Dealing with the proliferation of metastable states in glasses and optimization problems

Marc Mézard, LPTMS, Université de Paris Sud, Orsay, France

16th September 2002

Collaborators:

- GLASSES: G. Biroli, B. Coluzzi, V. Martin-Mayor, G. Parisi, P. Verrocchio
- OPTIMIZATION: S. Franz, G. Parisi, F. Ricci-Tersenghi, M. Weigt, R. Zecchina

The 'many states' problem

1- Simple example: Random Energy Model (Derrida)

2^N energy levels E_i, i.i.d., $P(E) = \exp\left(-\frac{E^2}{N}\right)$,

Boltzmann probability $P_i \propto \exp(-\beta E_i)$.

Solution: Average number of states with $E/N = \epsilon$:

$$\mathcal{N}(E) \simeq 2^N P(E) \simeq \exp\left(N\left[\ln 2 - \epsilon^2\right]\right)$$

$\rightarrow$ No state at $\epsilon < \epsilon_c = -\sqrt{\ln 2}$.

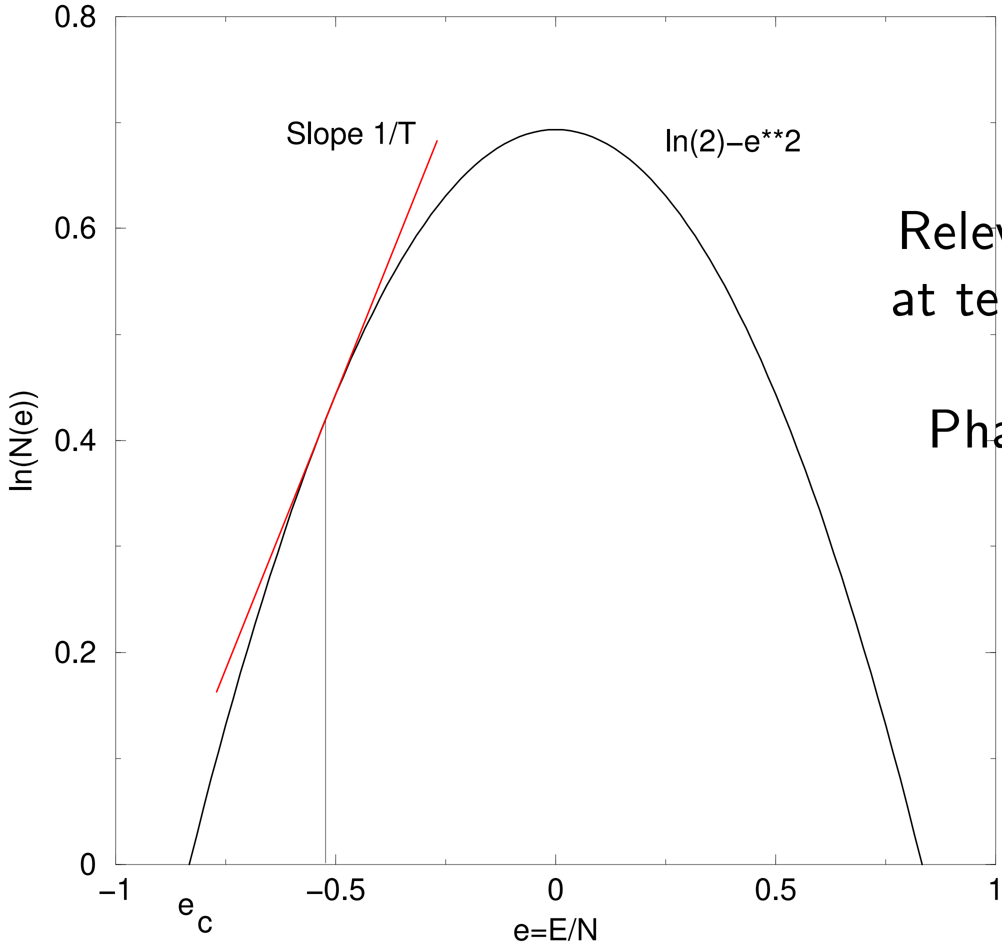
Boltzmann equilibrium:

$$Z = \int_{-\epsilon_c}^{\infty} d\epsilon\, \mathcal{N}(E)\ \exp(-\beta N \epsilon)$$

Phase transition when $\beta = \beta_c$.

High temperature: $< \sum_i P_i^2 >= 0$ (many configurations)

Low temperature: $< \sum_i P_i^2 > > 0$ (few configurations)

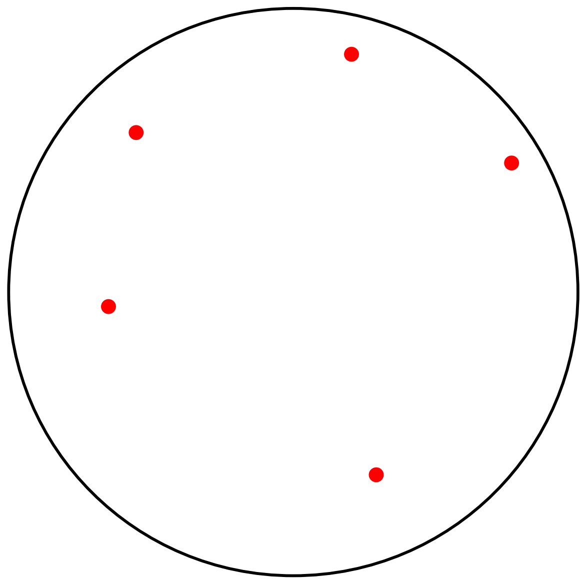


Relevant configurations at temperature $T > T_c$:

$$1/T = dS/dE$$

Phase transition at T_c

Phase space at $T < T_c$:

