Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Robustness of Pósa's conjecture

Master Thesis

Manuela Fischer

October 19, 2016

Advisors: Prof. Dr. Angelika Steger, Nemanja Škorić

Department of Computer Science, ETH Zürich

Abstract

The kth power of a cycle is obtained by adding an edge between all pairs of vertices whose distance on the cycle is at most k. In 1962, Pósa conjectured that a graph G on n vertices contains a square of a Hamilton cycle if it has minimum degree $\delta(G) \geq \frac{2}{3}n$, and, 11 years later, Seymour claimed that $\delta(G) \geq \frac{k}{k+1}n$ is sufficient for the appearance of a kth power of a Hamilton cycle for any $k \geq 2$. The so-called Pósa-Seymour conjecture could be resolved for sufficiently large graphs by Komlós, Sárközy, and Szemerédi in 1998 and the complete $(k+1)$-partite graph having k sets of size $\frac{n-1}{k+1}$ witnesses optimality of the constant $\frac{k}{k+1}$.

The main purpose of the thesis is to extend this statement to the random graph setting. We prove that for every $\alpha, \beta, \varepsilon \in (0,1)$ and $p^2 \geq n^{-1+\varepsilon}$ a.a.s. every subgraph of $\mathcal{G}(n,p)$ with minimum degree $\delta(G) \geq \left(\frac{2}{3}+\alpha\right)np$ and minimum codegree $\delta^2(G) \geq \beta n p^2$ contains a square of a Hamilton cycle.

A simple application of the First Moment method, which yields that $\mathcal{G}(n,p)$ a.a.s. does not contain any square of a Hamilton cycle for $p^2 \ll n^{-1}$, reveals almost optimality of the edge probability in this result and the tightness of Pósa's conjecture shows that the constant in the minimum degree is asymptotically best possible. Moreover, the assumption on the codegrees cannot be omitted, as without this condition every square of a Hamilton cycle can easily be obstructed by deleting all edges inside the neighborhood of one vertex.

Acknowledgements

The author is greatly indepted to Prof. Dr. Angelika Steger and Nemanja Škorić for suggesting this problem, many enlightening discussions, and the fruitful workshop in Buchboden, where a non-negligible part of the ideas emerged. It is their guidance as well as their seemingly infinite supply of knowledge and patience that made this thesis possible.

Contents

Chapter 1

Introduction

Ever since the inception of graph theory, the question whether a host graph G contains a certain graph H as subgraph has been intensively investigated. In the study of random graphs, the analysis of thresholds for the appearance of constant-size graphs H in a random graph $G \sim \mathcal{G}(n, p)$ has been initiated by Erdös and Rényi [ER60] in 1960 and settled in full generality by Bollobás [Bol81] in 1981. Whereas this is by now considered a standard result, the containment problem for large graphs turned out to be surprisingly difficult. The answer is known for very few graphs H_n, both for dense [KO09] as well as for sparse random host graphs [JKV08, Kri10].

1.1 Hamilton cycles

One of the simplest spanning subgraph and simultaneously also one of the most central notions in graph theory is the Hamilton cycle, a cycle passing through every vertex exactly once. Even before Sir William Rowan Hamilton invented the icosian game in 1857, the problem of Hamiltonicity had been studied [Big81], and ever since then has attracted a lot of attention. It is one of the NP-complete problems listed by Karp [Kar72] in 1972, exposing the elusiveness of finding a general classification of Hamiltonian graphs. Nevertheless, several sufficient conditions [BC76, Bon96, CE72, Chv72, Ore60, Tut56] have been established. A classic result of Dirac [Dir52] (with a surprisingly simple proof provided in [Wes01]) states that every graph on $n \geq 3$ vertices with minimum degree at least $\frac{n}{2}$ contains a Hamilton cycle. The complete bipartite graph on $\frac{n-1}{2}$ and $\frac{n+1}{2}$ vertices (assuming n is odd) reveals optimality of the constant $\frac{1}{2}$, as it has minimum degree $\frac{n-1}{2}$ but none of its Hamilton paths can be closed to a cycle. Actually, there exist graphs with minimum degree $\frac{n}{2} - 1$ that are not even connected (consider the union of two cliques of size $\frac{n}{2}$ for even n).

The random graph $\mathcal{G}(n,p)$ a.a.s. contains a Hamilton cycle [Bol84, KS83] if $p \geq \frac{\log n + \log\log n + \omega(n)}{n}$. This is best possible, as for $p \leq \frac{\log n + \log\log n - \omega(n)}{n}$ a.a.s. $\mathcal{G}(n,p)$ has a vertex of degree at most one, which forms a local obstruction for Hamiltonicity.

1.2 Powers of Hamilton cycles

For $k \in \mathbb{N}$, the kth power of a cycle, also called k-cycle, is obtained by adding an edge between all pairs of vertices whose distance on the cycle is at most k. A graph is called k-Hamiltonian if it contains a kth power of a Hamilton cycle. An overview over several sufficient conditions for k-Hamiltonicity can be found in [Gou03].

Pósa [Erd64] conjectured in 1962 that a graph G on n vertices contains a square of a Hamilton cycle if $\delta(G) \geq \frac{2}{3}n$ and, 11 years later, Seymour [Sey73] claimed for $k \geq 2$ that G is k-Hamiltonian if $\delta(G) \geq \frac{k}{k+1}n$. This so-called Pósa-Seymour conjecture thus can be seen as a generalization of Dirac's theorem to all powers of Hamilton cycles. The complete $(k+1)$-partite graph with k subsets of size $\frac{n-1}{k+1}$ and one of size $\frac{n+k}{k+1}$ (assuming $n-1$ is a multiple of $k+1$) does not contain a $(k+1)$-clique-covering and hence no Hamilton k-path, witnessing optimality of the constant $\frac{k}{k+1}$. It required the development of several powerful tools, such as Szemerédi's Regularity Lemma [Sze75a, Sze75b] and the Blow-Up Lemma [KSS97], before the Pósa-Seymour conjecture could be resolved. After an approximate version [KSS98a] and the special case for square-cycles [KSS96] could be shown, Komlós, Sárközy, and Szemerédi [KSS98b] provided a proof of the conjecture for sufficiently large graphs. Later, Levitt, Sárközy, and Szemerédi [LSS10] and Chau, DeBiasio, and Kierstead [CDK11] developed alternative proof techniques improving this result to smaller values of n for the case of Hamilton square-cycles.

The threshold for k-Hamiltonicity in random graphs, however, is not yet completely understood. As the expected number of Hamilton k-cycles in $\mathcal{G}(n,p)$ is $\frac{1}{2}(n-1)!p^{kn}$, a simple application of the First Moment method shows that for $p \ll n^{-1/k}$ a.a.s. there is no Hamilton k-cycle. The result of Nenadov and Škorić [NŠ16a], that for $p^k \geq \frac{C\log^8 n}{n}$ the random graph $\mathcal{G}(n,p)$ contains a kth power of a Hamilton cycle, hence determines the threshold up to a polylogarithmic factor. This approximate threshold is the best known.

1.3 Robustness measures

A different direction of research is concerned with the question of how strongly a graph possesses a certain property $\mathcal{P}$ or, put differently, how much a graph needs to be changed in order to destroy $\mathcal{P}$. This is particularly interesting (and closely related to the notion of fault tolerance [ACK$^+$00]) in the context of monotonically increasing graph properties $\mathcal{P}$ and alteration by edge deletion. Roughly speaking, robustness then indicates the number of edges that must be removed from G to obstruct $\mathcal{P}$.

1.3.1 Robustness with respect to random edge deletions

In the following, a graph G is said to robustly possess $\mathcal{P}$ if random removal of edges does not destroy it. More precisely, let $G_p \sim \mathcal{G}(G, p)$ be the random subgraph of G obtained by keeping each edge of G independently with probability p. If G_p a.a.s. has $\mathcal{P}$, then G is robust with respect to $\mathcal{P}$ for p.

Krivelevich, Lee, and Sudakov [KLS14] showed that Dirac graphs (that is, graphs on $n \geq 3$ vertices with minimum degree at least $\frac{n}{2}$) are robustly Hamiltonian for all $p \gg \frac{\log n}{n}$. This result extends Dirac's theorem (for $p = 1$, they are equivalent) to a robust version and establishes the correct order of magnitude for p due to the aforementioned threshold in $\mathcal{G}(n, p)$. As we will see, also Pósa's conjecture can be generalized to a robust version.

Theorem 1.1 *(Robust version of Pósa's conjecture)*

Let $n \in \mathbb{N}$ be sufficiently large, $\alpha, \varepsilon \in (0, 1)$, and G a graph on n vertices with minimum degree $\delta(G) \geq \left(\frac{2}{3} + \alpha\right) n$. Then $\mathcal{G}(G, p)$ a.a.s. contains a square of a Hamilton cycle, provided that $p^2 \geq n^{-1+\varepsilon}$.

For $p = 1$, Theorem 1.1 corresponds to an approximate version of Pósa's conjecture, as for example investigated in [KSS98a]. The parameter p is only slightly above its optimal value, since for $p^2 \ll \frac{1}{n}$ even $\mathcal{G}(n, p)$ is a.a.s. not 2-Hamiltonian.

Another way to interpret the setting of Theorem 1.1 is the following. Start with a complete graph K_n on n vertices and allow an adversary to delete some of the edges without violating the assumptions of (an approximate version of) Pósa's conjecture. In other words, let him remove a subgraph H with $\Delta(H) \leq \left(\frac{2}{3} - \alpha\right) n$, and then consider a random subgraph of the resulting graph $G := K_n \setminus H$.

In a similar manner, our definition of robustness can be restated. Allow an adversary to remove a subgraph H from K_n without destroying property $\mathcal{P}$. The resulting graph $G := K_n \setminus H$ then robustly possesses $\mathcal{P}$ if a random subgraph of G satisfies $\mathcal{P}$.

1.3.2 Local resilience

A slightly stronger version of this notion of robustness can be obtained by reversing the order of adversarial and random removal of edges. First consider a random subgraph $G_0 \sim \mathcal{G}(K_n, p) = \mathcal{G}(n, p)$ of K_n and then let the adversary delete some of the edges by removing a graph H from G_0. Does the resulting graph $G' := G_0 \setminus H$ still possess $\mathcal{P}$? This alternative form of robustness measure is also known under the name *local resilience* of $\mathcal{G}(n, p)$.

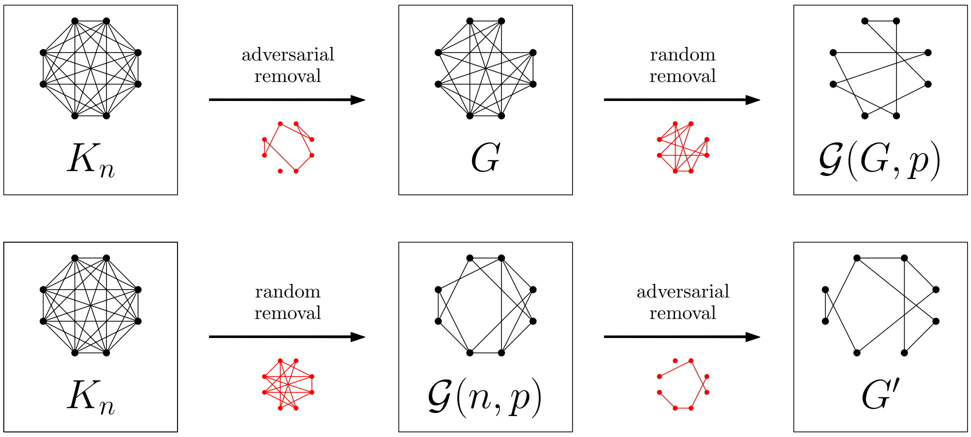

Figure 1.1: Comparison of the two mentioned settings for measuring robustness. Starting from a complete graph K_n, either first an adversary is allowed to delete a subgraph and then some of the edges are randomly removed, or, vice versa, an adversary can delete a subgraph from a random subgraph of K_n. The first setting is mainly used for robustness of $G := K_n \setminus H$, while the second one comes into play when studying (local) resilience of $\mathcal{G}(n, p)$.

The concept of resilience was introduced by Sudakov and Vu [SV08a] and has been studied on many different graph models (besides deterministic graphs and the binomial random graphs mainly on random regular graphs [BSKS11a], random directed graphs [FNP+15, HSS15], and random uniform hypergraphs [FNP+15]) with respect to various conditions, especially the containment of (almost) spanning structures (for instance perfect matchings [SV08b], packings of triangles [BLS12], directed Hamilton cycles [FNP+15], and many more [BCS11, BKT09, DKMS08, KLS10, LS12a]).

The local resilience of a graph G with respect to $\mathcal{P}$ is the minimum number r such that by deleting at each vertex v of G at most an r-fraction of the edges incident to v one can obtain a graph that does not satisfy $\mathcal{P}$. For random graphs $G \sim \mathcal{G}(n, p)$, the local resilience is defined in terms of the edge probability p. It is natural to expect that there exists a threshold p_0 such that for $p \gg p_0$ the local resilience of $\mathcal{G}(n, p)$ with respect to $\mathcal{P}$ is a.a.s. equal to the local resilience of the complete graph K_n, while for $p \ll p_0$ the random graph $\mathcal{G}(n, p)$ a.a.s. does not satisfy $\mathcal{P}$ at all.

For Hamilton cycles, Lee and Sudakov [LS12b] indeed showed that the local resilience of $\mathcal{G}(n,p)$ is $\frac{1}{2} - o(1)$ for $p \gg \frac{\log n}{n}$, which determines the threshold up to polynomial factors. For $p = 1$, their result corresponds to Dirac's theorem. This connection is very natural and actually most of the resilience statements can be viewed as a generalization of some classic result from graph theory to random graphs.

Thus, one might be tempted to assume that this also holds true for the Pósa-Seymour conjecture, namely that for $p^k \geq \frac{\text{polylog} n}{n}$ the local resilience of $\mathcal{G}(n,p)$ with respect to k-Hamiltonicity is $\frac{1}{k+1} - o(1)$. However, this is not anywhere near the case. In fact, for any $p = o(1)$, the resilience for k-Hamiltonicity is $o(1)$, as by deleting all edges inside the neighborhood of a vertex, thus by locally removing at most np^2 edges, one can obstruct any spanning k-path. Figure 1.2 provides more details on how this can be achieved.

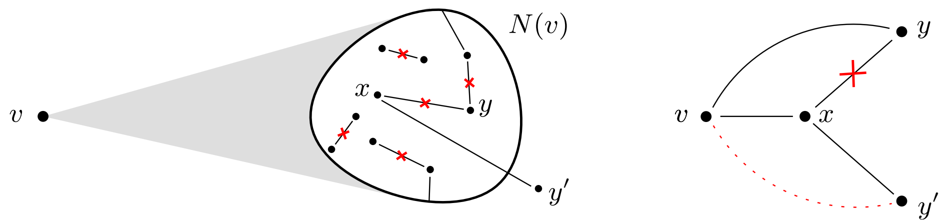

Figure 1.2: The occurrence of any Hamilton k-cycle can be prevented by deleting all edges inside the neighborhood $N(v)$ of an arbitrary vertex v. Indeed, whichever $x \in N(v)$ is chosen, there cannot exist a vertex y simultaneously connected to v and x, continuing a power of a path. Moreover, as $N(v)$ has size roughly np, for each vertex there are removed no more than np^2 incident edges.

The main issue hence seems to be the missing guarantee on the number of common neighbors. Whereas in the setting of Theorem 1.1 the joint neighbhorhood of k vertices is implicitly given to be non-empty (the minimum degree assumption ensures a lower bound of $\frac{1}{k+1}n$ on the codegree (that is, the size of the common neighborhood) of k vertices in G, which cannot drop far below $\frac{1}{k+1}np^k \gg 1$ after random removal), it can be completely erased in a subgraph of $\mathcal{G}(n,p)$. The crucial difference is that it is not possible for the adversary to delete all edges inside the neighborhood of a single vertex if he has to decide on the edges before the neighborhood is determined. The permission of removal after exposure of all randomness, however, gives more choice and control to the adversary, making the occurrence of powers of Hamilton cycles significantly harder.

In order to obtain a non-trivial result for $\mathcal{G}(n,p)$, a weaker notion of resilience is required where the aforementioned problem is not faced anymore. One possibility is to, additionally to the minimum degree condition, explicitly impose a lower bound on the size of the joint neighborhood of any k vertices. In the main result of the thesis, we consider this modified resilience notion of $\mathcal{G}(n,p)$ with respect to square-Hamiltonicity, where we let p range over (almost) all values for which $\mathcal{G}(n,p)$ is expected to contain a square of a Hamilton cycle.

Theorem 1.2 *(Pósa's conjecture for random graphs)*

Let $n \in \mathbb{N}$ be sufficiently large and $\alpha, \beta, \varepsilon > 0$. Consider a random graph $G_0 \sim \mathcal{G}(n,p)$ for $p^2 \geq n^{-1+\varepsilon}$ and a subgraph $H \subseteq G_0$ with maximum degree $\Delta(H) \leq \left(\frac{1}{3} - \alpha\right) np$ and maximum codegree $\Delta^2(H) \leq (1-\beta)np^2$. Then a.a.s. the graph $G := G_0 \setminus H$ contains a square of a Hamilton cycle.

Due to the existence of graphs with minimum degree only slightly below $\frac{2}{3}np$ without a square of a cycle (for instance, a random subgraph of the complete tripartite graph on $\frac{n-1}{3}$, $\frac{n-1}{3}$, and $\frac{n+2}{3}$ vertices, assuming $n-1$ is a multiple of 3), the minimum degree condition is almost optimal, and the non-existence of Hamilton square-cycles in $\mathcal{G}(n,p)$ for $p^2 \ll \frac{1}{n}$ shows that p is almost best possible. Moreover, the assumption on the codegree cannot be omitted by the above observation, or at least would have to be replaced by some other condition ruling out the possibility of square-isolated vertices.

An attentive reader might have realized that Theorem 1.1 is implied by Theorem 1.2. Indeed, $\mathcal{G}(G,p)$ can be interpreted as a subgraph of $\mathcal{G}(n,p)$ and the minimum degree condition $\delta(G) \geq \left(\frac{2}{3} + \alpha\right) n$ implies a lower bound $\delta^2(G) \geq \left(\frac{1}{3} + 2\alpha\right) n$ on the minimum codegree in G, which, by Chernoff's inequality and a union bound over all vertices, results in lower bounds $\delta(\mathcal{G}(G,p)) \geq \left(\frac{2}{3} + \alpha - o(1)\right) np$ and $\delta^2(\mathcal{G}(G,p)) \geq \left(\frac{1}{3} + 2\alpha - o(1)\right) np^2$ on the minimum degree and codegree in $\mathcal{G}(G,p)$. A different, more intuitive and less technical, way to see this is that random removal after adversarial deletion can be emulated by just letting an adversary remove all the edges from $\mathcal{G}(n,p)$ that are not present in $\mathcal{G}(G,p)$. This is precisely due to the aforementioned advantage an adversary is given by allowing him to remove edges after the exposure of all randomness.

1.4 Proof techniques

For a random subgraph $\mathcal{G}(G,p)$ it is convenient to first study the quantity of interest in G (employing methods for dense graphs) and then apply probabilistic tools to draw conclusions for $\mathcal{G}(G,p)$. However, this is not possible for subgraphs of $\mathcal{G}(n,p)$, since the adversary can interfere after random removal. That is why the arguments for the resilience result are almost entirely of a combinatorial nature. The terminology and the proof setup are strongly based on the work of Nenadov and Škorić [NŠ16a], also borrowing some of the ideas from [FNP$^+$15].

The most crucial ingredient thereby is the *absorber method* (also known under the name *reservoir method*), a concept for extending almost spanning to spanning structures. It was introduced as a general method by Rödl, Runcinski, and Szemerédi [RRS06], but also implicitly used before [EGP91, Kri97], and has been employed for the containment problem of various types of spanning subgraphs in dense graphs [KSS98a, LSS10, RRS09]. Although additional difficulties arise in the context of sparse graphs, the method is believed to have significant further potential even in this setting, and several results of this type are known [ABKP15, Mon14].

We adopt the absorbing structures from [NŠ16a], stemming from Kühn and Osthus [KO12] who first studied the absorber method for powers of paths. In this context, an absorbing structure in a graph G for a reservoir set $X \subseteq V(G)$ is a k-path $P_X \subseteq V(G) \setminus X$ which can absorb any subset $X' \subseteq X$ of vertices into it without changing the endpoints of the path, resulting in a k-path $P^*_{X'}$ with $V(P^*_{X'}) = V(P_X) \cup X'$. The k-Hamiltonicity problem of G thus can be reduced to two easier ones: finding an absorbing structure P_X *(Absorbing Lemma)* and an almost spanning k-path P covering all vertices in $V(G) \setminus X \setminus V(P_X)$ *(Covering Lemma)*. Employing the *Connecting Lemma*, which enables the disjoint connection of pairs of vertices by powers of paths, these two subgraphs can be connected to a k-cycle C. Absorbing the uncovered vertices $X' := V(G) \setminus V(C) \subseteq X$ into P_X and hence C, this k-cycle can be extended to a Hamilton k-cycle. While it is typical for proofs based on the absorber method to have some form of Connecting Lemma (in order to connect the almost spanning and the absorbing structure), we also make use of the Connecting Lemma to find the absorbing structure, as opposed to many other proofs relying on probabilistic arguments only.

1.5 Structure of the thesis

Chapter 2 introduces the notation, basic definitions as well as tools of broad applicability, while Chapter 3 serves the purpose of presenting techniques and auxiliary results specially tailored to our setting and proof method. In Chapter 4, first the absorber method is formally introduced and then a proof of the Absorbing Lemma relying on the Connecting Lemma is given. The main technical ingredient, that is, the proof of the Connecting Lemma, is presented in Chapter 5. After having settled all the required auxiliary results, Theorem 1.2 is proved in Chapter 6, putting everything together. Finally, in Chapter 7, our findings are summarized and potential extensions and variations including possibly arising problems are discussed.

Chapter 2

Preliminaries

2.1 Notation and definitions

For a natural number $n \in \mathbb{N} := \{1, 2, \dots\}$, we let $[n]$ denote the set $\{1, \dots, n\}$. We abbreviate $x \in [(a-b)c, (a+b)c]$ as $x = (a \pm b)c$ and $x \notin [(a-b)c, (a+b)c]$ as $x \neq (a \pm b)c$ for $x, a, b, c \in \mathbb{R}$. All logarithms are with natural base e. We employ the standard asymptotic notation $\mathcal{O}, o, \Omega, \omega$, and Θ, following [JŁR11], and write $f \ll g$ if $f = o(g)$ and $f \gg g$ if $g \ll f$.

For a set A and a natural number $k \in \mathbb{N}$, we use $\binom{A}{k}$ to denote the set of all ordered k-element subsets of A and $A^k := A \times \cdots \times A$ for the k-fold Cartesian product of A. The members of $\binom{A}{k}$ and A^k are called k-tuples. We say that a set S of k-tuples is disjoint if no element appears twice in S, neither inside a tuple nor among different tuples. Where it simplifies notation, we think of a tuple as a set and apply set operations to it. For a subset $B \subseteq A$ and an element $x \in A$, we write (B, x) to denote the set $\{(b, x) \mid b \in B\}$.

We adopt the standard graph theory notation of [Die00]. In the following, let G be a simple undirected graph with vertex set $V(G)$ and edge set $E(G)$. We use $v(G) := |V(G)|$ and $e(G) := |E(G)|$ to denote the size of the vertex and edge set, respectively.

The restriction $G[X]$ of G to a subset $X \subseteq V(G)$ is the subgraph induced by X, that is, a graph with vertex set X and edge set $E(G) \cap \binom{X}{2}$. By a copy of a given graph H inside G we mean any (not necessarily induced) subgraph of G that is isomorphic to H. If such a copy exists, we then say that there is an embedding of H in G and write $H \subseteq G$.

