Sciencexpress Research Article

# Analytic and Algorithmic Solution of Random Satisfiability Problems

M. Mézard,[1] G. Parisi,[1,2] R. Zecchina[1,3]*

[1]Laboratoire de Physique Théorique et Modèles Statistiques, CNRS and Université Paris Sud, Bât. 100, 91405 Orsay cedex, France. [2]Sezione INFN, SMC and UdRm1 of INFM, Università "La Sapienza," P.le A. Moro 2, 00185 Roma, Italy. [3]The Abdus Salam International Centre for Theoretical Physics Str. Costiera 11, 34100 Trieste, Italy.

*To whom correspondence should be addressed; E-mail: zecchina@ictp.trieste.it

**We study the satisfiability of random Boolean expressions built from many clauses with *K* variables per clause (K-satisfiability). Expressions with a ratio $\alpha$ of clauses to variables less than a threshold $\alpha_c$ are almost always satisfiable, while those with a ratio above threshold are almost always unsatisfiable. We show the existence of an intermediate phase below $\alpha_c$, where the proliferation of metastable states is responsible for the onset of complexity in search algorithms. We introduce a new class of optimization algorithms which can deal with these metastable states; one such algorithm has been tested successfully on the largest existing benchmark of K-satisfiability.**

The K-satisfiability problem (Ksat) asks whether one can satisfy simultaneously a set of *M* constraints between *N* Boolean variables, where each constraint is a clause built as the logical OR involving *K* variables (or their negations). Ksat is at the core of combinatorial optimization theory (*1*), and often serves as a benchmark to search algorithm in artificial intelligence and computer science. An efficient algorithm for solving Ksat for $K \geq 3$ would immediately lead to other algorithms for solving efficiently thousands of different hard combinatorial problems. The class of combinatorial problems sharing such a crucial feature is called NP-complete (*2*) and it is a basic conjecture of modern computer science that no such efficient algorithm exists. However, on the more practical side, algorithms which are used to solve real-world NP-complete problems display a huge variability of running times, ranging from linear to exponential. A theory for the typical-case behaviour of algorithms, on classes of random instances chosen from a given probability distribution, is therefore the natural complement to the worst-case analysis (*3–5*). While 1sat and 2sat are solved efficiently by polynomial time algorithms (*6*), $K > 2$ randomly generated Boolean formulae may become extraordinarily difficult to solve: it has been observed numerically (*7*, *8*) that computationally hard random instances are generated when the problems are critically constrained, i.e. close to the SAT/UNSAT phase boundary. The study of critical instances represents a theoretical challenge towards an understanding of the onset of computational complexity and the analysis of algorithms. Moreover, such hard instances are a popular test-bed for the performance of search algorithms (*9*).

The random Ksat problem has close similarities with models of complex materials such as spin glasses which are fundamental systems in the statistical physics of disordered systems (*10*). Spin glasses deal with binary variables ('spins'), interacting with random exchange couplings. Each pair of interacting spins can be seen as a constraint, and finding the state of minimal energy in a spin glass amounts to minimizing the number of violated constraints. Although the precise form of the constraints in spin glasses and Ksat differ, in both cases the difficulty comes from the possible existence of 'frustration' (*11*) which makes it difficult to find the global optimal state by a purely local optimization procedure. Links between combinatorial optimization and statistical physics have been known for long (*10*, *12*, *13*). Techniques taken from statistical physics are particularly useful in the case in which the size of the instance is large.

Two main categories of questions can be addressed. One type is algorithmic, e.g. finding an algorithm that decides whether an instance is SAT or UNSAT (or that tries to minimize the number of violated constraints). Another one is more theoretical and deals with random instances, for which one wants to predict the typical behaviour, e.g. phase transitions and structure of the solution space.

