

Random 2-XORSAT/MAX-2-XORSAT and their phase transitions

VLADY RAVELOMANANA

LIAFA – UMR CNRS 7089 .
Université Denis Diderot.
vlad@liafa.jussieu.fr

joint work with

HERVÉ DAUDÉ (LATP – Université de Provence)

&

VONJY RASENDRAHASINA (LIPN – Université de Paris-Nord)

Outline of the talk

- Introduction & motivations.
- The 2-XORSAT phase transition.
- MAX-2-XORSAT.
- Conclusion and perspectives.

Introduction & Motivations

- Decision and optimization problems play central key rôle in CS (cf. [GAREY, JOHNSON 79] , [AUSIELLO *et al.* 03])
 - 1 A **decision problem** is a question in some formal system with a **yes/no** answer :
$$\left\{ \begin{array}{l} \text{INPUT : an instance } \mathcal{I} \text{ and a property } \mathcal{P}. \\ \text{OUTPUT : yes or no } \mathcal{I} \text{ satisfies } \mathcal{P}. \end{array} \right.$$
 - 2 An **optimization problem** is the problem of finding the **best solution** from all feasible solutions.
- In this talk, we consider two such problems :
2-XORSAT and **MAX-2-XORSAT**.

- Random k -SAT formulas ($k > 2$) are subject to phase transition phenomena [FRIEDGUT, BOURGAIN 1999] .
- Main research tasks include
 - 1 **Localization** of the threshold (ex. **3-SAT** 4.2...? **3-XORSAT** 0.91... [DUBOIS, MANDLER 03])
 - 2 Nature of the phenomena : **sharp/coarse**. [CREIGNOU, DAUDÉ 2000++] .
 - 3 Details inside the **window of transition** (ex. **2-SAT** [BOLLOBÀS, BORGS, KIM, WILSON 01])
 - 4 **Space** of solutions (ex. [ACHLIOPTAS, NAOR, PERES 07] or [MONASSON *et al.* 07])

- An instance : $(v_1 \vee v_2) \wedge (\neg v_1 \vee v_3) \wedge (\neg v_1 \vee \neg v_2)$
- A solution : SAT with $(v_1 = 1, v_2 = 0, v_3 = 1)$.
- Localization of the threshold : n variables, $m = c \times n$ clauses randomly picked from the set of $4 \binom{n}{2}$ clauses.
 $c < 1$ Proba SAT $\rightarrow 1$, $c > 1$ Proba SAT $\rightarrow 0$.

Underlying combinatorial structures : directed graphs.

$$\text{Write } x \vee y \quad \text{as} \quad \begin{cases} \neg x = 1 \implies y = 1 \\ \neg y = 1 \implies x = 1 \end{cases}$$

Characterization : SAT iff no directed path between x and $\neg x$ (and vice-versa).

Proof. First and second moments method [GOERDT 92, DE LA VEGA 92, CHVÁTAL, REED 92] .

Main motivations

- Since the empirical results ([KIRKPATRICK, SELMAN 90] about k -SAT, rigorous results are quite **limited!**
- What are the contributions of **ENUMERATIVE/ANALYTIC COMBINATORICS** to SAT/CSP-like problems?
- **MONASSON** (2007) inferred that (statistical physics) :

$$\lim_{n \rightarrow +\infty} n^{\text{critical exponent}} \times \text{Pr}[\text{2XORSAT}(n, \frac{n}{2})] = O(1),$$

where “critical exponent” = **1/12** .

- We will **show** that “critical exponent” = **1/12** and will **explicit** the hidden constant behind the $O(1)$.
- We will **quantify** the **MAXIMUM** number of satisfiable clauses in random formula.

The 2-XORSAT phase transition

- **Ex :**

$$x_1 \oplus x_2 = 1, x_2 \oplus x_3 = 0, x_1 \oplus x_3 = 0, x_3 \oplus x_4 = 1, \dots$$

- **General form :** $AX = C$ where A has m rows and 2 columns and C is a m -dimensional 0/1 vector.

- **Distribution :** uniform. We pick m clauses of the form $x_i \oplus x_j = \varepsilon \in \{0, 1\}$ from the set of $n(n-1)$ clauses.

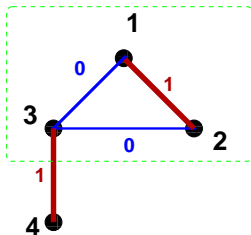
- **Underlying structures :** **graphs with weighted edges**
 $x \oplus y = \varepsilon \iff$ edges of weight $\varepsilon \in \{0, 1\}$.

