

RESEARCH ARTICLE

WILEY

Limiting probabilities of first order properties of random sparse graphs and hypergraphs

Alberto Larrauri¹ | Tobias Müller² | Marc Noy³

¹Department of Computer Science, Universitat Politècnica de Catalunya, Barcelona, Spain

²Department of Mathematics, University of Groningen, Groningen, Netherlands

³Department of Mathematics, Universitat Politècnica de Catalunya, Barcelona, Spain

Correspondence

Marc Noy, Department of Mathematics, Universitat Politècnica de Catalunya, Barcelona, Spain.

Email: marc.noy@upc.edu

Funding information

the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme, Grant/Award Number: ERC-2014-CoG 648276 AUTAR; the Ministerio de Economía y Competitividad, Grant/Award Number: MTM2017-82166-P.

Abstract

Let G_n be the binomial random graph $G(n, p = c/n)$ in the sparse regime, which as is well-known undergoes a phase transition at $c = 1$. Lynch (*Random Structures and Algorithms*, 1992) showed that for every first order sentence ϕ , the limiting probability that G_n satisfies ϕ as $n \rightarrow \infty$ exists, and moreover it is an analytic function of c . In this paper we consider the closure $\overline{L_c}$ in the interval $[0, 1]$ of the set L_c of all limiting probabilities of first order sentences in G_n . We show that there exists a critical value $c_0 \approx 0.93$ such that $\overline{L_c} = [0, 1]$ when $c \geq c_0$, whereas $\overline{L_c}$ misses at least one subinterval when $c < c_0$. We extend these results to random sparse d -uniform hypergraphs, where the probability of a d -edge is $p = c/n^{d-1}$.

KEYWORDS

first order logic, logical limit laws, sparse random graphs and hypergraphs

1 | INTRODUCTION

We consider properties of random graphs expressible in the first order (FO) language of graphs, which is ordinary first order logic enriched with an adjacency relation $E(x, y)$ assumed to be symmetric and antireflexive. Our model is the binomial random graph $G(n, p)$ with vertex set $\{1, \dots, n\}$ and in which every edge is present independently with probability p . We focus on the so-called sparse regime $p = c/n$ with $c > 0$. It is well-known that $G(n, c/n)$ undergoes a phase transition at $c = 1$, corresponding to the emergence of the giant component [2]. The model studied in [2] was the uniform model $G(n, M)$ on graphs with n vertices and M edges. However, the results we need in this work can be translated into the $G(n, p)$ model with the relation $M = p \binom{n}{2}$.

Our starting point is the following result by Lynch [9]. The notation $G \models \phi$ means that the graph G satisfies the sentence ϕ . We recall that a sentence is a FO formula without free variables, thus expressing a graph property closed under isomorphism.

Theorem. *For each FO sentence ϕ , the following limit exists:*

$$p_c(\phi) = \lim_{n \rightarrow \infty} \mathbf{P}\{G_n \models \phi\}.$$

Moreover, $p_c(\phi)$ is an expression built using c , rational constants, sums, products and exponentiation with base e , hence it is an analytic function of c .

The previous result shows in a strong form that FO logic does not capture the phase transition (see also [13] for a discussion including monadic second order logic).

Instead of considering limiting probabilities of individual sentences, in this paper we consider the set of all limiting probabilities

$$L_c = \{p_c(\phi) : \phi \text{ FO sentence}\},$$

and its topological closure $\overline{L_c}$ in $[0, 1]$. Our main result is that there is a transition in the structure of $\overline{L_c}$ at a particular value of c . We say that $\overline{L_c}$ contains a *gap* if there is at least one subinterval $(a, b) \subseteq [0, 1]$ with $a < b$ such that $\overline{L_c} \cap (a, b) = \emptyset$.

Theorem 1. *Let $\overline{L_c}$ be the closure of the of limiting probabilities of first order sentences in $G(n, c/n)$. Let $c_0 \approx 0.93$ be the unique positive solution of*

$$e^{\frac{c}{2} + \frac{c^2}{4}} \sqrt{1 - c} = \frac{1}{2}. \tag{1}$$

Then for every $c > 0$ the set $\overline{L_c}$ is a finite union of closed intervals. Moreover, the following holds:

1. $\overline{L_c} = [0, 1]$ for $c \geq c_0$.
2. $\overline{L_c}$ has at least one gap for $0 < c < c_0$.

Remark. This line of research was considered in [4] for minor-closed classes of graphs under the uniform distribution. For instance, it was shown there that for the class of acyclic graphs (forests), the set $\overline{L_c}$ is the union of 4 disjoint intervals. It was also shown that for every proper minor-closed class of graphs whose forbidden minors are all 2-connected, $\overline{L_c}$ is always a finite union of at least two intervals.

We extend the previous result to random sparse hypergraphs. We consider the model $G^d(n, p)$ of random d -uniform hypergraphs, where every d -edge has probability p of being in $G^d(n, p)$ independently. When $p = c/n^{d-1}$ the expected number of edges $p \binom{n}{d}$ is linear in n , justifying the qualifier “sparse.” A phase transition also occurs in $G^d(n, c/n^{d-1})$ when $c = (d - 2)!$, as shown in [12].

Theorem 2. *Let $d \geq 3$ be fixed and let $\overline{L_c}$ be the closure of the of limiting probabilities of first order sentences in $G^d(n, c/n^{d-1})$. Let c_0 be the unique positive solution of*

$$\exp\left(\frac{c}{2(d-2)!}\right) \sqrt{1 - \frac{c}{2(d-2)!}} = \frac{1}{2}. \tag{2}$$

Then for every $c > 0$ the set \overline{L}_c is a finite union of intervals. Moreover, the following holds:

1. $\overline{L}_c = [0, 1]$ for $c \geq c_0$.
2. \overline{L}_c has at least one gap for $0 < c < c_0$.

We remark that $c_0 = r(d-2)!$, where $r \approx 0.898$ is the positive solution of $\exp(r/2)\sqrt{1-r} = 1/2$. As we will see one of the differences between Equations (1) and (2) comes from the fact that in graphs we consider cycles of length at least 3, whereas in hypergraphs we have to consider cycles of length at least 2.

Here is a summary of the paper. In Section 2 we review several preliminaries we need on probability, random graphs and logic. In Section 3 we prove Theorem 1, and in Section 4 we prove Theorem 2. To avoid repetition the preliminary results needed for hypergraphs in Section 4 are only sketched.

2 | PRELIMINARIES

We start with Brun’s sieve for obtaining limiting Poisson distributions (see [1, Theorem 1.23]).

Lemma 1. *Let $X_{n,1}, \dots, X_{n,k}$ be nonnegative integer valued random variables defined over the same probability space. Let $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ be nonnegative. Suppose that for all $a_1, \dots, a_k \geq 0$ it holds that*

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\prod_{i=1}^k \binom{X_{n,i}}{a_i} \right] = \prod_{i=1}^k \frac{\lambda_i^{a_i}}{a_i!}.$$

Then the $X_{n,i}$ converge in distribution to independent Poisson variables whose respective means are the λ_i . That is, for any $b_1, \dots, b_k \geq 0$

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\bigwedge_{i=1}^k X_{n,i} = b_i \right) = \prod_{i=1}^k e^{-\lambda_i} \frac{\lambda_i^{b_i}}{b_i!}.$$

Next we present several results on the number of cycles in random sparse graphs. Let $X_{n,k}$ be the number of k -cycles in $G(n, c/n)$. It is easy to show that

$$\mathbb{E}[X_{n,k}] \leq \frac{c^k}{2k},$$

and

$$\lim_{n \rightarrow \infty} \mathbb{E}[X_{n,k}] = \frac{c^k}{2k}.$$

The first part of the next lemma appears already in [2] for the $G(n, M)$ model. The second part is easily proved using the method of moments [1, Theorem 1.23].

