Are Copying and Innovation Enough?

Doug Plato, Institute for Mathematical Science
Tim Evans, Theoretical Physics
Tevong You, Physics Dept.

ECMI2008
UCL

Outline

- Part 1: Simple Copying Model
  - Why Copying?
  - Introduce a simple network model with only copying and innovation
  - Show how this model can be solved exactly for all times
  - Properties of the network: Homogeneity
- Part 2: Generalisations
  - Additional vertex graphs
  - Update Dynamics
  - Additional Vertex types (time permitting!)

Why Copying?

Local process yet produces macroscopic features

- Copying can lead to preferential attachment i.e. with probability proportional to degree k [e.g. T.S. Evans & J.P. Sarämaki (2005)]→ scale free behaviour

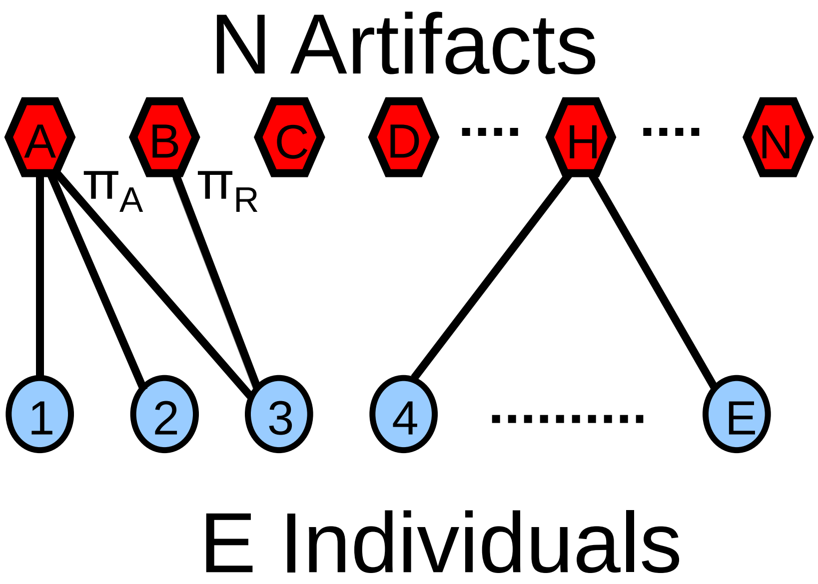
Copying and Innovation natural mechanisms in many models

- Gene Frequencies: Inheritance and Mutation [M. Kimura & J.F. Crow (1964)]
- Urn models
- Voter Models: e.g. Language Extinction [Stauffer et al. (2007)]
- Cultural Transmission: Baby names, dog breeds, pop music… [e.g. Bentley et al. 2007]

Interesting from a mathematical point of view – simple models with just copying and innovation can be exactly solved.

Model - Simple non-growing bi-partite network

- Edges fixed to Individuals but Artifact side free to rewire
- Choose an Individual to remove edge, with probability π_R
- Choose Artifact to attach edge to, with probability π_A
- Only then perform rewiring



Model - Mean Degree Distribution Master Equation

- Interested in degree distribution n(k,t) (and p(k,t)=n(k,t)/N) of Artifact Vertices.

Mean Field Equation:

$$
\begin{aligned}
n(k,t+1) - n(k,t) = &+ n(k+1,t)\Pi_R(k+1)[1-\Pi_A(k+1)] \\
&- n(k,t)\Pi_A(k)[1-\Pi_A(k)] \\
&- n(k,t)\Pi_R(k)[1-\Pi_R(k)] \\
&+ n(k-1,t)\Pi_A(k-1)[1-\Pi_R(k-1)]
\end{aligned}
$$

Number of edges attaching to vertex of degree (k-1)

Probability of NOT reattaching to same vertex

Model - Mean Degree Distribution Master Equation

- Good approximation when vertex correlations small

- For Probabilities $\propto k^{\beta}/z_{\beta}$, exact when normalisations constant, i.e.

$$\sum_k n_i(k) k^{\beta} = \left\langle \sum_k n_i(k) k^{\beta} \right\rangle$$

- ONLY when β=0 or β=1, thus most general π_R and π_A

$$\Pi_R = \frac{k}{E}, \qquad \Pi_A = p_r \frac{1}{N} + p_p \frac{k}{E}$$

Only Random or Preferential Attachment ($p_r + p_p = 1$)

Model - Solutions

- Mean-field equation is linear
  $\rightarrow$ Encode degree distribution in Generating function

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

- Re-write mean-field equation as differential equation:

$$z(1-z)G^{(m)''}(z) + \left[c - (a+b+1)z\right]G^{(m)'}(z) - \left[ab - \frac{(\lambda_m - 1)}{1-z} b(c-a-1)\right]G^{(m)}(z) = 0$$

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

Model - Solutions

- Expand $G^{(m)}(z)$ as a polynomial in (1-z). Leads to Hypergeometric differential equations with solutions

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

and Eigenvalues given by

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

$$\lambda_m \geq \lambda_{m+1}$$

Model - Equilibrium Distribution

- Can recover degree distribution from generating function.

$$n(k) = \frac{1}{k!} \frac{d^k G(z)}{dz^k}\bigg|_{z=0} = A \frac{\Gamma(k+\overline{K})}{\Gamma(k+1)} \frac{\Gamma(E-\overline{E}-\overline{K}-k)}{\Gamma(E+1-k)}$$

$$\overline{K} = \tfrac{p_r}{p_p}\langle k \rangle \qquad \overline{E} = \tfrac{p_r}{p_p} E$$

A – Normalisation. Ratio of Gamma functions.