Characterisation :

SAT iff no elementary cycle of odd weight.

SAT iff no elementary cycle of odd weight

$$\begin{cases} x_1 \oplus x_2 = 1 \\ x_2 \oplus x_3 = 0 \\ x_1 \oplus x_3 = 0 \\ x_3 \oplus x_4 = 1 \end{cases}$$



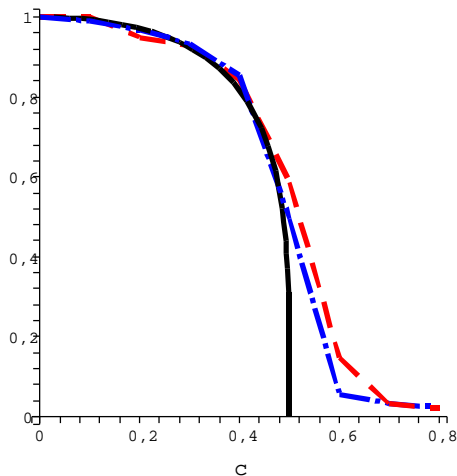
- UNSAT \iff Fix a cycle of odd weight ...
- SAT \iff No cycles of odd weight. DFS affectation based proof.

A basic scheme

- 1 **Enumeration** of “SAT”-graphs (graphs without cycles of odd weight) by means of generating functions.
- 2 Use the obtained results with **analytic combinatorics** to compute :

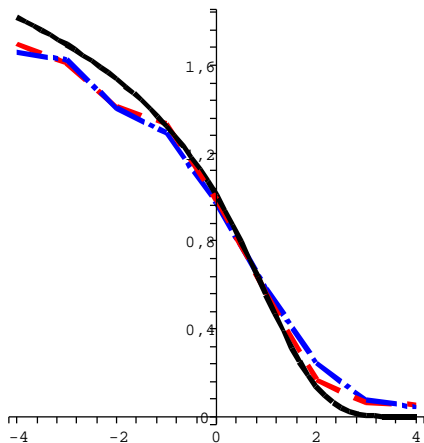
$$\text{Prob. SAT} = \frac{\text{Nbr of configurations without cycles of odd weight}}{\text{Nbr total of configurations}} .$$

Taste of our results : the whole window



$p(n, cn) \stackrel{\text{def}}{=} \text{Proba} [2\text{-XOR with } n \text{ variables, } cn \text{ clauses}] \text{ is SAT}$
for $n = 1000$, $n = 2000$ and the **theoretical** function : $e^{c/2}(1 - 2c)^{1/4}$.

Taste of our results: rescaling the critical window



Rescaling at the point “zero”, i.e $c = 1/2$: $n = 1000$, $n = 2000$ and $\lim_{n \rightarrow \infty} \underbrace{n^{1/12}}_{\times} p(n, n/2 + \mu n^{2/3})$ as a **function of μ** .

Enumerating graphs of 2-XORSAT.

We will **enumerate** the **connected graphs without cycles of odd weight** according to two parameters: **number of vertices n** and **number of edges $n + \ell$** . $\ell \stackrel{\text{def}}{=} \mathbf{excess}$.

Let

$$C_\ell(z) = \sum_{n>0} c_{n,n+\ell} \frac{z^n}{n!}.$$

What are the series C_ℓ ?

Enumerating graphs of 2-XORSAT.

We will **enumerate** the **connected graphs** without cycles of **odd weight** according to two parameters: **number of vertices** n and **number of edges** $n + \ell$. $\ell \stackrel{\text{def}}{=} \mathbf{excess}$.

Let

$$C_\ell(z) = \sum_{n>0} c_{n,n+\ell} \frac{z^n}{n!}.$$

What are the series C_ℓ ?

Th.

$$C_\ell(z) = \frac{1}{2} W_\ell(2z)$$

with $W_\ell =$ Exponential generating functions of connected graphs
WRIGHT (1977).