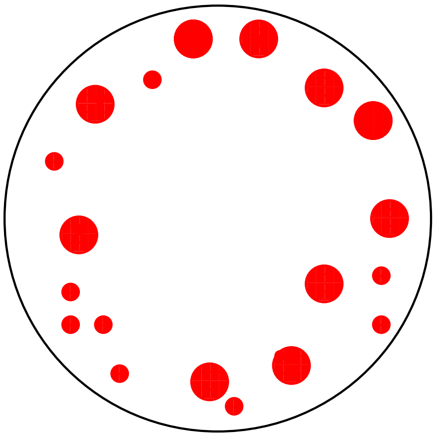


“Random first order” (or “one step RSB”) glass transitions

One well known scenario for a glass transition at the mean field level

Two transitions

- Dynamical transition T_d= ergodicity breaking

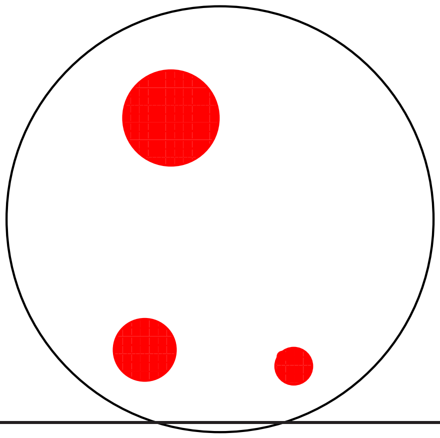


$$Z_\alpha = e^{-\beta F_\alpha} = \sum_{C \in \alpha} e^{-\beta F_c}$$

$$P_\alpha = Z_\alpha / \sum_\gamma Z_\gamma$$

$$\sum_\alpha P_\alpha^2 = 0$$

- Static transition $T_s < T_d$= weight condensation in few states

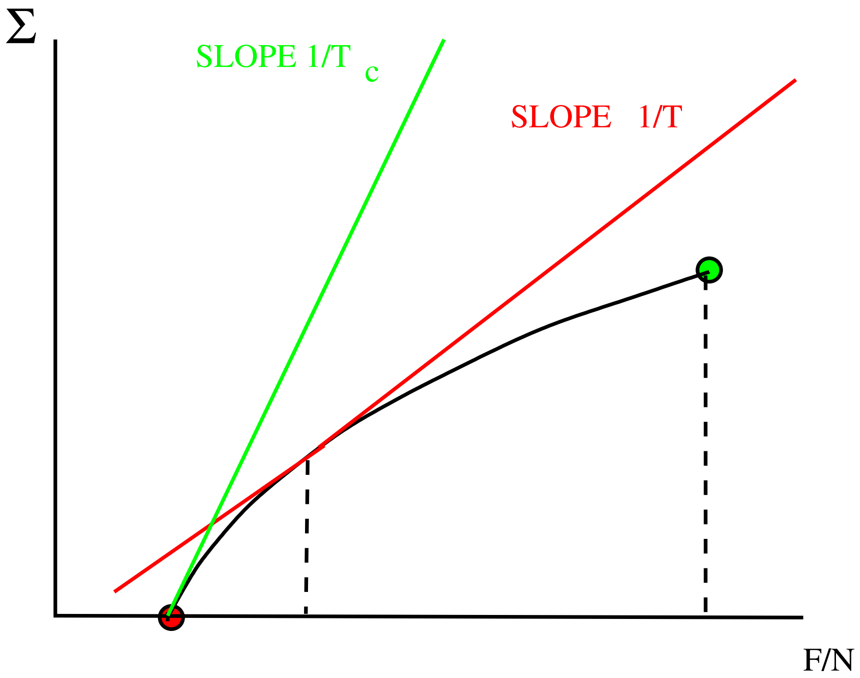


$$\sum_\alpha P_\alpha^2 > 0$$

Complexity

For $T_s < T < T_d$: Many states.

$$\mathcal{N}(F) \propto e^{N\Sigma_T(F/N)}$$



Boltzmann weights (hyp.):

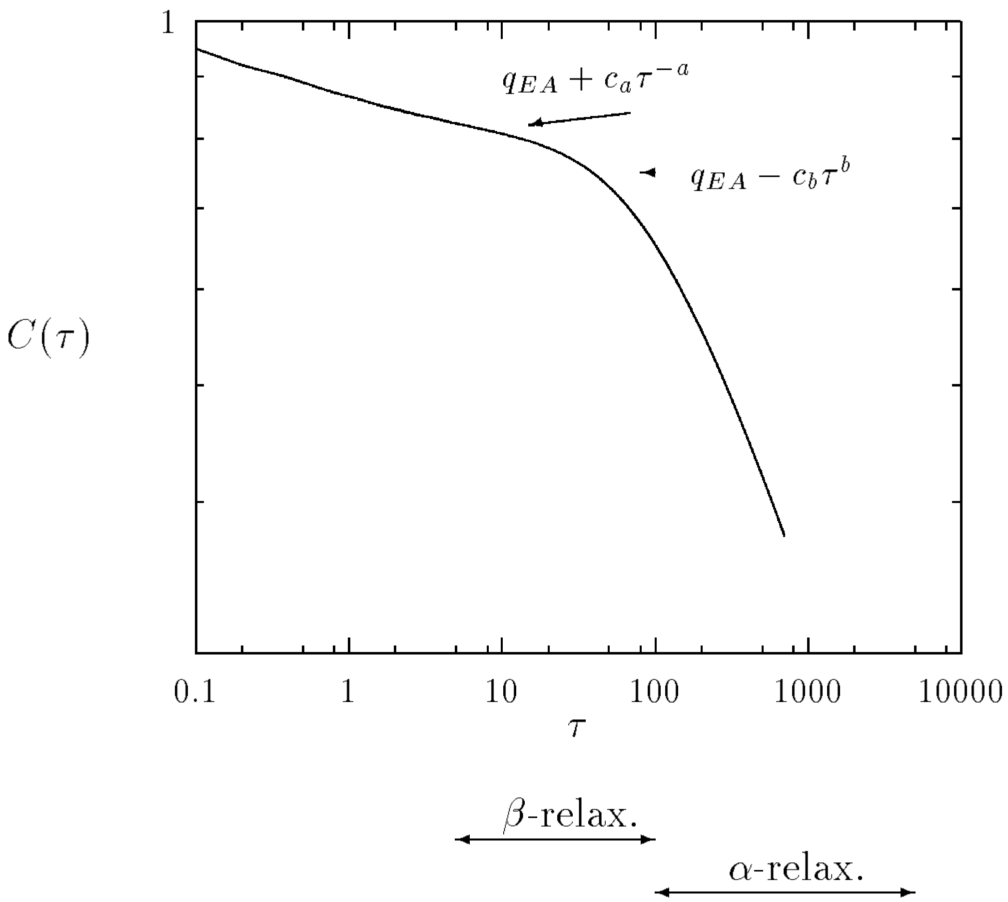
$$Z = \sum_\alpha \left(\sum_c e^{-\beta E_c} \right) = \sum_\alpha e^{-\beta F_\alpha} = \int df \, \exp\left(N\left[\Sigma_T(f) - \beta f\right]\right)$$

Dominated by s.p.: $d\Sigma_T/df = \beta$. $\rightarrow$ REM-like transition when $T < T_c$, where $1/T_c =$ slope at origin.

