Algorithmica (2011) 59: 48–65
DOI 10.1007/s00453-009-9308-1

Random 2 XORSAT Phase Transition

Daudé Hervé · Ravelomanana Vlady

Received: 10 July 2008 / Accepted: 12 April 2009 / Published online: 5 May 2009

Abstract We consider the 2-XOR satisfiability problem, in which each instance is a formula that is a conjunction of Boolean equations of the form $x \oplus y = 0$ or $x \oplus y = 1$. Formula of size m on n Boolean variables are chosen uniformly at random from among all $\binom{n(n-1)}{m}$ possible choices. When $c < 1/2$ and as n tends to infinity, the probability $p(n, m = cn)$ that a random 2-XOR formula is satisfiable, tends to the threshold function $\exp(c/2) \cdot (1 - 2c)^{1/4}$. This gives the asymptotic behavior of random 2-XOR formula in the SAT/UNSAT subcritical phase transition. In this paper, we first prove that the error term in this subcritical region is $O(n^{-1})$. Then, in the critical region $c = 1/2$, we prove that $p(n, n/2) = \Theta(n^{-1/12})$. Our study relies on the symbolic method and analytical tools coming from generating function theory which also enable us to describe the evolution of $n^{1/12}\ p(n, \frac{n}{2}(1 + \mu n^{-1/3}))$ as a function of μ. Thus, we propose a complete picture of the finite size scaling associated to the subcritical and critical regions of 2-XORSAT transition.

Keywords XORSAT · Satisfiability · Constraint Satisfaction Problem · Phase transition · Random graph · Generating functions · Finite size scaling

D. Hervé
Laboratoire d'Analyse, Topologie et Probabilités, UMR CNRS 6632, Université de Provence, 39, rue Joliot Curie, 13453 Marseille Cedex 13, France
e-mail: herve.daude@gyptis.univ-mrs.fr

R. Vlady (✉)
Laboratoire d'informatique de Paris Nord, UMR CNRS 7030, Institut Galilée, Université Paris Nord, 99, Av. Clément, 93430 Villetaneuse, France
e-mail: vlad@lipn.univ-paris13.fr

1 Introduction

1.1 Context

The last decade has seen a growth of interest in the phase transition for Boolean Satisfiability (SAT) and more generally for Constraint Satisfaction Problems (CSP). For any $k \geq 2$, the random version of the famous k-SAT problem is known to exhibit a sharp [14] phase transition: as the density c of formula (where the number of clauses is c times the number of variables) is increased, formula abruptly change from being satisfiable to being unsatisfiable at a critical threshold point. For general CSP, determining the nature of the SAT/UNSAT phase transition (sharp or coarse), locating it, determining a precise scaling window and better understanding the structure of the space of solutions turn out to be very challenging tasks. These have aroused a lot of interest in different communities, namely mathematics, computer science and statistical physics (see e.g. [1, 13]). It is well known that a random 2-CNF formula with density $c < 1$ is satisfiable with probability tending to 1 as the number n of variables tends to infinity, while for $c > 1$, the probability of satisfiability tends to 0 as n tends to infinity [10, 15]. Indeed there is now, [3], a detailed picture of the transition yielding a scaling window of size $\Theta(n^{2/3})$. For greater values of k much less is known about the precise behaviour of random k-SAT near the threshold point whose exact location is still an open problem.

The difficulty of identifying transition factors and of performing theoretical explorations of the SAT transition has incited many researchers to turn to a variant of the SAT problem: the k-XORSAT problem. One is given a linear system over n Boolean variables, composed of m equations modulo 2, each involving exactly $k \geq 2$ variables. This problem introduced in [7] has contributed to develop or validate techniques, thus revealing typical behaviors of both random instances and their space of solutions for SAT-type problems (see, e.g., [5, 12, 24]). Particularly 2-XORSAT appears to be a seed of coarseness for the transition of a wide class of CSP [9].

Our main goal is to give a precise description of the SAT/UNSAT phase transition associated to random 2-XOR-formula. We propose a new and detailed analysis of this transition, based entirely on generating functions and analytic methods.

1.2 Main Contribution and Organization of the Paper

In this work, we consider random formula with n variables and m clauses chosen uniformly, independently and without replacement from the $n(n-1)$ possible 2-XOR-clauses. In using the so-called symbolic method from generating function theory [18], we give new enumerative and analytic results on the probability $p(n,m)$ that such random formula is satisfiable. Let us recall that in [8], Creignou and Daudé have shown that $p(n,m)$ verifies:

$$\lim_{n\to+\infty} p(n,cn) = \exp(c/2)\cdot(1-2c)^{1/4} \quad \text{for } 0 < c < 1/2 \text{ and } 0 \text{ if } c \geq 1/2. \qquad (1)$$

We will provide more accurate results about the evolution of the function $p(n,m)$ as n tends to infinity. First, we will consider formula in a subcritical region, including $m = cn$ with $0 < c < 1/2$. Then, we will focus on formula in the critical region, that is when $m = \frac{n}{2}(1+\mu n^{-1/3})$. Our main contribution is condensed in the following:

Theorem 1.1 *The probability of satisfiability of random* 2*-XOR-formula with n variables and m clauses satisfies*:

- *for any m such that* $n - 2m \gg n^{2/3}$,

$$p(n,m) = \exp\left(\frac{m}{2n}\right)\left(1 - \frac{2m}{n}\right)^{1/4} + O\left(\frac{n^2}{(n-2m)^3}\right), \tag{2}$$

- *for any real* μ, *there exists* $\Psi(\mu) > 0$ *such that*

$$\lim_{n\to\infty} n^{1/12}\, p\left(n, \frac{n}{2}(1 + \mu n^{-1/3})\right) = \Psi(\mu). \tag{3}$$

For the first item, let us point out that the threshold function $\exp(c/2) \cdot (1 - 2c)^{1/4}$ mentioned in (1) was obtained by probabilistic methods. Here we show that the generating function approach is very well suited to obtain stringant results about finite size effects. In particular, when $m = cn$ with $c < 1/2$, the error term in (2) becomes $O(n^{-1})$ and this high rate of convergence is nicely illustrated by the first Fig. 1. Observe that a lower rate of convergence clearly takes place at the critical ratio $c = 1/2$. In Sect. 4.1, we study the subcritical regime and (2) is proven in Theorem 4.1.

