Cluster expansions in dilute systems: Applications to satisfiability problems and spin glasses

Guilhem Semerjian[1,*] and Leticia F. Cugliandolo[1,2,†]

[1]*Laboratoire de Physique Théorique de l'École Normale Supérieure, 24 rue Lhomond, 75231 Paris Cedex 05, France*
[2]*Laboratoire de Physique Théorique et Hautes Énergies, Jussieu, 1er étage, Tour 16, 4 Place Jussieu, 75252 Paris Cedex 05, France*

We develop a systematic cluster expansion for dilute systems in the highly dilute phase. We first apply it to the calculation of the entropy of the K-satisfiability problem in the satisfiable phase. We derive a series expansion in the control parameter, the average connectivity, that is identical to the one obtained by using the replica approach with a replica symmetric (RS) ansatz, when the order parameter is calculated via a perturbative expansion in the control parameter. As a second application we compute the free energy of the Viana-Bray model in the paramagnetic phase. The cluster expansion allows one to compute finite-size corrections in a simple manner, and these are particularly important in optimization problems. Importantly enough, these calculations prove the exactness of the RS ansatz below the percolation threshold, and might require its revision between this and the easy-to-hard transition.

I. INTRODUCTION

Very few analytical tools have been successfully employed to study disordered systems beyond mean field. Mainly, one can mention the functional renormalization group analysis [1], high-temperature expansions of finite dimensional systems [2], expansion in the concentration of disordered models defined on finite dimensional lattices [3], and expansions around mean-field theories [4]. The replica method has been used to study dilute spin-glass models [5–9] and, even if it allows one to obtain a number of analytical results, it has been particularly difficult to implement when applied to dilute systems. It is then desirable to develop other analytical methods to treat these problems, at least in their simplest phase. In this paper we investigate an independent analytical approach that is based on a cluster expansion. It allows one to compute several "additive" quantities of interest in dilute systems such as the energy density, the entropy, the free energy, etc. We shall apply this tool to the study of two standard problems, with definitions recalled below, that are random K satisfiability (K-sat) [10] and the Viana-Bray dilute spin glass [5]. The method is similar to one of the techniques used by Weigt and Hartmann [11] in their study of the vertex-cover problem on a random graph. The application to other dilute systems is straightforward. Some of the advantages of this method with respect to replicas are the following: it allows one to compute the corrections to the thermodynamic limit in a simple way; it allows one to pinpoint a possible limitation of the replica symmetric (RS) ansatz in the satisfiable and paramagnetic (PM) phases of dilute disordered systems; and, not less importantly, it can be straightforwardly adapted to study the evolution of some of the algorithms developed to analyze K-sat numerically [12].

The paper is organized as follows. In Sec. II we define the random K-sat problem and recall its main properties. In Sec. III we define the clusters, as well as several useful notions associated with them, and we introduce the cluster expansion. Section IV is devoted to an explicit calculation of the entropy of K-sat. The finite N corrections are also described. In Sec. V we discuss the interplay between the percolation and the easy-to-hard transition. We underline the consequences of this calculation in regard to the validity of the RS ansatz in the satisfiable and PM phases. As an application to a physical system, we discuss the calculation of the paramagnetic PM free energy of the Viana-Bray [5] dilute spin glass in Sec. VI. Finally, in Sec. VII we present our conclusions.

II. K SATISFIABILITY

The theory of complexity has been developed to characterize *worst*-case instances of hard computational problems [13]. A classification scheme, according to the time needed to find solutions with the best performing algorithms, or to prove that a problem is not solvable, is one of the outcomes of these studies. Of particular importance is the problem of K satisfiability [10,14,15] (K-sat), that has been used as a testing ground for these theories.

However, it was recently realized that in many interesting cases in computer science, it is more relevant to determine the properties of *typical*, and not worst, realizations of a given problem [16]. *Random K-sat*, defined as the ensemble of randomly generated instances of K-sat, is the paradigm and the goal is now to predict the behavior of a typical element of the ensemble.

The relation between phase transitions, or threshold phenomena, and intractability in random combinatorial problems, has been stressed by several authors [17]. Problems that are very hard to solve in the worst case are not so in the typical case, unless the control parameter takes values within a finite interval that defines the critical region. Away from the critical region, simple algorithms are capable of finding a solution, or showing that there is no solution, in polynomial time. Random K-sat has a well-defined threshold phenomenon.

Random K-sat, as well as other random optimization problems, can be mapped onto disordered spin models. The mapping is done by associating the cost function in the optimization problem to an energy density in the physical sys-

*Email address: guilhem@lpt.ens.fr
†Email address: leticia@lpt.ens.fr

tem [8]. Consequently, the random character associated with the choice of different instances in the optimization problem translates into quenched disordered interactions in the physical system. The most interesting optimization problems like K-sat become spin-glass models of a particularly difficult type, where each spin interacts with a finite fraction of other spins in the sample. These are called dilute spin glasses, and they are interesting *per se* since they appear as a case of intermediate difficulty between solvable mean-field spin glasses and realistic finite-dimensional ones.

The quest for a threshold value of the control parameter then becomes a search for a phase transition. Thus all tools developed to treat disordered physical systems in statistical mechanics [18] can be adapted to study random optimization problems. In the context of random K-sat, two main techniques have been used so far: the replica approach in the thermodynamic limit [8,9,19,20] and numerical simulations complemented by finite-size scaling when the number of variables remains finite [21]. The same two techniques are used in the study of dilute spin glasses.

Random K-sat is defined as follows. Consider N Boolean variables, $\{x_1, \ldots, x_N\}$, that can take two logical values: x_i is either true or false, for each i. First, choose K indices from the set of N elements, $i=1, \ldots, N$. Second, assign to each of these indices the *literal* x_j, or its negation $\bar{x}_j$, with equal probability $p=1/2$. Third, construct a *clause* C_1 as the logical "or" ($\vee$) of the K previously determined literals. If $K=3$ and $N=10$, a possible clause is $x_1 \vee \bar{x}_5 \vee x_7$. New clauses are generated in an identical manner, independently of the previous ones. One usually calls M the total number of clauses. A *formula F* is the logical "and" ($\wedge$) of M such clauses. It reads

$$F = \wedge_{l=1}^{M} C_l = \wedge_{l=1}^{M} (\vee_{i=1}^{K} z_i^l), \tag{1}$$

where $z_i^l \in \{x_1, \bar{x}_1, \ldots, x_N, \bar{x}_N\}$. A solution, if it exists, is an assignment of the N variables that satisfies F, that is to say, for which all clauses are verified simultaneously.

Note that in the process of generation of a clause, two random processes intervene. In the first process, one selects the variables; in the second process, once the variables have been chosen, one determines the requirements that will be imposed on them. We shall later take advantage of this two-step process to perform an average over disorder in a convenient order.

