Online Ramsey Games and Some Classical Problems in Graph Theory

Diss. ETH No. 23882

Online Ramsey Games and Some Classical Problems in Graph Theory

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH Zurich

(Dr. sc. ETH Zurich)

presented by

Andreas Claudio Noever

Master of Science ETH in Computer Science

born on 3.11.1988

citizen of Germany

accepted on the recommendation of

Prof. Dr. Angelika Steger, examiner

Prof. Dr. Oleg Pikhurko, co-examiner

2016

To my family

Contents

Abstract

The topics of this thesis are Online Ramsey Games, resilience of random graphs and two problems from classical graph theory. *Ramsey theory* goes back to a result of the British mathematician and philosopher Frank. P. Ramsey from 1930. In its simplest form Ramsey's theorem states that for any graph F if one colors the edges of a sufficiently large complete graph with two colors, then one cannot avoid a monochromatic copy of F. In other words: complete disorder is an impossibility.

Over the last 80 years many generalizations and variations of this statement have been proven. In this thesis we consider an *online* version of the problem. *Online Ramsey Games*, introduced by Friedgut, Kohayakawa, Rödl, Ruciński and Tetali in 2003, are *one-player games* and are played on an initially empty graph on n vertices. In every round we insert a new edge chosen uniformly at random among all non-edges of the graph. The player (called Painter) must immediately color this edge with one of r available colors. Her objective is to avoid a monochromatic copy of some fixed graph F for as long as possible. What can we say about the typical duration of this game provided that Painter follows an optimal strategy?

For the case where the forbidden graph is a triangle and Painter is given only two colors Friedgut, Kohayakawa, Rödl, Ruciński and Tetali have shown the answer to be $\Theta(n^{4/3})$. Their result was significantly extended in 2009 when Marciniszyn, Spöhel and Steger determined the correct answer for a large class of graphs, which in particular includes cycles and cliques. The game has proven to be much more difficult to analyze when more than two colors are in play. For triangles the corresponding offline game (where Painter gets to see all N edges at once) has a threshold given by $N = n^{3/2}$ and so far the only improvement is due to Balogh and Butterfield: Painter can avoid a monochromatic triangle for at most $n^{3/2-c_r}$ rounds for some constant $c_r > 0$ depending on the number of colors. For more general graphs only the trivial offline bound is available.

Our first contribution is an analysis of the Online Ramsey Game for three or more colors. We use the sparse regularity lemma as well as the recently introduced hypergraph container theorem to determine the threshold for the typical duration of the game with an arbitrary number of colors for a large class of graphs (in particular our result applies to cycles and cliques).

Our second result concerns the resilience of random graphs for squares of long cycles. Random graphs where introduced by Erdős and Rényi, and independently by Gilbert, in 1959. Since its inception the theory has been an interesting and fruitful branch of combinatorics and probability theory that has attracted the attention of both mathematicians and theoretical computer scientists. To motivate the concept of resilience it is useful to introduce a classical result of Dirac from 1952. Dirac's result states that any graph G on $n \geq 3$ vertices with minimum degree $\delta(G) \geq n/2$ contains a Hamiltonian cycle. A modern interpretation of this result is that the complete graph K_n contains a Hamilton cycle resiliently. That is, it remains Hamiltonian even if we remove up to $n/2 - 1$ edges incident to every vertex. A generalization of Dirac's result is the following conjecture of Pósa from 1962. Pósa conjectured that a graph G contains a square of a Hamiltonian cycle if $\delta(G) \geq 2n/3$. A square of a cycle C is the cycle C together with all edges between vertices that have distance 2 in C. Proving this conjecture turned out to be a difficult problem: it took almost 35 years and the development of powerful tools, like Szemerédi's Regularity Lemma and the so-called Blow-Up Lemma, before Pósa's conjecture was proven, at least for all sufficiently large n.

It is natural to ask whether these resilience results continue to hold when one replaces the complete graph with the random graph $G_{n,p}$. Indeed Sudakov and Vu have shown that for Hamiltonicity this is the case if $p \geq \log^4 n/n$. If one considers instead the square of a Hamilton cycle then one quickly realizes that one cannot hope to find the square of a cycle covering more than $n - \Theta(1/p^2)$ vertices. As our second result we show an approximation of such a best possible result: namely that the random graph resiliently contains the square of a cycle covering at least $(1 - \varepsilon)n$ vertices if $p \geq n^{\varepsilon}/\sqrt{n}$.

The second part of the thesis focuses on classical graph theory. A reoccurring pattern in combinatorics is that the absence of some object implies the presence of another and vice versa. Consider for example Menger's theorem or the more general max-flow min-cut theorem: either there exist many s-t-paths or there exists a small s-t-cut. A classic result of Erdős and Pósa states that any graph either contains k vertex-disjoint cycles or can be made acyclic by deleting at most $\mathcal{O}(k \log k)$ vertices. This seminal result has spawned a long line of papers proving similar duality results between covering and packing of different families of graphs, directed graphs, hypergraphs, rooted graphs and other combinatorial objects. Our third contribution is a proof of the following theorem: for every k and l every graph either contains k vertex-disjoint cycles of length at least l or is such that one can destroy all cycles of length at least l by removing at most $\mathcal{O}(kl + k\log k)$ vertices. This is asymptotically optimal.

The final result relates to a more than 25 years old conjecture of Erdős. In 1990 Erdős conjectured that every graph on n vertices with $2n-1$ edges contains a subgraph of minimal degree 3 of order at most $(1-\varepsilon)n$. Towards this conjecture Erdős, Faudree, Rousseau and Schelp proved that one can always remove $\Omega(\sqrt{n})$ vertices. We improve this to $\Omega(n/\log n)$.

Zusammenfassung

Diese Arbeit befasst sich mit Online-Ramsey-Spielen, der Widerstandsfähigkeit von Zufallsgraphen sowie zwei Problemen aus der klassischen Graphentheorie. Die *Ramseytheorie* hat ihren Ursprung in einem Satz, den der britischen Mathematiker und Philosoph Frank. P. Ramsey im Jahre 1930 bewies. In seiner einfachsten Form besagt Ramsey's Satz, dass, wenn man für einen beliebigen Graphen F die Kanten eines genügend grossen vollständigen Graphen mit zwei Farben färbt, man gezwungen ist, eine monochromatische Kopie von F zu erstellen. In anderen Worten: Völlige Unordnung ist unmöglich.

Im Verlauf der letzten 80 Jahre wurden viele Verallgemeinerungen und Variationen dieses Satzes bewiesen. In dieser Arbeit beschäftigen wir uns mit einer *online* Version des Problems. *Online-Ramsey-Spiele* wurden erstmals 2003 von Friedgut, Kohayakawa, Rödl, Ruciński und Tetali betrachtet. Diese Spiele werden alleine und auf einem zu Beginn leeren Graphen mit n Knoten gespielt. In jeder Runde wird eine zufällig ausgewählte neue Kante hinzugefügt, die der Spieler mit einer von r verfügbaren Farben färben muss. Ziel des Spiels ist es, eine monochromatische Kopie eines vorher festgelegten Graphen F möglichst lange zu vermeiden. Was können wir über die typische Länge dieses Spiels sagen, wenn wir annehmen, dass der Spieler optimal spielt?

Für den Fall, dass der verbotene Graph F ein Dreieck ist und dem Spieler nur zwei Farben zur Verfügung stehen, haben Friedgut, Kohayakawa, Rödl, Ruciński und Tetali bereits bewiesen, dass die gesuchte Antwort $\Theta(n^{4/3})$ ist. Dieses Resultat wurde 2009 signifikant erweitert als Marciniszyn, Spöhel und Steger die korrekte Antwort für eine grosse Klasse von Graphen bestimmten (insbesondere für Kreise und Cliquen). Es hat sich herausgestellt, dass die Analyse des Spiels bedeutend schwieriger ist, wenn mehr als zwei Farben im Spiel sind. Für Dreiecke hat das entsprechende Offline-Spiel (in dem der Spieler alle N Kanten auf einmal sieht) einen Schwellwert

bei $N = n^{3/2}$. Diese Schranke wurde bisher nur einmal verbessert: Balog und Butterfield haben gezeigt, dass der Spieler ein einfarbiges Dreieck für höchstens $n^{3/2-c_r}$ Runden vermeiden kann, wobei die Konstante $c_r > 0$ von der Anzahl verfügbarer Farben abhängt. Für andere Graphen ist kein besseres Ergebnis als die triviale obere Schranke des offline Spiels bekannt.

Unser erster Beitrag ist eine Analyse des Online-Ramsey-Spiels für mehrere Farben. Wir verwenden das 'sparse regularity lemma' sowie den erst kürzlich eingeführten Hypergraphenkontainersatz, um den Schwellwert für die typische Länge des Spiels mit einer beliebigen Anzahl Farben für eine grosse Klasse von Graphen zu bestimmen (insbesondere für Kreise und Cliquen).

Unser zweites Ergebnis bezieht sich auf die Widerstandsfähigkeit von Zufallsgraphen. Zufallsgraphen wurden von Erdős und Rényi sowie von Grilbert im Jahre 1959 eingeführt. Die Theorie der Zufallsgraphen ist ein interessantes und ertragreiches Gebiet der Kombinatorik und Wahrscheinlichkeitstheorie, welches die Aufmerksamkeit sowohl von Mathematikern als auch von Informatikern auf sich gezogen hat. Um das Konzept der Widerstandsfähigkeit zu motivieren, macht es Sinn, ein klassisches Resultat Dirac's aus dem Jahr 1952 vorzustellen. Der Satz besagt, dass jeder Graph G auf $n \geq 3$ Knoten mit Minimalgrad $\delta(G) \geq n/2$ einen Hamiltonkreis enthält. Eine moderne Interpretation dieses Ergebnisses besagt, dass der vollständige Graph K_n einen Hamiltonkreis widerstandsfähig enthält. Das heisst, K_n bleibt hamiltonisch, auch wenn wir bis zu $n/2 - 1$ Kanten an jedem Knoten entfernen. Eine Verallgemeinerung von Dirac's Satz ist die folgende Vermutung von Pósa aus dem Jahre 1962. Pósa vermutete, dass ein Graph G das Quadrat eines Hamiltonkeises enthält, falls $\delta(G) \geq 2n/3$. Das Quadrat eines Kreises C ist der Kreis C zusammen mit allen Kanten zwischen Knoten mit Distanz 2 in C. Diese Vermutung zu beweisen, stellte sich als schwieriges Unterfangen heraus. Es dauerte fast 35 Jahre und benötigte die Entwicklung neuer Werkzeuge, wie Szemerédi's Regularitäts-Lemma sowie das sogennante Blow-Up-Lemma, bis Pósa's Vermutung bewiesen werden konnte - zumindest für n gross genug.

Eine interessante Frage ist, ob diese Widerstandsfähigkeit erhalten bleibt, wenn man den vollständigen Graphen mit dem Zufallsgraphen $G_{n,p}$ ersetzt. In der Tat haben Sudakov und Vu gezeigt, dass dies für Hamiltonkreise der Fall ist wenn $p \geq \log^4 n/n$. Interessiert man sich stattdessen für das Quadrat eines Hamiltonkreises, dann erkennt man schnell, dass es nicht möglich ist, mehr als $n - \Theta(1/p^2)$ Knoten mit dem Quadrat eines Kreises abzudecken. Unser zweites Resultat ist eine Annäherung an ein solches bestmögliches Ergebnis: ein Zufallsgraph enthält widerstandsfähig das Quadrat eines Kreises, der mindestens $(1 - \varepsilon)n$ Knoten beinhaltet wenn $p \geq n^{\varepsilon}/\sqrt{n}$.

In der zweiten Hälfte dieser Arbeit konzentrieren wir uns auf Probleme der klassischen Graphentheorie. Ein wiederkehrendes Phänomen in der Kombinatorik ist, dass die Abwesenheit eines Objekts die Existenz eines anderen impliziert und umgekehrt. Betrachten wir zum Beispiel den Satz von Menger oder allgemeiner das Max-Flow-

Min-Cut-Theorem: In jedem Graphen finden wir entweder viele s-t-Pfade oder einen kleinen s-t-Schnitt. Ein weiteres klassisches Theorem aus diesem Bereich ist das folgende von Erdős und Pósa. Es besagt, dass jeder Graph entweder k knotendisjunkte Kreise enthält oder dass durch das Entfernen von höchstens $\mathcal{O}(k \log k)$ Knoten alle Kreise zerstört werden können. Aufbauend auf diesem Resultat wurde eine lange Reihe von Sätzen über ähnliche Dualitäten zwischen Abdeckung und Packungen von verschiedenen Familien von Graphen, gerichteten Graphen, Hypergraphen, gewurzelten Graphen und anderen kombinatorischen Objekten bewiesen. Unser dritter Beitrag ist ein Beweis des folgenden Satzes: Für jedes k und l enthält jeder Graph entweder k knotendisjunkte Kreise der Länge mindestens l oder man kann alle Kreise der Länge mindestes l zerstören, indem man höchstens $\mathcal{O}(kl + k \log k)$ Knoten entfernt. Das Resultat ist asymptotisch optimal.

Der letzte Beitrag dieser Arbeit bezieht sich auf eine mehr als 25 Jahre alte Vermutung von Erdős. 1990 vermutete Erdős, dass jeder Graph auf n Knoten mit $2n - 1$ Kanten einen Teilgraphen mit Minimalgrad 3 auf höchstens $(1 - \varepsilon)n$ Knoten enthält. Eine Verallgemeinerung für Subgraphen mit Minimalgrad k wurde ebenfalls angegeben. Erdős, Faudree, Rousseau und Schelp bewiesen, dass immer $\Omega(\sqrt{n})$ Knoten entfernt werden können. Wir verbessern diese Schranke auf $\Omega(n/\log n)$.

Acknowledgments

First and most important, I want to express my deep gratitude to everybody who supported me during the years as a PhD student. I am especially grateful to my supervisor Angelika Steger for her guidance, advice and for giving me the opportunity to work in her excellent research group over the past three and a half years. Doing a PhD with Angelika was probably one of the best decisions of my life.

I am grateful to Oleg Pikhurko, who agreed to co-examine my thesis.

I sincerely appreciated the collaboration with my co-authors: Frank Mousset, Rajko Nenadov, Nemanja Škorić, Asaf Ferber, Yann Disser, Felix Weissenberger, Yury Person and Ueli Peter.

Also my special thanks go to all past and present members of the ever-growing CSA familiy. The atmosphere in our group, especially during the many joint lunches and during the coffee breaks was always a highlight and our regular workshops in Buchboden offered excellent research opportunities. Thank you for making this possible: Angelika Steger, Frank Mousset, Rajko Nenadov, Nemanja Škorić, Johannes Lenger, Felix Weissenberger, Ralph Keusch, Asaf Ferber, Wojciech Samotij, Konstantinos Panagiotou, Hafsteinn Einarsson, Marcelo Gauy, Pascal Pfister, Florian Meier, Miloš Trujić, Manuela Fischer, Karl Bringmann, Luca Gugelmann, Yann Disser, Yury Person, Thomas Rast and Ueli Peter.

My work was supported by a grant from the Swiss National Science Foundation, which I gratefully acknowledge.

Finally, I want to thank my Mum, Dad and sisters for all their love and support.
Thank you!

Chapter 1

Introduction and results

In *Ramsey theory* one studies conditions under which order must appear. A simple and very early example of this kind of condition is the so called *pigeonhole principle*: if 10 pigeons fly into 9 holes than at least on hole must contain more than one pigeon. The following special case of the original theorem of Ramsey can be seen as a profound generalization of this principle: in every group of at least six people there are either three people who know each other or three people who are mutual strangers.

Statements which answer questions of the form "does every large structure necessarily contain an orderly substructure?" are often called *Ramsey-type* statements. The seminal theorem of this field, proved in 1930 by Frank P. Ramsey [Ram30], implies that for any integer k there exists a large integer $N(k)$ such that if one colors the edges of the complete graph on N vertices with two colors then one always finds a monochromatic clique of size k. Phrased another way, Ramsey theory states that complete disorder is an impossibility. Every sufficiently large structure must contain an orderly substructure.

Determining just how large a graph needs to be for Ramsey's theorem to hold has been one of the central question in the study of Ramsey theory. The Ramsey number $r(K_k)$ is defined to be the smallest number $N(k)$ for which Ramsey's theorem holds. Even for small values of k the only two exactly known Ramsey numbers are $r(K_3)$

and $r(K_4)$. For general k it is known that

$$\frac{\sqrt{2}}{e} 2^{k/2} k \leq r(K_k) \leq k^{-C\frac{\log k}{\log \log k}} \binom{2k-2}{k-1},$$

where the lower bound is due to Spencer [Spe75] and the upper bound was recently obtained by Conlon [Con09]. Determining the exact asymptotic growth rate of $r(K_k)$ remains a notoriously difficult open problem in combinatorics.

Ramsey's original theorem has sparked a large amount of research into various generalizations and variations of this problem. For example one can replace the clique K_k with an arbitrary graph F or allow for more than two colors. Another venue is to exchange the host graph K_N with some other graph and investigate which conditions are necessary for Ramsey's theorem to hold. One might believe that the underlying reason for Ramsey's theorem is the large number of edges in K_N and that every host graph must be quite dense. This turns out to be false. In fact it is possible to obtain a Ramsey-type statement where the complete graph has been replaced by a much sparser random graph.

In the following sections we introduce the concept of a random graph and the Rödl-Ruciński theorem which extends Ramsey's theorem to random graphs. We then introduce *Online Ramsey Games* which transfer the Ramsey problem from an 'offline' setting into an 'online' setting where the host graph is revealed piece by piece instead of being known from the start. These games will be the focus of the first part of the thesis. We also briefly introduce the other contributions of the thesis and the notions surrounding them. In particular we will present the concept of *local resilience*, the *Erdős-Pósa-property* and a conjecture of Erdős on *subgraphs of minimal degree k*.

1.1 Random Graphs

The systematic study of random graphs was pioneered by Erdős and Rényi in 1959 [ER59, ER60] and independently by Gilbert [Gil59]. They introduced two closely related models of random graphs: the binomial random graph $G_{n,p}$ and the uniform random graph $G_{n,m}$.

For two integers n and m, where $0 \leq m \leq \binom{n}{2}$, we define the uniform random graph model $G_{n,m}$ by assigning every graph G on n vertices with exactly m edges the same probability

$$\Pr[G_{n,m} = G] = \frac{1}{\binom{\binom{n}{2}}{m}}.$$

The closely related binomial random graph model is defined as follows: for a given integer n and a probability $p \in [0,1]$ define $G_{n,p}$ by assigning to every graph G on n vertices the probability

$$\Pr\left[G_{n,p} = G\right] = p^{e_G}(1-p)^{\binom{n}{2}-e_G},$$

where e_G denotes the number of edges of G. Equivalently $G_{n,p}$ may be viewed as the product space of $\binom{n}{2}$ independent Bernoulli random variables, where every random variable controls the presence of a single edge.

One of the most characteristic properties of both models is that they exhibit a threshold behavior. That is the probability that a certain property holds changes rapidly with small changes in the edge probability p. For example it is well-known that a.a.s. $G_{n,p}$ is connected if $p \gg n/\log n$, and disconnected if $p \ll n/\log n$. Connectivity is not the only property which exhibits this phenomena. In fact Bollobás and Thomason [BT87] have shown that that for every monotone graph property $\mathcal{P}$ (monotone in the sense that it cannot be destroyed by adding more edges) there exists a threshold $p_0(n)$ such that

$$\lim_{n\to\infty} \Pr\left[G_{n,p} \in \mathcal{P}\right] = \begin{cases} 0 & \text{if } p \ll p_0(n), \\ 1 & \text{if } p \gg p_0(n). \end{cases}$$

Examples of such monotone properties include 'containing a triangle', connectedness, Hamiltonicity or even the property 'every 2-edge-coloring contains a monochromatic triangle'. Since the inception of the field of random graphs the thresholds for many interesting properties have been determined. For a full introduction into the theory of random graphs we refer to the classic books of Bollobás [Bol01] and Janson, Ruciński and Łuczak [JŁR11].

1.2 Online Ramsey Games

Random graphs and Ramsey theory are intricately related. Many constructions in the field are based on random graphs. In fact, if one wants to color the edges of K_n while avoiding a large monochromatic clique the best known construction is to color randomly.

In 1992 Łuczak, Ruciński and Voigt [ŁRV92] initiated the study of Ramsey properties of random graphs by proving a threshold result for the property 'every r-vertex coloring of $G_{n,p}$ contains a monochromatic copy of F'. In the same paper they also gave initial results for the edge coloring version of the problem and solved the case when F is a triangle. This result was later generalized to arbitrary graphs by Rödl and Rucinśki [RR93, RR94, RR95] who determined the threshold for the property 'every r-edge coloring of $G_{n,p}$ contains a monochromatic copy of F'.

An *online* version of the above problem called *Online Ramsey Games* was introduced Friedgut, Kohayakawa, Rödl, Ruciński and Tetali in 2003 [FKR$^+$03]. This version of the problem is phrased in terms of a one-player game. The online F-avoidance game with r colors is played on an initially empty graph on n vertices. In every round we insert a new edge chosen uniformly at random among all non-edges of the graph. The player (called Painter) must immediately color this edge with one of r available colors. Her objective is to avoid a monochromatic copy of some fixed graph F for as long as possible.

We call $N_0(F, r, n)$ a threshold function for the online F-avoidance game with r colors if for every $N \ll N_0(F, r, n)$ there exists a strategy for Painter that survives for N rounds with high probability and if for every $N \gg N_0(r, n)$ every strategy fails to survive for N rounds with high probability. Such a threshold function always exists [MSS09a].

Already in their original paper Friedgut, Kohayakawa, Rödl, Ruciński and Tetali have shown that $N_0(K_3, 2, n) = n^{4/3}$ is a threshold for the online triangle-avoidance game with two colors [FKR+03]. Later, in 2009, Marciniszyn, Spöhel and Steger analyzed a greedy strategy and showed its optimality for two colors for a class of graphs which includes cycles and cliques [MSS09a, MSS09b]. They conjectured that for cycles and cliques the greedy strategy is also tight when more than two colors are available to Painter.

Getting tight results when more than two colors are involved has proven to be quite difficult. The corresponding offline triangle avoidance game (where Painter gets to see all N edges at once) has a threshold given by $N = n^{3/2}$ [RR95]. Clearly this upper bound also applies to the online game. In [BMS12] Belfrage, Mütze and Spöhel connected the probabilistic one player game to a deterministic two player game. This technique was used by Balogh and Butterfield in [BB10] to improve the upper bound for the online triangle avoidance game to $n^{3/2-c_r}$ for some constant $c_r > 0$. Thus the thresholds of the online and offline games differ when F is a triangle. Still their upper bound of $n^{2/3-c_r}$ does not match the lower bound provided by Marciniszyn, Spöhel and Steger.

In Chapter 3 we prove an upper bound, which matches the known lower bounds from [MSS09a], for an arbitrary number of colors and large a class of graphs. In particular we show that

Theorem. *Let F be a cycle or clique. Then the threshold for the online F-avoidance game with r colors is*

$$N_0(F, r, n) = n^{2-1/\overline{m}_2^r(F)}.$$

where $\overline{m}_r^2(F)$ is given by

$$\overline{m}_2^r(F) := \begin{cases} \max_{H \subseteq F} \frac{e_H}{v_H} & \text{if } r = 1, \\ \max_{H \subseteq F} \frac{e_H}{v_H - 2 + 1/\overline{m}_2^{r-1}(F)} & \text{if } r \geq 2. \end{cases}$$

1.3 Resilience of Random Graphs

A classical result in the area of Hamiltonicity is the following one of Dirac [Dir52]. It states that any graph G on $n \geq 3$ vertices with minimum degree $\delta(G) \geq n/2$

contains a Hamilton cycle. This result is not difficult and a proof can be found in most text books on graph theory, see e.g. [Bol98, Wes00]. One also easily checks that the constant $1/2$ is best possible: the complete bipartite graph on $(n-1)/2$ and $(n+1)/2$ vertices (assuming n odd) has minimum degree $n/2 - 1/2$ but does not contain a Hamilton cycle.

A modern interpretation of this result is that the complete graph K_n contains a Hamilton cycle resiliently. That is it remains Hamiltonian even if we remove up to $n/2 - 1$ edges incident to every vertex. This type of resilience, where we limit the number of edge deletions at each vertex, is called *local* resilience. For many natural problems this type of resilience gives more interesting results than the maybe more intuitive variant of *global* resilience (where we count the total number of edge deletions). For example it is easily seen that the removal $n-2$ edges suffices to destroy the Hamiltonian property of K_n.

A generalization of Dirac's result is the following conjecture of Pósa. In 1962, Pósa conjectured that $G(V,E)$ contains a square of a Hamiltonian cycle if $\delta(G) \geq 2n/3$. A square of a cycle C is the cycle C together with all edges between vertices that have distance 2 in C. Again it is not difficult to see that the constant $2/3$ is best possible, just consider the complete tripartite graph on $(n-1)/3$, $(n-1)/3$ and $(n+2)/3$ vertices (assuming 3 divides $n-1$). However proving that minimum degree $2n/3$ actually suffices turned out to be a difficult problem. It required the development of powerful tools, most notably Szemerédi's Regularity Lemma [Sze75, Sze76] and the so-called Blow-Up Lemma [KSS97], before Pósa's conjecture was proven, at least for all sufficiently large n [KSS96].

It is natural to ask whether these resilience results continue to hold when one replaces the complete graph with the random graph $G_{n,p}$. The systematic study of resilience of random graphs was initiated by Sudakov and Vu in 2008 [SV08] and has attracted considerable research interest [DKMS08, KLS10, BCS11, BSKS11, LS12, BKT13, KLS14, HSS15, FNP$^+$15]. The problem of Hamiltonicity was already considered in the original paper of Sudakov and Vu who have shown that if $p > \log^4 n/n$ then $G_{n,p}$ a.a.s. has local resilience $1/2 - o(1)$. This was later improved to the asymptotically optimal value of $p = C \log n/n$ by Lee and Sudakov [LS12].

In Chapter 4 we study the resilience of the random graph for the property 'containing a square of a Hamilton cycle'. As it turns out for this problem already the threshold for existence is a hard problem that, till today, is not yet completely understood. As a square of a Hamilton cycle contains many triangles, $p \geq c/\sqrt{n}$ is certainly necessary, and it is conceivable that this may also be the true answer. The best bound known is $p \geq C \log^4 n/\sqrt{n}$ [NŠ16a].

For the corresponding resilience problem one is thus tempted to speculate that at least for $p \geq \mathrm{poly}(\log n)/\sqrt{n}$, for an appropriate polylog-factor, we have that the resilience of $G_{n,p}$ with respect to the property 'containing a square of a Hamilton cycle' is $1/3 - o(1)$. However, for this property it is easy to see that this is far too optimistic. By deleting all edges in the neighborhood of a vertex v we can ensure that v cannot be part of any square of a cycle. Thus for any $p = o(1)$ we have that the resilience for 'containing a square of a Hamilton cycle' is $o(1)$. In order to obtain a non-trivial result we thus need to weaken the required property. One easily checks that for constant resilience the best one can hope for is to find a square of a cycle that covers all but $\Theta(1/p^2)$ vertices. In Chapter 4 we show an approximate version of such a best possible result.

Theorem. *For every $\gamma, \nu > 0$ and $p = n^{-\frac{1}{2}+\gamma}$ a.a.s. every subgraph of $G_{n,p}$ with minimum degree at least $(2/3 + \nu)np$ contains a square of a cycle on at least $(1 - \nu)n$ vertices.*

1.4 The Erdős-Pósa property

A reoccurring pattern in combinatorics is that the absence of some object implies the presence of another and vice versa. Consider Menger's theorem: two vertices s, t of a graph can either be separated by removing k vertices or there exist k internally vertex-disjoint s-t-paths. In a graph theoretic context such a relation can often be expressed via the notions of a packing and a covering of some structures in a graph. For example in Menger's theorem we are trying to either cover all s-t-paths with k vertices or to pack k internally disjoint s-t-paths into the graph.

A classic result of Erdős and Pósa in the area of packing/covering dualities states that any graph either contains k vertex-disjoint cycles or can be made acyclic by deleting at most $\mathcal{O}(k \log k)$ vertices [EP65]. Erdős and Pósa also showed that this bound is essentially tight. Their result has spawned a long line of papers proving similar duality results between covering and packing of different families of graphs, directed graphs, hypergraphs, rooted graphs and other combinatorial objects (see a recent survey of Raymond and Thilikos [RT16] for more information).

In Chapter 5 we investigate the Erdős-Pósa property for the family $\mathcal{F}_l = \{C_m \mid m \geq l\}$ of cycles of length at least l. Given numbers l and k, what is the optimal function $f(l, k)$ such that every graph G either contains k vertex-disjoint cycles of length at least l or contains a set X of $f(l, k)$ vertices that meets all cycles of length at least l? A result of Thomassen [Tho88] implies that for every l, the family $\mathcal{F}_l$ has the Erdős-Pósa property with a function $f(l, k) \in 2^{l^{\mathcal{O}(k)}}$ (though recent results of Chekuri and Chuzhoy make it possible to substantially improve the dependency on k in this bound [CC13]). This result was sharpened by Birmelé, Bondy, and Reed [BBR07] to $f(l, k) \in \mathcal{O}(lk^2)$ in 2007 and by Fiorini and Herinckx [FH14] to $f(l, k) \in \mathcal{O}(lk \log k)$ in 2014. In Chapter 5 we improve these results to the asymptotically optimal bound $f(l, k) \in \mathcal{O}(kl + k \log k)$, thus settling the question asked in [BBR07].

1.5 Small subgraphs of minimal degree k

It is easy to show that every graph on n vertices and at least $2n-2$ edges contains a subgraph of minimal degree 3. More generally for $k \geq 2$ any graph on n vertices and at least $t_k(n) := (k-1)(n-k+2) + \binom{k-2}{2}$ edges contains a subgraph of minimal degree k. This statement is best possible in two ways: (1) there exist graphs with $t_k(n) - 1$ edges which do not contain a subgraph of minimal degree k, and (2) there exist graphs with $t_k(n)$ edges without a *proper* induced subgraph of minimal degree k. For example the wheel $W(1,n) = K_1 + C_{n-1}$ has exactly $2n-2$ edges and minimal degree 3, but contains no proper induced subgraph with minimal degree 3. A similar construction is available for all k (consider the generalized wheel $W(k-2,n) = K_{k-2} + C_{n-k+2}$).

In 1990 Erdős, Faudree, Rousseau and Schelp proved that for $k \geq 2$ every graph with n vertices and $(k-1)(n-k+2) + \binom{k-2}{2} + 1$ edges contains a subgraph of minimal degree k on at most $n - \sqrt{n}/\sqrt{6k^3}$ vertices. They conjecture that it is possible to remove at least $\varepsilon_k n$ many vertices and remain with a subgraph of minimal degree k. We make progress towards their conjecture by showing that one can remove at least $\Omega(n/\log n)$ vertices.

1.6 Thesis overview

The thesis contains four new results from the area of Online Ramsey Games, Random Graphs and General Graph Theory.

The first two results heavily involve the use of the sparse regularity lemma and related techniques. An introduction into these topics is given in Chapter 2.

The main result is a tight upper bound for the duration of Online Ramsey Games with arbitrarily many colors. It is presented in Chapter 3. The result will also appear in the journal Random Structures & Algorithms [Noe].