The second item shows that in the critical region, the probability of satisfaction is of order of magnitude $O(n^{-1/12})$. R. Monasson [25] pointed out to us that this scaling can also be obtained by statistical physics methods. There is an explicit connection

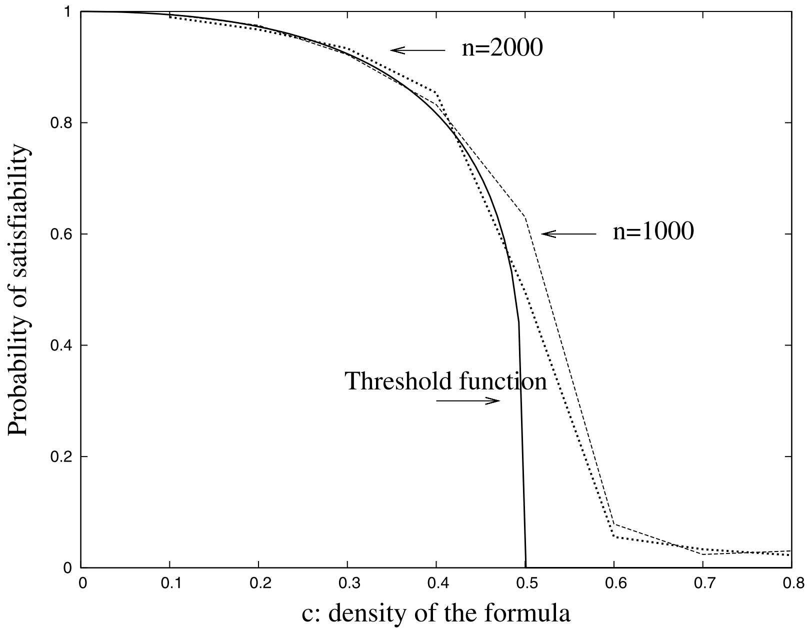

Fig. 1 Probability $p(n, cn)$ that a random 2-XOR formula with cn equations and n variables is satisfiable as a function of the ratio c, for various size of and the asymptotic threshold function

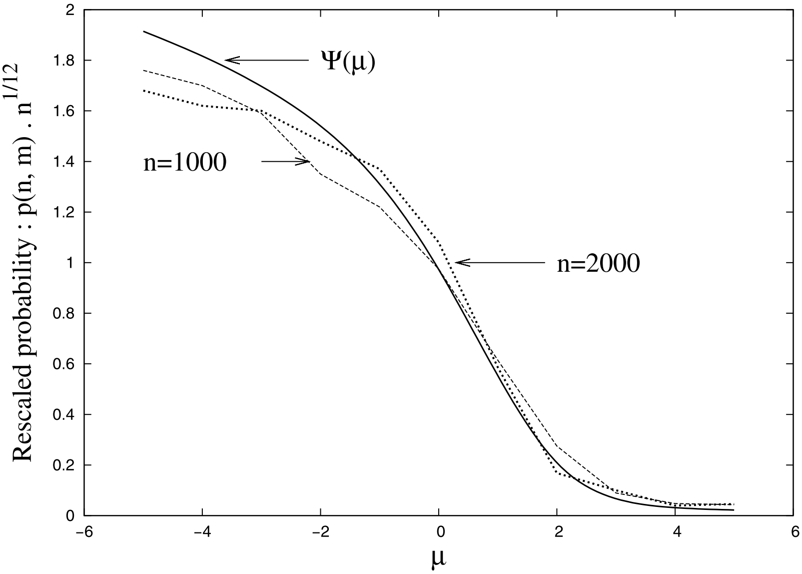

Fig. 2 Rescaled probability at the zero threshold point $c = 1/2$: $n^{1/12}\, p(n, n/2 + \mu/2n^{2/3})$ as a function of μ, for $n = 1000$ and 2000 with the theoretical limit given by Theorem 4.2

between so-called Airy functions and the decreasing function Ψ mentioned above. In Sect. 4.2, Theorem 4.2 gives the expression of Ψ in terms of the Airy function. In Fig. 2, observe that after the rescaling, our theoretical result fits quite well with the empirical results. Moreover, their accuracy can be compared to the one on the finite size scaling for the core of large random hypergraphs and recently obtained via probabilistic methods by A. Dembo and A. Montanari [16].

In Fig. 2, we give empirical results corresponding to random 2-XOR formula with $n = 1000$ (resp. $n = 2000$) variables around the point phase transition $m = \frac{n}{2} + \mu \frac{n^{2/3}}{2}$ with $\mu \in [-4, 4]$.

Our main references for this work will be the pioneering work of E.M. Wright [29–31] together with the "giant paper" of Janson, Knuth, Luczak and Pittel [22].

Section 2 gives the enumerative background of our investigation. Theorem 2.1 will show that $p(n, m)$ can be expressed in terms of the nth coefficient of some generating function associated to edge (0/1) weighted graphs. We will see that this key generating function is composed of generating functions associated to the enumeration of connected graphs by their excess (number of edges minus number of vertices).

Section 3 will provide algebraic and analytic tools. They will point out the role of each generating function appearing in the expression for $p(n, m)$. In Sect. 3.1 we will first extract dominant terms, and Lemma 3.1 from [22] will explain the connection between enumerative results and the so-called Airy functions. Then, in Theorem 3.2 we make precise earlier investigations done in [19] about the probability that a random graph has complex components with more edges than vertices.

In Sect. 4 we establish our main results on the 2-XORSAT phase transition.

2 Enumerative Results

In order to study random 2-XOR-formula, we will use enumerative methods by means of generating functions (see for instance [18]). In this section, we first show that a 2-XOR-formula is satisfiable if and only if its associated edge-weighted (0/1) graph has no cycle of odd weight. Then, in following the pioneering work of E.M. Wright [29–31], we will show how these graph configurations can be enumerated with exponential generating functions.

2.1 2 XORSAT and Weighted Graphs

A 2-*XOR-clause* or shortly a *clause* is a linear equation over the finite field $GF(2)$ using exactly 2 variables, $C = ((x_1 \oplus x_2) = \varepsilon)$ where $\varepsilon = 0$ or 1. A 2-*XOR-formula* or simply a *formula* is a conjunction of distinct 2-XOR-clauses.

A *truth assignment* I is a mapping that assigns 0 or 1 to each variable in its domain. It satisfies a 2-XOR-clause $C = ((x_1 \oplus x_2) = \varepsilon)$ iff $(I(x_1) + I(x_2)) \bmod 2 = \varepsilon$. A truth assignment satisfies a formula s iff it satisfies every clause in s. The set of 2-XOR-formula that are satisfiable (resp., unsatisfiable) will be denoted by 2*XORSAT* or shortly *SAT* (resp. UNSAT). Observe that every formula that contains an unsatisfiable sub-formula is itself unsatisfiable. In the terminology of random graph theory, *SAT* (resp. *UNSAT*) is a decreasing (resp. increasing) property. Throughout this paper we reserve n for the number of variables ($\{x_1, \dots, x_n\}$ denotes the set of variables). There are $N = 2\binom{n}{2} = n(n-1)$ different 2-XOR-clauses over n variables. We consider random formula obtained by choosing uniformly, independently and without replacement m clauses from the N possible 2-clauses. This defines a discrete probability space $(\Omega(n,m), P_{n,m})$ whose associated probability is the uniform law:

$$\forall s \in \Omega(n,m) \quad P_{n,m}(s) = \binom{n(n-1)}{m}^{-1}.$$

We are interested in estimating the probability that a formula drawn at random from $\Omega(n,m)$ is satisfiable, that is in estimating

$$p(n,m) := P_{n,m}(2XORSAT).$$

Each formula s in $\Omega(n,m)$ can be represented by a weighted graph $G(s)$ with n vertices (one for each variable) and m weighted (0 or 1) edges. For each equation $x_i \oplus x_j = \varepsilon$, we add the edge $\{x_i, x_j\}$ weighted by ε. The *weight* of a graph is the sum of the weights of its edges. Observe that our weighted graphs are without self-loops and that the underlying graph of a satisfiable formula is always simple (without multiple edges). We have, as noticed in [8]:

Proposition 2.1 *A formula s is satisfiable if and only if $G(s)$ does not contain any cycle of odd weight*, *i.e.*,

$$p(n,m) = P_{n,m}\Big(G(s) \text{ has no cycle of odd weight}\Big). \tag{4}$$

Proof Suppose that we have an elementary cycle of odd weight in $G(s)$. To such a cycle corresponds a subsystem $s' = (x_{i_1} \oplus x_{i_2} = \varepsilon_1, x_{i_2} \oplus x_{i_3} = \varepsilon_2, \ldots, x_{i_{\ell-1}} \oplus x_{i_\ell} = \varepsilon_{l-1}, x_{i_\ell} \oplus x_{i_1} = \varepsilon_l)$ in s. For any assignment I, $I(x_{i_\ell}) + I(x_{i_1}) + \sum_{j=1}^{\ell-1} I(x_{i_j}) + I(x_{i_{j+1}})$ is even whereas $\sum_{j=1}^{\ell} \varepsilon_j$ is odd, hence s' and so s are unsatisfiable.

Conversely, suppose there is no elementary cycle of odd weight in $G(s)$. We can sequentially assign a value to each vertex in proceeding a depth-first search in the graph $G(s)$. Since it does not contain any elementary cycle with odd weight, the assignment so obtained satisfies all the corresponding equations. □

2.2 Exponential Generating Functions

Let us recall that a connected (weighted) graph has *excess* ℓ if it has $\ell(\geq -1)$ more (weighted) edges than vertices. A connected component of excess ℓ is also called ℓ-*component*. If $\ell > 0$, ℓ-components are called *complex*. The *total excess* r of a graph is the number of edges plus the number of acyclic components, minus the number of vertices (see [22, Sect. 13]). Note that the total excess of a tree component is equal to 0 whereas its excess is equal to -1 and the total excess of a graph is non negative. Obviously, we get from (4):

Proposition 2.2 *For any $r \geq 0$, let $p_r(n,m)$ be the probability that a weighted graph with m weighted* $(0/1)$ *edges, n vertices and a total excess equals to r, has no cycle of odd weight. Then,*

$$p(n,m) = \sum_{r \geq 0} p_r(n,m). \tag{5}$$

The number of trees on n labelled vertices is given by Cayley's formula: n^{n-2}. Let $W_{-1}(z)$ be the exponential generating function (EGF for short) of labelled trees and T be the EGF of rooted labelled trees, we know from [4] that:

$$W_{-1}(z) = T(z) - \frac{T^2(z)}{2} \quad \text{where } T(z) = ze^{T(z)} = \sum_{n=1}^{\infty} n^{n-1}\frac{z^n}{n!}. \tag{6}$$

In our case, observe that from each labelled tree on n vertices one can construct 2^{n-1} weighted trees corresponding to the choices of the weight of each of the $n-1$ edges. Hence, the EGF of labelled weighted $(0/1)$ trees is given by:

$$\sum_{n=1}^{\infty} 2^{n-1} n^{n-2}\frac{z^n}{n!} = \frac{W_{-1}(2z)}{2} = \frac{T(2z)}{2} - \frac{T^2(2z)}{4}. \tag{7}$$

Note that our combinatorial structures are constrained since the considered weighted graphs are without cycles of odd weights. As shown by the longstanding open problem of enumerating exactly triangle-free graphs [20], it is in general extremely difficult to derive EGFs of such constrained structures [28]. However, we have a crucial enumerative result generalizing (7):

Proposition 2.3 *For any $\ell \geq -1$, let W_ℓ be the EGF of connected graphs with excess ℓ (or ℓ-connected graphs), then the EGF of ℓ-connected weighted $(0/1)$ graphs without cycles of odd weight is given by*:

$$\frac{W_\ell(2z)}{2}. \tag{8}$$

Proof Let G be a connected graph on n vertices with excess ℓ. We will show that among the $2^{n+\ell}$ ways to put $(0/1)$ weights on its $n+\ell$ edges only 2^{n-1} of them give a weighted graph without cycles of odd weight.

Let T be a spanning tree of G, and a be any of the $(\ell+1)$ edges in $G \setminus T$. The graph $T \cup a$ has a unique cycle C_a and in graph terminology (see for instance [11]) the $(\ell+1)$ cycles C_a form a basis of the cycle space associated to G. Observe now that as soon as we have chosen the weight of the $(n-1)$ edges in T there is a unique way to put a weight on edge a closing the fundamental cycle C_a in order to produce an even weight for it. Having done so, the fact that any (weighted) cycle in G can be expressed as a linear combination of the C_a gives the assurance that G has no cycles of odd weight. □

Following the terminology of [22], a graph is called *complex* when all its components are complex. Let $F_r(z)$ be the EGF of all complex weighted labelled graphs (connected or not), with a positive total excess r and without cycles of odd weight. The empty graph gives $F_0(z) = 1$. Observe that a non connected complex graph has a total excess greater than 2, this shows that $F_1(z) = \frac{W_1(2z)}{2}$. More generally, as detailed in [22, Sect. 8], the formal enumerative method gives

$$\sum_{r\geq 0} F_r(z) = \exp\left(\sum_{k\geq 1} \frac{W_k(2z)}{2}\right) \tag{9}$$

and for any $r \geq 1$

$$r F_r(z) = \sum_{k=1}^{r} k \frac{W_k(2z)}{2} F_{r-k}(z). \tag{10}$$

Observe that weighted graphs on n labelled vertices, m edges and with total excess r have exactly $n-m+r$ acyclic components (or tree components). Thus, such graphs without cycles of odd weight are enumerated by the following EGF:

$$\frac{1}{(n-m+r)!}\left(\frac{W_{-1}(2z)}{2}\right)^{n-m+r} \exp\left(\frac{W_0(2z)}{2}\right) F_r(z). \tag{11}$$

Let us recall (see [22, Equation (3.5)]) that

$$W_0(z) = \frac{1}{2}\ln\frac{1}{1-T(z)} - \frac{T(z)}{2} - \frac{T^2(z)}{4}. \tag{12}$$

Therefore, with (7) we get:

Theorem 2.1 *The probability $p_r(n,m)$ that a weighted graph with m weighted* $(0/1)$ *edges, n vertices and a total excess equal to r, has no cycle of odd weight is exactly*

$$p_r(n,m) = \frac{n!}{\binom{n(n-1)}{m}} \left[z^n\right] \frac{(\frac{T(2z)}{2} - \frac{T^2(2z)}{4})^{n-m+r}}{(n-m+r)!} \frac{e^{-\frac{T(2z)}{4} - \frac{T^2(2z)}{8}}}{(1-T(2z))^{1/4}} F_r(z), \quad (13)$$

where $[z^n]f(z)$ denotes the n-th coefficient in the series $f(z)$.

