J. Phys. A: Math. Gen. **31** (1998) 9209–9217. Printed in the UK PII: S0305-4470(98)96172-8

Tricritical points in random combinatorics: the $(2+p)$-SAT case

Rémi Monasson†§ and Riccardo Zecchina‡‖

† CNRS, Laboratoire de Physique Théorique de l'ENS, 24 rue Lhomond, 75231 Paris Cedex 05, France

‡ International Centre for Theoretical Physics, Strada Costiera 11, PO Box 586, 34100 Trieste, Italy

Received 22 July 1998

Abstract. The $(2+p)$-satisfiability (SAT) problem interpolates between different classes of complexity theory and is thought to be of basic interest in understanding the onset of typical case complexity in random combinatorics. In this paper, a tricritical point in the phase diagram of the random $(2+p)$-SAT problem is analytically computed using the replica approach and found to lie in the range $\frac{2}{5} \leqslant p_0 \leqslant 0.416$. These bounds on p_0 are in agreement with previous numerical simulations and rigorous results.

1. Introduction

The satisfiability (SAT) problem [1] is the prototype of NP-complete combinatorial decision problems arising in theoretical computer science. Such decision problems are, by definition, the most difficult problems solvable in polynomial time by some ideal non-deterministic algorithm [1]. In practice, however, real algorithms may drastically change their performance depending on whether the instances of the problem are highly constrained or not. Therefore, the worst-case classification on which complexity theory is founded does not necessarily capture the behaviour of search algorithms in specific applications. For example, random instances of NP-complete decision problems undergo a dramatic change in the median time required for their solution when the instances are generated at the boundary of a critical region in the parameter space (for an introduction to these issues, see [2]).

A paradigm for such a behaviour is provided by the random K-satisfiability (K-SAT) problem. Briefly speaking, one is given N Boolean variables and a set of M clauses to be satisfied simultaneously. A clause refers to a logical constraint on K Boolean variables, randomly chosen among the N ones. For large instances ($M, N \to \infty$), K-SAT exhibits a striking threshold phenomenon as a function of the intensive ratio $\alpha = M/N$. Numerical simulations show that the probability of finding an assignment of the Boolean variables satisfying all clauses falls abruptly from one to zero when α crosses a critical value $\alpha_c(K)$ of the number of clauses per variable [3]. This scenario is rigorously established in the (polynomial) $K = 2$ case, where $\alpha_c(2) = 1$ [4]. For $K \geqslant 3$, much less is known; $K(\geqslant 3)$-SAT belongs to the NP-complete class, which means that running times of search algorithms

§ E-mail address: monasson@lpt.ens.fr

‖ E-mail address: zecchina@ictp.trieste.it

0305-4470/98/469209+09$19.50

are thought to scale exponentially in N when the problem instances are critically constrained. Recent numerical works have provided an estimate for $\alpha_c(3) \simeq 4.2$–4.3 [3].

A statistical mechanics approach has been attempted to gain insight into the K-SAT problem [5–7]. These studies rely on the correspondence between solutions and ground states of diluted spin-glass-like energy-cost functions. Threshold phenomena therefore correspond to zero temperature critical points in the phase diagram of the associated spin-glass model. Replica symmetric (RS) theory gives the correct value of the threshold for $K = 2$ but fails in predicting the critical α_c for $K \geqslant 3$ [6, 7]. This stems from the nature of the transition taking place at α_c, which is continuous for $K = 2$ and appears discontinuous when $K \geqslant 3$. In the latter case, the precise location of the critical point for the first-order transition would require an appropriate replica symmetry-breaking (RSB) scheme. For interacting models with finite connectivity, the latter issue is still an open problem in many respects [8].

Recently [9], it has been suggested that the particular nature—continuous or discontinuous—of the phase transition taking place at the threshold could be strictly connected with the appearance of computationally hard instances, and hence to the onset of exponential regimes in search algorithms [10]. Numerical studies on the so-called $(2+p)$-SAT problem [9] (that smoothly interpolates between 2-SAT ($p = 0$) and 3-SAT ($p = 1$) [7]) strongly support this statement. It follows that interest in the precise analytical localization of discontinuous transitions in random SAT models goes far beyond the purely technical aspects of the replica formalism.

