Random Graphs with Bounded Degrees

Eli Ben-Naim

Los Alamos National Laboratory

with: Paul Krapivsky (Boston)

E. Ben-Naim and P.L. Krapivsky

Phys. Rev. E **83**, 061102 (2011) & J. Stat. Mech. P11008 (2011) & EPL **97**, 48003 (2012)

arXiv:1102.5348 & arXiv:1110.1134 & arXiv:1112.0049

Talk, paper available from: http://cnls.lanl.gov/~ebn

IEEE NetSciCom 2012, March 30, 2012

Menu

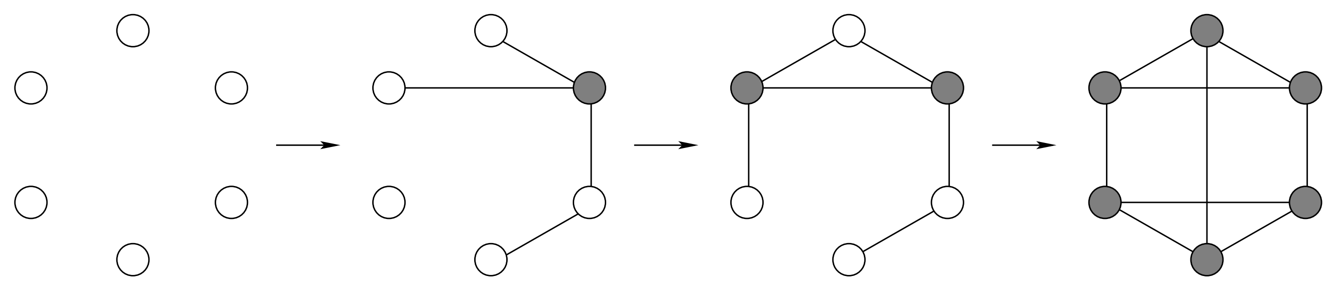
- Aperitif: Introduction
  - Motivation
  - Problem: Giant component in a regular random graph
  - Tutorial: Classical Random Graphs
- Main Course: Regular Random Graphs
  - Degree distribution
  - Emergence of Giant Component
  - Finite-size scaling Laws
- Dessert: Shuffling Algorithms and Rings

Motivation

In many network problems, the degree is bounded

- Social network: bounded number of friends Wasserman 88
- Power grids: transmission lines
- Communication networks: cellphone towers
- Computer networks Peleg 88
- Physics: bounded number of neighbors in a bead pack Girvan 10
- Chemistry: bounded number of chemical bonds in branched polymers Ballinska 91

Problem: Generating Regular Random Graphs

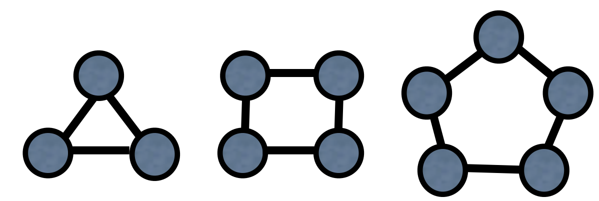


- Initial state: regular random graph (degree = 0)
- Define two classes of nodes
  - Active nodes: degree $< \mathrm{d}$
  - Inactive nodes: degree $= \mathrm{d}$
- Sequential linking
  - Pick two active nodes
  - Draw a link
- Final state: regular random graph (degree = d)

Erdos 60, Balinska 81, Wormald 84

Emergence of the Giant Component

✓ d=1 microscopic graphs, dimers 

✓ d=2 mesoscopic graphs, rings  $N_k = k^{-1}$

? d>2 one macroscopic graph = "giant component"

- *Dwarf component* phase: microscopic graphs only
- *Giant component* phase: one macroscopic component coexists with many microscopic graphs

Question

How many links (per node) are needed for the giant component to emerge?

Answer

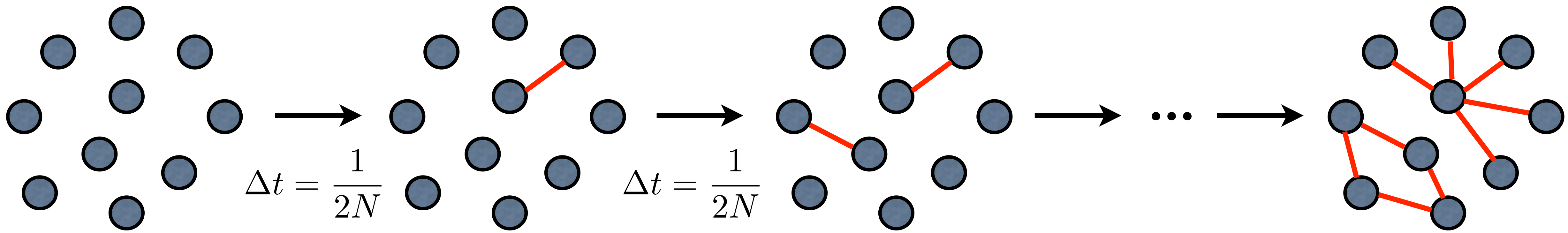
0.577200 (when d=3)

The kinetic approach

- Implicitly take the infinite system size limit
- Implicitly take an average over all realizations of the stochastic process
- Introduce the notion of continuous time variable
- For evolving graphs, time=number of links per node
- Describe the time evolution of probability distributions through differential equations
- Heavily used in physics, chemistry, biology

Discrete mathematics gone continuous!

Evolving Classical Random Graphs



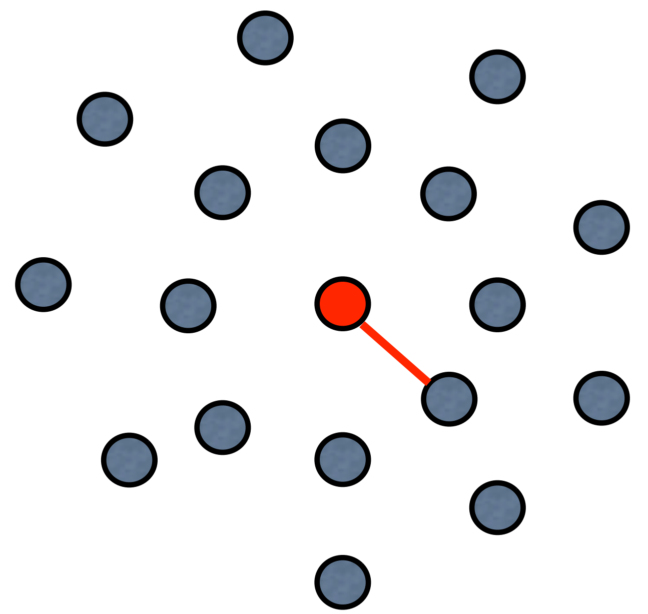
- Initial state: N isolated nodes
- Dynamical linking
    1. Pick 2 nodes at random
    2. Connect the 2 nodes with a link
    3. Augment time $t \to t + \frac{1}{2N}$
- Each node experiences one linking event per unit time

Chemical Physics: Flory, Stockmeyer 43
Graph Theory: Erdos, Renyi 60

Degree Distribution

- Distribution of nodes with degree *j* at time *t* is $n_j(t)$
- Linking Process is very simple

$$j \to j+1$$



- Linear evolution equation

$$\frac{dn_j}{dt} = n_{j-1} - n_j$$

- Initial condition: all nodes are isolated

$$n_j(t=0) = \delta_{j,0}$$

- Degree distribution is Poissonian

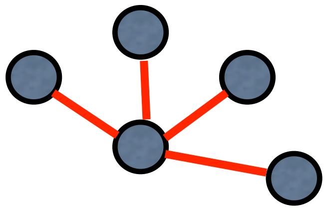
$$n_j(t) = \frac{t^j}{j!} e^{-t}$$

- Average degree characterizes the entire distribution

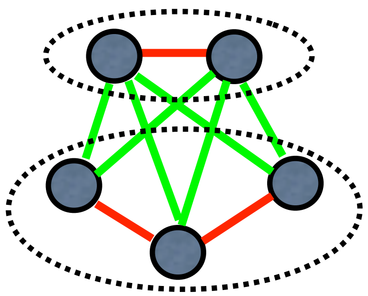
Random Process, Random Distribution

Aggregation Process

- Cluster = a connected graph component



- Aggregation rate = product of cluster sizes



$$K_{ij} = ij$$

- Master equation

$$\frac{dc_k}{dt} = \frac{1}{2}\sum_{i+j=k} ij c_i c_j - k c_k \qquad c_k(t=0) = \delta_{k,1}$$

- Cluster size density

$$c_k(t) = \frac{1}{k \cdot k!}(kt)^{k-1} e^{-kt}$$

- Divergent second moment = emergence of giant component

$$M_2 = \sum_k k^2\, c_k \qquad M_2 = (1-t)^{-1} \qquad t_g = 1$$

Ziff 82
Leyvraz 03

Detecting the giant component

- Master equation

$$\frac{dc_k}{dt} = \frac{1}{2}\sum_{i+j=k} ijc_ic_j - kc_k \qquad c_k(t=0) = \delta_{k,1}$$

