TOWARDS THE KOHAYAKAWA–KREUTER CONJECTURE ON ASYMMETRIC RAMSEY PROPERTIES

FRANK MOUSSET, RAJKO NENADOV, AND WOJCIECH SAMOTIJ

Abstract. For fixed graphs $F_1, \ldots, F_r$, we prove an upper bound on the threshold function for the property that $G(n, p) \to (F_1, \ldots, F_r)$. This establishes the 1-statement of a conjecture of Kohayakawa and Kreuter.

1. Introduction

Given a graph $G$, a positive integer $r$, and graphs $F_1, \ldots, F_r$, we write

$$G \to (F_1, \ldots, F_r)$$

if for every colouring of the edges of $G$ using colours from the set $[r] := \{1, \ldots, r\}$, there exists a copy of $F_i$ in $G$ whose all edges have been coloured $i$, for some $i \in [r]$. We study the asymptotic probability that

$$G(n, p) \to (F_1, \ldots, F_r)$$

for fixed graphs $F_1, \ldots, F_r$, where $G(n, p)$ is the binomial random graph with $n$ vertices and edge probability $p$.

An important special case of this problem, known as the symmetric case, arises when the graphs $F_1, \ldots, F_r$ are all the same. The study of symmetric Ramsey properties in random graphs was initiated by Łuczak, Ruciński, and Voigt [15], who proved that $p = n^{-1/2}$ is a threshold for the property

$$G(n, p) \to (K_3, K_3).$$

(The earlier work of Frankl and Rödl [4] established that $G(n, p) \to (K_3, K_3)$ under the stronger assumption that $p \geq n^{-1/2}$.) This was followed by a series of papers by Rödl and Ruciński [19, 20, 21] that culminated in the following statement. For a nonempty graph $F$, let $d_2(F) := 1/2$ if $F = K_2$ and $d_2(F) := \frac{r-1}{2(r-2)}$ otherwise and define the 2-density of $F$ by

$$m_2(F) := \max \{d_2(F') : F' \subseteq F \text{ with } e_{F'} \geq 1\}.$$  \hspace{1cm} (1)

Theorem 1 [21]. Let $r \geq 2$ and suppose that $F$ is a nonempty graph such that at least one component of $F$ is not a star or (in the case $r = 2$) a path of length three. Then there exist positive constants $c$ and $C$ such that

$$\lim_{n \to \infty} \mathbb{P}(G(n, p) \to \underbrace{F, \ldots, F}_{r \text{ times}}) = \begin{cases} 0 & \text{if } p \leq cn^{-1/m_2(F)} \\ 1 & \text{if } p \geq Cn^{-1/m_2(F)}. \end{cases}$$

One usually refers to the assertion of Theorem 1 for $p \leq cn^{-1/m_2(F)}$ as the 0-statement and to the assertion for $p \geq Cn^{-1/m_2(F)}$ as the 1-statement. It is worth pointing out that the assumption on the structure of $F$ is necessary. Indeed, if every component of $F$ is a star, then it is easy to see that $G \to (F, \ldots, F)$ as soon as $G$ has sufficiently many vertices of degree $r(\Delta_F - 1) + 1$. The function $n^{-1-1/(r(\Delta_F - 1)+1)}$ is a threshold for this property in $G(n, p)$; on the other hand, $m_2(F) = 1$ for every such $F$. In the case where $r = 2$ and at least one component of $F$ is a path of length three while the others are stars, the 0-statement of Theorem 1 is no longer true. For example, if $p = cn^{-1/m_2(F_3)} = cn^{-1}$ for some $c > 0$, then the probability that $G(n, p)$ contains a cycle of length five with an edge pending...
at every vertex is bounded from below by a positive constant (that depends on \(c\)); it is easy to check that every colouring of the edges of this graph with two colours yields a monochromatic path of length three. This exceptional case, originally missed by Rödl and Ruciński in [24], was eventually noticed and corrected by Friedgut and Krivelevich [5]; the corrected version of the 0-statement requires the assumption that \(p = o(n^{-1/m_2(F)})\). A short proof of Theorem 1 was given by Nenadov and Steger [18].

In the case where \(F\) is a tree (other than a star or the path of length three) [5], a triangle [6] or, more generally, a strictly 2-balanced graph that can be made bipartite by removing some edge [24], it is known that the property \(G(n, p) \rightarrow (F, F)\) has a sharp threshold:

**Theorem 2** ([5] [6] [24]). Suppose that \(F\) is either (i) a tree that is not a star or the path of length three or (ii) a strictly 2-balanced graph with \(e_F \geq 2\) edges that can be made bipartite by removing some edge. Then there exist \(c_0, c_1\), and a function \(c : \mathbb{N} \rightarrow [c_0, c_1]\) such that

\[
\lim_{n \to \infty} \mathbb{P}(G(n, p) \rightarrow (F, F)) = \begin{cases} 
0 & \text{if } p \geq (1 + \varepsilon)c(n)n^{2 - 1/m_2(F)} \\
1 & \text{if } p \leq (1 - \varepsilon)c(n)n^{2 - 1/m_2(F)}
\end{cases}
\]

for every positive constant \(\varepsilon\).

It is widely believed that one can choose \(c(n)\) to be a constant function; however proving this and, what is more, determining the value of the constant, remains a formidable challenge.

The main topic of this paper is the asymmetric case of the Ramsey problem in \(G(n, p)\), where the graphs \(F_1, \ldots, F_r\) are allowed to be different. This problem was first considered by Kohayakawa and Kreuter [12]. For nonempty graphs \(F_1\) and \(F_2\) with \(m_2(F_1) \geq m_2(F_2)\), we define the asymmetric 2-density

\[
m_2(F_1, F_2) := \max \left\{ \frac{e_{F_1'}}{v_{F_1'} - 2 + 1/m_2(F_2)} : F_1' \subseteq F_1 \text{ with } e_{F_1'} \geq 1 \right\}.
\]

The following generalisation of Theorem 1 is a slight rephrasing of a conjecture made by Kohayakawa and Kreuter [12]. (The original conjecture was stated only for two colours and it lacked the assumption that the graphs \(F_1\) and \(F_2\) are not forests, which was later added by Kohayakawa, Schacht, and Spöhel [13].)

