Contents

Introduction

Notation

Chapter 1

Preliminaries

1.1 Models of Random Graphs

1.2 First Order Logic

1.3 Ehrenfeucht Fraisse Games and the Logic of Random Graphs

Chapter 2

Probabilities of Sentences about Very Sparse Random Graphs

In this chapter we will review the results obtained in the paper with the same name by James F. Lynch [1]. In there, limit probabilities of sentences in the first order language of graphs $\mathcal{L}$ are discussed for the binomial model $G(n, p)$ in the cases $p = \beta/n$ and $p = \beta n^{-\alpha}$ with $\alpha = (l+1)/l$.

More precisely, it is proven that in those cases the probability of every sentence converges and it is shown that for any of those sentences, its limit probability is among the values taken by some analytic formulas with parameter β.

We are interested in the case $p = \frac{\beta}{n}$, which is the one discussed more extensively in [1]. According to the author, the relevant theorems for the other case can be proven analogously. From now on we will only refer as random graphs to the ones in $G(n, \beta/n)$

From now on we will denote by Poi_λ the probability function of the Poisson distribution with mean λ. That is, the one given by $Poi_\lambda(n) = e^{-\lambda}\lambda^n/n!$ for any $n \in \mathbb{N}$. Also, we define $Poi_\lambda(\leq n)$ and $Poi_\lambda(> n)$ as $\sum_{i=0}^{n} Poi_\lambda(n)$ and $1 - Poi_\lambda(\leq n)$ respectively. Notice that for a fixed n, both $Poi_\lambda(\leq n)$ and $Poi_\lambda(> n)$ can be considered real functions of parameter λ.

We define the following sets of functions. Let Λ be the smallest set of expressions with parameter β such that:

- $1 \in \Lambda$,
- For any $\lambda \in \Lambda$ and $i \in \mathbb{N}$, both $Poi_{\beta\lambda}(n)$ and $Poi_{\beta\lambda}(> n)$ are in Λ.
- For any $\lambda_1, \lambda_2 \in \Lambda$, $\lambda_1\lambda_2$ belongs to Λ as well.

And let Θ be the smallest set of functions with parameter β such that:

- For any $\lambda \in \Lambda$ and $n, a, i \in \mathbb{N}$, with $i \geq 3$, both $Poi_{\beta^i\lambda/a}(\leq n)$ and $Poi_{\beta^i\lambda/a}(> n)$ are in Θ.

The main result is the following:

Theorem 2.1 (Lynch, 1992). *Let ϕ be a sentence in the first order theory of graphs. Then the limit $\lim_{n \to \infty} P(G(n, \beta/n) \models \phi)$ exists for all positive real numbers β, and it is a finite sum of expressions in Θ.*

We show now an outline of the proof.

We show that for any quantifier rank k there are some classes of graphs $C_1^k, \ldots, C_{n_k}^k$ 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 expression in Θ.

After this is archived the theorem follows easily. Indeed, let ϕ be a sentence in the first order language $\mathcal{L}$ of graphs whose quantifier rank is k, and denote by G a random graph in $G(n, \beta/n)$. We define the events $E_1, \ldots, E_{n_k}$ as

$$E_i := (G \models \phi) \wedge (G \in C_i),$$

and the event F as

$$F := (G \models \phi) \bigwedge_{i=1}^{n_k} (G \notin C_i).$$

Then, for any $n \in \mathbb{N}$

$$P(G \models \phi) = \sum_{i=1}^{n_k} P(E_i) + P(F), \tag{1}$$

as the events E_i together with F form a partition of all the cases where G satisfies ϕ.

Fix and index $i \in \{1, \ldots, n_k\}$. From the property (1) of the classes $C_1, \ldots, C_{n_k}$ it follows that if G and H are random graphs, then

$$\lim_{n \to \infty} P((G \models \phi) \wedge \neg(H \models \phi) \,|\, G \in C_i \wedge H \in C_i) = 0.$$

This is because G and H share a.a.s the same rank k type if they both belong to C_i. In consequence the limit

$$\lim_{n \to \infty} P(G \models \phi \,|\, G \in C_i)$$

is either zero or one, and

$$\lim_{n \to \infty} P(E_i) = \lim_{n \to \infty} P(G \in C_i) \cdot P(G \models \phi \,|\, G \in C_i) = \text{ either } 0 \text{ or } \lim_{n \to \infty} P(G \in C_i). \tag{2}$$

Also, as a consequence of property (2) we obtain

$$\lim_{n \to \infty} P(\bigwedge_{i=1}^{n_k} G \notin C_i) = 0,$$

6

so

$$\lim_{n \to \infty} P(F) = \lim_{n \to \infty} P(\bigwedge_{i=1}^{n_k} G \notin C_i) \cdot P(G \models \phi \mid \bigwedge_{i=1}^{n_k} G \notin C_i) = 0. \qquad (3)$$

Taking limits in equation 1 and using equations 2 and 3 we get

$$\lim_{n \to \infty} P(G \models \phi) = \sum_{C_i \in \mathcal{C}} \lim_{n \to \infty} P(G \in C_i),$$

where $\mathcal{C}$ is a (possibly empty) subset of $\{C_1, \ldots, C_{n_k}\}$. Finally, because of property (3) for each i the limit $\lim_{n \to \infty} P(G \in C_i)$ is an expression in Θ. Thus $\lim_{n \to \infty} P(G \models \phi)$ is a finite sum of expressions in Θ and the theorem follows.

The objective of next sections will be to define the classes $C_1, \ldots, C_{n_k}$ and to show that they satisfy properties (1), (2) and (3). Later we will prove a stronger result, so we will allow ourselves to just sketch some of the proofs during this chapter.

2.1 Agreability Classes

It is known that $n^{-v/e}$ is the t hreshold probability for the appearance of a balanced graph of density v/e. In our case $v/e = 1$, so in consequence any connected graph H with $e(H) < v(H)$ a.a.s will not appear as a subgraphs of $G(n, \beta/n)$. It can be easily shown that such graphs H are exactly the ones containing more than one cycle.

If H is a connected graph with $v = e$, then H is an uni-cyclic graph. In this case, the number X_H of copies of H in $G(n, \beta/n)$ will asymptotically have non-zero bounded expectancy m. It does not take much work to prove, using Brun's sieve, that X_H converges in distribution to a Poisson random variable with mean m as n goes to infinity.

Finally, if H is a connected graph with $v > e$ then it must be a tree. Here the expected number of copies of H in $G(n, \beta/n)$ diverges asymptotically. Informally, trees of any kind will occur arbitrarily often.

This all means, in a sense, that a.a.s the only difference between large graphs in $G(n, \beta/n)$ lies in their uni-cyclic subgraphs. More precisely, because of the "locality" of first order logic of quantifier rank k we will only be interested in the "small" neighborhoods of the "short" cycles. Thus, our goal will be to classify uni-cyclic graphs in a way that respects equivalence under first order logic of quantifier rank k.

To make our classification suitable for proofs involving E.F. games we need to work graphs to which we "attach" labels. We define the set of symbols $Const = \{c_i\}_{i \in \mathbb{N}}$ as the set of constants. Also, we will denote by $Const_n$ the set $\{c_1, \ldots, c_n\}$.

Definition 2.1. A **co-labeling** of a graph $G = (V, E)$ is a map $\sigma : D \to V$, where $D \subset C$ is a finite set of constant symbols. Given $c_i \in D$, we will say that the vertex $\sigma(c_i)$ is labeled c_i. Equivalently, we can denote a labeling σ as a tuple $(c_{i_1}[x_1], \ldots, c_{i_m}[x_m])$ where each c_{i_j} is a constant symbol, and x_j is the vertex in V labeled c_{i_j}.

Definition 2.2. A **graph with constants**[1] $G(c_{i_1}[x_1], \ldots, c_{i_m}[x_m])$ is a graph G together with a co-labeling $(c_{i_1}[x_1], \ldots, c_{i_m}[x_m])$.

To keep our notation compact we will often drop he x_i's and say $G(c_{i_1}, \ldots, c_{i_m})$.

Definition 2.3. Let G be a graph with constants. A subgraph H of G is a graph with constants such that $V(H) \subseteq V(G)$, $E(H) \subseteq E(G)$ and all vertices in $V(H)$ have the same labels in H and G.

An important abuse of notation we are going to make will be to identify the constants c_i with their labeled vertices $\sigma(c_i)$. This way things like $c_i \sim c_j$ will make sense. In this context, notice that the expression $c_i = c_j$ is ambiguous because the vertices labeled c_i and c_j may be the same for some $i \neq j$, but the constant symbols c_i and c_j will be equal only if $i = j$. We will make sure to leave no room for ambiguity in this situations.

Proposition / Definition 2.1. Let $G = (V, E, c_{i_1}, \cdots, c_{i_m})$ be a connected graph with constants. Then it has a unique minimal connected subgraph H containing all its constants and cycles. We will call the **center** of G to such subgraph and denote it by $Center(G)$. If $\bar{G}$ is an arbitrary graph with constants, then its center $Center\bar{G}$ will be the union of the centers of its connected components.