Lemma 2. *For fixed $k \geq 3$, the number of k -cycles $X_{n,k}$ in $G(n, c/n)$ is distributed asymptotically as $n \rightarrow \infty$ as a Poisson law with parameter $\lambda_k = \frac{c^k}{2k}$. Moreover, for fixed k the random variables $X_{n,3}, \dots, X_{n,k}$ are asymptotically independent.*

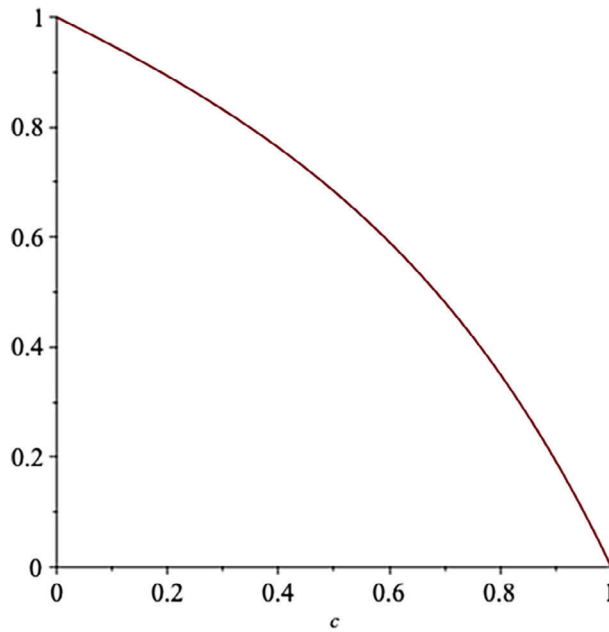


FIGURE 1 The probability that $G(n, c/n)$ has no cycles as a function of c

We set

$$f(c) = \frac{1}{2} \ln \frac{1}{1-c} - \frac{c}{2} - \frac{c^2}{4}. \tag{3}$$

This is a function defined on $(0, 1)$ that plays an important role in our results. The function $e^{-f(c)}$ is the limiting probability that $G(n, c/n)$ is acyclic; see Figure 1 for a plot of it.

Corollary 1. *When $c < 1$ the expected number of cycles in $G(n, c/n)$ is $f(c)$. Moreover, the limiting probability as $n \rightarrow \infty$ that $G(n, c/n)$ contains no cycle is*

$$e^{-f(c)} = e^{\frac{c}{2} + \frac{c^2}{4}} \sqrt{1-c}.$$

Proof. We have

$$\lim_{n \rightarrow \infty} \mathbf{E} \left[\sum_{k \geq 3} X_{n,k} \right] = \sum_{k \geq 3} \frac{c^k}{2k} = f(c).$$

The second statement follows from Lemma 2. ■

A graph is unicyclic, or a unicycle, if it is connected and has a unique cycle. A unicyclic graph has a simple structure: it consists of a cycle of length at least 3 and a collection of rooted trees attached to the vertices of the unique cycle. The *size* of a unicyclic graph is the number of edges, which is equal to the number of vertices (we use this convention because for hypergraphs it is more convenient to define the size of a unicycle as the number of edges). We denote by \mathcal{U} the family of unlabeled graphs whose connected components are all unicyclic. We let $\mathcal{U}_n = \{H \in \mathcal{U} : |H| = n\}$ and $\mathcal{U}_{\leq n} = \bigcup_{i=1}^n \mathcal{U}_i$.

The following is well-known; see [3, Lemma 2.10] for a proof. We say that a property holds asymptotically almost surely (a.a.s.) if it holds with probability tending to 1 as $n \rightarrow \infty$. The notation $a_n \sim b_n$ means that $\lim_{n \rightarrow \infty} a_n/b_n = 1$.

Lemma 3. *Let $p(n) \sim c/n$ with $0 < c < 1$. Then a.a.s all the connected component of $G(n, p)$ are either trees or unicycles.*

Given a graph G , we define its *fragment* $\text{Frag}(G)$ as the union of all the unicyclic components in G . We will write Frag_n to denote the fragment of $G(n, p)$. The following result states that below the critical value $c = 1$ the expected size of Frag_n is asymptotically bounded.

Lemma 4. *Let $p(n) \sim c/n$ with $0 < c < 1$. Then $\lim_{n \rightarrow \infty} E[|\text{Frag}_n|]$ exists and is a finite quantity.*

Proof. This is done in [2, Theorem 5d] for the uniform model and in greater detail in [3, Lemma 2.11] for the binomial model. For future reference we sketch the main ingredients in the proof.

Let $Y_{n,i}$ be the random variable equal to the number of unicyclic components in G_n that contain exactly i edges. Then one proves that for k large enough and $n \geq 0$

$$E[Y_{n,k}] \leq (ce^{1-c})^k e^{c/2}.$$

Furthermore, for all $k \geq 3$

$$\lim_{n \rightarrow \infty} E[Y_{n,k}] = C(k, k)(ce^{-c})^k,$$

where $C(k, k)$ denotes the number of labeled unicyclic graphs on k vertices. Then the statement follows from the Dominated Convergence Theorem. ■

Let ϕ be a first order sentence. We recall that the quantifier rank $\text{qr}(\phi)$ is the maximum number of nested quantifiers in ϕ . It is shown in [9] that, a.a.s., whether G_n satisfies ϕ or not depends only on the induced unicycles in G_n of diameter at most 3^k , where $k = \text{qr}(\phi)$ (see Theorems 4.7–4.9 in [9]). This, together with Lemma 3), implies the following:

Lemma 5. *Let $p(n) \sim c/n$ with $0 < c < 1$. Let ϕ be a FO sentence and let $H \in \mathcal{U}$. Then*

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(G(n, p) \models \phi \mid \text{Frag}_n \simeq H \right) = 0 \text{ or } 1.$$

Moreover, the value of the limit depends only on ϕ and H , and not on c .

Because of Lemma 3, when $0 < c < 1$ a.a.s. all cycles in G_n are contained in unicyclic components. Since the expected number of cycles in G_n is asymptotically bounded (Corollary 1) we obtain the following.

Corollary 2. *Let $p(n) \sim c/n$ with $0 < c < 1$, and let Z_n be the random variable equal to the number of cycles in $G(n, p)$ that belong to connected components that are not trees or unicycles. Then*

$$\lim_{n \rightarrow \infty} E[Z_n] = 0.$$

Let $\text{aut}(H)$ denote the number of automorphisms of a graph H .

Lemma 6. Let $p(n) \sim c/n$ with $c > 0$. Let T be a finite set of unlabeled unicycles. For each $H \in T$ let $X_{n,H}$ be the random variable equal to the number of connected components in $G(n, p)$ isomorphic to H , and let $\lambda_H = \frac{(e^{-c})^{|H|}}{\text{aut}(H)}$. Then

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\bigwedge_{H \in T} X_{n,H} = a_H \right) = \prod_{H \in T} e^{-\lambda_H} \frac{\lambda_H^{a_H}}{a_H!}.$$

In other words, the $X_{n,H}$ converge in distribution to independent Poisson variables with respective means λ_H .

Proof. The proof is a slight modification of Theorem 4.8 in [1]. It follows from a straightforward application of Lemma 1. ■

We also need a classical result conjectured by Kakeya [6] and later proven in [10] on the set of subsums of a convergent series of nonnegative terms.

Lemma 7. Let $\sum_{n \geq 0} p_n$ be a convergent series of nonnegative real numbers. Then the following are equivalent:

- (1) $p_i \leq \sum_{j>i} p_j$ for all $i \geq 0$.
- (2)

$$\left\{ \sum_{i \in A} p_i : A \subset \mathbb{N} \right\} = \left[0, \sum_{n=0}^{\infty} p_n \right].$$

Moreover, if the condition $p_i \leq \sum_{j>i} p_j$ holds for all values of i large enough, then the set $\left\{ \sum_{i \in A} p_i : A \subset \mathbb{N} \right\}$ is a finite union of intervals.

3 | PROOF OF THEOREM 1

For the sake of exposition we outline here the main components in the proof:

- For $c \geq 1$ we can approximate any $p \in [0, 1]$ with probabilities of statements of the form “there are at most l cycles of length at most k in $G(n, c/n)$ ” (Section 3.1).
- We prove that the constant c_0 given by Equation (1) satisfies that \overline{L}_c contains at least one gap whenever $c < c_0$ (Section 3.2).
- For $c < 1$ we show that

$$\overline{L}_c = \left\{ \sum_{H \in T} p_H(c) : T \subseteq \mathcal{U} \right\},$$

where $p_H(c) = \lim_{n \rightarrow \infty} \mathbf{P}(\text{Frag}_n \simeq H)$ is given by Equation (5). In other words, \overline{L}_c coincides with the set of subsums of the convergent series $\sum_{H \in \mathcal{U}} p_H(c) = 1$.

