UNIVERSALITY OF GRAPHS WITH FEW TRIANGLES AND ANTI-TRIANGLES

DAN HEFETZ AND MYKHAYLO TYOMKYN

Abstract. We study *3-random-like* graphs, that is, sequences of graphs in which the densities of triangles and anti-triangles converge to 1/8. Since the random graph $\mathcal{G}_{n,1/2}$ is, in particular, 3-random-like, this can be viewed as a weak version of quasirandomness. We first show that 3-random-like graphs are 4-universal, that is, they contain induced copies of all 4-vertex graphs. This settles a question of Linial and Morgenstern [10]. We then show that for larger subgraphs, 3-random-like sequences demonstrate a completely different behaviour. We prove that for every graph H on $n \geq 13$ vertices there exist 3-random-like graphs without an induced copy of H. Moreover, we prove that for every ℓ there are 3-random-like graphs which are ℓ-universal but not m-universal when m is sufficiently large compared to ℓ.

1. Introduction

A graph is called ℓ-universal if it contains every ℓ-vertex graph as an induced subgraph. Universality is a well-studied graph property, for instance, the famous Erdős-Hajnal conjecture [7] can be formulated in the following form.

Conjecture 1.1 (Erdős-Hajnal). *For every integer ℓ there exists an $\varepsilon > 0$ such that every n-vertex graph G with no clique or independent set of size n^ε is ℓ-universal.*

Recently Linial and Morgenstern [10] asked a question of a similar flavour. Instead of forbidding large cliques and independent sets (anti-cliques) they asked, what happens if the graph G contains only *few* cliques and anticliques of a certain order m. The present paper addresses this question.

First, let us introduce some useful notation and terminology, most of which is standard (see e.g. [4]). For a graph G write $V(G)$ and $E(G)$ for its sets of vertices and edges, respectively. Let $|G| = |V(G)|$ denote the *order* of G and let $e(G) = |E(G)|$ denote its *size*. The *complement* of G is denoted by $\overline{G}$. For a set $S \subseteq V(G)$ put $G[S]$ for the subgraph of G induced on the set S. For a set $S \subseteq V(G)$ and a vertex $u \in V(G)$, let $N_G(u, S) = \{w \in S : uw \in E(G)\}$ denote the set of neighbours of u in S and let $d_G(u, S) = |N_G(u, S)|$ denote the *degree* of u into S. We abbreviate $N_G(u, V(G))$ to $N_G(u)$ and $d_G(u, V(G))$ to $d_G(u)$. The former is referred to as the *neighbourhood* of u in G and the latter as its *degree*. We use $d_G(u, v)$ to denote the *co-degree* of u and v, that is, $|N_G(u) \cap N_G(v)|$ and the somewhat less standard $d_G(u, -v)$ to denote $|N_G(u) \setminus N_G(v)|$. Often, when there is no risk of confusion, we omit the subscript G from the notation above.

For graphs G and H, put $D_H(G)$ for the number of induced copies of H in G and $p_H(G)$ for the corresponding density:

$$p_H(G) = \binom{n}{|H|}^{-1} \cdot D_H(G)\,.$$

The quantity $p_H(G)$ can be also interpreted as the probability that a randomly picked set of $|H|$ vertices of G induces a copy of H.

For $H = K_2$, a single edge, $D_H(G)$ is simply $e(G)$ and thus we write $p_e(G)$ for $p_{K_2}(G)$, the *edge density* of G. For graphs of order 3, since they are determined up to isomorphism by their size, we write $D_i(G)$ for $D_H(G)$ and $p_i(G)$ for $p_H(G)$, where $i = e(H)$. The vector $(p_0(G), \ldots, p_3(G))$ is called the 3-*local profile* of G.

2000 *Mathematics Subject Classification.* 05C35, 05C60.

Research supported by EPSRC grant EP/K033379/1.

Let $\mathcal{G} = (G_k)_{k=1}^{\infty}$ be a sequence of graphs, where $G_k = (V_k, E_k)$ is of order $n_k := |V_k|$ and n_k tends to infinity with k. If for some graph parameter λ the limit $\lim_{k\to\infty} \lambda(G_k)$ exists, we denote it by $\lambda(\mathcal{G})$. A sequence $\mathcal{G}$ is said to be *ℓ-universal* if G_k is ℓ-universal for every sufficiently large k.

Linial and Morgenstern proved in [10] that there exists a constant $\rho = 0.159181\ldots$ such that every $\mathcal{G}$ with $p_0(\mathcal{G}), p_3(\mathcal{G}) < \rho$ is 3-universal and asked whether an analogous result holds for higher universalities.

Question 1.2 ([10])**.** *Given $\ell \geq 4$, is there some $\varepsilon > 0$ such that every graph sequence $\mathcal{G}$ with $p_0(\mathcal{G}), p_3(\mathcal{G}) < \frac{1}{8} + \varepsilon$ is ℓ-universal?*

Note that our definition of *ℓ-universal* sequences is slightly different from the one given in [10]. The latter required additionally that $p_H(G_k)$ be bounded away from 0 for each H of order ℓ. However for our purposes (i.e. answering Question 1.2) these definitions are equivalent due to the induced graph removal lemma of Alon, Fischer, Krivelevich and Szegedy [1].

It was pointed out by the second author that for every $\ell \geq 5$ the answer to Question 1.2 is negative. Though his counterexample has already appeared in [10], for the sake of completeness we will repeat it in the next section of the present paper.

This leaves $\ell = 4$ as the only remaining open case of Question 1.2. Our first main result in this paper, Theorem 1.3, answers it in the affirmative, thereby settling Question 1.2 in full.

Let us define a sequence of graphs $\mathcal{G}$ to be *t-random-like*, or *tRL* for brevity, if $p_{K_t}(\mathcal{G}) = p_{\overline{K_t}}(\mathcal{G}) = 2^{-\binom{t}{2}}$. Our choice of terminology stems from the fact that such a sequence has approximately the same number of t-cliques and t-anticliques, that is, independent sets of size t, as the random graph $\mathcal{G}_{n,1/2}$. Note that for $\mathcal{G}$ to be 2RL it is sufficient to have $p_e(\mathcal{G}) = 1/2$. We will be mostly interested in 3RL sequences; in our terminology $\mathcal{G}$ is 3RL if and only if $p_0(\mathcal{G}) = p_3(\mathcal{G}) = 1/8$.

A standard diagonalisation argument shows that in order to answer Question 1.2 for $\ell = 4$ affirmatively, it suffices to prove the following assertion.

Theorem 1.3. *Every 3RL sequence is 4-universal.*

Theorem 1.3 is related to the *quasirandomness* of graphs as well. This is a central notion in extremal and probabilistic graph theory. It was introduced by Thomason in [14] and was extensively studied in many subsequent papers. In particular, it was proved by Chung, Graham and Wilson [5] (see also [2] for more details) that if $p_H(\mathcal{G}) = p_H(\mathcal{G}_{n,1/2})$ holds for every graph H of order 4, then the same equality holds for every graph H of *any* fixed size. In the terminology of [5] this fact is denoted by $P_1(4) \Rightarrow P_1(s)$. On the other hand, it was pointed out in [5] that the property $P_1(3)$, that is, containing the "correct" number of induced copies of every 3-vertex graph, is not sufficient to ensure quasirandomness. As we shall see in Section 2, $P_1(3)$ is in fact equivalent to 3RL. Thus, our results in this paper can be viewed as the study of $P_1(3)$. Under this viewpoint Theorem 1.3 shows that, while 3RL graphs need not satisfy $P_1(4)$, they still must contain a positive density of every possible induced 4-vertex graph.