In Chapter 4 we present a resilience result in random graphs for squares of almost spanning cycles. This is joint work with Angelika Steger [NS16b].

In the second half of the thesis we depart from the topic of random graphs and enter the area of pure graph theory. Chapter 5 is devoted to the Erdős-Pósa property of long cycles. The result is joint work with Frank Mousset, Nemanja Škorić, and Felix Weissenberger [MNŠW16] and establishes an asymptotically tight Erdős-Pósa function for cycles with prescribed minimum length.

Finally, in Chapter 6, we present a short proof related to the existence of small subgraphs of minimal degree k. This marks the first progress on a more than 25 years old conjecture of Erdős. The bound obtained is only a logarithmic factor away from the conjectured (and best possible) value. Joint work with Frank Mousset, Rajko

Nenadov, Yury Person and Nemanja Škorić [MNN+16].

Chapter 2

A short introduction to Szemerédi's regularity lemma for sparse graphs

Szemerédi's regularity lemma is one of the most powerful tools in graph theory and extremal combinatorics. It was introduced by Szemerédi in 1975 to prove a celebrated result concerning arithmetic progression in subsets of the integers [Sze75, Sze76] . The lemma has since proven to be extremely useful in a variety of contexts and has established itself as a standard tool in the area of graph theory.

The statement of the lemma is, in rough terms, that the vertices of every sufficiently large graph may be partitioned such that the edges between most parts behave roughly like a random graph. In other words every large graph is essentially made up of a small number of random-like graphs.

The proofs in Chapter 3 and Chapter 4 rely heavily on the sparse regularity lemma, which is a generalization of the original lemma to sparse graphs (see [Koh97]), and related concepts. In this chapter we give a short introduction to the required definitions and theorems. For a more in depth introduction to the topic see for example [GS05].

As mentioned before the sparse regularity lemma allows one to partition a graph into a small number of random-like parts. Our first definition makes precise what we mean by a "random-like" graph.

Definition 2.1. *A bipartite graph $B = (U \cup W, E)$ is called (ε, p)-regular if for all $U' \subseteq U$ and $W' \subseteq W$ with $|U'| \geq \varepsilon|U|$ and $|W'| \geq \varepsilon|W|$,*

$$\left| \frac{|E(U', W')|}{|U'||W'|} - \frac{|E|}{|U||W|} \right| \leq \varepsilon p.$$

We write (ε)-regular in case p equals the density $|E|/(|U||W|)$.

B is called (ε, p)-lower-regular if for all $U' \subseteq U$ and $W' \subseteq W$ with $|U'| \geq \varepsilon|U|$ and $|W'| \geq \varepsilon|W|$,

$$\frac{|E(U', W')|}{|U'||W'|} \geq (1 - \varepsilon)p.$$

The original regularity lemma of Szemerédi allows us to partition arbitrary graphs into a constant number of $(\varepsilon, 1)$-regular pairs. Kohayakawa [Koh97] and Rödl (unpublished) independently introduced an analogue of Szemerédi's regularity lemma which gives meaningful results for $p \to 0$. The generalization works for a class of graph which do not contain large dense spots.

Definition 2.2. *Let $G = (V, E)$ be a graph and let $0 < \eta, p \leq 1$. We say that G is (η, p)-upper-uniform if for all disjoint sets $U, W \subseteq V$ with $|U| \geq \eta|V|$ and $|W| \geq \eta|V|$*

$$|E(U, W)| \leq (1 + \eta)p|U||W|.$$

We can now state Szemerédi's regularity lemma for sparse graphs. We use the second version presented in [Koh97].

Definition 2.3. *A partiton $(V_i)_{i=0}^k$ of the vertex set V is called an (ε, p)-regular partition with exceptional class V_0 if $|V_1| = |V_2| = \cdots = |V_k|$, $|V_0| \leq \varepsilon n$, and, with the exception of at most εk^2 pairs, the pairs (V_i, V_j), $1 \leq i \leq j \leq k$ are (ε, p)-regular.*

Theorem 2.4 (sparse regularity lemma [Koh97])**.** *For any $\varepsilon > 0$ and $m_0 \geq 1$, there are constants $\eta = \eta(\varepsilon, m_0) > 0$ and $M_0 = M_0(\varepsilon, m_0) \geq m_0$ such that for any $p > 0$, any (η, p)-upper-uniform graph with at least m_0 vertices admits an (ε, p)-regular partition $(V_i)_{i=0}^k$ with exceptional class V_0 such that $m_0 \leq k \leq M_0$.*

It is well-known (and an easy consequence of Chernoff's inequality) that a random graph $G_{n,p}$ with $p \gg 1/n$ is a.a.s. (η, p)-upper-uniform. The Sparse Regularity Lemma can thus be applied a.a.s. to every spanning subgraph of $G_{n,p}$. But the (ε, p)-regularity of a pair alone does not imply any lower bounds on its density. In fact the empty graph is (ε, p)-regular for all $\varepsilon > 0$, $0 \leq p \leq 1$. Still it is not hard to show that if $G \subseteq G_{n,p}$ has density at least αp then, for η, ε small enough, we find at least one pair V_i, V_j which is (ε, p)-regular with density at least $\alpha p/2$ (and thus $(2\varepsilon/\alpha)$-regular).

Even though the definition or regularity talks only about linear sized sets of vertices it in fact implies a great deal about the structure of the graph and even about individual vertices. Below we introduce a selection of these properties (we omit proofs) to give the reader some intuition about the behavior of an (ε, p)-regular graph.

The first observation is that almost all vertices have the degree that one would expect if one was dealing with a random bipartite graph:

Lemma 2.5. *Let $B = (U \cup W, E)$ be an (ε, p)-lower-regular bipartite graph. Then for every $W' \subseteq W$ of size at least $\varepsilon|W|$ all but at most $\varepsilon|U|$ vertices in U have at least $(1-\varepsilon)|W'|p$ neighbours in W'. If B is (ε)-regular with density p, then all but at most $2\varepsilon|U|$ vertices in U have at least $(1-\varepsilon)|W'|p$ and at most $(1+\varepsilon)|W'|p$ neighbours in W'.*

Finally we state two properties about spanning subgraphs of (ε)-regular graphs. Firstly it immediately follows from the definition that deleting a small number of edges does not destroy the regularity property, as we can accommodate the loss by a slightly increase in the error term.

Lemma 2.6. *Let $B = (U \cup W, E)$ be an (ε)-regular bipartite graph for some $\varepsilon < 1/3$. Then every subgraph $(U \cup W, E')$ of B such that $|E \setminus E'| \leq \varepsilon^4|E|$ is (2ε)-regular.*

Secondly it is possible to find an (ε)-regular subgraph with a prescribed number of edges (in fact a random subgraph on m edges will w.h.p. inherit regularity).

Lemma 2.7 (Lemma 4.3 in [GS05]). *For every $0 < \varepsilon \leq 1/6$, there exists a constant $C = C(\varepsilon)$ such that any (ε)-regular bipartite graph B contains a (2ε)-regular spanning subgraph with m edges for all m satisfying $C|V(B)| \leq m \leq |E(B)|$.*

Chapter 3

Online Ramsey Games for more than two colors

The contents of this chapter will appear in the journal Random Structures & Algorithms [Noe].

Consider the following one-player game played on an initially empty graph on n vertices. In every round we insert a new edge chosen uniformly at random among all non-edges of the graph. The player, henceforth called Painter, must immediately color this edge with one of r available colors. Her objective is to avoid a monochromatic copy of some fixed graph F for as long as possible. We refer to this game as the online F-avoidance game with r colors.

We call $N_0(F, r, n)$ a threshold function for the online F-avoidance game with r colors if for every $N \ll N_0(F, r, n)$ there exists a strategy for Painter that survives for N rounds with high probability and if for every $N \gg N_0(r, n)$ every strategy fails to survive for N rounds with high probability. Note that such a threshold function always exists [MSS09a, Lemma 2.1].

This game was first studied by Friedgut, Kohayakawa, Rödl, Ruciński and Tetali, who have shown in [FKR$^+$03] that $N_0(K_3, 2, n) = n^{4/3}$ is a threshold for the online triangle-avoidance game with two colors. In 2009 Marciniszyn, Spöhel and Steger proved the following lower bound.

Theorem 3.1 ([MSS09a])**.** *Let F be a graph that is not a forest, and let $r \geq 1$. Then the* *online F-avoidance game with r colors has a threshold $N_0(F,r,n)$ that satisfies*

$$N_0(F,r,n) \geq n^{2-1/\overline{m}_2^r(F)},$$

where $\overline{m}_r^2(F)$ is given by

$$\overline{m}_2^r(F) := \begin{cases} \max_{H\subseteq F} \frac{e_H}{v_H} & \text{if } r = 1, \\ \max_{H\subseteq F} \frac{e_H}{v_H - 2 + 1/\overline{m}_2^{r-1}(F)} & \text{if } r \geq 2. \end{cases}$$

In an accompanying paper they provide matching upper bounds in the two color case for a large class of graphs, which includes cycles and cliques:

Theorem 3.2 ([MSS09b]). *Let F be a graph that is not a forest which has a subgraph $F_- \subset F$ with $e_F - 1$ edges satisfying*

$$m_2(F_-) \leq \overline{m}_2^2(F),$$

where

$$m_2(F) := \max_{H\subseteq F} \frac{e_H - 1}{v_H - 2}.$$

Then the threshold for the online F-avoidance coloring game with two colors is

$$N_0(F,2,n) = n^{2-1/\overline{m}_2^2(F)}.$$

They conjecture that a similar result is true for all $r \geq 3$.

For three or more colors no tight upper bounds are known. The corresponding offline triangle avoidance game (where Painter gets to see all N edges at once) has a threshold given by $N = n^{3/2}$ [RR95]. Clearly this upper bound also applies to the online game. In [BMS12] Belfrage, Mütze and Spöhel connected the probabilistic one player game to a deterministic two player game originally introduced by Kurek and Ruciński in [KR05]. In this version of the game the edges are no longer presented in a random order but can be chosen by a second player called Builder. They show that if there exists a winning strategy for Builder which only creates subgraphs of density at most d, then $n^{2-1/d}$ is an upper bound for the threshold of the original probabilistic game.

This technique was used by Balogh and Butterfield in [BB10] to improve the upper bound for the online triangle avoidance game to $n^{3/2-c_r}$ for some constant $c_r > 0$. Thus the thresholds of the online and offline games differ. Still their upper bound of $n^{2/3-c_r}$ does not match the lower bound provided by Marciniszyn, Spöhel and Steger.

Our contribution is an upper bound, which matches the lower bound Marciniszyn, Spöhel and Steger, for an arbitrary number of colors. That is we show the following:

Theorem 3.3. *Let F be a 2-balanced graph that is not a tree which has a subgraph $F_- \subset F$ with $e_F - 1$ edges satisfying*

$$m_2(F_-) \leq \overline{m}_2^2(F).$$

Then the threshold for the online F-avoidance game with r colors is

$$N_0(F, r, n) = n^{2-1/\overline{m}_2^r(F)}.$$

The premise of our theorem is satisfied by a large class of graphs, which includes cycles and cliques. The condition that F is 2-balanced is used only for technical reasons. On the other hand the second condition is (in general) necessary. In [MSS09a] the authors give an example of a graph (two triangles intersecting in a single vertex) for which the above threshold is incorrect.

To go from two to more colors we prove a generalization of the KŁR-conjecture. For two colors the (unmodified) KŁR-conjecture immediately tells us that Painter has to color on the order of $n^{v_{F_-}} p^{e_{F_-}}$ copies of F_- with the majority color (where $p = \Theta(n^{-1/\overline{m}_2^2})$). In expectation a p-fraction of those copies of F_- will form a copy of F which contains one edge colored in the secondary color. The density of those edges is roughly $n^{v_F-2}p^{e_F} = n^{-1/m(F)}$ so we may expect them to form a copy of F. A.a.s. this is indeed the case and thus Painter looses the game after $\omega\left(n^{2-1/\overline{m}_2^2(F)}\right)$ edges have been presented.

To generalize this argument to three colors we want to show that there exist copies of F_- in the primary and secondary colors which share their missing (non)-edge. These non-edges, if they appear later, will have to be colored with the tertiary color and as before the $\overline{m}_2^r$ density is large enough to guarantee that a.a.s. Painter will have to close a copy of F in the tertiary color.

To find the aforementioned copies we prove a variant of the KŁR-conjecture, which allows us find copies of F_- where the missing edge lies in some fixed set of (non-)edges. The proof of this statement uses the container theorem introduced by Saxton and Thomason [ST15] and independently by Balogh, Morris and Samotij [BMS15].

Preliminaries and Notation

For $r \in \mathbb{N}$ let $[r] = \{1, \ldots, r\}$. For sets V, V' and $\varepsilon \in [0, 1]$ we write $V' \subseteq_\varepsilon V$ to denote that V' is a subset of V of cardinality at least $\varepsilon|V|$. We say that a statement holds asymptotically almost surely (a.a.s.) if it holds with probability $1 - o(1)$. The underlying uncolored graph of the game follows the random graph process $(G(n, N))_{1 \leq N \leq \binom{n}{2}}$, where the edges are added in an order selected uniformly at random from the $\binom{n}{2}!$ possible permutations. If $N = (1 + o(1))p\binom{n}{2}$ then the the random graph process and the random graph $G_{n,p}$ are equivalent in terms of asymptotic properties [JŁR11]. We will thus mostly work with $G_{n,p}$.

Let G be a graph and let $R \subseteq V(H)$ denote an ordered subset of the vertices. We call the pair (R, G) a rooted graph. We denote the number of vertices of G with v_G and the number of edges with e_G. For a rooted graph we set $\overline{v}_{R,G} = v_G - |R|$ and $\overline{e}_{R,G} = e_G - e_{G[R]}$. For convenience we drop the dependence on R if the set of roots is obvious from the context. That is $\overline{v}_G = \overline{v}_{R,G}$ and $\overline{e}_G = \overline{e}_{R,G}$. We write $H \subseteq_R G$ to denote that H is a subgraph of G with $R \subseteq V(H)$ and $H[R] = G[R]$. For a rooted graph (R, F) we denote with F_- the subgraph of F obtained by removing all edges of $F[R]$. We write (e, F) to indicate that the set of roots has cardinality two (this notation does not imply that $e \in E(F)$).

For two rooted graphs (R, G), (e, F) we denote with $(R, G) \times (e, F)$ the graph obtained by attaching to every non root edge e' of (R, G) a new copy of F rooted in e' (possibly removing the edge e' if it is not present in F). In general one can choose to orient the attached copies of F in two different ways. For our purposes the actual choice does not matter so we fix one based on the lexicographic ordering of the vertices.

For a collection of graphs $(R, G_1), \ldots, (R, G_k)$ which agree on R we denote with $\bigsqcup_{i=1}^k (R, G_i)$ the rooted graph obtained by joining pairwise disjoint copies of $G_1, \ldots, G_k$ together at their roots.

For graphs we define the following densities (by convention $0/0 = 0$):

$$d(G) = \frac{e_G}{v_G} \qquad\qquad m(G) = \max_{H \subseteq G} d(H)$$
$$d_1(G) = \frac{e_G}{v_G - 1} \qquad\qquad m_1(G) = \max_{H \subseteq G} d_1(H)$$
$$d_2(G) = \frac{e_G - 1}{v_G - 2} \qquad\qquad m_2(G) = \max_{H \subseteq G} d_2(H)$$
$$\overline{d}_2^r(G, H) = \begin{cases} \frac{e_H}{v_H} & \text{if } r = 1, \\ \frac{e_H}{v_H - 2 + 1/\overline{m}_2^{r-1}(G)} & \text{if } r \geq 2. \end{cases} \qquad \overline{m}_2^r(G) = \max_{H \subseteq G} \overline{d}_2^r(G, H)$$

We say that a graph G is (strictly) balanced with respect to a density function if the maximum is attained (uniquely) by G. We say that G is balanced (1-balanced, 2-balanced) if it is balanced with respect to m (m_1, m_2).

One can check (see [MSS09a]) that for every graph G we have

$$m(G) = \overline{m}_2^1(G) < \overline{m}_2^2(G) < \cdots < \overline{m}_2^r(G) < \cdots < m_2(G).$$

Furthermore if G is 2-balanced then it is also balanced with respect to $\overline{m}_2^r$ for all r. It is also easy to check that for every graph G which is not a forest

$$m_1(G) \leq \overline{m}_2^2(G),$$

and that if G is additionally 2-balanced then it is also strictly 1-balanced.

The density of a rooted graph is defined by

$$d(R, G) = \frac{\overline{e}_G}{\overline{v}_G} \qquad\qquad m(R, G) = \max_{H \subseteq G} d(R \cap V(H), H).$$

As in the unrooted case we call a rooted graph balanced if it is balanced with respect to m.

Assume that there exists $G' \subseteq G_{n,p}$ such that G' is isomorphic to $G - E(G[R])$. We say that G' is a copy of $G - E(G[R])$ in $G_{n,p}$ and that the vertices of G' which correspond to the roots of (R, G) span a copy of (R, G). Observe that the edges between root vertices are immaterial. We will make heavy use of the following upper bound due to Spencer on the number of rooted graphs spanned by vertices of the random graph.

Theorem 3.4 ([Spe90])**.** *Let (R, G) be a rooted graph and suppose that $t > m(R, G)$ and $p(n) = \Omega(n^{-1/t})$. Then a.a.s. in $G_{n,p}$ every $|R|$-tuple of vertices spans $(1 \pm o(1))\mu$ copies of (R, G) where $\mu \asymp n^{\bar{v}_G} p^{\bar{e}_G}$ is the expected number of such copies.*

If p is below the density of (R, G) then the following easy to show upper bound will suffice:

Lemma 3.5. *Suppose that (R, G) is a balanced rooted graph and that $t < m(R, G)$. Then there exists a constant $D(t)$ such that for $p \le n^{-1/t}$ with probability $1 - o(1)$ no set of $|R|$ vertices in $G_{n,p}$ spans more than D copies of G.*

Proof. Let us first prove that the balancedness of (R, G) implies that $t < d(R', G)$ for all R' with $R \subseteq R' \subsetneq V(G)$. As (R, G) is balanced we have for $q = n^{-1/m(R,G)}$

$$n^{v_G - |R'|} q^{e_G - e_{G[R']}} = \frac{n^{v_G - |R|} q^{e_G - e_{G[R]}}}{n^{|R'| - |R|} q^{e_{G[R']} - e_{G[R]}}} = \frac{\Theta(1)}{\Omega(1)} = O(1),$$

and thus $d(R', F) \ge m(R, F) > t$.

Now the probability that a fixed set of roots spans C pairwise edge disjoint copies of (R', G) is at most

$$(n^{\bar{v}_{R',G}} p^{\bar{e}_{R',G}})^C = \left(n^{-\Theta(1)}\right)^C = o\left(n^{-v_G}\right),$$

provided that C is large enough depending on t and v_G. Using the union bound we conclude that a.a.s. for every $R' \supseteq R$ no set of $|R'|$ vertices spans more than C pairwise edge disjoint copies of (R', G).

Fix a set of roots $R \subseteq V(G_{n,p})$ and a maximal set of edge disjoint copies of (R, G) spanned by R. Every other copy of (R, G) spanned by R must intersect these copies in some set $R' \supsetneq R$. By induction R' spans at most a constant number of copies of (R', G) and since the number of choices for R' is a constant the total number of copies of (R, G) spanned by R is a constant as well. □

The final density of interest is a generalization of the 2-density to rooted graphs. For $t > 0$ we define

$$d_2(R, G, t) = \begin{cases} \frac{\bar{e}_G - 1}{v_G - 2 - te_{G[R]}} & \text{if } v_G - 2 - te_{G[R]} > 0, \\ \infty & \text{otherwise.} \end{cases}$$

And

$$m_2(R,G,t) = \max_{\substack{H\subseteq G\\ e_H - e_{H[R]} > 1}} d_2(R\cap V(H), H, t).$$

The motivation for this definition will become apparent in the next section.

A KŁR type statement for rooted graphs

Fix a graph F and let $(V_i)_{i\in V(F)}$ denote pairwise disjoint sets of size n. We call a graph G on the vertex set $\cup_{i\in V(F)} V_i$ *(F,ε)-regular* if for every $\{i,j\}\in E(F)$ the pair (V_i,V_j) is (ε)-regular. We denote with $\mathcal{G}(F,n,m,\varepsilon)$ the class of (F,ε)-regular graphs G for which for every $i,j\in V(F)$

$$\left|E(V_i,V_j)\right| = \begin{cases} m & \text{if } \{i,j\}\in E(F),\\ 0 & \textit{otherwise.}\end{cases}$$

A *partite copy* of F in G is a set of vertices $\{v_i\in V_i\colon i\in V(F)\}$ such that $\{v_i,v_j\}\in E(G)$ whenever $\{i,j\}\in E(F)$. In [KŁR97] Kohayakawa, Łuczak and Rödl conjectured that almost all graphs in $\mathcal{G}(F,n,m,\varepsilon)$ contain a partite copy of F. This conjecture, known as the KŁR-conjecture, was recently proven in full by Saxton and Thomason [ST15] and independently by Balogh, Morris and Samotij [BMS15]. The following counting version is due to Saxton and Thomason.

Theorem 3.6 (KŁR conjecture, weak counting version [ST15])**.** *Let F be a graph and let $\beta>0$. There exists $\mu(\beta)>0$ such that for n sufficiently large and $m\geq \mu^{-1}n^{2-1/m_2(F)}$ the number of graphs in $\mathcal{G}(F,n,m,\mu)$ which do not contain at least $\mu(m/n^2)^{e_F}n^{v_F}$ partite copies of F is at most*

$$\beta^m\binom{n^2}{m}^{e_F}.$$

Our main tool will be a slight generalization of this theorem: we want to count only those copies of F which satisfy some additional constraints. These constraints take the form of a partite hypergraph on a subset of the vertex partitions. For a rooted graph (R,F) we denote with $\mathcal{R}(R,n)$ the class of partite $|R|$-uniform hypergraphs on the partitions $V_1,\dots,V_{|R|}$. Fix $G_R\in\mathcal{R}(R,n)$ and $G\in\mathcal{G}(F-E(F[R]),n,m,\varepsilon)$. We denote with $T(G,G_R)$ the multi-hypergraph on $V_1,\dots,V_{|R|}$ which contains an edge $e\in E(G_R)$ with multiplicity k if G contains exactly k partite copies of $F-E(F[R])$ which contain all vertices from e.

For our theorem to work we require that the edges of G_R are roughly distributed like partite copies of $F[R]$ in a random $|R|$-partite graph. This notion is formalized in the following two definitions.

Definition 3.7. *We say that $G_R\in\mathcal{R}(R,n)$ is (F,q,ε)-lower-regular if all tuples of subsets $V_1'\subseteq_\varepsilon V_1,\dots,V_{|R|}'\subseteq_\varepsilon V_{|R|}$ induce at least $q^{e_{F[R]}}\prod_{i\in R}|V_i'|$ edges.*

Definition 3.8. *We say that $G_R \in \mathcal{R}(R,n)$ is (F,q)-upper-extensible if for every induced subgraph $F' \subseteq F[R]$ the degree of all tuples from $\bigtimes_{i \in V(F')} V_i$ is at most*

$$q^{e_{F[R]}-e_{F'}} n^{|R|-v_{F'}}.$$

With these definitions at hand we can state our generalization of Theorem 3.6.

Theorem 3.9. *Let (R,F) be a rooted graph. For every $\beta > 0$, $A \geq 1$ there exist $\alpha(A,\beta), \mu(\beta) > 0$ such that for every $q(n) = o(1)$ the following holds:*

For n large enough suppose that $m \geq \alpha^{-1} n^{2-1/m_2(R,F,-\log_n q)}$ and that $G_R \in \mathcal{R}(R,n)$ is (F,Aq)-upper-extensible as well as (F,q,μ)-lower-regular. Then the number of graphs G in $\mathcal{G}(F_-,n,m,\mu)$ for which $T(G,G_R)$ contains fewer than $\alpha(m/n^2)^{e_{F_-}} q^{e_F[R]} n^{v_F}$ edges is at most

$$\beta^m \binom{n^2}{m}^{e_F - e_{F[R]}}.$$

The proof follows the proof of the KŁR conjecture presented in [ST15] and is deferred to Section 3.2.

3.1 Proof of Main Theorem

We will assume that F is a fixed 2-balanced graph which contains an edge $e \in E(F)$ such that $m_2(F-e) \leq \overline{m}_2^2(F)$. This fixes a rooted graph (e,F). Based on the choice of e we now define the classes $\mathcal{F}^1, \mathcal{F}^2, \ldots$ of rooted graphs. $\mathcal{F}^1$ consists of a singular rooted graph: an edge rooted in its endpoints. Recall that $\bigsqcup_{i=1}^k (R,G_i)$ is the rooted graph obtained by joining pairwise disjoint copies of $G_1, \ldots, G_k$ together at their roots. For $k \geq 2$ we define

$$\mathcal{F}^k := \left\{ \bigsqcup_{i<k} (e,F) \times (e_i, F_i^*) \mid \forall i \colon (e_i, F_i^*) \in \mathcal{F}^{\leq i} \right\},$$

where $\mathcal{F}^{\leq i} := \bigcup_{j \leq i} \mathcal{F}^j$.

It is useful to observe that every $F^* \in \mathcal{F}^k$, $k \geq 2$ can be built by starting with a copy of F and then repeatedly attaching a copy of (e,F) to some edge. Since F is 2-balanced this implies that F^* is 2-balanced with the same 2-density (see Lemma 3.25).

If Painter employs the greedy strategy (where she only uses color k if all colors $k' < k$ would close a monochromatic copy of F) then all edges colored with color k will span a copy of the densest graph from $\mathcal{F}^k$. We will ultimately show that, for some $F^* \in \mathcal{F}^{\leq k}$, Painter will have to color a linear fraction of all edges which span a copy of F^* with the k-color (up to a permutation of the colors). To this end let us define the notion of a *dangerous* copy of F_-^*.

Definition 3.10. *Let $F^* \in \mathcal{F}^k$. We say that an r-coloring of F^*_- is dangerous if the $k-1$ copies of F_- whose roots were identified in the construction of F^*_- are all monochromatic and colored with pairwise different colors. We say that F^*_- is dangerous if it is colored with a dangerous coloring. We say that $F \times (e, F^*_-)$ is dangerous if all attached copies of F^*_- are colored according to the same dangerous coloring.*

Assume that Painter has crated a dangerous copy of (e, F^*_-) where $F^* \in \mathcal{F}^k$. If e appears as an edge in a later round then Painter will be forced to color it with one of the remaining $r-k+1$ colors or close a monochromatic copy of F. In particular if $F^* \in \mathcal{F}^r$ and Painter creates a dangerous copy of $F \times (e, F^*_-)$ then Painter cannot color the inner copy of F (if it were to appear) without creating a monochromatic copy of F. The following Lemma states that Painter cannot avoid such dangerous copies of $F \times (e, F^*_-)$.

Lemma 3.11. *Fix a function $p = p(n)$ satisfying $n^{-1/\overline{m}_2^r(F)} \ll p \ll n^{-1/\overline{m}_2^r(F)} \log n$. Then there exists constants $c, C > 0$ such that a.a.s. after Cn^2p rounds there either exists a monochromatic copy of F or we find a graph $(e, F^*) \in \mathcal{F}^r$ such that Painter has created $cn^{v_F}\left(n^{v_{F^*_-}-2}p^{e_{F^*_-}}\right)^{e_F}$ dangerous copies of $F \times (e, F^*_-)$.*

In other words at least a constant fraction of the copies of $F \times (e, F^*_-)$ are dangerous (for some $F^* \in \mathcal{F}^r$). Assuming the above Lemma the main result follows from a second moment argument similarly to the one presented in [MSS09b]. For completeness we restate the proof below. We shall also require the following proposition, whose proof we defer to Section 3.1.2.

Proposition 3.12. *All rooted graphs $(e, F^*) \in \mathcal{F}^r$ satisfy*

$$m(F \times (e, F^*)) \leq \overline{m}_2^r(F).$$

Proof of main theorem (Theorem 3.3). We pause the game after $m = \Theta(n^2p)$ rounds. Exploiting the asymptotic equivalence between $G_{n,m}$ and $G_{n,p}$ we consider the resulting graph to be distributed like a $G_{n,p}$.

For a graph G let the random variable X_G denote the number of copies of G in $G_{n,p}$. Let F^* denote the graph guaranteed by Lemma 3.11 and define the unrooted graphs $\tilde{F} := (\emptyset, F) \times (e, F^*)$ and $\tilde{F}_- := (\emptyset, F) \times (e, F^*_-)$. By Lemma 3.11 Painter has created $M = \Omega\left(X_{\tilde{F}_-}\right)$ dangerous copies of $\tilde{F}_-$. We now consider these M copies to be fixed and for $i \in [M]$ denote with F_i the missing inner copy of F of the i-th copy of $\tilde{F}_-$. Observe that F_i and F_j are not required to be disjoint and may in fact be identical.

Observe that if Painter is forced to color one of the F_i in a future round then she must close a monochromatic copy of F and thus loose the game. We now show that indeed a.a.s. one of the F_i appears within the next $\Theta(n^2p)$ rounds.

Let Z_i denote the event that F_i appears and let $Z = \sum_{i=1}^M Z_i$. We have

$$\mathrm{E}[Z] = Mp^{e_F} = \Omega\left(\mathrm{E}\left[X_{\tilde{F}_-}\right]\right)p^{e_F} = \Omega(\mathrm{E}[X_{\tilde{F}}]) \stackrel{(*)}{=} \omega(1),$$

where $(*)$ follows from Proposition 3.12. Furthermore

$$\begin{aligned}\mathrm{Var}[Z] &= \mathrm{E}\left[Z^2\right] - \mathrm{E}[Z]^2 \\ &= \sum_{i,j} \mathrm{E}\left[Z_i Z_j\right] - \mathrm{E}[Z_i]\mathrm{E}\left[Z_j\right] \\ &= \sum_{\substack{G \subseteq F \\ e_G \geq 1}} \sum_{\substack{i,j \\ F_i \cap F_j \sim G}} p^{2e_F - e_G} - p^{2e_F} \\ &\leq \sum_{\substack{G \subseteq F \\ e_G \geq 1}} M_G p^{2e_F - e_G},\end{aligned}$$

where M_G denotes the number of pairs of F_-^* whose (missing) inner copies of F intersect in a copy of G.

Fix $G \subseteq F$ and let H denote a graph obtained as the union of two copies of $\tilde{F}_-$ whose (missing) inner copies intersect in a copy of G. Let T denote their intersection and T_+ the graph obtained by adding the missing edges of the inner copy of G to T. Observe that $T_+ \subseteq \tilde{F}$ and thus by Proposition 3.12 $\mathrm{E}[X_{T_+}] = \omega(1)$.

We have

$$\mathrm{E}[X_H] = \Theta\left(\frac{\mathrm{E}\left[X_{\tilde{F}_-}\right]^2}{\mathrm{E}[X_T]}\right) = \Theta\left(\frac{\mathrm{E}\left[X_{\tilde{F}_-}\right]^2 p^{e_G}}{\mathrm{E}[X_{T_+}]}\right) = o\left(\mathrm{E}\left[X_{\tilde{F}_-}\right]^2 p^{e_G}\right).$$

Since for every $G \subseteq F$ the number of choices for H is constant we have (over the first $\Theta(n^2 p)$ rounds) $\mathrm{E}[M_G] = o\left(\mathrm{E}\left[X_{\tilde{F}_-}\right]^2 p^{e_G}\right)$ and thus by first moment method $M_G = o\left(\mathrm{E}\left[X_{\tilde{F}_-}\right]^2 p^{e_G}\right)$ a.a.s.

