Entropy for complex and non-ergodic systems – derivation from first principles

Stefan Thurner & R. Hanel & M. Gell-Mann

SECTION FOR SCIENCE OF COMPLEX SYSTEMS

SANTA FE INSTITUTE

IIASA

www.complex-systems.meduniwien.ac.at www.santafe.edu

What does it need for a theory?

- Take a few – but good assumptions (principles, postulates or axioms)

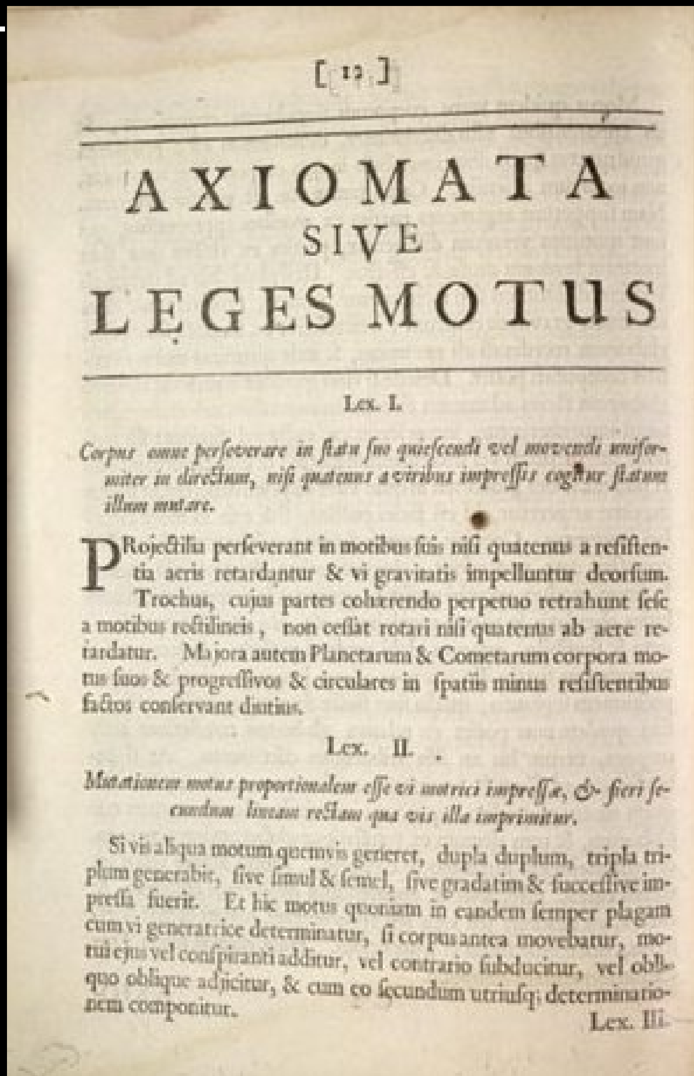
→ derive and explain many observable facts

→ bring order to hitherto unexplained separated facts

→ unite hitherto separated areas

→ find the sub-world where the theory applies and where not

- A theory is something highly coherent and not a collection of laws take …



[12]

AXIOMATA
SIVE
LEGES MOTUS

Lex. I.

Corpus omne perseverare in statu suo quiescendi vel movendi uniformiter in directum, nisi quatenus a viribus impressis cogitur statum illum mutare.

PRojectilia perseverant in motibus suis nisi quatenus a resistentia aeris retardantur & vi gravitatis impelluntur deorsum. Trochus, cujus partes cohærendo perpetuo retrahunt sese a motibus rectilineis, non cessat rotari nisi quatenus ab aere retardatur. Majora autem Planetarum & Cometarum corpora motus suos & progressivos & circulares in spatiis minus resistentibus factos conservant diutius.

Lex. II.

Mutationem motus proportionalem esse vi motrici impressæ, & fieri secundum lineam rectam qua vis illa imprimitur.

Si vis aliqua motum quemvis generet, dupla duplum, tripla triplum generabit, sive simul & semel, sive gradatim & successive impressa fuerit. Et hic motus quoniam in eandem semper plagam cum vi generatrice determinatur, si corpus antea movebatur, motui ejus vel conspiranti additur, vel contrario subducitur, vel obliquo oblique adjicitur, & cum eo secundum utriusq; determinationem componitur.

Lex. III.

… and do 300 years of physics

We try to present a theory of complex systems

- **The input: Three axioms**

→ derive the possible forms of entropy

→ bring order in the zoo of entropies

→ understand the consequences of extensivity

→ when does the MEP exist?

→ understand the relation of the MEP and extensive entropy

→ understand the possible types of constraints in MEP

→ understand why trace-forms in MEP

→ find those systems where entropies for CS apply (process→entropy)
. aging
. path-dependent
. out-of-equilibrium

→ understand the entropies of superstatistics

- **The rule of the game:** no assumptions – just derivations – just math

Entropy

$$S[p] = \sum_{i=1}^{W} g(p_i)$$

p_i ... probability for a particular (micro) state of the system, $\sum_i p_i = 1$

W ... number of states

g ... some function. What does it look like?

The Shannon-Khinchin axioms

Measure for the amount of uncertainty S

- SK1: S depends continuously on p
- SK2: S maximum for equi-distribution $p_i = 1/W$
- SK3: $S(p_1, p_2, \cdots, p_W) = S(p_1, p_2, \cdots, p_W, 0)$
- SK4: $S(A + B) = S(A) + S(B|A)$