Having resolved Question 1.2, we know that 3RL implies 4-universality, but is not enough to ensure ℓ-universality for any larger ℓ. A natural follow up question to ask is, whether there still exist infinite classes of graphs H that must be contained in every 3RL sequence $\mathcal{G}$. Cliques, paths, cycles and stars are natural candidates for such classes. We shall answer this question in the negative by providing counterexamples for each of these classes. In fact, our second main result, Theorem 1.4 provides, perhaps surprisingly, a counterexample for *any single* graph which is not too small.

Theorem 1.4. *For every graph H of order at least 13 there exists a 3RL sequence $\mathcal{G}$, where no $G_k \in \mathcal{G}$ contains a copy of H as an induced subgraph.*

Theorem 1.4 combined with Theorem 1.3 and the induced graph removal lemma [1] immediately give the following corollary.

Corollary 1.5. *There exists an $\varepsilon > 0$ such that for every graph H of order at least 13 there is a sequence $\mathcal{G}$, where no $G_k \in \mathcal{G}$ contains an induced copy of H, but $p_J(\mathcal{G}) > \varepsilon$ for every 4-vertex graph J.*

Theorem 1.4 and Corollary 1.5 show that for $\ell \geq 13$, having either the "correct" densities of triangles and anti-triangles or positive densities of every 4-vertex graph is far from being enough to ensure ℓ-universality. This goes in stark contrast with $\mathcal{G}$ having the "correct" densities of *all* induced 4-vertex graphs, which implies that $\mathcal{G}$ is quasirandom and therefore ℓ-universal for every ℓ.

Having constructions of 3RL sequences which are only ℓ-universal for very small values of ℓ on the one hand and the random graph $\mathcal{G}_{n,1/2}$ (which is ℓ-universal for every fixed ℓ) on the other hand, it is natural to ask, if for arbitrarily large ℓ there exists a 3RL sequence which is ℓ-universal but not $f(\ell)$-universal for some function f. This would show that no fixed universality is sufficient to ensure all other universalities. Our third theorem shows that this is indeed the case in the following strong sense.

Theorem 1.6. *For every ℓ there exists a 3RL sequence $\mathcal{G}_\ell$ such that $p_H(\mathcal{G}_\ell) > 0$ for every graph H of order 2^ℓ, but $\mathcal{G}_\ell$ is not $24\ell \cdot 2^\ell$-universal.*

The rest of this paper is organised as follows. In the next section we establish some basic properties of 3RL sequences and recall the construction of a 3RL sequence which is not 5-universal. In Section 3 we prove our first main result, Theorem 1.3, by considering each 'forbidden' 4-vertex subgraph individually and applying different methods in different cases. In Section 4 we prove our second main result, Theorem 1.4. This will be achieved through a probabilistic construction of a 3RL sequence in which no graph contains a clique of size 5. We also give a second construction of a 3RL sequence, in which no graph contains a clique of size 10; we believe this construction to be of independent interest, since it is entirely deterministic. In Section 5 we prove Theorem 1.6 by adapting a construction of Chung, Graham and Wilson from their seminal paper on quasirandomness [5]. Finally, in Section 6 we state a number of open questions and outline some possible extensions of our results.

2. Preliminaries

Goodman's Theorem [9] gives a formula for the number of triangles and anti-triangles in a graph $G = (V, E)$ of order n:

$$D_0(G) + D_3(G) = \frac{1}{2}\left[-\binom{n}{3} + \sum_{v\in V}\left[\binom{d_G(v)}{2} + \binom{d_{\overline{G}}(v)}{2}\right]\right]. \tag{1}$$

For densities this translates into

$$p_0(G) + p_3(G) = \frac{\sum_{v\in V}\left[\binom{d_G(v)}{2} + \binom{d_{\overline{G}}(v)}{2}\right]}{2\binom{n}{3}} - \frac{1}{2}.$$

Since $d_G(v) + d_{\overline{G}}(v) = n - 1$ for every v, and due to the convexity of binomial coefficients, the minimal value of $D_0 + D_3$ is achieved whenever the degree of each vertex v is as close to $n/2$ as possible, resulting in $p_0 + p_3$ being asymptotically $1/4$. This is known as *Goodman's bound*. Consequently, let us call $\mathcal{G}$ a *Goodman sequence* if $(p_0+p_3)(\mathcal{G}) = 1/4$; note that we do not require the existence of the individual limits $p_0(\mathcal{G})$ and $p_3(\mathcal{G})$. Needless to say that 3RL sequences are Goodman. Applying a common abuse of terminology, we will talk about Goodman and 3RL *graphs* referring to respective sequences of graphs. Notice that, since $D_H(G) = D_{\overline{H}}(\overline{G})$, a graph G is Goodman (respectively 3RL) if and only if $\overline{G}$ is Goodman (respectively 3RL).

A vertex v of a graph G on n vertices is said to be *ε-ordinary* if $|d_G(v) - n/2| < \varepsilon n$ and *ε-exceptional* otherwise. Occasionally we will suppress the ε in the above notation if there is no ambiguity. The following fact is an immediate consequence of Goodman's Theorem. It asserts that a Goodman graph is essentially $n/2$-regular.

Proposition 2.1. *For every $\varepsilon > 0$ and every Goodman sequence $\mathcal{G} = (G_k)_{k=1}^{\infty}$ there exists an integer $k_0(\varepsilon, \mathcal{G})$ such that for every $k \geq k_0$ at most εn_k vertices of G_k are ε-exceptional.*

Note that, in particular, Goodman graphs are 2RL, thus "Goodman" can be considered an intermediate level between 2RL and 3RL. This was already pointed out by Chung, Graham and Wilson [5] ("$P_1(3) \Rightarrow P_0 \Rightarrow P_1(2)$" in their terminology; Corollary 2.3 below states that $P_1(3)$ and 3RL are equivalent).

Proof. Consider a graph G of order n in which at least εn vertices are ε-exceptional. Due to the convexity of binomial coefficients, each exceptional vertex v contributes to the right hand side of (1) at least

$$\begin{aligned}\binom{d_G(v)}{2} + \binom{d_{\overline{G}}(v)}{2} &\geq \binom{(1/2-\varepsilon)n}{2} + \binom{(1/2+\varepsilon)n}{2} + O(n) \\ &= (1+4\varepsilon^2)\left[\binom{n/2}{2} + \binom{n/2}{2}\right] + O(n)\,.\end{aligned}$$

If this happens εn times, then $(D_0 + D_3)(G)$ exceeds its minimum possible value by at least $c\varepsilon^3 n^3$ for some constant $c > 0$. Therefore $(p_0 + p_3)(G) > 1/4 + c'\varepsilon^3 + o(1)$ for some absolute constant $c' > 0$. This can only happen finitely many times in a Goodman sequence. □

Conversely, it is easy to see that every $\mathcal{G}$ satisfying the above is Goodman. In other words, Proposition 2.1 gives an alternative characterisation of Goodman sequences.

The next lemma and its corollary can be viewed as a strengthening of the 3-universality result from [10] (although, unlike Linial and Morgenstern, we do not optimise the error term ε). It provides additional information about the 3-local profile of Goodman graphs, asserting that it is determined completely by p_0 (and, equivalently, by p_3).

