First order complexity of finite random structures

Danila Demin
Moscow Institute of Physics and Technology
Dolgoprudny, Russia
St Petersburg University
Saint Petersburg, Russia
demin.da@phystech.edu

Maksim Zhukovskii
The University of Sheffield
Sheffield, UK
m.zhukovskii@sheffield.ac.uk

ABSTRACT

For a sequence of random structures with n-element domains over a relational signature, we define its FO complexity as a certain subset in the Banach space ℓ^{∞}/c_0. The well-known FO zero-one law and FO convergence law correspond to FO complexities equal to $\{0,1\}$ and a subset of $\mathbb{R}$, respectively. We present a hierarchy of FO complexity classes, introduce a stochastic FO reduction that allows to transfer complexity results between different random structures, and deduce using this tool several new logical limit laws for binomial random structures. Finally, we introduce a conditional distribution on graphs, subject to a FO sentence φ, that generalises certain well-known random graph models, show instances of this distribution for every complexity class, and prove that the set of all φ validating 0–1 law is not recursively enumerable.

CCS CONCEPTS

• **Theory of computation → Finite Model Theory**; Complexity theory and logic; • **Mathematics of computing → Random graphs**.

KEYWORDS

random structures, first order logic, logical limit laws, zero-one laws, random graphs, logical complexity, recursively enumerable languages

ACM Reference Format:
Danila Demin and Maksim Zhukovskii. 2024. First order complexity of finite random structures. In *39th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS '24), July 8–11, 2024, Tallinn, Estonia.* ACM, New York, NY, USA, 20 pages. https://doi.org/10.1145/3661814.3662127

1 INTRODUCTION

Let σ be a relational signature. A *random n-structure* over the signature σ is a random element of the set of all structures with domain $[n] := \{1,\ldots,n\}$ and over σ. Commonly it is supported by a set of σ-axioms $\mathcal{F}$, i.e. it is assumed that with probability 1 the random n-structure models $\mathcal{F}$. Let $\sigma = \{P_1,\ldots,P_s\}$, where P_i has arity d_i (we adopt a usual convention that σ includes the equality relation whose axioms are part of the logic). We denote the random n-structure uniformly distributed over all n-structures over σ by $D^{\sigma}(n \mid \mathcal{F})$ or $D^{(d_1,\ldots,d_s)}(n \mid \mathcal{F})$.

For example, the well known binomial random graph $G(n,p)$ is a random n-structure over $\sigma = \{=,\sim\}$, where $=$ represents the coincidence of vertices, and $\sim$ represents the graph adjacency relation. The axioms in

$$\mathcal{F} = \{\forall x \forall y\ (x \sim y) \Leftrightarrow (y \sim x),\ \forall x\ \neg(x \sim x)\}$$

allow to define the distribution on the set of undirected graphs without loops. The distribution of $G(n,p)$ is defined as follows: for a fixed n-structure G that models $\mathcal{F}$ (that is, G is an undirected graph without loops), its probability equals $p^{|E(G)|}(1-p)^{\binom{n}{2}-|E(G)|}$, where $E(G)$ is the set of unordered pairs $\{x,y\}$ such that $x \sim y$, (that is, the set of edges of G). In other words, edges appear independently with probability p.

Another example is a binomial random n-structure

$$D^{\sigma}(n,p_1,\ldots,p_s) = D^{(d_1,\ldots,d_s)}(n,p_1,\ldots,p_s)$$

over $\sigma = \{=,P_1,\ldots,P_s\}$ that does not have any predefined axioms: each interpretation appears with probability $p_1^{N_1}\ldots p_s^{N_s}(1-p_1)^{n^{d_1}-N_1}\ldots(1-p_s)^{n^{d_s}-N_s}$, where N_i is the number of d_i-tuples satisfying P_i. In other words, $D^{(d_1,\ldots,d_s)}(n,p_1,\ldots,p_s)$ is the binomial random hypergraph with hyperedges (ordered multisets) of cardinalities $d_1,\ldots,d_s$, where each hyperedge of "type P_i" appears with probability p_i independently of all the others. In particular, for a signature consisting of one predicate P of arity d, we get a binomial random directed d-uniform hypergraph $D^{(d)}(n,p)$.

The following fundamental result in finite model theory, known as the first order (FO) 0–1 law, was proven by Glebskii, Kogan, Liogonkii, Talanov [7] and independently by Fagin [5]: for a fixed finite relational signature σ, any FO sentence φ over σ is either true on asymptotically almost all n-structures or false. In other words, $\Pr(D^{(d_1,\ldots,d_s)}(n,1/2,\ldots,1/2) \models \varphi)$ converges either to 0 or to 1 as $n \to \infty$. Same arguments can be used to show that $D^{(d_1,\ldots,d_s)}(n,p_1,\ldots,p_s)$ obeys the FO 0–1 law for all constant $p_1,\ldots,p_s \in (0,1)$. Fagin [5] also proved that the same is true for graphs, i.e. $G(n,p)$ obeys the FO 0–1 law.

Since then, the validity of the FO 0–1 law was studied for many other random structures (binomial random graphs with $p = p(n)$ [21, 29, 31], random regular graphs [9], random trees [26], recursive random graphs [15, 22], random geometric graphs [25], random graphs embeddable on a surface [11], and many others [30, 32, 35]). However, no methods have been developed to transfer logical limit laws between different random structures. In particular, Fagin applied the same proof as for $D^{(d_1,\ldots,d_s)}(n,p_1,\ldots,p_s)$ to $G(n,p)$

LICS '24, July 8–11, 2024, Tallinn, Estonia

ACM ISBN 979-8-4007-0660-8/24/07
https://doi.org/10.1145/3661814.3662127

instead of transferring the law. One of the main contributions of our paper is such a transferring tool. In particular, it allows to transfer the FO 0–1 law from $D^{(2)}(n,p)$ to the binomial random graph as well as to the binomial random *directed* graph without loops.

Before moving on to a more detailed discussion of our results, let us notice an important application of logical limit laws to the study of hierarchy of logics and time complexity. Indeed, if a random structure satisfies the 0-1 (or convergence) law for a language $\mathcal{L}_1$ while it does not satisfy the law for $\mathcal{L}_2 \supseteq \mathcal{L}_1$, then the inclusion $\mathcal{L}_2 \supset \mathcal{L}_1$ is strict. This simple observation was used in [33, 34] to show the lower bound on the minimum quantifier depth and the minimum number of variables of a FO sentence describing the property of containing an induced subgraph isomorphic to a fixed given graph F. The latter fact implies certain limitations of the respective method of solving the induced subgraph isomorphism decision problem: the validity of a FO sentence with k variables on an n-vertex graph is decidable in time $O(n^k)$, see details in [33, 34].

For $G(n,p)$ with $p = p(n) = o(1)$, a breakthrough achievement was obtained by Shelah and Spencer. First of all note that the above mentioned classical FO 0–1 law for $G(n,p)$ can be generalised to all $p = p(n)$ such that $\min\{p, 1-p\}n^{\alpha} \to \infty$ as $n \to \infty$ for all $\alpha > 0$ (see [31]). So it is natural to further consider $p = n^{-\alpha}$, $\alpha > 0$. Shelah and Spencer [29] proved the following:

- If $\alpha \in (0,1]$ is rational, then $G(n,n^{-\alpha})$ does not obey the FO 0–1 law.
- If $\alpha \in (0,1]$ is irrational, then $G(n,n^{-\alpha})$ obeys the FO 0–1 law.
- If $\alpha = 1 + \frac{1}{k}$ for some $k \in \mathbb{N}$, then $G(n,n^{-\alpha})$ does not obey the FO 0–1 law.
- If either $p = o(n^{-2})$, or for some $k \in \mathbb{N}$, $n^{-1-1/k} \ll p \ll n^{-1-1/(k+1)}$, then $G(n,p)$ obeys the FO 0–1 law.

Our tool can be also used to show that the FO 0-1 law does not hold: for example, for $\alpha > 1$, the failure of the FO 0–1 law for $G(n,n^{-\alpha})$ transfers to the failure of the FO 0–1 law for $D^{(d+1)}(n,(d+\alpha-2)\frac{\ln n}{n})$ for every integer $d \geq 2$.

Remark 1.1. The random graph $G(n,n^{-\alpha})$ for $\alpha > 1$ is very sparse: with asymptotical probability 1 (*with high probability* or, for brevity, *whp* in what follows) it is a forest consisting only of tree components of bounded sizes (see, e.g., [12]). In contrast, $D^{(d+1)}(n,\Theta(\ln n/n))$ is weakly connected whp [28]. Since the validity of the FO 0–1 law for $G(n,n^{-\alpha})$ in this case follows immediately from standard properties of the logical equivalence (see, e.g., [27, Section 4.2]), so it is not surprising that our tool does not transfer the validity of FO 0–1 laws from $G(n,n^{-\alpha})$ to $D^{(d+1)}(n,\Theta(\ln n/n))$.

A special interest in combinatorial and probabilistic community was chained to properties of specifically $G(n,c/n)$ for constant c because of so-called phase transition phenomenon [4] — in particular, the emergence of a giant component. Lynch [21] proved that $G(n,c/n)$ obeys the *FO convergence law* (i.e. for every FO sentence the probability of its truth on $G(n,c/n)$ converges to some number in $[0,1]$ as $n \to \infty$), however the FO 0–1 law does not hold. Note that Shelah and Spencer [29] disproved even the FO convergence law for $G(n,n^{-\alpha})$ when $\alpha \in (0,1)$ is rational. Larrauri, Müller, and Noy [18] described possible limits of truth probabilities of FO sentences on $G(n,c/n)$. They proved that the closure of the set of limits in $[0,1]$ consists of a finite number of segments and determined the minimum constant c_0 for which this segment is unique and coincides with $[0,1]$. Also, they generalised this result to d-uniform unoriented hypergraphs.

We transfer the upper bound for this threshold c_0 from d-uniform unoriented hypergraphs to d-uniform H-hypergraphs, where H is an arbitrary subgroup of the symmetry group S_d (i.e. a hyperedge is an orbit under the action of H on d-tuples $(x_1,\ldots,x_d)$ of distinct elements from a fixed d-element set). For example, unoriented hypergraphs correspond to the case $H = S_d$. Regarding the lower bound, which coincides with the upper bound, it can be transferred from oriented hypergraphs (i.e. $H = \{\mathrm{id}\}$) to H-hypergraphs for any H. Luckily, the same proof method as from [18] can be applied to prove the tight lower bound for oriented hypergraphs as well. Thus, the problem of finding the threshold c_0 for any way of assigning an orientation to hyperedges can be reduced to two extreme cases $H = S_d$ and $H = \{\mathrm{id}\}$.

Our tool is a certain preorder on sequences of random n-structures, $n \in \mathbb{N}$, which we call *the stochastic FO reduction*. This preorder expresses a hierarchy of sequences of random structures: for higher sequences in this preorder, their FO limit behaviour is more complex. In particular, for each pair of sequences A and B such that A is reducible to B and B obeys the FO 0–1 law (convergence law), A obeys FO 0–1 law (convergence law) as well. Besides the above mentioned applications of the stochastic FO reduction to transferring logical laws, we prove that the stochastic FO reduction preorder defines *stable* equivalence classes of $G(n,p)$: a little change of p does not change the equivalence class. For example, this observation allows to transfer the absence of the FO 0–1 law from $D^{(1)}(n,\frac{c}{n})$ to $D^{(d+1)}(n,\frac{d\ln n}{n})$. The stochastic FO reduction is defined in Section 3.

The notion of stochastic FO reduction as well as different asymptotic logical behaviour of different random structures naturally lead to a concept of *FO complexity* of a sequence of random structures D_n. The FO complexity has to generically describe the limit behaviour of $\Pr(D_n \models \varphi)$ over all FO sentences φ in such a way that if A stochastically reduces to B, then B is at least as complex as A.

The FO complexity is defined in Section 2. Formally, we define it as $\mathcal{D}/c_0$, where $\mathcal{D} \subset \ell^{\infty}$ is exactly the set of all sequences $(\Pr(D_n \models \varphi))_{n\in\mathbb{N}}$, and c_0 is the set of sequences converging to 0. In particular, if a sequence of random structures obeys the FO 0–1 law, then its FO complexity is $\{0,1\}$ (for brevity, we identify a constant sequence with its element), and, if it obeys the FO convergence law, then its FO complexity is exactly the set of all limits of probabilities of FO sentences. Note that the above mentioned result of Larrauri, Müller, and Noy [18] guarantees that the closure of the FO complexity of $G(n,c/n)$ consists of finitely many segments. However, it is not hard to see that there are binomial random structures such that their FO complexities are even not totally bounded. In particular, this is the case for the binomial random graph $G(n,n^{-\alpha})$ for rational $\alpha < 1$ as we show in Section 2.2. It is also possible to define $p(n)$ in a way such that the FO complexity of $G(n,p(n))$ spans an infinite-dimensional subspace as well, but the complexity is totally bounded.

Finally, we consider the random graph

$$G(n \mid \varphi) := D^2(n \mid \varphi \wedge \text{symmetric} \wedge \text{anti-reflexive})$$

chosen uniformly at random from the set of all undirected graphs without loops that satisfy a given FO axiom φ. This model generalises the well-studied binomial random graph $G(n,1/2)$, random regular graphs [9, 36], random union of disjoint cliques [8], and random permutations [16]. For every complexity class, we show the existence of the respective axiom φ: the FO complexity of $G(n\mid\varphi)$ may be trivial (i.e. the FO 0–1 law holds), may span a 1-dimensional subspace of ℓ^∞/c_0 (the FO convergence law holds), may span a k-dimensional subspace for every positive integer k, may be totally bounded but span an infinite-dimensional subspace, and may not be totally bounded. These examples are given in Sections 2.3, 5.

Note that the FO 0–1 law holds for $G(n\mid\varphi)$ whenever φ is true on $G(n,1/2)$ with probability bounded away from 0. The latter may only happen when probability $\Pr(G(n,1/2)\models\varphi)$ approaches 1. Since the FO almost sure theory of $G(n,1/2)$ can be axiomatised by *extension axioms* (see, e.g., [30]) and probability that $G(n,1/2)$ does not satisfy the k-th extension axiom is at most $n^k(1-2^{-k})^{n-k}$, it is easy to see that the problem of determining whether $\Pr(G(n,1/2)\models\varphi)$ approaches 1 is decidable. Nevertheless, we show in Section 4 that the problem of determining, for an arbitrary input φ, whether $G(n\mid\varphi)$ obeys the FO 0–1 law is even not recursively enumerable.

2 FO COMPLEXITY

In Section 2.1, we define *the FO complexity* of a sequence of random structures that generalises the FO 0–1 law and the FO convergence law. After that, we show a strict hierarchy of FO complexity classes of binomial random graphs (Section 2.2) and conditional random graphs subject to FO sentences (Section 2.3). The most essential part of the main theorem in Section 2.3 that asserts the existence of a FO φ such that the complexity of $G(n\mid\varphi)$ is totally bounded and infinite dimensional is proved in Section 5. This proof develops a method of constructing FO sentences that define isomorphism classes of certain asymmetric graphs and may be of its own interest.

2.1 Definition of complexity and hierarchy of random structures

Let us recall necessary definitions. The Banach space ℓ^∞ is the linear space of all bounded sequences of real numbers $x=(x_1,x_2,\dots)$ with the norm $\|x\|=\sup_{n\in\mathbb{N}}|x_n|$. The Banach space c_0 is a subspace of ℓ^∞, which consists of all vectors x such that $\lim_{n\to\infty}|x_n|=0$. The Banach space c is a subspace of ℓ^∞, which consists of all vectors x such that $\lim_{n\to\infty}|x_n|$ exists. Norms on c_0 and c are induced from ℓ^∞. The Banach space ℓ^∞/c_0 is a quotient space, which consists of classes $x+c_0$ with the norm $\limsup_{n\to\infty}|x_n|$. The canonical projection $\pi:\ell^\infty\to\ell^\infty/c_0$ maps each $x\in\ell^\infty$ to $x+c_0\in\ell^\infty/c_0$. We denote by X/c_0 the image of a subset $X\subset\ell^\infty$ under π.

Definition 2.1. Let D_n, $n\in\mathbb{N}$, be a sequence of random n-structures. The *FO complexity* $\mathrm{FOC}(D_n)$ of D_n is the set $\mathcal{D}/c_0$, where $\mathcal{D}\subset\ell^\infty$ is the set of all sequences $(\Pr(D_n\models\varphi))_{n\in\mathbb{N}}$ over all FO sentences φ.

Due to the next straightforward proposition (we give a proof for the sake of completeness), the validity of the FO 0–1 law or the validity of the FO convergence law are the cases of the smallest FO complexities. For brevity, for any real λ, we denote vectors $(\lambda,\lambda,\dots)\in\ell^\infty$ and $\pi(\lambda,\lambda,\dots)\in\ell^\infty/c_0$ by λ. Note that $c/c_0\cong\mathbb{R}$, and there exists an isomorphism that maps $\lambda\in c/c_0$ to $\lambda\in\mathbb{R}$. Therefore, we identify any subset of c/c_0 with the set of respective numbers in $\mathbb{R}$.

Claim 2.2. *Let D_n, $n\in\mathbb{N}$, be a sequence of random structures.*

(i) *D_n satisfies the FO 0–1 law iff* $\mathrm{FOC}(D_n)=\{0,1\}$.
(ii) *D_n satisfies the FO convergence law iff* $\mathrm{FOC}(D_n)\subset\mathbb{R}$.
(iii) *The set of limits of sequences $(\Pr(D_n\models\varphi))_{n\in\mathbb{N}}$ coincides with* $\mathrm{FOC}(D_n)\cap\mathbb{R}$.

Proof. Let $x\in\ell^\infty$. From the definition, $\pi(x)-\lambda=0$ is equivalent to $\lim_{n\to\infty}|x_n-\lambda|=0$, i.e. $\lim_{n\to\infty}x_n=\lambda$. Since there is an isomorphism between c/c_0 and $\mathbb{R}$ that maps any $\pi(x)$, $x\in c$, to $\lim_{n\to\infty}x_n$, we have (iii). Therefore, $\mathrm{FOC}(D_n)\subset c/c_0$ is equivalent to the fact that each sequence $(\Pr(D_n\models\varphi))_{n\in\mathbb{N}}$ converges, i.e. we have (ii). Similarly, we get (i). □

Examples of random structures with these smallest FO complexities are well known: (1) $\mathrm{FOC}(D_n)=\{0,1\}$ for the uniformly random structure $D_n=D^\sigma(n,1/2,\dots,1/2)$, as well as for the uniformly chosen random graph $D_n=G(n,1/2)$ [5, 7]; (2) $\mathrm{FOC}(D_n)=\{0,1\}$ for the binomial random graph $D_n=G(n,p)$ with $pn^\alpha\to\infty$ for all $\alpha>0$ [31] or $p=n^{-\alpha}$, where α is either irrational or bigger than 1 and not equal $1+1/m$ for any positive integer m [29]; (3) $\mathrm{FOC}(D_n)\subset\mathbb{R}$ for the binomial random graph $D_n=G(n,n^{-\alpha})$ for $\alpha=1+\frac{1}{m}$ [29]; (4) $\mathrm{FOC}(D_n)\subset\mathbb{R}$ for the binomial random graph $D_n=G(n,\frac{c}{n})$ for $c>0$ [21], and the closure of $\mathrm{FOC}(D_n)$ consists of finitely many segments [18]. In the next section, we present random structures with d-dimensional, infinite dimensional but totally bounded, as well as not totally bounded FO complexities.

2.2 Complexity of $G(n,p)$

We first show that it is possible to construct very sparse binomial random graphs (consisting of only isolated vertices and isolated edges whp) $G(n,p)$ with all the properties of FO complexities mentioned in Section 1. However, the respective sequences $p=p(n)$ are quite artificial and far from being 'smooth'. So, later in this section we show that the FO complexity of $G(n,n^{-\alpha})$ for rational $\alpha\in(0,1)$ is not totally bounded, and that all the properties are achievable by $G(n\mid\varphi)$ for appropriately chosen FO sentences φ. We shall use the following technical claim that follows from the fact that whp $G(n,p=o(n^{-3/2}))$ consists of isolated vertices and isolated edges, and the number of isolated edges can be approximated by Poisson random variables $\mathrm{Pois}(\lambda_n=p\binom{n}{2})$ (see, e.g., [12, 30]).

