Abstract

We consider a finite relational vocabulary σ and a first order theory T for σ composed of symmetry and anti-reflexivity axioms. We define a binomial random model of finite σ-structures that satisfy T and show that first order properties have well defined asymptotic probabilities in the sparse case. It is also shown that those limit probabilities are well-behaved with respect to some parameters that represent edge densities. An application of these results to the problem of random Boolean satisfiability is presented afterwards. We show that there is no first order property of k-CNF formulas that implies unsatisfiability and holds for almost all typical unsatisfiable formulas when the number of clauses is linear.

Introduction

Since the work of Erdös and Rènyi on the evolution of random graphs [1] the study of the asymptotic properties of random structures has played a relevant role in combinatorics and computer science. A central theme in this topic is, given a succession $(G_n)_n$ of random structures of some sort and a property P, to determine the limit probability that G_n satisfies P or to determine whether that limit exists.

One approach that has proven to be useful is to classify the properties P according to the logical languages they can be defined in. We say that the succession $(G_n)_n$ obeys a convergence law with respect to some logical language $\mathscr{L}$ if for any given property P expressible in $\mathscr{L}$ the probability that G_n satisfies P tends to some limit as n grows to infinity. We say that $(G_n)_n$ obeys a zero-one law with respect to $\mathscr{L}$ if that limit is always either zero or one. The seminal theorem on this topic, due to Fagin [2] and Glebskii et al. [3] independently, states that if G_n denotes a labeled graph with n vertices picked uniformly at random among all $2^{\binom{n}{2}}$ possible then $(G_n)_n$ satisfies a zero-one law with respect to the first order (FO) language of graphs.

Originally this result was proven in the broader context of relational structures but it was in the theory of random graphs were the study of other zero-one and convergence laws became more prominent. In particular, the asymptotic behavior of FO logic in the binomial model of random graphs $G(n, p)$ has been extensively studied. In this model, introduced by Gilbert [4], a random graph is obtained from n labeled vertices by adding each possible edge with probability p independently. When $p = 1/2$ this distribution of random graphs coincides with the uniform one, mentioned above. In general, for the case where p is a constant probability a slight generalization of the proofs in [2] and [3] works and $G(n, p)$ satisfies a zero-one law for FO logic. If we consider $p(n)$ a decreasing function of the form $n^{-\alpha}$ we can ask the question of what are the values of α for which $G(n, p(n))$ obeys a zero-one or a convergence law for FO logic. In [5] Shelah and Spencer gave a complete answer for the range $\alpha \in (0, 1)$. Among other results, they proved that if α is an irrational number in this interval then $G(n, p(n))$ obeys a zero-one law for FO logic, while if α is a rational number in the same range then $G(n, p(n))$ does not even satisfy a convergence law for FO logic. The case $\alpha = 1$ was later solved by Lynch in [6]. A weaker form of the main theorem in that article states the following:

Theorem 0.1. *For any FO sentence ϕ, the function $F_\phi : (0, \infty) \to [0, 1]$ given by*

$$F_\phi(\beta) = \lim_{n\to\infty} \Pr\big(\mathrm{G}(n, \beta/n) \textit{ satisfies } \phi\big)$$

is well defined and analytic. In particular, for any $\beta \geq 0$ the model $\mathrm{G}(n, \beta/n)$ obeys a convergence law for FO logic.

The analyticity of these asymptotic probabilities with respect to the parameter β implies that FO properties cannot "capture" sudden changes that occur in the random graph $\mathrm{G}(n, \beta/n)$ as β changes. Given $p(n)$ a probability, P a property of graphs, and Q a sufficient condition for P - i.e., a property that implies P -, we say that Q explains P if $\mathrm{G}(n, p(n))$ satisfies the converse implication $P \implies Q$ asymptotically almost surely (a.a.s.). A notable example of this phenomenon happens in the range $p(n) = \log(n)/n + \beta/n$ with β constant. Erdös and Rènyi [1] showed that for probabilities of this form $\mathrm{G}(n, p(n))$ a.a.s. is disconnected only if it contains an isolated vertex. An observation by Albert Atserias is the following:

Theorem 0.2. *Let c be a real constant such that $\lim_{n\to\infty} \Pr\big(G(n, c/n)$ is not 3-colorable $\big) > 0$. Then there is no FO graph property that explains non-3-colorability for $G(n, c/n)$.*

The short proof of this theorem is as follows: It is a known fact that there are positive constants $c_0 \leq c_1$ such that $G(n,c/n)$ is a.a.s 3-colorable if $c < c_0$ and it is a.a.s non 3-colorable if $c > c_1$ REFERENCES NEEDED. Suppose that P is a FO graph property that implies non-3-colorability. Then, because of this implication, for all values of c

$$\lim_{n\to\infty} \Pr\big(G(n,c/n) \text{ satisfies } P\big) \leq \lim_{n\to\infty} \Pr\big(G(n,c/n) \text{ is not 3-colorable}\big).$$

In consequence the asymptotic probability that $G(n,c/n)$ satisfies P is zero when $c < c_0$. By Lynch's theorem, if P is definable in FO logic then this asymptotic probability varies analytically with c. Using the fact that any analytic function that takes value zero in a non-empty interval must equal zero everywhere, we obtain that $G(n,c/n)$ a.a.s does not satisfy P for any value of c. As a consequence the theorem follows.

The aim of this work is to extend Lynch's result to arbitrary relational structures were the relations are subject to some predetermined symmetry and anti-reflexivity axioms. This was originally motivated by an application to the study of random k-CNF formulas. Since [7] it is known that for each k there are constants c_0, c_1 such that a random k-CNF formula with cn clauses over n variables

1 Preliminaries

1.1 General notation

Given a positive natural number n, we will write $[n]$ to denote the set $1,2,\dots,n$.

Given a set S and a natural number $k \in \mathbb{N}$ we will use $\binom{S}{k}$ to denote the set of subsets of S whose size is k.

Given numbers, $n,m \in \mathbb{N}$ with $m \leq n$ we define $(n)_m := n\cdot(n-1)\cdots(n-m+1)$. Given a set S and a number $n \in \mathbb{N}$ with $n \leq |S|$ we define $(S)_n$ as the subset of S^n consisting of the n-tuples whose coordinates are all different. We also define $S^* := \bigcup_{n=0}^{\infty} S^n$ and $(S)_* = \bigcup_{n\leq|S|}(S)_n$. Given a tuple $\overline{x} \in S^*$ and an element $x \in S$ the expression $x \in \overline{x}$ will mean that x appears as some coordinate in $\overline{x}$. We will at times make an abuse of notation and treat the tuple $\overline{x}$ as the set of elements $x \in \overline{x}$.

We will use the convention that over-lined variables , like $\overline{x}$, denote ordered tuples of arbitrary length. For example, given a set S, if we write $\overline{x} \in S^*$ then $\overline{x} = (x_1,\dots,x_a)$ for some $a \in \mathbb{N}$ and some $x_1,\dots,x_a \in S$. Given an ordered tuple $\overline{x}$ we define the number $len(\overline{x})$ as its length. Given a map $f : X \to Y$ between two sets X,Y and an ordered tuple $\overline{x} := (x_1,\dots,x_a) \in X^*$ we define $f(\overline{x}) \in Y^*$ as the tuple $(f(x_1),\dots,f(x_a))$.

Let S be a set, a a positive natural number, and Φ a group of permutations over $[a]$. Then Φ acts naturally over S^a in the following way: Given $g \in \Phi$ and $\overline{x} = (x_1,\dots,x_a) \in S^a$ we define $g\cdot(x_1,\dots,x_a)$ and $g\cdot\overline{x}$ as the tuple $(x_{g(1)},\dots,x_{g(a)})$. We will denote by S^a/Φ to the quotient of the set S^a by this action. Given an element $\overline{x} := (x_1,\dots,x_a) \in S^a$ we will denote its equivalence class in S^a/Φ by $[x_1,\dots,x_a]$ or $[\overline{x}]$. Thus, for any $g \in \Phi$, by definition $[x_1,\dots,x_a] = [x_{g(1)},\dots,x_{g(a)}]$. The notations $\overline{x}$ and $(x_1,\dots,x_a)$ will be reserved to ordered tuples while $[\overline{x}]$ and $[x_1,\dots,x_a]$ will denote ordered tuples modulo the action of some arbitrary group of permutations. Which group is this will depend on the ambient set where $[x_1,\dots,x_a]$ belongs and it should either be clear from context or not be relevant.

Given two real functions over the natural numbers $f,g:\mathbb{N}\to\mathbb{R}$ we will write $f=O(g)$ to mean that there exists some constant $C\in\mathbb{R}$ such that $f(n)\le Cg(n)$ for n sufficiently large, as usual. If $g(n)\neq 0$ for n large enough then we will write $f\sim g$ when $\lim_{n\to\infty}\frac{f(n)}{g(n)}=1$.

1.2 Logical preliminaries

We will assume a certain degree of familiarity with the concepts. For a more complete exposition of the topics presented here one can consult [8].

A relational vocabulary σ is a collection of relation symbols $\{R_1,\dots,R_m,\dots\}$ where each relation symbol R_i has associated a natural number $ar(R_i)$ called its arity. A σ-structure $\mathfrak{A}$ is composed of a set A, called the universe of $\mathfrak{A}$, equipped with relations $R_1^{\mathfrak{A}}\subseteq A^{a_1},\dots,R_m^{\mathfrak{A}}\subseteq A^{a_m}$. When σ is understood we may refer to σ-structures as relational structures or simply as structures. A structure is called finite if its universe is a finite set.

In the first order language $FO[\sigma]$ with signature σ formulas are formed by variables $x_1,\dots,$ $x_i,\dots$, the relation symbols in σ, the equal symbol $=$, the usual Boolean connectives $\neg,\wedge,\vee,\dots$, the existential and universal quantifiers $\exists,\forall$, and the parentheses $),($. Then formulas in $FO[\sigma]$ are defined as follows.

- The expression $R(y_1,\dots,y_a)$, where the y_i's are variables and R is a relation symbol in σ such that $ar(R)=a$, belongs to $FO[\sigma]$.
- The expression $y_1=y_2$, where y_1,y_2 are variables, belongs to $FO[\sigma]$.
- Given formulas $\phi,\psi\in FO[\sigma]$, any Boolean combination of them $\neg(\phi),(\phi\wedge\psi),(\phi\vee\psi),\dots$ belongs to $FO[\sigma]$ as well.
- Given a formula $\phi\in FO[\sigma]$ and x a variable that does not appear bounded by a quantifier in ϕ, the expressions $\forall x(\phi)$ and $\exists x(\phi)$ belong both to $FO[\sigma]$.

We will write $\forall y_1,y_2,\dots,y_m$ or simply $\forall\bar{y}$ instead of $\forall y_1,\forall y_2,\dots,\forall y_m$ and likewise for the quantifier $\exists$. Also, if $\bar{y}:=(y_1,\dots,y_a)$ is a tuple of variables we may write simply $R(\bar{y})$ instead of $R(y_1\dots,y_a)$.

For the remaining of this article we will reserve the names x,y,z for the variables in our first order formulas.

We define the set of free variables of a formula as usual. Given a formula $\phi\in FO[\sigma]$ we will use the notation $\phi(\bar{y})$ to denote that $\bar{y}$ is a tuple of (different) variables that contains all free variables in ϕ and none of its bounded variables, although it may contain variables which not appear in ϕ.

Formulas with no free variables are called sentences and formulas with no quantifiers are called open formulas.

The quantifier rank of a formula ϕ, denoted by $qr(\phi)$, is defined as the maximum number of nested quantifiers in ϕ.

Sentences in $FO[\sigma]$ are interpreted over σ-structures in the natural way. Given an structure $\mathscr{A}$, and a sentence $\phi\in FO[\sigma]$ we write $\mathscr{A}\models\phi$ to denote that $\mathscr{A}$ satisfies ϕ. If $\psi(\bar{y})$ is a formula, $\bar{a}$ are elements in the universe of $\mathscr{A}$, and $\bar{y}$ and $\bar{a}$ are lists of the same size, then we write $\mathscr{A}\models\psi(\bar{a})$ to mean that $\mathscr{A}$ satisfies ψ when the free variables in $\bar{y}$ are interpreted as the elements in $\bar{a}$ (i.e., the i-th element of $\bar{y}$ is interpreted as the i-th element of $\bar{a}$).