Lemma 2.2. *If $\mathcal{G}$ is Goodman then $(p_1 - 3p_3)(\mathcal{G}) = (p_2 - 3p_0)(\mathcal{G}) = 0$.*

Proof. Counting vertex-edge pairs (v, e) of $G \in \mathcal{G}$, where $v \notin e$ in two different ways, we obtain

$$3D_3(G) + 2D_2(G) + D_1(G) = (n-2)e(G)\,. \tag{2}$$

Similarly, counting such vertex-edge pairs in $\overline{G}$, we obtain

$$3D_0(G) + 2D_1(G) + D_2(G) = 3D_3(\overline{G}) + 2D_2(\overline{G}) + D_1(\overline{G}) = (n-2)e(\overline{G})\,. \tag{3}$$

Since, by Proposition 2.1, $|e(H) - e(\overline{H})| = o(|H|^2)$ holds for every Goodman graph H, we obtain asymptotic equality between the left hand sides of (2) and (3). Passing to densities, this translates into

$$3p_3 + p_2 = 3p_0 + p_1 = \frac{1}{2}(3p_3 + p_2 + p_1 + 3p_0) = \frac{1}{2}\left[(p_0 + p_1 + p_2 + p_3) + 2(p_0 + p_3)\right] = \frac{3}{4}\,,$$

hence

$$p_1 = \frac{3}{4} - 3p_0 = 3p_3\,.$$

Similarly, $p_2 = 3p_0$. □

As an immediate consequence, we determine the 3-local profile of 3RL graphs.

Corollary 2.3. *If $\mathcal{G}$ is 3RL then $p_1(\mathcal{G}) = p_2(\mathcal{G}) = 3/8$.*

In other words, the 3-local profile of a 3RL graph mirrors that of the random graph $\mathcal{G}_{n,1/2}$, justifying our choice of terminology.

The following construction from [10] is known as the *iterated blow-up* (see e.g. [8]) and demonstrates that 3RL graphs need not be 5-universal. Let $G_1 \cong C_5$ be a 5-cycle. Given G_k, construct G_{k+1} as follows. Take a 5-*blow-up* of G_k (that is, replace every vertex v of G_k by 5 new vertices $v_1, \ldots, v_5$ and draw an edge between u_i and v_j if and only if there was an edge between u and v), and add a 5-cycle within each set $v_1, \ldots, v_5$. Alternatively, G_{k+1} can be constructed by taking a 5^k-blow-up of C_5 and adding a copy of G_k on each partition class. It is not hard to check that no G_k contains an induced path on 5 vertices. In order to see that $\mathcal{G}$ is 3RL one can either calculate the densities directly (as in [10]) or observe that $\mathcal{G}$ is a sequence of self-complementary $\lfloor n/2 \rfloor$-regular graphs, which by Goodman's Theorem yields $p_0(\mathcal{G}) = p_3(\mathcal{G}) = 1/8$.

Note that this construction also shows that for every $\ell \geq 5$ and every $r \geq 6$ there exist arbitrarily large 3RL graphs which do not contain the path P_ℓ of length $\ell - 1$ or the cycle C_r as induced subgraphs. This is because any graph which contains an induced P_ℓ for some $\ell \geq 5$ or an induced C_r for some $r \geq 6$ contains an induced P_5. The case of the 5-cycle remains open.

3. Proof of Theorem 1.3

We have to show that a sufficiently large 3RL graph G contains each graph of order 4 as an induced subgraph. Note that, in contrast to Corollary 2.3, we cannot expect the density of H in G to be random-like for every graph H on 4 vertices. Indeed, it is well-known (see e.g. Theorem 9.3.1 in [2]) that such graphs are quasirandom and thus, in particular, ℓ-universal for any fixed ℓ.

Since G is 3RL if and only if $\overline{G}$ is, and the induced subgraphs of the latter are precisely the complements of induced subgraphs of the former, it suffices to split all 4-vertex graphs into complementary pairs (the graph P_4, the path of length three, is self complementary) and prove containment for one graph H from each pair. Thus we need only consider the following 6 cases:

- $H = K_4$, the complete 4-vertex graph
- $H = K_4^-$, the complete graph with one edge missing
- $H = C_4$, the 4-cycle
- $H = T^+$, a triangle with a pendant edge
- $H = K_{1,3}$, the star (also known as the claw)
- $H = P_4$, the path of length 3

While the graphs above are listed in order of decreasing number of edges, we will consider them in a different order, starting from what we believe is the simplest case and finishing with the most difficult. In each of the cases the containment of H is proved by contradiction, assuming initially that $\mathcal{G}$ is 3RL and H-free (remember that we are always looking for an *induced* copy of H).

Case 1: $H = T^+$**.** It follows by Proposition 2.1 that, for every $\varepsilon > 0$ and sufficiently large n, if $G \in \mathcal{G}$ is a graph on n vertices, then it contains at most εn exceptional vertices. The set of all exceptional vertices of G can intersect at most εn^3 triangles. Since G is 3RL, it contains $(1/48 + o(1))n^3$ triangles and so, for sufficiently large n, there must exist ε-ordinary vertices u, v and w which form a triangle T in G.

Since G is T^+-free, for any $x \in V(G) \setminus \{u, v, w\}$ we must have $d_G(x, T) \in \{0, 2, 3\}$. We partition the vertices of $V(G) \setminus \{u, v, w\}$ into two sets $X = \{x \in V(G) \setminus \{u, v, w\} : d_G(x, T) = 0\}$ and $Y = \{x \in V(G) \setminus \{u, v, w\} : d_G(x, T) \in \{2, 3\}\}$. Since u, v and w are ordinary, on average, a vertex $x \in V(G) \setminus \{u, v, w\}$ will have $3/2 + o(1)$ neighbours in T. Therefore, we must have $n/4 - o(n) \leq |X| \leq n/2 + o(n)$ and $n/2 - o(n) \leq |Y| \leq 3n/4 + o(n)$. Let $x \in X$ and $y \in Y$ be arbitrary vertices. Assume without loss of generality that $\{u, v\} \subseteq N_G(y, T)$. We conclude that x and y are not adjacent in G as otherwise the vertices x, y, u and v would form an induced copy of T^+ in G.

Since x and y were arbitrary, it follows that there are no edges of G between X and Y. Since $|X| \geq n/4 - o(n)$ and at most εn vertices of G are exceptional, there exists some ordinary $x \in X$. Because of $N_G(x) \subseteq X$, it follows that $|X| = n/2 + o(n)$. Finally, since all but at most εn vertices of X are ordinary and each of them has degree $n/2 + o(n)$ in X, we conclude that $e(G[X]) \geq \binom{n/2}{2} - o(n^2)$.

A similar argument shows that $|Y| = n/2 + o(n)$ and that $e(G[Y]) \geq \binom{n/2}{2} - o(n^2)$. Counting anti-triangles in G, it follows that $D_0(G) = o(n^3)$ and thus $p_0(G) = 0$, contrary to our assumption that G is 3RL.