- **Rooted and unrooted trees** (excess = -1)

$$T(z) = ze^{2T(z)} = \sum_{n>0} (2n)^{n-1} \frac{z^n}{n!}, \quad C_{-1}(z) = T - T^2.$$

- **Unicyclic components** (excess = 0)

- 1 Number of labellings of a *smooth* cycle (i.e. without vertices of degree 1) using $n > 2$ vertices :

$$\frac{2^n n!}{2n}.$$

- 2 Thus, the EGF of smooth unicyclic components

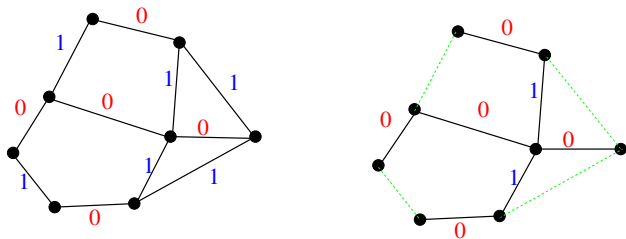
$$\tilde{C}_0(z) = -\frac{1}{4} \log(1 - 2z) - z/2 - z^2/2.$$

- 3 Substituting each vertex with a full rooted tree, we get

$$C_0(z) = -\frac{1}{4} \log(1 - 2T) - T/2 - T^2/2.$$

- **What about multicyclic components?** (excess > 0)

Enumerations: connected multicyclic components



On a connected “SAT”-graph with n vertices and $n + \ell$ edges, the edges of a spanning tree can be colored in 2^{n-1} ways. The colors of the other edges are “determined”.

Enumerations: general multicyclic components

Let $F_r(z)$ be the EGF of all complex weighted labelled graphs (**connected** or **not**), with a positive *total excess*¹ r and without cycles of odd weight (“SAT-graph”).

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

and for any $r \geq 1$

$$rF_r(z) = \sum_{k=1}^r k \frac{W_k(2z)}{2} F_{r-k}(z), \quad F_0(z) = 1.$$

Since $W_k(x) \asymp \frac{w_k}{(1-T(x))^{3r}}$ [WRIGHT 80], we also have $F_k(x) \asymp \frac{f_k}{(1-T(2x))^{3r}}$ with $2rf_r = \sum_{k=1}^r kb_k f_{r-k}$, $r > 0$.

¹total excess of the random graphs $\stackrel{\text{def}}{=} \text{nbr of edges} + \text{number of trees} - \text{number of vertices}$

Th.

The probability that a random formula with n variables and m clauses is SAT satisfies the following :

(i) **Sub-critical phase** : As $0 < n - 2m \ll n^{2/3}$,

$$\Pr(n, m) = e^{m/2n} \left(1 - 2\frac{m}{n}\right)^{1/4} + O\left(\frac{n^2}{(n-2m)^3}\right).$$

(ii) **Critical phase** : As $m = \frac{n}{2} + \mu n^{2/3}$, $\mu \in \mathbb{R}$ fixed

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

where Ψ can be expressed in terms of the Airy function.

(iii) **Super-critical phase** : As $m = \frac{n}{2} + \mu n^{2/3}$ with $\mu = o(n^{1/12})$

$$\Pr\left(n, \frac{n}{2}(1 + \mu n^{-1/3})\right) = \text{Poly}(n, \mu) e^{-\frac{\mu^3}{6}}.$$

Proof of (i) : the sub-critical phase

- 1 As $0 < n - 2m \ll n^{2/3}$, the probability that a Erdős-Rényi random graph $\mathbb{G}(n, m)$ has NO MULTICYCLIC COMPONENTS is

$$1 - o\left(\frac{n^2}{(n-2m)^3}\right) \begin{cases} \text{if } m = cn \text{ with } \limsup c < 1/2, \text{ BigOh} = O(1/n) \\ \text{if } m = \frac{n}{2} - \mu(n)n^{2/3}, \text{ BigOh} = O(1/\mu^3) \end{cases}$$

- 2 Then, the probability that the graph associated to random 2-XORSAT formula is SAT (conditionally that there is no multicyclic components) is given by

$$\frac{n!}{\binom{n(n-1)}{m}} [z^n] \underbrace{\frac{C_{-1}(z)^{n-m}}{(n-m)!}}_{\text{unrooted trees}} \times \underbrace{e^{C_0(z)}}_{\text{set of even weighted unicyclic components}}$$

$$m \leq \frac{n}{2} - \mu n^{2/3}, \quad 1 \ll \mu$$

- 1 Cauchy integral formula leads to

$$\text{coeff}(n, m) \times \frac{1}{2\pi i} \oint \frac{e^{-T(2z)/4 - T(2z)^2/8}}{(1 - T(2z))^{1/4}} \left(\frac{T(2z)}{2} - \frac{T(2z)^2}{4} \right)^{n-m} \frac{dz}{z^{n+1}}$$

- 2 “Lagrangian” substitution $u = T(2z)$.