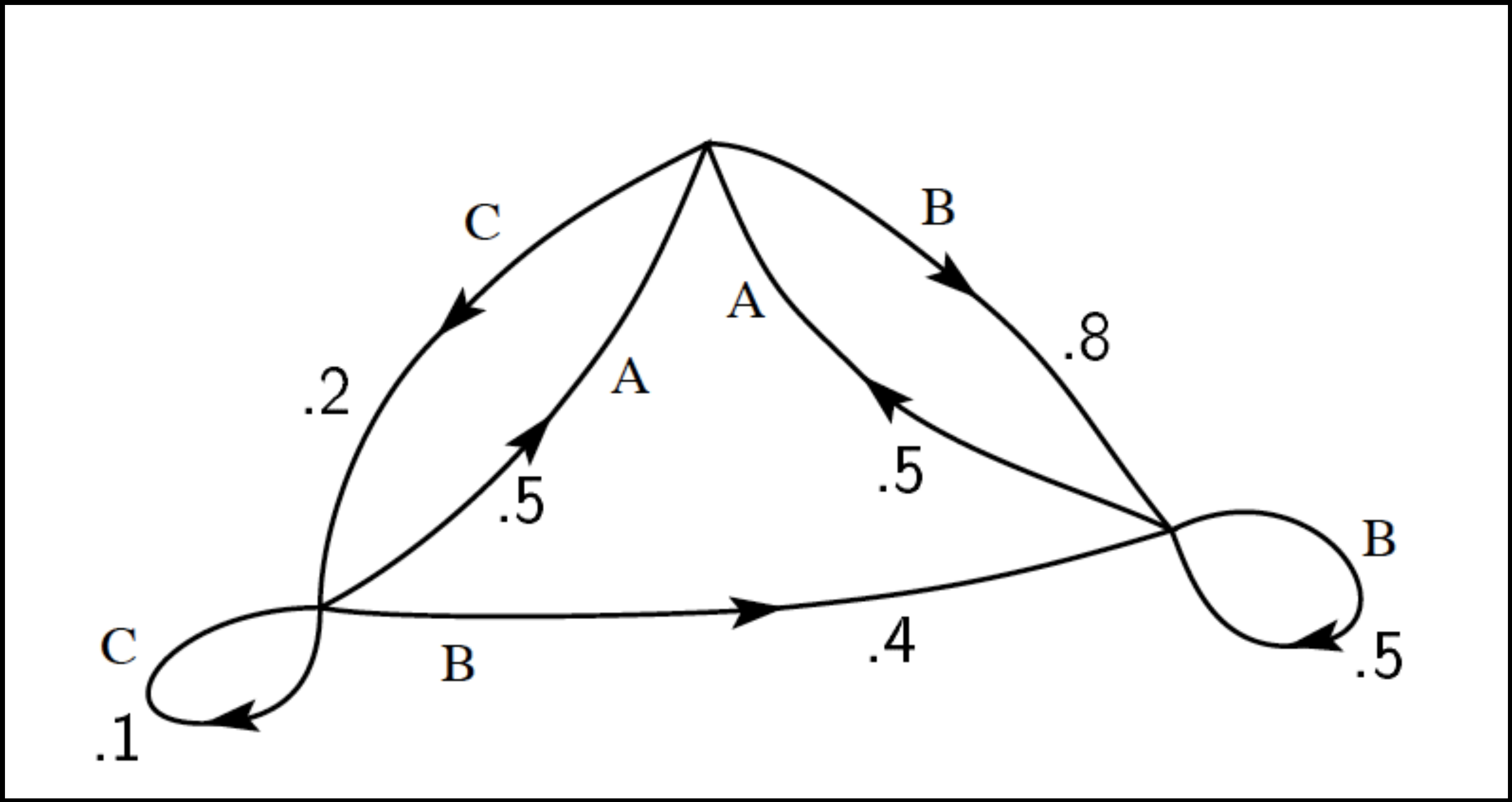
Theorem: If SK1- SK4 hold, the only possibility is $S[p] = -\sum_{i=1}^{W} p_i \ln p_i$

Appendix 2, Theorem 2, C.E. Shannon, The Bell System Technical Journal **27**, 379-423, 623-656, 1948.

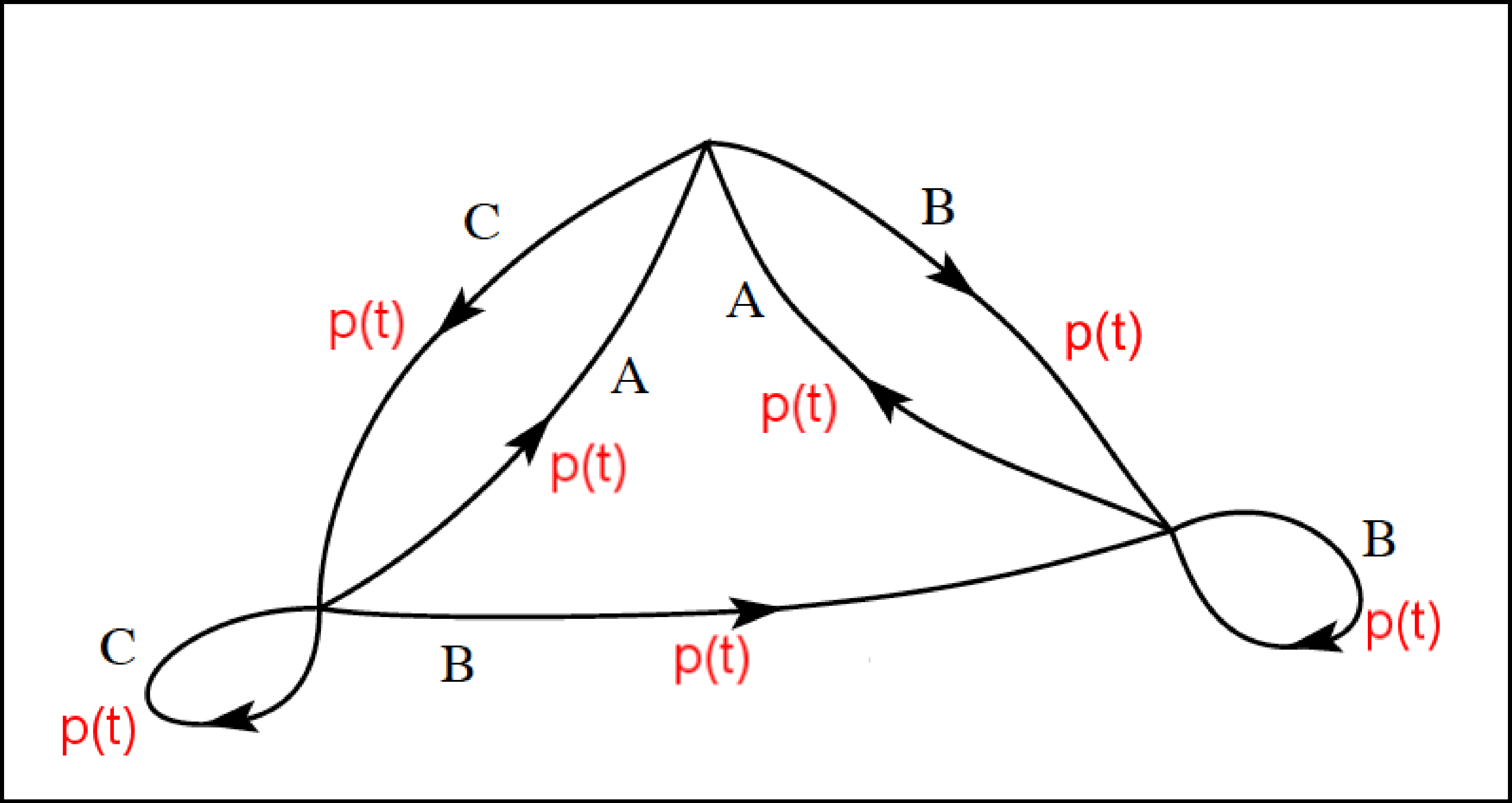
Where does this apply?