In this paper we present the analytical calculation of the tricritical point p_0 of the $(2+p)$-SAT model, separating second-order phase transitions ($0 \leqslant p < p_0$) from first-order ones ($p_0 < p \leqslant 1$). In section 2, we recall the definition of the $(2+p)$-SAT model and the main steps of the statistical physics analysis are exposed in section 3. In section 4, we study the critical region and establish the self-consistent equations fulfilled by the order parameter at threshold. We analyse these equations and show that $\frac{2}{5} \leqslant p_0 \leqslant 0.416$. In conclusion, we underline the agreement between our result and another recent mathematical study of the $(2+p)$-SAT model.

2. Presentation of the $(2+p)$-SAT model

The $(2+p)$-SAT model is a mixed version of 2-SAT and 3-SAT including $(1-p)M$ (resp. pM) clauses constraining two (resp. three) Boolean variables [7].

To start with, we consider a set of N Boolean variables $\{x_i = 0, 1\}_{i=1,\ldots,N}$. We first randomly choose two among the N possible indices i and then, for each of them, a literal z_i that is the corresponding x_i or its negation $\bar{x}_i$ with equal probabilities one half. A clause C is the logical OR of the two previously chosen literals: that is C will be true (or satisfied) if and only if at least one literal is true. Next, we repeat this process to obtain $(1-p)M$ independently chosen clauses $\{C_\ell\}_{\ell=1,\ldots,(1-p)M}$ and ask for all of them to be true at the same time: i.e. we take the logical AND of the M clauses thus obtaining a Boolean expression in the so-called conjunctive normal form (CNF). The resulting 2-CNF formula F_2 may be written as

$$F_2 = \bigwedge_{\ell=1}^{(1-p)M} C_\ell = \bigwedge_{\ell=1}^{(1-p)M} \left(\bigvee_{i=1}^{2} z_i^{(\ell)} \right) \tag{1}$$

where $\bigwedge$ and $\bigvee$ stand for the logical AND and OR operations respectively.

Then, using the above prescription, we generate a 3-CNF, hereafter called F_3 including pM clauses of length 3. The resulting Boolean formula F that we shall analyse, reads

$F = F_2 \wedge F_3$. A logical assignment of the $\{x_i\}$'s satisfying all clauses, that is evaluating F to be true, is called a solution of the satisfiability problem. If no such assignment exists, F is said to be unsatisfiable. It is worth noting that as far as the complexity classification of this problem is concerned, for any $p > 0$ any instance of the model contains a 3-CNF subformula, therefore proving that the problem itself belongs to the NP-complete class.

This model has the usual threshold behaviour of K-SAT instances [9, 11] at a critical ratio $M/N = \alpha_c(p)$, with $\alpha_c(0) = 1$ and $\alpha_c(1) = \alpha_c^{3\text{sat}} \simeq 4.2$–$4.3$. The critical ratio is obviously bounded from above by $\alpha_c(p) \leqslant 1/(1-p)$, obtained from the requirement that F_2 is almost certainly satisfiable. We shall show in the following that

$$\alpha_c(p) = \frac{1}{1-p} \qquad (0 \leqslant p < p_0) \tag{2}$$

i.e. that the upper bound is reached when p is smaller than a value p_0 lying in the range

$$0.4 \leqslant p_0 \leqslant 0.416. \tag{3}$$

Most remarkably, since an earlier presentation of our result [9], a rigorous proof of the equality (2) has been derived for $p \leqslant \frac{2}{5}$ based on an analysis of the so-called unit clause algorithm [11].