3

$$\text{coeff}(n, m) \times \frac{1}{2\pi i} \oint g(u) \exp(nh(u)) du$$

- 4 $h(u) = u - \frac{m}{n} \log u + (1 - \frac{m}{n}) \log(2 - u)$.

Saddle-points at $u_0 = 2m/n < 1$ and $u_1 = 1$.

$$h''(1) = 2m/n - 1 < 0 \text{ and } h''(2m/n) = \frac{n(n-2m)}{4m(n-m)} > 0.$$

Saddle-point method applies on circular path $|z| = 2m/n \dots$

Proof of (ii) : Inside the critical phase (1/2)

$$m = \frac{n}{2} \pm \mu n^{2/3}, \quad |\mu| = O(n^{1/12})$$

Some MULTICYCLIC COMPONENTS (can) appear and the general formula for the integral becomes

1

$$\text{coeff}(n, m, r) \times \frac{1}{2\pi i} \oint \frac{e^{-T(2z)/4 - T(2z)^2/8}}{(1 - T(2z))^{1/4+3r}} \left(\frac{T(2z)}{2} - \frac{T(2z)^2}{4} \right)^{n-m+r} \frac{dz}{z^{n+1}}$$

2

$$\text{coeff}(n, m, r) e^n \times \frac{1}{2\pi i} \oint g_r(u) \exp(nh(u)) du$$

3

$$h(u) = u - 1 - \frac{m}{n} \log u + \left(1 - \frac{m}{n}\right) \log(2 - u).$$

Saddle-points at $u_0 = 2m/n = 1 + 2\mu n^{-1/3}$ and $u_1 = 1$.

BUT at the critical point $m = 2n$ ($\mu = 0$), we have $u_0 = u_1 = 1$ with triple zero

$$h(1) = h'(1) = h''(1) = 0.$$

Integral representation on the complex plane

The Airy function is given by

$$\text{Ai}(z) = \frac{1}{2\pi i} \int_C \exp\left(\frac{t^3}{3} - zt\right) dt,$$

where the integral is over a path C starting at the point at infinity with argument $-\pi/3$ and ending at the point at infinity with argument $\pi/3$.

Integral representation on the complex plane

The Airy function is given by

$$\text{Ai}(z) = \frac{1}{2\pi i} \int_C \exp\left(\frac{t^3}{3} - zt\right) dt,$$

where the integral is over a path C starting at the point at infinity with argument $-\pi/3$ and ending at the point at infinity with argument $\pi/3$.

Well suited for our purpose (see also [FLAJOLET, KNUTH, PITTEL 89], [JANSON, KNUTH, ŁUZCAK, PITTEL 93], [FLAJOLET, SALVY, SCHAEFFER 02], [BANDERIER, FLAJOLET, SCHAEFFER, SORIA 01])!

Integrating on a path $\mathbf{z} = \mathbf{e}^{-(\alpha+it)n^{-1/3}}$, we get

$$\frac{e^{-\mu^3/6-n}}{2^{2m-n-2r}} \times \frac{1}{2\pi i} \oint \frac{e^{-T(2z)/4 - T(2z)^2/8}}{(1 - T(2z))^{1/4+3r}} \left(\frac{T(2z)}{2} - \frac{T(2z)^2}{4} \right)^{n-m+r} \frac{dz}{z^{n+1}}$$

$$\sim e^{-3/8} A(1/4 + 3r, \mu) n^{r-7/12},$$

$$\text{where } A(y, \mu) = \frac{e^{-\mu^3/6}}{3^{(y+1)/3}} \sum_{k \geq 0} \frac{\left(\frac{1}{2} 3^{2/3} \mu\right)^k}{k! \Gamma((y+1-2k)/3)}$$

Define $p_r(n, m) =$ Proba to have SAT-graph of excess r . The proba.

that a random formula is given by $p(n, m) = \sum_{r \geq 0} p_r(n, m)$.

The proof of part (ii) can now be completed by means of the following facts

- 1 Using the Airy stuff, we compute for fixed r

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

- 2 Bounding the magnitude of the integral, it can be proved that there exist $R, C, \epsilon > 0$ such that for all $r \geq R$ and all n :

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

(dominated convergence theorem applies).