- Markov processes (no memory)
- Ergodic processes (probabilities coincide with experiment)
- Processes must be stationary

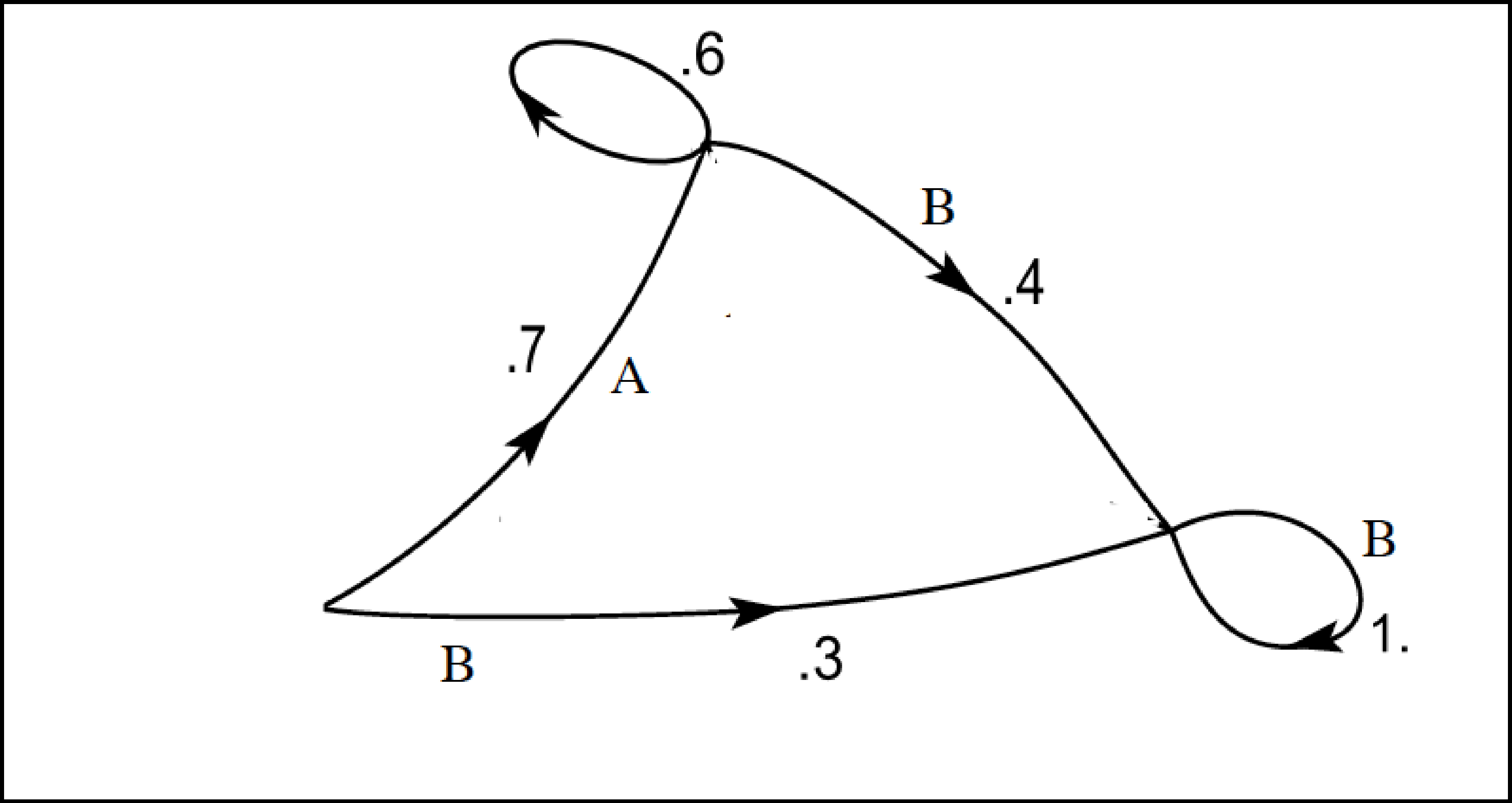
stationary and ergodic



No! no! no!



No! no! no!



We are not interested in information theory

We are interested in complex systems

What are Complex Systems ?

- CS are made up from many elements
- These elements are in strong contact with each other

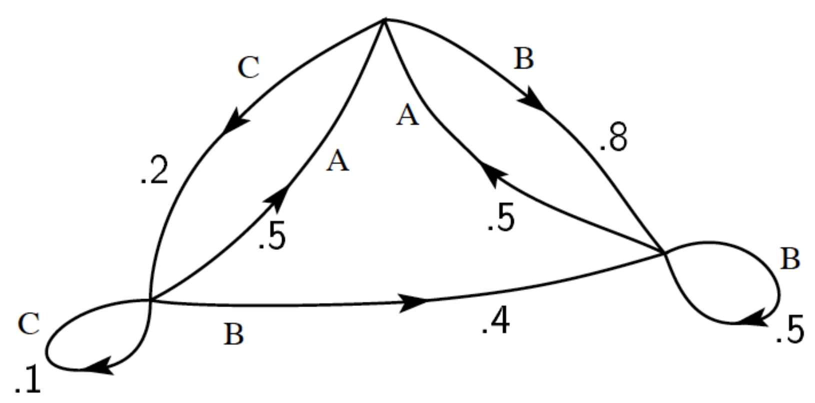
As a consequence

- CS are intrinsically non-ergodic
- CS are most often intrinsically non-Markovian

What are Complex Systems ?

- evolutionary
- path-dependent
- long-memory
- out-of-equilibrium

all of this violates ergodic, Markov, stationary



Why SM of Complex Systems ?

• Central concept: understanding macroscopic system behavior on the basis of microscopic elements and interactions → *entropy*

• Entropy relates number of states to an extensive quantity, plays fundamental role in the thermodynamical description

• Hope: From 'thermodynamical' relations of CS → phase diagrams, etc.

• Dream: reduce number of parameters → understand and handle CS ?

How should this be done?

What is the entropy of CS ?

Remember Shannon-Khinchin axioms

- SK1: S depends continuously on $p \to g$ is continuous
- SK2: S maximal for equi-distribution $p_i = 1/W \to g$ is concave
- SK3: $S(p_1, p_2, \cdots, p_W) = S(p_1, p_2, \cdots, p_W, 0) \to g(0) = 0$
- SK4: $S(A + B) = S(A) + S(B|A)$

note: $S[p] = \sum_i^W g(p_i)$. If SK1-SK4 $\to g(x) = -kx \ln x$

Shannon-Khinchin axiom 4 is non-sense for CS

SK4 ensures that system is Markovian and ergodic

$\rightarrow$ SK4 violated for non-ergodic systems

$\rightarrow$ nuke SK4

The 3 Complex Systems axioms

- SK1 holds
- SK2 holds
- SK3 holds
- $S_g = \sum_i^W g(p_i)$, $W \gg 1$

Theorem: All systems for which these axioms hold

(1) can be uniquely classified by 2 numbers, c and d

(2) have the entropy

$$S_{c,d} = \frac{e}{1-c+cd} \left[\sum_{i=1}^{W} \Gamma\left(1+d\,,\,1-c\ln p_i\right) - \frac{c}{e} \right] \qquad e \cdots \mathrm{Euler\,const}$$

TABLES

OF

THE INCOMPLETE Γ-FUNCTION

COMPUTED BY THE STAFF OF THE DEPARTMENT OF APPLIED STATISTICS, UNIVERSITY OF LONDON, UNIVERSITY COLLEGE

EDITED BY

KARL PEARSON, F.R.S.