We address the two types of questions in the 3sat problem. When the number of variables *N* and of clauses *M* both increase at a fixed value of $\alpha = M/N$, random Ksat problems become generically SAT at small $\alpha$, and generically UNSAT at large $\alpha$. The existence of a SAT-UNSAT phase transition in the infinite *N* limit has been established rigorously for any *K* (*14*), but the critical value $\alpha_c$ (that separates the two phases) has been found only in the (polynomial) $K = 2$ problem where $\alpha_c = 1$ (*15–17*). For the NP-complete case $K \geq 3$, much less is known. The present best numerical estimate for $\alpha_c$ at $K = 3$ is 4.26 (*18*), and the rigorous bounds are(*19*, *20*) $3.42 < \alpha_c < 4.506$ , while previous statistical mechanics analysis using the replica method, has found $\alpha_c(3) \sim 4.48$ (*21*)

and $\alpha_c(3) \sim 4.396$ (*22*) in the framework of variational approximations. The SAT-UNSAT decision problem is also known experimentally to be algorithmically harder to solve in the neighborhood of $\alpha_c$ depending on the characteristics of the SAT/UNSAT phase transition. Indeed 2sat and 3sat are different in this respect (*23*).

**Setting out the statistical physics problem.** The Ksat problem deals with $N$ Boolean variables $x_i$, $i \in \{1,..., N\}$ . Each clause $a \in \{1,..., M\}$ involves $K$ variables $\{x_{i1(a)},..., x_{iK(a)}\}$ . Each such variable can be negated or not, and the clause is built as the OR function of the K resulting variables. In physical terms, the variable $x_i$ can be represented by a 'spin' $s_i = \pm 1$ through the one-to-one mapping $s_i = -1$ (respectively +1) if $x_i$ is false (resp. true). For each variable $x_{ir(a)}$ appearing in clause $a$, one introduces a 'coupling' $J_a^r = -1$ if the variable appears negated in the clause, otherwise the coupling is $J_a^r = 1$. The set of indices $i_1(a),..., i_K(a)$ and of 'couplings' $\mathbf{J_a} = \{J_a^1,..., J_a^K\}$ define an instance of the problem under study. Given a spin configuration, the 'energy' $\varepsilon_{\mathbf{Ja}}(s_{i_1(a)},...,s_{i_k(a)})$ of clause $a$ is taken equal to 0 if the clause is satisfied, equal to 1 if it is violated (*24*). The total energy $E$ equals the number of violated clauses.

In statistical physics one assigns to each of the $2^N$ spin configurations a Boltzmann probability $\exp(-\beta E)/Z$ where $\beta$ is an auxiliary parameter playing the role of the inverse of temperature and $Z$ is a normalization term; here we shall be interested in the $\beta \to \infty$ 'zero temperature' limit, where Boltzmann's law selects optimal states.

**The spin glass approach.** We first study the large $N$ limit of the random 3sat problem, where the indices in each clause are chosen randomly, as well as the sign of each coupling, with uniform distributions. Our approach to these problems uses a general strategy initiated years ago in spin glass theory (*10*). The first concept we need to introduce is that of a state. Roughly speaking states correspond to connected regions of configurations, such that one must cross energy-barriers that diverge when $N \to \infty$ in order to go from one state to another. The archetype of such a situation is the ferromagnetic transition where the spins collectively polarize, either towards an 'up' state or towards a 'down' state. In frustrated systems such as satisfiability problems there can exist many states: the number of states with energy $E$ behaves as $\exp(N\,\Sigma(e))$, where $e \equiv E/N$ and the function $\Sigma(e)$, called the complexity, is a crucial concept in the studies of structural glasses. The ground state energy density $e$ can be found by the condition $\Sigma(e) = 0$. Here we choose a restricted zero temperature definition which applies to random Ksat: a state is simply a cluster of configurations of equal energy related by single spin flip moves, such that the energy cannot be decreased by any sequence of single spin flips (*25*). Generalizing the approach of (*26*), one can develop a whole 'zero temperature thermodynamics' of the states by introducing a 'free energy' function $\Phi(y)$ defined from:

$$\exp(-Ny\Phi(y)) = \int de \exp\left(N\left[\Sigma(e) - ye\right]\right) \ . \tag{1}$$

The reweighting $y$ is a Lagrange parameter (similar to an inverse temperature) which allows to fix the energy of the states. Larger reweighting selects states of lower energies, until one reaches $y = y^*$ corresponding to the lowest energy states ($y^* = \infty$ in the SAT region).