- Moments of the size distribution

$$M_n(t) = \sum_k k^n c_k(t)$$

- First moment is conserved

$$\frac{dM_1}{dt} = M_2(M_1 - 1) = 0 \quad \text{when} \qquad M_n(t=0) = 1$$

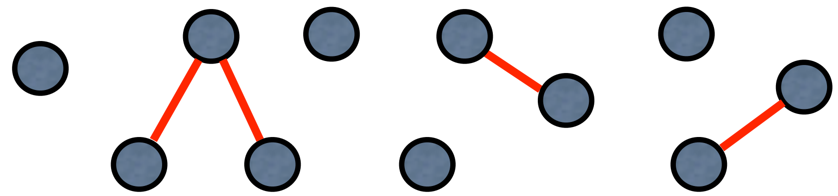
- Second moment obeys closed equation

$$\frac{dM_2}{dt} = M_2^2 \qquad M_2(0) = 1$$

- Finite-time singularity signals emergence of infinite cluster

$$M_2 = (1-t)^{-1}$$

Dwarf Component Phase (t<1)



- Microscopic clusters, tree structure
- Cluster size distribution contains <u>entire</u> mass

$$M(t) = \sum_{k=1}^{\infty} k\, c_k = 1$$

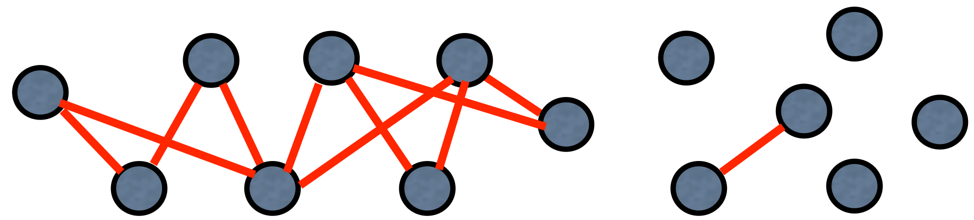
- Typical cluster size diverges near percolation point

$$k_* \sim (1-t)^{-2}$$

- Critical size distribution has power law tail

$$c_k(1) \simeq \frac{1}{\sqrt{2\pi}} k^{-5/2}$$

Giant Component Phase (t>1)



- Macroscopic component exist, complex structure
- Cluster size distribution contains fraction of mass

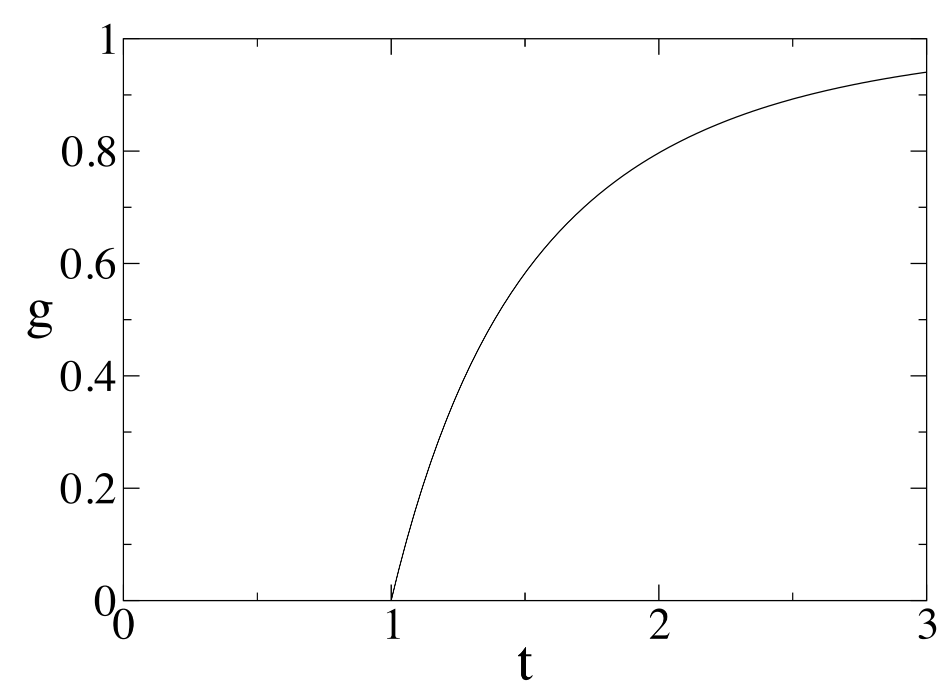
$$M(t) = \sum_{k=1}^{\infty} k c_k = 1 - g$$

- Giant component accounts for "missing" mass

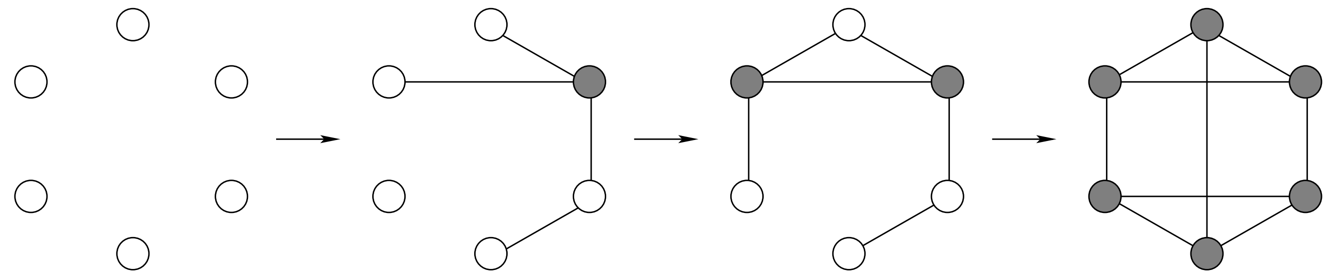
$$g = 1 - e^{-gt}$$

- Giant component takes over entire system

$$g \to 1$$



Generating a Regular Random Graph

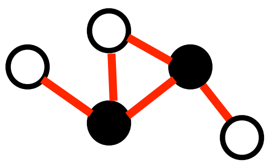


- Initial state: regular random graph (degree = 0)
- Define two classes of nodes
  - Active nodes: degree $< d$
  - Inactive nodes: degree $= d$
- Sequential linking
  - Pick two active nodes
  - Draw a link
- Final state: regular random graph (degree = d)

Erdos 60, Balinska 81, Wormald 84

Degree Distribution

- Distribution of nodes with degree *j* is n_j
- Density of active nodes $\nu = n_0 + n_1 + \cdots + n_{d-1}$ $\nu = 1 - n_d$
- Linking Process



$$(i,j) \to (i+1, j+1) \qquad i,j < d$$

- Active nodes control linking process, effectively <u>linear</u> evolution equation

$$\frac{dn_j}{dt} = \nu\,(n_{j-1} - n_j) \xrightarrow{\tau = \int_0^t dt'\,\nu(t')} \frac{dn_j}{d\tau} = n_{j-1} - n_j$$

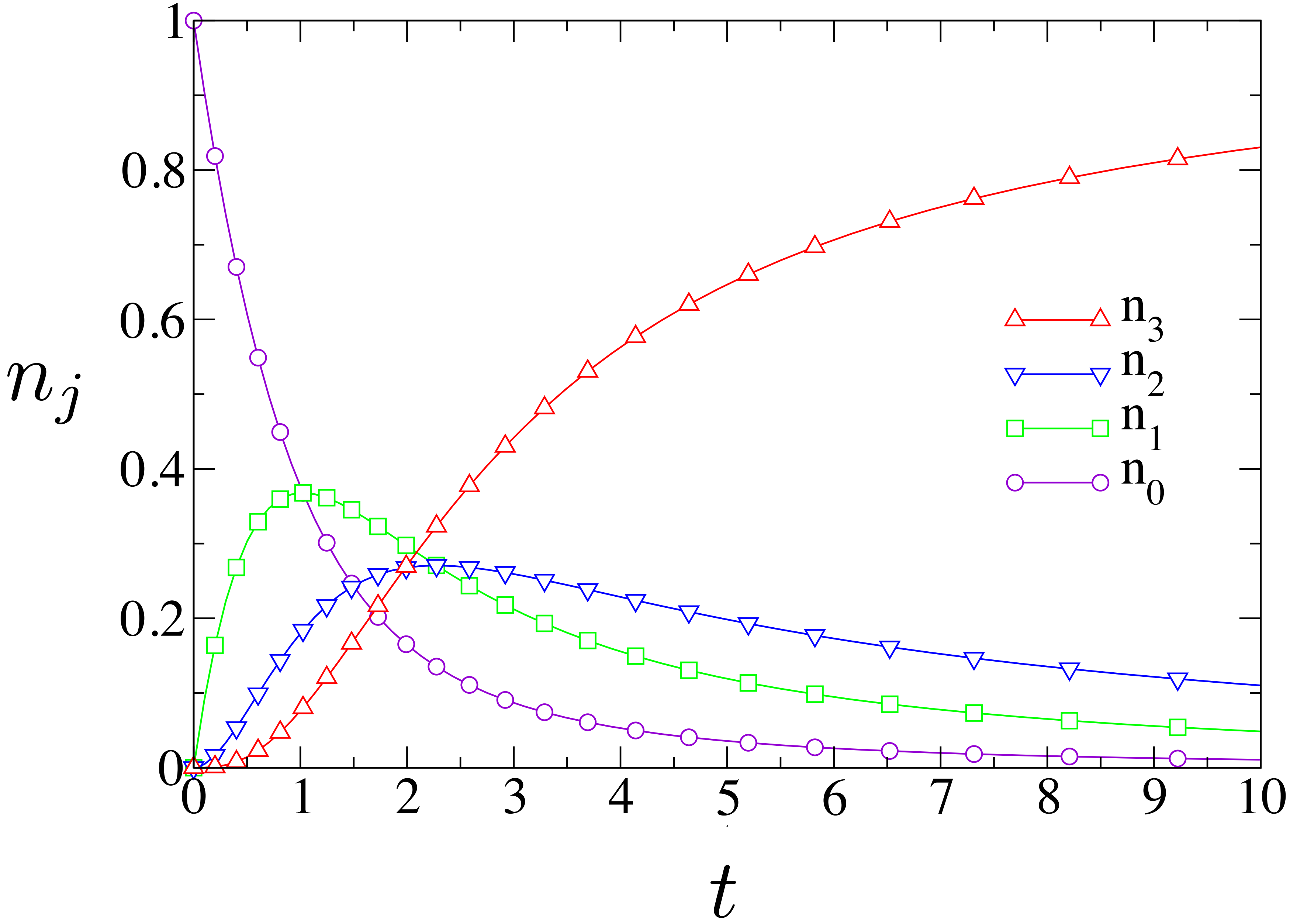
- Solve using an effective time variable

