ALTERNATIVE SOLUTIONS TO DILUTED P-SPIN MODELS AND RANDOM XORSAT PROBLEMS

Federico RICCI TERSENGHI
Dep. Physics, Univ. "La Sapienza" Roma

In collaboration with

Marc Mézard (Orsay)
Riccardo Zecchina (ICTP)

Alfredo Braunstein (SISSA/ICTP)
Silvio Franz (ICTP)
Michele Leone (SISSA/ICTP)
Andrea Montanari (ENS, Paris)
Martin Weigt (Göttingen)

- Motivations
- Definition of the model
- Solution with the cavity approach } comparison
- Solution with rigorous methods } comparison
- Generalized XORSAT

Independent work on rigorous results by Cocco, Dubois, Mandler and Monasson

Motivations

Original motivation: simplify random 3-SAT keeping relevant features of the phase transition.

$$\mathcal{H}_{\text{3SAT}} \propto \alpha N - \sum_{i=1}^{N} H_i s_i + \sum_{ij} T_{ij} s_i s_j - \sum_{ijk} J_{ijk} s_i s_j s_k$$

$$\Downarrow$$

$$\mathcal{H}_{\text{3XORSAT}} \propto \alpha N - \sum_{ijk} J_{ijk} s_i s_j s_k$$

- Exactly solvable model with a hard-SAT phase, very similar to random K-SAT.

- Show correctness of replica and cavity calculations on a non-trivial model.

Definition of the model

1) Diluted p-spin model at $T=0$

$$\mathcal{H} = \sum_{\{i,j,k\}\in G} \left(1 - J_{ijk}\, s_i\, s_j\, s_k\right)$$

G: set of αN triples randomly chosen among $\binom{N}{3}$ (random hypergraph with average degree 3α)

$J_{ijk} = \pm 1$ quenched random variables

Find $s_i = \pm 1$ **such that** $s_i s_j s_k = J_{ijk} \quad \forall \{i,j,k\} \in G$

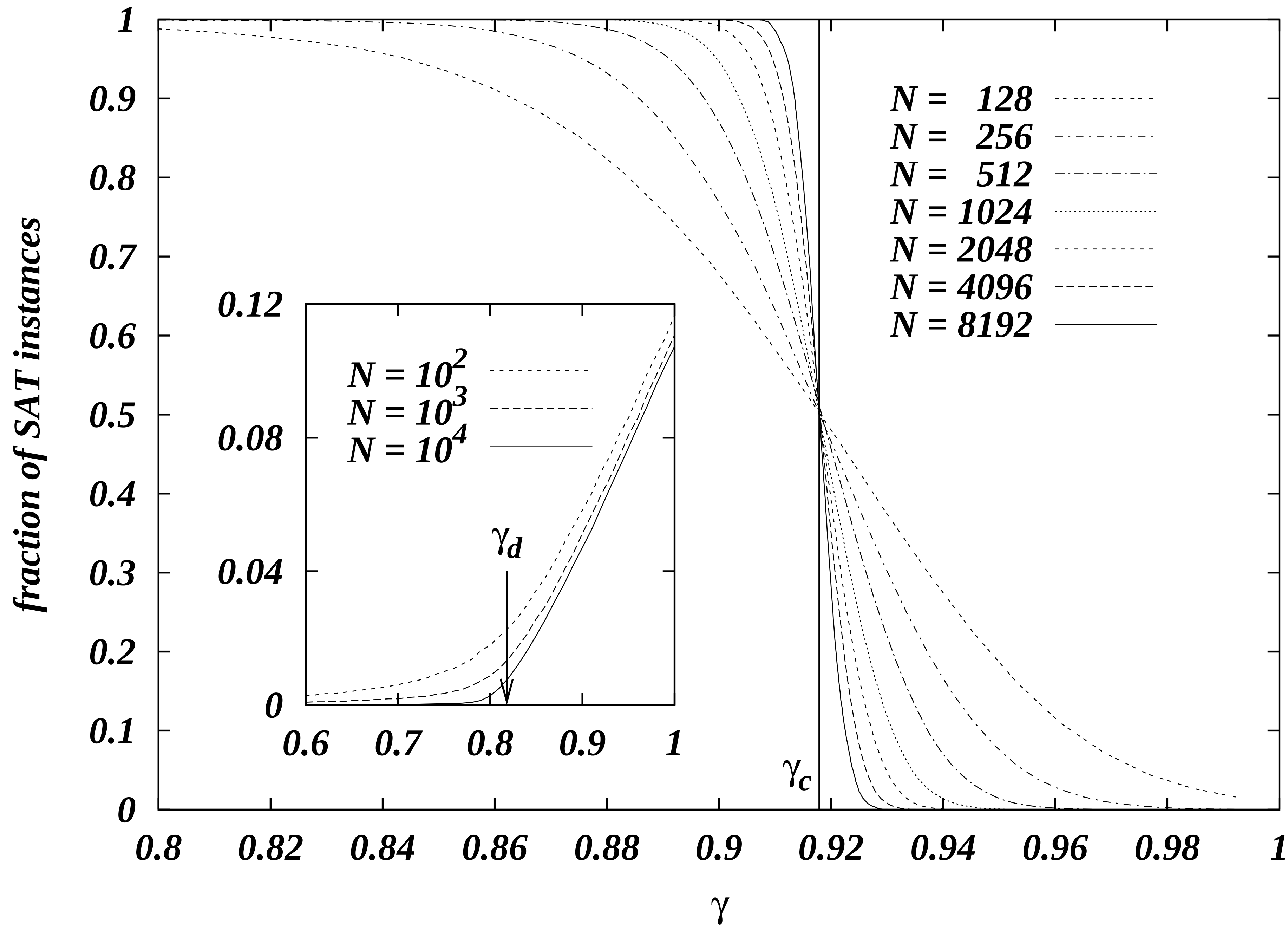
2) Set of αN linear equations mod 2 in N variables