G R

LONDON

PUBLISHED FOR THE DEPARTMENT OF SCIENTIFIC AND INDUSTRIAL RESEARCH BY

HIS MAJESTY'S STATIONERY OFFICE

1922

The argument: generic mathematical properties of g

- Scaling transformation $W \to \lambda W$: how does entropy change ?

Mathematical property I: an unexpected scaling law !

$$\lim_{W\to\infty} \frac{S_g(W\lambda)}{S_g(W)} = ... = \lambda^{1-c}$$

Define $f(z) \equiv \lim_{x\to 0} \frac{g(zx)}{g(x)}$ with $(0 < z < 1)$
Theorem 1: For systems satisfying SK1, SK2, SK3: $f(z) = z^c$, $0 < c \leq 1$

Theorem 1

Let g be a continuous, concave function on $[0,1]$ with $g(0)=0$ and let $f(z)=\lim_{x\to 0^+} g(zx)/g(x)$ be continuous, then f is of the form $f(z)=z^c$ with $c\in(0,1]$.

Proof. note

$$f(ab)=\lim_{x\to 0}\frac{g(abx)}{g(x)}=\lim_{x\to 0}\frac{g(abx)}{g(bx)}\frac{g(bx)}{g(x)}=f(a)f(b)$$

$c>1$ explicitly violates SK2, $c<0$ explicitly violates SK3.

□

Mathematical properties II: yet another one !!

$$\lim_{W\to\infty} \frac{S(W^{1+a})}{S(W)W^{a(1-c)}} = ... = (1+a)^{d}$$

Theorem 2: Define $h_c(a) = \lim_{x\to 0} \frac{g(x^{a+1})}{x^{ac}g(x)}$...

Theorem 2

Let g be as before and $f(z) = z^c$ then $h_c(a) = (1+a)^d$ for d constant.

Proof. We determine $h_c(a)$ again by a similar trick as we have used for f.

$$h_c(a) \quad = \lim_{x \to 0} \frac{g(x^{a+1})}{x^{ac} g(x)} = \frac{g\left((x^b)^{\left(\frac{a+1}{b}-1\right)+1} \right)}{(x^b)^{\left(\frac{a+1}{b}-1\right)c} g(x^b)} \frac{g(x^b)}{x^{(b-1)c} g(x)} = h_c\left(\frac{a+1}{b}-1\right) h_c(b-1)$$

for some constant b. By a simple transformation of variables, $a = bb' - 1$, one gets $h_c(bb'-1) = h_c(b-1)h_c(b'-1)$. Setting $H(x) = h_c(x-1)$ one again gets $H(bb') = H(b)H(b')$. So $H(x) = x^d$ for some constant d and consequently $h_c(a)$ is of the form $(1+a)^d$. □

Summary

CS (non-ergodic) systems $\to$ SK1-SK3 hold $\to$ 2 laws

$$\to \lim_{W\to\infty} \frac{S_g(W\lambda)}{S_g(W)} = \lambda^{1-c} \qquad 0 \le c < 1$$

$$\to \lim_{W\to\infty} \frac{S(W^{1+a})}{S(W)W^{a(1-c)}} = (1+a)^d \qquad d \text{ real}$$

Remarkable:

- all systems are characterized by 2 exponents: (c,d) – universality class
- Which S fulfills above? $\to S_{c,d} = \sum_{i=1}^{W} re\,\Gamma\,(1+d\,,\,1-c\ln p_i) - rc$
- Which distribution maximizes $S_{c,d} \to p_{c,d}(x) = e^{-\frac{d}{1-c}\left[W_k\left(B(1+\frac{ex}{r})^{\frac{1}{d}}\right)-W_k(B)\right]}$

$r = \frac{e}{1-c+cd}$, $B = \frac{1-c}{cd}\exp\left(\frac{1-c}{cd}\right)$; Lambert-$W$: solution to $x = W(x)e^{W(x)}$

Examples

- $S_{1,1} = \sum_i g_{1,1}(p_i) = -\sum_i p_i \ln p_i + 1$ (BG entropy)
- $S_{q,0} = \sum_i g_{q,0}(p_i) = \frac{1-\sum_i p_i^q}{q-1} + 1$ (Tsallis entropy)
- $S_{1,d>0} = \sum_i g_{1,d}(p_i) = \frac{e}{d} \sum_i \Gamma\left(1+d\,,\, 1-\ln p_i\right) - \frac{1}{d}$ (AP entropy)
- ...

Classification of entropies: order in the zoo

entropy		c	d
$S_{BG} = \sum_i p_i \ln(1/p_i)$		1	1
• $S_{q<1} = \frac{1-\sum p_i^q}{q-1}$	$(q < 1)$	$c = q < 1$	0
• $S_\kappa = \sum_i p_i(p_i^\kappa - p_i^{-\kappa})/(-2\kappa)$	$(0 < \kappa \leq 1)$	$c = 1 - \kappa$	0
• $S_{q>1} = \frac{1-\sum p_i^q}{q-1}$	$(q > 1)$	1	0
• $S_b = \sum_i (1 - e^{-bp_i}) + e^-b - 1$	$(b > 0)$	1	0
• $S_E = \sum_i p_i (1 - e^{\frac{p_i - 1}{p_i}})$		1	0
• $S_\eta = \sum_i \Gamma(\frac{\eta+1}{\eta}, -\ln p_i) - p_i \Gamma(\frac{\eta+1}{\eta})$	$(\eta > 0)$	1	$d = 1/\eta$
• $S_\gamma = \sum_i p_i \ln^{1/\gamma}(1/p_i)$		1	$d = 1/\gamma$
• $S_\beta = \sum_i p_i^\beta \ln(1/p_i)$		$c = \beta$	1

$S_{c,d} = \sum_i er\Gamma(d+1, 1 - c\ln p_i) - cr$	c	d

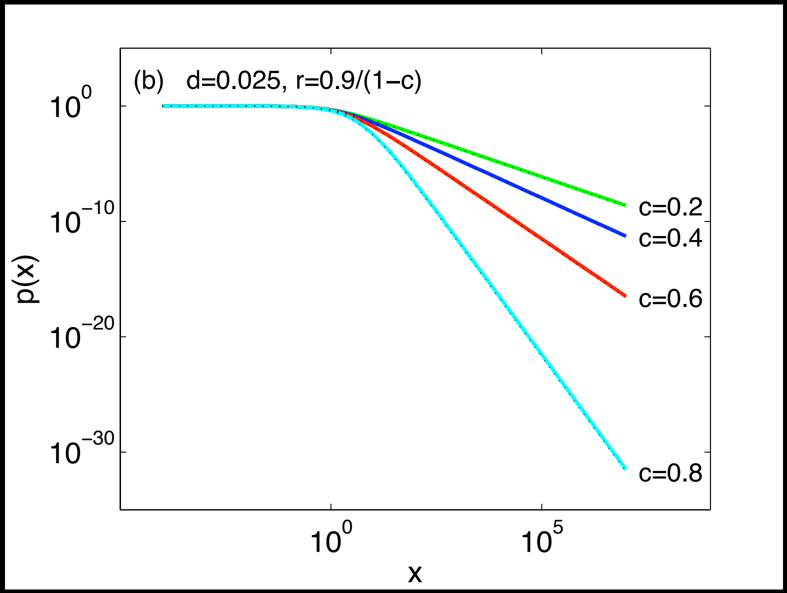
Theorem: all (c, d) entropies are Lesche stable