In the following, let $X, Y \subseteq V(G)$ be (not necessarily disjoint) non-empty vertex sets. We let $E_G(X, Y) := \{(x, y) \in E(G) \mid x \in X, y \in Y\}$ and $e_G(X, Y) := |E_G(X, Y)|$. Moreover, we let $N_G(X, Y) := \{y \in Y \mid \exists x \in X\colon \{x, y\} \in E(G)\}$ stand for the set of all neighbors of X in Y, write $N_G(X) := N_G(X, V(G))$, and make use of the abbreviations $N_G(v, X) := N_G(\{v\}, X)$ and $N_G(v) :=$

$N_G(\{v\})$ for the neighborhood of a vertex $v \in V(G)$. The sizes of the neighborhoods of a single vertex are denoted by $d_G(v, Y) := |N_G(v, Y)|$ and $d_G(v) := |N_G(v)|$, and are called degrees. We use $\delta(G, X) := \min\{d_G(v, X) \mid v \in V(G)\}$, $\delta(G) := \delta(G, V(G))$ for the minimum degree and $\Delta(G, X) := \max\{d_G(v, X) \mid v \in V(G)\}$, $\Delta(G) := \Delta(G, V(G))$ for the maximum degree of G. For a tuple $v := (v_1, v_2) \in \binom{V(G)}{2}$, we make use of the notation $N^2_G(v, X) := \bigcap_{i \in \{1,2\}} N_G(v_i, X)$, $N^2_G(v) := N^2_G(v, V(G))$, $d^2_G(v, X) := |N^2_G(v, X)|$, $d^2_G(v) := |N^2_G(v)|$, $\delta^2(G, X) := \min\{d^2_G(v, X) \mid v \in \binom{V(G)}{2}\}$, $\delta^2(G) := \delta^2(G, V(G))$, $\Delta^2(G, X) := \max\{d^2_G(v, X) \mid v \in \binom{V(G)}{2}\}$, as well as $\Delta^2(G) := \Delta^2(G, V(G))$ to extend the notions of neighborhood and degree to 2-neighborhood and codegree. For a set $Z \subseteq \binom{V(G)}{2}$ of tuples, we additionally introduce $E^2_G(Z, X) := \bigcup_{z \in Z}\{\{z, x\} \mid x \in N^2_G(z, X)\}$, $e^2_G(Z, X) := |E^2_G(Z, X)| = \sum_{z \in Z} d^2_G(z, X)$, as well as $N^2_G(Z, X) := \bigcup_{z \in Z} N^2_G(z, X)$. We call $e \in E^2_G(Z, X)$ a 2-edge or a square-edge. This previously introduced notation is also used replacing G by an edge set $E \subseteq \binom{V}{2}$ for some vertex set V, meaning that the graph G is implicitly given by $V(G) = V$ and $E(G) = E$.

Let $k, l \geq 1$ be natural numbers, and consider a path P of length $v(P) = l$. Note that the length of a path uncommonly is defined as the number of vertices in it. The kth power P^k of P, also called k-path, is obtained by adding an edge between all pairs of vertices whose distance on P is at most k. For $k = 1$, the k-path is just P itself. A natural order of vertices of P^k is an order $V(P) = \{v_1, \ldots, v_l\}$ of the vertices such that for all $i \neq j \in [l]$ with $|i - j| \leq k$ we have $\{v_i, v_j\} \in E(P^k)$. A k-path thus can be defined by listing the vertices in a natural order. In this context, we slightly abuse notation by interpreting the vertex set as ordered. We say that P^k connects the k-tuples $a := (a_1, \ldots, a_k), b := (b_1, \ldots, b_k) \in \binom{V(P)}{k}$ if there is a natural order of $V(P)$ such that the k left-most and the k right-most vertices are the ones in a and b following the same left-to-right order, respectively. We say that a path P avoids the set X if $V(P) \cap X = \emptyset$.

For a positive integer n and a $p := p(n) \in (0, 1]$, we let $\mathcal{G}(n, p)$ denote the binomial random graph introduced by Erdös and Rényi [ER60], that is, the probability space consisting of graphs on a vertex set $[n]$ where each pair of vertices forms an edge independently with probability p. We write $H \sim \mathcal{G}(n, p)$ to indicate that H is one realization of such a random graph model. A random graph $\mathcal{G}(n, p)$ possesses a graph property $\mathcal{P}$ asymptotically almost surely, or a.a.s. for brevity, if the probability that it satisfies $\mathcal{P}$ tends to 1 as its order tends to infinity.

For $n \in \mathbb{N}$ and $p := p(n) \in (0, 1]$, we use $\text{Bin}(n, p)$ to denote the binomial distribution with n trials and success probability p. We write $X \sim \text{Bin}(n, p)$ to express that X is a random variable with such a distribution.

2.2 Basic inequalities

For binomial coefficients, we repeatedly use the following well-known upper bounds.

Claim 2.1 *For all $n \in \mathbb{N}$ and $k \in [n]$, we have $\binom{n}{k} \leq \left(\frac{ne}{k}\right)^k$ and $\binom{n}{k} \leq n^k$.*

2.3 Probabilistic tools

This section introduces inequalities used in the context of discrete probability spaces.

A very important tool is the so called union bound, providing an upper bound on the probability of a union of events.

Claim 2.2 *(Union bound)*

Let $n \in \mathbb{N}$ be an integer and $A_1, \ldots, A_n$ be events on a discrete probability space. Then $\Pr\left[\bigcup_{i\in[n]} A_i\right] \leq \sum_{i\in[n]} \Pr[A_i]$.

A similar lower bound or even the exact probability can be computed exploiting the Bonferroni inequalities [Com74] and the inclusion-exclusion principle, respectively.

Claim 2.3 *(Inclusion-exclusion principle and Bonferroni inequalities)*

Let $n \in \mathbb{N}$ be an integer and $A_1, \ldots, A_n$ be events on a discrete probability space. Then

$$\begin{aligned} \Pr\left[\bigcup_{i\in[n]} A_i\right] &= \sum_{k=1}^{n} (-1)^{k+1} \left(\sum_{1\leq i_1<\cdots<i_k\leq n} \Pr\left[\bigcap_{j\in[k]} A_{i_j}\right]\right) \\ &\geq \sum_{i\in[n]} \Pr[A_i] - \sum_{1\leq i_1<i_2\leq n} \Pr[A_{i_1} \cap A_{i_2}]. \end{aligned}$$

In the following, we provide tail estimates, that is, upper bounds on the probability of a random variable taking values far away from its mean.

The famous Chernoff's inequalities [AS04, FK15, JŁR11] provide exponentially decreasing tail estimates for binomial random variables and easily can be extended to hypergeometric distributions (see Theorem 2.10 in [JŁR11]).

Lemma 2.4 *(Chernoff's inequalities)*

Let X be binomially or hypergeometrically distributed with $\mu := \mathbb{E}[X]$. Then for any $0 < \gamma < 1$,

(i) $\Pr[X < (1-\gamma)\mu] \leq e^{-\frac{\gamma^2\mu}{2}}$, *and*

(ii) $\Pr[X > (1+\gamma)\mu] \leq e^{-\frac{\gamma^2\mu}{3}}$.

For a random variable that counts the number of events, Janson's inequality can be exploited to derive an upper bound on the lower tail. A proof can be found in [AS04] or [FK15].

Lemma 2.5 *(Janson's inequality)*

Let $m \in \mathbb{N}$ be an integer, $X_1, \ldots, X_m$ indicator random variables, and $X := \sum_{i\in[m]} X_i$. For each tuple $(i,j) \in \binom{[m]}{2}$, write $X_i \sim X_j$ if the random variables X_i and X_j are dependent. Furthermore, let

$$\mu := \mathbb{E}[X] = \sum_{i\in[m]} \Pr[X_i = 1]$$

and

$$\Delta := \sum_{i\neq j\colon X_i \sim X_j} \Pr[X_i = 1, X_j = 1].$$

Then

$$\Pr[X \leq (1-\gamma)\mu] \leq e^{-\frac{\gamma^2\mu^2}{2(\mu+\Delta)}}$$

for any $0 \leq \gamma \leq 1$.

2.4 Matchings

In this section, we collect sufficient conditions for the existence of saturating matchings, that is, independent edge sets completely covering a particular vertex set.

We first present a classic result for bipartite graphs of Hall [Hal35] (also see Theorem 3.1.11 in [Wes01]).

Theorem 2.6 *(Hall's marriage theorem)*

Let B be a bipartite graph on $X \cup Y$. Then there exists a matching entirely covering X if and only if for every subset $X' \subseteq X$ we have $|X'| \leq |N_B(X')|$.

For uniform hypergraphs, a criteria for matchability was found by Haxell (Theorem 3 in [Hax95]). In the following, we provide a slightly modified (but equivalent) version. Given a family $\mathcal{E}$ of subsets of some ground set V, we denote with $\tau(\mathcal{E})$ the size of a smallest subset $X \subseteq V$ such that $\varepsilon \cap X \neq \emptyset$ for every $\varepsilon \in \mathcal{E}$.

Theorem 2.7 *Let $r \in \mathbb{N}$ and $\mathcal{I}$ be an index set. Consider a family $\{H_i = (V, \mathcal{E}_i)\}_{i\in\mathcal{I}}$ of r-uniform hypergraphs on a vertex set V. If*

$$\tau\left(\bigcup_{i\in\mathcal{I}'} \mathcal{E}_i\right) > (2r-1)\left(|\mathcal{I}'|-1\right)$$

for every $\mathcal{I}' \subseteq \mathcal{I}$, then there exists a family of hyperedges $\{\varepsilon_i\}_{i\in\mathcal{I}}$ such that $\varepsilon_i \in \mathcal{E}_i$ and $\varepsilon_i \cap \varepsilon_j = \emptyset$ for every $i \neq j \in \mathcal{I}$.

Chapter 3

Basic tools and properties

This chapter aims to introduce basic techniques and auxiliary results more specifically tailored to our setting. In particular, we analyze crucial properties of the graphs and structures of interest and present helpful tools applicable in this context. More concretely, since we are studying the appearance of squares of cycles, Section 3.1 investigates notable variations of square-paths. Furthermore, as the considered graphs all bring certain lower bounds on the minimum degree and codegree, Section 3.2 is dedicated to all kinds of results involving minimum degree conditions. In Section 3.3, basic results about binomial random graphs are stated. Important methods primarily employable in the context of subgraphs of binomial random graphs are then presented in Section 3.4. As most of the proofs follow standard arguments, they are postponed to Section 8.1 in the appendix.

3.1 Powers of paths

We introduce two variations of square-paths, *almost square-paths* and *alternating square-paths*, that are of great importance for us, especially in the context of the Connecting Lemma.

3.1.1 Almost k-paths

A k-path without edges in-between the first k and the last k vertices is called an almost k-path.

Definition 3.1 *(Almost k-path)*

Let $k \geq 2$ and $l \geq 0$ be integers. An almost k-path is a graph Q_l^k on the vertex set $U_l = \{u_1, \ldots, u_{l+2k}\}$ and the edge set consisting of all $\{u_i, u_j\}$ such that $0 \leq i \leq k + l$ and $0 < j - i \leq k$ for $j > k$.

3.1.2 Alternating square-paths

Definition 3.2 *(Alternating square-path)*

Let $l \in \mathbb{N}$ be an integer. Let R_l denote a graph on the vertex set $U_l = \{u_1, \dots, u_{l+2}\}$ and the edge set

$$\{u_1, u_3\} \cup \bigcup_{i \in \{4, \dots, l+2\}} \{\{u_{i-1}, u_i\}, \{u_{2\lceil \frac{i}{2} \rceil - 3}, u_i\}\}.$$

For l even and $\frac{l}{2}$ odd, an alternating square-path is a graph denoted by $\widehat{R}_{l-2}$ on $l+2$ vertices $V_{l-2} = \{v_1, \dots, v_{l+2}\}$, emerging from the union of two copies of the graph $R_{\frac{l}{2}}$ with vertex sets $\{u_1, \dots, u_{\frac{l}{2}+2}\}$ and $\{u'_1, \dots, u'_{\frac{l}{2}+2}\}$, respectively, identifying the vertices $v_i = u_i$ for $i \le \frac{l}{2} - 1$, $v_{\frac{l}{2}} = u_{\frac{l}{2}} = u'_{\frac{l}{2}+2}$, $v_{\frac{l}{2}+1} = u_{\frac{l}{2}+1}$, $v_{\frac{l}{2}+2} = u_{\frac{l}{2}+2} = u'_{\frac{l}{2}}$, $v_{\frac{l}{2}+3} = u'_{\frac{l}{2}+1}$, as well as $v_i = u'_{l-i+3}$ for $\frac{l}{2} + 4 \le i \le l+2$.

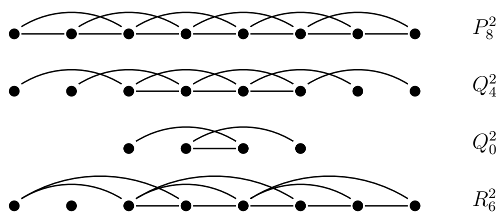

Figure 3.1: Comparison of the square-path P_8^2 to the almost square-paths Q_4^2 and Q_0^2 as well as R_6.

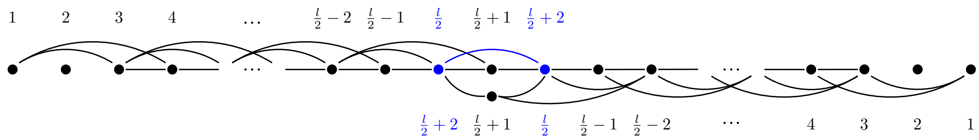

Figure 3.2: Alternating square-path $\widehat{R}_{l-2}$. The two joint vertices as well as the joint edge are depicted in blue.

3.2 Minimum degree and codegree conditions

In this section, we analyze the influence of (restricted) edge removal on the minimum degree of a graph and show how degree conditions can be transferred to induced subgraphs.

3.2.1 Subgraphs of binomial random graphs

When an adversary removes a subgraph H from a binomial random graph in the resilience setting, the degree of the resulting graph directly follows from the lower bounds on the degrees in the random graph.

Claim 3.3 *Let $n \in \mathbb{N}$ be sufficiently large, $\alpha, \beta \in (0,1)$, and H a graph on n vertices such that $\Delta(H) \leq \left(\frac{1}{3} - \alpha\right) np$ and $\Delta^2(H) \leq (1-\beta)np^2$. Then there exist constants $C, \alpha', \beta' > 0$ such that a.a.s. $\delta(G) \geq \left(\frac{2}{3} + \alpha'\right) np$ and $\delta^2(G) \geq \beta' np^2$ for $G_0 \sim \mathcal{G}(n,p)$ and $G := G_0 \setminus H$, provided that $p^2 \geq \frac{C \log n}{n}$.*

3.2.2 Graph partitioning

Given a multiplicative lower bound on the minimum degree, a graph can be partitioned into several sets such that the minimum degree condition transfers to each set.

Lemma 3.4 *Let $n \in \mathbb{N}$ be sufficiently large, $r \in [n]$, $\beta_j \in (0,1]$ for $j \in \{1,2\}$, and $p := p(n) \in (0,1]$. Consider a graph G on n vertices and a subset $U \subseteq V(G)$. Then for any $0 < \gamma < 1$ there exists a constant C such that for all $u_1, \dots, u_r \in [n]$ with $\sum_{i \in [r]} u_i \leq |U|$ and $u_i \geq \frac{C \log n}{p^2}$ for $i \in [r]$ there exist disjoint subsets $U_1, \dots, U_r \subseteq U$ with $|U_i| = u_i$ for $i \in [r]$ such that for all $i \in [r]$ a.a.s.*

(i) $d_G(v, U_i) \geq (1-\gamma)\beta_1 u_i p$ for all $v \in V(G)$ with $d_G(v,U) \geq \beta_1 |U| p$, and

(ii) $d_G^2(v, U_i) \geq (1-\gamma)\beta_2 u_i p^2$ for all $v \in \binom{V(G)}{2}$ with $d_G^2(v,U) \geq \beta_2 |U| p^2$.

3.3 Properties of binomial random graphs

In this section, some well-known properties of binomial random graphs are stated.

The following claim shows that in binomial graphs the degrees and codegrees into large sets are concentrated around their expected value. As it follows by a simple application of Lemmas 2.4 and 2.2, its proof is omitted.

Lemma 3.5 *(Degree Concentration)*

Let $\tau \in (0,1]$ and $n \in \mathbb{N}$ be sufficiently large. Consider a random graph $G \sim \mathcal{G}(n,p)$ and a vertex set $S \subseteq V(G)$ of size $|S| = \tau n$. Then for any $0 < \gamma < 1$ there exists a constant C such that a.a.s. $d_G^j(v,S) = (1 \pm \gamma)\, \tau n p^j$ for all $j \in \{1,2\}$ and $v \in \binom{V(G) \setminus S}{j}$, provided that $p^2 \geq \frac{C \log n}{n}$.

Random graphs have the property that they are not too dense, that is, between any two (large enough) sets there cannot be more edges than expected. This does not only hold for simple edges but also for square-edges.

Lemma 3.6 *(Edge Concentration)*

Let $n \in \mathbb{N}$ be sufficiently large and $p := p(n) \in (0,1]$. Then for any $0 < \gamma < 1$ there exists a constant C such that the random graph $G \sim \mathcal{G}(n,p)$ a.a.s. has $e^j_G(X^j, Y) = (1 \pm \gamma)\,|X^j||Y|p^j$ for all $j \in \{1,2\}$, $Y \subseteq V(G)$, and sets $X^j \subseteq \binom{V(G)\setminus Y}{j}$ of disjoint j-tuples, provided that $|X^j|, |Y| \geq \frac{C \log n}{p^2}$.

In the following, we provide three claims analyzing the number of vertices in a designated set that are reachable from a set of edges (or pairs of vertices) by square-edges under different assumptions.

Claim 3.7 *Let $n \in \mathbb{N}$ be sufficiently large, $\delta \in \left(0, \frac{1}{2}\right)$, and $p := p(n) \in (0,1]$. Consider $G \sim \mathcal{G}(n,p)$, a set $W \subseteq V(G)$, and a vertex $v \in V(G) \setminus W$. Then for all sets $U \subseteq V(G) \setminus (W \cup \{v\})$ of size $|U| \leq \frac{\delta}{p}$ and any $\frac{\delta}{1-\delta} < \gamma < 1$ there exists a constant C such that a.a.s. $|N^2_G((U,v),W)| = (1 \pm \gamma)|U||W|p^2$, provided that $|W| \geq \frac{C \log n}{p^2}$.*

Claim 3.8 *Let $n \in \mathbb{N}$ be sufficiently large, $p := p(n) \in (0,1]$, and $G \sim \mathcal{G}(n,p)$. Then for any $0 < \delta < 1$ there exists a constant C such that for all $W \subseteq V(G)$ with $|W| \geq \frac{C \log n}{p^2}$ and all sets $T \subseteq \binom{V(G)\setminus W}{2}$ of tuples with $|T| \leq \frac{\delta}{p^2}$ and maximum degree $\Delta(T) \leq \frac{\delta}{p}$ a.a.s. $|N^2_G(T,W)| = (1 \pm 3\delta)\,|T||W|p^2$.*

Claim 3.9 *Let $n \in \mathbb{N}$ be sufficiently large, $\tau \in (0,1]$, and $p := p(n) \in (0,1]$ such that $p^2 = \omega\left(\frac{1}{n}\right)$. Consider $G \sim \mathcal{G}(n,p)$, disjoint subsets $U, V, W \subseteq V(G)$ with $|W| = \tau n$, as well as a set $T \subseteq U \times V$. For $W' \subseteq W$, let $X^{W'}$ be the number of triangles in G induced by T and W', that is,*

$$X^{W'} := \left|\left\{(u,v,w) \mid u \in U, v \in V, w \in W' \colon \{u,v\} \in T, \{u,w\}, \{v,w\} \in E(G)\right\}\right|.$$

Then for any $0 < \gamma < 1$, all $W' \subseteq W$ of size $|W'| = \Theta(n)$, and sets T with $|T| = \omega\left(\frac{n}{\log^2 n}\right)$ and $\Delta(T) = \mathcal{O}\left(\frac{|W'|p}{\log^2 n}\right)$, a.a.s. $X^{W'} \geq (1-\gamma)|T||W'|p^2$.

3.4 Subgraphs of binomial random graphs

3.4.1 Density arguments

Binomial random graphs do not contain dense subgraphs, that is, they cannot have significantly more edges than expected between any two (not too small) sets. This property, which directly transfers to subgraphs of random graphs, together with the minimum degree assumption, allows us to conclude good expansion of such graphs. The rough idea is as follows. Consider a set X of vertices with large degree into another set Y. Then the neighborhood of X in Y must not be too small, as otherwise the subgraph between X and its neighborhood in Y would be too dense.

Claim 3.10 *(Density Lemma)*

Let $n \in \mathbb{N}$ be sufficiently large, $\alpha \in (0,1)$, $p := p(n) \in (0,1]$, and G a subgraph of $\mathcal{G}(n,p)$. Then for any $0 < \gamma < 1$ there exists a constant C such that for all disjoint sets $X, Y \subseteq V(G)$ with $d_G(x,Y) \geq \alpha|Y|p$ for all $x \in X$, we a.a.s. have $|N_G(X,Y)| \geq \frac{\alpha}{1+\gamma}|Y|$, provided that $|X|, |Y| \geq \frac{C \log n}{p}$.

Very similarly, one also can prove the following related result. It says that there can be only few vertices having degree much larger than expected.

Claim 3.11 *Let $n \in \mathbb{N}$ be sufficiently large, $\beta > 1$, and $p := p(n) \in (0,1]$. Consider a subgraph $G \subseteq \mathcal{G}(n,p)$. Then there exists a constant C such that for all sets $X \subseteq V(G)$ and $Y \subseteq V(G) \setminus X$ with $|X|, |Y| \geq \frac{C \log n}{p}$ a.a.s. the set $X_\beta := \{x \in X \colon d_G(x,Y) \geq \beta|Y|p\}$ must have size $|X_\beta| < \frac{C \log n}{p}$.*

Analogously, also a bound which shows that a set having degree above or only slightly below the expected value must be large.

Claim 3.12 *Let $n \in \mathbb{N}$ be sufficiently large, $p := p(n) \in (0,1]$, and G a subgraph of $\mathcal{G}(n,p)$. Moreover, let $\alpha > 0$ and $0 < \beta < \alpha$ be constants. Then for any small enough $\gamma > 0$ there exist constants C and $\kappa := \frac{(1-\gamma)\alpha - \beta}{1+\gamma-\beta} > 0$ such that for all sets $X, Y \subseteq V(G)$ with $|X|, |Y| \geq \frac{C \log n}{p}$ and $e_G(X,Y) \geq \alpha|X||Y|p$, the set $Y_\beta := \{y \in Y \mid d_G(y,X) \geq \beta|X|p\}$ a.a.s. must have size $|Y_\beta| \geq \kappa|Y|$.*

The following claim shows that if the average and maximum degree are close, then a large fraction of the vertices must have almost maximum degree, thus establishes a relation between the maximum and the average degree. Note that this result holds for arbitrary graphs. Nevertheless, we state it in this section, as we only apply it in the context of subgraphs of random graphs.

Claim 3.13 *Let $n \in \mathbb{N}$ be sufficiently large, $\alpha \in (0,1]$, and G a graph on n vertices. Consider subsets $X, Y \subseteq V(G)$ such that $e_G(X,Y) \geq \alpha \max\{d_G(x,Y) \colon x \in X\}$. Then, for any $0 < \beta < \alpha$ there exists a constant $\kappa > 0$ such that the set $X_\beta := \{x \in X \mid d_G(x,Y) \geq \beta \max\{d_G(x,Y) \colon x \in X\}\}$ has size $|X_\beta| \geq \frac{\alpha - \beta}{1-\beta}|X|$.*

3.4.2 Appearance of short square-paths

In the following, we show that several short (that is, constant-length) 2-paths can be simultaneously embedded into subgraphs of binomial random graphs, applying the Expansion Lemma and Hall's marriage theorem.

Claim 3.14 *Let $n \in \mathbb{N}$ be sufficiently large, $l \geq 2$ an integer, $\beta_j \in (0,1)$ for $j \in \{1,2\}$, and $p := p(n) \in (0,1)$. Consider a subgraph G of $\mathcal{G}(n,p)$. Then there exists a constant C such that for all sets $X = \{x_1, \ldots, x_{|X|}\} \subseteq V(G)$ and $Y \subseteq V(G) \setminus X$ with $|Y| > \max\left\{\frac{(l-1)}{\min\{\beta_1,\beta_2\}}|X|, \frac{C \log n}{p^2}\right\}$, $\delta(G,Y) \geq \beta_1|Y|p$, and $\delta^2(G,Y) \geq \beta_2|Y|p^2$, there a.a.s. exist $|X|$ many vertex-disjoint square-paths $\{L_r\}_{r\in[|X|]}$ of length l in G such that L_r has x_r as endpoint and $V(L_r) \setminus \{x_r\} \subseteq Y$ for all $r \in [|X|]$.*

Proof Use Lemma 3.4 to subdivide Y into $l-1$ many sets $Y_2, \ldots, Y_l$ of size $\frac{|Y|}{l-1}$ each such that $d^j_G(v, Y_i) \geq (1-\gamma)\beta_j|Y_i|p^j$ for all $j \in \{1,2\}$, $i \in \{2, \ldots, l\}$, $v \in \binom{V(G)}{j}$ and any $0 < \gamma < 1$. To simplify notation, let $Y_1 := X$ and use $Y_2 := Y$ in the case that $l = 2$.

