PACKINGS AND COVERINGS WITH HAMILTON CYCLES AND ON-LINE RAMSEY THEORY

by

JOHN ADAM LAPINSKAS

A thesis submitted to
The University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

School of Mathematics
The University of Birmingham
October 2014

Abstract

A major theme in modern graph theory is the exploration of maximal packings and minimal covers of graphs with subgraphs in some given family. We focus on packings and coverings with Hamilton cycles, and prove the following results in the area.

- Let $\varepsilon > 0$, and let G be a large graph on n vertices with minimum degree at least $(1/2+\varepsilon)n$. We give a tight lower bound on the size of a maximal packing of G with edge-disjoint Hamilton cycles.
- Let T be a strongly k-connected tournament. We give an almost tight lower bound on the size of a maximal packing of T with edge-disjoint Hamilton cycles.
- Let $\log^{117} n/n \leq p \leq 1 - n^{-1/8}$. We prove that $G_{n,p}$ may a.a.s. be covered by a set of $\lceil \Delta(G_{n,p})/2 \rceil$ Hamilton cycles, which is clearly best possible.

In addition, we consider some problems in on-line Ramsey theory. Let $\tilde{r}(G, H)$ denote the on-line Ramsey number of G and H. We conjecture the exact values of $\tilde{r}(P_k, P_\ell)$ for all $k \leq \ell$. We prove this conjecture for $k = 2$, prove it to within an additive error of 10 for $k = 3$, and prove an asymptotically tight lower bound for $k = 4$. We also determine $\tilde{r}(P_3, C_\ell)$ exactly for all ℓ.

Acknowledgements

First of all, I would like to thank my supervisors, Daniela Kühn and Deryk Osthus. They are the best supervisors I could possibly have hoped for. They are the ones who taught me how to do research, how to write a clear paper, and how to give an engaging talk. (And on a more prosaic note, they are the ones who provided me with travel funding!) It's a sentiment I've expressed before, but that doesn't make it any less true – they are the ones who made me a mathematician.

I would therefore also like to thank Imre Leader, who suggested I apply to them in the first place.

I would also like to thank the other people I have worked with closely: Ben Barber, Dan Hefetz, Allan Lo and Viresh Patel. They have all been a joy to collaborate with, and they have all taught me a great deal about the way mathematics is done.

I would also like to thank my friends, both within Birmingham and without. They have all been truly wonderful, helping me celebrate in the good times and vent in the bad. Any list of names would be far too short, so I will refrain. You all know who you are. Thank you.

Finally, I would like to thank my family. They have been constant and unwavering bastions of love and support, and without them none of this would ever have been possible.

Contents

Chapter 1

Introduction

1.1 Extremal graph theory

Extremal graph theory is fundamentally concerned with the dependence of graph properties on graph parameters. This is perhaps best explained through an example. Intuitively, we would expect it to be easier to find a triangle in a dense graph than a sparse one. We may therefore ask: how many edges may a graph G on n vertices have and still remain triangle-free? Certainly the complete bipartite graph $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$ is triangle-free, so the answer is at least $\lfloor n^2/4 \rfloor$. One of the foundational results in extremal graph theory, first proved by Mantel [65] in 1907, is as follows.

Theorem 1.1 *If G is a graph on n vertices with $e(G) > \lfloor n^2/4 \rfloor$, then G contains a triangle.*

Thus a triangle-free graph may contain up to $\lfloor n^2/4 \rfloor$ edges, but no more. Interestingly, we can say more. The following result is due to Erdős and Simonovits [29, 30, 85]. Here the symmetric difference of two sets X and Y is denoted by $X \triangle Y$.

Theorem 1.2 *For all $\varepsilon > 0$, there exists some $\delta > 0$ such that the following holds. If G is a triangle-free graph with $e(G) \geq (1/4 - \delta)n^2$, then $|E(G) \triangle E(K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil})| \leq \varepsilon n^2$.*

Thus in some sense, dense triangle-free graphs are the exception rather than the rule – any sufficiently dense triangle-free graph must be very close to $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$. This phenomenon is known as *stability*, and it is common in extremal problems.

In general, writing $\mathcal{G}$ for the class of all graphs, we may define a *graph parameter* to be a function $f : \mathcal{G} \to \mathbb{R}$ which respects isomorphisms. We may also define a *graph property* to be a subclass $\mathcal{P} \subseteq \mathcal{G}$ which is closed under isomorphisms. Thus edge count is a graph parameter, and the class of all triangle-free graphs is a graph property. (We will generally abuse our notation slightly, and refer simply to e.g. the property of being triangle-free.) Then much of extremal graph theory is concerned with the quantities $\max\{f(G) : G \in \mathcal{P}\}$ and $\min\{f(G) : G \in \mathcal{P}\}$. In other words, we ask the question: how large (or small) may $f(G)$ be in a graph satisfying $\mathcal{P}$? We may also ask: if $G \in \mathcal{P}$ with $f(G)$ almost maximal (or minimal), what restrictions are there on the structure of G?

Extremal graph theory is a major focus of modern combinatorics, and many of the results in this thesis are of an extremal nature.

1.2 Hamilton cycles

A *Hamilton cycle* of G is a cycle $C \subseteq G$ with $V(C) = V(G)$. We say a graph G is *Hamiltonian* if it contains a Hamilton cycle.

This concept was first introduced by Hamilton in 1857 in the context of a (succesfully marketed) solitaire game, in which players were asked to find Hamilton cycles in the graph of the dodecahedron. In fact, when viewed in a more general context, this simple game is incredibly difficult. In 1972, Karp [47] demonstrated that determining the Hamiltonicity of an arbitrary graph is an NP-complete problem. In fact, the problem remains NP-complete even under highly restrictive conditions on the input graph. For example, Akiyama, Nishizeki and Saito [4] proved the following.

Theorem 1.3 *Determining whether a cubic, planar, bipartite graph is Hamiltonian is an NP-complete problem.*

Despite this, there are a wealth of sufficient conditions for the existence of a Hamilton cycle which may be checked in polynomial time. For example, the following is a well-known extremal result of Dirac [28].

Theorem 1.4 *If G is a graph on $n \geq 3$ vertices, and $\delta(G) \geq n/2$, then G is Hamiltonian.*

If $\delta(G) < n/2$ then G may be a disconnected graph, or a bipartite graph on an odd number of vertices. In either case, it is immediate that G does not contain a Hamilton cycle. Thus Dirac's theorem says that the trivially necessary condition (in terms of minimum degree) is also sufficient. This phenomenon repeats itself in many other settings, as we will see over the course of the thesis, and forms an important part of the subject's allure.

It is important to note that while the study of Hamilton cycles has led to many beautiful results, it is by no means purely theoretical. Indeed, it has yielded algorithms applicable to many real-world problems. The most famous example is the Travelling Salesman Problem, which dates back to the 1930s. To solve the problem, given an edge-weighted complete graph, one must find a Hamilton cycle of minimal weight. We may imagine a travelling salesman who wishes to visit several cities and then return home, while travelling as little distance as possible – the cities correspond to vertices, and the edge weights correspond to distances between cities.

Unfortunately the Travelling Salesman Problem is NP-hard, but in practice an exact solution is rarely needed. There are many algorithms for finding solutions which, while not optimal, are "good enough for the task at hand". These algorithms have been applied not just to transportation problems, but also to problems in areas as wide-ranging as circuit design, scheduling, and X-ray crystallography. (See e.g. Matai, Singh and Mittal [68] for a more detailed survey.)

1.3 Packing Hamilton cycles in graphs of high minimum degree

Recall Dirac's theorem from the previous section. While Dirac's result is best possible in the sense that the minimum degree bound may not be weakened, Nash-Williams [72] proved in 1971 that it can be dramatically strengthened in another direction.

Theorem 1.5 *If G is a graph on n vertices, and $\delta(G) \geq n/2$, then G contains $\lfloor 5n/224 \rfloor$ edge-disjoint Hamilton cycles.*

In other words, the trivially necessary condition does not simply guarantee that G is Hamiltonian – it guarantees that a constant proportion of G's edge set can be decomposed into Hamilton cycles!

As beautiful as this result is, it raises a natural question: is the bound optimal? Can we hope to pack even more Hamilton cycles into every graph on n vertices with minimum degree $n/2$? Nash-Williams [71] conjectured that if $\delta(G) \geq n/2$ then G contains at least $\lfloor (n+2)/8 \rfloor$ edge-disjoint Hamilton cycles, and Babai gave a conjectured extremal example in the same paper. In Chapter 2 we prove this conjecture for all sufficiently large graphs which are not close to the extremal cases of Dirac's theorem.

Interestingly, the example Babai gave not only fails to contain $\lfloor (n+2)/8 \rfloor$ edge-disjoint Hamilton cycles, but also fails to contain any other spanning $2\lfloor (n+2)/8 \rfloor$-regular subgraph (also known as a $2\lfloor (n+2)/8 \rfloor$-*factor*). Thus our ability to guarantee r edge-disjoint Hamilton cycles in G is limited only by our ability to guarantee any $2r$-factor at all.

In Chapter 2, we prove that this is not a coincidence – the pattern extends to graphs with minimum degree at least αn for all fixed $\alpha > 1/2$. To make this formal, we define the following notation.

Definition 1.6 *Let G be a graph and let $n, \delta \in \mathbb{N}$. Then we define*

$$\mathrm{reg}_{\mathrm{even}}(G) := \max\{r \text{ even} : G \text{ contains an } r\text{-factor}\},$$
$$\mathrm{reg}_{\mathrm{even}}(n,\delta) := \min\{\mathrm{reg}_{\mathrm{even}}(H) : |H| = n,\ \delta(H) \geq \delta\}.$$

Thus $\mathrm{reg}_{\mathrm{even}}(n,\delta)$ is the degree of the densest even-regular spanning subgraph we can guarantee in G given only that $|G| = n$ and $\delta(G) \geq \delta$. Our main result in chapter 2 is as follows.

Theorem 1.7 *For all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that the following holds for all $n \geq N$. Let G be a graph on n vertices with $\delta := \delta(G) \geq n/2$. Then one of the following holds.*

(i) G contains at least $\mathrm{reg}_{\mathrm{even}}(n,\delta)/2$ edge-disjoint Hamilton cycles.

(ii) $\delta \leq (1/2 + \varepsilon)n$ and G is "close" to an extremal construction for Dirac's theorem.

In particular, this proves Nash-Williams' conjecture for large graphs in the non-extremal case. In subsequent work, Csaba, Kühn, Lo, Osthus and Treglown [25, 26, 57, 58] have proved the conjecture in the extremal case. Taken together with Theorem 1.7, this implies the following result.

Theorem 1.8 *There exists $N \in \mathbb{N}$ such that the following holds for all $n \geq N$. Let G be a graph on n vertices with minimum degree $\delta \geq n/2$. Then G contains $\mathrm{reg}_{\mathrm{even}}(n,\delta)/2$ edge-disjoint Hamilton cycles.*

Note that due to independent work by Christofides, Kühn and Osthus [21] and Hartke, Martin and Seacrest [42], the value of $\mathrm{reg}_{\mathrm{even}}(n,\delta)$ is known to within an absolute error of 1 for all $n, \delta \in \mathbb{N}$. Thus Theorem 1.8 is quantitative as well as qualitative.

Theorem 1.8 implies that our ability to guarantee r edge-disjoint Hamilton cycles in a graph of high minimum degree is limited only by our ability to guarantee any $2r$-factor at all. We conjecture that this idea extends further.

Conjecture 1.9 *Let G be a graph on n vertices with $\delta(G) \geq n/2$. Then G contains* $\mathrm{reg}_{\mathrm{even}}(G)/2$ *edge-disjoint Hamilton cycles.*

Thus if G's minimum degree is large enough to guarantee any Hamilton cycles at all, we believe that G contains r edge-disjoint Hamilton cycles if and only if G contains a $2r$-factor. This conjecture was already known for large graphs with $\delta(G) \geq (2-\sqrt{2})n$ due to work by Kühn and Osthus [59], and substantial progress has been made subsequently. The conjecture has been proved for large even-regular graphs by Csába, Kühn, Lo, Osthus and Treglown [25, 26, 57, 58], and an approximate version for all large graphs has been proved by Ferber, Krivelevich and Szabó [33].

1.4 Packing Hamilton cycles in highly connected tournaments

We now consider Hamilton cycles in another setting. A *tournament* is an orientation of a complete graph. Hamilton cycles are ubiquitous in a complete graph – indeed, any ordering of the vertices of K_n will yield a Hamilton cycle. We may therefore expect to find many (directed) Hamilton cycles in a tournament, but this is not the case. A tournament need not even be Hamiltonian! Indeed, consider the *transitive tournament* T on $[n]$, in which every edge ij with $i < j$ is oriented towards j. (Here $[n]$ is shorthand for $\{1, \ldots, n\}$.) Then a Hamilton cycle in T would contain an edge leaving n, which is clearly impossible. More generally, we make the following definition.

Definition 1.10 *A digraph G is* strongly connected *if for all distinct vertices $u, v \in V(G)$, there exist paths in G from u to v and from v to u.*

Since a Hamilton cycle is strongly connected, it is immediate that any Hamiltonian tournament must also be strongly connected. Fortunately, a classical result of Camion [20] tells us that this is the only obstacle.

Theorem 1.11 *A tournament T is Hamiltonian if and only if it is strongly connected.*

But what happens if we try to find multiple edge-disjoint Hamilton cycles in a tournament? Certainly a strongly connected tournament need not contain two edge-disjoint Hamilton cycles. Indeed, by taking the transitive tournament on $[n]$ and reversing the edge between 1 and n, we obtain a strongly connected tournament with a unique Hamilton cycle. We might, however, hope that a "more strongly connected" tournament may contain more Hamilton cycles. There are two natural ways of formalising this requirement, both by analogy with connectivity in undirected graphs.

Definition 1.12 *Let $k \in \mathbb{N}$, and let G be a digraph with $|G| \geq k+1$. Then G is* strongly k-connected *if $G - U$ is strongly connected for all $U \subseteq V(G)$ with $|U| \leq k-1$. Alternatively, G is* strongly k-edge-connected *if $G - F$ is strongly connected for all $F \subseteq E(G)$ with $|F| \leq k-1$.*

Strong edge-connectivity turns out to be of little use to us – Thomassen [89] proved in 1982 that there exist tournaments with arbitrarily high edge connectivity which fail to contain two edge-disjoint Hamilton cycles. However, strong k-connectivity is a stronger notion – indeed, any strongly k-connected digraph is also strongly k-edge-connected – and so we should not lose hope. In fact, Thomassen conjectured the following.

Conjecture 1.13 *For all $k \in \mathbb{N}$, there exists $f(k) \in \mathbb{N}$ such that any strongly $f(k)$-connected tournament must contain k edge-disjoint Hamilton cycles.*

Unfortunately, finding two edge-disjoint Hamilton cycles is dramatically harder than finding a single Hamilton cycle. Indeed, removing a Hamilton cycle from a tournament destroys n edges, an effect which seems to dwarf any constant connectivity. Consequently, until our research in the area, it was an open problem to prove even the existence of $f(2)$. In Chapter 3, we prove Thomassen's conjecture in its entirety.