- Using Kakeya’s Criterion (Lemma 7) we show that the set \overline{L}_c of subsums of $\sum_{H \in \mathcal{U}} p_H(c)$ is always a finite union of intervals (Section 3.4).
- Using Kakeya Criterion (Lemma 7) once again we prove that \overline{L}_c is the whole interval $[0, 1]$ for $c_0 \leq c < 1$ (Section 3.5).

3.1 | No gap when $c \geq 1$

Let X_k be as before the number of cycles of length k in $G(n, c/n)$, which is asymptotically $Po(c^k/(2k))$. Moreover, for fixed k , the random variables X_3, \dots, X_k are asymptotically independent by Lemma 2. Hence for fixed k ,

$$X_{\leq k} = X_3 + \dots + X_k \xrightarrow[n \rightarrow \infty]{d} Po\left(\sum_{i=3}^k \frac{c^i}{2i}\right).$$

Since $c \geq 1$ the mean $\sum_{i=3}^k c^i/2i$ is not bounded as k grows to infinity so we can pick k such that this mean is as large as we like. Note that for any k and a the property that $X_{\leq k} \leq a$ can be expressed in FO logic. By the central limit theorem we have, for any fixed $x \in \mathbb{R}$

$$\mathbf{P}(Po(\mu) \leq \mu + x\sqrt{\mu}) \xrightarrow{\mu \rightarrow \infty} \Phi(x),$$

where $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$ is the c.d.f. of the standard normal law.

For $0 < p < 1$ and $\epsilon > 0$ we can find x such that $\Phi(x) = p$, a value μ_0 such that $\mathbf{P}(Po(\mu) \leq \mu + x\sqrt{\mu}) \in (p - \epsilon, p + \epsilon)$ for all $\mu \geq \mu_0$, and then finally a k such that $\sum_{i=3}^k \frac{c^i}{2i} \geq \mu_0$. Hence there exists a FO property ϕ with limiting probability within ϵ of p .

3.2 | At least one gap when $c < c_0$

Let $e^{-f(c)}$ be as in Corollary 1 the limiting probability that $G_n = G(n, c/n)$ is acyclic. By elementary calculus $e^{-f(c)}$ is strictly decreasing for $c \in [0, 1]$ (see Figure 1), and by definition of c_0 we have $e^{-f(c_0)} = 1/2$.

Fix $c < c_0$, and let $p(n) \sim c/n$. We are going to show that $\overline{L_c}$ has a gap around $1/2$. Let A_n be the event that G_n is acyclic and let ϕ be a FO sentence and let $p(\phi) = \lim_{n \rightarrow \infty} \mathbf{P}(G_n \models \phi)$. Then

$$\begin{aligned} p(\phi) &= \lim_{n \rightarrow \infty} \mathbf{P}\left(G_n \models \phi \mid A_n\right) \mathbf{P}(A_n) + \mathbf{P}\left(G_n \not\models \phi \mid \neg A_n\right) \mathbf{P}(\neg A_n) \\ &= \lim_{n \rightarrow \infty} \mathbf{P}\left(G_n \models \phi \mid A_n\right) e^{-f(c)} + \mathbf{P}\left(G_n \not\models \phi \mid A_n\right) (1 - e^{-f(c)}). \end{aligned} \tag{4}$$

By Lemma 5 we have

$$\lim_{n \rightarrow \infty} \mathbf{P}\left(G_n \not\models \phi \mid A_n\right) = 0 \text{ or } 1.$$

If the last limit equals 0 then from Equation (4) we obtain that $p(\phi) \leq 1 - e^{-f(c)}$. Otherwise, if the limit is 1, we have $p(\phi) \geq e^{-f(c)}$. Since $e^{-f(c)}$ is strictly decreasing, $c < c_0$ and $e^{f(c_0)} = 1/2$ it follows that $e^{-f(c)} > 1/2$. As a consequence $1 - e^{-f(c)} < 1/2 < e^{-f(c)}$ and $(1 - e^{-f(c)}, e^{-f(c)})$ is a gap of $\overline{L_c}$.

3.3 | Asymptotic distribution of the fragment for $c < 1$ and its consequences

We compute below the asymptotic probability that the fragment Frag_n is isomorphic to a given union H of unicycles.

Lemma 8. *Let $p(n) \sim c/n$ with $0 < c < 1$, and let $H \in \mathcal{U}$. Then*

$$\lim_{n \rightarrow \infty} \mathbf{P}\left(\text{Frag}_n \simeq H\right) = e^{-f(c)} \frac{(e^{-c}c)^{|H|}}{\text{aut}(H)}. \tag{5}$$

Proof. Fix such an H . Let $U_1, U_2, \dots, U_i, \dots$ be an enumeration of all unlabeled unicycles ordered by nondecreasing size. For each i let a_i be the number of connected components of H that are copies of U_i , and let $W_{n,i}$ be the random variable equal to the number of connected components in G_n that are isomorphic to U_i . Clearly $\text{Frag}_n \simeq H$ if and only if $W_{n,i} = a_i$ for all i . Thus,

$$\lim_{n \rightarrow \infty} \mathbf{P}(\text{Frag}_n \simeq H) = \lim_{n \rightarrow \infty} \mathbf{P}\left(\bigwedge_{i=1}^{\infty} W_{n,i} = a_i\right).$$

First, we are going to show that

$$\lim_{n \rightarrow \infty} \mathbf{P}\left(\bigwedge_{i=1}^{\infty} W_{n,i} = a_i\right) = \lim_{j \rightarrow \infty} \lim_{n \rightarrow \infty} \mathbf{P}\left(\bigwedge_{i=1}^j W_{n,i} = a_i\right). \tag{6}$$

Fix $\epsilon > 0$. For each k let $X_{n,k}$ be the random variable that counts the unicyclic connected components of G_n with exactly k edges. By Lemma 4 we have that for some k_0

$$\lim_{n \rightarrow \infty} \sum_{l=k_0}^{\infty} \mathbf{E}[X_{n,l}] \leq \epsilon.$$

Let k_1 be the maximum number of edges in a connected component of H , and let $k = \max(k_0, k_1 + 1)$. Finally, fix j_0 such that $e(U_j) > k_1$ for any $j \geq j_0$. Let $j \geq j_0$. Then $\text{Frag}_n \simeq H$ if and only if

$$\bigwedge_{i=1}^j W_{n,i} = a_i, \quad \text{and} \quad \sum_{\ell=k}^{\infty} X_{n,\ell} = 0.$$

Using Markov's inequality we get

$$\lim_{n \rightarrow \infty} \mathbf{P}\left(\sum_{\ell=k}^{\infty} X_{n,\ell} \geq 1\right) \leq \epsilon.$$

And in consequence,

$$\left| \lim_{n \rightarrow \infty} \mathbf{P}\left(\bigwedge_{i=1}^{\infty} W_{n,i} = a_i\right) - \lim_{j \rightarrow \infty} \lim_{n \rightarrow \infty} \mathbf{P}\left(\bigwedge_{i=1}^j W_{n,i} = a_i\right) \right| \leq \epsilon,$$

proving Equation (6).

Using Lemma 6 we get

$$\lim_{j \rightarrow \infty} \lim_{n \rightarrow \infty} \mathbf{P}\left(\bigwedge_{i=1}^j W_{n,i} = a_i\right) = \prod_{i=1}^{\infty} e^{-\lambda_i} \frac{\lambda_i^{a_i}}{a_i!},$$

where $\lambda_i = \frac{(ce^{-c})^{U_i}}{\text{aut}(U_i)} = \lim_{n \rightarrow \infty} \mathbf{E}[W_{n,i}]$. Using Corollary 2 together with the Dominated Convergence Theorem as in the proof of Lemma 4 we obtain

$$\sum_{i=1}^{\infty} \lambda_i = f(c),$$

and as a consequence

$$\prod_{i=1}^{\infty} e^{-\lambda_i} = e^{-f(c)}.$$

Since $\sum_{i=1}^{\infty} |U_i| a_i = |H|$ and $\prod_{i=1}^{\infty} \text{aut}(U_i)^{a_i} a_i! = \text{aut}(H)$, we finally get

$$\prod_{i=1}^{\infty} \frac{\lambda_i^{a_i}}{a_i!} = \frac{(ce^{-c})^{|H|}}{\text{aut}(H)},$$

and the result follows. ■

Given $H \in \mathcal{U}$ we define $p_H = p_H(c) = \lim_{n \rightarrow \infty} \mathbf{P}(\text{Frag}_n \simeq H)$. The following is a direct consequence of the fact that the expected size of Frag_n is bounded.

