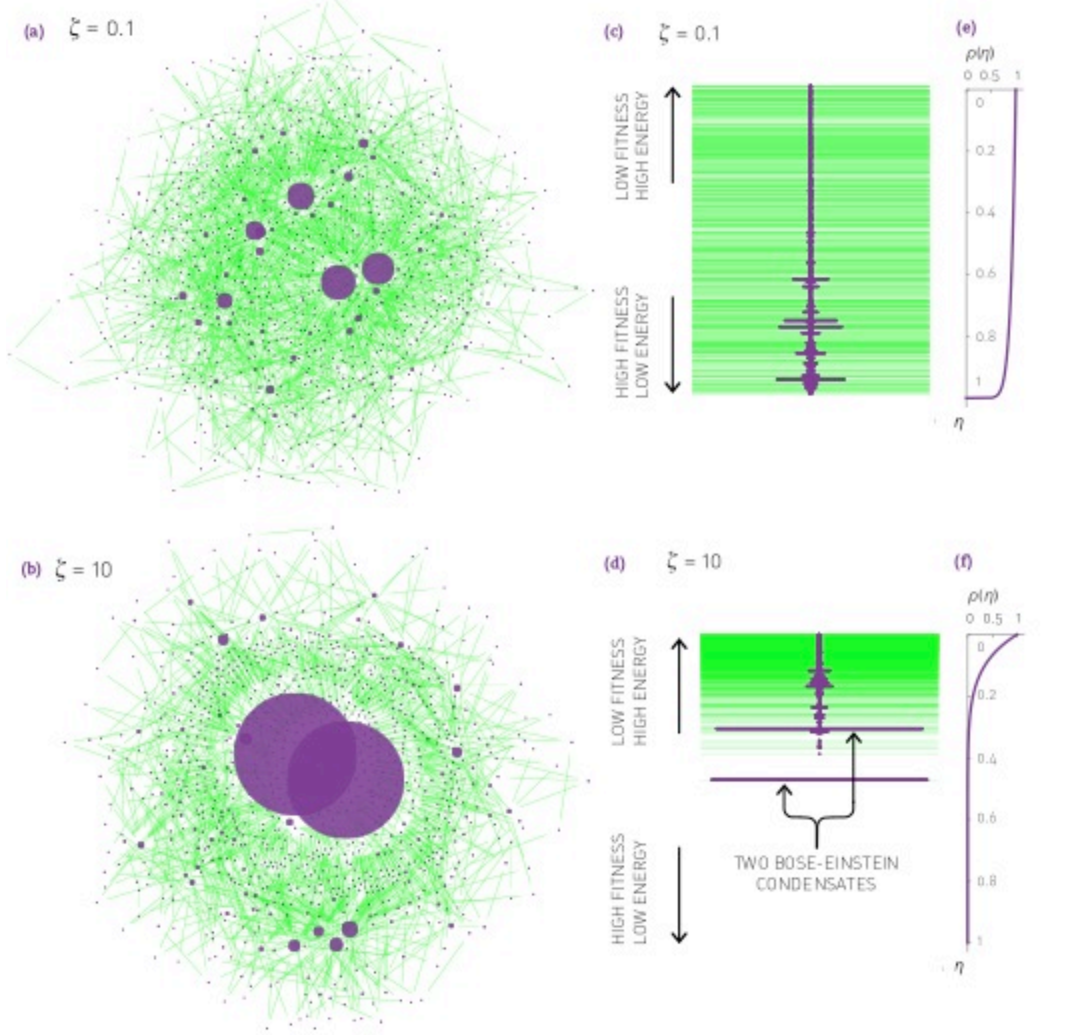
$$I(\beta, \mu) = \int d\varepsilon \, p(\varepsilon) \frac{1}{e^{\beta(\varepsilon - \mu)} - 1} = 1. \qquad \int d\varepsilon g(\varepsilon) n(\varepsilon) = 1$$