This implies (over the second set of $\Theta(n^2 p)$ rounds) $\mathrm{Var}[Z] = o\left(\mathrm{E}[Z]^2\right)$ and thus $Z \geq 1$ a.a.s. □

3.1.1 Proof of Lemma 3.11

Fix a function $p = p(n)$ which satisfies $n^{-1/\overline{m}_2^r(F)} \ll p \ll n^{-1/\overline{m}_2^r(F)} \log n$. We will divide the game into a constant number of phases. In each phase we sample a copy of the binomial random graph $G_{n,p}$ and present its edges to Painter in random order (edges already presented in a previous phase are ignored). A.a.s. in each phase at most $\Theta(n^2 p)$ edges are presented. Denote with $G_{n,p}^k$ the colored graph after k phases. We implicitly assume that $G_{n,p}^k$ does not contain a monochromatic copy of F.

As a main step in the proof we will show that for every set S of at most $r-2$ colors Painter must create a graph $G \in \mathcal{G}(K_t, \tilde{n}, m, \varepsilon)$ monochromatic in some color from $[r] \setminus S$ after a constant number of phases. In general we cannot expect that $m = \Omega(n^2 p)$ (for example a greedy Painter will only produce a single color class with this density). Instead we will require that $m = \Theta(n^{v_{F^*}} p^{e_{F^*}})$ for some $F^* \in \mathcal{F}^{|S|+1}$. To retain some control on these graphs we introduce the concept of an F^*-spanning subgraph.

Definition 3.13. *For a rooted graph $(e, F^*) \in \mathcal{F}^k$ we say that a subgraph $G \subseteq G^k_{n,p}$ is F^*-spanning if for every edge $e \in E(G)$ there exists $F^*(e) \sim F^*$ in $G^k_{n,p}$ such that*

1. *the endpoints of e are the roots of $(e, F^*(e))$,*
2. *all non root vertices of $F^*(e)$ lie outside of $V(G)$ and,*
3. *$F^*(e)$ and $F^*(e')$ are edge disjoint for all $e' \in E(G) \setminus \{e\}$.*

We shall see later that one can find an F^*-spanning subgraph that behaves like a $G_{n,q}$ with $q = n^{v_{F^*}-2} p^{e_{F^*}}$ in the sense that we obtain bounds on its maximum degree as well as exponential upper bounds on the number of edges between linear sized vertex sets.

We are now in a position to state the main Lemma of this subsection.

Lemma 3.14. *Fix a set S of at most $r-2$ colors and an integer $t \geq 2$. Then there exist a positive integer k and a constant $\delta > 0$ such that for every $\varepsilon > 0$ there exists $\eta > 0$ such that for $p = \omega\left(n^{-1/\overline{m}_2^r(F)}\right)$ a.a.s. in $G^k_{n,p}$ we find a subgraph $G \in \mathcal{G}(K_t, \tilde{n}, m, \varepsilon)$ which is monochromatic in some color from $[r] \setminus S$ and F^*-spanning in $G^k_{n,p}$ where $F^* \in \mathcal{F}^{\leq|S|+1}$, $m \geq \eta n^{v_{F^*}} p^{e_{F^*}}$ and $\tilde{n} \geq \eta n$.*

Furthermore for every choice of $\tilde{n}$, m and graphs $F^* \in \mathcal{F}^{\leq|S|+1}$ and G' with $|E(G')| = \omega(n)$ the probability that the statement nominates $\tilde{n}$, m, F^* and $G \supseteq G'$ is at most

$$\left(\frac{m}{\tilde{n}^2 \delta}\right)^{|E(G')|}.$$

It is crucial that in the probability bound we loose only a constant factor (the δ) independently of the requested regularity (as opposed to the density of G which depends on $\eta(\varepsilon)$).

The next lemma states that this is the density guaranteed by Lemma 3.14 has the right order of magnitude in the sense that if we forbid $|S|$ colors then the resulting graph should have high enough density for Painter to loose the game with $r - |S|$ colors.

Lemma 3.15. *Every $F^* \in \mathcal{F}^k$ satisfies*

$$n^{v_{F^*}-2} n^{-e_{F^*}/\overline{m}_2^r(F)} \geq n^{-1/\overline{m}_2^{r-k+1}(F)},$$

provided that $k \leq r$.

Proof. The proof proceeds by induction on k. The singular graph in $\mathcal{F}^1$ consists of a single edge and thus the statement holds for $k = 1$.

For $k \geq 2$ let $F_1^*, \dots, F_{k-1}^*$ denote the graphs used during the construction of $F^* \in \mathcal{F}^k$. Writing $p_i = n^{-1/\overline{m}_2^i(F)}$ we have

$$n^{v_{F^*}-2} p_r^{e_{F^*}} = p_r \prod_{i<k} n^{v_F - 2} \left(n^{v_{F_i^*}-2} p_r^{e_{F_i^*}} \right)^{e_F - 1} \overset{(*)}{\geq} p_r \prod_{i<k} n^{v_F-2} p_{r-i+1}^{e_F-1}, \tag{3.1}$$

where (*) follows from the induction hypothesis.

By definition of $\overline{m}_2^i$ we have

$$n^{v_F-2} p_{r-i+1}^{e_F} \geq p_{r-i}$$

and thus (3.1) is at least

$$p_r \prod_{i<k} p_{r-i}/p_{r-i+1} = p_{r-k+1}.$$

□

We will give a detailed proof of Lemma 3.14 below. Before that we will walk through the main argument and state a number of auxiliary lemmas.

Assume that (by induction) we have found graphs $G_1, G_2, \dots, G_k$ with

$$G_i \in \mathcal{G}\left(F_-, \tilde{n}, \Theta\left(n^{-1/\overline{m}_2^{r-i+1}(F)}\right), \varepsilon\right),$$

which are monochromatic in pairwise different colors. Assume furthermore that G_i is F^i-spanning for some $F^i \in \mathcal{F}^i$ and that the partitions V_a, V_b corresponding to the missing edge of $F_- = F - \{\{a,b\}\}$ are the same for all G_i. Through repeated application of Theorem 3.9 we will be able to count the number of copies of $\bigsqcup_{i \leq k}(e, F_-)$ in $\bigcup_{i \leq k} G_i$ (where the i-th copy of F_- is to be from G_i).

We expect to find roughly

$$n^2 \prod_{i \leq k} n^{v_F-2} n^{-(e_F-1)/\overline{m}_2^{r-i+1}(F)} = n^2 \prod_{i \leq k} \frac{n^{-1/\overline{m}_2^{r-i}(F)}}{n^{-1/\overline{m}_2^{r-i+1}(F)}} = n^{2-1/\overline{m}_2^{r-k}(F)+1/\overline{m}_2^{r}(F)} \tag{3.2}$$

such graphs. The 2-density of F is strictly above $\overline{m}_2^r(F)$. Since F is 2-balanced we have $m_2(F) = m(e, F) = m(e, F_-)$ and Lemma 3.5 implies that every pair from $V_a \times V_b$ spans at most a constant number of copies of (e, F_-). Thus the number of pairs in $V_a \times V_b$ which span a copy of F_- in each of the graphs G_i is of the same order of magnitude as (3.2).

Out of these pairs roughly $n^{2-1/\overline{m}^{r-k}(F)}$ will appear as actual edges if we present Painter with another set of $n^2 p$ edges. If Painter wants to avoid a monochromatic copy of F then she is forced to color these edges with colors distinct from those used in $G_1, \dots, G_k$. Furthermore since all the G_i were F^i-spanning these edges all span a copy of

$$\bigsqcup_{i \leq k} (e, F) \times (e, F^i) = F^* \in \mathcal{F}^{k+1}.$$

We are below the 2-density of F^* (which equals that F) and therefore the following Lemma tells us that this edge set can be turned into an F^*-spanning subgraph by discarding a negligible number of edges.

Lemma 3.16. *Suppose that F is a 2-balanced graph and that $F_1, F_2 \sim F$ intersect in at least one, but not all edges. Then $n^{v_F} p^{e_F} \gg n^{v_{F_1 \cup F_2}} p^{e_{F_1 \cup F_2}}$ provided that $p = o\left(n^{-1/m_2(F)}\right)$.*

Proof. For a graph H write $X_H = n^{v_H} p^{e_H}$. Let $G = F_1 \cap F_2$. Since F is 2-balanced and $v_F, v_G \geq 2$ we have

$$\frac{X_F}{n^2 p} = \left(\frac{p}{n^{-1/m_2(F)}}\right)^{e_F - 1} \qquad \text{and} \qquad \frac{X_G}{n^2 p} \geq \left(\frac{p}{n^{-1/m_2(F)}}\right)^{e_G - 1}.$$

For $p = o\left(n^{-1/m_2(F)}\right)$ we obtain

$$\frac{X_F}{X_{F_1 \cup F_2}} = \frac{X_F}{\frac{X_F^2}{X_G}} = \frac{X_G}{X_F} \geq \left(\frac{p}{n^{-1/m_2(F)}}\right)^{e_G - e_F} = \omega(1),$$

as desired. □

Finally we will want to apply the sparse regularity lemma to this F^*-spanning subgraph. For this we need it to be upper-uniform, which is confirmed in the following lemma.

Lemma 3.17. *Suppose that $p = \omega\left(n^{-1/\overline{m}_2^r(F)}\right)$. Let $(e, F^*) \in \mathcal{F}^{\leq r-1}$. Then for every $\eta > 0$ a.a.s. every F^*-spanning subgraph G of $G(n,p)$ with at least ηn vertices is $\left(\eta, n^{v_{F^*}-2} p^{e_{F^*}}\right)$-upper-uniform.*

Proof. The lemma follows from the following extension of the standard Chernoff bound:

Theorem 3.18 ([PS97]). *Let $X_1, \ldots, X_n$ be a sequence of not necessarily independent Bernoulli-distributed random variables which satisfy $\Pr[\bigwedge_{i \in S} X_i] \leq q^{|S|}$ for all subsets $S \subseteq [n]$. Then for $0 < \varepsilon \leq 1$*

$$\Pr\left[\sum_{i=1}^{n} X_i \geq (1+\varepsilon) q n\right] \leq e^{-nq\varepsilon^2/3}.$$

For fixed vertex sets $V_1, V_2 \subseteq_{\eta^2} V$ let $G \subseteq G(G_{n,p})$ denote a (canonical) (e, F^*)-spanning graph in $G_{n,p}$ which maximizes the number of edges between V_1 and V_2. For $e \in E(K_n[V_1, V_2])$ let X_e denote the indicator random variable for the event $e \in G$. We have for every set S

$$\Pr\left[\bigwedge_{e \in S} X_e\right] \leq \Pr\left[S \text{ is } (e, F^*)\text{-spanning in } G_{n,p}\right] \leq \left(n^{v_{F^*}-2} p^{e_{F^*}}\right)^{|S|}.$$

And thus by Theorem 3.18

$$\Pr\left[|E_G(V_1, V_2)| \geq (1+\eta)|V_1||V_2| n^{v_{F^*}-2} p^{e_{F^*}}\right] \leq e^{-\Theta(n^{v_{F^*}} p^{e_{F^*}})}.$$

Since

$$n^{v_{F^*}} p^{e_{F^*}} \overset{\text{Lemma 3.15}}{\gg} n^{2-1/\overline{m}_2^1(F)} \geq n$$

a union bound over at most 4^n choices for V_1 and V_2 proves the Lemma. □

We thus obtain an F^*-spanning graph $G \in \mathcal{G}(K_2, \tilde{n}, \Theta(n^{v_{F^*}} p^{e_{F^*}}), \varepsilon)$. Repeating the argument a constant number of times (by exposing more edges inside one of the two partitions of G) one can obtain a monochromatic graph $G_{k+1} \in \mathcal{G}(K_t, \tilde{n}, \Theta(n^{v_{F^*}} p^{e_{F^*}}), \varepsilon)$ as required to finish the induction.

This argument can be repeated as long as $|S| \leq r-2$. One could hope to iterate one more time and find a graph $G_r \in \mathcal{G}\Big(F_-, \tilde{n}, \Theta\Big(n^{2-1/\overline{m}_2^1(F)}\Big), \varepsilon\Big)$. This approach is bound to fail. The density of G_r is (in general) not above the 2-density of F_- so we cannot hope to find copies of F_- in G_r. Instead we find graphs $G_1, \dots, G_{r-1}$, where $G_i \in \mathcal{G}\Big(F \times (e, F_-), \tilde{n}, \Theta\Big(n^{2-1/\overline{m}^{r-i+1}(F)}\Big), \varepsilon\Big)$, whose *inner* partitions agree and use Theorem 3.9 to show directly that many v_F-tuples span copies of $F \times (e, F_-)$ in all the G_i.

Before formalizing the above we need two more auxiliary lemmas. The first one asserts that the density of G_i is indeed large enough to apply Theorem 3.9.

Lemma 3.19. *Suppose that F is a 2-balanced graph, which contains an edge e that satisfies $m_2(F - \{e\}) \leq \overline{m}_2^2(F)$. Then for all $r \geq k \geq 2$*

$$\overline{m}_2^k(F) \geq m_2(e, F, +1/\overline{m}_2^k(F) - 1/\overline{m}_2^r(F)),$$
$$\overline{m}_2^k(F) \geq m_2(V(F), F \times (e, F), +1/\overline{m}_2^k(F) - 1/\overline{m}_2^r(F)).$$

Secondly Theorem 3.9 requires the (hyper)-graph to be upper-extensible. For us this hypergraph will consist of all pairs (all v_F-tuples) which already span a copy of F_- (of $F \times (e, F_-)$) in all graphs $G_1, \dots, G_i$. By Theorem 3.4 it suffices to show that we are above the rooted density of the corresponding graphs:

Lemma 3.20. *Let $F^* \in \mathcal{F}^k$ where $k \geq 2$. Then*

$$m_1(F_-^*) < \overline{m}_2^k(F)$$

and for every $V_0 \subsetneq V(F)$

$$m(V_0, (V_0, F) \times (e, F_-^*)) < \overline{m}_2^k(F).$$

We can now state the proof of Lemma 3.14.

Lemma 3.14. *Fix a set S of at most $r-2$ colors and an integer $t \geq 2$. Then there exist a positive integer k and a constant $\delta > 0$ such that for every $\varepsilon > 0$ there exists $\eta > 0$ such that for $p = \omega\Big(n^{-1/\overline{m}_2^r(F)}\Big)$ a.a.s. in $G_{n,p}^k$ we find a subgraph $G \in \mathcal{G}(K_t, \tilde{n}, m, \varepsilon)$ which is monochromatic in some color from $[r] \setminus S$ and F^*-spanning in $G_{n,p}^k$ where $F^* \in \mathcal{F}^{\leq |S|+1}$, $m \geq \eta n^{v_{F^*}} p^{e_{F^*}}$ and $\tilde{n} \geq \eta n$.*

Furthermore for every choice of $\tilde{n}$, m and graphs $F^* \in \mathcal{F}^{\leq |S|+1}$ and G' with $|E(G')| = \omega(n)$ the probability that the statement nominates $\tilde{n}$, m, F^* and $G \supseteq G'$ is at most

$$\left(\frac{m}{\tilde{n}^2 \delta}\right)^{|E(G')|}.$$

Proof. The proof follows by induction on $|S|$ and t. For $t = 2$ and $S = \emptyset$ we apply the sparse regularity lemma (Theorem 2.4) to the majority color class (in that case $F^* = (e, e)$ is just an edge - the unique graph in $\mathcal{F}^1$).

t **step:** Fix $t > 2$ and a set S. We will apply the induction hypothesis (for $t \leftarrow 2$ and $S \leftarrow S$) $K = rt|\mathcal{F}^{\leq |S|+1}|$ times. Let k', δ' denote the absolute constants guaranteed for $t \leftarrow 2$. Denote with ε_i the value which we will use for ε in the i-th application of the induction hypothesis and let $\eta_i(\varepsilon_i)$ denote the guaranteed constant. Our choice for ε_i will depend only on the constants ε_j, η_j where $j > i$ and on the requested ε.

Apply the induction hypothesis once to $G^{k'}_{n,p}$ for $t \leftarrow 2$ and obtain an (ε_1)-regular graph $G_1 \in \mathcal{G}(K_2, \tilde{n}_1, m_1, \varepsilon_1)$. Let $V_1 \subset V(G)$ denote one of its vertex partitions. We then ask painter to color another $G^{k'}_{n,p}$ but we look only at the subgraph induced by V_1 (which is distributed like a $G^{k'}_{|V_1|,p}$). Since $|V_1| \geq \eta_1 n$ we have $p = \omega\left(|V_1|^{-1/\overline{m}_2^r(F)}\right)$ and thus we can apply the induction hypothesis a second time to obtain an (ε_2)-regular graph G_2 whose edges are fully contained in V_1. We repeat this procedure K times and obtain a sequence of sets $V_1 \supset V_2 \supset \cdots \supset V_K$ and nested graphs $G_1, \ldots, G_K$, where $G_i \in \mathcal{G}(K_2, \tilde{n}_i, m_i, \varepsilon_i)$. Every such G_i nominates a color and a graph $F^*_i \in \mathcal{F}^{\leq |S|+1}$ out of at most $r|F_{\leq |S|+1}|$ choices. By the pigeonhole principle we may thus fix a subset $T \subseteq [K]$ of size t such that all graphs G_i with $i \in T$ nominate the same color and the same graph $F^* \in \mathcal{F}^{\leq |S|+1}$.

Let $\tilde{n} := \tilde{n}_K$ denote the size of the vertex partitions of G_K. We arbitrarily pick sets $\overline{V_i}$ of size $\tilde{n}$ such that

$$\begin{aligned} \overline{V_1} &\subseteq V(G_1) \setminus V_1, \\ &\vdots \\ \overline{V_K} &\subseteq V(G_K) \setminus V_K. \end{aligned}$$

Finally set $\overline{V_{K+1}} = V_K$. These sets are pairwise disjoint and for every pair $i < j$ we have $\overline{V_j} \subseteq V_{j-1} \subseteq V_i$. Therefore the sets $\overline{V_i}, \overline{V_j}$ are subsets of the two partitions of G_i and for ε_i small enough, depending on ε, $\eta_{i+1}, \ldots, \eta_K$, the induced bipartite graph $G_i\left[\overline{V}_i, \overline{V}_j\right]$ is $(\varepsilon/2)$-regular with at least half the density. Let

$$m = \min_{\substack{i,j \in T \\ i<j}} \left|E\left(G_i\left[\overline{V}_i, \overline{V}_j\right]\right)\right|$$

and pick for every $i, j \in T$, $i < j$ a subgraph $G_{i,j} \subset G_i\left[\overline{V}_i, \overline{V}_j\right]$ with exactly m edges u.a.r. among all subgraphs with m edges. By Lemma 3.15 we have

$$m = \Omega(n^{v_{F^*}} p^{e_{F^*}}) \gg n^{2-1/\overline{m}_2^{r-|S|+1}(F)} \geq n$$

and thus these graphs $G_{i,j}$ will be (ε)-regular with high probability. Since

$$\tilde{n} \geq n \prod_{i \in [K]} \eta_i \quad \text{and} \quad m \geq \eta_k \tilde{n}^{v_F^*} p^{e_F^*}$$

we may set

$$G := \bigcup_{\substack{i,j \in T \\ i<j}} G_{i,j} \in \mathcal{G}(K_t, \tilde{n}, m, \varepsilon).$$

Furthermore we claim that the graphs $F^*(e)$ guaranteed by the invocations of the induction hypothesis are pairwise edge disjoint. This is because for $e \in E(G_i)$ the graph $F^*(e)$ has no edges inside V_i, but for $j > i$ the graphs $F^*(e')$, $e' \in E(G_j)$ lie completely inside V_i. Thus G is F^*-spanning.

Finally we have to calculate the probability that $G' \subseteq G$ for some graph G' with $\omega(n)$ edges. To do so fix $\tilde{n}$, m and the sets $\overline{V}_i$ among $2^{\Theta(n)}$ possibilities. We may assume that all edges of G' go between the sets $\overline{V}_i$. Since G is the (disjoint) union of random subgraphs of $G_i[\overline{V}_i, \overline{V}_j]$ and $\left|E(G_i[\overline{V}_i, \overline{V}_j])\right| \geq \tilde{n}^2(m_i/2\tilde{n}_i^2)$whose density is at least half as large as the density of G_i the probability that $G' \subseteq G$ is then at most

$$\begin{aligned}
&\prod_{i<j \in T} \Pr\left[G'[\overline{V}_i, \overline{V}_j] \subseteq G_i[\overline{V}_i, \overline{V}_j]\right] \left(\frac{m}{\left|E(G_i[\overline{V}_i, \overline{V}_j])\right|}\right)^{\left|E(G'[\overline{V}_i, \overline{V}_j])\right|} \\
&\leq \prod_{i<j \in T} \left(\frac{m_i}{\tilde{n}_i^2 \delta'} \frac{m}{\left|E(G_i[\overline{V}_i, \overline{V}_j])\right|}\right)^{\left|E(G'[\overline{V}_i, \overline{V}_j])\right|} \\
&\leq \prod_{i<j \in T} \left(\frac{2m}{\delta' \tilde{n}^2}\right)^{\left|E(G'[\overline{V}_i, \overline{V}_j])\right|} = \left(\frac{2m}{\delta' \tilde{n}^2}\right)^{|E(G')|}.
\end{aligned}$$

Allowing some room for a union bound over the choices for $\tilde{n}$, m and the sets $\overline{V}_i$ we may fix $\delta = \delta'/3$, $\eta = \eta_K \prod_{i \in [K]} \eta_i^{v_{F^*}}$ and $k = Kk'$.

S **step:** Fix a nonempty set S of at most $r-2$ colors and assume that the statement holds for all sets containing fewer than $|S|$ colors. Our goal is to show that then the statement holds for S and $t = 2$.

Similarly to what we did in the induction step for t we apply the induction hypothesis $|S|$ times in a nested fashion. As before let k', δ' denote absolute constants for which the induction hypothesis holds for $t \leftarrow v_F - 1$ and all subsets of S. Denote with ε_i the value which we will use for ε in the i-th application of the induction hypothesis and let $\eta_i(\varepsilon_i)$ denote the guaranteed constant. Again ε_i will depend only on ε_j, η_j, where $j > i$. Crucially ε_i will not depend on the requested ε and $\varepsilon_{|S|}$ will be an absolute constant. Finally for the i-th invocation we will pick S as the set of colors of $G_1, \ldots, G_{i-1}$ (thus $S \leftarrow \emptyset$ for $i = 1$).

As before we obtain a collection of monochromatic graphs $G_1, \dots, G_{|S|}$, such that $G_i \in \mathcal{G}(K_{v_F-1}, \tilde{n}_i, m_i, \varepsilon_i)$ and $V(G_i) \subseteq V_{i-1}$ for $V_0 = V$ and where V_i is an arbitrary partition of G_i.

Assume that one of the G_i is monochromatic in a color from $[r] \setminus S$. The density of G_i is in $\Theta\left(n^{v_{F_i^*}-2} p^{e_{F_i^*}}\right)$, where the constant does not depend on ε (since $\varepsilon_1, \dots, \varepsilon_{|S|}$ do not depend on ε). Furthermore by Lemma 3.17 G_i is $\left(o(1), n^{v_{F_i^*}-2}(|S|k'p)^{e_{F_i^*}}\right)$-upper-uniform. Thus we can apply the sparse regularity lemma (Theorem 2.4) to G_i and obtain a graph from $\mathcal{G}(K_2, \tilde{n}, m, \varepsilon)$ whose density is of the same order as the density of G_i and we are done.

Otherwise all of the G_i are monochromatic in distinct colors of S. We want to show that in $V_{|S|}$ there are many pairs of vertices which span a copy of (e, F_-) in each of the G_i. To this end we define the auxiliary directed graphs A_i. Let A_0 denote the complete directed graph on V. For $i = 1, \dots, |S|$ the vertex set of A_i is V_i and we connect two vertices $x \neq y \in V_i$ if $(x, y) \in E(A_{i-1})$ and if (x, y) span a partite copy of F_- in G_i (partite with respect to the non root vertices, x and y lie in the same partition).

Define $(e, F^*) := \bigsqcup_{j \in [|S|]} (e, F) \times (e, F_i^*) \in \mathcal{F}^{|S|+1}$. By definition of $A_{|S|}$ every edge $e \in E(A_{|S|})$ spans a copy of (e, F_-) in each of the G_i. Furthermore since every edge of G_i spans a copy of (e, F_i^*) the edge e spans a copy of (e, F_-^*). Finally observe that since the G_i are monochromatic in pairwise different colors of S the edge e (if it would be presented to Painter in some later round) has to be colored with some color from $[r] \setminus S$.

We will need the following auxiliary claim about the density of $A_{|S|}$ whose proof we defer.

Claim 3.21. *For every integer $i \leq |S|$ and $\kappa > 0$ and small enough $\varepsilon_1, \dots, \varepsilon_i$ there exists $\gamma(\kappa, \varepsilon_1, \dots, \varepsilon_i) > 0$ such that a.a.s. for all disjoint and equi-sized subsets $X, Y \subseteq_\kappa V_i$ the induced bipartite subgraph $A_i[X, Y]$ contains at least $\gamma |X||Y| \prod_{j=1}^{i} p_j$ edges, where*

$$p_j = n^{v_F - 2}\left(n^{v_{F_j^*}-2} p^{e_{F_j^*}}\right)^{e_F - 1}.$$

We invoke the claim for $i \leftarrow |S|$ and $\kappa \leftarrow 1/4$ to lower bound the number of pairs $x, y \in V_{|S|}$ which span a dangerous copy of F^* by

$$\frac{\gamma}{2}\left(\frac{|V_{|S|}|}{2}\right)^2 \prod_{j=1}^{|S|} p_i \geq \gamma' n^{v_{F^*}} p^{e_{F^*}-1} \overset{\text{Lemma 3.15}}{\gg} \frac{n^{2-1/\overline{m}_2^{r-|S|}(F)}}{p} \gg \frac{n}{p},$$

where γ is the constant guaranteed by the claim and $\gamma' = \gamma(\prod_i \eta_i)^2/8$ is an absolute constant, which in particular does not depend on ε.

We then present another $G_{n,p}$ to Painter. Painter will be forced to color at least a p/r-fraction of the edges in $A_{|S|}$ with some color from $[r] \setminus S$ (or create a monochromatic copy of F). Thus we obtain a monochromatic set of $\gamma' n^{v_{F^*}} p^{e_{F^*}}/r$ edges which all span a copy of F^*. Next we remove all edges whose copies of F^* intersect. Lemma 3.16 together with Markov's inequality implies that with probability $1 - o(1)$ we remove only $o(n^{v_{F^*}} p^{e_{F^*}})$ edges.

So we are left with a (e, F^*)-spanning set of at least $\gamma' n^{v_{F^*}} p^{e_{F^*}}/2 \gg n$ edges E'. By Lemma 3.17 E' is $\big(o(1), n^{v_{F^*}-2}((|S|k'+1)p)^{e_{F^*}}\big)$-upper-uniform. Therefore we may apply the sparse regularity lemma to E' and obtain a graph from $\mathcal{G}(K_2, \tilde{n}, m, \varepsilon)$ whose density is a constant fraction of the density of E'.

Finally the probability that a fixed set of s edges is (e, F^*) spanning is at most $\big(n^{v_{F^*}-2}((|S|k'+1)p)^{e_{F^*}}\big)^s$. Since $m/\tilde{n}^2 = \Omega\big(n^{v_{F^*}-2} p^{e_{F^*}}\big)$ (not depending on ε) the probability bound holds for some δ.

It remains to prove Claim 3.21. We proceed by induction on i. A_0 is complete and thus the base case $i = 0$ holds vacuously. So let $i \geq 1$ and fix some $\kappa > 0$. Denote the lower bound on the density of A_{i-1} guaranteed by the induction by

$$q = \Omega\left(\prod_{j<i} p_j\right) \gg \frac{n^{-1/\overline{m}_2^{r-i+1}(F)}}{p}.$$

We define $\beta = (\delta'/(3e))^{e_{F_-}}$ and

$$A = \frac{\prod_{j=1}^{i-1} p_j}{q\kappa \prod_{j=1}^{i} \eta_j} = \Theta(1).$$

Let $\mu(\beta)$ and $\alpha(A, \beta)$ denote the constant guaranteed by Theorem 3.9 when invoked with $(R, F) \leftarrow (e, F)$, $\beta \leftarrow \beta$ and $A \leftarrow A$. Fix disjoint equi-sized sets $X, Y \subseteq_\kappa V_i$. Write $m = \lceil |X||Y|m_i/(2\tilde{n}_i^2) \rceil$ and pick any subgraph $G_i' \subseteq G_i$ from $\mathcal{G}(F_-, |X|, m, \mu)$ such that its partitions which correspond to the roots of F_- are X and Y (taking suitable vertex sets and a random subset of m edges from each partition succeeds with probability $1 - 2^{-\Theta(m)}$).

If $T(G', A_{i-1}[X, Y])$ contains at least

$$\alpha\left(\frac{m}{|X|^2}\right)^{e_{F_-}} |X|^{v_F} q = \Theta\left(n^2 \prod_{j \leq i} p_j\right)$$

edges, then since by Lemma 3.5 every edge e spans at most a constant number of copies of F_- the density of $A_i[X, Y]$ is of the correct order of magnitude.

Otherwise we want to apply Theorem 3.9 with $G \leftarrow G'$ and $G_R \leftarrow A_{i-1}[X, Y]$ (viewed as an undirected graph) and $q \leftarrow q$. To apply Theorem 3.9 it suffices to check the following:

1. The number of edges in G' is in

$$\Omega\left(n^{v_{F_i^*}} p^{e_{F_i^*}}\right) \overset{\text{Lemma 3.15}}{\gg} n^{2-1/\overline{m}_2^{r-i+1}(F)} \overset{\text{Lemma 3.19}}{\geq} n^{2-1/m_2(e,F,-\log_n q)},$$

   since for n large enough $-\log_n q \leq 1/\overline{m}_2^{r-i+1}(F) - 1/\overline{m}_2^{r}(F)$.

2. If we invoke the induction hypothesis with say $\kappa \leftarrow \kappa\eta_i\mu$ then $A_{i-1}[X, Y]$ is (F, μ, q)-lower-regular.
3. To see that it is also (F, Aq)-upper-extensible observe that every edge of A_{i-1} spans a copy of of $(e, F'^*_-) := \bigsqcup_{j<i}(e, F_-) \times (e, F^*_j)$. So for upper uniformity it suffices to bound the number of copies of F'^*_- spanned by a single vertex. But our p is such that we are above the 1-density of F'^*_- (Lemma 3.20). Thus this number is concentrated around its expectation (Theorem 3.4), which is upper bounded by

$$n^{v_{F'^*_-}-1}p^{e_{F'^*_-}} = n\prod_{j<i} p_j \leq |X|\frac{\prod_{j<i} p_j}{\kappa\prod_{j\leq i}\eta_j} = Aq|X|.$$

Therefore we can apply Theorem 3.9 and G' must come from a set of at most

$$\beta^m\binom{|X|^2}{m}^{e_{F_-}} \leq \beta^m\left(\frac{e|X|^2}{m}\right)^{e_{F_-}m} \leq \beta^m\left(\frac{2e\tilde{n}_i^2}{m_i}\right)^{e_{F_-}m}$$

graphs. But then for our choice of β the probability that $G' \subseteq G_i$ is in $o(1)$.

