The Logic of Random Regular Graphs

Simi Haber * Michael Krivelevich †

Abstract

The first order language of graphs is a formal language in which one can express many properties of graphs — known as first order properties. The classic Zero-One law for random graphs states that if p is some constant probability then for every first order property the limiting probability of the binomial random graph $G(n,p)$ having this property is either zero or one. The case of sparse random graphs has also been studied in detail for the binomial random graph model. We obtain results for random regular graphs that match the main results for $G(n,p)$. In particular we prove that if the degree d is linear in the number of vertices n, or if $d = n^\alpha$ for $0 < \alpha < 1$ irrational, then the random d-regular graph $G_{n,d}$ obeys the Zero-One law. On the contrary, if $d = n^\alpha$ for rational $0 < \alpha < 1$, then there is a (theoretically explicit) first order property with no limiting probability in $G_{n,d}$.

1 Introduction

In this paper we discuss the behavior of sets of properties in random regular graphs or, more generally, in random graphs with a given degree sequence. The properties under investigation are those describable in the standard first order language of graphs (to be defined in a few paragraphs). This blend of combinatorics and logic shed light on the connection between random regular graphs and binomial random graphs, and also on the descriptive power of the first order language of graphs.

1.1 Random regular graphs

Along the paper we use standard graph theory notions. We refer the reader to [46], [8] or [13] for recommended general graph theory monographs. A *graph* is a pair (V,E). The first element, V, is simply a set, called the *vertex set*. The members of the vertex set are called *vertices*. The second element, E, is a set

*Sackler School of Mathematical Sciences, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel. E-mail: habbersi@post.tau.ac.il

†Sackler School of Mathematical Sciences, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel. E-mail: krivelev@post.tau.ac.il.
Research supported in part by USA-Israel BSF Grant 2006322, by grant 1063/08 from the Israel Science Foundation, and by a Pazy memorial award.

of unordered pairs of elements of V. The elements of E are called *edges* and E is the *edge set*. We denote an edge by uv where u and v are vertices. Given a graph G we denote its vertex set by $V[G]$ and its edge set by $E[G]$. While in principle the vertex set, and hence the graph, can be infinite, we shall restrict the discussion to finite graphs solely. For a subset of vertices $V' \subset V$ we define $G[V']$ to be the *induced* subgraph of G on V', that is, the graph with vertex set V' and with the edges of E that have both endpoints in V' as the edge set. For a vertex v, we define the *degree* of v, denoted here $d(v)$, as the number of edges containing v. A *d-regular* graph is simply a graph with all degrees equal d. Let $\bar{d} = (d_1, \ldots, d_n)$ be a sequence of integers. A graph with degree sequence $\bar{d}$ is a graph on the vertex set $[n] = \{1, \ldots, n\}$, where the degree of i is d_i. The *random d-regular graph*, $G_{n,d}$, is a probability space on the set of all d-regular graphs in which every graph has the same probability. Similarly, for a degree sequence $\bar{d}$, the *random graph with degree sequence* $\bar{d}$, denoted $G_{n,\bar{d}}$, is a probability space composed of the set of all graphs on the vertex set $[n]$ having degree sequence $\bar{d}$ and endowed with the uniform distribution. For completeness we also define $G(n,p)$, the *binomial random graph model*, which is a probability space on the set of graphs with vertex set $[n]$. The probability in $G(n,p)$ of every graph with $|E|$ edges is $p^{|E|}(1-p)^{\binom{n}{2}-|E|}$. It is convenient and useful to think of $G(n,p)$ as a product space where for every pair of vertices u and v, the probability that uv is an edge is p, independently of other edges. As is now customary, we shall abuse the notation above and use $G_{n,d}$, $G_{n,\bar{d}}$ and $G(n,p)$ both for the probability spaces aforementioned and for samples from these spaces. As a last notational remark we mention that the degree d and degree sequence $\bar{d}$ must be integral, and although the functions we shall use for d and $\bar{d}$ will not be integral in general, we prefer to omit rounding notation in favor of readability. For general reference on random regular graphs we refer the reader to Wormald's survey [48].

While random regular graphs were extensively studied from the late 70's, this model has proven to be substantially more difficult than the binomial model of random graphs. $G_{n,d}$ is not a product space, there is no obvious efficient procedure to sample from it and even the very basic question of the number of such graphs has no simple answer, just to mention a few hurdles. Most of the (hundreds of) results for $G_{n,d}$ were obtained using the configuration or pairing model of random regular graphs introduced by Bender and Canfield in [3] and also by Bollobás in [6]. While this technique handles the intrinsic difficulties of $G_{n,d}$, results obtained by the configuration model are valid only for relatively sparse random regular graphs. While there are results for dense random regular graphs, that is, $G_{n,d}$ with d growing faster than $\sqrt{n}$, and while those results cover a large range of graph properties and invariants, they are still relatively scarce. See, for example. the results of Krivelevich et al. [26], [27], Cooper et al. [11], [12], Boldi and Vigna [5] and Kim et al. [22].

While demonstrating properties of $G_{n,d}$ is hard, the results for dense regular random graphs are very often quite similar to the parallel results for $G(n,p)$ with $p = d/n$. This is intuitive in a way, since when np is asymptotically

larger than $\log n$, then, with probability tending to one when n tends to infinity, $G(n,p)$ is nearly np-regular by Chernoff's bound. Thus, it is natural to look for connections between these models. One approach to the study of this connection was taken by Kim and Vu in [23] where they conjectured that there exists a joint distribution, or a coupling, between a random regular graph $G_{n,d}$ and two binomial random graphs $G(n,p_1)$ and $G(n,p_2)$ such that $p_1 = d/n(1-o(1))$, $p_2 = d/n(1+o(1))$ and asymptotically almost surely $G(n,p_1) \subseteq G_{n,d} \subseteq G(n,p_2)$. Thus it may be seen as the random regular graph is "sandwiched" between two binomial graphs. Unfortunately they were able to prove only a weaker version of their conjecture. Our results may also be seen as a connection between these models.

1.2 First order language of graphs

The first order theory of graphs is a language in which one can describe some, but certainly not all, properties of graphs. The alphabet is composed of the following symbols:

1. Variables, denoted along this paper by lower case Latin letters x, y, z. Variables stand for vertices solely. This is a crucial difference between first order and higher order logics in which one may quantify over relations as well.

2. Relations. There are exactly two of these: adjacency, denoted here by $\sim$, and equality, denoted as usual by $=$. Thus it is possible to write $x = y$ or $y \sim z$.

3. Quantifiers. Again, there are two, the existential $\exists$ and the universal $\forall$. These can be applied only on variables which means that quantification is only possible over vertices.

4. Boolean connectives, like $\wedge, \vee, \neg$ and $\rightarrow$.

Notice that there are no constants in this language. As usual, we shall also use parentheses and punctuation marks to the benefit of readability.

For example, we can write

$$\exists x \exists y \exists z \ \neg(x = y) \wedge \neg(x = z) \wedge \neg(y = z) \wedge x \sim y \wedge x \sim z \wedge y \sim z,$$

which we can interpret as "there exists a triangle". Another example might be "there are no isolated vertices":

$$\forall x \exists y \ \neg(x = y) \wedge x \sim y.$$

We shall use these examples along the introduction for demonstrating other notions.

There are a few limitations on the language. Every predicate must be of finite length, there are no constants in the language, variables stand only for

vertices (in particular, n, the number of vertices, is not in the language), and again, we may only quantify over vertices.

Connectivity, for example, is not first order expressible. It is quite easy to see that standard definitions of connectivity can not be formulated in the first order language. Proving that properties are not first order expressible usually requires some effort. The Ehrenfeucht game described below is often used for that purpose.

For a given predicate A, the *quantifier depth* of A, denoted here by $\mathrm{QD}(A)$, is defined by induction:

- For every two variables x, y, $\mathrm{QD}(x = y) = \mathrm{QD}(x \sim y) = 0$.
- For every predicate A, $\mathrm{QD}(\neg A) = \mathrm{QD}(A)$.
- For every two predicates A, B, $\mathrm{QD}(A \vee B) = \mathrm{QD}(A \wedge B) = \mathrm{QD}(A \to B) = \max(\mathrm{QD}(A), \mathrm{QD}(B))$.
- Finally, for every predicate A and variable x, $\mathrm{QD}(\forall x\ A) = \mathrm{QD}(\exists x\ A) = \mathrm{QD}(A) + 1$.

The quantifier depth of the sentence in the first example above is three, and for the second example it is two.

In expressions of the form $\forall x\ \varphi(x)$ and $\exists x\ \varphi(x)$, the universal and existential quantifiers serve as *bounding operators* for their first operand — the variable appearing immediately after the quantifier symbol. This means that they bound every unbound instance of the quantified variable in their range — the second operand of the quantifier ($\varphi(x)$ in the expressions above). In this case we say that the instances of the quantified variable are *bound*, or simply that the variable is bound. A variable that is not bound is called a *free* variable. Looking at the examples above one may notice that the instances of variables in both examples are always bound. This is no coincidence, a predicate may have free variables, but in that case its truth value may depend on the specific value of these variables. A predicate with no free variables is called a *sentence*, and given a graph it is either "true" or "false", that is, it has a truth value. Therefore, sentences correspond to properties of graphs and they are the object of our study. Predicates with free variables will only appear in order to simplify the writing of sentences. We shall identify a sentence with the corresponding property and call such a property *a first order property*. We denote the set of sentences in the first order language of graphs by FOL.

A collection of sentences that is closed under logical consequence is called a *theory*. In order to illustrate this notion, let T_0 be the minimal theory containing the two examples above. Then T_0 should also include the first order property "If there are exactly four distinct vertices, then the graph is connected", because a graph on four vertices containing a triangle and no isolated vertices is connected. A theory is called *consistent* if it does not contain the sentence "false". A *model* G of a theory T is simply a specific graph (possibly infinite) that satisfies all the properties $A \in T$. A triangle is a model of T_0 (and thus T_0 is a consistent theory). If G is a model of T and $A \in T$ we say that G *models* A and denote

it by $G \models A$ or $G \in A$. Let T be a theory. If for every first order sentence A, either $A \in T$ or $\neg A \in T$, we say that T is a *complete* theory. It is easy to observe that T_0 is not complete.

Let $\mathcal{G} = \mathcal{G}(n)$ be a probability space of graphs with vertex set $[n]$. For a (not necessarily first order) property A, we say that $\mathcal{G}$ has A *asymptotically almost surely* (abbreviated a. a. s.) if $\lim_{n\to\infty} \Pr[\mathcal{G}(n) \text{ has } A] = 1$. The set of all first order properties that a. a. s. hold in $\mathcal{G}(n)$ is the *almost sure theory* of $\mathcal{G}(n)$, denoted by $\mathrm{AST}(\mathcal{G}(n))$. This set is indeed a theory as it is closed under logical consequence. To see this notice that if B is a logical consequence of sentences from the AST, then by Gödel's compactness theorem there is a *finite* proof for B, that is, B is a logical consequence of a finite subset of them, and thus it also a. a. s. holds.

If $\mathrm{AST}(\mathcal{G}(n))$ is complete we say that $\mathcal{G}$ satisfies the *Zero-One law*. In other words, if for every first order property A one has $\lim_{n\to\infty} \Pr[\mathcal{G}(n) \models A] \in \{0, 1\}$ then $\mathcal{G}$ satisfies the Zero-One law. As a trivial example for a Zero-One law we may consider the case where the support of $G(n)$ contains only one graph (for example, the case of $G(n, p)$ with $p = o(n^{-2})$). Let $\mathcal{G}_1(n)$ and $\mathcal{G}_2(n)$ be two probability spaces of random graphs. If for every first order property A

$$\lim_{n\to\infty} \Pr[\mathcal{G}_1(n) \models A] = \lim_{n\to\infty} \Pr[\mathcal{G}_2(n) \models A] \in \{0, 1\},$$

then we say that $\mathcal{G}_1$ and $\mathcal{G}_2$ are *equivalent*.

Remark 1.1. Equivalence between $\mathcal{G}_1$ and $\mathcal{G}_2$ implies that $\mathrm{AST}(\mathcal{G}_1) = \mathrm{AST}(\mathcal{G}_2)$, or that they agree on every first order property. It also means that this $\mathrm{AST}(\mathcal{G}_1)$ is complete so both satisfy the Zero-One law.

Our notion of equivalence is clearly related to the notion of *elementary equivalence* in Model Theory, where two models are said to be elementary equivalent if they agree on every first order sentence. Here however we consider sequences of probability spaces of models.

1.3 Previous results

The first Zero-One law for random graphs was proven by Glebskii et al. in [19] and independently by Fagin in [15]. They considered the binomial random graph model $G(n, p)$ with $0 < p < 1$ constant, and showed that in this setting the Zero-One law holds for every p.

In 1988 Shelah and Spencer [40] studied $G(n, p)$ where $p = p(n)$. They proved that if p satisfies any one of the following conditions:

1. $p(n) \ll n^{-2}$,
2. $n^{-1-1/k} \ll p(n) \ll n^{-1-1/(k+1)}$ for some positive integer k,
3. $n^{-1-\epsilon} \ll p(n) \ll n^{-1}$ for all positive ϵ,
4. $n^{-1} \ll p(n) \ll n^{-1} \ln n$,

5. $n^{-1} \ln n \ll p(n) \ll n^{-1+\epsilon}$ for all positive ϵ,

6. $p(n) = n^{-\alpha}$ for some irrational $0 < \alpha < 1$,

then $G(n,p)$ satisfies the Zero-One law. Notice that if $p(n)$ equals to one of the functions $n^{-1-1/k}, n^{-1}, n^{-1} \ln n$ and $n^{-\alpha}$ for rational $0 < \alpha < 1$, then it is well known that $p(n)$ is a *threshold function*[1] for some first order property, and by 1988 it was common knowledge that there is no Zero-One law in these cases. As for the regions were the Zero-One law fails to hold, Shelah and Spencer proved that for every rational α there exists a first order sentence A such that $\Pr[G(n, p = n^{-\alpha}) \models A]$ has no limit as n grows to infinity. In the region of very sparse graphs (cases 1 – 5 above) the first order behavior is nicer. Lynch [29] demonstrated that for every positive constant λ, if $p = \lambda/n$ then for every first order property A, the limit $\lim \Pr[G(n,p) \models A]$ exists and can be expressed as a function of λ using λ, constants, addition, subtraction, multiplication and base e exponentiation. Spencer and Thoma [43] demonstrated that if $p = \ln n/n + c/n$, then for every first order property A, the limit $\lim \Pr[G(n,p) \models A]$ exists and it is a finite sum of summands of the form $e^{-e^{-c}}(e^{-c})^i/i!$, or one minus such a sum. Further, all such expressions occur as a limit for some first order property A. The behavior for $p = n^{-1-1/k}$ where k is a positive integer is similar. There are a few surveys (and one book) summarizing the results and the techniques used in the proofs of the binomial case. [41] is aimed for logicians while [47] and [45] may be more convenient for graph theorists. An almost complete answer for the question "when does the Zero-One law hold?" was given by Łuczak and Spencer in [28] for the binomial random graph model.

The results mentioned above (and other) motivated further research on other random graph models and with stronger languages. The first order behavior of random geometric graphs[2] were studied in [31] for the one dimensional case and in [1] for the two dimensional torus. In [24] the first order behavior of a preferential attachment model[3] was studied.

In 2005 Lynch [30] considered the behavior of first order properties of graphs in $G_{n,\bar{d}}$ where $\bar{d}$ is required to satisfy some requirements. When applied to regular graphs these requirements enforce d to be constant. His result was partially positive. While the zero-one law does not hold, he was able to show that if $\bar{d}$ satisfies some requirements, which are satisfied in particular by the random regular graph $G_{n,d}$ with degree $d \geq 3$ any constant, then for every first order property of graphs A, the limit $\lim \Pr[G_{n,\bar{d}} \models A]$ exists. The proof utilizes the aforementioned configuration model, and thus it is not suitable for

[1] $p_A(n)$ is a threshold function for a property A if whenever $p \ll p_A$ then $\lim \Pr[G(n,p) \in A] = 0$ and if $p \gg p_A$ then $\lim \Pr[G(n,p) \in A] = 1$. Usually it is the case that at the threshold function one has $0 < \lim \Pr[G(n,p_A) \in A] < 1$, and indeed, for all the functions above there were known first order properties for which the limiting probability was strictly between zero and one.

[2] See [37] for background on Random Geometric Graphs.

[3] See [9] for background on the Web Graph.

random regular graphs with high degree. Lynch asked what happens when the average degree tends to infinity with n.

1.4 Our results

Our first theorem deals with dense regular graphs, and it provides a matching result to the results of Glebskii et al. and of Fagin as a simple consequence. We formulate our result in the more general setting of random graphs with a given degree sequence, and we require the degree sequence to be, in a sense, close to regular.