3. Statistical mechanics analysis

3.1. The energy-cost function

The above mixed random SAT problem can be mapped onto a diluted spin energy-cost function upon introducing the spin variables, $S_i = 1$ if the Boolean variable x_i is true, $S_i = -1$ if x_i is false, and by taking into account the clauses through an $M \times N$ random matrix C where $C_{\ell,i} = -1$ (respectively $+1$) if clause C_l contains $\bar{x}_i$ (resp. x_i), 0 otherwise. It can easily be checked that $\sum_{i=1}^{N} C_{\ell i} S_i$ equals the number of wrong literals in clause ℓ. Then the energy-cost function

$$E[C, S] = \sum_{\ell=1}^{(1-p)M} \delta\Big[\sum_{i=1}^{N} C_{\ell i} S_i; -2\Big] + \sum_{\ell=(1-p)M+1}^{M} \delta\Big[\sum_{i=1}^{N} C_{\ell i} S_i; -3\Big] \tag{4}$$

where $\delta[.\,;.]$ denotes the Kronecker function, counts the number of violated clauses in the CNF Boolean expression F for logical assignment S. The ground state (GS) energy of the cost function (4), i.e. its minimum over S at fixed C, encodes for the existence of satisfying assignments (zero violated clauses, $E_{\text{GS}} = 0$) or, if not, for the minimum number ($E_{\text{GS}} > 0$) of violated clauses.

It is worth noting that, in addition to usual two-spin interactions that give rise to continuous phase transitions [13], the energy (4) involves three-spin interactions due to the presence of three-clauses. The latter can generate discontinuous phase transitions at sufficiently high concentration, i.e. for large p [14].

We now go on to calculate the value of the tricritical point p_0 separating the second-order phase transitions from the first-order ones on the threshold line $\alpha_c(p)$.

3.2. The average over the disorder

Resorting to the replica method for diluted spin glasses and following [7], one proceeds by computing the model 'free-energy' density at inverse temperature β, averaged over the clauses distribution $F(\beta) = -\frac{1}{\beta N}\overline{\ln Z[C]}$ where $Z[C]$ is the partition function. The overbar

denotes the average over the random clauses matrices C and is performed using the replica trick $\overline{\ln Z} = \lim_{n\to 0}(\overline{Z^n} - 1)/n$, starting from integer values of n. The typical properties of the ground state, i.e. the internal energy and the entropy, are recovered in the $\beta \to \infty$ limit.

To express the nth moment of the partition function, it is convenient to use the multilevel gas formalism proposed in [8]. The replicated theory is equivalent to a gas of N particles occupying 2^n levels labelled by n binary component vectors $\boldsymbol{\sigma} = (\sigma_1 = \pm 1, \sigma_2 = \pm 1, \ldots, \sigma_n = \pm 1)$. Calling $\rho(\boldsymbol{\sigma})$ the population, that is the fraction of particles on level $\boldsymbol{\sigma}$, the energy of the gas per particle reads after some simple algebra (detailed in the appendix),

$$E_{\rm gas}[\rho] = -\frac{\alpha}{\beta}(1-p)\ln\left[\sum_{\boldsymbol{\sigma},\boldsymbol{\tau}}\rho(\boldsymbol{\sigma})\rho(\boldsymbol{\tau})\exp\left(-\beta\sum_{a=1}^{n}\delta[\sigma_a;1]\delta[\tau_a;1]\right)\right] - \frac{\alpha}{\beta}p\ln\left[\sum_{\boldsymbol{\sigma},\boldsymbol{\tau},\boldsymbol{\omega}}\rho(\boldsymbol{\sigma})\rho(\boldsymbol{\tau})\rho(\boldsymbol{\omega})\exp\left(-\beta\sum_{a=1}^{n}\delta[\sigma_a;1]\delta[\tau_a;1]\delta[\omega_a;1]\right)\right] \quad (5)$$

with the symmetry constraint $\rho(\boldsymbol{\sigma}) = \rho(-\boldsymbol{\sigma})$. The stationary distribution ρ_s of the level populations ρ in the thermodynamic limit $N \to \infty$ is obtained by balancing the above energetic interactions and the mixing entropy (per particle) [8]

$$S_{\rm gas}[\rho] = -\sum_{\boldsymbol{\sigma}}\rho(\boldsymbol{\sigma})\ln\rho(\boldsymbol{\sigma}) \quad (6)$$

thus minimizing $E_{\rm gas}[\rho] - S_{\rm gas}[\rho]/\beta$. The dominant contribution to $\overline{Z^n}$ is then given by

$$\overline{Z^n} \simeq \exp\left(-\beta N\left[E_{\rm gas}[\rho_s] - \frac{1}{\beta}S_{\rm gas}[\rho_s]\right]\right). \quad (7)$$

The determination of the saddle point $\rho_s(\boldsymbol{\sigma})$ is very difficult in general but can be performed under some simplifying assumptions.