Claim 2.3. *Let $p=o(n^{-3/2})$, $\lambda_n=p(n)\binom{n}{2}$, and, for every $k\in\mathbb{Z}_{\geq 0}$, $x_k=\pi\left((\frac{\lambda_n^k}{k!}e^{-\lambda_n})_{n\in\mathbb{N}}\right)$. Then $\mathrm{FOC}(G(n,p))$ is a union of the set X of all finite sums of vectors x_k and the set $1-X:=\{1-x,\ x\in X\}$.*

Proof. Let a FO sentence $\varphi_{\geq k}$ express the property of being a disjoint union of at least k edges. The sentence $\varphi_k=\varphi_{\geq k}\wedge\neg\varphi_{\geq k+1}$ expresses the property of being a disjoint union of exactly k edges. Each FO sentence $\psi\wedge\varphi_{\geq 0}$ is tautologically equivalent either to $\varphi_{i_1}\vee\dots\vee\varphi_{i_s}$ or to $\varphi_{i_1}\vee\dots\vee\varphi_{i_s}\vee\varphi_{\geq i_{s+1}}$ for some non-negative integers $i_1<\dots<i_s<i_{s+1}$. Since $p=o(n^{-\frac{3}{2}})$, the sequence of probabilities $(\Pr(G(n,p)\models\varphi_{\geq 0}))_{n\in\mathbb{N}}$ converges to one (see [30,

Theorem 3.6.2]). Hence, $\Pr(G(n,p) \models \psi) - \Pr(G(n,p) \models \psi \wedge \varphi_{\geq 0})$ converges to zero.

Let us show that for φ_i, the equality $\pi(\Pr(G(n,p) \models \varphi_i)) = x_i$ holds. It is easier to compute the probability that the graph has exactly i edges without any restriction on overlapping edges. Luckily, these two probabilities are asymptotically equal. Let a FO sentence $\hat{\varphi}_i$ express the property of containing exactly i edges (not necessary disjoint). So, the FO sentence $\hat{\varphi}_i \wedge \varphi_{\geq 0}$ is tautologically equivalent to φ_i. Since $\varphi_{\geq 0}$ is true whp, we get

$$\begin{aligned}\pi(\Pr(G(n,p) \models \varphi_i)) &= \pi(\Pr(G(n,p) \models \hat{\varphi}_i \wedge \varphi_{\geq 0}))\\ &= \pi(\Pr(G(n,p) \models \hat{\varphi}_i)).\end{aligned}$$

Now, it is sufficient to prove that $\pi(\Pr(G(n,p) \models \hat{\varphi}_i)) = x_i$.

$$\begin{aligned}\Pr(G(n,p) \models \hat{\varphi}_i) &= \binom{\binom{n}{2}}{i} p^i (1-p)^{\binom{n}{2}-i}\\ &= (1+o(1))\frac{1}{i!}\binom{n}{2}^i p^i e^{\binom{n}{2}\ln(1-p)}(1-p)^{-i} = \frac{\lambda_n^i}{i!}e^{-\lambda_n} + o(1)\end{aligned}$$

as needed.

Since φ_i contradicts φ_j for each $j \neq i$, $\pi(\Pr(G(n,p) \models \varphi_{i_1} \vee \cdots \vee \varphi_{i_s})) = x_{i_1} + \cdots + x_{i_s}$. Therefore, we get $X \subset \mathrm{FOC}(G(n,p(n)))$. For the sentence $\psi = \varphi_{i_1} \vee \cdots \vee \varphi_{i_s} \vee \varphi_{\geq i_{s+1}}$, we consider a sentence $\neg\psi$, for which $\pi(\Pr(G(n,p) \models \neg\psi)) = 1 - \pi(\Pr(G(n,p) \models \psi))$. Note that the sentence $\neg\psi \wedge \varphi_{\geq 0}$ is equivalent to $\bigvee_{j<i_{s+1}, j\neq i_1,\dots,i_s} \varphi_j$. So, $\pi(\Pr(G(n,p) \models \psi)) \in 1-X$, and we have $\mathrm{FOC}(G(n,p)) = X \cup (1-X)$, completing the proof. □

Theorem 2.4. *For $p = p(n)$, we let $G_n \sim G(n,p)$.*

- (ii) *For each $d \geq 1$, there is a sequence $p(n) \in [0,1]$ such that* $\mathrm{FOC}(G_n)$ *spans a d-dimensional subspace of ℓ^∞/c_0.*
- (ii) *There is a sequence $p(n) \in [0,1]$ such that* $\mathrm{FOC}(G_n)$ *is totally bounded but spans an infinite-dimensional subspace of ℓ^∞/c_0.*
- (iii) *There is a sequence $p(n) \in [0,1]$ such that* $\mathrm{FOC}(G_n)$ *is not totally bounded.*

Proof. To prove (i), consider $p = \lambda_n/\binom{n}{2}$, where λ_n equals n modulo d. The sequence $(\frac{\lambda_n^k}{k!}e^{-\lambda_n})_{n\in\mathbb{N}}$ is d-periodic, i.e. its $(n+d)$-th element equals the n-th element. The subspace L_d of d-periodic sequences in ℓ^∞ is d-dimensional. Consider the following basis in L_d: let e_r be the vector with ones on $(dt+r)$-th positions, $t \in \mathbb{Z}_+$, and zeros on all others. Let us prove that the system of vectors $f_k = (\frac{\lambda_n^k}{k!}e^{-\lambda_n})_{n\in\mathbb{N}}$ for $0 \leq k \leq d-1$ is also a basis in L_d. The vector f_k equals $\sum_{r=0}^{d-1} \frac{r^k}{k!}e^{-r}e_r$, and

$$\begin{aligned}\det\left(\left(\frac{i^j}{j!}e^{-i}\right)_{0\leq i,j\leq d-1}\right) &= \det\left((i^j)_{0\leq i,j\leq d-1}\right)\prod_i e^{-i}\prod_j \frac{1}{j!}\\ &= \prod_{i<i'}(i'-i)\prod_i e^{-i}\prod_j \frac{1}{j!} \neq 0.\end{aligned}$$

Then, f_k, $0 \leq k \leq d-1$, is a basis in L_d. Since $x_k = \pi(f_k)$ for all k, we have that $\langle x_k, k \geq 0\rangle$ coincides with the subspace $\pi(L_d)$ in ℓ^∞/c_0. The intersection of L_d and c_0 is trivial because each d-periodic vector in c_0 is zero. Therefore, the space $\pi(L_d)$ is also d-dimensional. Note that $\langle X\rangle = \pi(L_d)$ and $1 \in \pi(L_d)$ implying that $\langle 1-X\rangle = \pi(L_d)$. Since $f_k \in X$ for all $0 \leq k \leq d-1$, by Claim 2.3, we get that $\langle \mathrm{FOC}(G(n,p))\rangle = \pi(L_d)$ completing the proof.

To prove (ii), consider $p = \frac{2\lambda_n}{n(n-1)}$, where λ_n equals $\frac{1}{r}$ if n is divisible by 2^r, but not divisible by 2^{r+1}. We will denote $r(n)$ the maximum r such that n is divisible by 2^r. Thus, $\lambda_n = \frac{1}{r(n)}$. Let e_r be the vector with ones on $2^r(2t+1)$-th positions, and zeros on all others. The vector $(\frac{\lambda_n^k}{k!}e^{-\lambda_n})_{n\in\mathbb{N}}$ equals $\sum_{r=0}^{\infty}\frac{1}{r^k k!}e^{-\frac{1}{r}}e_r$. Vectors $\pi(e_r)$ are nonzero and linearly independent. Therefore, $x_k = \sum_{r=0}^{\infty}\frac{1}{r^k k!}e^{-\frac{1}{r}}\pi(e_r)$. Similarly, as for (i), x_k are linearly independent. So, the set X spans an infinite-dimensional subspace.

Let us show that the sum $\sum_{k=0}^{\infty}\|x_k\|$ converges.

$$\|x_k\| = \limsup_n \frac{1}{(r(n))^k k!}e^{-\frac{1}{r(n)}} \leq \limsup_n \frac{1}{(r(n))^k k!} = \frac{1}{k!}.$$

Since the sum $\sum_{k=0}^{\infty}\frac{1}{k!}$ converges, we get the convergence of the considered sum. For each positive ε, we can choose an integer N such that $\sum_{k=N+1}^{\infty}\|x_k\| < \varepsilon$. Then, for each vector $v = x_{i_1}+x_{i_2}+\cdots+x_{i_s} \in X$ consider a vector $v' = x_{i_1}+x_{i_2}+\cdots+x_{i_{s'}}$, where $i_{s'}$ is the greatest number among i_j such that $i_j \leq N$. Since $v - v'$ is expressible as a sum of vectors x_k with $k > N$, we get $\|v - v'\| \leq \varepsilon$. Therefore, the set X can be covered by finitely many balls of radius ε because there is finitely many sums of vectors x_k with $k \leq N$. Similarly for $1-X$. Hence, $X \cup (1-X)$ is totally bounded.

To prove (iii), consider $p = 2\lambda_n/(n(n-1))$, where λ_n equals $m(r(n))$, and $m(r)$, $r \in \mathbb{Z}_+$, is defined in such a way that, for some $k(r)$, vectors $y_r = x_0 + \cdots + x_{k(r)}$ are at distances at least $\frac{1}{3}$ from each other. Let us construct such $m(r)$ and $k(r)$. For every non-negative integer k, consider a function $g_k : \mathbb{R} \to \mathbb{R}$ defined as follows: $g_k(\lambda) = (1+\lambda+\cdots+\lambda^k/k!)e^{-\lambda}$. This sequence of functions satisfies two properties: (a) $\lim_{\lambda\to\infty} g_k(\lambda) = 0$ for every fixed k; (b) $\lim_{k\to\infty} g_k(\lambda) = 1$ for every fixed λ. Let $m(0) = 0$ and $k(0) = 0$. For each $r \geq 1$, we define $m(r)$ as the least integer M such that $\forall m \geq M : g_{k(r-1)}(m) \leq \frac{1}{3}$. Such an integer M exists due to the property (a). Next, we define $k(r)$ as the least integer K such that $\forall k \geq K : g_k(m(r)) \geq \frac{2}{3}$. Such an integer K exists due to the property (b). The vector y_r equals $\pi\left((g_{k(r)}(m(r(n))))_{n\in\mathbb{N}}\right)$. So, y_r and y_s, $r \geq s+1$, are at distance

$$\begin{aligned}\|y_r - y_s\| &= \limsup_n \left|g_{k(r)}(m(r(n))) - g_{k(s)}(m(r(n)))\right|\\ &\geq g_{k(r)}(m(s+1)) - g_{k(s)}(m(s+1)) \geq \frac{2}{3} - \frac{1}{3} = \frac{1}{3}.\end{aligned}$$

Since there is a sequence of vectors $y_r \in X$ such that $\|y_r - y_s\| \geq \frac{1}{3}$ for each pair of distinct positive integers r and s, the set $X \cup (1-X)$ is not totally bounded. □

Theorem 2.6 stated below claims that $G(n, n^{-\alpha})$ has a not totally bounded complexity for every rational $\alpha \in (0,1)$. We derive it using a construction of a FO sentence introduced by Shelah and Spencer in [5] to disprove the convergence law.

Lemma 2.5 (Shelah, Spencer [29]). *Let $G_n \sim G(n, n^{-\alpha})$. For every integer $d \geq 100$, there exists a FO sentence φ_d such that*

(i) *if* $\log^* n \equiv \lfloor \frac{d}{4} \rfloor (\bmod\ d)$ *then* $G_n \models \varphi_d$ *whp;*
(ii) *if* $\log^* n \equiv \lfloor \frac{3d}{4} \rfloor (\bmod\ d)$ *then* $G_n \models \neg\varphi_d$ *whp,*

where $\log^*$ *denotes the iterated logarithm, i.e. the number of times the logarithm need to be applied to the number to make it one or less.*

In the original paper [29] Lemma 2.5 was formulated only for $d = 100$, but literally the same proof works for any $d \geq 100$ (the lower bound 100 could be sufficiently improved, though it is not important for us).

Theorem 2.6. *For any rational* $\alpha \in (0, 1)$, $\mathrm{FOC}(G(n, n^{-\alpha}))$ *is not totally bounded.*

Proof. For every prime number $p \geq 100$, let us consider the sentence φ_p whose existence is claimed by Lemma 2.5. Let $x_p(n)$ be the probability $\Pr(G(n, n^{-\alpha}) \models \varphi_p)$. Then, vectors $x_p = \pi\left((x_p(n))_{n\in\mathbb{N}}\right)$ are in $\mathrm{FOC}(G(n, n^{-\alpha}))$. Let $(y_p(n))_{n\in\mathbb{N}} \in \ell^\infty$ be the sequence such that

(i) if $\log^* n \equiv \lfloor \frac{p}{4} \rfloor (\bmod\ p)$ then $y_p(n) = 1$;
(ii) if $\log^* n \equiv \lfloor \frac{3p}{4} \rfloor (\bmod\ p)$ then $y_p(n) = 0$;
(iii) $y_p(n) = x_p(n)$, otherwise.

The sequence $x_p(n) - y_p(n)$ converges to zero because, for each n such that $\log^* n \equiv \lfloor \frac{p}{4} \rfloor (\bmod\ p)$ or $\log^* n \equiv \lfloor \frac{3p}{4} \rfloor (\bmod\ p)$, we have the required convergence due to the properties of φ_p given by Lemma 2.5, and $x_p(n) - y_p(n) = 0$ for all other n.

Thus, vectors x_p and $\pi\left((y_p(n))_{n\in\mathbb{N}}\right)$ are equal. By the Chinese remainder theorem, for each pair of distinct primes p and q there exists an integer number m such that $m \equiv \lfloor \frac{p}{4} \rfloor (\bmod\ p)$ and $m \equiv \lfloor \frac{3q}{4} \rfloor (\bmod\ q)$. Therefore, for each n such that $\log^* n \equiv m (\bmod\ pq)$, we have $y_p(n) - y_q(n) = 1$. Since there are infinitely many n such that $\log^* n \equiv m (\bmod\ pq)$, we have $\|x_p - x_q\| \geq 1$. Hence, we have infinitely many vectors in the set $\mathrm{FOC}(G(n, n^{-\alpha}))$, which are at distances at least 1 from each other, i.e. $\mathrm{FOC}(G(n, n^{-\alpha}))$ is not totally bounded. □

So, indeed, $\mathrm{FOC}(G(n, n^{-\alpha}))$ is not totally bounded when $\alpha \in (0, 1) \cap \mathbb{Q}$. We are not able to present a "nice" p so that $\mathrm{FOC}(G(n, p))$ is either d-dimensional, $d > 1$, or infinite-dimensional and totally bounded. However, this appears to be possible for $G(n \mid \varphi)$.

2.3 Complexity of $G(n \mid \varphi)$

First of all, note that for any FO sentence φ, if $\liminf_{n\to\infty} \Pr(G(n, \frac{1}{2}) \models \varphi) > 0$ (it actually may only happen when the limit is 1), then $\mathrm{FOC}(G(n \mid \varphi)) = \{0, 1\}$ due to the FO 0–1 law for $G(n, \frac{1}{2})$. Moreover, there is a FO sentence φ such that $G(n \mid \varphi)$ obeys the FO convergence law but not the FO 0–1 law. For example, consider φ which expresses the property of consisting of isolated vertices and exactly one connected component of size 3. Then, probability of containing a triangle converges to $\frac{1}{4}$.

Theorem 2.7. *For a FO sentence φ, we let $G_n \sim G(n \mid \varphi)$.*

(i) *There is a FO sentence φ such that* $\mathrm{FOC}(G_n)$ *is a dense subset of* $[0, 1]$.
(ii) *For each $d \geq 1$, there is a FO sentence φ such that* $\mathrm{FOC}(G_n)$ *spans a d-dimensional subspace of* ℓ^∞/c_0.
(iii) *There is a FO sentence φ such that* $\mathrm{FOC}(G_n)$ *is totally bounded but spans an infinite-dimensional subspace of* ℓ^∞/c_0.
(iv) *There is a FO sentence φ such that* $\mathrm{FOC}(G_n)$ *is not totally bounded.*

We postpone the proof of (iii) to Section 5: it is long enough to interrupt the flow of the paper and it requires an additional background that we outline in the beginning of Section 5.

Proof of parts (i), (ii) and (iv) of Theorem 2.7. **To prove (i),** consider a FO sentence φ which expresses the property of being 2-regular. For the respective random graph $G(n \mid \varphi)$, the FO convergence law was proven by Lynch in [20]. Then, $\mathrm{FOC}(G(n \mid \varphi))$ is a subset of $[0, 1]$. To prove that this subset is dense, we refer to the result proven by Bollobás and Wormald [1, 37, 38].

Lemma 2.8 (Bollobás, Wormald [1, 37, 38]). *Fix an integer $d \geq 2$ and $C > 0$. In random uniform d-regular graphs on $[n]$, the vectors of numbers of cycles of length $\ell \leq C$ converge in distribution to a vector of independent Poisson random variables* $\mathrm{Pois}\left((d-1)^\ell/(2\ell)\right)$.

Let a FO sentence ψ_ℓ express the property of containing a cycle of length ℓ. For a $\{0, 1\}$-word W of length w, let ψ_W be a conjunction of sentences ψ_ℓ if $W(\ell - 2) = 1$, and $\neg\psi_\ell$ if $W(\ell - 2) = 0$. For each pair of distinct words W and W' of length w, ψ_W contradicts $\psi_{W'}$. Also, the disjunction of ψ_W over all words W of length w is a tautology. Therefore, for an arbitrary labelling $W_1, W_2, \ldots, W_{2^w}$ of all such words, we have that

$$q_s := \lim_{n\to\infty} \Pr\left(G(n \mid \varphi) \models \bigvee_{i=1}^{s} \psi_{W_i}\right) = \sum_{i=1}^{s} \lim_{n\to\infty} \Pr(G(n \mid \varphi) \models \psi_{W_i}).$$

Note that $q_0 = 0$ and $q_{2^w} = 1$. Also, from Lemma 2.8, we have that

$$q_s - q_{s+1} = \lim_{n\to\infty} \Pr(G(n \mid \varphi) \models \psi_{W_s}) =$$

$$= \prod_{W_s(\ell-2)=1} (1 - e^{-\frac{1}{2\ell}}) \prod_{W_s(\ell-2)=0} e^{-\frac{1}{2\ell}} \leq \prod_{\ell=3}^{w+2} e^{-\frac{1}{2\ell}} < \frac{e^{\frac{3}{2}}}{\sqrt{w+2}}.$$

Therefore, q_s is an increasing sequence of numbers in $[0, 1]$ such that $q_0 = 0$, $q_{2^w} = 1$ and $q_s - q_{s-1} < e^{\frac{3}{2}}/\sqrt{w+2}$. Hence, for each number $x \in [0, 1]$, there is an element of this sequence such that $|x - q_s| < \frac{1}{2} e^{\frac{3}{2}}/\sqrt{w+2}$. Since q_s are limiting probabilities for FO sentences, and w can be chosen arbitrary large, $\mathrm{FOC}(G(n \mid \varphi))$ is a dense subset of $[0, 1]$ as needed.

To prove (ii), consider a FO sentence φ which expresses the property of being a disjoint union of d-cliques and at most one r-clique for some $0 \leq r < d$. Note that, for each n, this property defines a single isomorphism class. Fix a FO sentence ψ. Since for any graphs A, B, there exists $m_0 \in \mathbb{N}$ such that, for any $m \geq m_0$, graphs $m_0 A \sqcup B$ and $mA \sqcup B$ are not distinguishable by ψ (see [2]) we get that there exists t_0 such that for all $t \geq t_0$ graphs on $[dt_0 + r]$ and $[dt + r]$ that satisfy φ are not distinguishable by ψ. Then, the sequence $(\Pr(G(n \mid \varphi) \models \psi))_{n\in\mathbb{N}}$ consists only of zeros and ones and is d-periodic for n large enough. Therefore, $\mathrm{FOC}(G(n \mid \varphi))$ is a set of projections of d-periodic sequences of ones and zeros, i.e. is contained in the d-dimensional subspace $\pi(L_d) \subset \ell^\infty/c_0$, where L_d is the d-dimensional space of all d-periodic sequences in ℓ^∞.

For each $0 \leq r < d$ consider a FO sentence ψ_r which expresses the property of containing an isolated r-clique. The n-th element of the sequence $\Pr(G(n \mid \varphi) \models \psi_r)$ equals one if $n \equiv r (\mathrm{mod}\ d)$, and zero otherwise. Projections of $(\Pr(G(n \mid \varphi) \models \psi_r)_{n\in\mathbb{N}}$, $0 \le r < d$, generate the space $\pi(L_d)$. Thus, $\mathrm{FOC}(G(n \mid \varphi))$ does not span a $(d-1)$-dimensional subspace of ℓ^∞/c_0.

To prove (iv), we need the following definition.

Definition 2.9. Let G and H be two graphs. *The Cartesian product* $G\square H$ is the graph on $V(G)\times V(H)$ with adjacency relation $(u,v)\sim(u',v') \Leftrightarrow ((u\sim u')\wedge(v=v'))\vee((u=u')\wedge(v\sim v'))$. In other words, every edge of the Cartesian product belongs either to an induced subgraph $G_v \cong G$ on $\{(u,v) \mid u\in V(G)\}$ for some $v\in V(H)$ or to an induced subgraph $H_u \cong H$ on $\{(u,v)\mid v\in V(H)\}$ for some $u \in V(G)$.