Theorem 1.14 *There exists $C > 0$ such that for all $k \in \mathbb{N}$, any strongly $Ck^2 \log^2 k$-connected tournament must contain k edge-disjoint Hamilton cycles.*

Thus we prove not only that f exists, but also that $f(k) = O(k^2 \log^2 k)$ – surprisingly small. In fact, our value for f is almost best possible – we provide a construction showing that $f(k) = \Omega(k^2)$. We therefore make the natural conjecture that in fact $f(k) = \Theta(k^2)$. This conjecture has since been proved by Pokrovskiy [75]. Interestingly, just as in the setting of graphs with high minimum degree, our conjectured extremal construction fails to contain not just k edge-disjoint Hamilton cycles but any k-regular spanning subgraph.

As part of our proof, we prove another result which is of independent interest. The following theorem of Menger [69] is well-known.

Theorem 1.15 *Let G be a strongly k-connected digraph, and let $X = \{x_1, \ldots, x_k\}$ and $Y = \{y_1, \ldots, y_k\}$ be disjoint subsets of $V(G)$. Then there exist vertex-disjoint paths $P_1, \ldots, P_k \subseteq G$ and a permutation $\sigma : [k] \to [k]$ such that P_i is a path from x_i to $y_{\sigma(i)}$.*

In other words, in a strongly k-connected digraph, we can join up two vertex subsets of size k with vertex-disjoint paths, but with no control over the endpoints of each path. In our proof, however, we need to join pairs of sets in this way with full control over the endpoints. For this, we require the following property.

Definition 1.16 *Let $k \in \mathbb{N}$, and let G be a digraph with $|G| \geq 2k$. Then G is k-*linked *if whenever $X = \{x_1, \ldots, x_k\}$ and $Y = \{y_1, \ldots, y_k\}$ are disjoint subsets of $V(G)$, there exist vertex-disjoint paths $P_1, \ldots, P_k$ such that P_i is a path from x_i to y_i.*

Fortunately, as Thomassen [90] proved in 1984, for all $k \in \mathbb{N}$ there exists $g(k) \in \mathbb{N}$ such that any strongly $g(k)$-connected tournament is k-linked. (This is not the case for general digraphs, as proved by Thomassen [91].) We are therefore able to use linkedness in our proof of Theorem 1.14. However, Thomassen's upper bound on $g(k)$ is exponential in k. In Chapter 3, we improve this substantially.

Theorem 1.17 *There exists $C > 0$ such that for all $k \in \mathbb{N}$, any strongly $Ck \log k$-connected tournament is also k-linked.*

Thus we prove that $g(k) = O(k \log k)$. This is almost best possible, since k-linkedness trivially implies strong k-connectedness, and we conjecture that in fact $g(k) = \Theta(k)$. This conjecture has since been proved by Pokrovskiy [76].

Our method of proof is novel, and involves finding powerful "linking structures" in our tournament. Pokrovskiy [75, 76] used similar methods to prove our two conjectures. Similar methods were also used in subsequent work by by Kühn, Osthus and Townsend [62] to prove another conjecture of Thomassen (see [82]) – that if T is a sufficiently strongly connected tournament, then $V(T)$ can be partitioned into t vertex-disjoint strongly k-connected subtournaments.

1.5 Random graphs and the Erdös-Rényi-Gilbert model

Suppose we are studying a graph property $\mathcal{P}$. It is very natural to ask not only when a graph G satisfies $\mathcal{P}$, but also whether "most" graphs satisfy $\mathcal{P}$. We may make this idea precise with the following definition.

Definition 1.18 *We define $G_{n,1/2}$ to be a uniformly random (labelled) graph on n vertices. We say $G_{n,1/2}$ satisfies some property $\mathcal{P}$* asymptotically almost surely *(or* a.a.s.*) if*

$$\mathbb{P}(G_{n,1/2} \text{ satisfies } \mathcal{P}) \to 1 \text{ as } n \to \infty.$$

It is difficult to work with $G_{n,1/2}$ directly, but fortunately there is an easier way – we may model $G_{n,1/2}$ as a random graph on n vertices in which every possible edge is included independently with probability $1/2$. This independence between edges then gives us access to a wide array of powerful techniques from probability theory.

Unfortunately, $G_{n,1/2}$ often fails to capture the aspects of a problem we are most interested in. For example, we have seen that "most" dense graphs contain a triangle, in the sense that if $e(G) \approx n^2/4$ and G is triangle-free then G must be close to $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$. In fact, far more is true. Even if we restrict our attention to graphs with $e(G) = m(n)$ for some function $m : \mathbb{N} \to \mathbb{N}$, "most" such graphs will contain a triangle as long as $m(n) = \omega(n)$ as $n \to \infty$! This phenomenon is interesting, and similar behaviour occurs in many other extremal problems, but study of $G_{n,1/2}$ sheds little light on it. Indeed, $G_{n,1/2}$ is a.a.s. approximately $(n/2)$-regular, so it seems unlikely to be the best tool to study graphs with average degree $\Theta(\log \log n)$ (for example).

We would therefore like to define a random graph which allows us to talk about the properties of "most" sparse graphs, in the same way that $G_{n,1/2}$ allows us to talk about the properties of "most" arbitrary graphs. Perhaps the most natural approach, taken by Erdős and Rényi [32] in 1960, is the following model.

Definition 1.19 *Let $m : \mathbb{N} \to \mathbb{N}$ be a function. We define $G(n, m)$ to be a uniformly random (labelled) graph on n vertices with $m(n)$ edges. We say $G(n, m)$ satisfies some property $\mathcal{P}$* a.a.s. *if*

$$\mathbb{P}(G(n, m) \text{ satisfies } \mathcal{P}) \to 1 \text{ as } n \to \infty.$$

It is then indeed the case that $G(n, m)$ contains a triangle a.a.s. whenever $m = \omega(n)$. However, working with $G(n, m)$ is far less pleasant than working with $G_{n,1/2}$, since edges are no longer independent. In 1959, working independently from Erdös and Rényi, Gilbert [38] introduced the following alternative model – now widely known as the *Erdős-Rényi-Gilbert model.*

Definition 1.20 *Let $p : \mathbb{N} \to [0,1]$ be a function. We define $G_{n,p}$ to be a random (labelled) graph on n vertices in which each possible edge is included independently with probability $p(n)$. We say $G_{n,p}$ satisfies some property $\mathcal{P}$* a.a.s. *if*

$$\mathbb{P}(G_{n,p} \text{ satisfies } \mathcal{P}) \to 1 \text{ as } n \to \infty.$$

Intuitively, it is easy to see that $G_{n,p}$ models sparse graphs when p is small, and dense graphs when p is large. Moreover, since we have independence between edges, we are free to use the same powerful techniques that we used to study $G_{n,1/2}$.

It may therefore seem that we are left with a choice – we can work with $G(n,m)$, which perfectly captures the idea of a property holding for "most" graphs of a given density, or we can work with $G_{n,p}$, which is less exact but far easier to work with. Happily, this is not the case. There is a very strong correspondence between the two models, and in most circumstances a property will hold a.a.s. in $G_{n,p}$ iff it holds a.a.s. in $G(n, p\binom{n}{2})$. (For more details, see e.g. Janson, Łuczak and Ruciński [46].) Thus $G_{n,p}$ boasts both excellent modelling power and relative approachability, a combination which has cemented its study as a major focus of modern graph theory.

1.6 Covering random graphs with Hamilton cycles

We now turn to the study of Hamilton cycles in $G_{n,p}$. The most obvious question we can ask is: for which values of p is $G_{n,p}$ a.a.s. Hamiltonian? It is relatively easy to show that when $p = (\log n + 2\log\log n + O(1))/n$, we have $\delta(G_{n,p}) \le 1$ with probability bounded away from zero. Thus $G_{n,p}$ clearly cannot be a.a.s. Hamiltonian unless $p = (\log n + 2\log\log n + \omega(1))/n$. In fact, Ajtai, Komlós and Szemerédi [3] and Bollobás [14] proved that this is the only obstacle.

Theorem 1.21 *Suppose $p = (\log n + 2\log\log n + \omega(1))/n$. Then $G_{n,p}$ is a.a.s. Hamilto-*

nian.

In fact, they proved something substantially stronger. We may visualise the uniformly random graph $G(n, m)$ as a random process, obtaining $G(n, m+1)$ from $G(n, m)$ by adding an edge chosen uniformly at random. In this setting, a.a.s. the very same edge which raises the minimum degree of $G(n, m)$ to 2 will also introduce a Hamilton cycle! Such results are known as *hitting time results*, and are well-studied in the literature.

As we might expect, the more general problem of packing Hamilton cycles in $G_{n,p}$ has also been well-studied. Certainly $G_{n,p}$ cannot contain more than $\lfloor \delta(G_{n,p})/2 \rfloor$ edge-disjoint Hamilton cycles, since the cycles must be disjoint at any vertex of minimal degree. However, this turns out to be the only restriction. The following result is due to Knox, Krivelevich, Kühn, Osthus and Samotij [51, 54, 59].

Theorem 1.22 *Let $p : \mathbb{N} \to [0, 1]$ be an arbitrary function. Then a.a.s. $G_{n,p}$ contains $\lfloor \delta(G_{n,p})/2 \rfloor$ edge-disjoint Hamilton cycles.*

As an aside, note that this trivially implies that we can (a.a.s.) guarantee r edge-disjoint Hamilton cycles in $G_{n,p}$ as soon as we can (a.a.s.) guarantee any $2r$-factor at all – just as with graphs of high minimum degree and highly connected tournaments. Also note that when $p = \omega(\log n/n)$, we a.a.s. have $\delta(G_{n,p}) \sim \Delta(G_{n,p})$ and so this result implies that $G_{n,p}$ may be almost entirely decomposed into edge-disjoint Hamilton cycles!

It is natural to consider the dual problem of covering the edge set of $G_{n,p}$ with as few Hamilton cycles as possible. In other words, we seek a small set of Hamilton cycles such that every edge of $G_{n,p}$ is contained in at least one Hamilton cycle. We may see a large packing of Hamilton cycles into $G_{n,p}$ as approximating a decomposition of G "from below", and a small covering as approximating a decomposition of G "from above".

It is clear that we cannot hope to cover $G_{n,p}$ with fewer than $\lceil \Delta(G_{n,p})/2 \rceil$ Hamilton cycles, since the cycles must cover all edges incident to any vertex of maximal degree.

Unlike with packing, however, there are other barriers. The first is simple – we cannot cover $G_{n,p}$ with Hamilton cycles at all if it doesn't contain any Hamilton cycles! The second is less obvious. Suppose p is very large, so that with probability bounded away from zero $G_{n,p}$ is the complete graph with one edge removed. As we note in Chapter 4, a simple parity argument implies that this graph cannot be covered with $(n-1)/2$ Hamilton cycles when n is odd. We therefore do not expect to be able to cover $G_{n,p}$ with $\lceil \Delta(G_{n,p})/2 \rceil$ Hamilton cycles when $p = (\log n + 2 \log\log n + O(1))/n$, or when $p = 1 - \Omega(n^{-2})$.

Interestingly, these barriers may be the only ones. Glebov, Krivelevich and Sudakov [39] have proved that when $p = \Omega(n^{-1+\varepsilon})$ for any fixed $\varepsilon > 0$, $G_{n,p}$ can be covered with $(1 + o(1))\Delta(G_{n,p})/2$ Hamilton cycles. In Chapter 4, we improve this to an exact result with a sharper lower bound on p.

Theorem 1.23 *Suppose $G \sim G_{n,p}$. If $\log^{117} n/n \le p \le 1 - n^{-1/8}$, then a.a.s. the edges of $G_{n,p}$ can be covered with $\lceil \Delta(G_{n,p})/2 \rceil$ Hamilton cycles. If $p > 1 - n^{-1/8}$, then a.a.s. the edges of $G_{n,p}$ can be covered with $(1 + o(1))\Delta(G_{n,p})/2$ Hamilton cycles.*

It would be very interesting to know the exact range of p for which this result holds, and whether it can be generalised to a hitting time result in the same way that Ajtai, Komlós and Szemerédi's original result can be.

1.7 Ramsey theory

Ramsey theory may be thought of as the study of the inevitable appearance of order in large structures. As a more concrete example, consider the following simple question. If we colour the edges of K_n red and blue in an arbitrary fashion, must there always be a monochromatic triangle? We may think of the vertices of K_n as guests at a party, where an edge is coloured blue if two guests know each other and red if they do not. Then the question becomes: in a party with n guests, must there always be either three guests who are mutual acquaintances or three guests who have never met? It is relatively easy to show that the answer is yes, as long as $n \geq 6$. In 1928, Ramsey [81] proved the following substantially stronger result.

Theorem 1.24 *For any fixed graphs G and H, there exists $r(G, H) \in \mathbb{N}$ such that the following holds for all $n \geq r(G, H)$. Suppose the edges of K_n are coloured red and blue. Then K_n contains either a red copy of G or a blue copy of H.*

We call $r(G, H)$ the *Ramsey number* of G and H. Similar phenomena often arise in other settings, and also fall under the umbrella of Ramsey theory. For example, Ramsey's original result applied to uniform hypergraphs as well as graphs. We may also consider Ramsey theory on the integers – for example, the following foundational result is due to van der Waerden [94].

Theorem 1.25 *For any $k \in \mathbb{N}$, there exists $W(k) \in \mathbb{N}$ such that the following holds for all $n \geq W(k)$. Suppose the elements of $[n]$ are coloured red and blue. Then $[n]$ contains a monochromatic arithmetic progression of length k.*

This result has been substantially generalised. We may view an arithmetic progression of length k as a solution to a set of simultaneous linear equations. In 1933, Rado [79] gave a necessary and sufficient condition for an arbitrary set of simultaneous linear equations to satisfy a similar result. For the rest of this thesis, however, we shall be concerned only with Ramsey theory on graphs.

A central research problem in graph Ramsey theory is the exact and approximate determination of Ramsey numbers. The problem is famously difficult, and Burr [19] proved in 1984 that it is NP-hard. (Indeed, given a graph G on n vertices, one may determine the chromatic number of G from the value of $r(G, P_{n^3})$ in time polynomial in n.) Erdős once famously described the difficulty as follows. "Suppose aliens invade the earth and threaten to obliterate it in a year's time unless human beings can find [$r(K_5, K_5)$]. We could marshal the world's best minds and fastest computers, and within a year we could probably calculate the value. If the aliens demanded $[r(K_6, K_6)]$, however, we would have no choice but to launch a preemptive attack." The best known general bounds on $r(K_k, K_k)$ are as follows, due to Spencer [86] and Conlon [23].

Theorem 1.26 *There exists $C > 0$ such that for all $k \in \mathbb{N}$,*

$$\left(\frac{\sqrt{2}}{e} + o(1)\right)(k+1)2^{(k+1)/2} \le r(K_{k+1}, K_{k+1}) \le k^{-C\frac{\log k}{\log\log k}}\binom{2k}{k}$$

Note that there exists $C' > 0$ such that

$$k^{-C\frac{\log k}{\log\log k}}\binom{2k}{k} \ge 4^{k-C'\log^2 k},$$

so these bounds are still very far apart. However, for other graphs G and H, the problem often becomes easier – $r(G, H)$ has been determined exactly in many cases. See Radziszowski [80] for a detailed dynamic survey.

