Studi numerici di sistemi disordinati e complessi

Enzo Marinari

(Roma *La Sapienza*)

Grazie agli organizzatori per l'invito!

Studi numerici di sistemi disordinati e complessi.

CASPUR, Settembre 2006

La fisica dei sistemi disordinati
e/o complessi.

Il ruolo delle (difficili) simulazioni numeriche.

Una piccola quantità di disordine (rilevante) può cambiare completamente il comportamento di un sistema fisico.

Un sistema "complesso" (Hamiltoniana o funzione obiettivo "complicata") può avere un comportamento del tutto analogo a quello di un sistema caratterizzato da disordine congelato.

Classi grandi e molto diverse di problemi rientrano in questa categoria. Ottimizzazione, ad esempio. Si usa un approccio unificante per capire, migliorare, risolvere.

La Meccanica Statistica.

Mecc Stat $\rightarrow$ Mecc Stat dei sistemi disordinati.

Senza disordine:

$$P \sim \frac{e^{-\beta H}}{Z}$$

$$\langle O \rangle \equiv \int dC \; P(C) \; O(C)$$

Medie. Domande rilevanti. Non servono le condizioni iniziali né risolvere 10^{23} equazioni...

Quando il sistema è grande $P(O)$ ha un picco stretto. canonico $\simeq$ microcanonico.

$$H_{ISING} = -\sum{}' \sigma_i \sigma_j$$

$T ==$: stato fondamentale, minimo dell'energia

(Ising:± 1)

$T \neq 0$: energia libera (principio di minimo)

$$F = E - T\ S$$

L'entropia S ha a che fare con il logaritmo del numero di configurazioni di un certa energia.

Particelle, spin $\rightarrow$ "configurazioni"

Il costo computazionale K necessario a "risolvere" il problema (l'affermazione può avere vari significati).

N; numero delle componenti del sistema. Per N grande

$$K \sim N^c \; ? \qquad K \sim e^{AN^D} \; ?$$

Nel secondo caso, problemi molto difficili, ma interessanti.

In ogni caso assumiamo: calcolo di H per una data configurazione, tempo polinomiale.

Classificazione.

P: è noto un algoritmo che risolve il problema in tempo polinomiale.

NP: $\tau > N^{\alpha}$ (ad esempio l'enumerazione per Ising, 2^N). Polinomiale su un computer non standard.

Insieme NP-completo: problemi "equivalenti", classe di equivalenza.

Caso peggiore, caso tipico: quali domande?

Il problema di determinare lo stato fondamentale di un vetro di spin appartiene alla classe dei problemi NP-completi.

$$H = -\sum \sigma_i J_{ij} \sigma_j$$

somma su un reticolo in $D = 3$ o somma su tutte le coppie, campo medio.

Le costanti di accoppiamento J_{ij} sono variabili aleatorie congelate (fisse). Per esempio valgono ± 1 con probabilità uguale a $1/2$ o sono distribuite Gaussianamente (a media nulla).

$T = 0$: ricerca dello stato di energia minima del problema (o stato fondamentale).

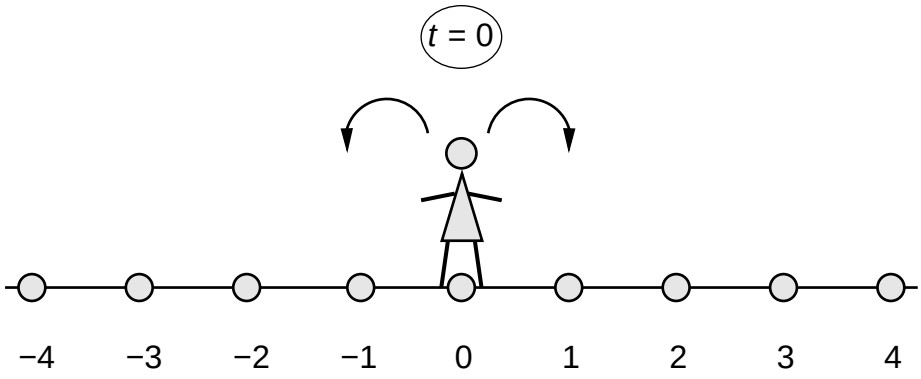
Vetri di spin come paradigma della NP-completezza.

Il disordine cambia tutto.

Cominciamo con un esempio semplice.

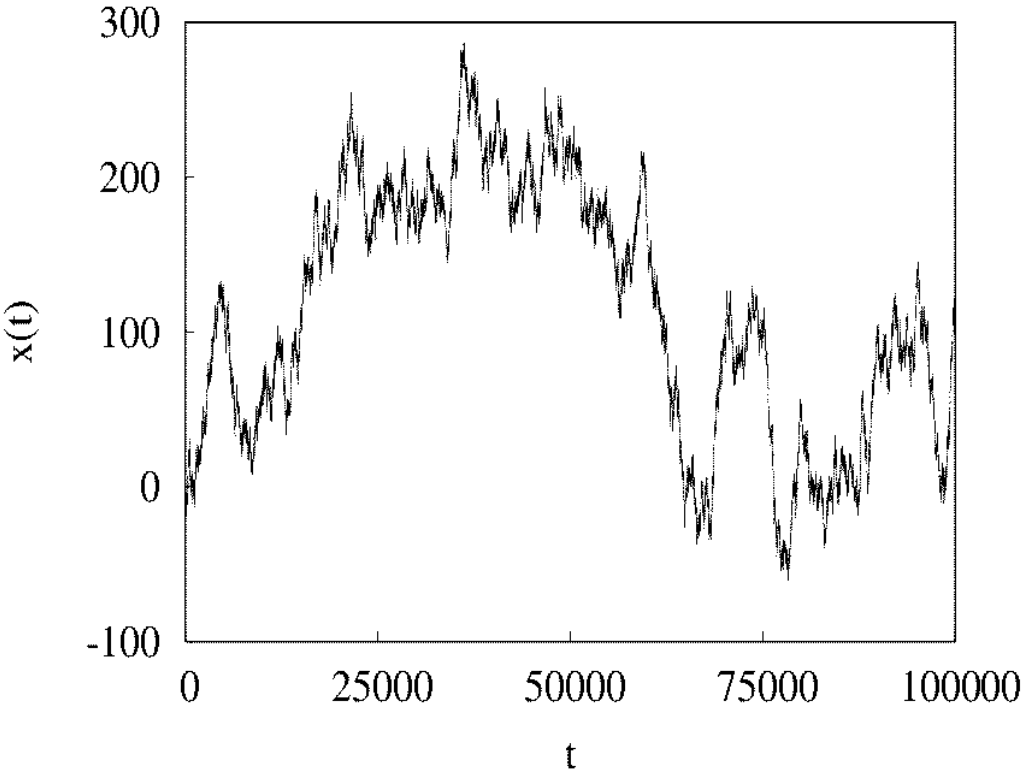
I cammini aleatori (random walk) senza disordine ambientale

Risultati cruciali di Einstein, calcolo del coefficiente di diffusione-



$$P(destra) = \frac{1}{2} \qquad P(sinistra) = \frac{1}{2}$$

(simulazioni: generatore di numeri pseudo random).


300
200
100
0
-100
x(t)
0
25000
50000
75000
100000
t

Media $\langle \cdot \rangle$, passo reticolare a.

$$\langle \Delta x \rangle = \frac{1}{2}\left((+a) + (-a)\right) = 0$$

$$\langle \Delta x^2 \rangle = \frac{1}{2}\left((+a)^2 + (-a)^2\right) = a^2$$

La posizione del viaggiatore si evolve come:

$$x_{t+1} = x_t + \Delta x_t$$

faccio il quadrato e medio:

$$\langle x_{t+1}^2 \rangle = \langle x_t^2 \rangle + \langle 2 x_t \Delta x_t \rangle + \langle \Delta x_t^2 \rangle$$

$$\langle x_{t+1}^2 \rangle = \langle x_t^2 \rangle + 0 + a^2$$

condizione iniziale ad esempio $x_0 = 0$:

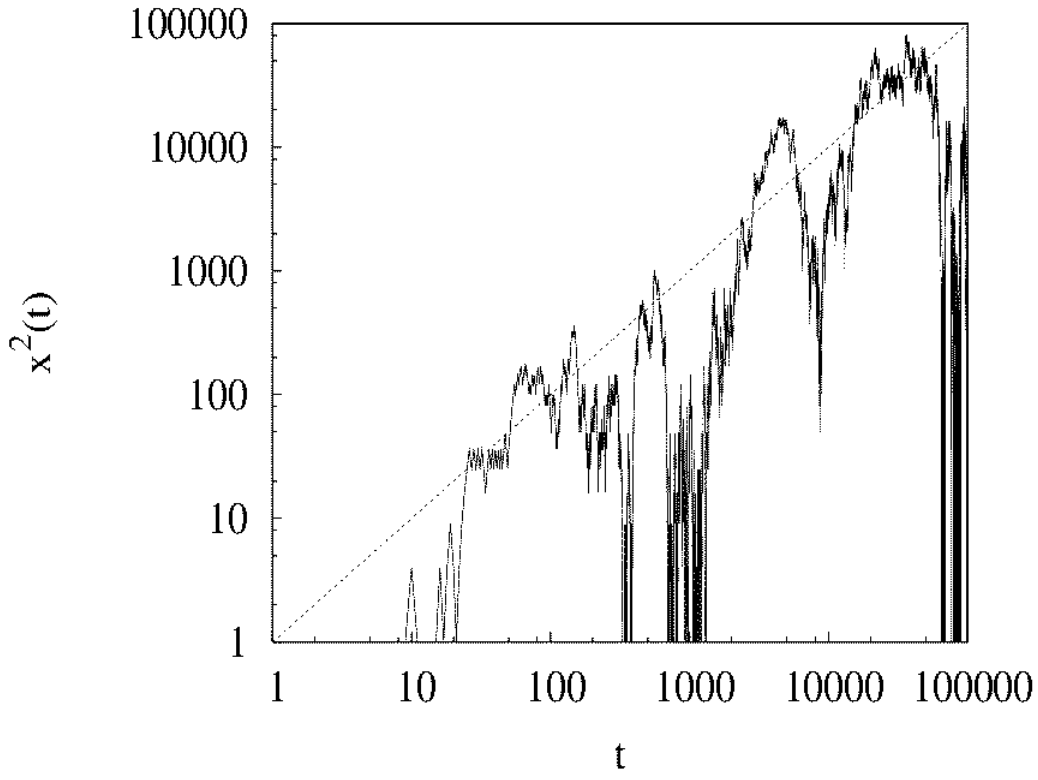
$$\langle x_t^2 \rangle = t\ a^2$$

t, non t^2!

distanza $\sim \sqrt{\text{tempo}}$

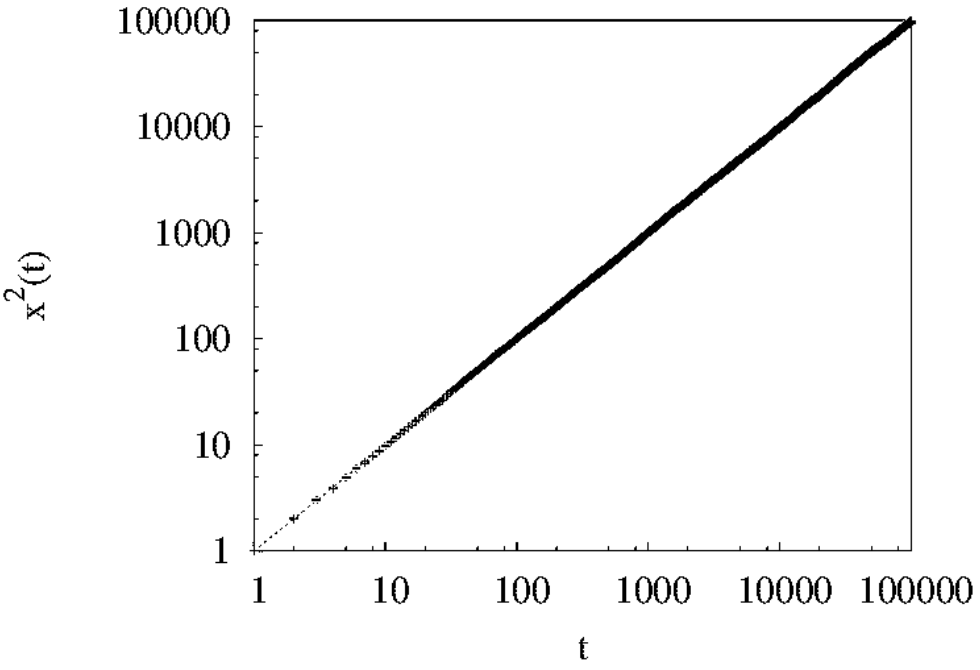
muoversi a caso rallenta ... (!!!)

numeri random



Le fluttuazioni...

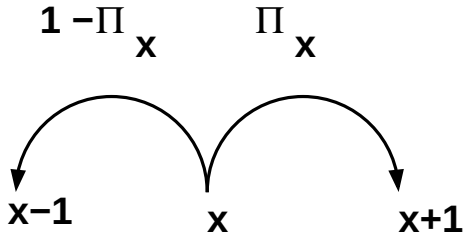
Tenerle sotto controllo, può essere difficile, vedi cosa succede quando si aggiunge disordine ambientale.



Il valore medio di x^2 fa quel che deve, e cresce come il tempo.

E se, oltre ad essere dissennato il viaggiatore, è tortuos(issima) anche la strada?

random walk in random environment (Sinai)



Le Π_x sono probabilità di salto
congelate (disordine quenched). Dipendono dal sito e non cambiano nel tempo

$$P(destra) = \Pi_x = P(\text{spostamento} = +a)$$

$$P(sinistra) = 1 - \Pi_x = P(\text{spostamento} = -a)$$

Π_x è una variabile aleatoria uniformemente distribuita fra 0 ed 1.

Adesso ci sono due tipi di rumore: il rumore termico (random walk)
ed il rumore ambientale congelato ($\{\Pi_x\}$).