Deberá al menos mencionar un par de trabajos que estudien random k-SAT y propiedades a.a.s suficientes para no-satisfacibilidad

1.3 Structures as multi-hypergraphs

For the rest of the article consider fixed:

- A relational vocabulary σ such that all the relations $R \in \sigma$ satisfy $ar(R) \geq 2$.
- Groups $\overline{\Phi} = \{\Phi_R\}_{R\in\sigma}$ such that each Φ_R is consists of permutations on $[ar(R)]$ with the usual composition as its operation.
- Sets $\{P_R\}_{R\in\sigma}$ satisfying that for all $R \in \sigma$, $P_R \subseteq \binom{[ar(R)]}{2}$

We will only consider relational structures where the relations are of arity at least two. This restriction is not necessary, but it makes notation easier.

Quizás debería añadir un anexo dando alguna indicación sobre cómo tratar las relaciones unarias?

We define the class $\mathscr{C}$ as the class of σ-structures that satisfy the following axioms:

- *Symmetry axioms*: For each $R \in \sigma$ and each $g \in \Phi_R$:

$$\forall \overline{x} := x_1, \ldots, x_{ar(R)} \big(R(\overline{x}) \iff R(g \cdot \overline{x}) \big)$$

- *Anti-reflexivity axioms*: For each $R \in \sigma$ and $\{i, j\} \in P_R$

$$\forall x_1, \ldots, x_{ar(R)} \big((x_i = x_j) \implies \neg R(x_1, \ldots, x_{a_s}) \big)$$

We can think any structure G in $\mathscr{C}$ as a "multi-hypergraph" whose vertices are the elements of the universe of G. Each relation $R \in \sigma$, can be represented over G as an "edge set" formed by tuples vertices of size $ar(R)$ modulo the action of Φ_R. Furthermore, repetitions of vertices in the positions given by P_R are not allowed in these tuples.

The following definitions make this ideas formal. They depend on our choices of σ, $\{\phi_R\}_{R\in\sigma}$ and $\{P_R\}_{R\in\sigma}$ but as those are fixed we can allow ourselves to omit those dependencies for the sake of readability.

Definition 1.1. Let V be a set, and let $R \in \sigma$. We define the **total edge set over V given by R** as

$$E_R[V] = (V^{ar(R)}/\Phi_R) \quad \setminus \quad X,$$

where

$$X = \Big\{ [x_1, \ldots, x_{ar(R)}] \;\Big|\; x_1, \ldots, x_{ar(R)} \in V, \text{ and } x_i = x_j \text{ for some } \{i, j\} \in P_R \Big\}.$$

Also, we will say that the **sort** of the elements $e \in E_R[V]$ is R.

That is, $E_R[V]$ contains all the "$ar(R)$-tuples of elements in V modulo the permutations in ϕ_R" excluding those that contain some repetition of elements in the positions given by P_R.

The fact that the elements $e \in E_R(V)$ are of sort R is a technical detail introduced so that for any different relation symbols R_1 and R_2 it holds $E_{R_1}(V) \cap E_{R_2}(V) = \emptyset$ even in the case that $ar(R_1) = ar(R_2)$, $\Phi_{R_1} = \Phi_{R_2}$ and $P_{R_1} = P_{R_2}$.

In the case where $V = [n]$ we will write simply $E_R[n]$ instead of $E_R[[n]]$.

Definition 1.2. We call $\mathscr{C}$-**hypergraph**, or simply hypergraph, to a pair $G = \Big(V(G), \{E_R(G)\}_{R\in\sigma}\Big)$, where for each R, $E_R(G) \subseteq E_R[V]$.

Hypergraphs, as we have defined them, can be naturally interpreted as structures from $\mathscr{C}$ in the following way: given $G = (V, \{E_R\}_{R\in\sigma})$, we consider V to be the universe of G, and for any $R \in \sigma$ we define $R^G \subseteq V^{ar(R)}$ as the set of tuples $\bar{x} \in V^{ar(R)}$ such that $[\bar{x}] \in E_R$. Under this interpretation hypergraphs satisfy by definition the symmetry and anti-reflexivity axioms given at the beginning of this section. It is also easy to see that this interpretation induces a one-to-one identification between structures in $\mathscr{C}$ and hypergraphs.

1.4 Hypergraph notation and nomenclature

We will use standard nomenclature and notation from graph theory with some additions. Given an hypergraph G we will call its **vertex set** to $V(G)$ and **vertices** to the elements $v \in V(G)$. Likewise, each of the $E_R(G)$'s will be called an **edge set** and its elements, **edges**. Given an edge set $E_R(G)$, the index R will be called its **relation**.

Given an hypergraph G we define the set $E(G)$ as the union $\cup_{R\in\sigma} E_R(G)$. Notice that this union is disjoint because elements from different $E_R(G)$'s are of different sorts. Thus, $|E(G)| = \sum_{R\in\sigma} |E_R(G)|$. Analogously, given a set V, we define $E[V] := \cup_{R\in\sigma} E_R[V]$. Given an edge $e \in E[V]$ we define $R(e)$ as the sort of e, i.e., the unique relation symbol $R(e) \in \sigma$ such that $e \in E_{R(e)}[V]$.

Given two hypergraphs H and G we say that H is a **sub-hypergraph** of G, which we write as $H \subset G$, if $V(H) \subset V(G)$ and $E(H) \subset E(G)$ (notice that this is equivalent to $E_R(H) \subset E_R(G)$ for all $R \in \sigma$, since the edges are sorted).

Given a set of vertices $U \subseteq V(G)$, we will denote by $G[U]$ the **hypergraph induced by** G **on** U. That is, $G[U]$ is an hypergraph $H = (V(H), \{E(H)_R\}_{R\in\sigma})$ such that $V(H) = U$ and for any $R \in \sigma$ each edge $e \in E_R(G)$ belongs to $E_R(H)$ as well if and only if the vertices involved in e are in U (i.e. $e \in E_R[U]$).

We define the **excess** $ex(G)$ of an hypergraph G as the number

$$ex(G) := \big(\sum_{R\in\sigma} (ar(R)-1)|E_R(G)|\big) - |V(G)|.$$

That is, the excess of G is its "weighted number of edges" minus its number of vertices.

Given an hypergraph G we define the following metric, d, over $V(G)$:

$$d^G(u,v) = \min_{\substack{H\subset G \\ H \text{ connected} \\ u,v\in V(H)}} |E(H)|.$$

That is, the distance between v and u is the minimum number of edges necessary to connect v and u. If such number does not exist we define $d^G(u,v) = \infty$.

As usual, define $d^G(X,Y)$ for sets $X,Y \subset V(G)$ as the minimum distance $d(u,v)$ where $u \in X$ and $v \in Y$. When $X = \{x\}$ we will omit the brackets and write $d^G(x,Y)$ instead of $d^G(\{x\},Y)$, for example. When G is understood or not relevant we will usually simply denote the distance by d instead of d^G.

Given set of vertices vertex, $X \subseteq V(G)$, we denote by $N^G(X;r)$ the r-**neighborhood** of X in G. That is, $N^G(X;r) = G[Y]$, where $Y \subseteq V(G)$ is the set:

$$Y := \{u \in V(G) \mid d(X,u) \leq r\}.$$

In particular, when X is a singleton $\{v\}$, we will write $N^G(v;r)$ instead of $N^G(\{v\};r)$. As before, we will usually drop the "G" from our notation when G is understood or not relevant.

We will often write tuples of vertices instead of sets inside of $d(\cdot,\cdot)$ and $N(\cdot;r)$. In those cases we are treating those tuples as sets as specified in section 1.1.

1.5 Colored Hypergraphs and Copies.

Definition 1.3. Let Σ be a set. A Σ**-hypergraph** is a pair (G,χ) where G is an hypergraph and $\chi : V(G) \to \Sigma$ is a map called Σ**-coloring** of G.

Given two Σ-hypergraphs (G^1,χ_1) and (G^2,χ_2), and **isomorphism** between them is a bijection $f : V(G^1) \to V(G^2)$ satisfying that f is an hypergraph isomorphism between G^1 and G^2 as well as $\chi_2(f(v)) = \chi_1(v)$ for all $v \in V(G^1)$.

Given a Σ-hypergraph (G,χ), an **automorphism** of (G,χ) is an isomorphism from it into itself. We will denote by $Aut(G,\chi)$ the group of such automorphisms.

1.6 The random model

For each $R \in \sigma$ let p_R be a real number between zero and one. Let $\overline{p} := \{p_R\}_{R\in\sigma}$. The random model $G^{\mathscr{C}}(n,\overline{p})$ is the discrete probability space that assigns to each hypergraph G whose vertex set $V(G)$ is $[n]$ the following probability:

$$\Pr(G) = \prod_{R\in\sigma} p_R^{|E_R(G)|}(1-p_R)^{\left|E_R[n]\right|-\left|E_R(G)\right|}.$$

Equivalently, this is the probability space obtained by assigning to each edge $e \in E_R[n]$ probability p_R independently.

As in the case of Lynch's theorem, we are interested in the "sparse regime" of $G^{\mathscr{C}}(n,\overline{p})$, were the expected number of edges each color is linear. This is achieved when each of the p_R's are of the form $\beta_R/n^{ar(R)-1}$ for some non-negative real numbers $\{\beta_R\}_{R\in\sigma}$. From now on we will write $G_n\left(\{\beta_R\}_{R\in\sigma}\right)$ to denote a random sample of $G^{\mathscr{C}}\left(n,\left\{\frac{\beta_R}{n^{ar(R)-1}}\right\}_{R\in\sigma}\right)$. When the choice of $\{\beta_R\}_{R\in\sigma}$ is not relevant we will write G_n instead of $G_n\left(\{\beta_R\}_{R\in\sigma}\right)$.

Our goal is to prove the following theorem:

Theorem 1.1. *Let ϕ be a sentence in $FO[\sigma]$. Then the function $F_\phi : [0,\infty)^{|\sigma|} \to \mathbb{R}$ given by*

$$\{\beta_R\}_{R\in\sigma} \mapsto \lim_{n\to\infty} Pr\big(G_n\left(\{\beta_R\}_{R\in\sigma}\right) \models \phi\big)$$

is well defined and analytic.

1.7 Ehrenfeucht-Fraisse Games

Let G^1 and G^2 be hypergraphs. We define the k round Ehrenfeucht-Fraisse game on G^1 and G^2, denoted by $\textsc{Ehr}_k(G^1;G^2)$, as follows: The game is played between two players, Spoiler and Duplicator, and the number of rounds, k, is known for both from the start. At the beginning of each round Spoiler chooses a vertex from either $V(G^1)$ or $V(G^2)$ and Duplicator responds by choosing a vertex from the other set. Let us denote by v_i, resp. u_i the vertex from G^1, resp. from G^2, chosen in the i-th round, for $i \in [k]$. At the end of the k-th round Duplicator wins if the following holds:

- For any $i,j \in [k]$, $v_i = v_j \iff u_i = u_j$.
- Given a relation symbol $R \in \sigma$ and indices $i_1,\dots,i_{ar(R)} \in [k]$, $[v_{i_1},\dots,v_{i_{ar(R)}}] \in E_R(G^1) \iff [u_{i_1},\dots,u_{i_{ar(R)}}] \in E_R(G^2)$.

We define the equivalence relation $=_k$ between hypergraphs as follows: We say that $G^1 =_k G^2$ if for any sentence $\phi \in FO[\sigma]$ with $qr(\phi) \leq k$ then "$G^1 \models \phi$ if and only if $G^2 \models \phi$".

The following is satisfied:

Theorem 1.2 (Ehrenfeut, 9). *Let G^1 and G^2 be hypergraphs. Then Duplicator wins $\textsc{Ehr}_k(G^1;G^2)$ if and only if $G^1 =_k G^2$.*

Let $\overline{v} \in V(G^1)^*$, and $\overline{u} \in V(G^2)^*$ be lists of vertices of the same length, $l = len(\overline{v}) = len(\overline{u})$. We define the k round Ehrenfeucht-Fraisse game on G^1 and G^2 with initial position given by $\overline{v}$ and $\overline{u}$, denoted by $\textsc{Ehr}_k(G^1,\overline{v};G^2,\overline{u})$, the same way as $\textsc{Ehr}_k(G^1;G^2)$, but in this case the game has l extra rounds at the beginning where the vertices in $\overline{v}$ and $\overline{u}$ are played successively. After this, k more rounds are played normally.

We also define the k-round distance Ehrenfeucht-Fraisse game on G^1 and G^2, denoted by $d\textsc{Ehr}_k(G^1;G^2)$, the same way as $\textsc{Ehr}_k(G^1;G^2)$, but now, in order for Duplicator to win the game, the following additional condition has to be satisfied at the end of the game:

- For any $i,j \in [k]$, $d^{G^1}(v_i,v_j) = d^{G^2}(u_i,u_j)$.

Given $\overline{v} \in V(G^1)^*$, and $\overline{u} \in V(G^2)^*$ lists of vertices of the same length, we define the game $d\textsc{Ehr}_k(G^1,\overline{v};G^2,\overline{u})$ analogously to $\textsc{Ehr}_k(G^1,\overline{v};G^2,\overline{u})$.

