Some aspects of information-driven networks

David Sherrington

University of Oxford

With help from

Ton Coolen, Tobias Galla, Heiko Bauke, Cris Moore

Some aspects of information-driven networks

dynamical

David Sherrington

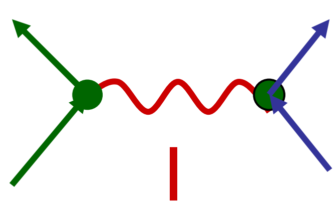
University of Oxford

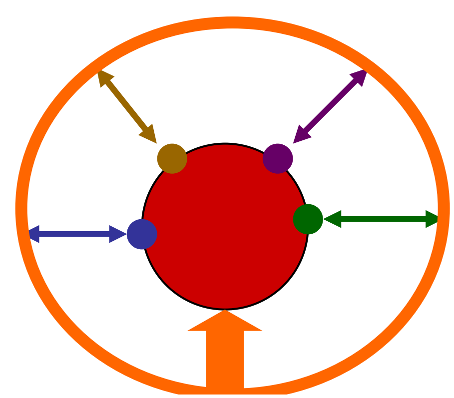
With help from

Ton Coolen, Tobias Galla, Heiko Bauke, Cris Moore

Types of problem

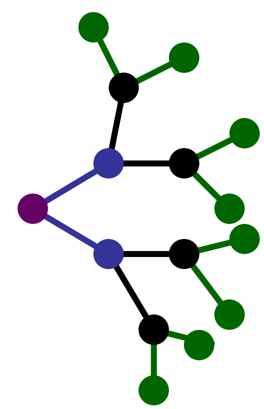
1. 'Information' as a connector
   - Many 'agents' with individual propensities
     – Abilities, inclinations, aversions, strategies
   - Not necessarily any direct interaction
   - Respond to 'common information'
     - Available equally to all
     - Some generated by the collection of agents (endogenous)
     - Some generated by external sources (exogenous)
   - Leads to effective interaction
     - *c.f.* bosons in QFT or maybe





2. Networks retrieving information by queries

- Minimise 'time'/ # steps to find someone with the answer
  - Scale-free networks
    - Search N nodes in $\ln N$ steps



- Dynamical networks
  - Growing
    - Much studied
  - Networks under churn
    - Nodes constantly entering and leaving
    - Topological transitions

1. Information as connector

- Many-body
- Quenched disorder
  - Different ‘agents’ ~ different abilities,strategies etc.
- Often frustration/competition
- Dynamical

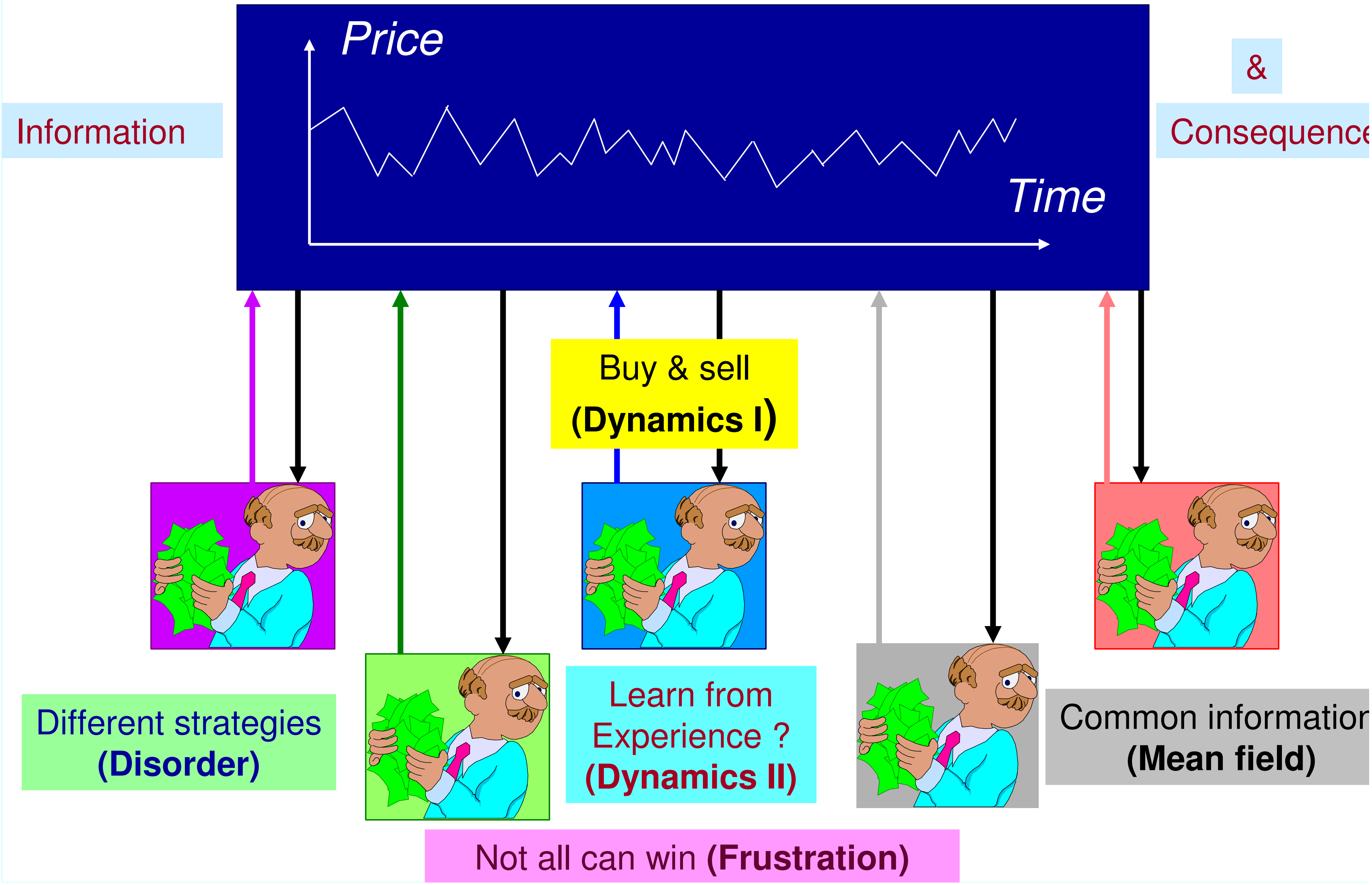
- Cooperative behaviour?
- Transitions?
- Complex?

