1-28-2011

On the Effect of Criticality and Topology on Learning in Random Boolean Networks

Alireza Goudarzi
Portland State University, alireza.goudarzi@gmail.com

On the Effect of Criticality and Topology on Learning in Random Boolean Networks

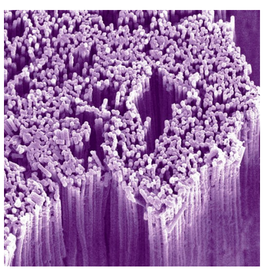
Alireza Goudarzi[1], Christof Teuscher, Natali Gulbahce, Thimo Rohlf

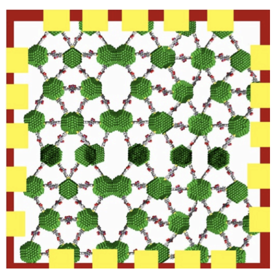
[1] System Science and Computer Science Department

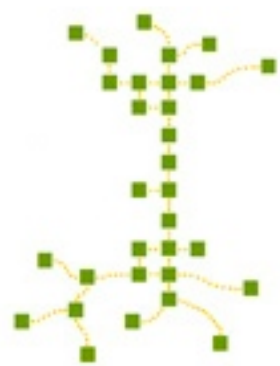
Portland State University (PSU), Portland, OR, USA

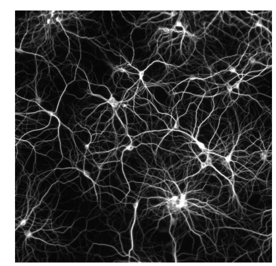
[1]alirezag@cecs.pdx.edu

Systems Science seminar, Jan 28, 2011









teuscher.:Lab
teuscher-lab.com
Emerging Computing Models and Technologies

Portland State
UNIVERSITY

Outline

- Introduction
  - Network information processing
  - Previous work
  - Random Automata Network
- Setup
  - GA implementation
  - Task solving
- Results
- Conclusion
  - Learning drives the network to a critical connectivity K_c=2 for large systems.
  - K_c scales for smaller systems as a power-law.

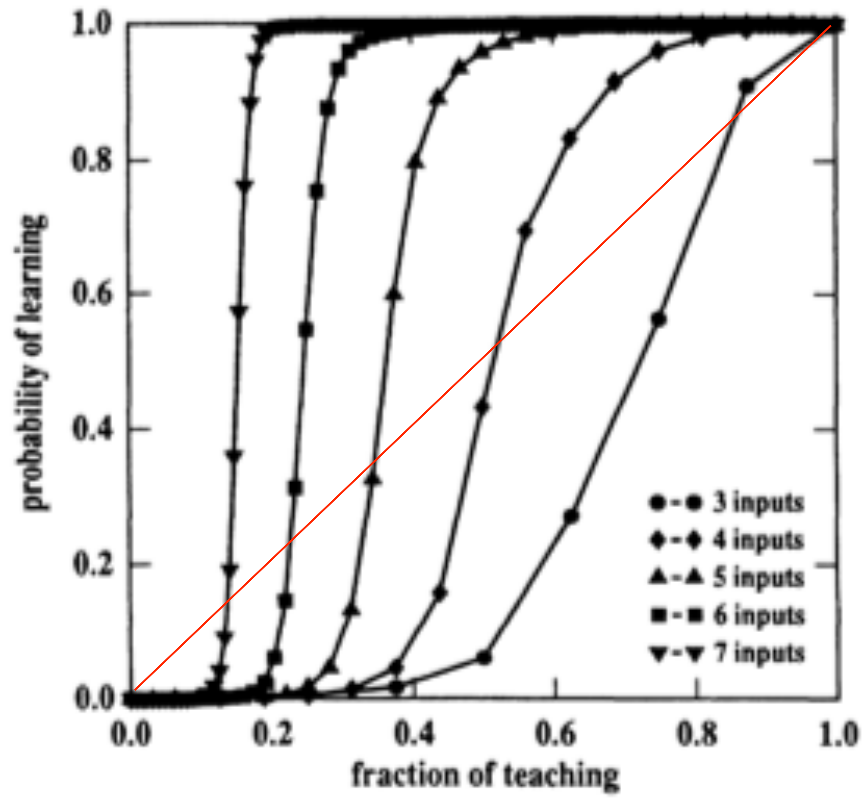
Introduction

- Imagine a random parallel machine, e.g., a number of compute elements with a given complexity that are interconnected with a random network.
- Why random parallel machines?
  - Self-assembled nanoelectronics have irregular structures.
  - We can build them very cheaply.
- How do we program our random parallel machine?
  - In this paper: How much connectivity do we need for effective learning?

K=?