□

The proof of Lemma 3.11 proceeds similarly to the proof of Claim 3.21 from the previous lemma. The only difference is that we replace A_i with the hypergraph of v_F-tuples which span a copy of $F \times (e, F_-)$ in each of the graphs G_i.

Lemma 3.11. *Fix a function $p = p(n)$ satisfying $n^{-1/\overline{m}_2^r(F)} \ll p \ll n^{-1/\overline{m}_2^r(F)}\log n$. Then there exists constants $c, C > 0$ such that a.a.s. after Cn^2p rounds there either exists a monochromatic copy of F or we find a graph $(e, F^*) \in \mathcal{F}^r$ such that Painter has created $cn^{v_F}\left(n^{v_{F^*_-}-2}p^{e_{F^*_-}}\right)^{e_F}$ dangerous copies of $F \times (e, F^*_-)$.*

Proof. Let $t = v_F - 1$ and let k, δ be such that Lemma 3.14 holds for all $S \subseteq [r]$, $|S| \leq r-2$. Let $\varepsilon_1, \ldots, \varepsilon_{r-1}$ denote constants whose value we will determine later in reverse order (that is ε_i will depend on $\varepsilon_{i+1}, \ldots, \varepsilon_{r-1}$.)

We ask Painter to color an instance of $G^k_{n,p}$. Applying Lemma 3.14 with $t \leftarrow t$, $S \leftarrow \emptyset$, $\varepsilon \leftarrow \varepsilon_1$ we obtain a constant $\eta_1(\varepsilon_1)$, a graph $F^*_1 \in \mathcal{F}^1$ and an F^*_1-spanning graph $G_1 \in \mathcal{G}(K_t, \tilde{n}_1, m_1, \varepsilon_1)$ monochromatic in some color s_1. Pick one of the vertex partitions of G_1 arbitrarily and call it V_1. We now present Painter with a second instance of $G^k_{n,p}$ but only consider the subgraph induced by V_1 which is distributed like a $G^k_{\tilde{n}_1,p}$. Invoking Lemma 3.14 a second time with $\varepsilon \leftarrow \varepsilon_2$ and $S \leftarrow \{s_1\}$ we obtain a second graph $G_2 \in \mathcal{G}(K_t, \tilde{n}_2, m_2, \varepsilon_2)$. We repeat this procedure $r-1$ times and obtain

1. sets $V = V_0 \supset V_1 \supset \cdots \supset V_{r-1}$ such that $|V_i| \geq \eta_i|V_{i-1}|$ for $i \in [r-1]$,
2. graphs $F^*_i \in \mathcal{F}^{\leq i}$ where $i \in [r-1]$,

3. monochromatic graphs $G_i \subseteq \mathcal{G}(K_t, \tilde{n}_i, m_i, \varepsilon_i) \subseteq G_{n,p}^{i\cdot k}[V_{i-1}]$ in pairwise different colors, where $\tilde{n}_i = |V_i|$, $m_i \geq \eta_i \tilde{n}^{v_{F_i^*}} p^{e_{F_i^*}}$ such that G_i is F_i^*-spanning in $G_{n,p}^{i\cdot k}[V_{i-1}]$.

Furthermore for every graph G' with $\omega(n)$ edges we have

$$\Pr\left[G' \subseteq G_i\right] \leq \left(\frac{m_i}{\delta \tilde{n}_i^2}\right)^{|E(G')|}.$$

Observe that this probability is over the phases $(i-1)\cdot k+1, \ldots, i\cdot k$ and that G_i is fixed after the first $i \cdot k$ phases.

Let A_0 denote the complete directed v_F-uniform hypergraph on $V = V_0$. We identify the edges of A_0 with a (hypothetical) copy of F in V (depending on the automorphisms of F different (directed) edges might represent the same copy of F). Now for $i \in [r-1]$ let A_i denote the directed v_F-uniform hypergraph on $V_i \subseteq V_{i-1}$ where $e \in E(A_i)$ if $e \in E(A_{i-1})$ and additionally the edges of e (when viewed as a graph $F(e) \sim F$) are the roots of pairwise edge disjoint copies of F_- in G_i.

Define

$$\begin{aligned}
\tilde{F}_i^* &= \bigsqcup_{j<i} (e, F) \times (e, F_j^*) \in \mathcal{F}^i, \\
p_i &= n^{v_F - 2}\left(n^{v_{F_i^*}-2} p^{e_{F_i^*}}\right)^{e_F - 1}, \\
q_i &= \prod_{j \in [i]} p_i = n^{v_{\tilde{F}_{i+1}^*}-2} p^{e_{\tilde{F}_{i+1}^*}-1}.
\end{aligned}$$

Since G_i is F_i^*-spanning the edges of $F(e)$ (for $e \in A_i$) are not only roots of pairwise edge disjoint copies of F_- but even of (still pairwise edge disjoint) copies of $F_- \times (e, F_i^*)$. Furthermore these copies are disjoint from those certifying membership in $A_1, \ldots, A_{i-1}$. Thus for every $e \in E(A_i)$ the edges of $F(e)$ are the roots of pairwise edge disjoint copies of $\left(\tilde{F}_{i+1}\right)_-$.

Claim 3.22. *Suppose that $X_1, \ldots, X_{v_F} \subseteq V_i$ are mutually disjoint and of size $\tilde{n}$. Then $A_i[X_1, \ldots, X_{v_F}]$, when viewed as a undirected hypergraph from $\mathcal{R}(V(F), \tilde{n})$, is $(F, (n/\tilde{n})^{v_F} q_i)$-upper-extensible provided that $i \leq r-1$.*

Proof of claim. $q_0 = 1$ and thus the claim holds vacuously for $i = 0$. For $i \geq 1$ fix some $\sigma \subseteq V(A_i)$. If there exists e such that $\sigma \subsetneq e \in E(A_i)$ then σ fixes some $V' \subsetneq V(F)$ and the degree of σ is at most the number of copies of $(V', F) \times (e, \left(\tilde{F}_{i+1}^*\right)_-)$ rooted in σ. By Lemma 3.20

$$m(V', (V', F) \times (e, \left(\tilde{F}_{i+1}^*\right)_-)) < \overline{m}_2^{i+1}(F) \leq \overline{m}_2^r(F)$$

and thus by Theorem 3.4 this number is concentrated around its expectation which is at most $n^{v_F - |V'|} q_i^{e_F - e_{F[V']}}$ as required for the upper-extensibility of A_i. □

Claim 3.23. *For every integer $i \leq r-1$ and $\kappa > 0$ and small enough $\varepsilon_1, \ldots, \varepsilon_i$ there exists $\gamma(\kappa, \varepsilon_1, \ldots, \varepsilon_i) > 0$ such a.a.s. for all pairwise disjoint equi-sized sets $X_1, \ldots, X_{v_F} \subseteq_\kappa V_i$ the number of directed edges in $E(A_i[X_1, \ldots, X_{v_F}])$ is at least $\gamma |X_1|^{v_F} q_i^{e_F}$.*

Invoking the claim for $i = r-1$ and say $\kappa = 1/(2v_F)$ proves the Lemma.

To prove the claim we will proceed by induction on i. For $i = 0$ the statement holds vacuously with $\gamma = 1$ since A_0 is complete. For $i \geq 1$ we fix equi-sized and pairwise disjoint sets $X_1, \ldots, X_{v_F} \subseteq_\kappa V_i$ of size $\tilde{n}$. Define

$$m = \left\lceil \frac{\tilde{n}^2 m_i}{2 e_F \tilde{n}_i^2} \right\rceil \qquad \text{and} \qquad \beta = \left(\frac{\delta}{4 e_F e} \right)^{e_F \cdot (e_F - 1)}.$$

Let $\mu = \mu(\beta)$ be given by Theorem 3.9 (invoked with $(R, F) \leftarrow (V(F), F \times (e, F_-))$) and let $\gamma' = \gamma(\kappa \eta_i \mu, \varepsilon_1, \ldots \varepsilon_{i-1})$ denote the constant guaranteed by the induction hypothesis. Finally define

$$A = \frac{1}{\gamma' \left(\kappa \prod_{j \in [i]} \eta_j \right)^{v_F}}.$$

$F \times (e, F_-)$ is t-partite and thus for ε_i small enough depending on μ we can, using standard techniques, find a graph $G_i' \subseteq G_i$ from $\mathcal{G}(F \times (e, F_-), \tilde{n}, m, \mu)$ such that the vertex partitions corresponding to to the vertices of the (missing) inner copy of F are the sets $X_1, \ldots, X_{v_F}$.

If the number of edges in $T(G_i', A_{i-1}[X_1, \ldots, X_{v_F}])$ is at least

$$\Theta\left(\left(\frac{m}{\tilde{n}^2} \right)^{e_F \cdot (e_F - 1)} \tilde{n}^{v_F + e_F \cdot (v_F - 2)} q_{i-1}^{e_F} \right) = \Theta\left(p_i^{e_F} n^{v_F} q_{i-1}^{e_F} \right) = \Theta\left(n^{v_F} q_i^{e_F} \right),$$

then we are done, since by Lemma 3.25 $(V(F), F \times (e, F_-))$ is balanced with density $m_2(F)$ and thus Lemma 3.5 the multiplicity of all edges is at most a constant.

Otherwise we invoke the induction hypothesis with $\kappa \leftarrow \kappa \eta_i \mu$ to deduce that $A_{i-1}[X_1, \ldots, X_{v_F}]$ is (F, q, μ)-lower-regular where $q = \gamma' q_{i-1}$. We have

$$\left(\frac{n}{\tilde{n}} \right)^{v_F} q_{i-1} \leq \frac{q_{i-1}}{(\kappa \prod_{j \in [i]} \eta_j)^{v_F}} = Aq$$

and therefore by Claim 3.22 the hypergraph $A_{i-1}[X_1, \ldots, X_{v_F}]$ is (F, Aq)-upper-extensible. Finally

$$m = \Omega\left(n^{v_{F_i^*}} p^{e_{F_i^*}} \right) \overset{\text{Lemma 3.15}}{\gg} n^{2 - 1/\overline{m}_2^{r-i+1}(F)} \overset{\text{Lemma 3.19}}{\geq} n^{2 - 1/m_2(V(F), F \times (e, F), q_i)}$$
$$\asymp n^{2 - 1/m_2(V(F), F \times (e, F), q)}$$

and thus we may apply apply Theorem 3.9 with $q \leftarrow q$, $G \leftarrow G_i'$ and $G_R \leftarrow$

$A_{i-1}[X_1, \ldots, X_{v_F}]$ to deduce that G_i' is from a set graphs of size at most

$$\begin{aligned}\beta^m \binom{\tilde{n}^2}{m}^{e_F \cdot (e_F - 1)} &\leq \beta^m \left(\frac{e\tilde{n}^2}{m}\right)^{m \cdot e_F \cdot (e_F - 1)} \\ &\leq \beta^m \left(\frac{e2e_F\tilde{n}_i^2}{m_i}\right)^{m \cdot e_F \cdot (e_F - 1)} \\ &= \left(\frac{\tilde{n}_i^2}{2\delta m_i}\right)^{m \cdot e_F \cdot (e_F - 1)},\end{aligned}$$

Since $m \gg n$ a union bound over the $2^{\Theta(n)}$ choices for the sets $X_1, \ldots, X_{v_F}$ together with the bound

$$\Pr\left[G' \subseteq G_i\right] \leq \left(\frac{m_i}{\tilde{n}_i^2 \delta}\right)^{|E(G')|}$$

from Lemma 3.14 guarantees that a.a.s. no such subgraph G_i' exists. □

3.1.2 Auxiliary Lemmas

In Section 3.1 we stated a number of auxiliary statements without proof (namely Proposition 3.12, Lemma 3.19 and Lemma 3.20). The proofs of these statements are somewhat technical and are given in this section.

We start with the proof of Lemma 3.19 for which we need the following simple bound.

Lemma 3.24. *Every graph F on at least 3 vertices and with at least one edge satisfies $m_2(F) \leq 2m(F)$.*

Proof. One checks that the statement holds for all graphs on 3 vertices. If F is 2-balanced and contains at least 4 vertices then

$$m_2(F) = \frac{e-1}{v-2} \leq \frac{e}{v-2} \leq 2\frac{e}{v} = 2d(F) \leq 2m(F).$$

Otherwise let $F' \subseteq F$ denote a graph that attains the m_2 density of F. By the above

$$m_2(F) = m_2(F') \leq 2m(F') \leq 2m(F).$$

□

Lemma 3.19. *Suppose that F is a 2-balanced graph, which contains an edge e that satisfies $m_2(F - \{e\}) \leq \overline{m}_2^2(F)$. Then for all $r \geq k \geq 2$*

$$\begin{aligned}&\overline{m}_2^k(F) \geq m_2(e, F, +1/\overline{m}_2^k(F) - 1/\overline{m}_2^r(F)), \\ &\overline{m}_2^k(F) \geq m_2(V(F), F \times (e, F), +1/\overline{m}_2^k(F) - 1/\overline{m}_2^r(F)).\end{aligned}$$

Proof. Write $p = n^{-1/\overline{m}_2^k(F)}$ and $q = n^{-1/\overline{m}_2^k(F)+1/\overline{m}_2^r(F)}$. The first inequality is equivalent to

$$\min_{\substack{(R,F')\subseteq(e,F)\\ \bar{e}_{F'}\geq 1}} n^{v_{F'}-2}p^{\bar{e}_{F'}-1}q^{e_{F'[R]}} \geq 1.$$

Fix such a rooted graph (R, F'). If $|R| \leq 1$ then $F' \subseteq F - \{e\}$ and $e_{F'[R]} = 0$. Since $p \geq n^{-1/\overline{m}_2^2(F)} \geq n^{-1/m_2(F-\{e\})}$ we have

$$n^{v_{F'}-2}p^{e_{F'}-1} \geq 1.$$

Otherwise $|R| = 2$ and without loss of generality $e_{F'[R]} = 1$ and $\bar{e}_{F'} = e_{F'} - 1$. We rewrite the above as

$$n^{v_{F'}-2}p^{e_{F'}-2}q = \frac{n^{v_{F'}-2}p^{e_{F'}}}{pn^{-1/\overline{m}_2^r(F)}} \geq \frac{n^{v_{F'}-2}p^{e_{F'}}}{n^{-2/m_2(F)}}.$$

By definition of $\overline{m}_2^k$ we have $n^{v_{F'}-2}p^{e_{F'}} \geq n^{-1/\overline{m}_2^{k-1}(F)}$. Together with

$$n^{2/m_2(F)} \overset{\text{Lemma 3.24}}{\geq} n^{1/m(F)} \geq n^{1/\overline{m}_2^{k-1}(F)}$$

this implies the desired bound.

The second inequality is equivalent to

$$\min_{\substack{(R,H)\subseteq(V(F),F\times(e,F))\\ \bar{e}_H\geq 1}} n^{v_H-2}p^{\bar{e}_H}q^{e_{H[R]}} \geq 1.$$

For $e \in E(F)$ let $F_e \subseteq H$ denote the graph isomorphic to a subgraph of F which is attached to the root e in H. There must exist at least on edge e' such that $F_{e'}$ contains at least one non root edge. For such an edge we apply the first inequality to obtain

$$n^{v_{F_{e'}}-2}p^{e_{F_{e'}}-e_{F_{e'}[e']}-1}q^{e_{F_{e'}[e']}} \geq 1.$$

Thus the minimization is at least

$$n^{v_{H[R]-2}}\prod_{e\in E(F)\setminus\{e'\}} n^{v_{F_e}-v_{F_e[e]}}p^{e_{F_e}-e_{F_e[e]}}q^{e_{F_e[e]}}.$$

If for some $e \in E(F)$ we have $v_{F_e[e]} \leq 1$ then F_e is isomorphic to a subgraph of F_- and

$$n^{v_{F_e}-v_{F_e[e]}}p^{e_{F_e}-e_{F_e[e]}}q^{e_{F_e[e]}} \geq n^{v_{F_e}-1}p^{e_{F_e}} \geq 1,$$

since we are above the 1-density of F_-. In particular if H does not contain at least two root vertices then we are done. If $v_{F_e[e]} = 2$ then

$$\begin{aligned} n^{v_{F_e}-v_{F_e[e]}}p^{e_{F_e}-e_{F_e[e]}}q^{e_{F_e[e]}} &\geq n^{v_{F_e}-2}p^{e_{F_e}-1}q = \frac{n^{v_{F_e}-2}p^{e_{F_e}}}{n^{-1/\overline{m}_2^r(F)}} \\ &\geq \frac{n^{-1/\overline{m}_2^{k-1}(F)}}{n^{-1/\overline{m}_2^r(F)}} \geq \frac{n^{-1/m(F)}}{n^{-1/m_2(F)}} \geq n^{-1/m_2(F)}. \end{aligned}$$

Thus the original minimization reduces to

$$\min_{\substack{F' \subseteq F \\ v_{F'} \geq 2}} n^{v_{F'}-2} n^{-(e_{F'}-1)/m_2(F)}$$

which is at least 1 by definition of $m_2(F)$. □

Proposition 3.12 concern the density of graphs in the class $\mathcal{F}^k$. Every graph $F^* \in \mathcal{F}^k$ (for $k \geq 2$) can be constructed by starting with a copy of F and repeatingly attaching copies of (e, F) to some edge. Since F is 2-balanced one may expect that graphs constructed by this procedue will be also be 2-balanced. The following lemma establishes that this is indeed the case.

Lemma 3.25. *Suppose that G and H are two 2-balanced graphs such that $G \cap H$ is a single edge. If G and H both have 2-density d then $G \cup H$ is also 2-balanced with density d.*

Similarly if (R, G) and (R, H) are balanced rooted graphs of density d with $V(G) \cap V(H) = R$ then $(R, G \cup H)$ is balanced with density d.

Proof. Let $p = n^{-1/d}$ and pick an induced subgraph $F \subseteq G \cup H$ with $e_F \geq 1$. Write $G' = F[V(G)]$ and $H' = F[V(H)]$. Without loss of generality we have $e_{G'} \geq 1$ and

$$n^{v_F-2} p^{e_F-1} = n^{v_{G'}-2} p^{e_{G'}-1} n^{v_{H'}-v_{H' \cap G'}} p^{e_{H'}-e_{H' \cap G'}} \geq 1,$$

since $d = m_2(H) \geq m_1(H) \geq m(H)$ and since $H' \cap G'$ is either an edge, a vertex or empty. Thus $m_2(G \cup H) \leq d$. Furthermore

$$n^{v_{G \cup H}-2} p^{e_{G \cup H}-1} = n^{v_G-2} p^{e_G-1} n^{v_H-2} p^{e_H-1} = 1 \cdot 1,$$

which implies $m_2(G \cup H) = d$ and that $G \cup H$ is balanced with respect to the 2-density.

The second claim can be proved in a similar fashion. □

Thus every $F^* \in \mathcal{F}^k$, where $k \geq 2$, is 2-balanced with 2-density $m_2(F)$. Similarly $F \times (e, F^*)$ is 2-balanced (and thus balanced)

Proposition 3.12. *All rooted graphs $(e, F^*) \in \mathcal{F}^r$ satisfy*

$$m(F \times (e, F^*)) \leq \overline{m}_2^r(F).$$

Proof. Fix $F^* \in \mathcal{F}^r$, where $r \geq 2$. As noted above $F \times (e, F^*)$ is balanced. Therefore it suffices to check that

$$\begin{aligned} n^{v_{F \times (e,F^*)}} n^{-e_{F \times (e,F^*)}/\overline{m}_2^r(F)} &= n^{v_F} \left(n^{v_{F^*}-2} n^{-e_{F^*}/\overline{m}_2^r(F)} \right)^{e_F} \\ &\overset{(3.15)}{\geq} n^{v_F} n^{-e_F/\overline{m}_2^1(F)} \geq 1. \end{aligned}$$

□

It remains to prove Lemma 3.20. To do so we require two more auxiliary lemmas.

Lemma 3.26. *Let $(R, G), (e, H)$ be rooted graphs. Suppose that (e, H) is balanced and that for some $t > 0$*

$$m_1(H - e) \leq t,$$
$$\overline{v}_H - \frac{\overline{e}_H}{t} \geq -\frac{1}{m(R, G)}.$$

Then

$$m(R, (R, G) \times (e, H - e)) \leq t.$$

Proof. Let $p = n^{-1/t}$ and let $(R, F) = (R, G) \times (e, H - e)$. It suffices to show that

$$\min_{(R,F') \subseteq (R,F)} n^{\overline{v}_{F'} - |R|} p^{\overline{e}_{F'}} \geq 1. \tag{3.3}$$

Fix a graph $(R, F') \subseteq (R, F)$ which attains the minimum. Let $H_1, \ldots, H_k \sim H - e$ denote the (canonical) copies of $H - e$ in F and write $H_i' = F' \cap H_i$. The above term can be rewritten as

$$n^{v_{F' \cap G} - |R|} \prod_{i \in [k]} n^{v_{H_i'} - v_{H_i' \cap G}} p^{e_{H_i'}}.$$

If for some i we have $v_{H_i' \cap G} \in \{0, 1\}$ then $t \geq m_1(H - e) \geq m(H - e)$ implies

$$n^{v_{H_i'} - v_{H_i' \cap G}} p^{e_{H_i'}} \geq 1.$$

We may thus assume that for such i we have $H_i' = H_i \cap G$.

Otherwise $v_{H_i' \cap G} = 2$. If $t \geq m(e, H)$ then the above bound holds as well and in particular (3.3) is satisfied. If $t < m(e, H)$ then, since (e, H) is balanced, the minimum of $n^{v_{H_i'} - 2} p^{e_{H_i'}}$ is attained for $H_i' = H_i$. Thus the minimization reduces to

$$\begin{aligned} \min_{(R,G') \subseteq (R,G)} n^{\overline{v}_{G'}} \left(n^{v_H - 2} p^{e_H - 1} \right)^{\overline{e}_{G'}} &= \min_{(R,G') \subseteq (R,G)} n^{\overline{v}_{G'}} \left(n^{v_H - 2 - (e_H - 1)/t} \right)^{\overline{e}_{G'}} \\ &\geq \min_{(R,G') \subseteq (R,G)} n^{\overline{v}_{G'}} n^{-\overline{e}_{G'}/m(R,G)} = 1. \end{aligned}$$

□

Lemma 3.27. *Let $r \geq 2$ and suppose that G is a 2-balanced graph with density at least 1. Then*

$$\max_{R \subsetneq V} m(R, G) \leq v - 1 < \left(\frac{1}{m(G)} - \frac{1}{\overline{m}_2^r(G)} \right)^{-1}.$$

Proof. $m(R, G)$ is monotone increasing under edge addition. Thus for the first inequality it suffices to consider the case $G = K_v$. We have

$$m(R, K_v) = \frac{\binom{v}{2} - \binom{|R|}{2}}{v - |R|} \leq \frac{\binom{v}{2} - \binom{v-1}{2}}{v - (v - 1)} = v - 1.$$

For the second inequality we use $\overline{m}_2^r(G) < m_2(G)$ and the fact that since G is 2-balanced it is also balanced to obtain

$$\frac{1}{m(G)} - \frac{1}{\overline{m}_2^r(G)} < \frac{1}{m(G)} - \frac{1}{m_2(G)} = \frac{v}{e} - \frac{v-2}{e-1}. \tag{3.4}$$

Maximizing (3.4) subject to $v \leq e$ we see that the maximum is attained whenever $v = e$. Thus the above is at most

$$1 - \frac{v-2}{v-1} = \frac{1}{v-1}.$$

□

Lemma 3.20. *Let $F^* \in \mathcal{F}^k$ where $k \geq 2$. Then*

$$m_1(F^*_-) < \overline{m}_2^k(F)$$

and for every $V_0 \subsetneq V(F)$

$$m(V_0, (V_0, F) \times (e, F^*_-)) < \overline{m}_2^k(F).$$

Proof. Let $F^* \in \mathcal{F}^k$. We have

$$n^{v_{F^*}-1} n^{-e_{F^*}/\overline{m}_2^k(F)} \overset{Lemma\ 3.15}{\geq} n^{1-1/\overline{m}_2^1(F)} \geq 1$$

and thus $d_1(F^*) \leq \overline{m}_2^k(F)$. F^* is 2-balanced and thus strictly 1-balanced. Therefore we obtain the inequality

$$m_1(F^*_-) < m_1(F^*) = d_1(F^*) \leq \overline{m}_2^k(F),$$

which proves the first part of Lemma 3.20.

For the second part we want to apply Lemma 3.26 with $(R, G) \leftarrow (V_0, F)$, $(e, H) \leftarrow (e, F^*)$ and $t \leftarrow \overline{m}_2^k(F) - \varepsilon$, where $\varepsilon > 0$ is a small constant such that $m_1(F^*_-) \leq \overline{m}_2^k - \varepsilon$. Since F^* is 2-balanced (e, F^*) is also balanced. We have choosen ε such that $m_1(F^*_-) \leq t$. The final premise of Lemma 3.26 is established by

$$v_{F^*} - 2 - \frac{e_{F^*} - 1}{\overline{m}_2^k(F)} \overset{Lemma\ 3.15}{\geq} -\frac{1}{\overline{m}_2^1(F)} + \frac{1}{\overline{m}_2^k(F)} \overset{Lemma\ 3.27}{>} -\frac{1}{m(V_0, G)}.$$

Thus we can apply Lemma 3.26 which proves the last property.

□

3.2 The KŁR statement for rooted graphs

The proof of the theorem follows the proof of the KŁR-conjecture by Saxton, Thomason in [ST15] and relies on their container theorem:

Definition 3.28. *Let G be an r-graph of order n and average degree d. Let $\tau > 0$. Given $v \in V(G)$ and $2 \le j \le r$, let*

$$d^{(j)}(v) = \max\{d(\sigma) : v \in \sigma \subset V(G), |\sigma| = j\}.$$

If $d > 0$ we define δ_j by the equation

$$\delta_j \tau^{j-1} nd = \sum_v d^{(j)}(v).$$

Then the co-degree function $\delta(G, \tau)$ is defined by

$$\delta(G,\tau) = 2^{\binom{r}{2}-1} \sum_{j=2}^{r} \delta_j 2^{-\binom{j-1}{2}}.$$

If $d = 0$ we define $\delta(G, \tau) = 0$.

Theorem 3.29 ([ST15], Corollary 3.6)**.** *Let $\mathcal{E}$ be an r-graph on the vertex set $[n]$. Let $0 < \varepsilon, \tau < 1/2$. Suppose that τ satisfies $\delta(\mathcal{E}, \tau) \le \varepsilon/12r!$. Then there exists a constant $c = c(r)$, and a function $C\colon \mathcal{P}([n])^s \to \mathcal{P}[n]$ where $s \le c\log(1/\varepsilon)$, with the following properties. Let $\mathcal{T} = \{(T_1, \dots, T_S) \in \mathcal{P}([n])^s \colon |T_i| \le c\tau n\}$, and let $\mathcal{C} = \{C(T)\colon T \in \mathcal{T}\}$. Then*

1. *for every $I \subset [n]$ for which $e(\mathcal{E}[I]) \le \varepsilon\tau^r e(\mathcal{E})$ there exists $T = (T_1, \dots, T_s) \in \mathcal{T} \cap P(I)^s$ with $I \subset C(T)$,*
2. *$e(\mathcal{E}[C]) \le \varepsilon e(\mathcal{E})$ for all $C \in \mathcal{C}$.*

For a graph F we denote with $K_{F,n}$ the v_F-partite graph with vertex partitions $V_1, \dots, V_{v_F}$ of size n, such that $K_{F,n}[V_i, V_j]$ is complete if $\{i, j\} \in E(F)$ and empty otherwise.

For a rooted graph (R, F) and $G_R \in \mathcal{R}(R, n)$ we denote with $\mathcal{E}(G_R, F)$ the hypergraph whose vertices are the edges of $K_{F_-,n}$ and whose edges form (when seen as subgraphs of $K_{F_-,n}$) a partite copy of F_- whose roots induce an edge in G_R.

To proof Theorem 3.9 we will apply Theorem 3.29 to $\mathcal{E}(G_R, F)$. The first step is to obtain a bound on the co-degree function.

Lemma 3.30. *Let (R, F) be a rooted graph with $e_{F_-} > 1$. Let $0 < \gamma, q(n) \le 1 \le A$.*

Then for n sufficiently large every hypergraph $G_R \in \mathcal{R}(R, n)$ which is (F, Aq)-upper-extensible satisfies

$$\delta\Big(\mathcal{E}(G_R, F), \gamma^{-1} n^{-1/m_2(R,F,-\log_n(q))}\Big) \le \gamma e_{F_-} 2^{e_{F_-}^2} \frac{n^{|R|}(Aq)^{e_{F[R]}}}{|E(G_R)|}.$$

Proof. Let σ denote a set of vertices of $\mathcal{E} = \mathcal{E}(G_R, F)$. We identify σ with the set of edges from $K_{F_-,n}$ which it represents. If the degree of σ is non zero this set of edges is a graph $F' \subset K_{F_-,n}$ which is isomorphic to some subgraph of F_-. The degree of F' is the number of ways we can extend F' to a partite copy of F_- in $K_{F_-,n}$ whose roots form an edge in G_R.

Since G is (F, Aq)-upper-extensible we have

$$d\big(F'\big) \le n^{v_F - v_{F'}} (Aq)^{e_{F[R]} - e_{F[R \cap V(F')]}}.$$

For $j \ge 2$ and an edge $e \in E(K_{F_-,n})$ the quantity $d^{(j)}(e)$ is the maximum of $d(F')$ over all F' with $e \in F'$ and $|F'| = j$. Thus

$$d^{(j)}(e) \le n^{v_F - v_{F_j}} (Aq)^{e_{F[R]} - e_{F[R \cap V(F_j)]}}$$

where

$$F_j = \operatorname*{arg\,max}_{\substack{F' \subseteq F \\ e(F') - e(F'[R]) = j}} n^{-v_{F'}} (Aq)^{-e_{F'[R]}}.$$

Observe that $F_j[R] = F[R \cap V(F_j)]$. Let $t = -\log_n(q)$ and $\tau = \gamma^{-1} n^{-1/m_2(R,F,t)}$. Using $m_2(R, F, t) \ge d_2\big(R \cap V(F_j), F_j, t\big)$ we obtain

$$\begin{aligned} \frac{1}{\tau^{j-1}} &= \gamma^{j-1} \Big(n^{1/m_2(R,F,t)}\Big)^{(j-1)} \\ &\le \gamma^{j-1} \Big(n^{1/d_2\left(R \cap V(F_j), F_j, t\right)}\Big)^{(j-1)} \\ &= \gamma^{j-1} n^{v_{F_j} - 2} q^{e_{F_j[R]}}. \end{aligned}$$

The number of edges in $\mathcal{E}$ is $|E(G_R)| n^{v_F - |R|}$. Thus for $j \ge 2$ we have

$$\delta_j = \frac{\sum_e d^{(j)}(e)}{\tau^{j-1} e_{F_-} |E(\mathcal{E})|} \le \frac{e_{F_-} n^2 n^{v_F - v_{F_j}} (Aq)^{e_{F[R]} - e_{F_j[R]}}}{\tau^{j-1} e_{F_-} |E(G_R)| n^{v_F - |R|}} \le \gamma^{j-1} \frac{n^{|R|} A^{e_{F[R]} - e_{F_j[R]}} q^{e_{F[R]}}}{|E(G_R)|}.$$

Finally we obtain

$$\delta(\mathcal{E}, \tau) = 2^{\binom{e_{F_-}}{2} - 1} \sum_{j=2}^{e_{F_-}} \delta_j 2^{-\binom{j-1}{2}} \le e_{F_-} 2^{e_{F_-}^2} \gamma \frac{n^{|R|} (Aq)^{e_{F[R]}}}{|E(G_R)|}.$$

as claimed. □

Having bounded the co-degree function we can obtain a collection of containers for $\mathcal{E}(G_R, F)$ which do not induce many edges in $\mathcal{E}(G_R, F)$. Viewing our contains as subgraphs of $K_{F_-,n}$ this means that they contain few copies of F_- whose roots induce an edge in G_R. To prove a KŁR-type statement we want our containers to be sparse subgraphs of $K_{F_-,n}$. The following two lemmas establish that if G_R is lower-regular then the containers obtained by Theorem 3.29 are indeed sparse.