**Conjecture 3.** Let \(r \geq 2\) and suppose that \(F_1, \ldots, F_r\) are graphs with \(m_2(F_1) \geq \cdots \geq m_r(F_r)\) and \(m_2(F_2) > 1\). Then there are positive constants \(c\) and \(C\) such that

\[
\lim_{n \to \infty} \mathbb{P}(G(n, p) \rightarrow (F_1, \ldots, F_r)) = \begin{cases} 
0 & \text{if } p \leq cn^{-1/m_2(F_1, F_2)} \\
1 & \text{if } p \geq Cn^{-1/m_2(F_1, F_2)}
\end{cases}
\]

The assumption that \(m_2(F_2) \geq 1\) is necessary, since otherwise we have \(m_2(F_2) = 1/2\) (i.e., \(F_2\) is a matching) and so \(m_2(F_1, F_2) = e_{F_1'}/v_{F_1'}\) for some nonempty subgraph \(F_1' \subseteq F_1\). In this case, for every constant \(C > 0\), the probability that \(G(n, p)\) with \(p = Cn^{-1/m_2(F_1, F_2)}\) contains no copies of \(F_1\) at all exceeds a positive constant (that depends on \(C\)); see, for example [11]. The assumption that \(m_2(F_2) > 1\) (which holds if and only if \(F_2\) is not a forest) is most likely not always necessary, but it precludes exceptional sequences \(F_1, \ldots, F_r\) such as a sequence of stars.

There have been several attempts at resolving Conjecture 3. Kohayakawa and Kreuter [12] proved it in the case where each \(F_i\) is a cycle. Marciniszyn, Skokan, Spöhel, and Steger [16] observed that the proof of the 1-statement of Conjecture 3 given in [12] for sequences of cycles extends to all sequences \(F_1, \ldots, F_r\) such that \(F_1\) is strictly 2-balanced, assuming the so-called KLR (Kohayakawa–Łuczak–Rödl) conjecture [13] holds. (The KLR conjecture has since been verified, see [2] [3] [23].) The main result of [16] however was a proof of the 0-statement of Conjecture 3 in the case where each \(F_i\) is a complete graph. A self-contained (i.e., not relying on the KLR conjecture) proof of the 1-statement of Conjecture 3 for \(r = 2\) that assumes a similar density condition, namely that \(F_1\) is strictly balanced w.r.t. \(m_2(\cdot, F_2)\), was given by Kohayakawa, Schacht, and Spöhel [13]. This result was generalised

---

1 A nonempty graph \(F\) is said to be 2-balanced if \(d_2(F) = m_2(F)\) and strictly 2-balanced if in addition \(d_2(F') < m_2(F)\) for every nonempty proper subgraph \(F' \subseteq F\).
by allowing \( F_1, \ldots, F_r \) to be uniform hypergraphs and extended from two to an arbitrary number of colours by Gugelmann, Nenadov, Person, Škorić, Steger, and Thomas. It was furthermore shown in [7] that the 1-statement of Conjecture 3 holds, with no additional conditions on the graphs \( F_1, \ldots, F_r \), under the stronger assumption that \( p \geq Cn^{-1/m_2(F_1,F_2)} \log n \). The proofs of both these statements employed the hypergraph container method developed by Balogh, Morris, and Samotij (see also [11]) and, independently, by Saxton and Thomason [23]. Let us also mention that the short argument of Nenadov and Steger [18] that establishes the \( (\mathrm{iii}) \)-statement of Conjecture 3 in the case where \( m_2(F_1) = m_2(F_2) \), with no further conditions on \( F_1 \) (this was explicitly observed in [8]).

In summary, all previous results related to the \( (\mathrm{iii}) \)-statement of Conjecture 3 require either some non-trivial assumptions on \( F_1 \) (or both \( F_1 \) and \( F_2 \)) or a stronger lower bound on \( p \). Our main contribution is a proof of this statement in its full generality.

**Theorem 4.** Let \( r \geq 2 \) and suppose that \( F_1, \ldots, F_r \) are graphs with \( m_2(F_1) \geq \cdots \geq m_2(F_r) \) and \( m_2(F_2) \geq 1 \). Then there exists a positive constant \( K \) such that if \( p = p(n) \geq Kn^{-1/m_2(F_1,F_2)} \), then

\[
\lim_{n \to \infty} \mathbb{P}(G(n,p) \not\rightarrow (F_1, \ldots, F_r)) = 1.
\]

### 2. Preliminaries

#### 2.1. Ramsey’s theorem

The following quantitative version of Ramsey’s theorem can be obtained from the usual version by a standard averaging argument (see, e.g., [7] Theorem 11).

**Lemma 5** (Ramsey’s theorem). For every positive integer \( r \) and all graphs \( F_1, \ldots, F_r \), there exists a positive \( \alpha \) and some \( n_0 \) such that the following holds for all \( n \geq n_0 \). For every colouring of the edges of \( K_n \) with colours from \([r]\), there exists a colour \( i \in [r] \) such that \( K_n \) contains at least \( \alpha n^{\frac{v_i}{m_i}} \) copies of \( F_i \) whose edges all have colour \( i \).

#### 2.2. Hypergraph containers

The following lemma is a well-known consequence of the hypergraph container theorems obtained by Balogh, Morris, and Samotij and independently by Saxton and Thomason [23].