$$s_i = (-1)^{x_i} \quad J_{ijk} = (-1)^{y_{ijk}} \quad x, y \in \{0,1\}$$

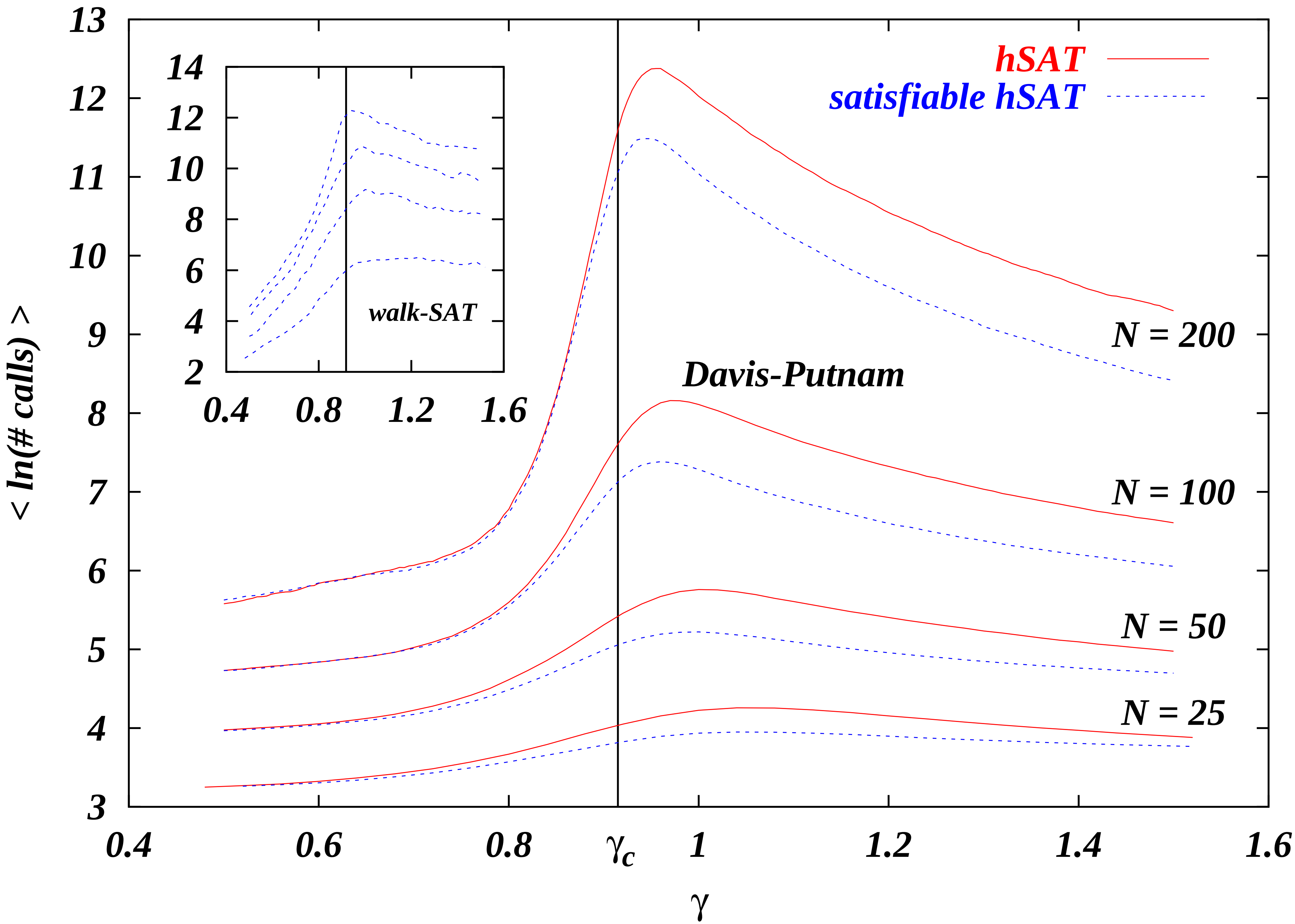
$$s_i\, s_j\, s_k = J_{ijk} \quad \Longleftrightarrow \quad x_i + x_j + x_k = y_{ijk} \mod 2$$

3) Random XORSAT problem:

$$F = \bigwedge_{\{i,j,k\}\in G} x_i \oplus x_j \oplus x_k \oplus y_{ijk}$$



fraction of SAT instances
γ
γ_c
γ_d
N = 128
N = 256
N = 512
N = 1024
N = 2048
N = 4096
N = 8192
N = 10^2
N = 10^3
N = 10^4



hSAT
satisfiable hSAT
Davis-Putnam
walk-SAT
N = 200
N = 100
N = 50
N = 25
< ln(# calls) >
γ
γ_c