$$n_j = \frac{\tau^j}{j!} e^{-\tau} \qquad j < d$$

Truncated Poisson Distribution

Wormald 99

Degree Distribution

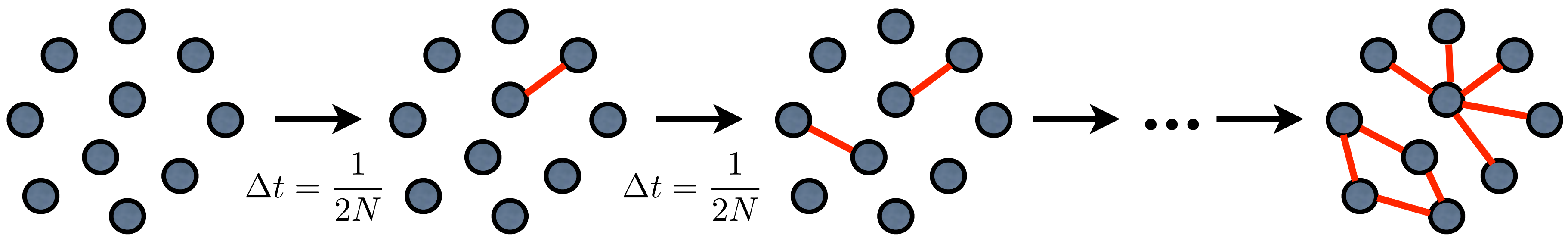


Isolated nodes dominate initially

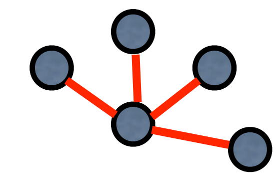
All nodes become inactive eventually

Unbounded Random Graphs

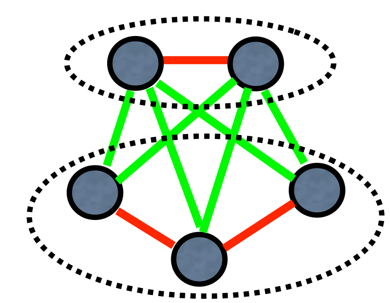
Erdos-Renyi



- Cluster = a connected graph component



- Links involving two separate components lead to merger
- Aggregation rate = product of cluster sizes



$$K_{ij} = ij$$

- Master equation for size distribution

$$\frac{dc_k}{dt} = \frac{1}{2}\sum_{i+j=k} ij\,c_i c_j - k c_k \qquad c_k(t=0) = \delta_{k,1}$$

- Master equation for generating function

$$\frac{\partial \mathcal{C}}{\partial t} + x\,\frac{\partial \mathcal{C}}{\partial x} = \frac{1}{2}\left(x\,\frac{\partial \mathcal{C}}{\partial x}\right)^2 \qquad \mathcal{C}(x,t) = \sum_k c_k(t)\,x^k$$

Hamilton-Jacobi Theory I

- Master equation is a first-order PDE

$$\frac{\partial \mathcal{C}}{\partial t} + x\,\frac{\partial \mathcal{C}}{\partial x} = \frac{1}{2}\left(x\,\frac{\partial \mathcal{C}}{\partial x}\right)^2$$

$\mathcal{C}(x,0) = x$

- Recognize as a Hamilton-Jacobi equation

$$\frac{\partial \mathcal{C}(x,t)}{\partial t} + H(x,p) = 0$$

- By identifying "momentum" and "Hamiltonian"

$$p = \frac{\partial \mathcal{C}}{\partial x} \qquad \text{and} \qquad H = xp - \frac{1}{2}(xp)^2$$

- Hamilton-Jacobi equations generate two coupled ODEs

$$\frac{dx}{dt} = \frac{\partial H}{\partial p}, \quad \frac{dp}{dt} = -\frac{\partial H}{\partial x} \quad \Longrightarrow \quad \frac{dx}{dt} = x(1-xp), \quad \frac{dp}{dt} = -p(1-xp)$$

$x(0) = 1 - g \quad p(0) = 1$

Initial coordinate unknown, final coordinate known!

Hamiltonian is a conserved quantity

Solution I

- Coordinate and momentum are immediate

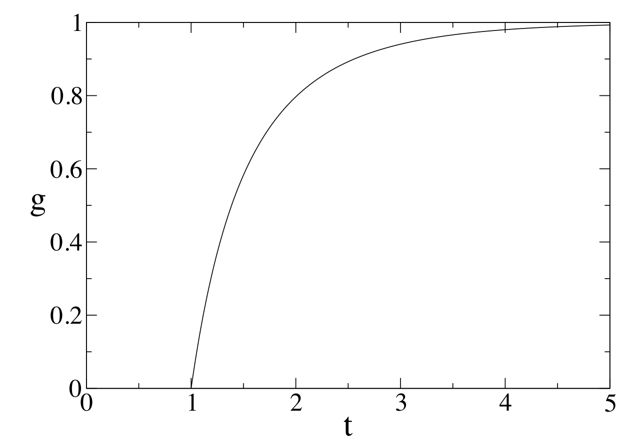
$$x = (1-g)e^{gt} \qquad p = e^{-gt}$$

- Size of giant component found immediately

$$g = 1 - \sum_k k\, c_k = 1 - p(0)$$

- Satisfies a closes equation

$$1 - g = e^{-gt}$$



- Nontrivial solution beyond the percolation threshold

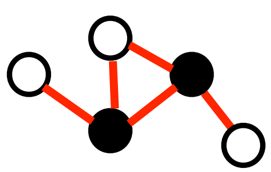
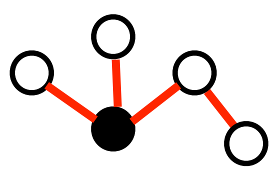
$$t_g = 1$$

The giant component emerges when the average degree equals one

Bounded Random Graphs

- Total size of components provides insufficient description
- Describe components by a d+1 dimensional vector whose components specify number of nodes with given degree

$$(k_0, k_1, \cdots, k_d) \qquad k = k_0 + k_1 + \cdots + k_d$$

$(0, 2, 1, 2)$   $(0, 3, 1, 1)$

- Multivariate aggregation process
- Aggregation rate is product of the number of active nodes

$$K(\mathbf{l}, \mathbf{m}) = (l - l_d)(m - m_d)$$

- Why can't we get away with two variables only?
- Node degrees are coupled!

$$0 \to 1 \to 2 \to 3$$

Hamilton-Jacobi Theory II

- Master equation is a first-order PDE

$$\frac{\partial C}{\partial \tau} = \frac{1}{2\nu}\left(\sum_{j=0}^{d-1} x_{j+1}\frac{\partial C}{\partial x_j}\right)^2 - \sum_{j=0}^{d-1} x_j \frac{\partial C}{\partial x_j} \qquad C(\mathbf{x},0) = x_0$$

- Recognize as a Hamilton-Jacobi equation

$$\frac{\partial C(\mathbf{x},\tau)}{\partial \tau} + H(\mathbf{x}, \nabla C, \tau) = 0$$

- By identifying "momentum" and "Hamiltonian"

$$H(\mathbf{x},\mathbf{p},\tau) = \sum_{j=0}^{d-1} x_j p_j - \frac{\Pi_1^2}{2\nu(\tau)} \qquad \Pi_j = \sum_{i=j}^{d} x_i\, p_{i-j}$$

- Hamilton-Jacobi equation give $2(\mathrm{d}+1)$ coupled ODEs

$$\frac{dx_j}{dt} = \frac{\partial H}{\partial p_j}, \quad \frac{dp_j}{dt} = -\frac{\partial H}{\partial x_j} \quad \Longrightarrow \quad \frac{dx_j}{dt} = x_j - \frac{\Pi_1}{\nu}x_{j+1}, \quad \frac{dp_j}{dt} = \frac{\Pi_1}{\nu}p_{j-1} - p_j$$

Initial coordinates unknown, final coordinates known!

Equations are now in $\mathrm{d}+1$ dimensions!

Hamiltonian no longer conserved!