Proof. TO DO $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

For an arbitrary graph with constants we define the metric $d(\cdot, \cdot)$ on $V(G)$ as the one such that $d(x, y)$ is the minimum length of a path connecting x and y in G or ∞ if such path does not exist. For any vertex $x \in V(G)$ and $r \in \mathbb{N}$ we define the co-labeled subgraph $N(x; r)$ as the ball of radius r centered at v. That is, the induced subgraph with vertex set

$$V(N(x; r)) = \{\, y \in V \,|\, d(x, y) \leq r \,\}.$$

In a similar vein, given $X \subseteq V(G)$ we define its neighborhood of radius r as the induced co-labeled subgraph $N(X; r)$ whose vertex set is

$$V(N(X; r)) = \{\, y \in V \,|\, \forall x \in X : \ d(x, y) \leq r \,\}.$$

Let $G = (V, E)$, and $V' \subseteq V$. Another important abuse of notation we will make is writing $H = (V', E)$ for a subgraph H to mean that the edge set of $E(H)$ is the one induced by $E(G)$ on V'.

---

[1] Compare with [1], where they are called "rooted graphs".

Definition 2.4. A **rooted tree** $T = (V, E, x)$ is a tree (V, E) with distinguished vertex $x \in V$ with we will call **root** of the tree.

Proposition / Definition 2.2. Let $G = (V, E, c_{i_1}, \cdots, c_{i_m})$ be a connected graph and $x \in V$. Then define $Tree(x, G)$ as the rooted tree

$$Tree(x, G) = (V_x, E, x),$$

where

$$V_x = \{\, y \in V \mid d(Center(G), y) = d(Center(G), x) + d(x, y)\}.$$

Proof. TO DO $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ □

The radius $r(T)$ of a rooted tree $T = (V, E, x)$ is the maximum distance between its root x and any other of its vertices. The branches of T are the rooted trees of the form $Tree(y, T)$, where $y \sim x$. We will denote by $Br(T)$ the set of branches of T.

We begin by defining an equivalence relation between rooted trees for each quantifier rank k.

Definition 2.5. Let $k \in \mathbb{N}$ with $k \geq 1$. The **k-morphism** equivalence relation $\overset{k}{\simeq}$ between graph with constantss is the one inductively defined as follows:

- If T_1, T_2 are rooted trees of radius 0 -i.e., they consist only of their roots- they are k-morphic.

- Let T_1, T_2 be rooted trees of radius r whose rots have the same label. Then $T_1 \overset{k}{\simeq} T_2$ if for any k-morphism class C of trees with radius less than r and root either

  "T_1 and T_2 have the same number of branches of type C"
  $$|Br(T_1) \cap C| = |Br(T_2) \cap C|,$$

  or

  "T_1 and T_2 have both more than k branches of type C"
  $$|Br(T_i) \cap C| \geq k + 1 \text{ for } i = 1, 2.$$

It follows from the definition that k-morphic trees have the same radius. It is also easy to check that the k-morphism relation is indeed an equivalence one.

Proposition 2.1. *For all $k, r \in N$ and with $k \geq 1$, the set of classes of k-morphic trees with radius lesser or equal than r is finite.*

Proof. TO DO $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ □

We define now the k-morphism relation for arbitrary graph with constantss.

Definition 2.6. Let $G^1 = (V^1, E^1, c_{i_1}[x_1^1], \ldots, c_{i_m}[x_m^1])$ and $G^2 = (V^2, E^2, c_{i_1}[x_2^1], \ldots, c_{i_m}[x_m^2])$ be graph with constantss with the same constant symbols. We will say that they are k-morphic (denoted by $G^1 \overset{k}{\simeq} G^2$) if there is a bijection $f : V(Center(G^1)) \to V(Center(G^2))$ such that

- "f preserves edges"

$$\forall x, y \in V(Center(G^1)): \quad x \sim y \iff f(x) \sim f(y).$$

- "f preserves labels"

$$\forall j \in \{1, \ldots, m\}: \quad f(x_j^1) = x_j^2.$$

- "f preserves k-morphism classes of trees"

$$\forall x \in V(Center(G^1)): \quad Tree(x, G^1) \overset{k}{\simeq} Tree(f(x), G^2).$$

In this case we will say that $G^1 \overset{k}{\simeq} G^2$ via f.

We are going to show that the rank k type of a random graph a.a.s only depends on the neighborhoods of its small cycles. In consequence the following definition is motivated:

Definition 2.7. Let $G = (V, E, c_{i_1}, \cdots, c_{i_m})$ be a graph with constants. Then its core of radius r, $Core(G, r)$ is the co-labeled subgraph $N(X; r)$, where X is the union of the (vertex sets of the) cycles in G with size at most $2r + 1$ and all of the labeled vertices in G.

Chapter 3

First Order Logic of Sparse Random General Hypergraphs

3.1 General Hypergraphs

$\forall n \in N : \quad [n] := \{1, \ldots, n\}$

S_n : symmetric group on n elements.

Let X be a set, then:

- $\Delta \subset X^2 : \Delta$ is the diagonal subset. $\Delta : \{(a, a) \in X^2\}$.

- S_n acts on X^n in the obvious way.

- When the subgroup ϕ is understood or (is not relevant) $\langle x_1, \ldots, x_n \rangle$ denotes the equivalence class of $(x_1, \ldots, x_n)$ in X^n/ϕ.

Definition 3.1. Let V be a finite set. Given a natural number $a \in \mathbb{N}$, a group $\phi \leq S_a$ and a subset $R \subseteq [a]^2 \setminus \delta$, the total (hiper-)edge set of size a, symmetry group ϕ and restrictions R is the set:

$$\mathcal{H}_{V,a}^{\phi,R} = V^n/\phi \ \setminus \{ \langle x_1, \ldots, x_n \rangle \in V^n/\phi \mid x_i = x_j \text{ for some } (i,j) \in R\}$$

Definition 3.2. An hypergraph $(V, H_1, \ldots, H_l)$ with edge sizes $a_1, \ldots, a_l$, symmetry groups $\phi_1, dots, \phi_1$ and restrictions $R_1, \ldots, R_l$ consists of

- A finite vertex set V.

- For $i = 1, \ldots, l$, an edge set $H_i \subseteq \mathcal{H}_{V,a_i}^{\phi_i, R_i}$.

Definition 3.3. Random model $G(n, p_1, \ldots, p_l)$ with edge sizes $a_1, \ldots, a_l$, symmetry groups $\phi_1, dots, \phi_1$ and restrictions $R_1, \ldots, R_l$:

- We obtain a hypergraph $(V, H_1, \ldots, H_l)$ with $V = [n]$

- The probability of each edge $e \in \mathcal{H}_{V,a_i}^{\phi_i,R_i}$ is p_i independently.

From now on we fix the edge sizes $a_1, \ldots, a_l$, symmetry groups $\phi_1, dots, \phi_1$ and restrictions $R_1, \ldots, R_l$.

In a graph $G(V, E_1, \ldots, E_l)$, the type of an edge e is the index i such that $e \in E_i$.

The language is the first order relational language with relations $R_1, ..., R_l$ of arities $a_1, \ldots, a_l$ respectively. The interpretation on $(V, H_1, \ldots, H_l)$ is the evident one.

Definition 3.4. The characteristic $ch(G)$ of an hypergraph $G = (V, E_1, \ldots, E_l)$ is

$$\left(\sum_{i=1}^{l} |E_i|(a_i - 1) \right) - |V|.$$

An hypergraph is c-balanced if it contains no subgraph of greater characteristic than itself.

Proposition 3.1. *A connected graph cannot have characteristic lesser than* -1.

Definition 3.5. A unicycle is a connected c-balanced graph with characteristic 0.

Definition 3.6. A cluster is a minimal c-balanced graph.

Remark: Note that clusters have to be connected.

Definition 3.7. A cycle is a minimal unicycle, or equivalently a characteristic 0 cluster.

Proposition 3.2. *Any unicycle contains exactly one cycle.*

Definition 3.8. A (hiper)forest is a graph with no clusters. A (hiper)tree is a connected forest.

Definition 3.9. A path between x and y is a connected graph containing both x and y that is minimal among the ones with those properties.

Proposition 3.3. *There is a unique path between any two vertices of a tree.*

We define the constant set $Const := \{c_i\}_{i \in \mathbb{N}}$.

Definition 3.10. As before, a (hiper)graph with constants $G(c_{i_1}, \cdots, c_{i_m})$ is a hypergraph G together with a co-labeling.

Definition 3.11. The center of a connected graph with constants $G(c_{i_1}, \cdots, c_{i_m})$ is its minimal connected subgraph containing all the constants and clusters. If $G(c_{i_1}, \cdots, c_{i_m})$ is not connected then its center is the union of the centers of its connected components.

We define a distance over $G(c_{i_1}, \cdots, c_{i_m})$.

$$d(x, y) = (\text{minimum size of a connected graph containing } x \text{ and } y) - 1$$

If such graph does not exist we define $d(x, y) = \infty$.