Previous work

- Alan Turing (1948) proposed unorganized machines with simple NAND gates and learning through artificial evolution.
- Many attempts to use random networks for computation: Martland (1987), Aleksander (1973), Amari (1972).
- Patarnello and Carnevali (1989) and den Broeck and Kawai (1990) conducted general study of *learning capability* of feedforward random Boolean nets.
- Beiu and Makaruk (1998) proved feedforward Boolean nets with K=2 are size optimal for VLSI implementation.
- Darabos et al. (2007) showed small-world automata nets have higher computational performance than lattice, ring, and scale-free nets.

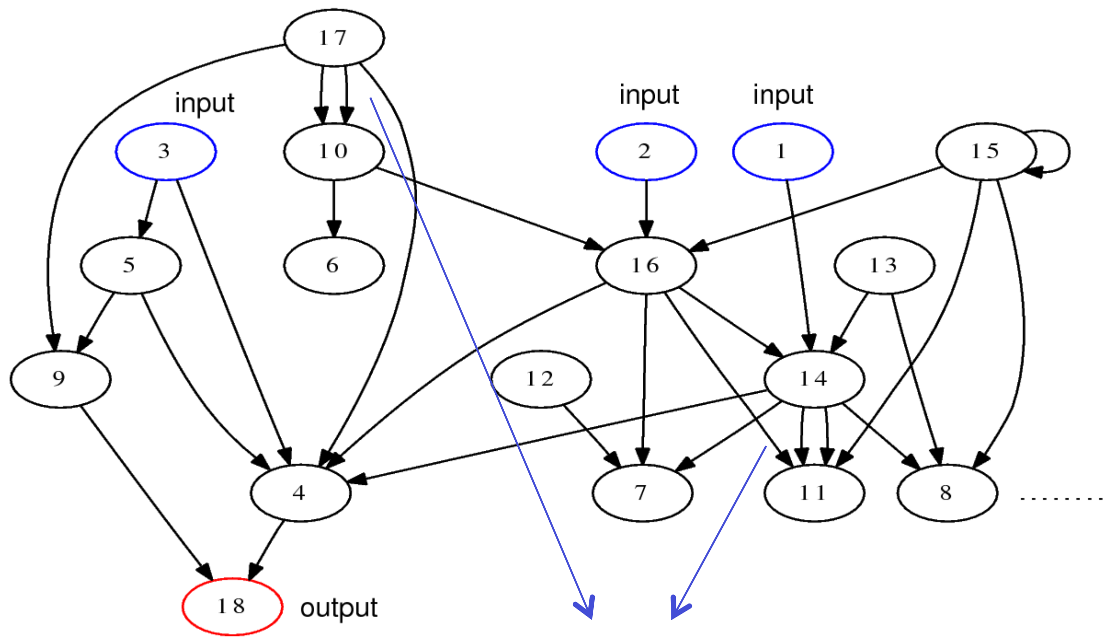


All of these evolutionary studies are under fixed topology and fixed connectivity K.

Random Automata Networks

- State transition $$\mathbf{F} : \{0,1\}^N \mapsto \{0,1\}^N$$
- Node states $$\sigma_i^{t+1} = f_i(x_{i_1}^t, x_{i_2}^t, \ldots, x_{i_{K_i}}^t) \quad \sigma_i^t \in \{0,1\}$$

Where K_i is the number of inputs to node *i.*



node LUT

in			out	
0	0	0	1	LSB
0	0	1	1	
0	1	0	0	
0	1	1	1	
1	0	0	0	
1	0	1	1	
1	1	0	0	
1	1	1	0	MSB

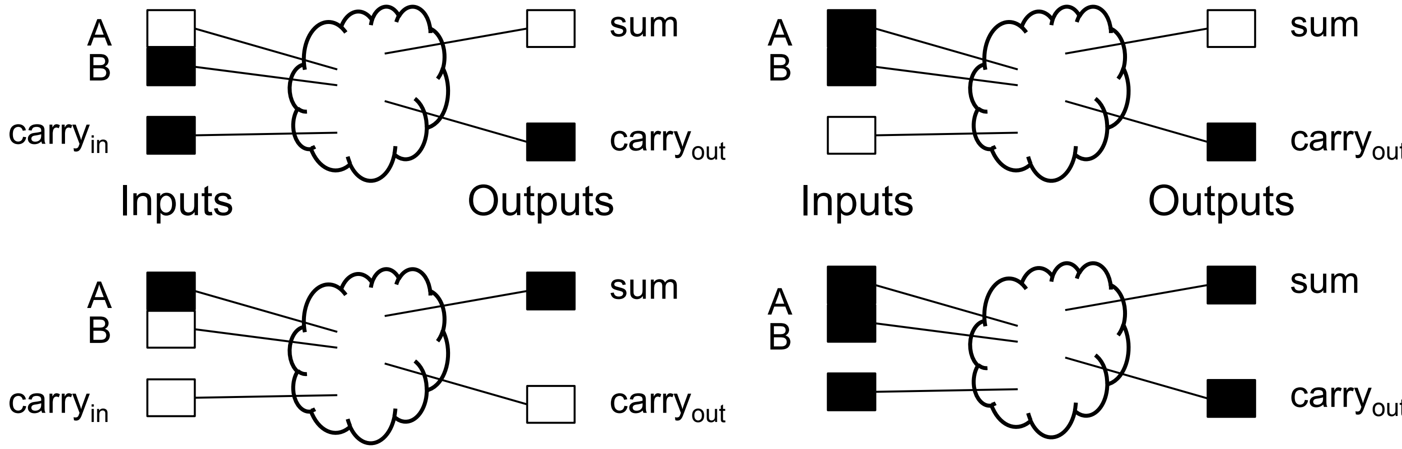
rule 43

Double connections are not allowed.