3.3. The replica symmetric theory

In the RS hypothesis, one looks for a stationary distribution $\rho_s(\sigma_1, \sigma_2, \ldots, \sigma_n)$ invariant under any permutation of the n replicas. Therefore, $\rho_s(\boldsymbol{\sigma})$ depends on its argument through $\sum_{a=1}^{n}\sigma_a$ only. This allows the introduction of a generating function $R(z)$,

$$\rho_s(\sigma_1, \sigma_2, \ldots, \sigma_n) = \int_{-\infty}^{\infty} {\rm d}z\, R(z)\prod_{a=1}^{n}\left(\frac{{\rm e}^{\beta z\sigma_a/2}}{{\rm e}^{\beta z/2} + {\rm e}^{-\beta z/2}}\right) \quad (8)$$

which becomes the Laplace transform of the populations ρ_s in the limit $n \to 0$ [7]. Note that, since the sum of the fractions ρ equals one, $R(z)$ is normalized to unity.

The minimization condition over ρ_s yields a self-consistent equation for the function $R(z)$. In the limits of interest $n \to 0$ and $\beta \to \infty$, this equation reads (see [7]),

$$R(z) = \int_{-\infty}^{\infty}\frac{{\rm d}u}{2\pi}\cos(uz)\exp\left\{-\alpha(1-p) + 2\alpha(1-p)\int_0^{\infty}{\rm d}z_1\, R(z_1)\cos(u\min(1,z_1)) - \tfrac{3}{4}\alpha p + 3\alpha p\int_0^{\infty}{\rm d}z_1\,{\rm d}z_2\, R(z_1)R(z_2)\cos(u\min(1,z_1,z_2))\right\}. \quad (9)$$

The interpretation of $R(z)$ is transparent within the cavity approach: it is the probability distribution of the effective fields z seen by the spins [7]. In other words, $R(z)$ accounts for the histogram $P(\langle\langle S\rangle\rangle)$ of the thermal average values of the variables through the relation $\langle\langle S_i\rangle\rangle = \tanh(\beta z_i/2)$.

4. Analysis of the critical region

4.1. The order parameter at threshold

When $\alpha < \alpha_c(p)$, that is for weakly constrained formulae, the stable solution of (9) is $R(z) = \delta(z)$ because the number of fully constrained spins in the ground state is not extensive in N [7, 9]. Let us now fix p to a small value. As discussed in the previous section, the SAT/UNSAT transition is thought to be of second order. We thus consider some small (and even) fluctuations $\mu(z) = R(z) - \delta(z)$ around the solution of the SAT phase. From (9), we find

$$\mu(z) = \int_0^\infty \mathrm{d}z_1\, \tilde{L}(z, z_1)\mu(z_1) + \int_0^\infty \mathrm{d}z_1\, \mathrm{d}z_2\, \tilde{M}(z, z_1, z_2)\mu(z_1)\mu(z_2) + \mathrm{O}(\mu^3) \tag{10}$$

for all z where

$$\begin{aligned} \tilde{L}(z, z_1) &= \alpha(1-p) \sum_{\sigma_1=\pm1} \delta(z - \sigma_1 \min(1, z_1)) \\ \tilde{M}(z, z_1, z_2) &= \tfrac{3}{2}\alpha p \sum_{\sigma_1=\pm1} \delta(z - \sigma_1 \min(1, z_1, z_2)) \\ &\quad + \tfrac{1}{2}\alpha^2(1-p)^2 \sum_{\sigma_1,\sigma_2=\pm1} \delta(z - \sigma_1 \min(1, z_1) - \sigma_2 \min(1, z_2)). \end{aligned} \tag{11}$$