**The cavity method: message passing procedures.** To compute $\Phi(y)$, we use the zero temperature cavity method (*27*), in which the basic ingredients are the cavity-fields and the cavity-biases, which are defined in each state. The cavity-field $h_{i\to a}$ measures the tendency of spin $i$ to be up, when one of the clauses, $a$, to which $i$ belongs, has been disconnected (Fig. 1). It is equal to the sum of cavity-biases $u_{b\to i}$, sent to site $i$ from all the other clauses $b$ to which it belongs. In computer science terminology, cavity-fields are messages sent from a variable node to a function node, while cavity-biases are messages sent from a function node to a variable node (*28*). The cavity-biases are determined by a local optimization procedure. Consider one clause $a$, involving $K$ variables $s_1,...s_K$, and a penalty function $\varepsilon_J(s_1,..,s_K)$. The optimization on the variables $s_2,..,s_K$

$$\min_{s_2,...,s_K}\left(\epsilon_J(s_1,..,s_K) - \frac{1}{2}\sum_{j=2}^{K} h_j s_j\right) = -\frac{1}{2}\left(a_J(h_2,..,h_K) + s_1 u_J(h_2,...,h_K)\right) \tag{2}$$

defines the mean energy shift $a_J$ and the cavity-bias $u_{a\to 1} = u_J(h_2,..,h_K)$ propagated from this clause to the variable $s_1$ (*29*).

The advantage of cavity-biases and cavity-fields in large ($N >> 1$) random Ksat and spin glass problems is the special structure of the interaction graph: It is locally tree-like, the connectivity fluctuates from site to site with a Poisson distribution of mean $K\alpha$ (see Fig. 1). On a more global scale these random graphs have loops with a typical length growing as $\log(N)$. As the cavity fields $h_2,..,h_K$ are defined in the absence of the clause, they correspond to far away variables (with a distance of order $\log(N)$). The 'clustering property', valid inside each state, implies that their correlations go to zero at large $N$ (on a real tree they would be fully uncorrelated). The topology of the graph implies that the cavity equations are exact on finite subgraphs.

In order to determine the statistical properties of the set of cavity-biases and how they change from state to state, we introduce the 'surveys' which are histograms of cavity-biases. For each state $\omega$, there is one cavity-bias $u^{\omega}_{a\to 1}$ propagated from one clause $a$ to site 1, it can be computed from Eq. 2 where the cavity-fields are those corresponding to the state $\omega$.

For a given value of the reweighting, the survey propagated to spin $s_1$ in Fig. 1 is defined as: $Q^{(y)}_{a\to 1}(u) = C \sum_{\omega} \delta(u - u^{\omega}_{a\to 1})$ where $C$ is a normalization constant insuring that $Q^{(y)}_{a\to 1}(u)$ is a probability distribution, and the sum over $\omega$ is restricted to the states having the energy density $e$ selected by the reweighting $y$.

The survey propagation rules on the graph of Fig. 1 take the precise form:

$$Q^{(y)}_{a\to 1}(u) = C' \int \prod_{r=1}^{q} dw_r Q^{(y)}_{b_r\to 2}(w_r) \int \prod_{s=1}^{q'} dv_s Q^{(y)}_{c_s\to 3}(v_s) \delta(u - u_J(W,V)) e^{y a_J(W,V)} \quad (3)$$

where $W = w_1+...+w_q$, $V = v_1+...+v_{q'}$, and $C'$ is a normalisation constant. The exponential term in Eq.3 takes care of the energy level crossings induced by the propagation. Once the surveys are known, the free energy $\Phi(y)$ can be computed using the formulae of (*27*), and the complexity can be deduced from Eq.1. The order parameter of the theory is the collection of the surveys.