3 Tools

In this section, we will provide key algebraic and analytic tools stemming from [19, 22].

3.1 Algebraic Tools

E.M. Wright has shown that the EGF of connected graphs with excess $\ell \geq 1$ can be expressed in terms of $T(z)$. More precisely, Wright has shown that for each $\ell \geq 1$ there exist rational coefficients $(w_{\ell,d})_{d\in\{0,\dots,3\ell+2\}}$ such that:

$$W_\ell(z) = \sum_{d=0}^{3\ell+2} \frac{w_{\ell,d}}{(1-T(z))^{3\ell-d}}. \quad (14)$$

Following [18, 22], we remark that the dominant asymptotic behavior of $[z^n]W_\ell(z)$ is governed by the leading coefficient $b_l := w_{l,0}$. These coefficients are known as Wright's constants (see Janson [21] and references therein). Then, starting with recurrence relation (10) and along the same lines as in [22, Sects. 7, 8], one deduces that the EGF of all weighted graphs with total excess $r \geq 1$ and without cycles of odd weight can also be expressed as a rational function of T:

$$F_r(z) = \sum_{d=0}^{2r} \frac{f_{r,d}}{(1-T(2z))^{3r-d}}. \quad (15)$$

Once more, for any $r \geq 1$, the dominant asymptotic behavior of $[z^n]F_r(z)$ will be governed by the leading coefficient $f_r := f_{r,0}$. Observe that from (10) and with $f_0 = 1$ we get:

$$2rf_r = \sum_{k=1}^{r} kb_k f_{r-k}, \qquad r > 0. \quad (16)$$

Note that these new coefficients f_r are strongly related to coefficients e_r introduced in [22] and satisfying

$$re_r = \sum_{k=1}^{r} kb_k e_{r-k}, \qquad \text{with } e_0 = 1. \quad (17)$$

Indeed, let us consider the (ordinary) generating functions:

$$B(z) := \sum_{k\ge 0} b_k z^k, \qquad F(z) := \sum_{k\ge 0} f_k z^k, \qquad E(z) := \sum_{k\ge 0} e_k z^k,$$

then from (17) and (16) we get $\exp(B(z)) = E(z)$ and $\exp(\frac{B(z)}{2}) = F(z)$, thus $F(z)^2 = E(z)$.

The coefficients e_r have a nice expression (see [22, (7.2)]):

$$e_r = \frac{(6r)!}{2^{5r} 3^{3r} (3r)!(2r)!}. \tag{18}$$

Thought we did not find an explicit expression similar to (18) for the sequence (f_r), the first few of them can be computed via the recurrence given by (16) and we find:

$$f_1 = \frac{5}{48}, \qquad f_2 = \frac{745}{4608}, \qquad f_3 = \frac{329165}{663552}, \qquad f_4 = \frac{289208785}{127401984},$$
$$f_5 = \frac{84326076625}{6115295232}, \qquad f_6 = \frac{183796927660325}{1761205026816} \cdots$$

From the above formulas, it is also easy to prove that

$$\text{for } r \ge 1 \quad f_r \le \frac{e_r}{2}. \tag{19}$$

Moreover, the EGFs F_r satisfy Wright's type inequalities. As in [22, 27] we have

$$\frac{f_r}{(1-T(2z))^{3r}} - \frac{f_{r,1}}{(1-T(2z))^{3r-1}} \preceq F_r(z) \preceq \frac{f_r}{(1-T(2z))^{3r}}, \quad r \ge 1, \tag{20}$$

where for formal series $A(z)$ and $B(z)$, $A(z) \preceq B(z)$ iff for all $n \ge 1$ $[z^n]A(z) \le [z^n]B(z)$.

3.2 Analytic Tools

Speaking about the phase transition in random graphs $G(n,m)$ or $G(n,p)$, Erdős and Rényi [17] suggested that a "*double jump*" occurs: the largest component changes its size (with respect to the number of vertices n) twice—from $O(\log n)$ to $O_p(n^{2/3})$—and then from $O_p(n^{2/3})$ to $O(n)$ (for the notation $X_n = O_p(a_n)$ see e.g. [23, p. 10]). During these changes, the structures of components as well as the size of the giant component have been studied by various authors [2, 6, 22, 26]. Often, the methods in use differ depending on the edge probability $p = p(n)$ or on the number of edges $m = m(n)$.

Inside the phase transition of random graph, the most significant results have been obtained by Janson, Knuth, Pittel and Łuczak in the giant paper [22] using generating functions and analytic combinatorics. In fact, the latter authors proved the following [22, Lemma 3] (with our notations):

Lemma 3.1 *If* $m = \frac{1}{2}n(1 + \mu n^{-1/3})$ *and if* y *is any real constant, we have*

$$\frac{2^m m!n!}{(n-m)!n^{2m}}[z^n]\frac{W_{-1}(z)^{n-m}}{(1-T(z))^y} = \sqrt{2\pi}A(y,\mu)n^{y/3-1/6} + O\big((1+|\mu|^B)n^{y/3-1/2}\big) \tag{21}$$

uniformly for $|\mu| \le n^{1/12}$, *where* $B = \max(4, \frac{9}{2} - y)$ *and*

$$A(y,\mu) = \frac{e^{-\mu^3/6}}{3^{(y+1)/3}} \sum_{k\ge 0} \frac{(\frac{1}{2}3^{2/3}\mu)^k}{k!\Gamma((y+1-2k)/3)}. \tag{22}$$

This non-obvious lemma, whose proof is based on analytic techniques will be also our key tool for the computation of probabilities of satisfiability inside the critical window $m \in [\frac{n}{2} - O(n^{2/3}), \frac{n}{2} + O(n^{2/3})]$ of the phase transition of random 2-XOR formula.

In the subcritical phase, we will use the following improvement of a result given Ph. Flajolet, D.E. Knuth and B. Pittel about the probability that a random (unweighted) graph with n vertices and m edges contains only trees and unicycles. In [19, Theorem 4] they prove that this probability is equal to $1 - O(n^{-1/2})$ when $\frac{m}{n} < \frac{1}{2}$ and as n tends to ∞. In modifying their proof, we obtain:

Theorem 3.2 *For any sequence of integers* $m(n)$ *such that* $n - 2m \gg n^{2/3}$, *let* $q_0(n,m)$ *be the probability that a random unweighted graph with* n *vertices and* m *edges contains only trees and unicycles. We have*

$$q_0(m,n) = 1 - O\left(\frac{n^2}{(n-2m)^3}\right), \tag{23}$$

as $n \to +\infty$.