1.8 Outline of the proof

We show now an outline of the proof.

We show that for any quantifier rank k there are some classes of hypergraphs $C^k_1,\dots,C^k_{n_k}$ such that

(1) a.a.s the rank k type of any two graphs in the same class coincide,
(2) a.a.s. any random graph belongs to some of them, and
(3) the limit probability of random graph belonging to any of them is an analytic expression on the parameters $\overline{\beta}$.

After this is archived the theorem follows easily.

The objective of next sections will be to define the classes $C_1,\dots,C_{n_k}$ and to show that they satisfy properties (1), (2) and (3).

Explicar esto mejor

2 Some winning strategies for Duplicator

The aim of this section is to show the winning strategy for Duplicator that is going to be used in our proofs.

Let G^1 and G^2 be hypergraphs, and let $V_1 := V(G^1), V_2 := V(G^2)$. Let $\overline{v} \in V_1^*, \overline{u} \in V_2^*$ be tuples of the same length. We say that $\overline{v}$ and $\overline{u}$ have k-**similar** r-neighborhoods, written as $(G^1,\overline{v}) \simeq_{k,r} (G^2\overline{u})$, if Duplicator wins $d\text{EHR}_k\big(N(\overline{v};r),\overline{v};N(\overline{u};r),\overline{u}\big)$. Given sets of vertices $X \subseteq V_1$ and $Y \subseteq V_2$ we say that X and Y have k-similar r-neighborhoods, written as $X \simeq_{k,r} Y$, if we can order their elements to form lists $\overline{v}$, resp. $\overline{u}$ such that $(G^1,\overline{v}) \simeq_{k,r} (G^2,\overline{u})$. Given sets of vertices $X \in V_1$, $Y \in V_2$ and tuples of the same length $\overline{v} \in V_1^*$ and $\overline{u} \in V_2^*$ we will say that $(X,\overline{v})$ and $(Y,\overline{u})$ have k-similar r-neighborhoods, written as $\big(G_1,(X,\overline{v})\big) \simeq_{k,r} \big(G_2,(Y,\overline{u})\big)$, if the elements of X and Y can be ordered in lists $\overline{w}$, $\overline{z}$ such that $(G_1,\overline{w}^\frown\overline{v}) \simeq_{k,r} (G_2,\overline{z}^\frown\overline{u})$.

Fix $r \in \mathbb{N}$. Suppose that $X \subseteq V_1$ and $Y \subseteq V_2$ can be partitioned into sets $X = X_1 \cup \cdots \cup X_a$ and $Y = Y_1 \cup \cdots \cup Y_b$ such that $N(X_i;r)$'s, and the $N(Y_i;r)$'s, are connected and disjoint. We say that X and Y have k-**agreeable** r-neighborhoods, written as $(G^1,X) \cong_{k,r} (G^2,Y)$, if for any set $Z \subset V_\delta$, with $\delta \in \{1,2\}$, among the X_i's or the Y_i's it is satisfied that "the number of X_i's such that $(G^\delta,Z) \simeq_{k,r} (G^1,X_i)$ and the number of Y_i's such that $(G^\delta,Z) \simeq_{k,r} (G^2,Y_i)$ are both equal or are both greater or equal than k" .

The main theorem of this section,which is a slight strengthening of Theorem 2.6.7 from [10], is the following:

Theorem 2.1. *Set $r = (3^k-1)/2$. Let G^1,G^2 be hypergraphs and let $V_1 := V(G^1)$, $V_2 := V(G^2)$. Suppose there exist sets $X \subseteq V_1$, $Y \subseteq V_2$ with the following properties:*

(1) $(G^1,X) \cong_{k,r} (G^2,Y)$.

(2)
- *Let $r' \le r$. Let $v \in V_1$ be a vertex such that $d(X,v) > 2r'+1$. Let $\overline{u} \in (V_2)^{k-1}$ be a tuple of vertices. Then there exists $u \in V_2$ such that $d(u,\overline{u}) > 2r'+1$, $d(Y,u) > 2r'+1$ and $(G^1,v) \simeq_{k,r'} (G^2,u)$.*
- *Let $r' \le r$. Let $u \in V_2$ be a vertex such that $d(Y,u) > 2r'+1$. Let $\overline{v} \in (V_1)^{k-1}$ be a tuple of vertices. Then there exists $v \in V_1$ such that $d(v,\overline{v}) > 2r'+1$, $d(X,v) > 2r'+1$ and $(G^1,v) \simeq_{k,r'} (G^2,u)$*

Then Duplicator wins $\text{EHR}_k\big(G^1;G^2\big)$.

In order to prove this theorem we need to make two observations and prove a previous lemma.

Observation 2.1. *Let H^1, H^2 be hypergraphs and $\overline{v} \in V(H^1)^*$, $\overline{u} \in V(H^2)^*$, be lists of vertices. Suppose that Duplicator wins $d\mathrm{EHR}_k(H^1,\overline{v}; H^2,\overline{u})$. Then, for any $r \in \mathbb{N}$, $(H^1,\overline{v}) \simeq_{k,r} (H^2,\overline{u})$. A direct consequence of this fact is that given hypergraphs G^1, G^2 and sets $X \subseteq V(G^1)$, $Y \subseteq V(G^2)$ such that $(G^1,X) \simeq_{k,r} (G^2,Y)$ for some $r \in \mathbb{N}$, then for any $r' \leq r$ it also holds $(G^1,X) \simeq_{k,r'} (G^2,Y)$*

Given a set S and tuples $\overline{s}, \overline{t} \in S^*$ we write $\overline{s}^\frown\overline{t}$ to denote their concatenation.

Observation 2.2. *Let H_1, H_2 be hypergraphs and $\overline{v}$, $\overline{u}$, be lists of vertices from $V(H_1)$ and $V(H_2)$ respectively. Suppose Duplicator wins $d\mathrm{EHR}_k(H_1,\overline{v}; H_2,\overline{u})$. Let $v \in V(H_1), u \in V(H_2)$ be vertices played in the first round of an instance of the game where Duplicator is following a winning strategy. Then Duplicator also wins $d\mathrm{EHR}_{k-1}(H_1,\overline{v_2}; H_2,\overline{u_2})$, where $\overline{v_2} := \overline{v}^\frown v$ and $\overline{u_2} := \overline{u}^\frown u$.*

Lemma 2.1. *Let G^1, G^2 be hypergraphs and let $V_1 := V(G^1), V_2 := V(G^2)$. Let $\overline{v} \in V_1^*$ and $\overline{u} \in V_2^*$ be lists of vertices. Let $r \in \mathbb{N}$ be greater than zero. Suppose that $(G^1,\overline{v}) \simeq_{k,3r+1} (G^2,\overline{u})$. Let $v \in V_1$ and $u \in V_2$ be vertices played in the first round of an instance of*

$$d\mathrm{EHR}_k\big(N(\overline{v};3r+1),\overline{v};\quad N(\overline{u};3r+1),\overline{u}\big)$$

where Duplicator is following a winning strategy. Further suppose that $d(\overline{v},v) \leq 2r+1$ (and in consequence $d(\overline{u},u) \leq 2r+1$ as well). Let $\overline{v_2} := \overline{v}^\frown v$ and $\overline{u_2} := \overline{u} \frown u$. Then $(G^1,\overline{v_2}) \simeq_{k-1,r} (G^2,\overline{u_2})$.

Proof. Using observation 2.2 we get that Duplicator wins

$$d\mathrm{EHR}_{k-1}\big(N(\overline{v};3r+1),\overline{v_2};\quad N(\overline{u};3r+1),\overline{u_2}\big)$$

as well. Call $H_1 = N(\overline{v};3r+1)$, $H_2 = N(\overline{u};3r+1)$. Then by observation 2.2 Duplicator wins

$$d\mathrm{EHR}_{k-1}\big(N^{H_1}(\overline{v_2};r),\overline{v_2};\quad N^{H_2}(\overline{u_2};r),\overline{u_2}\big).$$

Because of this if we prove $N^{G^1}(\overline{v_2};r) = N^{H_1}(\overline{v_2};r)$ and $N^{G^2}(\overline{u_2};r) = N^{H_2}(\overline{u_2};r)$, then we are finished. Let $z \in N^{G^1}(v';r)$. Then $d(z,\overline{v}) \leq d(z,v') + d(v',\overline{v}) = 3r+1$. In consequence, $N^{G^1}(v;r) \subset H_1$. Thus, $N^{G^1}(\overline{v_2};r) \subseteq H_1$, and $N^{G^1}(\overline{v_2};r) = N^{H_1}(\overline{v_2};r)$. Analogously we obtain $N^{G^2}(\overline{u_2};r) = N^{H_2}(\overline{u_2};r)$, as we wanted. □

Now we are in conditions to prove theorem 2.1.

Proof of theorem 2.1. Let $X_1,\dots,X_a$ and $Y_1,\dots,Y_b$ be partitions of X and Y respectively as in the definition of k-agreeability. Define $r_0 = (3^k-1)/2$ and $r_i = (r_{i-1}-1)/3$ for each $1 \leq i \leq k$. Let us denote by v_i^1 and v_i^2 the vertices played in G^1 and G^2 respectively during the i-th round of $\mathrm{EHR}_k(G^1,G^2)$. We will show a winning strategy for Duplicator in $d\mathrm{EHR}_k\big(G^1;G^2\big)$. For each $0 \leq i \leq k$, Duplicator will keep track of some marked sets of vertices $T \subset V_1$, $S \subset V_2$. For $\delta = 1,2$ each marked set $T \subset V_\delta$ will have associated a tuple of vertices $\overline{v}(T) \in V_\delta^*$ consisting of the vertices played in G_2 so far, ordered according to the rounds they where played, that were "appropriately close" to T when chosen. The game will start with no sets of vertices marked and at the end of the i-th round Duplicator will perform one of the two following operations:

- Mark two sets $S \subset V_1$ and $T \subset V_2$ and define $\overline{v}(S) = v_i^1$ and $\overline{v}(T) = v_i^2$.
- Given two sets $S \subset V_1, T \subset V_2$ that were previously marked during the same round, append v_i^1 and v_i^2 to $\overline{v}(S)$ and $\overline{v}(T)$ respectively.

In particular this means that at the end of the i-th round the marked sets $S \subset V_1$, $T \subset V_2$ and their respective lists $\overline{v}(S)$, $\overline{v}(T)$ satisfy

(i) For $\delta = 1,2$, each vertex played so far $v_j^\delta \in V_\delta$ belongs to $\overline{v}(S)$ for a unique marked set $S \subset V_\delta$.

(ii) Let $S \subset V_1$ and $T \subset V_2$ be sets marked during the same round. Then any previously played vertex v_j^1 occupies a position in $\overline{v}(S)$ if and only if v_j^2 occupies the same position in $\overline{v}(T)$.

The following conditions will also be satisfied at the end of the i-th round

(iii)
- Let $S \subset V_1$ be a marked set. Then for any different marked $S' \subset V_1$ of any different S' among $X_1, \dots, X_a$ it holds $d(S, S') > 2r_i + 1$.
- Let $T \subset V_2$ be a marked set. Then for any different marked $T' \subset V_2$ or any different T' among $Y_1, \dots, Y_b$ it holds $d(T, T') > 2r_i + 1$.

(iv) Let $S \subset V_1, T \subset V_2$ be sets marked during the same round. Then

$$\big(G^1, (S, \overline{v}(S))\big) \simeq_{k-i, r_i} \big(G^2, (T, \overline{v}(T))\big).$$

In particular, if conditions (i) to (iv) are satisfied this means that if $\overline{v}^1 = (v_1^1, \dots, v_i^1)$ and $\overline{v}^2 = (v_1^2, \dots, v_i^2)$ are the vertices played so far then Duplicator wins

$$d\text{EHR}_{k-i}\big(N(\overline{v}^1; r_i), \overline{v}^1; \quad N(\overline{v}^2; r_i), \overline{v}^2\big),$$

And at the end of the k-th round Duplicator will have won $\text{EHR}(G^1; G^2)$.