Dynamic transition

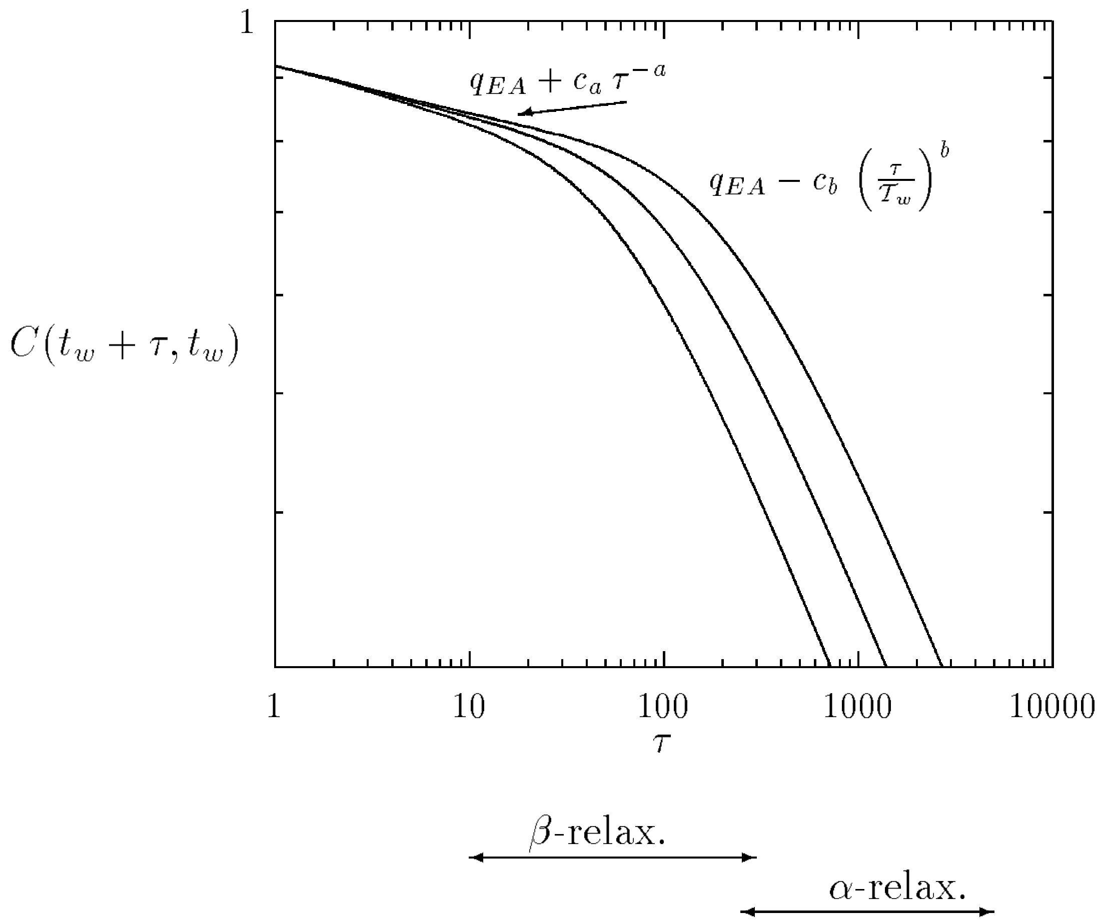
→ A major obstacle to equilibration of glasses and optimization problems.

$T \to T_d^+$: two steps relaxation:



"First order transition", but energy/entropy is continuous...

$T < T_d$: equilibrium not reached, AGING.



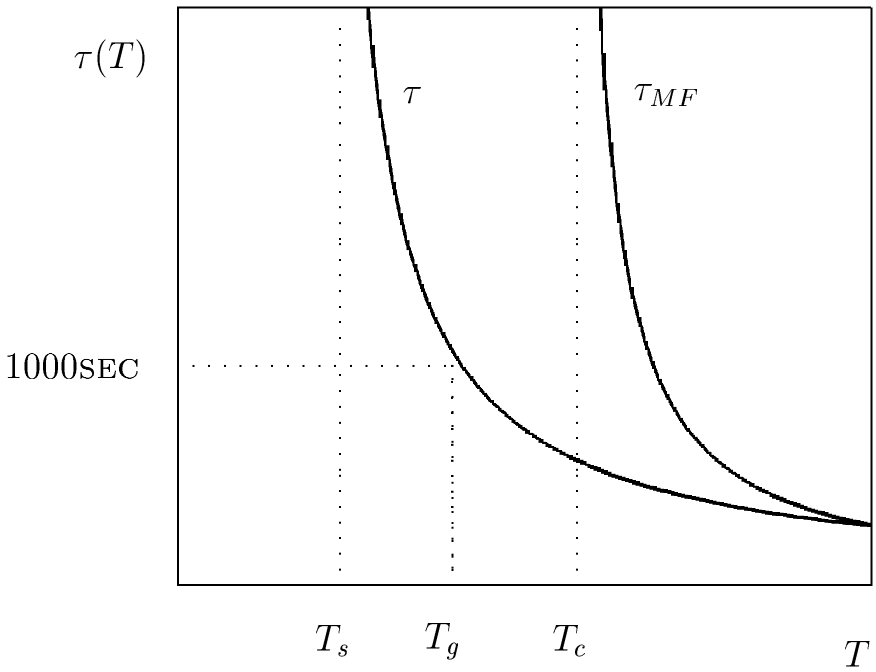
$\rightarrow$ FDT modification, effective temperature...

GLASSES

'Random first order' transition: a good candidate for a mean field description of structural glasses (KTW):

- First order (order parameter)
- Second order (U,S); jump in specific heat
- Two steps relaxation above T_d
- Aging at low temperatures, FDT modification with T_{eff}
- Entropy crisis at T_s.

Conjecture:



Dealing with many states

Two methods:

- Cloning ($\sim$ replicas, but without quenched disorder)
- Explicit assumption of many states in the cavity method

Cloning: m 'clones' of the original system constrained to stay in the same state (weak but extensive attraction).

$$Z_m = \sum_\alpha e^{-\beta m F_\alpha} = \int df \exp\left(N\left[\Sigma_T(f) - \beta m f\right]\right) \equiv e^{-\beta m \Phi(m)}$$

Saddle point $\rightarrow$ $\Phi(m)$ is the Laplace transform of $\Sigma(f)$ at a fixed given temperature.

Allows to compute low temperature properties of glasses by molecular liquid computations (bound states of m atoms).

Cavity method: Example of optimization problems.

Combinatorial optimization

- $\mathcal{S} = \{\text{Configuration} : \mathcal{C}\}$ described by N variables; $|\mathcal{S}| = e^{aN}$.

- Energy (cost) function $E(\mathcal{C})$ computable in $\sim N^b$ operations.

- Optimization Pb: Find $\mathcal{C} \in \mathcal{S}$ which minimizes $E(\mathcal{C})$.
  Decision Pb: Is there a $\mathcal{C} \in \mathcal{S}$ with $E(\mathcal{C}) < E_0$?

Boltzmann probabilities:

$$P(\mathcal{C}) = \frac{1}{Z} e^{-\beta E(\mathcal{C})}$$

Optimization = Find ground state ($\beta \to \infty$).

β finite: generalized problem, number of configuration at a given energy; entropy, free-energy, phase transitions etc...

Useful from the algorithmic point of view (simulated annealing+...)

Useful from an analytic point of view...

Computer Science: computational complexity

worst case analysis. Main Complexity Classes:

P $=$ polynomial$\leftrightarrow$ tractable, $t < N^{\alpha}$
NP $=$ non-deterministic polynomial (Given a 'yes' solution, it can be checked in polynomial time) $\leftrightarrow$ intractable $t \sim e^{aN}$.

P $=$ NP ?

NP-complete: the hardest **NP** problems: a problem is NPC iff all problems in **NP** are polynomially reducible to it. Theorem (Cook, 1971): The SATISFIABILITY problem is NP-complete.

From worst case to typical case: class of random instances

Two types of questions:

Algorithmic: one given sample, find optimal $\mathcal{C}$...

Theory: generic properties of typical sample...

Can the theory help the algorithms?

A core problem: SATISFIABILITY

"You are chief of protocol for the embassy ball. The crown prince instructs you either to invite Peru or to exclude Qatar. The queen asks you to invite either Qatar or Romania or both. The king, in a spiteful mood, wants to snub either Romania or Peru or both. Is there a guest list that will satisfy the whims of the entire royal family?" (from B.Hayes, American Scientist 1997).

Configurations= set of N Boolean variables

$$x_i \in \{0,1\} \quad i = 1,\ldots,N$$

M constraints =clauses like

$$x_1 \vee x_{27} \vee \bar{x}_3,\ \bar{x}_{11} \vee x_2,\ \ldots$$

Pb: is there a choice of the Boolean variables such that all constraints are satisfied (SAT)? (Generic: conjunctive normal form $(x_1 \vee x_{27} \vee \bar{x}_3) \wedge (\bar{x}_{11} \vee x_2) \wedge \ldots$).