It is clear that, if $M \ll N$, it will be very easy to find a solution to F. Conversely, if $M \gg N$, it will be extremely difficult to satisfy all requirements simultaneously. Indeed, a well-defined critical value $\alpha_c(K)$ of the parameter $\alpha \equiv M/N$ appears when $M \to \infty$ and $N \to \infty$, with their ratio α kept fixed. This limit corresponds to a thermodynamic limit, in the physical language. A threshold phenomenon, reminiscent of a phase transition, is observed: for $\alpha < \alpha_c(K)$, all formulas have at least one solution with probability 1, whereas for $\alpha > \alpha_c(K)$ any formula has no solution with probability 1.

Different values of K lead to different critical behaviors. When $K=1$, the model is unsatisfiable for all finite values of

α, i.e., $\alpha_c(K=1)=0$. When $K=2$, K-sat has a continuous phase transition at $\alpha_c(K=2)=1$. This is a rigorous result proven by using a mapping on a directed graph problem [14]. For $K \geqslant 3$ only numerical estimates for $\alpha_c(K \geqslant 3)$ and approximate results obtained with the replica method are available [8]; these yield $\alpha_c(K=3) \sim 4.2$.

The replica method is a powerful tool of statistical mechanics that allows one to compute the statistical properties of a disordered physical problem in equilibrium with a thermal environment. In order to use it to study optimization problems in general, and K-sat in particular, one first maps the optimization problem onto a statistical mechanics model. In the case of K-sat, the physical model is a spin-glass model with dilute interactions of random sign. Indeed, a natural representation of any K-sat formula is obtained by introducing an $M \times N$ matrix C_{li}, whose elements are

$$C_{li} = \begin{cases} 0 & \text{if neither } x_i \text{ nor } \bar{x}_i \in C_l \\ 1 & \text{if } x_i \in C_l \\ -1 & \text{if } \bar{x}_i \in C_l. \end{cases} \tag{2}$$

The random generation of clauses is equivalent to a uniform distribution of the matrices C_{li} that satisfy the constraints $\Sigma_i C_{li}^2 = K$, $\forall l$.

A cost function for K-sat is given by the number of unsatisfied clauses in a given formula. If one identifies the logical state where x_i is true, with a spin $S_i = 1$ and the logical state where x_i is false, with a spin $S_i = -1$, it is then easy to verify that the following expression counts the number of unsatisfied clauses,

$$E[\{C_{li}, x_i\}] = \sum_{l=1}^{M} \delta^{(K)} \left(\sum_{i=1}^{N} C_{li} S_i, -K \right), \tag{3}$$

where $\delta^{(K)}(a,b)$ is the Kronecker delta function. Using a polynomial representation of $\delta^{(K)}$, this expression can be rewritten as the total energy of a sum of dilute p spin-glass models in a random field (several values of p intervene, how many depends on the value of K) [8].

Once the energy function is identified, one introduces a fictive temperature T, then computes the average free energy with the help of the replica trick, and finally takes the limit $T \to 0$ to study the ground state properties of the physical model. This gives access to quantities such as, for example, the average entropy of the satisfiable phase. This is defined as the average over disorder of the logarithm of the number of solutions. One of the drawbacks of the use of the replica method is that an ansatz is necessary to pursue the calculation. Even in the simplest phases, the satisfiable one for K-sat, it is not obvious to show that the simplest ansatz, called replica symmetric, solves the problem exactly. Moreover, it has been proven that in the unsatisfiable phase one has to go beyond the RS ansatz and develop a replica symmetry breaking (RSB) scheme. This is indeed a very difficult task, since the order parameters for dilute systems have a much more intricate structure than for infinitely connected (IC) cases [5–7,22]. Recent progress in this direction was presented in Ref. [9].
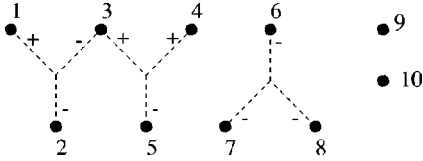
FIG. 1. A graphical representation of the formula $F = (x_1 \vee \bar{x}_2 \vee \bar{x}_3) \wedge (x_3 \vee x_4 \vee \bar{x}_5) \wedge (\bar{x}_6 \vee \bar{x}_7 \vee \bar{x}_8)$. See the text for details.

In this paper we rederive a generic expression for the entropy of K-sat using a very simple method that avoids the use of replicas. Furthermore, the method allows us to compute the finite-size corrections. Our derivation gives information about the domain of validity of Monasson and Zecchina's conjecture that the RS ansatz is exact in the satisfiable phase. We explain the expansion using the formalism of K-sat, but the line of reasoning can be applied to any dilute system in the dilute regime. In Sec. VI we shall analyze the Viana-Bray model [5] along the same lines.

III. METHOD

Let us start this section by setting the notation and defining a set of notions that will be used later. Given a formula F of K-sat, two variables x_i and x_j are called *adjacent* if there is at least one clause in F in which both x_i and x_j appear, irrespective of the fact that they are negated or not. Two variables are *connected* if and only if there is a path of adjacent variables between them. A *cluster* is a set of connected variables that are disconnected from all others. Let us label with an integer r the different clusters of the formula F, $r = 1, \ldots, \mathcal{N}_c(F)$, where $\mathcal{N}_c(F)$ is the total number of clusters in F. We shall call $n_0(F)$ the number of variables that do not belong to any cluster.

These definitions are very easy to picture. For instance, take a 3-sat problem with ten variables, $i = 1, \ldots, 10$, that is defined by the formula $F = (x_1 \vee \bar{x}_2 \vee \bar{x}_3) \wedge (x_3 \vee x_4 \vee \bar{x}_5) \wedge (\bar{x}_6 \vee \bar{x}_7 \vee \bar{x}_8)$. The variables $i = 9$ and 10 do not belong to any cluster; thus $n_0(F) = 2$. A graphical representation of each clause is very useful. We associate a point to each variable. Each clause is represented by a *star* with K legs, three in the example, with end points that represent the variables. In formula F there are two clusters, $\mathcal{N}_c = 2$, that link $i = 1, \ldots, 5$ and $i = 6$, 7, and 8, respectively. When a variable appears in two (or more) clauses, it will be shared by two stars. This is the case in the cluster on the left of Fig. 1. More complicated structures are possible, particularly when N and M are large. The assignment $x_i(\bar{x}_i)$ of the ith literal in a clause can be represented with a plus (minus) sign on its leg. These are the signs in Fig. 1. In this way, a one-to-one correspondence between formula and graphs is constructed.

When α is small, the typical cluster size is expected to be small as well, as there are much fewer clauses than variables. Indeed, for $K = 2$ this problem is the one of percolation in an infinite-dimensional space, also known as the random graph. Many properties of this model are known [23], among which the fact that for $\alpha < 1/2$ all variables belong to clusters of size at most proportional to $\ln N$ in the thermodynamic limit. When α crosses 1/2 a giant cluster containing a finite frac-tion of the variables grows continuously. For $K \geq 3$ the equivalent geometrical problem relies on the theory of hypergraphs, for which less is known. The percolation occurs at $\alpha = \alpha_p \equiv 1/[K(K-1)]$ [24]. In these two cases the system percolates much before it becomes unsatisfiable, $\alpha_p < \alpha_c$. Indeed, the appearance of a contradiction requires an intricate structure in the giant cluster.