The game $d\text{EHR}_k(G_1; G_2)$ proceeds as follows. Suppose that during the i-th round Spoiler chooses $v_i^1 \in V_1$ (the case where they play in V_2 is symmetric). There are three possible cases:

- For some unique previously marked set $S \subset V_1$ it holds that $d(S \cup \overline{v}, \quad v_i^1) \leq 2r_i + 1$. In this case let $T \subset V_2$ be the set in G_2 marked in the same round as T. By hypothesis

  $$\big(G^1, (S, \overline{v}(S))\big) \simeq_{k-i+1, 3r_i+1} \big(G^2, (T, \overline{v}(T))\big).$$

  Then, by definition, for some orderings $\overline{w}$, $\overline{z}$ of the vertices in S and T respectively it holds that Duplicator wins

  $$d\text{EHR}_{k-i+1}\big(N(\overline{w}^\frown\overline{v}(S); 3r_i + 1), \overline{w}^\frown\overline{v}(S); \quad N(\overline{z}^\frown\overline{v}(T); 3r_i + 1), \overline{z}^\frown\overline{v}(T)\big).$$

  Thus Duplicator can choose $v_i^2 \in V_2$ according to the winning strategy in that game. After this Duplicator sets $\overline{v}(S) := \overline{v}(S)^\frown v_i^1$, and $\overline{v}(T) := \overline{v}(T)^\frown v_i^2$. Notice that because of lemma 2.1 now

  $$\big(G^1, (S, \overline{v}(S))\big) \simeq_{k-i, r_i} \big(G^2, (T, \overline{v}(T))\big).$$

- For all marked sets $S \subset V_1$ it holds $d(S \cup \overline{v}(S), \quad v_i^1) > 2r_i + 1$, but there is a unique S among $X_1, \ldots, X_a$ such that $d(S, v_i^1) \leq 2r_i + 1$. In this case from condition (1) of the statement follows that there is some non-marked set T among $Y_1, \ldots, Y_b$ such that

  $$(G^1, S) \simeq_{k-i+1, 3r_i+1} (G^2, T).$$

  Thus, by definition, for some orderings $\overline{w}$, $\overline{z}$ of the vertices in S and T respectively it holds that Duplicator wins

  $$d\mathrm{EHR}_{k-i+1}\big(N(\overline{w}; 3r_i + 1), \overline{w}; \quad N(\overline{z}; 3r_i + 1), \overline{z}\big).$$

  Then Duplicator can choose $v_i^2 \in V_2$ according to a winning strategy for this game. After this Duplicator marks both S and T and sets $\overline{v}(S) := v_i^1$, and $\overline{v}(T) := v_i^2$. Notice that because of lemma 2.1 now

  $$\big(G^1, (S, \overline{v}(S))\big) \simeq_{k-i, r_i} \big(G^2, (T, \overline{v}(T))\big).$$

- For all marked sets $S \subset V_1$ it holds $d(S \cup \overline{v}(S), \quad v_i^1) > 2r_i + 1$, and for all sets S among $X_1, \ldots, X_a$ it also holds $d(S, v_i^1) > 2r_i + 1$. In this case from condition (2) of the statement follows that Duplicator can choose $v_i^2 \in V_2$ such that (A) $d(T \cup \overline{v}(T), \quad v_i^2) > 2r_i + 1$ for all marked sets $T \subset V_2$, (B) $d(T, v_i^2) > 2r_i + 1$ for all sets T among $Y_1, \ldots, Y_b$, and (C) $(G^1, v_i^1) \simeq_{k-i, r_i} (G^2, v_i^2)$. After this Duplicator marks both $S = \{v_i^1\}$ and $T = \{v_i^2\}$ and sets $\overline{v}(S) := v_i^1$, and $\overline{v}(T) := v_i^2$.

The fact that conditions (i) to (iv) still hold at the end of the round follows from comparing r_{i-1} and r_i as well as applying observation 2.1 and observation 2.2.

□

2.1 Types of trees

We define tree T as a connected hypergraph such that $ex(T) = -1$. We define a vertex-rooted tree (T, v) as a tree T with a distinguished vertex $v \in V(T)$ called its root. We will usually omit the root when it is not relevant and write just T instead of (T, v). We define the set of initial edges of a vertex-rooted tree (T, v) as the set of edges in T that contain v.

Given a rooted tree (T, v), and a vertex $u \in V(T)$, we define $\tau_{(T,v)}(u)$ as the tree $T[X]$ induced on the set $X := \{\, w \in V(T) \,|\, d(v, w) = d(v, u) + d(u, w) \,\}$, to which we assign u as the root. That is, $\tau_{(T,v)}(u)$ is the tree consisting of those vertices whose only path to v contains u.

We define the radius of a vertex-rooted, or edge-rooted, tree as the maximum distance between its marked vertex and any other one.

Fix a natural number k. We will define two equivalence relations, one between rooted trees and another between pairs (T, e) of rooted trees T and initial edges $e \in E(T)$. We will name both relations k-equivalence relations and denote them by $\simeq_k$. They are defined recursively as follows:

- Any two trees with radius zero are k-equivalent. Notice that those trees consist only of one vertex: their respective roots.

- Suppose that the k-equivalence relation has been defined for rooted trees with radius at most r. Let Σ_r be the set consisting of the root symbol τ and the k-equivalence classes of trees with radius lesser than r. Given a rooted tree (T,v) with whose radius is lesser than r we define its canonical Σ_r-coloring as the map $\chi_{(T,v)} : V(T) \to \Sigma_r$ satisfying that $\chi_{(T,v)}(u)$ is the k-equivalence class of $\mathrm{Tr}(u,T;v)$ for any $u \neq v$ and $\chi_{(T,v)}(v) = \tau$.

  Let T_1 and T_2 be rooted trees with radius at most $r+1$. We say that $(T_1,v_1) \simeq_k (T_2,v_2)$ if given any Σ_r-pattern (e,χ) the "quantity of initial edges $e_1\prime \in E(T_1)$ such that $(e,\chi) \simeq (e_1,\chi_{(T_1,v_1)})$" and the "quantity of initial edges $e_2\prime \in E(T_2)$ such that $(e,\chi) \simeq (e_2,\chi_{(T_2,v_2)})$" are the same or are both greater than $k-1$.

We want prove the following

Theorem 2.2. *Let (T_1,v_1) and (T_2,v_2) be rooted trees. Then, if they are k-equivalent Duplicator wins $d\text{EHR}_k(T_1,v_1,T_2,v_2)$.*

Before proceeding with the proof that we need an auxiliary result. Let (T,v) be a rooted tree and e an initial edge of T. We define $Tree_{(T_v)}(e)$ as the induced tree $T[X]$ on the set $X := \{v\} \cup \{\, u \in V(T) \,|\, d(v,u) = |e| + d(e,v)\,\}$, to which we assign v as the root. In other words, $Tree_{(T_v)}(e)$ is the tree formed of v and all the vertices in T whose only path to v contain e. Now we can check the following:

Lemma 2.2. *Fix $r > 0$. Suppose that theorem 2.2 holds for rooted trees with radii at most r. Let (T_1,v_1) and (T_2,v_2) be rooted trees with radii at most $r+1$. Let e_1 and e_2 be initial edges of T_1 and T_2 respectively satisfying $(T_1,e_1) \simeq_k (T_2,e_2)$. Name $T_1' = Tree_{(T_1,v_1)}(e_1)$ and $T_2' = Tree_{(T_2,v_2)}(e_2)$. Then Duplicator wins $d\text{EHR}k(T_1',v_1,T_2',v_2)$.*

Proof. We show a winning strategy for Duplicator. Suppose that in the i-th round of the game Spoiler plays on T_1'. The other case is symmetric. Let $f : e_1 \to e_2$ be a bijection as in the definition of $(T_1,e_1) \simeq_k (T_2,e_2)$. There are two possibilities:

- If Spoiler plays a vertex v on e_1 then Duplicator can play $f(v)$ on e_2.
- Otherwise, Spoiler plays a vertex v that belongs to some $Tree_{(T_1',v_1)}(u)$ for a unique $u \in e_1$ different from the root v_1. By the definition of $(T_1,e_1) \simeq_k (T_2,e_2)$, $Tree_{(T_1',v_1)}(u) \simeq_k Tree_{(T_2',v_2)}(f(u))$. As both these trees have radii at most r, by assumption Duplicator has a winning strategy between them and they can follow it.

$\square$

Now we can prove the main theorem of this section:

Proof of theorem 2.2.

Notice that, as $T_1 \simeq_k T_2$, both T_1 and T_2 have the same radius r. We prove the result by induction on r. If $r = 0$ then both T_1 and T_2 consist of only one vertex and we are done.

Now let $r > 0$ and assume that the statement is true for all lesser values of r. We will show that there is a winning strategy for Duplicator in $d\text{EHR}_k(T_1,v_1,T_2,v_2)$. At the start of the game, set all the initial edges in T_1 and T_2 as non-marked. Suppose that in the i-th round Spoiler plays in T_1. The other case is symmetric.

- If Spoiler plays v_1 then Duplicator plays v_2.
- Otherwise, the vertex played by Spoiler belongs to $Tree_{(T_1,v_1)}(e_1)$ for a unique initial edge e_1 of T_1. There are two possibilities:
  - If e_1 is not marked yet, mark it with the index i. In this case, there is a non-marked initial edge e_2 in T_2 satisfying $(T_1,e_1) \simeq_k (T_2,e_2)$. Mark e_2 with the index i as well. Because of lemma 2.2, Duplicator has a winning strategy in

    $$d\text{EHR}k(Tree_{T_1}(e_1), v_1, Tree_{T_1}(e_2), v_2)$$

    and can play according to it.
  - If e_1 is already marked then there is a unique initial edge e_2 in T_2 marked with the same mark as e_1 and $(T_1,e_1) \simeq_k (T_2,e_2)$. Again, Because of lemma 2.2, Duplicator has a winning strategy in

    $$d\text{EHR}k(Tree_{T_1}(e_1), v_1, Tree_{T_1}(e_2), v_2)$$

    and can continue playing according to it.

  Then Duplicator can find an initial edge e_2 of T_2 such that $(T_1,e_1) \simeq_k (T_2,e_2)$. Because of lemma 2.2, Duplicator has a winning strategy in $d\text{EHR}k(Tree_{T_1}(e_1), v_1, Tree_{T_1}(e_2), v_2)$ and can play according to it.

□

probablemente con algún dibujo sencillo esta demostración se entienda mejor

3 Probabilistic results

3.1 Convergence to Poisson variables

Given a natural numbers n and l we will use $(n)_l$ to denote $n(n-1)\cdots(n-l+1)$ or 1 if $l=0$.

Our main tool for computing probabilities will be the following multivariate version of Brun's Sieve (Theorem 1.23, [11]).

Theorem 3.1. *Fix $k \in \mathbb{N}$. For each $n \in \mathbb{N}$, let $X_{n,1},\dots,X_{n,k}$ be non-negative random integer variables over the same probability space. Let $\lambda_1,\dots,\lambda_l$ be real numbers. Suppose that for any $r_1,\dots,r_l \in \mathbb{N}$*

$$\lim_{n\to\infty} \mathrm{E}[\prod_{i=1}^{k} \binom{X_{n,i}}{r_i}] = \prod_{i=1}^{k} \frac{\lambda_i}{r_i!}.$$

Then the $X_{n,1},\dots,X_{n,k}$ converge in distribution to independent Poisson variables with means $\lambda_1,\dots,\lambda_k$ respectively.

3.2 Almost all hypergraphs are simple

We say that a connected hypergraph G is **dense** if $ex(G) > 0$. Given $r \in \mathbb{N}$, we say that G is r**-simple** if G does not contain any dense subgraph H such that $diam(H) \leq r$. The goal of this section is to show that, for any fixed r, a.a.s G_n is r-simple.

Lemma 3.1. *Let H be an hypergraph. Then* $\mathrm{E}\big[\#$ *copies of H in* $G_n\big] = \Theta(n^{-ex(H)})$ *as n tends to infinity.*

Proof. It holds

$$\mathrm{E}\big[\#\text{ copies of } H \text{ in } G_n\big] = \sum_{H' \in Copies(H,[n])} \Pr\big(H' \subset G_n\big).$$

We have that $|Copies(H,[n])| = \frac{(n)_{v(H)}}{|Aut(H)|}$. Also, for any $H' \in Copies(H,[n])$ it is satisfied

$$\Pr\big(H' \subset G_n\big) = \prod_{R \in \sigma} \left(\frac{\beta_R}{n^{ar(R)-1}} \right)^{e_R(H)}.$$

Substituting in the first equation we get

$$\mathrm{E}\big[\#\text{ copies of } H \text{ in } G_n\big] = \frac{(n)_{v(H)}}{|Aut(H)|} \cdot \prod_{R \in \sigma} \left(\frac{\beta_R}{n^{ar(R)-1}} \right)^{e_R(H)} \underset{n \to \infty}{\sim} n^{-ex(H)} \cdot \frac{\prod_{R \in \sigma} \beta_R^{e_R H}}{|Aut(H)|}.$$

□

As a corollary of last result we get the following:

Lemma 3.2. *Let H be an hypergraph such that $ex(H) > 0$. Then a.a.s there are no copies of H in G_n.*

Proof. Because of the previous fact, $\mathrm{E}\big[\#\text{ copies of } H \text{ in } G_n\big] \xrightarrow{n \to \infty} 0$. An application of the first moment method yields the desired result. □

A similar result that will be useful later is the following:

3.2.1 Rooted hypergraphs

A rooted hypergraph $(G,\overline{u})$ is an hypergraph H together with an ordered sequence of distinguished vertices $\overline{u} \in (V(H))_*$. An isomorphism between two rooted hypergraphs $(G,\overline{u})$ and $(H,\overline{v})$ is a map $f : V(G) \to V(H)$ such that f is an isomorphism between G and H that satisfies the additional condition $f(\overline{u}) = \overline{v}$. An automorphism of $(G,\overline{u})$ is an isomorphism from $(G,\overline{u})$ to itself. We write $Aut(G,\overline{u})$ to denote the group of automorphisms of $(G,\overline{u})$.

