UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
DEPARTAMENTO DE MATEMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

WALNER MENDONÇA DOS SANTOS

ON RAMSEY PROPERTY FOR RANDOM GRAPHS

FORTALEZA
2016

WALNER MENDONÇA DOS SANTOS

ON RAMSEY PROPERTY FOR RANDOM GRAPHS

Dissertação apresentada ao Programa de Pós-graduação em Matemática do Departamento de Matemática da Universidade Federal do Ceará, como parte dos requisitos necessários para a obtenção do título de Mestre em Matemática. Área de concentração: Combinatória.

Orientador: Prof. Dr. Fabricio Siqueira Benevides

FORTALEZA

2016

WALNER MENDONÇA DOS SANTOS

ON RAMSEY PROPERTY FOR RANDOM GRAPHS

Dissertação apresentada ao Programa de Pós-graduação em Matemática do Departamento de Matemática da Universidade Federal do Ceará, como parte dos requisitos necessários para a obtenção do título de Mestre em Matemática. Área de concentração: Combinatória.

Aprovoda em: 16/ 08 / 2016.

BANCA EXAMINADORA

Prof. Dr. Fabricio Siqueira Benevides (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Rudini Menezes Sampaio
Universidade Federal do Ceará (UFC)

Prof. Dr. Guilherme Oliveira Mota
Universidade de São Paulo (USP)

Dedico este trabalho à minha mãe, a quem devo tudo que sou.

AGRADECIMENTOS

À minha família. Em especial, à minha mãe, por tudo que sou hoje e pelo exemplo de pessoa a qual me espelho.

Ao professor Fabricio Benevides, pela orientação acadêmica e pelos conselhos pessoais e por ser sempre prestativo, esclarecedor, e racional e preciso em suas observações.

A todos os professores do Departamento de Matemática com os quais tive aula, por terem compartilhado o amor pela matemática e serem grande mestres dos quais tenho como exemplos. Em especial, aos professores Ana Shirley, Antônio Caminha, José Fábio e Lev Birbrair, por terem permitido que cada uma de suas aulas fossem além dos livros.

Aos grandes amigos e colegas do Departamento de Matemática. Em especial, ao Esdras Muniz, Janielly Araújo, José Danuso, Neilha Pinheiro, Rafael Alves, Renan Santos, Rodrigo Matos e Rosa Tayane, pela amizade e por terem tornado o ambiente acadêmico bem mais agradável.

Ao Guilherme Mota e Rudini Sampaio, por aceitarem participar da banca examinadora, por lerem a dissertação em tão pouco tempo e pelas correções as quais, sem dúvida alguma, ajudaram a tornar este trabalho bem mais interessante.

Ao ParGO, pelos seminários interessantes e conferências divertidas.

Ao Artur de Oliveira, por ser o amigo mais engraçado que eu já tive. Ao Isaac Farias, pela amizade e discussões filosóficas. Ao Caíque Pimentel, Laudiana Gomes e Luiz Neto, pela longa amizade e por terem feito parte daquele que é, de longe, o melhor momento da minha vida.

Ao Professor José Wagner, quem me inspirou a seguir na matemática.

À Andrea Dantas, pelo suporte técnico nessa fase tão importante da minha carreira. Aos funcinários da Biblioteca da Matemática, em especial ao Erivan, pela disponibilidade e simpatia.

À CAPES, pelo apoio financeiro com a manutenção da bolsa de auxílio.

"Knowledge is a correspondence between idea and fact."

– Frank Plumpton Ramsey

ABSTRACT

A graph G is Ramsey for a pair of graphs (F_1, F_2) if in every 2-edge-colouring of G, one can find a monochromatic copy of F_1 with the first colour or a monochromatic copy of F_2 with the second colour. The random binomial graph $\mathbf{G}_{n,p}$ is a subgraph of K_n, the complete graph on n vertices, obtained choosing each edge of K_n independently at random with probability p to belong to $\mathbf{G}_{n,p}$. For a graph F, let $m_2(F)$ be the maximum of $d_2(F') = (e(F') - 1)/(v(F') - 2)$ over all the subgraphs $F' \subseteq F$ with $v(F') \geq 3$. If this maximum is reached for $F' = F$, then we say that F is 2-balanced. Furthermore, we say that F is strictly 2-balanced if $d_2(F) > d_2(F')$, for all proper subgraph F' of F with $v(F') \geq 3$. For a pair of graphs (F_1, F_2), let $m_2(F_1, F_2)$ be the maximum of $e(F_1')/(v(F_1') - 2 + 1/m_2(F_2))$ over all the subgraphs $F_1' \subseteq F_1$ with $v(F_1') \geq 3$. This dissertation aims to present a proof that for every pair of graphs (F_1, F_2) such that F_1 is 2-balanced and $m_2(F_1) > m_2(F_2) > 1$ or F_1 is strictly 2-balanced and $m_2(F_1) \geq m_2(F_2) > 1$, there exists a positive constant C for which asymptotically almost surely $\mathbf{G}_{n,p}$ is Ramsey for the pair (F_1, F_2), whenever that $p \geq Cn^{-1/m_2(F_1,F_2)}$. This result was conjectured by Kohayakawa and Kreuter in 1997 without the balancing condition over F_1. The proof of the main theorem must uses a recently developed technique known as hypergraph containers.

Keywords: Ramsey property, binomial random graph, threshold function.

RESUMO

Um grafo G é Ramsey para um par de grafos (F_1, F_2) se em toda 2-aresta-coloração de G for possível encontrar cópias monocromáticas de F_1 com a primeira cor ou cópias monocromáticas de F_2 com a segunda cor. O grafo aleatório binomial $\mathbf{G}_{n,p}$ é um subgrafo de K_n, o grafo completo com n vértices, obtido escolhendo cada aresta de K_n independentemente e aleatoriamente com probabilidade p para pertencer à $\mathbf{G}_{n,p}$. Para um grafo F, seja $m_2(F)$ o valor máximo de $d(F') = (e(F') - 1)/(v(F') - 2)$ dentre todos os subgrafos $F' \subseteq F$ com $v(F') \geq 3$. Se tal máximo é atingido por $F' = F$, então dizemos que F é 2-balanceado. Ademais, dizemos que F é estritamente 2-balanceado se $d_2(F) > d_2(F')$ para todo subgrafo próprio F' de F com $v(F') \geq 3$. Para um par de grafos (F_1, F_2), seja $m_2(F_1, F_2)$ o valor máximo de $e(F_1')/(v(F_1') - 2 + 1/m_2(F_2))$ dentre todos os subgrafos $F_1' \subseteq F_1$ com $v(F_1') \geq 3$. Esta dissertação objetiva-se em apresentar uma prova de que para todo par de grafos (F_1, F_2) tais que F_1 é 2-balanceado e $m_2(F_1) > m_2(F_2) > 1$ ou F_1 é estritamente 2-balanceado e $m_2(F_1) \geq m_2(F_2) > 1$, existe uma constante positiva C para o qual assimptoticamente quase certamente, $\mathbf{G}_{n,p}$ é Ramsey para o par (F_1, F_2), sempre que $p \geq Cn^{-1/m_2(F_1,F_2)}$. Este resultado foi conjeturado por Kohayakawa and Kreuter em 1997 sem a condição de balanceamento sobre F_1. A prova do principal teorema nesta dissertação deverá usar técnicas desenvolvidas recentementes e conhecidas como *hypergraph containers*.

Palavras-chave: propriedade de Ramsey, grafo aleatório binomial, função limiar.

CONTENTS

1 INTRODUCTION

The classical Ramsey's Theorem states that for every $k \geq 3$, there exists n such that every 2-edge-colouring of K_n (not necessarily a proper colouring) contains a copy of K_k with all the edges with the same colour. Putting it in other words, we can find monochromatic copies of K_k in any 2-edge-colouring of K_n, provided that n is big enough. We could naturally generalize this result for other graphs rather than complete graphs. In order to do so, let us introduce some definitions and notations.

Given graphs G and H, we say that a graph F is *Ramsey* for the pair (G, H) if every 2-edge-colouring of F contains a copy (not necessarily induced) of G with all the edges being of the first colour or a copy of H with all the edges being of the second colour. We write $F \to (G, H)$ to mean that F is Ramsey for the pair (G, H). If $G = H$ and $F \to (G, H)$, then we write $F \to G$ and we say that F is Ramsey for G. In this language, the Ramsey theorem states that for every $k \geq 3$, there exists some n for which $K_n \to K_k$. This implies that for every pair of graphs (G, H), there exists some n for which $K_n \to (G, H)$, since we may take $k = \max\{v(G), v(H)\}$. Furthermore, if $K_n \subseteq F$, then $F \to (G, H)$.

It is interesting to determine the smallest n for which $K_n \to (G, H)$, for a given pair of graphs (G, H). Such n is called *Ramsey number* of (G, H) and it is denoted by $r(G, H)$. In the case $G = H$, we write $r(G)$ only. Many results has been established about $r(G, H)$ (see Radziszowski (1994) for a survey). And yet many are unknown even for complete graphs. For instance, the best known bounds for $r(K_k)$ is

$$\frac{k}{2} \cdot 2^{(k+1)/2} \leq r(K_k) \leq k^{-c\frac{\log k}{\log\log k}} \binom{2k-2}{k-1},$$

for some positive constant c. The lower bound is due to Spencer (1975) while the upper bound is due to Conlon (2009). Any improvement in the bounds above would be of significant interest. See Conlon, Fox, and Sudakov (2015) for a survey on some recent developments related to $r(G)$.

Notice that if we have a graph F such that $F \to G$, then we must have $K_{v(F)} \to G$, since $F \subseteq K_{v(F)}$. Therefore $r(G) \leq v(F)$, for every F that it is Ramsey for G. This provides us another way of defining $r(G)$ as the minimum of $v(F)$ over all graphs F that it is Ramsey for G. So, we could also ask about the minimum of $f(F)$ over all graphs F that are Ramsey for G, for some real-valued graph invariant f. We call this minimum the *f-Ramsey number* of G and we denote it by $r_f(G)$. Much attention has been given to the case where $f(F)$ is the number of edges (see Erdős *et al.* (1978)), the chromatic number (see Burr, Erdős, and Lovász (1976)), and the maximum degree, $\Delta(F)$ (see Kinnersley, Milans, and West (2012)). In the case $f(F)$ is the clique number, $\omega(F)$ (the number of vertex in the largest complete subgraph contained in F), it was first asked by Erdős and Hajnal (1967) for which integers k we have $r_\omega(K_k) = k$ and it was proved by Folkman (1970) that we have this for all $k \geq 2$. Later, it was proved by Nešetřil and Rödl (1976) that $r_\omega(G) = \omega(G)$, for any graph G. This result was revisited by Rödl and Ruciński (1993, 1995) who proved that for a suitable choice of p, the random binomial graph $\mathbf{G}_{n,p}$ (see the next section for a definition of $\mathbf{G}_{n,p}$) is almost surely Ramsey for G and it does not have a clique of size greater than $\omega(G)$. We shall discuss this result in the next section.

1.1 Ramsey theory for random graphs

Let $\mathbf{G}_{n,p}$ be the Erdős-Renyi binomial model of a random graph. That is, $\mathbf{G}_{n,p}$ is a subgraph of K_n where each edge is chosen independently at random with probability p. Much research has been made in order to understand what kind of properties we can expect from $\mathbf{G}_{n,p}$. We shall briefly discuss this here. In order to do this, let us introduce some definitions.

A *property* $\mathcal{P}$ is understood as a family of graphs (the family of graphs that satisfy that property). For instance, we call *containment property* the family of graphs F that contains a fixed graph G. And we call *Ramsey property* the family of graphs F such that $F \to (G, H)$, for a fixed pair of graphs (G, H). A property $\mathcal{P}$ is *monotone increasing* if $F \in \mathcal{P}$ and $F \subseteq H$ implies that $H \in \mathcal{P}$. And $\mathcal{P}$ is a *monotone decreasing* property if $\overline{\mathcal{P}}$, the family of all graphs not belonging to $\mathcal{P}$, is monotone increasing. We say that a property is *monotone* if it is monotone increasing or monotone decreasing. For instance, both containment and Ramsey properties are monotone increasing properties.

We say that an event happens *with high probability* (w.h.p.) if the probability of have it happening tends to one as the order of the sample space grows to infinity (usually, the order of the sample space grows as the number of vertices of a random graph grows). A *threshold* for a property $\mathcal{P}$ in $\mathbf{G}_{n,p}$ is a function $p_0 : \mathbb{N} \to [0,1]$ for which the following holds:

$$\lim_{n\to\infty} \mathbb{P}\big[\mathbf{G}_{n,p(n)} \in \mathcal{P}\big] = \begin{cases} 0, & \text{if } p(n) \ll p_0(n); \\ 1, & \text{if } p(n) \gg p_0(n). \end{cases} \tag{1}$$

A threshold is a *strong threshold* if there are c and C positive constants for which

$$\lim_{n\to\infty} \mathbb{P}\big[\mathbf{G}_{n,p(n)} \in \mathcal{P}\big] = \begin{cases} 0, & \text{if } p(n) \leq cp_0(n); \\ 1, & \text{if } p(n) \geq Cp_0(n). \end{cases} \tag{2}$$

And a threshold is a *sharp threshold* if for all $\varepsilon > 0$, we have

$$\lim_{n\to\infty} \mathbb{P}\big[\mathbf{G}_{n,p(n)} \in \mathcal{P}\big] = \begin{cases} 0, & \text{if } p(n) \leq (1-\varepsilon)p_0(n); \\ 1, & \text{if } p(n) \geq (1+\varepsilon)p_0(n). \end{cases} \tag{3}$$

In any of the three kinds of threshold above, we call 0-*statement* the statement that

$$\lim_{n\to\infty} \mathbb{P}\big[\mathbf{G}_{n,p(n)} \in \mathcal{P}\big] = 0.$$

While the 1-*statement* corresponds to

$$\lim_{n\to\infty} \mathbb{P}\big[\mathbf{G}_{n,p(n)} \in \mathcal{P}\big] = 1.$$

Determining threshold for properties in $\mathbf{G}_{n,p}$ has been an interesting research subject. Bollobás and Thomason (1987) showed that any monotone property has a threshold function. Bollobás (1981) determined the threshold for the containment property. He showed the following.

Notation 1. For a graph F with at least one vertex, we define the *density* of F as

$$m(F) = \max\left\{\frac{e(F')}{v(F')} : F' \subseteq F, v(F') \geq 1\right\}.$$

Theorem 1.1 (Bollobás, 1981)**.** *Let F be an arbitrary graph with at leas one edge. Then, for $p = p(n)$, we have*

$$\lim_{n\to\infty} \mathbb{P}\big[F \subseteq \mathbf{G}_{n,p}\big] = \begin{cases} 0, & \textit{if } p \ll n^{-1/m(F)}; \\ 1, & \textit{if } p \gg n^{-1/m(F)}. \end{cases}$$

We write $F \to (G, H)^v$ if every 2-vertex-colouring of F contains a copy of G with all the vertices with the first colour or contains a copy of H with all the vertices with the second colour. This is the Ramsey property with respect to vertex colouring. Kreuter (1996) determined the threshold for the Ramsey property with respect to vertex colouring. He showed the following.

Notation 2. For a graph F with at least two vertices, let

$$m_1(F) = \max\left\{\frac{e(F')}{v(F')-1} : F' \subseteq F, v(F') \geq 2\right\}$$

be the 1-*density* of F.

Theorem 1.2 (Kreuter, 1996)**.** *Let F_1 and F_2 be graphs with at least one edge and suppose F_1 is not a matching and that $m_1(F_2) \geq m_1(F_1)$. Then there exist positive constants c and C such that, for $p = p(n)$,*

$$\lim_{n\to\infty} \mathbb{P}\big[\mathbf{G}_{n,p} \to (F_1, F_2)^v\big] = \begin{cases} 0, & \textit{if } p \leq cn^{-1/m_1(F_1,F_2)}; \\ 1, & \textit{if } p \geq Cn^{-1/m_1(F_1,F_2)}, \end{cases}$$

where

$$m_1(F_1, F_2) = \max\left\{\frac{e(F') + m_1(F_2)}{v(F')} : F' \subseteq F, v(F') \geq 1\right\}.$$

About the threshold for Ramsey properties with respect to edge colouring, Frankl and Rödl (1986) were the first to investigate them. The first remarkable result is due to Rödl and Ruciński(1993 1995) who determined the threshold for the symmetric Ramsey property on $\mathbf{G}_{n,p}$.

Before we state precisely the theorem of Rödl & Ruciński, let us give an intuition that suggest a lower bound for a possible threshold for the event that $\mathbf{G}_{n,p} \to F$. So, given a graph F with at least three vertices, we must determine a large p for which $\mathbf{G}_{n,p} \nrightarrow F$ almost surely. First, notice that the expected number of copies of F in $\mathbf{G}_{n,p}$ containing a given edge e of K_n is at most

$$2e(F) \cdot \frac{(v(F)-2)!}{\mathrm{aut}(F)} \cdot \binom{n-2}{v(F)-2} p^{e(F)-1} \leq n^{v(F)-2} p^{e(F)-1},$$

where $\mathrm{aut}(F)$ denotes the *number of automorphism* of F. Therefore, if we have $p = p_F = cn^{-(v(F)-2)/(e(F)-1)}$, then we can choose c small in order to have the bound above smaller than 1. This way, we expect no more than one copy of F for most of the edges of $\mathbf{G}_{n,p}$. Therefore, if we give any 2-edge-colouring of $\mathbf{G}_{n,p}$ and we have a monochromatic copy $\tilde{F}$ of F in $\mathbf{G}_{n,p}$ with this colouring, then in most cases, we can change the colour of an edge e of $\tilde{F}$ making $\tilde{F}$ no more a monochromatic copy and we will not be worried with creating a new monochromatic copy of F, since the only copy of F in $\mathbf{G}_{n,p}$ containing e is $\tilde{F}$. We can keep changing the colour of edges in $\mathbf{G}_{n,p}$ like above eliminating, thus, all the monochromatic copies of F in $\mathbf{G}_{n,p}$. With a careful analysis of the edges for which we can not change immediately its color, we may find a coloring of $\mathbf{G}_{n,p}$ which does not contain monochromatic copies of F. This way, we would have that $\mathbf{G}_{n,p} \nrightarrow F$.

Instead of eliminating the monochromatic copies of F, we could try to do the same as above but for a subgraph F' of F eliminating monochromatic copies of F', which in turn, ends up eliminating monochromatic copies of F as well. Therefore, if for some $F' \subseteq F$, we have

$$\frac{e(F)-1}{v(F)-2} \leq \frac{e(F')-1}{v(F')-2},$$

then it should be more profitable to eliminate monochromatic copies of F' instead of F, since we would have $p_F \leq p_{F'}$ and yet $\mathbf{G}_{n,p_{F'}} \nrightarrow F$. This suggest us to consider the following definition.

Notation 3. For a graph F on at least three vertices, we set $d_2(F) = (e(F)-1)/(v(F)-2)$. The 2-*density* of F, denoted by $m_2(F)$, is the number

$$m_2(F) = \max\{d_2(F') : F' \subseteq F, v(F') \geq 3\}.$$

We say F is 2-*balanced* if $m_2(F) = d_2(F)$, and *strictly* 2-*balanced* if in addition $m_2(F) > d_2(F')$, for every $F' \subset F$ with $v(F') \geq 3$.

The Theorem of Rödl & Ruciński states that the lower bound suggested above give us, indeed, the threshold for the symmetric Ramsey property on $\mathbf{G}_{n,p}$ for any graph

F that is not a forest.

Theorem 1.3 (Rödl & Ruciński, 1995)**.** *For every graph F with $m_2(F) > 1$, there exist positive constants c and C such that, for $p = p(n)$,*

$$\lim_{n\to\infty} \mathbb{P}\big[\mathbf{G}_{n,p} \to F\big] = \begin{cases} 0, & \text{if } p \le cn^{-1/m_2(F)}; \\ 1, & \text{if } p \ge Cn^{-1/m_2(F)}. \end{cases}$$

Nenadov and Steger (2016) gave a proof of Theorem 1.3 using the hypergraph container method. We present this proof with more details in Chapter 3.

About an asymmetric version, we have the following insight for a lower bound. We want to determine p for which $\mathbf{G}_{n,p} \nrightarrow (F_1, F_2)$, for a given pair of graphs (F_1, F_2) with $m_2(F_1) \ge m_2(F_2)$. We start colouring the edges of $\mathbf{G}_{n,p}$ that do not belong to a copy of F_1 with colour 1 and, then, we try to find a colouring of the remaining edges without creating monochromatic copies of F_1 and F_2 with the colour 1 and 2, respectively. We can bound the expected number of edges that belong to a copy of F_1 by an union bound as

$$m = \binom{n}{2} \cdot 2e(F_1) \cdot \frac{(v(F_1)-2)!}{\mathrm{aut}(F_1)} \cdot \binom{n-2}{v(F_1)-2} \cdot p^{e(F_1)} = O\big(n^{v(F_1)} p^{e(F_1)}\big),$$

Therefore, taking $p = p_{F_1} = cn^{-(v(F_1)-2+1/m_2(F_2))/e(F_1)}$, for some positive constant c, we must have $m = O\big(n^{2-1/m_2(F_2)}\big)$. Now, we hope that the distribution of those edges behaves like one of a random graph $\mathbf{G}_{n,q}$ with $q = c'n^{-1/m_2(F_2)}$, for a positive constant $c' = c'(c)$, as suggested by the value of m. As we discussed earlier when we were giving an intuition to suggest the lower bound for the threshold of the symmetric version, the expected number of copies of F_2 in $\mathbf{G}_{n,q}$ containing a fixed edge is bounded by a constant that can be as small as we want. Therefore, when c is small (so is c'), it should be easy to colour the edges in $\mathbf{G}_{n,q}$ avoiding copies of F_1 with colour 1 and copies of F_2 with colour 2 just by assign the colour 2 to most of the edges. This give us, heuristically, a colouring for the remaining edges of $\mathbf{G}_{n,p}$. Then we have a 2-edge-colouring of $\mathbf{G}_{n,p}$ with no copies of F_1 with the first colour and no copies of F_2 with the second colour.