**Lemma 6.** For every graph \( F \) and every positive \( \varepsilon \), there exists a positive \( C = C(F, \varepsilon) \) such that the following holds for all \( n \in \mathbb{N} \). Let \( F(n) \) be the family of all \( F \)-free graphs with vertex set \([n]\). Then there exist functions

\[
g : F(n) \to 2^{E(K_n)} \quad \text{and} \quad f : 2^{E(K_n)} \to 2^{E(K_n)}
\]

such that for every \( G \in F(n) \),

(i) \( g(G) \) has at most \( Cn^{2-1/m_2(F)} \) edges,

(ii) \( f(g(G)) \) contains at most \( \varepsilon n^{v_i} \) copies of \( F_i \), and

(iii) \( g(G) \subseteq G \subseteq f(g(G)) \).

### 3. Proof overview

Suppose that \( G \not\rightarrow (F_1, \ldots, F_r) \), that is, that there exists a colouring \( c : E(G) \to [r] \) such that for each \( i \in [r] \), the graph \( c^{-1}(i) \) of edges in colour \( i \) is \( F_i \)-free. It follows from Lemma 6 that there are ‘signatures’ \( S_2 = g_2(c^{-1}(2)), \ldots, S_r = g_r(c^{-1}(r)) \), each with \( |S_i| = O(n^{2-1/m_2(F_i)}) \leq O(n^{2-1/m_2(F_1)}) \) edges, such that \( S_i \subseteq c^{-1}(i) \subseteq f_i(S_i) \) for each \( i \in \{2, \ldots, r\} \), where the graph \( f_i(S_i) \) contains \( o(n^{v_i}) \) copies of \( F_i \). Since each edge of \( G \) that lies outside \( f(S_2) \cup \ldots \cup f(S_r) \) is coloured 1, the intersection of \( G \) and the graph

\[
K(S_2, \ldots, S_r) := K_n \setminus \big( f(S_2) \cup \ldots \cup f(S_r) \big)
\]

must be \( F_1 \)-free. In particular, the event \( G(n,p) \not\rightarrow (F_1, \ldots, F_r) \) is contained in the union of the events

\[
S_2 \cup \cdots \cup S_r \subseteq G(n,p) \quad \text{and} \quad G(n,p) \cap K(S_2, \ldots, S_r) \text{ is } F_1 \text{-free,}
\]
where \((S_2, \ldots, S_r)\) ranges over all sequences of ‘signatures’. Since the property “\(S_2 \cup \cdots \cup S_r \subseteq G\)” is increasing in \(G\) and the property “\(G \cap K(S_1, \ldots, S_r)\) is \(F_1\)–free” is decreasing in \(G\), Harris’ inequality and the union bound yield
\[
P(G(n, p) \rightarrow (F_1, \ldots, F_r)) \leq \sum_{(S_2, \ldots, S_r)} P(S_2 \cup \cdots \cup S_r \subseteq G(n, p)) \cdot P(G(n, p) \cap K(S_2, \ldots, S_r) \text{ is } F_1\text{-free}).
\]
Lemma 5 implies that the graph \(K(S_2, \ldots, S_r)\) has at least \(\delta n^{v_1}\) copies of \(F_1\), for some constant \(\delta > 0\) that is independent of \((S_2, \ldots, S_r)\), and consequently, following [10], one can derive the bound
\[
P(G(n, p) \cap K(S_2, \ldots, S_r) \text{ is } F_1\text{-free}) \leq \exp(-\delta' \cdot \min\{n^{v_1}/p^{\delta'} : \emptyset \neq I \subseteq F_1\}) \tag{3}
\]
from Janson’s inequality.

If one assumes that \(p \gg n^{-1/m_2(F_1,F_2)} \log n\), then the right-hand side of (3) can be bounded from above by \(\exp(-\omega(n^{2-1/m_2(F)}) \log n))\), whereas there are only \(\exp\left(O(n^{2-1/m_2(F)} \log n)\right)\) sequences \((S_2, \ldots, S_r)\); thus one may conclude that \(G(n, p) \rightarrow (F_1, \ldots, F_r)\) with probability very close to 1, without any further assumptions on \((F_1, \ldots, F_r)\). The weaker assumption that \(p \gg n^{-1/m_2(F_1,F_2)}\) implies only the upper bound \(\exp(-\omega(n^{2-1/m_2(F)})\)) on the right-hand side of (3) and the challenge is to obtain the estimate
\[
\sum_{(S_2, \ldots, S_r)} P(S_2 \cup \cdots \cup S_r \subseteq G(n, p)) = \exp\left(O(n^{2-1/m_2(F)})\right).
\]
Unfortunately, this estimate is valid only if \(p = O(n^{-1/m_2(F_1,F_2)})\), which we can assume only if \(m_2(F_1) = m_2(F_2)\). This is the essence of why proving the 1-statement of Conjecture 3 is significantly more difficult than proving the 1-statement of Theorem 4.

