

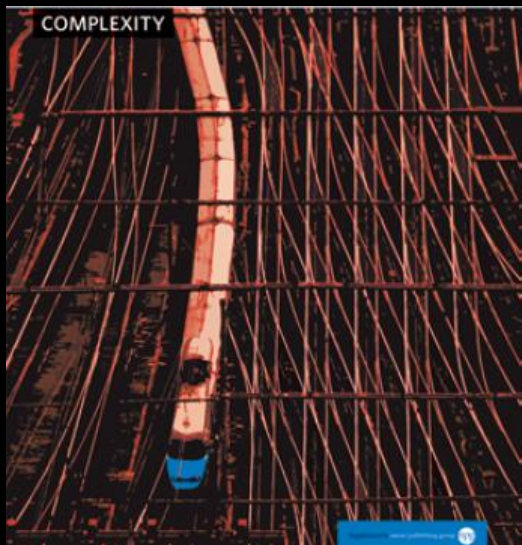
# Characterizing and modeling citation dynamics

**Young-Ho Eom**

Institute for Scientific Interchange

Turin, Italy

predate comprehensive traditional game  
 coupled represent many behaviors references  
 insect concrete relatively probably dynamical examples agent-based ecosystems simulations devoted  
 precise foundation flocking new understand foundational form Monte economy involves scientists materials  
 course Experimental individual Internet ecosystems form Monte economy involves scientists materials  
 solvable usually different studies markets particularly emergent include system  
 observations disciplines basic serve computational Economy complexity economies watch name  
 appear simplified may called sum immune automata simulation theory theoretical Emergent colonies fish  
 grown increasing techniques parts industry elements simulation theory theoretical Emergent colonies fish  
 mimic techniques parts industry elements simulation theory theoretical Emergent colonies fish  
 create Carlo composed cellular tools also collective try area computer realistic qualitative approaches software  
 condensed review community Complex divided advances modeling approach field surge traffic interest agree behavior often focuses realm  
 exactly schooling Unfortunately makes scientific interdisciplinary field surge traffic interest agree behavior often focuses realm  
 road clear created numerical primarily brain part  
 sophisticated granular birds models study Society  
 current Collective networks possible technical entire science including although substantial academia gain financial thus