1.8 On-line Ramsey numbers of paths and cycles

In 1983, Beck [11] considered the following question. Suppose we are given a colouring of the edges of K_n, where n is substantially larger than $r(K_k, K_k)$, and we wish to find a monochromatic K_k (as opposed to proving that one exists). How many edges must we examine in order to do so, in the worst case? We define the *kth on-line Ramsey number* $\tilde{r}(K_k, K_k)$ to be the answer to this question. Beck was able to prove that $\tilde{r}(K_k, K_k) \ge 2^{k/2}$, ruling out the existence of a deterministic algorithm to find a monochromatic K_k in time polynomial in k.

Alternatively, as emphasised in Beck's seminal paper [12] on the subject, we may define $\tilde{r}(K_k, K_k)$ in terms of a combinatorial game. We consider a game played by two players, Builder and Painter, on the infinite clique $K_{\mathbb{N}}$. In each round of the game Builder chooses an edge of $K_{\mathbb{N}}$, and Painter colours it red or blue. Builder wins when a monochromatic clique has been formed. Builder wishes to win in as little time as possible, and Painter wishes to draw the game out for as long as possible. We may then define $\tilde{r}(K_k, K_k)$ to be the duration of the game, assuming that both Builder and Painter play optimally. Note that $\tilde{r}(K_k, K_k)$ is well-defined, since Builder can always win by uncovering all the edges of a clique on $r(K_k, K_k)$ vertices. We may likewise define $\tilde{r}(G, H)$ for general graphs G and H by requiring Builder to construct either a red copy of G or a blue copy of H.

As Alon observed (see [12]), Beck's result has an easy proof in this setting. Indeed, for any graphs G and H, by definition there exists an edge-colouring of $K_{r(G,H)-1}$ containing no red copy of G and no blue copy of H. If Painter simply copies this colouring, it is immediate that she cannot lose the game until Builder has uncovered edges incident to $r(G, H)$ distinct vertices. This strategy allows Painter to survive for at least $\lfloor (r(G, H) - 1)/2 \rfloor$ rounds, so $\tilde{r}(G, H) \geq r(G, H)/2$. In particular, $\tilde{r}(K_k, K_k) \geq 2^{k/2}$, and so Beck's result follows. In general this setting is substantially easier to work with, and has been widely adopted in the literature.

As with classical Ramsey numbers, determining the exact values of on-line Ramsey numbers is an extremely difficult problem. In fact, determining on-line Ramsey numbers seems to be even harder than determining classical Ramsey numbers. For example, while the exact value of $r(P_{k+1}, P_{\ell+1})$ was determined for all $k, \ell \in \mathbb{N}$ by Gerencsér and Gyárfás in 1967, determining the values of $\tilde{r}(P_{k+1}, P_{\ell+1})$ remains an unsolved problem despite repeated attempts. Until our work in the area, only very loose general bounds (and a few exact values for small k and ℓ) were known.

In Chapter 5, we conjecture the following exact values for $\tilde{r}(P_{k+1}, P_{\ell+1})$.

Conjecture 1.27 *For all $k, \ell \in \mathbb{N}$ with $k \leq \ell$, we have*

$$\tilde{r}(P_{k+1}, P_{\ell+1}) = \begin{cases} \ell & \text{if } k = 1, \\ \lceil 5\ell/4 \rceil & \text{if } k = 2, \\ \lceil (7\ell+2)/5 \rceil & \text{if } k = 3, \\ \lceil 3\ell/2 \rceil + k - 3 & \text{if } k \geq 4. \end{cases}$$

We prove our conjecture exactly for $k = 2$ and to within an additive error of 10 for $k = 3$. We also prove an asymptotically tight lower bound for all fixed k as $\ell \to \infty$. Our error is linear in k, and our lower bound improves on the previous state of the art unless k is very close to ℓ. The general problem of determining $\tilde{r}(P_{k+1}, P_{\ell+1})$ for all $k, \ell \in \mathbb{N}$ remains open, however.

We also prove that $\tilde{r}(P_3, C_\ell) = \tilde{r}(P_3, P_{\ell+1})$ for all $\ell \geq 5$. This is somewhat counter-intuitive – it seems as though it should be easier for Builder to extend a path than to close a cycle. It would be interesting to know for which other graphs G we have $\tilde{r}(G, P_{\ell+1}) = \tilde{r}(G, C_\ell)$ for sufficiently large ℓ.

Surprisingly, all our lower bounds follow from considering so-called $\mathcal{F}$*-blocking* strategies for Painter, in which she colours each edge (wlog) red unless doing so would create a red graph in some forbidden family $\mathcal{F}$. Thus an $\mathcal{F}$-blocking strategy is essentially an enlightened greedy strategy, in which Painter is allowed to avoid "dangerous" graphs other than G. We believe this is the first non-trivial case in which such strategies have given tight lower bounds. Note that they are certainly not optimal in general – for example, $\tilde{r}(K_3, K_3) = 8$ (see Kurek and Ruciński [64]) but it is easy to show via case analysis that no $\mathcal{F}$-blocking strategy will allow Painter to survive longer than 7 rounds. However, it would be fascinating to know for which graphs an optimal $\mathcal{F}$-blocking strategy does exist.

1.9 Attribution

Chapter 2 is joint work with Daniela Kühn and Deryk Osthus, and has been accepted for publication in Combinatorica [55]. Chapter 3 is joint work with Daniela Kühn, Deryk Osthus and Viresh Patel, and has been accepted for publication in the Proceedings of the London Mathematical Society [56]. Chapter 4 is joint work with Dan Hefetz, Daniela Kühn and Deryk Osthus, and has been accepted for publication in Combinatorica [45]. Chapter 5 is joint work with Joanna Cyman, Tomasz Dzido and Allan Lo, and has been submitted for publication [27].

Chapter 2

Optimal packings of Hamilton cycles in graphs of high minimum degree

2.1 Introduction

Dirac's theorem [28] states that any graph on $n \geq 3$ vertices with minimum degree at least $n/2$ contains a Hamilton cycle. This degree condition is best possible. Surprisingly, though, the assertion of Dirac's theorem can be strengthened considerably: Nash-Williams [72] proved that the conditions of Dirac's theorem actually guarantee linearly many edge-disjoint Hamilton cycles.

Theorem 2.1.1 *Every graph on n vertices with minimum degree at least $n/2$ contains at least $\lfloor 5n/224 \rfloor$ edge-disjoint Hamilton cycles.*

Nash-Williams [71] initially conjectured that such a graph must contain at least $\lfloor n/4 \rfloor$ edge-disjoint Hamilton cycles, which would clearly be best possible. However, Babai observed that this trivial bound is very far from the truth (see [71]). Indeed, the following construction (which is based on Babai's argument) gives a graph G which contains at most

$\lfloor (n+2)/8 \rfloor$ edge-disjoint Hamilton cycles. The graph G consists of one empty vertex class A of size $2m$, one vertex class B of size $2m+2$ containing a perfect matching and no other edges, and all possible edges between A and B. Thus G has order $n = 4m+2$ and minimum degree $2m+1$. Any Hamilton cycle in G must contain at least two edges of the perfect matching in B, so G contains at most $\lfloor (m+1)/2 \rfloor$ edge-disjoint Hamilton cycles.

The above question of Nash-Williams naturally extends to graphs of higher minimum degree: suppose that $n/2 \leq \delta \leq n-1$. *How many edge-disjoint Hamilton cycles can one guarantee in a graph G on n vertices with minimum degree δ?*

Clearly, as δ increases, one expects to find more edge-disjoint Hamilton cycles. However, the above construction shows that the trivial bound of $\lfloor \delta/2 \rfloor$ cannot always be attained. A less trivial bound is provided by the largest even-regular spanning subgraph in G. More precisely, let $\mathrm{reg}_{\mathrm{even}}(G)$ be the largest degree of an even-regular spanning subgraph of G. Then let

$$\mathrm{reg}_{\mathrm{even}}(n, \delta) := \min\{\mathrm{reg}_{\mathrm{even}}(G) : |G| = n,\ \delta(G) = \delta\}.$$

Clearly, in general we cannot guarantee more than $\mathrm{reg}_{\mathrm{even}}(n, \delta)/2$ edge-disjoint Hamilton cycles in a graph of order n and minimum degree δ. In fact, we conjecture this bound can always be attained.

Conjecture 2.1.2 *Suppose G is a graph on n vertices with minimum degree $\delta \geq n/2$. Then G contains at least* $\mathrm{reg}_{\mathrm{even}}(n, \delta)/2$ *edge-disjoint Hamilton cycles.*

Our main result confirms this conjecture exactly, as long as n is large and δ is slightly larger than $n/2$.

Theorem 2.1.3 *For every $\varepsilon > 0$, there exists an integer $n_0 = n_0(\varepsilon)$ such that every graph G on $n \geq n_0$ vertices with $\delta(G) \geq (1/2+\varepsilon)n$ contains at least* $\mathrm{reg}_{\mathrm{even}}(n, \delta(G))/2$ *edge-disjoint Hamilton cycles.*

In fact, we even show that if G is not close to the extremal example, then G contains significantly more than the required number of edge-disjoint Hamilton cycles (see Lemma 2.5.15). Our proof of Theorem 2.1.3 is based on a recent result (Theorem 2.3.2) of Kühn and Osthus [61, 59], which states that every "robustly expanding" regular (di)graph has a Hamilton decomposition. In [59], a straightforward argument was already used to derive Conjecture 2.1.2 for $\delta \geq (2 - \sqrt{2} + \varepsilon)n$ (see Section 2.3.2). Our extension of this result to $\delta \geq (1/2 + \varepsilon)n$ involves new ideas.

Subsequently, Csaba, Kühn, Lo, Osthus and Treglown [58, 25, 26, 57] have proved Conjecture 2.1.2 for large n, by solving the case when δ is allowed to be close to $n/2$. The proof relies on Theorem 2.1.3 and Theorem 2.1.6. (The latter provides a stability result when δ is close to $n/2$.)

Earlier, Christofides, Kühn and Osthus [21] used the regularity lemma to prove an approximate version of Theorem 2.1.3. Hartke and Seacrest [43] were able improve this result while avoiding the use of the regularity lemma (but still with the same restriction on δ). This enabled them to omit the condition that G has to be very large. They also gave significantly better error bounds.

Accurate bounds on $\mathrm{reg}_{\mathrm{even}}(n, \delta)$ are known. Note that the complete bipartite graph whose vertex classes are almost equal shows that $\mathrm{reg}_{\mathrm{even}}(n, \delta) = 0$ for $\delta < n/2$. Katerinis [48] considered the case when $\delta = n/2$. His result was independently generalised to larger values of δ in [21] (see [59] for a summarised version) and by Hartke, Martin and Seacrest [42]. The following bounds are from [42].

Theorem 2.1.4 *Suppose that $n, \delta \in \mathbb{N}$ and $n/2 \leq \delta < n$. Then*

$$\frac{\delta + \sqrt{n(2\delta - n) + 8}}{2} - \varepsilon \leq \mathrm{reg}_{\mathrm{even}}(n, \delta) \leq \frac{\delta + \sqrt{n(2\delta - n)}}{2} + \frac{4}{\sqrt{n(2\delta - n)} + 4}. \tag{2.1.5}$$

where $0 < \varepsilon \leq 2$ is chosen to make the left hand side of (2.1.5) an even integer.

Note that (2.1.5) always yields at most two possible values for $\mathrm{reg}_{\mathrm{even}}(n, \delta)$ and even determines it exactly for many values of the parameters n and δ. For example, (2.1.5) determines $\mathrm{reg}_{\mathrm{even}}(n, n/2)$ (e.g. in the case when n is divisible by 8 it is $n/4$). The bounds in [21] also give at most two possible values. The lower bound in (2.1.5) is based on Tutte's factor theorem [92]. The upper bound is obtained by a natural generalization of Babai's construction (see Section 2.3.1 for a description).

Our second result concerns the case of Conjecture 2.1.2 where we allow δ to be close to $n/2$. In this case, we obtain the following 'stability result': if $\delta(G) = (1/2 + o(1))n$, then Conjecture 2.1.2 holds for large n as long as G has suitable expansion properties. In this case, we even obtain significantly more than the required number of edge-disjoint Hamilton cycles again. These expansion properties fail only when G is very close to the extremal examples for Dirac's theorem.

Theorem 2.1.6 *For every $0 < \eta < 1/8$, there exist $\varepsilon > 0$ and an integer n_0 such that every graph G on $n \geq n_0$ vertices with $(1/2 - \varepsilon)n \leq \delta(G) \leq (1/2 + \varepsilon)n$ satisfies one of the following:*

(i) There exists $A \subseteq V(G)$ with $|A| = \lfloor n/2 \rfloor$ and such that $e(A) \leq \eta n^2$.

(ii) There exists $A \subseteq V(G)$ with $|A| = \lfloor n/2 \rfloor$ and such that $e(A, \overline{A}) \leq \eta n^2$.

(iii) G contains at least $\max\{\mathrm{reg}_{\mathrm{even}}(n, \delta(G))/2, n/8\} + \varepsilon n$ *edge-disjoint Hamilton cycles.*

Note that if G satisfies (i) then $e(A, \overline{A})$ must be roughly $n^2/4$, i.e. G is close to $K_{n/2,n/2}$ with possibly some edges added to one of the vertex classes. If G satisfies (ii), then both $e(A)$ and $e(\overline{A})$ must be roughly $n^2/8$, i.e. G is close to the union of two equal-sized cliques.

Although Conjecture 2.1.2 is optimal for the class of graphs on n vertices and minimum degree δ, it will not be optimal for every graph in the class – some graphs G will contain far more than $\mathrm{reg}_{\mathrm{even}}(n, \delta)/2$ edge-disjoint Hamilton cycles. The following conjecture accounts for this and would be best possible for every single graph G. Note that it is far stronger than Conjecture 2.1.2.

Conjecture 2.1.7 *Suppose G is a graph on n vertices with minimum degree $\delta(G) \geq n/2$. Then G contains at least $\mathrm{reg}_{\mathrm{even}}(G)/2$ edge-disjoint Hamilton cycles.*

For $\delta \geq (2-\sqrt{2}+\varepsilon)n$, this conjecture was proved in [59], based on the main result of [61]. It would already be very interesting to obtain an approximate version of Conjecture 2.1.7, i.e. a set of $(1-\varepsilon)\mathrm{reg}_{\mathrm{even}}(G)/2$ edge-disjoint Hamilton cycles under the assumption that $\delta(G) \geq (1+\varepsilon)n/2$.

As a very special case, Conjecture 2.1.7 would imply the long-standing 'Hamilton factorization' conjecture of Nash-Williams [71, 73]: any d-regular graph on at most $2d$ vertices contains $\lfloor d/2 \rfloor$ edge-disjoint Hamilton cycles. Jackson [73] raised the same conjecture independently, and proved a partial result. This was improved to an approximate version of the conjecture in [21]. The best current result towards the Hamilton factorization conjecture is due to Kühn and Osthus [59] (again as a corollary of their main result in [61]). Note that the set of Hamilton cycles guaranteed by Theorem 2.1.8 actually forms a Hamilton decomposition.