We show by induction that for every $i \in [l]$ there exist t many vertex-disjoint square-paths $L^i_1, \ldots, L^i_t$ of length i in G such that $\left|V(L^i_r) \cap Y_{i'}\right| = 1$ for all $r \in [t]$ and $i' \in [i]$. For $i = l$, this gives a family of vertex-disjoint square-paths $L_1, \ldots, L_t$ of length l, as desired.

Let L^1_r be the path on vertex x_r (that is, a 2-path of length 1) and introduce $y^1_r := x_r$ for all $r \in [t]$. Let $X' \subseteq X$ be an arbitrary subset of vertices in X. Lemma 3.10 yields $|N_G(X', Y_2)| > \frac{\beta_1}{1+\gamma}|Y_2| > |X'|$, thus Lemma 2.6 implies the existence of a matching saturating X. Let $y^2_r \in Y_2$ be the vertex matched to x_r.

Now, suppose that the induction hypothesis is true for some $i \in \{2, \ldots, l-1\}$. To construct paths for $i+1$, we consider the auxiliary bipartite graph B_{i+1} with vertex classes being square-paths $\{L^i_r\}_{r\in[t]}$ (that is, each path L^i_r represents a single vertex in B_{i+1}) and Y_{i+1}. The edge set

$$E(B_{i+1}) = \{\{L^i_r, y\} \mid \forall i' \in [i] \colon \{V(L^i_r) \cap Y_{i'}, y\} \in E(G)\}$$

consists of edges in $E(G)$ that continue a 2-path L^i_r of length i to a 2-path of length $i+1$. Note that, by definition of these bipartite graphs,

$$d_{B_{i+1}}(L^i_r) = d^2_G\left(\left(y^{i-1}_r, y^i_r\right), Y_{i+1}\right) \geq \beta_2|Y_{i+1}|p^2.$$

Let $\mathcal{I} \subseteq [t]$ be an arbitrary index set and define $N^{\mathcal{I}}_{i+1}$ to be the set of vertices $y \in Y_{i+1}$ reached by edges in B_{i+1} originating from such vertices in $\{L^i_r\}_{r\in\mathcal{I}}$, that is, $N^{\mathcal{I}}_{i+1} := \left\{y \in Y_{i+1} \mid \exists r \in \mathcal{I} \colon \{L^i_r, y\} \in E(B_{i+1})\right\}$. Lemma 3.10 then implies that $|N^{\mathcal{I}}_{i+1}| \geq \frac{\beta_2}{1+\gamma}|Y_{i+1}| > |\mathcal{I}|$. Lemma 2.6 thus shows the existence of a matching in B_{i+1} saturating $\{L^i_r\}_{r\in[t]}$. Writing $y^{i+1}_r \in Y_{i+1}$ for the vertex matched to the path L^i_r, the extension of L^i_r to y^{i+1}_r for each $r \in [t]$ gives the desired family of 2-paths for $i+1$. This concludes the proof of the induction step, and hence of the lemma. □

Chapter 4

Absorber Method

In the scope of k-Hamiltonicity, an absorber for a reservoir set $X \subseteq V(G)$ is a k-path P in G with $V(P) \subseteq V(G) \setminus X$ that can absorb any subset $X' \subseteq X$ into it without changing the endpoints and without deleting any vertices on the path, resulting in a k-path $P^*_{X'}$ with $V(P^*_{X'}) = V(P) \cup X'$.

In the following, we first introduce a special absorbing structure in Section 4.1 and then show that it can indeed be found in the considered graph in Section 4.2. We thereby closely follow [NŠ16a], adopting their notation and some of their proofs.

4.1 Absorbing structure for powers of spanning cycles

The following definition specifies the meaning of a k-path 'absorbing' a set of vertices into it, and is directly borrowed from Definition 4.1 in [NŠ16a].

Definition 4.1 *(Absorber, absorbing structure, k-absorbing path)*

Let $k \geq 2$ be an integer, G a graph, $X \subseteq V(G)$ a subset of vertices in G, and $P_X \subseteq G$ a k-path in G avoiding set X. We say that P_X is a k-absorbing path (or, equivalently, absorber or absorbing structure) for X if for any subset $X' \subseteq X$ there exists a k-path $P^*_{X'} \subseteq G$ such that

(i) $V(P^*_{X'}) = V(P_X) \cup X'$, and

(ii) $P^*_{X'}$ and P_X connect the same pair of k-tuples of vertices.

An absorbing k-path for X can be found by concatenating absorbing k-paths $P_{\{x\}}$ for all $x \in X$. That is why in the following only absorbing structures for a single vertex $x \in X$ are considered. We introduce a graph $A^k_l(x)$ and then show that indeed it is an absorber for x. The parameter l in $A^k_l(x)$ roughly corresponds to its size, since $v\left(A^k_l(x)\right) = 1 + 2kl + l(l-1) \approx l^2$.

Definition 4.2 *Let $k \geq 2$ and $l > k$ be integers. The graph $A_l^k(x)$ has vertex set*

$$V(A_l^k(x)) = \{x\} \cup \bigcup_{i \in [l], j \in [2k]} \{w_{i,j}\} \cup \bigcup_{i \in [l-1], j \in [l]} \{u_{i,j}\}$$

and an edge set consisting of the union of edge-disjoint graphs $R, J_0, \ldots, J_{l-1}$, where R is isomorphic to a k-path with natural order

$$\begin{aligned} V(R) := \{&w_{1,1}, \ldots, w_{1,k}, x, w_{1,k+1}, \ldots, w_{1,2k}, u_{1,1}, \ldots, u_{1,l}, \\ &w_{2,1}, \ldots, w_{2,2k}, u_{2,1}, \ldots, u_{2,l}, \\ &\ldots \\ &w_{l-1,1}, \ldots, w_{l-1,2k}, u_{l-1,1}, \ldots, u_{l-1,l}, \\ &w_{l,1}, \ldots, w_{l,2k}\}, \end{aligned}$$

and J_i for $i \in \{0, \ldots, l-1\}$ are isomorphic to an almost k-path Q_0^k with natural order

$$\begin{aligned} V(J_0) &:= \{w_{1,1}, \ldots, w_{1,k}, w_{2,k}, \ldots, w_{2,1}\}, \\ V(J_i) &:= \{w_{i,2k}, \ldots, w_{i,k+1}, w_{i+2,k}, \ldots, w_{i+2,1}\}, \quad \text{for } i \in [l-2], \text{ and} \\ V(J_{l-1}) &:= \{w_{l-1,2k}, \ldots, w_{l-1,k+1}, w_{l,k+1}, \ldots, w_{l,2k}\}. \end{aligned}$$

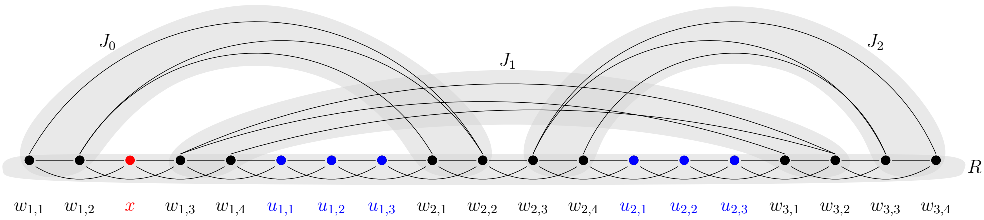

Figure 4.1: The absorber $A_l^k(x)$ for $k = 2$ and $l = 3$ with its constituents R, J_0, J_1, and J_2.

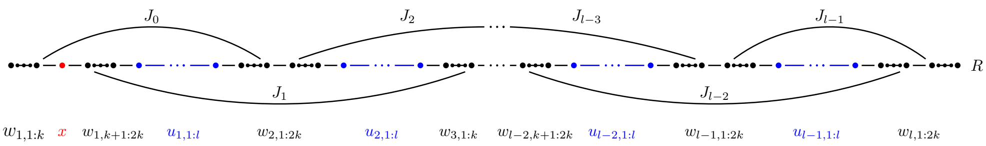

Figure 4.2: The abstract skeleton (that is, a graph where powers of paths are replaced by the underlying simple paths) of an absorber $A_l^k(x)$ with its constituents $R, J_0, \ldots, J_{l-1}$. The dashed lines represent powers of paths of indicated length. For any $i \in [l]$, we use $u_{i,1:l}$ to denote the sequence $u_{i,1}, \ldots, u_{i,l}$ and analogously define $w_{i,1:2k}$, $w_{i,1:k}$ as well as $w_{i,k+1:2k}$.

In the following, the important role given to $x \in V(A_l^k(x))$ is justified by showing that indeed $A_l^k(x)$ is an absorber for the set $\{x\}$. As the proof is identical to the one of Claim 4.3 in [NŠ16a], it is omitted.

Claim 4.3 *The graph $A_l^k(x)$ contains k-paths P_x and P_x^* such that*

(i) both P_x and P_x^* connect $(w_{1,1}, \dots, w_{1,k})$ to $(w_{l,k+1}, \dots, w_{l,2k})$, and

(ii) $V(P_x) = V(A_l^k(x)) \setminus \{x\}$ and $V(P_x^*) = V(A_l^k(x))$.

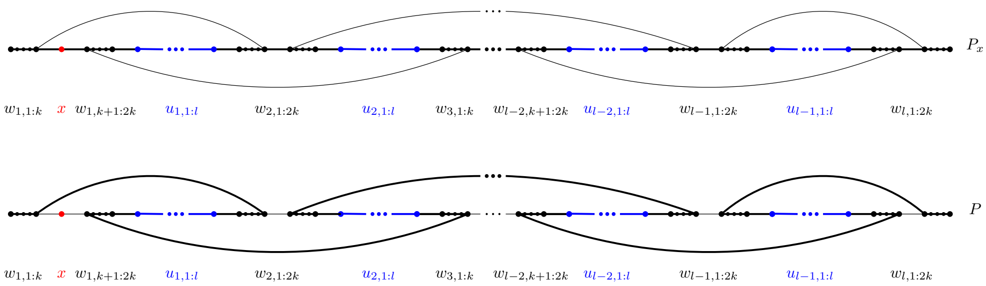

Figure 4.3: Abstract depiction of the two paths P_x and P_x^* (drawn with bold lines each) in the skeleton of the absorber $A_l^k(x)$. Light lines stand for edges not used in the corresponding path. The dashed lines represent powers of paths of indicated length.

4.2 Absorbing Lemma

Lemma 4.4 *(Absorbing Lemma)*

Let $n \in \mathbb{N}$ be sufficiently large, $\alpha, \beta, \varepsilon \in (0,1)$, and $G \subseteq \mathcal{G}(n,p)$ for $p^2 \geq n^{-1+\varepsilon}$ such that $\delta(G) \geq \left(\frac{2}{3} + \alpha\right) np$ and $\delta^2(G) \geq \beta n p^2$. Then there exist a constant $c_1 \in (0,1)$ and a set $V_1 \subseteq V(G)$ of size $n_1 := |V_1| = c_1 n$ chosen uniformly at random such that the graph G a.a.s. contains an absorbing square-path P_{V_1} for V_1.

Proof By Lemma 3.4, there exists a partition $V(G) = V_0^1 \cup V_0^2 \cup \bigcup_{i=1}^4 V_i$ of the vertices in G such that $|V_1| = c_1 n$, $|V_0^1| = |V_0^2| \geq \max\{10, \frac{3}{\beta}\} c_1 n$, $|V_2|, |V_3|, |V_4| \geq \sqrt{c_1} n$, as well as $\delta(G, X) \geq \left(\frac{2}{3} + \alpha'\right) |X| p$ and $\delta^2(G, X) \geq \beta' |X| p^2$ for all $X \in V_0^1 \cup V_0^2 \cup V_2 \cup V_3 \cup V_4$, $\alpha' < \alpha$, and $\beta' < \beta$. Let l and c_o be the constants from Lemma 5.1 (for $\overline{\alpha} \mapsto \alpha'$ and $\overline{\beta} \mapsto \beta'$), pick c_1 sufficiently small such that $\sqrt{c_1} \leq \frac{c_o}{100 l^3}$, say, and note that therefore $|V_i| \geq \sqrt{c_1} n \geq \frac{1}{\sqrt{c_1}} |V_1| \geq \frac{100 l^3}{c_o} |V_1|$ for all $i \in \{2, 3, 4\}$.

We first find disjoint absorbing paths P_x with $V(P_x) \subseteq V_0^1 \cup V_0^2 \cup V_2 \cup V_3$ for all $x \in V_1$ and then connect them, using vertices from V_4, to one absorbing path P_{V_1} for the set V_1. Let us arbitrarily enumerate the vertices in V_1 as $V_1 = \{x_1, \dots, x_{n_1}\}$.

Applying Claim 3.14 twice, with $X \mapsto V_1$, $l \mapsto 3$, as well as $Y \mapsto V_0^1$ and $Y \mapsto V_0^2$, respectively, we can find $2n_1$ disjoint 2-paths $L_1^1, \dots, L_{n_1}^1, L_1^2, \dots, L_{n_1}^2$ of length 3, extending x_r to the sets V_0^1 and V_0^2, respectively, for all $r \in [n_1]$. Give names $z_r^{1,1}, z_r^{1,2} \in V_0^1$ and $z_r^{2,1}, z_r^{2,2} \in V_0^2$ to the vertices in V_0^1 and V_0^2, respectively, connected to x_r.

We use Lemma 5.1 with $S \mapsto V_2$ as connecting set to connect the family $\left\{\left((z_r^{1,1}, z_r^{1,2}), (z_r^{2,1}, z_r^{2,2})\right)\right\}_{r\in[n_1]}$ by alternating square-paths $\widehat{R}_{4(l-1)}$ on $4l$ vertices $\{v_r^1, \dots, v_r^{4l}\}$, where $v_r^1 = z_r^{1,1}$, $v_r^2 = z_r^{1,2}$, $v_r^{4l-1} = z_r^{2,1}$, and $v_r^{4l} = z_r^{2,2}$, for every $r \in [n_1]$.We call the resulting structure together with the 2-paths L_r^1 and L_r^2 (thus also adding the vertex x_r) H_r. Give the vertices $w_{i,j}$ and $u_{i,j}$ from the absorber $A_l^2(x_r)$ the names $w_{i,j}^r$ and $u_{i,j}^r$, respectively. By rearranging the vertices in such a way that

$$v_r^{4i-1} \mapsto w_{2i,2}^r, v_r^{4i} \mapsto w_{2i,1}^r, v_r^{4i+1} \mapsto w_{2i,3}^r, v_r^{4i+2} \mapsto w_{2i,4}^r$$

for $i \in \{1, \dots, \lfloor \frac{l}{2} \rfloor\}$,

$$v_r^{4i-1} \mapsto w_{2l+1-2i,2}^r, v_r^{4i} \mapsto w_{2l+1-2i,1}^r, v_r^{4i+1} \mapsto w_{2l+1-2i,3}^r, v_r^{4i+2} \mapsto w_{2l+1-2i,4}^r$$

for $i \in \{\lfloor \frac{l}{2} \rfloor + 1, \dots, l-1\}$, and

$$v_r^1 \mapsto w_{1,2}^r, v_r^2 \mapsto w_{1,1}^r, v_r^{4l-1} \mapsto w_{1,4}^r, v_r^{4l} \mapsto w_{1,3}^r,$$

it is easy to check that H_r is isomorphic to the vertex-induced subgraph of $A_l^2(x_r)$ restricted to the vertex set $\{w_{i,j}^r \mid i \in [l], j \in [4]\} \cup \{x_r\}$.

In the next step, we aim to add the $l-1$ missing almost square-paths with natural ordering $\{w_{i,3}^r, w_{i,4}^r, u_{i,1}^r, \dots, u_{i,l}^r, w_{i+1,1}^r, w_{i+1,2}^r\}$ (including the vertices $\bigcup_{i\in[l-1]} u_{i,l}^r$) for $i \in [l-1]$ to H_r, resulting in the absorbing structure $A_l^2(x_r)$. This is achieved by connecting

$$\bigcup_{r\in n_1, i\in[l-1]} \left\{\left((w_{i,3}^r, w_{i,4}^r), (w_{i+1,1}^r, w_{i+1,2}^r)\right)\right\}$$

by Lemma 5.1 with almost square-paths of length l and connecting set $S \mapsto V_3$, yielding disjoint absorbers $A_l^2(x_r)$ for all $r \in [n_1]$. Let $s^r := (w_{1,1}^r, w_{1,2}^r)$ and $e^r := (w_{l,3}^r, w_{l,4}^r)$ denote the first and the last two vertices of $A_l^2(x_r)$, respectively.

These absorbers then can be connected to an absorber P_{V_1} for the set V_1 by applying Lemma 5.1 connecting the pairs $\bigcup_{r\in[n_1-1]} \left\{(e^r, s^{r+1})\right\}$ with almost square-paths of length l using the connecting set $S \mapsto V_4$.

It is now easy to see that P_{V_1} has the required properties. For each $X \subseteq V_0$ we show the existence of a 2-path P_X^* which connects s^1 to e^{n_1}. For each $x_r \in V_1$, define F_r as $F_r := \begin{cases} P_{x_r} & \text{if } x_r \notin X \text{, and} \\ P_{x_r}^* & \text{if } x_r \in X. \end{cases}$

Let K_r be the path used by Connecting Lemma before to connect e^r to s^{r+1}. Note that for any $r \in [n_1-1]$, the graph $F_r \cup K_r \cup F_{r+1}$ is a 2-path connecting s^r to e^{r+1}. Therefore, the graph $P_X^* := F_1 \cup K_1 \cup F_2 \cup \ldots \cup K_{n_1-1} \cup F_{n_1}$ is a 2-path connecting s^1 to e^{n_1}. Thus, P_X^* satisfies the required properties, which finishes the proof of the lemma. □

Chapter 5

Connecting Lemma

The Connecting Lemma enables the disjoint connection of many pairs of tuples by either almost or alternating square-paths, using vertices from a dedicated set. This is particularly helpful to join several square-paths to longer square-paths or to square-cycles, which is achieved by connecting the corresponding endpoints.

Lemma 5.1 *(Connecting Lemma)*

Let $n \in \mathbb{N}$ be sufficiently large and $\overline{\alpha}, \overline{\beta}, \varepsilon \in (0,1)$. Consider a random graph $G_0 \sim \mathcal{G}(n,p)$ for $p^2 \geq n^{-1+\varepsilon}$ and a subgraph $H \subseteq G_0$ with maximum degree $\Delta(H) \leq \left(\frac{1}{3} - \overline{\alpha}\right) np$ and maximum codegree $\Delta^2(H) \leq (1-\overline{\beta})np^2$. Then there exist $l \in \mathbb{N}$ (a multiple of 4) and $c_o \in (0,1)$ such that for $L \in \{l, 4(l-1)\}$, $t \leq \frac{c_o}{100 l^2} n$, a family $\{T_i := (a_i, b_i)\}_{i \in [t]} \subseteq \binom{V(G)}{2}^2$ of pairs of disjoint tuples (where a_i and b_i might overlap or even coincide), and a subset $S \subseteq [n] \setminus \bigcup_{i \in [t]} T_i$ of size $|S| \geq \frac{100 l^2}{c_o} t$ such that $\delta(G,S) \geq \left(\frac{2}{3} + \overline{\alpha}\right) |S| p$ and $\delta^2(G,S) \geq \overline{\beta} |S| p^2$, there a.a.s. exist t internally disjoint almost or alternating square-paths $P_1, \ldots, P_t \subseteq G$ of length L such that P_i connects a_i to b_i and satisfies $V(P_i) \setminus T_i \subseteq S$ for all $i \in [t]$.

An idea of general applicability by Nenadov and Škorić [NŠ16a], based on the hypergraph matching criteria of Haxell, reduces the problem of finding many disjoint connecting structures to finding one connecting structure in a slightly smaller set. We perform this reduction and introduce further notation in Section 5.1. Section 5.2 then serves the purpose of stating the main result, that is, Claim 5.2, and providing the proof, relying on the aforementioned claim, of the Connecting Lemma. In Section 5.3, a simple proof of Claim 5.2 (i) is given. The auxiliary results presented in Section 5.4 are then used to prove Claim 5.2 (ii) and (iii) in Sections 5.5 and 5.6, respectively.

5.1 Setup and notation

We assume without loss of generality that $L = l$, as a connecting path of length l implies the existence of a connecting path of length $4(l-1)$ recalling that, by the assumption on l, both numbers are multiples of 4. For each $i \in [t]$, define the l-uniform hypergraph $H_i = (S, \mathcal{E}_i)$ with the family $\mathcal{E}_i$ consisting of all subsets $\varepsilon_i \subseteq S$ of size l such that there exists a structure (that is, an almost square-path or an alternating square-path) of length l connecting a_i to b_i using the vertices from ε_i. To prove the lemma, it suffices to find pairwise disjoint subsets $\varepsilon_i \in \mathcal{E}_i$ for $i \in [t]$, which can be done by verifying the conditions of Theorem 2.7. Towards a contradiction, assume that there exists a $\mathcal{I}^* \subseteq [t]$ such that $\tau\left(\bigcup_{i\in\mathcal{I}^*} \mathcal{E}_i\right) \leq 2l|\mathcal{I}^*|$. By definition, there exists an obstruction set $O \subseteq S$ of size $o := |O| = 2l|\mathcal{I}^*|$, that is, a set O such that $\varepsilon_i \cap O \neq \emptyset$ for every $i \in \mathcal{I}^*$ and $\varepsilon_i \in \mathcal{E}_i$. We derive a contradiction by showing the existence of an index $i \in \mathcal{I}^*$ for which there exists a structure connecting a_i to b_i only using vertices from $S \setminus O$.

By Lemma 3.4, we can partition S into sets $S_1, \ldots, S_l$ of size $s := \frac{|S|}{l}$ each such that a.a.s. for all $i \in [l]$ we have $\delta(G, S_i) \geq \left(\frac{2}{3} + \overline{\alpha}'\right) sp$ for any $0 < \overline{\alpha}' < \overline{\alpha}$ and $\delta^2(G, S_i) \geq \overline{\beta}' sp^2$ for any $0 < \overline{\beta}' < \overline{\beta}$. To simplify notation, we use $\alpha := \overline{\alpha}'$ and $\beta := \overline{\beta}'$ and let $S_i' := S_i \setminus O$ for all $i \in [l]$.

In the following, we introduce sequences of bipartite graphs and corresponding quantities of interest with respect to an arbitrary set $\mathcal{I} \subseteq \mathcal{I}^*$. For the sake of brevity, we only add the superscript $\mathcal{I}$ to the quantities when it is not clear from the context with respect to which set $\mathcal{I}$ the graph sequence is considered. Intuitively speaking, we inductively keep track of the edges that can be reached by square-paths in $S_1, S_2, \ldots, S_l$ when starting from the tuples $\{a_i\}_{i\in\mathcal{I}}$ only using vertices from $S \setminus O$, and the sequence of edge sets that can be reached in $S_l, S_{l-1}, \ldots, S_1$ when starting from the tuples $\{b_i\}_{i\in\mathcal{I}}$ only using vertices from $S \setminus O$. As these notions are symmetric, we only introduce the definitions and prove the results for the first sequence.

Let $R := \bigcup_{i\in\mathcal{I}}\{a_i\}$ be the set of all tuples induced by $\mathcal{I}$, give the vertices in a_i the names a_i^1 and a_i^2, and let $R_1 := \bigcup_{i\in\mathcal{I}}\{a_i^1\}$ as well as $R_2 := \bigcup_{i\in\mathcal{I}}\{a_i^2\}$.