**Phase diagram.** In the zero temperature 3sat problem, one sees from the definition (*2*) that a given cavity-bias $u_{a\to i}$ takes, either the values {0,1} (if the variable $x_i$ appears unnegated in clause $a$), or the values {0,−1}. The corresponding survey $Q^{(y)}_{a\to i}(u)$ is thus characterized by a single number, the probability that $u = 0$. Using this simplification, we have been able to compute (*30*) the statistical distribution of surveys in random graphs in the infinite volume limit. We find two critical values of $\alpha$ at $\alpha_d \sim 3.921$ and $\alpha_c \sim 4.256$. For $\alpha < \alpha_d$, the solution is of a paramagnetic type (all the surveys equal $\delta(u)$), a generic instance is satisfiable, and the solution can be found even by a simple zero temperature Metropolis algorithm (ZTMA) (*31*). For $\alpha_d < \alpha < \alpha_c$, the space of configurations breaks up into many states, and there exists a non trivial complexity (*32*). Some of the states have zero energy, therefore we are still in the SAT phase. It can be argued that algorithms like ZTMA will generically get trapped into the most numerous states, which have an extensive (proportional to $N$) energy $E_{th}$.

At $\alpha = 4.2$ we find analytically $E_{th} \sim .003N$, and we have checked that ZTMA converges to a similar value of energy. The fact that $e_{th} = E_{th}/N$ is small explains the good performance of smarter algorithms on instances involving a few thousand variables. At $\alpha > \alpha_c$, the system is in its UNSAT phase, the lowest possible energy is positive. The phase diagram is summarized in Fig. 2.

**Survey propagation algorithm.** We now consider one given instance (*30*), i.e. one fixed large graph. We have seen experimentally that in the glassy region $\alpha > \alpha_d$, the standard ($y = 0$) iteration of cavity-biases either ceases to converge or it converges to the trivial paramagnetic solution where all $u_{a\to i} = 0$. If $i$ is the $r$-th site connected to the function node $a$, we introduce a survey $Q^{(y)}_{a\to i}(u) = \eta_{a\to i}\delta(u) + (1 - \eta_{a\to i})\delta(u + J_a^r)$ which is characterized by the single number $\eta_{a\to i}$. The survey propagation of Eq. 3 performed with random sequential updating is a message passing procedure which defines a dynamical process in the space of the $K N$ variables $\eta_{a\to i}$. We have implemented it on large random instances in the hard part of the SAT phase, with $\alpha \sim 4.2$–$4.25$, using sufficiently large value of $y$ (typically $y \sim 4$–$6$). The process is found to converge to a unique non-trivial solution. We expect that this survey propagation technique can be of interest in many problems of statistical inference.

The set of all surveys $Q^{(y)}_{a\to i}(u)$ found after convergence provides a nontrivial information on the structure of the states. From all the surveys sent onto one site $i$, we reconstruct through a reweighted convolution (*33*) the probability distribution of local fields on this site, $P_i(H)$. This is a distribution on integers ($P_i(H) = \sum_r \delta(H - r) w_i^r$). The total weight $w_i^+ = \sum_{r=1}^{\infty}$ (resp. $w_i^- = \sum_{r=-\infty}^{-1}$) of $P_i(H)$ on positive (resp. negative) integers gives the fraction of zero energy states where $s_i$ is equal to 1 (resp. to −1). We have checked numerically on single instances with $N = 10000$, that these fractions predicted from survey propagation agree with those obtained by averaging on a few hundreds of ground states.