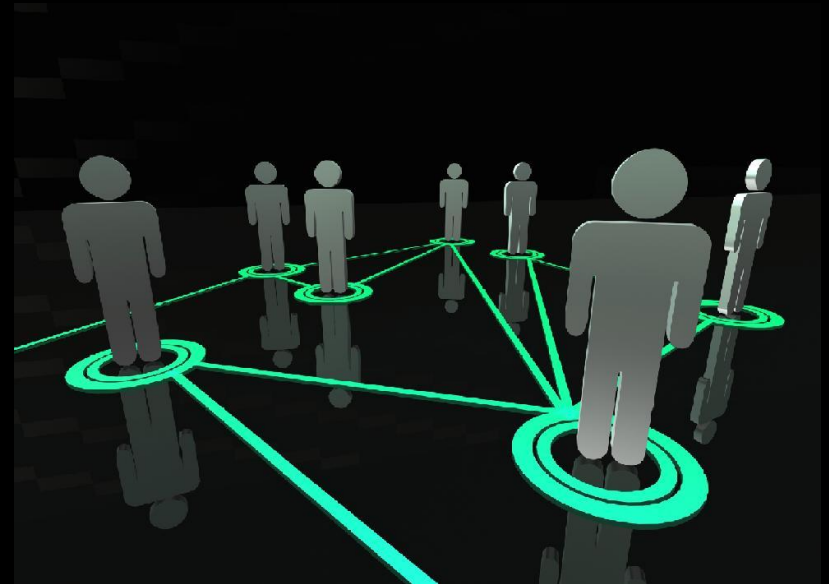


# Complex networks

The parts → Nodes



The interaction → Links



System-level perspective

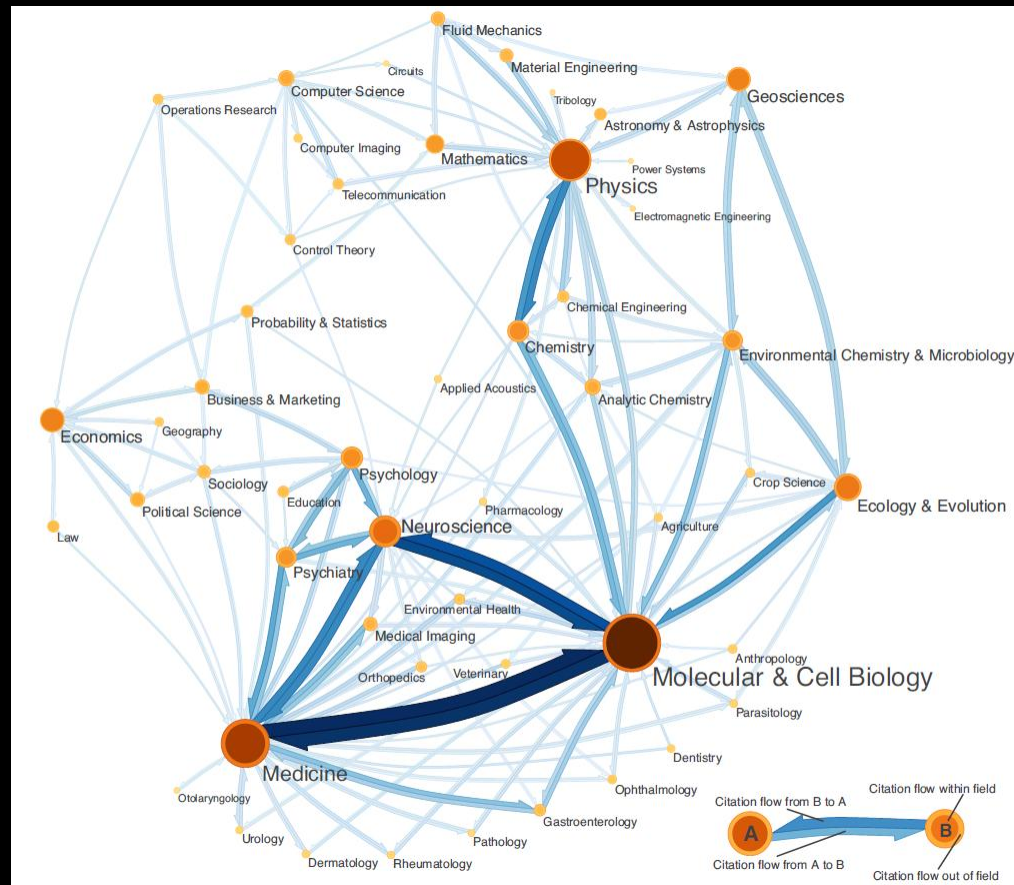


# Citation Dynamics

Evolution of science

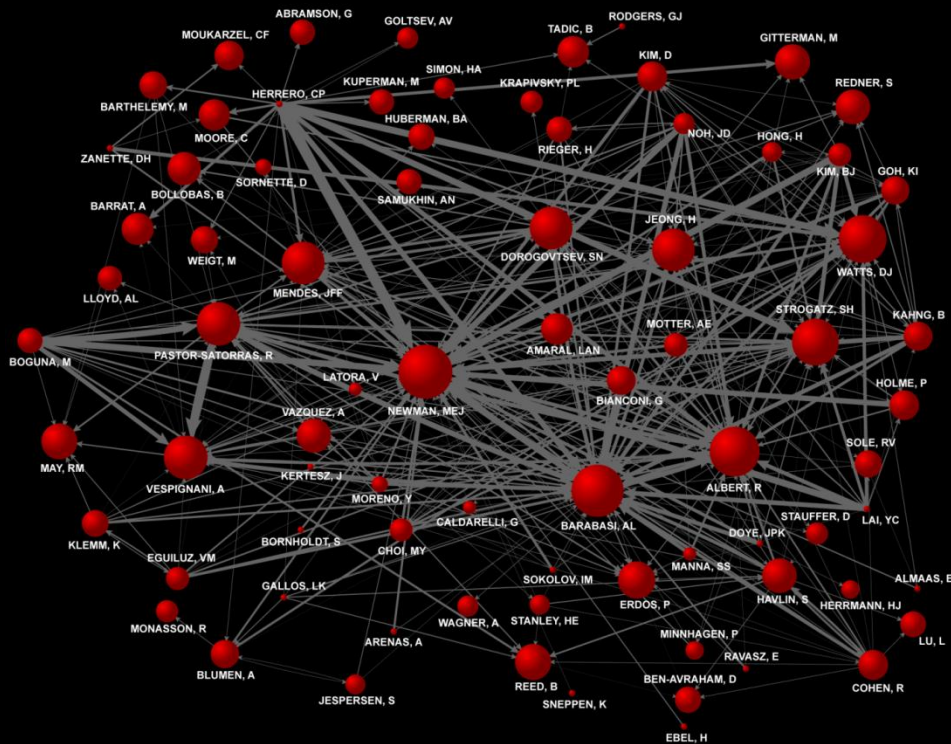
Spreading of scientific knowledge

Relationship between papers, scientists, and fields



# Motivation

## Evolution of citation network



2006



