

# **On some problems in extremal hypergraph theory**

Amedeo Sgueglia

A thesis submitted for the degree of  
*Doctor of Philosophy*



Department of Mathematics  
The London School of Economics  
and Political Science

London, October 2022

---

## Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work, with the exceptions outlined below.

The copyright of this thesis rests with the author. Quotation from it is permitted, provided that full acknowledgement is made. This thesis may not be reproduced without my prior written consent. I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that this thesis consists of 56267 words.

## Statement of co-authored work

The contents of Chapter 2 are well-known results of various authors not including myself.

I confirm that Section 1.1.4 and Chapter 3 are based on joint work with Julia Böttcher, Olaf Parczyk, and Jozef Skokan [27].

I confirm that Section 1.1.3 and Chapter 4 are based on joint work with Julia Böttcher, Olaf Parczyk, and Jozef Skokan [28].

I confirm that Sections 1.2.2 and 1.2.3 and Chapter 5 are based on joint work with Pranshu Gupta, Fabian Hamann, Alp Müyesser, and Olaf Parczyk [54].

I confirm that Sections 1.3.2 and 1.3.3 and Chapter 6 are based on joint work with Jury Barkey, Dennis Clemens, Fabian Hamann, and Mirjana Mikalački [12].

---

## Abstract

In this thesis, we study several problems in extremal (hyper)graph theory.

We begin by investigating the problem of subgraph containment in the model of randomly perturbed graphs. In particular, we study the perturbed threshold for the appearance of the square of a Hamilton cycle and the problem of finding pairwise vertex-disjoint triangles. We provide a stability version of these results and we discuss their implications on the perturbed thresholds for 2-universality and for a triangle factor.

We then turn to the notion of threshold in the context of transversals in hypergraph collections. Here the question is, given a fixed  $m$ -edge hypergraph  $F$ , how large the minimum degree of each hypergraph  $H_i$  needs to be, so that the hypergraph collection  $(H_1, \dots, H_m)$  necessarily contains a transversal copy of  $F$ . We prove a widely applicable sufficient condition on  $F$  such that the following holds. The needed minimum degree is asymptotically the same as the minimum degree required for a copy of  $F$  to appear in each  $H_i$ . The condition is general enough to obtain transversal variants of various classical Dirac-type results for (powers of) Hamilton cycles.

Finally, we initiate the study of a new variant of the Maker-Breaker positional game, which we call the  $(1:b)$  multistage Maker-Breaker game. Starting with a given hypergraph, we play several stages of a usual  $(1:b)$  Maker-Breaker game where, in each stage, we shrink the board by keeping only the elements that Maker claimed in the previous stage and updating the collection of winning sets accordingly. The game proceeds until no winning sets remain, and the goal of Maker is to prolong the duration of the game for as many stages as possible. We estimate the maximum duration of the  $(1:b)$  multistage Maker-Breaker game for several standard graph games played on the edge set of  $K_n$  with biases  $b$  subpolynomial in  $n$ .

---

## Acknowledgements

This journey would not have been possible without a number of people who, in a way or another, contributed to it. It is likely that I forgot someone, but I blame the approaching submission deadline for that.

Above all, I would like to thank my supervisors, Julia Böttcher and Jozef Skokan, who went above and beyond their duties. Thank you for your guidance towards exciting research problems, your encouraging words and your sincere enthusiasm. Thank you for creating a friendly atmosphere within the graph theory group at LSE, this has been valuable for me.

I would also like to thank Peter Allen who, despite not being my formal supervisor, has been a third person on which I could always count on.

Special thanks go to Olaf Parczyk for his constant patience in answering all my questions, his energy and passion in discussing maths, and the great time in London and Berlin.

I had the pleasure of being involved in several collaborations with amazing coauthors. They are Alp, Dennis, Fabian, Jozef, Julia, Juri, Mirjana, Olaf and Pranshu. Thank you for our enjoyable meetings and for everything I have learnt from you.

The last four years in London are full of great memories and the main responsible for that is the LSE Maths PhD gang. In particular, I would like to thank Cedric, Domenico, Francesco, Jan, Joanna, Johannes, Justin, Karl, Keat, Mahsa, Nóra, Raymond, Rennie, Sahar, Stan and Xinyi.

I wish to thank all my *Patavini* friends: Alice, Ariane, Elisa, Francesca, Francesco, Laura, Lina, Luca, Michele, Saeda, Sara and Simona. Thank you for your company, your optimism and precious advice, and for bringing serious and deep topics into our conversations (after all, it is only because of you that I now know the word ‘*albedo*’).

And of course, I would like to thank my more dated friends from Caiazzo and surroundings: Alessia Pia, Angela Myriam, Mena, Myriam, and all the *eruditi* Antonella, Antonio, Claudia, Giovanna, Giuseppe, Ilaria, Irene, Luca, Mariangela, Mario, Piero and Serena. Thank you for all the fun time together and our board-games nights. Antonio and Piero, thank you for not complaining too much (anymore, you understood it was hopeless) when I reply to your messages a week later (though, you should admit, I am super efficient if you send e-mails).

I would like to mention my indebtedness to Prof Candida Vetrone and all the other inspiring professors I was lucky to meet over my educational path at school and university: you truly made a difference in my life.

Finally, I would like to express my gratitude to my parents, Egidio and Maria, my brother Arnaldo, my sister Palma, my uncles and cousins, and my grandparents. You deserve much more thanks than I will be ever capable of expressing: I would not have made it so far without you.

---

# Contents

---

<b>Declaration</b>	<b>ii</b>
<b>Abstract</b>	<b>iii</b>
<b>Acknowledgements</b>	<b>iv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Randomly perturbed graphs . . . . .	4
1.1.1 Thresholds . . . . .	5
1.1.2 Overview of known results . . . . .	7
1.1.3 New results: triangles . . . . .	9
Stability . . . . .	10
1.1.4 New results: the square of a Hamilton cycle . . . . .	11
Lower bounds and stability . . . . .	12
2-universality . . . . .	13
1.2 Transversal versions of Dirac-type theorems . . . . .	14
1.2.1 Overview of known results . . . . .	15
1.2.2 Applications . . . . .	17
1.2.3 An introduction to the main theorem . . . . .	17
1.3 Multistage Maker-Breaker game . . . . .	19
1.3.1 Classical $(1: b)$ Maker-Breaker game . . . . .	19
1.3.2 Multistage games . . . . .	21
1.3.3 New results . . . . .	22
1.4 Notation . . . . .	24
<b>2 Methods and tools</b>	<b>28</b>
2.1 The regularity method . . . . .	28
2.2 The absorption method . . . . .	30
2.3 The container method . . . . .	31
2.4 Concentration inequalities . . . . .	31
2.5 Embedding results . . . . .	33
2.6 Preserving minimum degrees . . . . .	36

2.7	<i>K</i> -collections	38
<b>3</b>	<b>The square of a Hamilton cycle in randomly perturbed graphs</b>	<b>40</b>
3.1	Proof overview of the extremal and the non-extremal case	41
3.1.1	Strategy	41
3.1.2	Extremal case	42
3.1.3	Non-extremal case	43
3.2	Proof of the extremal Theorem 1.1.17	45
3.3	Proof of the stability Theorem 1.1.16	50
3.4	Regularity in auxiliary graphs	59
3.5	Proof of the embedding Lemmas 3.1.3 and 3.1.4	63
3.6	Proof of the embedding Lemma 3.1.1	73
3.6.1	Case $k = 2$	73
3.6.2	Case $k \geq 3$	78
3.7	Supplementary proofs	81
<b>4</b>	<b>Triangles in randomly perturbed graphs</b>	<b>88</b>
4.1	Proof overview of the extremal and the non-extremal case	89
4.1.1	Extremal case.	90
4.1.2	Non-extremal case.	90
4.1.3	Proof of Lemma 4.1.2	92
4.2	Proof of the extremal Theorem 1.1.12	93
4.3	Proof of the stability Theorem 1.1.11	96
<b>5</b>	<b>Transversals in hypergraph collection</b>	<b>103</b>
5.1	Setting and terminology	103
5.2	Main theorem	105
5.3	Proof overview	106
5.4	Proof of main theorem	109
5.5	Applications	114
5.5.1	Powers of Hamilton cycles	114
5.5.2	Hamilton $\ell$ -cycles in $k$ -uniform hypergraphs	116
<b>6</b>	<b>Multistage Maker-Breaker game</b>	<b>118</b>
6.1	Winning Criteria	119
6.2	Proof of Theorem 1.3.4 (Hamilton cycle game)	123
6.2.1	Maker's strategy	124
6.3	Proof of Theorem 1.3.6 (Non- $k$ -colourability game)	125
6.3.1	Maker's strategy	126
6.3.2	Breaker's strategy	127

6.4	Proof of Theorem 1.3.7 ( <i>H</i> -game) . . . . .	127
6.4.1	Maker’s strategy . . . . .	127
6.4.2	Breaker’s strategy . . . . .	128
6.5	Proof of Theorem 1.3.8 (Pancyclicity game) . . . . .	130
6.5.1	Maker’s strategy . . . . .	130
<b>7</b>	<b>Conclusion and open problems</b>	<b>132</b>
7.1	Other spanning structures in randomly perturbed graphs . . . . .	132
7.1.1	Larger powers of Hamilton cycles . . . . .	132
7.1.2	Larger clique factors . . . . .	134
7.2	Transversals in hypergraph collections . . . . .	135
7.2.1	Vertex degree for tight Hamilton cycles . . . . .	135
7.2.2	Exact results and stability . . . . .	135
7.2.3	Other potential applications . . . . .	136
7.3	Multistage positional games . . . . .	136
7.3.1	Connection to the classical setting . . . . .	136
7.3.2	Additional games . . . . .	137
7.3.3	Multistage Maker-Breaker game with stop . . . . .	137

---

## Introduction

---

Broadly speaking, extremal combinatorics is concerned with the following question.

**Question 1.0.1.** *How large (with respect to some parameter  $P$ ) can a structure be without containing a forbidden substructure?*

The structures we are interested in here are graphs. A *graph* is a pair  $(V, E)$ , where  $V$  is a set of *vertices* and  $E$  is a set of pairs of vertices called *edges*. A generalisation of this notion allows edges to join more than two vertices, in which case we call the pair  $(V, E)$  a *hypergraph*. For further explanations concerning notation, see Section 1.4. The origin of (hyper)graph theory within mathematics dates back at least to Euler's problem of the seven bridges of Königsberg in the 18th century [49], but experienced a significant growth in the last century, with a wealth of applications in computer science, engineering and social sciences. An early appearance of an extremal problem is given by Mantel's Theorem.

**Theorem 1.0.2** ([87]). *Let  $G$  be a triangle-free  $n$ -vertex graph. Then the number of edges of  $G$  is at most  $\lfloor n^2/4 \rfloor$ .*

Comparing this theorem with Question 1.0.1, the structure is a graph on  $n$  vertices, the forbidden substructure is a triangle, and the parameter  $P$  is the number of edges. This result was later greatly generalised by Turán.

Answers to Question 1.0.1 lead to some additional natural questions: among all structures forbidding a given substructure, what are the ones maximising  $P$ ? These structures are usually called the *extremal structures*. Going even further, we can ask if *stability* holds, that is, if structures which do not contain the forbidden substructure and almost maximise  $P$  are *close* to one of the extremal structures. Going back to Mantel's theorem, it is easy to show that the complete balanced bipartite graph  $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$  with parts of size  $\lfloor n/2 \rfloor$  and  $\lceil n/2 \rceil$  is the unique triangle-free  $n$ -vertex graph with precisely  $\lfloor n^2/4 \rfloor$  edges. Therefore the graph  $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$  is the unique extremal graph for this problem. A stability version of Mantel's Theorem is given by the famous Stability Theorem of Erdős and Simonovits [45, 104]: if the number of edges of a triangle-free  $n$ -vertex graph is close to  $n^2/4$ , then the graph must be close to the complete balanced bipartite graph. More precisely, the theorem



states that if  $G$  is a triangle-free  $n$ -vertex graph with  $\lfloor n^2/4 \rfloor - o(n^2)$  edges, then it can be transformed into the graph  $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$  by changing only  $o(n^2)$  edges.

Question 1.0.1 investigates when the behaviour of a structure changes with respect to the containment of a forbidden substructure. Therefore, a convenient way to describe and answer it is through the notion of a *threshold*. Thresholds form the common thread in the different parts of this thesis, though it will depend on the context what precisely we mean by a threshold (which we shall make precise). Our contributions can be grouped in three areas: the subgraph containment problem in randomly perturbed graphs (Chapters 3 and 4), the existence of transversals in hypergraph collections (Chapter 5), and the study of a new multistage variant of the classical Maker-Breaker game (Chapter 6). We will now briefly introduce each of these topics, before delving into more details in the subsequent sections.

The *randomly perturbed graph*  $G_\alpha \cup G(n, p)$  is a graph obtained by taking a graph  $G_\alpha$  on  $n$  vertices with minimum degree at least  $\alpha n$  and adding the edges of the random graph  $G(n, p)$  on the same vertex set. This model was introduced by Bohman, Frieze, and Martin [22], who, by adding random edges to a dense graph, were interested in understanding the interplay between the extremal and the probabilistic setting. The same idea, though in a different form, also appeared around the same time in Computer Science: Spielman and Teng [105] coined the notion of smoothed analysis of algorithms and, by randomly perturbing an input to an algorithm, they could interpolate between a worst-time case analysis and an average case analysis.

For a fixed  $\alpha \in (0, 1)$ , we investigate under which conditions on  $p$  we have that, for any graph  $G_\alpha$  with minimum degree  $\alpha n$ , asymptotically almost surely the perturbed graph  $G_\alpha \cup G(n, p)$  contains a given subgraph, where, by *asymptotically almost surely* (a.a.s.), we mean that the statement holds with probability tending to 1 as  $n$  tends to infinity. We first study the appearance of the square of a Hamilton cycle in Chapter 3. This is known when  $\alpha > 1/2$ , and we determine the exact perturbed threshold in all the remaining cases, i.e., for each  $\alpha \leq 1/2$ . Our result has implications on the perturbed threshold for 2-universality and for triangle factors, where we also fully address all open cases. Moreover, in Chapter 4, we study the problem of finding pairwise vertex-disjoint triangles in the same graph model. We prove that a.a.s.  $G \cup G(n, p)$  contains at least  $\min\{\delta(G), \lfloor n/3 \rfloor\}$  pairwise vertex-disjoint triangles, provided  $p \geq C \log n/n$ , where  $C$  is a large enough constant. This can be seen as a perturbed version of an old result of Dirac. Additionally, we prove a stability version of all the results mentioned above.

We then consider transversals in hypergraph collections. A *hypergraph collection* is a collection of hypergraphs  $\mathbf{H} = (H_1, \dots, H_m)$  with the same vertex set. An  $m$ -edge

hypergraph  $F \subset \cup_{i \in [m]} H_i$  is a *transversal* if there is a bijection  $\phi: E(F) \rightarrow [m]$  such that  $e \in H_{\phi(e)}$  for each  $e \in E(F)$ .

For a fixed a hypergraph  $F$ , we want to estimate how large the minimum  $d$ -degree of each  $H_i$  needs to be, so that  $\mathbf{H}$  necessarily contains a transversal copy of  $F$ . There is a trivial lower bound: each  $H_i$  in the collection could be the same hypergraph, hence the minimum  $d$ -degree of each  $H_i$  needs to be large enough to ensure that  $H_i$  contains  $F$ . When  $F$  is a (2-uniform) Hamilton cycle, Joos and Kim [70] showed that this lower bound is tight. After their result, a growing body of work has shown that this is the case for several other hypergraphs  $F$ . However, each such work uses specific and ad-hoc methods, which does not allow much flexibility, hence is often not useful when trying to prove the same behaviour for other hypergraphs.

In Chapter 5, we provide a unified approach to this problem by giving a widely applicable sufficient condition on the hypergraph  $\mathcal{F}$  for the trivial lower bound to be asymptotically tight. The condition is general enough to recover many previous results in the area and obtain novel transversal variants of several classical Dirac-type results for (powers of) Hamilton cycles. For example, it can be used to prove a rainbow version of the minimum  $d$ -degree threshold for the containment of a  $k$ -uniform Hamilton  $\ell$ -cycle, for several ranges of the parameters  $k, d, \ell \in \mathbb{N}$  where  $1 \leq d, \ell \leq k - 1$ . Moreover, it can also be used to prove a rainbow version of the Pósa-Seymour conjecture on the minimum degree threshold for the containment of the  $r$ -th power of a Hamilton cycle in graphs.

Finally, we introduce and discuss a new variant of the Maker-Breaker positional game, which we call the multistage Maker-Breaker game. Given a hypergraph  $\mathcal{H} = (X, \mathcal{F})$  and a *bias*  $b \geq 1$ , the classical  $(1: b)$  Maker-Breaker game on  $\mathcal{H}$  is played as follows. The players, Maker and Breaker, alternately claim elements of  $X$ , with Maker moving first. Maker claims one element per turn, while Breaker can claim up to  $b$  elements. If Maker manages to claim all the elements of set  $F \in \mathcal{F}$ , she is declared the winner. Otherwise, Breaker wins.

The goal of our new variant is to initiate a different perspective on the so-called random graph intuition: while in the  $(1: b)$  Maker-Breaker game, Maker's graph is forced to be sparse when Breaker's bias is large, we want to force sparseness by playing multiple games consecutively. More precisely, the  $(1: b)$  multistage Maker-Breaker game on  $\mathcal{H}$  is played in several stages as follows. Each stage is played as a usual  $(1: b)$  Maker-Breaker game, until all the elements of the board get claimed by one of the players, with the first stage being played on  $\mathcal{H}$ . In every subsequent stage, the game is played on the board reduced to the elements that Maker claimed in the previous stage, and with the winning sets reduced to those fully contained in the new board. The game proceeds until no winning sets remain, and the goal of Maker is to prolong the duration of the game for as many stages as possible.

In Chapter 6, we estimate the maximum duration of the  $(1: b)$  multistage Maker-Breaker

game for some standard graph games played on the edge set of  $K_n$  and with biases  $b$  sub-polynomial in  $n$ : the connectivity game, the Hamilton cycle game, the non- $k$ -colourability game, the pancyclicity game and the  $H$ -game. We show that, while the first three games adhere to a probabilistic intuition, it turns out that the last two games fail to do so.

**Structure of the thesis.** The rest of this chapter is intended to provide the relevant history, formalise the concepts introduced above, add some more motivations, and state our results for each of the three areas mentioned above. A collection of the general definitions and the conventions for notation used throughout the thesis is included for an easier consultation in Section 1.4. Chapter 2 discusses the relevant tools and methods from the literature. We then proceed to the proofs of our results in Chapters 3 to 6, which will be respectively concerned with the square of a Hamilton cycle in randomly perturbed graphs, triangles in randomly perturbed graphs, transversals in hypergraph collections, and the Maker-Breaker multistage game. This thesis is then concluded in Chapter 7 with a discussion of some of the open questions arising from our work.

## 1.1 Randomly perturbed graphs

Mantel's Theorem answers Question 1.0.1 when we ask to maximise the number of edges in a triangle-free  $n$ -vertex graph. When the forbidden substructure is spanning, it is natural to consider maximising the minimum degree, rather than the number of edges. An easy corollary of Mantel's Theorem gives the corresponding version for the minimum degree: any  $n$ -vertex graph  $G$  with minimum degree larger than  $\lfloor n/2 \rfloor$  contains a triangle. This minimum degree condition cannot be lowered as the complete balanced bipartite graph  $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$  has minimum degree  $\lfloor n/2 \rfloor$ , but it does not contain any triangle.

So far, we have been focusing on deterministic graphs, but Question 1.0.1 can be considered for random structures as well, in which case we formulate it as follows.

**Question 1.1.1.** *How dense can a random structure be before it contains a forbidden substructure with high probability?*

As a random structure, we will consider the binomial random graph  $G(n, p)$ . Already Erdős and Rényi [47], in one of the early papers on random graphs, proved that if the edge density  $p$  is at least  $Cn^{-1}$ , with  $C$  being a large constant, then  $G(n, p)$  contains a triangle with high probability. This is asymptotically best possible by an easy first moment calculation: the expected number of triangles in  $G(n, p)$  is  $\binom{n}{3}p^3$ , which goes to zero if  $p = o(n^{-1})$ . Therefore, if  $p = o(n^{-1})$ , with high probability,  $G(n, p)$  does not contain any triangle.

The study of the randomly perturbed graph model introduced above combines Questions 1.0.1 and 1.1.1. The natural question in this setting then is how large one can choose

the minimum degree of the deterministic graph and the edge probability of the random graph while avoiding the perturbed graph to contain a given forbidden substructure with high probability.

**Question 1.1.2.** *Given a real number  $\alpha \in [0, 1]$ , how dense can  $G(n, p)$  be before, for each  $n$ -vertex graph  $G_\alpha$  with  $\delta(G_\alpha) \geq \alpha n$ , we have that a.a.s.  $G_\alpha \cup G(n, p)$  contains a forbidden substructure?*

It is easy to answer Question 1.1.2 if the forbidden substructure is a triangle. When  $\alpha = 0$ , the graph  $G_0$  can be the empty graph, and thus we can only rely on random edges coming from  $G(n, p)$ . Therefore the model coincides with  $G(n, p)$  and a triangle appears with high probability for  $p \geq Cn^{-1}$ , as shown by Erdős and Rényi [47]. When  $0 < \alpha \leq 1/2$ , it is easy to see that we only need to add to  $G_\alpha$  a constant number of random edges and thus  $p \geq Cn^{-2}$  suffices, where again  $C$  is a large constant. This is asymptotically optimal, as  $G(n, p)$  with  $p = o(n^{-2})$  a.a.s. is empty, while  $G_\alpha$  might contain no triangles. When  $\alpha > 1/2$ , by Mantel's Theorem, the graph  $G_\alpha$  already contains a triangle and there is no need to add random edges, which means we can take  $p = 0$ .

We will consider Question 1.1.2 for several forbidden substructures: the square of a Hamilton cycle,  $m$  pairwise vertex-disjoint triangles, and the simultaneous containment of every graph with maximum degree at most 2 (i.e. the property of being 2-universal). We state our results in Sections 1.1.3 and 1.1.4. However, before doing that, we formalise the question by introducing a notion of *threshold* in Section 1.1.1, and we overview the known results in the area in Section 1.1.2.

### 1.1.1 Thresholds

As already mentioned, it is convenient to introduce a notion of threshold in order to have a suitable language for Questions 1.0.1 to 1.1.2. We define such a notion for each of the three models considered above (dense graphs, random graphs, and randomly perturbed graphs), and we do that for a general graph property, rather than only when we want to forbid a given substructure. We will say that a property is *monotone* if the addition of edges cannot destroy the property. Many natural properties are monotone: for example, the property of containing a given subgraph.

We start with the notion of a minimum degree threshold in dense graphs.

**Definition 1.1.3** (Minimum degree threshold). Let  $\mathcal{P}$  be a graph property. The minimum degree threshold for  $\mathcal{P}$  is defined as the infimum of the set of the real numbers  $\delta$  such that the following holds: let  $G$  be any  $n$ -vertex graph with  $\delta(G) \geq \delta n$ , then  $G \in \mathcal{P}$ .

Similarly, we can ask for which values of  $p$  we have a.a.s. that  $G(n, p)$  has property  $\mathcal{P}$ . Observe that for monotone properties, the probability of the event ' $G(n, p) \in \mathcal{P}$ ' increases

with  $p$ . Therefore we might want to ask for a value  $\hat{p}$  such that if  $p$  is much larger than  $\hat{p}$  a.a.s.  $G(n, p) \in \mathcal{P}$ , while if  $p$  is much smaller than  $\hat{p}$  a.a.s.  $G(n, p) \notin \mathcal{P}$ . This justifies the following notion of threshold.

**Definition 1.1.4** (Threshold in  $G(n, p)$ ). We say that  $\hat{p} = \hat{p}(n, \mathcal{P})$  is a threshold for property  $\mathcal{P}$  if the following two conditions hold:

- (i) for any  $p = \omega(\hat{p})$ , a.a.s.  $G(n, p) \in \mathcal{P}$ , and
- (ii) for any  $p = o(\hat{p})$ , a.a.s.  $G(n, p) \notin \mathcal{P}$ .

Strictly speaking, a threshold is not uniquely determined and really is defined here only within constant factors. Nevertheless, we will always talk about *the* threshold. It was shown by Bollobás and Thomason [24] that all nontrivial monotone properties admit a threshold function, where by nontrivial we mean that the property is not satisfied by the empty graph, and is satisfied by the complete graph, while by monotone we mean that the addition of edges preserves the property.

Finally, Definition 1.1.4 can be adapted to the setting of randomly perturbed graphs as follows.

**Definition 1.1.5** (Perturbed threshold). Let  $\alpha \in [0, 1]$ . We say that  $\hat{p}_\alpha = \hat{p}_\alpha(n, \mathcal{P})$  is the perturbed threshold for property  $\mathcal{P}$  at  $\alpha$  if the following two conditions hold:

- (i) for any  $p = \omega(\hat{p}_\alpha)$  and for any sequence of  $n$ -vertex graphs  $(G_{\alpha, n})_{n \in \mathbb{N}}$  with  $\delta(G_{\alpha, n}) \geq \alpha n$ , a.a.s.  $G_{\alpha, n} \cup G(n, p) \in \mathcal{P}$ , and
- (ii) for any  $p = o(\hat{p}_\alpha)$ , there exists a sequence of  $n$ -vertex graphs  $(G_{\alpha, n})_{n \in \mathbb{N}}$  with  $\delta(G_{\alpha, n}) \geq \alpha n$  such that a.a.s.  $G_{\alpha, n} \cup G(n, p) \notin \mathcal{P}$ .

Observe that the property of containing a given subgraph in  $G_\alpha \cup G(n, p)$  is monotone, and thus the threshold exists. We remark that for all the results we will mention later, a sharper threshold condition holds. Namely, there exist constants  $C = C(\alpha, \mathcal{P}) > 0$  and  $c = c(\alpha, \mathcal{P}) > 0$  such that (i) already holds for  $p \geq C\hat{p}_\alpha$  and (ii) already holds for  $p \leq c\hat{p}_\alpha$ . For  $\alpha = 0$ , the perturbed threshold is simply the usual threshold for purely random graphs, and the perturbed threshold is 0 for any  $\alpha$  such that all graphs with minimum degree  $\alpha n$  are in  $\mathcal{P}$ . In this sense, randomly perturbed graphs interpolate between questions from extremal graph theory and questions concerning random graphs. For small  $\alpha > 0$ , we are trying to compare the threshold in  $G(n, p)$  with the one in  $G_\alpha \cup G(n, p)$  and we are asking how much random graph theory results are influenced by the fact that in a random graph there may be vertices with relatively few neighbours. For small  $p$  (of order  $n^{-1}$ ) and hence values of  $\alpha$  close to the minimum degree threshold, we are asking how ‘atypical’ extremal graphs for the property are. However, more generally, we would like to determine the evolution of the perturbed threshold, as  $\alpha$  ranges over the whole interval  $[0, 1]$ .

### 1.1.2 Overview of known results

Several strands of results regarding the properties of randomly perturbed graphs have been studied over the years, particularly Ramsey properties [3, 4, 40, 41, 96], Schur properties [39], and the containment of spanning structures. Here we focus on the latter and we now highlight some of the known results, where we remark that much of the research so far has been focusing on small  $\alpha$  and small  $p$ .

The main result of the original paper of Bohman, Frieze, and Martin [22] deals with Hamiltonicity and shows that for every  $\alpha > 0$  there exists  $C = C(\alpha) > 0$  such that, if  $G$  is an  $n$ -vertex graph with  $\delta(G) \geq \alpha n$  and we add  $Cn$  random edges to  $G$ , then a.s. the resulting graph is Hamiltonian. This is optimal for  $\alpha \in (0, 1/2)$ : indeed if we add less than  $(1 - 2\alpha)n$  edges to the complete bipartite graph  $H_\alpha = K_{\alpha n, (1-\alpha)n}$ , then the resulting graph is not Hamiltonian. Therefore at least a linear number of additional edges is needed to make  $H_\alpha$  Hamiltonian, and thus for  $\alpha \in (0, 1/2)$  the perturbed threshold  $\hat{p}_\alpha$  for Hamiltonicity is  $n^{-1}$ . For  $\alpha \geq 1/2$ , the perturbed threshold is zero, as any  $n$ -vertex graph with minimum degree at least  $n/2$  is already Hamiltonian. For  $\alpha = 0$ , we are in the purely random model and the threshold is  $n^{-1} \log n$ , as proved by Pósa [95] and Koršunov [77], independently. Comparing the threshold for  $\alpha = 0$  and for small  $\alpha > 0$ , we can observe that the former requires an additional  $(\log n)$ -factor. This is necessary to guarantee that the minimum degree in  $G(n, p)$  is at least 2, which is not needed in the randomly perturbed model, as already the deterministic graph  $G_\alpha$  guarantees this minimum degree.

Krivelevich, Kwan, and Sudakov [79] extended this result on Hamilton cycles to bounded-degree spanning trees, which was in turn extended to a universality result by Böttcher, Han, Kohayakawa, Montgomery, Parczyk, and Person [25]. Böttcher et al. showed that for each  $\alpha > 0$  and  $\Delta \in \mathbb{N}$ , there exists  $C = C(\alpha, \Delta) > 0$  such that the following holds. If  $G$  is an  $n$ -vertex graph with minimum degree at least  $\alpha n$ , then a.s.  $G \cup G(n, p)$  contains every  $n$ -vertex tree with maximum degree at most  $\Delta$ , provided  $p \geq Cn^{-1}$ . The order of magnitude of the bound on  $p$  is optimal for  $\alpha \in (0, 1/2)$ , as the complete bipartite graph  $H_\alpha$  requires a linear number of additional edges already to contain a perfect matching. This shows the perturbed threshold  $\hat{p}_\alpha$  is  $n^{-1}$  for  $\alpha \in (0, 1/2)$ . On the other hand, the threshold in  $G(n, p)$  is again  $n^{-1} \log n$ , as proved by Montgomery [89], and, again, we see a decrease of a  $(\log n)$ -factor in the perturbed threshold. Moreover, when  $\alpha > 1/2$ , there is no need to add random edges and the threshold is zero, as proved by Komlós, Sárközy, and Szemerédi [75].

Another spanning structure which has received considerable attention is the  $r$ -th power of a Hamilton cycle, with the case  $r = 1$  having been the main focus of [22] as explained above. These are relevant structures in extremal graph theory: indeed the square of a Hamilton cycle often serves as a good concrete but reasonably complex special case when results about the appearance of a more general class of structures are still out of reach. For

example, a well-known conjecture by Pósa [44] from the 1960s states that any  $n$ -vertex graph  $G$  of minimum degree at least  $2n/3$  contains the square of a Hamilton cycle. This was solved in the 1990s for large  $n$  by Komlós, Sarközy, and Szemerédi [73], demonstrating the power of the then new and by now celebrated Blow-Up Lemma. Only more than 10 years later the analogous problem was settled for a more general class of spanning subgraphs [30], using the result for squares of Hamilton cycles as a fundamental stepping stone. For  $r \geq 3$ , an analogous conjecture was formulated by Seymour [103] and solved for large  $n$  again by Komlós, Sarközy, and Szemerédi [74], who proved that the minimum degree threshold for the  $r$ -th power of a Hamilton cycle is  $rn/(r+1)$ .

In random graphs, the threshold for the containment of the  $r$ -th power of a Hamilton cycle is  $n^{-1/r}$ . For  $r \geq 3$ , this follows from a more general result of Riordan [98] based on the second moment method. However, for  $r = 2$  the problem proved to be much harder, and the threshold was determined only very recently by Kahn, Narayanan, and Park [71], using tools from the pioneering work of Frankston, Kahn, Narayanan and Park [52] on fractional expectation-thresholds.

In the perturbed model a full picture is still missing. In the regime  $\alpha \geq r/(r+1)$ , where  $G_\alpha$  alone contains the  $r$ -th power of a Hamilton cycle, Dudek, Reiher, Ruciński, and Schacht [43] showed that adding a linear number of random edges suffices to enforce the  $(r+1)$ -st power of a Hamilton cycle. This was improved by Nenadov and Trujić [92] who showed that one can indeed enforce the  $(2r+1)$ -st power of a Hamilton cycle with these parameters. When  $\alpha > 1/2$ , even higher powers of Hamilton cycles have been studied by Antoniuk, Dudek, Reiher, Ruciński, and Schacht [6] and by Antoniuk, Dudek, and Ruciński [7], although this last result is concerned with a different notion of threshold.

We conclude this section by discussing  $H$ -factors. Balogh, Treglown, and Wagner [10] proved that, for a fixed a graph  $H$  and  $\alpha \in (0, 1/v(H))$ , the perturbed threshold for an  $H$ -factor satisfies  $\hat{p}_\alpha = n^{-1/m_1(H)}$ . If  $H$  is a strictly-balanced graph, then the threshold for an  $H$ -factor in  $G(n, p)$  is  $n^{-1/m_1(H)} \log^{1/e(H)} n$  (this follows from Johansson, Kahn, and Vu [69], see Theorem 2.5.2), and thus we save a  $(\log^{1/e(H)} n)$ -factor. However, if  $H$  is non-vertex-balanced, then the threshold for an  $H$ -factor in  $G(n, p)$  is  $n^{-1/m_1(H)}$  (this follows from Gerke and McDowell [53]), and thus having a few deterministic edges does not provide any benefit: the perturbed threshold at  $\alpha = 0$  and the one at small positive  $\alpha$  coincide.

More recently, Han, Morris, and Treglown [57] started a more thorough investigation of the intermediate regime for  $\alpha$  when  $H$  is a clique. The perturbed threshold for a  $K_r$ -factor in the purely random case follows from [69] as above, and gives  $\hat{p}_0 = n^{-2/r} \log^{-2/(r^2-r)} n$ . In the extremal setting, the theorems of Corrádi and Hajnal [38] and Hajnal and Szemerédi [55] show that we have  $\hat{p}_\alpha = 0$  for any  $\alpha \geq 1 - 1/r$ . Han, Morris, and Treglown, improving on [10], proved the following.

**Theorem 1.1.6** (Han, Morris, and Treglown [57]). *Let  $k$  and  $r$  be integers with  $2 \leq k \leq r$ . Then given any  $1 - \frac{k}{r} < \alpha < 1 - \frac{k-1}{r}$ , the perturbed threshold for a  $K_r$ -factor satisfies  $\hat{p}_\alpha = n^{-2/k}$ .*

Observe that in Theorem 1.1.6 the value of  $\hat{p}_\alpha$  demonstrates a *jumping* phenomenon: for a fixed  $k$  with  $2 \leq k \leq r$ , its value is the same for all  $\alpha \in (1 - \frac{k}{r}, 1 - \frac{k-1}{r})$ ; however if  $\alpha$  is just above this interval its value is significantly smaller. Moreover, Theorem 1.1.6 almost bridges the gap between Hajnal-Szemerédi and Johansson–Kahn–Vu for clique-factors, leaving open only the *boundary* cases  $\alpha = 1 - k/r$  for each  $k$  with  $2 \leq k \leq r - 1$ .

### 1.1.3 New results: triangles

We now discuss our results related to the appearance of pairwise vertex-disjoint copies of triangles.

As already mentioned in the previous section, the perturbed threshold for a triangle factor is known for all  $\alpha \in [0, 1] \setminus \{1/3\}$ , with the boundary case  $\alpha = 1/3$  being still open. Observe that for  $G_{1/3} \cup G(n, p)$  to a.a.s. contain a triangle factor, we need  $p \geq Cn^{-1} \log n$ . To see this consider the complete bipartite graph  $G = H_{1/3}$ , and denote the partition classes by  $A$  and  $B$  with  $|A| < |B|$ . By Markov’s inequality, if  $p \leq \frac{1}{2}n^{-1} \log n$ , then a.a.s. there are  $O(\log^4 n)$  triangles within  $B$  and a.a.s. there is a polynomial number of vertices in the class  $B$  without any neighbours in  $B$  [68, Theorem 6.36]. However, for a triangle factor to exist, for each triangle with at most one vertex in  $B$ , there must be at least one triangle fully contained in  $B$ . In conclusion, a.a.s.  $G \cup G(n, p)$  does not contain a triangle factor. We can show that  $n^{-1} \log n$  is indeed the perturbed threshold at  $\alpha = 1/3$ .

**Theorem 1.1.7** (Triangle factor at  $\alpha = 1/3$ ). *Let  $\hat{p}_\alpha$  be the perturbed threshold for a triangle factor. Then we have  $\hat{p}_{1/3} = n^{-1} \log n$ .*

This result closes the problem of determining, given  $\alpha \in [0, 1]$ , the threshold for a triangle factor in  $G_\alpha \cup G(n, p)$  and we refer to Table 1.1 for a summary.

$\alpha$	$\alpha = 0$	$0 < \alpha < 1/3$	$\alpha = 1/3$	$1/3 < \alpha < 2/3$	$2/3 \leq \alpha$
$\hat{p}_\alpha$	$n^{-2/3} \log^{1/3} n$	$n^{-2/3}$	$n^{-1} \log n$	$n^{-1}$	0

Table 1.1: Triangle factor containment in  $G_\alpha \cup G(n, p)$ , where  $\delta(G_\alpha) \geq \alpha n$ .

Theorem 1.1.7 is a special case of a more general result, where we answer the question which minimum degree condition is needed in the randomly perturbed graph model with  $p$  of order  $n^{-1} \log n$  to enforce  $m$  vertex-disjoint triangles for any  $1 \leq m \leq \lfloor n/3 \rfloor$ .

**Theorem 1.1.8** (Triangles in randomly perturbed graphs). *There exists  $C > 0$  such that for any  $n$ -vertex graph  $G$  we can a.a.s. find at least  $\min\{\delta(G), \lfloor n/3 \rfloor\}$  pairwise vertex-disjoint triangles in  $G \cup G(n, p)$ , provided that  $p \geq Cn^{-1} \log n$ .*



Observe that Theorem 1.1.8 can be seen as a perturbed version of the following result by Dirac.

**Theorem 1.1.9** (Dirac [42]). *Any  $n$ -vertex graph  $G$  with  $n/2 \leq \delta(G) \leq 2n/3$  contains at least  $2\delta(G) - n$  pairwise vertex-disjoint triangles.*

Given an integer  $m$  with  $n/2 \leq m \leq 2n/3$ , the tripartite complete graph with parts of size  $2m - n$ ,  $n - m$ , and  $n - m$  shows the result is best possible. Moreover, a stability version of Theorem 1.1.9 was proved by Hladký, Hu, and Piguet [66]. We are not aware of other results in the randomly perturbed graph model that consider large but not spanning structures. Theorem 1.1.8 is basically optimal in terms of the number of triangles, because given  $1 \leq m < n/3$ , then  $G = K_{m,n-m}$  has minimum degree  $\delta(G) = m$ , and there can be at most  $m$  pairwise vertex-disjoint triangles using each at least one edge of  $G$ , and at most  $O(\log^4 n)$  additional triangles solely coming from  $G(n, p)$ . The bound on  $p$  is asymptotically optimal as it is in Theorem 1.1.7, but we remark that when  $m$  is ‘significantly smaller’ than  $n/3$ , then already  $p \geq Cn^{-1}$  is sufficient to a.a.s. find  $m$  pairwise vertex-disjoint triangles in  $K_{m,n-m} \cup G(n, p)$ . We will provide the proof of Theorem 1.1.8 in Chapter 4.

### Stability

We already discussed how the probability  $p \geq Cn^{-1} \log n$  can not be significantly lowered in Theorem 1.1.8. However, we are able to show that when the minimum degree of  $G$  is linear in  $n$ , then with  $m = \min\{\delta(G), n/3\}$ , the complete bipartite graph  $K_{m,n-m}$  is the unique extremal graph for Theorem 1.1.8, in the sense that if the graph  $G$  is not ‘close’ to  $K_{m,n-m}$  then a.a.s.  $G \cup G(n, p)$  contains  $m$  pairwise vertex-disjoint triangles already at probability  $p \geq Cn^{-1}$  and we can even assume a slightly smaller minimum degree on  $G$ . The next definition formalises what we mean by *close*.

**Definition 1.1.10** ( $(\alpha, \beta)$ -stable). For  $\alpha, \beta > 0$  we say that an  $n$ -vertex graph  $G$  is  $(\alpha, \beta)$ -stable if there exists a partition of  $V(G)$  into two sets  $A$  and  $B$  of sizes  $|A| = (\alpha \pm \beta)n$  and  $|B| = (1 - \alpha \pm \beta)n$  such that the minimum degree of the bipartite subgraph  $G[A, B]$  is at least  $\alpha n/4$ , all but at most  $\beta n$  vertices from  $A$  have degree at least  $|B| - \beta n$  into  $B$ , all but at most  $\beta n$  vertices from  $B$  have degree at least  $|A| - \beta n$  into  $A$ , and  $G[B]$  contains at most  $\beta n^2$  edges.

We can show a stability version of Theorem 1.1.8.

**Theorem 1.1.11** (Stability Theorem for triangles). *For  $0 < \beta < 1/12$  there exist  $\gamma > 0$  and  $C > 0$  such that for any  $\alpha$  with  $4\beta \leq \alpha \leq 1/3$  the following holds. Let  $G$  be an  $n$ -vertex graph with minimum degree  $\delta(G) \geq (\alpha - \gamma)n$  that is not  $(\alpha, \beta)$ -stable. With  $p \geq Cn^{-1}$ , a.a.s. the perturbed graph  $G \cup G(n, p)$  contains at least  $\min\{\alpha n, \lfloor n/3 \rfloor\}$  pairwise vertex-disjoint triangles.*

The result is best possible as  $G$  can be bipartite and have no triangles, in which case we need at least a linear number of edges from the random graph to find a linear number of pairwise vertex-disjoint triangles in  $G \cup G(n, p)$ . On the other hand, the logarithmic factor is needed for the extremal graph.

**Theorem 1.1.12** (Extremal Theorem for triangles). *For  $0 < \alpha_0 \leq 1/3$  there exist  $\beta, \gamma > 0$  and  $C > 0$  such that for any  $\alpha$  with  $\alpha_0 \leq \alpha \leq 1/3$  the following holds. Let  $G$  be an  $n$ -vertex graph with minimum degree  $\delta(G) \geq (\alpha - \gamma)n$  that is  $(\alpha, \beta)$ -stable. With  $p \geq Cn^{-1} \log n$ , a.a.s. the perturbed graph  $G \cup G(n, p)$  contains at least  $\min\{\delta(G), \lfloor \alpha n \rfloor\}$  pairwise vertex-disjoint triangles.*

Indeed our argument will give slightly more. If  $G$  is  $(\alpha, \beta)$ -stable with partition  $V(G) = A \cup B$  and  $|A| \geq \alpha n$ , then we can a.a.s. find  $\lceil \alpha n \rceil$  if  $\alpha < 1/3$  and  $\lfloor n/3 \rfloor$  if  $\alpha = 1/3$  pairwise vertex-disjoint triangles in  $G \cup G(n, p)$  (even when the minimum degree in  $G$  is smaller than  $\alpha n$ ).

Moreover we prove the following when the minimum degree gets smaller than that in Theorems 1.1.11 and 1.1.12

**Theorem 1.1.13** (Sublinear Theorem for triangles). *There exists  $C > 0$  such that the following holds for any  $1 \leq m \leq n/512$  and any  $n$ -vertex graph  $G$  of minimum degree  $\delta(G) \geq m$ . With  $p \geq C \log n/n$  a.a.s. the perturbed graph  $G \cup G(n, p)$  contains at least  $m$  pairwise vertex-disjoint triangles.*

Observe that Theorem 1.1.8 easily follows from Theorems 1.1.11, 1.1.12, and 1.1.13. We will prove Theorem 1.1.8 and Theorem 1.1.13 at the beginning of Chapter 4, and Theorems 1.1.11 and 1.1.12 in Sections 4.3 and 4.2, respectively.

#### 1.1.4 New results: the square of a Hamilton cycle

We now discuss our results related to the appearance of the square of a Hamilton cycle.

We determine the perturbed threshold for the containment of the square of a Hamilton cycle for  $0 < \alpha \leq 1/2$ , which answers a question of Antoniuk, Dudek, Reiher, Ruciński and Schacht [6] in a strong form. In the range  $\alpha \in (1/2, 2/3)$  the perturbed threshold for squares of Hamilton cycles was determined by Dudek, Reiher, Ruciński, and Schacht [43]. The case  $\alpha = 0$ , on the other hand, is the purely random graph case addressed in [71], and the range  $\alpha \geq 2/3$  is the purely extremal scenario addressed in [73]. Therefore, our result completely settles the question of determining the perturbed threshold for the square of a Hamilton cycle for the whole range of  $\alpha$ .

**Theorem 1.1.14** (Square of a Hamilton cycle). *Let  $\hat{p}_\alpha$  be the perturbed threshold for the*

containment of the square of a Hamilton cycle. Then

$$\hat{p}_\alpha = \begin{cases} 0 & \text{if } \alpha \geq \frac{2}{3}, \\ n^{-1} & \text{if } \alpha \in [\frac{1}{2}, \frac{2}{3}), \\ n^{-(k-1)/(2k-3)} & \text{if } \alpha \in (\frac{1}{k+1}, \frac{1}{k}) \text{ for } k \geq 2, \\ n^{-(k-1)/(2k-3)} (\log n)^{1/(2k-3)} & \text{if } \alpha = \frac{1}{k+1} \text{ for } k \geq 2, \\ n^{-1/2} & \text{if } \alpha = 0. \end{cases}$$

The probabilities appearing in Theorem 1.1.14 have the following nice interpretation:  $n^{-(k-1)/(2k-3)}$  is the threshold in  $G(n, p)$  for a linear number of copies of  $P_k^2$  (by a standard application of Janson's inequality, see also Lemma 2.5.1), while  $n^{-(k-1)/(2k-3)} (\log n)^{1/(2k-3)}$  is the threshold in  $G(n, p)$  for the existence of a  $P_k^2$ -factor (this follows from [69], see Theorem 2.5.2). The appearance of such probabilities will become clearer while justifying the lower bounds and explaining our proof strategy in Section 3.1.1. We will provide the proof of Theorem 1.1.14 in Chapter 3.

Observe that Theorem 1.1.14 implies that, as long as  $\alpha \in (1/3, 2/3)$  it suffices to add a linear number of random edges to the deterministic graph  $G_\alpha$  for enforcing the square of a Hamilton cycle, and for  $\alpha \leq 1/3$  the perturbed threshold  $\hat{p}_\alpha(n)$  exhibits 'jumps' at  $\alpha = \frac{1}{k+1}$  for each integer  $k \geq 2$ , where an extra  $((\log n)^{1/(2k-3)})$ -factor is needed at  $\alpha$  precisely equal to  $\frac{1}{k+1}$ . To the best of our knowledge Theorem 1.1.14 is the first result exhibiting a countably infinite number of jumps. Moreover, for  $\alpha$  tending to zero, the threshold  $\hat{p}_\alpha$  tends to  $n^{-1/2}$ , which is precisely  $\hat{p}_0$ , i.e. the threshold is continuous at zero.

As already explained, the cases  $\alpha = 0$  and  $\alpha > 1/2$  follow from known results. Moreover, the case  $\alpha = 1/2$  will follow from the monotonicity of the perturbed threshold, once we will have determined the perturbed threshold in the range  $\alpha < 1/2$ . Therefore, in the remainder of this section, we can fix an integer  $k \geq 2$  and assume  $\alpha \in [\frac{1}{k+1}, \frac{1}{k})$ .

### Lower bounds and stability

The lower bounds in Theorem 1.1.14 on  $\hat{p}_\alpha$  for  $\alpha \in (\frac{1}{k+1}, \frac{1}{k})$  and  $\alpha = \frac{1}{k+1}$  follows by taking the graph  $H_\alpha$  as the deterministic graph  $G_\alpha$ , where we recall that  $H_\alpha$  is the complete bipartite graph with parts of size  $\alpha n$  and  $(1 - \alpha)n$ .

**Proposition 1.1.15** (Lower bounds for the square of a Hamilton cycle).

- (i) Let  $\alpha \in (\frac{1}{k+1}, \frac{1}{k})$ . Then there exists  $c \in (0, 1)$  such that  $H_\alpha \cup G(n, p)$  a.a.s. does not contain a copy of  $C_n^2$ , provided  $p \leq cn^{-(k-1)/(2k-3)}$ .
- (ii) Let  $\alpha = \frac{1}{k+1}$ . Then  $H_\alpha \cup G(n, p)$  a.a.s. does not contain a copy of  $C_n^2$ , provided  $p \leq \frac{1}{2} n^{-(k-1)/(2k-3)} (\log n)^{1/(2k-3)}$ .

Proposition 1.1.15 shows that, compared to the threshold for  $\alpha \in (\frac{1}{k+1}, \frac{1}{k})$ , the additional  $((\log n)^{1/(2k-3)})$ -factor in the threshold at  $\alpha = \frac{1}{k+1}$  is needed in order to a.a.s. find the square of a Hamilton cycle in  $H_{1/(k+1)} \cup G(n, p)$ . We can say more and prove a more general stability version of Theorem 1.1.14: namely, that the graph  $H_{1/(k+1)}$  is the ‘only’ reason we need the extra term, in the sense that this extra factor at  $\alpha = \frac{1}{k+1}$  is only necessary when the deterministic graph  $G_\alpha$  is close to  $H_{1/(k+1)}$ , where being close is again formalised by the definition of  $(\alpha, \beta)$ -stable graphs introduced above (see Definition 1.1.10). The following stability theorem treats the non-extremal case of Theorem 1.1.14.

**Theorem 1.1.16** (Stability Theorem for the square of a Hamilton cycle). *For every  $k \geq 2$  and every  $0 < \beta < \frac{1}{6k}$ , there exist  $\gamma > 0$  and  $C > 0$  such that the following holds. Let  $G$  be any  $n$ -vertex graph with minimum degree at least  $(\frac{1}{k+1} - \gamma)n$  that is not  $(\frac{1}{k+1}, \beta)$ -stable. Then  $G \cup G(n, p)$  a.a.s. contains the square of a Hamilton cycle, provided that  $p \geq Cn^{-(k-1)/(2k-3)}$ .*

Only when the graph  $G$  is stable we need the  $((\log n)^{1/(2k-3)})$ -factor. This case is treated by the following theorem.

**Theorem 1.1.17** (Extremal Theorem for the square of a Hamilton cycle). *For every  $k \geq 2$  there exist  $\beta > 0$  and  $C > 0$  such that the following holds. Let  $G$  be any  $n$ -vertex graph with minimum degree at least  $\frac{1}{k+1}n$  that is  $(\frac{1}{k+1}, \beta)$ -stable. Then  $G \cup G(n, p)$  a.a.s. contains the square of a Hamilton cycle, provided that  $p \geq Cn^{-(k-1)/(2k-3)}(\log n)^{1/(2k-3)}$ .*

Together with the lower bounds provided in Proposition 1.1.15, Theorems 1.1.16 and 1.1.17 imply Theorem 1.1.14 for  $\alpha \in [\frac{1}{k+1}, \frac{1}{k})$  with  $k \geq 2$ . We will prove Theorem 1.1.14 and Proposition 1.1.15 at the beginning of Chapter 3, and Theorems 1.1.16 and 1.1.17 in Sections 3.3 and 3.2, respectively.

## 2-universality

Observe that the square of a Hamilton cycle is 2-universal, i.e. it contains every graph with maximum degree at most 2. Therefore, as a corollary of Theorem 1.1.14 we also get the following result, establishing the perturbed threshold for 2-universality for all  $\alpha$ .

**Corollary 1.1.18** (2-universality). *Let  $\hat{p}_\alpha$  be the perturbed threshold for 2-universality. Then*

$$\hat{p}_\alpha = \begin{cases} 0 & \text{if } \alpha \geq \frac{2}{3}, \\ n^{-1} & \text{if } \alpha \in (\frac{1}{3}, \frac{2}{3}), \\ n^{-1} \log n & \text{if } \alpha = \frac{1}{3}, \\ n^{-2/3} & \text{if } \alpha \in (0, \frac{1}{3}), \\ n^{-2/3}(\log n)^{1/3} & \text{if } \alpha = 0. \end{cases}$$

The upper bound for the range  $1/3 \leq \alpha < 2/3$  is a corollary of our Theorem 1.1.14, while the lower bound can be justified again by taking the graph  $H_\alpha$  as the deterministic graph with minimum degree  $\alpha n$ . This indeed follows from the corresponding discussion for triangle factors in Section 1.1.3, as if a graph is 2-universal, then it has to contain a triangle factor.

The remaining cases of Corollary 1.1.18 follow from known results. For  $\alpha = 0$ , this is due to Ferber, Kronenberg, and Luh [51], for  $\alpha \geq 2/3$  to Aigner and Brandt [2], and for  $\alpha \in (0, 1/3)$  to Parczyk [93].

Observe that the perturbed threshold for the containment of the square of a Hamilton cycle and the perturbed threshold for 2-universality differ for  $\alpha < 1/3$ . This is due to the fact that in this regime the structure of the deterministic graph  $G_\alpha$  may force us to find many copies of the square of a short path in  $G(n, p)$  if we want to find the square of a Hamilton cycle in  $G_\alpha \cup G(n, p)$ .

## 1.2 Transversal versions of Dirac-type theorems

A *hypergraph collection* on vertex set  $V$  is a collection of hypergraphs  $H_1, \dots, H_m$ , each with vertex set  $V$ , and it is denoted by  $\mathbf{H} = (H_1, \dots, H_m)$ . We call the collection a *graph collection* if each hypergraph in the collection has uniformity two. Given an  $m$ -edge hypergraph  $F$  on  $V$ , we say that  $\mathbf{H}$  has a *transversal* copy of  $F$  if there is a bijection  $\phi: E(F) \rightarrow [m]$  such that  $e \in H_{\phi(e)}$  for each  $e \in E(F)$ . In other words, a transversal copy of  $F$  is a copy of  $F$  in  $\bigcup_{i \in [m]} H_i$  obtained by selecting exactly one edge from each hypergraph. In addition, a transversal copy of  $F$  can be thought as a rainbow copy of  $F$ . Indeed, we can think of the edges of hypergraph  $H_i$  to be coloured with colour  $i$  and, in this framework, a transversal copy of  $F$  is a copy of  $F$  in  $\bigcup_{i \in [m]} H_i$  with edges of pairwise distinct colours.

We are interested in the following general question formulated originally by Joos and Kim [70].

**Question 1.2.1.** *Let  $F$  be an  $m$ -edge hypergraph with vertex set  $V$ ,  $\mathcal{H}$  be a family of hypergraphs, and  $\mathbf{H} = (H_1, \dots, H_m)$  be a hypergraph collection on vertex set  $V$  with  $H_i \in \mathcal{H}$  for each  $i \in [m]$ . Which conditions on  $\mathcal{H}$  guarantee a transversal copy of  $F$  in  $\mathbf{H}$ ?*

By taking  $H_1 = H_2 = \dots = H_m$ , it is clear that such a property needs to guarantee that each hypergraph in  $\mathcal{H}$  contains  $F$  as a subhypergraph. However, this alone is not always sufficient, not even asymptotically. For example, Aharoni, DeVos, de la Maza, Montejano and Šámal [1] showed that if  $\mathbf{G} = (G_1, G_2, G_3)$  is a graph collection on  $[n]$  with  $e(G_i) > (\frac{26-2\sqrt{7}}{81})n^2$  for each  $i \in [3]$ , then  $\mathbf{G}$  contains a transversal which is a triangle. As shown in [1], the constant  $\frac{26-2\sqrt{7}}{81} > 1/4$  is optimal. On the other hand, Mantel's theorem states that already any graph with at least  $n^2/4$  edges must contain a triangle.

Instead of a lower bound on the total number of edges in each hypergraph, it is also natural to investigate what can be guaranteed with a lower bound on the minimum degree. It turns out that even in this more restrictive setting, there can be a discrepancy between the uncoloured and the rainbow versions of the problem (see next section). On the other hand, there are many natural instances where they coincide (at least asymptotically), e.g. [32, 33, 70, 84, 85, 86, 90]. Although such results all show the same behaviour (the asymptotic equivalence between the uncoloured and the rainbow problem), each of them requires an ad-hoc proof. Therefore, it is natural to ask for a more flexible machinery to translate uncoloured results into rainbow ones.

We make progress on this question and give a widely applicable sufficient condition for the uncoloured and the rainbow versions to coincide asymptotically. This offers a unified proof of several known rainbow Dirac-type results, as well as many new ones. We formalise the problem and give an overview of the known results in Section 1.2.1. The precise statement of the main theorem is technical and we postpone it to Chapter 5, together with its proof. However, we state its applications in Section 1.2.2 and give some intuition behind our main theorem in Section 1.2.3.

### 1.2.1 Overview of known results

We have been quite vague on what we mean by the uncoloured and the rainbow minimum degree conditions being (asymptotically) equivalent. To make this more precise, we give the following two definitions. We have already discussed the problem for the uncoloured setting in Section 1.1, where we defined the minimum degree threshold for a graph property to hold (Definition 1.1.3). In Definition 1.1.3 we require the minimum degree threshold to be *exact*: any  $n$ -vertex graph with  $\delta(G) \geq \delta n$  has to satisfy the given property. However, in view of the results of this section, it is more convenient to consider the weaker notion of the *asymptotic* minimum degree threshold, where we only require that the given property is satisfied by the  $n$ -vertex graphs with minimum degree at least  $(\delta + \alpha)n$ , where  $\alpha$  is a strictly positive number which can be chosen arbitrarily small. We formalise that for families of hypergraphs in the next definition. To avoid any confusion, we refer to this definition as the *uncoloured minimum degree threshold*.

**Definition 1.2.2** (Uncoloured minimum degree threshold). Let  $\mathcal{F}$  be an infinite family of  $k$ -uniform hypergraphs. By  $\delta_{\mathcal{F},d}$  we denote, if it exists, the smallest real number  $\delta$  such that for all  $\alpha > 0$  the following holds for all but finitely many  $F \in \mathcal{F}$ . Let  $n = |V(F)|$  and  $H$  be any  $n$ -vertex  $k$ -uniform hypergraph with  $\delta_d(H) \geq (\delta + \alpha)n^{k-d}$ . Then  $H$  contains a copy of  $F$ .

Observe again that, if we allow  $\alpha = 0$ , Definition 1.2.3 is equivalent to Definition 1.1.3 with the (hyper)graph property being the containment of a hypergraph from the family

$\mathcal{F}$ . For example, if  $\mathcal{F}$  is the family of graphs consisting of a cycle on  $n$  vertices for each  $n \in \mathbb{N}$ , then we have  $\delta_{\mathcal{F},1} = 1/2$ . Indeed, this follows from Dirac's theorem (which in fact gives the exact threshold). Similarly, we define the rainbow minimum degree threshold in hypergraph collections.

**Definition 1.2.3** (Rainbow minimum degree threshold). Let  $\mathcal{F}$  be an infinite family of  $k$ -uniform hypergraphs. By  $\delta_{\mathcal{F},d}^{\text{rb}}$  we denote, if it exists, the smallest real number  $\delta$  such that for all  $\alpha > 0$  the following holds for all but finitely many  $F \in \mathcal{F}$ . Let  $n = |V(F)|$  and  $\mathbf{H}$  be any  $k$ -uniform hypergraph collection on  $n$  vertices with  $|\mathbf{H}| = |E(F)|$  and  $\delta_d(\mathbf{H}) \geq (\delta + \alpha)n^{k-d}$ . Then  $\mathbf{H}$  contains a transversal copy of  $F$ .

If the two values are well-defined, it must be that  $\delta_{\mathcal{F},d}^{\text{rb}} \geq \delta_{\mathcal{F},d}$ . Indeed, if  $H$  contains no copy of  $F$ , the collection  $\mathbf{H}$  consisting of  $|E(F)|$  copies of  $H$  does not contain a transversal copy of  $H$  either. However, Montgomery, Műyesser, and Pehova [90] made the following observation which shows that  $\delta_{\mathcal{F},d}^{\text{rb}}$  can be much larger than  $\delta_{\mathcal{F},d}$ . Set  $\mathcal{F} = \{k \times (K_{2,3} \cup C_4) : k \in \mathbb{N}\}$ , where  $K_{2,3} \cup C_4$  denotes the graph obtained by taking the disjoint union of a copy of the complete bipartite graph  $K_{2,3}$  and a copy of the cycle  $C_4$ , while  $k \times (K_{2,3} \cup C_4)$  denotes the graph obtained by taking  $k$  vertex-disjoint copies of the graph  $K_{2,3} \cup C_4$ . It follows from a result of Kühn and Osthus [80] that  $\delta_{\mathcal{F},1} \leq 4/9$ . Consider the graph collection  $\mathbf{H} = (H_1, \dots, H_m)$  on  $V$  obtained in the following way. Partition  $V$  into two almost equal vertex subsets, say  $A$  and  $B$ , and suppose that  $H_1 = H_2 = \dots = H_{m-1}$  are all disjoint unions of a clique on  $A$  and a clique on  $B$ . Suppose that  $H_m$  is a complete bipartite graph between  $A$  and  $B$ . Observe that each  $H_i$  in this resulting graph collection has minimum degree  $\lfloor |V|/2 \rfloor$ . Further observe that if  $\mathbf{H}$  contains a transversal copy of some  $F \in \mathcal{F}$ , there is an edge of a copy of  $K_{2,3}$  or  $C_4$  that gets copied to an edge of  $H_m$ . As any other edge of  $F$  has both endpoints in either  $A$  or  $B$ , this edge has to be a bridge (an edge whose removal disconnects the graph) of that copy of  $K_{2,3}$  or  $C_4$ . However, neither  $K_{2,3}$  nor  $C_4$  contains a bridge. Hence,  $\delta_{\mathcal{F},d}^{\text{rb}} \geq 1/2$ .

Nevertheless, there are several families  $\mathcal{F}$  for which  $\delta_{\mathcal{F},d}^{\text{rb}}$  and  $\delta_{\mathcal{F},d}$  coincide.

**Definition 1.2.4** ( $d$ -colour blind families). Let  $\mathcal{F}$  be a family of  $k$ -uniform hypergraphs and  $1 \leq d < k$ . We say that the family  $\mathcal{F}$  is  $d$ -colour-blind if  $\delta_{\mathcal{F},d}^{\text{rb}} = \delta_{\mathcal{F},d}$ , provided they are both well-defined. In the case  $\mathcal{F}$  is a family of graphs (and  $d = 1$ ), we just say that  $\mathcal{F}$  is *colour-blind*.

For example, Joos and Kim [70], improving a result of Cheng, Wang, and Zhao [34] and confirming a conjecture of Aharoni [1], showed that, if  $n \geq 3$ , then any  $n$ -vertex graph collection  $\mathbf{G} = (G_1, \dots, G_n)$  with  $\delta(G_i) \geq n/2$  for each  $i \in [n]$  has a transversal copy of a Hamilton cycle. This generalises Dirac's classical theorem and implies that the family  $\mathcal{F}$  of cycles is colour-blind<sup>1</sup>. There are many more families of colour-blind (hyper)graphs.

<sup>1</sup>In fact, in this particular case, the corresponding thresholds are *exactly* the same, and there is no need for an error term. We discuss this aspect of the problem further in Section 7.2.

In particular, matchings [32, 84, 85, 86], Hamilton  $\ell$ -cycles [33], factors [32, 90], and spanning trees [90] have been extensively studied.

Building on techniques introduced by Montgomery, Mütters, and Pehova [90], in our main theorem we give a general sufficient condition for a family of hypergraphs  $\mathcal{F}$  to be colour-blind. The formal statement requires some more terminology and thus is postponed to Chapter 5.

### 1.2.2 Applications

The power of our main theorem is its flexibility, which allows to prove rainbow Dirac-type results in a unified and short way. The following theorem lists the applications we derive, though we believe that our setting can capture even more families of hypergraphs.

**Theorem 1.2.5.** *The following families of hypergraphs are all  $d$ -colour-blind.*

- (A) *The family of the  $r$ -th powers of Hamilton cycles for a fixed  $r \geq 2$  (and  $d = 1$ ).*
- (B) *The family of  $k$ -uniform Hamilton  $\ell$ -cycles for the following ranges of  $k$ ,  $\ell$ , and  $d$ .*
  - (B1)  $1 < \ell < k/2$  and  $d = k - 2$ ;
  - (B2)  $1 \leq \ell < k/2$  or  $\ell = k - 1$ , and  $d = k - 1$ ;
  - (B3)  $\ell = k/2$  and  $k/2 < d \leq k - 1$  with  $k$  even.

Theorem 1.2.5 (B2) when  $\ell = k - 1$  was originally proven by Cheng, Han, Wang, Wang, and Yang [33], who raised the problem of obtaining the rainbow minimum degree threshold for a wider range of  $\ell \in [k - 2]$ . Moreover, the case of Hamilton cycles in graphs (i.e.  $k = 2$  and  $d = \ell = 1$ ) was previously proven by Cheng, Wang, and Zhao [34] (and their result was sharpened by Joos and Kim [70]). Theorem 1.2.5 is derived from our main theorem in Section 5.5.

### 1.2.3 An introduction to the main theorem

As explained before, our main theorem concerning this topic (Theorem 5.2.1) is technically complicated. Hence, we do not provide it in this introduction. Instead, we now outline what it states and give some intuition for our approach. Firstly, Theorem 5.2.1 is concerned with hypergraph families  $\mathcal{F}$  with a ‘cyclic’ structure. That is, we assume there exists a hypergraph  $\mathcal{A}$  such that all  $F \in \mathcal{F}$  can be obtained by gluing several copies of  $\mathcal{A}$  in a Hamilton cycle fashion (see Definition 5.1.3). For example, for  $k$ -uniform Hamilton cycles,  $\mathcal{A}$  would be a single  $k$ -uniform edge (see Figure 5.1), whereas for the  $r$ -th power of a Hamilton cycle,  $\mathcal{A}$  would be a clique on  $r$  vertices (see Figure 5.2). Most of the well-studied problems in the uncoloured setting, including everything listed in Theorem 1.2.5, fit into this framework.



A common framework for embedding such hypergraphs with cyclic structure is the absorption method. We will present this method in Section 2.2, showing the typical structure of an absorption-based proof (see Step 1. to Step 5.). Our main theorem essentially states that if there is such a proof that  $\delta$  is the uncoloured minimum degree threshold for some  $\mathcal{F}$  with cyclic structure, then the rainbow minimum degree threshold of  $\mathcal{F}$  is equal to  $\delta$ . Some partial progress towards such an abstract statement was already made in [90]. In fact, in [90], it was remarked that, plausibly, transversal versions of other Dirac-type results can be shown, provided that one can prove strengthened versions of the non-transversal (uncoloured) embedding problem. In [90], this strengthening took the form of embedding trees with the location of a single vertex being specified adversarially (see Theorem 4.4 in [90]). Using such a strengthening, one can translate each step of the uncoloured absorption-based proof into a coloured setting. On the other hand, the main advantage of our main theorem is that it eliminates the need to make ad-hoc strengthenings to the uncoloured version of the result, allowing for a very short proof of Theorem 1.2.5. To achieve this, we codify, through what we call properties **Ab** and **Con**, what it means for there to be streamlined absorption proof for the uncoloured result, and we use the existence of such a proof as a black-box. In our applications, to ensure that the relevant properties hold, we rely on existing lemmas in the literature without having to do any extra work (see Table 5.1).

In addition to properties **Ab** and **Con** which guarantee we can rely on a streamlined absorption proof for the uncoloured result, we have one more hypothesis in the main theorem, which we call property **Fac**. One reason why transversal versions of Dirac-type results are more difficult is that every single hypergraph in the collection as well as every single vertex of the host graph needs to be utilised in the target spanning structure (the transversal). This is crucial as demonstrated by the construction given after Definition 1.2.3. In this construction, the possibility of finding a transversal copy of  $\mathcal{F}$  is ruled out by showing that a particular graph in the collection (namely the hypergraph  $H_m$ ) cannot be used in a transversal copy of a  $K_{2,3}$  or  $C_4$ . Therefore, in addition to some properties which are uncoloured, we require a coloured property which we call **Fac**. This roughly states that, in any fixed, adversarially specified small set of hypergraphs from the collection, we can find vertex-disjoint copies of  $\mathcal{A}$  (the building block of the hypergraph we are trying to find) in a rainbow fashion. This ensures that we never get stuck while trying to use up every single colour/hypergraph that we start with. When  $\mathcal{A}$  is just a single edge (as it will be the case for Theorem 1.2.5 (B)), the property **Fac** is essentially trivial to check (see Observation 5.1.5). For powers of Hamilton cycles, however, this property is more delicate and, to verify **Fac**, we rely on a result from [90].

### 1.3 Multistage Maker-Breaker game

A *positional game* is a perfect information game, played by two players, which can be represented by a hypergraph  $\mathcal{H} = (X, \mathcal{F})$  and a winning rule. The set  $X$  is called the *board* of the game, while the sets of the family  $\mathcal{F} \subseteq \mathcal{P}(X)$  are called the *winning sets* of the game. Both players alternate in claiming unclaimed elements of  $X$  and the winner of the game is determined according to the winning rule and the family  $\mathcal{F}$ . Different winning rules produce different games, and many variants have been considered over the time. One of the most studied one is the so called *Maker-Breaker game*, which attracted a lot of attention in the last few decades, starting with the seminal papers of Hales and Jewett [56], Erdős and Selfridge [48], Chvátal and Erdős [35] and Beck [17, 19]. The two players are called Maker and Breaker, and, for simplicity, we assume them to be female and male, respectively. They alternately claim one unclaimed element of  $X$ , with Maker moving first. If Maker succeeds in claiming all the elements of one of the winning sets  $F \in \mathcal{F}$ , she is declared the winner of the game (we also say that she occupied this winning set). Otherwise, Breaker is declared the winner. Observe that for Breaker to win the game, he has to claim at least one element from each winning set.

Here we will formulate and study a new variant of the Maker-Breaker game, which we call the *multistage Maker-Breaker game*. However, before doing that, we will give a short history and an overview of the known results related to the (classical) Maker-Breaker game in Section 1.3.1. This will also motivate our new variant, which we introduce in Section 1.3.2. Finally, we state our results in Section 1.3.3.

#### 1.3.1 Classical (1: $b$ ) Maker-Breaker game

Maker-Breaker games are played on various boards, but some of the most studied ones are those played on the edge set of the complete graph on  $n$  vertices. They will also be our focus, and we will mainly look at the games defined by the hypergraphs  $C_n$ ,  $\mathcal{HAM}_n$ ,  $\mathcal{H}_{H,n}$ ,  $\mathcal{PAN}_n$  and  $\mathcal{COL}_{n,k}$  on the vertex set  $X = E(K_n)$ , with the hyperedges being the edge sets of all spanning trees of  $K_n$ , all Hamilton cycles of  $K_n$ , all copies of a fixed graph  $H$  in  $K_n$ , all pancyclic spanning subgraphs of  $K_n$  and all subgraphs of  $K_n$  with chromatic number larger than  $k$ , respectively. We refer to these games as the connectivity game, the Hamiltonicity game, the  $H$ -game, the pancyclicity game, and the non- $k$ -colourability game, respectively.