Let us restrict to $z \in [0; 1[$†. The inspection of the linear term in (11) shows that the threshold is given by (2). Next, we expand around the latter by posing $\alpha = \alpha_c(p) + x$, $\mu(z) = x\ \eta(z) + \mathrm{O}(x^2)$ and obtain, when $x \to 0$,

$$0 = (1-p)\eta(z) + \int_0^\infty \mathrm{d}z_1\, \mathrm{d}z_2\, M(z, z_1, z_2)\eta(z_1)\eta(z_2) \tag{12}$$

where the kernel of the quadratic form reads

$$\begin{aligned} M(z, z_1, z_2) &= \frac{3p}{2(1-p)} \sum_{\sigma_1=\pm1} \delta(z - \sigma_1 \min(1, z_1, z_2)) \\ &\quad + \tfrac{1}{2} \sum_{\sigma_1,\sigma_2=\pm1} \delta(z - \sigma_1 \min(1, z_1) - \sigma_2 \min(1, z_2)). \end{aligned} \tag{13}$$

Note that the positivity of the probability distribution R imposes $\eta(z) \geqslant 0$ for $z \neq 0$. Furthermore, the normalization of R implies that

$$\int_{-\infty}^{\infty} \mathrm{d}z\, \eta(z) = 0. \tag{14}$$

Consequently, $\eta(z)$ includes a Dirac peak in $z = 0$ with a negative weight $-\eta_0$, $\eta_0 \geqslant 0$.

4.2. Discretization of the self-consistent equations

Within the iterative scheme for the RS solution discussed in [7], we can discretize the above equation and look for an exact solution of the form

$$\eta(z) = -\eta_0\delta(z) + \sum_{\ell\neq0} \eta_\ell \delta\left(z - \frac{\ell}{q}\right). \tag{15}$$

† Equation (9) is indeed a self-consistent constraint on $R(z)$ in this range only, see [7].

In the above equation, $1/q$ is the resolution of the effective field which eventually goes to zero. The self-consistent equations for the coefficients, η_ℓ's ($\ell = 0, 1, \ldots, q-1$) are easily obtained from (12),

$$(1-p)\eta_0 = \frac{3}{4}\frac{1-2p}{1-p}\eta_0^2 - \eta_0 \sum_{j=1}^{q-1}\eta_j + \sum_{j=1}^{q-1}\eta_j^2 + \left(\sum_{j=1}^{q-1}\eta_j\right)^2 \tag{16}$$

and, for $\ell = 1, \ldots, q-1$,

$$(1-p)\eta_\ell = \eta_\ell\left\{\eta_0 + \frac{3}{2}\frac{p}{1-p}\left[-\eta_0 + 2\sum_{j=1}^{\ell-1}\eta_j + \eta_\ell\right]\right\} - \tfrac{1}{2}\sum_{j=1}^{\ell-1}\eta_j\eta_{\ell-j} - \sum_{j=1}^{q-\ell-1}\eta_j\eta_{\ell+j} + \eta_{q-\ell}\left(\sum_{j=1}^{q-1}\eta_j - \tfrac{1}{2}\eta_0\right). \tag{17}$$

4.3. Homogeneous equations at tricriticallity

The onset of first-order transition corresponds to the smallest value of p for which $\eta(z)$ diverges. Let us call $p_0(q)$ the tricritical point for a resolution of the field $1/q$. When $q = 1$, equation (16) gives

$$\eta_0 = \frac{4(1-p)^2}{3(1-2p)} \tag{18}$$

leading to $p_0(1) = \frac{1}{2}$. When increasing q, one gets smaller and smaller values for $p_0(q)$: e.g. $p_0(2) = 0.4614$, $p_0(3) = 0.4484, \ldots$. When approaching $p_0(q)$ from below, the weights of the Dirac peaks always diverge according to

$$\eta_\ell(p) \simeq \frac{\Omega_\ell}{p_0(q) - p} \qquad p \to p_0(q)^- \tag{19}$$

as can be explicitly checked with (18) for $q = 1$ and $\ell = 0$. Therefore, the amplitudes Ω_ℓ have to satisfy the *homogeneous* versions of equations (16) and (17),