2012

**With what?**

**Large scale data**

**+**

**Statistical analysis**

**+**

**Descriptive model**

# Large scale data



**All papers** published in journals  
of the American Physical Society  
(**APS**) from 1893 to 2008.

**414947** papers and  
**3992736** citations among them

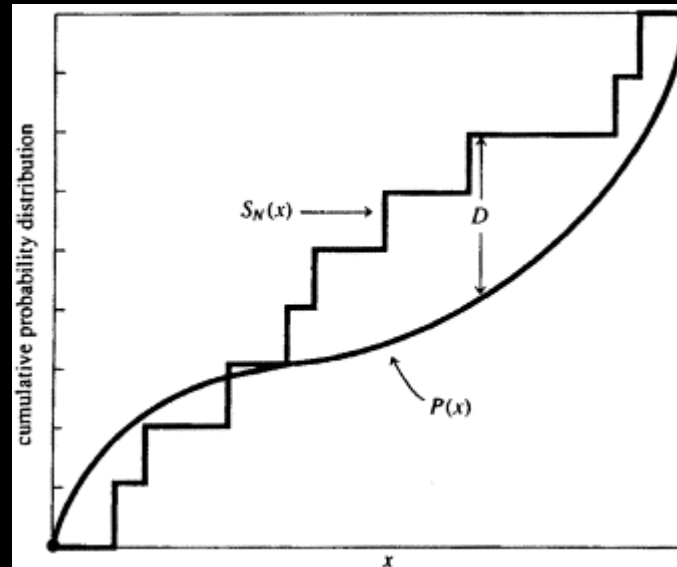
# **Evolution of citation distribution**