A first step towards making the above union bound argument more efficient, already taken in [7], is to restrict the family of ‘non-Ramsey’ colourings that are being considered. To this end, observe that every colouring \(c: E(G) \rightarrow [r]\) such that
\[
c^{-1}(i) \text{ is } F_1\text{-free for every } i \in [r] \tag{4}
\]
may be altered as follows: Every edge of \(G\) that is not contained in a copy of \(F_1\) in \(G\) is recoloured 1. This way we obtain a new coloring \(c'\) but now each edge of \(c^{-1}(2) \cup \cdots \cup c^{-1}(r)\) lies in a copy of \(F_1\) in \(G\). We may thus replace the event “\(S_2 \cup \cdots \cup S_r \subseteq G(n, p)\)” in the above argument with the event “each edge of \(S_2 \cup \cdots \cup S_r\) lies in a copy of \(F_1\) in \(G(n, p)\)”, which we will abbreviate by \(S_2 \cup \ldots \cup S_r \subseteq_{F_1} G(n, p)\), and conclude that the event \(G(n, p) \rightarrow (F_1, \ldots, F_r)\) is contained in the union of the events
\[
S_2 \cup \ldots \cup S_r \subseteq_{F_1} G(n, p) \quad \text{and} \quad G(n, p) \cap K(S_2, \ldots, S_r) \text{ is } F_1\text{-free},
\]
where again \((S_2, \ldots, S_r)\) ranges over all sequences of ‘signatures’. Since the property \(S_2 \cup \cdots \cup S_r \subseteq_{F_1} G\) is still increasing in \(G\), in order to complete the argument, it would suffice to prove that
\[
\sum_{(S_2, \ldots, S_r)} P(S_2 \cup \cdots \cup S_r \subseteq_{F_1} G(n, p)) = \exp\left(O(n^{2-1/m_2(F)})\right). \tag{5}
\]
Unfortunately, nothing in the spirit of (5) can be true in general. Indeed, there are pairs of graphs \(F_1, F_2\) such that \(m_2(F_1) > m_2(F_1, F_2) > m_2(F_2)\), but typically every edge of \(G(n, p)\) lies in a copy of \(F_1\) (for example, this is the case where \(F_1\) contains an isolated edge), which contradicts (5). However, in the case where \(F_1\) is strictly balanced w.r.t. \(m_2(\cdot, F_2)\), then one can prove a version of (5) that complements an argument similar to the one we have outlined above; this was achieved in [17].

In order to dispose of the balancedness assumption, we shall restrict our attention only to a subcollection of all copies of \(F_1\) in \(K_n\). More precisely, we will require \(c^{-1}(1)\) to avoid only a certain family \(\mathcal{F}\) of copies of \(F_1\), which we term typed copies of \(F_1\). Moreover, we will replace \(G(n, p)\) with a sparser random subgraph of it. (We will define both these notions formally in Section 4.) We are going to do

\footnote{When \(m_2(F_1) > m_2(F_2)\), even if one assumes that \(F_1\) is strictly balanced w.r.t. \(m_2(\cdot, F_2)\), there is an event of very small probability in \(G(n, p)\) that nevertheless blows up the left-hand side of (5) above \(\exp\left(O(n^{2-1/m_2(F_2)})\right)\). However, after conditioning on the complement of this event, the inequality (5) becomes true.}
both of these things in such a way that: (i) the left-hand side of (3) can be still bounded from above by \( \exp(-\omega(n^{2-1/m}\bar{H})) \) and (ii) a version of (3) holds true. Since the left-hand side of (3) is decreasing in both \( G(n,p) \) and the family \( \mathcal{F} \), whereas each term in the left-hand side of (3) is increasing in both \( G(n,p) \) and \( \mathcal{F} \), we will have to strike a delicate balance in order to achieve (i) and (ii) simultaneously.

4. \( H \)-typed graphs

For the entirety of this section, suppose that \( H \) and \( F \) are fixed nonempty graphs.

An important role in our proof is played by weight functions \( w: E(H) \to [1,\infty) \). Given such a function \( w \) and a subgraph \( I \subseteq H \), we shall use the shorthand \( w_I := \sum_{e \in E(I)} w(e) \).

**Definition 7** \((w, F)\)-balanced graphs. Given a function \( w: E(H) \to [1,\infty) \), we say that \( H \) is \((w, F)\)-balanced if for every edge \( e \in E(H) \), we have

\[
\min \left\{ v_I - w_I/m_2(H, F) : I \subseteq H \text{ with } e \in E(I) \right\} = 2 - 1/m_2(F).
\]

**Lemma 8.** If \( m_2(H) \geq m_2(F) \), then there exists a function \( w: E(H) \to [1,\infty) \) such that \( H \) is \((w, F)\)-balanced.

**Proof.** Given a function \( w: E(H) \to [1,\infty) \) and an edge \( e \in E(H) \), let us write

\[
r_e(w) := \min \left\{ v_I - w_I/m_2(H, F) : I \subseteq H \text{ with } e \in E(I) \right\} - 2 + 1/m_2(F).
\]

Our goal is then to show that there is a function \( \hat{w} \) such that \( r_e(\hat{w}) = 0 \) for all \( e \in E(H) \). To this end, consider the set \( \mathcal{W} \) of all functions \( w: E(H) \to [1,\infty) \) that satisfy \( r_e(w) \geq 0 \) for all \( e \in E(H) \). Note that the constant function \( w \equiv 1 \) belongs to \( \mathcal{W} \), by the definition of \( m_2(H, F) \), and that \( 0 \leq r_e(w) \leq 1/m_2(F) - w(e)/m_2(H, F) \) for every \( e \in E(H) \) and all \( w \in \mathcal{W} \). In particular, \( \mathcal{W} \) is nonempty and compact and thus the (continuous) map \( \mathcal{W} \ni \hat{w} \mapsto w_H \in [0,\infty) \) achieves its maximum at some \( \hat{w} \in \mathcal{W} \).

We claim that \( r_e(\hat{w}) = 0 \) for each \( e \in E(H) \). It this were not true and there was an \( e \in E(H) \) satisfying \( r_e(\hat{w}) > 0 \), then for some sufficiently small \( \varepsilon > 0 \), the function \( \tilde{w} \) defined by \( \tilde{w}(f) = \hat{w}(f) + \varepsilon \cdot 1[f = e] \) would belong to \( \mathcal{W} \), contradicting the maximality of \( \hat{w} \).