Given $x \in V(G)$ and $X \subseteq V(G)$ we define $N(x; r)$ and $N(X; r)$ as

$$N(x; r) = G[S], \quad \text{where } S = \{\, y \in V(G) \,|\, dist(x, y) \le r \,\}$$

$$N(X; r) = G[S], \quad \text{where } S = \{\, y \in V(G) \,|\, dist(X, y) \le r \,\}.$$

Definition 3.12. A rooted (hiper)tree is a tree with a distinguished vertex called root. The radius of the tree is the maximum distance between its root and any of its vertices.

In a rooted tree (T, x) all the (hiper)edges can be rooted in a canonical way. The root of an edge e is the vertex $y \in e$ such that $dist(x, y) = dist(x, e)$.

Definition 3.13. Let $G(c_{i_1}, \cdots, c_{i_m})$ be a graph with constants and $x \in V(G)$. Then $Tree(x, G) = (G[V_x], x)$, where

$$V_x = \{\, y \in V \,|\, dist(Center(G), y) = dist(Center(G), x) + dist(x, y) \,\}.$$

Definition 3.14. The k-morphism equivalence relation $\overset{k}{\simeq}$ between rooted trees, and the $\overset{k}{\sim}$ relation for rooted (hyper)edges are defined inductively as follows:

- If T_1 and T_2 are radius 0 trees then they consist only of their roots and $T_1 \overset{k}{\simeq} T_2$.

- If e_1 and e_2 are edges in rooted trees T_1, T_2 of radii at most r, then $e_1 \overset{k}{\sim} e_2$ if e_1 and e_2 are edges of the same type and $e_1 = [x_1, \ldots, x_m]$, $e_2 = [y_1, \ldots, y_m]$ in a way that

  - x_i is the root of e_1 if and only if y_i is the root of e_2.
  - If x_j is not the root of e_1, then $Tree(x_j, T_1) \overset{k}{\simeq} Tree(y_j, T_2)$.

- If T_1 and T_2 are trees of radii at most r and roots x and y, then $T_1 \overset{k}{\simeq} T_2$ means that for any $\overset{k}{\sim}$ class of edges C either

  - the number of edges in T_1 of type C containing x is the same as the number of edges in T_2 of tipe C containing y, or
  - both T_1 and T_2 contain no less than $k + 1$ edges of type C containing x and y respectively.

Definition 3.15. Let $G_1(c_{i_1}, \cdots, c_{i_m})$ and $G_2(c_{i_1}, \cdots, c_{i_m})$ be graphs with the same constants. Then we say $G_1(c_{i_1}, \cdots, c_{i_m}) \overset{k}{\simeq} G_2(c_{i_1}, \cdots, c_{i_m})$ if there is a bijection $f : V(Center(G_1)) \to V(Center(G_2))$ such that

- $f(c_i^1) = c_i^2$ for all constants, and

- $Tree(x, G_1) \overset{k}{\simeq} Tree(f(x), G_2)$ for all $x \in V(Center(G_1))$.

The diameter of a (hiper)graph is the maximum distance between any two of its vertices.

Definition 3.16. Let $G(c_{i_1}, \cdots, c_{i_m})$ be a graph with constants. We define $Core(G, r)$ to be $N(X, r)$ where X is the union of constants and clusters of diameter at most r.

Definition 3.17. A (hyper)graph G is r-simple if $Core(G, r)$ is a union of unicycles.

Definition 3.18. A (hyper)graph G is j, r-rich for some $j, r \in \mathbb{N}$, if for any rooted tree T of radius at most r there are $x_1, \ldots, x_j \in V(G)$ such that

- The $N(x_l; r)$'s are disjoint.

- The $N(x_l; r)$'s do not intersect $Core(G, r)$

- $N(x_l; r)(x_l) \overset{j}{\simeq} T$.

Theorem 3.1. *If $G_1(c_1, \ldots, c_m) \overset{k}{\simeq} G_2(c_1, \ldots, c_m)$, then for any $x \in V(G_1)$ such that $d(x, Center(G_1)) \neq \infty$ there is a $y \in V(G_2)$ such that $G_1(c_1, \ldots, c_m, c_{m+1}[x]) \overset{k-1}{\simeq} G_2(c_1, \ldots, c_m, c_{m+1}[y])$.*

Theorem 3.2. *If $Core(G_1, r) \overset{k}{\simeq} Core(G_2, r)$ then for any $s < r$, $Core(G_1, s) \overset{k}{\simeq} Core(G_2, s)$.*

Definition 3.19. Two graphs with constants G_1 and G_2 are k-agreeable if for each k-morphism equivalence class C they either have the same number of connected components of type C or they both have no less than $k + 1$ components of type C.

Theorem 3.3. *Let G_1, G_2 be graphs with constants satisfying:*

- *They are both $(k, k, 3^k)$-rich.*

- *$Core(G_1, 3^k)$ is k-agreeable with $Core(G_2, 3^k)$.*

Then the G_1 and G_2 share the same rank k type.

From now on fix $\beta_1, \ldots, \beta_l$ positive real numbers.

Lemma 3.1. *Let H be a cluster with $ch(H) > 0$, and let X_n be the random variable that counts the number of times that H appears as a subgraph of $G(n, p(n))$. Then*

$$\lim_{n \to \infty} Pr(X_n > 0) = 0.$$

14

Proof. Let $v = |V(H)|$ and $e_i = |E_i(H)|$ for $i = 1, \ldots, l$. The criticality condition on H means

$$\langle e, a - 1 \rangle - v > 0.$$

Chose a ordering of the vertices in H. For any ordered sequence of vertices $S = (x_1, \ldots, x_v)$, let $X_{n,S}$ be the indicator variable that equals 1 if H is a subgraph of $G[S]$ (in a way that respects the ordering) and 0 otherwise. Clearly X_n is the sum of all the $X_{n,S}$'s, so

$$E(X_n) = \frac{n(n-1)\cdots(n-v+1)}{b} \prod_{i=1}^{l} \left(\frac{c_i}{n^{a_1-1}}\right)^{e_i},$$

where b is the carnality of H's group of isomorphisms. Then because of the criticality hypothesis

$$\lim_{n\to\infty} E(X_n) = 0,$$

and using the first moment method the result follows. $\qquad\square$

Lemma 3.2. *Let H be a critical graph of characteristic r. Then H contains a critical subgraph 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 H edges e such that $dist(x, e)$ is maximum until the resulting graph H' has characteristic no greater than 0. We have two cases:

- $ch(H') = 1$. Let $e = [x_1, \ldots, x_b]$ be the last removed edge and $e \cap H' = \{x_{i_1}, \ldots, x_{i_c}\}$. For any $j = 1, \ldots, c$ choose P_j a path of size no greater than $r + 1$ joining x and x_{i_j} in H'. Then $P_1 \cup \ldots P_c \cup e$ is a critical subgraph of H of size less than $a(r+1) + a < (a+2)(r+1) + 2a$.

- $ch(H') = 0$. Let $e_1 = [x_1, \ldots, x_{b_1}]$ be the last removed edge. Continue removing the edges of H' that are at maximum distance from x until you obtain H'' with $ch(H') = -1$. Let $e_2 = [y_1, \ldots, y_{b_1}]$ be the last removed edge. As before, let $e_1 \cap H' = \{x_{i_1}, \ldots, x_{i_c}\}$ and for $j = 1, \ldots, c$ 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 characteristic 0 and size less than $2r+2+a$, and $Q_1 \cup Q_2 \cup P_1 \cup \cdots \cup P_c \cup e_1 \cup e_2$ is a critical graph with size less than $(2 + a)(r + 1) + 2a$

$\qquad\square$

Corollary 3.1. *Let A_n be the event that $G(n, p(n))$ contains critical subgraph with diameter no greater than r. Then*

$$\lim_{n\to\infty} Pr(A_n) = 0.$$

Proof. If a random graph contains G such critical graph, then by the previous lemma it has to contain a critical graph of size less than some constant M. The number of critical graphs of such size is finite and the probability that any one of those appears as a subgraph of G is asymptotically zero. $\qquad\square$

Theorem 3.4. *For any r,*

$$\lim_{n\to\infty} Pr(G(n,\beta_1/n^{a_1-1},\dots,\beta_l/n^{a_l-1}) \text{ is } r\text{-simple}) = 1.$$

Lemma 3.3. *Let (T_1,x) and (T_2,y) be rooted trees such that for some $k > 0$ $T_1 \overset{k}{\simeq} T_2$. Then, for any $0 \le j < k$, it is satisfied $T_1 \overset{j}{\simeq} T_2$.*

Proof. Fix $j < k$. The proof is by induction on the radius of T_1 and T_2.

- If they have radius at most 0 they are j-morphic by definition.

- Suppose now that they have radii at most r. Let e_1 be an edge with root x in T_1 and e_2 be an edge of root y in T_2 such that $e_1 \overset{k}{\simeq} e_2$. Then $e_1 = [x_1,\dots,x_a]$ and $e_2 = [y_1,\dots,y_a]$ for some vertices such that for all $i = 1,\dots,a$ either

  - $Tree(x_i) \overset{k}{\simeq} Tree(y_i)$, and because of the induction hypothesis $Tree(x_i) \overset{j}{\simeq} Tree(y_i)$.
  - Or $x_i = x$ and $y_i = y$.