- Exact for all parameter values of E, N and p_r.

Model - Equilibrium Distribution

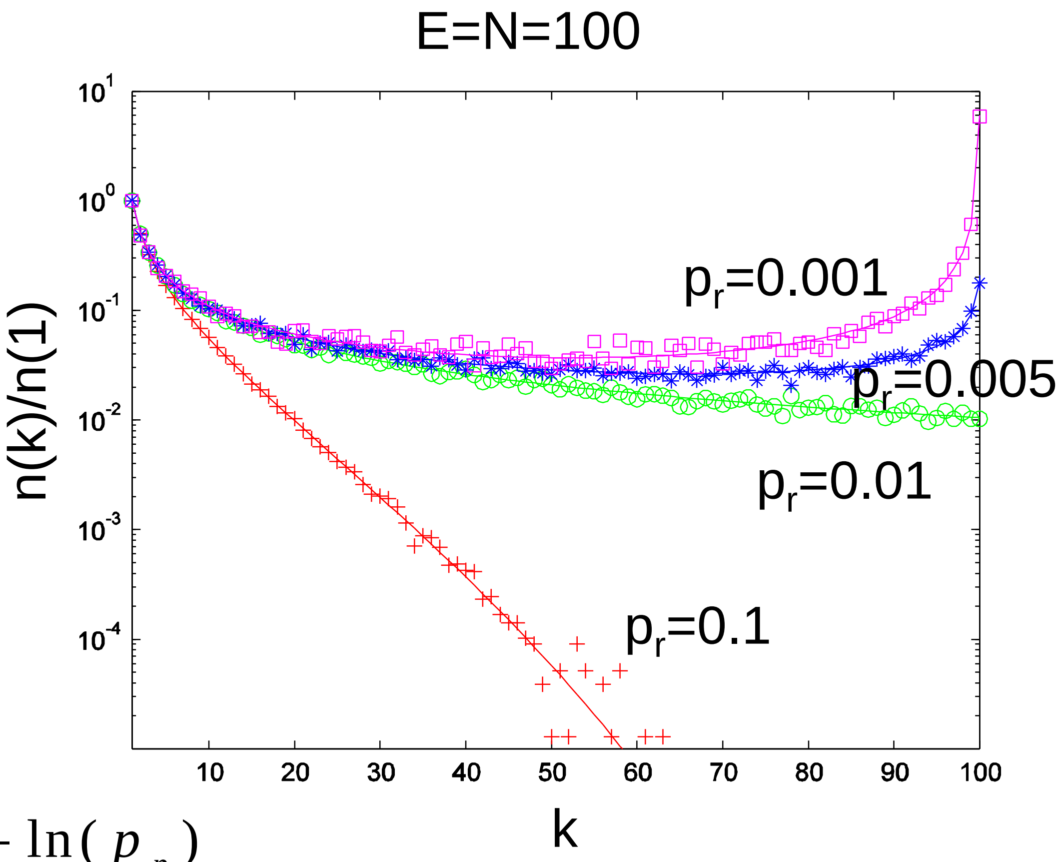
- Numerical: Averaged over 10^5 realisations
  Lines: Analytic Results

- $p_r \geq 1/E$: power law

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

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

and exponential cutoff $\zeta = -\ln(p_p)$



Homogeneity Measures

- Call the probability that two *distinct* edges chosen at random connected to same Artifact, $F_2(t)$

$$F_2(t) = \sum_{k=0}^{E} \frac{k(k-1)}{E(E-1)} n(k,t) = \frac{1}{E(E-1)} \left. \frac{d^2 G(x,t)}{dx^2} \right|_{x=1}$$

- E fixes $c_1=0$

$$\longrightarrow \quad F_2(t) = F_2(\infty) + (\lambda_2)^t \left(F_2(0) - F_2(\infty) \right)$$

$$F_2(\infty) = \frac{1 + p_r \left(\langle k \rangle - 1 \right)}{1 + p_r (E-1)}$$

$F_2(0)$ fixed by initial conditions

- Generalise to *n* distinct edges

$$F_n(t) = \frac{\Gamma(E+1-n)}{\Gamma(E+1)} \left. \frac{d^n G(x,t)}{dx^n} \right|_{x=1}$$