- Models
- Methodology

- Range-free information
- Some solutions
- Some concepts

Stockmarket

Many speculators; buy low, sell high



Minority game

N agents 2 choices

Aim to be in minority

Individual strategies ➡ Collective consequence

act on common in turn

preferences modified by experience

⇩

Correlated behaviour through common purpose

Phase transitions, fruitful irrationality, 'non-ergodicity'

Original MG model

(Challet & Zhang '97)

- Information: Minority choice last m steps
- Strategies: Boolean functions
  - (few each, random quenched, different for each agent)
- Points: decide which strategy to use (t)

  updated by performance (t)

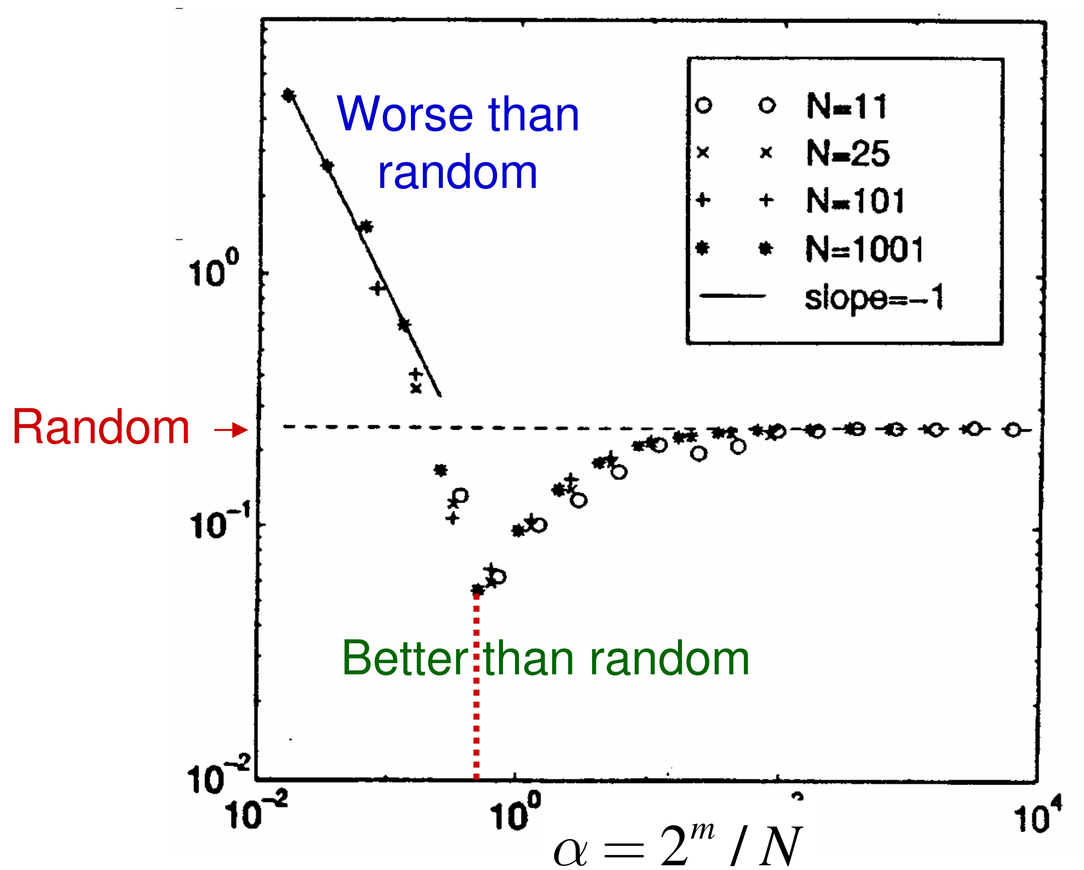
  best strategy used (t)

Simplify to 2 strategies per agent

‘Volatility’

a ‘natural’ relevant macroscopic observable

Standard deviation of #‘buy’ versus #‘sell’



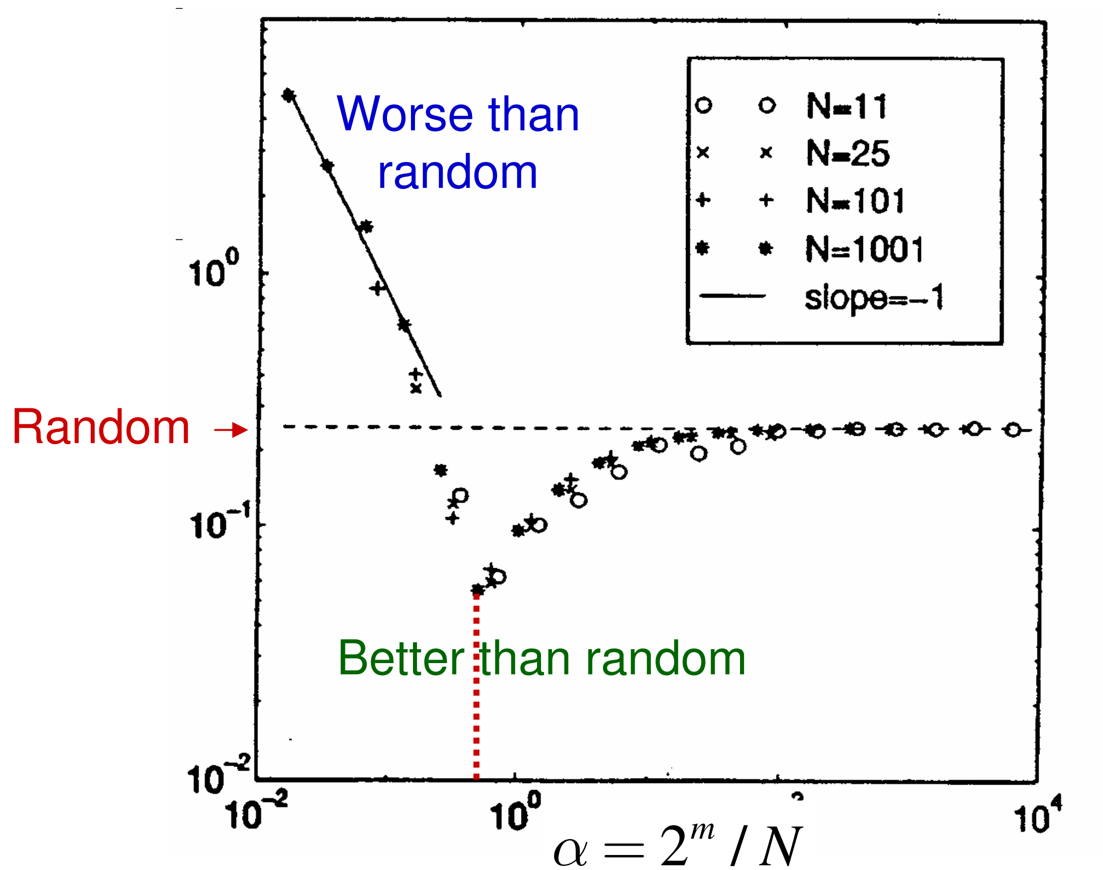
Savit, Manuca, Riola 99

- *Correlations*
- *Scaling parameters:* $\alpha=2^m/N$, $\sigma/\sqrt{N}$
- *Phase transition:* α_c
  minimum in volatility

‘Volatility’

a ‘natural’ relevant macroscopic observable

Standard deviation of #‘buy’ versus #‘sell’



Similar for ‘random histories’

