

# Technical Report

TR-2004-016

**A Sharp Threshold for Random Graphs with a Monochromatic Triangle in  
Every Edge Coloring**

by

Ehud Friedgut, Vojtech Rodl, Andrzej Rucinski, Prasad Tetali

**MATHEMATICS AND COMPUTER SCIENCE**

**EMORY UNIVERSITY**

# A Sharp Threshold for Random Graphs with a Monochromatic Triangle in Every Edge Coloring

Ehud Friedgut      Vojtech Rödl      Andrzej Ruciński  
Prasad Tetali

October 17, 2004

## Abstract

Let  $\mathcal{R}$  be the set of all finite graphs  $G$  with the Ramsey property that every coloring of the edges of  $G$  by two colors yields a monochromatic triangle. In this paper we establish a sharp threshold for random graphs with this property. Let  $G(n, p)$  be the random graph on  $n$  vertices with edge probability  $p$ . We prove that there exists a function  $\hat{c} = \hat{c}(n) = \Theta(1)$  such that for any  $\varepsilon > 0$ , as  $n$  tends to infinity,

$$Pr [G(n, (1 - \varepsilon)\hat{c}/\sqrt{n}) \in \mathcal{R}] \rightarrow 0$$

and

$$Pr [G(n, (1 + \varepsilon)\hat{c}/\sqrt{n}) \in \mathcal{R}] \rightarrow 1.$$

A crucial tool that is used in the proof and is of independent interest is a generalization of Szemerédi's Regularity Lemma to a certain hypergraph setting.

---

AMS 2000 Mathematics Subject Classification: 05C15, 05C55, 05C80.

Key words and phrases: Ramsey Theory, Random Graphs, Threshold Phenomena, Szemerédi's Regularity Lemma.

Research supported by grants of the ISF, BSF, NSF and KBN. For details see section 7.1

# Contents

<b>1</b>	<b>Introduction</b>	<b>4</b>
1.1	Overview . . . . .	4
1.2	Ramsey Properties of Random Graphs . . . . .	5
1.3	Sharp Thresholds of Increasing Graph Properties . . . . .	7
1.4	Regularity . . . . .	10
1.5	Outline of the Paper . . . . .	12
<b>2</b>	<b>Outline of the Proof</b>	<b>13</b>
2.1	Main steps . . . . .	13
2.2	Overview of the Proof Strategy . . . . .	18
2.2.1	An Illustration . . . . .	19
<b>3</b>	<b>Tepees and Constellations</b>	<b>21</b>
3.1	Tepees . . . . .	21
3.2	Constellations . . . . .	26
<b>4</b>	<b>Regularity</b>	<b>31</b>
4.1	Classical Regularity . . . . .	31
4.1.1	Dense Regular Graphs . . . . .	31
4.1.2	Sparse Regular Graphs . . . . .	33
4.2	The Subgraph Regularity Lemma . . . . .	39
4.2.1	The Hypergraph Regularity of Frankl and Rödl . . . . .	40
4.2.2	General Subgraph Framework . . . . .	41
4.2.3	Measuring Density . . . . .	42
4.2.4	Partitions and Polyads . . . . .	45
4.2.5	Regularity . . . . .	46
4.3	Proof of the Subgraph Regularity Lemma . . . . .	49
4.3.1	Pumping the Index . . . . .	49
4.3.2	Refining a Partition . . . . .	55
4.4	No Dense Patches . . . . .	56
<b>5</b>	<b>The Core Section (Proof of Lemma 2.4)</b>	<b>67</b>
5.1	What is a Core? . . . . .	68
5.1.1	Applying the Regularity Lemma . . . . .	68
5.1.2	Anatomy and Polyads . . . . .	70

5.1.3	The Definition of a Core . . . . .	71
5.2	Every Hitting Set of $\mathcal{S}$ Contains a Core . . . . .	72
5.3	There Are Few Cores . . . . .	74
5.4	Cores Are Large . . . . .	75
5.4.1	Perfect Olympiads . . . . .	76
5.4.2	Precores Are Large . . . . .	80
<b>6</b>	<b>Random Graphs</b>	<b>82</b>
6.1	The Graph Family $\mathcal{G}$ . . . . .	82
6.2	Proof of Lemma 2.3 . . . . .	85
<b>7</b>	<b>Summary, Further Remarks, Glossary</b>	<b>89</b>
7.1	Acknowledgments . . . . .	90
7.2	Glossary . . . . .	90

# 1 Introduction

## 1.1 Overview

This paper brings together several important themes of combinatorics: Ramsey properties, threshold phenomena of random graphs, and Szemerédi-type regularity.

Ramsey properties guarantee, for an arbitrary partition of a given structure, that a highly organized substructure can be found in at least one part of the partition. During the last decade of the last century there has been extensive study of Ramsey properties of random structures, see e.g. [11, 27, 29, 30, 31, 32, 13, 33]. These papers were all concerned with establishing a threshold function for various Ramsey-type properties, either of random graphs, random hypergraphs or random sets of integers. For a binomial random graph  $G(n, p(n))$ , for instance, they provide a critical edge probability  $\hat{p} = \hat{p}(n)$  such that the limiting probability that every coloring of a random graph  $G(n, p)$  contains certain monochromatic structures depends on the asymptotic ratio between  $p$  and  $\hat{p}$ .

In all the above papers there is a multiplicative gap left between the upper and lower bound on the threshold edge probability  $\hat{p}$  (see Theorem 1.2 below). It is therefore not surprising that the natural question of whether the gap can be closed has been around for some time. In other words, does there exist a constant  $\hat{c}$  such that the asymptotic probability that  $G(n, c\hat{p})$  has a Ramsey property is either 0 or 1, depending only on whether  $c > \hat{c}$  or  $c < \hat{c}$ ? This question is usually phrased in the specialized jargon as “does there exist a *sharp threshold*?”

Sharp thresholds have been established for many random graph properties, like connectivity, hamiltonicity and perfect matchings. However, such precise results for Ramsey properties seemed out of hand until 1999, when a general technique for settling these questions was introduced in [8]. Loosely speaking, the main theorem in [8] showed that the question of sharpness of threshold for a random graph property is determined by whether the property is related to local or rather global graph phenomena.

Two papers exploiting the technique from [8] for coloring questions are [1] and [10]. The latter paper (as well as the earlier [31]) states as the next natural candidate for attack, the problem of sharpness of the threshold for property  $\mathcal{R}$  consisting of all graphs  $G$  such that in every blue-red coloring of

the edges of  $G$  there exists a monochromatic triangle. However, this problem turned out to be more difficult than those in [1] and [10] and required some new combinatorial approach.

In this paper we add the missing tool that enables us to crack this nut. It is a regularity lemma for a certain class of hypergraphs, whose edges consist of small subgraphs of a fixed underlying sparse graph (see Lemma 4.13). Our lemma is a generalization of the celebrated Szemerédi Regularity Lemma for graphs [35], and follows in the footsteps of the regularity lemma for sparse graphs presented in [20] (see also [22] and [16], Section 8.3) and of the hypergraph regularity lemma of Frankl and Rödl [12].

The proof of our regularity lemma, Lemma 4.13, provides for a considerable portion of the bulk of the proof of the following sharp threshold theorem, which is the ultimate result of our paper.

**Theorem 1.1** *There exists a function  $\widehat{c} = \widehat{c}(n) = \Theta(1)$  such that for every  $\varepsilon > 0$*

$$\lim_{n \rightarrow \infty} \Pr[G(n, p) \in \mathcal{R}] = \begin{cases} 1 & \text{if } p > (1 + \varepsilon)\widehat{c}/\sqrt{n} \\ 0 & \text{if } p < (1 - \varepsilon)\widehat{c}/\sqrt{n}. \end{cases} \quad (1)$$

There is a slightly disappointing aspect of this result: although we prove that  $\widehat{c}(n)$  is bounded, the natural conjecture is that  $\widehat{c}(n)$  converges to a positive limit, and this does not follow from our theorem. Unfortunately, an inherent property of the technique we use is that it can only supply such existence theorems but no new information as to the exact threshold probability. We discuss various possible extensions of Theorem 1.1 in Section 7.

## 1.2 Ramsey Properties of Random Graphs

Let us introduce the arrow notation, commonly used in Ramsey theory. For two graphs,  $H$  and  $G$ , and an integer  $r \geq 2$ , we write  $G \rightarrow (H)_r$  if for every coloring of the edges of  $G$  by  $r$  colors there exists a monochromatic copy of  $H$ . For example, it is well known that  $K_6 \rightarrow (K_3)_2$ . Let  $\mathcal{R}$  be the set of all graphs  $G$  such that  $G \rightarrow (K_3)_2$ .

A basic question studied in Ramsey theory is, given a graph  $H$  and an integer  $r \geq 2$ , when is  $G$  “rich” enough for  $G \rightarrow (H)_r$ ? Here richness can be interpreted either as the number of edges of  $G$ , or as the ratio of edges to vertices.

In modern graph theory, problems of this type are often studied via random graphs. The theory of random graphs addresses questions concerning typical graphs, or graphs “on average”. The standard model for a random graph is  $G(n, p)$ , a graph on  $n$  vertices, where every one of the  $\binom{n}{2}$  edges of the complete graph belongs to  $G(n, p)$  independently, with probability  $p$ . When studying random graphs a natural problem is: Given  $H$ , find a threshold function  $\hat{p}(n)$  such that  $G(n, p) \rightarrow (H)_r$  with probability tending to 1 when  $\hat{p}(n) \gg p$  and  $G(n, p) \rightarrow (H)_r$  with probability tending to 0 when  $\hat{p}(n) \ll p$ . (The existence of such a threshold function is guaranteed by a general result of Bollobás and Thomason [4].)

In a series of papers [11, 27, 29, 30, 31], a threshold function  $\hat{p}(n)$  is determined for all graphs  $H$ . Its culmination, paper [31], establishes  $\hat{p} = n^{-1/m^{(2)}(H)}$  as a threshold for  $G(n, p) \rightarrow (H)_r$ , regardless of  $r$ , for all  $H$  which are not star forests. Here  $m^{(2)}(H) = \max_{F \subseteq H} (|E(F)| - 1) / (|V(F)| - 2)$ . An analogous result in the case when the vertices (and not the edges) are colored is given in [27]. In the edge-coloring setting, the first case to be settled was that of triangles (not counting star forests which are rather trivial).

**Theorem 1.2 ([30])** *For every integer  $r \geq 2$  there exist constants  $c_r$  and  $C_r$  such that*

$$\lim_{n \rightarrow \infty} \Pr[G(n, p) \rightarrow (K_3)_r] = \begin{cases} 1 & \text{if } p > C_r/\sqrt{n} \\ 0 & \text{if } p < c_r/\sqrt{n}. \end{cases} \quad (2)$$

**Remark:** In the above theorem, and throughout the paper  $G(n, p)$  is usually meant to denote  $G(n, p(n))$ , hence the limits, and asymptotic notations. As we can see, for a range of  $p$ , namely for  $c_r/\sqrt{n} \leq p \leq C_r/\sqrt{n}$  this statement is inconclusive. Similarly, in all the papers mentioned above there is a multiplicative gap left between the upper and lower bound on the threshold edge probability  $\hat{p}(n)$ .

In a recent paper [10] it was shown that in many cases this gap can be closed, using a general technique from [8] for proving sharpness of thresholds. The cases treated in [10] cover vertex-coloring when  $H$ , the graph defining the Ramsey property, belongs to a wide family of graphs including, for example, all complete graphs. Also the case of edge-coloring when  $H$  is a tree is dealt with. In all these instances it is shown that there exists a function  $\hat{p}(n)$  such that for every  $\varepsilon > 0$ ,  $\lim_{n \rightarrow \infty} \Pr[G(n, p) \rightarrow (H)] = 1$ , if  $p > (1 + \varepsilon)\hat{p}(n)$  and 0, if  $p < (1 - \varepsilon)\hat{p}(n)$ .

It is worthwhile pointing out here a difference between this kind of a “sharp threshold” statement and the previous Ramsey threshold results. Although the results from [10] show that the transition from the non-Ramsey region of the values of  $p$  to the Ramsey region is swift and sharper than what was previously proven, these are only an existence results. They give no further information on the critical value of  $p$  which was calculated in previous works.

The most basic case not treated in [10] is the case of graphs with a monochromatic triangle in every edge coloring, for which a (weak) threshold is given by Theorem 1.2. In the present paper we prove a sharp threshold theorem for this Ramsey property, with two colors. This is our Theorem 1.1 stated above.

### 1.3 Sharp Thresholds of Increasing Graph Properties

In this subsection we introduce the notion of a *sharp threshold* for a graph property, as well as a technique for proving sharpness of thresholds. Consider a property  $Q$  of graphs on  $n$  vertices. We identify  $Q$  with the set of graphs having this property, and use the notation  $G \in Q$  to denote the fact that  $G$  has property  $Q$ . We will restrict ourselves to properties which are invariant under a graph automorphism and also distinguish an important class of *increasing* graph properties, i.e. those that are preserved under edge addition.

**Definition 1.1** We say that  $\hat{p} = \hat{p}(n)$  is a *threshold function* for an increasing graph property  $Q$  if

$$\lim_{n \rightarrow \infty} \Pr[G(n, p) \in Q] = \begin{cases} 1 & \text{if } \hat{p} = o(p) \\ 0 & \text{if } p = o(\hat{p}). \end{cases}$$

Bollobás and Thomason proved in [4] the existence of threshold functions for all increasing set properties, and in particular for all graph properties.

Some properties do not have sharper thresholds in the sense that for all  $p = p(n)$  which are of the same order as  $\hat{p}$ , we have  $0 < \lim_{n \rightarrow \infty} \Pr[G(n, p) \in Q] < 1$ . E.g., this is the case of the (increasing) property of containing a copy of a given, balanced graph  $H$ , the threshold for which has been established by Erdős and Rényi [7] at  $n^{-1/\rho(H)}$ . Here  $\rho(H)$  is the edge to vertex ratio in

*H.* For example,  $\rho(K_k) = (k - 1)/2$ , so the thresholds for the appearance in  $G(n, p)$  of a copy of  $K_3$ ,  $K_5$ , and  $K_6$ , respectively, are  $n^{-1}$ ,  $n^{-1/2}$  and  $n^{-2/5}$ .

But there are other properties, like connectivity, which enjoy much sharper thresholds. Indeed, it has been proved by Erdős and Rényi in [6] that

$$\lim_{n \rightarrow \infty} \Pr[G(n, p) \text{ is connected}] = \begin{cases} 1 & \text{if } np - \log n \rightarrow \infty \\ 0 & \text{if } np - \log n \rightarrow -\infty. \end{cases}$$

**Definition 1.2** We say that an increasing property  $Q$  has a *sharp threshold* if there exists a function  $\hat{p} = \hat{p}(n)$ , such that for every  $\varepsilon > 0$ :

$$\lim_{n \rightarrow \infty} \Pr[G(n, p) \in Q] = \begin{cases} 1 & \text{if } p > (1 + \varepsilon)\hat{p} \\ 0 & \text{if } p < (1 - \varepsilon)\hat{p}. \end{cases}$$

Otherwise we say that  $Q$  has a *coarse threshold*.

Thus, the property of being connected has a sharp threshold at  $\hat{p} = \log n/n$ . Our main theorem in this paper, Theorem 1.1, states that the Ramsey property  $\mathcal{R}$  has a sharp threshold.

In [8] the first author gives a necessary and sufficient condition for an increasing property to have a sharp threshold. This condition will be the central tool used in this paper. Roughly stated, it says that a property with a sharp threshold is not statistically determined by a simple local reason, that is, by the presence or absence of finitely many edges. For example, the property of having a triangle as a subgraph is obviously local, and indeed has a coarse threshold (both the critical probability and the length of the threshold interval equal  $\Theta(1/n)$ ), whereas it seems obvious that connectivity is a non-local property, even though it is statistically equivalent to the absence of isolated vertices.

For the Ramsey property  $\mathcal{R}$ , the condition in [8] asserts that in order to establish the sharpness of its threshold, one has to show that  $\mathcal{R}$  is not influenced by the appearance of any fixed subgraph, which is likely to be contained in  $G(n, p)$ , with the range of  $p = p(n)$  limited by Theorem 1.2 to  $p = \Theta(1/\sqrt{n})$ . Of course,  $\mathcal{R}$  is extremely influenced by the appearance of  $K_6$ , which, however, is very unlikely to be present in such a  $G(n, p)$ .

In this paper we will use a version of the sharpness criterion from [8], which follows readily from the original statement but is more suitable for applications. Given a graph  $M$  and a disjoint set of  $n$  vertices, let  $M^*$  be an

ordered copy of  $M$  placed uniformly at random on one of the  $n!/(n-|V(M)|)!$  possible locations.

**Theorem 1.3** *Let  $Q$  be an increasing graph property, with a coarse threshold. Then there exist real constants  $0 < c < C$ ,  $\beta > 0$ , a rational  $\rho$ , and a sequence  $p = p(n)$  satisfying  $cn^{-1/\rho} < p(n) < Cn^{-1/\rho}$ , such that  $\beta < \Pr[G(n, p) \in Q] < 1 - \beta$  for all  $n$ . Furthermore, there exist  $\alpha, \xi > 0$  and a balanced graph  $M$  with density  $\rho$  for which the following holds: For every graph property  $\mathcal{G}$  such that  $G(n, p) \in \mathcal{G}$  a.a.s. there are infinitely many values of  $n$  for which there exists a graph  $G$  on  $n$  vertices for which the following holds:*

- (i)  $G \in \mathcal{G}$  ,
- (ii)  $G \notin Q$  ,
- (iii)  $\Pr[(G \cup M^*) \in Q] > 2\alpha$ ,
- (iv)  $\Pr[(G \cup G(n, \xi p)) \in Q] < \alpha$  .

What the theorem says, intuitively, is that in the case of a coarse threshold one can find two graphs,  $G$  and  $M$  as follows:  $G$  is a fixed graph on  $n$  vertices that is not a random graph but rather a pseudorandom graph, typical of  $G(n, p)$  (actually a random choice of  $G \in G(n, p) \setminus Q$  will work with probability close to 1);  $M$  is a “magical” balanced graph such that it is often the case that adding a random copy of  $M$  to  $G$  induces the property in question, whereas increasing the number of edges in  $G$  randomly by a constant proportion  $\xi$  does not induce the property. The addition of a copy of  $M$  corresponds roughly to inducing a local property, in contrast to increasing the number of edges which corresponds roughly to increasing the global density of a random graph. Therefore the conclusion of the hypotheses of the theorem is that the property  $Q$  is “statistically local”.

The typical way in which this theorem is used to prove that a property  $Q$  has a sharp threshold involves two steps:

- The first step is usually easy. For a coarse property  $Q$ , the theorem guarantees the existence of  $M$ . A possible explanation to this would be that  $M$  itself has the property  $Q$ . In that case, since  $Q$  is a monotone increasing property it would no longer seem “magical” that adding

a copy of  $M$  induces property  $Q$ . In other words, if  $M \in Q$  then Assumption (iii) in the theorem is a triviality and does not enable one to deduce anything. Therefore, as a start, one has to rule out this possibility by showing that a balanced graph with the prescribed density can not have the property.

- The second step is typically more involved. One chooses an appropriate property  $\mathcal{G}$  that is typical of  $G(n, p)$  and then shows that a graph  $G \in \mathcal{G}$  with  $\Pr[(G \cup M^*) \in Q] > 2\alpha$  is quite “saturated”, i.e. is close to having property  $Q$ . Therefore adding a random copy of  $G(n, \xi p)$  should induce the property  $Q$  with probability much larger than  $\alpha$ , contradicting condition (4) of the theorem.

In [8], [1], and [10] one can see variations of this scheme used to prove that various graph properties have a sharp threshold. Although the basic technique is similar, each property presents its own difficulties and requires a special approach. The case of property  $\mathcal{R}$ , handled in this paper, has by far been the most difficult and involved; a key technique in our approach turns out to be *regularization*.

## 1.4 Regularity

One of the fundamental tools in asymptotic graph theory is the well-known regularity lemma of Szemerédi [35] (see also [34]). Indeed, since its discovery in the 70s, this lemma has been instrumental in the study of the structure of large graphs. The reader is referred to the excellent survey [24] for a thorough introduction to the wide range of applications of this result.

In essence, the regularity lemma tells us that any large graph may be decomposed into a bounded number of quasi-random, induced bipartite graphs. Thus, this lemma is a powerful tool for detecting and making transparent the *random-like behavior* of large *deterministic* graphs. What makes the lemma such a powerful tool is that it reveals a quasi-random structure that enables one to carry out a deep quantitative analysis.

The precise formulation of the regularity lemma is somewhat technical (see Section 4.1.1 for details). In this short section, we only discuss some points in broad terms.

The quasi-random bipartite graphs that Szemerédi’s lemma uses in its decomposition are graphs in which the edges are uniformly distributed. The

uniformity is measured by the ratio of edges to potential edges (pairs), and so this concept becomes trivial for graphs of vanishing density. To manage sparse graphs, one may adjust the notion of quasi-randomness by a natural rescaling, and it is a routine matter to check that the original proof extends to this notion, provided we restrict ourselves to graphs of vanishing density that do not contain ‘dense patches’ (see Section 4.1.2). However, the quasi-random structure that the lemma reveals in this case is harder to exploit than in the dense case, and one needs to work harder when applying the lemma to such ‘sparse graphs’. Nevertheless, there have been some successful applications of the lemma in this context (see [20, 23]).

The idea of regularity has also been extended to uniform hypergraphs. The version that is most relevant to us is the one in Frankl and Rödl [12], which makes it possible to decompose triple systems into quasi-random structures made up of triples together with an ‘underlying’ quasi-random tripartite graph. In that setting, the density is measured by the ratio of triples to the triangles in the underlying graph (see Section 4.2.1). Moreover, the concept of quasi-randomness here is strong enough to allow one to prove that these quasi-random pieces contain the same number of finite substructures as they would had they been truly random pieces (see [12] and Nagle and Rödl [28]).

In this paper, we shall introduce yet another concept of regularity, which expands and melts together the notions of sparse graph regularity and hypergraph regularity. In the usual context of graph regularity, and in some more delicate versions of it there is an invisible underlying graph behind the graph we look at, and the regularity expresses the distribution of specified edges with respect to the edges of the underlying graph. Similarly, a 3-uniform hypergraph can be viewed as a distinguished collection of triangles among all triangles of an underlying 3-partite graph. Here, we shall be interested in investigating the structure of sparse graphs with respect to some other fixed family of small subgraphs. Viewing these subgraphs as edges of a hypergraph, the lemma we prove (Lemma 4.13) may be interpreted as a sparse hypergraphs version of the regularity lemma. Our approach is partly based on methods from [20, 23] and [12], but it faces a further difficulty: the assumption of ‘no dense patches’ in the standard case (see [20, 23]) was an easy consequence of properties of random graphs and therefore did not play any significant role; the proof of the analogous fact in the setting of this paper, however, requires a fairly complex argument (see Section 4.4).

## 1.5 Outline of the Paper

- In Section 2 we present the skeleton of the proof of the main theorem. It is actually a formal proof which is made extremely short and compact by relying on lemmas which will be proven in the rest of the paper. An overview of the forthcoming proofs and an (oversimplified) illustration will also be given there.
- In Section 3, assuming the hypothesis of Theorem 1.3, we construct a family  $\mathcal{S}$  of special subgraphs of  $G$  (called special constellations) and show how every triangle-free coloring of  $G$  defines a set of edges of  $G$  that intersects all special constellations (that is, defines a hitting set for the family  $\mathcal{S}$ ).
- In Section 4 we prove a regularity theorem in the spirit of Szemerédi's Regularity Lemma that provides a partition of both the vertices and the edges of  $G$  such that the special constellations are uniformly distributed with respect to this partition.
- In Section 5, based on the regular partition found in the previous section, we define a core which is a central notion of this proof, and show some crucial properties of cores.
- In Section 6 we show various properties of random graphs that are needed throughout the paper. It is there where the family  $\mathcal{G}$  is defined, and an important lemma, Lemma 2.3, is proved.
- We conclude with open questions and possible extensions.
- At the end of the paper we include a glossary of symbols and definitions as well as a flowchart of constants exhibiting their mutual dependencies. We strongly encourage the reader to make use of both when struggling through our proof.

### Notation:

In Sections 4-6 we use the following notation. For  $0 < \varepsilon < 1$ , and positive reals  $x, y$ ,

$$x \overset{\varepsilon}{\sim} y \quad \text{denotes that} \quad (1 - \varepsilon)y \leq x \leq (1 + \varepsilon)y.$$

We will often abbreviate it further as follows: if  $\varepsilon'$  is any function of  $\varepsilon$  that tends to zero with  $\varepsilon$ , and  $x \overset{\varepsilon'}{\sim} y$ , then we will simply write  $x \overset{\varepsilon}{\approx} y$ .

Let  $G = (V, E)$  be a graph,  $v, u \in V$  and  $W \subseteq V$ . We write  $e_G(W) = |E(G[W])|$  for the number of edges in the subgraph of  $G$  induced by  $W$ ,  $\deg_G(v, W) = \deg(v, W)$  for the number of neighbors of  $v$  in  $W$ ,  $\deg_G(v) = \deg(v) = \deg(v, V)$ , for the degree of  $v$  in  $G$ , and  $\text{codeg}(v, u)$  for the number of common neighbors of  $u$  and  $v$  in  $G$ , called the *co-degree* of  $v$  and  $u$ . The set of neighbors of  $v$  in  $G$  is denoted by  $N_G(v) = N(v)$ , while  $N_G(W)$  stands for the set of vertices outside  $W$ , each having at least one neighbor in  $W$  (so called *open* neighborhood of  $W$ ).

For a family of sets  $(A_i)_{i \in I}$ , we call a set  $T$  a *hitting set* if  $T \cap A_i \neq \emptyset$  for all  $i \in I$ . We will often use set partitions  $V = V_1 \cup \dots \cup V_t$ , where  $|V_1| \leq \dots \leq |V_t| \leq |V_1| + 1$ . Such partitions will be called here *equipartitions*. Finally, all logarithms are natural and will be denoted by  $\log$ .

## 2 Outline of the Proof

### 2.1 Main steps

Recall that  $\mathcal{R}$  is the graph property that for every blue-red coloring of the edges of a graph there exists a monochromatic triangle. Graphs that have this property will be called Ramsey graphs. For a non-Ramsey graph, we call a coloring that does not have a monochromatic triangle a *triangle-free coloring*.

We wish to prove that  $\mathcal{R}$  has a sharp threshold. By Theorem 1.2, there exist constants  $c_2$  and  $C_2$  such that any threshold  $\hat{p}$  for  $\mathcal{R}$  satisfies

$$c_2/\sqrt{n} < \hat{p} < C_2/\sqrt{n}.$$

This means that when applying Theorem 1.3 to Property  $\mathcal{R}$  we may restrict ourselves to sequences  $p = p(n)$  falling into this range, and consequently to balanced graphs  $M$  with  $\rho(M) = 2$  (i.e. average degree 4.) Thus it suffices to prove the following result, which is a mere adaptation of Theorem 1.3 to our case. (In fact, it is slightly stronger, since the inequality in (iv) is replaced in (4) below by convergence to 1.)

**Theorem 2.1** For all  $\alpha, \xi > 0$ , all sequences

$$c_2/\sqrt{n} < p = p(n) < C_2/\sqrt{n}$$

and all balanced graphs  $M$  with  $\rho(M) = 2$ , there exists a graph property  $\mathcal{G}$  with  $\lim_{n \rightarrow \infty} \Pr[G(n, p) \in \mathcal{G}] = 1$ , and an integer  $n_1$ , such that for all  $G \in \mathcal{G} \setminus \mathcal{R}$  with  $|V(G)| = n > n_1$ , if

$$\Pr[G \cup M^* \in \mathcal{R}] > 2\alpha \tag{3}$$

then

$$\Pr[(G \cup G(n, \xi p)) \in \mathcal{R}] = 1 - o(1). \tag{4}$$

Note that if the assumption  $G \notin \mathcal{R}$  (Assumption (ii) in Theorem 1.3) were false then (4) would be trivial. A finer point is that if  $M \in \mathcal{R}$  then (3) yields no information on  $G$ , and hence is useless. Fortunately the following lemma rules out this possibility. This is the typical “easy step” mentioned in the introduction.

**Lemma 2.2** If  $M$  is balanced and  $\rho(M) = 2$  then  $M \notin \mathcal{R}$ .

**Proof:** It is a well known fact from the theory of random graphs (see [16], page 66) that for any *balanced* graph  $H$  with  $\rho(H) = \rho$  and  $p = \Theta(n^{-1/\rho})$  there exists a constant  $\beta > 0$  such that the probability that  $H$  appears in  $G(n, p)$  is at least  $\beta$ . Hence, for any  $b > 0$ , the probability that  $M$  appears as a subgraph of  $G(n, b/\sqrt{n})$  is bounded away from zero. If  $M \in \mathcal{R}$  then, by the monotonicity of  $\mathcal{R}$ , this would mean that  $\Pr[G(n, b/\sqrt{n}) \in \mathcal{R}]$  is also bounded away from zero. But by Theorem 1.2, for all  $b < c_2$ ,  $\Pr[G(n, b/\sqrt{n}) \in \mathcal{R}] \rightarrow 0$ . Therefore  $M \notin \mathcal{R}$ .

Alternatively, and without the use of random graphs, Lemma 2.2 follows from a result in [25] which shows that any Ramsey graph (for  $K_3$ ) must have a subgraph  $H$  for which  $\rho(H) \geq 5/2$ .  $\square$

The rest of the paper is devoted to proving that (3) implies (4). An approach to statements like (4), which has become standard by now, is via the so called two round exposure. This technique originated in the seventies, in work of Posa, Ajtai, Komlos, Szemerédi, Fernandez de la Vega, Fenner and Frieze, devoted to the existence of Hamilton paths and cycles in random graphs (see [3], Chapter VIII, for references). In the context of Ramsey

properties, it was explored already in [29] and [30] (see also [15], Sections 1.1 and 8.4). For a graph  $G = (V, E)$  and a set of edges  $F \subseteq E$ , let  $\text{Base}(F)$  be the set of edges in the complete graph on  $V$  that form a triangle with two edges of  $F$ , formally:

$$\text{Base}(F) = \{uv : uw, vw \in F \text{ for some } w \in V\}.$$

We often identify a subset of edges of a graph with the spanning subgraph consisting of them. So, in the above, both  $F$  and  $\text{Base}(F)$  can be viewed as graphs on the same vertex set  $V$ .

Suppose we want to show that  $G_1 \cup G_2 \in \mathcal{R}$  with high probability, where  $G_i = G(n, b_i/\sqrt{n})$ ,  $i = 1, 2$ , and the two random graphs are independent. (This is, in fact, our case with  $b_2 = \xi b_1$ , except for two minor points: that  $G_1$  is not random but pseudorandom, and that  $b_1$  may depend on  $n$ .)

First generate the edges of  $G_1$ , and let them be colored by an adversary. Suppose that at least half of them are blue and call the set of blue edges *Blue*. Clearly, if  $\text{Base}(\text{Blue})$  contains a triangle  $\Delta$ , no proper coloring of  $G_1 \cup \Delta$  can extend the adversarial coloring. Therefore, it will suffice to show that  $\text{Base}(\text{Blue}) \cap G_2$  contains a triangle. (See Figure 1.)

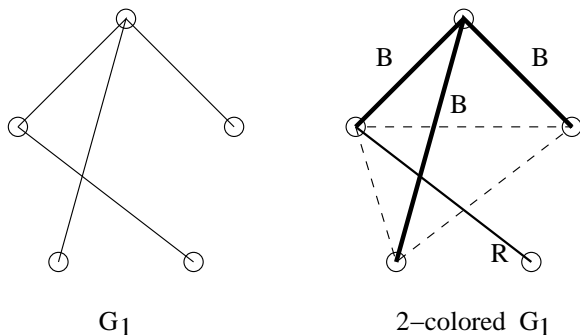


Figure 1: No triangle-free coloring of  $G_1 \cup \Delta$  can extend the given one

To this end we utilize the following lemma, proved in Section 6. Given two real numbers  $0 < \lambda < 1$  and  $0 < a < 1/6$ , we say that a graph  $G$  has property  $\mathcal{T}(\lambda, a)$ , if for any subgraph  $F$  of  $G$  with at least  $\lambda|E(G)|$  edges, the graph  $\text{Base}(F)$  contains at least  $a|V(G)|^3$  triangles.