Ci sono quindi due operazioni di media: $\overline{\langle O \rangle}$

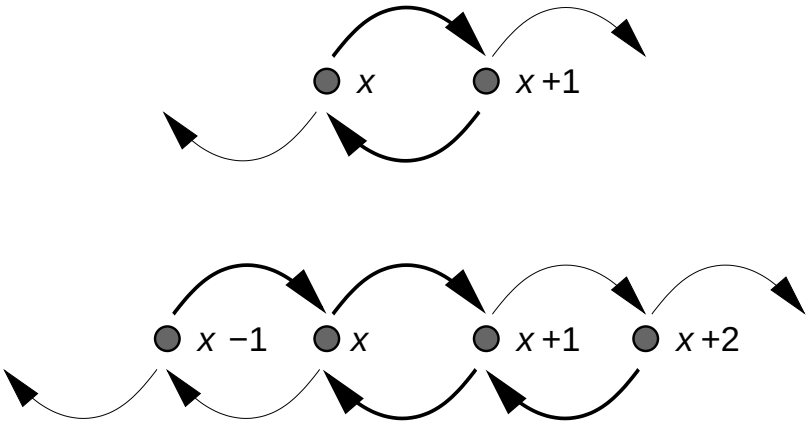
$\langle \cdot \rangle$ media termica usuale, $\overline{\cdot}$ media sul disordine.

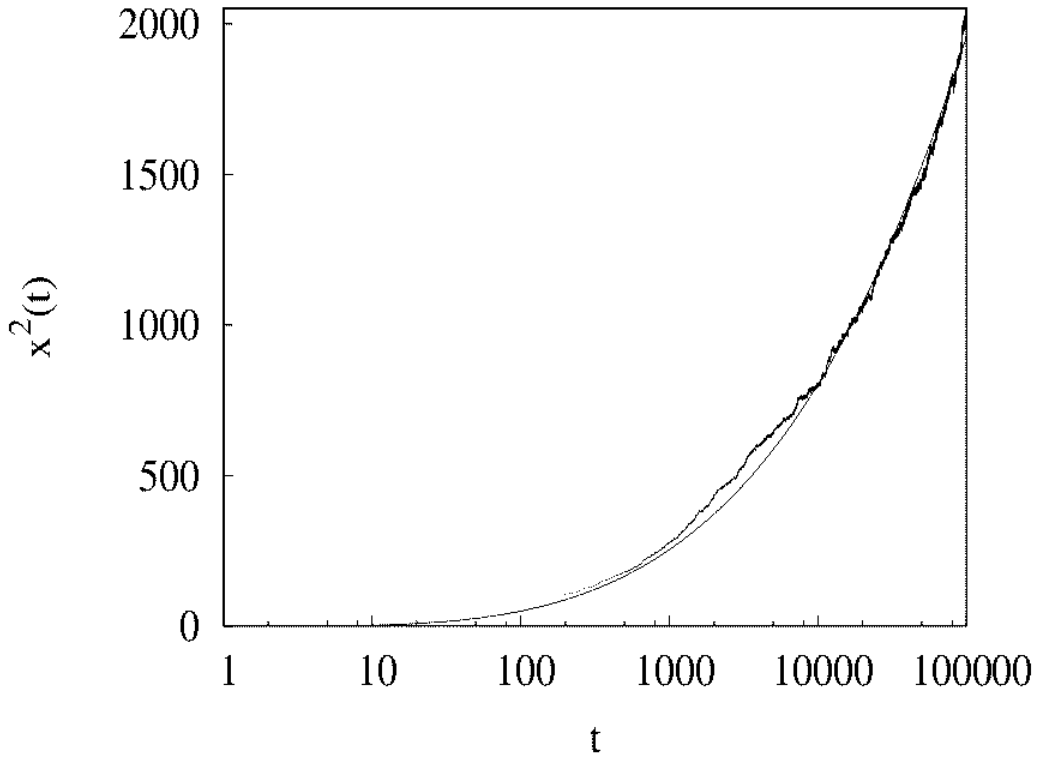
Sinai:

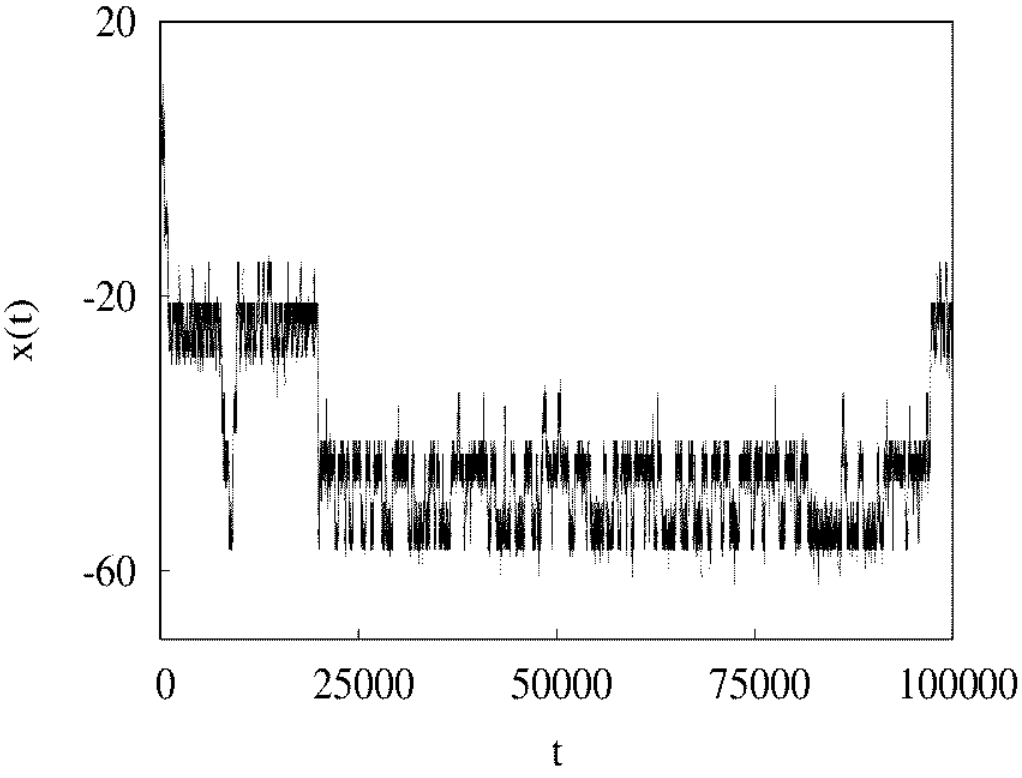
$$\overline{\langle x^2(t)\rangle} \sim (\log t)^4$$

Molto lento! Il disordine congelato ha reso la dinamica ancora piú lenta.

Motivo: trappole di grandezza arbitraria.




2000
1500
1000
500
0
x^2(t)
1
10
100
1000
10000
100000
t


20
-20
-60
x(t)
0
25000
50000
75000
100000
t

Dinamica lenta

Concetto molto generale e fondamentale

Anche qui: quello che nei sistemi fisici usuali avviene su scale di tempi esponenziali

$$C \sim e^{-\frac{t}{\tau}}$$

a causa del disordine congelato può essere logaritmicamente lento o convergere solo a potenza.

Ad esempio:

$$C \sim 1 - At^{\gamma}$$

Esempio semplice. Sistema diluito. (disordine congelato semplice). Isole e canali.

Vetri di spin (spin glasses)

Interessanti soprattutto come paradigma di sistemi in cui il disordine introduce cambiamenti drammatici.

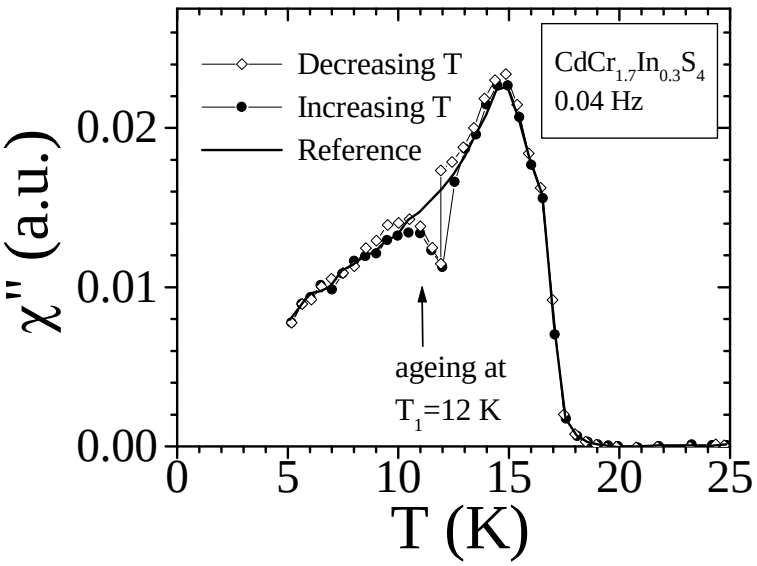
(mercati finanziari: differenze di atteggiamenti fra gli agenti
sistemi biologici: componenti eterogenee
ottimizzazione: condizioni logiche assegnate in modo [almeno apparantemente] aleatorio)

Soluzione di Parisi: risultato analitico cruciale per l'apporssimazione di campo medio (realistica in molti contesti).

Spin glasses have been, in the last years, lucky materials.

I mean: they are not very interesting from any technological point of view. Still, some people care a lot about them.

But: very interesting and new phenomena. And a very interesting theory. From material physics to structural glasses, and a paradigmatic role leading to optimization,



These features are natural in a hierarchical valley approach (but very possibly are not explained by models, see later...).

The standard model (maybe not really correct...):

$$H_{EA} \equiv - \sum_{\text{i nn j}} \sigma_i J_{ij} \sigma_j \ ,$$

Edwards-Anderson spin glass. $\sigma_i = \pm 1$ (or Heisenberg), sum is over first nearest neighboring sites, in D spatial dimensions.

$P(J) \sim e^{-J^2}$ or $J = \pm 1$ with probability $\frac{1}{2}$.
Random couplings are quenched: that gives the huge complexity.

One cannot solve this model (one cannot even more or less understand it...). So define mean field theory:

$$H_{SK} \equiv - \sum_{\text{all couples i j}} \sigma_i J_{ij} \sigma_j \ ,$$

where the sum runs now over all spin couples (Sherrington and Kirkpatrick).

The mean field theory can be solved, and it shows a really peculiar behavior. Recently a mathematical proof of the correctness of the solution has been obtained (Talagrand building on Guerra work).

Parisi mean field solution. Use replicas.

$$\log Z = \lim_{n \longrightarrow 0} \frac{Z^n - 1}{n} \ ,$$

and defining with the over-line $\bar{\cdot}$ the (quenched) disorder average:

$$\overline{\log Z} = \lim_{n \longrightarrow 0} \frac{\overline{Z^n - 1}}{n} \ ,$$

So compute $Z(n)$ from Z, Z^2, Z^3, ..., $Z^{(1)} \cdot Z^{(2)} \cdot Z^{(2)} \cdot ...$, and define $\lim_{n \longrightarrow 0}$ from analytic continuation of $Z(n)$.

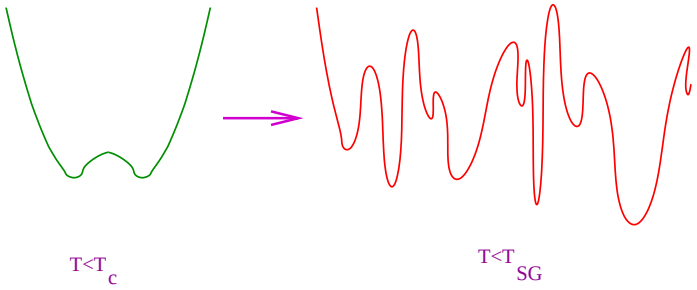
Natural guess: in $Z^{(1)} \cdot Z^{(2)} \cdot Z^{(3)} \cdot ...$ all replicas are symmetric.

Parisi: replica Symmetry undergoes spontaneous symmetry breaking.

A complex phase space emerges.

Main features of Sherrington-Kirkpatrick, Mean Field Spin Glasses phase space.

1. Complex phase space:

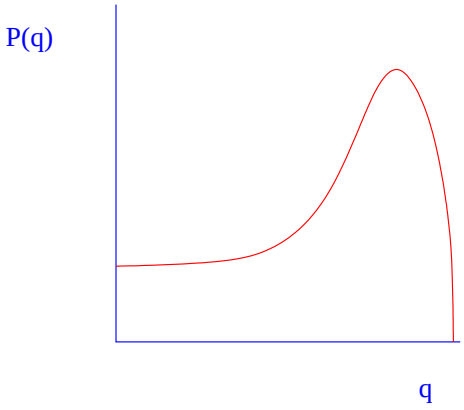


2. This situation implies the phase transition also exists for $h > 0$.

3. Two microscopic configurations of the system at equilibrium can be very similar or very different. Measure the overlap

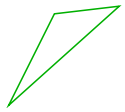
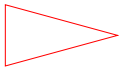
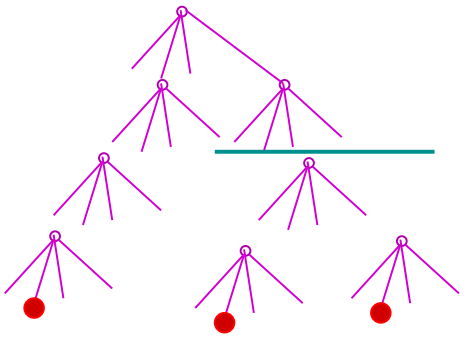
$$q \equiv \frac{1}{N} \sum_i \sigma_i^{(\alpha)} \sigma_i^{(\beta)} \ .$$

$P(q)$ has measurable support when $N \longrightarrow \infty$ for $T < T_c$.



4. States that characterize the system at equilibrium obey an ultrametric inequality for the distance:

$$\delta^{\alpha\beta} \simeq 1 - q^{\alpha\beta} \equiv 1 - \frac{1}{N}\sum_i \sigma_i^{(\alpha)}\sigma_i^{(\beta)}.$$

So, not only triangular inequality:  (obvious since δ is a distance) but also UM inequality. All triangles are isosceles with two equal sides longer than the shortest side. They can be equilateral (same state).  States turn out to be organized on a ultrametric tree.  How much of this is shared from 3D spin glass? This is not clear (I would say much) but maybe it is not too relevant (very relevant systems in many contexts are mean field like in nature).

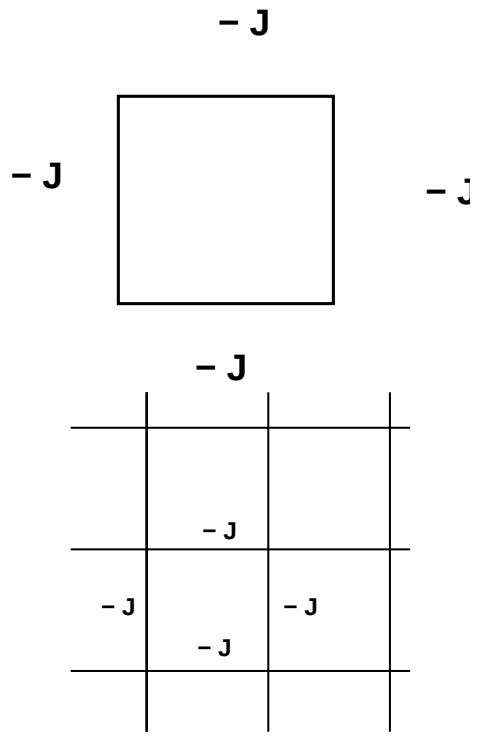
A theoretical "very complex" structure is nowadays well understood.