4.1. \( H \)-typed graphs. By an \( H \)-typed graph we mean a graph \( \mathcal{G} \) equipped with a type function \( \tau_H: E(\mathcal{G}) \to E(H) \). If \( H \) is clear from the context, we shall just say that \( \mathcal{G} \) is a typed graph. We shall mostly use calligraphic letters to refer to typed graphs. However, every subgraph \( I \subseteq H \) will be treated as a typed graph in the natural way, by taking its type function to be the inclusion map from \( E(I) \) into \( E(H) \). We write \( \mathcal{G}' \cong \mathcal{G} \) if there is a graph isomorphism between \( \mathcal{G}' \) and \( \mathcal{G} \) that preserves the type of every edge (in this case, we say that \( \mathcal{G} \) and \( \mathcal{G}' \) are typomorphic). We write \( \mathcal{G} \preceq \mathcal{G} \) if \( \mathcal{G}' \subseteq \mathcal{G} \) and \( \tau_H(e) = \tau_H(e) \) for all \( e \in E(\mathcal{G}') \). A typed copy of a subgraph \( I \subseteq H \) in \( \mathcal{G} \) is a typed graph \( \mathcal{I} \cong I \) where \( I \) is treated as a typed graph. Finally, for a set \( \mathcal{H} \) of (untyped) copies of \( H \) in \( \mathcal{K}_n \) and a typed graph \( \mathcal{G} \) on the vertex set \([n]\), we write \( \mathcal{H}(\mathcal{G}) \) for the set of all \( \tilde{H} \in \mathcal{H} \) such that \( \tilde{H} \subseteq E(\mathcal{G}) \) and the typed graph \( \mathcal{H} \) obtained by equipping \( \tilde{H} \) with the type function \( \tau_H|_{E(\tilde{H})} \) is typomorphic to \( H \).

4.2. Random \( H \)-typed graphs. We define a random \( H \)-typed graph \( \mathcal{G}(n, p, w) \) as follows.

**Definition 9** \((\mathcal{G}(n, p, w))\). Given \( n \in \mathbb{N} \), \( p \in (0,1) \), and \( w: E(H) \to [1,\infty) \), we define \( \mathcal{G}(n, p, w) \) to be the random \( H \)-typed graph \( \mathcal{G} \) on the vertex set \([n]\) constructed by the following two-step procedure

(i) choose a function \( \tau: E(K_n) \to E(H) \) uniformly at random,

(ii) include every \( e \in E(K_n) \) independently with probability \( p^w(\tau(e)) \) and set \( \tau_H = \tau|_{E(\mathcal{G})} \).

We shall be using the following estimate of the upper tail of the number of typed copies of \( H \) and its subgraphs in \( \mathcal{G}(n, p, w) \); the proof is a straightforward modification of the classical argument of Ruciński and Vince [22].
It is easy to check that
\[ \mathbb{P}(X_I \geq 2\mathbb{E}[X_I]) \leq c \cdot \left( \min_{\emptyset \neq I' \subseteq I} n^{v_{I'}^r} p^{w_{I'}^r} \right)^{-1}, \]
for some positive constant \( c \) depending only on \( H \).

**Proof.** It is easy to see that
\[ \mathbb{E}[X_I] = \Theta(n^{v_I^r} p^{w_I^r}) \]
and
\[ \text{Var}[X_I] = O \left( \max_{\emptyset \neq I' \subseteq I} n^{2v_{I'}^r - v_I^r} p^{2w_{I'}^r - w_I^r} \right) = O \left( \frac{\mathbb{E}[X_I]^2}{\min_{\emptyset \neq I' \subseteq I} n^{v_{I'}^r} p^{w_{I'}^r}} \right). \]
The assertion now follows from Chebyshev’s inequality \( \mathbb{P}(X_I \geq 2\mathbb{E}[X_I]) \leq \text{Var}[X_I]/\mathbb{E}[X_I]^2 \). \( \square \)

**Lemma 10.** Fix a nonempty subgraph \( I \subseteq H \) and let \( X_I \) denote the number of typed copies of \( I \) in \( \mathcal{G}(n,p,w) \). We have
\[ \mathbb{P}(X_I \geq 2\mathbb{E}[X_I]) \leq c \cdot \left( \min_{\emptyset \neq I' \subseteq I} n^{v_{I'}^r} p^{w_{I'}^r} \right)^{-1}, \]
for some positive constant \( c \) depending only on \( H \).

**Proof.** For a given copy \( C \subseteq H \), let us write \( \mathbb{1}_C \) for the indicator variable of the event that \( C \in \mathbb{H}(\mathcal{G}(n,p,w)) \). Thus \( |\mathbb{H}(\mathcal{G}(n,p,w))| = \sum_{C \in \mathbb{H}} \mathbb{1}_C \). Observe that \( X_C \) and \( X_{C'} \) are independent if \( C \) and \( C' \) are edge-disjoint.

For every \( C \in \mathbb{H} \), we have \( \mathbb{E}[\mathbb{1}_C] \geq p^{u_C^r} / e_H^{u_C} \), and so
\[ \mu := \mathbb{E}[|\mathbb{H}(\mathcal{G}(n,p,w))|] \geq |\mathbb{H}| p^{u_H^r} / e_H^{u_H} \geq \alpha n^{v_H^r} p^{w_H^r} / e_H^{u_H}. \]
Define
\[ \Delta := \sum_{\{C,C'\} \subseteq \mathbb{H}} \mathbb{E}[\mathbb{1}_C \mathbb{1}_{C'}], \quad \delta := \max_{C \in \mathbb{H}} \sum_{C' \in \mathbb{H}} \mathbb{E}[\mathbb{1}_{C'}], \]
It is easy to check that
\[ \Delta = O \left( \frac{\mu^2}{\min_{\emptyset \neq I' \subseteq I} n^{v_{I'}^r} p^{w_{I'}^r}} \right), \]
where the constants implicit in the \( O \)-notation may depend on \( \alpha \) and on \( H \). It is also easy to see that
\[ \delta \leq e_H n^{v_H^r - 2} p^{w_H}. \]
The claim then follows from the following version of Suen’s inequality due to Janson [9, Theorem 3]:
\[ \mathbb{P}(|\mathbb{H}(\mathcal{G}(n,p,w))| = 0) \leq \exp \left( -\min \left( \frac{\mu^2}{8\Delta}, \frac{\mu}{6\delta}, \frac{\mu}{2} \right) \right). \]
Note in particular that \( \mu/\delta \geq \alpha n^2 / e_H^{u_H} = \Omega(n^{v_H^r} p^{w_H^r}) \) for every subgraph \( I \subseteq H \) consisting of two vertices and one edge. \( \square \)

5. PROOF OF THEOREM

Fix nonempty graphs \( F_1, \ldots, F_r \) with \( m_2(F_1) \geq \cdots \geq m_2(F_r) \) and \( m_2(F_2) \geq 1 \). For the sake of brevity, we shall write \( H = F_1 \) and \( F = F_2 \). For each \( i \in \{2, \ldots, r\} \), let \( C_i, f_i, \) and \( g_i \) be as given by Lemma 11 applied to the graph \( F_i \) and to some sufficiently small positive constant \( \varepsilon = \varepsilon(r, F_1, \ldots, F_r) \); let \( C := \max \{ C_2, \ldots, C_r \} \).