Lemma 9. *Let $p(n) \sim c/n$ with $0 < c < 1$, and let $\mathcal{T} \subset \mathcal{U}$. Then*

$$\lim_{n \rightarrow \infty} \mathbf{P}\left(\bigvee_{H \in \mathcal{T}} \text{Frag}_n \simeq H\right) = \sum_{H \in \mathcal{T}} p_H = 1.$$

Proof. If \mathcal{T} is finite then the statement is clearly true, since the events $\text{Frag}_n \simeq H$ are disjoint for different H . Suppose otherwise. Let H_1, \dots, H_i, \dots be an enumeration of \mathcal{T} by nondecreasing size. Fix $\epsilon > 0$. Let $m = \lim_{n \rightarrow \infty} \mathbf{E}[|\text{Frag}_n|]$, and let $M = m/\epsilon$. Then there exists j_0 such that $\mathbf{E}(H_j) \geq M$ for all $j \geq j_0$. Using Markov's inequality we obtain that for any $j \geq j_0$

$$\begin{aligned} \lim_{n \rightarrow \infty} \left| \mathbf{P}\left(\bigvee_{H \in \mathcal{T}} \text{Frag}_n \simeq H\right) - \sum_{i=1}^j p_{H_i} \right| &= \lim_{n \rightarrow \infty} \left| \mathbf{P}\left(\bigvee_{H \in \mathcal{T}} \text{Frag}_n \simeq H\right) - \mathbf{P}\left(\sum_{i=1}^j \text{Frag}_n \simeq H\right) \right| \\ &\leq \lim_{n \rightarrow \infty} \mathbf{P}(|\text{Frag}_n| > M) \leq \epsilon. \end{aligned}$$

As our choice of ϵ was arbitrary this proves the statement. ■

We define S_c as the set of subsums of $\sum_{H \in \mathcal{U}} p_H(c)$,

$$S_c = \left\{ \sum_{H \in \mathcal{T}} p_H(c) : \mathcal{T} \subseteq \mathcal{U} \right\}.$$

Theorem 3. *Let $0 < c < 1$. Then $\overline{L_c} = S_c$.*

Proof. We first show that $\overline{L_c} \subseteq S_c$. It is a known fact [6], [5,10] that S_c is a perfect set, in particular S_c is closed and $\overline{S_c} = S_c$. Thus it is enough to show that $L_c \subset S_c$.

Fix a FO property P . We want to prove that $\lim_{n \rightarrow \infty} \mathbf{P}(P(G_n))$ lies in S_c . That is, we want to show that for some $\mathcal{T} \subseteq \mathcal{U}$

$$\lim_{n \rightarrow \infty} \mathbf{P}(P(G_n)) = \sum_{H \in \mathcal{T}} p_H. \tag{7}$$

Let $H \in \mathcal{U}$. First we will prove that

$$\lim_{n \rightarrow \infty} \mathbf{P}(P(G_n)) = \sum_{H \in \mathcal{U}} \lim_{n \rightarrow \infty} \mathbf{P}(\text{Frag}_n \simeq H) \mathbf{P}(P(G_n) \mid \text{Frag}_n \simeq H). \tag{8}$$

Let H_1, \dots, H_i, \dots be an enumeration of \mathcal{U} . Fix an arbitrarily small real constant $\epsilon > 0$. Notice that the events of the form $F_n \simeq H_i$ are disjoint for each i . So we obtain:

$$\lim_{n \rightarrow \infty} \mathbf{P}(P(G_n)) = \lim_{n \rightarrow \infty} \sum_{i=1}^{\infty} (\text{Frag}_n \simeq H_i) \mathbf{P}(P(G_n) \mid \text{Frag}_n \simeq H_i).$$

Let $m = \lim_{n \rightarrow \infty} \mathbb{E}[|\text{Frag}_n|]$ and let $M = m/\epsilon$. There exists some $j_0 \in \mathbb{N}$ such that $|H_i| \geq M$ for all $i \geq j_0$. As a consequence, for all $j \geq j_0$,

$$\lim_{n \rightarrow \infty} \sum_{i=j}^{\infty} \mathbf{P}(\text{Frag}_n \simeq H_i) \leq \epsilon.$$

And we obtain

$$\begin{aligned} & \left| \lim_{n \rightarrow \infty} \mathbf{P}(P(G_n)) - \sum_{i=1}^j \lim_{n \rightarrow \infty} \mathbf{P}(\text{Frag}_n \simeq H_i) \mathbf{P}(P(G_n) \mid \text{Frag}_n \simeq H_i) \right| \\ &= \lim_{n \rightarrow \infty} \sum_{i=j}^{\infty} \mathbf{P}(\text{Frag}_n \simeq H_i) \leq \epsilon. \end{aligned}$$

This proves Equation (8). Because of Lemma 5 for all $H \in \mathcal{U}$

$$\lim_{n \rightarrow \infty} \mathbf{P}(P(G_n) \mid \text{Frag}_n \simeq H) = 0 \text{ or } 1.$$

Let $T = \{i \in \mathbb{N} \mid \lim_{n \rightarrow \infty} \mathbf{P}(P(G_n) \mid \text{Frag}_n \simeq H_i) = 1\}$. Using Equation (8) we obtain

$$\sum_{i=1}^{\infty} \lim_{n \rightarrow \infty} \mathbf{P}(\text{Frag}_n \simeq H_i) \mathbf{P}(P(G_n) \mid \text{Frag}_n \simeq H_i) = \sum_{i \in T} p_{H_i}. \tag{9}$$

This proves Equation (7) and as a consequence $\overline{L_c} \subset S_c$.

Now we proceed to prove $S_c \subset \overline{L_c}$. Let $T \subset \mathcal{U}$, and let $\epsilon > 0$ be an arbitrarily small real number. We will show that there exists a FO property P such that

$$\left| \lim_{n \rightarrow \infty} \mathbf{P}(P(G_n)) - \sum_{H \in T} p_H \right| \leq \epsilon. \tag{10}$$

First, notice that Lemma 9 implies that

$$\sum_{H \in T} p_H = \lim_{n \rightarrow \infty} \mathbf{P}\left(\bigvee_{H \in T} p_H\right).$$

We define the property Q in the following way

$$Q(G) := \bigvee_{H \in T} \text{Frag}(G) \sim H.$$

Let $m := \lim_{n \rightarrow \infty} \mathbb{E}[|\text{Frag}_n|]$, and let $M = 2m/\epsilon$. Then using Markov's inequality:

$$\lim_{n \rightarrow \infty} \mathbf{P}(\text{Frag}_n \in \mathcal{U}_{\leq M}) \geq 1 - \epsilon/2. \quad (11)$$

Also, using that $\text{Frag}_n \neq \text{Frag}_n^M$ implies that $|\text{Frag}_n| \geq M$,

$$\lim_{n \rightarrow \infty} \mathbf{P}(\text{Frag}_n \neq \text{Frag}_n^M) \leq \mathbf{P}(\text{Frag}_n \in \mathcal{U}_{\leq M}) \leq \epsilon/2. \quad (12)$$

Let $T' = T \cap \mathcal{U}_{\leq M}$. As $\mathcal{U}_{\leq M}$ is a finite set so is T' . Define the properties Q' and P as

$$Q'(G) = \bigvee_{H \in T'} \text{Frag}(G) \sim H, \quad \text{and} \quad P(G) = \bigvee_{H \in T'} \text{Frag}^M(G) \sim H.$$

Notice that P can be expressed in FO logic. Also, the implications $Q' \Rightarrow Q$ and $Q' \Rightarrow P$ hold. If $Q(G)$ holds and $\neg Q'(G)$ holds as well then in particular $|\text{Frag}(G)| \geq M$. The same happens if $P(G)$ and $\neg Q'(G)$ hold at the same time. Joining everything we get

$$\begin{aligned} \left| \lim_{n \rightarrow \infty} \mathbf{P}(P(G_n)) - \sum_{H \in T} p_H \right| &= \left| \lim_{n \rightarrow \infty} \mathbf{P}(P(G_n)) - \lim_{n \rightarrow \infty} \mathbf{P}(Q(G_n)) \right| \\ &\leq \left| \lim_{n \rightarrow \infty} \mathbf{P}(P(G_n)) - \lim_{n \rightarrow \infty} \mathbf{P}(Q'(G_n)) \right| \\ &\quad + \left| \lim_{n \rightarrow \infty} \mathbf{P}(Q'(G_n)) - \lim_{n \rightarrow \infty} \mathbf{P}(Q(G_n)) \right| \\ &\leq 2\mathbf{P}(|\text{Frag}_n| \geq M) \leq \epsilon. \end{aligned}$$