Now, just like in the symmetric case, we can consider subgraphs F_1' of F_1 creating a colouring of $\mathbf{G}_{n,p}$ just like above, but avoiding copies of F_1' with first colour and copies of F_2 with the second colour. This way, we avoid monochromatic copies of F_1 with first colour as well. Therefore, we must choose $F_1' \subseteq F$ for which $p_{F_1'}$ is the greatest as possible, which it should be the one that maximizes the value of $e(F_1')/(v(F_1') - 2 + 1/m_2(F_2))$. This lead us to the following notation.

Notation 4. For two graphs F_1 and F_2, with $v(F_1) \ge 3$, we set

$$d_2(F_1, F_2) = e(F_1)/(v(F_1) - 2 + 1/m_2(F_2)).$$

The 2-*density* of the pair (F_1, F_2) is the number

$$m_2(F_1, F_2) = \max\{d_2(F_1, F_2) : F_1' \subseteq F_1, v(F_1') \geq 3\}.$$

We say that F_1 is *balanced* w.r.t. $m_2(F_1, F_2)$ if $m_2(F_1, F_2) = d_2(F_1, F_2)$, and *strictly balanced* w.r.t. $m_2(F_1, F_2)$ if in addition $m_2(F_1, F_2) > d_2(F_1', F_2)$ for all $F_1' \subset F_1$, with $v(F_1') \geq 3$. Notice that if $m_2(F_1) \geq m_2(F_2)$, then $m_2(F_1) \geq m_2(F_1, F_2) \geq m_2(F_2)$. In particular, $m_2(F, F) = m_2(F)$. And if $m_2(F_1) > m_2(F_2)$, then $m_2(F_1) > m_2(F_1, F_2) > m_2(F_2)$.

Following the insight given above, we hope that $p = cn^{-1/m_2(F_1,F_2)}$ give us a lower bound for the threshold for the event that $\mathbf{G}_{n,p} \to (F_1, F_2)$. It was conjectured by Kohayakawa and Kreuter (1997) that such p give us, in fact, the threshold for the asymmetric Ramsey property on $\mathbf{G}_{n,p}$.

Conjecture 1.4 (Kohayakawa & Kreuter, 1997)**.** *Let F_1 and F_2 be graphs with $m_2(F_1) \geq m_2(F_2) > 1$. Then there exist positive constants c and C such that, for $p = p(n)$,*

$$\lim_{n\to\infty} \mathbb{P}\big[\mathbf{G}_{n,p} \to (F_1, F_2)\big] = \begin{cases} 0, & \textit{if } p \leq Cn^{-1/m_2(F_1,F_2)}; \\ 1, & \textit{if } p \geq cn^{-1/m_2(F_1,F_2)}. \end{cases}$$

Kohayakawa and Kreuter (1997) proved the conjecture above for pairs of cycles (with distinct size) and they proved the upper bound for pairs (F, C_k), where C_k is a cycle of size k and F is a graph under some constraints. Precisely, they proved the following.

Theorem 1.5 (Kohaykawa & Kreuter, 1997)**.** *Let F be a 2-balanced graph and $k \geq 3$ be an integer such that $m_2(F) > m_2(C_k) = 1 + 1/(k-2)$. Then there exists a positive constant C such that, for $p = p(n) \geq Cn^{-1/m_2(F_1,F_2)}$,*

$$\lim_{n\to\infty} \mathbb{P}\big[\mathbf{G}_{n,p} \to (F, C_k)\big] = 1.$$

Furthermore, the same conclusion holds for a graph F and an integer $k \geq 3$ with $m_2(F_1) \geq m_2(C_k)$, provided that F is strictly 2-balanced. If in addition F is a cycle C_ℓ, then there is a positive constant c such that, for $p = p(n) \leq cn^{-1/m_2(F_1,F_2)}$,

$$\lim_{n\to\infty} \mathbb{P}\big[\mathbf{G}_{n,p} \to (C_\ell, C_k)\big] = 0.$$

Marciniszyn *et al.* (2009) proved the lower bound as in the Conjecture (1.4) for (F_1, F_2) being a pair of cliques. They observed that the upper bound follows from an important conjecture of Kohayakawa, Łuczak and Rödl known as KŁR conjecture (we shall state it precisely in Chapter 2).

Theorem 1.6 (Marciniszyn, Skokan, Spöhel & Steger, 2009)**.** *Let k_1 and k_2 be positive in-*

tegers with $k_1 \geq k_2 \geq 3$. Then there exist $c > 0$ such that, for $p = p(n) \leq cn^{-1/m_2(K_{k_1}, K_{k_2})}$,

$$\lim_{n\to\infty} \mathbb{P}\big[\mathbf{G}_{n,p} \to (K_{k_1}, K_{k_2})\big] = 0.$$

Furthermore, the KŁR conjecture implies that there exists $C > 0$ such that, for $p = p(n) \geq Cn^{-1/m_2(K_{k_1}, K_{k_2})}$,

$$\lim_{n\to\infty} \mathbb{P}\big[\mathbf{G}_{n,p} \to (K_{k_1}, K_{k_2})\big] = 1.$$

In fact, the observation from Marciniszyn *et al.* (2009) was that the KŁR conjecture implies in the upper bound in Conjecture (1.4) for pairs of graph (F_1, F_2) with $m_2(F_1) \geq m_2(F_2) > 1$ and with an additional constraint that F_1 is 2-balanced. The KŁR conjecture was proved only recently by Balogh, Morris, and Samotij (2015) and Saxton and Thomason (2015), independently, using the *hypergraph containers method* (which is addressed in Chapter 2). Therefore, now that the KŁR conjecture is proved, the following result holds.

Theorem 1.7. *Let F_1 and F_2 be graphs with $m_2(F_1) > m_2(F_2) > 1$ such that F_1 is 2-balanced. Then there exists a positive constant C such that for $p = p(n) \geq Cn^{-1/m_2(F_1,F_2)}$,*

$$\lim_{n\to\infty} \mathbb{P}\big[\mathbf{G}_{n,p} \to (F_1, F_2)\big] = 1.$$

Furthermore, the same conclusion holds for pairs of graphs (F_1, F_2) with $m_2(F_1) \geq m_2(F_2) > 1$ provided that F_1 is strictly 2-balanced and.

Just to illustrate how deep is the theorem above, let see some examples of pairs of graphs (F_1, F_2) for which Theorem 1.7 can be applied:

1. Pairs of cycles (C_s, C_t) with $s \geq t \geq 3$.
2. Pairs of graphs $(K_{a,b}, G)$ with $G \subseteq K_{a,b}$, $v(G) \geq 3$ and $m_2(G) > 1$ (this includes the pairs of complete bipartite graphs $(K_{a,b}, K_{s,t})$ with $a \geq s$, $b \geq t$ and $s + t \geq 3$).
3. Pairs of graph (K_s, G) with $G \subseteq K_s$, $v(G) \geq 3$ and $m_2(G) > 1$ (this includes the pairs of complete graph (K_s, K_t) with $s \geq t \geq 3$).

It is worth mentioning the following result from Kohayakawa, Schacht, and Spöhel (2014) which was proved without using any sparse-regularity technique, in special, without using the KŁR conjecture.

Theorem 1.8 (Kohaykawa, Schacht & Spöhel, 2014)**.** *Let F_1 and F_2 be graphs with $m_2(F_1) > m_2(F_2) > 1$ such that F_1 is strictly balanced w.r.t. $m_2(F_1, F_2)$. Then there exists a positive constant C such that, for $p = p(n) \geq Cn^{-1/m_2(F_1,F_2)}$,*

$$\lim_{n\to\infty} \mathbb{P}\big[\mathbf{G}_{n,p} \to (F_1, F_2)\big] = 1.$$

One of the aims of this dissertation is to give a complete proof of Theorem 1.7. To do so, we first discuss in Chapter 2 some techniques (such as sparse version for the regularity lemma, KŁR conjecture and the hypergraph container method) that will be useful in our proofs. Then in Chapter 3, we present the proof of the symmetric version

(Theorem 1.3) using the hypergraph containers method. Following, we present the proof of the asymmetric version (Theorem 1.7) using the KŁR conjecture in Chapter 4. For the last, we present a proof of the KŁR conjecture using the hypergraph containers method in Chapter 5.

2 TECHNIQUES

In this section, we make an exposition of those well established results in the literature which will be essential tools in our proofs along this text.

2.1 Regularity lemma for sparse graphs

Let G be a graph and U and V be disjoint subsets of $V(G)$. The *density* $d(U,V)$ of the pair (U,V) is the number $|U||V|/e(U,V)$, where $e(U,V)$ is the number of edges between U and V. We say that (U,V) is an *ε-regular pair* if, for every $U' \subseteq U$ and $V' \subseteq V$ with $|U'| \geq \varepsilon|U|$ and $|V'| \geq \varepsilon|V|$, we have

$$|d(U,V) - d(U',V')| \leq \varepsilon.$$

A partition $V_1, \ldots, V_k$ of $V(G)$ is an *equipartition* if $|V_1| \leq |V_2| \leq \cdots \leq |V_k| \leq |V_1| + 1$. If at least $(1-\varepsilon)\binom{k}{2}$ pairs (V_i, V_j), with $1 \leq i, j \leq k$, are ε-regular, then we say that the equipartition is *ε-regular*.

Roughly speaking, the celebrated Szemerédi's regularity lemma states that every large enough graph admits an ε-regular equipartition with few parts. Szemerédi (1975) first applied this lemma to prove a conjecture from Erdős and Turán (1936) about arithmetic progressions in dense subsets of $\mathbb{N}$ – nowadays this result is known as Szemerédi's Theorem and it has fostered a large amount of research in additive number theory and related areas. The lemma itself became one of the most important tool in modern combinatorics being already applied to many other problems (see Komlós *et al.* (2002) for a survey on Szemerédi's regularity lemma and some of its classical applications). Here is the precise statement of the lemma. Here we prefer to refer to it as a theorem.

Theorem 2.1 (Szemerédi's Regularity Lemma, 1975)**.** *For every $\varepsilon > 0$ and every positive integer k_0, there exists a positive integer K such that every graph G with at least K vertices has an ε-regular equipartition $V_1, \ldots, V_k$ of $V(G)$ with $k_0 \leq k \leq K$.*

It is worth noting that for graphs with $o(n^2)$ edges (we say that those graphs are *sparse graphs*), the Szemerédi's regularity lemma tell us nothing, since the error part encapsulated by the regularity has $o(n^2)$ edges and therefore, we could have the whole graph covered by the error part. Kohayakawa (1997) and Rödl, independently, observed that the regularity lemma can still be useful for sparse graphs under some reasonable conditions. In the following, we develop some definitions in order to state precisely the sparse version of Szemerédi's regularity lemma.

We say that a pair (U,V) of disjoint subsets of $V(G)$ is *(ε,p)-regular* if, for every $U' \subseteq U$ and $V' \subseteq V$ with $|U'| \geq \varepsilon|U|$ and $|V'| \geq \varepsilon|V|$, we have

$$|d(U,V) - d(U',V')| \leq \varepsilon p.$$

An equipartition $V_1, \dots, V_k$ of $V(G)$ is (ε, p)-*regular* if at least $(1-\varepsilon)\binom{k}{2}$ pairs (V_i, V_j), with $1 \leq i, j \leq k$, are (ε, p) regular. A graph G is (η, D, p)-*upper-uniform* if, for all disjoint subsets U and V of $V(G)$ with $|U|, |V| \geq \eta |V(G)|$, we have $d(U, V) \leq Dp$.

Now, we can state the sparse version of Szemerédi's regularity lemma due to Kohayakawa and Rödl.

Theorem 2.2 (Kohayakawa & Rödl's Sparse Regularity Lemma, 1997)**.** *For every positive constants ε and D and every positive integer k_0, there exist $\eta > 0$ and a positive integer K_0 such that for any $p \in [0, 1]$, every (η, D, p)-upper-uniform graph G with at least k_0 vertices has an (ε, p)-regular equipartition $V_1, \dots, V_k$ of $V(G)$ with $k_0 \leq k \leq K_0$.*

The reader is referred to Gerke and Steger (2005) for some classical applications of Theorem 2.2 in extremal graph theory. It is worth mentioning that Scott (2011) proved a sparse version of Szemerédi regularity lemma in which the upper-uniformity assumption is dropped. In the following, an (ε)-regular equipartition of a graph G is an (ε, p)-regular equipartition of G with $p = e(G)/\binom{n}{2}$. Here is the precise result established by Scott.

Theorem 2.3 (Scott's Sparse Regularity Lemma, 2011)**.** *For every $\varepsilon > 0$ and every positive integer k_0, there exists a positive integer K such that every graph G with at least k_0 vertices has an (ε)-regular equipartition $V_1, \dots, V_k$ of $V(G)$, with $k_0 \leq k \leq K$.*

In the following, we show how the (ε, p)-regularity is related to random bipartite graphs. This example will be important for the subsequent examples we have in this text.

Example 1. Consider a random bipartite graph $\mathbf{G}_m[U, V]$ chosen uniformly at random from all the bipartite graphs on (U, V) with m edges. Notice that an edge uv with $u \in U$ and $v \in V$ belongs to $\mathbf{G}_m[U, V]$ with probability $p = m/n^2 = d(U, V)$. Now, let ε be a positive constant. We want to bound the probability of $\mathbf{G}_m[U, V]$ being (ε, p)-regular. For this purpose, fix $U' \subseteq U$ and $V' \subseteq V$ with $|U'| \geq \varepsilon |U|$ and $|V'| \geq \varepsilon |V|$. Let X be the random variable $e(U', V')$. Notice that X has distribution Binomial$(|U'||V'|, p)$ with mean $\mu = p|U'||V'| \geq \varepsilon^2 p |U||V|$. Thus, by the Chernoff's inequality (see Corollary B.4), for $\varepsilon \leq 3/2$, we have

$$\begin{aligned}
\mathbb{P}\big[|d(U', V)' - d(U, V)| \geq \varepsilon p\big] &= \mathbb{P}\big[\,|\, e(U', V') - p|U'||V'|\,| \geq \varepsilon p |U'||V'|\,\big] \\
&= \mathbb{P}\big[|X - \mu| \geq \varepsilon \mu\big] \\
&\leq 2 \cdot \exp\left\{-\frac{\varepsilon^2}{3}\mu\right\} \\
&\leq 2 \cdot \exp\left\{-\frac{\varepsilon^4}{3} p |U||V|\right\}.
\end{aligned}$$

Now, by applying the union bound for all the choices of U' and V',

$$\mathbb{P}\big[\mathbf{G}_m[U, V] \text{ is not } (\varepsilon, p)\text{-regular}\big] \leq 2^{|U|+|V|+2} \exp\left\{-\frac{\varepsilon^4}{3} p |U||V|\right\}.$$

In particular, if $|U| = |V| = n$ and if $p \geq Cn^{-1}$, for some positive constant $C = C(\varepsilon)$

big enough, the right hand side of the last inequality goes to zero when n goes to infinity and, thus, $\mathbf{G}_m[U, V]$ is w.h.p. (ε, p)-regular with density p. $\square$

2.2 Embedding lemma and KŁR conjecture

In most applications of the Szemerédi's regularity lemma, we apply an embedding lemma or even a counting lemma together. An embedding lemma is a result that explores the structure of a regular equipartition in order to ensure the existence of copies of a certain graph. A counting lemma, meanwhile, does the same to count the number of such copies. Roughly speaking, the embedding lemma states that if we do a *blow-up* of a graph H, meaning to replace its vertices by large independent sets and its edges by ε-regular bipartite graphs with positive density, then we can find a copy of H in this blown-up graph. The counting lemma ensures that we can find as many copies of H as we would find in a random blow-up of H with the same density (that is, a blow-up just like before, but instead of adding ε-regular bipartite graphs, we add random bipartite graphs with the same density). In the following, we formalize this.

Let H be a graph on $\{1, \ldots, v(H)\}$. Consider $\varepsilon > 0$, $p \in [0, 1]$, and $m \leq n^2$ positive integers. Let $V = V_1 \cup \ldots \cup V_{v(H)}$ be a disjoint union of independent sets of size n. For each $ij \in E(H)$, we add m edges between the pair (V_i, V_j) in a such way that (V_i, V_j) is an (ε, p)-regular pair. Let $G = (V, E)$ be the resulting graph. We denote by $\mathcal{G}\left(H, n, m, p, \varepsilon\right)$ the collection of all graphs obtained in this way. A canonical copy of H in $G \in \mathcal{G}\left(H, n, m, p, \varepsilon\right)$ is a copy of H in G with exactly one vertex in each V_i. We denote by $\mathcal{G}^*\left(H, n, m, p, \varepsilon\right)$ the set of all graphs $G \in \mathcal{G}\left(H, n, m, p, \varepsilon\right)$ that do not contain a canonical copy of H.

Example 2. Let $\mathbf{G}[H, n, m]$ be the random blow-up of H obtained replacing each vertex v_i of H by an independent set V_i with n vertices and each edge v_iv_j of H by a random bipartite graph $\mathbf{G}_m[V_i, V_j]$ (see Example 1). Let $p = m/n^2$. As was shown in Example 1, each of those bipartite random graphs $\mathbf{G}_m[V_i, V_j]$ are, w.h.p., (ε, p)-regular provided that $p \geq Cn^{-1}$, for some positive constant C. Then for every $\varepsilon > 0$, w.h.p., $\mathbf{G}[H, n, p]$ belongs to $\mathcal{G}\left(H, n, m, p, \varepsilon\right)$ if $p \geq Cn^{-1}$. $\square$

The following theorem is what is known as embedding lemma (see Komlós *et al.* (2002)).

Theorem 2.4 (The embedding lemma)**.** *For every graph H and every positive d, there exist a positive ε and an integer n_0 such that for every n and m with $n \geq n_0$ and $m \geq dn^2$, every graph $G \in \mathcal{G}\left(H, n, m, 1, \varepsilon\right)$ contains a canonical copy of H.*

The counting lemma associated to the embedding lemma above states that there are $(d^{e(H)} + o(1))n^{v(H)}$ canonical copies of H in a such graph G as above, for $d = m/n^2$. Notice that the expect number of canonical copies of H in $\mathbf{G}[H, n, m]$ is $d^{e(F)}n^{v(H)}$. Therefore, the counting lemma states that we should find roughly the same number of canonical copies of H in a graph $G \in \mathcal{G}\left(H, n, m, 1, \varepsilon\right)$ as in the random blow-up $\mathbf{G}[H, n, m]$.

Theorem 2.5 (Counting lemma)**.** *For every graph H and every positive $\delta > 0$, there exist a positive ε and an integer n_0 such that for every n and m with $n \geq n_0$ and $d = m/n^2 > \delta$, every graph $G \in \mathcal{G}(H, n, m, 1, \varepsilon)$ contains $\left(d^{e(H)} \pm \delta\right) n^{v(H)}$ canonical copies of H.*

We would like to have a sparse version of the embedding and counting lemma. That is, we would like to consider $p = p(n)$ being any number in $[0, 1]$ tending to zero, as n grows to infinity, instead of just having p being a constant. Furthermore, the conclusion should be that every graph $G \in \mathcal{G}(H, n, m, p, \varepsilon)$ contains $\left(d^{e(H)} \pm \delta\right) p^{e(H)} n^{v(H)}$ canonical copies of H. However, this does not happen if p is too small, as shown in the following example.

Example 3. Let $\mathbf{G}[H, n, m]$ be the random graph defined in the Example 2. We saw that for any $\varepsilon > 0$, w.h.p., $\mathbf{G}[H, n, m]$ belongs to $\mathcal{G}(H, n, m, p, \varepsilon/2)$, for $p = m/n^2$. Let us put $p = cn^{-\frac{v(H)-2}{e(H)-1}}$, for some small positive constant c. Notice that we have $p \gg n^{-1}$ if $(e(H)-1)/(v(H)-2) > 1$; so let us say that this is the case. The expected number of canonical copies of H in $\mathbf{G}[H, n, m]$ is $n^{v(H)} p^{e(H)} = n^{v(H)-2} p^{e(H)-1} m = c^{e(H)-1} m$. Therefore, there is a graph $G' \in \mathcal{G}(H, n, m, p, \varepsilon/2)$ with at most $c^{e(H)-1} m$ canonical copies of H. Let $\delta = c^{e(H)-1}/e(H)$ and take $c = c(H, \varepsilon)$ small enough in order to have $\delta \leq \varepsilon^3/4$. This way, the number of canonical copies of H in G' is $e(H)\delta m$. So we can chose δm edges from each bipartite subgraph (V_i, V_j) of G', with $v_i v_j \in E(H)$, in a way that all canonical copies of H contain at least one of those edges. Then, by deleting those edges, we obtain a graph G with $(1-\delta)m$ edges between (V_i, V_J) with no canonical copies of H. Furthermore, we have that (V_i, V_j) is an (ε, p)-regular pair in G. In fact, if $U \subseteq V_i$ and $V \subseteq V_j$ are such that $|U|, |V| \geq \varepsilon n$, then

$$\begin{aligned}
|d_G(U, V) - d_G(V_i, V_j)| &= \left| \frac{e_G(U, V)}{|U||V|} - \frac{(1-\delta)m}{n^2} \right| \\
&\leq \left| \frac{e_{G'}(U, V)}{|U||V|} - \frac{m}{n^2} \right| + \left| \frac{e_{G'}(U, V) - e_G(U, V)}{|U||V|} - \frac{\delta m}{n^2} \right| \\
&\leq |d_{G'}(U, V) - d_{G'}(V_i, V_j)| + \frac{\delta m}{|U||V|} + \frac{\delta m}{n^2} \\
&\leq \frac{\varepsilon}{2} p + \frac{\delta m}{\varepsilon^2 n^2} + \frac{\delta m}{n^2} \\
&\leq \frac{\varepsilon}{2} p + \frac{2\delta}{\varepsilon^2} p \\
&\leq \frac{\varepsilon}{2} p + \frac{\varepsilon}{2} p = \varepsilon p,
\end{aligned}$$

where we use the fact that (V_i, V_j) is also $(\varepsilon/2, p)$ regular in G'. Therefore, G belongs to $\mathcal{G}(H, n, (1-\delta)m, p, \varepsilon)$ and has no canonical copies of H. In other words, G belongs to $\mathcal{G}^*(H, n, (1-\delta)m, p, \varepsilon)$. $\square$

Therefore, if we expect to have a sparse version of an embedding lemma, a necessary condition, as shown by the construction above, is to have $p \geq Cn^{-\frac{v(H)-2}{e(H)-1}}$, for some positive constant $C = C(H)$, whenever $\frac{e(H)-1}{v(H)-2} > 1$. Since removing all canonical copies of a subgraph $H' \subset H$ would still remove all the canonical copies of H, we actually need to take $p \geq Cn^{-1/m_2(H)}$, whenever $m_2(H) > 1$. However, as the next example illustrates, we still have another fundamental difficulty in determining a sparse embedding lemma.