Definition 1.2. Let $\bar{d} = \bar{d}(n) = (d_1(n), d_2(n), \ldots, d_n(n))$ be a sequence of integers such that for every $1 \leq i \leq n$ one has $0 \leq d_i(n) < n$. Let λ be the *normalized average* of $\bar{d}$, that is, $\lambda = (\sum_{i=1}^{n} d_i)/(n(n-1))$. We say that $\bar{d}$ is *dense and nearly regular* degree sequence (with average degree λ) if the following conditions hold:

1. $\sum_{i=1}^{n} d_i$ is an even integer,
2. $\min\{\lambda, 1-\lambda\} > c/\log n$ for some $c > \frac{2}{3}$,
3. $|\lambda n - d_i| = O\left(n^{1/2+\epsilon}\right)$ uniformly over i for sufficiently small fixed $\epsilon > 0$.

Theorem 1.3. *Let $\bar{d}$ be a dense and nearly regular degree sequence. Then for every constant $0 < p < 1$, the random graphs $G_{n,\bar{d}}$ and $G(n,p)$ are equivalent.*

Notice that the average degree of the degree sequence and the exact value of p play no role in the statement of the theorem.

The following corollary is an immediate consequence of Theorem 1.3 above (see Remark 1.1):

Corollary 1.4. *Let $0 < \delta < 1$ be fixed and let $d = \delta n$. Then $G_{n,d}$ satisfies the Zero-One law.*

The fact that for every $0 < p_1, p_2 < 1$, the binomial random graph models $G(n,p_1)$ and $G(n,p_2)$ are equivalent is a straightforward corollary of Fagin's proof.

Definition 1.5. Let $d = d(n)$ be a function of n tending to infinity with n such that $d = o(n)$. If $\bar{d} = \bar{d}(n) = (d_1(n), d_2(n), \ldots, d_n(n))$ is a sequence of integers such that $|d_i(n) - d| = o(d)$ for every i then we say that $\bar{d}$ is a *sparse and nearly regular* degree sequence around d. When d is clear from the context we may simply call $\bar{d}$ nearly regular.

Theorems 1.6 and 1.8 below deal with sparse regular graphs and yield matching results to the aforementioned theorems of Shelah and Spencer [40].

Theorem 1.6. *Let $0 < \alpha < 1$ be irrational, let $d = n^{1-\alpha}$ and let $p = n^{-\alpha}$. Then for every sparse and nearly regular degree sequence $\bar{d}$ around d it is the case that $G_{n,\bar{d}}$ and $G(n,p)$ are equivalent.*

In a similar manner to Corollary 1.4 we formulate a Zero-One law for sparse random regular graphs:

Corollary 1.7. *Let $0 < \alpha < 1$ be irrational and let $d = n^{1-\alpha}$. Then $G_{n,d}$ satisfies the Zero-One law.*

Theorem 1.8. *Let $0 < \alpha < 1$ be rational and let $d = n^{1-\alpha}$. There exists a first order sentence A such that $\Pr[G_{n,d} \models A]$ has no limit as $n \to \infty$.*

The first two theorems above may be viewed as indistinguishability results: Assume that $d = \delta n$ for constant $0 < \delta < 1$ or that $d = n^{1-\alpha}$ for constant irrational $0 < \alpha < 1$. Then there is no first order property A such that $\Pr[G(n, p = d/n) \models A] \to 1$ and $\Pr[G(n, d) \models A] \to 0$. Simply put, for a range of values of d, no first order property can distinguish between $G(n, p)$ and $G_{n,d}$ when $p = d/n$. Notice that Theorem 1.3 does not require $p = d/n$ and indeed, no first order property can distinguish between $G_{n,\delta n}$ and $G(n, p)$ for any two constants $0 < \delta, p < 1$.

1.5 Notation

Let $G = (V, E)$ be a graph. For a vertex v we denote its neighborhood by $N(v)$, that is, $N(v) = \{u \in V | uv \in E\}$. Let U be a set of vertices. We denote the common neighborhood of U by $N^*(U) = \bigcap_{u \in U} N(u)$. We also use $N^*(u_1, u_2, \dots)$ for $N^*(\{u_1, u_2, \dots\})$ where $u_1, u_2, \dots$ are vertices, and sometimes the hybrid notation $N^*(U, u_1, u_2, \dots)$ for $N^*(U) \cap N^*(u_1, u_2, \dots)$.

We use the standard "Big O" asymptotic notation of Bachmann and Landau. Let $f(n), g(n)$ be two positive functions whose domain is $\mathbb{N}$. We say that $f = O(g)$ if there is a constant C such that $f(n) \leq Cg(n)$ for every integer n. We say that $f = \Omega(g)$ if $g = O(f)$, and that $f = \Theta(g)$ if both $f = O(g)$ and $f = \Omega(g)$. If $f/g \to 0$ as $n \to \infty$ we write $f = o(g)$, if $f = o(g)$ then we also write $g = \omega(f)$. In particular, we may use $f = o(1)$ and $f = \omega(1)$ to denote functions tending to zero and to infinity respectively. If $f/g \to 1$ as $n \to \infty$ we say that $f \approx g$. Finally, we write $f = (1 \pm \epsilon)g$ if there exists a constant N such that $(1 - \epsilon)g(n) \leq f(n) \leq (1 + \epsilon)g(n)$ for every $n > N$.

For a set B we denote the set of all subsets of B by $2^B = \{A | A \subseteq B\}$. The set $\{1, 2, \dots, n\}$ is denoted by $[n]$.

We implicitly assume that n is large enough whenever necessary. As already mentioned, we omit rounding notation. All logarithm are to the natural base $e = 2.718281828\dots$

The rest of the paper is organized as follows: In Section 2 we describe the combinatorial tools that we will use, which are some useful facts about random regular graphs. These are the McKay-Wormald formula for enumeration of dense graphs with a prescribed degree sequence and a lemma about the probability of occurrence of a fixed, small sized set of edges which is proved using a switching argument. In Section 3 we present the logical tools that we need, namely the Ehrenfeucht Game and a corollary of it dealing with sequences of models. In Section 4 we consider dense random regular graphs and prove

Theorem 1.3. Sections 5 and 6 are devoted to the proofs of Theorems 1.6 and 1.8 respectively. The last section contains concluding remarks and open problems.

2 Random regular graphs

This section contains a few properties of random regular graphs that will be used later.

We denote the number of graphs with vertex set $[n]$ and degree sequence $\bar{d}$ by $N(\bar{d})$. Before quoting the McKay-Wormald formula, we wish to mention that the formulation below is similar to Proposition 3.1 of [26] which, in turn, is based on Theorem 3 of [33]. The later also contains an elegant and interesting probabilistic interpretation of this formula.

Lemma 2.1 ([33])**.** *Assume that $\bar{d} = (d_j)_{j=1}^n$ is a dense and nearly regular degree sequence and let $\lambda = (\sum_{j=1}^n d_j)/(n(n-1))$ be the normalized average of $\bar{d}$. Then,*

$$N(d_1, \ldots, d_n) = f(d_1, \ldots, d_n) \left(\lambda^\lambda (1-\lambda)^{(1-\lambda)}\right)^{\binom{n}{2}} \prod_{j=1}^n \binom{n-1}{d_j}$$

where $f(d_1, \ldots, d_n) = O(1)$.

Another technique proven to be useful when dealing with random regular graphs is the so called *edge switching* technique introduced to the random regular graph scene by McKay in [32] in order to obtain general bounds on the probability of a subgraph occurrence. Another, more modern kind of switchings, was introduced by McKay and Wormald in [34] and [35], and this is the kind we apply here. While there are enumeration formulas parallel to Lemma 2.1 also for moderate degrees ([34]), we prefer to apply the edge switching technique directly. We do it in the following lemma, which is a slightly stronger version of a result of Kim et al. (Lemma 2.3. of [22]). Again, we shall formulate both the tools and the results for random graphs with a prescribed degree sequence.

Lemma 2.2. *Assume that $d = d(n)$ is $\omega(1)$ and $o(n)$. Let $\bar{d} = (d_i)$, $1 \le i \le n$, be a degree sequence such that for every i one has $|d_i - d| \le \epsilon d$ where $\epsilon = \epsilon(n)$ is some function tending to zero. Let $\mathscr{E}$ be a fixed collection of edges on the vertex set $[n]$ with $|\mathscr{E}| = t$ such that $t = o(d)$ and let uw be an edge in $\mathscr{E}$. Then*

$$\Pr\left[\mathscr{E} \subset G_{n,\bar{d}}\right] = \frac{d}{n}\left(1 + O\left(\frac{d}{n}\right) + O\left(\frac{t}{d}\right) + O(\epsilon)\right) \cdot \Pr\left[\mathscr{E} \setminus \{uw\} \subset G_{n,\bar{d}}\right].$$

Proof. Let $\mathscr{C}_1$ be the set of all graphs on vertex set $[n]$ with degree sequence $\bar{d}$ that contain $\mathscr{E}$ and let $\mathscr{C}_0$ be the set of all graphs on vertex set $[n]$ with degree sequence $\bar{d}$ containing $\mathscr{E} \setminus \{uw\}$ and not containing the edge uw.

Given a graph $G \in \mathscr{C}_1$ we define an operation called *forward switching* as follows. We choose two edges u_1w_1 and u_2w_2 of G, delete them together with uw and insert three new edges uw_2, u_1w and u_2w_1 instead. We only allow choosing distinct vertices such that the edges u_1w_1 and u_2w_2 are not in $\mathscr{E}$ and the edges uw_2, u_1w and u_2w_1 are not in G. It is not hard to see that the graph obtained from G by forward switching belongs to $\mathscr{C}_0$.

We wish to estimate the number of possible forward switchings. For counting purposes u_1w_1 and w_1u_1 are considered different. We have approximately $n-d$ ways to choose u_1 — any vertex except w and its neighbors in G will do (u is a neighbor of w). Given u_1 we have $d(1+O(\epsilon))$ ways to pick w_1 as any neighbor of u_1, except w (and maybe u), is an option. Therefore there are $nd(1+O(\epsilon)) - d^2(1+O(\epsilon))$ ways of choosing u_1w_1. Since we also require $u_1w_1 \notin \mathscr{E}$ we need to subtract at most $t = |\mathscr{E}|$ choices which we may neglect. All in all we have $nd(1-O(d/n)+O(\epsilon))$ relevant choices for u_1w_1.

Given uw and u_1w_1 we want to choose a vertex u_2 that is not a neighbor of w_1 in G and we can do so in about $n-d$ ways. Given u, w, u_1, w_1 and u_2 there are $d(1+O(\epsilon))$ ways of choosing w_2. Again, we also need to subtract at most t ways resulting in $u_2w_2 \in \mathscr{E}$, and also at most $O(d)$ ways resulting in $w_2u \in G$. Summarizing the above we conclude that the number of forward switchings is $(nd)^2(1-O(d/n)+O(\epsilon))$.

Given a graph G' in $\mathscr{C}_0$ we now define the *reverse switching* operation by deleting edges uw_2, u_1w and u_2w_1 from G' and inserting the edges uw, u_1w_1 and u_2w_2. Again, we allow only choices where all six vertices are distinct, uw_2, u_1w and u_2w_1 are not in $\mathscr{E}$ and also u_1w_1 and u_2w_2 are not edges of G'. Notice that the resulting graph will be a member of $\mathscr{C}_1$.

Now we wish to count reverse switchings. Assume $G' \in \mathscr{C}_0$. Let t_u and t_w be the number of edges from $\mathscr{E}$ containing u or w respectively. Following the same considerations as above we get that the number of ways to choose u_1 and w_2 is approximately $(d(1+O(\epsilon)) - t_w)(d(1+O(\epsilon)) - t_u) - O(d) = d^2(1-O(t/d)+O(\epsilon))$. Given u, w, u_1 and w_2, there are $n-O(d)$ ways of picking w_1 and then there are $d(1+O(\epsilon))$ ways of picking u_2. Subtracting choices that end with u_2 adjacent to w_2 (there are at most d of these) and choices that pick w_1u_2 from $\mathscr{E}$ (no more than t) we get that the number of reverse switchings is $d^3n(1-O(\frac{t}{d})-O(\frac{d}{n})+O(\epsilon))$.

Knowing the number of forward and reverse switchings we get

$$|\mathscr{C}_1|\left(\frac{n}{d}\right)^2\left(1-O\left(\frac{d}{n}\right)+O(\epsilon)\right) = |\mathscr{C}_0|d^3n\left(1-O\left(\frac{d}{n}\right)-O\left(\frac{t}{d}\right)+O(\epsilon)\right),$$

which implies

$$|\mathscr{C}_1|/|\mathscr{C}_0| = \frac{d}{n}\left(1+O\left(\frac{d}{n}\right)+O\left(\frac{t}{d}\right)+O(\epsilon)\right).$$

Now we can estimate $\Pr[uw \in G_{n,\bar{d}} | \mathscr{E} \setminus \{uw\} \subset G_{n,\bar{d}}]$,

$$\Pr[uw \in G_{n,\bar{d}} | \mathscr{E} \setminus \{uw\} \subset G_{n,\bar{d}}] = \frac{|\mathscr{C}_1|}{|\mathscr{C}_0| + |\mathscr{C}_1|} =$$
$$= \frac{|\mathscr{C}_1|/|\mathscr{C}_0|}{1 + |\mathscr{C}_1|/|\mathscr{C}_0|} = \left(1 + O\left(\frac{d}{n}\right) + O(\epsilon)\right) \frac{|\mathscr{C}_1|}{|\mathscr{C}_0|} =$$
$$= \frac{d}{n}\left(1 + O\left(\frac{d}{n}\right) + O\left(\frac{t}{d}\right) + O(\epsilon)\right),$$

and the proof is complete. □

Corollary 2.3. *Let $d, \bar{d}, \epsilon, \mathscr{E}$ and t be as in Lemma 2.2, with t restricted even more by $t = o(\sqrt{\bar{d}})$, $t = o(n/d)$ and $t = o(1/\epsilon)$. Then*

$$\Pr\left[\mathscr{E} \subset G_{n,\bar{d}}\right] = \left(\frac{d}{n}\right)^t (1 + o(1)).$$

Proof. Applying Lemma 2.2 recursively t times we get

$$\Pr\left[\mathscr{E} \subset G_{n,\bar{d}}\right] = \left(\frac{d}{n}\right)^t \left(1 + O\left(\frac{d}{n}\right) + O\left(\frac{t}{d}\right) + O(\epsilon)\right)^t =$$
$$= \left(\frac{d}{n}\right)^t \left(1 + O\left(\frac{dt}{n}\right) + O\left(\frac{t^2}{d}\right) + O(t\epsilon) + \cdots\right),$$

where the next terms are higher powers of the previous terms. By our choice of bounds for t it is easy to see that the whole error tends to zero with n. □

Notice that in particular, when $d = n^{1-\alpha}$ for some $0 < \alpha < 1$, we can apply Corollary 2.3 for sets of edges of polynomial size $t = n^\beta$ for some $\beta < \alpha$ and $\beta < (1-\alpha)/2$, provided that $\epsilon = o(n^{-\beta})$.

For convenience we also state a corollary for a specific setting that we shall use later — the case of $t = O(1)$, that is, the case of $\mathscr{E}$ being of constant size:

Corollary 2.4. *Let $\bar{d}$ be a sparse and nearly regular degree sequence and let $\mathscr{E}$ be a fixed set of edges of constant size. Then*

$$\Pr\left[\mathscr{E} \subset G_{n,\bar{d}}\right] = \left(\frac{d}{n}\right)^t (1 + o(1)).$$

Proof. Simply notice that the requirements of Corollary 2.3 are satisfied. □

Remark 2.5. The last corollaries allow us to transfer first moment arguments from the binomial model of random graphs into the regular and nearly regular cases, given that the underlying events depend on fixed sets of at most t edges. Letting $p = d/n$, the probability for a fixed set of t edges to appear in $G(n,p)$ is p^t. Here we get asymptotically the same result. Thus, if X is a random variable that can be represented as a sum of indicator random variables for events depending on at most t edges, then we can estimate EX in $G(n,p)$ and transfer the result to $G_{n,\bar{d}}$.

Having the last remark in mind, we shall not give detailed proofs when the only change from a standard first moment argument is the use of Corollary 2.3.

We close this section with a simple and useful fact connecting similar probability spaces of graphs.

Fact 2.6. *Consider $G_{n,\bar{d}}$ and let $V' \subset V$ be a set of labelled vertices of size n'. In addition, let $\bar{d}'$ be a degree sequence of length n'. Then*

$$G_{n,\bar{d}|\bar{d}'} = G_{n',\bar{d}'},$$

where $G_{n,\bar{d}|\bar{d}'}$ is the random graph we get by picking a graph from $G_{n,\bar{d}}$ conditioned on the event "the degree sequence of the subgraph spanned by V' is $\bar{d}'$".