Hamilton-Jacobi Equations

- Coupled differential equations for coordinate and momenta

$$\frac{dx_j}{d\tau} = \begin{cases} x_j - \frac{\Pi_1}{\nu}\, x_{j+1} & j < d \\ 0 & j = d \end{cases} \quad \text{and} \quad \frac{dp_j}{d\tau} = \frac{dx_j}{d\tau} = \begin{cases} x_j - \frac{\Pi_1}{\nu}\, x_{j+1} & j < d \\ 0 & j = d \end{cases}$$

- Initial conditions: (i) known for momenta (ii) unknown coordinates!

$$p_j(0) = \delta_{j,0} \qquad \text{and} \qquad x_j(0) = y_j \qquad\qquad C(\mathbf{x}, 0) = x_0$$

- Identify conservation laws!

$$\frac{d\Pi_0}{d\tau} = 0 \qquad \text{and} \qquad \frac{dx_d}{d\tau} = 0 \qquad\qquad \Pi_j = \sum_{i=j}^{d} x_i\, p_{i-j}$$

- Backward evolution equations for the initial coordinates!

$$\frac{dy_j}{d\tau} = \sum_{i=0}^{d-j-1} \left[\frac{du}{d\tau}\, x_{i+j+1} - x_{i+j} \right] p_i \qquad\qquad u = \int_0^{\tau} d\tau' \, \frac{\Pi_1(\tau')}{\nu(\tau')}$$

Solution II

- Find hidden <u>conservation laws</u> and explicit backward equations
- reduce 2(d+1) first order ODE to 1 second order ODE

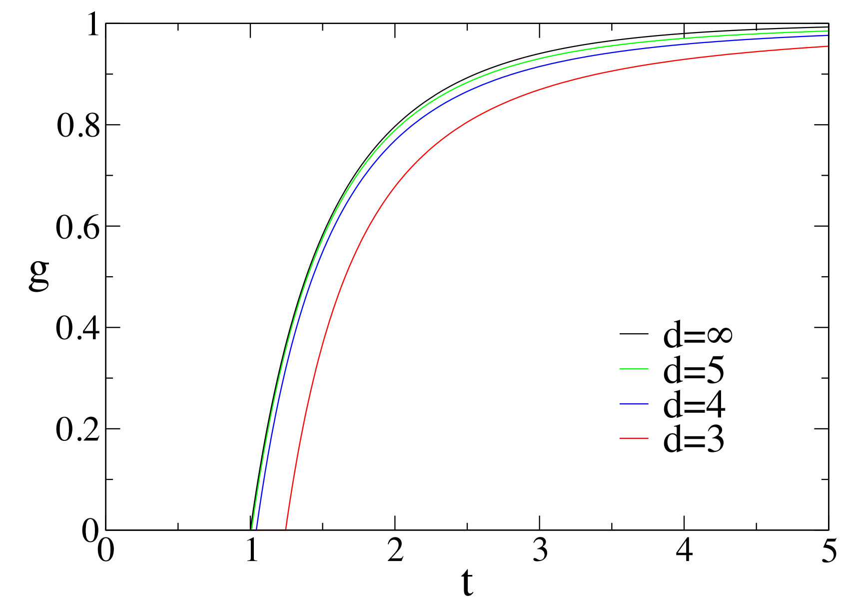
$$\frac{d^2u}{d\tau^2} + \frac{n_{d-1}}{\nu}\frac{du}{d\tau} - x_d\frac{p_{d-1}}{\nu} = 0 \qquad u = \int_0^\tau d\tau' \, \frac{\Pi_1(\tau')}{\nu(\tau')}$$

- Nontrivial solution when d>2
- Numerical solution gives percolation threshold (d=3)

$$t_g = 1.243785, \quad L_g = 0.577200$$

- The size distribution of components at the critical point

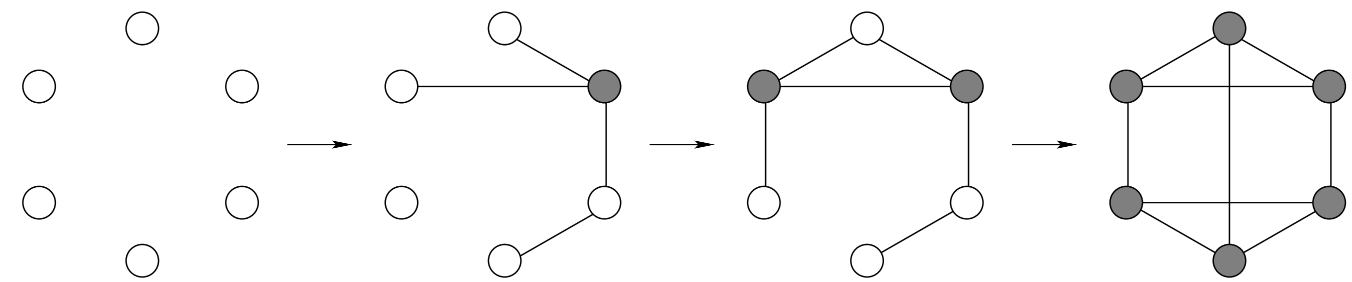
$$c_k \simeq A\, k^{-5/2}$$



- Mean-field percolation

Hamilton-Jacobi theory gives all percolation parameters

Finite-size scaling



Degree distribution

$$n_j \simeq \frac{(d-1)!}{j!} t^{-1} (\ln t)^{-(d-1-j)}$$

Regular random graph emerges in several steps

1. Giant component emerges at finite time

$$t_1 = 1.243785$$

deterministic

2. Graph becomes fully connected emerges at time

$$N n_0 \sim 1 \Longrightarrow \qquad t_2 \sim N (\ln N)^{-(d-1)}$$

stochastic

3. Regular random graph emerges at time

$$N n_{d-1} \sim 1 \Longrightarrow \qquad t_3 \sim N$$

stochastic

Giant fluctuations in completion time

General Random Graphs

- Theory straightforward to generalize
- Degree controls linking process

$$(i, j) \to (i+1, j+1) \text{ with rate } C_{i,j}$$

- Connection rate is arbitrary
- Equation for generating function

$$\frac{\partial C}{\partial t} = \frac{1}{2} \sum_{i,j} C_{i,j} \left[\left(x_{i+1} \frac{\partial C}{\partial x_i} \right) \left(x_{j+1} \frac{\partial C}{\partial x_j} \right) - 2n_i \left(x_j \frac{\partial C}{\partial x_j} \right) \right]$$

- Hamiltonian

$$H(\mathbf{x}, \mathbf{p}, t) = \sum_j \nu_j(t) x_j p_j - \frac{1}{2} \sum_{i,j} C_{i,j} (x_{i+1} p_i)(x_{j+1} p_j)$$

Conservation laws neither obvious nor guaranteed

Multi-dimensional Newton solver

Summary

- Dynamic formation of regular random graphs
- Degree distribution is truncated Poissonian
- Hamilton-Jacobi formalism powerful
- Percolation parameters with essentially arbitrary precision
- Mean-field percolation universality class
- A multitude of finite-size scaling properties
- Giant fluctuations in completion time

Theory applicable to broader set of evolving graphs

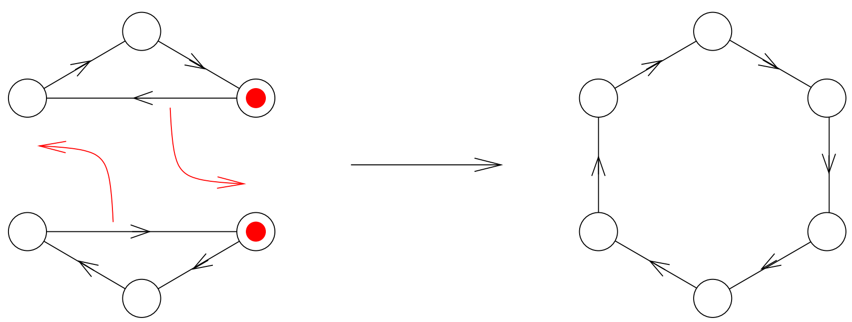
Shuffling Algorithm

$$1\,\underline{2}\,3\,4\,\underline{5}\,6 \rightarrow 1\,5\,\underline{3}\,\underline{4}\,2\,6 \rightarrow 1\,5\,4\,3\,2\,6 \rightarrow \cdots$$

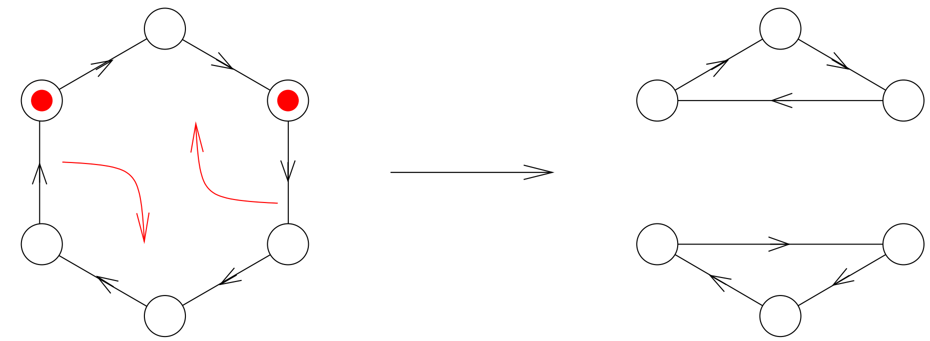
- Initial configuration: N ordered integers
- Pairwise shuffling:
  1. Pick 2 numbers at random
  2. Exchange positions
  3. Augment time $t \rightarrow t + \frac{1}{2N}$
- Each integer is shuffled once per unit time
- Efficient algorithm, computational cost is $\mathcal{O}(N)$

Isomorphic to dynamical regular random graph with d=2!