■

3.4 | \overline{L}_c is always a finite union of intervals

Since we have shown that there is no gap for $c \geq 1$ we only need to consider $c < 1$. Let H_1, \dots, H_n, \dots be an enumeration of \mathcal{U} such that $p_{H_i}(c) \leq p_{H_j}(c)$ for all $i \leq j$. We shorten p_{H_i} to p_i . Because of Lemma 7 proving that \overline{L}_c is a finite union of intervals amounts to showing that for all i large enough

$$p_i \leq \sum_{j>i} p_j. \quad (13)$$

Let $f = f(c)$ be as defined in Equation (3), and let $s = ce^{-c}$, and notice that as $c < 1$ we have $s < 1$ as well. We can rewrite the p_i given by Equation (5) as

$$p_i = e^{-f} \frac{s^{|H_i|}}{\text{aut}(H_i)}.$$

For $i \geq 1$ let $k(i)$ be the least integer such that

$$e^{-f} s^{k(i)-1} \geq p_i > e^{-f} s^{k(i)}.$$

Notice if $k \geq k(i)$ and $H_j \in \mathcal{U}_k$ then $p_j < e^{-f} s^k < p_i$ since $\text{aut}(H_j) \geq 1$. For the same reason we also obtain that $|H_i| \leq k(i) - 1$. Hence to prove (13) it is sufficient to show that

$$p_i \leq \sum_{k \geq k(i)} \sum_{H_j \in \mathcal{U}_k} p_j. \tag{14}$$

Let $C_{x,y}$ denote the graph in \mathcal{U} consisting of a cycle of length x with a path of length y attached to one of its vertices. If $y = 0$ then $\text{aut}(C_{x,y}) = 2x$, and $\text{aut}(C_{x,y}) = 2$ otherwise. Let $T_{x,y,z}$ be the graph consisting of a triangle with paths of length $x, y,$ and z attached to its three vertices. Note that $\text{aut}(T_{x,y,z}) = 1$ if x, y, z are distinct, $\text{aut}(T_{x,y,z}) = 6$ if $x = y = z$, and $\text{aut}(T_{x,y,z}) = 2$ otherwise. It is easy to see that $C_{3,k-3}, C_{4,k-4}, \dots, C_{k-1,1}$ together with $T_{0,1,k-4}, T_{0,2,k-5}, \dots, T_{0,[(k-3)/2],[(k-3)/2]}$ form a family of different elements of \mathcal{U}_k . We have that for $k \geq 3$

$$\sum_{i=3}^{k-1} p_{C_{i,k-i}} = e^{-f} s^k \frac{k-3}{2}.$$

If k is odd $T_{0,[(k-3)/2],[(k-3)/2]}$ has two automorphisms, and the remaining $T_{i,k-3-i}$ with $i \geq 1$ each have only one automorphism. If k is even then all of $T_{0,1,k-4}, T_{0,2,k-5}, \dots, T_{0,[(k-3)/2],[(k-3)/2]}$ have exactly one automorphism. This gives

$$\sum_{i=1}^{[(k-3)/2]} p_{T_{0,i,k-3-i}} = e^{-f} s^k \frac{k-4}{2}, \quad \text{for } k \geq 4.$$

Using the last two equations it follows that for $k \geq 4$

$$\sum_{H \in \mathcal{U}_k} \frac{1}{\text{aut}(H)} \geq e^{-f} s^k \frac{2k-7}{2}. \tag{15}$$

Hence if i is such that $(2k(i) - 7)/2 > 1/s$ (i.e., $k(i) > 1/s + 7/2$) then

$$\sum_{j>i} p_j \geq \sum_{H_j \in \mathcal{U}_{k(i)}} p_j \geq e^{-f} s^{k(i)} \frac{2k-7}{2} > e^{-f} s^{k(i)-1} \geq p_i.$$

Note that $k(i) > 1/s + 7/2$ whenever $|H_i| + 1 \geq 1/s + 7/2$, and this is true for sufficiently large i . We have seen that, for any $0 < c < 1$, it is indeed the case that $p_i < \sum_{j>i} p_j$ for all sufficiently large i , as was to be proved.

3.5 | No gap when $c_0 \leq c < 1$

Fix $c \in [c_0, 1)$. Notice that in this case then $s = ce^{-c}$ satisfies

$$\frac{1}{3} < s < \frac{1}{e}.$$

Let i be such that $k(i) \geq 4$. Then, using (15) we obtain

$$\sum_{j>i} p_j \geq \sum_{k \geq k(i)} \sum_{H_j \in \mathcal{U}_k} p_j \geq \sum_{k \geq k(i)} e^{-fs^k} \frac{2k-7}{2}.$$

And using $\sum_{k=0}^\infty a^k(b+ck) = \frac{b}{1-a} + \frac{ca}{(1-a)^2}$ together with $s > 1/3$ we obtain that

$$\sum_{j>i} p_j \geq e^{-fs^{k(i)}} \left(\frac{2k(i)-7}{2(1-s)} + \frac{s}{(1-s)^2} \right) \geq e^{-fs^{k(i)}} \frac{3k(i)-9}{2}.$$

In particular, since $\frac{3k-9}{2} \geq 3 > 1/s$ for all $k \geq 5$, if $p_i \leq s^4$ then $p_i < \sum_{j>i} p_j$. As a consequence, if $|H_i| \geq 4$ then $p_i < \sum_{j>i} p_j$.

The only two cases left to consider are the ones when H_i is either the empty graph or the triangle. If H_i is the empty graph then necessarily $i = 1$ because the empty graph is the most likely fragment. By the definition of p_0 critically we have $p_1 \leq 1/2$ if $c \geq c_0$, hence $p_1 \leq \sum_{j>1} p_j$. If H_i is the triangle graph, then $p_i = e^{-fs^3}/6$ and

$$\sum_{j>i} p_j = \sum_{k \geq 4} \sum_{H_j \in \mathcal{U}_k} p_j \geq \sum_{k \geq 4} e^{-fs^k} \frac{2k-7}{2} \geq e^{-fs^4} \frac{3}{2} \geq e^{-fs^3} \frac{1}{6} = p_i,$$

as needed.

Thus $p_i \leq \sum_{j>i} p_j$ for every i , as we needed to prove.

4 | PROOF OF THEOREM 2

Recall that we consider the model $G_n^d = G^d(n, p = c/n^{d-1})$ of random d -uniform hypergraphs where each d -edge has probability $p = c/n^{d-1}$ of being in G_n^d independently, with $c > 0$. Throughout this section we consider $d \geq 3$ as being fixed and we will refer to “ d -uniform hypergraphs” simply as hypergraphs. The FO language of d -uniform hypergraphs is the FO language with a d -ary relation which is anti-reflexive and completely symmetric. Analogously to the case of graphs, this relation symbolizes the adjacency relation in the context of d -uniform hypergraphs.

The following is an analog of Lynch’s convergence law for random hypergraphs and can be found in [11, Proposition 6.4] and in more detail for more general relational structures in [8].

Theorem. *Let $p(n) \sim c/n^{d-1}$. Then for each FO sentence ϕ , the following limit exists:*

$$p_c(\phi) = \lim_{n \rightarrow \infty} \mathbf{P}(G_n^d \models \phi).$$

Moreover, $p_c(\phi)$ is a combination of sums, products, exponentials and a set of constants Λ_c , hence it is an analytic function of c .