Proof. Simply notice that if we pick a graph from $G_{n,\bar{d}|\bar{d}'}$ and remove $V \setminus V'$ we are left with a graph from $G_{n',\bar{d}'}$, and that this mapping is measure preserving. □

3 The Ehrenfeucht game and model equivalence

The Ehrenfeucht Game (or the Ehrenfeucht-Fraïssé Game[4]) is a technique for determining whether two graphs (or, generally, two structures) are elementary equivalent, that is, determining if the two structures agree on exactly the same set of first order sentences. The definition of the game uses no logic and its analysis is combinatorial, which makes it very useful for combinatorialists who study logic.

The Ehrenfeucht game is played by two players called Spoiler and Duplicator which are usually thought of as male and female respectively in order to further distinguish them. The game has k rounds and the number of rounds is known to both players before the game begins. The gameboard is made of two graphs, G_1 and G_2, on disjoint vertex sets. These parameters determine the game which we denote here by $\mathrm{EHR}[G_1, G_2; k]$. Each round has two parts, Spoiler's move followed by Duplicator's move. On the i-th move Spoiler selects the graph on which he wants to play in this round, then he picks a vertex from that graph and marks it i. Now Duplicator picks a vertex from the other graph and marks it also by i. There is no difference between marks made by Spoiler and those made by Duplicator. Also, a vertex may receive more than one mark.

At the end of the game there are k vertices of G_1 that are marked $1, \ldots, k$ which we denote by $x_1, \ldots, x_k$. Notice that the x_i's are not necessarily distinct and that it does not matter who marked them. Similarly, let $y_1, \ldots, y_k$ be the vertices in G_2 that are marked $1, \ldots, k$. To determine the winner we define $f\colon \{x_1, \ldots, x_k\} \to \{y_1, \ldots, y_k\}$ by $f(x_i) = y_i$. If f is an isomorphism then

[4]The back-and-forth method used in the game was developed by Roland Fraïssé in his thesis [16], [17]. It was formulated as a game by Andrzej Ehrenfeucht [14]. The names Spoiler and Duplicator are due to Joel Spencer.

Duplicator wins and otherwise Spoiler wins. In other words, if for every[5] pair of indices $1 \leq i \leq j \leq k$ one has $x_i = x_j \Leftrightarrow y_i = y_j$ and $x_i \sim x_j \Leftrightarrow y_i \sim y_j$, then Duplicator wins. Notice that in order for duplicator to win it is in general not enough for her to make sure that $G_1[\{x_1, \ldots, x_k\}]$ and $G_2[\{y_1, \ldots, y_k\}]$ will be isomorphic — she needs to guarantee that the induced labelled graphs will be *identical*, that is, she needs the specific mapping $f(x_i) = y_i$ to be an isomorphism.

Given the graphs of the gameboard, G_1, G_2, and the number of rounds k, the Ehrenfeucht Game is a finite perfect information game with no draws. Therefore, one of the players has a winning strategy — a metaphorical instruction book containing a response for every move the opponent makes, ending with victory for that player. If Duplicator has a winning strategy for $\mathrm{EHR}[G_1, G_2; k]$ we say that Duplicator wins $\mathrm{EHR}[G_1, G_2; k]$, or that $\mathrm{EHR}[G_1, G_2; k]$ is a win for Duplicator (and similarly for Spoiler). Of course, the fact that a player has a winning strategy does not mean that finding this strategy is easy.

The game is interesting and useful because of the following theorem of Ehrenfeucht [14]:

Theorem 3.1. *If Duplicator wins EHR$[G_1, G_2; k]$ then G_1 and G_2 agree on every sentence in FOL of quantifier depth at most k. That is, if $A \in$ FOL and $\mathrm{QD}(A) \leq k$ then $G_1 \models A \iff G_2 \models A$.*

In this paper we are interested in the first order asymptotic behavior of models of random graphs. The following theorem bridges between the Ehrenfeucht game and the notion of model equivalence defined earlier.

Theorem 3.2. *Let $\mathcal{G}_1(n)$ and $\mathcal{G}_2(m)$ be two probability spaces of random graphs. Assume that for every positive integer k*

$$\lim_{m,n\to\infty} \Pr\left[\textit{Duplicator wins EHR}\left[\mathcal{G}_1(n), \mathcal{G}_2(m); k\right]\right] = 1.$$

Then $\mathcal{G}_1$ and $\mathcal{G}_2$ are equivalent.

Proof. Let A be a sentence in the first order language of graphs, and let $k = \mathrm{QD}(A)$. Fix $\epsilon > 0$. By our assumption there exists N s. t. if $n, m > N$ then

$$\Pr\left[\text{Duplicator wins } \mathrm{EHR}[\mathcal{G}_1(n), \mathcal{G}_2(m); k]\right] > 1 - \epsilon.$$

Therefore, with probability greater than $1 - \epsilon$, either both $\mathcal{G}_1(n) \models A$ and $\mathcal{G}_2(m) \models A$, or $\mathcal{G}_1(n) \not\models A$ and $\mathcal{G}_2(m) \not\models A$. Thus either $\Pr[\mathcal{G}_1(n) \models A] > 1 - \epsilon$ and $A \in \mathrm{AST}(\mathcal{G}_1)$ or $\Pr[\mathcal{G}_1(n) \models A] < \epsilon$ and $\neg A \in \mathrm{AST}(\mathcal{G}_1)$. Now clearly $\mathrm{AST}(\mathcal{G}_1)$ is complete and equals $\mathrm{AST}(\mathcal{G}_2)$. □

[5] Since we consider here simple graphs without loops it is enough to require the above for $1 \leq i < j \leq k$.

4 The dense case

As mentioned in the introduction, the first Zero-One law for random graphs ([19], [15]) dealt with dense random graphs, i. e., graphs with $|E| = \Theta\left(\binom{n}{2}\right)$. In this situation a. a. s., for every two finite sets of vertices we can find a vertex that is adjacent to all the vertices in one set and not the adjacent to any vertex in the other. This fact enables Duplicator to use a simple strategy, all she needs to do is pick a vertex with the same adjacency pattern as the vertex chosen by Spoiler. Essentially, in order for Duplicator to win $\mathrm{EHR}[G_1, G_2; k+1]$ it is enough to require that both graphs of the gameboard will have the following property:

Definition 4.1. We say that a graph G has the *Alice*(k) *property* if for every two disjoint sets of vertices X, Y such that $|X| + |Y| = k$ there is a vertex $v \notin X \cup Y$ adjacent to every vertex of X and not adjacent to any vertex of Y.

The term "Alice's Restaurant Property" was coined by Peter Winkler [47], in reference to a popular song by Arlo Guthrie. The refrain of the song — You can get anything you want at Alice's restaurant — catches the spirit of this property. In [45] the term "Alice's restaurant property" was used to describe graphs having Alice(k) for every k, and the axiom $\Psi(k)$ of [47] is exactly what we call property Alice$(2k)$.

Claim 4.2. *If both G_1 and G_2 have Alice(k) then EHR$[G_1, G_2; k+1]$ is a win for Duplicator.*

Proof. At every round Duplicator picks a vertex with the same pattern of adjacency to the previous vertices as the vertex that Spoiler picked. She can hold this strategy up to round $k+1$. At the end of every round the two subgraphs are isomorphic with $f(x_i) = y_i$ as the isomorphism, and thus Duplicator wins the game. □

Corollary 4.3. *If for every k both $\mathcal{G}_1$ and $\mathcal{G}_2$ a. a. s. have Alice(k) then they have the same Almost Sure Theory and it is complete.*

Proof. This is an immediate consequence of Claim 4.2 combined with Theorem 3.2. □

For constant p it is known (and not too hard to see) that for every constant integer k a. a. s. $G(n,p)$ has Alice(k). We wish to show that the same holds for $G_{n,d=\delta n}$ with constant δ.

Theorem 4.4. *Let $\bar{d}$ be a dense and nearly regular degree sequence. For every fixed integer k, a. a. s. $G_{n,\bar{d}}$ has Alice(k).*

Remark 4.5. For the benefit of clearer presentation we shall prove Theorem 4.4 above for random *regular* graphs, i. e., we shall assume that $d_j = d$ for every $1 \le j \le n$, where $d = \delta n$ and $0 < \delta < 1$ is constant. Nevertheless, the proof will hold, mutatis mutandis, when $\bar{d}$ is any dense and nearly regular degree sequence

since all we require is that the McKay-Wormald formula will be valid. In fact, the requirements in Definition 1.2 come from the requirements of Lemma 2.1. Notice that the almost sure theory is the same for every random graph with dense and nearly regular degree sequence, and that it is the same theory as the almost sure theory of, say, $G(n, 1/2)$.

In order to prove Theorem 4.4 we estimate the size of the common neighborhood of every set of vertices of size k and then use the Inclusion–Exclusion Principle to complete the proof. The first part requires most of the effort and is the content of the following lemma:

Lemma 4.6. *Let $0 < \delta < 1$ and $\epsilon > 0$ be two real numbers and let k be an integer. Then a. a. s. the common neighborhood of every set of k vertices in $G_{n,\delta n}$ is of size $(1 \pm \epsilon)\delta^k n$.*

Proof of Theorem 4.4. Fix $k \in \mathbb{N}$ and $\epsilon > 0$. By Lemma 4.6 a. a. s. for every set of vertices X, where $|X| = s \leq k$, one has $|N^*(X)| = (1 \pm \epsilon)n\delta^s$. Let X, Y be two sets of vertices such that $|X| = s$, $|Y| = t$ and $s + t = k$.

Denote $U = N^*(X)$, then by the Inclusion–Exclusion Principle and Lemma 4.6 we have that

$$\left|\bigcap_{v\in X} N(v) \cap \bigcap_{u\in Y} \overline{N(u)}\right| = |U| - \sum_{u\in Y} |N^*(U,u)| + \sum_{u_1,u_2\in Y} |N^*(U,u_1,u_2)| - \dots$$

$$= (1 \pm \epsilon)n\left(\delta^s - t\delta^{s+1} + \binom{t}{2}\delta^{s+2} - \dots + (-1)^t\binom{t}{t}\delta^{s+t}\right) = (1 \pm \epsilon)n\delta^s(1-\delta)^t.$$

Therefore, a. a. s. there is a vertex adjacent to all the vertices of X and not adjacent to any of Y's vertices. □

For the benefit of clearer presentation we prove Lemma 4.6 for graphs on $n + k$ vertices, that is for $G_{n+k,\delta n}$. Since k is a fixed constant, this change is entirely cosmetic.

Lemma. *Let $0 < \delta < 1$ and $\epsilon > 0$ be two real numbers and let k be an integer. Then a. a. s. the common neighborhood of every set of k vertices in $G_{n+k,\delta n}$ is of size $(1 \pm \epsilon)\delta^k n$.*

Proof of Lemma 4.6. Let $d = \delta n$ and let $K = \{n + 1, \dots, n + k\}$. In order to estimate the number of d-regular graphs on $[n + k]$ where the size of the common neighborhood of K deviates from $\delta^k n$ by more than $\epsilon\delta^k n$, we map d-regular graphs into another space, one that admits easier enumeration. Let G be a d-regular graph on $[n + k]$.

The mapping is done by labelling every vertex v from the first n vertices of G with the set $L(v) = N(v) \cap K$ and then removing K. The result is a pair (G_0, L) consisting of a graph $G_0 = G[[n]]$ and the labelling function $L : [n] \to 2^K$ described above.

Notice that for each vertex in $[n]$, say v, we have $d(v) = d - |L(v)|$. Given such a pair (G_0, L), we denote the number of vertices with label A by x_A and

use $\bar{x} = \bar{x}(G_0, L)$ for the vector $(x_\varnothing, x_{\{n+1\}}, \ldots, x_K) \in \{0, 1, \ldots, d\}^{2^k}$. We also use the term $\bar{x}$-graph for a pair (G_0, L) with $\bar{x}(G_0, L) = \bar{x}$.

We shall assume from now on that there are no edges between the last k vertices of the original graph G. Arguments similar to the following may be stated for every one of the $\binom{k}{2}$ possible induced graphs on K. Formally we condition on $G[K]$ and apply the law of total probability.

Given our last assumption, it is rather easy to observe that this mapping is reversible and thus it is an injection. To make it a surjection as well we need to restrict its range. For every $i \in K$ let $\mathcal{A}_i = \{A \in 2^K | i \in A\}$. The mapping maps G into a pair (G_0, L) satisfying

$$\sum_{A \in 2^K} x_A = n, \tag{1a}$$

and

$$\sum_{A \in \mathcal{A}_i} x_A = d, \tag{1b}$$

for every $i \in K$. Equation (1b) holds since G is d-regular (in the general case we will have d_{n+i} instead). On the other hand, if $\bar{x}$ satisfies the constraints above then taking an $\bar{x}$-graph (G_0, L), adding K to its vertex set and joining every vertex $v \in [n]$ to every vertex in $L(v)$ gives the preimage of (G_0, L). Thus we have a bijection between the set of d-regular graphs on $n + k$ vertices with no edges connecting vertices in K and the set of $\bar{x}$-graphs on n vertices satisfying (1a) and (1b).

Note that under this mapping the size of the common neighborhood of K is exactly x_K. From now on we shall work only with vectors satisfying (1a). We denote the set of vectors $\bar{x}$ satisfying (1b) by R.

In order to count such $\bar{x}$-graphs we apply the McKay-Wormald formula for asymptotic enumeration of graphs with a prescribed degree sequence. Denote the number of $\bar{x}$-graphs on vertex set $[n]$ (not necessarily satisfying (1b)) by $f(\bar{x})$. Then enumerating over the choices for labelling and applying[6] Lemma 2.1 gives

$$f(\bar{x}) \approx \binom{n}{x_\varnothing, \ldots, x_K} \sqrt{2} e^{1/4} \left(\lambda^\lambda (1-\lambda)^{(1-\lambda)}\right)^{\binom{n}{2}} \prod_{A \in 2^K} \binom{n-1}{d-|A|}^{x_A}.$$

where $\lambda = \delta n/(n-1)$.

Define now

$$g(\bar{x}) = \binom{n}{x_\varnothing, \ldots, x_K} \prod_{A \in 2^K} \binom{n-1}{d-|A|}^{x_A},$$

and note that $g(\bar{x}) \approx C f(\bar{x})$ where $C = C(\delta, n)$ is asymptotically a constant not depending on $\bar{x}$.

[6]The $O(1)$ factor $f(d_1, \ldots, d_n)$ appearing in the statement of Lemma 2.1 is known to be $\sqrt{2}e^{1/4}$ when the degree sequence obeys some demands stricter than those of Definition 1.2. These stricter demands are met by the regular degree sequence $\bar{d} = (d, d, \ldots, d)$.

Let D be the set of vectors $\bar{x}$ with $|x_K - \delta^k n| > \epsilon\delta^k n$. Then,

$$\Pr_{G_{n+k,d}}\left[\left||N^*(K)| - \delta^k n\right| > \epsilon\delta^k n \middle| G[K] = \varnothing\right] = \frac{\sum_{\bar{x}\in R\cap D} f(\bar{x})}{\sum_{\bar{x}\in R} f(\bar{x})} \approx \frac{\sum_{\bar{x}\in R\cap D} g(\bar{x})}{\sum_{\bar{x}\in R} g(\bar{x})} \leq \frac{\sum_{\bar{x}\in D} g(\bar{x})}{\sum_{\bar{x}\in R} g(\bar{x})}. \tag{2}$$

At this point we have bounded the sought-after probability by the last term in equation (2). Now we can leave aside the probabilistic and combinatorial background and deal with this quotient.

First we wish to give an upper bound for the numerator. Define $s = \sum_{A\in 2^K} \binom{n-1}{d-|A|}$ and note that

$$s = \sum_{i=0}^{k} \binom{k}{i}\binom{n-1}{d-i} = \binom{n-1+k}{d}.$$

Let $p_i = \binom{n-1}{d-i}/s$. Consider n independent trials where at every trial a set $A \in 2^K$ is randomly chosen, independently and with probability $p_{|A|}$. Let Y_A be a random variable counting the number of trials resulting with A and let $\bar{Y} = (Y_\varnothing, \ldots, Y_K)$. Then, for a vector $\bar{x}$, the probability for $\bar{Y} = \bar{x}$ is exactly $g(\bar{x})/s^n$. Notice that Y_K follows a Binomial distribution with parameters (n, p_k). Now,

$$E[Y_K] = n\binom{n-1}{d-k}/s = n\frac{d(d-1)\ldots(d-k+1)}{n(n+1)\ldots(n+k-1)} = \left(1 + O\left(\frac{1}{n}\right)\right)\delta^k n.$$

Applying a Chernoff type bound ([2, Corollary A.1.14]) we get

$$\frac{1}{s^n}\sum_{\bar{x}\in D} g(\bar{x}) = \Pr\left[\left|Y_K - \delta^k n\right| > \epsilon\delta^k n\right] < e^{-cn},$$

where $c > 0$ is some constant depending on ϵ, δ and k. Thus we have an exponential bound for the numerator[7] of the last term of equation (2), and we move on to estimate its denominator.