Proof Throughout this proof, we can and will assume that m satisfies $\delta n < m < \frac{n}{2}$ for some small constant $\delta \in]0, \frac{1}{2}[$. In fact, the probability that a graph with $m = o(n)$ edges does not contain any multicyclic components is greater than the probability that a graph with $m = \Theta(n)$ edges contains only trees and unicyclic components. Thus, it suffices to prove (23) for m in the range $\delta n < m < \frac{n}{2}$.

As in Theorem (2.1) we get:

$$q_0(m,n) = \frac{n!}{\binom{\binom{n}{2}}{m}}\left[z^n\right]\frac{(T(z) - T(z)^2/2)^{n-m}}{(n-m)!}\exp\left(W_0(z)\right), \tag{24}$$

where T and W_0 are given by (6) and (12). Notice that T is the inverse of the function $z\exp(-z)$.

We split this formula in two parts : $q_0(m,n) = St(m,n) \cdot Ca(m,n)$ with

$$St(m,n) = \frac{n!}{\binom{\binom{n}{2}}{m}(n-m)!} \quad \text{and}$$
$$Ca(m,n) = \left[z^n\right](T(z) - T(z)^2/2)^{n-m}\exp\left(W_0(z)\right). \tag{25}$$

Using Stirling's formula, we have for the stated range of m

$$\frac{n!m!}{(n-m)!} = \sqrt{2\pi}\,\frac{n^{n+1/2}m^{m+1/2}}{(n-m)^{n-m+1/2}}e^{-2m}\left(1+O\left(\frac{1}{n}\right)\right). \tag{26}$$

Next, we compute

$$\binom{\binom{n}{2}}{m} = \frac{n^{2m}}{2^m m!}\exp\left(-\frac{m}{n}-\frac{m^2}{n^2}+O\left(\frac{m}{n^2}\right)+O\left(\frac{m^3}{n^4}\right)\right). \tag{27}$$

Therefore, we get

$$St(m,n) = \left(\frac{2\pi nm}{n-m}\right)^{1/2}\frac{2^m n^n m^m}{n^{2m}(n-m)^{n-m}}\exp\left(-2m+\frac{m}{n}+\frac{m^2}{n^2}\right)\left(1+O\left(\frac{1}{n}\right)\right). \tag{28}$$

Next, in using Cauchy integral's formula and substituting z by ze^{-z}, we obtain:

$$Ca(m,n) = \frac{2^{m-n}}{2\pi i}\oint (2T(z)-T(z)^2)^{n-m}\exp\left(W_0(z)\right)\frac{dz}{z^{n+1}} = \frac{2^{m-n}}{2\pi i}\oint g(z)e^{nh(z)}\frac{dz}{z} \tag{29}$$

where

$$\begin{aligned} g(z) &= (1-z)^{1/2}e^{-z/2-z^2/4}, \\ h(z) &= z-\frac{m}{n}\log z+\left(1-\frac{m}{n}\right)\log(2-z). \end{aligned} \tag{30}$$

$h'(z)=0$ for $z=1$ or $z=2m/n$. $h''(1)=2m/n-1<0$ and $h''(2m/n)=\frac{n(n-2m)}{4m(n-m)}$ >0. As in [19], we can apply the saddle-point method integrating around a circular path $|z|=2m/n$. Let $\Phi(\theta)$ be the real part of $h(2m/ne^{i\theta})$. We have

$$\begin{aligned} \Phi(\theta) &= 2\frac{m}{n}\cos\theta+\left(1-2\frac{m}{n}\right)\log 2-\frac{m}{n}\log\left(\frac{m}{n}\right) \\ &\quad +\frac{\left(1-\frac{m}{n}\right)}{2}\log\left(1+\frac{m^2}{n^2}-2\frac{m}{n}\cos\theta\right) \end{aligned} \tag{31}$$

and

$$\Phi'(\theta) = -2\frac{m}{n}\sin\theta+\frac{(1-m/n)m}{n\left(1+m^2/n^2-2m/n\cos\theta\right)}\sin\theta. \tag{32}$$

We note that $\Phi(\theta)$ is a symmetric function of θ. Fix sufficiently small positive constant θ_0. Then, $\Phi(\theta)$ takes its maximum value at $\theta=\theta_0$ as $\theta\in[-\pi,-\theta_0]\cup[\theta_0,\pi]$. In fact,

$$\Phi(\theta)-\Phi(\pi) = 4\frac{m}{n}+\left(1-\frac{m}{n}\right)\log\left(\frac{n-m}{n+m}\right)+O(\theta^2). \tag{33}$$

Therefore, if $\theta \to 0$ $\Phi(\theta) > \Phi(\pi)$. Also, $\Phi'(\theta) = 0$ for $\theta = 0$ and $\theta = \theta_1$ (for some $\theta_1 > 0$). Standard calculus show that $\Phi(\theta)$ is decreasing from 0 to θ_1 and then increasing from θ_1 to π. We also have

$$h^{(p)}(z) = (p-1)!\left((-1)^p \frac{m}{nz^p} - \frac{(n-m)}{n(2-z)^p}\right), \quad p \geq 2. \tag{34}$$

Hence,

$$h(2me^{i\theta}/n) = h(2m/n) + \sum_{p\geq 2} \xi_p (e^{i\theta} - 1)^p, \tag{35}$$

where $\xi_p = \frac{(2m/n)^p}{p!} h^{(p)}(2m/n)$ and $|\xi_p| \leq \frac{m}{np}(\frac{2m}{n})^p + \frac{n-m}{np}$. We then have

$$\left|\sum_{p\geq 4} \xi_p (e^{i\theta} - 1)^p\right| = O(\theta^4). \tag{36}$$

This allows us to write

$$h(2m/ne^{i\theta}) = h(2m/n) - \frac{m(n-2m)}{2n(n-m)}\theta^2 - i\frac{(n^2 - 5nm + 2m^2)m}{6n(n-m)^2}\theta^3 + O(\theta^4). \tag{37}$$

Let $\tau = n(n-m)/(m(n-2m))$ and

$$\theta_0 = \left(\frac{(n-m)}{(n-2m)m}\right)^{1/2} \cdot \omega(n) = \sqrt{\frac{\tau}{n}} \cdot \omega(n) \tag{38}$$

where we need a function $\omega(n)$ satisfying $n\theta_0^2 \gg 1$ but $n\theta_0^3 \ll 1$ as $n \to \infty$. We choose

$$\omega(n) = \frac{(n-2m)^{1/4}}{n^{1/6}}. \tag{39}$$

We can now use the magnitude of the integrand at θ_0 to bound the error and our choice of θ_0 verifies

$$|g(2m/ne^{i\theta_0})(\exp(nh(2m/ne^{i\theta_0})) - \exp(nh(2m/n)))| = O(e^{-\omega(n)^2/2}). \tag{40}$$