Cycles and Permutations



$$(1\underline{2}3)(4\underline{5}6) \to (156423)$$



$$(1\underline{5}64\underline{2}3) \to (123)(456)$$

- Cycle structure of a permutation

$$134265 \quad \Longrightarrow \quad (1)(234)(56)$$

- **Aggregation**: <u>inter</u>-cycle shuffling

$$i, j \xrightarrow{K_{ij}} i + j \qquad \text{with} \qquad K_{ij} = ij$$

- **Fragmentation**: <u>intra</u>-cycle shuffling

$$i + j \xrightarrow{F_{ij}} i, j \qquad \text{with} \qquad F_{ij} = \frac{i+j}{N}$$

Identical aggregation and fragmentation rates

Steady-State Distribution

- Steady-state size distribution satisfies

$$0 = \frac{1}{2} \sum_{i+j=k} K_{ij}\, c_i\, c_j - c_k \sum_{j \geq 1} K_{kj}\, c_j + \sum_{j \geq 1} F_{kj}\, c_{j+k} - \frac{1}{2} c_k \sum_{i+j=k} F_{ij}$$

- Detailed balance condition

$$K_{ij}\, c_i\, c_j = F_{ij}\, c_{i+j}$$

Lowe 95

- Substitute aggregation and fragmentation rates

$$K_{ij} = ij \qquad F_{ij} = \frac{i+j}{N}$$

- Steady-state solution

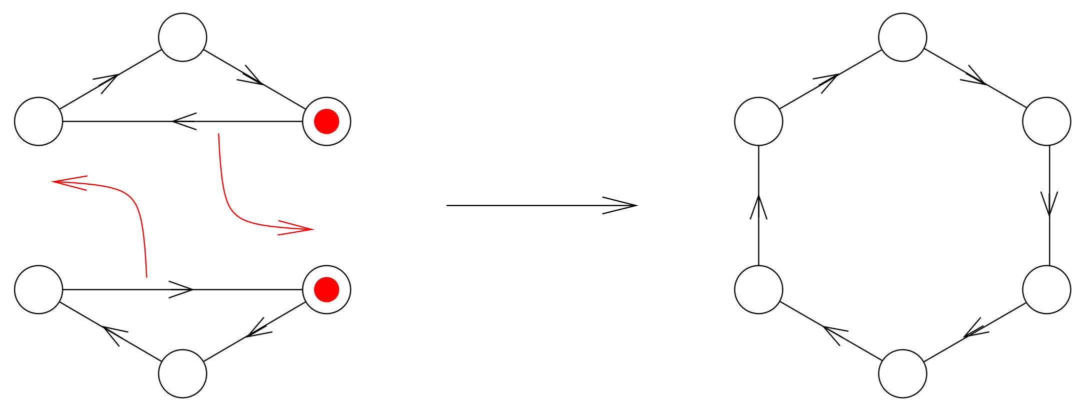
$$(ic_i)(jc_j) = \frac{1}{N}(i+j)c_{i+j} \quad \Longrightarrow \quad Nc_k = \frac{1}{k}$$

- Average number of cycles

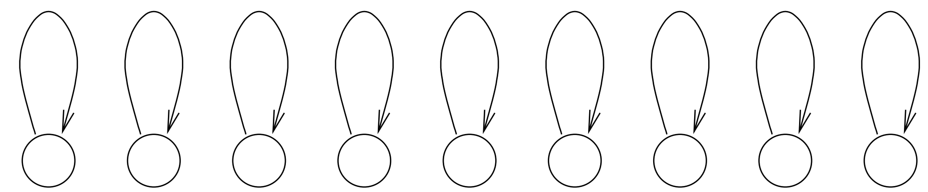
$$N_k = \frac{1}{k}$$

Kunth, art of comp. prog. vol 3

Redirection Process

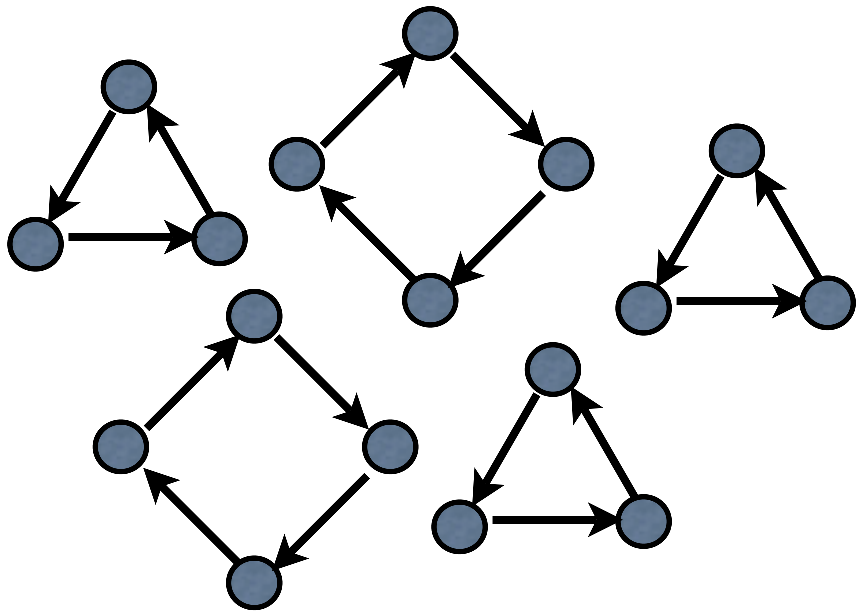


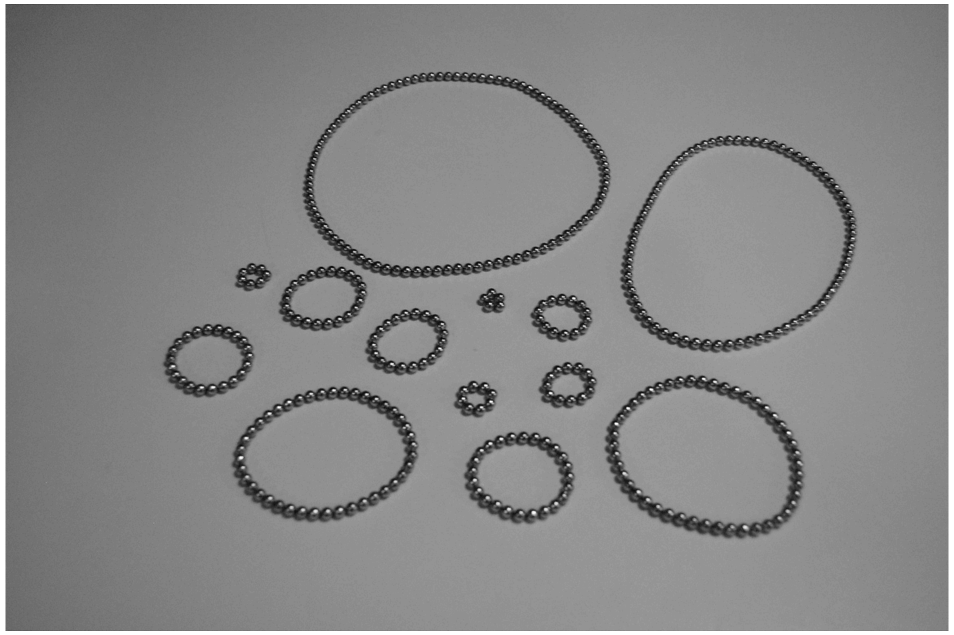
- Dynamical redirection
  1. Pick 2 nodes at random
  2. Connect 2 nodes by redirecting 2 associated links
  3. Augment time $t \to t + \dfrac{1}{2N}$
- A node experiences one redirection event per unit time
- Initial condition: isolated nodes, each has a self-link



Redirection process maintains ring topology

Rings





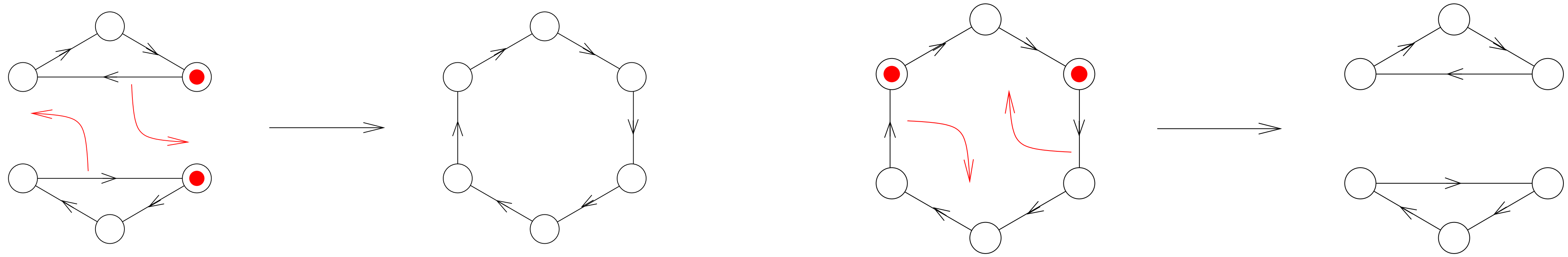
Talia Ben-Naim, age 9

- All nodes have identical degree
- Motivation: rings of magnetic particles
- Consider simplest case: rings; all nodes have degree 2
- Consider directed links (without loss of generality)
- In a system of N nodes, there are exactly N links

Number of links is conserved!