NP complete if clauses of length ≥ 3, but:

For $\alpha = \frac{M}{N}$ small, often SAT; For α large, often UNSAT.

Typical complexity: random 3SAT

Three variables per clause, chosen randomly in $\{x_1, .., x_N\}$, negated randomly with probability $1/2$:

$$(x_1 \vee x_{27} \vee \bar{x}_3) \wedge (\bar{x}_{11} \vee x_3 \vee x_2) \wedge \ldots \wedge (x_9 \vee \bar{x}_8 \vee \bar{x}_{30})$$

Properties of typical sample?

Control parameter: $\alpha = \frac{M}{N} =$ Constraints/Variables.

Numerically: Threshold phenomenon at $\alpha_c \sim 4.26$:

- Generically SAT for $\alpha < \alpha_c$, generically UNSAT for $\alpha > \alpha_c$
- Computer time much larger around α_c.

Phase transition, complexity much bigger near to the phase transition.

Statistical physics of the random 3SAT problem

Monasson, Zecchina, Weigt, Biroli, ...

Similar to disordered spin glass on a random (hyper-)graph with finite connectivity. $x_i = \frac{1+\sigma_i}{2}$, with $\sigma_i \in \{-1, 1\}$.

$$(x_1 \vee x_{27} \vee \bar{x}_3) \wedge (\bar{x}_{11} \vee x_3 \vee x_2) \wedge \ldots \wedge (x_9 \vee \bar{x}_8 \vee \bar{x}_{30})$$

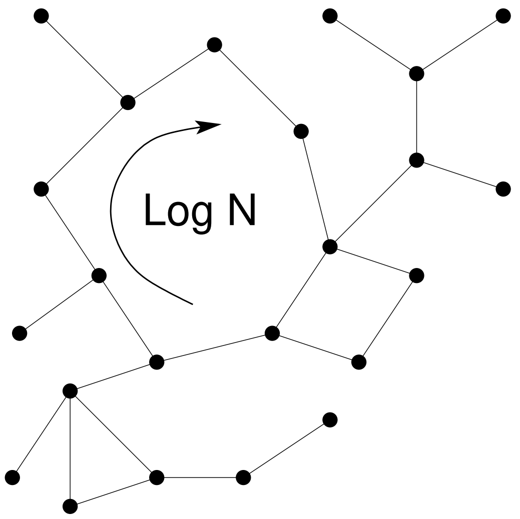
$$\Downarrow$$

$$E = \frac{1+\sigma_1}{2}\frac{1+\sigma_{27}}{2}\frac{1-\sigma_3}{2} + \frac{1-\sigma_{11}}{2}\frac{1+\sigma_3}{2}\frac{1+\sigma_2}{2} + \ldots$$

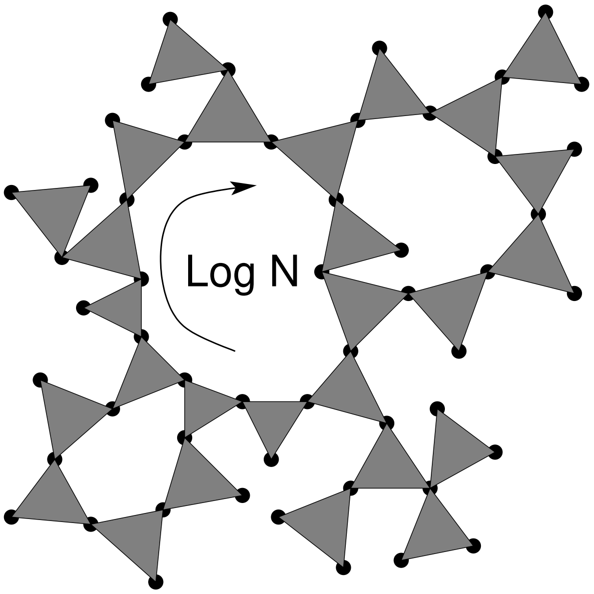
$$\ldots + \frac{1+\sigma_9}{2}\frac{1-\sigma_8}{2}\frac{1-\sigma_{30}}{2}$$

$\rightarrow$ 3-spin interactions on a random hypergraph

Geometry of random hypergraphs



Graph (2–spin interactions)



Hypergraph (3–spin interactions)

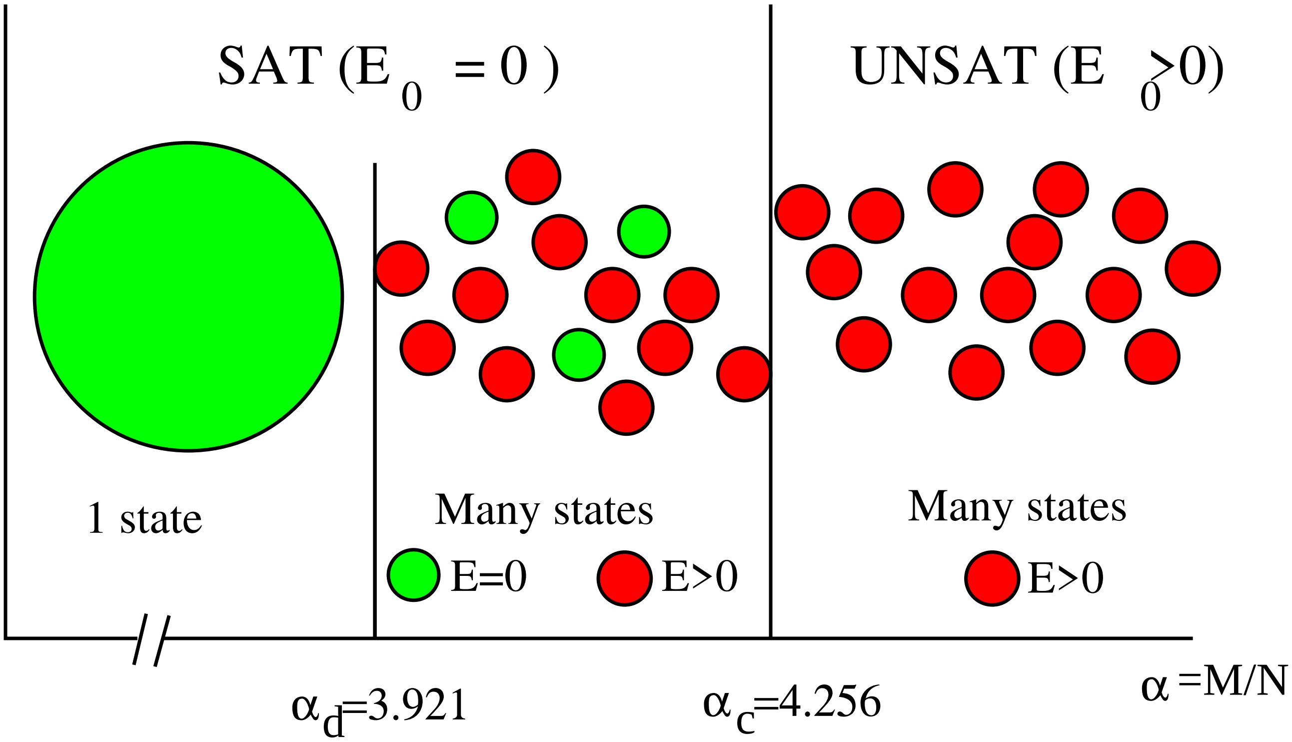
Locally tree-like, but loops of order $\log N$.

