Imperial College London

ECCS06
Oxford, 26th September 2006
Tim Evans
Theoretical Physics

Exact Solutions for Models of Network Rewiring

T.S.Evans

- *"Exact Solutions for Network Rewiring Models"* [cond-mat/0607196]

T.S.Evans, **A.D.K.Plato**

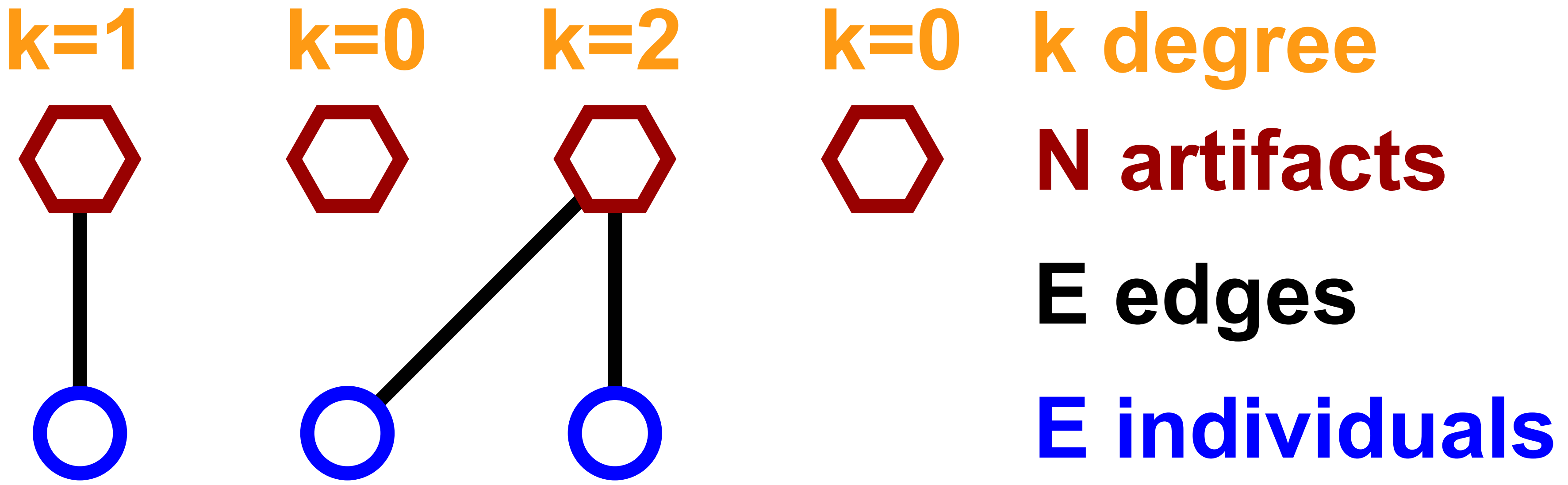
- *"Exact Solutions for Models of Cultural Transmission and Network Rewiring"* [physics/0608052]
- *"Exact Solution for the Time Evolution of Network Rewiring Models"*

© Imperial College London

The Model

- Bipartite network
  E individual vertices each with one edge
  connected to N individual vertices

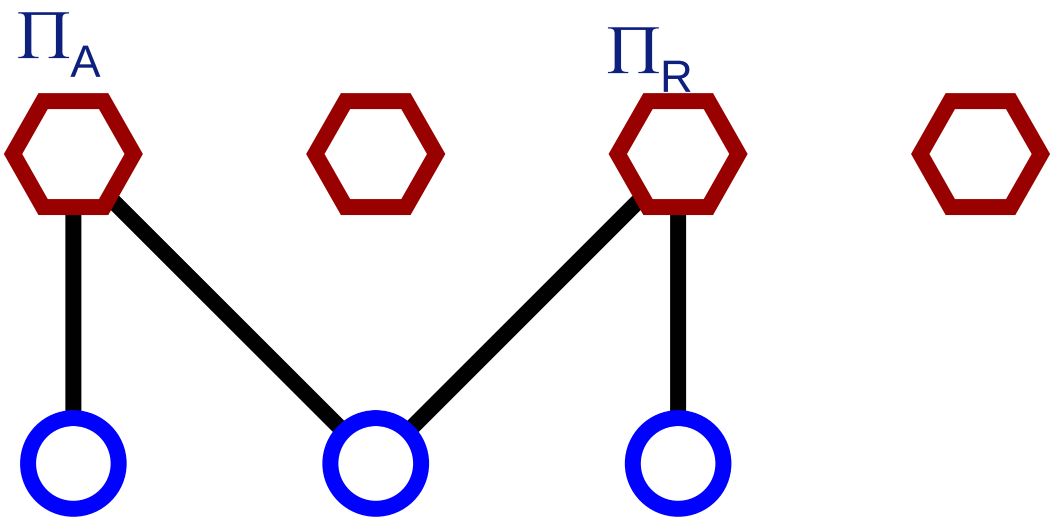
- Study degree k of artifact vertices
  n(k) = degree distribution,
  p(k) = n(k)/N = degree probability distribution



© Imperial College London

The Model - rewiring

- **Removal**: Choose an artifact with probability Π_R and select one of its edges at random for rewiring.
- **Attachment**: Choose an artifact with probability Π_A ready to accept edge.
- **Rewire**: Only after these choices are made is the rewiring performed.