Thus, from (29) we obtain

$$Ca(m,n) = \frac{2^{m-n}}{2\pi} \int_{-\theta_0}^{\theta_0} g\left(2\frac{m}{n}e^{i\theta}\right) \exp\left(nh(2m/ne^{i\theta})\right) d\theta \times \left(1 + O\left(e^{-\omega(n)^2/2}\right)\right). \tag{41}$$

We replace θ by $\frac{\tau^{1/2}}{n^{1/2}}t$. The integral in the above equation leads to

$$\left(\frac{\tau}{n}\right)^{1/2} \int_{-\omega(n)}^{\omega(n)} g\left(\frac{2m}{n}\exp(it\sqrt{\tau/n})\right) \exp\left(nh\left(\frac{2m}{n}\exp(it\sqrt{\tau/n})\right)\right) dt. \tag{42}$$

Expanding $g(2m/ne^{it\sqrt{\tau/n}})$, we obtain

$$\left(\frac{\tau}{n}\right)^{1/2}\int_{-\omega(n)}^{\omega(n)} g\,(2m/n)\left(1 - i\frac{2m\tau^{1/2}(n^2-2m^2)}{n^{5/2}(n-2m)}t + O\left(\frac{n^2}{(n-2m)^3}t^2\right)\right)$$
$$\times \exp\left(nh\left(\frac{2m}{n}\exp\left(it\sqrt{\tau/n}\right)\right)\right)dt. \tag{43}$$

Observe that our choice of $\omega(n)$ in (39) and the hypothesis $n-2m \gg n^{2/3}$ justify such an expansion. Similarly, using the expansion of $h(2m/ne^{it\sqrt{\tau/n}})$ yields

$$\left(\frac{\tau}{n}\right)^{1/2}\int_{-\omega(n)}^{\omega(n)} g\,(2m/n)\left(1 - i\frac{2m\tau^{1/2}(n^2-2m^2)}{n^{5/2}(n-2m)}t + O\left(\frac{n^2}{(n-2m)^3}t^2\right)\right)$$
$$\times \exp\left(nh\left(\frac{2m}{n}\right)-\frac{1}{2}t^2\right)\left(1 - i\frac{(n^2-5nm+2m^2)}{6(n-m)^{1/2}m^{1/2}(n-2m)^{3/2}}t^3\right.$$
$$\left.+\,O\left(\frac{n}{(n-2m)^2}t^4\right)\right)dt. \tag{44}$$

Using the symmetry of the function, we can cancel terms such as it and it^3 (in fact all odd powers of t). Standard calculations show also that for m in the stated ranges, the multiplication of the factors of it and it^3 leads to a term of order of magnitude $O(n^2/(n-2m)^3t^4)$. Therefore we obtain,

$$\begin{aligned} Ca(m,n) &= \frac{2^{m-n}}{2\pi}\left(\frac{\tau}{n}\right)^{1/2} g\,(2m/n)\,e^{nh(2m/n)} \\ &\quad\times\int_{-\omega(n)}^{\omega(n)} e^{-t^2/2}\left(1 - O\left(\frac{n^2}{(n-2m)^3}t^4\right)\right)dt \\ Ca(m,n) &= 2^{m-n}\left(\frac{\tau}{2\pi n}\right)^{1/2} g\,(2m/n)\,e^{nh(2m/n)} \\ &\quad\times\left(1 - e^{-O(\omega(n)^2)} - O\left(\frac{n^2}{(n-2m)^3}\right)\right). \end{aligned} \tag{45}$$

Multiplying (28) and (45) leads to the result after nice cancellations. □

4 Analytic Results

In this section, we shall use analytic combinatorics in order to compute the probability of satisfiability of random 2-XOR formula.

4.1 From SAT to UNSAT

In the subcritical region, namely as $m < \frac{n}{2}$ and n becomes large, we give the evolution from near 1 to near 0 of the probability $p(n,m)$.

Theorem 4.1 *For any m such that $n-2m \gg n^{2/3}$, the probability that a random 2 XOR formula with n variables and m clauses is SAT verifies*

$$p(n,m)=\exp\left(\frac{m}{2n}\right)\left(1-\frac{2m}{n}\right)^{1/4}+O\left(\frac{n^2}{(n-2m)^3}\right). \tag{46}$$

Proof By Theorem 3.2, the probability that the underlying support graph contains only trees and unicyclic components is $1-O(n^2/(n-2m)^3)$. To prove the theorem, we then need to show that the probability that the random formula contains only trees and even weighted unicycles is about $e^{m/2n}(1-2m/n)^{1/4}$ with error terms of order $O(n^2/(n-2m)^3)$. This probability is exactly

$$\frac{n!}{\binom{n(n-1)}{m}}\left[z^n\right]\frac{(\frac{T(2z)}{2}-\frac{T(2z)^2}{4})^{n-m}}{(n-m)!}\frac{e^{-T(2z)/4-T(2z)^2/8}}{(1-T(2z))^{1/4}}. \tag{47}$$

By using Stirling formula, we find ($m<n/2$)

$$\begin{aligned}\frac{n!}{(n-m)!\binom{n(n-1)}{m}}&=\sqrt{2\pi}n^{n-2m+1/2}\exp\left(\frac{m}{2n}+\frac{m^2}{2n^2}-2m\right)\frac{m^{m+1/2}}{(n-m)^{n-m+1/2}}\\&\quad\times\left(1+O\left(\frac{1}{n}\right)\right).\end{aligned} \tag{48}$$

After replacing z by ze^{-2z} in $[z^n](T(2z)/2-T(2z)^2/4)^{n-m}e^{-T(2z)/4-T(2z)^2/8}$ $\times(1-T(2z))^{-1/4}$ we obtain

$$\frac{2^{2m-n-1}e^n}{\pi}\oint g_1(z)\exp\left(nh_1(z)\right)\frac{dz}{z}, \tag{49}$$

where

$$\begin{aligned}h_1(z)&=z-1-\log z-\left(1-\frac{m}{n}\right)\log\frac{1}{1-(z-1)^2},\\g_1(z)&=(1-z)^{3/4}e^{-z/4-z^2/8}.\end{aligned} \tag{50}$$

In the same vein as the proof of Theorem 3.2, the saddle-point method applies. This time we integrate around a circular path $|z|=\frac{2m}{n}$. As in the proof of the previous theorem, we find that the integral given by (49) is

$$2^{2m-n-1}e^n\sqrt{\frac{2}{\pi}\frac{(n-m)}{m(n-2m)}}g_1\left(\frac{2m}{n}\right)\exp\left(nh_1\left(\frac{2m}{n}\right)\right)\left(1-O\left(\frac{n^2}{(n-2m)^3}\right)\right). \tag{51}$$

Multiplying (48) and (51) ends the proof of Theorem (4.1) after nice cancellations. □

4.2 Finite Size Scaling

In the critical region, our main result is a direct consequence of the dominated convergence theorem. Indeed, from (5) in Proposition 2.2 we have $p(n,m) = \sum_{r\geq 0} p_r(n,m)$. Then, the two following facts will prove the theorem given below.