For a graph G, we denote by G^n the *complete blow-up* of G of order n that is the graph obtained from G replacing each vertex v_i of G by an independent set V_i with n elements and adding a complete bipartite graph on (V_i, V_j) whenever that $v_i v_j \in E(G)$.

Example 4. Given N and p with $p(N) = o(1)$ and $p \gg N^{-1/m_2(H)}$, take $n = (c^{-1}p)^{-m_2(H)}$, where c is a small positive constant. Notice that $N \gg n$ and since p goes to zero as function of N, we have that n goes to infinity as a function of N. Furthermore, our choice of n implies that $p = cn^{-1/m_2(H)}$. We know from Example 3 that if c is small enough, then there exists $\tilde{G} \in \mathcal{G}(H, n, m, p, \varepsilon)$, for $m = pn^2$, with no canonical copy of H. Now, take $G = \tilde{G}^{\frac{N}{n}}$ the complete blow-up of $\tilde{G}$ of order N/n. So we have $G \in \mathcal{G}(H, N, M, p, \varepsilon)$, with $M = \left(\frac{N}{n}\right)^2 \cdot m = pN^2$. However, G belongs to $\mathcal{G}^*(H, N, M, p, \varepsilon)$, since it has no canonical copies of H. □

Despite of this, it is plausible to hope that those examples for which a sparse embedding lemma fails are rare. This was conjectured by Kohayakawa, Łuczak, and Rödl (1997) and it became known as the famous KŁR conjecture. Several special cases of this conjecture were verified over the years. In special, a random version was established by Conlon *et al.* (2014). But only recently, with the development of the technique known as *hypergraph containers method* due to Balogh, Morris, and Samotij (2015) and Saxton and Thomason (2015) a complete proof of the KŁR conjecture is known. Here is the precise statement of the conjecture.

Theorem 2.6 (The KŁR conjecture)**.** *For every graph H and every positive β, there exist positive constants C, n_0, and ε such that the following holds. For every $n \in \mathbb{N}$ with $n \geq n_0$ and $m \in \mathbb{N}$ with $m \geq Cn^{2-1/m_2(H)}$,*

$$\left|\mathcal{G}^*\left(H, n, m, m/n^2, \varepsilon\right)\right| \leq \beta^m \binom{n^2}{m}^{e(H)}.$$

The KŁR conjecture has been already applied in many problems even before it has been completely established. For instance, the random version of Turán's theorem was known to follow from the KŁR conjecture (see Kohayakawa, Łuczak, and Rödl (1997) and Gerke and Steger (2005)). Recently, Conlon and Gowers (2016) proved the random version of Turán's theorem for graphs under a certain balancing condition without using the KŁR conjecture, while Schacht (2016) proved it for any graph. Also, an upper bound for the Ramsey property for random graphs conjectured by Kohayakawa and Kreuter (1997) (Conjecture 1.4) was known to follow from the KŁR conjecture under some balancing condition (see Marciniszyn *et al.* (2009) and Theorem 1.7), while Kohayakawa, Schacht, and Spöhel (2014) proved this upper bound without using the KŁR conjecture, though under another balancing condition (see Theorem 1.8). We give the proof of Theorem 1.7 using the KŁR conjecture in Chapter 4.

2.3 Hypergraph containers

A hypergraph is a pair $\mathcal{H} = (V(\mathcal{H}), E(\mathcal{H}))$ where $E(\mathcal{H}) \subseteq \mathcal{P}(V(\mathcal{H}))$. We say that the elements in $V(\mathcal{H})$ are the *vertices* of $\mathcal{H}$ and $V(\mathcal{H})$ is the *vertex set* of $\mathcal{H}$. The elements in $E(\mathcal{H})$ are called *hyperedges* of $\mathcal{H}$ and the set $E(\mathcal{H})$ is the *hyperedge set* of $\mathcal{H}$. The number of vertices and hyperedges in $\mathcal{H}$ are denoted by $v(\mathcal{H})$ and $e(\mathcal{H})$, respectively. The hypergraph $\mathcal{H}$ is k*-uniform* if every hyperedge of $\mathcal{H}$ has exactly k vertices.

Given $A \subseteq V(\mathcal{H})$, the *hypergraph induced by* A is the hypergraph $\mathcal{H}[A] = (A, E')$, where $E' = \{e \in E : e \subseteq A\}$. A subset $I \subseteq V(\mathcal{H})$ of vertices is said *independent* if $E(\mathcal{H}[I]) = \emptyset$, that is, if there is no hyperedge of $\mathcal{H}$ contained in I. We denote by $\mathcal{I}(\mathcal{H})$ the collection of all independent sets in $\mathcal{H}$. The number of independent sets in $\mathcal{H}$ is denoted by $i(\mathcal{H})$. And the maximum size of an independent set in $\mathcal{H}$ is denoted by $\alpha(\mathcal{H})$. A trivial relation between the last two definend parameters is that $i(\mathcal{H}) \geq 2^{\alpha(\mathcal{H})}$, since any subset of an independent set is still a independent set.

Many extremal problems in combinatorics can be reduced to the analysis of independent sets in hypergraphs. This is naturally done by considering the hyperedges of a hypergraph as the set of elements which generates some forbidden configuration. Let us illustrate this with an example.

Example 5. Given a graph F, let $\mathcal{H}_{n,F}$ be the hypergraph (V, E) with $V = E(K_n)$ and $E = \{E(F') : F' \subseteq K_n \text{is a copy of F}\}$. Therefore, $\mathcal{H}$ is an $e(F)$-uniform hypergraph with $\binom{n}{2}$ vertices and with

$$\frac{v(F)!}{\mathrm{aut}(F)}\binom{n}{v(F)}$$

edges. An independent set I in $\mathcal{H}_{n,F}$ corresponds to the edge set of a graph $G \subseteq K_n$ which is F*-free* (that is, G has no subgraph isomorphic to F). Therefore, $\alpha(\mathcal{H}_{n,F})$ is equal to $\mathrm{ex}(n, F)$, the maximum number of edges in a F-free graph on n vertices. A classical result due to Erdős and Stone (1946) (see also Bollobás (2013)) states that

$$\mathrm{ex}(n, F) = \left(1 - \frac{1}{\chi(F) - 1} + o(1)\right)\binom{n}{2},$$

where $\chi(F)$ is the chromatic number of F. Furthermore, any extremal graph is isomorphic to the *Turán graph* $T(n, \chi(F))$, the graph obtained by partitioning a set on n vertices into $\chi(F)$ subsets, with size as equal as possible, and connecting two vertices by an edge if, and only if, they belong to different such subsets.

Let $\pi(F)$ be the *Turán density* of F, which is the number

$$\pi(F) = \lim_{n\to\infty} \mathrm{ex}(n, F)\binom{n}{2} = 1 - \frac{1}{\chi(F) - 1}.$$

An important result of Erdős and Simonovits (1983) implies that, for every $\delta > 0$, there exist $\varepsilon > 0$ and n_0 such that for $n \geq n_0$, if G is a graph on n vertices with $e(G) \geq (\pi(F)+\delta)\binom{n}{2}$, then G has at least $\varepsilon n^{v(F)}$ many copies of F. Back to $\mathcal{H}_{n,F}$, this says that if $A \subseteq V(\mathcal{H}_{n,F})$ is such that $|A| \geq (\pi(F)+\delta)v(\mathcal{H}_{n,F})$, then $e(\mathcal{H}_{n,F}[A]) \geq \varepsilon e(\mathcal{H})$. This result is what we call a *supersaturation* for $\mathcal{H}_{n,F}$.

Now, $i(\mathcal{H}_{n,F})$ corresponds to the number of F-free graphs on n vertices (which we are going to denote it by $f(n,F)$, though there is no a standard notation for that). Notice that $f(n,F) \geq 2^{\mathrm{ex}(n,F)}$, since any subgraph of a F-free graph is still F-free. Erdös, Frankl, and Rödl (1986) proved that

$$f(n,F) = 2^{\mathrm{ex}(n,F)+o(n^2)}.$$

Balogh, Bollobás, and Simonovits (2004) improved this result by showing that there is a positive constant ε which depends only on F such that

$$f(n,F) = 2^{\mathrm{ex}(n,F)+O(n^{2-\varepsilon})}.$$

□

The general ideal is to model forbidden configuration as independent sets. Therefore, we would like to have a method which allows us to describe approximately the independent sets in a hypergraph. In this direction, the containers method developed by Balogh, Morris, and Samotij (2015) and, independently, by Saxton and Thomason (2015) has excelled. The containers method goes back to a work of Kleitman and Winston (1982), where they proved that $f(n,C_4) = 2^{O(ex(n,C_4))}$. The idea behind the method relies on determining a *small family* $\mathcal{C}$ of subsets of $V(\mathcal{H})$ which are *almost independent* and such that $\mathcal{C}$ forms a *collection of containers* for $\mathcal{I}(\mathcal{H})$.

To be a little bit more precise (this should be just a warmout for the actual container method), let ε be (as always) a small positive constant. Then $\mathcal{C}$ must be a family of subsets of $V(\mathcal{H})$ satisfying the following:

(i) $\mathcal{C}$ has at most $2^{\varepsilon v(\mathcal{H})}$ elements;

(ii) for every $C \in \mathcal{C}$, we have $e(\mathcal{H}[C]) < \varepsilon e(\mathcal{H})$;

(iii) for every $I \in \mathcal{I}(\mathcal{H})$, there exists $C \in \mathcal{C}$ such that $I \subseteq C$.

These three items above are what we meant by *small family*, *almost independent* and *collection of containers*, respectively. We call a family $\mathcal{C}$ like above a *good collection of containers*xGood collection of containers for $\mathcal{I}(\mathcal{H})$ and ε.

Notice that it is straightforward to obtain a family $\mathcal{C}$ satisfying any two among the three items above. For instance, for items (i) and (ii), take $\mathcal{C} = \emptyset$; for items (i) and (iii), take $\mathcal{C} = V(\mathcal{H})$; and for items (ii) and (iii), take $\mathcal{C} = \mathcal{I}(\mathcal{H})$. And the last example also works for the three items at the same time if we have $i(\mathcal{H}) \leq 2^{\varepsilon v(\mathcal{H})}$. We could also take $\mathcal{C}$ to be the set of maximal independent sets in $\mathcal{H}$ in order to have the three conditions above. Of course, this family satisfy (ii) and (iii). However, it is not always true that such family will satisfy (i). For example, the graph $(n/2)K_2$, for n even, has $2^{n/2}$ maximal independent sets.

Therefore, the hardness on finding a good collection of containers relies on having the three items at the same time for a hypergraph with many (maximal) independent sets. This observation suggest us to consider hypergraph which are not so dense. Let us see how a good collection of containers could be applied in an extremal problem like that one we discussed in Example 5.

Example 6. Let $\mathcal{H}_n$ be a sequence of hypergraphs (in Example 5, $\mathcal{H}_n = \mathcal{H}_{n,F}$) and let us say that $v(\mathcal{H}_n)$ goes to infinity as n. Let $\mathcal{C}_n$ be a good collection of containers for $\mathcal{I}(\mathcal{H}_n)$ and a fixed $\varepsilon > 0$. We can bound $i(\mathcal{H}_n)$ by

$$\begin{aligned} i(\mathcal{H}_n) &\leq \sum_{C \in \mathcal{C}_n} i(\mathcal{H}_n[C]) \\ &\leq 2^{\varepsilon v(\mathcal{H}_n)} \max\{i(\mathcal{H}_n[C]) : C \in \mathcal{C}_n\} \\ &\leq 2^{\varepsilon v(\mathcal{H}_n) + \max\{|C| : C \in \mathcal{C}_n\}}. \end{aligned}$$

Actually, so far we only used items (i) and (iii) from the good collection of containers $\mathcal{C}_n$ to get to the inequality above. Item (ii) becomes useful when we have an additional property on $\mathcal{H}_n$: the *supersaturation*. A supersaturation is, in general, a statement ensuring that for some constant $\pi \in [0,1]$ (see $\pi(F)$ in Example 5), the following holds: for all $\delta > 0$, there is an $\varepsilon > 0$ and n_0 such that for $n \geq n_0$, if $A \subseteq V(\mathcal{H}_n)$ is such that $|A| \geq (\pi + \delta) v(\mathcal{H}_n)$, then $e(\mathcal{H}_n[A]) \geq \varepsilon e(\mathcal{H}_n)$. Therefore, since $C \in \mathcal{C}_n$ is such that $e(\mathcal{H}_n[C]) < \varepsilon e(\mathcal{H}_n)$, we must have $|C| < (\pi + \delta) v(\mathcal{H}_n)$. Thus, we can improve the previous bound on $i(\mathcal{H}_n)$ by

$$i(\mathcal{H}_n) \leq 2^{(\pi + \delta + \varepsilon) v(\mathcal{H}_n)}.$$

Since ε and δ are small constants, the bound above give us that $i(\mathcal{H}_n) \leq 2^{(\pi + o(1)) v(\mathcal{H}_n)}$ (in Example 5, it give us that $f(n, F) \leq 2^{\mathrm{ex}(n,F) + o(n^2)}$). In general, we can guarantee that for some π, we have $i(\mathcal{H}_n) \geq 2^{\pi v(\mathcal{H}_n)}$, by simply taking all the subsets of an independent set of sizer $\pi v(\mathcal{H}_n)$ (in Example 5, it corresponds to taking all the subgraphs of the Turán graph $T(n, \chi(F))$). $\square$

Therefore, the method relies on finding a good collection of containers and in most of the cases, it is useful when we have a supersaturation result together. In the following, we state the hypergraph container theorem due to Balogh, Morris, and Samotij (2015) which establishes a good collection of containers for k-uniform hypergraphs with some control under the degree distribution and we give a proof for the container theorem for F-free graphs. Before that, let us introduce some definitions.

Let $\mathcal{H}$ be a hypergraph. A family $\mathcal{F}$ of subsets of $V(\mathcal{H})$ is *increasing* in $\mathcal{H}$ if for every $A, B \subseteq V(\mathcal{H})$ with $A \in \mathcal{F}$ and $A \subseteq B$ implies that $B \in \mathcal{F}$. So, let $\mathcal{F}$ be an increasing family in $\mathcal{H}$ and let ε be a positive constant. We say that $\mathcal{H}$ is $(\mathcal{F}, \varepsilon)$-*dense* if $e(\mathcal{H}[A]) \geq \varepsilon e(\mathcal{H})$, for every $A \in \mathcal{F}$.

For a subset $T \subseteq V(\mathcal{H})$, we define the *degree* of T in $\mathcal{H}$ as

$$\deg_{\mathcal{H}}(T) = |\{E \in E(\mathcal{H}) : T \subseteq E\}|.$$

And the *maximum* ℓ*-degree* of $\mathcal{H}$ is defined as

$$\Delta_\ell(\mathcal{H}) = \max\{\deg_H(T) : T \subset V(\mathcal{H}) \text{ and } |T| = \ell\}.$$

The containers lemma roughly states that if a k uniform hypergraph $\mathcal{H}$ is $(\mathcal{F}, \varepsilon)$-dense and it has the edge distribution controlled by certain natural bound, then $\mathcal{I}(\mathcal{H})$ can be partitioned into few parts in a way that all the independent sets in the same part are essentially contained in a single set $A \notin \mathcal{F}$. Here is the precise statement of it.

Theorem 2.7 (Hypergraph Containers Theorem)**.** *For every $k \in \mathbb{N}$ and all positive c and ε, there exists a positive constant C such that the following holds. Let $\mathcal{H}$ be a k-uniform hypergraph and let $\mathcal{F} \subseteq \mathcal{P}(V(\mathcal{H}))$ be an increasing family of sets such that $|A| \geq \varepsilon v(\mathcal{H})$ for all $A \in \mathcal{F}$. Suppose that $\mathcal{H}$ is $(\mathcal{F}, \varepsilon)$-dense and $p \in (0,1)$ is such that, for every $\ell \in [k]$,*

$$\Delta_\ell(\mathcal{H}) \leq c\, p^{\ell-1} \frac{e(\mathcal{H})}{v(\mathcal{H})}. \tag{4}$$

Then there exist a family $\mathcal{S} \subseteq \mathcal{I}(\mathcal{H})$ with $|S| \leq Cp \cdot v(\mathcal{H})$, for all $S \in \mathcal{S}$, and functions $f : \mathcal{S} \to \overline{\mathcal{F}}$ and $g : \mathcal{I}(\mathcal{H}) \to \mathcal{S}$ such that for every $I \in \mathcal{I}(\mathcal{H})$, we have $g(I) \subseteq I$ and $I \setminus g(I) \subseteq f(g(I))$.

The families $\mathcal{S}$ and $f(\mathcal{S})$ above are called *source set* (or *fingerprint set*) and *container set* for the independent sets in $\mathcal{H}$, respectively.

Let $\mathcal{H}_n$ be a sequence of hypergraph. Let us see how Theorem 2.7 give us a good collection of containers for $\mathcal{I}(\mathcal{H}_n)$. Let $\mathcal{F}_n = \{A \subseteq V(\mathcal{H}_n) : e(\mathcal{H}_n[A]) \geq \varepsilon e(\mathcal{H}_n)\}$. Of course, $\mathcal{F}_n$ is an increasing family and $\mathcal{H}_n$ is $(\mathcal{F}_n, \varepsilon)$-dense. Suppose that $\mathcal{H}_n$ is such that $|A| \geq \varepsilon v(\mathcal{H}_n)$, for all $A \in \mathcal{F}_n$ and that for some $p_n \in (0,1)$, the inequality (4) holds for every $\ell \in [k]$. Suppose that, in addition, p_n goes to zero when n grows. Theorem 2.7 give us a family $\mathcal{S}_n \subseteq \mathcal{I}(\mathcal{H})$ with $|S| \leq Cp_n v(\mathcal{H}_n)$, for all $S \in \mathcal{S}_n$, and functions $f_n : \mathcal{S}_n \to \overline{\mathcal{F}}_n$ and $g_n : \mathcal{I}(\mathcal{H}_n) \to \mathcal{S}_n$ such that for every $I \in \mathcal{I}(\mathcal{H}_n)$, we have $g_n(I) \subseteq I$ and $I \setminus g_n(I) \subseteq f_n(g_n(I))$. Let $\mathcal{C}_n = \{f_n(S) \cup S : S \in \mathcal{S}_n\}$. Then $\mathcal{C}_n$ is a collection of containers for $\mathcal{I}(\mathcal{H}_n)$, since if $I \in \mathcal{I}(\mathcal{H}_n)$, then $I \subseteq f_n(S) \cup S$, for $S = g_n(I)$. Notice that, since $f_n(S) \in \overline{\mathcal{F}}_n$, we have $\mathrm{e}(\mathcal{H}_n[f_n(S)]) < \varepsilon e(\mathcal{H}_n)$. Therefore, the number of edges in $f_n(S) \cup S$ is bounded by

$$\begin{aligned} e(f_n(S) \cup S) &\leq e(\mathcal{H}_n[f_n(\mathcal{S})]) + |S| \cdot \Delta_1(\mathcal{H}_n) \\ &\leq \varepsilon e(\mathcal{H}_n) + Cp_n v(\mathcal{H}_n) \cdot c\, \frac{e(\mathcal{H}_n)}{v(\mathcal{H}_n)} \\ &\leq 2\varepsilon e(\mathcal{H}_n), \end{aligned}$$

for n large. Thus, $e(\mathcal{H}_n[C]) \leq 2\varepsilon e(\mathcal{H}_n)$, for all $C \in \mathcal{C}_n$ (that is, each set in $\mathcal{C}_n$ is almost independent). Now, for some $B > 0$, we have

$$\begin{aligned} |\mathcal{C}_n| &\leq |\mathcal{S}_n| \\ &\leq \sum_{i=1}^{Cp_n v(\mathcal{H}_n)} \binom{v(\mathcal{H}_n)}{i} \\ &\leq B \binom{v(\mathcal{H}_n)}{Cp_n v(\mathcal{H}_n)} \\ &\leq B \left(\frac{1}{Cp_n} \right)^{Cp_n v(\mathcal{H}_n)} \\ &\leq 2^{\varepsilon v(\mathcal{H}_n)}, \end{aligned}$$

by taking n larger. Therefore, $\mathcal{C}_n$ is a good collection of containers for $\mathcal{I}(\mathcal{H}_n)$ and 2ε.