Case 2: $H = K_4^-$**.** As in Case 1, consider a triangle $T = \{u, v, w\}$ where u, v and w are ordinary vertices. Since G is K_4^--free, for any $x \in V(G) \setminus \{u, v, w\}$ we must have $d_G(x, T) \in \{0, 1, 3\}$. We partition the vertices of $V(G)$ into two sets $X = \{x \in V(G) \setminus \{u, v, w\} : d_G(x, T) \in \{0, 1\}\}$

and $Y = V(G) \setminus X$. Since u, v and w are ordinary, on average a vertex $x \in V(G) \setminus \{u, v, w\}$ will have $3/2 + o(1)$ neighbours in T. Therefore, we must have $n/2 - o(n) \leq |X| \leq 3n/4 + o(n)$ and $n/4 - o(n) \leq |Y| \leq n/2 + o(n)$. Considering u, v and arbitrary $x, y \in Y \setminus \{u, v\}$, we deduce that x and y are adjacent in G, for otherwise u ,v, x and y would form an induced copy of K_4^- in G. It follows that $G[Y]$ is a clique.

Let $z \in X$ be an arbitrary vertex and assume without loss of generality that $\{z, u\} \notin E(G)$. Let $y_1, y_2 \in Y \setminus \{u\}$ be arbitrary vertices. If $\{z, y_1\} \in E(G)$ and $\{z, y_2\} \in E(G)$, then $G[\{z, y_1, y_2, u\}]$ is an induced copy of K_4^- in G. It follows that $d_G(x, Y) \leq 1$ for every $x \in X$. By Proposition 2.1 G is essentially $n/2$-regular and thus we must have $n/2 - o(n) \leq |X|, |Y| \leq n/2 + o(n)$ and $e(G[X]) \geq \binom{n/2}{2} - o(n^2)$. Similarly to Case 1, it follows that $p_0(G) = 0$, contrary to our assumption that G is 3RL.

Case 3: $H = K_4$**.** Let $\mathcal{G}'$ be any Goodman (not necessarily 3RL) sequence of K_4-free graphs. Observe that G' being K_4-free is equivalent to $N_{G'}(v)$ being triangle-free for every $v \in V(G')$. Therefore, by Mantel's Theorem, $e(G'[N(v)]) \leq d(v)^2/4$ for every $v \in V(G')$. Since $\mathcal{G}'$ is Goodman, it follows by Proposition 2.1 that the neighbourhoods of all but $o(n)$ vertices of $G' \in \mathcal{G}'$, span at most $(n/2 + o(n))^2/4 = n^2/16 + o(n^2)$ edges. Since $e(G'[N_{G'}(v)])$ is precisely the number of triangles of G' that include v, a double counting of the edges in all neighbourhoods shows that

$$3D_3 = \sum_{v \in V(G')} e(G'[N(v)]) \leq \sum_{v \in V(G')} d(v)^2/4 = n \cdot \frac{n^2}{16} + o(n^3) = \left(\frac{3}{8} + o(1)\right)\binom{n}{3}.$$

Hence, for any K_4-free Goodman sequence $\mathcal{G}'$ the value $p_3(\mathcal{G}')$, if it exists, is at most 1/8, with equality attained only when $e(G'[N(v)]) = (1 - o(1))d(v)^2/4$ holds for all but $o(n)$ vertices $v \in V(G')$. Conversely, $\mathcal{G}$ being 3RL implies that equality must be attained, whence we conclude that $e(G[N(v)]) = (1 - o(1))d(v)^2/4$ holds for all but $o(n)$ vertices $v \in V(G)$.

Structural information on such 'nearly extremal' graphs is provided by the Erdős-Simonovits stability Theorem [13] which, in this particular case and combined with the above, asserts that there exists a set $U \subseteq V(G)$ of order $(1 - o(1))n$ such that for every $v \in U$ we have $d_G(v) = n/2 + o(n)$ and the neighbourhood $N(v)$ admits a bipartition into parts $N_1(v)$ and $N_2(v)$ such that $|N_1(v)|, |N_2(v)| = n/4 + o(n)$, there are $o(n^2)$ edges within each partition class and $(1 - o(1))n^2/16$ edges between the two classes.

Let $v \in U$ be an arbitrary vertex. It follows by the above that there exists a vertex $u \in U \cap N_1(v)$ such that $d_G(u, N_2(v)) = n/4 + o(n)$ and $d_G(u, N_1(v)) = o(n)$ (recall that almost every vertex has these properties). Let $B = N_G(u) \setminus N_G(v)$; note that $|B| = n/4 + o(n)$. Since $u \in U$, its neighbourhood $N_G(u)$ must induce an essentially complete bipartite graph with both parts of order $n/4 + o(n)$. Since $N_2(u) := N_2(v) \cap N_G(u)$ is of order $n/4 + o(n)$ and contains $o(n^2)$ edges, up to $o(n)$ changes, the only way to achieve this is by taking the bipartition to be $N_G(u) = B \cup N_2(u)$. Let $w \in U \cap N_2(u)$ be a vertex such that $d_G(w, N_1(v)) = n/4 + o(n)$ and $d_G(w, B) = n/4 + o(n)$; by the above, almost every vertex of $U \cap N_2(u)$ has these properties. Since $w \in U$, its neighbourhood $N_G(w)$ must induce an essentially complete bipartite graph with both parts of order $n/4 + o(n)$. Up to $o(n)$ changes, the only way to achieve this is by taking the bipartition to be $N_G(w) = B \cup N_1(v)$. We conclude that the sets $N_1(v)$, $N_2(u)$ and B are of size $n/4 + o(n)$ each and $G[N_1(v) \cup N_2(u) \cup B]$ is essentially an $n/2$-regular tripartite graph.

Let $X = N_1(v) \cup N_2(u) \cup B$ and let $Y = U \setminus X$. On the one hand, $|Y| = n/4 + o(n)$ and the degree of every vertex in Y is $n/2 + o(n)$ entailing that there are $\Omega(n^2)$ edges between X and Y. On the other hand, all but $o(n)$ vertices of X have degree $n/2 + o(n)$ in G and in $G[X]$ entailing that there are $o(n^2)$ edges between X and Y. This is clearly a contradiction.

Case 4: $H = K_{1,3}$**.** Let $\mathcal{G}'$ be a Goodman sequence of $K_{1,3}$-free graphs. Note that G' being $K_{1,3}$-free is equivalent to $N_{G'}(v)$ being anti-triangle-free for every $v \in V(G')$, that is, the non-edges in $N_{G'}(v)$ must not form a triangle. Similarly to Case 3, it follows from Mantel's Theorem and Proposition 2.1 that the neighbourhoods of all but $o(n)$ vertices of G', span at most $(n/2 +$

$o(n))^2/4 = n^2/16 + o(n^2)$ non-edges. Double counting of the non-edges in all neighbourhoods shows that

$$D_2 = \sum_{v \in V(G')} e(\overline{G'}[N_{G'}(v)]) = n \cdot \left(n^2/16 + o(n^2)\right) = \left(\frac{3}{8} + o(1)\right)\binom{n}{3}.$$

Hence, for any Goodman sequence $\mathcal{G}'$, the value $p_2(\mathcal{G}')$, if it exists, is at most $3/8$, where by the Erdős-Simonovits stability Theorem, equality is attained only when almost all neighbourhoods are close to being disjoint unions of two complete graphs of order $n/4$ each. Since G is 3RL, it follows by Corollary 2.3 that this must indeed be the case.

Let $U \subseteq V(G)$ be a set of order $(1 - o(1))n$ such that for every $v \in U$ we have $d_G(v) = n/2 + o(n)$ and the neighbourhood $N(v)$ admits a bipartition into parts $N_1(v)$ and $N_2(v)$ such that $|N_1(v)|, |N_2(v)| = n/4 + o(n)$, there are $(1 - o(1))n^2/32$ edges within each partition class and $o(n^2)$ edges between the two classes.