Remark

- On the first hand, writing $m = \frac{n}{2} - \mu n^{2/3}$ the probability is about :

$$e^{m/2n} \left(1 - \frac{2m}{n}\right)^{1/4} \sim e^{1/4} \mu^{1/4} n^{-1/12}.$$

- On the other hand, the Airy stuff are valid for $m = \frac{n}{2} + \mu n^{2/3}$, $|\mu| = O(n^{1/12})$. Using

$$A(r, \mu) = \frac{1}{\sqrt{2\pi} |\mu|^{y-1/2}} \left(1 - \frac{3y^2+3y-1}{6|\mu|^3} + O(|\mu|^{-6})\right) \text{ as } \mu \rightarrow -\infty \text{ we get}$$

$$\sum_r p_r(n, m) \sim n^{-1/12} \left(\sum_{r=0}^{\infty} \frac{\sqrt{2\pi} e^{1/4} f_r}{2^r} A(3r + 1/4, \mu) \right) \sim e^{1/4} \mu^{1/4} n^{-1/12}.$$

For the case (iii) of the theorem, we use

$$A(y, \mu) = \frac{e^{-\mu^{3/6}}}{2^{y/2} \mu^{1-y/2}} \left(\frac{1}{\Gamma(y/2)} + \frac{4\mu^{-3/2}}{3\sqrt{2}\Gamma(y/2 - 3/2)} + O(\mu^{-2}) \right).$$

Random MAX-2-XORSAT

- **MAX-2-XORSAT** is an NP-optimization problem (NPO). The corresponding decision problem is in NP (deciding if the size of the MAX is k ...).
- **MAX/MIN** problems are interesting (and difficult) in randomness context.
- **PREVIOUS WORKS** : [COPPERSMITH, GAMARNIK, HAJIAGHAYI, SORKIN 04] **Expectations** of the **Maximum** number of satisfiable clauses in MAX-2-SAT and MAX-CUT for the subcritical phases. **Bounds** of these expectations for some cases (namely for the critical and supercritical phases of random graphs)!
- **OUR WORK** : Quantification of the **Minimum** number of clauses to remove in order to get satisfiable formula.

Th.

Let $X_{n,m}$ be the minimum number of clauses UNSAT in a random 2-XOR formula with n variables and m clauses. We have :

(i) **Sub-critical phase** : If $\limsup \frac{m}{n} < 1/2$ then

$$X_{n,m} \xrightarrow{\text{dist.}} \text{Poisson} \left(\frac{\log n - 3 \log \left(\frac{n-2m}{n^{2/3}} \right) - 3 \left(1 - \frac{2m}{n} \right)}{12} \right).$$

(ii) **Critical phase** : If $m = \frac{n}{2}(1 - \mu n^{-1/3})$, $1 \ll \mu \ll n^{1/3}$ then

$$\mathbb{P} \left(X_{n,m} - \frac{1}{4} \log(\mu n^{-1/3}) \leq x \sqrt{\frac{1}{4} \log(\mu n^{-1/3})} \right) \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du$$

(iii) **Supercritical phase** : If $m = \frac{n}{2} + \mu n^{2/3}$ with $\mu = o(n^{1/3})$ (resp. $m = \frac{n}{2}(1 + \varepsilon)$)

$$\frac{6 X_{n,m}}{\frac{(2m-n)^3}{n^2} + O(\log n)} \xrightarrow{\text{dist.}} \mathbf{1}. \quad (\text{resp. } \frac{8(1+\varepsilon)}{n(\varepsilon^2 - \sigma^2)} X_{n,m} \xrightarrow{\text{dist.}} \mathbf{1}.)$$

where σ is the solution of $(1 + \varepsilon)e^{-\varepsilon} = (1 - \sigma)e^{\sigma}$.

- $X_{n,m}$: minimum number of UNSAT clauses in random formula with n variables and m clauses.
- $Y_{n,m}$: minimum number of clauses to suppress in unicyclic components.
- $Z_{n,m}$: minimum number of clauses to suppress in multicyclic components.

$$X_{n,m} = Y_{n,m} + Z_{n,m}.$$

Proof of the sub-critical phase

In the sub-critical random graphs, we know that $Z_{n,m} = O_p(1)$.

- if $m = cn$, $c \in]0, \frac{1}{2}[$ $\forall R$ fixed, we have

$$\Pr(\mathbf{Y}_{n,m} = \mathbf{R}) = e^{-\alpha(\mathbf{c})} \frac{\alpha(\mathbf{c})^{\mathbf{R}}}{\mathbf{R}!} \left(1 + O\left(\frac{1}{n}\right)\right).$$

- If $m = \frac{n}{2}(1 - \mu n^{-1/3})$ with $\mu \rightarrow \infty$ but $\mu = o(n^{1/3})$, we get $\forall R \leq 4\beta(n)$

$$\Pr(\mathbf{Y}_{n,m} = \mathbf{R}) = e^{-\beta(n)} \frac{\beta(n)^{\mathbf{R}}}{\mathbf{R}!} \left(1 + O\left(\frac{1}{\mu^3}\right)\right).$$