Notice that $\sum_{\bar{x}\in R} g(\bar{x})$ is exactly s^n times the probability of the experiment above resulting with $\bar{Y}$ s. t. for every $i \in K$, $\sum_{A\in\mathcal{A}_i} Y_A = d$. Now, d (or, in general, d_{n+i}) is a specific integer close to the expectation of this sum. Therefore, in order for Equations (1b) to hold it is suffice to require that all the random variables Y will get a value close to the expectation.

The last argument suggests that the probability of the event "Equations (1b) hold" is not too small. Claim 4.7 formulates the last sentence in a more rigorous fashion. In particular, it gives that

$$s^{-n}\sum_{\bar{x}\in R} g(\bar{x}) \geq c' n^{-k/2}$$

[7] Up to the normalizing factor of s^n which is common to the denominator and thus cancels.

for some constant $c' > 0$.

Summarizing the above we get

$$\Pr_{G_{n+k,d}}\left[\left||N^*(K)| - \delta^k n\right| > \epsilon\delta^k n\right] \leq \frac{\sum_{\bar{x}\in D} g(\bar{x})}{\sum_{\bar{x}\in R} g(\bar{x})} < \frac{e^{-cn}}{c'n^{-k/2}} < e^{-c''n},$$

where $c'' > 0$ is constant.

As mentioned above, these arguments may be repeated for every graph spanned by K with minor changes. Specifically, for the general case equations (1b) need to be changed into

$$\sum_{A\in\mathcal{A}_i} x_A = d(i).$$

Again, by the law of total probability applied for the $\binom{k}{2}$ options for the spanned graph $G[K]$ we get

$$\Pr_{G_{n+k,d}}\left[\left||N^*(K)| - \delta^k n\right| > \epsilon\delta^k n\right] < e^{-an}$$

for some constant a. Applying the union bound over the $\binom{n}{k}$ choices for K we get exponentially small probability for deviation in the size of the common neighborhood of any k vertices. This completes the proof of the Lemma, and in fact proves that the statement of the Lemma holds with probability exponentially close to one. □

It remains to show that if $\bar{Y}$ follows a multinomial distribution as described above, then the probability of getting $\sum_{A\in\mathcal{A}_i} Y_A = d$ for all i is bounded from below by a polynomial. This is the content of the next claim, which is the last one in this section.

Claim 4.7. *Let k be a positive integer and assume that the random vector $\bar{Y} = (Y_\emptyset, Y_{\{1\}}, \ldots, Y_{[k]})$ follows a multinomial distribution with parameters n and $\bar{p} = (p_\emptyset, p_{\{1\}}, \ldots, p_{[k]})$, where $p_A = \Theta(1)$ for every index $A \subset [k]$. Let $\mathcal{A}_i = \{A \in 2^{[k]} | i \in A\}$ and let W be the event*

$$\forall i \in [k], \quad \sum_{A\in\mathcal{A}_i} Y_A = E\left[\sum_{A\in\mathcal{A}_i} Y_A\right] + C_i, \tag{3}$$

where $C_i = O(1)$ for every $i \in [k]$. Then $\Pr[W] = \Omega\left(n^{-k/2}\right)$.

Notice that the right hand side of all k equalities in Equation (3) must be integral. No effort was made to give the most general or tight statement of this claim.

Proof. Each Y_A is distributed binomially with parameters n and p_A and has mean $\mu_A = np_A$ and standard deviation $\sigma_A = \sqrt{np_A(1-p_A)}$. By Chebyshev's inequality there exists a constant c_A s. t. $\Pr[|Y_A - E[Y_A]| > c_A\sigma_A] < 1/(2^{k+1})$. Let $c_1 = \max_{A\subset[k]} c_A\sqrt{p_A(1-p_A)}$. Let B be the event that, simultaneously for

all A with $|A| > 1$, $|Y_A - \mu_A| \leq c_1\sqrt{n}$. By the above and by the union bound, $\Pr[B] > 1/2$.

Expose now the values of Y_A with $|A| > 1$, and assume that B holds. In order for the $Y_{\{i\}}$'s to satisfy the equalities of Equation (3), every $Y_{\{i\}}$ needs to get a specific value,

$$T_i = E\left[\sum_{A\in\mathcal{A}_i} Y_A\right] + C_i - \sum_{A\in\mathcal{A}_i\setminus\{i\}} Y_A \ .$$

This will guarantee that the target event W will hold. According to our assumption we know that $|T_i| - E[Y_{\{i\}}]$ is bounded from above by $C_i + 2^{k-1}c_1\sqrt{n} = O(\sqrt{n})$. Having this in mind we also define E_j, $0 \leq j < k$, to be the event that the last j equations from the set of equations (3) are satisfied. Notice that once the values of Y_A with $|A| > 1$ are exposed, E_j is simply the event that $Y_{\{i\}} = T_i$ for $k - j < i \leq k$.

Define the set

$$\mathcal{S} = \left\{ A \in 2^{[k]} \middle| \, |A| > 1 \text{ or } A = \{i\} \text{ for } k - j < i \leq k \right\}.$$

Assume now that we have exposed the values of Y_A with $A \in \mathcal{S}$ and that both B and E_j hold. By our assumptions, the values of Y_A with $A \in \mathcal{S}$ are all very close to their means. The rest of the random variables, $Y_\emptyset, Y_{\{1\}}, \ldots, Y_{\{k-j\}}$, are again distributed according to a multinomial distribution with parameters $n' = n - \sum_{A\in\mathcal{S}} Y_A$ and $(\rho_0, \rho_1, \ldots, \rho_{k-j})$ where $\rho_i = p_{\{i\}} / \left(p_\emptyset + p_{\{1\}} + \cdots + p_{\{k-j\}}\right)$ is a positive constant.

Specifically, if we let Y' be $Y_{\{k-j\}}$ given B and E_j, then Y' is distributed binomially, with expectation

$$\begin{aligned}
&E\left[Y'\right] = n'\rho_{k-j} = \left(n - \sum_{A\in\mathcal{S}} Y_A\right)\rho_{k-j} = \\
&n\left(1 - \sum_{A\in\mathcal{S}} \left(p_A \pm c_1 n^{-1/2}\right)\right) \frac{p_{\{k-j\}}}{p_\emptyset + p_{\{1\}} + \cdots + p_{\{k-j\}}} = \\
&np_{\{k-j\}}\left(1 \pm c_2 n^{-1/2}\right) = \mu_{\{k-j\}}\left(1 \pm c_2 n^{-1/2}\right)
\end{aligned}$$

for some constant $c_2 > 0$.

Being a binomially distributed random variable, the standard deviation of Y' is given by $\sqrt{n'\rho_{k-j}(1-\rho_{k-j})} \leq c_3\sqrt{n}$ where c_3 is some positive constant. Now, by the definition of Y'

$$\Pr[Y_{\{k-j\}} = T_{k-j} | B \wedge E_j] = \Pr[Y' = T_{k-j}],$$

but, summing the deviations above we get

$$T_{k-j} \leq E[Y'] + C_{k-j} + 2^{k-1}c_1\sqrt{n} + c_2\sqrt{n}.$$

Neglecting C_{k-j} we may write

$$T_{k-j} = E[Y'] \pm \left((2^{k-1}c_1 + c_2)/c_3\right) \sigma[Y'],$$

and thus, by [7, Theorem 1.5], $\Pr[Y' = T_{k-j}] \geq \gamma n^{-1/2}$ for some positive constant $\gamma > 0$.

Applying the above argument for $j = 0, \ldots, k-1$ we get

$$\Pr[W] \geq \frac{1}{2}\gamma^k n^{-k/2},$$

which completes the proof. □

5 The sparse case – irrational α

This section contains the proof of Theorem 1.6 dealing with equivalence of sparse random regular graphs and sparse binomial random graphs. The logic aspect of the proof relies on the Ehrenfeucht game and on Theorem 3.2. We shall show that Duplicator can utilize the strategy described in [45, Chapter 6]. Duplicator's strategy works if certain properties are satisfied by the gameboard — the two graphs. Before we describe the properties let us mention that $G(n,p)$ a. a. s. has them — this is the combinatorial part of the proof of the Zero-One law in the binomial model $G(n,p)$.

5.1 Rooted graphs

As mentioned above, Duplicator's strategy relies on some extension properties, which may be conveniently described using the terminology of rooted graphs. The terms and definitions that we present here are based on Spencer's book [45]. Formally, a *rooted graph* (R, H) is a pair, consisting of a labelled graph H and a subset of the vertex set $V[H]$ denoted by R, the *root set*. The members of R are naturally named *roots* and the other vertices are sometimes referred to as *nonroots*. Informally, a rooted graph is simply a graph in which some of the vertices are designated as roots. The set of roots $\varnothing \subseteq R \subsetneq V[H]$ may be empty, but it must not contain all $V[H]$. Rooted graphs are interesting for themselves, but here we shall mainly use rooted graphs as *extensions*. Formal definition of extensions appears in the next paragraph, here we merely mention that we shall use the notions rooted graphs and extensions interchangeably.

Along this section we shall use r for the number of roots, v for the number of nonroots and e for the number of edges having at least one nonroot endpoint. The pair (v, e) is called the *type* of the rooted graph. For example, if we take K_6 — the complete graph on six vertices — and designate three vertices as roots, we get a rooted graph of type $(3, 12)$.

Let (R, H) be a rooted graph of type (v, e) with a labelling $R = (r_1, \ldots, r_r)$ and $V[H] \setminus R = (v_1, \ldots, v_v)$. Let G be a graph, let $\mathbf{x} = (x_1, \ldots, x_r)$ be an r-tuple of distinct vertices of G and let $\mathbf{y} = (y_1, \ldots, y_v)$ be a v-tuple of G's

vertices, distinct from each other and from the vertices in $\mathbf{x}$. We say that $\mathbf{y}$ is an (R,H)-*extension* of $\mathbf{x}$ if $y_i \sim y_j$ whenever $v_i \sim v_j$ and $x_i \sim y_j$ whenever $r_i \sim y_j$. Note that we ignore edges between roots. Simply put, $\mathbf{y}$ is an (R,H)-extension of $\mathbf{x}$ if thinking of $\mathbf{x}$ as the roots and of $\mathbf{y}$ as the nonroots we get a (labelled) copy of (R,H), not necessarily induced and with edges between roots ignored.

We say that G has $\mathrm{Ext}(R,H)$ if every r-tuple of its vertices has an (R,H)-extension. Note that $\mathbf{y}$ being an (R,H)-extension of $\mathbf{x}$ depends on the labelling given to the vertices of H, however $\mathrm{Ext}(R,H)$ is invariant of the labelling. Clearly, $\mathrm{Ext}(R,H)$ is describable in the first order language of graphs. For example, if (R,H) is the rooted graph from the previous example, then $\mathrm{Ext}(R,H)$ is given by

$$\forall x_1,x_2,x_3,\ \left(\bigwedge_{1\le i<j\le 3} x_i \neq x_j\right) \rightarrow \left[\exists y_1,y_2,y_3 \ \text{ s.t. } \left[\left(\bigwedge_{1\le i,j\le 3} x_i \neq y_j\right) \wedge \left(\bigwedge_{1\le i<j\le 3} y_i \neq y_j\right) \wedge \left(\bigwedge_{1\le i,j\le 3} x_i \sim y_j\right) \wedge \left(\bigwedge_{1\le i<j\le 3} y_i \sim y_j\right)\right]\right].$$

The destiny of $\mathrm{Ext}(R,H)$ depends on the density of (R,H). The *density* of a rooted graph on r vertices and of type (v,e) is the ratio e/v and is denoted by $d(R,H)$ (notice that the number of roots does not play a role here). The expected number of (R,H)-extensions of a fixed r-tuple in $G(n,p)$ is $\binom{n-r}{v}p^e$. If $p = n^{-\alpha}$ then the expected number of (R,H)-extensions becomes $n^{v-\alpha e} = n^{v(1-\alpha d)}$, thus it is $o(1)$ if $d(R,H) > 1/\alpha$.

Let S be a set of vertices such that $R \subsetneq S \subseteq V[H]$. We call $(R,H[S])$ a *subextension* of (R,H). If $R \subseteq S \subsetneq V[H]$ we call (S,H) a *nailextension* of (R,H). Notice that (R,H) is both a subextension and a nailextension of itself (the *trivial* subextension/nailextension).

We say that (R,H) (with type (v,e)) is a *balanced* rooted graph if for every subextension with type (v',e') one has $e/v \ge e'/v'$. If in addition $e/v > e'/v'$ for every nontrivial subextension then the rooted graph is said to be *strictly balanced*.

Let $0 < \alpha < 1$ be some irrational number, which we fix for the rest of this section. We say that (R,H) is *sparse* if $e/v < 1/\alpha$ and *dense* if $e/v > 1/\alpha$. Clearly, the irrationality of α creates a dichotomy so every rooted graph is either sparse or dense. A rooted graph is called *safe* if all its subextensions are sparse and *minimally safe* if, in addition, for every vertex set S satisfying $R \subsetneq S \subsetneq V[H]$, the nailextension (S,H) is not safe. If all nailextensions of a rooted graph are dense, it is said to be a *rigid* rooted graph. Since the notions of sparse, dense, safe, rigid and minimally safe depend on α we shall only use them when it is clear what α is.

Properties 5.2 and 5.3 below show a close connection between two of the rooted graph notions defined above, namely strictly balanced and minimally safe. Some of the following properties and many more, together with a further development of the theory of rooted graphs may be found in [45, Chapter 4].

Besides of Property 5.3, we only state here the properties that we shall need later.

Property 5.1. *If (R, H) is sparse then it has a safe nailextension.*

Proof. Assume that (R, H) is not safe, then it has a dense subextension. Let $(R, H[S])$ be a maximal dense subextension. In this case (S, H) is safe, since otherwise it has a dense subextension, $(S, H[T])$, but then it is easy to see that $(R, H[T])$ is dense since it is a dense extension of a dense extension. This implies a contradiction to the maximality of $(R, H[S])$. □

Property 5.2. *If (R, H) is a minimally safe rooted graph (relative to some α) then it is also strictly balanced.*

Proof. Let (R, H) be a minimally safe extension with type (v, e) and assume that $R \subset S \subset V(H)$. Denote the types of $(R, H[S])$ and (S, H) by (v_1, e_1) and (v_2, e_2) respectively. First we claim that (S, H) is dense. Otherwise it is sparse but not safe and by Property 5.1 there is a set of vertices T s.t. (T, H) is safe, contradicting the minimality of (R, H).

Now, (S, H) is dense and by definition $v_2 - \alpha e_2 < 0$. Since $v = v_1 + v_2$ and $e = e_1 + e_2$ we have $v - \alpha e = (v_1 - \alpha e_1) + (v_2 - \alpha e_2)$. But then,

$$
\begin{aligned}
& v - \alpha e < v_1 - \alpha e_1 \\
& e(\frac{v}{e} - \alpha) < e_1(\frac{v_1}{e_1} - \alpha) \\
& \frac{e}{e_1}(\frac{v}{e} - \alpha) < (\frac{v_1}{e_1} - \alpha) \qquad \text{but } 0 < e_1 < e \text{ so} \\
& \frac{v}{e} - \alpha < \frac{v_1}{e_1} - \alpha \\
& e/v > e_1/v_1,
\end{aligned}
$$

which is what we wanted to prove. □

Property 5.2 states that if (R, H) is a minimally safe rooted graph with respect to some α, then it is also strictly balanced. The next property can be seen as a converse, saying that if (R, H) is strictly balanced then it is also minimally safe with respect to some α^*. Property 5.3 is not used later and it is stated solely for completeness. Along the statement and proof of the next property the notions of sparse, safe and minimally safe are with respect to α^* and not to α that was set at the beginning of the section.

Property 5.3. *If (R, H) is a strictly balanced rooted graph of type (v, e), then there are rational numbers $0 < \alpha_1 < \alpha_2 < 1$ s.t. if $\alpha_1 < \alpha^* < \alpha_2$ then (R, H) is minimally safe with respect to this α^*.*

Proof. Take $\alpha_2 = v/e$ and note that if $\alpha^* < \alpha_2$ then (R, H) is sparse. Moreover, no subextension of (R, H) is denser than (R, H), due to (R, H) being balanced, thus (R, H) is safe.

Set

$$\alpha_1 = \max_{R \subset S \subset V(H)} \inf\{\alpha' | (S,H) \text{ is not safe with respect to } \alpha'\}.$$

Clearly, if $\alpha^* > \alpha_1$ then for every $R \subset S \subset V[H]$ the nailextension (S,H) is not safe. For every (S,H) one has

$$\inf\{\alpha | (S,H) \text{ is not safe with respect to } \alpha\} \leq 1/d(S,H).$$

On the other hand $d(S,H) > d(R,H)$ because $d(R, H|_S) < d(R,H)$ (since (R,H) is strictly balanced). Therefore we have $\alpha_1 < \alpha_2$ and the property proof is complete. □

When two extensions intersect, the intersection may be thought of as a subextension of these rooted graphs. If two copies of a strictly balanced rooted graph intersect, the intersection is sparser then the original rooted graph. This fact is useful when applying the second moment method, as seen in Proposition 5.5, but first we need to define the following:

Definition 5.4. For a rooted graph (R,H) we define:

$$\gamma(R,H) = \min_{(R,S)} (v' - \alpha e'),$$

where the minimum is taken over all subextensions (R,S) of (R,H), and the type of (R,S) is (v',e')..