Let $v \in U$ be an arbitrary vertex. It follows by the above that there exists a vertex $u \in A(u) := U \cap N_1(v)$ such that $|A(u)| = n/4 + o(n)$, $d_G(u, A(u)) = n/4 + o(n)$ and $d_G(u, N_2(v)) = o(n)$. Let $B = N_G(u) \setminus N_G(v)$; note that $|B| = n/4 + o(n)$. Since $u \in U$, its neighbourhood $N_G(u)$ must be close to a union of two complete graphs of order $n/4$ each. Since $G[A(u)]$ is essentially a complete graph on $n/4 + o(n)$ vertices, it follows that $G[B]$ is essentially a complete graph on $n/4 + o(n)$ vertices as well. Moreover, there are $o(n^2)$ edges of G between $A(u)$ and $N_2(v) \cup B$.

Let $X = U \setminus (A(u) \cup N_2(v) \cup B)$; note that $|X| = n/4 + o(n)$. Since $d_G(w) = n/2 + o(n)$ holds for every $w \in A(u)$, it follows that $\{x, y\} \in E(G)$ for all but $o(n^2)$ pairs $(x, y) \in A(u) \times X$. Let $z \in A(u) \setminus \{u\}$ be an arbitrary vertex. Up to $o(n)$ vertices, its neighbourhood is $A(u) \cup X$ and so is far from being the disjoint union of two cliques of order $n/4$ each, contrary to the definition of U.

Case 5: $H = P_4$**.** For graphs with no induced P_4, also known as *cographs*, we have the following structural characterisation due to Seinsche [12]: if G is induced P_4-free then either G or $\overline{G}$ is disconnected (the other one is thereby forced to be connected). Let $W \subseteq V(G)$ be an arbitrary set of order at least 2. Clearly $G[W]$ is induced P_4-free and so, by the above characterisation, either $G[W]$ or $\overline{G}[W]$ is disconnected (note that it might be $G[W]$ for certain $W \subseteq V(G)$ and $\overline{G}[W]$ for others).

Seinsche's characterisation allows us to construct a sequence $\mathcal{P}_0, \mathcal{P}_1, \ldots$ of partitions of $V(G)$ as follows. $\mathcal{P}_0 = \{V(G)\}$ and, for every $i \geq 0$, $\mathcal{P}_{i+1}$ is obtained through partitioning each $W \in \mathcal{P}_i$ with $|W| \geq 2$ into the connected components of either $G[W]$ or $\overline{G}[W]$, depending which of the two is disconnected. For every $i \geq 0$, let V_i denote a largest set in $\mathcal{P}_i$. For an arbitrarily small $\varepsilon > 0$ and sufficiently large n let $j \geq 0$ denote the smallest index for which $|V_{j+1}| < (1 - \varepsilon)n$; clearly such an index j must exist. Since G is Goodman, it follows by Proposition 2.1 that at most εn vertices of G are ε-exceptional. Since $|V_{j+1}| < (1 - \varepsilon)n$, the only way to ensure that there will not be too many exceptional vertices in G is to split V_j in $\mathcal{P}_{j+1}$ into two sets W_1 and W_2 of size $n/2 - 2\varepsilon n \leq |W_1|, |W_2| \leq n/2 + 2\varepsilon n$, and possibly some additional small sets. Indeed, otherwise every vertex of $V(G) \setminus V_{j+1}$ would be exceptional. Assume first that $G[V_j]$ is disconnected. Then there are at most $3\varepsilon n^2$ pairs $\{x, y\} \subseteq W_1$ and at most $3\varepsilon n^2$ pairs $\{x, y\} \subseteq W_2$ which are not adjacent in G. It follows that $D_0(G) \leq c\varepsilon n^3$ for some absolute constant c and thus $p_0(G) = 0$. Similarly, if $\overline{G}[V_j]$ is disconnected then $p_3(G) = 0$. This contradicts our assumption that G is 3RL.

Case 6: $H = C_4$**.** Consider the expression

$$\sum_{\{u,v\} \in E(\overline{G})} \binom{d_G(u, v)}{2}.$$

On the one hand, it counts $D_{K_4^-}(G) + 2D_{C_4}(G)$, and since, by assumption, $D_{C_4} = 0$, it must equal $D_{K_4^-}$. Now, since by Corollary 2.3

$$\sum_{\{u,v\}\in E(\overline{G})} d_G(u,v) = D_2(G) = \frac{3}{8}\binom{n}{3} + o(n^3) = \frac{n^3}{16} + o(n^3)\,,$$

using the convexity of binomial coefficients we obtain

$$\begin{aligned} D_{K_4^-} &= \sum_{\{u,v\}\in E(\overline{G})} \binom{d_G(u,v)}{2} \geq e(\overline{G})\binom{\frac{1}{e(\overline{G})}\sum_{\{u,v\}\in E(\overline{G})} d_G(u,v)}{2} \\ &= \frac{n^2}{4}\binom{n/4}{2} + o(n^4) = \frac{n^4}{128} + o(n^4)\,. \end{aligned} \tag{4}$$

Thus

$$p_{K_4^-}(G) \geq 3/16\,. \tag{5}$$

Now consider the expression

$$\sum_{\{u,v\}\in E(G)} d_G(u,-v)\cdot d_G(v,-u)\,.$$

On the one hand, it counts $D_{P_4} + 4D_{C_4}$, which by assumption equals D_{P_4}. On the other hand, since G is Goodman, by Proposition 2.1 we have $d(u) = d(v) + o(n)$ for all but at most $o(n^2)$ pairs of vertices, hence, $d(u,-v) = d(v,-u) + o(n)$ for almost all pairs. As a result, we obtain

$$\sum_{\{u,v\}\in E(G)} d(u,-v)\cdot d(v,-u) = \frac{1}{4}\sum_{\{u,v\}\in E(G)} (d(u,-v)+d(v,-u))^2 + o(n^4)\,.$$

Now, since

$$\sum_{\{u,v\}\in E(G)} (d(u,-v)+d(v,-u)) = 2D_2(G) = \frac{3}{4}\binom{n}{3} + o(n^3) = \frac{n^3}{8} + o(n^3)\,,$$

the Cauchy-Schwarz inequality yields

$$\begin{aligned} D_{P_4} &= \frac{1}{4}\sum_{\{u,v\}\in E(G)} (d(u,-v)+d(v,-u))^2 + o(n^4) \\ &\geq \frac{1}{4}\cdot e(G)\left[\frac{1}{e(G)}\sum_{\{u,v\}\in E(G)} (d(u,-v)+d(v,-u))\right]^2 + o(n^4) \\ &= \frac{1}{4}\cdot\frac{n^2}{4}\cdot\left(\frac{n}{2}\right)^2 + o(n^4) \\ &= \frac{n^4}{64} + o(n^4)\,. \end{aligned} \tag{6}$$