For hypergraphs defined on  $E(K_n)$  with winning sets corresponding to the subgraphs of  $K_n$  satisfying a certain graph property, it often happens that for  $n$  large enough Maker has an easy and fast winning strategy. In the  $K_3$ -game, Maker is easily seen to win in 4 moves. In the connectivity game, Maker wins as well: by increasing the size of the connected component of her graph by one at each move, Maker can build a spanning tree

in  $n - 1$  moves. For these reasons, in order to even out this advantage of Maker, Chvátal and Erdős [35] introduced some game bias. Given an integer  $b \geq 1$ , the biased  $(1:b)$  Maker-Breaker game on  $(X, \mathcal{H})$  is the same game as the Maker-Breaker game on  $(X, \mathcal{H})$ , except that, in each move, Breaker can claim up to  $b$  unclaimed elements of the board, while Maker is still only allowed to claim one element per move. Note that these rules imply monotonicity in Breaker's bias. In particular, there must be a *threshold bias*  $b_{\mathcal{H}}$  such that Breaker wins if and only if  $b \geq b_{\mathcal{H}}$ . The size of such threshold biases has been investigated for many standard graph games in recent years (see the books [15, 63]).

Already Chvátal and Erdős [35] proved that the bias threshold for the connectivity game is of order  $\frac{n}{\log n}$ , a quantity which is closely related to the concept of thresholds in random graphs (see Definition 1.1.4). Namely, if Maker and Breaker are replaced by 'random' players who select their edges uniformly at random, the final subgraph of  $K_n$  consisting only of Maker's edges is a random graph  $G(n, M)$  chosen uniformly among all graphs on  $n$  vertices with  $M = \lfloor \frac{1}{b+1} \binom{n}{2} \rfloor$  edges, which is tightly linked to the  $G(n, p)$  model with  $p = \frac{1}{b+1}$ . As the threshold for connectivity in  $G(n, p)$  is  $(1 + o(1)) \frac{\log n}{n}$ , Maker wins the random game if and only if  $\frac{1}{b+1}$  is significantly larger than  $\frac{\log n}{n}$ , i.e. if  $b$  is significantly smaller than  $\frac{n}{\log n}$ , which is exactly the quantity found by Chvátal and Erdős [35]. This observation kicked off what is nowadays known as the *random graph intuition*: for many values of the bias  $b$ , the outcome of the clever game is a.a.s. predicted by the corresponding game where both Maker and Breaker play randomly.

As mentioned, the above intuition holds for the connectivity game. This was extended by Krivelevich [78] to the Hamilton cycle game, by showing that  $b_{\mathcal{H}\mathcal{A}\mathcal{M}_n} = (1 + o(1)) \frac{n}{\log n}$ , with the threshold probability for the existence of a Hamilton cycle being  $(1 + o(1)) \frac{\log n}{n}$ , as proved by Pósa [95] and Koršunov [77], independently. Up to constant factors the same holds for the game  $\text{COL}_{n,k}$  with  $k$  being a constant, where Hefetz, Krivelevich, Stojaković, and Szabó [62] proved that  $b_{\text{COL}_{n,k}}$  is of the order  $\frac{n}{k \log k}$ , while the probability threshold for non- $k$ -colourability is  $(2 + o(1)) \frac{k \log k}{n}$ . Further instances of the random graph intuition can be found in e.g. [16, 36].

On the other hand, the pancyclicity game and the  $H$ -game fail to satisfy the above intuition. Indeed, Ferber, Krivelevich, and Naves [50] proved that the threshold bias  $b_{\mathcal{P}\mathcal{A}\mathcal{N}_n}$  of the pancyclicity game is close to  $\sqrt{n}$ , while Cooper and Frieze [37] proved that the threshold probability for the pancyclicity property is  $(1 + o(1)) \frac{\log n}{n}$ . Bednarska and Łuczak [21] proved that the threshold bias  $b_{\mathcal{H}_{H,n}}$  of the  $H$ -game is of the order  $n^{1/m_2(H)}$  if  $H$  has at least two edges, where  $m_2(H) = \max \left\{ \frac{e(F)-1}{v(F)-2} : F \subseteq H \text{ with } v(F) \geq 3 \right\}$ . However it is well known [23] that the threshold probability for the random graph  $G(n, p)$  to contain a copy of  $H$  is  $n^{-1/m(H)}$ , where  $m(H) = \max \left\{ \frac{e(F)}{v(F)} : F \subseteq H \text{ with } v(F) \geq 1 \right\}$ . Although these two games provide examples for which the random graph intuition is not true, it should be noted that a deep connection to random graphs still exists. In fact, in both cases

it turns out that the results on the threshold biases are linked to *resilience properties* of  $G(n, p)$ .

In the  $(1: b)$  Maker-Breaker game, large Breaker's bias force Maker's graph to be sparse. We now force sparseness by playing multiple games consecutively, where each new game shrinks the board on which the next game is allowed to be played. We provide the precise rules of the game in the next section.

### 1.3.2 Multistage games

Given a hypergraph  $\mathcal{H} = (X, \mathcal{F})$  and a bias  $b \geq 1$ , we define the multistage  $(1: b)$  Maker-Breaker game on  $\mathcal{H}$  as follows.

**Definition 1.3.1** (Multistage  $(1: b)$  Maker-Breaker game). The game proceeds in several stages, with each stage being played as a usual  $(1: b)$  Maker-Breaker game. For convenience we define  $X_0 = X$ ,  $\mathcal{F}_0 = \mathcal{F}$  and  $\mathcal{H}_0 = \mathcal{H}$ . Then, for  $i \geq 1$ , in the  $i$ -th stage, Maker and Breaker play on the board  $X_{i-1}$ , consider the hypergraph  $\mathcal{H}_{i-1} = (X_{i-1}, \mathcal{F}_{i-1})$ , and alternate in turns in which Maker occupies exactly one unclaimed element of  $X_{i-1}$ , and afterwards Breaker occupies up to  $b$  unclaimed elements of  $X_{i-1}$ , with Maker moving first. Once all the elements of  $X_{i-1}$  have been distributed among both players, we let  $X_i \subset X_{i-1}$  be the set of all elements claimed by Maker in stage  $i$ , and we let  $\mathcal{F}_i = \{F \in \mathcal{F}_{i-1} : F \subset X_i\}$  be the set of all remaining winnings sets (from  $\mathcal{F}$ ) that Maker managed to fully occupy in stage  $i$ . Observe that this defines a new hypergraph  $\mathcal{H}_i = (X_i, \mathcal{F}_i)$ . We stop the game the first time that there are no winning sets left anymore.

We stress that, as stated in Definition 1.3.1, each stage is played until all the elements of the board have been claimed by either Maker or Breaker. In order to clarify the rules of the game, we provide an example of a  $(1: 1)$   $K_3$ -game starting on the complete graph on 6 vertices in Table 1.2. The initial hypergraph is  $\mathcal{H}_0 = (X_0, \mathcal{F}_0)$ , where  $X_0 = E(K_6)$  and  $\mathcal{F}_0$  is the collection of all triangles in  $K_6$ , i.e.  $\mathcal{F}_0 = \binom{[6]}{3}$ . The first stage is played as a classical  $(1: 1)$  game on  $\mathcal{H}_0$ . Suppose that, after all elements have been claimed, Maker has claimed (in some order) the edges 12, 15, 16, 23, 24, 26, 34, 56. Then the set of these element is the new board  $X_1$ , while  $\mathcal{F}_1$  is the collection of triangles whose edges are all in  $X_1$ , i.e.  $\mathcal{F}_1 = \{126, 156, 234\}$ . Now the second stage starts and it is played as a classical  $(1: 1)$  game on  $\mathcal{F}_1 = (X_1, \mathcal{F}_1)$ . Suppose that, after all elements have been claimed, Maker has claimed (in some order) the edges 12, 15, 16, 34. Then, the collection of these edges is the new board  $X_2$ . However, there are no triangles with all edges in  $X_2$  and thus  $\mathcal{F}_3 = \emptyset$ . As there are no winning sets left anymore, the multistage game ends. Therefore this game has lasted two stages.

The natural question arising in this new multistage setting is the following one.

**Question 1.3.2.** *How long can Maker delay the stop of a given game?*

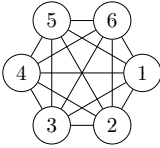
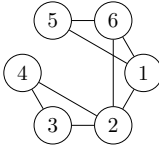
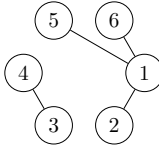
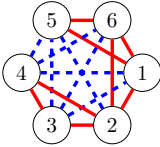
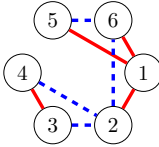
Stage 1	Stage 2	
<p>Beginning</p>  <p><math>\mathcal{F}_0 = \{123, 124, \dots\}</math></p>	<p>Beginning</p>  <p><math>\mathcal{F}_2 = \{126, 156, 234\}</math></p>	 <p><math>\mathcal{F}_3 = \emptyset</math></p>
<p>End</p> 	<p>End</p> 	

Table 1.2: A multistage  $(1:1)$   $K_3$ -game on the edges of  $K_6$  lasting two stages. The red (resp. blue) edges are the ones claimed by Maker (resp. Breaker).

We can formalise Question 1.3.2 by asking to estimate the following threshold parameter.

**Definition 1.3.3** (Threshold parameter  $\tau$ ). Given a hypergraph  $\mathcal{H}$  and an integer  $b \geq 1$ , the threshold parameter  $\tau(\mathcal{H}, b)$  is defined to be the largest number  $s$  such that, in the  $(1:b)$  multistage Maker-Breaker game on the hypergraph  $\mathcal{H}$ , Maker has a strategy to ensure  $\mathcal{F}_s \neq \emptyset$  and thus to play at least  $s$  stages.

### 1.3.3 New results

We answer Question 1.3.2 for several multistage games, by providing the asymptotic value of the threshold  $\tau(\mathcal{H}, b)$  defined in Definition 1.3.3. Before stating our results, we remark that a random graph intuition in this setting corresponds to the assumption that a random game is likely to last as long as a perfectly played game. If a multistage  $(1:b)$  Maker-Breaker game on  $K_n$  is played by two random players, then  $X_i$  is the edge set of a uniform random graph  $G(n, M)$  with roughly  $\left(\frac{1}{b+1}\right)^i \binom{n}{2}$  edges. Therefore, if true, the random graph intuition would suggest that Maker can maintain a spanning connected subgraph, a Hamilton cycle, or a non- $k$ -colourable graph for roughly  $\log_{b+1}(n) - \log_{b+1}(\log n)$  stages. We show this is indeed asymptotically the best that Maker can do.

**Theorem 1.3.4** (Multistage Hamilton cycle game). *If  $b$  is subpolynomial in  $n$ , then*

$$\tau(\mathcal{HAM}_n, b) = (1 + o(1)) \log_{b+1}(n).$$

**Corollary 1.3.5** (Multistage connectivity game). *If  $b$  is subpolynomial in  $n$ , then*

$$\tau(C_n, b) = (1 + o(1)) \log_{b+1}(n).$$

**Theorem 1.3.6** (Multistage non- $k$ -colourability game). *If  $b$  and  $k$  are subpolynomial in  $n$ , and  $k \geq 2$ , then*

$$\tau(\mathcal{COL}_{n,k}, b) = (1 + o(1)) \log_{b+1}(n).$$

However the random graph intuition fails for both the multistage  $H$ -game and the multistage pancyclicity game. If played randomly, they would typically last  $\left(\frac{1}{m(H)} + o(1)\right) \log_{b+1}(n)$  and  $(1 + o(1)) \log_{b+1}(n)$  stages, respectively, while we can show the following, where we recall that  $m(H)$  and  $m_2(H)$  denote the maximum density and the maximum 2-density of  $H$ , respectively.

**Theorem 1.3.7** (Multistage  $H$ -game). *Let  $H$  be a graph. If  $b$  is subpolynomial in  $n$ , then*

$$\tau(\mathcal{H}_{H,n}, b) = \left(\frac{1}{m_2(H)} + o(1)\right) \log_{b+1}(n).$$

**Theorem 1.3.8** (Multistage pancyclicity game). *If  $b$  is subpolynomial in  $n$ , then*

$$\tau(\mathcal{PAN}_n, b) = \left(\frac{1}{2} + o(1)\right) \log_{b+1}(n).$$

We will prove Theorem 1.3.4 and Corollary 1.3.5 in Section 6.2, and Theorems 1.3.6, 1.3.7, and 1.3.8 in Sections 6.3, 6.4, and 6.5, respectively.

## 1.4 Notation

Here we discuss our conventions for notation, together with some general definitions. We may introduce further notation in the subsequent chapters, but this will be used in the respective chapter only.

### Elementary notation

For real numbers  $a, b, c$ , we write  $a = b \pm c$  for  $b - c \leq a \leq b + c$ . Given  $n \in \mathbb{N}$ , we denote by  $[n] = \{1, 2, \dots, n\}$  the set of the first  $n$  positive integers. We denote by  $\log$  the logarithm with base  $e$ . Given a set  $V$  and an integer  $k \geq 1$ , we let  $\binom{V}{k} = \{U \subseteq V : |U| = k\}$ . Finally, given  $s \geq 1$  and sets  $V_1, \dots, V_s$ , when we say that a tuple belongs to  $\prod_{i=1}^s V_i^k$ , we mean that the tuple belongs to  $V_1^k \times \dots \times V_s^k$ , i.e. it is of the form  $(v_{i,j} : 1 \leq i \leq s, 1 \leq j \leq k)$  with  $v_{i,j} \in V_i$  for  $i = 1, \dots, s$ , and  $j = 1, \dots, k$ .

### Graphs

We use standard graph theory notation. A *graph*  $G = (V, E)$  is a pair consisting of a set  $V$ , whose elements are called the *vertices* of  $G$ , and a set  $E \subseteq \binom{V}{2}$ , whose elements are called the *edges* of  $G$ . Although edges are subsets of  $V(G)$  of size two, we write  $uv$  for the edge  $\{u, v\}$ , and we call  $u$  and  $v$  the *endpoints* of the edge. We write  $V(G)$  and  $E(G)$  for the vertex-set and the edge-set of  $G$ , and we denote their sizes by  $v(G)$  and  $e(G)$ , respectively. Furthermore, given two disjoint sets  $A, B \subseteq V(G)$ , we let  $e_G(A)$  be the number of edges of  $G$  with both endpoints in  $A$  and  $e_G(A, B)$  be the number of edges of  $G$  with one endpoint in  $A$  and the other one in  $B$ . When the graph  $G$  is clear from the context, we drop the subscript  $G$ . Given  $I \subset V(G)$ , we say that  $I$  is an *independent set* if  $e(I) = 0$ .

**Neighbourhoods and degrees.** Given a graph  $G$ , and two vertices  $u, v \in V(G)$ , we say that  $v$  is a neighbour of  $u$  if  $uv \in E(G)$ . Given a set  $U \subseteq V(G)$ , we define the *neighbourhood* of  $v$  in  $U$  to be the set of all neighbours of  $v$  which belong to  $U$ , and we denote it by  $N_G(v, U) = \{u \in U : uv \in E(G)\}$ . Furthermore, the degree of  $v$  in  $U$  is given by  $\deg_G(v, U) = |N_G(v, U)|$ . Given a set of vertices  $T \subseteq V(G)$ , we define its *common neighbourhood* in the set  $U$  as  $N_G(T, U) = \bigcap_{v \in T} N_G(v, U)$ , and we denote its size by  $\deg_G(T, U) = |N_G(T, U)|$ . If  $U$  is the whole vertex-set  $V(G)$ , we omit the set  $U$  in the notation above. Similarly, if  $G$  is clear from context, we drop the subscript  $G$ .

**Minimum and maximum degree.** Given a graph  $G$ , we define its *minimum* and *maximum degree* by  $\delta(G) = \min_{v \in V(G)} \{\deg(v)\}$  and  $\Delta(G) = \max_{v \in V(G)} \{\deg(v)\}$ , respectively.

**Subgraphs.** Given two graphs  $G$  and  $H$ , we say that  $H$  is a *subgraph* of  $G$  if  $V(H) \subseteq V(G)$  and  $E(H) \subseteq E(G)$ . We denote this by  $H \subseteq G$ . The subgraph  $H$  is called *spanning* if  $V(H) = V(G)$ . Given two disjoint sets  $A, B \subset V$ , we let  $G[A]$  be the subgraph of  $G$  induced by  $A$ , i.e.  $V(G[A]) = A$  and  $E(G[A]) = \{e \in E(G) : e \subseteq A\}$ . Similarly, we let  $G[A, B]$  be the bipartite subgraph of  $G$  induced by sets  $A$  and  $B$ , i.e.  $V(G[A, B]) = A \cup B$

and  $E(G[A, B]) = \{e \in E(G) : |e \cap A| = 1 = |e \cap B|\}$ . More generally, given pairwise-disjoint sets  $U_1, \dots, U_h \subseteq V$ , we let  $G[U_1, \dots, U_h]$  be the induced  $h$ -partite subgraph  $\bigcup_{1 \leq i < j \leq h} G[U_i, U_j]$  of  $G$ .

**Complete graphs, paths, cycles, stars, cherries.** Given an integer  $r \geq 3$ , we let  $K_r$  denote the *complete graph* on  $r$  vertices, i.e. the graph on vertices  $v_1, \dots, v_r$  and edges  $v_i v_j$  for each  $i, j \in [r]$  with  $i \neq j$ . The graph  $K_r$  is also called the *clique* on  $r$  vertices. Given an integer  $k \geq 2$ , we let  $P_k$  denote the *path* on  $k$  vertices, i.e. the graph on vertices  $v_1, v_2, \dots, v_k$  and edges  $v_1 v_2, v_2 v_3, \dots, v_{k-1} v_k$ . Moreover, for an integer  $k \geq 3$ , we let  $C_k$  denote the *cycle* on  $k$  vertices, i.e. the graph on vertices  $v_1, v_2, \dots, v_k$  and edges  $v_1 v_2, v_2 v_3, \dots, v_{k-1} v_k, v_k v_1$ . An *Hamilton path* (resp. *cycle*) in a graph is a spanning path (resp. cycle). Given an integer  $g \geq 2$ , we define the *star* on  $g + 1$  vertices to be the graph with one vertex of degree  $g$  (this vertex is called the *centre*) and the other vertices of degree one (these vertices are called *leaves*). In particular, a star with  $g = 2$  will be called *cherry*.

**(Complete) Partite graphs.** Given an integer  $k \geq 2$ , we say that a graph  $G$  is *k-partite* with parts  $V_1, \dots, V_k$  if  $V_1, \dots, V_k$  partition  $V(G)$  into  $k$  independent sets. We say that  $G$  is a *k-partite complete graph* if for each  $i, j \in [k]$  with  $i \neq j$  and for each  $v_i \in V_i$  and  $v_j \in V_j$ , the edge  $v_i v_j$  belongs to  $G$ . Further, we say that  $G$  is the complete *balanced k-partite graph* if  $||V_i| - |V_j|| \leq 1$  for each  $i, j \in [k]$ . Given  $\alpha \in (0, 1)$ , we let  $H_\alpha$  denote the complete bipartite graph with parts of size  $\alpha n$  and  $(1 - \alpha)n$ .

**Powers of graphs.** The distance of two vertices  $u$  and  $v$  in a graph  $G$  is the length of the shortest path from  $u$  to  $v$  in  $G$ . Given an integer  $r \geq 1$ , we denote by  $G^r$  the *r-th power* of  $G$ , i.e. the graph obtained from  $G$  by adding edges between all vertices of distance at most  $r$  in  $G$ . Given a copy  $F$  of the square of a path on  $k$  vertices  $P_k^2$ , we let  $v_1, v_2, \dots, v_k$  be an ordering of the vertices of  $F$  such that its edges are precisely  $v_i v_j$ , for each  $i, j$  with  $1 \leq |i - j| \leq 2$ . We call  $(v_2, v_1)$  and  $(v_{k-1}, v_k)$  the *end-tuples* of  $F$ , we refer to  $v_i$  as the *i-th vertex* of  $F$ , and we refer to  $F$  as the square of the path  $v_1, v_2, \dots, v_k$ . For simplicity we will talk about tuples  $(u, v)$  belonging to a set  $V$ , when implicitly meaning to  $V^2$ . Observe that the choice of taking  $(v_2, v_1)$  rather than  $(v_1, v_2)$  as end-tuple is intentional. This is to ensure that for both the end-tuples  $(v_2, v_1)$  and  $(v_{k-1}, v_k)$ , it is always the second vertex, i.e.  $v_1$  and  $v_k$  respectively, to be an endpoint of the path  $v_1, v_2, \dots, v_k$ .

**Factors.** For a fixed graph  $H$ , we say that a graph  $G$  contains an *H-factor* if it contains  $\lfloor v(G)/v(H) \rfloor$  pairwise vertex-disjoint copies of  $H$ .

**Universality.** Given an integer  $k \geq 2$ , let  $\mathcal{F}(n, k)$  denote the family of all graphs on  $n$ -vertices with maximum degree at most  $k$ . We say that a graph  $G$  is *k-universal* if it contains every graph of  $\mathcal{F}(n, k)$  as a subgraph.

**Densities.** Given a graph  $G$ , its *density* is defined by  $d(G) = \frac{e(G)}{v(G)}$  and its *maximum density* by  $m(G) = \max \{d(F) : F \subseteq G \text{ with } v(F) \geq 1\}$ . Moreover, its *1-density* is defined by  $d_1(G) = \frac{e(G)}{v(G)-1}$ , and its *maximum 1-density* by  $m_1(G) = \max \{d_1(F) : F \subseteq G \text{ with } v(F) \geq 2\}$ . Similarly, its *2-density* is defined by  $d_2(G) =$



$\frac{e(G)-1}{v(G)-2}$ , and its *maximum 2-density* by  $m_2(G) = \max \{d_2(F) : F \subseteq G \text{ with } v(F) \geq 3\}$ .

**Graph properties.** A *graph property*  $\mathcal{P}$  is a family of graphs, and thus the notation ' $G \in \mathcal{P}$ ' means that the graph  $G$  is one of the graphs in  $\mathcal{P}$ .

## Hypergraphs

The notion of hypergraph generalises that of graphs, and allows edges to join more than two vertices. More formally, a *hypergraph*  $\mathcal{H} = (V, \mathcal{F})$  is a pair consisting of a set of vertices  $V$  and a set of edges  $\mathcal{F} \subseteq 2^V$ . For an integer  $k \geq 2$ , we say that the hypergraph  $\mathcal{H}$  is *k-uniform* if  $|e| = k$  for each  $e \in \mathcal{F}$ . Observe that a graph is a 2-uniform hypergraph. The definitions given for graphs easily generalise to hypergraphs, but we list below those which require more caution.

**(Minimum) degrees.** Given  $f \subseteq V(\mathcal{H})$  with  $|f| \in [k-1]$ , the *degree* of  $f$ , denoted by  $\deg_{\mathcal{H}}(f)$ , is the number of edges of  $\mathcal{H}$  that  $f$  is contained in, i.e.  $\deg_{\mathcal{H}}(f) = |\{e : e \in E(\mathcal{H}) \text{ and } f \subseteq e\}|$ . For a  $k$ -uniform hypergraph  $\mathcal{H}$  and an integer  $d \in [k-1]$ , we let the *minimum d-degree* of  $\mathcal{H}$  be  $\delta_d(\mathcal{H}) = \min\{\deg_{\mathcal{H}}(f) : f \subseteq V(\mathcal{H}) \text{ and } |f| = d\}$ . Moreover, given two subsets of vertices  $T, U \subseteq V(\mathcal{H})$  we denote by  $\deg(T, U)$  the number of  $U' \subseteq U$  such that  $T \cup U'$  is an edge of the hypergraph.

**$\ell$ -cycles.** Given integers  $k \geq 2$  and  $1 \leq \ell < k$ , a  $k$ -uniform hypergraph is called an  *$\ell$ -cycle* if its vertices can be ordered cyclically such that each of its edges consists of  $k$  consecutive vertices and every two consecutive edges (in the natural order of the edges) share exactly  $\ell$  vertices. In particular,  $(k-1)$ -cycles and 1-cycles are known as *tight cycles* and *loose cycles* respectively.

**Hypergraph collections.** A *hypergraph collection*  $\mathbf{H} = (H_1, \dots, H_m)$  on vertex set  $V$  is a collection of (not necessarily distinct) hypergraphs  $H_i$ ,  $i \in [m]$ , all with  $V$  as vertex set. We call the collection a *k-uniform hypergraph collection* if each hypergraph in the collection is  $k$ -uniform; if  $k = 2$ , we simply call it a *graph collection*. We set  $|\mathbf{H}|$  to denote the number of hypergraphs in the collection  $\mathbf{H} = (H_1, \dots, H_m)$ , that is  $|\mathbf{H}| = m$ . Moreover we denote the *minimum d-degree* of the collection by  $\delta_d(\mathbf{H}) = \min_{i \in [m]} \delta_d(H_i)$ . Given a hypergraph collection  $\mathbf{H} = (H_1, \dots, H_m)$  with vertex set  $V$ , and a set  $U \subset V$ , the collection of hypergraphs  $H_i[U]$ ,  $i \in [m]$  induced on the vertex set  $U$  is denoted by  $\mathbf{H}[U]$ . Similarly, by  $\mathbf{H} \setminus U$  we denote the hypergraph collection  $\mathbf{H}[V \setminus U]$ . We will think of the edges of different hypergraphs in a collection as having different colours. In particular, given a hypergraph collection  $\mathbf{H} = (H_1, \dots, H_m)$ , we consider each hypergraph  $H_i$  to have a *colour i*. Given a subgraph  $H \subset \cup_{i \in [m]} H_i$ , an edge  $e \in E(H)$  can be assigned colour  $i$  if  $e \in E(H_i)$ . When we say  $H \subset \cup_{i \in [m]} H_i$  is *uncoloured*, we emphasise that a colouring has not yet been assigned.

**Transversals.** Given an  $m$ -edge hypergraph  $F$  on  $V$ , we say that a hypergraph collection  $\mathbf{H}$  on vertex set  $V$  has a *transversal copy* of  $F$  if there is a bijection  $\phi: E(F) \rightarrow [m]$  such

that  $e \in H_{\phi(e)}$  for each  $e \in E(F)$ .

### Random graphs

Given  $p \in [0, 1]$ , we let  $G(n, p)$  denote the *binomial random graph* model, which is the probability space of all labeled graphs on the vertex set  $[n]$ , where the probability for such a graph  $G$  to be chosen is  $p^{e(G)}(1-p)^{\binom{n}{2}-e(G)}$ . We can think of  $G(n, p)$  model as the outcome of the following random experiment: each edge of  $K_n$  is present with probability  $p$  independently of all other choices. Similarly, given  $p \in [0, 1]$ , an integer  $k \geq 1$  and  $k$  pairwise-disjoint sets of vertices  $V_1, \dots, V_k$ , we denote by  $G(V_1, \dots, V_k, p)$  the *random  $k$ -partite graph* with parts  $V_1, \dots, V_k$ , where each pair of vertices in two different parts forms an edge with probability  $p$ , independently of the other pairs. Moreover, for  $p \in [0, 1]$  and a set of vertices  $V$ , we denote by  $G(V, p)$  the random graph on  $V$ , where each pair of vertices in  $V$  forms an edge with probability  $p$ , independently of the other pairs. Further, given an integer  $M \geq 1$ , we let  $G(n, M)$  denote the probability space of all labeled graphs with vertex set  $[n]$  and exactly  $M$  edges, together with the uniform distribution. Next, given  $p \in [0, 1]$ , we let  $\vec{G}(n, p)$  denote the *binomial random directed graph* on the vertex set  $[n]$ , where each tuple  $(u, v) \in [n]^2$  of distinct vertices  $u \neq v$  is a directed edge with probability  $p$  independently of all the other choices. If a graph  $G$  is sampled according to one of these models, we write  $G \sim G(n, p)$ ,  $G \sim G(n, M)$  and  $G \sim \vec{G}(n, p)$ , respectively. Finally, for an event  $A = A(n)$  depending on  $n \in \mathbb{N}$ , we say that  $A$  happens *asymptotically almost surely* (a.a.s.) if  $\mathbb{P}[A] \rightarrow 1$  when  $n \rightarrow \infty$ .

### Asymptotic notation

We use standard Landau notation for functions  $f, g : \mathbb{N} \rightarrow \mathbb{R}_{>0}$ . We say that

- $f = O(g)$  if and only if there exist  $n_0 \in \mathbb{N}$  and  $C > 0$  such that  $f(n) \leq Cg(n)$  for all  $n \geq n_0$ ;
- $f = o(g)$  if and only if for every  $\varepsilon > 0$  there exist  $n_0 \in \mathbb{N}$  such that  $f(n) \leq \varepsilon g(n)$  for all  $n \geq n_0$ ;
- $f = \Omega(g)$  if and only if  $g = O(f)$ ;
- $f = \omega(g)$  if and only if  $g = o(f)$ .

We use standard hierarchical notation for constants, writing  $0 < x \ll y < 1$  to mean that there is a fixed positive non-decreasing function on  $(0, 1]$  such that the subsequent statements hold for  $x \leq f(y)$ . Where multiple constants appear in a hierarchy, they are chosen from right to left.

---

## Methods and tools

---

The proofs presented in the thesis use a variety of techniques, and in this chapter we discuss these tools. The majority of the results are well-known in the literature and are stated without proofs. However, those who deviate from the literature or whose exact precise statement has not appeared anywhere are given with a proof.

In Section 2.1, we present Szemerédi's Regularity Lemma together with a few consequences of the definition of regularity. Then we discuss the absorption method in Section 2.2. We state some concentration inequalities in Section 2.4, and we discuss some applications in Sections 2.5 and 2.6. Finally, in Section 2.7, we discuss a notion to describe a collection of copies of a given graph  $K$  which are clustered together.

### 2.1 The regularity method

The regularity method is currently one of the most powerful tools in extremal graph theory. Its key ingredient is Szemerédi's Regularity Lemma, which was originally motivated by a famous conjecture of Erdős and Turán concerning the existence of long arithmetic progressions in sets with positive (upper) density [107, 108]. It was first explicitly stated in its current form in [109], and, since then, it turned out to be essential in the resolution of several long-standing open problems, see [76] for a survey on the topic. We will use the regularity method in Chapters 3 and 4.

The Regularity Lemma relies on the concept of regular pairs, for which we now introduce the relevant terminology. Let  $G = (V, E)$  be a graph and  $A, B \subseteq V$  be two disjoint sets of vertices. Then the *density* of the pair  $(A, B)$  is defined by  $d(A, B) = \frac{e(A, B)}{|A| \cdot |B|}$ . The pair  $(A, B)$  is called  $\varepsilon$ -regular if for all sets  $X \subseteq A$  and  $Y \subseteq B$  with  $|X| \geq \varepsilon|A|$  and  $|Y| \geq \varepsilon|B|$  we have  $|d(A, B) - d(X, Y)| \leq \varepsilon$ . We will often need a better control over the degree of all the vertices. With  $d \in [0, 1]$ , a pair  $(A, B)$  of disjoint sets of vertices is called  $(\varepsilon, d)$ -super-regular if

- for all sets  $X \subseteq A$  and  $Y \subseteq B$  with  $|X| \geq \varepsilon|A|$  and  $|Y| \geq \varepsilon|B|$ , we have  $d(X, Y) \geq d$ ;
- for all  $a \in A$ , we have  $\deg(a) \geq d|B|$ ;
- for all  $b \in B$ , we have  $\deg(b) \geq d|A|$ .

We begin with some simple well-known facts about (super-)regular pairs. The first lemma states that in an  $\varepsilon$ -regular pair of density  $d$ , most vertices in one part have close to the expected number of neighbours in any large enough subset of the other part.

**Lemma 2.1.1** (Minimum Degree Lemma). *Let  $(A, B)$  be an  $\varepsilon$ -regular pair with  $d(A, B) = d$ . Then, for every  $Y \subseteq B$  with  $|Y| \geq \varepsilon|B|$ , the number of vertices from  $A$  with degree into  $Y$  less than  $(d - \varepsilon)|Y|$  is at most  $\varepsilon|A|$ .*

The second lemma states that an  $\varepsilon$ -regular pair of density  $d$  contains a large super-regular pair with not too different parameters.

**Lemma 2.1.2.** *Let  $(A, B)$  be an  $\varepsilon$ -regular pair with  $d(A, B) = d$ . Then there exist  $A' \subseteq A$  and  $B' \subseteq B$  with  $|A'| \geq (1 - \varepsilon)|A|$  and  $|B'| \geq (1 - \varepsilon)|B|$  such that  $(A', B')$  is a  $(2\varepsilon, d - 3\varepsilon)$ -super-regular pair.*

The third lemma, which can be easily proved using Hall's Theorem, states that a super-regular pair with parts of the same size contains a perfect matching.

**Lemma 2.1.3.** *For any  $d > 0$  there exists  $\varepsilon > 0$  such that any  $(\varepsilon, d)$ -super-regular pair  $(U, V)$  with  $|U| = |V|$  contains a perfect matching.*

We are now ready to state the degree form of the Szemerédi's Regularity Lemma, which can be derived from the original version [109].

**Lemma 2.1.4** (Degree form of the Szemerédi's Regularity Lemma [76]). *For every  $\varepsilon > 0$  and integer  $t_0$  there exists an integer  $T > t_0$  such that for any graph  $G$  on at least  $T$  vertices and  $d \in [0, 1]$  there is a partition of  $V(G)$  into  $t_0 < t + 1 \leq T$  sets  $V_0, \dots, V_t$  and a subgraph  $G'$  of  $G$  such that*

- (P1)  $|V_i| = |V_j|$  for all  $1 \leq i, j \leq t$  and  $|V_0| \leq \varepsilon|V(G)|$ ;
- (P2)  $\deg_{G'}(v) \geq \deg_G(v) - (d + \varepsilon)|V(G)|$  for all  $v \in V(G)$ ;
- (P3) the set  $V_i$  is independent in  $G'$  for  $1 \leq i \leq t$ ;
- (P4) for  $1 \leq i < j \leq t$  the pair  $(V_i, V_j)$  is  $\varepsilon$ -regular in  $G'$  and has density either 0 or at least  $d$ .

The sets  $V_1, \dots, V_t$  are also called clusters and we refer to  $V_0$  as the set of exceptional vertices. A partition  $V_0, \dots, V_t$  which satisfies (P1)–(P4) is called an  $(\varepsilon, d)$ -regular partition of  $G$ . Given this partition, we define the  $(\varepsilon, d)$ -reduced graph  $R$  for  $G$ , that is, the graph on vertex set  $[t]$ , in which  $ij$  is an edge if and only if  $(V_i, V_j)$  is an  $\varepsilon$ -regular pair in  $G'$  and has density at least  $d$ .

## 2.2 The absorption method

The absorption method has its origin in a work by Erdős, Gyárfás, and Pyber [46], but was codified as a versatile technique only later by Rödl, Ruciński, and Szemerédi [101]. Nowadays it is an extremely useful tool in packing problems and several variants have been developed to face some of the challenges that have emerged. We will make use of it in Chapter 5.

At a very general high level, the method builds on finding some *absorbing structure* and an *almost perfect packing* in the host graph. Then, whatever is leftover, together with the absorber, can be perfectly packed into. The last step heavily depends on the special property we require from the absorber, as well as on the properties of the leftover maintained through the almost perfect packing process.

We illustrate the method supposing we wish to prove Dirac's theorem: any  $n$ -vertex graph  $G$  with minimum degree at least  $n/2$  contains a Hamilton cycle. The key steps would roughly be as follows.

**Step 1. Set aside a vertex reservoir.**

Let  $R \subseteq V(G)$  be a small subset so that each vertex of  $G$  has about  $|R|/2$  many neighbours in  $R$ .

**Step 2. Find an absorber.**

Show that  $V(G) \setminus R$  contains a small subset  $A$  and vertices  $v, w \in A$  so that for any small enough subset  $L \subseteq V(G) \setminus A$ , we have that  $G[L \cup A]$  contains a spanning path with endpoints  $v$  and  $w$ . The set  $A$  is called a *vertex absorber*.

**Step 3. Almost cover the remainder with long paths.**

Show that all but few vertices of  $V(G) \setminus (A \cup R)$  can be covered by pairwise disjoint long paths in  $G$ .

**Step 4. Build an almost spanning path.**

Using the good minimum degree to  $R$ , connect the paths found in Step 3 via short paths to build a single path  $P$  with endpoints  $v$  and  $w$ , which is vertex-disjoint with  $A \setminus \{v, w\}$ , and covers all but a few vertices of  $V(G) \setminus A$ .

**Step 5. Use the absorber.**

Using the special property of the absorber  $A$ , the set  $L = V(G) \setminus (A \cup V(P))$  can be used together with  $A$  to connect  $v$  and  $w$  via a path  $P'$ . Then  $P \cup P'$  is the desired Hamilton cycle.

The exact way Step 1. to Step 3. are performed is not relevant in this high level description, and thus we do not give further details. However, we remark that it is often the case that the challenge of the method is to define an absorber which has the right property and, at the same time, can be found in the host graph.

## 2.3 The container method

The hypergraph container method can be used to bound the number of finite objects with forbidden substructures, and was developed by Balogh, Morris and Samotij [8], and independently by Saxton and Thomason [102]. It has been used to prove a wealth of results in extremal graph theory, Ramsey theory, additive combinatorics, number theory and discrete geometry. We will use it in Chapter 6 while dealing with the multistage Maker-Breaker  $H$ -game. We recall that the container method has been already used in the context of (classical) Maker-Breaker  $H$ -games, for the first time in [91].

The essence of the method is that the family of independent sets has some structure. Indeed, roughly speaking, a version of the container method proves that, given a hypergraph  $\mathcal{H}$  whose edges are sufficiently evenly distributed, there exists a small number of almost independent sets  $C_1, \dots, C_t$  in  $\mathcal{H}$  such that every independent set  $I$  of  $\mathcal{H}$  is contained in one of the  $C_i$ 's. Here small means that  $t$  is small, and almost independent means that  $e(\mathcal{H}[C_i])$  is small. Of course, there is some interplay between how small  $t$  and  $e(\mathcal{H}[C_i])$  can be: for a better bound on  $t$ , we have to lose on  $e(\mathcal{H}[C_i])$ . We use one of the formulations in [102], which needs the following notation. Given a set  $S$ , we define  $\mathcal{T}_{k,s}(S)$  as the following family of  $k$ -tuples of subsets of  $S$ ,

$$\mathcal{T}_{k,s}(S) = \left\{ (S_1, \dots, S_k) \mid S_i \subseteq S \text{ for } 1 \leq i \leq k \text{ and } \left| \bigcup_{i=1}^k S_i \right| \leq s \right\}.$$

**Theorem 2.3.1** (Theorem 2.3 in [102]). *For any graph  $H$  there exist constants  $n_0, r \in \mathbb{N}$  and  $\delta \in (0, 1)$  such that the following is true. For every  $n \geq n_0$  there exist  $t = t(n)$ , pairwise distinct tuples  $T_1, \dots, T_t \in \mathcal{T}_{r, rn^{2-1/m_2(H)}}(E(K_n))$  and sets  $C_1, \dots, C_t \subseteq E(K_n)$ , such that*

(C1) *each  $C_i$  contains at most  $(1 - \delta) \binom{n}{2}$  edges,*

(C2) *for every  $H$ -free graph  $G$  on  $n$  vertices there exists  $1 \leq i \leq t$  such that  $T_i \subseteq E(G) \subseteq C_i$ , where  $T_i \subseteq E(G)$  means that all sets contained in  $T_i$  are subsets of  $E(G)$ .*

We remark that Theorem 2.3.1 is obtained by applying the container method described above to the  $e(H)$ -uniform hypergraph  $\mathcal{H}$  whose vertices are the edges of  $K_n$  and whose edges are the  $e(H)$ -subsets of  $V(\mathcal{H})$  spanning a copy of  $H$  in  $K_n$ . Indeed, the  $H$ -free graphs on  $n$  vertices correspond to the independent sets of  $\mathcal{H}$ .

## 2.4 Concentration inequalities

In Chapters 3 to 5, we will often need to show that certain random variables are concentrated around their expected value. We start by stating well-known concentration inequalities due to Chernoff (see e.g. [68, Corollaries 2.3 and 2.4] and [67]).

**Lemma 2.4.1** (Chernoff's inequality). *Let  $X$  be the sum of independent Bernoulli random variables, then for any  $\delta \in (0, 1)$  we have*

$$\mathbb{P}\left[|X - \mathbb{E}[X]| \geq \delta \mathbb{E}[X]\right] \leq 2 \exp\left(-\frac{\delta^2}{3} \mathbb{E}[X]\right)$$

*and for any  $k \geq 7 \cdot \mathbb{E}[X]$  we have  $\mathbb{P}[X > k] \leq \exp(-k)$ . More precisely, if  $p$  is the success probability and there are  $n$  summands we get*

$$\mathbb{P}[X \leq \mathbb{E}[X] - \delta n] \leq \exp(-D(p - \delta || p) n),$$

*where  $D(x||y) = x \log(\frac{x}{y}) + (1-x) \log(\frac{1-x}{1-y})$  is the relative entropy.*

Lemma 2.4.1 is particularly useful when the random variable  $X$  is of the form  $X = \sum_{F \in \mathcal{F}} \chi_F$ , where  $\mathcal{F}$  is a collection of pairwise edge-disjoint subgraphs of  $K_n$  and  $\chi_F$  denotes the indicator variable of the event ' $F \subseteq G(n, p)$ '. However, if the subgraphs in  $\mathcal{F}$  are not edge-disjoint, then  $\{\chi_F\}_{F \in \mathcal{F}}$  is not a family of independent random variables anymore and Lemma 2.4.1 is not applicable. In that case, we can use the following Janson's inequality.

**Lemma 2.4.2** (Janson's inequality, Theorem 2.14 in [68]). *Let  $p \in (0, 1)$  and consider a family  $\{H_i\}_{i \in \mathcal{I}}$  of subgraphs of the complete graph on the vertex set  $[n] = \{1, \dots, n\}$ . For each  $i \in \mathcal{I}$ , let  $X_i$  denote the indicator random variable for the event that  $H_i \subseteq G(n, p)$  and, write  $H_i \sim H_j$  for each ordered pair  $(i, j) \in \mathcal{I} \times \mathcal{I}$  with  $i \neq j$  if  $E(H_i) \cap E(H_j) \neq \emptyset$ . Then, for  $X = \sum_{i \in \mathcal{I}} X_i$ ,  $\mathbb{E}[X] = \sum_{i \in \mathcal{I}} p^{e(H_i)}$ ,*

$$\delta = \sum_{H_i \sim H_j} \mathbb{E}[X_i X_j] = \sum_{H_i \sim H_j} p^{e(H_i) + e(H_j) - e(H_i \cap H_j)}$$

*and any  $0 < \gamma < 1$  we have*

$$\mathbb{P}[X \leq (1 - \gamma)\mathbb{E}[X]] \leq \exp\left(-\frac{\gamma^2 \mathbb{E}[X]^2}{2(\mathbb{E}[X] + \delta)}\right).$$

Moreover, if we know both the expected value and the variance of a random variable and they are finite, then we can use the Chebyshev's inequality.

**Lemma 2.4.3** (Chebyshev's inequality [68]). *Let  $X$  be a random variable with finite expected value and variance. Then for any  $a > 0$  we have*

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq a) \leq \frac{\text{Var}[X]}{a^2}.$$

We also state a concentration inequality due to McDiarmid [88], whose present formulation can be found in [83].

**Lemma 2.4.4** (Lemma 6.1 in [83]). *Let  $c > 0$  and let  $f$  be a function defined on the set of subsets of some set  $U$  such that  $|f(U_1) - f(U_2)| \leq c$  whenever  $|U_1| = |U_2| = m$  and  $|U_1 \cap U_2| = m - 1$ . Let  $A$  be a uniformly random  $m$ -subset of  $U$ . Then for any  $\alpha > 0$  we have*

$$\mathbb{P} [ |f(A) - \mathbb{E}[f(A)]| \geq \alpha c \sqrt{m} ] \leq 2 \exp(-2\alpha^2).$$

## 2.5 Embedding results

We give some embedding results that we will use in Chapters 3 and 4. A standard and well-known application of Janson's inequality gives the threshold for the property that any linear sized set of vertices induces a copy of a given graph  $F$  in  $G(n, p)$ .

**Lemma 2.5.1.** *For any graph  $F$  and any  $\delta > 0$ , there exists  $C > 0$  such that the following holds for  $p \geq Cn^{-1/m_1(F)}$ . In the random graph  $G(n, p)$  a.a.s. any set of  $\delta n$  vertices contains a copy of  $F$ .*

Note that  $m_1(P_{k-1}) = 1$  and  $m_1(P_{k-1}^2) = \frac{2k-3}{k-1}$  and, therefore, the bounds on  $p$  given by Lemma 2.5.1 for the containment of a copy of  $P_{k-1}$  and  $P_{k-1}^2$  in any linear sized set are  $p \geq C/n$  and  $p \geq Cn^{-(k-1)/(2k-3)}$ , respectively. Observe that a repeated application of Lemma 2.5.1 allows to find a collection of vertex-disjoint copies of  $F$  covering  $(1 - o(1))n$  vertices of  $G(n, p)$ , i.e.  $n^{-1/m_1(F)}$  is the threshold in  $G(n, p)$  for an almost  $F$ -factor. If we ask for an  $F$ -factor in  $G(n, p)$ , rather than an almost  $F$ -factor, the situation is much more complicated. The corresponding threshold is only known for some graphs  $F$  (see [53, 69]). In particular we state a breakthrough result of Johansson, Kahn, and Vu [69], who determined the threshold for covering all vertices of  $G(n, p)$  with pairwise vertex-disjoint copies of  $F$ , for any strictly 1-balanced graph  $F$ , i.e. those graphs with 1-density strictly larger than that of any proper subgraph.

**Theorem 2.5.2** (Johansson, Kahn, and Vu [69]). *Let  $F$  be a graph such that  $m_1(F') < m_1(F)$  for all  $F' \subseteq F$  with  $F' \neq F$  and  $v(F') \geq 2$ . Then the threshold for an  $F$ -factor in  $G(n, p)$  is  $n^{-1/m_1(F)} (\log n)^{1/e(F)}$ .*

Note that, for any  $r \geq 3$ , the clique  $K_r$  is a strictly 1-balanced graph, and so is the square of a path on  $k + 1$  vertices. In particular, Theorem 2.5.2 holds for both of them.

While embedding the square of a Hamilton cycle in the randomly perturbed graph, we will rely on a decomposition in random and deterministic edges (see Section 3.1.1). In particular, in the proof of Theorem 1.1.16, we will find the squares of short paths in  $G(n, p)$ . However, it will be often the case that we want to find combinations of these squares in  $G(n, p)$  such that their vertices satisfy some additional constraints; for that, we will use the following lemma.



**Lemma 2.5.3.** *For all integers  $s \geq 1$  and  $k \geq 2$ , and any  $0 < \eta \leq 1$ , there exists  $C > 0$  such that the following holds for  $p \geq Cn^{-(k-1)/(2k-3)}$ . Let  $V$  be a vertex set of size  $n$ ,  $V_1, \dots, V_s$  not necessarily disjoint subsets of  $V$ , and  $H$  be a collection of pairwise distinct tuples from  $\prod_{i=1}^s V_i^k$ . Then a.a.s. revealing  $\Gamma = G(n, p)$  on  $V$  gives the following. For any choice of  $W_i \subseteq V_i$  with  $i = 1, \dots, s$  such that  $H' = H \cap \prod_{i=1}^s W_i^k$  has size at least  $\eta n^{sk}$ , there is a tuple  $(v_{i,j} : 1 \leq i \leq s, 1 \leq j \leq k)$  in  $H'$  with pairwise distinct vertices  $v_{i,j} \in W_i$  for  $i = 1, \dots, s, j = 1, \dots, k$ , such that in  $\Gamma$  for  $i = 1, \dots, s$  we have the square of a path on  $v_{i,1}, \dots, v_{i,k}$  and for  $i = 1, \dots, s-1$  we have the edge  $v_{i,k}v_{i+1,1}$ .*

Observe that the structure we get from Lemma 2.5.3 in  $G(n, p)$  is given by  $s$  copies of the square of a path on  $k$  vertices and  $s-1$  additional edges joining two consecutive such copies. Moreover when  $k = 2$  the structure is a path on  $2s$  vertices. In applications, we will often define several collections of tuples  $H_j \subseteq \prod_{i=1}^s V_{j,i}^k$  for  $j = 1, \dots, m$ , and apply Lemma 2.5.3 to  $H = \bigcup_{j=1}^m H_j$ , where it is implicit that we apply it with  $V_i = \bigcup_{j=1}^m V_{j,i}$ . Also, we stress that, for a fixed  $H$  and a typical revealed  $G(n, p)$ , the conclusion of the lemma holds for any large enough subset of the form  $H \cap \prod_{i=1}^s W_i^k$  with  $W_i \subseteq V_i$ . In particular, we will be able to claim the existence of a tuple in each subcollection  $H_j$ , again provided they have the right size. This lemma is again a standard application of Janson's inequality. As this precise formulation is not stated in the literature, we include its proof.

*Proof of Lemma 2.5.3.* Let  $s \geq 1$  and  $k \geq 2$  be integers and  $0 < \eta \leq 1$ . Moreover let  $C \geq 2^6(sk)^{2sk}\eta^{-2}$  and  $p \geq Cn^{-(k-1)/(2k-3)}$ .

Let  $V$  be a vertex set of size  $n$ ,  $V_i$  be a subset of  $V$  for  $i = 1, \dots, s$ , and  $H$  be a collection of pairwise distinct tuples from  $\prod_{i=1}^s V_i^k$ . Let  $W_i \subseteq V_i$  for each  $i = 1, \dots, s$  and assume  $H' = H \cap \prod_{i=1}^s W_i^k$  has size at least  $\eta n^{sk}$ . Since the number of tuples from  $\prod_{i=1}^s V_i^k$  which contain a vertex more than once is  $O(n^{sk-1})$ , there are at least  $\frac{\eta}{2}n^{sk}$  tuples of  $H'$  such that their vertices are pairwise distinct. We restrict our analysis to the set of those tuples, which, abusing notation, we still denote by  $H'$ .

For each tuple  $(v_{i,j} : 1 \leq i \leq s, 1 \leq j \leq k)$  in  $H'$ , we consider the graph with vertex set  $V$  and the following edges. For  $i = 1, \dots, s$  we have the square of the path on  $v_{i,1}, \dots, v_{i,k}$  and for  $i = 1, \dots, s-1$  we have the edge  $v_{i,k}v_{i+1,1}$ . This gives a family  $\{H_i\}_{i \in [|H'|]}$  of graphs with vertex set  $V$  and, using the same notation as in Lemma 2.4.2, a collection of random variables  $\{X_i\}_{i \in [|H'|]}$ . Note that for each  $i = 1, \dots, s$ , we have  $e(H_i) = s(2k-3) + (s-1) = 2s(k-1) - 1$ , and thus, for  $X = \sum_{i \in [|H'|]} X_i$ , we have  $\mathbb{E}[X] = |H'|p^{2s(k-1)-1} \geq \sqrt{C}n$ . To compute the quantity  $\Delta[X] = \sum_{H_i \sim H_j} p^{e(H_i)+e(H_j)-e(H_i \cap H_j)}$ , we split the sum according to the number of vertices in the intersection  $E(H_i \cap H_j)$ . Suppose  $H_i$  and  $H_j$  intersect in  $m$  vertices. Then  $2 \leq m \leq sk-1$  and the largest size  $\tilde{e}(m)$  of the intersection  $E(H_i \cap H_j)$

is

$$\tilde{e}(m) = \begin{cases} \frac{m}{k}(2k-3) + \frac{m}{k} - 1, & \text{if } m \equiv 0 \pmod{k} \\ \lfloor \frac{m}{k} \rfloor (2k-3) + \lfloor \frac{m}{k} \rfloor, & \text{if } m \equiv 1 \pmod{k} \\ \lfloor \frac{m}{k} \rfloor (2k-3) + \lfloor \frac{m}{k} \rfloor + 2(m - k \lfloor \frac{m}{k} \rfloor) - 3, & \text{otherwise.} \end{cases}$$

In particular, observing that  $\tilde{e}(m) = 2m - 3$  if  $m < k$  (as  $m \geq 2$  we are in the third case) and  $\tilde{e}(m) \geq 2m - 2\frac{m}{k} - 1$  if  $m \geq k$  (the inequality follows from  $\lfloor \frac{m}{k} \rfloor \geq \frac{m}{k} - 1$ ), we can conclude that  $p^{-\tilde{e}(m)} n^{-m} \leq C^{-1} n^{-1}$  for each  $2 \leq m \leq sk - 1$ . Therefore,

$$\begin{aligned} \Delta[X] &\leq \sum_{m=2}^{sk-1} m! \binom{sk}{m}^2 n^{2sk-m} p^{[2s(k-1)-1] + [2s(k-1)-1] - \tilde{e}(m)} \\ &\leq \sum_{m=2}^{sk-1} (sk)^{2m} \frac{n^{2sk-m}}{|H'|^2} \mathbb{E}^2[X] p^{-\tilde{e}(m)} \\ &\leq 4(sk)^{2sk-2} \eta^{-2} \sum_{m=2}^{sk-1} \mathbb{E}^2[X] p^{-\tilde{e}(m)} n^{-m} \\ &\leq 4(sk)^{2sk-1} \eta^{-2} C^{-1} \mathbb{E}^2[X] n^{-1} \leq \frac{1}{8} s^{-1} \mathbb{E}^2[X] n^{-1}, \end{aligned}$$

where in the first inequality we used that there are at most  $m! \binom{sk}{m}^2 n^{2sk-m}$  choices for  $H_i$  and  $H_j$  intersecting in  $m$  vertices, in the second we used  $\mathbb{E}[X] = |H'| p^{2s(k-1)-1} \geq \sqrt{C} n$ , and in the third we used  $n^{sk}/|H'| \leq 2/\eta$ . Then with Lemma 2.4.2 applied with  $\gamma = 2^{-1/2}$ , we get that the probability that none of the graphs of the family  $\{H_i\}_{i \in [|H'|]}$  appears in  $G(n, p)$  is bounded from above by

$$\begin{aligned} \exp\left(-\frac{\mathbb{E}^2[X]}{4(\mathbb{E}[X] + \Delta[X])}\right) &\leq \exp\left(-\frac{1}{8} \min\left\{\mathbb{E}[X], \frac{\mathbb{E}^2[X]}{\Delta[X]}\right\}\right) \\ &\leq \exp\left(-\frac{1}{8} \min\{\sqrt{C} n, 8sn\}\right) \leq \exp(-sn). \end{aligned}$$

Using a union bound over the at most  $2^{sn}$  choices for the  $s$  subsets  $W_i$  with  $i = 1, \dots, s$ , we conclude that the lemma holds.  $\square$

Next, we state a theorem which will allow us to find a directed Hamilton cycle in the binomial random directed graph  $\vec{G}(n, p)$  and will be used in the proof of Theorem 1.1.17 in Chapter 3.

**Theorem 2.5.4** (Angluin and Valiant [5]). *There exists  $C > 0$  such that for  $p \geq C \log n/n$  a.a.s.  $\vec{G}(n, p)$  has a directed Hamilton cycle.*

Finally, we prove the following lemma which allows us to find specific triangles in a dense graph with additional random edges. This will be used in Chapter 4.

**Lemma 2.5.5.** *For any  $d > 0$  there exists  $C > 0$  such that the following holds. Let  $U, V, W$  be three sets of vertices of size  $n$ ,  $G$  be a bipartite graph on  $(U, W)$  with  $e(G) \geq dn^2$ ,*

$p > C/n$  and  $G(U \cup W, V, p)$  be the random bipartite graph. Then with probability at least  $1 - 2^{-4n/d}$  there is a triangle in  $G \cup G(U \cup W, V, p)$  with one vertex in each of  $U, V, W$ .

*Proof of Lemma 2.5.5.* Let  $d > 0$ ,  $C > 68/d^3$ , and  $p > C/n$ . Let  $\mathcal{I} = E_G(U, W) \times V$  and, for each  $i = (uw, v) \in \mathcal{I}$ , let  $H_i$  be the path  $uvw$  on the three vertices  $u \in U$ ,  $v \in V$  and  $w \in W$ . We want to apply Lemma 2.4.2 to the family  $\{H_i\}_{i \in \mathcal{I}}$ . Using the same notation, we have  $\mathbb{E}[X] = |\mathcal{I}|p^2 \geq dn^3p^2$  and  $\delta \leq n^2(2n)np^3$ , as for  $i = (uw, v)$  and  $j = (u'w', v')$  with  $i, j \in \mathcal{I}$  and  $i \neq j$  we have  $H_i \sim H_j$  if and only if  $v = v'$  and precisely one of the equalities  $u = u'$  and  $w = w'$  holds. Then the Janson's inequality with  $\gamma = 1/2$  gives

$$\mathbb{P}\left[X \leq \frac{\mathbb{E}[X]}{2}\right] \leq \exp\left(-\frac{\gamma^2}{2} \frac{dn^3p^2}{1+2np/d}\right) \leq \exp\left(-\frac{1}{17}d^2n^2p\right) \leq 2^{-4n/d},$$

as  $p > C/n$  and  $C > 68/d^3$ . Thus with probability at least  $1 - 2^{-4n/d}$  we have  $X > \mathbb{E}[X]/2$  and there is at least one path  $H_i$  on vertices  $u, v, w$  for some  $i = (uw, v) \in \mathcal{I}$  in  $G(n, p)$ . As  $uw$  is an edge of  $G$  by definition of  $\mathcal{I}$ , we get a triangle in  $G \cup G(U \cup W, V, p)$  with one vertex in each of  $U, V, W$ , as required.  $\square$

## 2.6 Preserving minimum degrees

In this section, we collect all the results which show that performing certain operations on a (hyper)graph (collection) does not change its minimum degree too much. We will use them in Chapter 5.

We begin with a trivial observation and a simple proposition.

**Observation 2.6.1.** *Let  $\mathbf{H}$  be a  $k$ -uniform hypergraph collection on  $V$  with  $\delta_d(\mathbf{H}) \geq \delta n^{k-d}$ . Let  $S \subseteq V$  with  $|S| \leq \zeta n$ . Then  $\mathbf{H} \setminus S$  has minimum  $d$ -degree at least  $(\delta - \zeta)n^{k-d}$ .*

*Proof of Observation 2.6.1.* Let  $H$  be a hypergraph in  $\mathbf{H}$  and let  $D$  be a set of size  $d$  disjoint with  $S$ . For any vertex  $v$  not in  $D$ , the set  $D \cup \{v\}$  can have degree at most  $n^{k-d-1}$ . Therefore, there are at most  $\zeta n^{k-d}$  edges containing  $D$  which are also incident to  $S$ . Hence,  $H \setminus S$  has minimum  $d$ -degree at least  $(\delta - \zeta)n^{k-d}$ , implying the observation.  $\square$

**Proposition 2.6.2.** *Let  $0 \leq \alpha, \delta \leq 1$ , and  $d, k, m, n \in \mathbb{N}$ , with  $1 \leq d \leq k - 1$ . Let  $\mathbf{H}$  be a  $k$ -uniform hypergraph collection on vertex set  $[n]$  with  $|\mathbf{H}| = m$  and  $\delta_d(\mathbf{H}) \geq \delta n^{k-d}$ . Let  $\mathcal{K}$  be the  $k$ -uniform hypergraph with vertex set  $[n]$ , where  $e$  is an edge of  $\mathcal{K}$  if  $e \in E(H_i)$  for at least  $\alpha m$  values of  $i \in [m]$ . Then,  $\delta_d(\mathcal{K}) \geq (\delta - \alpha)n^{k-d}$ .*

*Proof of Proposition 2.6.2.* For each  $d$  pairwise distinct vertices  $v_1, \dots, v_d \in [n]$ , we have

$$m \cdot \delta n^{k-d} \leq \sum_{i \in [m]} \deg_{H_i}(v_1, \dots, v_d) \leq m \cdot \deg_{\mathcal{K}}(v_1, \dots, v_d) + n^{k-d} \cdot \alpha m,$$

and therefore  $\deg_{\mathcal{K}}(v_1, \dots, v_d) \geq (\delta - \alpha)n^{k-d}$ . Thus,  $\delta_d(\mathcal{K}) \geq (\delta - \alpha)n^{k-d}$ , as wanted.  $\square$

Next we show that the vertex set of a hypergraph collection can be partitioned into linear sized sets, each preserving good minimum degree conditions in each of the hypergraphs of the collection.

**Lemma 2.6.3.** *Let  $1/n \ll \alpha, \beta$  and  $m \leq n^2$ . Let  $t \in \mathbb{N}$  and  $n_1, \dots, n_t \geq \beta n$  be integers such that  $\sum_{i=1}^t n_i = n$ . Let  $\mathbf{H}$  be a  $k$ -uniform hypergraph collection on vertex set  $[n]$  with  $|\mathbf{H}| = m$  and  $\delta_d(\mathbf{H}) \geq (\delta + \alpha)n^{k-d}$ . Then there exists a partition of  $[n]$  into  $V_1, \dots, V_t$  with  $|V_i| = n_i$  for  $i \in [t]$  such that any  $S \in \binom{[n]}{d}$  has degree at least  $(\delta + \alpha/2)n_i^{k-d}$  into  $V_i$  with respect to any of the  $m$  hypergraphs in  $\mathbf{H}$ .*

We will show that a partition chosen uniformly at random has the properties required from Lemma 2.6.3 with high probability. For that, we use the following consequence of the McDiarmid inequality.

**Lemma 2.6.4.** *Let  $k, \ell \in \mathbb{N}$ ,  $0 < \delta' < \delta < 1$  and  $1/n, 1/\ell \ll 1/k, \delta - \delta'$ . Let  $H$  be a  $k$ -uniform  $n$ -vertex hypergraph with vertex set  $V$  and suppose that  $\deg(S, V) \geq \delta n^{k-d}$  for each  $S \in \binom{V}{d}$ . Let  $A \subseteq V$  be a vertex set of size  $\ell$  chosen uniformly at random. Then, for every  $T \in \binom{V}{d}$  we have*

$$\mathbb{P}[\deg(T, A) < \delta' \ell^{k-d}] \leq 2 \exp(-\ell(\delta - \delta')^2/2).$$

*Proof.* Let  $f : \mathcal{P}(V) \rightarrow \mathbb{R}$  be defined by  $f(X) = \deg(T, X)$  for each  $X \subseteq V$ , and set  $\varepsilon = (\delta - \delta')/(2\delta) < 1/2$ . Observe that  $|f(U_1) - f(U_2)| \leq \ell^{k-d-1}$  for any  $U_1, U_2 \in \mathcal{P}(V)$  with  $|U_1| = |U_2| = \ell$  and  $|U_1 \cap U_2| = \ell - 1$ . Given an edge  $e$  with  $T \subseteq e$ , the probability that  $e \setminus T$  is contained in  $A$  is at least  $\frac{\binom{n-k}{\ell-k+d}}{\binom{n}{\ell}} \geq (1-\varepsilon) \frac{\ell^{k-d}}{n^{k-d}}$ , where we used  $1/n, 1/\ell \ll 1/k, \delta - \delta'$ . So by linearity of expectation we have  $\mathbb{E}[f(A)] \geq \delta(1 - \varepsilon)\ell^{k-d} = \frac{\delta + \delta'}{2}\ell^{k-d}$ . We can then apply Lemma 2.4.4 with  $c = \ell^{k-d-1}$ ,  $m = \ell$  and  $\alpha = \sqrt{\ell}(\delta - \delta')/2$ , and get that  $\mathbb{P}[f(A) < \delta' \ell^{k-d}] \leq 2 \exp(-\ell(\delta - \delta')^2/2)$ , as desired.  $\square$

We are now ready to prove Lemma 2.6.3.

*Proof of Lemma 2.6.3.* Pick a partition of  $[n]$  into  $V_1 \cup \dots \cup V_t$  uniformly at random from all partitions which satisfy  $|V_i| = n_i$  for all  $i \in [t]$ . Then by Lemma 2.6.4, we have the probability that there are  $i \in [t]$ ,  $j \in [m]$  and  $S \in \binom{[n]}{d}$  such that  $\deg_{H_j}(S, V_i) < (\delta + \alpha/2)n_i^{k-d}$  is at most  $t \cdot m \cdot \binom{n}{d} \cdot 2 \cdot \exp(-\alpha^2 \beta n/8) = o(1)$ , where we have used that  $n_i \geq \beta n$  for each  $i \in [t]$ ,  $1/n \ll \alpha, \beta$ , and  $m \leq n^2$ . Therefore there exists a partition with the desired properties.  $\square$

## 2.7 $K$ -collections

While proving Theorem 1.3.7 in Chapter 6, and given a graph  $K$ , we will need a notion to describe a collection of copies of  $K$  in a graph  $G$ , which are *clustered* together. Stojaković and Szabó [106], while analysing games on random graphs, introduced this notion when  $K$  is a clique. Here we generalise their definition and extend their related observations to general graphs.

**Definition 2.7.1** ( $K$ -collection). Let  $G$  and  $K$  be graphs. We define the auxiliary graph  $G_K$  to be the graph with vertices corresponding to the copies of  $K$  in  $G$ , and two vertices being adjacent if the corresponding copies of  $K$  have at least two vertices in common. Let  $\mathcal{K} = \{K_1, \dots, K_s\}$  be the family of copies of  $K$  in  $G$  corresponding to a connected component of  $G_K$ . Then the subgraph of  $G$  induced by  $\bigcup_{i \in [s]} K_i$  is called a  $K$ -collection, and we denote its vertex set and its number of vertices by  $V(\mathcal{K}) = \bigcup_{i \in [s]} V(K_i)$  and  $v(\mathcal{K}) = |V(\mathcal{K})|$ , respectively.

**Definition 2.7.2** ( $s$ -bunch). Let  $(K_1, \dots, K_s)$  be a sequence of copies of  $K$ . Then  $\bigcup_{i \in [s]} K_i$  is called an  $s$ -bunch if  $V(K_i) \setminus \left(\bigcup_{j \in [i-1]} V(K_j)\right) \neq \emptyset$  and  $\left|V(K_i) \cap \left(\bigcup_{j \in [i-1]} V(K_j)\right)\right| \geq 2$ , for each  $i = 2, \dots, s$ .

It is easy to observe that every large enough collection contains a large bunch.

**Claim 2.7.3.** Let  $G$  be a graph and  $t \in \mathbb{N}$ . Then every  $K$ -collection  $\mathcal{K}$  of  $G$  on at least  $tv(K)$  vertices contains an  $s$ -bunch  $B$  of copies of  $K$  with  $s \geq t$  and  $tv(K) \leq v(B) \leq (t+1)v(K)$ .

*Proof of Claim 2.7.3.* We start by taking any copy of  $K$  in  $\mathcal{K}$  and then construct the bunch recursively as follows. If  $\bigcup_{i \in [m]} K_i$  is an  $m$ -bunch of copies of  $K$ , we select another copy  $K_{m+1}$  of  $K$  in the collection  $\mathcal{K}$ , such that  $V(K_{m+1}) \setminus \left(\bigcup_{j \in [m]} V(K_j)\right) \neq \emptyset$  and  $\left|V(K_{m+1}) \cap \left(\bigcup_{j \in [m]} V(K_j)\right)\right| \geq 2$ . Note that this will give an  $(m+1)$ -bunch of copies of  $K$ . Since the family  $\mathcal{K}$  corresponds to a connected component of the auxiliary graph  $G_K$ , we are able to find such new copy of  $K$  if  $V(\mathcal{K}) \setminus \left(\bigcup_{i \in [m]} V(K_i)\right) \neq \emptyset$ , i.e. until we cover all the vertices of  $\mathcal{K}$ . In particular, since  $v(\mathcal{K}) \geq tv(K)$ , we can construct an  $s$ -bunch  $B = \bigcup_{i \in [s]} K_i$  of copies of  $K$  with  $tv(K) \leq v(B) < (t+1)v(K)$ . Moreover, since  $tv(K) \leq v(B) \leq v(K) + (s-1)(v(K)-2)$ , we get  $s \geq t$ .  $\square$

Observe that for the  $s$ -bunch where any two copies of  $K$  intersect in the same two adjacent vertices, we have  $d(B) = \frac{e(K)+(s-1)(e(K)-1)}{v(K)+(s-1)(v(K)-2)}$ , which tends to  $\frac{e(K)-1}{v(K)-2} = m_2(K)$  as  $s$  tends to infinity. Using a similar argument as in [106], we show that this is best possible in the following sense.

**Claim 2.7.4.** Let  $K$  be a graph such that  $d_2(K) = m_2(K)$ ,  $s \in \mathbb{N}$  and  $\delta > 0$  such that for all  $x \geq s-1$  we have  $\frac{e(K)+m_2(K)x}{v(K)+x} \geq m_2(K) - \delta$ . Then for any  $s$ -bunch  $B$  of copies of  $K$ , we have  $d(B) \geq m_2(K) - \delta$ .

*Proof of Claim 2.7.4.* Let  $B = \bigcup_{i \in [s]} K_i$  be any  $s$ -bunch of copies of  $K$ . First observe that for every  $S \subset K$  we have  $\frac{e(S)-1}{v(S)-2} \leq \frac{e(K)-1}{v(K)-2}$ , which can be rearranged as  $\frac{e(K)-e(S)}{v(K)-v(S)} \geq \frac{e(K)-1}{v(K)-2} = m_2(K)$ , which in turn gives

$$e(K) - e(S) \geq m_2(K) \cdot (v(K) - v(S)). \quad (2.7.1)$$

Setting  $S_i = K_i \cap (\bigcup_{j \in [i-1]} K_j)$  for each  $i = 2, \dots, s$ , we have

$$\begin{aligned} d(B) &= \frac{e(B)}{v(B)} = \frac{e(K) + \sum_{i \geq 2} (e(K_i) - e(S_i))}{v(K) + \sum_{i \geq 2} (v(K_i) - v(S_i))} \\ &\stackrel{(2.7.1)}{\geq} \frac{e(K) + m_2(K) \sum_{i \geq 2} (v(K_i) - v(S_i))}{v(K) + \sum_{i \geq 2} (v(K_i) - v(S_i))} \geq m_2(K) - \delta, \end{aligned}$$

where the last inequality follows from the assumption on  $\delta$  and as  $\sum_{i \geq 2} (v(K) - v(S_i)) \geq s - 1$ . □

## The square of a Hamilton cycle in randomly perturbed graphs

In this chapter, we discuss our results related to the perturbed threshold for the containment of the square of a Hamilton cycle, namely Theorems 1.1.14, 1.1.16 and 1.1.17. The main Theorem 1.1.14 follows easily from the stability Theorem 1.1.16, the extremal Theorem 1.1.17, and the lower bounds in Proposition 1.1.15. Therefore we prove Theorem 1.1.14 already here.