N artifacts

E edges

E individuals

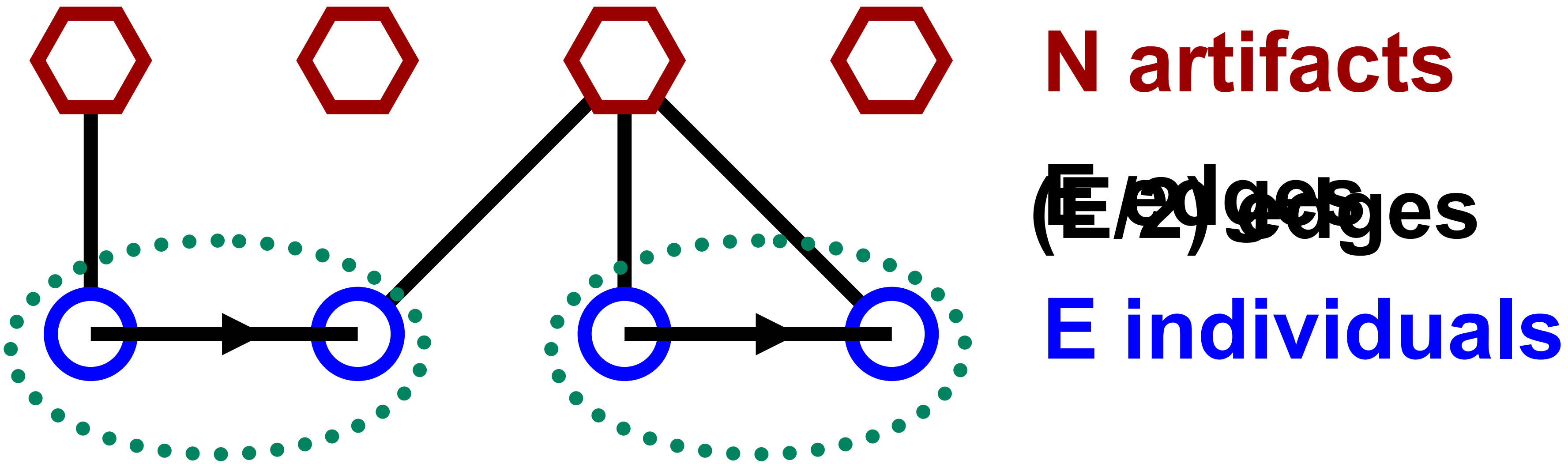
© Imperial College London

Equivalence to other network rewiring models

- Directed/Undirected Network:

  Join edges of individual vertices (2i) and (2i+1).

  [Watts and Strogatz, 1998]



- Directed Network version 2:

  (N=E) Merge each individual vertex with one artifact vertex
  and let edges point from the individual to the artifact end. [Park et al. 2005]

© Imperial College London

Relationship to Statistical Physics Models

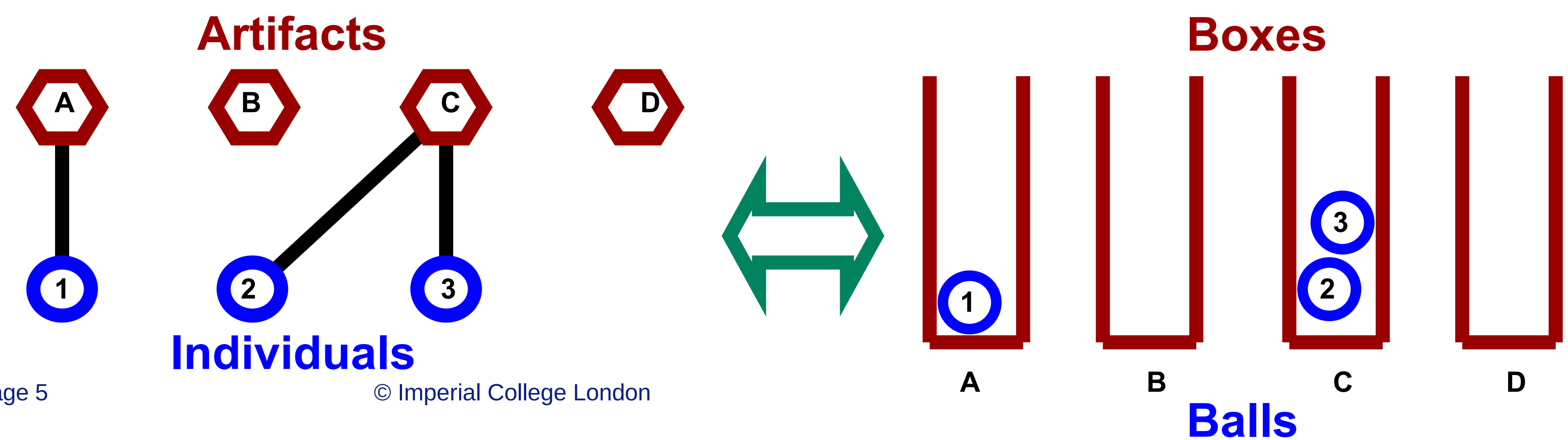
Some parameter values of other models are equivalent to our model:

- Backgammon/Balls-in-Boxes

  applied to glasses [Ritort 1995], wealth distributions, simplicical gravity

- Urn Models [Ohkubo et al. 2005]
- Zero Range Processes (Misanthrope version)

[M.R.Evans & Hanney 2005]



© Imperial College London

Relationship to Other Systems

- Minority Game variant [Anghel et al, 2004]
  Agents (individual vertices) ***copy*** best strategy (artifacts) of their neighbours in an additional individual network.
  Number of people following a given strategy is effectively n(k) of our model.

- Gene Frequencies [Kimura and Crow, 1964]
  Organisms (individuals) ***inherit*** (preferential attachment) copy of a gene (alleles = artifacts) leading to ***drift*** in genetic frequencies. Alternatively they gain a new ***mutation*** (random attachment).
- Family Names [Zanette and Manrubia, 2001]

© Imperial College London

Relationship to Other Systems

- Cultural Transmission [Bentley et al., 2004]

  Individuals ***copy*** (p_p) the choice of artifact made by others or ***innovate*** (p_r)

  e.g. choice of pedigree dog,

  baby names,

  archaeological pottery ty

  tennis star celebration a

J.Connors 1991

Mean Field Degree Distribution Master Equation

Mean field approximation very accurate for many models (low vertex correlations)

$$n(k,t+1)-n(k,t) = +n(k+1,t)\Pi_R(k+1)[1-\Pi_A(k+1)]$$
$$-n(k,t)\Pi_A(k)[1-\Pi_A(k)]$$
$$-n(k,t)\Pi_R(k)[1-\Pi_R(k)]$$
$$+n(k-1,t)\Pi_A(k-1)[1-\Pi_R(k-1)]$$

(1-Π) terms Invariably ignored

Probability of attaching to a vertex of degree (k-1)

Probability of NOT reattaching to same vertex

© Imperial College London

Solvable Case

To be able to solve exactly we limit the attachment and removal probabilities, Π_R and Π_A, to be **linear** in degree.

- $\Pi_R(k) = k / E$

  Choose edge to be removed, or an individual, at random

- $\Pi_A(k) = [\, p_p k + p_r \langle k \rangle \,] / E$

  Fraction p_p of the time use preferential attachment, i.e. choose an artifact with probability proportional to its popularity (degree)

  Fraction p_r of the time choose attach to an artifact chosen randomly

© Imperial College London

Exact Equilibrium Solution

$$n(k) = A \frac{\Gamma(k+\overline{K})}{\Gamma(k+1)} \frac{\Gamma(E-\overline{E}-\overline{K}-k)}{\Gamma(E+1-k)}$$

$$\overline{K} = \frac{p_r}{p_p}\langle k \rangle$$

$$\overline{E} = \frac{p_r}{p_p} E$$

A is ratio of four Γ functions

- Simple ratios of Γ functions
- Similar to those found for growing networks but second fraction is only found for network rewiring
- Only approximate solutions known previously

© Imperial College London

Large Degree Equilibrium Behaviour – Large p_r Case

For $p_r > p_* \sim 1/E$

(on average at least one edge attached to a randomly chosen artifact per generation)

$$\lim_{k\to\infty}\left[n(k)\right] = k^{-\gamma}\exp(-\zeta k)$$

$$\gamma = 1 - \frac{p_r}{p_p}\langle k \rangle$$

Power below one but in data indistinguishable from one

$$\zeta = -\ln(p_p)$$

Exponential Cutoff

© Imperial College London

Large Degree Equilibrium Behaviour – Small p_r Case

$$n(k) = A \frac{\Gamma(k+\overline{K})}{\Gamma(k+1)} \frac{\Gamma(E-\overline{E}-\overline{K}-k)}{\Gamma(E+1-k)}$$

For $p_r < p_* \sim 1/E$

(on average if all edges have been rewired once no edge is attached to a randomly chosen artifact per generation)