Set $S_{-1} := R_1$, $N_{-1} := R_1$, $S_0 := R_2$, and $N_0 := R_2$. Thus $n_0 := |N_0| = |\mathcal{I}|$ by disjointness of the tuples. Define B_0 on $V(B_0) := N_{-1} \cup N_0$ with $E(B_0) := \{\{u,v\} \mid u \in S_{-1}, v \in S_0, \{u,v\} \in E(G)\}$. Let $f(1) := 0$ and

$$f(i) := \begin{cases} i-1, & \text{if } i \text{ even or almost square-path,} \\ i-2, & \text{if } i \text{ odd and alternating square-path,} \end{cases}$$

for $2 \leq i \leq l$ as well as

$$g(i) := \begin{cases} i-2, & \text{if almost square-path,} \\ i-1, & \text{if } i \text{ odd and alternating square-path,} \\ i-3, & \text{if } i \text{ even and alternating square-path,} \end{cases}$$

for $i \in [l]$. We inductively define auxiliary bipartite graphs $B_i = B_{i,f(i)} = B_{f(i),i}$ on the vertex sets $S_{f(i)}$ and S_i with

$$E(B_i) := \Big\{\{v,w\} \mid \exists u \in S_{g(i)} \colon \{u,v\} \in E(B_{g(i),f(i)}), \\ \{u,w\} \in E(G), \{v,w\} \in E(G)\Big\}$$

for all $i \in [l]$. Let $\widehat{N}^+(v) := N_{B_i}(v)$, $N^+(v) := N_{B_i}(v, S_i')$, $\overline{N}^+(v) := N_{B_i}(v, O)$, $\widehat{d}^+(v) := |\widehat{N}^+(v)|$, $d^+(v) := |N^+(v)|$, and $\overline{d}^+(v) := |\overline{N}^+(v)|$ for $v \in S_{f(i)}$ as well as $\widehat{N}^-(w) := N_{B_i}(w)$, $N^-(w) := N_{B_i}(w, S'_{f(i)})$, $\overline{N}^-(w) := N_{B_i}(w, O)$, $\widehat{d}^-(w) := |N_{B_i}(w)|$, $d^-(w) := |N^-(w)|$, and $\overline{d}^-(w) := |\overline{N}^-(w)|$ for $w \in S_i$. Moreover, let $\widehat{e}_i(X) := e_{B_i}(X)$, $e_i(X) = e_{B_i}(X, S_i')$, and $\overline{e}_i(X) := e_{B_i}(X, O)$ for all $X \subseteq N_{f(i)}$ as well as $\widehat{e}_i(Y) := e_{B_i}(Y)$, $e_i(Y) := e_{B_i}(Y, S'_{f(i)})$, and $\overline{e}_i(Y) := e_{B_i}(Y, O)$ for $Y \subseteq N_i$. Furthermore, let $N_i := \bigcup_{v \in N_{f(i)}} N^+_{f(i)}(v)$ be the set of vertices reached in S_i' and $n_i := |N_i|$ its size. We call $d^+(v), \widehat{d}^+(v)$, and $\overline{d}^+(v)$ the out-degrees of vertex $v \in S_{f(i)}$ and $d^-(w), \widehat{d}^-(w)$, and $\overline{d}^-(w)$ the in-degrees of a vertex $w \in S_i$. In certain cases, we add the subscript $(i, f(i))$ or $(f(i), i)$ to the quantities to emphasize that the underlying bipartite graph has vertex sets S_i and $S_{f(i)}$. We call $\{(g(k), f(k), k)\}_{3 \leq k \leq l}$ a sequence of steps.

Let $c_v, c_e, c_{\mathcal{L}}, c_e' \in (0,1)$ be such that $c_v \geq \frac{10}{\beta} c_e, \frac{5}{\beta} c_{\mathcal{L}}$, and $c_{\mathcal{L}} \geq \frac{2}{\beta} c_e$. We will choose c_e' later to be sufficiently smaller than c_e and c_o to be arbitrarily smaller than all the other constants. Note that by choosing c_o arbitrarily small, we also have that $\frac{o}{s}$ is as small as we want, since $o \leq 2lt \leq \frac{c_o}{50l}|S| = \frac{c_o}{50}s$.

5.2 Proof of the Connecting Lemma

To prove the Connecting Lemma, we repeatedly make use of the following main result that provides upper bounds on the number of steps needed to reach a prescribed number of vertices or edges.

Claim 5.2 *For all $\mathcal{I} \subseteq \mathcal{I}^*$ there exist constants $l_1, l_2, l_3 \in \mathbb{N}$ such that a.a.s.*

(i) $n^{\mathcal{I}}_{l_1+j} \geq c_v s$ and $e^{\mathcal{I}}_{l_1+j} = \Omega(n^{1+\varepsilon})$ for all $0 \leq j \leq l - l_1$ if $|\mathcal{I}| \geq \lfloor \frac{|\mathcal{I}^*|}{2} \rfloor - 1$;

(ii) $e^{\mathcal{I}}_{i+l_2} \geq c_e s^2 p$ if $e^{\mathcal{I}}_i \geq n^{1+\eta}$ for some $0 \leq i \leq l - l_2$ and $\eta > \frac{\varepsilon}{8}$; and

(iii) $e^{\mathcal{I}}_{i+l_3+j} \geq \frac{2}{3} s^2 p$ for all $0 \leq j \leq l - i - l_3$ if $e^{\mathcal{I}}_i \geq c_e s^2 p$ for some $0 \leq i \leq l - l_3$.

By iteratively exploiting Claim 5.2 in the next claim, we show that there exists one tuple in R from which at least $\frac{2}{3}s^2p$ many edges can be reached.

Claim 5.3 *Let $\mathcal{I} \subseteq \mathcal{I}^*$ be of size $r := |\mathcal{I}| \geq \lfloor \frac{|\mathcal{I}^*|}{2} \rfloor - 1$. Then there exists a $\lambda \in \mathbb{N}$ and an index $i \in \mathcal{I}$ such that a.a.s. $e_\lambda^{\{i\}} \geq \frac{2}{3}s^2p$.*

Proof Claim 5.2 (i), (ii), and (iii) imply $e_{\lambda_1}^{\mathcal{I}} \geq \frac{2}{3}s^2p$ for some $\lambda_1 \in \mathbb{N}$. Thus, there must exist a subset $\mathcal{I}' \subseteq \mathcal{I}$ of size $|\mathcal{I}'| \leq \lceil \frac{|\mathcal{I}|}{n^{\frac{\varepsilon}{4}}} \rceil$ such that $e_{\lambda_1}^{\mathcal{I}'} \geq \frac{\frac{2}{3}s^2p}{n^{\frac{\varepsilon}{4}}}$, as otherwise we could take a partition $\mathcal{I} = \bigcup_{j \in [\lceil n^{\frac{\varepsilon}{4}} \rceil]} \mathcal{I}_j$ with sets $|\mathcal{I}_j| \leq \lceil \frac{r}{n^{\frac{\varepsilon}{4}}} \rceil$ for all $j \in [\lceil n^{\frac{\varepsilon}{4}} \rceil]$, leading to $e_{\lambda_1}^{\mathcal{I}} \leq \sum_{j \in [\lceil n^{\frac{\varepsilon}{4}} \rceil]} e_{\lambda_1}^{\mathcal{I}_j} < \frac{2}{3}sp^2$, contradicting the assumption. For this $\mathcal{I}'$, we have $e_{\lambda_1}^{\mathcal{I}'} \geq \frac{\frac{2}{3}s^2p}{n^{\frac{\varepsilon}{4}}} = \Theta\left(n^{\frac{3}{2}+\frac{\varepsilon}{4}}\right)$, hence Claim 5.2 (ii) and (iii) with $\mathcal{I} \mapsto \mathcal{I}'$ yield $e_{\lambda_1+\lambda_2^1}^{\mathcal{I}'} \geq \frac{2}{3}s^2p$ for some $\lambda_2^1 \in \mathbb{N}$. Iteratively applying this argument a small number $m := \lceil \frac{4 \log r}{\varepsilon \log n} \rceil$ of times until $|\mathcal{I}'| \leq \lceil \frac{r}{n^{\frac{\varepsilon}{4}m}} \rceil \leq 1$ gives us $i \in \mathcal{I}$ and $\{\lambda_2^j\}_{j \in [m]}$ such that $e_{\lambda_1+\sum_{j=1}^m \lambda_2^j}^{\{i\}} \geq \frac{2}{3}s^2p$. □

We exploit the aforementioned result to conclude the existence of an index $i \in \mathcal{I}^*$ such that a_i and b_i can be connected by a structure only using vertices from $S \setminus O$. Claim 5.3 with $\mathcal{I} \mapsto \mathcal{I}^*$ gives us a $\lambda_1^* \in \mathbb{N}$ and an index i_1 for which $e_{\lambda_1^*}^{\{i_1\}} \geq \frac{2}{3}s^2p$. After removal of this tuple $\{a_{i_1}^1, a_{i_1}^2\}$, Claim 5.3 with $\mathcal{I} \mapsto \mathcal{I}^* \setminus \{i_1\}$ yields $e_{\lambda_2^*}^{\{i_2\}} \geq \frac{2}{3}s^2p$ for a $\lambda_2^* \in \mathbb{N}$ and an index $i_2 \in \mathcal{I}^* \setminus \{i_1\}$. Repeating this argument $\lceil \frac{|\mathcal{I}^*|}{2} \rceil + 1$ many times, we get indices $i_1, \ldots, i_{\lceil \frac{|\mathcal{I}^*|}{2} \rceil + 1}$ and $\{\lambda_j^*\}_{1 \leq j \leq \lceil \frac{|\mathcal{I}^*|}{2} \rceil + 1}$ with $e_{\lambda_j^*}^{\{i_j\}} \geq \frac{2}{3}s^2p$ for $j \in \left[\lceil \frac{|\mathcal{I}^*|}{2} \rceil + 1\right]$. Let $\lambda^* := \max\left\{\lambda_j^* \mid j \in \left[\lceil \frac{|\mathcal{I}^*|}{2} \rceil + 1\right]\right\}$. We have $e_{\lambda^*}^{\{i_j\}} \geq \frac{2}{3}s^2p$ for all $j \in \left[\lceil \frac{|\mathcal{I}^*|}{2} \rceil + 1\right]$. By symmetry, this also holds for the graph sequence starting from the other side (in the following denoted with tilde), that is, $\widetilde{e}_{\lambda^*}^{\{\widetilde{i}_j\}} \geq \frac{2}{3}s^2p$ for all $j \in \left[\lceil \frac{|\mathcal{I}^*|}{2} \rceil + 1\right]$ and a set $\left\{\widetilde{i}_1, \ldots, \widetilde{i}_{\lceil \frac{|\mathcal{I}^*|}{2} \rceil + 1}\right\} \subseteq \mathcal{I}^*$. Hence, there must exist an $i \in \mathcal{I}^*$ such that $e_{\lambda^*}^{\{i\}} \geq \frac{2}{3}s^2p$ and $\widetilde{e}_{\lambda^*}^{\{i\}} \geq \frac{2}{3}s^2p$. Then for any l such that $\frac{l}{2} + 1 \geq \lambda^*$ and l is a multiple of 4, at least $\frac{2}{3}s^2p$ many edges in $E\left(B_{\frac{l}{2}+1}\right)$ can be reached from a_i and, symmetrically, at least $\frac{2}{3}s^2p$ many edges in $E\left(\widetilde{B}_{\frac{l}{2}+1}\right) = E\left(B_{\frac{l}{2}+1}\right)$ can be reached from b_i, by Claim 5.2 (iii). Therefore, there must exist an edge that can be reached both from a_i and b_i (as otherwise at least $\frac{4}{3}s^2p$ edges would be present in $B_{\frac{l}{2}+1} \subseteq G \subseteq G_0$, contradicting the Edge Concentration Lemma), implying the existence of a connecting structure of length l and hence concluding the proof of the Connecting Lemma.

5.3 Proof of Claim 5.2 (i)

In this section, we provide a proof of Claim 5.2 (i), that is, we show that there exists a $l_1 \in \mathbb{N}$ such that $n_{l_1+j} \geq c_v s$ for all $j \geq 0$, provided that we are starting with a set $\mathcal{I} \subseteq \mathcal{I}^*$ of size $|\mathcal{I}| \geq \lfloor \frac{|\mathcal{I}^*|}{2} \rfloor - 1$. To this end, we inductively keep track of (a lower bound on) the number of reachable vertices in every step.

Claim 5.4 *For all $i \in [l]$ there exist $n'_i \leq n_i$, $N'_i \subseteq N_i$ of size $|N'_i| = n'_i$, and $E'_i \subseteq E(B_i) \cap (N'_{f(i)} \times N'_i)$ of size $|E'_i| = n'_i$ such that a.a.s.*

(i) either $n'_i = c_v s$ or $n'_i \geq csp^2 n'_{i-1}$ for some constant $c > 0$,

(ii) $\Delta(E'_i) = o\left(\frac{1}{p^2}\right)$, and

(iii) $\Delta(E'_i, N'_i) = \delta(E'_i, N'_i) = 1$.

Proof Set $E'_0 := R$, $N'_0 := N_0$, and $n'_0 := n_0 = |E'_0| = |\mathcal{I}|$. As all the edges in R are disjoint, (ii) and (iii) are trivially satisfied for $i = 0$. We prove the claim by induction on i, that is, for any $i \in [l]$ we assume that (ii) and (iii) hold for $i-1$ and show that (i), (ii), and (iii) are true for i as well. Let $N'_{i-1} \subseteq N_{i-1}$ be a subset of the vertices reached, $n'_{i-1} \leq n_{i-1}$ its size, and $E'_{i-1} \subseteq E(B_{g(i),f(i)})$ the edge set. Let $0 < \delta < 1$ be such that $\delta \leq \frac{\beta}{3+4\delta}$. If $n'_{i-1} \geq \frac{\delta}{p^2}$, take an arbitrary subset $E \subseteq E'_{i-1}$ of size $|E| = \lfloor \frac{\delta}{p^2} \rfloor$. Otherwise, let $E := E'_{i-1}$. Claim 3.8 applied to $W \mapsto S_i$ and $E \mapsto E$ yields $|N^2_{G_0}(E, S_i)| \geq (1-3\delta)sp^2|E|$. Moreover, observe that $|N^2_G(E, O \cap S_i)| \leq o$. Taken together, also using

$$|N^2_H(E, S_i)| \leq e^2_H(E, S_i) \leq |E|(1-\beta)sp^2$$

due to the Degree Concentration Lemma on G_0 and the minimum codegree assumption in G, we have

$$\begin{aligned} n_i &\geq |N^2_G(E(B'_{i-1}, S'_i)| \geq |N^2_G(E, S'_i)| = |N^2_G(E, S_i)| - |N^2_G(E, O \cap S_i)| \\ &= |N^2_{G_0}(E, S_i)| - |N^2_H(E, S_i)| - |N^2_G(E, O \cap S_i)| \\ &\geq (\beta - 3\delta)sp^2|E| - o. \end{aligned}$$

If $n'_{i-1} \geq \frac{\delta}{p^2}$ and $|E| = \lfloor \frac{\delta}{p^2} \rfloor$, we thus have $n_i \geq \frac{\beta-3\delta}{\delta}s - o \geq \frac{\beta-3\delta}{4\delta}s \geq c_v s$ by the choice of δ and choosing c_o (hence $\frac{o}{s}$) sufficiently small. Set $n'_i := c_v s$.

On the other hand, if $n'_{i-1} < \frac{\delta}{p^2}$ and hence $|E| = |E'_{i-1}| = n'_{i-1}$, we have $o \leq 5l|\mathcal{I}| \leq 5l|E'_{i-1}|$, observing that $|E'_{i-1}| \geq |\mathcal{I}|$ in every step i due to (i) and $n'_0 = |\mathcal{I}|$. Thus $n_i \geq \left((\beta-3\delta)sp^2 - 5l\right) n'_{i-1} \geq csp^2 n'_{i-1}$ for some constant $c > 0$. In this case, set $n'_i := csp^2 n'_{i-1}$.

It remains to show the existence of an edge set E'_i. Take an arbitrary subset $N'_i \subseteq N_i$ of size $|N'_i| = n'_i$, and define $B'_i \subseteq B_i$ on $V(B'_i) = N'_{f(i)} \cup N'_i$ with

$$E(B'_i) := \left\{ \{v, w\} \mid v \in N'_{f(i)}, w \in N^2_G \left(\left(N_{E'_{i-1}}(v), v \right), N'_i \right) \right\}.$$

Since by the induction hypothesis every vertex $v \in N'_{i-1}$ has degree one in E'_{i-1}, we have $\left| N^2_G \left((N_{E'_{i-1}}(v), v), N'_i \right) \right| \leq (1+\gamma)sp^2$ by an upper bound on the codegree due to the Degree Concentration Lemma, hence $\Delta(B'_i, N'_{f(i)}) \leq (1+\gamma)sp^2 = o\left(\frac{1}{p^2}\right)$ (implicitly assuming without loss of generality that $\varepsilon < \frac{1}{2}$, as for larger ε the result follows by monotonicity). Now pick a subset $E'_i \subseteq E(B'_i)$ such that each $v \in N'_i$ has exactly one incident edge. This ensures that $\Delta(E'_i, N'_i) = \delta(E'_i, N'_i) = 1$ and, since $E'_i \subseteq E(B'_i)$, we have $\Delta(E'_i) = o\left(\frac{1}{p^2}\right)$, which concludes the proof of the induction step and hence of the claim. □

This result in particular implies that if $n'_i \geq c_v s$, then $n'_{i+j} \geq c_v s$ for all $j \geq 0$. Therefore, if $n'_i < c_v s$, then $n'_i \geq (csp^2)^i n'_0 \geq (cn^\varepsilon)^{l_1}$ by Claim 5.4 (i). Thus, $n_{l_1^*} \geq n'_{l_1^*} \geq c_v s$ for a $l_1^* > \frac{1}{\varepsilon}$ and hence $n_{l_1^*+j} \geq n'_{l_1^*+j} \geq c_v s$ for all $j \geq 0$. Finally, note that $n'_i \geq c_v s$ implies $e_{i+1} \geq n'_i \beta sp^2 - (1+\gamma)osp = \Omega(n^{1+\varepsilon})$, observing that every vertex in N'_i has degree 1 in E'_i by property (iii) in Claim 5.4 and using the lower bound on the codegree as well as a trivial upper bound on the edges between $S_{\min\{f(i),g(i)\}}$ and $S_{i+1} \cap O$ due to the Edge Concentration Lemma. Claim 5.2 (i) thus follows for a constant $\geq l_1^* + 1$.

5.4 Auxiliary results

This section aims to introduce results repeatedly used throughout the next sections. Relations between the in-degrees and out-degrees of a vertex as well as conjunctions between in-edges and out-edges are established. To this end, different types of vertices, based on their in-degree, are introduced, and their behaviors, that is, their out-degrees, are analyzed. For the rest of this chapter, let $i, j, k \in [l]$ be such that $(i, j, k) = (g(k), f(k), k)$.

5.4.1 Degree evolution

For a constant C, which we will choose later to be sufficiently large, define

$$\mathcal{T}_{j,i} := \left\{ v \in N_j \colon d^-_{j,i}(v) < \frac{C \log n}{p} \right\}$$

as the set of all vertices in N_j with tiny in-degree from set N_i,

$$\mathcal{M}_{j,i} := \left\{ v \in N_j \colon \frac{C \log n}{p} \leq d^-_{j,i}(v) < \left(\frac{1}{3} + \frac{\alpha}{2} \right) sp \right\}$$

the set of all vertices with medium in-degree, and

$$\mathcal{H}_{j,i} := \left\{ v \in N_j \colon d^-_{j,i}(v) \geq \left(\frac{1}{3} + \frac{\alpha}{2} \right) sp \right\}$$

as the set of all huge-in-degree vertices in N_j. Moreover, we aggregate the tiny set $\mathcal{T}_{j,i}$ and the medium set $\mathcal{M}_{j,i}$ to small-in-degree vertices $\mathcal{S}_{j,i} := \mathcal{T}_{j,i} \cup \mathcal{M}_{j,i}$ as well as $\mathcal{M}_{j,i}$ and $\mathcal{H}_{j,i}$ to large-in-degree vertices $\mathcal{L}_{j,i} := \mathcal{M}_{j,i} \cup \mathcal{H}_{j,i}$.

The next result shows how vertices with tiny, medium, and large in-degree, respectively, behave.

Claim 5.5 *The following holds asymptotically almost surely.*

(i) We have $\widehat{d}^+_{j,k}(v) \geq d^-_{j,i}(v) \frac{sp^2}{C_1 \log n}$ *for some constant* $C_1 > 0$ *and all* $v \in \mathcal{T}_{j,i}$.

(ii) If $|\mathcal{M}_{j,i}| = \Theta(s)$*, then all except* $o(s)$ *many vertices* $v \in \mathcal{M}_{j,i}$ *satisfy* $\widehat{d}^+_{j,k}(v) \geq \left(\frac{1}{3} + \frac{2\alpha}{3}\right) sp$.

(iii) If $|\mathcal{H}_{j,i}| = \Theta(s)$*, then all except* $o(s)$ *many vertices* $v \in \mathcal{H}_{j,i}$ *satisfy* $\widehat{d}^+_{j,k}(v) \geq d_G(v, S_k) - \frac{C \log n}{p}$.

Proof (i) Let $v \in \mathcal{T}_{j,i}$ be arbitrary, choose $\delta \in (0,1)$ such that $\frac{\delta}{1-\delta} < \beta$ and fix a subset $X \subseteq N^-_{j,i}(v)$ with $|X| = \min\{\frac{\delta}{p}, d^-_{j,i}(v)\}$, thus such that $|X| \geq \frac{\delta}{C \log n} d^-_{j,i}(v)$. Claim 3.7 with $U \mapsto X$ and $W \mapsto S_k$ then yields $d^2_{G_0}((X,v), S_k) \geq (1-\gamma)|X|sp^2$ for a $\gamma < \beta$. Using the upper bound on the maximum codegree in H, we thus get

$$\begin{aligned}
\widehat{d}^+_{j,k}(v) &= d^2_G \left((N^-_{j,i}(v), v), S_k \right) \geq d^2_G ((X,v), S_k) \\
&= d^2_{G_0} ((X,v), S_k) - d^2_H ((X,v), S_k) \\
&\geq (1-\gamma)|X|sp^2 - (1-\beta)|X|sp^2 = (\beta - \gamma)|X|sp^2 \\
&\geq \frac{(\beta-\gamma)\,\delta}{C \log n} sp^2 d^-_{j,i}(v) = \frac{1}{C_1 \log n} sp^2 d^-_{j,i}(v),
\end{aligned}$$

for some constant $C_1 > 0$.

(ii) First note that it is enough to show that for any $c > 0$ and for all $\mathcal{M}_{j,i} \subseteq N_j$ with $|\mathcal{M}_{j,i}| \geq cs$ there exists a $v^* \in \mathcal{M}_{j,i}$ such that $\widehat{d}^+_{j,k}(v^*) \geq \left(\frac{1}{3} + \frac{2\alpha}{3}\right) sp$. By successively removing the vertices v^* until a set smaller than $c's$ for any $c' > 0$ is reached, indeed at least $|\mathcal{M}_{j,i}| - o(s)$ many such vertices v^* can be found.