Questions

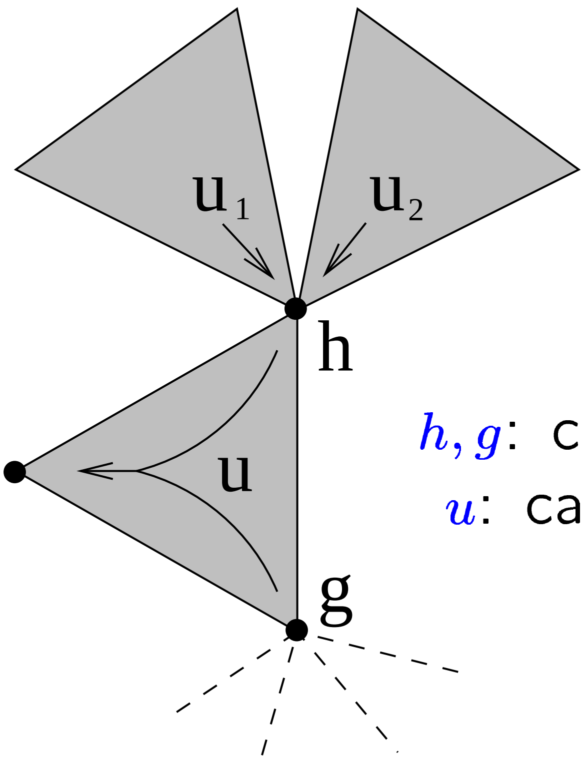
Can we locate the phase transition?

Which kind of phase transition is it?

Is there any hard-SAT region?

Why finding solutions is hard in this region?

Solution with cavity



h, g: cavity fields

u: cavity biases

$$P(h) \propto \int \prod_{i=1}^{k} dQ_i(u_i)\; \delta\Big(h - \sum u_i\Big)\; e^{-\mu(\sum |u_i| - |\sum u_i|)}$$

$$Q(u) = \int dP(h)\, dP(g)\; \delta\Big(u - \mathrm{sign}(hg)\Big)$$

Solution with cavity

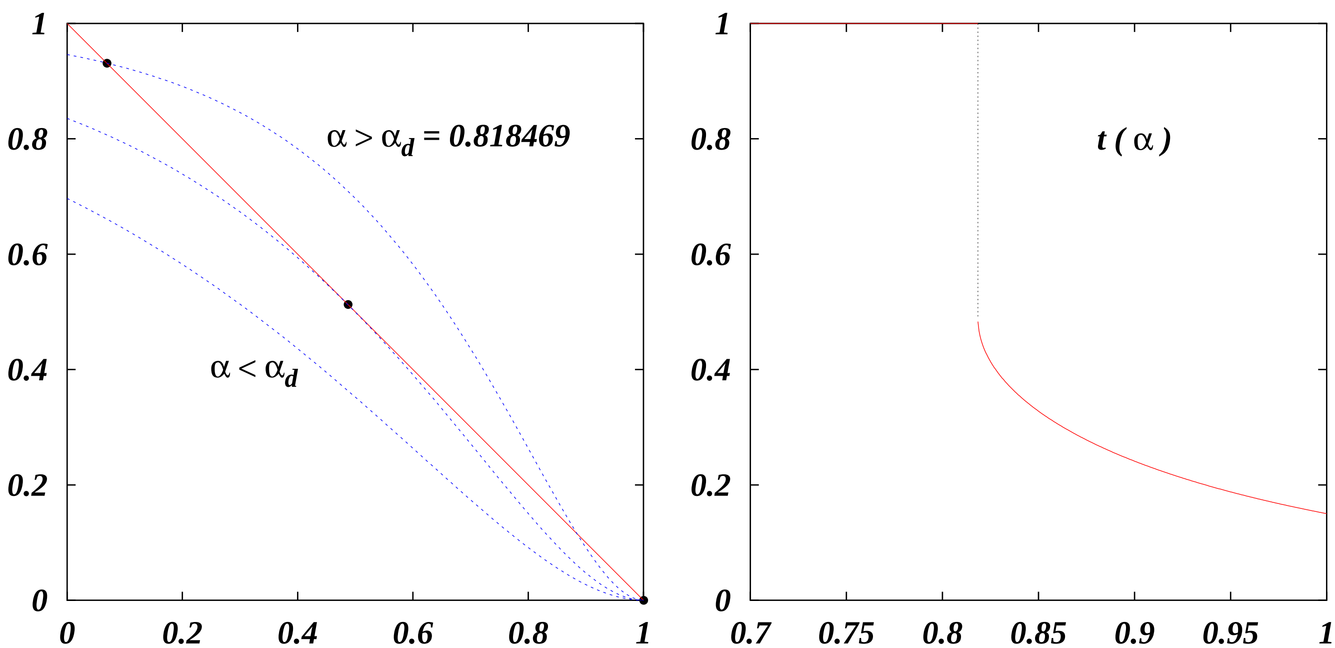
For $\mu \to \infty$

$$Q(u) = \begin{cases} \delta(u) & \text{prob.} \quad t \\ \frac{1}{2}\Big[\delta(u+1)+\delta(u-1)\Big] & \text{prob.} \quad 1-t \end{cases}$$