Computational Tasks (difficult)

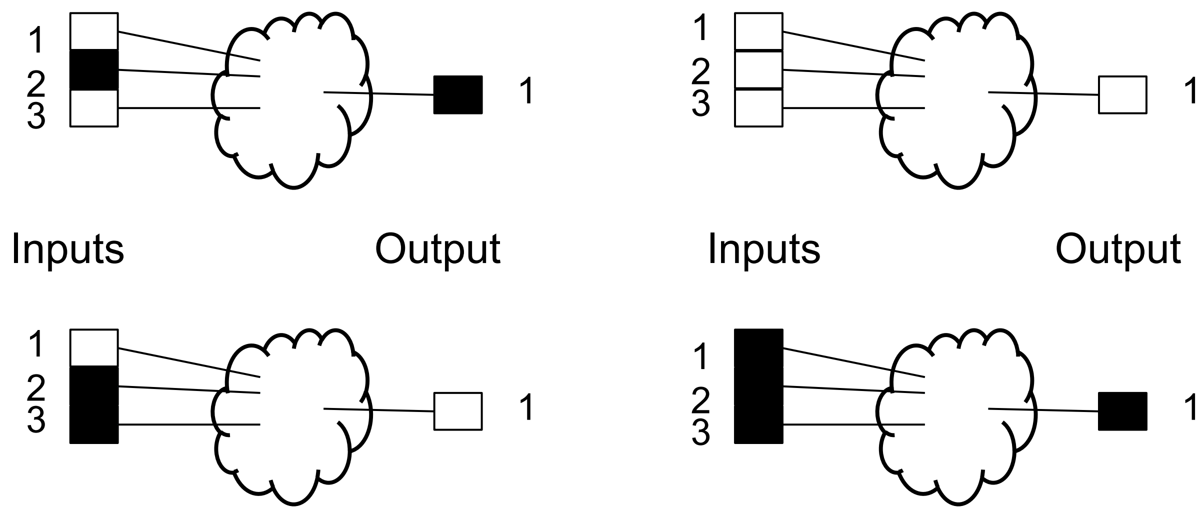
Full Adder

l=3



Even-odd

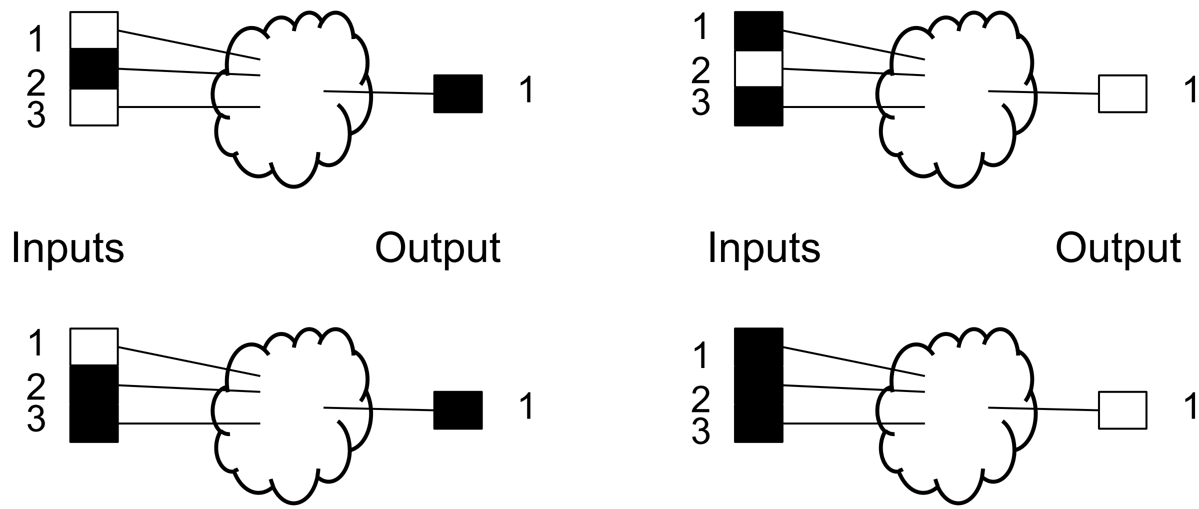
l=3



Computational Tasks (easy)