MG output

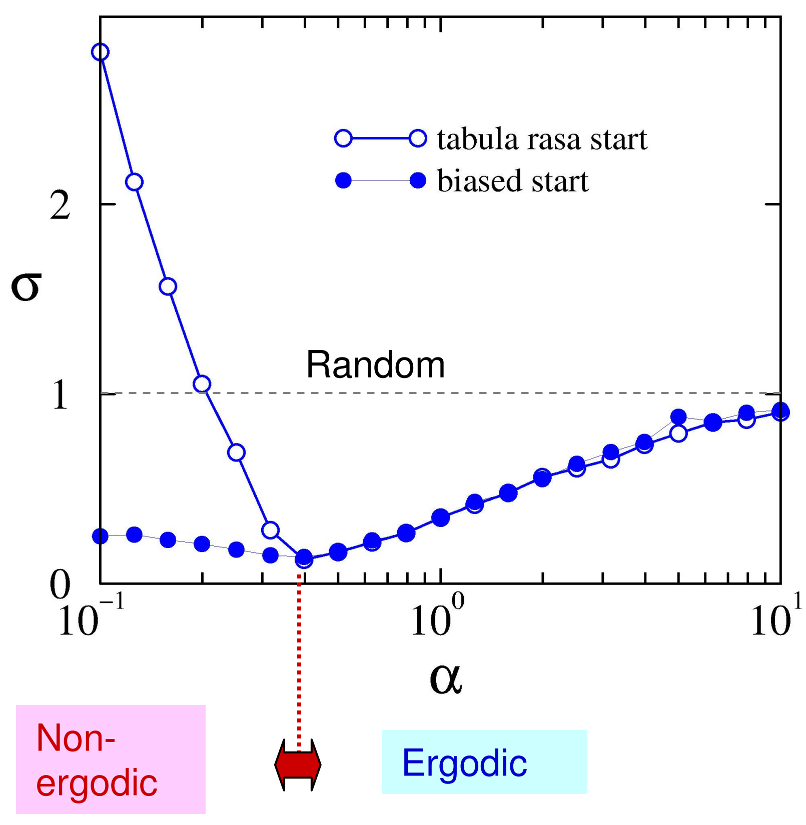
~

interaction effectuator

Cavagna

Ergodicity-breaking

Generalized batch model



Phase transition

$$\alpha = \alpha_c$$

Minimum in volatility
&
Ergodic/ non-ergodic

Recall: $\alpha = 2^m / N$

$= D / N$

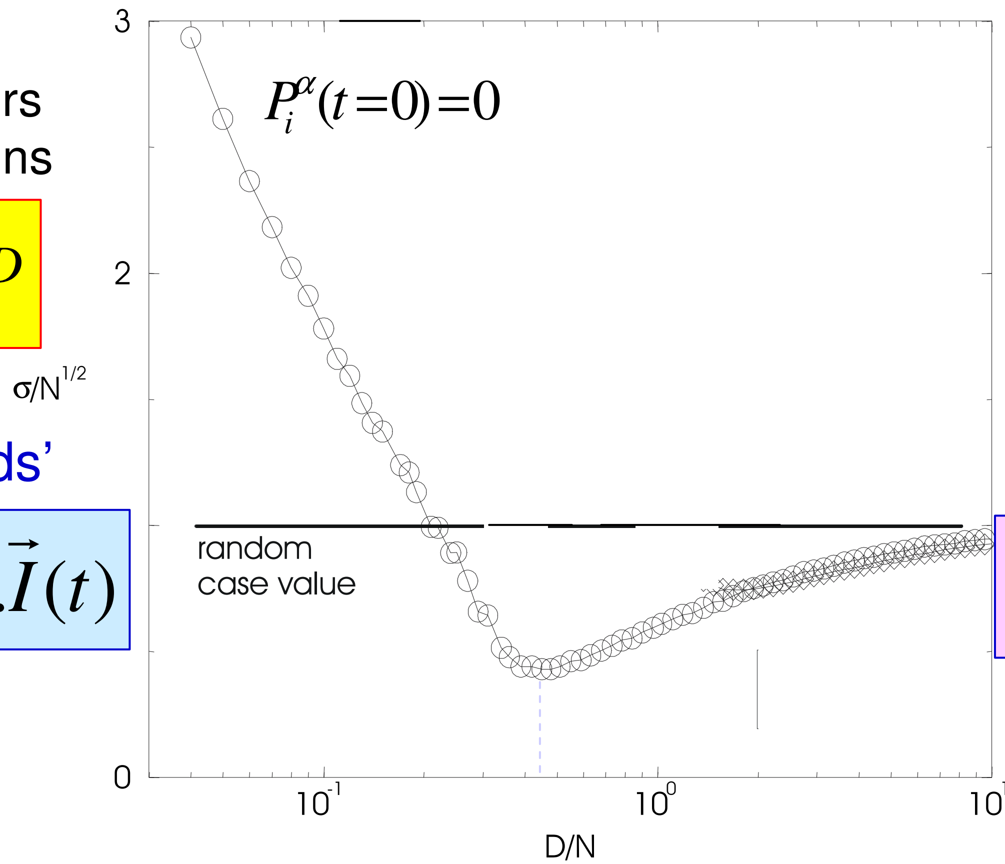
MG with 'random information'

Strategies:
quenched-random vectors in D dimensions

$$\vec{R}_i^{\alpha} \in \mathbf{R}^D$$

Individual 'bids'

$$b_i^{\alpha}(t) = \vec{R}_i^{\alpha} . \vec{I}(t)$$



Information:
stochastically random vectors in D dimensions

$$\vec{I}(t) \in \mathbf{R}^D$$

Average bid

$$a(t) = N^{-1} \sum_i b_i^{*}(t)$$

* denotes used

Points update:

$$P_i^{\alpha}(t+1) = P_i^{\alpha}(t) - b_i^{\alpha}(t) a(t)$$

Difference equation

- Relative point-score: $p_i(t) = P_i^1(t) - P_i^2(t)$
- Dynamics: $p_i(t+1) = p_i(t) - N^{-1}\sum_j [\overrightarrow{R^*{}_j}.\overrightarrow{I(t)}][\overrightarrow{I(t)}.\vec{\xi}_i].$
- Strategy vectors: $\overrightarrow{\varpi_i},\ \vec{\xi}_i = \overrightarrow{R_i^1} \pm \overrightarrow{R_i^2}$

Coarse-grained time-average over I(t)

Effective interaction between agents

$$H = \sum_{ij} J_{ij} s_i s_j + \sum_i h_i s_i$$

'Range-free'

$$J_{ij} = \sum_{\mu} \xi_i^{\mu} \xi_j^{\mu}, \quad h_i = \sum_{j,\mu} \varpi_j^{\mu} \xi_i^{\mu}$$

Batch

equivalent to updating points only after time O(N); averaging over 'common information'