- $g_n^{(m)}$: n^{th} derivatives of $G^{(m)}(z)$ at z=1. Only non zero for m≤n so contributions only from first n+1 eigenfunctions:

Homogeneity Measures

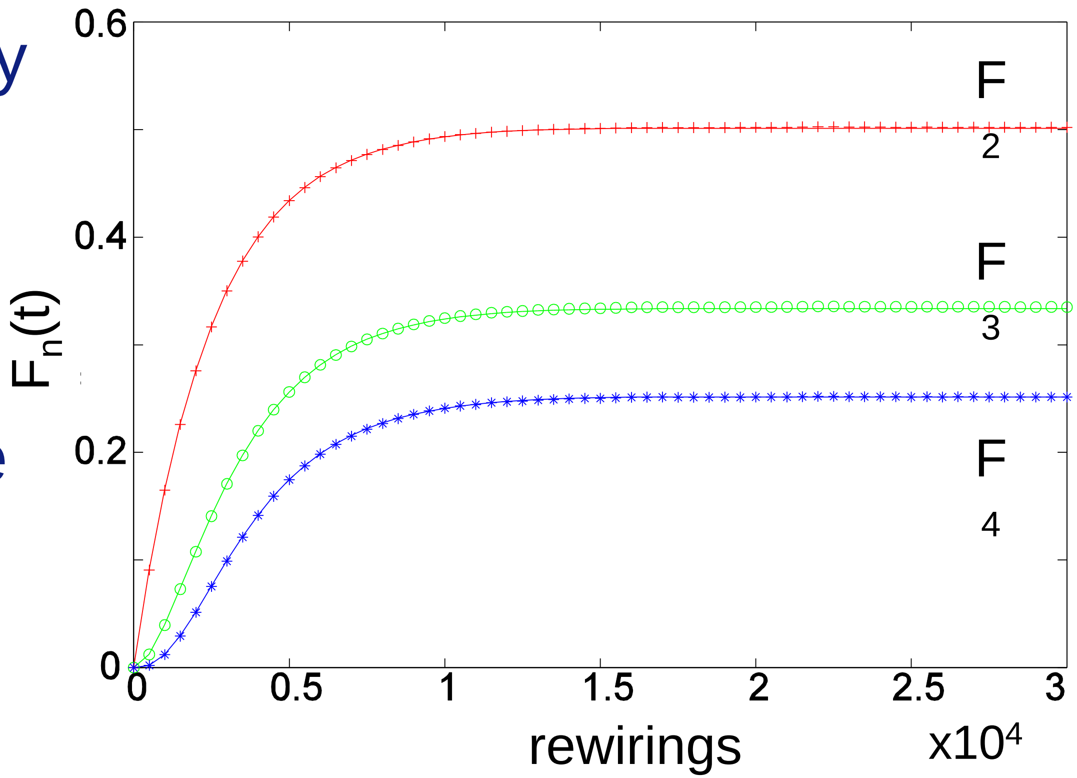
- F_n measures can fix coefficients c_m iteratively

$$\frac{\Gamma(E+1-n)}{\Gamma(E+1)} \sum_{m=0}^{n} c_m g_n^{(m)} = F_n(0)$$

$$g_n^{(m)} = 0, \quad m > n$$

- Network becomes more homogenous with time

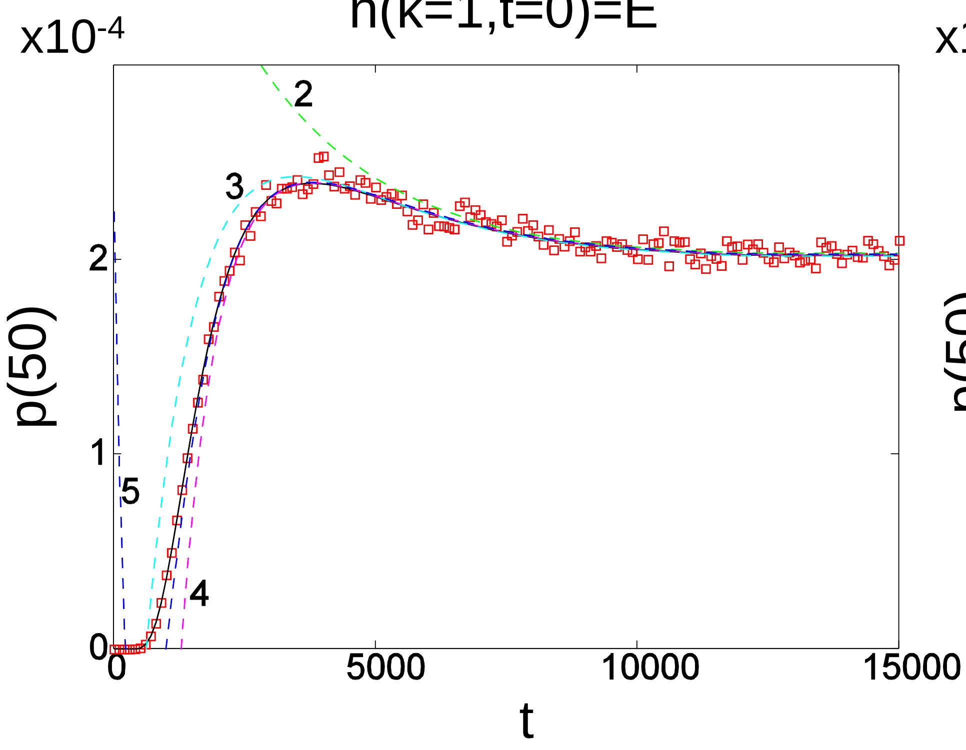
E=N=100, p_r=1/E. n(k)=$\delta_{k,1}$ averaged over 10^5 runs

