Phase transitions in optimization problems

Viljo Petäjä

June 17, 2002

1 Introduction

A common everyday example of a phase transition is the melting of a solid with increasing temperature. As temperature increases the atomic vibrations increase gradually. At the phase transition (that is melting) there is a sudden change in the properties of the substance leading to the appearance of a qualitatively different phase – a liquid.

Similar behaviour has been also observed in computational problems. While we slowly vary a control parameter of the problem a sharp transition may occur in the solvability. One can also observe the change in the average computational complexity. There appears an 'easy-hard-easy' pattern in the average complexity as a function of the control parameter. These similarities motivate to tackle computational problems with the tools of statistical physics. One can draw a table of the equivalences between two languages:

OPTIMIZATION	STATISTICAL PHYSICS
instance	sample
cost function	energy
optimal configuration	ground state
minimal cost	ground state energy

In this paper we describe the analysis of the K-SAT problem including the replica method. We discuss the properties of the solutions and how these

are related to the sat/unsat transition. The analysis of the K-SAT problem is based on [1]. In addition we briefly discuss the threshold behaviour of random NK landscapes.

2 Statistical physics and the replica method

When the system is in equilibrium it will be in configuration C with probability $p(C)$:

$$p(C) = \frac{1}{Z} \exp\left(-\frac{1}{T} E(C)\right) \tag{1}$$

where $E(C)$ is the energy corresponding to configuration C and the normalization factor Z is the partition function

$$Z = \sum_{C} \exp\left(-\frac{1}{T} E(C)\right) \tag{2}$$

Knowing the partition function allows us to calculate the free energy of the system:

$$F(T) = -T \ln Z(T) \tag{3}$$

For a mechanically isolated systems kept at constant temperature the state of equilibrium is the state of minimum free energy. Thermodynamic equilibrium prevails when the thermodynamic state (configuration) of the system does not change with time.

Consider a generic model with N spins σ_i and energy function $E(C, J)$ depending on random couplings J. In addition, assume that the free-energy $F(J)$ of this model is self-averaging [1] and one would like to compute its quenched averaged value $\overline{F(J)}$. That is, first calculate F for each randomly drawn J and then calculate the average. According to (3) we need to calculate $\overline{\ln Z(J)}$. Its computation is generally a very hard task from the analytic point of view. However, this can be done with the non-rigorous replica method. Start by writing the Taylor expansion of Z^n with respect to n (recall that $\frac{dZ^n}{dn} = Z^n \ln Z$):

$$Z(J)^n = 1 + n \ln Z(J) + O(n^2). \tag{4}$$

[1] The distributions of some random variables become highly concentrated as the size of the system grows. This is self-averaging.

This is valid for any set of couplings J and small real n. Now insert $\ln Z(J) = (Z(J)^n - 1)/n$ into (3) and in order to get ride of n take the limit as $n \to 0$:

$$\overline{F(J)} = -T \lim_{n \to 0} \left(\frac{\overline{Z(J)^n} - 1}{n} \right) \tag{5}$$

If we restrict to integer n, the n^{th} moment of the partition function Z can be written as

$$\overline{Z(J)^n} = \overline{\left[\sum_C \exp\left(-E(C,J)/T\right) \right]^n} = \sum_{C^1,...,C^n} \overline{\exp\left(\frac{1}{T} \sum_{a=1}^{n} E(C^a, J) \right)} \tag{6}$$

In the rightmost expression we have n copies, or replicas, of the initial problem. The random couplings disappear once the average over the quenched couplings has been carried out. Finally, we must compute the partition function of an abstract system of N vectorial spins $\vec{\sigma}_i = (\sigma_i^1, ..., \sigma_i^n)$ with the non random energy function

$$E_{eff}(\{\vec{\sigma}_i\}) = -T \ln \left[\overline{\exp\left(-\frac{1}{T} \sum_{a=1}^{n} E(C^a, J) \right)} \right] \tag{7}$$

The corresponding partition function $\overline{Z(J)^n} = \sum_{\{\vec{\sigma}_i\}} \exp(-E_{eff}(\{\vec{\sigma}_i\})/T)$ can be estimated analytically in some cases by means of the saddle-point method. The result may be written formally as

$$\overline{Z(J)^n} = \exp(-N\tilde{f}(n)/T) \tag{8}$$

to leading order in N where $\tilde{f}(n)$ is a function of integer n satisfying $\tilde{f}(0) = 0$. Now, continue analytically $\tilde{f}$ to the set of real n and obtain $\overline{F(J)} = TNd\tilde{f}/dn$ [2] evaluated at $n = 0$.