Lemma 3.31. *Let (R, F) denote a rooted graph. For every $\delta > 0$ there exists $\varepsilon > 0$ such that for all $p \ge \delta$ the following holds. Suppose that $G_R \in \mathcal{R}(R, n)$ is (F, q, ε)-lower-regular and that the bipartite graphs of $G \subseteq K_{F_-,n}$ are (ε)-regular with density at least p then*

$$|E(T(G, G_R))| \ge (1 - \delta) p^{e_{F_-}} q^{e_{F[R]}} n^{v_F}.$$

Proof. Observe that the density p is at least δ, which is a constant. Therefore G is a (dense) regular graph and standard counting arguments apply. We only sketch of the proof: using standard arguments we find roughly $n^{v_F-|R|}p^{e_{F[V(F)\setminus R]}}$ tuples in $\bigtimes_{i\in V(F)\setminus R} V_i$ whose common neighborhoods into the partitions G_R are roughly as large as expected (in particular they are of linear size). Thus for ε small enough the (F, q, ε)-lower-regularity of G_R guarantees that every one of these tuples extends to roughly $n^{|R|}p^{e_F-e_{F[V(F)\setminus R]}}q^{e_{F[R]}}$ copies of F_- whose roots form an edge in G_R. Every such copy of F_- contributes on edge to the multi-hypergraph $T(G, G_R)$ and we obtain the desired bound. □

Lemma 3.32. *Let (R, F) be a rooted graph with $e_{F_-} > 1$. Let $\delta > 0$ be small enough and let $A \geq 1$. Then there exists $c, \varepsilon(\delta), R(\delta), \gamma(\delta, A)$ such that the following is true. Suppose that $\tau(n), q(n) \in o(1)$ satisfy $\tau \geq \gamma^{-1}n^{-1/m_2(R,F,-\log_n q)}$. If $G_R \in \mathcal{R}(R, n)$ is (F, Aq)-upper-extensible and (F, q, ε)-lower-regular then for n large enough there exists a collection $\mathcal{C}$ of subgraphs of $K_{F_-,n}$ such that*

1. *for every $G \subseteq K_{F_-,n}$ for which $e(T(G, G_H)) \leq \varepsilon\tau^{e_{F_-}}q^{e_{F[R]}}n^{v_F}$ there exists $T_1, \ldots, T_s \subseteq G$ with $G \subset C(T_1, \ldots, T_s) \in \mathcal{C}$, $e(T_i) \leq c\tau n^2$ and $s \leq c\log(A^{e_{F[R]}}/\varepsilon)$,*
2. *for every $C \in \mathcal{C}$ there exists $\{i, j\} \in E(F_-)$ and equitable partitions $V_i = V_{i,1} \cup \cdots \cup V_{i,r}$ and $V_j = V_{j,1} \cup \cdots \cup V_{j,r}$ where $r \leq R(\delta)$ such that for at least $r^2/2e_{F_-}$ pairs $x, y \in [r]$ we have $e\big(C[V_{i,x}, V_{j,y}]\big) \leq \delta|V_{i,x}||V_{j,y}|$.*

Proof. The constants $c, \mu(\delta), R(\mu), \varepsilon(\delta, \mu, R)$ and $\gamma(\varepsilon, A)$ will be determined later. Let $\varepsilon' = \varepsilon q^{e_{F[R]}}n^{|R|}/e(G_R)$ and $\mathcal{E} = \mathcal{E}(G_R, F)$. Since G_R is (F, Aq)-upper-extensible we can invoke Lemma 3.30 to obtain the bound

$$\delta(\mathcal{E}, \tau) \leq \gamma e_{F_-}2^{e_{F_-}^2}\frac{n^{|R|}(Aq)^{e_{F[R]}}}{e(G_H)} = \gamma e_{F_-}2^{e_{F_-}^2}\frac{\varepsilon' A^{e_{F[R]}}}{\varepsilon}.$$

For $\gamma(\varepsilon, A)$ small enough we obtain

$$\delta(\mathcal{E}, \tau) \leq \frac{\varepsilon'}{12r!},$$

which is what we need to apply Theorem 3.29 with $\mathcal{E} \leftarrow \mathcal{E}$, $\varepsilon \leftarrow \varepsilon'$, $\tau \leftarrow \tau, r \leftarrow e_{F_-}$ to obtain a collection of containers $\mathcal{C}$. We will now show that these containers (when viewed as subgraphs of $K_{F_-,n}$) satisfy the conditions of our Lemma.

So let $G \subseteq K_{F_-,n}$ with

$$e(T(G, G_R)) \leq \varepsilon\tau^{e_{F_-}}q^{e_{F[R]}}n^{v_F} = \varepsilon'\tau^{e_{F_-}}n^{v_F-|R|}e(G_R) = \varepsilon'\tau^{e_{F_-}}e(\mathcal{E}).$$

Define $I = E(G)$ and observe that $e(\mathcal{E}[I]) = e(T(G, G_R))$ and thus $e(\mathcal{E}[I]) \leq \varepsilon'\tau^{e_{F_-}}e(\mathcal{E})$. Therefore we obtain $T_1, \ldots, T_s \subseteq G \subseteq C(T_1, \ldots, T_s) \in \mathcal{C}$ with $e(T_i) \leq c'\tau v(\mathcal{E}) = c\tau n^2$ and $s \leq c\log(1/\varepsilon')$. Since G_R is (F, Aq)-upper-extensible we have

$$\log\left(\frac{1}{\varepsilon'}\right) = \log\left(\frac{e(G_R)}{\varepsilon q^{e_{F[R]}}n^{|R|}}\right) \leq \log\left(\frac{A^{e_{F[R]}}}{\varepsilon}\right),$$

which proves the bound on s.

It remains to show that every $C \in \mathcal{C}$ contains a sparse partition (in the sense of 2.). To this end, for μ small enough depending on δ, consider a μ-regular partition of C which refines the initial partition $(V_i)_{i \in V(F)}$. For every i we obtain a partition $V_i = V_{i,1} \cup \cdots \cup V_{i,r}$ for some $r \leq R(\mu)$. Now consider $C_x = C\Big[V_{1,x_1} \cup \cdots \cup V_{v_F, x_{v_F}}\Big]$ for some $x \in [r]^{v_F}$. For ε small enough depending on R, μ the $|R|$-graph $G_{R,x} = G_R[V(C_x)]]$ is (F, μ, q)-lower-regular. Thus if all pairs in C_x are μ-regular with density at least δ then by Lemma 3.31 for μ small enough depending on δ

$$e(T(C_x, G_{R,x})) \geq (1 - \delta)\delta^{e_{F_-}} q^{e_{F[R]}} \left(\frac{n}{2r}\right)^{v_F}.$$

But $e(T(C_x, G_{R,x}))$ is at most

$$e(\mathcal{E}[C]) \leq \varepsilon' e(\mathcal{E}) = \varepsilon' n^{v_F - |R|} e(G_R) = \varepsilon n^{v_F} q^{e_{F[R]}},$$

which is a contradiction for ε small enough depending on R, δ.

Thus for every x there exists $\{i, j\} \in E(F)$ such that $C\Big[V_{i,x_i}, V_{j,x_j}\Big]$ is either sparse or not μ-regular. By the pigeonhole principle at least an $1/e_F$-fraction of the x nominate the same edge $\{i, j\}$ and every pair $V_{i,a}, V_{j,b}$ is nominated by at most $r^{v_F - 2}$ different x. Finally at most an μ-fraction of these pairs is not μ-regular. Therefore we have found i, j such that at least $r^2/(2e_{F_-})$ of the pairs $V_{i,\cdot}, V_{j,\cdot}$ have density at most δ. □

The proof of Theorem 3.9 now follows from a standard counting argument:

Theorem 3.9. *Let (R, F) be a rooted graph. For every $\beta > 0$, $A \geq 1$ there exist $\alpha(A, \beta), \mu(\beta) > 0$ such that for every $q(n) = o(1)$ the following holds:*

For n large enough suppose that $m \geq \alpha^{-1} n^{2 - 1/m_2(R, F, -\log_n q)}$ and that $G_R \in \mathcal{R}(R, n)$ is (F, Aq)-upper-extensible as well as (F, q, μ)-lower-regular. Then the number of graphs G in $\mathcal{G}(F_-, n, m, \mu)$ for which $T(G, G_R)$ contains fewer than $\alpha(m/n^2)^{e_{F_-}} q^{e_F[R]} n^{v_F}$ edges is at most

$$\beta^m \binom{n^2}{m}^{e_F - e_{F[R]}}.$$

Proof. The proof will require a number of constants which will be fixed during the proof. Their dependencies are as follows: $\delta(\beta)$, $\varepsilon(\delta)$, $R(\delta)$, $\gamma(\delta, A)$, $\hat{s}(A, \varepsilon)$, $\eta(\hat{s}, \gamma)$, $\alpha(\varepsilon, \gamma, \eta)$, $\mu(\varepsilon, R)$.

We invoke Lemma 3.32 with $\delta \leftarrow \delta$, $A \leftarrow A$ and obtain constants c, $\varepsilon(\delta)$, $R(\delta)$ and $\gamma(\delta, A)$.

Fix $\tau = \eta m/n^2$. For α small enough depending on γ and η we have

$$\tau \geq \gamma^{-1} n^{-1/m_2(R, F, -\log_n q)}$$

and for α small enough depending on ε and η we have

$$\varepsilon \tau^{e_{F_-}} q^{e_{F[R]}} n^{v_F} \geq \alpha(m/n^2)^{e_{F_-}} q^{e_{F[R]}} n^{v_F}.$$

Therefore for $\mu \leq \varepsilon$ small enough Lemma 3.32 guarantees the existence of a container $T_1, \ldots, T_s \subseteq G \subseteq C(T_1, \ldots, T_s)$ with $s \leq c \log(A^{e_{F[R]}}/\varepsilon) =: \hat{s}$ whenever $G \in \mathcal{G}(F_-, n, m, \mu)$ does not satisfy $e(T(G, G_R)) > \alpha\big(m/n^2\big)^{e_{F_-}} q^{e_{F[R]}} n^{v_F}$.

To count all such graphs G we fix $T = (T_1, \ldots, T_s)$ and then pick $G \subseteq K_{F_-,n}$ u.a.r. among all graphs with exactly m edges in each bipartite graph. Following [ST15] we define the following events

$$\begin{aligned} E_T &: T_1 \cup \cdots \cup T_s \subseteq G \subseteq C(T) \text{ and } G \in \mathcal{G}(F_-, n, m, \mu), \\ F_T &: T_1 \cup \cdots \cup T_s \subseteq G, \\ G_T &: G \subseteq C(T) \text{ and } G \in \mathcal{G}(F_-, n, m, \mu). \end{aligned}$$

We firstly show that $\sum_T \Pr[F_T] \leq 2^m$. Note that this is the expected number of tuples $T \subseteq G$. The maximum number of tuples $T \subseteq G$ is at most

$$\begin{aligned} \sum_{|T_1|,\ldots,|T_{\hat{s}}|} \prod_{i \leq \hat{s}} \binom{e_{F_-} m}{|T_i|} &\leq \big(c\tau n^2\big)^{\hat{s}} \binom{e_{F_-} m}{c\tau n^2}^{\hat{s}} \\ &\leq \big(c\tau n^2\big)^{\hat{s}} \left(\frac{e e_{F_-} m}{c\tau n^2}\right)^{\hat{s} c\tau n^2} \\ &= (c\eta m)^{\hat{s}} \left(\frac{e e_{F_-}}{c\eta}\right)^{\hat{s} c\eta m} \\ &\leq 2^m, \end{aligned}$$

for η small enough depending on $\hat{s}$.

Secondly we show that $\Pr[G_T \mid F_T] \leq (\beta/2)^m$ and thus

$$\sum_T \Pr[E_T] = \sum_T \Pr[G_T \mid F_T] \Pr[F_T] \leq \beta^m,$$

which implies the Theorem.

For fixed T and $C(T)$ let $\{i, j\} \in E(F_-)$ and $V_i = V_{i,1} \cup \cdots \cup V_{i,r}$ and $V_j = V_{j,1} \cup \cdots \cup V_{j,r}$ be given by property (2) of Lemma 3.32. For $\mu(R)$ small enough we use the (μ)-regularity of $G\big[V_i, V_j\big]$ to require

$$\big|G\big[V_{i,x}, V_{j,x}\big]\big| \geq (1-\mu)\Big(\frac{n}{r}\Big)^2 \frac{m}{n^2} \geq m/2r^2$$

for every $x, y \in [r]$ while $\big|C\big[V_{i,x}, V_{j,x}\big]\big| \leq \delta\big(\frac{n}{r}\big)^2$ for at least $r^2/2e_{F_-}$ choices of x, y. Let $C' = \bigcup C\big[V_{i,x}, V_{j,x}\big]$ where the union runs over $r^2/2e_{F_-}$ sparse pairs. We have $|C'| \leq \delta n^2/2e_{F_-}$ and for G to be (μ)-regular we require $|G \cap C'| \geq m/4e_{F_-}$. We conclude for $\eta(\gamma, c)$ small enough

$$\begin{aligned} \Pr[G_T \mid F_T] &\leq \Pr\big[\big|G \cap C'\big| \geq m/4e_{F_-} \mid F_T\big] \\ &\leq \Pr\big[\big|(G-T) \cap C'\big| \geq m/4e_{F_-} - \hat{s} c\tau n^2\big] \\ &\leq \Pr\big[\big|(G-T) \cap C'\big| \geq m/6e_{F_-}\big] \\ &\leq \binom{\delta \frac{n^2}{2e_{F_-}}}{\frac{m}{6e_{F_-}}} \left(\frac{m}{n^2}\right)^{m/6e_{F_-}} \leq (3e\delta)^{m/6e_{F_-}} \leq \left(\frac{\beta}{2}\right)^m, \end{aligned}$$

for $\eta(\hat{s})$ and $\delta(\beta)$ small enough. □

Chapter 4

Local resilience for squares of almost spanning cycles in random graphs

This chapter is based on a joint work with Angelika Steger [NS16b].

4.1 Introduction

A classical result of Dirac [Dir52] states that any graph G on $n \geq 3$ vertices with minimum degree $\delta(G) \geq n/2$ contains a Hamilton cycle. This result is not difficult and a proof can be found in most text books on graph theory, see e.g. [Bol98, Wes00]. One also easily checks that the constant $1/2$ is best possible: the complete bipartite graph on $(n-1)/2$ and $(n+1)/2$ vertices (assuming n odd) has minimum degree $n/2 - 1/2$ but does not contain a Hamilton cycle.

In 1962, Pósa conjectured that $G(V, E)$ contains a square of a Hamiltonian cycle if $\delta(G) \geq 2n/3$. A square of a cycle C is the cycle C together with all edges between vertices that have distance 2 in C. As before it is not difficult to see that the constant $2/3$ is best possible, just consider the complete tripartite graph on $(n-1)/3$, $(n-1)/3$ and $(n+2)/3$ vertices (assuming 3 divides $n-1$). However proving that minimum degree $2n/3$ actually suffices turned out to be a difficult problem. It required the development of powerful tools, most notably Szemerédi's Regularity Lemma [Sze75, Sze76] and the so-called Blow-Up Lemma [KSS97], before Pósa's conjecture was proven, at least for all sufficiently large n [KSS96].

Theorem 4.1 (Komlós, Sárközy, Szemerédi)**.** *There exists a natural number n_0 such that if G has order n with $n \geq n_0$ and*

$$\delta(G) \geq \frac{2}{3}n,$$

then G contains the square of a Hamiltonian cycle.

In modern terminology the above results can also be stated as resilience statements. For a monotone increasing graph property $\mathcal{P}$ the (local) resilience of a graph $G = (V, E)$ with respect to $\mathcal{P}$ is the minimum $r \in \mathbb{R}$ such that by deleting at each vertex $v \in V$ at most an r-fraction of the edges incident to v one can obtain a graph that does not have property $\mathcal{P}$. Dirac's theorem implies that the local resilience of the property 'containing a Hamilton cycle' of the complete graph is $1/2$, while the proof of Posa's conjecture implies that the property 'containing a square of a Hamilton cycle' of the complete graph is $1/3$.

A natural extension for resilience results is to consider instead of the complete graph the random graph $G_{n,p}$ and ask for the resilience as a function of the edge probability p. It is natural to expect that there exists a threshold p_0 so that for $p \gg p_0$ the local resilience of $G_{n,p}$ of a property $\mathcal{P}$ is w.h.p. equal to the local resilience of the complete graph, while for $p \ll p_0$ the random graph $G_{n,p}$ w.h.p. does not satisfy the property $\mathcal{P}$ at all. Indeed, such a result is known, up to polynomial factors, for the property 'contains a Hamilton cycle'. The threshold for existence of a Hamilton cycle is $p = (\log n + \log\log n + \omega(n))/n$ [KS83, Bol83], while Lee and Sudakov [LS12] showed that for $p \gg \log n/n$ the local resilience is $1/2 - o(1)$.

Our aim is to study the local resilience for the property 'containing a square of a Hamilton cycle'. As it turns out for this problem already the threshold for existence is a hard problem that, till today, is not yet completely understood. As a square of a Hamilton cycle contains many triangles, $p \geq c/\sqrt{n}$ is certainly necessary, and it is conceivable that this may be also be the true answer. The best bound known is $p \geq C\log^4 n/\sqrt{n}$ [NŠ16a].

For the corresponding resilience problem one is thus tempted to speculate that at least for $p \geq \mathrm{poly}(\log n)/\sqrt{n}$, for an appropriate polylog-factor, we have that the resilience of $G_{n,p}$ with respect to the property 'containing a square of a Hamilton cycle' is $1/3 - o(1)$. However, for this property it is easy to see that this is far too optimistic. By deleting all edges in the neighborhood of a vertex v we can ensure that v cannot be part of any square of a cycle. Thus for any $p = o(1)$ we have that the resilience for 'containing a square of a Hamilton cycle' is $o(1)$. In order to obtain a non-trivial result we thus need to weaken the required property. One easily checks that for constant resilience the best one can hope for is to find a square of a cycle that covers all but $\Theta(1/p^2)$ vertices. Here we show an approximate version of such a best possible result.

Theorem 4.2. *For every $\gamma, \nu > 0$ and $p = n^{-\frac{1}{2}+\gamma/2}$ a.a.s. every subgraph of $G_{n,p}$ with minimum degree at least $(2/3+\nu)np$ contains a square of a cycle on at least $(1-\nu)n$ vertices.*

Our result should be compared to a recent result of Peter Allen, Julia Böttcher, Julia Ehrenmüller and Anusch Taraz [ABET]. The authors prove a sparse version of the bandwidth theorem, which in particular implies that $G_{n,p}$ has local resilience for the property 'contains a square of a cycle on $n - C/p^2$ vertices' as long as $p \gg (\log n/n)^{1/4}$. Their proof technique as well as previous universality results hit a natural barrier around $p = n^{-1/\Delta}$ where Δ denotes the maximum degree of the embedded graph. Note that this density is required for any greedy / sequential type of embedding scheme, as for $p \ll n^{-1/\Delta}$ the typical neighbourhood of any Δ vertices is empty and one thus need to design more sophisticated look-ahead schemes. We achieve this by designing some pruning process that identifies edges that satisfy some good expansion properties. In this way we obtain the first nontrivial resilience result for almost spanning subgraphs that achieves, up to polylog factors, the optimal density.

A natural question is whether a similar result holds for higher powers of a cycle. Applying the aforementioned bandwidth theorem gives a bound of $p \gg (\log n/n)^{1/2k}$ for the k-th power of an almost spanning cycle. Similarly to the case $k = 2$ this does not match the obvious lower bound of $p \geq n^{-1/k}$. Our approach does not generalize easily to this setting. This is mostly due to our reliance on sparse regularity techniques which yield very strong statements about the distribution of the edges, but not on larger structures like triangles or larger complete graphs.

4.2 Proof

The proof makes heavy use of the sparse regularity lemma (see [Koh97]) and related techniques. The lemma was introduced in Chapter 2 and we have seen that a.a.s. $G_{n,p}$ is upper-uniform and that therefore the spare regularity lemma can be applied to every subgraph of $G_{n,p}$. For our proof we need the slightly stronger statement that a.a.s. every subgraph of $G_{n,p}$ that satisfies some minimum degree condition contains a particularly nice regular partition. The proof follows routinely by standard arguments. We include it for convenience of the reader.

Lemma 4.3. *For every small enough $\mu, \nu, \varepsilon > 0$ and every positive integer $r_{\min}$ there exists $\alpha(\nu)$ and an integer $r_{\max}(\nu, \varepsilon, r_{\min})$ such that for $p \gg 1/n$ a.a.s. every spanning subgraph $\tilde{G} \subseteq G_{n,p}$ with minimum degree at least $(\mu+\nu)np$ contains a partition of the vertices $V = V_0 \cup V_1 \cup \cdots \cup V_r$, where $r \in [r_{\min}, r_{\max}]$, such that $|V_0| \leq \varepsilon n$, $|V_1| = \cdots = |V_r|$ and such that for every i there exist at least μr indices $j \in [r] \setminus \{i\}$ such that $\tilde{G}[V_i, V_j]$ contains a spanning (ε)-regular subgraph with $\lfloor |V_i||V_j|\alpha p \rfloor$ edges.*

Proof. We choose α, $k_{\min}$ and ε_0 such that inequality $(*)$ from below and the following inequalities are satisfied simultaneously:

$$\sqrt{\varepsilon_0} \le \frac{\nu}{2}, \qquad 2\varepsilon_0/\alpha < \varepsilon, \qquad (1-\sqrt{\varepsilon_0})k_{\min} \ge r_{\min}, \qquad \varepsilon_0 + 2\sqrt{\varepsilon_0} \le \varepsilon.$$

Suppose that $\tilde{G}$ is a spanning subgraph of $G_{n,p}$ with $\delta(\tilde{G}) \ge (\mu+\nu)np$. For every $\eta > 0$ a.a.s. every subgraph of the random graph $G_{n,p}$ is (η, p)-upper-uniform and thus the sparse regularity lemma can be applied to every subgraph of $G_{n,p}$. The sparse regularity lemma (Theorem 2.4) gives us a constant $k_{\max}(\varepsilon_0, k_{\min})$ such that we find an (ε_0, p)-regular partition $(V_i)_{i=0}^k$ with exceptional class V_0 of $\tilde{G}$ for some $k \in [k_{\min}, k_{\max}]$.

Denote with $\tilde{n} \in [(1-\varepsilon_0)n/k, n/k]$ the size of the partition classes. For $i \in [k]$ define

$$d_i := \{j \in [k] \setminus \{i\} \mid V_i,\, V_j \text{ has density at least } \alpha p \text{ in } \tilde{G}\}.$$

Note that a.a.s. in $G_{n,p}$ we have that $\left|E(V_i, V_j)\right| \le (1+\varepsilon_0)\tilde{n}^2 p$ for all $i, j \in [k]$ and, with room to spare, $|E(V_i, V_0 \cup V_i)| \le 2(\varepsilon_0 n + \tilde{n})\tilde{n}p$ for all $i \in [k]$. The minimum degree condition of $\tilde{G}$ thus implies that for all $i \in [k]$

$$\tilde{n}(\mu+\nu)np \le E(V_i, V_0 \cup V_i) + \sum_{j \in [k]\setminus i} E(V_i, V_j) \le 2(\varepsilon_0 n + \tilde{n})\tilde{n}p + k\alpha p\tilde{n}^2 + d_i(1+\varepsilon_0)\tilde{n}^2 p$$

and thus

$$d_i \ge \frac{(\mu+\nu)n - 2(\varepsilon_0 n + \tilde{n}) - k\alpha\tilde{n}}{(1+\varepsilon_0)\tilde{n}} \ge \frac{(\mu+\nu)k - 2(\frac{\varepsilon_0}{1-\varepsilon_0}k + 1) - k\alpha}{(1+\varepsilon_0)} \overset{(*)}{\ge} \Big(\mu + \frac{\nu}{2}\Big)k.$$

Observe that the fact that at most $\varepsilon_0\binom{k}{2}$ pairs (V_i, V_j) are not (ε_0, p)-regular implies that there are at most $\sqrt{\varepsilon_0}k$ indices in $[k]$ for which the set

$$\{j \in [k] \setminus \{i\} \mid (V_i, V_j) \text{ is not } (\varepsilon_0, p)\text{-regular graph}\}$$

has size at least $\sqrt{\varepsilon_0}k$. Every (ε_0, p)-regular graph with density at least αp is (ε_0/α)-regular and thus contains a $(2\varepsilon_0/\alpha)$-regular subgraph with $\lfloor \tilde{n}^2\alpha p \rfloor$ edges (see Lemma 2.7). As $\sqrt{\varepsilon_0} \le \frac{\nu}{2}$, we can thus find a subset $R \subseteq [k]$ of at least $(1-\sqrt{\varepsilon_0})k$ indices such that for every $i \in R$ the set

$$\{j \in R \setminus \{i\} \mid V_i,\, V_j \text{ contain an } (2\varepsilon_0/\alpha)\text{-regular graph with } \lfloor \tilde{n}^2\alpha p \rfloor \text{ edges}\}$$

is of size at least μk. Wlog. we may assume that $R = \{1, \dots, r\}$, where $r \ge (1-\sqrt{\varepsilon_0})k \ge r_{\min}$. By choice of ε_0 we know that the cardinality of $V_0 \cup \bigcup_{i>r} V_i$ is at most $\varepsilon_0 n + \sqrt{\varepsilon_0}k\tilde{n} \le \varepsilon n$. Thus $V_0 \cup \bigcup_{i>r} V_i, V_1, \dots, V_r$ is the desired partition and the corollary thus holds for $r_{\max} = k_{\max}$. □

With Lemma 4.3 at hand, an alert reader will certainly be able to guess our proof strategy: apply the Komlós, Sárközy, Szemerédi Theorem to the partition guaranteed by Lemma 4.3 in order to find a square of a cycle for this partition and then find a long square of a cycle within this structure. The next definitions provide the notations to make this idea precise. First we define the notion of a regular blow-up of a graph.

Definition 4.4. *Denote with $\mathcal{G}(F, n, p, \varepsilon)$ the class of graphs that consist of $|V(F)|$ pairwise disjoint vertex sets of size n. Each vertex set represents a vertex of F, and two vertex sets span an (ε)-regular graph with density $(1 \pm \varepsilon)p$ whenever the corresponding vertices are adjacent in F.*

With P_k we denote a path of length k, with vertex set $\{1, \ldots, k+1\}$ and edges $\{i, i+1\}$ for $1 \leq i \leq k$. The cycle C_k is obtained from P_k by identifying the vertices 1 and $k+1$. The square of a path is denoted with P_k^2 and the square of a cycle with C_k^2. For a graph $F \in \{P_k^2, C_k^2\}$ and its blow-up $G \in \mathcal{G}(F, n, p, \varepsilon)$ we denote the vertex sets of G by $V_1, \ldots, V_{|F|}$ and assume that the vertex set V_i represents the i-th vertex along the path (or cycle). We collect some additional properties of $\mathcal{G}(P_k^2, n, p, \varepsilon)$ in the following definition:

Definition 4.5. *Denote with $\tilde{G}(P_k^2, n, p, \varepsilon) \subseteq \mathcal{G}(P_k^2, n, p, \varepsilon)$ the class of graphs in which for $i \in \{1, \ldots, k-1\}$*

(i) every edge spanned by V_i, V_{i+1} closes a triangle with at least $(1-\varepsilon)np^2$ vertices in V_{i+2}, and

(ii) all but εn vertices $v \in V_{i+1}$ have neighboorhoods into V_i and V_{i+2} which induce an (ε, p)-lower-regular subgraph.

We also need two auxiliary lemmas related to the above definition. Their proofs are deferred to Section 4.2.1 and Section 4.2.2. The first lemma states that in the random graph the restrictions imposed by Definition 4.5 are easy to satisfy:

Lemma 4.6. *For every $\alpha, \varepsilon, \gamma, k_{\max} > 0$ there exists $\varepsilon' > 0$ such that for every $\eta > 0$ a.a.s. in $G_{n,p}$ with $p = n^{-1/2+\gamma/2}$ every subgraph $G \subseteq G_{n,p}$, where $G \in \mathcal{G}(P_k^2, n_0, \alpha p, \varepsilon')$, $n_0 \geq \eta n$, $k \leq k_{\max}$ contains a spanning subgraph from $\tilde{\mathcal{G}}(P_k^2, n_0, \alpha p, \varepsilon)$.*

For $G \in \tilde{\mathcal{G}}(P_k^2, n, p, \varepsilon)$ with vertex partitions $V_1, \ldots, V_k$ we say that an edge $e \in E(V_1, V_2)$ *expands* to a set of edges $E' \subseteq E(V_i, V_{i+1})$ if there exists a square of a path of length i in G between e and every edge of E'. The next lemma asserts that in the random graph every edge in $E(V_1, V_2)$ expands to a majority of the edges in $E(V_k, V_{k+1})$, whenever k is sufficiently large. Property (1) of Definition 4.5 already guarantees that every edge spanned by V_1, V_2 is contained in many squares of a path ending in $E(V_k, V_{k+1})$. But since these paths may overlap this alone does not imply that every edge expands to a large portion of $E(V_k, V_{k+1})$.

Lemma 4.7. *Let $\eta, \alpha, \gamma > 0$ and $k \geq \frac{3}{\gamma}$ be fixed and let $p = n^{-1/2+\gamma/2}$. For ε small enough depending on α a.a.s. every subgraph $G \subseteq G_{n,p}$ which is from $G \in \tilde{\mathcal{G}}(P_{k+4}^2, n_0, \alpha p, \varepsilon)$ for some*

$n_0 \geq \eta n$ satisfies the following: all edges in $G[V_1 \cup V_2]$ expand to at least a 0.52-fraction of the edges in $G[V_{k+4} \cup V_{k+5}]$.

With these two lemmas at hand we can prove our main result.

Theorem 4.2. *For every $\gamma, \nu > 0$ and $p = n^{-\frac{1}{2}+\gamma/2}$ a.a.s. every subgraph of $G_{n,p}$ with minimum degree at least $(2/3+\nu)np$ contains a square of a cycle on at least $(1-\nu)n$ vertices.*

Proof. Suppose that $\tilde{G}$ is a spanning subgraph of $G_{n,p}$ with $\delta(\tilde{G}) \geq (\frac{2}{3}+\nu)np$. We will the fix constants $\alpha(\nu), \varepsilon(\nu,\alpha), \varepsilon'(\nu,\alpha,\varepsilon,\gamma) > 0$ throughout the proof.