'Equation of motion'

$$p_i(t+1) = p_i(t) - h_i - \sum_j J_{ij} \operatorname{sgn} p_j(t)$$

$$= p_i(t) - \partial H / \partial s_i \Big|_{\{s_i = \operatorname{sgn} p_i(t)\}}$$

c.f. Anti-Hopfield in field

Effective Hamiltonian

$$H = \sum_{ij} J_{ij} s_i s_j + \sum_i h_i s_i$$

$$J_{ij} = \sum_{\mu} \xi_i^{\mu} \xi_j^{\mu}, \quad h_i = \sum_{j,\mu} \varpi_j^{\mu} \xi_i^{\mu}$$

Repellors

c.f. Hopfield model

$$J_{ij} = -\sum_{\mu} \xi_i^{\mu} \xi_j^{\mu}$$

Attractors

Recall

$$\overrightarrow{\varpi}_i, \ \overrightarrow{\xi_i} = \overrightarrow{R_i^1} \pm \overrightarrow{R_i^2}$$

$$\therefore \{h_i = 0\} \ \equiv \ \{\overrightarrow{R_i^1} = -\overrightarrow{R_i^2}\}$$

Anti-correlated strategies

Full macrodynamics

equilibrium or non-equilibrium

Starting point: generating functional

$$Z = \int \prod_t d\vec{p}(t) W\left(\vec{p}(t+1) \middle| \vec{p}(t)\right) P_0(\vec{p}(0))$$

Updates: $p_i(t+1) = p_i(t) - h_i - \sum_j J_{ij} \,\mathrm{sgn}\, p_j(t)$ $\longrightarrow W$

Batch: $h_i = N^{-1} \sum_j \vec{\xi}_i . \vec{\omega}_j \;;\; J_{ij} = N^{-1} \vec{\xi}_i . \vec{\xi}_j$

(Coolen & Heimel)

Micro $\rightarrow$ Macro

- Introduce auxiliary ‘macrofields’ (x 1)

$$1 = \int DC(t,t')\Pi_{t,t'}\delta(C(t,t') - N^{-1}\sum_i p_i(t)p_i(t'))$$ etc.

- Exponentiate delta functions e.g.

$$\int d\widehat{C}(t,t')\exp\{-i\widehat{C}(t,t')[C(t,t') - N^{-1}\sum_i p_i(t)p_i(t')]\}$$

- Disorder average (over strategies)
- Substitute for many microvariables
- Gaussian in explicit microvariables: integrate out

Micro $\rightarrow$ Macro

Now macrovariables only

$$\bar{Z} = \int [DCD\hat{C}][DKD\hat{K}][DLD\hat{L}] \exp\{N[\Psi + \Phi + \Omega]\}$$

$$C(t,t') = N^{-1} \sum_i s_i(t) s_i(t')$$

$$K(t,t') = N^{-1} \sum_i s_i(t) \hat{p}_i(t'); \quad \hat{p} \sim \partial / \partial s$$

$$L(t,t') = N^{-1} \sum_i \hat{p}_i(t) \hat{p}_i(t'), \quad etc.$$

Large *N*: extremally dominated

Saddle-point $\rightarrow$ effective single particle dynamics

Effective single-agent ensemble

Non-Markovian stochastic process

$$p(t+1) = p(t) - \alpha \sum_{t' \le t} (\mathbf{1}+\mathbf{G})^{-1}{}_{tt'}\,\mathrm{sgn}\, p(t') + \theta(t) + \sqrt{\alpha}\,\eta(t)$$

$$\text{where } \langle \eta(t)\eta(t')\rangle = [(\mathbf{1}+\mathbf{G})^{-1}(1+\mathbf{C})(\mathbf{1}+\mathbf{G}^{\mathbf{T}})^{-1}]_{tt'}$$

with coloured noise, memory, self-consistent correlation & response functions

$$C_{tt'} = \langle \mathrm{sgn}\, p(t)\,\mathrm{sgn}\, p(t')\rangle_* \equiv N^{-1}\sum_i \langle \mathrm{sgn}\, p_i(t)\,\mathrm{sgn}\, p_i(t')\rangle$$

$$G_{tt'} = \frac{\partial}{\partial\theta(t')}\langle \mathrm{sgn}\, p(t)\rangle_* \equiv N^{-1}\sum_i \frac{\partial}{\partial\theta_i(t')}\langle \mathrm{sgn}\, p_i(t)\rangle$$

where $\langle f \rangle_*$ is an effective average involving $P_0(p(0))$, $\mathbf{G}$, $\mathbf{C}$.

Exact but non-trivial

Simulations & iterated theory

Initial bias

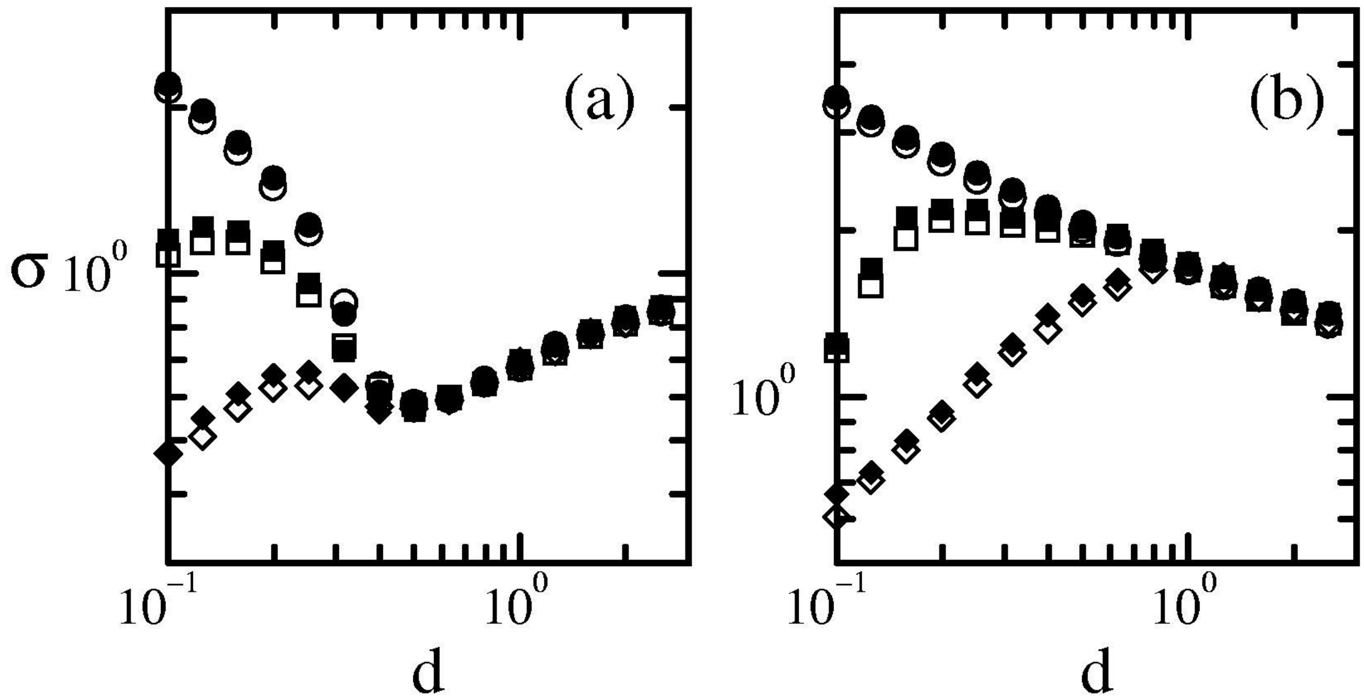
$p_i(0)=0$

$p_i(0)=0.5$

$p_i(0)=1$

ρ=.5: uncorrelated

ρ=0: anti-correlated



Open = simulations Solid = numerical iteration of analytic effective agent equations