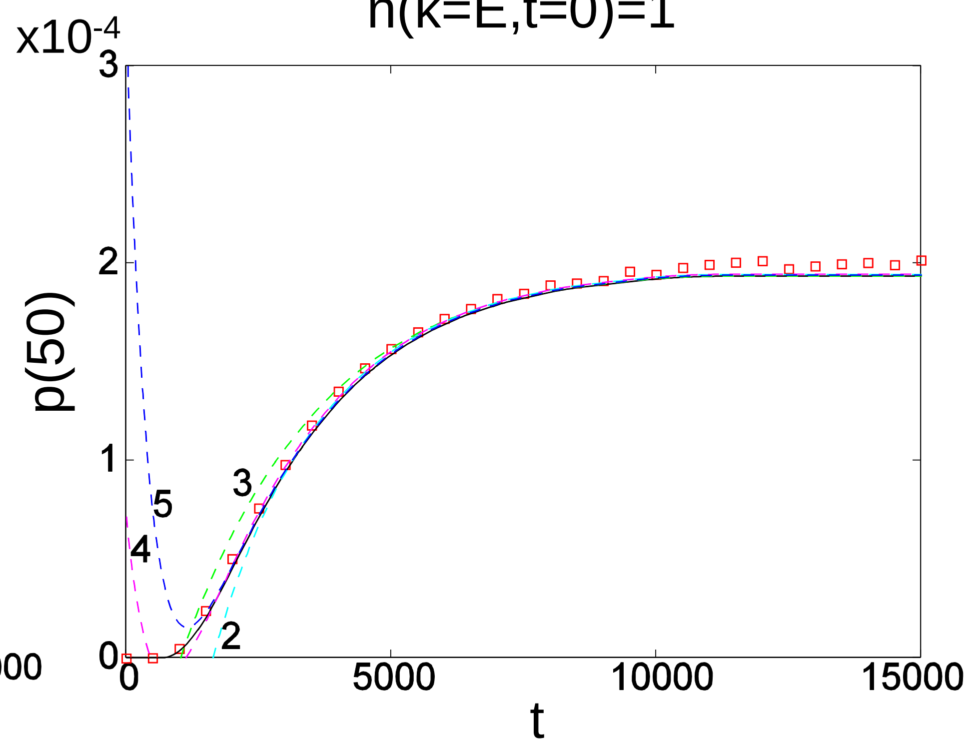


Time Dependence

- Low Eigenvalue contribution to the degree distribution evolution

N=E=100, pr=1/E





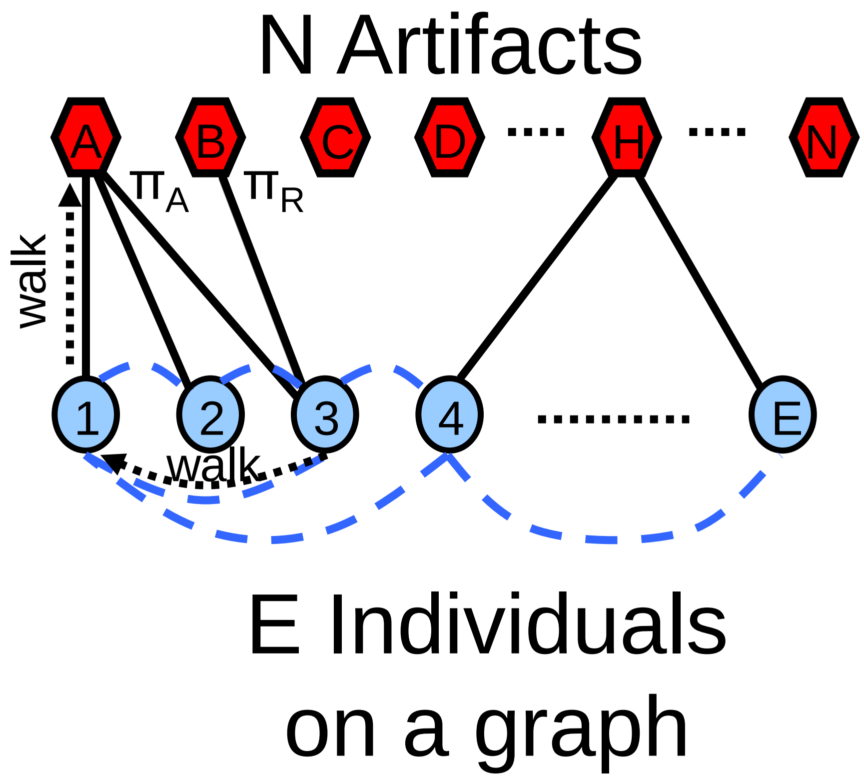
Part II:

Generalisations

- Add a graph to the Individuals Vertices
- Different update dynamics
  - Simple example: 2 step update
  - Multi edge rewiring
- Different Types of Individuals

Network of Individuals

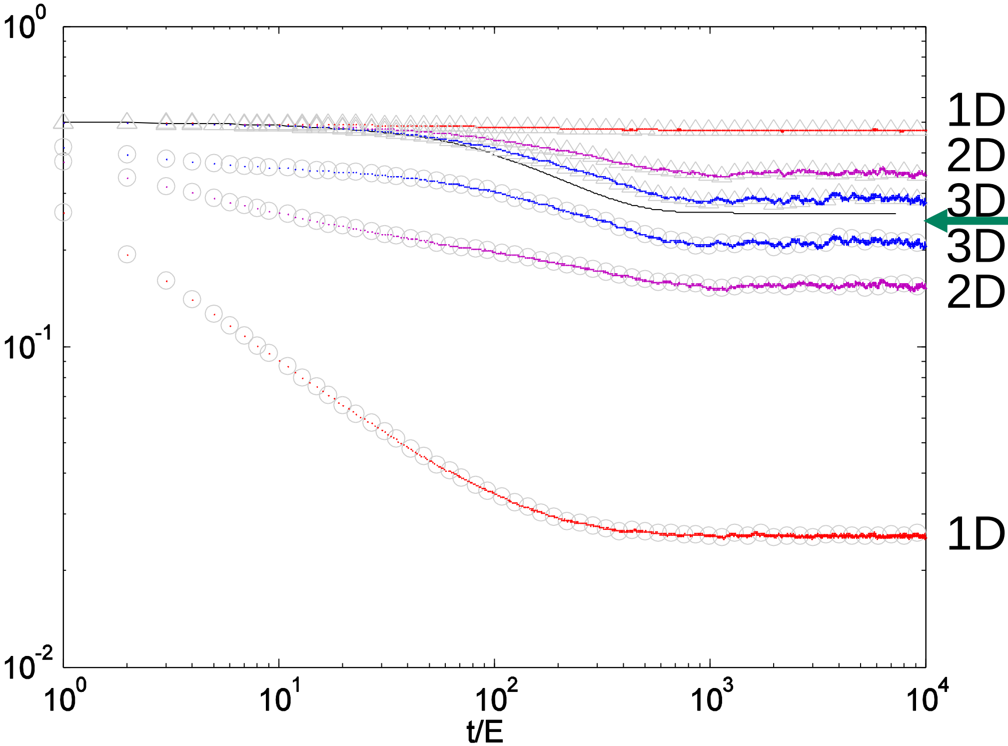
- Individuals vertices now connected by a graph
- Choose removal edge with π_R=k/E
- Preferential Attachment is replaced by a random walk on Individuals network
- Copy vertex choice
- Rewire



Network of Individuals – Homogeneity

- N=2: traditional Voter Model with randomness
- Average interface density $\langle\rho\rangle_t$
- $(1-F_2(t))$ vs. $\langle\rho\rangle_t \leftrightarrow$ Global vs. local