Rule 85

l=3

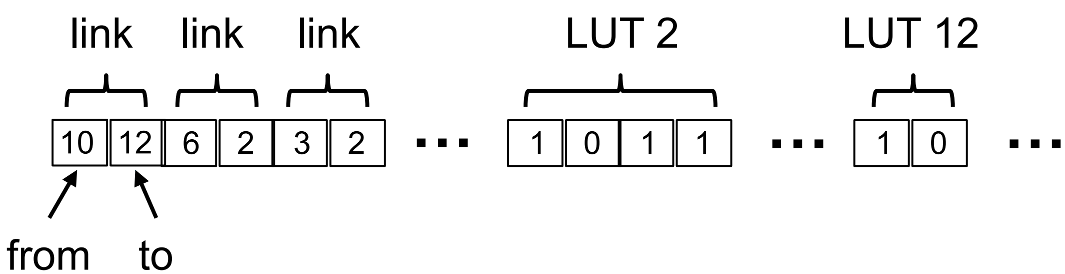


- The input 1 alone can identify the output completely (canalizing). The phase volume is more than 200 times larger than that of even-odd.

Genetic Algorithm

- Genetic representation

  Our networks are mostly sparsely connected. We use an adjacency list to encode the links in the network.



- Genetic operator: adding or deleting P links with probability 0.5^P randomly, or a bit flip on the look-up table. Note that K will be changing. This turns out to be very crucial.

- Fitness: $$f = 1 - \frac{1}{|M|}\sum_{M}(expected - actual)^2$$

  Where *M* is the set of training samples

Performance Measures Example

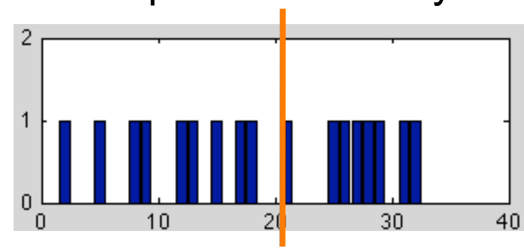
Input space	i1	i2	i3	O
pattern 1	0	0	0	0
pattern 2	0	0	1	1
pattern 3	0	1	0	1
pattern 4	0	1	1	0
pattern 5	1	0	0	1
pattern 6	1	0	1	0
pattern 7	1	1	0	0
pattern 8	1	1	1	1

training sample	i1	i2	i3	O
pattern 1	0	0	1	1
Pattern 2	0	1	0	1
Pattern 3	1	0	1	0
Pattern 4	1	1	0	0

Training sample size (T) = 4

Randomized for each generation

Output node activity



settling

measuring
8 “1”
12 “0”
output = “0”

training (f)

	expected	actual	expected	actual	expected	actual	expected	actual
pattern 1	1	1	1	1	1	1	1	1
pattern 2	1	0	0	0	1	1	0	0
pattern 3	1	1	0	1	0	0	0	0
pattern 4	0	0	1	1	0	0	0	0

generalization (g)