**A decimation algorithm.** This information can be exploited in order to invent new types of algorithms (*30*), or improve existing ones. We have worked out one such application, the survey inspired decimation (SID), which shows very promising performance, but probably other algorithms could be found using the same type of information. Given an instance, we first compute all the surveys by the survey propagation algorithm with a sufficiently large value of $y$ (e.g. $y = 6$). Then we deduce the distribution of local fields, and in particular their weights $w_i^{\pm}$ on positive and negative integers. We then fix the variable $i$ with largest $|w_i^+ - w_i^-|$ to the value $s_i = Sign(w_i^+ - w_i^-)$. Satisfied clauses are eliminated, unsatisfied K-clauses involving $i$ are transformed into $K$−1 clauses, leading to a new instance with a reduced number of variables, and of clauses. The surveys can be propagated again on this new instance (starting from the previous ones) until convergence, and the procedure is iterated. Whenever a paramagnetic state is found (signaled by all $\eta_{a\to i} = 1$) or at some intermediate

steps, a rapid search process like simulated annealing at a fixed cooling rate is run.

This SID algorithm has been tested successfully on the largest (up to $N = 2000$) existing benchmarks (*9*) of random 3sat instances in the hard regime. Satisfying assignments have been found for all benchmarks. We have applied the SID to much larger instances, increasing $N$ up to $N = 10^5$ at a fixed $\alpha = 4.2$. The algorithm is very efficient: it always finds a SAT configuration, and its apparent complexity scales like $N^2$, although more systematic studies with higher statistics will be necessary to establish this behaviour. For the very same large instances, the only existing algorithm able to find solutions, at a considerable computational cost, is a highly optimized version of the walksat algorithm (*9*, *34*).

**Conclusion.** We have proposed an analytical method that predicts quantitatively the phase diagram of the random 3sat problem in the limit of infinite number of clauses, and opens the way to other types of algorithms. The existence of an intermediate phase with many metastable states close to the SAT-UNSAT transition explains the slowing down of algorithms in this region. We would like to stress that the solution which we propose is typical of a 'one step replica symmetry breaking' solution, as it is called in spin glasses (*10*). All the consistency checks of the analytic results lead us to believe that this solution is exact for the 3sat problem. From the strict mathematical point of view, the phase diagram which we propose should be considered as a conjecture, as for the great majority of the theoretical results in statistical physics. Our computation implies that a way to provide a fully rigorous proof of the transition behaviour in random Ksat problems could be based on the study of the decomposition of the probability measure into states endowed with the clustering property (*35*). On the other hand, the predictions of our theory can be compared with numerical experiments, and our first such tests have confirmed its validity. Based on the analytical study, our algorithm looks promising in that it can solve large instances exploring a rather small number of spin configurations. It will be very interesting to explore its application to other optimization problems.