Kun 01

Aggregation-Fragmentation Process



- **Aggregation**: inter-ring redirection

Identical to random graph process

$$i, j \xrightarrow{K_{ij}} i + j \qquad \text{with} \qquad K_{ij} = ij$$

- **Fragmentation**: intra-ring redirection

Fragmentation rate depends on system size!

$$i + j \xrightarrow{F_{ij}} i, j \qquad \text{with} \qquad F_{ij} = \frac{i + j}{N}$$

- Total fragmentation rate is quadratic

$$F_k = \sum_{i+j=k} F_{ij} = \frac{k(k-1)}{2N}$$

Reversible process

Rate Equations

- Size distribution satisfies

$$\frac{dr_k}{dt} = \frac{1}{2}\sum_{i+j=k} ij\, r_i r_j - k\, r_k + \frac{1}{N}\left[\sum_{j>k} j\, r_j - \frac{k(k-1)}{2} r_k\right]$$

- Rate equation includes explicit dependence on N
- Perturbation theory

$$r_k = \underbrace{f_k}_{\text{finite rings}} + \frac{1}{N}\underbrace{g_k}_{\text{giant rings}}$$

- Fragmentation irrelevant for finite rings $F_k \sim \dfrac{k^2}{N}$

$$\frac{df_k}{dt} = \frac{1}{2}\sum_{i+j=k} ij\, f_i f_j - k\, f_k$$

Recover random graph equation

Finite Rings Phase (t<1)

- All rings are finite in size

$$M(t) = \sum_{k=1}^{\infty} f_k = 1$$

- Size distribution

$$f_k(t) = \frac{1}{k \cdot k!}(kt)^{k-1} e^{-kt}$$

- Second moment diverges in finite time $M_2 = \sum_k k^2 f_k$

$$\frac{dM_2}{dt} = M_2^2 \quad \Longrightarrow \quad M_2 = (1-t)^{-1}$$

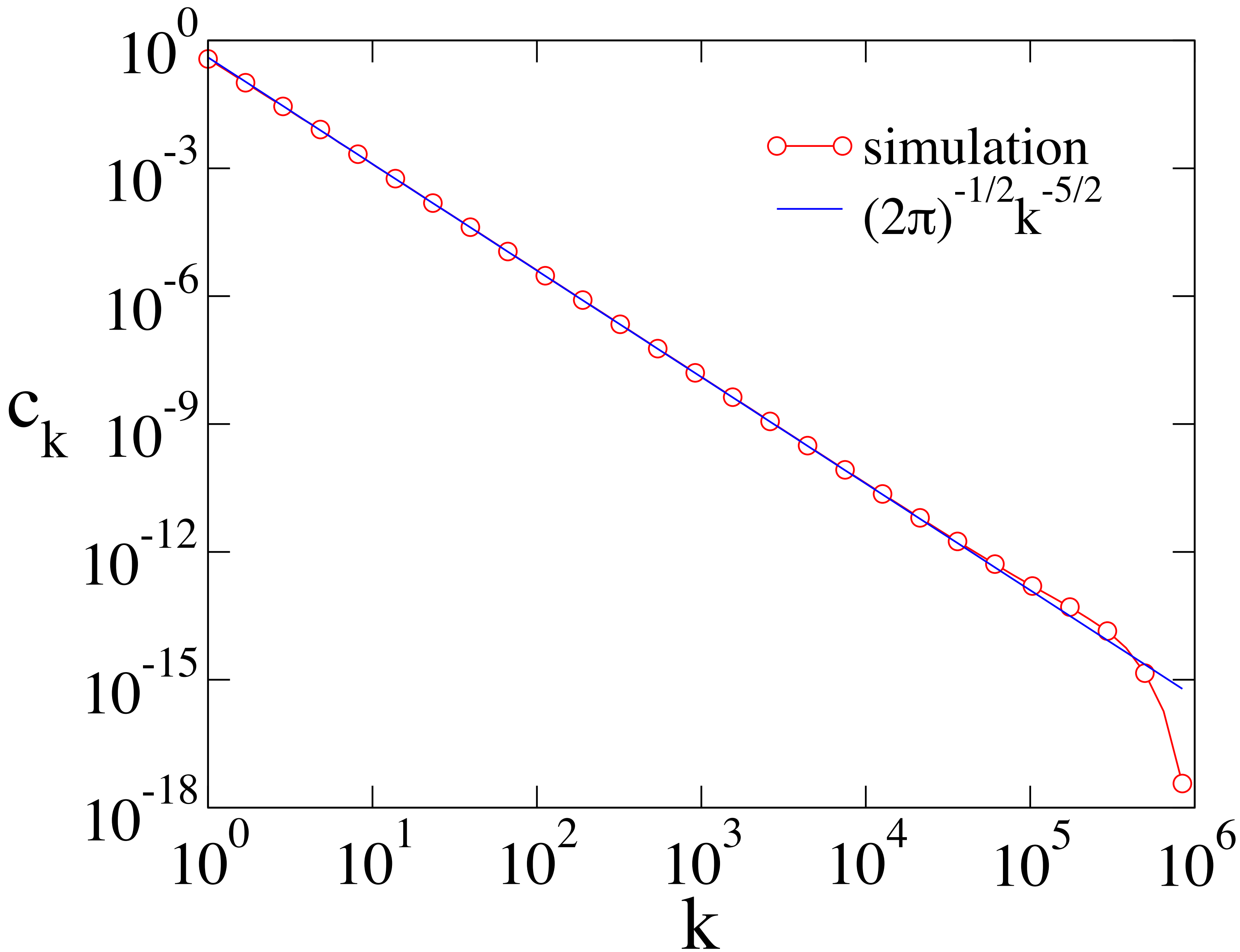
- Critical size distribution

$$f_k(1) \simeq \frac{1}{\sqrt{2\pi}} k^{-5/2}$$

Identical behavior to good-old random graph

Critical Size Distribution

Simulation results



Excellent agreement between theory and simulation

Giant Rings Phase (t>1)

- Finite rings contain only a fraction of g all mass

$$M(t) = \sum_{k=1}^{\infty} k f_k = 1 - g$$

- "Missing Mass" 1-g must be found in giant rings

$$g = 1 - e^{-gt}$$

- Expect giant, macroscopic rings
- Very fast aggregation and fragmentation processes

$$F_k \sim \frac{k^2}{N} \sim N \quad \text{when} \quad k \sim N$$

Fragmentation comparable to aggregation
No longer negligible

Distribution of giant rings

- Quantify giant rings by normalized size $\ell = \dfrac{k}{N}$
- Average number of giant rings of normalized size ℓ

$$g(t) = \int_0^{g(t)} d\ell\, \ell\, G(\ell, t)$$

- Rate equation

agg gain= $\ell/2$ agg loss= $g - \ell$

$$\cancel{\frac{1}{N}\frac{\partial G(\ell,t)}{\partial t}} = \frac{1}{2}\int_0^{\ell} ds\, s(\ell - s)G(s,t)G(\ell - s, t) - \ell(g - \ell)G(\ell, t)$$