**Lemma 2.3** *For all  $\lambda > 0$  and  $c > 0$ , there exists  $a > 0$ , such that if  $p \geq c/\sqrt{n}$  then, with probability  $1 - o(1)$ , the random graph  $G(n, p)$  has property  $\mathcal{T}(\lambda, a)$ .*

Applying the above lemma with  $F = \text{Blue}$ ,  $\lambda = 1/2$  and  $c = b_1$  would yield  $\Theta(n^3)$  triangles in  $\text{Base}(\text{Blue})$ . In the second round, by a standard application of a correlation estimate from [15] (see [16], Section 2.2), often called Janson's inequality, at least one of these triangles will be included in  $G_2$  with very high probability. If we were allowed to take  $b_2$  sufficiently large, then we could make the reciprocal of the error probability larger than the exponential number of all bi-colorings of  $G_1$ , proving  $G_1 \cup G_2 \in \mathcal{R}$  with probability  $1 - o(1)$ .

Unfortunately, in our case  $b_2 = \xi b_1$  depends on a given a priori  $\xi$  and can be much smaller than  $b_1$ . This major difficulty demonstrates a more general problem of establishing a sharp threshold *without knowing up front what the exact value of the critical constant  $\hat{c}$  should be*.

Hence we must seek a refinement of the above approach, making use of the assumption (3). And indeed, through a special regularization of  $G$  we will be able to construct a family  $\text{CORE}$  of subgraphs of  $G$  such that for every triangle-free coloring of  $G$  at least one of these subgraphs is monochromatic. Moreover, the size of each subgraph  $K \in \text{CORE}$  will be large enough to yield, via Lemma 2.3 and Janson's inequality, at least one triangle in  $\text{Base}(K) \cap G(n, \xi p)$  with probability very close to one, but at the same time the size of this family,  $|\text{CORE}|$ , multiplied by the error probability will tend to 0. This is the content of the following two lemmas which together imply Theorem 2.1. They are preceded by a *setup*, which we will often be referring to in the paper.

**Setup:**

For the rest of the paper, let us fix constants  $\alpha, \xi > 0$ , a sequence  $c/\sqrt{n} < p = p(n) < C/\sqrt{n}$ , where  $c = c_2$  and  $C = C_2$ , and a graph  $M \notin \mathcal{R}$  as in Theorem 2.1 (it is no longer relevant that  $M$  is balanced). We will define a graph property  $\mathcal{G}$  in Definition 6.1 so that, in particular, each  $G \in \mathcal{G}$  has Property  $\mathcal{T}(\lambda, a)$  with  $\lambda = \lambda(\alpha, c, C, M)$  to be specified later (or see the Glossary now) and  $a = a(\lambda, c)$  determined by Lemma 2.3.

We do not attempt to compute explicitly the integer  $n_1$ , promised in Theorem 2.1 and appearing in the next lemma. In principle, it is the maximum

of all values of  $n_0$  encountered throughout the proof, most notably the  $n_0$ 's in Theorem 4.10 and Lemma 4.13, as well as of several implicit lower bounds on  $n$  hidden in our calculations.

**Lemma 2.4** *Let  $G \in \mathcal{G} \setminus \mathcal{R}$  be a graph with  $|V(G)| = n \geq n_1$  for which the assumption (3) of Theorem 2.1 holds. Then for every  $\tau > 0$  there exists a family CORE of subgraphs of  $G$  such that*

- (a) *For every triangle-free coloring  $\chi$  of  $E(G)$ , there exists  $K \in \text{CORE}$  which is monochromatic under  $\chi$ .*
- (b)  $|\text{CORE}| \leq \exp(\tau n^{3/2})$ ,
- (c) *For every  $K \in \text{CORE}$  we have  $|K| > \lambda|E(G)|$ .*

After thorough preparations in Sections 3 and 4, the family CORE will be constructed in Section 5. Also there, all three conclusions of the above lemma will be proven in the following manner:

- (a) follows from Lemma 3.5 and Lemma 5.1,
- (b) is Lemma 5.2, and
- (c) follows from Lemma 3.5, Lemma 5.3 and Property (P3) of  $\mathcal{G}$ .

A specific value of  $\tau$ , with which we apply Lemma 2.4, is provided by another application of Janson's inequality. We say that a subgraph  $F \subseteq G$  *survives* if  $F \cup G(n, \xi p)$  has a triangle-free coloring in which  $F$  is monochromatic.

**Lemma 2.5** *For every  $K \in \text{CORE}$ , the probability that  $K$  survives is at most  $\exp\{-2\tau_0 n^{3/2}\}$ , where*

$$\tau_0 = \frac{a^2 \xi^6 c^6}{2(a\xi^3 c^3 + 2\xi^5 C^5)}.$$

**Proof:** By Lemma 2.4(c) and the fact that  $G \in \mathcal{T}(\lambda, a)$  there are at least  $an^3$  triangles in  $\text{Base}(K)$ . Let us number some  $an^3$  of them by  $1, 2, \dots, an^3$  and let  $I_i$  be the indicator random variable for the event that the  $i$ th triangle

is contained in  $G(n, \xi p)$ . Clearly, if  $K$  survives then  $\sum_i I_i = 0$ . But, by [16], Theorem 2.18(ii),

$$\Pr\left(\sum_i I_i = 0\right) \leq \exp\left\{-\frac{(\sum_i \mathbf{E}I_i)^2}{\sum \sum \mathbf{E}(I_i I_j) + \sum_i \mathbf{E}I_i}\right\},$$

where the double sum is over all ordered pairs of distinct, edge-intersecting triangles in  $\text{Base}(K)$ . Note that  $\sum_i \mathbf{E}I_i = an^3(\xi p)^3 \geq a\xi^3 c^3 n^{3/2}$  and that the double sum contains at most  $n^4$  summands, each equalling at most  $(\xi p)^5 \leq (\xi C)^5 n^{-5/2}$ . This completes the proof.  $\square$

Let us now show how Lemmas 2.4 and 2.5 imply Theorem 2.1, and consequently Theorem 1.1. We refer to the members of CORE by the name *cores*.

**Proof of Theorem 2.1:** Lemma 2.4(a) implies that if  $G \cup G(n, \xi p)$  has a proper coloring  $\chi$  (which, of course, induces a triangle-free coloring of  $G$ ) then there exists a core which survives. But, using Lemma 2.4(b) with  $\tau = \tau_0$ , Lemma 2.5 and a simple union bound, we deduce that with probability  $1 - o(1)$  no core survives. Indeed,

$$\Pr[\text{Any core survives}] \leq \exp\{\tau_0 n^{3/2}\} \cdot \exp\{-2\tau_0 n^{3/2}\} = o(1).$$

$\uparrow$   $\uparrow$   
 $\# \text{ of cores}$   $\text{survival probability}$

Hence  $\Pr[G \cup G(n, \xi p) \in \mathcal{R}] = 1 - o(1)$  as required.  $\square$

## 2.2 Overview of the Proof Strategy

Clearly, at such an early point in the paper, the main idea of the proof may yet be obscure. Specifically, we have given no hint as to the connection between the existence of a magical graph  $M$  having the property described in assumption (3) of Theorem 2.1 and the existence of a family CORE satisfying the three conclusions of Lemma 2.4. In order to shed some light on this connection (albeit it will still be a dim light), we give here a short explanation of the logic and motivation behind the construction of CORE. The next three sections of the paper are devoted to the constructions that underlie the following scheme:

- The existence of  $M$  such that

$$Pr[G \cup M^* \in \mathcal{R}] > 2\alpha$$

implies that the set of triangle-free colorings of  $G$  is very restricted; there are many sets of vertices in  $G$  such that planting a copy of  $M$  on them kills all triangle-free colorings, i.e. no triangle-free coloring of  $G$  can be extended to  $G \cup M$  when  $M$  is placed on one of the aforementioned “bad” sets.

- We will associate every such “bad” set with a set of edges in  $G$ , a union of stars which we will call a special constellation. We will fix a coloring of these special constellations in such a way that every proper coloring of  $G$  must agree with every such colored constellation on at least one edge (see Lemma 3.2).
- The above can be translated to the language of hypergraphs: we will construct a hypergraph whose hyperedges are the edge sets of the special constellations, and it will turn out that every triangle-free coloring gives rise to a hitting set of the hyperedges of this hypergraph (see Lemma 3.5).
- We will show that every such hitting set (and hence every triangle-free coloring) may be associated with a large monochromatic set called a core. CORE is the family of cores (see Lemma 2.4 above).
- The key to associating every hitting set with a core is in showing that our hypergraph has an inherent regular structure that may be revealed by a Szemerédi type partition (see Lemma 4.13).

### 2.2.1 An Illustration

To get a better feeling of how regularity helps in creating the family CORE, let us consider a simpler analogue that takes place in the well understood setting of graphs, in which special constellations will be replaced simply by edges. Hopefully this will give some clue as to what we are doing in the forthcoming sections. We refer the reader to Section 4 for the notion of an  $\varepsilon$ -regular graph and the statement of the Szemerédi Regularity Lemma.

Let  $H = H(n)$  be a sequence of graphs on  $n$  vertices. A *cover set* in a graph is a set of vertices that intersects every edge. In other words, it is a hitting set for the family of all edges of the graph. In general, the number of cover sets in an  $n$ -vertex graph may be exponential in  $n$ , and our goal is to “capture” all cover sets of  $H$  by a smaller family of large subsets of vertices which we will call *cores*. (Elsewhere in the paper the term *core* is used in our larger setting, cf. Lemma 2.4.) We want the following properties to hold for cores:

- Every cover set contains a core,
- The number of cores is  $2^{o(n)}$ ,
- Every core is of size linear in  $n$ .

In general, this is not an easy task, as the last two properties seem to contradict each other.

Before describing how to construct the cores, here is a hint as to *why* one would want the last two conditions to hold simultaneously: in our general setting we wish to capture the large family of all triangle-free colorings of our graph  $G$  by a smaller family of partial colorings, hence a *small* family of cores. On the other hand, we wish every partial coloring to be sufficiently large to ensure that the probability of being able to extend it to a larger graph  $G \cup G(n, \xi p)$  is very small, hence cores should be *large*.

Returning to the graph  $H$ , suppose that  $H$  is a bipartite  $\varepsilon$ -regular graph on vertex sets  $V_1, V_2$ , with  $|V_1| = |V_2| = n$  and density  $d(V_1, V_2) = d > 0$ . Recall that this means that for every  $W_1 \subseteq V_1$  and  $W_2 \subseteq V_2$  such that  $|W_1| > \varepsilon n$  and  $|W_2| > \varepsilon n$ , the density of the subgraph between  $W_1$  and  $W_2$ , is “ $\varepsilon$ -close” to  $d$ , and consequently, there is at least one edge between such  $W_1$  and  $W_2$ . Hence any cover set must necessarily include at least  $(1 - \varepsilon)n$  vertices from either  $V_1$  or  $V_2$ . Then our family of cores may be formed by all sets of size  $\lceil (1 - \varepsilon)n \rceil$  that are subsets of either  $V_1$  or  $V_2$ . It is easy to see that every cover set contains a core, and the number of cores is  $O\left(\binom{n}{\varepsilon n}\right)$  which indeed is  $2^{o(n)}$ , for  $\varepsilon = o(1)$ .

It is not hard to generalize this construction to the case of a multi-partite graph on vertex sets  $V_1, \dots, V_k$  such that for most pairs  $V_i, V_j$  the spanned bipartite graph is dense and  $\varepsilon$ -regular. Now comes the use of the Szemerédi Regularity Lemma: since every sufficiently large and dense graph is very

close to being of this type, the original problem can essentially be reduced to this case.

After seeing this example, hopefully, the reader may have a feeling as to why we will eventually devote much energy to exposing the hidden regular structure in the hypergraph which expresses the restrictions on the triangle-free colorings of  $G$ .

### 3 Tepees and Constellations

Assumption (3) of Theorem 2.1 should imply that  $G$  is close to being Ramsey in the sense that its triangle-free colorings are quite restricted. In this section we set up a family of subgraphs of  $G$  called special constellations that help to capture these restrictions. We will show that every triangle-free coloring of  $G$  may be associated with a hitting set of this family.

#### 3.1 Tepees

Our analysis of the restrictions imposed by (3) on the colorings of  $G$  will lead us rather naturally to the structures we shall call *tepees*.

Assume that  $M$ , an arbitrary balanced graph with  $\rho(M) = 2$ , has  $\nu$  vertices, and thus  $2\nu$  edges, and fix a generic copy of  $M$  with vertices labeled by  $x_1, \dots, x_\nu$ . We begin by defining a copy of  $M$  “planted” on an ordered subset of vertices of  $G$ . For every sequence  $X = (v_1, \dots, v_\nu)$  of distinct vertices of  $G$ , let  $M_X$  be the copy of  $M$  with vertex  $x_a$  mapped onto  $v_a$  for each  $a = 1, \dots, \nu$ . We will often identify a sequence  $X$  with the set  $\{v_1, \dots, v_\nu\}$  of vertices making up  $X$ .

The family of all sequences  $X$  satisfying  $G \cup M_X \in \mathcal{R}$  will be denoted by  $\mathcal{X}$ . Note that assumption (3) of Theorem 2.1 implies that

$$|\mathcal{X}| \geq 2\alpha n(n-1) \cdots (n-\nu+1) = (2-o(1))\alpha n^\nu.$$

Let  $\mathcal{X}_1 \subseteq \mathcal{X}$  be the family of all  $X \in \mathcal{X}$  such that

- (i)  $X$  is an independent set in  $G$ , i.e. the vertices in  $X$  span no edges of  $G$ , and
- (ii) every vertex of  $G$  has at most two neighbors in  $X$ .

Since  $G \in \mathcal{G}$  (see Definition 6.1, parts (P1) and (P2)), it follows that almost all  $X$ 's have properties (i) and (ii) above, and so we still have

$$|\mathcal{X}_1| \geq (2 - o(1))\alpha n^\nu.$$

Because of (i), the property  $G \cup M_X \in \mathcal{R}$  indicates that there should be some triangles in  $G \cup M_X$  with one edge in  $M_X$  and two edges in  $G$  (see the proof of Lemma 3.2 below). We now give a name to such structures.

**Definition 3.1** For a pair of vertices  $\{u, v\}$  of  $G$  a *tepee over  $\{u, v\}$*  is a pair of edges of  $G$  of the form  $\{uw, wv\}$ . Then vertex  $w$  is called it the tip of the tepee  $\{u, v\}$ .

In other words, a tepee over  $\{u, v\}$  is any path of length two in  $G$  with endpoints in  $u$  and  $v$ . (Clearly, the vertices of a tepee over  $\{u, v\}$  form a triangle in  $G + uv$ .) We will denote such a tepee by  $uwv$  for short. For a sequence of vertices  $X \in \mathcal{X}_1$ , let  $T(X)$  denote the set of all tepees in  $G$  over those pairs of vertices from  $X$  which are edges of  $M_X$ . By properties (i) and (ii) above, all tepees in  $T(X)$  are pairwise edge-disjoint and have distinct tips.

Furthermore, let  $\widehat{X}$  be the graph formed by the edges of tepees in  $T(X)$ . The vertex set of  $\widehat{X}$  consists of  $X$  and the set of tips of all tepees in  $T(X)$ . (This notation is supposed to be suggestive of the tepees formed over  $X$ .) Set  $t(X) = |T(X)|$ . Given  $t(X) = \phi$ ,  $0 \leq \phi \leq n - \nu$ , the graph  $\widehat{X}$  has  $\nu + \phi$  vertices and  $2\phi$  edges, and is isomorphic to a graph which can be obtained from  $M_X$  by replacing some of its edges by multiple edges, erasing others, and finally replacing all  $\phi$  edges of the obtained multigraph by internally disjoint paths of length two. So, there are many isomorphism types of  $\widehat{X}$  possible. The next lemma states that a positive fraction of  $X$ 's have the same isomorphism type of the graph  $\widehat{X}$ , and, moreover, the number of vertices in  $\widehat{X}$  is bounded. (We have sacrificed  $\alpha_1$  for the sake of global harmony.)

**Lemma 3.1** *Let  $q = \lceil 10C^2\nu/\alpha \rceil$  and  $\alpha_2 = \frac{\alpha}{(2\nu+q)^q}$ . There exists an integer  $\phi \leq q$ , a graph  $\widehat{M}$  on  $\nu + \phi$  vertices, and a family  $\mathcal{X}_2 \subseteq \mathcal{X}_1$  of size  $|\mathcal{X}_2| = \alpha_2 n^\nu$  such that for all  $X \in \mathcal{X}_2$ , the graph  $\widehat{X}$  is isomorphic to  $\widehat{M}$ .*

**Proof:** We will first show that there are at most  $\frac{4}{5}\alpha n^\nu$  sequences  $X \in \mathcal{X}_1$  such that  $t(X) > q$ . For every pair of vertices  $\{u, v\} \subset V$ , let  $T(u, v)$  be the set of tepees over  $\{u, v\}$ .

For a vertex  $w \in V$  of degree  $\deg(w)$  there are exactly  $\binom{\deg(w)}{2}$  tepees of the form  $uvw$ . This and the fact that  $G \in \mathcal{G}$  and, in particular, that the degrees of all vertices in  $G$  are bounded from above by  $2C\sqrt{n}$  (see Definition 6.1 (P3)), yields

$$\sum_w \binom{\deg(w)}{2} \leq n \binom{2C\sqrt{n}}{2} < 2C^2 n^2.$$

To apply a simple counting argument note that for all  $\{u, v\} \subset V$

$$|\{X : uv \in M_X\}| \leq 4\nu n^{\nu-2}.$$

Consequently, by reversing the order of summation,

$$\sum_{X \in \mathcal{X}_1} t(X) = \sum_{X \in \mathcal{X}_1} \sum_{uv \in M_X} |T(u, v)| \leq 4\nu n^{\nu-2} \sum_{u, v} |T(u, v)| \leq 8C^2 \nu n^\nu.$$

This immediately implies that  $t(X) > q$  holds for at most  $\frac{4}{5}\alpha n^\nu$  sets  $X \in \mathcal{X}_1$ , and, therefore, at least

$$|\mathcal{X}_1| - \frac{4}{5}\alpha n^\nu > \alpha n^\nu$$

sets  $X \in \mathcal{X}_1$  have  $t(X) \leq q$ .

Given that  $t(X) \leq q$ , the number of isomorphism types possible for the graph  $\widehat{X}$  is no more than the number of ordered partitions of the integer  $q$  into  $2\nu + 1$  nonnegative parts, corresponding to the decisions of how many tepees span over each edge of  $M_X$ . (The  $(2\nu + 1)$ -st part is the difference between  $q$  and the actual number of tepees present.) There are  $\binom{2\nu+q}{q} < (2\nu + q)^q$  such partitions. Take as  $\widehat{M}$  the most common isomorphism type among  $\{\widehat{X} : X \in \mathcal{X}_1, t(X) \leq q\}$  and set  $\phi = |V(\widehat{M})| - \nu$  and

$$\mathcal{X}_2 = \{X \in \mathcal{X}_1 : \widehat{X} \text{ is isomorphic to } \widehat{M}\}.$$

Then

$$|\mathcal{X}_2| \geq \frac{\alpha}{(2\nu + q)^q} n^\nu = \alpha_2 n^\nu.$$

□

Let us summarize the properties of the family  $\mathcal{X}_2$ . For every  $X \in \mathcal{X}_2$

- $G \cup M_X \in \mathcal{R}$ ,
- $X$  is an independent set,
- No three vertices in  $X$  share a common neighbor in  $G$ , and hence  $T(X)$  is composed of edge-disjoint tepees with distinct tips,
- $\widehat{X}$  is isomorphic to  $\widehat{M}$ , a graph with  $\nu + \phi$  vertices and  $2\phi$  edges, for some  $\phi \leq q$ .

By Lemma 2.2,  $M$  has a triangle-free coloring, that is, a blue-red coloring of its edges with no monochromatic triangle. Fix one such coloring and call it  $\sigma'$ . For each  $X \in \mathcal{X}_2$ , color  $M_X$ , the copy of  $M$  planted on  $X$ , by  $\sigma'$ , and denote the so colored copy by  $M'_X$ . This partitions the tepees in  $T(X)$  into two sets,  $RT(X)$  and  $BT(X)$ , as follows:  $RT(X)$  is the set of tepees in  $G$  over the red edges of  $M'_X$  and  $BT(X)$  is the set of tepees over the blue edges of  $M'_X$ .

We broaden the definition of property  $\mathcal{R}$  to partially colored graphs by saying that a partially colored graph  $F$  belongs to  $\mathcal{R}$  if there is no triangle-free coloring of all edges of  $F$  consistent with the partial coloring. Clearly, if  $F \in \mathcal{R}$  then  $F$  with any partial coloring of  $F$  belongs to  $\mathcal{R}$  too. Hence, by the first property of  $\mathcal{X}_2$  listed above, for all  $X \in \mathcal{X}_2$

$$G \cup M'_X \in \mathcal{R}. \tag{5}$$

We now make a simple but crucial observation which captures in terms of the tepees the restrictions on proper colorings of  $G$  imposed by (3).

**Lemma 3.2** *For every triangle-free coloring  $\chi$  of  $E(G)$  and every  $X \in \mathcal{X}_2$ , there is either a tepee in  $RT(X)$  colored red or a tepee in  $BT(X)$  colored blue.*

**Proof:** Let  $X \in \mathcal{X}_2$ . By (5), for every triangle-free coloring  $\chi$  of  $E(G)$  there is a monochromatic triangle in  $G \cup M'_X$ . By the definition of  $\sigma'$ , we know that  $M'_X$  itself does not contain such a triangle. On the other hand, by the definition of  $\mathcal{X}_2$ , the vertices in  $X$  do not span any edges of  $G$ . Therefore, we conclude that the monochromatic triangle contains exactly one edge of  $M_X$ . The other two edges form a tepee guaranteed by the lemma.  $\square$

We would now like to pre-color a subset of edges of  $G$  according to whether they belong to the tepees from  $RT(X)$  or  $BT(X)$  for some  $X \in \mathcal{X}_2$ . Based on the above lemma, we could then conclude that every proper coloring of  $E(G)$  agrees with the pre-coloring on a set of edges which intersects every subgraph  $\widehat{X}$  on at least two edges, namely, the two edges making up a tepee.

The problem with this approach is that we cannot assign a color to an edge  $e$  in  $G$  once and for all, because  $e$  may belong to a tepee in  $RT(X_1)$  and at the same time to a tepee in  $BT(X_2)$  for some  $X_1 \neq X_2$ . To remedy this obstacle we now move to a more restricted family  $\mathcal{X}_3 \subseteq \mathcal{X}_2$  for which no such clashes occur.

Recall that  $V(M) = \{x_1, \dots, x_\nu\}$ . Let us label the remaining vertices of the prototype graph  $\widehat{M}$  by  $u_1, \dots, u_\phi$ . For each  $X = (v_1, \dots, v_\nu) \in \mathcal{X}_2$  choose an isomorphism  $f$  between  $\widehat{M}$  and  $\widehat{X}$  which maps  $x_a$  onto  $v_a$  for all  $a = 1, \dots, \nu$ , and then set  $w_b = f(u_b)$  for all  $b = 1, \dots, \phi$ .

Assuming that  $n$  is divisible by  $\nu + \phi$ , let  $\pi = \{V_1, \dots, V_\nu, W_1, \dots, W_\phi\}$  be a partition of  $V = V(G)$  into  $\nu + \phi$  parts of equal size  $m = n/(\nu + \phi)$ . (In reality, we put aside an arbitrary set of less than  $\nu + \phi$  vertices, which will have only a negligible effect on the estimates in our proofs.)

We will call a subgraph  $\widehat{X}$  *consistent* with  $\pi$  if, under the above notational convention,  $x_a \in V_a$  for  $1 \leq a \leq \nu$ , and  $w_b \in W_b$  for  $1 \leq b \leq \phi$ . (See, e.g. Figure 2.) Our next lemma establishes the existence of a partition with respect to which a positive fraction of subgraphs  $\widehat{X}$  will be consistent. The attached degree constraint will be utilized only in Sections 4 and 5.

**Lemma 3.3** *There exists a partition  $\pi$  as above and a family  $\mathcal{X}_3 \subseteq \mathcal{X}_2$  with  $|\mathcal{X}_3| = \alpha_3 n^\nu$ , where  $\alpha_3 = \alpha_2/2(\nu + \phi)^{(\nu+q)}$ , such that for every  $X \in \mathcal{X}_3$ , the subgraph  $\widehat{X}$  is consistent with  $\pi$ . Moreover, for every vertex  $v \in V_1 \cup \dots \cup V_\nu$  and for each  $b = 1, \dots, \phi$ , we have  $\frac{3}{4}mp \leq \deg_G(v, W_b) \leq \frac{3}{2}mp$ , where  $m = n/(\nu + \phi)$ .*

**Proof:** Choose an ordered partition of the vertices of  $V$  into  $\nu + \phi$  parts, uniformly at random from all such partitions. For a given subgraph  $\widehat{X}$  the probability that it is consistent with the chosen partition is precisely

$$\frac{m^{\nu+\phi}}{n(n-1)\dots(n-\nu-\phi+1)} \geq (\nu + \phi)^{-(\nu+\phi)}.$$

Therefore the expected number of subgraphs  $\widehat{X}$  with  $X \in \mathcal{X}_2$  that are *not* consistent with the random partition is at most

$$|\mathcal{X}_2|(1 - (\nu + \phi)^{-(\nu + \phi)}) \leq (\alpha_2 - 2\alpha_3)n^\nu.$$

By Markov's inequality, with probability at least  $\alpha_3/(\alpha_2 - \alpha_3)$  at least  $\alpha_3 n^\nu$  such subgraphs are consistent with the random partition. On the other hand, Chernoff's inequality for hypergeometric distributions (see, e.g., [16], Theorem 2.10) and Property (P3) yield that the degree constraint is satisfied with probability  $1 - o(1)$ . Hence, the existence of a required partition follows.  $\square$

Let  $\pi_0 = \{V_1, \dots, V_\nu, W_1, \dots, W_\phi\}$  be a partition guaranteed by Lemma 3.3 and let  $\mathcal{X}_3 \subseteq \mathcal{X}_2$  be the corresponding family. We have now finished refining the initial family  $\mathcal{X}$ . The new subfamily  $\mathcal{X}_3$  is the final one we will work with. Also, the partition  $\pi_0$  will now be fixed for the rest of the paper providing a starting point for our further constructions.

### 3.2 Constellations

Note that for all distinct  $X_1, X_2 \in \mathcal{X}_3$  we have  $RT(X_1) \cap BT(X_2) = \emptyset$ . In view of this and of Lemma 3.2, we are now in a position to assign a pre-coloring to the edges of  $\bigcup_{X \in \mathcal{X}_3} \widehat{X}$  in such a way that every triangle-free coloring of  $G$  agrees with the pre-coloring on at least two edges (which form a tepee) of  $\widehat{X}$ , for every  $X \in \mathcal{X}_3$ . However, for our purposes it is sufficient to concentrate on such an agreement even on one edge of each  $\widehat{X}$ . Because of this excess, we now simplify the structure of subgraphs we will be dealing with through the rest of the paper.

For every  $X \in \mathcal{X}_3$ , instead of  $\widehat{X}$  we will consider a certain subgraph  $S(X)$  of  $\widehat{X}$ , called a *special constellation*. (See Figure 3.) We will shortly define this new notion formally, but for now we note that  $S(X)$  is a star forest (a disjoint union of stars) obtained by erasing one edge of every tepee in  $T(X)$ . The reason we now shift from tepees to constellations is that we will need various estimates on their number, and this boils down to counting star forests, a task much simpler than counting copies of  $\widehat{M}$  which may contain many cycles.

Let  $X = (v_1, \dots, v_\nu) \in \mathcal{X}_3$  and  $v_a v_{a'}$  be an edge in  $M_X$  with  $a < a'$ . For a tepee  $v_a w v_{a'}$  over  $\{v_a, v_{a'}\}$  in  $G$  we will call  $v_a w$  the *left leg* of the tepee

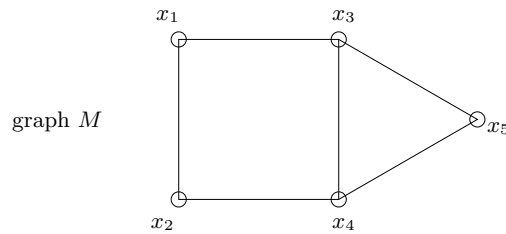
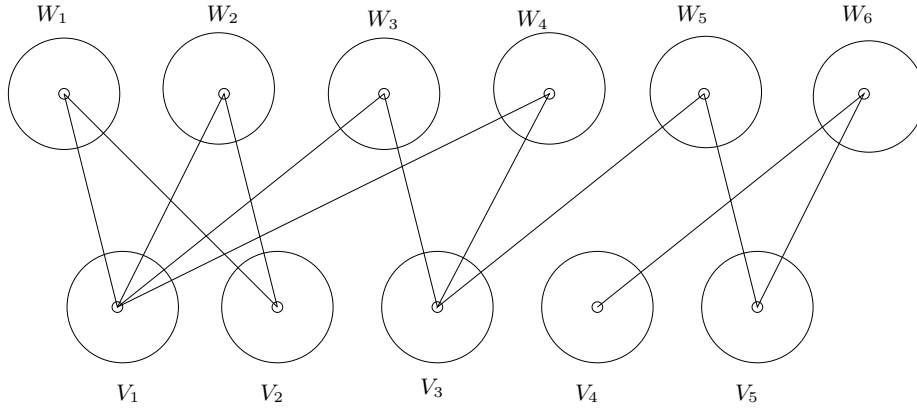


Figure 2: subgraphs  $\widehat{X}$  consistent with  $\pi_0$

and  $v_a w$  the *right leg* of the tepee. Once we can distinguish between the two edges of every tepee in  $T(X)$ , it is easy to define  $S(X)$ .

**Definition 3.2** For each  $X \in \mathcal{X}_3$ , let  $S(X)$  be the spanning subgraph of  $\widehat{X}$  whose edges are the left legs of all tepees in  $T(X)$ . The subgraphs  $S(X)$ ,  $X \in \mathcal{X}_3$ , will be called *special constellations*.

Since for all  $X \in \mathcal{X}_3$  the subgraph  $\widehat{X}$  is isomorphic to  $\widehat{M}$ , it follows that for all  $X \in \mathcal{X}_3$ , constellations  $S(X)$  have the same isomorphism type.

**Definition 3.3** Let  $S$  denote the common isomorphism type of all special constellations  $S(X)$ . That is,  $S$  is a subgraph of  $\widehat{M}$  with  $V(S) = V(\widehat{M})$ , which is a union of  $\nu$  vertex disjoint stars, each centered at one of the vertices

of  $M$ . Their degrees will be denoted by  $\phi_1, \dots, \phi_\nu$ , where  $\phi = \sum \phi_a$ , and  $\phi_\nu = 0$  (because  $x_\nu$  is never adjacent to the left leg of a tepee containing it).