## References and Notes

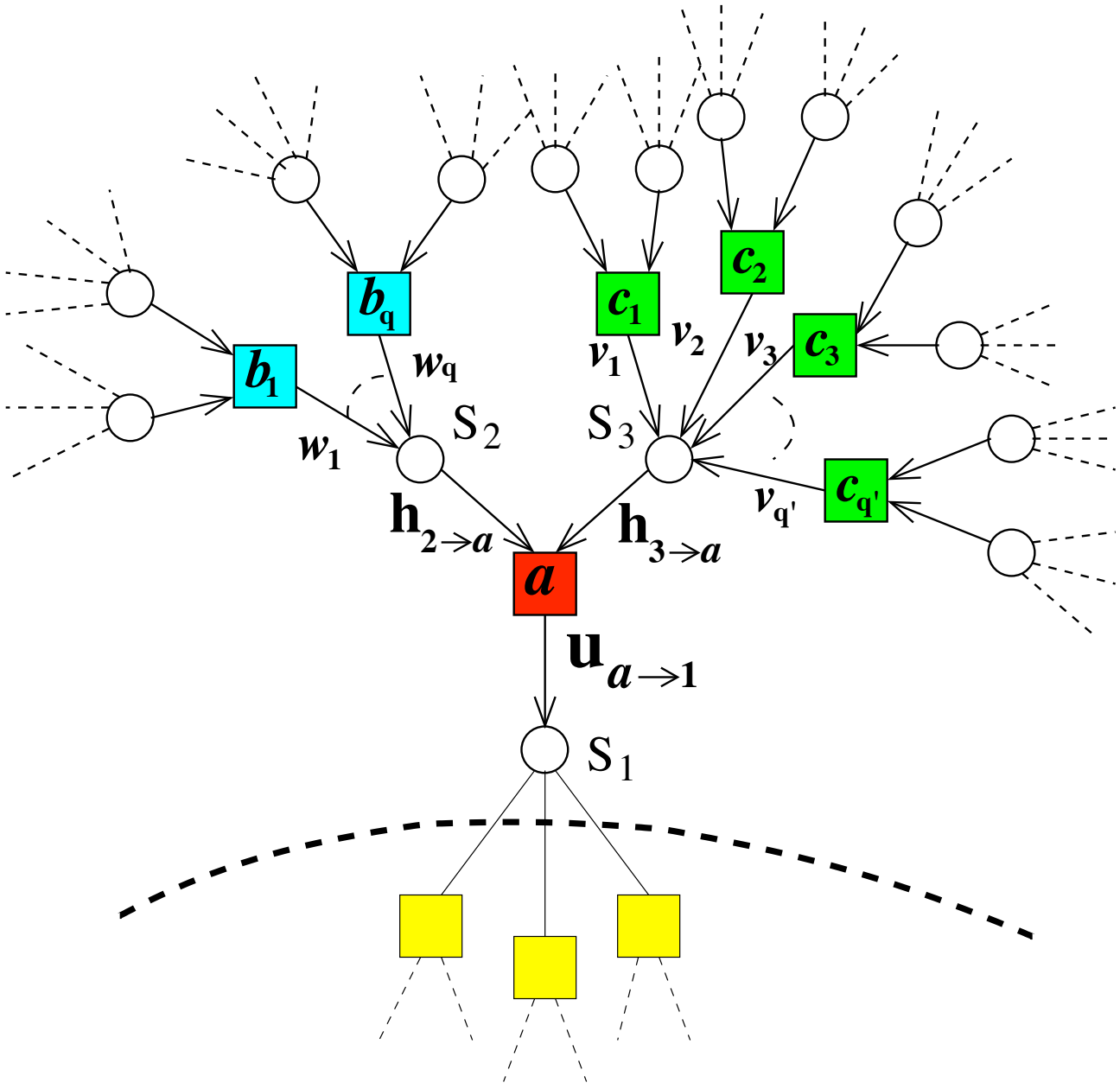

1. S. Cook, in: *Proc. 3rd Ann. ACM Symp. on Theory of Computing* (Assoc. Comput. Mach., New York, 1971) p. 151.
2. M. R. Garey, D. S.Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness* (Freeman, New York, 1979).
3. L. A. Levin, *SIAM J. Comput.*, **14**, 285 (1986).
4. S.Ben-David, B. Chor, O. Goldreich, M. Luby, *JCSS* **44**, 193 (1992).
5. Y. Gurevich, *JCSS* **42**, 246 (1991).
6. B. Aspvall, M. F. Plass, R. E. Tarjan, in *Process. Lett.* **8**, 121 (1979).
7. T. Hogg, B. A. Huberman, C. Williams, Eds., *Artif. Intell.* **81** (1996).
8. O. Dubois, R. Monasson, B. Selman, R. Zecchina, (Eds.), *Theoret. Comp. Sci.* **265** (2001).
9. Satisfiability Library: www.satlib.org/.
10. M. Mézard, G. Parisi, M.A. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987).
11. A subset of variables is frustrated when it is impossible to assign these variables in such a way that the constraints are satisfied [see (*10*)].
12. Y. Fu, P. W. Anderson , Combinatorial Optimization, *J. Phys. A 19*, 1605 (1986).
13. O. C. Martin, R. Monasson, R. Zecchina, *Theor. Comp. Sci.* **256**, 3 (2001).
14. E. Friedgut, *J. Amer. Math. Soc.* **12**, 1017 (1999).
15. A. Goerdt, in *Proc. of the 17th Int. Symposium on Mathematical Foundations of Computer Science* (Prague, 1992).
16. V. Chvàtal, B. Reed, in *Proc. 33rd IEEE Symp. on Foundations of Computer*, 620 (1992).