Set $k_0 = \lceil \frac{3}{\gamma} \rceil + 4$ and let $r_{\min} = \max\{3k_0, n_0\}$ where n_0 denotes the constant from Theorem 4.1. For $\alpha(\nu)$ small enough we may invoke Lemma 4.3 with $\mu \leftarrow 2/3$, $\nu \leftarrow \nu$, $\varepsilon \leftarrow \varepsilon'$ to obtain, for some $r \in [r_{\min}, r_{\max}(\nu,\varepsilon')]$, a partition $V = V_0 \cup \cdots \cup V_r$ of the vertices of $\tilde{G}$ such that $|V_1| = \cdots = |V_r| = \tilde{n} \in [(1-\varepsilon')n/r, n/r]$ and for every $i \in [r]$ the number of indices $j \in [r] \setminus \{i\}$ such that $\tilde{G}[V_i, V_j]$ contains a spanning (ε')-regular subgraph with $\lfloor \tilde{n}^2 \alpha p \rfloor$ edges is at least $\frac{2}{3}r$.

Since $r \geq n_0$ Theorem 4.1 tells us that then $\tilde{G}$ must contain a subgraph $G \in \mathcal{G}(C_r^2, \tilde{n}, \alpha p, \varepsilon')$ on the partitions $V_1, \ldots, V_r$. We shall assume wlog. that V_i represents the i-th vertex of the cycle. Furthermore for ease of notation we identify V_{r+i} with V_i.

For integers $i \in [r]$ and $t \in [k_0, 2k_0]$ consider a collection of subsets $V_i' \subseteq V_i, \ldots, V_{i+t}' \subseteq V_{i+t}$ each of size $n' \geq \varepsilon\tilde{n}$. By definition of (ε')-regularity the sets $V_i', \ldots, V_{i+t}'$ induce a subgraph $G' \in \mathcal{G}(P_t^2, n', \alpha p, 2\varepsilon'/\varepsilon)$ in G. By Lemma 4.6 we may pick ε' small enough depending on $\alpha, \varepsilon, 2k_0$ such that a.a.s. G' contains a spanning subgraph $G_0 \subseteq G'$ with $G_0 \in \tilde{\mathcal{G}}(P_t^2, n', \alpha p, \varepsilon)$.

We call an edge $e \in E(V_i', V_{i+1}')$ *good* (w.r.t. $V_i', \ldots, V_{i+t}'$) if it expands to at least a 0.51-fraction of the edges in $E(V_{i+t-1}', V_{i+t}')$ (through $V_i', \ldots, V_{i+t}'$ in G_0). Lemma 4.7 with $\eta \leftarrow \varepsilon/2r_{\max}$, $\alpha \leftarrow \alpha$, $\gamma \leftarrow \gamma$, $k \leftarrow t-4 \geq k_0 - 4 \geq \frac{\gamma}{3}$ tells us that all edges in $G_0[V_i', V_{i+1}']$ expand to at least a 0.52-fraction of the edges in $G_0[V_{i+t-1}', V_{i+t}']$. Since $G_0 \subseteq G$ and since the density of G_0 and G differs by at most $2\varepsilon\alpha p$ this implies that say a 0.99-fraction of the edges in $E(V_i', V_{i+1}')$ are good.

We can now find a long square of a path as follows: Fix an edge $e \in E(V_1, V_2)$ which is good with respect to $V_1, \ldots, V_{k_0+1}$. e expands to a 0.51 fraction of the edges spanned by V_{k_0}, V_{k_0+1} and a 0.99 fraction of the edges in $E(V_{k_0}, V_{k_0+1})$ are good with respect to $V_{k_0}, \ldots, V_{2k_0}$. Thus we may fix a square of a path $P \subseteq G[V_1 \cup \cdots \cup V_{k_0+1}]$ of length k_0 from e to some edge edge $e' \in E(V_{k_0}, V_{k_0+1})$ which is good with respect to $V_{k_0}, \ldots, V_{2k_0}$.

Now remove $V(P) \setminus e'$ from G. Observe that since $r \geq 3k_0$ we did not remove any vertices from $V_{k_0}, \ldots, V_{2k_0}$ and therefore e' is still good (w.r.t. $V_{k_0}, \ldots, V_{2k_0}$). Thus we may extend P to end in an edge $e'' \in E(V_{2k_0-1}, V_{2k_0})$ which is good w.r.t. to $V_{2k_0-1}, \ldots, V_{3k_0-1}$. This procedure can be continued for as long as $|V_i| \geq \varepsilon\tilde{n}$ for every

$i \in [r]$. Thus we obtain a square of a path P which uses at least $(1-2\varepsilon)\tilde{n}$ vertices from each partition of G.

This construction can be generalized to obtain a square of the cycle: before fixing the first edge e set aside sets $\tilde{V}_1 \subseteq V_1, \ldots, \tilde{V}_r \subseteq V_r$ each of size $\varepsilon\tilde{n}$. Pick $e \in E(\tilde{V}_1, \tilde{V}_2)$ such that it expands (backwards) to at least a 0.51-fraction of the edges in $E(\tilde{V}_{r+1-k_0}, \tilde{V}_{r+2-k_0})$. Then embed a long path starting with e in $V(G) \setminus \bigcup \tilde{V}_i$. At any point we may decide to close it by picking the next edge such that it expands to a 0.51-fraction of the edges in $E(\tilde{V}_{r+1-k_0}, \tilde{V}_{r+2-k_0})$. With this construction we find a square of a cycle of length at least

$$r \cdot (1-3\varepsilon)\tilde{n} \geq (1-3\varepsilon)(1-\varepsilon')\frac{n}{r} \geq (1-\mu)n$$

provided that $\varepsilon', \varepsilon$ are chosen small enough depending on μ. □

4.2.1 Proof of Lemma 4.6

For the proof of Lemma 4.6 we need the following sparse regularity inheritance theorem:

Theorem 4.8 (Gerke, Kohayakawa, Rödl, and Steger [GKRS07]). *For $0 < \beta, \varepsilon' < 1$, there exists $\varepsilon_0 = \varepsilon_0(\beta, \varepsilon') > 0$ and $C = C(\varepsilon')$ such that, for any $0 < \varepsilon \leq \varepsilon_0$ and $0 < p < 1$ every (ε, p)-lower-regular graph $G = (V_1 \cup V_2, E)$ satisfies that, for every $q_1, q_2 \geq Cp^{-1}$, the number of pairs of sets (Q_1, Q_2) with $Q_i \subseteq V_i$ and $|Q_i| = q_i$ $(i = 1, 2)$ that induce an (ε', p)-lower-regular graph is at least*

$$\left(1 - \beta^{\min\{q_1, q_2\}}\right)\binom{|V_1|}{q_1}\binom{|V_2|}{q_2}.$$

From Theorem 4.8 we easily deduce a bound on the number of graphs in $\mathcal{G}(K_3, n, p, \varepsilon_0)$ for which there exist 'many' vertices in V_1 whose neighborhood does not induce a lower regular graph of 'roughly' the expected size.

Lemma 4.9. *For every $\beta, \varepsilon > 0$ there exists $\varepsilon_0 > 0$ and $C > 0$ such that for all integers $n, m \in \mathbb{N}$ that satisfy $3^n n^{2n} \leq 2^{n^{3/2}}$ and $m \geq Cn^{3/2}$ the following holds. The number of graphs $G = (V_1 \cup V_2 \cup V_3, E)$ in $\mathcal{G}(K_3, n, m/n^2, \varepsilon_0)$ with m edges between each two partitions for which more than εn vertices in $v \in V_1$ have neighborhoods in V_2, V_3 which are not of size $(1 \pm \varepsilon)m/n$ or which do not induce an $(\varepsilon, m/n^2)$-lower-regular subgraph in $G[V_2, V_3]$ is at most*

$$\beta^m \binom{n^2}{m}^3.$$

Proof. Write $p = m/n^2$ and define β_0 by $\beta_0^{(1-\varepsilon)\varepsilon/2} = \beta/2$. Let $\varepsilon_0 = \min\{\varepsilon/4, \varepsilon_0(\beta_0, \varepsilon)\}$ and $C = \max\{1, C(\varepsilon)/(1-\varepsilon)\}$, where $\varepsilon_0(\cdot,\cdot)$ and $C(\cdot)$ are the functions given by Theorem 4.8.

Consider a graph $G = (V_1 \cup V_2 \cup V_3, E)$ for which the statement fails. For any such graph we may partition V_1 into three sets $V_D, V_B, V_G \subseteq V_1$ as follows: the set V_D contains all vertices whose degree into at least one of V_2 or V_3 is not within $(1 \pm \varepsilon)np$. V_B contains all vertices whose degree into both V_2 and V_3 is within $(1 \pm \varepsilon)np$ but whose neighborhoods in V_2 and V_3 do not induce an (ε, p)-lower-regular graph in $G[V_2, V_3]$. Finally set $V_G = V_1 \setminus (V_D \cup V_B)$. Lemma 2.5 implies that $|V_D| \leq 2\varepsilon_0 n \leq \varepsilon n/2$. Therefore it suffices to enumerate graphs G with $|V_B| \geq \varepsilon n/2$.

We now construct all graphs G which produce a partition with $|V_B| \geq \varepsilon n/2$. First we pick the (ε_0)-regular graph spanned by V_2, V_3. There are at most $\binom{n^2}{m}$ choices.

Second we pick a partition $V = V_D \cup V_B \cup V_G$ and fix the degrees of all vertices $v \in V_1$ in $G[V_1, V_2]$ and $G[V_1, V_3]$. The number of choices is at most $3^n \cdot n^{2n}$. Finally we fix the actual neighborhoods of the vertices of V_1. For a vertex $v \in V_D \cup V_G$ the number of choices is at most

$$\binom{n}{\deg_{G[V_1,V_2]}(v)}\binom{n}{\deg_{G[V_1,V_3]}(v)}.$$

For $v \in V_B$ we have to select a neighborhood which does not induce an (ε, p)-lower-regular subgraph in $G[V_2, V_3]$. Since $\deg_{G[V_1,V_i]}(v) \geq (1-\varepsilon)np \geq C(\varepsilon)p^{-1}$ Theorem 4.8 tells us that the number of such neighborhoods is at most

$$\beta_0^{(1-\varepsilon)np}\binom{n}{\deg_{G[V_1,V_2]}(v)}\binom{n}{\deg_{G[V_1,V_3]}(v)}.$$

Since $|V_B| \geq \varepsilon n/2$ the total number of choices for the neighborhoods of the vertices in V_1 is bounded by

$$\beta_0^{(1-\varepsilon)np|V_B|}\prod_{v\in V_1}\binom{n}{\deg_{G[V_1,V_2]}(v)}\binom{n}{\deg_{G[V_1,V_3]}(v)} \leq \beta_0^{(1-\varepsilon)\varepsilon m/2}\binom{n^2}{m}^2 = \left(\frac{\beta}{2}\right)^m\binom{n^2}{m}^2,$$

where the inequality follows from Vandermonde's identity. Since $3^n \cdot n^{2n} \leq 2^{n^{3/2}} \leq 2^m$ by assumption on n and m, this completes the proof. □

Lemma 4.9 immediately implies the following corollary about the number of triangles spanned by almost all edges:

Corollary 4.10. *For every $\beta, \delta > 0$ there exists $\varepsilon_0 > 0$ and $C > 0$ such that for $m \geq Cn^{3/2}$ and n sufficiently large the following holds. The number of graphs $G = (V_1 \cup V_2 \cup V_3, E)$ in $\mathcal{G}(K_3, n, m/n^2, \varepsilon_0)$, with m edges between each two partitions, for which more than δm edges of $G[V_1, V_2]$ are contained in fewer than $(1-\delta)n(m/n^2)^2$ triangles is at most*

$$\beta^m\binom{n^2}{m}^3.$$

Proof. Let $p = m/n^2$ and choose ε small enough for $(1-\varepsilon)^3 \geq (1-\delta)$ to hold. Denote with $V_1' \subseteq V_1$ the set of vertices $v \in V_1$ whose neighborhoods in V_2, V_3 are of size $(1 \pm \varepsilon)np$ and induce an (ε, p)-lower-regular subgraph in $G[V_1, V_2]$. From Lemma 2.5 we deduce that every vertex in V_1' is incident to at least $(1-\varepsilon)^2 np$ edges (with endpoint in V_2) which are each contained in at least $(1-\varepsilon)^2 np^2 \geq (1-\delta)np^2$ triangles. Therefore, the total number of edges which are contained in fewer than $(1-\delta)np^2$ triangles is at most $m - |V_1'|(1-\varepsilon)^2 np$. By choice of δ, the only possibility that the desired condition is not fulfilled is thus that $|V_1'| \leq (1-\varepsilon)n$. Lemma 4.9 handles exactly this case – and thus concludes the proof if we choose C and ε_0 as in this lemma. □

With Corollary 4.10 at hand we are now ready to prove Lemma 4.6.

Proof of Lemma 4.6. Observe first that it suffices to consider a fixed integer $k \leq k_{\max}$, as the lemma then follows by choosing the minimum ε' for all $k \leq k_{\max}$ and a trivial union bound argument.

We first consider property (i) of Definition 4.5. The key fact here is that we want the property to hold for *every* edge, while Corollary 4.10 guarantees this only for *'almost all'* edges. It is thus obvious what we need to do: show that we can remove edges appropriately. To this end Lemma 2.6 will come in very handy, as it shows that if we choose ε' small enough then we can take away edges repeatedly, while still keeping some regularity properties.

Define constants as follows: $\varepsilon_0(\alpha, \varepsilon)$ will be fixed at the end of the proof, but will be small enough to satisfy $(1-\varepsilon) \leq (1-\varepsilon_0)^3$. Set $\beta = (\alpha/(4e))^3$ and for $i \in [k-1]$ define $\delta_i = (\varepsilon_{i-1}/4)^4/2$ and $\varepsilon_i = \min\{\delta_i/4, \varepsilon_{cor}(\beta, \delta_i)\}$, where $\varepsilon_{cor}(\cdot,\cdot)$ denotes the function $\varepsilon(\beta, \delta)$ defined in Corollary 4.10. Finally define $m_i = \lceil (1-\varepsilon_i) n_0^2 p_0 \rceil$ and $\varepsilon' = \varepsilon_{k-1}/4$. Observe that $\varepsilon_0 > \delta_1 > \varepsilon_1 > \cdots > \delta_{k-1} > \varepsilon_{k-1} > \varepsilon'$.

We now proceed as follows: for $i = k-1$ down to $i = 1$ we remove edges from $E(V_i, V_{i+1})$ if they are not contained in enough triangles with V_{i+2} *with respect* to the edge set that survived the removal process in the previous round. That is, we remove all edges in $E(V_i, V_{i+1})$ which are contained in fewer than $(1-\varepsilon)n_0 p_0^2$ triangles with V_{i+2}, only taking in account edges that are still present.

Assume first that for all $1 \leq i < k$ we remove at most $2\delta_i m_i$ edges. Then the resulting subgraph $\tilde{G}$ is, by construction, such that the graph satisfies property (i) of Definition 4.5. We claim that we also have that all pairs are (ε_0)-regular with density at least $(1-\varepsilon_0)p_0$. Note that this implies that $\tilde{G} \in G(P_k^2, n_0, p_0, \varepsilon_0)$. By definition of $\mathcal{G}(P_k^2, n_0, p_0, \varepsilon')$ we know that (before removing any edges) the pair (V_i, V_{i+1}) is (ε')-regular (and thus $(\varepsilon_0/2)$-regular) with density at least $(1-\varepsilon')p_0$. If we remove at most $2\delta_i m_i \leq (\varepsilon_0/2)^4 \left| E(V_i, V_{i+1}) \right|$ edges from $E(V_i, V_{i+1})$, then Lemma 2.6 implies that the remaining graph is (ε_0)-regular with density at least $(1-\varepsilon'-2\delta_i)p_0 \geq (1-3\delta_i)p_0 \geq (1-\varepsilon_0)p_0$.

So assume the above condition does not hold. Let i denote the largest $1 \leq i < k$ such that when processing $E(V_i, V_{i+1})$ we have to remove more than $2\delta_i m_i$ edges.

We claim that then $G[V_i \cup V_{i+1} \cup V_{i+2}]$ contains one of the subgraphs enumerated by Corollary 4.10.

Indeed, let $G'_{i+1,i+2} \subseteq G[V_{i+1}, V_{i+2}]$ denote the subgraph obtained after removing the edges which do not satisfy property (i). If $i = k-1$ then, since we do not touch the last partition, we have $G'_{i+1,i+2} = G[V_k, V_{k+1}]$ which is trivially $(\varepsilon_i/2)$-regular with density at least $(1-\varepsilon_i)p_0$. If $i < k-1$ then by maximality of i the graph $G'_{i+1,i+2}$ is obtained by removing at most $2\delta_{i+1}m_{i+1} \leq (\varepsilon_i/4)^4|E(G[V_{i+1}, V_{i+2}])|$ edges from the $(\varepsilon_i/4)$-regular graph $G[V_{i+1}, V_{i+2}]$. Therefore by Lemma 2.6 $G'_{i+1,i+2}$ is $(\varepsilon_i/2)$-regular with density at least $(1-\varepsilon' - 2\delta_{i+1})p_0 \geq (1-\varepsilon_i)p_0$. Let $G_{i,i+1} \subseteq G[V_i, V_{i+1}]$, $G_{i,i+2} \subseteq G[V_i, V_{i+2}]$, $G_{i+1,i+2} \subseteq G'_{i+1,i+2}$ denote spanning (ε_i)-regular subgraphs with exactly m_i edges each (for $m_i \gg n_0$ and $\varepsilon' \leq \varepsilon_i/2$ such subgraphs always exists, see Lemma 2.7). Observe that to obtain $G_{i,i+1}$ we removed at most $2\varepsilon_i n_0^2 p_0 \leq \delta_i m_i$ edges from $G[V_i, V_{i+1}]$.

Thus by choice of i there have to be at least $\delta_i m_i$ more edges in $G_{i,i+1}$ which are each contained in fewer than $(1-\varepsilon)n_0p_0^2$ triangles with V_{i+2} in $G_i := G_{i,i+1} \cup G_{i,i+2} \cup G_{i+1,i+2}$. Since $(1-\varepsilon)n_0p_0^2 \leq (1-\varepsilon_0)^3 n_0 p_0^2 \leq (1-\delta_i)n_0(m_i/n_0^2)^2$ and $\varepsilon_i \leq \varepsilon_{cor}\big((\alpha/(4e))^3, \delta_i\big)$ we may apply Corollary 4.10 (with $\beta \leftarrow \beta = (\alpha/(4e))^3$, $\delta \leftarrow \delta_i$, $m \leftarrow m_i$) to conclude that G_i must be one of at most

$$\beta^{m_i}\binom{n_0^2}{m_i}^3,$$

graphs enumerated by Corollary 4.10. The probability that $G_{n,p}$ contains any of these graphs as a subgraph is at most

$$\binom{n}{n_0}^3 \cdot \beta^{m_i}\binom{n_0^2}{m_i}^3 \cdot p^{3m_i} \leq 2^{3n}\beta^{m_i}\left(\frac{en_0^2p}{m_i}\right)^{3m_i} \leq 2^{3n}\beta^{m_i}\left(\frac{2e}{a}\right)^{3m_i} = 2^{3n-3m_i}.$$

Since $n_0 \geq \eta n$ we have $m_i \gg n$ and may additionally union bound over all choices for n_0 and conclude that a.a.s. no such graph appears in $G_{n,p}$. In particular our procedure never removes more than $2\delta_i m_i$ edges and produces a graph $\tilde{G} \in \mathcal{G}\big(P_k^2, n_0, p_0, \varepsilon_0\big)$ which satisfies property (i).

It remains to show that additionally $\tilde{G}$ also satisfies property (ii). Fix the constant $\varepsilon_0 = \min\{\varepsilon/3, \varepsilon_{lem}(\beta, \varepsilon/2)\}$, where $\varepsilon_{lem}(\cdot,\cdot)$ denotes the function $\varepsilon(\beta,\varepsilon)$ defined in Lemma 4.9. This choice is valid since $(1-\varepsilon) \leq (1-(\varepsilon/3))^3$. Suppose that $\tilde{G}$ fails property (ii) for some $i \in [k-1]$. We claim that then $\tilde{G}[V_i \cup V_{i+1} \cup V_{i+2}]$ must contain a subgraph enumerated by Lemma 4.9. To this end let $\tilde{G}' \subseteq \tilde{G}[V_i \cup V_{i+1} \cup V_{i+2}]$ denote a spanning subgraph in which every partition contains exactly $m_0 = \lceil (1-\varepsilon_0)n_0^2p_0 \rceil$ edges and is $(2\varepsilon_0)$-regular (as before see Lemma 2.7).

For $\varepsilon_0 \leq \varepsilon/2$ we have $(1-\varepsilon/2)m_0/n_0^2 \geq (1-\varepsilon)p$ and thus every $\big(\varepsilon/2, m_0/n_0^2\big)$-lower-regular graph is also (ε, p_0)-lower-regular. It follows that $\tilde{G}'$ must be among the at most

$$\beta^{m_0}\binom{n_0^2}{m_0}^3$$

graphs enumerated by Lemma 4.9 (with $\varepsilon \leftarrow \varepsilon/2$, $\beta \leftarrow \beta$ and $m \leftarrow m_0$). As before a union bound shows that a.a.s. $G_{n,p}$ does not contain such a graph and thus $\tilde{G}$ also satisfies property (ii). □

4.2.2 Proof of Lemma 4.7

Definition 4.5 already implies that every edge spanned by V_1, V_2 is contained in many copies of P^2_{k+4}. The goal of this section is to show that these paths also reach a majority of the edges spanned by V_{k+4}, V_{k+5}. As a first step we show that two edges cannot span too many copies of P^2_k.

Lemma 4.11. *For every $\gamma > 0$, $k > \frac{3}{2} \cdot \frac{1+\gamma}{\gamma}$ and $p = n^{-(1-\gamma)/2}$ a.a.s. no pair of edges in $G_{n,p}$ is connected by more than $2n^{k-3}p^{2k-3}$ squares of paths of length k.*

Proof. This follows directly from a theorem of Spencer on the number of graph extensions in the random graph [Spe90]. P^2_k rooted at both of its end-edges is strictly rooted balanced and $n^{k-3}p^{2(k-3)+3} \gg \log n$. Thus the number of squares of paths of length k connecting any two edges is concentrated around its expectation. □

The previous lemma easily implies a weaker version of Lemma 4.7 where the constant 0.52 has to be replaced with a small value that depends on α and η. And this will indeed be the first step in the proof. Next we prove a small lemma which contains two ad hoc arguments that will allow us to go from a small set of the edges in V_i, V_{i+1} to a slightly larger set of edges in V_{i+1}, V_{i+2}.

Lemma 4.12. *Let $\eta, \alpha, \gamma, \varepsilon > 0$ be fixed. For $p = n^{-1/2+\gamma/2}$ in $G_{n,p}$ a.a.s. every copy of $G = (V_1 \cup V_2 \cup V_3, E) \in \tilde{\mathcal{G}}(K_3, n_0, p_0, \varepsilon)$, where $n_0 \geq \eta n$ and $p_0 = \alpha p$, satisfies the following: for every set of edges $\tilde{E} \subseteq E(V_1, V_2)$ denote with $\triangle(\tilde{E})$ the number of edges in $E(V_2, V_3)$ which form a triangle with some edge of $\tilde{E}$ in G. Let $\tilde{V}_2 \subseteq V_2$ denote the vertices which are incident to some edge of $\tilde{E}$ and set $s = \min_{v \in \tilde{V}_2} \deg_{\tilde{E}}(v)$. Then*

$$\triangle(\tilde{E}) \geq \begin{cases} |\tilde{V}_2| n_0 p_0^2/(2p) & \text{if } s \geq \log^2(n) n_0 p/n^\gamma, \\ (1-\varepsilon)^2 \big(|\tilde{V}_2| - 5\varepsilon n_0\big) n_0 p_0 & \text{if } s \geq 2\varepsilon n_0 p_0. \end{cases}$$

Proof. Fix $v \in \tilde{V}_2$. Write $V_1^v = \Gamma_{\tilde{E}}(v)$ and let $V_3^v = V_3 \cap \Gamma(v) \cap \Gamma(V_1^v)$ denote the subset of vertices of V_3 which form a triangle with v and some edge from $\tilde{E}$. By definition of G every edge in $\tilde{E}$ is contained in at least $(1-\varepsilon) n_0 p_0^2$ triangles with V_3 and therefore

$$E(V_1^v, V_3^v) \geq (1-\varepsilon)|V_1^v| n_0 p_0^2. \tag{4.1}$$

A.a.s. in $G_{n,p}$ all disjoints sets A, B of sizes $|A| \geq \log^2(n) n_0 p_0/n^\gamma$ and $|B| \leq b = n_0 p_0^2/(2p)$ have at most $(1+\varepsilon)|A|bp$ edges between them. To see this, observe that $|A|bp \gg \log n \max\{|A|, b\}$ and the claim thus follows from Chernoff's inequality together with a straightforward union bound argument over all sets of size $|A|$ and at most b.

We now consider the two cases. If $s \geq \log^2(n)n_0p_0/n^\gamma$ then (4.1) implies $E(V_1^v, V_3^v) \geq (1-\varepsilon)|V_1^v|n_0p_0^2 = 2(1-\varepsilon)|V_1^v|bp$. For $|V_3^v| \leq b$ this would contradict the bounds from the Chernoff inequality in the previous paragraph; thus $|V_3^v| \geq b$, implying the desired bound.

Now suppose that $s \geq 2\varepsilon n_0 p_0$ and that v is such that its neighboorhoods in V_1 and V_3 are of size $(1 \pm \varepsilon)n_0p_0$ and induce a (ε, p)-lower-regular subgraph. As $|V_1^v| \geq s$ the assumption on p and v imply $|V_1^v| \geq \varepsilon|\Gamma(v) \cap V_1|$. Thus we can apply Lemma 2.5 to deduce that at most $\varepsilon|\Gamma(v) \cap V_3|$ vertices in $\Gamma(v) \cap V_3$ have no neighboor in V_1^v. Thus $|V_3^v| \geq (1-\varepsilon)|\Gamma(v) \cap V_3| \geq (1-\varepsilon)^2 n_0p_0$. By Lemma 2.5 at most $4\varepsilon n_0$ vertices do have neighborhoods of the wrong size in either V_1 or V_3. By definition of $\tilde{\mathcal{G}}$ at most εn_0 vertices have neighborhoods which do not induce an (ε, p)-lower-regular subgraph. Thus $\triangle(\tilde{E}) \geq \left(|\tilde{V}_2| - 5\varepsilon n_0\right)(1-\varepsilon)^2 n_0 p_0$, as claimed. □

With these two lemmas at hand the proof of Lemma 4.7 can be summarized as follows: Use the weak version implied by Lemma 4.11 to expand to a small fraction of the edges spanned by V_k, V_{k+1}. Then invoke Lemma 4.12 (four times!) to expand to a 0.52-fraction of the edges in V_{k+4}, V_{k+5}. The details are given below.

Lemma 4.7. *Let $\eta, \alpha, \gamma > 0$ and $k \geq \frac{3}{\gamma}$ be fixed and let $p = n^{-1/2+\gamma/2}$. For ε small enough depending on α a.a.s. every subgraph $G \subseteq G_{n,p}$ which is from $G \in \tilde{\mathcal{G}}\left(P_{k+4}^2, n_0, \alpha p, \varepsilon\right)$ for some $n_0 \geq \eta n$ satisfies the following: all edges in $G[V_1 \cup V_2]$ expand to at least a 0.52-fraction of the edges in $G[V_{k+4} \cup V_{k+5}]$.*

Proof. Write $p_0 = \alpha p$. Since $G \in \tilde{\mathcal{G}}\left(P_{k+4}^2, n_0, p_0, \varepsilon\right)$ every edge in $G[V_i \cup V_{i+1}]$ forms at least $(1-\varepsilon)n_0p_0^2$ triangles with V_{i+2}. In particular every edge spanned by $V_1 \cup V_2$ is connected to $E(V_k, V_{k+1})$ by $((1-\varepsilon)n_0p_0^2)^{k-1}$ copies of P_k^2. Furthermore since $k \geq \frac{3}{\gamma} > \frac{3}{2} \cdot \frac{1+\gamma}{\gamma}$ by Lemma 4.11 every edge in $E(V_1, V_2)$ is connected to at least

$$\frac{((1-\varepsilon)n_0p_0^2)^{k-1}}{2n^{k-3}p^{2k-3}} = \Theta(n^2p) \gg n^2p/n^{\gamma/2}.$$

distinct edges in $E(V_k, V_{k+1})$ via the square of a path.

Therefore it suffices to show that every set $E_0 \subseteq E(V_k, V_{k+1})$ of size at least $n_0^2p_0/n^{\gamma/2}$ expands to at least a 0.52-fraction of $E(V_{k+4}, V_{k+5})$. To this end we will apply Lemma 4.12 four times to $G \leftarrow G[V_{k+i} \cup V_{k+i+1} \cup V_{k+i+2}]$ for $i \in \{0, 1, 2, 4\}$.

Set $s = \log^2(n)n_0p/n^\gamma$. In $G_{n,p}$ a.a.s. all degrees are bounded by $(1+o(1))np$. Therefore we can find a set of vertices $\tilde{V}_{k+1} \subseteq V_{k+1}$ of size

$$\frac{|E_0| - n_0 s}{\Theta(np)} = \Theta(n^{1-\gamma/2})$$

and an edge set $\tilde{E}_0 \subseteq E_0$ such that each $v \in \tilde{V}_{k+1}$ is incident to exactly s edges of $\tilde{E}_0$.