It will now be convenient to relabel the vertices of  $S$  and denote the neighbors of  $x_a$  in  $S$  by  $u_{ab}$ , where  $b = 1, \dots, \phi_a$ . Accordingly, we modify the convention from the previous subsection; namely, we label the vertices of any special constellation  $S(X)$  by  $v_a$  and  $w_{ab}$  in a such a way that there is an isomorphism between  $S$  and  $S(X)$  which maps  $x_a$  onto  $v_a$  and  $u_{ab}$  onto  $w_{ab}$  for all  $a = 1, \dots, \nu$  and  $b = 1, \dots, \phi_a$ . Finally, also to conform with this new notation, let us relabel the sets of the partition  $\pi_0$ :

$$\pi_0 = \{V_1, \dots, V_\nu, W_{11}, \dots, W_{1\phi_1}, \dots, W_{\nu-1,1}, \dots, W_{\nu-1,\phi_{\nu-1}}\},$$

so that for every  $X \in \mathcal{X}_3$ , the special constellation  $S(X)$  satisfies  $v_a \in V_a$  and  $w_{ab} \in W_{ab}$  for all  $a = 1, \dots, \nu - 1$  and  $b = 1, \dots, \phi_a$ .

**Definition 3.4** We say that a copy  $S_0$  of  $S$  is consistent with  $\pi_0$  if there is an isomorphism  $f : S \rightarrow S_0$  for which  $f(x_a) \in V_a$  and  $f(u_{ab}) \in W_{ab}$  for all  $a = 1, \dots, \nu$  and  $b = 1, \dots, \phi_a$ .

Clearly, all special constellations  $S(X)$  are consistent with  $\pi_0$ , but there may be many other copies of  $S$  in  $G$  which are consistent with  $\pi_0$  too. We will simply call them *constellations*.

**Definition 3.5** A *constellation* is any copy  $S_0$  of  $S$  in  $G$  which is consistent with the partition  $\pi_0$ . A *special constellation* is a copy  $S_0$  of  $S$  in  $G$  for which there exists  $X \in \mathcal{X}_3$  such that  $S(X) = S_0$ . The set of all constellations in  $G$  will be denoted by  $\mathcal{C}$ , and the set of all special constellations will be denoted by  $\mathcal{S}$ .

Note that the above definition of special constellations is consistent with the previously stated Definition 3.2.

Although we claim that removing the right leg of each tepee will make our lives easier, we will need to recall them several times in the paper. This will be done via the following “Missing Leg Property”. To state it formally, we introduce another piece of handy notation related to the indexing of the vertices of every  $S(X)$ , or equivalently, of  $S$ . Define  $a'(ab)$  to be the unique index  $a'$  such that  $\widehat{x_a u_{ab} x_{a'}}$  is a tepee in  $\widehat{M}$ . This is well defined because the tepees composing  $\widehat{M}$  are pairwise edge disjoint.

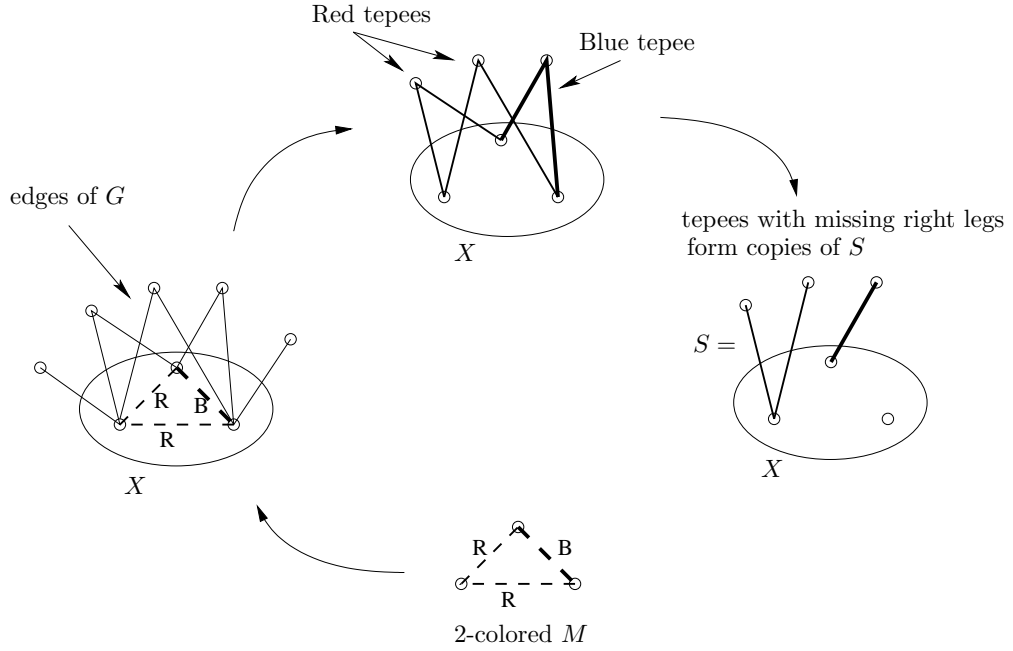


Figure 3: Tepees over  $X$ , with and without right legs – creation of a special constellation  $S$

**Observation 3.4 (“The Missing Leg Property”)** *Let  $X \in \mathcal{X}_3$ .*

- (a) *For every edge  $v_a w_{ab}$  of  $S(X)$  the pair  $w_{ab} v_{a'(ab)}$  is an edge of  $G$  (“the missing leg”).*
- (b)  *$v_a$  and  $v_{a'(ab)}$  are the only vertices in  $X$  which are neighbors of  $w_{ab}$  in  $G$  (only one missing leg per edge in  $S(X)$ ).*
- (c) *The edges of  $S(X)$  together with the missing legs form all the tepees in  $T(X)$ . In other words there are no other tepees over the edges of  $M_X$ .*

So, the structural property that distinguishes special constellations from the ordinary ones is the “Missing Leg Property” that they inherit as offsprings of the  $\hat{X}$ ’s. The most important inheritance, however, is the color-related property to be formulated in Lemma 3.5 below.

But first let us define a partial coloring  $\sigma$  of  $E(G)$  which will play a crucial role in our proof. It will assign colors only to the edges of  $\bigcup_{X \in \mathcal{X}_3} S(X)$ . To begin with, recall that  $\sigma'$  is a fixed, triangle-free coloring of  $E(M)$ , and define an auxiliary coloring of the pairs  $(a, b)$ ,  $a = 1, \dots, \nu$ ,  $b = 1, \dots, \phi_a$ , by  $\sigma(a, b) = \sigma'(x_a x_{a'(ab)})$ .

Now, for every edge  $vw \in \bigcup_{X \in \mathcal{X}_3} S(X)$  with  $v \in V_a$  and  $w \in W_{ab}$ , let  $\sigma(vw) = \sigma(a, b)$ . Note that for any  $X \in \mathcal{X}_3$  and any edge  $e \in S(X)$ ,  $\sigma(e)$  is red or blue according to whether  $e$  belongs to a tepee in  $RT(X)$  or  $BT(X)$ . Note also that the color  $\sigma(e)$  is determined by the partition classes of  $\pi_0$  to which the endpoints of  $e$  belong.

The partial coloring  $\sigma$  is a lens through which we can study complete colorings of  $G$ . For a coloring  $\chi$  of the edges of  $G$ , let

$$\text{Agree}(\chi, \sigma) = \left\{ e \in \bigcup_{X \in \mathcal{X}_3} S(X) : \chi(e) = \sigma(e) \right\}. \quad (6)$$

Now comes a crucial observation that justifies the whole construction erected in this section.

**Lemma 3.5** *Let  $\chi$  be a triangle-free coloring of the edges of  $G$ . Then  $\text{Agree}(\chi, \sigma)$  is a hitting set of  $\mathcal{S}$ , the family of special constellations.*

**Proof:** We must show that for every  $X \in \mathcal{X}_3$  there exists an edge  $e \in S(X)$  such that  $\sigma(e) = \chi(e)$ . Lemma 3.2 states that for every  $X \in \mathcal{X}_3$  and for every triangle-free coloring  $\chi$ , there exists either a tepee in  $RT(X)$  whose edges are colored red by  $\chi$ , or a tepee in  $BT(X)$  whose edges are colored blue. The left leg of such a tepee is an edge of  $S(X)$  for which  $\sigma$  and  $\chi$  agree, by the definition of  $\sigma$ .

□

We may summarize this section as follows: we have shown that the triangle-free colorings of  $G$  are somewhat structured, in the sense that any such coloring must “contain” a hitting set of the family of special constellations (see Lemma 3.5). As it turns out, this structure imposed on triangle-free colorings is quite restrictive; in Section 5, we will capture this restriction in a convenient way. (We are not able to say at this point what we mean by this precisely; let us simply appeal to an analogy and refer the reader to

the “cores” from Section 2.2.1, which capture a strong structural property of cover sets of  $\varepsilon$ -regular bipartite graphs.) In what follows, it will be convenient to consider the hypergraph whose vertices are the edges of  $G$  and the hyperedges are the edge sets of the special constellations. Lemma 3.5 tells us that triangle-free colorings of  $G$  give rise to hitting sets of this hypergraph (or cover sets in the terminology of Section 2.2.1). In Section 4, we will discuss how to “regularize” this hypergraph; from this regularization we will be able to deduce that the structural property imposed on triangle-free colorings of  $G$  is indeed quite strong (again, recall the discussion in Section 2.2.1).

## 4 Regularity

In this section we state and prove a regularity theorem that is the central technical tool in this paper. It is a generalization of the celebrated Szemerédi Regularity Lemma to a setting in which the density is measured not merely by the number of edges but with respect to the number of copies of certain subgraphs.

### 4.1 Classical Regularity

In this subsection we introduce all the necessary results concerning the background on standard graph regularity. See Section 1.4 for more information on regularity.

#### 4.1.1 Dense Regular Graphs

In the classical regularity lemma of Szemerédi [35] a crucial parameter is the density of a bipartite graph defined as a ratio of the number of edges of the graph to the number of all potential edges.

*In the following,  $B$  stands for a bipartite graph with bipartition  $(U, V)$ .*

**Definition 4.1** Let  $U' \subseteq U$  and  $V' \subseteq V$ . The *density* of the pair  $(U', V')$  is defined as

$$d(U', V') = \frac{e_B(U', V')}{|U'| |V'|},$$

where  $e_B(U', V')$  is the number of edges of  $B$  with one endpoint in  $U'$  and the other in  $V'$  (such edges are said to *belong to the pair*  $(U', V')$ ). We will sometimes write  $d(B)$  for  $d(U, V)$ .

The  $\varepsilon$ -regularity of  $B$  reflects low discrepancy between the densities of large, induced subgraphs of  $B$ .

**Definition 4.2** Let  $\varepsilon > 0$ . A bipartite graph  $B$  is called  $\varepsilon$ -regular, if for each choice of  $U' \subseteq U$  and  $V' \subseteq V$ , with  $|U'| \geq \varepsilon|U|$  and  $|V'| \geq \varepsilon|V|$ , we have

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

Sometimes the pair  $(U, V)$  itself is called  $\varepsilon$ -regular.

The following result from [5] and [2] gives a sufficient criterion for regularity in terms of vertex degrees and co-degrees only.

**Lemma 4.1** Let  $\varepsilon > 0$  and  $d(U, V) = d > 2\varepsilon$ . Assume further that  $|U| \geq 2/\varepsilon$  and denote by  $W$  the set of all pairs  $\{v, v'\}$  of vertices of  $U$  for which

- (i)  $\deg(v), \deg(v') > (d - \varepsilon)|V|$ , and
- (ii)  $\text{codeg}(v, v') < (d + \varepsilon)^2|V|$ .

If  $|W| > \frac{1}{2}(1 - 5\varepsilon)|U|^2$ , then  $B$  is  $\varepsilon'$ -regular, where  $\varepsilon' = (16\varepsilon)^{1/5}$ .

The next, simple fact is a direct consequence of Definition 4.2.

**Observation 4.2** Let  $\varepsilon' > \varepsilon > 0$ . If  $B$  is  $\varepsilon$ -regular, and  $V' \subset V$ , with  $|V'| = \varepsilon'|V|$ , then the pair  $(U, V')$  is  $\varepsilon''$ -regular, where  $\varepsilon'' = \varepsilon/\varepsilon'$ .

One of the basic applications of  $\varepsilon$ -regularity is for counting copies of a given graph in a structure composed of several  $\varepsilon$ -regular and dense pairs of vertex sets. We do not give references, as the result seems to belong to the “local folklore”.

**Proposition 4.3** For every  $\varepsilon > 0$ ,  $0 < d < 1$  and an integer  $k$  such that  $\varepsilon < d^k$  the following is true. Let  $V_i$ ,  $i = 1, \dots, k$ , be disjoint vertex sets with  $\min |V_i| > n_0 = (2/d)^k$  and let  $B_{ij}$ ,  $1 \leq i < j \leq k$ , be  $\varepsilon$ -regular graphs with bipartitions  $(V_i, V_j)$  and densities  $d(B_{ij}) = d_{ij}$ , where  $\min d_{ij} > d$ . Then the

number  $\#(K_k, \bigcup_{i,j} B_{ij})$  of complete graphs  $K_k$  contained in the union  $\bigcup_{i,j} B_{ij}$  satisfies, with  $\varepsilon' = \varepsilon d^{-k}$ ,

$$\#(K_k, \bigcup_{i,j} B_{ij}) \stackrel{\varepsilon'}{\sim} \prod_{i=1}^k |V_i| \prod_{i,j} d_{ij},$$

If a less precise estimate of the copies of  $K_k$  is needed, then the  $\varepsilon$ -regularity can be replaced with a less restrictive notion. We say that a graph  $G$  is  $(\varrho, d)$ -dense if for all  $U \subseteq V(G)$  with  $|U| \geq \varrho|V(G)|$ , we have  $e_G(U) \geq d \binom{|U|}{2}$ . When verifying this property it suffices to check only all subsets  $U$  of size  $|U| = \lfloor \varrho|V(G)| \rfloor$ . The next result comes from [31].

**Proposition 4.4** *For all  $d > 0$  and  $k$  there exist  $\varrho, n_0, c_0 > 0$  such that for every  $(\varrho, d)$ -dense graph  $G$  with  $n = |V(G)| \geq n_0$  we have  $\#(K_k, G) \geq c_0 n^k$ .*

Although we do not need it in our proofs, we now state the celebrated Szemerédi Regularity Lemma. This is not exactly the original version, but one which can be compared with the sparse version from the next subsection. Recall the definition of equipartition given at the end of Section 1.

**Theorem 4.5 (Szemerédi Regularity Lemma)** *For all  $\varepsilon > 0$  and integers  $t_0$  and  $r$ , there exist  $T_0$  and  $n_0$  such that if  $G_1, \dots, G_r$  are graphs with a common vertex set  $V$ ,  $|V| = n > n_0$ , then every equipartition of  $V$  into  $t_0$  parts has a refined equipartition  $V = W_1 \cup \dots \cup W_t$ , where  $t_0 \leq t \leq T_0$ , and all but at most  $\varepsilon n^2$  edges of  $G_1 \cup \dots \cup G_r$  belong to the pairs  $\{W_i, W_j\}$  that are  $\varepsilon$ -regular with respect to all graphs  $G_i$ ,  $i = 1, \dots, r$ .*

#### 4.1.2 Sparse Regular Graphs

The graph  $G$ , one of the main constituents of the proof of Theorem 2.1, is relatively sparse, with density of order  $1/\sqrt{n}$ . Below we present a version of the regularity lemma designed for sequences of graphs with  $n$  vertices and density  $p = o(1)$ .

As before, let  $B$  be a bipartite graph with bipartition  $(U, V)$ , and let  $0 < p \leq 1$ .

**Definition 4.3** Let  $U' \subseteq U$  and  $V' \subseteq V$ . The  $p$ -density of the pair  $(U', V')$  is defined as

$$d_p(U', V') = \frac{e_B(U', V')}{p|U'||V'|}.$$

We will sometimes write  $d_p(B)$  for  $d_p(U, V)$ .

As in the dense case, the  $(\varepsilon, p)$ -regularity of  $B$  (as defined below) reflects low discrepancy between the  $p$ -densities of large, induced subgraphs of  $B$ .

**Definition 4.4** Let  $\varepsilon > 0$  and  $0 < p \leq 1$ . A bipartite graph  $B$  is  $(\varepsilon, p)$ -regular if for each choice of  $U' \subseteq U$  and  $V' \subseteq V$ , with  $|U'| \geq \varepsilon|U|$  and  $|V'| \geq \varepsilon|V|$ , we have

$$|d_p(U', V') - d_p(U, V)| < \varepsilon.$$

Sometimes the pair  $(U, V)$  itself is called  $(\varepsilon, p)$ -regular.

There is an analog of Observation 4.2 in the sparse case.

**Observation 4.6** Let  $\varepsilon' > \varepsilon > 0$  and  $0 < p \leq 1$ . If  $B$  is  $(\varepsilon, p)$ -regular, and  $V' \subset V$ , with  $|V'| = \varepsilon'|V|$ , then the pair  $(U, V')$  is  $(\varepsilon'', p)$ -regular, where  $\varepsilon'' = \varepsilon/\varepsilon'$ .

A basic feature of  $\varepsilon$ -regular graphs is that most of the vertex degrees are under control. Here is a useful fact exploited in Section 6.

**Proposition 4.7** If  $(U, V)$  is  $(\varepsilon, p)$ -regular and  $W \subseteq V$  satisfies  $|W| \geq \varepsilon|V|$ , then at least  $(1 - \varepsilon)|U|$  vertices of  $U$  have each at least  $(d_p(U, V) - \varepsilon)p|W|$  neighbors in  $W$ .

The above notions of  $p$ -density and  $(\varepsilon, p)$ -regularity extend naturally to pairs of disjoint subsets of vertices of non-bipartite graphs, for which we now introduce one more notion, specific just to the sparse case.

**Definition 4.5** For  $D > 0$  and  $0 < p \leq 1$  we say that a graph  $F$  on  $n$  vertices has no  $(D, p)$ -dense patches, if for every two disjoint sets of vertices  $W_1, W_2 \subset V(F)$  with  $|W_1|, |W_2| > n/\log n$ , we have

$$d_p(W_1, W_2) < D.$$

This property, called  $(p; D, 1/\log n)$ -boundedness in [16], page 215, guarantees that the so called partition index is bounded from above, and thus allows to repeat the proof of the Szemerédi Regularity Lemma *mutatis mutandis*.

But there is another profit from this assumption. In general, there is no simple analog of Proposition 4.3 in the sparse case. However, there are several places in this paper where we need estimates of the number of stars consistent with a certain family of disjoint subsets of vertices. Fortunately, in this simple case, a similar counting result is true, provided some kind of boundedness is assumed. We prove it here in a specific form suitable for our purposes.

**Proposition 4.8** *For every  $\varepsilon > 0$ ,  $D > 0$ , and integer  $k \geq 2$  such that  $\varepsilon \leq \min\{k^{-4}, 2^{-6}D^{-6k}\}$  there exists  $n_0$  for which the following is true. Let  $V_0, W_j, j = 1, \dots, k$ , be disjoint vertex sets of sizes  $n_0 \leq |W_j|/2 \leq |V_0| \leq |W_j|$ ,  $n = |V_0| + |W_1| + \dots + |W_k|$  and  $0 < p = p(n) \leq 1$ . Let  $B_j, 1 \leq j \leq k$ , be  $(\varepsilon, p)$ -regular graphs with bipartitions  $(V_0, W_j)$  and  $p$ -densities  $d_p(B_j) = d_j$ , where  $\min d_j \geq \varepsilon^{1/2k}$ . Assume further that the graphs  $B_j$  have no  $(D, p)$ -dense patches and that the maximum degree in the graph  $\bigcup_j B_j$  is at most  $Dnp$ . Let  $\#(S_k, \bigcup_j B_j)$  be the number of copies of the star  $S_k$  contained in  $\bigcup_j B_j$  and such that the vertex of degree  $k$  is in  $V_0$  and the vertices of degree one each belong to a distinct set  $W_j$ . Then we have,*

$$\#(S_k, \bigcup_j B_j) \stackrel{\varepsilon^{1/3}}{\sim} |V_0| p^k \prod_{j=1}^k |W_j| d_j.$$

**Proof:** For any subset  $V'$  of  $V_0$ , set  $S(V') = \#(S_k, \bigcup_j B'_j)$ , where  $B'_j$  is the subgraph of  $B_j$  induced by the vertex set  $V' \cup W_j$ . Let  $\deg_j(v) = \deg_{B_j}(v)$  be the number of neighbors of vertex  $v$  in graph  $B_j$ . Clearly,

$$S(V') = \sum_{v \in V'} \prod_{j=1}^k \deg_j(v),$$

so we would be done if we knew that all, not just *almost* all, vertices of  $V_0$  satisfy  $\deg_j(v) \stackrel{\varepsilon}{\approx} d_j p |W_j|$  for all  $j$ . Unfortunately, this is not so, and we have to somehow handle the exceptional degrees. Of course, for a lower bound on  $S(V_0)$  we can just leave them out. Let

$$V_{small}^j = \{v \in V_0 : \deg_j(v) < (d_j - \varepsilon)p|W_j|\}.$$

By the  $(\varepsilon, p)$ -regularity of  $B_j$  it follows that  $|V_{small}^j| < \varepsilon|V_0|$ . Let

$$V_{small} = \bigcup_{j=1}^k V_{small}^j.$$

Then

$$\begin{aligned} S(V_0) &\geq S(V_0 \setminus V_{small}) = \sum_{v \in V_0 \setminus V_{small}} \prod_{j=1}^k \deg_j(v) \geq \sum_{v \in V_0 \setminus V_{small}} \prod_{j=1}^k (d_j - \varepsilon)p|W_j| \\ &\stackrel{*}{\geq} |V_0|(1 - k\varepsilon)p^k \prod_{j=1}^k d_j(1 - \varepsilon^{3/4})|W_j| \geq |V_0|(1 - k\varepsilon)(1 - \varepsilon^{3/4})^k \prod_{j=1}^k d_j|W_j| \\ &\geq |V_0|(1 - (k+1)\varepsilon^{3/4}) \prod_{j=1}^k d_j|W_j| \stackrel{**}{\geq} (1 - \varepsilon^{1/3})|V_0| \prod_{j=1}^k d_j|W_j|. \end{aligned}$$

Note that for inequality (\*) above we use  $d_j \geq \varepsilon^{1/4}$ , which follows from  $d_j \geq \varepsilon^{1/2k}$  and  $k \geq 2$ . The assumption  $1/k \geq \varepsilon^{1/4}$  validates inequality (\*\*).

We now proceed to prove a corresponding upper bound on  $S(V_0)$ . We will partition the vertices of  $V_0$  into three classes  $V_{med}$ ,  $V_{big}$  and  $V_{huge}$ , according to their degrees, and will see that, as expected, the majority of the contribution to  $S(V_0)$  comes from  $V_{med}$ , whereas the contribution from the other two classes is negligible. Let

$$V_{med} = \{v \in V_0 : \forall j \deg_j(v) \leq (d_j + \varepsilon)p|W_j|\},$$

$$V_{huge} = \{v \in V_0 : \exists j \deg_j(v) > Dp|W_j|\}$$

and

$$V_{big} = V \setminus (V_{med} \cup V_{huge}).$$

Note that  $S(V) = S(V_{med}) + S(V_{big}) + S(V_{huge})$ . As in the case of the lower bound, it is immediate that

$$S(V_{med}) \leq |V_0| \prod_j (d_j + \varepsilon)p|W_j| \stackrel{***}{\leq} |V_0|p^k(1 + \varepsilon^{1/3}/2) \prod_j d_j|W_j|, \quad (7)$$

where, as above, (\*\*\*) follows from the assumptions that  $d_j \geq \varepsilon^{1/4}$  and  $k \leq \varepsilon^{-1/4}$ .

Next, observe that by the  $(\varepsilon, p)$ -regularity of  $B_j$ 's we have  $|V_{big}| < \varepsilon k |V_0|$ . Hence

$$\begin{aligned} S(V_{big}) &< |V_{big}| (Dp)^k \prod_j |W_j| < \varepsilon k |V_0| (Dp)^k \prod_j |W_j| \\ &\leq \sqrt{\varepsilon} |V_0| (\varepsilon^{\frac{1}{2k}} Dp)^k \prod_j |W_j| \leq \sqrt{\varepsilon} |V_0| (Dp)^k \prod_j d_j |W_j|. \end{aligned} \quad (8)$$

Finally, by the assumptions on the maximum degree and the lack of  $(D, p)$ -dense patches, for all  $v \in V_0$  and for all  $j = 1, \dots, k$  we have  $d_j(v) < Dnp$  and  $|V_{huge}| < kn/\log n$ . Hence

$$S(V_{huge}) < k \frac{n}{\log n} (Dnp)^k = o(|V_0| p^k \prod_j d_j |W_j|). \quad (9)$$

Putting (7), (8) and (9) together gives

$$S(V) \leq (1 + \varepsilon^{1/3}) |V_0| p^k \prod_j d_j |W_j|.$$

□

Of course, the above result can easily be adopted to asymptotically count copies of star forests as well.

**Example 4.9** Consider a graph  $G \in \mathcal{G}$ , a star forest  $S$ , and a partition  $\pi_0$  of  $V = V(G)$  as in Section 3. For each  $a = 1, \dots, \nu$  and  $b = 1, \dots, \phi_a$ , let  $V'_a \subseteq V_a$  and  $W'_{ab} \subseteq W_{ab}$  be sets of size at least  $n/\log n$ . By Property (P5) of  $\mathcal{G}$  there are no  $(2, p)$ -dense patches in  $G$ . Moreover, by the same property, the bipartite subgraphs  $F'_{ab} = G[V'_a, W'_{ab}]$  are all  $(2n^{-1/5}, p)$ -regular. Thus, by  $\nu$  simultaneous applications of Proposition 4.8 (one for each star of  $S$ ), we obtain an asymptotic formula for the number  $\kappa'$  of the constellations in  $G$  contained in the union of  $F'_{ab}$ 's:

$$\kappa' \sim \prod_{a=1}^{\nu} |V'_a| \prod_{b=1}^{\phi_a} |W'_{ab}|.$$

In particular, the number  $\kappa = |\mathcal{C}|$  of all constellations in  $G$  consistent with  $\pi_0$  satisfies:

$$\kappa = |\mathcal{C}| \sim m^{\nu+\phi} p^\phi \quad \text{where} \quad m = \frac{n}{\nu + \phi}.$$

(See, e.g., Figure 4)

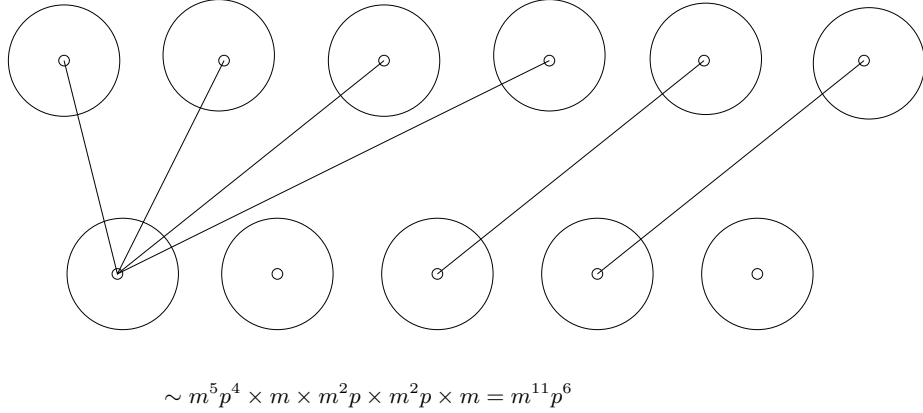


Figure 4: Illustration of the example

Finally, we state a sparse version of the Szemerédi Regularity Lemma. Theorem 4.10 below was independently observed by Y. Kohayakawa and V. Rödl and, in fact, its proof is a simple modification of Szemerédi’s proof. For applications of this variant of the regularity lemma see [20], [22] and [16], Section 8.3.

**Theorem 4.10 (Sparse Regularity Lemma)** *For all  $\varepsilon > 0$ ,  $D > 0$ , and integers  $t_0$  and  $r$ , there exist  $T_0$  and  $n_0$  such that if  $G_1, \dots, G_r$  are graphs with a common vertex set  $V$ ,  $|V| = n > n_0$ , and  $p = p(n)$  is such that  $G_i$  does not have  $(D, p)$ -dense patches for all  $i = 1, 2, \dots, r$ , then every equipartition of  $V$  into  $t_0$  parts has a refined equipartition  $V = W_1 \cup \dots \cup W_t$ , where  $t_0 \leq t \leq T_0$ , and all but at most  $\varepsilon n^2 p$  edges of  $G_1 \cup \dots \cup G_r$  belong to the pairs  $\{W_i, W_j\}$  that are  $(\varepsilon, p)$ -regular with respect to all graphs  $G_i$ ,  $i = 1, \dots, r$ .*

This theorem is essentially the same as the one presented in [22]. The theorem in [22] guarantees that all but an  $\varepsilon$ -proportion of the *pairs*  $(W_i, W_j)$  in the partition are  $(\varepsilon, p)$ -regular with respect to all  $G_i$ . Given the fact that there is an upper bound on the  $p$ -density of every pair (i.e., no  $(D, p)$ -dense patches), we have that at most an  $\varepsilon D$ -proportion of the *edges* belong to such pairs. Hence the theorems are equivalent (after re-scaling the constants).

## 4.2 The Subgraph Regularity Lemma

Let us summarize the state-of-art at the end of Section 3. Given a graph  $G \in \mathcal{G} \setminus \mathcal{R}$ , as a far reaching consequence of assumption (3) of Theorem 2.1, we have constructed a star forest  $S$  and a partition  $\pi_0$  of the vertices of  $G$  into  $|V(S)|$  sets of equal size  $m = \lfloor n/|V(S)| \rfloor$  and, possibly, one set of less than  $|V(S)|$  vertices. (For clarity, we will be further assuming that  $n$  is divisible by  $|V(S)|$ .) Based on this partition, we have focused on the set  $\mathcal{C}$  of all constellations, that is all copies of  $S$  in  $G$  that are consistent with  $\pi_0$ . Among the constellations we have distinguished a set  $\mathcal{S}$  of special constellations. These are the ones that correspond to selected copies  $M_X$  of a certain graph  $M$ , such that  $G \cup M_X \in \mathcal{R}$ . Later in the proof of our main result, we are going to apply a regularity lemma to the family  $\mathcal{S}$ . The purpose of this section is to prove such a regularity lemma in a more abstract setting, not using the concrete definition of the family  $\mathcal{S}$ .

Consider a hypergraph whose vertices are the edges of  $G$  and whose edges are the members of  $\mathcal{S}$ . By abuse of notation we may also denote the hypergraph itself by  $\mathcal{S}$ . As we roughly explained in Section 2.2, we wish to apply a regularity lemma to  $\mathcal{S}$ , and partition its vertices, i.e. the edges of  $G$ . Our objective will be to find such a partition so that various parts of the partition will carry the edges of  $\mathcal{S}$  (special constellations) in a fairly regular manner with respect to all constellations, just as regular pairs in a usual Szémeredi partition of a graph have edge density that is, in a sense, smoothly distributed.

This partition of  $\mathcal{S}$  will rely heavily on the underlying graph  $G$ , and we will start out by partitioning the vertices of  $G$ , thus inducing a partition of the edges of  $G$  (which are the vertices of  $\mathcal{S}$ !). Only in the next step will we focus on the sets thus formed and introduce a further partition of the *edges* of  $G$  that is *not* induced simply by partitioning the vertices of  $G$ . We will then iterate this procedure in a manner similar to the one introduced by Frankl and Rödl in [12], see the next subsection. We are aware that the two roles of edges of  $G$  as vertices of  $\mathcal{S}$  may be confusing, and hence, having pointed out the hypergraph structure, we will now stick to the language of subgraphs, thus avoiding ambiguity.

Let  $F$  be the set of all edges of  $G$  with one endpoint in  $V_a$  and the other in  $W_{ab}$  for all  $a = 1, \dots, \nu$  and  $b = 1, \dots, \phi_a$ . We will identify  $F$  with the spanning subgraph of  $G$  consisting of all these edges.