N=2, E=729, pr=1/E



Global $(1-F_2)$

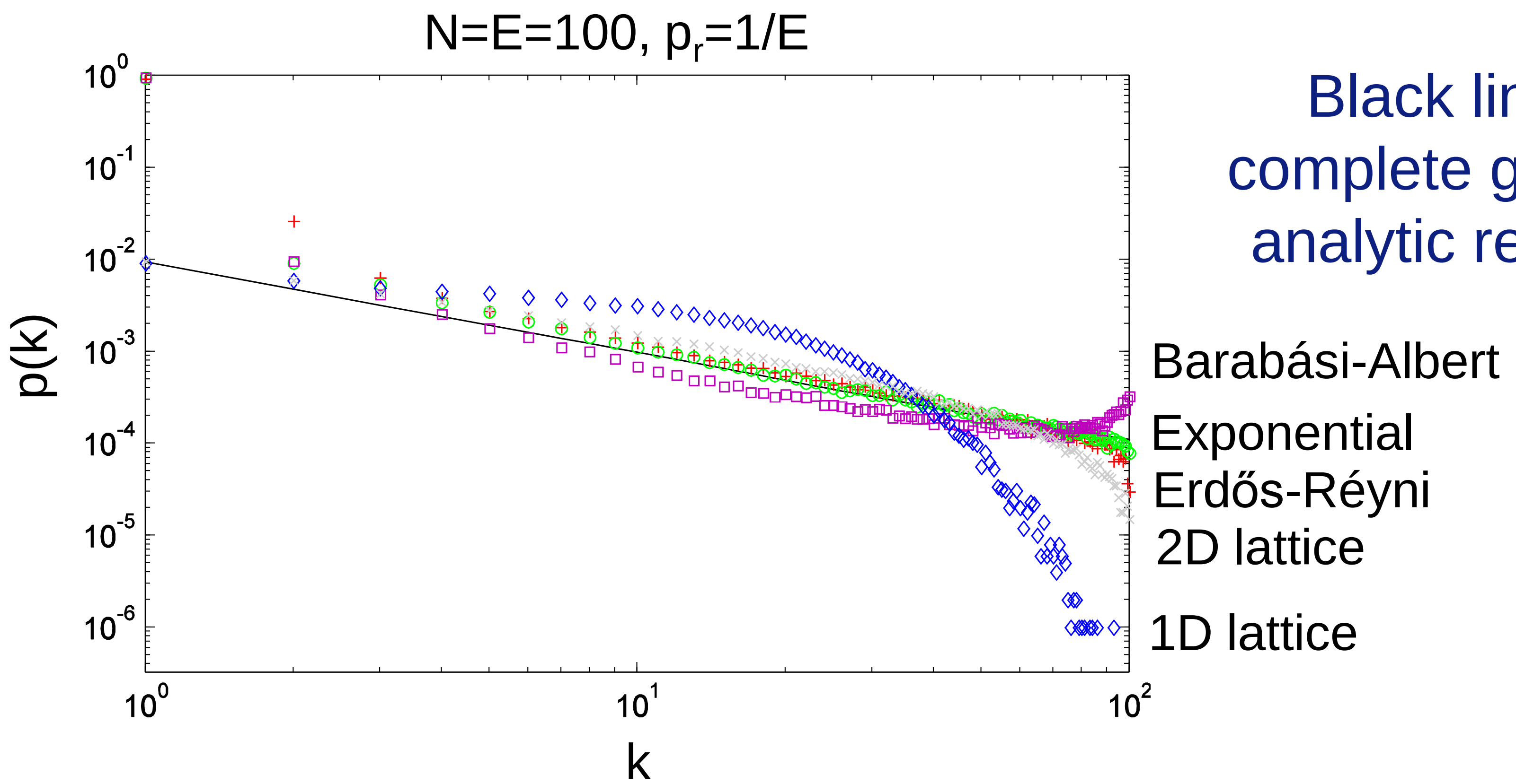
Complete graph

Local $\langle\rho\rangle$

- Randomness prevents consensus

Network of Individuals – Degree Distribution

- Degree Distribution similar except for 1D lattice



Black line complete graph analytic result

Alternate Update dynamics – two step update

Simple Example. Step 1: Remove edge

Step 2: Attach new edge

- New attachment probability: $\tilde{\Pi}_A = p_r \frac{1}{N} + p_p \frac{k}{E-1}$
- Modified master equation.

However…

Solution still similar to before

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

Minor change in parameters $a \to \tilde{a} = \frac{p_r}{p_p}\frac{E-1}{N}$, $b = -E$,

$c \to \tilde{c} = 1 + \tilde{a} + b - \frac{p_r}{p_p} E$

Alternate Update dynamics – Multi Edge Rewiring

In more realistic models up to date information may not be available for each entity

- Parameter X controls number of distinct edges to be rewired at each timestep
- Now choose edges at Random or Sequentially
- Extremes: X=1 → Simple Copy Model already discussed