In consequence $e_1 \overset{j}{\simeq} e_2$. Then, any class C of j-morphic edges in trees of radius at most r is union of classes $D_1,\dots,D_s$ of k-morphic edges in trees of radius at most r. Then

$$(T_1, C) = \sum_{i=1}^{s}(T_1, D_i), \text{ and } (T_2, C) = \sum_{i=1}^{s}(T_2, D_i).$$

Both of the above sums coincide if all the (T_1, D_i)'s are not greater than k or both are greater than k otherwise. As $j \le k$ this implies

$$\text{Either } (T_1, C) = (T_2, C) \text{ or } (T_1, C), (T_2, C) \ge j + 1,$$

and $T_1 \overset{j}{\simeq} T_2$.

$\qquad\square$

Corollary 3.2. *Let $G_1(c_1,\dots,c_t)$, and $G_2(c_1,\dots,c_t)$ be graphs with constants such that $G_1(c_1,\dots,c_t) \overset{k}{\simeq} G_2(c_1,\dots,c_t)$. Then, for any $j < k$, $G_1(c_1,\dots,c_t) \overset{j}{\simeq} G_2(c_1,\dots,c_t)$.*

Lemma 3.4. *Let (T_1,x) and (T_2,y) be rooted trees such that for some $k > 0$ $T_1 \overset{k}{\simeq} T_2$. Then, for any $r \ge 0$, $N(x;r) \overset{k}{\simeq} N(y;r)$.*

Proof. The proof is by induction on the radius of T_1 and T_2.

- If they have at most radius 0 the statement is vacuously true.

- Suppose now that they have radii at most r. Let e_1 be an edge with root x in $N(x; r)$ and e_2 be an edge of root y in $N(y; r)$ such that $e_1 \overset{k}{\simeq} e_2$ when both edges are considered in T_1 and T_2 respectively. We want to prove that $e_1 \overset{k}{\simeq} e_2$ as well when they are considered as edges in $N(x; r)$ and $N(y; r)$. By definition we can write $e_1 = [x_1, \ldots, x_a]$ and $e_2 = [y_1, \ldots, y_a]$ for some vertices such that for all $i = 1, \ldots, a$ either

  – $Tree(x_i) \overset{k}{\simeq} Tree(y_i)$, and because of the induction hypothesis

  $$Tree(x_i) \cap N(x_i; r - a) \overset{k}{\simeq} Tree(y_i) \cap N(y_i; r - a).$$

  – Or $x_i = x$ and $y_i = y$.

  In consequence $e_1 \overset{k}{\simeq} e_2$ as edges in $N(x; r)$ and $N(y; r)$. This implies that for any k-morphism class of rooted edges C of trees with radii less than r there are k-morphism classes $D_1, \ldots, D_c$ of rooted edges of trees with radii at most r such that

  $$(N(x; r), C) = (T_1, D_1) \cup \cdots \cup (T_1, D_a),$$

  and

  $$(N(y; r), C) = (T_2, D_1) \cup \cdots \cup (T_2, D_a).$$

  And we can easily obtain that either $|(N(x; r), C)| = |(N(y; r), C)|$ or $|(N(x; r), C)|$, $|(N(y; r), C)| \geq k + 1$. Thus $N(x; r) \overset{k}{\simeq} N(y; r)$.

  $\square$

Lemma 3.5. *Let* $(T_1, x), (T_1', x), (T_2, y)$ *and* (T_2', x) *be rooted trees satisfying* $T_1 \overset{k}{\simeq} T_2$. $T_1' \overset{k}{\simeq} T_2'$ *for some* $k \geq 0$ *and* $V(T_1) \cap V(T_1') = x$, $V(T_2) \cap V(T_2') = y$. *Then* $T_1 \cup T_1' \overset{k}{\simeq} T_2 \cup T_2'$.

Proof. Let C be a k-morphism class of rooted edges. Then

$$|(T_1 \cup T_1', C)| = |(T_1, C)| + |(T_1', C)|, \text{and}$$

$$|(T_2 \cup T_1', C)| = |(T_2, C)| + |(T_2', C)|.$$

And it follows easily that either

$$|(T_1 \cup T_1', C)| = |(T_2 \cup T_2', C)|,$$

or both quantities are greater than k. $\square$

Lemma 3.6. *Let (T_1, x), (T_2, y) be rooted trees and let (T_1', x) and (T_2', y) be sub-trees of T_1 and T_2 respectively. Let $k \geq 0$. If there is an isomorphism $f : T_1' \to T_2'$ such that $f(x) = y$ and for all $x_1 \in V(T_1')$*

$$Tree(x_1, T_1') \overset{k}{\simeq} Tree(f(x_1), T_2'),$$

then $T_1 \overset{k}{\simeq} T_2$.

Proof. The proof is, again, by induction on the radius of T_1' and T_2'.

- If T_1' and T_2' have radius zero, then they consist only of x and y respectively.

- Suppose that T_1' and T_2' have at most r. Let $x_1 \neq x$ be a vertex in an edge e rooted at x such that $e \in T_1'$. Then $Tree(x_1, T_1)$ and $Tree(x_1, T_1')$ satisfy the hypothesis of the lemma, and the radius of $Tree(x_1, T_1')$ is strictly less than r. Then by the induction hypothesis $Tree(x_1, T_1) \overset{k}{\simeq} Tree(f(x_1), T_2)$. In consequence $e \overset{k}{\simeq} f(e)$ and using the previous lemma successively we get $T_1 \overset{k}{\simeq} T_2$.

$\square$

Theorem 3.5. *Let $G_1(c_1, \ldots, c_t)$, and $G_2(c_1, \ldots, c_t)$ be graphs with constants such that for some $r \in N$*

$$Core(G_1(c_1, \ldots, c_t); r) \overset{k}{\simeq} Core(G_2(c_1, \ldots, c_t); r)$$

through f. Then, for any $s < r$

$$Core(G_1(c_1, \ldots, c_t); s) \overset{k}{\simeq} Core(G_2(c_1, \ldots, c_t); s).$$

Proof. Fix $s < r$. Let us introduce some notation

$$H_i := Core(G_i(c_1, \ldots, c_t); r) \quad \text{for } i = 1, 2.$$

$$H_i' := Core(G_i(c_1, \ldots, c_t); s) \quad \text{for } i = 1, 2.$$

One can check $H_i' \subseteq H_i$ for $i = 1, 2$, and that the isomorphism $f : Center(H_1) \to Center(H_2)$ restricts to one between $Center(H_1')$ and $Center(H_2')$. Let $v \in Center(H_1')$. Let $T = Tree(v, H_1') \cap Center(H_1)$. It is not hard to see that T is connected and in consequence is a tree. The following identity holds:

$$Tree(v, H_1') = (Tree(v, H_1) \cap N(v; s)) \cup Tree(T, H_1')$$

Also, for any $w \in V(T)$

$$Tree(w, Tree(v, H_1'), T) = Tree(w, H_1) \cap N(w; s - dist(Center(H_1'), w)).$$

So, by lemma REF

$$Tree(w, Tree(v, H'_1), T) \stackrel{k}{\simeq} Tree(f(w), Tree(f(v), H'_2), f(T)).$$

In consequence, by lemma REF,

$$Tree(T, H'_1) \stackrel{k}{\simeq} Tree(f(T), H'_2).$$

Using that

$$Tree(v, H_1) \cap N(v; s) \stackrel{k}{\simeq} \stackrel{k}{\simeq} Tree(v, H_1) \cap N(v; s),$$

and lemma REF we obtain

$$Tree(v, H'_1) \stackrel{k}{\simeq} Tree(f(v), H'_2).$$

Hence, $H'_1 \stackrel{k}{\simeq} H'_2$, as desired. $\qquad\square$

Lemma 3.7. *Let* (T_1, x), (T_2, y) *be rooted trees such that for some* $k \geq 0$ $T_1 \stackrel{k}{\simeq} T_2$. *Let* e_1 *and* e_2 *be initial edges of* T_1 *and* T_2 *such that* $e_1 \stackrel{k}{\simeq} e_2$. *Then* $T_1 \backslash e_1 \stackrel{k-1}{\simeq} T_2 \backslash e_2$.

Proof. For any k-morphism class C of rooted edges clearly either

$$|(T_1 \backslash e_1, C)| = |(T_2 \backslash e_2, C)|,$$

or both quantities are greater than $k - 1$. Now, using that any $k - 1$ morphism class of rooted edges is union of k morphism classes the result follows. $\qquad\square$

Lemma 3.8. *Let* (T_1, x_1), (T_2, x_2) *be rooted trees such that for some* $k \geq 0$ $T_1 \stackrel{k}{\simeq} T_2$. *For* $i = 1, 2$, *given a vertex* $v \in V(T_i)$ *let us denote by* $P(v)$ *the unique path between* v *and* x_i. *Then for any* $v \in V(T_1)$ *there is a vertex* $w \in V(T_2)$ *and an isomorphism* $f : P(v) \to P(w)$ *such that*

(1) $f(x_1) = x_2$ *and* $f(v) = w$.