Proposition 5.5. *Let (R,H) be a sparse strictly balanced rooted graph with type (v,e). Let $\mathbf{x}$ be an r-tuple of distinct vertices and let X be the number of (R,H)-extensions of $\mathbf{x}$ in $G(n,p=n^{-\alpha})$. Then $\gamma > 0$ and*

$$\mathrm{Var}[X] = \Theta\left(n^{2(v-\alpha e)-\gamma}\right).$$

Proof. For a v-tuple of distinct vertices $\mathbf{y}$ that are also distinct from the vertices in $\mathbf{x}$, let $I_{\mathbf{y}}$ be the indicator random variable for the event "$\mathbf{y}$ is an (R,H)-extension of $\mathbf{x}$". For a pair of v-tuples of vertices $\mathbf{y}', \mathbf{y}''$ we use $\mathbf{y}' D \mathbf{y}''$ to denote the fact that the (R,H)-extensions of $\mathbf{x}$ by $\mathbf{y}'$ and $\mathbf{y}''$ have a common edge. It is easy to observe that $I_{\mathbf{y}'}$ and $I_{\mathbf{y}''}$ are independent if and only if $\neg(\mathbf{y}' D \mathbf{y}'')$, because $I_{\mathbf{y}'}$ and $I_{\mathbf{y}''}$ are indicator random variables for sets of edges. Now, since $X = \sum I_{\mathbf{y}}$, we have

$$\mathrm{Var}[X] = \sum_{\mathbf{y}',\mathbf{y}''} \mathrm{Cov}[I_{\mathbf{y}'}, I_{\mathbf{y}''}] = \sum_{\mathbf{y}' D \mathbf{y}''} \left[E\left[I_{\mathbf{y}'} I_{\mathbf{y}''}\right] - E[I_{\mathbf{y}'}] E[I_{\mathbf{y}''}]\right] =$$

$$= \Theta\left(\sum_{R \subsetneq V[H_1] \cap V[H_2] \subseteq V[H_1]} n^{2v-v'} \left(p^{2e-e'} - p^{2e}\right) \right),$$

where v' denotes the number of vertices common to $\mathbf{y}'$ and $\mathbf{y}''$, and e' denotes the number of common edges in the (R,H)-extension of $\mathbf{x}$ by $\mathbf{y}'$ and $\mathbf{y}''$. Rewriting the last expression, we may sum over S, the set of common vertices (including roots), to get

$$\mathrm{Var}[X] = \Theta\left(\sum_{R\subsetneq S\subseteq V[H_1]} n^{2v-v'-\alpha(2e-e')}\right) = \Theta\left(n^{2(v-\alpha e)-\min_S(v'-\alpha e')}\right).$$

Since $e'/v' < e/v < 1/\alpha$ implies $v'-\alpha e' > 0$ for every $R \subsetneq S \subseteq V[H_1]$ we get that $\gamma > 0$ and the proof is complete. □

Proposition 5.5 above shows that $\mathrm{Var}[X]/(EX)^2 = o(1)$ which means that X is concentrated around its mean. Since we want to show that this typically happens simultaneously for every r-tuple of vertices $\mathbf{x}$, we need to work a bit harder.

For a set of graphs $\mathscr{G} = \{G_1, \ldots, G_l\}$ we define the *intersection graph* of $\mathscr{G}$, denoted by $\mathrm{IG}(\mathscr{G})$, as follows. The vertex set of $\mathrm{IG}[\mathscr{G}]$ is $\mathscr{G}$ and there is an edge between G_i and G_j if $|V[G_i] \cap V[G_j]| \geq 2$ (in particular, G_i is adjacent to G_j if they have an edge in common).

Consider the following situation. Start with a rooted graph (R,H). At the i'th step, pick a nontrivial nailextension of (R,H) (denoted by (R_i,H_i)) and extend the current graph with a copy of this nailextension. That is, pick $|R_i|$ vertices from the current graph and then add $|V[H_i]| - |R_i|$ vertices and complete the edges according to H_i. We can think of the graph resulting from l such steps as an intersection of l copies of (R,H) or as a sequence of copies of (R,H) extending each other. We are interested in the sparsest configuration of this structure.

Definition 5.6. Consider a sequence of l copies of a rooted graph (R,H) as above, and denote them by $(R,H_1),\ldots,(R,H_l)$. Let J be the intersection graph of the graphs induced on the nonroot vertices of every rooted graph (R,H_i). Now define

$$\beta_l(R,H) = \min\left\{d(R,\bigcup_{i=1}^{l} H_i)\right\}$$

where the minimum is over all arrangements of $(R,H_1),\ldots,(R,H_l)$ for which J is connected.

Property 5.7. *If (R,H) is a strictly balanced rooted graph then for every l*

$$\beta_l(R,H) < \beta_{l+1}(R,H).$$

Proof. Assume that the rooted graph (R,H) has type (v,e) and is extended by a sequence of l copies of nontrivial nailextensions of itself denoted again by $(R_1,H_1),\ldots,(R_l,H_l)$ with types $(v_1,e_1),\ldots,(v_l,e_l)$ respectively. This means that the root set of every (R_i,H_i) is a subset of $V[H_{i-1}]$ and the nonroot vertices

of (R_i, H_i) are outside of H_{i-1}. Then the density of the resulting rooted graph is given by the ratio

$$\frac{e + e_1 + \cdots + e_l}{v + v_1 + \cdots + v_l} = \frac{vd + v_1 d_1 + \cdots + v_l d_l}{v + v_1 + \cdots + v_l}.$$

Thus the density of the resulting graph is simply a weighted average of the densities of the components, where the weight of the i'th component is given by $v_i/(v + v_1 + \cdots + v_l)$.

Assume that the sparsest nontrivial nailextension among all possible nontrivial nailextensions of (R, H) has type (v', e') (here we use the fact that J is connected — we know that the nailextensions are proper). Then

$$\beta_{l+1} \leq \frac{e + le'}{v + lv'} < \frac{e'}{v'}.$$

The first inequality comes from the definition of β_{l+1}. As for the second:

$$ev' + le'v' < e'v + le'v' \Leftrightarrow \frac{e}{v} < \frac{e'}{v'},$$

where in the second inequality here we use the assumption that (R, H) is strictly balanced, and thus every proper nailextension of (R, H) is denser than (R, H) itself. In other words, β_{l+1} is smaller than the density of the sparsest nailextension. Assume that taking (R, H) and extending it by $l+1$ nailextensions gives a rooted graph with density exactly β_{l+1}. Then any of these nailextensions must have density greater than β_{l+1}, and thus removing it will give a rooted graph with lower density, implying $\beta_l < \beta_{l+1}$. □

When we shall use Property 5.7 later, the root set of all of the nailextensions will be the same. That is, we shall have $R_1 = R_2 = \cdots = R_l = R$. Still, the sparsest structure will have the same density as it only depends on the densities of the nailextensions used and not on the way they connect to each other (but it still may be that there will be less structures attaining this minimum density when requiring that all the root sets shall be the same).

The next property is part of the proof of Theorem 5.2.1 of [45] and we shall use it in a similar manner.

Property 5.8. *Every safe extension may be decomposed into a sequence of minimally safe extensions. That is, if (R, H) is a safe rooted graph then there is a chain of sets $R = R_0 \subset R_1 \subset \cdots \subset R_t = V[H]$ s.t. for every $1 \leq i \leq t$ the rooted graph $(R_{i-1}, H|_{R_i})$ is minimally safe.*

Proof. If for every vertex set S, $R \subset S \subset V[H]$, the rooted graph (S, H) is not safe, then (R, H) is minimally safe. Otherwise, let S be a vertex set such that (S, H) is safe. $(R, H[S])$ is a subextension of a safe rooted graph, so it is also safe. Thus we have got a finer chain. □

We end this subsection by citing a few more definitions from [45]. These terms will be used to describe the gameboard properties required for Duplicator's strategy, and are given using the notions defined above.

Fix $0 < \alpha < 1$. The terms "safe" and "rigid" in the following definitions are relative to this α.

Definition 5.9. Let t be a fixed positive integer. Let G be a graph and let U be a subset of its vertices. The *t-closure* of U, denoted $\mathrm{cl}_t(U)$, is the minimal set of vertices X satisfying the following requirements:

- $U \subseteq X$,
- Let (R, H) be a rigid rooted graph with r roots and $v \leq t$ nonroot vertices. Let $\mathbf{x} = (x_1, \ldots, x_r)$ be an r-tuple of distinct vertices of X. If $\mathbf{y} = (y_1, \ldots, y_v)$ is an (R, H)-extension of $\mathbf{x}$ then all $y_1, \ldots, y_v \in X$.

Why is Duplicator interested in the t-closure? Because she needs to make sure that the t-closure[8] of the marked vertices is the same in both graphs of the gameboard. This is best demonstrated by example. Let $\alpha = 2/\pi$ and consider $G(n, p)$ with $p = n^{-\alpha}$. Most pairs of vertices do not have a common neighbor, but some do. Now consider $\mathrm{EHR}[G(n, p), G(n, p); 3]$. Spoiler may choose two nonadjacent vertices in G_1 that have a common neighbor. If Duplicator will just pick two nonadjacent vertices in G_2, she will most likely lose the game in the next round when Spoiler will pick the common neighbor and she will not be able to respond. Thus, when Spoiler picks a vertex, Duplicator needs to look for any "special" situation that can occur in the future (in our example she will take special care for vertices having a common neighbor with the vertex picked by Spoiler). The vertices that can be used to create such a special structure are precisely the vertices forming the t-closure.

Definition 5.10. Let G be a graph, let (R, H) be a rooted graph and let t be a fixed positive integer. We say that a v-tuple of vertices of G, denoted $\mathbf{y} = (y_1, \ldots, y_v)$, is a *$t$-generic (R, H)-extension* of the r-tuple $\mathbf{x} = (x_1, \ldots, x_r)$ of vertices of G if

- $\mathbf{y}$ is an (R, H)-extension of $\mathbf{x}$,
- There are no additional edges of G between vertices of $\mathbf{y}$ or between vertices of $\mathbf{x}$ and vertices of $\mathbf{y}$, besides the edges required for (R, H)-extension,
- If any $\mathbf{z} = (z_1, \ldots, z_s)$ with $s \leq t$ forms a rigid extension of H then there are no edges of G between the z_i and the y_j. That is, if there are $v_1, \ldots, v_q \in H$ such that the rooted graph $(\{v_1, \ldots, v_q\}, G[\mathbf{z}])$ is rigid, then there is no edge of the form $y_i z_j$.

[8] The value of t depends on the situation in the game (the gameboard, the vertices that were marked, the number of rounds that have passed and so on). Still, given α and the number of rounds, t is bounded by some constant.

This definition will be used in the other part of Duplicator's strategy. If Spoiler picks an ordinary vertex — a vertex that is a safe extension of the marked vertices, then Duplicator will most probably be able to find a vertex with the same adjacency pattern and the same t-closure in the other graph. Generic extensions are just extensions with nothing special about them. Notice that while we have defined t-generic (R, H)-extensions for any rooted graph (R, H), the last remark means that we shall use it only when (R, H) is a safe extension.

5.2 Duplicator's strategy

Let $0 < \alpha < 1$ be an irrational number and let k be an integer. Define $p = n^{-\alpha}$ and $d = n^{1-\alpha}$ and let $\bar{d}$ be a nearly regular degree sequence around d. Recall that we wish to show that the probability that Duplicator wins $\mathrm{EHR}[G_1, G_2; k]$, where G_1 is drawn from $G(n, p)$ and G_2 from $G_{n,\bar{d}}$, tends to one as n tends to infinity. Having the notions of rooted graphs we can describe the graph properties Duplicator needs for her strategy. While the strategy itself is rather complicated, it relies on the gameboard having the following properties for every $0 \leq r \leq k$ and $0 \leq t \leq f(\alpha, k)$, where $f(\alpha, k)$ is some constant depending only on α and k.

Finite Closure(r, t) There exists an integer K s.t. for every m-tuple of vertices $\mathbf{x}$ of length at most r one has $|\mathrm{cl}_t(\mathbf{x})| < r + K$.

Generic Extension(r, t) If (R, H) is safe, then every r-tuple of vertices $\mathbf{x}$ has a t-generic (R, H)-extension.

We will show that for every fixed integers r, t both properties above are a. a. s. satisfied by $G_{n,\bar{d}}$. This, together with the known fact that this is also the case in $G(n, p)$, will mean that for every k Duplicator a. a. s. will be able to follow her strategy and to win $\mathrm{EHR}[G(n, p), G_{n,\bar{d}}; k]$. Applying Theorem 3.2 this completes the proof of Theorem 1.6.

The next lemma states that a. a. s. $G_{n,\bar{d}}$ has Finite Closure. We only sketch the lemma's proof.

Lemma 5.11. *For every two integers r, t a. a. s. $G_{n,\bar{d}}$ has Finite Closure with parameters r and t.*

Showing that $G_{n,\bar{d}}$ has Finite Closure is relatively straightforward. Since it follows the same reasoning as in the binomial model up to the replacement of direct probabilistic computations with applications of Corollary 2.4, we will not describe it in full detail (see Remark 2.5). Instead we give a short informal description of the proof.

The closure of a set of vertices is made of extensions by rigid rooted graphs. Each such extension makes the closure a bit denser until, at some point — say after K/t extension, the expected number of copies of the closure in the graph tends to zero. At this point we split from the binomial model, where a

direct computation of the probability is made, and instead use Corollary 2.4 to estimate this probability and show that it tends to zero as a function of n.

To complete the proof one needs to notice that there are only finitely many rigid extensions of r vertices using at most t vertices, and that there are only finitely many ways to combine K/t rigid extensions. Since every combination of that sort appears with probability tending to zero, we can deduce that a. a. s. the size of the closure is less than $r + K$.

Showing that $G_{n,\bar{d}}$ has Generic Extension is more involved and it is done in the next subsection.

5.3 Generic Extension in random regular graphs

In this subsection we aim to prove a regular parallel to the Generic Extension Theorem of [45, Theorem 5.3.1]. Let α be an irrational number between zero and one, let $p = n^{-\alpha}, d = np$ and let $\bar{d}$ be a nearly regular degree sequence around d.

Theorem 5.12. *Let (R, H) be a safe extension with r roots and let $t > 0$ be an integer. If $G \sim G_{n,\bar{d}}$ then asymptotically almost surely every r-tuple of vertices $\mathbf{x}$ has a t-generic (R, H)-extension.*

Proof. Let (R, H) be a safe extension with r roots and of type (v, e) and let also $\mathbf{x}$ be an r-tuple of distinct vertices. By Property 5.8 we can find $s + 1$ subsets of $V[H]$, denoted $R = R_0 \subset R_1 \subset \cdots \subset R_s = V[H]$ s. t. for every $1 \leq i \leq s$ the rooted graph $(R_{i-1}, H[R_i])$ is minimally safe. By property 5.2 every rooted graph $(R_{i-1}, H[R_i])$ is also strictly balanced. Thus we have a finite sequence of strictly balanced extensions and we can add in the main ingredient of the proof — Lemma 5.13 below — which allows us to estimate the number of copies of each strictly balanced extension. On a more concrete level, if the type of $(R_{i-1}, H[R_i])$ is (v_i, e_i) then by Lemma 5.13, a. a. s. the number of $(R_0, H[R_1])$-extensions of $\mathbf{x}$ is $\Theta\left(n^{v_1 - \alpha e_1}\right)$. Now, looking at a specific extension of $\mathbf{x}$, we have $|R_1|$ vertices and we can find $\Theta\left(n^{v_2 - \alpha e_2}\right)$ copies of $(R_1, H[R_2])$ to extend it (again, by Lemma 5.13). Continuing in this manner for s iterations we end up with

$$\Theta\left(n^{\sum (v_i - \alpha e_i)}\right) = \Theta\left(n^{v - \alpha e}\right)$$

(R, H)-extensions of $\mathbf{x}$. It may be that some of these extensions were counted more than once in this process, but the number of ways to build an extension is bounded by some constant depending on (R, H) solely.