frag gain= $g - \ell$ frag loss= $\ell/2$

$$+ \int_\ell^g ds\, s\, G(s,t) - \frac{1}{2}\ell^2 G(\ell, t)$$

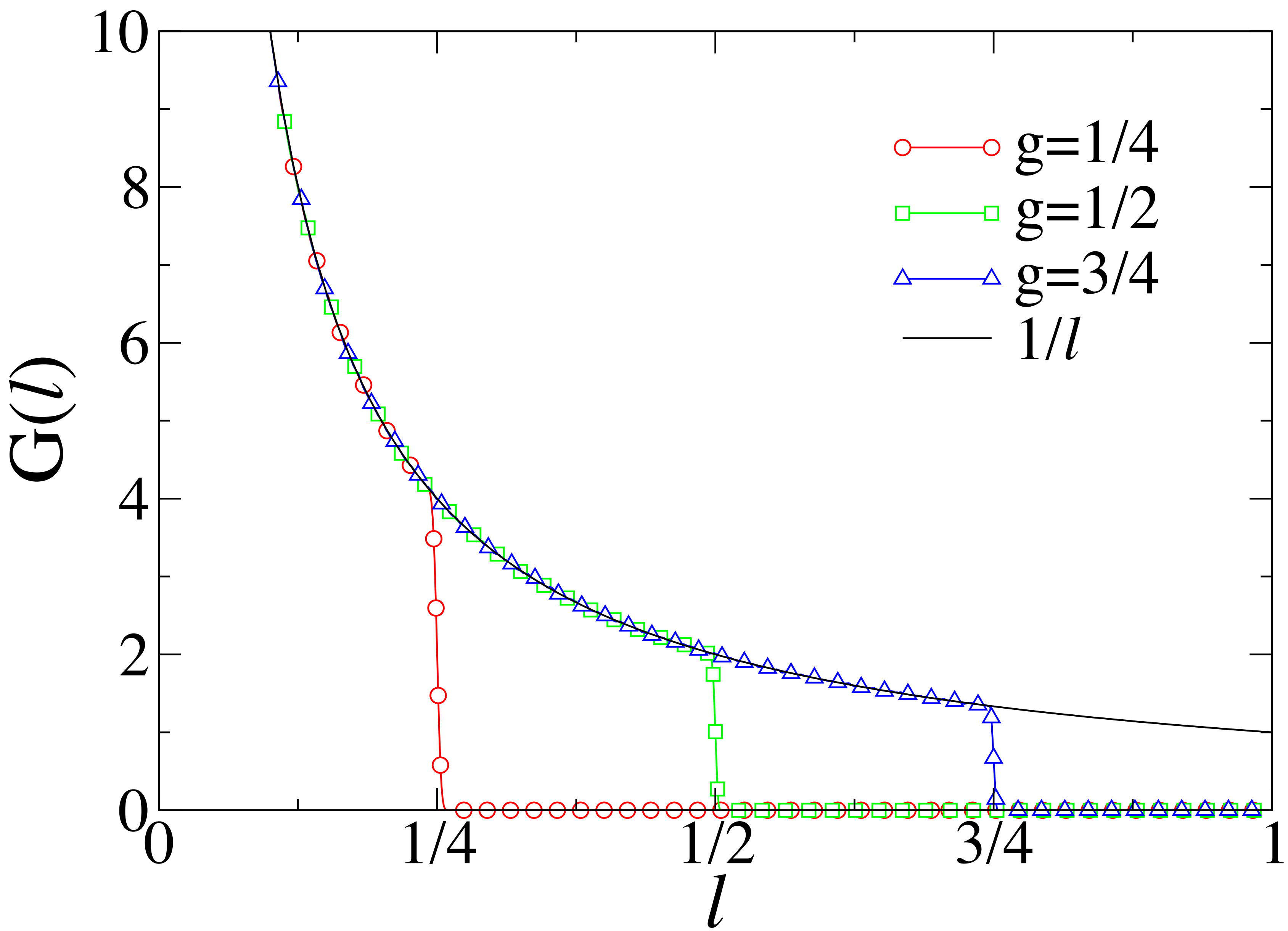
- Quasi steady-state

$$G(\ell, t) = \begin{cases} \ell^{-1} & \ell < g(t), \\ 0 & \ell > g(t). \end{cases}$$

Universal distribution, span grows with time

Average Number of Giant Rings

Simulation results



$$G(\ell, t) = \begin{cases} \ell^{-1} & \ell < g(t), \\ 0 & \ell > g(t). \end{cases}$$

Comments

- Rate equation for <u>average number</u> of giant rings

$$\cancel{\frac{1}{N}\frac{\partial G(\ell,t)}{\partial t}} = \frac{1}{2}\int_0^\ell ds\, s(\ell-s)G(s,t)G(\ell-s,t) - \ell(g-\ell)G(\ell,t)$$
$$+ \int_\ell^g ds\, s\, G(s,t) - \frac{1}{2}\ell^2 G(\ell,t)$$

- Practically closed equation; coupling to finite rings only through total mass $g(t)$

- Steady flux $N\, dg/dt$ from finite rings to giant rings

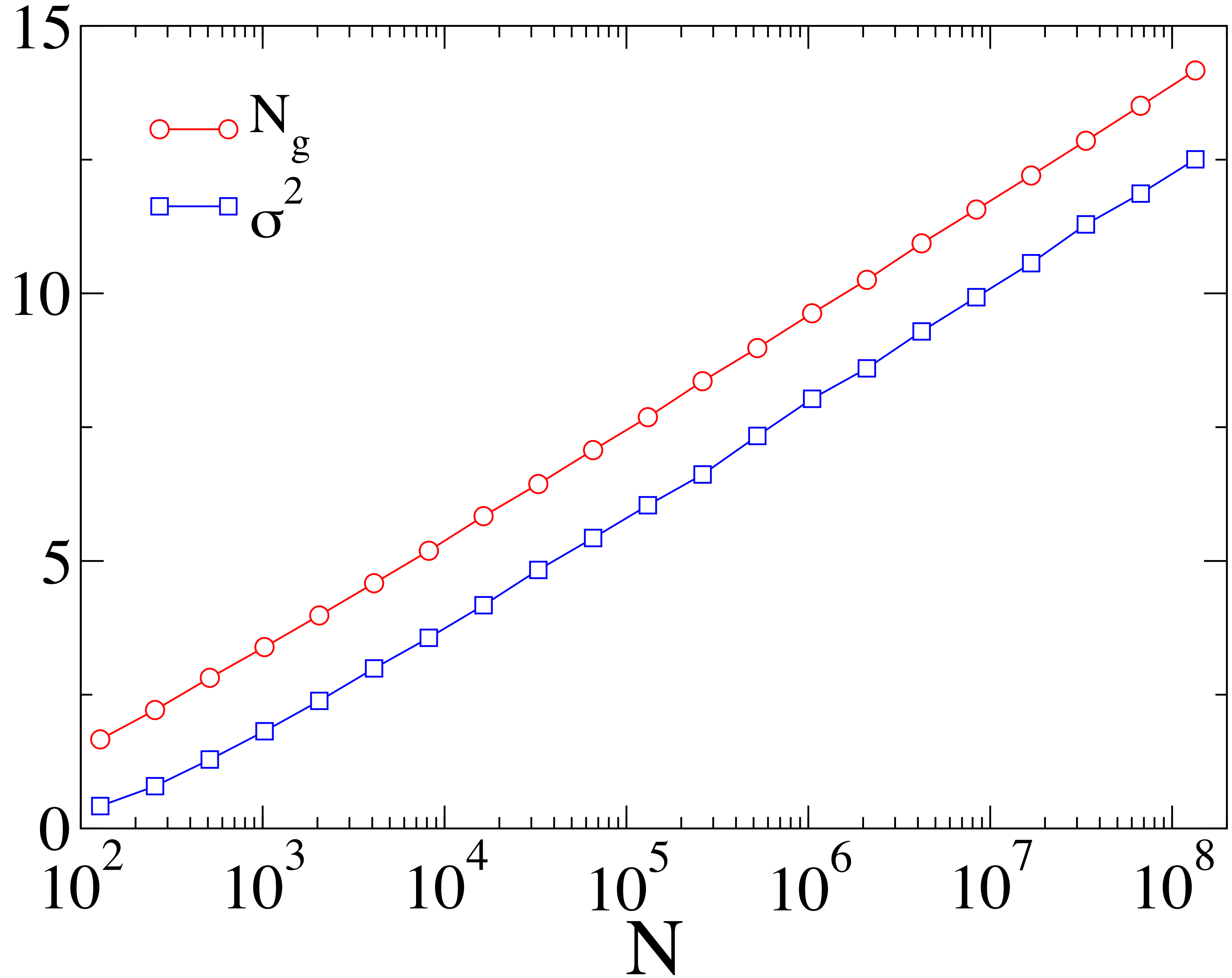
- Number of giant rings is not proportional to N!

$$N_g \simeq \ln N$$

Number of microscopic rings proportional to N
Number of macroscopic rings logarithmic in N