Theorem 2.1.8 *For every $\varepsilon > 0$ there exists an integer n_0 such that every d-regular graph on $n \geq n_0$ vertices for which $d \geq (1/2+\varepsilon)n$ is even contains $d/2$ edge-disjoint Hamilton cycles.*

Frieze and Krivelevich conjectured that the trivial bound of $\lfloor \delta(G)/2 \rfloor$ edge-disjoint Hamilton cycles is in fact correct for random graphs. Indeed, the results of several authors (mainly Krivelevich and Samotij [54] as well as Knox, Kühn and Osthus [51]) can be combined to show that for all $0 \leq p \leq 1$, the binomial random graph $G_{n,p}$ contains $\lfloor \delta(G_{n,p})/2 \rfloor$ edge-disjoint Hamilton cycles with high probability. Some further related results can be found in [45, 59, 61].

2.2 Notation

Given a graph G, we write $V(G)$ for its vertex set, $E(G)$ for its edge set, $e(G) := |E(G)|$ for the number of its edges and $|G|$ for the number of its vertices. Given $X \subseteq V(G)$, we write $G - X$ for the graph formed by deleting all vertices in X and $G[X]$ for the subgraph of G induced by X. We will also write $\overline{X} := V(G) \setminus X$ when it is unambiguous to do so. Given disjoint sets $X, Y \subseteq V(G)$, we write $G[X, Y]$ for the bipartite subgraph induced by X and Y. If G and G' are two graphs, we write $G \dot{\cup} G'$ for the graph on $V(G) \dot{\cup} V(G')$ with edge set $E(G) \dot{\cup} E(G')$. If $V(G) = V(G')$, we also write $G + G'$ for the graph on $V(G)$ with edge set $E(G) \cup E(G')$. An *r-factor* of a graph G is a spanning r-regular subgraph of G. If H is an r-factor of G and r is even then we also call H an *even factor* of G.

If G is an undirected graph, we write $\delta(G)$ for the minimum degree of G, $\Delta(G)$ for the maximum degree of G and $d(G)$ for the average degree of G. Whenever $X, Y \subseteq V(G)$, we write $e_G(X, Y)$ for the number of all those edges which have one endvertex in X and the other in Y. We write $e_G(X)$ for the number of edges in $G[X]$, and $e'_G(X, Y) := e_G(X, Y) + e_G(X \cap Y)$. Thus $e'_G(X, Y)$ is the number of ordered pairs (x, y) of vertices such that $x \in X$, $y \in Y$ and $xy \in E(G)$. Given a vertex x of G, we write $d_G(x)$ for the degree of x in G. We often omit the subscript G if this is unambiguous. Also, if $A \subseteq V(G)$ and the graph G is clear from the context, we sometimes write $d_A(x)$ for the number of neighbours of x in A. If G is a digraph, we write $\delta^+(G)$ for the minimum outdegree of G and $\delta^-(G)$ for the minimum indegree of G.

In order to simplify the presentation, we omit floors and ceilings and treat large numbers as integers whenever this does not affect the argument. The constants in the hierarchies used to state our results have to be chosen from right to left. More precisely, if we claim that a result holds whenever $0 < 1/n \ll a \ll b \ll c \leq 1$ (where n is the order of the graph or digraph), then this means that there are non-decreasing functions

$f : (0,1] \to (0,1]$, $g : (0,1] \to (0,1]$ and $h : (0,1] \to (0,1]$ such that the result holds for all $0 < a, b, c \le 1$ and all $n \in \mathbb{N}$ with $b \le f(c)$, $a \le g(b)$ and $1/n \le h(a)$. We will not calculate these functions explicitly. Hierarchies with more constants are defined in a similar way. Note that this is distinct from the other common definition of $\ll$ as a substitute for Landau O-notation.

Whenever $x \in \mathbb{R}$ we shall write $x_+ := \max\{x, 0\}$. We will write $a = x \pm \varepsilon$ as shorthand for $x - \varepsilon \le a \le x + \varepsilon$, and $a \ne x \pm \varepsilon$ as shorthand for the statement that either $a < x - \varepsilon$ or $a > x + \varepsilon$.

2.3 Proof outline and further notation

2.3.1 The extremal graph

We start by defining a graph $G_{n,\delta,\mathrm{ext}}$ on n vertices which is extremal for Theorem 2.1.4 in the sense that $G_{n,\delta,\mathrm{ext}}$ has minimum degree δ but the largest degree of an even factor of $G_{n,\delta,\mathrm{ext}}$ is at most the right hand side of (2.1.5). Given $\delta > n/2$, let Δ be the smallest integer such that $\Delta(\delta + \Delta - n)$ is even and $\Delta \ge (n + \sqrt{n(2\delta - n)})/2$. Partition the vertex set of $G_{n,\delta,\mathrm{ext}}$ into two classes A and B, with $|B| = \Delta$ and $|A| = n - \Delta$. Let $G_{n,\delta,\mathrm{ext}}[A]$ be empty, let $G_{n,\delta,\mathrm{ext}}[B]$ be any $(\delta + \Delta - n)$-regular graph, and let $G_{n,\delta,\mathrm{ext}}[A, B]$ be the complete bipartite graph. Clearly $\delta(G_{n,\delta,\mathrm{ext}}) = \delta$. Moreover, if H is a factor of $G_{n,\delta,\mathrm{ext}}$, then one can show that $d(H)$ is at most the right hand side of (2.1.5) (see [42] for details). In particular, $G_{n,\delta,\mathrm{ext}}$ contains at most $d(H)/2$ Hamilton cycles. Essentially the same construction was given in [21].

2.3.2 Tools and proof overview

An important concept in our proofs of Theorems 2.1.3 and 2.1.6 will be the notion of robust expanders. This concept was first introduced by Kühn, Osthus and Treglown [63] for directed graphs. Roughly speaking, a graph is a robust expander if for every set S

which is not too small and not too large, its "robust" neighbourhood is at least a little larger than S.

Definition 2.3.1 *Let G be a graph on n vertices. Given $0 < \nu \leq \tau < 1$ and $S \subseteq V(G)$, we define the ν-*robust neighbourhood *$RN_{\nu,G}(S)$ of S to be the set of all vertices $v \in V(G)$ with $d_S(v) \geq \nu n$. We say that G is a* robust (ν, τ)-expander *if for all $S \subseteq V(G)$ with $\tau n \leq |S| \leq (1-\tau)n$, we have $|RN_{\nu,G}(S)| \geq |S| + \nu n$.*

The main tool for our proofs is the following result of Kühn and Osthus [61] which states that every even-regular robust expander G whose degree is linear in $|G|$ has a Hamilton decomposition.

Theorem 2.3.2 *For every $\alpha > 0$, there exists $\tau > 0$ such that for every $\nu > 0$, there exists $n_0(\alpha, \tau, \nu)$ such that the following holds. Suppose that*

(i) G is an r-regular graph on $n \geq n_0$ vertices, where $r \geq \alpha n$ and r is even;

(ii) G is a robust (ν, τ)-expander.

Then G has a Hamilton decomposition.

Let G be a graph on n vertices as in Theorem 2.1.3. Let $\delta := \delta(G) = (1/2 + \alpha)n$. (So $\alpha \geq \varepsilon$.) As observed in [59], every graph on n vertices whose minimum degree is at least slightly larger than $n/2$ is a robust expander (see Lemma 2.5.2). Thus our given graph G is a robust expander. Let G^* be an even factor of largest degree in G. So $d(G^*) \geq \mathrm{reg}_{\mathrm{even}}(n, \delta)$. If G^* is still a robust expander, then we can apply Theorem 2.3.2 to obtain a Hamilton decomposition of G^* and thus at least $\mathrm{reg}_{\mathrm{even}}(n, \delta)/2$ edge-disjoint Hamilton cycles in G. The problem is that if α is small, then we could have $d(G^*) \leq n/2$. So we cannot guarantee that G^* is a robust expander. (However, this approach works if α is at least slightly larger than $3/2 - \sqrt{2}$. Indeed, in this case Theorem 2.1.4 implies that $d(G^*)$ will be slightly larger than $n/2$ and so G^* will be a robust expander. This observation was used in [59] to prove Theorem 2.1.3 for such values of α.)

So instead of using this simple strategy, in the proof of Theorem 2.1.3 we will distinguish two cases depending on whether our graph G contains a subgraph which is close to $G_{n,\delta,\mathrm{ext}}$. Suppose first that G contains such a subgraph, G_1 say. We can choose G_1 in such a way that $\delta(G_1) = \delta$, so G_1 must have an even factor G_2 of degree at least $\mathrm{reg}_{\mathrm{even}}(n, \delta)$. We will then use the fact that G_1 is close to $G_{n,\delta,\mathrm{ext}}$ in order to prove directly that G_2 is a robust expander. As before, this yields a Hamilton decomposition of G_2 by Theorem 2.3.2. This part of the argument is contained in Section 2.4.

If G does not contain a subgraph close to $G_{n,\delta,\mathrm{ext}}$, then we will first find a sparse even factor H of G which is still a robust expander and remove it from G. Call the resulting graph G'. We will then use the fact that G is far from containing $G_{n,\delta,\mathrm{ext}}$ to show that G' still contains an even factor H' of degree at least $\mathrm{reg}_{\mathrm{even}}(n, \delta)$. Since robust expansion is a monotone property, it follows that $H + H'$ is still a robust expander and may therefore be decomposed into Hamilton cycles by Theorem 2.3.2. So in this case we even find slightly more than $\mathrm{reg}_{\mathrm{even}}(n, \delta)/2$ edge-disjoint Hamilton cycles. This part of the argument is contained in Section 2.5. Altogether this will then imply Theorem 2.1.3.

In order to prove Theorem 2.1.6, we first show that every graph G whose minimum degree is close to $n/2$ either satisfies conditions (i) and (ii) of Theorem 2.1.6 or is a robust expander which does not contain a subgraph close to $G_{n,\delta,\mathrm{ext}}$. So suppose G does not satisfy (i) and (ii). We will use the fact that G is a robust expander to find a sparse robustly expanding even factor of G, and then argue similarly to the second part of the proof of Theorem 2.1.3 to find slightly more than $\mathrm{reg}_{\mathrm{even}}(n, \delta)/2$ edge-disjoint Hamilton cycles in G. This proof is contained in Section 2.6.

2.3.3 η-extremal graphs

The following definition formalises the notion of "containing a subgraph close to $G_{n,\delta,\mathrm{ext}}$". For technical reasons we extend the definition to the case where α is negative – this will be used in Section 2.6 (with $|\alpha|$ very small). Note that if $\delta = (1/2 + \alpha)n$, then the vertex classes A and B of $G_{n,\delta,\mathrm{ext}}$ have sizes roughly $(1/2 - \sqrt{\alpha/2})n$ and $(1/2 + \sqrt{\alpha/2})n$ respectively, and that $G_{n,\delta,\mathrm{ext}}[B]$ is regular of degree roughly $(\alpha + \sqrt{\alpha/2})n$.

Definition 2.3.3 *Let $\eta > 0$ and $-1/2 \leq \alpha \leq 1/2$, and let G be a graph on n vertices with $\delta(G) = (1/2 + \alpha)\,n$. Recall that $\alpha_+ = \max\{\alpha, 0\}$. We say that G is η-*extremal *if there exist disjoint $A, B \subseteq V(G)$ such that*

(E1) $|A| = (1/2 - \sqrt{\alpha_+/2} \pm \eta)n$;

(E2) $|B| = (1/2 + \sqrt{\alpha_+/2} \pm \eta)n$;

(E3) $e(A, B) > (1 - \eta)|A||B|$;

(E4) $e(B) < (\alpha_+ + \sqrt{\alpha_+/2} + \eta)n|B|/2$.

Note that (E1) and (E2) together imply

(E5) $n - |A \cup B| \leq 2\eta n$.

Note also that we allow $G[A]$ to be arbitrary – we do not force A to be close to an independent set, for example. This is necessary, since adding edges internal to A does not disrupt the extremality of $G_{n,\delta,\mathrm{ext}}$.

The following result states that if G is η-extremal, then $G[B]$ is "almost regular".

Lemma 2.3.4 *Suppose $0 < \eta \ll \alpha, 1/2 - \alpha < 1/2$. Suppose G is an η-extremal graph on n vertices, with $\delta(G) = (1/2 + \alpha)\,n$, and let $A, B \subseteq V(G)$ be as in Definition 2.3.3.*

(i) For all vertices $v \in B$, we have $d_B(v) \geq (\alpha + \sqrt{\alpha/2} - 3\eta)n$.

(ii) For all but at most $2\sqrt{\eta}n$ vertices $v \in B$, we have $d_B(v) \le (\alpha + \sqrt{\alpha/2} + 2\sqrt{\eta})n$.

Proof. (i) immediately follows from (E1) and (E5). Indeed, for all $v \in B$, we have

$$\begin{aligned} d_B(v) &\ge \delta(G) - d_A(v) - d_{\overline{A \cup B}}(v) \overset{(E5)}{\ge} \delta(G) - |A| - 2\eta n \\ &\overset{(E1)}{\ge} \left(\alpha + \sqrt{\frac{\alpha}{2}} - 3\eta\right) n, \end{aligned} \tag{2.3.5}$$

as desired.

Suppose (ii) fails. Then there exist at least $2\sqrt{\eta}n$ vertices in B with degree greater than $(\alpha + \sqrt{\alpha/2} + 2\sqrt{\eta})n$ in B. We therefore have

$$\begin{aligned} e_G(B) &= \frac{1}{2}\sum_{v \in B} d_B(v) \overset{(2.3.5)}{>} \frac{1}{2}\left(\left(\alpha + \sqrt{\frac{\alpha}{2}} - 3\eta\right) n|B| + 2\sqrt{\eta}n \cdot 2\sqrt{\eta}n\right) \\ &\ge \frac{1}{2}\left(\alpha + \sqrt{\frac{\alpha}{2}} + \eta\right) n|B|. \end{aligned}$$

But this contradicts (E4), so (ii) must hold. □

2.4 The near-extremal case

Suppose that $0 < 1/n \ll \eta \ll \alpha < 1/2$, and that G is an η-extremal graph on n vertices with $\delta(G) = (1/2 + \alpha)n$. Recall that our aim in this case is to show that G contains a factor of degree $\mathrm{reg}_{\mathrm{even}}(n, \delta)/2$ which is a robust expander. Let $A, B \subseteq V(G)$ be as in Definition 2.3.3. We will first show that G contains a spanning subgraph G_1 which is close to $G_{n,\delta,\mathrm{ext}}$ and satisfies $\delta(G_1) = \delta(G)$.

Lemma 2.4.1 *Suppose $0 < 1/n \ll \eta \ll 1/C \ll 1/2 - \alpha \le 1/2$, so that in particular $0 \le \alpha < 1/2$. Let G be an η-extremal graph on n vertices with $\delta := \delta(G) = (1/2 + \alpha)\, n$, and let $A, B \subseteq V(G)$ be as in Definition 2.3.3. Then there exists a spanning subgraph G_1 of G which satisfies the following properties:*

(i) A and B satisfy (E1)–(E4) for the graph G_1. In particular, G_1 is η-extremal.

(ii) $\delta(G_1) = \delta$.

(iii) $e_{G_1}(A) < C\eta|A|^2$.

Proof. We will define G_1 using a greedy algorithm. Initially, let $G_1 := G$. Suppose that $G_1[A]$ contains an edge xy such that $d_{G_1}(x), d_{G_1}(y) > \delta$. Then remove xy from G_1, and continue in this way until G_1 contains no such edge. Note that we have $\delta(G_1) = \delta$, and (E1)–(E4) are not affected by these edge deletions, so G_1 satisfies (i) and (ii).