- There are $R_0, C, \varepsilon > 0$, s. t. $\forall R > R_0$

$$\Pr(\mathbf{Y}_{n,m} = \mathbf{R}) \leq C e^{-\varepsilon \mathbf{R}}.$$

with

$$\beta(n) = \frac{1}{12} \log(n) - \frac{1}{4} \log(\mu) - \frac{1}{4} + \frac{1}{4} \mu n^{-1/3}, \quad \alpha(\mathbf{c}) = -\frac{1}{4} \log(1 - 2\mathbf{c}) - \frac{\mathbf{c}}{2}$$

Corollary. As $n \rightarrow \infty$ and $m = cn, c \in]0, 1/2[$, for all $k \geq 1$,

$$\mathbf{E}(X_{n,m})_k \sim \alpha(\mathbf{c})^k.$$

If $m = \frac{n}{2}(1 - \mu n^{-1/3})$ but $\mu = o(n^{1/3})$

$$\mathbf{E}(X_{n,m})_k \sim \beta(\mathbf{n})^k.$$

with

$$\beta(\mathbf{n}) = \frac{1}{12} \log(\mathbf{n}) - \frac{1}{4} \log(\mu) - \frac{1}{4} + \frac{1}{4} \mu \mathbf{n}^{-1/3}$$

and

$$\alpha(\mathbf{c}) = -\frac{1}{4} \log(1 - 2\mathbf{c}) - \frac{\mathbf{c}}{2}$$

Theorem. As $n, m \rightarrow \infty$ and $m = \frac{n}{2} \pm O(1)n^{2/3}$, then for all k

$$E(X_{n,m})_k \sim \frac{1}{12^k} \log(n)^k.$$

- Here again, by [JANSON, KNUTH, ŁUCZAK, PITTEL 93] $Z_{n,m} = O_p(1)$.
- The k -th factorial moment of $Y_{n,m}$ is $\sum_r s_{r,k}$ where

$$s_{r,k} = \frac{1}{\binom{n(n-1)}{m}} \frac{n!}{2\pi i} \oint \frac{\partial^k}{\partial u^k} S_r(u, z)|_{u=1} \frac{dz}{z^{n+1}}$$

$$S_r(u, z) = \frac{(T(z) - T(z)^2)^{n-m+r}}{(n-m+r)!} \exp(C_{0,0}(z) + uC_{1,0}(z)) F_r(z),$$

$F_r(z)$: EGF of multicyclic components.

Lemma. As $\ell \rightarrow \infty$, the probability that the number of edges to suppress in order to obtain a (weighted) connected graph without cycles of odd weight from a (weighted) connected graph of excess ℓ is larger than

$$\frac{\ell}{4} - o(\ell)$$

is at least

$$1 - e^{-o(\ell)} - e^{-4c(\ell)^2 + \frac{1}{2} \log(\ell)}$$

where $c(\ell)^2 \gg \log(\ell)$

Lemma. As $\ell \rightarrow \infty$, the probability that the number of edges to suppress in order to obtain a (weighted) connected graph without cycles of odd weight from a (weighted) connected graph of excess ℓ is larger than

$$\frac{\ell}{4} - o(\ell)$$

is at least

$$1 - e^{-o(\ell)} - e^{-4c(\ell)^2 + \frac{1}{2} \log(\ell)}$$

where $c(\ell)^2 \gg \log(\ell)$

To prove this lemma, we need another one!

Lower bound of the probability (super-critical phase)

Let $C_{s,\ell}$ be the EGFs of connected components of EXCESS ℓ and where **at LEAST** s edges have to be suppressed to obtain components without cycles of odd weight.

Lemma. For all $s \geq 0$, we have

$$C_{s,\ell}(\mathbf{z}) \prec \left(\sum_{i=s}^{2s} \binom{\ell+1}{i} \right) C_{0,\ell}(\mathbf{z})$$

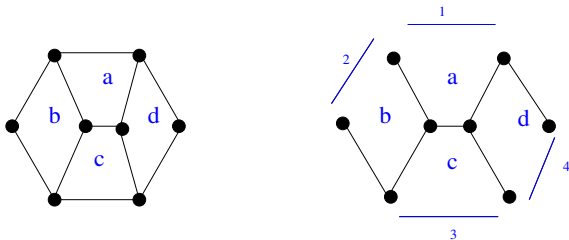
Lower bound of the probability (super-critical phase)