(2) *For any edge* $e \in E(P(v))$, $e \stackrel{k}{\simeq} f(e)$.

(3) *For any vertex* $y \in V(P(v))$, $Tree(y, T_1) \stackrel{k}{\simeq} Tree(f(y), T_2)$.

Proof. The proof is by induction on $d(x_1, v)$.

- If $d(x_1, v) = 0$ then $x_1 = v$ and the statement is true taking $w = x_2$.

- Suppose now that $d(x_1, v) = r$. Then one can write the path $P(v)$ as a succession of edges $e_1, e_2, \ldots, e_s$, where $v \in e_s$. Let v' be the root of e_s. Then $d(x_1, v') = r - a(e_s) + 1$. Thus, by the induction hypothesis there exists w' such that there is an isomorphism $f : P(v') \to P(w')$ with the

required properties. In particular, $Tree(v', T_1) \overset{k}{\simeq} Tree(w', T_2)$, so there is an edge e' rooted at w' with the same k-type as e_s. Then one can write

$$e_s = [v_1, \ldots, v_a], \quad e' = [w_1, \ldots, w_a],$$

in a such a way that for some i, $v_i = v'$ and $w_i = w'$ and for all $j \neq i$ $Tree(v_j, T_1) \overset{k}{\simeq} Tree(w_j, T_2)$. Let j be such that $v = v_j$. Then we can take $w = w_j$, and extend the isomorphism f in the evident way.

$\square$

Theorem 3.6. *Let $G_1(c_1, \ldots, c_t)$, and $G_2(c_1, \ldots, c_t)$ be graphs with constants such that*

$$Core(G_1(c_1, \ldots, c_t); r) \overset{k}{\simeq} Core(G_2(c_1, \ldots, c_t); r)$$

by means of f. Then, for any vertex $x \in Core(G_1(c_1, \ldots, c_t); r)$ there is a vertex $y \in Core(G_2(c_1, \ldots, c_t); r)$ such that

$$Core(G_1(c_1, \ldots, c_t, c_{t+1}[v_1]); r) \overset{k-1}{\simeq} Core(G_2(c_1, \ldots, c_t, c_{t+1}[v_2]); r)$$

Proof. Let us introduce some notation:

$$H_i := Core(G_i(c_1, \ldots, c_t); r) \quad \text{for } i = 1, 2.$$

$$H_1' := Core(G_1(c_1, \ldots, c_t, c_{t+1}[v_1]); r).$$

The vertex v_1 belongs to $Tree(x_1, H_1)$ for a unique $x_1 \in Center(H_1)$. By the previous lemma there exist a vertex v_2 in $Tree(f(x_1), H_2)$ such that the path $P(v_1)$, joining v_1 and x_1, is isomorphic to the path $P(v_2)$, joining v_2 and $f(x_1)$, through an isomorphism $f' : P(v_1) \to P(v_2)$ satisfying properties (1), (2) and (3). Let

$$H_2' := Core(G_1(c_1, \ldots, c_t, c_{t+1}[v_2]); r).$$

We are going to show that $H_1' \overset{k-1}{\simeq} H_2'$. Clearly $Center(H_i') = Center(H_i) \cup P(v_1)$ for $i = 1, 2$, so we can glue f and f' into an isomorphism $g : Center(H_1') \to Center(H_2')$. Let $w_1 \in H'$. Then either

- $Tree(w_1, H_1)$ contains no edges in $P(v_1)$. In this case $Tree(w_1, H_1') = Tree(w_1, H_1)$. Thus, $Tree(w_1, H_1') \overset{k}{\simeq} Tree(g(w_1), H_1')$ and $Tree(w_1, H_1') \overset{k-1}{\simeq} Tree(g(w_1), H_1')$ in consequence by lemma REF.

- $Tree(w_1, H_1)$ contains edges from $P(v_1)$. In this case, $Tree(w_1, H_1)$ contains exactly one initial edge e_1 in $P(v_1)$, and $Tree(w_1, H_1') = Tree(w_1, H_1) \setminus e_1$. One can check that $Tree(g(w_1), H_2)$ contains exactly one edge in $P(v_2)$, namely $g(e_1)$, and $Tree(g(w_1), H_2') = Tree(g(w_1), H_2) \setminus g(e_1)$. We had $Tree(w_1, H_1) \overset{k}{\simeq} Tree(g(w_1), H_2)$, so by lemma REF $Tree(w_1, H_1') \overset{k-1}{\simeq} Tree(g(w_1), H_2')$.

□

Definition 3.20. Let G_1, G_2 be graphs with constants. We say that G_1 and G_2 are k-agreeable if for each $\overset{k}{\simeq}$ class C of connected graphs either

- G_1 and G_2 have the same number of connected components of type C, or

- Both G_1 and G_2 have no less than k connected components of type C.

Theorem 3.7. *Let G_1, G_2 be $k, 3^k$-rich graphs such that $Core(G_1; 3^k)$ is k-agreeable with $Core(G_2; 3^k)$. Then Duplicator has a winning strategy in the E.F game on G_1 and G_2.*

Proof. We will show, by induction on i, that Duplicator can play in such a way that in the i-th round $Core(G_1(c_1[x_1], \ldots, c_i[x_i]); 3^{k-i})$ is $k - i$-agreeable with $Core(G_2(c_1[y_1], \ldots, c_i[y_i]); 3^{k-i})$, where for each $j = 1, \ldots, i, 0$ x_j, y_j are the j-th chosen vertex in the graphs G_1 and G_2 respectively. After this the theorem follows because $Core(G_1(c_1[x_1], \ldots, c_i[x_k]); 1)$ being 0-agreeable with $Core(G_2(c_1[y_1], \ldots, c_i[y_k]); 1)$ implies that the map given by $x_i \mapsto y_i$ defines a partial isomorphism.

- For $i = 0$ the statement is true by hypothesis.

- Assume now that $Core(G_1(c_1[x_1], \ldots, c_i[x_{i-1}]); 3^{k-i+1})$ is $k - i + 1$-agreeable with $Core(G_2(c_1[y_1], \ldots, c_i[y_{i-1}]); 3^{k-i+1})$. Without loss of generality we can suppose that Spoiler chooses a vertex x_i in G_1 in the i-th round. We have two cases.

Case 1. $N(x_i; 3^{k-i})$ is contained in $Core(G_1(c_1[x_1], \ldots, c_i[x_{i-1}]); 3^{k-i+1})$. Let H_1 be the connected component of $Core(G_1(c_1[x_1], \ldots, c_i[x_{i-1}]); 3^{k-i+1})$ containing $N(x_i; 3^{k-i})$. Then there is a connected component H_2 in $Core(G_2(c_1[y_1], \ldots, c_i[y_{i-1}]); 3^{k-i+1})$ such that $H_1 \overset{k-i+1}{\simeq} H_2$. Applying theorems REF and REF successively Duplicator can choose $y_i \in V(H_2)$ such that

$$Core(H_1(c_i[x_i]); 3^{k-1}) \overset{k-i}{\simeq} Core(H_2(c_i[y_i]); 3^{k-1}).$$

Using Theorem Ref and counting the connected components in each $k - i$-morphism class one can check that now $Core(G_1(c_1[x_1], \ldots, c_i[x_i]); 3^{k-i})$ is $k - i$-agreeable with $Core(G_2(c_1[y_1], \ldots, c_i[y_i]); 3^{k-i})$.

Case 2. $N(x_i; 3^{k-i})$ is contained in $Core(G_1(c_1[x_1], \ldots, c_i[x_{i-1}]); 3^{k-i+1})$. Then $N(x_i; 3^{k-i})$ is disjoint from $Core(G_1(c_1[x_1], \ldots, c_i[x_{i-1}]); 3^{k-i})$ and in particular, $N(x_i; 3^{k-i})$ is a tree. As G_2 was originally $k, 3^k$-rich Duplicator can choose y_i in G_2 such that $N(y_i; 3^{k-i})$ is disjoint from $Core(G_2(c_1[y_1], \ldots, c_i[y_{i-1}]); 3^{k-i})$ and the tree $N(x_i; 3^{k-i})$ rooted at x_i is $k - i$-morphic to the tree $N(y_i; 3^{k-i})$ rooted at y_i. Counting connected components of each type we can conclude that $Core(G_1(c_1[x_1], \ldots, c_i[x_i]); 3^{k-i})$ is $k-i$-agreeable with $Core(G_2(c_1[y_1], \ldots, c_i[y_i]); 3^{k-i})$.

□

Define the families of expressions with arguments $\beta_1, \ldots, \beta_l$: Λ, M and Θ.

For any sentence ϕ we use the notation $Pr_n(\phi) := Pr(G \models \phi)$.

For any formula ϕ with free variables $x_1, \ldots, x_l$, we define $Pr_n(\phi(x_1, \ldots, x_l)) = \sum_{G \models \phi(a_1, \ldots, a_l)} Pr_n(G)$, where $a_1, \ldots, a_l$ are fixed **different** natural numbers in $[n]$. If ϕ and σ are formulas with possibly some free variables, then we define $Pr_n(\phi \,|\sigma) = Pr_n(\phi \wedge \sigma)/Pr_n(\sigma)$.