2nd Γ function blows up for large degree k

$\Rightarrow$Degree distribution rises near k=E

$\Rightarrow$In extreme case $p_r=0$ all the edges are attached to ONE artifact

- a CONDENSATION or FIXATION

© Imperial College London

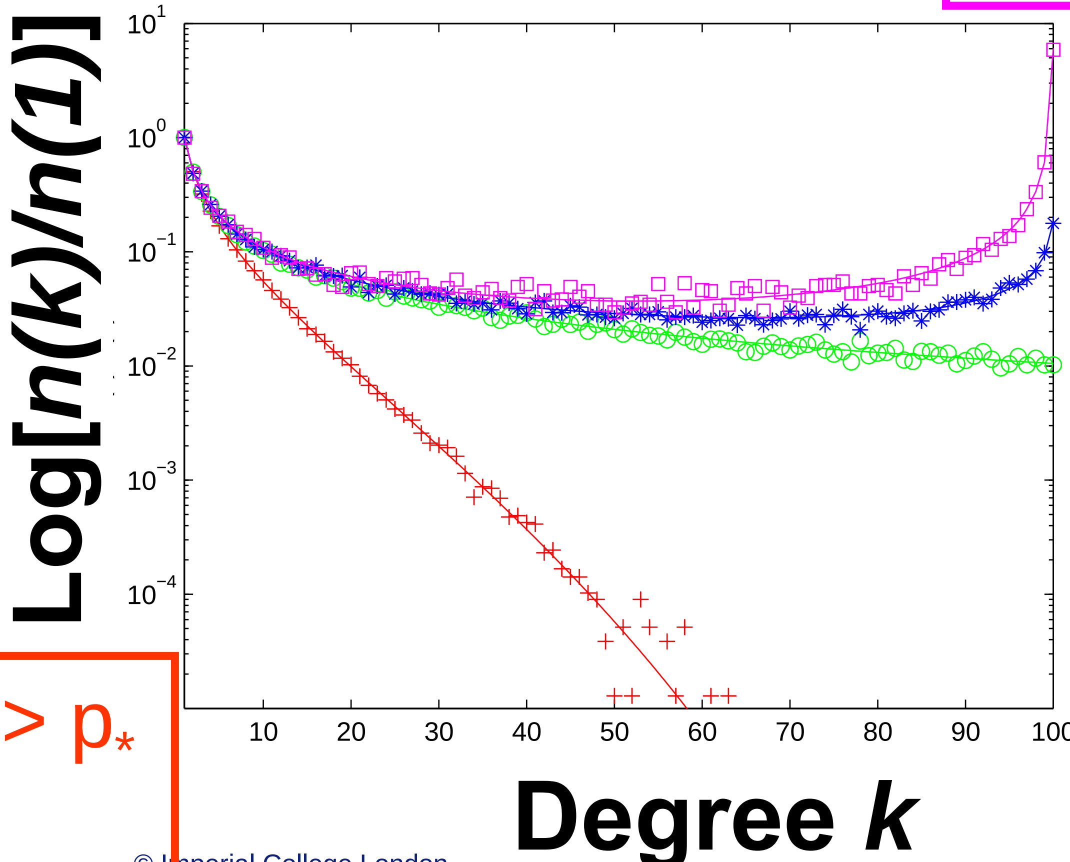
Equilibrium Behaviour Numerical Results

N=E=100

Points: 10^5 data runs

Lines: exact mean field solution

$p_r = 0.001 < p_*$



$p_r = 0.005 < p_*$

$p_r = 0.01 \cong p_*$

Almost pure power law

$p_r = 0.1 > p_*$

$\zeta \cong 10$

© Imperial College London

Time Dependence

Can solve these mean field equations ***exactly*** for all times!

- Equivalent to a Markov process for vector

$$\underline{\boldsymbol{n}}(k) = (n(0), n(1), \ldots, n(E))$$

- Evolves as

$$\underline{\boldsymbol{n}}(k,t+1) = \mathrm{M}\, \underline{\boldsymbol{n}}(k,t)$$

  where M is an (E+1)-dimensional tridiagonal matrix

- M is constant *only* if Π_R and Π_A are both of form (ak+b)

© Imperial College London

Solution

Best solved using the generating function

$$G(z,t) = \sum_{k=0}^{E} (z)^k \, n(k,t) = \sum_{m=0}^{E} c_m (\lambda_m)^t \, G^{(m)}(z)$$

where:-

- **Eigenfunctions** $G^{(m)}(z) = (1-z)^m \, F(a+m, b+m; c; z)$

  Hypergeometric function

  $$a = \frac{p_r}{p_p}\langle k \rangle, \quad b = -E, \quad c = 1 + a + b - \frac{p_r}{p_p} E$$