Towards a contradiction, assume that $\widehat{d}^+_{j,k}(v) < \left(\frac{1}{3} + \frac{2\alpha}{3}\right) sp$ for all $v \in \mathcal{M}_{j,i}$. Let $N^0(v) := N_G(v, S_k) \setminus \widehat{N}^+_{j,k}(v)$ and note that, by definition of these sets, $e_G(N^-_{j,i}(v), N^0(v)) = 0$ and

$$|N^0(v)| = d_G(v, S_k) - \widehat{d}^+_{j,k}(v) \geq \left(\frac{1}{3} + \frac{\alpha}{3}\right) sp,$$

using the Degree Concentration Lemma for a lower bound on $d_G(v, S_k)$. The Edge Concentration Lemma thus yields

$$\begin{aligned} e_{G_0}(N^-_{j,i}(v), N^0(v)) &\geq (1-\gamma_1)|N^-_{j,i}(v)||N^0(v)|p \\ &= (1-\gamma_1)\left(\frac{1}{3} + \frac{\alpha}{3}\right) d^-_{j,i}(v) sp^2 \\ &\geq \left(\frac{1}{3} + \frac{\alpha}{4}\right) d^-_{j,i}(v) sp^2. \end{aligned}$$

Moreover,

$$\begin{aligned} \max\{d_{G_0}(u, N^0(v)) \colon u \in N^-_{j,i}(v)\} &\leq \max\{d_{G_0}(u, N_G(v, S_k)) \colon u \in N^-_{j,i}(v)\} \\ &\leq \max\{d^2_{G_0}((u,v), S_k) \colon u \in N^-_{j,i}(v)\} \\ &\leq (1+\gamma_2) sp^2 \end{aligned}$$

for any $0 < \gamma_2 < 1$ using a trivial bound on the maximum codegree by the Degree Concentration Lemma. Therefore, Claim 3.13 (with $X \mapsto N^-_{j,i}(v)$ and $Y \mapsto N^0(v)$) lets us deduce the existence of a subset $U_v \subseteq N^-_{j,i}(v)$ of size $|U_v| = \kappa d^-_{j,i}(v) \geq \kappa \frac{C \log n}{p}$ for some $\kappa > 0$ such that

$$d_{G_0}(u, N^0(v)) \geq \left(\frac{1}{3} + \frac{\alpha}{8}\right) sp^2$$

for all $u \in U_v$. Let $U_{\mathcal{M}} := \bigcup_{v \in \mathcal{M}_{j,i}} U_v$ and define the auxiliary bipartite graph F_1 between $U_{\mathcal{M}}$ and $\mathcal{M}_{j,i}$ with $V(F_1) := U_{\mathcal{M}} \cup \mathcal{M}_{j,i}$ and

$$E(F_1) := \{\{u, v\} \mid v \in \mathcal{M}_{j,i}, u \in U_v\}.$$

We have

$$e(F_1) = \sum_{v \in \mathcal{M}_{j,i}} d_{F_1}(v, U_{\mathcal{M}}) = \sum_{v \in \mathcal{M}_{j,i}} |U_v| \geq |\mathcal{M}_{j,i}| \kappa \frac{C \log n}{p} \geq \kappa cs \frac{C \log n}{p}.$$

As $|U_{\mathcal{M}}| \leq |S_i| \leq s$, the average degree in F_1 of a vertex in $\mathcal{M}_{j,i}$ is lower bounded by $\kappa c \frac{C \log n}{p}$. Hence there must exist a $u^* \in U_{\mathcal{M}}$ with $d_{F_1}(u^*, \mathcal{M}_{j,i}) \geq \kappa c \frac{C \log n}{p}$. By definition of the auxiliary graph F_1 this means that u^* has at least $\kappa c \frac{C \log n}{p}$ many vertices $v \in \mathcal{M}_{j,i}$ for which $u^* \in U_v$. For each of those vertices $v \in N_{F_1}(u^*, \mathcal{M}_{j,i})$ we thus have

$$d_{G_0}(u^*, N^0(v)) \geq \left(\frac{1}{3} + \frac{\alpha}{8}\right) sp^2.$$

Let $W_{u^*} := \bigcup_{v \in N_{F_1}(u^*, \mathcal{M}_{j,i})} N_{G_0}(u^*, N^0(v))$ and introduce a second auxiliary bipartite graph F_2, with $V(F_2) := N_{F_1}(u^*, \mathcal{M}_{j,i}) \cup S_k$ and $E(F_2) := \left\{\{v, w\} \mid w \in N_{G_0}(u^*, N^0(v))\right\}$. Then $d_{F_2}(v, S_k) \geq \left(\frac{1}{3} + \frac{\alpha}{8}\right) sp^2$ and thus Claim 3.10 with $G \mapsto F_2$, $X \mapsto N_F(u^*, \mathcal{M}_{j,i})$, and $Y \mapsto S_k$ yields

$$|W_{u^*}| = \left|N_{F_2}\left(N_{F_1}(u^*, \mathcal{M}_{j,i}), S_k\right)\right| \geq \left(\frac{1}{3} + \frac{\alpha}{10}\right) sp.$$

We claim that $d_G(u^*, W_{u^*}) = 0$. Towards a contradiction, suppose there is an edge from u^* to some vertex $w \in W_{u^*}$ in G. By definition of W_{u^*}, there must exist a $v \in N_F(u^*, \mathcal{M}_{j,i})$ such that $w \in N_{G_0}(u^*, N^0(v)) \subseteq N^0(v) \subseteq N_G(v, S_k)$. So there is an edge from v to w in G as well. But this means that $w \in N_G^2((u, v), S_k) \subseteq \widehat{N}_{j,k}^+(v)$, violating the assumption $w \in N^0(v)$, hence proving $d_G(u^*, W_{u^*}) = 0$. Moreover, observe that $d_{G_0}(u^*, W_{u^*}) \geq |W_{u^*}| \geq \left(\frac{1}{3} + \frac{\alpha}{10}\right) sp$. Taken together,

$$d_H(u^*, W_{u^*}) = d_{G_0}(u^*, W_{u^*}) \geq \left(\frac{1}{3} + \frac{\alpha}{10}\right) sp,$$

which contradicts the upper bound on the maximum degree of H, and therefore concludes the proof.

(iii) The proof is very similar to the one of (ii). For more details, we refer to Section 8.2 in the appendix. □

The following claim provides a similar result valid for all $v \in \mathcal{L}_{j,i}$.

Claim 5.6 *We a.a.s. have $\widehat{d}_{j,k}^+(v) \geq \frac{7}{12}\beta sp$ for all $v \in \mathcal{L}_{j,i}$.*

Proof We have $d_G\left(u, N_G(v, S_k)\right) = d_G^2((u, v), S_k) \geq \beta sp^2$ for all $u \in N_{j,i}^-(v)$ and $v \in \mathcal{L}_{j,i}$ by the lower bound on the codegree. Hence, Claim 3.10, applied to $\alpha \mapsto \beta$, $X \mapsto N_{j,i}^-(v)$, and $Y \mapsto N_G(v, S_k)$, yields

$$\widehat{d}_{j,k}^+(v) = \left|N_G\left(N_{j,i}^-(v), N_G(v, S_k)\right)\right| \geq \frac{\beta}{1+\gamma}|N_G(v, S_k)| \geq \frac{\left(\frac{2}{3} + \alpha\right)\beta}{1+\gamma} sp,$$

using the lower bound on the degree. This can be bounded by $\frac{7}{12}\beta sp$. □

5.4.2 Edge evolution

The classification and the results from the previous section are used to find relations between the number of edges in the bipartite graphs.

Claim 5.7 *If $e_{j,i}(\mathcal{T}_{j,i}) \geq n^{1+\eta}$ for some $\eta > \frac{\varepsilon}{8}$, then there exists a constant $c > 0$ such that a.a.s. $e_{j,k} \geq \frac{1}{c \log^3 n} sp^2 e_{j,i}(\mathcal{T}_{j,i})$.*

Proof Let $0 < \delta < 1$ and $\frac{\delta}{1-\delta} < \gamma < \frac{\beta}{3}$ be arbitrarily small. Consider a subset E of the edges $E_{j,i}(\mathcal{T}_{j,i})$ obtained as follows. For each vertex $v \in \mathcal{T}_{j,i}$, delete all but $\frac{\delta}{p}$ many edges and for each vertex $u \in N^-_{j,i}(\mathcal{T}_{j,i})$, delete all but $\frac{sp}{\log^2 n}$ edges. Let $N'(v)$ denote the set of neighbors of v in E and $d'(v) := |N'(v)|$ its degree. Since $d^-_{j,i}(v) \leq \frac{C \log n}{p}$ for all $v \in \mathcal{T}_{j,i}$ and $d^+_{i,j}(u) \leq (1+\gamma)sp$ for any $u \in N_i$ by the Degree Concentration Lemma, we have

$$\frac{|E|}{e_{j,i}(\mathcal{T}_{j,i})} \geq \frac{\frac{\delta}{p}}{\frac{C \log n}{p}} \frac{\frac{sp}{\log^2 n}}{(1+\gamma)sp} = \frac{\delta}{(1+\gamma)C \log^3 n} \geq \frac{1}{C' \log^3 n}$$

for some $C' > 0$. Claim 3.7 applied to $U \mapsto N'(v)$ and $W \mapsto S_k$ implies $|N^2_{G_0}((N'(v), v), S_k)| \geq (1-\gamma)\, d'(v) sp^2$ for all $v \in \mathcal{T}_{j,i}$, thus $e^2_{G_0}(E, S_k) \geq (1-\gamma)sp^2|E|$. Observe that the number $e^2_{G_0}(E, S_k \cap O)$ of square-edges connecting edges from E to vertices in $S_k \cap O$ can be upper bounded by the number

$$X(S_k \cap O) = |\{(\{u,v\}, w) \in E \times S_k \cap O \colon \{u,w\}, \{v,w\} \in E(G_0)\}|$$

of triangles in G_0 into $S_k \cap O$ induced by edges in E. Let

$$X(S_k) = |\{(\{u,v\}, w) \in E \times S_k \colon \{u,w\}, \{v,w\} \in E(G_0)\}|$$

and

$$X(S'_k) = \left|\{(\{u,v\}, w) \in E \times S'_k \colon \{u,w\}, \{v,w\} \in E(G_0)\}\right|.$$

Then $X(S_k \cap O) = X(S_k) - X(S'_k)$. Note that $X(S_k) = \sum_{e \in E} d^2_{G_0}(e, S_k) \leq (1+\gamma)sp^2|E|$ using a bound on the maximum codegree derived from the Degree Concentration Lemma. Moreover, Claim 3.9 with $T \mapsto E$ and $W' \mapsto S'_k$ yields $X(S'_k) \geq (1-\gamma)|E||S'_k|p^2 \geq (1-\gamma)(s-o)p^2|E|$. Putting everything together,

$$\begin{aligned} e^2_{G_0}\left(E, S'_k\right) &\geq e^2_{G_0}\left(E, S_k\right) - X(S_k \cap O) = e^2_{G_0}\left(E, S_k\right) - X(S_k) + X(S'_k) \\ &\geq \left((1-\gamma) - (1+\gamma) + (1-\gamma)\left(1 - \frac{o}{s}\right)\right) sp^2|E| \\ &\geq \left(1 - 3\gamma - \frac{o}{s}\right) sp^2|E|. \end{aligned}$$

Thus, using the upper bound on the maximum codegree in H,

$$\begin{aligned} e_{j,k} &\geq e^2_G\left(E_{j,i}(\mathcal{T}_{j,i}), S'_k\right) \geq e^2_G\left(E, S'_k\right) = e^2_{G_0}\left(E, S'_k\right) - e^2_H\left(E, S'_k\right) \\ &\geq \left(1 - 3\gamma - \frac{o}{s}\right) sp^2|E| - (1-\beta)sp^2|E| = \left(\beta - 3\gamma - \frac{o}{s}\right) sp^2|E| \\ &\geq \frac{(\beta - 3\gamma - \frac{o}{s})}{C' \log^3 n} e_{j,i}(\mathcal{T}_{j,i}) \geq \frac{1}{c' \log^3 n} e_{j,i}(\mathcal{T}_{j,i}), \end{aligned}$$

for some $c' > 0$ by choosing γ and c_o sufficiently small. $\square$

If there are enough large vertices, then the following lower bound on the edges associated to large vertices can be found.

Claim 5.8 *If $|\mathcal{L}_{j,i}| \geq \frac{C \log n}{p}$, then a.a.s. $e_{j,k}(\mathcal{L}_{j,i}) \geq \frac{\beta}{2} sp|\mathcal{L}_{j,i}|$.*

Proof Let $Y \supseteq S_k \cap O$ be a superset of size $|Y| = c_o s > \frac{C \log n}{p}$. The Edge Concentration Lemma thus yields $e_{j,k}(\mathcal{L}_{j,i}, Y) \leq (1+\gamma)|\mathcal{L}_{j,i}|c_o sp$, and hence

$$\begin{aligned} e_{j,k}(\mathcal{L}_{j,i}) &\geq \widehat{e}_{j,k}(\mathcal{L}_{j,i}) - e_{j,k}(\mathcal{L}_{j,i}, Y) = \sum_{v \in \mathcal{L}_{j,i}} \widehat{d}^{+}_{j,k}(v) - e_{j,k}(\mathcal{L}_{j,i}, Y) \\ &\geq \frac{7}{12}\beta sp|\mathcal{L}_{j,i}| - (1+\gamma)c_o sp|\mathcal{L}_{j,i}| \geq \frac{\beta}{2} sp|\mathcal{L}_{j,i}|, \end{aligned}$$

employing Claim 5.6 and choosing c_o sufficiently small. □

The next claim is a corollary of Claim 5.5.

Claim 5.9 *We a.a.s. have*

(i) $\widehat{e}_{j,k}(\mathcal{T}_{j,i}) \geq \frac{sp^2}{C_1 \log n} e_{j,i}(\mathcal{T}_{j,i})$ *for some constant* $C_1 > 0$,

(ii) $\widehat{e}_{j,k}(\mathcal{M}_{j,i}) \geq \left(1 + \frac{\alpha}{10}\right) e_{j,i}(\mathcal{M}_{j,i})$ *if* $|\mathcal{M}_{j,i}| = \Theta(s)$, *and*

(iii) $\widehat{e}_{j,k}(\mathcal{H}_{j,i}) \geq (1-\delta)e_{j,i}(\mathcal{H}_{j,i})$ *for any* $0 < \delta < 1$ *if* $|\mathcal{H}_{j,i}| = \Theta(s)$.

Proof The statement in (i) is a direct consequence of Claim 5.5 (i). If $|\mathcal{M}_{j,i}| = \Theta(s)$, then Claim 5.5 (ii) implies

$$\widehat{e}_{j,k}(\mathcal{M}_{j,i}) \geq \left(|\mathcal{M}_{j,i}| - o(s)\right)\left(\frac{1}{3} + \frac{2\alpha}{3}\right) sp \geq \left(1 + \frac{\alpha}{10}\right) e_{j,i}(\mathcal{M}_{j,i}),$$

which proves (ii). For $|\mathcal{H}_{j,i}| = \Theta(s)$, the observation that due to the uniform partitioning of S into $\{S_r\}_{r \in [l]}$ we have $d_G(v, S_{i'}) = (1 - o(1))d_G(v, S_{j'})$ for all $i', j' \in [l]$ and Claim 5.5 (iii) yield $\widehat{d}_{j,k}(\mathcal{H}_{j,i}) = (1 - o(1))e_{j,i}(\mathcal{H}_{j,i})$. □

The following claim establishes a relation between $e_{j,i}$ and $e_{j,k}$.

Claim 5.10 *If $e_{j,i} \geq c'_e s^2 p$, then for any $0 < \delta < 1$ we a.a.s. have*

$$e_{j,k} \geq \left(1 + \frac{\alpha}{20}\right) e_{j,i}(\mathcal{S}_{j,i}) + (1-\delta)e_{j,i}(\mathcal{H}_{j,i}) \geq (1-\delta)e_{j,i}.$$

Proof If $|\mathcal{S}_{j,i}| \geq 100c_o s$, say, then $\widehat{e}_{j,k}(\mathcal{S}_{j,i}) \geq \left(1 + \frac{\alpha}{10}\right) e_{j,i}(\mathcal{S}_{j,i})$ by Claim 5.9 (i) and (ii). If both $|\mathcal{S}_{j,i}|, |\mathcal{H}_{j,i}| \geq 100c_o s$, Claim 5.9 (iii) thus yields

$$\begin{aligned} e_{j,k} &= \widehat{e}_{j,k}(\mathcal{S}_{j,i}) + \widehat{e}_{j,k}(\mathcal{H}_{j,i}) - \bar{e}_{j,k} \\ &\geq \left(1 + \frac{\alpha}{10}\right) e_{j,i}(\mathcal{S}_{j,i}) + \left(1 - \frac{\delta}{2}\right) e_{j,i}(\mathcal{H}_{j,i}) - (1+\gamma)osp \\ &\geq \left(1 + \frac{\alpha}{20}\right) e_{j,i}(\mathcal{S}_{j,i}) + (1-\delta)\, e_{j,i}(\mathcal{H}_{j,i}) \end{aligned}$$

for any $0 < \delta < 1$ using a trivial upper bound on $\widehat{e}_{j,k}$ and choosing c_o sufficiently small.

If $|\mathcal{H}_{j,i}| < 100c_o s$ (and thus $|\mathcal{S}_{j,i}| > 100c_o s$ in order to have $e_{j,i} \geq c_e' s^2 p$), then $e_{j,i}(\mathcal{H}_{j,i}) \leq (1+\gamma)100c_o s^2 p$ by the Edge Concentration Lemma, thus

$$\begin{aligned}
e_{j,k} &\geq \left(1+\frac{\alpha}{10}\right) e_{j,i}(\mathcal{S}_{j,i}) - (1+\gamma)osp \\
&= \left(1+\frac{\alpha}{10}\right) e_{j,i}(\mathcal{S}_{j,i}) + (1-\delta)\, e_{j,i}(\mathcal{H}_{j,i}) - (1+\gamma)osp - (1-\delta)\, e_{j,i}(\mathcal{H}_{j,i}) \\
&\geq \left(1+\frac{\alpha}{20}\right) e_{j,i}(\mathcal{S}_{j,i}) + (1-\delta)\, e_{j,i}(\mathcal{H}_{j,i})
\end{aligned}$$

observing that $e_{j,i}(\mathcal{S}_{j,i}) \geq (c_e' - (1+\gamma)100c_o)s^2 p$ (by the Edge Concentration Lemma) and choosing c_o sufficiently small. The case $|\mathcal{S}_{j,i}| < 100c_o s$ can be proved analogously. □

5.5 Proof of Claim 5.2 (ii)

In this section, we prove Claim 5.2 (ii) by first deriving a sufficient condition for reaching at least $c_e s^2 p$ edges in the next step and then showing that indeed after some number of steps this condition must be satisfied. The main idea is as follows. Since vertices in $\mathcal{L}_{j,i}$ have lots of outgoing edges in $B_{j,k}$, it is enough to reach lots of large vertices in order to reach the prescribed amount of edges. In every step, either the number of edges increases or, otherwise, there must be a lot of large vertices for which the out-degree cannot grow by too much compared to the in-degree. In any case, this brings us another step closer to our goal.

To simplify the notation, we assume without loss of generality that $e_0^{\mathcal{I}} \geq n^{1+\eta}$ for some $\eta > \frac{\varepsilon}{8}$ and show that $e_{l_2}^{\mathcal{I}} \geq c_e s^2 p$ for some $l_2 \in \mathbb{N}$.

Claim 5.11 *If $|\mathcal{L}_{j,i}| \geq c_{\mathcal{L}} s$, then a.a.s. $e_{j,k} \geq c_e s^2 p$.*

Proof Claim 5.8 yields $e_{j,k} \geq e_{j,k}(\mathcal{L}_{j,i}) \geq \frac{\beta}{2} sp|\mathcal{L}_{j,i}| \geq \frac{\beta c_{\mathcal{L}}}{2} s^2 p \geq c_e s^2 p$. □

We call a step (i, j, k) good if $e_{j,k} \geq n^{\frac{\varepsilon}{4}} e_{j,i}$ and bad if it is not good. Intuitively, as tiny in-degree vertices $\mathcal{T}_{j,i}$ grow by a lot, that is, give much more out-edges (in $B_{j,k}$) than they have in-edges (in $B_{j,i}$), in a bad step we must have a lot of vertices with large in-degree. The next claim shows that indeed in a bad step almost all edges are incident to $\mathcal{L}_{j,i}$ and $\mathcal{L}_{j,i}$ must be large.

Claim 5.12 *Let (i, j, k) be bad with $e_{j,i}(\mathcal{T}_{j,i}) \geq n^{1+\eta}$ for some $\eta > \frac{\varepsilon}{8}$. Then a.a.s.*

(i) $(1-o(1))e_{j,i} = e_{j,i}(\mathcal{L}_{j,i}) \geq n^{\frac{\varepsilon}{2}} e_{j,i}(\mathcal{T}_{j,i})$, *and*

(ii) $|\mathcal{L}_{j,i}| \geq \frac{p}{n^{\frac{\varepsilon}{2}}} n_j$.

Proof For a bad step, we have

$$e_{j,i}(\mathcal{L}_{j,i}) \geq \frac{e_{j,k} - n^{\frac{\varepsilon}{4}} e_{j,i}(\mathcal{T}_{j,i})}{n^{\frac{\varepsilon}{4}}} \geq \frac{\frac{sp^2}{c' \log^3 n} - n^{\frac{\varepsilon}{4}}}{n^{\frac{\varepsilon}{4}}} e_{j,i}(\mathcal{T}_{j,i}) \geq n^{\frac{\varepsilon}{2}} e_{j,i}(\mathcal{T}_{j,i})$$

by Claim 5.7. Thus $e_{j,i} \geq \left(1 + n^{\frac{\varepsilon}{2}}\right) e_{j,i}(\mathcal{T}_{j,i})$, hence $e_{j,i}(\mathcal{T}_{j,i}) = o(e_{j,i})$, which implies $e_{j,i}(\mathcal{L}_{j,i}) = (1 - o(1))e_{j,i}$ and therefore proves (i). Using $|\mathcal{T}_{j,i}| \leq e_{j,i}(\mathcal{T}_{j,i})$, a trivial upper bound on the maximum degree yields

$$|\mathcal{L}_{j,i}| \geq \frac{e_{j,i}(\mathcal{L}_{j,i})}{(1+\gamma)sp} \geq \frac{\frac{sp^2}{c' \log^3 n} - n^{\frac{\varepsilon}{4}}}{(1+\gamma)spn^{\frac{\varepsilon}{4}}} |\mathcal{T}_{j,i}| \geq \frac{p}{(1+2\gamma)c' \log^3 n n^{\frac{\varepsilon}{4}}} |\mathcal{T}_{j,i}| \geq \frac{p}{n^{\frac{3}{8}\varepsilon}} |\mathcal{T}_{j,i}|.$$

Thus $\left(1 + \frac{p}{n^{\frac{3}{8}\varepsilon}}\right) |\mathcal{L}_{j,i}| \geq \frac{p}{n^{\frac{3}{8}\varepsilon}} n_j$, which implies $|\mathcal{L}_{j,i}| \geq \frac{p}{n^{\frac{\varepsilon}{2}}} n_j$. □

Next, a sequence of good steps surrounded by a bad step each is considered.

Claim 5.13 *Let $j \geq 0$ and consider an almost square-path for arbitrary $k \in [l]$ or an alternating square-path with odd $k \in [l]$. Suppose that $(g(k), f(k), k)$ is a bad step, $\{(g(i), f(i), i)\}_{k \leq i \leq k+j}$ are good steps, and $(g(k+j+1), f(k+j+1), k+j+1)$ is a bad step. Then a.a.s. either $|\mathcal{L}_{k+j,f(k+j)}| \geq n^{\nu} |\mathcal{L}_{f(k),g(k)}|$ or $e_j \geq c_e s^2 p$ for a $j \in \{k+1, \ldots, k+1+j\}$.*

Proof As for $e_{f(k),g(k)}(\mathcal{T}_{f(k),g(k)}) = \Theta(e_{f(k),g(k)})$ Claim 5.12 (i) also implies that $e_{f(k),g(k)}(\mathcal{L}_{f(k),g(k)}) = \Theta(e_{f(k),g(k)})$, we have $|\mathcal{L}_{f(k),g(k)}| = \Omega\left(\frac{e_{f(k),g(k)}}{sp}\right) \gg \frac{\log n}{p}$ by a trivial upper bound on the maximum degree. Therefore, by Claim 5.8, $e_k = e_{f(k),k} \geq e_{f(k),k}(\mathcal{L}_{f(k),g(k)}) \geq \frac{\beta}{2} sp |\mathcal{L}_{f(k),g(k)}|$. As the number of edges increases in good steps, hence $e_{k+j'} \geq e_k$ for any $j' \in [j]$, we can assume without loss of generality that we do not have any good step, thus suppose $j = 0$, which means that $(g(k+1), f(k+1), k+1)$ is a bad step as well. Let $X := \left\{ v \in N_k \colon d^-_{k,f(k)}(v) \geq \frac{1}{10} |\mathcal{L}_{f(k),g(k)}| p \right\}$. Claim 3.12 with $X \mapsto \mathcal{L}_{f(k),g(k)}$ and $Y \mapsto N_k$ implies $|X| \geq \frac{2\beta c_v}{5} s$. If all except for at most $\frac{c_v \beta}{5} s$, say, vertices in X are in $\mathcal{L}_{k,f(k)}$, we have $|\mathcal{L}_{k,f(k)}| \geq \frac{c_v \beta}{5} s > c_{\mathcal{L}} s$, and hence, by Claim 5.11, with $(i, j, k) \mapsto (f(k), k, k+1)$ we have $e_{k+1} = e_{k,k+1} \geq c_e s^2 p$. Otherwise, at least $\frac{c_v \beta}{5} s$ of the vertices in X are in $\mathcal{T}_{k,f(k)}$, in which case $|\mathcal{T}_{k,f(k)} \cap X| \geq \frac{c_v \beta}{5} s$, and thus

$$e_{k,f(k)}(\mathcal{T}_{k,f(k)}) \geq \sum_{v \in \mathcal{T}_{k,f(k)} \cap X} d^-_{k,f(k)}(v) \geq \frac{c_v \beta^2}{50} |\mathcal{L}_{f(k),g(k)}| sp.$$

Since $(f(k), k, k+1)$ is a bad step as well, Claim 5.12 (i) yields

$$e_{k,f(k)}(\mathcal{L}_{k,f(k)}) \geq n^{\frac{\varepsilon}{2}} e_{k,f(k)}(\mathcal{T}_{k,f(k)}) \geq \frac{c_v \beta^2}{50} n^{\frac{\varepsilon}{2}} |\mathcal{L}_{f(k),g(k)}| sp,$$

and thus, by a trivial upper bound on the maximum degree,

$$|\mathcal{L}_{k,f(k)}| \geq \frac{c_v \beta^2}{50(1+\gamma)} n^{\frac{\varepsilon}{2}} |\mathcal{L}_{f(k),g(k)}| > n^{\frac{\varepsilon}{4}} |\mathcal{L}_{f(k),g(k)}|,$$

which concludes the proof. □

We now put these results together to prove Claim 5.2 (ii). Let l_g be even such that $l_g > \lceil \frac{2}{\frac{\varepsilon}{4}} \rceil$. Then we cannot have l_g successive good steps, as this would lead to $e_{l_g} \geq n^{\frac{\varepsilon}{4} l_g} e_0 \gg n^2$ (observing that for alternating square-paths we have $e_i = e_{i,i-1} \geq n^{\frac{\varepsilon}{4}} e_{i-1,i-3} \geq n^{\frac{\varepsilon}{2}} e_{i-3,i-2} = n^{\frac{\varepsilon}{2}} e_{i-2,i-3}$ for odd i) which obviously is not possible. We thus can have at most $l_g - 1$ successive good steps, followed by at least one bad. Since good steps increase the edge count and our argument only depends on the number of edges (see the observation in the proof of Claim 5.13), we can assume without loss of generality that all steps are bad, counting each bad step as l_g many.