$$0 = \frac{3}{4}\frac{1-2p}{1-p}\Omega_0^2 - \Omega_0 \sum_{j=1}^{q-1}\Omega_j + \sum_{j=1}^{q-1}\Omega_j^2 + \left(\sum_{j=1}^{q-1}\Omega_j\right)^2 \tag{20}$$

and, for $\ell = 1, \ldots, q-1$,

$$0 = \Omega_\ell\left\{\Omega_0 + \frac{3}{2}\frac{p}{1-p}\left[-\Omega_0 + 2\sum_{j=1}^{\ell-1}\Omega_j + \Omega_\ell\right]\right\} - \tfrac{1}{2}\sum_{j=1}^{\ell-1}\Omega_j\Omega_{\ell-j} - \sum_{j=1}^{q-\ell-1}\Omega_j\Omega_{\ell+j} + \Omega_{q-\ell}\left(\sum_{j=1}^{q-1}\Omega_j - \tfrac{1}{2}\Omega_0\right). \tag{21}$$

The tricritical point p_0 is the smallest value of p for which the quadratic forms in (20) and (21) have a non-zero solution Ω_ℓ. In the above equations, we can choose $\Omega_0 = 1$ arbitrarily and we are left with q coupled equations for p_0 and the $q-1$ amplitudes Ω_ℓ, $\ell = 1, \ldots, q-1$.

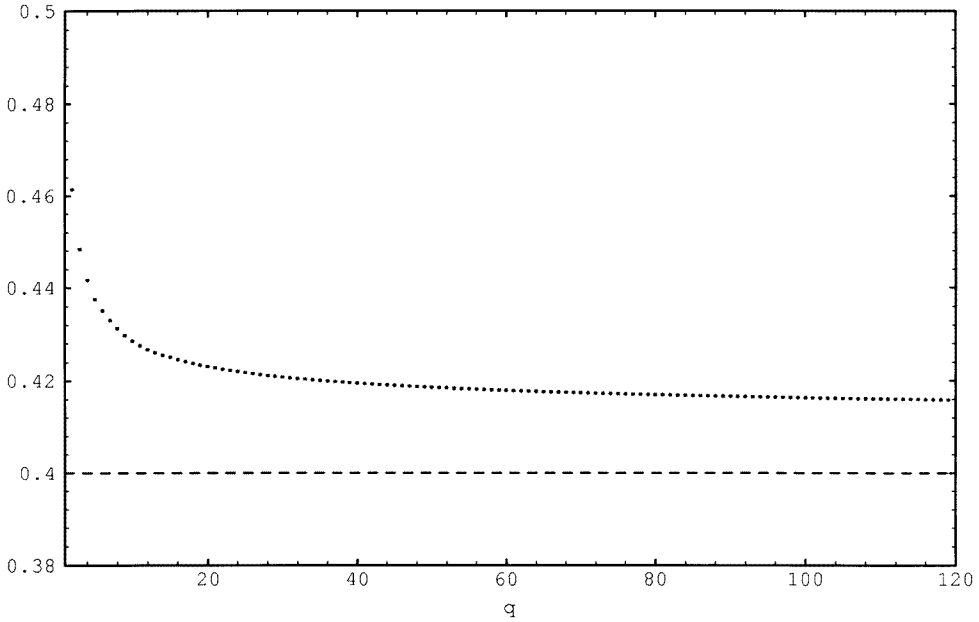

Figure 1. Plot of $p_0(q)$ versus q. The broken line is the lower bound $p_0 = \frac{2}{5}$.

4.4. Lower bound to the tricritical point

We now focus upon the self-consistent equation (20) that we rewrite as follows:

$$\frac{5p-2}{4(1-p)}\Omega_0^2 = \left(\frac{1}{2}\Omega_0 - \sum_{j=1}^{q-1}\Omega_j\right)^2 + \sum_{j=1}^{q-1}\Omega_j^2 \tag{22}$$

from which the lower bound $\frac{2}{5} \leqslant p_0$ is immediately derived. Furthermore, this lower bound can be reached if and only if $\eta(z)$ at the tricritical point vanishes outside of the interval $]-1;1[$ and there is no Dirac distribution are present in the continous limit $q \to \infty$.