Given a typed graph \( \mathcal{G} \) on \([n]\), we write \( \mathcal{G} \rightarrow (H, F_2, \ldots, F_r) \) if there exists a colouring \( c: E(\mathcal{G}) \rightarrow [r] \) such that there is neither a typed copy of \( H \) in \( \mathcal{G} \) nor an (untyped) copy of \( F_i \) in colour \( i \), for any \( i \in \{2, \ldots, r\} \). Note that if \( \mathcal{G} \rightarrow (H, F_2, \ldots, F_r) \), then there also exists such a colouring \( c \) where \( c \) additionally every edge of \( \mathcal{G} \) that is not contained in a typed copy of \( H \) has colour 1. Indeed, we can always recolour such edges in colour 1 without creating any copies of \( H \) in \( c^{-1}(1) \).

Now, assume that \( \mathcal{G} \rightarrow (H, F_2, \ldots, F_r) \) and let \( c: E(\mathcal{G}) \rightarrow [r] \) be a colouring that satisfies all of the above conditions. For each \( i \in \{2, \ldots, r\} \), let \( S_i := g_i(c^{-1}(i)) \) and recall that \( S_i \) has at most \( Ch^{2-1/m_2(F)} \) edges. By the definition of \( f_i \) and \( g_i \), we have \( S_i \subseteq c^{-1}(i) \subseteq f_i(S_i) \) for every
implies that the probability that witnesses $W \in I$. We call such a pair a pair of $G$. If $\mathcal{S} = \{S_2, \ldots, S_r\}$, with $S_2, \ldots, S_r$ pairwise disjoint, and some $W \in \mathcal{W}_n$ such that:

1. $S_2 \sqcup \cdots \sqcup S_r \subseteq \mathcal{W}$ and $W \not\in \mathcal{G}$ and
2. $\text{HS}(\mathcal{G}) = \emptyset$.

We call such a pair $(\mathcal{S}, W)$ a witness for the fact that $\mathcal{G} \not\rightarrow (H, F_2, \ldots, F_r)$. Now, suppose that $p \geq Kn^{-1/m_2(H,F)}$ for a sufficiently large constant $K$. Our goal is to prove that

$$\mathbb{P}(G(n,p) \not\rightarrow (H, F_2, \ldots, F_r)) = o(1).$$

As the property $G \not\rightarrow (H, F_2, \ldots, F_r)$ is monotone in $G$, we may assume without loss of generality that $p = Kn^{-1/m_2(H,F)}$. Since there is a natural coupling of $G(n,p)$ and $\mathcal{G}(n,p,w)$ such that $E(G(n,p,w)) \subseteq E(G(n,p))$, we may conclude that

$$\mathbb{P}(G(n,p) \not\rightarrow (H, F_2, \ldots, F_r)) \leq \mathbb{P}(\mathcal{G}(n,p,w) \not\rightarrow (H, F_2, \ldots, F_r)).$$

In particular, it suffices to show that the probability that $\mathcal{G}(n,p,w)$ admits a witness $(\mathcal{S}, W)$ for the fact that $\mathcal{G}(n,p,w) \not\rightarrow (H, F_2, \ldots, F_r)$ is small. This is what we are going to do in the remainder of the proof.

Let $\mathcal{G} \sim \mathcal{G}(n,p,w)$. For a subgraph $I \subseteq H$, we write $X_I$ for the number of typed copies of $I$ contained in $\mathcal{G}$. We shall now split the proof into two cases, depending on whether or not $X_I$ exceeds $2\mathbb{E}[X_I]$ for some nonempty $I \subseteq H$.

**Case 1.** There is a nonempty subgraph $I \subseteq H$ such that $\mathcal{G}$ has more than $2\mathbb{E}[X_I]$ typed copies of $I$.

The probability that $\mathcal{G}$ is in this case tends to zero by Lemma 10, because $n^{r1}p^{w1} \geq Kn^{-1/m_2(F)} \to \infty$ for every nonempty $I \subseteq H$, by the definition of $(w,F)$-balancedness.

**Case 2.** For every nonempty subgraph $I \subseteq H$, $\mathcal{G}$ contains at most $2\mathbb{E}[X_I]$ typed copies of $I$.

Let us write $\mathcal{U}_n \subseteq \mathcal{W}_n$ for the subset comprising all $W \in \mathcal{W}_n$ that contain at most $2\mathbb{E}[X_I]$ typed copies of every nonempty $I \subseteq H$. Note that if $(\mathcal{S}, W)$ is a witness for $\mathcal{G} \not\rightarrow (H, F_2, \ldots, F_r)$, then necessarily $W \in \mathcal{U}_n$, since otherwise $\mathcal{G}$ would fall into the first case. Let $Z$ denote the number of witnesses $(\mathcal{S}, W)$ with $W \in \mathcal{U}_n$. We shall show that $\mathbb{E}[Z] = o(1)$, which, by Markov’s inequality, will imply that the probability that $\mathcal{G} \not\rightarrow (H, F_2, \ldots, F_r)$ tends to zero.