3 Random K-SAT

A random version of the K-SAT problem is defined as follows. Consider N boolean variables x_i, $i = 1, ..., N$. Clauses C are formed of K randomly

[2] Using (5) one gets

$$\overline{F(J)} = -T \lim_{n \to 0} \frac{e^{-N\tilde{f}(n)} - e^{-N\tilde{f}(0)}}{n} = -T \frac{de^{-N\tilde{f}(n)}}{dn}\bigg|_{n=0}$$

chosen variables, each of them negated with probability 1/2. This is repeated by drawing independently M random clauses C_l, $l = 1, ..., M$. In numerical experiments it has been studied the probability $P_N(\alpha, K)$ that a randomly chosen formula having $M = \alpha N$ clauses is satisfiable. For $N \to \infty$, $P_N \to 1$ for $\alpha < \alpha_c(K)$ and $P_N \to 0$ for $\alpha > \alpha_c(K)$. For $K = 2$ there is a known exact result $\alpha_c(2) = 1$ [2, 3] , whereas for $K = 3$ there are only estimatas. From numerical studies one has $\alpha_c(3) \simeq 4.3$ [4]. It has been observed numerically that hard random instances are generated when the problems are close to the sat/unsat phase boundary.

3.1 Application of methods of statistical physics

To apply the statistical physics approach one has to identify the energy function corresponding to the K-SAT problem. Consider a K-SAT problem instance with M random clauses C_l and with N Boolean variables x_i.

First, map the Boolean variables x_i to the spin variables S_i:

$$\begin{aligned} S_i &= +1 \quad \text{if } x_i \text{ is } true \\ S_i &= -1 \quad \text{if } x_i \text{ is } false \end{aligned} \tag{9}$$

The random clauses can be encoded into an $M \times N$ matrix C_{li} in the following way:

$$\begin{aligned} C_{li} &= +1 \quad \text{if the clause } C_l \text{ includes } x_i \\ C_{li} &= -1 \quad \text{if the clause } C_l \text{ includes } \bar{x}_i \\ C_{li} &= 0 \quad \text{otherwise} \end{aligned} \tag{10}$$

One can check that $\sum_{i=1}^{N} C_{li} S_i$ equals to $-K$ if all literals have a wrong [3] assignment in clause l. The cost-function $E[\mathbf{C}, \mathbf{S}]$ is defined as the number of clauses that are not satisfied by the logical assignment corresponding to configuration $\mathbf{S}$:

$$E[\mathbf{C}, \mathbf{S}] = \sum_{l=1}^{M} \delta \left(\sum_{i=1}^{N} C_{li} S_i + K \right) , \tag{11}$$

where $\delta()$ denotes the Kronecker delta function. The minimum $E[\mathbf{C}]$ of $E[\mathbf{C}, \mathbf{S}]$ is the lowest number of unsatisfied clauses that can be achieved by

[3] For example with $x_1 = T$, $x_2 = T$ and $x_3 = F$ in $(x_1 \vee \overline{x_2} \vee x_3)$, x_2 and x_3 have 'wrong' assignments.

the best possible logical assignment. $E[\mathbf{C}]$ is a random variable (because $\mathbf{C}$'s are random) that becomes highly concentrated around its average value $E_{GS} \equiv \overline{E[\mathbf{C}]}$ as $N \to \infty$ [4]. Since E_{GS} is the minimal number of violated clauses averaged over all $\mathbf{C}$, $E_{GS} = 0$ in the sat region and $E_{GS} > 0$ in the unsat phase. The knowledge of E_{GS} as a function of α therefore determines the threshold ratio $\alpha_c(K)$. E_{GS} can be calculated by statistical physics means from

$$E_{GS} = -T\overline{\ln Z(\mathbf{C})} + O(T^2) \tag{12}$$

when the auxiliary parameter T is sent to zero. As in the previous section:

$$Z = \sum_{\mathbf{S}} \exp\left(-\frac{1}{T}E[\mathbf{C},\mathbf{S}]\right). \tag{13}$$

In calculating of E_{GS} one applies the replica method by exploiting (4). By replicating n times the sum over the spin configurations $\mathbf{S}$ and averaging over the clause distribution one obtains:

$$\overline{Z[\mathbf{C}]^n} = \sum_{\mathbf{S}^1,...,\mathbf{S}^n} \overline{\exp\left(-\sum_{a=1}^{n} E[\mathbf{C},\mathbf{S}^a]/T\right)} \tag{14}$$