Given a rooted hypergraph $(G,\overline{u})$, a set of vertices V and a list $\overline{v} \in (V)_*$ such that $len(\overline{u}) = len(\overline{v})$ we define the set $Copies((G,\overline{u}),(V,\overline{v}))$ as the set of rooted hypergraphs $(H,\overline{v})$ isomorphic to $(G,\overline{u})$ such that $V(H) \subset V$.

Lemma 3.3. *Let $(H,\overline{u})$ be a rooted hypergraph. Let $\overline{v} \in (\mathbb{N})_*$ be a list of vertices satisfying $len(\overline{u}) = len(\overline{v})$. For each $n \in \mathbb{N}$ let X_n be the random variable that counts the copies $(H',\overline{v}) \in Copies((H,\overline{u}),([n],\overline{v}))$ that are contained in G_n. Then* $\mathrm{E}\big[X_n\big] = \Theta(n^{-ex(H)-len(\overline{u})})$.

Proof. It holds

$$\mathrm{E}\big[X_n\big] = \sum_{H' \in (H',\overline{v}) \in Copies((H,\overline{u})} \Pr\big(H' \subset G_n\big) = \frac{(n)_{(v(H)-len(\overline{u}))}}{|Aut(H,\overline{u})|} \cdot \prod_{R\in\tau} \left(\frac{\beta_R}{n^{ar(R)-1}}\right)^{e_R(H)}$$

□

The main theorem of this section is the following

Theorem 3.2. *Let $r \in \mathbb{N}$. Then a.a.s G_n is r-simple.*

The first moment method alone is not sufficient to prove our claim because the amount of dense hypergraphs H such that $diam(H) \leq r$ is not finite in general. Thus, we need to prove that it suffices to prohibit a finite amount of dense sub-hypergraphs in order to guarantee that G_n is r-simple.

Lemma 3.4. *Let H be a dense hypergraph of radius r. Then H contains a dense sub-hypergraph H' with size no greater than $(a+2)(r+1)+2a$, where a is the largest edge size in H.*

Proof. Choose $x \in V(H)$. Successively remove from G edges e such that $d(x,e)$ is maximum until the resulting graph H' has excess no greater than 0. We have two cases:

- $ex(H') = -1$. Let $e = [x_1,\dots,x_b]$ be the last removed edge and $e \cap H' = \{x_{i_1},\dots,x_{i_d}\}$. For any $j = 1,\dots,d$ choose P_j a path of size no greater than $r+1$ joining x and x_{i_j} in H'. Then $P_1 \cup \dots P_d \cup e$ is a dense sub-hypergraph of H of size less than $a(r+1)+a < (a+2)(r+1)+2a$.

- $ex(H') = 0$. Let $e_1 = [x_1,\dots,x_{b_1}]$ be the last removed edge. Continue removing the edges of G' that are at maximum distance from x until you obtain H'' with $ex(H'') = -1$. Let $e_2 = [y_1,\dots,y_{b_2}]$ be the last removed edge this time. As before, let $e_1 \cap H' = \{x_{i_1},\dots,x_{i_d}\}$ and for $j = 1,\dots,d$ let P_j a path of size no greater than $r+1$ joining x and x_{i_j} in H' Then $e_2 \cup H'' = \{y_{i_1}, y_{i_2}\}$. Let Q_1, Q_2 be paths size no greater than $r+1$ from x to y_{i_1} and y_{i_2} in H''. Then $Q_1 \cup Q_2 \cup e_2$ is a graph of likelihood 0 and size less than $2r+2+a$, and $Q_1 \cup Q_2 \cup P_1 \cup \dots \cup P_d \cup e_1 \cup e_2$ is a critical graph with size less than $(2+a)(r+1)+2a$

□

Now we are in conditions to prove theorem 3.2.

Proof. Because of last lemma there is a constant R such that "G does not contain dense hypergraphs of size bounded by R" implies that "G is r-simple". Thus,

$$\lim_{n\to\infty} \Pr\big(G_n \text{ is } r\text{-simple}\big) \geq \lim_{n\to\infty} \Pr\big(G_n \text{ does not contain dense hypergraphs of size bounded by } R\big).$$

Because of lemma 3.2, given any individual dense hypergraph, the probability that there are no copies of it in G_n tends to 1 as n goes to infinity. Using that there are a finite number of dense hypergraphs of size bounded by R we deduce that the RHS of last inequality tends to 1. □

3.3 Probabilities of trees

During this section we want to study the asymptotic probability that the r-neighborhood of a given vertex $v \in \mathbb{N}$ in G_n is a tree that belongs to a given k-equivalence class of trees $\mathscr{T}$ with radius at most r. That is, we want to know

$$\lim_{n\to\infty} \Pr\big(T := N^{G_n}(v;r) \text{ is a tree, and } (T,v) \in \mathscr{T}\big).$$

Denote this limit by $\Pr[r,\mathscr{T}]$. Notice that the definition of $\Pr[r,\mathscr{T}]$ does not depend by the choice of v.

We define Λ and M as the minimal families of expressions with arguments $\overline{\beta}$ that satisfy the conditions: **(1)** $1 \in \Lambda$, **(2)** for any $b,i \in \mathbb{N}$ with $1 \leq i \leq c$, $b > 0$, and $\lambda_1,\dots,\lambda_{a_i-1} \in \Lambda$, the expression $(\beta_i/b)\prod_{j=1}^{a_i-1}\lambda_j$ belongs to M, **(3)** for any $\mu \in M$ and any $n \in \mathbb{N}$ both $\text{Poiss}_\mu(n)$ and $\text{Poiss}_\mu(\geq n)$ are in Λ, and **(4)** for any $\lambda_1,\lambda_2 \in \Lambda$, the product $\lambda_1\lambda_2$ belongs to Λ as well.

The goal of this section is to show that $\Pr[r,\mathscr{T}]$, as an expression with parameters $\overline{\beta}$, belongs to Λ for any choice of r and $\mathscr{T}$.

Lemma 3.5. *Let $\overline{v} \subset \mathbb{N}*$ be a finite set of fixed vertices and let $\sigma(\overline{x})$ be an open formula with no equality such that $length(\overline{x}) = length(\overline{v})$. Define $G'_n = G_n \setminus E(\overline{v})$. Fix $R \in \mathbb{N}$.*

- *Let A_n be the event that G'_n contains a path of size at most $R+1$ between any two vertices $u,w \in \overline{v}$.*
- *Let B_n be the event that G'_n contains a cycle of size at most $R+1$ that contains a vertex $u \in \overline{v}$.*

Then $\lim_{n\to\infty} \Pr(A_n \,|\, \sigma(\overline{v})) = 0$, and $\lim_{n\to\infty} \Pr(B_n \,|\, \sigma(\overline{v}) = 0$.

Proof. Notice that the events A_n and B_n do not concern the possible edges induced over $\overline{v}$. In consequence, because edges are independent in our random model, $\Pr(A_n \,|\, \sigma(\overline{v})) = \Pr(A_n)$ and $\Pr(B_n \,|\, \sigma(\overline{v})) = \Pr(B_n)$.

The facts that $\lim_{n\to\infty} \Pr(A_n) = 0$ and $\lim_{n\to\infty} \Pr(B_n) = 0$ follow from lemma 3.3 using that (1) the excess of any path is greater or equal than -1, (2) the amount of paths of size at most $R+1$ is finite, (3) the excess of any cycle is zero, and (4) the amount of cycles of size at most $R+1$ is finite. $\square$

Definition 3.1. We call an hypergraph G **saturated** if any proper sub-hypergraph $G' \subset H$ satisfies $ex(G') < ex(G)$.

The **center** of a connected hypergraph G is its maximal saturated sub-hypergraph and it is denoted by $Center(G)$. In the general case the center of an hypergraph is the union of the centers of its connected components.

Definition 3.2. Let G be a connected hypergraph and let $\overline{v} \in V(G)^*$. Then we call $Center(G,\overline{v})$ to the minimal connected hypergraph that contains $Center(G)$ and the vertices $\overline{v}$. In general, if G is an arbitrary hypergraph with connected components $G^1,\dots,G_k$, and $\overline{v}$ are vertices $V(G)$, then we call $Center(G,\overline{v})$ to the union of $Center(G_i, V(G_i) \cap \overline{v})$ for all the connected components G_i.

Definition 3.3. Let G be an hypergraph G, let $\overline{u} \in V(G)^*$ and let $v \in \overline{u}$. Consider the graph $G' = G \setminus E\big(Center(G,\overline{u})\big)$. Then the connected components of G' are all trees. We call the **tree of** v **in** $G(\overline{u})$, denoted by $Tr\big(G(\overline{u}),v\big)$, to the connected component of G' to which v belongs with v as its root.

In this same situation, let $r \in \mathbb{N}$ and $H := N^G(\overline{u};r)$. We call the r**-tree of** v **in** $G(\overline{u})$, denoted by $Tr\big(G(\overline{u}),v;r\big)$ to $Tr\big(H(\overline{u}),v\big)$.

Theorem 3.3. *Fix $r \in \mathbb{N}$. The following are satisfied:*

(1) Let $\mathscr{T}$ be a k-equivalence class for trees with radii at most r. Then $\Pr[r, \mathscr{T}]$ exists and is an expression in Λ.

(2) Let $\overline{u} \in (\mathbb{N})_*$ be a list of different fixed vertices, and let $\phi[\overline{x}] \in FO[\sigma]$ be a consistent edge sentence such that $len(\overline{x}) = len(\overline{u})$. Let $\overline{v} \in (\mathbb{N})_*$ be vertices contained in $\overline{u}$. For each $v \in \overline{v}$ let $\mathscr{T}_v$ be a k-equivalence class of trees with radii at most r. Then

$$\lim_{n\to\infty} \Pr\Big(\bigwedge_{v\in\overline{v}} Tr\big(G_n, \overline{u}, v; r\big) \in \mathscr{T}_v \,|\, \sigma(\overline{w})\Big) = \prod_{v\in\overline{v}} \Pr[r, \mathscr{T}_v].$$

Proof. We will prove (1) and (2) together by induction on r.

Assume $r = 0$. We start by showing that (1) holds. Recall that all trees with radius zero are k-equivalent. Thus, if $\mathscr{T}$ is the unique k-equivalence class of trees with radius zero and $v \in \mathbb{N}$ is a fixed vertex then

$$\Pr[0; \mathscr{T}] = \lim_{n\to\infty} \Pr\big(T := N^{G_n}(v;0) \text{ is a tree, and } (T, v) \in \mathscr{T}\big) = 1,$$

Indeed, $N^{G_n}(v;0)$ consists of a single vertex for all $n \geq v$, and the above equation follows. The expression 1 belongs to Λ, so (1) holds.

The case of (2) is analogous. As $r = 0$, then $\mathscr{T}_1 = \cdots = \mathscr{T}_k$ are the unique k-equivalence class of trees with radius zero. Then, given $\sigma, \overline{u}, v_1, \ldots, v_k$ as in the statement,

$$\lim_{n\to\infty} \Pr\Big(\bigwedge_{i=1}^{l} Tr\big(G_n, \overline{u}, v_i; 0\big) \in \mathscr{T}_i \,|\, \sigma(\overline{w})\Big) = \prod_{i=1}^{l} \Pr[0, \mathscr{T}_i] = 1.$$

Because of (1), $\Pr[0, \mathscr{T}_i] = 1$ for all i's, and (2) holds.

Now let $r > 0$ and assume that both (1) and (2) hold for all

Via similar computations we can show that for any fixed $v \in \overline{v}$.