Let us define the ground state entropy of a formula $S_{\mathrm{GS}}(F)$ as the logarithm of the number of assignments of the variables that minimize the number of violated clauses. If F is satisfiable, $S_{\mathrm{GS}}(F)$ is the logarithm of the number of solutions of F. It is clear from the cluster definition that S_{GS} is the sum of contributions of the different independent subformulas:

$$S_{\mathrm{GS}}(F) = n_0(F) \ln 2 + \sum_{r=1}^{\mathcal{N}_c(F)} S_{\mathrm{GS}}(F_r). \tag{4}$$

We are interested in the entropy averaged over the ensemble of formulas, $\bar{S}_{\mathrm{GS}}$. We shall henceforth denote ensemble averages with an overbar. As stressed in Sec. II, this average is twofold. Indeed, clusters can be separated into ensembles with the same topology, ignoring for the moment the sign assignment of the literals. Thus the averaging proceeds in two steps: one first chooses the topology of the cluster, with its associated probability, and then one averages over the two possibilities for each literal in the cluster. For a given cluster, once the latter average is performed, the entropy depends only on the topology of the cluster. This remark allows us to rewrite the average of the sum in Eq. (4) in a more convenient manner. If we introduce a new integer t that labels all possible topologies, and $n_t(F)$ and $\langle S_t \rangle$ are the numbers of t-like clusters in formula F and the average over the sign assignment of the entropy of the t-like clusters, respectively, we arrive at the following expression for the averaged entropy:

$$\overline{S}_{\mathrm{GS}} = \sum_t [n_t]\langle S_t \rangle. \tag{5}$$

Here we have included the isolated variables in the sum, associating them with the index $t = 0$, $S_0 = \ln 2$ and we denote the average number of t-like clusters with $[n_t]$.

A more convenient expression for $[n_t]$ can now be worked out. Let us call $X_t^i(F)$ the function which takes the value 1 if the variable i belongs to a t-like cluster of the formula F and 0 otherwise; let L_t be the number of variables in such a cluster. Then

$$n_t(F) = \frac{1}{L_t} \sum_i X_t^i(F), \tag{6}$$

which implies

$$\frac{1}{N}[n_t] = \frac{1}{L_t}[X_t^1] = \frac{P_t}{L_t}, \tag{7}$$

where $P_t \equiv [X_t^1]$ is the probability that a given variable belongs to a t-like cluster. Finally,

$$\frac{1}{N}\,\bar{S}_{GS}=\sum_t \frac{1}{L_t}P_t\langle S_t\rangle. \tag{8}$$

P_t and $\langle S_t\rangle$ can now be obtained using elementary combinatorial arguments and simple enumeration.

This formulation can be adapted to any quantity for which the clusters contribute additively, the free energy for instance, and to other dilute problems where there is also a decoupling in the randomness between a geometrical part and an interaction one, as in the Viana-Bray model [5].

IV. CLUSTER EXPANSION OF THE K-SAT ENTROPY

In this section we shall apply the cluster expansion to the calculation of the average entropy of random K-sat. For our present purposes $K=1$-sat is not interesting, since it is unsatisfiable for all finite values of α. We shall then start by analyzing in detail $K=2$-sat. Afterwards, we shall discuss how the approach generalizes to larger values of K.

A. $K=2$-sat in the thermodynamic limit

For a cluster of n variables connected by p distinct clauses, the probability P_t reads

$$P_t = p!\binom{M}{p}\left(\frac{2}{N(N-1)}\right)^p\left(\frac{(N-n)(N-n-1)}{N(N-1)}\right)^{M-p}$$

$$\times (n-1)!\binom{N-1}{n-1}K_t. \tag{9}$$

Let us briefly describe the origin of the factors in this equation. Each of the p clauses is chosen with probability $2/[N(N-1)]$ at each of the M steps in the formula generation process. For the variables belonging to the cluster to be disconnected from all other sites, the $M-p$ other clauses must belong to the set of the $(N-n)(N-n-1)/2$ clauses connecting the other sites. The first two factors come from the possible permutation of the p steps where the considered clauses appear. The last three factors arise from the freedom in the choice of $n-1$ sites connected to the chosen site. In particular, K_t is a symmetry factor that is equal to the number of distinct labelings of the n sites, divided by $(n-1)!$. Note that two labelings which lead to the same set of clauses are *not* distinct: for the linear three site cluster 1-2-3 and 3-2-1 correspond to the same labeling, with clauses (12) and (23). But 1-2-3 and 2-1-3 are distinct.

In the thermodynamic limit $N\rightarrow\infty$ and α fixed, and for n and p finite, this expression is proportional to N^{n-1-p} ($p \geq n-1$). It is then finite only if $p=n-1$, that is to say for treelike clusters. This justifies the choice of distinct clauses. In this limit, for $p=L_t-1$ and $n=L_t$, this expression simplifies greatly:

$$P_t = (2\alpha)^{L_t-1}e^{-2L_t\alpha}K_t. \tag{10}$$

The different clusters considered in the expansion are represented in Fig. 2. For each type, the relevant quantities, obtained by basic enumeration, are given in Table I. For instance, the average entropy of the linear three site cluster is
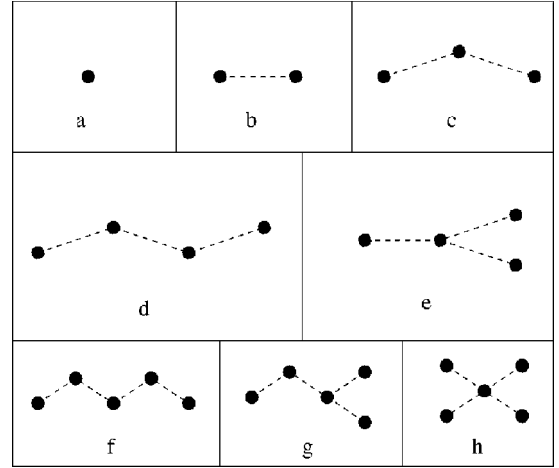

FIG. 2. Treelike clusters that contribute to $K=2$-sat.

made up of two parts: if the clauses require the same sign for the central variable, one can find five solutions of the formula; if the clauses are contradictory for the central variable, there are only four solutions.