Galla & S

Solutions

to

Effective single agent equations

Any α: Numerically soluble for finite number of time-steps,
but increasingly computer-expensive as t increases

$\alpha \geq \alpha_c$: Analytically soluble for certain quantities
with ansätze whose breakdown signals α_c

$\alpha < \alpha_c$: Not yet solved

Further ansätze for equilibrium analysis: $\alpha \geq \alpha_c$

- Stationarity: $C_{tt'} = C(t-t'),\ G_{tt'} = G(t-t')$
- Finite integrated response
- Weak long term memory: $\lim_{t\to\infty} G_{tt'} = 0 \text{ for all finite } t'$

→ Order parameters in stationary state

- Persistent correlation function: $Q = \lim_{\tau\to\infty} C(\tau)$
- Integrated response: $\chi = \sum_{\tau} G(\tau)$

- Breakdown of theory: one of these assumptions violated

Persistent correlations

Correlated strategies:ρ: $P(R_{i1}^{\mu} = R_{i2}^{\mu}) = \rho$

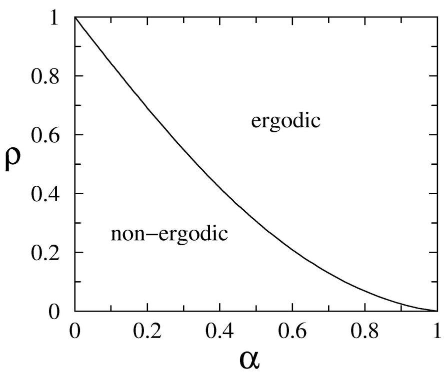
ρ=0: anti-correlated

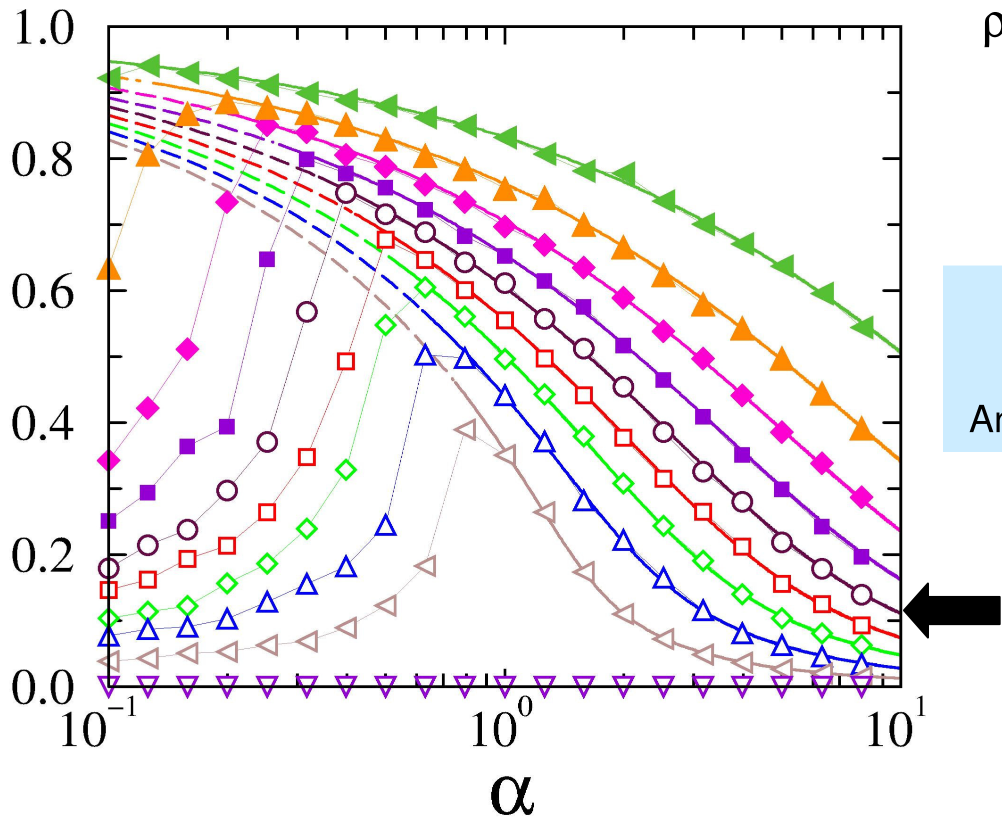
ρ=0.5: uncorrelated

ρ=1:identical

Simulations & analytic theory

Batch





ρ=0,0.1,..,0.9
bottom to top
Anti-corr. to highly corr.

Possible extensions

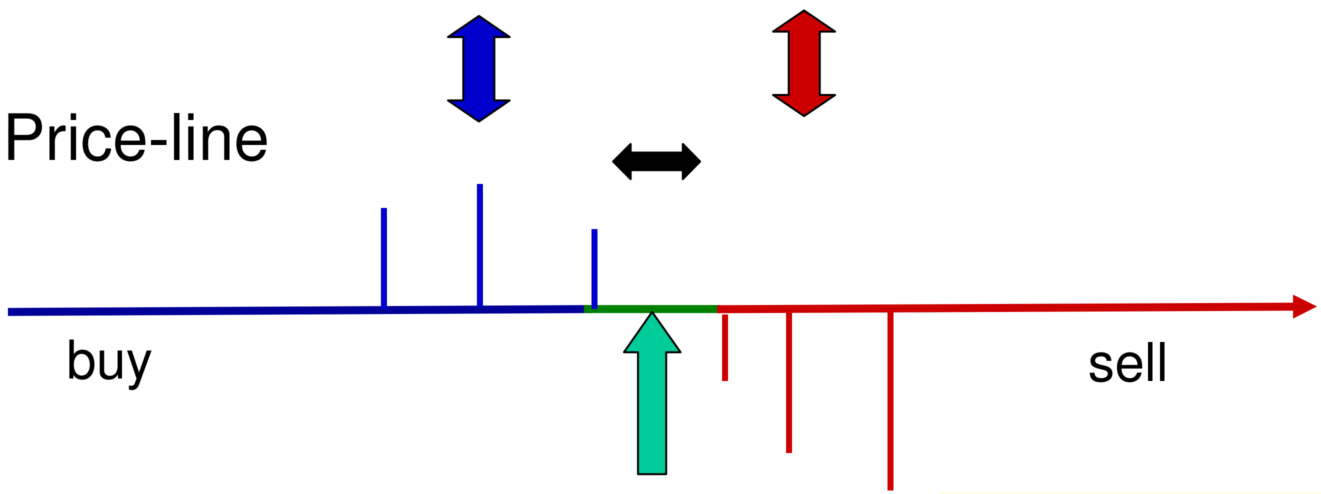
within econophysics

- Systems with more features but still range-free
  - i.e. more ‘local’ variables and couplings but still global interactions
- Dynamical strategies: still need heterogeneity
- Liquidity providers: *c.f.* ATP

More realistic extension of minority game?