Suppose $e_{G_1}(A) \geq C\eta|A|^2$, and note that we have

$$\delta = \left(\frac{1}{2} + \alpha\right) n \leq \left(\frac{1}{2} + \sqrt{\frac{\alpha}{2}}\right) n \overset{\text{(E2)}}{\leq} |B| + \eta n.$$

(Indeed, $x \leq \sqrt{x/2}$ for all $0 \leq x \leq 1/2$.) If $v \in A$ is a vertex with $d_{G_1}(v) = \delta$, we therefore have

$$d_{G[A,B]}(v) = d_{G_1[A,B]}(v) \leq \delta - d_{G_1[A]}(v) \leq |B| + \eta n - d_{G_1[A]}(v).$$

Each edge in $G_1[A]$ must have at least one endpoint with degree δ in G_1, so

$$\begin{aligned} e_G(A,B) = \sum_{v \in A} d_{G[A,B]}(v) &\leq |A||B| - \sum_{v \in A,\, d_{G_1}(v) = \delta} \big(d_{G_1[A]}(v) - \eta n\big) \\ &\leq |A||B| + \eta n^2 - e_{G_1}(A) \leq |A| \left(|B| + \eta \frac{n^2}{|A|} - C\eta|A|\right). \end{aligned}$$

Since $1/C \ll 1/2 - \alpha$, we have $C|A| \geq 2|B| + n^2/|A|$ by (E1) and (E2). Hence

$$e_G(A,B) \leq |A|(|B| - 2\eta|B|) = (1 - 2\eta)|A||B|,$$

which contradicts (E3). We therefore have $e_{G_1}(A) < C\eta|A|^2$, and so G_1 satisfies (iii) as desired. $\square$

Let G_1 be as in Lemma 2.4.1, and let G_2 be a degree-maximal even factor of G_1. (So G_2 is an even-regular spanning subgraph of G_1 whose degree is as large as possible.) By Theorem 2.1.4, we have that

$$d(G_2) \geq \mathrm{reg}_{\mathrm{even}}(n, \delta) \geq \frac{n}{4} + \frac{\alpha n}{2} + \sqrt{\frac{\alpha}{2}}n - 2. \tag{2.4.2}$$

It can be shown that any degree-maximal even factor of $G_{n,\delta,\mathrm{ext}}$ contains almost all edges inside the larger vertex class B. The following lemma uses a similar argument to prove a similar statement for G_1.

Lemma 2.4.3 *Suppose* $0 < 1/n \ll \eta \ll 1/C \ll \alpha, 1/2 - \alpha < 1/2$. *Suppose that* G *is an* η*-extremal graph on* n *vertices with* $\delta(G) = (1/2 + \alpha)n$. *Let* G_1 *be the graph obtained by applying Lemma 2.4.1 to* G*, and let* G_2 *be a degree-maximal even factor of* G_1. *Let* $A, B \subseteq V(G)$ *be as in Definition 2.3.3. Then for all but at most* $3\eta^{1/4}n$ *vertices* $v \in B$*, we have*

$$d_{G_2[B]}(v) \geq \left(\alpha + \sqrt{\frac{\alpha}{2}} - 3\eta^{\frac{1}{4}}\right) n.$$

Proof. Let r be the degree of G_2. Suppose that $d_{G_2[B]}(v) < (\alpha + \sqrt{\alpha/2} - 3\eta^{1/4})n$ for more than $3\eta^{1/4}n$ vertices. Then by Lemma 2.3.4(ii), we have

$$\begin{aligned}
r|B| &= \sum_{v \in B} d_{G_2}(v) = e_{G_2}(A, B) + 2e_{G_2}(B) \\
&\leq r|A| + \left(\alpha + \sqrt{\frac{\alpha}{2}} + 2\sqrt{\eta}\right) n|B| + 4\sqrt{\eta}n^2 - 3\eta^{\frac{1}{4}}n \cdot 3\eta^{\frac{1}{4}}n \\
&\leq r|A| + \left(\alpha + \sqrt{\frac{\alpha}{2}} - 3\sqrt{\eta}\right) n|B|.
\end{aligned}$$

Since $|B|-|A| \geq (\sqrt{2\alpha}-2\eta)n$ by (E1) and (E2), it follows that

$$\sqrt{2\alpha} rn \leq \left(\alpha + \sqrt{\frac{\alpha}{2}} - 3\sqrt{\eta}\right) n|B| + 2\eta n^2,$$

and hence

$$\begin{aligned} r &\leq \left(\sqrt{\frac{\alpha}{2}} + \frac{1}{2} - 3\sqrt{\frac{\eta}{2\alpha}}\right)|B| + \eta\sqrt{\frac{2}{\alpha}}n \\ &\overset{\text{(E2)}}{\leq} \left(\sqrt{\frac{\alpha}{2}} + \frac{1}{2} - 3\sqrt{\frac{\eta}{2\alpha}}\right)\left(\frac{1}{2} + \sqrt{\frac{\alpha}{2}} + \eta\right) n + \eta\sqrt{\frac{2}{\alpha}}n \\ &\leq \left(\frac{1}{4} + \frac{\alpha}{2} + \sqrt{\frac{\alpha}{2}} - \eta^{3/4}\right) n. \end{aligned}$$

(In the last inequality we used that $\eta \ll \alpha$.) It therefore follows from (2.4.2) that $r < \mathrm{reg}_{\mathrm{even}}(n, \delta)$. But G_2 was chosen to be degree-maximal, so this is a contradiction. $\square$

We now collect some robust expansion properties of G_2. For convenience, if $X \subseteq V(G_2)$, we shall write $X_A := X \cap A$ and $X_B := X \cap B$. In particular, if $S \subseteq V(G)$ then (for example) $RN_\nu(S_A)_B = RN_\nu(S \cap A) \cap B$.

Lemma 2.4.4 *Suppose that* $0 < 1/n \ll \nu \ll \eta \ll \mu \ll \tau \ll \lambda \ll 1/C \ll \alpha, 1/2 - \alpha < 1/2$. *Suppose that* G *is an* η*-extremal graph on* n *vertices with* $\delta(G) = (1/2 + \alpha)\, n$. *Let* G_1 *be the graph obtained by applying Lemma 2.4.1 to* G, *and let* G_2 *be a degree-maximal even factor of* G_1. *Let* $A, B \subseteq V(G)$ *be as in Definition 2.3.3. Then in the graph* G_2, *the following properties all hold.*

(i) If $S \subseteq A$ *with* $|S| \geq |A|/2$, *then* $|RN_\nu(S)_B| \geq (1-\mu)|B|$.

(ii) If $S \subseteq B$ *with* $|S| \geq |B|/2$, *then* $|RN_\nu(S)_A| \geq (1-\mu)|A|$.

(iii) If $S \subseteq A$ *with* $|S| \geq \tau n/3$, *then* $|RN_\nu(S)_B| \geq |B|/2 + \lambda n$.

(iv) If $S \subseteq B$ *with* $|S| \geq \tau n/3$, *then* $|RN_\nu(S)_A| \geq |A|/2 + \lambda n$.

(v) If $S \subseteq B$, then $|RN_\nu(S)_B| \geq |S| - \mu n$.

Proof. Write $d := d(G_2)$. We first prove (i). Suppose $S \subseteq A$ with $|S| \geq |A|/2$. Lemma 2.3.4(ii) implies that in G_2 all but at most $2\sqrt{\eta}n \leq \mu|B|$ vertices $v \in B$ satisfy

$$\begin{aligned} d_A(v) &= d - d_{\overline{A \cup B}}(v) - d_B(v) \\ &\overset{(2.4.2),(E5)}{\geq} \left(\frac{1}{4} + \frac{\alpha}{2} + \sqrt{\frac{\alpha}{2}}\right)n - 2 - 2\eta n - \left(\alpha + \sqrt{\frac{\alpha}{2}} + 2\sqrt{\eta}\right)n \\ &\geq \left(\frac{1}{4} - \frac{\alpha}{2} - 3\sqrt{\eta}\right)n \geq \left(\frac{1}{4} - \frac{1}{2}\sqrt{\frac{\alpha}{2}} + \eta\right)n + \nu n \overset{(E1)}{\geq} \frac{|A|}{2} + \nu n, \end{aligned}$$

where the third inequality follows since $x < \sqrt{x}/2$ for all $0 < x < 1/4$. Thus in the graph G_2 we have $|N_A(v) \cap S| \geq \nu n$, and hence $v \in RN_\nu(S)$, for each such v. The result therefore follows.

We now prove (ii). Suppose $S \subseteq B$ with $|S| \geq |B|/2$. Let $A' \subseteq A$ be the set of vertices which in G_2 have less than $|B|/2 + \nu n$ neighbours inside B. Each vertex $v \in A'$ must satisfy

$$\begin{aligned} d_A(v) &= d - d_{\overline{A \cup B}}(v) - d_B(v) \\ &\overset{(2.4.2),(E5)}{\geq} \left(\frac{1}{4} + \frac{\alpha}{2} + \sqrt{\frac{\alpha}{2}}\right)n - 2 - 2\eta n - \frac{|B|}{2} - \nu n \\ &\overset{(E2)}{\geq} \left(\frac{1}{4} + \frac{\alpha}{2} + \sqrt{\frac{\alpha}{2}} - 2\eta - \nu\right)n - 2 - \left(\frac{1}{4} + \frac{1}{2}\sqrt{\frac{\alpha}{2}} + \eta\right)n \\ &\geq \frac{\alpha}{2}n, \end{aligned}$$

and so we have $e_{G_2}(A) \geq \alpha n|A'|/4$. But by Lemma 2.4.1(iii) we have $e_{G_2}(A) \leq e_{G_1}(A) < C\eta|A|^2$. Therefore

$$|A'| \leq \frac{4C\eta}{\alpha} \cdot \frac{|A|^2}{n} \leq \sqrt{\eta}\frac{|A|^2}{n} \leq \sqrt{\eta}|A| \leq \mu|A|.$$

However, our assumption that $|S| \geq |B|/2$ and the definition of A' together imply that every vertex $v \in A \setminus A'$ satisfies $|N_B(v) \cap S| \geq \nu n$. Therefore $|RN_\nu(S)| \geq |A \setminus A'| \geq (1-\mu)|A|$, as required.

We now prove (iii). Suppose $S \subseteq A$ with $|S| \geq \tau n/3$. Then we double-count the edges between S and B in G_2. The definition of a robust neighbourhood implies that

$$e_{G_2}(S,B) = e_{G_2}(S, RN_\nu(S)_B) + e_{G_2}(S, B \setminus RN_\nu(S)_B) \leq |S||RN_\nu(S)_B| + \nu n^2.$$

On the other hand, Lemma 2.4.1(iii) implies that

$$\begin{aligned} e_{G_2}(S,B) &\geq d|S| - 2e_{G_2}(S,A) - e_{G_2}(S, \overline{A \cup B}) \overset{\text{(E5)}}{\geq} d|S| - 2C\eta|A|^2 - 2\eta n^2 \\ &\geq d|S| - 3C\eta n^2. \end{aligned}$$

Combining the two inequalities yields

$$\begin{aligned} |RN_\nu(S)_B| &\geq d - 3C\eta\frac{n^2}{|S|} - \nu\frac{n^2}{|S|} \\ &\overset{(2.4.2)}{\geq} \left(\frac{1}{4} + \frac{\alpha}{2} + \sqrt{\frac{\alpha}{2}}\right)n - 2 - \frac{9C\eta}{\tau}n - \frac{3\nu}{\tau}n \\ &\overset{\text{(E2)}}{\geq} \frac{|B|}{2} + \left(\frac{\alpha}{2} + \frac{1}{2}\sqrt{\frac{\alpha}{2}} - \eta - \frac{9C\eta}{\tau} - \frac{3\nu}{\tau}\right)n - 2 \geq \frac{|B|}{2} + \frac{\alpha}{2}n, \end{aligned}$$

and so the result follows.

We now prove (iv). Suppose $S \subseteq B$ with $|S| \geq \tau n/3$. Then we double-count the edges between S and A in G_2. As before, we have

$$e_{G_2}(S,A) \leq |S||RN_\nu(S)_A| + \nu n^2. \tag{2.4.5}$$

On the other hand,

$$e_{G_2}(S,A) \geq d|S| - \sum_{v\in S} d_B(v) - \sum_{v\in S} d_{\overline{A\cup B}}(v) \overset{\text{(E5)}}{\geq} d|S| - \sum_{v\in S} d_B(v) - 2\eta n^2.$$

Lemma 2.3.4(ii) implies that

$$\sum_{v\in S} d_B(v) \leq 2\sqrt{\eta}n^2 + \left(\alpha + \sqrt{\frac{\alpha}{2}} + 2\sqrt{\eta}\right) n|S|,$$

and so

$$\begin{aligned}
e_{G_2}(S,A) &\overset{(2.4.2),\text{(E5)}}{\geq} \left(\frac{1}{4} + \frac{\alpha}{2} + \sqrt{\frac{\alpha}{2}}\right) n|S| - 2|S| \\
&\qquad - \left(\alpha + \sqrt{\frac{\alpha}{2}} + 2\sqrt{\eta}\right) n|S| - (2\eta + 2\sqrt{\eta})n^2 \\
&\geq \left(\frac{1}{4} - \frac{\alpha}{2}\right) n|S| - 5\sqrt{\eta}n^2.
\end{aligned}$$

Combining this with (2.4.5) shows that in G_2 we have

$$\begin{aligned}
|RN_\nu(S)_A| &\geq \left(\frac{1}{4} - \frac{\alpha}{2}\right) n - 6\sqrt{\eta} \cdot \frac{n^2}{|S|} \geq \left(\frac{1}{4} - \frac{\alpha}{2}\right) n - \frac{18\sqrt{\eta}}{\tau} n \\
&= \left(\frac{1}{4} - \frac{1}{2}\sqrt{\frac{\alpha}{2}}\right) n + \left(\frac{1}{2}\sqrt{\frac{\alpha}{2}} - \frac{\alpha}{2}\right) n - \frac{18\sqrt{\eta}}{\tau} n \\
&\overset{\text{(E1)}}{\geq} \frac{|A|}{2} + \frac{1}{2}\left(\frac{1}{2}\sqrt{\frac{\alpha}{2}} - \frac{\alpha}{2}\right) n \geq \frac{|A|}{2} + \lambda n,
\end{aligned}$$

and so the result follows. (Here we used that $\sqrt{x}/2 > x$ for all $0 < x < 1/4$.)

Finally, we prove (v). Suppose $S \subseteq B$. Recall that $e'_{G_2}(S, RN_\nu(S)_B)$ denotes the number of ordered pairs (u,v) of vertices of G_2 such that $uv \in E(G_2)$, $u \in S$ and $v \in RN_\nu(S)_B$. (Note that this may not equal $e(S, RN_\nu(S)_B)$, as we may have $S \cap RN_\nu(S)_B \neq$

$\emptyset$.) By Lemma 2.3.4(ii), counting from $RN_\nu(S)_B$ yields

$$e'_{G_2}(S, RN_\nu(S)_B) \leq \left(\alpha + \sqrt{\frac{\alpha}{2}} + 2\sqrt{\eta}\right) n|RN_\nu(S)_B| + 2\sqrt{\eta}n^2.$$

By Lemma 2.4.3, counting from S yields

$$e'_{G_2}(S, RN_\nu(S)_B) \geq \left(\alpha + \sqrt{\frac{\alpha}{2}} - 3\eta^{\frac{1}{4}}\right) n|S| - 3\eta^{\frac{1}{4}}n^2 - \nu n^2.$$

Combining the two inequalities yields $|RN_\nu(S)_B| \geq |S| - \mu n$ as desired. □

We are now in a position to prove Theorem 2.1.3 for η-extremal graphs.