Apply Lemma 4.12 with $G \leftarrow G[V_k \cup V_{k+1} \cup V_{k+2}]$, $\tilde{V}_2 \leftarrow \tilde{V}_{k+1}$, $\tilde{E} \leftarrow \tilde{E}_0$ and denote the set of edges which form a triangle with some edge of $\tilde{E}_0$ with $E_1 \subseteq E(\tilde{V}_{k+1}, V_{k+2})$. We have $|E_1| \geq \frac{p_0}{2p}|\tilde{V}_{k+1}|n_0 p_0$. Denote with $\tilde{V}_{k+2} \subseteq V_{k+2}$ the set of vertices which are incident to more than s edges in E_1. A.a.s. in $G_{n,p}$ there exists no set S of size $|\tilde{V}_{k+1}| = \Theta(n^{1-\gamma/2})$ such that more than $\sqrt{n}$ vertices have degree at least $2|\tilde{V}_{k+1}|p$ into S. Therefore for ε small enough depending on α

$$|\tilde{V}_{k+2}| \geq \frac{|E_1| - \sqrt{n} \cdot \Theta(np) - n_0 s}{2|\tilde{V}_{k+1}|p} \geq \frac{\frac{p_0}{3p}|\tilde{V}_{k+1}|n_0 p_0}{2|\tilde{V}_{k+1}|p} \geq \frac{p_0^2}{6p^2} n_0 \geq 100\varepsilon n_0.$$

As before pick a subset $\tilde{E}_1 \subseteq E_1$ such that every $v \in \tilde{V}_{k+2}$ is incident to exactly s edges of $\tilde{E}_1$. Apply Lemma 4.12 a second time with $G \leftarrow G[V_{k+1} \cup V_{k+2} \cup V_{k+3}]$, $\tilde{V}_2 \leftarrow \tilde{V}_{k+2}$, $\tilde{E} \leftarrow \tilde{E}_1$ and denote the set of edges which form a triangle with some edge of $\tilde{E}_1$ with $E_2 \subseteq E(V_{k+2}, V_{k+3})$. We have $|E_2| \geq \frac{p_0}{2p}|\tilde{V}_{k+2}|n_0 p_0$. Let $\tilde{V}_{k+3} \subseteq V_{k+3}$ denote the subset of vertices which are incident to more than $s' = 2\varepsilon n_0 p_0$ edges in E_2. As $E(V_{k+2}, V_{k+3})$ is (ε)-regular, by Lemma 2.5 there are at most εn_0 vertices in V_{k+3} with more than $(1+\varepsilon)|\tilde{V}_{k+2}|p_0$ neighbours in $\tilde{V}_{k+2}$. And by definition of (ε)-regularity these vertices are in total incident to at most $\varepsilon n_0(1+\varepsilon)|\tilde{V}_{k+2}|p_0$ edges. Therefore for ε sufficiently small

$$|\tilde{V}_{k+3}| \geq \frac{|E_2| - \varepsilon n_0(1+\varepsilon)|\tilde{V}_{k+2}|p_0 - n_0 s'}{(1+\varepsilon)|\tilde{V}_{k+2}|p_0} \geq \frac{\frac{p_0}{3p}|\tilde{V}_{k+2}|n_0 p_0}{(1+\varepsilon)|\tilde{V}_{k+2}|p_0} \geq \frac{p_0}{4p} n_0 \geq 100\varepsilon n_0.$$

As before pick a subset $\tilde{E}_2 \subseteq E_2$ such that every $v \in \tilde{V}_{k+3}$ is incident to exactly s' edges of $\tilde{E}_2$. Apply Lemma 4.12 a third time with $G \leftarrow G[V_{k+2} \cup V_{k+3} \cup V_{k+4}]$, $\tilde{V}_2 \leftarrow \tilde{V}_{k+3}$, $\tilde{E} \leftarrow \tilde{E}_2$ and denote the set of edges which form a triangle with some edge of $\tilde{E}_2$ with $E_3 \subseteq E(V_{k+3}, V_{k+4})$. We have $|E_3| \geq (1-\varepsilon)^2(|\tilde{V}_{k+2}| - 5\varepsilon n_0)n_0 p_0 \geq \frac{6}{7}|\tilde{V}_{k+2}|n_0 p_0$. Let $\tilde{V}_{k+4} \subseteq V_{k+4}$ denote the subset of vertices which are incident to more than $s' = 2\varepsilon n_0 p_0$ edges in E_3. As before we obtain

$$|\tilde{V}_{k+4}| \geq \frac{|E_3| - \varepsilon n_0(1+\varepsilon)|\tilde{V}_{k+3}|p_0 - n_0 s'}{(1+\varepsilon)|\tilde{V}_{k+3}|p_0} \geq \frac{\frac{5}{7}|\tilde{V}_{k+3}|n_0 p_0}{(1+\varepsilon)|\tilde{V}_{k+3}|p_0} \geq \frac{4}{7} n_0.$$

Applying Lemma 4.12 a fourth time we see E_3 expands to $(1-\varepsilon)^2(|\tilde{V}_{k+4}| - 5\varepsilon n_0)n_0 p_0 > 0.53 n_0^2 p_0$ edges in $E(V_{k+4}, V_{k+5})$. This concludes the proof. □

Chapter 5

A tight Erdős-Pósa function for long cycles

This chapter is joint work with Frank Mousset, Nemanja Škorić, and Felix Weissenberger [MNŠW16].

A classic result of Erdős and Pósa states that any graph either contains k vertex-disjoint cycles or can be made acyclic by deleting at most $\mathcal{O}(k \log k)$ vertices. Birmelé, Bondy, and Reed (2007) raised the following more general question: given numbers l and k, what is the optimal function $f(l,k)$ such that every graph G either contains k vertex-disjoint cycles of length at least l or contains a set X of $f(l,k)$ vertices that meets all cycles of length at least l? In this chapter, we answer that question by proving that $f(l,k) = \mathcal{O}(kl + k \log k)$, which is optimal up to constant factors. As a corollary, the tree-width of any graph G that does not contain k vertex-disjoint cycles of length at least l is of order $\mathcal{O}(kl + k \log k)$. This is also optimal up to constant factors and answers another question of Birmelé, Bondy, and Reed from 2007 [BBR07].

Let $\mathcal{F}$ be any family of graphs. Given a graph G, a subset $X \subseteq V(G)$ is called a *transversal* (of $\mathcal{F}$) if the graph $G - X$ obtained by deleting X does not contain any member of $\mathcal{F}$. We say that $\mathcal{F}$ has the *Erdős-Pósa property* if there exists a function $f \colon \mathbb{N} \to \mathbb{R}$ such that every graph G which does not contain k vertex-disjoint members of $\mathcal{F}$ contains a transversal of size at most $f(k)$.

The study of this property dates back to 1965 when Erdős and Pósa [EP65] showed the following:

Theorem. *Every graph contains either k vertex-disjoint cycles or a set of at most $f(k) = (4+o(1))k\log k$ vertices meeting all its cycles.*

The value of $f(k)$ in this theorem is optimal up to the leading constant. The Erdős-Pósa property is closely related to classical 'covering vs. packing' results in graph theory, such as Kőnig's theorem or Menger's theorem. For example, Kőnig's theorem can be stated as follows: every bipartite graph contains either k vertex-disjoint edges or a set of $f(k) = k$ vertices meeting all the edges. The above result has spawned a long line of papers about the duality between packing and covering of different families of graphs, directed graphs, hypergraphs, rooted graphs, and other combinatorial objects (see a recent survey of Raymond and Thilikos [RT16] for more information).

We are interested in the Erdős-Pósa property for the family $\mathcal{F}_l = \{C_m \mid m \geq l\}$ of cycles of length at least l. A 1988 result of Thomassen [Tho88] implies that for every l, the family $\mathcal{F}_l$ has the Erdős-Pósa property with a function $f(l,k) \in 2^{l^{\mathcal{O}(k)}}$ (though recent results of Chekuri and Chuzhoy make it possible to substantially improve the dependency on k in this bound [CC13]). This result was sharpened by Birmelé, Bondy, and Reed [BBR07] to $f(l,k) \in \mathcal{O}(lk^2)$ in 2007 and by Fiorini and Herinckx [FH14] to $f(l,k) \in \mathcal{O}(lk\log k)$ in 2014. We improve these results to the asymptotically optimal bound $f(l,k) \in \mathcal{O}(kl + k\log k)$, thus settling the question asked in [BBR07] and [FH14].

Theorem 5.1. *For every integer $l \geq 3$, the family $\mathcal{F}_l$ of cycles of length at least l has the Erdős-Pósa property with the function*

$$f(l,k) = \begin{cases} 6kl + 10k\log_2 k + 40k + 10k\log_2\log_2 k & \text{if } k \geq 2, \\ 0 & \text{if } k = 1. \end{cases}$$

Certainly the constant factors in Theorem 5.1 are not optimal. Birmelé, Bondy, and Reed [BBR07] conjectured the correct function in the case $k = 2$ to be $f(l,2) = l$. The complete graph on $2l-1$ vertices shows that, if true, this bound would be tight. Lovász [Lov65] confirmed the conjecture for $l = 3$, while Birmelé [Bir03a] confirmed the cases $l = 4$ and $l = 5$. For larger l, Birmelé, Bondy, and Reed [BBR07] proved that the optimal function satisfies $f(l,2) \leq 2l+3$. This was recently improved by Meierling, Rautenbach and Sasse [MRS14] to $f(l,2) \leq 5l/3 + 29/2$.

There are two constructions which together imply that the function in Theorem 5.1 is asymptotically optimal for large k and l. On the one hand, for all k and l we must have $f(l,k) \geq (k-1)l$, as can be seen from the example of a complete graph on $kl-1$ vertices: this graph does not contain k vertex-disjoint cycles of length at least l, but to remove all cycles of length at least l one must delete $kl-1-(l-1) = (k-1)l$ vertices. This construction also gives the lower bound $f(l,k) \geq \frac{1}{2}(k-1)\log_2 k$ whenever $l \geq \frac{1}{2}\log_2 k$.

On the other hand, for $l < \frac{1}{2}\log_2 k$ we can obtain the lower bound $f(l,k) \geq \frac{1}{8}k\log_2 k$ using the fact that there exist 3-regular graphs on n vertices with girth at least $(1-o(1))\log_2 n$ [ES63]. Indeed for n large enough, let G denote such a graph with girth $g(G)$. Clearly G contains at most $n/g(G)$ vertex disjoint cycles. So fix $k = \lfloor n/g(G) \rfloor + 1 > n/g(G)$ and observe that for n large enough $n \geq \frac{1}{2}k\log_2 k$. All cycles in G have length at least $g(n) > \frac{1}{2}\log_2 k > l$. Thus if X is a transversal of all cycles of length at least l then $G - X$ is a forest. Because G is 3-regular, removing $|X|$ vertices leaves at least $3n/2 - 3|X|$ edges. Since the resulting graph should be a forest, we need $\frac{3}{2}n - 3|X| \leq n - |X|$, and therefore every transversal must have size $|X| \geq \frac{n}{4} \geq \frac{1}{8}k\log_2 k$. This gives the desired lower bound $f(l,k) \geq \frac{1}{8}k\log_2 k$.

Notation All graphs are assumed to be simple unless stated otherwise. However, multigraphs do make an appearance in the proof. We define a multigraph M in the standard way, that is, as an ordered pair (V, E), where V denotes the vertex set of M and E is the multiset of edges of M. For a (multi-)graph G we denote by $V(G)$ and $E(G)$ the vertex set and the edge (multi-)set of G, respectively. Given two multigraphs M_1 and M_2 we write $M_1 \cup M_2$ for the multigraph $M = (V, E)$ where $V = V(M_1) \cup V(M_2)$ and $E = E(M_1) \cup E(M_2)$. In particular, the multiplicity of an edge e in $M_1 \cup M_2$ is equal to the sum of multiplicities of e in M_1 and M_2. We use the standard asymptotic notation $\mathcal{O}$, o, ω and Ω.

Tree-width Our results also imply an asymptotically optimal upper bound on the tree-width $\mathrm{tw}(G)$ of every graph G that does not contain k vertex-disjoint cycles of length at least l. We need the following theorem.

Theorem 5.2 (Birmelé [Bir03b]). *Suppose that G does not contain a cycle of length at least l. Then* $\mathrm{tw}(G) \leq l - 2$.

Generalizing this, Birmelé, Bondy, and Reed proved that any graph G not containing k vertex-disjoint cycles of length at least l has tree-width in $\mathcal{O}(k^2 l)$ [BBR07]. Theorem 5.1 allows us to improve this bound:

Corollary 5.3. *Assume that G does not contain k vertex-disjoint cycles of length at least l. Then* $\mathrm{tw}(G) \in \mathcal{O}(kl + k\log k)$.

Proof. Assume that G does not contain k vertex-disjoint cycles of length at least l. By Theorem 5.1 there is a set $X \subseteq V(G)$ of size $|X| \leq 6kl + 10k\log_2 k + 40k + 10\log_2\log_2 k$ such that $G - X$ does not contain a cycle of length at least l. By Theorem 5.2 we have $\mathrm{tw}(G - X) \leq l - 2$. We can turn a tree-decomposition of $G - X$ into a tree-decomposition of G by adding X to each bag, which gives the bound $\mathrm{tw}(G) \leq \mathrm{tw}(G - X) + |X| \leq (6k+1)l + 10k\log_2 k + 40k + 10\log_2\log_2 k - 2$. □

This is tight in the sense that there are examples of graphs that do not contain k disjoint cycles of length at least l and whose tree-widths are in $\Omega(kl + k \log k)$.

In fact, similar constructions as above work. An example where $\mathrm{tw}(G) \geq kl - 2$ is provided by the complete graph on $kl - 1$ vertices. For $l \leq c \log k$, for sufficiently small positive constant c, we can use fact that there exist constant-degree expander graphs G on n vertices with $g(G) \in \Omega(\log n)$ and $\mathrm{tw}(G) \in \Omega(n)$ (using for example the results in [LPS88] and [BPTW10]). Choosing k such that $k \cdot g(G) \in [n+1, 2n]$ one obtains a graph which does not contain k vertex-disjoint cycles (of any length) but whose tree-width is in $\Omega(k \log k)$.

5.1 Proof of Theorem 5.1

We will use the following lemma.

Lemma 5.4 (Diestel [Die05]). *For each natural number k, let*

$$s_k := \begin{cases} 4k(\log_2 k + \log_2 \log_2 k + 4) & \text{if } k \geq 2, \\ 1 & \text{if } k = 1. \end{cases}$$

Then every 3-regular multigraph on at least s_k vertices contains a set of k vertex-disjoint cycles.

Fix $l \geq 3$ and a graph G. We say that a cycle in G is *long* if it has length at least l, and otherwise we say that it is *short*. By *disjoint*, we always mean *vertex-disjoint*. We assume that G does not contain k disjoint long cycles and show that G contains a transversal of $\mathcal{F}_l$ of size at most $f(l,k)$, where

$$f(l,k) = \begin{cases} 6kl + 10k \log_2 k + 40k + 10k \log_2 \log_2 k & \text{if } k \geq 2, \\ 0 & \text{if } k = 1. \end{cases}$$

The proof is by induction on k, where the base case $k = 1$ is obvious.

If G contains a long cycle C of length at most $6l$ then by induction, $G - V(C)$ contains either $k-1$ disjoint long cycles or a transversal X of size $f(l, k-1)$. In the first case G contains k disjoint long cycles and in the second case $X \cup V(C)$ is a transversal of size $f(l, k-1) + 6l \leq f(l,k)$. Therefore we may assume that every long cycle in G contains strictly more than $6l$ vertices.

Let H denote a maximal subgraph of G with the following properties:

1. all vertices of H have degree 2 or 3 in H;

2. H contains no short cycle.

Similarly as in [EP65], observe that H is the union of a subdivision of a 3-regular multigraph and at most $k-1$ disjoint long cycles. If H contains at least s_k vertices of degree 3 then by Lemma 5.4, it contains k disjoint cycles, which by definition of H are all long. So from now on, we can assume that H contains fewer than s_k vertices of degree 3.

Definition 5.5. *We say that a path P is an H-*path *if its endpoints are distinct vertices of H and if it is internally vertex-disjoint from H. Observe that we allow for $P \subseteq H$ if the length of P is one. We say that P is a* proper H-path *if none of its edges are contained in H.*

For each $i \in \{2,3\}$ let $V_i \subseteq V(H)$ denote the set of vertices with degree i in H. We modify G by removing all edges from $E(G) \setminus E(H)$ that are incident to a vertex from V_3. Note that any transversal of the modified graph can be turned into a transversal of the original graph by additionally removing V_3. Furthermore H is still maximal in the modified graph. From now on we assume that all vertices of V_3 have degree 3 in G. In particular the endpoints of every proper H-path lie in V_2.

This implies that every H-path has length at most l, as otherwise we could add the path to H without violating either the degree or the cycle condition, contradicting the maximality. For the same reason, if P is an H-path with endpoints $s, t \in V(H)$, then there exists a path between s and t in H of length at most l. In fact, as H contains no cycles of length at most $2l$, this path is unique. Thus the following notion is well-defined.

Definition 5.6 (Projection)**.** *Suppose that P is an H-path with endpoints $s, t \in V(H)$. The* projection *of P, denoted by $\pi(P)$, is defined to be the unique path of length at most l between s and t in H.*

Let $C \subseteq G$ be a cycle in G that intersects H in at least two vertices. We define the projection $\pi(C)$ of C as follows. Let $C = P_1 \cup \ldots \cup P_m$ be a decomposition of C into distinct H-paths. Then we define the projection of C to be the multigraph $\pi(C) = \pi(P_1) \cup \cdots \cup \pi(P_m)$.

If P is a path in G with distinct endpoints in H (not necessarily an H-path), then we define the projection analogously: let $P = P_1 \cup \cdots \cup P_m$ be a decomposition into H-paths and define $\pi(P) = \pi(P_1) \cup \cdots \cup \pi(P_m)$.

We remark that in the definition above, the decomposition of a cycle or path into distinct H-paths is unique up to permutation, so that the projection is in fact well-defined.

We claim that information about the length of a cycle can be recovered by looking at the following property of its projection:

Definition 5.7. *A multigraph M is called* even *if the multiplicity of every edge in M is even.*

Lemma 5.8. *Suppose that C is a cycle in G which intersects H in at least two vertices. If $V(\pi(C))$ induces a tree in H, then C is short. If $\pi(C)$ is not even, then C is long.*

Proof. Among all cycles for which the lemma fails, we may pick a cycle C whose decomposition into H-paths minimizes the number of proper H-paths. If none of these H-paths is proper, then $C \subseteq H$ and thus $\pi(C) = C$ is not even and long. Therefore we may assume that C contains at least one proper H-path P.

Note that $\pi(P)$ is a path in H with the same endpoints as P. If C contains all edges of $\pi(P)$, then we actually have $C = P \cup \pi(P)$ and so the projection of C is even. Moreover, the length of C is $|P| + |\pi(P)| \leq 2l$. Since all long cycles have length greater than $6l$ we see that C must be short and we are done. Therefore, we can assume that at least one edge of $\pi(P)$ does not belong to C. Since $\pi(P)$ is a path with endpoints in $V(C)$, there exists a path $P' \subseteq \pi(P)$ with endpoints $s, t \in C$ which is internally vertex-disjoint from C and whose edges are not edges of C. Let $P_1, P_2 \subseteq C$ denote the two internally disjoint s,t-paths in C and consider the two cycles $C_1 := P_1 \cup P'$ and $C_2 := P_2 \cup P'$. Observe that since $\pi(P') \subseteq \pi(C)$ we have $V(\pi(C)) = V(\pi(C_1)) \cup V(\pi(C_2))$. Additionally, the parity of each edge in $\pi(C)$ is equal to the parity of the same edge in $\pi(C_1) \cup \pi(C_2)$.

We are now ready to prove the first claim. Assume that C is long and that $H[V(\pi(C))]$ is a tree. Then the projections of C_1 and C_2 induce trees as well and therefore both cycles contain at least one proper H-path. In particular both cycles contain strictly fewer proper H-paths than C. Furthermore, at least one of the two cycles has length at least $|C|/2$. Since C has length at least $6l$, this cycle is still long, which contradicts the minimality of C.

For the second claim, assume that C is short and that $\pi(C)$ is not even. Observe that we have $|C_i| \leq |C| + |P'| \leq l + |P'|$ for each $i \in \{1, 2\}$. As P' is a subgraph of $\pi(P)$ we know that $|C_i| \leq 2l$, which implies that both cycles C_i are short. Since H contains only long cycles, this means that both C_1 and C_2 contain at least one proper H-path, so both C_1 and C_2 contain fewer proper H-paths than C. Finally, as $\pi(C)$ is not even and since each edge in $\pi(C_1) \cup \pi(C_2)$ has the same parity as the same edge in $\pi(C)$, at least one of $\pi(C_1)$ and $\pi(C_2)$ is not even. This contradicts the minimality of C. □

Definition 5.9. *Let $X \subseteq V(H) \cup E(H)$ be a set of vertices and edges of H. Let us denote by $G - X$ and $H - X$ the graphs obtained by deleting from G and H the edges and vertices in X.*

Then we say that X is π-preserving *if $P \subseteq G - X$ implies $\pi(P) \subseteq H - X$ for all H-paths P.*

The following lemma is the crucial ingredient of the proof of Theorem 5.1. For now, we only state the lemma; the proof is given in the next subsection.

Lemma 5.10. *There exists a π-preserving set X for which $H - X$ is a forest and such that $|X \cap (V_2 \cup E(H[V_2]))| \leq 3|V_3|/2 + k$.*

We can now finish the proof of Theorem 5.1. By the lemma, there exists a π-preserving set X for which $H - X$ is a forest and such that $|X \cap (V_2 \cup E(H[V_2]))| \leq 3|V_3|/2 + k$.

Suppose that C is a cycle in $G - X$ that intersects H at least twice. Then because X is π-preserving, we know in particular that all vertices of $V(\pi(C))$ are contained in the same component of $H - X$. Since each component of $H - X$ is a tree, the graph $H[V(\pi(C))]$ must also be a tree. Thus, by Lemma 5.8, the cycle C is short. It follows that every long cycle in $G - X$ intersects H at most once.

To construct the transversal we define the following two sets.

1. Let $X' \subseteq V(H)$ be a set containing the vertices in $X \cap V_2$ and also containing one endpoint of each edge in $X \cap E(H[V_2])$.
2. Let $Z \subseteq V(H)$ denote the set of all vertices $z \in V(H)$ for which there exists some long cycle C_z such that $V(C_z) \cap V(H) = \{z\}$.

We now claim that $V_3 \cup X' \cup Z$ is a transversal of all long cycles. Every long cycle C in G intersects H at least once since otherwise C could be added to H. If C intersects H exactly once than it intersects Z. If C intersects H at least twice then, by the observation above, this means that C intersects X. But then C must intersect either V_3 or X'. So $V_3 \cup X' \cup Z$ is a transversal in G. Recall that in the beginning, we modified the graph G by removing all edges of $E(G) \setminus E(H)$ that are incident to a vertex of V_3. Since we remove V_3 anyway, $V_3 \cup X' \cup Z$ is also a transversal in the original graph.

It remains to bound the size of this transversal. If for some $z \neq z' \in Z$ the cycles C_z, $C_{z'}$ intersect, then one can see that the assumption $|C_z|, |C_{z'}| \geq 6l$ implies that $C_z \cup C_{z'}$ contains a z-z'-path of length at least l, which contradicts the fact that there are no H-paths of length l or longer. Therefore $\{C_z \mid z \in Z\}$ is a collection of vertex disjoint long cycles and in particular $|Z| < k$. Furthermore, we have $|X'| \leq 3|V_3|/2 + k$. Since $|V_3| < s_k$, we get

$$|V_3 \cup X' \cup Z| < |V_3| + \frac{3}{2}|V_3| + 2k < \frac{5}{2}s_k + 2k \leq f(l,k).$$

This completes the proof of Theorem 5.1. However, we still need to prove Lemma 5.10.

5.2 Finding a small π-preserving set

Let us call a π-preserving set *valid* if it contains at most one edge or vertex from every component of $H[V_2]$. There exists at least one valid π-preserving set: the empty set. Let X denote a valid π-preserving set of maximal size. Since H is the disjoint union of a subdivision of a 3-regular graph on $|V_3|$ vertices and fewer than k cycles, the fact that X is valid immediately implies that $|X \cap (V_2 \cup E(H[V_2]))| \leq 3|V_3|/2 + k$. We will show that $H - X$ is a forest. Assume towards a contradiction that $H - X$ contains a cycle. Among all such cycles let $C = (x_0 \ldots, x_n)$ denote one of minimum length. Since H only contains long cycles, we have $|C| > 6l$. Let $B_{H-X}(C, l)$ denote the ball of radius l around C in $H - X$, i.e., the set of vertices that have distance at most l to a vertex of C in $H - X$. Let H_C be the subgraph of $H - X$ induced by $B_{H-X}(C, l)$. We now have some claims about the structure of H_C. First of all, it is clear that $C \subseteq H_C$. Moreover:

Claim 5.11. *In H_C, every vertex v has a unique nearest vertex in C (which may be v itself if $v \in V(C)$).*

Proof. This is clear if $v \in V(C)$, so assume otherwise. By the definition of H_C, the distance from v to any nearest vertex on C is at most l. If there are two nearest vertices of v on C, then using the shortest path between them on C, we obtain a cycle of length at most $|C|/2 + 2l < |C|$ in H_C, where the inequality follows from $|C| > 4l$. But this would contradict the minimimality of C. □

This claim shows in particular that the projection map $p\colon V(H_C) \to V(C)$ taking vertices of H_C to their unique nearest vertex in C is well-defined. In fact, the preimages of this projection have rather nice properties:

Claim 5.12. *The following hold:*

(i) for every $x \in V(C)$, the graph $H_C[p^{-1}(x)]$ is a tree;

(ii) for distinct vertices $x, y \in C$, there are no edges between $p^{-1}(x)$ and $p^{-1}(y)$ in $H_C \setminus C$;

Proof. For (i) simply observe that the diameter of $H_C[p^{-1}(x)]$ is at most $2l$, so any cycle would be of length at most $4l$; however H does not contain cycles that are this short. The argument for (ii) is similar to the argument in the proof of Claim 5.11: if such an edge exists, then H_C contains a cycle of length at most $|C|/2 + 2l + 1 < |C|$, using $|C| > 6l > 4l + 2$. This would contradict the minimality of C. □

By the above claim, the graph H_C is just the cycle C with trees attached at every vertex. In particular, $H_C - e$ is a tree for any edge $e \in E(C)$.

We will now construct a larger graph G_C where $H_C \subseteq G_C \subseteq G - X$ as follows. Let $D = \{x_l, x_{l+1}, \dots, x_{n-l}\}$. For each $x \in D$, let $P(x)$ be the set of all H-paths P in $G - X$ such that $x \in V(\pi(P))$. Since X is π-preserving, every such path satisfies $\pi(P) \subseteq H - X$. We then define

$$G_C = H_C \cup \bigcup_{x \in D} P(x).$$

It is worth noting that $G_C \cap H = H_C$. Let us further denote by e^* the edge $\{x_0, x_n\}$. We have the following claim, whose proof we postpone to the end of the section.

Claim 5.13. *For every proper H-path $P \subseteq G_C$, we have $\pi(P) \subseteq H_C - e^*$. In particular, every path in $G_C - e^*$ with endpoints in $V(H_C)$ projects to a subgraph of $H_C - e^*$.*

Using the claim, we can now finish the proof of Lemma 5.10. Let $A = \{x_0, \ldots, x_l\}$ and $B = \{x_{n-l}, \ldots, x_n\}$. We claim that $G_C - e^*$ does not contain two internally disjoint A-B-paths. Suppose for a contradiction that P_1 and P_2 are two such paths, where we can assume that both P_1 and P_2 intersect both A and B in exactly one vertex. By Claim 5.13 both paths project onto a subgraph of $H_C - e^*$, which implies in particular that $\pi(P_1) \cup \pi(P_2)$ does not contain e^*. We can combine P_1 and P_2 into a cycle in $G_C - e^*$ by adding the shortest paths between the endpoints of P_1 and P_2 in $C[A]$ and $C[B]$, respectively. The projection of this cycle lies in $H_C - e^*$ which is a tree. Thus, by Lemma 5.8, this cycle is short. In particular we have $|P_1| < l$. However, as the projection of P_1 avoids e^*, connecting the endpoints of P_1 using the shortest path in $C[A \cup B]$ results in a cycle C_1 whose projection is not even, since the multiplicity of e^* in $\pi(C_1)$ is one. So by Lemma 5.8 the cycle C_1 is long. Since any path in $C[A \cup B]$ is of length at most $2l + 1$, the length of C_1 is bounded by $|P_1| + 2l + 1 \leq 3l$. This is a contradiction since G does not contain a long cycle of length at most $3l$. We conclude that there are no two internally disjoint A-B-paths in $G_C - e^*$.

By Menger's theorem, $G_C - e^*$ contains a single-vertex A-B-cut. Denote the cut vertex by x. Because C contains an A-B-path from x_l to x_{n-l}, we must have $x \in D$. Thus x has degree at least two in $H_C - e^*$. We distinguish two cases, depending on whether x has degree two or three in $H_C - e^*$.

First, assume that x is a vertex of degree two in $H_C - e^*$. By the maximality of X, we know that $X \cup \{x\}$ is not a valid π-preserving set. One possibility is that $X \cup \{x\}$ is π-preserving, but it is not valid. Since X is valid on its own, this means that x belongs to a component of $H[V_2]$ which intersects X. But since x belongs to the cycle $C \subseteq H - X$, this is impossible. Thus it must be that $X \cup \{x\}$ is not a π-preserving set. Since X by itself is π-preserving, the definition implies that there exists an H-path P in $G - X$ such that $x \in V(\pi(P))$ and $x \notin V(P)$. Note that $P \subseteq G_C - e^*$ by the definition of G_C. Since x has degree two in $H_C - e^*$ and since $H_C - e^*$ is a tree, we know that $H_C - e^* - x$ breaks into exactly two trees T_1 and T_2. As $x \in V(\pi(P))$ and $\pi(P) \subseteq H_C - e^*$, it must be that P has one endpoint in T_1 and the other in T_2. It follows that $T_1 \cup T_2 \cup P$ is a connected graph and thus $T_1 \cup T_2 \cup P$ must contain an A-B-path. This is a contradiction with the fact that x separates A and B in $G_C - e^*$ and completes the proof in this case.

Next, assume that x is a vertex of degree three in $H_C - e^*$. Recall that then x has degree 3 in G as well, and in particular no proper H-path has x as an endpoint. As in the previous case, $H_C - e^* - x$ breaks into exactly three trees T_1, T_2 and T_3. Let x^- and x^+ be the neighbours of x on C. Without loss of generality assume that $x^- \in T_1$ and $x^+ \in T_3$. Consider any H-path $P \subseteq G - X$ such that $x \in V(\pi(P))$ (so in particular $P \subseteq G_C - e^*$). Since $\pi(P)$ is a path contained in $H_C - e^*$, the endpoints of P must be in two different trees. However, it cannot happen that one endpoint of P is in T_1 and the other in T_3, since then $T_1 \cup T_3 \cup P$ would contain an A-B-path, contradicting the fact that x is an A-B-cut vertex in $G_C - e^*$. For the same reason there are no two H-paths $P_1, P_2 \subseteq G - X$ such that $x \in V(\pi(P_1)) \cap V(\pi(P_2))$ and such that P_1 has endpoints in T_1 and T_2 and P_2 has endpoints in T_2 and T_3. We conclude that there is some $j \in \{1,3\}$ such that all H-paths P with $x \in V(\pi(P))$ have one endpoint in T_2 and the other endpoint in T_j. Without loss of generality, assume $j = 1$. We now claim that by adding the edge $\{x, x^+\}$ to X we get again a π-preserving set. If we assume otherwise, then there exists a H-path P in $G - (X \cup \{\{x, x^+\}\})$ whose projection contains $\{x, x^+\}$. This is not possible, as such a path satisfies $x \in V(\pi(P))$ and has one endpoint lying in T_3. Thus $X \cup \{\{x, x^+\}\}$ is π-preserving. In fact, since the edge $\{x, x^+\}$ is incident to the degree-three vertex x, it is automatically a valid π-preserving set. This contradicts the maximality of X and completes the proof of Lemma 5.10.

We now present the missing proof of Claim 5.13.

Proof of Claim 5.13. The second statement clearly follows from the first.

We start the proof with an observation about the H-paths in $G - X$ whose endpoints lie in $V(H_C)$. Let $P \subseteq G - X$ be such a path with endpoints $a, b \in V(H_C)$. Since X is π-preserving, we have $\pi(P) \subseteq H - X$. Note that we know of at least one path in $H - X$ from a to b: the path $Q = Q_1Q_2Q_3$, where Q_1 is the shortest path in H_C from a to $p(a)$, Q_2 is the shortest path from $p(a)$ to $p(b)$ on C, and Q_3 is the shortest path from $p(b)$ to b in H_C. This path Q is contained in H_C and it has length at most $2l + |C|/2$. The important observation is that $\pi(P) = Q$. Indeed, if this were not so, then $Q \cup \pi(P) \subseteq H - X$ would contain a cycle of length at most $2l + |C|/2 + |\pi(P)| \leq 3l + |C|/2 < |C|$ (using $|C| > 6l$), contradicting the minimality of C.