Distribution functions of CS

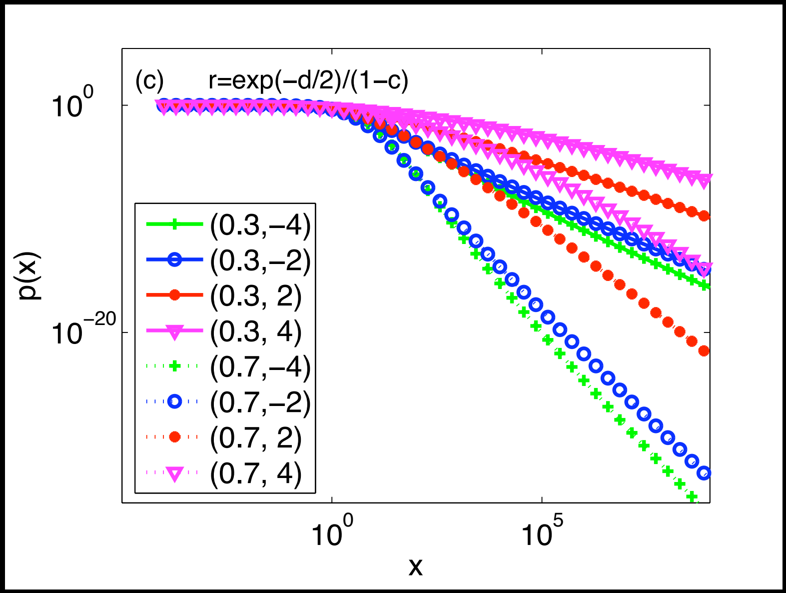
- $p_{(1,1)} \rightarrow$ exponentials (Boltzmann distribution) $p \sim e^{-ax}$
- $p_{(q,0)} \rightarrow$ power-laws (q-exponentials) $p \sim \frac{1}{(a+x)^b}$
- $p_{(1,d>0)} \rightarrow$ stretched exponentials $p \sim e^{-ax^b}$
- $p_{(c,d)}$ all others $\rightarrow$ Lambert-W exponentials $p \sim e^{aW(x^b)}$

NO OTHER POSSIBILITIES EXIST

q-exponentials

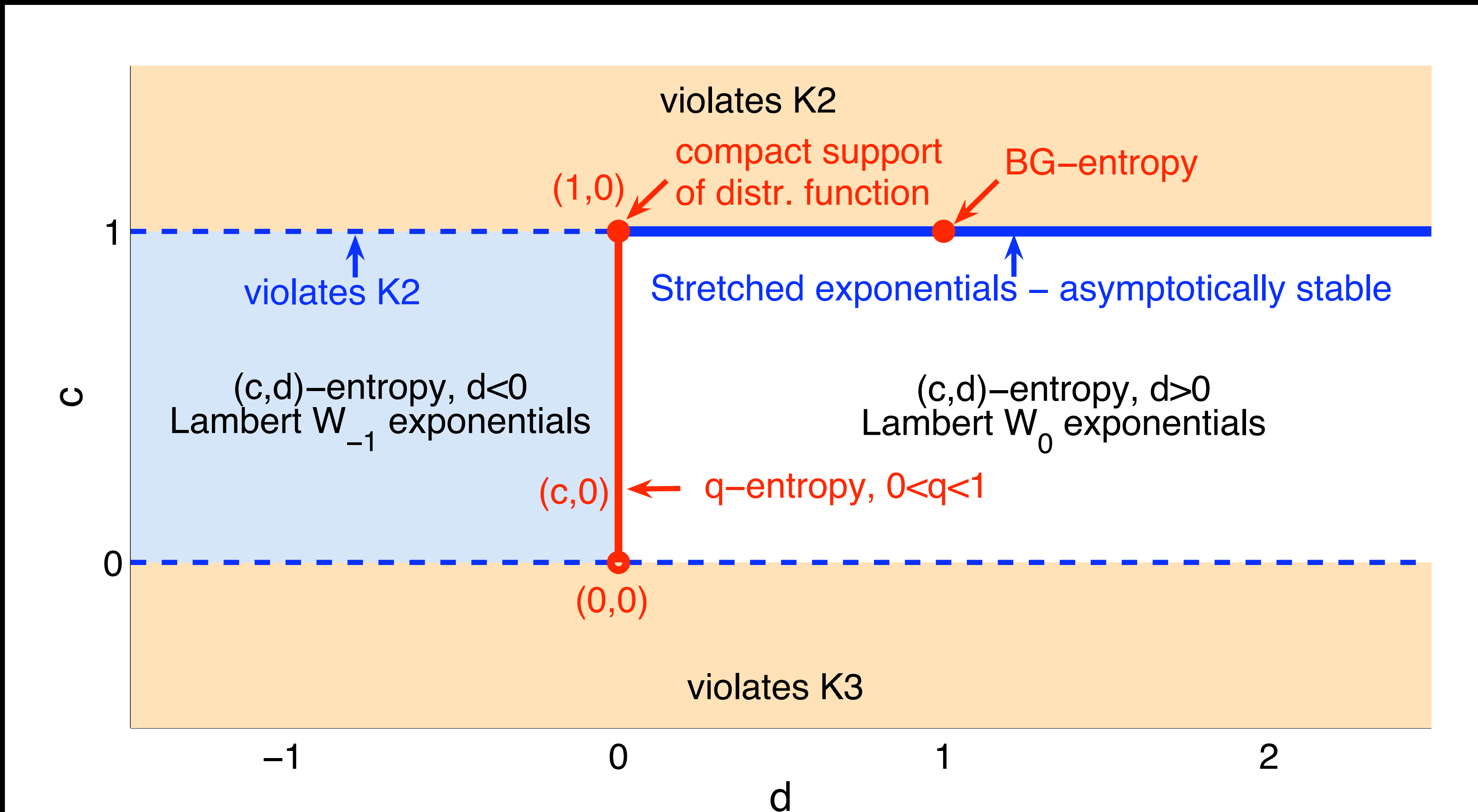


Lambert-exponentials



The Lambert-W: a reminder

- solves $x = W(x)e^{W(x)}$
- inverse of $p \ln p = [W(p)]^{-1}$
- delayed differential equations $\dot{x}(t) = \alpha x(t - \tau) \rightarrow x(t) = e^{\frac{1}{\tau} W(\alpha\tau)t}$



violates K2
compact support
of distr. function
BG−entropy
(1,0)
1
violates K2
Stretched exponentials − asymptotically stable
c
(c,d)−entropy, d<0
Lambert W_−1 exponentials
(c,d)−entropy, d>0
Lambert W_0 exponentials
(c,0)
q−entropy, 0<q<1
0
(0,0)
violates K3
−1
0
1
2
d

Relaxing ergodicity opens door to ...