Expanding in α up to $O(\alpha^4)$, we obtain

$$\frac{1}{N}\,\overline{S}_{GS}=\ln 2+\alpha\ln\left(\frac{3}{4}\right)+\alpha^2\ln\left(\frac{80}{81}\right)+\frac{\alpha^3}{3}\ln\left(\frac{3^{29}7^6}{5^{15}2^{28}}\right)$$

$$+\frac{\alpha^4}{12}\ln\left(\frac{2^{225}5^{160}11^{36}13^{24}17}{3^{216}7^{168}}\right). \tag{11}$$

Monasson and Zecchina obtained this series by using the replica trick, with a RS ansatz, to average the free-energy of the physical model related to 2-sat [8]. The averaged entropy follows from the averaged free-energy density, that itself depends on the probability distribution of the local fields. This quantity is determined by an integrodifferential equation that cannot be solved analytically. Monasson and Zecchina developed a perturbative solution in α that allowed them to derive a series for $\overline{S}_{SG}/N$ that coincides, up to $O(\alpha^4)$, with the one in Eq. (11). The perturbative nature of this result is now clarified from the cluster analysis. Note that we performed two expansions to obtain this series: the cluster enumeration and an expansion in powers of α of the exponentials in P_t. We shall discuss this issue further in Sec. V.

TABLE I. The contributions of the clusters in Fig. 2.

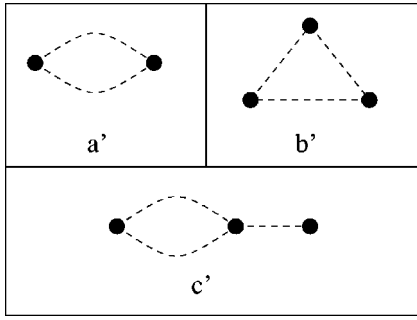| Type	L_t	K_t	$\langle S_t\rangle$
a	1	1	$\ln 2$
b	2	1	$\ln 3$
c	3	3/2	$(2\ln 2+\ln 5)/2$
d	4	2	$(3\ln 2+\ln 5+2\ln 7)/4$
e	4	2/3	$(3\ln 2+5\ln 3)/4$
f	5	5/2	$(4\ln 2+6\ln 3+\ln 5+2\ln 11+\ln 13)/8$
g	5	5/2	$(9\ln 2+2\ln 5+2\ln 7+\ln 11+\ln 13)/8$
h	5	5/24	$(13\ln 2+4\ln 5+\ln 17)/8$

FIG. 3. Loop diagrams contributing to the $1/N$ corrections.

B. Finite-size corrections to the entropy of $K=2$-sat

There are two kinds of finite-size corrections to the expansion presented in Eq. (11). On the one hand, the probability P_t of a variable belonging to a treelike cluster has $1/N$ corrections that can be simply computed from the general expression (9). On the other hand, clusters that include loops also contribute to the $1/N$ corrections.

The expansion of expression (9) up to order $1/N$ for treelike clusters, with $n=L_t$ and $p=L_t-1$, yields

$$P_t=(2\alpha)^{L_t-1}e^{-2L_t\alpha}K_t\left\{1+\frac{1}{N}\left[2L_t(L_t-1)-\alpha L_t(L_t+1)\right.\right.$$
$$\left.\left.-\frac{(L_t-1)(L_t-2)}{2}\left(1+\frac{1}{\alpha}\right)\right]\right\}. \tag{12}$$

Clusters with l loops contribute to the order $1/N^l$. Hence, if we only wish to compute the $1/N$ corrections, we can content ourselves with clusters that have only one loop. These have L_t variables and also L_t clauses. One obtains

$$P_t=\frac{1}{N}(2\alpha)^{L_t}e^{-2L_t\alpha}K_t, \tag{13}$$

with K_t defined as before, and multiplied by $1/2$ if there is a repeated clause. The one-loop clusters that we considered are represented in Fig. 3.

Including the $1/N$ corrections in Eq. (12) and the ones stemming from the new diagrams in Fig. 3 and Eq. (13) calculated with the results of Table II, the correction to $\bar{S}_{GS}/N$ reads

$$\frac{1}{N}\left[\alpha\ln\left(\frac{3^4}{2^4 5}\right)+\frac{\alpha^2}{4}\ln\left(\frac{2^{107}5^{56}}{3^{107}7^{24}}\right)\right.$$
$$\left.+\frac{\alpha^3}{2}\ln\left(\frac{3^{193}7^{156}}{2^{199}5^{141}11^{36}13^{24}17}\right)\right]. \tag{14}$$

TABLE II. The contributions of the clusters represented in Fig. 3.

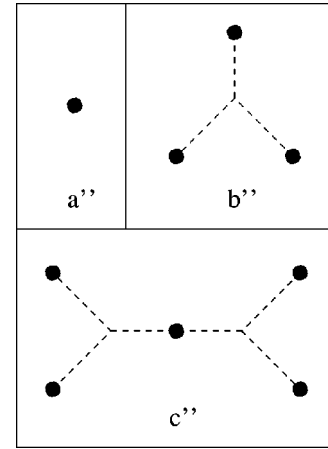| Type	L_t	K_t	$\langle S_t\rangle$
a'	2	1/2	$(3\ln 2+\ln 3)/4$
b'	3	1/2	$(9\ln 2+3\ln 3)/8$
c'	3	3/2	$(5\ln 2+4\ln 3+\ln 5)/8$

FIG. 4. Clusters yielding the leading contributions to $K=3$-sat.

This result has to be checked against exhaustive numerical evaluation of the entropy for small systems.

C. $K\geq 3$-sat in the thermodynamic limit

The method described in detail for $K=2$ can be used for any value of K. As the graphical representation and the enumeration of clusters are, however, more cumbersome than for $K=2$, we shall present less detailed results for the case $K\geq 3$.

The probability for a given variable to be present in a cluster of L_t variables that are linked by p clauses is of order 1 in the thermodynamic limit only if $p(K-1)=L_t-1$, which is the treelike condition for these hypergraphs. If this holds,

$$P_t=(\alpha K!)^p e^{-L_t\alpha K}K_t. \tag{15}$$

In Fig. 4 we have drawn the diagrams leading to the main contributions for $K=3$. In the text we give the analytic expression for general K (also see Table III).

With these values we obtain

$$\frac{1}{N}\overline{S}_{GS}=\ln 2+\alpha\ln\left(1-\frac{1}{2^K}\right)+\frac{\alpha^2 K^2}{2}\left[-\ln\left(1-\frac{1}{2^K}\right)\right.$$
$$\left.+\frac{1}{2}\ln\left(1-\frac{1}{2^K-1}\right)+\frac{1}{2}\ln\left(1-\frac{2^{K-1}-1}{2^{K-1}(2^K-1)}\right)\right]. \tag{16}$$

Again we recover the RS result of Ref. [8]. The contributions to the finite-size corrections are similar to the ones discussed for 2-sat; we obtain, at the leading order in α,

$$\frac{1}{N}\frac{\alpha K^2}{2}\left[\ln\left(1-\frac{1}{2^K}\right)-\frac{1}{2}\ln\left(1-\frac{1}{2^K-1}\right)\right.$$
$$\left.-\frac{1}{2}\ln\left(1-\frac{2^{K-1}-1}{2^{K-1}(2^K-1)}\right)\right]. \tag{17}$$

TABLE III. The contributions of the clusters represented in Fig. 4.