Thus

$$p_{P_4}(G) \geq 3/8\,. \tag{7}$$

Finally, double counting pairs of edges in G not sharing a vertex, we obtain

$$\frac{e(G)^2}{2} + o(n^4) = \binom{n^2/4}{2} + o(n^4) = D_{2K_2} + D_{P_4} + D_{T^+} + 2D_{C_4} + 2D_{K_4^-} + 3D_{K_4}\,,$$

where $2K_2$ is the complement of C_4. Thus

$$p_{2K_2} + p_{P_4} + p_{T^+} + 2p_{K_4^-} + 3p_{K_4} = \frac{3}{4}\,.$$

Since from (5) and (7) we know that $2p_{K_4^-}(G) + p_{P_4}(G) \geq 3/4$, we deduce that $p_{T^+}(G) = p_{K_4}(G) = 0$, $p_{K_4^-}(G) = 3/16$, and $p_{P_4}(G) = 3/8$. The last two identities can only hold if in (4) and (6) we have equality up to $o(n^4)$. The former would imply that $d(u,v)$ must be close to its average $n/4+o(n)$ for all but $o(n^2)$ pairs $\{u,v\} \in E(\overline{G})$. Similarly, an equality up to $o(n^4)$ in (6) implies that $d(u,-v) = d(v,-u)+o(n) = n/4+o(n)$ for almost every pair $\{u,v\} \in E(G)$, which, given Proposition 2.1, also means that $d(u,v) = n/4 + o(n)$ for every such pair. In total, we obtain that all but $o(n^2)$ pairs of vertices $\{u,v\}$ have $n/4+o(n)$ joint neighbours, and therefore

$$\sum_{u,v\in V} \left| d(u,v) - \frac{n}{4} \right| = o(n^3)\,. \tag{8}$$

However, equation (8) is one of several equivalent definitions of a quasirandom graph (see e.g. Theorem 9.3.1 in [2]) and thus all induced densities in G are random-like. In particular, contrary to our assumption, G cannot be induced C_4-free.

With contradiction obtained for each 4-vertex graph H, the proof of Theorem 1.3 is concluded.

4. Large induced subgraphs

Our aim in this section is to prove Theorem 1.4. The following probabilistic construction, communicated to us by Paul Balister [3], gives a 3RL sequence $\mathcal{H}^1 = (H^1_n)_{n=1}^\infty$ such that H^1_n is K_5-free for all n. Let $V(H^1_n)$ consist of 8 pairwise disjoint sets of size n each, which we shall call $V_1, \dots V_4$ and $W_1, \dots, W_4$. For every $1 \le i < j \le 4$ we place a complete bipartite graph between V_i and V_j and between W_i and W_j. Moreover, for every $1 \le i \ne j \le 4$, we place an edge between $v \in V_i$ and $w \in W_j$ randomly with probability $1/3$, where all choices are mutually independent. Note that H^1_n contains no edges between V_i and W_i.

It is easy to check that, even if all randomly selected edges are present, H^1_n is K_5-free. Furthermore, a straightforward calculation shows that the expected 3-local profile of $\mathcal{H}^1$ is $(1/8, 3/8, 3/8, 1/8)$. Since the number of randomly selected edges is binomially distributed, standard bounds on the tail probabilities yield that $\mathcal{H}^1$ is a.a.s. 3RL.

Before we proceed with the proof of Theorem 1.4, we shall give a second construction of a 3RL sequence $\mathcal{H}^2$ with no cliques of order at least 10. Despite being "only" K_{10}-free rather than K_5-free, $\mathcal{H}^2$ has the advantage of being constructed deterministically. Moreover, we believe this construction to be interesting in its own right.

Given any graph G of order n, we construct an $(n-1)$-regular graph $H = f(G)$ of order $2n$ as follows. Let $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ be two disjoint copies of $G = (V, E)$, where $V = \{u_1, \dots, u_n\}$, $V_1 = \{v_1, \dots, v_n\}$ and $V_2 = \{v'_1, \dots, v'_n\}$. Set $V(H) = V_1 \cup V_2$ and $E(H) = E_1 \cup E_2 \cup F$, where $F = \Big\{\{v_i, v'_j\} \colon 1 \le i \ne j \le n \text{ and } \{u_i, u_j\} \notin E\Big\}$. That is, we take two identical copies of G and connect two distinct vertices, one from each copy, by an edge of H if and only if they are *not* adjacent in G.

The above construction is very similar to the *tensor product* of G and K_2; the sole difference is that we exclude the "vertical" edges, that is, edges between v_i and v'_i. Tensor products of graphs were first defined by Thomason in [15]. See Section 6 for more details.

Note that for any sequence $\mathcal{G} = (G_k)_{k=1}^\infty$, the corresponding sequence $f(\mathcal{G}) = (f(G_k))_{k=1}^\infty$ is automatically Goodman. The next question to ask is, under what conditions is $f(\mathcal{G})$ 3RL.

Lemma 4.1. *$\mathcal{H}^2 = f(\mathcal{G})$ is 3RL if and only if $(p_0 + p_2)(\mathcal{G}) = (p_1 + p_3)(\mathcal{G}) = 1/2$.*

Proof. For every $0 \le i \le 3$, let T_i denote the number of triangles of H with exactly i vertices in V_1. It is evident that $D_3(H) = T_0 + T_1 + T_2 + T_3 = 2(T_0 + T_1)$. Clearly $T_0 = D_3(G)$. Moreover, every triangle with exactly one vertex in V_1 corresponds to three vertices which induce precisely one edge in G and thus $T_1 = D_1(G)$. It follows that $D_3(H) = 2(D_3(G) + D_1(G))$ and thus

$$p_3(H) = \frac{1}{4}\left[p_1(G) + p_3(G)\right] + o(1)\,.$$

We conclude that $p_3(H) = 1/8$ if and only if $(p_1 + p_3)(G) = 1/2$. An analogous argument shows that $p_0(H) = 1/8$ if and only if $(p_0 + p_2)(G) = 1/2$. □

Given Lemma 4.1, our aim is to construct a sequence $\mathcal{G}$ with $p_1(\mathcal{G}) + p_3(\mathcal{G}) = 1/2$ such that $f(\mathcal{G})$ does not contain a clique of some fixed size. Given a positive integer k and a real number $r \geq 2$ that might depend on k, let $G_k^r = (V_k, E_k)$, where $V_k = \{0, 1, \ldots, k-1\}$ and $E_k = \{\{i, j\} : i - j \mod k > k/r \text{ and } j - i \mod k > k/r\}$; let $\mathcal{G} = (G_k^r)_{k=1}^{\infty}$. Since any $\lceil r \rceil$ vertices of V_k must contain two whose distance in the cyclic group C_k is at most k/r, the largest clique of G_k^r is of order at most $\lceil r \rceil - 1$. We claim that a similar statement holds in H.

Claim 4.2. *For every $r \geq 5$ and sufficiently large k, the largest clique in $H = f(G_k^r)$ is of order at most $\lceil r \rceil - 1$.*

Proof. For the sake of clarity of presentation let us assume that r is an integer. Suppose for a contradiction that ϕ is an embedding of K_r in H. Let J denote the resulting copy; let $U_1 = V_1 \cap V(J)$ and $U_2 = V_2 \cap V(J)$. Since, as observed above, G_k^r does not contain K_r as a subgraph, it follows that $U_1 \neq \emptyset$ and $U_2 \neq \emptyset$. Let $t \geq s \geq 1$ be integers such that $s = |U_1|$ and $t = |U_2|$; note that $t \geq \lceil r/2 \rceil \geq 3$. Since every $u \in U_1$ and every $v \in U_2$ are joined by an edge of H and yet $\{v_i, v_i'\} \notin E(H)$ for every $1 \leq i \leq n$, it follows that $\overline{G_k^r}$ contains the complete bipartite graph $K_{s,t}$ as an induced subgraph. Let $u \in U_1$ and let w_1, w_2 and w_3 be three of its neighbours in U_2. These four vertices correspond to four vertices i and $j_1 < j_2 < j_3$ of V_k such that j_1, j_2 and j_3 are pairwise far in C_k but i is close to all of them. This is clearly impossible. □

It remains to find a value of r for which $(p_1 + p_3)f(\mathcal{G}) = 1/2$, where $\mathcal{G} = (G_k^r)_{k=1}^{\infty}$. Observe that for $r > k$ the graph G_k^r is a complete graph, whereas for $r = 2$ the graph G_k^r is empty. Hence there must exist a real number r for which $(p_1 + p_3)(G_k^r) = 1/2 + o(1)$. A straightforward calculation shows that $p_3(G_k^r) = \left(\frac{r-3}{r}\right)^2 + o(1)$ and $p_1(G_k^r) = \frac{3}{r^2} + o(1)$, whence the desired value of r is achieved at $2\sqrt{3} + 6 \approx 9.46$. For this value of r the sequence $\mathcal{H}^2 = f(\mathcal{G})$ is 3RL but does not contain a clique of order 10.