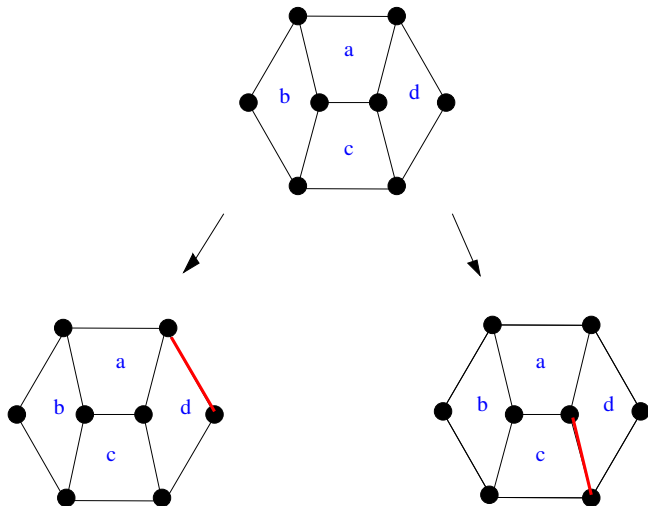
Let $C_{s,\ell}$ be the EGFs of connected components of EXCESS ℓ and where **at LEAST** s edges have to be suppressed to obtain components without cycles of odd weight.

Lemma. For all $s \geq 0$, we have

$$C_{s,\ell}(z) \prec \left(\sum_{i=s}^{2s} \binom{\ell+1}{i} \right) C_{0,\ell}(z)$$

Idea of the proof.





Lemma. If in a connected component of excess ℓ we have to suppress at least s edges to obtain a SAT-graph then this component has at most s fundamental and distinct cycles of **odd weight**.

Idea of the proof. Immediate.

As a crucial **consequence**, such a connected component has a **cactus** (as a subgraph) with at most s cycles of odd weight.

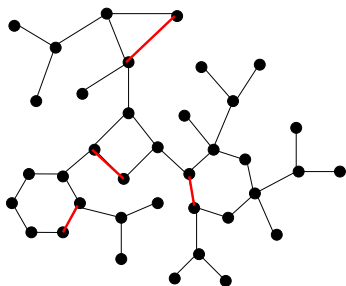
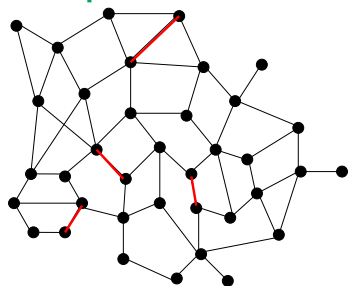
Upper-bound of the probability (super-critical phase)

Lemma. If in a connected component of excess ℓ we have to suppress at least s edges to obtain a SAT-graph then this component has at most s fundamental and distinct cycles of **odd weight**.

Idea of the proof. Immediate.

As a crucial **consequence**, such a connected component has a **cactus** (as a subgraph) with at most s cycles of odd weight.

Example.



Lemma. Let $\tilde{\Xi}_s(z)$ be the EGF of smooth cactii (Husimi trees) with s cycles, we have :

$$\partial_z \tilde{\Xi}_s + (s-1) \tilde{\Xi}_s = \frac{1}{2} \sum_{i=1}^{s-1} (\partial_z \tilde{\Xi}_i) (\partial_z \tilde{\Xi}_{s-i}) (\partial(P) - P) + \sum_{k=1}^{s-1} z^k \frac{\partial^k}{\partial z^k} \partial_z \tilde{\Xi}_1$$

$$\times \sum_{\substack{l_1+2l_2+\dots+(s-1)l_{s-1}=s-1 \\ l_1+l_2+\dots+l_{s-1}=k, l_j \in \mathbb{N}}} \frac{(\partial_z \tilde{\Xi}_1)^{l_1}}{l_1!} \dots \frac{(\partial_z \tilde{\Xi}_{s-1})^{l_{s-1}}}{l_{s-1}!} \left(\frac{1}{z} + \frac{P}{z^2} \right)^k$$

, with $P \equiv P(z) = \frac{z^2}{1-z}$.