Type	L_t	K_t	$\langle S_t \rangle$
a''	1	1	$\ln 2$
b''	K	$K/K!$	$\ln(2^K - 1)$
c''	$2K-1$	$K^2(2K-1)/[2(K!)^2]$	$\frac{1}{2}\ln[1+2^K(2^{K-1}-1)]+\frac{1}{2}\ln[2^K(2^{K-1}-1)]$

We have shown that the perturbative analysis of the RS ansatz leads to a series in α which leading orders coincide with the ones stemming from the *finite* cluster expansion for all K. In Sec. V we shall discuss the relevance of the contributions from the infinite cluster that appears at the percolation transition $\alpha_p < \alpha_c$.

V. DISCUSSION

The domain of validity of our expansion, and of the results obtained with the replica method [8], can be enlightened by studying the percolation phenomenon in detail. The series expansion in Eq. (8) is ordered following the index t, that is directly related to the size of the clusters. The dependence on α is involved, since the coefficients P_t are α dependent via an exponential times a power. The expansion of the exponential factors in powers of α leads to a rearrangement of the series in powers of α.

The range of validity of both series is not obvious. We can start by analyzing the simpler series $\Sigma_t P_t$ that should count the total fraction of sites and be identical to 1. For $K=2$, its direct summation yields 1 for $\alpha < \alpha_p$ and $1-P$ for $\alpha > \alpha_p$, where P is the solution to $1-P=e^{-2\alpha P}$ [see Eq. (18) below]. Thus it fails above α_p, because of the emergence of a giant cluster at the percolation transition. Instead, if one expands the exponentials in powers of α, the result is $\Sigma_t P_t = 1 + 0 \times \alpha + 0 \times \alpha^2 + \cdots$. The sum yields 1, independently of α, even beyond the percolation threshold. The rearrangement in powers of α captures the correct behavior of this quantity.

One can conjecture that the rearrangement yields the exact result for all α for all quantities that depend mainly on the locally treelike structure of the percolating cluster, and only weakly on its loops, which show up only on a scale of order $\ln N$. As can be seen in Eq. (9), the exponentials in P_t arise from the requirement that the considered cluster is disconnected from the rest of the sites. Expanding the exponentials amounts basically to assuming that a subtree of the giant cluster gives the same contribution as its disconnected counterpart. The alternative signs arise from the need to avoid the double counting of the smaller clusters contained in the giant cluster.

As regards the calculation of the averaged entropy, if one could compute all the terms in Eq. (8) and sum the series, the result would be exact under the percolation threshold, $\alpha_p = 1/[K(K-1)]$. Indeed all sites belong to clusters of size at most proportional to $\ln N$ in this regime. As soon as α goes beyond α_p, the direct summation of Eq. (8) should fail.

In spite of the discussion of the next to last paragraph, we do not expect that reordering the series in powers of α gives

the correct result for the entropy. Our argument is based on the drastic influence of loops on this quantity. Let us compare the entropy of a loop and of a linear cluster of the same size, for 2-sat. There are roughly twice as many solutions for the linear cluster as for the loop, as one does not require the ending variables to be the same. The difference of entropy between the two should then be finite. As there are an extensive number of loops in the percolating cluster, one can expect a finite deviation in the average entropy per site between the result assuming a tree like structure of the giant cluster (i.e., the expansion in α of the original series) and the correct one.

We have examined these issues with the help of numerical simulations of systems with small sizes. First, we compare the averaged total entropy per degree of freedom, $\bar{S}/N$, to the value predicted by the series expansion once reordered in powers of α. In Fig. 5 we plot $\bar{S}/N$ against $1/N$ with lines points $+$ for $K=2$-sat problems with $N=16$ (50000), 20 (50000), 24 (50000), 28 (10000), 32 (15000), 36 (10000), and $\alpha_p < \alpha = 0.75 < \alpha_c$. For each sample we computed the entropy by exhaustive enumeration. The numbers between parentheses are the numbers of realizations of random instances of K-sat used to compute the averages. The accord with the analytical prediction of the truncated series expansion in the thermodynamic limit (horizontal line below), and including $1/N$ corrections (tilted line above), is very good within 0.3%. We have also computed the variance $1/N(\bar{S}^2 - \bar{S}^2)$, and checked that it is in good accord with the analytical result, $0.025\alpha^2$, that we obtained with an extension of the
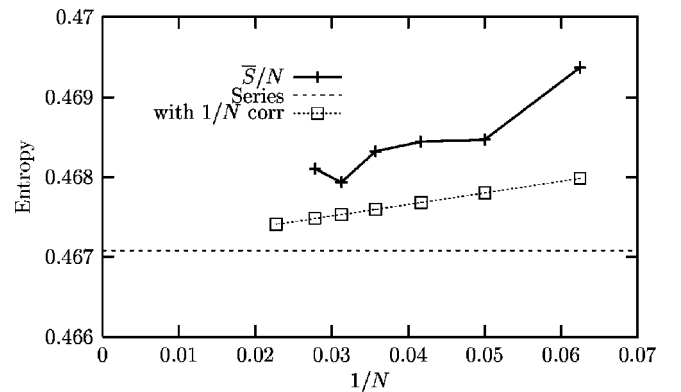

FIG. 5. The averaged entropy per site for $K=2$ and $\alpha=0.75$. The sizes are $N=16$, 20, 24, 28, 32, and 36. The lower straight line indicates the value given by the truncated series expansion found with the replica method in the thermodynamic limit [up to $O(\alpha^8)$], and the upper curve with line points includes the $1/N$ corrections for each value of N.
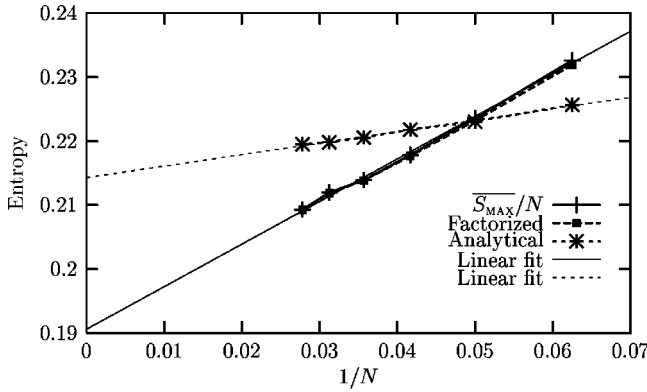
FIG. 6. The averaged entropy of the maximum cluster per variable, $\overline{S_{\mathrm{MAX}}}/N$, the factorized average $\overline{S_{\mathrm{MAX}}/L_{\mathrm{MAX}}L_{\mathrm{MAX}}}/N$ and the semianalytical prediction $P_{\mathrm{MAX}}\overline{S_{\mathrm{MAX}}/L_{\mathrm{MAX}}}$ as functions of $1/N$ for $\alpha=0.75$ and $K=2$.