In order to develop the containers for F-free graphs, let $\mathcal{H}$ be the hypergraph $\mathcal{H}_{n,F}$ defined in Example 5 and let $\mathcal{F}$ be the family of graphs $G \subseteq K_n$ with at least $\varepsilon n^{v(F)}$ copies of F. So, $\mathcal{H}$ is $(\mathcal{F}, \varepsilon)$-dense. Therefore, if we want to apply Theorem 2.7, we have to choose p in order to have inequality (4). Thus, the containers for F-free graphs can be established proving that we can chose $p = n^{-1/m_2(F)}$. What we get from the hypergraph containers theorem can, then, be stated as the following.

Theorem 2.8 (Containers lemma for F-free Graphs)**.** *For any graph F and $\varepsilon > 0$, there exist n_0 and $D > 0$ such that the following is true. For every $n > n_0$, there exist $t = t(n)$ pairwise distinct subsets $S_1, \ldots, S_t \subseteq E(K_n)$ of edges of K_n and t subsets $C_1, \ldots, C_t \subseteq E(K_n)$ such that*

1. *each S_i, $i \in [t]$, contains at most $Dn^{2-1/m_2(F)}$ elements,*
2. *each C_i, $i \in [t]$, forms at most $\varepsilon n^{v(F)}$ copies of F in K_n,*
3. *if $G \subseteq K_n$ is an F-free graph, then there exists $i \in [t]$ such that $S_i \subseteq E(G) \subseteq C_i$.*

The sets $C_1, \ldots, C_t$ and $S_1, \ldots, S_t$ in the theorem above are called *containers* and *sources*, respectively.

Proof of Theorem 2.8. We shall apply Theorem 2.7. In order to do that, let $\mathcal{H}$ be the $e(F)$-uniform hypergraph with vertex set $V(\mathcal{H}) = E(K_n)$ and hyperedge set

$$E(\mathcal{H}) = \{E(F') : F' \subseteq K_n \text{ and } F' \text{ is isomorphic to } F\}.$$

First, we will show that inequality (4) holds for $p = n^{-1/m_2(F)}$. Let T be a subset of $V(\mathcal{H})$ with $\ell \leq e(F)$ elements. Thus, $\deg_{\mathcal{H}}(T)$ corresponds to the number of subgraphs $F' \subseteq K_n$ isomorphic to F for which $T \subseteq E(F')$. For a graph $H \subseteq F$, let us denote by $F(H,n)$ the number of subgraphs $F' \subseteq K_n$ isomorphic to F with $H \subseteq F'$. Thus, $F(H,n) \leq c(F,H) \cdot \binom{n}{v(F)-v(H)}$, where $c(F,H)$ is a positive constant. Therefore, we can bound $\deg_{\mathcal{H}}(T)$ by

$$\begin{aligned}\deg_{\mathcal{H}}(T) &\leq \max\{F(H,n) : H \subseteq F \text{ and } e(H) = \ell\}\\ &\leq c(F) \cdot \max\{n^{v(F)-v(H)} : H \subseteq F \text{ and } e(H) = \ell\},\end{aligned}$$

where $c(F) = \max\{c(F,H) : H \subseteq F\}$.

Now, let $p = n^{-1/m_2(F)}$. By the definition of $m_2(F)$, we have, for all $H \subseteq F$, that

$$p^{e(H)-1}n^{v(H)-2} \geq 1.$$

So, for all $H \subseteq F$ with $e(H) = \ell$, we have

$$n^{v(F)-v(H)} \leq p^{\ell-1}n^{v(F)-2}.$$

Thus,

$$\deg_{\mathcal{H}}(T) \leq c(F) \cdot p^{\ell-1}n^{v(F)-2}.$$

On the other hand, as $v(\mathcal{H}) = \Theta(n^2)$ and $e(\mathcal{H}) = \Theta(n^{v(F)})$, we have $e(\mathcal{H})/v(\mathcal{H}) = \Theta(n^{v(F)-2})$. Therefore, for a large enough constant $c = c(F)$, we must have

$$\deg_{\mathcal{H}}(T) \leq c \cdot p^{\ell-1}\frac{e(\mathcal{H})}{v(\mathcal{H})}.$$

This shows that inequality (4) holds.

Now, let $\mathcal{F}$ be the family

$$\mathcal{F} = \{E(G) : G \subseteq K_n \text{ and } G \text{ has more than } (\varepsilon/2)n^{v(F)} \text{ copies of F}\}.$$

Of course, $\mathcal{F}$ is an increasing family. We claim that $\mathcal{H}$ is $(\mathcal{F}, \varepsilon/2)$-dense. Indeed, if $A = E(G) \in \mathcal{F}$, then G has at least $(\varepsilon/2)n^{v(F)}$ copies of F and, therefore, $\mathcal{H}[A]$ has at least $(\varepsilon/2)n^{v(F)}$ hyperedges. Now, once $e(\mathcal{H}) \leq n^{v(F)}$, it follows that $e(\mathcal{H}[A]) \geq (\varepsilon/2)e(\mathcal{H})$. Also, we have $|A| = e(G) \geq \varepsilon v(\mathcal{H})$. In fact, since the number of copies of F in G is at most $e(G)n^{v(F)-2}$ and at least $(\varepsilon/2)n^{v(F)}$, it follows that $e(G) \geq \varepsilon\binom{n}{2} = \varepsilon v(\mathcal{H})$.

Therefore, $\mathcal{H}$ and $\mathcal{F}$ satisfy the conditions on the hypotheses of Theorem 2.7. The conclusion give us the existence of a family $\mathcal{S} = \{S_1, \ldots, S_t\}$ of independent subsets in $\mathcal{H}$ for which, for some positive constant D, we have $|S_i| \leq Cp \cdot v(\mathcal{H}) \leq Dn^{2-1/m_2(F)}$, for all $i \in [t]$. This give us the first item in the statement of the theorem.

The sets $C_1, \dots, C_t$ come from the function f given by Theorem 2.7. Indeed, for each $i \in [t]$, take $C_i = f(S_i) \cup S_i$. As $f(S_i) \in \overline{\mathcal{F}}$, we have that $f(S_i)$ has at most $(\varepsilon/2)n^{v(F)}$ copies of F. Notice that for a fixed edge e, the number of copies of F in K_n containing that edge is at most $n^{v(F)-2}$. Thus, each edge in S_i is contained in at most $n^{v(F)-2}$ copies of F in $f(S_i) \cup S_i$ and, therefore, there exist at most $n^{v(F)-2}|S_i| \leq Dn^{v(F)-1/m_2(F)} = o(n^{v(F)})$ copies of F in C_i that are not contained in $f(S_i)$. This way, we have no more than $\varepsilon n^{v(F)}$ copies of F in C_i.

And finally, for the third item, just notice that if G is an F-free graph, then $I = E(G)$ is an independent set in $\mathcal{H}$. From Thereom 2.7, $g(I) \subset I$ and $I \setminus g(I) \subseteq f(g(I))$. As $g(I) = S_i$, for some $i \in [t]$, and $C_i = f(S_i) \cup S_i$, we must then have $S_i \subseteq E(G) \subseteq C_i$. This concludes our proof. □

We shall apply Theorem 2.8 to prove the 1-statement of Theorem 1.3 in Chapter 3. And Thoerem 2.7 will be used to prove Theorem 2.6 in Chapter 5.

3 THE SYMMETRIC CASE

In this chapter, we prove Theorem 1.3. The proof will follow the one given in Nenadov and Steger (2016) and it is divided into two parts: the 1-statement and the 0-statement.

3.1 The 1-statement

In order to show the 1-statement, the following folkloric saturated version of Ramsey theorem will be useful.

Lemma 3.1. *Let $F_1, \ldots, F_r$ be any graphs. There exist $\varepsilon > 0$ and $n_0 \in \mathbb{N}$ such that if $n > n_0$, then any r-colouring of the edges of K_n contains at least $\varepsilon n^{v(F_i)}$ copies of F_i with the colour i, for some $i \in [r]$.*

Proof. Notice that we may assume that $F_1, \ldots, F_r$ are complete graphs. Take

$$k = \max\{v(F_1), \ldots, v(F_r)\}.$$

Let $t = R_r(k)$ be the k-th Ramsey number for r colours. Thus, for an arbitrary r-colouring of K_n, each t-subset of the vertices contains at least one monochromatic copy of K_k. Since a copy of K_k is contained in $\binom{n-k}{t-k}$ copies of K_t, there exist

$$\frac{\binom{n}{t}}{\binom{n-k}{t-k}} = \frac{\Theta\left(n^t\right)}{\Theta(n^{t-k})} = \Theta(n^k)$$

monochromatic copies of K_k. Thus we have at least $\Omega(n^k)$ copies of K_k for some colour $i \in [r]$. Now, notice that there exist

$$\binom{n - v(F_i)}{k - v(F_i)} = \Theta(n^{k-v(F_i)})$$

copies of K_k in K_n containing a fixec copy of F_i. Therefore, there exist at least

$$\frac{\Omega(n^k)}{\Theta(n^{k-v(F_i)})} = \Omega(n^{v(F_i)})$$

monochromatic copies of F_i of colour i. □

Now, we can prove the 1-statement in the Theorem 1.3 with the help of Lemma 3.1 and Theorem 2.8. The idea of the proof is the following. Suppose that $\mathbf{G}_{n,p} \nrightarrow F$. Then we can split the edges of $\mathbf{G}_{n,p}$ into two disjoint edge sets E_1 and E_2 that induce two F-free graphs G_1 and G_2, respectively. Then, by Theorem 2.8, there are $C_1, C_2 \subseteq E(K_n)$ containing, respectively, G_1 and G_2 and such that each of them generates few copies of F. We then consider the following 3-edge-colouring of K_n: we colour and edge e with colour

1, if $e \in C_1$; with colour 2, if $e \in C_2 \setminus C_1$; and with colour 3, if $e \in E(K_n) \setminus (C_1 \cup C_2)$. In particular, edges in K_n with colour 3 are non-edges. Applying Lemma 3.1 for $F_1 = F_2 = F$ and $F_3 = K_2$, since we have few copies of F with the colour 1 and 2, we must have many edges (or copies of K_2) with colour 3. This implies that $\mathbf{G}_{n,p}$ has many non-edges. Then we proceed showing that this happens with small probability, since $\mathbf{G}_{n,p}$, for $p \gg n^{-1/m_2(F)}$, is not too sparse in order to have this happening.

Proof of the 1*-statement in the Theorem 1.3.* Let $\varepsilon > 0$ and n_0 given by the Lemma 3.1 with $F_1 = F_2 = F$ and $F_3 = K_2$. Applying Theorem 2.8 for F-graphs, we get, for $n > n_0$ sufficiently large, $t(n)$ many containers $C_1, \dots, C_{t(n)}$ and sources $S_1, \dots, S_{t(n)}$. Let $G = \mathbf{G}_{n,p}$. Suppose that $G \nrightarrow F$. Then we can partition G into two edge disjoint F-free graphs G_1 and G_2. Therefore there are $i_1, i_2 \in [t(n)]$ such that $S_{i_j} \subseteq G_j \subseteq C_{i_j}$, for $j = 1, 2$. Then we have

$$
\begin{aligned}
\mathbb{P}\big[G \nrightarrow F\big] &\leq \sum_{i_1,i_2=1}^{t(n)} \mathbb{P}\big[(S_{i_1} \subseteq G_1 \subseteq C_{i_1}) \wedge (S_{i_2} \subseteq G_2 \subseteq C_{i_2})\big] \\
&\leq \sum_{i_1,i_2=1}^{t(n)} \mathbb{P}\big[(S_{i_1} \cup S_{i_2} \subseteq G) \wedge (G \subseteq C_{i_1} \cup C_{i_2})\big] \\
&= \sum_{i_1,i_2=1}^{t(n)} \mathbb{P}\big[(S_{i_1} \cup S_{i_2} \subseteq G) \wedge \big(E(K_n) \setminus (C_{i_1} \cup C_{i_2}) \subseteq \overline{G}\big)\big].
\end{aligned}
$$

Since, for each $i_1, i_2 \in [t(n)]$, we have $S_{i_1} \cup S_{i_2} \subseteq C_{i_1} \cup C_{i_2}$, it follows that $S_{i_1} \cup S_{i_2}$ and $E(K_n) \setminus (C_{i_1} \cup C_{i_2})$ are disjoint. Therefore, the events $S_{i_1} \cup S_{i_2} \subseteq G$ and $E(K_n) \setminus (C_{i_1} \cup C_{i_2}) \subseteq \overline{G}$ are independent. Thus,

$$
\mathbb{P}\big[G \nrightarrow F\big] \leq \sum_{i_1,i_2=1}^{t(n)} \mathbb{P}\big[S_{i_1} \cup S_{i_2} \subseteq G\big] \mathbb{P}\big[E(K_n) \setminus (C_{i_1} \cup C_{i_2}) \subseteq \overline{G}\big].
$$

Consider the follow 3-edge-colouring of K_n: colour an edge e with colour 1 if $e \in C_{i_1}$; with colour 2 if $e \in C_{i_2} \setminus C_{i_1}$; and with colour 3 if $e \in E(K_n) \setminus (C_{i_1} \cup C_{i_2})$. Since C_{i_1} and C_{i_2} contain at most $\varepsilon n^{v(F)}$ copies of F, applying Lemma 3.1 for $F_1 = F_2 = F$ and $F_3 = K_2$ to this 3-edge-colouring of K_n, it follows that $E(K_n) \setminus (C_{i_1} \cup C_{i_2})$ contains at least εn^2 edges. Therefore the event $E(K_n) \setminus (C_{i_1} \cup C_{i_2}) \subseteq \overline{G}$ happens with probability at most $(1-p)^{\varepsilon n^2}$. Thus,

$$
\begin{aligned}
\mathbb{P}\big[G \nrightarrow F\big] &\leq \sum_{i_1,i_2=1}^{t(n)} \mathbb{P}\big[S_{i_1} \cup S_{i_2} \subseteq G\big] (1-p)^{\varepsilon n^2} \\
&\leq \exp\left\{-\varepsilon p n^2\right\} \sum_{i_1,i_2=1}^{t(n)} \mathbb{P}\big[S_{i_1} \cup S_{i_2} \subseteq G\big],
\end{aligned}
$$

where we use the estimative $(1-p) \leq \exp\{-p\}$ (see Proposition A.2).

Now as $|S| \leq Dn^{2-1/m_2(F)}$ for every $S \in \{S_1, \ldots, S_{t(n)}\}$, we have

$$\sum_{i_1,i_2=1}^{t(n)} \mathbb{P}\big[S_{i_1} \cup S_{i_2} \subseteq G\big] = \sum_{m=1}^{2Dn^{2-1/m_2(F)}} \sum_{\substack{S_{i_1},S_{i_2} \in \{S_1,\ldots,S_{t(n)}\} \\ |S_{i_1} \cup S_{i_2}|=m}} \mathbb{P}\big[S_{i_1} \cup S_{i_2} \subseteq G\big]$$
$$\leq \sum_{m=1}^{2Dn^{2-1/m_2(F)}} \binom{\binom{n}{2}}{m} 4^m p^m,$$

where the term $\binom{\binom{n}{2}}{m} 4^m$ bounds with room to spare the number of pairs of sets $S_{i_1}, S_{i_2} \in \{S_1, \ldots, S_{t(n)}\}$ such that $|S_{i_1} \cup S_{i_2}| = m$. Using the estimative from Proposition A.3 to the binomial term, we get

$$\sum_{i_1,i_2=1}^{t(n)} \mathbb{P}\big[S_{i_1} \cup S_{i_2} \subseteq G\big] \leq \sum_{m=1}^{2Dn^{2-1/m_2(F)}} \binom{\frac{n^2}{2}}{m} (4p)^m$$
$$\leq \sum_{m=1}^{2Dn^{2-1/m_2(F)}} \left(\frac{2epn^2}{m}\right)^m.$$

Now, by Proposition A.5, the function $f(x) = (a/x)^x$, defined for $x > 0$ (for some given $a > 0$), is increasing for $x \leq a/e$. Let us set $a = 2epn^2 = 2eCn^{2-1/m_2(F)}$, where C is a very large constant to be determined later. So $f(x)$ is increasing for $x \leq 2Cn^{2-1/m_2(F)}$. Once we have $C > D$, the following holds for sufficiently large n.

$$\sum_{i_1,i_2=1}^{t(n)} \mathbb{P}\big[S_{i_1} \cup S_{i_2} \subseteq G\big] \leq 2Dn^{2-1/m_2(F)} \left(\frac{2epn^2}{2Dn^{2-1/m_2(F)}}\right)^{2Dn^{2-1/m_2(F)}}$$
$$\leq n^2 \left(\frac{eC}{D}\right)^{\left(\frac{2D}{C}\right)pn^2}.$$

Now, since $\frac{1}{x}\ln(ex)$ tends to zero when x tends to infinity, we can chose C large enough such that $(D/C)\ln(eC/D) \leq \varepsilon/4$. And since $\ln n \ll pn^2$, we have

$$\sum_{i_1,i_2=1}^{t(n)} \mathbb{P}\big[S_{i_1} \cup S_{i_2} \subseteq G\big] \leq \exp\left\{2\ln n + pn^2\left(\frac{D}{C}\right)\ln\left(\frac{eC}{D}\right)\right\}$$
$$\leq \exp\left\{\frac{\varepsilon}{2}pn^2\right\},$$

for n large. Since $pn^2 \gg 1$, we have

$$\mathbb{P}\big[G \nrightarrow F\big] \leq \exp\left\{-\frac{\varepsilon}{2}pn^2\right\} = o(1),$$

as we wanted. $\square$

3.2 The 0-statement

In this section, we want to prove the 0-statement in Theorem 1.3. Therefore, we need to find a 2-colouring of the edges of $\mathbf{G}_{n,p}$ avoiding monochromatic copies of F, provided that $p = cn^{1/m_2(F)}$ for some $c > 0$ small enough. We say that such a colouring is a *valid colouring*.

We say that $e \in E(G)$ is a *closed edge* if there are $F_1, F_2 \subseteq G$ distinct copies of F with $E(F_1) \cap E(F_2) = \{e\}$. Otherwise, we say that e is an *open edge*. An important insight that will guide us along the proof is the following. Suppose that e is a open edge in G and that $G - e$ has a valid colouring. Then such colouring can be extended to a valid colouring of G, for otherwise when we assign the red colour to e we must find $F_1 \subseteq G$ as a red copy of F containing e, while when we assign the blue colour to e, we must find $F_2 \subseteq G$ as a blue copy of F containing e. But then $E(F_1) \cap E(F_2) = \{e\}$, contradicting the fact that e is an open edge.

Therefore, in order to find a valid colouring of G, we may remove open edges from G until we get a graph $\widehat{G}$ with the property that every edge in $\widehat{G}$ is closed. So, if we find a valid colouring for $\widehat{G}$, then we can extend such colouring to the removed open edges and get a valid colouring for the whole graph G. Notice that it does not matter in which order we remove the open edges, we will always end up with the same graph $\widehat{G}$. The reason for this is that if an edge is open in a graph, then it is open in any of the subgraphs containing that edge. Furthermore, let us say that we removed the open edges $e_1, \ldots, e_{k_1}$ and got the graph $\widehat{G}_1$ and by removing the edges open edges $f_1, \ldots, f_{k_2}$ we got the graph $\widehat{G}_2$. Let us say that $k_1 \geq k_2$. Let i_0 be the minimum of $\{i : e_i \in \widehat{G}_2\}$ (notice that under the assumption that $\widehat{G}_1 \neq \widehat{G}_2$, such i_0 must exist). Then e_{i_0} is a closed edge in $\widehat{G}_2$; let $F_1, F_2 \subseteq \widehat{G}_2$ be copies of F with $E(F_1) \cap E(F_2) = \{e\}$. Thus e_{i_0} is a closed edge in G and it was removed in the process that generated $\widehat{G}_1$. Then for some $j < i_0$, we have $F_1 \cup F_2 \subseteq G_1 \setminus \{e_1, \ldots, e_{j-1}\}$ and $F_1 \cup F_2 \nsubseteq G_1 \setminus \{e_1, \ldots, e_j\}$. Thus $e_j \in E(F_1) \cup E(F_2) \subseteq \widehat{G}_2$, which contradicts the choice of i_0. Therefore, there is only one subgraph $\widehat{G}$ generated by this removing process. We say that such graph is the *F-core* of G.

So now, let us focus on finding a valid colouring for the F-core $\widehat{G}$ of $G = \mathbf{G}_{n,p}$. Let us say we have a 2-edge-colouring of $\widehat{G}$ that is not a valid colouring yet. A natural move would be to take a monochromatic copy of F in such a colouring, let us say F_1, and change the colour of an edge $e_1 \in E(F_1)$. If with this move we create a new copy of F, let us say F_2, then we change the colour of an edge $e_2 \in E(F_2) \setminus E(F_1)$. We keep doing this hoping to obtain a valid colouring. Of course, we can not guarantee that this will produce a valid colouring, since we can be caught in a circular changing returning to e_1 or, even worse, we can create many copies of F and not be able to deal with them. But we can try to distinguish those edges in $\widehat{G}$ which could possibly be affected by this sequence of moves starting from e_1. We formalize this with the following equivalence relation on $E(\widehat{G})$: for two edges $e, f \in E(\widehat{G})$, we have $e \sim f$ if there is a sequence $F_1, \ldots, F_q$ of copies of F in $\widehat{G}$, with $E(F_i) \cap E(F_{i+1}) \neq \emptyset$, for $i \in [q-1]$, and $e \in F_1$ and $f \in F_q$. This way, if $e \not\sim f$, then a move like above starting from e will never affect the colour of f. Each equivalence class in $E(\widehat{G})/\sim$ is said to be an *F-component* of $\widehat{G}$.