Prob. $P\Big(h = \sum_{i=1}^{k} u_i\Big)$ is non trivial $= 1 - t^k$

$$e^{-3\alpha}\sum_{k=0}^{\infty}\frac{(3\alpha)^k}{k!}(1-t^k) = 1 - e^{-3\alpha(1-t)}$$

$$1 - t = \Big(1 - e^{-3\alpha(1-t)}\Big)^2$$



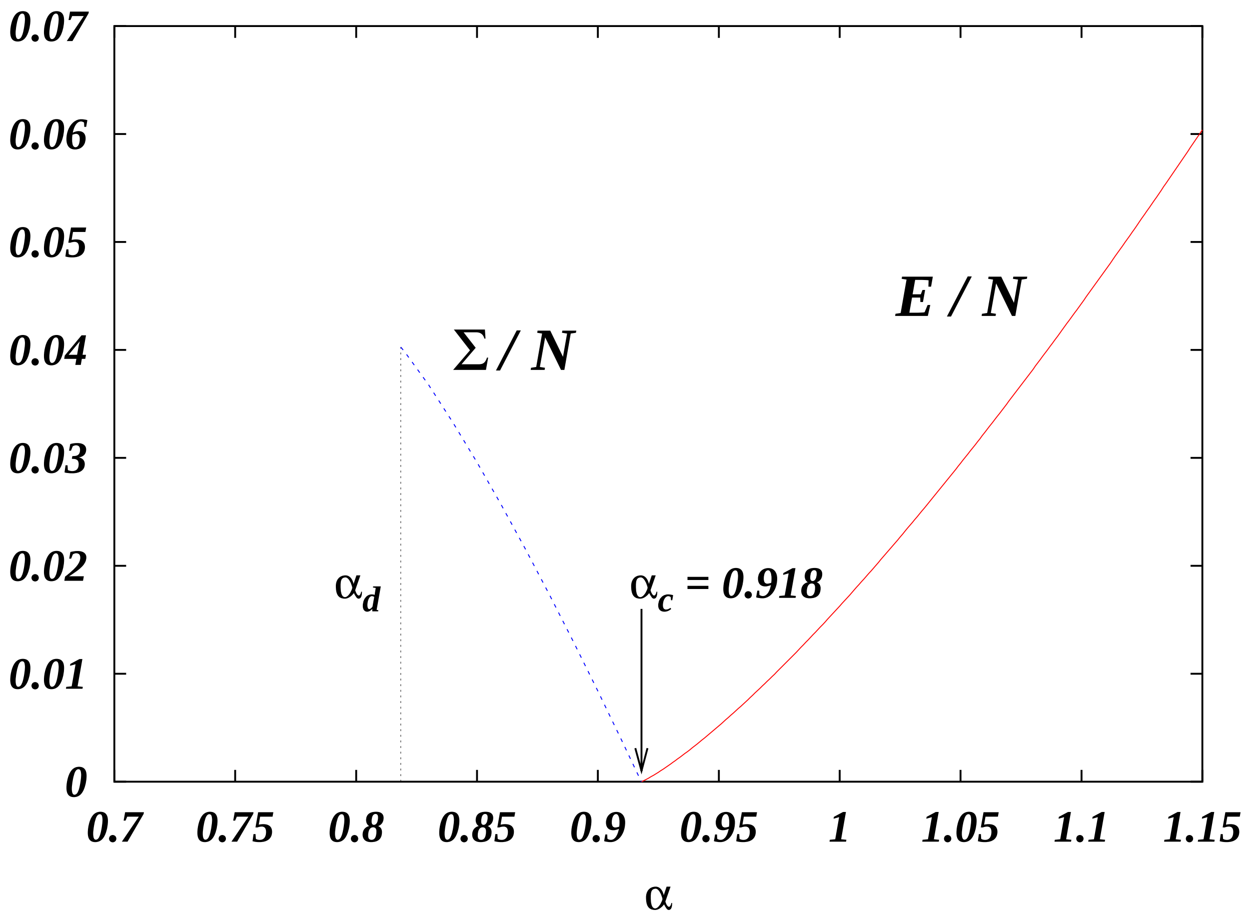
Solution with cavity

$$\lambda = 3\alpha(1-t) \qquad \lambda = 3\alpha\left(1-e^{-\lambda}\right)^2$$

For $\mu \gg 1$ $\quad \Phi(\mu) \simeq \frac{1}{\mu}\left(\psi - \omega\, e^{-2\mu}\right)$ with

$$\psi = \log(2)\left[\frac{\lambda}{3} - 1 + e^{-\lambda}\left(1+\frac{2}{3}\lambda\right)\right]$$

$$\omega = \frac{\lambda}{6}\left[2 - e^{-\lambda}(2+3\lambda)\right]$$



Rigorous solution

Physical idea

The hypergraph can be divided in 2 parts:
1) a central core, where $Q(u)$ are non-trivial and variables are more constrained;
2) an external part, made of dangling ends, where $Q(u)$ are trivial and variables much less constrained.

Variables in the non-core part induce large fluctuations in the number of solutions $\mathcal{N}$, such that

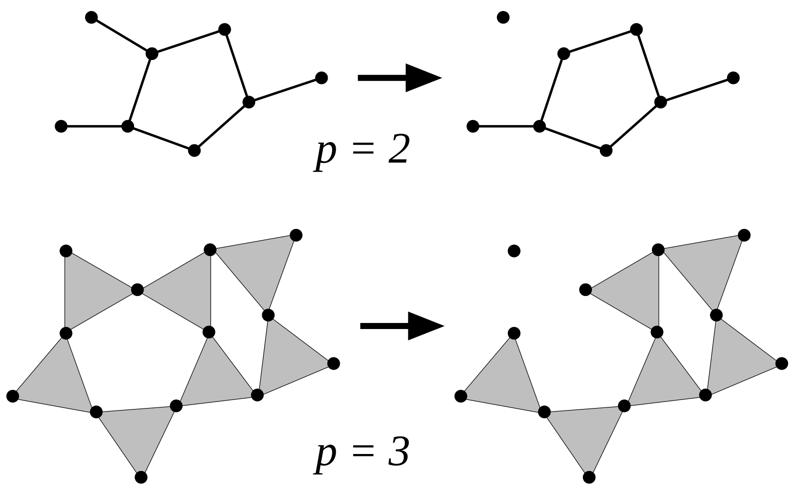
$$\overline{\log(\mathcal{N})} \neq \log(\overline{\mathcal{N}})$$