- **Eigenvalues** $\lambda_m = 1 - m(m-1)\frac{p_p}{E^2} - m\frac{p_r}{E}$

- **c_m are constants fixed by initial conditions**

© Imperial College London

Normalisation

- Only eigenfunction number zero (*m=0*) contributes to the zero-th moment, i.e. the normalisation of the degree distribution is time independent as $\lambda_0=1$

$$N = \mathrm{G}(z=1,t)$$

$$= \sum_{k=0}^{E} \mathrm{n}(k,t) = \sum_{m=0}^{E} c_m (\lambda_m)^t \, \mathrm{G}^{(m)}(1)$$

$$= c_0 \, \mathrm{F}(a,b;c;1)$$

i.e. Simple ratio of Γ functions

© Imperial College London

First Moment - E

- Only eigenfunction number zero and one ($m=0,1$) contribute to the first moment, i.e. the number of edges

$$E = \left.\frac{d\,G(z,t)}{dz}\right|_{z=1} = c_0 \left.\frac{d\,G^{(0)}}{dz}\right|_{z=1} + c_1 (\lambda_1)^t \left.\frac{d\,G^{(1)}}{dz}\right|_{z=1}$$

To ensure E is time independent we *must* set

$c_1 = 0$

$\Rightarrow$ eigenfunction m=1 *never* contributes

© Imperial College London

Homegeneity Measures F_n

- *n*-th derivatives of generating function gives measures of homegeneity related to *n*-th moment of degree distribution

$$\mathrm{F}_n(t) := \frac{\Gamma(E+1-n)}{\Gamma(E+1)} \frac{d^n\, \mathrm{G}(z,t)}{dz^n}\bigg|_{z=1} = \sum_{k=0}^{E} \frac{k}{E} \frac{(k-1)}{(E-1)} \cdots \frac{(k-n+1)}{(E-n+1)} \mathrm{n}(k)$$

- **These are simple known ratios of Γ functions**

- **Equals the probability of choosing *n* different individuals connected to the same artifact**

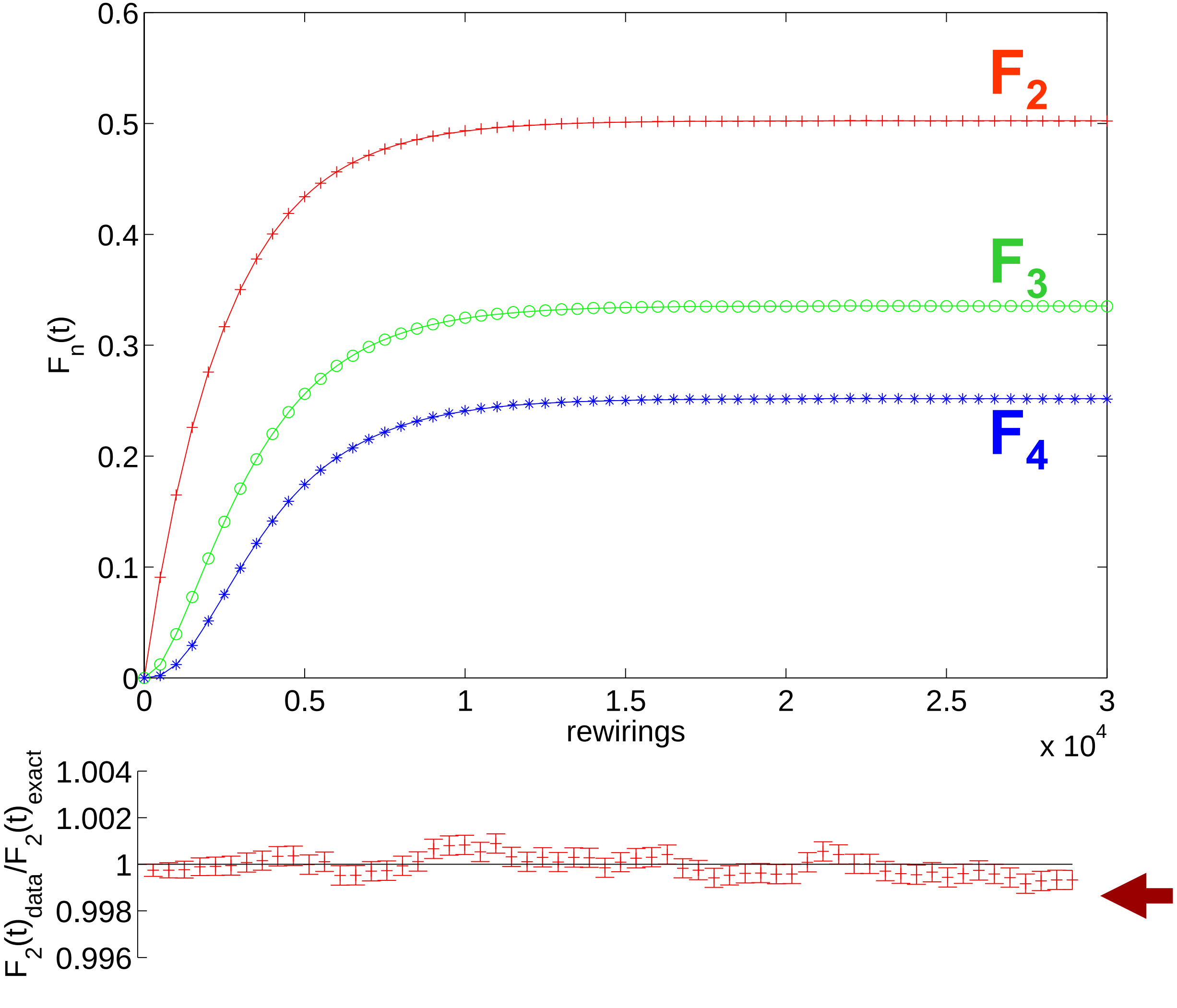
⇒ $F_n = 0$ if no artifact chosen more than once