**Which function describe empirical distribution best?**

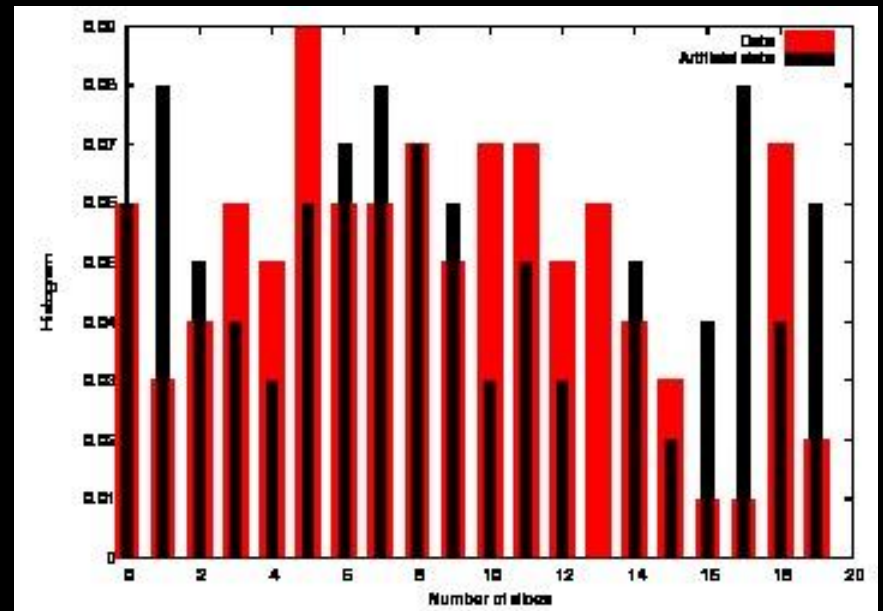
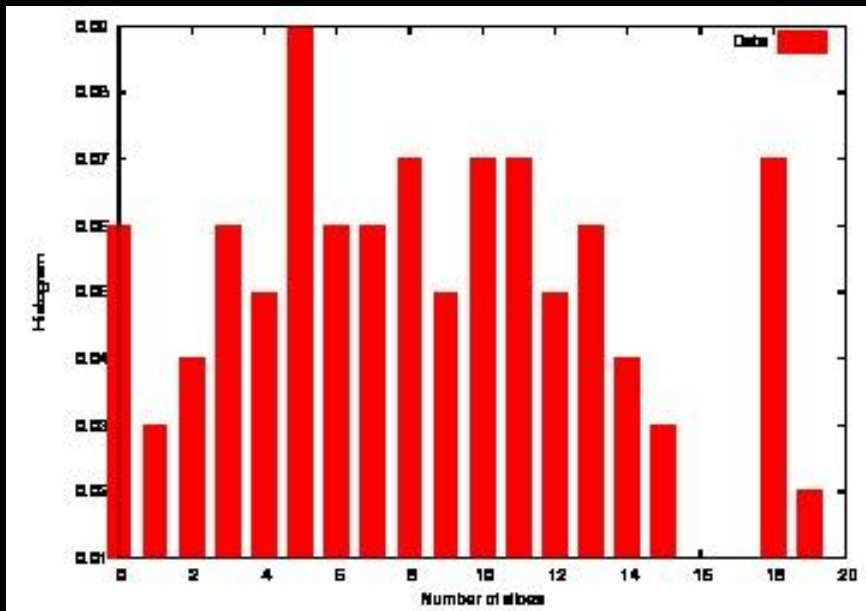
**Fitting analysis with  
goodness of fit test with  
Kolmogorov-Smirnov static**

# Komogorov-Smirnov Static



**Distance between empirical  
cdf and cdf of fit function**

# Goodness of fit test : compare with artificial samples from best fit fn.



# Candidate distributions

$$P(k_{in}) \sim \frac{1}{k_{in} \sqrt{2\pi\sigma}} \exp\{-[\ln(k_{in}) - \mu]^2 / (2\sigma^2)\}$$