*Proof of Theorem 1.1.14.* As already explained in the introduction, the cases  $\alpha = 0$  and  $\alpha > 1/2$  follow from known results, and the case  $\alpha = 1/2$  will follow from monotonicity of the perturbed threshold, once we will have determined the threshold in the range  $\alpha < 1/2$ . Therefore we can assume  $\alpha \in \left[\frac{1}{k+1}, \frac{1}{k}\right)$  for some integer  $k \geq 2$ .

Let  $\alpha \in \left(\frac{1}{k+1}, \frac{1}{k}\right)$ , then  $\alpha = \frac{1}{k+1} + \eta$  for some  $0 < \eta < \frac{1}{k(k+1)}$ . Let  $C$  be given by Theorem 1.1.16 on input  $k$  and  $\beta = \eta/5$ . Let  $G$  be an  $n$ -vertex graph with  $\delta(G) \geq \alpha n$  and observe that  $G$  cannot be  $(1/(k+1), \beta)$ -stable, due to its minimum degree condition. Therefore, by Theorem 1.1.16,  $G \cup G(n, p)$  a.a.s. contains the square of a Hamilton cycle, provided  $p \geq Cn^{-(k-1)/(2k-3)}$ . Together with the lower bound in Proposition 1.1.15(i), this shows  $\hat{p}_\alpha = n^{-(k-1)/(2k-3)}$ .

Now we turn to the case  $\alpha = \frac{1}{k+1}$ . Let  $\beta > 0$  and  $C > 0$  be given by Theorem 1.1.17 on input  $k$ ; we can assume  $\beta < 1/(6k)$ . Let  $G$  be an  $n$ -vertex graph with  $\delta(G) \geq \frac{1}{k+1}n$ . If  $G$  is  $(\frac{1}{k+1}, \beta)$ -stable, then, by Theorem 1.1.17,  $G \cup G(n, p)$  a.a.s. contains the square of a Hamilton cycle, provided  $p \geq Cn^{-(k-1)/(2k-3)} \log^{1/(2k-3)} n$ . Otherwise  $G$  is not  $(\frac{1}{k+1}, \beta)$ -stable and, under the same bound on  $p$ , the same conclusion holds by Theorem 1.1.16. Together with the lower bound in Proposition 1.1.15(ii), this shows  $\hat{p}_{\frac{1}{k+1}} = n^{-(k-1)/(2k-3)} \log^{1/(2k-3)} n$ .  $\square$

The proof of Proposition 1.1.15 is standard, and thus we only give a sketch.

*Sketch of the proof Proposition 1.1.15.* Let  $A$  and  $B$  be the vertex classes of  $H_\alpha$  of size  $\alpha n$  and  $(1 - \alpha)n$ , respectively.

- (i) Let  $\frac{1}{k+1} < \alpha < \frac{1}{k}$ , take  $0 < c < \frac{1/k-\alpha}{2}$ , and observe that in  $B$  there are a.a.s. at most  $2cn$  copies of  $P_k^2$  (by an upper tail bound on the distribution of small subgraphs [110]). Assume for a contradiction that there is an embedding of  $C_n^2$  into  $H_\alpha \cup G(n, p)$ . Then  $H_\alpha \cup G(n, p)$  must contain  $\frac{n}{k}$  vertex-disjoint copies of  $P_k^2$ , and only at most  $|A| = \alpha n$  of them have a vertex in  $A$ . Therefore there must be at least  $\frac{n}{k} - \alpha n > 2cn$  copies of  $P_k^2$  in  $B$ , where the inequality follows from the choice of  $c$ . This gives a contradiction.
- (ii) Let  $\alpha = \frac{1}{k+1}$  and  $c = \frac{1}{4k}$ . A.a.s. (by the first moment method)  $B$  contains at most  $n^{1-2c}$  copies of  $P_{k+1}^2$ , and a.a.s. (by the second moment method) at least  $n^{1-c}$  vertices from  $B$  are not contained in any copy of  $P_k^2$  within  $B$ . Assume for a contradiction that there is an embedding of  $C_n^2$  into  $H_\alpha \cup G(n, p)$ . Then  $H_\alpha \cup G(n, p)$  must contain a  $P_{k+1}^2$ -factor. Since  $|B| = k|A|$ , the average size of the intersection of a copy of  $P_{k+1}^2$  in such a factor with  $B$  would be  $k$ . However, given the restrictions above, it is not possible to cover the vertices of  $G(n, p)[B]$  with a family of squares of paths whose average size is  $k$ . This gives a contradiction.  $\square$

We are left to prove Theorems 1.1.16 and 1.1.17. Before doing that in Sections 3.2 and 3.3 respectively, we discuss the idea of our embedding strategy and give an overview of both proofs, together with some auxiliary lemmas. An important novel ingredient in our proofs is an embedding lemma in randomly perturbed graphs (Lemma 3.1.3).

### 3.1 Proof overview of the extremal and the non-extremal case

We recall that, given a copy  $F$  of the square of a path on  $k$  vertices, we let  $v_1, v_2, \dots, v_k$  be an ordering of the vertices of  $F$  such that its edges are precisely  $v_i v_j$ , for each  $i, j$  with  $1 \leq |i - j| \leq 2$ . We call  $(v_2, v_1)$  and  $(v_{k-1}, v_k)$  the *end-tuples* of  $F$ , we refer to  $v_i$  as the  $i$ -th vertex of  $F$ , and we refer to  $F$  as the square of the path  $v_1, v_2, \dots, v_k$ . For simplicity we will talk about tuples  $(u, v)$  belonging to a set  $V$ , when implicitly meaning to  $V^2$ . We remark again that the choice of taking  $(v_2, v_1)$  rather than  $(v_1, v_2)$  as end-tuple is intentional. This is to ensure that for both the end-tuples  $(v_2, v_1)$  and  $(v_{k-1}, v_k)$ , it is always the second vertex, i.e.  $v_1$  and  $v_k$  respectively, to be an endpoint of the path  $v_1, v_2, \dots, v_k$ .

#### 3.1.1 Strategy

Let  $G$  be any  $n$ -vertex graph with minimum degree  $\alpha n$  and  $\alpha \in \left[\frac{1}{k+1}, \frac{1}{k}\right)$ . Our goal is to find the square of a Hamilton cycle  $C_n^2$  in the perturbed graph  $G \cup G(n, p)$  and therefore we will use a *decomposition* of  $E(C_n^2)$  into ‘deterministic edges’ (to be embedded to  $G$ ) and ‘random edges’ (to be embedded to  $G(n, p)$ ). To get the square of a path we would like vertex-disjoint copies  $F_1, \dots, F_t$  of  $P_k^2$  in the random graph  $G(n, p)$  such that the following holds. For each  $i = 1, \dots, t-1$ , if we denote by  $(y_i, x_i)$  and  $(u_i, w_i)$  the end-tuples of  $F_i$ , then  $w_i x_{i+1}$  is also an edge in  $G(n, p)$ . Moreover, there exist  $t-1$  additional vertices



$v_1, \dots, v_{t-1}$  such that, for  $i = 1, \dots, t-1$ , all four edges  $v_i u_i, v_i w_i, v_i x_{i+1}, v_i y_{i+1}$  are edges in  $G$ . This gives the square of a path on  $t(k+1) - 1$  vertices with edges from  $G \cup G(n, p)$  (c.f. Figure 3.1).

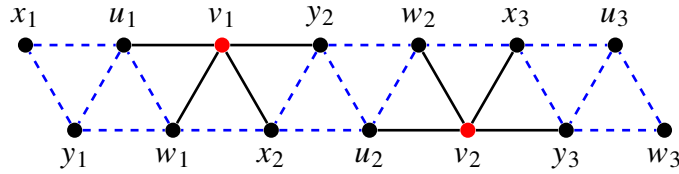


Figure 3.1: The square of a path with end-tuples  $(x_1, y_1)$  and  $(u_3, w_3)$  with our decomposition into random (dashed blue) and deterministic (black) edges for  $k = 4$  and  $t = 3$ .

Note that by requiring the edge  $w_t x_1$  from  $G(n, p)$  and adding another vertex  $v_t$  joined to  $u_t, w_t, x_1, y_1$  in  $G$ , we get the square of a cycle on  $t(k+1)$  vertices. However for parity reasons this may not cover all  $n$  vertices. Hence in order to find the square of a Hamilton cycle and for some additional technical reasons, our proof(s) will allow some of  $F_1, \dots, F_t$  to be the squares of paths of different lengths.

### 3.1.2 Extremal case

For the extremal case (Theorem 1.1.17), suppose that  $G$  is an  $n$ -vertex  $(\alpha, \beta)$ -stable graph with  $\alpha = \frac{1}{k+1}$ , and let  $p \geq C(\log n)^{1/(2k-3)} n^{-(k-1)/(2k-3)}$ . The definition of stability (Definition 1.1.10) gives a partition  $A \cup B$  of  $V(G)$  in which the size of  $B$  is roughly  $k$  times the size of  $A$ , the minimum degree of  $G[A, B]$  is at least  $\alpha n/4$ , and all but few vertices of  $A$  (resp.  $B$ ) are adjacent to all but few vertices of  $B$  (resp.  $A$ ). Our proof will follow three steps.

In the first step, we would like to embed copies  $F_i$  of  $P_k^2$  into  $B$  and vertices  $v_i$  into  $A$ , following the decomposition described above. However, this is only possible if  $|B| = k|A|$  and, therefore, we first embed squares of short paths of different lengths to ensure this parity condition holds in the remainder. We find a family  $\mathcal{F}_1$  of copies of squares of paths with end-tuples in  $B$ , such that after removing the vertices  $V_1 = \bigcup_{F \in \mathcal{F}_1} V(F)$ , we are left with two sets  $A_1 = A \setminus V_1$  and  $B_1 = B \setminus V_1$  with  $|B_1| = k(|A_1| - |\mathcal{F}_1|)$ . Note that we construct the family  $\mathcal{F}_1$  in such a way that we have  $|B_1| = k(|A_1| - |\mathcal{F}_1|)$  rather than  $|B_1| = k|A_1|$ , because each square path in  $\mathcal{F}_1$  still needs to be connected into the final square of a Hamilton cycle, and for each of these connections we shall use one vertex in  $A_1$ . The precise way we find  $\mathcal{F}_1$  depends on the sizes of  $A$  and  $B$ , but in all cases we will ensure that the vertices in the end-tuples of each  $F \in \mathcal{F}_1$  are neighbours of all but few vertices of  $A$ . When  $|B| > \frac{k}{k+1}n$ , the family  $\mathcal{F}_1$  consists of copies of  $P_{k+1}^2$  inside of  $B$ . Its existence is guaranteed by the following lemma using the minimum degree of  $G[B]$ .

**Lemma 3.1.1** (Embedding lemma in randomly perturbed graphs). *For all integers  $k \geq 2$  and  $t \geq 1$ , there exists  $C > 0$  such that the following holds for any  $0 \leq m \leq n/(32kt)$  and any  $n$ -vertex graph  $G$  of minimum degree  $\delta(G) \geq m$  and maximum degree  $\Delta(G) \leq n/(32kt)$ . For  $p \geq C(\log n)^{1/(2k-3)}n^{-(k-1)/(2k-3)}$ , a.a.s. the perturbed graph  $G \cup G(n, p)$  contains  $tm + t$  pairwise vertex-disjoint copies of the square of a path on  $k + 1$  vertices.*

When  $|B| \leq \frac{k}{k+1}n$ , the family  $\mathcal{F}_1$  consists of copies of  $P_{2k+3}^2$  with all vertices in  $B$ , except the  $(k + 1)$ -st and  $(k + 3)$ -rd, that belong to  $A$ .

Our second step is to cover the vertices in  $A_1$  and  $B_1$  that do not have a high degree to the other part. For this we will find another family  $\mathcal{F}_2$  of copies of squares of paths with end-tuples in  $B$ . For any vertex  $v$  in  $A_1$  with small degree into  $B_1$ , we find a copy of  $P_{2k+1}^2$  with  $v$  being the  $(k + 1)$ -st vertex and all the remaining vertices belonging to  $B_1$ . Similarly, for any vertex  $v$  in  $B_1$  with small degree into  $A_1$ , we find a copy of  $P_{3k+2}^2$  consisting of three copies of  $P_k^2$  in  $B$  connected by edges and two vertices from  $A_1$ , where  $v$  is in the middle copy of  $P_k^2$ . We need that  $v$  is in the middle copy, because then we can again ensure that the end-tuples of each  $F \in \mathcal{F}_2$  see all but few vertices of  $A$ . Moreover, with  $V_2 = \bigcup_{F \in \mathcal{F}_2} V(F)$  and  $A_2 = A_1 \setminus V_2$  and  $B_2 = B_1 \setminus V_2$ , we have  $|B_2| = k(|A_2| - |\mathcal{F}_1| - |\mathcal{F}_2|)$ .

At this point, each of the vertices in  $A_2$  (resp.  $B_2$ ) is adjacent to all but few vertices of  $B_2$  (resp.  $A_2$ ) and we kept the parity intact. In the third step, we let  $\mathcal{F}_3$  be pairwise disjoint random copies of  $P_k^2$  covering  $B_2$ , which is possible by Theorem 2.5.2 with our  $p$  and because  $|B_2|$  is divisible by  $k$ .

We let  $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3$  and, for each  $F \in \mathcal{F}$ , denote its end-tuples by  $(y_F, x_F)$  and  $(u_F, w_F)$ . We now reveal additional edges of  $G(n, p)$  and encode their presence in an auxiliary directed graph  $\mathcal{T}$  on vertex set  $\mathcal{F}$  as follows. There is a directed edge  $(F, F')$  if and only if the edge  $w_F x_{F'}$  appears in  $G(n, p)$ . It is easy to see that all directed edges in  $\mathcal{T}$  are revealed with probability  $p$  independently of all the others and, therefore, we can find a directed Hamilton cycle  $\vec{C}$  in  $\mathcal{T}$  with Theorem 2.5.4. We finally match to each edge  $(F, F')$  of  $\vec{C}$  a vertex  $v \in A_2$  such that  $u_F, w_F, x_{F'}, y_{F'}$  are all neighbours of  $v$  in the graph  $G$ . Owing to the high minimum degree conditions, that this is possible easily follows from Hall's matching theorem. Thus we get the square of a Hamilton cycle, as wanted.

We will prove Lemma 3.1.1 in Section 3.6.

### 3.1.3 Non-extremal case

For the non-extremal case (Theorem 1.1.16), assume that  $G$  is not  $(\frac{1}{k+1}, \beta)$ -stable and let  $p \geq Cn^{-(k-1)/(2k-3)}$ . Then we apply the regularity lemma to  $G$  and we obtain the reduced graph  $R$ . By adjusting an argument of Balogh, Mousset, and Skokan [9], we can prove the following stability result.

**Lemma 3.1.2.** *For any integer  $k \geq 2$ , and  $0 < \beta < 1/12$  there exists  $d > 0$  such that the following holds for any  $0 < \varepsilon < d/4$ ,  $4\beta \leq \alpha \leq 1/3$ , and  $t \geq 10/d$ . Let  $G$  be an  $n$  vertex*

graph with minimum degree  $\delta(G) \geq (\alpha - d/2)n$  that is not  $(\alpha, \beta)$ -stable and let  $R$  be the  $(\varepsilon, d)$ -reduced graph for some  $(\varepsilon, d)$ -regular partition  $V_0, \dots, V_t$  of  $G$ . Then  $R$  contains a matching  $M$  of size  $(\alpha + 2kd)t$ .

For completeness, we give the proof in a supplementary section of this chapter, Section 3.7. With Lemma 3.1.2, it is not hard to show that the reduced graph  $R$  can be vertex-partitioned into copies of stars  $K_{1,k}$  and matching edges  $K_{1,1}$ , such that there are not too many stars<sup>2</sup>. We would like to cover the clusters corresponding to each such star and matching edge with the square of a Hamilton path, and then connect these square paths to get the square of a Hamilton cycle. However, since we want to avoid the additional log-term in the probability we are working with, for this strategy to work in the randomly perturbed graph, we need that in each star the centre cluster is larger than the other clusters. Moreover, to ensure that we can connect the Hamilton paths, we need to setup some connections between the stars and matching edges.

Therefore, we first remove some vertices from the leaf clusters of each star to make it unbalanced and ensure that all pairs are super-regular. Then we label the stars and matching edges arbitrarily as  $Q_1, \dots, Q_s$ , and for  $i = 1, \dots, s$  find a copy  $F_i$  of  $P_6^2$  to connect  $Q_i$  and  $Q_{i+1}$  (where indices are modulo  $s$ ). More precisely, for each star  $Q_i$ , one end-tuple of  $F_{i-1}$  and one of  $F_i$  belong to the centre cluster of  $Q_i$ , while the other two end-tuples belong to the same leaf cluster of  $Q_i$ . Moreover, for each matching edge  $Q_i$ , each of its two clusters contains exactly one end-tuple, one from  $F_{i-1}$  and the other from  $F_i$ . We will refer to these squares of paths as the connecting (squares of) paths. Let  $V_0$  be the set of vertices no longer contained in any of the stars or matching edges. We cover each  $v \in V_0$  by appending  $v$  to one of the connecting paths. Here we use that any vertex  $v \in V_0$  has degree at least  $(\frac{1}{k+1} - \alpha)n$  and, as we do not have too many stars, the vertex  $v$  has also many neighbours in clusters which are not centres of stars. This is crucial, because it allows us to ensure that the relations between the sizes of the sets in any star are suitable for an application of the following lemma, which is the main technical ingredient in the proof.

**Lemma 3.1.3** (Embedding lemma in randomly perturbed graphs). *For any  $k \geq 2$  and any  $0 < \delta' \leq d < 1$  there exist  $\delta_0, \delta_1, \varepsilon > 0$  with  $\delta' \geq \delta_0 > 2\delta_1 > \varepsilon$  and  $C > 0$  such that the following holds. Let  $V, U_1, \dots, U_k$  be pairwise disjoint sets such that  $|V| = n + 4$ ,  $(1 - \delta_0)n \leq |U_1| = \dots = |U_k| \leq (1 - \delta_1)n$  and  $n - |U_1| \equiv -1 \pmod{3k - 1}$ . Suppose that  $(V, U_i)$  is an  $(\varepsilon, d)$ -super-regular pair with respect to a graph  $G$  for each  $i = 1, \dots, k$ , and  $(x, x')$  and  $(y, y')$  are two tuples from  $V$  such that both tuples have  $d^2n/2$  common neighbours in  $U_i$  for each  $i = 1, \dots, k$ . Furthermore, let  $G(V, p)$  and  $G(U_1, \dots, U_k, p)$  be random graphs with  $p \geq Cn^{-(k-1)/(2k-3)}$ .*

*Then a.a.s. there exists the square of a Hamilton path in  $G[V, U_1, \dots, U_k] \cup G(V, p) \cup$*

<sup>2</sup>We remark that covers of the reduced graph by stars were used in [10, 79]. However, in contrast to [10], for our purposes it is necessary that we cover the reduced graph by matching edges as well.

$G(U_1, \dots, U_k, p) \cup \{xx', yy'\}$  covering  $V, U_1, \dots, U_k$ , and with end-tuples  $(x, x')$  and  $(y, y')$ .

It is worth to explain why the size of the leaf clusters needs to be slightly smaller than the size of the centre cluster. Proposition 1.1.15 shows that if  $G$  is the complete bipartite graph with parts of size  $\frac{1}{k+1}n$  and  $\frac{k}{k+1}n$ , then, in order for  $G \cup G(n, p)$  to a.a.s. contain the square of a Hamilton cycle, we need  $p \geq Cn^{-(k-1)/(2k-3)} \log^{1/(2k-3)}$  for some constant  $C > 0$ . Suppose now that  $V, U_1, \dots, U_k$  are pairwise disjoint sets all with the same size  $n$  and that  $(V, U_i)$  is an  $(\varepsilon, d)$ -super-regular pair with respect to a graph  $G$  for each  $i = 1, \dots, k$ . If we want  $G[V, U_1, \dots, U_k] \cup G((k+1)n, p)$  to a.a.s. contain the square of a Hamilton cycle, we need again  $p \geq C'n^{-(k-1)/(2k-3)} \log^{1/(2k-3)} n$  for some constant  $C' > 0$ . In fact, while we are only assuming that the pairs  $(V, U_i)$  are super-regular, the same bound on  $p$  is needed even when  $G[V, U_i]$  is complete bipartite for each  $i \in [k]$  since, in that case,  $G$  is the complete bipartite graph with parts of size  $n$  and  $kn$  and the lower bound on  $p$  follows from Proposition 1.1.15. The unbalanced setting is therefore crucial to avoid the log-term in  $p$  and allows us to prove the stability theorem with a lower edge-density.

Lemma 3.1.3 implies that for any star  $Q_i$  we can connect the end-tuples of  $F_{i-1}$  and  $F_i$  which belong to the centre cluster of  $Q_i$ , while covering all vertices in the clusters of  $Q_i$ . Similarly for the matching edges we use the following lemma.

**Lemma 3.1.4** (Embedding lemma in randomly perturbed graphs). *For any  $0 < d < 1$  there exist  $\varepsilon > 0$  and  $C > 0$  such that the following holds for sets  $U, V$  with  $|V| = n$  and  $\frac{3}{4}n \leq |U| \leq n$ . Let  $(U, V)$  be an  $(\varepsilon, d)$ -super-regular pair with respect to a graph  $G$ , and  $(x, x')$  and  $(y, y')$  be tuples from  $V$  and  $U$ , respectively, such that the vertices from the tuples have  $\frac{1}{2}d^2n$  common neighbours in  $U$  and  $V$ , respectively. Furthermore let  $G(U, p)$  and  $G(V, p)$  be random graphs with  $p \geq Cn^{-1}$ .*

*Then a.a.s. there exists the square of a Hamilton path in  $G[U, V] \cup G(U, p) \cup G(V, p) \cup \{xx', yy'\}$  covering  $U, V$  with end-tuples  $(x, x')$  and  $(y, y')$ .*

Together this gives the square of a Hamilton cycle in  $G \cup G(n, p)$ . We will give the proofs of Lemmas 3.1.3 and 3.1.4 in Section 3.5.

### 3.2 Proof of the extremal Theorem 1.1.17

*Proof of Theorem 1.1.17.* Given an integer  $k \geq 2$ , define  $\gamma = \frac{1}{32k(k+1)}$ . Then let  $C_2$  be given by Lemma 3.1.1 on input  $k$  and  $t = k + 1$ , and let  $C_4$  be given by Theorem 2.5.4. Finally set  $C = 4C_2 + 8kC_4 + 8$ . Next, we let  $0 < \beta \leq \frac{1}{100k^3}\gamma$ . Given  $n$ , let  $0 \leq a \leq k$  be such that  $n = (k+1)\lfloor \frac{1}{k+1}n \rfloor + a$  and  $p \geq C(\log n)^{1/(2k-3)}n^{-(k-1)(2k-3)}$ . We reveal a subgraph of  $G(n, p)$  in four rounds  $G_i \sim G(n, p/4)$  for  $i = 1, 2, 3, 4$ . By Lemma 2.5.1 and

a union bound over all graphs on at most  $4k$  vertices, we can a.a.s. assume that

$$G_1 \text{ contains a copy of the graph } F \text{ in any vertex-set of size at least } \beta n, \quad (3.2.1)$$

where  $F$  is any graph on at most  $4k$  vertices with  $m_1(F) \leq m_1(P_k^2) = \frac{2k-3}{k-1}$ .

Let  $G$  be an  $n$ -vertex graph with minimum degree at least  $\frac{1}{k+1}n$  that is  $(\frac{1}{k+1}, \beta)$ -stable. Then there exists a partition of  $V(G)$  into  $A$  and  $B$  that satisfies Definition 1.1.10. As outlined in Section 3.1.2 our proof will consist of three steps. We will successively build parts of the square of a Hamilton cycle, first covering some vertices to balance the partition, secondly covering vertices of low degree to the other side, and then covering the remaining vertices. Finally we will connect these parts into the square of a Hamilton cycle.

**Balancing the partition.** Our goal is to find a family  $\mathcal{F}_1$  of pairwise disjoint copies of squares of paths with end-tuples in  $B$ , such that the size of the set  $V_1 = \bigcup_{F \in \mathcal{F}_1} V(F)$  is smaller than  $3k^2\beta n$ , and after removing the vertices of  $V_1$  we are left with two sets

$$A_1 = A \setminus V_1 \quad \text{and} \quad B_1 = B \setminus V_1 \quad \text{such that} \quad |B_1| = k(|A_1| - |\mathcal{F}_1|). \quad (3.2.2)$$

We distinguish between the cases  $|A| = \lfloor \frac{n}{k+1} \rfloor + m$  with  $1 \leq m \leq \beta n$  and  $|A| = \lfloor \frac{n}{k+1} \rfloor - m$  with  $0 \leq m \leq \beta n$ . Suppose first that  $|A| = \lfloor \frac{n}{k+1} \rfloor + m$  for some  $1 \leq m \leq \beta n$ . In this case we want  $|\mathcal{F}_1| = m$  and the family  $\mathcal{F}_1$  will consist of  $m - 1$  copies of  $P_{3k+2}^2$  and one copy of  $P_{3k+2+a}^2$ , such that for each of these  $m$  copies, exactly three vertices are in  $A$ , both end-tuples are in  $B$  and each end-tuple has at least  $|A| - 2\beta n$  common neighbours in  $A$ .

We can do this greedily in  $G \cup G_1$ . Assume that during this process we have to find a copy of  $P_{3k+2}^2$  or  $P_{3k+2+a}^2$ , i.e. a copy of  $P_{3k+b}^2$  for some  $2 \leq b \leq k + 2$ , such that the above conditions are satisfied. There are three vertices  $v_1, v_2$  and  $v_3$  in  $P_{3k+b}^2$ , such that none of them is in an end-tuple of  $P_{3k+b}^2$ , they do not induce a triangle in  $P_{3k+b}^2$ , and the subgraph  $H = P_{3k+b}^2 \setminus \{v_1, v_2, v_3\}$  satisfies  $m_1(H) \leq m_1(P_k^2)$  (see Figure 3.2). We can avoid an induced triangle because there are at least  $3k + b - 4 \geq 4$  vertices to choose from that are not in end-tuples. Moreover we can guarantee the bound on the density because, when distributing the three vertices evenly, the longest square of a path in  $H$  has at most  $\lceil (3k + b - 3)/4 \rceil \leq k$  vertices. We remark that we can always ask  $\{v_1, v_2, v_3\}$  to be an independent set in  $P_{3k+b}^2$  when  $k > 2$ .

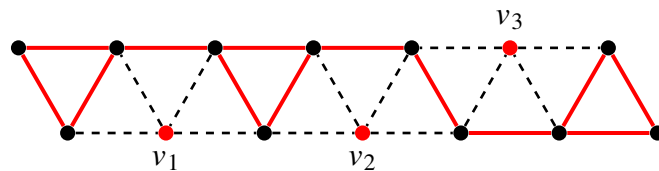


Figure 3.2: Subgraph  $H$  (solid red) obtained from  $P_{3k+b}^2$  after removing three vertices  $v_1, v_2, v_3$  (red) for  $b = k + 2$  and  $k = 3$ . The 1-density of  $H$  is the same as  $P_k^2$ .

Then we find a copy of  $P_{3k+b}^2$  by embedding the vertices  $v_1, v_2, v_3$  in  $A$  and the other vertices in  $B$  in the following way. Let  $A' \subseteq A$  be the set of vertices of  $A$  that have degree at least  $|B| - \beta n$  into  $B$  and have not yet been covered, and observe that  $|A'| \geq |A| - \beta n - 3|\mathcal{F}_1| \geq \beta n$ . Then, since  $m_1(P_3) = 1 \leq m_1(P_k^2)$  and given (3.2.1), the random graph  $G_1[A']$  contains a path on three vertices  $u_1, u_2, u_3$ . Next, let  $B' \subseteq B$  be the set of vertices of  $B$  that have degree at least  $|A| - \beta n$  into  $A$ , are common neighbours of  $u_1, u_2, u_3$  and have not yet been covered, and observe that  $|B'| \geq |B| - \beta n - 3\beta n - (3k + k + 2 - 3)|\mathcal{F}_1| \geq \beta n$ . Since  $m_1(H) = m_1(P_k^2)$ , using again (3.2.1), the random graph  $G_1[B']$  contains a copy of  $H$ , that together with the vertices  $u_1, u_2, u_3$  and some edges from  $G$  gives a copy of  $P_{3k+b}^2$ . In particular, since each vertex in the end-tuples of  $P_{3k+b}^2$  is embedded in  $B'$ , then, by definition of  $B'$ , each end-tuple has at least  $|A| - 2\beta n$  common neighbours in  $A$ . In view of (3.2.2), we get that  $|A_1| = \lfloor \frac{n}{k+1} \rfloor - 2m$ ,

$$|B_1| = n - |A| - (3k - 1)m - a = k \left\lfloor \frac{n}{k+1} \right\rfloor - 3km = k(|A_1| - |\mathcal{F}_1|),$$

and  $|V_1| = (m - 1)(3k + 2) + 3k + 2 + a \leq (4k + 3)m \leq 3k^2\beta n$ .

Now suppose that  $|A| = \lfloor \frac{n}{k+1} \rfloor - m$  for some  $0 \leq m \leq \beta n$ . In this case, the family  $\mathcal{F}_1$  will consist of some copies of  $P_{k+1}^2$  and some copies of  $P_{2k+1}^2$ , such that, for each copy, all vertices are in  $B$  and each end-tuple has at least  $|A| - 8k^2\beta n$  common neighbours in  $A$ ; in particular, we do not touch the set  $A$ . We start from  $\mathcal{F}_1 = \emptyset$  and let  $B^\star = \{v \in B : \deg(v, B) \geq 4k^2\beta n\}$  be the set of vertices in  $B$  with high degree to  $B$  in  $G$ . Since by Definition 1.1.10 we have  $e(G[B]) \leq \beta n^2$ , then  $|B^\star| \leq \frac{n}{2k^2}$ . Moreover if  $v \in B \setminus B^\star$ , then  $\deg(v, A) \geq \delta(G) - \deg(v, B) \geq \frac{n}{k+1} - 4k^2\beta n \geq |A| - 4k^2\beta n$ . With  $m_0 = \max\{m - |B^\star|, 0\}$ , we have  $\delta(G[B \setminus B^\star]) \geq m_0$ , as

$$\delta(G[B \setminus B^\star]) \geq \frac{n}{k+1} - |A| - |B^\star| = \frac{n}{k+1} - \left\lfloor \frac{n}{k+1} \right\rfloor + m - |B^\star| \geq m - |B^\star|.$$

Moreover, by definition of  $B^\star$ , we have  $\Delta(G[B \setminus B^\star]) \leq 4k^2\beta n \leq \gamma|B \setminus B^\star|$ , where the last inequality follows from the choice of  $\beta$  and  $|B \setminus B^\star| \geq (\frac{k}{k+1} - \frac{1}{2k^2})n$ . Therefore we can use Lemma 3.1.1 with parameters  $k$  and  $t = k + 1$  and we a.a.s. find  $(k + 1)m_0 + a$  pairwise disjoint copies of  $P_{k+1}^2$  in  $(G \cup G_2)[B \setminus B^\star]$ , which we add to  $\mathcal{F}_1$ . All the vertices of such copies belong to  $B \setminus B^\star$  and thus in particular each end-tuple has at least  $|A| - 8k^2\beta n$  common neighbours in  $A$ .

Next we want to find  $m - m_0 \leq \min\{|B^\star|, m\}$  copies of  $P_{2k+1}^2$  in  $B$  disjoint from any graph already in  $\mathcal{F}_1$ . First observe that the graph  $H$  obtained by taking the disjoint union of two copies of  $P_k^2$  with the addition of the edge between the last vertex of the first copy and the first vertex of the second copy satisfies  $m_1(H) \leq m_1(P_k^2)$ . The graph  $H$  will be embedded in  $G_1$  and we will turn that into an embedding of  $P_{2k+1}^2$ , by adding a vertex and four edges of  $G$ . We can again do this greedily in  $G \cup G_1$ . First we remove vertices from

$B^*$  such that  $|B^*| = m - m_0 \leq m \leq \beta n$ . Then we let  $B'$  be the set of vertices in  $B$  with less than  $|A| - \beta n$  neighbours in  $A$ , and note that  $|B'| \leq \beta n$ . We then pick a vertex  $w$  from  $B^*$ , not yet covered, and denote by  $N_w$  the set of neighbours of  $w$  in  $B \setminus (B^* \cup B')$  in the graph  $G$ , that have not yet been covered. Then

$$|N_w| \geq 4k^2\beta n - |B'| - |B^*| - ((k+1)m_0 + a)(k+1) - (m - m_0)(2k+1) \geq \beta n.$$

Therefore, since  $m_1(H) \leq m_1(P_k^2)$  and using (3.2.1), the random graph  $G_1[N_w]$  contains a copy of this  $H$ , that together with  $w$  and four edges from  $G$ , gives a copy of  $P_{2k+1}^2$  as desired. The end-tuples of this copy belong to  $B \setminus B'$  and thus each of them has at least  $|A| - 2\beta n$  common neighbours in  $A$ . Once this is done, in view of (3.2.2), we indeed get  $|A_1| = |A| = \lfloor \frac{1}{k+1}n \rfloor - m$ ,

$$\begin{aligned} |B_1| &= n - |A| - (k+1)((k+1)m_0 + a) - (2k+1)(m - m_0) \\ &= k \left( \left\lfloor \frac{n}{k+1} \right\rfloor - a - 2m - km_0 \right) = k(|A_1| - |\mathcal{F}_1|), \end{aligned}$$

and  $|V_1| = ((k+1)m_0 + a)(k+1) + (m - m_0)(2k+1) \leq (k+1)^2m + k(k+1) \leq 3k^2\beta n$ . This finishes the first step of our proof.

**Covering low degree vertices.** In this step the goal is to find a family  $\mathcal{F}_2$  of pairwise disjoint copies of  $P_{2k+1}^2$  and  $P_{3k+2}^2$  that cover all vertices from  $A$  (respectively  $B$ ) that do not have high degree to  $B$  (respectively  $A$ ) and such that, for each copy, each end-tuple is in  $B$  and has at least  $|A| - 16k^2\beta n$  common neighbours in  $A$ . We will do this such that after removing  $V_2 = \bigcup_{F \in \mathcal{F}_2} V(F)$  we are left with two sets  $A_2 = A_1 \setminus V_2$  and  $B_2 = B_1 \setminus V_2$  such that

$$|B_2| = k(|A_2| - |\mathcal{F}_1| - |\mathcal{F}_2|). \quad (3.2.3)$$

We let  $A' = \{v \in A_1 : \deg(v, B) \leq |B| - \beta n\}$  and  $B' = \{v \in B_1 : \deg(v, A) \leq |A| - 8k^2\beta n\}$ , and note that  $|A'| \leq \beta n$  and  $|B'| \leq \beta n$ . Let  $\mathcal{F}_2 = \emptyset$ . We start by taking care of the vertices in  $A'$  and we cover each of them with a copy of  $P_{2k+1}^2$  with all other vertices in  $B_1 \setminus B'$ . For  $u \in A'$ , let  $N_u$  be the set of neighbours of  $u$  in  $B_1 \setminus B'$  in  $G$  that are not yet covered by any graph in  $\mathcal{F}_2$  and observe that

$$|N_u| \geq \deg(u, B) - |B \setminus B_1| - |B'| - 2k|\mathcal{F}_2| \geq \frac{n}{4(k+1)} - 3k^2\beta n - \beta n - 2k\beta n \geq \beta n,$$

where we used that  $\deg(u, B) \geq \frac{n}{4(k+1)}$  as  $G$  is  $(\frac{1}{k+1}, \beta)$ -stable (see Definition 1.1.10). Note that the graph  $H$  obtained by taking the union of two copies of  $P_k^2$  with the addition of an edge between two end-vertices satisfies  $m_1(H) \leq m_1(P_k^2)$ . Therefore, using (3.2.1), the random graph  $G_1[N_u]$  contains a copy of  $H$  and together with  $u$  and four edges of  $G$ , this gives the desired copy of  $P_{2k+1}^2$ . Both end-tuples of this copy of  $P_{2k+1}^2$  belong to  $B_1 \setminus B'$

and, thus, have at least  $|A| - 16k^2\beta n$  common neighbours in  $A$ . We add this copy of  $P_{2k+1}^2$  to  $\mathcal{F}_2$  and we continue until we cover all vertices of  $A'$ .

Now we cover each vertex from  $B'$  with a copy of  $P_{3k+2}^2$ , where each copy uses one vertex from  $B'$ , two vertices from  $A_1 \setminus A'$  and the other  $3k - 1$  vertices from  $B_1 \setminus B'$ . Let  $w \in B'$  and  $u_1, u_2 \in A_1 \setminus A'$  be vertices not yet covered. We denote by  $N_w$  the subset of  $B_1 \setminus B'$  which contains the common neighbours of  $w, u_1, u_2$  in  $G$  that are not yet covered. Observe that the definitions of  $A'$  and  $B'$  give

$$\begin{aligned} |N_w| &\geq (\delta(G) - \deg(w, A)) - (|B| - \deg(u_1, B)) - (|B| - \deg(u_2, B)) + \\ &\quad - |B \setminus B_1| - 3k|B'| - 2k|A'| \\ &\geq \frac{n}{k+1} - (|A| - 8k^2\beta n) - \beta n - \beta n - 3k^2\beta n - 3k\beta n - 2k\beta n \geq \beta n. \end{aligned}$$

Similarly as above, using (3.2.1), the random graph  $G_1[N_w]$  contains a copy of a graph  $H$  on  $3k - 1$  vertices, that together with  $w, u_1, u_2$  and some edges from  $G$ , gives a copy of  $P_{3k+2}^2$ . The end-tuples of the copy of  $P_{3k+2}^2$  belong to  $B_1 \setminus B'$  and, thus, have at least  $|A| - 16k^2\beta n$  common neighbours in  $A$ . We add the copy of  $P_{3k+2}^2$  to  $\mathcal{F}_2$  and repeat until all of  $B'$  is covered. We then get (3.2.3), because of (3.2.2) and since for each graph added to  $\mathcal{F}_2$  the ratio of vertices removed from  $A_1$  and  $B_1$  is one to  $2k$  or two to  $3k$ . Moreover, we have  $\deg(v, B_2) \geq |B_2| - \beta n$  for  $v \in A_2$ ,  $\deg(v, A_2) \geq |A_2| - 8k^2\beta n$  for  $v \in B_2$ , and  $|V_2| \leq 6k\beta n$ , which implies  $|A_2| \geq |A| - |V_1| - |V_2| \geq \frac{n}{2(k+1)}$ .

**Covering everything and connecting.** In this step we first cover  $B_2$  with copies of  $P_k^2$ . Then, using the uncovered vertices in  $A_2$ , we connect all the copies of squares of paths found so far, to get the square of a Hamilton cycle. Observe that after the cleaning steps,  $k$  divides  $|B_2|$  by (3.2.3) and thus Theorem 2.5.2 implies that a.a.s the random graph  $G_3[B_2]$  has a  $P_k^2$ -factor. We denote the family of such copies of  $P_k^2$  by  $\mathcal{F}_3$ , and observe that (3.2.3) implies  $|\mathcal{F}_3| = |A_2| - |\mathcal{F}_1| - |\mathcal{F}_2|$ .

We let  $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3$  be the family of all the squares of paths that we have constructed and, for each  $F \in \mathcal{F}$ , denote the end-tuples of  $F$  by  $(y_F, x_F)$  and  $(u_F, w_F)$ . Note that by construction, each pair  $x_F, y_F$  and  $u_F, w_F$  has at least  $|A_2| - 16k^2\beta n$  common neighbours in  $A_2$ . We now reveal the edges of  $G_4$  and construct an auxiliary directed graph  $\mathcal{T}$  on vertex set  $\mathcal{F}$  as follows. Given any two  $F, F' \in \mathcal{F}$ , there is a directed edge  $(F, F')$  if and only if the edge  $w_F x_{F'}$  appears in  $G_4$ . Since all directed edges in  $\mathcal{T}$  are revealed with probability  $\frac{1}{4}p$  independently of all the others,  $\mathcal{T}$  is distributed as  $\vec{G}(|\mathcal{F}|, \frac{1}{4}p)$ . Then, as  $|\mathcal{F}| \geq \frac{1}{2k}n$  and  $\frac{1}{4}p \geq C_4 \frac{\log |\mathcal{F}|}{|\mathcal{F}|}$ , there a.a.s. is a directed Hamilton cycle  $\vec{C}$  in  $\mathcal{T}$  by Theorem 2.5.4.

In order to get the desired square of a Hamilton path, it remains to match the edges  $(F, F')$  of  $\vec{C}$  to the vertices  $v \in A_2$  such that  $u_F, w_F, x_{F'}, y_{F'}$  are all neighbours of  $v$  in the graph  $G$ . Observe indeed that  $|E(\vec{C})| = |\mathcal{F}_1| + |\mathcal{F}_2| + |\mathcal{F}_3| = |A_2|$ , the cycle  $\vec{C}$  gives a cyclic ordering of the squares of paths in  $\mathcal{F}$ , and between consecutive copies  $(F, F')$  we have the edge  $w_F x_{F'}$  (by definition of the graph  $\mathcal{T}$ ) and the matching will give a unique vertex  $v$



incident to  $u_F, w_F, x_{F'}$  and  $y_{F'}$ . For that we define the following auxiliary bipartite graph  $\mathcal{B}$  with classes  $E(\vec{C})$  and  $A_2$ . There is an edge between  $(F, F') \in E(\vec{C})$  and  $v \in A_2$  if and only if  $u_F, w_F, x_{F'}, y_{F'}$  are all neighbours of  $v$  in the graph  $G$ . The existence of a perfect matching in  $\mathcal{B}$  easily follows from Hall's condition, as the minimum degree of  $\mathcal{B}$  is large: by the choice of  $\beta$  and the lower bound on  $|A_2|$ , the degree of  $(F, F') \in E(\vec{C})$  in  $\mathcal{B}$  is at least  $|A_2| - 32k^2\beta n \geq \frac{1}{2}|A_2|$  and the degree of  $v \in A_2$  in  $\mathcal{B}$  is at least  $|A_2| - 4\beta n \geq \frac{1}{2}|A_2|$ . This finishes the proof.  $\square$

### 3.3 Proof of the stability Theorem 1.1.16

*Proof of Theorem 1.1.16.* Let  $k \geq 2$  and  $0 < \beta < 1/6k$ . We obtain  $d > 0$  from Lemma 3.1.2 with input  $k$  and  $\beta$ , and let  $0 < \delta' < 2^{-12}(k-1)^{-2}d^{2k^2}$  and  $\gamma = d/(k-1)$ . Next we obtain  $\delta_0, \delta, \varepsilon' > 0$  and  $C_1$  from Lemma 3.1.3 with input  $k, \delta'$ , and  $d/2$  such that  $\delta' \geq \delta_0 > 2\delta$  ( $\delta$  plays the role of  $\delta_1$  in the lemma). We additionally assume that  $\varepsilon'$  is small enough for Lemma 3.1.4 with input  $d/2$  and also obtain  $C_2$  from this. Then we let  $0 < \varepsilon < \varepsilon'/8$ . The constant dependencies can be summarised as follows:

$$\varepsilon \ll \varepsilon' \ll \delta \ll \delta_0 \ll \delta' \ll d \ll \beta < 1/6k.$$

We now apply Lemma 2.1.4 with input  $\varepsilon$  and  $t_0 = d/10$ , and get  $T$ . Further, let the following parameters be given by Lemma 2.5.3:  $C_3$  for input 3 (in place of  $s$ ), 2 (in place of  $k$ ), and  $\delta(kT)^{-6}$  (in place of  $\eta$ );  $C_4$  for input  $2k, k$ , and  $\delta(kT)^{-2k^2}$ ;  $C_5$  for input 4, 2, and  $\delta(kT)^{-8}$ ; and  $C_6$  for input 1, 2, and  $\delta(kT)^{-2}$ . Then let  $C$  be large enough such that with  $p \geq Cn^{-(k-1)/(2k-3)}$  the random graph  $G(n, p)$  contains the union  $\bigcup_{i=1}^6 G_i$ , where  $G_1 \sim G(n, C_1 \cdot 2(k-1)Tn^{-(k-1)/(2k-3)})$ ,  $G_2 \sim G(n, C_2 \cdot 2(k-1)Tn^{-1})$ ,  $G_3 \sim G(n, C_3n^{-1})$ ,  $G_4 \sim G(n, C_4n^{-(k-1)/(2k-3)})$ ,  $G_5 \sim G(n, C_5n^{-1})$ , and  $G_6 \sim G(n, C_6n^{-1})$ .

Let  $G$  be an  $n$ -vertex graph with vertex set  $V$  and minimum degree  $\delta(G) \geq (\frac{1}{k+1} - \gamma)n$  that is not  $(\frac{1}{k+1}, \beta)$ -stable. We apply the regularity lemma (Lemma 2.1.4) to  $G$  and get a subgraph  $G'$  of  $G$ , a constant  $t$  with  $3 < t+1 \leq T$ , and an  $(\varepsilon, d)$ -regular partition  $V'_0, \dots, V'_t$  of  $V$ , satisfying (P1) – (P4). Consider the  $(\varepsilon, d)$ -reduced graph  $R$  for  $G$  and observe  $\delta(R) \geq (\frac{1}{k+1} - 2d)t$ , as, otherwise, in  $G'$  there would be a vertex of degree at most  $(\frac{1}{k+1} - 2d)t \frac{n}{t} + \varepsilon n < (\frac{1}{k+1} - \gamma)n - (d + \varepsilon)n$  in contradiction to (P2). As outlined in Section 3.1.3, the proof will consist of four steps. We will cover the reduced graph with copies of stars isomorphic to  $K_{1,k}$  and  $K_{1,1}$ , connect those stars with the squares of short paths, cover the exceptional vertices, and finally cover the whole graph with the square of a Hamilton path.

**Covering  $R$  with stars.** We start by covering the vertices of  $R$  with vertex-disjoint stars, each with at most  $k$  leaves; then we will turn this into a cover with copies of stars isomorphic to  $K_{1,k}$  and  $K_{1,1}$ . We let  $M_1$  be a largest matching in  $R$  and, with Lemma 3.1.2,

we get that  $|M_1| \geq \left(\frac{1}{k+1} + 2kd\right)t$  since  $G$  is not  $(\frac{1}{k+1}, \beta)$ -stable. By the maximality of  $M_1$ , the remaining vertices  $X_1 = V(R) \setminus V(M_1)$  form an independent set in  $R$ ; moreover only one endpoint of each edge in  $M_1$  can be adjacent to more than one vertex from  $X_1$  and the endpoints of each edge in  $M_1$  cannot have different neighbours in  $X_1$ . For each  $i = 2, \dots, k$ , we greedily pick a maximal matching  $M_i$  between  $X_{i-1}$  and  $V(M_1)$  and we set  $X_i = X_{i-1} \setminus V(M_i)$ . Observe that, using the properties coming from the maximality of  $M_1$  outlined above, the matching  $M_i$  covers at least  $\min\{|X_{i-1}|, \delta(R)\}$  vertices of  $X_{i-1}$ . Since  $2|M_1| + (k-1)\delta(R) \geq t$ , we have  $X_k = \emptyset$  and the union  $\bigcup_{i=1}^k M_i$  covers all vertices of  $R$  with stars, each isomorphic to one of  $K_{1,1}, \dots, K_{1,k}$ . Moreover note that the number of  $K_{1,k}$  is at most  $|M_k| \leq t - 2|M_1| - (k-2)\delta(R) \leq \left(\frac{1}{k+1} - 2kd\right)t$ .

For simplicity we only want to work with stars isomorphic to  $K_{1,k}$  and  $K_{1,1}$ . To obtain this, we split each cluster  $V$  arbitrarily into  $k-1$  parts  $V^1, \dots, V^{k-1}$  of the same size, where we move at most  $t(k-1)$  vertices to  $V'_0$  for divisibility reasons. Note that from any  $(\varepsilon, d)$ -regular pair we get  $(k-1)^2$  pairs that are  $(k\varepsilon, d-\varepsilon)$ -regular. We denote this new partition by  $V_0 = V'_0, V_1, \dots, V_{t'}$  with  $t' = t(k-1)$  and denote the reduced graph for this partition by  $R'$ . We now show that we can cover  $R'$  with copies of stars isomorphic to  $K_{1,k}$  and  $K_{1,1}$ . Any copy of  $K_{1,1}$  or  $K_{1,k}$  in  $R$  immediately gives  $k-1$  copies of  $K_{1,1}$  or  $K_{1,k}$  in  $R'$ , respectively. Moreover given any copy  $V, U_1, \dots, U_i$  of  $K_{1,i}$  in  $R$ , with  $V$  being the centre cluster and  $2 \leq i \leq k-1$ , we find  $i-1$  copies of  $K_{1,k}$  and  $k-i$  copies of  $K_{1,1}$  in  $R'$  in the following way. For each  $j = 1, \dots, i-1$ , the clusters  $V^j, U_j^1, \dots, U_j^{k-1}, U_i^j$  give a copy of  $K_{1,k}$  in  $R'$ , and for each  $j = i, \dots, k-1$  the clusters  $V^j, U_i^j$  give a copy of  $K_{1,1}$  in  $R'$ . Therefore, we have covered the vertices of  $R'$  with a collection  $\mathcal{K}$  of copies of  $K_{1,k}$  and  $K_{1,1}$ . We remark that we can still upper bound the number of copies of  $K_{1,k}$  in  $\mathcal{K}$  as follows. Since each copy of  $K_{1,i}$  in the original cover gives  $i-1$  copies of  $K_{1,k}$  in  $\mathcal{K}$ , we get the largest number of copies of  $K_{1,k}$  in  $\mathcal{K}$  when the total number of stars in the original cover is minimal. The original cover of  $R$  had at most  $(\frac{1}{k+1} - 2kd)t$  copies of  $K_{1,k}$ , and the remaining  $t - (k+1)(\frac{1}{k+1} - 2kd)t = 2k(k+1)dt$  vertices can give at most  $\frac{2k(k+1)dt}{k} = 2(k+1)dt$  copies of  $K_{1,k-1}$ . Therefore the collection  $\mathcal{K}$  has at most  $(k-1)(\frac{1}{k+1} - 2kd)t + (k-2)2(k+1)dt \leq (\frac{1}{k+1} - \frac{4d}{k-1})t'$  copies of  $K_{1,k}$ . We let  $n_0 = |V_1| = \lfloor |V'_1|/(k-1) \rfloor$  be the size of the clusters in  $R'$  and observe that  $(1-2\varepsilon)n/t' \leq n_0 \leq n/t'$ .

For convenience we relabel the clusters as follows. We let  $\mathcal{I} \subseteq [t']$  be the set of indices of those clusters of  $R'$  that are the centre cluster in a copy of  $K_{1,k}$  in  $\mathcal{K}$ , and for  $i \in \mathcal{I}$  we denote by  $U_{i,1}, \dots, U_{i,k}$  the clusters of  $R'$  that, together with  $V_i$ , create a copy of  $K_{1,k}$  in  $\mathcal{K}$ . Then we let  $\mathcal{J} \subseteq [t']$  be any set of indices with the following property: each index in  $\mathcal{J}$  corresponds to a cluster of a copy of  $K_{1,1}$  in  $\mathcal{K}$ , and for each copy of  $K_{1,1}$  in  $\mathcal{K}$ , exactly one of its clusters has its index in  $\mathcal{J}$ . Moreover, for each  $i \in \mathcal{J}$ , we let  $U_{i,1}$  be the cluster of  $R'$  that, together with  $V_i$ , creates a copy of  $K_{1,1}$  in  $\mathcal{K}$ .

We would like to apply Lemma 3.1.3 and Lemma 3.1.4 to the clusters corresponding to

the copies of  $K_{1,k}$  and  $K_{1,1}$ , respectively. However we first need to make each regular pair super-regular and unbalance some of the clusters to allow an application of Lemma 3.1.3. For that we arbitrarily move  $\delta n_0 + 4(1 - \delta)$  vertices from each cluster  $U_{i,j}$  with  $i \in \mathcal{I}$  and  $j = 1, \dots, k$  to  $V_0$ . Observe this ensures that  $|U_{i,k}| \leq (1 - \delta)(|V_i| - 4)$ . Next we repeatedly use Lemma 2.1.2 and move at most  $k^2 \varepsilon n_0$  vertices from each cluster to  $V_0$ , to ensure that all edges of  $R'$  within a copy of  $K_{1,k}$  or  $K_{1,1}$  from  $\mathcal{K}$  are  $(2k\varepsilon, d - 4k\varepsilon)$ -super-regular. When doing this we can ensure that for each  $i \in \mathcal{I}$  all clusters  $V_i$  have the same size, and that for  $j = 1, \dots, k$  all clusters  $U_{i,j}$  have the same size, except the cluster  $U_{i,1}$  that has two more vertices than the other  $U_{i,j}$ 's. Additionally, we can ensure that for each  $i \in \mathcal{J}$ , the clusters  $V_i$  and  $U_{i,1}$  have the same size. Moreover, by moving only at most  $3k$  additional vertices from each  $U_{i,j}$  to  $V_0$  (which do not harm the bounds above), we can ensure that for  $i \in \mathcal{I}$  we have  $|V_i| - 4 - |U_{i,k}| \equiv -1 \pmod{3k - 1}$ , again in view of a later application of Lemma 3.1.3. Note that at this point we have  $|V_0| \leq \varepsilon n + t(k - 1) + t'\delta n_0 + t'k^2 \varepsilon n_0 + t'3k \leq 2\delta n$ .

**Connecting the stars.** In this step, we fix an arbitrary cyclic ordering of the copies of  $K_{1,k}$  and  $K_{1,1}$ , and we connect each consecutive pair using the square of a short path. For the rest of the proof, we will refer to these (squares of) short paths as the connecting paths. We first explain the connection between two copies of  $K_{1,k}$  and assume without loss of generality that  $1, 2 \in \mathcal{I}$ ; when at least one of the copies is  $K_{1,1}$ , the connection is similar and will be explained later. We use the square of a path on six vertices with end-tuples within  $V_1$  and  $V_2$ , such that, again in view of Lemma 3.1.3, the end-tuples have many common neighbours into the other clusters  $U_{1,j}$  and  $U_{2,j}$  for  $j = 1, \dots, k$ , respectively. Recall that both  $U_{1,1}$  and  $U_{2,1}$  contain two more vertices than the other leaf clusters, one of which will be used for this connection.

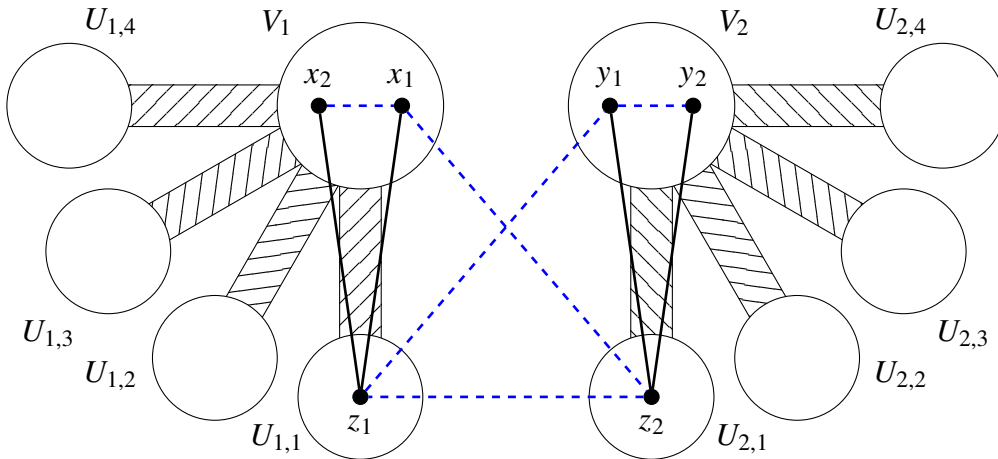


Figure 3.3: Construction of the square of a path with end-tuples  $(x_1, x_2)$  and  $(y_1, y_2)$  that connects two copies of  $K_{1,k}$  in the cluster graph  $R'$ . The dashed blue edges come from the random graph and the black edges from the deterministic graph.

We want to find  $x_1, x_2 \in V_1, z_2 \in U_{2,1}, z_1 \in U_{1,1}$ , and  $y_1, y_2 \in V_2$  such that the following

holds (see Figure 3.3):

- (A1) the tuples  $(x_1, x_2)$  and  $(y_1, y_2)$  have at least  $\frac{3}{4}d^2n_0$  common neighbours in each of  $U_{1,1}, \dots, U_{1,k}$  and  $U_{2,1}, \dots, U_{2,k}$  in the graph  $G$ , respectively;
- (A2)  $x_1z_1, x_2z_1, y_1z_2,$  and  $y_2z_2$  are edges of  $G$ ;
- (A3)  $x_2, x_1, z_2, z_1, y_1, y_2$  is a path in  $G_3$ .

For that we use Lemma 2.5.3 on the following collection  $H$  of tuples. We pick subsets  $X$  in  $V_1, Y$  in  $V_2, Z_1$  in  $U_{1,1}$ , and  $Z_2$  in  $U_{1,2}$ , all of size  $n_0/3$ , and we let  $H$  be the set of those tuples  $(x_2, x_1, z_2, z_1, y_1, y_2) \in X \times X \times Z_2 \times Z_1 \times Y \times Y$  which satisfy the properties (A1) and (A2). Note that  $H$  contains enough tuples for an application of Lemma 2.5.3 as, from regularity, we get that for all but at most  $2\epsilon n_0$  vertices  $x_1 \in X$ , all but  $4k\epsilon n_0$  vertices  $x_2 \in X$  have at least  $\frac{3}{4}d^2n_0$  common neighbours in  $U_{1,1}, \dots, U_{1,k}$  and at least  $\frac{1}{4}d^2n_0$  common neighbours in  $Z_1$  with respect to  $G$  and similarly for  $Y$  with  $U_{2,1}, \dots, U_{2,k}$  and  $Z_2$ . Therefore,  $H$  has size at least  $d^4 2^{-12} n_0^6 \geq \delta(kT)^{-6} n^6$ . Now we reveal the edges of  $G_3[X \cup Z_2 \cup Z_1 \cup Y]$  and with Lemma 2.5.3 and given the choice of  $C_3$ , we a.a.s. find a tuple  $(x_2, x_1, z_2, z_1, y_1, y_2) \in H$  satisfying the property (A3) as well. This will give the desired connecting square of a path. We then remove the vertices  $z_1$  and  $z_2$  from  $U_{1,1}$  and  $U_{2,1}$ , respectively.

Given an arbitrary cycling ordering of the indices from  $\mathcal{I}$  and  $\mathcal{J}$ , we use this construction to connect all neighbouring pairs  $(i, j)$ . If  $i, j \in \mathcal{I}$  we proceed as described above for the pair  $(1, 2)$ ; if  $i \in \mathcal{J}$ , we let  $U_{i,1}$  take the role of all  $U_{1,1}, \dots, U_{1,k}$ ; and if  $j \in \mathcal{J}$ , we let  $U_{j,1}$  take the role of  $V_2$  and  $V_j$  the role of all  $U_{2,1}, \dots, U_{2,k}$ . As we pick sets  $X, Y, Z_1$  and  $Z_2$  of size  $n_0/3$  and each cluster is involved in at most two distinct connections, we can choose disjoint sets for each connection and avoid clashes. Therefore each edge of  $G_3$  is revealed at most once and, as there are at most  $t'$  connections, we a.a.s. get all the desired edges of  $G_3$ . Observe that the choices done so far are needed for a later application of Lemma 3.1.3 and Lemma 3.1.4. Indeed for  $i \in \mathcal{I}$  there are two end-tuples of connecting paths in  $V_i$ , and for  $i \in \mathcal{J}$  there is one end-tuple of a connecting path in each of  $V_i$  and  $U_{i,1}$ ; moreover, for  $i \in \mathcal{I}$ , since  $U_{i,1}$  is involved in exactly two connections, during this construction we removed exactly two vertices from  $U_{i,1}$ , and now all  $U_{i,j}$ 's have the same size for  $j = 1, \dots, k$ .

**Covering  $V_0$ .** In the next step we cover all vertices of  $V_0$  by extending the connecting paths that we have already constructed, where we recall that  $|V_0| \leq 2\delta n$ . It is crucial for the rest of the argument (in particular for the applications of Lemma 3.1.3 and 3.1.4) that the conditions on the relation between the sizes of the clusters are still satisfied and that the end-tuples remain in the same clusters, i.e. if we extend the square of a path with one end-tuple in a cluster  $V_i$ , then the extended path needs to have the new end-tuple in the same cluster  $V_i$ . We will again make use of Lemma 2.5.3, with a suitable collection of tuples  $H$ . Before giving a precise description, we refer to Figure 3.4 and illustrate the

extension when  $k = 3$  and we want to cover a vertex  $v \in V_0$  by extending the square of a connecting path with end-tuple  $(x_1, x_2)$  in the centre cluster  $V_i$  of a copy of  $K_{1,3}$ . We will find 24 additional vertices as drawn in Figure 3.4, where we stress the following conditions. The vertices  $v_{3,1}, \dots, v_{3,6}$  are all neighbours of  $v$  in  $G$  (which will be guaranteed by the minimum degree condition), the blue edges are random edges (which will be guaranteed by Lemma 2.5.3), while the black edges are from the graph  $G$  (which will be guaranteed by regularity). Thus we extend the connecting path into the square of a path with end-tuples  $(x_2, x_1)$  and  $(x'_2, x'_1)$  still in  $V_i$ . Moreover, since six new vertices have been covered from each cluster, the relation between their sizes still holds.

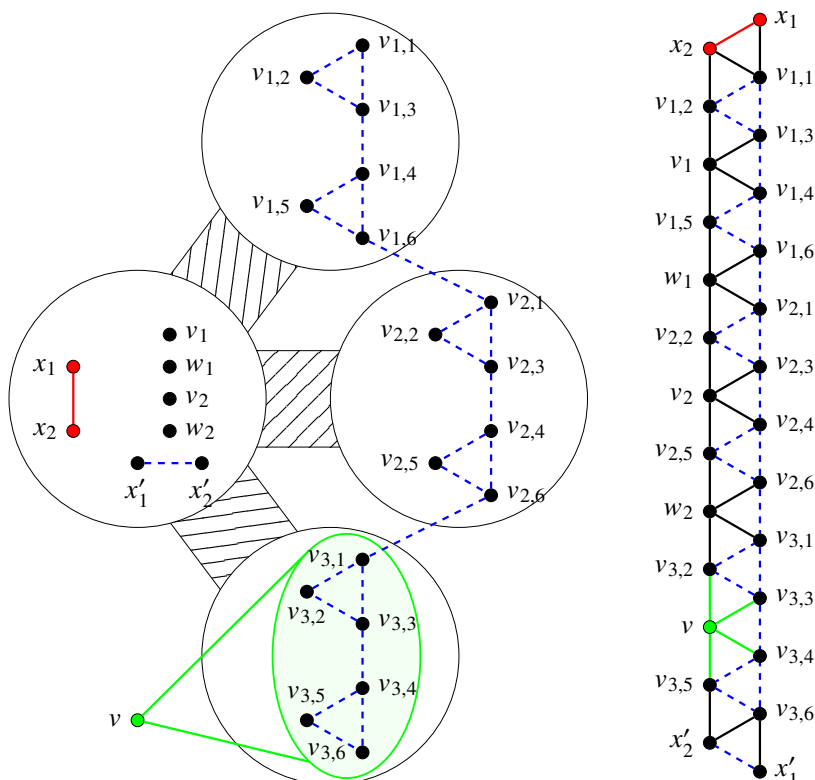


Figure 3.4: Covering one vertex  $v \in V_0$  by the square of a path for a copy of  $K_{1,k}$  in  $\mathcal{K}$  for  $k = 3$ . We keep the position of the end-tuples in  $V_i$  and the sizes of the clusters balanced. The edges within the clusters (dashed blue) come from the random graph and the edges between the clusters (black) come from regularity.

We will now give the details of these constructions and we start by defining the collections of tuples we will use for the applications of Lemma 2.5.3. For  $i \in \mathcal{I}$ , we let  $H_{1,i}$  be the set of those tuples in  $\prod_{j=1}^k (U_{i,j}^k \times U_{i,j}^k)$  such that the  $2k^2$  vertices in each tuple have at least  $\frac{1}{2}d^{2k^2}n_0$  common neighbours in  $V_i$  in  $G$ . Then we let  $H_1 = \bigcup_{i \in \mathcal{I}} H_{1,i}$ . Similarly, for  $j \in \mathcal{J}$ , we let  $H_{2,j}$  be the set of those tuples in  $V_j^8 \cup U_{j,1}^8$  such that the 8 vertices in each tuple have at least  $\frac{1}{2}d^8n_0$  common neighbours in the other set in  $G$ . Then we let  $H_2 = \bigcup_{j \in \mathcal{J}} H_{2,j}$ . Moreover, for  $i \in \mathcal{I}$ , we let  $H_{3,i}$  be the set of those tuples in  $V_i^2$  such that

the 2 vertices in each tuple have at least  $\frac{3}{4}d^2n_0$  common neighbours in each of the sets  $U_{i,j}$  for  $j = 1, \dots, k$ . For  $j \in \mathcal{J}$ , we let  $H_{3,j}$  be the set of those tuples in  $V_j^2 \cup U_{j,1}^2$  such that the 2 vertices in each tuple have at least  $\frac{3}{4}d^2n_0$  common neighbours in the other set in  $G$ . Then we let  $H_3 = \bigcup_{i \in \mathcal{I}} H_{3,i} \cup \bigcup_{j \in \mathcal{J}} H_{3,j}$ . With the constants specified in the beginning of the proof for obtaining  $C_4$ ,  $C_5$ , and  $C_6$ , we apply Lemma 2.5.3 to  $H_1$  with the random graph  $G_4$ , to  $H_2$  with  $G_5$ , and to  $H_3$  with  $G_6$ . Given the choice of the constant  $C$  done at the beginning, we can a.a.s. assume that

$$G_4, G_5 \text{ and } G_6 \text{ are all in the good event of Lemma 2.5.3 for the application above.} \quad (3.3.1)$$

Now we explain how we cover the vertices of  $V_0$  and we show that suitable subsets of  $H_1$ ,  $H_2$  and  $H_3$  are large enough for Lemma 2.5.3, i.e. larger than  $\delta(kT)^{-sk}n^{sk}$  where  $sk = 2k^2, 8, 2$ , respectively.

Given a vertex  $v \in V_0$ , we insist that the neighbours of  $v$  that we use to cover  $v$  do not come from any of the centre clusters  $V_i$  with  $i \in \mathcal{I}$ , because in this case we could not ensure to use the same number of vertices from each cluster in the copy of  $K_{1,k}$  and we would unbalance the star, creating issues for a later application of Lemma 3.1.3. Observe that, as we have  $|V_0 \cup (\bigcup_{i \in \mathcal{I}} V_i)| \leq 2\delta n + (\frac{1}{k+1} - \frac{4d}{k-1})t'n_0 \leq (\frac{1}{k+1} - \gamma)n - \frac{2d}{k-1}n$ , every vertex in  $G$  has at least  $\frac{2d}{k-1}n$  neighbours outside of  $V_0 \cup (\bigcup_{i \in \mathcal{I}} V_i)$  in the graph  $G$ . Moreover, for each vertex  $v \in V_0$ , we will use at most  $\max\{2k(k+1), 22\} \leq 6k^2$  vertices outside of  $V_0$ . During the process of covering  $V_0$ , we let  $V^* \subseteq V_0$  be the set of vertices of  $V_0$  already covered and  $W$  be the set of vertices outside of  $V_0$  that we already used to cover  $V^*$ ; at the beginning  $V^* = \emptyset$  and  $W = \emptyset$ . Note that we have  $|W| \leq 6k^2|V^*| \leq 12k^2\delta n$ . Then we let  $\mathcal{T} \subseteq \mathcal{I} \cup \mathcal{J}$  be the set of those indices  $i \in \mathcal{I} \cup \mathcal{J}$  such that  $U_{i,1}$  intersects  $W$  in at least  $\sqrt{\delta}n_0$  vertices and note that the bound on  $|W|$  implies that  $|\mathcal{T}| \leq 12k^2\sqrt{\delta}t'$ . We recall that at the beginning of the process for each  $i \in \mathcal{I}$  all the clusters  $U_{i,j}$ 's had the same size for  $j = 1, \dots, k$ . Similarly, for each  $i \in \mathcal{J}$  the clusters  $V_i$  and  $U_{i,1}$  had the same size as well. Since throughout the process we always cover the same number of vertices in each cluster of a copy of  $K_{1,1}$  or  $K_{1,k}$ , if  $i \notin \mathcal{T}$ , then  $W$  intersects each cluster of the copy corresponding to the index  $i$  in less than  $\sqrt{\delta}n_0$  vertices. Moreover, notice that as each  $v \in V_0$  has at least  $\frac{2d}{k-1}n$  neighbours outside of  $V_0 \cup (\bigcup_{i \in \mathcal{I}} V_i)$ , there are at least  $\frac{d}{k-1}t'$  clusters that are not the centre cluster of a copy of  $K_{1,k}$  and such that  $v$  has at least  $\frac{d}{k-1}n_0$  neighbours in it. Therefore, for every  $v \in V_0 \setminus V^*$  there exists  $i(v) \in (\mathcal{I} \cup \mathcal{J}) \setminus \mathcal{T}$  and some  $j(v)$  such that  $v$  has at least  $\frac{d}{k-1}n_0$  neighbours in  $U_{i(v),j(v)} \setminus W$  with respect to  $G$ .

Fix any  $v \in V_0 \setminus V^*$  and let  $i = i(v)$ . We start discussing the case  $i \in \mathcal{I}$ ; the case  $i \in \mathcal{J}$  is conceptually simpler and will be treated afterwards. Without loss of generality we can assume  $j(v) = k$ . Let  $(x_1, x_2)$  and  $(y_1, y_2)$  be the end-tuples in  $V_i$  of the connecting paths found in the previous step, and recall that  $x_1$  and  $x_2$  have at least  $\frac{3}{4}d^2n_0$  common neighbours in  $U_{i,j}$  for each  $j = 1, \dots, k$ . We will extend the square of the connecting path

ending in  $(x_1, x_2)$ , without using neither  $y_1$  nor  $y_2$ . Moreover, we will make sure that the new end-tuple  $(x'_2, x'_1)$  will still belong to  $V_i$ , and that  $x'_1$  and  $x'_2$  will have at least  $\frac{3}{4}d^2n_0$  common neighbours in  $U_{i,j}$  for each  $j = 1, \dots, k$ .