To this end, we have

$$\mathbb{E}[Z] \leq \sum_{W \in \mathcal{U}_n} \sum_{S \in \mathcal{S}(W)} \mathbb{P}((\mathcal{S}, W) \text{ is a witness for } \mathcal{G} \not\rightarrow (H, F_2, \ldots, F_r)).$$
where we write $\mathcal{S}(\mathcal{W})$ for the set of all sequences $\mathbf{S} = (S_1, \ldots, S_r) \in \mathcal{S}$ such that $S_2, \ldots, S_r$ are pairwise edge-disjoint and $S_2 \cup \cdots \cup S_r \subseteq \mathcal{W}$ (in particular, note that $|\mathcal{S}(\mathcal{W})| \leq r^e(\mathcal{W})$). If a pair $(\mathbf{S}, \mathcal{W})$ is a witness for $G \mapsto (H, F_2, \ldots, F_r)$, then $\mathcal{W} \subseteq G$ and $H_{\mathbf{S}}(G) = \emptyset$. Thus

$$E[Z] \leq \sum_{W \in \mathcal{U}_n} \sum_{\mathbf{S} \in \mathcal{S}(\mathcal{W})} P(W \cap G) \cdot P(H_{\mathbf{S}}(G) = \emptyset).$$

Given a $W \in \mathcal{U}_n$ and an $\mathbf{S} \in \mathcal{S}(\mathcal{W})$, let $H_{\mathbf{S}}^W$ denote the collection of all (untyped) copies of $H$ in $H_{\mathbf{S}}$ that are edge-disjoint from $\mathcal{W}$. Since $H_{\mathbf{S}}^W(\mathcal{G}) \leq H_{\mathbf{S}}(\mathcal{G})$ and, crucially, the events $H_{\mathbf{S}}^W(\mathcal{G}) = \emptyset$ and $W \subseteq G$ are independent, we obtain

$$E[Z] \leq \sum_{W \in \mathcal{U}_n} \sum_{\mathbf{S} \in \mathcal{S}(\mathcal{W})} P(W \subseteq G) \cdot P(H_{\mathbf{S}}^W(\mathcal{G}) = \emptyset) = \sum_{W \in \mathcal{U}_n} P(W \subseteq G) \cdot \sum_{\mathbf{S} \in \mathcal{S}(\mathcal{W})} P(H_{\mathbf{S}}^W(\mathcal{G}) = \emptyset).$$

Since each $f_i(S_i)$ contains at most $\varepsilon n^w \log n$ copies of $F_i$, it follows from Ramsey’s theorem (Lemma 5) that if $\varepsilon = \varepsilon(r, H, F_2, \ldots, F_r)$ is sufficiently small, then $|H_{\mathbf{S}}| \geq 2\varepsilon n^w$. Consequently,

$$|H_{\mathbf{S}}^W| \geq |H_{\mathbf{S}}| - e(W) \cdot e_H \cdot n^w \geq 2\varepsilon n^w - O(n^{w-1/m_2(F)}) \geq \varepsilon n^w.$$

It thus follows from Lemma 11 and the fact that $H$ is $(w, F)$-balanced that

$$P(H_{\mathbf{S}}^W(\mathcal{G}) = \emptyset) \leq \exp \left( -c \cdot \min_{\mathbf{S} \neq \mathbf{S}'} n^w \cdot p^w \right) \leq e^{-cK n^{2-1/m_2(F)}},$$

for some positive constant $c$ that depends only on $H$ and $\varepsilon$. Therefore,

$$E[Z] \leq \sum_{W \in \mathcal{U}_n} P(W \subseteq G) \sum_{\mathbf{S} \in \mathcal{S}(\mathcal{W})} e^{-cK n^{2-1/m_2(F)}} \leq \sum_{W \in \mathcal{U}_n} P(W \subseteq G) \cdot r^e(\mathcal{W}) \cdot e^{-cK n^{2-1/m_2(F)}} \leq e^{-cK n^{2-1/m_2(F)}/2} \cdot \sum_{W \in \mathcal{U}_n} P(W \subseteq G),$$

where we have used that every $W \in \mathcal{U}_n$ has at most $T(n)$ edges and thus for sufficiently large $K$,

$$e(W) \log r \leq T(n) \log r \leq e_H(r - 1) C n^{2-1/m_2(F)} \log r \leq cK n^{2-1/m_2(F)}/2.$$
the typed graph $W_i = \bigcup_{j \in \{I \setminus i\}} H_j$ belongs to $\mathcal{U}_{n,k-1}$. Moreover, all these typed graphs are distinct, because every $H_j$ contains an edge that is not covered by the union of all the other $H_j$ for $j \neq i$. In other words, for every $W \in \mathcal{U}_{n,k}$, there are at least $k$ distinct typed graphs $W_i \in \mathcal{U}_{n,k-1}$ such that $W = W_i \cup H_i$ for some typed copy $H_i$ of some $H_f$ with $f \in E(H)$. Denoting for each $I \subseteq H$ the set of all typed graphs with vertices from $[n]$ that are isomorphic to $I$ by $\mathcal{C}_n(I)$, have

$$k \cdot \mathbb{E}[U_k] = k \sum_{W \in \mathcal{U}_{n,k}} \mathbb{P}(W \leq \mathcal{G}) \leq \sum_{W' \in \mathcal{U}_{n,k-1}} \sum_{f \in E(H)} \sum_{H' \in \mathcal{C}(H_f)} \mathbb{P}(W' \cup H' \leq \mathcal{G}).$$

where we further require that the type function of each $H'$ in the last inequality uses $\mathcal{G}$ into (11) and dividing through by the proof.

$$\square$$

Because for every $W \in \mathcal{U}(H_f)$, every $I \subseteq H_f$, and every $\mathcal{I} \in \mathcal{C}_n(I)$ with $I \subseteq H'$, we have $\mathbb{P}(H' \setminus I \leq \mathcal{G}) = p^{w_{H_f} - w_I}$, and since by the definition of $\mathcal{U}_n$, there are at most $2^{|\mathcal{I}|} |\mathcal{I}|^{p w_I}$ typed copies of $I$ in $W'$, we get

$$\sum_{H' \in \mathcal{C}(H_f)} \mathbb{P}(W' \cup H' \leq \mathcal{G}) \leq \sum_{I \subseteq H_f} \sum_{\mathcal{I} \in \mathcal{C}_n(I)} \sum_{I \cap \mathcal{I} \cap \mathcal{H'} = \mathcal{I}} \mathbb{P}(W' \leq \mathcal{G}) \cdot \mathbb{P}(H' \setminus I \leq \mathcal{G}),$$

which last inequality utilizes $p = Kn^{-1/m_2(H,F)}$ and the definition of $H_f$. Substituting this bound into (11) and dividing through by $k$, we obtain (10).