Il ruolo della frustrazione.

Cruciale (sistemi complessi e frustrazione auto-oindotta).

Complesso = Disordinato.

AF quadrato, non complesso

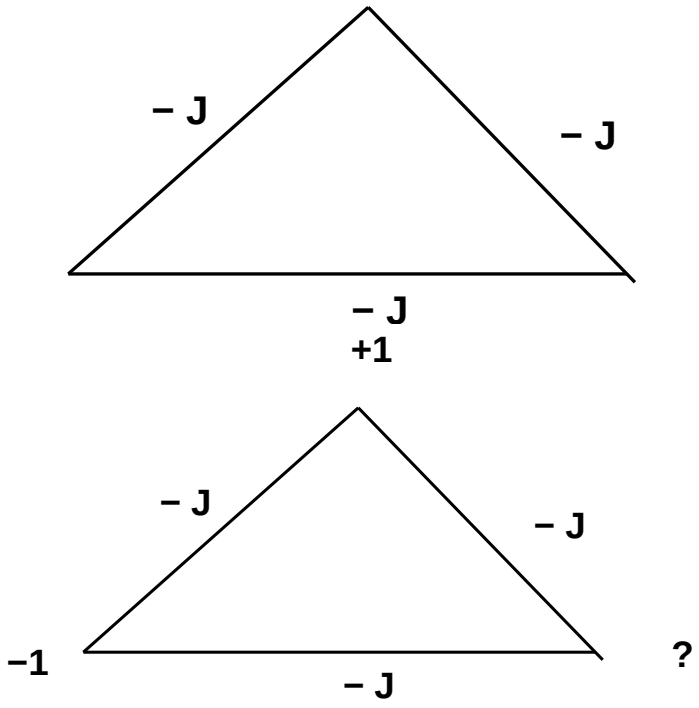


ferromagnete: $\{+1\} \longrightarrow \{-1\}$, che succede nel AF quadrato?

staggered, equivalente.

Perchè?

AF triangolare

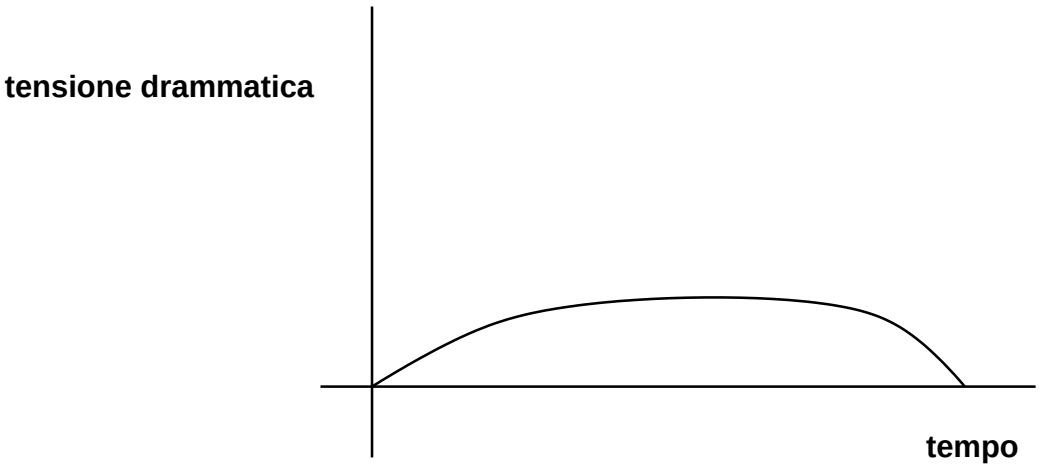


Problema diverso

Degenerazione (molti più stati fondamentali e molti piú stati)

Sfere in una scatola (vetro)

Shakespeare:



tensione drammatica = frustrazione

Vetri di spin in 3D: i "veri" materiali.

$$H \equiv -\sum \prime \sigma_i J_{ij} \sigma_j$$

somma su un reticolo $3D$.

Non è ancora del tutto chiaro cosa succeda!
(cosa succede quando in un problema cambio la struttura della matrice di connessione)

Due ipotesi predominanti:

1. MF, Parisi. MF è $D = \infty$.
Valido in $D > D_c^u$.
Simile in $D \geq D_c^L$.

2. Droplet, Moore, Bray, Fisher, Huse.
D finita o $D = 3$ del tutto diverso da $D = \infty$.
Eccitazioni frattali.
Domini complessi che però non si estendono in tutto il volume decidono la dinamica del sistema.

Abbiamo già visto. Osservabili diverse che per un ferromagnete. Ripetiamo per completezza. Scopi (non ancora raggiunti) delle simulazioni numeriche in D finita (dove al contrario che per il campo medio di SK non sono disponibili risultati analitici esatti).

$P_J(q)$, $P(q) = \overline{P_J(q)}$

$P_J^E(q_E)$, $P^E(q_E) = \overline{P_J^E(q_E)}$

UM? Organizzazione gerarchica? Transizione in campo?

3-Sat

Here one (very important!) example among many: K-SAT (much more, all very connected: matching problems, traveling salesman, ...).

N Boolean variables $\{x_i\}_{i=1,N} \quad x_i \in \{0,1\}$

M clauses (constraints) $C_1, \ldots C_M$ consisting of the OR of 3 distinct variables (or of their negations).

Example: $x_3 \vee \overline{x_{12}} \vee \overline{x_{24}}$

For a given assignment of the x_i, a clause is TRUE or FALSE

Decision problem (SAT)

$\exists$ assignment $\{x_i\}_{i=1,N}$ such that $F(\{x_i\})$ is TRUE?

Optimization problem (MAX-SAT)

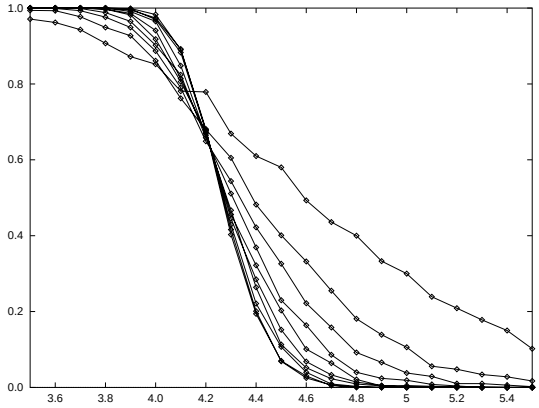
Find an assignment $\{x_i\}_{i=1,N}$ which minimizes E_{GS} = the number of FALSE clauses

The problem can be written in terms of a Statistical Mechanical Hamiltonian, making it very similar to a Spin Glass.

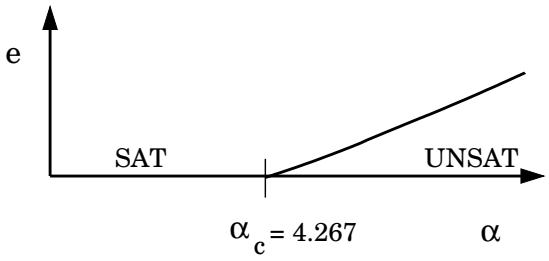
Random 3-Max-Sat

Each clause is chosen at random with the flat distribution among the $2^3 N(N-1)(N-2)/3!$ possibilities.

SAT-UNSAT phase transition

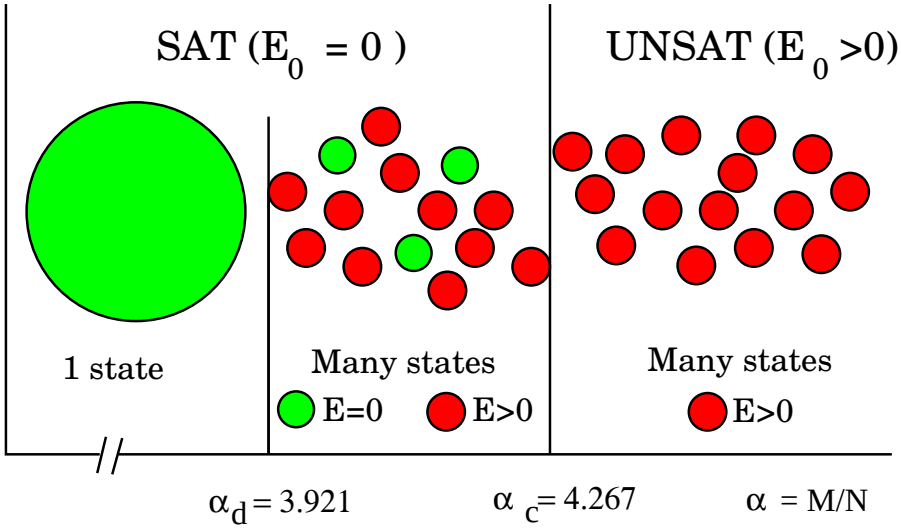


Probability that a random sample is satisfiable versus the control parameter $\alpha = M/N$



Ground state energy density $e = E_{GS}/N$ versus α

Statistical description of the clustering



Mézard, Parisi, Zecchina and coworkers.

- Order parameters
- Organization of the ground states (para, spin glass,...)
- Energy landscape ($E > E_{GS}$, barriers, ...)
- Thermodynamics ($T > 0$)

Da questo approccio si deduce un nuovo algoritmo (Mézard, Parisi e Zecchina).

Molto efficiente.

Survey propagation (da belief propagation, approccio probabilistico).

In molte situazioni batte di ordini di grandezza gli algoritmi noti.

Interessante: si usa tutto quello che si sa di un problema per scrivere il miglior algoritmo possibile.

Monte Carlo e Tempi di Correlazione

Si veda il capitolo 3 in "Monte Carlo Methods for the Self-Avoiding Walk" di Alan D. Sokal, preprint hep-lat/9405016, pubblicato in "Monte Carlo and Molecular Dynamics Simulations in Polymer Science", edito da Kurt Binder, Oxford University Press.

Annealing

Che cosa è.

Ricerca dello stato fondamentale. Tanti minimi.

MC

Alte T. N_T passi.

Faccio scendere $T \to T' \to T''$ con N_T, $N_{T'}$, $N_{T''}$ passi.

Ogni tanto faccio una discesa a $T = 0$ (Lo stato fondamentale non si trova partendo da T basse! Infatti lí sono già congelato, quasi certamente in una valle sbagliata).

Annoto tutti gli stati di energia piú bassa che trovo.

Ma perchè trovo lo stato fondamentale?
Sembra impossibile.

$$N_S \sim e^{kN}$$

C'è un motivo profondo.

Nella gran parte dei sistemi fisici complessi gli stati profondi sono larghi.

Contro-esempio: Bernasconi o p-spin.

La schedule (agenda)

lineare

logaritmica

Comunque decisa a mano...

MC ottimizzato: metodo equivalente per l'energia libera invece che per l'energia (conta anche quanto una certa classe di configurazioni appare).

verso il tempering... che fornisce anc he uno scheduling automatico ed un nuovo metodo di ottimizzazione.

Metodi Monte Carlo Ottimizzati

si veda il preprint cond-mat/9612010, E. Marinari, "Optimized Monte Carlo Methods", pubblicato in "Advances in Computer Simulation", edito da J. Kertész e I. Kondor (Springer-Verlag, Berlino 1998), p. 50,

Computing Ground States.

Here ground states are (very) interesting.

- Optimization: solution.
- Statistical Mechanics: interesting physics, connected to low T physics.

Different approaches to ground state computation:

- Exact. Very effective in $2D$ with at least one open boundary. Very limited in $3D$.
- Heuristics: typically population dynamics, genetic algorithms, many scales and renormalization.

L'algoritmo, alla lavagna.

Heuristic approach to GS computation

For example Pal, Houdayer-Martin, EM-Parisi, Palassini-Young, building on Kernighan-Lin, Kawashima-Suzuki.

- Local Search.
- Families $\Longrightarrow$ genetic selection.
- Renormalization.

Local Search Flip if "gain". Single spin flip, double spin flip...

Renormalization

$$H \equiv -\sum \sigma_i J_{ij} \sigma_j - \sum h_i \sigma_i \ .$$

Consider K such spin configurations (tentative ground states...): $\sigma_i^{(\alpha)}$. Compare these configurations, use block spins, and use their similarities/differences to update them.

Population Dynamics A genetic algorithm allows to select/improve/variate the K configurations.

Children tend to be better than parents (but keep variability!).

Calcolo esatto della funzione di partizione di vetri di spin in 2D. Un algoritmo efficiente ed elegante.

Ottimizzazione.

Vari risultati fisici.

A. Galluccio, J. Lukic, E.M., O. C. Martin and G. Rinaldi, Phys. Rev. Lett. **92** (2004) 117202.
Additional details in Biophys. Chem. **115** (2005) 109 J. Lukic, E.M. and O. C. Martin, Phys. Rev. Lett. **92** (2004) 117202;
and in J. Lukic, E.M. and O. C. Martin, Progress of Theoretical Physics Supplement No.157 (2005) 17.
Also J. Lukic, E.M. and O. C. Martin, Finite size scaling in Villain's fully frustrated model and singular effects of plaquette disorder, cond-mat/0507530 and EPL in press;
and T. Jörg, J. Lukic, E.M. and O. C. Martin, Scaling behavior of diluted $2D$ spin glasses, PRL 2006
J. Lukic, E.M. and O. C. Martin, Chaos in $2D$ spin glasses, JSTAT 2006

Compute (many) exact partition functions of (large) 2D Ising SG, $J = \pm 1$ with PBC. Apply the same method to understand fully frustrated $2D$ finite size scaling and what happens when adding small amount of random quenched unfrustrated plaquettes.

Regge and Zecchina and Galluccio, Löbl and Vondrák: Pfaffians, modular arithmetics, Chinese remainder theorem.

Results:

- Believe you solve dispute, but reopen the issue right ahead in different terms. Physical scaling as $\beta \longrightarrow \infty$ c_V does not behave as $e^{-A\beta}$ with $A = 4$ as from naive scaling. Find $A < 4$.
- hyperscaling works.