Let  $Z_1$  be the set  $(N_G(x_1, U_{i,1}) \cap N_G(x_2, U_{i,1})) \setminus (W \cup \{y_1, y_2\})$ , let  $Z_j$  be the set  $U_{i,j} \setminus W$  for  $j = 2, \dots, k-1$ , and let  $Z_k$  be the set  $N_G(v, U_{i,k}) \setminus W$ . Observe that  $|Z_1| \geq \frac{1}{2}d^2n_0$  since  $x_1$  and  $x_2$  have at least  $\frac{3}{4}d^2n_0$  common neighbours in  $U_{i,1}$ ,  $|Z_j| \geq (1 - \sqrt{\delta})n_0 \geq dn_0$  for  $j = 2, \dots, k-1$ , and  $|Z_k| \geq \frac{d}{k-1}n_0$  by the choice of  $i = i(v)$ . Using the regularity properties, analogously as above, it is easy to see that there are at least  $\delta(kT)^{-2k^2}n^{2k^2}$  tuples in  $\prod_{j=1}^k (Z_j^k \times Z_j^k)$  such that the  $2k^2$  vertices in each tuple have at least  $\frac{1}{2}d^{2k^2}n_0$  common neighbours in  $V_i$  in  $G$ . These tuples are in  $H_{1,i} \subseteq H_1$  as well and thus (3.3.1) guarantees that we find one of such tuples  $(v_{j,j'} : 1 \leq j \leq k, 1 \leq j' \leq 2k)$  in  $\prod_{j=1}^k (Z_j^k \times Z_j^k)$  where  $v_{j,j'} \in Z_j$ , such that in  $G_4$  we have the square of a path on  $v_{j,1}, \dots, v_{j,k}$ , the square of a path on  $v_{j,k+1}, \dots, v_{j,2k}$ , and the edges  $v_{j,k}v_{j,k+1}$  and  $v_{j,2k}v_{j+1,1}$ .

Let  $Z \subseteq V_i \setminus (W \cup \{y_1, y_2\})$  be the common neighbourhood of the vertices  $v_{j,j'}$  with  $1 \leq j \leq k$  and  $1 \leq j' \leq 2k$  in  $V_i$ , and observe  $|Z| \geq \frac{1}{4}d^{2k^2}n_0$ . Again using regularity, there are at least  $\delta(kT)^{-2}n^2$  tuples in  $Z^2 \cap H_3$ , and thus (3.3.1) guarantees that there is a tuple  $(x'_1, x'_2) \in Z^2 \cap H_3$  such that  $x'_1x'_2$  is an edge of  $G_6$  and  $x'_1$  and  $x'_2$  have at least  $\frac{3}{4}d^2n_0$  common neighbours in each  $U_{i,j}$  for  $j = 1, \dots, k$ . We then greedily pick additional  $2(k-1)$  vertices  $v_1, w_1, \dots, v_{k-1}, w_{k-1}$  in  $Z$  and we claim that

$$x_1, x_2, v_{1,1}, \dots, v_{1,k}, v_1, v_{1,k+1}, \dots, v_{1,2k}, w_1, \\ v_{2,1}, \dots, v_{2,k}, v_2, \dots, w_{k-1}, v_{k,1}, \dots, v_{k,k}, v, v_{k,k+1}, \dots, v_{k,2k}, x'_2, x'_1$$

is the square of a path with end-tuples  $(x_2, x_1)$  and  $(x'_2, x'_1)$  that contains  $v$ . Indeed,  $x_1$  and  $x_2$  are common neighbours of  $v_{1,1}$  and  $v_{1,2}$ , while  $x'_1$  and  $x'_2$  are common neighbours of  $v_{k,2k-1}$  and  $v_{k,2k}$ ; moreover the vertex  $v$  is a common neighbour of  $v_{k,k-1}, v_{k,k}, v_{k,k+1}$  and  $v_{k,k+2}$ , the vertex  $v_j$  is a common neighbour of  $v_{j,k-1}, v_{j,k}, v_{j,k+1}$  and  $v_{j,k+2}$ , and the vertex  $w_j$  is a common neighbour of  $v_{j,2k-1}, v_{j,2k}, v_{j+1,1}$  and  $v_{j+1,2}$ . This fills the gaps left after the initial construction above where we used Lemma 2.5.3 with the graph  $G_4$ . We now add  $v$  to  $V^*$  and all the other used vertices  $v_{j,j'}$  for  $1 \leq j \leq k$  and  $1 \leq j' \leq 2k$ ,  $v_i, w_i$  for  $1 \leq i \leq k-1$ , and  $x_1$  and  $x_2$  to  $W$ . Note we do not add  $x'_1$  and  $x'_2$  to  $W$  as  $(x'_2, x'_1)$  is the end-tuple of the extended square of a path. Observe that we used  $2k(k+1)$  vertices to cover  $v$  and that we covered exactly  $2k$  vertices from each cluster of the copy of  $K_{1,k}$ .

Now we move to the construction for the case  $i = i(v) \in \mathcal{J}$ . We will use always the same construction regardless of the value of  $k$  as illustrated in Figure 3.5 in a similar way as Figure 3.4 earlier. By assumption we have  $j(v) = 1$  and thus  $|N_G(v, U_{i,1} \setminus W)| \geq \frac{d}{k-1}n_0$ . Let  $(x_1, x_2) \in V_i^2$  and  $(y_1, y_2) \in U_{i,1}^2$  be the end-tuples of the connecting paths found in the previous step, and recall that  $x_1$  and  $x_2$  have at least  $\frac{3}{4}d^2n_0$  common neighbours in  $U_{i,1}$ . We will extend the square of the connecting path ending in  $(x_1, x_2)$ , without using neither  $y_1$

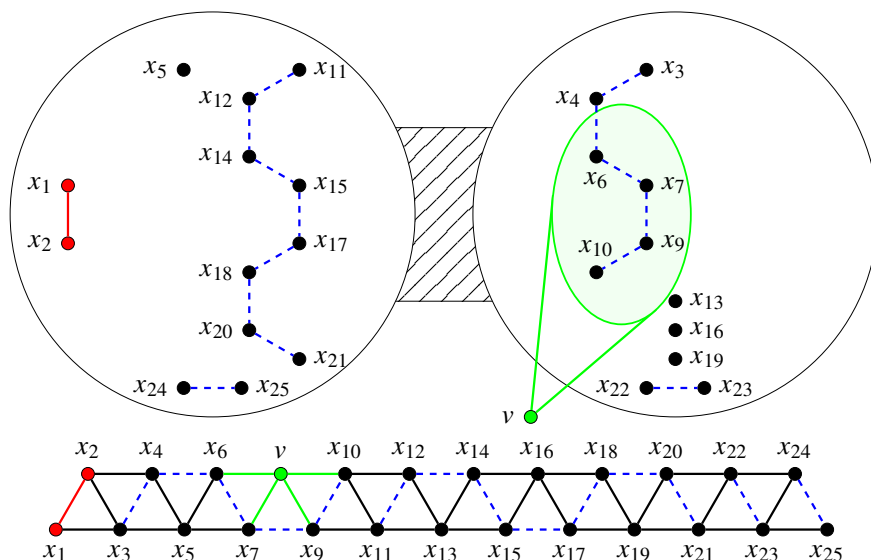


Figure 3.5: Covering one vertex  $v \in V_0$  by the square of a path for a copy of  $K_{1,1}$  in  $\mathcal{K}$ . We keep the position of the end-tuples in  $V_i$  and the sizes of the clusters balanced. The edges within the clusters (dashed blue) come from the random graph and the edges between the clusters (green and black) come from regularity.

nor  $y_2$ , by constructing the square of a path on pairwise distinct vertices  $x_1, x_2, \dots, x_{24}, x_{25}$ , where  $x_8 = v$ . Moreover, we will make sure that the new end-tuple  $(x_{24}, x_{25})$  will still belong to  $V_i$ , and that  $x_{24}$  and  $x_{25}$  will have at least  $\frac{3}{4}d^2n_0$  common neighbours in  $U_{i,1}$ . Let  $Z_3 = Z_4$  be the common neighbourhood of  $x_1$  and  $x_2$  in  $U_{i,1} \setminus (W \cup \{y_1, y_2\})$  and  $Z_6 = Z_7 = Z_9 = Z_{10}$  be the neighbourhood of  $v$  in  $U_{i,1} \setminus (W \cup \{y_1, y_2\})$ , and note that  $|Z_3| \geq \frac{1}{2}d^2n_0$  and  $|Z_6| \geq \frac{d}{k-1}n_0 - 2$ . Using regularity, it is easy to see that for at least  $\delta(kT)^{-8}n^8$  tuples in  $Z_3 \times Z_4 \times Z_6 \times Z_7 \times Z_9 \times Z_{10} \times U_{i,1} \times U_{i,1}$  the common neighbourhood of their vertices in  $V_i$  has size at least  $\frac{1}{2}d^8n_0$  (where we added the set  $U_{i,1}$  twice only for the following application of Lemma 2.5.3). Thus (3.3.1) guarantees we find a path in  $G_5$  on six vertices  $x_3, x_4, x_6, x_7, x_9, x_{10}$  (we ignore the two vertices in  $U_{i,1}$ ) with  $x_j \in Z_j$  and such that the set  $Z$  of the common neighbours of  $x_3, x_4, x_6, x_7, x_9, x_{10}$  in  $V_i \setminus (W \cup \{x_1, x_2\})$  has size  $|Z| \geq \frac{1}{4}d^8n_0$ . We let  $x_5$  be any vertex of  $Z$ . Again using regularity, there are at least  $\delta(kT)^{-8}n^8$  tuples in  $(Z \setminus \{x_5\})^8$  such that the eight vertices in the tuple have at least  $\frac{1}{2}d^8n_0$  common neighbours in  $U_{i,1}$ . These tuples are in  $H_2$  as well and thus (3.3.1) guarantees we find in  $G_5$  a path on eight vertices  $x_{11}, x_{12}, x_{14}, x_{15}, x_{17}, x_{18}, x_{20}, x_{21}$  such that these eight vertices belong to  $Z \setminus \{x_5\}$  and the set  $Z'$  of their common neighbours in  $U_{i,1} \setminus (W \cup \{y_1, y_2, x_3, x_4, x_5, x_6, x_7, x_9, x_{10}\})$  has size  $|Z'| \geq \frac{1}{4}d^8n_0$ . We let  $x_{13}, x_{16}$ , and  $x_{19}$  be any three vertices of  $Z'$ . Again with (3.3.1), we find an edge  $x_{22}x_{23}$  of  $G_6$  such that  $x_{22}, x_{23} \in Z' \setminus \{x_{13}, x_{16}, x_{19}\}$  and their common neighbourhood  $Z''$  in  $V_i \setminus (W \cup \{y_1, y_2, x_1, x_2, x_5, x_{11}, x_{12}, x_{14}, x_{15}, x_{17}, x_{18}, x_{20}, x_{21}\})$  has size at least  $\frac{1}{2}d^2n_0$ . With another application of (3.3.1) we find an edge  $x_{24}x_{25}$  of  $G_6$  such that  $x_{24}, x_{25} \in Z''$  and they have



at least  $\frac{3}{4}d^2n_0$  common neighbours in  $U_{i,1}$ . This gives the square of a path on the vertices  $x_1, \dots, x_7, v, x_9, \dots, x_{25}$ . We add  $v$  to  $V^*$  and  $x_1, \dots, x_7, x_9, \dots, x_{23}$  to  $W$  and remark that we do not remove neither  $x_{24}$  nor  $x_{25}$  as  $(x_{24}, x_{25})$  is the end-tuple of the extended square of a path. Observe that we used 22 vertices to cover  $v$  and that we covered exactly 11 vertices from each cluster  $V_i$  and  $U_{i,1}$ .

We keep covering the vertices of  $V_0$  in this way until  $V^* = V_0$ , then we remove the vertices in  $W$  from the clusters.

**Completing the square of the Hamilton cycle.** Before finishing the proof, we summarise what we have done so far, and, abusing notation, we still denotes the clusters by  $V_i, U_{i,1}, \dots, U_{i,k}$  for  $i \in \mathcal{I}$  and  $V_i, U_{i,1}$  for  $i \in \mathcal{J}$ , even if we removed several vertices from them in the previous steps of the proof while connecting stars and covering  $V_0$ . We have covered all vertices of  $G$ , except those that are still in the clusters of a copy of  $K_{1,1}$  or  $K_{1,k}$ , with the squares of short paths with the following properties. Their end-tuples belong to some  $V_i$  with  $i \in \mathcal{I}$ , or some  $V_j$  or some  $U_{j,1}$  with  $j \in \mathcal{J}$ . Moreover, for each  $i \in \mathcal{I}$ , the cluster  $V_i$  contains exactly two of such tuples, say  $(x_1, x_2)$  and  $(y_1, y_2)$ , such that  $x_1$  and  $x_2$  have at least  $\frac{3}{4}d^2n_0 - \sqrt{\delta}n_0 \geq \frac{1}{2}d^2n_0$  common neighbours in  $U_{i,j}$  for  $j = 1, \dots, k$ , and the same holds for  $y_1$  and  $y_2$ . Additionally, for each  $j \in \mathcal{J}$ , the clusters  $V_j$  and  $U_{j,1}$  contain exactly one such tuple each, say  $(x_1, x_2)$  and  $(y_1, y_2)$  respectively, such that  $x_1$  and  $x_2$  have at least  $\frac{1}{2}d^2n_0$  common neighbours in  $U_{j,1}$ , and  $y_1$  and  $y_2$  have at least  $\frac{1}{2}d^2n_0$  common neighbours in  $V_j$ . Therefore, it remains to cover the clusters by extending the squares of paths we already have.

Let  $i \in \mathcal{I}$ . Note that we still have that  $|U_{i,1}| = \dots = |U_{i,k}|$  and  $|V_i| - 4 - |U_{i,k}| \equiv -1 \pmod{3k-1}$  for  $i \in \mathcal{I}$ . Recall that we removed  $\delta n_0$  vertices from each  $U_{i,j}$  at the beginning. Moreover, while making the regular pairs super-regular and during the previous step, we removed the same number of vertices from each cluster of a copy of  $K_{1,k}$ , and at most  $2\sqrt{\delta}n_0$  vertices from each. Therefore, using that  $2\delta < \delta_0$ , we get  $|U_{i,j}| \geq (1 - \delta_0)(|V_i| - 4)$  for  $j = 1, \dots, k$ . Note that also  $|U_{i,k}| \leq (1 - \delta)(|V_i| - 4)$  still holds. Now let  $(x_1, x_2)$  and  $(y_1, y_2)$  be the two end-tuples in  $V_i$  of connecting paths and recall they both have at least  $\frac{1}{2}d^2n_0$  common neighbours in  $U_{i,j}$  for each  $j = 1, \dots, k$ . We apply Lemma 3.1.3 with the random graph  $G_1$  to  $V_i, U_{i,1}, \dots, U_{i,k}$  and get the square of a path with end-tuples  $(x_2, x_1)$  and  $(y_2, y_1)$ , covering all vertices in  $V_i \cup U_{i,1} \cup \dots \cup U_{i,k}$ .

For  $i \in \mathcal{J}$  we proceed similarly. We have  $|V_i| = |U_{i,1}|$  and two end-tuples  $(x_1, x_2)$  in  $V_i$  and  $(y_1, y_2)$  in  $U_{i,1}$  of connecting paths. Since  $x_1$  and  $x_2$  have at least  $\frac{1}{2}d^2n_0$  common neighbours in  $U_{i,1}$ , and  $y_1$  and  $y_2$  have at least  $\frac{1}{2}d^2n_0$  common neighbours in  $V_i$ , we can apply Lemma 3.1.4 with the random graph  $G_2$  to  $V_i, U_{i,1}$ , and get the square of a path with end-tuples  $(x_2, x_1)$  and  $(y_2, y_1)$ , covering all vertices in  $V_i \cup U_{i,1}$ .

This completes the square of a Hamilton cycle covering all vertices of  $G$ , and finishes the proof.  $\square$

### 3.4 Regularity in auxiliary graphs

The aim of this section is to prove the main technical lemma behind the embedding Lemma 3.1.3, whose proof is provided in the next section. In Lemma 3.1.3, we have to find the square of a Hamilton path (subject to additional conditions), where we can use both deterministic and random edges. Here we look at the edges coming from the random graph and show that we can find many disjoint copies of the square of a short path in the random graph, with some additional properties with respect to the deterministic graph, which guarantee we can then nicely connect them and build the desired structure. We state a more general version of the result we need for Lemma 3.1.3, as we believe it might be of independent interest and helpful in other problems of similar flavour. Before stating the lemma we introduce some definitions.

**Definition 3.4.1** (*H-transversal family*). Let  $H$  be a graph and set  $h = |V(H)|$ . Let  $V, U_1, \dots, U_h$  be disjoint sets of vertices and  $G$  be a graph on vertex set  $V \cup U_1 \cup \dots \cup U_h$ . A family  $\mathcal{H}$  of pairwise vertex-disjoint copies of  $H$  in  $G[U_1, \dots, U_h]$  is said to be a  $H$ -transversal family if, in addition, there exists a labelling  $v_1, \dots, v_h$  of the vertices of  $H$  such that, for each copy  $H \in \mathcal{H}$ , the vertex  $v_i$  is embedded into  $U_i$  for each  $i = 1, \dots, h$ .

With  $\mathcal{H}$  being an  $H$ -transversal family in  $G$ , we define the following auxiliary bipartite graph  $\mathcal{T}_G(\mathcal{H}, V)$ .

**Definition 3.4.2** (*Auxiliary graph  $\mathcal{T}_G(\mathcal{H}, V)$* ). We define  $\mathcal{T}_G(\mathcal{H}, V)$  to be the bipartite graph with partition classes  $\mathcal{H}$  and  $V$ , where the edge between  $H \in \mathcal{H}$  and  $v \in V$  appears if and only if all vertices of  $H$  are incident to  $v$  in  $G$ .

We prove the following general lemma.

**Lemma 3.4.3.** *Let  $H$  be a graph on  $h$  vertices. For any  $d, \delta, \varepsilon' > 0$  with  $2\delta \leq d$  there exist  $\varepsilon > 0$  and  $C > 0$  such that the following holds. Let  $V, U_1, \dots, U_h$  be sets with  $|V| = n$  and  $|U_i| = m = (1 \pm \frac{1}{2})n$  for  $i = 1, \dots, h$  such that  $(V, U_i)$  is  $(\varepsilon, d)$ -super-regular with respect to a graph  $G$  for  $i = 1, \dots, h$ . Furthermore, suppose that  $p \geq Cn^{-1/m_1(H)}$ .*

*Then a.a.s. there exists an  $H$ -transversal family  $\mathcal{H} \subseteq G(U_1, \dots, U_h, p)$  of size  $|\mathcal{H}| \geq (1 - \delta)m$  such that the pair  $(\mathcal{H}, V)$  is  $(\varepsilon', d^{h+1}2^{-h-3})$ -super-regular with respect to the auxiliary graph  $\mathcal{T}_G(\mathcal{H}, V)$ .*

The lemma shows that not only there is a large  $H$ -transversal family  $\mathcal{H}$ , but that we can additionally require that  $(\mathcal{H}, V)$  is a super-regular pair in  $\mathcal{T}_G(\mathcal{H}, V)$ . The proof of Lemma 3.1.3 will use the special case of  $H$  being the square of a path on  $k$  vertices.

Before giving a proof of Lemma 3.4.3, we introduce an auxiliary  $h$ -partite  $h$ -uniform hypergraph  $F = F_{G,V}(U_1, \dots, U_h)$  to encode the potential tuples in  $U_1 \times \dots \times U_h$  that we would like to use for building the copies of  $H$  for the family  $\mathcal{H}$ .

**Definition 3.4.4** (Auxiliary hypergraph  $F$ ). Let  $h \geq 1$  be an integer,  $V, U_1, \dots, U_h$  be pairwise disjoint sets with  $|V| = n$  and  $|U_i| = m = (1 \pm \frac{1}{2})n$  for  $i = 1, \dots, h$ . We define  $F = F_{G,V}(U_1, \dots, U_h)$ <sup>3</sup> to be the  $h$ -partite  $h$ -uniform hypergraph on  $U_1 \times \dots \times U_h$  where a tuple  $(u_1, \dots, u_h) \in U_1 \times \dots \times U_h$  is an edge of  $F$  if and only if the vertices  $u_1, \dots, u_h$  have at least  $\frac{1}{2}d^h n$  common neighbours in the set  $V$  in the graph  $G$ , i.e.  $|\bigcap_{i=1}^h N_G(u_i, V)| \geq \frac{1}{2}d^h n$ .

Similarly, given a set  $X \subseteq V$ , we call an edge  $(u_1, \dots, u_h) \in E(F)$  *good for  $X$*  if and only if there are at least  $\frac{1}{2}d^h |X|$  vertices in  $X$  that are incident to all of  $u_1, \dots, u_h$  in  $G$ . We denote by  $F_X$  the spanning subgraph of  $F$  with edges that are good for  $X$ .

However, we can only use those copies of  $H$  which actually do appear in the random graph, and we encode that by defining the random spanning subgraph  $\tilde{F}$  of  $F$ , as follows.

**Definition 3.4.5** (Auxiliary hypergraph  $\tilde{F}$ ). Let  $H$  be a graph on  $h$  vertices. Let  $V, U_1, \dots, U_h$  be pairwise disjoint sets and  $F$  be the hypergraph defined in Definition 3.4.4. After revealing the edges of the random  $h$ -partite graph  $G(U_1, \dots, U_h, p)$ , we denote by  $\tilde{F}$  the (random) spanning subhypergraph of  $F$  formed by those edges  $(u_1, \dots, u_h)$  of  $F$  for which the vertices  $u_1, \dots, u_h$  give a copy of  $H$  in the revealed random graph  $G(U_1, \dots, U_h, p)$ . We will say that  $\tilde{F}$  is supported by  $G(U_1, \dots, U_h, p)$ .

We remark that if  $(u_1, \dots, u_h) \in U_1 \times \dots \times U_h$  is an edge of  $\tilde{F}$ , then the vertices  $u_1, \dots, u_h$  give a copy of  $H$  in  $G(U_1, \dots, U_h, p)$  and have at least  $\frac{1}{2}d^h n$  common neighbours in the set  $V$  in the graph  $G$ . We state some additional properties of  $F$  below.

**Lemma 3.4.6.** *Let  $H$  be a graph on  $h$  vertices. Let  $0 < d < 1$  and  $\varepsilon \leq \min\{\frac{1}{2}d^{h-1}, d(1 - 2^{-1/h})\}$ . Let  $G$  be a graph on vertex set  $V \cup U_1 \cup \dots \cup U_h$  with  $|V| = n$  and  $|U_1| = \dots = |U_h| = m = (1 \pm \frac{1}{2})n$ , and assume  $(V, U_i)$  is a  $(\varepsilon, d)$ -super-regular pair with respect to  $G$  for each  $i = 1, \dots, h$ . Let  $F = F_{G,V}(U_1, \dots, U_h)$  be the hypergraph defined in Definition 3.4.4. Then the following holds:*

- (i) *The minimum degree of  $F$  is at least  $(1 - h\varepsilon)m^{h-1}$ .*
- (ii) *If  $|X| \geq 2\varepsilon nd^{1-h}$ , all but at most  $\varepsilon m$  vertices from each  $U_i$  have degree at least  $(1 - h\varepsilon)m^{h-1}$  in  $F_X$ .*

Moreover the subgraph  $\tilde{F}$  keeps roughly the expected number of edges of  $F$ .

**Lemma 3.4.7.** *For any graph  $H$  on  $h \geq 2$  vertices and any  $\delta > 0$  there exist  $\varepsilon > 0$  and  $C > 0$  such that the following holds for  $p \geq Cn^{-1/m_1(H)}$ . Let  $F = F_{G,V}(U_1, \dots, U_h)$  be the hypergraph defined in Definition 3.4.4. Then a.a.s. for any sets  $U'_i \subseteq U_i$  of size at least  $\delta m$  for  $i = 1, \dots, h$ , we have that  $\tilde{F}' = \tilde{F}[U'_1, \dots, U'_h]$  satisfies*

$$e(\tilde{F}') = (1 \pm \sqrt{\varepsilon}) \prod_{i=1}^h |U'_i| p^{e(H)}. \quad (3.4.1)$$

<sup>3</sup>We remark that the definition of  $F$  depends on  $d$  as well. However, as this will always be clear from the context, we omit writing  $d$  explicitly in  $F_{G,V}(U_1, \dots, U_h)$ .

Moreover, if  $|X| \geq 2\epsilon nd^{1-h}$ , then with probability at least  $1 - e^{-n}$ , for any choice of  $U'_1, \dots, U'_h$  as above and with  $\tilde{F}'_X = \tilde{F}_X[U'_1, \dots, U'_h]$ , we have

$$e(\tilde{F}'_X) \geq (1 - \sqrt{\epsilon}) \prod_{i=1}^h |U'_i| p^{e(H)}. \quad (3.4.2)$$

We remark that we will show a more general version of Lemma 3.4.7. Indeed our proof will only use that  $F$  satisfies (i) and that  $F_X$  satisfies (ii) for all  $X$  with  $|X| \geq 2\epsilon nd^{1-h}$ . Thus (3.4.1) holds for any  $h$ -partite  $h$ -uniform hypergraph  $F$  on partition classes  $U_1, \dots, U_h$  of size  $m = (1 \pm \frac{1}{2})n$  with minimum degree  $(1 - h\epsilon)m^{h-1}$ . Similarly (3.4.2) holds for all subgraphs  $F_X$  of such  $F$ , such that all but  $\epsilon m$  vertices in each class have degree at least  $(1 - h\epsilon)m^{h-1}$  in  $F_X$ .

The proof of Lemma 3.4.6 relies on a standard application of the regularity method, and that of Lemma 3.4.7 follows from an application of Chebyshev's and Janson's inequalities. Therefore we postpone them to a supplementary section of this chapter (Section 3.7), and we now turn to the main proof of this section.

*Proof of Lemma 3.4.3.* Given a graph  $H$  on  $h \geq 2$  vertices and  $d, \delta, \epsilon' > 0$  with  $2\delta \leq d$ , suppose that

$$\epsilon \ll \eta \ll \gamma \ll d, \delta, \epsilon'$$

are positive real numbers such that

$$\gamma \leq \frac{\epsilon'(1-\delta)}{2}, \quad \eta \log\left(\frac{1}{\eta}\right) \leq \frac{1}{4}\gamma, \quad \epsilon \leq \min\left\{\frac{1}{2}\eta d^{h-1}, d(1 - 2^{-1/h}), \frac{1}{14^2}\gamma^2\right\},$$

where additionally we require that  $\epsilon$  is small enough for Lemma 3.4.7 with input  $H$  and  $\delta$ . Furthermore, let  $C > 0$  be large enough for Lemma 3.4.7 with the same input.

Given  $n$ , let  $V, U_1, \dots, U_h$  be sets of size  $|V| = n$  and  $|U_i| = m = (1 \pm \frac{1}{2})n$  for  $i = 1, \dots, h$  such that  $(V, U_i)$  is  $(\epsilon, d)$ -super-regular with respect to a graph  $G$  for  $i = 1, \dots, h$ . To find the family  $\mathcal{H}$  we will now reveal edges of  $G(U_1, \dots, U_h, p)$  with probability  $p \geq Cn^{-1/m_1(H)}$  and consider the spanning subgraph  $\tilde{F}$  of  $F = F_{G,V}(U_1, \dots, U_h)$  as defined in Definitions 3.4.4 and 3.4.5. Let  $X \subseteq V$  be of size  $\eta n$ . With Lemma 3.4.7, we can assume that for all  $U'_i \subseteq U_i$  of size  $\delta m$  for  $i = 1, \dots, h$  we have that (3.4.1) and (3.4.2) hold, where the latter holds for all  $X$  as above by a union bound.

Using a random greedy process we now choose a family of transversal copies  $\mathcal{H}$  of  $H$  of size  $(1 - \delta)m$  in  $\tilde{F}$  as follows. Having chosen copies  $H_1, \dots, H_t \in \tilde{F}$  with  $t < (1 - \delta)m$ , we pick  $H_{t+1}$  uniformly at random from all edges of  $\tilde{F}$  that do not share an endpoint with any of  $H_1, \dots, H_t$ . This is possible since by (3.4.1) there is always an edge in  $\tilde{F}[U'_1, \dots, U'_h]$  for any subsets  $U'_i \subseteq U_i$  of size at least  $\delta m$  for  $i = 1, \dots, h$ , and thus a transversal copy of  $H$  in  $G(U_1, \dots, U_h, p)$ . For  $i = 1, \dots, t$ , we denote the  $i$ -th chosen copy of  $H$  for  $\mathcal{H}$  by  $H_i$ , and by  $\mathcal{H}_i$  the history  $H_1, \dots, H_i$ . It remains to show that a.a.s.  $(\mathcal{H}, V)$  is

$(\varepsilon', 2^{-h-3}d^{h+1})$ -super-regular with respect to the auxiliary graph  $\mathcal{T} = \mathcal{T}_G(\mathcal{H}, V)$ .

Observe that any  $H \in \mathcal{H}$  has  $|N_{\mathcal{T}}(H)| \geq \frac{1}{2}d^h n \geq d^{h+1}2^{-h-3}|V|$  by construction. Moreover for any  $v \in V$  we have  $|N_{\mathcal{T}}(v)| \geq 2^{-h-3}d^{h+1}m$ ; this can be shown as follows. Consider the first  $\frac{1}{2}dm$  chosen copies, then for  $i = 1, \dots, \frac{1}{2}dm$ , by (3.4.1), there are at most  $(1 + \sqrt{\varepsilon}) \prod_{i=1}^h |U_i| p^{e(H)}$  available copies to chose  $H_i$  from. On the other hand, as long as  $i < \frac{1}{2}dm$ , the vertex  $v$  has at least  $\frac{1}{2}dm \geq \delta m$  neighbours  $U'_i \subseteq U_i$  for  $i = 1, \dots, h$  that are not covered by the edges in  $\mathcal{H}_{i-1}$ . Therefore, by (3.4.1) there are at least  $(1 - \sqrt{\varepsilon}) \prod_{i=1}^h |U'_i| p^{e(H)}$  choices for  $H_i$  such that  $H_i \in N_{\mathcal{T}}(v)$ .

Hence, for  $i = 1, \dots, \frac{1}{2}dn$ , we get

$$\mathbb{P}[H_i \in N_{\mathcal{T}}(v) | \mathcal{H}_{i-1}] \geq \frac{(1 - \sqrt{\varepsilon}) \prod_{i=1}^h |U'_i| p^{e(H)}}{(1 + \sqrt{\varepsilon}) \prod_{i=1}^h |U_i| p^{e(H)}} \geq \frac{(1 - \sqrt{\varepsilon})(\frac{1}{2}d)^h}{(1 + \sqrt{\varepsilon})} \geq 2^{-h-1}d^h.$$

As this holds independently of the history of the process, this process dominates a binomial distribution with parameters  $\frac{1}{2}dm$  and  $2^{-h-1}d^h$ . Therefore, even though the events are not mutually independent, we can use Chernoff's inequality (Lemma 2.4.1) to infer that  $|N_{\mathcal{T}}(v)| \geq 2^{-h-3}d^{h+1}m$  with probability at least  $1 - n^{-2}$ . Then, by applying the union bound over all  $v \in V$ , we obtain that a.a.s.  $|N_{\mathcal{T}}(v)| \geq 2^{-h-3}d^{h+1}m \geq 2^{-h-3}d^{h+1}|\mathcal{H}|$  for all  $v \in V$ .

Next let  $X \subseteq V$  be any subset with  $|X| = \eta n$ , and let  $t = (1 - \delta)m$ . For  $i = 0, 1, \dots, t-1$ , we obtain from (3.4.1) that there are at most  $(1 + \sqrt{\varepsilon}) \prod_{i=1}^h |U_i| p^{e(H)}$  edges in  $\tilde{F} \setminus \mathcal{H}_i$  available for choosing  $H_{i+1}$ , of which, by (3.4.2), at least  $(1 - \sqrt{\varepsilon}) \prod_{i=1}^h |U'_i| p^{e(H)}$  are in  $\tilde{F}_X$ . Then

$$\mathbb{P}\left[H_i \text{ good for } X \mid \mathcal{H}_{i-1}\right] \geq \frac{(1 - \sqrt{\varepsilon}) \prod_{i=1}^h |U'_i| p^{e(H)}}{(1 + \sqrt{\varepsilon}) \prod_{i=1}^h |U_i| p^{e(H)}} \geq (1 - 2\sqrt{\varepsilon}).$$

Again, as the lower bound on the probability holds independently of the history of the process, this process dominates a binomial distribution with parameters  $(1 - \delta)m$  and  $(1 - 2\sqrt{\varepsilon})$ . We let  $B_X \subseteq \mathcal{H}$  be the copies in  $\mathcal{H}$  that are not good for  $X$  and deduce

$$\mathbb{E}[|B_X|] \leq (1 - \delta)m 2\sqrt{\varepsilon} \leq 2\sqrt{\varepsilon}m.$$

Then we get from Chernoff's inequality (Lemma 2.4.1) that, since  $\gamma m \geq 14\sqrt{\varepsilon}m \geq 7\mathbb{E}[|B_X|]$ , we have

$$\mathbb{P}[|B_X| > \gamma m] \leq \exp(-\gamma m).$$

There are at most  $\binom{n}{\eta n} \leq \left(\frac{e}{\eta}\right)^{\eta n} \leq \exp(\eta \log(\frac{1}{\eta})n) \leq \exp(\frac{1}{2}\gamma m)$  choices for  $X$  and, thus, with the union bound over all these choices, we obtain that a.a.s. there are at most  $\gamma m$  bad copies in  $\mathcal{H}$  for any  $X \subseteq V$  with  $|X| = \eta n$ .

Fix a choice of  $\mathcal{H}$  such that there are at most  $\gamma m$  bad copies for any  $X \subseteq V$  with  $|X| = \eta n$ .

Then for any set  $X' \subseteq V$  and  $\mathcal{H}' \subseteq \mathcal{H}$  with  $|X'| \geq \varepsilon'n$  and  $|\mathcal{H}'| \geq \varepsilon'|\mathcal{H}|$  we find

$$e_{\mathcal{T}}(\mathcal{H}', X') \geq (|\mathcal{H}'| - \gamma m) \frac{d^h \eta n |X'|}{2 \cdot 2\eta n} \geq \frac{d^h}{8} |\mathcal{H}'| |X'|$$

by partitioning  $X'$  into pairwise disjoint sets of size  $\eta n$ .

We conclude that a.a.s. the pair  $(\mathcal{H}, V)$  is  $(\varepsilon', d^{h+1}2^{-h-3})$ -super-regular with respect to  $\mathcal{T}_G(\mathcal{H}, V)$ .  $\square$

### 3.5 Proof of the embedding Lemmas 3.1.3 and 3.1.4

In this section we prove Lemma 3.1.3 and then derive Lemma 3.1.4. Throughout the section we denote the square of a path on  $k$  vertices by  $H^{(k)}$ , we list their vertices as  $u_1, \dots, u_k$ , meaning that the edges of the square are  $u_i u_j$  for each  $1 \leq |i - j| \leq 2$ , and we recall that its end-tuples are  $(u_2, u_1)$  and  $(u_{k-1}, u_k)$ . We start with a short overview of our argument for the proof of Lemma 3.1.3, where we want to find the square of a Hamilton path covering  $V, U_1, \dots, U_k$  and with end-tuples  $(x, x')$  and  $(y, y')$ . Our proof will follow four steps, and the decomposition of this square of a Hamilton path in random and deterministic edges is outlined in Figure 3.6.

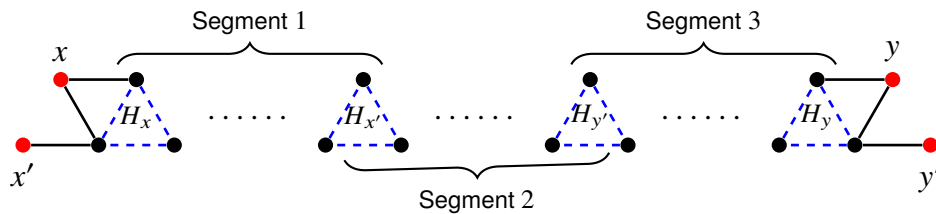


Figure 3.6: The square of a Hamilton path with end-tuples  $(x, x')$  and  $(y, y')$  in Lemma 3.1.3 and its decomposition into edges from  $G$  (black) and from  $G(n, p)$  (dashed blue). Each dotted  $H_x, H_{x'}, H_{y'}, H_y$  stands for a copy of  $P_k^2$  with edges all from  $G(n, p)$ . Segment 1 and 3 (resp. segment 2) are realised through several copies of the structure in Figure 3.8 (resp. Figure 3.7).

To ensure that  $(x, x')$  and  $(y, y')$  are the end-tuples of the square of the path, we will first find copies  $H_x$  and  $H_y$  of  $H^{(k)}$  that are connected to the tuples  $(x, x')$  and  $(y, y')$  (c.f. Figure 3.6). Moreover, with Lemma 3.4.3, we will find a large family  $\mathcal{H}$  of transversal copies of  $H^{(k)}$  in  $U_1, \dots, U_k$  (c.f. Definition 3.4.1) such that  $(\mathcal{H}, V)$  is super-regular with respect to the auxiliary graph  $\mathcal{T}_G(\mathcal{H}, V)$  (c.f. Definition 3.4.2). In particular, this will ensure that most pairs  $(H, H') \in \mathcal{H}^2$  have many common neighbours in  $V$  in the graph  $G$ .

The next step is to find random edges between the copies in  $\mathcal{H}$ . For that we will consider a directed auxiliary graph  $\bar{F}$  with vertex set  $\mathcal{H}$  where, given  $H, H' \in \mathcal{H}$ , with  $H$  on  $u_1, \dots, u_k$  and  $H'$  on  $u'_1, \dots, u'_k$ , the pair  $(H, H')$  is an edge of  $\bar{F}$  if and only if  $u_k u'_1$  is a random edge and  $V(H) \cup V(H')$  have many common neighbours in  $V$  in the graph  $G$ . This

will allow us to connect  $H$  to  $H'$  with a random edge, while also giving many choices for vertices from  $V$  to turn this into the square of a path on  $2k + 1$  vertices (c.f. Figure 3.7).

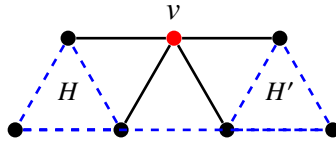


Figure 3.7: Connecting two copies  $H$  and  $H'$  of  $P_k^2$ , using one vertex  $v \in V$  (red), edges from  $G(n, p)$  (dashed blue), and edges from  $G$  (black).

We will use a random greedy procedure to find a long directed path  $D$  in  $\overline{F}$  that covers most of  $\mathcal{H}$ , which is possible by the properties of  $\mathcal{H}$  and our choice of  $p \geq Cn^{-(k-1)/(2k-3)} \geq Cn^{-1}$ . Additionally, we can guarantee that we will later be able to extend this into the square of a path using any subset of vertices  $V' \subset V$  of the right size. We denote the first and last copy of the path  $D$  by  $H'_x$  and  $H'_y$ , respectively.

In the next step, we take care of the set  $Z$  of those vertices in  $U_1 \cup \dots \cup U_k$  that are not covered by any copy of  $H^{(k)}$  from  $\mathcal{H}$ . We will absorb each vertex  $z \in Z$  into the square of a short path, using four vertices from  $V$ , two copies of  $H^{(k)}$  in  $\mathcal{H} \setminus V(D)$ , and random edges within  $V$  (c.f. Figure 3.8). In fact, we will be able to do that simultaneously for each vertex in  $Z$ , by constructing two squares of paths, one from  $H_x$  to  $H'_x$  and one from  $H'_y$  to  $H_y$ , which contain all vertices of  $Z$  and all copies of  $H^{(k)}$  in  $\mathcal{H} \setminus V(D)$ .

In the final step, we will find a perfect matching between the edges  $(H, H')$  of  $D$  and the remaining vertices of  $V$ , while making sure that the sizes of the two sets we want to match are the same. A vertex  $v \in V$  can be matched to  $(H, H')$  if and only if  $u_k, u_{k-1}, u'_1, u'_2$  are neighbours of  $v$  in  $G$  (with the labelling of the vertices of  $H, H'$  as above). This matching will close the gap between the two copies  $H$  and  $H'$  for each edge  $(H, H')$  of  $D$  with a vertex  $v$  from  $V$ . This will give the square of a path from  $H'_x$  to  $H'_y$  and, together with the other pieces from  $H_x$  to  $H'_x$  and from  $H_y$  to  $H'_y$ , we will get the square of a Hamilton path with the correct end-tuples. Ultimately, the shape of this square of the path is as illustrated in Figure 3.6, where the segments between  $H_x$  and  $H_{x'}$  and between  $H_{y'}$  and  $H_y$  (resp. between  $H_{x'}$  and  $H_{y'}$ ) are obtained by repeatedly inserting Figure 3.8 (resp. Figure 3.7) several times. We will now turn to the details of the argument.

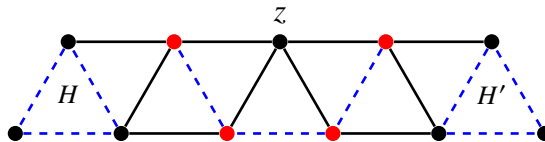


Figure 3.8: Absorbing a vertex  $z \in Z \subseteq U_1 \cup \dots \cup U_k$ , using two copies  $H$  and  $H'$  of  $P_k^2$ , four vertices from  $V$  (red), edges from  $G(n, p)$  (dashed blue), and edges from  $G$  (black).

*Proof of Lemma 3.1.3.* Given an integer  $k \geq 2$ , let  $H^{(k)}$  be the square of a path on  $k$  vertices and observe that  $m_1(H^{(k)}) = \frac{2k-3}{k-1}$ . Given  $0 < \delta' \leq d \leq 1$ , let  $\delta_1, \delta_0, \varepsilon' > 0$  with  $2\delta_1 < \delta_0 < \min\{\delta', d^{3k+3}2^{-3k-20}\}$  and  $\varepsilon' < \delta_1^8$ . Let  $C_1$  be given by Lemma 2.5.1 for input  $2\delta_1$  and  $F$ , where  $F$  is the path on four vertices. Then let  $\varepsilon_2$  and  $C_2$  be given by Lemma 3.4.3 for input  $H^{(k)}$ ,  $d$ ,  $\varepsilon'$ , and  $\varepsilon'$ , where  $\varepsilon'$  plays also the role of  $\delta$  in the statement of Lemma 3.4.3. Let  $\varepsilon_3$  and  $C_3$  be given by Lemma 3.4.7 with input  $H^{(k)}$  and  $\delta_0$ . Finally let  $\varepsilon < \min\{\varepsilon_2, \varepsilon_3, \varepsilon'/4\}$  and  $C = \max\{C_1, 2C_2, 2C_3, 48\varepsilon'^{-1}\delta^{-1}\}$ . Observe that  $\varepsilon < 2\delta_1$ , as required.

Let  $G$  be a graph on  $V \cup U_1 \cup \dots \cup U_k$ , where  $V, U_1, \dots, U_k$  are pairwise disjoint sets of size  $|V| = n+4$  and  $(1-\delta_0)n \leq |U_i| = m \leq (1-\delta_1)n$  for  $i = 1, \dots, k$  such that  $n-m \equiv -1 \pmod{3k-1}$ . Suppose that  $(V, U_i)$  is a  $(\varepsilon, d)$ -super-regular pair for  $i = 1, \dots, k$ . Further let  $(x, x')$  and  $(y, y')$  be two tuples from  $V$  such that both tuples have  $d^2m/2$  common neighbours in  $U_i$  for  $i = 1, \dots, k$  in the graph  $G$ . Let  $\delta$  be such that  $m = (1-\delta)n + 1$  and observe that  $\delta_1 \leq \delta \leq \delta_0 + 1/n$  and that the divisibility condition on  $n-m$  implies that  $\delta n \equiv 0 \pmod{3k-1}$ . For later convenience, we define  $m_0 = \left(1 - \frac{3k}{3k-1}\delta\right)n - 1$ ,  $t = \left(1 - \frac{4k}{3k-1}\delta\right)n + 1$ , and  $s = \frac{k}{3k-1}\delta n$ . Since  $\delta n \equiv 0 \pmod{3k-1}$ , we observe that  $m_0$ ,  $t$ , and  $s$  are positive integers. Moreover, we have that

$$m_0 - t = s - 2, \tag{3.5.1}$$

$$(1-\delta)n - 1 - m_0 = \frac{s}{k}, \tag{3.5.2}$$

$$n - 4s = t - 1. \tag{3.5.3}$$

We let  $p \geq Cn^{-(k-1)/(2k-3)}$  and reveal  $G(n, p)$  in three rounds  $G_0 \sim G(U_1, \dots, U_k, \frac{1}{2}p)$ ,  $G_1 \sim G(U_1, \dots, U_k, \frac{1}{2}p)$ , and  $G_2 \sim G(V, p)$ . Moreover we assume that a.a.s. the events of Lemma 2.5.1 and 3.4.7 hold in  $G_2$  and  $G_0$ , respectively.

**Finding transversal copies of  $H^{(k)}$ .** We start by ensuring that  $(x, x')$  and  $(y, y')$  can be the end-tuples of the square of a path. Fix  $i = 1, \dots, k$ . Recall that each of the tuples  $(x, x')$  and  $(y, y')$  has  $d^2m/2$  common neighbours in  $U_i$ . Thus we can pick disjoint sets  $U_{i,x}, U_{i,y} \subset U_i$  of size  $d^2m/4$  such that  $U_{i,x}$  and  $U_{i,y}$  are in the common neighbourhoods of  $(x, x')$  and  $(y, y')$  in  $U_i$ , respectively. Let  $F$  and  $\tilde{F}$  be the hypergraphs defined in Definition 3.4.4, and with  $H^{(k)}$  and  $\tilde{F}$  supported by  $G_0$ . Then with Lemma 3.4.7 we a.a.s. find an edge in  $\tilde{F}[U_{1,x}, \dots, U_{k,x}]$  and an edge in  $\tilde{F}[U_{1,y}, \dots, U_{k,y}]$ . Given the definition of  $\tilde{F}$ , these edges correspond to copies of  $H^{(k)}$  in  $G_0$  and we denote them by  $H_x = (x_1, \dots, x_k)$  and  $H_y = (y_1, \dots, y_k)$ .

For  $i = 1, \dots, k$ , let  $U'_i = U_i \setminus \{x_i, y_i\}$  and observe that  $|U'_i| = (1-\delta)n - 1$  and  $(U'_i, V)$  is  $(2\varepsilon, d/2)$ -super-regular. Then we apply Lemma 3.4.3 with  $G_1[U'_1, \dots, U'_k]$  to a.a.s. obtain a family  $\mathcal{H}$  of transversal copies of  $H^{(k)}$ , of size  $|\mathcal{H}| \geq (1-\varepsilon')(m-2) \geq \left(1 - \frac{3k}{3k-1}\delta\right)n - 1 = m_0$  and such that the pair  $(\mathcal{H}, V)$  is  $(\varepsilon', d^{k+1}2^{-k-3})$ -super-regular



with respect to  $\mathcal{T}_G(\mathcal{H}, V)$ , where  $\mathcal{T}_G(\mathcal{H}, V)$  is the graph defined in Definition 3.4.2. By removing some copies of  $H^{(k)}$ , we can assume that  $|\mathcal{H}| = m_0$  and still have that the pair  $(\mathcal{H}, V)$  is  $(\varepsilon', d^{k+1}2^{-k-4})$ -super-regular with respect to  $\mathcal{T}_G(\mathcal{H}, V)$ .

**Building the directed path  $D$ .** Ultimately we want to find a directed path  $D$ , that has some of the copies of  $H^{(k)}$  in  $\mathcal{H}$  as vertices. As we later would like to connect two copies of  $H^{(k)}$  by one random edge and a vertex from  $V$  to get the square of a path on  $2k + 1$  vertices (c.f. Figure 3.7), we want them to appear consecutively in  $D$  only if all their vertices have enough common neighbours in  $V$  in the graph  $G$ . We encode this condition in the auxiliary graph  $F^*$  with vertex set  $\mathcal{H}$ , and where, given  $H, H' \in \mathcal{H}$ , the edge  $HH'$  is in  $F^*$  if and only if the vertices in  $V(H) \cup V(H')$  have at least  $d^{2k+2}2^{-2k-7}n$  common neighbours in  $V$  in the graph  $G$ .

**Claim 3.5.1.** *The minimum degree of  $F^*$  is at least  $(1 - 2\varepsilon')m_0$ .*

*Proof of Claim 3.5.1.* Any copy of  $H \in \mathcal{H}$  has degree at least  $d^{k+1}2^{-k-3}n$  into  $V$  in the graph  $\mathcal{T}_G(\mathcal{H}, V)$ . Then, by super-regularity in  $\mathcal{T}_G(\mathcal{H}, V)$ , all but  $2\varepsilon'm_0$  copies of  $H' \in \mathcal{H}$  have at least  $d^{2k+2}2^{-2k-7}n$  common neighbours with  $H$ . This implies that  $H$  has degree  $(1 - 2\varepsilon')m_0$  in  $F^*$ .  $\square$

For a set  $X \subseteq V$ , we call an edge  $HH' \in E(F^*)$  *good for  $X$*  if there is at least one vertex  $v \in X$  that is incident to  $H$  and  $H'$  in  $\mathcal{T}_G(\mathcal{H}, V)$ . We denote the subgraph of  $F^*$  with edges that are good for  $X$  by  $F_X^*$ .

**Claim 3.5.2.** *If  $|X| \geq d^{-k-1}2^{k+4}\varepsilon'n$ , then all but at most  $\varepsilon'n$  vertices of  $\mathcal{H}$  have degree at least  $(1 - 4\varepsilon')m_0$  in  $F_X^*$ .*

*Proof of Claim 3.5.2.* All but at most  $\varepsilon'n$  copies  $H \in \mathcal{H}$  have degree at least  $d^{k+1}2^{-k-4}|X| \geq \varepsilon'n$  into  $X$  in the graph  $\mathcal{T}_G(\mathcal{H}, V)$ . Fixing any  $H \in \mathcal{H}$  with this property, all but at most  $\varepsilon'm_0$  copies  $H' \in \mathcal{H} \setminus \{H\}$  have at least one common neighbour with  $H$  in  $X$ . From Claim 3.5.1 we know that  $\delta(F^*) \geq (1 - 2\varepsilon')m_0$  and, therefore, all but at most  $\varepsilon'n$  vertices from  $\mathcal{H}$  have degree at least  $(1 - 4\varepsilon')m_0$  in  $F_X^*$ .  $\square$

We now define an auxiliary directed graph  $\bar{F}$  on vertex set  $\mathcal{H}$  as follows. Given any  $H$  and  $H' \in \mathcal{H}$  with  $H = u_1, \dots, u_k$  and  $H' = u'_1, \dots, u'_k$ , the pair  $(H, H')$  is a directed edge of  $\bar{F}$  if and only if  $HH'$  is an edge of  $F^*$  (which means that  $H$  and  $H'$  have many common neighbours in  $V$  in  $G$ ) and  $u_k u'_1$  is an edge of  $G_1$ . Note that for each edge  $HH'$  of  $F^*$  we get the directed edge  $(H, H')$  with probability  $p/2$  independently of all the other edges (in particular, independently of the edge  $(H', H)$ ). Therefore, we can also reveal each edge  $(H, H')$  in  $F^*$  with probability  $p/2$  independently of all the others and every edge of  $G_1$  will be revealed with probability at most  $p/2$  independently of all the others. We denote the resulting directed graph by  $\bar{F}$  and we want to find a long directed path  $D$  in  $\bar{F}$ .

For this we will use a random greedy process that explores the graph using a depth-first search algorithm. We do not reveal all the edges of  $F^*$  at the beginning, but, at each step of the algorithm, we only reveal those edges that are relevant for that step, and add those which are successful to  $\bar{F}$ . In each step the algorithm maintains a directed path  $H_1, \dots, H_r$  in  $\bar{F}$  and the set  $B$  of the vertices in  $\mathcal{H} \setminus \{H_1, \dots, H_r\}$ , whose neighbours have already been all revealed (we call them “dead-ends”). Additionally, we keep track of the vertices which have already been visited at least twice (due to backtracking of the algorithm) and denote their set by  $A$ . We initialise  $r = 0$ ,  $A = \emptyset$ , and  $B = \emptyset$ .

The algorithm proceeds as follows. If  $r = 0$ , then we choose an arbitrary vertex  $H_1 \in \mathcal{H} \setminus B$  and increase  $r$  by one. If  $r > 0$ , we let  $\mathcal{H}' = \mathcal{H} \setminus (\{H_1, \dots, H_r\} \cup B)$  be the vertices that have not been used and that are not “dead-ends”. If  $H_r \notin A$ , then from  $F^*$  we reveal all directed edges from  $H_r$  to  $\mathcal{H}'$  with probability  $p/2$ . If possible, we pick one neighbour uniformly at random among all those that are successful, denote it by  $H_{r+1}$ , and increase  $r$  by one. If none of them is successful, we add  $H_r$  to  $B$ ,  $H_{r-1}$  to  $A$ , and decrease  $r$  by one. If  $H_r \in A$ , then the directed edges from  $H_r$  to  $\mathcal{H}'$  in  $F^*$  have already been revealed earlier. If there is an edge from  $H_r$  to some  $H' \in \mathcal{H}'$  in  $\bar{F}$ , we let  $H_{r+1} = H'$  and increase  $r$  by one. Otherwise, we add  $H_r$  to  $B$ ,  $H_{r-1}$  to  $A$ , and decrease  $r$  by one. The algorithm stops if  $r = (1 - \frac{4k}{3k-1}\delta)n + 1$  or when  $|B| \geq \varepsilon'n$ , whichever happens first. We claim that the algorithm terminates and the latter does not happen.

**Claim 3.5.3.** *The graph  $\bar{F}$  a.a.s. contains a directed path  $D$  on  $t$  vertices (with  $t$  being  $(1 - \frac{4k}{3k-1}\delta)n + 1$ , as defined above).*

*Proof of Claim 3.5.3.* First, we observe that the algorithm terminates. Indeed, if  $|B| < \varepsilon'n$  then  $|\mathcal{H}'| \geq |\mathcal{H}| - t - |B| > 2\varepsilon'm_0$  and with Claim 3.5.1 there is at least one edge of  $F^*$  from  $H_r$  to  $\mathcal{H}'$ . Secondly, we claim that a.a.s.  $|B| < \varepsilon'n$ . Assume that at some point we have  $|B| = \varepsilon'n$  and  $r < t = (1 - \frac{4k}{3k-1}\delta)n + 1$ . Since at least  $|\mathcal{H}| - r - |B| \geq \delta n/4$  vertices of  $\mathcal{H}$  are not covered by the path or a vertex from  $B$ , we can pick a set  $\mathcal{H}'$  of exactly  $\delta n/4$  of them. This implies that all edges from  $B$  to  $\mathcal{H}'$  that are in  $F^*$  have been revealed but none is present in  $\bar{F}$ . However, with Claim 3.5.1, the expected number of edges in  $\bar{F}$  from  $B$  to  $\mathcal{H}'$  with  $|B| = \varepsilon'n$  and  $|\mathcal{H}'| = \frac{1}{4}\delta n$  is  $\frac{1}{8}p\varepsilon'\delta n^2$  and by Chernoff’s inequality (Lemma 2.4.1) the probability that none of the edges in  $F^*$  from  $B$  to  $\mathcal{H}'$  appears in  $\bar{F}$  is at most  $2 \exp(-\frac{1}{24}\varepsilon'p\delta n^2) \leq 2 \exp(-2n)$ . A union bound over the at most  $2^{2n}$  choices for  $B$  and  $\mathcal{H}'$  implies that the probability that there exist  $B$  and  $\mathcal{H}'$  as above is  $o(1)$ . Therefore the process stops when  $r = t$  and we obtain a directed path  $D$  on  $t$  vertices in  $\bar{F}$ .  $\square$

**Preparing the final matching.** The previous claim already guarantees a long directed path, but we need some additional properties that enable us to finish the proof later. An edge of  $D$  corresponds to an edge  $HH'$  of  $F^*$  and, for each of them, there are many choices for a vertex  $v \in V$  that turns this into the square of a path on  $2k + 1$  vertices. We will need

to do this simultaneously for all edges of  $D$  in the last step of the proof. However, before the last step, we have to cover the leftover of  $U_1 \cup \dots \cup U_k$  and  $\mathcal{H}$ , which will be possible by using some vertices of  $V$ . This will leave a subset  $V' \subseteq V$  of size  $t - 1$  to match to the edges of  $D$  in the last step, where we remark that the path  $D$  has exactly  $t - 1$  edges. We now show that this is possible for any subset  $V' \subseteq V$  of size  $t - 1$ . We encode this task as follows. Given a vertex  $v \in V$ , we define  $\mathcal{F}_v$  to be the set of all pairs  $(H, H') \in \mathcal{H}^2$  such that both  $H$  and  $H'$  are adjacent to  $v$  in  $\mathcal{T}_G(\mathcal{H}, V)$ ; note that this means that  $v$  is adjacent to all vertices in  $V(H) \cup V(H')$  in the graph  $G$ . Then we define an auxiliary bipartite graph  $\mathcal{F}_D$  with partition  $E(D)$  and  $V$ , where for  $e \in E(D)$  and  $v \in V$ , the pair  $ev$  is an edge of  $\mathcal{F}_D$  if and only if  $e \in \mathcal{F}_v$ .

**Claim 3.5.4.** *A.a.s. for any  $V' \subseteq V$  of size  $t - 1$  the graph  $\mathcal{F}_D[V', E(D)]$  contains a perfect matching.*

*Proof of Claim 3.5.4.* It suffices to show that for each  $X \subseteq V$  of size at most  $t - 1$  we have  $|\bigcup_{v \in X} N_{\mathcal{F}_D}(v)| \geq |X|$  and then the result follows by Hall's condition. We will use the notation from the algorithm introduced above.

Let  $X \subseteq V$  of size at most  $t - 1$  be given. First suppose that  $|X| > (1 - d^{k+2}2^{-2k-8})n$  and let  $e = (H, H')$  be any edge of  $D$ . Since  $HH'$  is in particular an edge of  $F^*$ , the vertices  $V(H) \cup V(H')$  have at least  $d^{2k+2}2^{-2k-7}n$  common neighbours in  $V$ . As  $|V \setminus V'| \leq 2\delta n \leq d^{2k+2}2^{-2k-8}n$ ,  $e$  has a neighbour in  $X$  with respect to  $\mathcal{F}_D$ . Since this is true for any edge  $e$  of  $D$ , we conclude that  $|\bigcup_{v \in X} N_{\mathcal{F}_D}(v)| = t - 1 \geq |X|$ .

Secondly suppose that  $|X| < \delta n$ . Here it suffices to show that for any  $v \in V$  we have  $|N_{\mathcal{F}_D}(v)| \geq \delta n$ . Fix any  $v \in V$  and let  $\ell = d^{k+1}2^{-k-6}n$ . As the pair  $(\mathcal{H}, V)$  is  $(\varepsilon', d^{k+1}2^{-k-4})$ -super-regular with respect to  $\mathcal{T}_G(\mathcal{H}, V)$ , we have that  $v$  has degree at least  $d^{k+1}2^{-k-4}m_0$  into  $\mathcal{H}$  with respect to  $\mathcal{T}_G(\mathcal{H}, V)$ . Consider any point during the first  $\ell$  steps of the algorithm and let  $H_1, \dots, H_r$  be the current path. We assume that  $H_r$  is not (yet) in  $A$  and denote by  $\mathcal{M}$  the history of the algorithm until this point. Now we will look into the next two steps of the algorithm and estimate the probability that two vertices  $H_{r+1}$  and  $H_{r+2}$  are added to the path and the edge  $(H_{r+1}, H_{r+2})$  is in  $\mathcal{F}_v$ .

The first part is equivalent to  $H_r \notin B$  and  $H_{r+1} \notin B$  and with  $\mathbb{P}[H_r \in B] \leq (1 - \frac{\rho}{2})^{\delta n} \leq \exp(-\frac{\rho}{2}\delta n) \leq \varepsilon'$  we get that  $\mathbb{P}[H_r, H_{r+1} \notin B] \geq (1 - \varepsilon')^2$ . Next, we want to bound the number of valid choices for  $H_{r+1}$  in  $\mathcal{H}'$  that are neighbours of  $v$  in  $\mathcal{T}_G(\mathcal{H}, V)$ . From the neighbours of  $v$  in  $\mathcal{T}_G(\mathcal{H}, V)$ , we have to exclude those  $H'$  such that  $H_r H'$  is not an edge of  $F^*$ , and those  $H'$  that are "dead-ends": in the first case their number is at most  $2\varepsilon' m_0$  by Claim 3.5.1, in the second case their number is at most  $|B| < \varepsilon' n$ . Therefore there are at least  $d^{k+1}2^{-k-4}m_0 - 2\varepsilon' m_0 - \varepsilon' n - \ell \geq d^{k+1}2^{-k-5}m_0$  valid choices for  $H_{r+1}$  in  $\mathcal{H}'$  that are neighbours of  $v$  in  $\mathcal{T}_G(\mathcal{H}, V)$ . Repeating the same argument in the next step of the algorithm, there are at least  $d^{k+1}2^{-k-5}m_0$  valid choices for  $H_{r+2}$  in  $\mathcal{H}'$  that are neighbours

of  $v$  in  $\mathcal{T}_G(\mathcal{H}, V)$ . In particular for such choices of  $H_{r+1}$  and  $H_{r+2}$ , the edge  $(H_{r+1}, H_{r+2})$  is in  $\mathcal{F}_v$ .

We have that  $H_r$  (resp.  $H_{r+1}$ ) is not in  $A$  and we condition on the event that it is also not in  $B$ . Since we revealed the edges of  $F^*$  in each step separately the vertex  $H_{r+1}$  (resp.  $H_{r+2}$ ) is then chosen uniformly at random from the at most  $|\mathcal{H}| = m_0$  available possibilities. Therefore,

$$\mathbb{P}[(H_{r+1}, H_{r+2}) \in \mathcal{F}_v | \mathcal{M} \wedge H_r, H_{r+1} \notin B] \geq \frac{(d^{k+1}2^{-k-5}m_0)^2}{m_0^2} \geq d^{2k+2}2^{-2k-10}$$

and together with the bound on  $\mathbb{P}[H_r, H_{r+1} \notin B]$  we get

$$\mathbb{P}[(H_{r+1}, H_{r+2}) \in \mathcal{F}_v | \mathcal{M}] \geq (1 - \varepsilon')^2 d^{2k+2}2^{-2k-10} \geq d^{2k+2}2^{-2k-11}.$$

Crucially, this lower bound holds independently of the history  $\mathcal{M}$ . As among the first  $\ell$  steps we can have at most  $\varepsilon'n$  many steps in which  $H_r \in A$ , this process dominates a binomial distribution with parameters  $\ell - \varepsilon'n$  and  $d^{2k+2}2^{-2k-11}n$ . Therefore, even though the events are not mutually independent, we can use Chernoff's inequality (Lemma 2.4.1) to infer that with probability at least  $1 - n^{-2}$  at least  $d^{3k+3}2^{-3k-18}n$  of these edges are in  $\mathcal{F}_v$ . Some of these edges might not appear in the final path  $D$ , because of the "dead-ends" and the backtracking of the algorithm, but their number is at most  $\varepsilon'n$ . Thus we get that  $|N_{\mathcal{F}_D}(v)| \geq d^{3k+3}2^{-3k-18}n - \varepsilon'n \geq \delta n$  with probability at least  $1 - n^{-2}$ . By applying the union bound over all  $v \in V$ , we obtain that a.s.  $|N_{\mathcal{F}_D}(v)| \geq \delta n$  for all  $v \in V$ , as desired.

Finally, assume that  $\delta n \leq |X| \leq (1 - d^{k+2}2^{-2k-8})n$ . Here it suffices to show that, for any  $X \subseteq V$  with  $|X| = \delta n$ , we have  $|\bigcup_{v \in X} N_{\mathcal{F}_D}(v)| \geq (1 - d^{k+2}2^{-2k-8})n$ . We use a similar argument as above, but this time we need to give more precise estimates. Consider any step of the algorithm where the current path is  $H_1, \dots, H_r$  for some  $r < (1 - \frac{4k}{3k-1}\delta)n + 1$ , assume that  $H_r \notin A$ , and denote by  $\mathcal{M}$  the history of the algorithm until this point. We want to bound the number of valid choices for  $H_{r+1}$  in  $\mathcal{H}'$  that are neighbours of some  $v \in X$  in  $\mathcal{T}_G(\mathcal{H}, V)$ . With Claim 3.5.1, there are at least  $m_0 - 2\varepsilon'm_0 - \varepsilon'n - r$  choices for  $H_{r+1} \in \mathcal{H}'$  such that  $H_r H_{r+1}$  is an edge of  $F^*$  and  $H_{r+1}$  is not a "dead-end". Using Claim 3.5.2, for at least  $m_0 - r - 4\varepsilon'n$  of these choices,  $H_{r+1}$  has degree at least  $(1 - 4\varepsilon')m_0$  in  $F_X^*$ . Then there are at least  $m_0 - r - 4\varepsilon'm_0 - \varepsilon'n \geq m_0 - r - 5\varepsilon'n$  choices for  $H_{r+2}$ , such that  $(H_{r+1}, H_{r+2})$  is in  $\bigcup_{v \in X} \mathcal{F}_v$ .

On the other hand, there are at most  $(m_0 - r)$  choices for each of  $H_{r+1}$  and  $H_{r+2}$  and as above we have that at least one neighbour of  $H_r$  or  $H_{r+1}$  appears with probability at least  $1 - 2\varepsilon'$ . Using that  $m_0 - r \geq \delta n/3$  we get

$$\mathbb{P}\left[(H_{r+1}, H_{r+2}) \in \bigcup_{v \in X} \mathcal{F}_v \mid \mathcal{M} \wedge H_r \notin A\right] \geq \frac{(1 - 2\varepsilon')^2 (m_0 - r - 5\varepsilon'n)^2}{(m_0 - r)^2} \geq 1 - 20 \frac{\varepsilon'}{\delta}.$$

Again, as the lower bound holds independently of the history  $\mathcal{M}$  and as there are at most  $\varepsilon'n$  steps with  $H_r \in A$ , this process dominates a binomial distribution with parameters  $t - \varepsilon'n$  and  $1 - 20\frac{\varepsilon'}{\delta}$ . Therefore, the number  $Y$  of these edges that are in  $\bigcup_{v \in X} \mathcal{F}_v$  is in expectation at least  $(t - \varepsilon'n)(1 - 20\frac{\varepsilon'}{\delta}) \geq (1 - \delta)t$  and we get from the more precise version of Chernoff's inequality (Lemma 2.4.1) that

$$\begin{aligned} \mathbb{P} \left[ Y < (1 - 2\delta)t \leq \mathbb{E}[Y] - \left( \delta - 20\frac{\varepsilon'}{\delta} \right) (t - \varepsilon'n) \right] \\ \leq \exp \left( -D \left( (1 - \delta) \parallel 1 - 20\frac{\varepsilon'}{\delta} \right) (t - \varepsilon'n) \right) \\ \leq \exp \left( -\delta \left( \log \left( \frac{\delta^2}{20\varepsilon'} \right) - 2 \right) (t - \varepsilon'n) \right) \\ \leq \exp \left( -\delta \log \left( \frac{1}{\varepsilon'} \right) \frac{1}{2} t \right). \end{aligned}$$

There are at most  $\binom{n}{\delta n} \leq \left(\frac{e}{\delta}\right)^{\delta n} \leq \exp(\delta \log(\frac{1}{\delta})n) \leq \exp(\delta \log(\frac{1}{\varepsilon'})\frac{1}{4}t)$  choices for  $X$  and, thus, with the union bound over all these choices, we obtain that a.a.s. at least  $(1 - 2\delta)t$  of the edges are in  $\bigcup_{v \in X} \mathcal{F}_v$  for every  $X \subseteq V$  with  $|X| = \delta n$ . At most  $\varepsilon'n$  of these edges do not belong to  $D$  and putting this together we a.a.s. have

$$\left| \bigcup_{x \in X} N_{\mathcal{F}_D}(x) \right| \geq (1 - 2\delta)t - \varepsilon'n \geq (1 - 4\delta)n \geq (1 - d^{k+2}2^{-2k-8})n$$

for any  $X \subseteq V$  with  $|X| = \delta n$ , as wanted.  $\square$

Let  $D$  be the directed path in  $F^*$  given by Claim 3.5.3 and assume that the assertion of Claim 3.5.4 holds. We denote the first vertex of  $D$  by  $H'_x$  and the last by  $H'_y$ . Before dealing with the next step, we summarise what we have so far. We have several copies of  $H^{(k)}$ :  $H_x, H_y$  and those in  $\mathcal{H}$ . The vertices  $x$  and  $x'$  (resp.  $y$  and  $y'$ ) are adjacent in  $G$  to all vertices of  $H_x$  (resp.  $H_y$ ), and thus  $(x, x')$  and  $(y, y')$  can be end-tuples of the square of a Hamilton path we want to construct. Moreover, we have an ordering (given by the directed path  $D$ ) of  $t$  copies of  $H^{(k)}$  in  $\mathcal{H}$ , such that if  $H = u_1, \dots, u_k$  and  $H = u'_1, \dots, u'_k$  appear consecutively, then  $u_k u'_1$  is an edge of the random graph, and all their vertices  $u_1, \dots, u_k, u'_1, \dots, u'_k$  have many common neighbours in  $V$  in the graph  $G$ .

**Covering the left-over vertices from  $U_1 \cup \dots \cup U_k$ .** Let  $\mathcal{H}'$  be the copies of  $H^{(k)}$  in  $\mathcal{H}$  not used for the path  $D$  and observe that  $|\mathcal{H}'| = |\mathcal{H} \setminus V(D)| = m_0 - t = s - 2$ , where the last equality follows from (3.5.1). Further observe that the number of vertices in  $U_i$  not in any copy of  $H^{(k)}$  in  $\mathcal{H}$  is  $|U_i| - 2 - |\mathcal{H}| = (1 - \delta)n - m_0 = \frac{s}{k}$ , where the last equality follows from (3.5.2). Therefore we have exactly  $s$  vertices in total in  $U_1 \cup \dots \cup U_k$  to absorb; let  $Z$  be the set of these vertices. We want to cover the  $s$  vertices in  $Z$  with the square of two paths connecting  $H_x$  to  $H'_x$  and  $H'_y$  to  $H_y$  respectively, while using all copies of  $H^{(k)}$  in  $\mathcal{H}'$  and exactly  $4s$  vertices from  $V$  (c.f. Figure 3.8).