Lemma 2.4.6 *Suppose* $0 < 1/n \ll \eta \ll \alpha, 1/2 - \alpha < 1/2$. *If* G *is an* η*-extremal graph on* n *vertices with* $\delta := \delta(G) = (1/2 + \alpha)n$, *then* G *contains at least* $\mathrm{reg}_{\mathrm{even}}(n, \delta)/2$ *edge-disjoint Hamilton cycles.*

Proof. Let $\tau_0 := \tau(1/4)$ be the constant returned by Theorem 2.3.2. Choose additional constants ν, μ, τ, λ and C such that

$$0 < \nu \ll \eta \ll \mu \ll \tau \ll \lambda \ll 1/C \ll \alpha, 1/2 - \alpha, \tau_0.$$

Take $A, B \subseteq V(G)$ as in Definition 2.3.3. Apply Lemma 2.4.1 to G and C to obtain a spanning subgraph G_1. Let G_2 be a degree-maximal even factor of G_1. Note that Lemma 2.4.4 may be applied to G_2.

Claim: *G_2 is a robust (ν, τ)-expander.*

Note that the claim immediately implies the desired result. Indeed, any robust (ν, τ)-expander is also a robust (ν, τ_0)-expander, and so Theorem 2.3.2 implies that G_2 may be decomposed into Hamilton cycles. Moreover, Lemma 2.4.1 implies that $\delta(G_1) = \delta$ and so $d(G_2) \geq \mathrm{reg}_{\mathrm{even}}(n, \delta)$. Hence the Hamilton decomposition of G_2 yields the desired collection of $d(G_2)/2 \geq \mathrm{reg}_{\mathrm{even}}(n,\delta)/2$ edge-disjoint Hamilton cycles.

To prove the claim, consider $S \subseteq V(G)$ with $\tau n \leq |S| \leq (1-\tau)n$. We will use Lemma 2.4.4 to show that in G_2 we have $|RN_\nu(S)| \geq |S| + \nu n$. We will split the proof into cases depending on the sizes of $S_A = S \cap A$ and $S_B = S \cap B$. Note that $|S_{\overline{A\cup B}}| \leq 2\eta n$ by (E5).

Case 1: $|S_A| \leq \tau n/3$, $|S_B| \leq \tau n/3$.

In this case, we have

$$|S| \stackrel{\text{(E5)}}{\leq} \frac{2\tau}{3} + 2\eta n < \tau n,$$

which is a contradiction.

Case 2: $\tau n/3 \leq |S_A| \leq |A|/2$, $|S_B| \leq \tau n/3$.

In this case, by Lemma 2.4.4(iii) we have

$$\begin{aligned}|RN_\nu(S)| \geq |RN_\nu(S_A)_B| \geq \frac{|B|}{2} + \lambda n &\stackrel{\text{(E5)}}{\geq} \frac{|A|}{2} + \sqrt{\frac{\alpha}{2}}n - 2\eta n + \lambda n \\ &\geq \frac{|A|}{2} + \frac{\tau}{3}n + 2\eta n + \nu n \geq |S| + \nu n,\end{aligned}$$

as desired.

Case 3: $|S_A| \geq |A|/2$, $|S_B| \leq \tau n/3$.

In this case, by Lemma 2.4.4(i) we have

$$\begin{aligned}|RN_\nu(S)| \geq |RN_\nu(S_A)_B| \geq (1-\mu)|B| &\geq |A| + 2\sqrt{\frac{\alpha}{2}}n - 2\eta n - \mu n \\ &\geq |A| + \frac{\tau}{3}n + 2\eta n + \nu n \geq |S| + \nu n,\end{aligned}$$

as desired.

Case 4: $|S_A| \leq |A|/2$, $|S_B| \geq \tau n/3$.

In this case, by Lemma 2.4.4(iv) and (v), we have

$$\begin{aligned}|RN_\nu(S)| &\geq |RN_\nu(S_B)_A| + |RN_\nu(S_B)_B| \geq \frac{|A|}{2} + \lambda n + |S_B| - \mu n \\ &\geq |S_A| + |S_B| + 2\eta n + \nu n \geq |S| + \nu n,\end{aligned}$$

as desired.

Case 5: $|S_A| \geq |A|/2$, $\tau n/3 \leq |S_B| \leq |B|/2$.

In this case, by Lemma 2.4.4(i) and (iv), we have

$$\begin{aligned}|RN_\nu(S)| &\geq |RN_\nu(S_A)_B| + |RN_\nu(S_B)_A| \geq |B| + \frac{|A|}{2} + (\lambda - \mu)n \\ &\geq \frac{|B|}{2} + |A| + (\lambda - \mu)n \geq |S_B| + |S_A| + 2\eta n + \nu n \geq |S| + \nu n,\end{aligned}$$

as desired, where the third inequality follows since $|B| \geq |A|$ by (E1) and (E2).

Case 6: $|S_A| \geq |A|/2$, $|S_B| \geq |B|/2$.

In this case, by Lemma 2.4.4(i) and (ii), we have

$$\begin{aligned}|RN_\nu(S)| &\geq |RN_\nu(S_A)_B| + |RN_\nu(S_B)_A| \geq |B| + |A| - 2\mu n \\ &\stackrel{\text{(E5)}}{\geq} n - (2\eta + 2\mu)n \geq (1 - \tau)n + \nu n \geq |S| + \nu n,\end{aligned}$$

as desired.

Thus in all cases we have $|RN_\nu(S)| \geq |S| + \nu n$. Indeed, if $|S_B| \leq \tau n/3$ this follows by Cases 1, 2 and 3; if $\tau n/3 \leq |S_B| \leq |B|/2$ this follows by Cases 4 and 5; and if $|S_B| \geq |B|/2$ this follows by Cases 4 and 6. Hence G_2 is a robust (ν, τ)-expander as desired. This proves the claim and hence the lemma. □

2.5 The non-extremal case

Suppose now that G is not η-extremal. Our first aim is to find a sparse even factor H of G which is a robust expander. This has essentially already been done in [59], but for digraphs. We first require the following definition, which generalises the notion of robust expanders to digraphs.

Definition 2.5.1 *Let D be a digraph on n vertices. Given $0 < \nu \leq \tau < 1$, we define the* ν-robust outneighbourhood $RN^+_{\nu,D}(S)$ *of S to be the set of all vertices $v \in V(D)$ which have at least νn inneighbours in S. We say that D is a* robust (ν, τ)-outexpander *if for all $S \subseteq V(D)$ with $\tau n \leq |S| \leq (1-\tau)n$, we have $|RN^+_{\nu,D}(S)| \geq |S| + \nu n$.*

We will now quote three lemmas from [59]. Lemma 2.5.2 implies that our given graph G is a robust expander. We will use Lemmas 2.5.3 and 2.5.4 to deduce Lemma 2.5.5, which together with Lemma 2.5.2 implies that G contains a sparse even factor H which is still a robust expander.

Lemma 2.5.2 *Suppose $0 < \nu \leq \tau \leq \varepsilon < 1/2$ and $\varepsilon \geq 2\nu/\tau$. Let G be a graph on n vertices with minimum degree $\delta(G) \geq (1/2 + \varepsilon)\, n$. Then G is a robust (ν, τ)-expander.*

Lemma 2.5.3 *Suppose $0 < 1/n \ll \eta \ll \nu, \tau, \alpha < 1$. Suppose that G is a robust (ν, τ)-expander on n vertices with $\delta(G) \geq \alpha n$. Then the edges of G can be oriented in such a way that the resulting oriented graph G' satisfies the following properties:*

(i) G' is a robust $(\nu/4, \tau)$-outexpander.

(ii) $d^+_{G'}(x), d^-_{G'}(x) = (1 \pm \eta) d_G(x)/2$ for every vertex x of G.

An *r-factor* of a digraph G is a spanning subdigraph of G in which every vertex has in- and outdegree r.

Lemma 2.5.4 *Suppose* $0 < 1/n \ll \nu' \ll \xi \ll \nu \leq \tau \ll \alpha < 1$*. Let* G *be a robust* (ν,τ)*-outexpander on* n *vertices with* $\delta^+(G), \delta^-(G) \geq \alpha n$*. Then* G *contains a* ξn*-factor which is still a robust* (ν',τ)*-outexpander.*

Lemma 2.5.5 *Suppose* $0 < 1/n \ll \nu' \ll \varepsilon \ll \nu \ll \tau \ll \alpha < 1$*, and suppose in addition that* εn *is an even integer. If* G *is a robust* (ν,τ)*-expander on* n *vertices with* $\delta(G) \geq \alpha n$*, then there exists an* εn*-factor* H *of* G *which is a robust* (ν',τ)*-expander.*

Proof. We apply Lemma 2.5.3 to orient the edges of G, forming an oriented graph G' which is a robust $(\nu/4,\tau)$-outexpander and which satisfies $\delta^+(G'), \delta^-(G') \geq \alpha n/3$. We then apply Lemma 2.5.4 to find an $\varepsilon n/2$-factor H of G' which is a robust (ν',τ)-outexpander. Now remove the orientation on the edges of H to obtain a robust (ν',τ)-expander which is an εn-factor of G, as desired. $\square$

We will now show that even after removing a sparse factor H, our given graph G still contains an even factor of degree at least $\mathrm{reg}_{\mathrm{even}}(n,\delta)$. To do this, we first show that $G - H$ is still non-extremal.

Lemma 2.5.6 *Suppose* $0 < 1/n \ll \varepsilon \ll \eta \ll 1/2 - \alpha$*, and that* $-\varepsilon \leq \alpha < 1/2$*. Let* G *be a graph on* n *vertices with* $\delta(G) = (1/2+\alpha)\,n$ *which is not* 2η*-extremal. Suppose* H *is an* εn*-factor of* G*. Then* $G - H$ *is not* η*-extremal.*

Proof. Suppose $A, B \subseteq V(G)$ are disjoint with $|A|$ and $|B|$ satisfying (E1) and (E2) of Definition 2.3.3. Let $G' := G - H$. Since G is not 2η-extremal, we must have either $e_G(A,B) \leq (1-2\eta)|A||B|$ or $e_G(B) \geq (\alpha_+ + \sqrt{\alpha_+/2} + 2\eta)n|B|/2$. In the former case we have

$$e_{G'}(A,B) \leq e_G(A,B) < (1-\eta)|A||B|,$$

and in the latter case we have

$$\begin{aligned} e_{G'}(B) &\geq e_G(B) - \varepsilon n|B| \geq \frac{1}{2}\left(\alpha_+ + \sqrt{\frac{\alpha_+}{2}} + 2\eta - 2\varepsilon\right) n|B| \\ &\geq \frac{1}{2}\left((\alpha-\varepsilon)_+ + \sqrt{\frac{(\alpha-\varepsilon)_+}{2}} + \frac{3\eta}{2}\right) n|B|. \end{aligned}$$

Since $\delta(G-H) = (1/2+\alpha-\varepsilon)\,n$, it follows that $G-H$ is not η-extremal. □

We now show that $G-H$ contains a large even factor. We will do this using the well-known result of Tutte [92], given below.

Theorem 2.5.7 *Let G be a graph. Given disjoint $S, T \subseteq V(G)$ and $r \in \mathbb{N}$, let $Q_r(S,T)$ be the number of connected components C of $G-(S\cup T)$ such that $r|C| + e(C,T)$ is odd, and let*

$$R_r(S,T) := \sum_{v\in T} d(v) - e(S,T) + r(|S|-|T|). \tag{2.5.8}$$

Then G contains an r-factor if and only if $Q_r(S,T) \leq R_r(S,T)$ for all disjoint $S,T \subseteq V(G)$.

In proving the following lemma, we follow a similar approach to that used in [21]. We will also make frequent and implicit use of the inequality $\sqrt{x} \leq \sqrt{x+h} \leq \sqrt{x}+\sqrt{h}$ for $x, h \geq 0$.

Lemma 2.5.9 *Suppose $0 < 1/n \ll \varepsilon \ll \eta \ll 1/2-\alpha$ and that $-\varepsilon \leq \alpha < 1/2$. Let G be a graph on n vertices with minimum degree $\delta := \delta(G) = (1/2+\alpha)\,n$, and suppose that G is not η-extremal. Let*

$$r := \frac{n}{4} + \frac{(\alpha+\varepsilon)n}{2} + \sqrt{\frac{\alpha+\varepsilon}{2}}n,$$

and suppose that r is an even integer. Then G contains an r-factor.

Proof. Let S, T be two arbitrary disjoint subsets of $V(G)$. We will show that $Q_r(S,T) \leq R_r(S,T)$, from which the result follows by Theorem 2.5.7. We first note a useful bound on $Q_r(S,T)$. If $\delta \geq |S|+|T|$ then every vertex outside $S \cup T$ has at least $\delta - |S| - |T|$ neighbours outside $S \cup T$, so every component of $G-(S\cup T)$ contains at least $\delta-|S|-|T|+1$ vertices. Thus

$$Q_r(S,T) \leq \frac{n-|S|-|T|}{\delta-|S|-|T|+1} \quad \text{if } \delta \geq |S|+|T|. \tag{2.5.10}$$

Also, note that we always have

$$\delta - r = \left(\frac{1}{4} + \frac{\alpha-\varepsilon}{2} - \sqrt{\frac{\alpha+\varepsilon}{2}}\right) n = \left(\frac{1}{4} + \frac{\alpha+\varepsilon}{2} - \sqrt{\frac{\alpha+\varepsilon}{2}} - \varepsilon\right) n \geq \varepsilon n, \tag{2.5.11}$$

since $1/4 + x - \sqrt{x} = (1/2-\sqrt{x})^2 > 0$ for all $0 \leq x < 1/4$ and since $\varepsilon \ll 1/2 - \alpha$. We will now split the proof into cases depending on $|S|$ and $|T|$.

Case 1: $|T| \leq r-1$, $|S| \leq \delta - r$, and $|S|+|T| \geq 3$.

We have

$$\begin{aligned} R_r(S,T) &\overset{(2.5.8)}{=} \sum_{v\in T}(d(v)-r) + \sum_{v\in S}(r-d_T(v)) \geq |T|(\delta-r) + \sum_{v\in S} 1 \\ &\overset{(2.5.11)}{\geq} |S|+|T|. \end{aligned} \tag{2.5.12}$$

Let $k := |S|+|T|$. By (2.5.10) and (2.5.12) it suffices to show that $k \geq (n-k)/(\delta-k+1)$. This is equivalent to showing that

$$\delta k - k^2 + 2k - n = (k-2)(\delta-k) + 2\delta - n \geq 0.$$

We have $3 \leq k \leq \delta - 1$ and the function $(k-2)(\delta-k)$ is concave, so it must be minimised in this range when $k=3$ or when $k = \delta-1$. In either case, we have

$$(k-2)(\delta-k) + 2\delta - n = \delta - 3 + 2\delta - n \geq \delta - 3 - 2\varepsilon n \geq 0$$

as desired.