Fact (i) Let A be defined as in (22). For all integer $r \geq 0$

$$n^{1/12} p_r(n,m) \sim \frac{\sqrt{2\pi} e^{1/4} f_r}{2^r} A(3r + 1/4, \mu). \tag{52}$$

Fact (ii) There exist $R, C, \epsilon > 0$ such that for all $r \geq R$ and all n:

$$n^{1/12}\ p_r(n,m) \leq C\ e^{-\epsilon r.} \tag{53}$$

Theorem 4.2 *Let μ be any real constant. The probability $p(n,m)$ that a random 2-XORSA-formula with n variables and $m = \frac{n}{2}(1 + \mu n^{-1/3})$ equations is satisfiable verifies*:

$$\lim_{n\to\infty} n^{1/12} p(n,m) = \left(\sum_{r=0}^{\infty} \frac{\sqrt{2\pi} e^{1/4} f_r}{2^r} A(3r + 1/4, \mu) \right), \tag{54}$$

where the sequence $(f_r)_{r\in\mathbb{N}}$ is given by (16) and A is defined by (22).

Proof We have to prove facts (i) and (ii).

Proof of Fact (i): Following [22, Lemma 3], let $r \in \mathbb{N}$ be fixed. If $m = \frac{1}{2}n(1 + \mu n^{-1/3})$ and if y is any real constant, we have

$$\frac{e^{-\mu^3/6-n}}{2^{2m-n-2r}} [z^n] \frac{(\frac{T(2z)}{2} - \frac{T(2z)^2}{4})^{n-m+r} e^{-T(2z)/4-T(2z)^2/8}}{(1-T(2z))^y} \sim e^{-3/8} A(y,\mu) n^{y/3-2/3} \tag{55}$$

where $A(y,\mu)$ is given by (22).

This formula is a key tool for the computation of probabilities $p_r(n,m)$. In order to prove it, we first use Cauchy's integral formula. Then the substitution $\tau = \frac{z}{2}e^{-z}$ gives $T(\tau) = \frac{z}{2}$ and we get

$$[z^n] \frac{(\frac{T(2z)}{2} - \frac{T(2z)^2}{4})^{n-m+r} e^{(-\frac{T(2z)}{4} - \frac{T(2z)^2}{8})}}{(1-2T))^y} = \frac{e^n 2^{2m-n-2r-1}}{\pi i} \oint (1-z)^{1-y} e^{-z/4-z^2/8} e^{nh_1(z)} \frac{dz}{z}, \tag{56}$$

where h_1 is given by (50). The proof of (55) can now be completed following the one of [22, Lemma 3]. We choose the path of integration $z = e^{-(\alpha+it)n^{-1/3}}$ where t runs from $-\pi n^{1/3}$ to $\pi n^{1/3}$ and α is the positive solution of $\mu = \frac{1}{\alpha} - \alpha$. As in [22, Equation (10.12)], we have $h_1(1) = h_1'(1) = 0$ and if $m = \frac{n}{2}$, $h_1''(1) = 0$.

Next, using Stirling approximation yields

$$\frac{n!}{\binom{n(n-1)}{m}}\frac{1}{(n-m+r)!}\sim\pi^{1/2}\frac{2^{-\mu n^{2/3}+r+1/2}}{n^{r-1/2}}e^{-\mu^3/6-n+5/8}. \tag{57}$$

The term '$n^{y/3}$' in (55) tells us that we have only to consider the term $\frac{f_r}{(1-2T)^{3r}}$ from F_r (the other term in the lower-bound of F_r in the inequalities (20) can be neglected). In taking $y=3r+1/4$ and multiplying (55), (57) and f_r, we establish Fact (i).

Proof of Fact (ii): With the inequalities (20) and (13), and in using the change of variable $\tau=\frac{z}{2}e^{-z}$, we prove that the probability $p_r(n,m)$ verifies

$$\begin{aligned}
p_r(n,m) &\le \frac{n!}{\binom{n(n-1)}{m}}\left[z^n\right]\frac{(T(2z)/2-T(2z)^2/4)^{n-m+r}}{(n-m+r)!}\frac{f_r e^{-T(2z)/4-T(2z)^2/8}}{(1-T(2z))^{3r+1/4}}\\
&=\frac{n!}{\binom{n(n-1)}{m}}\frac{f_r}{(n-m+r)!}\frac{e^n 2^{2m-n-r-1}}{\pi i}\\
&\quad\times\oint\left(\frac{z(2-z)}{1-z}\right)^r(1-z)^{3/4-2r}e^{-z/4-z^2/8}e^{nh_1(z)}dz,
\end{aligned}\tag{58}$$

with $h_1(z)$ as in (50) and where the contour is a circle $z=\rho e^{i\theta}$ with $0<\rho<1$. On this circle, $|(2-z)/(1-z)|$ and $1/|1-z|$ attain their maxima at $z=\rho$. When $r\ge 1$, the contour is less than

$$\begin{aligned}
&\frac{\rho}{2\pi}\left(\frac{\rho(2-\rho)}{1-\rho}\right)^r(1-\rho)^{3/4-2r}e^{-\rho/4-\rho^2/8}e^{nh_1(\rho)}\int_{-\pi}^{\pi}\exp\left(\frac{-4n\rho(1-\rho)}{9\pi^2}\right)d\theta\\
&\quad<\frac{3}{4}\sqrt{\frac{\pi}{n}}\rho^{r+1/2}(2-\rho)^r(1-\rho)^{1/4-3r}e^{nh_1(\rho)}.
\end{aligned}\tag{59}$$

Let $\rho=1-\frac{r^{1/3}}{n^{1/3}}$. Note that if $r>m$, the probability $p_r(n,m)=0$. Therefore, we can restrict our attention to $r\le\frac{n}{2}+\frac{\mu n^{2/3}}{2}$. First, we have

$$\frac{\rho^r(2-\rho)^r}{(1-\rho)^{3r}}=\left(\frac{n}{r}\right)^r\left(1-\frac{r^{2/3}}{n^{2/3}}\right)^r<\left(\frac{n}{r}\right)^r. \tag{60}$$