This shows in particular that for every H-path $P \subseteq G - X$ with endpoints in $V(H_C)$, we have $\pi(P) \subseteq H_C$. It remains to show the stronger statement that if additionally $P \subseteq G_C$, then $\pi(P) \subseteq H_C - e^*$.

To obtain a contradiction, assume that $P \subseteq G_C$ is an H-path with endpoints $a, b \in V(H_C)$ such that $\{x_0, x_n\} \in E(\pi(P))$. Since the projection of P has length at most l, and by the observation above, we can assume without loss of generality that $p(a) \in \{x_0, \ldots, x_l\}$ and $p(b) \in \{x_{n-l}, \ldots, x_n\}$. Now let c be the neighbour of a on P. The edge $\{a, c\}$ belongs to G_C but not to H_C, so there is some $x \in D = \{x_l, \ldots, x_{n-l}\}$ and some H-path $P' \in P(x)$ such that $\{a, c\} \in E(P')$. Let d be the other endpoint of P' (one endpoint is a). Since $|C| > 6l$, $x \in D$, $x \in V(\pi(P'))$, and $p(a) \in \{x_0, \ldots, x_l\}$, we must have $p(d) \in \{x_l, \ldots, x_{2l}\}$. The union of P and P' contains an H-path with endpoints d and b. However, the projection of this H-path has length at least $\min\{l + 1, |C| - 3l\} > l$, which is impossible. □

5.3 Some open questions

The main contribution of this chapter is showing the asymptotically optimal Erdős-Pósa function for the case of long cycles and growing k and ℓ and thus answering the question asked in [BBR07] and [FH14]. We conclude by mentioning some open problems:

- ***S*-cycles:** Kakimura, Kawarabayashi, and Marx [KKM11] introduced a different generalization of the standard Erdős-Pósa theorem. They considered the family of S-cycles, i.e., all cycles of a graph which intersect a specified set S, and proved that such a family of cycles has the Erdős-Pósa property. Their result was later improved by Pontecorvi and Wollan [PW12], resulting in the following theorem:

  Theorem 5.14. *For any graph and any vertex subset S, the graph either contains k vertex-disjoint S-cycles or a vertex set of size $\mathcal{O}(k \log k)$ that meets all S-cycles.*

  Since the vertex set S can be the vertex set of the whole graph, this result is asymptotically tight. In 2014, Bruhn, Joos, and Schaudt [BJS14] combined the family of S-cycles with the family of long cycles and proved that the family of S-cycles of length at least ℓ has the Erdős-Pósa property with $f(k, \ell) = O(\ell k \log k)$. Thus, it is natural to ask if Theorem 5.1 generalizes to S-cycles as well.

- **Edge-version:** A family of graphs $\mathcal{F}$ is said to have the *edge-Erdős-Pósa* property if there exists a function $f\colon \mathbb{N} \to \mathbb{R}$ such that every graph G which does not contain k edge-disjoint members of $\mathcal{F}$ contains a set of $f(k)$ edges which meets all copies of members of $\mathcal{F}$ in G. Since the pioneering papers by Erdős and Pósa [EP62, EP65] it has been known that the family of cycles has the edge-Erdős-Pósa property as well. Namely, the following is true:

  Theorem. *Any graph G contains either k edge-disjoint cycles or a set of $(2 + o(1))k \log k$ edges meeting all its cycles.*

  Pontecorvi and Wollan [PW12] generalized this result to the case of S-cycles by a clever reduction to the standard vertex-version of the problem. Already Birmelé et al. [BBR07] asked if the family of long cycles has the edge-Erdős-Pósa property. Unfortunately, the gadget trick from [PW12] breaks down in the case of long cycles and thus nothing is known in this scenario. It would be interesting to see if our approach could be applied for proving that the family of long cycles has the edge-Erdős-Pósa property.

Chapter 6

Smaller subgraphs of minimal degree k

The results of this chapter are based on a joint work with Frank Mousset, Rajko Nenadov, Yury Person and Nemanja Škorić [MNN+16].

It is easy to show that every graph on n vertices and at least $2n-2$ edges contains a subgraph of minimal degree 3. More generally for $k \geq 2$ any graph on n vertices and at least $t_k(n) := (k-1)(n-k+2) + \binom{k-2}{2}$ edges contains a subgraph of minimal degree k. This statement is best possible in two ways: (1) there exists graphs with $t_k(n)-1$ edges which do not contain a subgraph of minimal degree k, and (2) there exists graphs with $t_k(n)$ edges without a *proper* induced subgraph of minimal degree k. For example the wheel $W(1,n) = K_1 + C_{n-1}$ has exactly $2n-2$ edges and minimal degree 3, but contains no proper induced subgraph with minimal degree 3. A similar construction is available for all k (consider the generalized wheel $W(k-2,n) = K_{k-2} + C_{n-k+2}$).

Erdős conjectured that the presence of even a single additional edge allows one to find a much smaller subgraph:

Conjecture 6.1 ([EFRS90, Erd93]). *For $k \geq 2$ there exist an $\varepsilon_k > 0$ such that every graph on n vertices and $t_k(n)+1$ edges contains a subgraph of minimal degree k with at most $(1-\varepsilon_k)n$ vertices.*

As far as we know the only progress on this conjecture is the following theorem due to Erdős, Faudree, Rousseau and Schelp from 1990:

Theorem 6.2 ([EFRS90]). *For $k \geq 2$, let G be a graph on n vertices and $t_k(n) + 1$ edges. Then G contains a subgraph of order at most $n - \left\lfloor \sqrt{n}/\sqrt{6k^3} \right\rfloor$ with $\delta(G) \geq k$.*

Our contribution is a short argument that one can remove $\Omega(n/\log n)$ vertices and remain with a subgraph of minimal degree k.

Theorem 6.3. *For $k \geq 2$, let G be a graph on n vertices and $t_k(n) + 1$ edges. Then G contains a subgraph of order at most $n - n/(12(k+1)^5 \log n)$ with $\delta(G) \geq k$.*

Proof

Our proof is split into two cases depending on the number of degree k vertices. Already in the proof of Theorem 6.2 the authors observed that the conjecture holds if the number of degree k is not too large.

Lemma 6.4 (Lemma 4 in [EFRS90]). *For $k \geq 2$, let G be a graph on n vertices and $t_k(n) + 1$ edges with $\delta(G) \geq k$. If for some positive $\alpha < 1/(2k)$, G has at most αn vertices of degree k, then G has a subgraph H of order at most $n - (1 - 2\alpha k)n/(8k^2)$ with $\delta(H) \geq k$.*

We will not improve on this result. Instead we handle the complementary case where the number of degree k vertices is large (say at least $n/3k$).

We use standard graph notation: for a graph H we write v_H for the number of vertices and e_H for the number of edges in H. $V_i(H)$ denotes the set of vertices of H with degree exactly i. Similarly $V_{\leq i}(H)$ denotes the set of vertices of degree at most i. For a vertex v we denote its neighborhood by $\Gamma_H(v)$.

We now introduce the notion of a *cover*. Suppose that H is a graph and that $S \subseteq V(H)$ is a subset of its vertices. A supergraph $\tilde{H} \supseteq H$ is called an (H, S)*-cover* if every vertex $v \in S \cup (V(\tilde{H}) \setminus V(H))$ has degree at least k in $\tilde{H}$. The use of this somewhat strange definition will become apparent later. For now we will just show the following lemma:

Lemma 6.5. *Suppose that H is a graph which does not contain a subgraph of minimal degree k. Then there exists a subset of the vertices S with $S \subseteq V_{\leq k-1}(H)$ of cardinality at most $2((k-1)v_H - e_H)$ such that every (H, S)-cover contains a subgraph of minimal degree k.*

Proof. The idea of the proof is simple. We reconstruct H from the empty graph by repeatedly joining a vertex of degree at most $k - 1$. This is possible since H does not contain a subgraph of minimal degree k. A new vertex v is added to S if it has degree at most $k - 2$ into the graph constructed so far or if one of its neighbors was in S before. A vertex is removed from S once it reaches degree k. We now show that this construction has the desired properties.

For a graph G define the potential function

$$\varphi(G) := 2((k-1)v_G - e_G) - \sum_{i=0}^{k-2}(k-1-i)|V_i(G)|.$$

We will use induction on the number of vertices to prove there exists such a set S of size at most $\varphi(H)$. For the base case, in which H is the empty graph, we have $\varphi(H) = 0$ and can set $S = \emptyset$. Otherwise H is non-empty and since H is not of minimal degree k it must contain a vertex v of degree at most $k-1$. By induction there exists a set S' of cardinality at most $\varphi(H-v)$ such that every $(H-v, S')$-cover contains a subgraph of minimal degree k.

We now define the set S based on S' by

$$S := (S' \cup I_v) \setminus V_k(H),$$

where

$$I_v = \begin{cases} \{v\} & \text{if } \deg(v) \leq k-2 \text{ or } \Gamma(v) \cap S' \neq \emptyset, \\ \emptyset & \text{otherwise.} \end{cases}$$

Observe that $S \subseteq V_{\leq k-1}(H)$ is satisfied by construction. To check that S is not too big note that

$$\varphi(H) - \varphi(H-v) = (k-1-\deg(v)) + |V_{\leq k-1}(H) \cap \Gamma(v)| \geq 0$$

and therefore if $v \notin S$ then $|S| \leq |S'| \leq \varphi(H') \leq \varphi(H)$. If $v \in S$ then we distinguish two cases:

1. If $\deg(v) = k-1$ and $|V_{\leq k-1}(H) \cap \Gamma(v)| = 0$ then there must exists $u \in S' \cap \Gamma(v)$ with degree at least k. By definition of S', u has degree at most $k-1$ in $H-v$ and thus u has degree exactly k in H. In particular $u \in S' \setminus S$ and thus $|S| \leq |S'|$.

2. If $\deg(v) \leq k-2$ or $\deg(v) = k-1$ and $|V_{\leq k-1}(H) \cap \Gamma(v)| \geq 1$ then $\varphi(H) - \varphi(H-v) \geq 1$. Since $|S| \leq |S'| + 1$ always holds we are done.

Now suppose that $\tilde{H} \supseteq H$ is an (H, S)-cover. We claim that then either $\tilde{H}$ is an $(H-v, S')$-cover or $\tilde{H} - v$ is an $(H-v, S')$-cover.

Indeed if $\deg_{\tilde{H}}(v) \geq k$ then all vertices in $V(\tilde{H}) \setminus (V(H) - v)$ have degree at least k in $\tilde{H}$. Furthermore since $\tilde{H}$ is an (H, S)-cover all vertices in S have degree at least k in $\tilde{H}$. Finally $S' \setminus S \subseteq V_k(H) \cap \Gamma(v)$ and thus all of S' has degree at least k in $\tilde{H}$ and $\tilde{H}$ is an $(H-v, S')$-cover.

If $\deg_{\tilde{H}}(v) \leq k-1$ then $v \notin S$ and thus $\deg_H(v) = k-1 = \deg_{\tilde{H}}(v)$ and $\Gamma_H(v) \cap S' = \emptyset$. In particular $S' = S$ and v is not incident to any vertex of $S \cup (V(\tilde{H} - v) \setminus V(H - v))$. Therefore all vertices of $S' \cup (V(\tilde{H} - v) \setminus V(H-v))$ have degree at least k in $\tilde{H} - v$ and $\tilde{H} - v$ is an $(H-v, S')$-cover.

In both cases we obtain a subgraph of minimal degree k by induction. $\square$

With Lemma 6.5 at hand we can give a short outline of the proof of Theorem 6.3 We will pick a suitable subgraph $H \subset G$ of size at most $n - \Omega(n/\log n)$ and obtain a set S as described by Lemma 6.5. Intuitively since S is small we will be able to find an (H, S)-cover $\tilde{H}$, $H \subseteq \tilde{H} \subseteq G$ which is not much larger than H. $|G \setminus \tilde{H}|$ is then a lower bound on the number of vertices that can be removed from G.

Proof of Theorem 6.3. The proof will use induction on the number of vertices. Since $t_k(n-1) = t_k(n) - (k-1)$ we may assume that G does not contain a vertex of degree at most $k-1$.

If G contains fewer than $n/3k$ vertices of degree k then by Lemma 6.4 G contains a subgraph of order at most $(1 - 1/(24k^2))n \leq (1 - 1/(12(k+1)^5 \log n)n$ with minimum degree k. So from now on assume that G contains at least $n/3k$ vertices of degree k.

We define the notion of a *good* set of vertices. A set of vertices is deemed to be a good if it can be constructed from the following recursive procedure: Initially all sets consisting of a single degree k vertex are good sets. For a good set A and a vertex $v \notin A$ we deem $A \cup \{v\}$ to be good if v is incident to at most $k-1$ edges which do not end in A. Additionally for any two good sets A, B their union $A \cup B$ is good if at least one edge is incident to both A and B (that is the case if $A \cap B \neq \emptyset$ or if $E(A, B) \neq \emptyset$).

It is easy to prove by induction that every good set touches at most $(k-1)|C| + 1$ edges. Furthermore by construction if C is a good set then there exists a good set of size at most $\lceil |C|/2 \rceil$. Thus if we find some good set containing more than $n/\log n$ vertices then there exists a good set C satisfying $n/\log n \leq |C| \leq 3n/\log n$. For such a set $G - C$ consists of at least two vertices and at least $t_k(n - |C|)$ edges. But then $G - C$ would contain a (not necessarily proper) subgraph of minimal degree k and we would be done. Finally if there exists a good set which is incident to at most $(k-1)|C|$ edges then $G - C$ has at least $t_k(n - |C|) + 1$ edges and by induction on the number of vertices $G - C$ contains a small subgraph of minimal degree k. Therefore we may assume that every good set C is incident to exactly $(k-1)|C| + 1$ edges.

Let $\mathcal{C} = \{C_1, C_2, \ldots, \}$ denote a maximal collection of maximal good sets. This implies in particular that the C_i are pairwise disjoint and no edges lie between two such sets. Since G contains at least $n/3k$ degree k vertices we have $\sum_{C \in \mathcal{C}} |C| \geq n/3k$. For $1 \leq i \leq \log n$ let $\mathcal{C}_i \subseteq \mathcal{C}$ denote the subfamily of sets with sizes between 2^{i-1} and 2^i. By the pidgeonhole principle there exists i such that $\sum_{C \in \mathcal{C}_i} |C| \geq n/(3k \log n)$.

Define $H = G - \bigcup_{C \in \mathcal{C}_i} C$. The number of edges in H is exactly

$$e_H = t_k(n) + 1 - \sum_{C \in \mathcal{C}_i} [(k-1)|C| + 1] = t_k(v_H) - |\mathcal{C}_i| + 1.$$

If H contains a subgraph of minimal degree k then we are done. Otherwise we invoke

Lemma 6.5 to obtain a set $S \subseteq V_{\leq k-1}(H)$ of size at most

$$2((k-1)v_H - e_H) = 2((k-1)(k-2) - \binom{k-2}{2} + |\mathcal{C}_i| - 1) \leq 2|\mathcal{C}_i| + k^2,$$

such that every (H, S)-cover contains a subgraph of minimal degree k.

For every $s \in S$ we have $\deg_H(s) \leq k-1$, but $\delta(G) \geq k$. Therefore the family $\mathcal{C}(s) \subseteq \mathcal{C}_i$ of good sets whose neighborhood contains s is nonempty. The maximality of the good sets implies that there cannot be $C \in \mathcal{C}(s)$ such that the degree of s in $G - C$ is at most $k-1$. Therefore we must have $\deg_G(s) \geq k+1$ and $|\mathcal{C}(s)| \geq 2$ for every $s \in S$.

Now let $\mathcal{C}'(s) \subseteq \mathcal{C}(s)$ be a minimal subfamily with the property that for every $C^* \in \mathcal{C}'(s)$ the degree of s in $G - \bigcup_{C \in \mathcal{C}_i \setminus (\mathcal{C}'(s) - C^*)} C$ is at least k. By the previous observation the set $\mathcal{C}'(s) = \mathcal{C}(s)$ satisfies the condition so such a family always exists. Furthermore it is easy to see that $|\mathcal{C}'(s)| \leq k+1$.

Let A denote an auxiliary graph on the vertex set $\mathcal{C}_i$ with edges obtained by adding a clique on $\mathcal{C}'(s)$ for every $s \in S$. Since $|\mathcal{C}'(s)| \leq k+1$ the number of edges in A is at most

$$|S| \binom{k+1}{2} \leq (2|\mathcal{C}_i| + k^2) \frac{(k+1)^2}{2}$$

and thus the average degree of A is bounded by $2(k+1)^2 + k^2(k+1)^2 \leq 2(k+1)^4 - 1$. Using Turán's theorem we see that A must contain an independent set I of size at least $|\mathcal{C}_i|/(2(k+1)^4)$. We claim that $G' := G - \bigcup_{C \in I} C$ is an (H, S)-cover and thus must contain a subgraph of minimal degree k. Indeed for every $s \in S$ the independent set I contains at most one good set from $\mathcal{C}'(s)$. By definition of $\mathcal{C}'(s)$ this implies that s has minimum degree k in G'. Furthermore since there are no edges between two maximal good sets $C, C' \in \mathcal{C}$ all vertices in $\bigcup_{C \in \mathcal{C}_i \setminus I} C$ have degree at least k in G' and G' is a (H, S)-cover.

Therefore we can remove at least a $1/(2(k+1)^4)$ fraction of the good sets in $\mathcal{C}_i$ and remain with a graph containing a subgraph of minimal degree k. Since the sizes of the good sets in $\mathcal{C}_i$ differ by at most a factor of 2 and, by choice of i, contain in total at least $n/(3k \log n)$ vertices we have removed at least $n/(12(k+1)^5 \log n)$ vertices. □

Bibliography

[ABET] Peter Allen, Julia Böttcher, Julia Ehrenmüller, and Anusch Taraz, *The bandwidth theorem in sparse graphs*, in preperation. 47

[BB10] József Balogh and Jane Butterfield, *Online Ramsey games for triangles in random graphs*, Discrete Mathematics **310** (2010), no. 24, 3653 – 3657. 4, 14

[BBR07] Etienne Birmelé, J. Adrian Bondy, and Bruce A. Reed, *The Erdős-Pósa property for long circuits*, Combinatorica **27** (2007), no. 2, 135–145. 6, 59, 60, 61, 69

[BCS11] József Balogh, Béla Csaba, and Wojciech Samotij, *Local resilience of almost spanning trees in random graphs*, Random Structures & Algorithms **38** (2011), no. 1-2, 121–139. 5

[Bir03a] Etienne Birmelé, Ph.D. thesis, Université Lyon 1, 2003. 60

[Bir03b] ______, *Tree-width and circumference of graphs*, Journal of Graph Theory **43** (2003), 24–25. 61

[BJS14] Henning Bruhn, Felix Joos, and Oliver Schaudt, *Long cycles through prescribed vertices have the Erdős-Pósa property*, arXiv preprint arXiv:1412.2894 (2014). 69

[BKT13] Julia Böttcher, Yoshiharu Kohayakawa, and Anusch Taraz, *Almost spanning subgraphs of random graphs after adversarial edge removal*, Combinatorics, Probability and Computing **22** (2013), no. 05, 639–683. 5

[BMS12] Michael Belfrage, Thorsten Mütze, and Reto Spöhel, *Probabilistic one-player Ramsey games via deterministic two-player games*, SIAM Journal on Discrete Mathematics **26** (2012), no. 3, 1031–1049. 4, 14

[BMS15] József Balogh, Robert Morris, and Wojciech Samotij, *Independent sets in hypergraphs*, Journal of the American Mathematical Society **28** (2015), no. 3, 669–709. MR 3327533 15, 18

[Bol83] Béla Bollobás, *Almost all Regular Graphs are Hamiltonian*, European Journal of Combinatorics **4** (1983), no. 2, 97 – 106. 46

[Bol98] ______, *Modern Graph Theory*, Graduate Texts in Mathematics, vol. 184, Springer-Verlag, New York, 1998. 5, 45

[Bol01] ______, *Random graphs*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2001. 3

[BPTW10] Julia Böttcher, Klaas P. Pruessmann, Anusch Taraz, and Andreas Würfl, *Bandwidth, expansion, treewidth, separators and universality for bounded-degree graphs*, European Journal of Combinatorics **31** (2010), no. 5, 1217–1227. 62

[BSKS11] Sonny Ben-Shimon, Michael Krivelevich, and Benny Sudakov, *Local resilience and Hamiltonicity Maker–Breaker games in random regular graphs*, Combinatorics, Probability and Computing **20** (2011), no. 02, 173–211. 5

[BT87] Béla Bollobás and Arthur G. Thomason, *Threshold functions*, Combinatorica **7** (1987), no. 1, 35–38. 3

[CC13] Chandra Chekuri and Julia Chuzhoy, *Large-treewidth graph decompositions and applications*, Proceedings of the forty-fifth annual ACM symposium on Theory of Computing (STOC'13), 2013, pp. 291–300. 6, 60

[Con09] David Conlon, *A new upper bound for diagonal Ramsey numbers*, Annals of Mathematics (2009), 941–960. 2

[Die05] Reinhard Diestel, *Graph theory*, third ed., Graduate texts in mathematics, Springer Heidelberg, 2005. 62

[Dir52] Gabriel A. Dirac, *Some Theorems on Abstract Graphs*, Proceedings of the London Mathematical Society **s3-2** (1952), no. 1, 69–81. 4, 45

[DKMS08] Domingos Dellamonica, Yoshiharu Kohayakawa, Martin Marciniszyn, and Angelika Steger, *On the resilience of long cycles in random graphs*, Electronic Journal of Combinatorics **15** (2008), R32. 5

[EFRS90] Paul Erdős, Ralph J. Faudree, Cecil C. Rousseau, and Richard H. Schelp, *Subgraphs of minimal degree k*, Discrete Mathematics **85** (1990), no. 1, 53 – 58. 71, 72

[EP62] Paul Erdős and Lajos Pósa, *On the maximal number of disjoint circuits of a graph*, Publ. Math. Debrecen **9** (1962), 3–12. 69

[EP65] ______, *On independent circuits contained in a graph*, Canad. J. Math **17** (1965), 347–352. 6, 59, 63, 69

[ER59] Paul Erdős and Alfréd Rényi, *On random graphs I*, Publ. Math. Debrecen **6** (1959), 290–297. 2

[ER60] ______, *On the evolution of random graphs*, Publ. Math. Inst. Hungar. Acad. Sci **5** (1960), 17–61. 2

[Erd93] Paul Erdős, *Some of my favorite solved and unsolved problems in graph theory*, Quaestiones Mathematicae **16** (1993), no. 3, 333–350. 71

[ES63] Paul Erdös and Horst Sachs, *Reguläre Graphen gegebener Taillenweite mit minimaler Knotenzahl*, Wiss. Z. Martin-Luther-Univ. Halle-Wittenberg Math.-Natur. Reihe **12** (1963), 251–257. 61

[FH14] Samuel Fiorini and Audrey Herinckx, *A tighter Erdős-Pósa function for long cycles*, J. Graph Theory **77** (2014), no. 2, 111–116. 6, 60, 69

[FKR$^+$03] Ehud Friedgut, Yoshiharu Kohayakawa, Vojtěch Rödl, Andrzej Ruciński, and Prasad Tetali, *Ramsey games against a one-armed bandit*, Combinatorics, Probability and Computing **12** (2003), no. 6, 515–545. 3, 4, 13

[FNP$^+$15] Asaf Ferber, Rajko Nenadov, Ueli Peter, Andreas Noever, and Nemanja Škorić, *Robust hamiltonicity of random directed graphs*, Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, 2015, pp. 1752–1758. 5

[Gil59] Edgar N. Gilbert, *Random graphs*, The Annals of Mathematical Statistics **30** (1959), no. 4, 1141–1144. 2

[GKRS07] Stefanie Gerke, Yoshiharu Kohayakawa, Vojtěch Rödl, and Angelika Steger, *Small subsets inherit sparse ε-regularity*, Journal of Combinatorial Theory Series B **97** (2007), no. 1, 34–56. 51

[GS05] Stefanie Gerke and Angelika Steger, *The sparse regularity lemma and its applications*, Surveys in Combinatorics 2005, Cambridge University Press, 2005. 9, 11

[HSS15] Dan Hefetz, Angelika Steger, and Benny Sudakov, *Random directed graphs are robustly Hamiltonian*, Random Structures & Algorithms (2015). 5

[JŁR11] Svante Janson, Tomasz Łuczak, and Andrzej Ruciński, *Random graphs*, John Wiley & Sons, 2011. 3, 15

[KKM11] Naonori Kakimura, Ken-ichi Kawarabayashi, and Dániel Marx, *Packing cycles through prescribed vertices*, J. Combin. Theory Ser. B **101** (2011), no. 5, 378–381. MR 2802885 69

[KŁR97] Yoshiharu Kohayakawa, Tomasz Łuczak, and Vojtěch Rödl, *On K_4-free subgraphs of random graphs*, Combinatorica **17** (1997), 173–213. 18

[KLS10] Michael Krivelevich, Choongbum Lee, and Benny Sudakov, *Resilient pancyclicity of random and pseudorandom graphs*, SIAM Journal on Discrete Mathematics **24** (2010), no. 1, 1–16. 5

[KLS14] M. Krivelevich, C. Lee, and B. Sudakov, *Robust Hamiltonicity of Dirac graphs*, Transactions of the American Mathematical Society **366** (2014), no. 6, 3095–3130. 5

[Koh97] Yoshiharu Kohayakawa, *Szemerédi's regularity lemma for sparse graphs*, Foundations of Computational Mathematics (1997), 216–230. 9, 10, 47

[KR05] Andrzej Kurek and Andrzej Ruciński, *Two variants of the size Ramsey number*, Discussiones Mathematicae Graph Theory **25** (2005), 141–149. 14

[KS83] János Komlós and Endre Szemerédi, *Limit distribution for the existence of Hamiltonian cycles in a random graph*, Discrete Mathematics **43** (1983), no. 1, 55 – 63. 46

[KSS96] János Komlós, Gábor N. Sárközy, and Endre Szemerédi, *On the square of a Hamiltonian cycle in dense graphs*, Random Structures & Algorithms **9** (1996), no. 1-2, 193–211. 5, 46

[KSS97] János Komlós, Gábor N. Sárközy, and Endre Szemerédi, *Blow-up Lemma*, Combinatorica **17** (1997), no. 1, 109–123. 5, 46

[Lov65] László Lovász, *On graphs not containing independent circuits*, Mat. Lapok **16** (1965), 289–299. 60

[LPS88] Alexander Lubotzky, Ralph Phillips, and Peter Sarnak, *Ramanujan graphs*, Combinatorica **8** (1988), no. 3, 261–277. 62

[ŁRV92] Tomasz Łuczak, Andrzej Ruciński, and Bernd Voigt, *Ramsey properties of random graphs*, J. Combin. Theory Ser. B **56** (1992), no. 1, 55–68. 3

[LS12] Choongbum Lee and Benny Sudakov, *Dirac's theorem for random graphs*, Random Structures & Algorithms **41** (2012), no. 3, 293–305. 5, 46

[MNN$^+$16] Frank Mousset, Rajko Nenadov, Andreas Noever, Nemanja Škorić, and Yury Person, *Smaller subgraphs of minimal degree k*, in preparation. 8, 71

[MNŠW16] Frank Mousset, Andreas Noever, Nemanja Škorić, and Felix Weissenberger, *A tight Erdős-Pósa function for long cycles*, ArXiv e-prints (2016). 7, 59

[MRS14] Dirk Meierling, Dieter Rautenbach, and Thomas Sasse, *The Erdős-Pósa property for long circuits*, J. Graph Theory **77** (2014), no. 4, 251–259. 60

[MSS09a] Martin Marciniszyn, Reto Spöhel, and Angelika Steger, *Online Ramsey games in random graphs*, Combinatorics, Probability and Computing **18** (2009), 271–300. 4, 13, 15, 16

[MSS09b] ______, *Upper bounds for online Ramsey games in random graphs*, Combinatorics, Probability and Computing **18** (2009), 259–270. 4, 14, 20

[Noe] Andreas Noever, *Online Ramsey Games for more than two colors*, Random Structures & Algorithms, to appear. 7, 13

[NŠ16a] Rajko Nenadov and Nemanja Škorić, *Powers of cycles in random graphs and hypergraphs*, ArXiv e-prints (2016). 5, 46

[NS16b] Andreas Noever and Angelika Steger, *Local resilience for squares of almost spanning cycles in sparse random graphs*, ArXiv e-prints (2016). 7, 45

[PS97] Alessandro Panconesi and Aravind Srinivasan, *Randomized distributed edge coloring via an extension of the Chernoff–Hoeffding bounds*, SIAM Journal on Computing **26** (1997), no. 2, 350–368. 24

[PW12] Matteo Pontecorvi and Paul Wollan, *Disjoint cycles intersecting a set of vertices*, J. Combin. Theory Ser. B **102** (2012), no. 5, 1134–1141. MR 2959394 69

[Ram30] Frank P. Ramsey, *On a problem of formal logic*, Proceedings of the London Mathematical Society **2** (1930), no. 1, 264–286. 1

[RR93] Vojtěch Rödl and Andrzej Ruciński, *Lower bounds on probability thresholds for Ramsey properties*, Combinatorics, Paul Erdős is eighty, Vol. 1, Bolyai Soc. Math. Stud., 1993, pp. 317–346. 3

[RR94] ______, *Random graphs with monochromatic triangles in every edge coloring*, Random Structures Algorithms **5** (1994), no. 2, 253–270. 3

[RR95] ______, *Threshold functions for Ramsey properties*, Journal of the American Mathematical Society **8** (1995), no. 4, 917–942 (English). 3, 4, 14

[RT16] Jean-Florent Raymond and Dimitrios. M. Thilikos, *Recent techniques and results on the Erdős Pósa property*, ArXiv e-prints (2016). 6, 60

[Spe75] Joel Spencer, *Ramsey's theorem – a new lower bound*, Journal of Combinatorial Theory, Series A **18** (1975), no. 1, 108–115. 2

[Spe90] Joel Spencer, *Counting extensions*, Journal of Combinatorial Theory, Series A **55** (1990), no. 2, 247 – 255. 17, 55

[ST15] David Saxton and Andrew Thomason, *Hypergraph containers*, Inventiones mathematicae (2015), 1–68 (English). 15, 18, 19, 37, 38, 42

[SV08] Benny Sudakov and V.H. Vu, *Local resilience of graphs*, Random Structures & Algorithms **33** (2008), no. 4, 409–433. 5

[Sze75] Endre Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 299–345. 5, 9, 46

[Sze76] ———, *Regular partitions of graphs*, Problèmes combinatoires et théorie des graphes, Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976, pp. 399–401. 5, 9, 46

[Tho88] Carsten Thomassen, *On the presence of disjoint subgraphs of a specified type*, J. Graph Theory **12** (1988), no. 1, 101–111. 6, 60

[Wes00] Douglas B. West, *Introduction to graph theory*, 2 ed., Prentice Hall, 2000. 5, 45