We introduce some notation now. For any numbers $l, r \in \mathbb{N}$ we denote by $\phi_r(x_1, \ldots, x_l)$ the formula with free variables $x_1, \ldots, x_l$ satisfying $G \models \phi_r(x_1, \ldots, x_l)$ if for any y there is a unique (minimal) path from y to the set of x_i's.

Lemma 3.9. *Let σ be an open formula (i.e., a formula with no quantifiers) with free variables $x_1, \ldots, x_l$. Then, for any $r \in \mathbb{N}$*

$$\lim_{n \to \infty} Pr_n(\phi_r(x_1, \ldots x_l) \,|\sigma) = 1.$$

Proof. We will se that

$$\lim_{n \to \infty} Pr_n(\neq \phi_r(x_1, \ldots, x_l) \,|\sigma) = 0.$$

Fix n and $x_1, \ldots, x_l \in [n]$. Notice that without loss of generality we can assume that σ Boolean combination of atomic formulas of the form $R_i(y_1, \ldots, y_{a_i})$, where the y_i's are among the free variables $x_1, \ldots, x_l$. In particular $G \models \sigma(x_1, \ldots, x_l)$ and $G \models \phi_r(x_1, \ldots, x_l)$ are independent events, because ϕ_r depends only on the edges not in $x_1, \ldots, x_l$. Thus, $Pr_n(\neg\phi_r \,|\sigma) = Pr_n(\neg\phi_r)$. Consider G a random graph in $G(n, p(n))$. If ϕ_r is not satisfied in G then there exists a y different from the x_i's and two paths P_1, P_2 between y and the set of x_i's with $V(P_i) \leq r+1$. We have two cases

- P_1 and P_2 contain some vertex $z \neq y$ in their intersection. Then, using that $ch(P_1), ch(P_2) \geq -1$ and counting we get that $P_1 \cup P_2$ is a critical graph of size no greater than $2r+1$.

- The union $P_1 \cup P_2$ is a path between some x_{i_1}, x_{i_2} of size no greater than $2r+1$ that contains vertices different from the x_j's. Let us denote by A the event that such a path exists. One can check that $Pr_n(A) \leq C/n$ for some fixed C.

Thus, if we denote by B the event that G contains some super-critical subgraph of size no greater than $2r+1$, by the union bound:

$$Pr_n(\neq \phi_r) \leq Pr_n(B) + Pr_n(A) \overset{n \to \infty}{\longrightarrow} 0.$$

$\square$

Theorem 3.8. *(Multivatiate Brun's Sieve) Let $r \in \mathbb{N}$ and for each $i = 1, \ldots, r$ let $\{X_i(n)\}_{n \in \mathbb{N}}$ be a succession of random variables such that for each $n \in \mathbb{N}$, $X_i(n)$ is a sum of random indicator variables (i.e. variables taking only values 0 and 1)*

$Y_{i,1}(n), \ldots, Y_{i,s_n}(n)$. Let $\lambda_1, \ldots, \lambda_r$ be real numbers. If for each r-tuple of natural numbers $a_1, \ldots, a_r$ is satisfied

$$\lim_{n\to\infty} E\left[\prod_{i=1}^{r}\binom{X_i}{a_i}\right] = \prod_{i=1}^{r}\frac{\lambda_i^{a_i}}{a_i!},$$

then the random variable $(X_1, \ldots, X_r)$ converges in distribution to a tuple of independent Poisson variables with means $\lambda_1 \ldots \lambda_r$. That is,

$$\forall x_1, \ldots, x_r \in \mathbb{N}: \quad \lim_{n\to\infty} Pr(\wedge_{i=1}^{r}X_i = x_i) = \prod_{i=1}^{r} Poi_{\lambda_i}(x_i).$$

Remark 3.1. Let X be a random variable sum of indicator variables $Y_1, \ldots, Y_s$. For each $i \in \mathbb{N}$, let X_i be the random variable

$$X_i = |\{(j_1, \ldots, j_l) \mid j_1 < \cdots < j_i, \, Y_{j_1} = \cdots = Y_{j_i} = 1\}|.$$

That is, X_i counts the unordered i-tuples of Y_j's that take value 1. Then it is not difficult to check that

$$\binom{X}{i} = X_i, \text{ for all } i \in \mathbb{N}.$$

Let $x_1, \ldots, x_l$ be vertices of a random graph G. For each $y \in V(G)$ we abbreviate by $T_r(y; x, 1, \ldots, x_l)$ the rooted tree

$$Tree(y, Core(G(c_1[x_1], \ldots, c_l[x_l]); r)).$$

Theorem 3.9. *Let $\sigma(x_1, \ldots, x_l)$ be a consistent open formula, and let $k, r, s \in \mathbb{N}$ $s \leq l$. Then, for any k-morphism classes $C_1, \ldots, C_s$ of trees with radii at most r it is satisfied*

$$\lim_{n\to\infty} Pr_n(\bigwedge_{i=1}^{s} Tree_r(x_i; x_1, \ldots, x_l) \in C_i \mid \sigma(x_1, \ldots, x_l)) = \prod_{i=1}^{s}\lambda_{C_i, r}.$$

Proof. The proof is by induction on r.

For $r = 0$ there is only one class of k-morphic trees and the probability in the statement is always 1 for all n.

Fix $r > 0$ and assume that the statement is true for all lesser values of r.

Let $\mathcal{E}$ be the set of k-types of edges E of radius at most $r - 1$. For each $E \in \mathcal{E}$ pick a representative $(C_{E,1}, \ldots, C_{E,j-1}, r, C_{E,j} \ldots, C_{E,a_E-1})$, and denote by j_E the index of the root in that representative. Denote by a_E and c_E the arity and color of E, and denote by ψ_E the subgroup of the symmetry group of ϕ_{c_E} that fixes the index j_E. Consider, for each $i = 1, \ldots, s$ and each the random variables

$$X_{i,E}(n) = \text{number of initial edges of type } E \text{ in } T_r(x_i; x_1, \ldots, x_l).$$

Given $e = [y_1, \ldots, y_{j_E-1}, x_i, y_{j_E}, \ldots, y_{a_E-1}] \in \mathcal{H}_{c_E}(n)$ we can define the indicator random variable $X_{i,E,e}$ that takes value 1 if the following are all satisfied

- $e \in H_{c_E}$,

- e belongs to $T_r(x_i; x_1, \ldots, x_l)$, and

- the k-type of e is E.

One can check that for fixed i, E

$$X_{i,E}(n) = \sum_{e=[y_1, \ldots, y_{j_E-1}, x_i, y_{j_E}, \ldots, y_{a_E-1}] \in \mathcal{H}_{c_E}(n)} X_{i,E,e}(n).$$

Thus we can apply the multivariate Brun's Sieve to the variables $X_{i,E}$.

Let $(b_{i,E})_{\substack{i=1,\ldots,s \\ E \in \mathcal{E}}}$ be natural numbers. We want to compute

$$\lim_{n \to \infty} E\left[\prod_{i=0}^{s} \prod_{E \in \mathcal{E}} \binom{X_{i,E}(n)}{b_{i,E}} \,\middle|\, \sigma \right].$$

Define by Ω the set

$$\Omega := \{(i, E, b, j) \mid i, b, j \in \mathbb{N}, \, E \in \mathcal{E},$$
$$1 \leq i \leq s, \, b = 1 \leq b \leq b_{i,E}, \, 1 \leq j \leq a_E - 1\}.$$

And let $\hat{\Omega}$ be the projection of Ω onto its first three coordinates. That is,

$$\widehat{\Omega} := \{(i, E, b) \mid i, b \in \mathbb{N}, \, E \in \mathcal{E},$$
$$1 \leq i \leq s, \, b = 1 \leq b \leq b_{i,E}\}.$$

Denote by X be the set $\{x_1, \ldots, x_l\}$. Choose a function $y : \Omega \to [n] \setminus X$. Informally, $y()$ represents a choice of edges in G. We say that $y()$ satisfies the property P if for any fixed $1 \leq i \leq s$, $E \in \mathcal{E}$ and $1 \leq b_1 < b_2 \leq b_{i,E}$, and $t = 1, 2$, the tuples

$$[y(i, E, b_t, 1), \ldots, y(i, E, b_t, j_E - 1), x_i, y(i, E, b_t, j_E), \ldots y(i, E, b_t, a_E - 1)]$$

represent different elements in $\mathcal{H}_{c_E}(n)$. In other words, $y()$ is a choice of different edges.