Log normal

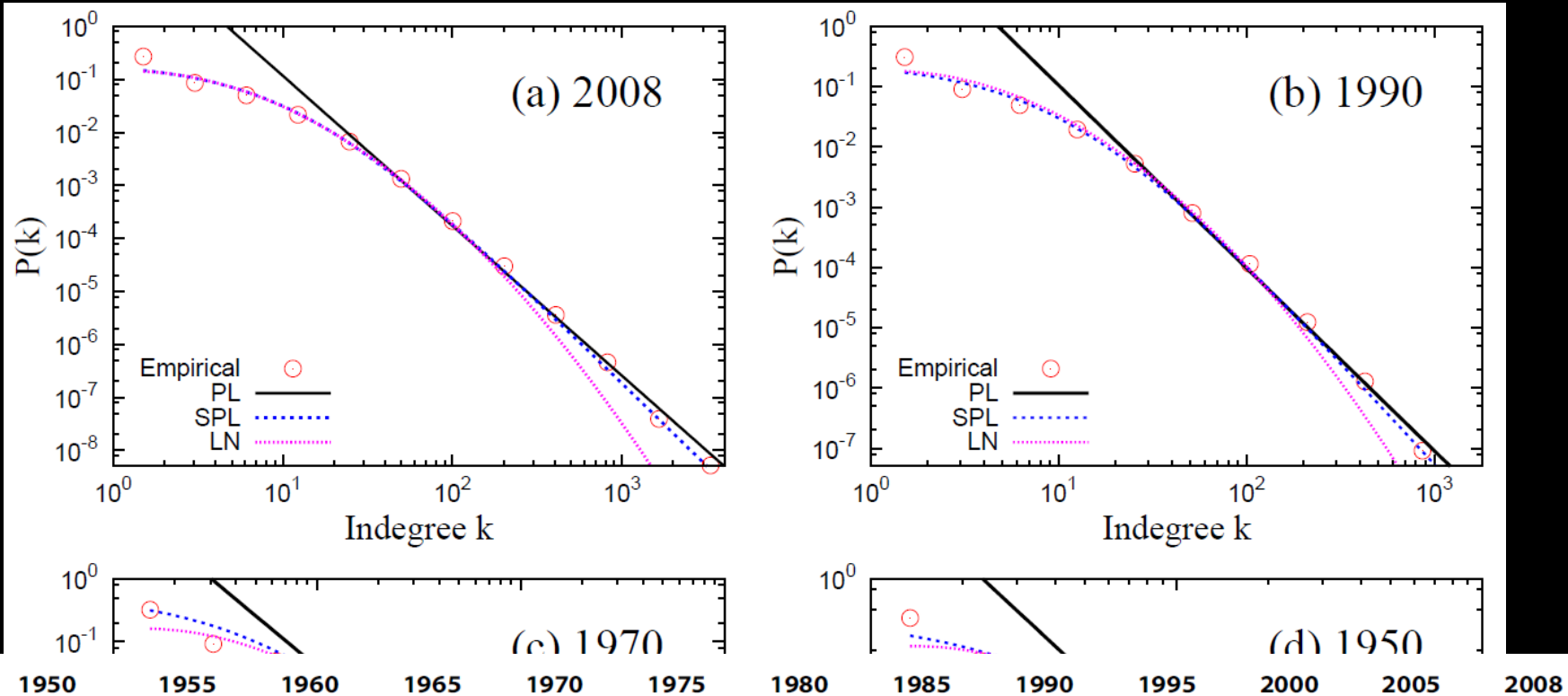
$$P(k_{in}) \sim k_{in}^{-\gamma}$$

Power law

$$P(k_{in}) \sim (k_{in} + k_0)^{-\gamma}$$

Shifted power law

# Fitting results



$$P(k_{in}) \sim (k_{in} + k_0)^{-\gamma}$$

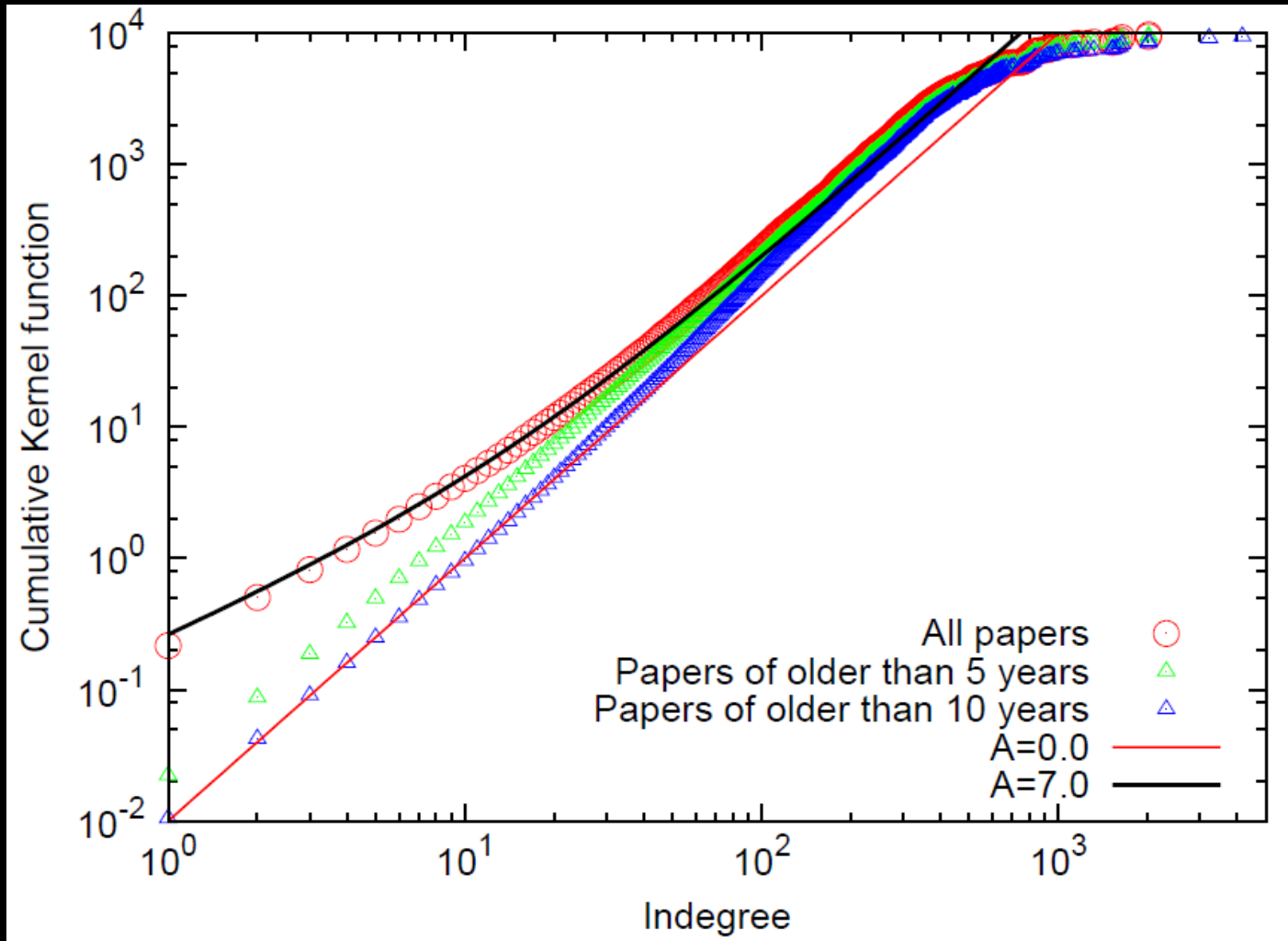
	1950	1955	1960	1965	1970	1975	1980	1985	1990	1995	2000	2005	2008
LN													
p-value	0.717											0.05	0.064
$k_{min}$	2											2	2
PL													
p-value	0.001											0.619	0.44
$k_{min}$	6	16	9	19	12	17	20	39	46	39	43	47	47
SPL													
p-value	0.832	0.777	0.49	1.00	0.943	0.958	0.49	0.728	0.909	1.00	0.797	0.989	0.99
$k_{min}$	2	2	2	14	9	12	2	2	2	2	3	6	5