$$\boxed{n(\varepsilon) = \frac{1}{e^{\beta(\varepsilon - \mu)} - 1}}$$

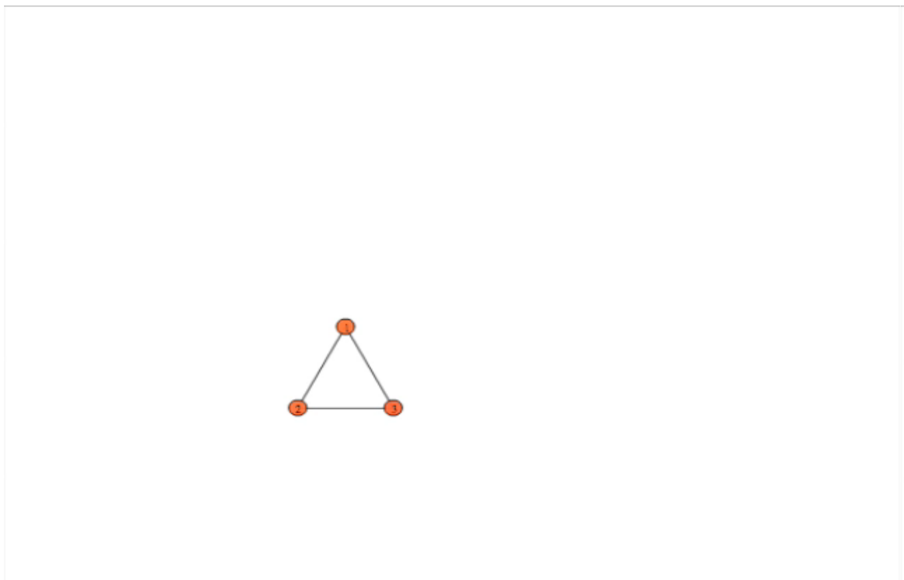

FERMI GAS
BOSE GAS
(a)
(b)
(c)
(d)
HIGH TEMPERATURE
LOW TEMPERATURE

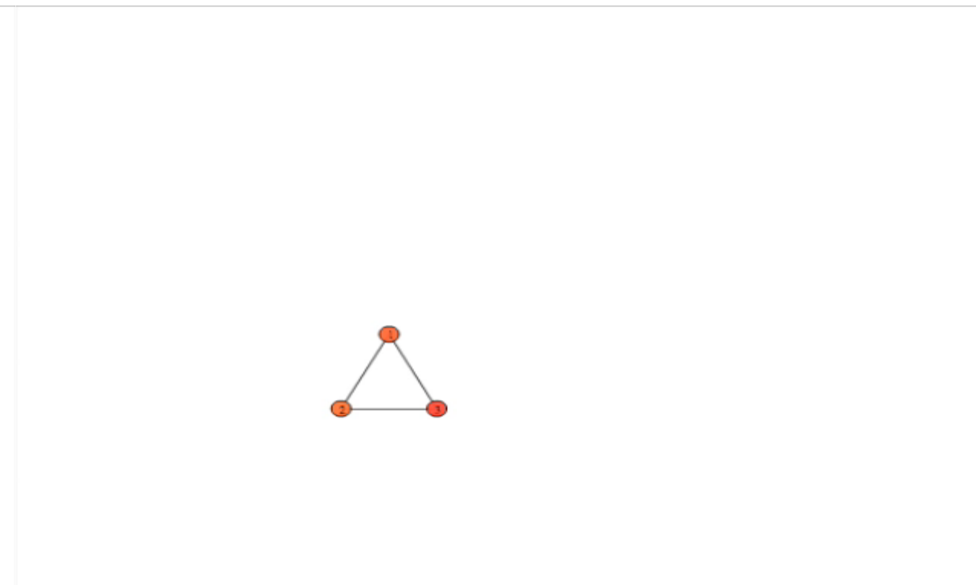
Section 4 Bose-Einstein Condensation

$$\rho(\eta) = (1 - \eta)^{\zeta}$$



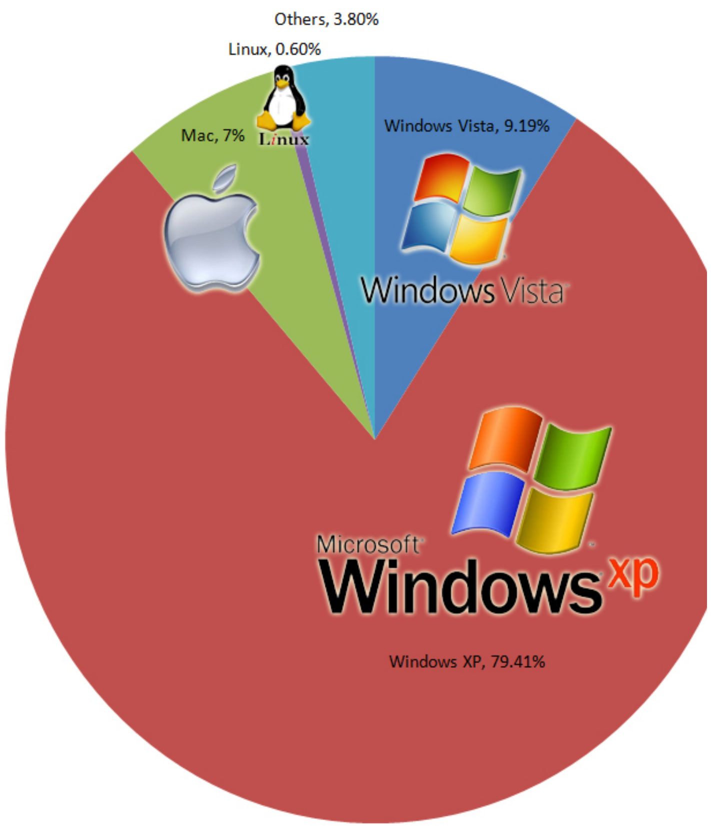
Bose-Einstein Condensation

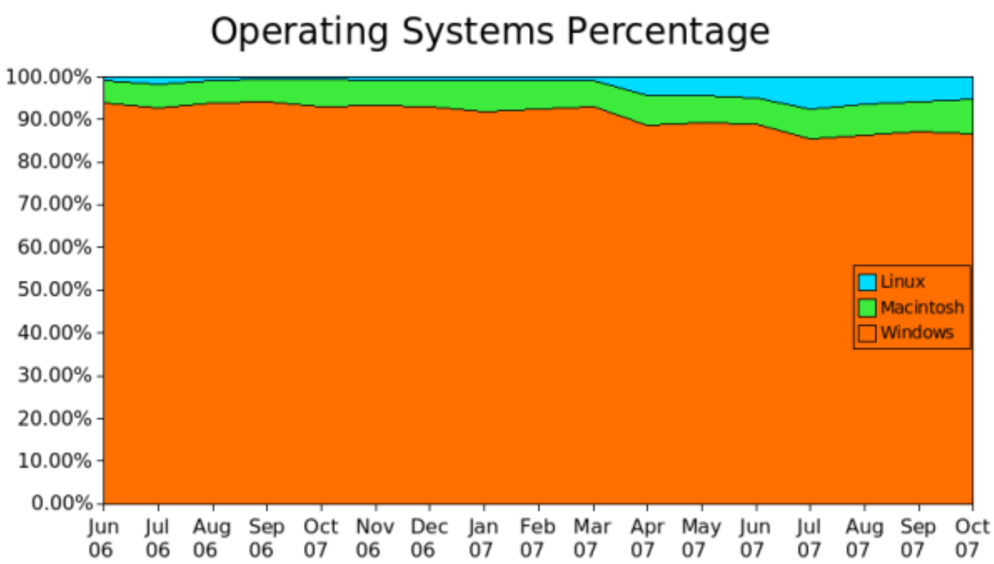




Bianconi & Barabási, *Physical Review Letters* 2001; *Europhys. Lett.* 2001.

FITNESS MODEL: Bose-Einstein Condensation





Bianconi & Barabási, *Physical Review Letters* 2001; *Europhys. Lett.* 2001.

Evolving Networks

Limitations

(i) The model predicts $\gamma = 3$, while the experimentally observed degree exponents vary between 2 and 5 (**Table 4.1**).

(ii) The model predicts a power-law degree distribution, while in real systems we observe systematic deviations from a pure power-law function, like small-degree saturation or high-degree cutoff (**BOX 4.8**).

(iii) The model ignores a number of elementary processes that are obviously present in many real networks, like the addition of internal links and node or link removal.