Plan

1. remove these variables (low connectivity)
2. calculate annealed averages $\overline{\mathcal{N}}$ and $\overline{\mathcal{N}^2}$ on the remaining hypergraph

Leaf removal algorithm

Rule: As long as there are variables of connectivity less than 2, remove them.



$$\frac{\partial f_0(t)}{\partial t} = (p-1)\frac{f_1(t)}{m(t)} + 1$$

$$\frac{\partial f_1(t)}{\partial t} = (p-1)\frac{2f_2(t) - f_1(t)}{m(t)} - 1$$

$$\frac{\partial f_k(t)}{\partial t} = (p-1)\frac{(k+1)f_{k+1}(t) - kf_k(t)}{m(t)}$$

where $m(t) = \sum_k k f_k(t) = p(\alpha - t)$.

$$f_k(t) = e^{-\lambda(t)}\frac{\lambda(t)^k}{k!} \qquad \forall k \geq 2$$

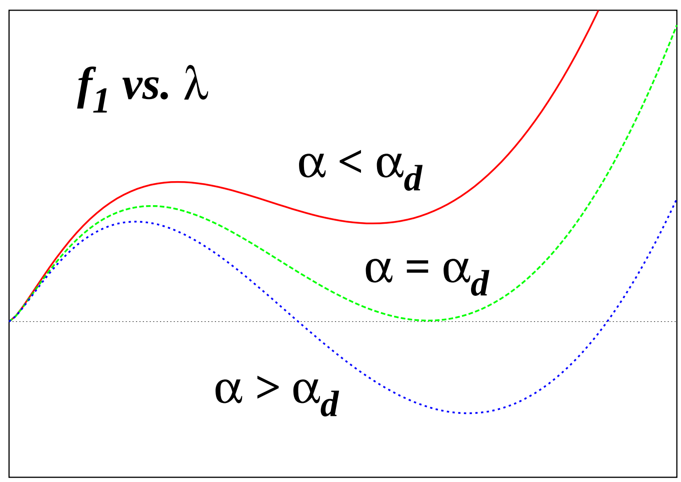
Leaf removal algorithm

Solution for $p = 3$

$$\lambda(t) = 3\sqrt[3]{\alpha(\alpha - t)^2} \qquad f_1(\lambda) = \lambda\left(e^{-\lambda} - 1 + \sqrt{\frac{\lambda}{3\alpha}}\right)$$

First order transition at $\alpha = \alpha_d = 0.818469$

The core appears in a discontinuous way



On the core:

$$f_0 = f_1 = 0 \qquad f_k = \frac{\lambda^k}{k!}e^{-\lambda} \quad \forall k \geq 2$$

where λ solves $\lambda = 3\alpha\left(1 - e^{-\lambda}\right)^2$

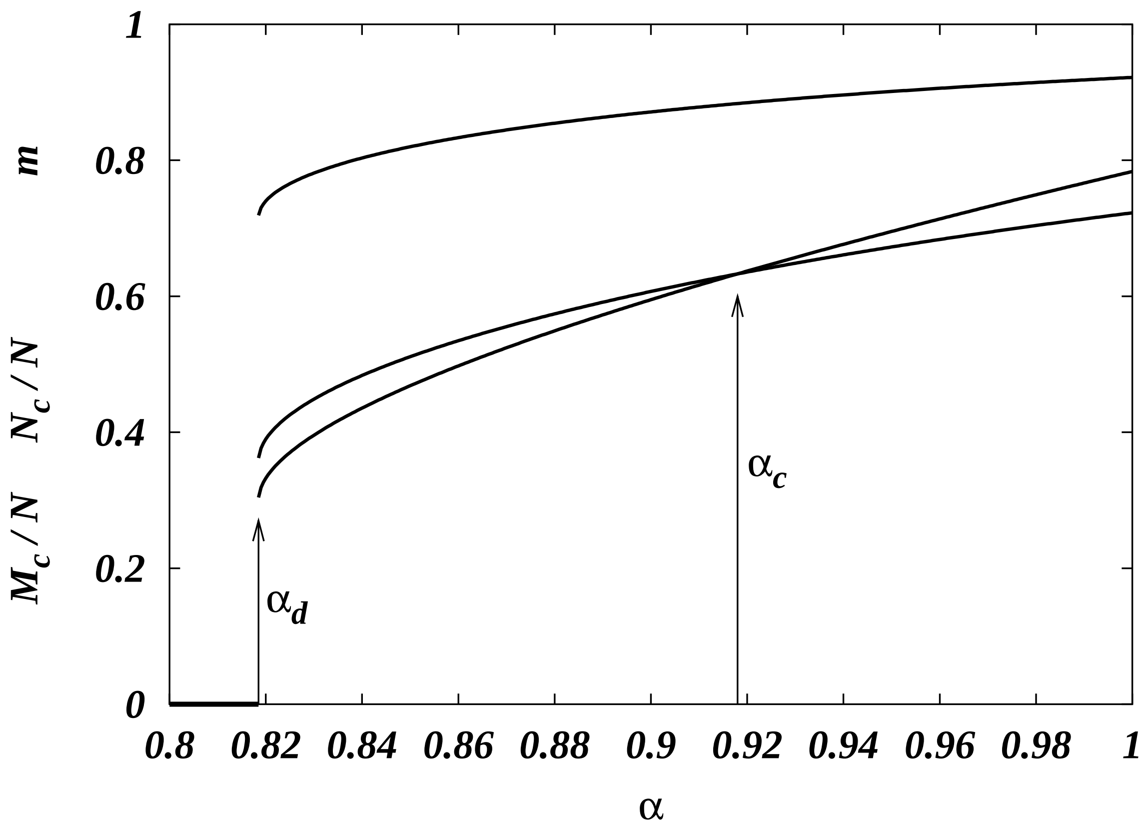
$\mathcal{N}_c$: number of solution on the core