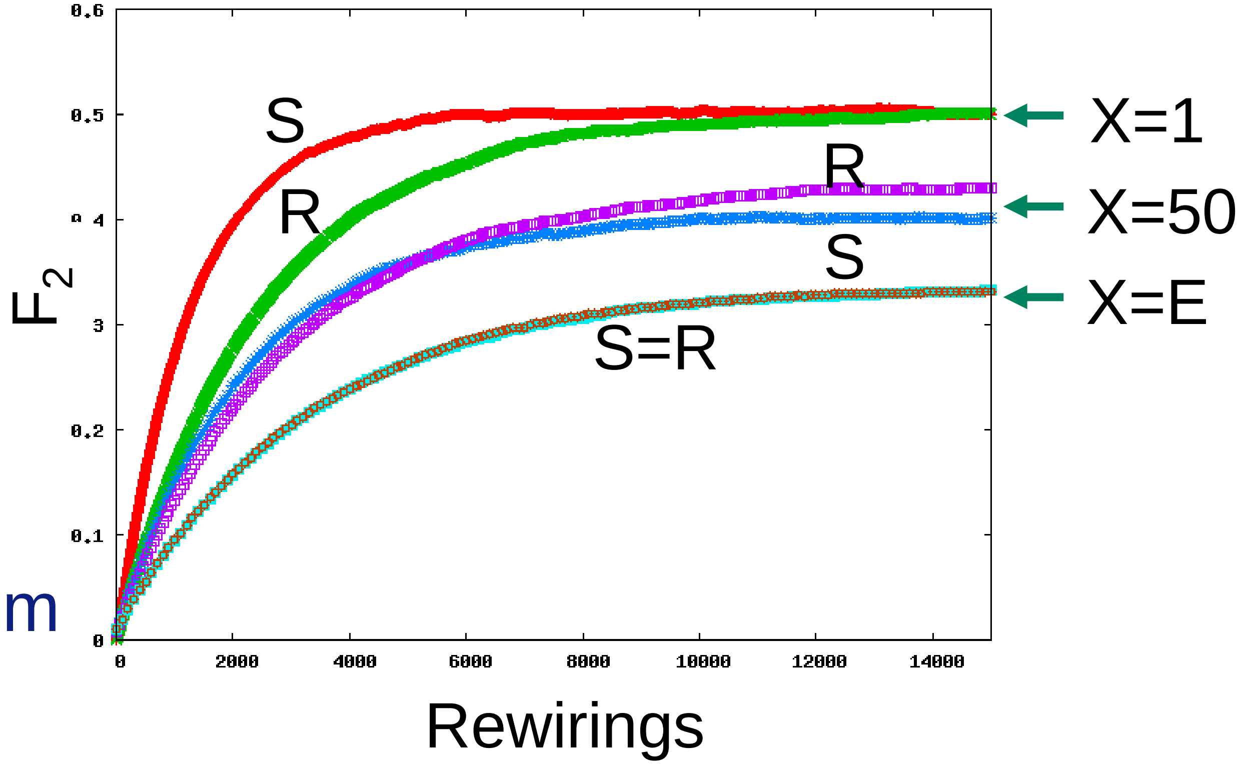
  X=E → Bentley *et al*. generational rewiring,

  Fisher-Wright Model

Alternate Update dynamics – Multi Edge Rewiring

- Numerical results:
- X=1: Equilibrium identical, but sequential update faster
- X=E/2: Similar timescales, but different Equilibrium F_2
- X=E: Sequential =Random

Sequential and Random updates.
$E=N=100$, $p_r=1/E$

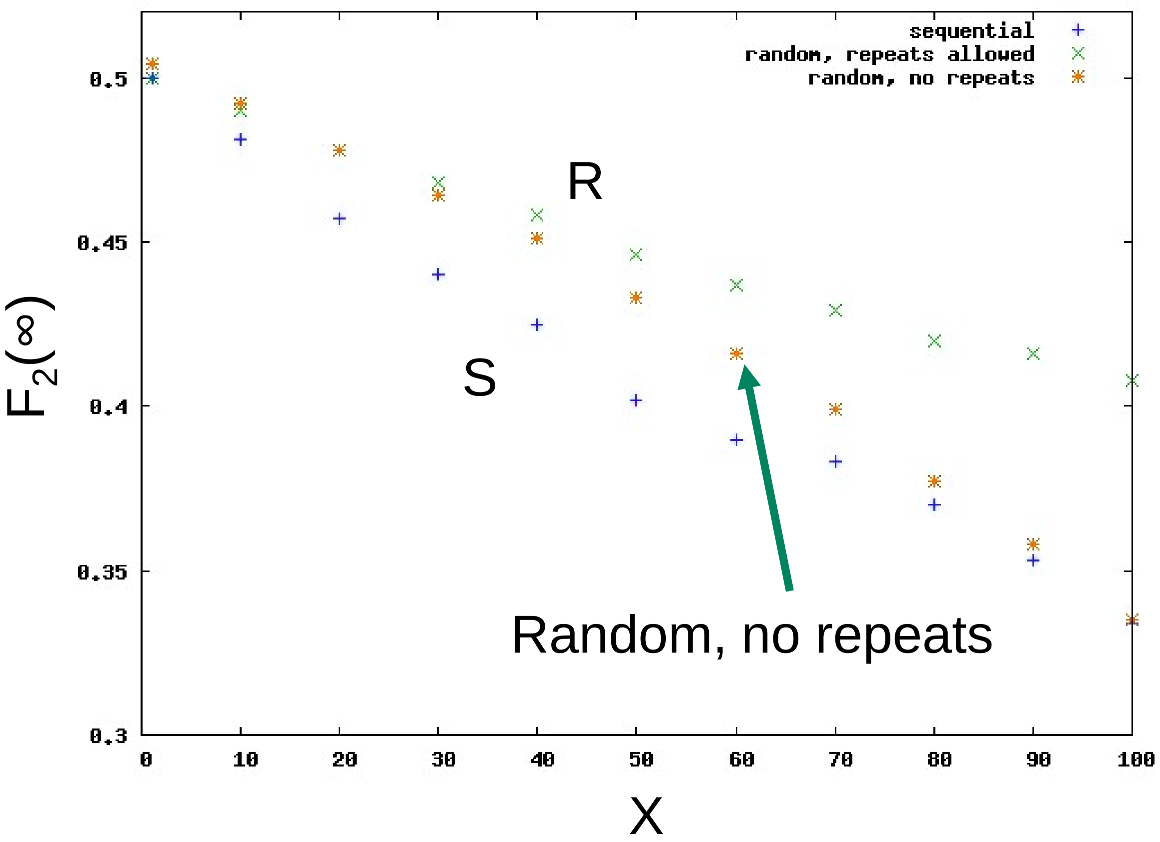


$$F_2(\infty) = \frac{p_p^2 + (1-p_p^2)\langle k \rangle}{p_p^2 + (1-p_p^2)E}, \qquad \lambda_2 = \frac{p_p^2(E-1)}{E}$$