INITIAL ATTRACTIVENESS

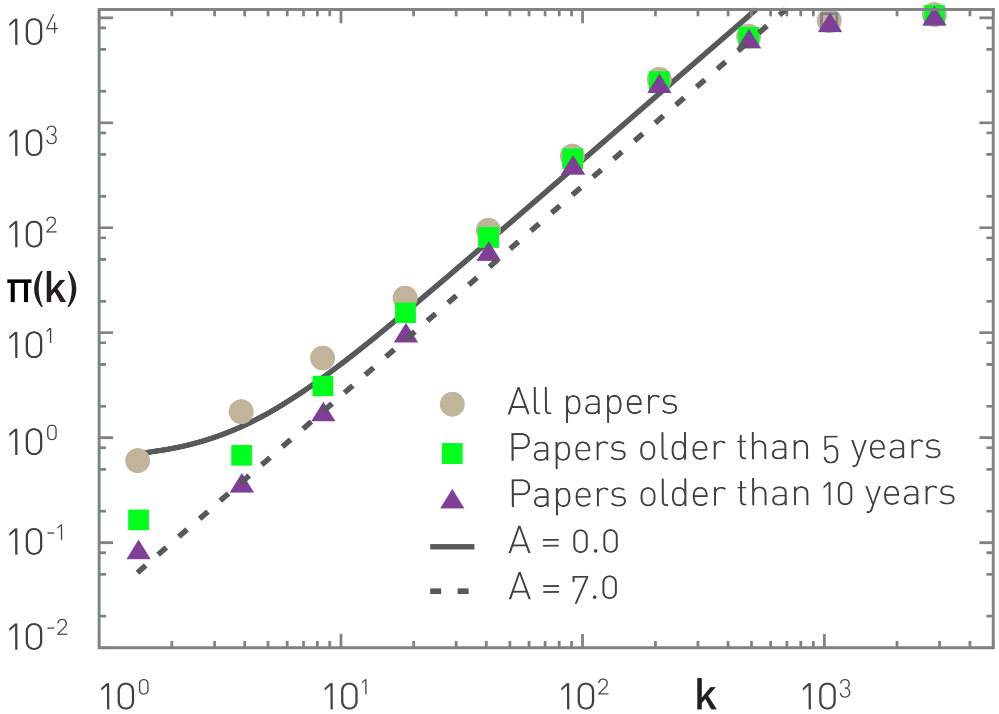
$$\Pi(k) \sim A + k$$

Increases the degree exponent.

$$\gamma = 3 + \frac{A}{m}$$

Generates a small-degree cutoff.

$$p_k = C(k + A)^{-\gamma}$$



INTERNAL LINKS

$$\Pi(k,k')\sim(A+Bk)(A'+B'k').$$

Double preferential attachment (A=A'=0).

$$\gamma = 2 + \frac{m}{m+2n}$$

Random attachment (B=B'=0).

$$\gamma = 3 + \frac{2n}{m}$$

NODE DELETION

r<1: *Scale-free phase*

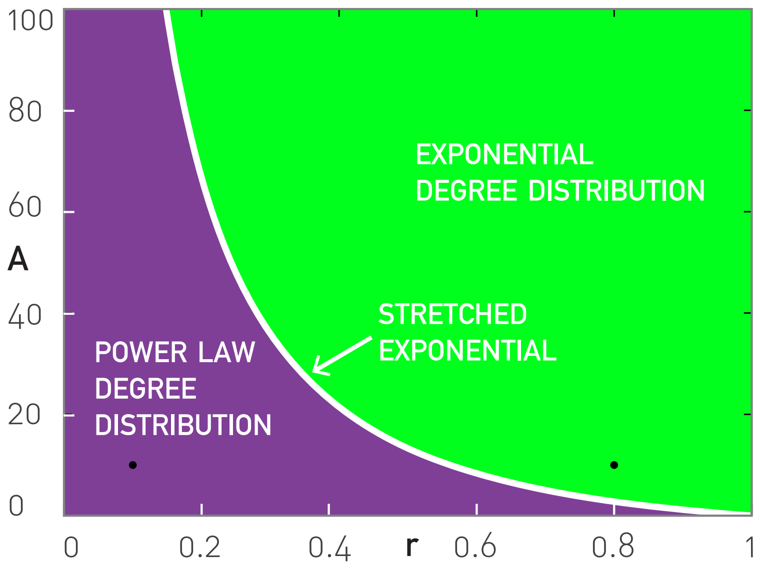
$$\gamma = 3 + \frac{2r}{1-r}$$

- Start with the Barabási-Albert model.
- In each time step:
    - add a new node with *m links*
    - with probability *r* remove a node.