Besides the necessary change described in Remark 2.5 the rest is exactly as in the proof of Theorem 5.3.1. of [45] and we shall not repeat it here in full detail. Instead, we limit ourselves to the following outline description of the rest of the proof. It remains to show that at least one of the $\Theta\left(n^{v - \alpha e}\right)$ extensions is generic, so we need to show that some (finitely many) denser structures do not occur too often. This is done using the above argument (the fact that we can estimate the number of safe extensions) and by several first moment considerations which, in our case, are based on Corollary 2.4. □

Lemma 5.13. *Let (R,H) be a strictly balanced sparse extension of type (v,e) and with $r=|R|$ roots, and also let $\epsilon>0$. For an r-tuple of vertices $\mathbf{x}$ let $N_{\mathbf{x}}$ denote the number of v-tuples of vertices $\mathbf{y}$ such that $\mathbf{y}$ is an (R,H)-extension of $\mathbf{x}$. Then a. a. s. for every r-tuple $\mathbf{x}=(x_1,\ldots,x_r)$ of distinct vertices in $G_{n,\bar{d}}$ one has*

$$N_{\mathbf{x}}=(1\pm\epsilon)EN_{\mathbf{x}}=(1\pm\epsilon)\binom{n-r}{v}p^e.$$

Proof. Let $\epsilon>0$ be a positive small number and let $\mathbf{x}$ be an r-tuple of distinct vertices. For every v-tuple of distinct vertices $\mathbf{y}$, distinct from $\mathbf{x}$ as well, let $X_{\mathbf{y}}$ be the indicator random variable of the event "$\mathbf{y}$ is an (R,H)-extension of $\mathbf{x}$". Clearly $N_{\mathbf{x}}=\sum_{\mathbf{y}}X_{\mathbf{y}}$ and $\mu=EN_{\mathbf{x}}=\binom{n-r}{v}p^e=\Theta(n^{v-\alpha e})$. Let m be the smallest even integer greater or equal to $2r/\gamma+1$ (γ is defined in Definition 5.4). We are interested in μ_m, the m'th central moment of $N_{\mathbf{x}}$, defined by

$$\mu_m=E(N_{\mathbf{x}}-\mu_1)^m.$$

By Lemma 5.14 and Property 5.5, $\mu_m=O((\mathrm{Var}[N_{\mathbf{x}}])^{m/2})=O(n^{m(v-\alpha e)-\gamma m/2})$. Applying Fact 5.15 for the m'th moment we get

$$\Pr[|N_{\mathbf{X}}-\mu|>\epsilon\mu]<\frac{\mu_m}{(\epsilon\mu)^m}=o(n^{-r}).$$

By the union bound over all choices of r-tuple of distinct vertices, the proof is complete. □

Lemma 5.14. *Let $d=n^\alpha$ for some $0<\alpha<1$ and let $\bar{d}$ be a nearly regular degree sequence around d. Let (R,H) be a strictly balanced sparse extension having r roots and being of type (v,e), and let X_d be the number of (R,H)-extensions of some fixed r-tuple of distinct vertices in $G_{n,\bar{d}}$. For a constant integer $m\geq 1$ let $\mu_m=\mu_m[X_d]=E[(X_d-EX_d)^m]$ be the m'th central moment of X_d. Then*

$$\mu_m=O(\mathrm{Var}[X_d]^{m/2})\ .$$

Proof. Let $\mathbf{x}=(x_1,\ldots,x_r)$ be a fixed r-tuple of distinct vertices. We want to estimate μ_m. By definition

$$\mu_m[X_d]=E\left[(X_d-EX_d)^m\right]=\sum_{i=0}^{m}\binom{m}{i}(-1)^i(EX_d)^{m-i}E\left[X_d^i\right],$$

so by estimating the moments EX_d^m we also get an estimation for the central moments. As in the proof of Lemma 5.13 above, we define for every v-tuple $\mathbf{y}$ of distinct vertices from $V[G_{n,\bar{d}}]\setminus\{x_1,\ldots,x_r\}$, a random variable $X_{d,\mathbf{y}}$, which is the indicator random variable of the event "$\mathbf{y}$ is an (R,H)-extension of $\mathbf{x}$ in $G_{n,\bar{d}}$". Notice that every $X_{d,\mathbf{y}}$ depends only on a finite number of edges. Clearly $X_d=\sum X_{d,\mathbf{y}}$, and by definition,

$$EX_d^m=E\left(\sum X_{d,\mathbf{y}}\right)^m=\sum\left[\prod_{i=1}^{m}X_{d,\mathbf{y}_i}\right],$$

where the sum is over all m-tuples of v-tuples of distinct vertices $\mathbf{y}$. The $X_{d,\mathbf{y}}$ are indicator random variables, and thus the product equals one exactly when all the events occur simultaneously. The event "For every $1 \leq i \leq m$, $\mathbf{y}_i$ is an (R,H)-extension of $\mathbf{x}$ in $G_{n,\bar{d}}$" depends on at most $m \cdot e$ edges, which is constant. Thus Corollary 2.4 is applicable and therefore

$$EX_d^m = \sum \Pr\left[\bigwedge_{i=1}^{m} \text{“}\mathbf{y}_i \text{ is an } (R,H)\text{-extension of } \mathbf{x} \text{ in } G_{n,\bar{d}}\text{”}\right] =$$
$$= \sum \Pr\left[\bigwedge_{i=1}^{m} \text{“}\mathbf{y}_i \text{ is an } (R,H)\text{-extension of } \mathbf{x} \text{ in } G(n,p=d/n)\text{”}\right].$$

Hence we conclude that $\mu_m[X_d] = (1+o(1))\mu_m[X_p]$, where X_p is the number of (R,H)-extensions of $\mathbf{x}$ in $G(n,p)$ with $p = d/n$.

Next we claim that $\mu_m[X_p] = (m-1)!!(\mu_2[X_p])^{m/2}(1+o(1))$ for even m (and $o\left((\mathrm{Var}X_p)^{m/2}\right)$ for odd m). In order to prove the last claim we borrow ideas from the proof of Theorem 6.5 of [20] and, when we can, take advantage of the fact that (R,H) is strictly balanced.

Similarly to the above we define $X_{p,\mathbf{y}}$ to be an indicator random variable for the event "$\mathbf{y}$ is an (R,H)-extension of $\mathbf{x}$ in $G(n,p)$". Now,

$$E[(X_p - EX_p)^m] = \sum_{X_{p,\mathbf{y}_1},\ldots,X_{p,\mathbf{y}_m}} E\left[(X_{p,\mathbf{y}_1} - EX_{p,\mathbf{y}_1})\cdots(X_{p,\mathbf{y}_m} - EX_{p,\mathbf{y}_m})\right], \tag{4}$$

summing over all m-tuples of v-tuples of vertices as before.

Define

$$T(X_{p,\mathbf{y}_1},\ldots,X_{p,\mathbf{y}_m}) = E\left[(X_{p,\mathbf{y}_1} - EX_{p,\mathbf{y}_1})\cdots(X_{p,\mathbf{y}_m} - EX_{p,\mathbf{y}_m})\right],$$

that is, the T's are the terms in the sum in Equation (4). For each such term consider a graph, $L = L(X_{p,\mathbf{y}_1},\ldots,X_{p,\mathbf{y}_m})$, whose vertices are the v-tuples $\mathbf{y}_i$. There is an edge in L between $\mathbf{y}_i$ and $\mathbf{y}_j$ if the labelled copies of (R,H) onto the vertices of $\mathbf{x}$ and $\mathbf{y}_i$ and onto the vertices of $\mathbf{x}$ and $\mathbf{y}_j$ have a common edge. In fact, L is the dependency graph of the random variables $X_{p,\mathbf{y}_1},\ldots,X_{p,\mathbf{y}_m}$, that is, there is an edge between $\mathbf{y}_i$ and $\mathbf{y}_j$ if and only if $X_{p,\mathbf{y}_i}$ and $X_{p,\mathbf{y}_j}$ are dependent random variables.

Group the terms in Equation (4) according to the structure of L. Consider first the case where m is even and L is a matching. In this case there are $m/2$ disjoint edges in L. There are exactly $(m-1)!! = m!/(2^{m/2}(m/2)!)$ labelled matchings, each contributing $(\mathrm{Var}X_p)^{m/2}(1+o(1))$ to the sum aforementioned. Every other graph L contributes $o((\mathrm{Var}X_p)^{m/2})$, as shown below, regardless of the parity of $|L|$.

If L has an isolated vertex, say vertex $\mathbf{y}_i$, then the term $E(X_{p,\mathbf{y}_i} - EX_{p,\mathbf{y}_i})$ in T is independent in the rest of the terms in T, which makes T vanish. If there are no isolated vertices in L and L is not a perfect matching, then the number of edges is greater than $m/2$. In order to complete the proof we will

show, bearing Property 5.7 in mind, that there exists a strictly monotonically increasing sequence (β_i) such that the contribution to the sum in equation (4) of a connected component of L having l elements is $O\left(n^{l(v-\alpha e-\beta_l)}\right)$. Thus, if L has connected components of sizes $l_1, \ldots, l_c$ then the contribution of terms in the group of this L is $O\left(n^{m(v-\alpha e)-\sum l_i\beta_{l_i})}\right)$. Since the sequence (β_i) is monotonically increasing, this expression is maximal when all the connected components are of size two. There are only finitely many graphs L to consider and hence the total contribution of all the graphs that are not matchings is negligible.

We will finish the proof by showing the existence of a sequence (β_i) as described above. For every vertex set $R \subsetneq S \subsetneq V[H]$ let (v_S, e_S) be the type of the proper nailextension (S, H). Define $\dot{S}$ to be the vertex set that makes the difference $(v_S - \alpha e_S)$ maximal. Denote the type of $(\dot{S}, H)$ by $(\dot{v}, \dot{e})$ and the type of $(R, H[\dot{S}])$ by (v', e').

Consider a connected component of L having size l. We can think of it as (R, H) being extended by nailextensions of itself $(S_2, H), (S_3, H), \ldots, (S_l, H)$. The expected number of extensions corresponding to this structure is

$$\Theta\left(n^{v+\sum v_{S_i}-\alpha(e+\sum e_{S_i})}\right) = \Theta\left(n^{v-e+\sum(v_{S_i}-\alpha e_{S_i})}\right).$$

The exponent is maximal when every summand is maximal, that is, when all the nailextensions are of type $(\dot{v}, \dot{e})$. In this case the exponent becomes

$$\begin{aligned} v - \alpha e + (l-1)(\dot{v} - \alpha\dot{e}) = \\ v - \alpha e + (l-1)(v - v' - \alpha(e - e')) = \\ l\left(v - \alpha e - \frac{l-1}{l}(v' - \alpha e')\right). \end{aligned}$$

Define $\beta_l = (1 - \frac{1}{l})(v - \alpha e')$, and recall that $v' - \alpha e' > 0$ because (v', e') is a type of a subextension of a strictly balanced sparse rooted graph. Thus $\beta_1 = 0$ and (β_i) is a strictly monotonically increasing sequence. □

The following fact is a straightforward generalization of the Chebyshev inequality for higher moments.

Fact 5.15. *Let X be a random variable with expectation μ and m'th central moment μ_m. If m is even then*

$$\Pr[|x - \mu| \geq \alpha] \leq \mu_m/\alpha^m.$$

Proof. Set $Y = (X - \mu)^m$ and apply Markov's inequality for Y. That is,

$$\Pr[|x - \mu| \geq \alpha] = \Pr[Y \geq \alpha^m] \leq \frac{EY}{\alpha^m} = \frac{\mu_m}{\alpha^m}.$$

□

6 The sparse case – rational α

In this section we describe the proof of Theorem 1.8. The general idea is to interpret a segment of arithmetic using properties of the underlying structure — the regular random graph in our case. Our construction is similar (but not identical) to the constructions of [40] and [45]. Taking advantage of this similarity, we shall not describe the construction in full detail but rather mention the differences between our construction and the construction of [45]. Showing that the constructed sentence has no limiting probability relies on properties of the random regular graph model. Taking clarity of presentation into consideration, we prove the special case where $d = n^{2/3}$. This also corresponds to the binomial random graph model $G(n,p)$ with $p = n^{-1/3}$, which is the study case of [45] used here as reference. The proofs of the existence of the relevant graph properties are given in full detail. The generalization of the proof to any rational $0 < \alpha < 1$ requires some effort, that is, it requires to define a different sentence A depending on α. Still, once we have such a sentence for $G(n, p = n^{-\alpha})$, we can easily define a similar sentence A' that will have no limiting probability in $G_{n,d=n^{1-\alpha}}$. Moreover, the combinatorial and probabilistic techniques that are given here are all that one needs in order to demonstrate the nonconvergence of the probability of such a sentence.

This section is organized as follows. First we describe the method of encoding of small graphs that we use in the construction. Next we state and prove Proposition 6.1 saying that given a set of vertices U, we can encode any graph H on $|U|$ vertices using U as its vertex set. This proof requires several technical lemmas appearing next. Then we show that a. a. s. there exist three vertices with common neighborhood of size $\ln \ln n$. Finally we describe the construction up to the point where it coincide with the original construction of Shelah and Spencer.

Let $G = (V, E)$ be a graph and let U be a set of vertices. Given a vertex $w \in V \setminus U$ we define the graph $H[G, U, w]$ in the following way. The vertex set of $H[G, U, w]$ is U, and there is an edge between two vertices $v_1, v_2 \in U$ if and only if there exists another vertex t not in U that is adjacent to w, v_1 and v_2, and is not adjacent to any other vertex of U. Formally

$$v_1 v_2 \in H[G, U, w] \iff \exists t \in G \setminus U \ (t \sim w \wedge N(t, U) = \{v_1, v_2\}) .$$

Notice that if being in U is first order describable then we can also express the fact that there is an edge between v_1 and v_2 in first order language. A key point in the construction is the representation of general graphs in the form of $H[G, U, w]$. In other words, we show that we can encode any graph on $|U|$ vertices in this form. The fact that we can encode any graph enables us to enrich the language. For example, parity of a set is in general not a first order property, but given a first order describable set U we express the fact that there

is a perfect matching on $|U|$ vertices (which is, of course, equivalent) by saying

$$\exists w \in G \setminus U,\, \forall v_1 \in U,\, \exists! v_2 \in U,\, v_1 v_2 \in H\left[G, U, w\right].$$

That is, saying that there exists a vertex w encoding a perfect matching on U. If $|U|$ is even, there is such a graph and by our assumption there is such a vertex w. If $|U|$ is odd then clearly there is no perfect matching on U. Therefore, assuming that we can encode graphs (using $H\left[G, U, w\right]$ for example) gives us enough expressive power to talk about parity.

In fact, using this method we can express any existential second order property dealing with the vertices of U, as long as we can define it by a first order predicate as we did here. Thus we can interpret a segment of arithmetic and state a sentence with no limiting probability.

The way we represent graphs, that is, the way we defined $H\left[G, U, w\right]$, differs from the matching definition of Shelah and Spencer since in the model studied here we are in a much greater need of avoiding dependencies. Still, this is merely a matter of convenience, and we believe that using the techniques presented in this paper we can prove that there is no limiting probability also for the sentences constructed in [40] and in [45]. On the other hand, showing that the sentence constructed here has no limiting probability in the binomial random graph model is easy, in fact, in a sense, it is easier to demonstrate the nonconvergence of the sentence presented here than doing the same for the original sentence of [40].

We begin the construction with a proposition stating that a. a. s. for every small set U and for every graph H on $|U|$ vertices there is a vertex w in $V \setminus U$ such that $H\left[G_{n,\bar{d}}, U, w\right] \cong H$ as labelled graphs. The proof of the following lemma uses some concentration results that are afterwards stated and proved as Lemmas 6.2 – 6.7.

Proposition 6.1. *Let $d = n^{2/3}$ and consider the random regular graph $G_{n,d}$. Then a. a. s. for every set of vertices U of size $u \leq 2 \ln\ln\ln n$ and for every graph H on u vertices, there exist a vertex $w \notin U$ such that $H\left[G_{n,\bar{d}}, U, w\right]$ is isomorphic to H as labelled graphs.*

Proof. Fix U. For every unordered pair of distinct vertices $\{v_1, v_2\} \subset U$ we define the set C_{v_1,v_2} as the set of vertices out of U that are adjacent to v_1 and v_2 and not adjacent to any other vertex of U. That is,

$$C_{v_1,v_2} = \{t \notin U | N(t, U) = \{v_1, v_2\}\}.$$

Lemma 6.2 states that the codegree of every two vertices is $n^{1/3} \pm n^{1/5}$. We want to estimate the size of C_{v_1,v_2}, so we remove vertices of $N(v_1, v_2)$ that are adjacent to other vertices of U. By Lemma 6.4 we removed at most $u \cdot (\ln n)$ vertices. We conclude that with probability tending to one, for every two vertices $\{v_1, v_2\} \subset U$ one has $|C_{v_1,v_2}| = n^{1/3} \pm n^{1/4}$. In this process we exposed the edges between U and its complement. Now we remove U and by Fact 2.6 we can think of the resulting graph as $G_{n-u,\bar{d}_1}$ where $\max \bar{d} - \bar{d}_1 \leq u$. In addition we also have $\binom{u}{2}$ labelled sets of vertices, all of size $n^{1/3} \pm n^{1/4}$.

Fix a graph H on $[u]$. We wish to show that with probability high enough there is a vertex w that has a neighbor in C_{v_i,v_j} if and only if ij is an edge of H. At this point we can forget about H and rephrase our problem. We are given the random graph $G_{n_1=n-u,\bar{d}_1}$ and $s = \binom{u}{2}$ sets of size $n^{1/3} \pm n^{1/4}$ denoted $C_1, \ldots, C_s$. We need to show that a. a. s. there is a vertex w with neighbors in each of the first k sets $C_1, \ldots, C_k$ and with no neighbor in $C = \bigcup_{i=k+1}^{s} C_i$.