- ... bring order in the zoo of entropies through universality classes
- ... understand ubiquity of power laws (and extremely similar functions)
- ... understand how Tsallis entropy emerges from non-ergodicity

$$c = ?\ d = ?$$

and the requirement of extensivity

Entropy is an extensive quantity

extensive: double the system $\rightarrow$ value of an extensive quantity doubles
intensive: double the system $\rightarrow$ quantity stays the same (e.g. temperature)

imagine 2 systems A and B. W_A is the number of states in A. W_{A+B} is the number of states in the combined system

extensive entropy means: $S(W_{A+B}) = S(W_A) + S(W_B)$

Don't confuse with additive: $S(W_A.W_B) = S(W_A) + S(W_B)$

For SM program to work: need extensive entropies

System has N elements $\to W(N)$... phase-space volume (system property)

Extensive: $S(W_{A+B}) = S(W_A) + S(W_B) = \cdots$ [use scaling laws] $\to$

Theorem: Extensivity is equivalent to $W(N) = \exp\left[\frac{d}{1-c} W_k\left(\mu(1-c)N^{\frac{1}{d}}\right)\right]$

$$
\begin{aligned}
c &= \lim_{N\to\infty} 1 - 1/N \frac{W'(N)}{W(N)} \\
d &= \lim_{N\to\infty} \log W \left(\frac{1}{N}\frac{W}{W'} + c - 1\right)
\end{aligned}
$$

Message: Growth of phase-space volume determines entropy and *vice versa*

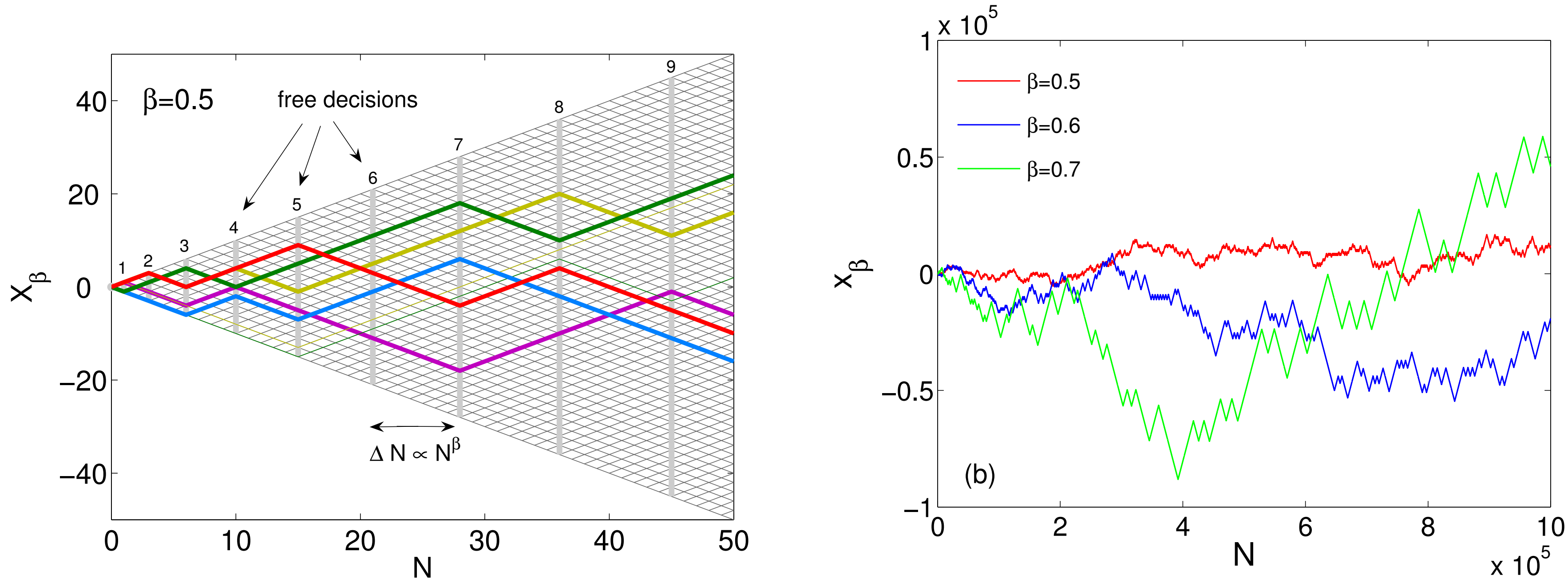
Examples

- $W(N) = 2^N \to (c,d) = (1,1)$ and system is BG
- $W(N) = N^b \to (c,d) = (1-\frac{1}{b}, 0)$ and system is Tsallis
- $W(N) = \exp(\lambda N^\gamma) \to (c,d) = (1, \frac{1}{\gamma})$
- ...

you give us your phase-space volume $\to$ we tell you the extensive entropy

Examples for extensive entropies

Example: Super-diffusion: Accelerating random walks



- up-down decision of walker is followed by $[N^{\beta}]_{+}$ steps in same direction
- $k(N)$ number of random decisions up to step $N \to k(N) \sim N^{1-\beta}$
- number of all possible sequences $W(N) \sim 2^{N^{1-\beta}} \to (c,d) = (1, \frac{1}{1-\beta})$
- note that continuum limit of such processes is well defined

Example: Join-a-club spin system

• NW growth: new node links to $\alpha N(t)$ random neighbors, $\alpha < 1$ constant connectency network A (e.g. person joining club)