***r*=1**: *Exponential phase*

***r*>1:** *Declining network*

NODE DELETION



- Start with the Initial Attractiveness model:

  $$\Pi(k) \sim A + k$$

- In each time step:
  - add a new node with *m links*
  - with probability *r* remove a node.

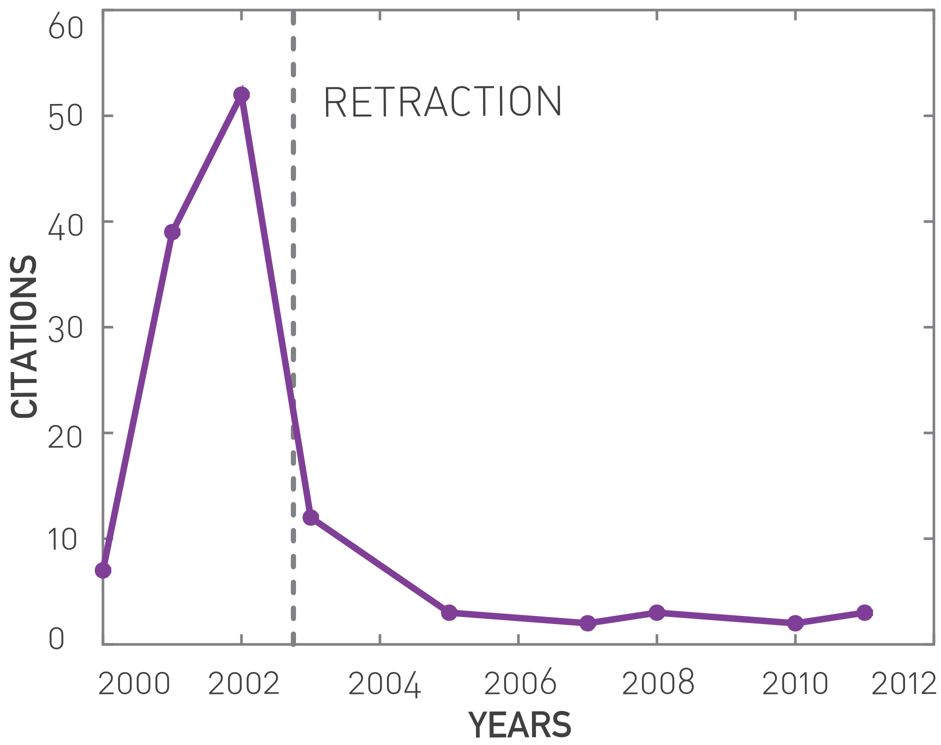
The coexistence of node removal with other elementary processes can lead to interesting topological phase transitions. This is illustrated by a simple model in which the network's growth is governed by **(6.23)**, and we also remove nodes with rate *r* [30]. The network displays three distinct phases, captured by the phase diagram shown above, whose axes are the node removal rate *r* and initial attractiveness *A*:

Subcritical Node Removal: $r < r^*(A)$
If the rate of node removal is under a critical value $r^*(A)$, shown as the white line on the figure, the network will be scale-free.

Critical Node Removal: $r=r^*(A)$
Once r reaches a critical value $r^*(A)$, the degree distribution turns into a stretched exponential (**SECTION 4.A**).

Exponential Networks: $r> r^*(A)$
The network looses its scale-free nature, developing an exponential degree distribution.

The Impossibility of Node deletion



[23] J.H. Schön, Ch. Kloc, R.C. Haddon, and B. Batlogg. A superconducting field-effect switch. Science, 288: 656–8. 2000.