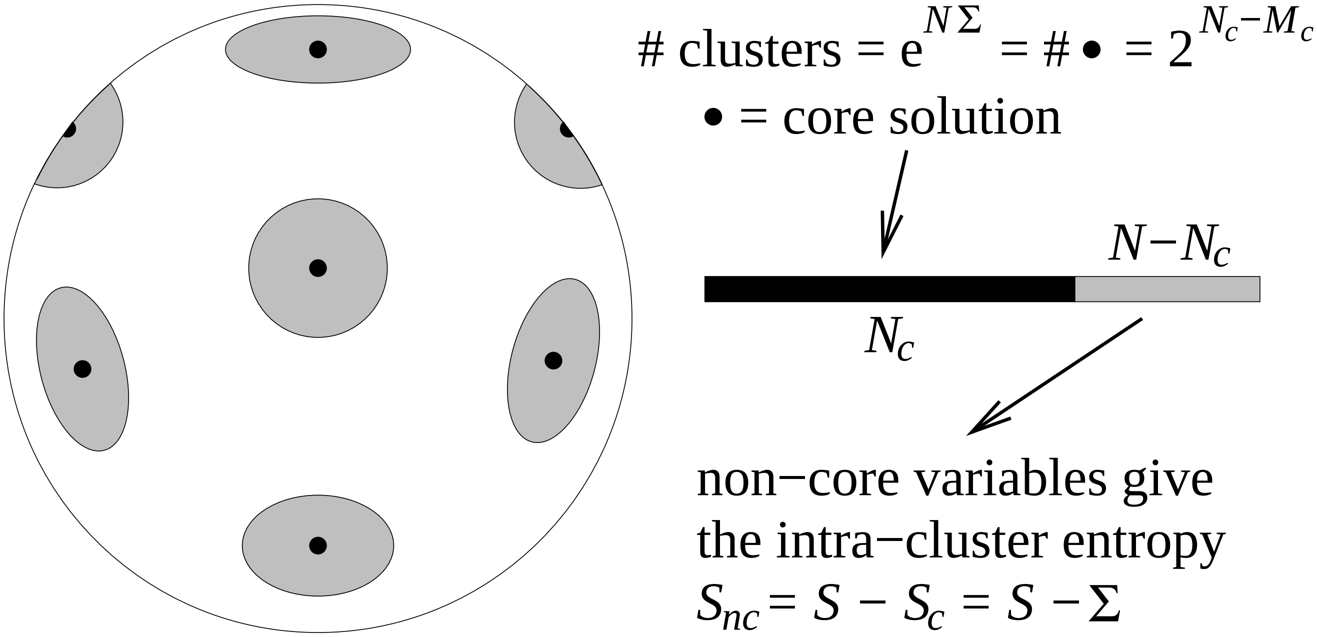
For $N \to \infty$,

$$\overline{\mathcal{N}_c} = 2^{N_c - M_c} \quad , \quad \frac{\overline{\mathcal{N}_c^2} - \overline{\mathcal{N}_c}^2}{\overline{\mathcal{N}_c}^2} \to 0$$

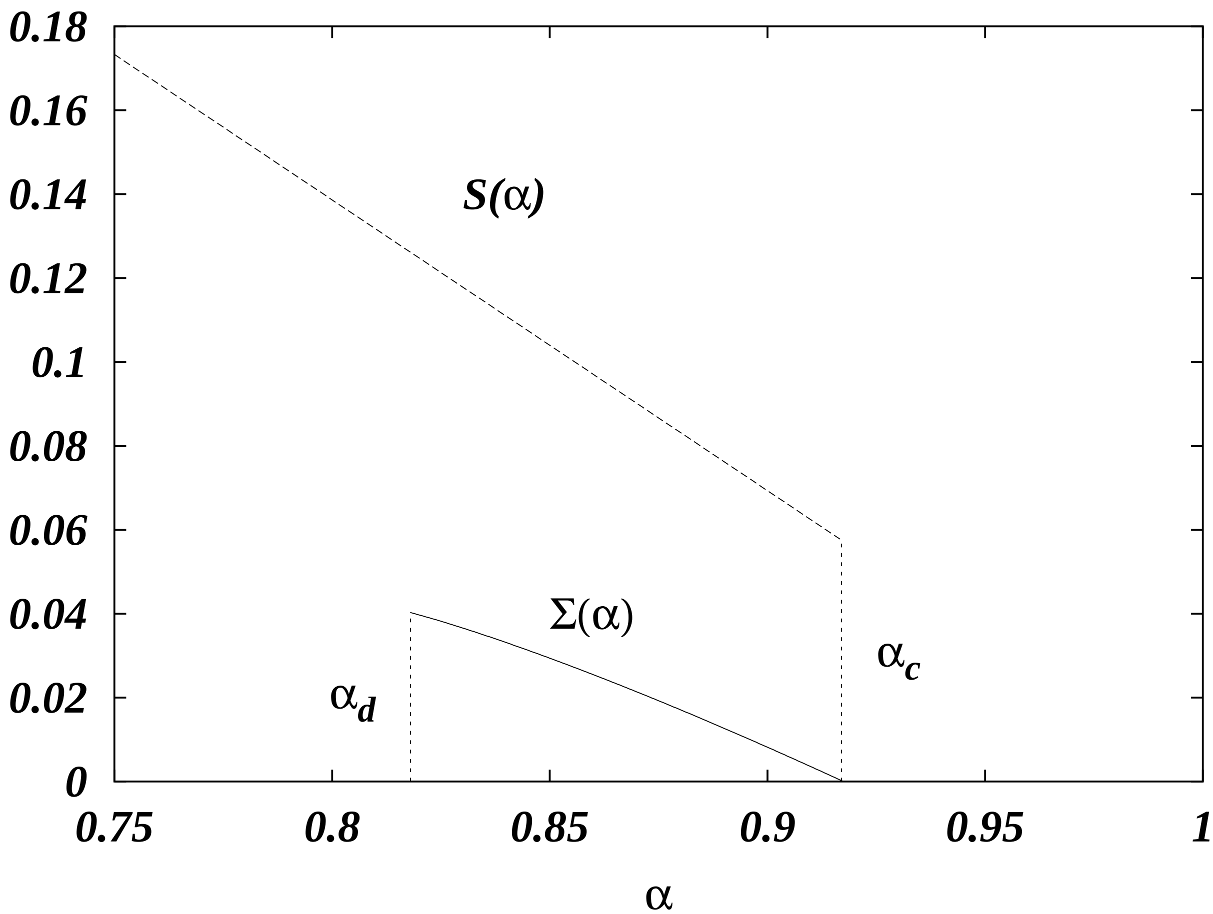
$$\overline{\log \mathcal{N}_c} = \log \overline{\mathcal{N}_c} = \log(2)\left(N_c - M_c\right)$$