**What can makes  
shifted power law?**

# **Evolution mechanism**

# PA with attractiveness



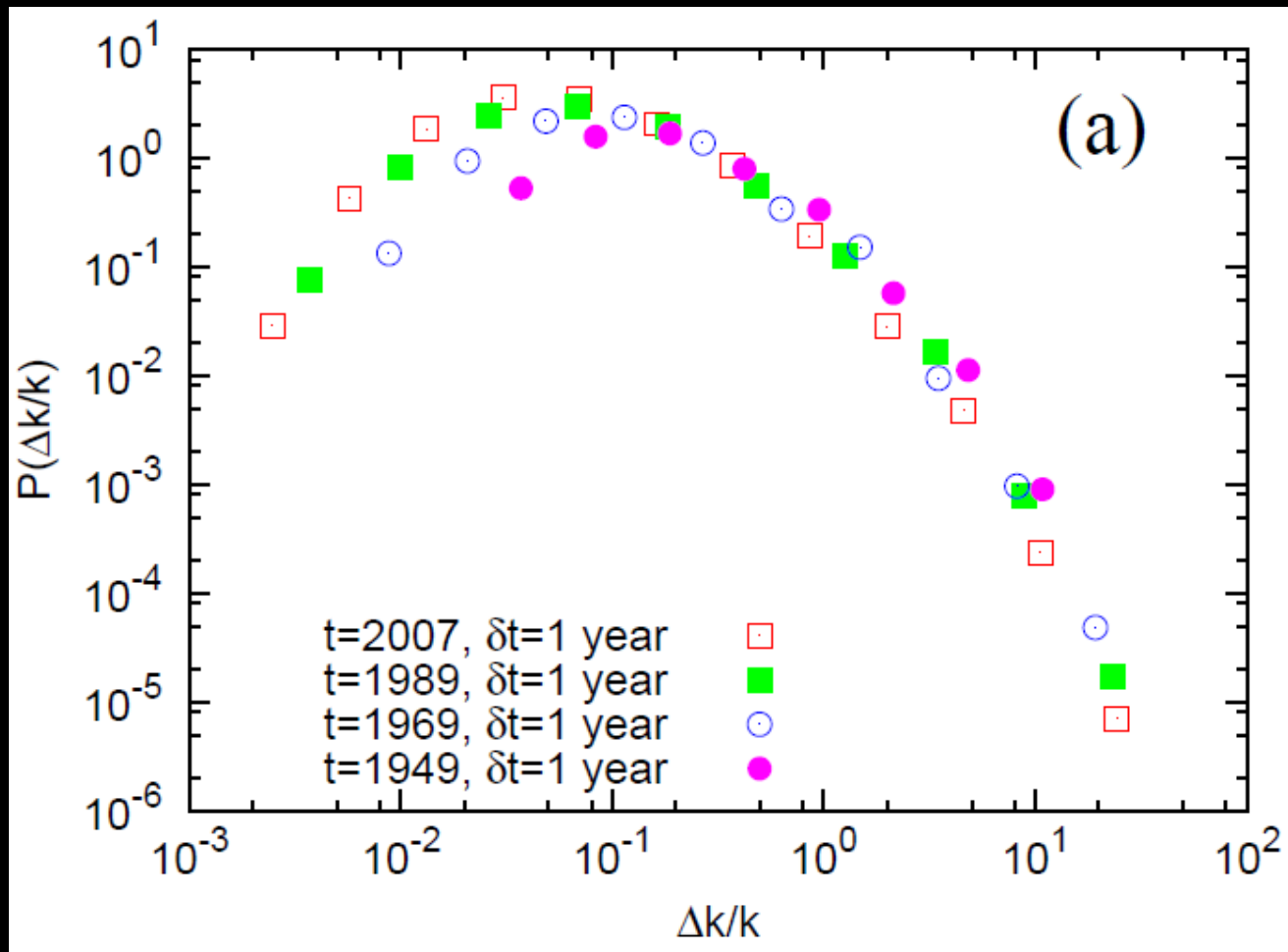
$$P(k) \sim [k + A]$$

**Other dynamics?**

**Citation burst**

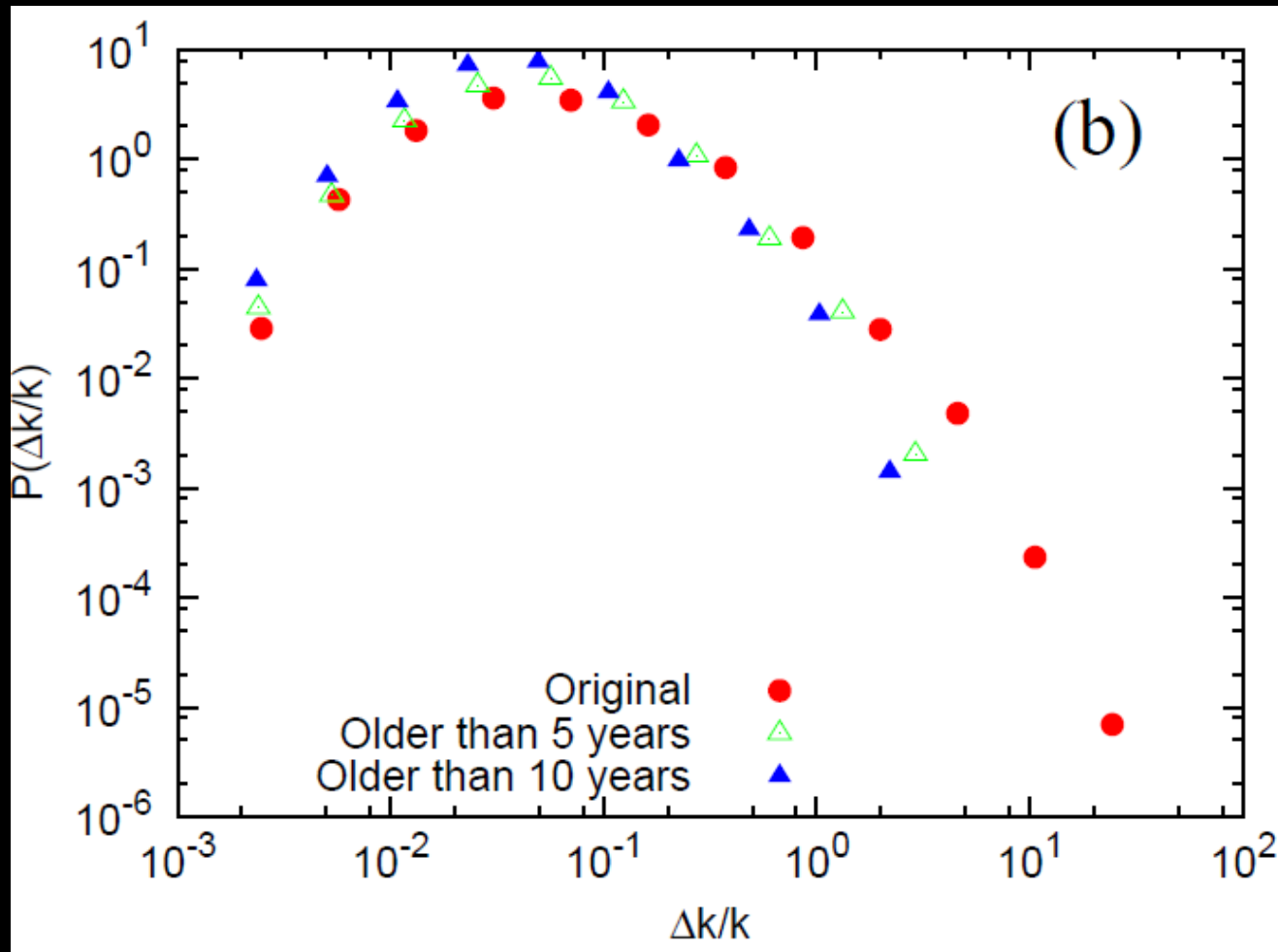
$$\Delta k/k = [k(t + \delta t)_{in}^i - k(t)_{in}^i] / k(t)_{in}^i]$$