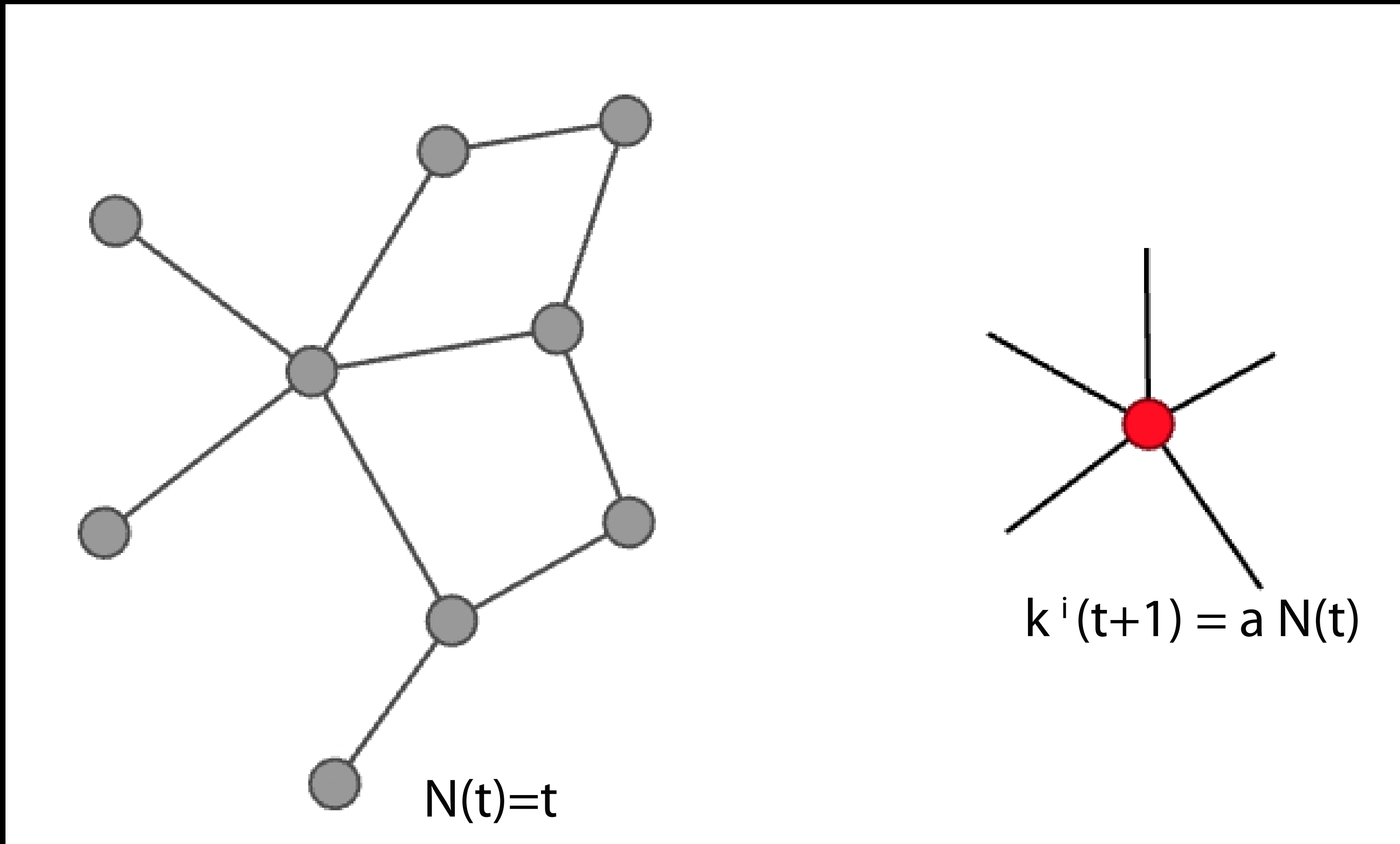
• each node i has 2 states: $s_i = \pm 1$; YES / NO (e.g. opinion)

• each node i has initial ('kinetic') energy ϵ_i (e.g. free will)

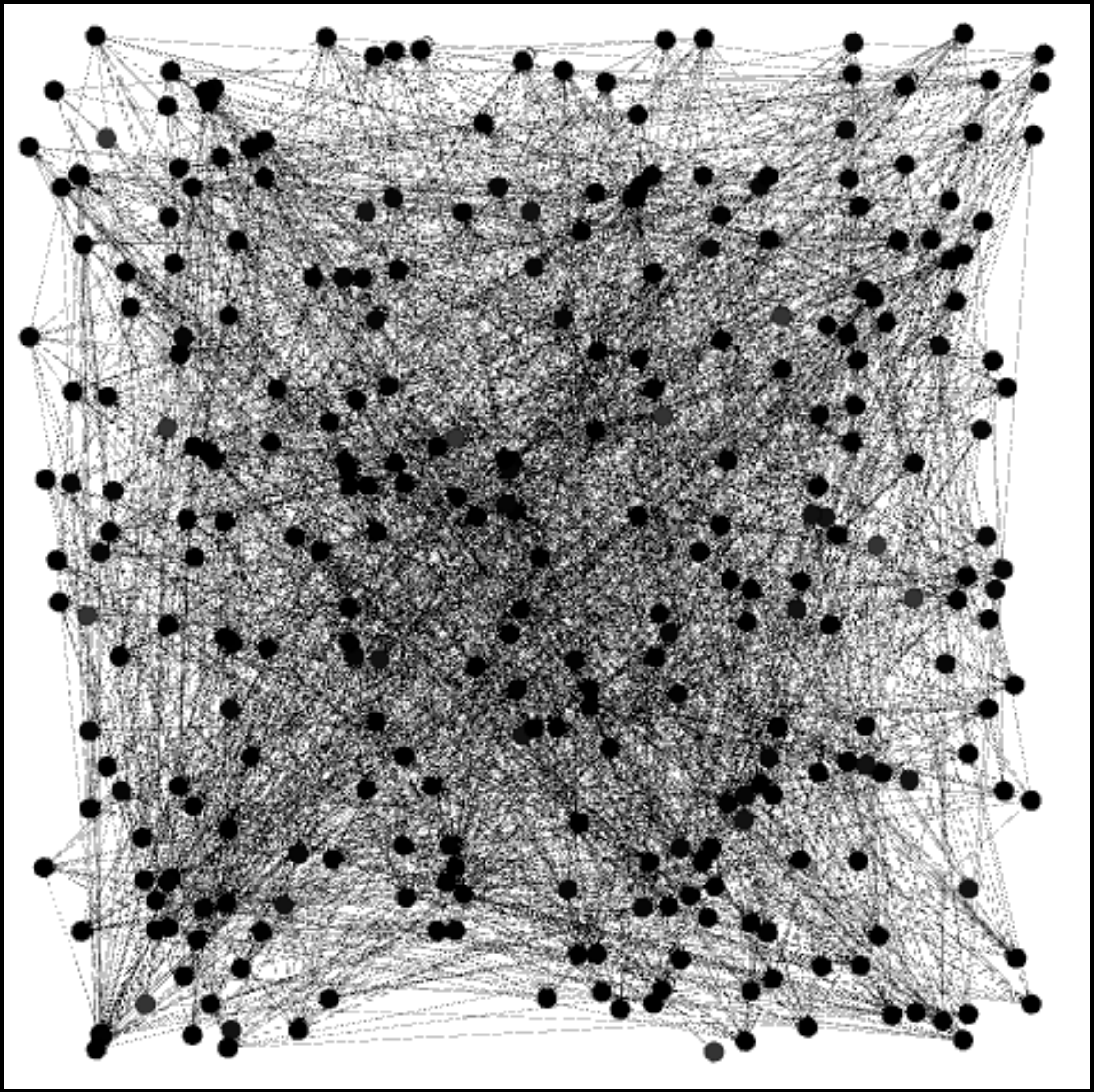
• interaction $H_{ij} = -J A_{ij} s_i s_j$