The regularity lemma we are aiming at will later be applied only to subgraphs  $F$  of graphs  $G \in \mathcal{G} \setminus \mathcal{R}$ , created in the above way by selecting first a star forest  $S$  and an initial partition  $\pi_0$ . However, it will be convenient to formulate it more generally for all graphs like  $F$ , that is, for graphs structured by  $S$  and  $\pi_0$  but with no reference to  $\mathcal{G}$  and to assumption (3) of Theorem 2.1. Moreover, with no additional effort, our regularity lemma can be proved with respect to subgraphs other than  $S$ . Therefore, we will soon set up a general framework in which the regularity lemma will be stated and proved, bearing in mind that the only application we need is that for the star forest  $S$ . But first, in order to give stronger foundations for our generalization, we describe one more instance of regularity.

#### 4.2.1 The Hypergraph Regularity of Frankl and Rödl

Recently, Frankl and Rödl [12] proved a powerful regularity lemma for 3-uniform hypergraphs. A simplified version of that lemma is described briefly here as another piece of motivation for the general framework we are up to.

Given three disjoint sets  $V_1, V_2, V_3$ , a *triad* is a triple  $F = (F^{12}, F^{23}, F^{13})$  of bipartite graphs with vertex sets  $V_1 \cup V_2$ ,  $V_2 \cup V_3$  and  $V_1 \cup V_3$ . We will identify  $F$  with the union of its three component graphs. We refer to a 3-partite 3-uniform hypergraph  $\mathcal{H}$  with a 3-partition  $(V_1, V_2, V_3)$  as a *3-hypergraph*. For a triad  $F$  with the same vertex partition as  $\mathcal{H}$ , we say that  $F$  *underlies*  $\mathcal{H}$  if  $\mathcal{H} \subseteq \text{Tr}(F)$ , where  $\text{Tr}(F)$  is the family of the vertex sets of the triangles in the graph  $F$ .

Let  $\mathcal{H}$  be a 3-hypergraph with an underlying triad  $F$ . The *density* of  $\mathcal{H}$  with respect to  $F$  is defined by

$$d_{\mathcal{H}}(F) = \frac{|\mathcal{H} \cap \text{Tr}(F)|}{|\text{Tr}(F)|}.$$

In other words, the density measures the proportion of triangles of  $F$  which are triples of  $\mathcal{H}$ .

Let  $\delta > 0$ . A 3-hypergraph  $\mathcal{H}$  is said to be  *$\delta$ -regular* with respect to an underlying triad  $F$  if for every subtriad  $Q = (Q^{12}, Q^{23}, Q^{13})$  of  $F$ ,  $Q^{12} \subset F^{12}$ ,  $Q^{23} \subset F^{23}$ ,  $Q^{13} \subset F^{13}$ , with  $|\text{Tr}(Q)| > \delta |\text{Tr}(F)|$  we have

$$|d_{\mathcal{H}}(Q) - d_{\mathcal{H}}(F)| < \delta.$$

Thus, the  $\delta$ -regularity of  $\mathcal{H}$  reflects low discrepancy between the densities of  $\mathcal{H}$  with respect to those subtriads of  $F$  which are relatively rich in triangles. With a fixed  $\mathcal{H}$  in mind, we call the triad itself  $\delta$ -regular. A triad which is *not*  $\delta$ -regular is called  $\delta$ -irregular.

Let  $V$  be a set, and let  $[V]^2$  denote the set of all 2-element subsets of  $V$ . An  $(l, t, \epsilon)$ -partition  $\Pi$  of  $[V]^2$  consists of an equipartition  $V = V_1 \cup \dots \cup V_t$ , together with a system of edge-disjoint bipartite graphs  $F_k^{ij}$  with bipartitions  $(V_i, V_j)$ ,  $1 \leq i < j \leq t$ ,  $1 \leq k \leq l_{ij} \leq l$ , such that

- $\left| \bigcup_{k=1}^{l_{ij}} F_k^{ij} \right| = |V_i||V_j|$  for all  $i, j$ ,  $1 \leq i < j \leq t$ , and
- all but at most  $\epsilon \binom{t}{2} m^2$  pairs  $\{v_i, v_j\}$ ,  $v_i \in V_i$ ,  $v_j \in V_j$ ,  $1 \leq i < j \leq t$ , belong to graphs  $F_k^{ij}$  which are  $\epsilon$ -regular.

Note that there are at most  $\binom{t}{3} l^3$  triads  $F$  made up of the graphs of  $\Pi$ . Let  $\mathcal{P}_{\Pi}^{\text{irr}}$  be the set of all  $\delta$ -irregular triads among them.

For a 3-uniform hypergraph  $\mathcal{H} = (V, E)$ , with  $|V| = n$ , we say that an  $(l, t, \epsilon)$ -partition  $\Pi$  of  $[V]^2$  is  $\delta$ -regular if

$$\sum_{F \in \mathcal{P}_{\Pi}^{\text{irr}}} |\text{Tr}(F)| < \delta n^3.$$

**Theorem 4.11** *For every  $\delta > 0$ , integers  $t_0$  and  $l_0$ , and all decreasing functions  $\epsilon(\ell)$ , there exist  $T_0$ ,  $L_0$ , and  $N_0$  such that any 3-uniform hypergraph  $\mathcal{H} = (V, E)$  with  $|V| > N_0$  admits a  $\delta$ -regular,  $(l, t, \epsilon(l))$ -partition for some  $t$  and  $l$  satisfying  $t_0 \leq t \leq T_0$  and  $l_0 \leq l \leq L_0$ .*

**Remark 4.1** In [12] a more general and powerful hypergraph lemma guarantees, for every integer function  $r = r(\ell)$ , a stronger,  $(\delta, r)$ -regular partition. Above we discussed the case  $r \equiv 1$  only.

## 4.2.2 General Subgraph Framework

In what follows  $H$  always stands for a fixed graph with respect to copies of which other graphs are regularized. The reader can bear in mind that the cases of  $H = K_2$  and  $H = K_3$  correspond, respectively, to the Szemerédi and Frankl-Rödl regularity schemes, and that  $H = S$ , the star forest from Definition 3.3, will be our principal (and only) application.

Given a graph  $H$  with  $h$  vertices and an integer  $m$ , let  $H^m$  be the order  $m$  blow-up of  $H$ , i.e., the  $h$ -partite graph obtained by replacing each vertex  $x$  of  $H$  by a set  $V_x$  of  $m$  vertices, and by replacing every edge  $xy$  of  $H$  by the complete bipartite graph  $K_{m,m}$  spanned between  $V_x$  and  $V_y$ . Let  $\mathcal{F}(H, m)$  denote the family of all subgraphs  $F$  of  $H^m$ . The graphs  $F \in \mathcal{F}(H, m)$  will be called *H-graphs* and a copy  $H'$  of  $H$  in  $F$  will be declared *consistent* if for each  $x \in V(H)$ , its image  $x'$  belongs to  $V_x$  (see Figure 5). By a copy of  $H$  in an  $H$ -graph we will always mean one which is consistent. Let  $\mathcal{C}(F) = \mathcal{C}_H(F)$  be the set of all copies of  $H$  in  $F$ .

Besides graph  $F$ , the other important input to our regularity lemma is an arbitrarily specified subfamily of  $\mathcal{C}(F)$ , the elements of which will be called *special copies*. The subfamily itself will be denoted by  $\mathcal{S}$ . By choosing this notation we put emphasis on our primary application with  $H = S$  and  $\mathcal{S}$  – the family of special constellations. To obtain the standard Szemerédi setting take  $H = K_2$ ,  $F = K_{m,m}$  and  $\mathcal{S}$  – the set of edges of a bipartite graph. For the hypergraph setting of Frankl and Rödl take  $H = K_3$ ,  $F = K_{m,m,m}$  and  $\mathcal{S}$  – the set of triples of a 3-hypergraph.

### 4.2.3 Measuring Density

In the classical case of graphs, regularity is related to the density measured by the ratio of the number of edges to all potential edges. For 3-uniform hypergraphs, regularity will be with respect to the ratio of triples to the triangles in the underlying triad. Similarly, we will measure the density of various (spanning) subgraphs  $R$  of  $F$  by the ratio of the number of special copies of  $H$  to all copies of  $H$  contained in  $R$ . Since our primary application is in the sparse case, the density will have to be normalized by a scaling factor  $p^*$ , which will typically depend both on  $H$  and on the average density of the sparse graph in question.

Having fixed  $H$ , let us also fix  $F \in \mathcal{F}(H, m)$  with  $V(F) = \bigcup_{x \in V(H)} V_x$ ,  $\mathcal{S} \subseteq \mathcal{C}_H(F)$  and  $0 < p^* \leq 1$ . Any subgraph  $R$  of  $F$  is itself a member of  $\mathcal{F}(H, m)$  and thus  $\mathcal{C}(R) = \mathcal{C}_H(R)$  is already defined, meaning the family of copies of  $H$  in  $R$ . As for the special copies, though, they can only be inherited from  $F$ . We set

$$c_R = |\mathcal{C}(R)| \text{ and } s_R = |\mathcal{S} \cap \mathcal{C}(R)|.$$

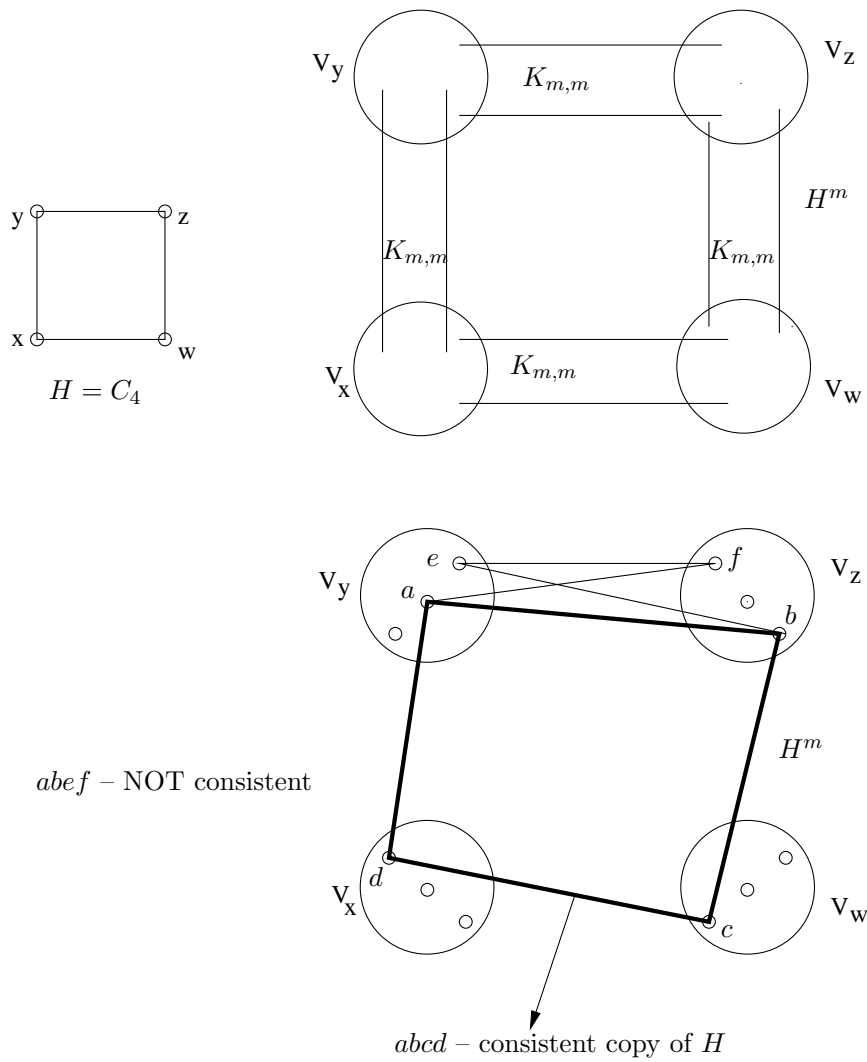


Figure 5: consistent copy of  $H = C_4$

For all subgraphs  $R \subseteq F$ , with  $c_R > 0$ , we define their  $(\mathcal{S}, p^*)$ -density by

$$d_R = \frac{s_R}{p^* c_R}. \quad (10)$$

The normalizing factor  $p^*$  helps in bounding the density from below. It should be related closely to the ratio  $s_F/c_F$  between the numbers of special copies and all copies of  $H$  in the entire graph  $F$ .

The price we pay for using a normalized density is the potential presence of some kind of dense patches. It is unavoidable to have small subgraphs  $R$  with unbounded  $d_R$  (e.g., if  $c_R = s_R = 1$  and  $p^* = o(1)$ ). Thus, all we can do is to control the  $(\mathcal{S}, p^*)$ -density of large subgraphs of  $R$ , where “large” means “rich in copies of  $H$ ”. Set

$$\kappa = |\mathcal{C}(F)|$$

for the total number of copies of  $H$  in  $F$  (“total volume” of  $F$ ).

**Definition 4.6** For fixed  $H$  and  $D^* > 0$ , an  $F \in \mathcal{F}(H, m)$  with  $n = hm$  vertices, together with a nonempty family  $\mathcal{S} \subseteq \mathcal{C}(F)$ , is said to *have no  $(D^*, p^*)$ -dense  $H$ -patches*, if for every subgraph  $R \subseteq F$  with  $c_R \geq \kappa/\log^2 n$  we have  $d_R \leq D^*$ .

Note that the assumption of no dense patches applied to  $R = F$  and the definition of density in (10) imply together that

$$p^* \geq \frac{s_F}{D^* c_F} \geq \frac{h^h}{D^*} n^{-h}, \quad (11)$$

unless  $s_F = 0$ .

The property of having no dense patches, together with an upper bound on the number of copies that one edge may belong to, implies that subgraphs of  $F$  with few copies of  $H$  must also have few special ones.

**Definition 4.7** An  $H$ -graph  $F$  with  $n$  vertices is called  *$H$ -uniform* if no edge of  $F$  belongs to more than  $\kappa/\log^3 n$  copies of  $H$ .

**Proposition 4.12** *Assume that an  $H$ -graph  $F$  (together with  $\mathcal{S}$ ) has no  $(D^*, p^*)$ -dense  $H$ -patches and is  $H$ -uniform. Let  $R \subseteq F$  with  $c_R \leq \kappa/\log^2 n$ . Then  $s_R \leq (1 + o(1))D^* p^* \kappa/\log^2 n$ .*

**Proof:** Given  $R$  with  $c_R \leq \kappa/\log^2 n$ , we add to it extra edges of  $F$ , enlarging  $R$  to obtain a supergraph  $R'$  with  $\kappa/\log^2 n \leq c_{R'} = (1 + o(1))\kappa/\log^2 n$ . This can easily be accomplished by adding one edge at a time, and noting that each edge can increase the number of copies of  $H$  by at most  $\kappa/\log^3 n$ . Thus, because there are no  $(D^*, p^*)$ -dense patches in  $F$ , it follows that  $d_{R'} \leq D^*$ , and hence  $s_R \leq s_{R'} \leq (1 + o(1))D^*p^*c_{R'}$ .  $\square$

#### 4.2.4 Partitions and Polyads

The subgraph of an  $H$ -graph  $F$ , induced by a pair  $(V_x, V_y)$  corresponding to the edge  $xy \in H$ , will be denoted by  $F_{xy}$ . The sets  $V_x$  for all  $x \in V(H)$  and the graphs  $F_{xy}$  for all  $xy \in E(H)$  form *the initial partition*  $\Pi_0$  of a fixed  $H$ -graph  $F$ . The regularity lemma we prove in this section involves refinements of the sets and graphs of  $\Pi_0$  and towards this we define a  $(t, l)$ -partition, by generalizing the definition of  $\Pi_0$ .

Informally speaking, a  $(t, l)$ -partition is an equipartition obtained by splitting each set  $V_x$  into  $t$  disjoint subsets  $V_x^1, \dots, V_x^t$ . This partitions every subgraph  $F_{xy}$  into  $t^2$  tubes

$$F_{xy}^{i,j} = F_{xy}[V_x^i, V_y^j], \quad 1 \leq i, j \leq t,$$

which are induced bipartite subgraphs spanned between the partition sets. We then partition every tube into  $l(xy, i, j) \leq l$  edge-disjoint subgraphs.

**Definition 4.8** A  $(t, l)$ -partition  $\Pi$  consists of the following ingredients:

$$\text{equipartitions} \quad V_x = V_x^1 \cup \dots \cup V_x^t, \quad x \in V(H) \quad (12)$$

and

$$\text{partitions} \quad F_{xy}^{i,j} = \bigcup_{k=1}^{l(xy,i,j)} F_{xy}^{i,j,k}, \quad xy \in E(H), \quad i, j = 1, \dots, t. \quad (13)$$

A *refinement* of  $\Pi$  is any  $(t', l')$ -partition, with  $t' \geq t$ ,  $l' \geq l$ , which consists of refinements of (12) and (13).

In particular, any  $(t, l)$ -partition is a refinement of the initial  $(1, 1)$ -partition  $\Pi_0$ , and a  $(t, l)$ -partition divides the vertex set into  $th$  classes and the edge set into  $\sum_{xy \in E(H)} l(xy, i, j) \leq |E(H)|t^2l$  classes.

We now define a basic notion whose role is similar to that of bipartite subgraphs in the classical Szemerédi Lemma, and to that of triads in the hypergraph regularity lemma of Frankl and Rödl.

**Definition 4.9** A *polyad*  $P$  consistent with a  $(t, l)$ -partition  $\Pi$  is a subgraph of  $F$  obtained by selecting one subclass  $V_x^{i(x)}$  for each  $x \in V(H)$ , and then one subgraph  $F_{xy}^{i(x), j(y), k(xy)}$  for each  $xy \in E(H)$ . More formally, given  $i(x)$  for  $x \in V(H)$  and  $k(xy)$  for  $xy \in E(H)$ ,

$$P = \bigcup_{xy} F_{xy}^{i(x), j(y), k(xy)}$$

is the corresponding polyad.

Observe that there is only one polyad consistent with the original partition  $\Pi_0$ , which is  $F$  itself, and, more generally, there are at most  $s = t^h l^{|E(H)|}$  polyads consistent with a  $(t, l)$ -partition (with equality if every tube is partitioned into exactly  $l$  parts). Let  $\mathcal{P} = \mathcal{P}_\Pi$  be the set of all polyads consistent with a  $(t, l)$ -partition  $\Pi$ . Note also that every copy of  $H$  belongs to exactly one polyad. Consequently, recalling that  $\kappa = |\mathcal{C}(F)|$  is the total number of copies of  $H$  in  $F$ , we have

$$\kappa = \sum_{P \in \mathcal{P}} c_P. \tag{14}$$

Finally, notice that if  $\Psi$  is a refinement of  $\Pi$  then for every polyad  $R \in \mathcal{P}_\Psi$  there exists a unique polyad  $P \in \mathcal{P}_\Pi$  such that  $R \subseteq P$ .

#### 4.2.5 Regularity

Polyads are the substructures of this construction which play the same role as  $\epsilon$ -regular pairs in a Szemerédi partition of a graph.

**Definition 4.10** Let  $\delta > 0$ . Given an  $H$ -graph  $F$  and a family  $\mathcal{S}$  of special copies of  $H$  in  $F$ , a polyad  $P$  is said to be  $\delta$ -regular, if for every subgraph  $Q \subseteq P$  with  $c_Q \geq \delta c_P$ , we have

$$|d_P - d_Q| \leq \delta.$$

In other words,  $\delta$ -regularity of  $P$  reflects low discrepancy between the  $(\mathcal{S}, p^*)$ -densities of those subpolyads of  $F$  which are relatively rich in ordinary, consistent copies of  $H$ . A polyad which is not  $\delta$ -regular will be called  $\delta$ -irregular, or just *irregular*. The set of all  $\delta$ -irregular polyads of  $\mathcal{P}_\Pi$  will be denoted by  $\mathcal{P}_\Pi^{\text{irr}}$  or just  $\mathcal{P}^{\text{irr}}$ .

We are now two notions away from our regularity lemma. The first of them ensures  $(\varepsilon, p)$ -regularity of most subgraphs constituting a  $(t, l)$ -partition.

**Definition 4.11** Let  $\varepsilon > 0$  and  $0 < p \leq 1$ . A  $(t, l)$ -partition  $\Pi$  is called  $(\varepsilon, p)$ -uniform if all but  $\varepsilon|F|$  edges of  $F$  are in  $(\varepsilon, p)$ -regular subgraphs  $F_{xy}^{i,j,k}$ .

The next definition specifies what kind of regularity we impose on partitions in terms of  $\delta$ -regular polyads. Roughly, the irregular polyads together cannot contain too many copies of  $H$ . Recall that  $\kappa = |\mathcal{C}(F)|$ .

**Definition 4.12** Given an  $H$ -graph  $F$  and a family  $\mathcal{S}$  of special copies of  $H$  in  $F$ , a  $(t, l)$ -partition  $\Pi$  is  $\delta$ -regular if

$$\sum_{P \in \mathcal{P}_\Pi^{\text{irr}}} c_P \leq \delta \kappa.$$

We now come to the central lemma of this section, which, when applied with  $H = S$ , the star forest from Section 3, puts into action the mechanism that makes the whole proof tick. It describes, in the spirit of the Szemerédi Regularity Lemma, a decomposition of an  $H$ -graph  $F$  in which the special copies of  $H$  are evenly distributed, i.e. it guarantees the existence of a  $\delta$ -regular partition. In addition, at the same time the partition that is obtained is  $(\varepsilon, p)$ -uniform. This extra feature is very useful in our application.

**Lemma 4.13 (Subgraph Regularity Lemma)** *For all graphs  $H$ , constants  $\delta > 0$ ,  $D > 0$ ,  $D^* \geq 1$ , and for all functions  $\varepsilon(\ell) > 0$ , there exist integers  $T_0, L_0, n_0$  such that*

1. for every  $n = hm > n_0$ ,  $0 < p = p(n) < 1$ , a graph  $F \in \mathcal{F}(H, m)$  with  $|V(F)| = n$  such that
  - (a)  $F$  is  $H$ -uniform
  - (b)  $F$  has no  $(D, p)$ -dense patches, and

2. for every  $0 < p^* = p^*(n) < 1$  and  $\mathcal{S} \subseteq \mathcal{C}(F)$  such that  $F$  (together with  $\mathcal{S}$ ) has no  $(D^*, p^*)$ -dense  $H$ -patches,

there exists a refinement  $\Pi_1$  of the initial partition  $\Pi_0$  which is a  $\delta$ -regular,  $(\varepsilon(l), p)$ -uniform,  $(t, l)$ -partition for some  $l \leq L_0$  and  $t \leq T_0$ .

**Remark 4.2** In both classic cases,  $H = K_2$  and  $H = K_3$ , all copies of  $H$  in  $H^m$  are consistent, and consequently, the corresponding regularity lemmas are true for arbitrary graphs  $F$ , not just for bipartite or tripartite graphs  $F \in \mathcal{F}(H, m)$ . However, for general  $H$  the “partiteness” seems necessary. Consider, e.g.,  $H$  being a path on vertices 1, 2, 3, and let  $F = K_{3m}$  be a complete graph on vertex set  $V = V_1 \cup V_2 \cup V_3$ . Further, let  $c_P$  count all copies of  $H$ , while  $s_P$  – only those consistent with the partition  $(V_1, V_2, V_3)$ . Then no regularity lemma may guarantee a  $\delta$ -regular partition of  $F$ . Indeed, every polyad  $P$  with large  $s_P$  must contain a bipartite subgraph  $R \subset F[V_1, V_2]$  with still large  $c_R$  but, clearly, with  $s_R = 0$ , making  $P$   $\delta$ -irregular.

**Remark 4.3** Our lemma admits edge partitions besides the more traditional vertex partitions. Consequently, non-induced subgraphs are considered. It is then closer in spirit to the hypergraph regularity lemma from [12], than the classical Szemerédi lemma, where pairs of vertex partition classes determine induced subgraphs.

**Remark 4.4** If both  $p$  and  $p^*$  are constants, say 1, then the assumptions about the absence of dense patches are vacuously satisfied with  $D = D^* = 1$ . Indeed, e.g., by (10) we have then  $d_R = s_R/c_R \leq 1$  and no  $(1, 1)$ -dense  $H$ -patches exist (see Definition 4.6). In the sparse case, however, the assumption of no  $(D^*, p^*)$ -dense  $H$ -patches, which plays a key role in the proof, can be very hard to verify. For the application in this paper, i.e. for  $H$  being a star forest  $S$  with  $\phi$  edges, and  $p^* = p^\phi$ , where  $p = p(n)$  is as in the Setup on page 16, this is done in Section 4.4 with considerable technical effort.

**Remark 4.5** Using the relevant definitions, it is easy to see that the condition of not having any  $(D, p)$ -dense patches bounds  $p$  from below:  $p \geq 1/(n^2 D)$ ; similarly, as observed above (see (11)), not having  $(D^*, p^*)$ -dense  $H$ -patches yields  $p^* = \Omega(n^{-h})$ .

**Remark 4.6** We will use this lemma with  $\varepsilon(\ell) \leq \ell^{-4\phi}$ . The reason for having  $\varepsilon$  decrease with  $\ell$  so quickly is that we will need  $\varepsilon$  to be much smaller than  $1/\ell$ , a lower bound on the average  $p$ -density of the resulting subgraphs (see proof of Lemma 5.4).

### 4.3 Proof of the Subgraph Regularity Lemma

The proof is based on a technique from the original Szemerédi Regularity Lemma: refining the initial partition  $\Pi_0$  until it becomes  $\delta$ -regular. This procedure will be monitored by the *index* of a partition. We will see that as long as a current partition is not  $\delta$ -regular, it can be refined, resulting in a partition with a substantially larger index (this is Lemma 4.15). A priori bounds on the index given in Lemma 4.16 will guarantee that this process terminates successfully after a bounded number of steps. Throughout, the  $(\varepsilon, p)$ -uniformity of the partitions will be maintained as well.

#### 4.3.1 Pumping the Index

Let

$$\sigma_R = \frac{c_R}{\kappa} \tag{15}$$

be the *relative volume* of a subgraph  $R$  of  $F$ . It measures what percentage of the copies of  $H$  which can be found in  $F$  are contained in  $R$ . Note that in view of (14),

$$\sum_{P \in \mathcal{P}_\Pi^{\text{irr}}} \sigma_P = 1. \tag{16}$$

Let

$$\text{Index}(\Pi) = \sum_{P \in \mathcal{P}_\Pi} \sigma_P d_P \log d_P$$

be the *index of a partition*  $\Pi$ . Thus, the index is a weighted sum of a convex function  $(x \log x)$  of the  $(\mathcal{S}, p^*)$ -densities of the polyads consistent with  $\Pi$ . The merit of this convex function is that whenever a  $(t, l)$ -partition is *not*  $\delta$ -regular it is possible to refine it into a  $(t', l')$ -partition with a substantially larger index.

To start with, let  $\Psi$  be *any* refinement of a current partition  $\Pi$  as defined in Definition 4.8, and, for a given polyad  $P \in \mathcal{P}_\Pi$ , let  $\Lambda(P) = \Lambda(\Psi, P)$  be the

set of all polyads  $R$  consistent with  $\Psi$  and such that  $R \subseteq P$ . Clearly,

$$\mathcal{P}_\Psi = \bigcup_{P \in \mathcal{P}_\Pi} \Lambda(P).$$

Setting

$$\mu_R = \mu_R(P) = \frac{c_R}{c_P} \quad \left( = \frac{\sigma_R}{\sigma_P} \right),$$

and using the identity  $\sigma_R = \sigma_P \mu_R$  we have

$$\text{Index}(\Psi) = \sum_{P \in \mathcal{P}_\Pi} \sum_{R \in \Lambda(P)} \sigma_R d_R \log d_R = \sum_P \sigma_P \sum_{R \in \Lambda(P)} \mu_R d_R \log d_R.$$

Observe that  $d_P = \sum_{R \in \Lambda(P)} \mu_R d_R$  and  $\sum \mu_R = 1$ . Hence, by Jensen's inequality, for each  $P \in \mathcal{P}_\Pi$ ,

$$\sum_{R \in \Lambda(P)} \mu_R d_R \log d_R \geq \left( \sum_R \mu_R d_R \right) \log \left( \sum_R \mu_R d_R \right) = d_P \log d_P,$$

and consequently  $\text{Index}(\Psi) \geq \text{Index}(\Pi)$ .

The above application of Jensen's inequality can be sharpened significantly in the case when  $P$  is a  $\delta$ -irregular polyad, provided we impose an extra constraint on partition  $\Psi$ . Given a subgraph  $Q$  of  $F$ , we say that a partition  $\Psi$  *refines*  $Q$  if its edge partition (13) refines the partition  $(Q, F-Q)$ , that is, every subgraph from (13) is either contained in or disjoint from  $Q$ .

**Lemma 4.14 (Defect Lemma)** *Let  $P \in \mathcal{P}_\Pi^{\text{irr}}$  be a  $\delta$ -irregular polyad and  $Q \subseteq P$  be such that*

$$(i) \quad c_Q \geq \delta c_P \text{ and}$$

$$(ii) \quad |d_P - d_Q| > \delta.$$

*If  $\Psi$  is a refinement of  $\Pi$  which refines  $Q$  then*

$$\sum_{R \in \Lambda(P)} \mu_R d_R \log d_R \geq d_P \log d_P + \frac{2\delta^4}{d_P}.$$

For the sake of readability, we defer for a moment the technical proof of Lemma 4.14. An immediate consequence of Lemma 4.14 is a jump in the value of the index when a partition has many  $\delta$ -irregular polyads and its refinement is properly chosen. A subgraph  $Q$  which satisfies conditions (i) and (ii) of Lemma 4.14 with respect to a given  $\delta$ -irregular polyad  $P$  will be called a *witness of irregularity for  $P$* .

**Lemma 4.15 (Index Pumping Lemma)** *If  $\Pi$  is a  $\delta$ -irregular  $(t, l)$ -partition of  $F$  and  $\Psi$  is a refinement of  $\Pi$  which also refines at least one witness of irregularity for each polyad  $P \in \mathcal{P}^{\text{irr}}$ , then*

$$\text{Index}(\Psi) \geq \text{Index}(\Pi) + \frac{\delta^5}{D^*}.$$

**Proof:** As before, but using Lemma 4.14 instead of Jensen's inequality,

$$\begin{aligned} \text{Index}(\Psi) &= \sum_P \sigma_P \sum_{R \in \Lambda(P)} \mu_R d_R \log d_R \\ &\geq \sum_P \sigma_P d_P \log d_P + \sum_{P \in \mathcal{P}^{\text{irr}}} \sigma_P \frac{2\delta^4}{d_P} \\ &\geq \text{Index}(\Pi) + \sum_{P \in \mathcal{P}^{\text{irr}}, c_P \geq \kappa / \log^2 n} \sigma_P \frac{2\delta^4}{D^*}, \end{aligned}$$

For the last inequality above we used the assumption that there are no  $(D^*, p^*)$ -dense  $H$ -patches in  $F$ , and thus, if  $c_P \geq \kappa / \log^2 n$  then  $d_P \leq D^*$ . By the  $\delta$ -irregularity of  $\Pi$  (see Definition 4.12) and by (15), we have

$$\sum_{P \in \mathcal{P}^{\text{irr}}} \sigma_P > \delta.$$

Moreover, for sufficiently large  $n$ , recalling that  $|\mathcal{P}| \leq s = t^h l^{|E(H)|}$ ,

$$\sum_{P: c_P < \kappa / \log^2 n} \sigma_P < \sum_P \log^{-2} n \leq s \log^{-2} n < \frac{\delta}{2},$$

because  $t, l$  and  $H$  do not depend on  $n$ . Hence, in view of (16),

$$\sum_{P \in \mathcal{P}^{\text{irr}}, c_P \geq \kappa / \log^2 n} \sigma_P \frac{2\delta^4}{D^*} \geq \frac{\delta^5}{D^*}.$$

□

The Index Pumping Lemma works hand in hand with the following estimates.