Define the event $A(y)$ as

$$\bigwedge_{(i,E,b,j) \in \Omega} y(i, E, b, j) \in T_r(x_i; x_1, \ldots, x_l).$$

Define also the event $B(y)$ as

$$\bigwedge_{\omega=(i,E,b) \in \widehat{\Omega}} [y(\omega, 1), \ldots, y(\omega, j_E - 1), x_i, y(\omega, j_E), \ldots, y(\omega, a_E - 1)] \in H_{c_E}.$$

Finally, let $T(y)$ be the event that

$$\bigwedge_{(i,E,b,j) \in \Omega} T_{(r-a_E+1)}(y(i, E, b, j); x_1 \ldots, x_l) \in C_{E,j}.$$

That is,

- $A(y)$ is the event that for any fixed (i, E, b) the vertices $y(i, E, b, j)$ belong to the tree of x_i,

- $B(y)$ is the event that for any fixed (i, E, b) the vertices $y(i, E, b, j)$ together with x_i form an edge in H_{c_E} when ordered in a particular way, and

- $T(y)$ is the event that the tree hanging from each vertex $y(i, E, b, j)$ belongs to the particular k-morphism class given by the edge type E and the position j.

Then it is satisfied

$$
E\left[\prod_{i=0}^{s}\prod_{E\in\mathcal{E}}\binom{X_{i,E}}{b_{i,E}}\,\middle|\,\sigma\right] =
$$

$$
=\prod_{i=1}^{s}\prod_{E\in\mathcal{E}}\left(\frac{1}{|\psi_E|^{b_{i,E}}\cdot b_{i,E}!}\right)\cdot\sum_{\substack{y:\Omega\to[n]\setminus X \\ y \text{ satisfies } P}} Pr(T(y)\wedge A(y)\wedge B(y)\,|\,\sigma). \quad (4)
$$

Notice that $A(y)$ implies that y is injective. Indeed, if a vertex v belongs to two edges incident to some x_i then both edges cannot belong to the tree of x_i because they would form a cycle (or a super-critical graph). Also, if v belongs to the edges e_1, e_2 incident to x_i and x_j respectively then it cannot happen that e_1 is in the tree of x_i and e_2 is in the tree of x_j at the same time, because $e_1 \cup e_2$ would belong to the center of $Core(G(c_i[x_i], c_j[x_j]); r)$. In consequence we only need to take in consideration injective y's in last equation. Also, by the symmetry of the random model, the probability written in that equation is equal for all injective y's. Hence we have

$$
\prod_{i=1}^{s}\prod_{E\in\mathcal{E}}\left(\frac{1}{|\psi_E|^{b_{i,E}}\cdot b_{i,E}!}\right)\cdot\sum_{\substack{y:\Omega\to[n]\setminus X \\ y \text{ injective}}} Pr(T(y)\wedge A(y)\wedge B(y)\,|\,\sigma) =
$$

$$
\prod_{i=1}^{s}\prod_{E\in\mathcal{E}}\left(\frac{1}{|\psi_E|^{b_{i,E}}\cdot b_{i,E}!}\right)\cdot Pr(T(z)\wedge A(z)\wedge B(z)\,|\,\sigma)\cdot\sum_{\substack{y:\Omega\to[n]\setminus X \\ y \text{ injective}}} 1, \quad (5)
$$

where z is an arbitrary injective map from Ω to $[n]\setminus X$.

We can write

$$
Pr(T(z)\wedge A(z)\wedge B(z)\,|\,\sigma) = Pr(T(z)\wedge A(z)\,|\,B(z)\wedge\sigma)\cdot Pr(B(z)\,|\,\sigma)
$$

Notice that $\phi_r(x_1,\ldots,x_l)\wedge B(z)$ implies $A(z)\wedge B(z)$, and in consequence the following chain of inequalities holds

$$
Pr(T(z)\,|\,B(z)\wedge\sigma) \geq Pr(T(z)\wedge A(z)\,|\,B(z)\wedge\sigma) \geq
$$
$$
\geq Pr(T(z)\wedge\phi_r(x_1,\ldots,x_l)\,|\,B(z)\wedge\sigma).
$$

But using lemma REF we get

$$\lim_{n\to\infty} Pr_n(T(z) \wedge \phi_r(x_1,\dots,x_l) \,|\, B(z) \wedge \sigma) = \lim_{n\to\infty} Pr_n(T(z) \,|\, B(z) \wedge \sigma),$$

so

$$\lim_{n\to\infty} Pr_n(T(z) \wedge A(z) \,|\, B(z) \wedge \sigma) = \lim_{n\to\infty} Pr_n(T(z) \,|\, B(z) \wedge \sigma). \qquad (6)$$

Notice that $B(z)$ can be written in terms of a purely relational open formula with free variables the $y(i,E,b,j)$'s. Thus by the induction hypothesis we have

$$\lim_{n\to\infty} Pr_n(T(z) \,|\, B(z) \wedge \sigma) = \Gamma,$$

where

$$\Gamma := \prod_{\substack{1\le i\le s \\ E\in\mathcal{E} \\ 1\le j\le a_E-1}} (\lambda_{C_{E,j}})^{b_{i,E}}. \qquad (7)$$

In particular Γ is different from 0. Hence, last term in eq. (5) is equal to

$$\lim_{n\to\infty} \prod_{i=1}^{s} \prod_{E\in\mathcal{E}} \left(\frac{1}{|\psi_E|^{b_{i,E}} \cdot b_{i,E}!} \right) \cdot \Gamma \cdot Pr_m(B(z)\,|\,\sigma) \cdot \sum_{\substack{y:\Omega\to[n]\setminus X \\ y \text{ injective}}} 1. \qquad (8)$$

Also,

$$\sum_{\substack{y:\Omega\to[n]\setminus X \\ y \text{ injective}}} 1 = |[n]\setminus X| \cdot (|[n]\setminus X|) \cdots (|[n]\setminus X| - |\Omega| + 1),$$

and using that X and Ω are constant in size,

$$\sum_{\substack{y:\Omega\to[n]\setminus X \\ y \text{ injective}}} 1 \simeq n^{|\Omega|}, \qquad (9)$$

where $f(n) \simeq g(n)$ means that $\lim_{n\to\infty} f(n)/g(n) = 1$. Finally, as σ only affects the edges between the x_i's , $B(z)$ and σ are independent. Hence,

$$Pr_n(B(z)\,|\,\sigma) = \prod_{i=1}^{s} \prod_{E\in\mathcal{E}} \left(\frac{\beta_{c_E}}{n^{(a_E-1)}} \right)^{b_{i,E}},$$

and using that

$$n^{|\Omega|} = \prod_{i=1}^{s} \prod_{E\in\mathcal{E}} (n^{(a_E-1)})^{b_{i,E}}$$

and eq. (9) we obtain

$$Pr_n(B(z)\,|\,\sigma) \cdot \sum_{\substack{y:\Omega\to[n]\setminus X \\ y \text{ injective}}} 1 \simeq \prod_{E\in\mathcal{E}} \beta_{c_E}^{b_{i,E}}. \qquad (10)$$

In consequence, using eq. (10) and eq. (7) in eq. (8) we get

$$\lim_{n\to\infty} E\left[\prod_{i=0}^{s}\prod_{E\in\mathcal{E}}\binom{X_{i,E}}{b_{i,E}}\,\bigg|\,\sigma\right] = \prod_{i=0}^{s}\prod_{E\in\mathcal{E}}\left[\left(\frac{\beta_{c_E}\prod_{j=1}^{a_E-1}\lambda_{C_{E,j}}}{|\psi_E|}\right)^{b_{i,E}}\frac{1}{b_{i,E}!}\right].$$

And using the multivariate Brun's Sieve we get that for each choice of natural numbers $\{b_{i,E}\}_{\substack{1\le i\le s \\ E\in\bar{\mathcal{E}}}}$ it is satisfied

$$\lim_{n\to\infty} Pr_n\left(\bigwedge_{i=0}^{s}\bigwedge_{E\in\mathcal{E}} X_{i,E}=b_{i,E}\,\bigg|\,\sigma\right) = \prod_{i=0}^{s}\prod_{E\in\mathcal{E}} Poi_{\mu_E}(b_{i,E}),$$

where

$$\mu_E = \frac{\beta_{c_E}\prod_{j=1}^{a_E-1}\lambda_{C_{(E,j)},(r-a_E+1)}}{|\psi_E|}.$$

The k-morphism class of $T_r(x_j;x_1,\ldots,x_l)$ depends exclusively on the number, up to $k+1$, of its initial edges of each type. More explicitly

$$T_r(x_j;x_1,\ldots,x_l) \in C \iff$$
$$\bigwedge_{E\in\mathcal{E}}\left(X_{j,E}=(C,E) \text{ if } (C,E)\le k\right),\ \text{or}\ \left(X_{j,E}\ge k+1 \text{ otherwise}\right).$$

In consequence

$$\lim_{n\to\infty} Pr_n(\bigwedge_{i=1}^{s} Tree_r(x_i;x_1,\ldots,x_l)\in C_i\,|\,\sigma) =$$

$$= \prod_{i=0}^{s}\left[\left(\prod_{\substack{E\in\mathcal{E} \\ (C_i,E)<k+1}} Poi_{\mu_E}((C_i,E))\right)\left(\prod_{\substack{E\in\mathcal{E} \\ (C_i,E)\ge k+1}} Poi_{\mu_E}(\ge(k+1))\right)\right] =$$

$$= \prod_{i=0}^{s}\lambda_{C_i,r},$$

as we wanted.