Case 2: $0 \leq |S| + |T| \leq 2$.

If $S = T = \emptyset$, then we have $Q_r(S,T) = R_r(S,T) = 0$ (since r is even). So suppose that $|S| + |T| > 0$. Then it follows from (2.5.10) that

$$Q_r(S,T) < \frac{n}{\delta - 1} \leq \frac{3n}{n} = 3.$$

If $T \neq \emptyset$, we have

$$R_r(S,T) \overset{(2.5.8)}{\geq} \delta|T| - 1 - r|T| \overset{(2.5.11)}{\geq} 3.$$

If $T = \emptyset$, we have $|S| \geq 1$ and so by (2.5.8) we have $R_r(S,T) \geq r \geq 3$. We therefore have $Q_r(S,T) \leq R_r(S,T)$ in all cases.

Case 3: $|T| \geq r$ or $|S| \geq \delta - r$, but not both.

We have

$$\begin{aligned} R_r(S,T) &\overset{(2.5.8)}{\geq} (\delta - r)|T| - |S||T| + r|S| \\ &= (|T| - r)(\delta - r - |S|) + r(\delta - r) \qquad (2.5.13) \\ &\geq r(\delta - r) \overset{(2.5.11)}{\geq} \frac{\varepsilon}{4}n^2. \end{aligned}$$

(Note that (2.5.13) holds regardless of the values of $|S|$ and $|T|$.) Moreover, we have $Q_r(S,T) \leq n$. Hence $Q_r(S,T) \leq R_r(S,T)$ as desired.

Case 4: $|T| \geq r$, $|S| \geq \delta - r$, and $|T| \neq (n + 2r - \delta)/2 \pm 3\sqrt{\varepsilon}n$.

The right hand side of (2.5.13) is clearly minimised when $|S| + |T| = n$. It therefore suffices to consider this case alone, yielding

$$\begin{aligned} R_r(S,T) - Q_r(S,T) &\geq (\delta - r)|T| - (n - |T|)|T| + r(n - |T|) - n \\ &= |T|^2 + (\delta - 2r - n)|T| + n(r-1). \end{aligned}$$

Define a polynomial $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = x^2 + (\delta - 2r - n)x + n(r-1).$$

Suppose this quadratic has real zeroes at τ_1 and τ_2, with $\tau_1 < \tau_2$. Then for $|T| \leq \tau_1$ and $|T| \geq \tau_2$, we must have $R_r(S,T) - Q_r(S,T) \geq 0$. The discriminant D of f is given by

$$\begin{aligned} D &= (n + 2r - \delta)^2 - 4n(r-1) \\ &= (n + 2r - \delta)^2 - \left(1 + 2(\alpha+\varepsilon) + 2\sqrt{2(\alpha+\varepsilon)} - \frac{4}{n}\right) n^2. \end{aligned}$$

But

$$n + 2r - \delta = \left(1 + \varepsilon + \sqrt{2(\alpha+\varepsilon)}\right) n, \qquad (2.5.14)$$

so

$$(n + 2r - \delta)^2 = \left(1 + \varepsilon^2 + 2(\alpha+\varepsilon) + 2\varepsilon + 2(1+\varepsilon)\sqrt{2(\alpha+\varepsilon)}\right) n^2$$

and

$$D = \varepsilon\left(\varepsilon + 2 + 2\sqrt{2(\alpha+\varepsilon)}\right) n^2 - 4n.$$

Hence $0 < D \leq 5\varepsilon n^2$. In particular, the quadratic does indeed have two real zeroes $\tau_1 < \tau_2$, and from the quadratic formula we have

$$\tau_1 \geq \frac{n + 2r - \delta - 3\sqrt{\varepsilon} n}{2}, \qquad \tau_2 \leq \frac{n + 2r - \delta + 3\sqrt{\varepsilon} n}{2}.$$

Since we are in Case 4, we therefore have either $|T| \leq \tau_1$ or $|T| \geq \tau_2$, and the result follows.

Case 5: $|T| = (n + 2r - \delta)/2 \pm 3\sqrt{\varepsilon}n$ and $\delta - r \leq |S| \leq (n - 2r + \delta)/2 - 3\sqrt{\varepsilon}n$.

(Note that our condition on $|T|$ implies that we cannot have $|S| > (n - 2r + \delta)/2 + 3\sqrt{\varepsilon}n$.) Let $x_0 := (n + 2r - \delta)/2 + 3\sqrt{\varepsilon}n \geq |T|$. We then have

$$R_r(S,T) \overset{(2.5.13)}{\geq} (x_0 - r)(\delta - r - |S|) + r(\delta - r).$$

Since $x_0 + |S| \leq n$, we may now argue exactly as in Case 4 (with x_0 in place of $|T|$) to show that $R_r(S,T) \geq Q_r(S,T)$.

Case 6: $|T| = (n + 2r - \delta)/2 \pm 3\sqrt{\varepsilon}n$ and $|S| = (n - 2r + \delta)/2 \pm 3\sqrt{\varepsilon}n$.

In this case, we will use the fact that G is not η-extremal. From (2.5.14), we have

$$\left| \frac{n + 2r - \delta}{2} - \left(\frac{1}{2} + \sqrt{\frac{\alpha_+}{2}} \right) n \right| \leq \left(\frac{\varepsilon}{2} + \sqrt{\frac{\varepsilon}{2}} \right) n.$$

Since $\varepsilon \ll \eta$, we may conclude that

$$\left| |T| - \left(\frac{1}{2} + \sqrt{\frac{\alpha_+}{2}} \right) n \right| < \eta n.$$

A similar argument shows that

$$\left| |S| - \left(\frac{1}{2} - \sqrt{\frac{\alpha_+}{2}} \right) n \right| < \eta n.$$

Since G is not η-extremal, this implies that either $e(S,T) \leq (1 - \eta)|S||T|$ or

$$e(T) \geq \frac{1}{2} \left(\alpha_+ + \sqrt{\frac{\alpha_+}{2}} + \eta \right) n|T|.$$

Case 6a: $e(S,T) \leq (1-\eta)|S||T|$.

Then we have

$$\begin{aligned} R_r(S,T) - Q_r(S,T) &\overset{(2.5.8)}{\geq} (\delta - r)|T| - (1-\eta)|S||T| + r|S| - n \\ &\geq (\delta - r)|T| - (1-\eta)(n - |T|)|T| + r(n - |T| - 6\sqrt{\varepsilon} n) - n \\ &= (1-\eta)|T|^2 + (\delta - 2r - (1-\eta)n)|T| + (1 - 6\sqrt{\varepsilon})nr - n. \end{aligned}$$

Write this quadratic as $a|T|^2 + b|T| + c$, and let the discriminant be D. We then have

$$\begin{aligned} b^2 = ((1-\eta)n + 2r - \delta)^2 &\overset{(2.5.14)}{=} \left(1 - \eta + \varepsilon + \sqrt{2(\alpha+\varepsilon)}\right)^2 n^2 \\ &= \left((1-\eta)^2 + \varepsilon^2 + 2(\alpha+\varepsilon) + 2(1-\eta)\varepsilon + 2(1-\eta)\sqrt{2(\alpha+\varepsilon)} + 2\varepsilon\sqrt{2(\alpha+\varepsilon)}\right) n^2 \\ &\leq \left((1-\eta)^2 + 2\alpha + 2(1-\eta)\sqrt{2(\alpha+\varepsilon)} + \varepsilon^{\frac{1}{3}}\right) n^2 \end{aligned}$$

and

$$\begin{aligned} 4ac &= 4(1-\eta)(1-6\sqrt{\varepsilon})nr - 4(1-\eta)n \\ &\geq (1-\eta)(1-6\sqrt{\varepsilon})\left(1 + 2(\alpha+\varepsilon) + 2\sqrt{2(\alpha+\varepsilon)}\right) n^2 - 4n \\ &\geq (1-\eta)\left(1 + 2\alpha + 2\sqrt{2(\alpha+\varepsilon)}\right) n^2 - \varepsilon^{\frac{1}{3}} n^2. \end{aligned}$$

Thus

$$\begin{aligned} D = b^2 - 4ac &\leq \left((1-\eta)^2 - (1-\eta) + 2\eta\alpha + 2\varepsilon^{\frac{1}{3}}\right) n^2 \\ &= \left(-\eta\,(1 - \eta - 2\alpha) + 2\varepsilon^{\frac{1}{3}}\right) n^2 < 0, \end{aligned}$$

where the last line follows since $\varepsilon \ll \eta \ll 1/2 - \alpha$ and $\alpha < 1/2$. Hence this quadratic has no real zeroes, and $R_r(S,T) - Q_r(S,T) \geq 0$ as desired.

Case 6b: $e(T) \geq (\alpha_+ + \sqrt{\alpha_+/2} + \eta)n|T|/2$ and $e(S,T) \geq (1-\eta)|S||T|$.

Then we have

$$\begin{aligned}
\sum_{v\in T} d(v) &\geq e(S,T) + 2e(T) \\
&\geq \left((1-\eta)|S| + \left(\alpha_+ + \sqrt{\frac{\alpha_+}{2}} + \eta\right) n\right)|T| \\
&\geq \left((1-\eta)\left(\frac{n-2r+\delta}{2} - 3\sqrt{\varepsilon}n\right) + \left(\alpha + \sqrt{\frac{\alpha_+}{2}} + \eta\right) n\right)|T| \\
&\overset{(2.5.14)}{\geq} \left((1-\eta)\left(\frac{1}{2} - \frac{\varepsilon}{2} - \sqrt{\frac{\alpha+\varepsilon}{2}} - 3\sqrt{\varepsilon}\right) + \alpha + \sqrt{\frac{\alpha_+}{2}} + \eta\right) n|T| \\
&\geq \left(\frac{1}{2} - \sqrt{\frac{\alpha+\varepsilon}{2}} - 4\sqrt{\varepsilon} - \frac{\eta}{2} + \alpha + \sqrt{\frac{\alpha_+}{2}} + \eta\right) n|T| \\
&\geq \left(\frac{1}{2} + \frac{\eta}{3} + \alpha\right) n|T|.
\end{aligned}$$

Hence

$$\begin{aligned}
R_r(S,T) - Q_r(S,T) &\overset{(2.5.8)}{\geq} \sum_{v\in T} d(v) - (n-|T|)|T| + r(|S|-|T|) - n \\
&\geq \sum_{v\in T} d(v) + |T|^2 - n|T| + r(n-|T|-6\sqrt{\varepsilon}n) - r|T| - n \\
&= \sum_{v\in T} d(v) + |T|^2 - (n+2r)|T| + (1-6\sqrt{\varepsilon})nr - n \\
&\geq |T|^2 - \left(\left(\frac{1}{2} - \frac{\eta}{3} - \alpha\right) n + 2r\right)|T| + (1-6\sqrt{\varepsilon})nr - n \\
&\geq |T|^2 - \left(1 + \varepsilon + \sqrt{2(\alpha+\varepsilon)} - \frac{\eta}{3}\right) n|T| + (1-6\sqrt{\varepsilon})nr - n \\
&\geq |T|^2 - \left(1 + \sqrt{2(\alpha+\varepsilon)} - \frac{\eta}{4}\right) n|T| + (1-7\sqrt{\varepsilon})nr.
\end{aligned}$$

Write this quadratic as $|T|^2 + b|T| + c$, and let the discriminant be D. We then have

$$b^2 \leq \left(1 + 2\alpha + 2\varepsilon + \frac{\eta^2}{16} + 2\sqrt{2(\alpha+\varepsilon)} - \frac{\eta}{2}\right) n^2 \leq \left(1 + 2\alpha + 2\sqrt{2(\alpha+\varepsilon)} - \frac{\eta}{3}\right) n^2$$

and

$$
\begin{aligned}
4c &= 4(1-7\sqrt{\varepsilon})nr = (1-7\sqrt{\varepsilon})\left(1+2(\alpha+\varepsilon)+2\sqrt{2(\alpha+\varepsilon)}\right)n^2 \\
&\geq \left(1+2\alpha+2\sqrt{2(\alpha+\varepsilon)}\right)n^2 - \varepsilon^{\frac{1}{3}}n^2.
\end{aligned}
$$

Thus

$$
D = b^2 - 4c \leq \left(\varepsilon^{\frac{1}{3}} - \frac{\eta}{3}\right)n^2 < 0
$$

since $\varepsilon \ll \eta$. Hence this quadratic has no real zeroes, and $R_r(S,T) - Q_r(S,T) \geq 0$ as desired. This completes the proof. □

It is now simple to prove that every non-extremal graph G whose minimum degree δ is slightly larger than $n/2$ contains significantly more than $\mathrm{reg}_{\mathrm{even}}(n,\delta)/2$ edge-disjoint Hamilton cycles.

Lemma 2.5.15 *Suppose* $0 < 1/n \ll c \ll \eta \ll \alpha, 1/2-\alpha < 1/2$. *Let G be a graph on n vertices with* $\delta := \delta(G) = (1/2+\alpha)n$ *such that G is not η-extremal. Then G contains at least* $\mathrm{reg}_{\mathrm{even}}(n,\delta)/2 + cn$ *edge-disjoint Hamilton cycles.*

Proof. Let $\tau_0 := \tau(1/4)$ be as defined in Theorem 2.3.2. Choose new constants $\varepsilon, \varepsilon', \nu, \nu', \tau$ such that

$$
0 < 1/n \ll \nu', c \ll \varepsilon, \varepsilon' \ll \eta \ll \nu \ll \tau \ll \alpha, 1/2-\alpha, \tau_0.
$$

Let

$$
r := \left(\frac{1}{4} + \frac{\alpha+\varepsilon'}{2} + \sqrt{\frac{\alpha+\varepsilon'}{2}}\right)n.
$$

By reducing ε' and ε slightly if necessary we may assume that both r and εn are even integers. By Lemmas 2.5.2 and 2.5.5, G contains an εn-factor H which is a robust (ν',τ)-expander. Let $G' := G - H$. By Lemma 2.5.6, G' is not $(\eta/2)$-extremal. Since also

$\delta(G') = (1/2 + \alpha - \varepsilon)n$, we can apply Lemma 2.5.9 with $\varepsilon + \varepsilon'$ and $\alpha - \varepsilon$ playing the roles of ε and α to find an r-factor H' of G'.

Since H is a robust (ν', τ)-expander (and thus also a robust (ν', τ_0)-expander), the same holds for $H + H'$. Hence by Theorem 2.3.2, $H + H'$ can be decomposed into $d(H + H')/2$ edge-disjoint Hamilton cycles. By Theorem 2.1.4 we have $r \geq \mathrm{reg}_{\mathrm{even}}(n, \delta)$, and so

$$\frac{1}{2}d(H + H') \geq \frac{1}{2}\left(\mathrm{reg}_{\mathrm{even}}(n, \delta) + \varepsilon n\right) \geq \frac{1}{2}\mathrm{reg}_{\mathrm{even}}(n, \delta) + cn$$

as desired. $\square$

2.6 Proof of Theorems 2.1.3 and 2.1.6

We first combine Lemmas 2.4.6 and 2.5.15 to prove Theorem 2.1.3.