**Lemma 4.16** *For all partitions  $\Pi$  of  $F$ ,*

$$-\frac{1}{e} \leq \text{Index}(\Pi) \leq D^* \log D^* + o(1).$$

**Proof:** The lower bound follows easily from (16) and the fact that the function  $x \log x$  achieves its minimum at  $x = 1/e$ .

For the upper bound, we split

$$\text{Index}(\Pi) = \sum_{P:c_P \geq \kappa/\log^2 n} \sigma_P d_P \log d_P + \sum_{P:c_P < \kappa/\log^2 n} \sigma_P d_P \log d_P.$$

Because  $F$  has no  $(D^*, p^*)$ -dense  $H$ -patches, we can bound the first term by  $D^* \log D^*$ . As for the second term, by Proposition 4.12,  $c_P < \frac{\kappa}{\log^2 n}$  implies that  $s_P \leq \frac{(1+o(1))D^* p^* \kappa}{\log^2 n}$ . Hence,

$$\sigma_P d_P = s_P / (\kappa p^*) \leq (1 + o(1)) D^* / (\log^2 n).$$

Moreover, for all polyads  $P$  we have  $\log d_P = O(\log n)$ , because by (11) (or see Remark 4.5)

$$d_P = \frac{s_P}{p^* c_P} \leq \frac{1}{p^*} = O(n^h).$$

Hence,

$$\sum_{P:c_P < \kappa/\log^2 n} \sigma_P d_P \log d_P \leq s \frac{D^*}{\log^2 n} O(\log n) = o(1),$$

also using the fact that the number of polyads is at most  $s = t^h l^{|E(H)|} = O(1)$ .  
□

**Remark 4.7** Note that we would not have been able to bound the second term above as easily, had we used the more traditional functional,  $\sum_P d_P^2$ , in the definition of index. The index function was introduced by Szemerédi in [Sz] as the main tool in the proof of his regularity lemma (Theorem 4.5 above), and later used also in [FR] to prove the hypergraph regularity lemma

(Theorem 4.11 above). The idea of replacing  $d^2$  by any convex function in this context was suggested by Ajtai, Komlos, Szemerédi [E. Szemerédi, oral communication, 1978]. Such an alternative index function first appeared in print in [26].

The above two lemmas basically complete the proof of Lemma 4.13, provided we can construct a required refinement which in addition is  $(\varepsilon(l), p)$ -uniform. Then, iterating the process of refining a current partition, we must reach a  $\delta$ -regular  $(t, l)$ -partition. The bounds in these lemmas imply an upper bound on the number of iterations, solely in terms of  $D^*$  and  $\delta$ .

In order to conclude the existence of constants  $T_0$  and  $L_0$  such that  $t \leq T_0$  and  $l \leq L_0$ , we need to show how the refinement is created. Likewise, although the asymptotic estimates we encountered so far contribute toward the value of  $n_0$ , the most demanding constraint on  $n_0$  comes from the refinement which invoke Theorem 4.10. A description of how the refinement is constructed is deferred to the next subsection. We complete this subsection by giving a proof of the Defect Lemma.

**Proof of Lemma 4.14:** Let  $\Lambda(Q)$  be the set of all polyads  $R$  consistent with  $\Psi$  and such that  $R \subseteq Q$ . Recall that  $\mu_R = c_R/c_P$  and set

$$\mu = \sum_{R \in \Lambda(Q)} \mu_R.$$

Since  $\Psi$  refines  $Q$ , we have  $\sum_{R \in \Lambda(Q)} c_R = c_Q$  and  $\sum_{R \in \Lambda(Q)} s_R = s_Q$ . Thus, by (i),  $\mu = c_Q/c_P \geq \delta > 0$  and

$$d_Q = \frac{1}{\mu} \sum_{R \in \Lambda(Q)} \mu_R d_R.$$

Moreover,  $\mu = 1$  then  $c_Q = c_P$ ,  $s_Q = s_P$ , and consequently,  $d_P = d_Q$ , contradicting (ii). Hence, we may assume that  $\mu < 1$ , and by Jensen's inequality,

$$\sum_{R \in \Lambda(Q)} \mu_R d_R \log d_R \geq \mu \cdot d_Q \log d_Q$$

and

$$\sum_{R \in \Lambda(P) \setminus \Lambda(Q)} \mu_R d_R \log d_R \geq (d_P - \mu d_Q) \log \left( \frac{d_P - \mu d_Q}{1 - \mu} \right).$$

So, setting  $d = d_P$ ,  $d' = d_Q$ , and  $d'' = \frac{d - \mu d'}{1 - \mu}$  for convenience,

$$\sum_{R \in \Lambda(P)} \mu_R d_R \log d_R \geq \mu d' \log d' + (1 - \mu) d'' \log d''. \quad (17)$$

Consider the function

$$f(x) = x \log \left( \frac{x}{\mu} \right) + (d - x) \log \left( \frac{d - x}{1 - \mu} \right).$$

Clearly,  $f(\mu d) = d \log d$ , and further, since  $\mu d' + (1 - \mu) d'' = d$ ,

$$\begin{aligned} f(\mu d') &= \mu d' \log d' + (d - \mu d') \log \left( \frac{d - \mu d'}{1 - \mu} \right) \\ &= \mu d' \log d' + (1 - \mu) d'' \log d''. \end{aligned} \quad (18)$$

Thus in view of (17),

$$\sum_{R \in \Lambda(P)} \mu_R d_R \log d_R - d \log d \geq f(\mu d') - f(\mu d),$$

and we will be done if

$$f(\mu d') - f(\mu d) \geq \frac{\delta^4}{2d}. \quad (19)$$

Towards this, compute

$$f'(x) = \log \left( \frac{x}{\mu} \right) - \log \left( \frac{d - x}{1 - \mu} \right),$$

and

$$f''(x) = \frac{1}{x} + \frac{1}{d - x},$$

and note that  $f'(\mu d) = 0$ , and that  $f''(x) \geq 4/d$ , for all  $0 < x < d$ . Thus, by the Lagrange formula for the remainder, we have, for some  $0 < \mu d < x_0 < \mu d' < d$ ,

$$f(\mu d') - f(\mu d) = f''(x_0) \frac{(\mu d' - \mu d)^2}{2} \geq \frac{4}{d} \frac{\delta^4}{2} = \frac{2\delta^4}{d},$$

since  $\mu \geq \delta$  and  $|d' - d| > \delta$ . This proves (19), completing the proof of Lemma 4.14.  $\square$

### 4.3.2 Refining a Partition

We will now finish the proof of Lemma 4.13. As already observed prior to the proof of Lemma 4.14, it remains to construct a  $(t', l')$ -partition  $\Psi$  with the following properties. Recall that a subgraph  $Q$  which satisfies conditions (i) and (ii) of Lemma 4.14 with respect to a given  $\delta$ -irregular polyad  $P$  is called a *witness of irregularity for  $P$* . The new partition  $\Psi$  must

1. be a refinement of a current  $\delta$ -irregular  $(t, l)$ -partition  $\Pi$ ,
2. refine at least one witness of irregularity for each irregular polyad of  $\Pi$ ,
3. be  $(\varepsilon(l'), p)$ -uniform.

Recall that a partition  $\Pi$  gives rise to at most  $s = t^h l^{|E(H)|}$  different polyads. Suppose that polyads  $P_1, \dots, P_{s'}$  are  $\delta$ -irregular with witnesses of their irregularity  $Q_1, \dots, Q_{s'}$ . The requirement that the new partition must refine each  $Q_g$ ,  $g = 1, \dots, s'$ , is easy to fulfill. Let  $\text{Venn}\{Q_1, \dots, Q_{s'}\} = \{R_1, \dots, R_{s''}\}$ ,  $s'' \leq 2^{s'}$ , be the family of all graphs of the form  $\bigcap_{g=1}^{s'} Q_i^{\epsilon_g}$ , for  $\epsilon_g \in \{0, 1\}$ , where

$$Q_i^{\epsilon_g} = \begin{cases} Q_g & \epsilon_g = 1 \\ F - Q_g & \epsilon_g = 0. \end{cases}$$

The graphs  $R_g$ ,  $g = 1 \dots s''$ , partition (by intersection) every subgraph  $F_{xy}^{i,j,k}$  into smaller subgraphs  $F_{xy}^{i,j,k,g} = F_{xy}^{i,j,k} \cap R_g$ , that is,

$$F_{xy}^{i,j,k} = \bigcup_{g=1}^{s''} F_{xy}^{i,j,k,g}.$$

This new partition  $F_{xy}^{i,j,k,g}$ , clearly, refines every subgraph  $Q_g$ ,  $g = 1, \dots, s'$ . Note also that the previous partition of every tube into at most  $l$  subgraphs is now refined into at most  $l'$  finer subgraphs, with  $l' = ls'' \leq l2^{s'} \leq l2^s$ .

So far we have only partitioned *the edge sets* of  $\Pi$ . With the help of Theorem 4.10 (the Sparse Regularity Lemma) we will now refine the vertex equipartition of  $\Pi$  to ensure that the new equipartition is  $(\varepsilon(l'), p)$ -uniform (and still refines every  $Q_g$ ). By the assumption on  $F$ , none of the resulting graphs  $F_{xy}^{i,j,k,g}$  has  $(D, p)$ -dense patches ( $F_{xy}^{i,j,k,g}$  is a subgraph of  $F$ ). Thus, we

may apply Theorem 4.10 with  $\varepsilon = \varepsilon(t')$ ,  $D = 2$ ,  $t_0 = t$  and  $r = |E(H)|t^2l2^s$ , to the graphs  $\{F_{xy}^{i,j,k,g}\}$  with the vertex equipartition given by  $\Pi$ , where  $xy \in E(H)$ ,  $i, j = 1, \dots, t$ ,  $k = 1, \dots, l(xy, i, j) \leq l$  and  $g = 1, \dots, s''$ . This way we obtain a finer equipartition of the vertices which together with the described above partition of the subgraphs, defines an  $(\varepsilon(t'), p)$ -uniform  $(t', l')$ -partition  $\Psi$ . This new partition consists of the following parts:

- Vertex sets defined by the equipartition given by Theorem 4.10:

$$\{\bar{V}_x^i, i \in [t']\} \prec \{V_x^i, i \in [t]\},$$

where  $\prec$  denotes refinement.

- Edge sets

$$\bar{F}_{xy}^{i,j,k,g} = F_{xy}^{i,j,k,g}[\bar{V}_x^i, \bar{V}_y^j].$$

The new parameter  $t'$  is bounded by  $T_0$  of Theorem 4.10 with the above inputs, and so depends on  $l$ ,  $t$ ,  $H$  and the function  $\varepsilon(\ell)$ . Starting from  $t = l = 1$  and iterating for a bounded number of times, as determined in the previous subsection, yields the promised constants  $T_0$  and  $L_0$  of Lemma 4.13. We do not attempt to compute any of these constants explicitly. This completes the proof of Lemma 4.13.

## 4.4 No Dense Patches

In the previous section we proved Lemma 4.13. Now we are going to verify that it can be applied in the setting relevant to the proof of Lemma 2.4. Let us summarize again the state-of-art at the end of Section 3. As a consequence of assumption (3) of Theorem 2.1 we have selected a star forest  $S$  consisting of  $\nu$  stars with, respectively,  $\phi_1, \dots, \phi_\nu$  arms. We set

$$|E(S)| = \phi = \phi_1 + \dots + \phi_\nu$$

and so  $|V(S)| = \nu + \phi$ . For clarity, we will be further assuming that  $n$  is divisible by  $\nu + \phi$ . Under this assumption, Lemma 3.3 provided us with a partition of the vertices of  $G$  into sets of equal size  $m = n/(\nu + \phi)$ , further denoted

$$\pi_0 = \{V_1, \dots, V_\nu, W_{11}, \dots, W_{1\phi_1}, \dots, W_{\nu,1}, \dots, W_{\nu,\phi_\nu}\}.$$

Based on this partition, we have focused on the set  $\mathcal{C}$  of all constellations, that is all copies of  $S$  in  $G$  that are consistent with  $\pi_0$ . Among the constellations we have distinguished the set  $\mathcal{S}$  of special constellations, related to the special role of certain additions of a copy of  $M$  to  $G$ .

Recall that  $M$  is an arbitrary (balanced) graph with  $\rho(M) = 2$  with vertices labeled  $x_1, \dots, x_\nu$ , and that  $\widehat{M}$  is a graph built upon  $M$  and consisting of a number of tepees spreading above the edges of  $M$ . It is  $\widehat{M}$  which gave rise to the star forest  $S$  by deleting one leg of each tepee (see The Missing Leg Property – Observation 3.4). The vertices of  $S$  not belonging to  $M$  (i.e. the tips of tepees) are denoted by  $u_{ab}$ , where  $a = 1, \dots, \nu$  and  $b = 1, \dots, \phi_a$ , in such a way that  $u_{ab}$  are the neighbors of  $x_a$ . The graph  $G$  is an arbitrary member of family  $\mathcal{G}$  defined in Definition 6.1. For all  $a = 1, \dots, \nu$  and  $b = 1, \dots, \phi_a$ , set  $F_{ab} = G[V_a, W_{ab}]$  and

$$F = \bigcup_{ab} F_{ab}.$$

A copy of  $S$  in  $F$  is consistent if each  $x_a$  is mapped onto a vertex of  $V_a$  and each  $u_{ab}$  is mapped onto a vertex of  $W_{ab}$ .

Note that  $F \in \mathcal{F}(S, m)$ . We want to apply the Subgraph Regularity Lemma 4.13, with  $H = S$ , to the graph  $F$ , family  $\mathcal{S}$ , and sequences:  $p = p(n)$  as in the Setup preceding Lemma 2.4 and  $p^* = p^\phi$ . This normalizing factor  $p^*$  is to be expected because there are  $\phi$  edges that must appear in  $G$  to turn a constellation into a special constellation (the missing legs of the tepees), and informally (if  $G$  were random), each edge would appear with probability  $p$ .

In order to verify the assumptions of Lemma 4.13 for  $F$ , we have to show that for some  $D > 0$  and  $D^* \geq 1$  (independent of  $G$ , and thus independent of  $S$  and  $\phi$  in particular),

1(a):  $F$  has no  $(D, p)$ -dense patches

1(b):  $F$  is  $S$ -uniform

2:  $F$ , together with given  $\mathcal{S}$ , has no  $(D^*, p^\phi)$ -dense  $S$ -patches.

All these properties of  $F$  will be derived from Definition 6.1 of family  $\mathcal{G}$  to which  $G$ , a supergraph of  $F$ , belongs. Property 1(a) follows immediately with  $D = 2$  from Property (P5) of  $\mathcal{G}$ , while property 1(b) is an easy consequence

of Property (P3), that is, of the bounds on the vertex degrees in  $G$ . Indeed, by Example 4.9, there are

$$\kappa \sim m^{\nu+\phi} p^\phi$$

(consistent) copies of  $S$  in  $F$ , while a given edge belongs to at most

$$n^{\nu-1} (2pn)^{\phi-1} = O\left(\frac{\kappa}{n^{3/2}}\right) = o\left(\frac{\kappa}{\log^3 n}\right) \quad (20)$$

of them, for large  $n$ . The next lemma establishes assumption 2 with  $D^* = 2^q + 1$ , where, recall,  $q = \lceil 10C^2\nu/\alpha \rceil$  is an upper bound on  $\phi$  (see Lemma 3.1). Recall also that  $c_R$  is the number of constellations (i.e., consistent with  $\pi_0$  copies of  $S$ ) contained in a subgraph  $R$ .

**Lemma 4.17** *For every subgraph  $R \subseteq F$  with  $c_R \geq \kappa/\log^2 n$  we have  $d_R \leq 2^q + 1$ .*

Let us set  $U_a = V(R) \cap V_a$ ,  $R_{ab} = R \cap F_{ab}$  and  $\deg_{ab}(v) = \deg_{R_{ab}}(v)$  for all  $a = 1, \dots, \nu$ ,  $b = 1, \dots, \phi_a$ . It is straightforward to express  $c_R$  in terms of the degrees  $\deg_{ab}(v)$ . Indeed, we have

$$c_R = |U_\nu| \prod_{a=1}^{\nu-1} \left( \sum_{v \in U_a} \prod_{b=1}^{\phi_a} \deg_{ab}(v) \right). \quad (21)$$

This simple formula plays a key role in our proof. First, it trivially implies that

$$c_R \geq \prod_{a=1}^{\nu} \left( |U_a| \prod_{b=1}^{\phi_a} \delta_{ab} \right), \quad (22)$$

where  $\delta_{ab} = \min_v \deg_{ab}(v)$ . Second, in connection with the estimate of  $\kappa$ , (21) yields lower bounds on  $|U_a|$  and  $\Delta_{ab} = \max_v \deg_{ab}(v)$  in terms of  $c_R$ . We will use them later in the proof.

**Proposition 4.18** *For all  $1 \leq a \leq \nu$ ,*

- (a)  $|U_a| = \Omega\left(\frac{c_R}{\kappa} n\right)$ , and
- (b)  $\Delta_{ab} = \Omega\left(\frac{c_R}{\kappa} np\right)$  for each  $b = 1, \dots, \phi_a$ .

**Proof.** Using Property (P3) of  $G$ , we can further estimate (21) as follows. For each  $1 \leq a \leq \nu$  and  $1 \leq b \leq \phi_a$  we have

$$\begin{aligned} c_R &= |U_\nu| \prod_{a'=1}^{\nu-1} \left( \sum_{v \in U_{a'}} \prod_{b'=1}^{\phi_{a'}} \deg_{a'b'}(v) \right) \\ &\leq \begin{cases} m \prod_{a'=1}^{\nu-1} \left( \sum_{v \in U_{a'}} (2np)^{\phi_{a'}} \right) = O(|U_a| n^{\nu-1} (np)^\phi) \\ m^\nu \prod_{a'=1}^{\nu-1} \left( \prod_{b'=1}^{\phi_{a'}} \Delta_{a'b'} \right) = O(n^\nu (np)^{\phi-1} \Delta_{ab}). \end{cases} \end{aligned}$$

By Example 4.9, we have  $\kappa \sim m^{\nu+\phi} p^\phi$ , which completes the proof.  $\square$

In view of (22) to prove Lemma 4.17 we need to prove a matching upper bound on  $s_R$ . However, our current proof techniques allow us only to set a bound of the form

$$s_R \leq (1 + o(1)) p^\phi \prod_{a=1}^{\nu} \left( |U_a| \prod_{b=1}^{\phi_a} \Delta_{ab} \right), \quad (23)$$

(see later in this section). So, we can only obtain the desired bound on  $d_R$  if  $\delta_{ab}$  and  $\Delta_{ab}$  are at most a factor of two apart. This leads to the following definition. A subgraph  $R$  of  $F$  will be called *degree sandwiched* if there exist numbers  $r_{ab}$ ,  $a = 1, \dots, \nu$ ,  $b = 1, \dots, \phi_a$ , such that for all  $v \in U_a$ , we have  $r_{ab} \leq \deg_{ab}(v) \leq 2r_{ab}$ .

To put ourselves in a more comfortable position, we will cover  $R$  with its degree sandwiched subgraphs in such a way that each constellation in  $R$  is contained in exactly one of them. Remembering that  $\phi \leq q$ , the following lemma will imply Lemma 4.17.

**Lemma 4.19** *For every degree sandwiched subgraph  $R \subseteq F$  with  $c_R \geq \kappa / \log^{\phi+3} n$  we have  $d_R \leq 2^\phi + o(1)$ .*

In order to see how Lemma 4.19 implies Lemma 4.17 we need one more fact.

**Proposition 4.20** *For every degree sandwiched subgraph  $R$  with  $c_R < \kappa / \log^{\phi+3} n$  there exists a degree sandwiched subgraph  $R'$  such that  $R \subset R' \subseteq F$  and*

$$\kappa / \log^{\phi+3} n \leq c_{R'_i} = (1 + o(1)) \kappa / \log^{\phi+3} n.$$

**Proof.** We will construct a sequence of subgraphs  $R = R^1 \subset R^2 \subset \dots \subset F$  such that each  $R^j$  is degree sandwiched and is obtained from  $R^{j-1}$  by adding at most  $2np = O(\sqrt{n})$  edges. Note that by Lemma 3.3,  $F$  itself is degree sandwiched between  $\frac{3}{4}mp$  and  $\frac{3}{2}mp$ . By (20), the increments in  $c_{R^j}$  are at most  $O(\kappa/n) = o(k/\log^{\phi+3} n)$ , and so there must be some  $j$  such that  $R^j = F$  is the desired subgraph. (Note that it is certain that such an index  $j$  will be found many steps before reaching  $F$ . Indeed, by (22), we will be done when, for example, all  $U_a$  will be of order  $\Omega(n)$  and all degrees of order at least  $np/\log n$ .)

Our construction is performed in two phases. In phase one, for each  $a$  and  $b$ , we turn  $R \cap F_{ab}$  into an induced subgraph of  $F_{ab}$ , and then in phase two we will add one by one all remaining vertices of  $\bigcup_a V_a$ . A single step of phase one consists of fixing a vertex  $v$  in  $U_a$  of minimum degree  $r$  in  $R^{j-1} \cap F_{ab}$  and adding some  $\min(r, \deg_{F_{ab}}(v) - r)$  edges  $vu \in F_{ab} \setminus R^{j-1}$ . This way the process maintains the “degree-sandwichness” throughout, and is bound to arrive at  $R^j = F_{ab}[U_a]$  for some  $j$ . A single step of phase two consists of adding a vertex  $v \in \bigcup_a V_a \setminus V(R^j)$  together with all its neighbors in  $F$ .  $\square$

**Proof of Lemma 4.17:** Let  $R \subseteq F$  with  $|C(R)| = c_R \geq \kappa/\log^2 n$ . Without loss of generality assume that for all  $a = 1, \dots, \nu$ ,  $b = 1, \dots, \phi_a$ , and all  $v \in U_a$ , we have  $\deg_R(v, W_{ab}) > 0$ , and define for each  $j = 0, \dots, \lfloor \log \Delta_{ab} \rfloor$

$$U_{ab}^j = \left\{ v \in U_a : \frac{\Delta_{ab}}{2^{j+1}} \leq \deg_{ab}(v) \leq \frac{\Delta_{ab}}{2^j} \right\}.$$

This subdivides every set  $U_a$  into  $O(\log^{\phi_a} n)$  classes

$$U_a^{j_1, \dots, j_{\phi_a}} = \bigcap_{b=1}^{\phi_a} U_{ab}^{j_b}$$

so that the degrees in  $R$  of all vertices from one class, along every subgraph  $F_{ab}$  are sandwiched between a number and its double. By selecting one partition class from each set  $V_a$  and taking all the edges of  $R$  with endpoints in these classes, this yields a covering of  $R$  by  $O(\log^{\phi} n)$  subgraphs  $R_1, R_2, \dots$  which are all degree sandwiched. Set  $d_i = d_{R_i}$ , etc., for convenience. Clearly, we have  $c_R = \sum_i c_i$ ,  $s_R = \sum_i s_i$ , and

$$d_R = \sum_i \frac{s_i}{p^{\phi} c_R} = \sum_{c_i \geq \kappa/\log^{\phi+3} n} d_i \frac{c_i}{c_R} + \sum_{c_i < \kappa/\log^{\phi+3} n} \frac{s_i}{p^{\phi} c_R}.$$

Applying Lemma 4.19 in the first sum, we bound it by  $2^\phi + o(1)$ . For the second sum, we use a similar idea as in the proof of Proposition 4.12. For each  $i$ , let  $R'_i$  be the degree sandwiched supergraph of  $R_i$  as in Proposition 4.20. Applying Lemma 4.19 to  $R'_i$ , we obtain

$$s_i \leq s_{R'_i} \leq (2^\phi + o(1))p^*c_{R'_i} \leq (1 + o(1))\frac{2^\phi p^* \kappa}{\log^{\phi+3} n}$$

and so

$$\sum_{c_i < \kappa / \log^{\phi+3} n} \frac{s_i}{p^\phi c_R} = O\left(\frac{\kappa \log^\phi n}{c_R \log^{\phi+3} n}\right) = O(\log^{-1} n) = o(1),$$

by the assumption that  $c_R \geq \kappa / \log^2 n$ . □

The rest of this section is devoted to proving Lemma 4.19. Let  $R$  be a degree sandwiched subgraph of  $F$  with  $c_R \geq \kappa / \log^{\phi+3} n$ . As for its vertex set, we have, for each  $a$  and  $b$ ,  $V(R) \supset W_{ab}$ . and  $R_{ab} = R \cap F_{ab}$ . The degrees in  $R$  are such that for all  $a$  and  $b$ , and all  $v \in U_a$

$$r_{ab} \leq \deg_{R_{ab}}(v) \leq 2r_{ab},$$

for some  $r_{ab} > 0$ , where recall  $\deg_{R_{ab}}(v) = \deg_R(v, W_{ab}) = \deg_{R_{ab}}(v)$ .

In view of (22), to prove Lemma 4.19, it suffices to show that

$$s_R \leq (1 + o(1))p^\phi 2^\phi \prod_{a=1}^\nu \left( |U_a| \prod_{b=1}^{\phi_a} r_{ab} \right). \quad (24)$$

The following corollary of Proposition 4.12 will be of some use.

**Corollary 4.21** *If  $c_R \geq \kappa / \log^{\phi+3} n$ , then for all  $1 \leq a \leq \nu$ ,*

$$(a) |U_a| \geq \frac{n}{\log^{\phi+4} n}, \text{ and}$$

$$(b) r_{ab} \geq \frac{np}{\log^{\phi+4} n} \text{ for each } b = 1, \dots, \phi_a.$$

□

Let us now briefly outline the rest of the proof of Lemma 4.19: we want to bound the number of special constellations in  $R$ , or more specifically we want to prove (24), where  $U_a = V(R) \cap V_a$ . Recall from Section 3 that every special constellation is of the form  $S(X)$  for a set of vertices  $X \in \mathcal{X}_3$  which contains one vertex from each set  $V_a$ . Such a constellation corresponds via the Missing Leg Property (Observation 3.4) to a set of  $\phi$  edges of  $G$  (the missing legs.) These edges, together with the edges of the constellation itself form a set of  $\phi$  tepees over the edges of  $M_X$ , a copy of  $M$  planted at  $X$  (see Figure 3).

We will construct a  $\nu$ -partite auxiliary graph  $A$  on the vertex sets  $U_1, \dots, U_\nu$  that actually contains these graphs  $M_X$  for which  $S(X) \subset R$ , but possibly also other copies of  $M$ . Hence, the number of consistent copies of  $M$  in  $A$  (such that  $x_a$  is mapped onto a vertex of  $U_a$ ) will be an upper bound on the number of special constellations in  $R$ .

More precisely, for every pair  $a, a'$  for which  $x_a x_{a'} \in E(M)$  define the set  $\Gamma_{aa'}$  of indices  $b$  such that  $x_a u_{ab} x_{a'}$  is a tepee in  $\widehat{M}$ . Formally, using notation  $a'(a, b)$  from Section 3,

$$\Gamma_{aa'} = \{b : a' = a'(a, b)\}. \quad (25)$$

We will draw an edge in  $A$  joining  $v \in U_a$  with  $v' \in U_{a'}$ , if for each  $b \in \Gamma_{aa'}$  the vertex  $v'$  is connected by an edge of  $G$  with a vertex of  $W_{ab}$  which is a neighbor of  $v$  in  $R$ . If each bipartite subgraph  $A_{aa'} = A[U_a, U_{a'}]$  was highly regular, then using Proposition 4.3 it would be easy to count the number of copies of  $M$  in  $A$ .

To achieve this high regularity of  $A_{aa'}$  we will consider its supergraph  $A'_{aa'}$  with the entire  $V(G)$  in place of  $U_{a'}$ . Then the degree of  $v \in U_a$  in  $A'_{aa'}$  will be asymptotically determined, via Property (P6) of  $G$ , by the sizes of its neighborhoods in  $W_{ab}$  for all  $b \in \Gamma_{aa'}$ . We know that for distinct  $v$  they are all sandwiched between  $r_{ab}$  and  $2r_{ab}$ , but may vary, obstructing the desired regularity. To remedy this problem we will enlarge these neighborhoods so that they are all equal  $2r_{ab}$ . Since we still want to keep the co-degrees fairly low (as they are by Property (P4) of  $G$ ), we will make these enlargements randomly. Only after this is done, we will formally define the auxiliary graph, prove that it is highly regular and finally count copies of  $M$  in it.

Fix  $a, b$  and set  $r = r_{ab}$ . Let  $k = |U_a|$ ,  $U_a = \{v_1, \dots, v_k\}$ , and  $N_i = N_i(b)$  be the neighborhood of  $v_i$  in  $R_{ab}$ . Note that  $r \leq |N_i| \leq 2r$  and, by Definition

6.1 (P4), for all  $i \neq j$ ,  $|N_i \cap N_j| \leq 3 \log n$ . Note also that by Corollary 4.21 we have, say,  $r \geq n^{0.49}$ . The following lemma gives us exactly what we need.

**Lemma 4.22** *For a sufficiently large integer  $n$ , let  $N_1, \dots, N_k$ ,  $k \leq n$ , be a set system of subsets of  $W$ , where  $\theta_1 n \leq |W| \leq n$ , for all  $i$ ,  $n^{0.49} \leq r \leq |N_i| \leq 2r \leq \theta_2 \sqrt{n}$ , and for all  $i \neq j$ ,  $|N_i \cap N_j| \leq 3 \log n$ . Then there exist sets  $\bar{N}_i = \bar{N}_i$ ,  $i = 1, \dots, k$ , such that for all  $i$*

- (a)  $N_i \subseteq \bar{N}_i$
- (b)  $|\bar{N}_i| = 2r$
- (c) for all  $i \neq j$ ,  $|\bar{N}_i \cap \bar{N}_j| \leq 12 \log n$ .

**Proof:** Pick  $\hat{N}_1, \dots, \hat{N}_k$  randomly, uniformly and independently from  $[W]^{2r}$ , the family of  $2r$ -element subsets of  $W$ . Then with probability  $1 - o(1)$ ,

$$|\hat{N}_i \cap \hat{N}_j| \leq 3 \log n, \quad \forall i < j \quad \text{and} \quad |\hat{N}_i \cap N_j| \leq 3 \log n \quad \forall i, j. \quad (26)$$

This is because for a fixed set  $N \subset W$ , and a random set  $\hat{N} \subset W$ ,  $|\hat{N}| = 2r$ , the random variable  $Z = |N \cap \hat{N}|$  has hypergeometric distribution with expectation  $|N| |\hat{N}| / |W| \leq 4\theta_2^2 / \theta_1$ . By Chernoff's inequality (e.g., (2.11) and Thm. 2.10 from [16]),

$$\Pr(Z > 3 \log n) < e^{-3 \log n} = n^{-3}.$$

As there are fewer than  $2n^2$  pairs of sets  $(\hat{N}_i, \hat{N}_j)$ ,  $i < j$ , and  $(\hat{N}_i, N_j)$ , the probability that at least one pair will violate (26) is less than  $\frac{2n^2}{n^3} \rightarrow 0$ . From (26), clearly,

$$|(N_i \cup \hat{N}_i) \cap (\hat{N}_j \cup N_j)| \leq 12 \log n.$$

Since  $3r - 3 \log n > 2r$  for large  $n$ , the required sets  $\bar{N}_1, \dots, \bar{N}_k$  can now be chosen arbitrarily in such a way that  $|\bar{N}_i| = 2r$  and that  $N_i \subseteq \bar{N}_i \subseteq N_i \cup \hat{N}_i$ . Then

$$\bar{N}_i \cap \bar{N}_j \subseteq (N_i \cup \hat{N}_i) \cap (N_j \cup \hat{N}_j),$$

and thus (c) holds as well.  $\square$

Let  $\hat{N}_i(b)$  be the extention of neighborhoods  $N_i(b)$  of vertices  $v_i \in U_a$  in  $W_{ab}$  whose existence is guaranteed by Lemma 4.22. We now define the auxiliary graphs. For every  $aa'$  such that  $x_a x_{a'} \in E(M)$  let  $A'_{aa'} = (U_a, V(G), E'_{aa'})$ , where for  $v_i \in U_a, v' \in V(G)$ :

$$v_i v' \in E'_{aa'} \Leftrightarrow |\bar{N}_i(b) \cap N_G(v')| \geq 1 \text{ for all } b \in \Gamma_{aa'}.$$

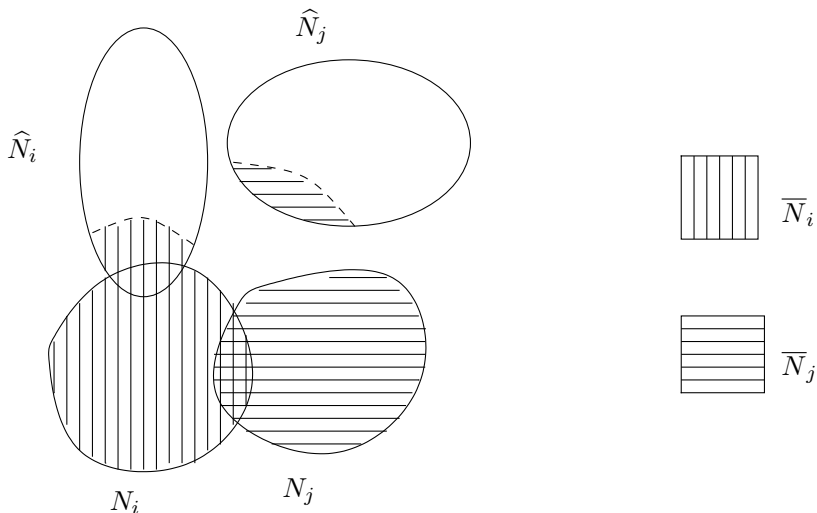


Figure 6: Enlarging  $N_i$ 's

**Remark 4.8** Since  $U_a \subseteq V(G)$ , the above definition involves a slight abuse of notation. Formally we are creating a graph between a copy of  $U_a$  and  $V(G)$ . We use  $V(G)$  and not  $V(G) \setminus U_a$  for technical reasons. However, we will momentarily consider the subgraph of  $A'_{aa'}$  induced between  $U_a$  and  $U'_a$ , so that this conflict will be resolved.