Combining (9) and the claim, we obtain

$$\sum_{W \in \mathcal{U}_n} \mathbb{P}(W \leq \mathcal{G}) \leq \sum_{k \in T(n)} \left( \frac{e_H \cdot K w_H n^{2-1/m_2(F)}}{k} \right)^k \leq \left( T(n) + 1 \right) \cdot \left( \frac{e_H \cdot K w_H n^{2-1/m_2(F)}}{T(n)} \right)^{T(n)}$$

where the last inequality follows from the fact that the function $x \mapsto (c/x)^x$ is increasing for $0 < x \leq c/e$ and that $T(n) = e_H (r-1) C n^{2-1/m_2(F)} \leq c_H \cdot K w_H n^{2-1/m_2(F)}/e$ if $K$ is large. This yields

$$\sum_{W \in \mathcal{U}_n} \mathbb{P}(W \leq \mathcal{G}) \leq \frac{2 c H K w_H}{e_H (r-1) C} \leq e^{O(\log K) n^{2-1/m_2(F)}},$$

which, together with (8), implies that $\mathbb{E}[Z] \to 0$, provided that $K$ is sufficiently large. This completes the proof.

$$\square$$

6. Concluding remarks

While, even before our work, the 1-statement of Conjecture 3 was known to be true up to a log $n$ factor, the situation with the 0-statement is quite different. So far it has only been verified in the case where all the graphs $F_1, \ldots, F_r$ are either cycles [12] or complete graphs [16]. The general case, however, seems to be rather difficult. A criterion which reduces the 0-statement of Conjecture 3 to a purely deterministic question, a potentially fruitful approach, was given in [7]. We now present this reduction.

Given graphs $F_1$ and $F_2$, let $F(F_1, F_2)$ be the family of all graphs $F$ with the following property: There exists a copy $F'_0$ of $F_2$ in $F$ and an edge $e_0 \in E(F'_2)$ such that for each $e \in E(F'_2) \setminus \{e_0\}$, there is a copy $F'_1$ of $F_1$ in $F$ containing $e$ and

$$E(F) = E(F'_2) \cup \bigcup_{e \in E(F'_2) \setminus \{e_0\}} E(F'_1);$$

$\text{At least}$ since, again, $W$ can have more than one representation as $H_1 \cup \cdots \cup H_k$. 

$\square$
Theorem 13. Let $F_1$ and $F_2$ be graphs such that $m_2(F_1) > m_2(F_2) > 1$ and suppose that the following holds:

(i) $F_1$ and $F_2$ are strictly 2-balanced,
(ii) $F_1$ is strictly balanced w.r.t. $m_2(\cdot, F_2)$,
(iii) $\mathbf{F}(F_1, F_2)$ is asymmetric-balanced, and
(iv) for every graph $G$ such that

\[
\max_{G' \subseteq G} \frac{e(G')}{v(G')} \leq m_2(F_1, F_2)
\]

we have $G \not\rightarrow (F_1, F_2)$.

Then there exists $c > 0$ such that if $p \leq cn^{-1/m_2(F_1,F_2)}$, then

\[
\lim_{n \to \infty} \mathbb{P}(G(n, p) \rightarrow (F_1, F_2)) = 0.
\]

A minor modification of the proof of Theorem 13 shows that one can further drop the requirement in (i) that $F_1$ is strictly 2-balanced. Therefore, in order to prove the 0-statement of Conjecture 3 it is enough to consider a strictly 2-balanced subgraph $F_2' \subseteq F_2$ with $m_2(F_2') = m_2(F_2)$ and a subgraph $F_1' \subseteq F_1$ that is strictly balanced w.r.t $m_2(\cdot, F_2')$ and satisfies $m_2(F_1', F_2') = m_2(F_1, F_2')$, and show that conditions (iii) and (iv) in Theorem 13 hold. As an exercise, we invite the reader to show this in the case where $m_2(F_1) = m_2(F_2)$ (for part (iv) see, e.g., the appendix of [18]). In this case, it turns out that if $F_1' \subseteq F_1$ is chosen in the manner described above, then it is also strictly 2-balanced and $m_2(F_1') = m_2(F_1)$; in particular, one can use Theorem 13 without any modifications. Unfortunately, the general case remains wide open.

Finally, let us mention that the proof of Theorem 14 transfers to the setting of uniform hypergraphs with almost no changes. However, unlike for graphs, in the case of hypergraphs of uniformity larger than two, even in the symmetric case (i.e., $F_1 = \cdots = F_r$) a complete characterisation of the threshold functions is not known. We refer the interested reader to [9, 17] for further details.

References


Frank Mousset, School of Mathematical Sciences, Tel Aviv University, Tel Aviv 6997801, Israel

Email address: moussetfrank@gmail.com

Rajko Nenadov, Department of Computer Science, ETH Zurich, 8092 Zürich, Switzerland

Email address: rnenadov@inf.ethz.ch

Wojciech Samotij, School of Mathematical Sciences, Tel Aviv University, Tel Aviv 6997801, Israel

Email address: samotij@post.tau.ac.il