For convenience we relabel the vertices so that the vertices of $C_1, \ldots, C_k$ and of C will be first. We look at the last vertex, n_1, and if it has a neighbor in C_1 we mark it. By Lemma 6.6 n_1 will be marked with some constant positive probability greater than half. In both cases, we put this vertex aside for a moment, and consider the remaining graph. This graph is a random graph on $n_1 - 1$ vertices and with a degree sequence that is very close to $\bar{d}_1$ — every coordinate is either the same or differs by one. Now we look at the last vertex of this new graph, n_2, and again ,with probability greater than half, it has a neighbor in C_1. We continue in this manner for n^ϵ vertices. The number of marked vertices has a binomial distribution with parameters n^ϵ and q, where $q > 1/2$ by Lemma 6.6. Thus, by Chernoff inequality, at the end of the process, with probability tending to one exponentially fast, we have marked more than $\frac{1}{2}n^\epsilon$ vertices. In the process we exposed the edges touching C_1. We remove C_1 and do the same for C_2. Similarly to the above, we will be left with $\frac{1}{4}n^\epsilon$ vertices that have neighbors both in C_1 and C_2. After repeating this k times we shall have $\left(\frac{1}{2}\right)^k n^\epsilon$ vertices with neighbors in every C_i for $i \leq k$. We finish by applying Lemma 6.7 in the same manner — checking each of the $\left(\frac{1}{2}\right)^k$ vertices one by one to see if they have no neighbor in C. For each vertex the probability of success is greater than $1/(10 \ln \ln n)$, and thus the probability of failure — not finding such a vertex — is bounded from above by

$$\left(1 - \frac{1}{10 \ln \ln n}\right)^{n^\epsilon / 2^k} \leq e^{-n^{\epsilon'}}$$

where ϵ' is some positive constant.

We have shown that with probability $1 - e^{-n^{\epsilon'}}$ there is a witness w for H, that is, a vertex w such that $H\left[G_{n,\bar{d}}, U, w\right] \cong H$. Therefore, by a union bound over all the $2^{\binom{u}{2}} \leq e^{(\ln \ln \ln n)^2}$ possible graphs H on $[u]$, we conclude that with probability at least $1 - e^{-0.9n^{\epsilon'}}$ we can find such a witness for U for every graph on $|U|$ vertices. We finish the proof by another union bound, this time over all choices of U. Clearly there are $\sum_{u=1}^{2 \ln \ln \ln n} \binom{n}{u}$ possibilities to choose U. Now, since

$$\sum_{u=1}^{2 \ln \ln \ln n} \binom{n}{u} \leq 2 \ln \ln \ln n \binom{n}{2 \ln \ln \ln n} \leq n^{4 \ln \ln \ln n} \leq e^{\ln n 4 \ln \ln \ln n} \leq e^{(\ln n)^2},$$

the proof is complete. □

The next lemmas contain the technical part of Proposition 6.1 above. Along the proof of the aforementioned proposition, the underlying graph was changed

a few times when several sets of vertices were removed. This changes the degree sequence of the remaining graph. Still, the change in the degree of every vertex is bounded by the number of vertices removed. All in all, $O(\ln n)n^{1/3}$ vertices were removed along the proof, and thus we shall state the following lemmas for graphs with degree sequence $\bar{d} = (d_i)_{i=1}^n$ satisfying that for every i one has $|d_i - n^{2/3}| < n^{1/3+1/100}$. This means that, when applying Corollary 2.3, we are limited to sets of edges of size $o\left(n^{1/3-1/100}\right)$. In practice however, the sets of edges on which we shall apply Corollary 2.3 will be much smaller.

Lemma 6.2. *A. a. s. for every two vertices $v_1, v_2 \in G_{n,\bar{d}}$ one has*

$$codegree(v_1, v_2) = n^{1/3} \pm n^{1/5}.$$

Proof. The proof of this lemma demonstrates our general technique. Fix two vertices, say the first two. For every $3 \leq i \leq n$ let X_i be an indicator random variable for the event "i is adjacent to the first two vertices". Clearly $X = \text{codegree}(1,2) = \sum_{i=3}^n X_i$. Applying Corollary 2.3 with $d = n^{2/3}, e = O(n^{-1/3+1/100})$ and $t = 124$, we can estimate the first moments of X as if it was defined in $G(n, p = d/n)$, up to a multiplicative factor $(1 + o(1))$. The codegree of two fixed vertices in $G(n, p = d/n)$ is a binomial random variable with parameters $n' = n - 2$ and $p' = p^2$. By [38] the k'th central moment of a binomial random variable with parameters n' and p' is $N_k(n'p'(1-p'))$, where N_k is some polynomial[9] of degree $\lfloor k/2 \rfloor$. Therefore, for constant k, the k'th central moment is bounded from above by $c\left(np^2(1-p^2)\right)^{k/2}$, where c is some constant. Thus, the same bound also hold for $E(X - EX)^k$. Taking $k = 62$ and applying Fact 5.15 we get

$$\Pr[|X - n^{1/3}| \geq n^{1/5}] \leq \frac{c\left(np^2(1-p^2)\right)^{\frac{k}{2}}}{(n^{1/5})^k} \leq cn^{62(1/6-1/5)} = o\left(n^{-2}\right),$$

and by a union bound over all pairs of vertices the proof is complete. □

Remark 6.3. In [26] the authors estimate the codegree in random regular graphs. In particular, from the first part of Theorem 2.1 of [26] one can conclude that with probability tending to one, the maximal deviation of the codegree of any two vertices in $G_{n,d=n^{2/3}}$ from $d^{1/3}$ is bounded by $6n^{1/6}\sqrt{\ln n} + C$ where C is constant.

Lemma 6.4. *Let $\bar{d}$ be an almost regular and sparse degree sequence around $d = n^{2/3}$ as above. Then, a. a. s. for every three vertices $u_1, u_2, u_3 \in G_{n,\bar{d}}$ one has*

$$|N^*(u_1, u_2, u_3)| = O(\ln n).$$

Proof. Fix three vertices, say the first three. Let X_i be an indicator random variable for the event "i is adjacent to the first three vertices" and let $X = \sum X_i$. In $G(n,p)$ the distribution of X is Binomial with parameters $n-3$ and $1/n$ which

[9] When k is even the leading term of N_k is $k!!$.

is asymptotically Poisson with $\lambda = 1$. The factorial moments of X in $G(n,p)$ are thus simply $E_k(X) = \lambda^k(1+o(1)) = 1+o(1)$. According to Corollary 2.3, the factorial moments in $G_{n,\bar{d}}$ are of the form $E_k(X) = 1+o(1)$ as well. To bound the probability that X is large we apply a corollary of Bonferroni's inequality (here we use the formulation of [7, Corollary 1.14]): If X is a random variable with values in $\{0,1,\ldots,n\}$ then

$$\Pr[X \geq k] = \sum_{r=k}^{n} (-1)^{k+r} \frac{E_r(X)}{(k-1)!(r-k)!r},$$

and, in addition, if $k+s$ is odd and $k+t$ is even then

$$\sum_{r=k}^{k+s} (-1)^{k+r} \frac{E_r(X)}{(k-1)!(r-k)!r} \leq \Pr[X \geq k] \leq \sum_{r=k}^{k+t} (-1)^{k+r} \frac{E_r(X)}{(k-1)!(r-k)!r}. \quad (5)$$

Take k to be the smallest even integer greater than $\ln n$ and take $t = 0$. Then from the right hand side of the inequality above we get

$$\Pr[X \geq \ln n] \leq \frac{1+o(1)}{(\ln n - 1)!(\ln n)} \leq \frac{n}{(\ln n)^{\ln n}} = n^{-\ln\ln n - 1}.$$

By a union bound over all triplets of vertices the proof is complete. □

Lemma 6.5. *Fix three vertices $u_1, u_2, u_3 \in G_{n,\bar{d}}$. The probability of them having a common neighborhood of size $|N^*(u_1,u_2,u_3)| \geq \ln\ln n$ is bounded from below by*

$$\Pr[|N^*(u_1,u_2,u_3)| \geq \ln\ln n] \geq (\ln n)^{-\ln\ln\ln n}.$$

Proof. In the previous proof we have estimated the factorial moments of X, the number of common neighbors of some fixed three vertices in $G_{n,\bar{d}}$, as $E_k(X) = 1+o(1)$. Using the left hand side of Equation (5) we get that

$$\Pr\left[X \geq \ln\ln n\right] \geq \sum_{r=k}^{k+s} (-1)^{k+r} \frac{E_r(X)}{(k-1)!(r-k)!r},$$

where $k \leq \ln\ln n$ and $k+s$ odd. Take $s = 0$ and take k to be the largest odd integer smaller or equal to $\ln\ln n$. Then

$$\Pr\left[X \geq \ln\ln n\right] \geq (-1)^{2k} \frac{E_k(X)}{(k-1)!k} \geq \frac{1+o(1)}{(\ln\ln n)!}.$$

Using Stirling's approximation we conclude

$$\Pr\left[X \geq \ln\ln n\right] \geq \frac{1+o(1)}{((\ln\ln n)/e)^{\ln\ln n+1}} \geq$$
$$\geq \frac{e}{\ln\ln n} \frac{\ln n}{(\ln n)^{\ln\ln\ln n}} \geq (\ln n)^{-\ln\ln\ln n},$$

as required. □

The next two lemmas deal with the probability of a vertex to send an edge to a vertex set of size $n^{1/3}$ or to avoid a slightly larger set of vertices. In the binomial random graph model this probability can be computed directly, here we utilize a mechanism similar to the one used in the proofs above.

Lemma 6.6. *Fix a vertex v and a set of vertices C of size $n^{1/3} \pm n^{1/4}$. The probability in $G_{n,\bar{d}}$ that there is an edge between v and at least one vertex from C is at least $1 - \frac{1.1}{e}$.*

Proof. Let $X_1 = d(v, C)$ be a random variable counting the number of edges between v and vertices from C in $G_{n,\bar{d}}$. Define X_2 for the same quantity in $G(n, p = d/n)$. Notice that in both probability spaces $X_i = \sum X_{i,u}$ where $X_{1,u}$ is an indicator random variable for the event "uv is an edge in $G_{n,\bar{d}}$" and $X_{2,u}$ is the number of such edges in $G(n, p)$. Thus, by Corollary 2.3, if we let $\mu_k(X_i)$ denote the k'th moment of X_i, we may write

$$\mu_k(X_1) = \mu_k(X_2)(1 + o(1)),$$

for $k = O(\ln n)$. The first k factorial moment of a random variable is a linear function of the first k moments, and thus for every constant k we also have

$$E_k(X_1) = E_k(X_2)(1 + o(1)),$$

where $E_k(X_i)$ is the k'th factorial moment of X_i. The number of edges between v and C has a Poisson distribution in $G(n, p = n^{-1/3})$ with expectation $\lambda = np = 1 + o(1)$. The factorial moments of a random variable with a Poisson distribution have a nice form, $E_k(X_2) = \lambda^k$, and by the above we also have $E_k(X_1) = \lambda^k(1 + o(1))$. Applying Bonferroni's inequality we get that if k is an even and large enough constant then

$$\Pr[X_1 = 0] \leq \sum_{r=0}^{k} (-\lambda)^r / r! \leq \frac{1.1}{e},$$

which is what we wanted to show. □

Lemma 6.7. *Fix a vertex v and a set of vertices C of size $O(n^{1/3} \ln\ln\ln n)$. The probability in $G_{n,\bar{d}}$ that there is no edge between v and at least one vertex from C is at least $(10 \ln\ln n)^{-1}$.*

Proof. We mimic the proof of the previous lemma. This time X_2 has a Binomial distribution with parameters $m = |C| = O(n^{1/3} \ln\ln\ln n)$ and $p = n^{-1/3}$. The k'th factorial moment of X_2 is then $E_k(X_p) = p^k m(m-1)\cdots(m-k+1) = (pm)^k(1 + O(1/m)) = (\ln\ln\ln n)^k(1 + O(1/m))$. This time, instead of applying Corollary 2.3 directly, we plug in the parameters of the current proof into the displayed equation in the proof of Corollary 2.3. This yields $E_k(X_1) = (\ln\ln\ln n)^k(1 + O(1/m))\left(1 + o\left(n^{-1/4}\right)\right)$. Applying the other side of Bonferroni's inequality we get

$$\Pr[X_1 = 0] \geq \left(1 + o\left(n^{-1/4}\right)\right) \sum_{r=0}^{k} (-\ln\ln\ln n)^r / r!.$$

Take $k \geq 10 \ln\ln n \ln\ln\ln n$ to be odd. Then

$$\sum_{r=k}^{\infty} \frac{(-\ln\ln\ln n)^r}{r!} \leq \sum_{r=k}^{\infty} \frac{(\ln\ln\ln n)^r}{(r/e)^r} \leq$$
$$\sum_{r=k}^{\infty} \left(\frac{e\ln\ln\ln n}{r}\right)^r \leq \sum_{r=k}^{\infty} \left(\frac{e}{10\ln\ln n}\right)^r \leq \frac{2e}{10\ln\ln n}.$$

Which implies $\Pr[X_1 = 0] \geq \frac{1}{10\ln\ln n}$. □

The following lemma describes the last graph property needed for the construction to work.

Lemma 6.8. *Let $\bar{d}$ be a degree sequence as above, then a. a. s. there exist three vertices with common neighborhood of size N, where $N \geq \ln\ln n$ and $N = O(\ln n)$.*

Proof. Let $\epsilon > 0$ be a small constant and fix $K = n^\epsilon$ triplets of vertices. By Lemma 6.5, the probability that the common neighborhood of the first triplet is of size at least $\ln\ln n$ is at least $\frac{1}{2}(\ln n)^{-\ln\ln\ln n}$. If this is the case we have succeeded and we may stop, otherwise we remove the first triplet and continue. By Fact 2.6 we are left with another random graph with a given degree sequence. Clearly, the degree sequence of the remaining graph still satisfies the requirements of Lemma 6.5 so we may continue the process for the next triplet. We stop when a triplet having a common neighborhood of size at least $\ln\ln n$ is found, and then the process succeed, or when we run out of triplets and the process fails. The probability of the process to fail is

$$\left(1 - \frac{1}{2}(\ln n)^{-\ln\ln\ln n}\right)^{n^\epsilon} \leq e^{-cn^\epsilon/(\ln n)^{\ln\ln\ln n}} \leq e^{-n^{\epsilon'}}$$

for some constants c and $0 < \epsilon' < \epsilon$. We conclude — the probability of failure is $o(1)$ and so the process a. a. s. succeeds. It may be that some of the vertices that were removed during the process are also in the common neighborhood of the triplet that stopped the process, but this will only increase the size of its common neighborhood. On the other hand, by Lemma 6.4 this size is $O(\ln n)$. □

Now we can describe the nonconverging sentence and sketch the proof of Theorem 1.8. The sentence deals with common neighborhoods of triplets, which are first order describable sets, since we can write $v \in N^*(x, y, z)$ as $(v{\sim}x) \wedge (v{\sim}y) \wedge (v{\sim}z)$. We would like to be able to compare the sizes of two sets, so we define the following predicate (PB stands for "Pseudo Bigger"). $\mathrm{PB}(S, T)$ means that there exists a vertex $w \in G_{n,d}$ such that $H[G_{n,d}, S \cup T, w]$ has the following properties:

1. Every vertex $t \in T$ has a neighbor $s \in S$.
2. The common neighborhoods in S of any two distinct vertices $t, t' \in T$ is empty.

3. There exists a vertex $s \in S$ who has no neighbor in T.

If such a vertex w exists, it of course means that S is larger than T. The contrapositive is not correct, but if in addition to $|S| > |T|$ it is also true that both sets are smaller than $\ln\ln\ln n$, then by Proposition 6.1 such a witness a. a. s. exists. In particular, we can talk about a "pseudo maximal" triplet, that is, a triplet u_1, u_2, u_3 such that for no other triplet v_1, v_2, v_3 one has $\mathrm{PB}(N^*(v_1, v_2, v_3), N^*(u_1, u_2, u_3))$. Soon we shall talk about triplets having some additional property, but in the meantime note that the size of the common neighborhood of a pseudo maximal triplet is at least $\ln\ln\ln n$ and at the same time also $O(\ln n)$.

The next part of the construction enables us to refer to the size of certain first order describable sets. This is again done by using Proposition 6.1 to encode some second order properties. Since this part is identical to the matching part of the proof of Shelah and Spencer, we refer the reader to [40] or [45] where it is brightly explained in full detail. All one needs to know is the fact that there is a first order predicate $\mathrm{ARITH}(u_1, u_2, u_3)$ enabling us to talk about the size of the common neighborhood of u_1, u_2, u_3, and also about arithmetical relations involving this quantity and constants. Given ARITH we can define (as [40] does) another first order predicate, $\mathrm{GAP}(u_1, u_2, u_3)$ that can be defined whenever ARITH holds. $\mathrm{GAP}(u_1, u_2, u_3)$ is true if and only if $1 \leq \log^* |N^*(u_1, u_2, u_3)| \leq 50 \pmod{100}$, where $\log^*(n)$ is the smallest number of base two logarithms needed to be taken in order to reduce n below one.