We are going to show that the variables $X_{n,i,\varepsilon}$ converge, as n tends to infinity, to independent Poisson variables $\text{Poiss}(\mu_{r_i,\varepsilon})$ whose means $\mu_{r_i,\varepsilon}$ are expressions in the family M. Let $[\Sigma_{r_i}]$ be the set of Σ_{r_i}-patterns. We want to prove:

$$\lim_{n\to\infty} \Pr\Big(\bigwedge_{i=1}^{k} \bigwedge_{\varepsilon\in[\Sigma_{r_i}]} X_{n,i,\varepsilon} = a_{i,\varepsilon} \,|\, \sigma(\overline{w})\Big) = \prod_{i=1}^{k} \prod_{\varepsilon\in[\Sigma_{r_i}]} e^{-\mu_{r_i,\varepsilon}} \frac{(\mu_{r_i,\varepsilon})^{a_{i,\varepsilon}}}{a_{i,\varepsilon}!} \tag{1}$$

Furthermore, for each i and ε we will prove that the mean $\mu_{r_i,\varepsilon}$ does only depend on r_i and ε. This proves both (1), and (2).

Given an Σ_{r_i}-pattern ε and any $(e, \chi) \in Copies(\varepsilon, [n])$ we say that $(e, \chi) \in T_{n,i}$ if the following are satisfied: (1) e is an initial edge of $T_{n,i}$, and (2) that for any $v \in e$ such that $v \neq v_i$, it holds that $\chi(v)$ is the $\simeq_k$ class of $Tr(T_{n,i}, v)$.

Given $r \in \mathbb{N}$ and $\varepsilon \in \Sigma_r$ we define $\mu_{r,\varepsilon}$ as follows. Let (e, χ) be any representative of ε. Then

$$\mu_{r,\varepsilon} = \frac{\beta_{R(e)}}{|Aut(e, \chi)|} \prod_{\substack{v\in e \\ \chi(v)\neq\tau}} \Pr[r, \chi(v)].$$

For each $i \in [l]$ and $\varepsilon \in [\Sigma_{r_i}]$ let $b_{i,\varepsilon} \in \mathbb{N}$ be fixed.

We want to prove

$$\lim_{n\to\infty} \mathrm{E}\left[\prod_{\substack{i\in[l]\\ \varepsilon\in[\Sigma_{r_i}]}} \binom{X_{n,i,\varepsilon}}{b_{i,\varepsilon}}\right] = \prod_{\substack{i\in[l]\\ \varepsilon\in[\Sigma_{r_i}]}} \frac{(\mu_{r_i,\varepsilon})^{b_{i,\varepsilon}}}{b_{i,\varepsilon}!}.$$

Because theorem 3.1 this is sufficient to prove eq. (1).

For each $n \in \mathbb{N}$ define

$$\Omega_n = \Bigg\{ (E_{v,\varepsilon})_{\substack{v\in\overline{v}\\ \varepsilon\in[\Sigma_{r(v)}]}} \;\Bigg|\; \forall v \in \overline{v}, \forall \varepsilon \in [\Sigma_{r(v)}] \quad E_{v,\varepsilon} \subset Copies(\varepsilon, [n], \overline{w}; v) \quad \wedge \quad |E_{v,\varepsilon}| = b_{v,\varepsilon} \Bigg\}.$$

Informally, the elements $(E_{v,\varepsilon})_{v,\varepsilon}$ of Ω_n are represent all choices of possible initial edges for the $T_{n,v}$'s: For each $v \in \overline{v}$ and each $\varepsilon \in [\Sigma_{r(v)}]$, $E_{n,v,\varepsilon}$ selects $b_{v,\varepsilon}$ possible initial edges of $T_{n,v}$ with pattern ε.

Using observation REF we obtain that

$$\mathrm{E}\Big[\prod_{\substack{i\in[l]\\ \varepsilon\in[\Sigma_{r_i}]}} \binom{X_{n,i,\varepsilon}}{b_{i,\varepsilon}}\Big] = \sum_{(E_{v,\varepsilon})_{v,\varepsilon}\in\Omega_n} \Pr\left(\bigwedge_{\substack{v\in\overline{v}\\ \varepsilon\in[\Sigma_{r(v)}]\\ (e,\chi)\in E_{v,\varepsilon}}} \Bigg(e \in E(T_{n,i}) \bigwedge_{u\in V(e), u\neq v} Tr(T_{n,v}; u) \in \chi(u) \Bigg) \right)$$

Let $(E_{v,\varepsilon})_{v,\varepsilon} \in \Omega_n$. In order for $e \in E(T_{n,v})$ to be possible for all $v \in \overline{v}$, $\varepsilon \in [\Sigma_{r(v)}]$, $e \in E_{v,\varepsilon}$ it is needed that each vertex in $[n] \setminus \overline{w}$ belongs at most to one edge $(e,\chi) \in \cup_{v,\varepsilon} E_{v,\varepsilon}$. This is because if for some

$$\bigwedge_{\substack{i\in[l]\\ \varepsilon\in[\Sigma_{r_i}]\\ (e,\chi)\in O_{i,\varepsilon}}} \Bigg(e \in E(T_{n,i}) \bigwedge_{v\in e, v\neq v_i} Tr(T_{n,i}, v) \in \chi(v) \Bigg)$$

to be possible it is needed that each $v \in$ □

Lemma 3.6. *Let $r \in \mathbb{N}$, $r > 0$. Let $\mathscr{T}$ be a $\simeq_k$ class of trees with radii at most r. Suppose that theorem 3.3 holds for $r-1$. Then* $\Pr\left[r, \mathscr{T}\right]$ *exists and is and expression in* Λ.

Proof. Fix a vertex $v \in \mathbb{N}$. For each n let $T_n := Tr(G_n, v; r)$. We are going to show that $\lim_{n\to\infty} \Pr\big(T_n \in \mathscr{T}\big)$ exists and it is an expression in Λ.

For any (k,r)-pattern ε let $X_{n,\varepsilon}$ be the random variable that counts the initial edges in T_n whose k-pattern is ε. In other words, $X_{n,\varepsilon}$ counts the colored edges $(e,\tau) \in Copies(\varepsilon, [n], v)$ such that e is an initial edge in T_n satisfying that for any $u \in V(e)$ with $u \neq v$, it holds $Tr(T_n(v), u) \in \tau(u)$. Thus,

$$\mathrm{E}\big[X_{n,\varepsilon}\big] = \sum_{(e,\tau)\in Copies(\varepsilon,[n];v)} \Pr\left(e \in E(T_n) \bigwedge_{\substack{u\in V(e)\\ u\neq v}} Tr(T_n, v; u) \in \tau(u) \right).$$

Because of the symmetry of our random model the probability in the RHS of last equation is the same for all $(e,\tau) \in Copies(\varepsilon, [n], ; v)$. Let $(e,\tau) \in Copies(\varepsilon, \mathbb{N}; v)$ be fixed. Using that $|Copies(\varepsilon, [n], v)| = \frac{(n)_{|e|-1}}{|Aut(\varepsilon)|}$ we obtain

$$\mathrm{E}\big[X_{n,\varepsilon}\big] = \frac{(n)_{|e|-1}}{|Aut(\varepsilon)|} \Pr\left(e \in E(T_n) \bigwedge_{\substack{u\in V(e)\\ u\neq v}} Tr(T_n, v; u) \in \tau(u) \right).$$

Also, it is satisfied

$$\Pr\left(e \in E(T_n) \bigwedge_{\substack{u\in V(e)\\ u\neq v}} Tr(T_n, v; u) \in \tau(u) \right) =$$

$$\Pr\big(e \in E(G_n)\big) \cdot \Pr\left(e \in E(T_n) \bigwedge_{\substack{u\in V(e)\\ u\neq v}} Tr(T_n, v; u) \in \tau(u) \,\middle|\, e \in E(G_n) \right)$$

Using REF and $\Pr\big(e \in E(G_n)\big) = \frac{\beta_{R(e)}}{n^{|e|-1}}$, the RHS of last equation is asymptotically equivalent to

$$\frac{\beta_{R(e)}}{n^{|e|-1}} \cdot \Pr\left(\bigwedge_{\substack{u\in V(e)\\ u\neq v}} Tr(T_n, v; u) \in \tau(u) \,\middle|\, e \in E(G_n) \right)$$

Fix $\overline{u} \in (\mathbb{N})_*$ a list that contains exactly the vertices in e. Then it holds that $Tr(T_n, v, u) = Tr(G_n(\overline{u}), u; r-1)$. Thus,

$$\Pr\left(\bigwedge_{\substack{u\in V(e)\\ u\neq v}} Tr(T_n, v; u) \in \tau(u) \,\middle|\, e \in E(G_n) \right) = \Pr\left(\bigwedge_{\substack{u\in V(e)\\ u\neq v}} Tr(G_n(\overline{u}), u; r-1) \in \tau(u) \,\middle|\, e \in E(G_n) \right)$$

The event $e \in E(G_n)$ can be written as an edge sentence whose variables are interpreted as vertices in $\bar{u}$. Thus, by hypothesis, the RHS of last equality is asymptotically equivalent to $\prod_{\substack{u\in V(e)\\ u\neq v}} \Pr\left[r-1,\tau(u)\right]$. Finally, joining everything we obtain

$$\lim_{n\to\infty} \mathrm{E}\big[X_{n,\varepsilon}\big] = \lim_{n\to\infty} \frac{(n)_{|e|-1}}{|Aut(\varepsilon)|}\cdot\frac{\beta_{R(e)}}{n^{|e|-1}} \prod_{\substack{u\in V(e)\\ u\neq v}} \Pr\left[r-1,\tau(u)\right] = \frac{\beta_{R(e)}}{|Aut(\varepsilon)|} \prod_{\substack{u\in V(e)\\ u\neq v}} \Pr\left[r-1,\tau(u)\right].$$

For each $\varepsilon \in [(k,r)]$ we define $\mu_{r,\varepsilon}$ as follows: let (e,τ) be a representative of ε whose root is v. Then

$$\mu_{r,\varepsilon} = \frac{\beta_{R(e)}}{|Aut(\varepsilon)|} \prod_{\substack{u\in V(e)\\ u\neq v}} \Pr\left[r-1,\tau(u)\right].$$

Notice that $\mu_{r,\varepsilon}$ depends only on r and ε and it is an expression belonging to M.

We are going to prove that the variables $X_{n,\varepsilon}$ converge in distribution to independent Poisson variables with mean values $\mu_{r,\varepsilon}$ respectively. For each $\varepsilon \in [(k,r)]$ let $b_\varepsilon \in \mathbb{N}$. We want to show that

$$\lim_{n\to\infty} \mathrm{E}\left[\prod_{\varepsilon\in[(k,r)]} \binom{X_{n,\varepsilon}}{b_\varepsilon}\right] = \prod_{\varepsilon\in[(k,r)]} \frac{(\mu_{r,\varepsilon})^{b_\varepsilon}}{b_\varepsilon!}. \tag{2}$$

For each $n \in \mathbb{N}$ define

$$\Omega_n := \left\{ (E_\varepsilon)_{\varepsilon\in[(k,r)]} \;\middle|\; \forall\varepsilon \in [(k,r)] \quad E_\varepsilon \subset Copies(\varepsilon,[n],v), \quad |E_\varepsilon| = b_\varepsilon \right\}$$

Informally, elements of Ω_n represent choices of b_ε possible initial edges of T_n whose k- pattern is ε for all (k,r)-patterns ε. Using observation REF we obtain

$$\mathrm{E}\left[\prod_{\varepsilon\in[(k,r)]} \binom{X_{n,\varepsilon}}{b_\varepsilon}\right] = \sum_{(E_\varepsilon)_{\varepsilon\in[(k,r)]}\in\Omega_n} \Pr\left(\bigwedge_{\substack{\varepsilon\in[(k,r)]\\ (e,\tau)\in E_\varepsilon}} \left(e\in E(T_n) \bigwedge_{\substack{u\in V(e)\\ u\neq v}} Tr(T_n,v,u)\in\tau(u) \right)\right).$$

We say that an element $(E_\varepsilon)_{\varepsilon\in[(k,r)]}$ of Ω_n is **disjoint** each vertex $w \in [n]\setminus\{v\}$ belongs to at most one edge $(e,\tau) \in \bigcup_{\varepsilon\in[(k,r)]} E_\varepsilon$ Notice that if we want the probability in the last sum to be greater than 0 for a particular $(E_\varepsilon)_{\varepsilon\in[(k,r)]} \in \Omega_n$ then necessarily $(E_\varepsilon)_{\varepsilon\in[(k,r)]}$ is disjoint. Indeed, suppose that a vertex $w \in [n]\setminus\{v\}$ belongs to two different edges $(e_1,\tau_1),(e_2,\tau_2) \in \bigcup_{\varepsilon\in[(k,r)]} E_\varepsilon$. In consequence e_1 and e_2 form a cycle, as they both contain v and w. This implies that $e_1,e_2 \notin E(T_n)$.