The averaged term in the r.h.s. of (14) depends on the $n \times N$ spin values only through the 2^n occupation fractions $x(\vec{\sigma})$ labeled by the vectors $\vec{\sigma}$ with n binary components; $x(\vec{\sigma})$ equals the number (divided by N) of labels i such that $S_i^a = \sigma^a$, $\forall a = 1,...,n$. It follows that $\mathbf{x} \to \{\mathbf{S}^a\}$. To leading order in N the n^{th} moment of Z can be written as

$$\overline{Z[\mathbf{C}]^n} \simeq \exp(-N f_{opt}/T) \tag{15}$$

where f_{opt} is the optimum over all possible $\mathbf{x}$s of the functional $f[\mathbf{x}]$ [5]. f_{opt} can also be interpreted as a free-energy density, since from (15) one gets $f_{opt}N \simeq -T\ln\overline{Z[\mathbf{C}]^n}$.

The optimization conditions over $f[\mathbf{x}]$ gives 2^n coupled equations for $\mathbf{x}$s. In this case f is a symmetric functional [5], invariant under any permutation of replicas. The optimum may thus be sought in the so-called replica symmetric (RS) subspace. In the limit $T \to 0$ and with the RS subspace, the occupation fractions may by expressed as the moments of a probability density $P(m)$ over the range $-1 \le m \le 1$ [5]. In the limit $n \to 0$ $P(m)$ is the probability density of the expectation values of the spin variables over the set of ground states. Consider all the spin configurations $\mathbf{S}^{(j)}$, $j = 1,...,Q$ realizing the

[4] That is, $E[\mathbf{C}]$ is self-averaging. GS stands for 'ground state'

minimum $E[\mathbf{C}]$ of the cost-function $E[\mathbf{CS}]$ of the MAX-SAT problem. Define the average magnetizations of the spins over the set of optimal configurations:

$$m_i = \frac{1}{Q} \sum_{j=1}^{Q} S_i^{(j)}. \tag{16}$$

Call $H(\mathbf{C}, m)$ the histogram of the m_is and $H(m)$ its quenched average. $H(m)$ is a probability density over the interval $-1 \leq m \leq 1$. $H(m \simeq 0)$ corresponds to weakly constrained variables and $H(m = \pm 1)$ represents a 'backbone' of completely constrained variables, i.e. variables that have the same value in all ground state configurations. If the RS solution is the global optimum of $f(\mathbf{x})$ then $H(m)$ equals to $P(m)$ in the limit $N \to \infty$ [5].

3.2 Sat phase

The typical number of solutions can be obtained from the ground state entropy density $s_{GS}(\alpha)$ [5] that is given by $-f_{opt}/T$ in the $T \to 0$ limit. If there are no clauses all assignments are solutions: $s_{GS}(\alpha = 0) = \ln 2$. It has been computed [6] the Taylor expansion of $s_{GS}(\alpha)$ in the vicinity of $\alpha = 0$. Results are shown in Figure 1. It is found that $s_{GS}(\alpha_c = 1) = 0.38$ and $s_{GS}(\alpha_c = 4.2) = 0.1$ for 2-SAT and 3-SAT respectively. The fact that $s_{GS} = const. > 0$ just below the threshold means there are exponentially many solutions as a function of N ($s_{GS}N = \ln \#$).

The RS calculations are believed to be exact at low α ratios. According to analytical and numerical calculations for 3-SAT [7] the RS theory breaks down at a definite ratio $\alpha_{RSB} < \alpha_c$, where the solutions start to be organized into distinct clusters (RSB – replica symmetry breaking). Consider the space of N-spin configurations as the 2^N vertices on the N-dimensional hyper cube. Optimal assignments are a subset of the vertices. Effectively RS assumes that any pair of vertices are separated by the same Hamming distance d, defined as the fraction of distinct spins in the corresponding configurations. Solutions are in a single cluster of diameter Nd. This holds for $\alpha < \alpha_{RBS}$, where the space of solutions is replica symmetric and the solutions are characterized by a single probability distribution of local magnetizations. At $\alpha_{RSB} \simeq 4.0$, the space of solutions breaks into a large number of different clusters. Now there are two typical Hamming distances: the distance between solutions

[5] The entropy S = Log of the number of configurations which contribute at a fixed energy. On the other hand $F = E - TS$, now the internal energy $E = 0$.