For $p = cn^{-1/m_2(F)}$, the expected number of copies of F in $G = \mathbf{G}_{n,p}$ containing a fixed edge e of $\mathbf{G}_{n,p}$ is bounded by

$$2e(F)\frac{(v(F)-2)!}{\mathrm{aut}(F)} \cdot \binom{n-2}{v(F)-2} \cdot p^{e(F)-1} \leq \frac{2e(F)}{\mathrm{aut}(F)} \cdot n^{v(F)-2} \cdot p^{e(F)-1} = \frac{2e(F)}{\mathrm{aut}(F)} \cdot c^{e(F)-1}.$$

Thus, for $c = c(F)$ small enough, we expect to find no more than one such a copy. Furthermore, we hope to find many open edges in G in a such way that $\widehat{G}$ should be very small. More precisely, we hope that the F-components of $\widehat{G}$ are very small. This is formalized in the following key lemma.

Lemma 3.2. *(Key Lemma) Let F be a strictly 2-balanced graph with $e(F) \geq 3$. There exist $c = c(F) > 0$ and $L = L(F) > 0$ such that if $p \leq cn^{-1/m_2(F)}$, then w.h.p., every F-component of the F-core of $\mathbf{G}_{n,p}$ has size at most L.*

We will postpone the proof of Lemma 3.2 for a while. It is worthy mentioning that Lemma 3.2 was generalized by Nenadov, Škorić, and Steger (2015) to prove results on anti-Ramsey properties on random hypergraphs. In order to prove the 0-statement in Theorem 1.3, the following theorem due to Rödl and Ruciński (1995) will be useful. We will not present a proof of this fact, since the proof is a little bit technical and it is beyond the scope of this text. The reader can also find the proof of this fact in Nenadov and Steger (2016).

Theorem 3.1 (Rödl and Ruciński, 1995)**.** *Let G be any graph and F be a graph with $m_2(F) > 1$. If $m(G) \leq m_2(F)$, then $G \not\to F$.*

Let's see how can we combine Lemma 3.2 and Theorem 3.1 to obtain the 0-statement of Theorem 1.3.

Proof of 0-statement of Theorem 1.3. We must prove that if F is a graph with $m_2(F) > 1$, then there exists c such that w.h.p., $\mathbf{G}_{n,p} \not\to F$ for $p \leq cn^{-1/m_2(F)}$. First note that we can assume that F is a strictly 2-balanced graph, since if $F' \subset F$ is such that $\mathbf{G}_{n,p} \not\to F'$, then $\mathbf{G}_{n,p} \not\to F$, and therefore, if F was not a 2-balanced graph, we could replace F by a 2-balanced subgraph of F with the same 2-density. Let c and L be the constants given by Lemma 3.2. Then w.h.p., every F-component of the F-core of $\mathbf{G}_{n,p}$ has size at most L, for $p \leq cn^{-1/m_2(F)}$. Our goal is to show that w.h.p., $\mathbf{G}_{n,p}$ does not contain a subgraph G with at most L vertices and $m(G) > m_2(F)$. Once we have this done, then by Theorem 3.1, it follows that w.h.p., every F-component of the F-core of $\mathbf{G}_{n,p}$ is not Ramsey to F. Furthermore, we have $\mathbf{G}_{n,p} \not\to F$.

Let G be a graph with at most L vertices and $m(G) > m_2(F)$. Consider $G' \subseteq G$

such that $m(G) = e(G')/v(G')$. Then the probability of the event $G \subseteq \mathbf{G}_{n,p}$ is bounded by the probability of the event $G' \subseteq \mathbf{G}_{n,p}$. Thus, by Markov's inequality (Theorem B.1), we have

$$\begin{aligned}
\mathbb{P}\big[G \subseteq \mathbf{G}_{n,p}\big] &\leq \mathbb{P}\big[G' \subseteq \mathbf{G}_{n,p}\big] \\
&\leq \mathbb{E}\big[\#\{G' \subseteq \mathbf{G}_{n,p}\}\big] \\
&\leq n^{v(G')} \cdot p^{e(G')} \\
&\leq c^{e(G')} \cdot n^{v(G') - e(G')/m_2(F)} \\
&= c^{e(G')} \cdot n^{v(G')(1 - m(G)/m_2(F))} \\
&= o(1).
\end{aligned}$$

Therefore, w.h.p., there is no copy of G in $\mathbf{G}_{n,p}$. Since there are $O(1)$ many graphs G with L vertices (and $m(G) > m_2(F)$), it follows that, w.h.p., $\mathbf{G}_{n,p}$ does not contain a graph G with at most L vertices and $m(G) > m_2(F)$. This concludes the proof. $\square$

3.3 Proof of the Key Lemma

Now, we are going to prove the Key Lemma (Lemma 3.2). We want to prove that the F-components of the F-core $\widehat{G}$ of $G = \mathbf{G}_{n,p}$, are typically small. For this purpose, we start describing a procedure that generates a sequence $(F_0, F_1, \ldots, F_\ell)$ of copies of F in a way that $G' = \cup_{i \leq \ell} F_i$ and $F_{i-1} \cap (\cup_{j \leq i} F_j) \neq \emptyset$, for each $i \in [\ell]$. Thereby, we should analyse how this sequence is typically build in an F-component of the F-core of $\mathbf{G}_{n,p}$.

Algorithm 1

procedure
    Let F_0 be a copy of F in G';
    $\ell \leftarrow 0$; $G_0 \leftarrow F_0$;
    while $G_\ell \neq G'$ **do**
        $\ell \leftarrow \ell + 1$;
        if $G_{\ell-1}$ contains an open edge **then**
            let $\ell' < \ell$ be the smallest index such that $F_{\ell'}$ contains an open edge;
            let e be any open edge in $F_{\ell'}$;
            let F_ℓ be a copy of F in G' that contains e but is not contained in $G_{\ell-1}$;
        else
            let F_ℓ be a copy of F in G' that is not contained in $G_{\ell-1}$ and intersects $G_{\ell-1}$ in at least one edge;
        $G_\ell \leftarrow G_{\ell-1} \cup F_\ell$;

As the copies F_ℓ of F, $\ell \geq 1$, are being added to $G_{\ell-1}$ in the ℓ-th step in Algorithm 1, we can have two types of copies of F: one that intersects $G_{\ell-1}$ in exactly two vertices, and one that intersects $G_{\ell-1}$ in some subgraph J with $v(J) \geq 3$. We say that F_ℓ

is a *regular* copy if it is of the first type, and F_ℓ is a *degenerate* copy if it is of the second type. Now, F_0 is not considered a regular copy, neither a degenerate copy. We denote by $\text{reg}(\ell)$ and $\deg(\ell)$, respectively, the number of regular and degenerate copies among $F_1, \ldots, F_\ell$.

From the way we determined Algorithm 1, some regular copies F_i, for $0 \leq i \leq \ell$, has all of those vertices that was added in the step i not yet touched by the subsequent copies F_j in the step ℓ. Those copies F_i are said to be *fully-open* at time ℓ. More precisely, for $0 \leq i \leq \ell$, we say that F_i is fully-open at time ℓ if F_i is a regular copy (or $i = 0$) and $(V(F_i) \setminus V(G_{i-1})) \cap (V(G_\ell) \setminus V(G_i)) = \emptyset$. We denote by $f_o(\ell)$ the number of fully-open copies at time ℓ. From the way we have defined the process, if the process has not stopped yet, and many regular copies were added during the earlier steps of the process, then we hope to find many fully-open copies at that time. The following lemma states this formally.

Lemma 3.3. *For every $\ell \geq 1$, assuming the process does not stop before adding the ℓ-th copy, we have*

$$f_o(\ell) \geq \text{reg}(\ell)\left(1 - \frac{1}{e(F) - 1}\right) - \deg(\ell) \cdot v(F).$$

In particular, after adding L copies, with at most τ being degenerated, we will still have at least

$$(L - \tau)\left(1 - \frac{1}{e(F) - 1}\right) - \tau \cdot v(F) \tag{5}$$

fully-open copies at time L. Then F_L can not be the last copy in the process.

In order to prove the lemma above, the following lemma will be necessary. It essentially says that, because F is strictly 2-balanced, when we add the regular copy F_ℓ to the graph $G_{\ell-1}$ by attaching to an edge xy, the only open edge in $G_{\ell-1}$ which turns close is the edge xy.

Lemma 3.4. *Let F be a strictly 2-balanced graph let G be an arbitrary graph. Suppose G' is a graph obtained from G attaching a F to a single edge xy of G. Then if F' is a copy of F in G' which contains at least one element from $V(F) \setminus \{x, y\}$, then $F' = F$.*

Proof. First, notice that F is connected and it has no vertex with degree one. Let us show that F is 2-connected. Indeed, suppose we have a cut-vertex v in F. Then there exist $F_1, F_2 \subseteq F$ connected such that $V(F_1) \cap V(F_2) = \{v\}$ and we must have $v(F_1), v(F_2) \geq 3$. Since F is 2-balanced, we have

$$\frac{e(F_1) - 1}{v(F_1) - 2} < m_1(F) \quad \text{and} \quad \frac{e(F_2) - 1}{v(F_2) - 2} < m_1(F).$$

Therefore,

$$\frac{e(F) - 2}{v(F) - 3} = \frac{(e(F_1) - 1) + (e(F_2) - 1)}{(v(F_1) - 2) + (v(F_2) - 2)} < m_2(F) = \frac{e(F) - 1}{v(F) - 2}.$$

This implies that $e(F) < v(F) - 1$, which is a contradiction with the fact that F is connected. Therefore F is 2-connected.

Now let us say F' is a copy of F which contradicts the lemma. Then F' contains at least one vertex from $V(F) \setminus \{x, y\}$ and one vertex from $G \setminus \{x, y\}$. Let $F_1 = F'[V(G)]$ and $F_2 = F'[V(F)]$. Since F' is 2-connected, it follows that $\{x, y\} \subseteq V(F')$. Then $v(F) = v(F_1) + v(F_2) - 2$ and $v(F_1), v(F_2) \geq 3$. If xy is not an edge in F' then we add xy to the edge-set of F_2. Therefore, $e(F) = e(F_1) + e(F_2) - 1$, regardless whether xy is an edge in F' or not. Since F is strictly 2-balanced, we have

$$\frac{e(F_1) - 1}{v(F_1) - 2} < m_2(F) \quad \text{and} \quad \frac{e(F_2) - 1}{v(F_2) - 2} < m_2(F).$$

On the other hand,

$$m_2(F) = \frac{e(F) - 1}{v(F) - 2} = \frac{e(F_1) - 1 + e(F_2) - 1}{v(F_1) - 2 + v(F_2) - 2} < m_2(F).$$

A contradiction.

$\square$

Proof of Lemma 3.3. Let $\phi(\ell) = \text{reg}(\ell)(1 - 1/(e(F) - 1)) - \deg(\ell) \cdot v(F)$ be the right hand side of the inequality in the statement. We will prove by induction the following stronger statement:

$$f_o(\ell) \geq \begin{cases} \phi(\ell) & \text{if } F_\ell \text{ is a regular copy;} \\ \phi(\ell) + 1 & \text{if } F_\ell \text{ is a degenerate copy.} \end{cases}$$

The initial case when $\ell = 1$ follows trivially. So suppose we have $\ell \geq 2$. For $1 \leq \ell' \leq \ell$, let

$$\kappa(\ell') := |\{0 \leq i < \ell' : \ F_i \text{ is fully-open only at time } \ell' - 1 \text{ but not at time } \ell'\}|.$$

So, $\kappa(\ell')$ counts the number of fully-open copies at time $\ell' - 1$ which are destroyed when $F_{\ell'}$ is added. In particular, $\kappa(\ell') \leq v(F) - 1$, since we can not destroy more than $v(F) - 1$ copies. And also, if $F_{\ell'}$ is a regular copy, then $\kappa(\ell') \leq 1$.

If F_ℓ is a degenerate copy, then $\text{reg}(\ell) = \text{reg}(\ell - 1)$, $\deg(\ell) = \deg(\ell - 1) + 1$ and

$\kappa(\ell) = f_o(\ell-1) - f_o(\ell) \leq v(F) - 1$. Thus

$$\begin{aligned} f_o(\ell) &\geq f_o(\ell-1) - v(F) + 1 \\ &\geq \phi(\ell-1) - v(F) + 1 \\ &= \mathrm{reg}(\ell-1)(1 - 1/(e(f)-1)) - \deg(\ell-1)\cdot v(F) - v(F) + 1 \\ &= \mathrm{reg}(\ell)(1 - 1/(e(f)-1)) - (\deg(\ell)-1)\cdot v(F) - v(F) + 1 \\ &= \mathrm{reg}(\ell)(1 - 1/(e(f)-1)) - \deg(\ell)\cdot v(F) + 1 \\ &\geq \phi(\ell) + 1. \end{aligned}$$

Therefore, we can assume F_ℓ is a regular copy and let

$$\ell' := \max\{1 \leq \ell' < \ell : \kappa(\ell') > 0 \text{ or } F_{\ell'} \text{ is a degenerate copy}\}.$$

Thereby,

$$\phi(\ell) = \phi(\ell') + (\ell - \ell')(1 - 1/(e(F)-1)),$$

since for $\ell' < i \leq \ell$, F_i is a regular copy. Also, as $\kappa(i) = 0$, for $\ell' < i < \ell$, we have that F_i is a fully-open copy at time i. Thus

$$f_o(\ell) = f_o(\ell') + \ell - \ell' - \kappa(\ell).$$

If $F_{\ell'}$ is a degenerate copy, then by the induction hypothesis, $f_o(\ell') \geq \phi(\ell') + 1$. Then, since $\kappa(\ell) \leq 1$, we have

$$\begin{aligned} f_o(\ell) &= f_o(\ell') + \ell - \ell' - \kappa(\ell) \\ &\geq \phi(\ell') + 1 + \ell - \ell' - \kappa(\ell) \\ &\geq \phi(\ell') + \ell - \ell' \\ &= \phi(\ell) - (\ell - \ell')(1 - 1/(e(F)-1)) + \ell - \ell' \\ &= \phi(\ell) + 1/(e(F)-1) \\ &\geq \phi(\ell). \end{aligned}$$

Let us assume $F_{\ell'}$ is a regular copy. If $\kappa(\ell) = 0$, then the same argument as above works to show that $f_o(\ell) \geq \phi(\ell)$. So, let us say that $\kappa(\ell) = 1$. We claim that in this case we must have $\ell \geq \ell' + e(F) - 1$. In fact, we are going to prove that actually we must have $\kappa(\ell'+1) = \cdots = \kappa(\ell' + e(F) - 2) = 0$.

Claim 3.1. If $F_i, \ldots, F_{i+e(F)-2}$ is a sequence of consecutive regular copies of F such that $\kappa(i) = 1$, then $\kappa(i+1) = \cdots = \kappa(i + e(F) - 2) = 0$.

Indeed, if F_i is a regular copy and $\kappa(i) = 1$, then F_i intersects some copy $F_{i'}$, $i' < i$, in exactly one edge and we know that $F_{i'}$ was fully-open at time $i-1$. Thus, at time $i-1$,

$F_{i'}$ had $e(F)-1$ open edges (or $e(F)$, if $i'=0$) and F_i intersects $F_{i'}$ in one of those open edges. At time $i+1$, for the wat the process was defined, the copy F_{i+1} must to choose one of the $e(F)-2$ remaining open edges from $F_{i'}$ to attach itself. From Lemma 3.4, it remains $e(F)-3$ open edges in $F_{i'}$ at step $i+2$. Proceeding at this way, $F_{i+2},\ldots,F_{i+e(F)-2}$ must all be attached to $F_{i'}$ by some open edge without closing other edges than that one which the copy was attached. This implies that $\kappa(i+1)=\cdots=\kappa(i+e(F)-2)=0$. Therefore, $\ell \geq \ell' + e(F) - 1$. This way, we have

$$\begin{aligned}
f_o(\ell) &= f_o(\ell') + \ell - \ell' - 1 \\
&\geq f_o(\ell') + (\ell - \ell')(1 - 1/(e(F)-1)) \\
&\geq \phi(\ell') + (\ell - \ell')(1 - 1/(e(F)-1)) \\
&= \phi(\ell).
\end{aligned}$$

This finishes the proof of the lemma. □

Now, we can finally prove the Key Lemma (Lemma 3.2). Our aim is to chose a very large L for which the expected number of sequences $(F_0, F_1, \ldots)$ that generate an F-component in $\mathbf{G}_{n,p}$ with more than L vertices throughout Algorithm 1 is $o(1)$. Then, the result will follow by Markov's inequality (Theorem B.1).

First, given $G_{\ell-1}$, the expected number of choices for F_ℓ being a regular copy attached to an open edge in $G_{\ell-1}$ in Algorithm 1 is bounded by

$$2e(F)^2 \cdot n^{v(F)-2} \cdot p^{e(F)-1} \leq 2e(F)^2 \cdot c \leq \frac{1}{2}, \tag{6}$$

for $0 < c < 1/(4\,\mathrm{aut}(F)e(F)^2)$. The term $2e(F)^2$ comes from the choice of an edge uv in F_ℓ and an edge $u'v'$ in $F_{\ell'}$ (where ℓ' is given by the Algorithm 1) and a choice between $u=u'$ or $u=v'$. The term $n^{v(F)-2}$ bounds the number of choices of copies of F in K_n, and $p^{e(F)-1}$ bounds the probability of each of this copies being in $\mathbf{G}_{n,p}$.

Given the $G_{\ell-1}$, the expected number of choices for F_ℓ being a regular copy attached to a closed edge in $G_{\ell-1}$ is bounded by

$$2\ell e(F)^2 \cdot n^{v(F)-2} \cdot p^{e(F)-1} \leq 2\ell e(F)^2 \cdot c \leq \frac{\ell}{2}. \tag{7}$$

The term $2\ell e(F)^2$ comes from the choice of an edge uv in F_ℓ and an edge $u'v'$ in $G_{\ell-1}$ (there are at most $\ell e(F)$ such edges) and a choice between $u=u'$ or $u=v'$. The remaining terms are considered for the same reasons from the last case.

Now let us consider the case when F_ℓ forms a degenerate copy. First, since F is strictly 2-balanced, for every $J \subsetneq F$ with $v(J) \geq 3$ we have

$$m_2(F) = \frac{e(F)-1}{v(F)-2} > \frac{e(J)-1}{v(J)-2}.$$

Thus we have

$$\frac{e(F) - e(J)}{v(F) - v(J)} > m_2(F).$$

Then we may choose an $\alpha > 0$ with

$$(v(F) - v(J)) - \frac{e(F) - e(J)}{m_2(F)} < -\alpha, \quad \text{for all } J \subsetneq F \text{ with } v(J) \geq 3.$$

Now we can bound the expected number of choices from F_ℓ being a degenerate copy attached to $G_{\ell-1}$ by

$$\sum_{\substack{J \subsetneq F \\ v(J) \geq 3}} (\ell \cdot v(F))^{v(J)} \cdot n^{v(F)-v(J)} \cdot p^{e(F)-e(J)} < (\ell v(F) 2^{e(F)})^{v(F)} n^{-\alpha}, \tag{8}$$

taking $c = c(F)$ small enough. The first term comes from the number of choices of $v(J)$ vertices in $G_{\ell-1}$ to form a copy of J in $G_{\ell-1}$ which F_ℓ will be attached on. The last term $n^{v(F)-v(J)} \cdot p^{e(F)-e(J)}$ bounds the expected number of choices of $v(F) - v(J)$ vertices in $\mathbf{G}_{n,p}$ to form a copy of $F - J$.

Now let τ be such that $\tau \cdot \alpha > v(F) + 1$ and take $L = L(F)$ large enough such that the term in (5) is positive. Take $\ell_0 = (v(F) + 1) \log n + \tau$.

We can bound the expected number of sequences $(F_0, F_1, \ldots, F_{\ell'})$ of copies of F in $\mathbf{G}_{n,p}$ such that $\ell 1 \leq \ell_0$ and $F_{\ell'}$ is the xi-th degenerate copy of F by

$$\sum_{\ell' \leq \ell_0} \binom{\ell' - 1}{\tau - 1} \cdot \left[(\ell v(F) 2^{e(F)})^{v(F)} n^{-\alpha}\right]^\tau \cdot n^{v(F)} \cdot L^L \cdot 2^{-(\ell' - L - \tau)}$$
$$\leq n^{v(F)} \cdot o(n) \cdot n^{-\alpha \cdot \tau} = o(1).$$

The binomial term in the summation above comes from the choice of $\tau - 1$ copies of F in the set $F_1, \ldots, F_{\ell'-1}$ to be the degenerate copies in the sequence. The term inside the brackets comes from the expected number of the choices of τ degenerate copies of F in the sequence (look at inequality (8)). The term $n^{v(F)}$ corresponds to the choice of the first copy F_0 in the sequence. Now, according to the inequalities 6 and 7, we can bound the expected number of choices for a regular copy F_i for $i \leq L$ by $L/2$ or $1/2$, depending on whether it is attached to an open or closed edge. Bounding both of them simply by L, we have the expected number of regular copies F_i for $i < L$ bounded by L^L; this is where the term L^L in the summation above comes from. Since $\deg(i) \leq \tau$, for every $i \leq \ell'$, we have by the inequality (5) an by the choice of L that after L steps, all the regular copies are attached to open edges. Therefore, the remaining regular copies F_i for $i > L$ (we have at least $\ell' - L - \tau$ of them) are all of them attached to an open edge. Thus, their probability can be bound by $1/2$ and that give us the term $(1/2)^{\ell' - L - \tau}$.