$\square$

Theorem 3.10. *Let O be a simple k-agreeability class. Then it is satisfied*

$$\lim_{n\to\infty} Pr_n(Core(G;r)\in O) = \theta,$$

for some $\theta\in\Theta$.

Proof. This is an easier version of last theorem.

Let $\mathcal{U}$ be the set of all k-morphism classes of unicycles with radius at most r. For each class $U\in\mathcal{U}$ choose a representative $rep(U)\in U$, and let $cycle(U)$ be

the cycle in $rep(U)$, whose number of vertex will be denoted by n_U. Choose an ordering $x_{U,1}, \ldots, x_{U,n_U}$ of the vertices in $cycle(U)$ and for each $1 \le i \le n_U$ denote by $C_{U,i}$ the k-morphism class of $Tree(x_i, rep(U))$. One can consider $cycle(U)$ to be a vertex colored graph where the color assigned to each vertex x_i is $C_{U,i}$. Isomorphisms of the colored cycle $cycle(U)$ induce permutations of $[n_U]$ via this ordering. Let us denote by ψ_U that group of permutations.

For any $U \in \mathcal{U}$ we define the random variable

$$X_U(n) = \text{number of connected components of } Core(G;r) \text{ in } U,$$

and for each element $g \in V(G)^{n_U}/\psi_U$ we define the indicator variable $X_{U,g}(n)$ that equals 1 if

- $g = [x_1, \ldots, x_{n_U}]$, for some vertices $x_1, \ldots, x_{n_U}$ such that the map $f : G[X] \to cycle(U)$,where $X = \{x_1, \ldots x_{n_U}\}$, defined by $x_i \mapsto x_{U,i}$ is an isomorphism.

- $N(X;r)$ is a connected component of $Core(G;r)$.

- $N(X;r) \overset{k}{\simeq} rep(U)$ via f. In particular this means that $Tree(x_i, G)$ belongs to $C_{U,i}$.

In other words, $X_{U,g}$ indicates if there is a graph in the k-morphism class U embedded in G in a particular way represented by g. One can check that for all $U \in \mathcal{U}$

$$X_U(n) = \sum_{g \in [n]/\psi_U} X_{U,g}(n),$$

so we can apply the multivariate Brun's Sieve to the X_U's.

Let $(b_U)_{U \in \mathcal{U}}$ be fixed natural numbers. We are interested in obtaining

$$\lim_{n \to \infty} E\left[\prod_{U \in \mathcal{U}} \binom{X_U(n)}{b_U} \right].$$

Let Ω be the set defined as

$$\Omega := \{ (U, b, i) \,|\, U \in \mathcal{U}, \, b, i \in \mathbb{N}, \, 1 \le b \le b_U, \, 1 \le i \le n_U \}$$

and let $\widehat{\Omega}$ be the projection of Ω onto its two first coordinates. That is,

$$\widehat{\Omega} := \{ (U, b) \,|\, U \in \mathcal{U}, \, b \in \mathbb{N}, \, 1 \le b \le b_U \}.$$

Let $y : \Omega \to [n]$ be a map. Informally, $y()$ represents a choice of embeddings of graphs in with the appropriate k-morphism classes. We will say that $y()$ satisfies the property P if for any fixed $U \in \mathcal{U}$ and $1 \le b_1 < b_2 \le b_U$ the tuples

$$[y(U, b_1, 1), \ldots, y(U, b_1, n_U)], \text{ and } [y(U, b_2, 1), \ldots, y(U, b_2, n_U)]$$

represent different elements in $[n]^{n_U}/\psi_U$. That is, $y()$ is a choice of different embeddings.

Define for any $(U, b) \in \widehat{\Omega}$ the set $Y(U, b) = \{\, y(U, b, i) \mid 1 \leq i \leq n_U \,\}$.

We define the following events for a given $y : \Omega \to [n]$.

- Let $A(y)$ be the event that for each $(U, b) \in \widehat{\Omega}$, the map $f_{U,b} : cycle(U) \to G[Y(U, b)]$ given by $x_{U,i} \to y(U, b, i)$ is an embedding.

- Let $B(y)$ be the event that $Center(N(Y(U, b); r)$ is the image of $f_{U,b}$, for each $(U, b) \in \widehat{\Omega}$.

- Let $T(y)$ be the event that

$$\bigwedge_{(U,b,i) \in \Omega} T_r(y(U, b, i); Y) \in C_{U,i},$$

where Y denotes set of vertices in the image of y.

Then,

$$E\left[\prod_{U \in \mathcal{U}} \binom{X_U(n)}{b_U}\right] = \prod_{U \in \mathcal{U}} \frac{1}{|\psi_U|^{b_U} b_U!} \cdot \sum_{\substack{y:\Omega \to [n] \\ y \text{ satisfies } P}} Pr(A(y) \wedge B(y) \wedge T(y)).$$

Property P, together with events $A(y)$ and $B(y)$ imply that y is injective, so we can consider only such y's in last equation. Again, by the symmetry of the random model the probability appearing there is the same for all injective y's. Hence,

$$\prod_{U \in \mathcal{U}} \frac{1}{|\psi_U|^{b_U} b_U!} \cdot \sum_{\substack{y:\Omega \to [n] \\ y \text{ satisfies } P}} Pr(A(y) \wedge B(y) \wedge T(y)) =$$

$$= \prod_{U \in \mathcal{U}} \frac{1}{|\psi_U|^{b_U} b_U!} \cdot Pr(A(z) \wedge B(z) \wedge T(z)) \cdot \sum_{\substack{y:\Omega \to [n] \\ y \text{ injective}}} 1, \tag{11}$$

where z is an arbitrary injective map $z : \Omega \to [n]$.

We can write

$$Pr(A(z) \wedge B(z) \wedge T(z)) = Pr(B(z) \wedge T(z) \mid A(z)) \cdot Pr(A(z)).$$

Let τ_r be the event that G is r-simple. One can check that $A(y) \wedge \tau_r$ implies $A(y) \wedge B(y)$. In consequence the following chain of inequalities holds

$$Pr(T(z) \mid B(z)) \geq Pr(T(z) \wedge A(z) \mid B(z)) \geq Pr(T(z) \wedge \tau_r \mid B(z)).$$

Notice that $B(z)$ can be expressed as a purely relational open formula with free variables the elements indexed by z, because it only depends on the edges between vertices in the image of z. Using the previous theorem and LEMMA we obtain

$$\lim_{n \to \infty} Pr_n(T(z) \wedge \tau_r \,|\, B(z)) = \lim_{n \to \infty} Pr_n(T(z)\,|\,|\,B(z)) = \Gamma,$$

where

$$\Gamma := \prod_{\substack{U \in \mathcal{U} \\ 1 \leq i \leq n_U}} (\lambda_{C_{(U,i)},r})^{b_U}.$$

Because the probability of each edge is independent, one obtains

$$Pr_n(B(z)) = \prod_{(U,b) \in \widehat{\Omega}} \frac{\prod_{i=1}^{c} \beta_i^{|H_i(cycle(U))|}}{n^{n_U}}.$$

Also,

$$\sum_{\substack{y:\Omega \to [n] \\ y \text{ injective}}} 1 \simeq \prod_{(U,b) \in \widehat{\Omega}} n^{n_U}.$$

This way substituting REF, REF and REF in EQ we get

$$\lim_{n \to \infty} E\left[\prod_{U \in \mathcal{U}} \binom{X_U(n)}{b_U}\right] = \prod_{U \in \mathcal{U}} \left(\frac{\prod_{i=1}^{c} \beta_i^{|H_i(cycle(U))|} \prod_{i=1}^{n_U} \lambda_{U,i}}{|\psi_U|}\right)^{b_U} \cdot \frac{1}{b_U!}.$$

Applying the multivatiate Brun's Sieve we obtain that for any fixed natural numbers $(b_U)_{U \in \mathcal{U}}$

$$\lim_{n \to \infty} Pr_n\left(\bigwedge_{U \in \mathcal{U}} X_U = b_U\right) = \prod_{U \in \mathcal{U}} Poi_{\xi_U}(b_i),$$

where

$$\xi_U = \frac{\prod_{i=1}^{c} \beta_i^{|H_i(cycle(U))|} \prod_{i=1}^{n_U} \lambda_{U,i}}{|\psi_U|}.$$

The class of k-agreeability of a graph depends only on the number of connected components of each k-morphism class. More explicitly, if O is a k-agreeability class of radius r,

$$Core(G;r) \in O \iff \bigwedge_{U \in \mathcal{U}} (X_U = (O,U) \text{ if } (O,U) \leq k), \text{ or } (X_U \geq k+1 \text{ otherwise}).$$

In consequence,

$$\lim_{n \to \infty} Pr_n(Core(G;r) \in O) = \left(\prod_{\substack{U \in \mathcal{U} \\ (O,U) \leq k}} Poi_{\xi_U}((O,U))\right)\left(\prod_{\substack{U \in \mathcal{U} \\ (O,U) \geq k+1}} Poi_{\xi_U}(\geq (k+1))\right),$$

as we wanted. $\qquad\square$

Bibliography

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