Finally we can write A, the nonconverging first order sentence.

$$\begin{aligned}&\exists u_1, u_2, u_3,\ \mathrm{ARITH}(u_1, u_2, u_3) \wedge \\ &\quad [\forall v_1, v_2, v_3,\ \mathrm{ARITH}(v_1, v_2, v_3) \rightarrow \neg \mathrm{PB}(N^*(v_1, v_2, v_3), N^*(u_1, u_2, u_3))] \wedge \\ &\quad \mathrm{GAP}(u_1, u_2, u_3).\end{aligned}$$

That is, A states that the pseudo maximal triplet among the triplets that satisfy ARITH also satisfies GAP. Why does A not converge? First notice that if u_1, u_2, u_3 is a maximal triplet satisfying ARITH, then $\ln\ln\ln n \leq |N^*(u_1, u_2, u_3)| = O(\ln n)$. This is true since by Lemma 6.5 there are triplets with common neighborhood of size $\ln\ln n$, and by Proposition 6.1, every vertex set of size at most $2\ln\ln\ln n$ satisfies ARITH (satisfying ARITH means that we can find a few witnesses $w_1, \ldots, w_m$ for some $m \leq 10$ graphs on this set). On the other hand no triplet has a common neighborhood of size $\omega(\ln n)$. Now, if $\log^*(n) = 50 \pmod{100}$, then a. a. s. the size of the common neighborhood of a maximal triplet u_1, u_2, u_3 satisfying ARITH will be between $\ln\ln\ln n$ and $C\ln n$. Therefore we shall have $47 \leq \log^*(|N^*(u_1, u_2, u_3)|) \leq 49 \pmod{100}$, and $\mathrm{GAP}(u_1, u_2, u_3)$ together with A will hold. On the other hand, if $\log^*(n) = 0 \pmod{100}$, then the aforementioned quantity will a. a. s. satisfy $97 \leq \log^*(|N^*(u_1, u_2, u_3)|) \leq 99 \pmod{100}$, then $\mathrm{GAP}(u_1, u_2, u_3)$ will be false and so will be A. Since $\log^*(n) \pmod{100}$ will visit both 0 and 50 infinitely many times as n grows to infinity, A has no limiting probability. Moreover, its infimum limit is zero and its supremum limit is one. The proof is complete.

7 Concluding remarks and open questions

As already mentioned in the introduction, results for random regular graphs with degree asymptotically higher than $\ln n$ are scarce and require much effort (this paper being another example of this phenomenon). Still, when examining these results we see that they are very similar to the corresponding results from the binomial random graph model. This should not come as a surprise, since when $d = \omega(\ln n)$ a similarity between $G_{n,d}$ and $G(n, p = d/n)$ may be expected, as asymptotically the latter becomes typically close to being regular.

One result that may be seen as an attempt to make the last reasoning into a concrete theorem was made by Kim and Vu in [23]. Their result, and even more, their Sandwich Conjecture, describe a useful connection between the two models. While our research also studies the relationship between random regular graphs and random binomial graphs, the results are not comparable. For example, the sandwich conjecture, if proven, can be used for a one line derivation of the asymptotic value of the chromatic number of $G_{n,d}$, a graph invariant which is not describable in the first order language. On the other hand, containing some small graph as an induced subgraph is generally not a monotone property which prevents application of the Sandwich Conjecture, but it is a first order property, so one can learn its behavior using our results.

The sentence A constructed in Section 6 does not have a limiting probability both in the random binomial model and in the random regular model. Still, it may be the case that the connection between the models remains. Is it true that $\lim[\Pr[G(n,p) \models A] - \Pr[G_{n,d} \models A]] = 0$? A more general question was asked by Joel Spencer [44]: Fix a rational $0 < \alpha < 1$. Is it true that for every first order sentence A one has

$$\lim \left[\Pr\left[G(n, p = n^{-\alpha}) \models A\right] - \Pr\left[G_{n,d=n^{1-\alpha}} \models A\right]\right] = 0?$$

Or, in Spencer's words — We know that in this region there is a mess both in $G(n,p)$ and in $G_{n,d}$, but is it the *same* mess?

As mentioned earlier, the first order behavior of very sparse random regular graphs was studied by Lynch [30]. His result was in the positive, to some extent, showing that every first order sentence A has a limiting probability. Can we say anything regarding the general form of this limiting probability? Results of this type were obtained for several regions of the binomial model by Lynch [29] and by Spencer and Thoma [43].

In this paper we have considered degree sequences that are nearly regular. What happens if we change the underlying degree sequence? More specifically, random graphs with a degree sequence obeying a power law were well-studied in recent years since they resemble properties of real-world complex networks more closely then regular or nearly regular graphs. Two of the more studied models are the "Preferential attachment" model and the "copying" model. Is it true that every first order property has a limiting probability in these models? It is worth mentioning that there are many flavors of, say, preferential attachment graphs. It will be satisfying to know that convergence of first order properties is independent in the specific implementation of the preferential attachment rule

and the base setting. We mention again the paper of Kleinberg and Kleinberg dealing with limits of first order properties in preferential attachment random graphs [24].

One conclusion naturally arising from this direction of research is that the first order language is weak. This is not surprising as the following example demonstrates. Consider the binomial model of random graphs with $p = \lambda/n$. As mentioned in the introduction, Lynch showed in [29] that every first order property A has a limiting probability. Moreover, that probability is of a known form and is continuous in λ. On the other hand, for every $\epsilon > 0$, there is a vast difference in the appearance and behavior of $G(n, p = (1-\epsilon)/n)$ and $G(n, p = (1+\epsilon)/n)$. In the first case all the components of the graphs are small (the largest being of size $O(\ln n)$), and they are all trees or unicyclic. The situation is very different when $p = (1+\epsilon)/n$, where the graph contains a giant component of linear size having much more edges than vertices. The remaining vertices lie in small components of size $O(\ln n)$ which are all trees or unicyclic. This phenomenon was dubbed "the double jump" by Erdős and Rényi and today it is often thought of as a percolation phenomenon or phase transition. Still, our point remains, the famous Erdős - Rényi phase transition of the binomial model is completely unseen when looking through first order lens, as, for example, nothing is special when $\lambda = 1$. The following question arises naturally: Can we find a stronger language of graphs being able to express "natural" graph properties and the same time having theorems like Theorem 1.3 or Zero-One Laws like Corollary 1.4 still holding? Variants of this question were asked by many (e. g. Shelah [39], Lynch [29] and Compton [10]) and is still open.

A few attempts were made in this direction, but, to the extent of our knowledge, all ended in languages too expressive in which one can describe properties with no limiting probability (in $G(n,p)$ with constant p, for example), or languages too weak unable to express, say, k-colorability. It worth mentioning that $G(n,p)$ with constant $0 < p < 1$ is asymptotically almost surely connected, Hamiltonian and can not be colored with k colors for every fixed integer k. In the following we shall describe some of these endeavours, focusing in those that are relevant to out study.

The first strengthening that comes to one's mind us moving from first order logic to second order logic. Already Fagin discussed second order language in his paper [15], showing that it is too expressive to obey even a convergence law. In [21] Kaufmann and Shelah demonstrated that there exists an *existential monadic* sentence A such that $\Pr[G(n, \frac{1}{2}) \models A]$ does not converge. In existential monadic language one may quantify over relations, as in second order logic, but with severe restriction — second order quantification is limited to existential quantifiers quantifying only over unary relations (i. e. over sets). This shows that even a small fraction of the power of second order language is enough to ruin convergence.

In [39] Shelah studied the *ordered* first order language, i. e., the first order language strengthened by an order relation over the vertices. He has shown that while there are sentences with no limit probability in this langauge, the

following holds for every *ordered* first order sentence A:

$$\lim_{n\to\infty}\left(\Pr\left[G\left(n+1,\frac{1}{2}\right)\models A\right]-\Pr\left[G\left(n,\frac{1}{2}\right)\models A\right]\right)=0.$$

Can similar results be obtained for other situations in which there are sentences with no limiting probabilities? In particular, can such a "weak Zero-One law" be proven for sentences in the standard first order language where the graphs are distributed according to binomial model $G(n,p=n^{\alpha})$ or the regular random graph model $G(n,d=n^{1-\alpha})$ for rational $0<\alpha<1$?

In [25] Kolaitis and Kopparty considered the language resulted by adding a *parity* quantifier (or, more generally, a Mod_k quantifier) to the first order language of graphs. Take $p\in(0,1)$ constant. The authors showed that for every sentence φ in that language there are two rational numbers a_0,a_1 such that

$$\lim_{n\to\infty}\Pr[G(2n+i,p)\models\varphi]=a_i\quad;\quad i=0,1.$$

Their result extends naturally to the language we get by adding a modulus k quantifier instead of parity, for any fixed integer k.

Another approach was taken by Naor et. al. in [36] where they let the quantifier depth to go to infinity with n and considered the maximal quantifier depth for which the Zero-One law for $G(n,p)$ with constant p still holds. This gives rise to a few more questions. For example. what quantifier depth will enable us to spot the Erdős - Rényi phase transition? More relevant to our study is the quantifier depth that will let us distinguish between, say, $G(n,p=n^{-1/\pi})$ and $G_{n,d=np}$.

References

[1] A. Agrawal and S. Joel. Undecidable statements and the zero-one law in random geometric graphs. *Discrete and Computational Geometry*. accepted.

[2] N. Alon and J. H. Spencer. *The Probabilistic Method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley, New Jersey, Third edition, Aug. 2008.

[3] E. A. Bender and E. R. Canfield. The asymptotic number of labeled graphs with given degree sequences. *Journal of Combinatorial Theory, Series A*, 24(3):296–307, May 1978.

[4] E. A. Bender, K. J. Compton, and L. B. Richmond. 0-1 laws for maps. *Random Struct. Algorithms*, 14(3):215–237, 1999.

[5] P. Boldi and S. Vigna. More lower bounds for weak sense of direction: The case of regular graphs. (manuscript).

[6] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combinatorics*, 1:311–316, 1980.

[7] B. Bollobás. *Random Graphs*. Number 73 in Cambridge studies in advanced mathemtics. Cambridge University Press, Cambridge, UK, Second edition, Jan. 2001.

[8] B. Bollobás. *Modern Graph Theory*. Graduate Texts in Mathematics. Springer, New York, Aug. 2002. Corrected Edition.

[9] A. Bonato. *A Course on the Web Graph*. Number 89 in Graduate Studies in Mathematics. American Mathematical Society, 2008.

[10] K. J. Compton. 0-1 laws in logic and combinatorics. In R. I., editor, *Algorithms and Order*, Advanced Study institute, pages 353–383. Kluwer Academic Publishers, Dordrecht, 1989.

[11] C. Cooper, A. M. Frieze, and B. A. Reed. Random regular graphs of non-constant degree: Connectivity and hamiltonicity. *Combinatorics, Probability & Computing*, 11(3):249–262, 2002.

[12] C. Cooper, A. M. Frieze, B. A. Reed, and O. Riordan. Random regular graphs of non-constant degree: Independence and chromatic number. *Combinatorics, Probability & Computing*, 11(4):323–342, 2002.

[13] R. Diestel. *Graph Theory*. Graduate Texts in Mathematics. Springer, New York, Third edition, Feb. 2006.

[14] A. Ehrenfeucht. An application of games to the completeness problem for formalized theories. *Fundamenta Mathematicae*, 49:129–141, 1961.

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

[16] R. Fraïssé. Sur une nouvelle classification des systèmes de relations. *Comptes Rendus*, 230:1022–1024, 1950.

[17] R. Fraïssé. *Sur quelques classifications des systèmes de relations*. PhD thesis, University of Paris, 1953. Published in [18].

[18] R. Fraïssé. Sur quelques classifications des systèmes de relations. *Publications Scientifiques de l'Université d'Alger, series A*, 1:35–182, 1954.

[19] Y. V. Glebskii, D. I. Kogan, M. I. Liogon'kii, and V. A. Talanov. Range and degree of realizability of formulas in the restricted predicate calculus. *Cybernetics and Systems Analysis*, 5(2):142–154, Mar. 1969. (Russian original: Kibernetika 5(2):17-27, March-April 1969).

[20] S. Janson, T. Łuczak, and A. Ruciński. *Random Graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley, New York, 2000.

[21] M. Kaufmann and S. Shelah. On random models of finite power and monadic logic. *Discrete Mathematics*, 54(3):285–293, 1985.

[22] J. H. Kim, B. Sudakov, and V. Vu. Small subgraphs of random regular graphs. *Discrete Mathematics*, 307(15):1961–1967, July 2007.

[23] J. H. Kim and V. H. Vu. Sandwiching random graphs. *Advances in Mathematics*, 188(2):444–469, Nov. 2004.

[24] R. D. Kleinberg and J. M. Kleinberg. Isomorphism and embedding problems for infinite limits of scale-free graphs. In *Proceedings of the 16th ACM-SIAM Symposium on Discrete Algorithms*, pages 277–286, 2005.

[25] P. G. Kolaitis and S. Kopparty. Random graphs and the parity quantifier. In M. Mitzenmacher, editor, *STOC*, pages 705–714. ACM, 2009.

[26] M. Krivelevich, B. Sudakov, V. H. Vu, and N. C. Wormald. Random regular graphs of high degree. *Random Structures and Algorithms*, 18(4):346–363, July 2001.

[27] M. Krivelevich, B. Sudakov, and N. C. Wormald. Regular induced subgraphs of a random graph. *Preprint*.

[28] T. Łuczak and J. Spencer. When does the zero-one law hold? *Journal of the American Mathematical Society*, 4(3):451–468, July 1991.

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

[30] J. F. Lynch. Convergence law for random graphs with specified degree sequence. *ACM Transactions on Computational Logic*, 6(4):727–748, 2005.

[31] G. L. McColm. First order zero-one laws for random graphs on the circle. *Random Structures and Algorithms*, 14(3):239–266, 1999.

[32] B. D. McKay. Subgraphs of random graphs with specified degrees. *Congressus Numerantium*, 33:213–223, 1981.

[33] B. D. McKay and N. C. Wormald. Asymptotic enumeration by degree sequence of graphs of high degree. *European J. Combinatorics*, 11:565–580, 1990.

[34] B. D. McKay and N. C. Wormald. Uniform generation of random regular graphs of moderate degree. *Journal of Algorithms*, 11:52–67, 1990.

[35] B. D. McKay and N. C. Wormald. Asymptotic enumeration by degree sequence of graphs with degress $o\left(n^{1/2}\right)$. *Combinatorica*, 11(4):369–382, 1991.

[36] M. Naor, A. Nussbaum, and E. Tromer. Efficiently constructible huge graphs that preserve first order properties of random graphs. *Preprint*.

[37] M. Penrose. *Random Geometric Graphs*. Number 5 in Oxford Studies in Probability. Oxford University Press, USA, July 2003.

[38] V. Romanovsky. Note on the moments of a binomial $(p + q)^n$ about its mean. *Biometrika*, 15(3/4):410–412, Dec. 1923.

[39] S. Shelah. Very weak zero one law for random graphs with order and random binary functions. *Random Structures and Algorithms*, 9(4):351–358, 1996.

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

[41] J. Spencer. Zero-one laws with variable probability. *Journal of Symbolic Logic*, 58(1):1–14, 1993.

[42] J. Spencer and G. Tardos. Ups and downs of first order sentences on random graphs. *Combinatorica*, 20(2):263–280, 2000.

[43] J. Spencer and L. Thoma. On the limit values of probabilities for the first order properties of graphs. In R. L. Graham, J. Kratochvíl, J. Nešetřil, and F. S. Roberts, editors, *Contemporary Trends in Discrete Mathematics*, volume 49 of *DIAMACS*, pages 317–336. American Mathematical Society, 1999.

[44] J. H. Spencer. private communication.

[45] J. H. Spencer. *The Strange Logic of Random Graphs*. Number 22 in Algorithms and Combinatorics. Springer-Verlag, Berlin, 2001.

[46] D. B. West. *Introduction to Graph Theory*. Prentice Hall, New Jersy, Second edition, Sept. 2000.

[47] P. Winkler. Random structures and zero-one laws. In S. N. W., W. R. E., and S. B., editors, *Finite and Infinite Combinatorics in Sets and Logic*, Advanced Science Institutes, pages 399–420. Kluwer Academic Publishers, Dordrecht, 1993.

[48] N. C. Wormald. Models of random regular graphs. In J. Lamb and D. Preece, editors, *Surveys in Combinatorics, 1999*, volume 276 of *London Mathematical Society Lecture Note Series*, pages 239–298. Cambridge University Press, Cambridge, 1999.