$F_n = 1$ if all individuals attached to same artifact

© Imperial College London

F_n numerical results

$E=N=100$, $p_r=0.01 \cong p_*$,
Points: average of 10^5 simulations
Lines: exact mean field prediction
Start: $n(k)=\delta_{k,1}$



F increases as homegeneity increases with time

Time dependence of averages predicted very accurately, deviations less than 1%

Does the distribution change definition of an artifact altered for the same data set?

Example [Morgan & Swanell]:

- The shoes of 200 male physicists are photographed as they leave a lecture
- These are put into categories by 6 different people e.g. by TSE using colour, type, fastening
- Even with the same categorisation different people will produces different results
  - what is blue, what is a trainer?
- Small data set yet the results all seem to be consistent with the model
  – long tails at least power law and cutoff fit reasonably

© Imperial College London

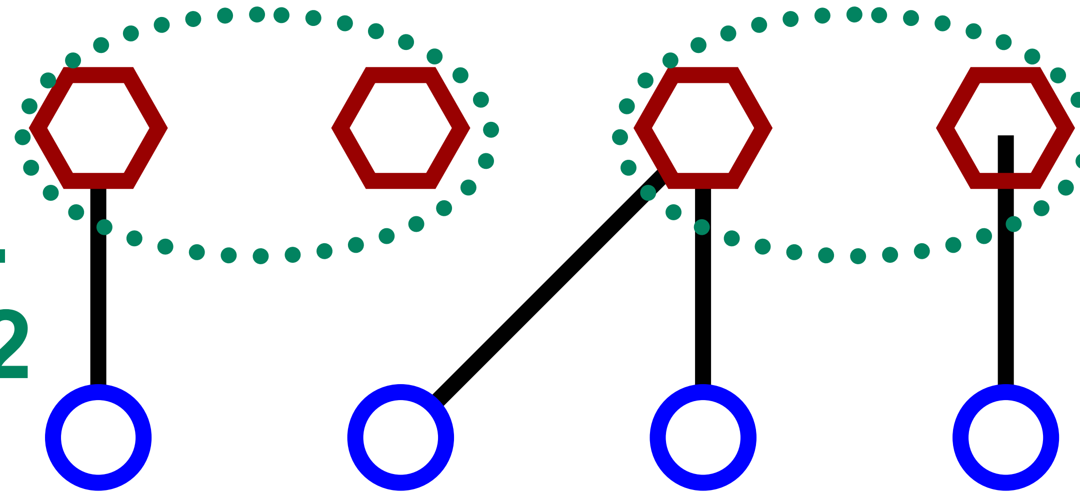
Scaling

- Rewire network as before
- Study the distribution of artifact pairs

  Pair artifacts at random.

  Consider degree distribution of artifact pairs $n_2(k)$.