17. B. Bollobàs, C. Borgs, J. T. Chayes, J. Han Kim, D. B. Wilson, *Rand. Struct. Alg.* **18**, 301 (2001).
18. J. A. Crawford, L. D. Auton, *Artif. Intell.* **81**, 31 (1996).
19. A. Kaporis, L. Kirousis, E. Lalas, in *Proceedings of the 4th European Symposium on Algorithms* (ESA 2002), to appear in series: Lecture Notes in Computer Science, Springer.
20. O. Dubois, Y. Boufkhad, J. Mandler, in *Proc. 11th ACM-SIAM Symp. on Discrete Algorithms*, 124 (San Francisco, CA, 2000).
21. G. Biroli, R. Monasson, M. Weigt, *Euro. Phys. J.* **B 14** 551 (2000).
22. S. Franz, M. Leone, F. Ricci-Tersenghi, R. Zecchina, *Phys. Rev. Lett.* **87**, 127209 (2001).
23. R. Monasson, R. Zecchina, S. Kirkpatrick, B. Selman, L. Troyansky, *Nature* **400**, 133-137 (1999).
24. The energy of a clause can be written explicitly as: $\varepsilon_{\mathbf{Ja}}(s_{i_1(a)},\ldots,s_{i_k(a)}) = 2^{-K}\prod_{q=1}^{K}(1 + J_a^q\, s_{i_q(a)})$.
25. In Ksat, it turns out that each state involves an exponentially large (in $N$) number of configurations.
26. R. Monasson, *Phys. Rev. Lett.* **75**, 2847 (1995).
27. M. Mézard, G. Parisi, *Eur.Phys. J. B* **20**, 217-233 (2001); *The cavity method at zero temperature*, preprint (2002)
28. F.R. Kschischang, B.J. Frey, H.-A. Loeliger, *IEEE Trans. Infor. Theory* **47**, 498 (2002).
29. For $K = 3$ one has $a_J(h_2,h_3) = |h_2| + |h_3| - \theta(J_2h_2)\theta(J_3h_3)$ and $u_J(h_2,..,h_K) = -J_1\theta(J_2h_2)\theta(J_3h_3)$, with $\theta(x) = 1$ if $x > 0$ and $\theta(x) = 0$ if $x \le 0$.