**Remark 4.9** In the case that  $\Gamma_{aa'} = \emptyset$  we take  $A'_{aa'}$  to be the complete bipartite graph, as the condition in the definition is fulfilled vacuously.

Next define the bipartite subgraph  $A_{aa'} \subset A'_{aa'}$  as the induced graph between the vertex sets  $U_a$  and  $U'_a$ , and finally let

$$A = \bigcup_{x_a x_{a'} \in E(M)} A_{aa'}.$$

See Figure 7 below for an illustration of  $A$ . Our definition of  $A_{aa'}$  guarantees that every special constellation in  $G$  yields a (not necessarily induced) subgraph of  $A$  isomorphic to  $M$ . More formally let

$$\#(M, A) = \text{the number of (consistent) copies of } M \text{ in } A.$$

**Proposition 4.23**

$$s_R \leq \#(M, A)$$

**Proof:** This follows immediately from the definition of  $A$  and the Missing Leg Property. Every special constellation  $S(X)$  has one vertex exactly in each of the sets  $U_a$ . The definition of  $A$  is designed so that a copy of  $M$  is placed on  $X$  if condition (a) of the Missing Leg Property holds for  $S(X)$ . Note also that every copy of  $M$  corresponds to at most one special constellation, else the corresponding copy of  $S(X)$  would violate condition (b) or (c) of the Missing Leg Property.  $\square$

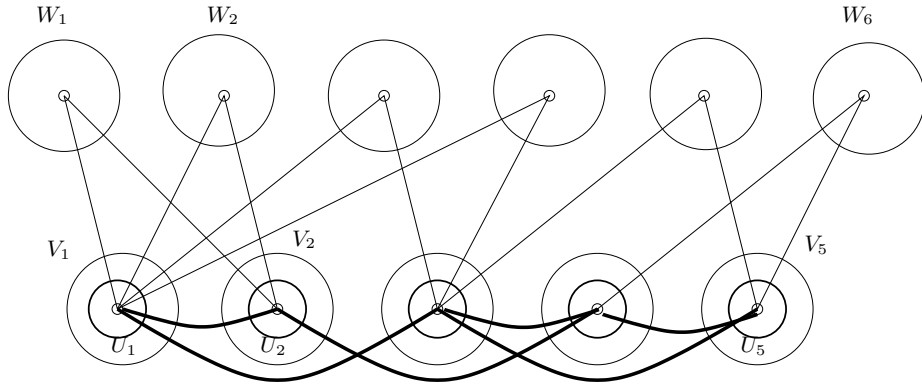


Figure 7: The auxiliary graph  $A$

**Remark 4.10** Note that most likely there will be many “fake” copies of  $M$  in  $A$ , that is, copies of  $M$  which do not correspond to any special constellation  $S(X)$ . There are two reasons for this overcount. First, we may create a copy of  $M$  in  $A$  over a set  $X \notin \mathcal{X}_3$ , or for which  $\widehat{X}$  is not isomorphic to  $\widehat{M}$ , but rather contains  $\widehat{M}$  as a proper subgraph. Second, in verifying the definition of an edge in  $A$  we may have used the vertices of  $\widehat{N}_i(b) \setminus N_i(b)$ , which, in fact, are not neighbors of  $v_i$ .

We will show now that  $A_{aa'}$  inherits from  $G$  some random-like properties. This in turn will allow us to use Proposition 4.3 to easily estimate the number

of copies of  $M$  in  $A$ . Targeting Property (P6) of  $G$ , for each  $aa'$  with  $x_a x_{a'} \in E(M)$ , define

$$\gamma_{aa'} = \prod_{b \in \Gamma_{aa'}} (1 - (1 - p)^{2r_{ab}}),$$

and note that

$$(1 + 2C^2)^{-\phi} \prod_{b \in \Gamma_{aa'}} 2pr_{ab} \leq \prod_{b \in \Gamma_{aa'}} \frac{2pr_{ab}}{1 + 2pr_{ab}} \leq \gamma_{aa'} \leq \prod_{b \in \Gamma_{aa'}} 2pr_{ab}, \quad (27)$$

where the leftmost inequality follows from the fact that  $2pr_{ab} \leq p(2np) \leq 2C^2$ . For our next lemma recall notation  $x \stackrel{\varepsilon}{\sim} y$  defined at the end of Section 1.

**Lemma 4.24** *For each  $aa'$  such that  $x_a x_{a'} \in E(M)$ , the graph  $A_{aa'}$  is  $n^{-1/30}$ -regular with density*

$$d(A_{aa'}) \stackrel{n^{-1/30}}{\sim} \gamma_{aa'}.$$

**Proof:** Our main tools will be Lemma 4.1 and Property (P6) of Definition 6.1. This property implies that for all  $v_i, v_j \in U_a$ , the degrees  $\deg(v_i)$  and co-degrees  $\text{codeg}(v_i, v_j)$  in  $A'_{aa'}$  are quite close to their averages. More precisely, let  $\Gamma_{aa'} = \{b_1, \dots, b_g\}$ . Applying Property (P6) first with  $k = g$ , and  $S_l = \overline{N}_i(b_l)$ ,  $l = 1, \dots, k$ , and then again with  $k = 2g$ ,  $S_l = \overline{N}_i(b_l)$  and  $S_{k+l} = \overline{N}_j(b_l)$ ,  $l = 1, \dots, k$ , we get, respectively,

$$(I) \quad \forall v \in U_a \quad |\deg(v) - n\gamma_{aa'}| < n^{4/5}, \quad \text{and}$$

$$(II) \quad \forall v, v' \in U_a \quad |\text{codeg}(v, v') - n\gamma_{aa'}^2| < n^{4/5}.$$

Properties (I) and (II) imply, respectively, (i) and (ii) of Lemma 4.1 with  $d = \gamma_{aa'}$ , and, say,  $\varepsilon = n^{-11/60}$  (and with  $W = [U_a]^2$ ). Indeed, by (I),

$$\deg(v) > n\gamma_{aa'} - n^{4/5} = n(\gamma_{aa'} - n^{-1/5}) > n(d - \varepsilon),$$

which is precisely condition (i) of Lemma 4.1. Also, by Corollary 4.21(b), and the above lower bound (27) on  $\gamma_{aa'}$ ,

$$\gamma_{aa'} \geq (1+2C^2)^{-\phi} \prod_{b \in \Gamma_{aa'}} 2pr_{ab} > \frac{1}{\log n} \left( \frac{np^2}{(\log n)^{\phi+4}} \right)^g > \frac{1}{\log n} \frac{c^{2\phi}}{(\log n)^{\phi(\phi+4)}} > n^{-1/60},$$

and so

$$\gamma_{aa'}^2 + n^{-1/5} < (\gamma_{aa'} + n^{-11/60})^2.$$

Thus, by (II),

$$\text{codeg}(v, v') < n\gamma_{aa'}^2 + n^{4/5} = n(\gamma_{aa'}^2 + n^{-1/5}) < n(d + \epsilon)^2,$$

and so, condition (ii) of Lemma 4.1 is true for *all* pairs  $v, v'$ . Moreover, Corollary 4.21(a) implies that for all  $1 \leq a \leq \nu$

$$|U_a| \geq \frac{n}{\log^{\phi+4} n} > 2n^{11/60} = \frac{2}{\epsilon}.$$

Hence, by Lemma 4.1 the graph  $A'_{aa'}$  is  $(16n^{-11/60})^{1/5}$ -regular. Recall that  $A_{aa'} \subset A'_{aa'}$  is the induced (bipartite) subgraph of  $A'_{aa'}$  on the vertex set  $U_a \cup U_{a'}$ . Therefore, we infer by Observation 4.2 with  $\epsilon' = |U_{a'}|/n \geq (\log n)^{-\phi-4}$ , that  $A_{aa'}$  is  $n^{-1/30}$ -regular.  $\square$

**Proof of Lemma 4.19:** In view of Lemma 4.24, we apply Proposition 4.3 with  $k = \nu$  and  $B_{aa'} = A_{aa'}$  if  $x_a x_{a'} \in E(M)$  and otherwise  $B_{aa'}$  being the bipartite complete graph between  $U_a$  and  $U_{a'}$ , so that  $\#(M, A) = \#(K_\nu, \bigcup_{a,a'} B_{aa'})$ . Hence,

$$\#(M, A) = (1 + o(1)) \prod_{a=1}^{\nu} |U_a| \prod_{aa'} \gamma_{aa'} \leq (1 + o(1))(2p)^\phi \prod_{a=1}^{\nu} |U_a| \prod_{b=1}^{\phi_a} r_{ab},$$

which, by Proposition 4.23 proves (24) and thus completes the proof of Lemma 4.19.  $\square$

## 5 The Core Section (Proof of Lemma 2.4)

In this section we will define the notion of a core, show that every hitting set of the family  $\mathcal{S}$  of all special constellations contains a core, prove that there are few cores and that every core is large. In other words, we shall prove Lemma 2.4. Fix an arbitrary  $\tau > 0$ .

## 5.1 What is a Core?

Cores will be defined through a  $\delta$ -regular partition  $\Pi_1$  of  $G$  guaranteed by Lemma 4.13. To make this definition more user-friendly, after applying the regularity lemma we will give suggestive names to all the parts of the final partition, and only then define cores formally.

### 5.1.1 Applying the Regularity Lemma

Let  $\alpha_3$  and  $q$  be as defined in Section 3, and constants  $\nu, c, C$  as in the Setup on page 16 (see also the Glossary). For an arbitrary  $\tau > 0$ , let  $\eta$  be any constant satisfying  $0 < \eta < 1$  and

$$\eta \log(1/\eta) \leq \frac{\tau}{4C\nu}.$$

Furthermore, let us set

$$\alpha_4 = \frac{(\nu + q)^{\nu+q}}{3C^{2q}} \alpha_3,$$

$$\delta = \min \left\{ 0.4\eta^q, \left( \frac{\alpha_4}{8D^*} \right)^2 \right\},$$

and

$$\varepsilon(\ell) = \min \left\{ \ell^{-4q}, \left( \frac{\delta \alpha_4}{D^* 2^{q+3} q (\nu + q)^{\nu+q-2}} \right)^{4q}, q^{-4}, 2^{-6-6q} \right\}.$$

For each star forest  $S$  with at most  $q$  edges there are constants  $T_0 = T_0(S)$ ,  $L_0 = L_0(S)$  and  $n_0 = n_0(S)$  determined by Lemma 4.13 with  $D = 2$ ,  $D^* = 2^q + 1$ , the above defined  $\delta$  and  $\varepsilon(\ell)$ , and with  $H = S$ . The meaning of these constants is that for every  $n > n_0$ , for all  $0 < p, p^* < 1$ , all  $S$ -uniform,  $S$ -graphs  $F$  with  $|V(F)| = n$ , and no  $(2, p)$ -dense patches, and for all  $\mathcal{S} \subseteq \mathcal{C}(F)$  with no  $(2^q + 1, p^*)$ -dense  $S$ -patches, there exists a refinement  $\Pi_1$  of the initial partition  $\Pi_0$ . This refined partition  $\Pi_1$  is a  $\delta$ -regular,  $(\varepsilon(\ell), p)$ -uniform,  $(t, l)$ -partition for some  $l \leq L_0$  and  $t \leq T_0$ . Set  $T_1 = \max_S T_0(S)$ ,  $L_1 = \max_S L_0(S)$  and  $n_1 = \max_S n_0(S)$ .

Let  $G \in \mathcal{G}$  be as in Lemma 2.4 (in particular,  $|V(G)| = n > n_1$ ) and let  $F$  be a subgraph of  $G \in \mathcal{G}$  determined by the partition  $\pi_0$ . Furthermore, let  $\mathcal{S}$  be the family of special constellations in  $F$  generated by assumption (3) of Theorem 2.1 as shown in Section 3. Recall that  $S$ , a star forest with  $\nu$  stars

and  $\phi \leq q$  arms altogether, is the isomorphism type of the constellations. For clarity, we will be further assuming that  $n$  is divisible by  $\nu + \phi$ , so that the initial equipartition  $\Pi_0$  has, in fact, all sets of equal size. We will apply Lemma 4.13 to the pair  $(F, \mathcal{S})$  with  $p = p(n)$  as in the Setup on page 16 and  $p^* = p^\phi$ . We have proved in Section 4.4 that in such a setting  $(F, \mathcal{S})$  does fulfill all the above assumptions of Lemma 4.13. For the purpose of regularization, set  $F_{ab} = G[V_a, W_{ab}]$  and let  $\Pi_0$  denote the *initial partition* made of  $\pi_0$ , the partition of vertices, and of the edge sets between the corresponding vertex sets. In other words, the elements of  $\Pi_0$  are:

- vertex sets:  $V_a, a = 1, \dots, \nu; W_{ab}, b = 1, \dots, \phi_a$ , for  $a = 1, \dots, \nu - 1$ .
- edge sets:  $E(F_{ab}), a = 1, \dots, \nu, b = 1, \dots, \phi_a$ .

Finally, let

$$F = \bigcup_{ab} F_{ab}$$

and note that  $F \in \mathcal{F}(S, m)$ , where

$$m = \frac{n}{\nu + \phi} = |V_a| = |W_{ab}|.$$

Let  $\Pi_1$  be the final  $(\varepsilon(l), p)$ -uniform,  $\delta$ -regular  $(t, l)$ -partition,  $t \leq T_1$ ,  $l \leq L_1$ , guaranteed by Lemma 4.13. It consists of vertex equipartitions  $V_a = V_a^1 \cup \dots \cup V_a^t$ ,  $a = 1, \dots, \nu$ , and  $W_{ab} = W_{ab}^1 \cup \dots \cup W_{ab}^t$ ,  $a = 1, \dots, \nu - 1$ ,  $b = 1, \dots, \phi_a$ , and graph partitions

$$F_{ab}^{i,j} = G[V_a^i, W_{ab}^j] = \bigcup_{k=1}^{l(ab,i,j)} F_{ab}^{i,j,k}, \quad a = 1, \dots, \nu - 1, \quad b = 1, \dots, \phi_a,$$

$$l(ab, i, j) \leq l, \quad i, j = 1, \dots, t.$$

For simplicity, we assume that  $m$  is divisible by  $t$  and thus for all  $a, b, i$  and  $j$  we have

$$|V_a^i| = |W_{ab}^j| = \frac{m}{t}.$$

The number of edges in different graphs  $F_{ab}^{i,j,k}$  may vary. Note that  $F = \bigcup_{a,b,i,j,k} F_{ab}^{i,j,k}$ .

**Remark 5.1** The assumption that  $t$  divides  $m$  is not very serious. Since  $t \leq T_1$ , every  $m$  which is divisible by  $T_1!$  is also divisible by  $t$ , no matter what  $t$  turns out to be. Thus, our proof is absolutely formal for all  $m$  which are the multiples of  $T_1$ , that is, for all  $n$  which are the multiples of  $T_1!(\nu + \phi)$ . The argument for other values of  $n$  is essentially the same with only some obvious, though tedious, notational complications.

### 5.1.2 Anatomy and Polyads

We use suggestive terminology throughout this section. The definitions of tubes and polyads given below are equivalent to those introduced in Section 4 in a more general setting of  $H$ -graphs. Figure 8 illustrates these notions.

- The sets  $V_a^i$  will be called *palms*.
- The sets  $W_{ab}^j$  will be called *fingernails*.
- The subgraphs  $F_{ab}^{ijk}$  will be called *fingers*.
- The induced subgraphs  $F_{ab}^{ij} = \bigcup_{k=1}^{l(ab,i,j)} F_{ab}^{ijk}$  will be called *tubes*.
- A *polyad*  $P$  consists of  $\nu$  palms  $V_a^P$  (one for each  $a$ ),  $\phi$  fingernails  $W_{ab}^P$  ( $\phi_a$  for each  $a$ ), and  $\phi$  fingers  $F_{ab}^P$  (one from each tube  $F_{ab}^{ij}$ , where  $V_a^P = V_a^i$  and  $W_{ab}^P = W_{ab}^j$ ).
- An *olympiad*  $\mathcal{O}$  consists of  $\nu$  palms  $V_a^{\mathcal{O}}$  (one for each  $a$ ),  $\phi$  fingernails  $W_{ab}^{\mathcal{O}}$  ( $\phi_a$  for each  $a$ ), and  $\phi$  tubes  $F_{ab}^{\mathcal{O}} = F_{ab}^{ij}$ , where  $V_a^{\mathcal{O}} = V_a^i$  and  $W_{ab}^{\mathcal{O}} = W_{ab}^j$ .
- The partition  $\Pi_1$  will be called *myriad* as it contains all the elements of the puzzle, that is, it consists of all  $t\nu$  palms,  $t\phi$  fingernails, and at most  $t^2l\nu\phi$  fingers.

Note that there are precisely  $t^{\nu+\phi}$  olympiads in the myriad, and at most  $l^\phi$  polyads in any given olympiad. Throughout, we will treat the olympiads (and the myriad) both as a set of polyads and as a subgraph of  $F$  (which is the union of those polyads.)

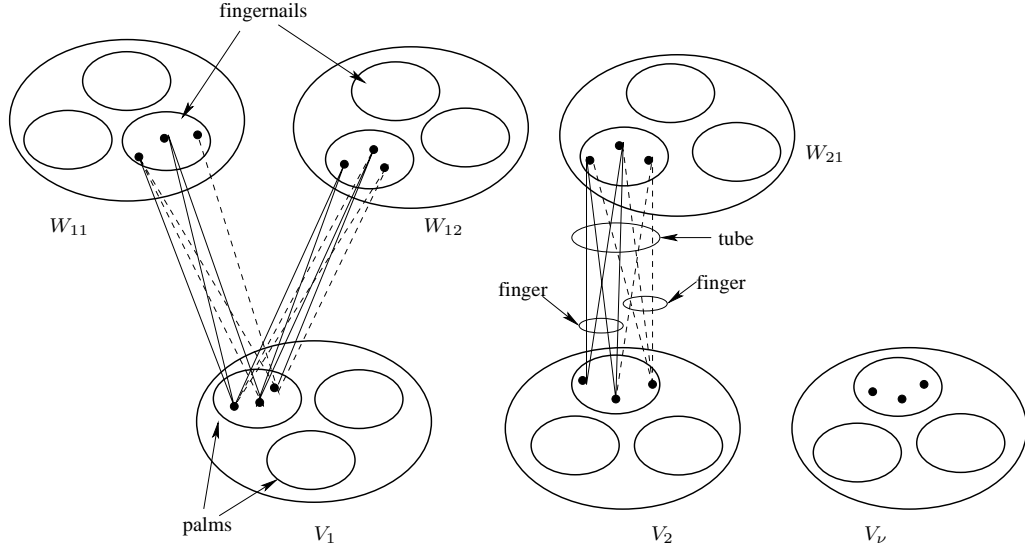


Figure 8: The Anatomy

### 5.1.3 The Definition of a Core

Let  $\varepsilon = \varepsilon(l)$ . We say that a polyad  $P$  is *nice* if all the fingers in  $P$  are  $(\varepsilon, p)$ -regular, and have  $p$ -density at least  $3\varepsilon^{\frac{1}{2\phi}}$ . We say that a polyad  $P$  is *healthy* if the following three conditions hold:

- $P$  is nice,
- $P$  is  $\delta$ -regular,
- $d_P > \delta$ .

We define precores first. A *precore* will be a subgraph which contains a substantial part of every healthy polyad  $P$ . Roughly speaking these parts will be formed by choosing at least a half of the vertices of just one palm  $V_a^P$ , and for each vertex  $v \in V_a^P$  choosing a finger  $F_{ab}^P$ , where  $b = b(v)$ , and taking more than a  $(1 - \eta)$ -fraction of the  $\deg_P(v, W_{a,b}^P) = \deg_{F_{ab}^P}(v)$  edges incident to  $v$  in that finger, where  $\eta$  is defined at the beginning of Section 5.1.1.

More formally, denote the set of all healthy polyads by  $\mathcal{P}^{healthy}$ . For each  $P \in \mathcal{P}^{healthy}$  define a family  $\text{PRECORE}^P$  of subgraphs of  $P$  as follows:  $J \in \text{PRECORE}^P$  if  $J \subseteq P$  and

- (i)  $\exists 1 \leq a_J \leq \nu : |V(J) \cap V_a^P| \begin{cases} = 0 & \text{for } a \neq a_J \\ \geq \frac{m}{2t} = 0.5|V_a^P| & \text{for } a = a_J \end{cases}$
- (ii)  $\forall v \in V(J) \cap V_a^P \exists b(v) \in \{1, \dots, \phi_a\}$  such that
- $$\deg_J(v, W_{a_J, b}^P) \begin{cases} = 0 & \text{for } b \neq b(v) \\ > (1 - \eta)\deg_P(v, W_{a_J, b}^P) & \text{for } b = b(v) \end{cases}.$$

Now we are ready to define the family  $\text{PRECORE}$  of subgraphs of  $F$ :

$$\text{PRECORE} = \left\{ J : J = \bigcup_{P \in \mathcal{P}^{healthy}} J^P, \quad J^P \in \text{PRECORE}^P \right\} \quad (28)$$

The graphs in  $\text{PRECORE}$  are called *precores*. Thus, a precore  $J$  is a union of subgraphs  $J^P$ , one for every healthy polyad  $P$ .

We are one step from the definition of a core.

Recall that the (pre)coloring  $\sigma$  of all the edges of  $\bigcup_{X \in \mathcal{X}_3} S(X) \subseteq F$ , that is, of all the edges of special constellations, is determined by a fixed triangle-free coloring  $\sigma'$  of the graph  $M$ . For any precore  $J$  contained in  $\bigcup_{X \in \mathcal{X}_3} S(X)$ , the coloring  $\sigma$  partitions its edges into  $J_{red}$  and  $J_{blue}$ . Define  $\text{Maj}(J)$  to be the bigger of the two sets (ties are broken arbitrarily). Then

$$\text{CORE} = \left\{ \text{Maj}(J) : J \in \text{PRECORE}, J \subseteq \bigcup_{X \in \mathcal{X}_3} S(X) \right\}.$$

The graphs in  $\text{CORE}$  are called *cores*. The reader is encouraged to draw an analogy between the above definition and that given in Section 2.2.1.

In the forthcoming subsections we will verify that the family  $\text{CORE}$  defined above satisfies conditions (a-c) of Lemma 2.4.

## 5.2 Every Hitting Set of $\mathcal{S}$ Contains a Core

We have come to the heart of our construction. The following lemma, together with Lemma 3.5, immediately implies Lemma 2.4(a) – see below. Recall that a hitting set of a family of graphs is a set of edges which intersects the edge set of every graph in the family.

**Lemma 5.1** *Let  $T$  be a hitting set of the family  $\mathcal{S}$  of special constellations. Then there exists a precore  $J \in \text{PRECORE}$  such that  $J \subseteq T$ .*

**Proof:** Consider an arbitrary subset  $T$  of  $E(F)$  not containing any precore  $J \in \text{PRECORE}$ . We are going to show that there is a special constellation disjoint from  $T$ . Because  $T$  contains no precore, there exists  $P_0 \in \mathcal{P}^{\text{healthy}}$  such that for all  $J \in \text{PRECORE}^{P_0}$  we have  $J \not\subseteq T$  (otherwise,  $T$  would contain a union as in (28)). That is, for all  $a = 1, \dots, \nu$ , the set

$$\tilde{V}_a^{P_0} = \{v \in V_a^{P_0} : \forall b = 1, \dots, \phi_a : \deg_{T \cap F_{ab}^{P_0}}(v) \leq (1 - \eta) \deg_{F_{ab}^{P_0}}(v)\}$$

has size

$$|\tilde{V}_a^{P_0}| > \frac{m}{2t}.$$

In other words, more than half of the vertices of each palm retain in  $T$  at most  $(1 - \eta)$ -fraction of their neighbors they had within each finger of  $P_0$ . Defining  $Q_0 = P_0 - T$ , we see that for all  $a = 1, \dots, \nu$ , if  $v \in \tilde{V}_a^{P_0}$  then for all  $b = 1, \dots, \phi_a$

$$\deg_{Q_0 \cap F_{ab}^{P_0}}(v) \geq \eta \deg_{F_{ab}^{P_0}}(v).$$

Recall that for a subgraph  $R$  of  $F$ , symbols  $c_R$  and  $s_R$  stand, respectively, for the numbers of constellations and special constellations contained in  $R$ . Let  $\tilde{Q}_0$  and  $\tilde{P}_0$  be induced subgraphs of, respectively,  $Q_0$  and  $P_0$  with sets  $V_a^{P_0}$  restricted to  $\tilde{V}_a^{P_0}$ . Then, trivially  $c_{Q_0} \geq c_{\tilde{Q}_0}$  and, by formula (21),  $c_{\tilde{Q}_0} \geq \eta^\phi c_{\tilde{P}_0}$ .

Since  $P_0$  is nice, all fingers  $F_{ab}^{P_0}$  are  $(\varepsilon, p)$ -regular with  $p$ -density at least  $3\varepsilon^{\frac{1}{2\phi}}$ . Moreover, by Observation 4.6, all subfingers  $F_{ab}^{P_0} \cap \tilde{P}_0$  are  $(2\varepsilon, p)$ -regular with  $p$ -density at least  $3\varepsilon^{\frac{1}{2\phi}} - \varepsilon > (2\varepsilon)^{\frac{1}{2\phi}}$  and also  $d_p(F_{ab}^{P_0}) \stackrel{\varepsilon}{\approx} d_p(F_{ab}^{P_0} \cap \tilde{P}_0)$ . Hence, by two applications of Proposition 4.8, to  $P_0$  and to  $\tilde{P}_0$ ,

$$\frac{c_{\tilde{P}_0}}{c_{P_0}} > (1 - \varepsilon^{1/4})(0.5)^\nu > (0.4)^\nu,$$

and consequently,

$$c_{Q_0} \geq c_{\tilde{Q}_0} \geq \eta^\phi c_{\tilde{P}_0} \geq (0.4)^\nu \eta^\phi c_{P_0} > \delta c_{P_0},$$

by our definition of  $\delta$ . Therefore, since  $P_0$  is a  $\delta$ -regular polyad of  $(S, p^*)$ -density greater than  $\delta$ ,

$$\frac{s_{Q_0}}{p^\phi c_{Q_0}} = d_{Q_0} > d_{P_0} - \delta > 0,$$

implying that  $s_{Q_0} \geq 1$ . This means that there exists a promised special constellation  $S \in \mathcal{S}$  such that  $S \subseteq Q_0 \subseteq F - T$ .  $\square$

**Proof of Lemma 2.4(a):** Note that for every triangle-free coloring  $\chi$  of  $G$  the set  $\text{Agree}(\chi, \sigma)$  is a hitting set of the family  $\mathcal{S}$ . Thus, by Lemma 5.1, there is a precore  $J$  contained in  $\text{Agree}(\chi, \sigma)$ . Since the latter set is contained in  $\bigcup_{X \in \mathcal{X}_3} S(X) \subseteq F$ , the set  $\text{Maj}(J)$  is a monochromatic core under  $\chi$ .  $\square$

### 5.3 There Are Few Cores

Now we give a quick proof of Lemma 2.4(b).

**Lemma 5.2**  $|\text{CORE}| < \exp\{\tau n^{3/2}\}$ .

**Proof:** We will actually show that  $|\text{PRECORE}| < \exp\{\tau n^{3/2}\}$ , which is stronger. It follows from the definition of precores that for all  $J \in \text{PRECORE}$ , all  $v \in V(J) \cap \bigcup_{a=1}^{\nu} V_a$ , and for all  $a, b, i, j$  and  $k$

$$\deg_{J \cap F_{ab}^{ijk}}(v) \begin{cases} = 0, & \text{or} \\ \geq (1 - \eta) \deg_{F_{ab}^{ijk}}(v). \end{cases}$$

We will bound the number of *all* subgraphs of  $F$  with that property. For every vertex  $v \in V_a$  there are at most  $2^{\phi t l}$  choices of the “substantial” fingers  $F_{ab}^{ijk}$  along which  $v$  has positive degree. For each choice of substantial fingers, if the degree of  $v$  in their union is  $r$ , then the number of choices of neighbors of  $v$  in  $J$  (which is also the number of choices of the non-neighbors) is at most

$$\sum_{k=0}^{\eta r} \binom{r}{k} \leq \eta r \binom{r}{\eta r} < n \binom{2pn}{\eta 2pn},$$

by Property (P3) of  $G$ . Thus

$$|\text{PRECORE}| < \left( 2^{\phi t l} n \binom{2pn}{\eta 2pn} \right)^{\nu n} < (2^{\phi t l} n)^{\nu n} \left( \frac{e}{\eta} \right)^{2C\eta\nu} < e^{\tau n^{3/2}}$$

for  $n$  sufficiently large, by the definition of  $\eta$ .  $\square$

## 5.4 Cores Are Large

In this section we will show the most difficult part of Lemma 2.4, that is, part (c). We will prove it with

$$\lambda = \frac{1}{160} \left( \frac{\alpha_4}{D^* q (\nu + q)} \right)^2,$$

where

$$\alpha_4 = \frac{(\nu + q)^{\nu+q}}{3C^{2q}} \alpha_3 \quad \text{and} \quad D^* = 2^q + 1.$$

Recall that olympiads were defined in Section 5.1.2 and can be viewed as sets of polyads. The plan is to cover a precore  $J$  by subgraphs arising from different olympiads:

$$J = \bigcup_{\mathcal{O} \in \Pi_1} J_{\mathcal{O}}, \quad \text{where} \quad J_{\mathcal{O}} = \bigcup_{P \in \mathcal{O}^{\text{healthy}}} J^P \quad \text{and} \quad \mathcal{O}^{\text{healthy}} = \mathcal{O} \cap \mathcal{P}^{\text{healthy}},$$

and then to show (see Lemma 5.6 below) that for a certain type of olympiads, called perfect, we have

$$|J_{\mathcal{O}}| \geq \frac{\alpha_4 p m^2}{10 D^* \phi^2 t^2}.$$

We will prove that there are at least  $(\alpha_4/8D^*)t^{\nu+\phi}$  perfect olympiads (see Lemma 5.4 below). This and the obvious fact that an edge may belong to at most  $t^{\nu+\phi-2}$  olympiads, yields the following result.