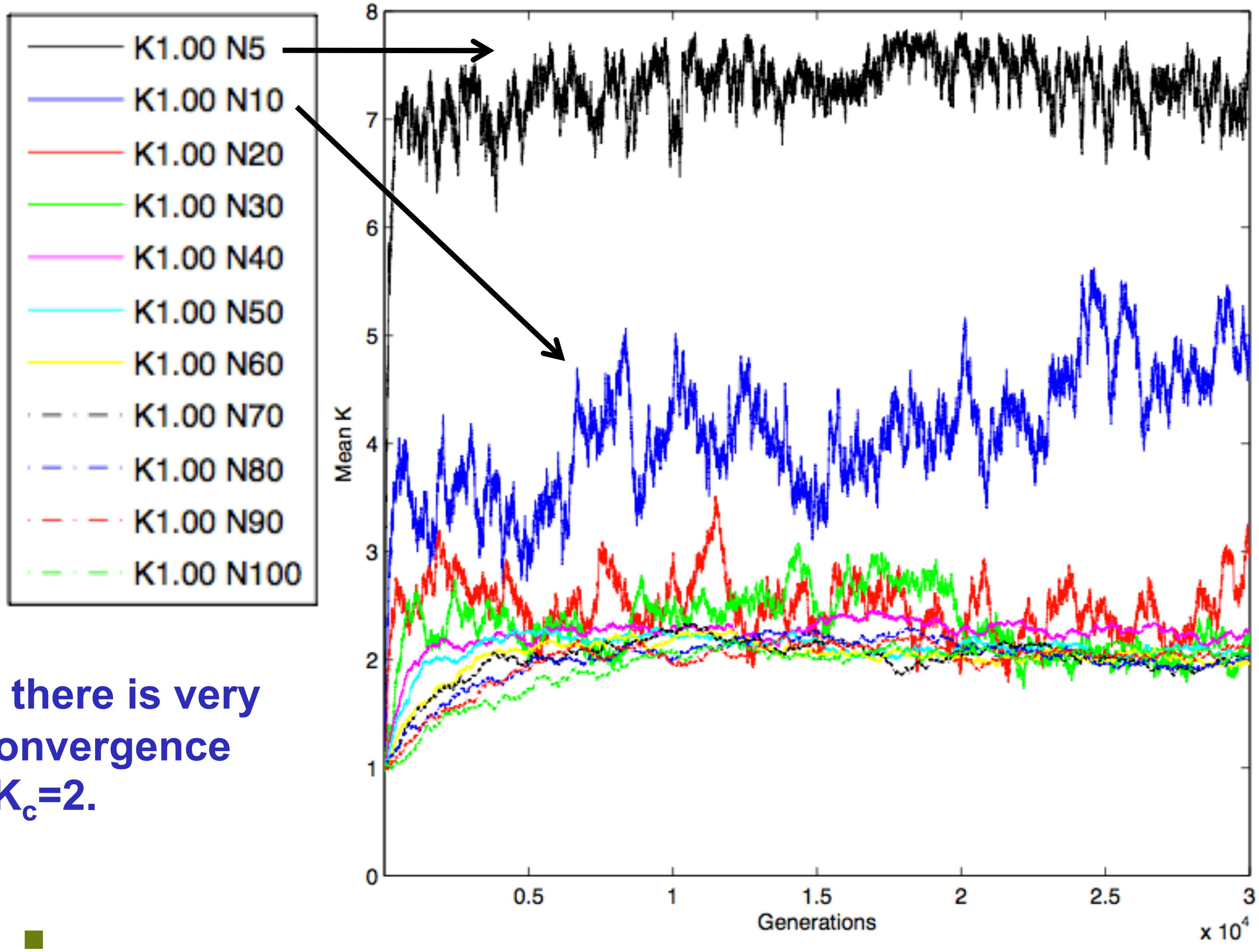
	expected	actual	expected	actual	expected	actual	expected	actual
pattern 1	0	0	0	1	0	0	0	0
pattern 2	1	1	1	0	1	1	1	1
pattern 3	1	1	1	1	1	1	1	1
pattern 4	0	0	0	0	0	0	0	0
pattern 5	1	0	1	0	1	1	1	1
pattern 6	0	0	0	0	0	0	0	0
pattern 7	0	0	0	0	0	0	0	0
pattern 8	1	1	1	0	1	1	1	0

runs	1	2	3	4
	f= 0.75	f= 0.75	f= 1	f= 1
	g= 0.875	g= 0.5	g= 1	g= 0.875

$$\beta'(s) = \frac{1}{r} \sum_r g_{final}$$

$$\beta'(s) = 0.8125$$

Finding the Critical K



For N>30 there is very smooth convergence towards K_c=2.

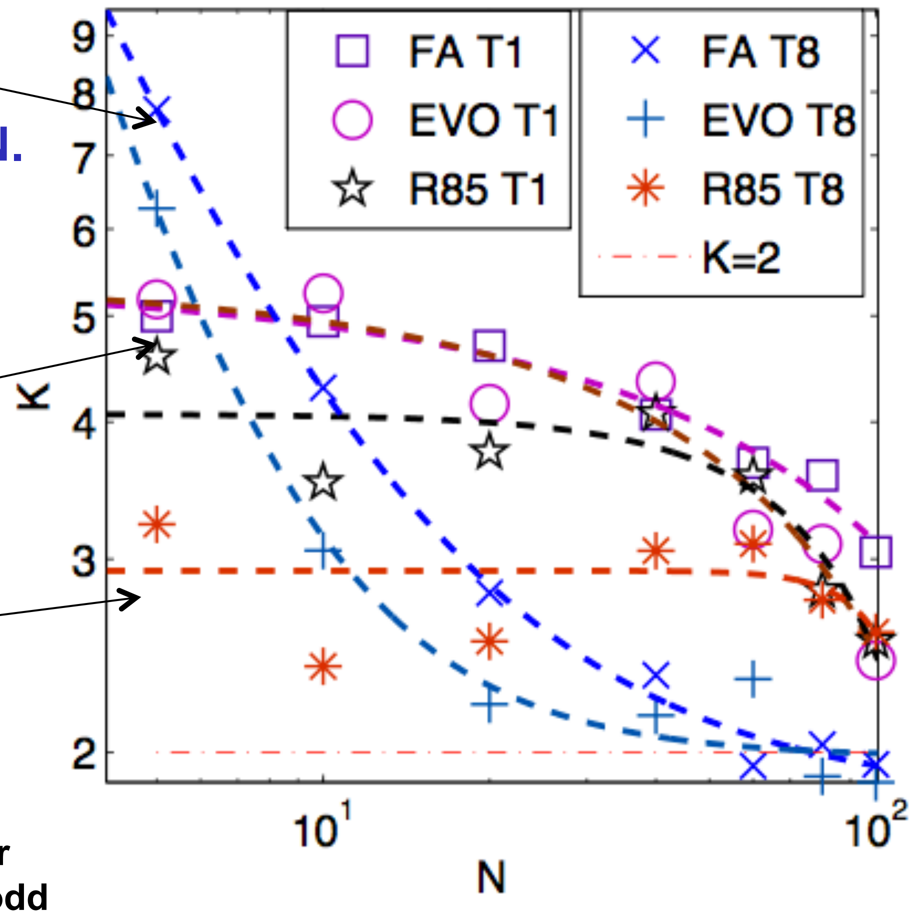
Power-Law Scaling of the K_c with T and N

Asymptotic fit: $\alpha N^{\beta} + \theta$

High info scaling regime,
Converges to K_c=2 for large N.