For each $n \in \mathbb{N}$ let $\Omega'_n \subset \Omega_n$ be the set of disjoint elements in Ω_n. Then

$$\mathrm{E}\left[\prod_{\varepsilon\in[(k,r)]} \binom{X_{n,\varepsilon}}{b_\varepsilon}\right] = \sum_{(E_\varepsilon)_{\varepsilon\in[(k,r)]}\in\Omega'_n} \Pr\left(\bigwedge_{\substack{\varepsilon\in[(k,r)]\\ (e,\tau)\in E_\varepsilon}} \left(e\in E(T_n) \bigwedge_{\substack{u\in V(e)\\ u\neq v}} Tr(T_n,v,u)\in\tau(u) \right)\right).$$

Also, because of the symmetry of the random model, for all disjoint elements $(E_\varepsilon)_{\varepsilon\in[(k,r)]}$ the probability in last sum is the same. In consequence, if we fix $(E_\varepsilon)_{\varepsilon\in[(k,r)]} \in \Omega'_{\mathbb{N}}$ we obtain

$$\mathrm{E}\left[\prod_{\varepsilon\in[(k,r)]} \binom{X_{n,\varepsilon}}{b_\varepsilon}\right] = |\Omega'_n|\cdot \Pr\left(\bigwedge_{\substack{\varepsilon\in[(k,r)]\\ (e,\tau)\in E_\varepsilon}} \left(e\in E(T_n) \bigwedge_{\substack{u\in V(e)\\ u\neq v}} Tr(T_n,v,u)\in\tau(u) \right)\right).$$

Let $\overline{w} \in (\mathbb{N})_*$ be a list containing exactly the vertices $u \in V(e)$ for all $e \in \bigcup_{\varepsilon\in[(k,r)]} E_\varepsilon$. Then, for any $e \in \bigcup_{\varepsilon\in[(k,r)]} E_\varepsilon$ and any any $V(e)$ with $u \neq v$ it holds that if $e \in E(T_n)$ then $Tr(T_n, v, u) = Tr(G_n, \overline{w}, u; r-1)$. Then

$$\mathrm{E}\left[\prod_{\varepsilon\in[(k,r)]} \binom{X_{n,\varepsilon}}{b_\varepsilon}\right] = |\Omega_n'| \cdot \Pr\left(\bigwedge_{\substack{\varepsilon\in[(k,r)]\\(e,\tau)\in E_\varepsilon}} \left(e \in E(T_n) \bigwedge_{\substack{u\in V(e)\\u\neq v}} Tr(G_n, \overline{w}, u; r-1) \in \tau(u)\right)\right).$$

Counting vertices and automorphisms we get that

$$|\Omega_n'| = (n)_{\sum_{\varepsilon\in[(k,r)]}(|\varepsilon|-1)\cdot b_\varepsilon} \prod_{\varepsilon\in[(k,r)]} \frac{1}{b_\varepsilon!} \cdot \left(\frac{1}{|Aut(\varepsilon)|}\right)^{b_\varepsilon}.$$

Also, using REF

$$\Pr\left(\bigwedge_{\substack{\varepsilon\in[(k,r)]\\(e,\tau)\in E_\varepsilon}} \left(e \in E(T_n) \bigwedge_{\substack{u\in V(e)\\u\neq v}} Tr(G_n, \overline{w}, u; r-1) \in \tau(u)\right)\right) \sim$$

$$\Pr\left(\bigwedge_{\substack{\varepsilon\in[(k,r)]\\(e,\tau)\in E_\varepsilon}} e \in E(G_n)\right) \cdot \Pr\left(\bigwedge_{\substack{\varepsilon\in[(k,r)]\\(e,\tau)\in E_\varepsilon\\u\in V(e)\\u\neq v}} Tr(G_n, \overline{w}, u; r-1) \in \tau(u) \,\middle|\, \bigwedge_{\substack{\varepsilon\in[(k,r)]\\(e,\tau)\in E_\varepsilon}} e \in E(G_n)\right).$$

The event $\bigwedge_{\substack{\varepsilon\in[(k,r)]\\(e,\tau)\in E_\varepsilon}} e \in E(G_n)$ clearly can be described via an edge sentence whose variables are interpreted as vertices in $\overline{w}$. Thus, by hypothesis last product of probabilities is asymptotically equivalent to

$$\prod_{\varepsilon\in[(k,r)]} \left(\frac{\beta_{R(\varepsilon)}}{n^{|\varepsilon|-1}}\right)^{b_\varepsilon} \cdot \prod_{\varepsilon\in[(k,r)]} (\lambda_{r,\varepsilon})^{b_\varepsilon}.$$

Joining everything we obtain

$$\begin{aligned}\lim_{n\to\infty} \mathrm{E}\left[\prod_{\varepsilon\in[(k,r)]} \binom{X_{n,\varepsilon}}{b_\varepsilon}\right] &= \lim_{n\to\infty} \frac{(n)_{\sum_{\varepsilon\in[(k,r)]}(|\varepsilon|-1)\cdot b_\varepsilon}}{n^{\sum_{\varepsilon\in[(k,r)]}(|\varepsilon|-1)\cdot b_\varepsilon}} \cdot \prod_{\varepsilon\in[(k,r)]} \frac{1}{b_\varepsilon!} \cdot \left(\frac{\beta_{R(\varepsilon)}}{|Aut(\varepsilon)|}\right)^{b_\varepsilon} \cdot (\lambda_{r,\varepsilon})^{b_\varepsilon} \\ &= \prod_{\varepsilon\in[(k,r)]} \frac{(\mu_{r,\varepsilon})^{b_\varepsilon}}{b_\varepsilon!},\end{aligned}$$

as we wanted. In consequence, by theorem 3.1, given $a_\varepsilon \in \mathbb{N}$ for all $\varepsilon \in [(k,r)]$ it holds

$$\lim_{n\to\infty} \Pr\left(\bigwedge_{\varepsilon\in[(k,r)]} X_{n,\varepsilon} = a_\varepsilon\right) = \prod_{\varepsilon\in[(k,r)]} e^{-\mu_{r,\varepsilon}} \frac{(\mu_{r,\varepsilon})^{a_\varepsilon}}{a_\varepsilon!}.$$

Notice that, because of the definition of $\simeq_k$, the event $T_n \in \mathscr{T}$ is equivalent to

$$\left(\bigwedge_{\varepsilon\in E^1_{\mathscr{T}}} X_{n,\varepsilon} \geq k\right) \wedge \left(\bigwedge_{\varepsilon\in E^2_{\mathscr{T}}} X_{n,\varepsilon} = a_\varepsilon\right),$$

for some partition $E^1_{\mathscr{T}}, E^2_{\mathscr{T}}$ of $[(k,r)]$ that only depends on $\mathscr{T}$ and some natural numbers $a_\varepsilon < k$ for each $\varepsilon \in E^2_{\mathscr{T}}$ that only depend on $\mathscr{T}$ as well. In consequence

$$\lim_{n\to\infty} \Pr\big(T_n \in \mathscr{T}\big) = \left(\prod_{\varepsilon\in E^1_{\mathscr{T}}} \left(1 - \sum_{i=0}^{k-1} e^{-\mu_{r,\varepsilon}} \frac{(\mu_{r,\varepsilon})^i}{i!}\right)\right) \cdot \left(\prod_{\varepsilon\in E^2_{\mathscr{T}}} e^{-\mu_{r,\varepsilon}} \frac{(\mu_{r,\varepsilon})^{a_\varepsilon}}{a_\varepsilon!}\right)$$

And last expression belongs to Λ as we wanted to prove. $\square$

4 Probabilities of cycles

Theorem 4.1. *Let $\mathscr{O}$ be a simple k-agreeability class of hypergraphs. Then $\lim_{n\to\infty} \Pr\big(G_n \in \mathscr{O}\big)$ exists and is an expression in Θ.*

Proof. Define $r := 3^k$. For each $O \in C(k,r)$ let $X_{n,O}$ be the random variable that counts the number of cycles in $Core(G_n;r)$ whose k-type is O. Fix $O \in C(k,r)$. It holds

$$\mathrm{E}\big[X_{n,O}\big] = \sum_{(H,\tau)\in Copies(0,[n])} \Pr\left(H \subset G_n \bigwedge_{v\in V(H)} Tree(G_n,v;r) \in \tau(v)\right)$$

Because of the symmetry of the random model last probability is the same for all $(H,\tau) \in Copies(O,[n])$. Fix $(H,\tau) \in Copies(O,\mathbb{N})$. Then

$$\begin{aligned}
\mathrm{E}\big[X_{n,O}\big] &= \frac{(n)_{|V(H)|}}{|Aut(H,\tau)|} \cdot \Pr\left(H \subset G_n \bigwedge_{v\in V(H)} Tree(G_n,v;r) \in \tau(v)\right)\\
&= \frac{(n)_{|V(H)|}}{|Aut(H,\tau)|} \cdot \frac{\prod_{R\in\sigma} \beta_R^{|E_R(H)|}}{n^{|V(H)|}} \cdot \Pr\left(\bigwedge_{v\in V(H)} Tr(G_n,v;r) \in \tau(v) \;\middle|\; H \subset G_n\right)\\
&\sim \frac{\prod_{R\in\sigma} \beta_R^{|E_R(H)|}}{|Aut(H,\tau)|} \cdot \Pr\left(\bigwedge_{v\in V(H)} Tr(G_n,v;r) \in \tau(v) \;\middle|\; H \subset G_n\right)
\end{aligned}$$

Let $\overline{v} \in (\mathbb{N})_*$ be a list containing exactly the vertices in $V(H)$. If $H \subset G_n$ then $Tr(G_n,v;r) = Tr(G_n,\overline{v},v;r)$. Also, the event $H \subset G_n$ clearly can be described via an edge sentence concerning the vertices in $\overline{v}$. In consequence, using theorem 3.3, last expression is asymptotically equivalent to

$$\frac{\prod_{R\in\sigma} \beta_R^{|E_R(H)|}}{|Aut(H,\tau)|} \cdot \prod_{v\in V(H)} \Pr\big[r,\tau(v)\big].$$

For any $O \in C(k,r)$ we define λ_O and ω_O in the following way. Let (H,τ) be a representative of O. Then

$$\lambda_O := \prod_{v\in V(H)} \Pr\big[r,\tau(v)\big],$$

and

$$\omega_O := \frac{\prod_{R\in\sigma} \beta_R^{|E_R(H)|}}{|Aut(H,\tau)|} \cdot \lambda_O.$$

We are going to prove that the variables $X_{n,O}$ converge in distribution as n tends to infinity to independent Poisson variables whose respective means are the ω_O. For that we are going to use again the factorial moments method. For each $O \in C(k,r)$ fix a number $b_O \in \mathbb{N}$. We want to prove

$$\lim_{n\to\infty} \mathrm{E}\left[\prod_{O\in C(k,r)} \binom{X_{n,O}}{b_O}\right] = \prod_{O\in C(k,r)} \frac{(\omega_O)^{b_O}}{b_O!}.$$

For each $n \in \mathbb{N}$ we define

$$\Omega_n := \left\{ (F_O)_{O\in C(k,r)} \;\middle|\; \forall O \in C(k,r) \quad F_O \subset Copies(O,[n]), \quad |F_O| = b_O \right\}.$$

We also define $\Omega_{\mathbb{N}}$ by substituting $[n]$ for $\mathbb{N}$ in the definition of Ω_n. Informally, an element of Ω_n represents a choice of an unordered b_O-tuple of possible cycles over $[n]$ whose (k,r)-type is O, for each (k,r) type O. Using observation REF we obtain

$$\mathrm{E}\left[\prod_{O\in C(k,r)} \binom{X_{n,O}}{b_O}\right] = \sum_{(F_O)_{O\in C(k,r)}\in\Omega_n} \Pr\left(\bigwedge_{\substack{O\in C(k,r)\\ (H,\tau)\in F_O}} \left(H \subset G_n \bigwedge_{v\in V(H)} Tr(G_n,v;r) \in \tau(v) \right)\right).$$

Consider the subset $\Omega'_n \subset \Omega_n$ that contains the elements $(F_O)_{O\in C(k,r)} \in \Omega_n$ such that there exists some vertex $v \in [n]$ contained in two graphs $(H_1,\tau_1),(H_2,\tau_2) \in \bigcup_{O\in C(k,r)} F_O$. We want to argue that

$$\lim_{n\to\infty} \sum_{(F_O)_{O\in C(k,r)}\in\Omega'_n} \Pr\left(\bigwedge_{\substack{O\in C(k,r)\\ (H,\tau)\in F_O}} \left(H \subset G_n \bigwedge_{v\in V(H)} Tr(G_n,v;r) \in \tau(v) \right)\right) = 0. \tag{3}$$

Given an element $(F_O)_{O\in C(k,r)} \in \Omega_n$ we define the hypergraph $G\Big((F_O)_{O\in C(k,r)}\Big)$ as follows:

$$G\Big((F_O)_{O\in C(k,r)}\Big) := \bigcup_{H\in F} H,$$

where

$$F := \left\{ H \;\middle|\; (H,\tau) \in \bigcup_{O\in C(k,r)} F_O \right\}.$$

That is, $G\Big((F_O)_{O\in C(k,r)}\Big)$ is the union of all hypergraphs chosen in $(F_O)_{O\in C(k,r)}$. Then, for all $(F_O)_{O\in C(k,r)} \in \Omega_n$ it is satisfied

$$\Pr\left(\bigwedge_{\substack{O\in C(k,r)\\ (H,\tau)\in F_O}} \left(H \subset G_n \bigwedge_{v\in V(H)} Tr(G_n,v;r) \in \tau(v) \right)\right) \leq \Pr\left(\bigwedge_{\substack{O\in C(k,r)\\ (H,\tau)\in F_O}} H \subset G_n \right) =$$
$$\Pr\left(G\Big((F_O)_{O\in C(k,r)}\Big) \subset G_n \right).$$

Let

$$t = \sum_{O\in C(k,r)} |V(O)| \cdot b_O.$$

Then $V\Big(G\Big((F_O)_{O\in C(k,r)}\Big)\Big) \le t$ for any $(F_O)_{O\in C(k,r)} \in \Omega_n$.