- Ground state properties ($\theta^E = \theta_{DW} = 0$ etc...).
- Number of excitations (possible mechanism for anomalous scaling).
- MKA anomalous scaling.
- To understand better go through fully frustrated Villain model finite size scaling.
- A small amount of quenched random unfrustrated plaquettes completely changes the scaling behavior.
- Go back to $2D$ spin glasses with binary couplings, on larger lattices, by using dilution and joining forces of Monte Carlo and partition function computation: get new point of view.

Summary

- Spin glasses, $2D$ Ising spin glasses. The quenched physics.
- $T = 0$ and low T physics. Choice of couplings and "universality".
- Monte Carlo versus ground states computations. Computations of Z_β.
- The dispute: Swendsen and Wang with (optimized) Monte Carlo versus Kardar-Saul with exact transfer matrix.
- The Galluccio-Löbl-Vondrák algorithm.
- Our findings. The anomalous scenario.

SG: frustration + disorder (complexity)

Quenched averages

$$H = -\sum_{\texttt{nn}\ i,j} J_{ij}\sigma_i\sigma_j$$

couplings J_{ij} are random quenched.

Huge interest:

1. Parisi solution of mean field SK theory.
2. paradigmatic role (boring as materials, as such).

Open debate on behavior in finite D.

"For sure": $D_c^L \geq 2$, no transition in $2D$ for $T > 0$.

As $T \longrightarrow 0$: scaling theory. Coarse graining and scaling Ansatz.

$$\tilde{J}(l) \sim l^{\theta}$$

effective coupling among (block) spins at large distance.

$2D$: $\theta < 0$ (and/or zero, see later). Coupling becomes weaker at large scale, and the ordered state is unstable and breaks down.

Typical choices for the probability distribution of quenched couplings: $P(J) \sim \exp(-J^2/2)$ or $J = \pm 1$ with uniform probabilities, or many other possibilities (but equivalent, see later).

In $2D$ this can play (and does play) a role in deciding the "critical behavior" as $T \longrightarrow 0$ (see C. Amoruso, EM, O. Martin and A. Pagnani, PRL **91** (2003) 087201).

Starting point is

$$\delta E \equiv E_{GS}^{(P)} - E_{GS}^{(AP)}$$

and as $L \longrightarrow \infty$

$$\overline{\left(\delta E - \overline{\delta E}\right)^2} \sim L^{2\theta}$$

While for Gaussian J one has $\theta = -0.28$ for binary couplings one finds clearly (Hartmann and Young) $\theta = 0$.

We find that all $P(J)$ which can only produce quantized energies give $\theta = 0$, while all distribution that can generate continuous energies without a gap give $\theta \simeq -0.28$ (even if for example are built on only two coupling values, but with an irrational ratio).

1. $D \leq D_c^L \Longrightarrow$ small δE values are relevant for the large distance behavior of the system $\Longrightarrow$ gap in coupling distribution can play a role.

2. $\theta_{(D)} = 0$ does only mean $D \leq D_c^L$, not $D = D_C^L$.

Monte Carlo versus Ground State computations.

1. MC for Spin Glasses is very difficult.

   "Naive" MC is basically of no use. High free energy barriers make impossible exploring the full phase space.

   Optimized Monte Carlo (multicanonical, replica MC, parallel tempering) helps.

   Still: it is difficult to go at low T. You are never sure you thermalized...

2. Computing GS you study directly T=0 physics. No problems with thermalization.

   Main problem: what do you learn, say, about finite T physics? (it seems it works...).

3. A third approach: compute directly the full partition function. "Best of both worlds" (but: depending on the algorithm only reach some observables) (but: can only do it in some models, see later...).

The dispute

Saul and Swendsen (PRL **38** (1988) 4840) after a very accurate optimized MC simulation claimed to detect an anomalous scaling behavior.

2D Ising Spin Glass, $J = \pm 1$, Periodic Boundary Conditions. V=128^2

$$c_V \sim \beta^2 e^{-A\beta} \quad , \qquad A = 2 \; .$$

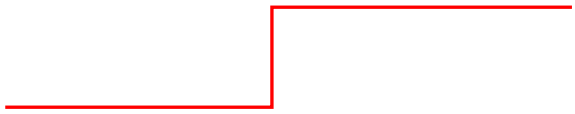
Would expect $A = 4$, since minimal excitation costs $4J$.

Periodic Boundary Conditions 1D Ising model analogy. Minimal excitation is $4J$, since $\Downarrow \Uparrow \Downarrow$

kink − antikink

Still, an easy computation gives $c_V \sim \beta^2 e^{-2\beta}$.

Now. With fixed boundary conditions minimal excitation only costs $2J$.



kink

But infinite volume limit does not depend on boundary conditions... Answer: kink-antikink excitation is no elementary. Notice that there are too many of them, $O(V^2)$.

$T \longrightarrow 0$, V fixed: eventually find $e^{-4\beta}$. But scaling limit, small T and large V: $e^{-2\beta}$.

Kardar and Saul, NP B **432** (1994) 641. They reanalyzed the problem by computing exactly the full partition function. $2D$ Ising Spin Glass with PBC, $J = \pm 1$.

They follow Kac and Ward:

1. From high T expansion, in terms of closed graphs on the square lattice (including graphs wrapping around the lattice).

$$Z = 2^V \left(\cosh\left(\beta J\right)\right)^{2V} \sum_{c:B} A_B \tanh\left(\beta J\right)^B$$

where the sum is on closed graphs with B bonds.

2. Kac-Ward $\longrightarrow$ the problem is rephrased in a local random walk with non-trivial weights. $4V \times 4V$ hopping matrix.

PBC: need four matrices (see later Regge-Zecchina and Galluccio-Löbl theorem for graphs of bounded genus). In this case one finds (Potts-Ward, 1955):

$$Z = \frac{1}{2}\left(-Z_1 + Z_2 + Z_3 + Z_4\right)$$

$$Z_\lambda = 2^V \left(\cosh\left(\beta J\right)\right)^V \sqrt{\det\left(1 - U_\lambda \tanh\left(\beta J\right)\right)}$$

U_λ: 4 different hopping matrices, of size $4V \times 4V$.

So: $\{J_{ij}\} \longrightarrow$ four matrices $U_\lambda \longrightarrow$ traces of U_λ^W for $W \leq V \longrightarrow$ polynomial in $e^{-\beta}$, density of states $\longrightarrow$ $Z = \sum_E N(E) e^{-\beta E}$.

Lot of precautions to deal with large numbers (Kardar and Saul also compute zeroes of Z).

Polynomial time estimated roughly as $\sim V^{3.2}$.

They have basically:

L	S
4-8	8000
10-14	2000
16-18	800

and few samples for larger lattices.

This turns out to be too small...

So, they disagree with Swendsen-Wang, and claim

$$c_V \sim \beta^3 e^{-4\beta}$$

(note the anomalous power, see fully frustrated Ising model in $2D$).

Number of excitations looks smaller than in $1D$ Ising. Claim is here that

$$\log V < S_1 - S_0 < \log V^2$$

But, again, the authors notice (as a sign of severe warning) that they cannot clearly detect the asymptotic behavior.

Our approach (Galluccio, Löbl and Vondrák PRL **84** (2000) 5924)

Similar to Kardar-Saul, but many further:

1. theoretical results
2. technical improvements

Summary:

Z_β^{ISG2D} $\longrightarrow$ generating function of cuts

Galluccio-Löbl: it is possible to solve the Max Cut problem in polynomial time for any graph of genus bounded by a constant. The method provides directly the generating function of cuts.

$\longrightarrow$ Eulerian subgraphs

$\longrightarrow$ perfect matching

$\longrightarrow$ (on graphs of bounded genus) Pfaffian computation (square root of the determinant of an antisymmetric matrix). Need 4^g Pfaffians.

$\longrightarrow$ compute Pfaffian by using modular arithmetics (no need for infinite precision).

$\longrightarrow$ use the Chinese Remainder Theorem to reconstruct the exact partition function.

Cut of a graph $G = (V, E)$ (vertices, edges) is a partition of its vertices into two disjoint subsets $V_1, V_2 \subset V$ and the implied set of edges between the two parts (each edge can carry a weight w_e, and the total weight of the cut is $w(C)$).

Max Cut (min Cut): divide vertices in two parts so that total weight of edges between the two parts is max (min).

Generating function of cuts: polynomial

$$\sum_{\text{over all cuts}} x^{w(C)} .$$

Eulerian subgraph: set of edges U such that each vertex of V is incident with an even number of edges from U.

Perfect matching: set of edges P such that each vertex of V is incident with exactly one edge from P.

From Ising to Cuts

Assign spins to $+1$ or -1. $V_+ = \{i \in V | \sigma_i = +1\}$ $V_- = \{i \in V | \sigma_i = -1\}$. Let $C(V_+, V_-)$ be the cut of spins $+1$ and -1. $W \equiv \sum_{\{i,j\} \in E} J_{ij}$ is the sum of all edge weights in G.

$$H = \sum_{\{i,j\} \in C} J_{ij} - \sum_{\{i,j\} \in (E-C)} J_{ij} = 2w(C) - W$$

Let the generating function of cuts be

$$\mathcal{C}(G, x) = \sum_{\text{cuts in G}} c_k x^{w(C)} ,$$

where c_k is the number of cuts with weight k.

$$Z(\beta) = \sum_{\{\sigma\}} e^{-\beta H} \simeq \sum_{\text{cuts}} e^{-2\beta w(C) + \beta V} \simeq e^{\beta V} \mathcal{C}(G, e^{-2\beta})$$

From cuts to Eulerian subgraphs

$$\mathcal{C}(G, e^{-2\beta}) \sim \quad x^{\frac{V}{2}} \quad \prod_{\{i,j\} \in E} \left(\frac{x^{\frac{w_{ij}}{2}} + x^{-\frac{w_{ij}}{2}}}{2} \right)$$

$$\mathcal{E} \quad \left(G, \frac{x^{\frac{w_{ij}}{2}} - x^{-\frac{w_{ij}}{2}}}{x^{\frac{w_{ij}}{2}} + x^{-\frac{w_{ij}}{2}}} \right)$$

$\mathcal{E}$: generating function of Eulerian subgraphs.

By the Fischer construction Eulerian subgraphs can be rewritten as a perfect matching problem.

- Planar graphs
  and graphs of bounded genus
  Perfect matching can be translated to a Pfaffian computation (of 4^g Pfaffian).

- Modular arithmetics. Work modulo some given prime number.
  Theorem: Let $P(x)$ be a polynomial of degree n with integer coefficients, $\Phi(p)$ a finite field of size $p > n$, and x_0, x_1, ... x_n distinct elements of $\Phi(p)$. Then there exists a unique polynomial of degree n over $\Phi(p)$ such that

  $$Q(x_i) = P(x_i) \bmod p, \quad i = 0, \ldots, n \ .$$

  The coefficients of $Q(x)$ are equal to the coefficients of $P(x)$ mod p.

- The Chinese Remainder Theorem.
  If we work in a number large enough of fields, i.e. p_1, p_2, ..., p_k such that

  $$\prod_{i=1}^{k} p_i > 2^n$$

  we can reconstruct the exact polynomial, i.e. the exact partition function. Great!

Summary of the Algorithm

1. Find prime numbers p_i such that

$$\prod_{i=1}^{k} p_i > 2^V \ .$$

   For each of them repeat steps 2, 3, 4 performing all operations in $\Phi(p_i)$.

2. Select $(m+1)$ distinct elements x_j of $\Phi(p_i)$. For each of them repeat step 3.

3. Write the 4^g matrices encoding the relevant orientations of the modified graph. This gives Z_β (in the point $e_\beta = x_j$).

4. From these values of $Z_\beta(\ \mathrm{mod}\ p_i)$ in given points interpolate in $\Phi(p_i)$ and get the coefficients of the polynomial.

5. Apply the Chinese Remainder Theorem: compose the results from each $\Phi(p_i)$ to get the full Z_β.

Complexity: $O(V)$ finite fields, $O(V)$ evaluations in each field (for edge weights bounded by a constant), $O(V^{\frac{3}{2}})$ operations for a single evaluation of a polynomial $\Longrightarrow$ Total $O(V^{\frac{7}{2}})$.

Technically this approach and implementation by Galluccio, Löbl, Rinaldi and Vondrák looks full of very brilliant ideas.

Main features:

- parallel;
- no problems with precision;
- basically only bound by CPU time, not by memory or word length;
- scaling $V^{\frac{7}{2}}$.

Our work. J. Lukic, A. Galluccio, EM, O. Martin, G. Rinaldi.

$2D$ Ising Spin Glass, PBC, $J = \pm 1$.

For example:

L	S
6	400000
10	100000
30	10000
40	1000
50	300

(and similar values for different L values).

$$F_J(\beta) = -\frac{1}{\beta} \log Z_J(\beta) \quad , \; U_J(\beta) = \langle H_J \rangle \quad ,$$

$$c_V = L^{-2} \frac{dU_J}{dT} \; ,$$

and average over samples. We mainly look at c_V (irrelevant constants are already subtracted).

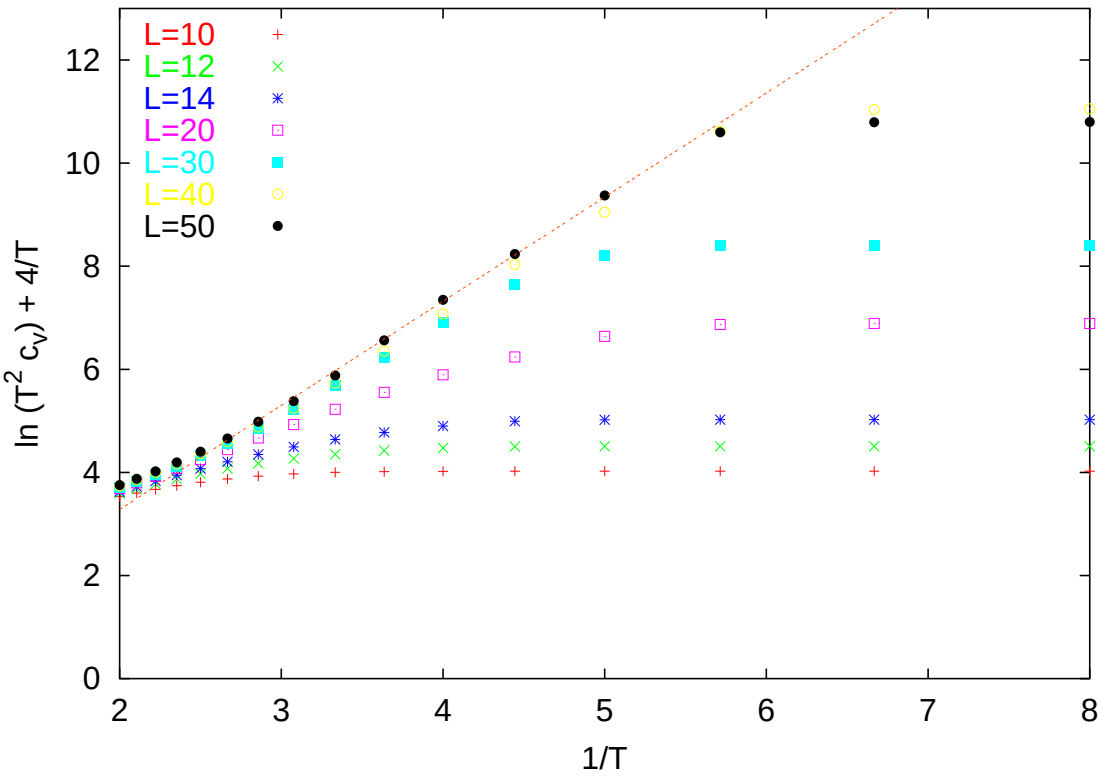
$$c_V \sim \beta^2 e^{-A\beta}$$

(we have checked that $p = 2$ is the best available choice for power corrections).

$$\log \frac{c_V}{\beta^2} \sim -A\beta$$

$$y \equiv \left(\log \frac{c_V}{\beta^2} + 4\beta \right) = (4 - A)\,\beta$$

So if we have naive scaling $y \sim$ constant in the scaling regime. If not: slope is $(4 - A)$.