The property of being isomorphic to $K_s\square K_t$, for some $s,t>0$, is FO. Let φ be a FO sentence that describes this property. Let, for $n\in\mathbb{Z}_{>0}$, $D(n)$ be the set of all divisors of n. Let $\mu_d = 2$, if $d=\sqrt{n}$, and $\mu_d = 1$, otherwise. For each $d\in D(n)$ there are $\mu_d d!\left(\frac{n}{d}\right)!$ automorphisms of the graph $K_d\square K_{\frac{n}{d}}$. Therefore, there are $\frac{n!}{\mu_d d!(\frac{n}{d})!}$ graphs on $[n]$ isomorphic to $K_d\square K_{\frac{n}{d}}$. The probability that $G(n\mid\varphi)$ is isomorphic to $K_d\square K_{\frac{n}{d}}$ equals

$$\frac{n!/\mu_d}{d!\left(\frac{n}{d}\right)!}\Bigg/\left(\sum_{d'\in D(n)}\frac{n!}{2d'!\left(\frac{n}{d'}\right)!}\right)=\frac{1/\mu_d}{d!\left(\frac{n}{d}\right)!}\Bigg/\left(\sum_{d'\in D(n)}\frac{1}{2d'!\left(\frac{n}{d'}\right)!}\right)$$

where we have 2 in the denominator of the normalisation factor instead of $\mu_{d'}$ since we count twice every graph $K_{d'}\square K_{n/d'}$ when $d'\neq\sqrt{n}$.

Let ψ_d be a FO sentence which expresses the property of containing an inclusion-maximal clique of size d, i.e. a clique of size d which is not included in any clique of size $d+1$. If a graph on $[n]$ satisfies φ then it is isomorphic to $K_d\square K_{\frac{n}{d}}$ if and only if it satisfies ψ_d. Consider $d_1<d_2$, let us prove that, for each $\varepsilon>0$ there are infinitely many numbers n such that $\Pr(G(n\mid\varphi)\models\psi_{d_1})-\Pr(G(n\mid\varphi)\models\psi_{d_2})>1-\varepsilon$. Let $n=d_1p$, where p is a prime number bigger than d_2. Therefore, n is not divisible by d_2, and $\Pr(G(n\mid\varphi)\models\psi_{d_2})=0$. Since $d_1<d_2<p$, each divisor m of n such that $m\le\sqrt{n}$ cannot be divisible by $p>\sqrt{n}$. Hence, such divisors are divisors of d_1. Thus, we can estimate the normalisation factor for the probability $\Pr(G(n\mid\varphi)\models\psi_{d_1})$ in the following way:

$$\sum_{d'\in D(n)}\frac{1/2}{d'!\left(\frac{n}{d'}\right)!}=\sum_{d'\in D(d_1)}\frac{1}{d'!\left(\frac{n}{d'}\right)!}\le\frac{1}{d_1!\left(\frac{n}{d_1}\right)!}+\frac{|D(d_1)|-1}{\left(\frac{2n}{d_1}\right)!}\le$$
$$\frac{1}{d_1!\left(\frac{n}{d_1}\right)!}\left(1+\frac{(d_1+1)!p!}{(2p)!}\right)\le\frac{1}{d_1!\left(\frac{n}{d_1}\right)!}\left(1+\frac{p!p!}{(2p)!}\right)\le\frac{1+2^{-p}}{d_1!\left(\frac{n}{d_1}\right)!}.$$

Therefore, for all primes p such that $(1+2^{-p})^{-1}>1-\varepsilon$, we have $\Pr(G(n\mid\varphi)\models\psi_{d_1})>1-\varepsilon$. So, each pair ψ_{d_1} and ψ_{d_2} defines a pair of vectors in $\mathrm{FOC}(G(n\mid\varphi))$ at distance at least 1, and $\mathrm{FOC}(G(n\mid\varphi))$ is not totally bounded. □

3 STOCHASTIC FO REDUCTION

In this section we define a *stochastic FO reduction* and describe its useful properties (in Section 3.1). Then we show its effectiveness by using it to derive certain logical limit laws for dense (in Section 3.2) and sparse (in Section 3.3) relational structures as well as to transfer higher FO complexities between random relational structures. Finally, in Section 3.4 we use the stochastic FO reduction to generalise the result of Larrauri, Müller, and Noy about the closure of FO complexity of binomial random d-hypergraphs with $p=c/n^{d-1}$ from undirected hypergraphs to directed hypergraphs for any possible way to choose an orientation of hyperedges.

3.1 Definition and main properties

Let σ,σ' be two signatures; $\mathcal{D}_n$ and $\mathcal{D}'_n$ be the sets of all finite structures on $[n]$ over σ and σ' respectively; $\mathcal{D}=\sqcup_{n\in\mathbb{N}}\mathcal{D}_n$ and $\mathcal{D}'=\sqcup_{n\in\mathbb{N}}\mathcal{D}'_n$; D_n, D'_n be random relational n-structures over σ,σ' respectively. Moreover, for any two FO sentences φ,φ' over σ,σ' respectively, let $\mathcal{D}(\varphi)\subset\mathcal{D}$ and $\mathcal{D}'(\varphi')\subset\mathcal{D}'$ be sets of all structures satisfying φ and φ' respectively. Finally, let us consider algebras $\mathcal{A}=\{\mathcal{D}(\varphi)\}$, $\mathcal{A}'=\{\mathcal{D}'(\varphi')\}$ (recalling that an *algebra* is closed under finite unions in contrast to a σ-algebra). For an $(\mathcal{A}\mid\mathcal{A}')$-measurable function $f:\mathcal{D}\to\mathcal{D}'$ and a FO sentence φ', we denote by $f^{-1}(\varphi')=:\varphi$ a FO sentence such that $\mathcal{D}(\varphi)=f^{-1}(\mathcal{D}'(\varphi'))$.

Definition 3.1. **A stochastic FO reduction** from $D'=(D'_n)_{n\in\mathbb{N}}$ to $D=(D_n)_{n\in\mathbb{N}}$ is an $(\mathcal{A}\mid\mathcal{A}')$-measurable function $f:\mathcal{D}\to\mathcal{D}'$ such that, for every $n\in\mathbb{N}$, f maps n-structures to n-structures, and $\lim_{n\to\infty}|\Pr(D_n\models f^{-1}(\varphi'))-\Pr(D'_n\models\varphi')|=0$ for every FO sentence φ' over σ'.

If there is a stochastic FO reduction from D' to D, we say that D' is *reducible* to D (or sometimes we say that D'_n is reducible to D_n meaning of course a reduction of the entire sequences) and denote it as $D'\le D$ (or $D'_n\le D_n$). Let us first observe that the basic property of being a preorder, that holds for reductions in the computational complexity theory, holds for our reduction as well.

Claim 3.2. *The stochastic FO reduction relation* $\le$ *is a preorder.*

The proof is straightforward, it can be also found in the extend version of the paper [**?**].

Next, we show key property of stochastic FO reductions which allow us to transfer FO complexities between different random relational structures.

Claim 3.3. *Suppose* $D'_n\le D_n$. *Then,* $\mathrm{FOC}(D'_n)\subseteq\mathrm{FOC}(D_n)$.

Proof. Let $f:\mathcal{D}\to\mathcal{D}'$ be a reduction from D'_n to D_n. Consider a vector $v\in\mathrm{FOC}(D'_n)$. There is a FO sentence φ in the signature of D'_n such that $v=(\Pr(D'_n\models\varphi))_{n\in\mathbb{N}}+c_0$. Since f is the stochastic FO reduction, there is a FO sentence $f^{-1}(\varphi)$ in the signature of D_n such that $\lim_{n\to\infty}|\Pr(D_n\models f^{-1}(\varphi))-\Pr(D'_n\models\varphi)|=0$. Therefore, $v=(\Pr(D'_n\models\varphi))_{n\in\mathbb{N}}+c_0=(\Pr(D_n\models f^{-1}(\varphi)))_{n\in\mathbb{N}}+c_0\in\mathrm{FOC}(D_n)$. □

Corollary 3.4. *Suppose* $D'_n\le D_n$.

(i) *If* D_n *obeys the FO 0–1 law, then* D'_n *obeys the FO 0–1 law as well.*
(ii) *If* D_n *obeys the FO convergence law, then* D'_n *obeys the FO convergence law as well.*

Proof. By Claim 2.2 and Claim 3.3, we have that $\mathrm{FOC}(D'_n)\subseteq\mathrm{FOC}(D_n)\subseteq\{0,1\}$, for the case (i), and $\mathrm{FOC}(D'_n)\subseteq\mathrm{FOC}(D_n)\subseteq\mathbb{R}$, for the case (ii). Then, by Claim 2.2, we obtain the assertion of the corollary. □

3.2 Application to FO zero-one laws for dense structures

Let us now use stochastic FO reductions to transfer FO 0-1 laws. We consider the signature $\{=, \to\}$, where $\to$ has arity 2, and denote by $\vec{G}(n, p)$ the binomial random directed graph without loops on the set of vertices $[n]$, i.e. every directed edge (out of the set of $n(n-1)$ possible edges) appears independently of the others with probability p.

Theorem 3.5. *Let $p \in (0, 1)$ (not necessarily a constant). There are stochastic FO reductions $G(n, p^2) \preceq \vec{G}(n, p) \preceq D^\sigma(n, p)$, where $\sigma = \{=, P\}$. In particular, for a constant $p \in (0, 1)$, $\vec{G}(n, p)$ obeys the FO 0–1 law.*

Proof. Define the reduction from $\vec{G}(n, p)$ to $D^\sigma(n, p)$ as the function f which deletes loops from a directed graph. This mapping can be "defined" by the FO formula $\psi_\to(x, y) = P(x, y) \wedge \neg(x = y)$. So, for each FO sentence φ over signature $\{=, \to\}$, the FO sentence $f^{-1}(\varphi)$ is constructed from φ by replacing each $x \to y$ by $\psi_\to(x, y)$. Note that $f(D^\sigma(n, p)) \overset{d}{=} \vec{G}(n, p)$. Thus, $\Pr(D^\sigma(n, p) \models f^{-1}(\varphi)) = \Pr(\vec{G}(n, p) \models \varphi)$ and f is indeed a stochastic FO reduction.

From Corollary 3.4 and the fact that $D^\sigma(n, p)$ obeys the FO 0–1 law, we have that $\vec{G}(n, p)$ obeys the FO 0–1 law as well.

Let g be a mapping from directed graphs without loops to undirected graphs which replaces pairs of directed edges (x, y) and (y, x) by an undirected edges $\{x, y\}$ and deletes all the other directed edges. The mapping g can be defined by the FO formula $\psi_\sim(x, y) = (x \to y) \wedge (y \to x)$. Then, as in the previous case, the FO sentence $g^{-1}(\varphi)$ is obtained by replacing all $x \sim y$ by $\psi_\sim$. Furthermore, $g(\vec{G}(n, p)) \overset{d}{=} G(n, p)$ because the presence of an undirected edge $\{x, y\}$ depends only on the presence of two directed edges (x, y) and (y, x), and the probability that both of them are presented is p^2. So, we have the equality $\Pr(\vec{G}(n, p) \models g^{-1}(\varphi)) = \Pr(G(n, p^2) \models \varphi)$ which finishes the proof. □

Let us stress once again that in [5] 0-1 law for the binomial random graph and for $D^\sigma(n, p)$ are proven separately. Due to Theorem 3.5, the validity of the FO 0-1 law for $D^\sigma(n, p)$ for all constant p implies its validity for $G(n, p)$ for all constant p as well.

Actually, Theorem 3.5 admits a generalisation to arbitrary signatures and distributions. We state it below, and then use for other particular reductions. It is straightforward that Theorem 3.5 follows immediately from this claim.

As above, we consider two relational signatures σ, σ' and respective sequences of random structures D, D'. For every $P \in \sigma'$ of arity a, assume that we are given with a FO formula ψ_P of arity a over σ. Let $f : \mathcal{D} \to \mathcal{D}'$ be defined as follows: for every $P \in \sigma'$ and every $X \in \mathcal{D}$, we set $f(X) \models P(x_1, \dots, x_a)$ if and only if $X \models \psi_P(x_1, \dots, x_a)$. We call f a *reduction defined by* $(\psi_P,\ P \in \sigma')$.

Claim 3.6. *Let D be a random structure over σ, and let f be a reduction defined by $(\psi_P,\ P \in \sigma')$. Letting $D' \overset{d}{=} f(D)$, we get that $D' \preceq D$ and f reduces D' to D.*

Proof. In order to see that f is $(\mathcal{A} \mid \mathcal{A}')$-measurable, it is sufficient to observe that, for every FO φ' over σ', $f^{-1}(\varphi')$ is obtained from φ' by replacing each P by ψ_P. Since the distribution of D' coincides with the distribution of $f(D)$, we have the equality $\lim_{n\to\infty} |\Pr(D \models f^{-1}(\varphi)) - \Pr(D' \models \varphi)| = 0$ which finishes the proof. □

A similar to f object which is called *a FO translation* appeared in [14] and was used for reductions between languages (for more details, we refer a reader to the book [13], where the concept of *FO reductions* is introduced and its properties are described).

We denote by $G_{loop}(n, p)$ the binomial random undirected graph which allows loops on the set of vertices $[n]$, i.e. every edge (out of the set of $\frac{n(n+1)}{2}$ possible edges) appears independently of the others with probability p. We consider this structure over the signature $\{=, \sim\}$. The proof by reduction of the FO 0–1 law for this random structure (with constant p) is not as straightforward as for $\vec{G}(n, p)$. Indeed, if we apply a reduction to the $D^\sigma(n, \sqrt{p})$ defined by the FO formula $P(x, y) \wedge P(y, x)$, then we get a random undirected graph which allows loops, but the probability of the presence of a loop is p, while a non-loop edge has the emergence probability $p^2 \neq p$.

Let us denote by $G'_{loop}(n, p, q)$ the random undirected graph which allows loops on the set of vertices $[n]$, but over the signature $\{=, \sim', L\}$, where $\sim'$ has arity 2 and L has arity 1. The predicate $\sim'$ expresses the presence of non-loop undirected edges, the predicate L expresses the presence of loops. The distribution of $G'_{loop}(n, p, q)$ is such that each edge appears independently with probability q if it is a loop, and p otherwise. To prove the FO 0–1 law for $G_{loop}(n, p)$, we prove the equivalence between this structure and $G'_{loop}(n, p, p)$. The next claim is a generalisation of this statement.

Since stochastic FO reductions define a preorder, they induce an equivalence relation on random relational structures: we call D_n and D'_n *equivalent* if $D_n \preceq D'_n$ and $D'_n \preceq D_n$. For our purposes, it is useful to have a "loopless" representative in every equivalence class, which is defined below.

Definition 3.7. Let $\sigma = \{=, P_1, \dots, P_s\}$ be a signature, where P_i is a predicate symbol of arity a_i. Let D_n be a random n-structure over the signature σ. The random structure D_n is called **loopless** if, for all $1 \le i \le s$ and $1 \le j < k \le a_i$, the following equality holds:

$$\Pr(D_n \models \forall x_1 \dots \forall x_{a_i}\ (x_j = x_k) \Rightarrow \neg P_i(x_1, \dots, x_{a_i})) = 1.$$

Claim 3.8. *For each random relational structure D_n over the signature σ, there is a loopless random relational structure D^*_n which is equivalent to D_n. Moreover, $G_{loop}(n, p)$ is equivalent to $G'_{loop}(n, p, p)$.*

The proof of Claim 3.8 is technical and based on a explicit construction of the distribution of D^*_n, stochastic FO reductions, and application of Claim 3.6. Its proof can be found in the extended version of the paper [?].

Let τ be a signature $\{=, P, L\}$ where P has arity 2 and L has arity 1. There is a stochastic FO reduction $G'_{loop}(n, p, q)$ to $D^\tau(n, \sqrt{p}, q)$ defined by the FO formulae $P(x, y) \wedge P(y, x) \wedge \neg(x = y)$ and $L(x)$. Hence, by Corollary 3.4, Claim 3.2 and Claim 3.8, we have

Theorem 3.9. *Let $p \in (0,1)$ be a constant. The binomial undirected random graph which allows loops $G_{loop}(n,p)$ obeys the FO 0–1 law.*

We now switch to random d-uniform hypergraphs. It is known [5, 7] that oriented random hypergraphs obey the FO 0-1 law. Using the stochastic FO reduction, we show that this is also the case for unoriented random hypergraphs. Note that an unoriented hypergraph can be considered as an oriented hypergraph where all hyperedges that can be obtained from each other via a permutation $\sigma \in S_d$ are identified. Naturally, one may consider partially oriented hypergraphs by identifying hyperedges that belong to an H-orbit of an oriented hyperedge for some subgroup H of S_d. Using reductions we prove the FO 0–1 law for all these hypergraphs too. Let us give an accurate definition.

Let S_d be the group of permutations of $[d]$. We define the action of S_d on a hyperedge $\{v_1,\ldots,v_d\}$ in the natural way: $g(v_1,\ldots,v_d) = (v_{g^{-1}(1)},\ldots,v_{g^{-1}(d)})$, $g \in S_d$. Let H be a subgroup of S_d.

Definition 3.10. A d-**uniform** H-**hypergraph** on $[n]$ is an oriented hypergraph without loops which is invariant under the action of H on $[n]_d$, where $[n]_d$ is the set of all d-tuples of distinct vertices from $[n]$. An H-**hyperedge** of this graph is an orbit of its hyperedge under the action of H. A **binomial** d-**uniform** H-**hypergraph** $G^H(n,p)$ is a random structure over the signature $\sigma_H = \{=, P_H\}$, where P_H has arity d, where every H-hyperedge is presented independently with probability p.

A generalisation of the notion of H-oriented hypergraphs for structures with arbitrary number of relations was introduced in [17] and studied in the context of convergence laws in sparse regimes.

Theorem 3.11. *Let H be a subgroup of S_d, K be a subgroup of H, and $p \in (0,1)$ (not necessarily a constant). There is a stochastic FO reduction $G^H(n,p^{[H:K]}) \leq G^K(n,p)$. In particular, for a constant $p \in (0,1)$, $G^H(n,p)$ obeys the FO 0–1 law.*

Proof. A reduction f is defined by the formula

$$\bigwedge_{g\in H} P_K(x_{g^{-1}(1)},\ldots,x_{g^{-1}(d)}).$$

Indeed, $f(G^K(n,p)) \stackrel{d}{=} G^H(n,p^{[H:K]})$

- because the presence of an H-hyperedge $(x_1,\ldots,x_d)$ in $f(G^K(n,p))$ depends only on the presence of K-hyperedges $(x_{g^{-1}(1)},\ldots,x_{g^{-1}(d)})$ in $G^K(n,p)$, for $g \in H$;
- the probability that all of them are presented is $p^{[H:K]}$ as, for each coset $gK \subseteq H$, the hyperedge $(x_{g^{-1}(1)},\ldots,x_{g^{-1}(d)})$ is presented with probability p, and there are exactly $[H:K]$ such cosets.

Moreover, we have a reduction $G^{\{id\}}(n,p) \leq D^\sigma(n,p)$, where $\sigma = \{=, P\}$ and P has an arity d. In the same way, as in the proof of Theorem 3.5, this reduction is defined by the formula

$$P(x_1,\ldots,x_d) \wedge \bigwedge_{i\neq j}(x_i \neq x_j).$$

Therefore, for each subgroup H of S_d, there is a reduction from $G^H(n,p)$ to $D^\sigma(n,p^{|H|^{-1}})$. Since $D^\sigma(n,p^{|H|^{-1}})$ obeys the FO 0–1 law, for constant p, by the Corollary 3.4, $G^H(n,p)$ obeys the FO 0–1 law as well. □

Remark 3.12. In order to give a better flavour of the FOC-hierarchy of random H-hypergraphs, we note that conjugate subgroups of S_d define the same random structure up to equivalence. That is, for any $p \in (0,1)$ (not necessarily a constant), any subgroup H of S_d, and any $g \in S_d$, random structures $G^H(n,p)$ and $G^{gHg^{-1}}(n,p)$ are equivalent.

3.3 Application to FO limit laws for sparse structures

For technical reasons, we will require a claim stating that a small shift of the probability parameter of a binomial random structure does not affect its equivalence class. Let σ be a relational signature, and recall that $\mathcal{D}_n$ is the set of all finite structures on $[n]$ over σ. For every n, consider a non-negative integer s_n and an arbitrary mapping $r_n : \{0,1\}^{s_n} \to \mathcal{D}_n$. Let $B(r_n,p)$ be a random structure over the signature σ defined as $B(r_n,p) = r_n(\xi_1,\ldots,\xi_{s_n})$, where ξ_i are independent random variables with Bernoulli distribution with the parameter p. Note that the binomial random graph $G(n,p)$ is distributed as $B(r_n,p)$, where r_n maps a sequence of $\binom{n}{2}$ ones and zeros into the graph with edges corresponding to ones in the sequence. Similarly, binomial random d-uniform H-oriented hypergraphs and binomial random oriented hypergraphs with loops are distributed as $B(r_n,p)$ with an appropriately chosen r_n. We shall use the following assertion about the total variation distance between Bernoulli random variables (though we believe that it might be known, its proof can be found in the extended version of the paper [?]).