If $|\mathcal{L}_{f(k),g(k)}| \geq c_{\mathcal{L}} s$ for any step $(g(k), f(k), k)$, then $e_k \geq c_e s^2 p$ by Claim 5.11. Hence, we can assume without loss of generality that $|\mathcal{L}_{f(k),g(k)}| < c_{\mathcal{L}} s$ for all k. In the case of an almost square-path, if we do not have $e_i \geq c_e s^2 p$ for any i, then $|\mathcal{L}_{i,i-1}| \geq n^{(i-2)\frac{\varepsilon}{4}} |\mathcal{L}_{2,1}| \geq n^{(i-2)\frac{\varepsilon}{4}}$ by Claim 5.13, making use of the fact that $|\mathcal{L}_{2,1}| \geq 1$ for a bad step $(1, 2, 3)$, as only large vertices can decrease the edge count. For an alternating square-path, if we do not have $e_i \geq c_e s^2 p$ for any i, then $|\mathcal{L}_{k,k-2}| \geq n^{\frac{\varepsilon}{4}} |\mathcal{L}_{k-2,k-1}|$ for odd k by Claim 5.13. Observe that $\mathcal{L}_{k,k+1} \supseteq \mathcal{L}_{k,k-2}$, therefore $|\mathcal{L}_{k,k+1}| \geq n^{\frac{\varepsilon}{4}} |\mathcal{L}_{k-2,k-1}|$. As $(k-3, k-4, k-2)$ and $(k-4, k-2, k-1)$ are bad steps (and $k-2$ is odd) as well, iteratively applying the same argument yields $|\mathcal{L}_{k,k+1}| \geq n^{\frac{k-1}{2}\frac{\varepsilon}{4}} |\mathcal{L}_{1,2}| \geq n^{\frac{k-1}{2}\frac{\varepsilon}{4}}$. In both cases, for j and k sufficiently large, respectively, at least $c_e s^2 p$ edges must be reached, as otherwise there would be too many large vertices. Thus, we can find an l_2 such that after l_2 (non-summarized) steps at least $c_e s^2 p$ many edges can be reached, concluding the proof of Claim 5.2 (ii).

5.6 Proof of Claim 5.2 (iii)

To simplify notation, let us assume without loss of generality that $e_0^{\mathcal{I}} \geq c_e s^2 p$ and prove that $e_{l_3+j}^{\mathcal{I}} \geq \frac{2}{3} s^2 p$ for a $l_3 \in \mathbb{N}$ and all $j \geq 0$.

We first show that once we have reached a lot of huge vertices, then we also have reached the prescribed number of edges.

Claim 5.14 *If $|\mathcal{H}_{j,i}| \geq \left(\frac{2}{3} + \frac{\alpha}{5}\right) s$, then a.a.s. $e_{k+2+r} \geq \frac{2}{3} s^2 p$ for all $r \geq 0$.*

Proof As having fewer huge vertices leads to more out-edges, we can assume without loss of generality that $|\mathcal{H}_{j,i}| = \left(\frac{2}{3} + \frac{\alpha}{5}\right) s$, thus $|S_j \setminus \mathcal{H}_{j,i}| =$

$\left(\frac{1}{3} - \frac{4\alpha}{5}\right) s$. Let $X := \left\{v \in S_k \colon d_{G_0}(v, S_j \setminus \mathcal{H}_{j,i}) \geq \left(\frac{1}{3} + \frac{\alpha}{5}\right) sp\right\}$. Claim 3.11 applied to $X \mapsto X$ and $Y \mapsto S_j \setminus \mathcal{H}_{j,i}$ implies $|X| < \frac{C \log n}{p}$. So $s - |X| = s(1 - o(1))$ vertices $v \in S_k \setminus X$ have degree

$$\begin{aligned} d_G(v, \mathcal{H}_{j,i}) &\geq d_{G_0}(v, S_j) - d_{G_0}(v, S_j \setminus \mathcal{H}_{j,i}) - d_H(v, S_j) \\ &\geq (1-\gamma)sp - \left(\frac{1}{3} + \frac{\alpha}{5}\right) sp - \left(\frac{1}{3} - \alpha\right) sp \\ &\geq \left(\frac{1}{3} + \frac{4\alpha}{5} - \gamma\right) sp \end{aligned}$$

by the Degree Concentration Lemma in G_0 and the upper bound on the maximum degree in H. Claim 5.5 (iii) implies that at most $|\mathcal{H}_{j,i}| \frac{C \log n}{p} + o(s^2 p)$ many edges can be part of G but not of $B_{j,k}$ between $H_{j,i}$ and S_k, since for $|\mathcal{H}_{j,i}| - o(s)$ many vertices $\frac{C \log n}{p}$ many edges could be lost, and, due to a trivial upper bound on the degree, for the remaining $o(s)$ many vertices possibly $(1+\gamma)sp$ many edges could be lost. Thus,

$$e_G(\mathcal{H}_{j,i}, S_k \setminus O \setminus X) - e_{j,k}(\mathcal{H}_{j,i}, S_k \setminus O \setminus X) \leq |\mathcal{H}_{j,i}| \frac{C \log n}{p} + o(s^2 p) \tag{5.1}$$

in the worst case when we suppose that all these lost edges are incident to $S_{i+1} \setminus O \setminus X$. Let $Y := \left\{v \in S_k \setminus O \setminus X \mid d_G(v, \mathcal{H}_{j,i}) - d^-_{k,j}(v, \mathcal{H}_{j,i}) \geq \frac{\alpha}{5} sp\right\}$ be the set of vertices in $S_k \setminus O \setminus X$ for which a large fraction of the incident edges in G are not part of $B_{k,j}$. A comparison of

$$\begin{aligned} e_G(\mathcal{H}_{j,i}, S_k \setminus O \setminus X) - e_{B_{j,k}}(\mathcal{H}_{j,i}, S_k \setminus O \setminus X) &\geq e_G(\mathcal{H}_{j,i}, Y) - e_{B_{j,k}}(\mathcal{H}_{j,i}, Y) \\ &= \sum_{v \in Y} \left(d_G(v, \mathcal{H}_{j,i}) - d_{B_{j,k}}(v, \mathcal{H}_{j,i})\right) \\ &\geq \frac{\alpha}{5} |Y| sp \end{aligned}$$

to (5.1) yields

$$|Y| \leq \frac{|\mathcal{H}_{j,i}| \frac{C \log n}{p} + o(s^2 p)}{\frac{\alpha}{5} sp} = \Theta(n^{1-\varepsilon} \log n) = o(s).$$

Therefore, the set $Z := S_k \setminus O \setminus X \setminus Y$ still has size $|Z| \geq (1 - \frac{o}{s} - o(1))s$. For all these vertices $v \in Z$ we have lost at most $\frac{\alpha}{5} sp$ many edges into $\mathcal{H}_{j,i}$, hence

$$d^-_{k,j}(v) \geq d^-_{k,j}(v, \mathcal{H}_{j,i}) \geq d_G(v, \mathcal{H}_{j,i}) - \frac{\alpha}{5} sp \geq \left(\frac{1}{3} + \frac{3\alpha}{5} - \gamma\right) sp.$$

In particular, for $0 < \gamma < \frac{\alpha}{20}$, we have $d^-_{k,j}(v) \geq \left(\frac{1}{3} + \frac{\alpha}{2}\right) sp$ for all $v \in Z \subseteq S_k$, which implies that $Z \subseteq \mathcal{H}_{k,j}$.

Claim 5.5 (iii) then yields

$$e_{k+2} \geq (|Z| - o(s)) \left(\left(\frac{2}{3} + \alpha \right) sp - \frac{C \log n}{p} \right) - (1+\gamma)osp \geq \frac{2}{3}s^2 p,$$

recalling the lower bound on the degree and the fact that the number of edges into $O \cap S_{k+2}$ can be bounded $(1+\gamma)osp$, which is sufficiently small for small enough c_o.

Note that $Z \subseteq \mathcal{H}_{k,j}$ implies $|\mathcal{H}_{k,j}| \geq |Z| \geq \left(1 - \frac{o}{s} - o(1)\right) s > \left(\frac{2}{3} + \frac{\alpha}{5}\right) s$. Iteratively applying the same argument again, it follows that $e_{k+2+r} \geq \frac{2}{3}s^2 p$ for all $r \geq 0$, which concludes the proof. □

The rest of this section is designated to the proof that indeed lots of huge vertices can be reached. After having established the auxiliary claims, in the end everything is put together to prove Claim 5.2 (iii).

In the following, we say that a step (i, j, k) is good if $e_{j,k} \geq (1 + \alpha^{20}) e_{j,i}$ and bad otherwise.

Claim 5.15 *In a bad step (i, j, k) with $e_{j,i} \geq c'_e s^2 p$, we a.a.s. have*

(i) $e_{j,i}(\mathcal{H}_{j,i}) \geq \frac{1}{\alpha^{10}} e_{j,i}(\mathcal{S}_{j,i})$, *and*

(ii) $|\mathcal{H}_{j,i}| \geq \frac{c'_e}{1+\alpha^9} s$.

Proof Claim 5.10 implies

$$\begin{aligned}\left(1 + \frac{\alpha}{20}\right) e_{j,i}(\mathcal{S}_{j,i}) + (1-\delta) e_{j,i}(\mathcal{H}_{j,i}) \leq e_{j,k} &< (1+\alpha^{20}) e_{j,i} \\ &= (1+\alpha^{20}) e_{j,i}(\mathcal{S}_{j,i}) + (1+\alpha^{20}) e_{j,i}(\mathcal{H}_{j,i}),\end{aligned}$$

hence

$$e_{j,i}(\mathcal{H}_{j,i}) \geq \frac{\frac{\alpha}{20} - \alpha^{20}}{\alpha^{20} + \delta} e_{j,i}(\mathcal{S}_{j,i}) \geq \frac{1}{\alpha^{10}} e_{j,i}(\mathcal{S}_{j,i}),$$

picking δ sufficiently small. This proves (i) and implies

$$e_{j,i}(\mathcal{H}_{j,i}) \geq \frac{1}{\alpha^{10}\left(1 + \frac{1}{\alpha^{10}}\right)} e_{j,i} = \frac{1}{1+\alpha^{10}} e_{j,i}.$$

Using a trivial upper bound on the degree, we thus get

$$|\mathcal{H}_{j,i}| \geq \frac{1}{(1+\gamma)sp} e_{j,i}(\mathcal{H}_{j,i}) \geq \frac{1}{(1+\gamma)sp} \frac{1}{1+\alpha^{10}} e_{j,i} \geq \frac{c'_e}{1+\alpha^9} s,$$

which concludes the proof of (ii). □

The following claim shows that in two consecutive bad steps the number of huge vertices must grow.

Claim 5.16 *Consider an almost square-path for arbitrary $k \in [l]$ or an alternating square-path with odd $k \in [l]$ such that $(g(k), f(k), k)$ and $(f(k), k, k+1)$ are bad steps with $e_{f(k),g(k)}, e_{k,g(k)} \geq c'_e s^2 p$. Then a.a.s. either $|\mathcal{H}_{k,f(k)}| \geq \left(\frac{2}{3} + \frac{\alpha}{5}\right) s$ or $|\mathcal{H}_{k,f(k)}| \geq \frac{1}{\alpha} |\mathcal{H}_{f(k),g(k)}|$.*

Proof Let $X := \left\{ v \in S_k : d^-_{k,f(k)}(v) \geq \kappa |\mathcal{H}_{f(k),g(k)}| p \right\}$ for some $0 < \kappa < 1$. As $|\mathcal{H}_{f(k),g(k)}| = \Theta(s)$ by Claim 5.15 (ii), Claim 5.5 (iii) and the minimum degree condition imply

$$\begin{aligned} \widehat{e}_{f(k),k}(\mathcal{H}_{f(k),g(k)}) &\geq \left(|\mathcal{H}_{f(k),g(k)}| - o(s) \right) \left(\left(\frac{2}{3} + \alpha \right) sp - \frac{C \log n}{p} \right) \\ &\geq |\mathcal{H}_{f(k),g(k)}| \left(\frac{2}{3} + \frac{\alpha}{2} \right) sp. \end{aligned}$$

Claim 3.12 with $X \mapsto \mathcal{H}_{f(k),g(k)}$ and $Y \mapsto S_k$ thus yields $|X| \geq \left(\frac{2}{3} + \frac{\alpha}{10}\right) s$, say. If all except at most $\frac{\alpha}{20} s$ vertices from X are in $\mathcal{H}_{k,f(k)}$, then $|\mathcal{H}_{k,f(k)}| \geq |X| - \frac{\alpha}{20} s \geq \left(\frac{2}{3} + \frac{\alpha}{20}\right) s$, which proves the claim for this case. On the other hand, if more than $\frac{\alpha}{20} s$ of the vertices in X are not in $\mathcal{H}_{k,f(k)}$, then, as $\mathcal{S}_{k,f(k)} \subseteq \mathcal{S}_{k,f(k)} \cap X = X \setminus \mathcal{H}_{k,f(k)} \setminus O$, we have $|\mathcal{S}_{k,f(k)}| \geq \frac{\alpha}{20} s - o \geq \frac{\alpha}{40} s$, by choosing c_o sufficiently small. Thus,

$$e^-_{k,f(k)}(\mathcal{S}_{k,f(k)}) = \sum_{v \in \mathcal{S}_{k,f(k)}} d^-_{k,f(k)}(v) \geq \sum_{v \in \mathcal{S}_{k,f(k)} \cap X} d^-_{k,f(k)}(v) \geq \frac{\alpha}{10} s\kappa |\mathcal{H}_{f(k),g(k)}| p,$$

and, since it is a bad step, by Claim 5.15 (i),

$$e_{k,f(k)}(\mathcal{H}_{k,f(k)}) \geq \frac{1}{\alpha^{10}} e_{k,f(k)}(\mathcal{S}_{k,f(k)}) \geq \frac{\alpha\kappa}{10\alpha^{10}} sp |\mathcal{H}_{f(k),g(k)}| \geq \frac{1}{\alpha^2} sp |\mathcal{H}_{f(k),g(k)}|,$$

say. A trivial upper bound on the degrees yields

$$|\mathcal{H}_{k,f(k)}| \geq \frac{1}{\alpha^2(1+\gamma)} |\mathcal{H}_{f(k),g(k)}| \geq \frac{1}{\alpha} |\mathcal{H}_{f(k),g(k)}|,$$

which finishes the proof. □

We now exploit the aforementioned results to prove Claim 5.2 (iii). Let $L \in \mathbb{N}$ be sufficiently large such that $\left(\frac{1}{\alpha}\right)^{\frac{L-4}{2}} \frac{c'_e}{1+\alpha^9} \geq 2$, choose δ small enough such that $(1-\delta)^L \geq \frac{1}{1+\alpha^{20}}$, $j^* \in \mathbb{N}$ large enough such that $\left(\frac{1}{1-\delta}\right)^{j^*} \geq \frac{2}{c_e}$, and c'_e sufficiently small such that $c'_e \leq (1-\delta)^{Lj^*} c_e$. In a sequence of Lj^* steps, we must have at least L successive bad steps as otherwise, if we only have at most $L-1$ successive bad steps, hence at least one good step in every L steps, by definition of a good step and Claim 5.10,

$$e_{j^*L} = e_{(j^*L, f(j^*L))} \geq (1-\delta)^{j^*(L-1)} (1+\alpha^{20})^{j^*} e_0 \geq \left(\frac{1}{1-\delta}\right)^{j^*} e_0 \geq 2s^2 p,$$

which contradicts the maximum possible number of edges between S_{j^*L} and $S_{f(j^*L)}$ due to the Edge Concentration Lemma. Note that $e_i \geq c'_e s^2 p$ for all $i \in [Lj^*]$, since

$$e_i \geq (1-\delta)^i e_0 \geq (1-\delta)^i c_e s^2 p \geq (1-\delta)^{Lj^*} c_e s^2 p \geq c'_e s^2 p.$$

Let $\{(g(k), f(k), k)\}_{k^* \leq k \leq k^*+L-1}$ be the L successive bad steps among the first Lj^* steps. We can suppose $|\mathcal{H}_{f(k),g(k)}| < \left(\frac{2}{3} + \frac{\alpha}{5}\right) s$ for all $k \in \{k^*, \dots, k^* + L-1\}$, as otherwise $e_{k+2+r} \geq \frac{2}{3} s^2 p$ for all $r \geq 0$ by Claim 5.14.

In the case of an almost square-path, iterative application of Claim 5.16 and Claim 5.15 (ii) applied for $(i,j,k) \mapsto (k^*-2, k^*-1, k^*)$ imply

$$|\mathcal{H}_{k^*+L-2,k^*+L-3}| \geq \left(\frac{1}{\alpha}\right)^{L-1} |\mathcal{H}_{k^*-1,k^*-2}| \geq \left(\frac{1}{\alpha}\right)^{L-1} \frac{c'_e}{1+\alpha^9} s \geq 2s,$$

leading to a contradiction.

Now, consider the case of an alternating square-path and let $k \in \{k^* + L - 2, k^* + L - 1\}$ be odd. Claims 5.16 applied to $(i,j,k) \mapsto (k-1, k-2, k)$ yields implies $|\mathcal{H}_{k,k-2}| \geq \frac{1}{\alpha} |\mathcal{H}_{k-2,k-1}|$. Claim 5.15 (ii) thus tells us that $|\mathcal{H}_{k,k-2}|, |\mathcal{H}_{k-2,k-1}| = \Theta(s)$. Claim 5.5 (iii) for $(i,j,k) \mapsto (k-2, k, k+1)$ implies $d^+_{k,k+1}(v) \geq d_G(v, S_{k+1}) - \frac{C \log n}{p} > \left(\frac{1}{3} + \frac{\alpha}{2}\right) sp$ for all $v \in \mathcal{H}_{k,k-2}$, hence $\mathcal{H}_{k,k+1} \supseteq \mathcal{H}_{k,k-2}$ and therefore $|\mathcal{H}_{k,k+1}| \geq \frac{1}{\alpha} |\mathcal{H}_{k-2,k-1}|$. Iteratively applying the same argument and also using Claim 5.15 (ii), we get, similarly as for the almost square-path, $|\mathcal{H}_{k,k+1}| \geq \left(\frac{1}{\alpha}\right)^{\frac{L-4}{2}} \frac{c'_e}{1+\alpha^9} s \geq 2s$, say, leading to a contradiction.

Thus, after at most $Lj^* + 2$ steps we must have reached at least $\frac{2}{3} s^2 p$ edges, and the observation in the proof of Claim 5.14 shows that we cannot drop below $\frac{2}{3} s^2 p$ afterwards, which concludes the proof of Claim 5.2 (iii).

Chapter 6

Proof of Theorem 1.2

In this chapter, we provide a proof of Pósa's conjecture for random graphs.

Theorem 1.2 *(Pósa's conjecture for random graphs)*

Let $n \in \mathbb{N}$ be sufficiently large and $\alpha, \beta, \varepsilon > 0$. Consider a random graph $G_0 \sim \mathcal{G}(n, p)$ for $p^2 \geq n^{-1+\varepsilon}$ and a subgraph $H \subseteq G_0$ with maximum degree $\Delta(H) \leq \left(\frac{1}{3} - \alpha\right) np$ and maximum codegree $\Delta^2(H) \leq (1-\beta)np^2$. Then a.a.s. the graph $G := G_0 \setminus H$ contains a square of a Hamilton cycle.

We make use of the following result of Noever and Steger [NS16b].

Lemma 6.1 *(Covering Lemma)*

Let $n \in \mathbb{N}$ be sufficiently large and $\alpha, \varepsilon > 0$. Then a.a.s. every subgraph $G \subseteq \mathcal{G}(n, p)$ with minimum degree $\delta(G) \geq \left(\frac{2}{3} + \alpha\right) np$ contains a square of a cycle on at least $(1-\alpha)n$ vertices, provided that $p^2 \geq n^{-1+\varepsilon}$.

Proof of Theorem 1.2 First note that Claim 3.3 implies $\delta(G) \geq \left(\frac{2}{3} + \alpha'\right) np$ for any $\alpha' < \alpha$ as well as $\delta^2(G) \geq \beta' np^2$ for any $\beta' < \beta$. Randomly split G into two subgraphs G_1 and G_2 of equal size (to simplify the notation, we assume that n is even) and note that an argument similar as in the proof of Lemma 3.4 yields $\delta(G_1), \delta(G_2) \geq \left(\frac{2}{3} + \alpha''\right) \frac{n}{2} p$ and $\delta^2(G_2), \delta^2(G_2) \geq \beta'' \frac{n}{2} p^2$ for any $\alpha'' < \alpha'$ and $\beta'' < \beta'$. Moreover, observe that G_1 and G_2 are subgraphs of $\mathcal{G}(\frac{n}{2}, p)$ satisfying $p^2 \geq n^{-1+\varepsilon} \geq \left(\frac{n}{2}\right)^{-1+\varepsilon'}$ for an arbitrary $\varepsilon' > \varepsilon$.

The Absorbing Lemma, applied to $n \mapsto \frac{n}{2}$, $G \mapsto G_1$, $\alpha \mapsto \alpha''$, $\beta \mapsto \beta''$, and $\varepsilon \mapsto \varepsilon'$, yields a constant $c_1 \in (0,1)$, a random subset $V_1 \subseteq V(G_1)$ of size $|V_1| = \frac{c_1}{2} n$, and an absorbing path $P_{V_1} \subseteq G_1$ for V_1. Note that, as V_1 is a random subset of G_1 and hence of G, the argument in the proof of Lemma 3.4 implies $\delta(G, V_1) \geq \left(\frac{2}{3} + \overline{\alpha}\right) |V_1| p$ for any $\overline{\alpha} < \alpha'$ as well as $\delta^2(G, V_1) \geq \overline{\beta} |V_1| p^2$ for any $\overline{\beta} < \beta'$.

Let $U := V(G) \setminus V_1 \setminus V(P_{V_1})$ be the vertices not used in the absorbing structure. We have

$$\frac{1}{2}n \leq |U| = n - |V_1|\,(1 + (4l + l(l-1)) + l) + l \leq \left(1 - \frac{l^2 c_1}{2}\right) n$$

and

$$\delta(G[U]) \geq \delta(G, V(G_2)) \geq \delta(G_2) \geq \left(\frac{2}{3} + \alpha''\right)\frac{n}{2} \geq \left(\frac{2}{3} + \alpha_0\right)|U|$$

for a sufficiently small $\alpha_0 > 0$.

Lemma 6.1 with $\alpha \mapsto \alpha_0$, $\varepsilon \mapsto \varepsilon'$, $n \mapsto |U|$, and $G \mapsto G[U]$ yields an almost spanning square-cycle $Q \subseteq G[U]$ using all but $\alpha_0|U|$ many vertices. Let W be the set of uncovered vertices in U, that is, $W := U \setminus V(Q)$ of size $|W| \leq \alpha_0|U|$, and let $s(Q)$ be the first two and $t(Q)$ the last two vertices on Q (for an arbitrary natural ordering and arbitrarily breaking the cycle). By Claim 3.4, we can split V_1 into two sets V_1^1 and V_1^2 of equal size such that the relative degrees are transferred, that is, $\delta(G, X) \geq \left(\frac{2}{3} + \frac{\bar{\alpha}}{2}\right)|X|p$ and $\delta^2(G, X) \geq \frac{\bar{\beta}}{2}|X|p^2$, say, for $X \in \{V_1^1, V_1^2\}$. Apply Claim 3.14 with $X \mapsto W$, $Y \mapsto V_1^1$, and $l \mapsto 4$ to find a sequence of $|W|$ many square-paths $L_1, \dots, L_{|W|}$ of length 4 starting at the vertices in W (which is possible by choosing α_0 sufficiently small compared to c_1). Let $s(L_i)$ be the first two vertices on L_i and $t(L_i)$ the last two vertices on L_i. Similarly, let $s(P_{V_1})$ be the first two and $t(P_{V_1})$ be the last two vertices on P_{V_1}, respectively. Connect

$$\begin{aligned}&\{(t(P_{V_1}), s(Q))\} \cup \{t(Q), s(L_1)\}\\ &\quad \cup \bigcup_{i \in [|W|-1]} \{(t(L_i), s(L_{i+1}))\} \cup \left\{\left(t(L_{|W|}), s(V_{P_1})\right)\right\}\end{aligned}$$

by Lemma 5.1 with $S \mapsto V_1^2$ with almost square-paths, which is possible by choosing α_0 and hence $|W|$ sufficiently small compared to $|V_1^2| = \frac{c_1}{4}n$. This provides us with a cycle C. Let $V_1' := V_1 \setminus V(C)$ be the set of vertices in V_1 not occurring in the cycle and note that $V_1' \cup V(C) = V(G)$. Exploiting the absorbing property of P_{V_1} for V_1', that is, replacing P_{V_1} by $P^*_{V_1'}$, we can absorb V_1' into P_{V_1}, and thus, since $P_{V_1} \subseteq C$, the square-cycle C can be extended to a square of a Hamilton cycle. This concludes the proof of Theorem 1.2. □

Chapter 7

Conclusion

In this thesis, we revisit Pósa's conjecture and strengthen it regarding its robustness in two ways.