# Citation burst





# Citation bursts emerge among only young paper



# The model

$$\Pi(i \rightarrow j, t) \sim [k_{in}^j + A_j(t)]$$

**PA with attractiveness**

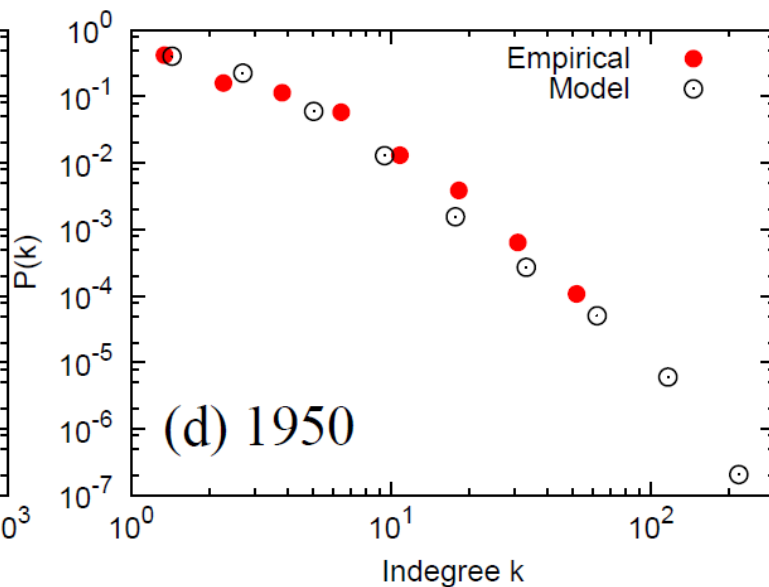
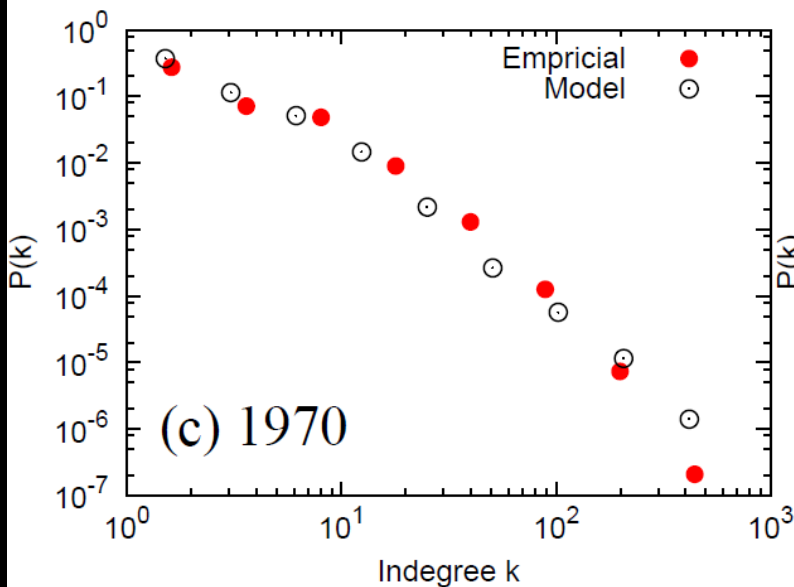
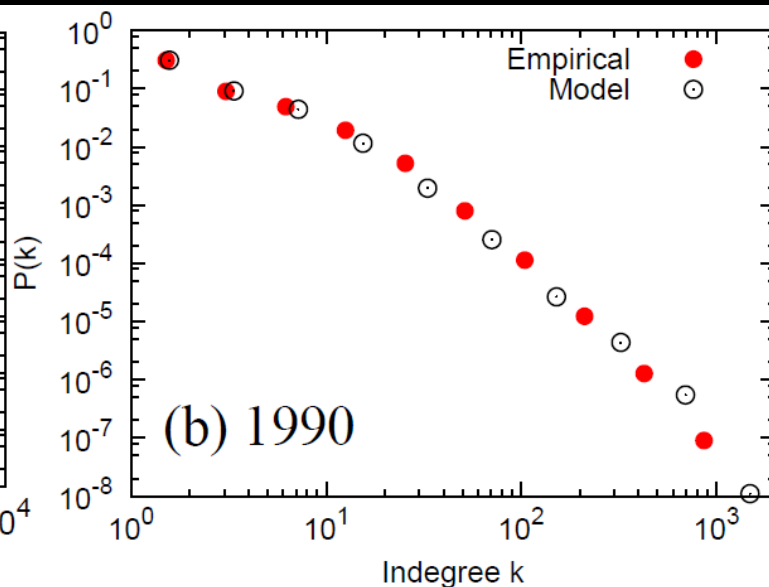
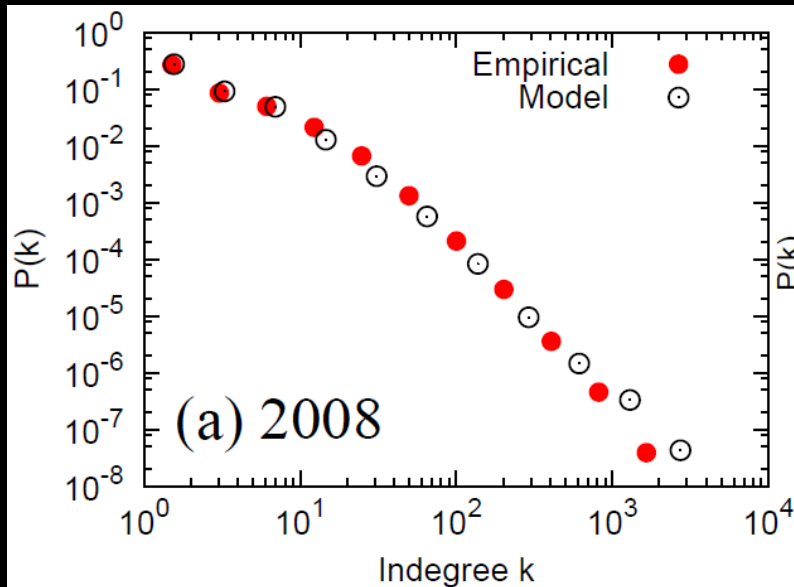
$$A(t) = A_0 \exp[-(t - t_0)/\tau]$$