We start from connecting  $H_x$  to  $H'_x$ , while absorbing two vertices of  $Z$ . We pick  $H' \in \mathcal{H}'$  and  $z_x, z'_x \in Z$  such that the vertices in  $H_x \cup \{z_x\} \cup H'$  and  $H' \cup \{z'_x\} \cup H'_x$  each have at least  $2\delta n$  common neighbours in  $V$ . This is possible by using Claim 3.5.1 and the regularity property of  $G$ . Then by Lemma 2.5.1, a.a.s. the random graph  $G_2$  has a path on four vertices within each of these two sets of  $2\delta n$  vertices, that gives the desired connection (c.f. Figure 3.8).

Now we connect  $H'_y$  to  $H_y$ , while absorbing the other  $s - 2$  vertices of  $Z \setminus \{z_x, z'_x\}$ . Let  $H'_1 = H_y, H'_{s-1} = H'_y$  and  $H'_2, \dots, H'_{s-2}$  be a labelling of the remaining  $s - 3$  copies of  $H^{(k)}$  in  $\mathcal{H}' \setminus \{H'\}$  such that for  $j = 1, \dots, s - 2$  we have that all vertices in  $V(H_j) \cup V(H_{j+1})$  have at least  $d^{2k+2}2^{-2k-7}n$  common neighbours in  $V$  in  $G$ . This is possible by Dirac's Theorem and because, by Claim 3.5.1, for each  $H \in \mathcal{H}' \cup \{H_y, H'_y\}$  all but  $6\frac{\varepsilon'}{\delta}s$  choices  $H' \in \mathcal{H}'$  are such that the vertices  $V(H) \cup V(H')$  have  $d^{2k+2}2^{-2k-7}n$  common neighbours.

Next, we want to find a labelling  $z_1, \dots, z_{s-2}$  of the vertices from  $Z' = Z \setminus \{z_x, z'_x\}$  such that for  $j = 1, \dots, s - 2$  the vertices  $V(H_j) \cup \{z_j\} \cup V(H_{j+1})$  have at least  $2\delta n$  common neighbours in  $V$ . This again follows easily from Hall's condition for perfect matchings and because, by Claim 3.5.1, for each  $j = 1, \dots, s - 2$  all but  $6\varepsilon's/\delta$  choices  $z \in Z$  are such that  $V(H_j) \cup V(H_{j+1}) \cup \{z\}$  have  $2\delta n$  common neighbours and, similarly, vice versa. Then by Lemma 2.5.1, a.a.s. we can greedily choose a path on four vertices in the common neighbourhood of the vertices from  $V(H_j) \cup V(H_{j+1}) \cup \{z_j\}$  in  $V$  for  $j = 1, \dots, s - 2$ , with all the edges coming from the random graph  $G_2$ . This again gives the desired connection (c.f. Figure 3.8).

This completes the square of two paths from  $H_x$  to  $H'_x$  and from  $H'_y$  to  $H_y$ . These two cover exactly  $4s$  vertices of  $V$ . Therefore, there are precisely  $|V| - 4 - 4s = n - 4s = t - 1$  vertices of  $V \setminus \{x, x', y, y'\}$  not yet covered by the square of a path, where we used (3.5.3); we let  $V'$  be the set of such vertices. Observe that  $|V'| = t - 1 = |E(D)|$ .

**Finishing the square of a path.** We finish the proof by constructing the square of a path with  $H'_x$  and  $H'_y$  at the ends using precisely the vertices of  $V'$  and the copies of  $H^{(k)}$  that are vertices of  $V(D)$ . For this we use that by Claim 3.5.4 there is a perfect matching in  $\mathcal{F}_D[V', E(D)]$ . For  $i = 1, \dots, t - 1$ , let  $v_i$  be the vertex of  $V$  matched to the edge  $(H_i, H_{i+1}) \in E(D)$  in  $\mathcal{F}_D$ . With  $H_i = u_1, \dots, u_k$  and  $H_{i+1} = u'_1, \dots, u'_k$ , we then have that  $v_i$  is incident to  $u_{k-1}u_k, u'_1, u'_2$  by definition of  $\mathcal{F}_D$ . This completes the construction of the square of the path with  $H'_x$  to  $H'_y$  at the ends. By adding the two connections found above from  $H_x$  to  $H'_x$  and from  $H'_y$  to  $H_y$  and the initial tuples  $(x, x')$  and  $(y, y')$ , we get the square of a Hamilton path with end-tuples  $(x, x')$  and  $(y, y')$  as desired (c.f. Figure 3.6). This finishes the proof of the lemma.  $\square$

We end this section by giving the proof of Lemma 3.1.4, that follows from Lemma 3.1.3, once we split appropriately the super-regular regular pair  $(U, V)$  into two copies of super-regular  $K_{1,2}$ , both suitable for an application of Lemma 3.1.3 with  $k = 2$ .

*Proof of Lemma 3.1.4.* Let  $0 < d < 1$ , choose  $\delta'$  with  $0 < \delta' \leq d/8$  and apply Lemma 3.1.3 with  $k = 2$ ,  $\delta'$ , and  $d/8$  to obtain  $\delta_0, \delta, \varepsilon'$  with  $\delta' \geq \delta_0 > 2\delta > \varepsilon' > 0$  and  $C' > 0$ . Then let  $0 < \varepsilon \leq \varepsilon'/8$ ,  $C \geq 4C'$ , and  $p \geq Cn^{-1}$ . Next let  $U$  and  $V$  be vertex-sets of size  $|V| = n$  and  $3n/4 \leq |U| = m \leq n$  and assume that  $(U, V)$  is an  $(\varepsilon, d)$ -super-regular pair. Let  $(x, x')$  and  $(y, y')$  be tuples from  $V$  and  $U$ , respectively, such that they have  $\frac{1}{2}d^2n$  common neighbours into the other set. We will reveal  $G(V, p)$  and  $G(U, p)$  both in two rounds as  $G_1, G_3 \sim G(V, \frac{1}{2}p)$ , and  $G_2, G_4 \sim G(U, \frac{1}{2}p)$ .

We partition  $V$  into  $V_1, U_2, W_2$  and  $U$  into  $V_2, U_1, W_1$  such that for  $i = 1, 2$  the pairs  $(U_i, V_i)$  and  $(W_i, V_i)$  are  $(\varepsilon', \frac{1}{8}d)$ -super-regular pairs and  $(1 - \delta_0)|V_i| \leq |U_i| = |W_i| \leq (1 - \delta)|V_i|$ . Additionally, we require that  $(x, x')$  is in  $V_1$  and that  $(y, y')$  is in  $V_2$  and that they have at least  $\frac{1}{2}(\frac{d}{8})^2n$  common neighbours in  $U_1, W_1$  and in  $U_2, W_2$ , respectively. To obtain this we split the sets according to the following random distribution. We put any vertex of  $V$  into each of  $U_2$  and  $W_2$  with probability  $q_1$  and into  $V_1$  with probability  $1 - 2q_1$ . Similarly, we put any vertex of  $U$  into each of  $U_1$  and  $W_1$  with probability  $q_2$  and into  $V_2$  with probability  $1 - 2q_2$ . We choose  $q_1$  and  $q_2$  such that the expected sizes satisfy for  $i = 1, 2$

$$\mathbb{E}[|U_i|] = \mathbb{E}[|W_i|] = \left(1 - \frac{\delta_0 + \delta}{2}\right) \mathbb{E}[|V_i|].$$

This is possible since such conditions give a linear system of two equations in two unknowns  $q_1$  and  $q_2$ , and, as  $3n/4 \leq m \leq n$ , the solution satisfies  $1/7 \leq q_1, q_2 \leq 3/7$ . Then by Chernoff's inequality (Lemma 2.4.1) and with  $n$  large enough there exists a partition such that for  $i = 1, 2$  we have that  $|W_i|$ ,  $|U_i|$ , and  $|V_i|$  are all within  $\pm n^{2/3}$  of their expectation and the minimum degree within both pairs  $(U_i, V_i)$  and  $(W_i, V_i)$  is at least a  $d/4$ -fraction of the other set. For  $i = 1, 2$  we redistribute  $o(n)$  vertices between  $U_i$  and  $W_i$  and move at most one vertex from or to  $V_i$  to obtain

$$(1 - \delta_0)|V_i| \leq |U_i| = |W_i| \leq (1 - \delta)|V_i|$$

with minimum degree within both pairs  $(U_i, V_i)$  and  $(W_i, V_i)$  at least a  $d/8$ -fraction of the other set. Moreover, for  $i = 1, 2$ , we can ensure that with  $n_i = |V_i| - 4$  we have  $n_i - |U_i| \equiv -1 \pmod{5}$ .

From this we get that for  $i = 1, 2$  the pairs  $(U_i, V_i)$  and  $(W_i, V_i)$  are  $(\varepsilon', \frac{1}{8}d)$ -super-regular. With  $G_1$  and  $G_2$  we reveal random edges within  $V_1$  and  $V_2$  with probability  $p/2$  to find tuples  $(z, z')$  in  $V_1$  and  $(w, w')$  in  $V_2$  such that together they give a copy of  $K_4$  and  $(z, z')$  and  $(w, w')$  have at least  $\frac{1}{2}(\frac{d}{8})^2n$  common neighbours in  $U_1, W_1$  and in  $U_2, W_2$ , respectively. Then we use Lemma 3.1.3 and  $G_3, G_4$  with  $Cn^{-1} \geq C' \min\{|V_1|, |V_2|\}^{-1}$  to a.a.s. find the square of a Hamilton path on  $V_i, U_i, W_i$  for  $i = 1, 2$  with end-tuples  $(x, x')$ ,  $(z, z')$  and  $(y, y')$ ,  $(w, w')$ , respectively. Together with the edges between  $(z, z')$  and  $(w, w')$  this gives the square of a Hamilton path covering  $U$  and  $V$  with end-tuples  $(x, x')$  and  $(y, y')$ .  $\square$

### 3.6 Proof of the embedding Lemma 3.1.1

The proof Lemma 3.1.1 for  $k \geq 3$  is a not too difficult application of Janson's inequality, while the case  $k = 2$  requires more ad-hoc arguments. Therefore, though the two proof strategies share some ideas, we treat the case  $k = 2$  separately.

#### 3.6.1 Case $k = 2$

We want to show that if  $G$  is an  $n$ -vertex graph with minimum degree  $\delta(G) \geq m$  and maximum degree  $\Delta(G) \leq n/(64t)$ , then a.a.s. the perturbed graph  $G \cup G(n, p)$  contains  $tm+t$  pairwise vertex-disjoint triangles, provided  $p \geq Cn^{-1} \log n$ . We split the proof in three ranges for the value of  $m$ :  $0 \leq m \leq (\log n)^3$ ,  $(\log n)^3 \leq m \leq \sqrt{n}$ , and  $\sqrt{n} \leq m \leq n/(64t)$ . If  $0 \leq m \leq (\log n)^3$  a.a.s.  $tm+t$  pairwise vertex-disjoint triangles already exist in  $G(n, p)$  (see [68, Theorem 3.29]). If  $(\log n)^3 \leq m \leq \sqrt{n}$  we will find many large enough vertex-disjoint stars in  $G$  (see Lemma 3.6.3) and a.a.s. at least  $tm+t$  of them will be completed to triangles using edges of  $G(n, p)$  (see Proposition 3.6.1). However if  $m > \sqrt{n}$  we cannot hope to find  $tm+t$  large enough vertex-disjoint stars and instead we will apply a greedy strategy using that a.a.s. every vertex has an edge in its neighbourhood (see Proposition 3.6.2).

**Proposition 3.6.1.** *For any integer  $t \geq 1$  and real number  $0 < \gamma < 1/2$ , there exists  $C > 0$  such that for any  $(\log n)^3 \leq m \leq \sqrt{n}$  and any  $n$ -vertex graph  $G$  with maximum degree  $\Delta(G) \leq \gamma n$  and minimum degree  $\delta(G) \geq m$  the following holds. With  $p \geq Cn^{-1}$  there are a.a.s. at least  $tm+t$  pairwise vertex-disjoint triangles in  $G \cup G(n, p)$ .*

**Proposition 3.6.2.** *For any integer  $t \geq 1$ , there exists  $C > 0$  such that for any  $\sqrt{n} \leq m \leq n/(64t)$  and any  $n$ -vertex graph  $G$  with maximum degree  $\Delta(G) \leq n/(64t)$  and minimum degree  $\delta(G) \geq m$  the following holds. With  $p \geq Cn^{-1} \log n$  there are a.a.s. at least  $tm+t$  pairwise vertex-disjoint triangles in  $G \cup G(n, p)$ .*

With this at hand, Lemma 3.1.1 is now obvious, and it remains to prove Proposition 3.6.1 and Proposition 3.6.2.

For Proposition 3.6.1, which deals with the cases  $(\log n)^3 \leq m \leq \sqrt{n}$ , we first need to find many large enough vertex-disjoint stars in  $G$ . These can be found deterministically with Lemma 3.6.3 below and afterwards we will show that a.a.s. at least  $tm+t$  of them can be completed to triangles with the help of  $G(n, p)$ .

Recall that, give any integer  $g \geq 2$ , the *star* on  $g+1$  vertices is the graph with one vertex of degree  $g$  (this vertex is called the *centre*) and the other vertices of degree one (these vertices are called *leaves*). Given a star  $K$ , we denote the number of its leaves by  $g_K$ . Moreover, given a family of vertex-disjoint stars  $\mathcal{K}$ , we denote the set of all their centre vertices by  $\mathcal{K}_C$  and the set of all their leaf vertices by  $\mathcal{K}_L$ .



**Lemma 3.6.3.** *For every  $0 < \gamma < 1/2$  and integer  $s$  there exists  $\varepsilon > 0$  such that for  $n$  large enough and any  $m$  with  $2/\varepsilon \leq m \leq \sqrt{n}$  the following holds. In every  $n$ -vertex graph  $G$  with minimum degree  $\delta(G) \geq m$  and maximum degree  $\Delta(G) \leq \gamma n$  there exists a family  $\mathcal{K}$  of vertex-disjoint stars in  $G$  such that every  $K \in \mathcal{K}$  has  $g_K$  leaves with  $\varepsilon m \leq g_K \leq \varepsilon \sqrt{n}$  and*

$$\sum_{K \in \mathcal{K}} g_K^2 \geq s\varepsilon^2 nm.$$

*Proof of Lemma 3.6.3.* Given  $0 < \gamma < 1/2$  and an integer  $s$  we let  $\varepsilon > 0$  such that  $\varepsilon \leq 1/(6s)$  and  $\varepsilon < 1/2 - s\gamma$ . Moreover, we let  $n$  be large enough for our calculations and, for simplicity, assume that  $\varepsilon \sqrt{n}$  is an integer. Then let  $2/\varepsilon \leq m \leq \sqrt{n}$  and  $G$  be an  $n$ -vertex graph on vertex set  $V$  with  $\delta(G) \geq m$  and  $\Delta(G) \leq \gamma n$ .

Let  $\mathcal{K}$  be a family of vertex-disjoint stars in  $G$  with  $\varepsilon m \leq g_K \leq \varepsilon \sqrt{n}$  for all  $K \in \mathcal{K}$ , that maximizes the sum

$$\sum_{K \in \mathcal{K}} g_K^2 \tag{3.6.1}$$

among all such families. Note that each star in  $\mathcal{K}$  has at least 2 leaves because  $g_K \geq \varepsilon m$  and  $m \geq 2/\varepsilon$ .

If the sum in (3.6.1) is bigger than  $s\varepsilon^2 nm$  we are done. So we assume the family  $\mathcal{K}$  satisfies

$$\sum_{K \in \mathcal{K}} g_K^2 < s\varepsilon^2 nm. \tag{3.6.2}$$

We are going to prove that then there exists a vertex of degree larger than  $\gamma n$ , contradicting our assumption on the maximum degree.

For this we split  $\mathcal{K}$  into two subfamilies

$$\mathcal{M} = \{K \in \mathcal{K} : \varepsilon m \leq g_K < \varepsilon \sqrt{n}\} \quad \text{and} \quad \mathcal{H} = \{K \in \mathcal{K} : g_K = \varepsilon \sqrt{n}\}$$

and we let  $R$  be the set of vertices not covered by the stars in  $\mathcal{K}$ , that is  $R = V(G) \setminus (\mathcal{H}_C \cup \mathcal{H}_L \cup \mathcal{M}_C \cup \mathcal{M}_L)$ , where  $\mathcal{H}_C, \mathcal{H}_L, \mathcal{M}_C,$  and  $\mathcal{M}_L$  are obtained from  $\mathcal{M}$  and  $\mathcal{H}$  as defined above.

For all stars  $K \in \mathcal{H}$  we have  $g_K^2 = \varepsilon^2 n$ . From (3.6.2) we get that the subfamily  $\mathcal{H}$  contains at most  $sm$  stars and hence

$$|\mathcal{H}_C| \leq sm \leq s\sqrt{n} \quad \text{and} \quad |\mathcal{H}_L| = |\mathcal{H}_C| \varepsilon \sqrt{n} \leq sm \varepsilon \sqrt{n} \leq s\varepsilon n, \tag{3.6.3}$$

because  $m \leq \sqrt{n}$ .

As each star in  $\mathcal{M}$  has at least  $\varepsilon m$  leaves we have  $\sum_{K \in \mathcal{M}} g_K \geq |\mathcal{M}| \varepsilon m$ . Using the

Cauchy-Schwarz inequality, we then get

$$\left( \sum_{K \in \mathcal{M}} g_K \right)^2 \leq \left( \sum_{K \in \mathcal{M}} g_K^2 \right) |\mathcal{M}| \stackrel{(3.6.2)}{\leq} s\varepsilon^2 nm |\mathcal{M}| \leq \left( \sum_{K \in \mathcal{M}} g_K \right) s\varepsilon n,$$

which implies

$$|\mathcal{M}_L| = \sum_{K \in \mathcal{M}} g_K \leq s\varepsilon n. \quad (3.6.4)$$

Therefore,  $|\mathcal{M}_C| \leq |\mathcal{M}_L|/2 \leq s\varepsilon n/2$  since each star has at least 2 leaves.

These bounds on  $\mathcal{M}_L$  and  $\mathcal{M}_C$  together with (3.6.3) immediately imply that  $|R| \geq (1 - 3s\varepsilon)n \geq n/2$ . We are going to show that there are many edges between  $R$  and  $\mathcal{H}_C$  and from that we derive the existence of a high degree vertex, giving the desired contradiction.

A vertex in  $R$  cannot have at least  $\varepsilon m$  neighbours inside  $R$ , because otherwise we could create a new star and increase the sum in (3.6.1). Therefore,  $e(R) < |R|\varepsilon m/2$ . We also have  $e(R, \mathcal{M}_C) = 0$  since otherwise we could add an edge to one of the existing stars in  $\mathcal{K}$  increasing the sum in (3.6.1) (recall that stars in  $\mathcal{M}$  have less than  $\varepsilon\sqrt{n}$  leaves).

Given a leaf  $v \in \mathcal{M}_L$  that belongs to a star  $K$  with  $g_K$  leaves, we must have  $\deg(v, R) < g_K + 1$ . Otherwise, we could take  $V' \subset N_R(v)$  of size  $|V'| = g_K + 1 \leq \varepsilon\sqrt{n}$  and create a new family of vertex-disjoint stars, given by  $\mathcal{K} \setminus \{K\}$  and the star on  $v \cup V'$ , to increase the sum in (3.6.1). Therefore,

$$e(R, \mathcal{M}_L) \leq \sum_{K \in \mathcal{M}} g_K(g_K + 1) \leq \sum_{K \in \mathcal{K}} g_K^2 + \sum_{K \in \mathcal{M}} g_K \stackrel{(3.6.2), (3.6.4)}{\leq} s\varepsilon^2 nm + s\varepsilon n.$$

Similarly, given  $v \in \mathcal{H}_L$ , we must have  $\deg(v, R) < \varepsilon\sqrt{n}$ . Otherwise, we could take  $V' \subset N_R(v)$  of size  $|V'| = \varepsilon\sqrt{n}$  and create a new family of vertex-disjoint stars, given by  $\mathcal{K} \setminus \{K\}$ , the star  $K \setminus \{v\}$ , and the star on  $v \cup V'$ , to increase the sum in (3.6.1). Therefore,  $e(R, \mathcal{H}_L) \leq |\mathcal{H}_L|\varepsilon\sqrt{n} \leq s\varepsilon^2 nm$  by (3.6.3).

On the other hand,  $\delta(G) \geq m$  implies  $e(R, V) \geq m|R|$ , where the edges inside of  $R$  are counted twice. Then we can lower bound the number of edges between  $R$  and  $\mathcal{H}_C$  by

$$\begin{aligned} e(R, \mathcal{H}_C) &\geq m|R| - 2e(R) - e(R, \mathcal{M}_C) - e(R, \mathcal{M}_L) - e(R, \mathcal{H}_L) \\ &\geq m|R| - \varepsilon m|R| - 0 - 2s\varepsilon^2 nm - s\varepsilon n \\ &\geq m|R| - \varepsilon m|R| - 4s\varepsilon^2 m|R| - s\varepsilon^2 m|R| \\ &\geq (1 - \varepsilon - 5s\varepsilon^2)|R|m \geq |R|(1 - 2\varepsilon)m. \end{aligned}$$

where we used the bounds on  $e(R)$ ,  $e(R, \mathcal{M}_C)$ ,  $e(R, \mathcal{M}_L)$ , and  $e(R, \mathcal{H}_L)$  we found above, together with  $|R| \geq n/2$ ,  $\varepsilon m \geq 2$  and the choice of  $\varepsilon < 1/(6s)$ . In particular, as  $|\mathcal{H}_C| \leq sm$

and using  $\varepsilon < 1/2 - s\gamma$ , and  $|R| > n/2$ , there exists a vertex  $v \in \mathcal{H}_C$  of degree

$$\deg(v) \geq \deg(v, R) \geq \frac{|R|(1-2\varepsilon)m}{sm} \geq |R|2\gamma > \gamma n.$$

This contradicts the maximum degree of  $G$ .  $\square$

*Proof of Proposition 3.6.1.* Let  $n$  be sufficiently large for the following arguments. Let  $t$  be an integer and  $0 < \gamma < 1/2$ . Let  $G$  be an  $n$ -vertex graph with maximum degree  $\Delta(G) \leq \gamma n$  and minimum degree  $\delta(G) \geq m$ . With  $(\log n)^3 \leq m \leq \sqrt{n}$ , we first find many vertex-disjoint stars in  $G$  and then complete at least  $m$  of them to triangles with the help of  $G(n, p)$ . We apply Lemma 3.6.3 with  $\gamma$  and  $s = 8t + 4$  to get  $0 < \varepsilon < 1/2$  and, as  $n$  is large enough and  $m \geq 2/\varepsilon$ , we get a family  $\mathcal{K}$  of vertex-disjoint stars on  $V(G)$  such that  $\varepsilon m \leq g_K \leq \varepsilon\sqrt{n}$  for  $K \in \mathcal{K}$  and  $\sum_{K \in \mathcal{K}} g_K^2 \geq (8t + 4)\varepsilon^2 nm$ .

As we have stars of different sizes, we split  $\mathcal{K}$  into  $t = \lceil \log(\sqrt{n}/m)/\log 2 \rceil + 1$  subfamilies

$$\mathcal{K}_i = \{K \in \mathcal{K} : 2^{i-1}\varepsilon m \leq g_K < 2^i\varepsilon m\}, \quad 1 \leq i \leq t,$$

and set  $k_i = |\mathcal{K}_i|$ .

By deleting leaves, we may assume that all stars in  $\mathcal{K}_i$  have exactly  $\lceil 2^{i-1}\varepsilon m \rceil$  leaves. Denote by  $\mathcal{I}$  the set of indices  $i \in [t]$  such that  $k_i (2^{i-1}\varepsilon m)^2 \geq \varepsilon^2 nm/t$ . Next we prove that  $\sum_{i \in \mathcal{I}} k_i (2^{i-1}\varepsilon m)^2 \geq \varepsilon^2 nm$ .

Observe first that  $\sum_{i \notin \mathcal{I}} k_i (2^{i-1}\varepsilon m)^2 \leq t(\varepsilon^2 nm/t) = \varepsilon^2 nm$ . It follows that

$$\begin{aligned} \sum_{i \in \mathcal{I}} k_i (2^{i-1}\varepsilon m)^2 &= \frac{1}{4} \sum_{i \in \mathcal{I}} k_i (2^i\varepsilon m)^2 = \frac{1}{4} \sum_{i=1}^t |\mathcal{K}_i| (2^i\varepsilon m)^2 - \sum_{i \notin \mathcal{I}} k_i (2^{i-1}\varepsilon m)^2 \\ &\geq \frac{1}{4} \sum_{i=1}^t \sum_{K \in \mathcal{K}_i} g_K^2 - \varepsilon^2 nm \geq (2t + 1)\varepsilon^2 nm - \varepsilon^2 nm \geq 2t\varepsilon^2 nm. \end{aligned}$$

Now we reveal random edges on  $V(G)$  with probability  $p \geq C/n$  where  $C$  is large enough for the Chernoff bounds and inequalities below. We shall show that this allows us to find at least  $m$  triangles a.a.s.. Indeed, for each  $i \in \mathcal{I}$ , we find many pairwise vertex-disjoint triangles in  $\mathcal{K}_i$  using random edges.

**Claim 3.6.4.** *For any  $i \in \mathcal{I}$ , after revealing edges of  $G(n, p)$  with  $p \geq C/n$  we have with probability at least  $1 - 1/n$  at least  $k_i(2^{i-1}m)^2/n$  pairwise vertex-disjoint triangles within  $(G \cup G(n, p))[\cup_{K \in \mathcal{K}_i} V(K)]$ .*

Having this claim and since  $|\mathcal{I}| \leq t = o(n)$ , with a union bound over  $i \in \mathcal{I}$ , there are a.a.s. at least

$$\sum_{i \in \mathcal{I}} \frac{k_i(2^{i-1}m)^2}{n} = \frac{1}{\varepsilon^2 n} \sum_{i \in \mathcal{I}} k_i (2^{i-1}\varepsilon m)^2 \geq \frac{2t\varepsilon^2 nm}{\varepsilon^2 n} = 2tm \geq tm + t$$

pairwise vertex-disjoint triangles in  $G \cup G(n, p)$ . It remains to prove Claim 3.6.4.

*Proof of Claim 3.6.4.* Fix  $i \in \mathcal{I}$  and let  $k = k_i$  and  $g = \lceil 2^{i-1} \varepsilon m \rceil$ . We reveal random edges with probability  $p$  within each set of leaves of the  $k$  stars in  $\mathcal{K}_i$ . We recall that these  $k$  sets are pairwise disjoint and each has size  $g$ . Let  $X_j$  be the indicator variable of the event that the  $j$ -th of these sets contains at least one edge for  $1 \leq j \leq k$ , and set  $X = \sum_{j=1}^k X_j$ . Then  $\mathbb{P}[X_j = 1] = 1 - (1-p)^{\binom{g}{2}}$  and  $\mathbb{E}[X] = k \left(1 - (1-p)^{\binom{g}{2}}\right)$ . We have that  $\mathbb{E}[X] \geq 2kg^2/(\varepsilon^2 n)$ . Indeed,

$$k \left(1 - (1-p)^{\binom{g}{2}}\right) \geq 2kg^2/(\varepsilon^2 n) \quad \Leftrightarrow \quad 1 - 2g^2/(\varepsilon^2 n) \geq \left(1 - \frac{C}{n}\right)^{\binom{g}{2}},$$

and the later holds for large enough  $C$  and  $n$  using the inequality  $1 - x \leq e^{-x} \leq 1 - \frac{x}{2}$  valid for  $x < 3/2$ .

From Chernoff's inequality (Lemma 2.4.1) and from the fact that  $kg^2/(\varepsilon^2 n) \geq m/t$  by the definition of  $\mathcal{I}$ , it follows that with probability at most

$$2 \exp\left(-\frac{1}{6} \frac{kg^2}{\varepsilon^2 n}\right) \leq 2 \exp\left(-\frac{1}{6} \frac{m}{t}\right) \leq \frac{1}{n}$$

there are less than  $kg^2/(\varepsilon^2 n)$  triangles, where the last inequality holds as  $t \leq \log n$ ,  $m \geq (\log n)^3$  and  $n$  is large enough.  $\square$

$\square$

*Proof of Proposition 3.6.2.* Let  $t \geq 1$  be an integer,  $\sqrt{n} \leq m \leq n/(64t)$ , and  $G$  be an  $n$ -vertex graph with maximum degree  $\Delta(G) \leq n/(64t)$  and minimum degree  $\delta(G) \geq m$ . We can greedily obtain a spanning bipartite subgraph  $G' \subseteq G$  of minimum degree  $\delta(G') \geq m/2$  by taking a partition of  $V(G)$  into sets  $A$  and  $B$  such that  $e_G(A, B)$  is maximised and letting  $G' = G[A, B]$ . Indeed, a vertex of degree less than  $m/2$  can be moved to the other class to increase  $e_G(A, B)$ . W.l.o.g. we assume  $|B| \geq n/2 \geq |A|$ . Moreover, we have  $|A| \geq 16tm$ , as otherwise with  $e(A, B) \geq nm/4$  there is a vertex of degree larger than  $n/(64t)$ , a contradiction.

**Claim 3.6.5.** For every  $A' \subseteq A$ ,  $B' \subseteq B$  with  $|A'| < 4tm$ ,  $|B'| \leq n/4$  we have  $e(A \setminus A', B \setminus B') \geq nm/16$ .

*Proof.* If  $e(A \setminus A', B \setminus B') < nm/16$ , it follows from  $e(A, B \setminus B') \geq |B \setminus B'|m/2 \geq nm/8$  that we have  $e(A', B \setminus B') \geq nm/16$ . Since  $|A'| < 4tm$ , there must be a vertex of degree at least  $n/(64t)$  in  $A'$ , a contradiction.  $\square$

From this claim it follows that there are many vertices of high degree in  $A \setminus A'$ .

**Claim 3.6.6.** *Suppose that  $A' \subseteq A$ ,  $B' \subseteq B$  with  $|A'| < 4tm$ ,  $|B'| \leq n/4$ . Let  $A^* = \{v \in A \setminus A' : \deg(v, B \setminus B') \geq m/(32t)\}$ . Then  $|A^*| \geq (4t - 1)m$ .*

*Proof.* We have  $|A^*| \frac{n}{64t} + |A| \frac{m}{32t} \geq e(A^*, B \setminus B') + e(A \setminus (A' \cup A^*), B \setminus B') = e(A \setminus A', B \setminus B') \geq \frac{nm}{16}$ , where the last inequality uses Claim 3.6.5. Since  $|A| \leq n/2$ , we get

$$|A^*| \geq \frac{nm/16 - nm/(64t)}{n/(64t)} = (4t - 1)m.$$

□

Let  $s = \lceil \frac{4tn}{m} \rceil$  and  $r = \lceil \frac{m^2}{2n} \rceil$ . We will now iteratively construct our  $tm + t$  triangles in  $r$  rounds of  $s$  triangles each. In each round we will reveal  $G(n, q)$  with  $q = \frac{C \log n}{m^2}$ , where  $C$  is large enough for the Chernoff bound below. For the start we set  $A' = B_0 = \emptyset$ .

Let  $i = 1, \dots, r$ , suppose that before the  $i$ -th round we have

$$|A'| = (i - 1)s \leq \left( \left\lceil \frac{m^2}{2n} \right\rceil - 1 \right) \left\lceil \frac{4tn}{m} \right\rceil \leq \frac{m^2}{2n} \left( \frac{4tn}{m} + 1 \right) < 4tm$$

and  $|B_0| = (i - 1)(2s) < 5tm$ , and note this is true for  $i = 1$ . In the  $i$ -th round we pick vertices  $v_1, \dots, v_s \in A \setminus A'$  and pairwise disjoint sets  $B_1, \dots, B_s \subseteq B \setminus B_0$ , each of size  $\lceil m/(32t) \rceil$ , such that  $B_j \subset N_{G'}(v_j)$  for each  $j = 1, \dots, s$ . We can do this greedily, where for  $j = 1, \dots, s$  we set  $B^j = B_0 \cup B_1 \cup \dots \cup B_{j-1}$  and apply Claim 3.6.6 to obtain a vertex  $v_j \in A \setminus A'$  together with a set  $B_j \subseteq B \setminus B^j$  of  $\lceil m/(32t) \rceil$  neighbours of  $v_j$ . We can do this as  $|A'| < 4tm$  and  $|B^j| \leq s \lceil m/(32t) \rceil + |B_0| \leq n/4$  as  $m \leq n/(64t)$ . Now we reveal additional edges at random with probability  $q$ . Then with probability at least  $1 - 1/n^2$  we have at least one edge in each set  $B_1, \dots, B_s$ . Indeed the probability that there is no edge in a set  $B_i$  is at most  $(1 - q)^{\binom{|B_i|}{2}} \leq \exp\left(-C \frac{\log n}{m} \left(\frac{m}{32t}\right)^2\right) \leq n^{-3}$  as  $C$  is large enough. Therefore the probability that there is a set without any edge is at most  $sn^{-3} \leq n^{-2}$  by a union bound. We fix an arbitrary edge from each  $B_i$  and together with  $v_1, \dots, v_s$  this gives us  $s$  triangles. We add the vertices  $v_1, \dots, v_s$  to  $A'$  and the vertices of the edges that we chose to  $B_0$ . Notice that  $|A'| = is$  and  $|B_0| = i(2s)$ , as required at the beginning of next round.

We can repeat the above  $r$  times because with  $m \geq \sqrt{n}$  we get  $rq \leq \frac{C \log n}{n} = p$ . By a union bound over the  $r = \lceil \frac{m^2}{2n} \rceil \leq n$  rounds, we get that we succeed a.a.s. and find  $ts \geq 2tm \geq tm + t$  triangles. □

### 3.6.2 Case $k \geq 3$

The case  $k \geq 3$  of Lemma 3.1.1 is much easier. The proof is still split into two parts. When  $m$  is small, we can find  $mt + t$  pairwise vertex-disjoint copies of the square of a path on  $k + 1$  vertices already in the random graph  $G(n, p)$ . When  $m$  is large, we need to use

the edges of  $G$  as well, and the reader will recognise a similar strategy to the one used for Proposition 3.6.2.

*Proof of Lemma 3.1.1* ( $k \geq 3$ ). Let  $k \geq 3$  and  $t \geq 1$  be integers and  $C$  be large enough for the inequalities below to hold. For convenience, we set  $\gamma = 1/(16kt)$ . Further let  $p \geq C(\log n)^{1/(2k-3)}n^{-(k-1)/(2k-3)}$ ,  $1 \leq m \leq \gamma n$ , and let  $G$  be an  $n$ -vertex graph with vertex set  $V$ , minimum degree  $\delta(G) \geq m$  and maximum degree  $\Delta(G) \leq \gamma n$ .

We distinguish two cases. If  $m \leq (\log n)^{2/(2k-3)}n^{(2k-4)/(2k-3)}$ , we only need Janson's inequality (Lemma 2.4.2) and we will greedily find  $tm + t$  copies of  $P_{k+1}^2$ , using only edges from the random graph  $G(n, p)$ . Let  $V' \subseteq V$  be the set of vertices used in this greedy construction. As long as we have not found  $tm + t$  copies of  $P_{k+1}^2$ , we have  $|V'| \leq (tm + t)(k + 1)$  and thus  $|V \setminus V'| \geq n/2$ . We let  $\{H_i\}_{i \in \mathcal{I}}$  be the family of copies of  $P_{k+1}^2$  with vertices in  $V \setminus V'$  and note  $|\mathcal{I}| \geq 2^{-k-2}n^{k+1}$ . Then, using the notation of Lemma 2.4.2, we observe that the expected number of these copies appearing as subgraphs of  $G(n, p)$  is

$$\begin{aligned} \mathbb{E}[X] &= |\mathcal{I}|p^{2k-1} \geq 2^{-k-2}n^{k+1}p^{2k-1} \\ &\geq C(\log n)^{(2k-1)/(2k-3)}n^{(2k-4)/(2k-3)} \geq 32ktm \log n. \end{aligned}$$

On the other hand, we have

$$\begin{aligned} \Delta[X] &= \sum_{H_i \sim H_j} p^{e(H_i)+e(H_j)-e(H_i \cap H_j)} \leq \sum_{r=2}^k O(n^{2(k+1)-r} p^{2(2k-1)-(2r-3)}) \\ &\leq \mathbb{E}^2[X] \sum_{r=2}^k O(p^{3-2r} n^{-r}) \leq \mathbb{E}^2[X] o(n^{-1}), \end{aligned}$$

where in the first inequality we split the sum according to the value of  $r = v(H_i \cap H_j)$  and used that then  $e(H_i \cap H_j) \leq 2r - 3$ . Then with Lemma 2.4.2 we get that the probability that there is no copy of  $P_{k+1}^2$  is bounded from above by  $\exp(-\mathbb{E}[X]/8) \leq n^{-4ktm}$ . We conclude with a union bound over the at most  $\binom{n}{(k+1)(tm+t)} \leq n^{3ktm}$  possible choices for  $V'$  that we can a.a.s. find  $tm + t$  copies of  $P_{k+1}^2$  in  $G(n, p)$ .

For  $m \geq (\log n)^{2/(2k-3)}n^{(2k-4)/(2k-3)}$  we need to use the edges of  $G$ . We will find copies of  $P_{k+1}^2$ , where all edges incident to one vertex come from  $G$  and the remaining edges come from  $G(n, p)$ , where we will need to distinguish between  $k = 3$  and  $k \geq 4$ . First, we greedily obtain a spanning bipartite subgraph  $G' \subseteq G$  of minimum degree  $\delta(G') \geq m/2$  by taking a partition of  $V(G)$  into sets  $A$  and  $B$  such that  $e_G(A, B)$  is maximised and letting  $G' = G[A, B]$ . Indeed, a vertex of degree less than  $m/2$  can be moved to the other class to increase  $e_G(A, B)$ . W.l.o.g. we assume  $|B| \geq n/2 \geq |A|$ . Moreover, we have  $|A| \geq m/(4\gamma)$ , as otherwise with  $e(A, B) \geq nm/4$  there is a vertex of degree at least  $\gamma n$ , a contradiction.

Then we observe that given any sets  $A' \subseteq A$  and  $B' \subseteq B$  such that  $|A'| \leq m/(16\gamma)$  and  $|B'| \leq n/4$ , we also have  $e(A \setminus A', B \setminus B') \geq nm/16$ . Otherwise, from  $e(A, B \setminus B') \geq |B \setminus B'|m/2 \geq nm/8$ , it would follow that  $e(A', B \setminus B') \geq nm/16$  and thus, since  $|A'| \leq m/(16\gamma)$ , we would have a vertex of degree at least  $\gamma n$  in  $A'$ , a contradiction to the maximum degree of  $G$ .

We will greedily find  $tm + t$  copies of  $P_{k+1}^2$  with one vertex in  $A$  and  $k$  vertices in  $B$ . Let  $A' \subseteq A$  and  $B' \subseteq B$  be the set of vertices used in this greedy construction. As long as we have not found  $tm + t$  copies of  $P_{k+1}^2$ , we have  $|A'| \leq tm + t \leq m/(16\gamma)$  and  $|B'| \leq k(tm + t) \leq n/4$ , and thus  $e(A \setminus A', B \setminus B') \geq nm/16$ . Therefore, using that  $|A| \leq n/2$ , there is a vertex  $v \in A \setminus A'$  with degree at least  $m/8$  into  $B \setminus B'$ . We let  $B^*$  be a set of  $m/8$  neighbours of  $v$  in  $B \setminus B'$ .

When  $k = 3$ , we will find a path on three vertices in  $B^*$  in the random graph, which will give, together with the three edges of  $G$  between  $v$  and those vertices, a copy of  $P_4^2$ . We argue as follows. If such a path does not appear, then there are less than  $m$  edges of  $G(n, p)$  in  $B^*$ . However the expected number of random edges within  $B^*$  is at least  $p \binom{m/8}{2} \geq 8ktm \log n$ , and therefore, by Chernoff's inequality (Lemma 2.4.1), with probability at least  $1 - n^{-4ktm}$  there are more than  $m$  edges of  $G(n, p)$  in  $B^*$ . We conclude by union bound over the at most

$$\binom{|A|}{tm + t} \binom{|B|}{k(tm + t)} \leq n^{3ktm}$$

choices for  $A'$  and  $B'$ .

For  $k \geq 4$  we let  $B_1, \dots, B_4$  be pairwise disjoint sets of size  $m/32$  in  $B^*$ . Moreover, we let  $B_5, \dots, B_k$  be pairwise disjoint sets of size  $n/(4k)$  in  $B \setminus B'$ , each disjoint from  $B_1, \dots, B_4$ . This is possible as  $|B \setminus B'| \geq n/4$ .

**Claim 3.6.7.** *With probability at least  $1 - n^{-\omega(m)}$  there exists vertices  $b_1, \dots, b_k$  with  $b_i \in B_i$  for  $i = 1, \dots, k$  such that in  $G(n, p)$  we have the edges  $b_i b_{i+1}$  for  $i = 1, \dots, k - 1$  and  $b_i b_{i+2}$  for  $i = 3, \dots, k - 2$ .*

Observe that, together with  $v$  and the edges  $vb_i$  for  $i = 1, \dots, 4$ , this gives a copy of  $P_{k+1}^2$  with vertices  $b_1, b_2, v, b_3, \dots, v_k$ . As there are at most  $n^{O(m)}$  choices for  $A'$  and  $B'$ , by a union bound and Claim 3.6.7, we a.a.s. find  $tm + t$  copies of  $P_{k+1}^2$ . It remains to prove the claim.

*Proof of Claim 3.6.7.* We denote by  $\{H_i\}_{i \in I}$  the graphs on  $k$  vertices  $b_1, \dots, b_k$  with  $b_i \in B_i$  for  $i = 1, \dots, k$  and edges  $b_i b_{i+1}$  for  $i = 1, \dots, k - 1$  and  $b_i b_{i+2}$  for  $i = 3, \dots, k - 2$ . Then, using the notation of Lemma 2.4.2, the expected number of those graphs appearing

in  $G(n, p)$  is

$$\begin{aligned} \mathbb{E}[X] &= |I|p^{2k-5} \geq \Omega(m^4 n^{k-4} p^{2k-5}) \\ &\geq \Omega\left(m(\log n)^{\frac{6}{2k-3} + \frac{2k-5}{2k-3}} n^{(k-4) + \frac{3(2k-4)}{2k-3} - \frac{(2k-5)(k-1)}{2k-3}}\right) = \omega(m \log n), \end{aligned}$$

where we used the bounds on the sizes of the sets  $B_i$  for  $i = 1, \dots, k$  in the first inequality and the bounds on  $m$  and  $p$  in the second inequality. On the other hand we get

$$\begin{aligned} \Delta[X] &= \sum_{H_i \sim H_j} p^{e(H_i) + e(H_j) - e(H_i \cap H_j)} \\ &\leq \sum_{r,s} O(m^{8-r} n^{2k-8-s} p^{2(2k-5) - (2s + \min\{2r-3, r-1\})}) \\ &\leq \mathbb{E}^2[X] \sum_{r,s} O(m^{-r} n^{-s} p^{-2s - \min\{2r-3, r-1\}}) \\ &\leq \mathbb{E}^2[X] O(n^{-2} p^{-1}) \leq \mathbb{E}^2[X] o(n^{-1}), \end{aligned}$$

where we split the sum according to the value of  $r$  and  $s$ , with  $0 \leq r \leq 4$ ,  $0 \leq s \leq k-4$ , and  $2 \leq r+s \leq k-1$ , where  $r$  and  $s$  are the number of common vertices of  $H_i$  and  $H_j$  in  $B_1, \dots, B_4$  and  $B_5, \dots, B_k$ , respectively. In the first inequality we used that  $e(H_i \cap H_j) \leq 2s + \min\{2r-3, r-1\}$ , and in the third inequality we used that  $O(m^{-r} n^{-s} p^{-2s - \min\{2r-3, r-1\}})$  is maximised for  $r = 0$  and  $s = 2$  with the given bounds on  $m$  and  $p$ . The claim follows by Lemma 2.4.2, as in the application above.  $\square$

$\square$

### 3.7 Supplementary proofs

In the final section of this chapter, we prove Lemmas 3.1.2, 3.4.6 and 3.4.7, whose proofs are either standard or a close adaptation of results from the literature.

We begin by proving Lemma 3.1.2, following the argument of [9, Lemma 12 and 13]. For this we consider a largest matching  $M$  in the reduced graph  $R$  and assume that  $|M| < (\alpha + 2kd)t$ . Then we will find a set  $I \subset V(R)$  of size roughly  $(1 - \alpha)t$  which contains very few edges. With the properties of the reduced graph, we conclude that the original graph  $G$  has to be  $(\alpha, \beta)$ -stable.

*Proof of Lemma 3.1.2.* Given an integer  $k \geq 2$  and  $0 < \beta < 1/12$ , we let  $0 < d < 10^{-4}k^{-2}\beta^6$ ,  $0 < \varepsilon < d/4$ ,  $4\beta \leq \alpha \leq 1/3$ , and  $t \geq 10/d$ . Then we let  $G$  be an  $n$ -vertex graph on vertex set  $V$  with minimum degree  $\delta(G) \geq (\alpha - d/2)n$  that is not  $(\alpha, \beta)$ -stable and we let  $R$  be the  $(\varepsilon, d)$ -reduced graph for some  $(\varepsilon, d)$ -regular partition  $V_0, \dots, V_t$  of  $G$ . We observe for the minimum degree of  $R$  that  $\delta(R) \geq (\alpha - 2kd)t$  because, otherwise,



there would be vertices with degree at most  $(\alpha - 2kd)t(n/t) + \varepsilon n < (\alpha - d/2)n - (d + \varepsilon)n$  in  $G'$  contradicting (P2).

Let  $M$  be a matching in  $R$  of maximal size. Observe that  $|M| \geq \min\{\delta(R), \lfloor t/2 \rfloor\} \geq (\alpha - 2kd)t$ . We assume  $|M| < (\alpha + 2kd)t$  and show that  $G$  must then be  $(\alpha, \beta)$ -stable, which is a contradiction. Let  $U = V(R) \setminus V(M)$ . We shall first show that there exists a set  $I \subset V(R)$  of size  $|U| + |M|$  that contains only few edges.

Since  $M$  is a matching of maximal size in  $R$ ,  $U$  is independent. Moreover, given an edge  $xy \in M$ , either  $x$  or  $y$  has at most one neighbour in  $U$ . Then we can split  $V(M)$  into two disjoint subsets  $X$  and  $Y$  by placing for each matching edge  $xy$  of  $M$  one of its endpoints with at most one neighbour in  $U$  into the subset  $X$ , and the other endpoint into the subset  $Y$ . We claim that  $I = U \cup X$  contains only few edges. We have  $e(U) = 0$ ,  $e(X, U) \leq |X|$ , and we can upper bound  $e(X)$  as follows. Let  $xy \in E(X)$  and denote by  $x'$  and  $y'$  the vertices matched to  $x$  and  $y$  in  $M$  respectively. Then  $x', y' \in Y$  and either  $x'$  or  $y'$  has at most one neighbour in  $U$ . Otherwise, there would be two distinct vertices  $x'', y'' \in U$  such that  $x'x''$  and  $y'y''$  are edges of  $R$ , and we could apply the rotation  $M \setminus \{xx', yy'\} \cup \{xy, x'x'', y'y''\}$  and get a larger matching, contradicting the maximality of  $M$ . Therefore,  $e(X) \leq |X||Z|$ , where  $Z = \{v \in Y \mid \deg(v, U) < 2\}$ . Observe that

$$e(Y, U) \leq (|Y| - |Z|)|U| + |Z|$$

and

$$e(Y, U) \geq |U|\delta(R) - e(X, U) \geq |U|\delta(R) - |X|$$

where we use that since  $U$  is independent, a vertex in  $U$  can have neighbours only in  $X$  and  $Y$ . We get

$$|Z| \leq \frac{(|Y| - \delta(R))|U| + |X|}{|U| - 1} < \frac{4kdt|U| + |X|}{|U| - 1} \leq 5kdt,$$

where the first inequality comes from the upper and lower bound on  $e(Y, U)$ , the second one from  $|Y| = |M| < (\alpha + 2kd)t$  and  $\delta(R) \geq (\alpha - 2kd)t$ , and the last one from  $|U| = t - 2|M| \geq t/4$ ,  $|X| = |M| < t/2$  and  $10/t \leq d$ . Hence,  $e(X) \leq |X|5kdt$ .

Therefore, the set  $I = U \cup X$  has size

$$|I| = |V(R)| - |M| = (1 - \alpha \pm 2kd)t$$

because  $|Y| = |M| = (\alpha \pm 2kd)t$  and contains at most

$$\begin{aligned} e(I) &\leq e(X) + e(X, U) + e(U) \leq |X|5kdt + |X| \\ &\leq (5kdt + 1)(\alpha + 2kd)t \leq 6\alpha kdt^2 \end{aligned}$$

edges, where we use  $|X| = |M| = (\alpha \pm 2kd)t$ , and  $d \leq \alpha/(20k)$  and  $10/t \leq d$  in the last

inequality.

We now move to the original graph  $G$  and prove that the existence of such set  $I$  in  $R$  implies that  $G$  is  $(\alpha, \beta)$ -stable. Let  $B'' = \bigcup_{i \in I} V_i$  be the union of the clusters  $I$ . Then  $|B''| = (1 - \alpha \pm 3kd)n$  and  $e(B'') \leq 6\alpha kdn^2$ . Let

$$B' = \{v \in B'' \mid \deg(v, B'') \leq \sqrt{kdn}\}.$$

Then  $e(B'') \geq (|B''| - |B'|)\sqrt{kdn}$  and, therefore, all but at most  $6\alpha\sqrt{kdn}$  vertices of  $B''$  belong to  $B'$  and, thus,  $|B'| = (1 - \alpha \pm 4\sqrt{kdn})n$ . Let

$$A' = \{v \in V \mid \deg(v, B') \geq (1 - \beta/4)|B'|\}$$

and note that  $A' \cap B' = \emptyset$ . Observe that if  $v \in B'$ , then

$$\begin{aligned} \deg(v, V \setminus B') &\geq \delta(G) - \deg(v, B') \\ &\geq (\alpha - d/2 - \sqrt{kdn})n \geq (\alpha - 2\sqrt{kdn})n. \end{aligned} \tag{3.7.1}$$

With  $|V \setminus B'| \leq (\alpha + 4\sqrt{kdn})n$  this implies

$$e(B', V \setminus B') \geq (\alpha - 2\sqrt{kdn})n|B'| \geq (|V \setminus B'| - 6\sqrt{kdn})|B'|$$

and with the definition of  $A'$  and the fact that  $|V \setminus (B' \cup A')| = |V \setminus B'| - |A'|$ , we get

$$\begin{aligned} e(B', V \setminus B') &\leq |V \setminus (B' \cup A')|(1 - \beta/4)|B'| + |A'||B'| \\ &= (|V \setminus B'| - |V \setminus (B' \cup A')|\beta/4)|B'|. \end{aligned}$$

The last two inequalities imply that all but at most  $24\sqrt{kdn}/\beta$  vertices of  $V \setminus B'$  belong to  $A'$ . Therefore we can bound the size of  $A'$  as follows

$$\begin{aligned} |A'| &\geq |V \setminus B'| - 24\sqrt{kdn}/\beta \\ &\geq \alpha n - 4\sqrt{kdn} - 24\sqrt{kdn}/\beta \geq \alpha n - \beta^2 n, \end{aligned}$$

and

$$|A'| \leq |V \setminus B'| \leq \alpha n + 4\sqrt{kdn} \leq \alpha n + \beta^2 n.$$

where we used in both inequalities that  $4\sqrt{kdn} + 24\sqrt{kdn}/\beta \leq \beta^2 n$ , as  $d \leq 10^{-4}k^{-2}\beta^6$ .

It follows that we have built two sets  $A'$  and  $B'$  such that  $|A' \cup B'| \geq n - \beta^2 n$ ,  $|A'| = \alpha n \pm \beta^2 n$  and  $|B'| = (1 - \alpha)n \pm \beta^2 n$ . Moreover each vertex of  $A'$  has at least  $(1 - \beta/4)|B'|$  neighbours in  $B'$  by the definition of  $A'$ , and each vertex of  $B'$  has at least  $(1 - \beta/2)|A'|$

neighbours in  $A'$ . This can be justified as follows. Given  $v \in B'$ ,

$$\begin{aligned} \deg(v, A') &= \deg(v, V \setminus B') - \deg(v, V \setminus (A' \cup B')) \\ &\geq \deg(v, V \setminus B') - |V \setminus (A' \cup B')| \\ &\geq (\alpha - 2\sqrt{kd})n - 24\sqrt{kd}n/\beta \geq (\alpha - \beta^2)n \\ &\geq \frac{\alpha - \beta^2}{\alpha + \beta^2}|A'| \geq (1 - \beta/2)|A'|, \end{aligned}$$

where we used that  $A'$  and  $B'$  are disjoint, the inequalities (3.7.1),  $|V \setminus (A' \cup B')| \leq 24\sqrt{kd}n/\beta$  and  $2\sqrt{kd} + 24\sqrt{kd}/\beta \leq \beta^2$ , the upper bound on  $|A'|$  and the inequality  $\alpha \geq 4\beta$ .

Now we need to take care of the vertices of  $G$  not yet covered by  $A' \cup B'$ , i.e. the at most  $\beta^2 n$  vertices in  $V \setminus (A' \cup B')$ . Let  $v$  be one such vertex. Then  $\deg(v, A' \cup B') \geq \delta(G) - |V \setminus (A' \cup B')| \geq (\alpha - d/2)n - \beta^2 n \geq \alpha n/2$ . Therefore, it is possible to add these vertices to  $A'$  and  $B'$  to obtain  $A \supseteq A'$  and  $B \supseteq B'$  such that each vertex of  $B$  has at least  $\alpha n/4$  neighbours in  $A$ , and each vertex of  $A$  has at least  $\alpha n/4$  neighbours in  $B$ . As we add at most  $\beta^2 n$  vertices, we have  $|A| = (\alpha \pm 2\beta^2)n$  and  $|B| = (1 - \alpha \pm 2\beta^2)n$ . Moreover, all but at most  $\beta^2 n \leq \beta n$  vertices from  $A$  have degree at least

$$(1 - \beta/4)|B'| \geq (1 - \beta/4)(|B| - \beta^2 n) \geq (1 - \beta)|B|$$

into  $B$ , where we used that  $|B| \leq |B'| + \beta^2 n$ ,  $|B| \geq (1 - \alpha - 2\beta^2)n \geq (2/3 - 2\beta^2)n$  and  $\beta < 1/12$ . Similarly, all but at most  $\beta^2 n \leq \beta n$  vertices from  $B$  have degree at least  $(1 - \beta/2)|A'| \geq (1 - \beta)|A|$  into  $A$ . Moreover, as  $B'$  is a subset of  $B''$  and we add at most  $\beta^2 n$  vertices to  $B'$  to get  $B$ , we have  $e(B) \leq e(B'') + \beta^2 n^2 \leq (6\alpha kd + \beta^2)n^2 \leq \beta n^2$ . Therefore,  $G$  is  $(\alpha, \beta)$ -stable according to Definition 1.1.10.  $\square$

We now prove Lemma 3.4.6, which is a standard application of the regularity method.

*Proof of Lemma 3.4.6.* To prove (i), without loss of generality, it suffices to show that the degree of every vertex in  $U_1$  is at least  $(1 - h\varepsilon)m^{h-1}$ . Fix any  $u_1 \in U_1$  and set  $N_1 = N_G(u_1, V)$ . Notice that since  $(V, U_1)$  is  $(\varepsilon, d)$ -super-regular, we have  $|N_1| \geq d|V| \geq \varepsilon|V|$ . Since  $(V, U_2)$  is  $(\varepsilon, d)$ -super-regular, there are at least  $(1 - \varepsilon)m$  vertices  $u_2 \in U_2$  such that the set  $N_2 = N_G(u_2, N_1)$  of neighbours of  $u_2$  in  $N_1$  has size at least  $(d - \varepsilon)|N_1| \geq (d - \varepsilon)d|V| \geq \varepsilon|V|$ . Continuing in the same way, by applying Lemma 2.1.1 to the  $(\varepsilon, d)$ -super-regular pair  $(V, U_j)$  for  $j = 3, \dots, h$ , we get that there are at least  $((1 - \varepsilon)m)^{j-1}$  choices of  $(u_2, \dots, u_j) \in U_2 \times \dots \times U_j$  such that the vertices  $u_1, u_2, \dots, u_j$  have at least  $(d - \varepsilon)^{j-1}|N_1| \geq (d - \varepsilon)^{j-1}d|V| \geq (d - \varepsilon)^{h-1}d|V| \geq \varepsilon|V|$  common neighbours in the set  $V$ . Since  $(d - \varepsilon)^{h-1}|N_1| \geq \frac{1}{2}d^h n$  and  $((1 - \varepsilon)m)^{h-1} \geq (1 - h\varepsilon)m^{h-1}$ , the first part of the lemma follows.

Without loss of generality, it suffices to prove (ii) for  $U_1$ . If  $|X| \geq 2\varepsilon nd^{1-h}$ , then, by applying Lemma 2.1.1, for all but at most  $\varepsilon m$  vertices  $u_1 \in U_1$ , the set  $N_1 = N(u_1, X)$  of neighbours of  $u_1$  in  $X$  is of size at least  $(d - \varepsilon)|X|$ . Fix any such  $u_1$  and proceed in the same way as in the proof of (i). We get that there are at least  $((1 - \varepsilon)m)^{h-1} \geq (1 - h\varepsilon)m^{h-1}$  choices of  $(u_2, \dots, u_h) \in U_2 \times \dots \times U_h$  such that the vertices  $u_1, u_2, \dots, u_h$  have at least  $(d - \varepsilon)^{h-1}|N_1| \geq (d - \varepsilon)^h|X| \geq \frac{1}{2}d^h|X|$  common neighbours in the set  $X$ , and the second part of the lemma follows.  $\square$

Finally, we move to the proof of Lemma 3.4.7, which follows from the Chebyshev's and Janson's inequalities.

*Proof of Lemma 3.4.7.* Given any graph  $H$  on  $h \geq 2$  vertices and any  $\delta > 0$ , we fix  $\varepsilon > 0$  with  $\varepsilon < 2^{-4h-24}h^{-8}\delta^{4h}$ , and we let  $\delta' = 2^{-3}h^{-1}\varepsilon^{1/4}$  and large enough for the inequalities indicated below to hold. Observe that the maximum degree of  $F$  is  $m^{h-1}$  and, by Lemma 3.4.6(i), the minimum degree of  $F$  is at least  $(1 - h\varepsilon)m^{h-1}$ . Therefore

$$\mathbb{E}[e(\tilde{F})] = e(F)p^{e(H)} = (1 \pm h\varepsilon)m^h p^{e(H)}$$

and

$$\begin{aligned} \text{Var}[e(\tilde{F})] &= O_{h,\delta} \left( \sum_{H' \subseteq H, e(H') > 0} m^{2v(H) - v(H')} \left( p^{2e(H) - e(H')} - p^{2e(H)} \right) \right) \\ &= O_{h,\delta} \left( \mathbb{E}[e(\tilde{F})]^2 \sum_{H' \subseteq H, e(H') > 0} m^{-v(H')} p^{-e(H')} \right) \\ &= O_{h,\delta} \left( \mathbb{E}[e(\tilde{F})]^2 \sum_{H' \subseteq H, e(H') > 0} n^{-v(H')} p^{-e(H')} \right) \\ &= O_{h,\delta} \left( \mathbb{E}[e(\tilde{F})]^2 C^{-1} n^{-1} \right), \end{aligned}$$

where we used that  $n^{-v(H')} p^{-e(H')} \leq C^{-e(H')} n^{-v(H') + e(H')/m_1(H)} \leq C^{-e(H')} n^{-1}$  in the last step. Using Chebyshev's inequality (Lemma 2.4.3), we have

$$\begin{aligned} \mathbb{P}[e(\tilde{F}) \neq (1 \pm \varepsilon)\mathbb{E}[e(\tilde{F})]] &= O_{h,\delta,\varepsilon} \left( \frac{\text{Var}[e(\tilde{F})]}{\mathbb{E}[e(\tilde{F})]^2} \right) \\ &= O_{h,\delta,\varepsilon}(C^{-1}n^{-1}), \end{aligned}$$

and thus a.a.s.

$$e(\tilde{F}) = (1 \pm \varepsilon)\mathbb{E}[e(\tilde{F})] = (1 \pm \varepsilon)(1 \pm h\varepsilon)m^h p^{e(H)}. \quad (3.7.2)$$

Similarly as above, given  $U'_i \subseteq U_i$  of size at least  $\delta m$  for  $i = 1, \dots, h$ , we have

$$\begin{aligned} \mathbb{E}[e(\tilde{F}')] &= e(F')p^{e(H)} \\ &= (1 \pm h\frac{\varepsilon}{\delta}) \prod_{i=1}^h |U'_i| p^{e(H)} \\ &= \Omega_{h,\delta,\varepsilon}(n^h p^{e(H)}) = \Omega_{h,\delta,\varepsilon}(Cn) \end{aligned}$$

and  $\Delta[e(\tilde{F}')] = O_{h,\delta,\varepsilon}(\mathbb{E}[e(\tilde{F}')]^2 C^{-1} n^{-1})$ . Then with Janson's inequality (Lemma 2.4.2) we have

$$\begin{aligned} \mathbb{P}[e(\tilde{F}') < (1 - \varepsilon)\mathbb{E}[e(\tilde{F}')] ] &\leq \exp\left(-\frac{\varepsilon^2 \mathbb{E}[e(\tilde{F}')]^2}{2\Delta[e(\tilde{F}')] + 2\mathbb{E}[e(\tilde{F}')]}\right) \\ &\leq \exp(-hn), \end{aligned}$$

where the last inequality holds for large enough  $C$ , and we conclude with a union bound that a.a.s.

$$\begin{aligned} e(\tilde{F}') &\geq (1 - \varepsilon)\mathbb{E}[e(\tilde{F}')] \\ &\geq (1 - \varepsilon)(1 - h\frac{\varepsilon}{\delta}) \prod_{i=1}^h |U'_i| p^{e(H)} \\ &\geq (1 - \sqrt{\varepsilon}) \prod_{i=1}^h |U'_i| p^{e(H)} \end{aligned} \tag{3.7.3}$$

for all choices of  $U'_i \subseteq U_i$  of size at least  $\delta m$  for  $i = 1, \dots, h$  and using the choice of  $\varepsilon$ . This proves the lower bound of (3.4.1). Note that (3.7.2) and (3.7.3) hold also with  $\delta$  replaced by  $\delta'$ .

Next we upper bound  $e(\tilde{F}')$  by taking  $e(\tilde{F})$  and subtracting those edges of  $\tilde{F}$  that are not in  $\tilde{F}'$ , i.e. the edges that contain at least one vertex  $v_i$  that belongs to  $U \setminus U'_i$ . We only need a lower bound on their number and we will see that it is enough to lower bound those for which  $|U_i \setminus U'_i| \geq \delta' m$  (which can be done using (3.7.3)), and simply ignore the others. For this we let  $J \subseteq [h]$  be the set of those indices  $j \in [h]$  such that  $|U'_j| \leq (1 - \delta')m$ , and for any  $\emptyset \neq I \subseteq J$  we let  $F_I$  be the subgraph of  $F$  induced by the sets  $U_i \setminus U'_i$  for  $i \in I$  and  $U'_i$  for  $i \notin I$ . If  $J = \emptyset$ , then the inequality  $e(\tilde{F}') \leq e(\tilde{F})$  already gives the desired upper bound on  $e(\tilde{F}')$ . Otherwise, using (3.7.2) and (3.7.3), we get

$$\begin{aligned} e(\tilde{F}') &\leq e(\tilde{F}) - \sum_{\emptyset \neq I \subseteq J} e(\tilde{F}_I) \\ &\leq (1 + \varepsilon)\mathbb{E}[e(\tilde{F})] - \sum_{\emptyset \neq I \subseteq J} (1 - \varepsilon)\mathbb{E}[e(\tilde{F}_I)], \end{aligned}$$

that we can further upper bound by

$$\begin{aligned}
 & (1 + \varepsilon)(1 + h\varepsilon) \prod_{i=1}^h |U_i| p^{e(H)} - \sum_{\emptyset \neq I \subseteq J} \left[ (1 - \varepsilon)(1 - h\frac{\varepsilon}{\delta'}) \prod_{i \in I} |U_i \setminus U'_i| \prod_{i \notin I} |U'_i| p^{e(H)} \right] \\
 \leq & \prod_{j \notin J} |U_j| \prod_{j \in J} |U'_j| p^{e(H)} + 2h\varepsilon \prod_{i=1}^h |U_i| p^{e(H)} + 2h\frac{\varepsilon}{\delta'} \sum_{\emptyset \neq I \subseteq J} \left( \prod_{i \in I} |U_i \setminus U'_i| \prod_{i \notin I} |U'_i| p^{e(H)} \right) \\
 \leq & (1 - \delta')^{|J|-h} \prod_{i=1}^h |U'_i| p^{e(H)} + 2h\varepsilon \delta^{-h} \prod_{i=1}^h |U'_i| p^{e(H)} + 2h\frac{\varepsilon}{\delta'} 2^{|J|} \delta^{-|J|} \prod_{i=1}^h |U'_i| p^{e(H)} \\
 \leq & \left( 1 + 2h\delta' + 2h\varepsilon \delta^{-h} + 2^{h+1} h \delta^{-h} \frac{\varepsilon}{\delta'} \right) \prod_{i=1}^h |U'_i| p^{e(H)} \leq (1 + \sqrt{\varepsilon}) \prod_{i=1}^h |U'_i| p^{e(H)}.
 \end{aligned}$$

To get the second line, we used  $(1 + \varepsilon)(1 + h\varepsilon) \leq 1 + 2h\varepsilon$ ,  $(1 - \varepsilon)(1 - h\frac{\varepsilon}{\delta'}) \geq 1 - 2h\frac{\varepsilon}{\delta'}$  and

$$\prod_{i=1}^h |U_i| - \sum_{\emptyset \neq I \subseteq J} \prod_{i \in I} |U_i \setminus U'_i| \prod_{i \notin I} |U'_i| = \prod_{j \notin J} |U_j| \prod_{j \in J} |U'_j|$$

as we are left with those edges that have vertices in  $U_j \setminus U'_j$  for  $j \notin J$  and in  $U'_j$  for  $j \in J$ . To get to the third line, we used that for  $j \notin J$  we have  $|U'_j| \geq (1 - \delta')m = (1 - \delta')|U_j|$  and that for each  $i$  we have  $|U_i \setminus U'_i| \leq |U_i| \leq \delta^{-1}|U'_i|$ . In the last estimate we use the bound on  $\varepsilon$  and the choice of  $\delta'$ . This finishes the proof of (3.4.1).

For (3.4.2), we repeat essentially the same argument we used for the lower bound of (3.4.1). Observe that from (ii) of Lemma 3.4.6, if  $|X| \geq 2\varepsilon n d^{1-h}$ , all but at most  $\varepsilon m$  vertices from each  $U_i$  have degree at least  $(1 - h\varepsilon)m^{h-1}$  in  $F_X$ . Therefore

$$\mathbb{E}[e(\tilde{F}'_X)] = e(F'_X) p^{e(H)} = (1 \pm h\frac{\varepsilon}{\delta}) \prod_{i=1}^h |U'_i| p^{e(H)} = \Omega_{h, \delta, \varepsilon}(Cn) \quad (3.7.4)$$

and again  $\Delta[e(\tilde{F}'_X)] = O_{h, \varepsilon, \delta}(\mathbb{E}[e(\tilde{F}'_X)]^2 C^{-1} n^{-1})$ . Then with Lemma 2.4.2 we have

$$\mathbb{P}[e(\tilde{F}'_X) < (1 - \varepsilon)\mathbb{E}[e(\tilde{F}'_X)]] \leq \exp\left(-\frac{\varepsilon^2 \mathbb{E}[e(\tilde{F}'_X)]^2}{2\Delta[e(\tilde{F}'_X)] + 2\mathbb{E}[e(\tilde{F}'_X)]}\right) \leq \exp(-hn),$$

where the last inequality holds for large enough  $C$ . Then with a union bound over all choices of  $U'_1, \dots, U'_h$  we conclude that  $e(\tilde{F}'_X) \geq (1 - \varepsilon)\mathbb{E}[e(\tilde{F}'_X)]$  with probability  $1 - e^{-n}$ , and using (3.7.4), we finish the proof.  $\square$

## Triangles in randomly perturbed graphs

In this chapter, we discuss our results related to the containment of  $m$  vertex-disjoint triangles in the perturbed graph model, namely Theorems 1.1.8 and 1.1.11 to 1.1.13. These theorems borrow a lot from the strategies used while dealing with the perturbed threshold for the square of a Hamilton cycle in Chapter 3. In fact, using that the square of a cycle on three vertices is a triangle, we do not need to develop any further technical result. The main Theorem 1.1.8 follows easily from the stability Theorem 1.1.11, the extremal Theorem 1.1.12, and the sublinear Theorem 1.1.13. The sublinear Theorem 1.1.13 is essentially a consequence of Lemma 3.1.1. Therefore we prove Theorem 1.1.8 and Theorem 1.1.13 already here. We prove the extremal Theorem 1.1.12 and the stability Theorem 1.1.11 in Sections 4.2 and 4.3 respectively, after giving a brief overview of both in Section 4.1.

*Proof of Theorem 1.1.8.* Let  $\beta_1, \gamma_1 > 0$  and  $C_1 > 0$  be given by Theorem 1.1.12 on input  $\alpha_0 = 1/512$ . Then let  $\gamma_2 > 0$  and  $C_2 > 0$  be given by Theorem 1.1.11 on input  $\beta = \min\{\alpha_0/4, \beta_1\}$ . Moreover, let  $C_3 > 0$  be given by Theorem 1.1.13. Define  $C = \max\{C_1, C_3\}$  and  $\gamma = \min\{\gamma_1, \gamma_2\}$ .