As before we consider the set

$$L_c = \left\{ \lim_{n \rightarrow \infty} \mathbf{P}(G_n^d \models \phi) : \phi \text{ FO sentence, and } p(n) \sim c/n^{d-1} \right\}.$$

4.1 | Hypergraph preliminaries

Given a hypergraph H , we denote the set of vertices by $V(H)$ and the set of edges by $E(H)$. As in [7] we define the *excess* $\text{ex}(H)$ of H as the quantity

$$\text{ex}(H) = (d - 1)|H| - |V(H)|.$$

It is easily seen that the minimum excess of a connected hypergraph is -1 . A *tree* T is a connected hypergraph satisfying $\text{ex}(T) = -1$. Equivalently, a tree is a hypergraph that can be obtained gluing edges repeatedly to an initial vertex in such a way that each new edge intersects the hypergraph obtained so far in exactly one vertex. A *unicycle* is a connected hypergraph of excess 0, and a *cycle* is a minimal unicycle. Equivalently a cycle is a connected hypergraph H where every edge shares exactly two vertices with the remaining edges, and a unicycle is a cycle with disjoint trees attached to each vertex. A k -cycle is a cycle with k edges.

It is shown in [12] that a phase transition in the structure of G_n^d occurs when $c = (d - 2)!$, similar to the one for random graphs. In particular, we have the following results [12, Theorem 3.6].

Lemma 10. *Let $p(n) \sim c/n^{d-1}$ with $0 < c < (d - 2)!$. Then a.a.s. all connected components of G_n^d are either trees or unicycles.*

The proofs of the next results are very similar to those for graphs presented in Section 2 and are omitted.

Lemma 11. *Let $p \sim c/n^{d-1}$ with $c > 0$. For each $k \geq 2$, let $X_{n,k}$ be the random variable equal to the number of k -cycles in G_n^d , and let $\lambda_k = \left(\frac{c}{(d-2)!}\right)^k$. Then for fixed $k \geq 2$*

- (1) $E[X_{n,k}] \leq \lambda_k$,
- (2) $\lim_{n \rightarrow \infty} E[X_{n,k}] = \lambda_k$,
- (3) $X_{n,k}$ converges in distribution to a Poisson variable with mean λ_k as $n \rightarrow \infty$.

Furthermore, for any fixed $k \geq 2$ the variables $X_{n,2}, \dots, X_{n,k}$ are asymptotically independent.

Corollary 3. *Let $p \sim c/n^{d-1}$ with $c > 0$. Set*

$$f(c) = \sum_{k \geq 2} \left(\frac{c}{(d-2)!}\right)^k \frac{1}{2k} = \frac{1}{2} \ln \frac{1}{1 - \frac{c}{(d-2)!}} - \frac{c}{2(d-2)!}. \tag{16}$$

Let X_n be the random variable equal to the total number of cycles in G_n^d . Then

$$\lim_{n \rightarrow \infty} E[X_n] = f(c),$$

and

$$\lim_{n \rightarrow \infty} \mathbf{P}(G_n^d \text{ contains no cycles}) = e^{-f(c)} = \exp\left(\frac{c}{2(d-2)!}\right) \sqrt{1 - \frac{c}{2(d-2)!}}.$$

Lemma 12. *Let $p \sim c/n^{d-1}$ with $0 < c < (d - 2)!$. Let Z_n be the random variable equal to the number of cycles in G_n^d that belong to connected components that are not unicycles. Then*

$$\lim_{n \rightarrow \infty} E[Z_n] = 0.$$

Let \mathcal{U} be the family of unlabeled d -hypergraphs whose connected components are unicyclic.

Lemma 13. *Let $p \sim c/n^{d-1}$ with $c > 0$. Let $\mathcal{T} \subset \mathcal{U}$ be a finite set of unicycles. For each $H \in \mathcal{T}$ let $X_{n,H}$ be the random variable that counts the connected components in G_n^d that are isomorphic to H , and set*

$$\lambda_H = \frac{(ce^{-c/(d-2)!})^{|H|}}{\text{aut}(H)}.$$

Then $X_{n,H}$ converges in distribution to a Poisson variable with mean λ_H as $n \rightarrow \infty$ and the $X_{n,H}$ are asymptotically independent, that is

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\bigwedge_{H \in \mathcal{T}} X_{n,H} = a_H \right) = \prod_{H \in \mathcal{T}} e^{-\lambda_H} \frac{\lambda_H^{a_H}}{a_H!}.$$

Define the fragment of hypergraph as the collection of components that are unicycles and let Frag_n be the fragment of the random hypergraph G_n^d .

Lemma 14. *Let $p \sim c/n^{d-1}$ with $0 < c < (d - 2)!$. Then the limit*

$$\lim_{n \rightarrow \infty} \mathbf{E} [|\text{Frag}_n|]$$

exists and is a finite quantity.

Lemma 15. *Let $p \sim c/n^{d-1}$ with $0 < c < (d - 2)!$. Let ϕ be a FO sentence and let $H \in \mathcal{U}$. Then*

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(G_n^d \models \phi \mid \text{Frag}_n \simeq H \right) = 0 \text{ or } 1.$$

Moreover, the value of the limit depends only on ϕ and H , and not on c .

As for graphs, we divide the proof of Theorem 2 into several cases. Along the way we analyze the distribution of the fragment and the number of automorphisms in hypergraphs.

4.2 | No gap when $c \geq (d - 2)!$

The arguments here mirror exactly those in Section 3.1. For each k let $X_{n,k}$ be the random variable equal to the number of k -cycles in G_n^d . Then

$$X_{n, \leq k} = X_{n,2} + \dots + X_{n,k} \xrightarrow[n \rightarrow \infty]{d} \text{Po} \left(\sum_{i=2}^k \frac{(c/(d-2)!)^i}{2^i} \right).$$

If $c \geq (d - 2)!$ then $\sum_{i=2}^k \frac{(c/(d-2)!)^i}{2^i}$ tends to infinity and we can use the Central Limit Theorem to approximate any $p \in (0, 1)$ with FO statements of the form “ $X_{n, \leq k} \leq a$.”

4.3 | At least when gap for $c < c_0$

Let $f(c)$ be as in Equation (16), and let c_0 be the unique solution of $e^{-f(c)} = 1/2$ lying in $[0, (d - 2)!$. By Lemma 3, c_0 is the only value for which

$$\lim_{n \rightarrow \infty} \mathbf{P}(G_n^d) = 1/2,$$

where $p(n) \sim c/n^{d-1}$. One can check that this is achieved when the expected degree $c/(d - 2)!$ is $\simeq 0.898$, independently of d .

Let $p(n) \sim c/n^{d-2}$ with $0 < c < c_0$. Because of Lemma 15, for any FO sentence ϕ

$$\lim_{n \rightarrow \infty} \mathbf{P}\left(G_n^d \models \phi \mid G_n^d \text{ contains no cycles}\right) = 0 \text{ or } 1.$$

From this point we continue as in Section 3.2 to show that $(1 - e^{-f(c)}, e^{-f(c)})$ is a gap of \overline{L}_c .

4.4 | Asymptotic distribution of the fragment for $c < (d - 2)!$

The same proof of Lemma 8 can be used to prove the following result.

Theorem 4. *Let $p(n) \sim c/n^{d-1}$ with $0 < c < (d - 2)!$. Let $H \in \mathcal{U}$. Then*

$$\lim_{n \rightarrow \infty} \mathbf{P}(\text{Frag}_n \simeq H) = e^{-f(c)} \frac{(e^{-c/(d-2)!}c)^{|H|}}{\text{aut}(H)}.$$

For each $H \in \mathcal{U}$ define $p_H(c) = p_H = \lim_{n \rightarrow \infty} \mathbf{P}(\text{Frag}_n \simeq H)$. Consider the set

$$S_c = \left\{ \sum_{H \in \mathcal{T}} p_H(c) : \mathcal{T} \subseteq \mathcal{U} \right\}.$$

One can proceed exactly as in Theorem 3 to prove the following:

Theorem 5. *Let $0 < c < (d - 2)!$. Then $\overline{L}_c = S_c$.*

4.5 | A lower bound on the number of automorphisms of unicyclic hypergraphs

Let H be a hypergraph and $h \in E(H)$ an edge. We call a vertex v lying in e *free* if e is the only edge that contains v . We denote by $\text{free}(h)$ the number of free vertices in e . Notice that

$$\text{aut}(H) \geq \prod_{h \in E(H)} \text{free}(h)!,$$

because free vertices inside an edge can be permuted without restriction. Given a unicycle H we define the *leaves* of H as the edges $e \in E(H)$ that contain only one nonfree vertex.