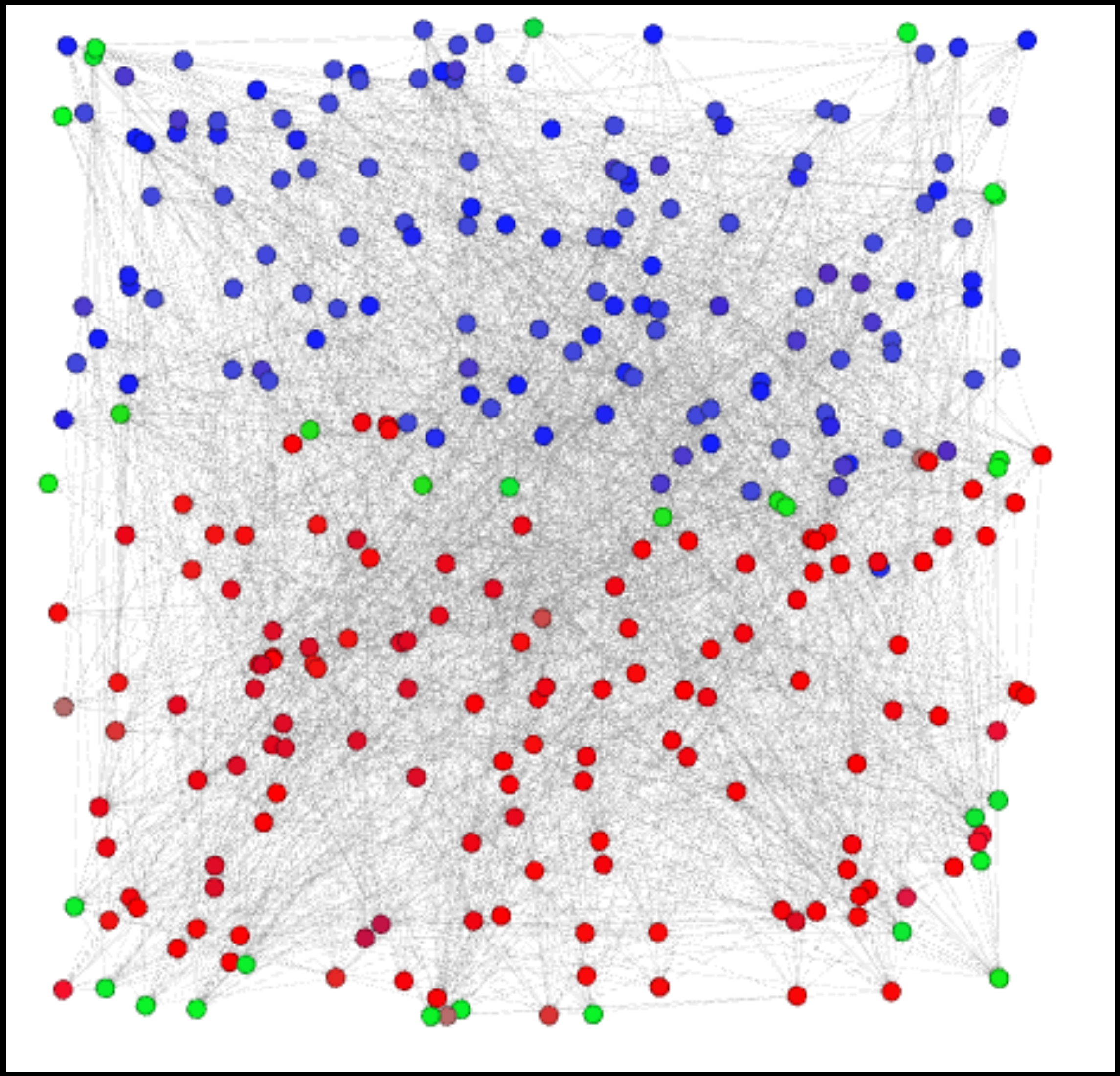
• spin-flip of node can occur if node has enough energy for it (microcanonic)

$\rightarrow$ **Can show** extensive entropy is Tsallis entropy $(c,d) = (q,0)$, $S_{c,d} = S_{q,0}$



N(t)=t
k i (t+1) = a N(t)





Example from physics: Black hole entropy

$$\log W_{\text{black-hole}} \propto \text{area}$$

- Extensive entropy is $(c, d) = (0, 3/2)$-entropy

Details, see C. Tsallis L.J.L. Cirto, arxiv 1202.2154 [cond-mat.stat-mech]

Now, what to do with extensive (c,d)-entropy?

- If you maximize it – will you get the right distribution functions?

$\rightarrow$ in general NO !

- Can the Maximum Entropy Principle be derived from the three axioms?
- How is extensive entropy related to the Maximum Entropy Principle?

$\rightarrow$ see next talk

- **The input: Three axioms**

✓ derive the possible forms of entropy

✓ bring order in the zoo of entropies

✓ understand the consequences of extensivity

→ when does the MEP exist?

→ understand the relation of the MEP and extensive entropy

→ understand the possible types of constraints in MEP see PNAS 2012

→ why trace-forms?

→ find those systems where entropies for CS apply (process→entropy)
. aging
. path-dependent
. out-of-equilibrium

→ understand the entropies of superstatistics see PNAS 2011

- **The rule of the game:** no assumptions – just derivations – just math

Conclusions

- Complex Systems are non-ergodic by nature

- Hope: describe CS with a few parameters a là thermodynamics

- Interpret CS as those where Shannon axioms 1-3 hold

- Showed: all macroscopic statistical systems can be uniquely classified in terms of 2 scaling exponents (c, d)

- Single entropy covers all systems: $S_{c,d} = re\sum_i \Gamma\left(1+d\,,\,1-c\ln p_i\right) - rc$

- All known entropies of SK1-SK3 systems are special cases

- Distribution functions of *all* systems are Lambert-W exponentials. There are no other options

- Phasespace growth determines entropy

- Systems with such entropies are related to surface effects: SOC, ...

with R. Hanel and M. Gell-Mann

RH, ST, Europhysics Letters 93 (2011) 20006
RH, ST, MGM, PNAS 108 (2011) 6390-6394
RH, ST, Europhysics Letters 96 (2011) 50003
RH, ST, MGM, PNAS 109 (2012) 19151-19154
ST, RH, Bentham e-book (2013)
RH, ST, arXiv:1310.5959

A note on Rényi entropy

It is it not sooo relevant for CS. Why?

- Relax Khinchin axiom 4:

$S(A+B)=S(A)+S(B|A)\to S(A+B)=S(A)+S(B)\to$ Rényi entropy

- $S_R=\frac{1}{\alpha-1}\ln\sum_i p_i^\alpha$ violates our $S=\sum_i g(p_i)$

But: our above argument also holds for Rényi-type entropies !!!

$$S=G\left(\sum_{i=1}^{W} g(p_i)\right)$$

$$\lim_{W\to\infty}\frac{S(\lambda W)}{S(W)}=\lim_{R\to\infty}\frac{G\left(\frac{f_g(z)}{z}G^{-1}(R)\right)}{R}=[\text{for } G\equiv\ln]=1$$