Let  $G$  be any  $n$ -vertex graph and  $p \geq C \log n/n$ , and define  $m = \min\{\delta(G), \lfloor n/3 \rfloor\}$ . If  $m \leq n/512$ , then we get from Theorem 1.1.13 that a.a.s.  $G \cup G(n, p)$  contains at least  $m$  pairwise vertex-disjoint triangles, as  $C \geq C_3$ . Otherwise,  $m > n/512$  and we can choose  $\alpha \in (\alpha_0, 1/3]$  such that  $(\alpha - \gamma)n \leq m \leq \alpha n$ . If  $G$  is  $(\alpha, \beta)$ -stable, then  $G$  is also  $(\alpha, \beta_1)$ -stable and, by Theorem 1.1.12, there are a.a.s. at least  $\min\{\delta(G), \lfloor \alpha n \rfloor\} \geq m$  pairwise vertex-disjoint triangles in  $G \cup G(n, p)$ , as  $\alpha_0 < \alpha \leq 1/3$  and  $C \geq C_1$ . Otherwise,  $G$  is not  $(\alpha, \beta)$ -stable and, by Theorem 1.1.11, a.a.s.  $G \cup G(n, p)$  contains at least  $\min\{\alpha n, \lfloor n/3 \rfloor\} \geq m$  pairwise vertex-disjoint triangles, as  $4\beta \leq \alpha \leq 1/3$  and  $p = \omega(1/n)$ .  $\square$

We now move to the proof of the sublinear Theorem 1.1.13 which uses Lemma 3.1.1. For readability, we restate Lemma 3.1.1 in the specific case  $k = 2$  and  $t = 1$ , which will be enough for our scope.

**Lemma 4.0.1** (Lemma 3.1.1 restated with  $k = 2$  and  $t = 1$ ). *There exists  $C > 0$  such that the following holds for any  $1 \leq m \leq n/64$  and any  $n$ -vertex graph  $G$  of minimum degree*

$\delta(G) \geq m$  and maximum degree  $\Delta(G) \leq n/64$ . For  $p \geq Cn^{-1} \log n$ , a.a.s. the perturbed graph  $G \cup G(n, p)$  contains  $m$  pairwise vertex-disjoint triangles.

*Proof of Theorem 1.1.13.* We let  $C$  be large enough such that with  $p \geq C \log n/n$  we can expose  $G(n, p)$  in three rounds as  $\bigcup_{i=1}^3 G_i$  with  $G_i \sim G(n, C_i \log n/n)$  for  $i = 1, \dots, 3$  such that the following hold. We let  $C_1 = 1$  and observe that, by a union bound, a.a.s. for any set of vertices  $U$  of size at least  $n/512$  there is at least one edge in  $G_1[U]$ . Next, we let  $C_2$  be large enough such that for a set of vertices  $U$  of size at least  $n/2$  there are a.a.s. at least  $\log^3 n$  pairwise vertex-disjoint triangles in  $G_2[U]$ , which is possible by [68, Theorem 3.29]. Finally, let  $C_3$  be such that we can apply Lemma 4.0.1 with  $C_3/2$ . We expose  $G_1$  already now and assume that the described property holds, while we leave  $G_2$  and  $G_3$  until we need them.

Let  $1 \leq m \leq n/512$  and let  $G$  be an  $n$ -vertex graph on vertex set  $V$  with minimum degree  $\delta(G) \geq m$ . To apply Lemma 4.0.1 to a large subgraph  $G'$  of  $G$ , we need  $\Delta(G') \leq v(G')/64$ . For this let  $V'$  be the set of vertices from  $G$  of degree at least  $n/128$ . If  $|V'| \geq m$ , then we let  $V''$  be any subset of  $V'$  of size  $m$  and we greedily find  $m$  pairwise vertex-disjoint triangles in  $G \cup G_1$ , each containing exactly one vertex from  $V''$ . Indeed, as long as we have less than  $m$  triangles there is a vertex  $v \in V''$  not yet contained in a triangle. Then there is a set  $U \subseteq N_G(v) \setminus V''$  of at least  $n/128 - 3m \geq n/512$  vertices not covered by triangles, and we can find an edge within  $G_1[U]$  that gives us a triangle containing  $v$  and two vertices from  $U$ .

Otherwise,  $|V'| < m$  and we remove  $V'$  from  $G$  to obtain  $G' = G[V \setminus V']$ . Note that we have  $v(G') = n - |V'| \geq n/2$ , minimum degree  $\delta(G') \geq m' = m - |V'|$ , and maximum degree  $\Delta(G') < n/128 \leq v(G')/64$ . If  $m' < (\log n)^3$ , then we a.a.s. find  $m'$  pairwise vertex-disjoint triangles within  $G_2[V(G')]$ . Otherwise,  $(\log v(G'))^3 \leq (\log n)^3 \leq m' \leq n/512 \leq v(G')/64$ , and, by Lemma 4.0.1 and as  $C_3 \log n/n \geq \frac{C_3}{2} \log v(G')/v(G')$ , there are a.a.s. at least  $m'$  pairwise vertex-disjoint triangles in  $G' \cup G_3[V(G')]$ .

Now that we found  $m'$  pairwise vertex-disjoint triangles, we can greedily add triangles by using the  $m - m'$  vertices from  $V'$  and an edge in their neighbourhood until we have  $m$  triangles. Analogous to above, as long as we have less than  $m$  triangles, for each available vertex  $v \in V'$ , there is a set  $U \subset N_G(v) \setminus V'$  of at least  $n/512$  vertices not covered by triangles, and we find an edge within  $G_1[U]$ .  $\square$

## 4.1 Proof overview of the extremal and the non-extremal case

The extremal Theorem 1.1.12 and the stability Theorem 1.1.11 can be proved along the same lines as the corresponding statements for the square of a Hamilton cycle in Chapter 3. Nevertheless, we provide a very brief overview and, for convenience, we restate the technical lemmas we need from Chapter 3, in the specific case of interest for this chapter. For simplicity, when outlining the proofs of Theorems 1.1.11 and 1.1.12, we assume  $\alpha = 1/3$ ,



$n$  is a multiple of 3, and  $G$  is an  $n$ -vertex graph with minimum degree  $\delta(G) \geq n/3$ , in which case both theorems give a triangle factor in  $G \cup G(n, p)$ .

#### 4.1.1 Extremal case.

Assume that  $G$  is  $(1/3, \beta)$ -stable and let  $p \geq C \log n/n$ . We want to show that a.a.s.  $G \cup G(n, p)$  contains a triangle factor. The definition of stability (Definition 1.1.10) gives a partition of  $V(G)$  into  $A \cup B$  where the size of  $B$  is roughly double the size of  $A$ , there is a minimum degree condition between  $A$  and  $B$ , and in each part all but at most a few vertices see all but at most a few vertices of the other part. Our proof will follow three steps. Firstly, we find a collection of pairwise vertex-disjoint triangles  $\mathcal{T}_1$ , such that after removing the triangles of  $\mathcal{T}_1$ , we are left with two sets  $A_1 = A \setminus V(\mathcal{T}_1)$  and  $B_1 = B \setminus V(\mathcal{T}_1)$  with  $|B_1| = 2|A_1|$ . The way we find these triangles depends on the sizes of  $A$  and  $B$  and we will use two different approaches when  $|B| > 2n/3$  and  $|B| \leq 2n/3$ . In particular when  $|B| > 2n/3$  we need to find some triangles entirely within  $B$ , just using the minimum degree  $n/3 - |A|$  and random edges. For that we will use Theorem 1.1.13.

Our second step is to cover the vertices in  $A_1$  and  $B_1$  that do not have a high degree to the other part; this will give two collections of pairwise vertex-disjoint triangles  $\mathcal{T}_2$  and  $\mathcal{T}_3$ . Each such triangle has one vertex in  $A_1$  and two vertices in  $B_1$  so that we still have  $|B_2| = 2|A_2|$ , where  $A_2 = A_1 \setminus V(\mathcal{T}_2 \cup \mathcal{T}_3)$  and  $B_2 = B_1 \setminus V(\mathcal{T}_2 \cup \mathcal{T}_3)$ . Moreover at this point, each vertex sees all but at most a few vertices of the other part. Given this high minimum degree condition between  $A_2$  and  $B_2$  and using Hall's matching theorem, we can find a perfect matching between the vertices in  $A_2$  and the edges of  $G(n, p)[B_2]$  in the following sense: we can match a vertex  $v \in A_2$  to an edge  $uw \in E(G(n, p)[B_2])$  if and only if  $uv$  and  $vw$  are both edges of  $G$ . Note that if  $v$  is matched to  $uw$ , then the vertices  $u, v, w$  induce a triangle in  $G \cup G(n, p)$ . This perfect matching corresponds to a collection of pairwise vertex-disjoint triangles  $\mathcal{T}_4$ , covering all the vertices of  $A_2$  and  $B_2$ . We conclude observing that the collection of triangles  $\mathcal{T}_1 \cup \mathcal{T}_2 \cup \mathcal{T}_3 \cup \mathcal{T}_4$  gives a triangle factor in  $G \cup G(n, p)$ .

#### 4.1.2 Non-extremal case.

Assume that  $G$  is not  $(1/3, \beta)$ -stable and let  $p \geq C/n$ . We want to show that a.a.s.  $G \cup G(n, p)$  contains a triangle factor. We apply the regularity lemma to  $G$  and obtain the reduced graph  $R$ , to which we apply Lemma 3.1.2 (with  $k = 1$ ). We restate it for an easier consultation.

**Lemma 4.1.1** (Lemma 3.1.2 restated with  $k = 1$ ). *For any  $0 < \beta < 1/12$  there exists  $d > 0$  such that the following holds for any  $0 < \varepsilon < d/4$ ,  $4\beta \leq \alpha \leq 1/3$ , and  $t \geq 10/d$ . Let  $G$  be an  $n$ -vertex graph with minimum degree  $\delta(G) \geq (\alpha - d/2)n$  that is not  $(\alpha, \beta)$ -stable and let  $R$  be the  $(\varepsilon, d)$ -reduced graph for some  $(\varepsilon, d)$ -regular partition  $V_0, \dots, V_t$  of  $G$ . Then  $R$  contains a matching  $M$  of size  $(\alpha + 2d)t$ .*

It follows that we can cover the vertices of  $R$  with cherries  $K_{1,2}$  and matching edges  $K_{1,1}$ , such that there are not too many cherries. Using the tools developed for dealing with the square of a Hamilton cycle (Lemmas 3.1.3 and 3.1.4), we can already find a triangle factor in each super-regular cherry and super-regular matching edge. Indeed, given an unbalanced super-regular cherry  $V,U,W$  of  $G$ , under certain additional assumptions, Lemma 3.1.3 with  $k = 2$  guarantees the existence of the square of a Hamilton path in  $G[V, U, W] \cup G(V, p) \cup G(U, W, p)$ , covering  $V$ ,  $U$ , and  $W$ , provided  $p \geq Cn^{-1}$ . Assuming that  $|V| + |U| + |W| \equiv 0 \pmod{3}$ , we can then easily extract a triangle factor. However, given the way we build the square of a Hamilton path in Lemma 3.1.3, this lemma requires an extra divisibility condition (possibly an artefact of our proof), which we are able to avoid here. Therefore we use the following lemma, for which we give a short proof in the next subsection.

**Lemma 4.1.2.** *For any  $0 < \delta' \leq d < 1$  there exist  $\delta_0, \delta_1, \varepsilon$  with  $\delta' \geq \delta_0 > 2\delta_1 > \varepsilon > 0$  and  $C > 0$  such that the following holds. Let  $U, V, W$  be pairwise disjoint sets such that  $|V| = n$  and  $(1 - \delta_0)n \leq |U| = |W| \leq (1 - \delta_1)n$ , where  $|V| + |U| + |W| \equiv 0 \pmod{3}$ . Suppose that  $(V, U)$  and  $(V, W)$  are  $(\varepsilon, d)$ -super-regular pairs with respect to a graph  $G$  and let  $G(V, p)$  and  $G(U, W, p)$  be random graphs with  $p \geq Cn^{-1}$ .*

*Then a.a.s. there exists a triangle factor in  $G[V, U, W] \cup G(V, p) \cup G(U, W, p)$  covering  $V, U$ , and  $W$ .*

The following lemma deals with regular matching edges.

**Lemma 4.1.3.** *For any  $0 < d < 1$  there exist  $\varepsilon > 0$  and  $C > 0$  such the following holds for sets  $U, V$  of size  $|V| = n$  and  $3n/4 \leq |U| \leq n$  where  $|V| + |U| \equiv 0 \pmod{3}$ . If  $(U, V)$  is an  $(\varepsilon, d)$ -super-regular pair and  $G(U, p)$  and  $G(V, p)$  are random graphs with  $p \geq C/n$ , then a.a.s. there exists a triangle factor.*

We omit the proof as it can be easily derived from Lemma 4.1.2 by appropriately splitting the super-regular pair  $(U, V)$  into two copies of super-regular cherries, in the exact same way as we derived Lemma 3.1.4 from Lemma 3.1.3 in Chapter 3.

However, before we can apply Lemma 4.1.2 to each cherry and Lemma 4.1.3 to each matching edge, some preliminary steps are needed. We remove some vertices from each cherry to make it unbalanced and ensure that both edges are super-regular. Then we cover all vertices that are not contained in any of the cherries or edges by finding a collection of triangles  $\mathcal{T}_1$ . We construct another collection of triangles  $\mathcal{T}_2$  to ensure that in each cherry the relations between the three sets are as required by Lemma 4.1.2. For constructing  $\mathcal{T}_1$  and  $\mathcal{T}_2$  we will mainly rely on the minimum degree condition of  $G$  and the fact that in the probability  $p \geq C/n$ , the constant  $C$  can be chosen large enough so that a.a.s. the following holds: each linear-sized set contains a random edge, and for any not too small part of a regular pair and any linear-sized set there is a triangle containing an edge from the pair

and the third vertex from the set. Finally, we can use Lemma 4.1.2 and Lemma 4.1.3 to cover the remaining vertices with a collection of triangles  $\mathcal{T}_3$ . Together  $\mathcal{T}_1 \cup \mathcal{T}_2 \cup \mathcal{T}_3$  gives a triangle factor in  $G \cup G(n, p)$ .

We mention already now that when  $\alpha$  is sufficiently smaller than  $1/3$  and the condition on the minimum degree of  $G$  reads as  $\delta(G) \geq (\alpha - \gamma)n$ , many of the steps outlined above are not necessary. In this case indeed we do not have to cover all graph with triangles and we only want to find  $\alpha n$  pairwise vertex-disjoint triangles in  $G \cup G(n, p)$ . We will see that an application of Lemma 4.1.2 and Lemma 4.1.3 to the cherries and the matching edges found at the beginning (after having made them super-regular and suitable for Lemma 4.1.2) is already enough to find these  $\alpha n$  triangles.

### 4.1.3 Proof of Lemma 4.1.2

*Proof of Lemma 4.1.2.* Given  $0 < \delta' \leq d < 1$ , let  $\varepsilon' > 0$  be given by Lemma 2.1.3 on input  $d^3/64$  and let  $0 < \delta < \delta_0 \leq \min\{\delta', d^3/22\}$ . Furthermore, let  $C \geq 8\delta^{-2}$ , let  $0 < \varepsilon < \delta/4$  be given by Lemma 3.4.3 on input  $H$  being an edge,  $d$ ,  $\frac{\delta}{3(1-\delta)}$  (in place of  $\delta$ ), and  $\varepsilon'/2$ , and let  $p \geq C/n$ .

Suppose  $U, V, W$  are disjoint sets of size  $|V| = n$  and  $(1 - \delta_0)n \leq |U| = |W| \leq (1 - \delta)n$  with  $|V| + |U| + |W| \equiv 0 \pmod{3}$ , and  $G$  is a graph with vertex set  $U \cup V \cup W$  such that the pairs  $(V, U)$  and  $(V, W)$  are  $(\varepsilon, d)$ -super-regular with respect to  $G$ . Let  $\delta_1$  be such that  $|U| = |W| = (1 - \delta_1)n$  and observe that  $\delta \leq \delta_1 \leq \delta_0$ . We reveal random edges  $G_1 \sim G(U, W, p)$  and  $G_2 \sim G(V, p)$  and we have that a.a.s. any set of size at least  $\delta n$  in  $V$  contains an edge of  $G_2$ . Indeed, for a fixed set of size at least  $\delta n$ , the probability that it does not contain an edge of  $G_2$  is at most  $(1 - p)^{\binom{\delta n}{2}} \leq \exp(-p \binom{\delta n}{2}) \leq \exp(-2n)$  since  $C \geq 8\delta^{-2}$ , and we conclude by a union bound over the at most  $2^n$  choices of such set. Then we apply Lemma 3.4.3 to  $G_1$  and we obtain a matching  $M \subseteq G_1$  of size  $|M| = \left(1 - \frac{\delta}{3(1-\delta)}\right) |W| = \left(1 - \frac{\delta}{3(1-\delta)}\right) (1 - \delta_1)n$  such that the pair  $(M, V)$  is  $(\varepsilon'/2, d^3/32)$ -super-regular with respect to the graph  $\mathcal{T}_G(M, V)$  defined in Definition 3.4.2. As for  $x \in (0, 1)$  the function  $x \rightarrow x/(1 - x)$  is increasing and  $\delta \leq \delta_1$ , we have  $\left(\frac{\delta_1}{3} - \frac{\delta}{3(1-\delta)}(1 - \delta_1)\right) \geq \left(\frac{\delta_1}{3} - \frac{\delta_1}{3(1-\delta_1)}(1 - \delta_1)\right) \geq 0$ . Thus by ignoring  $\left(\frac{\delta_1}{3} - \frac{\delta}{3(1-\delta)}(1 - \delta_1)\right)n \leq d^3n/64$  edges of  $M$ , we get a subset  $M' \subseteq M$  with  $|M'| = (1 - 4\delta_1/3)n$ .

Next, let  $U' = U \setminus V(M')$  and  $W' = W \setminus V(M')$  be the sets of vertices in  $U$  and  $W$ , respectively, that are not incident to edges of  $M'$ . Note that both  $U'$  and  $W'$  have size  $\delta_1 n/3$ . We want to cover these vertices with triangles having the other two vertices in  $V$ . Any vertex  $v \in U' \cup W'$  has degree at least  $dn$  into  $V$  and as  $d > 2\delta_0 \geq 2\delta_1$  we can pick these triangles greedily for each  $v \in U' \cup W'$  using  $G_2$ . Let  $V' \subseteq V$  be the vertices that were used for these triangles and observe  $|V \setminus V'| = |M'| = (1 - 4\delta_1/3)n$ .

To obtain the triangle factor it remains to find a perfect matching in  $H_G(M', V \setminus V')$ . By Lemma 2.1.3 it is sufficient to observe that the pair  $(M', V \setminus V')$  is  $(\varepsilon', d^3/64)$ -super-regular

with respect to  $H_G(M', V \setminus V')$ , which holds because  $(M, V)$  is  $(\varepsilon'/2, d^3/32)$ -super-regular with respect to  $H_G(M, V)$ .  $\square$

## 4.2 Proof of the extremal Theorem 1.1.12

*Proof of Theorem 1.1.12.* Let  $0 < \alpha_0 \leq 1/3$  and choose  $d = 1/2$ . Let  $0 < \varepsilon < 1/5$ , and choose  $\beta$  and  $\gamma$  with  $0 < \beta < \alpha_0 \varepsilon/36$  and  $0 < \gamma < \beta/11$ . Let  $C_1$  be given by Theorem 1.1.13, and choose  $C_3$  such that a.a.s. the random graph  $G(n, C_4 \log n/n)$  contains a perfect matching. Finally, let  $C \geq 2C_1 + 1 + 2C_3/\alpha_0$  and  $\alpha_0 \leq \alpha \leq 1/3$ .

Let  $p \geq C \log n/n$ . With our choice of  $C$ , we can reveal  $G(n, p)$  in three rounds  $G_1 \sim G(n, 2C_1 \log n/n)$  and  $G_2 \sim G(n, \log n/n)$ , and  $G_3 \sim G(n, \frac{2C_3}{\alpha_0} \log n/n)$ . We will only know later in which subset we will use  $G_1$  and  $G_3$ , but we have that a.a.s. there is an edge of  $G_2$  between any two not necessarily disjoint sets of size  $\beta n$ . Indeed, fixed two such sets, the probability that there is no edge of  $G_2$  is at most  $(1 - \log n/n)^{(\beta n)^2} \leq \exp(-\beta^2 n \log n)$ , and we conclude by a union bound over the at most  $2^{2n}$  choices for the two sets. Now let  $G$  be an  $n$ -vertex graph with minimum degree  $\delta(G) \geq (\alpha - \gamma)n$  that is  $(\alpha, \beta)$ -stable and define  $m_0 = \max\{n/3 - \delta(G), n/3 - \lfloor \alpha n \rfloor\}$ . Our goal is to a.a.s. find pairwise vertex-disjoint triangles in  $G \cup G(n, p)$  such that at most  $3m_0$  vertices are left uncovered.

To aid with calculations we let  $\kappa = n/3 - \lfloor \alpha n \rfloor$  and observe that  $\kappa \in \{0, -1/3, -2/3\}$  if  $\alpha = 1/3$  and that  $\kappa > 0$  if  $\alpha < 1/3$  and  $n$  large enough. Also note that  $m_0 - \kappa$  is an integer and that  $3\kappa = \lfloor (1 - \alpha)n \rfloor - 2\lfloor \alpha n \rfloor$ . With this we set  $w = \max\{3\kappa, 0\}$ . As  $G$  is  $(\alpha, \beta)$ -stable we get a partition of  $V(G)$  into sets  $A$  and  $B$  satisfying the conditions of Definition 1.1.10.

**Claim 4.2.1.** *There a.a.s. are a collection of triangles  $\mathcal{T}_1$  in  $G \cup G_1 \cup G_2$  with  $|\mathcal{T}_1| \leq \beta n$  and a set  $W \subseteq V(G) \setminus V(\mathcal{T}_1)$  with  $|W| \leq 3m_0 - w$  such that the following holds. For  $A_1 = A \setminus (V(\mathcal{T}_1) \cup W)$  and  $B_1 = B \setminus (V(\mathcal{T}_1) \cup W)$ , we have that  $|A_1| \leq \lfloor \alpha n \rfloor$ ,  $|B_1| = 2|A_1| + w$ , the minimum degree between  $A_1$  and  $B_1$  is at least  $\alpha n/5$ , all but at most  $\beta n$  vertices of  $A_1$  have degree at least  $|B_1| - \beta n$  into  $B_1$ , and all but at most  $\beta n$  vertices of  $B_1$  have degree at least  $|A_1| - \beta n$  into  $A_1$ .*

The sets  $A_1$  and  $B_1$  partition  $V(G) \setminus (V(\mathcal{T}_1) \cup W)$  and, after proving Claim 4.2.1, we will cover all but  $w$  vertices from  $A_1 \cup B_1$  with additional triangles. Hence, if we manage to find these triangles, we have covered all but  $|W| + w \leq 3m_0$  vertices, as desired. We remark for later that  $|W| \leq 3m_0 - w \leq 4\gamma n$ .

*Proof of Claim 4.2.1.* We have either  $|B| > \lfloor (1 - \alpha)n \rfloor$  or  $|A| \geq \lfloor \alpha n \rfloor$ . First suppose that we are in the first case, where  $|B| = \lfloor (1 - \alpha)n \rfloor + m$  for some  $1 \leq m \leq \beta n$  (and  $|A| = \lfloor \alpha n \rfloor - m$ ), and note that

$$|B| - 2|A| = n - 3\lfloor \alpha n \rfloor + 3m = 3m + 3\kappa > 0.$$

If  $1 \leq m \leq m_0 - \kappa$ , then  $0 < 3m \leq 3m_0 - 3\kappa$  and we let  $W$  be any set with  $\min\{3m, 3m+3\kappa\} \leq 3m_0 - w$  vertices from  $B$ . Then with the choice of  $\mathcal{T}_1 = \emptyset$ , we have that the sets  $A_1 = A$  and  $B_1 = B \setminus W$  partition  $V(G) \setminus W$ , and  $|A_1| = \lceil \alpha n \rceil - m$  and  $|B_1| = |B| - \min\{3m, 3m+3\kappa\} = 2|A_1| + w$ . If on the other hand  $m_0 < m + \kappa$ , then

$$\delta(G[B]) \geq \delta(G) - |A| \geq (n/3 - m_0) - (\lceil \alpha n \rceil - m) = m - m_0 + \kappa > 0$$

and we observe that  $m - m_0 + \kappa$  is an integer. Moreover  $m - m_0 + \kappa \leq m \leq |B|/256$ , where we use  $m_0 - \kappa \geq 0$ ,  $m \leq \beta n \leq \alpha_0 \varepsilon n / 36 \leq n / (3 \cdot 5 \cdot 36)$  and  $|B| = \lfloor (1 - \alpha)n \rfloor + m \geq n/2 + m$ . Thus, by Theorem 1.1.13 and as  $2C_1 \log n/n \geq C_1 \log |B|/|B|$  we a.a.s. find  $m - m_0 + \kappa$  pairwise vertex-disjoint triangles in  $(G \cup G_1)[B]$ . Denote by  $\mathcal{T}_1$  the collection of these  $m - m_0 + \kappa$  triangles. Let  $W$  be any set of  $3m_0 - w$  vertices from  $B$  not covered by any triangle in  $\mathcal{T}_1$ . Then the sets  $A_1 = A$  and  $B_1 = B \setminus (V(\mathcal{T}_1) \cup W)$  partition  $V(G) \setminus (V(\mathcal{T}_1) \cup W)$ , and we have  $|A_1| = \lceil \alpha n \rceil - m$  and

$$|B_1| = |B| - 3(m - m_0 + \kappa) - (3m_0 - w) = 2|A_1| + w.$$

It remains to consider the second case, where  $|A| = \lceil \alpha n \rceil + m$  for some  $0 \leq m \leq \beta n$ . First, we greedily pick  $m$  pairwise vertex-disjoint triangles in  $G \cup G_2$  each with two vertices in  $A$  and one vertex in  $B$ . Indeed during the process, there is always a vertex  $v$  in  $B$ , not yet contained in a triangle, with at least  $\deg(v, A) - 2m \geq (\alpha/4 - 2\beta)n \geq \beta n$  uncovered neighbours in  $A$  in the graph  $G$ . By the property assumed in  $G_2$  we can then find an edge within these neighbours of  $v$  to get a triangle. Denote by  $\mathcal{T}_1$  the collection of these  $m$  triangles.

If  $\kappa \geq 0$ , then, with the choice of  $W = \emptyset$ , we have that  $A_1 = A \setminus V(\mathcal{T}_1)$  and  $B_1 = B \setminus V(\mathcal{T}_1)$  partition  $V(G) \setminus V(\mathcal{T}_1)$  and

$$|B_1| = |B| - |\mathcal{T}_1| = \lfloor (1 - \alpha)n \rfloor - 2m = 2(\lceil \alpha n \rceil - m) + 3\kappa = 2|A_1| + w.$$

If  $\kappa < 0$ , we additionally pick a set  $W$  of vertices not covered by triangles from  $\mathcal{T}_1$ , such that  $|W| = 1$ ,  $|W \cap A| = 1$ ,  $|W \cap B| = 0$  if  $\kappa = -2/3$ , and  $|W| = 2$ ,  $|W \cap A| = |W \cap B| = 1$  if  $\kappa = -1/3$ . Then, the sets  $A_1 = A \setminus (V(\mathcal{T}_1) \cup W)$  and  $B_1 = B \setminus (V(\mathcal{T}_1) \cup W)$  partition  $V(G) \setminus (V(\mathcal{T}_1) \cup W)$ , and  $|B_1| = 2|A_1| + w$ . Indeed, if  $\kappa = -2/3$  we have  $|A_1| = |A| - 2|\mathcal{T}_1| - 1 = \lceil \alpha n \rceil - m - 1$  and

$$|B_1| = |B| - |\mathcal{T}_1| = \lfloor (1 - \alpha)n \rfloor - 2m = 2(\lceil \alpha n \rceil - m) + 3\kappa = 2|A_1|$$

and if  $\kappa = -1/3$  we have  $|A_1| = |A| - 2|\mathcal{T}_1| - 1 = \lceil \alpha n \rceil - m - 1$  and

$$|B_1| = |B| - |\mathcal{T}_1| - 1 = \lfloor (1 - \alpha)n \rfloor - 2m - 1 = 2(\lceil \alpha n \rceil - m) + 3\kappa - 1 = 2|A_1|.$$

Observe, that in both the first and the second case  $|B_1| = 2|A_1| + w$  and  $|W| \leq 3m_0 - w$ . Moreover, as we remove at most  $3m_0 - w \leq 4\gamma n \leq \alpha n/20$  vertices from each  $A$  and  $B$ , the minimum degree between  $A_1$  and  $B_1$  is at least  $\alpha n/4 - \alpha n/20 = \alpha n/5$ . The other conditions on the degrees between  $A_1$  and  $B_1$  are clearly satisfied, because for all but at most  $\beta n$  vertices from each set there are still at most  $\beta n$  non-neighbours in the other set. The bounds  $|\mathcal{T}| \leq \beta n$  and  $|A_1| \leq \lceil \alpha n \rceil$  also hold in all cases.  $\square$

We want to cover all but  $w$  vertices in  $A_1 \cup B_1$  and we start from those vertices in  $A_1$  and  $B_1$  that do not have a high degree to the other part. We will always cover them with triangles with one vertex in  $A_1$  and two vertices in  $B_1$  to ensure that the relation between the number of vertices remaining in  $A_1$  and  $B_1$  does not change. Let

$$\tilde{A}_1 = \{v \in A_1 : \deg(v, B_1) \leq |B_1| - 9\beta n\}$$

and

$$\tilde{B}_1 = \{v \in B_1 : \deg(v, A_1) \leq |A_1| - 9\beta n\},$$

and observe that  $|\tilde{A}_1|, |\tilde{B}_1| \leq \beta n$ .

We claim that a.a.s. we can greedily pick pairwise vertex-disjoint triangles in  $(G \cup G_2)[A_1 \cup B_1]$  that cover all vertices of  $\tilde{A}_1$ , with each triangle having one vertex in  $\tilde{A}_1$  and two vertices in  $B_1 \setminus \tilde{B}_1$ . Indeed, at each step during the process, an uncovered vertex  $v$  in  $\tilde{A}_1$  has at least  $\deg(v, B_1) - |\tilde{B}_1| - 2|\tilde{A}_1| \geq (\alpha/5 - 3\beta)n \geq \beta n$  uncovered neighbours in  $B_1 \setminus \tilde{B}_1$  in the graph  $G$ . We then find an edge of  $G_2$  within these neighbours of  $v$  and build a triangle. Denote by  $\mathcal{T}_2$  the collection of these triangles and note that  $|\mathcal{T}_2| \leq \beta n$ .

Observe that at this point  $2|\mathcal{T}_2| \leq 2\beta n$  vertices of  $B_1 \setminus \tilde{B}_1$  have already been covered. We claim that a.a.s. we can greedily pick pairwise vertex-disjoint triangles in  $(G \cup G_2)[(A_1 \cup B_1) \setminus V(\mathcal{T}_2)]$  that cover all vertices of  $\tilde{B}_1$ , where each triangle has one vertex in  $A_1 \setminus \tilde{A}_1$ , one vertex in  $\tilde{B}_1$  and one vertex in  $B_1 \setminus \tilde{B}_1$ . Indeed, at each step during the process, an uncovered vertex  $v$  in  $\tilde{B}_1$  has at least  $\deg(v, A_1) - |\tilde{A}_1| - |\tilde{B}_1| \geq (\alpha/5 - 2\beta)n \geq \beta n$  uncovered neighbours in  $A_1 \setminus \tilde{A}_1$  in the graph  $G$  and at least

$$\begin{aligned} \delta(G) - 3|\mathcal{T}_1| - |W| - \deg(v, A_1) - 2|\mathcal{T}_2| - 2|\tilde{B}_1| \\ \geq (\alpha - \gamma)n - 3\beta n - 4\gamma n - (\lceil \alpha n \rceil - 9\beta n) - 4\beta n \geq \beta n \end{aligned}$$

uncovered neighbours in  $B_1 \setminus \tilde{B}_1$  in the graph  $G$ . We then find an edge of  $G_2$  between these two neighbourhood sets to get a triangle. Denote by  $\mathcal{T}_3$  the collection of these triangles and note that  $|\mathcal{T}_3| \leq \beta n$ .

The sets  $A_2 = A_1 \setminus V(\mathcal{T}_2 \cup \mathcal{T}_3)$  and  $B_2 = B_1 \setminus V(\mathcal{T}_2 \cup \mathcal{T}_3)$  give a partition of the remaining

vertices in  $V(G) \setminus (V(\mathcal{T}_1) \cup V(\mathcal{T}_2) \cup V(\mathcal{T}_3) \cup W)$ . We have

$$|A_2| \geq |A| - 2|\mathcal{T}_1| - |\mathcal{T}_2| - |\mathcal{T}_3| - |W| \geq \alpha n - 5\beta n - 4\gamma n \geq \alpha n/2$$

and  $|B_2| = 2|A_2| + w$ . Moreover, the degree from  $A_2$  to  $B_2$  is at least  $|B_1| - 9\beta n - 2|\mathcal{T}_2 \cup \mathcal{T}_3| = |B_2| - 9\beta n$  and the degree from  $B_2$  to  $A_2$  is at least  $|A_1| - 9\beta n - |\mathcal{T}_2 \cup \mathcal{T}_3| = |A_2| - 9\beta n$ . Let  $W'$  be any subset of  $B_2$  of size 2 and let  $B'_2 = B_2 \setminus W'$ . Observe that  $|B'_2| = 2|A_2|$ . Since  $\frac{2C_3}{\alpha_0} \log n/n \geq C_3 \log |B'_2|/|B'_2|$ , a.a.s.  $G_3[B'_2]$  contains a perfect matching  $M$ .

Let  $\mathcal{B}$  be the following auxiliary bipartite graph with classes  $A_2$  and  $E(M)$ . There is an edge between a vertex  $v \in A_2$  and an edge  $uw \in E(M)$  if and only if the vertices  $u$  and  $w$  are both neighbours of the vertex  $v$  in the graph  $G$ . Observe that if  $v$  is connected with  $uw$  in  $\mathcal{B}$ , then  $\{v, u, w\}$  induces a triangle in  $G \cup G_3$ . Using Hall's condition, the graph  $\mathcal{B}$  has a perfect matching. Indeed, the degree of each  $v \in A_2$  in  $\mathcal{B}$  is at least  $|E(M)| - 18\beta n \geq |E(M)|/2$ , and the degree of each  $uw \in E(M)$  in  $\mathcal{B}$  is at least  $|A_2| - 18\beta n \geq |A_2|/2$ . Therefore there exists a perfect matching in  $\mathcal{B}$ , which gives a triangle factor  $\mathcal{T}_4$  in  $(G \cup G_3)[V']$ , where  $V' = V(G) \setminus (V(\mathcal{T}_1) \cup V(\mathcal{T}_2) \cup V(\mathcal{T}_3) \cup W \cup W')$ .

Then  $\mathcal{T}_1 \cup \mathcal{T}_2 \cup \mathcal{T}_3 \cup \mathcal{T}_4$  contains at least

$$(n - |W| - |W'|)/3 \geq n/3 - m_0 \geq \min\{\delta(G), \lfloor \alpha n \rfloor\}$$

pairwise vertex-disjoint triangles covering  $V(G) \setminus (W \cup W')$ .  $\square$

We point out that under certain conditions our proof of Theorem 1.1.12 gives more triangles. When  $\alpha < 1/3$  and  $|A| \geq \alpha n$ , as  $|W| \leq 3m_0 - w$ , we get  $\lfloor \alpha n \rfloor$  pairwise vertex-disjoint triangles in  $G \cup G(n, p)$ , even when  $\delta(G) < \alpha n$ . Similarly, when  $\alpha = 1/3$  and  $|A| \geq n/3$ , as  $|W| \leq 2$ , we get  $\lfloor n/3 \rfloor$  pairwise vertex-disjoint triangles in  $G \cup G(n, p)$ , even when  $\delta(G) < n/3$ .

### 4.3 Proof of the stability Theorem 1.1.11

*Proof of Theorem 1.1.11.* We start by defining necessary constants. Given  $0 < \beta < 1/12$ , let  $d > 0$  be obtained from Lemma 4.1.1, and set  $\gamma = d/2$  and  $t_0 = 11/d$ . Next, we take any  $0 < \delta' < 160^{-2}d^2$  and use Lemma 4.1.2 on input  $d/2$  and  $\delta'$  to obtain  $\delta_0, \delta, \varepsilon'$  with  $\delta' \geq \delta_0 > \delta > \varepsilon' > 0$  and  $C_1$ . Additionally we assume that  $C_1$  is large enough and  $\varepsilon'$  is small enough for Lemma 4.1.3 to hold with input  $d/2$ . Finally, let  $C_2$  be given by Lemma 2.5.5 on input  $d/2$ . We let  $0 < \varepsilon \leq \varepsilon'/2$ . In summary, the dependencies between our constants are as follows:

$$\varepsilon \ll \varepsilon' < \delta < \delta_0 \leq \delta' \ll d \ll \beta < \frac{1}{12} \quad \text{and} \quad \frac{1}{t_0}, \gamma \ll d.$$

We apply Lemma 2.1.4 with  $\varepsilon$  and  $t_0$  to obtain  $T$ . We take  $C$  large enough such that, for  $p \geq C/n$ , the random graph  $G(n, p)$  contains the union  $G_1 \cup G_2 \cup G_3$ , where  $G_1 \sim G(n, 2C_1T/n)$ ,  $G_2 \sim G(n, 4C_2T/(dn))$ , and  $G_3 \sim G(n, 96T^2/(d^2n))$ .

Now, for any  $\alpha$  with  $4\beta \leq \alpha \leq 1/3$ , let  $G$  be an  $n$ -vertex graph on the vertex set  $V$  with minimum degree  $\delta(G) \geq (\alpha - \gamma)n$  that is not  $(\alpha, \beta)$ -stable. With the regularity lemma (Lemma 2.1.4) applied to  $G$ , we get  $G'$ ,  $t_0 < t + 1 \leq T$  and a partition  $V_0, \dots, V_t$  of  $V(G)$  such that (P1) – (P4) hold. Define  $n_0 = |V_1| = |V_2| = \dots = |V_t|$  and observe that  $(1 - \varepsilon)n/t \leq n_0 \leq n/t$ . We denote by  $R$  the  $(\varepsilon, d)$ -reduced graph for  $G$ , that is, the graph on the vertex set  $[t]$  with edges  $ij$  corresponding to  $\varepsilon$ -regular pairs  $(V_i, V_j)$  of density at least  $d$  in  $G'$ . We observe that the minimum degree of  $R$  satisfies  $\delta(R) \geq (\alpha - 2d)t$  because, otherwise, there would be vertices with degree at most  $(\alpha - 2d)t(n/t) + \varepsilon n < (\alpha - \gamma)n - (d + \varepsilon)n$  in  $G'$ , contradicting (P2).

The purpose of  $G_1$  will become clear later, but we describe some useful properties of  $G_2$  and  $G_3$  now. Let  $U$  and  $W$  be any two clusters that give an edge in  $R$ ,  $V$  any cluster, and  $U' \subseteq U$ ,  $W' \subseteq W$ ,  $V' \subseteq V$  three pairwise disjoint subsets each of size  $dn_0/2$ . Then, with  $G_2$  and as  $4C_2T/(dn) \geq 2C_2/(dn_0)$ , by Lemma 2.5.5 we have that with probability at least  $1 - 2^{-4(dn_0/2)/(d/2)} = 1 - 2^{-4n_0}$

$$\text{there is a triangle in } G \cup G_2 \text{ with one vertex in each set } U', W', V'. \quad (4.3.1)$$

With a union bound over the at most  $t^3 2^{3n_0}$  choices for  $U, W, V$  and  $U', W', V'$ , we conclude that a.a.s. (4.3.1) holds for all choices as above.

With  $G_3$  we a.a.s. have that

$$\text{any set } A \text{ of size at least } dn_0/2 \text{ contains an edge of } G_3. \quad (4.3.2)$$

In fact, given any set  $A$  of size at least  $dn_0/2$ , the expected number of edges of  $G_3$  in  $A$  is

$$\binom{|A|}{2} \cdot \frac{96T^2}{d^2n} \geq \frac{1}{3} \cdot \frac{d^2n_0^2}{4} \cdot \frac{96T^2}{d^2n} = 8T^2 \frac{n_0^2}{n} \geq 2n,$$

where we used that  $n/n_0 \leq t/(1 - \varepsilon) \leq 2T$ . Therefore the probability that the set  $A$  does not contain an edge of  $G_3$  is at most  $(1 - \frac{96T^2}{d^2n})^{\binom{|A|}{2}} \leq \exp\left(-\binom{|A|}{2} \cdot \frac{96T^2}{d^2n}\right) \leq \exp(-2n)$  and (4.3.2) follows from a union bound over the at most  $2^n$  choices for  $A$ .

Now let  $M_1$  be a largest matching in  $R$ . Since  $G$  is not  $(\alpha, \beta)$ -stable, using Lemma 4.1.1, we conclude that  $|M_1| \geq (\alpha + 2d)t$ . At this point, for the sake of clarity, we split our proof into two cases  $-0 < \alpha < 1/3 - d/3$  and  $1/3 - d/3 \leq \alpha \leq 1/3$  – although some steps will be the same. The first case is indeed much easier, as we do not need to cover all the graph with triangles, while in the second case we are looking for a spanning structure and we want to find  $\lfloor n/3 \rfloor$  pairwise vertex-disjoint triangles.



**Case**  $0 < \alpha < 1/3 - d/3$ . As  $M_1$  is a largest matching in  $R$ , the set  $V(R) \setminus V(M_1)$  is independent and only one endpoint of each edge of  $M_1$  can be adjacent to more than one vertex from  $V(R) \setminus V(M_1)$ . Therefore, we can greedily pick a second matching  $M_2$  such that each edge of  $M_2$  contains a vertex of  $V(R) \setminus V(M_1)$  and a vertex of  $V(M_1)$ , and  $M_2$  covers at least  $\min\{|V(R) \setminus V(M_1)|, \delta(R)\}$  vertices of  $V(R) \setminus V(M_1)$ . The two matchings  $M_1$  and  $M_2$  together cover a subset  $V(M_1 \cup M_2) \subset V(R)$  of

$$2|M_1| + \min\{|V(R) \setminus V(M_1)|, \delta(R)\} \geq \min\{t, (3\alpha + 2d)t\} \geq (3\alpha + d)t$$

vertices, and we can extract a collection of  $|M_2|$  vertex-disjoint cherries and a disjoint matching that cover such vertices. This gives a subgraph  $R' \subseteq R$  consisting of cherries and a matching such that for all edges  $ij \in E(R')$  the pair  $(V_i, V_j)$  is  $\varepsilon$ -regular of density at least  $d$  in  $G'$ , and therefore in  $G$  as well. We denote by  $\mathcal{J} \subset [t]$  the indices of the clusters  $V_i$  of the cherries and the matching edges in  $R'$  and we observe from above that  $|\mathcal{J}| \geq (3\alpha + d)t$ . We add to  $V_0$  all the vertices of  $G$  that are in the clusters  $V_j$  for  $j \notin \mathcal{J}$ .

Then we make all pairs associated with the edges of  $R'$  super-regular. Given a pair  $(A, B)$ , by Lemma 2.1.1, all but at most  $\varepsilon n_0$  vertices of  $A$  (resp.  $B$ ) have degree at least  $(d - \varepsilon)n_0$  to  $B$  (resp.  $A$ ). For every such pair we remove these vertices from  $A$  and  $B$ , and remove additional vertices to ensure all clusters have the same size. As  $R'$  only contains vertex-disjoint cherries and a disjoint matching, we can achieve that by removing a total of at most  $2\varepsilon n_0$  vertices from each cluster. We add all the removed vertices to  $V_0$ . Observe that afterwards all the pairs  $(A, B)$  associated with the edges of  $R'$  are  $(2\varepsilon, d - 3\varepsilon)$ -super-regular, because every vertex  $a \in A$  has degree at least  $(d - \varepsilon)n_0 - 2\varepsilon n_0 \geq (d - 3\varepsilon)|B|$  into  $B$ , and every vertex  $b \in B$  has degree at least  $(d - 3\varepsilon)|A|$  into  $A$ .

Recall that for a later application of Lemma 4.1.2 we need that for each cherry the sizes of the leaf-clusters are smaller than the size of the centre-cluster. Thus for each cherry  $ijk$  of  $R'$ , with  $j$  being the centre, we additionally remove  $\delta|V_j| \leq \delta n_0$  vertices from the leaves  $V_i$  and  $V_k$ , and add them to  $V_0$ . We have  $|V_i| = |V_k| = (1 - \delta)|V_j|$  that implies  $|V_i| = |V_k| \geq (1 - \delta_0)|V_j|$ , as  $\delta_0 > \delta$ . We have that all edges of  $R'$  still give  $(2\varepsilon, d - 3\varepsilon - \delta)$ -super-regular pairs. Moreover

$$\left| \bigcup_{j \in \mathcal{J}} V_j \right| \geq (1 - 2\varepsilon - \delta)n_0|\mathcal{J}| \geq (1 - \varepsilon)(1 - 2\varepsilon - \delta)(3\alpha + d)n \geq 3\alpha n.$$

We can assume (by moving only a few additional vertices to  $V_0$  that do not harm the bounds above) that for all cherries and matching edges in  $R'$  the number of vertices in the clusters together is divisible by three.

For each such super-regular cherry  $ijk$  of  $R'$ , after revealing  $G_1[V_i \cup V_j \cup V_k]$  we find by Lemma 4.1.2 a.s. a triangle factor covering all the vertices in  $V_i \cup V_j \cup V_k$ . Similarly

for any matching edge  $ij$  of  $R'$ , after revealing  $G_1[V_i \cup V_j]$  we find by Lemma 4.1.3 a.a.s. a triangle factor covering all the vertices in  $V_i \cup V_j$ . Note that we apply Lemma 4.1.2 and Lemma 4.1.3 only constantly many times and thus a.a.s. we get a triangle factor in all such applications. Let  $\mathcal{T}$  be the union of all such triangle factors. Then  $\mathcal{T}$  covers  $|\bigcup_{j \in \mathcal{J}} V_j| \geq 3\alpha n$  vertices and gives at least  $\alpha n = \min\{\alpha n, \lfloor n/3 \rfloor\}$  pairwise vertex-disjoint triangles in  $G \cup G(n, p)$ .

**Case  $1/3 - d/3 \leq \alpha \leq 1/3$ .** As discussed in the overview, here we cannot directly apply Lemma 4.1.2 and Lemma 4.1.3 as in the case  $0 < \alpha < 1/3 - d/3$ , but we need additional steps. However even with a lower minimum degree, we will cover all vertices of  $G$  and find  $\lfloor n/3 \rfloor$  pairwise vertex-disjoint triangles. Recall that  $M_1$  is a largest matching and that  $|M_1| \geq (\alpha + 2d)t$ . Then the set  $V(R) \setminus V(M_1)$  is independent, has size

$$|V(R) \setminus V(M_1)| = t - 2|M_1| \leq (1 - 2\alpha - 4d)t \leq (\alpha - 3d)t$$

and only one endpoint of each edge of  $M_1$  can be adjacent to more than one vertex from  $V(R) \setminus V(M_1)$ . Given that  $\delta(R) \geq (\alpha - 2d)t$ , we can greedily pick a second matching  $M_2$  such that each edge of  $M_2$  contains a vertex of  $V(R) \setminus V(M_1)$  and a vertex of  $V(M_1)$ , and  $M_2$  covers the remaining vertices  $V(R) \setminus V(M_1)$  completely. Therefore, the two matchings  $M_1$  and  $M_2$  together cover the vertex set  $V(R)$  and we can extract a collection of  $\ell = |M_2| \leq (\alpha - 3d)t$  vertex-disjoint cherries and a disjoint matching that cover  $V(R)$ .

This gives a spanning subgraph  $R' \subseteq R$  on vertex set  $[t]$  containing  $\ell \leq (\alpha - 3d)t$  cherries and a matching of size  $(t - 3\ell)/2 \geq (1 - 3\alpha + 9d)t/2 \geq 9dt/2$  such that for all edges  $ij \in E(R')$  the pair  $(V_i, V_j)$  is  $\varepsilon$ -regular of density at least  $d$  in  $G'$ , and therefore in  $G$  as well. We denote by  $\mathcal{I} \subseteq [t]$  the indices of the clusters  $V_i$  that are not the centre of a cherry in  $R'$ . As above, with Lemma 2.1.1, we can make the pairs associated with the edges of  $R'$   $(2\varepsilon, d - 3\varepsilon)$ -super-regular, while keeping the clusters all of the same size. For this we have to remove at most  $t2\varepsilon n_0 \leq t2\varepsilon n/t = 2\varepsilon n$  vertices, which we add to  $V_0$ . Next, as for a later application of Lemma 4.1.2 we need that for each cherry the sizes of the leaf-clusters are smaller than the size of the centre-cluster, we remove for each  $i \in \mathcal{I}$  additionally  $\delta|V_i| \leq \delta n_0 \leq \delta n/t$  vertices from  $V_i$  and add them to  $V_0$ . Note that we remove vertices from the clusters of matching edges as well, although this is not necessary. We then get  $|V_0| \leq \varepsilon n + 2\varepsilon n + t\delta n/t \leq 2\delta n$ . We can assume (by moving only a few additional vertices to  $V_0$  that do not harm the bounds above) that for all cherries and matching edges in  $R'$  the number of vertices in the clusters together is divisible by three. By removing  $n \pmod{3} \in \{0, 1, 2\}$  vertices from  $V_0$  we also have  $|V_0| \equiv 0 \pmod{3}$ ; note that this only happens when  $n$  is not divisible by 3 and we can discard these vertices.

**Covering  $V_0$  with triangles.** We now want to cover the exceptional vertices in  $V_0$  by triangles. It would be easy to do this greedily by just using (4.3.2), but it might happen that afterwards in many of the cherries the number of vertices is not divisible by three or that

the centre cluster gets too small. To avoid both these issues, we will cover  $V_0$  while using the same number of vertices from clusters that are together in a cherry or matching edge. For this we will always cover three vertices at a time and combine (4.3.1) with (4.3.2) to find additional triangles. Observe that  $|V_0 \cup \bigcup_{i \notin \mathcal{I}} V_i| \leq 2\delta n + \ell n/t \leq (\alpha - \gamma)n - 2dn$  and, therefore, any  $v \in V_0$  has at least  $2dn$  neighbours in  $\bigcup_{i \in \mathcal{I}} V_i$ .

Assume we have already covered  $V' \subseteq V_0$  vertices of  $V_0$  using at most  $5|V'|$  triangles in total. Let  $W'$  be the set of vertices from  $\bigcup_{i \in \mathcal{I}} V_i$  used for the triangles covering  $V'$  and note that  $|W'| \leq 30\delta n$ . Then let  $\mathcal{I}' \subseteq \mathcal{I}$  be the set of indices of clusters  $V_i$  with  $i \in \mathcal{I}$  which intersect  $W'$  in at least  $\sqrt{\delta}n_0$  vertices and note that  $|\mathcal{I}'| \leq |W'|/(\sqrt{\delta}n_0) \leq 30\sqrt{\delta}t/(1-\varepsilon) \leq 40\sqrt{\delta}t \leq dt/4$ . Moreover, notice that as for each  $v \in V_0$  we have  $\deg_G(v, \bigcup_{i \in \mathcal{I}} V_i) \geq 2dn$ , there are at least  $dt$  indices  $i \in \mathcal{I}$  such that  $v$  has at least  $dn_0$  neighbours in  $V_i$ . In particular, as  $|\mathcal{I}'| \leq dt/4$  and  $t \geq 10/d$ , there are at least  $dt - |\mathcal{I}'| \geq 3dt/4 \geq 7$  indices  $i \in \mathcal{I} \setminus \mathcal{I}'$  such that  $v$  has at least  $dn_0$  neighbours in  $V_i$ . Therefore we can pick three vertices  $v_1, v_2, v_3 \in V_0 \setminus V'$  and three indices  $i_1, i_2, i_3$  in  $\mathcal{I} \setminus \mathcal{I}'$  such that  $v_j$  has  $dn_0$  neighbours in  $V_{i_j}$  for  $j = 1, 2, 3$  and the clusters  $V_{i_1}, V_{i_2}, V_{i_3}$  belong to pairwise different cherries or matching edges. For  $j = 1, 2, 3$  with (4.3.2) we find an edge  $e_j$  in  $G_3[N(v_j, V_{i_j}) \setminus W']$  and we cover the three vertices with triangles. It is easy to show that we can find at most 10 additional triangles with the help of (4.3.1) and (4.3.2), in such a way that, overall, for each cherry and matching edge, we use the same number of vertices from each of their clusters; in particular, the number of vertices used from each cherry and matching edge is divisible by three. The clusters  $V_{i_1}, V_{i_2}, V_{i_3}$  can belong to three cherries, two cherries and one matching edge, one cherry and two matching edges, or three matching edges. We give details in the case where they are all leaves of (different) cherries, and we refer to Figure 4.1 for the other three cases. With (4.3.1) we find four triangles: two with a vertex in each of the other cluster of the cherry containing  $V_{i_1}$  and the third vertex in one of the other clusters of the cherry containing  $V_{i_2}$ , and other two triangles with one vertex in each of the other cluster of the cherry containing  $V_{i_3}$  and the third vertex in the remaining cluster of the cherry containing  $V_{i_2}$ . When a  $V_{i_j}$  belongs to a matching edge of  $R'$ , we first find with (4.3.2) two triangles inside this matching edge each with one vertex in the cluster  $V_{i_j}$  and the other two vertices in the other cluster of the matching edge, then we proceed as before (see Figure 4.1). Note that we cover three vertices of  $V_0$  using at most 13 triangles, and thus to cover  $V' \subseteq V_0$  we use at most  $13|V'|/3 \leq 5|V'|$ , as claimed above. Therefore we can repeat this procedure until  $V' = V_0$ .

Let  $\mathcal{T}_1$  be the set of triangles we found above to cover  $V_0$  and keep the divisibility condition. We now update the regularity partition by deleting  $V(\mathcal{T}_1)$  from each  $V_i$  for  $i \in [t]$  and note that for all cherries and matchings from  $R'$  the number of vertices in the clusters together is divisible by three. We recall that so far we removed at most  $(2\varepsilon + \delta + 2\sqrt{\delta})n_0$  vertices from each cluster, where the first (resp. second, third) term bounds the number of vertices removed for making each pair super-regular (resp. for a

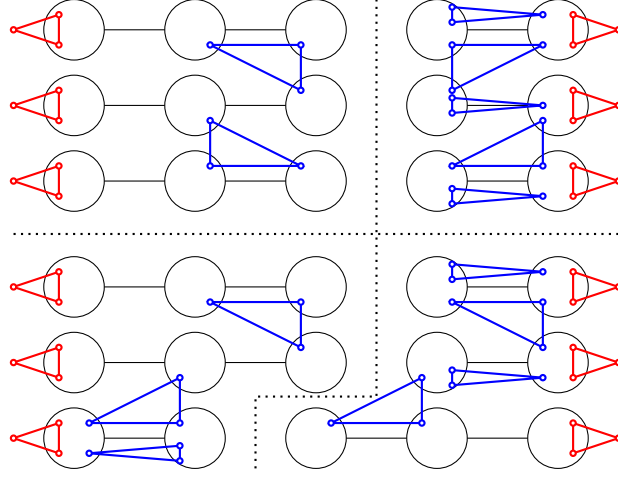


Figure 4.1: Embeddings of triangles for absorbing  $V_0$  while using the same number of vertices from each cluster within a cherry or a matching edge. Each red triangle covers a vertex of  $V_0$ . Each blue triangle stands for two triangles with end-points in the same clusters; we only draw one for simplicity.

later application of Lemma 4.1.2, for covering  $V_0$ ).

**Balancing the partition.** Now the matching edges in  $R'$  are already ready for an application of Lemma 4.1.3 and we will not modify the corresponding clusters anymore. However, before an application of Lemma 4.1.2 to the cherries in  $R'$ , we need to ensure that the ratio between their size and the size of the centre-cluster satisfies the hypotheses of the lemma. This is what we are going to do now. For  $i \notin \mathcal{I}$  we denote by  $W_i$  and  $U_i$  the leaf-clusters of the cherry centred at  $V_i$ . Before covering the vertices of  $V_0$ , we had  $|W_i| = |U_i| \leq (1 - \delta)|V_i|$  for  $i \notin \mathcal{I}$ , which still holds as we removed the same number of vertices from each cluster of a cherry.

However we still need to guarantee the other inequality  $|W_i|, |U_i| \geq (1 - \delta_0)|V_i|$ . For that, we find  $2m$  triangles with two vertices in  $V_i$ , of which one half has the third vertex in  $U_i$  and the other half in  $W_i$ , where  $m$  is the smallest integer such that

$$|U_i| - m \geq (1 - \delta_0)(|V_i| - 4m). \quad (4.3.3)$$

Then after removing these  $2m$  triangles, we will have precisely  $(1 - \delta_0)|V_i| \leq |U_i| = |W_i|$ . Observe that the inequality (4.3.3) implies that  $m \geq \frac{(1 - \delta_0)|V_i| - |U_i|}{4(1 - \delta_0) - 1}$  and, as we chose the smallest such  $m$ , we get  $m \leq \lceil \frac{(1 - \delta_0)|V_i| - |U_i|}{4(1 - \delta_0) - 1} \rceil$ . Moreover as  $\delta < \delta_0$ ,  $|V_i| \leq n_0$  and  $|U_i| \geq (1 - 2\varepsilon - \delta - 2\sqrt{\delta})n_0$ , we have  $\frac{(1 - \delta_0)|V_i| - |U_i|}{4(1 - \delta_0) - 1} < \frac{(1 - \delta_0) - (1 - 2\varepsilon - \delta - 2\sqrt{\delta})}{2}n_0 < 2\sqrt{\delta}n_0$ . Therefore, for  $n$  (and thus  $n_0$ ) large enough,  $m \leq 2\sqrt{\delta}n_0$ . We can find these at most  $4\sqrt{\delta}n_0$  triangles, by iteratively picking them with (4.3.2) and removing the corresponding vertices from  $U_i$ ,  $W_i$ , and  $V_i$ . Indeed, for any  $v \in W_i \cup U_i$  we have degree into  $V_i$  at least  $(d - 3\varepsilon - \delta - 10\sqrt{\delta})n_0 \geq dn_0/2$ , as we started from  $(2\varepsilon, d - 3\varepsilon)$ -super-regular pairs and  $\delta < \delta' < 160^{-2}d^2$ .

Note that afterwards we still have  $|U_i| \leq (1 - \delta)|V_i|$  as for large enough  $n$  and with  $\delta < \delta_0$  we have  $m \leq \lceil \frac{(1-\delta_0)|V_i|-|U_i|}{4(1-\delta_0)-1} \rceil \leq \frac{(1-\delta)|V_i|-|U_i|}{4(1-\delta)-1}$ . Therefore, we have  $(1 - \delta_0)|V_i| \leq |U_i| = |W_i| \leq (1 - \delta)|V_i|$ . Moreover with  $d - 3\varepsilon - \delta - 10\sqrt{\delta} \geq d/2$  and  $2\varepsilon \leq \varepsilon'$ , we get that the pairs  $(U_i, V_i)$  and  $(W_i, V_i)$  are  $(\varepsilon', d/2)$ -super-regular. Let  $\mathcal{T}_2$  be the set of triangles we removed during this phase.

**Completing the triangles.** Now for any  $i \notin \mathcal{I}$ , after revealing  $G_1[V_i \cup W_i \cup U_i]$ , we a.a.s. find a triangle factor covering the vertices of  $U_i, W_i$ , and  $V_i$  by Lemma 4.1.2. Similarly for any matching edge  $ij$  of  $R'$  observe that  $(V_i, V_j)$  is a  $(\varepsilon', d/2)$ -super-regular pair. Then after revealing  $G_1[V_i \cup V_j]$ , we a.a.s. find a triangle factor covering the vertices of  $V_i$  and  $V_j$  by Lemma 4.1.3. Note that we apply Lemma 4.1.2 and Lemma 4.1.3 only constantly many times and thus a.a.s. we get a triangle factor in all such applications. Let  $\mathcal{T}_3$  be the union of the triangle factors we obtain for each  $i \notin \mathcal{I}$  and each matching edge  $ij$  from  $R'$ . Then  $\mathcal{T}_1 \cup \mathcal{T}_2 \cup \mathcal{T}_3$  gives  $\lfloor n/3 \rfloor$  pairwise vertex-disjoint triangles in  $G \cup G(n, p)$ .  $\square$

---

## Transversals in hypergraph collection

---

In this chapter, we discuss our general sufficient condition for a family of hypergraphs to be  $d$ -colour-blind. We start by explaining the relevant terminology in Section 5.1, then we state our result in Section 5.2. We overview its proof in Section 5.3 and we give a full proof in Section 5.4. Finally we prove the applications listed in Theorem 1.2.5 in Section 5.5.

### 5.1 Setting and terminology

This section introduces the relevant terminology. An *ordered hypergraph* is a hypergraph equipped with a linear order of its vertex set. For convenience, we often index the vertices of an ordered  $n$ -vertex hypergraph with  $\{1, 2, \dots, n\}$  so that  $v_i < v_j$  if and only if  $i < j$ . A subgraph of an ordered hypergraph inherits an ordering from the parent hypergraph in the obvious way. Whenever we state that two ordered hypergraphs are isomorphic, we mean that they are isomorphic as ordered hypergraphs.

**Definition 5.1.1** ( $\ell$ -link). Let  $k, \ell, m \in \mathbb{N}$  with  $\ell \leq m$ . Let  $\mathcal{A} = (V, E)$  be an ordered  $k$ -uniform hypergraph on  $m$  vertices. We call  $\mathcal{A}$  an  $\ell$ -link of uniformity  $k$  if  $\mathcal{A}_s$  and  $\mathcal{A}_t$  are isomorphic, where  $\mathcal{A}_s = \mathcal{A}[\{v_1, \dots, v_\ell\}]$ , and  $\mathcal{A}_t = \mathcal{A}[\{v_{m-\ell+1}, \dots, v_m\}]$ . We refer to  $m$  as the order of  $\mathcal{A}$  and we call the ordered hypergraphs  $\mathcal{A}_s$  and  $\mathcal{A}_t$  the start and the end of  $\mathcal{A}$ , respectively.

**Definition 5.1.2** ( $\mathcal{A}$ -chain). Let  $k, \ell, m \in \mathbb{N}$  with  $\ell \leq m$ , and  $\mathcal{A}$  be an  $\ell$ -link of uniformity  $k$  and order  $m$ . We say that an ordered hypergraph  $\mathcal{P}$  is an  $\mathcal{A}$ -chain if the following properties hold.

- (i)  $v(\mathcal{P}) = n = (m - \ell)t + \ell$  for some  $t \in \mathbb{N}$ .
- (ii) Set  $S_1 = \{1, \dots, m\}$  and for  $1 < q \leq t$  define  $S_q \subseteq [n]$  recursively as follows. For  $1 < q \leq t$ , if  $S_{q-1} = \{s, \dots, s + m - 1\}$ , define  $S_q = \{s + m - \ell, \dots, s + 2m - \ell - 1\}$ . Then, for each  $1 \leq q \leq t$ , the hypergraph  $\mathcal{P}_q = \mathcal{P}[\{v_i : i \in S_q\}]$  is isomorphic to  $\mathcal{A}$ .
- (iii) Each edge of  $\mathcal{P}$  is contained in  $\mathcal{P}_q$  for some  $q \in [t]$ .

We refer to  $t$  as the length of the  $\mathcal{A}$ -chain and we call  $\mathcal{P}_1$  and  $\mathcal{P}_t$  the first and the last links of  $\mathcal{P}$ , respectively. Moreover, we call the start of  $\mathcal{P}_1$  and the end of  $\mathcal{P}_t$  the start and the end of  $\mathcal{P}$ , respectively, and refer to them collectively as the ends of  $\mathcal{P}$ .

**Definition 5.1.3** ( $\mathcal{A}$ -cycle). Let  $\mathcal{P}$  be an  $\mathcal{A}$ -chain. Let  $\mathcal{S}$  and  $\mathcal{T}$  be the start and the end of  $\mathcal{P}$ , respectively. Let  $\phi$  be the isomorphism between the ordered hypergraphs  $\mathcal{S}$  and  $\mathcal{T}$ , and identify  $x \in \mathcal{S}$  with  $\phi(x) \in \mathcal{T}$  for each  $x \in \mathcal{S}$ . We call the resulting (unordered) hypergraph an  $\mathcal{A}$ -cycle.

We remark that with  $\mathcal{A}$  being an  $\ell$ -link of order  $m$ , if  $\mathcal{P}$  is an  $\mathcal{A}$ -chain and  $C$  is an  $\mathcal{A}$ -cycle, then the following holds:  $v(C) \in (m - \ell)\mathbb{N}$  and  $e(C) = \frac{e(\mathcal{A}) - e(\mathcal{A}_s)}{m - \ell} v(C)$ , while  $v(\mathcal{P}) \in (m - \ell)\mathbb{N} + \ell$  and  $e(\mathcal{P}) = \frac{e(\mathcal{A}) - e(\mathcal{A}_s)}{m - \ell} v(\mathcal{P}) - O(1)$ , where  $O(1)$  stands for a constant which only depends on  $\mathcal{A}$ .

Observe that for each  $1 \leq \ell \leq k$ , a single  $k$ -uniform edge induces an  $\ell$ -link of uniformity  $k$  and order  $k$ , and its chain (resp. cycle) corresponds to a  $k$ -uniform  $\ell$ -path (resp. cycle). Figure 5.1 shows the case  $k = 5$  and  $\ell = 2$ . Similarly, the complete graph on  $r$  vertices induces a  $(r - 1)$ -link of uniformity 2 and order  $r$ , and its chain (resp. cycle) corresponds to the  $(r - 1)$ -th power of a path (resp. cycle). Figure 5.2 illustrates the case  $r = 3$ . Finally, Figure 5.3 shows that a pillar can also be obtained as an  $\mathcal{A}$ -chain.

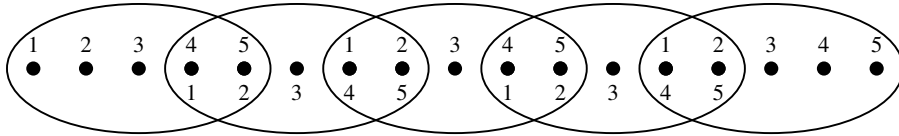


Figure 5.1: A 5-uniform 2-path is an  $\mathcal{A}$ -chain, with  $\mathcal{A}$  being (any ordering of) a single 5-uniform edge. The numbering of the vertices in an edge denotes the (ordered) isomorphism between that edge and  $\mathcal{A}$ .

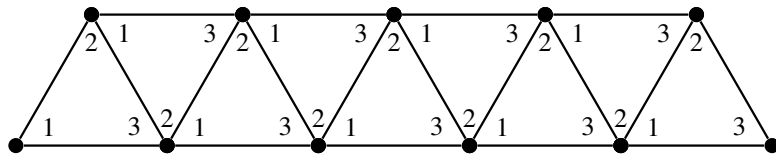


Figure 5.2: The square of a path is an  $\mathcal{A}$ -chain, with  $\mathcal{A}$  being (any ordering of) a triangle.

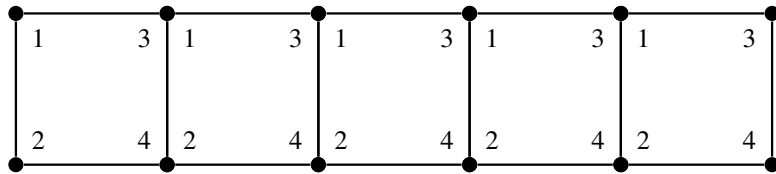


Figure 5.3: A pillar is an  $\mathcal{A}$ -chain, with  $\mathcal{A}$  being the above ordering of a cycle on 4 vertices.

We now state the properties we require from the link  $\mathcal{A}$  for our main theorem to hold.

**Definition 5.1.4.** Let  $k, \ell, m \in \mathbb{N}$  with  $\ell \leq m$ ,  $\mathcal{A}$  be an  $\ell$ -link of order  $m$  and uniformity  $k$ , and  $d \in [k - 1]$ . We say that  $\mathcal{A}$  is  $(\delta, d)$ -good if the following three properties hold.

**Ab.** For any  $\alpha > 0$ , there exist  $0 < \tau, \eta \leq \alpha$  and  $n_0 \in \mathbb{N}$  so that if  $\mathcal{H}$  is a  $k$ -uniform hypergraph on  $n \geq n_0$  vertices with  $\delta_d(\mathcal{H}) \geq (\delta + \alpha)n^{k-d}$ , then there exists  $A \subseteq V(\mathcal{H})$  of size at most  $\tau n$  with the following property.

For any  $L \subseteq V(\mathcal{H}) \setminus A$  of size at most  $\eta n$  with  $|L| \in (m - \ell)\mathbb{N}$ , there exists an embedding of an  $\mathcal{A}$ -chain to  $\mathcal{H}$  with vertex set  $A \cup L$ . Furthermore, the embedding of the start and the end of the  $\mathcal{A}$ -chain does not depend on the subset  $L$ .

**Con.** For any  $\alpha > 0$ , there exist a positive integer  $c$  and  $n_0 \in \mathbb{N}$  so that if  $\mathcal{H}$  is a  $k$ -uniform hypergraph  $\mathcal{H}$  on  $n \geq n_0$  vertices with  $\delta_d(\mathcal{H}) \geq (\delta + \alpha)n^{k-d}$ , the following holds.

Let  $\mathcal{S}$  and  $\mathcal{T}$  be vertex-disjoint copies of  $\mathcal{A}_s$  in  $\mathcal{H}$ . Then,  $\mathcal{H}$  contains an embedding of an  $\mathcal{A}$ -chain of length at most  $c$  with start  $\mathcal{S}$  and end  $\mathcal{T}$ .

**Fac.** For any  $\alpha > 0$ , there exist  $\beta_0 > 0$  and  $n_0 \in \mathbb{N}$  so that the following holds for any  $n \geq n_0$  and  $\beta \leq \beta_0$ .

Let  $\mathbf{H}$  be a hypergraph collection on vertex set  $[n]$  with  $|\mathbf{H}| \leq \beta n$  and  $\delta_d(\mathbf{H}) \geq (\delta + \alpha)n^{k-d}$ . Moreover, suppose  $e(\mathcal{A})$  divides  $|\mathbf{H}|$ . Then  $\mathbf{H}$  contains a transversal which consists of  $|\mathbf{H}|/e(\mathcal{A})$  vertex-disjoint copies of  $\mathcal{A}$ .

We remark that the property **Fac** easily holds when  $\mathcal{A}$  consists of a single edge, as stated in the following observation.