Lemma 16. *Let $H \in \mathcal{U}$ be a d -hypergraph. Then,*

$$\frac{(d - 2)!^{|H|}}{\text{aut}(H)} \leq \frac{(d - 2)^2}{(d - 1)^2}.$$

Proof. It suffices to prove the statement for unicycles, since

$$\frac{(d - 2)!^{|H|}}{\text{aut}(H)} \leq \prod_i \frac{(d - 2)!^{|H_i|}}{\text{aut}(H_i)},$$

where the H_i are the connected components of H .

Let λ be the number of leaves in H . We show by induction that

$$\prod_{h \in E(H)} \frac{(d-2)!}{\text{free}(h)!} \leq \left(\frac{d-2}{d-1}\right)^\lambda. \tag{17}$$

If $\lambda = 0$ then H is a cycle and each of its edges contains exactly $d - 2$ free vertices, so that

$$\prod_{h \in E(H)} \frac{(d-2)!}{\text{free}(h)!} = 1,$$

and H satisfies (17). Now let H be a unicycle satisfying (17). Add a new edge h' to H to obtain another unicycle H' . Since h' intersects H in only one vertex v , it follows that h' is a leaf of H' . There are two possibilities:

- $\lambda(H') = \lambda(H)$. In this case no new leaves are created with the addition of h' . This means that v is a free vertex in one leaf g of H (that is, h' “grows” out of g), and

$$\prod_{h \in E(H')} \frac{(d-2)!}{\text{free}(h)!} = \prod_{h \in E(H)} \frac{(d-2)!}{\text{free}(h)!}.$$

- $\lambda(H') = \lambda(H) + 1$. In this case h' intersects an edge of H that is not a leaf. The case that maximizes $\prod_{h \in E(H')} \frac{(d-2)!}{\text{free}(h)!}$ is when h' grows out of a free vertex of an edge in H with exactly $d - 2$ free vertices. In this case

$$\prod_{h \in E(H')} \frac{(d-2)!}{\text{free}(h)!} = \frac{d-2}{d-1} \prod_{h \in E(H)} \frac{(d-2)!}{\text{free}(h)!},$$

and H' satisfies (17) as well.

Finally, as all unicycles can be obtained adding edges to a cycle successively, (17) holds for all unicycles.

To prove the original statement consider the cases $\lambda = 0$, $\lambda = 1$ and $\lambda \geq 2$.

- If $\lambda = 0$ then H is a cycle of length $l \geq 2$ and $\text{aut}(H) = (d - 2)!2l$, yielding

$$\frac{(d-2)!^{|H|}}{\text{aut}(H)} = \frac{1}{2l} \leq \frac{(d-2)^2}{(d-1)^2},$$

since $1/2l \leq 1/4 \leq (d - 2)^2/(d - 1)^2$ for all $l \geq 2, d \geq 3$.

- If $\lambda = 1$ then H is a cycle with a path attached to it. In this case, H has a nontrivial automorphism (a reflection of the cycle) and as a consequence $2 \prod_{h \in E(H)} \text{free}(h)! \leq \text{aut}(H)$. Using this and (17) we get

$$\frac{(d-2)!^{|H|}}{\text{aut}(H)} \leq \frac{1}{2} \prod_{h \in E(H)} \frac{(d-2)!}{\text{free}(h)!} \leq \frac{1}{2} \left(\frac{d-2}{d-1}\right) \leq \left(\frac{d-2}{d-1}\right)^2,$$

as we wanted.

- Finally, when $\lambda \geq 2$ the relation (17) suffices, since

$$\prod_{h \in E(H)} \frac{(d-2)!}{\text{free}(h)!} \leq \left(\frac{d-2}{d-1}\right)^\lambda \leq \left(\frac{d-2}{d-1}\right)^2.$$

■

4.6 | Some families of unicycles

In this section, we introduce three families of hypergraphs having a small number of automorphisms, and that will be used in the subsequent proofs.

- Let $T_{\alpha,\beta}$ denote the hypergraph consisting of a triangle (as a d -hypergraph) with two paths of length α and β respectively attached to two of its free vertices, each one from a different edge. One can check that

$$\frac{(d-2)!^{|T_{\alpha,\beta}|}}{\text{aut}(T_{\alpha,\beta})} = \frac{(d-2)^{\alpha+\beta+3}}{\text{aut}(T_{\alpha,\beta})} = \begin{cases} \left(\frac{d-2}{d-1}\right)^2 & \text{for } \alpha \neq \beta, \\ \frac{1}{2} \left(\frac{d-2}{d-1}\right)^2 & \text{otherwise.} \end{cases}$$

Let \mathcal{T} be the family of hypergraphs $\{T_{\alpha,\beta} : \alpha, \beta > 0\}$. Then for $k \geq 4$

$$\sum_{H \in \mathcal{T}, |H|=k} \frac{(d-2)!^{|H|}}{\text{aut}(H)} = \sum_{\alpha=1}^{\lfloor \frac{k-3}{2} \rfloor} \frac{(d-2)!^k}{\text{aut}(T_{\alpha,k-3-\alpha})} = \frac{k-4}{2} \left(\frac{d-2}{d-1}\right)^2. \tag{18}$$

- Let $B_{\alpha,\beta}$ denote the hypergraph consisting of a two-cycle with two paths of length α and β respectively attached to two of its free vertices, each one from a different edge. In this case

$$\frac{(d-2)!^{|B_{\alpha,\beta}|}}{\text{aut}(B_{\alpha,\beta})} = \frac{(d-2)^{\alpha+\beta+2}}{\text{aut}(B_{\alpha,\beta})} = \begin{cases} \frac{1}{2} \left(\frac{d-2}{d-1}\right)^2 & \text{for } \alpha \neq \beta, \\ \frac{1}{4} \left(\frac{d-2}{d-1}\right)^2 & \text{otherwise.} \end{cases}$$

Let $\mathcal{B} = \{B_{\alpha,\beta} : \alpha, \beta > 0\}$. Then for $k \geq 3$

$$\sum_{H \in \mathcal{B}, |H|=k} \frac{(d-2)!^{|H|}}{\text{aut}(H)} = \sum_{\alpha=1}^{\lfloor \frac{k-2}{2} \rfloor} \frac{(d-2)!^k}{\text{aut}(B_{\alpha,k-2-\alpha})} = \frac{k-3}{4} \left(\frac{d-2}{d-1}\right)^2. \tag{19}$$

- We denote by $O_{\alpha,\beta}$, the hypergraph formed by attaching a path of length β to a free vertex of a cycle of length α . One can check that $|O_{\alpha,\beta}| = \alpha + \beta$ and that $\frac{(d-2)!^{\alpha+\beta}}{\text{aut}(O_{\alpha,\beta})} = \frac{1}{2} \left(\frac{d-2}{d-1}\right)$.

Let $\mathcal{O} = \{O_{\alpha,\beta} : \alpha > 1, \beta > 0\}$. Then for $k \geq 2$

$$\sum_{H \in \mathcal{O}, |H|=k} \frac{(d-2)!^{|H|}}{\text{aut}(H)} = \sum_{\alpha=2}^{k-1} \frac{(d-2)!^k}{\text{aut}(O_{\alpha,k-\alpha})} = \frac{k-2}{2} \left(\frac{d-2}{d-1}\right). \tag{20}$$

4.7 | \overline{L}_c is always a finite union of intervals

Fix $0 < c < (d-2)!$. Let H_1, \dots, H_n, \dots be an enumeration of \mathcal{U} such that $p_{H_i} \leq p_{H_j}$ for all $i \leq j$. As before we shorten p_{H_i} to p_i . Analogously to Section 3.4 we need to prove that for i large enough

$$p_i \leq \sum_{j>i} p_j.$$

Let $f = f(c)$ be as defined in Equation (16), and let $s = \frac{c}{(d-2)!} e^{-c/(d-2)!}$. By Lemma 4 we have that

$$p_i = e^{-f} s^{|H_i|} \frac{(d-2)!^{|H_i|}}{\text{aut}(H_i)}. \tag{21}$$

For $i > 0$ we define $k(i)$ as the unique integer such that

$$e^{-f} s^{k(i)-1} \left(\frac{d-2}{d-1}\right)^2 \geq p_i > e^{-f} s^{k(i)} \left(\frac{d-2}{d-1}\right)^2$$

Notice that because of Lemma (16), we have $|H_i| \leq k(i) - 1$.