FIG. 7. $K=2$. The fraction of sites belonging to the largest cluster. The asymptotic limit $N\rightarrow\infty$ is approximately reached for $N\sim10000$, where $\overline{L_{\mathrm{MAX}}}/N$ is "finite" only above the percolation threshold $\alpha_p(K=2)=0.5$. The analytical expression is given by the solution to Eq. (18) and it fits the data very accurately for $\alpha>0.6$ as soon as $N>1000$. We have verified that all other variables are distributed among finite clusters.

cluster expansion described in preceding sections.

Even if the accord between numerical results and theory is almost perfect for these small sizes, a more careful inspection of the different contributions to the total entropy shows that one could expect important deviations for increasing N. We have separately computed the contribution of the largest cluster. Figure 6 represents its study. Here we plot, with crosses, the averaged entropy of the largest cluster, per degree of freedom, $\overline{S_{\mathrm{MAX}}}/N$. Its contribution, even for these small sizes, represents roughly half of the total entropy (note that below the percolation threshold the contribution of the largest cluster is much smaller), and this does not seem to vanish in the thermodynamic limit. A simple linear fit of $\overline{S_{\mathrm{MAX}}}/N$ yields, when $N\rightarrow\infty$, the finite limit 0.1906 ± 0.0008. This may be taken as a guess of a lower limit for the contribution of the largest cluster.

In the same figure we compare the averaged entropy of the largest cluster $\overline{S_{\mathrm{MAX}}}/N$ to the "factorized" quantity $\overline{L_{\mathrm{MAX}}/N}\overline{S_{\mathrm{MAX}}/L_{\mathrm{MAX}}}$, where L_{MAX} is the number of variables in the largest cluster, that is represented with squares. For finite sizes we have shown that these two quantities coincide. The good accord between the two curves suggests that the factorization also holds when $N\rightarrow\infty$.

This observation suggests an improvement of the finite-size numerical study. If we *assume* that the factorization holds in the limit $N\rightarrow\infty$ for the percolating cluster, we can then replace the fraction of sites that belong to the largest cluster, $P_{\mathrm{MAX}}\equiv\overline{L_{\mathrm{MAX}}}/N$, by its analytical value in the thermodynamic limit. This is given by

$$1-P_{\mathrm{MAX}}=\exp\{\alpha K[(1-P_{\mathrm{MAX}})^{K-1}-1]\}. \quad (18)$$

This result is obtained using a self-consistent equation on the generating function that counts the number of sites in finite size clusters [25]. For $K=2$ and $\alpha=0.75$, $P_{\mathrm{MAX}}\sim0.5828$. Figure 7 displays P_{MAX} as a function of α for $K=2$ and different system sizes, $N=25$, 100, 1000, and 10000. One sees that the percolation transition is reached only for too large a size, $N\sim10000$, that is far beyond the largest sample for which one can compute the entropy by exhaustive enu-
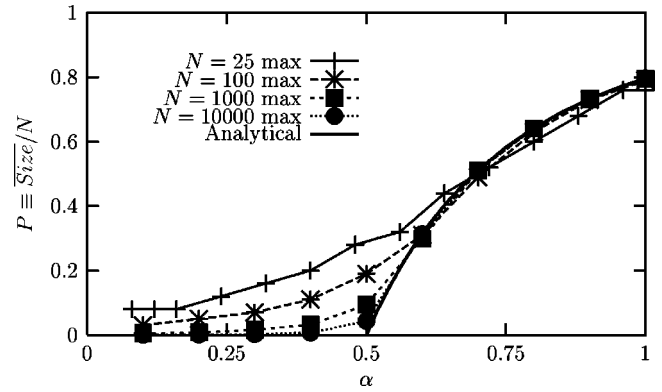
meration. Moreover, it is important to note that the approach to the asymptotic value is nonmonotonic since, for these values of N, the approach to the asymptote comes from below, while one can easily prove that $P_{\mathrm{MAX}}(N=4,\alpha=0.75)\sim0.75$.

Finally, the third curve in Fig. 6 (stars) represents the "improved" contribution of the largest cluster, $0.5828\overline{S_{\mathrm{MAX}}/L_{\mathrm{MAX}}}$. For $N>20$ the improved curve is still higher than the actual one, and a linear fit yields the limiting value $\lim_{N\rightarrow\infty}\overline{S_{\mathrm{MAX}}}/N\sim0.2143\pm0.0003$.

This numerical analysis, even if it is based on simulations of very small systems and highly speculative, suggests that the contribution of the percolating cluster is finite in the thermodyamic limit. On the analytical side, it might be possible to compute the entropy of the percolating cluster, at least for 2-sat, by taking profit of the numerous mathematical results on the structure of random graphs [23]. Such a study is necessary to settle the problem.

This picture could also explain the discrepancy between numerical studies and a rigorous bound for the value of the exponent governing the size of the scaling window of the satisfiability transition [26–28]. For the relatively small system sizes first studied, the observed exponent may have been altered by percolation, as its asymptotic regime was not yet reached [19]. At larger sizes, and in mathematical studies, the exponent is purely due to the satisfiability threshold inside the giant cluster; one might observe the true exponent for lower system sizes by studying the probability of the largest cluster to be satisfiable.

Let us briefly summarize the route followed in studies based on the replica trick to attempt to identify the effect of the percolation transition within this calculation [8]. In the RS ansatz, an intricate integral equation over the probability of local fields, $P(h)$, that is the order parameter of the system, is obtained. From $P(h)$ all thermodynamic quantities follow, including the entropy of the satisfiable phase. As an analytical resolution of the equation that determines $P(h)$ seems out of reach, it has been solved order by order in α.

This yields an α expansion for the entropy which is in exact coincidence with ours up to $O(\alpha^4)$ for $K=2$ and up to $O(\alpha^2)$ for $K\geqslant 3$. It seems natural to assume that the two expansions coincide to all orders. The RS ansatz is thus proven to be exact for all α such that $\alpha < \alpha_p = 1/[K(K-1)]$. Beyond this value if, as we discussed above, the influence of the loops on the entropy of the percolating cluster is not negligible, two possibilities arise: either the RS ansatz is wrong, or a more refined handling of the integral equation on $P(h)$ is required [29]. A careful analysis of this problem is worthwhile. One could, for instance, investigate the presence of a singularity at the percolation threshold. In any case, the fact is confirmed that the entropy of the satisfiable phase remains finite up to the satisfiability transition: at the threshold a finite, even if small (~ 0.2 for $K=2$), fraction of sites are in finite-size clusters, and their contribution provides a lower bound for the entropy of the system.

VI. DILUTE SPIN GLASSES: AN EXACT SOLUTION FOR THE PARAMAGNETIC PHASE OF THE VIANA BRAY MODEL

Spin glasses are magnetic systems where the interactions between degrees of freedom (spins) are disordered. The key quantity that determines all the statistical properties of such systems is the free energy averaged over the distribution of the interactions. In most cases this average can only be computed with the replica trick [18], often involving technically subtle replica symmetry breaking (RSB). For IC models in which each spin interacts with all others in the sample, such as the Sherrington-Kirkpatrick model [30], the RSB scheme that yields the exact solution in both high- and low-temperature phases is well understood. The final aim, far from being reached, is to determine the nature of the spin-glass phase of disordered models on a d-dimensional lattice with only short range interactions. Dilute spin glasses can be viewed as an intermediate step between these two limits: each spin interacts with a finite number of other spins, but the model includes no notion of distance since the "neighbors" are randomly chosen from the whole set of spins of the system.