Limit-order book

Agents place or remove orders: buy, sell, market. May be executed.
Speculators gain on price changes. Manufacturers must absorb → liquidity.



But how do they choose what to do?
Evolution of strategies?
Driven by individual attitudes, co-operative actions, learning?

c.f. Evaporation-deposition-annihilation

More generally

Dynamical generating functionals

$$Z = \int DS\ DJ\ \ \delta(\text{Equations of motion})\ \delta(\text{constraints})\ Jacobian$$

Microscopic variables, all times
Fast & slow
microscopic "attempt" times in eqns. of motion

Also generating term
$\exp\{i(\lambda S + \mu J)\}$

Include real endogenous information and exogenous influences, agent-differences & stochasticity/uncertainty

Micro $\rightarrow$ Macro-variables: multi-time

Peer-to-peer networks

- Computer connectivity networks
  - Operational connections: e.g. file-sharing
    - Distinct from physical connection network
  - Nodes constantly leaving and joining the network
    - Under churn

- Need fast file-finding
  - Scale-free structure: $p(k) \sim k^{-\gamma}$
    - Local search strategies scale sub-linearly with size

Can we devise

easily-implemented "networks under churn"

with power-law connectivity distributions ?

Preferential attachment

Barabasi-Albert

- Addition of a new node
  - Assign to each node an attractiveness: $A_i \propto k_i$
  - Connect new node to *m* existing nodes chosen randomly with probabilities proportional to their attractivenesses
  - Needs information about connectivity of all nodes
- Growing network
  - Yields power-law distribution: $p(k) \propto k^{-\gamma}$
- Network under churn
  - $P(k)$ decays faster than power-law

Local attachment

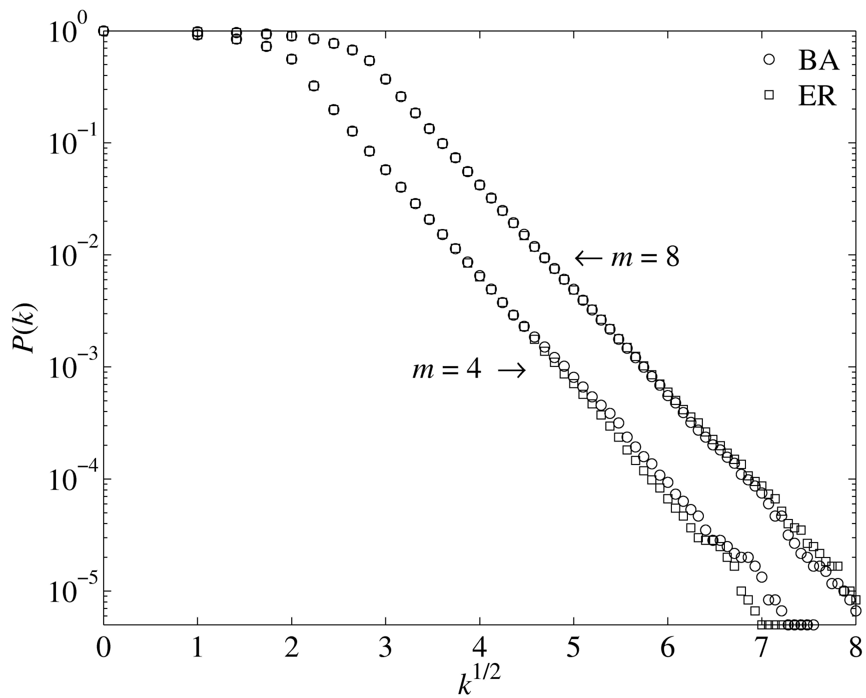
Bauke-S

- Addition of a new node, two-step procedure
  - Pick any other node randomly: no preference
  - Do not connect to that node!
  - Connect to a nearest neighbour of that node
  - Repeat *m* times
- Yields power-law connectivity for both growing networks and networks under churn
- Needs only local information