Proba(connectivity$=k$) $= \frac{(3\alpha)^k}{k!} e^{-3\alpha}$.

$\rightarrow$ Possible to use iterative methods (Bethe lattice) to solve the statistical physics problem.

RESULTS (T=0 study)

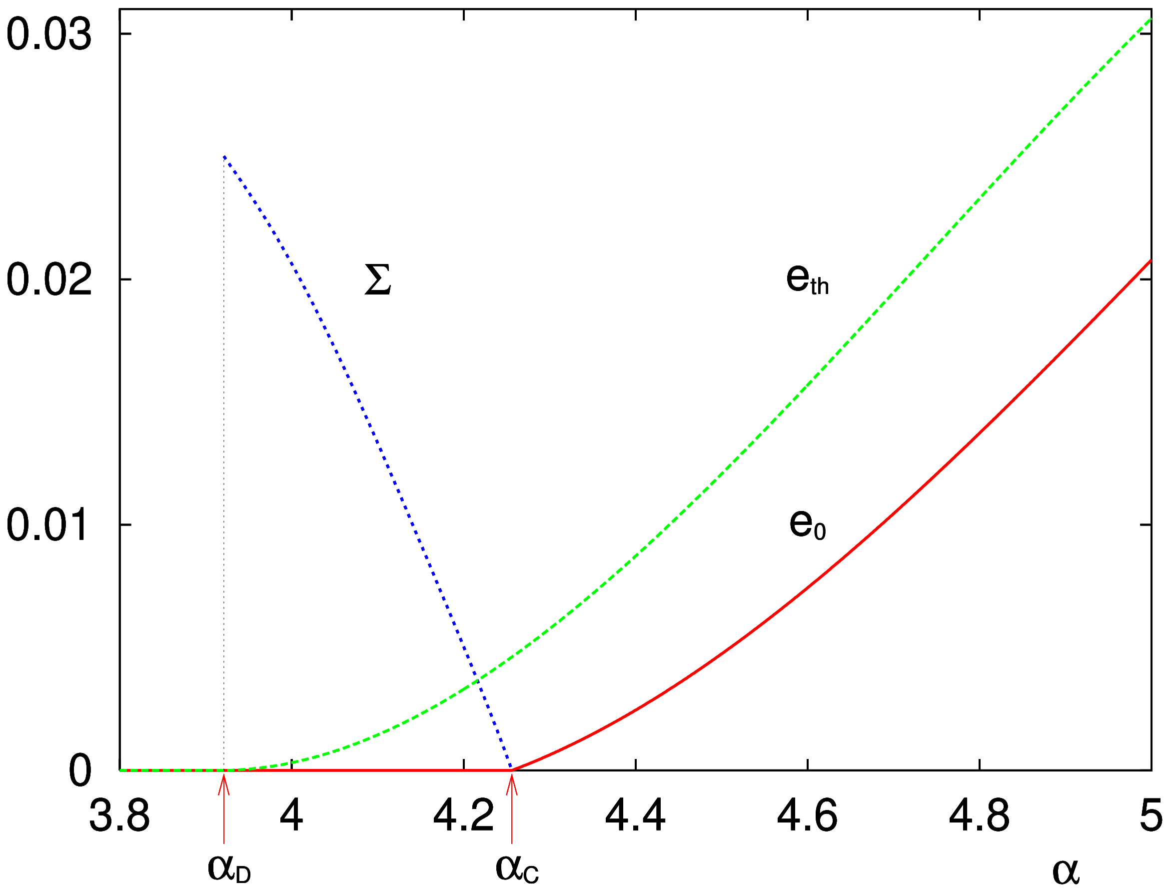
1- Analytic result: Three phases



- Easy SAT: Solution found by $T = 0$ Metropolis
- Hard SAT: Proliferation of local ground states, stable to finite number of spin flips $\rightarrow$ local algorithms (simulated annealing) trapped in excited states.
- UNSAT: Ground state has $E_0 > 0$, not found by local algorithms

2- A new class of algorithms dealing with the many states

THERMODYNAMICS



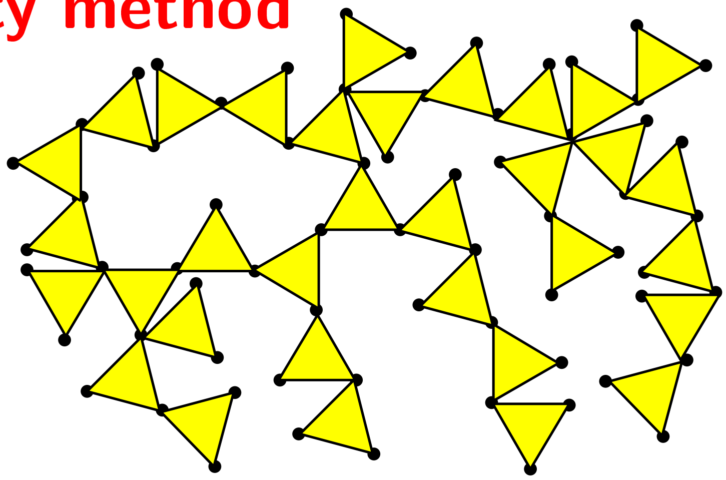
e_0 = Ground state energy density (minimal number of violated clauses per variable)

e_{th}= Energy density of the most numerous states

Σ = Complexity of states with $E = 0$: $\log(\mathcal{N}) \sim \exp(N\Sigma)$

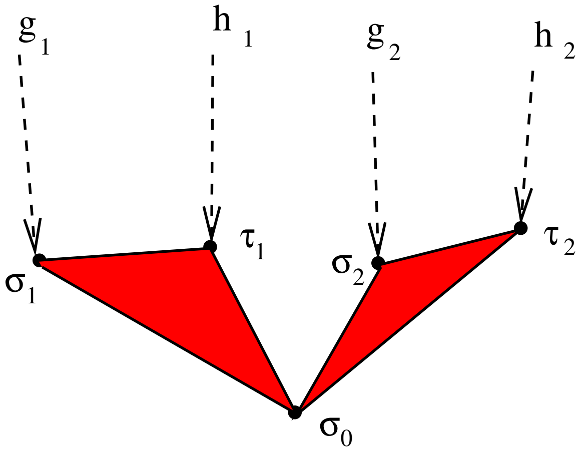
The cavity method

Initial set: N points,
M clauses

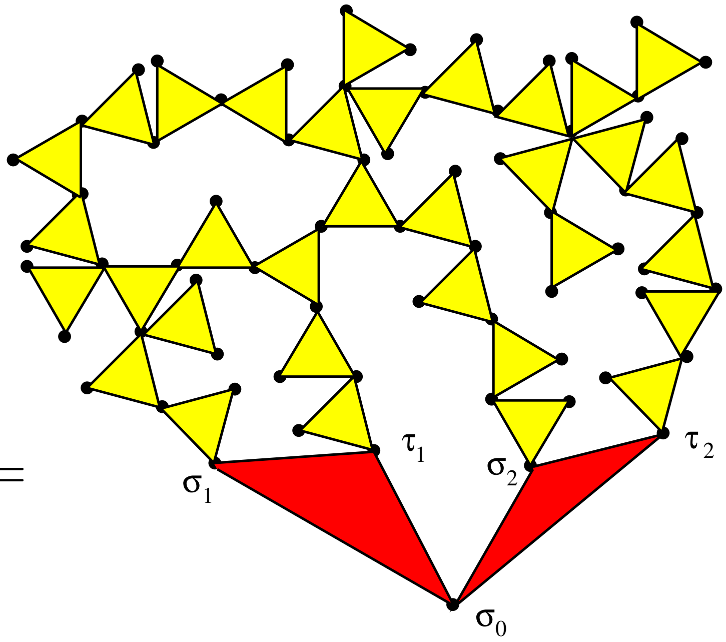


Add one spin:

$\Downarrow$



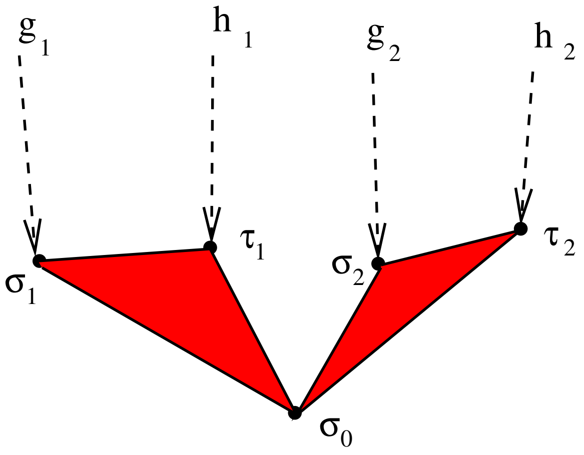
$=$



$$E_0^{(N)}(\sigma_1,\tau_1,\sigma_2,\tau_2) = A - g_1\sigma_1 - h_1\tau_1 - g_2\sigma_2 - h_2\tau_2$$

$$E_0^{(N+1)}(\sigma_0) = \min_{\sigma_1,\tau_1,\sigma_2,\tau_2}\Big(E_0^{(N)}(\sigma_1,\tau_1,\sigma_2,\tau_2) + \epsilon_1(\sigma_1,\tau_1,\sigma_0) + \epsilon_2(\sigma_2,\tau_2,\sigma_0)\Big)$$

One state



$$\rightarrow h_0 = f(g_1, h_1, g_2, h_2, ...)$$

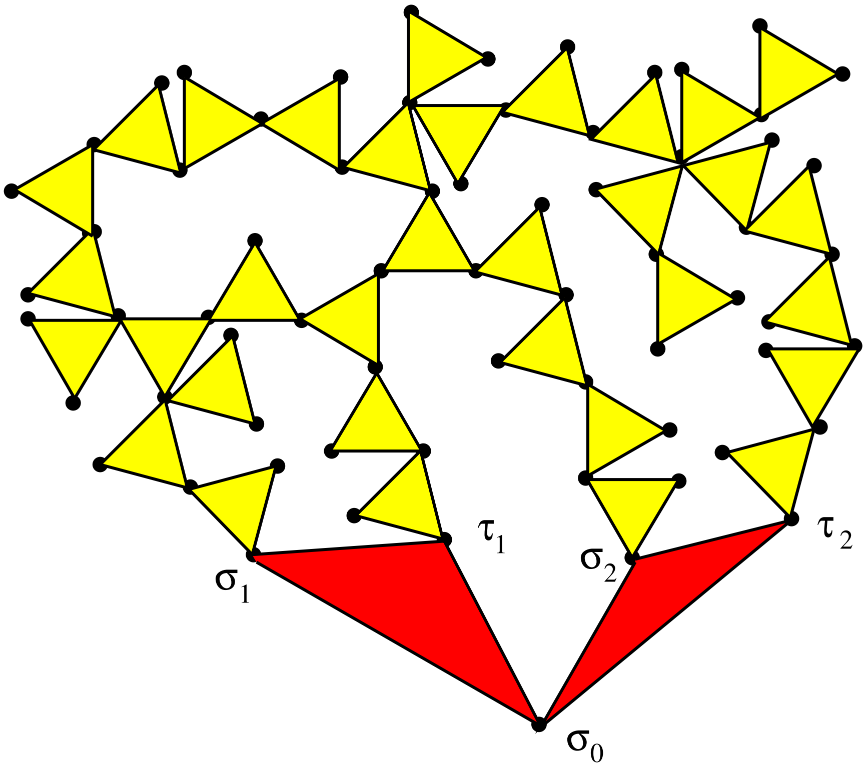
$$P(h) = \int P(g_1) dg_1 \; P(h_1) dh_1 \; ... \; \delta \left(h_0 - f(g_1, h_1, g_2, h_2, ...) \right)$$

$T = 0$: h integer, easily solved, but...

correct only for α small,
in the 'paramagnetic' region with $E_0 = 0$

in general: wrong.

Many states, clustering



If $\sigma_1, \tau_1, \sigma_2, \tau_2$ uncorrelated, then:

$$E_0^{(N)}(\sigma_1, \tau_1, \sigma_2, \tau_2) = A - g_1\sigma_1 - h_1\tau_1 - g_2\sigma_2 - h_2\tau_2$$

Far away spins (in absence of σ_0) $\rightarrow$ uncorrelated inside one state: $h_0^\alpha = f(g_1^\alpha, h_1^\alpha, g_2^\alpha, h_2^\alpha, ...)$ true for one given state.

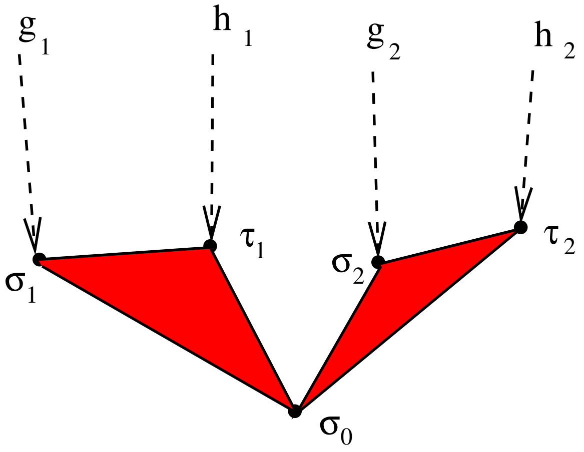
State proliferation; hypothesis:

$$\mathcal{N}(E) \sim \exp\left(\Sigma\left(\frac{E}{N}\right)\right)$$

Order parameter with many states

Order parameter = Survey of the local magnetic fields on the new spin σ_0 in the states of energy density e:

$$P_0^e(h) = C^t \sum_\alpha \delta\left(h_0^\alpha - h\right) \delta\left(\frac{E^\alpha}{N} - e\right)$$