Small T: saturation at constant value.

Intermediate T: $A \sim 2$.

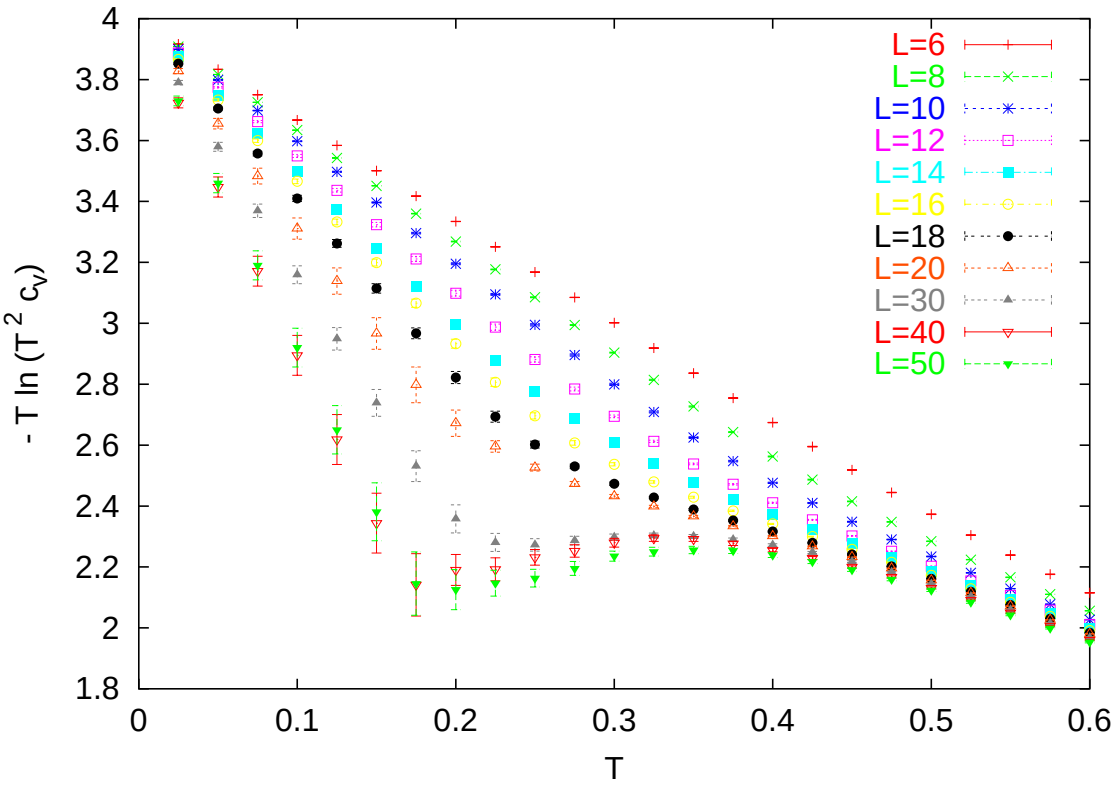
Straight line: best fit $\beta \in [2.5, 5.5]$ gives $A = 2.02 \pm 0.03$.

Clearly not 4, obviously decreasing, but asymptotic behavior not emerging.

$$-T \log \left(T^2 c_V \right) \sim A$$

So look at limit $T \longrightarrow 0$.

Very interesting scaling pattern.



Three regions:

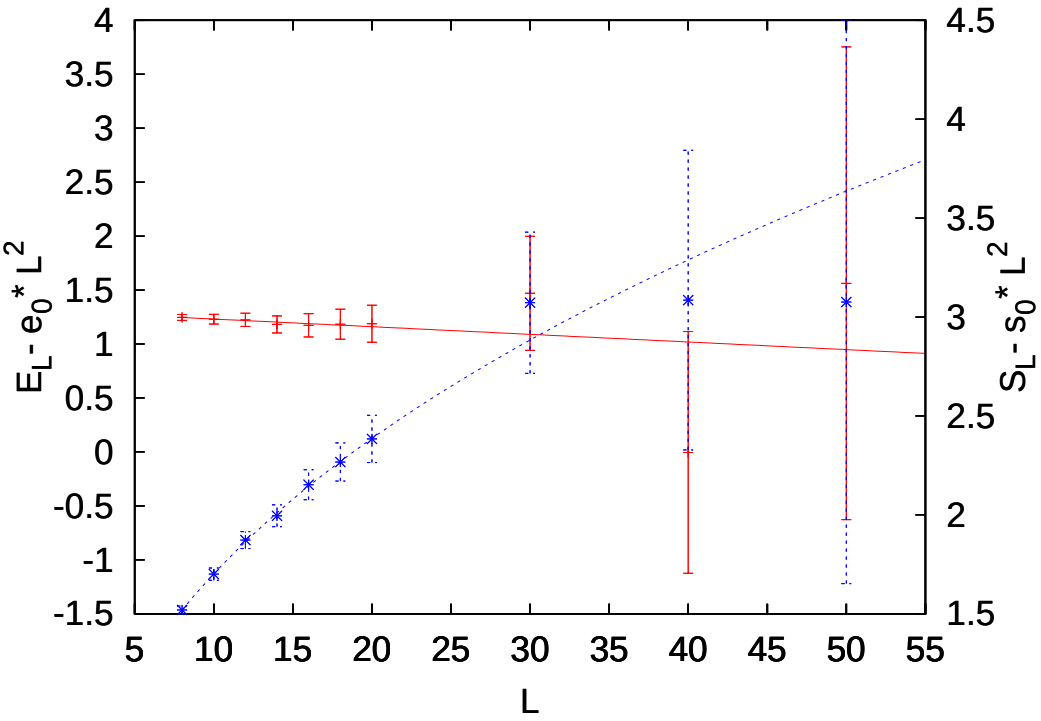
- high T "no scaling";
- low T $A = 4$ naive behavior;
- intermediate T, large lattices: $A \sim 2$, decreasing.

$T = 0$ properties.

Lines in the plot are best fits.

$$e_0(L) = e_0^* + aL^{-2+\theta^e}$$

$e_0^* = -1.4017(3)$, $\theta^e = -0.08(7)$. We see that as good evidence that $\theta^e = \theta_{DW} = 0$ (Hartmann-Young).



$$s_0(L) = s_0^* + aL^{-2+\theta^s}$$

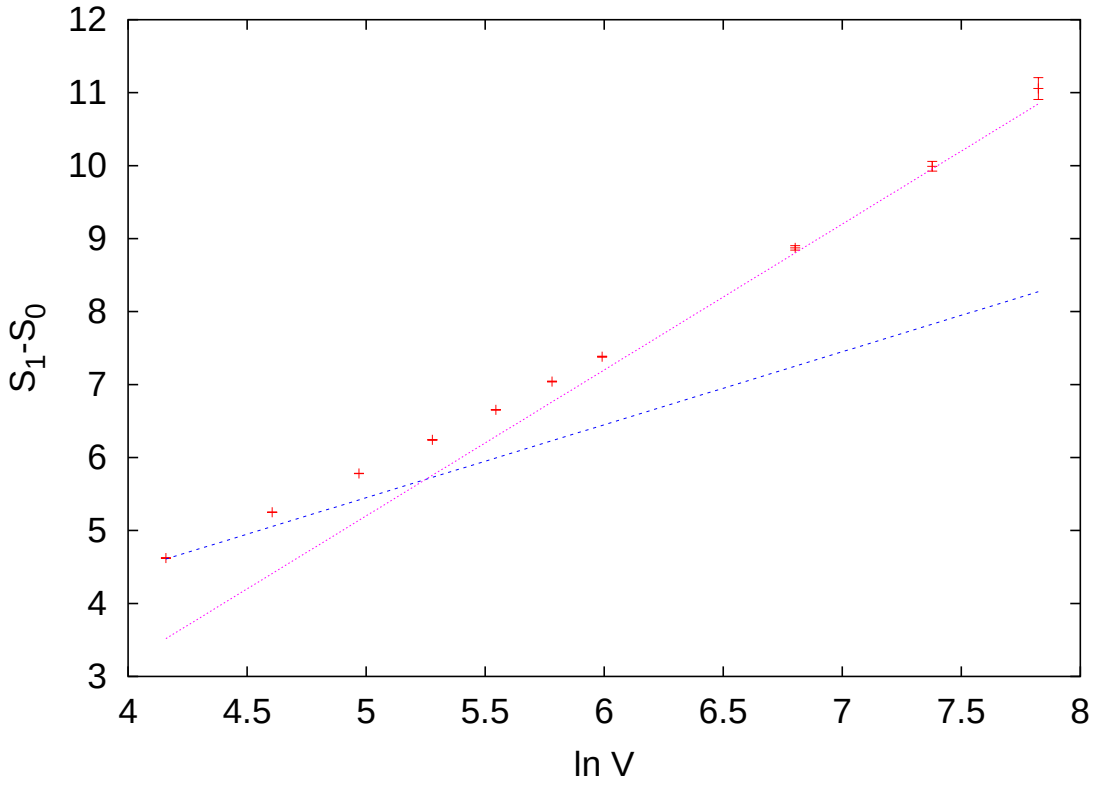
$s_0^* = 0.0714(2)$ (most precise estimate available), $\theta^s = 0.42(2)$. Could be that $\theta^s = 0.5$.

Anomalous density of excitations.

$$S_1 - S_0 = S(E_0 + 4J) - S(E_0)$$

Straight lines: $\log V$, $2 \log V$.

On large lattices: $2 \log V$ (Kardar-Saul could only see the transient behavior on smaller lattices).



In $1D$ Ising model equivalent of $4J$ excitations are "not elementary": here similar but more complex basic excitations?

Finite Size Scaling.

Difficult to fit from the numerical data the exact scaling law. We use two approaches.

1. For each L value we determine $T^*(L)$ as the temperature where "something happens" (where the data separate from the envelope).

   Scaling of such $\xi(T)$ obtained by inverting $T^*(L)$ prefers

   $$\xi \sim e^{\beta}$$

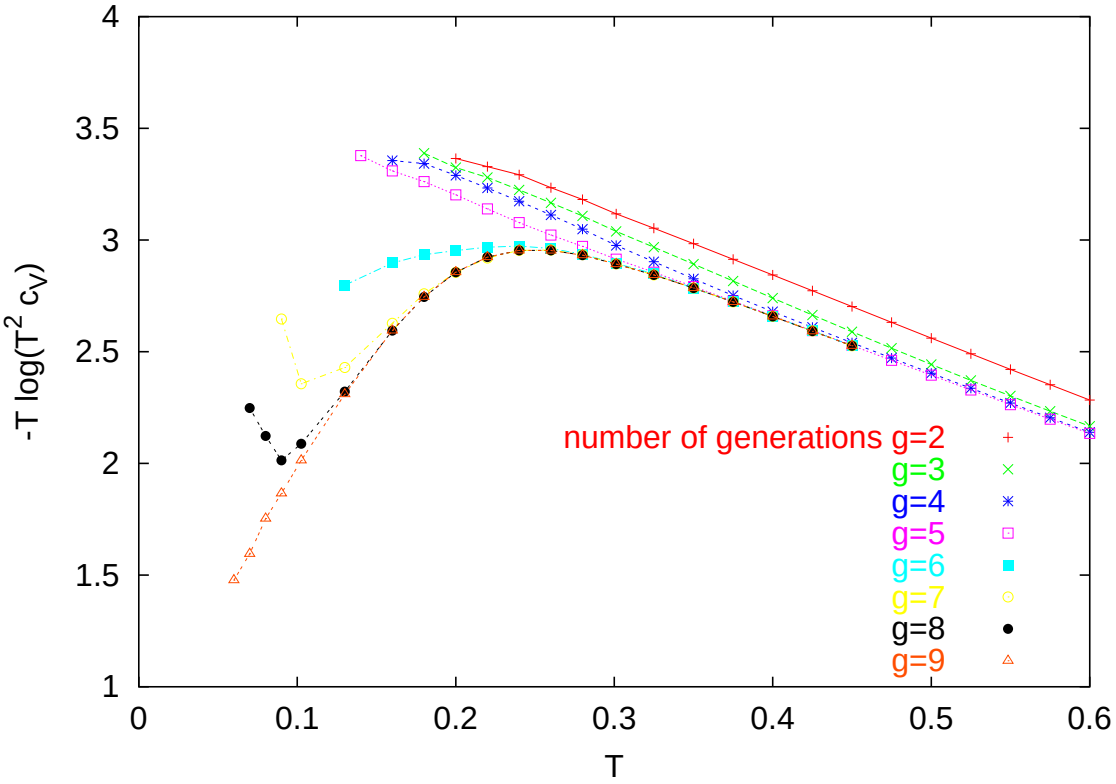
   far over $\xi \sim e^{2\beta}$, but again, not asymptotic. Behavior in the transient region is reminiscent of hyperscaling.

2. We can use a simple scaling argument based on the finding $S_1 - S_0 \sim 2 \log V$ to find the same behavior.

MKA approximation.

Very similar scaling pattern!

But lower A, maybe going to zero...? Here: MKA, $b = 3$ branches, $s = 3$ segments.



10^4 samples for 3 generations. 200 samples for 9 generations.