c.f. Gnutella cache-ponging

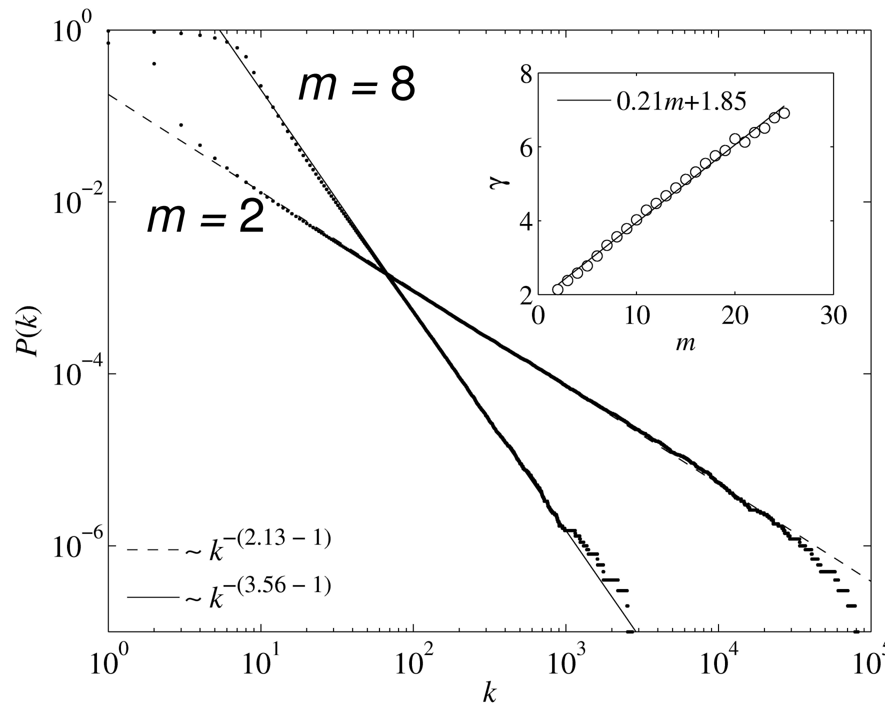
Cumulative degree distributions

Preferential attachment



Exponential

Local attachment



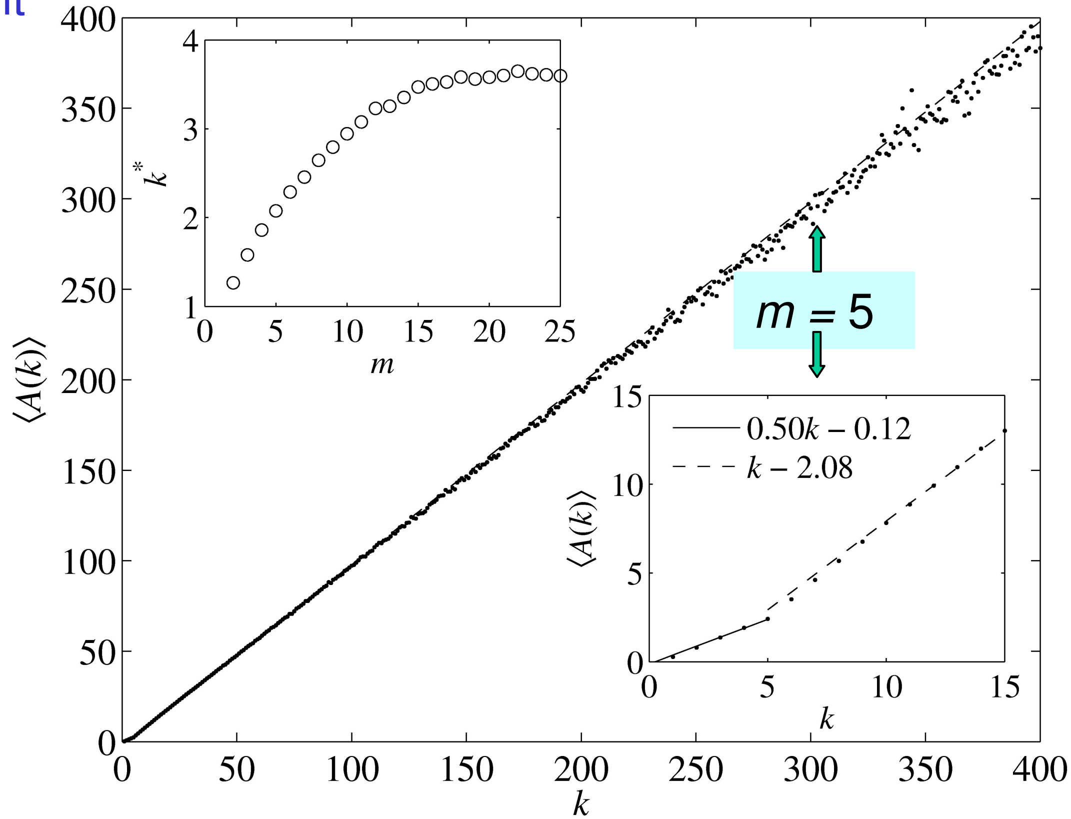
Power-law

Both are for networks of mean connectivity *m* under churn

Hereafter consider just local attachment

Mean attractiveness

Local attachment networks



$$\langle A(k) \rangle = (1 - \frac{k^*}{m})k;\ 0 \le k \le m$$

$$k - k^*;\quad k \ge m$$

Top left inset shows *k*(m)*

Conclusion so far

- 2-stage local attachment
  - Gives power-law scale-free networks
    - With their search-speed advantages
  - Without needing data on all peers
    - Recall
      - (i) random unbiased connection to peer A
      - (ii) ask who are his neighbours
      - (iii) connect randomly without bias to one of them
- Offers possibilities as a practical protocol

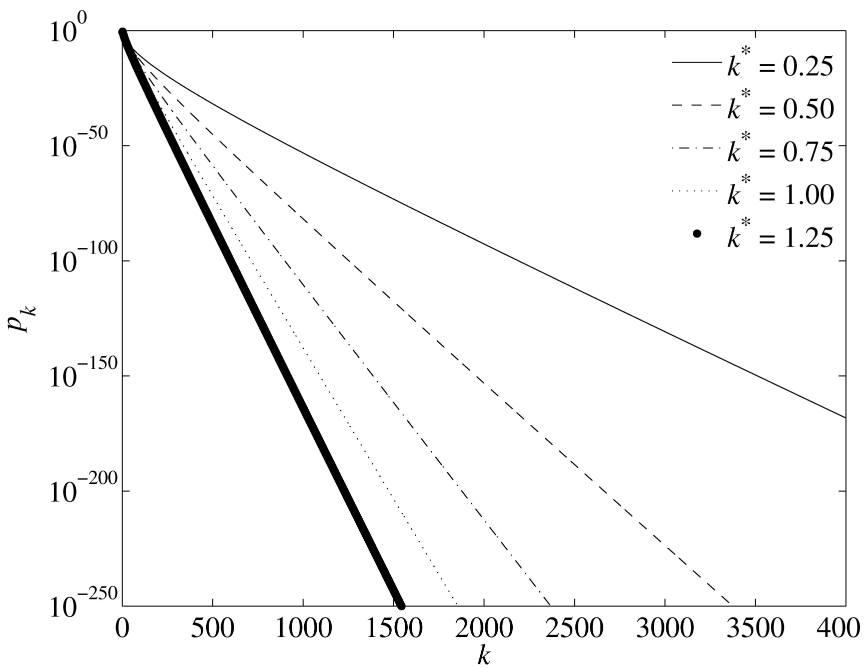
Topological transitions

- Networks under churn
  - E.g. At each time step:
    - Prob r: remove randomly chosen node
    - Prob 1: add new node and from it *m* new links
    - Choose the nodes to connect to randomly with attractiveness

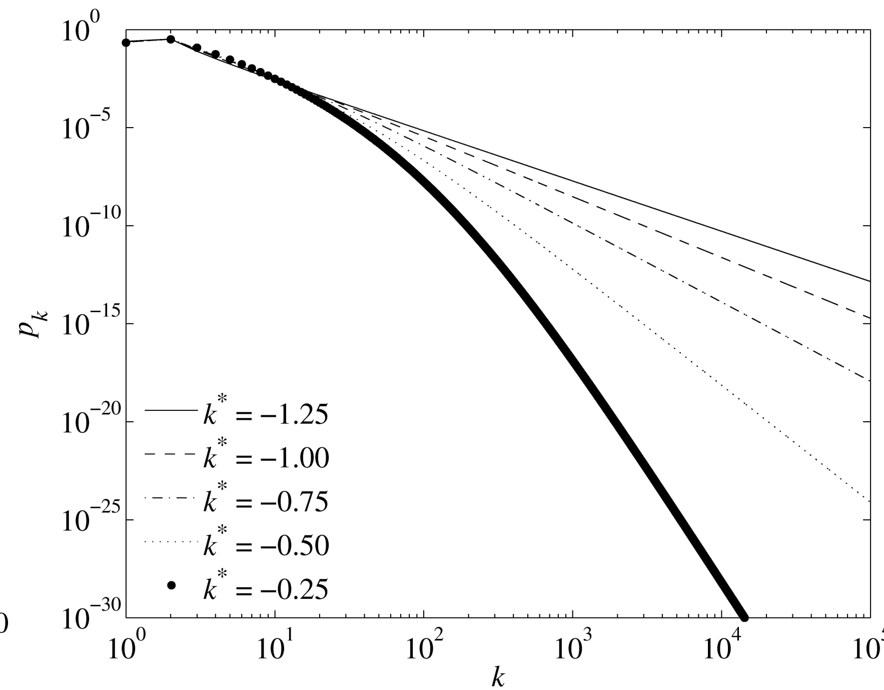
$$A_k = k + \varepsilon(k, k^*); \quad \varepsilon(k, k^*) = \begin{cases} kk^*/m & \text{if } k \le m \\ k^* & \text{else} \end{cases}$$