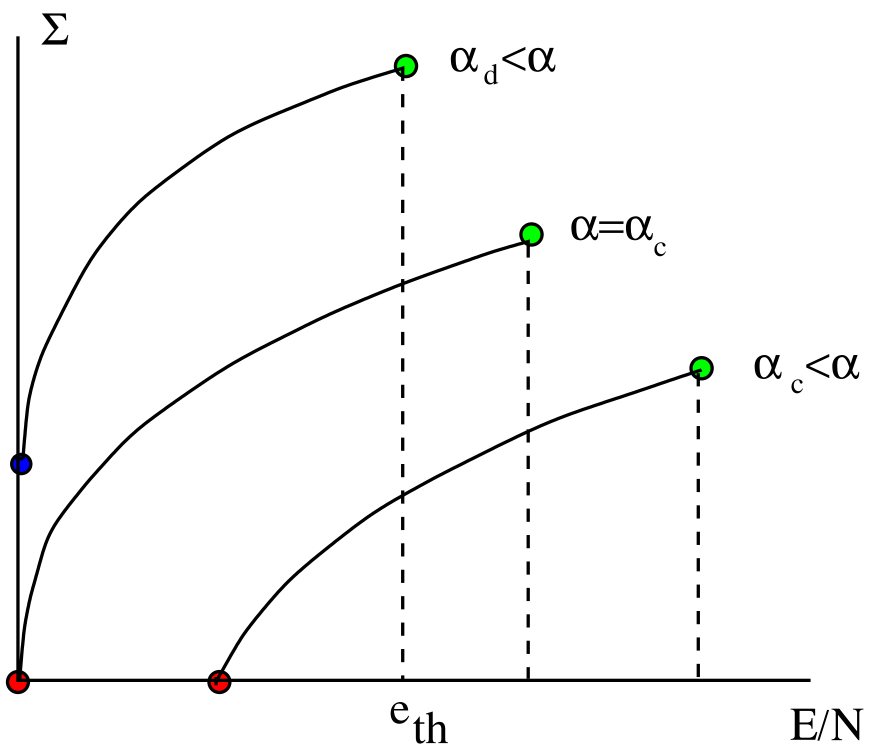
→ recursion relation between the surveys:

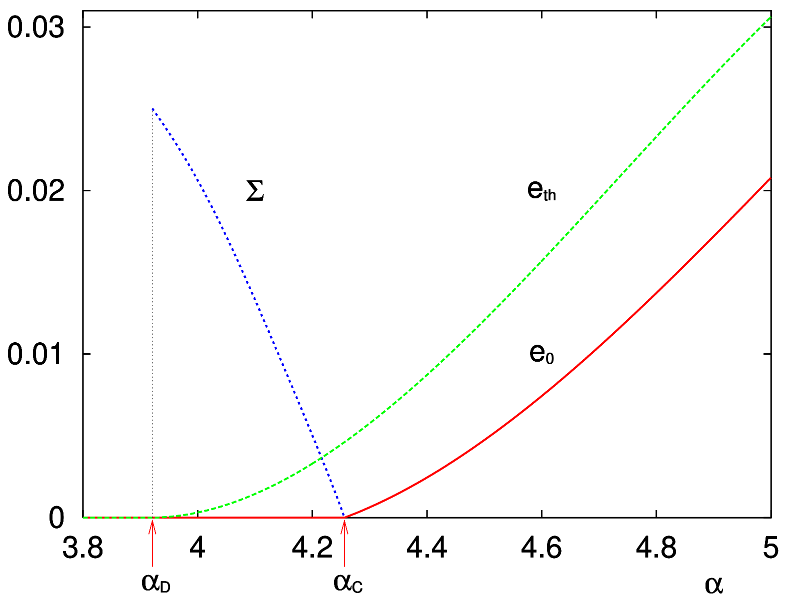
$$P_0^e(h) = C^t \int P_{\sigma_1}^e(g_1)dg_1 \; P_{\tau_1}^e(h_1)dh_1 \; ...\delta(h - f(g_1, h_1, g_2, h_2, ...))$$
$$\exp\left(\frac{d\Sigma}{de} w(g_1, h_1, g_2, h_2, ...)\right)$$

Distributed on integers: possible to solve the equation and compute the 'complexity curve' $\Sigma(E)$.

Complexity

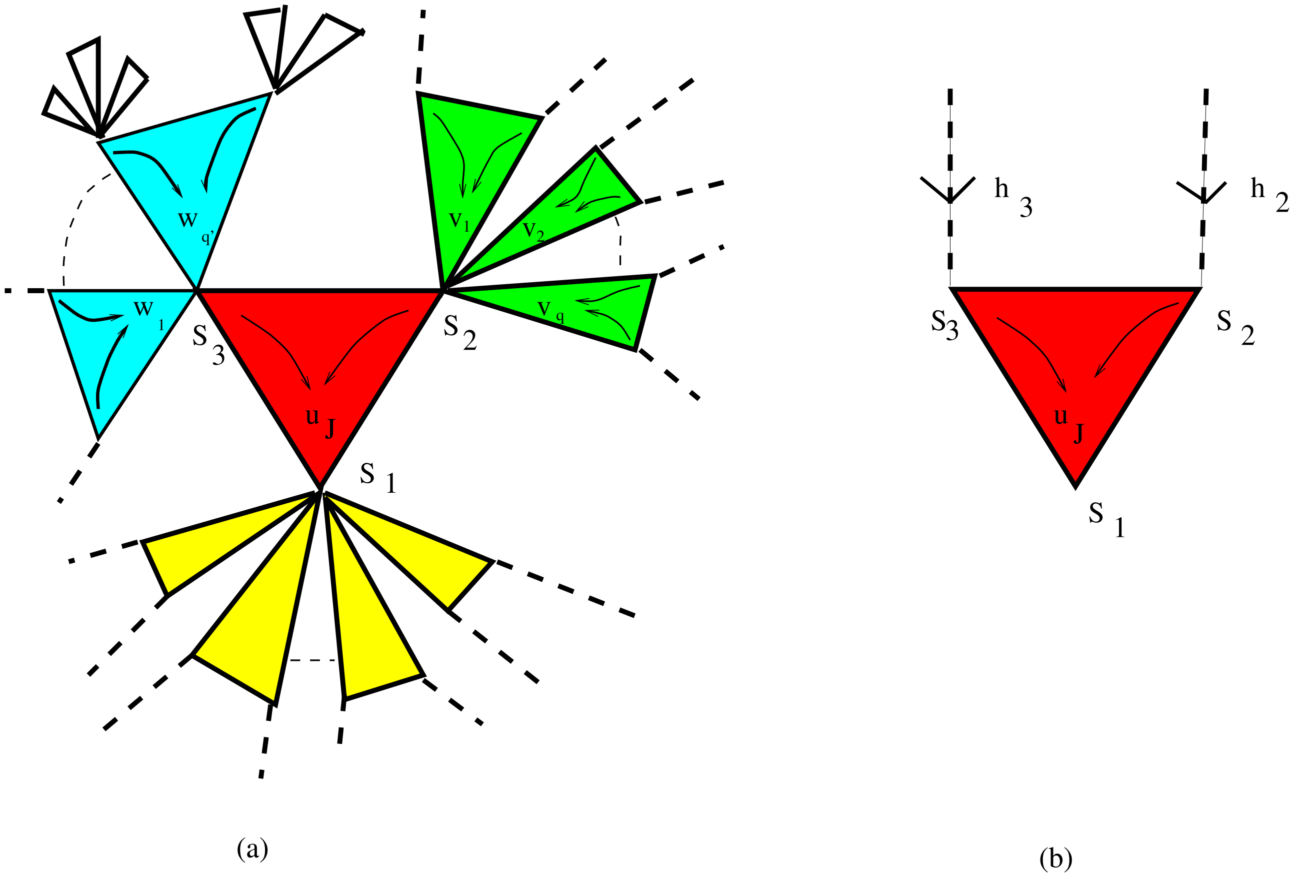
Qualitative behaviour of the complexity as the function of the number of constraints per variable:





Single sample analysis

Order parameter= Survey of local magnetic fields. Propagated along the graph.