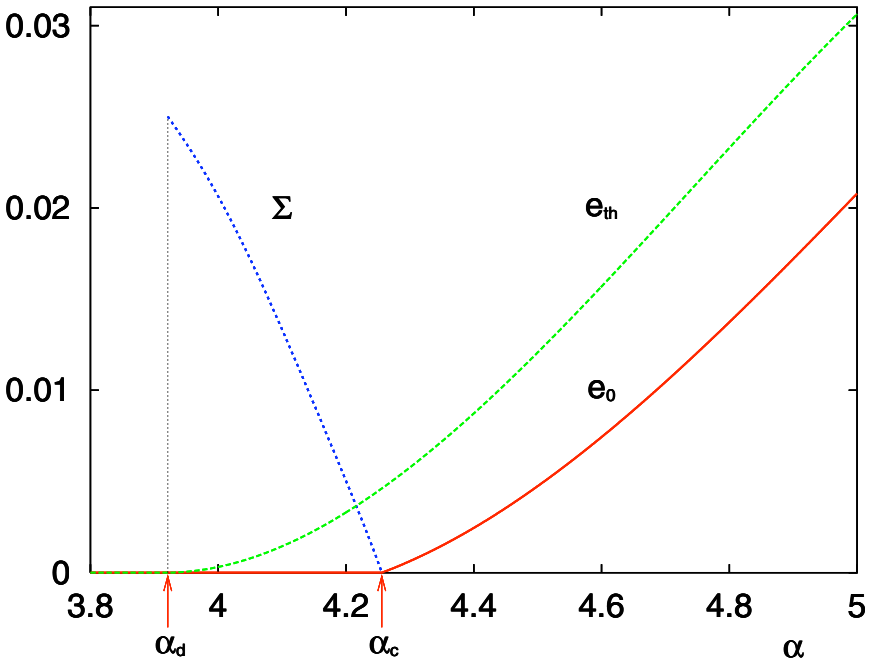
30. M. Mézard, R. Zecchina, *The random K-satisfiability problem: from an analytic solution to an efficient algorithm*, preprint (2002)
31. In the ZTMA, one starts from a given configuration of energy $E$, picks up randomly one variable and computes the new energy $E'$ if this variable is flipped. The move is accepted (i.e. the variable is flipped) whenever $E' \leq E$, otherwise the move is rejected. The procedure is then iterated.
32. This situation is familiar in mean field models of glasses, where $\alpha_d$ corresponds to a density of dynamical arrest, while $\alpha_c$ is the true transition point. For a review, see J.P. Bouchaud, L. Cugliandolo, J. Kurchan, M. Mézard, in *Spin glasses and Random fields*, A.P. Young ed., (World Scientific, Singapore,1997).
33. The explicit form is: $P_i(H) = \int\prod_a[du_a Q_{a\to i}(u_a)]\delta(H - \sum_a u_a)\exp(y|\sum_a u_a|)$, where the index $a$ spans all the function nodes connected to the variable $i$.
34. B. Selman, H. Kautz, B. Cohen, in: *Proceedings of DIMACS*, p. 661 (1993).
35. M. Talagrand, *Prob. Theory and Related Fields***117**, 303-360 (2000).
36. We thank B. Selman and J.S. Yedidia for useful exchanges. This research has been supported in part by the European Science Fundation under the SPHINX program.

25 April 2002; accepted 19 June 2002
Published online 27 June 2002; 10.1126/science.1073287
Include this information when citing this paper.

**Fig. 1.** In the random 3sat problem, the graph of clauses is locally isomorphic to a tree. Variables (spins) are depicted by a circle, and clauses by a square. The cavity-bias $u_{a\to 1}$ sent from the red clause $a$ to the variable $s_1$ summarizes the effects of optimizing clause $a$ on $s_2$,$s_3$, taking into account all the blue+green (top) part of the graph, when the yellow (bottom) part has been taken away. The cavity-field $h_{2\to a}$ (resp. $h_{3\to a}$) sums up all cavity-biases $w_1,...,w_q$ ($v_1,...,v_{q'}$) arriving onto $s_2$ ($s_3$) from the blue (green) clauses, in the absence of the red clause.

**Fig. 2.** The phase diagram of the random 3sat problem. Plotted is $e_0$, the number of violated clauses per variable (red), versus the control parameter $\alpha$ which is the number of clauses per variable. The sat-unsat transition occurs at $\alpha = \alpha_c \sim 4.256$. The green line is $e_{th}$, the threshold energy per variable, where local algorithms get trapped. The blue line is the complexity $\Sigma$ of satisfiable states, equal to $1/N$ times the logarithm of their number.




0.03
0.02
0.01
0
3.8
4
4.2
4.4
4.6
4.8
5
Σ
e_th
e_0
α_d
α_c
α