The standard dilute spin-glass model was introduced by Viana and Bray [5]. It is defined by the following Hamiltonian involving N classical Ising spins, $S_i = \pm 1$:

$$H = -\sum_{i \neq j} J_{ij} S_i S_j. \qquad (19)$$

The interactions J_{ij} are independently distributed with the same probability law:

$$P(J_{ij}) = \left(1 - \frac{c}{N}\right)\delta(J_{ij}) + \frac{c}{N}\rho(J_{ij}). \qquad (20)$$

ρ is normalized to 1, and has average and mean square deviation of order 1 to obtain a sensible thermodynamic limit. c is the mean connectivity per spin.

Despite numerous studies [5–7], a complete understanding of this model has not yet been reached. The main diffi-

culty in the study of dilute models is that even without RSB, the order parameter is a function instead of a number as in the IC case. In order to introduce RSB, one has to cope with an order parameter which is at least a functional, leading to very difficult calculations.

The mean connectivity per spin c plays here the same role as α in our expansion of the 2-sat entropy. As opposed to K-sat, for the Viana-Bray model the percolation and the paramagnetic (PM) to spin-glass transition occur at the same critical value $c=1$. In the dilute phase, $c<1$, the clusters have not percolated and the model is in the PM phase [22]. The statistical properties in this phase can be studied with the cluster expansion. The average free energy per site reads

$$-\beta f = \sum_t \frac{1}{L_t} c^{L_t - 1} e^{-L_t c} K_t \overline{\ln Z_t}, \qquad (21)$$

where the overline denotes an average with respect to ρ, and only tree clusters contribute in the thermodynamic limit. This is the analog of Eq. (8) with a slightly different expression for P_t where 2α is replaced by c. Z_t is the partition function of the cluster. One can easily prove that, for trees,

$$\overline{\ln Z_t} = L_t \ln 2 + (L_t - 1)\overline{\ln \cosh \beta J}, \qquad (22)$$

irrespectively of the topology. Then the sum of symmetry factors for all trees of L_t sites is $L_t^{L_t - 2}/(L_t - 1)!$ (a well-known result from graph theory [31]), and results in

$$-\beta f = \ln 2 \left(\sum_{k=1}^{\infty} \frac{(e^{-c})^k}{k!} (ck)^{k-1} \right) + \overline{\ln \cosh \beta J}$$

$$\times \left(c \sum_{k=1}^{\infty} \frac{(e^{-c})^k}{k!} (k-1)(ck)^{k-2} \right). \qquad (23)$$

The two sums can be evaluated (cf. the Appendix) to yield

$$-\beta f = \ln 2 + \frac{c}{2} \overline{\ln \cosh \beta J}. \qquad (24)$$

This result has a simple interpretation: from Eq. (22), each site contributes with $\ln 2$ to the free energy, each link with $\overline{\ln \cosh \beta J}$ for treelike structures. As there are N sites and $cN/2$ links on average, the result follows. One obtains exactly the same free energy following the replica method with a RS ansatz. The exactness of the RS ansatz in this phase is then proven.

The cluster expansion allows one to compute finite-size corrections in a rather simple manner. These read, up to $O(c^3/N)$,

$$\frac{1}{N}\left[-\frac{c}{2}\overline{\ln \cosh \beta J} + c^3\left(\frac{4}{3}\ln 2 + \overline{\ln \cosh \beta J} \right.\right.$$

$$\left.\left. + \frac{1}{6}\overline{\ln(1 + \tanh \beta J_1 \tanh \beta J_2 \tanh \beta J_3)} \right) \right],$$

where J_1, J_2, and J_3 are three independent couplings taken from the probability distribution $\rho(J_{ij})$. It will be very inter-

esting to confront this result with the finite-size corrections to the replica calculation using the RS ansatz. As in the K-sat problem, the expansion cannot be used beyond $c=1$, since it does not take into account the giant cluster appearing at the percolation transition.

VII. CONCLUSIONS AND PERSPECTIVES

The cluster expansion relies on very simple combinatorial arguments. It allows one to solve optimization problems in the "easy" phase, avoiding the introduction of replicas, and it is a general method to obtain finite N corrections. Most importantly, it has allowed us to signal the possible need for a revision of the replica solution of the K-sat, and similar problems, with two successive percolation and easy-to-hard transition.

For spin glasses without a difference in these two transitions, the interest of the method is less apparent, as highly diluted systems are in the less interesting PM phase. Still it would be interesting to test if the RS ansatz is exact for any spin-glass model under its percolation threshold, as was proven here for the Viana-Bray (VB) spin glass in the thermodynamic limit.

Let us note, however, the difference in the percolating and critical behavior of K-sat and the Viana-Bray dilute spin glass. In the former, percolation occurs before the satisfiability transition ($\alpha_p < \alpha_c$); in the latter both phenomena arise at the same value of the control parameter ($c=1$). This can be understood as being due to the fact that 2-sat is, in a way, less frustrated than the VB spinglass. Two manifestations of this fact are given by the behavior of a single loop and a linear cluster. A single loop in 2-sat is always satisfiable while in the VB spinglass it is frustrated each time there is an odd number of antiferromagnetic couplings on it. A linear cluster in 2-sat can be satisfied by a large number of configurations while, in the VB spinglass only two-spin configurations satisfy all bonds.

The cluster expansion can be applied to a variety of interesting problems. For instance, algorithms that solve satisfiability problems through local search, like walk-sat [32], can also be studied with this method [12]. The number of steps needed to solve a formula is the sum of the number of steps needed to solve each cluster. Improvements of the algorithms by means of better heuristics can thus be quantified [12].

ACKNOWLEDGMENTS

APPENDIX

In this Appendix we shall derive a proof of the two summations used in Eq. (23):

$$A = \sum_{k=1}^{\infty} \frac{(e^{-c})^k}{k!} (ck)^{k-1} = 1, \qquad (A1)$$

$$B = \sum_{k=1}^{\infty} \frac{(e^{-c})^k}{k!} (k-1)(ck)^{k-2} = \frac{1}{2}. \qquad (A2)$$

The proof relies on a mathematical identity proven with the help of analytical function tools [33]. Let $w(z)$ be a given function which can be inverted to $z(w)$. Then the coefficients of the series expansion of $z(w)$ are obtained via the following expressions:

$$z(w) = \sum_{k=1}^{\infty} \frac{1}{k!} b_k w^k, \qquad (A3)$$

$$b_k = \frac{d^{k-1}}{dt^{k-1}} \left[\left(\frac{t}{w(t)} \right)^k \right]_{t=0}. \qquad (A4)$$

Let us consider the function $w(z) = z \exp(-cz)$, which is a bijection from $[0,1]$ to $[0,e^{-c}]$ if $c<1$. The coefficients of the series expansion of the reciprocal $z(w)$ are

$$b_k = \frac{d^{k-1}}{dt^{k-1}} \left[\left(\frac{t}{te^{-ct}} \right)^k \right]_{t=0} = (kc)^{k-1}. \qquad (A5)$$

Thus $A = z(e^{-c}) = 1$, as $w(1) = e^{-c}$.