Next, we find

$$nh_1(\rho)<\frac{13}{12}r+\frac{11}{6}\mu r^{2/3}. \tag{61}$$

To get (61), we bound $h_1(\rho)$ as follows

$$\begin{aligned}
h_1(\rho)&=-\frac{r^{1/3}}{n^{1/3}}-\log\left(1-\frac{r^{1/3}}{n^{1/3}}\right)+\left(1-\frac{m}{n}\right)\log\left(1-\frac{r^{2/3}}{n^{2/3}}\right)\\
&<-\frac{r^{1/3}}{n^{1/3}}+f_1\left(\frac{r^{1/3}}{n^{1/3}}\right)+\left(1-\frac{m}{n}\right)f_2\left(\frac{r^{2/3}}{n^{2/3}}\right),
\end{aligned}\tag{62}$$

where $f_1(x) = x + x^2/2 + x^3/3 + x^4$ and $f_2(x) = -x - x^2/2 - x^3/3$. The upper bound for $h_1(\rho)$, (62), is justified since for $x \in [0, 0.7]$, we have $-\log(1-x) < f_1(x)$ and $\log(1-x) < f_2(x)$. After a bit of algebra and in assuming that $r/n < 1$ we get (61). In using (57) and $f_r \leq e_r/2$, we prove that there exists an absolute constant $\alpha_1 > 0$ such that

$$\frac{n!}{\binom{n(n-1)}{m}} \frac{f_r}{(n-m+r)!} \frac{e^n 2^{2m-n-2r-1}}{\pi} < \alpha_1 \left(\frac{r}{n}\right)^{r-1/2} \left(\frac{3}{4e}\right)^r e^{-\mu^3/6}. \tag{63}$$

Taking into account (59), (60), (61) and (63), yields

$$p_r(n,m) < \alpha_2 r^{-5/12} n^{-1/12} \exp\left(-\frac{\mu^3}{6} + \frac{11}{6}\mu r^{2/3} + \left(\log 3 - 2\log 2 + \frac{1}{12}\right) r\right), \tag{64}$$

for some constant $\alpha_2 > 0$. Note that, the quantity $(\log 3 - 2\log 2 + \frac{1}{12}) = -0,2043\cdots < 0$. Therefore, for r sufficiently large we have (53). □

5 Conclusion

In this paper, we have studied the probability that a random 2-XOR formula is satisfiable. By means of enumerative and analytic combinatorics we have shown how one can quantify this probability when the number of variables goes to infinity. With the choices of appropriate paths of integration in Cauchy's integral formulas we gave an accurate description for the scaling window associated to 2-XORSAT transition. Such proof techniques may be considered relatively new for the study of Constraint Satisfaction Problems. We hope this work will initiate innovative and promising techniques in this research field.

References

1. Achlioptas, D., Moore, C.: Random k-SAT: Two moments suffice to cross a sharp threshold. SIAM J. Comput. **36**, 740–762 (2006)
2. Barraez, D., Boucheron, S., De la Vega, F.: On the fluctuations of the giant component. Comb. Probab. Comput. **9**, 287–304 (2000)
3. Bollobás, B., Borgs, C., Chayes, J.T., Kim, J.H., Wilson, D.B.: The scaling window of the 2-SAT transition. Random Struct. Algorithms **18**, 201–256 (2001)
4. Cayley, A.: A Theorem on Trees. Quart. J. Math. Oxford Ser. **23**, 376–378 (1889)
5. Cocco, S., Dubois, O., Mandler, J., Monasson, R.: Rigorous decimation-based construction of ground pure states for spin glass models on random lattices. Phys. Rev. Lett. **90**, 047205 (2003)
6. O'Connell, N.: Some large deviations results for sparse random graphs. Probab. Theory Relat. Fields **110**, 277–285 (1998)
7. Creignou, N., Daudé, H.: Satisfiability threshold for random XOR-CNF formula. Discrete Appl. Math. **96–97**(1–3), 41–53 (1999)
8. Creignou, N., Daudé, H.: Coarse and sharp thresholds for random k-XOR-CNF. Satisfiability. Inform. Théor. Appl./Theor. Inform. Appl. **37**(2), 127–147 (2003)
9. Creignou, N., Daudé, H.: Coarse and sharp transitions for random generalized satisfiability problems. In: Proceedings of the Third Colloquium on Mathematics and Computer Science, Vienna (Austria), pp. 507–516. Birkhäuser, Basel (2004)

10. Chvátal, V., Reed, B.: Mick gets some (the odds are on his side). In: Proceedings of the 33rd Annual Symposium on Foundations of Computer Science, pp. 620–627. IEEE Press, New York (1992)
11. Diestel, R.: Graph Theory. Springer, Heidelberg (2000)
12. Dubois, O., Mandler, J.: The 3-XOR-SAT threshold. In: Proceedings of the 43th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2002), pp. 769–778 (2002)
13. Dubois, O., Monasson, R., Selman, B., Zecchina, R.: Editorial. Theor. Comput. Sci. **265**(1–2), (2001)
14. Friedgut, E.: Sharp thresholds of graph properties, and the k-sat problem. J. Am. Math. Soc. **12**(4), 1017–1054 (1999) (appendix by J. Bourgain)
15. Goerdt, A.: A threshold for unsatisfiability. J. Comput. Syst. Sci. **53**(3), 469–486 (1996)
16. Dembo, A., Montanari, A.: Finite size scaling for the core of large random hypergraphs. Ann. Appl. Probab. **18**, 1993–2040 (2008)
17. Erdös, P., Rényi, A.: On the evolution of random graphs. Publ. Math. Inst. Hung. Acad. Sci. **5**, 17–61 (1960)
18. Flajolet, P., Sedgewick, R.: Analytic Combinatorics. Cambridge University Press, Cambridge (2008)
19. Flajolet, P., Knuth, D.E., Pittel, B.: The first cycles in an evolving graph. Discrete Math. **75**, 167–215 (1989)
20. Harary, F., Palmer, E.: Graphical Enumeration. Academic Press, San Diego (1973)
21. Janson, S.: Brownian excursion area, Wright's constants in graph enumeration, and other Brownian areas. Probability Surveys **3**, 80–145 (2007)
22. Janson, S., Knuth, D.E., Luczak, T., Pittel, B.: The birth of the giant component. Random Struct. Algorithms **4**, 233–358 (1993)
23. Janson, S., Łuczak, T., Ruciński, A.: Random Graphs. Wiley, New York (2000)
24. Mézard, M., Ricci-Tersenghi, F., Zecchina, R.: Alternative solutions to diluted p-spin models and XORSAT problems. J. Stat. Phys. **111**, 505 (2003)
25. Monasson, R.: Introduction to phase transitons in random optimization problems. Personnal communication (2007)
26. Ravelomanana, V.: The average size of giant components between the double-jump. Algorithmica **46**(3–4), 529–555 (2006)
27. Ravelomanana, V.: Another proof of Wright's inequalities. Inf. Proc. Lett. **104**(1), 36–39 (2007)
28. Ravelomanana, V., Thimonier, L.: Forbidden subgraphs in connected graphs. Theor. Comput. Sci. **314**(1–2), 121–171 (2004)
29. Wright, E.M.: Asymptotic relations between enumerative functions in graph theory. Proc. Lond. Math. Soc. **20**, 558–572 (1970)
30. Wright, E.M.: The number of connected sparsely edged graphs. J. Graph Theory **1**, 317–330 (1977)
31. Wright, E.M.: The number of connected sparsely edged graphs. III. Asymptotic results. J. Graph Theory **4**, 393–407 (1980)