Low info scaling regime,
No real convergence.

Lowest info, constant,
No real convergence.



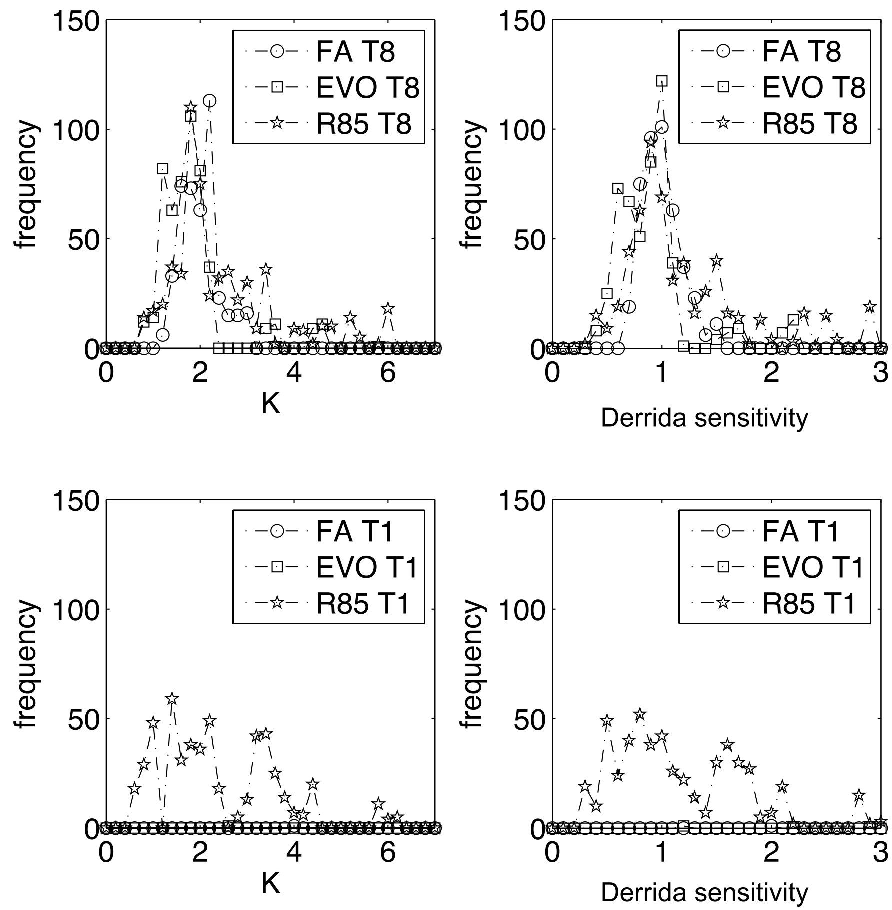
FA= full adder
EVO = even-odd
R85 = rule 85

Convergence, Criticality, Sensitivity for networks with generalization > 0.8

Derrida and Pomeau (1986) Introduced an annealed approximation method for calculating dynamical regime. DS < 1: frozen, DS > 1: chaotic, and DS = 1: complex.

Final K distribution shows sharp peak at critical connectivity K_c=2 and critical dynamics 1 (T8 high information).

Both K and DS have a wide distribution (T1 low information).



FA= full adder
EVO = even-odd
R85 = rule 85

Evolution of In-Degree Distribution

Erdős, P. and Rényi, A. (1959). Initial random graph with degree distribution:

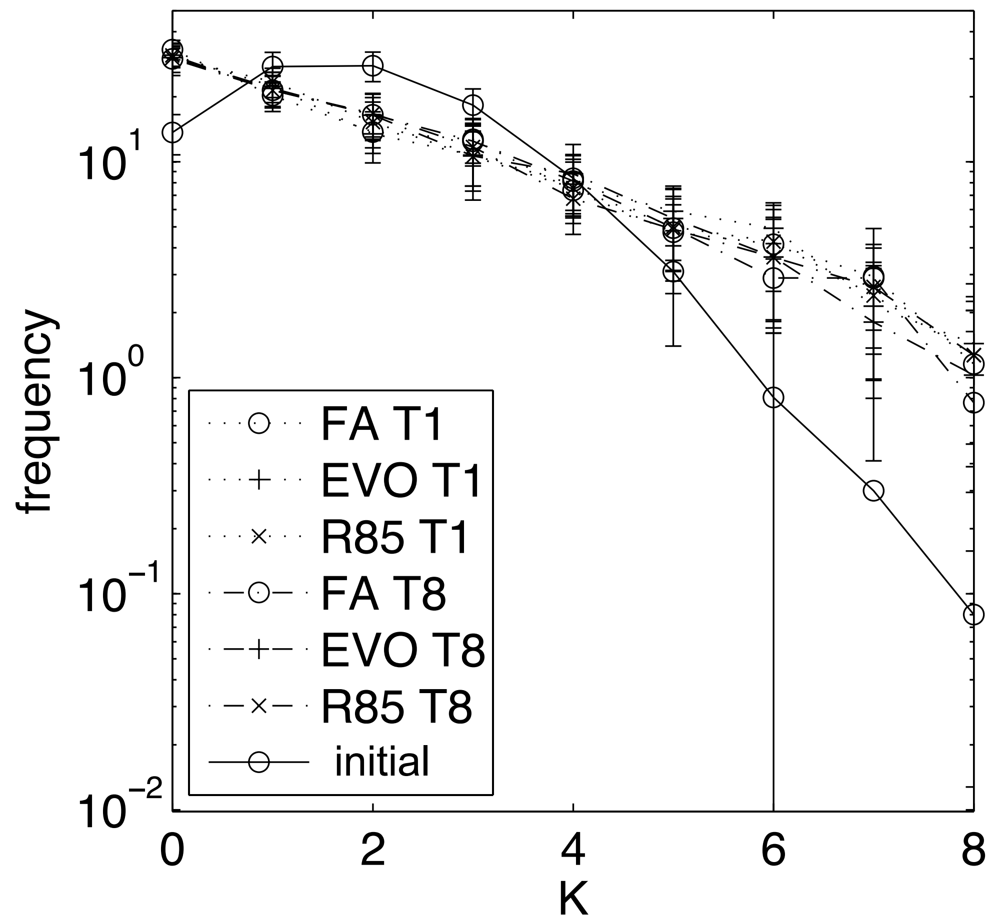
$$P(k;K) = \frac{K^k e^{-K}}{K!}$$

After evolution regardless of task and T, the distribution is:

$$P(k) = \alpha e^{-\beta k}$$

Due to maximum entropy principle. In the limit that would imply:

$$P(k;K) = \frac{e^{\frac{-k}{K}}}{K}$$



Evolution of In-Degree Distribution *(cont'd)*

Growing network by adding links to random nodes in the net result in exponential degree distribution.
S.N. Dorogovtsev and J.F.F. Mendes (2003), Alain Barrat et al. (2008)

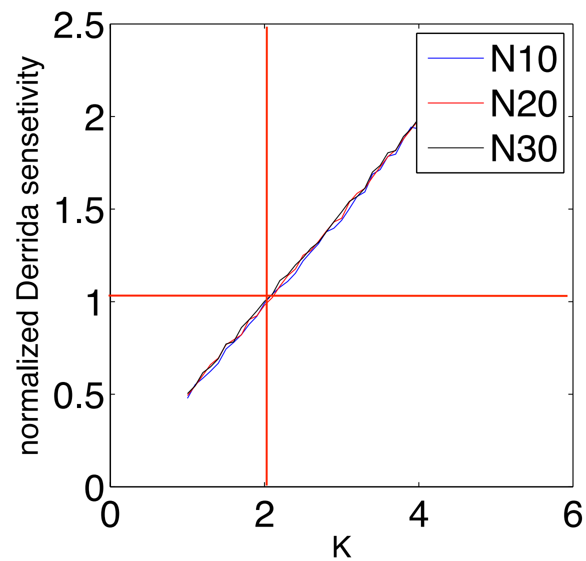
Local rewiring rules that drives networks to criticality also result in evolution of degree distribution from Poissonian to exponential.
Stefan Bornholdt and Thimo Rohlf (2000).

In biological networks such as gene regulatory networks or neural networks, where maintaining links impose cost the degree distribution is exponential not scale-free.
Amaral et al. (2000)

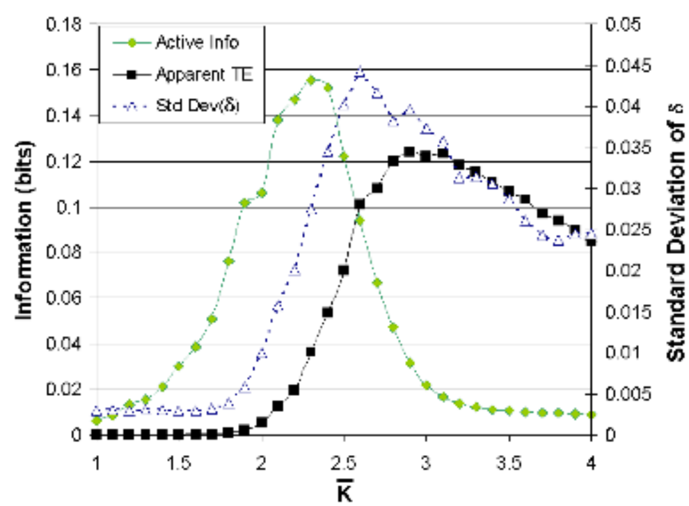
To summarize:
Unbiased evolution of networks towards maximum entropy subject to constraints such as K_c and k_{max} results in exponential degree distribution.