4.5. Upper bound to the tricritical point

If $\{\tilde{\Omega}_\ell\}$ is a solution of equations (20) and (21) for a given pair of parameters $(p, q = \tilde{q})$, then so also is $\{\Omega_{2\ell} = \tilde{\Omega}_\ell, \Omega_{2\ell+1} = 0\}$ for $(p, q = 2\tilde{q})$. Thus, $p_0(q) \geqslant p_0(2q) \geqslant \cdots \geqslant p_0$ for any finite q, defining a sequence of more and more refined upper bounds to p_0. We have then obtained the numerical values of $p_0(q)$ for $q = 1, \ldots, 120$. It appears that $p_0(q)$ indeed decreases with q and equals 0.4158 for $q = 120$, giving a numerical upper bound to p_0.

The convergence of $p_0(q)$ down to its limit value p_0 is very slow and seems to display some power-law effects (see figure 1). At first sight, the numerical prediction for p_0 is close to 0.41, a value close to but higher than the lower bound $\frac{2}{5}$.

5. Conclusion

To conclude, a few observations are in order. The above results have been derived within an iterative RS scheme allowing for more and more refined effective field resolutions. With the

simplest choice of integer fields, the value of p_0 would have been $\frac{1}{2}$, a wrong result which tells us that there must exist other non-integer contibutions to $R(z)$. The appearance of non-integer effective fields has recently been shown to reflect the existence of RSB. Further work will be necessary to elucidate the role of RSB effects on the structure of the solutions (in principle, even the calculation of p_0 could be affected). The rigorous results discussed in [11] show that the RS solution is exact at least up to $p < \frac{2}{5}$. Such probabilistic results are based on the convergency analysis of a simple algorithm which proceeds by successive simplifications of the Boolean formula originated by fixing at random one variable at a time. In [11] it is shown that for $\alpha < \alpha_c(p)$ and $p < \frac{2}{5}$, the above algorithm has a finite probability of finding a satisfying assignment and hence the starting formula has to be satisfiable with probability one in the limit $N \to \infty$. For $p < \frac{2}{5}$ the three-clauses are ineffective even for a rather trivial 'dynamical process' like the algorithm mentioned. Such a result is indeed consistent with the idea that the nature of the phase transition taking place at $\alpha_c(p)$ does not change at least up to $p < \frac{2}{5}$. In the case $p_0 > \frac{2}{5}$, as suggested by the RS solution, it would be of interest to understand how one should modify the algorithm in order to recover the statistical mechanics result.

Let us conclude by noting that from a physical point of view, the nature of the transition manifests itself through the appearance of a finite fraction of completely constrained variables when crossing the threshold [7, 12]. Above p_0, this fraction discontinuously blows up at α_c. The narrow correspondence between this fact and the onset of computational complexity shown by simulations [9] suggests that the underlying mechanisms causing the increase of the typical computational search cost could be related to the fact that search algorithms have to find the precise values of an $O(N)$ number of Boolean variables through extensive enumeration.

Acknowledgments

We thank D Achilioptas, O Dubois, S Kirkpatrick and B Selman for motivating discussions. We also thank the Microsoft Theory Group for hospitality.

Appendix. Calculation of the effective gas energy

Consider n Boolean assignments S^a, where $a = 1, \ldots, n$, each comprised of N binary spins. The replica method requires the computation of the average product of their Gibbs weights corresponding to energy (4).

$$z[S^a] = \overline{\exp\left(-\beta \sum_{a=1}^{n} E[C, S^a] \right)} \tag{A1}$$

factorizes over the sets of two- and three-clauses due to the absence of any correlation in their probability distribution. Thus,

$$z[S^a] = (\zeta_2[S^a])^{(1-p)M} (\zeta_3[S^a])^{pM}. \tag{A2}$$

The single-clause factors in the above formula are defined by (for $K = 2, 3$)

$$\zeta_K[S^a] = \overline{\exp\left(-\beta \sum_{a=1}^{n} \delta\left[\sum_{i=1}^{N} C_i S_i^a; -K \right] \right)} \tag{A3}$$