⇒ **Equations as before with no. vertices N→N/2**



Shape of distribution same, only cutoff altered

© Imperial College London

Minority Game Example - Leaders and Followers

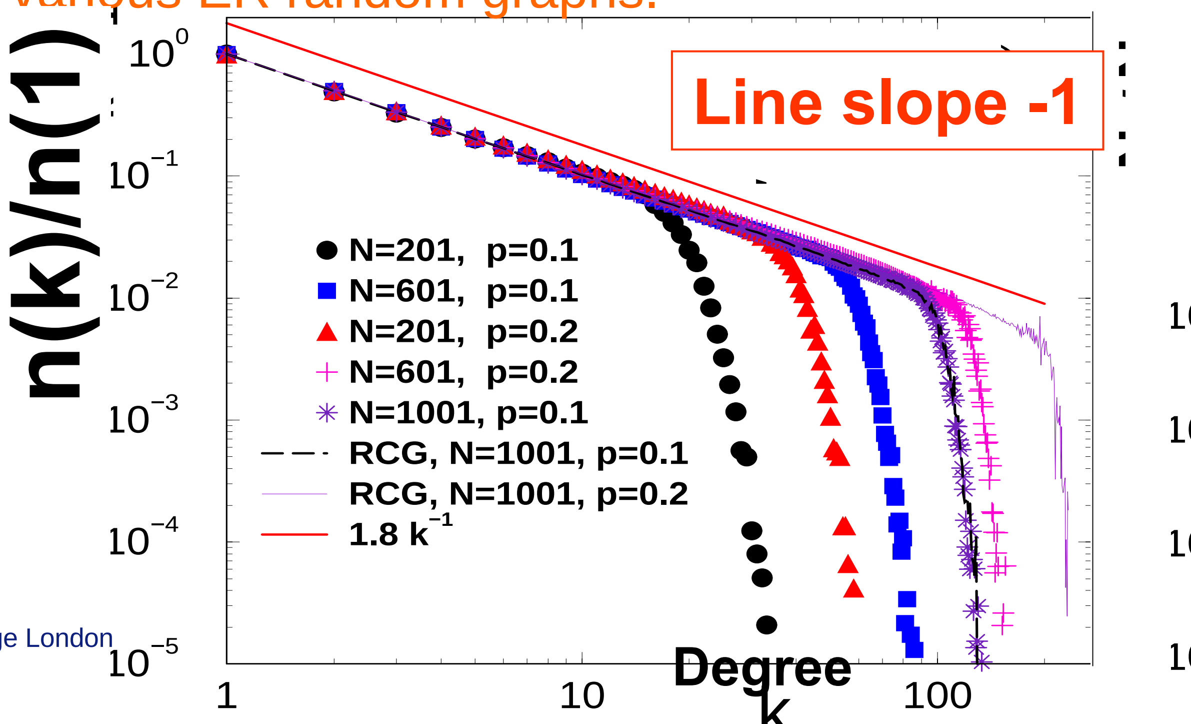
Individuals choose best strategy (each artifact is a strategy) from their neighbours in an ER random graph of individuals [Anghel et al. 2004]

Plot *n(k)* the average of the number of strategies (of some leader) used by *k* individuals (followers).

Various system sizes and various ER random graphs.

Result exactly as in our model

⇒ Random Copying



© Imperial College London

Summary

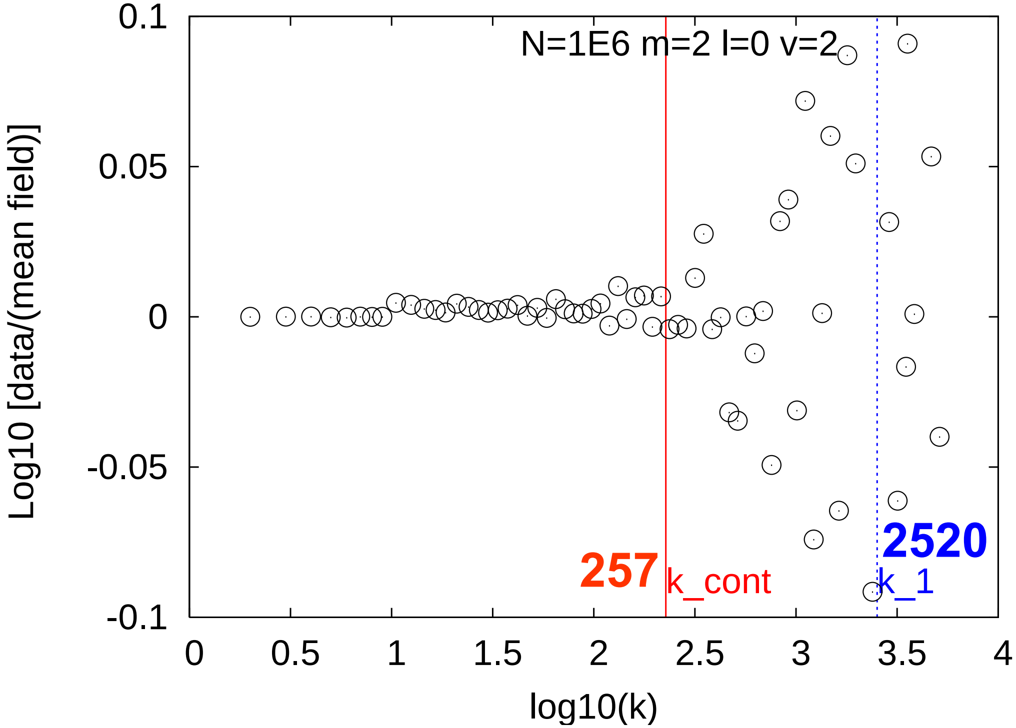
- Made connections between rewiring of bipartite network and many other network, statistical physics and social science models.
  Some connections made in some existing papers.
- More accurate mean field equation.
  Only now is behaviour at boundary k=E correct.
- Exact equilibrium solutions.
  Previous results for large degree k, large systems N,E.
- Exact solutions for all times in terms of standard functions.
  I know of no other network solutions for arbitrary time and arbitrary size.