Therefore, w.h.p., every sequence $(F_1, \ldots, F_\ell)$ with $\ell \leq \ell_0$ has less than τ degenerate copies F_i. Furthermore, we can not have a sequence $(F_1, \ldots, F_\ell)$ generated by Algorithm 1 with $L \leq \ell \leq \ell_0$, since this would imply by inequality (5) that the process did not end yet. Therefore, a sequence $(F_1, \ldots, F_\ell)$ generated by Algorithm 1 when applied in an F-component of $\mathbf{G}_{n,p}$ ends either with $\ell < L$, which is fine for us, or with $\ell > \ell_0$. Thus, it suffices to show that the expected number of subgraphs of $\mathbf{G}_{n,p}$ that can be generated by a sequence $(F_1, \ldots, F_\ell)$ with $\ell > \ell_0$ is $o(1)$. We, actually, are going to show that the expected number of subgraphs of $\mathbf{G}_{n,p}$ generated by a sequence $(F_1, \ldots, F_{\ell_0})$ is $o(1)$, what is stronger than what we wanted. Indeed, such expected value can be bounded by

$$\sum_{k<\tau} \binom{\ell_0}{k} n^{v(F)} \cdot \left[(\ell_0 v(F) 2^{e(F)})^{v(F)} \cdot n^{-\alpha}\right]^k \cdot L^L \cdot 2^{-(\ell_0 - L - k)}$$
$$\leq n^{v(F)} \cdot o(n) \cdot n^{-(v(F)+1)} = o(1).$$

The summation above counts the sequences with $k < \tau$ degenerate copies before F_{ℓ_0}. The binomial term corresponds to the choice of those copies before F_{ℓ_0} that form degenerate copies. The term inside the brackets corresponds to the expected value of choices of each degenerate copies in $G^r_{n,p}$. The term $n^{v(F)}$ corresponds to the number of choices for F_0. The term $L^L \cdot 2^{-(\ell_0 - k)}$ comes from the fact that the expected number of choices of regular copies at a step $i \leq L$ is bounded by $L/2 \leq L$, whereas after L (there is at least $\ell_0 - L - k$ such copies) is bounded by $1/2$. This concludes the proof of the Lemma 3.2. $\square$

4 THE ASYMMETRIC CASE

In this section, we will prove Theorem 1.7, which we rephrase it here with an alteration in the notation.

Theorem 4.1. *Let F and H be graphs with $1 < m_2(F) < m_2(H)$ and such that H is 2-balanced. Then there exists a positive constant C such that if $p \geq Cn^{-1/m_2(H,F)}$, then*

$$\lim_{n\to\infty} \mathbb{P}\big[\mathbf{G}_{n,p} \to (H,F)\big] = 1.$$

Furthermore, the same conclusion holds if, in instead, H is strictly 2-balanced and $1 < m_2(F) \leq m_2(H)$.

Kohayakawa and Kreuter (1997) proved Theorem 4.1 for cycles and we shall use the same idea from them to prove this more general version. The idea of the proof is roughly the following. We will look at the copies of H in $\mathbf{G}_{n,p}$ that do not share an edge with other copies of H. We say that those copies are isolated copies of H. Suppose we have an 2-edge-colouring of $\mathbf{G}_{n,p}$ avoiding monochromatic copies of H with the first colour. Then each of those isolated copies of H must have an edge with the second colour. Those edges are all distinct and generate a subgraph G' of $\mathbf{G}_{n,p}$. We will show that G' contains a copy of F, w.h.p.. This give to us the existence of a monochromatic copy of F in $\mathbf{G}_{n,p}$ with the second colour.

4.1 The isolated copies of H in $\mathbf{G}_{n,p}$

Let H be a given graph. An *isolated* copy of H in a graph G is a copy $H' \subseteq G$ of H in G for which there is no other copy of H in G with some edge in common with H'. Let $\{H_1, \dots, H_s\}$ be all the isolated copies of H in $\mathbf{G}_{n,p}$. Let $H(\mathbf{G}_{n,p})$ be the subgraph of $\mathbf{G}_{n,p}$ with $V(H(\mathbf{G}_{n,p})) = V(\mathbf{G}_{n,p})$ and $E(H(\mathbf{G}_{n,p})) = \cup_{i=1}^{s} E(H_i)$. Therefore, $H(\mathbf{G}_{n,p})$ is a random graph with n vertices and $s \cdot e(H)$ edges.

Let q be the expected number of copies of H in $\mathbf{G}_{n,p}$ containing a given edge xy of K_n. It is easy to see that

$$q = 2e(H) \cdot \frac{(v(H)-2)!}{\mathrm{aut}(H)} \cdot \binom{n-2}{v(H)-2} \cdot p^{e(H)} = \Theta\big(n^{v(H)-2}p^{e(H)}\big). \tag{9}$$

Therefore, we have for some positive constants $A = A(H)$ and $B = B(H)$ that

$$An^{v(H)-2}p^{e(H)} \leq q \leq Bn^{v(H)-2}p^{e(H)}. \tag{10}$$

In the following two lemmas, we shall show that $H(\mathbf{G}_{n,p})$ has some properties that remind us a random graph with density q. For an edge set $E \subseteq K_n$, we write $E \sqsubseteq H(\mathbf{G}_{n,p})$ if $E \subseteq E(H(\mathbf{G}_{n,p}))$ and each copy of H in $H(\mathbf{G}_{n,p})$ contains at most one edge from E.

Lemma 4.1. *Let H be a graph, p be the edge density of $\mathbf{G}_{n,p}$ and q be as in (9). Then, for any subset of edges $E \subseteq E(K_n)$, we have*

$$\mathbb{P}\big[E \sqsubseteq H(\mathbf{G}_{n,p})\big] \leq q^{|E|}.$$

Proof. Let us say $E = \{e_1, \dots, e_m\}$. If $E \sqsubseteq H(\mathbf{G}_{n,p})$, then we must have m distinct isolated copies $H_1, \dots, H_m$ of H contained in $\mathbf{G}_{n,p}$ with $H_i \supseteq e_i$, for all $i \in [m]$. Let Ω be the set of m-uples $(H_1, \dots, H_m)$ of distinct isolated copies of H contained in K_n with $e_i \subseteq H_i$, for each $i \in [m]$. And let X be the number of such m-uples contained in $\mathbf{G}_{n,p}$. Since those copies H_i are isolated and distinct, the events $H_i \subseteq \mathbf{G}_{n,p}$, for $i \in [m]$, are mutually independent. Therefore, writing $\sum_{H_i}$ for the sum over all copies H_i of H in K_n, we must have

$$\begin{aligned}
\mathbb{P}\big[E \sqsubseteq H(\mathbf{G}_{n,p})\big] &\leq \mathbb{E}\big[X\big] \\
&\leq \sum_{(H_1,\dots,H_m)\in\Omega} \mathbb{P}\big[H_1 \subseteq \mathbf{G}_{n,p}, \dots, H_m \subseteq \mathbf{G}_{n,p}\big] \\
&= \sum_{(H_1,\dots,H_m)\in\Omega} \mathbb{P}\big[H_1 \subseteq \mathbf{G}_{n,p}\big] \cdots \mathbb{P}\big[H_m \subseteq \mathbf{G}_{n,p}\big] \\
&\leq \left(\sum_{H_1} \mathbb{P}\big[H_1 \subseteq \mathbf{G}_{n,p}\big]\right) \cdots \left(\sum_{H_m} \mathbb{P}\big[H_m \subseteq \mathbf{G}_{n,p}\big]\right) \\
&= q^m,
\end{aligned}$$

as stated. $\square$

Lemma 4.2. *Let H be a graph with h vertices and α be a positive constant with $\alpha < (h-1)/e(H)$. Then taking $p = cn^{-\alpha}$, for some positive constant c, and q as in (9), we have that for any $\eta > 0$, w.h.p. the graph $H(\mathbf{G}_{n,p})$ is $(\eta, 6e(H), q)$-upper-uniform*

Proof. We have to show that w.h.p., the following happens. For all disjoint subsets U and V of $V(H(\mathbf{G}_{n,p}))$ with $|U|, |V| \geq \eta v(H(\mathbf{G}_{n,p}))$, we have $d(U,V) \leq 6e(H)q$. If this does not happen, them we must have some pair (U,V) of disjoint subset of $V(H(\mathbf{G}_{n,p}))$ with $|U|, |V| \geq \eta v(H(\mathbf{G}_{n,p}))$ but with $d(U,V) > 6e(H)q$, that is,

$$e(U,V) \geq 6e(H)q|U||V|.$$

Then, by the pigeonhole principle, there exists a set of edges $E \subseteq E(U,V)$ such that $E \sqsubseteq H(\mathbf{G}_{n,p})$ and $m := |E| \geq 6q|U||V| \geq 6\eta q n^2$. By Lemma 4.1 and by Markov's inequality (Theorem B.1), this happens with probability at most

$$\binom{|U||V|}{m} q^m \leq \left(\frac{e|U||V|q}{m}\right)^m \leq \left(\frac{e}{6}\right)^m.$$

Now, as there are at most 2^{2n} choices for the pair (U, V), by the union bound, we get that the probability of failing the required property is at most

$$2^{2n}\left(\frac{e}{6}\right)^m \leq 2^{2n-6\eta qn^2}.$$

Since $qn^2 \geq Ac^{e(H)}n^{h-\alpha e(H)}$, and as $h - \alpha e(H) > 1$, we have $qn^2 \gg n$. This way, the required property for the $(\eta, 6e(H), q)$-upper-uniformity fails with probability $o(1)$, as stated. □

Let H be a graph with vertex set $V(H) = [h]$. Consider a family $\mathbf{V} = (V_1, \ldots, V_h)$ of h disjoint subsets of $V(\mathbf{G}_{n,p})$. A copy H' of H contained in $\mathbf{G}_{n,p}$ is called a **V**-copy of H if $V(H') = \{v_1, \ldots, v_h\}$ with $v_i \in V_i$ and $v_iv_j \in E(H')$ whenever that $ij \in E(H)$. Let $Z_{\mathbf{V}}$ be the number of **V**-copies of H in $\mathbf{G}_{n,p}$ and let $Y_{\mathbf{V}}$ be the number of **V**-copies of H which are isolated copies of H in $\mathbf{G}_{n,p}$. The following lemma roughly states that if the colour classes are large enough, then we can expect at least a half of the **V**-copies of H being isolated copies of H.

Lemma 4.3. *Let H be a 2-balanced-graph with h vertices and α be a positive constant with*

$$\frac{1}{m_2(H)} < \alpha < \frac{h-1}{e(H)}. \tag{11}$$

Take $p = cn^{-1/\alpha}$, for some $c > 0$. Suppose that $\mathbf{V} = (V_1, \ldots, V_h)$ is a family of disjoint subsets of $V(\mathbf{G}_{n,p})$ such that for every $i \in [h]$, we have $|V_i| \geq n/\log n$. Then $Y_{\mathbf{V}} \geq (1/2)\mathbb{E}\big[Z_{\mathbf{V}}\big]$, w.h.p..

Proof. Let H, α and p be as in the statement. Let $\delta = \alpha - 1/m_2(H)$. As H is 2-balanced, for every $J \subseteq H$ with $v(J) \geq 3$, we have

$$m_2(H) = \frac{e(H)-1}{v(H)-2} \geq \frac{e(J)-1}{v(J)-2}.$$

Thus, by Proposition A.1, we have

$$\frac{v(H)-v(J)}{e(H)-e(J)} \leq \frac{1}{m_2(H)} = \alpha - \delta.$$

It follows that

$$2h - v(J) - \alpha(2e(H) - e(J)) \leq h - \alpha e(H) - \delta(e(H) - e(J)), \tag{12}$$

for all $J \subsetneq H$. We will use this later on.

Now, take a family $\mathbf{V} = (V_1, \ldots, V_h)$ with $|V_i| \geq n/\log n$, for all $i \in [h]$. Let us denote $V_1 \times \cdots \times V_h$ by Γ. For each h-tuple $\mathbf{v} = (v_1, \ldots, v_h) \in \Gamma$, let $H_{\mathbf{v}}$ be the copy of H such that $V(H) = \{v_1, \ldots, v_h\}$ and $E(H_{\mathbf{H}}) = \{v_iv_j : ij \in E(H)\}$ and let $Z_{\mathbf{v}}$ be the indicator random variable for the event $H_{\mathbf{H}} \in \mathbf{G}_{n,p}$. Thus $Z_{\mathbf{V}} = \sum_{\mathbf{v}\in\Gamma} Z_{\mathbf{v}}$ and, by the linearity of the expectation, we have

$$\begin{aligned}
\mathbb{E}\big[Z_{\mathbf{V}}\big] &= \sum_{\mathbf{v}\in\Gamma} \mathbb{E}\big[Z_v\big] \\
&= \sum_{\mathbf{v}\in\Gamma} p^{e(H)} \\
&= |V_1|\cdots|V_h| p^{e(H)} \\
&\geq \left(\frac{n}{\log n}\right)^h p^{e(H)} \\
&= \frac{c^{e(H)}}{(\log n)^h} \cdot n^{h-\alpha e(H)} =: \nu.
\end{aligned}$$

Let $\mathcal{A}$ be the family of graphs in K_n which are the union of two distinct copies of H which intersect at least in one edge. And let X be the number of elements in $\mathcal{A}$. Then we have that

$$\begin{aligned}
\mathbb{E}\big[X\big] &= \sum_{\substack{H_1,H_2\subseteq K_n \\ E(H_1\cap H_2)\neq\emptyset}} \mathbb{P}\big[H_1\cup H_2 \subseteq \mathbf{G}_{n,p}\big] \\
&\leq \sum_{K_2\subseteq J\subsetneq H} n^{2h-v(J)} p^{2e(H)-e(J)} \\
&\leq \sum_{K_2\subseteq J\subsetneq H} c^{2e(H)-e(J)} \cdot n^{2h-v(J)-\alpha(2e(H)-e(J))} \\
&\leq \sum_{K_2\subseteq J\subsetneq H} c^{2e(H)-e(J)} \cdot n^{h-\alpha e(H)-\delta(e(H)-e(J))} \qquad (13) \\
&= O\big(n^{h-\alpha e(H)-\delta}\big) \\
&= o\big(\nu\big) \\
&= o\big(\mathbb{E}\big[Z_{\mathbf{V}}\big]\big),
\end{aligned}$$

where we used the inequality (12). Therefore, by Markov's inequality (Theorem B.1), $X = o\big(\mathbb{E}\big[Z_H^\gamma\big]\big)$ w.h.p..

For two distinct $\mathbf{v}, \mathbf{w} \in \Gamma$, we write $\mathbf{v} \sim \mathbf{w}$ if $\mathbf{v} \neq \mathbf{w}$ and $E(H_{\mathbf{v}} \cap H_{\mathbf{w}}) \neq \emptyset$. In order to apply the Janson's inequality (Theorem B.5), let Δ be

$$\begin{aligned}
\Delta &= \sum_{\substack{\mathbf{v},\mathbf{w}\in\Gamma \\ \mathbf{v}\sim\mathbf{w}}} \mathbb{E}\big[Z_{\mathbf{v}} Z_{\mathbf{w}}\big] \\
&= \sum_{\substack{\mathbf{v},\mathbf{w}\in\Gamma \\ \mathbf{v}\sim\mathbf{w}}} \mathbb{P}\big[H_{\mathbf{v}} \cup H_{\mathbf{w}} \subseteq \mathbf{G}_{n,p}\big] \\
&\leq \mathbb{E}\big[X\big] \\
&= o\big(\mathbb{E}\big[Z_{\mathbf{V}}\big]\big),
\end{aligned}$$

since $H_{\mathbf{v}} \cup H_{\mathbf{w}} \in \mathcal{A}$, whenever that $\mathbf{v} \sim \mathbf{w}$.

Notice that if $A \in \mathcal{A}$, then as $e(A) \geq 2h-1$, there are at most $\binom{2h-1}{h} < 2^{2h}$ copies of H contained in A. This way, there are at most $2^{2h}X$ copies of H contained in some $A \in \mathcal{A}$, and if we remove all such copies, the remaining copies of H must be isolated copies. This implies that $Y_{\mathbf{V}} \geq Z_{\mathbf{V}} - 2^{2h}X$. Therefore, since $X = o\big(\mathbb{E}\big[Z_{\mathbf{V}}\big]\big)$ w.h.p., we must have $Y_{\mathbf{V}} \geq Z_{\mathbf{V}} - (1/4)\mathbb{E}\big[Z_{\mathbf{V}}\big]$ w.h.p.. Since $\Delta = o\big(\mathbb{E}\big[Z_{\mathbf{V}}\big]\big)$, applying the Janson's inequality, we get

$$\begin{aligned}
\mathbb{P}\big[Y_{\mathbf{V}} \leq \frac{1}{2}\mathbb{E}\big[Z_{\mathbf{V}}\big]\big] &\leq \mathbb{P}\big[Z_{\mathbf{V}} - \frac{1}{4}\mathbb{E}\big[Z_{\mathbf{V}}\big] \leq \frac{1}{2}\mathbb{E}\big[Z_{\mathbf{V}}\big]\big] \\
&= \mathbb{P}\big[Z_{\mathbf{V}} \leq \frac{3}{4}\mathbb{E}\big[Z_{\mathbf{V}}\big]\big] \\
&\leq 2\exp\left\{-\frac{\mathbb{E}\big[Z_{\mathbf{V}}\big]^2}{32(\mathbb{E}\big[Z_{\mathbf{V}}\big] + \Delta)}\right\} \\
&\leq 2\exp\left\{-\frac{1}{33}\mathbb{E}\big[Z_{\mathbf{V}}\big]\right\} \\
&\leq 2\exp\left\{-\frac{\nu}{33}\right\},
\end{aligned}$$

By the union bound for each choice of $\mathbf{V}$ (there are at most 2^{hn} choices, with room to spare), we have that the probability of having a family $\mathbf{V} = (V_1, \ldots, V_h)$ of disjoint subsets of $\mathbf{G}_{n,p}$ with $Y_{\mathbf{V}} \leq \frac{1}{2}\mathbb{E}\big[Z_{\mathbf{V}}\big]$ is at most

$$2^{hn}\exp\left\{-\frac{\nu}{33}\right\} = o\big(1\big),$$

since $h - \alpha e(H) > 1$. This finishes the proof of the lemma. $\square$

An observation is that we can relax inequality (11) to

$$\frac{1}{m_2(H)} \leq \alpha < \frac{h-1}{e(H)}, \tag{14}$$

provided that H is strictly 2-balanced. Indeed, if H is strictly 2-balanced, then there is some $\eta > 1/m_2(H)$ such that for every $J \subsetneq H$ with $v(J) \geq 3$ we have

$$\frac{1}{m_2(H)} = \frac{v(H)-2}{e(H)-1} < \eta < \frac{v(J)-2}{e(J)-1},$$

Then, for all $J \subsetneq H$ with $v(J) \geq 3$, it follows that

$$\frac{v(H)-v(J)}{e(H)-e(J)} < \eta < \alpha - \delta.$$

Therefore,

$$2h - v(J) - \alpha(2e(H) - e(J)) \leq h - \alpha e(H) + (\eta - \alpha)(e(H) - e(J)).$$

Since $\eta - \alpha < -\delta$,

$$\begin{aligned} n^{2h-v(J)-\alpha(2e(H)-e(J))} &\leq n^{h-\alpha e(H)+(\eta-\alpha)(e(H)-e(J))} \\ &= o(n^{h-\alpha e(H)+\delta(e(H)-e(J))}), \end{aligned}$$

which is stronger than inequality (13). Therefore, the proof follows the same.

In order to apply Lemma 4.2 and 4.3 for some $p = cn^{-\alpha}$, we must have H being 2-balanced and choose α such that

$$\frac{1}{m_2(H)} < \alpha < \frac{v(H)-1}{e(H)}. \tag{15}$$

Furthermore, if H is a 2-balanced graph and if F is graph with $1 < m_2(F) < m_2(H)$, then taking $\alpha = 1/m_2(H,F)$, since we have $m_2(H,F) < m_2(H)$ and

$$m_2(H,F) \geq \frac{e(H)}{v(H) - 2 + 1/m_2(F)} > \frac{e(H)}{v(H)-1},$$

it follows that such α satisfies the inequalities in (15).

Now, if H is strictly 2-balanced and $1 < m_2(F) \leq m_2(H)$, then $\alpha = 1/m_2(H,F)$ must satisfy the inequality (14), since $m_2(H,F) \leq m_2(H)$.

4.2 Proof of Theorem 4.1

So, let H be 2-balanced graph with h vertices and F a graph with $m_2(H) > m_2(F) > 1$ (or let H be strictly 2-balanced with $m_2(H) \geq m_2(F) > 1$). We will show that there exists a positive constant C such that if $p \geq Cn^{-1/m_2(H,F)}$, then $\mathbb{P}\big[\mathbf{G}_{n,p} \nrightarrow (H,F)\big] = o\big(1\big)$. To do so, suppose we have a colouring of $\mathbf{G}_{n,p}$ avoiding monochromatic copies of H of the first colour and monochromatic copies of F of the second colour. Therefore, in this colouring, each isolated copy of H must contain an edge of the second colour. Let J be the set of all the edges contained in some isolated copy of H and coloured with the second colour. In order to yield a contradiction, we shall show that w.h.p. we can find a copy of F using the edges from J.

Let q as in equation (9). As it was told before, by taking $\alpha = 1/m_2(H,F)$, we can apply both Lemma 4.2 and 4.3 for our choice of p, no matter which positive constant C we take.