Lemma 3.13. *Let $p,q \in [0,1]$, $p_{\min} = \min\{p,q,1-p,1-q\}$, and s_n be a sequence of non-negative integers. Consider random vectors $(\xi_1,\ldots,\xi_{s_n})$ and $(\eta_1,\ldots,\eta_{s_n})$, where ξ_i and η_i are independent random variables with Bernoulli distribution with the parameter p and q respectively. If $\lim_{n\to\infty}\sqrt{\frac{s_n}{p_{\min}}}|p-q| = 0$, then the total variation distance between the distributions of $(\xi_1,\ldots,\xi_{s_n})$ and $(\eta_1,\ldots,\eta_{s_n})$ converges to zero.*

Corollary 3.14. *Let $p,q \in [0,1]$, $p_{\min} = \min\{p,q,1-p,1-q\}$, and $r_n : \{0,1\}^{s_n} \to \mathcal{D}_n$ be a sequence of arbitrary mappings. If $\lim_{n\to\infty}\sqrt{\frac{s_n}{p_{\min}}}|p-q| = 0$, then $B(r_n,p)$ is equivalent to $B(r_n,q)$.*

Proof. Consider the identity mapping $id : \mathcal{D} \to \mathcal{D}$. Let us prove that it is a stochastic FO reduction. Consider a FO sentence φ. It is eligible to set $id^{-1}(\varphi) := \varphi$. Let $A_n = r_n^{-1}(\mathcal{D}(\varphi))$.Then

$$\begin{aligned}\Pr(B(r_n,p) \models \varphi) &= \Pr((\xi_1,\ldots,\xi_{s_n}) \in A_n),\\ \Pr(B(r_n,q) \models \varphi) &= \Pr((\eta_1,\ldots,\eta_{s_n}) \in A_n).\end{aligned}$$

Since the total variation distance between vertors $(\xi_1,\ldots,\xi_{s_n})$ and $(\eta_1,\ldots,\eta_{s_n})$ converges to zero, we immediately get the statement of the corollary. □

This corollary advances our tool to prove logical limit laws using reductions. In order to demonstrate its efficiency, we prove the following.

Theorem 3.15. *Let r be a nonnegative integer.*

(i) *The random structure $D^{(r+2)}(n,(r+1)\ln n/n)$ does not obey the FO 0–1 law.*

(ii) *For any positive integer* k, $D^{(r+3)}(n,(1+\frac{1}{k}+r)\ln n/n)$ *does not obey the FO 0–1 law.*
(iii) *For any rational* $\beta\in(\frac{2}{3},1)$, $\mathrm{FOC}(D^{(r+3)}(n,(\beta+r)\ln n/n))$ *is not totally bounded.*

Proof. We will use two FO reductions encapsulated in the two claims stated below that follow from Corollary 3.14 and reductions defined by formulae (see the extended version of the paper [?] for their proofs).

Claim 3.16. *For any* $\alpha>\frac{d}{3}$ *and each nonnegative integer* r, *we have* $D^{(d)}(n,\frac{c}{n^{\alpha}})\le D^{(d+r)}(n,\frac{c}{n^{\alpha+r}})$.

Claim 3.17. *For any* $\alpha>d-2$, $D^{(d)}(n,n^{-\alpha})\le D^{(d+1)}(n,\frac{\alpha\ln n}{n})$.

Let us finish the proof of Theorem 3.15. We start from (i). The FO 0–1 law fails for $D^{(1)}(n,\frac{1}{n})$, since the number of elements satisfying the unary predicate from the signature of this random structure converges in distribution to Pois(1). Using Claim 3.16 for parameter $\alpha=1>\frac{1}{3}=\frac{d}{3}$, we get a reduction $D^{(1)}(n,\frac{1}{n})\le D^{(r+1)}(n,\frac{1}{n^{r+1}})$. Using Claim 3.17 for parameter $\alpha=r+1>r-1=d-2$, we get a reduction $D^{(r+1)}(n,\frac{1}{n^{r+1}})\le D^{(r+2)}(n,\frac{(r+1)\ln n}{n})$. Then, we transfer the absence of the FO 0–1 law from $D^{(1)}(n,\frac{1}{n})$ to $D^{(r+2)}(n,\frac{(r+1)\ln n}{n})$.

Next, let us prove (ii). By Theorem 3.5, there is a reduction $G(n,(1-n^{-1-\frac{1}{k}})^2)\le D^{(2)}(n,1-n^{-1-\frac{1}{k}})$. Since the inversion of all edges defines equivalences between $G(n,p)$ and $G(n,1-p)$, and between $D^{(2)}(n,p)$ and $D^{(2)}(n,1-p)$, we have $G(n,2n^{-1-\frac{1}{k}}-n^{-2-\frac{2}{k}})\le D^{(2)}(n,n^{-1-\frac{1}{k}})$. Using Claim 3.16 for parameter $\alpha=1+\frac{1}{k}>\frac{2}{3}=\frac{d}{3}$, we get a reduction $D^{(2)}(n,n^{-1-\frac{1}{k}})\le D^{(r+2)}(n,n^{-1-\frac{1}{k}-r})$. Using Claim 3.17 for parameter $\alpha=1+\frac{1}{k}+r>r=d-2$, we get a reduction $D^{(r+2)}(n,n^{-1-\frac{1}{k}-r})\le D^{(r+3)}(n,(1+\frac{1}{k}+r)\ln n/n)$. Then, we obtain a reduction $G(n,2n^{-1-\frac{1}{k}}-n^{-2-\frac{2}{k}})\le D^{(r+3)}(n,(1+\frac{1}{k}+r)\ln n/n)$. From [29], we have an absence of the FO 0–1 law for $G(n,p)$ with $p\sim cn^{-1-\frac{1}{k}}$, and then for $D^{(r+3)}(n,(1+\frac{1}{k}+r)\ln n/n)$ as well.

Finally, we prove (iii). By Theorem 3.5, there is a reduction $G(n,2n^{-\beta}-n^{-2\beta})\le D^{(2)}(n,n^{-\beta})$. Using Claim 3.16 for parameter $\alpha=\beta>\frac{2}{3}=\frac{d}{3}$, we get a reduction $D^{(2)}(n,n^{-\beta})\le D^{(r+2)}(n,n^{-\beta-r})$. Using Claim 3.17 for parameter $\alpha=\beta+r>r=d-2$, we get a reduction $D^{(r+2)}(n,n^{-\beta-r})\le D^{(r+3)}(n,(\beta+r)\ln n/n)$. Then, we obtain a reduction $G(n,2n^{-\beta}-n^{-2\beta})\le D^{(r+3)}(n,(\beta+r)\ln n/n)$ which implies that $\mathrm{FOC}(G(n,2n^{-\beta}-n^{-2\beta}))\subseteq\mathrm{FOC}(D^{(r+3)}(n,(\beta+r)\ln n/n))$.

The proof of Theorem 2.6 and the proof of Lemma 2.5 in [29] do not rely on the exact equality $p=n^{-\alpha}$ but rather works for every $p\sim cn^{-\alpha}$. Then, $\mathrm{FOC}(G(n,2n^{-\beta}-n^{-2\beta}))$ is not totally bounded, and it finishes the proof of the theorem by Claim 3.3. □

3.4 First order complexity of random H-hypergraphs around the connectivity threshold

For a subgroup H of S_d, the random hypergraph $G^H(n,p)$ is defined in Section 3.2. In particular, $G^{S_d}(n,p)$ is the well-studied and commonly considered binomial unoriented hypergraph. Larrauri, Müller, and Noy [18] proved the following.

Theorem 3.18 (Larrauri, Müller, Noy [18]). *The set of limiting probabilities* $\lim_{n\to\infty}\Pr(G^{S_d}(n,\frac{c}{n^{d-1}})\models\varphi)$ *over all FO sentences* φ *is dense in* $[0,1]$ *if and only if* $c\ge c_0^{S_d}$, *where* $c_0^{S_d}$ *is the unique positive solution of the equation*

$$\frac{1}{2}\ln\frac{1}{1-\frac{c}{(d-2)!}}-\frac{c}{2(d-2)!}=\ln 2.$$

Using stochastic FO reductions, we generalise this result to all possible orientations.

In [17] it was proven that $G^H(n,\frac{c}{n^{d-1}})$ obeys the FO convergence law. For $c>0$, let L_c^H be the set of $\lim_{n\to\infty}\Pr(G^H(n,\frac{c}{n^{d-1}})\models\varphi)$ over all FO sentences φ. Let us denote by c_0^H the infimum of the set of positive numbers c such that L_c^H is dense in $[0,1]$. In this section, we prove the following.

Theorem 3.19. *Let* H *be a subgroup of* S_d, $c>0$, *and let* $d\ge 2$ *be an integer. Then,* $c_0^H=\frac{|H|}{d!}c_0^{S_d}$. *Moreover, for* $c\ge c_0^H$, L_c^H *is dense.*

Proof. Theorem 3.19 follows from the next two claims.

Claim 3.20. *Let* H *be a subgroup of* S_d, K *be a subgroup of* H, $c>0$, *and let* $d\ge 2$ *be an integer. Then,* $c_0^H\ge[H:K]c_0^K$. *Moreover, for any* $c\ge\frac{|H|}{d!}c_0^{S_d}$, *the set* L_c^H *is dense.*

Proof. Note that $G^H(n,p)$ is equivalent to $G^H(n,1-p)$ due to reductions defined by negations. From Theorem 3.11 and Corollary 3.14, we have the following reductions

$$G^H(n,[H:K]cn^{1-d})\le G^H(n,1-[H:K]cn^{1-d})\le$$
$$G^H(n,(1-cn^{1-d})^{[H:K]})\le G^K(n,1-cn^{1-d})\le G^K(n,cn^{1-d}).$$

We immediately get

Proposition 3.21. $G^H(n,[H:K]c/n^{d-1})\le G^K(n,c/n^{d-1})$.

By Proposition 3.21, for all $c>0$, $L_{[H:K]c}^H\subseteq L_c^K$. Therefore, if $L_{[H:K]c}^H$ is dense in $[0,1]$, then L_c^K is dense in $[0,1]$ as well. Hence, $c_0^H\ge[H:K]c_0^K$. Applying this inequality for S_d and H, we get $\frac{|H|}{d!}c_0^{S_d}=[S_d:H]^{-1}c_0^{S_d}\ge c_0^H$. Also, for each $c\ge\frac{|H|}{d!}c_0^{S_d}$, $L_{d!c/|H|}^{S_d}$ is dense in $[0,1]$ by Theorem 3.18. By Proposition 3.21, $G^{S_d}(n,d!c/(|H|n^{d-1}))\le G^H(n,c/n^{d-1})$. Therefore, the set L_c^H is dense in $[0,1]$ as well, completing the proof. □

The next claim can be proved in the same way as the analogous statement in [18] using a Poisson limit theorem for the number of small cycles in random hypergraphs and the validity of the FO 0–1 law subject to the absence of cycles, see the proof in the extended version of the paper [?].

Claim 3.22. *For* $c<c_0^{S_d}/d!$, *the set* $L_c^{\{id\}}$ *is not dense in* $[0,1]$.

By Claim 3.22, $c_0^{\{id\}}\ge c_0^{S_d}/d!$. Combining with Claim 3.20, we get $\frac{|H|}{d!}c_0^{S_d}\ge c_0^H\ge|H|c_0^{\{id\}}\ge\frac{|H|}{d!}c_0^{S_d}$. Thus, $c_0^H=\frac{|H|}{d!}c_0^{S_d}$. Moreover, due to Claim 3.20, for each $c\ge c_0$, L_c^H is dense, completing the proof of Theorem 3.19.

□

4 DECISION PROBLEM

In this section, we prove that the problem of determining whether, given an input FO sentence φ, $G(n \mid \varphi)$ obeys the FO 0–1 law is not recursively enumerable. In what follows, we refer to this decision problem as 0−1LAW.

Theorem 4.1. 0−1LAW *is not recursively enumerable.*

Proof. It is enough to reduce the problem of deciding whether a Diophantine equation has a solution (DE in what follows) to the complement of 0−1LAW. Indeed, DE is complete in the class of recursively enumerable languages due to Matiiasevich [24] (see also [23] for a survey on the negative solution of Hilbert's 10th problem), then its complement $\overline{\text{DE}}$ is not recursively enumerable, and therefore from the reduction we would get that 0−1LAW is not recursively enumerable as well.

So, for each integer polynomial $P(x_1, \ldots, x_k)$ we shall compute a FO sentence φ_P such that $P(x_1, \ldots, x_k) = 0$ has an integer solution if and only if $G(n \mid \varphi_P)$ does not obey the FO 0–1 law.

We first reduce the problem for integer solutions to a problem for positive integer solutions. For integer polynomial $P(x_1, \ldots, x_k)$, there is a solution of $P(x_1, \ldots, x_k) = 0$ in integers if and only if $P(y_1 - z_1, \ldots, y_k - z_k) = 0$ has a solution in positive integers. Also, we can move all monomials with negative coefficients in $P(y_1 - z_1, \ldots, y_k - z_k) = 0$ to the right-hand side of the equality. Then, we get an equation $Q(y_1, z_1, \ldots, y_k, z_k) = R(y_1, z_1, \ldots, y_k, z_k)$, where Q and R have nonnegative integer coefficients.

Let us now consider a system of equations $\mathcal{S}$ that has a solution in positive integers if and only if $Q = R$ has such a solution. Initially we denote occurrences of all variables in P or Q by $t_1, \ldots, t_s$ — here s is the total number of occurrences of variables, taking into account their powers. Then for every $i \in [2, s]$, we add $t_i = t_j$ to the system, if there exists $j < i$ such that t_i and t_j denote exactly the same variable. Then, for each of the two polynomials Q and P, we consider the sequence of arithmetic operations that are applied to compute them. Observe that the j-th operation can be written as either $t_{s+j} = t_i t_{i'}$ or $t_{s+j} = t_i + t_{i'}$ for certain $i, i' \leq s + j - 1$. For each of the operations, we add the respective equation to the system. Let us assume that Q and P are computed at steps q and p respectively (that is, $t_q = Q$ and $t_p = P$). The last equation in the system is $t_q = t_p$. Let us observe that indeed the initial Diophantine equation has a solution in integers if and only if the constructed system of equations $\mathcal{S}$ has a solution in positive integers.

Sequence of computations of P and Q encoded in $\mathcal{S}$ can be represented also by a FO sentence ψ_P (an explicit construction of this sentence is presented in the extended version of the paper [**?**]) with the following properties: (1) ψ_P has finite models if and only if $P = Q$ has solutions in positive integers; (2) if $G \models \psi_P$, then G has even number of vertices and a graph obtained from G by the addition of an isolated edge satisfies ψ_P as well.

We define the desired sentence φ_P as $\text{Empty} \vee \psi_P$, where $\text{Empty} = \forall x \forall y \neg(x \sim y)$ describes the property of being empty. If there are *no* integer solutions of $P(x_1, \ldots, x_k) = 0$, then there are no graphs satisfying ψ_P. Therefore, any G satisfying φ_P is empty. We immediately get that $G(n \mid \varphi_P)$ obeys the FO 0–1 law. On the other hand, if there is an integer solution of $P(x_1, \ldots, x_k) = 0$, then there is a solution of $\mathcal{S}$ in $\mathbb{Z}^s_{>0}$. Let G_0 be a graph satisfying ψ_P. Let G_i be obtained from G_0 by adding i isolated edges. Obviously, $G_i \models \psi_P$ for all i, and $|V(G_i)| = 2i + |V(G_0)|$. For odd n, $\Pr(G(n \mid \varphi_P) \models \text{Empty}) = 1$ because there are no graphs with odd number of vertices satisfying ψ_P. In contrast, for all even $n \geq |G_0|$, this probability is at most $1/2$, because there is at least one nonempty graph satisfying ϕ_P. Therefore, $\Pr(G(n \mid \varphi_P) \models \text{Empty})$ does not converge, and so $G(n \mid \varphi_P)$ does not obey the FO 0–1 law, completing the proof. ◻

Remark 4.2. Actually, we proved that 0−1LAW is Π_1-hard. On the other hand, $\varphi \in$ 0−1LAW if and only if

$$\forall M \forall \psi \ \exists N \ \forall n > N \ \Pr(G(n \mid \varphi) \models \psi) \notin (1/M, 1 - 1/M).$$

Since the property $\Pr(G(n \mid \varphi) \models \psi) \notin (1/M, 1 - 1/M)$ is decidable, we have that 0−1LAW $\in \Pi_3$. Unfortunately, we do not manage to find the level of 0−1LAW in the arithmetical hierarchy.

5 PROOF OF PART (III) OF THEOREM 2.7

This section is divided into four parts. In Section 5.1, we recall the necessary background used in the proof. In Section 5.2 we show the general scheme of the proof and reduce the theorem to a construction of two FO sentences φ_1 and φ_2 that have to satisfy certain properties. After that, in Section 5.3 and Section 5.4 we construct the desired FO sentences and verify their properties.

5.1 Preliminaries

Graphs G and H are *(elementary) k-equivalent* (we write $G \cong_k H$) if there is no FO sentence of *quantifier depth k* that distinguishes between G and H (quantifier depth is the maximum number of nested quantifiers in the sentence, see the formal definition in [19]). We have to recall a combinatorial approach to proving the elementary equivalence, this tool is widely known as the Ehrenfeucht–Fraïssé game. For simplicity of presentation (and this is also enough for our purposes), we recall the definition of this game for graphs; for arbitrary relational structures see [3, 6, 19].

Definition 5.1. In the *Ehrenfeucht–Fraïssé game* on graphs G and H with k rounds, there are two players, *Spoiler* and *Duplicator*. In round $i \in [k]$, Spoiler chooses a vertex in one of the graphs and then Duplicator chooses a vertex in the other graph. After k rounds are played, there are two tuples of k not necessarily different vertices $(x_1, \ldots, x_k)$ chosen in G and $(y_1, \ldots, y_k)$ chosen in H are chosen by both players (the i-th vertex is chosen in the i-th round). Duplicator wins if $(x_i = x_j) \Leftrightarrow (y_i = y_j)$ and $(x_i \sim x_j) \Leftrightarrow (y_i \sim y_j)$, for each pair $1 \leq i < j \leq k$, that is the map $x_i \to y_i$ is a partial isomorphism between G and H.

Theorem 5.2 (Ehrenfeucht, Fraïssé [3, 6]). *For two graphs G and H, Duplicator has a winning strategy in the Ehrenfeucht–Fraïssé game on graphs G and H in k rounds if and only if $G \cong_k H$.*

For constructing winning strategies in the Ehrenfeucht–Fraïssé game, we need the following auxiliary simple claims, their proofs are presented in the extended version of the paper [**?**].

Claim 5.3. *Let Spoiler has no winning strategy in k moves on graphs G_1 and G_2. Then, for each graph H, there is no winning strategy for Spoiler in k moves on graphs $G_1 \square H$ and $G_2 \square H$.*

Claim 5.4. *Let Y, G, H be a triple of graphs with induced subgraphs $X_0 \subset Y$, $X_1 \subset G$, $X_2 \subset H$ and isomorphisms $f : X_0 \to X_1$, g :*

$X_0 \to X_2$. Denote by G_X and H_X relational structures with the usual adjacency relation in G and H and a unary relation I_x for each vertex $x \in X_0$ such that $I_x(y) \Leftrightarrow (y = f(x))$ for $y \in G$ and $I_x(z) \Leftrightarrow (z = g(x))$ for $z \in H$. Let $G \cup_f Y$ be the graph obtained from G by identifying every $x \in X_0$ with $f(x) \in X_1$ and adding the rest of Y vertex-disjointly. Similarly define $H \cup_g Y$. Assume that Duplicator has a winning strategy in k moves on G_X and H_X, then Duplicator has a winning strategy in k moves on $G \cup_f Y$ and $H \cup_g Y$ as well.

Claim 5.5. *Let G, H be graphs with an additional unary relation I. Let G_I, H_I be graphs obtained from G, H by attaching exactly one leaf to each vertex x satisfying $I(x)$. Assume that Duplicator has a winning strategy in k rounds on G and H equipped with I, then Duplicator has a winning strategy in k rounds on G_I and H_I as well.*

Moreover, we need a well-known fact that FO sentences express only "local" properties in the following sense.

Theorem 5.6 (Hanf locality theorem [10]). *For a positive integer r and a graph G, denote by $n_r(x, G)$ the number of vertices x' in G such that r-neighbourhoods of x and x' are isomorphic. For every positive integer k there exist positive integers $r(k)$ and $s(k)$ such that the following implication holds true for any two graphs G and H. If, for every vertex $x \in G \sqcup H$, $\min\{s(k), n_r(x, G)\} = \min\{s(k), n_r(x, H)\}$, then $G \cong_k H$.*

Finally, we make use of the following theorem about the distribution of the number of cycles in a uniformly random permutation.