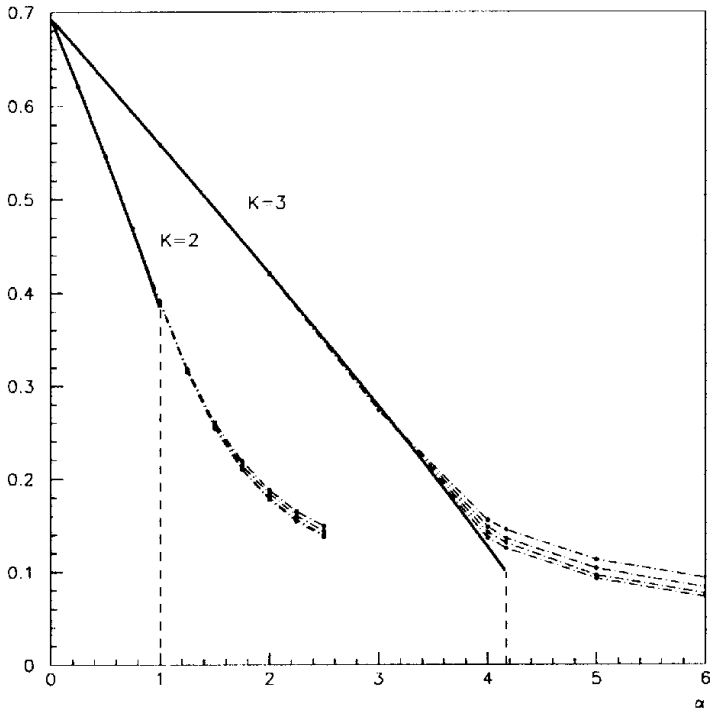

Figure 1: RS estimate for the entropy density (Y-axis) in random 2-SAT and 3-SAT (bold lines). The dots represent the results of exact enumeration in small systems ($N = 20...30$) [6].

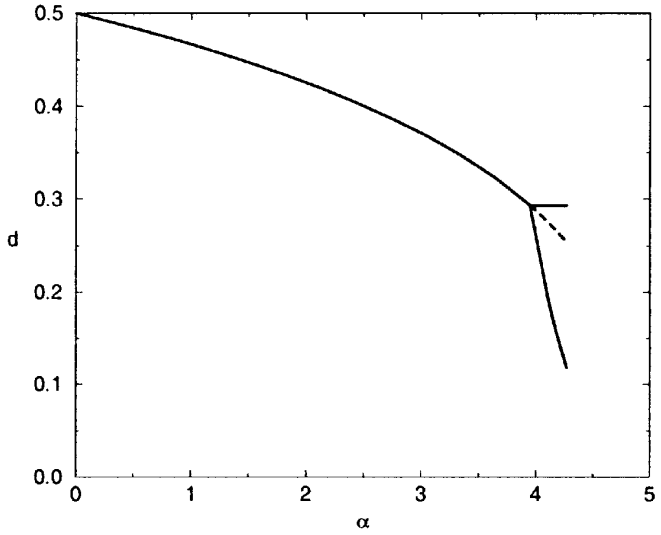

Figure 2: d is the typical Hamming distance between solutions. The splitting of curves at $\alpha \simeq 4$ corresponds to clustering.

belonging to different clusters and the intra-cluster distance. Figure 2 [1] shows qualitatively the clustering process. Thus, the space of solutions has a highly organized structure even in the sat phase.

3.3 Unsat phase

It is expected that $O(N)$ variables become totally constrained in the unsat phase. This hypothesis corresponds to a structural change in the probability distribution $P(m)$ which develops Dirac peaks at $m = \pm 1$. In the limit $T \to 0$ this is taken into account by introducing a new variable z, defined by the relation $m = \tanh(z/T)$. Instead of $P(m)$ now one works with the probability density $R(z)$ and can calculate the corresponding ground state energy density e_{GS}. The location of the sat/unsat threshold can be obtained for any K by looking at the value α beyond which e_{GS} becomes positive. For $K = 2$ one obtains correctly $\alpha_c(2) = 1$, for $K = 3$ $\alpha_c(3) \simeq 4.6$ which is slightly higher that the value estimated by numerical simulations $\alpha_c(3) \simeq 4.3$ [4]. The RS theory provides an upper bound for the thresholds for any $K > 2$.