On the one hand, we show that not only a graph satisfying the assumption of this conjecture but also a random subgraph of it contains a square of a Hamilton cycle. More concretely, we prove that for any $\alpha, \varepsilon \in (0,1)$ and all graphs G with $\delta(G) \geq \left(\frac{2}{3} + \alpha\right) n$ a random subgraph $\mathcal{G}(G,p)$ of G a.a.s. is 2-Hamiltonian, provided that $p^2 \geq n^{-1+\varepsilon}$. This generalizes a result of Krivelevich, Lee, and Sudakov [KLS14] about robust Hamiltonicity.

On the other hand, we prove that for all $\alpha, \beta, \varepsilon \in (0,1)$ and $p^2 \geq n^{-1+\varepsilon}$ every subgraph $G \subseteq \mathcal{G}(n,p)$ with minimum degree $\delta(G) \geq \left(\frac{2}{3} + \alpha\right) np$ contains a square of a Hamilton cycle, whenever we impose the supplemental condition on the codegree that $\delta^2(G) \geq \beta n p^2$. In other words, $\mathcal{G}(n,p)$ has local resilience $\frac{1}{3} - \alpha$ when additionally assuming a non-trivial lower bound on the minimum codegree. This constitutes an extension of the local resilience of $\mathcal{G}(n,p)$ with respect to Hamiltonicity [LS12b], and can be seen as a generalization of the threshold result for the appearance of squares of Hamilton cycles in random graphs by Nenadov and Škorić [NŠ16a], as it shows that $\mathcal{G}(n,p)$ not only is square-Hamiltonian but even resiliently possesses this property for comparable edge probability.

Summarized, we let the base graph (that is, G in the first and $\mathcal{G}(n,p)$ in the second case) be such that 2-Hamiltonicity is directly implied (either by Pósa's conjecture or by the threshold result in random graphs) and show that after edge deletions (random removal or (restricted) adversarial removal) the graph still possesses this property. Both results are almost optimal concerning the minimum degree condition, due to the tightness of Pósa's conjecture, as well as the range of p.

Besides its independent interest, this thesis provides intriguing techniques that are promising to further study in a broader context. Especially the proof of the Connecting Lemma, which constitutes the heart of the thesis, offers a lot of insights into how similar problems could be approached. For instance, the general idea of tracking the number of reached vertices and their degrees, which is inspired by the cascading approach in [LSS10], can be adopted for other (related) structures, not only for powers of paths. As also the applicability of the hypergraph matching theorem is not restricted to a particular type of connecting structure, a Connecting Lemma for many variations of paths could be proved analogously. Furthermore, absorbing structures for similar spanning graphs could be defined (and found in a subgraph of a random graph) alike.

7.1 Improvements

However, there are still several issues to tackle before the problem of determining the local resilience of $\mathcal{G}(n, p)$ with respect to the property of containing a kth power of a Hamilton cycle is completely resolved. In this section, we discuss the problems arising when trying to get rid of the slacks in the parameters.

For a tight minimum degree (that is, $\delta(G) = \frac{2}{3}np$), completely new strategies are needed, since the additional slack in the degree is crucial for the proof of the Connecting Lemma, hence implicitly also for the Absorbing Lemma. Furthermore, it is also used in the proof in [NS16b] for the almost spanning cycle. The usual approach for these kinds of problems is to distinguish several types of such graphs (for instance extremal and non-extremal, or another characterization), providing different proofs (exploiting new ideas) for each case.

For $p^2 \geq \frac{\text{polylog} n}{n}$, which is the (approximate) threshold, the Connecting Lemma, and hence also the Absorbing Lemma, can be proved in exactly the same way, when the length of the connecting paths is extended from constant to logarithmic length. It still remains to find an almost spanning square-cycle, though, which turns out to be the weak point of our proof regarding the strengthening of the edge probability, since the result in [NS16b] only works with the extra epsilon in the power of n.

Many of the auxiliary results and even main lemmas can be proved analogously for $k > 2$, only the Connecting Lemma is restricted to $k = 2$. This constraint is not for convenience but is crucial. The combinatorial proof of the Connecting Lemma indeed heavily relies on the fact that after some point almost all vertices in the cascade have out-degree $\approx \frac{2}{3}np$, yielding $\Theta(n^2 p)$ many reached edges, which are basically all the edges present in the graph. For $k > 2$, similar arguments would lead to degrees $\approx \frac{k}{k+1}np^{k-1}$ and hence

$\Theta(n^2 p^{k-1}) \ll n^2 p$ edges only. Opposed to the case for $k = 2$, this cannot be exploited to conclude the existence of a connecting structure. Thus, a completely new idea is needed for this generalization.

The condition imposed on the minimum codegree in Theorem 1.2 seems to be very natural. Nevertheless, it is a legitimate question to ask whether there could be other sufficient assumptions. For example, could it be replaced by a lower bound on the number of triangles a vertex is contained in? It would be interesting to find further and especially weaker notions implying k-Hamiltonicity.

7.2 Variations

There are many possible extensions and variations of the problem of robustness of the Pósa-Seymour conjecture. In this section, we briefly mention several other measures (besides random removal and local resilience) of robustness that have been studied for Dirac's theorem and would be interesting to generalize to the Pósa-Seymour conjecture.

Probably one of the most famous notions of robustness is *global resilience*, defined as the minimum integer r such that by removing r edges from G the property $\mathcal{P}$ can be destroyed. However, while for local properties (e.g. the containment of small subgraphs) global resilience appears to be appropriate, most global properties (e.g. the appearance of a spanning structure) can be violated by simple local changes (by deletion of all edges incident to one vertex of minimum degree, say), which is why for such properties the notion of global resilience does not convey what one would expect.

A more suitable robustness measure is based on the number of edge-disjoint copies of powers of Hamilton cycles. Intuitively, the more edge-disjoint Hamilton cycles one can find, the less removal of edges can harm the property. For Hamiltonicity, this has been investigated in [FK08].

Another possible variation could be the study of non-uniform constraints on the number of deleted edges incident to a single vertex. Ben-Shimon, Krivelevich, and Sudakov [BSKS11b] introduced the notion of $\mathbf{d}$-resilience for a sequence $\mathbf{d} = (d_1, \ldots, d_n)$ of integers. A graph G then is called $\mathbf{d}$-resilient with respect to a property $\mathcal{P}$ if for every subgraph $H \subseteq G$ with $d_H(i) \leq d_i$ for all $i \in [n]$ the graph $G \setminus H$ possesses $\mathcal{P}$. The authors have analyzed this robustness measure for Dirac's theorem.

A further measure of robustness is obtained by imposing the condition that even if not all subgraphs are admissible, that is, even if some of the spanning graphs have to be excluded from consideration, the graph must contain a power of a Hamilton cycle. This is investigated in [KLS16] for Hamilton cycles.

Chapter 8

Appendix

8.1 Proofs of basic tools and properties

This section provides proofs for the claims in Chapter 3. For the sake of readability, we restate all the results before proving them.

8.1.1 Minimum degree and codegree conditions

Claim 3.3 *Let $n \in \mathbb{N}$ be sufficiently large, $\alpha, \beta \in (0,1)$, and H a graph on n vertices such that $\Delta(H) \le \left(\frac{1}{3} - \alpha\right) np$ and $\Delta^2(H) \le (1-\beta)np^2$. Then there exist constants $C, \alpha', \beta' > 0$ such that a.a.s. $\delta(G) \ge \left(\frac{2}{3} + \alpha'\right) np$ and $\delta^2(G) \ge \beta' np^2$ for $G_0 \sim \mathcal{G}(n,p)$ and $G := G_0 \setminus H$, provided that $p^2 \ge \frac{C \log n}{n}$.*

Proof Fix an arbitrary graph H such that $\Delta(H) \le \left(\frac{1}{3} - \alpha\right) np$ and $\Delta^2(H) \le (1-\beta)np^2$. Let $G_0 \sim \mathcal{G}(n,p)$ and $G := G_0 \setminus H$. The Degree Concentration Lemma implies that for all $j \in \{1,2\}$, $v \in \binom{V(G_0)}{j}$, and $\gamma \in (0,1)$, we have $d^j_{G_0}(v) \ge (1-\gamma)np^j$. Hence,

$$d_G(v) = d_G(v) - d_H(v) \ge \left(1 - \gamma - \left(\frac{1}{3} - \alpha\right)\right) np = \left(\frac{2}{3} + \alpha'\right) np$$

for all $v \in V(G')$ and $\alpha' := \alpha - \gamma$, as well as

$$d^2_G(v) \ge (1 - \gamma - (1-\beta))\, np^2 = \beta' np^2$$

for all $v \in \binom{V(G)}{2}$ and $\beta' := \beta - \gamma$, what concludes the proof. □

Lemma 3.4 *Let $n \in \mathbb{N}$ be sufficiently large, $r \in [n]$, $\beta_j \in (0,1]$ for $j \in \{1,2\}$, and $p := p(n) \in (0,1]$. Consider a graph G on n vertices and a subset $U \subseteq V(G)$. Then for any $0 < \gamma < 1$ there exists a constant C such that for all $u_1, \dots, u_r \in [n]$ with $\sum_{i \in [r]} u_i \le |U|$ and $u_i \ge \frac{C \log n}{p^2}$ for $i \in [r]$ there exist disjoint subsets $U_1, \dots, U_r \subseteq U$ with $|U_i| = u_i$ for $i \in [r]$ such that for all $i \in [r]$ a.a.s.*

(i) $d_G(v, U_i) \geq (1-\gamma)\beta_1 u_i p$ *for all* $v \in V(G)$ *with* $d_G(v, U) \geq \beta_1 |U| p$*, and*

(ii) $d_G^2(v, U_i) \geq (1-\gamma)\beta_2 u_i p^2$ *for all* $v \in \binom{V(G)}{2}$ *with* $d_G^2(v, U) \geq \beta_2 |U| p^2$.

Proof Let $U = \bigcup_{i \in [r]} U_i \cup Z$ be a partition of U taken uniformly at random among all partitions for which $|U_i| = u_i$ for $i \in [r]$ and $|Z| = |U| - \sum_{i \in [r]} u_i$. Moreover, let $W_j := \{v \in \binom{V(G)}{j} \mid d_G^j(v, U) \geq \beta_j |U| p^j\}$ be the set of unordered tuples with multiplicative lower bound $\beta_j |U| p^j$ on the j-degree into U and consider such a tuple $v \in W_j$. The codegree of v in U_i for $i \in [r]$ follows a hypergeometric distribution with mean $\frac{d_G^j(v,U) u_i}{|U|}$, as the the joint neighborhood of v in the random subset $U_i \subseteq U$ is the intersection of U_i of size u_i with $N_G^j(v, U)$ of size $d_G^j(v, U)$. Therefore

$$\mathbb{E}[d_G^j(v, U_i)] = d_G^j(v, U) \frac{u_i}{|U|} \geq \beta_j u_i p^j \geq C \beta_j \log n.$$

Lemma 2.4 hence yields an upper bound

$$\Pr\left[d_G^j(v, U_i) \leq (1-\gamma) d_G^j(v, U) \frac{u_i}{|U|}\right] \leq e^{-\frac{C\gamma^2}{2}\beta_j \log n}.$$

The set W_j contains at most $\binom{n}{j}$ elements and, since the size of each part U_i is a positive integer, the number r of parts is at most n. Taking the union bound over all $j \in [k]$ and vertices in W_j, we thus get, for sufficiently large C, an upper bound

$$\sum_{j \in \{1,2\}} |W_j| r e^{-\frac{C\gamma^2}{2}\beta_j \log n} \leq 2 e^{3 \log n - \frac{C\gamma^2}{2} \min\{\beta_1, \beta_2\} \log n} = o(1)$$

on the probability that there exists a $j \in \{1, 2\}$, a tuple of vertices $v \in W_j$, and an $i \in [r]$ such that $d_G^j(v, U_i) < (1-\gamma)\beta u_i p^j$, which completes the proof. □

8.1.2 Properties of binomial random graphs

Lemma 3.6 *(Edge Concentration)*

Let $n \in \mathbb{N}$ *be sufficiently large and* $p := p(n) \in (0, 1]$*. Then for any* $0 < \gamma < 1$ *there exists a constant* C *such that the random graph* $G \sim \mathcal{G}(n, p)$ *a.a.s. has* $e_G^j(X^j, Y) = (1 \pm \gamma)\,|X^j||Y| p^j$ *for all* $j \in \{1, 2\}$*,* $Y \subseteq V(G)$*, and sets* $X^j \subseteq \binom{V(G) \setminus Y}{j}$ *of disjoint j-tuples, provided that* $|X^j|, |Y| \geq \frac{C \log n}{p^2}$.

Proof Let $j \in \{1, 2\}$, $Y \subseteq V(G)$ and $X^j \subseteq \binom{V(G) \setminus Y}{j}$ of sizes $|X^j|, |Y| \geq \frac{C \log n}{p^2}$ be arbitrary. The number $e_G^j(X^j, Y)$ of j-edges between X^j and Y follows a

Bin $\left(|X^j||Y|, p^j\right)$ distribution, since, due to the disjointness of the tuples, all the edges are distinct. By Lemma 2.4, we have

$$\Pr\left[e^j_G(X^j, Y) \neq (1 \pm \gamma)|X^j||Y|p^j\right] \leq e^{-c|X^j||Y|p^j}$$

for some $c > 0$. A union bound over all choices of $j \in \{1,2\}$, X^j, and Y thus implies, for sufficiently large C, an upper bound of

$$\begin{aligned}\sum_{x,y=\frac{C\log n}{p^2}}^{n} \sum_{j\in\{1,2\}} n^{jx+y} e^{-cxyp^j} &= \sum_{x,y=\frac{C\log n}{p^2}}^{n} \sum_{j\in\{1,2\}} e^{-C'\max\{x,y\}\log n} \\ &\leq \sum_{x,y=\frac{C\log n}{p^2}}^{n} \sum_{j\in\{1,2\}} e^{-C'\log n} = o(1)\end{aligned}$$

on the probability that such sets X^j, Y exist, when we choose C and C' appropriately. □

Claim 3.7 *Let $n \in \mathbb{N}$ be sufficiently large, $\delta \in \left(0, \frac{1}{2}\right)$, and $p := p(n) \in (0,1]$. Consider $G \sim \mathcal{G}(n,p)$, a set $W \subseteq V(G)$, and a vertex $v \in V(G) \setminus W$. Then for all sets $U \subseteq V(G) \setminus (W \cup \{v\})$ of size $|U| \leq \frac{\delta}{p}$ and any $\frac{\delta}{1-\delta} < \gamma < 1$ there exists a constant C such that a.a.s. $|N^2_G((U,v),W)| = (1 \pm \gamma)|U||W|p^2$, provided that $|W| \geq \frac{C\log n}{p^2}$.*

Proof First note that Claim 2.2 yields

$$\Pr\left[w \in N^2_G((U,v),W)\right] = \Pr\left[\exists u \in U\colon w \in N^2_G((u,v),W)\right] \leq |U|p^2.$$

Let U be arbitrary and enumerate the vertices in $U = \{u_1, \ldots, u_{|U|}\}$. We have

$$\begin{aligned}p_w &:= \Pr\left[w \in N^2_G((U,v),W)\right] = \Pr\left[\exists u \in U\colon w \in N^2_G((u,v),W)\right] \\ &= \sum_{1\leq i_1\leq |U|} \left|N^2_G((u_i,v),W)\right| - \sum_{1\leq i_1<i_2\leq |U|} \left|N^2_G((u_{i_1},v),W) \cap N^2_G((u_{i_2},v),W)\right| \\ &\quad + \sum_{1\leq i_1<i_2<i_3\leq |U^j|} \left|N^j_G(u_{i_1},W) \cap N^j_G(u_{i_2},W) \cap N^j_G(u_{i_3},W)\right| - \ldots \\ &\geq |U|p^2 - |U|^2p^3 - |U|^3p^4 - \ldots = |U|p^2\left(1 - \sum_{i=1}^{\infty}(|U|p)^i\right) \\ &\geq |U|p^2\left(1 - \sum_{i=1}^{\infty}\delta^i\right) = |U|p^2\left(1 - \frac{\delta}{1-\delta}\right)\end{aligned}$$

for the probability that a fixed vertex $w \in W$ is in $N^2_G((U,v),W)$, using Claim 2.3. Since the fact whether $w \in N^2_G((U,v),W)$ only depends on edges incident to w and U, these events are independent, which implies that $|N_G^2((U,v),W)| \sim \mathrm{Bin}(|W|, p_w)$. By Lemma 2.4, there exists a constant C' such that

$$\Pr\left[|N_G^2((U,v),W)| \neq (1\pm\gamma)|W||U|p^2\right] \leq e^{-C'|U|\log n}.$$

A union bound over sets U yields an upper bound

$$\sum_{|U|\in\left[\frac{\delta}{p}\right]} \binom{n}{|U|} e^{-C'|U|\log n} \leq \sum_{|U|\in\left[\frac{\delta}{p}\right]} e^{|U|(\log n - C'\log n)} = o(1)$$

on the probability of existence of such a set, making use of the fact that $\binom{n}{|U|} \leq n^{|U|}$ by Claim 2.1 and choosing C and hence C' large enough. □

Claim 3.8 *Let $n \in \mathbb{N}$ be sufficiently large, $p := p(n) \in (0,1]$, and $G \sim \mathcal{G}(n,p)$. Then for any $0 < \delta < 1$ there exists a constant C such that for all $W \subseteq V(G)$ with $|W| \geq \frac{C\log n}{p^2}$ and all sets $T \subseteq \binom{V(G)\setminus W}{2}$ of tuples with $|T| \leq \frac{\delta}{p^2}$ and maximum degree $\Delta(T) \leq \frac{\delta}{p}$ a.a.s. $|N_G^2(T,W)| = (1\pm 3\delta)\,|T||W|p^2$.*

Proof For any $w \in W$, we have

$$p_W := \Pr[w \in N_G^2(T,W)] = \Pr\left[w \in \bigcup_{e\in T} N_G^2(e,W)\right],$$

and thus

$$p_W \leq \sum_{e\in T} \Pr[w \in N_G^2(e,W)] \leq |T|p^2$$

by Claim 2.2 as well as

$$\begin{aligned} p_W &\geq \sum_{e\in T} \Pr[w \in N_G^2(e,W)] - \sum_{e\neq e'\in T} \Pr[w \in N_G^2(e,W) \cup N_G^2(e',W)] \\ &= |T|p^2 - \sum_{e\neq e'\in T:\, e\not\sim e'} p^4 - \sum_{e\neq e'\in T:\, e\sim e'} p^3 \\ &\geq |T|p^2 - |T|^2p^4 - 2|T|\Delta(E)p^3 = |T|p^2\left(1 - |T|p^2 - 2\Delta(T)p\right) \\ &\geq |T|p^2\,(1-\delta-2\delta) = |T|p^2\,(1-3\delta)\,, \end{aligned}$$

using Claim 2.3 and the fact that each edge can have at most $2\Delta(T)$ dependent tuples. Whether $w \in N_G^2(T,W)$, only depends on edges incident to w, which means that the events for $w \neq w'$ are independent, thus $|N_G^2(T,W)| \sim \mathrm{Bin}(|W|, p_W)$. Lemma 2.4 thus yields

$$\Pr[|N_G^2(T,W)| \neq (1\pm 3\delta)|W|p_W] \leq e^{-c|W||T|p^2}$$

for some constant c. A union bound over all such sets T then results in an upper bound

$$\sum_{|T|\in\left[\frac{\delta}{p^2}\right]} \binom{n}{|T|} e^{-c|W||T|p^2} \leq \sum_{|T|\in\left[\frac{\delta}{p^2}\right]} e^{|T|\log n - cC|T|\log n} = o(1)$$

for sufficiently large C, using $\binom{n}{|T|} \le n^{|T|}$ by Claim 2.1. □

Claim 3.9 *Let $n \in \mathbb{N}$ be sufficiently large, $\tau \in (0,1]$, and $p := p(n) \in (0,1]$ such that $p^2 = \omega\left(\frac{1}{n}\right)$. Consider $G \sim \mathcal{G}(n,p)$, disjoint subsets $U, V, W \subseteq V(G)$ with $|W| = \tau n$, as well as a set $T \subseteq U \times V$. For $W' \subseteq W$, let $X^{W'}$ be the number of triangles in G induced by T and W', that is,*

$$X^{W'} := \left|\{(u,v,w) \mid u \in U, v \in V, w \in W' \colon \{u,v\} \in T, \{u,w\}, \{v,w\} \in E(G)\}\right|.$$

Then for any $0 < \gamma < 1$, all $W' \subseteq W$ of size $|W'| = \Theta(n)$, and sets T with $|T| = \omega\left(\frac{n}{\log^2 n}\right)$ and $\Delta(T) = \mathcal{O}\left(\frac{|W'|p}{\log^2 n}\right)$, a.a.s. $X^{W'} \ge (1-\gamma)|T||W'|p^2$.

Proof For every tuple $e \in T$ and vertex $w \in W'$, let $X_{e,w}$ be the indicator random variable for the event that e together with w spans a triangle, that is,

$$X_{\{u,v\},w} := \begin{cases} 1, & \text{if } \{u,w\}, \{v,w\} \in E(G) \\ 0, & \text{otherwise.} \end{cases}$$

Then

$$\mu := \mathbb{E}\left[X^{W'}\right] = \sum_{e \in T, w \in W'} \Pr\left[w \in N_G^2(e, W')\right] = |T||W'|p^2.$$

Observe that we need $w = w'$ and $|e \cap e'| = 1$ in order to have $X_{e,w} \sim X_{e',w'}$, in which case

$$\Pr[X_{e,w} = 1, X_{e',w'} = 1] = \Pr[w = w' \in N_G^3((e \cap e', e \setminus e', e' \setminus e), W')] = p^3.$$

Since for every vertex $v \in V$ there are $\binom{d_T(v)}{2}$ many vertex pairs $(u,u') \in \binom{U}{2}$ for which $\{u,v\}, \{u',v\} \in T$, and similarly, for every $u \in U$ there are $\binom{d_T(u)}{2}$ such vertex pairs $(v,v') \in \binom{V}{2}$, we obtain

$$\begin{aligned}
\Delta &:= \sum_{(e,w) \sim (e',w') \in T \times W'} \Pr[X_{e,w} = 1, X_{e',w'} = 1] \\
&\le |W'| \left(\sum_{v \in V} \binom{d_T(v)}{2} p^3 + \sum_{u \in U} \binom{d_T(u)}{2} p^3 \right) \\
&\le |W'| \left(\sum_{v \in V} d_T(v)^2 + \sum_{u \in U} d_T(u)^2 \right) p^3 \\
&\le |W'| \Delta(T) \left(\sum_{v \in V} d_T(v) + \sum_{u \in U} d_T(u) \right) p^3 \le 2|W'|p^3|T|\Delta(T) \\
&= \mathcal{O}\left(\frac{|W'|^2 |T| p^4}{\log^2 n} \right) = \mathcal{O}\left(\frac{1}{|T| \log^2 n} \mu^2 \right).
\end{aligned}$$

Claim 2.5 thus implies

$$\Pr\left[X^{W'} \leq (1-\gamma)|W'||T|p^2\right] \leq e^{-\frac{\gamma^2\mu^2}{4\Delta}} = e^{-\Omega(|T|\log^2 n)}$$

and a union bound over all $W' \subseteq W$ and edge sets E yields an upper bound

$$\sum_{|W'|=\Theta(n)} \binom{\tau n}{|W'|} \sum_{|T|=0}^{n^2} \binom{n^2}{|T|} e^{-\Omega(|T|\log^2 n)} = e^{\mathcal{O}(n)+\Theta(|T|\log n)-\Omega(|T|\log^2 n)} = o(1),$$

as $\binom{\tau n}{|W'|} \leq \left(\frac{\tau n e}{|W'|}\right)^{|W'|} = e^{\mathcal{O}(n)}$ and $\binom{n^2}{|T|} \leq n^{2|T|} = e^{\Theta(|T|\log n)}$ by Claim 2.1. □

8.1.3 Subgraphs of binomial random graphs

Lemma 3.10 *(Density Lemma)*

Let $n \in \mathbb{N}$ be sufficiently large, $\alpha \in (0,1)$, $p := p(n) \in (0,1]$, and G a subgraph of $\mathcal{G}(n,p)$. Then for any $0 < \gamma < 1$ there exists a constant C such that for all disjoint sets $X, Y \subseteq V(G)$ with $d_G(x,Y) \geq \alpha|Y|p$ for all $x \in X$, we a.a.s. have $|N_G(X,Y)| \geq \frac{\alpha}{1+\gamma}|Y|$, provided that $|X|, |Y| \geq \frac{C\log n}{p}$.