Theorem 5.7 ([16], Theorem 8). *Let n be a positive integer, and let Y_i be the number of cycles of length i in a uniformly random permutation from S_n. Then, $(Y_1, \dots, Y_n) \xrightarrow{d} (\xi_1, \dots, \xi_n)$, where ξ_i are independent $\mathrm{Pois}(1/i)$ random variables.*

5.2 Overview of the proof and basic properties of φ

We define the desired sentence φ as a disjunction of three sentences φ_0, φ_1 and φ_2, such that the following holds:

(a) any graph satisfies at most one sentence from φ_0, φ_1 and φ_2;
(b) φ_0 expresses the property of graph being empty;
(c) for each n divisible by 6, there are $n!$ graphs on $[n]$ that satisfy φ_1, while, for any n which is not divisible by 6, any graph on $[n]$ does not satisfy φ_1;
(d) for each FO ψ, the limit $\lim_{m\to\infty} \Pr(G(6m \mid \varphi_1) \models \psi)$ exists;
(e) there is a family of FO sentences $\varphi_{2,d}$, $d \geq 3$, such that φ_2 and $\bigvee_{d\geq 3} \varphi_{2,d}$ are not distinguished by any finite graph, and $\varphi_{2,d} \wedge \varphi_{2,d'}$ are contradictions for all distinct d, d';
(f) for every $d \geq 3$ and every $n = 6dm^2$ for some integer $m \neq 0$, there are exactly $n!/d!$ graphs on $[n]$ that satisfy $\varphi_{2,d}$, while for all other n, any graph on $[n]$ does not satisfy $\varphi_{2,d}$;
(g) for all FO ψ and $d \geq 3$, $\lim_{m\to\infty} \Pr(G(6dm^2 \mid \varphi_{2,d}) \models \psi)$ is either 0 or 1.

Let us first show that these seven assumptions are enough for obtaining the desired properties of $\mathrm{FOC}(G(n \mid \varphi))$ and then construct this sentence.

Claim 5.8. *If $\varphi = \varphi_0 \vee \varphi_1 \vee \varphi_2$ satisfies assumptions (a)-(g), then $\mathrm{FOC}(G(n \mid \varphi))$ is totally bounded but spans an infinite-dimensional subspace of ℓ^∞/c_0.*

Proof. Consider any positive integer n. The number of graphs that satisfy φ equals

- 1, if n is not divisible by 6;
- $1 + n!\left(1 + \sum_{d\in\mathbb{Z},\, d\geq 3,\, d=\frac{n}{6m^2}} \frac{1}{d!}\right)$, if n is divisible by 6.

Consider a FO sentence ψ. Due to (b), (d), and (g), there exist

- $\beta_\psi := \lim_{n\to\infty} \Pr(G(n \mid \varphi_0) \models \psi) \in \{0, 1\}$,
- $p_{1,\psi} := \lim_{m\to\infty} \Pr(G(6m \mid \varphi_1) \models \psi) \in [0, 1]$,
- $\beta_{d,\psi} := \lim_{m\to\infty} \Pr(G(6dm^2 \mid \varphi_{2,d}) \models \psi) \in \{0, 1\}$.

Let us define the sequence $p_\psi(n)$, $n \in \mathbb{N}$, in the following way: if $6 \mid n$, then

$$p_\psi(n) = \left(p_{1,\psi} + \sum_{d\in\mathbb{Z},\, d\geq 3,\, d=\frac{n}{6m^2}} \frac{\beta_{d,\psi}}{d!}\right) \Bigg/ \left(1 + \sum_{d\in\mathbb{Z},\, d\geq 3,\, d=\frac{n}{6m^2}} \frac{1}{d!}\right);$$

otherwise, $p_\psi(n) = \beta_\psi$. Then, we have

$$\lim_{n\to\infty} \left|\Pr(G(n \mid \varphi) \models \psi) - p_\psi(n)\right| = 0. \tag{1}$$

Indeed, due to (b), (c), (e), and (f), the number of graphs on $[n]$ that satisfy $\varphi_0 \wedge \psi$ is β_ψ, for sufficiently large n; the number of graphs on $[n = 6m]$ that satisfy $\varphi_1 \wedge \psi$ is $p_{1,\psi} n!(1 + o(1))$; the number of graphs on $[n = 6dm^2]$ that satisfy $\varphi_{2,d} \wedge \psi$ is $\beta_{d,\psi} \frac{n!}{d!}(1 + o(1))$, and is bounded from above by $\frac{n!}{d!}$. Then, $6 \nmid n$ implies $\Pr(G(n \mid \varphi) \models \psi) = \beta_\psi$ for sufficiently large n, and $6 \mid n$ implies

$$\Pr(G(n \mid \varphi) \models \psi) = \frac{p_{1,\psi} + \sum_{d\in\mathbb{Z},\, d\geq 3,\, d=\frac{n}{6m^2}} \beta_{d,\psi} \frac{1}{d!}}{1 + \sum_{d\in\mathbb{Z},\, d\geq 3,\, d=\frac{n}{6m^2}} \frac{1}{d!}} + o(1). \tag{2}$$

For any $d' \geq 3$, let us consider $\psi_{d'} := \varphi_{2,d'}$. From (a), (c), and (f), observe that

$$p_{\psi_{d'}}(n) = \begin{cases} \left(d'!\left(1 + \sum_{d\in\mathbb{Z},\, d\geq 3,\, d=\frac{n}{6m^2}} \frac{1}{d!}\right)\right)^{-1}, & \text{if } \frac{n}{6d'} \text{ is a square;} \\ 0, & \text{otherwise.} \end{cases}$$

Let D be the set of all square-free integers $d' \geq 3$. Vectors $\pi\left(\left(p_{\psi_{d'}}(n)\right)_{n\in\mathbb{N}}\right)$, $d' \in D$, are linearly independent since, for each $d' \in D$, the sequence $n = 6d'm^2$ satisfies $p_{\psi_{d'}}(n) > \frac{1}{d'!e}$, and, for each $d'' \in D$ such that $d'' \neq d'$, and $n = 6d'm^2$, we have that $p_{\psi_{d''}}(n) = 0$. So, any finite non-trivial linear combination of $\left(p_{\psi_{d''}}(n)\right)_{n\in\mathbb{N}}$ involving $\left(p_{\psi_{d'}}(n)\right)_{n\in\mathbb{N}}$ does not equal 0. Due to (1), $\mathrm{FOC}(G(n \mid \varphi))$ spans an infinite-dimensional space.

Next, we prove that $\mathrm{FOC}(G(n \mid \varphi))$ is totally bounded. Consider $\varepsilon > 0$. Since all $\beta_{d,\psi}$ are at most 1, there is a d_0 such that, if we restrict the summation in the definition of $p_\psi(n)$ to $d \leq d_0$, then we get an $\frac{\varepsilon}{2}$-approximation, i.e.

$$\left|p_\psi(n) - \frac{p_{1,\psi} + \sum_{d\in\mathbb{Z},\ d_0\geq d\geq 3,\ d=n/(6m^2)} \frac{\beta_{d,\psi}}{d!}}{1 + \sum_{d\in\mathbb{Z},\ d\geq 3,\ d=n/(6m^2)} \frac{1}{d!}}\right| < \frac{\varepsilon}{2}.$$

Let us construct a finite ε-covering for the set of all sequences $p_\psi(n)$ in ℓ^∞. Let N be a positive integer such that $\frac{1}{N} < \frac{\varepsilon}{2}$. The desired ε-covering is the set of all sequences $v_{k,\mathbf{b}}(n)$ indexed by $k \in [N]$ and $\mathbf{b} = (b, b_3, b_4, \dots, b_{d_0}) \in \{0,1\}^{d_0-1}$ and defined in the following way: if $6 \mid n$, then

$$v_{k,\mathbf{b}}(n) = \left(\frac{k}{N} + \sum_{d_0 \geq d \geq 3,\ d = n/(6m^2)} \frac{b_d}{d!}\right) \Bigg/ \left(1 + \sum_{d \geq 3,\ d = n/(6m^2)} \frac{1}{d!}\right);$$

otherwise, $v_{k,\mathbf{b}}(n) = b$. Thus, this family of sequences is indeed the desired ε-covering. Due to (1), $\mathrm{FOC}(G(n \mid \varphi))$ is totally bounded, completing the proof. □

To finish the proof of (iii), we construct the FO sentences φ_1 in Section 5.3 and φ_2 in Section 5.4.

5.3 Definition of φ_1 and verification of its properties

Let Δ_m be the set of all digraphs on $[m]$ with all in-degrees and out-degrees equal 1. In particular, a loop contributes 1 both to the in-degree and the out-degree of the respective vertex. Such digraphs are disjoint unions of oriented cycles and there are exactly $m!$ digraphs in Δ_m. For each class of isomorphism D of digraphs on m vertices we construct a class of isomorphism $G(D)$ of simple graphs on $6m$ vertices in the following way. We replace every edge $(u,v) \in D$ with a graph on the set of vertices $\{u, v, w_i(u,v), i \in [5]\}$, where all vertices $w_i(u,v)$ are different, comprising edges $\{u, w_1(u,v)\}$, $\{w_1(u,v), w_3(u,v)\}$, $\{w_3(u,v), v\}$, $\{w_1(u,v), w_2(u,v)\}$, $\{w_3(u,v), w_4(u,v)\}$, and $\{w_4(u,v), w_5(u,v)\}$.

Let φ_1 be a FO sentence describing the property $\bigcup_m \Gamma_m$, where $\Gamma_m = \{G(D), D \in \Delta_m\}$. Let us check the condition (c). For each m, the set Γ_m consists only of graphs of size $6m$. Then, there are no graphs that satisfy φ_1 of size not divisible by 6. Let $D_1, \dots, D_t$ be all classes of isomorphism of digraphs in Δ_m. Then, we have $1 = \sum_{D_i} \frac{1}{|\mathrm{Aut}(D_i)|}$. Also, $|\mathrm{Aut}(D_i)| = |\mathrm{Aut}(G(D_i))|$. Then the number of graphs on $[6m]$ that satisfy φ_1 equals

$$\sum_{D_i} \frac{(6m)!}{|\mathrm{Aut}(G(D_i))|} = \sum_{D_i} \frac{(6m)!}{|\mathrm{Aut}(D_i)|} = (6m)! \sum_{D_i} \frac{1}{|\mathrm{Aut}(D_i)|} = (6m)!.$$

Let us finally verify the condition (d). Fix a FO sentence ψ of quantifier depth d. By Theorem 5.6, for $D, D' \in \Gamma_m$ with equal numbers of components of all sizes at most $6 \cdot 2^d$, ψ does not distinguish between D and D'. For every $i \leq 2^d$, let X_i be the number of components in $G(n \mid \varphi_1)$ of size $6i$; let Y_i be the number of components of size i in a uniformly random digraph from Δ_m. Note that $(X_1, \dots, X_{2^d}) \stackrel{d}{=} (Y_1, \dots, Y_{2^d})$, and that $(Y_1, \dots, Y_{2^d}) \xrightarrow{d} (\xi_1, \dots, \xi_{2^d})$ due to Theorem 5.7, where ξ_i are independent $\mathrm{Pois}(\frac{1}{i})$ random variables. Then,

$$\begin{aligned}\Pr(G(6m \mid \varphi_1) \models \psi) &= \sum \Pr(X_1 = x_1, \dots, X_{2^d} = x_{2^d}) \\ &= (1 + o(1)) \sum \Pr(\xi_1 = x_1, \dots, \xi_{2^d} = x_{2^d})\end{aligned}$$

completing the proof of (d).

5.4 Definition of φ_2 and verification of its properties

Next, we construct the FO sentence φ_2 and sentences $\varphi_{2,d}$. Let us first, for every integer $m \geq 2$, define an auxiliary graph L_m. This graph consists of vertices u_i, for $1 \leq i \leq 6m$; $v_{i,j}$, for $j \in [m-1]$, $6j + 1 \leq i \leq 6m$; and $w_{i,j}$, for $j \in [m-1]$, $6j+1 \leq i \leq 6m$; and edges

- $\{u_{i-1}, u_i\}$, for $2 \leq i \leq 6m$;
- $\{u_i, v_{i,j}\}$, $\{v_{i,j}, w_{i,j}\}$, for $j \in [m-1]$, $6j + 1 \leq i \leq 6m$;
- $\{v_{i-1,j}, v_{i,j}\}$, for $j \in [m-1]$, $6j + 2 \leq i \leq 6m$.

The graph L_m consists of $6m + 6\frac{m(m-1)}{2} + 6\frac{m(m-1)}{2} = 6m^2$ vertices and has no nontrivial automorphisms. Let φ_L be a FO sentence expressing the property of being isomorphic to L_m for some $m \geq 2$. Consider the family of graphs $K_d \square L_m$, for $d \geq 3$ and $m \geq 2$, where $G \square H$ is the Cartesian product of graphs, see Definition 2.9. We construct a FO sentence φ_2 that expresses the property of being isomorphic to $K_d \square L_m$ for some $d \geq 3$ and $m \geq 2$ and, for every $d \geq 3$, we construct a FO sentence $\varphi_{2,d}$ that expresses the property of being isomorphic to $K_d \square L_m$ for some $m \geq 2$. We first set

$$\varphi_2 = \mathrm{TEquiv} \wedge \mathrm{TGraph} \wedge \varphi_{TL} \wedge \mathrm{TCommute},$$

where the clauses are defined in the following way. Let

$$\mathrm{Triangle}(x,y) = (x = y) \vee (\exists z\ (x \sim y) \wedge (y \sim z) \wedge (z \sim x))$$

express the property of distinct x, y to belong to a triangle. For FO formulae ψ_1, ψ_2 with a free variable x, let

$$\begin{aligned}\mathrm{Matching}_x[\psi_1, \psi_2] = &\forall x\ \neg(\psi_1(x) \wedge \psi_2(x)) \wedge \\ &\big(\psi_1(x) \Rightarrow \exists!\ x'(x \sim x') \wedge \psi_2(x')\big) \wedge \\ &\big(\psi_2(x) \Rightarrow \exists!\ x'(x \sim x') \wedge \psi_1(x')\big)\end{aligned}$$

say that sets $\{x \mid \psi_1(x)\}$ and $\{x \mid \psi_2(x)\}$ are disjoint and edges between them form a perfect matching.

- TEquiv is a FO sentence saying that $\mathrm{Triangle}(x,y)$ is an equivalence relation and each equivalence class has size at least 3 (or, in other words, every vertex belongs to a triangle).
- TGraph is a FO sentence saying that for each pair y, y' such that $\neg\mathrm{Triangle}(y, y')$ holds, either there are no edges between their Triangle-equivalence classes, or $TEdge(y,y') := \mathrm{Matching}_x[\mathrm{Triangle}(x,y), \mathrm{Triangle}(x,y')]$ holds.
- φ_{TL} is the sentence φ_L with all predicates $x = y$ replaced by $\mathrm{Triangle}(x,y)$, and all $x \sim y$ replaced by $\mathrm{TEdge}(x,y)$.
- TCommute is a FO sentence saying that, for each four vertices x, x', y, y' such that $\mathrm{TEdge}(x,x') \wedge \mathrm{TEdge}(y,y')$ and $\mathrm{TEdge}(x,y) \wedge \mathrm{TEdge}(x',y')$, the subgraph induced on their Triangle-equivalence classes is a disjoint union of cycles of length 4 with one vertex from each class.

Finally, $\varphi_{2,d}$ is the conjunction of φ_2 and a FO sentence saying that there is a clique on d vertices but no cliques on $d+1$ vertices. So, we immediately have (e).

Let us show that φ_2 and $\varphi_{2,d}$ express the proper sets of graphs. As usual, we omit the straightforward verification that $K_d \square L_m \models \varphi_{2,d}$. Assume $G \models \varphi_2$. By TEquiv $\wedge$ TGraph, the set of vertices of G is partitioned into Triangle-equivalence classes; edges between vertices of two different equivalence classes B, C in G appear if and only if representatives $x \in B$, $y \in C$ satisfy $\mathrm{TEdge}(x,y)$. Let us consider an auxiliary graph $\tilde{G}$ whose vertices are the Triangle-equivalence classes of G; two vertices B, C are adjacent in $\tilde{G}$ if and only if there are edges between them in G (and, in this case, these edges in G between B and C compose a perfect matching). Due to φ_{TL}, $\tilde{G}$ is isomorphic to some L_m. Since L_m is connected, all Triangle-equivalence classes in G have the same size d. Moreover, for a certain $d \geq 3$, these classes induce cliques K_d because, for distinct vertices x, y, Triangle(x, y) implies $x \sim y$. So, $G \models \varphi_{2,d}$.

For two different Triangle-equivalence classes B and C, the perfect matching between them defines two bijections $f_{BC} : B \to C$ and $f_{CB} : C \to B$ in the natural way: for $x \in B$, $y \in C$, the adjacency $x \sim y$ implies $f_{BC}(x) = y$ and $f_{CB}(y) = x$. Note that $f_{BC} \circ f_{CB} = id_C$ and $f_{CB} \circ f_{BC} = id_B$. By TCommute we have that, for each cycle $BB'C'C$ of length 4 in $\tilde{G}$, $f_{CC'} \circ f_{BC} = f_{B'C'} \circ f_{BB'}$. Note that the 2-dimensional CW-complex obtained by "filling" all 4-cycles of L_m is simply connected. Or, in algebraic language, for a certain spanning subtree $\hat{L}_m \subset L_m$, the group presented by $\langle f_{BC}, BC \in E(L_m) \mid f_{BC}, BC \in E(\hat{L}_m), f_{BC}f_{CB}, BC \in E(L_m), f_{B'B}f_{C'B'}f_{CC'}f_{BC}, BB'CC'$ is 4-cycle in $L_m\rangle$ is trivial. So, for every two walks between $B, C \in V(\tilde{G})$, compositions of respective bijections along the walks are equal. This fact legitimises the definition of f_{BC} for any pair of $B, C \in V(\tilde{G})$: let $BP_0 \ldots P_kC$ be a path in $\tilde{G}$, then $f_{BC} := f_{P_kC} \circ f_{P_{k-1}P_k} \circ \cdots f_{P_0P_1} \circ f_{BP_0}$. Next, fix a Triangle-equivalence class $B = \{b_1, \ldots, b_d\}$. For each b_i, let $S_i = \{f_{BC}(b_i), C \in V(\tilde{G})\}$. By the definition of f_{BC}, there are no edges $\{f_{BC}(b_i), f_{BC'}(b_j)\}$ for $i \neq j$ and $C \neq C'$, i.e. each $\tilde{G}$-edge $\{C, C'\}$ is presented by G-edges $\{f_{BC}(b_i), f_{BC'}(b_i)\}$, $1 \leq i \leq d$. Therefore, the induced subgraph on S_i is isomorphic to L_m and, consequently, the graph G itself is isomorphic to $K_d \square L_m$. So, φ_2 and $\varphi_{2,d}$ express desired properties of finite graphs.

The number of vertices in $K_d \square L_m$ is $6dm^2$. For each $n \neq 6dm^2$, there are no graphs that satisfy $\varphi_{2,d}$. For $n = 6dm^2$, there is exactly one graph under isomorphism that satisfy $\varphi_{2,d}$. The graph $K_d \square L_m$ has exactly $d!$ automorphisms. We conclude that the number of graphs that satisfy $\varphi_{2,d}$ equals $\frac{n!}{d!}$, completing the proof of (f).

To finish the proof, it remains to show (g). Due to Ehrenfeucht's theorem [3, 6], we know that there is a FO sentence of quantifier depth k that distinguishes between two graphs G and H if and only if Spoiler has a winning strategy in k moves in the Ehrenfeucht–Fraïssé game on G and H. Suppose, there is a FO sentence ψ of the quantifier depth k such that $\Pr(G(6dm^2 | \varphi_{2,d}) \models \psi)$ does not converge to either 0 or 1. This means that there are infinitely many m such that $K_d \square L_m \models \psi$ and infinitely many m such that $K_d \square L_m \not\models \psi$. Therefore, it is enough to show that, for every fixed positive integer k and every fixed integer $d \geq 3$, in the Ehrenfeucht–Fraïssé game Spoiler has no winning strategy in k moves, on graphs $K_d \square L_m$ and $K_d \square L_{m'}$, for large enough m and m'. Due to Claim 5.3, we may get rid of the Cartesian product and prove the same fact for graphs L_m and $L_{m'}$.

For $a < b \leq m$, let us consider the following induced subgraphs of L_m:

- $L_{(a,b]} = L_m\left[\{u_i, v_{i,j}, w_{i,j} \mid 6a + 1 \leq i \leq 6b, 6j + 1 \leq i\}\right]$;
- $Z_{(a,b]} = L_m\left[\{u_i, v_{i,j}, w_{i,j} \mid 6a + 1 \leq i \leq 6b, j \leq a\}\right]$;
- $\tilde{Z}_{(a,b]} = L_m\left[\{u_i, v_{i,j} \mid 6a + 1 \leq i \leq 6b, j \leq a\}\right]$.