Here we know that $\theta = 0$ (Amoruso et al.).

Gaussian couplings: $c_V \sim T^{\alpha}$ as $T \longrightarrow 0$.

$J = \pm 1$: figure here. Very similar to $2D$ EA spin glass.

Analyze Villain fully frustrated 2D model.

For example all coupling equal to 1 but for even lines of vertical bonds equal to -1.

You have the analytic solution to look at. Here there is already a small mystery, i.e.

$$c_V \sim \beta^3 e^{-4\beta} \ .$$

Exact solution:

$$-\beta f_\infty(\beta) = \ln(2\cosh(\beta J)) + \frac{1}{16\pi^2}$$
$$\int_0^{2\pi} dh \int_0^{2\pi} dk \ \ln[(1+z^2)^2 - 2z^2(\cos 2h + \cos 2k)] \ .$$

Expand and find:

$$\beta f - \beta e_0 - s_0 \simeq c_1 \beta e^{-4\beta}$$

What happens in finite volume? Strip geometry computation (Mathematica, high precision).

Define ξ by

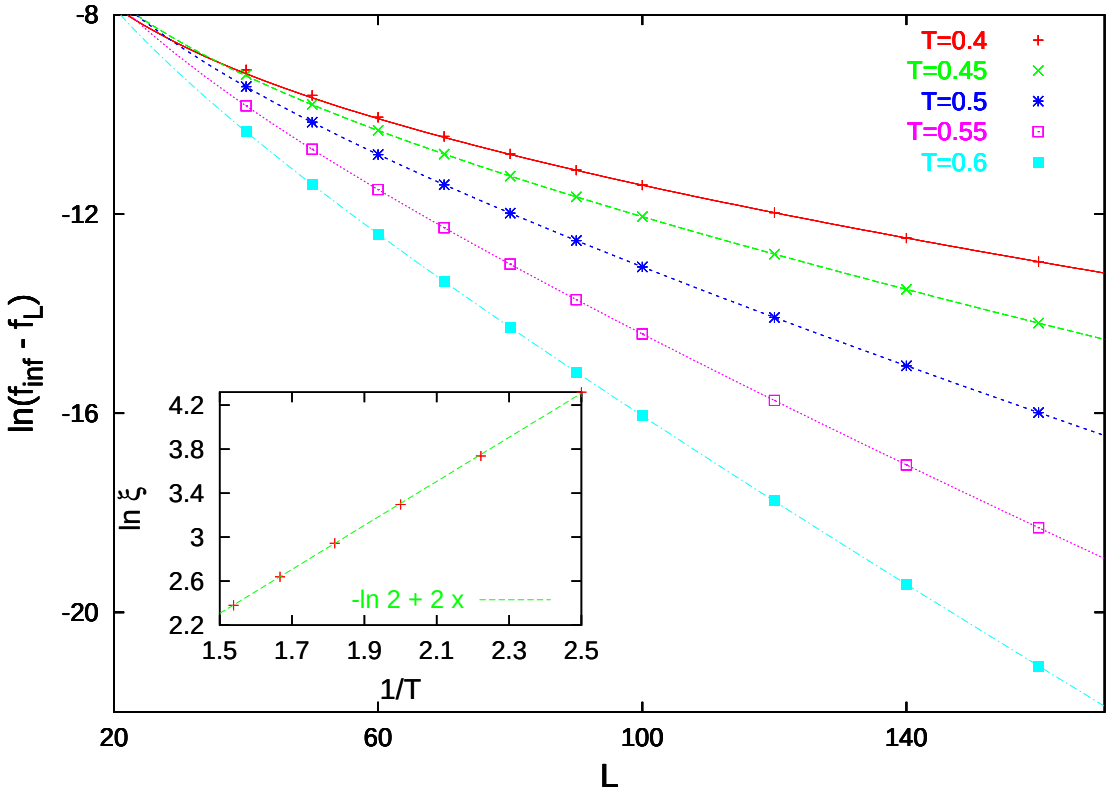
$$f_L - f_\infty \sim e^{-L/\xi} \ .$$

Figure: perfect best fits.

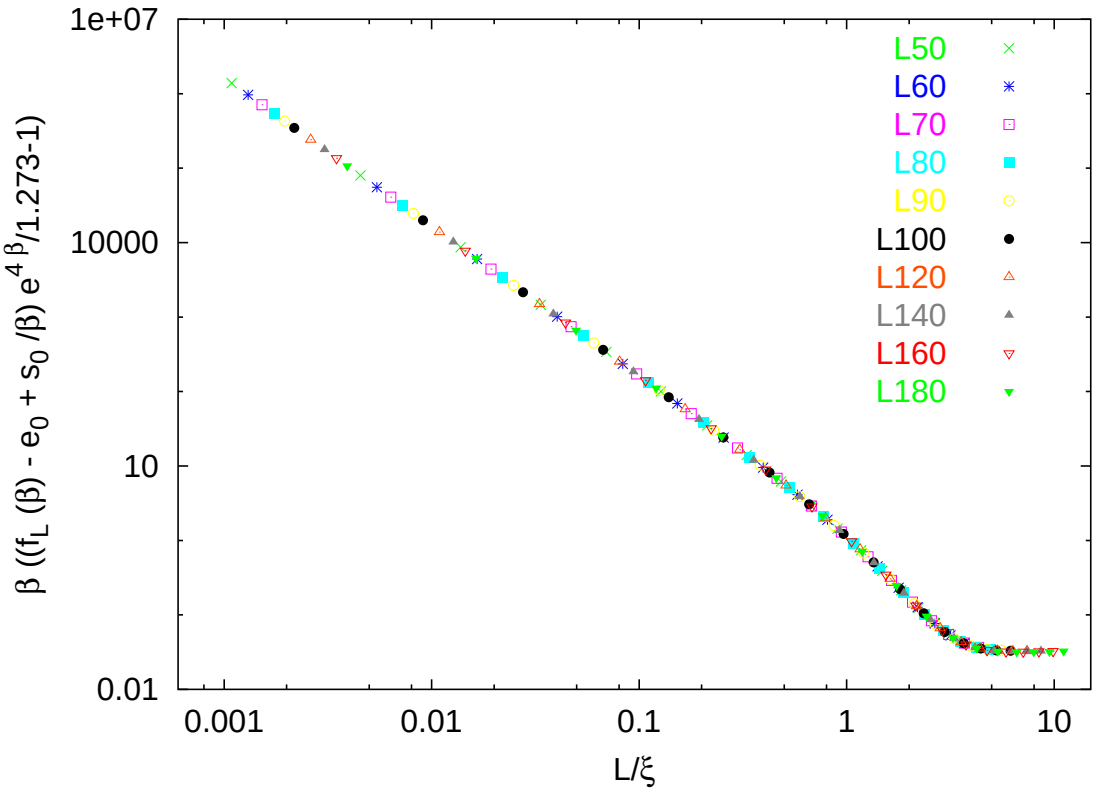
$$f_L - f_\infty = A(T)\ \exp\left(-\frac{L}{\xi(T)}\right)\ L^{-C(T)}$$

$A(T)$ smooth; $C = 1.5$, constant;

$\xi \sim e^{2\beta}$ (see Inset).



Scaling function (see y-axis):

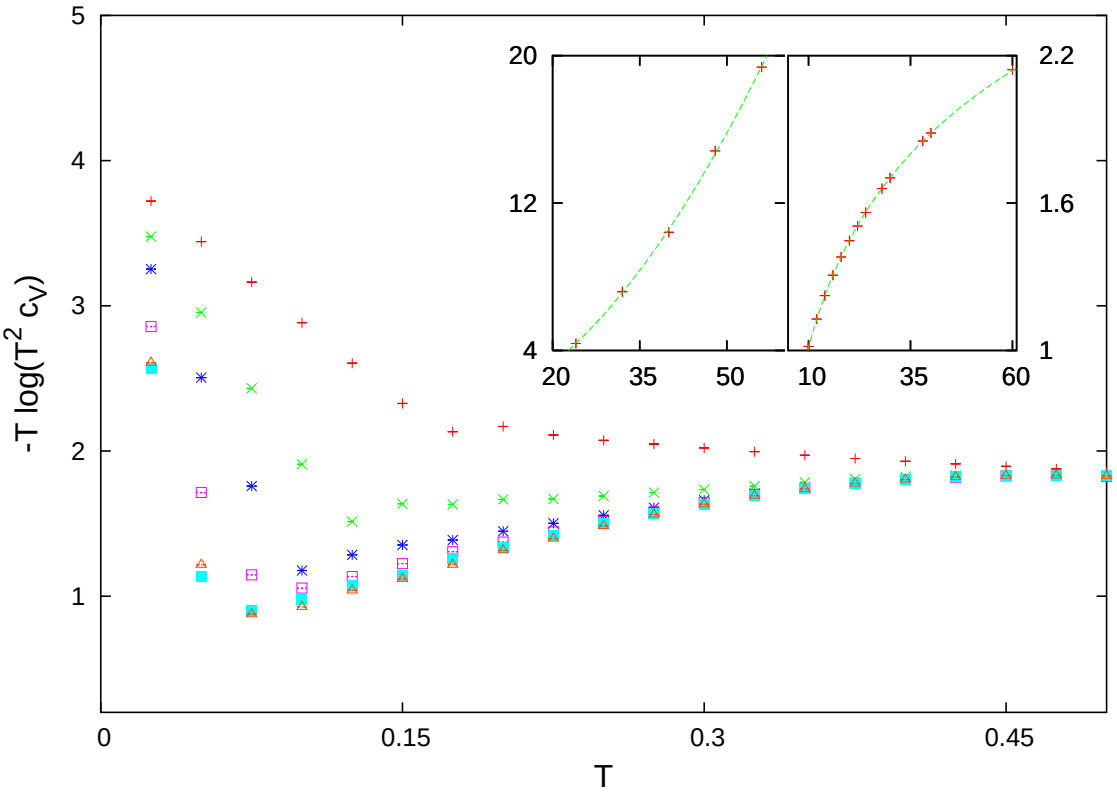


Count number of low energy states:

$$\frac{g_1}{g_0} = AL^2 + BL^2 \log(L)$$

that also implies this scaling law.

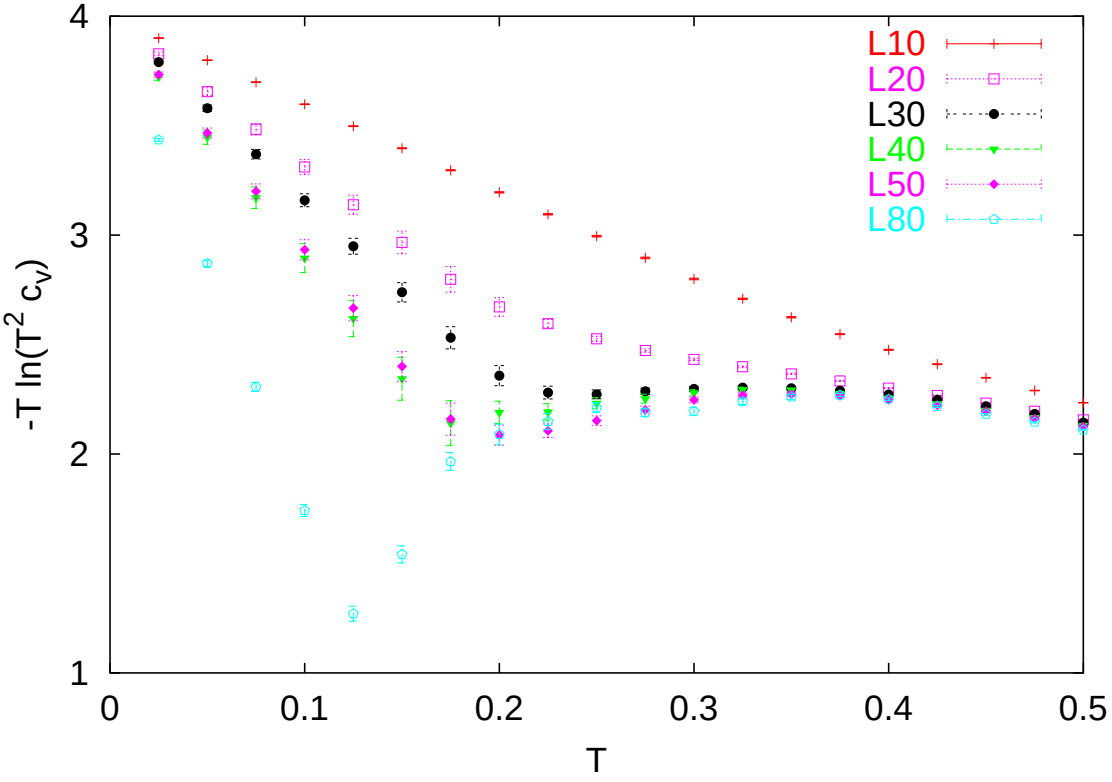
Quenched unfrustrated random plaquettes (here with density 1/8).



$-T\log(T^2 c_V)$ versus T in the PD model. In the inset, on the left: $\log(\frac{g_1}{g_0 L^2})$ versus L for the PD model. In the inset on the right: $\frac{g_1}{g_0 L^2}$ versus L for the FFM.

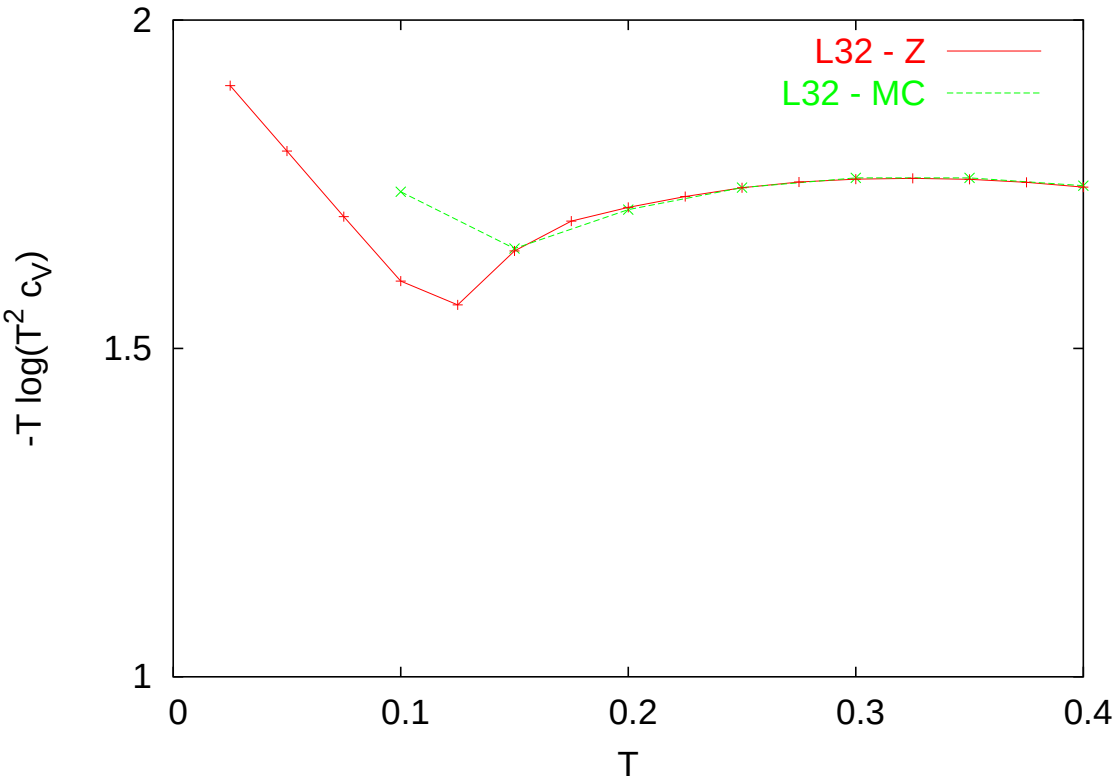
A=0 and algebraic scaling?