**Lemma 5.3** *For all  $J \in \text{PRECORE}$  we have  $|J| \geq 2\lambda p n^2$ .*

**Proof:** By Lemmas 5.4 and 5.6 we have

$$|J| \geq \frac{\alpha_4}{8D^*} t^{\nu+\phi} \times \frac{\alpha_4 p m^2}{10 D^* \phi^2 t^2} \times t^{-\nu-\phi+2} = 2\lambda p n^2.$$

□

Lemma 5.3, together with the fact that  $|E(G)| < p n^2$  (see Property (P3) of  $\mathcal{G}$ ), imply Lemma 2.4(c).

### 5.4.1 Perfect Olympiads

Recall that  $\kappa = \sum_{P \in \mathcal{P}_{\Pi_1}} c_P$  is the total volume of the myriad. By Proposition 4.8 (see Example 4.9)

$$\kappa \sim m^{\nu+\phi} p^\phi.$$

Let

$$\kappa_0 = \frac{\kappa}{t^{\nu+\phi}}$$

be the average volume of an olympiad. Setting

$$c_{\mathcal{O}} = \sum_{P \in \mathcal{O}} c_P,$$

we have, again by Proposition 4.8, that for all olympiads  $\mathcal{O}$

$$c_{\mathcal{O}} \sim \kappa_0.$$

Let  $\mathcal{O}^{\text{reg}}$  denote the set of  $\delta$ -regular polyads contained in  $\mathcal{O}$ , and  $\mathcal{O}^{\text{irr}} = \mathcal{O} \setminus \mathcal{O}^{\text{reg}}$  be the set of irregular ones. Write

$$c_{\mathcal{O}} = c_{\mathcal{O}}^{\text{reg}} + c_{\mathcal{O}}^{\text{irr}}$$

where  $c_{\mathcal{O}}^{\text{reg}} = \sum_{P \in \mathcal{O}^{\text{reg}}} c_P$  and  $c_{\mathcal{O}}^{\text{irr}} = \sum_{P \in \mathcal{O}^{\text{irr}}} c_P$ , and set

$$s_{\mathcal{O}} = \sum_{P \in \mathcal{O}} s_P.$$

Finally, let  $\mathcal{O}^{\text{nice}}$  denote the set of nice polyads contained in  $\mathcal{O}$  and

$$c_{\mathcal{O}}^{\text{nice}} = \sum_{P \in \mathcal{O}^{\text{nice}}} c_P.$$

**Definition 5.1** We call an olympiad  $\mathcal{O}$  *perfect* if the following three conditions hold:

$$c_{\mathcal{O}}^{\text{nice}} \geq (1 - \delta)\kappa_0 \tag{29}$$

$$c_{\mathcal{O}}^{\text{reg}} \geq (1 - \sqrt{\delta})\kappa_0, \tag{30}$$

and

$$s_{\mathcal{O}} \geq \alpha_4 p^\phi \kappa_0 \tag{31}$$

**Lemma 5.4** *At least  $\frac{\alpha_4}{8D^*}t^{\nu+\phi}$  olympiads are perfect.*

**Proof:** By Property (P5) of  $\mathcal{G}$ ,

$$|E(F)| \sim \phi m^2 p.$$

Let us begin with a count of the edges of  $F$  belonging to fingers which are either  $(\varepsilon, p)$ -irregular or of  $p$ -density smaller than  $3\varepsilon^{1/2\phi}$ . Let us call such fingers *nasty*. First, from the  $(\varepsilon, p)$ -uniformity of partition  $\Pi_1$ , there are at most  $\varepsilon|E(F)|$  edges belonging to fingers that are  $(\varepsilon, p)$ -irregular.

Next note that every tube is an induced subgraph of  $F$  split into at most  $l$  fingers, so even if they were all of  $p$ -density smaller than  $3\varepsilon^{1/2\phi}$ , they would contribute no more than a  $3l\varepsilon^{1/2\phi}$ -fraction of the edges. As we have  $l \leq \varepsilon^{-1/4\phi}$  and  $\varepsilon \leq \varepsilon^{1/4\phi}$ , altogether there are fewer than  $4\varepsilon^{1/4\phi}|E(F)|$  edges which belong to nasty fingers of  $F$ .

Furthermore, by Property (P3) of  $\mathcal{G}$ , any given edge may belong to at most  $n^{\nu-1}(2np)^{\phi-1}$  constellations. Call a constellation *spoiled* if at least one of its edges belongs to a nasty finger. Then, recalling the definition of a nice polyad from Section 5.1.3,  $c_{\mathcal{O}} - c_{\mathcal{O}}^{\text{nice}}$  counts precisely those constellations of olympiad  $\mathcal{O}$  which are spoiled. Thus,

$$\sum_{\mathcal{O}} (c_{\mathcal{O}} - c_{\mathcal{O}}^{\text{nice}}) \leq (1 + o(1))4\varepsilon^{\frac{1}{4\phi}}\phi m^2 p n^{\nu-1} (2np)^{\phi-1}.$$

On the other hand, denoting by  $\Upsilon_1^-$  the number of olympiads which *do not* satisfy condition (29), we have

$$\Upsilon_1^- \delta \kappa_0 \leq \sum_{\mathcal{O}} (\kappa_0 - c_{\mathcal{O}}^{\text{nice}}),$$

which, using the definitions of  $\varepsilon$ ,  $\kappa_0$ ,  $\kappa$ , and  $m$ , and the relation  $c_{\mathcal{O}} \sim \kappa_0$ , yields the bound  $\Upsilon_1^- \leq (1 + o(1))(\alpha_4/4D^*)t^{\nu+\phi}$ .

Next, let  $\Upsilon_2^-$  be the number of olympiads which *do not* satisfy condition (30). We claim that  $\Upsilon_2^- \leq \sqrt{\delta}t^{\nu+\phi} \leq (\alpha_4/8D^*)t^{\nu+\phi}$ , where the latter inequality follows just by the definition of  $\delta$ . Indeed, if  $\Upsilon_2^- > \sqrt{\delta}t^{\nu+\phi}$  then

$$\sum_{P \in \mathcal{P}_{\Pi_1}^{\text{irr}}} c_P = \sum_{\mathcal{O}} \sum_{P \in \mathcal{O}^{\text{irr}}} c_P > \sqrt{\delta}t^{\nu+\phi} \sqrt{\delta}\kappa_0 = \delta\kappa$$

– contradiction with the  $\delta$ -regularity of  $\Pi_1$  (compare Definition 4.12).

Finally, let  $\Upsilon_3^+$  be the number of olympiads fulfilling condition (31). We claim that  $\Upsilon_3^+ \geq \frac{2}{3}(\alpha_4/D^*)t^{\nu+\phi}$ . Indeed, by Proposition 4.12 and the fact that  $F$  has no  $(D^*, p^\phi)$ -dense  $S$ -patches, we have, for every olympiad  $\mathcal{O}$ ,

$$\begin{aligned} s_{\mathcal{O}} &= \sum_{c_P < \kappa/\log^2 n} s_P + \sum_{c_P \geq \kappa/\log^2 n} s_P \leq D^* p^\phi \kappa / \log^2 n + D^* p^\phi \sum_{P \in \mathcal{O}} c_P \\ &= o(p^\phi \kappa_0) + D^* p^\phi \kappa_0 < \frac{3}{2} D^* p^\phi \kappa_0. \end{aligned} \quad (*)$$

Recall from Section 3 that  $|\mathcal{S}| = \alpha_3 n^\nu$ , and suppose that  $\Upsilon_3^+ < \frac{2}{3}(\alpha_4/D^*)t^{\nu+\phi}$ . Then

$$|\mathcal{S}| < \frac{2}{3} \frac{\alpha_4 t^{\nu+\phi}}{D^*} \frac{3}{2} D^* p^\phi \kappa_0 + t^{\nu+\phi} \alpha_4 p^\phi \kappa_0 = 2\alpha_4 p^\phi \kappa < \alpha_3 n^\nu$$

– a contradiction. Hence, we conclude that there are at least

$$\Upsilon_3^+ - \Upsilon_1^- - \Upsilon_2^- \geq \left( \frac{2}{3} \frac{\alpha_4}{D^*} - \frac{\alpha_4}{4D^*} - \frac{\alpha_4}{8D^*} - o(1) \right) t^{\nu+\phi} > \frac{\alpha_4}{8D^*} t^{\nu+\phi}$$

perfect olympiads.  $\square$

Recall that  $\mathcal{O}^{\text{healthy}} = \mathcal{O} \cap \mathcal{P}^{\text{healthy}}$  is the set of healthy polyads contained in an olympiad  $\mathcal{O}$ . The next lemma will be used in the proof of Lemma 5.6.

**Lemma 5.5** *For a perfect olympiad  $\mathcal{O}$*

$$\sum_{P \in \mathcal{O}^{\text{healthy}}} c_P \geq \frac{\alpha_4}{2D^*} \kappa_0.$$

**Proof:** Because there are no  $(D^*, p^\phi)$ -dense  $S$ -patches in  $F$ , we have  $c_P \geq s_P / (D^* p^\phi)$  for each polyad  $P$  with  $c_P \geq \kappa / \log^2 n$ . Hence

$$\sum_{P \in \mathcal{O}^{\text{healthy}}} c_P \geq \frac{1}{D^* p^\phi} \sum_{\substack{P \in \mathcal{O}^{\text{healthy}} \\ c_P \geq \kappa / \log^2 n}} s_P.$$

Clearly,

$$\sum_{\substack{P \in \mathcal{O}^{\text{healthy}} \\ c_P \geq \kappa / \log^2 n}} s_P \geq s_{\mathcal{O}} - \sum_{\substack{P \notin \mathcal{O}^{\text{healthy}} \\ c_P \geq \kappa / \log^2 n}} s_P - \sum_{c_P < \kappa / \log^2 n} s_P.$$

Recalling the definition of a healthy polyad, to estimate further we must then subtract from  $s_{\mathcal{O}}$  the contribution coming from polyads with  $c_P \geq \kappa/\log^2 n$  that are either not nice, or  $\delta$ -irregular, or of  $(S, p^*)$ -density less than  $\delta$ , and also from polyads with  $c_P < \kappa/\log^2 n$ . First, since  $\mathcal{O}$  satisfies condition (29),

$$\sum_{\substack{P \notin \mathcal{O}^{\text{nice}} \\ c_P \geq \kappa/\log^2 n}} s_P \leq D^* p^\phi \sum_{P \notin \mathcal{O}^{\text{nice}}} c_P = D^* p^\phi (c_{\mathcal{O}} - c_{\mathcal{O}}^{\text{nice}}) \leq (1 + o(1)) \delta D^* p^\phi \kappa_0.$$

Second, using the fact that  $\mathcal{O}$  obeys condition (30),

$$\sum_{\substack{P \in \mathcal{O}^{\text{irr}} \\ c_P \geq \kappa/\log^2 n}} s_P < D^* p^\phi \sum_{P \in \mathcal{O}^{\text{irr}}} c_P = D^* p^\phi (c_{\mathcal{O}} - c_{\mathcal{O}}^{\text{reg}}) \leq (1 + o(1)) \sqrt{\delta} D^* p^\phi \kappa_0.$$

Third,

$$\sum_{P: d_P < \delta} s_P < \delta p^\phi \sum_{P: d_P < \delta} c_P \leq \delta p^\phi c_{\mathcal{O}} = (1 + o(1)) \delta p^\phi \kappa_0.$$

Last, by Proposition 4.12

$$\sum_{c_P < \kappa/\log^2 n} s_P \leq (1 + o(1)) D^* p^\phi \frac{\kappa}{\log^2 n} = o(1) p^\phi \kappa_0.$$

Putting all this together and using the fact that  $\mathcal{O}$  fulfills (gold) and thus  $s_{\mathcal{O}} \geq \alpha_4 p^\phi \kappa_0$ , we finally get

$$\begin{aligned} \sum_{\substack{P \in \mathcal{O}^{\text{healthy}} \\ c_P \geq \kappa/\log^2 n}} s_P &\geq \left( \alpha_4 - \delta D^* - \sqrt{\delta} D^* - \delta - o(1) \right) p^\phi \kappa_0 \\ &> \left( \alpha_4 - 3\sqrt{\delta} D^* - o(1) \right) p^\phi \kappa_0 > \left( \alpha_4 - \frac{3}{8} \alpha_4 - o(1) \right) p^\phi \kappa_0 \\ &> \frac{1}{2} \alpha_4 p^\phi \kappa_0 \end{aligned}$$

and consequently

$$\sum_{P \in \mathcal{O}^{\text{healthy}}} c_P \geq \frac{\alpha_4}{2D^*} \kappa_0.$$

□

## 5.4.2 Precores Are Large

We now concentrate on the contribution to the precore  $J$  coming from healthy polyads of one perfect olympiad. Recall our notation

$$J_{\mathcal{O}} = \bigcup_{P \in \mathcal{O}^{\text{healthy}}} J^P.$$

The following lemma, together with Lemma 5.4 completes the proof of Lemma 5.3, which, in turn, finishes off the proof of Lemma 2.4, yielding Theorem 2.1, and thus our main result – Theorem 1.1.

**Lemma 5.6** *For every precore  $J \in \text{PRECORE}$  and for every perfect olympiad  $\mathcal{O}$ , we have*

$$|J_{\mathcal{O}}| \geq \frac{\alpha_4 p m^2}{10 D^* \phi^2 t^2}.$$

**Proof.** The idea of the proof of Lemma 5.6 is to take only one finger from each polyad  $P$  – the one chosen by the majority of the vertices in the definition of  $J^P$  and then show, using Lemma 5.5 and Proposition 4.8, that many of those fingers will belong to the same tube. As such, they will be edge-disjoint and adding up their contributions to the respective  $J^P$ 's will give the desired lower bound on  $|J_{\mathcal{O}}|$ .

Recall the definition of the family  $\text{PRECORE}^P$  from Section 5.1.3. For every  $P \in \mathcal{O}^{\text{healthy}}$  and for every  $J^P \in \text{PRECORE}^P$ , let  $a_P$  and  $b_P$  be such that

$$|V(J^P) \cap V_{a_P}^{\mathcal{O}}| \geq \frac{m}{2t} \text{ and } |\{v \in V(J^P) \cap V_{a_P}^{\mathcal{O}} : b(v) = b_P\}| \geq \frac{m}{2\phi t}.$$

Then  $a_P$  and  $b_P$  determine a finger of  $P$ , denoted further by  $F^P = F_{a_P b_P}^P$ , belonging to the tube  $F_{a_P b_P}^{\mathcal{O}}$ . Obviously,

$$\left| \bigcup_{P \in \mathcal{O}^{\text{healthy}}} J^P \right| \geq \left| \bigcup_{P \in \mathcal{O}^{\text{healthy}}} (J^P \cap F^P) \right|.$$

Note that by the definition of  $J^P$ , by the  $(\varepsilon, p)$ -regularity of fingers and the inequality  $\varepsilon \leq 1/(2\phi)$ ,

$$|J^P \cap F^P| \geq \frac{1}{2\phi} (1 - \eta)(1 - \varepsilon) |F^P| > \frac{|F^P|}{3\phi}, \quad (32)$$

where the last inequality follows because, say, both  $\eta < 1/6$  and  $\varepsilon < 1/6$ . For every  $a, b$  consider the edge-disjoint union  $\overline{F}_{ab}^{\mathcal{O}}$  of selected fingers  $F^P$  belonging to the tube  $F_{ab}^{\mathcal{O}}$ , that is

$$\overline{F}_{ab}^{\mathcal{O}} = \bigcup_{a_P=a, b_P=b} F^P \subseteq F_{ab}^{\mathcal{O}}.$$

We claim that for some  $a, b$

$$|\overline{F}_{ab}^{\mathcal{O}}| \geq \frac{\alpha_4}{3D^*\phi} p \left(\frac{m}{t}\right)^2. \quad (33)$$

Suppose to the contrary that the opposite inequality is true for all  $a, b$ , and consider the subgraph  $R$  of  $F$  which is a union of the complements  $F_{ab}^{\mathcal{O}} - \overline{F}_{ab}^{\mathcal{O}}$  of  $\overline{F}_{ab}^{\mathcal{O}}$  within the respective tubes. Then, since every healthy polyad of  $\mathcal{O}$  must intersect some  $\overline{F}_{ab}^{\mathcal{O}}$ , the subgraph  $R$  is a union of unhealthy polyads, and by Lemma 5.5

$$c_R \leq c_{\mathcal{O}} - \frac{\alpha_4}{2D^*} \kappa_0 = \left(1 + o(1) - \frac{\alpha_4}{2D^*}\right) \kappa_0.$$

On the other hand, by property (P5) of Definition 6.1, tubes are  $(\varepsilon, p)$ -regular of  $p$ -density  $1 - o(1)$ , and the graphs  $F_{ab}^{\mathcal{O}} - \overline{F}_{ab}^{\mathcal{O}}$  are complements of unions of at most  $l$   $(\varepsilon, p)$ -regular fingers within the tube  $F_{ab}^{\mathcal{O}}$ . Thus, they are themselves  $(\varepsilon', p)$ -regular, where  $\varepsilon' \leq l\varepsilon \leq \varepsilon^{1-1/4\phi}$ , with  $p$ -densities at least  $1 - \alpha_4/(3D^*\phi) + o(1)$ . Thus, by Proposition 4.8 and by (33),

$$c_R \stackrel{\varepsilon}{\approx} \left(\frac{m}{t}\right)^{\nu+\phi} p^\phi \prod_{ab} d_p(F_{ab}^{\mathcal{O}} - \overline{F}_{ab}^{\mathcal{O}}) \geq \left(1 + o(1) - \frac{\alpha_4}{3D^*\phi}\right)^\phi \kappa_0,$$

contradicting the above upper bound.

Hence, let  $a_0, b_0$  be such that (33) holds. Then

$$\left| \bigcup_{P \in \mathcal{O}^{\text{healthy}}} (J^P \cap F^P) \right| \geq \left| \bigcup_{a_P=a_0, b_P=b_0} (J^P \cap F^P) \right| = \sum_{a_P=a_0, b_P=b_0} |J^P \cap F^P|.$$

Using (32) and (33) we easily estimate further to obtain the desired bound:

$$\sum_{a_P=a_0, b_P=b_0} |J^P \cap F^P| \geq \sum \frac{|F^P|}{3\phi} = \frac{|\overline{F}_{a_0 b_0}^{\mathcal{O}}|}{3\phi} \geq \frac{\alpha_4 p m^2}{10D^*\phi^2 t^2}.$$

□

## 6 Random Graphs

In this section we state and prove some assertions about random graphs which constitute the family  $\mathcal{G}$  of graphs playing an important role in our proof.

Let  $G = (V, E)$  be a graph, and let  $v, u \in V$ ,  $U, W \subseteq V$ , where  $U \cap W = \emptyset$ , and  $F \subset E$ . Recall our notation

$$\deg_G(v, W) = \deg(v, W) = |\{w \in W : vw \in E\}|,$$

$$\deg_G(v) = \deg(v) = \deg(v, V),$$

$$\text{codeg}(v, u) = |\{w : vw, uw \in E\}|,$$

and  $e_G(U)$  and  $e_G(U, W)$  for the number of edges of  $G$  with, resp., both endpoints in  $U$ , and one endpoint in  $U$  and the other in  $W$ . Furthermore,  $\text{Base}(F)$  is the set of edges in the complete graph on  $V = V(G)$  that form a triangle with two edges in  $F$ , formally:

$$\text{Base}(F) = \{uv : uv, vw \in F \text{ for some } w \in V\}.$$

### 6.1 The Graph Family $\mathcal{G}$

Given a constant  $\alpha$ , a sequence  $c/\sqrt{n} \leq p = p(n) \leq C/\sqrt{n}$  and an integer  $\nu \geq 5$ , let  $q$  and  $\lambda$  be as in The Glossary and let  $a = a(\lambda, c)$  be a constant determined by Lemma 2.3. We now define the property  $\mathcal{G}$  and show that it holds for  $G(n, p)$  asymptotically almost surely (a.a.s. for short), that is, with probability tending to 1 as  $n \rightarrow \infty$ .

**Definition 6.1** A graph  $G = (V, E)$  on  $n$  vertices is said to have property  $\mathcal{G} = \mathcal{G}(p, \nu, q, \lambda, a)$  if the following hold:

- (P1) The number of sets of  $\nu$  vertices that span an edge (i.e. are not independent sets) is  $o(n^\nu)$ .
- (P2) The number of sets of  $\nu$  vertices such that some three of them share a common neighbor is  $o(n^\nu)$ .
- (P3) For all  $v \in V$ ,

$$(1 - n^{-1/5})np \leq \deg(v) \leq (1 + n^{-1/5})np \leq 2np.$$

(P4) For all  $v, u \in V$   $\text{codeg}(v, u) \leq 3 \log n$ .

(P5) For all pairs of disjoint sets  $U, W \subset V$  with  $|U|, |W| \geq n/\log n$ ,

$$(1 - n^{-1/5})|U||W|p \leq e_G(U, W) \leq (1 + n^{-1/5})|U||W|p \leq 2|U||W|p,$$

and

$$e_G(U) < |U|^2 p.$$

(P6) For all  $1 \leq k \leq 2q$ , where  $q$  is as in the Glossary, and all choices of subsets  $S_1, \dots, S_k$  of  $V$ , such that  $|S_i| = s_i \leq 2np$ , for all  $i = 1, \dots, k$ , and  $|S_i \cap S_j| \leq 12 \log n$  for all  $1 \leq i < j \leq k$ , the number

$$Z = Z(S_1, \dots, S_k)$$

of vertices with at least one neighbor in each  $S_i$  satisfies

$$|Z - \mu| \leq n^{4/5},$$

where  $\mu = n \prod_{i=1}^k \mu_i$  and  $\mu_i = 1 - (1 - p)^{s_i}$ .

(P7)  $G$  has property  $\mathcal{T}(\lambda, a)$ , that is for any subgraph  $F$  of  $G$  with at least  $\lambda|E|$  edges, the set  $\text{Base}(F)$  contains at least  $a|V|^3$  triangles.

**Lemma 6.1** *For all sequences  $p = p(n)$  such that  $c/\sqrt{n} \leq p \leq C/\sqrt{n}$*

$$\lim_{n \rightarrow \infty} \Pr(G(n, p) \in \mathcal{G}) = 1.$$

**Proof:** We will prove that each of the above properties holds a.a.s., so, of course, their conjunction also holds a.a.s. The proofs of (P1,P2) rely on Markov's inequality, whereas those of (P3-P6) – on various versions of the Chernoff bound for which we refer the reader to Section 2.1 of [16]. Recall that the set of neighbors of  $v$  in  $G$  is denoted by  $N_G(v) = N(v)$ , while  $N_G(W)$  stands for the set of vertices *outside*  $W$ , each having at least one neighbor in  $W$ . Setting  $G = G(n, p)$ , note that  $|N_G(W)|$  is binomially distributed with expectation  $(n - |W|)(1 - (1 - p)^{|W|})$ .

(P1) The expected number of such sets is  $O(n^\nu p)$ . Thus, by Markov's inequality there are a.a.s. no more than, say,  $n^\nu \sqrt{p} = o(n^\nu)$  of them.

**(P2)** The expected number of such sets is  $O(n^{\nu+1}p^3)$ . Thus, by Markov's inequality there are a.a.s. no more than, say,  $n^{\nu+1}p^{5/2} = o(n^\nu)$  of them.

**(P3)** For any  $v \in V$ , the degree  $\deg(v)$  is a binomial random variable with expectation  $(n-1)p$ . Since  $c\sqrt{n} \leq np \leq C\sqrt{n}$ , the usual Chernoff bound ((2.9) in [16]) gives that

$$\Pr(\deg(v) \notin [(1-n^{-1/5})np, (1+n^{-1/5})np]) \leq \exp\{-\Theta(n^{1/10})\}.$$

A simple union bound shows that a.a.s.  $(1-n^{-1/5})np \leq \deg(v) \leq (1+n^{-1/5})np$  for all vertices  $v$  simultaneously.

**(P4)** As above, for fixed  $u, v$  the random variable in question,  $\text{codeg}(v, u)$ , is binomial with expectation  $(n-2)p^2 = \Theta(1)$ . From a handy version of Chernoff's bound ((2.11) in [16]) it follows that

$$\Pr(\text{codeg}(u, v) \geq 3 \log n) \leq e^{-3 \log n} = n^{-3}.$$

Hence, a.a.s. the required condition holds for all  $\binom{n}{2}$  pairs  $u, v$  simultaneously.

**(P5)** Again, given two disjoint sets of vertices  $U, W \subset V$  with  $|U|, |W| \geq n/\log n$ , the number  $e_G(U, W)$  is a binomial random variable with expectation  $|U||W|p = \Omega(n^{3/2}/\log^2 n)$ . And, again, by the usual Chernoff bound,

$$\Pr(d_p(U, W) \notin [(1-n^{-1/5}), (1+n^{-1/5})]) \leq \exp\{-\Theta(n^{12/11})\}.$$

Since there are less than  $2^{2n}$  choices of  $U, W$ , the required condition holds a.a.s. for all such pairs simultaneously. The second statement is proved similarly.

**(P6)** Split  $Z = Z' + Z''$ , where  $Z'$  is the number of vertices  $v$  counted by  $Z$  such that  $v \notin \bigcup_{i=1}^k S_i$  and  $\deg(v, S_i \cap S_j) = 0$  for all  $i < j$ . Set

$$\mu' = \mathbf{E}Z', \quad \mu'' = \mathbf{E}Z'' \quad \text{and} \quad \bar{\mu} = \mathbf{E}Z = \mu' + \mu''.$$

First, by the crude Chernoff bound ((2.12) in [16]) we have

$$\Pr(|Z' - \mu'| \geq n^{4/5}/2) \leq 2e^{-2n^{8/5}/(n-s)} \leq 2e^{-2n^{3/5}},$$

where  $s = \sum s_i$ . Observe also that there are at most

$$\left[2np \binom{n}{2np}\right]^{2q} \leq \left[4C\sqrt{n} \binom{n}{4C\sqrt{n}}\right]^{2q} = \exp\{O(\sqrt{n} \log n)\}$$

choices of  $S_1, \dots, S_k$ . Hence

$$|Z' - \mu'| \leq n^{4/5}/2$$

holds a.a.s for all such choices.

Second, by (P3) we have  $\mu'' = \mathbf{E}Z'' = O((\log n)\sqrt{n})$ . Let  $U_{ij} = S_i \cap S_j$ , for  $i, j = 1, \dots, k$ . Observe that if  $Z'' > n^{2/3}$  then for some  $i \neq j$ , we have

$$\left| \bigcup_{v \in U_{ij}} N_G(v) \right| > (n^{2/3} - s)/k^2 > n^{5/9}.$$

Consequently, by the handy Chernoff bound ((2.11) in [16])

$$\begin{aligned} P(\exists S_1, \dots, S_k : Z'' > n^{2/3}) \\ \leq P(\exists U \subset V, |U| \leq 12 \log n : |N_G(U)| > n^{5/9}) \leq n^{12 \log n} e^{-n^{5/9}} = o(1). \end{aligned}$$

Therefore, a.a.s. for all choices of  $S_1, \dots, S_k$ ,  $1 \leq k \leq 2q$ , we have

$$|Z - \bar{\mu}| \leq |Z' - \mu'| + Z'' + \mu'' \leq n^{4/5}/2 + n^{5/9} + O((\log n)\sqrt{n}) < 2n^{4/5}/3.$$

Finally note that  $|\mu - \mu'| \leq s = O(\sqrt{n})$ .

**(P7)** This is Lemma 2.3, proved in the next section.

We have thus completed the proof that  $G(n, p) \in \mathcal{G}$  a.a.s. □

## 6.2 Proof of Lemma 2.3

Here we prove Lemma 2.3, which turned out to be both difficult and interesting. Work on it led to a separate paper [9] where its extension is utilized in a wider context of dynamic Ramsey-type colorings of random graphs. The proof provided in [9] is based on a counting technique for sparse random graphs. In this paper we prefer our original approach which, besides a regularity lemma, uses an upper tail estimate established in [31]. We state it here in a general setting of random subsets.

Given a finite set  $\Gamma$  and  $0 < p < 1$ , we denote by  $\Gamma_p$  a random binomial subset of  $\Gamma$ . Note that  $([n]^2)_p$  is nothing else but a random graph  $G(n, p)$ . Let  $\mathcal{H} \subseteq [\Gamma]^h$ , where  $h$  is a positive integer, and set  $\mu = \mathbf{E}|\{H \in \mathcal{H} : H \subset \Gamma_p\}|$ .

**Lemma 6.2** *For all integers  $\beta > 0$ , with probability at least  $1 - \exp\{-\frac{\beta}{2h}\}$  there is  $E_0 \subset \Gamma_p$  with  $|E_0| = \beta$ , such that  $\Gamma_p \setminus E_0$  contains fewer than  $2\mu$  sets from  $\mathcal{H}$ .*

For more about upper tail estimates see Section 2.6 in [16], [18] and [17], and for further development of the above deletion technique see [19].

**Corollary 6.3** *A.a.s. for each pair  $U, W \subset [n]$ ,  $U \cap W = \emptyset$ , there exists  $E_0 \subset G(n, p)$  with  $|E_0| = n \log n$ , such that the bipartite subgraph of  $G(n, p) \setminus E_0$  spanned between the sets  $U$  and  $W$  contains no more than*

$$2 \binom{|U|}{2} \binom{|W|}{2} p^4$$

*copies of the 4-cycle  $C_4$ .*

**Proof:** Apply Lemma 6.2 to  $\Gamma = U \times W$  and  $\mathcal{H}$  – the family of the edge sets of all 4-cycles in  $\Gamma$ , with  $s = 4$ ,  $\beta = n \log n$  and  $\mu = \binom{|U|}{2} \binom{|W|}{2} p^4$ .

□

### Proof of Lemma 2.3

Given  $c$  and  $\lambda$ , let  $d = \min\{\lambda^3 c^2 / 270, \lambda^4 / 729\}$ , and let  $\varrho = \varrho(d)$  and  $c_0 = c_0(d)$  be determined via Proposition 4.4 (applied with  $k = 3$ ). Further, let  $\epsilon = \min\{\varrho, \lambda / 100\}$ , and let  $T_0 = T_0(\epsilon)$  be determined by Lemma 4.10 with  $D = 2$ ,  $t_0 = 3$  (say),  $r = 1$ , and the above  $\epsilon$  as inputs. We promise to prove Lemma 2.3 with  $a = c_0 / T_0^3$ .

Let  $G$  be a graph from the space  $G(n, p)$  which satisfies Properties (P3) and (P5) and the property stated in Corollary 6.3. Note that Property (P5) implies that  $G$ , as well as each of its subgraphs has no  $(2, p)$ -dense patches. Fix  $F \subseteq G$  with  $|F| > \lambda|G| > (\lambda/3)pn^2$  and apply Lemma 4.10 to  $F$  with  $D$ ,  $t_0$ ,  $r$ , and  $\epsilon$  as above, obtaining a partition  $V = W_1 \cup \dots \cup W_t$  with  $3 \leq t \leq T_0$  such that all but  $\epsilon n^2 p$  edges of  $F$  belong to the  $(\epsilon, p)$ -regular pairs  $\{W_i, W_j\}$ . Let  $F'$  be the subgraph of  $F$  consisting of those edges, and let  $x \binom{t}{2}$  be the number of pairs  $(i, j)$  such that  $e_{F'}(W_i, W_j) \geq \frac{1}{3} \lambda p (n/t)^2$ . For clarity of presentation, let us assume that  $t$  divides  $n$ , and so  $|W_i| = n/t$  for each  $i = 1, \dots, t$  (cf. Remark 5.1). We will now estimate  $x$  from below.