Theorem 1.4 is a simple consequence of the above constructions. Before showing this, let us remark that, for every $r \geq 10$, there exists a 3RL graph with no induced copy of $K_{1,r}$. This is simply because K_r is an induced subgraph of $\overline{K_{1,r}}$, so $\overline{\mathcal{H}^2}$, the sequence of complements of the graphs $f(G_k^r)$ does not contain any star of order 11 or greater as an induced subgraph.

Now let J be any graph of order $n \geq R(5,5)$, where $R(k, \ell)$ stands, as usual, for the corresponding Ramsey number (see [4] for more background details). By Ramsey's Theorem J must contain a K_5 or $\overline{K_5}$ as an induced subgraph. In the former case put $\mathcal{G} = \mathcal{H}^1$, and in the latter put $\mathcal{G} = \overline{\mathcal{H}^1}$. In either case we have found a sequence which is 3RL but without J as an induced subgraph. According to the best currently known bounds [11], we have $43 \leq R(5,5) \leq 49$, which proves Theorem 1.4 with a constant of 49 in place of 13.

This bound can be improved by considering the following construction, which was also communicated to us by Paul Balister [3]. Note that, by the above, it suffices to only consider graphs J containing no clique or independent set on 5 vertices.

Consider the following sequence $\mathcal{H}^3 = (H_n^3)_{n=1}^{\infty}$. Let $V(H_n^3) = V_1 \cup V_2 \cup V_3 \cup V_4$ be the disjoint union of 4 sets of size n each. Place complete graphs on V_1 and V_4 and empty graphs on V_2 and V_3. Moreover, place complete bipartite graphs on (V_1, V_2), (V_2, V_3) and (V_3, V_4). It is easily checked that $\mathcal{H}^3$ is Goodman and self-complementary, and is therefore also 3RL. Furthermore, it is easy to check that every induced subgraph of H_n^3 of order at least 13 must contain either a clique or an independent set on 5 vertices. Hence, for an arbitrary graph J of order at least 13 with $\alpha(J), \omega(J) \leq 4$ we can put $\mathcal{G} = \mathcal{H}^3$, obtaining a 3RL sequence without J as an induced subgraph.

This concludes the proof of Theorem 1.4. It is conceivable that the bound of 13 can be reduced even further.

5. A construction of high universality

In this section we prove Theorem 1.6, to which end we shall use the following construction.

Given vertex disjoint graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, where $|V_1| = |V_2|$, we construct a graph $H = G_1 \oplus G_2$ by joining G_1 and G_2 via a random bipartite graph. Formally, $V(H) =$

$V_1 \cup V_2$ and $E(H) = E_1 \cup E_2 \cup E_3$, where E_3 is formed by joining independently at random each pair $(x, y) \in V_1 \times V_2$ with probability 1/2.

The special case of this construction in which $G_1 = K_{n,n}$ and $G_2 = \overline{K_{n,n}}$ was used by Chung, Graham and Wilson [5] in order to demonstrate that a graph H that behaves random-like with respect to all 3-vertex subgraphs, that is, when $(p_0(H), p_1(H), p_2(H), p_3(H)) = (1/8, 3/8, 3/8, 1/8)$, need not be quasirandom. Note that, due to Corollary 2.3, being random-like with respect to all 3-vertex subgraphs is equivalent to being 3RL. The following two lemmas provide more information on the 3-local profile of $G_1 \oplus G_2$.

Lemma 5.1. *$H = G_1 \oplus G_2$ is a.a.s. Goodman if and only if G_1 and G_2 are Goodman.*

Proof. For every $0 \le i \le 3$ and sufficiently large n, the probabilities of a random vertex-triple of $V(H)$ having exactly i vertices in G_1 are roughly binomially distributed. It thus follows by the definition of $G_1 \oplus G_2$ and by standard bounds on the tails of the binomial distribution that a.a.s.

$$p_3(H) = \frac{1}{8}\left(p_3(G_1) + p_3(G_2)\right) + \frac{3}{8} \cdot \left(\frac{1}{2}\right)^2 \cdot \left(p_e(G_1) + p_e(G_2)\right) , \tag{9}$$

and similarly a.a.s.

$$p_0(H) = \frac{1}{8}\left(p_0(G_1) + p_0(G_2)\right) + \frac{3}{8} \cdot \left(\frac{1}{2}\right)^2 \cdot \left(p_e(\overline{G_1}) + p_e(\overline{G_2})\right) . \tag{10}$$

Since $p_e(G) + p_e(\overline{G}) = 1$, adding the above equations we obtain

$$p_0(H) + p_3(H) = \frac{1}{8}\left[(p_0(G_1) + p_3(G_1)) + (p_0(G_2) + p_3(G_2))\right] + \frac{3}{16} ,$$

which equals 1/4 if and only if

$$(p_0(G_1) + p_3(G_1)) + (p_0(G_2) + p_3(G_2)) = \frac{1}{2} . \tag{11}$$

By Goodman's bound, (11) can only hold when $p_0(G_1) + p_3(G_1) = p_0(G_2) + p_3(G_2) = \frac{1}{4}$, that is, when both G_1 and G_2 are Goodman. □

Lemma 5.2. *$H = G_1 \oplus G_2$ is a.a.s. 3RL if and only if G_1 and G_2 are Goodman and $p_i(G_1) = p_{3-i}(G_2)$ for all $0 \le i \le 3$.*

Proof. If G_1 and G_2 satisfy the conditions of the lemma, then $p_e(G_1) = p_e(G_2) = 1/2$ and $p_3(G_1) + p_3(G_2) = p_3(G_1) + p_0(G_1) = 1/4$. It follows from (9) that $p_3(H) = 1/8$. Similarly, the conditions of the lemma and (10) yield $p_0(H) = 1/8$, whence we conclude that H is 3RL.

Conversely, if H is 3RL, it is Goodman, so by Lemma 5.1 G_1 and G_2 must be Goodman as well, in which case $p_e(G_1) = p_e(G_2) = 1/2$, and the identities (9) and (10) transform into

$$\frac{1}{8} = p_3(H) = \frac{1}{8}\left(p_3(G_1) + p_3(G_2)\right) + \frac{3}{32}$$

and

$$\frac{1}{8} = p_0(H) = \frac{1}{8}\left(p_0(G_1) + p_0(G_2)\right) + \frac{3}{32} .$$

It follows that

$$p_0(G_1) + p_0(G_2) = p_3(G_1) + p_3(G_2) = \frac{1}{4} = p_0(G_1) + p_3(G_1) = p_0(G_2) + p_3(G_2).$$

Thus $p_0(G_1) = p_3(G_2)$ and $p_3(G_1) = p_0(G_2)$. By Lemma 2.2 we also have $p_1(G_1) = p_2(G_2)$ and $p_2(G_1) = p_1(G_2)$. □

Since two 3RL graphs satisfy the conditions of Lemma 5.2, we obtain the following useful fact as an immediate corollary.