Consider the following facts

(1) If $(F_O)_{O\in C(k,r)} \in \Omega'_n$ then $G\big((F_O)_{O\in C(k,r)}\big)$ is dense.

(2) Given an hypergraph H with $V(H) \subset \mathbb{N}$, the number of elements $(F_O)_{O\in C(k,r)} \in \Omega'_{\mathbb{N}}$ such that $H = G\big((F_O)_{O\in C(k,r)}\big)$ is finite and it is the same for all $H' \simeq H$ with $V(H') \subset \mathbb{N}$.

(3) There is a finite amount of unlabeled dense hypergraphs with size bounded by t.

Then it follows that

$$\sum_{(F_O)_{O\in C(k,r)}\in\Omega'_n} \Pr\Big(G\Big((F_O)_{O\in C(k,r)}\Big) \subset G_n\Big)$$
$$= \mathrm{O}\left(\mathrm{E}\left[\#\text{ of dense subgraphs in } G_n \text{ with size bounded by } t\right]\right).$$

And this, together with lemma 3.2 proves eq. (3).

For all n define $\Omega''_n = \Omega_n \setminus \Omega'_n$. That is, Ω''_n contains the elements $(F_O)_{O\in C(k,r)}$ in Ω_n such that all vertices $v \in [n]$ belong to at most one hypergraph $(H,\tau) \in \bigcup_{O\in C(k,r)} F_O$. We also define $\Omega''_{\mathbb{N}}$. Because of eq. (3) we have

$$\mathrm{E}\left[\prod_{O\in C(k,r)} \binom{X_{n,O}}{b_O}\right] = \sum_{(F_O)_{O\in C(k,r)}\in\Omega''_n} \Pr\left(\bigwedge_{\substack{O\in C(k,r)\\(H,\tau)\in F_O}} \left(H \subset G_n \bigwedge_{v\in V(H)} Tr(G_n,v;r) \in \tau(v)\right)\right) + o(1).$$

Because of the symmetry of the model the probability inside of last sum is the same for all elements $(F_O)_{O\in C(k,r)} \in \Omega''_n$. Also, counting all different vertices and automorphisms we obtain that

$$|\Omega''_n| = \frac{(n)_{\sum_{O\in C(k,r)} |V(O)|\cdot b_O}}{\prod_{O\in C(k,r)} b_O! \cdot |Aut(O)|^{b_O}}.$$

Fix $(F_O)_{O\in C(k,r)} \in \Omega''_{\mathbb{N}}$. Then

$$\lim_{n\to\infty} \mathrm{E}\left[\prod_{O\in C(k,r)} \binom{X_{n,O}}{b_O}\right] =$$
$$\lim_{n\to\infty} \frac{(n)_{\sum_{O\in C(k,r)} |V(O)|\cdot b_O}}{\prod_{O\in C(k,r)} b_O! \cdot |Aut(O)|^{b_O}} \cdot \Pr\left(\bigwedge_{\substack{O\in C(k,r)\\(H,\tau)\in F_O}} \left(H \subset G_n \bigwedge_{v\in V(H)} Tr(G_n,v;r) \in \tau(v)\right)\right).$$

It holds that the probability in last expression equals

$$\prod_{O\in C(k,r)} \left(\frac{\prod_{R\in\sigma} \beta_R^{|E_R(O)|}}{n^{|V(O)|}}\right)^{b_O} \cdot \Pr\left(\bigwedge_{\substack{O\in C(k,r)\\(H,\tau)\in F_O\\v\in V(H)}} Tr(G_n,v;r) \in \tau(v) \;\middle|\; \bigwedge_{\substack{O\in C(k,r)\\(H,\tau)\in F_O}} H \subset G_n\right).$$

Let $\overline{v} \in (\mathbb{N})_*$ be a list that contains exactly the vertices in $G\big((F_O)_{O\in C(k,r)}\big)$. Then the event

$$A_n := \bigwedge_{\substack{O\in C(k,r)\\(H,\tau)\in F_O}} H \subset G_n$$

can be written as an edge sentence concerning the vertices in $\overline{w}$. Also, if A_n holds then all vertices in $\overline{w}$ belong to $Core(G_n;r)$. Thus, for all $v \in \overline{v}$, $Tr(G_n,v;r) = Tr(G_n,\overline{w};r)$ and using theorem 3.3 we obtain

$$\Pr\left(\bigwedge_{\substack{O\in C(k,r)\\(H,\tau)\in F_O\\v\in V(H)}} Tr(G_n,v;r)\in\tau(v)\,\middle|\,\bigwedge_{\substack{O\in C(k,r)\\(H,\tau)\in F_O}} H\subset G_n\right) \sim \prod_{O\in C(k,r)} (\lambda_O)^{b_O}.$$

Joining everything together we obtain

$$\lim_{n\to\infty} \mathrm{E}\left[\prod_{O\in C(k,r)} \binom{X_{n,O}}{b_O}\right] =$$

$$\lim_{n\to\infty} \frac{(n)_{\sum_{O\in C(k,r)}|V(O)|\cdot b_O}}{\prod_{O\in C(k,r)} b_O!\cdot|Aut(O)|^{b_O}} \cdot \prod_{O\in C(k,r)} \left(\frac{\lambda_O\cdot\prod_{R\in\sigma}\beta_R^{|E_R(O)|}}{n^{|V(O)|}}\right)^{b_O} =$$

$$\prod_{O\in C(k,r)} \frac{1}{b_O!}\left(\frac{\lambda_O\cdot\prod_{R\in\sigma}\beta_R^{|E_R(O)|}}{|Aut(O)|}\right)^{b_O} = \prod_{O\in C(k,r)} \frac{(\omega_O)^{b_O}}{b_O!},$$

as we wanted. With this, because of theorem 3.1, it is proven that when n tends to infinity the $X_{n,O}$'s are asymptotically distributed like independent Poisson variables with the ω_O's as their respective means.

Given a simple k-agreeability class for hypergraphs $\mathscr{O}$ there is a partition $C_1,C_2\subset C(k,r)$, $C_1\cup C_2 = C(k,r)$ and there are natural numbers $a_O\le k-1$ for any $O\in C_2$ such that $C_1,C_2,(a_O)_{O\in C_2}$ depend only on $\mathscr{O}$ and the event $G_n\in\mathscr{O}$ is equivalent to

$$G_n \text{ is } r\text{-simple} \wedge \left(\bigwedge_{O\in C_1} X_{n,O}\ge k\right)\wedge\left(\bigwedge_{O\in C_1} X_{n,O}=a_O.\right).$$

In consequence

$$\lim_{n\to\infty}\Pr\big(G_n\in\mathscr{O}\big) =$$

$$\lim_{n\to\infty}\Pr\left(G_n \text{ is } r\text{-simple} \wedge \left(\bigwedge_{O\in C_1} X_{n,O}\ge k\right)\wedge\left(\bigwedge_{O\in C_1} X_{n,O}=a_O.\right)\right).$$

Because of theorem 3.2, last limit equals

$$\lim_{n\to\infty}\Pr\left(\left(\bigwedge_{O\in C_1} X_{n,O}\ge k\right)\wedge\left(\bigwedge_{O\in C_1} X_{n,O}=a_O.\right)\right) =$$

$$\left(\prod_{O\in C_1} 1-\sum_{i=0}^{k-1} e^{-\omega_O}\frac{(\omega_O)^i}{i!}\right)\cdot\left(\prod_{O\in C_2} e^{-\omega_O}\frac{(\omega_O)^{a_O}}{a_O!}\right).$$

This last expression belongs to Ω, so the theorem is proven. □

5 Proof of the main theorem

Theorem 5.1. *Let $\phi \in FO[\sigma]$. Then the function $F_\phi : [O,\infty)^{|\sigma|} \to [0,1]$ given by*

$$\overline{\beta} := (\beta_R)_{R\in\sigma} \mapsto \lim_{n\to\infty} \Pr\Big(G_n(\overline{\beta}) \models \phi\Big)$$

is well defined and it is given by a finite sum of expressions in Θ.

Proof. Let k be the quantifier rank of ϕ and let $r = 3^k$. Using theorem 3.2 we obtain

$$\lim_{n\to\infty} \Pr\Big(G_n(\overline{\beta}) \models \phi\Big) = \lim_{n\to\infty} \Pr\Big(G_n(\overline{\beta}) \models \phi \,\Big|\, G_n(\overline{\beta}) \text{ is } r\text{-simple}\Big).$$

Let S be the set of simple k-agreeability classes. Then the LHS of last equation equals

$$\lim_{n\to\infty} \sum_{\mathscr{O}\in S} \Pr\Big(G_n(\overline{\beta}) \in \mathscr{O}\Big) \cdot \Pr\Big(G_n(\overline{\beta}) \models \phi \,\Big|\, G_n(\overline{\beta}) \in \mathscr{O}\Big).$$

Notice that, because the set S is finite, this is the limit of a finite sum. Also, using REF, we obtain that for any $\mathscr{O} \in S$ it holds

$$\lim_{n\to\infty} \Pr\Big(G_n(\overline{\beta}) \models \phi \,\Big|\, G_n(\overline{\beta}) \in \mathscr{O}\Big) = 0 \text{ or } 1.$$

Let $S' \subset S$ be the set of classes $\mathscr{O}$ for which last limit equals 1. Then

$$\lim_{n\to\infty} \Pr\Big(G_n(\overline{\beta}) \models \phi\Big) = \sum_{\mathscr{O}\in S'} \lim_{n\to\infty} \Pr\Big(G_n(\overline{\beta}) \in \mathscr{O}\Big).$$

Because of theorem 4.1 we know that each of the limits inside last sum exists and is given by an expression that belongs to Θ. As a consequence the theorem follows. $\square$

6 Application to random SAT

References

[1] Paul Erdős and Alfréd Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60, 1960.

[2] Ronald Fagin. Probabilities on finite models 1. *The Journal of Symbolic Logic*, 41(1):50–58, 1976.

[3] Yu V Glebskii, Do I Kogan, MI Liogon'kiI, and VA Talanov. Range and degree of realizability of formulas in the restricted predicate calculus. *Cybernetics and Systems Analysis*, 5(2):142–154, 1969.

[4] Edgar N Gilbert. Random graphs. *The Annals of Mathematical Statistics*, 30(4):1141–1144, 1959.

[5] Saharon Shelah and Joel Spencer. Zero-one laws for sparse random graphs. *Journal of the American Mathematical Society*, 1(1):97–115, 1988.

[6] James F Lynch. Probabilities of sentences about very sparse random graphs. *Random Structures & Algorithms*, 3(1):33–53, 1992.

[7] Vašek Chvátal and Bruce Reed. Mick gets some (the odds are on his side)(satisfiability). In *Proceedings., 33rd Annual Symposium on Foundations of Computer Science*, pages 620–627. IEEE, 1992.

[8] H-D Ebbinghaus, Jörg Flum, and Wolfgang Thomas. *Mathematical logic*. Springer Science & Business Media, 2013.

[9] Andrzej Ehrenfeucht. An application of games to the completeness problem for formalized theories. *Fund. Math*, 49(129-141):13, 1961.

[10] Joel Spencer. *The strange logic of random graphs*, volume 22. Springer Science & Business Media, 2013.

[11] Béla Bollobás and Bollobás Béla. *Random graphs*. Number 73. Cambridge university press, 2001.