By (P5) and the definition of  $\epsilon$ ,

$$(\lambda/3)pn^2 \leq |F'| \leq 2x \binom{t}{2} (n/t)^2 p + \left( \binom{t}{2} - x \binom{t}{2} \right) (\lambda p/3)(n/t)^2$$

and so

$$\lambda \leq x + (1-x)\lambda/4 .$$

Solving the last inequality for  $x$ , we obtain  $x \geq \lambda/6$ . Hence, at least

$$(\lambda/6) \binom{t}{2} > (\lambda/10)t^2$$

pairs  $(W_i, W_j)$  are  $(\epsilon, p)$ -regular, each satisfying

$$e_F(W_i, W_j) \geq (\lambda p/3)(n/t)^2 .$$

By simple averaging, there must be an index  $i_0$  and a set  $J \subseteq [t] \setminus \{i_0\}$  of cardinality  $|J| = \lambda t/10$  such that  $(W_{i_0}, W_j)$  is as above. Set  $W = W_{i_0}$  and  $U = \bigcup_{j \in J} W_j$ , and note that  $|W| = n/t$  and

$$|U| = |J|(n/t) = \lambda n/10 .$$

Let  $B$  be the bipartite subgraph of  $F$  induced by all edges with one endpoint in  $U$  and the other in  $W$ , shortly  $B = F[U, W]$ . Then  $\text{Base}(B)[W] \subseteq \text{Base}(F)$ . We will show that  $\text{Base}(B)[W]$  is  $(\varrho, d)$ -dense, and hence, by Lemma 4.4, for  $n$  large enough, contains at least  $c_0(n/t)^3 \geq (c_0/T_0^3)n^3$  triangles, which will complete the proof.

To this end, assume for clarity that  $\varrho n/t$  is an integer, and pick any  $W' \subset W$  with  $|W'| = \varrho|W| = \varrho n/t$ . Our ultimate goal is to show that

$$N = |\text{Base}(B)[W']| \geq d \binom{|W'|}{2} .$$

Since for each  $j \in J$  the pair  $(W_j, W)$  is  $(\epsilon, p)$ -regular with density at least

$$\pi = \lambda p/3$$

and  $|W'| \geq \epsilon|W|$ , by Proposition 4.7, all but at most  $\epsilon|W_j|$  vertices of  $W_j$  have each at least  $(\lambda/3 - \epsilon)p|W'| = (1 - \epsilon')\pi|W'|$  neighbors in  $W'$ , where

$\epsilon' = 3\epsilon/\lambda > \epsilon$ . Consequently, all but at most  $\epsilon|U|$  vertices of  $U$  have each at least  $(1 - \epsilon')\pi|W'|$  neighbors in  $W'$ .

Let  $E_0 \subset G[U, W']$  be as in Corollary 6.3, i.e.  $|E_0| = n \log n$  and there are at most  $2\binom{|U|}{2}\binom{|W'|}{2}p^4$  copies of  $C_4$  in  $G_0 = G[U, W'] \setminus E_0$ . Clearly, the same is true for the subgraph  $B_0 = F[U, W'] \setminus E_0$  of  $G_0$ . As, say, only at most  $n^{2/3}$  vertices of  $U$  can be each incident to more than  $n^{1/3} \log n$  edges of  $E_0$ , still, say, all but at most  $2\epsilon|U|$  vertices of  $U$  have each at least  $(1 - 2\epsilon')\pi|W'|$  neighbors in  $W'$ . Put in another way, for all but at most  $2\epsilon|U|$  vertices  $u \in U$

$$d_u = \deg_{B_0}(u) > (1 - 2\epsilon')\pi|W'|.$$

Let  $[W']^2 = \{e_1, \dots, e_{\binom{|W'|}{2}}\}$ . Denote by  $x_i$  the number of vertices  $u \in U$  which are adjacent in  $B_0$  to both elements of the pair  $e_i$ , that is which are the tips of the tepees of  $B_0$  over  $e_i$ . Clearly,  $N = |\{i : x_i > 0\}|$ . For convenience, assume that  $x_i > 0$  for  $i = 1, \dots, N$ . Observe that for sufficiently large  $n$ , by the choice of  $\epsilon$ ,

$$\sum_{i=1}^N x_i = \sum_{u \in U} \binom{d_u}{2} > \frac{1}{2}(1 - O(n^{-1/2}))(1 - 2\epsilon')^3|U|\pi^2|W'|^2 > \frac{1}{3}\pi^2|U||W'|^2,$$

and that  $\sum_{i=1}^N \binom{x_i}{2}$  equals the number of copies of  $C_4$  in  $B_0$ . Consequently,

$$\sum_{i=1}^N \binom{x_i}{2} \leq 2\binom{|U|}{2}\binom{|W'|}{2}p^4 < \frac{1}{2}p^4|U|^2|W'|^2.$$

By the Cauchy-Schwarz inequality,

$$\sum_{i=1}^N \binom{x_i}{2} \geq N \binom{\sum x_i / N}{2} = \frac{\sum x_i}{2} \left( \frac{\sum x_i}{N} - 1 \right). \quad (34)$$

Consider two cases:

I.  $N \geq \sum x_i / 2$ . Then

$$N \geq \frac{1}{6}|U|\pi^2|W'|^2 \geq \frac{1}{3} \times \frac{\lambda n}{10} \times \frac{\lambda^2 p^2}{9} \binom{|W'|}{2} \geq d \binom{|W'|}{2},$$

as required.

II.  $N \leq \sum x_i/2$ . Then, by (34),

$$N \geq \frac{(\sum x_i)^2}{4 \sum \binom{x_i}{2}} \geq \frac{(1/9)\pi^4|U|^2|W'|^4}{2p^4|U|^2|W'|^2} \geq d \binom{|W'|}{2},$$

as needed. □

## 7 Summary, Further Remarks, Glossary

When embarking on the journey of solving this problem we did not realize the range of techniques which would be required, nor did we predict the extent of technical difficulties involved. Already one paper, [9], has sprung as a side effect, solving a problem which arose during our work, and addressing themes related to it. We hope in the future to follow the various trails that we have encountered here.

There are several directions in which the results of this paper may be generalized or extended. We state below the corresponding open problems. It seems that in order to solve some of them one will need to develop a better understanding of sparse regular graphs and “online” Ramsey theory as described in [9].

The natural questions that come to mind are with regard to generalizing the main theorem of this paper to the case of coloring with more than two colors, or the case of Ramsey properties where the defining forbidden monochromatic graph is not a triangle but some other graph. We conjecture that an analog of our main theorem in this paper holds for all these cases, that is, that for all such natural Ramsey properties there exists a sharp threshold. These problems seem to be within the grasp of the techniques of this paper but present some serious technical difficulties. It is our hope that we will be able to overcome these difficulties in future work.

It is also of interest to study thresholds of Ramsey properties of random sets of integers, such as having a monochromatic arithmetic progression in every bi-coloring of the set (see[31, 21, 32]).

A different question which seems at the moment quite hard is whether one can improve the result in this paper by establishing the exact threshold probability for the Ramsey property, or even proving that it exists:

**Question 7.1** *Does the function  $\widehat{c}(n)$  defined in Theorem 1.1 tend to a limit as  $n$  tends to infinity? If so what is the exact value of the constant it tends to?*

It is worthwhile noting that this natural question has not been answered in any of the cases where the existence of a sharp threshold has been established using the techniques from [8] (i.e. in [8], [1] and [10]).

## 7.1 Acknowledgments

The authors wish to thank the anonymous referee for his/her serious and professional effort, and the resulting suggestions.

The work on this paper was supported by the following grants:

- Ehud Friedgut: Research supported in part by the Israel Science Foundation grant no. 0329745, and by the Binational Science Fund grant 2002166.
- Vojtech Rödl: Research supported by NSF grant NSF Award DMS - 0300529, and by the Binational Science Fund grant 2002166.
- Andrzej Ruciński: Research supported by KBN grant 2 P03A 015 23; part of this research performed during several visits at Emory University
- Prasad Tetali: Research supported in part by NSF grants DMS-0100298 and DMS-0401239.

## 7.2 Glossary

Any fan of 19'th century Russian literature is familiar with the frustration of reading a novel with a huge number of characters, each having two or three names and nicknames. The timid western reader often finds comfort in a short list added by the editors to the translation, a list of all characters, each accompanied by a reminder of their family background. We present here such a list that may help the reader of this paper, who may refer back to it while reading. We also define here the appropriate choices to ensure the needed relations between various constants. We end the glossary with a diagram which gives the dependency between the constants.

## Constants

- $c, C$  are absolute constants given by Theorem 1.2; we have  $c = c_2 = 1/e$  and  $C = C_2 = 10^4$ ; their role is to bound the threshold probability  $p(n)$  scaled by  $\sqrt{n}$  (see Theorem 2.1);
- $\xi > 0$  and  $0 < \alpha < 1$  are arbitrary constants at the outset of Theorem 2.1;
- $\nu$  is the number of vertices of  $M$ , an arbitrary balanced graph with average degree 4 at the outset of Theorem 2.1;
- $q = \lceil 10C^2\nu/\alpha \rceil$  is an upper bound on the number  $\phi$  of edges in the star forest  $S$ , a prototype of constellations; used to define Properties (P6) and (P7) of  $\mathcal{G}$ ;
- $\alpha_1$  – no such thing;
- $\alpha_2 = \frac{\alpha}{(2\nu + q)^q}$  is an intermediate constant in Section 3;
- $\alpha_3 = \frac{\alpha_2}{2(\nu + q)^{(\nu+q)}}$  – there are at least  $\alpha_3 n^\nu$  special constellations in each  $G \in \mathcal{G}$ ;
- $\alpha_4 = \frac{(\nu + q)^{\nu+q}}{3C^{2q}} \alpha_3 = \frac{\alpha}{6C^{2q}(2\nu + q)^q}$  is an intermediate constant in Section 5;
- $D$  is a parameter bounding from above the  $p$ -density of patches (large subgraphs) of  $G$  (see Definition 4.5) which is a crucial assumption in all sparse regularity lemmas; in our application of these lemmas, owing to Property (P5) of  $\mathcal{G}$ , we take  $D = 2$ ;
- $D^*$  is a parameter bounding from above the  $(S, p^*)$ -density of  $S$ -patches of  $F$  (see Definition 4.6) which is a crucial assumption in the Subgraph Regularity Lemma 4.13; in our application of this lemma we take  $D^* = 2^q + 1$  (see Lemma 4.17);
- $\lambda = \frac{1}{160} \left( \frac{\alpha_4}{D^* q (\nu + q)} \right)^2$  – every core has size at least  $\lambda p n^2$ ;

- $a = a(\lambda)$  is a constant determined by Lemma 2.3; both,  $\lambda$  and  $a$  are used to define Property (P7) of  $\mathcal{G}$ ;
- $\tau_0 = \frac{a^2 \xi^6 c^6}{2(a \xi^3 c^3 + 2 \xi^5 C^5)}$  is a value of  $\tau$  with which we apply Theorem 2.1;
- $\tau > 0$  is an arbitrary constant at the outset of Theorem 2.1;
- $\eta > 0$  is chosen so that  $\eta \log(1/\eta) \leq \frac{\tau}{4C\nu}$ ; this parameter serves to define cores and the above bound guarantees that there are few of them;
- $\delta = \min \left\{ 0.4\eta^q, \left( \frac{\alpha_4}{8D^*} \right)^2 \right\}$ , is an input of the Subgraph Regularity Lemma 4.13; the final partition  $\Pi_1$  resulting from Lemma 4.13 with  $H$  being any star forest with at most  $q$  edges, is guaranteed to be  $\delta$ -regular;
- $\varepsilon(\ell) = \min \left\{ \ell^{-4q}, \left( \frac{\delta \alpha_4}{D^* 2^{q+3} q (\nu + q)^{\nu+q-2}} \right)^{4q}, q^{-4}, 2^{-6-6q} \right\}$ , is a function of variable  $\ell$ , at the outset of the Subgraph Regularity Lemma 4.13;
- $T_0, L_0$  are output constants of the Subgraph Regularity Lemma 4.13;
- $t, l$  are not really constants but vary with  $G$ ; the final partition  $\Pi_1$  resulting from the Subgraph Regularity Lemma 4.13 with  $H$  being any star forest with at most  $q$  edges, is a  $(t, l)$ -partition for some  $t \leq T_0$  and  $l \leq L_0$ ; do not confuse  $l$  with  $\ell$  – the argument of the function  $\varepsilon(\ell)$  defined above.
- $\varepsilon = \varepsilon(l)$  – the final partition  $\Pi_1$  resulting from the Subgraph Regularity Lemma 4.13 with  $H$  being any star forest with at most  $q$  edges, is guaranteed to be  $(\varepsilon, p)$ -uniform; also, this  $\varepsilon$  satisfies the assumptions of Proposition 4.8 with  $D = 2$  and  $k = q$ ;
- $n_0$  is a lower bound on  $n$ , different at different places in the paper, in particular it is an output constant of the Subgraph Regularity Lemma 4.13;

- $n_1$  is the lower bound on  $n$  promised in Theorem 2.1; it is the maximum of all values of  $n_0$  encountered throughout our proof, most notably the  $n_0$  of Lemma 4.13, as well as of several implicit lower bounds on  $n$  hidden in our calculations;

## Graphs

- $\rho(H) = \frac{|E|}{|V|}$  is half of the average degree in a graph  $H = (V, E)$ .
- $\mathcal{R}$  is the family of all graphs such that every bi-coloring of their edges results in a monochromatic triangle; they are referred to as Ramsey graphs;
- $\mathcal{G}$  is a family of graphs which depends on  $c, C, \alpha, \nu, p = p(n)$  and is defined in Definition 6.1; it is crucial for our proof that  $G(n, p) \in \mathcal{G}$  almost surely (see Theorem 2.1 and Lemma 6.1);
- $\mathcal{T}(\lambda, a)$  is a family of graphs  $G = (V, E)$  such that for any subgraph  $F$  of  $G$  with at least  $\lambda|E|$  edges, the set  $\text{Base}(F)$  contains at least  $a|V|^3$  triangles (for the definition of  $\text{Base}(F)$  see pages 15 or 82); we have  $\mathcal{G} \subseteq \mathcal{T}(\lambda, a)$  with  $\lambda$  as above;
- $M$  is an arbitrary, balanced graph with  $\nu$  vertices and  $2\nu$  edges, at the outset of Theorem 2.1; we put  $V(M) = \{x_1, \dots, x_\nu\}$ ;
- $G = (V, E)$  is an arbitrary member of  $\mathcal{G}$  with more than  $n_0$  vertices; it is assumed that  $G \notin \mathcal{R}$ , but  $G$  satisfies (3); the whole proof boils down to showing that it also satisfies (4);
- $M_X$  is an ordered copy of  $M$  with vertex set  $X = \{v_1, \dots, v_\nu\} \subset V$  and such that  $x_a$  is mapped onto  $v_a$ , for each  $a = 1, \dots, \nu$ ;
- $M^*$  is a random copy  $M_X$ , where  $X$  is chosen uniformly over all  $(n)_\nu$  possibilities;
- $M'_X$  is  $M_X$  with edges colored by a fixed coloring  $\sigma'$ ;
- $T(X)$  is the set of all tepees in  $G$  formed over all pairs of vertices from  $X$  which are edges of  $M_X$  (see Definition 3.1 for the definition of a tepee);

- $\widehat{X}$  is the subgraph of  $G$  formed by the edges of  $T(X)$ ;
- $\widehat{M}$  is a common isomorphic type of all  $M_X$  with  $X \in \mathcal{X}_2$ ; it has  $\nu + \phi$  vertices and  $2\phi$  edges,  $1 \leq \phi \leq q$ , that is, it consists of  $\phi$  tepees over  $M$ ;
- $S(X)$ , for  $X \in \mathcal{X}_3$ , called a special constellation, is a subgraph of  $\widehat{X}$  consisting of the left legs of all tepees from  $T(X)$  (see Section 3.2);
- $\mathcal{S}$  is the set of all special constellations;
- $S$  is the common isomorphism type of all  $S(X)$  with  $X \in \mathcal{X}_3$ ;  $S$  is a forest of  $\nu$  stars;
- $\phi$  is the number of edges of the star forest  $S$ ; the stars forming  $S$  have  $\phi_1, \dots, \phi_\nu$  edges, where  $\phi = \phi_1 + \dots + \phi_\nu$ ;
- $\pi_0 = \{V_1, \dots, V_\nu, W_{11}, \dots, W_{1\phi_1}, \dots, W_{\nu,1}, \dots, W_{\nu,\phi_\nu}\}$ , is an initial equipartition of  $V$ ; for clarity, we further assume that all the sets are of equal size  $m = n/(\nu + \phi)$ ;
- $\mathcal{C}$  is the set of all constellations, that is copies of  $S$  which are consistent with  $\pi_0$ ;
- $F_{ab} = G[V_a, W_{ab}]$ ,  $a = 1, \dots, \nu$ ,  $b = 1, \dots, \phi_a$ , is the bipartite subgraph of  $G$  consisting of all edges with one endpoint in  $V_a$  and the other in  $W_{ab}$ ;
- $F = \bigcup_{ab} F_{ab}$  is the subgraph of  $G$  consisting of all edges connecting sets  $V_a$  with  $W_{ab}$ ; in Section 4,  $F$  also stands for an arbitrary member of  $\mathcal{F}(H, m)$ ;
- $H$  always stands for an arbitrary fixed graph with  $h$  vertices, mostly in Section 4, where the general form of the Subgraph Regularity Lemma 4.13 is proved; it is later applied with  $H = S$ ;
- $\mathcal{F}(H, m)$  denotes the family of all spanning subgraphs of  $H^m$ , where  $H^m$  is the  $h$ -partite graph obtained by replacing each vertex  $x$  of  $H$  by a set  $V_x$  of  $m$  vertices, and by replacing every edge  $xy$  of  $H$  by the complete bipartite graph  $K_{m,m}$  spanned between  $V_x$  and  $V_y$ ;

- $\#(H, F)$  is our notation for the number of copies of  $H$  in  $F$ , sometimes it only counts copies which are consistent with a specified partition of  $V(F)$ ;
- $\Pi_0$  is an initial partition of  $F$  consisting of  $\pi_0$  and the (trivial partitions of) subgraphs  $F_{ab}$ ;
- $\Pi_1$  is a final partition of  $F$  resulting from the Subgraph Regularity Lemma 4.13; it consists of vertex equipartitions  $V_a = V_a^1 \cup \dots \cup V_a^t$  and  $W_{ab} = W_{ab}^1 \cup \dots \cup W_{ab}^t$  (under a simplifying assumption, all these sets are of equal size  $m/t$ ), and subgraph partitions  $F_{ab}^{i,j} = G[V_a^i, W_{ab}^j] = \bigcup_k F_{ab}^{i,j,k}$ ; note that  $F = \bigcup_{a,b,i,j,k} F_{ab}^{i,j,k}$ ;
- $\Pi$  stands for an arbitrary (or current)  $(t, l)$ -partition (see Definition 4.8);
- $\Psi$  stands for a refinement of  $\Pi$  (see Definition 4.8);
- $P$  always stands for a polyad – a special kind of a subgraph of  $F$ , always related to (consistent with) a partition  $\Pi$  (see Definition 4.9); in Section 5 all polyads are consistent with the final partition  $\Pi_1$ ;
- $\mathcal{P}_\Pi = \mathcal{P}$  is the set of all polyads consistent with a partition  $\Pi$ .
- $R$ , generally, is a subgraph of  $F$ , often a subgraph of a given polyad  $P$ ; in the proof of Lemma 4.13,  $R \subset P$  represents polyads which are consistent with a partition  $\Psi$  which refines  $\Pi$ ;

## Quantities depending on $n$

- $n > n_0$  is the number of vertices of  $G$ ;
- $m = n/(\nu + \phi)$  is the number of vertices in each part of the partition  $\pi_0$  of the vertices of  $G$  (we assume this is an integer);
- $p = p(n)$  is a sequence of edge probabilities satisfying  $c < p\sqrt{n} < C$ ; it is also used for an abstract scaling factor in the formula for  $p$ -density (see Section 4);
- $e_G(U, W)$  is the number of edges of a graph  $G$  with one endpoint in  $U$  and the other in  $W$ ; here  $U$  and  $W$  are disjoint subgraphs of  $V$ ;

- $d(B) = d(U, V) = \frac{e_G(U, V)}{|U||V|}$  is the density of the pair  $(U, V)$ ; also called the density of the bipartite graph  $B$ , where  $B = G[U, V]$ ;
- $d_p(B) = d_p(U, V) = \frac{e_G(U, V)}{p|U||V|}$  is, for  $0 < p \leq 1$ , the  $p$ -density of the pair  $(U, V)$ ; also called the  $p$ -density of the bipartite graph  $B$ , where  $B = G[U, V]$ ;
- $p^*$  is an abstract normalizing factor for the  $(H, p^*)$ -density  $d_R$  of subgraphs  $R$  of  $F$ ; for our application with  $H = S$  we take  $p^* = p^\phi$ ;
- $\kappa = |C|$  is the total number of constellations (consistent with  $\pi_0$  copies of  $S$ ) in  $F$ ; we have  $\kappa \sim m^{\nu+\phi} p^\phi$ ;
- $c_R = |C(R)|$ , where  $C(R)$  is the set of constellations in  $R$ ;
- $s_R = |S(R)|$ , where  $S(R)$  is the set of special constellations in  $R$ ;
- $d_R = \frac{s_R}{p^* c_R}$  is the (normalized)  $(H, p^*)$ -density of  $R$ ; for our application with  $H = S$  we take  $p^* = p^\phi$ ;

## The flowchart of constants

For the convenience of the reader we have included, in the following a chart indicating the interdependencies of various constants that we have used in our proofs.

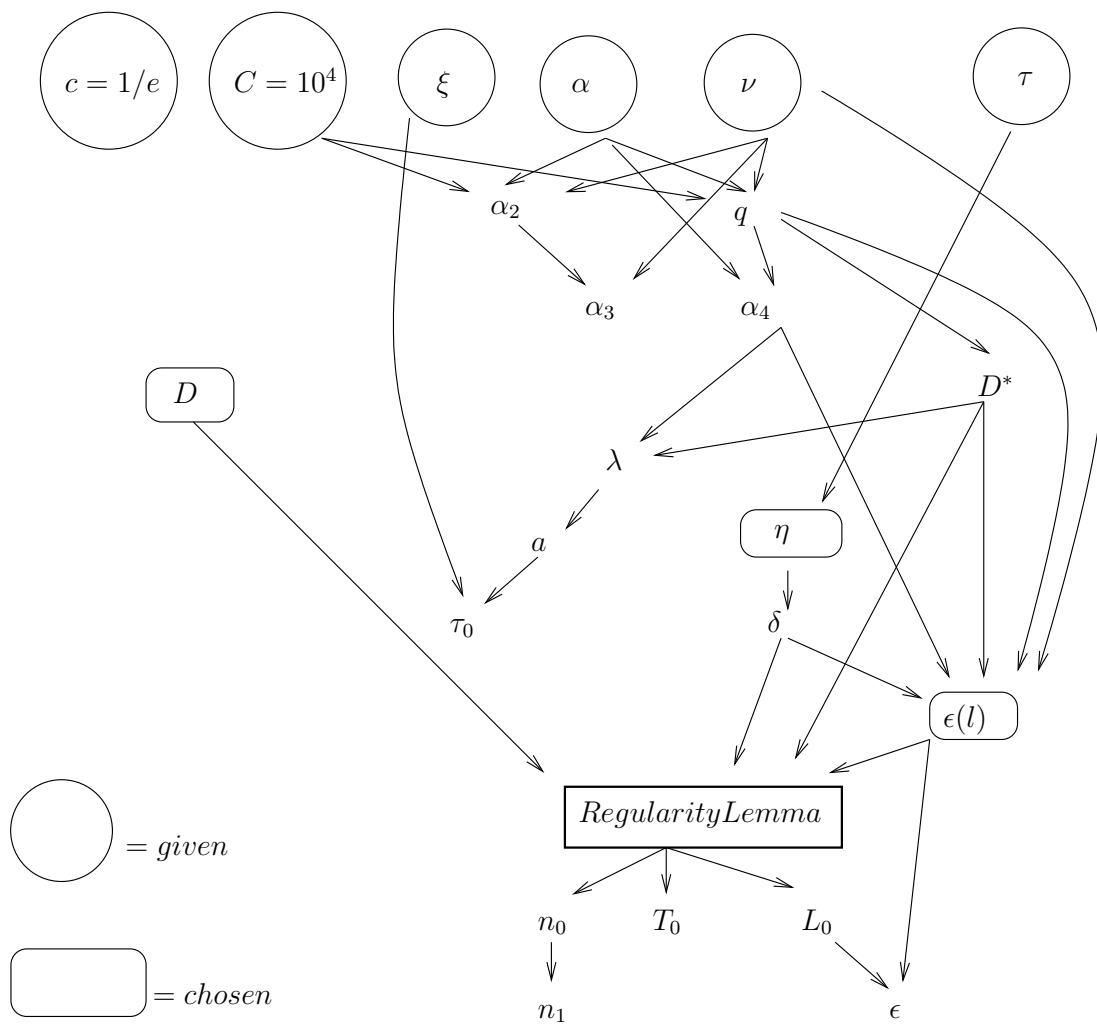


Figure 9: The flowchart of constants

## References

- [1] D. Achlioptas, E. Friedgut, A sharp threshold for  $k$ -colorability. *Random Structures and Algorithms*, 14 (1999), 63–70.
- [2] N. Alon, R. A. Duke, H. Lefmann, V. Rödl and R. Yuster, The algorithmic aspects of the regularity lemma, *J. Algorithms* 16(1) (1994), 80–109,
- [3] B. Bollobás, *Random Graphs*, Academic Press, London, 1985.
- [4] B. Bollobás and A. Thomason, Threshold functions, *Combinatorica* 7 (1987), 35–38.
- [5] R. Duke, H. Lefmann and V. Rödl, A fast approximation algorithm for computing the frequencies of subgraphs in a given graph, *SIAM J. Computing*, **24** (1995), 598–620.
- [6] P. Erdős and A. Rényi, On random graphs I, *Publ. Math. Debrecen* 6 (1959), 290–297.
- [7] P. Erdős and A. Rényi, On the evolution of random graphs, *Publ. Math. Inst. Hung. Acad. Sci.* 5 (1960), 17–61.
- [8] E. Friedgut, Sharp thresholds of graph properties, and the  $k$ -sat problem. *Journal of the American Math. Soc.*, 12 (1999), no. 4, 1017–1054.
- [9] E. Friedgut, Y. Kohayakawa, V. Rödl, A. Ruciński and P. Tetali, Ramsey games against one-armed bandit, *Combinatorics, Probability and Computing* (2003)
- [10] E. Friedgut and M. Krivelevich, Sharp Thresholds for Ramsey Properties of Random Graphs, *Random Structure and Algorithms* 17(1) (2000), 1–19.
- [11] P. Frankl and V. Rödl, Large triangle-free subgraphs in graphs without  $K_4$ , *Graphs and Combinatorics* 2 (1986), 135–144.
- [12] P. Frankl and V. Rödl, Extremal problems on set systems, *Random Structure and Algorithms* 20(2) (2002), 131–164.

- [13] R. Graham , V. Rödl and A. Ruciński, On Schur properties of random subsets of integers *Journal of Number Theory* 61(2) (1996), 388-408.
- [14] E. Gyóri, B. Rothschild and A. Ruciński, Every graph is contained in a sparsest possible balanced graph, *Math.Proc.Cambr.Phil.Soc.* 98 (1985), 397-401.
- [15] S. Janson, T. Łuczak and A.Ruciński, An exponential bound for the probability of nonexistence of a specified subgraph of a random graph, in: *Proceedings of Random Graphs'87*, Wiley, Chichester (1990), 73-87.
- [16] S. Janson, T. Łuczak and A. Ruciński, *Random Graphs*, Wiley 2000, 333+xi pages.
- [17] S. Janson, K. Oleszkiewicz and A. Ruciński, Upper tails for subgraphs counts in random graphs, *Israel J. Math*, to appear.
- [18] S. Janson and A. Ruciński, The infamous upper bound, *Random Structures Algorithms*, 20(3), (2002) 317-342.
- [19] S. Janson and A. Ruciński, The deletion method for upper tails, *Combinatorica*, to appear.
- [20] Y. Kohayakawa, Szemerédi's regularity lemma for sparse graphs, *Foundations of computational mathematics (Rio de Janeiro, 1997)*, Springer, Berlin, 1997, pp. 216–230.
- [21] Y. Kohayakawa, T. Łuczak and V. Rödl Arithmetic progressions of length three in subsets of a random set, *Acta Arithmetica* LXXV (1996), 133-163
- [22] Y. Kohayakawa and V. Rödl, Regular Pairs in Sparse Random Graphs, *Random Structure and Algorithms*, (2003), 22(4) 359-434.
- [23] Y. Kohayakawa and V. Rödl, Szemerédi's regularity lemma and quasi-randomness, *Recent Advances in Algorithmic Combinatorics (B. Reed and C. Linhares-Sales, eds.)*, CMS Books Math./Ouvrages Math. SMC, vol. 11, Springer, New York, 2003, pp. 289-351

- [24] J. Komlós and M. Simonovits, Szemerédi’s regularity lemma and its applications in graph theory, *Combinatorics, Paul Erdős is eighty*, Vol. 2 (Keszthely, 1993), János Bolyai Math. Soc., Budapest, 1996, pp. 295–352.
- [25] A. Kurek and A. Ruciński, Two variants of the size Ramsey number, submitted
- [26] T. Łuczak, On triangle-free random graphs, *Random Structures & Algorithms* (2000), 16(3), 260-276.
- [27] T. Łuczak, A. Ruciński and B.Voigt, Ramsey properties of random graphs, *J. Comb. Th. Ser. B* 56 (1992), 55–68.
- [28] B. Nagle and V. Rödl, Regularity properties for triple systems, *Random Structures & Algorithms* (2003), 23(3) 264-332.
- [29] V. Rödl and A. Ruciński, Lower bounds on probability thresholds for Ramsey properties, *Combinatorics, Paul Erdős is Eighty* (Vol.1), Keszthely(Hungary), Bolyai Soc. Math. Studies, 1993, pp.317–346.
- [30] V. Rödl and A. Ruciński, Random graphs with monochromatic triangles in every edge coloring, *Random Structures and Algorithms* 5 (1994), 253–270.
- [31] V. Rödl and A. Ruciński, Threshold functions for Ramsey properties, *Journal of Amer. Math. Soc.* Vol 8 (1995), 917–942.
- [32] V. Rödl and A. Ruciński, Rado partition theorem for random subsets of integers *Proceedings of the London Mathematical Society* 74(3) (1997) 481-502.
- [33] V. Rödl and A. Ruciński, Ramsey properties of random hypergraphs, *J. Comb. Th. Ser. A* 81 (1998), 1–33.
- [34] I. Z. Ruzsa and E. Szemerédi, Triple systems with no six points carrying three triangles, *Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976)*, Vol. II, North-Holland, Amsterdam, 1978, pp. 939–945.

- [35] E. Szemerédi, Regular partitions of graphs in *Problèmes combinatoires et théorie des graphes*, Orsay 1976, J.-C. Bermond, J.-C. Fournier, M. Las Vergnas, D. Sotteau, eds., Colloq. Internat. CNRS 260, Paris, 1978, 399–401.

Ehud Friedgut, Institute of Mathematics, Hebrew University, Jerusalem, Israel. Email: `ehudf@math.huji.ac.il`.

Vojtech Rödl, Emory University, Atlanta, Ga. U.S.A. Email: `rod1@mathcs.emory.edu`

Andrzej Ruciński Adam Mickiewicz University, Poznań, Poland. Email: `rucinski@amu.edu.pl`

Prasad Tetali School of Mathematics and College of Computing Georgia Institute of Technology Atlanta, GA 30332-0160, USA. Email: `tetali@math.gatech.edu`