Consider mappings $f_{a,b}$ of paths $X_{b-a} = u_1 \ldots u_{6(b-a)}$ to $Z_{(a,b]}$ such that $f_{a,b}(u_i) = u_{6a+i}$. Note that $L_{(a,b]} \cong Z_{(a,b]} \cup_{f_{a,b}} L_{(0,b-a]}$. Let Star$_r$ be a star graph with r leaves. So, $\tilde{Z}_{(a,b]} \cong X_{b-a} \square \text{Star}_a$. For two numbers $a, a' > k$, Duplicator has a winning strategy in the Ehrenfeucht–Fraïssé game with k moves on graphs Star$_a$ and Star$_{a'}$, this winning strategy preserves central vertices of stars. By Claim 5.3, Duplicator has a winning strategy in the Ehrenfeucht–Fraïssé game in k round on graphs $\tilde{Z}_{(a,b]}$ and $\tilde{Z}_{(a',b']}$, where $b - a = b' - a'$. Next, note that $Z_{a,b}$ is obtained from $\tilde{Z}_{(a,b]}$ by attaching a leaf to each vertex $v_{i,j}$. Since the winning strategy of Duplicator on stars is leaves-preserving, we can apply Claim 5.5 with I distinguishing between "leaves" and "non-leaves" and get a winning strategy of Duplicator in the game on $Z_{(a,b]}$ and $Z_{(a',b']}$. Finally, due to Claim 5.4, Duplicator has a winning strategy in the game on $L_{(a,b]}$ and $L_{(a',b']}$. This makes it possible to apply the well-known Duplicator's strategy on two long paths in the game on $L_m, L_{m'}$ since L_m and $L_{m'}$ can be represented as unions of three segments $L_{(a,b]}$ (the first one, for $a = 0$, an intermediate, and the last one, for $b \in \{m, m'\}$) such that the respective segments are elementary equivalent. For completeness, let us recall the strategy.

Let $p_m : V(L_m) \to [m]$ maps each vertex $u_i, v_{i,j}, w_{i,j}$ to the number $\lceil \frac{i}{6} \rceil$. Similarly define $p_{m'}$. Suppose that $m, m' > 3^{k+1}$ and in first $t - 1$ rounds vertices $x_1, \ldots, x_{t-1} \in L_m$ and $x'_1, \ldots, x'_{t-1} \in L_{m'}$ are chosen. Without loss of generality, Spoiler chooses a vertex $x_t \in L_m$ in round t.

- If $p_m(x_t) \leq 3^{k-t} + 1$, Duplicator chooses the same vertex x_t in $L_{m'}$.
- If $m - p_m(x_t) \leq 3^{k-t}$, Duplicator chooses the next vertex according to the strategy in the game on $L_{(m-3^{k-t}-1,m]}$ and $L_{(m'-3^{k-t}-1,m']}$.
- If there is x_s with $s < t$ such that $|p_m(x_t) - p_m(x_s)| \leq 3^{k-t}$, Duplicator chooses the next vertex according to the strategy in the game on $L_{(p_m(x_s)-3^{k-t}-1, p_m(x_s)+3^{k-t}]}$ and $L_{(p_{m'}(x'_s)-3^{k-t}-1, p_{m'}(x'_s)+3^{k-t}]}$.
- Otherwise, Duplicator chooses $3^{k-t} + 1 < a < m - 3^{k-t}$ such that $|a - p_{m'}(x'_s)| > 3^{k-t}$ for all $s < t$, and then chooses a vertex according to the strategy in the game on $L_{(p_m(x_s)-3^{k-t}-1, p_m(x_s)+3^{k-t}]}$ and $L_{(a-3^{k-t}-1, a+3^{k-t}]}$. It is possible to choose such an a since m' is large enough.

A straightforward inductive argument implies that, for every $t \leq k$, as soon as t rounds are played on $L_m, L_{m'}$, Duplicator has a winning strategy on pairs of graphs $\left(L_{(0,3^{k-t}+1]}, L_{(0,3^{k-t}+1]}\right)$, $\left(L_{(m-3^{k-t}-1,m]}, L_{(m'-3^{k-t}-1,m']}\right)$, $\left(L_{(p_m(x_s)-3^{k-t}-1, p_m(x_s)+3^{k-t}]}, L_{(p_{m'}(x'_s)-3^{k-t}-1, p_{m'}(x'_s)+3^{k-t}]}\right)$ for all $s \leq t$. It immediately implies that Duplicator wins the game on L_m, L'_m

Thus, (g) follows, completing the proof of part (iii) of Theorem 2.7.

ACKNOWLEDGEMENTS

The work of the first author is supported by RSF grant 22-11-00131.

REFERENCES

[1] B. Bollobás, *A probabilistic proof of an asymptotic formula for the number of labelled regular graphs*, European Journal of Combinatorics, **1**:4 (1980) 311–316.
[2] K. J. Compton, *A logical approach to asymptotic combinatorics II: monadic second-order properties*, Journal of Combinatorial Theory, Series A, **50**:1 (1989) 110–131.

[3] A. Ehrenfeucht, *An application of games to the completeness problem for formalized theories*, Fund. Math. **49** (1961) 129–141.
[4] P. Erdős, A. Rényi, *On the evolution of random graphs*, Publ. Math. Inst. Hungar. Acad. Sci, **5** (1960) 17–61.
[5] R. Fagin, *Probabilities in finite models*, J. Symbolic Logic, **41** (1976) 50–58.
[6] R. Fraïssé, *Sur quelques classifications des systèmes de relations*, Université d'Alger, Publications Scientifiques, SérieA(1): 35–182, 1954.
[7] Y. V. Glebskii, D.I. Kogan, M. I. Liogon'kii, V. A. Talanov, *Range and degree of realizability of formulas in the restricted predicate calculus*, Cybernetics, **5**:2 (1969) 142–154.
[8] M. M. Gösgens, L. Lüchtrath, E. Magnanini, M. Noy, É. De Panafieu, *The Erdős–Rényi graph conditioned on every component being fully connected*, The 12th International Conference on Complex Networks and Their Applications, 2023.
[9] S. Haber, M. Krivelevich, *The logic of random regular graphs*, J. Comb., **1**:3–4 (2010) 389–440.
[10] W. Hanf, *Model-theoretic methods in the study of elementary logic*, The theory of models, North-Holland, 2014, 132–145.
[11] P. Heinig, T. Müller, M. Noy, A. Taraz, *Logical limit laws for minor-closed classes of graphs*, Journal of Combinatorial Theory, Series B., **130** (2018) 158–206.
[12] S. Janson, T. Łuczak, A. Ruciński, Random Graphs, John Wiley & Sons, 2000.
[13] N. Immerman, *Descriptive complexity*, Springer Science & Business Media, 2012.
[14] N. Immerman, *Languages which capture complexity classes*, Proceedings of the 15th annual ACM symposium on Theory of computing (STOC'83), 1983, 347–354.
[15] R. D. Kleinberg, J. M. Kleinberg, *Isomorphism and embedding problems for infinite limits of scale-free graphs*, In Proc. of the 16th ACM-SIAM Symposium on Discrete Algorithms (SODA'05), 2005, 277–286.
[16] V. F. Kolchin, *A problem of the allocation of particles in cells and random mappings*, Theory of Probability & Its Applications, **21**:1 (1976) 48–63.
[17] L. A. Larrauri, *Probabilities of first-order sentences on sparse random relational structures: An application to definability on random CNF formulas*, Journal of Logic and Computation, **31**:2 (2021) 444–472.
[18] A. Larrauri, T. Müller, M. Noy, *Limiting probabilities of first order properties of random sparse graphs and hypergraphs*, Random Structures & Algorithms, **60**:3 (2022) 506–526.
[19] L. Libkin, Elements of Finite Model Theory. Texts in Theoretical Computer Science. An EATCS Series, Springer, 2004.
[20] J. F. Lynch, *Convergence law for random graphs with specified degree sequence*, ACM Transactions on Computational Logic, **6**:4 (2005) 727–748.
[21] J. F. Lynch, *Probabilities of sentences about very sparse random graphs*, Random Structures & Algorithms, **3**:1 (1992) 33–53.
[22] Y. Malyshkin, M. Zhukovskii, *γ-variable first-order logic of uniform attachment random graphs*, Discrete Mathematics, **345** (2022) 112802.
[23] Y. V. Matiiasevich, Hilbert's tenth problem, MIT press, 1993.
[24] Y. V. Matiiasevich, *The Diophantineness of enumerable sets*, Doklady of Russian Academy of Sciences, **191**:2 (1970) 279–282.
[25] G. L. McColm, *First order zero-one laws for random graphs on the circle*, Random Structures and Algorithms, **14**:3 (1999) 239–266.
[26] G. L. McColm, *MSO zero-one laws on random labelled acyclic graphs*, Discrete Mathematics, **254** (2002) 331–347.
[27] L. B. Ostrovsky, M. E. Zhukovskii, *Monadic second-order properties of very sparse random graphs*, Annals of pure and applied logic, **168**:11 (2017) 2087–2101.
[28] D. Poole, *On the strength of connectedness of a random hypergraph*, Electron. J. Combin., **22**:1 (2015) P1.69.
[29] S. Shelah, J. H. Spencer, *Zero-one laws for sparse random graphs*, J. Amer. Math. Soc., **1** (1988) 97–115.
[30] J. Spencer, *The strange logic of random graphs*, Springer, 2001.
[31] J. H. Spencer, *Threshold spectra via the Ehrenfeucht game*, Discrete Applied Math., **30** (1991) 235–252.
[32] N.M. Sveshnikov, M.E. Zhukovskii, *First order zero-one law for uniform random graphs*, Sbornik Mathematics, **211**:7 (2020) 956–966.
[33] O. Verbitsky, M. Zhukovskii, *On the First-Order Complexity of Induced Subgraph Isomorphism*, In 26th EACSL Annual Conference on Computer Science Logic (CSL 2017) **82**, 40:1–40:16.
[34] O. Verbitsky, M. Zhukovskii, *The Descriptive Complexity of Subgraph Isomorphism Without Numerics*, Theor. Comp. Sys., **63**:4 (2019) 902–921.
[35] P. Winkler, *Random Structures and Zero-One Laws*, Finite and Infinite Combinatorics in Sets and Logic, N.W. Sauer, R.E. Woodrow and B. Sands, eds., NATO Advanced Science Institute Series, Kluwer Academic Publishers, Dordrecht (1993) 399–420.
[36] N.C. Wormald, *Models of random regular graphs* In J. D. Lamb, D. A. Preece (Eds.), Surveys in Combinatorics, 1999 (Vol. 267, pp. 239–298). (London Mathematical Society Lecture Note Series).
[37] N. C. Wormald , *Some problems in the enumeration of labelled graphs*, Bulletin of the Australian Mathematical Society, **21**:1 (1980) 159–160.
[38] N. C. Wormald, *The asymptotic distribution of short cycles in random regular graphs*, Journal of Combinatorial Theory, Series B, **31**:2 (1981) 168–182.

A PROOFS OF FO EXPRESSIBILITY

Cartesian products. Here we present a FO sentence φ that describes the property of being isomorphic to $K_s \square K_t$, for some $s, t > 0$. We construct this sentence as a conjunction of three sentences describing the following properties:

(a) For each vertex v, its neighbourhood consists of two disjoint cliques A_v and B_v;
(b) For every pair of non-adjacent vertices u and v, there is a unique edge from u to A_v and a unique edge from u to B_v;
(c) For every vertex v and any two its non-adjacent neighbours $x \neq y$, there is a unique vertex $u \neq v$ adjacent to x and y.

Let us prove that this sentence expresses the desired property. Observe that $K_s \square K_t$ satisfies it. Let us then consider any graph G satisfying this sentence. Consider a vertex v, from (a) we have two cliques A_v and B_v. For any vertex x in A_v, we have two cliques $A_x = A_v \cup \{v\} \setminus \{x\}$ and B_x. The clique B_x does not intersect A_x by the property (a), and does not intersect B_v because there is no edges between x and B_v. Also, we conclude that each vertex adjacent to x and not adjacent to v is a vertex of B_x. Similarly, for $y \in B_v$, $B_y = B_v \cup \{v\} \setminus \{y\}$, A_y does not intersect both B_y and A_v, and each vertex adjacent to y and not adjacent to v is a vertex of A_y.

From (b) and (c), we have that vertices u which are not adjacent to v are in a one-to-one correspondence with pairs of vertices $x \in A_v$ and $y \in B_v$: each such u is adjacent to the respective x and y, and it is not adjacent to any other vertex in $A_v \cup B_v$. Therefore, for each pair $x \in A_v$ and $y \in B_v$, cliques B_x and A_y have a unique common vertex u. This implies that all cliques B_x have the same size as B_v, and all cliques A_y have the same size as A_v. Hence, the graph G consists of cliques $\{v\} \cup A_v$ and $\{y\} \cup A_y$, for all $y \in B_v$, and cliques $\{v\} \cup B_v$ and $\{x\} \cup B_x$, for all $x \in A_v$, i.e. isomorphic to $K_s \square K_t$, for some $s, t > 0$.

The property $\cup\Gamma_m$. Let us show that the property $\bigcup_m \Gamma_m$ is FO, and define φ_1 as a sentence describing this property. Let $Type_{i_1,\ldots,i_j}(x)$ be a FO formula saying that x has degree j and adjacent to vertices of degrees $i_1, \ldots, i_j$. The sentence φ_1 is conjunction of the following sentences:

- $\forall x\ Type_2(x) \vee Type_3(x) \vee Type_{1,3}(x) \vee Type_{3,3}(x) \vee$
  $\vee Type_{1,2,3}(x) \vee Type_{2,2,3}(x);$
- $\forall x\ Type_{1,2,3}(x) \Rightarrow \left(\exists y\ Type_{2,2,3}(y) \wedge (x \sim y)\right);$
- $\forall x\ Type_{2,2,3}(x) \Rightarrow \left(\exists y\ Type_{1,2,3}(y) \wedge (x \sim y)\right);$
- $\forall x\ (Type_{2,2,3}(x) \vee Type_{1,2,3}(x)) \Rightarrow$
  $\Rightarrow \left(\exists! y\ Type_{3,3}(y) \wedge (x \sim y)\right);$
- $\forall x\ Type_{3,3}(x) \Rightarrow$
  $\Rightarrow \left(\exists y \exists z\ Type_{1,2,3}(y) \wedge Type_{2,2,3}(z) \wedge (x \sim y) \wedge (x \sim z)\right).$

Let us prove that φ_1 indeed describes $\bigcup_m \Gamma_m$. Obviously, any graph having the property Γ_m satisfies φ_1. Let $G \models \varphi_1$. We have to show that $G \in \bigcup_m \Gamma_m$. Let us first show that all types listed in the first clause of of φ_1 are presented in G. All vertices of G have degree at most 3. Let us show that there is a vertex of degree exactly 3. Suppose the opposite, then there is no vertex of degree 2 because φ_1 forces each vertex of degree 2 to be adjacent to a vertex of degree 3. Therefore, all vertices have degree 1 that leads to contradiction because due to the definition of φ_1 vertices of degree 1 cannot be adjacent to other vertices of degree 1. Next, note that the presence of a degree 3 vertex of any of the two types implies the presence of a degree 3 vertex of the second type. Vertex that satisfies $Type_{1,2,3}$ has neighbours that satisfy $Type_3$ and $Type_{3,3}$. Vertex that satisfies $Type_{2,2,3}$ has exactly one neighbour that satisfies $Type_{3,3}$. Hence, another neighbour of degree 2 satisfies $Type_{1,3}$. Finally, a vertex that satisfies $Type_{1,3}$ is adjacent to a vertex with degree 1 that satisfies $Type_2$, i.e. all types allowed by φ_1 are indeed presented.

Let us reconstruct $D \in \Delta_m$ such that $G \cong G(D)$. It would imply that $G \in \Gamma_m$. We set the vertices of D to be the vertices of G that satisfy $Type_{3,3}$. We add an edge (u, v) between (not necessarily distinct) vertices of D if and only if there is a path $uxyv$ in G such that $Type_{1,2,3}(x) \wedge Type_{2,2,3}(y)$ holds. Due to φ_1, each vertex u that satisfies $Type_{3,3}(u)$ is adjacent to a unique vertex x that satisfies $Type_{1,2,3}(x)$; the vertex x is adjacent to a unique vertex y that satisfies $Type_{2,2,3}(y)$; the vertex y is adjacent to a unique vertex v (not necessarily $v \neq u$) that satisfies $Type_{3,3}(y)$. So, for each vertex of D its out-degree equals 1. Similarly, the in-degree of each vertex in D equals 1. Therefore, the reconstructed digraph is in Δ_m. It remains to show that the graph G is isomorphic to $G(D)$, and thus belongs to Γ_m. Indeed, vertices that satisfy $Type_{3,3}$ are the vertices of D. By the definition of edges in D, we have the respective vertices $w_1(u, v)$ and $w_3(u, v)$, and the conditions on types of these vertices guarantee the presence of properly connected vertices $w_2(u, v), w_4(u, v), w_5(u, v)$ in G. All vertices $w_i(u, v)$ are different because for different i they have different types, and u, v can be restored as the two (or one $u = v$) nearest vertices that satisfy $Type_{3,3}$. There are no other vertices due to the first clause in the definition of φ_1.

The sentence φ_L. Let us show that there is a FO sentence φ_L that expresses the property of being isomorphic to L_m for some $m \geq 2$. Let φ_L = Types ∧ UDeg ∧ VDeg ∧ VUEdges ∧ VUSquare ∧ $\text{UVPattern}_1 \wedge \text{UVPattern}_2 \wedge \text{UVPattern}_3$, where the clauses in the conjunction are defined as follows.

- $\text{Deg}_i(x)$ ($\text{Deg}_{\geq i}(x)$) is a FO formula that expresses the property of the vertex x to have degree i (at least i).
- $W(x) = \text{Deg}_1(x) \wedge \left(\exists y\ (x \sim y) \wedge \text{Deg}_{\geq 3}(y)\right)$.
- $V(x) = \exists! y\ (x \sim y) \wedge W(y)$.
- $U(x) = \neg W(x) \wedge \neg\exists y\ (x \sim y) \wedge W(y)$.
- $\text{Types} = \forall x\ W(x) \vee V(x) \vee U(x)$.
- UDeg is a FO sentence saying that the induced subgraph on $\{x : U(x)\}$ consists of vertices with degrees 1 or 2, and exactly two of them have degree 1.
- VDeg is a FO sentence saying that the induced subgraph on $\{x : V(x)\}$ consists of vertices of degrees 1 or 2.
- $\text{VUEdges} = \forall x\ (V(x) \Rightarrow (\exists! y\ U(y) \wedge (x \sim y)))$. This sentence defines a mapping $\text{UVMap} : \{x : V(x)\} \to \{x : U(x)\}$.
- VUSquare is a FO sentence saying that, for any pair of adjacent vertices x, x' satsifying $V(x) \wedge V(x')$, their images $\text{UVMap}(x)$ and $\text{UVMap}(x')$ are adjacent as well.

It remains to define UVPattern_1, UVPattern_2 and UVPattern_3. For FO formulae ψ_1, ψ_2 with a free variable x, let

$$\begin{aligned}\text{Matching}_x[\psi_1, \psi_2] =& \forall x\ \neg(\psi_1(x) \wedge \psi_2(x)) \wedge \\ & \left(\psi_1(x) \Rightarrow \exists!\ x'(x \sim x') \wedge \psi_2(x')\right) \wedge \\ & \left(\psi_2(x) \Rightarrow \exists!\ x'(x \sim x') \wedge \psi_1(x')\right)\end{aligned}$$

be the formula saying that the set A_1 of x satisfying $\psi_1(x)$ and the set A_2 of x satisfying $\psi_2(x)$ are disjoint and edges between them form a perfect matching.

- $\text{VMatching}_x(y, y') = \text{Matching}_x[V(x) \wedge (x \sim y), V(x) \wedge (x \sim y')]$.
- $\text{VAlmostMatching}(y, y')$ is the FO sentence

$$\begin{aligned}&\exists z'\ V(z') \wedge (y' \sim z') \wedge \\ &\quad \wedge \left(\forall z\ (V(z) \wedge (y \sim z)) \Rightarrow \neg(z \sim z')\right) \wedge \\ &\quad \wedge \text{Matching}_x[V(x) \wedge (x \sim y), V(x) \wedge (x \sim y') \wedge (x \neq z')].\end{aligned}$$

  In other words, the induced bipartite graph between $A = \{x : V(x) \wedge (x \sim y)\}$ and $B = \{x : V(x) \wedge (x \sim y')\}$ is a disjoint union of a matching and an isolated vertex $z' \in B$.
- UVPattern_1 is a FO sentence saying that, for each $y_0, y_1, \ldots, y_6$ satisfying

$$\bigwedge_{0 \leq i \leq 6} U(y_i) \wedge \bigwedge_{0 \leq i \leq 5} (y_i \sim y_{i+1}) \wedge \bigwedge_{0 \leq i \leq 4} (y_i \neq y_{i+2}), \qquad (3)$$

  there are five $i \in [6]$ such that $\text{VMatching}(y_{i-1}, y_i)$ holds and for the single remaining $i \in [6]$, at least one of the formulae $\text{VAlmostMatching}(y_{i-1}, y_i)$ and $\text{VAlmostMatching}(y_i, y_{i-1})$ is satisfied.