where the bar denotes the unbiased average over the set of $2^K\binom{N}{K}$ vectors of N components $C_i = 0, \pm 1$ and of squared norm equal to K. Using the identity,

$$\delta\left[\sum_{i=1}^{N} C_i S_i^a; -K\right] = \prod_{i/C_i\neq 0} \delta[S_i^a; -C_i] \tag{A4}$$

we carry out the averaging over in disorder in (A3) to obtain

$$\zeta_K[S^a] = \frac{1}{2^K} \sum_{C_1,\ldots,C_K=\pm 1} \frac{1}{N^K} \sum_{i_1,\ldots,i_K=1}^{N} \exp\left\{ -\beta \sum_{a=1}^{n} \prod_{\ell=1}^{K} \delta[S_{i_\ell}^a; -C_\ell]\right\} \tag{A5}$$

to the largest order in N. Defining $\rho(\boldsymbol{\sigma})$ as the fraction of spins $(S_i^1, \ldots, S_i^n)$ equal to $(\sigma^1, \ldots, \sigma^n)$ [8], we rewrite $\zeta_K[S^a] = \zeta_K[\rho]$ with

$$\zeta_K[\rho] = \frac{1}{2^K} \sum_{C_1,\ldots,C_K=\pm 1} \sum_{\boldsymbol{\sigma}_1,\ldots,\boldsymbol{\sigma}_K} \rho(-C_1\boldsymbol{\sigma}_1)\ldots\rho(-C_K\boldsymbol{\sigma}_K) \exp\left\{ -\beta \sum_{a=1}^{n} \prod_{\ell=1}^{K} \delta[\sigma_\ell^a; 1]\right\}. \tag{A6}$$

Note that $\rho(\boldsymbol{\sigma}) = \rho(-\boldsymbol{\sigma})$ due to the even distribution of the disorder C. The final expression of the effective gas energy per particle, defined as $E_{\text{gas}}[\rho] = -\log z[S^a]/\beta N$ is given in (5).

References

[1] Garey M and Johnson D S 1979 *Computers and Intractability; A Guide to the Theory of NP-completeness* (San Francisco, CA: Freeman)

[2] Hayes B 1997 *Am. Sci.* **85** 108

[3] Mitchell D, Selman B and Levesque H J 1992 *Proc. AAAI-92* p 456
Kirkpatrick S and Selman B 1994 *Science* **264** 1297

[4] Goerdt A 1992 *Proc. 7th Int. Symp. on Mathematical Foundations of Computer Science* p 264
Goerdt A 1996 *J. Comput. Syst. Sci.* **53** 469
Chvàtal V and Reed B 1992 *Proc. 33rd IEEE Symp. on Foundations of Computer Science* p 620

[5] Kirkpatrick S, Gyorgyi G, Tishby N and Troyansky L 1994 *Advances in Neural Information Processing Systems* vol 6, ed J Cowan, G Tesauro and J Alspector (San Mateo, CA: Morgan Kaufmann)

[6] Monasson R and Zecchina R 1996 *Phys. Rev. Lett.* **76** 3881

[7] Monasson R and Zecchina R 1997 *Phys. Rev.* E **56** 1357

[8] Monasson R 1998 *Phil. Mag.* B **77** 1515
Monasson R 1998 *J. Phys. A: Math. Gen.* **31** 513

[9] Monasson R, Zecchina R, Kirkpatrick S, Selmann B and Troyansky L 1996 *Phase Transition and Search Cost in the* $2 + p$ *SAT Problem (Proc. PhysComp96)* ed T Toffoli, M Biafore and J Leão (Boston, MA) p 229

[10] Hogg T, Huberman B A and Williams C (ed) 1996 *Artif. Intell.* **81**

[11] Achilioptas D, Kirousis L M, Kranakis E and Krinzac D 1997 Rigorous results for $(2 + p)$-SAT *Proc. RALCOM 97* p 1

[12] Gliozzi A S and Zecchina R On the structure of the space of solutions of P and NP-complete random decision problems *Preprint*

[13] Sherrington D and Kirkpatrick S 1975 *Phys. Rev. Lett.* **35** 1792

[14] Gardner E 1985 *Nucl. Phys.* B **257** 747