Jan Hendrik Schön

Declining Fashion: New York

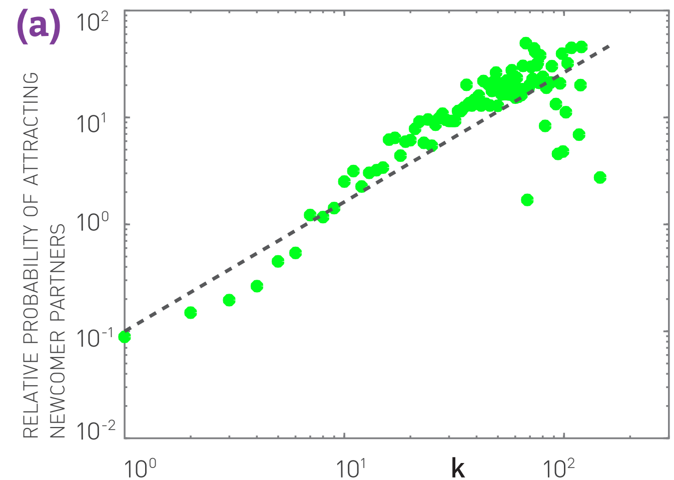
Declining Fashion

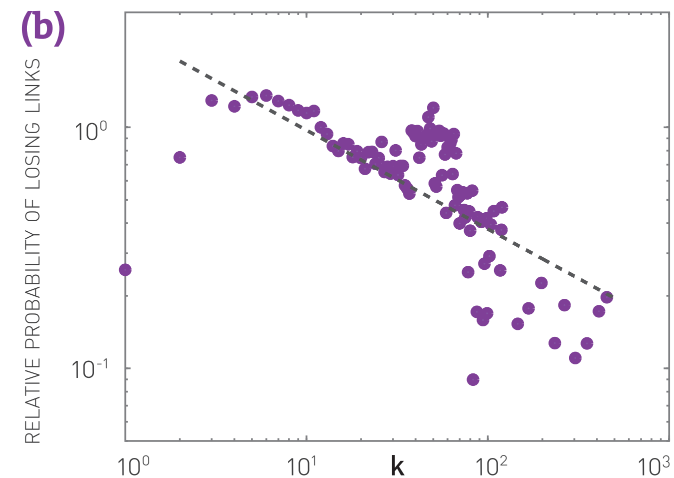
- **Preferential Attachment**

  While overall the network was shrinking, new nodes continued to arrive. The measurements indicate that the attachment probability of these new nodes follows $\Pi(k) \sim k^{\alpha}$ with α=1.20 ± 0.06 (**Figure 6.13a**), offering evidence of superlinear preferential attachment (SECTION 5.7).

- **Link Deletion**

  The probability that a firm lost a link follows $k(t)^{-\eta}$ with η = 0.41 ± 0.04, i.e. it decreased with the firms' degree (**Figure 6.13b**). This documents a *weak-gets-weaker* phenomenon, when the less connected firms are more likely to loose links.





Accelerated growth

we assumed that $L = \langle k \rangle N$, *where* $\langle k \rangle$ *is independent of time or N.*

•the average degree of the Internet increased from 3.42 (Nov. 1997) to 3.96 (Dec. 1998);

$$m(t) = m_0 t^{\theta}$$

•the WWW increased its average degree from 7.22 to 7.86 during five months;

$$\gamma = 3 + \frac{2\theta}{1-\theta}$$

•in metabolic networks the average degree of the metabolites grows approximately linearly with the number of metabolites [33].

$$\Pi(k, t - t_i) \sim k(t - t_i)^{-\nu}$$

v<0: *new nodes attach to older nodes → enhances the role of preferential attachment.*

v⟶ −∞ each new node will only connect to the oldest node → hub-and-spoke topology (Fig 6.10a).

v<0: *new nodes attach to younger nodes*

v⟶ +∞: each node will connect to its immediate predecessor (Fig. 6.10a).