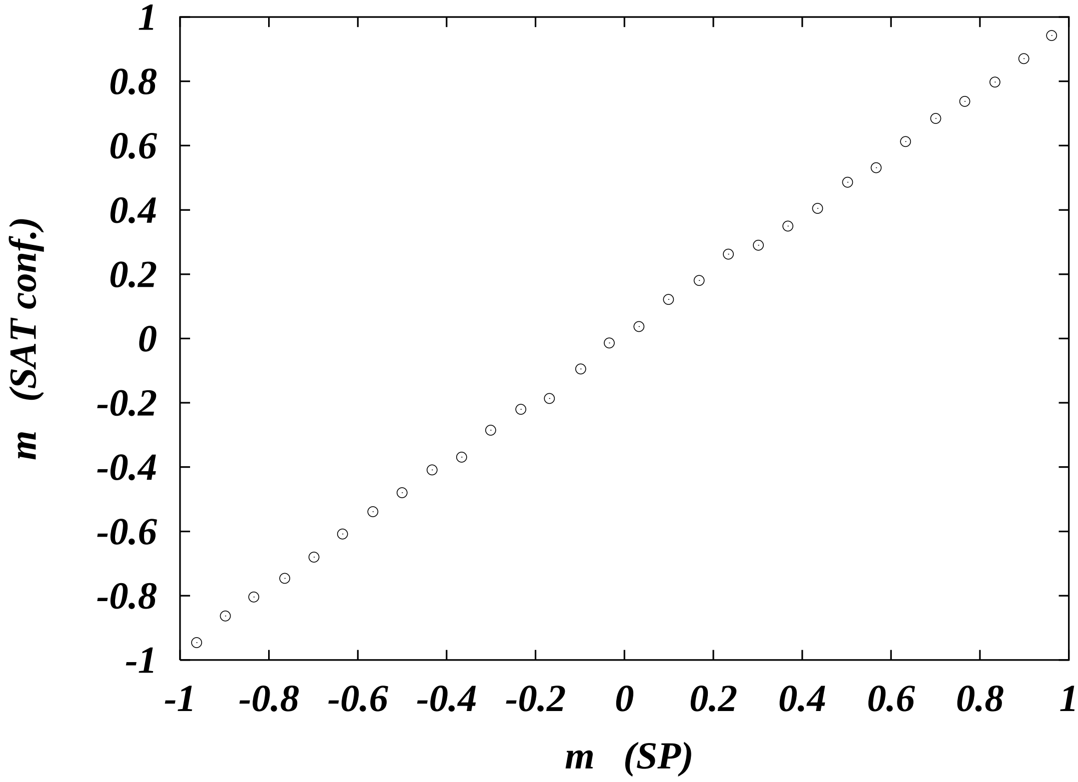


Message passing algorithm. Usually in computer science: “Belief propagation”=propagation of local magnetic fields. Converges only if unfrustrated. Here: Message = survey of magnetic fields in all states: CONVERGES.

Single sample result

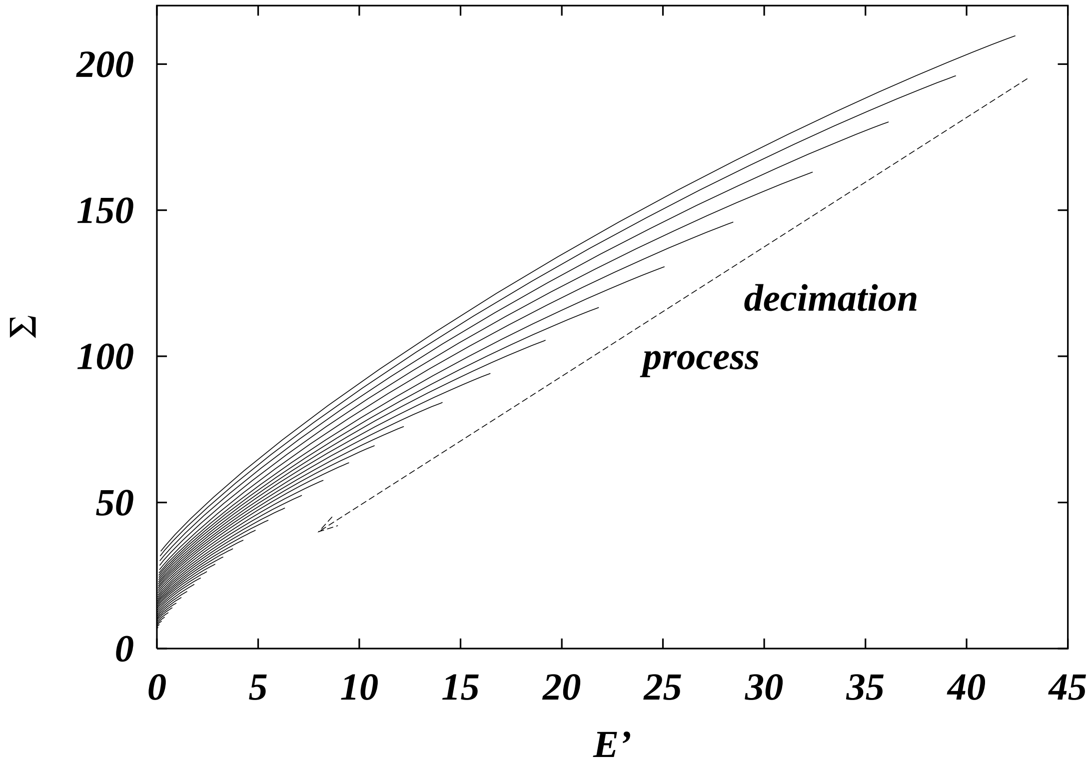
Result= Surveys of local magnetizations, in all states.

Check: compute many ground state with a standard algorithm, average the local spin in each state:



Survey Inspired Decimation

An algorithm to solve the Ksat problem: fix the spin which is most biased, rerun the survey propagation, iterate...



Solves the 'large' hard benchmarks of random 3sat at $\alpha = 4.2$ with $N = 1000, 2000$.

Solves typical random 3sat up to $N = 10^5$ at $\alpha = 4.2$.

Only competing algorithm: walksat35.

Local surveys of magnetic fields $\rightarrow$ a lot of information. Probably possible to invent many new algorithms.

Summary

- Analytic result on the generic samples of random 3sat: Phase diagram.

- Slowdown of algorithms near to $\alpha_c = 4.256$ due to the existence of a Hard SAT phase at $\alpha \in [3.921, 4.256]$, with exponentially many states.

- The whole construction can be checked versus rigorous computations on the '3spin model'.

- Single sample analysis: Survey propagation converges, yields very non trivial information on the sample (diversity of sites) $\rightarrow$ Survey Inspired Decimation: a very efficient algorithm for solving random 3sat problems.

References

- M. Mézard, G. Parisi and R. Zecchina, Science **297** (2002) 812.

- M. Mézard, R. Zecchina, *The random K-satisfiability problem: from an analytic solution to an efficient algorithm*, cond-mat/0207194.

- Mézard, M. & Parisi, G. *The cavity method at zero temperature*, cond-mat/0207121.

- M. Mézard, F. Ricci-Tersenghi, R. Zecchina, *Alternative solutions to diluted p-spin models and XOR-SAT problems*, cond-mat/0207140.