Corollary 5.3. *If G_1 and G_2 are 3RL then $G_1 \oplus G_2$ is a.a.s. 3RL.*

Lemma 5.2 and Corollary 5.3 allow us to iterate the construction $G_1 \oplus G_2$, taking the aforementioned example of Chung, Graham and Wilson as our starting point. Define $G_1 := K_{n,n} \oplus \overline{K_{n,n}}$ and having constructed G_ℓ, define $G_{\ell+1} := G_\ell \oplus G_\ell$. Let $\mathcal{G}_\ell = (G_\ell)_{n=1}^{\infty}$, that is, we fix the ℓ's iteration and let n go to infinity. By Lemma 5.2 the sequence $\mathcal{G}_1$ is a.a.s. 3RL and by Corollary 5.3 it follows inductively that for each $\ell > 1$ the sequence $\mathcal{G}_\ell$ is a.a.s. 3RL.

Each graph $G_\ell \in \mathcal{G}_\ell$ consists of 2^ℓ "deterministic" components, each of which is either a copy of $K_{n,n}$ or its complement. The edges connecting vertices from different components are picked independently at random with probability $1/2$ each. Now, if we select 2^ℓ vertices from G_ℓ uniformly at random, the probability of choosing precisely one vertex from each deterministic component is a positive function of ℓ. Since the obtained graph contains only randomly picked edges, the expected proportion of induced copies of *any* graph H of order 2^ℓ is also a positive function of ℓ. Hence, for a fixed ℓ and n tending to infinity we will have $p_H(\mathcal{G}_\ell) > 0$ for each such graph G.

On the other hand, any subgraph of G_ℓ of order $m = 24\ell \cdot 2^\ell$, contains by the pigeonhole principle either a clique or an independent set of size 12ℓ. So assuming that H_ℓ is m-universal, every graph of order m must contain such a set. However, the well-known lower bound on Ramsey numbers (see e.g. [4]) states that there exist graphs on $2^{12\ell/2} > m$ vertices without a clique or an independent set of size 12ℓ, a contradiction. Thus we conclude that $\mathcal{G}_\ell$ is not m-universal, thereby completing the proof of Theorem 1.6.

Remark. The standard proof of the bound $R(k,k) > 2^{k/2}$ has stronger consequences. Namely, it shows that for $n = 2^{k/2}$ with high probability a random graph on n vertices does not contain an induced copy of K_k or $\overline{K_k}$. Applying this fact to the proof of Theorem 1.6 shows that the proportion of graphs of order m contained in $\mathcal{H}_\ell$ vanishes as ℓ tends to infinity. In other words, not only is $\mathcal{H}_\ell$ not m-universal, it actually contains "very few" different induced subgraphs of order m.

6. Discussion

There are many intriguing open problems regarding random-like sequences and universality. Several of them, including some problems on universality of tournaments, can be found in [10].

It would be very interesting to generalise our results to m-random-like sequences, that is, to study universalities of sequences whose densities of m-cliques and m-anticliques is $2^{-\binom{m}{2}}$. However, this seems to be a much more difficult task, since for $m > 3$ we no longer have the analogue of Goodman's Theorem. In fact, in our proof of Theorem 1.3 we made use of the "lucky coincidence" that the random-like number of triangles and anti-triangles is also the minimal possible. Disproving a conjecture of Erdős [6], it was shown by Thomason [15] that this is no longer true for $m \geq 4$. It would therefore also be interesting to investigate universalities of graphs whose densities of cliques and anti-cliques are the *smallest possible* rather than random-like.

Since 3RL is not enough to ensure 5-universality, one might ask which stronger random-like properties suffice. We propose the following question.

Question 6.1. *Is it true that every sequence $\mathcal{G}$ that is m-random-like for every $m \leq M$ is M-universal?*

By [10] and Theorem 1.3 the answer to Question 6.1 is affirmative for $M \leq 4$, so $M = 5$ is the first open case. Note that the iterated blow-up construction used to prove that 3RL sequences are not necessarily 5-universal is not 4RL and thus does not provide a negative answer to Question 6.1.

Our proof of Theorem 1.3 established the existence of every possible 4-vertex induced subgraph H in a 3RL sequence $\mathcal{G}$. As noted in the Introduction, it follows from the induced graph removal lemma that in fact the corresponding density $p_H(\mathcal{G})$ must be bounded away from 0. It would be interesting to determine, for every 4-vertex graph H, the minimum density $p_H(\mathcal{G})$ over all 3RL sequences $\mathcal{G}$.

Acknowledgements

We thank Paul Balister for helpful comments which led to improvements in Section 4.

References

[1] N. Alon, E. Fischer, M. Krivelevich, and M. Szegedy. Efficient testing of large graphs. *Combinatorica*, 20(4):451–476, 2000.
[2] N. Alon and J.H. Spencer. *The Probabilistic Method, 3rd Ed.* Wiley, New York, 2008.
[3] P. Balister. Personal communication.
[4] B. Bollobás. *Modern Graph Theory.* Graduate texts in mathematics. Springer, Heidelberg, corrected edition, 1998.
[5] F. R. K. Chung, R. L. Graham, and R. M. Wilson. Quasi-random graphs. *Combinatorica*, 9(4):345–362, 1989.
[6] P. Erdős. On the number of complete subgraphs contained in certain graphs. *Magyar Tud. Akad. Mat. Kulató Int. Közl*, 7:459–464, 1962.
[7] P. Erdős and A. Hajnal. Ramsey-type theorems. *Discrete Applied Mathematics*, 25(12):37 – 52, 1989.
[8] V Falgas-Ravry and E.R. Vaughan. Applications of the semi-definite method to the Turán density problem for 3-graphs. *Combin. Probab. Comput.*, 22(1):21–54, 2013.
[9] A. W. Goodman. On sets of acquaintances and strangers at any party. *Amer. Math. Monthly*, 66:778–783, 1959.
[10] N. Linial and A. Morgenstern. Graphs with few 3-cliques and 3-anticliques are 3-universal. *Journal of Graph Theory*, to appear. http://arxiv.org/abs/1306.2020.
[11] S. Radziszowski. Small Ramsey numbers. *Electronic Jounal of Combinatorics*, DS1, 2011.
[12] D. Seinsche. On a property of the class of n-colorable graphs. *J. Combinatorial Theory Ser. B*, 16:191–193, 1974.
[13] M. Simonovits. A method for solving extremal problems in graph theory, stability problems. In *Theory of Graphs (Proc. Colloq., Tihany, 1966)*, pages 279–319. Academic Press, New York, 1968.
[14] A. Thomason. Pseudorandom graphs. In *Random graphs '85 (Poznań, 1985)*, volume 144 of *North-Holland Math. Stud.*, pages 307–331. North-Holland, Amsterdam, 1987.
[15] A. Thomason. A disproof of a conjecture of Erdős in Ramsey theory. *J. London Math. Soc*, 39:246–255, 1989.

School of Mathematics, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom
E-mail address: d.hefetz@bham.ac.uk, m.tyomkyn@bham.ac.uk