The second series can be transformed using $A=1$:

$$B = \frac{1-e^{-c}}{c} - \sum_{k=2}^{\infty} \frac{(e^{-c})^k}{k!} (ck)^{k-2}. \qquad (A6)$$

If we define

$$g(w) \equiv \sum_{k=2}^{\infty} \frac{w^k}{k!} (ck)^{k-2}, \qquad (A7)$$

then $B = (1-e^{-c})/c - g(e^{-c})$. To compute $g(w)$, we note that $g(0)=0$ and $g'(w) = z(w)/(cw) - 1/c$. By integration and with the change of variables $z=z(w)$, one obtains

$$g(e^{-c}) = -\frac{e^{-c}}{c} + \frac{1}{c} \int_0^1 dz \, w'(z) \frac{z}{w(z)} = \frac{1-e^{-c}}{c} - \frac{1}{2}. \qquad (A8)$$

This yields the final result $B=1/2$.

[1] D. S. Fisher, Phys. Rev. Lett. **56**, 1964 (1986); L. Balents and D. S. Fisher, Phys. Rev. B **48**, 5949 (1993); D. S. Fisher, P. Ledoussal, and C. Monthus, Phys. Rev. E **59**, 4795 (1999).

[2] R. R. P. Singh and S. Chakravarty, Phys. Rev. Lett. **57**, 245 (1986); Phys. Rev. B **36**, 559 (1987); R. R. P. Singh and M. E. Fisher, J. Appl. Phys. **63**, 3994 (1988).

[3] A. Aharony and K. Binder, J. Phys. C **13**, 4091 (1980).

[4] C. de Dominicis, I. Kondor, and T. Temesvári, in *Spin Glasses and Random Fields*, edited by A. P. Young (World Scientific, Singapore, 1997).

[5] L. Viana and A. J. Bray, J. Phys. C **18**, 3037 (1985).

[6] K. Y. M. Wong and D. S. Sherrington, J. Phys. A **21**, L459 (1988); P. Mottishaw, Europhys. Lett. **3**, 333 (1988); P. Mottishaw and C. de Dominicis, J. Phys. A **20**, L375 (1987); C. de Dominicis and P. Mottishaw, *ibid.* **20**, L1267 (1987).

[7] R. Monasson, J. Phys. A **31**, 513 (1998); D. S. Dean, Eur. Phys. J. B **15**, 493 (2000); A. Lefevre and D. S. Dean, e-print cond-mat/0011265.

[8] R. Monasson and R. Zecchina, Phys. Rev. Lett. **76**, 3881 (1996); Phys. Rev. E **56**, 1357 (1997).

[9] G. Biroli, R. Monasson, and M. Weigt, Eur. Phys. J. B **14**, 551 (2000).

[10] B. Hayes, Am. Sci. **85**, 108 (1996); A. Goerdt, J. Comput. Syst. Sci. **53**, 469 (1996); V. Chvàtal and B. Reed, in *Proceedings of the 33rd IEEE Symposium on Foundations of Computer Science* (IEEE, New York, 1992), p. 620.

[11] M. Weigt and A. K. Hartmann, Phys. Rev. Lett. **84**, 6118 (2000); e-print cond-mat/0006316.

[12] G. Semerjian, L. F. Cugliandolo, and R. Monasson (unpublished).

[13] M. Garey and D. S. Johnson, *Computers and Intractability; A Guide to the Theory of NP Completeness* (Freeman, New York, 1979).

[14] A. Goerdt, in *Proceedings of the 7th International Symposium on Mathematical Foundations of Computer Science* (Springer, Verlag, New York, 1992), p. 264.

[15] S. A. Cook, in *Proceedings of the 3rd Annual ACM Symposium on Theory of Computing* (Assoc. Comput. Machinery, New York, 1971), p. 151.

[16] Y.-T. Fu and P. W. Anderson, in *Lectures in the Science of Complexity*, edited by D. Stein (Addison-Wesley, Reading, MA, 1989).

[17] *Frontiers in Problem Solving: Phase Transitions and Complexity*, edited by T. Hogg, B. A. Huberman, and C. Williams, [Art. Intell. **81**, I+II (1996)]; O. Dubois, R. Monasson, B. Selman, and R. Zecchina, Theor. Comput. Sci. (to be published).

[18] M. Mezard, G. Parisi, and M. A. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987); K. H. Fischer and J. A. Hertz, *Spin Glasses* (Cambridge University Press, Cambridge, 1991).

[19] R. Monasson, R. Zecchina, S. Kirkpatrick, B. Selman, and L. Troyansky, Nature **400**, 133 (1999); *Proceedings of Phys-Comp96*, edited by T. Toffoli, M. Baifaore, and J. Leão (Publisher, Boston, 1996).

[20] F. Ricci-Tersenghi, M. Weigt and R. Zecchina, Phys. Rev. E **63**, 026702 (2001); e-print cond-mat/0011181.

[21] S. Kirkpatrick and B. Selman, Science **264**, 1297 (1994).

[22] I. Kanter and H. Sompolinsky, Phys. Rev. Lett. **58**, 164 (1987).

[23] B. Bollobas, *Random Graphs* (Academic Press, London, 1985).

[24] M. Karonsky and T. Luczak, Combinatorics, Probab. Comput. **2**, 283 (1993).

[25] A. J. Bray and G. J. Rodgers, Phys. Rev. B **38**, 11 461 (1988).

[26] D. B. Wilson, preprint math.PR/0005136.

[27] S. Cocco and R. Monasson, e-print cond-mat/0009410 [Phys. Rev. Lett. (to be published)]; e-print cond-mat/0012191.

[28] B. Bollobas, C. Borgs, J. T. Chayes, J. H. Kim, and D. B. Wilson, e-print math.co/9909031.

[29] It is interesting to note that in the variational calculation of Ref. [9], the replica symmetry breaking solution (Sc RSB) of $K \leq 3$-sat appears at a value α_{RSB} of α in between α_p and α_c. Due to the variational character of the solution in Ref. [9], α_{RSB} is an upper bound for the exact RSB threshold. Whether the appearance of the RSB solution is related to some percolation phenomenon remains an interesting open issue.

[30] S. Kirkpatrick and D. Sherrington, Phys. Rev. B **17**, 4384 (1978).

[31] B. Bollobas, *Modern Graph Theory* (Springer-Verlag, New York, 1998).

[32] B. Selman, H. A. Kautz, and B. Cohen (unpublished).

[33] G. Arfken, *Mathematical Methods for Physicists* (Academic Press, New York, 1966).