Lemma. We have

$$\Xi_s(z) \preceq \frac{\xi_s}{(1-t(z))^{3s-3}}, \quad s > 1$$

where $(\xi_s)_{s>1}$ satisfies $\xi_2 = \frac{1}{8}$, $\xi_3 = \frac{1}{12}$ and for $s \geq 3$, we have :

$$3(s-1)\xi_s = \frac{3}{2}(s-2)\xi_{s-1} + \frac{9}{2} \sum_{i=2}^{s-2} (i-1)(s-i-1)\xi_i \xi_{s-i} +$$

$$\frac{1}{2} \sum_{k=1}^{s-1} k! \left(\sum_{\substack{l_1+2l_2+\dots+(s-1)l_{s-1}=s-1 \\ l_1+l_2+\dots+l_{s-1}=k}} \frac{\left(\frac{1}{2}\right)^{l_1} (3\xi_2)^{l_2} \dots (3(s-2)\xi_{s-1})^{l_{s-1}}}{l_1! l_2! \dots l_{s-1}!} \right)$$

Lemma. We have

$$\Xi_s(z) \preceq \frac{\xi_s}{(1-t(z))^{3s-3}}, \quad s > 1$$

where $(\xi_s)_{s>1}$ satisfies $\xi_2 = \frac{1}{8}$, $\xi_3 = \frac{1}{12}$ and for $s \geq 3$, we have :

$$3(s-1)\xi_s = \frac{3}{2}(s-2)\xi_{s-1} + \frac{9}{2} \sum_{i=2}^{s-2} (i-1)(s-i-1)\xi_i \xi_{s-i} +$$

$$\frac{1}{2} \sum_{k=1}^{s-1} k! \left(\sum_{\substack{l_1+2l_2+\dots+(s-1)l_{s-1}=s-1 \\ l_1+l_2+\dots+l_{s-1}=k}} \frac{\left(\frac{1}{2}\right)^{l_1} (3\xi_2)^{l_2} \dots (3(s-2)\xi_{s-1})^{l_{s-1}}}{l_1! l_2! \dots l_{s-1}!} \right)$$

Lemma. As $s \rightarrow \infty$,

$$\xi_s = \frac{1}{6} \left(\frac{3}{2}\right)^{s-1} \frac{3^{s/2}}{\sqrt{2\pi s^3(s-1)}} \left(1 + \mathbf{O}\left(\frac{1}{s}\right)\right).$$

Corollary. The number of connected component of excess ℓ obtained by adding edges from cactii with s cycles can be neglected if

$$s > \frac{\ell}{2} + O\left(\frac{\ell}{\log(\ell)}\right).$$

Corollary. The number of connected component of excess ℓ obtained by adding edges from cactii with s cycles can be neglected if $s > \frac{\ell}{2} + O\left(\frac{\ell}{\log(\ell)}\right)$.

Idea of the proof.

- Pick a cactus with s cycles.
- Add $(\ell - s)$ edges to obtain a connected component of excess ℓ . The number of such constructions can be bounded by **pointing/depoining** the last added edge.
- The ratio of the number these objects over the number of all connected components of excess ℓ is exponentially small as $s > \frac{\ell}{2} + O(\ell/\log \ell)$.

- 1 On connected components of excess ℓ the number of edges to suppress lies w.h.p. between

$$\frac{\ell}{4} - O(\ell^{1/2}) \leq |\text{suppressions}| \leq \frac{\ell}{4} + O\left(\frac{\ell}{\log \ell}\right).$$

- 2 Now, we can use the result on random graphs from [PITTEL, WORMALD 05] to complete the proof of the theorem.

Conclusion and perspectives

Enumerative/Analytic approaches of

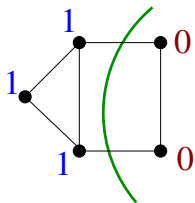
- 1 a decision problem and its phase transition
- 2 an NP-optimization problem.

Enumerative/Analytic approaches of

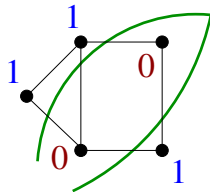
- 1 a **decision problem** and its **phase transition**
- 2 an **NP-optimization** problem.

Similar **methods** on other problems such as

- 1 bipartiness (or 2-COL),
- 2 MAX-2-COL, MAX-CUT, MIN-VERTEX-COVER, MIN-BISECTION
(all are **hard optimization problems** related to bipartiteness/2-COL).
- 3 2-QXORSAT (quantified formula).

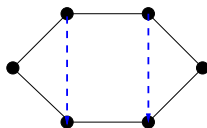
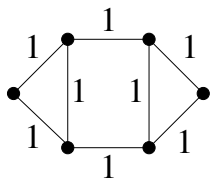


CUT



MAX-CUT

Graph \longrightarrow MAX-2-XORSAT



MAX-CUT