© Imperial College London

Mean Field Solutions

- Assume behaviour of the average number of vertices of degree k given by the average properties of the network
- These are excellent for pure preferential attachment (Simon/BA) ⇔ correlations in degrees of neighbouring vertices insignificant

Fractional deviation of data from one run of pure pref. attachment model against mean field solution



$n(k_{cont}) := 1$

$k_{cont} = O(N^{1/3})$

Limit of good data

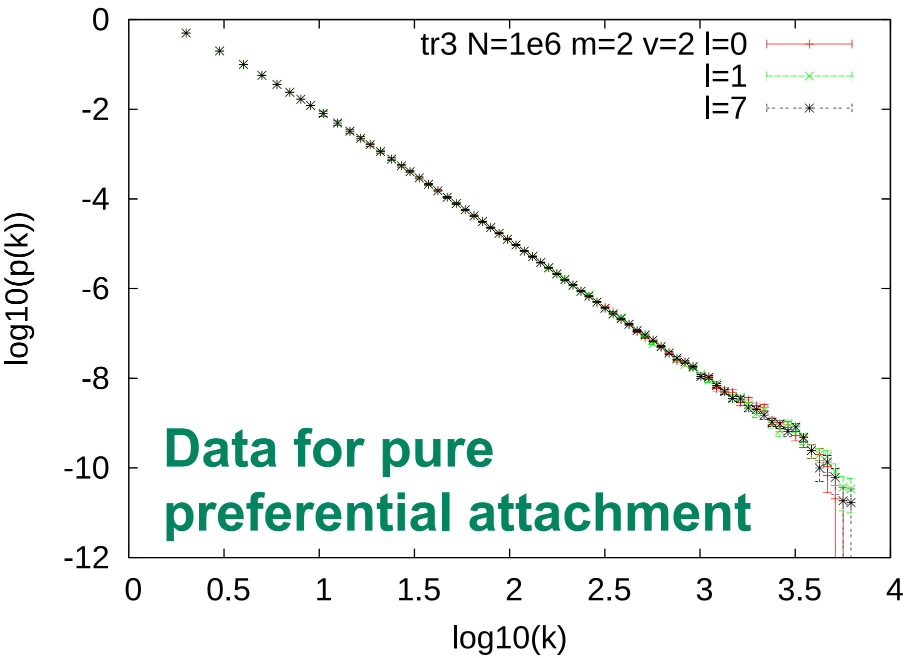
© Imperial College London

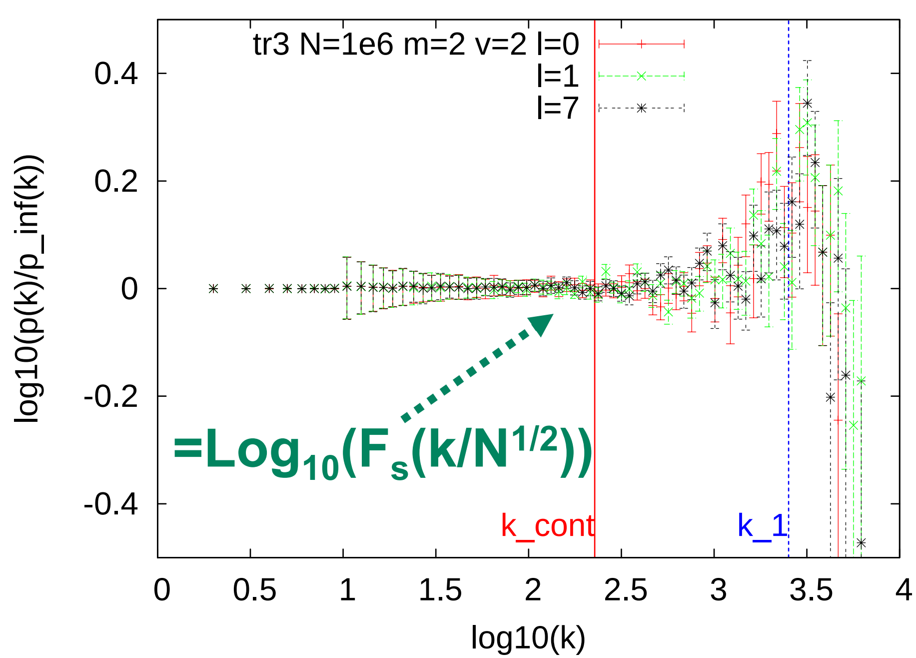
Finite Size Effects for pure preferential attachment

$$p(k) = p_\infty(k).F_S\left(\frac{k}{N^{1/2}}\right), \qquad p_\infty(k) = \frac{\langle k\rangle(\langle k\rangle+2)}{2k(k+1)(k+2)} \to \frac{1}{k^3}$$

Scaling Function F_s

$$F_S(x) \approx 1 \quad if \quad x < 1$$





100 runs to get enough data near k_1

© Imperial College London