In order to define UVPattern_2 and UVPattern_3, we need auxiliary formulae UStart and UEnd:

- $\text{UStart}(x) = U(x) \wedge Deg_1(x)$.
- $\text{UEnd}(x) = U(x) \wedge Deg_{\geq 2}(x) \wedge (\exists! y\ U(y) \wedge (x \sim y))$.
- Let UVPattern_2 be a FO sentence saying that, for each $y_0, y_1, \ldots, y_6$ satisfying (3),

$$\begin{aligned}&(\text{UStart}(y_0) \Rightarrow \text{VAlmostMatching}(y_5, y_6)) \wedge \\ &\qquad \wedge (\text{UEnd}(y_6) \Rightarrow \text{VAlmostMatching}(y_0, y_1)),\end{aligned}$$

  is satisfied, and there are vertices y and y' satisfying $\text{UStart}(y)$ and $\text{UEnd}(y')$.
- Let UVPattern_3 be a FO sentence saying that, for each $y_0, y_1, \ldots, y_7$ satisfying

$$\bigwedge_{0 \leq i \leq 7} U(y_i) \wedge \bigwedge_{0 \leq i \leq 6} (y_i \sim y_{i+1}) \wedge \bigwedge_{0 \leq i \leq 5} (y_i \neq y_{i+2}),$$

  the formula

$$\text{VAlmostMatching}(y_0, y_1) \Rightarrow \text{VAlmostMatching}(y_6, y_7)$$

  is satisfied.

Let us briefly verify that φ_L expresses the desired property of being isomorphic to some L_m. We skip a direct routine check that L_m satisfies all clauses in the definition of φ_L. Let $G \models \varphi_L$ and let us prove that $G \cong L_m$ for a certain m. By UDeg, we know that the subgraph G_U induced on vertices x that satisfy $U(x)$ is a disjoint union of a single path and cycles.

Suppose that there is a cycle $y_0 \ldots y_{s-1} y_0$ in G_U for some $s \geq 3$. Then, without loss of generality, the sentence UVPattern_1 implies that $\text{VAlmostMatching}(y_0, y_1)$ holds. For the sake of convenience, set $y_k := y_{k-s}$ for $k \geq s$. For every $i \in \mathbb{Z}_{\geq 0}$, denote by n_i the number of neighbours x of y_i that satisfy $V(x)$. Then, $\text{VMatching}(y_i, y_{i+1})$ implies $n_{i+1} = n_i$, and $\text{VAlmostMatching}(y_i, y_{i+1})$ implies $n_{i+1} = n_i + 1$. Due to $\text{UVPattern}_1 \wedge \text{UVPattern}_3$, for each i divisible by 6, $n_{i+1} = n_i + 1$, and $n_{i+1} = n_i$ for other. Therefore, the sequence

$n_0, n_6, n_{12}, \ldots, n_{6s}$ is strictly increasing. It leads to contradiction because $y_0 = y_{6s}$ and so $n_0 = n_{6s}$. Therefore, G_U is a path.

By VDeg, we have that there are no vertices of degree more than 2 in the subgraph G_V induced on $\{x : V(x)\}$. Hence, this graph is a disjoint union of paths and cycles. By VUEdges $\wedge$ VUSquare, we have the mapping UVMap : $G_V \to G_U$ that sends adjacent vertices to adjacent. Also, by UVPattern$_1$, UVMap sends two different vertices with a common neighbour in G_V to two different vertices in G_U. Indeed, if x_1, x_2 are neighbours of x in G_V and y is a common neighbour of x_1, x_2 from G_U, then there is a vertex $y' \in G_U$ which is a common neighbour of x and y. This contradicts UVPattern$_1$ since, if the latter holds, then either VMatching(y, y'), or VAlmostMatching(y, y'), or VAlmostMatching(y', y). All these predicates do not allow vertices of degree more than 1 in the bipartite graph between the neighbourhoods of y and y' in G_V. Since the graph induced on G_U is a path, there are no cycles in G_V, i.e. G_V is a disjoint union of paths, where each path is mapped by UVMap to a subpath in G_U. By UVPattern$_2$ $\wedge$ UVPattern$_3$, we have that these paths in G_V behave as in L_m in the following sense: the induced subgraph on $V(G_U \cup G_V)$ is isomorphic to the induced subgraph on u_i and $v_{i,j}$ in L_m; vertices of G_U correspond to vertices u_i and vertices of G_V correspond to vertices $v_{i,j}$. By Types, remaining vertices have degree one and adjacent to vertices in G_V. Since for each $x \in G_V$ there is exactly one neighbor y that satisfies $W(x)$, we have the correspondence between such vertices y and vertices $w_{i,j}$ in L_m that completes the isomorphism between G and L_m.

B WINNING STRATEGIES ON TRANSFORMED GRAPHS

Claim B.1. *Let Spoiler has no winning strategy in k moves on graphs G_1 and G_2. Then, for each graph H, there is no winning strategy for Spoiler in k moves on graphs $G_1 \square H$ and $G_2 \square H$.*

Proof. There is a winning strategy for Duplicator in k moves on G_1 and G_2. Let us construct a winning strategy for $G_1 \square H$ and $G_2 \square H$. Suppose in the first $t-1$ rounds vertices $(u_1, v_1), \ldots, (u_{t-1}, v_{t-1})$ from $G_1 \square H$ and $(u'_1, v'_1), \ldots, (u'_{t-1}, v'_{t-1})$ from $G_2 \square H$ are chosen. Without loss of generality, suppose that Spoiler chooses $(u_t, v_t) \in G_1 \square H$ in round t. Let $u'_t \in G_2$ be the Duplicator's choice in round t according to the strategy on G_1 and G_2 when $u_1, \ldots, u_t \in G_1$ and $u'_1, \ldots, u'_{t-1}$ are chosen. Then, Duplicator chooses (u'_t, v_t) in the game on $G_1 \square H$ and $G_2 \square H$.

Following this strategy, we have that $v'_i = v_i$, for all $i \leq k$ after k rounds. Also, vertices $u_1, \ldots, u_k \in G_1$ and $u'_1, \ldots, u'_k \in G_2$ are chosen according to the winning strategy of Duplicator in the game on G_1, G_2. Therefore,

$$\begin{aligned}(u_i, v_i) &\sim (u_j, v_j)\\ &\Leftrightarrow ((u_i \sim u_j) \wedge (v_i = v_j)) \vee ((u_i = u_j) \wedge (v_i \sim v_j))\\ &\Leftrightarrow ((u'_i \sim u'_j) \wedge (v'_i = v'_j)) \vee ((u'_i = u'_j) \wedge (v'_i \sim v'_j))\\ &\Leftrightarrow (u'_i, v'_i) \sim (u'_j, v'_j),\end{aligned}$$

so indeed Duplicator wins the game on $G_1 \square H$ and $G_2 \square H$. ◻

Claim B.2. *Let Y, G, H be a triple of graphs with induced subgraphs $X_0 \subset Y$, $X_1 \subset G$, $X_2 \subset H$ and isomorphisms $f : X_0 \to X_1$, $g : X_0 \to X_2$. Denote by G_X and H_X the relational structures with the usual adjacency relation in G and H and a unary relation I_x for each vertex $x \in X_0$ such that $I_x(y) \Leftrightarrow (y = f(x))$ for each $y \in G$ and $I_x(z) \Leftrightarrow (z = g(x))$ for each $z \in H$. Let $G \cup_f Y$ be the graph obtained from G and Y by identifying every $x \in X_0$ with $f(x) \in X_1$ and adding the rest of Y to G vertex-disjointly. Similarly define $H \cup_g Y$. Assume that Duplicator has a winning strategy in k moves on G_X and H_X, then Duplicator has a winning strategy in k moves on $G \cup_f Y$ and $H \cup_g Y$ as well.*

Proof. Denote by $Y_1 \subset G \cup_f Y$ and $Y_2 \subset H \cup_g Y$ the two copies of Y attached to G and H respectively, and let $h : Y_1 \to Y_2$ be an isomorphism between them such that its restriction to X_1 coincides with $g \circ f^{-1}$. Suppose that in first $t-1$ rounds vertices $u_1, \ldots, u_{t-1} \in G \cup_f Y$ and $u'_1, \ldots, u'_{t-1} \in H \cup_g Y$ are chosen. Without loss of generality, Spoiler chooses a vertex $u_t \in G \cup_f Y$ in round t. If $u_t \in G$, Duplicator chooses the vertex $u'_t \in H$ according to a winning strategy on G_X and H_X. If $u_t \in Y_1$, Duplicator chooses $h(u_t) \in Y_2$. Note that this strategy is well-defined since for $u_t \in X_1 = G \cap Y_1$, the vertex chosen according to the winning strategy on G_X and H_X has to be $h(u_t)$ because of the unary predicate $I_{f^{-1}(u_t)}$.

Let us prove that this is a winning strategy of Duplicator. Consider two chosen pairs of vertices $u_i, u_j \in G \cup_f Y$ and $u'_i, u'_j \in H \cup_g Y$. If $u_i, u_j \in G$, then $u'_i, u'_j \in H$ and $(u_i \sim u_j) \Leftrightarrow (u'_i \sim u'_j)$ since this vertices were chosen according to the winning strategy of Duplicator in the game on G_X and H_X. If $u_i, u_j \in Y_1$, then $u'_i, u'_j \in Y_2$ and $(u_i \sim u_j) \Leftrightarrow (u'_i \sim u'_j)$ because $u'_i = h(u_i)$ and $u'_j = h(u_j)$. If $u_i \in G \setminus Y_1, u_j \in Y_1 \setminus G$, then $u'_i \in H \setminus Y_2, u'_j \in Y_2 \setminus H$ and so there are no edges $\{u_i, u_j\}, \{u'_i, u'_j\}$. Therefore, Duplicator wins. ◻

Claim B.3. *Let G, H be a pair of graphs with an additional unary relation I. Denote by G_I and H_I graphs obtained from G and H by attaching exactly one leaf to each vertex x satisfying $I(x)$. Assume that Duplicator has a winning strategy in k rounds on G and H equipped with I, then Duplicator has a winning strategy in k rounds on G_I and H_I as well.*

Proof. Let $f_G : V(G_I) \to V(G)$ be a mapping that acts as identity on vertices of G and sends each vertex in $V(G_I) \setminus V(G)$ to its unique neighbour. Similarly, define f_H. Suppose that in the first $t-1$ rounds vertices $u_1, \ldots, u_{t-1} \in G_I$ and $u'_1, \ldots, u'_{t-1} \in H_I$ are chosen. Without loss of generality, Spoiler chooses a vertex $u_t \in G_I$ in round t. Let $v'_t \in H$ be the Duplicator's choice in round t according to the strategy on G and H with unary relation I when $f(u_1), \ldots, f(u_t) \in G$ and $f(u'_1), \ldots, f(u'_{t-1}) \in H$ are chosen. If $u_t \in G$, then Duplicator chooses $u'_t = v'_t$. If $u_t \notin G$, then Duplicator chooses the leaf u'_t attached to v'_t. Note that in both cases $f(u'_t) = v'_t$, so this strategy is well-defined.

Let us prove that this strategy is winning for Duplicator. Since $f(u'_t)$ is chosen according to the winning strategy in the game on G and H with unary relation I, $f(u_i) \sim f(u_j) \Leftrightarrow f(u'_i) \sim f(u'_j)$ and $f(u_i) = f(u_j) \Leftrightarrow f(u'_i) = f(u'_j)$, for $1 \leq i < j \leq k$. Also,

$u_i \in G \Leftrightarrow u_i' \in H$ for $1 \le i \le k$. Therefore,

$$\begin{aligned}(u_i \sim u_j) \Leftrightarrow &((u_i, u_j \in G) \wedge (f(u_i) \sim f(u_j)))\vee\\ &\vee ((u_i \neq u_j) \wedge (f(u_i) = f(u_j)))\\ \Leftrightarrow &((u_i', u_j' \in H) \wedge (f(u_i') \sim f(u_j')))\vee\\ &\vee ((u_i' \neq u_j') \wedge (f(u_i') = f(u_j'))) \Leftrightarrow (u_i' \sim u_j'),\end{aligned}$$

so indeed Duplicator wins the game on G_I and H_I. □

C PROOFS OF TECHNICAL PROPOSITIONS FROM SECTION 3

Claim C.1. *The stochastic FO reduction relation $\le$ is a preorder.*

Proof. Firstly, we prove that $\le$ is reflexive. Let we have a random relational structure D_n. We define a stochastic FO reduction by the identity mapping $id : \mathcal{D} \to \mathcal{D}$. It is clear that this mapping is $(\mathcal{A} \mid \mathcal{A})$-measurable, maps n-structures to n-structures, and satisfies $\lim_{n\to\infty} |\Pr(D_n \models id^{-1}(\varphi)) - \Pr(D_n \models \varphi)| = 0$, implying the reflexivity.

Next, we prove the transitivity. Let we have two stochastic FO reductions $f : \mathcal{D} \to \mathcal{D}'$ from D_n' to D_n and $g : \mathcal{D}' \to \mathcal{D}''$ from D_n'' to D_n'. As a reduction from D_n'' to D_n we use the composition $g \circ f$. A composition of $(\mathcal{A} \mid \mathcal{A}')$-measurable and $(\mathcal{A}' \mid \mathcal{A}'')$-measurable functions is $(\mathcal{A} \mid \mathcal{A}'')$-measurable. Also, a composition of two functions which map n-structures to n-structures, for each $n \in \mathbb{N}$, satisfies the same property. Finally,

$$\begin{aligned}0 &\le \lim_{n\to\infty} |\Pr(D_n \models (g\circ f)^{-1}(\varphi)) - \Pr(D_n'' \models \varphi)|\\ &\le \lim_{n\to\infty} |\Pr(D_n \models f^{-1}(g^{-1}(\varphi))) - \Pr(D_n' \models g^{-1}(\varphi))| +\\ &\qquad + \lim_{n\to\infty} |\Pr(D_n' \models g^{-1}(\varphi)) - \Pr(D_n'' \models \varphi)| = 0.\end{aligned}$$

Hence, we have

$$\lim_{n\to\infty} |\Pr(D_n \models (g\circ f)^{-1}(\varphi)) - \Pr(D_n'' \models \varphi)| = 0,$$

that finishes the proof. □

Claim C.2. *For each random relational structure D_n over the signature σ, there is a loopless random relational structure D_n^* which is equivalent to D_n. Moreover, $G_{loop}(n,p)$ is equivalent to $G_{loop}'(n,p,p)$.*

Proof. Let us give some auxiliary definitions.

- Let $\mathcal{B} = \{B_1, \dots, B_t\}$ be a partition of the set $[a_i]$, for some $1 \le t \le a_i$.
- Let $P_i^{\mathcal{B}}$ be a predicate symbol of arity t.
- Let $\beta : [a_i] \to [t]$ be the mapping such that $k \in B_{\beta(k)}$.
- Let $\beta' : [t] \to [a_i]$ be a mapping such that $\beta(\beta'(k)) = k$.
- Let signature $\bar\sigma$ consist of $=$ and $P_i^{\mathcal{B}}$ over all i and $\mathcal{B}$.
- Let $\bar{\mathcal{D}}$ be the set of all relational structures over the signature $\bar\sigma$.

Let $f : \mathcal{D} \to \bar{\mathcal{D}}$ be a function which maps each structure over the signature σ to a structure over the signature $\bar\sigma$ by assigning each $P_i^{\mathcal{B}}(x_1, \dots, x_t)$ the value $P_i(x_{\beta(1)}, \dots, x_{\beta(a_i)})$, if x_j are pairwise distinct, and zero, otherwise. We let the distribution of the random structure D_n^* be induced by f on $\bar{\mathcal{D}}$ from the probability distribution on $\mathcal{D}$. It is clear, that D_n^* is loopless because every $P_i^{\mathcal{B}}(x_1, \dots, x_t)$ in the image of f is zero on each tuple $(x_1, \dots, x_t)$ with two coinciding $x_j = x_k$. By Claim 3.6, f is a stochastic FO reduction defined by formulae

$$P_i(x_{\beta(1)}, \dots, x_{\beta(a_i)}) \wedge \bigwedge_{1\le i<j\le t} \neg(x_i = x_j).$$

Also, we have a reduction $g : \bar{\mathcal{D}} \to \mathcal{D}$ defined by formulae

$$P_i(x_1, \dots, x_{a_i}) = \\ = \bigvee_{\mathcal{B}} \left(P_i^{\mathcal{B}}(x_{\beta'(1)}, \dots, x_{\beta'(t)}) \wedge \bigwedge_{1\le i<j\le t} \neg(x_{\beta'(i)} = x_{\beta'(j)}) \right).$$

For $G_{loop}(n,p)$, it returns the structure $G_{loop}'(n,p,p)$. □

Lemma C.3. *Let $p, q \in [0,1]$, $p_{\min} = \min\{p, q, 1-p, 1-q\}$, and s_n be a sequence of non-negative integers. Consider random vectors $(\xi_1, \dots, \xi_{s_n})$ and $(\eta_1, \dots, \eta_{s_n})$, where ξ_i and η_i are independent random variables with Bernoulli distribution with the parameter p and q respectively. If $\lim_{n\to\infty} \sqrt{\frac{s_n}{p_{\min}}}|p-q| = 0$, then the total variation distance between the distributions of $(\xi_1, \dots, \xi_{s_n})$ and $(\eta_1, \dots, \eta_{s_n})$ converges to zero.*

Proof. Let $A_n \subset \{0,1\}^{s_n}$. Then,

$$\Pr((\xi_1, \dots, \xi_{s_n}) \in A_n) = \\ = \sum_{(a_1,\dots,a_{s_n})\in A_n} p^{a_1+\dots+a_{s_n}} (1-p)^{s_n - a_1 - \dots - a_{s_n}},$$

where $a_i \in \{0,1\}$, and the same equality with p replaced by q holds for $(\eta_1, \dots, \eta_{s_n})$. Then, letting

$$f_{a_1,\dots,a_n}(p) = p^{a_1+\dots+a_{s_n}} (1-p)^{s_n - a_1 - \dots - a_{s_n}},$$

we get

$$\begin{aligned}&|\Pr((\xi_1, \dots, \xi_{s_n}) \in A_n) - \Pr((\eta_1, \dots, \eta_{s_n}) \in A_n)|\\ &\qquad = \left| \sum_{(a_1,\dots,a_{s_n})\in A_n} (f_{a_1,\dots,a_n}(p) - f_{a_1,\dots,a_n}(q)) \right|\\ &\qquad \le \sum_{(a_1,\dots,a_{s_n})\in A_n} \left| f_{a_1,\dots,a_n}(p) - f_{a_1,\dots,a_n}(q) \right|\\ &\qquad \le \sum_{k=0}^{s_n} \binom{s_n}{k} \left| p^k (1-p)^{s_n-k} - q^k (1-q)^{s_n-k} \right|. \quad (4)\end{aligned}$$

We shall prove that this sum is at most $2C'\sqrt{\frac{s_n}{\pi(p_{\min} - \frac{1}{s_n})}}|p-q|$, for some constant C' and $p_{\min} \ge \frac{3}{2s_n}$. Without a loss of generality, we suppose that $p > q$ and $q \le \frac{1}{2}$. The term $p^k(1-p)^{s_n-k} - q^k(1-q)^{s_n-k}$ is positive if and only if $\left(\frac{p(1-q)}{q(1-p)}\right)^k > \left(\frac{1-q}{1-p}\right)^{s_n}$, i.e.

$$k > k_0 := \left\lfloor s_n \frac{\ln(1-q) - \ln(1-p)}{\ln p + \ln(1-q) - \ln q - \ln(1-p)} \right\rfloor.$$

Let us prove that

$$\lfloor q s_n \rfloor \le k_0 < p s_n. \quad (5)$$

Consider the function $g(x) = x^k(1-x)^{s_n-k}$. The derivative of this function is

$$g'(x) = kx^{k-1}(1-x)^{s_n-k} - (s_n-k)x^k(1-x)^{s_n-k-1} = (k - s_n x)x^{k-1}(1-x)^{s_n-k-1}.$$

For $p \le \frac{k}{s_n}$, $g'(x)$ is positive on the interval (q,p). Therefore, $g(p) - g(q) > 0$. This means that $ps_n > k_0$. Similarly, for $q \ge \frac{k}{s_n}$, $g(p) - g(q) < 0$, and then $\lfloor qs_n \rfloor \le k_0$, completing the proof of (5). We then get

$$\sum_{k=0}^{s_n}\binom{s_n}{k}\left|p^k(1-p)^{s_n-k} - q^k(1-q)^{s_n-k}\right| = \sum_{k=k_0+1}^{s_n}\binom{s_n}{k}\left(p^k(1-p)^{s_n-k} - q^k(1-q)^{s_n-k}\right) - \sum_{k=0}^{k_0}\binom{s_n}{k}\left(p^k(1-p)^{s_n-k} - q^k(1-q)^{s_n-k}\right). \quad (6)$$

Consider the function

$$f(x) = \sum_{k=k_0+1}^{s_n}\binom{s_n}{k}x^k(1-x)^{s_n-k} - \sum_{k=0}^{k_0}\binom{s_n}{k}x^k(1-x)^{s_n-k}.$$

By the Lagrange's mean value theorem, we have that there is a number $t \in (q,p)$ such that

$$\sum_{k=k_0+1}^{s_n}\binom{s_n}{k}\left(p^k(1-p)^{s_n-k} - q^k(1-q)^{s_n-k}\right) - \sum_{k=0}^{k_0}\binom{s_n}{k}\left(p^k(1-p)^{s_n-k} - q^k(1-q)^{s_n-k}\right) = f(p) - f(q) = (p-q)f'(t) = (p-q)\times \sum_{k=k_0+1}^{s_n}\binom{s_n}{k}\left(kt^{k-1}(1-t)^{s_n-k} - (s_n-k)t^k(1-t)^{s_n-k-1}\right) - (p-q)\sum_{k=0}^{k_0}\binom{s_n}{k}\left(kt^{k-1}(1-t)^{s_n-k} - (s_n-k)t^k(1-t)^{s_n-k-1}\right)$$

The last expression equals

$$(p-q)s_n\sum_{k=k_0}^{s_n-1}\binom{s_n-1}{k}t^k(1-t)^{s_n-1-k} - (p-q)s_n\sum_{k=k_0+1}^{s_n-1}\binom{s_n-1}{k}t^k(1-t)^{s_n-1-k} - (p-q)s_n\sum_{k=0}^{k_0-1}\binom{s_n-1}{k}t^k(1-t)^{s_n-1-k} + (p-q)s_n\sum_{k=0}^{k_0}\binom{s_n-1}{k}t^k(1-t)^{s_n-1-k} = 2(p-q)s_n\binom{s_n-1}{k_0}t^{k_0}(1-t)^{s_n-1-k_0}. \quad (7)$$

By Stirling's formula, there exists a constant $C' > 0$ such that, for all non-negative integers $a > b$ and any real $x \in (0,1)$,

$$\binom{a}{b}x^b(1-x)^{a-b} \le C'\sqrt{\frac{a}{2\pi(a-b)b}}\left(\frac{ax}{b}\right)^b\left(\frac{a(1-x)}{a-b}\right)^{a-b} \le C'\sqrt{\frac{a}{2\pi(a-b)b}}$$

since the function $x^b(1-x)^{a-b}$ achieves its maximum at $x = \frac{b}{a}$.