Now improve SG computation and go on larger lattices: beware of systematic errors and of underestimated error. (T. Jörg, J. Lukic, EM and O. Martin, in preparation)



A decreases... algebraic scaling?

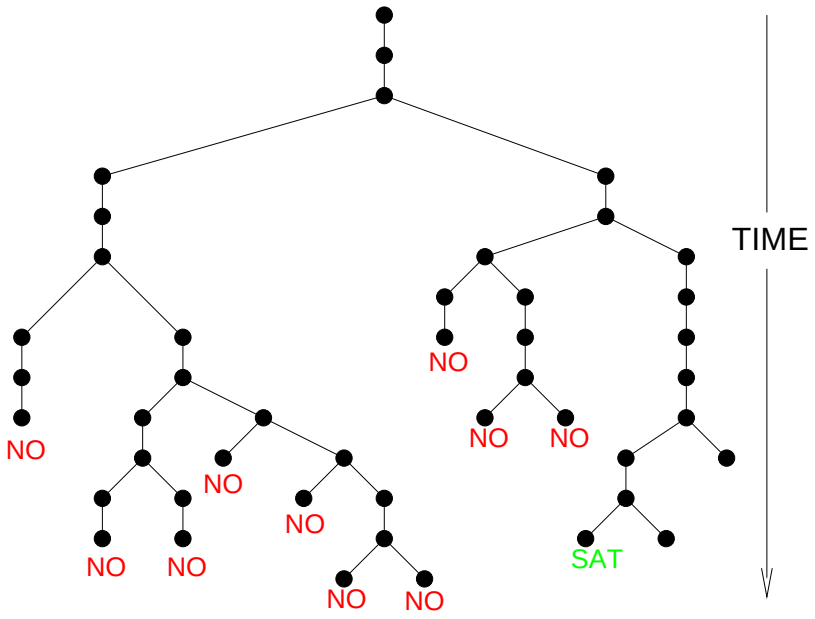
Also use diluted SG:



MC and exact Z fit very well.

Computational approaches

Checking for Sat or UNSAT:
Davis-Putnam-Loveland-Logemann algorithm



- Start with all x_i unassigned.
- BRANCHING: select one index i and branch ($x_i = 0$ and $x_i = 1$) and treat each sub-problem separately.
- PRUNING: if a node leads to one or more clauses being FALSE, label the node "NO" and backtrack. If no backtracking is possible, go to END.
- EXIT at a node if all clauses are TRUE (SAT sample).

END: If all leaves have been examined or pruned and are all labeled "NO", the sample is UNSAT.

Tool for solution of 3-SAT: belief propagation.

Can be improved dramatically, by understanding details of Replica Symmetry Breaking: survey propagation (by Mézard, Parisi and Zecchina, Science 2002).

Thousand-fold increase in the size of practical problems that can be solve (Mézard, Parisi and Zecchina, Mézard and Zecchina, PRE 2002, Brauenstein, Mézard and Zecchina, cs.cc/0212002).

Very interesting: use all what you know about the solution to derive an algorithm (see Swendsen and Wang with cluster algorithm based on properties of the Fortuin-Kasteleyn representation of the Ising model partition function).

Let $t_{i\to a}$ be the probability that variable x_i satisfies clause a in a solution.

Let $i_{a\to i}$ be the probability that clause a is satisfied by another variable than x_i.

Belief propagation gives in one direction the following iterative update equation:

$$i_{a\to i}^{(l)} = t_{j\to a}^{(l)} + t_{k\to a}^{(l)} - t_{j\to a}^{(l)} t_{k\to a}^{(l)} \ .$$

For the belief propagation equations in the other direction define

$$A_i^0 = \prod_{b\in V_i,\, y_{i,b}=\neg x_i} i_{b\to i} \ ,$$

$$A_i^1 = \prod_{b\in V_i,\, y_{i,b}=x_i} i_{b\to i} \ .$$

Update equations are

$$t_{i\to a}^{(l)} = \begin{cases} \dfrac{i_{a\to i}^{(l-1)} A_i^1}{i_{a\to i}^{(l-1)} A_i^1 + A_i^0} & \text{if } y_{i,b} = \neg x_i \\ \dfrac{i_{a\to i}^{(l-1)} A_i^0}{i_{a\to i}^{(l-1)} A_i^0 + A_i^1} & \text{if } y_{i,b} = x_i \end{cases}$$

Probabilistic interpretation:

$$\text{Prob}(x_i) = \frac{A_i^0}{A_i^0 + A_i^1}$$

$$\text{Prob}(\neg x_i) = \frac{A_i^1}{A_i^0 + A_i^1}$$

A BP-based decimation scheme consists in iteratively setting the variables with largest probability to be either true or false.

(This was probabilistic derivation as discussed and explained by Aurell, Gordon and Kirkpatrick in `cond-mat/0406217`).

To arrive at SP one introduces a modified system of beliefs, where every variable falls into one of three classes: TRUE in all solutions (1); FALSE in all solution (0); and TRUE in some and FALSE in other solutions (free).

Ultrametricity

I want to go back at last to ultrametricity, for sake of discussing a further technique (clustering).

Ultrametricity (UM) needs lot of space to emerge: it is really difficult to verify it on finite lattices.

Cacciuto, EM, Parisi 1996, Franz, Ricci-Tersenghi 1999.

Ultrametricity:

$$d_{13} \leq d_{12} + d_{23} \longrightarrow d_{13} \leq \max\left(d_{12}, d_{23}\right),$$

from triangular to (stronger) ultrametrical.

UM is an absolutely crucial feature of Parisi continuous RSB scheme. Consider two spin configurations α and β. Define a distance d from:

$$d^2_{\alpha,\beta} = \frac{1}{2}\left(1 - \frac{q_{\alpha,\beta}}{q_{EA}}\right)$$

equal to zero if $q = q_{EA}$, equal to 1 if $q = -q_{EA}$. Overlap:

$$q_{\alpha,\beta} = \frac{1}{V}\sum_{i=1}^{V} \sigma_i^\alpha \sigma_i^\beta \ .$$

Testing UM with Clustering

Clustering results of numerical simulations.
Domany, Hed, Hartmann, Stauffer 2001, Domany, Hed, Palassini, Young 2001.

We try to apply quantitative testing techniques Ciliberti, Marinari. We test MF: we know detecting UM is very difficult.

We find that the Z_2 symmetry has to be removed before any quantitative testing. This is very important: the ± 1 degeneracy completely obfuscates the results of the UM tests (see later).

Clustering (here for SK model with Gaussian couplings).

First you produce independent configurations (save $\forall$1000 full MC plus tempering sweeps) configurations at different (low) T values. N up to 512. T down to 0.2 (very low).

Set of configurations $\{C_t^{\tilde{T}}\}$. Compute overlaps at $T = \tilde{T}$ from $\sigma_{t'}^{\tilde{T}}(i)\sigma_{t''}^{\tilde{T}}(j)$, and since we are at equilibrium and configurations are uncorrelated this is a stationary sequence.

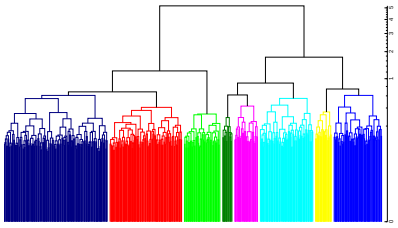
Clustering: partition data in "natural classes"':

- impose an ultrametric structure;
- check it it is natural.

Partition N objects into K clusters so that two points that belong to the same group are more similar than objects belonging to different groups.

Here we use, for the case $q \in (-1, +1)$ the definition $d \sim \frac{1-q}{2} \in (0, 1)$.

Hierarchical cluster algorithm $\longrightarrow$

dendogram.  There are many clustering algorithms one could use. Ward algorithm looks very suitable.

Fuse two clusters (individual objects are initial clusters).

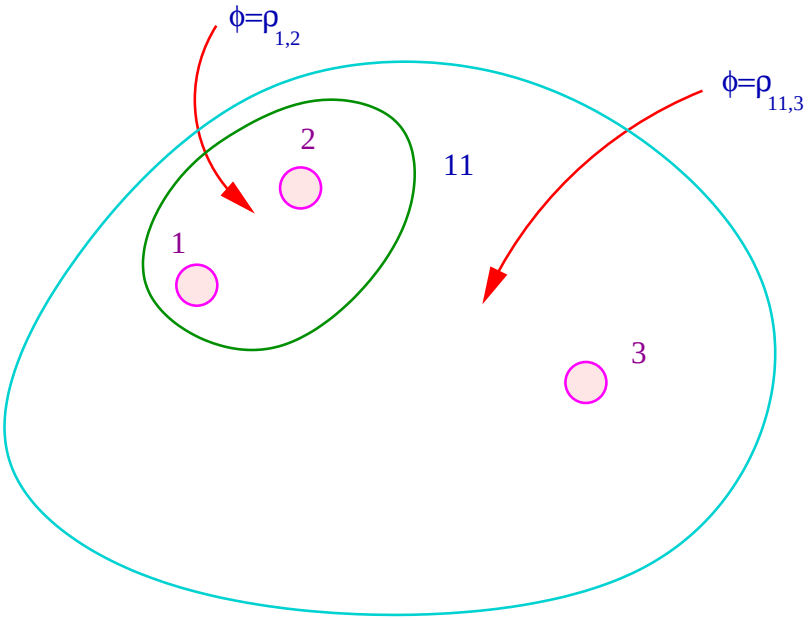
Initial partition: one object per cluster.
Compute $D_{\alpha,\beta}$ among all "clusters". Fuse the two closer clusters:

$$\gamma = \text{“}\alpha \cup \beta''$$

Now define effective distance from this cluster to other clusters. For the process $\alpha + \beta \longrightarrow \gamma$, let n_α be the number of objects in cluster α.
For all other cluster η we define
$\mathcal{N} \equiv n_\alpha + n_\beta + n_\eta$. and

$$d_{\gamma\eta} = \frac{n_\alpha + n_\eta}{\mathcal{N}} d_{\alpha\eta} + \frac{n_\beta + n_\eta}{\mathcal{N}} d_{\beta\eta} - \frac{n_\eta}{\mathcal{N}} d_{\alpha\beta}$$

$$\phi(\gamma) \equiv d_{\alpha,\beta} \quad (\phi_{\text{initial configuration}}\,(\alpha) = 0)$$



$$\delta(\gamma) = \sum_{\alpha,\beta \in \gamma} D^2_{\alpha,\beta}$$

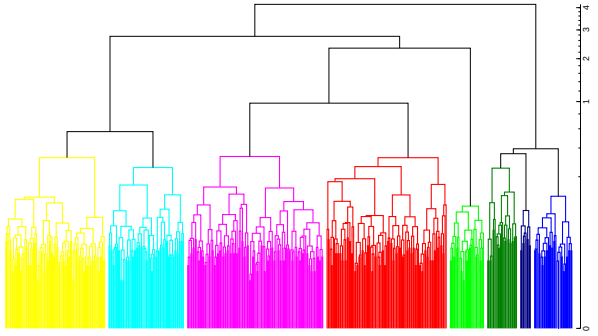
distance of all couples of configurations in a cluster. Clusters formed earlier have lower ϕ and δ.

Output of the procedure is a dendogram. Leaves are configurations. Ascending the tree you coarsen. UM is built in.

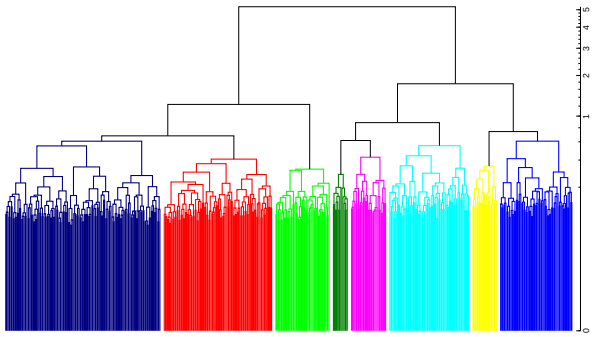
Testing: are we detecting a real UM? A valid clustering is equivalent to the presence of an ultrametric structure. So, we have to check validity of the clustering.

Visual observation does not help much... (but scale is different).