$$\begin{aligned} N_c(\alpha) &= N\left[1-(1+\lambda)e^{-\lambda}\right] \\ M_c(\alpha) &= N\frac{\lambda}{p}\left(1-e^{-\lambda}\right) \end{aligned}$$

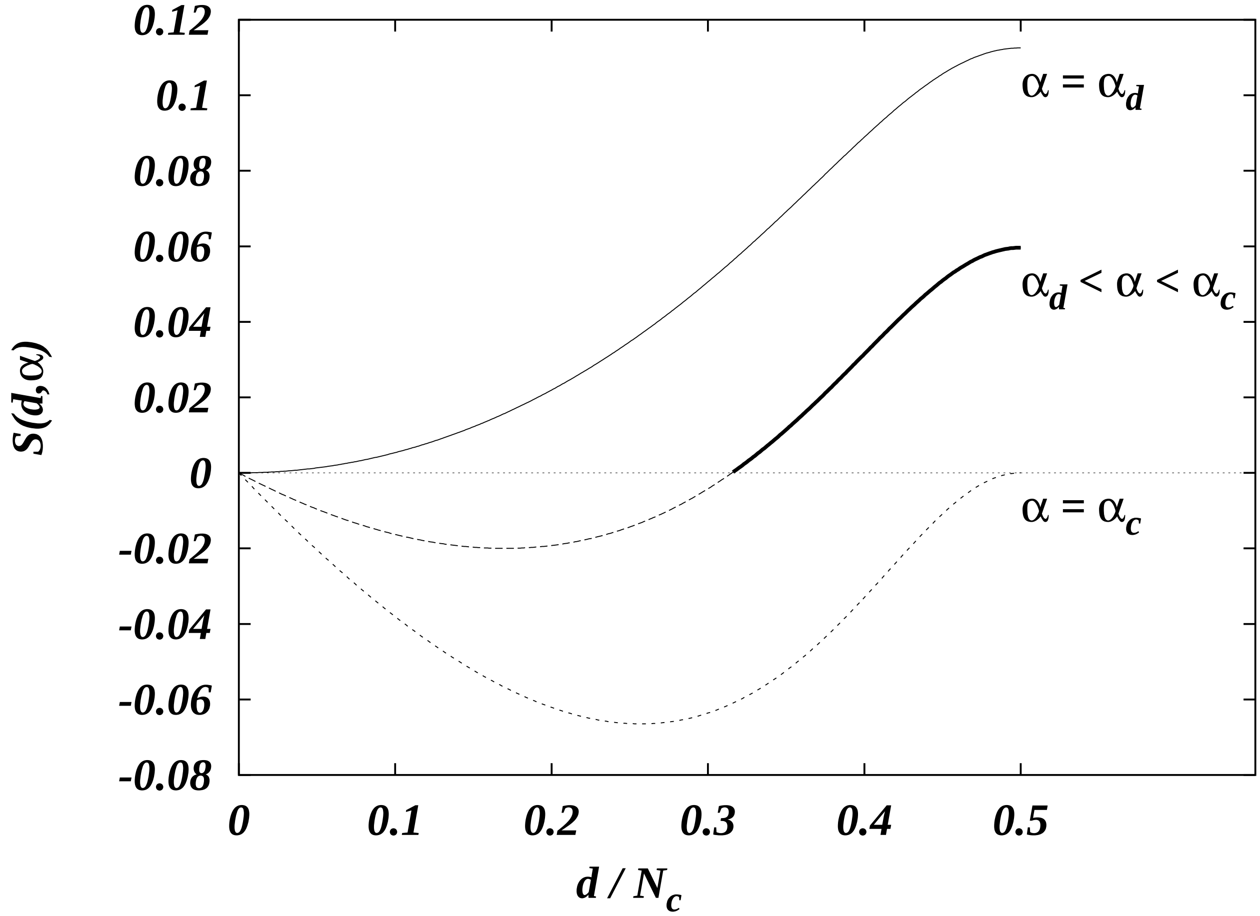


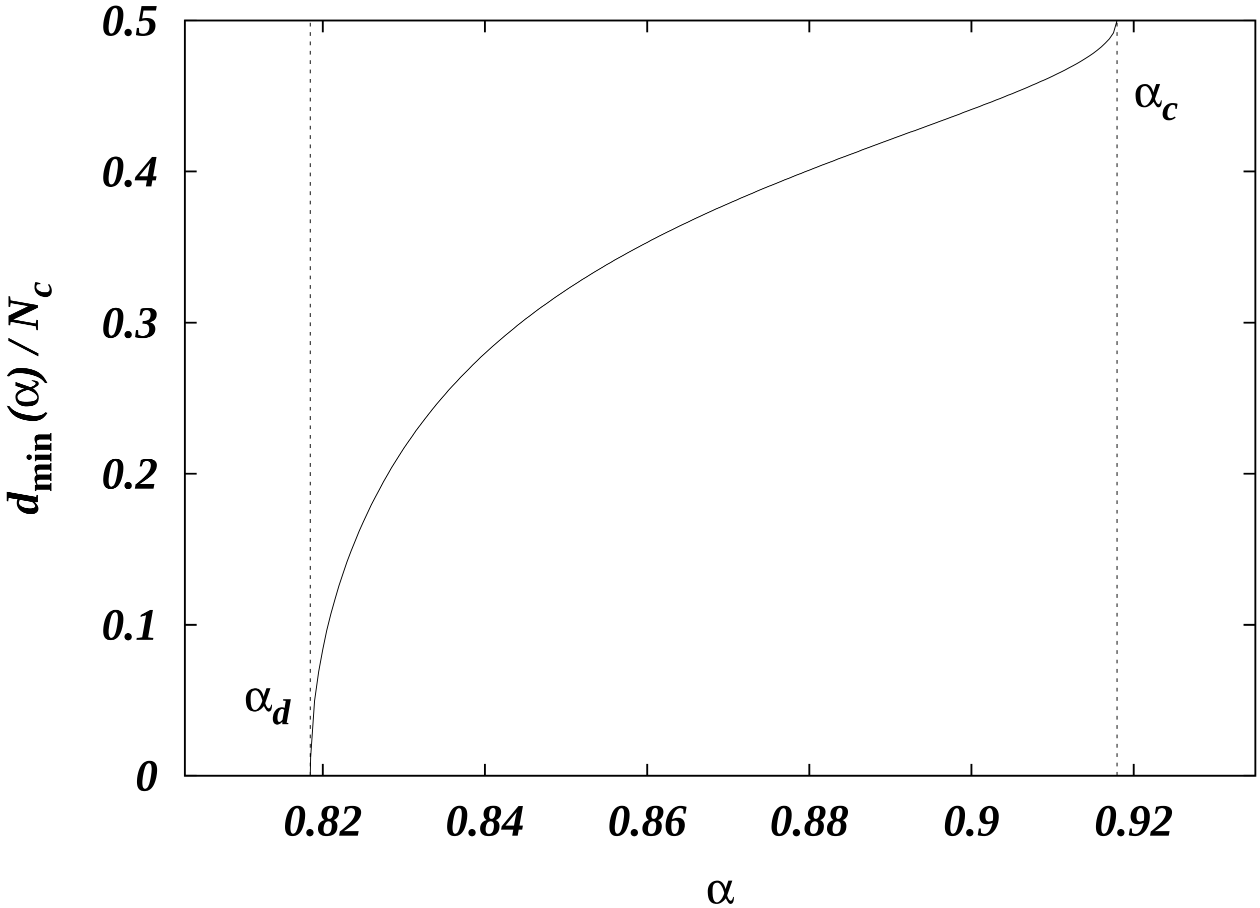


clusters = $e^{N\Sigma}$ = # • = $2^{N_c - M_c}$
• = core solution
$N - N_c$
N_c
non−core variables give
the intra−cluster entropy
$S_{nc} = S - S_c = S - \Sigma$



0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
0.75
0.8
0.85
0.9
0.95
1
$S(\alpha)$
$\Sigma(\alpha)$
α_d
α_c
α


0.12
0.1
0.08
0.06
0.04
0.02
0
-0.02
-0.04
-0.06
-0.08
S(d,α)
0
0.1
0.2
0.3
0.4
0.5
d / N_c
α = α_d
α_d < α < α_c
α = α_c


0.5
0.4
0.3
0.2
0.1
0
d_min (α) / N_c
0.82
0.84
0.86
0.88
0.9
0.92
α
α_d
α_c

Generalized XORSAT

f_k: fraction of variables with conn. k
g_p: fraction of clauses with p variables ($g_0 = g_1 = 0$)

$$f(z) = \sum_k f_k\, z^k \qquad g(z) = \sum_p g_p\, z^p \qquad f'(1) = \alpha\, g'(1)$$

α_d is determined by the appearance of positive solutions to the equation

$$1 - x = \frac{f'\left[1 - \frac{g'(x)}{g'(1)}\right]}{f'(1)} \qquad (1)$$

α_c is determined by the equation

$$\alpha\, g(x^*) = 1 - \alpha\,(1 - x^*)\, g'(x^*) - f\left[1 - \frac{g'(x)}{g'(1)}\right]$$

where $x^*(\alpha)$ is the largest solution to Eq.(1)