3.4 Nature of the phase transition

An qualitative difference between the sat/unsat transitions in 2-SAT and 3-SAT is the way $P(m)$ changes at the threshold. Instead of studying $P(m)$ itself let us consider the fraction of the fully constrained variables $\gamma(\alpha, K)$ which is directly computable in the RS theory. Clearly, $\gamma(\alpha, K)$ wanishes in the sat region both for $K = 2$ and $K = 3$. Two kinds of scenarios have been found when entering the unsat phase. For 2-SAT, $\gamma(\alpha, 2)$ changes smoothly, whereas for 3-SAT (and more generally $K \geq 3$), $\gamma(\alpha, 3)$ has a discontinuous jump to a finite value γ_c slightly above the threshold. Thus, one can say that for 2-SAT the phase transition is of the second order and for $K \geq 3$ of the first order. The RS theory gives $\gamma_c(K = 3) \simeq 0.94$. Subtracting the terms in $P(m)$ which appear to be splitting away from $m = \pm 1$ reduces that to an estimated $\gamma_c(3) \simeq 0.6$ in an approximation beyond replica symmetry, while experiments give $\gamma_c(3) \simeq 0.4$. Figure 3 presents numerical results which however are not very supportive due to relatively small N. It is interesting to notice that the phase transition is continuous in the entropy [6] as one can also see from the numerical data of Figure 2. This means that while an extensive number of degrees of freedom freeze at the transition, forming the backbone, the remaining ones can support critical fluctuations.

3.5 The random $2 + p$-SAT model

To understand the onset of exponential complexity that occurs when going from a problem in P (2-SAT) to a problem that is NP-complete (3-SAT), it has been introduced a mixed model which continuously interpolates between 2-SAT and 3-SAT [8]. Consider a random formula with M clauses, of which fraction $(1 - p)$ contains two variables and fraction p three variables. $p = 0$ corresponds to the pure 2-SAT and $p = 1$ to the 3-SAT problem. In [8] it has been studied how the nature of the sat/unsat transition changes with p.

Using the replica method it has been found a continuous transition at $\alpha_c(2 + p) = 1/(1 - p)$ for $p < p_0$, where $p_0 \simeq 0.4$. For $p > p_0$, the transition becomes discontinuous and the replica symmetric transition can only provide an upper bound for the true $\alpha_c(2+p)$. The RS theory predicts correctly a discontinuous appearance of a finite fraction of fully constrained variables which jumps from 0 to γ_c when crossing the threshold $\alpha_c(2 + p)$. However, the values of both $\gamma_c(2 + p)$ and α_c are overestimated as in the 3-SAT case.

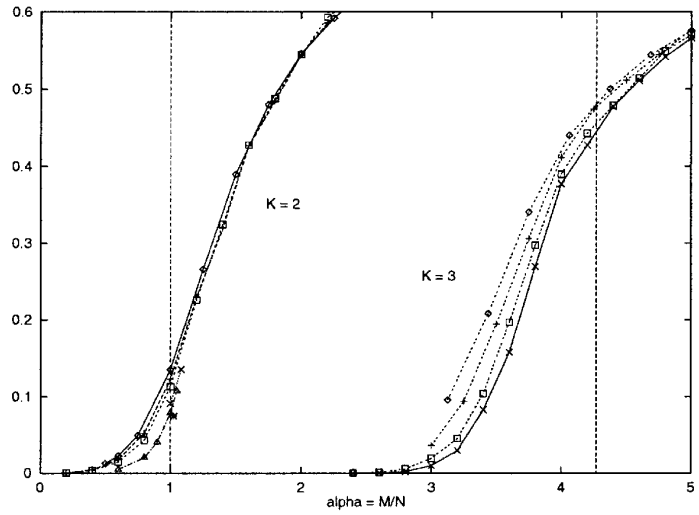

Figure 3: Numerical estimates of the value of the backbone order parameter in 2-SAT and 3-SAT. The data is from systems of sizes up to $N = 500$ variables for 2-SAT and $N = 30$ for 3-SAT.

4 Connection between computational complexity and phase transitions

The analytical and numerical results of K-SAT problem presented in the previous secction suggest that there is a connection between the order of the phase transition and the complexity of the computational problem. However, the recent results [9] give a counterexample to this interpretation. For 1-in-K SAT, K$\geq$ 3, (NP-complete problem) it has been obtained the exact location of the threshold $\alpha_c(1 - in - K) = 1/\binom{K}{2}$. In addition, it turns out that γ behaves continouusly at the threshold. This means that there is no direct connection between the order of phase transitions and computational costs. It would be interesting to compare also the behaviour of the entropy. Would it also be continuous as for K-SAT?