As a consequence, if $k = k(i) \geq 4$ then

$$\begin{aligned} \sum_{j>i} p_j &\geq s^k \sum_{H \in \mathcal{U}_k} \frac{(d-2)!^k}{\text{aut}(H)} \\ &\geq s^k \frac{k-4}{2} \left(\frac{d-2}{d-1}\right)^2. \end{aligned}$$

This is obtained taking into account only the hypergraphs in \mathcal{T} and using Equation (18). The last inequality implies that if $k(i)$ is such that $\frac{1}{s} \leq \frac{k(i)-4}{2}$ then $p_i \leq \sum_{j>i} p_j$. This clearly holds for i large enough, hence \overline{L}_c is a finite union of intervals, as needed to be proved.

4.8 | No gap when $c_0 \leq c < (d-2)!$

Fix $c_0 \leq c < (d-2)!$. Let H_1, \dots, H_n, \dots be an enumeration of \mathcal{U} satisfying the same conditions as before. We want to show that for all i

$$p_i \leq \sum_{j>i} p_j. \tag{22}$$

Notice that $s = \frac{c}{(d-2)!} e^{-c/(d-2)!}$ satisfies that

$$\frac{1}{3} < s < \frac{1}{e},$$

because $0.898 \leq c/(d-2)! < 1$. The following inequalities are obtained using Equations (18)–(20) respectively, together with the formula for the sum of an arithmetic–geometric series and the fact that $1/3 < s$.

$$\sum_{H \in \mathcal{T}, |H| \geq k} p_H \geq e^{-f} s^k \frac{6k-21}{8} \left(\frac{d-1}{d-2}\right)^2 \quad \text{for } k \geq 4. \tag{23}$$

$$\sum_{H \in \mathcal{B}, |H| \geq k} p_H \geq e^{-f} s^k \frac{6k-15}{16} \left(\frac{d-1}{d-2}\right)^2 \quad \text{for } k \geq 3. \tag{24}$$

$$\sum_{H \in \mathcal{O}, |H| \geq k} p_H \geq e^{-f} s^k \frac{6k-9}{8} \left(\frac{d-1}{d-2}\right)^2 \quad \text{for } k \geq 2. \tag{25}$$

Assume first that $k = k(i) \geq 5$. Then

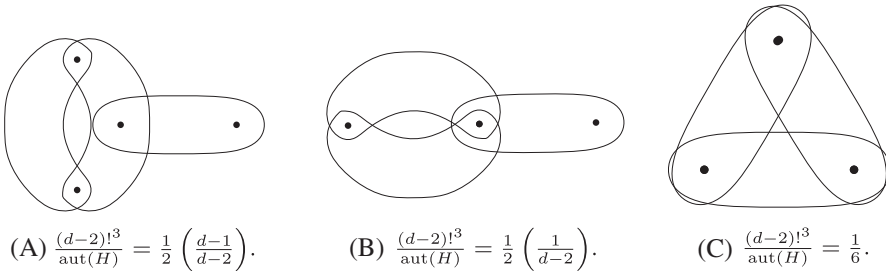
$$\begin{aligned} \sum_{j>i} p_j &\geq e^{-f} s^k \left[\frac{18k - 57}{16} \left(\frac{d-2}{d-1} \right)^2 + \frac{6k-9}{8} \left(\frac{d-2}{d-1} \right) \right] \\ &\geq e^{-f} s^k \frac{30k-75}{16} \left(\frac{d-2}{d-1} \right)^2 \geq e^{-f} s^k 3 \left(\frac{d-2}{d-1} \right)^2 \\ &\geq e^{-f} s^{k-1} \left(\frac{d-2}{d-1} \right)^2 \geq p_i, \end{aligned}$$

as was to be proven.

Otherwise, suppose that $k = k(i) \leq 4$. Notice that by Lemma 16 necessarily $|H_i| \leq 3$. We have three cases:

- $|H_i| = 3$. In this case, the following enumeration of all (unlabeled) unicycles of size 3 gives that

$$e^{-f} s^3 \frac{1}{2} \left(\frac{d-1}{d-2} \right) \geq p_i.$$



Proceeding as before we obtain

$$\begin{aligned} \sum_{j>i} p_j &\geq e^{-f} s^4 \left[\frac{184 - 57}{16} \left(\frac{d-2}{d-1} \right)^2 + \frac{64 - 9}{8} \left(\frac{d-2}{d-1} \right) \right] \\ &\geq e^{-f} s^4 \left[\frac{15}{16} \frac{1}{2} \left(\frac{d-2}{d-1} \right) + \frac{30}{8} \frac{1}{2} \left(\frac{d-2}{d-1} \right) \right] \\ &\geq e^{-f} s^4 \frac{3}{2} \left(\frac{d-2}{d-1} \right) \geq e^{-f} s^3 \frac{1}{2} \left(\frac{d-2}{d-1} \right) \geq p_i. \end{aligned}$$

- $|H_i| = 2$. In this case H_i is the 2-cycle, and $p_i = e^{-f} s^2 \frac{1}{4}$. Using Equations (24) and (25) we obtain

$$\begin{aligned} \sum_{j>i} p_j &\geq p_{C_3} + \sum_{H \in \mathcal{B}} p_H + \sum_{H \in \mathcal{O}} p_H \\ &\geq e^{-f} s^3 \left[\frac{1}{6} + \frac{3}{16} \left(\frac{d-2}{d-1} \right)^2 + \frac{9}{8} \left(\frac{d-2}{d-1} \right) \right] \\ &\geq e^{-f} s^3 \left[\frac{4}{6} \frac{1}{4} + \frac{3}{16} \frac{1}{4} + \frac{18}{8} \frac{1}{4} \right] \geq e^{-f} s^3 3 \frac{1}{4} \geq p_i. \end{aligned}$$

- $|H_i| = 0$. In this case H_i is the empty graph and $p_i \geq 1/2$ by hypothesis.

5 | CONCLUDING REMARKS

It can be shown that in Theorem 1 the number of intervals in which \bar{L}_c decomposes when $c < c_0$ is unbounded as $c \rightarrow 0$. It would be interesting to determine at which rate the number of intervals grows as $c \rightarrow 0$. It is a delicate issue, since the decreasing ordering of the possible fragments according to their limiting probabilities changes with c in an intricate way.

REFERENCES

1. B. Bollobás, *Random graphs*, Cambridge University Press, Cambridge, 2001.
2. P. Erdős and A. Rényi, *On the evolution of random graphs*, *Publ. Math. Inst. Hung. Acad. Sci.* **5** (1960), 17–60.
3. A. Frieze and M. Karoński, *Introduction to random graphs*, Cambridge University Press, Cambridge, 2016.
4. P. Heinig, T. Müller, M. Noy, and A. Taraz, *Logical limit laws for minor-closed classes of graphs*, *J. Combin. Theory Ser. B* **130** (2018), 158–206.
5. H. Hornich, *Über beliebige teilsommen absolut konvergenter reihen*, *Monatshefte Math. Phys.* **49** (1941), 316–320.
6. S. Kakeya, *On the partial sums of an infinite series*, *Tohoku Sci. Rep.* **3** (1914), 159–164.
7. M. Karoński and T. Łuczak, *The phase transition in a random hypergraph*, *J. Comput. Appl. Math.* **142** (2002), 125–135.
8. A. Larrauri, *On the limiting probabilities of sentences about sparse hypergraphs*, *J. Logic Comput.* **31** (2021), 444–472.
9. J. F. Lynch, *Probabilities of sentences about very sparse random graphs*, *Random Struct. Algorithms* **3** (1992), 33–53.
10. J. Nymann and R. Saenz, *On a paper of Guthrie and Nymann on subsums of infinite series*, *Colloquium Math.* **83** (2000), 1–4.
11. N. C. Saldanha and M. Telles, *Spaces of completions of elementary theories and convergence laws for random hypergraphs*. arXiv preprint arXiv:1602.06537, 2016.
12. J. Schmidt-Pruzan and E. Shamir, *Component structure in the evolution of random hypergraphs*, *Combinatorica* **5** (1985), 81–94.
13. S. Shelah and J. Spencer, *Can you feel the double jump?* *Random Struct. Algorithms* **5** (1994), 191–204.

How to cite this article: A. Larrauri, T. Müller, and M. Noy, *Limiting probabilities of first order properties of random sparse graphs and hypergraphs*, *Random Struct. Algorithms*. **60** (2022), 506–526. <https://doi.org/10.1002/rsa.21041>