Now, let us define the constants we will carry with us until the end of the proof. We start letting γ and k_0 be the constants corresponding, respectively, to ε and n_0 given by Theorem 3.1 when applied to the graphs K_2, H and F (and $r = 3$). Then, take

$$\delta = \frac{\gamma}{2^h B}, \ \beta = \left(\frac{\delta}{e^2}\right)^{e(F)} \text{ and } D = 6e(H).$$

Let C, n_0, and ε be the constants we get from Theorem 2.6 by means of F and β. We can assume that $\varepsilon < \gamma$. Let η and K_0 be the constants given by the regularity lemma for sparse graphs (Theorem 2.2) from the constants ε, k_0 and $D = 6e(H)$.

We know from Lemma 4.2, that w.h.p., the graph $H(\mathbf{G}_{n,p})$ is (η, D, q)-upper-uniform. So $G = ([n], J)$, is still (η, D, q)-upper-uniform, since $G \subseteq H(\mathbf{G}_{n,p})$. Therefore, w.h.p. there is an (ε, q)-regular partition $V_1, \ldots, V_k$ of G with $k_0 \leq k \leq K_0$. Now, consider the following 3-colouring, $c : E(K_K) \to [3]$, of the edges of the complete graph K_k. For each $ij \in E(K_k)$, put

(i) $c(ij) = 1$, if (V_i, V_j) *is not* an (ε, q)-regular pair in G;
(ii) $c(ij) = 2$, if (V_i, V_j) is an (ε, q)-regular pair with $d(V_i, V_j) < \delta q$;
(iii) $c(ij) = 3$, if (V_i, V_j) is an (ε, q)-regular pair with $d(V_i, V_j) \geq \delta q$.

By Theorem 3.1, since $k \geq k_0$, we must have at least one of the following:

(A.1) there exist γk^2 pairs (V_i, V_j) that are not (ε, q)-regular;
(A.2) there exist $\gamma k^{v(H)}$ monochromatic copies of H of colour 2;
(A.3) there exist $\gamma k^{v(F)}$ monochromatic copies of F of colour 3.

Since $V_1, \ldots, V_k$ is an (ε, q)-regular partition and $\varepsilon < \gamma$, (A.1) can not happen. Let us show that w.h.p. (A.2) can not happen as well. For this purpose, let S be the union of $E_G(V_i, V_j)$ for all $ij \in E(K_k)$ with $c(ij) = 2$. Since $e_G(V_i, V_j) < \delta q |V_i||V_j| \leq \delta q (n/k)^2$, whenever that $d(V_i, V_j) < \delta q$, we must have

$$\begin{aligned} |S| &= \sum_{\substack{ij \in E(K_n) \\ c(ij)=2}} e_G(V_i, V_j) \\ &< \binom{k}{2} \delta q \left(\frac{n}{k}\right)^2 \\ &\leq \frac{\delta}{2} q n^2 \\ &\leq \left(\frac{\delta B}{2}\right) n^h p^{e(H)}. \end{aligned} \tag{16}$$

On the other hand, if H' is a copy of H in K_k with $v(H') = \{v_1, \ldots, v_h\}$, then $\mathbf{V}_{H'} = (V_{v_1}, \ldots, V_{v_h})$ is a family of disjoint subsets of $V(\mathbf{G}_{n,p})$ with $|V_{v_i}| \geq n/(2k) \geq n/(2K_0) \geq n/\log n$, for each $i \in [h]$. Therefore, applying Lemma 4.3 to $\mathbf{V}_{H'}$, we get that w.h.p. $Y_{\mathbf{V}}$, the number of $\mathbf{V}_{H'}$-copies of H which are isolated copies in $\mathbf{G}_{n,p}$, is bounded by

$$\begin{aligned} Y_{\mathbf{V}_{H'}} &\geq \frac{1}{2} \mathbb{E}[Z_{\mathbf{V}_{H'}}] \\ &= \frac{1}{2} |V_{v_1}| \cdots |V_{v_h}| p^{e(H)} \\ &\geq \frac{1}{2} \left(\frac{n}{2k}\right)^h p^{e(H)}. \end{aligned}$$

This way, if H' is a monochromatic copy of H in K_k with colour 2, then for each $v_iv_j \in E(H')$, we have $E(V_{v_i}, V_{v_j}) \subseteq S$. Furthermore, the $\mathbf{V}_{H'}$-copies of H which are isolated do not share any edge with each other. Therefore, one edge from each $\mathbf{V}_{H'}$-copies of H which are isolated copies of H in $\mathbf{G}_{n,p}$ is an edge in G. Let $\mathcal{H}$ be the family of all monochromatic copies of H in K_k with the colour 2. Since monochromatic copies of H generate distinct $\mathbf{V}_{H'}$-copies, we must have

$$\begin{aligned} |S| &\geq \sum_{H' \in \mathcal{H}} Y_{\mathbf{V}_{H'}} \\ &\geq \gamma k^h \left(\frac{1}{2^{h+1}k^h} \right) n^h p^{e(H)} \\ &= \left(\frac{\delta B}{2} \right) n^h p^{e(H)}. \end{aligned} \tag{17}$$

This lead us to a contradiction between inequality (16) and (17). Therefore, we must have w.h.p. no monochromatic copy of H of colour 2.

Therefore, w.h.p., we must have case (A.3). In this case, one copy F' of F in K_k with the colour 3 must be enough. Let $v_1, \ldots, v_{v(F)}$ be the vertices of F' and let $\mathbf{V} = (V_{v_1}, \ldots, V_{v_{v(F)}})$ be the family of the correspondent parts in G. So we have $n/2K_0 \leq |V_{v_i}| \leq n/k_0$ and (V_{v_i}, V_{v_j}) is an (ε, q)-regular pair with $d(V_{v_i}, V_{v_j}) \geq \delta q$, whenever that $v_iv_j \in E(F)$. Therefore, there is a subgraph $G_0 \subseteq G[\mathbf{V}]$ that belongs to $\mathcal{G}(F, n_1, m, q, \varepsilon)$, for some $n_1 \geq n/2K_0$ and $m \geq \delta q n_1^2$. In order to finish the proof of the theorem, we shall show that w.h.p., there is no graph from $\mathcal{G}^*(F, n_1, m, q, \varepsilon)$ contained in $G[\mathbf{V}]$. In this way, we have that G_0 contains a copy of F. This copy, since $G_0 \subseteq G$, must have all the edges in J, as we wanted.

Let X be the number of subgraphs of $G[\mathbf{V}]$ that belongs to $\mathcal{G}^*(F, n_1, m, q, \varepsilon)$. As $m \geq \delta q n_1^2$, then by Theorem 2.6, we have

$$\begin{aligned} \mathbb{E}\big[X\big] &\leq \beta^m \binom{n_1^2}{m}^{e(F)} \cdot q^{m \cdot e(F)} \\ &\leq \left(\frac{\delta}{e^2} \right)^{m \cdot e(F)} \cdot \left(\frac{e n_1^2}{m} \right)^{m \cdot e(F)} \cdot q^{m \cdot e(F)} \\ &= \left(\frac{\delta q n_1^2}{e m} \right)^{m \cdot e(F)} \\ &\leq \exp\left\{-m \cdot e(F)\right\} \\ &= \exp\left\{-\Omega\big(qn^2\big)\right\}, \end{aligned}$$

since $m \geq \delta q n_1^2 = \Omega\big(qn^2\big)$. Thus, by Markov's inequality (Theorem B.1), w.h.p., there is no graph from $\mathcal{G}^*(F, n_1, m, q, \varepsilon)$ contained in G. Since we chose n_1 and m early, we should take in account the number of ways we can choose those parameters. Certainly, we have at most n choices for n_1 and $\binom{n}{2} = O(n^2)$ choices for m. This give us $O(n^3)$ choices for n_1 and m. So, the probability of having no $G_0 \subseteq G$ belonging to $\mathcal{G}(F, n_1, m, q, \varepsilon) \setminus \mathcal{G}^*(F, n_1, m, q, \varepsilon)$, for some $n_1 \geq n/2K_0$ and some $m \geq \delta q n_1^2$ is bounded by

$$O(n^3) \exp\left\{-\Omega\big(qn^2\big)\right\}.$$

Since $qn^2 \gg n$, we have that this bound is $o\big(1\big)$. Therefore, w.h.p., there is a graph $G_0 \subseteq G$ belonging to $\mathcal{G}(F, n_1, m, q, \varepsilon)$ but not belonging to $\mathcal{G}^*(F, n_1, m, q, \varepsilon)$. Furthermore, we have a copy of F contained in G, as stated. This concludes the proof of Theorem 4.1. $\square$

5 PROVING THE KŁR CONJECTURE

In this section, we show a proof for the KŁR conjecture (Theorem 2.6) using the hypergraph container lemma. We follow the proof given by Balogh, Morris, and Samotij (2015). The proof is essentially a typical application of the hypergraph container lemma, where we first establish a supersaturation result in order to guarantee the necessary conditions to apply the method for a particular hypergraph in which the graphs in $\mathcal{G}^*(H, n, m, m/n^2, \varepsilon)$ correspond to independent sets in such hypergraph, and then we count the number of independent sets in such hypergraph by means of the containers.

Before the proof, let us introduce a definition which will follow us throughout this proof. We say that a pair of disjoint sets of vertices (U, V) is (α, δ)-uniformly-dense if for every $U' \subseteq U$ and $V' \subseteq V$ with $|U'| \geq \alpha|U|$ and $|V'| \geq \alpha|V|$, we have $d(U', V') \geq \delta$.

The following lemma essentially says that, for a fixed graph H, if G is a subgraph of H^n, the complete blow-up of H of order n, and G has few copies of H, then there must exist some pair (V_i, V_j) among the independent parts of H^n which is not too uniformly-dense in G. Thus, this fact give us a characterization of those blown-up graphs with few copies of H. This lemma can be found in Balogh, Morris, and Samotij (2015). An observation is that this lemma is interesting when δ is a function for which $\delta(x)$ is much smaller than x.

Lemma 5.1. *Let H be a simple graph with vertex set $[h]$ and let $\delta : (0,1) \to (0,1)$ be an arbitrary function. There must exist positive constants α_0 and ξ and a positive integer n_0 such that the following must happens. Let H^n be the complete blow-up of H of order $n \geq n_0$ and let $V_1, \ldots, V_h$ be the independent parts of H^n. If $G \subseteq H^n$, then one of the following holds:*

- *(a) G contains at least ξn^h copies of H;*
- *(b) There exists some $\alpha \in (\alpha_0, 1)$ and some edge $ij \in E(H)$ such that (V_i, V_j) is not $(\alpha, \delta(\alpha))$-uniformly-dense with respect to G.*

Proof. The proof is by induction on h. For $h = 1$, there is nothing to prove. So let us say we have $h \geq 2$ and suppose the lemma is valid for $H' = H - h$, the simple graph on $[h-1]$ obtained removing the vertex h from H. Let $\alpha_1 = 1/h$. So we apply the lemma inductively for H' and for the function $\delta'(x) = \delta(\delta(\alpha_1) \cdot x)$, which return to us the constants α_0', ξ' and n_0'. Let

$$\alpha_0 = \min\{\alpha_1, \alpha_0' \cdot \delta(\alpha_1)\}, \quad \xi = \alpha_1 \cdot \delta(\alpha_1)^{h-1} \cdot \xi', \quad \text{and} \quad n_0 = \frac{n_0'}{\delta(\alpha_1)}.$$

Now let $n \geq n_0$, let $V_1, \ldots, V_h$ be the independent parts of H^n and let $G \subseteq H^n$. For each $i \in N_H(i)$, the neighborhood of h in H, consider the set

$$W_i = \{v \in V_h : \deg_G(v, V_i) < \delta(\alpha_1)|V_i|\},$$

where $\deg_G(v, V_i)$ is the number of neighbors of v in G contained in V_i. Since the density of the pair (V_i, W_i) in G is

$$d_G(V_i, W_i) = \frac{e_G(V_i, W_i)}{|V_i||W_i|} \leq \frac{\delta(\alpha_1)|V_i||W_i|}{|V_i||W_i|} = \delta(\alpha_1),$$

if we have $|W_i| \geq \alpha_1|V_h|$, then (V_i, V_h) is not $(\alpha_1, \delta(\alpha_1))$-uniformly-dense, and therefore we have the item (b). So we can assume that $|W_i| \leq \alpha_1|V_h|$, for each $i \in N(h)$. Let

$$W = V_h \setminus \bigcup_{i \in N_H(h)} W_i = \big\{v \in V_h : \deg(v, V_i) \geq \delta(\alpha_1)|V_i|, \forall i \in N_H(h)\big\},$$

Then we have $|W| \geq (1 - \Delta(H)\alpha_1)n \geq \alpha_1 n$, since $\alpha_1 = 1/h$.

Now, for each $v \in W$, consider the set

$$V_i(v) = \begin{cases} V_i \cap N_G(v), & \text{if } i \in N_H(h); \\ V_i, & \text{otherwise.} \end{cases}$$

Thus, we have $|V_i(v)| \geq \delta(\alpha_1)|V_i| = \delta(\alpha_1)n$, for each $i \in [h-1]$. Removing some vertices if necessary, we can assume that $|V_i(v)| = \delta(\alpha_1)n =: n'$, for each $i \in [h-1]$. Let $G(v)$ be the graph $G[V_1(v), \ldots, V_{h-1}(v)] \subseteq (H')^{n'}$. Since $\delta(\alpha_1)n_0 = n_0'$, we have $n' \geq n_0'$.

Therefore, for each $v \in W$, we can apply the induction hypothesis to $G(v)$. Suppose that for some $v \in W$, it happens that when we apply the induction hypothesis to $G(v)$, the item (b) holds. Then there exist $\alpha' \in (\alpha_0', 1)$ and some $ij \in E(H')$ such that $(V_i(v), V_j(v))$ is not $(\alpha', \delta'(\alpha'))$-uniformly-dense with respect to $G(v)$. Then for each $k \in \{i, j\}$, there must exist $\tilde{V}_k \subseteq V_k(v)$ with $|\tilde{V}_i| \geq \alpha'|V_k(v)|$ such that

$$d_{G(v)}(\tilde{V}_i, \tilde{V}_j) < \delta'(\alpha').$$

Since $|V_k(v)| \geq \delta(\alpha_1)|V_k|$ and $\alpha'\delta(\alpha_1) \geq \alpha_0'\delta(\alpha_1) \geq \alpha_0$, then taking $\alpha = \alpha'\delta(\alpha_1)$, we have $|\tilde{V}_k| \geq \alpha|V_k|$, for each $k \in \{i, j\}$. Since $V_k(v) \subseteq V_k$ and $e_G(\tilde{V}_i, \tilde{V}_j) = e_{G(v)}(\tilde{V}_i, \tilde{V}_j)$, we have

$$d_G(\tilde{V}_i, \tilde{V}_j) = d_{G(v)}(\tilde{V}_i, \tilde{V}_j) < \delta'(\alpha') = \delta(\delta(\alpha_1) \cdot \alpha') = \delta(\alpha).$$

Therefore, (V_i, V_j) is not an $(\alpha, \delta(\alpha))$-uniformly-dense graph with respect to G.

Thereby, we may assume that for each $v \in W$, $G(v)$ has at least $\xi'(n')^{h-1}$ copies of H'. Those copies must form a copy of H with v. Therefore, we have at least $|W|\xi'(n')^{h-1}$ copies of H in G. Since $n' \geq \delta(\alpha_1)n$, we have at least

$$|W|\xi'(n')^{h-1} \geq \alpha_1\xi'\delta(\alpha_1)^{h-1}n^h = \xi n^h$$

copies of H. This finishes the proof of the lemma. $\square$

Given $\beta > 0$, let us consider the function

$$\delta_\beta(x) = \frac{1}{4e}\left(\frac{\beta}{2}\right)^{2/x^2}.$$

The following lemma essentially tell us a good upper bound on how many (ε, p)-regular graphs containing all the edges of a small set S we can have between the parts V_i and V_j which is not $(\alpha, \delta_\beta(\alpha))$-uniformly-dense. This will allow us to bound the number of graphs covered by a container.

Lemma 5.2. *Let $\alpha_0 < 1$ and $\beta \leq 1$ be positive constants. Then there exists $\varepsilon > 0$ such that the following holds. Suppose $G = G[V_1, V_2]$ is a bipartite graph with class-sets V_1 and V_2 such that $|V_1| = |V_2| = n$ and let $m \leq n^2$. If for some $\alpha \in (\alpha_0, 1)$, the pair (V_1, V_2) is not $(\alpha, \delta_\beta(\alpha))$-uniformly-dense, then for any $S \subseteq E_G(V_1, V_2)$ with $|S| \leq \varepsilon m$, the number of subgraphs $G' \subseteq G$ belonging to $\mathcal{G}^*(K_2, n, m, p, \varepsilon)$, for $p = m/n^2$, and such that $S \subseteq E(G')$ is bounded by*

$$\beta^m \binom{n^2}{m - |S|}.$$

Proof. Given α_0 and β, set $\varepsilon = \min\{\alpha_0, 1/2\}$ and let G be a graph as in the statement of the lemma. Notice that we have $\delta(\alpha_0) < 1/8$. Since for some $\alpha \in (\alpha_0, 1)$ the pair (V_1, V_2) is not $(\alpha, \delta(\alpha))$-uniformly-dense, there exists a pair (U_1, U_2) with $U_i \subseteq V_i$ and $|U_i| \geq \alpha n$, for each $i \in \{1, 2\}$, such that $d_G(U_1, U_2) < \delta(\alpha)$. For each $i \in \{1, 2\}$, let U_i' be a subset of U_i with αn vertices chosen uniformly at random. The expected number of $e_G(U_1', U_2')$ is then given by

$$\begin{aligned}
\mathbb{E}\big[e_G(U_1', U_2')\big] &= \sum_{xy \in E_G(U_1, U_2)} \mathbb{P}\big[x \in U_1'\big]\mathbb{P}\big[y \in U_2'\big] \\
&= \sum_{xy \in E_G(U_1, U_2)} \frac{\binom{|U_1|-1}{\alpha n - 1}}{\binom{|U_1|}{\alpha n}} \cdot \frac{\binom{|U_2|-1}{\alpha n - 1}}{\binom{|U_2|}{\alpha n}} \\
&= e_G(U_1, U_2) \cdot \frac{\alpha n}{|U_1|} \cdot \frac{\alpha n}{|U_2|} \\
&= e_G(U_1, U_2) \cdot \frac{|U_1'||U_2'|}{|U_1||U_2|}.
\end{aligned}$$

Thus $\mathbb{E}\big[d_G(U_1', U_2')\big] = d_G(U_1, U_2)$. Therefore we must have, for each $i \in \{1, 2\}$, some $U_i' \subseteq U_i$ with $|U_i'| = \alpha n$ and such that $d_G(U_1', U_2') < \delta(\alpha)$. In particular, we have $e_G(U_1', U_2') < \delta(\alpha) \cdot \alpha^2 n^2$.

Now, let S be a set of edges of G as in the statement of the lemma. Let $\mathcal{G}^*$ be the set of graphs G' such that $S \subseteq E(G')$ and $G' \in \mathcal{G}^*(K_2, n, m, p, \varepsilon)$. For each $G' \in \mathcal{G}^*$, notice that we must have $d_{G'}(U_1', U_2') \geq (1 - \varepsilon)p$, since (V_1, V_2) is (ε, p)-regular with respect to G', $p = d_{G'}(V_1, V_2)$, and $|U_i'| = \alpha n \geq \alpha_0 n \geq \varepsilon |V_i|$, for each $i \in \{1, 2\}$. Let $\varepsilon' = \varepsilon(1 - 1/\alpha^2) \leq 1/2$. Then, $e_{G'}(U_1', U_2') \geq (1 - \varepsilon)p|U_1'||U_2'| = (1 - \varepsilon)\alpha^2 m$, and since

$|S| \le \varepsilon m$, (U_1', U_2') has at least $(1-\varepsilon-\varepsilon/\alpha^2)\alpha^2 m = (1-\varepsilon')\alpha^2 m \ge (\alpha^2/2)m$ edges of $E(G') \setminus S$. Furthermore, G' must have at least $m - |S| - e_{G'}(U_1', U_2')$ edges belonging to $E(G) \setminus (E_G(U_1', U_2') \cup S)$. Therefore, a graph G' in $\mathcal{G}^*$ can be built by choosing $\ell \ge (\alpha^2/2)m$ edges from $E_G(U_1', U_2')$ to belong to $E(G') \setminus S$ and choosing at least $m - |S| - \ell$ from $E(G) \setminus (E_G(U_1', U_2') \cup S)$ to belong to $E(G') \setminus (E_G(U_1', U_2') \cup S)$. This give us the following bound on the number of graphs in $\mathcal{G}^*$:

$$\begin{aligned} |\mathcal{G}^*| &\le \sum_{\ell \ge (\alpha^2/2)m} \binom{e_G(U_1', U_2')}{\ell} \binom{e(G) - e_G(U_1', U_2')}{m - |S| - \ell} \\ &\le \sum_{\ell \ge (\alpha^2/2)m} \binom{\delta(\alpha) \cdot \alpha^2 n^2}{\ell} \binom{n^2}{m - |S| - \ell} \\ &\le \sum_{\ell \ge (\alpha^2/2)m} \binom{\delta(\alpha) \cdot \alpha^2 n^2}{\ell} \left(\frac{m - |S|}{n^2 - m + |S|}\right)^{\ell} \cdot \binom{n^2}{m - |S|}, \end{aligned}$$

where we used Proposition A.4 in the last inequality. Notice that if we have $m > 2\delta(\alpha)n^2$, then the right-hand-side of the inequality above vanishes, once in this case we have $\ell \ge (\alpha^2/2)m > \delta(\alpha) \cdot \alpha^2 n^2$. Thus, let us assume $m \le 2\delta(\alpha)n^2$. In particular, $m - |S| \le m \le 2\delta(\alpha)n^2 \le n^2/2$, since $\delta(\alpha) < 1/4$ and $\alpha < 1$. Therefore, we can slightly relax the bound above to

$$\begin{aligned} |\mathcal{G}^*| &\le \sum_{\ell \ge (\alpha^2/2)m} \left(\frac{e\delta(\alpha)\alpha^2 n^2}{\ell}\right)^{\ell} \cdot \left(\frac{2m}{n^2}\right)^{\ell} \binom{n^2}{m - |S|} \\ &= \sum_{\ell \ge (\alpha^2/2)m} \left(\frac{2e\delta(\alpha) \cdot \alpha^2 m}{\ell}\right)^{\ell} \binom{n^2}{m - |S|}. \end{aligned}$$

Now, by Proposition A.5, the real-valued function $f(x) = (a/x)^x$, defined for $x > 0$ (for some given $a > 0$), is decreasing for $x > a/e$. In particular, for $a = 2e\delta(\alpha) \cdot \alpha^2 m$, we have $x = (\alpha^2/2)m > 2\delta(\alpha)\alpha^2 m = a/e$, once $\delta < 1/4$. Therefore, each summand in the right-hand-side of the last inequality is bounded by the first one. This way, we have

$$\begin{aligned} |\mathcal{G}^*| &\le m \left(4e\delta(\alpha)\right)^{\alpha^2 m/2} \binom{n^2}{m - |S|} \\ &\le \beta^m \binom{n^2}{m - |S|}, \end{aligned}$$

since $\beta/2 = (4e\delta(\alpha))^{\alpha^2/2}$. This finishes the proof of the lemma. $\square$

Now, we can prove the KŁR conjecture.