Proof Suppose $|N_G(X,Y)| < \frac{\alpha}{1+\gamma}|Y|$ for some $0 < \gamma < 1$ and sets X, Y as above. On the one hand, the minimum degree condition yields

$$e_G(X, N_G(X,Y)) = \sum_{x\in X} d_G(x,Y) \geq |X|\alpha|Y|p.$$

On the other hand, with a set Y' such that $N_G(X,Y) \subset Y' \subset Y$, we get

$$e_G(X, N_G(X,Y)) \leq e_G(X,Y') \leq \left(1+\frac{\gamma}{2}\right)|X||Y'|p < \alpha|X||Y|p,$$

by Claim 3.6 (with $\gamma \mapsto \frac{\gamma}{2}$), when we let $|Y'| \geq \frac{C'\log n}{p}$ for a constant C' as required in Claim 3.6 and choose C sufficiently large such that $C' < \frac{\alpha}{1+\frac{\gamma}{2}}C$. These two inequalities lead to a contradiction, concluding the proof. □

Claim 3.11 *Let $n \in \mathbb{N}$ be sufficiently large, $\beta > 1$, and $p := p(n) \in (0,1]$. Consider a subgraph $G \subseteq \mathcal{G}(n,p)$. Then there exists a constant C such that for all sets $X \subseteq V(G)$ and $Y \subseteq V(G) \setminus X$ with $|X|, |Y| \geq \frac{C\log n}{p}$ a.a.s. the set $X_\beta := \{x \in X\colon d_G(x,Y) \geq \beta|Y|p\}$ must have size $|X_\beta| < \frac{C\log n}{p}$.*

Proof Assume without loss of generality that $1 < \beta < 2$, observing that $X_\beta \subseteq X_{\beta'}$ for $\beta' < \beta$. Let X, Y as above be arbitrary and assume, towards a contradiction, that there exists a set $X_\beta \subseteq X$ of size $|X_\beta| \geq \frac{C\log n}{p}$. The lower bound on the minimum degree of the vertex tuples in X_β implies

$$e_G(X_\beta, Y) = \sum_{x\in X_\beta} d_G(x,Y) \geq \beta|X_\beta||Y|p.$$

By Claim 3.6, with $0 < \gamma < \beta - 1 < 1$, we have

$$e_G(X_\beta, Y) \leq (1+\gamma)\,|X_\beta||Y|p < \beta|X_\beta||Y|p,$$

which leads to a contradiction and therewith proves the claim. □

Claim 3.12 *Let $n \in \mathbb{N}$ be sufficiently large, $p := p(n) \in (0,1]$, and G a subgraph of $\mathcal{G}(n,p)$. Moreover, let $\alpha > 0$ and $0 < \beta < \alpha$ be constants. Then for any small enough $\gamma > 0$ there exist constants C and $\kappa := \frac{(1-\gamma)\alpha-\beta}{1+\gamma-\beta} > 0$ such that for all sets $X, Y \subseteq V(G)$ with $|X|, |Y| \geq \frac{C\log n}{p}$ and $e_G(X,Y) \geq \alpha|X||Y|p$, the set $Y_\beta := \{y \in Y \mid d_G(y,X) \geq \beta|X|p\}$ a.a.s. must have size $|Y_\beta| \geq \kappa|Y|$.*

Proof Let X and Y be arbitrary as above. First suppose that $|Y_\beta| \geq \frac{C\log n}{2p}$, say. We then have $e_G(X, Y_\beta) \leq (1+\gamma)|X||Y_\beta|p$ by Claim 3.6, thus

$$e_G(X,Y) = e_G(X,Y_\beta) + e_G(X, Y \setminus Y_\beta) \leq (1+\gamma)|X||Y_\beta|p + \beta|Y \setminus Y_\beta||X|p.$$

Together with the lower bound on the number of edges between X and Y in the assumptions, this yields $|Y_\beta| \geq \frac{\alpha-\beta}{1+\gamma-\beta}|Y|$.

Now, suppose that $|Y_\beta| < \frac{C\log n}{2p}$. Let $Y_0 := \{y \in Y \mid d_G(y,X) \geq (1+\gamma)|X|p\}$. By Claim 3.11, $|Y_0| < \frac{C\log n}{\gamma p}$, say, by choosing C appropriately. Extend it to $|Y_0'| = \frac{C\log n}{\gamma p}$. Set $Y_\beta' := Y_\beta \setminus Y_0'$ and $Y' := Y \setminus Y_\beta \setminus Y_0'$. Observe that $e_G(X, Y_0') \leq (1+\gamma)|X||Y_0'|p$ by Claim 3.6, thus

$$e_G(X, Y \setminus Y_0') \geq \alpha|X||Y|p - (1+\gamma)|X||Y_0'|p \geq (1-\gamma)\alpha|X||Y|p,$$

say, for γ small enough. Moreover,

$$e_G(X, Y \setminus Y \setminus Y_0') = e_G(X, Y_\beta') + e_G(X, Y') \leq (1+\gamma)|X||Y_\beta'|p + \beta|X||Y'|p.$$

Therefore,

$$|Y_\beta| \geq |Y_\beta'| \geq \frac{(1-\gamma)\alpha - \beta}{1+\gamma-\beta}|Y|,$$

which concludes the proof. □

Claim 3.13 *Let $n \in \mathbb{N}$ be sufficiently large, $\alpha \in (0,1]$, and G a graph on n vertices. Consider subsets $X, Y \subseteq V(G)$ such that $e_G(X,Y) \geq \alpha \max\{d_G(x,Y) \colon x \in X\}$. Then, for any $0 < \beta < \alpha$ there exists a constant $\kappa > 0$ such that the set $X_\beta := \{x \in X \mid d_G(x,Y) \geq \beta \max\{d_G(x,Y) \colon x \in X\}\}$ has size $|X_\beta| \geq \frac{\alpha-\beta}{1-\beta}|X|$.*

Proof We have

$$\begin{aligned}\alpha \max\{d_G(x,Y) \colon x \in X\}|X| \leq |X_\beta| \max\{d_G(x,Y) \colon x \in X\} \\ + |X \setminus X_\beta|\beta \max\{d_G(x,Y) \colon x \in X\}.\end{aligned}$$

Thus $|X_\beta| \geq \alpha|X| - \beta|X \setminus X_\beta|$, which lets us conclude that $|X_\beta| \geq \kappa|X|$ for some $\kappa > 0$. □

8.2 Proof of Claim 5.5 (iii)

First note that it is enough to show that for any $c > 0$ and for all $\mathcal{H}_{j,i} \subseteq S_j$ with $|\mathcal{H}_{j,i}| \geq cs$ there exists a $v^* \in \mathcal{H}_{j,i}$ such that $\widehat{d}^+_{j,k}(v^*) \geq d_G(v, S_k) - \frac{C \log n}{p}$ By successively removing the vertices v^* until a set smaller than $c's$ for any $c' > 0$ is reached, indeed at least $|\mathcal{H}_{j,i}| - o(s)$ many such vertices v^* can be found.

Towards a contradiction, assume that $\widehat{d}^+_{j,k}(v) < d_G(v, S_k) - \frac{C \log n}{p}$ for all $v \in \mathcal{H}_{j,i}$. Let $N^0(v) := N_G(v, S_k) \setminus \widehat{N}^+_{j,k}(v)$ and note that by definition of these sets, $e_G(N^-_{j,i}(v), N^0(v)) = 0$ and

$$|N^0(v)| = d_G(v, S_k) - \widehat{d}^+_{j,k}(v) \geq \frac{C \log n}{p} \quad .$$

Claim 3.6 thus implies

$$\begin{aligned} e_{G_0}(N^-_{j,i}(v), N^0(v)) &\geq (1 - \gamma_1)|N^-_{j,i}(v)||N^0(v)|p \\ &= (1 - \gamma_1)\left(\frac{1}{3} + \frac{\alpha}{2}\right)|N^0(v)|sp^2 \\ &\geq \left(\frac{1}{3} + \frac{\alpha}{4}\right)|N^0(v)|sp^2. \end{aligned}$$

Moreover,

$$\begin{aligned} \max\{d_{G_0}(w, N^-_{j,i}(v)) \colon w \in N^0(v)\} &\leq \max\{d_{G_0}(w, N^-_{j,i}(v)) \colon w \in N^0(v)\} \\ &\leq \max\{d^2_{G_0}((v, w), S_i) \colon w \in N^0(v)\} \\ &\leq (1 + \gamma_2)sp^2 \end{aligned}$$

for any $0 < \gamma_2 < 1$ using a trivial bound on the maximum codegree by the Degree Concentration Lemma. Therefore, Claim 3.13 (with $X \mapsto N^0(v)$ and $Y \mapsto N^-_{j,i}(v)$) lets us deduce the existence of a subset $W_v \subseteq N^0(v)$ of size $|W_v| = \kappa |N_0(v)| \geq \kappa \frac{C \log n}{p}$ for some $\kappa > 0$ such that

$$d_{G_0}(w, N^-_{j,i}(v)) \geq \left(\frac{1}{3} + \frac{\alpha}{8}\right) sp^2$$

for all $w \in W_v$. Let $W_{\mathcal{H}} := \bigcup_{v \in \mathcal{H}_{j,i}} W_v$ and define the auxiliary bipartite graph F_1 between $\mathcal{H}_{j,i}$ and $W_{\mathcal{H}}$ with $V(F_1) := \mathcal{H}_{j,i} \cup W_{\mathcal{H}}$ and

$$E(F_1) := \{\{v, w\} \mid v \in \mathcal{H}_{j,i}, w \in W_v\}.$$

We have

$$e(F_1) = \sum_{v \in \mathcal{H}_{j,i}} d_{F_1}(v, W_{\mathcal{H}}) = \sum_{v \in \mathcal{H}_{j,i}} |W_v| \geq \kappa cs \frac{C \log n}{p}.$$

As $|W_{\mathcal{H}}| \leq |S_k| \leq s$, the average degree in F_1 of a vertex in $\mathcal{H}_{j,i}$ is lower bounded by $\kappa c \frac{C \log n}{p}$. Hence there must exist a $w^* \in W_{\mathcal{H}}$ with

$$d_F(u^*, \mathcal{H}_i) \geq \kappa c \frac{C \log n}{p}.$$

By definition of the auxiliary graph F_1 this means that w^* has at least $\kappa c \frac{C \log n}{p}$ many vertices $v \in \mathcal{H}_i$ for which $w^* \in W_v$. For each of those vertices $v \in N_{F_1}(w^*, \mathcal{H}_i)$ we have

$$d_{G_0}(w^*, N^0(v)) \geq \left(\frac{1}{3} + \frac{\alpha}{8}\right) sp^2.$$

Let $U_{w^*} := \bigcup_{v \in N_F(w^*, \mathcal{H}_{j,i})} N_{G_0}(w^*, N^-_{j,i}(v))$ and introduce a second auxiliary bipartite graph F_2, with $V(F_2) := N_{F_1}(w^*, \mathcal{H}_{j,i}) \cup S_i$ and

$$E(F_2) := \left\{ \{u, v\} \mid u \in N_{G_0}(w^*, N^-_{j,i}(v)) \right\}.$$

Then $d_{F_2}(v, S_i) \geq \left(\frac{1}{3} + \frac{\alpha}{8}\right) sp^2$ and thus Claim 3.10 yields

$$|U_{w^*}| = \left| N_{F_2}\left(N_{F_1}(w^*, \mathcal{H}_{j,i}), S_i\right) \right| \geq \left(\frac{1}{3} + \frac{\alpha}{10}\right) sp.$$

We claim that $d_G(w^*, U_{w^*}) = 0$. Towards a contradiction, suppose there is an edge from w^* to some vertex $u \in U_{w^*}$ in G. By definition of U_{w^*}, there must exist a $v \in N_F(w^*, \mathcal{H}_{j,i})$ such that $u \in N_{G_0}(w^*, N^-_{j,i}(v)) \subseteq N^-_{j,i}(v) \subseteq N_G(v, S_i)$. So there is an edge from u to v in G as well. But this means that $w^* \in N^2((u,v), S_k) \subseteq \widehat{N}^+_{j,k}(v)$, contradicting $w^* \in N^0(v)$. Moreover, observe that $d_{G_0}(w^*, U_{w^*}) \geq |U_{w^*}| \geq \left(\frac{1}{3} + \frac{\alpha}{8}\right) sp$.. Taken together,

$$d_H(w^*, U_{w^*}) = d_{G_0}(w^*, U_{w^*}) \geq \left(\frac{1}{3} + \frac{\alpha}{8}\right) sp,$$

which contradicts the upper bound on the maximum degree of H, and therefore concludes the proof.

Bibliography

[ABKP15] Peter Allen, Julia Böttcher, Yoshiharu Kohayakawa, and Yury Person. Tight hamilton cycles in random hypergraphs. *Random Structures & Algorithms*, 46(3):446–465, 2015.

[ACK+00] Noga Alon, Michael Capalbo, Yoshiharu Kohayakawa, Vojtech Rodl, Andrzej Rucinski, and Endre Szemerédi. Universality and tolerance. In *Foundations of Computer Science, 2000. Proceedings. 41st Annual Symposium on*, pages 14–21. IEEE, 2000.

[AS04] Noga Alon and Joel H. Spencer. *The probabilistic method*. John Wiley & Sons, 2004.

[BC76] John A. Bondy and Vasek Chvátal. A method in graph theory. *Discrete Mathematics*, 15(2):111–135, 1976.

[BCS11] József Balogh, Béla Csaba, and Wojciech Samotij. Local resilience of almost spanning trees in random graphs. *Random Structures & Algorithms*, 38(1-2):121–139, 2011.

[Big81] Norman L. Biggs. T. p. kirkman, mathematician. *Bulletin of the London Mathematical Society*, 13(2):97–120, 1981.

[BKT09] Julia Böttcher, Yoshiharu Kohayakawa, and Anusch Taraz. Almost spanning subgraphs of random graphs after adversarial edge removal. *Electronic Notes in Discrete Mathematics*, 35:335–340, 2009.

[BLS12] József Balogh, Choongbum Lee, and Wojciech Samotij. Corrádi and hajnal's theorem for sparse random graphs. *Combinatorics, Probability and Computing*, 21(1-2):23–55, 2012.

[Bol81] Béla Bollobás. Random graphs. *Combinatorics, Proceedings, Swansea 1981, London Math. Soc. Lecture Note Ser. 52*, 1981.

[Bol84] Béla Bollobás. The evolution of sparse graphs, in "graph theory and combinatorics proceedings. In *Cambridge Combinatorial Conference in Honour of Paul Erdos*, pages 335–357, 1984.

[Bon96] John A. Bondy. Basic graph theory: Paths and circuits. In *Handbook of combinatorics (vol. 1)*, pages 3–110. MIT Press, 1996.

[BSKS11a] Sonny Ben-Shimon, Michael Krivelevich, and Benny Sudakov. Local resilience and hamiltonicity maker–breaker games in random regular graphs. *Combinatorics, Probability and Computing*, 20(02):173–211, 2011.

[BSKS11b] Sonny Ben-Shimon, Michael Krivelevich, and Benny Sudakov. On the resilience of hamiltonicity and optimal packing of hamilton cycles in random graphs. *SIAM Journal on Discrete Mathematics*, 25(3):1176–1193, 2011.

[CDK11] Phong Châu, Louis DeBiasio, and Hal A. Kierstead. Pósa's conjecture for graphs of order at least 2× 108. *Random Structures & Algorithms*, 39(4):507–525, 2011.

[CE72] Vašek Chvátal and Paul Erdös. A note on hamiltonian circuits. *Discrete Mathematics*, 2(2):111–113, 1972.

[Chv72] Václav Chvátal. On hamilton's ideals. *Journal of Combinatorial Theory, Series B*, 12(2):163–168, 1972.

[Com74] Louis Comtet. Advanced combinatorics: The art of finite and infinite expansions, enlarged ed. reidel, dordrecht. *Mathematical Reviews (MathSciNet): MR460128 Zentralblatt MATH*, 283, 1974.

[Die00] Reinhard Diestel. *Graph theory*. Springer-Verlag Berlin and Heidelberg GmbH, 2000.

[Dir52] Gabriel A. Dirac. Some theorems on abstract graphs. *Proceedings of the London Mathematical Society*, 3(1):69–81, 1952.

[DKMS08] Domingos Dellamonica, Yoshiharu Kohayakawa, Martin Marciniszyn, and Angelika Steger. On the resilience of long cycles in random graphs. *Electron. J. Combin*, 15:R32, 2008.

[EGP91] Paul Erdős, András Gyárfás, and László Pyber. Vertex coverings by monochromatic cycles and trees. *Journal of Combinatorial Theory, Series B*, 51(1):90–95, 1991.

[ER60] Paul Erdös and Alfréd Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(17-61):43, 1960.

[Erd64] Paul Erdős. Extremal problems in graph theory. In *"Theory of graphs and its applications," proc. sympos. smolenice*. Citeseer, 1964.

[FK08] Alan Frieze and Michael Krivelevich. On two hamilton cycle problems in random graphs. *Israel Journal of Mathematics*, 166(1):221–234, 2008.

[FK15] Alan Frieze and Michał Karoński. *Introduction to random graphs*. Cambridge University Press, 2015.

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

[Gou03] Ronald J. Gould. Advances on the hamiltonian problem–a survey. *Graphs and Combinatorics*, 19(1):7–52, 2003.

[Hal35] Philip Hall. On representatives of subsets. *J. London Math. Soc*, 10(1):26–30, 1935.

[Hax95] Penny E. Haxell. A condition for matchability in hypergraphs. *Graphs and Combinatorics*, 11(3):245–248, 1995.

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

[JKV08] Anders Johansson, Jeff Kahn, and Van Vu. Factors in random graphs. *Random Structures & Algorithms*, 33(1):1–28, 2008.

[JŁR11] Svante Janson, Tomasz Łuczak, and Andrzej Rucinski. *Random graphs*, volume 45. John Wiley & Sons, 2011.

[Kar72] Richard M. Karp. Reducibility among combinatorial problems. In *Complexity of computer computations*, pages 85–103. Springer, 1972.

[KLS10] Michael Krivelevich, Choongbum Lee, and Benny Sudakov. Resilient pancyclicity of random and pseudorandom graphs. *SIAM Journal on Discrete Mathematics*, 24(1):1–16, 2010.

[KLS14] Michael Krivelevich, Choongbum Lee, and Benny Sudakov. Robust hamiltonicity of dirac graphs. *Transactions of the American Mathematical Society*, 366(6):3095–3130, 2014.

[KLS16] Michael Krivelevich, Choongbum Lee, and Benny Sudakov. Compatible hamilton cycles in random graphs. *Random Structures & Algorithms*, 2016.

[KO09] Daniela Kühn and Deryk Osthus. Embedding large subgraphs into dense graphs. *arXiv preprint arXiv:0901.3541*, 2009.

[KO12] Daniela Kühn and Deryk Osthus. On pósa's conjecture for random graphs. *SIAM Journal on Discrete Mathematics*, 26(3):1440–1457, 2012.

[Kri97] Michael Krivelevich. Triangle factors in random graphs. *Combinatorics, Probability and Computing*, 6(03):337–347, 1997.

[Kri10] Michael Krivelevich. Embedding spanning trees in random graphs. *SIAM Journal on Discrete Mathematics*, 24(4):1495–1500, 2010.

[KS83] János Komlós and Endre Szemerédi. Limit distribution for the existence of hamiltonian cycles in a random graph. *Discrete Mathematics*, 43(1):55–63, 1983.

[KSS96] János Komlós, Gábor N. Sárközy, and Endre Szemerédi. On the square of a hamiltonian cycle in dense graphs. *Random Structures & Algorithms*, 9(1-2):193–211, 1996.

[KSS97] János Komlós, Gábor N. Sárközy, and Endre Szemerédi. Blow-up lemma. *Combinatorica*, 17(1):109–123, 1997.

[KSS98a] János Komlós, Gábor N. Sárközy, and Endre Szemerédi. On the pósa-seymour conjecture. *Journal of Graph Theory*, 29(3):167–176, 1998.

[KSS98b] János Komlós, Gábor N. Sárközy, and Endre Szemerédi. Proof of the seymour conjecture for large graphs. *Annals of Combinatorics*, 2(1):43–60, 1998.

[LS12a] Choongbum Lee and Wojciech Samotij. Pancyclic subgraphs of random graphs. *Journal of Graph Theory*, 71(2):142–158, 2012.

[LS12b] Choongbum Lee and Benny Sudakov. Dirac's theorem for random graphs. *Random Structures & Algorithms*, 41(3):293–305, 2012.

[LSS10] Ian Levitt, Gábor N. Sárközy, and Endre Szemerédi. How to avoid using the regularity lemma: Pósa's conjecture revisited. *Discrete Mathematics*, 310(3):630–641, 2010.

[Mon14] Richard Montgomery. Embedding bounded degree spanning trees in random graphs. *arXiv preprint arXiv:1405.6559*, 2014.

[NŠ16a] Rajko Nenadov and Nemanja Škorić. Powers of cycles in random graphs and hypergraphs. *arXiv preprint arXiv:1601.04034*, 2016.

[NS16b] Andreas Noever and Angelika Steger. Local resilience for squares of almost spanning cycles in sparse random graphs. *arXiv preprint arXiv:1606.02958*, 2016.

[Ore60] Oystein Ore. Note on hamilton circuits. *The American Mathematical Monthly*, 67(1):55–55, 1960.

[RRS06] Vojtech Rödl, Andrzej Rucinski, and Endre Szemerédi. A dirac-type theorem for 3-uniform hypergraphs. *Combinatorics, Probability and Computing*, 15(1-2):229–251, 2006.

[RRS09] Vojtech Rödl, Andrzej Ruciński, and Endre Szemerédi. Perfect matchings in large uniform hypergraphs with large minimum collective degree. *Journal of Combinatorial Theory, Series A*, 116(3):613–636, 2009.

[Sey73] Paul Seymour. Problem section. In *Combinatorics: Proceedings of the British Combinatorial Conference*, pages 201–202, 1973.

[SV08a] Benny Sudakov and Van H. Vu. Local resilience of graphs. *Random Struct. Algorithms*, 33(4):409–433, 2008.

[SV08b] Benny Sudakov and Van H. Vu. Local resilience of graphs. *Random Struct. Algorithms*, 33(4):409–433, 2008.

[Sze75a] Endre Szemerédi. On sets of integers containing no k elements in arithmetic progression. *Acta Arith*, 27(199-245):2, 1975.

[Sze75b] Endre Szemerédi. Regular partitions of graphs. Technical report, DTIC Document, 1975.

[Tut56] William T. Tutte. A theorem on planar graphs. *Transactions of the American Mathematical Society*, 82(1):99–116, 1956.

[Wes01] Douglas B. West. *Introduction to graph theory*, volume 2. Prentice hall Upper Saddle River, 2001.

ETH

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Declaration of originality

The signed declaration of originality is a component of every semester paper, Bachelor's thesis, Master's thesis and any other degree paper undertaken during the course of studies, including the respective electronic versions.

Lecturers may also require a declaration of originality for other written papers compiled for their courses.

I hereby confirm that I am the sole author of the written work here enclosed and that I have compiled it in my own words. Parts excepted are corrections of form and content by the supervisor.

Title of work (in block letters):

Robustness of Pósa's conjecture

Authored by (in block letters):
For papers written by groups the names of all authors are required.

Name(s):	**First name(s):**
Fischer	Manuela

With my signature I confirm that
- I have committed none of the forms of plagiarism described in the 'Citation etiquette' information sheet.
- I have documented all methods, data and processes truthfully.
- I have not manipulated any data.
- I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for plagiarism.

Place, date	**Signature(s)**
Zurich, 19.10.2016	M.Fischer

For papers written by groups the names of all authors are required. Their signatures collectively guarantee the entire content of the written paper.