Alternate Update dynamics – Multi Edge Rewiring

- Homogeneity decreases with number of edges rewired

E=N=100, pr=1/E



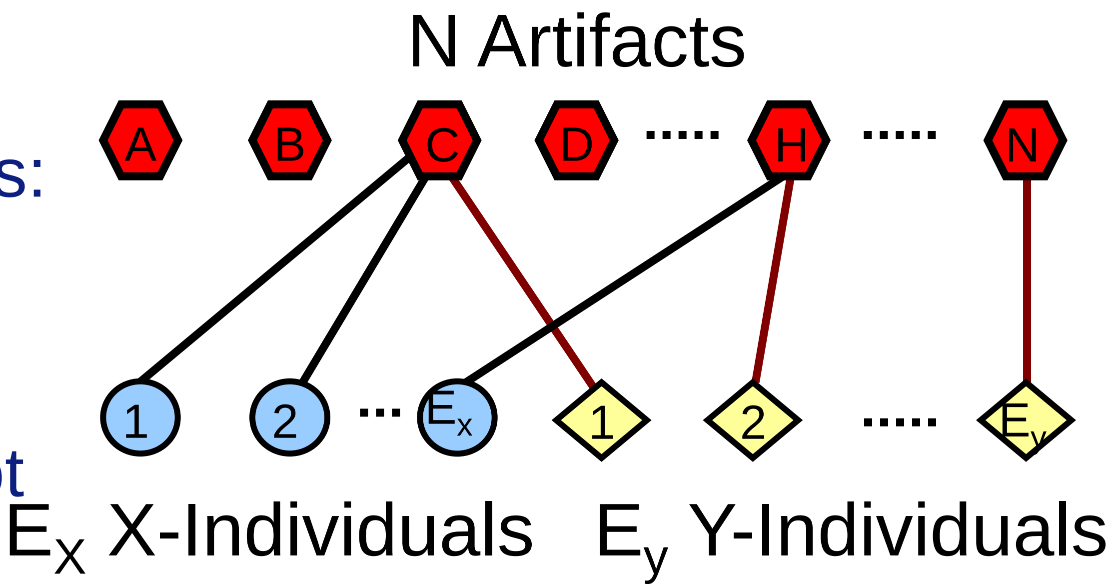
Two Individual Types

Artifacts have k_x and k_y degree and now four independent probabilities:

$$p_{rx}+p_{pxx}+p_{pxy}=1,$$

$$p_{ry}+p_{pyy}+p_{pyx}=1.$$

- Complete solutions not available



- Some progress can be made on F_{mn} measures, but lead to lengthy algebraic solutions

Summary

- Simple models can be solved Exactly
- Even with just copying and innovation, a large amount of variation in modelling is possible
- Rewiring dynamics is important
- Many other models can be mapped to a simple network model, and some properties studied analytically, e.g. Voter Model

References

- T.S.Evans & ADKP
  "Exact Solution for the Time Evolution of Network Rewiring Models"
  Phys. Rev. E **75** (2007) 056101
  [**`cond-mat/0612214`**]
- T.S.Evans & ADKP
  "Network Rewiring Models"
  Networks and Heterogeneous Media 3, 2 (2008) 221.
  `arXiv:0707.3783`
- T.S.Evans & ADKP
  "Exact Solutions for Models of Cultural Transmission and Network Rewiring" (for ECCS06)
  `[physics/0608052]`
- T.S.Evans, ADKP, & T.You (in prep)

Solutions – Time Dependence

- Now $n(k,t) = \frac{1}{k!} \frac{d^k G(z,t)}{dz^k}\bigg|_{z=0}$

- Must include contribution from all eigenvalues

$$G^{(m)}(z) = \sum_{k=0}^{E} z^k \omega^{(m)}(k)$$

- Cumbersome, but can in principle gives full time dependence exactly

$$\omega^{(m)}(k) = (-1)^m \frac{\Gamma(k+1)}{\Gamma(k+1-m)} \frac{\Gamma(c+k)}{\Gamma(c+k-m)} \frac{\Gamma(a)}{\Gamma(a+m)} \frac{\Gamma(b)}{\Gamma(b+m)}$$

$$\times {}_3F_2(-m, a+k, b+k; k+1-m, k+c-m; 1)\omega^{(0)}(k)$$