Proof of Theorem 2.6. Let us start defining some constants. Suppose $V(H) = [h]$ and let $\beta > 0$ be the constant given by hypothesis. So, applying Lemma 5.1 for H and the function

$$\delta_{\beta/2}(x) = \frac{1}{4e}\left(\frac{\beta}{4}\right)^{2/x^2},$$

we get $\alpha_0, \xi > 0$ and n_0. From Lemma 5.2 by means of α_0 and $\beta/2$, we get ε. We can assume ε is small enough so that

$$\left(\frac{2e \cdot e(H)}{\varepsilon}\right)^{\varepsilon} < \frac{4}{3}. \tag{18}$$

Let us leave C to be determined for the last. So, let $m \geq Cn^{2-1/m_2(H)}$.

In order to apply the hypergraph container theorem (Theorem 2.7), let $\mathcal{H}$ be the hypergraph with vertex-set being the edges of H^n and the hyperedges being the edge-set of canonical copies of H in H^n. So, suppose $n \geq n_0$ and let

$$\mathcal{F} = \big\{G \subseteq H^n : \ G \text{ contains at least } \xi n^h \text{ canonical copies of H}\big\}.$$

In particular, we have that $\mathcal{F}$ is an increasing family and that $\mathcal{H}$ is $(\mathcal{F}, \xi)$-dense. Taking $p = n^{-1/m_2(F)}$, there must exist a positive c for which

$$\Delta_\ell(\mathcal{H}) \leq c \cdot p^{\ell-1}\frac{e(\mathcal{H})}{v(\mathcal{H})},$$

for each $\ell \in [k]$. Indeed, the demonstration of this fact is similar to that from the proof of Theorem 2.8 and therefore we will omit it here (another way for checking the inequality above is just noticing that $\mathcal{H}$ is a sub-hypergraph of the hypergraph defined in the proof of Theorem 2.8 with a asymptotically positive fraction of the vertices).

Thereby, we can apply the hypergraph container theorem (Theorem 2.7). Thus we have a positive constant C', a family $\mathcal{S} \subseteq \mathcal{P}(V(\mathcal{H}))$ with $|S| \leq C'p \cdot v(\mathcal{H})$, for each $S \in \mathcal{S}$, and functions $f : \mathcal{S} \to \overline{\mathcal{F}}$ and $g : \mathcal{I}(\mathcal{H}) \to \mathcal{S}$ such that $g(I) \subseteq I$ and $I \setminus g(I) \subseteq f(g(I))$, for every $I \in \mathcal{I}(\mathcal{H})$. We, then, take C to be $C = C'/\varepsilon$.

Now, notice that $\mathcal{G}^* = \mathcal{G}^*(H, n, m, m/n^2, \varepsilon) \subseteq \mathcal{I}(\mathcal{H})$. For each $S \in \mathcal{S}$, let $\mathcal{G}^*_S = \{G \in \mathcal{G}^* : g(G) = S\}$. Let G_S be the subgraph of H^n with edge-set $f(S)$. Thus $G_S \in \overline{\mathcal{F}}$. From Lemma 5.1, there exists $\alpha \in (\alpha_0, 1)$ and an edge $ij \in E(H)$ such that (V_i, V_j) is not $(\alpha, \delta(\alpha))$-uniformly-dense with respect to G_S. Let $S_{i,j} = S \cap E(V_i, V_j)$. Notice that since we have $m \geq Cn^{2-1/m_2(H)}$, we also have

$$|S_{i,j}| \leq |S| \leq C'p \cdot v(\mathcal{H}) = C'n^{2-1/m_2(H)} = \varepsilon \cdot Cn^{2-1/m_2(H)} \leq \varepsilon m.$$

Therefore, we can apply Lemma 5.2 to $S_{i,j}$ and $G_S[V_i, V_j]$. We conclude, thus, that there exist at most

$$\left(\frac{\beta}{2}\right)^m \binom{n^2}{m-|S_{i,j}|}$$

choices from the edges of a graph $G' \in \mathcal{G}^*(K_2, n, m, m/n^2, \varepsilon)$ with $S_{i,j} \subseteq E(G')$ and $G' \subseteq G_S[V_i, V_j]$. It follows that

$$|\mathcal{G}^*_S| \le \left(\frac{\beta}{2}\right)^m \prod_{ij\in E(H)} \binom{n^2}{m-|S_{i,j}|}.$$

Then since $\mathcal{G}^* = \cup_{S\in\mathcal{S}}\mathcal{G}^*_S$, we have

$$\begin{aligned}
|\mathcal{G}^*| &\le \sum_{S\in\mathcal{S}} |\mathcal{G}^*_S| \\
&\le \sum_{S\in\mathcal{S}} \left(\frac{\beta}{2}\right)^m \prod_{ij\in E(H)} \binom{n^2}{m-|S_{i,j}|} \\
&\le \sum_{S\in\mathcal{S}} \left(\frac{\beta}{2}\right)^m \prod_{ij\in E(H)} \left(\frac{m}{n^2-m}\right)^{|S_{i,j}|} \binom{n^2}{m} \\
&= \left(\frac{\beta}{2}\right)^m \binom{n^2}{m}^{e(H)} \sum_{S\in\mathcal{S}} \left(\frac{m}{n^2-m}\right)^{|S|} \qquad (19)
\end{aligned}$$

Let us, now, give a bound on the summation in the right hand side of the inequality above.

$$\begin{aligned}
\sum_{S\in\mathcal{S}} \left(\frac{m}{n^2-m}\right)^{|S|} &\le \sum_{s\le\varepsilon m} \binom{e(H)n^2}{s} \left(\frac{2m}{n^2}\right)^s \\
&\le \sum_{s\le\varepsilon m} \left(\frac{e\cdot e(H)n^2}{s}\right)^s \left(\frac{2m}{n^2}\right)^s \\
&= \sum_{s\le\varepsilon m} \left(\frac{2e\cdot e(H)m}{s}\right)^s \\
&\le \varepsilon m \left(\frac{2e\cdot e(H)}{\varepsilon}\right)^{\varepsilon m},
\end{aligned}$$

where we used in the last inequality that the summand is increasing in s (see Proposition A.5). Now, once we have ε satisfying the inequality (18), and since $m \le (3/2)^m$, we get that

$$\sum_{S\in\mathcal{S}}\left(\frac{m}{n^2-m}\right)^{|S|} \leq \varepsilon m\left(\frac{2e\cdot e(H)}{\varepsilon}\right)^{\varepsilon m}$$
$$\leq \varepsilon\left(\frac{3}{2}\right)^{m}\cdot\left(\frac{4}{3}\right)^{m}$$
$$\leq 2^m.$$

Combining the inequality above with the inequality (19), we get that

$$|\mathcal{G}^*| \leq \beta^m \binom{n^2}{m}^{e(H)},$$

as stated. This finishes the proof of Theorem 2.6. □

6 CONCLUSION

Recent developments of techniques such as the hypergraph container method (Theorem 2.7) has allowed a major improvement towards the theory of Ramsey properties on random graphs. The method itself has a natural description and it is easily applicable to many problems in extremal and probabilistic combinatorics, as we outlined in Section 2 and as we illustrated in the proof of the 1-statement of Theorem 1.3 (which can be found in the beginning of Section 3). In Section 4, we showed that the lower bound on the threshold function for the Ramsey property conjectured by Kohayakawa & Kreuter (Conjecture 1.4) can be established for pairs of graphs under some balancing condition (see Theorem 1.7) by using the KŁR Conjecture (Theorem 2.6). And, finally, we saw in Section 5 how the KŁR Conjecture can be proven by using the hypergraph container method.

APPENDIX

A — BASIC INEQUALITIES

In the following, we state some basic inequalities we have used along the text. All of them can be easily proven through some basic calculation.

Proposition A.1. *For $a, b, c, d \in \mathbb{R}$ with $b > d > 0$, we have*

$$\frac{a}{b} \geq \frac{c}{d} \iff \frac{a-c}{b-d} \geq \frac{a}{b}.$$

Proposition A.2. *Let x be a real number. Then*

$$1 + x \leq \mathrm{e}^x.$$

Proposition A.3. *Let a and b be positive integers with $a \geq b$. Then*

$$\binom{a}{b} \leq \left(\frac{\mathrm{e}a}{b}\right)^b.$$

Proposition A.4. *Let a, b and c be positive integers with $a \geq b \geq c$. Then*

$$\binom{a}{b-c} \leq \left(\frac{b}{a-b}\right)^c \binom{a}{b}.$$

Proposition A.5. *Let $f : (0, +\infty) \to \mathbb{R}$ be the real function given by $f(x) = (a/x)^x$, for some positive real constant a. If $0 \leq x \leq y \leq a/\mathrm{e}$ or if $x \geq y \geq a/\mathrm{e}$, then $f(x) \leq f(y)$. In particular, the maximum of $f(x)$ on $(0, +\infty)$ is reached in $x = a/\mathrm{e}$.*

B — PROBABILISTIC INEQUALITIES

In the following, we state all the probabilistic inequalities we have used along the text. The proof for all of them can be found in details in Janson, Łuczak, and Ruciński (2011).

Theorem B.1 (Markov's Inequality)**.** *Let X be a non-negative random variable. Then, for every $t > 0$, we have*

$$\mathbb{P}\big[X \geq t\big] \leq \frac{\mathbb{E}\big[X\big]}{t}.$$

Theorem B.2 (Chebyshev's Inequality)**.** *Let X be a random variable with expectation μ and variance σ^2. Then, for every $t > 0$, we have*

$$\mathbb{P}\big[|X - \mu| \geq t\big] \leq \frac{\sigma^2}{t^2}.$$

In the following, let ϕ be the real value function defined for $x > -1$ by

$$\phi(x) = (1 + x)\log(1 + x) - x.$$

Theorem B.3 (Chernoff's Inequality)**.** *Let X be a binomial random variable and let μ be the expectation of X. Then, for every $t \geq 0$, we have*

$$\mathbb{P}\big[X \geq \mu + t\big] \leq \exp\left\{-\mu \cdot \phi\left(\frac{t}{\mu}\right)\right\} \leq \exp\left\{-\frac{t^2}{2(\mu + t/3)}\right\},$$

and

$$\mathbb{P}\big[X \leq \mu - t\big] \leq \phi \exp\left\{-\mu \cdot \phi\left(\frac{-t}{\mu}\right)\right\} \leq \exp\left\{-\frac{t^2}{2\mu}\right\}.$$

Corolary B.4. *Let X be a binomial random variable and let μ be the expectation of X. Then, for every $\varepsilon \geq 0$, we have*

$$\mathbb{P}\big[|X - \mu| \geq \varepsilon\mu\big] \leq 2\exp\left\{-\phi(\varepsilon)\mu\right\}.$$

In particular, if $\varepsilon \leq 3/2$, then

$$\mathbb{P}\big[|X - \mu| \geq \varepsilon\mu\big] \leq 2\exp\left\{-\frac{\varepsilon^2}{3}\mu\right\}.$$

Therefore, if X_n is a sequence of binomial random variables with expectation $\mu_n \gg 1$, then for any $\varepsilon > 0$, we must have $X_n = (1 \pm \varepsilon)\mu_n$ asymptotically almost surely.

Let E be a finite set. Suppose that we have a sequence of random variables i.i.d. $(X_e)_{e \in E}$ with distribution Bernoulli(p). We think in the sequence $(X_e)_{e \in E}$ as a sequence of indicator random variables for a random process in which each element $e \in E$ is chosen at random with probability p and independently from other elements of E. For each subset $A \subseteq E$, consider $X_A = \prod_{e\in A} X_e$ as the indicator random variable for the event that every element from A is chosen. Now, let $\mathcal{A}$ be a family of subsets of E and let X be the random variable $X = \sum_{A\in\mathcal{A}} X_A$. Thus, X counts how many sets in $\mathcal{A}$ has all of its elements chosen. For $A, B \in \mathcal{A}$, we denote $A \sim B$ meaning that $A \neq B$ and $A \cap B \neq \emptyset$. Let

$$\Delta = \sum_{A\sim B} \mathbb{E}\big[X_A X_B\big].$$

Theorem B.5 (Janson's Inequality)**.** *Let X and Δ be as above and let μ be the expectation of X. Then, for every $t \geq 0$, we have*

$$\mathbb{P}\big[X \leq \mu - t\big] \leq \exp\left\{-\frac{t^2}{2(\mu+\Delta)}\right\}.$$

In particular, if we consider a sequence of random variables X_n just as X in above, and if $\Delta_n \mu_n^2$, when $n \to \infty$, then, for any $a \in [0,1]$, we have $X \leq (1-a)\mu$ asymptotically almost surely.

We say that a random variable on a sample space Ω is *increasing* if for all $A \subseteq B \subseteq \Omega$ we have that $X(A) \geq X(B)$.

Theorem B.6 (FKG Inequality)**.** *If X and Y are two increasing random variables over the same sample space, then*

$$\mathbb{E}\big[XY\big] \geq \mathbb{E}\big[X\big]\mathbb{E}\big[Y\big].$$

REFERENCES

Balogh, József; Bollobás, Béla; Simonovits, Miklós. *The number of graphs without forbidden subgraphs.* Journal of Combinatorial Theory, Series B, v. 91, n. 1, p. 1–24, 2004.

Balogh, József; Morris, Robert; Samotij, Wojciech. *Independent sets in hypergraphs.* Journal of the American Mathematical Society, v. 28, n. 3, p. 669–709, 2015.

Bollobás, Béla. Random graphs. *Combinatorics (Swansea, 1981)*, Cambridge University Press, Cambridge-New York, v. 52 of *London Math. Soc. Lecture Note Ser.*, p. 80–102. 1981.

Bollobás, Béla. *Modern graph theory*, v. 184. Springer Science & Business Media, 2013.

Bollobás, Béla; Thomason, Arthur G. *Threshold functions.* Combinatorica, v. 7, n. 1, p. 35–38, 1987.

Burr, Stefan A; Erdős, Paul; Lovász, Laszlo. *On graphs of Ramsey type.* Ars Combinatoria, v. 1, n. 1, p. 167–190, 1976.

Conlon, David. *A new upper bound for diagonal Ramsey numbers.* Annals of Mathematics, p. 941–960, 2009.

Conlon, David; Fox, Jacob; Sudakov, Benny. *Recent developments in graph Ramsey theory.* Surveys in Combinatorics, v. 424, p. 49–118, 2015.

Conlon, David; Gowers, W Timothy. *Combinatorial theorems in sparse random sets.* Annals of Mathematics, v. 184, n. 2, p. 1–88, 2016.

Conlon, David; Gowers, W Timothy; Samotij, Wojciech; Schacht, Mathias. *On the KŁR conjecture in random graphs.* Israel Journal of Mathematics, v. 203, n. 1, p. 535–580, 2014.

Erdős, Paul; Hajnal, András. *Research problems.* Journal of Combinatorial Theory, v. 2, n. 1, p. 104 – 105, 1967.

Erdős, Paul; Stone, Arthur H. *On the structure of linear graphs.* Bulletin of the American Mathematical Society, v. 52, p. 1087–1091, 1946.

Erdős, Paul; Faudree, Ralph J; Rousseau, Cecil C; Schelp, Richard H. *The size Ramsey number.* Periodica Mathematica Hungarica, v. 9, n. 1-2, p. 145–161, 1978.

Erdös, Paul; Frankl, Peter; Rödl, Vojtech. *The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent.* Graphs and Combinatorics, v. 2, n. 1, p. 113–121, 1986.

Erdős, Paul; Simonovits, Miklós. *Supersaturated graphs and hypergraphs.* Combinatorica, v. 3, n. 2, p. 181–192, 1983.

Erdős, Paul; Turán, Paul. *On Some Sequences of Integers.* Journal London Mathematical Society, v. S1-11, n. 4, p. 261, 1936.

Folkman, Jon. *Graphs with monochromatic complete subgraphs in every edge coloring.* SIAM Journal on Applied Mathematics, v. 18, n. 1, p. 19–24, 1970.

Frankl, Peter; Rödl, Vojtech. *Large triangle-free subgraphs in graphs without K_4.* Graphs and Combinatorics, v. 2, n. 1, p. 135–144, 1986.

Gerke, Stefanie; Steger, Angelika. *The sparse regularity lemma and its applications.* Surveys in combinatorics, v. 327, p. 227–258, 2005.

Janson, Svante; Łuczak, Tomasz; Ruciński, Andrzej. *Random graphs*, v. 45. John Wiley & Sons, 2011.

Kinnersley, William B; Milans, Kevin G; West, Douglas B. *Degree Ramsey numbers of graphs.* Combinatorics, Probability and Computing, v. 21, n. 1-2, p. 229–253, 2012.

Kleitman, Daniel J; Winston, Kenneth J. *On the number of graphs without 4-cycles.* Discrete Mathematics, v. 41, n. 2, p. 167 – 172, 1982.

Kohayakawa, Yoshiharu. Szemerédi's regularity lemma for sparse graphs. *Foundations of computational mathematics*, Springer, p. 216–230. 1997.

Kohayakawa, Yoshiharu; Kreuter, Bernd. *Threshold functions for asymmetric Ramsey properties involving cycles.* Random Structures and Algorithms, v. 11, n. 3, p. 245–276, 1997.

Kohayakawa, Yoshiharu; Łuczak, Tomasz; Rödl, Vojtech. *On K_4-free subgraphs of random graphs.* Combinatorica, v. 17, n. 2, p. 173–213, 1997.

Kohayakawa, Yoshiharu; Schacht, Mathias; Spöhel, Reto. *Upper bounds on probability thresholds for asymmetric Ramsey properties.* Random Structures & Algorithms, v. 44, n. 1, p. 1–28, 2014.

Komlós, János; Shokoufandeh, Ali; Simonovits, Miklós; Szemerédi, Endre. *The Regularity Lemma and Its Applications in Graph Theory*, Springer, p. 84–112. 2002.

Kreuter, Bernd. *Threshold functions for asymmetric Ramsey properties with respect to vertex colorings.* Random Structures & Algorithms, v. 9, n. 3, p. 335–348, 1996.

Marciniszyn, Martin; Skokan, Jozef; Spöhel, Reto; Steger, Angelika. *Asymmetric Ramsey* *properties of random graphs involving cliques.* Random Structures & Algorithms, v. 34, n. 4, p. 419–453, 2009.

Nenadov, Rajko; Škorić, Nemanja; Steger, Angelika. An algorithmic framework for obtaining lower bounds for random Ramsey problems. *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms.* SIAM, 2015, p. 1743–1751.

Nenadov, Rajko; Steger, Angelika. *A short proof of the random Ramsey theorem.* Combinatorics, Probability and Computing, v. 25, n. 01, p. 130–144, 2016.

Nešetřil, Jaroslav; Rödl, Vojtěch. *The Ramsey property for graphs with forbidden complete subgraphs.* Journal of Combinatorial Theory, Series B, v. 20, n. 3, p. 243–249, 1976.

Radziszowski, Stanisław P. *Small ramsey numbers.* The Electronic Journal of Combinatorics, v. 1, n. 7, 1994.

Rödl, Vojtech; Ruciński, Andrzej. *Lower bounds on probability thresholds for Ramsey properties.* Combinatorics, Paul Erdős is eighty, v. 1, p. 317–346, 1993.

Rödl, Vojtěch; Ruciński, Andrzej. *Threshold functions for Ramsey properties.* Journal of the American Mathematical Society, v. 8, n. 4, p. 917–942, 1995.

Saxton, David; Thomason, Andrew. *Hypergraph containers.* Inventiones mathematicae, v. 201, n. 3, p. 925–992, 2015.

Schacht, Mathias. *Extremal results for random discrete structures.* Annals of Mathematics, v. 184, n. 2, p. 333–365, 2016.

Scott, Alexander. *Szemerédi's regularity lemma for matrices and sparse graphs.* Combinatorics, Probability and Computing, v. 20, n. 03, p. 455–466, 2011.

Spencer, Joel. *Ramsey's theorem – a new lower bound.* Journal of Combinatorial Theory, Series A, v. 18, n. 1, p. 108–115, 1975.

Szemerédi, Endre. *On sets of integers containing no k elements in arithmetic progression.* Acta Arithmetica, v. 27, n. 199-245, p. 2, 1975.

INDEX