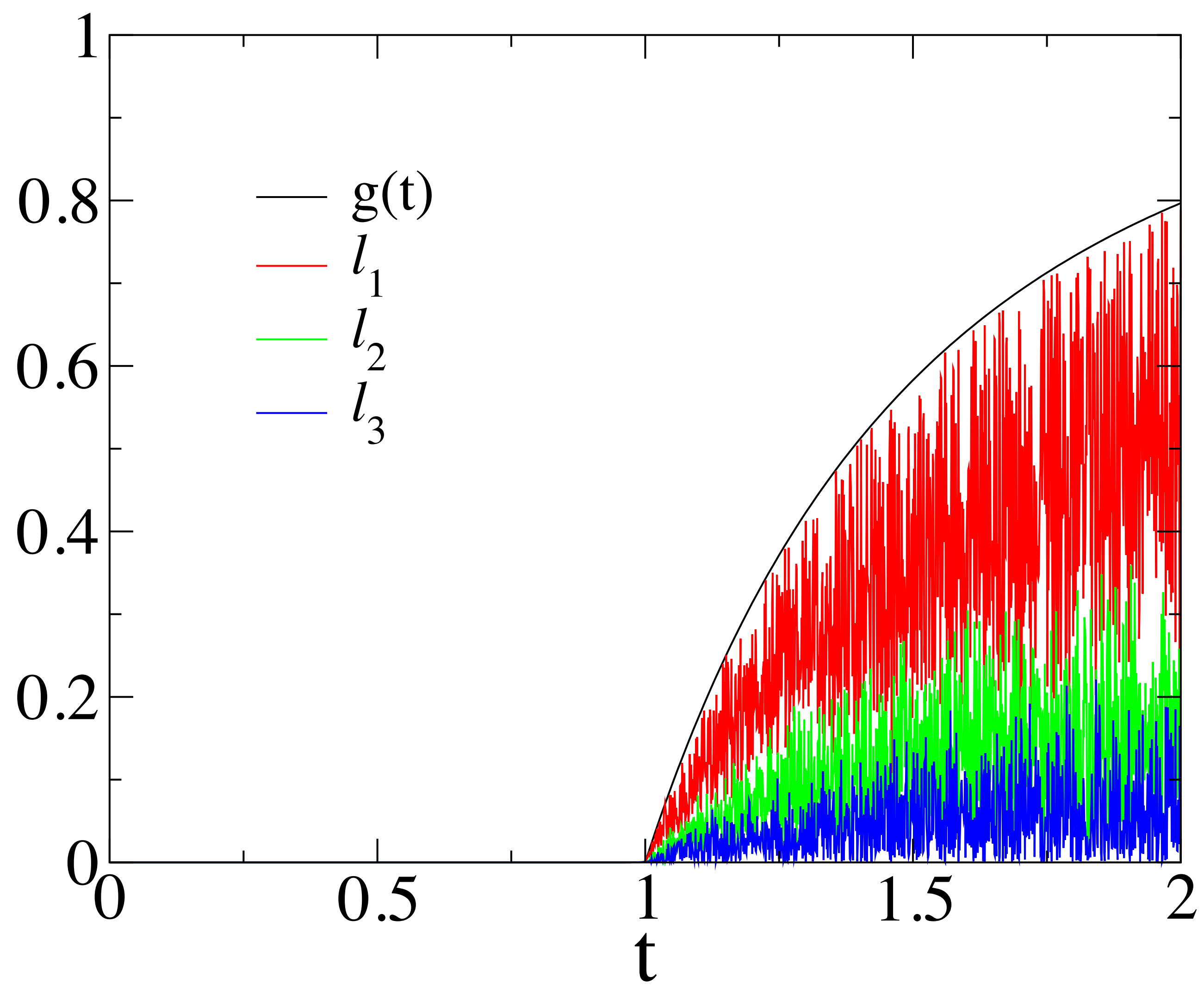
Total Number of Giant Rings

Simulation results



Law of large numbers

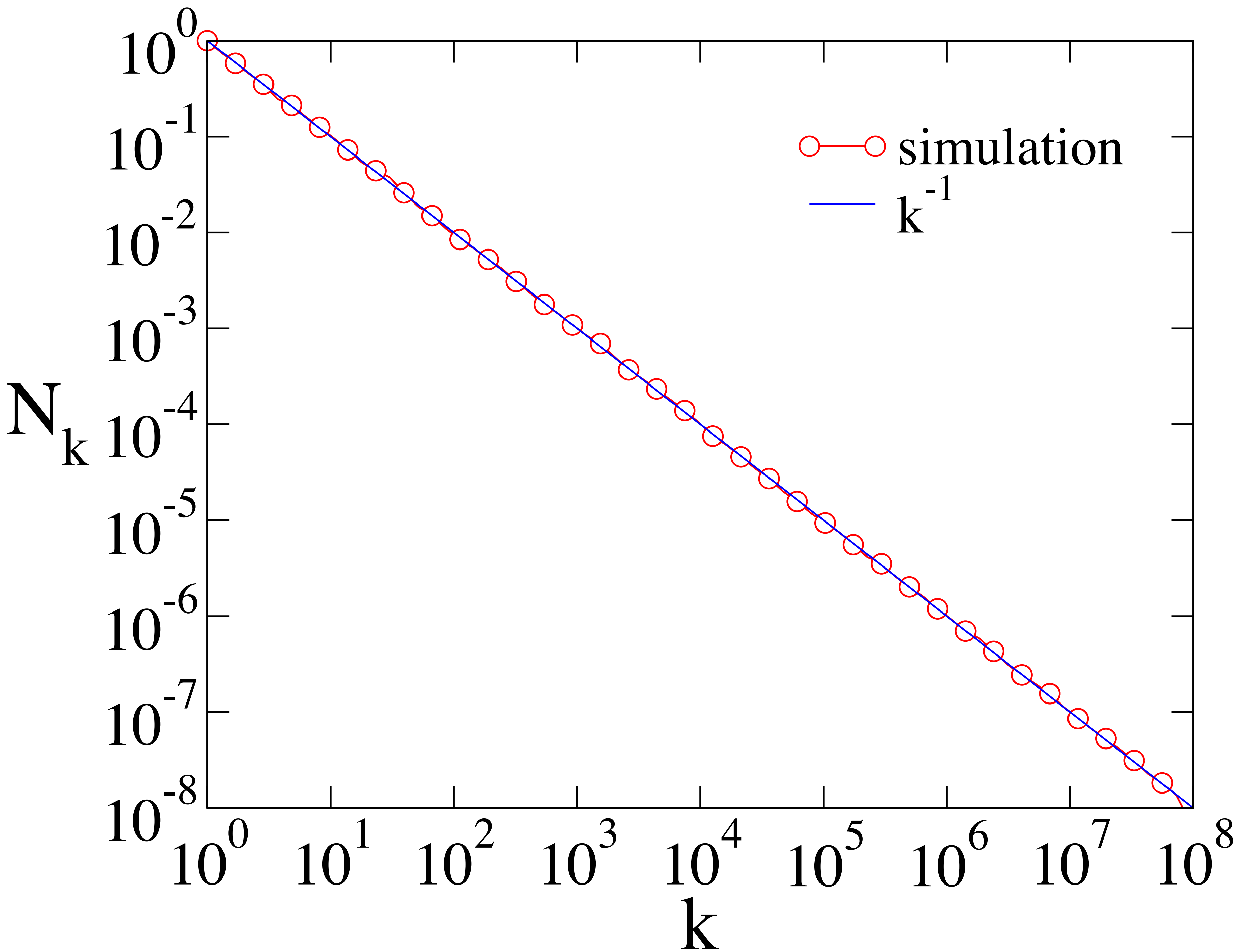
Multiple Coexisting Giant Rings



Total mass of giant rings is a deterministic quantity
Mass of an individual giant ring is a stochastic quantity!
Giant rings break and recombine very rapidly

Final Distribution

Simulation results



Implications to Shuffling

- N pairwise shuffles generate a giant cycle
- Size of emergent giant cycle is $N^{2/3}$
- $N \ln N$ pairwise shuffles generate random order

Golomb 61
Flatto 85
Diaconis 86

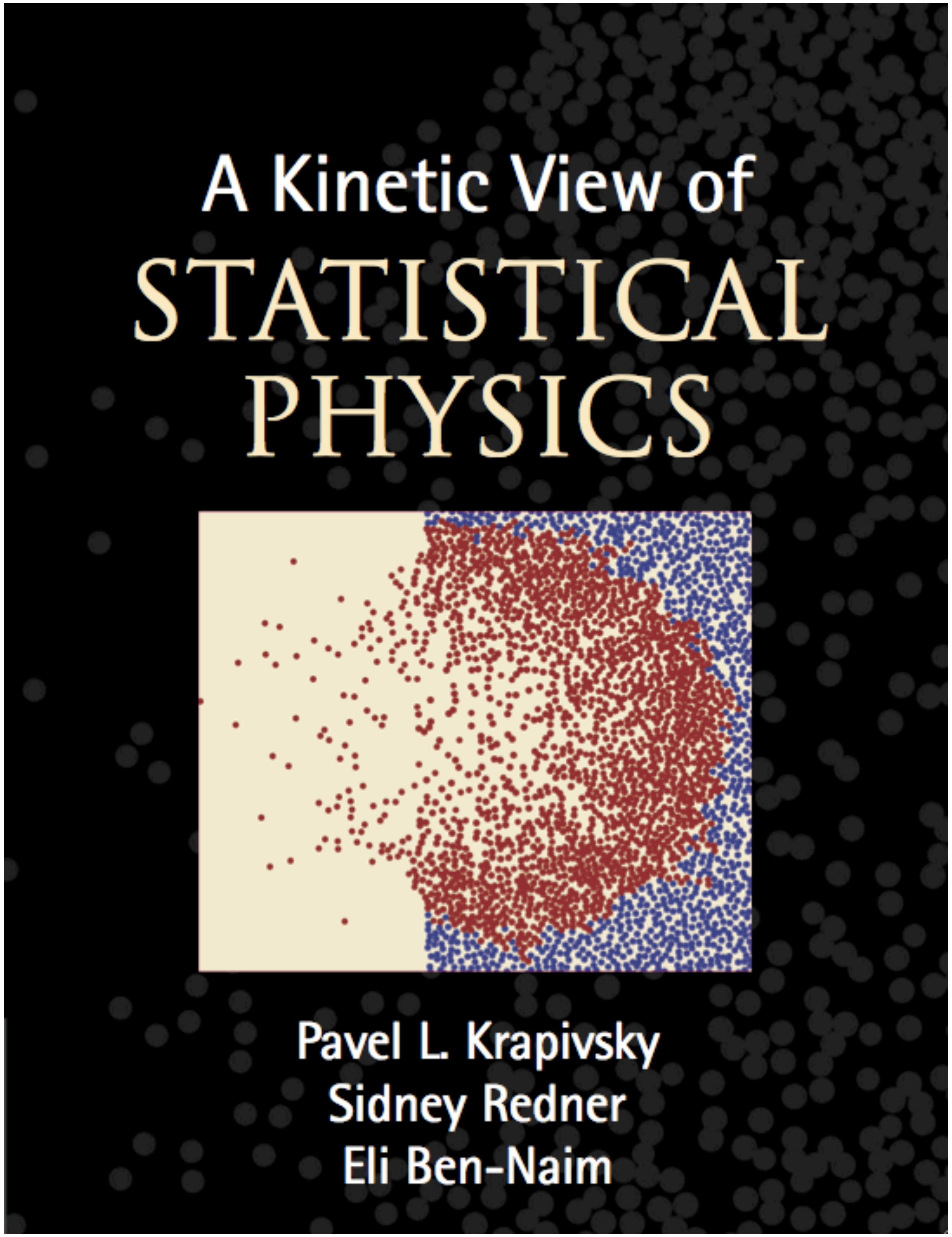
Summary

- Kinetic formulation of a regular random graph
- Equivalent to: (i) aggregation-fragmentation (ii) shuffling
- Finite rings phase: fragmentation is irrelevant
- Giant rings phase
  - Multiple giant rings coexist
  - Number of giant rings fluctuates
  - Total mass is a deterministic quantity
  - Very rapid evolution

chapter 5
aggregation

chapter 12
population dynamics

chapter 13
complex networks



Cambridge University Press 2010