5 NK landscapes

In [10] it has been analyzed two random models for the decision problem of NK landscapes from the perspective of threshold phenomena and phase transitions. An NK landscape $f(x) = \sum_{i=1}^{n} f_i(x_i, \Pi(x_i))$, is a realvalued function defined on binary strings $x = (x_1, ..., x_n) \in \{0,1\}^n$. Each local fitness function f_i depends on the main variable x_i and on its neighbourhood $\Pi(x_i) \subset \{x_1, ..., x_n\} \backslash \{x_i\}$, $k = |\Pi(x_i)|$. We restrict to random neighbourhoods. The corresponding NP-complete ($k \geq 2$) problem is 'Is the optimum of $f(x)$ equal to n?' Now one can distinguish two phases: 1) soluble phase (the answer is 'yes'), 2) insoluble phase (the answer is 'no').

The uniform probability model $\overline{N}(n,k,p)$ is defined as follows: $f_i(y) = 0$ with probability p and $f_i(y) = 1$ with probability $1-p$, where $y \in Dom(f_i) = \{0,1\}^{k+1}$. Now, the control parameter (like α in K-SAT) is p.

In the fixed ratio model $N(n,k,z)$ the control parameter z takes on values from $[0, 2^{k+1}]$.

- If z is integer specify f_i by choosing randomly without replacement z tuples of possible assignments $Y = (y_1, ..., y_z)$ from $Dom(f_i) = \{0,1\}^{k+1}$ and defining $f_i(y) = 0$ if $y \in Y$ and $f_i(y) = 1$ otherwise.
- For a non-integer $z = (1-\alpha)[z] + \alpha[z+1]$ choose randomly $[(1-\alpha)n]$ local fitness functions and determine their values according to $N(n,k,[z])$. The rest of f_i determine according to $N(n,k,[z]+1)$.

In the case of $\overline{N}(n,k,p)$ it can be shown [10] that a random instance of $\overline{N}(n,k,p)$ can be solved in polynomial time. For $N(n,2,z)$ it has been analytically determined upper bounds for z_c. It turns out that for $z > 2.837$ the fixed ratio model $N(n,2,z)$ is asymptotically polynomially solvable [6]. According to experiments $N(n,2,z)$ is also easy around and below the threshold.

The results given above do not contradict the fact that for $k \geq 2$ the discrete NK landscape is NP-complete. The results simply show that for restricted k the instances generated are easy with high probability.

[6] This result is based on showing that asymptotically $N(n,2,z)$ contains an unsatisfiable 2-SAT sub-problem with probability 1 for any $z > 2.837$

References

[1] O. C. Martin, R. Monasson, R. Zecchina, Statistical mechanics methods and phase transitions in optimization problems, cond-mat/0104428.

[2] A. Goerdt, A threshold for unsatisfiability, J. Comput. System Sci. 53 (1996) 469.

[3] V. Chvàtal, B. Reed, Mick gets some (the odds are on his side), in: Proc, 33rd IEEE Symp. on Foundations of Computer Science, 1992, p. 620

[4] S. Kirkpatrick, B. Selman, Critical behaviour in the satisfiability of random boolean expressions, Science 264 (1994) 1297.

[5] R. Monasson, R. Zecchina, Statistical mechanics of the random K-Sat problem, Phys. Rev. E 56 (1997) 1357-1361.

[6] R. Monasson, R. Zecchina, Entropy of the K-satisfiability problem, Phys. Rev. Lett. 76 (1996) 3881-3885.

[7] G. Biroli, R. Monasson, M. Weigt, A variational description of the ground stste structure in random satisfiability problems, Euro. Phys. J. B 14 (2000) 551.

[8] R. Monasson, R. Zecchina, S. Kirkpatrick, B.Selman, L. Troyansky, Determining computational complexity from characteristic 'phase transitions', Nature 400 (1999) 133-137.

[9] D. Achlioptas, A. Chtcherba, G. Istrate, C. Moore, The Phase Transition in 1-in-k SAT and NAE 3-SAT, Proceedings of the Twelfth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA-01), pp. 721-722, ACM Press, January 7-9 2001.

[10] Y. Gao, J. Culberson, The Phase Transition in NK Landscapes is Easy, in Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2001), pp. 1227-1234, Morgan Kaufmann, 7-11 July 2001.