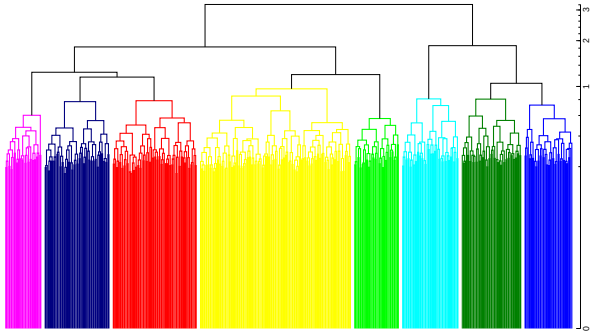
T=0.2



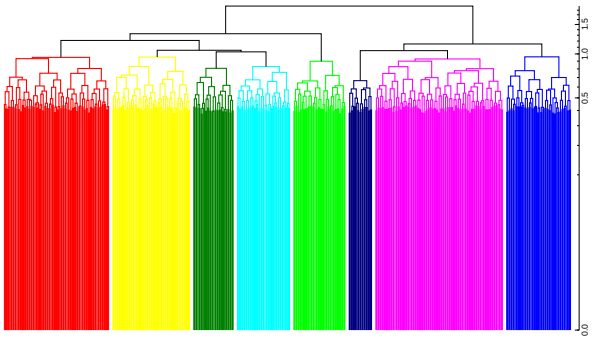
T=0.5



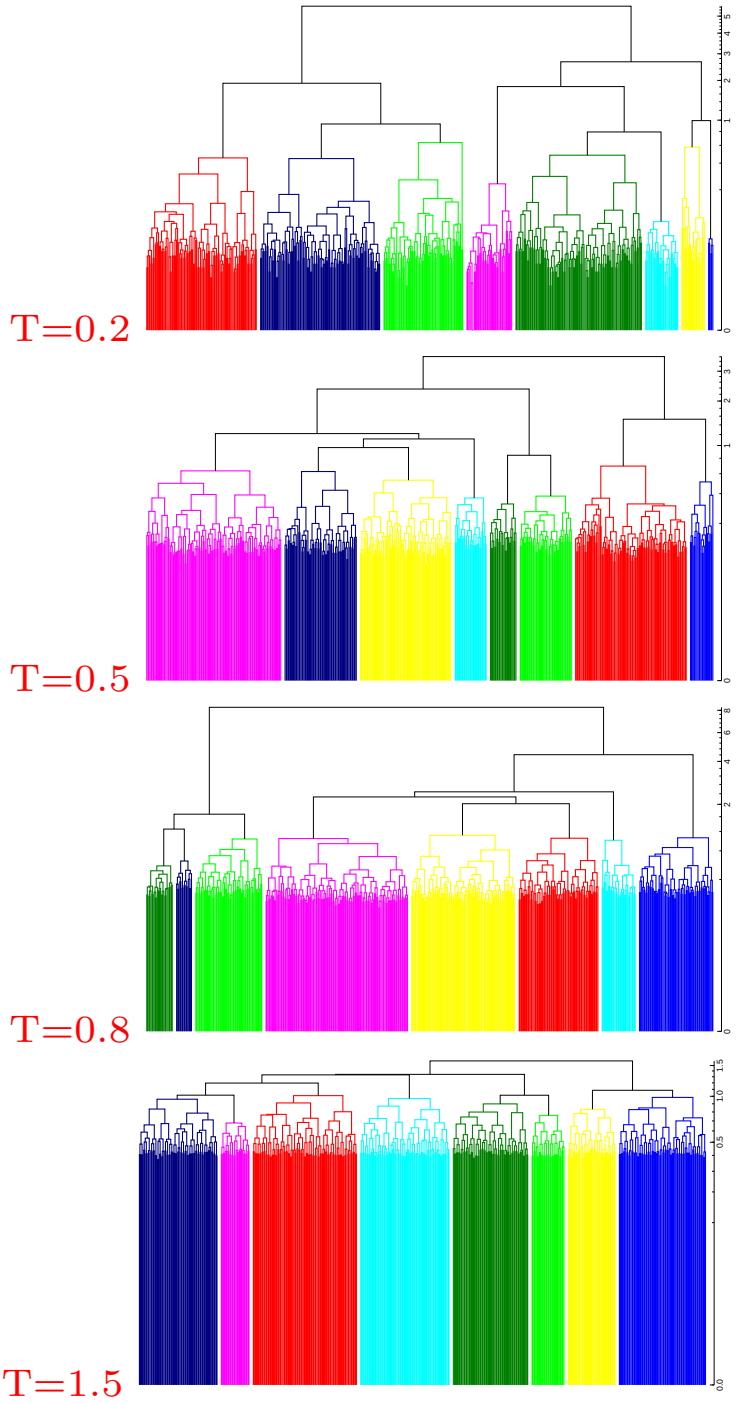
T=0.8



T=1.5



A different disorder sample.



We summarize:

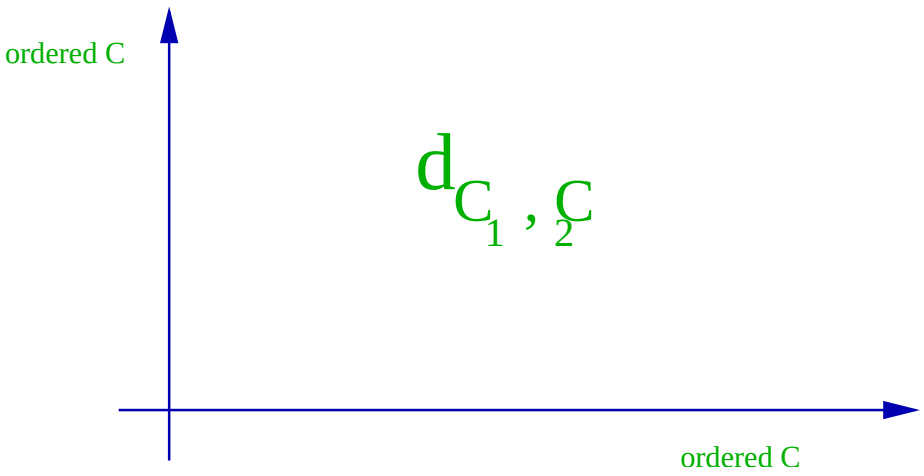
$$d_{\alpha,\beta} = \frac{1 - q_{\alpha,\beta}}{2}$$

by application of a cluster algorithm we obtain

a

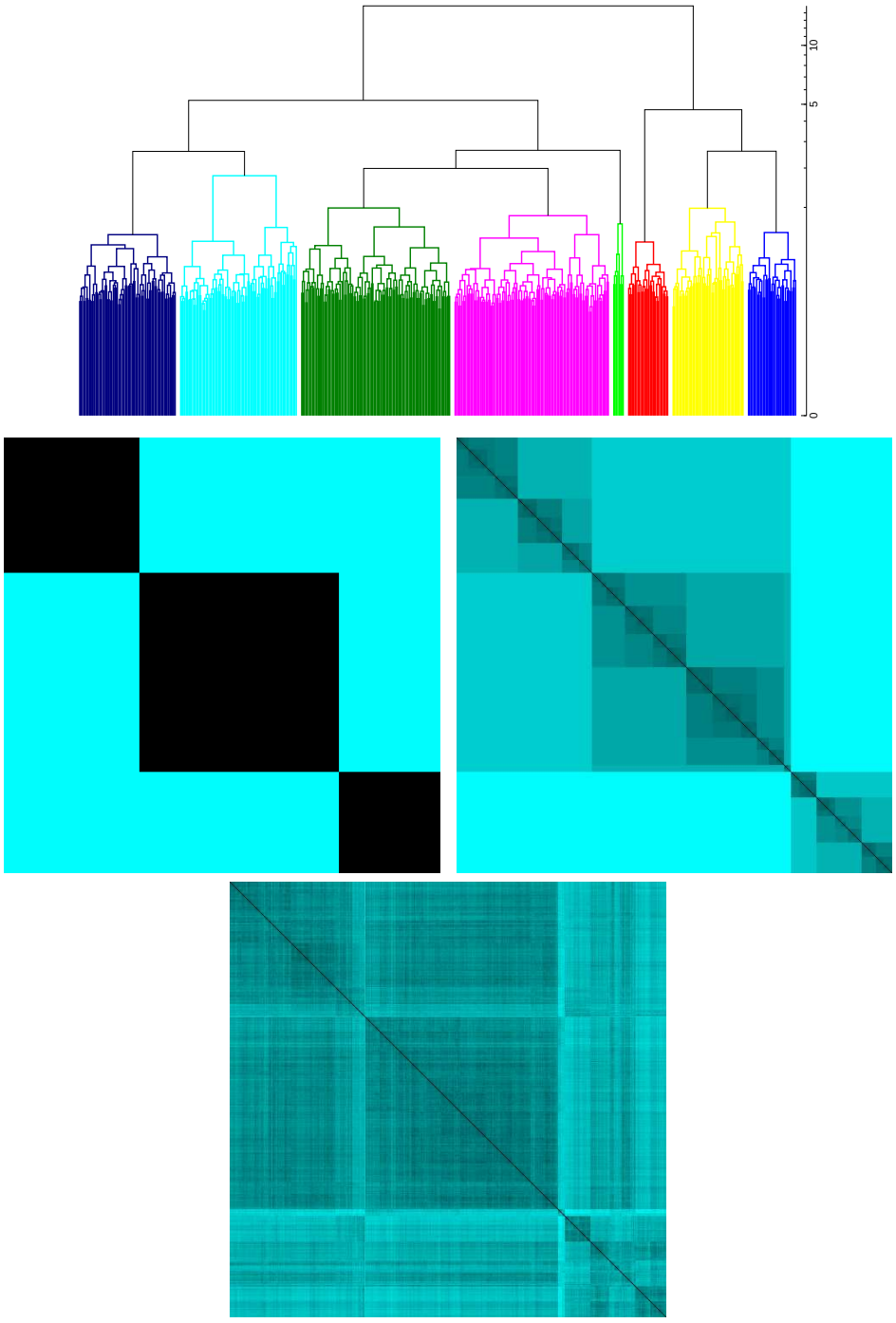
DENDOGRAM

i.e. an ordering of the configurations enriched by a (cophenetic) distance



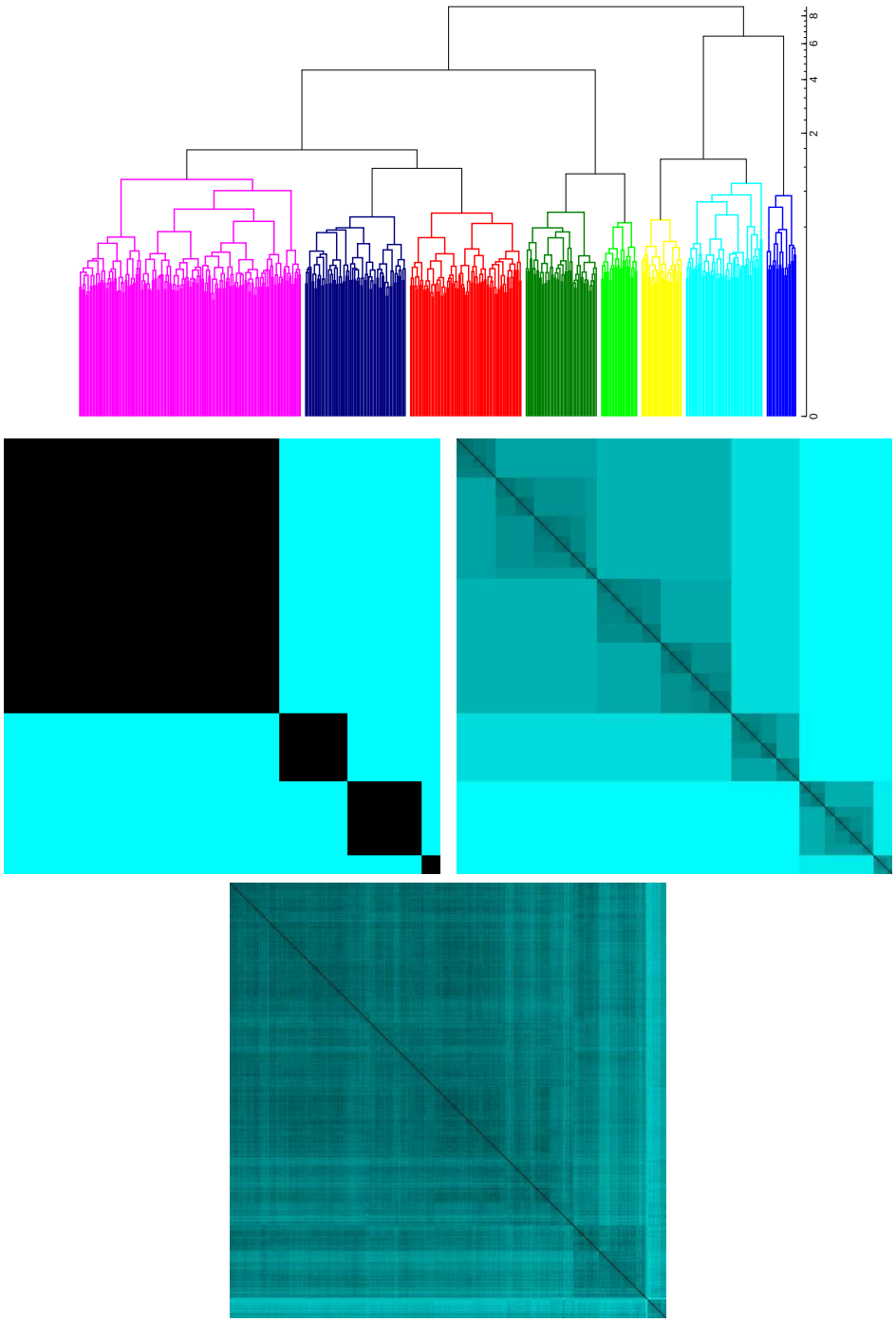
And now we can look at some figures.

SK, N=512, $T = T_c/2$



There is something good here.

SK, N=512, $T = T_c/2$



But much less here.

We discuss an important indicator, that tells us about the validity of the hierarchical structure.

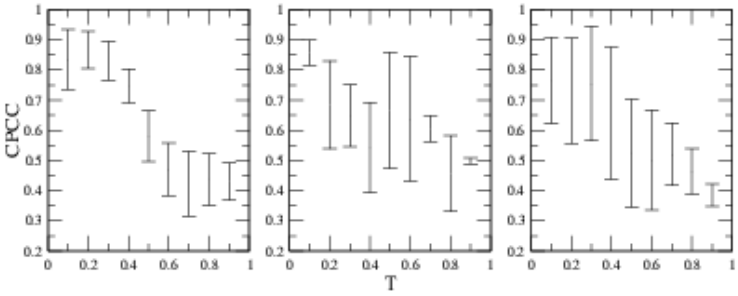
$d^C_{\alpha,\beta}$ cophenetic distance (UM by definition), i.e. the distance on the dendogram. $d_{\alpha,\beta}$ is the true distance.

$$\mathcal{K} \equiv \frac{\frac{1}{M^2}\sum_{\alpha,\beta} d^C_{\alpha,\beta} d_{\alpha,\beta} - \overline{d^C}\ \overline{d}}{\sigma_d \sigma_c}$$

It must be close to one to support the presence of a hierarchical structure (here there is not arbitrary threshold in the definition).

K is very used in numerical taxonomy. Empirically 0.9 is not enough (can establish accurately levels with MC as before).

Again, in presence of the Z_2 symmetry, K is very high and misleading. After removing it:



low, large errors, not very N dependent.

Detection on UM on "medium" size lattices is, even for MF models, very difficult or, better, impossible.