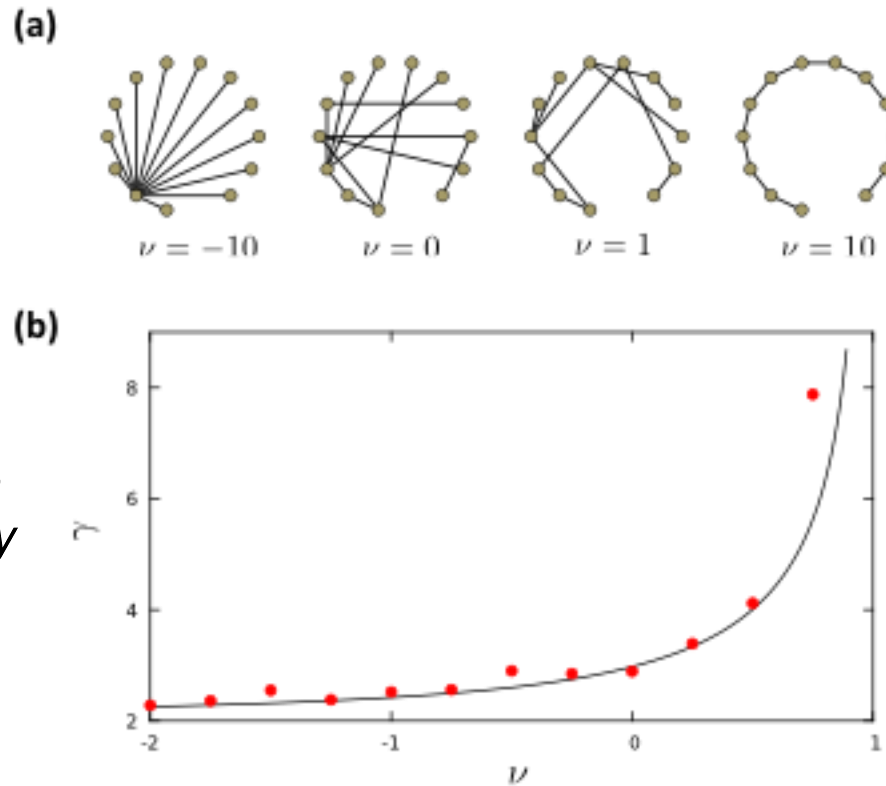
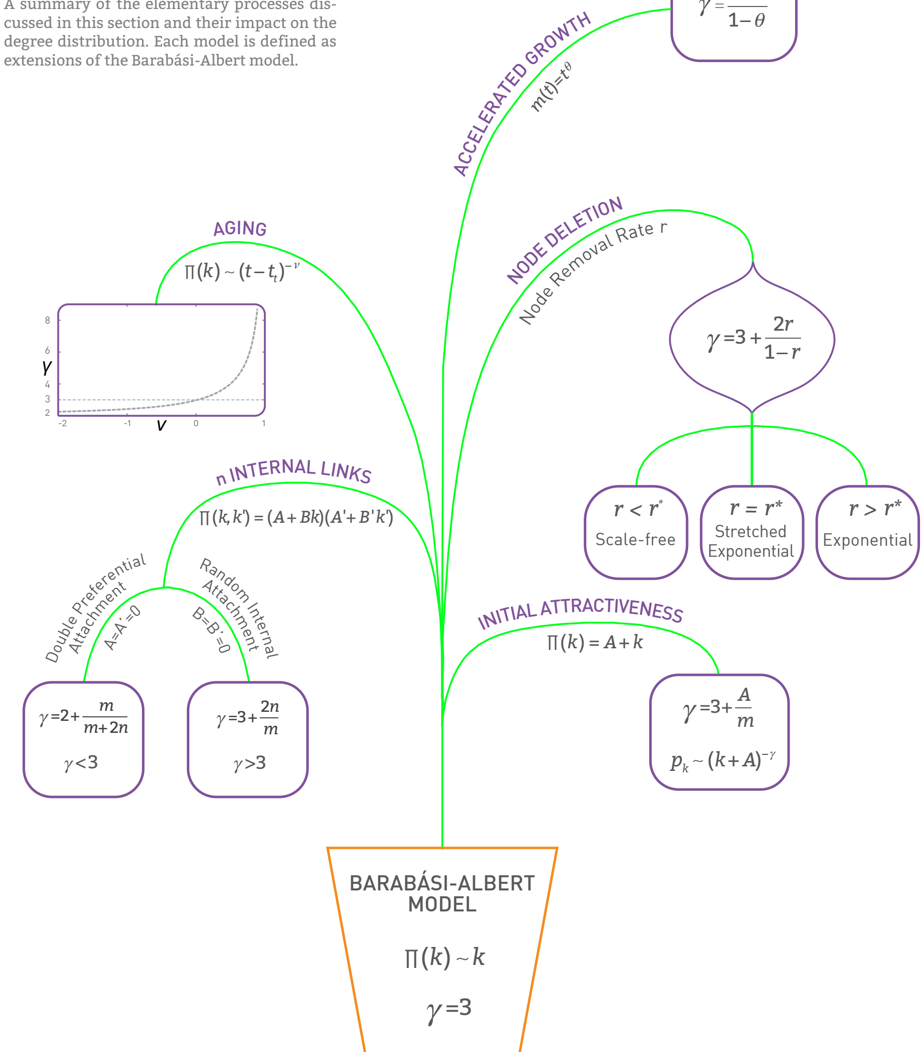