**Time decaying attractiveness**

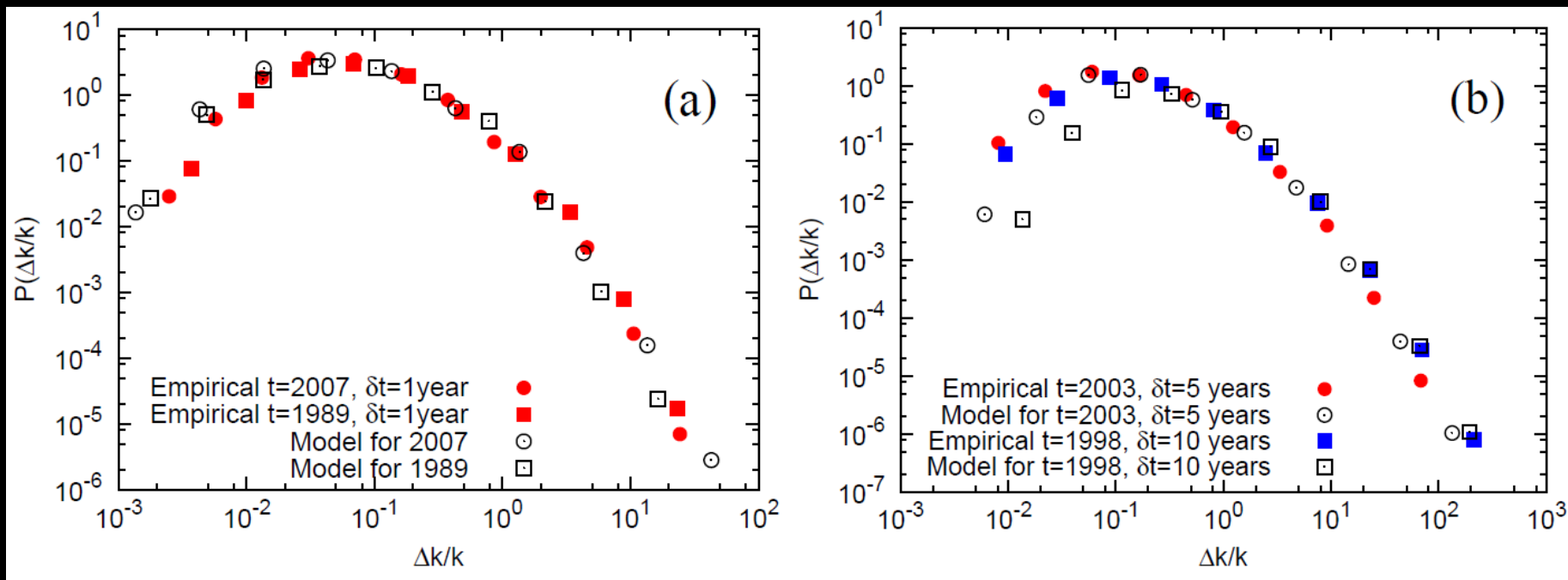
$$P(A_0) \sim A_0^{-\alpha}$$

**Broad attractiveness dist.**

# The model



# The model



# Conclusion

- Citation distribution is best described by shifted power law
- Big bursts emerged only from “young” papers
- Preferential attachment with time-decaying initial attractiveness