What is the Origin of K_c

The NK ensemble at K=2 is in a complex dynamical regime in which the information is preserve and is robust to perturbations.
Kauffman (1995), Shmulevich and Kauffman (2004)

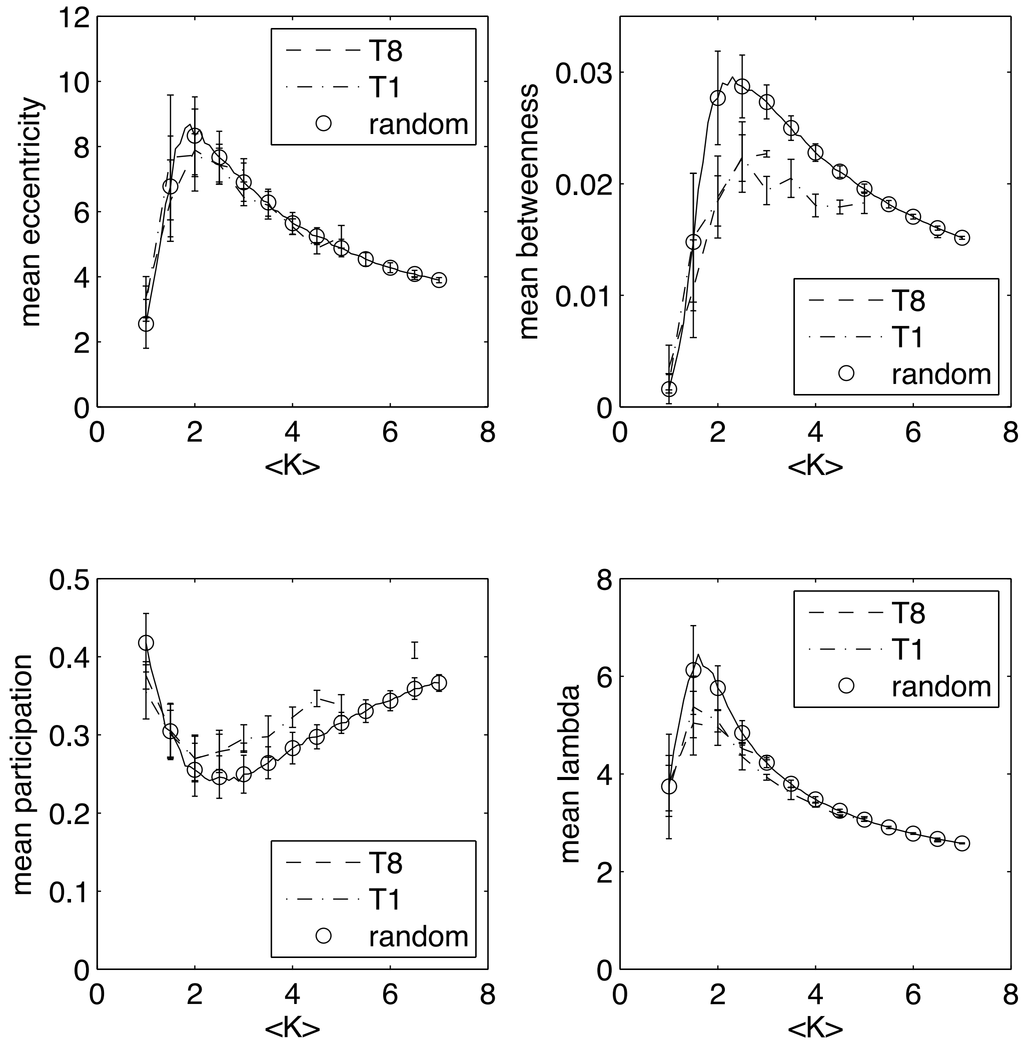


The NK ensemble at K=2 maintains a balance between information storage and information transfer.
Lizier et al. (2008)



What is the Origin of K_c *(cont'd)*

The variance of different topological measures is maximal near K_c=2. This Suggests more diversity and therefore improved search.



Conclusions

- Evolution of networks subject to generic fitness maximization results in a critical connectivity K_c in the network.
- $K_c = 2$ for large system size N and scales according to power-law for finite N.
- K_c corresponds to a region of space in the network ensemble with maximized topological diversity. This improves evolutionary search.
- During the evolution the degree distribution of the network evolves from the Poissonian to an exponential form due to entropy maximization subject to K_c and k_{max}.