FIG. 6.15

Elementary Processes Affecting the Network Topology

A summary of the elementary processes discussed in this section and their impact on the degree distribution. Each model is defined as extensions of the Barabási-Albert model.



Summary

- **Power-Law**

  A pure power-law emerges if a growing network is governed by linear preferential attachment only, as predicted by the Barabási-Albert model. It is rare to observe such a pure power law in real systems. This idealized model represents the starting point for understanding the degree distribution of real networks.

- **Stretched Exponential**

  If preferential attachment is sublinear, the degree distribution follows a stretched exponential (SECTION 5.7). A similar degree-distribution can also appear under node removal at the critical point (Figure 6.12).

- **Fitness-induced Corrections**
  In the presence of fitness the precise form of p_k depends on the fitness distribution $\rho(\eta)$, which determines p_k via (6.6). For example, a uniform fitness distribution induces a logarithmic correction in p_k as predicted by (6.8). Other forms of $\rho(\eta)$ can lead to rather exotic forms for p_k.

- **Small-degree Saturation**
  Initial attractiveness adds a random component to preferential attachment. Consequently, the degree distribution develops a small-degree saturation, as seen in (6.24).

- **High-degree Cutoffs**
  Node and link removal, present in many real systems, can induce exponential high-degree cutoffs in the degree distribution. Furthermore, random node-removal can deplete the small-degree nodes, inducing a peak in p_k.

In most real networks several of the elementary processes discussed in this chapter appear together. For example, in the scientific collaboration network we have sublinear preferential attachment with initial attractiveness and the links can be both external and internal. As researchers have different creativity, fitness also plays a role, hence an accurate model requires us to know the appropriate fitness distribution. Therefore, the degree distribution is expected to display small degree saturation (thanks to initial attractiveness), stretched exponential cutoff at high degrees (thanks to sublinear preferential attachment), and some unknown corrections due to the particular form of the fitness distribution $\rho(\eta)$.

In general if wish to obtain an accurate fit to the degree distribution, we first need to build a generative model that analytically predicts the functional form of p_k. Yet, in many systems developing an accurate theory for p_k may be an overkill. It is often sufficient, instead, to establish if we are dealing with a bounded or an unbounded degree distribution (SECTION 4.9), as the system's properties will be primarily driven by this distinction.

Section 6 summary

MODEL CLASS	EXAMPLES	CHARACTERISTICS
Static Models	Erdős–Rényi Watts-Strogatz	• N fixed • p_k bounded • Static, time independent topologies
Generative Models	Configuration Model Hidden Parameter Model	• Arbitrary pre-defined p_k • Static, time independent topologies
Evolving Network Models	Barabási–Albert Model Bianconi-Barabási Model Initial Attractiveness Model Internal Links Model Node Deletion Model Accelerated Growth Model Aging Model	• p_k is determined by the processes that contribute to the network's evolution. •Time-varying network topologies

LESSONS LEARNED: evolving network models

1. There is no universal exponent characterizing all networks.

2. Growth and preferential attachment are responsible for the emergence of the scale-free property.
3. The origins of the preferential attachment is system-dependent.
4. Modeling real networks:
    - identify the microscopic processes that take place in the system
    - measure their frequency from real data
    - develop dynamical models that capture these processes.

5. If the model is correct, it should correctly predict not only the degree exponent, but both small and large k-cutoffs.

The end