Power-law or exponential

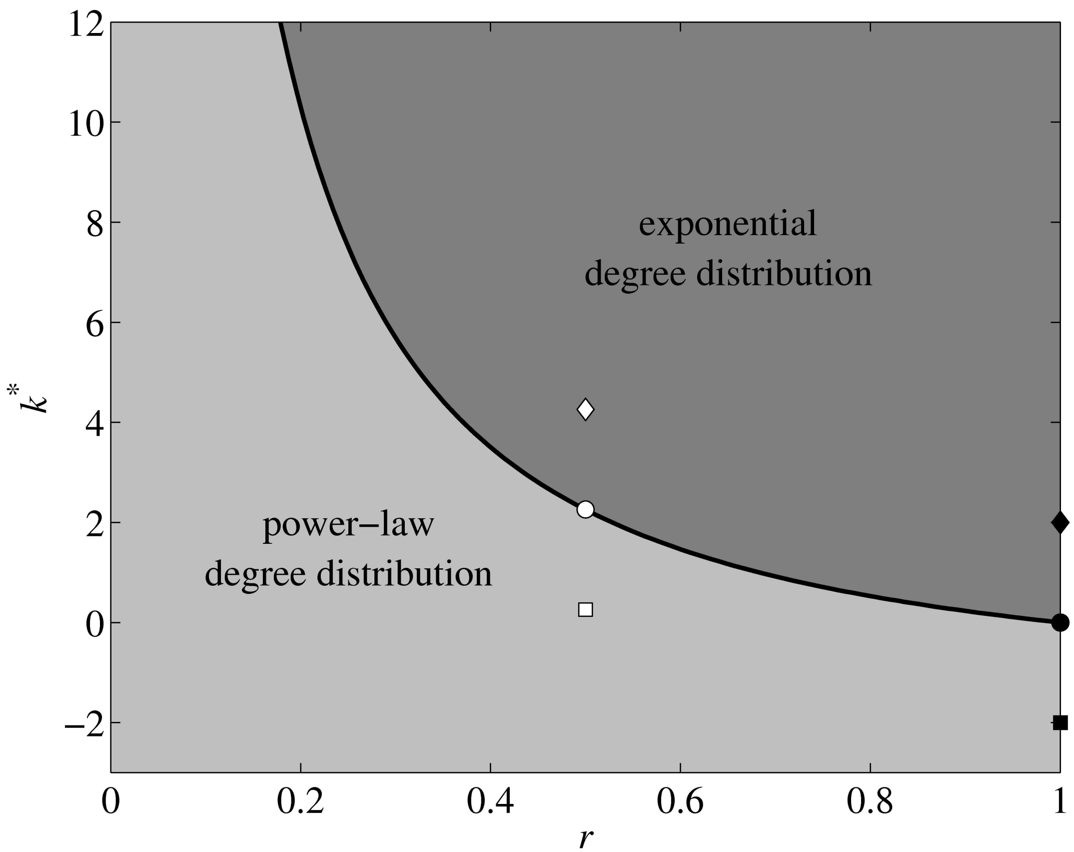
(a) exponential



(b) power-law



Phase diagram



Similar phase diagrams for other churn models

Analysis

$$|\delta_{k,m} + \frac{m}{\langle A \rangle}\{A_{k-1} p_{k-1} - A_k p_k\} + r(k+1) p_{k+1} - (rk+1) p_k = 0;$$

$$\langle A \rangle = \sum_k A_k p_k$$

Define

$$k^* = \langle A \rangle - \langle k \rangle$$

Transition

$$k_c^* = \frac{m(1-r)}{r(1-r)}$$

Can calculate behaviour of p_k

- $k^*<k_c^*$: power-law
- $k^*>k_c^*$: exponential with power-law corrections

$$p_k = Ck^\alpha \beta^k$$

$$k < k_c:\ \beta = 1;\ \alpha = 1 - m - \frac{\langle A \rangle (1-r) + m}{m - \langle A \rangle r} < 0$$

$$\rightarrow \infty \text{ as approach transition}$$

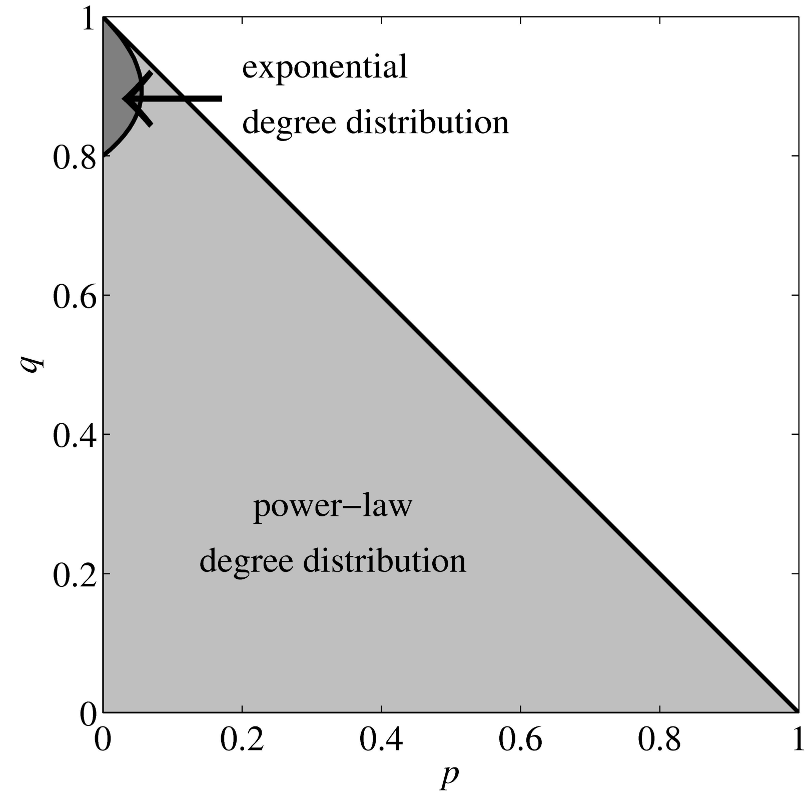
$$k > k_c:\ \alpha = -\frac{m(3-r) + k^*(1-r^2)}{m(1-r) - k^*} r(1+r)$$

$$\rightarrow -\frac{(3-r)}{(1-r)} \text{ as } k^* \rightarrow \infty$$

Another example

At each step:

- Insert *m* new edges with probability *p*
- Rewire *m* links randomly with probability *q*
- Add new node (*m* links) with probability *r*
- Preferential attachment



$p+q+r=1$

Conclusion

- Topological transitions as attachment rules varied

- Negative perturbations of linear attractiveness tend to stabilize power laws

- In view of ubiquity of power-laws in nature, do such pertubations occur in real world networks?