**Observation 5.1.5.** Let  $k \in \mathbb{N}$ ,  $d \in [k - 1]$  and  $\mathcal{A}$  be a  $k$ -uniform edge. Then, for any  $\delta > 0$ , property **Fac** holds for  $\mathcal{A}$  (with respect to minimum  $d$ -degree).

*Proof of Observation 5.1.5.* Let  $\delta, \alpha > 0$ , set  $\beta_0 = \alpha/(2k)$ , and let  $\beta \leq \beta_0$ . Let  $\mathbf{H}$  be a hypergraph collection on  $[n]$  with  $\delta_d(\mathbf{H}) \geq (\delta + \alpha)n^{k-d}$  and  $|\mathbf{H}| \leq \beta n$ . Suppose that we have found  $s < |\mathbf{H}|$  vertex-disjoint copies of  $\mathcal{A}$  on  $[n]$  together with a rainbow colouring (using  $s$  colours), and let  $S$  be the vertex set spanned by those copies. Observe that  $|S| = sk \leq \alpha n/2$ . Let  $H$  be a hypergraph in  $\mathbf{H}$  not yet used, then by Observation 2.6.1,  $H[V \setminus S]$  still contains an edge and thus a copy of  $\mathcal{A}$ . Hence we can extend the collection of copies of  $\mathcal{A}$  in a rainbow fashion. This proves the observation.  $\square$

## 5.2 Main theorem

We have now introduced all the necessary terminology to state our main theorem. Recall that, following Definition 1.2.2, the uncoloured minimum  $d$ -degree threshold for a Hamilton  $\mathcal{A}$ -cycle, with  $\mathcal{A}$  being a link of uniformity  $k$ , is the smallest real number  $\delta = \delta(\mathcal{A}, d)$  with the following property. For any  $\alpha > 0$ , there exists  $n_0 \in \mathbb{N}$  so that for any  $n \in (m - \ell)\mathbb{N}$



with  $n \geq n_0$ , every  $k$ -uniform hypergraph  $\mathcal{H}$  on  $n$  vertices with  $\delta_d(\mathcal{H}) \geq (\delta + \alpha)n^{k-d}$  contains a Hamilton  $\mathcal{A}$ -cycle.

**Theorem 5.2.1.** *Let  $k, \ell, m \in \mathbb{N}$  with  $\ell \leq m$ ,  $\mathcal{A}$  be an  $\ell$ -link of order  $m$  and uniformity  $k$ , and  $d \in [k - 1]$ . Let  $\delta = \delta(\mathcal{A}, d)$  be the uncoloured minimum  $d$ -degree threshold for the containment of a Hamilton  $\mathcal{A}$ -cycle and suppose that  $\mathcal{A}$  is  $(\delta_0, d)$ -good for some  $\delta_0 \geq \delta$ . Then, for any  $\alpha > 0$ , there exists  $n_0 \in \mathbb{N}$  so that for any  $n \in (m - \ell)\mathbb{N}$  with  $n \geq n_0$ , the following holds.*

*Let  $\mathbf{H}$  be a  $k$ -uniform hypergraph collection on vertex set  $[n]$  with  $|\mathbf{H}| = \frac{e(\mathcal{A}) - e(\mathcal{A}_s)}{m - \ell}n$  and  $\delta_d(\mathbf{H}) \geq (\delta_0 + \alpha)n^{k-d}$ . Then  $\mathbf{H}$  contains a transversal copy of a Hamilton  $\mathcal{A}$ -cycle.*

Observe that the quantity  $\frac{n}{m - \ell}(e(\mathcal{A}) - e(\mathcal{A}_s))$  appearing in Theorem 5.2.1 is precisely the number of edges in a Hamilton  $\mathcal{A}$ -cycle covering  $n$  vertices. Therefore, it is also the size of a hypergraph collection on  $[n]$  containing a transversal copy of a Hamilton  $\mathcal{A}$ -cycle. Moreover, if Theorem 5.2.1 holds with  $\delta_0 = \delta$ , then the family of Hamilton  $\mathcal{A}$ -cycles is  $d$ -colour-blind.

### 5.3 Proof overview

As previously mentioned, the framework of the proof of our main result borrows a lot from the work of Montgomery, Műyesser, and Pehova [90]. We will now attempt to give a self-contained account of the main ideas of our proof strategy. For the purposes of the proof sketch, it will be conceptually (and notationally) simpler to imagine that we are trying to prove that the family of (2-uniform) Hamilton cycles is colour-blind. Observe that a Hamilton cycle is an  $\mathcal{A}$ -cycle with  $\mathcal{A}$  being an edge.

**Proposition 5.3.1** (Theorem 2 in [34]). *For any  $\alpha > 0$ , there exists  $n_0 \in \mathbb{N}$  such that the following holds. Let  $\mathbf{G}$  be a graph collection on vertex set  $[n]$  with  $|\mathbf{G}| = n$  and  $\delta(\mathbf{G}) \geq (1/2 + \alpha)n$ . Then  $\mathbf{G}$  contains a transversal copy of a Hamilton cycle.*

**Colour absorption.** The basic premise of our approach, which is shared with [90], is that Proposition 5.3.1 becomes significantly easier to prove if we assume that  $|\mathbf{G}| = (1 + o(1))n$ , that is, if we have a bit more colours than we need to find a rainbow Hamilton cycle on  $n$  vertices. Thus, the starting goal of the proof is to somehow simulate having access to more colours than we need, while still starting with a graph collection of size exactly  $n$ . The way we achieve this is through the following lemma, which follows in a long tradition of absorption based ideas. Before stating it, we introduce the following terminology. Given a hypergraph collection  $\mathbf{H}$ , when we say that a hypergraph  $H \subset \cup_{i \in [m]} H_i$  is *uncoloured*, we mean that a colouring has not yet been assigned.

**Lemma 5.3.2.** *Let  $d, k, n \in \mathbb{N}$ ,  $1/n \ll \gamma \ll \beta \ll \alpha$  and  $\delta \geq 0$ . Let  $\mathcal{F}$  be a  $k$ -uniform hypergraph with  $e(\mathcal{F}) = \beta n$  and suppose that any  $n$ -vertex  $k$ -uniform hypergraph with*

minimum  $d$ -degree at least  $\delta n^{k-d}$  contains a copy of  $\mathcal{F}$ . Let  $\mathbf{H}$  be a  $k$ -uniform hypergraph collection on  $[n]$  with  $\delta_d(\mathbf{H}) \geq (\delta + \alpha)n^{k-d}$  and  $|\mathbf{H}| = m$  with  $m \geq \alpha n$ .

Then, there is an uncoloured copy  $\mathcal{S}$  of  $\mathcal{F}$  in  $\cup_{i \in [m]} H_i$  and disjoint sets  $A, C \subset [m]$ , with  $|A| = e(\mathcal{F}) - \gamma n$  and  $|C| \geq 10\beta m$  such that the following property holds. Given any subset  $B \subset C$  with  $|B| = \gamma n$ , there is a rainbow colouring of  $\mathcal{S}$  in  $\mathbf{H}$  using colours in  $A \cup B$ .

We remark that Lemma 5.3.2 is the hypergraph analogue of Lemma 3.4 from [90]. For the sake of completeness, we give its proof below. Before proving it, we need the following lemma from [90].

**Lemma 5.3.3** (Lemma 3.3 in [90]). *Let  $\alpha \in (0, 1)$  and let  $\ell, m, n \geq 1$  be integers satisfying  $\ell \leq \alpha^7 m / 10^5$  and  $\alpha^2 n \geq 8m$ . Let  $K$  be a bipartite graph on vertex classes  $A$  and  $B$  such that  $|A| = m$ ,  $|B| = n$  and, for each  $v \in A$ ,  $d_K(v) \geq \alpha n$ .*

*Then, there are disjoint subsets  $B_0, B_1 \subset B$  with  $|B_0| = m - \ell$  and  $|B_1| \geq \alpha^7 n / 10^5$ , and the following property. Given any set  $U \subset B_1$  of size  $\ell$ , there is a perfect matching between  $A$  and  $B_0 \cup U$  in  $K$ .*

*Proof of Lemma 5.3.2.* Let  $\mathcal{H}$  be the  $k$ -uniform hypergraph with vertex set  $[n]$ , where  $e$  is an edge of  $\mathcal{H}$  exactly when  $e \in E(H_i)$  for at least  $\alpha m$  values of  $i \in [m]$ . Then, by Proposition 2.6.2,  $\delta_d(\mathcal{H}) \geq \delta n^{k-d}$  and, therefore,  $\mathcal{H}$  contains a copy of  $\mathcal{F}$ , which we denote by  $\mathcal{S}$ . Observe that  $\mathcal{S}$  is an uncoloured copy of  $\mathcal{F}$  in  $\cup_{i \in [m]} H_i$ .

Let  $K$  be the bipartite graph with vertex classes  $E(\mathcal{S})$  and  $[m]$ , where  $ei$  is an edge of  $K$  exactly if  $e \in H_i$ . Note that, since each  $e \in E(\mathcal{S})$  is also an edge of  $\mathcal{H}$ , we have that  $d_K(e) \geq \alpha m$ . Then, as  $\gamma \ll \beta \ll \alpha$ , by Lemma 5.3.3 with  $\ell = \gamma n$ ,  $m = \beta n$ ,  $n = m$ , there are disjoint sets  $A, C \subset [m]$  with  $|A| = e(\mathcal{F}) - \gamma n$  and  $|C| \geq 10\beta m$ , such that, for any set  $B \subset C$  of size  $\gamma n$  there is a perfect matching between  $E(\mathcal{S})$  and  $A \cup B$ . Note that for such a matching  $M$ , the function  $\phi : E(\mathcal{S}) \rightarrow A \cup B$ , defined by  $e\phi(e) \in M$  for each  $e \in E(\mathcal{S})$ , gives a rainbow colouring of  $\mathcal{S}$  in  $\mathbf{H}$  using colours in  $A \cup B$ , as required.  $\square$

**Completing the cycle.** Lemma 5.3.2 provides us with a lot of flexibility, by finding a small subgraph that admits a rainbow colouring in many different ways. To prove Proposition 5.3.1, we will also need the following proposition.

**Proposition 5.3.4.** *Let  $1/n \ll \zeta \ll \kappa, \alpha$ . Let  $\mathbf{G}$  be a graph collection on  $[n]$  with  $|\mathbf{G}| = (1 + \kappa - \zeta)n$  and  $\delta(\mathbf{G}) \geq (1/2 + \alpha)n$ . Let  $a, b \in [n]$  be distinct vertices. Then,  $\mathbf{G}$  contains a rainbow Hamilton path with  $a$  and  $b$  as its endpoints, using every colour  $G_i$  with  $i \in [(1 - \zeta)n]$ .*

Proposition 5.3.4, in combination with Lemma 5.3.2, gives a proof of Proposition 5.3.1.

*Sketch of Proposition 5.3.1.* Let  $C$  denote the set of the  $n$  colours. Apply Lemma 5.3.2 with  $\mathcal{F}$  being a path of length  $\beta n$  (and some constant  $\gamma \ll \beta$ ). This gives a path  $\mathcal{S}$  in  $\mathbf{G}$  and

colour sets  $A$  and  $C$ . Let  $a$  and  $b$  be the endpoints of  $\mathcal{S}$ . Set  $\mathbf{G}'$  to be the graph collection obtained by restricting  $\mathbf{G}$  to the vertex set  $([n] \setminus V(\mathcal{S})) \cup \{a, b\}$  and colour set  $C \setminus A$ . Apply Proposition 5.3.4, labelling the colours in  $\mathbf{G}'$  so that the first  $(1 - \zeta)n$  colours correspond to those in  $C \setminus (A \cup C)$ . This way, we extend  $\mathcal{S}$  to a Hamilton cycle  $\mathcal{H}$ . While the edges in  $\mathcal{S}$  are still uncoloured, those in  $\mathcal{H} \setminus \mathcal{S}$  have been assigned a colour set using all colours in  $C \setminus (A \cup C)$  and exactly  $|C| - \gamma n$  colours from  $C$ . Using the absorption property of  $\mathcal{S}$ , the path  $\mathcal{S}$  can be given a colouring using all the colours in  $A$  and the remainder colours in  $C$ , thereby giving  $\mathcal{H}$  a rainbow colouring, as desired.  $\square$

Unfortunately, due to the technicalities present in the statement, Proposition 5.3.4 is far from trivial to show. Most of the novelty in the proof of our main theorem is the way we approach Proposition 5.3.4 for arbitrary  $\mathcal{A}$ -chains satisfying **Ab**, **Con**, and **Fac**. We now proceed to explain briefly how we achieve this, and how the three properties come in handy.

Firstly, in the setting of Proposition 5.3.4, it is quite easy to find a few rainbow paths using most of the colours from the set  $[(1 - \zeta)n]$ . Below is a formal statement of a version of this for arbitrary  $\mathcal{A}$ -chains, where we remark that  $\left(\frac{e(\mathcal{A}) - e(\mathcal{A}_s)}{m - \ell}\right)n$  is the number of edges of an  $\mathcal{A}$ -cycle on  $n$  vertices.

**Lemma 5.3.5.** *Let  $1/n \ll 1/T \ll \omega, \alpha$ . Let  $\mathcal{A}$  be an  $\ell$ -link of order  $m$  and uniformity  $k$ , and  $d \in [k - 1]$ . Let  $\delta$  be the minimum  $d$ -degree threshold for the containment of a Hamilton  $\mathcal{A}$ -cycle. Let  $\mathbf{H}$  be a  $k$ -uniform hypergraph collection on  $[n]$  with  $\delta_d(\mathbf{H}) \geq (\delta + \alpha)n^{k-d}$ , and suppose that  $|\mathbf{H}| \geq \left(\frac{e(\mathcal{A}) - e(\mathcal{A}_s)}{m - \ell}\right)n$ . Then  $\mathbf{H}$  contains a rainbow collection of  $T$ -many pairwise vertex-disjoint  $\mathcal{A}$ -chains covering all but at most  $\omega n$  vertices of  $\mathbf{H}$ .*

*Proof.* Choose  $\omega, T$  such that Lemma 2.6.3 holds with  $\beta = (1 - \omega/2)/T$ , and set  $t = \left(\frac{e(\mathcal{A}) - e(\mathcal{A}_s)}{m - \ell}\right)$ . Let  $\mathbf{H}$  be a  $k$ -uniform hypergraph collection on  $[n]$  with  $\delta_d(\mathbf{H}) \geq (\delta + \alpha)n^{k-d}$  and  $|\mathbf{H}| \geq tn$ .

By Lemma 2.6.3 applied with  $\beta = (1 - \omega/2)/T$ , there exists a partition of  $[n]$  into  $V_1, \dots, V_T, V_{T+1}$  with  $|V_1| = \dots = |V_T| = (1 - \omega/2)n/T$  and  $|V_{T+1}| = \omega n/2$ , such that for any  $1 \leq i \leq T + 1$  and any hypergraph  $\mathcal{H}$  of the collection  $\mathbf{H}$ , it holds that  $\delta_d(\mathcal{H}[V_i]) \geq (\delta + \alpha/2)|V_i|^{k-d}$ . We claim that we can greedily cover all but at most  $m \cdot T$  vertices of  $V_1, \dots, V_T$  with a rainbow collection of  $T$ -many pairwise vertex-disjoint  $\mathcal{A}$ -chains  $\mathcal{A}_1, \dots, \mathcal{A}_T$ , such that  $\mathcal{A}_i$  covers all but at most  $m$  vertices of the set  $V_i$  for each  $i \in [T]$ . Suppose we were able to do so for the sets  $V_1, \dots, V_i$  for some  $1 \leq i < T$ . Then the number of colours used so far is at most  $i \cdot (t|V_1|)$  and thus there are at least  $tn - (T - 1)t|V_1| = tn \frac{(T-1)\omega+2}{2T}$  available colours. Let  $C$  be the set of such colours. Observe that a rainbow  $\mathcal{A}$ -chain covering the vertices of  $V_{i+1}$  uses no more than  $t(1 - \omega)n/T = \eta|C|$  colours, where  $\eta = \frac{2-\omega}{(T-1)\omega+2} \leq \frac{\alpha}{4}$ , where we used  $1/T \ll \omega, \alpha$  for the last inequality. Let  $\mathcal{K}$  be the  $k$ -uniform hypergraph with vertex set  $V_{i+1}$ , where  $e$  is an edge of  $\mathcal{K}$  if  $e \in E(H_i)$  for at least  $\eta|C|$  colours  $i \in C$ . Then by Proposition 2.6.2, we have  $\delta_d(\mathcal{K}) \geq (\delta + \alpha/2 - \eta)|V_{i+1}|^{k-d} \geq (\delta + \alpha/4)|V_{i+1}|^{k-d}$ ,

where we used  $\eta \leq \alpha/4$  for the last inequality. Therefore  $\mathcal{K}$  contains a copy of a Hamilton  $\mathcal{A}$ -cycle, which in turn contains an  $\mathcal{A}$ -chain covering all but  $m$  vertices of  $V_{i+1}$ . Now we greedily assign colours from  $C$  to this  $\mathcal{A}$ -chain in a rainbow fashion.

This shows we can find a rainbow collection of  $T$ -many pairwise vertex-disjoint  $\mathcal{A}$ -chains covering all but at most  $m \cdot T + |V_{T+1}| \leq \omega n$  vertices of  $\mathbf{H}$ , as wanted.  $\square$

Although it is easy to use most of the colours coming from a colour set using the above result, a challenge in Proposition 5.3.4 is that we need to use all of the colours coming from the set  $[(1 - \zeta)n]$ . As we are currently concerned with the case when  $\mathcal{A}$  consists of a single edge, this will not be a major issue. Indeed, using the minimum degree condition on each of the colours, we can greedily find rainbow matchings using small colour subsets of  $[(1 - \zeta)n]$  (see Observation 5.1.5). For arbitrary  $\mathcal{A}$ , we would like to proceed in the same way; however, say when  $\mathcal{A}$  is a triangle, the situation becomes considerably more complicated. This is why the property **Fac** is built into the assumptions of the main theorem.

Our ultimate goal is to build a single  $\mathcal{A}$ -chain connecting specific ends, not just a collection of  $\mathcal{A}$ -chains. Hence, we rely on the property **Con** to connect the ends of the paths we obtained via Lemma 5.3.5 (as well as the greedy matching we found for the purpose of exhausting a specific colour set). An issue is that **Con** is an uncoloured property, whereas we would like to connect these ends in a rainbow manner. Here we rely on the trick offered by Proposition 2.6.2, which states that in hypergraph collections where each hypergraph has good minimum  $d$ -degree conditions, we can pass down to an auxiliary hypergraph  $\mathcal{K}$  which also has good minimum  $d$ -degree conditions. An edge appears in  $\mathcal{K}$  if and only if that edge has  $\Omega(n)$  many colours in the original hypergraph collection. We can use the property **Con** on  $\mathcal{K}$  to connect ends via short uncoloured paths, and later assign greedily one of the many available colours to the edges on this path.

As is the case with many absorption-based arguments, the short connecting paths we find will be contained in a pre-selected random set. After all the connections are made, there will remain many unused vertices inside this random set. To include these vertices inside a path, we use the property **Ab**. Similarly to **Con**, property **Ab** is an uncoloured property, but we can use again the trick of passing down to an appropriately chosen auxiliary graph.

## 5.4 Proof of main theorem

*Proof of Theorem 5.2.1.* Let  $k, \ell, m \in \mathbb{N}$  with  $\ell \leq m$ ,  $\mathcal{A}$  be an  $\ell$ -link of order  $m$  and uniformity  $k$ , and  $d \in [k - 1]$ . Let  $\delta = \delta(\mathcal{A}, d)$  be the minimum  $d$ -degree threshold for the containment of a Hamilton  $\mathcal{A}$ -cycle, and suppose that  $\mathcal{A}$  is  $(\delta_0, d)$ -good for some  $\delta_0 \geq \delta$ . In the following, the constant implicit in any  $O(\cdot)$  only depends on  $\mathcal{A}$  and, therefore, can be bounded in terms of  $m$ .

**Constants.** Let  $\alpha > 0$ , let  $c$  be given by **Con** with  $\alpha/10$  and let  $\beta_0$  be given by **Fac** applied with  $\alpha/6$ . Choose  $\beta < \beta_0$  such that  $0 < \beta \ll \alpha, 1/c, 1/m$ . Next choose  $\rho$  and  $\gamma$  such that  $0 < \gamma \ll \rho \ll \beta$  and the hierarchy in Lemma 5.3.2 is satisfied with  $\gamma, \beta, \alpha$ . Let  $\tau$  and  $\eta$  be given by **Ab** with  $\alpha = \rho$ , so that we have  $0 < \tau, \eta \leq \rho$ . Now choose  $T \in \mathbb{N}$  and  $\omega, \nu > 0$  with  $1/T \ll \omega \ll \nu \ll \eta$  so that the hierarchy in Lemma 5.3.5 is satisfied with  $T, \omega, \alpha$ . Finally, let  $n \in (m - \ell)\mathbb{N}$  be such that  $n \gg T$ , and  $n \gg n_0$  for any of the  $n_0$  coming from the applications of **Con**, **Fac** and **Ab** above. Without loss of generality we assume that  $\beta n$  is an integer and that there exists an  $\mathcal{A}$ -chain on  $\beta n$  edges. We summarise the dependency between the parameters as follows

$$1/n \ll 1/T \ll \omega \ll \nu \ll \eta, \tau, \gamma \ll \rho \ll \beta \ll \alpha, 1/c, 1/m.$$

**Set-up.** Let  $\mathbf{H}$  be a  $k$ -uniform hypergraph collection on vertex set  $[n]$  with  $|\mathbf{H}| = \frac{e(\mathcal{A}) - e(\mathcal{A}_s)}{m - \ell} n$  and  $\delta_d(\mathbf{H}) \geq (\delta + \alpha)n^{k-d}$ . Set  $t = \frac{e(\mathcal{A}) - e(\mathcal{A}_s)}{m - \ell}$  so that  $tn = |\mathbf{H}|$ . We will use  $[tn]$  to refer to our set of colours. Set  $V = V(\mathbf{H})$ .

**1. Setting up the colour absorber.** Let  $\mathcal{F}$  be an  $\mathcal{A}$ -chain on  $\beta n$  edges (and thus  $\frac{\beta n - e(\mathcal{A}_s)}{e(\mathcal{A}) - e(\mathcal{A}_s)}(m - \ell) + \ell = \beta n/t + O(1)$  vertices), which exists by our choice of  $\beta$ . As the minimum degree threshold for the containment of  $\mathcal{F}$  is at most  $\delta$ , the hypotheses of Lemma 5.3.2 are satisfied for  $\mathcal{F}$  with the hypergraph collection  $\mathbf{H}$  (observe also that we may assume without loss of generality that  $|\mathbf{H}| = tn \geq \alpha n$ ). Therefore, there exist disjoint colour sets  $A, C \subseteq [tn]$  with  $|A| = e(\mathcal{F}) - \gamma n = (\beta - \gamma)n$  and  $|C| \geq 10\beta tn$ , and an uncoloured copy  $\mathcal{S}_1$  of  $\mathcal{F}$  in  $\mathbf{H}$  such that the following holds.

Given any subset  $B \subseteq C$  with  $|B| = \gamma n$ , there is a rainbow colouring of  $\mathcal{S}_1$  in  $\mathbf{H}$  using colours in  $A \cup B$ . (5.4.1)

We denote the start and the end of  $\mathcal{S}_1$  by  $\mathcal{F}_1$  and  $\mathcal{F}_2$ , respectively, and define  $S'_1 = V(\mathcal{S}_1) \setminus (V(\mathcal{F}_1) \cup V(\mathcal{F}_2))$  to be the set of all vertices of  $\mathcal{S}_1$  except those in its ends. Note that, as  $\rho \ll \beta$ , we have  $10\beta tn \geq \rho n$  and, without relabelling, we can fix  $C$  to be a subset of the original set  $C$  of size exactly  $\rho n$ . For convenience, we split the set  $C$  arbitrarily into two subsets  $C_1$  and  $C_2$ , with  $|C_1| = \gamma n/2$  (and  $|C_2| = (\rho - \gamma/2)n$ ). Our goal in the remainder of the proof, in correspondence with Proposition 5.3.4 from the proof overview, is to find a rainbow  $\mathcal{A}$ -chain, vertex-disjoint with  $S'_1$ , starting in  $\mathcal{F}_2$  and ending in  $\mathcal{F}_1$ , using all colours in  $[tn] \setminus (C \cup A)$ , and some colours from  $C$ . Note that, similarly to the setting of Proposition 5.3.4, we have  $(t - (\beta - \gamma))n$  colours available compared to  $(t - \beta)n$  edges that we need to colour.

**2. Setting up the vertex absorber.** Since  $|V(\mathcal{S}_1)| = \beta n/t + O(1) \leq \alpha n/1000$ , where we used that  $\beta \ll \alpha$  in the last inequality, we have that  $\delta_d(\mathbf{H}[V \setminus V(\mathcal{S}_1)]) \geq (\delta + 9\alpha/10)n^{k-d}$  by Observation 2.6.1. We define an auxiliary graph  $\mathcal{K}_1$  to be the  $k$ -uniform graph on vertex set  $V_1 = V \setminus V(\mathcal{S}_1)$ , where  $e$  is an edge of  $\mathcal{K}_1$  if and only if  $e \in E(\mathcal{H}_i)$  for at least

$\alpha|C_1|/2 = \alpha\gamma n/4$  values of  $i \in C_1$ . Then, using Proposition 2.6.2 on  $\mathcal{K}_1$ , we get that  $\delta_d(\mathcal{K}_1) \geq (\delta + \alpha/2)n^{k-d} \geq (\delta + \rho)n^{k-d}$ . By the choice of the constants  $\eta$  and  $\tau$  for **Ab**, we have that there exists a set  $S_2 \subseteq V_1$  of size at most  $\tau n$  such that the following property holds.

For any set  $L \subseteq V_1 \setminus S_2$  of size at most  $\eta n$  with  $|L| \in (m-\ell)\mathbb{N}$ , there exists an embedding of an  $\mathcal{A}$ -chain in  $\mathcal{K}_1$  with vertex set  $S_2 \cup L$ . Furthermore, the embedding of the start and the end of this  $\mathcal{A}$ -chain does not depend on the subset  $L$ . (5.4.2)

In particular, by taking  $L = \emptyset$  in (5.4.2), there is a copy  $\mathcal{S}_2$  of an  $\mathcal{A}$ -chain in  $\mathcal{K}_1$  with vertex set  $S_2$ . We denote its ends by  $\mathcal{G}_1$  and  $\mathcal{G}_2$ , and define  $S'_2 = V(\mathcal{S}_2) \setminus (V(\mathcal{G}_1) \cup V(\mathcal{G}_2))$ .

**3. Setting up the reservoir connector.** By Observation 2.6.1, we have that  $\delta_d(\mathbf{H}[V \setminus (S'_1 \cup S'_2)]) \geq (\delta + \alpha/2)n^{k-d}$ , where we used that  $\tau, \beta \ll \alpha$ . We define another auxiliary graph  $\mathcal{K}_2$  as the  $k$ -uniform graph on vertex set  $V_2 = V \setminus (S'_1 \cup S'_2)$ , where  $e$  is an edge of  $\mathcal{K}_2$  if and only if  $e \in E(\mathcal{H}_i)$  for at least  $\alpha|C_1|/2 = \alpha\gamma n/4$  values of  $i \in C_1$ . By Proposition 2.6.2, we know that  $\delta_d(\mathcal{K}_2) \geq (\delta + \alpha/3)n^{k-d}$ . Using Lemma 2.6.3 on  $\mathcal{K}_2$  with  $t = 2$ ,  $n_1 = \nu n$  and  $n_2 = |V_2| - n_1$ , we get a set  $R_1$  of size  $\nu n$ , such that every subset of  $d$  vertices of  $V_2$  have  $d$ -degree at least  $(\delta + \alpha/6)n_1^{k-d}$  into  $R_1$  in the graph  $\mathcal{K}_2$ . Moreover, we can assume that  $R_1$  does not contain any of the vertices in  $V(\mathcal{F}_1) \cup V(\mathcal{F}_2) \cup V(\mathcal{G}_1) \cup V(\mathcal{G}_2)$ . From Observation 2.6.1, we have that for any two vertex-disjoint copies  $\mathcal{S}$  and  $\mathcal{T}$  of  $\mathcal{A}_s$  in  $V_2$  and any  $R' \subseteq R_1$  of size  $|R'| \leq \alpha n_1/50$ , we have that the minimum  $d$ -degree in  $(R_1 \setminus R') \cup V(\mathcal{S}) \cup V(\mathcal{T})$  in  $\mathcal{K}_2$  is at least  $(\delta + \alpha/10)n_1^{k-d}$ . Then, property **Con** applied to the hypergraph  $\mathcal{K}_2[(R_1 \setminus R') \cup V(\mathcal{S}) \cup V(\mathcal{T})]$  implies the following.

For any  $R' \subseteq R_1$  of size  $|R'| \leq \alpha n_1/50$  and any two vertex-disjoint copies  $\mathcal{S}$  and  $\mathcal{T}$  of  $\mathcal{A}_s$  in  $\mathcal{K}_2[V_2 \setminus R']$ , there is an  $\mathcal{A}$ -chain of length at most  $c$  in  $(R_1 \setminus R') \cup V(\mathcal{S}) \cup V(\mathcal{T})$  in  $\mathcal{K}_2$  with start  $\mathcal{S}$  and end  $\mathcal{T}$ . (5.4.3)

**4. Setting aside a random set to balance vertices and colours.** Define  $n_0 = n - \frac{|A|+|C|}{t}$  and  $r_2$  so that the equality below holds<sup>4</sup>

$$n - |V(\mathcal{S}_1)| - |S_2| - |R_1| - r_2 = n_0.$$

In particular, we have that

$$\begin{aligned} r_2 &= \frac{|A| + |C|}{t} - \frac{\beta n}{t} - |S_2| - \nu n + O(1) \\ &= \frac{\rho - \gamma}{t} n - |S_2| - \nu n + O(1). \end{aligned}$$

<sup>4</sup>Ignoring divisibility issues,  $n_0$  represents the number of vertices an  $\mathcal{A}$ -cycle on  $tn - |A| - |C|$  edges would have.

As  $0 \leq |S_2| \leq \tau n$ ,  $\tau, \nu \ll \rho$ , and  $\gamma \ll \rho$  we have that

$$\frac{\rho}{2t}n \leq r_2 \leq \frac{\rho - \gamma}{t}n + O(1). \quad (5.4.4)$$

As  $\nu, \tau, \beta \ll \alpha$ , by Observation 2.6.1, we have that  $\delta(\mathbf{H}[V \setminus (V(\mathcal{S}_1) \cup V(\mathcal{S}_2) \cup R_1)]) \geq (\delta + \alpha/3)n^{k-d}$ . Using Lemma 2.6.3 on  $\mathbf{H}[V \setminus (V(\mathcal{S}_1) \cup V(\mathcal{S}_2) \cup R_1)]$  with  $t = 2$ , we find a subset  $R_2$  of  $V \setminus (V(\mathcal{S}_1) \cup V(\mathcal{S}_2) \cup R_1)$  of size  $r_2$ , so that every subset of  $d$  vertices of  $V \setminus (V(\mathcal{S}_1) \cup V(\mathcal{S}_2) \cup R_1)$  has  $d$ -degree  $(\delta + \alpha/6)r_2^{k-d}$  into  $R_2$  with respect to each of the hypergraphs in the collection.

**5. Covering most of the leftover vertices via  $\mathcal{A}$ -chains using almost all the colours in  $[tn] \setminus (A \cup C)$ .** Set  $V' = V \setminus (V(\mathcal{S}_1) \cup V(\mathcal{S}_2) \cup R_1 \cup R_2)$  and note  $|V'| = n_0$ . Let  $\mathbf{H}'$  be the hypergraph collection obtained by restricting  $\mathbf{H}$  to the vertex set  $V'$  and colour set  $[tn] \setminus (A \cup C)$ . Using the upper bound from (5.4.4) and that  $\nu, \tau, \rho, \beta \ll \alpha$ , we have through Observation 2.6.1 that  $\delta(\mathbf{H}') \geq (\delta + \alpha/8)n_0^{k-d}$ . Moreover, by our definition of  $n_0$ , we have  $|\mathbf{H}'| = tn - |A| - |C| = tn_0 = \left(\frac{e(\mathcal{A}) - e(\mathcal{A}_s)}{m-\ell}\right)n_0$ . Therefore  $\mathbf{H}'$  satisfies the hypotheses of Lemma 5.3.5 and we find a rainbow collection  $\{\mathcal{P}_i : i \in [T]\}$  of  $T$ -many vertex-disjoint  $\mathcal{A}$ -chains in  $\mathbf{H}'$ , covering all but a vertex subset  $V_0$  of size at most  $\omega n_0$ , and using only colours from  $[tn] \setminus (A \cup C)$ . Moreover, observe that the set of colours from  $[tn] \setminus (A \cup C)$  unused by  $\bigcup_{i \in [T]} \mathcal{P}_i$ , which we denote by  $C_0$ , has size at most  $t\omega n_0 + O(T) \leq 2t\omega n_0$ .

**6. Exhaust  $C_0$  inside  $R_2$ .** Let  $C' \subseteq C_1$  be a minimal size subset of  $C_1$  such that  $|C' \cup C_0|$  is divisible by  $e(\mathcal{A})$ . Note this can be accomplished with a subset  $C'$  satisfying  $|C'| = O(1)$  and, since  $|C_0| \leq 2t\omega n_0$ , we have that  $|C_0 \cup C'| \leq 2t\omega n_0 + O(1)$ . Let  $\mathbf{H}''$  be the hypergraph collection obtained by restricting  $\mathbf{H}$  to the vertex set  $R_2$  and colour set  $C_0 \cup C'$ . Recall that, by property of the set  $R_2$ , we have that  $\delta(\mathbf{H}'') \geq (\delta + \alpha/6)r_2^{k-d}$ . As  $1/n \ll \rho$  and  $n \gg n_0$ , we have that  $r_2$  is sufficiently large to apply **Fac** and deduce that  $\mathbf{H}''$  contains  $|C_0 \cup C'|/e(\mathcal{A})$  vertex-disjoint rainbow copies of  $\mathcal{A}$  using all of the colours in  $C_0 \cup C'$  and  $|C_0 \cup C'| \leq 3t\omega n \leq (6t^2\omega/\rho)r_2 \leq (\alpha/1000)r_2$  vertices of  $R_2$ , where we used the lower bound on  $r_2$  in (5.4.4) and  $\omega \ll \rho$ .

**7. Shrink leftover vertices in  $R_2$  via  $C_2$ .** Let  $R'_2$  be the subset of  $R_2$  consisting of those vertices unused in the previous step and set  $r'_2 = |R'_2|$ . Since  $r_2 - r'_2 \leq (\alpha/1000)r_2$  and using Observation 2.6.1, we have that  $\delta(\mathbf{H}''[R'_2]) \geq (\delta + \alpha/30)r_2'^{k-d}$ . Note that by the upper bound in (5.4.4) we have that

$$tr'_2 \leq tr_2 \leq (\rho - \gamma)n + O(1) \leq (\rho - \gamma/2)n = |C_2|.$$

Let  $\mathbf{H}'''$  be the hypergraph collection obtained by restricting  $\mathbf{H}$  to the vertex set  $R'_2$  and colour set  $C_2$ . Then  $|\mathbf{H}'''| = |C_2| \geq \left(\frac{e(\mathcal{A}) - e(\mathcal{A}_s)}{m-\ell}\right)r_2'$ . Hence, similarly to Step 5, we can apply Lemma 5.3.5 to  $\mathbf{H}'''$  in order to find a rainbow collection  $\{\mathcal{P}'_i : i \in [T]\}$  of  $T$ -many vertex-disjoint  $\mathcal{A}$ -chains (with colours coming from  $C_2$ ) in  $\mathbf{H}'''$ , covering all but a vertex

subset  $V'_0 \subseteq R'_2$  of size at most  $\omega r'_2$ .

**8. Connect everything via  $C_1$  and  $R_1$  to build an almost spanning  $\mathcal{A}$ -cycle.** We recall that we built one uncoloured  $\mathcal{A}$ -chain in each of Step 1 and 2,  $|C_0 \cup C'|/e(\mathcal{A})$  rainbow  $\mathcal{A}$ -chains in Step 6 (indeed a copy of  $\mathcal{A}$  is trivially a rainbow  $\mathcal{A}$ -chain of length 1), and  $T$  rainbow  $\mathcal{A}$ -chains in each of Step 5 and 7. Therefore, at this point there are  $2 + 2T + |C_0 \cup C'|/e(\mathcal{A}) \leq 3t\omega n$  vertex-disjoint  $\mathcal{A}$ -chains, which we will now connect to build an  $\mathcal{A}$ -cycle, using additional vertices in  $R_1$  and colours in  $C_1$ . This can be done by repeatedly invoking property (5.4.3). Indeed, suppose that the chains are labelled  $\mathcal{J}_1, \dots, \mathcal{J}_z$  where  $z \leq 3t\omega n$ . Suppose that for some  $1 \leq z' \leq z$ , we found an  $\mathcal{A}$ -chain  $\mathcal{R}$  in  $\mathcal{K}_2$  such that the following properties all hold.

- $V(\mathcal{R}) \supseteq \bigcup_{i \in [z']} V(\mathcal{J}_i)$ ;
- With  $R' = V(\mathcal{R}) \setminus \bigcup_{i \in [z']} V(\mathcal{J}_i)$ , we have  $R' \subseteq R_1$  and  $|R'| \leq ((m - \ell)c + \ell)z'$ ;
- The start of  $\mathcal{R}$  is the start of  $\mathcal{J}_1$ , and its end is the end of  $\mathcal{J}_{z'}$ .

We remark that for  $z' = 1$ , the  $\mathcal{A}$ -chain  $\mathcal{R} = \mathcal{J}_1$  has the above properties. As  $|R'| \leq \alpha n_1/50 = \alpha \nu n/50$  (using that  $\omega \ll \alpha, 1/c, 1/m$ ), property (5.4.3) applies to show that there is an  $\mathcal{A}$ -chain in  $\mathcal{K}_2$  of length at most  $c$  starting in the end of  $\mathcal{J}_{z'}$  and ending in the start of  $\mathcal{J}_{z'+1}$  (where  $z' + 1 = 1$  if  $z' = z$ ), which uses vertices of  $R_1 \setminus R'$  and these shared ends. Observe that these chain uses less than  $(m - \ell)c + \ell$  vertices of  $R_1 \setminus R'$ . This shows that  $\mathcal{R}$  can be extended to satisfy the above properties with respect to  $z' + 1$ . Inductively, we obtain an  $\mathcal{A}$ -cycle  $C_0$  in  $\mathcal{K}_2$  covering  $\bigcup_{i \in [z]} \mathcal{J}_i$ , and we denote by  $R'_1$  the vertices from  $R_1$  unused by  $C_0$ .

Consider the set of edges of  $C_0$  not contained in some  $\mathcal{J}_i$ , i.e. the edges we have found in the previous steps to connect the various  $\mathcal{J}_i$ 's. Note that there are at most  $((m - \ell)c + \ell)zt \leq \alpha \gamma n/1000$  such edges, where we used  $\omega \ll \gamma, \alpha, 1/c, 1/m$ . Moreover, each such edge belongs to  $\mathcal{K}_2$  and thus has at least  $\alpha|C_1|/2 = \alpha \gamma n/4$  colours coming from  $C_1$ . Therefore we can greedily assign a distinct colour of  $C_1$  to each such edge.

**9. Absorb the leftover vertices.** Note that  $C_0$  covers everything in  $V$ , except the sets  $V_0, V'_0$ , and  $R'_1$  which are leftover from Steps 5, 7, and 8, respectively. Note that  $|R'_1| + |V_0| + |V'_0| \leq \nu n + \omega n_0 + \omega r'_2 \leq (\nu + 2\omega)n \leq \eta n$ , where we used that  $\omega, \nu \ll \eta$ . Therefore, by (5.4.2), there exists an embedding of an  $\mathcal{A}$ -chain  $\mathcal{S}'_2$  in  $\mathcal{K}_1$  with vertex set  $S_2 \cup R'_1 \cup V_0 \cup V'_0$ , and with the same ends as the  $\mathcal{A}$ -chain  $\mathcal{S}_2$ . As in the previous step, we can then colour the edges of  $\mathcal{S}'_2$  in a rainbow fashion, by assigning colours still available in  $C_1$ . This is possible as  $|V(\mathcal{S}'_2)| \leq (\tau + \eta)n$  and thus there are at most  $t(\tau + \eta)n \leq \alpha \gamma n/10$  new edges, where we used  $\tau, \eta \ll \gamma, \alpha$ . Moreover these edges belong to the hypergraph  $\mathcal{K}_1$  and appear in at least  $\alpha \gamma n/4$  colours in  $C_1$ , while we only used at most  $\alpha \gamma n/1000$  colours from  $C_1$  in the previous step. Therefore there are at least  $\alpha \gamma n/8$  available colours for each edge, and we can greedily assign distinct colours.



**10. Assign a colouring to  $\mathcal{S}_1$ .** Observe that we now have a Hamilton  $\mathcal{A}$ -cycle that is rainbow except for  $\mathcal{S}_1$ , which is still uncoloured. Moreover, we have used some colours in  $C$ , together with all colours outside  $A \cup C$ , but we have not used any of the colours in  $A$ . Therefore the unused colours must be those in  $A$  together with a subset  $B \subseteq C$  of size  $\gamma n$ . We can then assign colours to  $\mathcal{S}_1$  in a rainbow fashion by property (5.4.1). This completes the rainbow embedding and finishes the proof.  $\square$

## 5.5 Applications

In this section we discuss some applications of our main theorem and, in particular, we prove Theorem 1.2.5. The proofs of the statements of Theorem 1.2.5 all follow the same strategy. Suppose we want to prove  $d$ -colour-blindness of a family  $\mathcal{F}$ . We first identify a link  $\mathcal{A}$  such that each member of  $\mathcal{F}$  is an  $\mathcal{A}$ -cycle. Then we show that  $\mathcal{A}$  is  $(\delta, d)$ -good, with  $\delta$  being the uncoloured minimum degree threshold for the family  $\mathcal{F}$ . Once this is done, the  $d$ -colour-blindness of  $\mathcal{F}$  is a consequence of Theorem 5.2.1. We will give a full proof of the statement (A) of Theorem 1.2.5 with  $r = 2$ , while we will only sketch how to prove properties **Ab**, **Con**, and **Fac** for the statement (A) with  $r > 2$  and the statement (B). The reader can then easily complete a full proof, by mimicking the one given for the square of Hamilton cycles.

### 5.5.1 Powers of Hamilton cycles

The (uncoloured) minimum degree threshold for the containment of the  $r$ -th power of a Hamilton cycle in a 2-uniform graph was conjectured to be  $\frac{r}{r+1}n$  by Pósa (for  $r = 2$ ) and Seymour (for larger  $r$ ). This was proved by Komlós, Sárkozy, and Szemerédi [73, 74], using the regularity method and the Blow-Up Lemma. Later, Levitt, Sárkozy, and Szemerédi [82] obtained a proof for the case  $r = 2$  that avoids the regularity lemma and is instead based on the absorption method. More recently, Pavez-Signé, Sanhueza-Matamala, and Stein [94] generalised this to  $r \geq 2$ , while studying the hypergraph version of the problem. Both of these fit our framework and allow us to obtain part (A) of Theorem 1.2.5. We will first focus on the case  $r = 2$ , which we will use as a more detailed example and we first observe that it can be reformulated as follows.

**Theorem 5.5.1** (Rainbow version of Pósa’s conjecture). *For any  $\alpha > 0$  there exists  $n_0$  such that for  $n \geq n_0$  the following holds. Any graph collection  $\mathbf{G}$  on vertex set  $[n]$  with  $\delta(\mathbf{G}) \geq (2/3 + \alpha)n$  contains a transversal copy of the square of a Hamilton cycle.*

As mentioned above, in order to prove Theorem 5.5.1, it is enough to show that the square of a cycle is an  $\mathcal{A}$ -cycle for a suitable choice of a  $(2/3, 1)$ -good link  $\mathcal{A}$ . Towards that goal, we let  $\mathcal{A}$  be the 2-link coming from an arbitrary ordering of  $K_3$  (see Figure 5.2) and we prove that such  $\mathcal{A}$  is indeed  $(2/3, 1)$ -good. The properties **Ab** and **Con** for  $\mathcal{A}$  follow from

the proof of the (uncoloured) Pósa conjecture in [82]. In that proof, the authors give an exact version of the uncoloured threshold, by distinguishing an extremal and a non-extremal case. They say that a graph is extremal if it has two (not necessarily disjoint) sets each of size roughly  $n/3$  with few edges in between. However, for any  $\alpha > 0$ , a graph  $G$  with  $\delta(G) \geq (2/3 + \alpha)n$  cannot be extremal, thus we can use all lemmas from [82] dealing with the non-extremal case. We summarise the statements we use from [82] as follows.

**Theorem 5.5.2** (Lemma 3, Lemma 5, and Theorem 1 in [82]). *For any  $\alpha > 0$  there exists  $n_0$  such that for  $n \geq n_0$  the following holds for any  $n$ -vertex graph with minimum degree  $\delta(G) \geq (2/3 + \alpha)n$ .*

- (P1) *For any two disjoint ordered edges  $(a, b)$  and  $(c, d)$  there is a square of a path of length at most  $10\alpha^{-4}$ , with end-tuples  $(a, b)$  and  $(c, d)$ .*
- (P2) *There exists the square of a path  $P$  of length at most  $\alpha^9 n$  such that for every subset  $L \subseteq V(G) \setminus V(P)$  there exists a square of a path  $P_L$  with  $V(P_L) = V(P) \cup L$  that has the same end-tuples as  $P$ .*
- (P3) *There exists the square of a Hamilton cycle in  $G$ .*

Finally the property **Fac** for  $\mathcal{A}$  follows as a special case of a theorem in [90].

**Theorem 5.5.3** (Theorem 1.3 in [90]). *For any integer  $r \geq 1$  and any  $\alpha > 0$ , there exists  $n_0 \in \mathbb{N}$  such that for  $n \geq n_0$  the following holds. Any graph collection  $\mathbf{G}$  on  $[n]$  with  $\delta(\mathbf{G}) \geq \left(\frac{r}{r+1} + \alpha\right)n$  contains a transversal copy of a  $K_{r+1}$ -factor.*

We are now ready to give a full proof of Theorem 5.5.1.

*Proof of Theorem 5.5.1.* Let  $\alpha > 0$  and  $\mathcal{A}$  be the 2-link of order 3 and uniformity 2 coming from an arbitrary ordering of  $K_3$ . Note that an  $\mathcal{A}$ -chain is the square of a path (see Figure 5.2) and an  $\mathcal{A}$ -cycle is the square of a cycle.

The minimum degree threshold for a Hamilton  $\mathcal{A}$ -cycle is  $\delta = \delta(\mathcal{A}, 1) = 2/3$  by (P3) of Theorem 5.5.2<sup>5</sup>. Let  $n_0$  be large enough for Theorem 5.5.2 and 5.5.3 to hold. Then  $\mathcal{A}$  has property **Ab** with  $\tau = \alpha^9$  and  $\eta = \alpha^{20}$  by (P2) of Theorem 5.5.2, and it has property **Con** with  $C = 10\alpha^{-4}$  by (P1) of Theorem 5.5.2. Moreover,  $\mathcal{A}$  has property **Fac** with  $\beta_0 = 1$  by Theorem 5.5.3 (with  $r = 2$ ). Therefore  $\mathcal{A}$  is  $(\delta, d)$ -good.

Now let  $\mathbf{G}$  be a graph collection on  $[n]$  with  $\delta(\mathbf{G}) \geq (2/3 + \alpha)n$  with  $n \geq n_0$ . Then, by Theorem 5.2.1, there exists a rainbow Hamilton  $\mathcal{A}$ -cycle in  $\mathbf{G}$ , i.e. a transversal copy of the square of a Hamilton cycle, as desired.  $\square$

<sup>5</sup>The constant  $2/3$  is in fact best possible, as the complete tripartite graph with parts of size  $n/3 - 1$ ,  $n/3$ , and  $n/3 + 1$  has minimum degree  $2n/3 - 1$  and does not even contain a  $K_3$ -factor.

To obtain part (A) of Theorem 1.2.5 for  $r > 2$  we can proceed exactly as for  $r = 2$ , using the statements in [94]. However, these statements do not readily match our setup as those given in Theorem 5.5.2. Nevertheless, Lemma 4.3 in [94] implies property **Con** and it is straightforward to check that together with Lemma 7.2 in [94] this also gives property **Ab**. Indeed, Lemma 7.2 in [94] states that if  $G$  is a graph with  $\delta(G) \geq (r/(r+1) + \alpha)n$  and  $n$  is large enough, then there is a small set of pairwise vertex-disjoint  $r$ -th powers of short paths, such that every vertex of  $G$  can be absorbed into many of them (into the  $r$ -th power of a path). These paths can then be connected into the  $r$ -th power of a single path to fulfil property **Ab** (c.f. Step 1 of the proof of Theorem 1.1 in [94] for more details). As property **Fac** still holds by Theorem 5.5.3, we have that  $\mathcal{A}$  is  $(\delta, 1)$ -good, for  $\mathcal{A}$  being the  $(r-1)$ -link of order  $r$  and uniformity 2 coming from an arbitrary ordering of  $K_r$  and  $\delta = \delta(\mathcal{A}, 1) = r/(r+1)$ . The result follows by Theorem 5.2.1.

### 5.5.2 Hamilton $\ell$ -cycles in $k$ -uniform hypergraphs

The statements in (B) of Theorem 1.2.5 state  $d$ -colour-blindness of the family  $\mathcal{F}$  of  $k$ -uniform Hamilton  $\ell$ -cycles, for various ranges of  $d$ ,  $k$ , and  $\ell$ . Note that an  $\ell$ -cycle in a  $k$ -uniform hypergraph is an  $\mathcal{A}$ -cycle, with  $\mathcal{A}$  being the  $\ell$ -link of order  $k$  and uniformity  $k$  consisting of a single edge (see Figure 5.1). The result will follow from our main theorem, once we will have shown that such  $\mathcal{A}$  is  $(\delta, d)$ -good, with  $\delta$  being the uncoloured minimum degree threshold of the considered family  $\mathcal{F}$ .

We start by observing that, since  $\mathcal{A}$  consists of a single edge, Observation 5.1.5 guarantees that  $\mathcal{A}$  satisfies property **Fac** for any  $k \geq 3$ , and  $1 \leq \ell, d \leq k$ . The properties **Ab** and **Con** can be derived from the absorption-style proof of the uncoloured minimum degree threshold for  $\mathcal{F}$ . We summarise the precise reference for each property and each case of the statements in (B) of Theorem 1.2.5 in Table 5.1.

Family $\mathcal{F}$	Reference for $\delta_{\mathcal{F},d}$	Property <b>Ab</b>	Property <b>Con</b>	Property <b>Fac</b>
$1 < \ell < k/2$ and $d = k - 2$	Buß, Hàn, and Schacht [31]	Lemma 7 in [31]	Lemma 5 in [31]	Observation 5.1.5
	de Bastos, Mota, Schacht, Schnitzer, and Schenck [13]	Lemma 7 in [13]	Lemma 5 in [13]	
$1 \leq \ell < k/2$ and $d = k - 1$	Hàn and Schacht [61]	Lemma 5 in [61]	Lemma 6 in [61]	
$\ell = k - 1$ and $d = k - 1$	Rödl, Ruciński, and Szemerédi [99]	Lemma 2.1 in [99]	Lemma 2.4 in [99]	
$\ell = k/2$ and $k/2 < d \leq k - 1$ , with $k$ even	Hàn, Han, and Zhao [60]	Lemma 2.3 in [60]	Lemma 2.5 in [60]	

Table 5.1: References for the properties **Ab**, **Con**, and **Fac** for the families in the statement (B) of Theorem 1.2.5. The first row is split into two, as [31] deals with the case  $k = 3$  and [13] deals with the case  $k \geq 4$ .

Although some of these lemmas are not stated in the same exact form of the corresponding property, it is always straightforward to derive the properties from the lemmas.

Nevertheless, we clarify a few points. Firstly, we consider the second row of Table 5.1, where  $1 \leq \ell < k/2$  and  $d = k - 1$ . Lemma 6 in [61] states that for every integer  $k \geq 2$  and every pair of real numbers  $d, \varepsilon > 0$ , there exists an  $n_0$  such that for every  $k$ -uniform hypergraph  $\mathcal{H}$  on  $n$  vertices with  $\delta(\mathcal{H}) \geq dn$  the following holds. There is a set  $R$  of size at most  $\varepsilon n$  such that each set of  $k - 1$  vertices has degree at least  $d\varepsilon n/2$  into  $R$ . This implies property **Con** with  $c = 3$ . Indeed, given two edges  $\mathcal{S}$  and  $\mathcal{T}$  in  $\mathcal{H}$ , since  $2\ell \leq d$  and using the property of  $R$ , we can find an additional edge of  $\mathcal{H}$  and connect  $\mathcal{S}$  and  $\mathcal{T}$  into an  $\ell$ -path of length 3. Secondly, we consider the last row of Table 5.1, where  $\ell = k/2$  and  $k/2 < d \leq k - 1$  with  $k$  even. The authors of [60] prove an exact uncoloured minimum degree threshold, by distinguishing between an extremal and a non-extremal case. It is easy to see that any hypergraph  $\mathcal{H}$  with  $\delta_d(\mathcal{H}) \geq (\delta_{\mathcal{F},d} + \alpha)n^{k-d}$  is non-extremal, and thus we can use all lemmas from [60] dealing with the non-extremal case.

Statements in (B) of Theorem 1.2.5 can now be proved using the same arguments as in the proof of Theorem 5.5.1. Of course, Table 5.1 is not an exhaustive list of all Dirac-type results proven via the absorption method, rather it is only a small sample.

---

## Multistage Maker-Breaker game

---

In this chapter, we discuss our results related to the multistage Maker-Breaker game, namely Theorems 1.3.4 and 1.3.6 to 1.3.8 and Corollary 1.3.5. Before discussing more complex results, we illustrate an easy proof due to Barkey [11], which provides an exact result for the multistage duration threshold for the unbiased connectivity game (played on the edge set of the complete graph). We recall that  $C_n$  is the hypergraph on vertex set  $E(K_n)$  and with hyperedges corresponding to the edge sets of all spanning trees of  $K_n$ .

**Theorem 6.0.1** (Unbiased connectivity game [11]). *We have  $\tau(C_n, 1) = \lfloor \log_2(n) - 1 \rfloor$ .*

Theorem 6.0.1 is an application of Lehman's Theorem [81] of which we use the following formulation that can be found in e.g. [63].

**Theorem 6.0.2** (Theorem 1.1.3 in [63]). *Let  $G = (V, E)$  be a graph on  $n$  vertices which admits two edge-disjoint spanning trees. Then in the unbiased Maker-Breaker game on the edge set of  $G$ , Maker, even as a second player, has a strategy to build a connected spanning tree of  $G$  within  $n - 1$  moves.*

*Proof of Theorem 6.0.1.* The upper bound  $\tau(C_n, 1) \leq \lfloor \log_2(n) - 1 \rfloor$  is trivial. Indeed, after  $\log_2(n) - 1$  rounds, the board has fewer than  $n - 1$  edges and cannot contain a spanning tree.

The lower bound on  $\tau(C_n, 1)$  follows from Lehman's Theorem. Maker's goal is to ensure that the  $i$ -th stage is played on a board which contains  $\lfloor \frac{n}{2^i} \rfloor$  edge-disjoint spanning trees for every  $i \leq \lfloor \log_2(n) - 1 \rfloor$ . As  $\lfloor \frac{n}{2^t} \rfloor \geq 2$  for  $t = \lfloor \log_2(n) - 1 \rfloor$ , Theorem 6.0.2 then guarantees that Maker can still claim a spanning tree in round  $t$ , which yields  $\tau(C_n, 1) \geq \lfloor \log_2(n) - 1 \rfloor$ .

In order to achieve Maker's goal, notice first that it is a well known fact that the complete graph on  $n$  vertices can be edge-partitioned into  $\lfloor \frac{n}{2} \rfloor$  edge-disjoint spanning trees. Hence, the case  $i = 1$  of Maker's goal is obvious, and we can proceed by induction. Assume the board at stage  $i - 1$  has  $\lfloor \frac{n}{2^{i-1}} \rfloor$  spanning trees. Then Maker pairs them (arbitrarily) and forms  $\lfloor \frac{n}{2^i} \rfloor$  pairwise disjoint pairs of spanning trees (possibly ignoring one further spanning tree). She then plays on each pair separately using the strategy from Theorem 6.0.2, i.e. she always plays on the same pair that Breaker played on in his previous move. In this

way Maker occupies  $\lfloor \frac{n}{2^r} \rfloor$  edge-disjoint spanning trees, which then belong to the board on which the next stage is played, as wanted.  $\square$

The proofs of Theorems 1.3.4 and 1.3.6 to 1.3.8 are more involved, but they all share similar strategies, which we discuss in the next section.

## 6.1 Winning Criteria

We remark that, in order to show the equality  $\tau(\mathcal{H}, b) = T$ , we need to provide two strategies: one for Maker to claim at least one winning for at last  $T$  stages, and one for Breaker to destroy all the winning sets in at most  $T$  stages.

We start from Breaker's strategy. He wants to destroy all the winning sets in the least possible amount of stages. For that we will use the following well-known variant of Beck's Criterion [18]. It states that, under a certain Breaker's strategy, we can give an upper bound on the number of winning sets that Maker can completely occupy during one stage of a multistage Maker-Breaker game. Its proof, which can be found in [21], follows directly from the proof of that criterion.

**Theorem 6.1.1** (Beck's Criterion). *Let an integer  $b \geq 1$  and a hypergraph  $\mathcal{H} = (X, \mathcal{F})$  be given. Then, in the  $(1:b)$  Maker-Breaker game on  $\mathcal{H}$ , Breaker has a strategy which ensures that Maker occupies no more than*

$$\sum_{F \in \mathcal{F}} (1+b)^{-|F|+1}$$

*winning sets  $F \in \mathcal{F}$  completely.*

We now discuss Maker's strategy. She wants to claim at least one winning set for many stages. It might not be clear at this point, but it will often be the case that, while playing a  $(1:b)$  multistage Maker-Breaker game on  $K_n$ , Maker wants to claim at least one edge in each member of a suitable family of edge sets of  $K_n$ . We remark that this family is not the family of winning sets of the game. We provide a criterion under which Maker can achieve such a goal in Lemma 6.1.4. The methods used in the proof of Lemma 6.1.4 are similar to those in Chapters 17 and 20 of [15]. However the results in [15] only deal with uniform hypergraphs and a single stage game, and thus are not applicable in our setting, as we work with non-uniform hypergraphs and multistage games.

As a first step towards proving Lemma 6.1.4, we need to generalise a criterion of Beck [20] to biased games, which ensures that Maker can get an  $\alpha$ -fraction of each winning set in a biased Maker-Breaker game.

**Lemma 6.1.2.** *Given a hypergraph  $\mathcal{H} = (X, \mathcal{F})$ , a real  $0 \leq \alpha \leq 1$ , and an integer  $b \geq 1$ , if there exists  $\mu \in (0, 1)$  such that*

$$\sum_{F \in \mathcal{F}} \lambda_{\alpha, \mu, b}^{-|F|} < 1 \text{ with } \lambda_{\alpha, \mu, b} = (1 + \mu)^{\frac{1-\alpha}{b}} (1 - \mu)^\alpha,$$

*then Maker has a strategy to claim  $\alpha|F|$  elements of every winning set  $F \in \mathcal{F}$  in a  $(1:b)$  Maker-Breaker game.*

*Proof of Lemma 6.1.2.* Denote by  $X_i$  the elements which Maker took in the first  $i$  rounds. Denote by  $Y_{i,j}$  with  $0 \leq j \leq b$  the elements which Breaker took in the first  $i-1$  rounds plus the first  $j$  elements he took in the  $i^{\text{th}}$  round. We define the following potential for each  $F \in \mathcal{F}$ ,  $0 \leq j \leq b$ , and round  $i$ :

$$\phi_{i,j}(F) = (1 + \mu)^{\frac{1}{b}(|F \cap Y_{i,j}| - (1-\alpha)|F|)} (1 - \mu)^{|F \cap X_i| - \alpha|F|}.$$

We further define the potential of a vertex  $v \in X$  as

$$\phi_{i,j}(v) = \sum_{\substack{F \in \mathcal{F} \\ v \in F}} \phi_{i,j}(F),$$

and we let

$$\phi_{i,j} = \sum_{F \in \mathcal{F}} \phi_{i,j}(F)$$

denote the total potential of the game immediately after Breaker took his  $j^{\text{th}}$  element of round  $i$ . Note that  $\phi_{i,0}(v)$  and  $\phi_{i,0}$  then describe potentials immediately after Maker's  $i^{\text{th}}$  turn. We further let

$$\phi_{0,b} = \sum_{F \in \mathcal{F}} (1 + \mu)^{-\frac{1}{b}(1-\alpha)|F|} (1 - \mu)^{-\alpha|F|} = \sum_{F \in \mathcal{F}} \lambda_{\alpha, \mu, b}^{-|F|}$$

denote the potential at the start of the game, i.e. when no elements have been claimed yet. By assumption  $\phi_{0,b} < 1$ .

Maker's strategy in round  $i$  is to claim an element  $v \in X$  not yet claimed which maximises  $\phi_{i-1,b}(v)$ . We claim that, following this strategy, Maker can ensure that  $\phi_{i,j} < 1$  holds throughout the game. Indeed, for  $j \in [b]$ , let  $w_j$  be the  $j^{\text{th}}$  element Breaker claimed in round  $i$ . Then,  $\phi_{i-1,b}(v) \geq \phi_{i-1,b}(w_j) \geq \phi_{i,0}(w_j)$  where the first inequality holds by the maximality of  $v$ , and the second inequality holds since Maker's move can never increase the potential of any vertex. Moreover,  $\phi_{i,j}(w) \leq (1 + \mu)^{1/b} \phi_{i,j-1}(w)$  for all  $w \in X$ , as Breaker can increase the potential of any  $F \in \mathcal{F}$  at most by a factor  $(1 + \mu)^{1/b}$  when he

claims an element of  $X$ . In particular, this gives

$$\phi_{i,j-1}(w_j) \leq (1 + \mu)^{(j-1)/b} \phi_{i,0}(w_j) \leq (1 + \mu)^{(j-1)/b} \phi_{i-1,b}(v)$$

for all  $j \in [b]$ .

Therefore,

$$\begin{aligned} \phi_{i,b} &= \phi_{i-1,b} - \mu \phi_{i-1,b}(v) + ((1 + \mu)^{1/b} - 1) \sum_{j=1}^b \phi_{i,j-1}(w_j) \\ &\leq \phi_{i-1,b} - \mu \phi_{i-1,b}(v) + ((1 + \mu)^{1/b} - 1) \phi_{i-1,b}(v) \sum_{j=1}^b (1 + \mu)^{(j-1)/b} \\ &= \phi_{i-1,b} - \mu \phi_{i-1,b}(v) + \mu \phi_{i-1,b}(v) = \phi_{i-1,b}, \end{aligned}$$

where the first line uses that  $\mu \phi_{i-1,b}(v)$  describes the change of the total potential caused by Maker and  $((1 + \mu)^{1/b} - 1) \sum_{j=1}^b \phi_{i,j-1}(w_j)$  is the change caused by Breaker, and where we use the geometric sum to get the third line. Further, we have that  $\phi_{i,j} \leq \phi_{i,b}$  for all  $0 \leq j \leq b$ , because the potential can only increase when Breaker claims an element. We thus conclude that  $\phi_{i,j} \leq \phi_{i,b} \leq \phi_{0,b} < 1$  for all  $i$  and  $0 \leq j \leq b$ .

Now assume that there is some  $F \in \mathcal{F}$  such that Breaker has claimed at least  $(1 - \alpha)|F|$  elements of  $F$  after some round  $i$ . Note that this implies  $\phi_{i,b}(F) \geq 1$ , and therefore  $\phi_{i,b} \geq 1$ , which is a contradiction.  $\square$

Given any fixed winning set  $F \in \mathcal{F}$  and any Breaker's bias  $b$ , it is clear that Maker can guarantee to claim roughly a  $\frac{1}{b+1}$ -fraction of all elements of  $F$ . Our next aim towards proving Lemma 6.1.4 is to show that, under certain conditions, Maker can simultaneously ensure to get almost a  $\frac{1}{b+1}$ -fraction from each winning set  $F \in \mathcal{F}$ . The following lemma is obtained from Lemma 6.1.2 by a suitable choice of the parameters  $\alpha$  and  $\mu$ .

**Lemma 6.1.3.** *For every  $\delta \in (0, 1)$  the following holds. Let  $b, s \geq 1$  be integers,  $\mathcal{H} = (X, \mathcal{F} = \mathcal{F}_1 \cup \dots \cup \mathcal{F}_s)$  be a hypergraph, and  $k_i = \min_{F \in \mathcal{F}_i} |F|$  for  $i \in [s]$ . If  $k_i > 4\delta^{-2} \log(s|\mathcal{F}_i|)$  for every  $i \in [s]$ , then Maker has a strategy to claim at least  $(\frac{1}{b+1} - \delta)|F|$  elements of every winning set  $F \in \mathcal{F}$  in a  $(1:b)$  Maker-Breaker game.*

*Proof of Lemma 6.1.3.* With  $\mu = \delta/2$ , let  $\varepsilon > 0$  such that

$$e^{\mu^2} = (1 + \mu)^{\frac{1}{b+1} + \frac{\varepsilon}{b}} (1 - \mu)^{\frac{1}{b+1} - \varepsilon}. \quad (6.1.1)$$

Note that the existence of such  $\varepsilon$  is given by the fact that  $f(x) = (1 + \mu)^{\frac{1}{b+1} + \frac{x}{b}} (1 - \mu)^{\frac{1}{b+1} - x}$  defines a continuous function on  $\mathbb{R}$  with  $f(0) = (1 - \mu^2)^{\frac{1}{b+1}} < 1$  and  $\lim_{x \rightarrow +\infty} f(x) = +\infty$ .



Moreover, rearranging (6.1.1) gives

$$e^{\mu^2} = (1 - \mu^2)^{\frac{1}{b+1}} \left( \frac{(1 + \mu)^{\frac{1}{b}}}{1 - \mu} \right)^{\varepsilon} \Rightarrow \varepsilon = \frac{\mu^2 - \frac{1}{b+1} \log(1 - \mu^2)}{\frac{1}{b} \log(1 + \mu) - \log(1 - \mu)}.$$

Therefore, using that  $\frac{x}{1+x} < \log(1+x) < x$  holds for all  $x > -1, x \neq 0$ , we obtain

$$\varepsilon < \frac{\mu^2 + \frac{\mu^2}{(b+1)(1-\mu^2)}}{\frac{\mu}{b(1+\mu)} + \mu} < \frac{1 + \frac{2}{b+1}}{1 + \frac{1}{2b}} \cdot \frac{\mu^2}{\mu} \leq 2\mu.$$

We now apply Lemma 6.1.2 with  $\alpha = \frac{1}{b+1} - \varepsilon$ . By (6.1.1) and the choice of  $\mu$ , we get

$$\begin{aligned} \sum_{F \in \mathcal{F}} ((1 + \mu)^{\frac{1}{b+1} + \frac{\varepsilon}{b}} (1 - \mu)^{\frac{1}{b+1} - \varepsilon})^{-|F|} &= \sum_{i=1}^s \sum_{F \in \mathcal{F}_i} e^{-\mu^2 |F|} \leq \sum_{i=1}^s |\mathcal{F}_i| e^{-\frac{1}{4} \delta^2 k_i} \\ &< \sum_{i=1}^s |\mathcal{F}_i| e^{-\log(s|\mathcal{F}_i|)} = \frac{s}{s} = 1. \end{aligned}$$

Thus, Maker can claim at least

$$\alpha |F| = \left( \frac{1}{b+1} - \varepsilon \right) |F| > \left( \frac{1}{b+1} - 2\mu \right) |F| = \left( \frac{1}{b+1} - \delta \right) |F|$$

elements of every winning set  $F \in \mathcal{F}$ .  $\square$

Finally, we may apply Lemma 6.1.3 repeatedly over several stages in order to obtain that, even after some number of stages, Maker can make sure to get at least one element from every winning set.

**Lemma 6.1.4** (Multistage winning criteria). *For every  $\gamma \in (0, 1)$ , integers  $b, s \geq 1$ , and hypergraph  $\mathcal{H} = (X, \mathcal{F} = \mathcal{F}_1 \cup \dots \cup \mathcal{F}_s)$  with  $|\mathcal{F}| > 1$  the following holds. Let  $k_j = \min_{F \in \mathcal{F}_j} |F|$  for every  $j \in [s]$  and assume that*

$$\left( \frac{k_j}{\log(s|\mathcal{F}_j|)} \right)^{\gamma/2} \geq 20b \cdot \max \left\{ 1, \log_{b+1} \left( \frac{k_j}{\log(s|\mathcal{F}_j|)} \right) \right\} \quad (6.1.2)$$

for every  $j \in [s]$ . Then, in the  $(1 : b)$  Maker-Breaker multistage game on  $\mathcal{H}$ , Maker has a strategy to ensure that after  $(1 - \gamma) \min_{i \in [s]} \log_{b+1} \left( \frac{k_i}{\log(s|\mathcal{F}_i|)} \right)$  stages, she still claims at least one element in each  $F \in \mathcal{F}$ .

*Proof of Lemma 6.1.4.* Let  $t$  and  $\delta$  be defined by

$$t = (1 - \gamma) \min_{i \in [s]} \log_{b+1} \left( \frac{k_i}{\log(s|\mathcal{F}_i|)} \right) \quad \text{and} \quad \delta = 4 \max_{i \in [s]} \left( \frac{\log(s|\mathcal{F}_i|)}{k_i} \right)^{\frac{\gamma}{2}},$$

and observe that there exists some  $j \in [s]$  with

$$2b\delta t \leq 2b \cdot 4 \left( \frac{\log(s|\mathcal{F}_j|)}{k_j} \right)^{\frac{\gamma}{2}} \cdot (1-\gamma) \log_{b+1} \left( \frac{k_j}{\log(s|\mathcal{F}_j|)} \right) \stackrel{(6.1.2)}{<} \frac{1}{2}.$$

At the end of stage  $i$ , let  $\mathcal{X}^i$  denote the set of all elements of  $\mathcal{X}$  that belong to Maker, let  $\mathcal{F}_j^i = \{F \cap \mathcal{X}^i : F \in \mathcal{F}_j\}$  be the multiset of "left-overs" of the winning sets of  $\mathcal{F}_j$ , and set  $k_{i,j} = \min_{F \in \mathcal{F}_j^i} |F|$  for each  $j \in [s]$ . We aim to show that Maker can play in such a way that  $k_{t,j} \geq 1$  for all  $j \in [s]$ . To achieve that, in each stage  $i \leq t$ , we let Maker play according to the strategy of Lemma 6.1.3 with input  $\delta$  and hypergraph  $\mathcal{H}^i = (\mathcal{X}^i, \mathcal{F}^i = \mathcal{F}_1^i \cup \dots \cup \mathcal{F}_s^i)$ .