For $p_{\min} \ge \frac{3}{2s_n}$, we have $k_0 < ps_n \le (1-p_{\min})s_n < s_n - 1$. Hence, for $k_0 < s_n - 1$

$$2(p-q)s_n\binom{s_n-1}{k_0}t^{k_0}(1-t)^{s_n-1-k_0} \le 2(p-q)s_nC'\sqrt{\frac{s_n-1}{2\pi(s_n-1-k_0)k_0}} = 2(p-q)C'\sqrt{\frac{s_n-1}{2\pi(1-\frac{k_0+1}{s_n})\frac{k_0}{s_n}}}. \quad (8)$$

The inequality $\left(1-\frac{k_0+1}{s_n}\right)\frac{k_0}{s_n} \ge \frac{1}{2}\left(1-\frac{1}{s_n}\right)\min\left\{\left(1-\frac{k_0+1}{s_n}\right), \frac{k_0}{s_n}\right\}$ holds because the left side is a product of two factors that sum up to $1-\frac{1}{s_n}$, and then the largest one is at least half of this sum. Due to (5), $ps_n > k_0 \ge \lfloor qs_n \rfloor > qs_n - 1$. Therefore,

$$\min\left\{\left(1-\frac{k_0+1}{s_n}\right), \frac{k_0}{s_n}\right\} \ge p_{\min} - \frac{1}{s_n}.$$

Finally, from (4), (6), (7), (8), we get

$$|\Pr((\xi_1,\ldots,\xi_{s_n}) \in A_n) - \Pr((\eta_1,\ldots,\eta_{s_n}) \in A_n)| \le 2C'\sqrt{\frac{s_n}{\pi(p_{\min}-\frac{1}{s_n})}}|p-q|.$$

Since $p_{\min} - \frac{1}{s_n} \ge \frac{1}{3}p_{\min}$, we have the inequality

$$|\Pr((\xi_1,\ldots,\xi_{s_n}) \in A_n) - \Pr((\eta_1,\ldots,\eta_{s_n}) \in A_n)| \le 2C'\sqrt{\frac{3s_n}{\pi p_{\min}}}|p-q|. \quad (9)$$

If $p_{\min} < \frac{3}{2s_n}$, note that

$$|\Pr((\xi_1,\ldots,\xi_{s_n}) \in A_n) - \Pr((\eta_1,\ldots,\eta_{s_n}) \in A_n)| \le 1. \quad (10)$$

Hence, for $|p-q| \ge \frac{1}{s_n}$, we have

$$|\Pr((\xi_1,\ldots,\xi_{s_n}) \in A_n) - \Pr((\eta_1,\ldots,\eta_{s_n}) \in A_n)| \le \sqrt{\frac{3s_n}{2p_{\min}}}|p-q|.$$

For $|p-q|<\frac{1}{s_n}$, we can claim that $p,q\leq\frac{3}{s_n}$. Therefore,

$$\sum_{k=0}^{s_n}\binom{s_n}{k}\left|p^k(1-p)^{s_n-k}-q^k(1-q)^{s_n-k}\right|$$

$$\leq\sum_{k=0}^{s_n}\binom{s_n}{k}\left|p^k(1-p)^{s_n-k}-q^k(1-p)^{s_n-k}\right|+$$

$$+\sum_{k=0}^{s_n}\binom{s_n}{k}\left|q^k(1-p)^{s_n-k}-q^k(1-q)^{s_n-k}\right|$$

$$\leq\sum_{k=0}^{s_n}\binom{s_n}{k}|p^k-q^k|+\sum_{k=0}^{s_n}\binom{s_n}{k}s_nq^k|p-q|$$

$$\leq\sum_{k=0}^{s_n}\binom{s_n}{k}k|p-q|\left(\frac{3}{s_n}\right)^{k-1}+\sum_{k=0}^{s_n}3\binom{s_n}{k}\left(\frac{3}{s_n}\right)^{k-1}|p-q| \tag{11}$$

$$=\sum_{k=0}^{s_n}\binom{s_n}{k}|p-q|\left(\frac{3}{s_n}\right)^{k-1}(k+3)$$

$$\leq s_n|p-q|\sum_{k=0}^{s_n}\frac{3^{k-1}}{k!}(k+3)\leq 2e^3s_n|p-q|$$

$$<2e^3\sqrt{\frac{3s_n}{2p_{\min}}}|p-q|.$$

Let $C''=\max\left\{2C'\sqrt{\frac{3}{\pi}},\sqrt{\frac{3}{2}},2e^3\sqrt{\frac{3}{2}}\right\}$. Then, from (4), (9), (10), (11), we get

$$|\Pr((\xi_1,\ldots,\xi_{s_n})\in A_n)-\Pr((\eta_1,\ldots,\eta_{s_n})\in A_n)|\leq$$
$$\leq C''\sqrt{\frac{s_n}{p_{\min}}}|p-q|,$$

completing the proof of the lemma. □

Claim C.4. *For any $\alpha>\frac{d}{3}$ and each nonnegative integer r, we have $D^{(d)}(n,\frac{c}{n^\alpha})\leq D^{(d+r)}(n,\frac{c}{n^{\alpha+r}})$.*

Proof. It is sufficient for us to combine Corollary 3.14 and reduction defined by a formula, to show the reduction $D^{(d)}(n,\frac{c}{n^\alpha})\leq D^{(d+1)}(n,\frac{c}{n^{\alpha+1}})$, for $\alpha>\frac{d}{3}$. Let P and Q be predicates of arity d and $d+1$ from signatures of $D^{(d)}(n,p)$ and $D^{(d+1)}(n,p)$ respectively. The formula $\exists y\ Q(y,x_1,\ldots,x_d)$ reduces $D^{(d)}(n,\frac{c}{n^\alpha})$ to $D^{(d+1)}(n,1-\sqrt[n]{1-\frac{c}{n^\alpha}})$. Indeed, $P(x_1,\ldots,x_d)$ is true with the probability

$$\frac{c}{n^\alpha}=1-\left(1-\left(1-\sqrt[n]{1-\frac{c}{n^\alpha}}\right)\right)^n,$$

and all such events are independent. Also, we have

$$1-\sqrt[n]{1-\frac{c}{n^\alpha}}=\frac{c}{n^{\alpha+1}}+O(n^{-2\alpha-1}).$$

Then

$$\sqrt{\frac{n^{d+2+\alpha}}{c}}\left|\left(1-\sqrt[n]{1-\frac{c}{n^\alpha}}\right)-\frac{c}{n^{\alpha+1}}\right|=\frac{1}{\sqrt{c}}n^{\frac{d+2+\alpha}{2}}O(n^{-2\alpha-1})$$
$$=O(n^{\frac{d-3\alpha}{2}}).$$

Therefore, by Corollary 3.14, we have an equivalence between $D^{(d+1)}(n,1-\sqrt[n]{1-\frac{c}{n^\alpha}})$ and $D^{(d+1)}(n,\frac{c}{n^{\alpha+1}})$. Note that, if $\alpha>\frac{d}{3}$, then, for each nonnegative integer r, we have that $\alpha+r>\frac{d+r}{3}$. Therefore, in the same way we can apply other $r-1$ reductions. It proves the claim. □

Claim C.5. *For any $\alpha>d-2$, $D^{(d)}(n,\frac{1}{n^\alpha})\leq D^{(d+1)}(n,\frac{\alpha\ln n}{n})$.*

Proof. The formula $\forall y\ \neg Q(y,x_1,\ldots,x_d)$ reduces $D^{(d)}(n,\frac{c}{n^\alpha})$ to $D^{(d+1)}(n,1-\sqrt[n]{\frac{c}{n^\alpha}})$. Let us estimate

$$1-\sqrt[n]{\frac{c}{n^\alpha}}=1-e^{\frac{\ln c-\alpha\ln n}{n}}=\frac{\alpha\ln n-\ln c}{n}+O\left(\left(\frac{\ln n}{n}\right)^2\right).$$

Suppose ε is an arbitrary real number in the interval $(0,1)$. For $c<1-\varepsilon\frac{(\ln n)^3}{n}$ and sufficiently large n, we have $1-\sqrt[n]{\frac{c}{n^\alpha}}>\frac{\alpha\ln n}{n}$. For $c>1+\varepsilon\frac{(\ln n)^3}{n}$ and sufficiently large n, we have $1-\sqrt[n]{\frac{c}{n^\alpha}}<\frac{\alpha\ln n}{n}$. Note that the function $f(c)=\sqrt[n]{\frac{c}{n^\alpha}}$ is continuous on the interval $(0,2)$ and monotone. Then, for sufficiently large n there is unique solution $c=c_n$ of the equation $1-\sqrt[n]{\frac{c}{n^\alpha}}=\frac{\alpha\ln n}{n}$ in the interval $(0,2)$, and $\lim_{n\to\infty}(c_n-1)\frac{n}{(\ln n)^3}=0$. We get $D^{(d)}(n,\frac{c_n}{n^\alpha})\leq D^{(d+1)}(n,\frac{\alpha\ln n}{n})$. Thus, it remains to check the condition of Corollary 3.14 for $D^{(d)}(n,\frac{c_n}{n^\alpha})$ and $D^{(d)}(n,\frac{1}{n^\alpha})$:

$$\sqrt{n^{d+\alpha}}\left|\frac{c_n}{n^\alpha}-\frac{1}{n^\alpha}\right|=n^{\frac{d+\alpha}{2}}o\left(\frac{(\ln n)^3}{n^{\alpha+1}}\right)=o\left((\ln n)^3n^{\frac{d-2-\alpha}{2}}\right),$$

completing the proof of the claim. □

D PROOF OF CLAIM 3.22

We apply the strategy that was used in [18] to derive the analogous statement for unoriented hypergraphs. The proof is based on a Poisson limit theorem for the number of small cycles in random hypergraphs and the validity of the FO 0–1 law subject to the absence of cycles. Let us first recall these auxiliary results.

Claim D.1 (Larrauri, Müller, Noy [18]). *Let $p\sim\frac{c}{n^{d-1}}$ with $c>0$. Set*

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

Let X_n be the total number of cycles in $G^{S_d}(n,p)$. Then

$$\lim_{n\to\infty}\mathbb{E}(X_n)=f(c),$$

$$\lim_{n\to\infty}\Pr(X_n=0)=e^{-f(c)}=e^{\frac{c}{2(d-2)!}}\sqrt{1-\frac{c}{(d-2)!}}.$$

We define a cycle in an d-uniform oriented hypergraph as an oriented hypergraph with a set W of $(d-1)s$ vertices and s hyperedges such that there is no proper $W'\subset W$ inducing at least $\geq\frac{|W'|}{d-1}$ hyperedges in this hypergraph. Let us show that Claim D.1 implies

Corollary D.2. *Let $p \sim \frac{c}{n^{d-1}}$ with $c > 0$. Set*

$$f(c) = \sum_{k \geq 2} \frac{(d(d-1)c)^k}{2k} = \frac{1}{2} \ln \frac{1}{1 - d(d-1)c} - \frac{d(d-1)c}{2}. \quad (12)$$

Let X_n be the total number of cycles in $G^{\{id\}}(n,p)$. Then

$$\lim_{n \to \infty} \mathbb{E}(X_n) = f(c),$$

$$\lim_{n \to \infty} \Pr(X_n = 0) = e^{-f(c)} = e^{\frac{d(d-1)c}{2}} \sqrt{1 - d(d-1)c}.$$

Proof. For each oriented hypergraph G, we consider the unoriented hypergraph G' obtained by erasing orientations: there is a hyperedge $\{v_1, \dots, v_d\}$ in G' if and only if there is an oriented edge $(v_{\sigma(1)}, \dots, v_{\sigma(d)})$ in G, for some $\sigma \in S_d$. The corollary follows from the next two observations.

- For $G \sim G^{\{id\}}(n,p)$, the number of cycles in G equals to the number of cycles in G' whp, since whp G does not contain two different hyperedges on the same set of d vertices.
- If $G \sim G^{\{id\}}(n,p)$, then $G' \sim G^{S_d}(n,p')$, where $p' \sim \frac{d!c}{n^{d-1}}$.

Then, we conclude that the number of cycles in $G^{\{id\}}(n,p)$ has the same distribution as the number of cycles in $G^{S_d}(n,p')$ asymptotically. Then, replacing of c by $cd!$ in the statement of Claim D.1 yields Corollary D.2.

□

Finally, let us state the validity of the FO 0-1 law subject to the absence of cycles.

Lemma D.3. *Let $p \sim \frac{c}{n^{d-1}}$ with $0 < c < \frac{1}{d(d-1)}$. Let X_n be the random variable equal to the total number of cycles in $G^{\{id\}}(n,p)$. Let φ be a FO sentence. Then,*

$$\lim_{n \to \infty} \Pr(G^{\{id\}}(n,p) \models \varphi \mid X_n = 0) \in \{0,1\}.$$

The proof is literally the same as the proof of Lemma 4.7 in [18] that states a similar 0–1 law for unoriented hypergraphs. It is a direct corollary of the fact that the number of connected components in $G^{\{id\}}(n,p)$ isomorphic to any fixed tree is not bounded in probability. So, we omit this proof.

Proof of Claim 3.22. Note that $\frac{1}{d!}c_0^{S_d}$ is the unique positive solution of $e^{-f(c)} = \frac{1}{2}$, where f is defined in (12). By Corollary D.2, for each $c < c_0$, we have that

$$\lim_{n \to \infty} \Pr(X_n = 0) = e^{-f(c)} > \frac{1}{2}.$$

By Lemma D.3, we have that

$$\lim_{n \to \infty} \Pr\left(G^{\{id\}}\left(n, \frac{c}{n^{d-1}}\right) \models \varphi\right) \in [0, 1 - e^{-f(c)}] \cup [e^{-f(c)}, 1]$$

Then, the set $L_c^{\{id\}}$ is not dense in $[0,1]$. □

E COMPLEXITY OF DECIDING THE 0-1 LAW: CONSTRUCTION OF φ_P

We construct the sentence ψ_P explicitly. First of all, we consider

- a FO formula $\mathrm{Deg}_{\geq d}(v)$ saying that v has degree at least d;
- a FO formula $\mathrm{Leaf}_i(v)$ saying that v has exactly $2i+1$ neighbours of degree 1;
- $\mathrm{All} = \forall v \ \mathrm{Deg}_{\geq 1}(v) \wedge \left(\mathrm{Deg}_{\geq 2}(v) \Rightarrow \bigvee_{i=1}^{s} \mathrm{Leaf}_i(v)\right)$.

The sentence All divides vertices into $s+1$ types: vertices of the first type have degree 1, vertices of type $d \in [2, \dots, s+1]$ have exactly $2d+1$ neighbours of degree 1. We then encode the three types of equations from $\mathcal{S}$ by FO sentences in the following way.

- For $1 \leq i < j \leq s$, $\mathrm{Equal}_{i,j}$ states that for each vertex v that satisfies Leaf_i there is a unique vertex u such that $\mathrm{Leaf}_j(u) \wedge (u \sim v)$, and vice versa. That is, there is a perfect matching between the sets of vertices that satisfy Leaf_i and Leaf_j which forces the numbers of these vertices to be equal.
- For $1 \leq i < j < d \leq s$, $\mathrm{Sum}_{i,j,d}$ states that for each vertex v that satisfies Leaf_d there is a unique vertex u such that $(\mathrm{Leaf}_i(u) \vee \mathrm{Leaf}_j(u)) \wedge (u \sim v)$, and vice versa. That is, there is a perfect matching between the sets of vertices that satisfy Leaf_d and $\mathrm{Leaf}_i \vee \mathrm{Leaf}_j$ which forces the respective cardinalities to be equal.
- For $1 \leq i < j < d \leq s$, $\mathrm{Prod}_{i,j,d}$ states that for each vertex v that satisfies Leaf_d there is a unique pair of vertices u, w such that $\mathrm{Leaf}_i(u) \wedge \mathrm{Leaf}_j(w) \wedge (u \sim v) \wedge (u \sim w)$, and vice versa. That is, there is a bijection between the set of vertices that satisfy Leaf_d and the set of pairs of vertices u, w such that $\mathrm{Leaf}_i(u) \wedge \mathrm{Leaf}_j(w)$ which forces the cardinality of the first set to be equal to the product of cardinalities of sets $\{u \mid \mathrm{Leaf}_i(u)\}$ and $\{w \mid \mathrm{Leaf}_j(w)\}$.

Finally

$$\psi_P := \mathrm{All} \wedge \bigwedge_{t_i = t_j} \mathrm{Equal}_{i,j} \wedge \bigwedge_{t_d = t_i + t_j} \mathrm{Sum}_{i,j,d}$$

where the conjunctions are over the respective equations in $\mathcal{S}$. Note that if $G \models \psi_P$ for a certain G, then the system $\mathcal{S}$ has a solution $(t_1, \dots, t_s)$ in positive integers: t_i is the numbers of vertices in G that satisfy Leaf_i. On the other hand if $\mathcal{S}$ has a solution $(t_1, \dots, t_s) \in \mathbb{Z}^s_{>0}$, then a graph G that satisfies ψ_P can be constructed directly: first, consider a disjoint union of sets of vertices $T_1, \dots, T_s$, $|T_i| = t_i$. Then, for every $i \in [s]$, join $2i+1$ leaves to every vertex from T_i. Next, observe that every successive equation from $\mathcal{S}$, but the last one, contains a new variable — thus, we may recursively draw edges in the desired way defined by all equations, but the last one. Finally, if we get a graph that contains at least one edge between T_p and T_q, then it may only happen if a perfect matching between T_p and T_q is already drawn since, otherwise, $|T_p| \neq |T_q|$. On the other hand, if no edge has been added between T_p and T_q, then we add a perfect matching between them and finish the construction of G.

Let us finally observe that any graph that satisfies ψ_P must have even number of vertices: there is even number of vertices involved in isolated edges and any vertex satisfying Leaf_i has $2i+1$ neighbours of degree one.

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009