Notice that, as long as the assumptions of Lemma 6.1.3 hold at the beginning of some stage  $i \leq t$ , i.e. if  $k_{i-1,j} > 4\delta^{-2} \log(s|\mathcal{F}_j|)$  for all  $j \in [s]$ , we obtain  $k_{i,j} \geq k_{i-1,j}(\frac{1}{b+1} - \delta)$  and hence  $k_{i,j} \geq k_j(\frac{1}{b+1} - \delta)^i$  for every  $j \in [s]$ . In particular, we then conclude

$$\begin{aligned} k_{i,j} &\geq k_j \left( \frac{1}{b+1} - \delta \right)^t = k_j \left( \frac{1}{b+1} \right)^t \cdot (1 - (b+1)\delta)^t \\ &\geq \left( \frac{k_j}{\log(s|\mathcal{F}_j|)} \right)^\gamma \log(s|\mathcal{F}_j|) \cdot (1 - 2b\delta t) \geq \frac{1}{2} \left( \frac{k_j}{\log(s|\mathcal{F}_j|)} \right)^\gamma \log(s|\mathcal{F}_j|), \end{aligned}$$

where the second inequality holds since  $t \leq (1-\gamma) \log_{b+1} \left( \frac{k_j}{\log(s|\mathcal{F}_j|)} \right)$  and since  $(1-x)^t \geq 1-xt$  as long as  $x < 1$ , and the last inequality holds since  $2b\delta t < \frac{1}{2}$ . This in turn gives that

$$k_{i,j} \geq \frac{1}{2} \left( \frac{k_j}{\log(s|\mathcal{F}_j|)} \right)^\gamma \log(s|\mathcal{F}_j|) \geq 4\delta^{-2} \log(s|\mathcal{F}_j|) = 4\delta^{-2} \log(s|\mathcal{F}_j^i|),$$

where the second inequality holds by the definition of  $\delta$ . Now, this means that for the next stage  $i+1$  we can again apply Lemma 6.1.3. Inductively it follows that Maker can play in such a way that

$$k_{i,j} \geq \frac{1}{2} \left( \frac{k_j}{\log(s|\mathcal{F}_j|)} \right)^\gamma \log(s|\mathcal{F}_j|)$$

for all  $i \leq t$  and  $j \in [s]$ . In particular, using (6.1.2) again, we obtain  $k_{t,j} \geq 1$ .  $\square$

## 6.2 Proof of Theorem 1.3.4 (Hamilton cycle game)

In this section we prove Theorem 1.3.4 and Corollary 1.3.5. The upper bounds  $\tau(C_n, b), \tau(\mathcal{H}_n, b) \leq (1+o(1)) \log_{b+1}(n)$  are trivial, since after  $(1+o(1)) \log_{b+1}(n)$  stages, the board has fewer than  $n-1$  edges, and thus it can be neither Hamiltonian, nor connected. Thus it remains to provide a strategy for Maker to obtain a matching lower bound for the Hamilton cycle game, which immediately gives a lower bound for the connectivity game as well. For this section only, given any set  $S \subseteq V(G)$ , we denote the (joint) *external neighbourhood* of  $S$  by  $N_G(S) = \{u \in V(G) \setminus S : \text{there exists } x \in S \text{ with } ux \in E(G)\}$ .

We will use the following criterion for the existence of Hamilton cycles, which is obtained by choosing  $d = \log \log n$  in Theorem 1.1 in [65].

**Theorem 6.2.1** (Corollary from Theorem 1.1 in [65]). *For every large enough  $n$  the following holds. Let  $G = (V, E)$  be a graph on  $n$  vertices such that*

(P1) *for every  $S \subseteq V$ , if  $|S| \leq \frac{n}{\log n}$ , then  $|N(S)| \geq (\log \log n)|S|$ ;*

(P2) *there is an edge in  $G$  between any two disjoint subsets  $A, B \subseteq V$  such that  $|A| = |B| = \frac{n}{\log n}$ ,*

*then  $G$  is Hamiltonian.*

### 6.2.1 Maker's strategy

Choose

$$\varepsilon \in \left( \frac{4 \log \log n}{\log n}, \frac{5 \log \log n}{\log n} \right) \text{ and } \gamma = 2 \cdot \frac{\log b + \log \log_{b+1}(n) + 5}{(1 - \varepsilon) \log n - 2 \log \log n - \log 2}$$

such that  $\varepsilon^{-1} \in \mathbb{N}$ , and assume  $n$  to be large enough whenever needed. We will prove that  $\tau(\mathcal{HAM}_n, b) \geq (1 - \gamma - 2\varepsilon) \log_{b+1}(n)$ . Since  $b$  is subpolynomial, then  $\tau(\mathcal{HAM}_n, b) \geq (1 - o(1)) \log_{b+1}(n)$ , as required. Moreover observe that when  $b = 1$ , this gives  $\tau(\mathcal{HAM}_n, b) \geq \log_2(n) - \Theta(\log_2 \log(n))$  which is very close to the random graph intuition as discussed in the introduction.

In order to prove this bound, consider the family  $\mathcal{F} = \bigcup_{i=1}^s \mathcal{F}_i$  with  $s = 2/\varepsilon$  defined by

$$\mathcal{F}_j = \left\{ E_{K_n}(A, B) : A, B \subseteq V(K_n), A \cap B = \emptyset, |A| = n^{\frac{(j-1)\varepsilon}{2}}, |B| = n - \frac{1}{2}n^{\frac{(j+1)\varepsilon}{2}} \right\},$$

for  $1 \leq j \leq s - 1$ , and

$$\mathcal{F}_s = \left\{ E_{K_n}(A, B) : A, B \subseteq V(K_n), A \cap B = \emptyset, |A| = |B| = \frac{n}{\log n} \right\}.$$

Moreover set  $k_j = \min\{|F| : F \in \mathcal{F}_j\}$  for every  $j \in [s]$ .

Maker plays according to Lemma 6.1.4 with  $\gamma$  and  $s$  as defined above and hypergraph  $\mathcal{H} = (E(K_n), \mathcal{F})$ . For that we need to verify that the condition (6.1.2) holds. For  $j \in [s - 1]$ , we have  $\frac{1}{2}n^{1+\frac{(j-1)\varepsilon}{2}} \leq k_j \leq n^{1+\frac{(j-1)\varepsilon}{2}}$  and  $|\mathcal{F}_j| = \binom{n}{\frac{1}{2}n^{(j+1)\varepsilon/2}} \binom{\frac{1}{2}n^{(j+1)\varepsilon/2}}{n^{(j-1)\varepsilon/2}}$ . Hence

$$|\mathcal{F}_j| \leq n^{n^{(j+1)\varepsilon/2}} \text{ and } |\mathcal{F}_j| \geq \binom{n}{\frac{1}{2}n^{(j+1)\varepsilon/2}} \geq \left( 2n^{1-(j+1)\varepsilon/2} \right)^{\frac{1}{2}n^{(j+1)\varepsilon/2}} \geq 2^{\frac{1}{2}n^{(j+1)\varepsilon/2}}.$$

In particular, using that  $s = 2/\varepsilon \leq n$  and the bounds above on  $k_j$  and  $|F_j|$ , we have  $\frac{n^{1-\varepsilon}}{2\log^2 n} \leq \frac{k_j}{\log(s|\mathcal{F}_j|)} \leq \frac{k_j}{\log|\mathcal{F}_j|} \leq n$ , from which we can conclude (6.1.2) for  $j \in [s-1]$  as follows:

$$\begin{aligned} \left(\frac{k_j}{\log(s|\mathcal{F}_j|)}\right)^{\gamma/2} &\geq \left(\frac{n^{1-\varepsilon}}{2\log^2 n}\right)^{\gamma/2} = \exp\left[\frac{\gamma}{2} \cdot \left((1-\varepsilon)\log n - \log 2 - 2\log\log n\right)\right] \\ &= \exp(\log b + \log\log_{b+1}(n) + 5) > 20b \log_{b+1}(n) \\ &\geq 20b \max\left\{1, \log_{b+1}\left(\frac{k_j}{\log(s|\mathcal{F}_j|)}\right)\right\}. \end{aligned}$$

Similarly, for  $j = s$  we have  $k_s = \frac{n^2}{\log^2 n}$  and  $|\mathcal{F}_s| = \frac{1}{2} \binom{n}{n/\log n} \binom{n-n/\log n}{n/\log n}$ , hence  $|\mathcal{F}_s| \leq 4^n$  and  $|\mathcal{F}_s| \geq \exp\left(\frac{(2-o(1))n \log\log n}{\log n}\right)$ . In particular,  $\frac{n}{2\log^2 n} \leq \frac{k_s}{\log(s|\mathcal{F}_s|)} \leq \frac{k_s}{\log|\mathcal{F}_s|} \leq n$  from which we can conclude (6.1.2) analogously.

By Lemma 6.1.4 we now obtain that Maker can ensure to claim at least one edge in each  $F \in \mathcal{F}$  for at least

$$(1-\gamma) \log_{b+1}\left(\frac{n^{1-\varepsilon}}{2\log^2 n}\right) \geq (1-\gamma) \log_{b+1}(n^{1-2\varepsilon}) \geq (1-\gamma-2\varepsilon) \log_{b+1}(n)$$

stages, where we use the choice of  $\varepsilon$  in the first inequality.

It thus remains to check that if a graph  $G$  contains an edge from every  $F \in \mathcal{F}$ , then properties (P1) and (P2) hold. Using only  $\mathcal{F}_s$ , we immediately see that (P2) holds. For proving (P1), let  $S \subseteq V(K_n)$  with  $|S| \leq \frac{n}{\log n}$  be given, and let  $j \in [s-1]$  be largest such that  $n^{(j-1)\varepsilon/2} \leq |S|$ . If  $j \leq s-2$ , then  $|S| < n^{j\varepsilon/2}$ . Otherwise, if  $j = s-1$ , then  $|S| \leq \frac{n}{\log n} = \frac{n^{(j+1)\varepsilon/2}}{\log n}$ . In any case, choose any subset  $S' \subseteq S$  of size  $n^{(j-1)\varepsilon/2}$ , and observe that  $|N_G(S')| \geq \frac{1}{2}n^{(j+1)\varepsilon/2} - n^{(j-1)\varepsilon/2}$  since Maker claims an edge in every set  $E_{K_n}(S', B) \in \mathcal{F}_j$  with  $B \cap S' = \emptyset$  and  $|B| = n - \frac{1}{2}n^{(j+1)\varepsilon/2}$ . In particular,

$$\begin{aligned} |N_G(S)| &\geq |N_G(S')| - |N_G(S') \cap S| \geq \frac{1}{2}n^{(j+1)\varepsilon/2} - n^{(j-1)\varepsilon/2} - \frac{n^{(j+1)\varepsilon/2}}{\log n} \\ &\geq \frac{1}{4}n^{(j+1)\varepsilon/2} > (\log\log n)|S|, \end{aligned}$$

where we use the definition of  $j$  and that  $\varepsilon > \frac{4\log\log n}{\log n}$ . This proves (P1) and hence the lower bound of Theorem 1.3.4.  $\square$

### 6.3 Proof of Theorem 1.3.6 (Non-k-colourability game)

In this section we prove Theorem 1.3.6. Maker's strategy relies on Lemma 6.1.4, while Breaker's strategy relies on multiple applications of the following lemma.

**Lemma 6.3.1** (Corollary of Theorem 1.8 in [64]). *Let  $b \geq 1$  be an integer and  $G$  be the union of at most  $b + 1$  edge-disjoint forests. Then Breaker wins the  $(1:b)$  cycle game on  $G$ .*

### 6.3.1 Maker's strategy

Let  $n$  be large enough and consider the family

$$\mathcal{F} = \left\{ E_{K_n}(A) : A \subseteq V(K_n), |A| = \left\lceil \frac{n}{k} \right\rceil \right\}$$

Maker plays according to Lemma 6.1.4 with

$$\gamma = 2 \cdot \frac{\log b + \log \log_{b+1}(n) + 5}{\log n - 2 \log k - \log 4 - \log \log 2},$$

$s = 1$  and hypergraph  $\mathcal{H} = (E(K_n), \mathcal{F})$ . Note that  $\gamma = o(1)$  and  $\gamma > 0$  since  $b$  and  $k$  are subpolynomial in  $n$ . In order to apply Lemma 6.1.4, we need to check that condition (6.1.2) holds. Observe that  $|F| = \left\lceil \frac{n}{k} \right\rceil$  for every  $F \in \mathcal{F}$  and, with  $\ell = \left\lceil \frac{n}{k} \right\rceil$ , we have

$$\begin{aligned} \left( \frac{\ell}{\log |\mathcal{F}|} \right)^{\gamma/2} &\geq \left( \frac{\frac{n^2}{4k^2}}{\log(2^n)} \right)^{\gamma/2} = \exp(\log b + \log \log_{b+1}(n) + 5) \\ &> 20b \log_{b+1}(n) \geq 20b \max \left\{ 1, \log_{b+1} \left( \frac{\ell}{\log |\mathcal{F}|} \right) \right\}, \end{aligned}$$

where we use the definition of  $\gamma$  in the equality and that  $\frac{\ell}{\log(|\mathcal{F}|)} \leq n$  in the last inequality. Therefore, (6.1.2) holds and we conclude that Maker can claim an edge from each  $F \in \mathcal{F}$  for at least

$$\begin{aligned} (1 - \gamma) \log_{b+1} \left( \frac{\ell}{\log |\mathcal{F}|} \right) &\geq (1 - \gamma) \log_{b+1} \left( \frac{\frac{n^2}{4k^2}}{\log(2^n)} \right) \\ &\geq \left( 1 - \gamma - \frac{2 \log_{b+1}(2k) + \log_{b+1}(\log 2)}{\log_{b+1}(n)} \right) \log_{b+1}(n) \\ &= (1 - o(1)) \log_{b+1}(n) \end{aligned}$$

stages, where we use that  $\gamma = o(1)$  and  $k$  is subpolynomial in  $n$ .

It remains to show that if a graph  $G$  contains an edge from every  $F \in \mathcal{F}$ , then  $G$  does not admit a proper  $k$ -colouring. Observe indeed that if there is a proper  $k$ -colouring of  $G$ , then at least one colour would be assigned to at least  $\left\lceil \frac{n}{k} \right\rceil$  vertices. But  $G$  contains an edge in every set of  $\left\lceil \frac{n}{k} \right\rceil$  vertices, which is a contradiction.  $\square$

### 6.3.2 Breaker's strategy

Breaker wants to force the board to be a forest, so that it is bipartite and thus  $k$ -colourable for each  $k \geq 2$ . We explain below a strategy to achieve this in  $\log_{b+1}(n) + 1$  stages. In each stage we partition the board  $X_i = F_{i,1} \cup \dots \cup F_{i,k_i}$  into  $k_i$  edge-disjoint forests, where  $k_i$  is the smallest number such that such a partition exists. Let  $\mathcal{F}_i$  be the collection of forests from such a partition. We show that Breaker can ensure that  $k_{i+1} \leq \lceil \frac{k_i}{b+1} \rceil$ , by using the following strategy. Assume that in stage  $i$  the board  $X_i$  has  $k_i$  edge-disjoint forests. Then we split the board  $X_i$  into  $\lceil \frac{k_i}{b+1} \rceil$  edge-disjoint boards  $G_j$ , such that each forest  $F \in \mathcal{F}_i$  is contained in exactly one board  $G_j$ , and each board contains at most  $b + 1$  edge-disjoint forests  $F \in \mathcal{F}_i$ . Whenever Maker plays on some board  $G_j$ , Breaker plays on the same board according to the strategy given by Lemma 6.3.1. Thus, at the end of stage  $i$ , Maker has claimed an acyclic graph on each board  $G_j$ . We conclude that there is a partition of  $X_{i+1}$  into at most  $\lceil \frac{k_i}{b+1} \rceil$  edge-disjoint forests as wanted.

Using the fact that  $X_0 = E(K_n)$  and thus  $k_0 = \lceil \frac{n}{2} \rceil$ , we conclude that  $k_i \leq 1$  for  $i \geq \log_{b+1}(n) + 1$ , and thus the board becomes a forest.  $\square$

## 6.4 Proof of Theorem 1.3.7 ( $H$ -game)

In this section we prove Theorem 1.3.7. Maker's strategy relies on the container method (see Section 2.3) and Lemma 6.1.4. Breaker's strategy relies instead on the notion of  $K$ -collections introduced in Section 2.7.

### 6.4.1 Maker's strategy

Let  $H$  be a graph and assume  $n$  to be large enough whenever needed. We will prove  $\tau(\mathcal{H}_{H,n}, b) \geq \left( \frac{1}{m_2(H)} - o(1) \right) \log_{b+1}(n)$ , as required.

Let  $n_0, r \in \mathbb{N}$  and  $\delta \in (0, 1)$  be given by Theorem 2.3.1 on input  $H$ . Let  $n \geq n_0$  and denote the collection of containers by  $\{C_i : i \in [t(n)]\}$ . Maker plays according to Lemma 6.1.4 with

$$\gamma = 2 \cdot m_2(H) \cdot \frac{\log b + \log \log_{b+1}(n) + 5}{\log n - 2 \cdot m_2(H) \cdot \log \log n},$$

$s = 1$  and hypergraph  $\mathcal{H} = (E(K_n), \mathcal{F})$ , where

$$\mathcal{F} = \{E(K_n) \setminus C_i : i \in [t(n)]\}.$$

We let  $k = \min\{|F| : F \in \mathcal{F}\}$  and we check that condition (6.1.2) of Lemma 6.1.4

holds. Observe that  $\delta \binom{n}{2} \leq k \leq \binom{n}{2}$  and

$$|\mathcal{F}| \leq \binom{\binom{n}{2}}{rn^{2-1/m_2(H)}} \cdot (2^r)^{rn^{2-1/m_2(H)}} < n^{2rn^{2-1/m_2(H)}}.$$

In particular,  $\frac{n^{1/m_2(H)}}{\log^2 n} \leq \frac{k}{\log |\mathcal{F}|} \leq n^2$  and therefore

$$\begin{aligned} \left( \frac{k}{\log |\mathcal{F}|} \right)^{\gamma/2} &\geq \left( \frac{n^{1/m_2(H)}}{\log^2 n} \right)^{\gamma/2} = \exp \left[ \frac{\gamma}{2} \left( \frac{1}{m_2(H)} \log n - 2 \log \log n \right) \right] \\ &= \exp (5 + \log b + \log \log_{b+1}(n)) \\ &> 40b \log_{b+1}(n) > 20b \log_{b+1} \left( \frac{k}{\log |\mathcal{F}|} \right). \end{aligned}$$

Therefore Maker can claim an edge from each  $F \in \mathcal{F}$  for at least

$$\begin{aligned} (1 - \gamma) \log_{b+1} \left( \frac{k}{\log |\mathcal{F}|} \right) &\geq (1 - \gamma) \log_{b+1} \left( \frac{n^{1/m_2(H)}}{\log^2 n} \right) \\ &\geq \left( \frac{1 - \gamma}{m_2(H)} - \frac{2 \log_{b+1}(\log n)}{\log_{b+1}(n)} \right) \log_{b+1}(n) \\ &= \left( \frac{1}{m_2(H)} - o(1) \right) \log_{b+1}(n) \end{aligned}$$

stages, where we use that  $\gamma = o(1)$  since  $b$  is subpolynomial in  $n$ .

It remains to show that if a graph  $G$  contains an edge from every  $F \in \mathcal{F}$ , then  $G$  still contains a copy of  $H$ . Observe that if  $G$  is  $H$ -free then, by the container theorem (Theorem 2.3.1), there exists  $i \in [t]$  such that  $E(G) \subseteq C_i$ . But  $G$  contains an edge in  $E(K_n) \setminus C_i$ , which is a contradiction.

Therefore the lower bound of Theorem 1.3.7 is proven.  $\square$

#### 6.4.2 Breaker's strategy

We split Breaker's strategy in two phases. In the first phase, which will occupy the main part of the game, Breaker ensures that the board will not have many copies of  $H$  clustered together (see Definition 2.7.1). Then afterwards, in the second phase, Breaker can consider each cluster separately and destroy all remaining copies of  $H$  in a tiny number of stages.

Given any constant  $\varepsilon > 0$  and any  $n \in \mathbb{N}$  large enough, we show that Breaker can block all copies of  $H$  in at most  $\left( \frac{1}{m_2(H)} + \varepsilon \right) \log_{b+1}(n)$  stages. For this, let  $K$  be any subgraph of  $H$  such that  $\frac{e(K)-1}{v(K)-2} = m_2(H)$ , and notice that  $d_2(K) = m_2(K) = m_2(H)$ . Next, let  $\delta = \delta(H, \varepsilon) > 0$  be a constant such that

$$\frac{1}{m_2(K) - \delta} < \frac{1}{m_2(K)} + \frac{\varepsilon}{4},$$

and pick  $t = t(H, \varepsilon, \delta) \in \mathbb{N}$  such that for all  $x \geq t - 1$  it holds that

$$\frac{e(K) + m_2(K)x}{v(K) + x} \geq m_2(K) - \delta \quad \text{and} \quad \frac{(t+2)v(K)}{(m_2(K) - \delta) \cdot t \cdot v(K) - 1} < \frac{1}{m_2(K) - \delta} + \frac{\varepsilon}{4}.$$

As already pointed out, Breaker's strategy is based on two phases: first he blocks all  $K$ -collections on at least  $tv(K)$  vertices (see Claim 6.4.1), and then he blocks all remaining copies of  $K$ . Since  $K$  is a subgraph of  $H$ , at this point Breaker will have blocked all copies of  $H$  as well.

**Claim 6.4.1** (First phase of Breaker's strategy). *Breaker has a strategy so that after  $\left(\frac{1}{m_2(K)} + \frac{\varepsilon}{2}\right) \log_{b+1}(n)$  stages, the board does not contain a  $K$ -collection with at least  $tv(K)$  vertices.*

*Proof of Claim 6.4.1.* By Claim 2.7.3 we know that, if the board contains a  $K$ -collection on at least  $tv(K)$  vertices, then it must also contain an  $s$ -bunch  $B$  with  $s \geq t$  and  $tv(K) \leq v(B) \leq (t+1)v(K)$ . Hence, Breaker can concentrate on blocking such bunches, and we define

$$\mathcal{F} = \left\{ B = \bigcup_{i \in [s]} K_i : \begin{array}{l} B \text{ is an } s\text{-bunch of copies } K_i \text{ of } K \text{ in } K_n \text{ with} \\ s \geq t \text{ and } tv(K) \leq v(B) \leq (t+1)v(K) \end{array} \right\}.$$

Let  $\mathcal{F}_0 = \mathcal{F}$  and, throughout the game, denote by  $\mathcal{F}_i \subseteq \mathcal{F}_{i-1}$  the family of all elements  $F \in \mathcal{F}_{i-1}$  that Maker has fully occupied at the end of stage  $i$ . In order to prove the claim, we must show that Breaker has a strategy to ensure  $\mathcal{F}_k = \emptyset$  for  $k = \left(\frac{1}{m_2(K)} + \frac{\varepsilon}{2}\right) \log_{b+1}(n)$ .

Now, as any bunch in  $\mathcal{F}_0$  has at most  $(t+1)v(K)$  vertices, we have

$$|\mathcal{F}_0| \leq n^{(t+1)v(K)} \cdot 2^{((t+1)v(K))^2} < n^{(t+2)v(K)}. \quad (6.4.1)$$

Using Theorem 6.1.1, Breaker has a strategy to ensure that

$$|\mathcal{F}_i| \leq \sum_{B \in \mathcal{F}_{i-1}} (b+1)^{-e(B)+1} \leq |\mathcal{F}_{i-1}| \cdot (b+1)^{-(m_2(K)-\delta) \cdot t \cdot v(K)+1} \quad (6.4.2)$$

for each positive  $i \in \mathbb{N}$ , where in the last inequality we use that  $e(B) = d(B) \cdot v(B) \geq (m_2(K) - \delta) \cdot t \cdot v(K)$  for all  $B \in \mathcal{F}_{i-1} \subseteq \mathcal{F}_0$ , which follows from the Claim 2.7.4 (note its assumptions hold by the choice of  $t$ ). Combining (6.4.1) and (6.4.2), we observe that

$$|\mathcal{F}_k| < n^{(t+2)v(K)} \cdot (b+1)^{k \cdot [-(m_2(K)-\delta) \cdot t \cdot v(K)+1]} \leq 1$$

since, by our choice of  $\delta$  and  $t$ , we have  $k > \frac{(t+2)v(K)}{(m_2(K)-\delta) \cdot t \cdot v(K)-1} \log_{b+1}(n)$ . Therefore  $\mathcal{F}_k = \emptyset$  and this finishes the proof of the claim.  $\square$

Now, consider the first moment when Breaker made sure that every remaining  $K$ -



collection has fewer than  $tv(K)$  vertices, and denote them with  $\mathcal{K}_1, \dots, \mathcal{K}_\ell$ ,  $\ell \in \mathbb{N}_0$ . Since any two such collections are edge-disjoint by definition, we know then that each remaining copy of  $K$  must appear in a unique collection. From now on, in each further stage, Breaker plays as follows: whenever Maker claims an edge of  $E(\mathcal{K}_i)$  for some  $i \in [\ell]$ , Breaker claims as many edges as possible of the same collection. In all other cases, Breaker plays arbitrarily. Since each of the collections  $\mathcal{K}_i$  has fewer than  $(tv(K))^2$  edges, it takes fewer than  $\log_{b+1}((tv(K))^2)$  stages until from each collection there is at most one edge left and hence all copies of  $K$  are blocked. Combining the two phases, Breaker wins within

$$\begin{aligned} \left( \frac{1}{m_2(K)} + \frac{\varepsilon}{2} \right) \log_{b+1}(n) + 2 \log_{b+1}(tv(K)) &\leq \left( \frac{1}{m_2(K)} + \varepsilon \right) \log_{b+1}(n) \\ &= \left( \frac{1}{m_2(H)} + \varepsilon \right) \log_{b+1}(n) \end{aligned}$$

stages. Hence, the upper bound of Theorem 1.3.7 is proven.  $\square$

## 6.5 Proof of Theorem 1.3.8 (Pancyclicity game)

In this section we prove Theorem 1.3.8. We start observing that the upper bound  $\tau(\mathcal{PAN}_n, b) \leq \left( \frac{1}{2} + o(1) \right) \log_{b+1}(n)$  follows from Theorem 1.3.7. Indeed Breaker plays according to his strategy given by Theorem 1.3.7 with  $H = C_3$ . By doing this, he can ensure that the board does not contain any triangle after  $\left( \frac{1}{2} + o(1) \right) \log_{b+1}(n)$  stages and thus it cannot be pancyclic. For the lower bound, we will use the following criterion for a graph to be pancyclic, which is a corollary of Theorem 1.1 in [72].

**Theorem 6.5.1** (Corollary from Theorem 1.1 in [72]). *Let  $G$  be a graph on  $n$  vertices such that*

- (P1) *every independent set of  $G$  is of size at most  $\sqrt{n}$ ;*
- (P2)  *$G$  is  $600\sqrt{n}$ -vertex-connected,*

*then  $G$  is pancyclic.*

### 6.5.1 Maker's strategy

Let  $n$  be large enough and consider the family  $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3$  with

$$\begin{aligned} \mathcal{F}_1 &= \{E_{K_n}(A) : A \subset V(K_n), |A| = \sqrt{n}\}, \\ \mathcal{F}_2 &= \{E_{K_n}(A, B) : A, B \subset V(K_n), A \cap B = \emptyset, |A| = 1, |B| = n - 700\sqrt{n}\}, \\ \mathcal{F}_3 &= \{E_{K_n}(A, B) : A, B \subset V(K_n), A \cap B = \emptyset, |A| = \sqrt{n}, |B| = \sqrt{n}\}, \end{aligned}$$

and set  $k_j = \min\{|F| : F \in \mathcal{F}_j\}$  for  $j \in [3]$ . Maker plays according to Lemma 6.1.4 with

$$\gamma = 2 \cdot \frac{\log b + \log \log_{b+1}(n) + 5}{\log \sqrt{n} - \log \log n - \log 3000},$$

$s = 3$  and hypergraph  $\mathcal{H} = (E(K_n), \mathcal{F})$ . Note that  $\gamma = o(1)$  since  $b$  is subpolynomial in  $n$ .

We check that condition (6.1.2) of Lemma 6.1.4 holds. We have  $|\mathcal{F}_1| = \binom{n}{\sqrt{n}}$ ,  $|\mathcal{F}_2| = n \binom{n-1}{700\sqrt{n}-1}$ , and  $|\mathcal{F}_3| = \binom{n}{\sqrt{n}} \binom{n-\sqrt{n}}{\sqrt{n}}$ . Further, we have  $k_1 = \binom{\sqrt{n}}{2}$ ,  $k_2 = n - 700\sqrt{n}$ , and  $k_3 = n$ . Note that  $\frac{n}{4} \leq k_j \leq n$  and  $\sqrt{n}^{\sqrt{n}} \leq 3|F_j| = s|F_j| \leq n^{700\sqrt{n}}$  and thus  $\frac{\sqrt{n}}{3000 \log n} \leq \frac{k_j}{\log(s|\mathcal{F}_j|)} \leq 2 \frac{\sqrt{n}}{\log n}$  for every  $j \in [3]$ . Using this, we can estimate for each  $j \in [3]$ ,

$$\begin{aligned} \left( \frac{k_j}{\log(s|\mathcal{F}_j|)} \right)^{\gamma/2} &\geq \left( \frac{\sqrt{n}}{3000 \log(n)} \right)^{\gamma/2} = \exp(\log b + \log \log_{b+1}(n) + 5) \\ &> 20b \log_{b+1}(n) \geq 20b \max \left\{ 1, \log_{b+1} \left( \frac{k_j}{\log(s|\mathcal{F}_j|)} \right) \right\}, \end{aligned}$$

and thus (6.1.2) holds as well. We conclude that Maker can claim an edge of each  $F \in \mathcal{F}$  for at least

$$\begin{aligned} (1 - \gamma) \log_{b+1} \left( \frac{\sqrt{n}}{3000 \log(n)} \right) &\geq \left( \frac{1}{2} - \gamma - \frac{\log_{b+1}(3000 \log n)}{\log_{b+1}(n)} \right) \log_{b+1}(n) \\ &= \left( \frac{1}{2} - o(1) \right) \log_{b+1}(n) \end{aligned}$$

stages.

It remains to show that if a graph  $G$  contains an edge from every  $F \in \mathcal{F}$ , then  $G$  fulfills (P1) and (P2). Let  $A \subset V(K_n)$  be any vertex set of size  $|A| = \sqrt{n}$ . Then, using the subfamily  $\mathcal{F}_1$ , the graph  $G$  has at least one edge within  $A$ . Therefore, every independent set can be of size at most  $\sqrt{n}$  as required by (P1). Further, let  $V(K_n) = A \cup B \cup C$  be any partition of  $V(K_n)$  with  $|C| = 600\sqrt{n}$ . We show that  $G$  contains an edge between  $A$  and  $B$ , and thus it is  $600\sqrt{n}$ -vertex-connected as required by (P2). Assume  $|A| \geq \sqrt{n}$  and  $|B| \geq \sqrt{n}$ . In this case, using the subfamily  $\mathcal{F}_3$ , the graph  $G$  contains an edge between  $A$  and  $B$ . Assume otherwise without loss of generality that  $|A| < \sqrt{n}$ . Using  $\mathcal{F}_2$  instead, every vertex in  $A$  has at least  $700\sqrt{n} > |A| + |C|$  neighbours in  $G$ , so there needs to be a neighbour in  $B$ .  $\square$

---

## Conclusion and open problems

---

In this chapter, we discuss connections to other lines of research and highlight some open problems in the field. We first list the ones related to randomly perturbed graphs in Section 7.1, then those related to the appearance of transversals in hypergraph collections in Section 7.2, and finally those arising in the context of the multistage version of the Maker-Breaker game in Section 7.3.

### 7.1 Other spanning structures in randomly perturbed graphs

#### 7.1.1 Larger powers of Hamilton cycles

Theorem 1.1.14 completely determines the perturbed threshold for the containment of the square of a Hamilton cycle, and it is natural to investigate larger powers as well.

Let  $r \geq 3$  be fixed. In the extremal setting, Komlós, Sarközy, and Szemerédi [74] showed that for  $n$  large enough any  $n$ -vertex graph  $G$  with minimum degree  $\delta(G) \geq \frac{r}{r+1}n$  contains the  $r$ -th power of a Hamilton cycle. This result establishes that for the perturbed threshold  $\hat{p}_\alpha = 0$  holds for  $\alpha \geq \frac{r}{r+1}$ . In the random setting, i.e. when  $\alpha = 0$ , the threshold is understood as well and equal to  $n^{-1/r}$ , which follows from a result of Riordan [98]. Research in the perturbed setting has mainly focused on the range of small positive densities or densities close to the extremal threshold. In particular, it was shown by Böttcher, Montgomery, Parczyk, and Person [26] that for any  $\alpha \in (0, 1)$  there exists  $\eta = \eta(r, \alpha) > 0$  such that for any  $n$ -vertex graph  $G_\alpha$  with minimum degree  $\delta(G_\alpha) \geq \alpha n$  a.a.s.  $G_\alpha \cup G(n, p)$  contains the perturbed threshold for the containment of the  $r$ -th power of a Hamilton cycle, provided  $p \geq n^{-1/r-\eta}$ . At the other extreme, Nenadov and Trujić [92], improving on a result of Dudek, Reiher, Ruciński, and Schacht [43], showed that for  $\alpha \geq \frac{r}{r+1}$ , while  $G_\alpha$  alone contains the  $r$ -th power of a Hamilton cycle, adding a linear number of random edges suffices to enforce the  $(2r + 1)$ -st power of a Hamilton cycle. When  $\alpha > 1/2$ , even higher powers of Hamilton cycles have been studied by Antoniuk, Dudek, Reiher, Ruciński, and Schacht [6] and by Antoniuk, Dudek, and Ruciński [7], although this last result is concerned with a different notion of threshold. We remark that no exact results are known for  $\alpha \in (0, 1/2]$  and  $r \geq 3$ .

The first case that remains open is  $r = 3$ , where we recall that it is only known that the perturbed threshold is  $n^{-1/3}$  for  $\alpha = 0$ ,  $n^{-1}$  for  $1/2 < \alpha < 3/4$ , and 0 for  $\alpha \geq 3/4$ . We observe that we can obtain natural lower bounds by determining the sparsest possible structure that remains for  $G(n, p)$  after mapping the third power of a Hamilton cycle into the complete bipartite graph  $H_\alpha$  with parts of size  $\alpha n$  and  $(1 - \alpha)n$ . When  $\alpha = 1/4$ , this structure is essentially the square of a Hamilton cycle on  $3n/4$ , and is obtained by mapping every fourth vertex of the third power of a Hamilton cycle into the smaller part of  $H_{1/4}$ . Therefore, in order for  $H_{1/4} \cup G(n, p)$  to contain the third power of a Hamilton cycle, we need  $G(n, p)$  to contain the square of a Hamilton cycle on  $3n/4$  vertices. This gives  $\hat{p}_{1/4} \geq n^{-1/2}$  and we believe this is actually tight.

**Conjecture 7.1.1.** *Let  $\hat{p}_\alpha$  be the perturbed threshold for the containment of the third power of a Hamilton cycle. Then  $\hat{p}_{1/4} = n^{-1/2}$ .*

However, as discussed in the introduction, finding the square of a Hamilton cycle at this probability is a particularly challenging problem, and, additionally, it is not possible to first embed small parts arbitrarily and then connect them, as we do in the proof of our main result.

For each of the ranges  $0 < \alpha < 1/4$  and  $1/4 < \alpha < 1/2$ , it is not clear whether to expect a similar ‘jumping’ behaviour as the one proved for the square of a Hamilton cycle in the range  $0 < \alpha < 1/2$ . We can obtain natural lower bounds similarly as we did for  $\alpha = 1/4$ . Again, the sparsest structure that remains for  $G(n, p)$  is obtained essentially by mapping every  $1/\alpha$ -th vertex of  $C_n^3$  into the smaller part of  $H_\alpha$ . However, in contrast to the lower bounds in Proposition 1.1.15, the threshold for the appearance of this structure in  $G(n, p)$  is not determined by the second or third power of a short path. For example, when  $\alpha = 1/5$ , by doing as described above and mapping every fifth vertex of  $C_n^3$  into the smaller part of  $H_\alpha$ , we are left with copies of  $P_4^3$ , which are connected by three edges, cyclically. By a first moment argument, the threshold for this structure in  $G(n, p)$  is at least  $n^{-4/9}$ , and thus  $\hat{p}_{1/5} \geq n^{-4/9}$ , which is larger than the threshold for a  $P_4^3$ -factor in  $G(n, p)$ .

In addition, this lower bound does not seem to be attainable with our approach, because at this probability there is no small structure that we can find and then connect into the third power of a Hamilton cycle. Indeed at probability  $n^{-4/9}$  it is not possible to first find the copies of  $P_4^3$ s arbitrarily and then connect them. On the other hand, we believe our methods can give the following. We map every ninth and tenth vertex of  $C_n^3$  into the smaller part of  $H_{1/5}$ , and this leaves copies of  $P_8^3$  connected by single edges, cyclically. We expect that it is possible to extend our argument to this set-up, but this would only imply  $\hat{p}_{1/5} \leq n^{-7/18}$ . A similar discussion (with other edge-densities) holds for the other values of  $\alpha \in (0, 1/4) \cup (1/4, 1/2)$  and, if our argument goes through, it would improve on the bounds obtained in [26], but it would still be far from the lower bounds discussed above. The exact threshold remains a mystery.

### 7.1.2 Larger clique factors

As discussed in the previous section, the perturbed threshold for the  $(r - 1)$ -st power of a Hamilton cycle with  $r \geq 4$  is not entirely known. For certain values of  $\alpha$  and  $r \geq 4$ , the perturbed threshold is not even precisely known for  $K_r$ -factors. Indeed, as discussed in Section 1.1.2, the perturbed threshold is known for all  $\alpha \in [0, 1]$ , except the boundary cases  $\alpha \in \{1/r, 2/r, \dots, (r - 2)/r\}$ .

Let  $r \geq 4$  and  $\alpha = 1 - k/r$  with  $2 \leq k \leq r - 1$ . A natural extremal structure is given by the complete  $\lceil r/k \rceil$ -partite graph with  $\lfloor r/k \rfloor$  classes of size  $kn/r$  and possibly one class of size  $(1 - \lfloor r/k \rfloor k/r)n$  if  $k \nmid r$ . This implies that to get a  $K_r$ -factor in  $G_\alpha \cup G(n, p)$  we need to cover all but polylog  $n$  vertices of the sets of size  $kn/r$  with vertex-disjoint copies of  $K_k$ . It follows that the threshold is at least the threshold for a  $K_k$ -factor in  $G(n, p)$ , i.e.  $n^{-2/k}(\log n)^{2/(k^2-k)}$ . Surprisingly, this is not sufficient in the case when  $r > 3$  and  $k \neq 2$ ; in fact, for small  $\varepsilon$ , even  $n^{-2/k+\varepsilon}$  is not sufficient.

We briefly explain the counterexample for  $r = 4$  and  $k = 3$ , by constructing an  $n$ -vertex graph  $G$  with minimum degree  $\delta(G) \geq (1 - 3/4)n = n/4$  such that, even for small  $\varepsilon > 0$ , with  $p \geq n^{-2/3+\varepsilon}$  a.a.s. the graph  $G \cup G(n, p)$  does not contain a  $K_4$ -factor. Let  $0 < \varepsilon \leq 1/49$ ,  $p \geq n^{-2/3+\varepsilon}$ , and  $n^{7\varepsilon} \leq m \leq n^{1/7}$ . Then, for two sets  $A$  and  $B$  with  $|A| = n/4 - m$  and  $|B| = 3n/4 + m$ , we let  $G$  be the  $n$ -vertex graph on  $V(G) = A \cup B$  such that  $A$  is an independent set,  $G[B]$  is given by  $|B|/(2m)$  disjoint copies of  $K_{m,m}$ , and any pair of vertices  $(a, b)$  with  $a \in A$  and  $b \in B$  is an edge. Clearly  $G$  has minimum degree  $n/4$ . If  $G \cup G(n, p)$  contains a  $K_4$ -factor, since  $A$  only contains  $n/4 - m$  vertices, at least  $m$  copies of  $K_4$  must lie within  $B$ . However we claim that a.a.s. the perturbed graph  $G \cup G(n, p)[B]$  contains less than  $m$  copies of  $K_4$  and thus a.a.s.  $G \cup G(n, p)$  does not contain a  $K_4$ -factor. Denote by  $X$  the number of  $K_4$ 's in  $G \cup G(n, p)[B]$ . It is not hard to see that when  $m$  is not too small the best way to build a  $K_4$  in  $B$  is to choose a  $K_{m,m}$  in  $B$  and ask for an edge of  $G(n, p)$  on each side of  $K_{m,m}$ . We get  $\mathbb{E}[X] \lesssim \frac{n}{m} m^4 p^2 = o(m)$  and by Markov's inequality a.a.s.  $X < m$  as claimed.

**Problem 7.1.2.** *Determine the behaviour of the perturbed threshold for a  $K_4$ -factor and the extremal graphs at  $\alpha = 1/4$ .*

The counterexamples for other values of  $r > 3$  and  $k \neq 2$  can be constructed in a similar way, by slightly modifying the corresponding extremal graph defined above. In the case when  $k = 2$  this construction does not increase the lower bound  $n^{-1} \log n$  and, with Theorem 1.1.13 in mind, we believe that Theorem 1.1.7 generalises to  $K_r$ . However we believe that in all cases, using our methods, the non-extremal Theorem 1.1.11 can be extended to any  $K_r$ -factor: i.e. for all  $2 \leq k \leq r - 1$  and with  $\alpha = 1 - k/r$ , when  $G_\alpha$  is not close (with a similar condition as in Definition 1.1.10) to the extremal graph defined above, then  $p \geq Cn^{-2/k}$  is sufficient for a  $K_r$ -factor in  $G_\alpha \cup G(n, p)$ .

## 7.2 Transversals in hypergraph collections

### 7.2.1 Vertex degree for tight Hamilton cycles

Statement (B) of Theorem 1.2.5 proves  $d$ -colour-blindness of the family of  $k$ -uniform Hamilton  $\ell$ -cycles, for various ranges of  $d$ ,  $k$ , and  $\ell$ . However, there is a well-known (uncoloured) Dirac-type result whose rainbow version is missing there: the vertex minimum degree for tight Hamilton cycles in 3-uniform hypergraphs, corresponding to  $d = 1$ ,  $k = 3$  and  $\ell = k - 1$ .

The proof of the minimum vertex degree threshold for this family is due to Reiher, Rödl, Ruciński, Schacht, and Szemerédi [97], and it uses the absorption method, making this family an ideal candidate for our main theorem. However, it turns out that we cannot hope for the property **Con** to hold in this range of the parameters (see Section 2.1 in [97] for a discussion). Due to this additional complication, it would be an interesting challenge to obtain a transversal generalisation of the result in [97].

### 7.2.2 Exact results and stability

For Hamilton cycles in graphs, the exact rainbow minimum degree threshold is known [70], and the family of Hamilton cycles is *exactly* colour-blind, meaning that an error-term as in Definition 1.2.3 is not required. It is natural to ask whether exact results also hold for other structures in the rainbow setup and if a general statement similar to our Theorem 5.2.1 can be proved. For example, improving on the statement (A) of Theorem 1.2.5, it would be very interesting to show if  $\delta(\mathbf{G}) \geq rn/(r+1)$  is already sufficient for a transversal copy of the  $r$ -th power of a Hamilton cycle in graphs. Note that already for a rainbow  $K_r$ -factor it is not known whether  $\delta(\mathbf{G}) \geq rn/(r+1)$  suffices. Moreover, resolving this for the  $r$ -th power of a Hamilton cycles does not immediately imply the analogous result for a  $K_r$ -factors, even though the former contains the latter, because of the different number of colours needed for a rainbow embedding. A similar observation is true for tight Hamilton cycles and perfect matchings in hypergraphs. We remark that Lu, Wang, and Yu [84] showed that the family of  $k$ -uniform perfect matchings is *exactly*  $(k-1)$ -colour blind, proving that the rainbow minimum co-degree threshold essentially is  $n/2$ . Improving on one of the statements in (B) of Theorem 1.2.5, we can ask if the same condition  $\delta_{k-1}(\mathbf{H}) \geq n/2$  is sufficient for a transversal copy of a tight Hamilton cycle.

In the non-rainbow setup, exact results can typically be obtained by considering an extremal and non-extremal case separately, where the latter often gives stability. Lu, Wang, and Yu [84] give an exact result for perfect matchings in collections of  $k$ -uniform hypergraphs. Their arguments uses absorbers and distinguishes between an extremal and non-extremal case. Roughly speaking, they say that a hypergraph collection is extremal if essentially all of them are close to one of the extremal graphs for the uncoloured problem.

Working with a similar notion for an extremal collection, it would be interesting to prove an exact version for any of our results, as many of them hold in the uncoloured setup, e.g. [14, 58, 59, 60, 73, 100].

It seems to be too much to hope for a general theorem that covers all of these applications, because of the different extremal constructions in each case. But we remark that, besides the properties **Ab**, **Con**, and **Fac**, the additional  $(\alpha n^{k-d})$ -term for the minimum  $d$ -degree in our theorem is only necessary for the two applications of Lemma 5.3.5. Therefore, a major step would be a variant of this theorem (for specific  $\mathcal{A}$ ) which is applicable with a lower minimum  $d$ -degree, under the assumption that the hypergraph collection is not extremal. However, Hamilton cycles give new complications and in this setup it is harder to make a direct use of the results from the non-rainbow case, which was the main scope of our investigation.

### 7.2.3 Other potential applications

There are many more structures which can be represented as  $\mathcal{A}$ -cycles, e.g. copies of  $C_4$  glued as depicted in Figure 5.3. In the case of graphs, any  $\mathcal{A}$ -link forms an  $\mathcal{A}$ -cycle with bounded maximum degree and bounded bandwidth. Thus, the bandwidth theorem by Böttcher, Schacht and Taraz [29] immediately gives minimum degree thresholds for the existence of a Hamilton  $\mathcal{A}$ -cycle, but their proof relies on different techniques than we require for the application of Theorem 5.2.1. Hence, one would need to prove properties **Ab**, **Con**, and **Fac** for such structures in order to obtain the corresponding rainbow result using our method. More generally, a rainbow version of the bandwidth theorem would be very interesting. Note that the bandwidth theorem is not optimal for many graphs, so the minimum degree conditions for the containment of Hamilton  $\mathcal{A}$ -cycles is an interesting problem even in the non-rainbow setup.

## 7.3 Multistage positional games

In Section 1.3 we have introduced the multistage Maker-Breaker game and determined the duration of this game for several natural graph properties. Being a new set-up, there are many directions open for future work.

### 7.3.1 Connection to the classical setting

Observe that in all our results (Theorems 1.3.4 and 1.3.6 to 1.3.8) it happens that the threshold  $\tau(\mathcal{H}, b)$  is asymptotically the same as  $\log_{b+1}(b_{\mathcal{H}})$ , where  $b_{\mathcal{H}}$  denotes the threshold bias for the Maker-Breaker game on the hypergraph  $\mathcal{H}$  (see Section 1.3.1). We wonder whether this is always the case.

**Problem 7.3.1.** Does there exist a hypergraph  $\mathcal{H} = (X, \mathcal{F})$  and a positive integer  $b$ , for which  $\tau(\mathcal{H}, b)$  and  $\log_{b+1}(b_{\mathcal{H}})$  are not asymptotically the same?

### 7.3.2 Additional games

One could continue this line of research and estimate  $\tau(\mathcal{H}, b)$  for other starting hypergraphs  $\mathcal{H}$ . We believe it would be interesting to consider the case where  $\mathcal{H}$  is the hypergraph on vertex set  $E(K_n)$ , with hyperedges being the edge sets of (i) triangle factors, (ii) copies of a fixed spanning tree or (iii) powers of Hamilton cycles.

Additionally, as a natural next step one may also consider multistage variants of other positional games, for example Waiter-Client games and Client-Waiter game on  $K_n$ , or Maker-Breaker games on the random graph  $G(n, p)$ .

### 7.3.3 Multistage Maker-Breaker game with stop

The following variant of multistage games springs to mind, which we may call *multistage game with stop*. Let a hypergraph  $\mathcal{H} = (X, \mathcal{F})$  and a bias  $b \geq 1$  be given, and define  $X_0 = X$ . For  $i \geq 1$ , the stage  $i$  is played on the board  $X_{i-1} \subseteq X$ , and it ends the first time Maker claims a winning set from  $\mathcal{F}_{i-1} \subseteq \mathcal{F}$  completely. Then we define the next board  $X_i$  to consist of all the elements which have been claimed by Maker or are free by the end of stage  $i$ , and we let  $\mathcal{F}_i \subseteq \mathcal{F}_{i-1}$  be the family of those winning sets which are still fully contained in  $X_i$ . Similarly to the threshold  $\tau(\mathcal{H}, b)$ , we may define the duration of this game, and denote it by  $\tau^{\text{stop}}(\mathcal{H}, b)$ .

It is easy to see that  $\tau^{\text{stop}}(\mathcal{H}, b) \geq \tau(\mathcal{H}, b)$  always holds. Moreover, for the connectivity game and hence also the Hamilton cycle game, we obtain that this bound is asymptotically tight, as  $\tau^{\text{stop}}(C_n, b) \leq (1 + o(1)) \log_{b+1}(n)$  can be shown if  $b$  is subpolynomial in  $n$ . Indeed Breaker just needs to isolate a single vertex.

However, if we consider local properties instead, e.g. the  $H$ -game, there can be a huge difference between  $\tau^{\text{stop}}(\mathcal{H}, b)$  and  $\tau(\mathcal{H}, b)$ . Indeed, for  $b = 1$ , using that on dense graphs Maker can claim a copy of  $H$  fast and applying Turán's Theorem, it is straightforward to prove that  $\tau^{\text{stop}}(\mathcal{H}_{H,n}, 1) = \Theta(n^2)$ . It would be interesting to understand this variant much better. Hence, we suggest the following problem.

**Problem 7.3.2.** Given any graph  $H$  and any constant bias  $b$ , determine  $c = c(H, b)$  such that  $\tau^{\text{stop}}(\mathcal{H}_{H,n}, b) = (c \pm o(1))n^2$ .

Already the case when  $H = K_3$  and  $b = 1$  is open and would be of interest.



---

## Bibliography

---

- [1] R. Aharoni, M. DeVos, S. González Hermosillo de la Maza, A. Montejano and R. Šámal, *A rainbow version of Mantel's Theorem*, Adv. Comb. (2020), Paper No. 2, 12.
- [2] M. Aigner and A. Brandt, *Embedding arbitrary graphs of maximum degree two*, J. London Math. Soc. (2) 48.1 (1993), 39–51.
- [3] E. Aigner-Horev, O. Danon, D. Hefetz and S. Letzter, *Large Rainbow Cliques in Randomly Perturbed Dense Graphs*, SIAM J. Discrete Math. 36.4 (2022), 2975–2994.
- [4] E. Aigner-Horev, O. Danon, D. Hefetz and S. Letzter, *Small rainbow cliques in randomly perturbed dense graphs*, European J. Combin. 101 (2022), Paper No. 103452, 34.
- [5] D. Angluin and L. G. Valiant, *Fast probabilistic algorithms for Hamiltonian circuits and matchings*, J. Comput. System Sci. 18.2 (1979), 155–193.
- [6] S. Antoniuk, A. Dudek, C. Reiher, A. Ruciński and M. Schacht, *High powers of Hamiltonian cycles in randomly augmented graphs*, J. Graph Theory 98.2 (2021), 255–284.
- [7] S. Antoniuk, A. Dudek and A. Ruciński, ‘Powers of Hamiltonian cycles in randomly augmented Dirac graphs – the complete collection’, arXiv:2204.10738, 2022.
- [8] J. Balogh, R. Morris and W. Samotij, *Independent sets in hypergraphs*, J. Amer. Math. Soc. 28.3 (2015), 669–709.
- [9] J. Balogh, F. Mousset and J. Skokan, *Stability for vertex cycle covers*, Electron. J. Combin. 24.3 (2017), P3.56.
- [10] J. Balogh, A. Treglown and A. Z. Wagner, *Tilings in randomly perturbed dense graphs*, Combin. Probab. Comput. 28.2 (2019), 159–176.
- [11] J. Barkey, *Maker-Breaker Spiele über mehrere Runden*, Bachelor Thesis, Hamburg University of Technology (2021).
- [12] J. Barkey, D. Clemens, F. Hamann, M. Mikalački and A. Sgueglia, ‘Multistage Positional Games’, arXiv:2202.04344, 2022.

- 
- [13] J. d. O. Bastos, G. O. Mota, M. Schacht, J. Schnitzer and F. Schulenburg, *Loose Hamiltonian cycles forced by large  $(k - 2)$ -degree—approximate version*, SIAM J. Discrete Math. 31.4 (2017), 2328–2347.
- [14] J. d. O. Bastos, G. O. Mota, M. Schacht, J. Schnitzer and F. Schulenburg, *Loose Hamiltonian cycles forced by large  $(k - 2)$ -degree—sharp version*, Contrib. Discrete Math. 13.2 (2018), 88–100.
- [15] J. Beck, *Combinatorial games: Tic-Tac-Toe theory*, vol. 114, Cambridge university press Cambridge, 2008.
- [16] J. Beck, *Foundations of positional games*, Proceedings of the Seventh International Conference on Random Structures and Algorithms (Atlanta, GA, 1995), vol. 9, 1-2, 1996, 15–47.
- [17] J. Beck, *On positional games*, J. Combin. Theory Ser. A 30.2 (1981), 117–133.
- [18] J. Beck, *Random graphs and positional games on the complete graph*, Random graphs '83 (Poznań, 1983), vol. 118, North-Holland Math. Stud. North-Holland, Amsterdam, 1985, 7–13.
- [19] J. Beck, *Remarks on positional games. I*, Acta Math. Acad. Sci. Hungar. 40.1-2 (1982), 65–71.
- [20] J. Beck, *Van der Waerden and Ramsey type games*, Combinatorica 1.2 (1981), 103–116.
- [21] M. Bednarska and T. Łuczak, *Biased positional games for which random strategies are nearly optimal*, Combinatorica 20.4 (2000), 477–488.
- [22] T. Bohman, A. Frieze and R. Martin, *How many random edges make a dense graph Hamiltonian?*, Random Structures Algorithms 22.1 (2003), 33–42.
- [23] B. Bollobás, *Threshold functions for small subgraphs*, Mathematical Proceedings of the Cambridge Philosophical Society, vol. 90, 2, Cambridge University Press, 1981, 197–206.
- [24] B. Bollobás and A. Thomason, *Threshold functions*, Combinatorica 7.1 (1987), 35–38.
- [25] J. Böttcher, J. Han, Y. Kohayakawa, R. Montgomery, O. Parczyk and Y. Person, *Universality for bounded degree spanning trees in randomly perturbed graphs*, Random Structures Algorithms 55.4 (2019), 854–864.
- [26] J. Böttcher, R. Montgomery, O. Parczyk and Y. Person, *Embedding spanning bounded degree graphs in randomly perturbed graphs*, Mathematika 66.2 (2020), 422–447.
- [27] J. Böttcher, O. Parczyk, A. Sgueglia and J. Skokan, 'The square of a Hamilton cycle in randomly perturbed graphs', arXiv:2202.05215, 2022.

- 
- [28] J. Böttcher, O. Parczyk, A. Sgueglia and J. Skokan, *Triangles in randomly perturbed graphs*, *Combin. Probab. Comput.* 32.1 (2023), 91–121.
- [29] J. Böttcher, M. Schacht and A. Taraz, *Proof of the bandwidth conjecture of Bollobás and Komlós*, *Math. Ann.* 343.1 (2009), 175–205.
- [30] J. Böttcher, M. Schacht and A. Taraz, *Spanning 3-colourable subgraphs of small bandwidth in dense graphs*, *J. Combin. Theory Ser. B* 98.4 (2008), 752–777.
- [31] E. Buß, H. Hàn and M. Schacht, *Minimum vertex degree conditions for loose Hamilton cycles in 3-uniform hypergraphs*, *J. Combin. Theory Ser. B* 103.6 (2013), 658–678.
- [32] Y. Cheng, J. Han, B. Wang and G. Wang, ‘Rainbow spanning structures in graph and hypergraph systems’, arXiv:2105.10219, 2021.
- [33] Y. Cheng, J. Han, B. Wang, G. Wang and D. Yang, ‘Rainbow Hamilton cycle in hypergraph systems’, arXiv:2111.07079, 2021.
- [34] Y. Cheng, G. Wang and Y. Zhao, *Rainbow pancyclicity in graph systems*, *Electron. J. Combin.* 28.3 (2021), Paper No. 3.24, 9.
- [35] V. Chvátal and P. Erdős, *Biased positional games*, *Ann. Discrete Math.* 2 (1978), 221–229.
- [36] D. Clemens, H. Gebauer and A. Liebenau, *The random graph intuition for the tournament game*, *Combin. Probab. Comput.* 25.1 (2016), 76–88.
- [37] C. Cooper and A. M. Frieze, *Pancyclic random graphs*, *Random graphs ’87* (Poznań, 1987), Wiley, Chichester, 1990, 29–39.
- [38] K. Corrádi and A. Hajnal, *On the maximal number of independent circuits in a graph*, *Acta Math. Acad. Sci. Hungar.* 14 (1963), 423–439.
- [39] S. Das, C. Knierim and P. Morris, ‘Schur properties of randomly perturbed sets’, arXiv:2205.01456, 2022.
- [40] S. Das, P. Morris and A. Treglown, *Vertex Ramsey properties of randomly perturbed graphs*, *Random Structures Algorithms* 57.4 (2020), 983–1006.
- [41] S. Das and A. Treglown, *Ramsey properties of randomly perturbed graphs: cliques and cycles*, *Combin. Probab. Comput.* 29.6 (2020), 830–867.
- [42] G. A. Dirac, *On the maximal number of independent triangles in graphs*, *Abh. Math. Sem. Univ. Hamburg* 26 (1963), 78–82.
- [43] A. Dudek, C. Reiher, A. Ruciński and M. Schacht, *Powers of Hamiltonian cycles in randomly augmented graphs*, *Random Structures Algorithms* 56.1 (2020), 122–141.

- [44] P. Erdős, *Problem 9*, Theory of Graphs and its Applications (Proc. Sympos. Smolenice, 1963), ed. by M. Fiedler, Publ. House Czechoslovak Acad. Sci., Prague, 1964, 159.
- [45] P. Erdős, *Some recent results on extremal problems in graph theory. Results*, Theory of Graphs (Internat. Sympos., Rome, 1966), Gordon and Breach, New York; Dunod, Paris, 1967, 117–123 (English), pp. 124–130 (French).
- [46] P. Erdős, A. Gyárfás and L. Pyber, *Vertex coverings by monochromatic cycles and trees*, J. Combin. Theory Ser. B 51.1 (1991), 90–95.
- [47] P. Erdős and A. Rényi, *On the evolution of random graphs*, Magyar Tud. Akad. Mat. Kutató Int. Közl. 5 (1960), 17–61.
- [48] P. Erdős and J. L. Selfridge, *On a combinatorial game*, J. Combinatorial Theory Ser. A 14 (1973), 298–301.
- [49] L. Euler, *Solutio problematis ad geometriam situs pertinentis*, Commentarii academiae scientiarum Petropolitanae 8 (1741), 128–140.
- [50] A. Ferber, M. Krivelevich and H. Naves, *Generating random graphs in biased Maker-Breaker games*, Random Structures Algorithms 47.4 (2015), 615–634.
- [51] A. Ferber, G. Kronenberg and K. Luh, *Optimal threshold for a random graph to be 2-universal*, Trans. Amer. Math. Soc. 372.6 (2019), 4239–4262.
- [52] K. Frankston, J. Kahn, B. Narayanan and J. Park, *Thresholds versus fractional expectation-thresholds*, Ann. Math. (2) 194.2 (2021), 475–495.
- [53] S. Gerke and A. McDowell, *Nonvertex-balanced factors in random graphs*, J. Graph Theory 78.4 (2015), 269–286.
- [54] P. Gupta, F. Hamann, A. Müyesser, O. Parczyk and A. Sgueglia, ‘A general approach to transversal versions of Dirac-type theorems’, arXiv:2209.09289, 2022.
- [55] A. Hajnal and E. Szemerédi, *Proof of a conjecture of P. Erdős*, Combinatorial theory and its applications, II (Proc. Colloq., Balatonfüred, 1969), 1970, 601–623.
- [56] A. W. Hales and R. I. Jewett, *Regularity and positional games*, Trans. Amer. Math. Soc. 106 (1963), 222–229.
- [57] J. Han, P. Morris and A. Treglown, *Tilings in randomly perturbed graphs: bridging the gap between Hajnal-Szemerédi and Johansson-Kahn-Vu*, Random Structures Algorithms 58.3 (2021), 480–516.
- [58] J. Han and Y. Zhao, *Minimum codegree threshold for Hamilton  $\ell$ -cycles in  $k$ -uniform hypergraphs*, J. Combin. Theory Ser. A 132 (2015), 194–223.
- [59] J. Han and Y. Zhao, *Minimum vertex degree threshold for loose Hamilton cycles in 3-uniform hypergraphs*, J. Combin. Theory Ser. B 114 (2015), 70–96.

- 
- [60] H. Hàn, J. Han and Y. Zhao, *Minimum degree thresholds for Hamilton  $(k/2)$ -cycles in  $k$ -uniform hypergraphs*, J. Combin. Theory Ser. B 153 (2022), 105–148.
- [61] H. Hàn and M. Schacht, *Dirac-type results for loose Hamilton cycles in uniform hypergraphs*, J. Combin. Theory Ser. B 100.3 (2010), 332–346.
- [62] D. Hefetz, M. Krivelevich, M. Stojaković and T. Szabó, *Planarity, colorability, and minor games*, SIAM J. Discrete Math. 22.1 (2008).
- [63] D. Hefetz, M. Krivelevich, M. Stojaković and T. Szabó, *Positional games*, vol. 44, Oberwolfach Seminars, Birkhäuser/Springer, Basel, 2014.
- [64] D. Hefetz, M. Krivelevich and T. Szabó, *Avoider-enforcer games*, J. Combin. Theory Ser. A 114.5 (2007), 840–853.
- [65] D. Hefetz, M. Krivelevich and T. Szabó, *Hamilton cycles in highly connected and expanding graphs*, Combinatorica 29.5 (2009), 547–568.
- [66] J. Hladký, P. Hu and D. Piguet, *Komlós’s tiling theorem via graphon covers*, J. Graph Theory 90.1 (2019), 24–45.
- [67] W. Hoeffding, *Probability inequalities for sums of bounded random variables*, J. Amer. Statist. Assoc. 58 (1963), 13–30.
- [68] S. Janson, T. Łuczak and A. Rucinski, *Random graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000, xii+333.
- [69] A. Johansson, J. Kahn and V. Vu, *Factors in random graphs*, Random Structures Algorithms 33.1 (2008), 1–28.
- [70] F. Joos and J. Kim, *On a rainbow version of Dirac’s theorem*, Bull. Lond. Math. Soc. 52.3 (2020), 498–504.
- [71] J. Kahn, B. Narayanan and J. Park, *The threshold for the square of a Hamilton cycle*, Proc. Amer. Math. Soc. 149 (2021), 3201–3208.
- [72] P. Keevash and B. Sudakov, *Pancyclicity of Hamiltonian and highly connected graphs*, J. Combin. Theory Ser. B 100.5 (2010), 456–467.
- [73] J. Komlós, G. N. Sárközy and E. Szemerédi, *On the square of a Hamiltonian cycle in dense graphs*, Proceedings of the Seventh International Conference on Random Structures and Algorithms (Atlanta, GA, 1995), vol. 9, 1-2, 1996, 193–211.
- [74] J. Komlós, G. N. Sárközy and E. Szemerédi, *Proof of the Seymour conjecture for large graphs*, Ann. Comb. 2.1 (1998), 43–60.
- [75] J. Komlós, G. N. Sárközy and E. Szemerédi, *Spanning trees in dense graphs*, Combin. Probab. Comput. 10.5 (2001), 397–416.

- 
- [76] 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, 295–352.
- [77] A. D. Koršunov, *Solution of a problem of P. Erdős and A. Rényi on Hamiltonian cycles in undirected graphs*, Dokl. Akad. Nauk SSSR 228.3 (1976), 529–532.
- [78] M. Krivelevich, *The critical bias for the Hamiltonicity game is  $(1 + o(1))n/\ln n$* , J. Amer. Math. Soc. 24.1 (2011), 125–131.
- [79] M. Krivelevich, M. Kwan and B. Sudakov, *Bounded-degree spanning trees in randomly perturbed graphs*, SIAM J. Discrete Math. 31.1 (2017), 155–171.
- [80] D. Kühn and D. Osthus, *The minimum degree threshold for perfect graph packings*, Combinatorica 29.1 (2009), 65–107.
- [81] A. Lehman, *A solution of the Shannon switching game*, J. Soc. Indust. Appl. Math. 12 (1964), 687–725.
- [82] I. Levitt, G. N. Sárközy and E. Szemerédi, *How to avoid using the regularity lemma: Pósa's conjecture revisited*, Discrete Math. 310.3 (2010), 630–641.
- [83] A. Liebenau and N. Wormald, 'Asymptotic enumeration of graphs by degree sequence, and the degree sequence of a random graph', arXiv:1702.08373, 2017.
- [84] H. Lu, Y. Wang and X. Yu, 'Co-degree threshold for rainbow perfect matchings in uniform hypergraphs', arXiv:2111.00372, 2021.
- [85] H. Lu, Y. Wang and X. Yu, *Rainbow perfect matchings for 4-uniform hypergraphs*, SIAM J. Discrete Math. 36.3 (2022), 1645–1662.
- [86] H. Lu, X. Yu and X. Yuan, *Rainbow matchings for 3-uniform hypergraphs*, J. Combin. Theory Ser. A 183 (2021), 105489.
- [87] W. Mantel, *Problem 28*, Wiskundige Opgaven 10 (1907), 60–61.
- [88] C. McDiarmid, *On the method of bounded differences*, Surveys in combinatorics, 1989 (Norwich, 1989), vol. 141, London Math. Soc. Lecture Note Ser. Cambridge Univ. Press, Cambridge, 1989, 148–188.
- [89] R. Montgomery, *Spanning trees in random graphs*, Adv. Math. 356 (2019), 106793.
- [90] R. Montgomery, A. Müyesser and Y. Pehova, *Transversal factors and spanning trees*, Adv. Comb. (2022).
- [91] R. Nenadov, A. Steger and M. Stojaković, *On the threshold for the Maker-Breaker  $H$ -game*, Random Structures Algorithms 49.3 (2016), 558–578.
- [92] R. Nenadov and M. Trujić, *Sprinkling a few random edges doubles the power*, SIAM J. Discrete Math. 35.2 (2021), 988–1004.

- 
- [93] O. Parczyk, *2-universality in randomly perturbed graphs*, European J. Combin. 87 (2020), 103–118.
- [94] M. Pavez-Signé, N. Sanhueza-Matamala and M. Stein, ‘Towards a hypergraph version of the Pósa-Seymour conjecture’, arXiv:2110.09373, 2021.
- [95] L. Pósa, *Hamiltonian circuits in random graphs*, Discrete Math. 14.4 (1976), 359–364.
- [96] E. Powierski, ‘Ramsey properties of randomly perturbed dense graphs’, arXiv:1902.02197, 2019.
- [97] C. Reiher, V. Rödl, A. Ruciński, M. Schacht and E. Szemerédi, *Minimum vertex degree condition for tight Hamiltonian cycles in 3-uniform hypergraphs*, Proc. Lond. Math. Soc. (3) 119.2 (2019), 409–439.
- [98] O. Riordan, *Spanning subgraphs of random graphs*, Combin. Probab. Comput. 9.2 (2000), 125–148.
- [99] V. Rödl, A. Ruciński and E. Szemerédi, *An approximate Dirac-type theorem for  $k$ -uniform hypergraphs*, Combinatorica 28.2 (2008), 229–260.
- [100] V. Rödl, A. Ruciński and E. Szemerédi, *Dirac-type conditions for Hamiltonian paths and cycles in 3-uniform hypergraphs*, Adv. Math. 227.3 (2011), 1225–1299.
- [101] V. Rödl, A. Ruciński and E. Szemerédi, *A Dirac-type theorem for 3-uniform hypergraphs*, Combin. Probab. Comput. 15.1-2 (2006), 229–251.
- [102] D. Saxton and A. Thomason, *Hypergraph containers*, Invent. Math. 201.3 (2015), 925–992.
- [103] P. Seymour, *Problem section*, Combinatorics (Proc. British Combinatorial Conf. 1973), Cambridge University Press, London-New York, 1974, 201–204. London Math. Soc. Lecture Note Ser., No. 13.
- [104] M. Simonovits, *A method for solving extremal problems in graph theory, stability problems*, Theory of Graphs (Proc. Colloq., Tihany, 1966), Academic Press, New York, 1968, 279–319.
- [105] D. A. Spielman and S. Teng, *Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time*, J. ACM 51.3 (2004), 385–463.
- [106] M. Stojaković and T. Szabó, *Positional games on random graphs*, Random Structures Algorithms 26.1-2 (2005), 204–223.
- [107] E. Szemerédi, *On sets of integers containing no  $k$  elements in arithmetic progression*, Acta Arith. 27 (1975), 199–245.
- [108] E. Szemerédi, *On sets of integers containing no four elements in arithmetic progression*, Acta Math. Acad. Sci. Hungar. 20 (1969), 89–104.

- [109] E. Szemerédi, *Regular partitions of graphs*, Problèmes combinatoires et théorie des graphes (Colloques Internationaux CNRS, Univ. Orsay, Orsay, 1976), vol. 260, Colloq. Internat. CNRS, CNRS, Paris, 1978, 399–401.
- [110] V. H. Vu, *A large deviation result on the number of small subgraphs of a random graph*, *Combin. Probab. Comput.* 10.1 (2001), 79–94.