Proof of Theorem 2.1.3. Choose $n_0 \in \mathbb{N}$ and an additional constant η such that $1/n_0 \ll \eta \ll \varepsilon$. Define α by $\delta(G) = (1/2+\alpha)n$. Recall from Section 2.1 that Theorem 2.1.3 was already proved in [59] for the case when $\delta(G) \geq (2 - \sqrt{2} + \varepsilon)n$. So we may assume that $\alpha \leq 3/2 - \sqrt{2} + \varepsilon$ and so $\eta \ll \alpha, 1/2 - \alpha$. Thus we can apply Lemma 2.4.6 (if G is η-extremal) or Lemma 2.5.15 (if G is not η-extremal) to find $\mathrm{reg}_{\mathrm{even}}(n, \delta(G))/2$ edge-disjoint Hamilton cycles in G. $\square$

Let G be a graph on n vertices whose minimum degree is not much smaller than $n/2$. Before we can prove Theorem 2.1.6, we must first show that either G is a robust expander or it is close to either the complete bipartite graph $K_{n/2,n/2}$ or the disjoint union $K_{n/2}\dot\cup K_{n/2}$ of two cliques. The former case corresponds to (i) of Theorem 2.1.6, and the latter case corresponds to (ii).

Definition 2.6.1 *We say that a graph G is* ε-close to $K_{n/2,n/2}$ *if there exists $A \subseteq V(G)$ with $|A| = \lfloor n/2 \rfloor$ and such that $e(A) \leq \varepsilon n^2$. We say that G is* ε-close to $K_{n/2}\dot\cup K_{n/2}$ *if* *there exists $A \subseteq V(G)$ with $|A| = \lfloor n/2 \rfloor$ and such that $e(A, \overline{A}) \leq \varepsilon n^2$.*

Suppose that G is a graph of minimum degree roughly $n/2$. If G is ε-close to $K_{n/2,n/2}$ then the bipartite subgraph of G induced by A and $\overline{A}$ is almost complete. However, $\overline{A}$ may also contain many edges. If G is ε-close to $K_{n/2}\dot{\cup}K_{n/2}$ then both $G[A]$ and $G[\overline{A}]$ are almost complete.

Lemma 2.6.2 *Suppose $0 < 1/n \ll \kappa \ll \nu \ll \tau, \varepsilon < 1$. Let G be a graph on n vertices of minimum degree $\delta := \delta(G) \geq (1/2 - \kappa)n$. Then G satisfies one of the following properties:*

(i) G is ε-close to $K_{n/2,n/2}$;

(ii) G is ε-close to $K_{n/2}\dot{\cup}K_{n/2}$;

(iii) G is a robust (ν, τ)-expander.

Proof. Suppose $S \subseteq V(G)$ with $\tau n \leq |S| \leq (1 - \tau)n$. Our aim is to show that either $RN := RN_\nu(S)$ has size at least $|S| + \nu n$ or that G is close to either $K_{n/2,n/2}$ or $K_{n/2}\dot{\cup}K_{n/2}$. We will split the proof into cases depending on $|S|$.

Case 1: $\tau n \leq |S| \leq (1/2 - \sqrt{\nu})n$.

In this case, we have

$$\delta|S| \leq e'(S, V(G)) = e'(S, RN) + e'(S, \overline{RN}) \leq |S||RN| + \nu n^2 \leq |S||RN| + \nu n\frac{|S|}{\tau},$$

and so $|RN| \geq (1/2 - \kappa - \nu/\tau)n \geq |S| + \nu n$ as desired. (Recall that $e'(A, B)$ denotes the number of ordered pairs (a, b) with $ab \in E(G)$, $a \in A$ and $b \in B$.)

Case 2: $(1/2 + 2\nu)n \leq |S| \leq (1 - \tau)n$.

In this case, we have $RN = V(G)$ and so the result is immediate. Indeed, for all $v \in V(G)$, we have $d(v) \geq (1/2 - \kappa)n$ and so $|N(v) \cap S| \geq (2\nu - \kappa)n \geq \nu n$.

Case 3: $(1/2 - \sqrt{\nu})n \leq |S| \leq (1/2 + 2\nu)n$.

Suppose that $|RN| < |S| + \nu n$. We will first show that either $|S \setminus RN| < \sqrt{\nu}n$ or G is ε-close to $K_{n/2,n/2}$. Suppose $|S \setminus RN| \geq \sqrt{\nu}n$. Then

$$\begin{aligned}|S \setminus RN|(\delta - \nu n) &\leq e(S \setminus RN, \overline{S}) = e(S \setminus RN, \overline{S} \cap RN) + e(S \setminus RN, \overline{S} \setminus RN) \\ &\leq |S \setminus RN||\overline{S} \cap RN| + \nu n^2 \leq |S \setminus RN||\overline{S} \cap RN| + \sqrt{\nu}n|S \setminus RN|,\end{aligned}$$

and so $|\overline{S} \cap RN| \geq \delta - 2\sqrt{\nu}n$. But then together with our assumption that $|RN| < |S| + \nu n$, this implies $|S \cap RN| < 3\sqrt{\nu}n$. Hence $e(S) \leq 3\sqrt{\nu}n^2 + |S|\nu n < 4\sqrt{\nu}n^2$. By adding or removing at most $\sqrt{\nu}n$ arbitrary vertices to or from S, we can form a set A of $\lfloor n/2 \rfloor$ vertices with

$$e(A) < 4\sqrt{\nu}n^2 + \sqrt{\nu}n^2 = 5\sqrt{\nu}n^2 \leq \varepsilon n^2.$$

Thus G is ε-close to $K_{n/2,n/2}$.

We may therefore assume that $|S \setminus RN| < \sqrt{\nu}n$, from which it follows that $|\overline{S} \cap RN| < 2\sqrt{\nu}n$ (by our initial assumption that $|RN| < |S| + \nu n$). We will now show that G is ε-close to $K_{n/2} \dot{\cup} K_{n/2}$. We have $e(S, \overline{S} \cap RN) \leq |S||\overline{S} \cap RN| \leq 2\sqrt{\nu}n^2$, and hence $e(S, \overline{S}) \leq 3\sqrt{\nu}n^2$. As before, by adding or removing at most $\sqrt{\nu}n$ arbitrary vertices to or from S, we can therefore form a set A of $\lfloor n/2 \rfloor$ vertices with $e(A, \overline{A}) \leq e(S, \overline{S}) + \sqrt{\nu}n^2 \leq \varepsilon n^2$. Hence G is ε-close to $K_{n/2} \dot{\cup} K_{n/2}$.

If G is not ε-close to either $K_{n/2,n/2}$ or $K_{n/2} \dot{\cup} K_{n/2}$, we must therefore have $|RN| \geq |S| + \nu n$ for all $S \subseteq V(G)$ with $\tau n \leq |S| \leq (1 - \tau)n$, so that G is a robust (ν, τ)-expander as required. $\square$

We now have all the tools we need to prove Theorem 2.1.6.

Proof of Theorem 2.1.6. Let $\tau := \tau(1/4)$ be as defined in Theorem 2.3.2. Choose $n_0 \in \mathbb{N}$ and new constants $\varepsilon', \varepsilon'', \nu, \nu'$ such that

$$0 < 1/n_0 \ll \nu' \ll \varepsilon \ll \varepsilon', \varepsilon'' \ll \nu \ll \tau, \eta.$$

Consider any graph G on $n \geq n_0$ vertices as in Theorem 2.1.6. Let $\delta := \delta(G)$ and define α by $\delta = (1/2 + \alpha)n$. So $-\varepsilon \leq \alpha \leq \varepsilon$. Let

$$r := \left(\frac{1}{4} + \frac{\alpha + \varepsilon'}{2} + \sqrt{\frac{\alpha + \varepsilon'}{2}} \right) n.$$

By reducing ε' and ε'' slightly if necessary we may assume that both r and $\varepsilon'' n$ are even integers.

Suppose that G does not satisfy (i), i.e. $e(X) \geq \eta n^2$ for all $X \subseteq V(G)$ with $|X| = \lfloor n/2 \rfloor$. We claim that G is not $(\eta/4)$-extremal. To show this, consider any set $B \subseteq V(G)$ with

$$|B| = \left(\frac{1}{2} + \sqrt{\frac{\alpha_+}{2}} \pm \frac{\eta}{4} \right) n.$$

By adding or removing at most $\eta n/2$ arbitrary vertices to and from B, we obtain a set B' with $|B'| = \lfloor n/2 \rfloor$ and such that $e(B') \leq e(B) + \eta n^2/2$. Together with our assumption that (i) does not hold, this implies that

$$e(B) \geq \frac{\eta n^2}{2} \geq \frac{1}{2} \left(\alpha_+ + \sqrt{\frac{\alpha_+}{2}} + \frac{\eta}{4} \right) n|B|.$$

Hence G is not $(\eta/4)$-extremal.

Suppose moreover that (ii) does not hold, so that G fails to be η-close to $K_{n/2} \dot\cup K_{n/2}$. By Lemma 2.6.2, it follows that G is a robust (ν, τ)-expander. By Lemma 2.5.5, G therefore contains an $\varepsilon'' n$-factor H which is a robust (ν', τ)-expander. Let $G' := G - H$. By Lemma 2.5.6, G' is not $(\eta/8)$-extremal. Since also $\delta(G') = (1/2 + \alpha - \varepsilon'')n$, we can apply Lemma 2.5.9 with $\varepsilon' + \varepsilon''$ and $\alpha - \varepsilon''$ playing the roles of ε and α to find an r-factor H' of G'.

Since H is a robust (ν', τ)-expander, the same holds for $H + H'$. Hence by Theorem 2.3.2, $H + H'$ can be decomposed into $d(H + H')/2$ edge-disjoint Hamilton cycles. By Theorem 2.1.4 (and the fact that $\mathrm{reg}_{\mathrm{even}}(n, \delta) = 0$ if $\delta < n/2$) we have $r \geq \max\{\mathrm{reg}_{\mathrm{even}}(n, \delta), n/8\}$, and so

$$\frac{1}{2}d(H + H') \geq \frac{1}{2}\left(\max\{\mathrm{reg}_{\mathrm{even}}(n, \delta), n/8\} + \varepsilon'' n\right) \geq \frac{1}{2}\max\{\mathrm{reg}_{\mathrm{even}}(n, \delta), n/8\} + \varepsilon n,$$

as desired. □

Chapter 3

Proof of a conjecture of Thomassen on Hamilton cycles in highly connected tournaments

3.1 Introduction

3.1.1 Main result

A *tournament* is an orientation of a complete graph and a Hamilton cycle in a tournament is a (consistently oriented) cycle which contains all the vertices of the tournament. Hamilton cycles in tournaments have a long and rich history. For instance, one of the most basic results about tournaments is Camion's theorem, which states that every strongly connected tournament has a Hamilton cycle [20]. This is strengthened by Moon's theorem [70], which implies that such a tournament is even pancyclic, i.e. contains cycles of all possible lengths. Many related results have been proved; the monograph by Bang-Jensen and Gutin [7] gives an overview which also includes many recent results.

In 1982, Thomassen [89] made a very natural conjecture on how to guarantee not just one Hamilton cycle, but many edge-disjoint ones: he conjectured that for every k there is an $f(k)$ so that every strongly $f(k)$-connected tournament contains k edge-disjoint Hamilton cycles (see also the recent surveys [6, 60]). This turned out to be surprisingly difficult: not even the existence of $f(2)$ was known so far. Our main result shows that $f(k) = O(k^2 \log^2 k)$.

Theorem 3.1.1 *There exists $C > 0$ such that for all $k \in \mathbb{N}$ with $k \geq 2$ every strongly $Ck^2 \log^2 k$-connected tournament contains k edge-disjoint Hamilton cycles.*

In Proposition 3.5.1, we describe an example which shows that $f(k) \geq (k-1)^2/4$, i.e. our bound on the connectivity is asymptotically close to best possible. Thomassen [89] observed that $f(2) > 2$ and conjectured that $f(2) = 3$. He also observed that one cannot weaken the assumption in Theorem 3.1.1 by replacing strong connectivity with strong edge-connectivity; see Section 3.5.

To simplify the presentation, we have made no attempt to optimize the value of the constant C. Our exposition shows that one can take $C := 10^{12}$ for $k \geq 20$. Rather than proving Theorem 3.1.1 directly, we deduce it as an immediate consequence of two further results, which are both of independent interest: we show that every sufficiently highly connected tournament is highly linked (see Theorem 3.1.3) and show that every highly linked tournament contains many edge-disjoint Hamilton cycles (see Theorem 2.1.3).

3.1.2 Linkedness in tournaments

Given sets A, B of size k in a strongly k-connected digraph D, Menger's theorem implies that D contains k vertex-disjoint paths from A to B. In a k-linked digraph, we can even specify the initial and final vertex of each such path (see Section 3.2 for the precise definition).

Theorem 3.1.2 *There exists $C' > 0$ such that for all $k \in \mathbb{N}$ with $k \geq 2$ every $C'k^2 \log k$-linked tournament contains k edge-disjoint Hamilton cycles.*

The bound in Theorem 3.1.2 is asymptotically close to best possible, as we shall discuss below. We will show that $C' := 10^7$ works for all $k \geq 20$. (As mentioned earlier, we have made no attempt to optimise the value of this constant.)

It is not clear from the definition that every (very) highly connected tournament is also highly linked. In fact, for general digraphs this is far from true: Thomassen [91] showed that for all k there are strongly k-connected digraphs which are not even 2-linked. On the other hand, he showed that there is an (exponential) function $g(k)$ so that every strongly $g(k)$-connected tournament is k-linked [90]. The next result shows that we can take $g(k)$ to be almost linear in k. Note that this result together with Proposition 3.5.1 shows that Theorem 3.1.2 is asymptotically best possible up to logarithmic terms.

Theorem 3.1.3 *For all $k \in \mathbb{N}$ with $k \geq 2$ every strongly $10^4 k \log k$-connected tournament is k-linked.*

For small k, the constant 10^4 can easily be improved (see Theorem 3.4.5). The proof of Theorem 3.1.3 is based on a fundamental result of Ajtai, Komlós and Szemerédi [1, 2] on the existence of asymptotically optimal sorting networks. Though their result is asymptotically optimal, it is not clear whether this is the case for Theorem 3.1.3. In fact, for the case of (undirected) graphs, a deep result of Bollobás and Thomason [17] states that every $22k$-connected graph is k-linked (this was improved to $10k$ by Thomas and Wollan [87]). Thus one might believe that a similar relation also holds in the case of tournaments:

Conjecture 3.1.4 *There exists $C > 0$ such that for all $k \in \mathbb{N}$ every strongly Ck-connected tournament is k-linked.*

Similarly, we believe that the logarithmic terms can also be removed in Theorems 3.1.1 and 3.1.2:

Conjecture 3.1.5

(i) There exists $C' > 0$ such that for all $k \in \mathbb{N}$ every $C'k^2$-linked tournament contains k edge-disjoint Hamilton cycles.

(ii) There exists $C'' > 0$ such that for all $k \in \mathbb{N}$ every strongly $C''k^2$-connected tournament contains k edge-disjoint Hamilton cycles.

Note that Conjectures 3.1.4 and 3.1.5(i) together imply Conjecture 3.1.5(ii). Both conjectures have now been proved in subsequent work by Pokrovskiy [75, 76].

3.1.3 Algorithmic aspects

Both Hamiltonicity and linkedness in tournaments have also been studied from an algorithmic perspective. Camion's theorem implies that the Hamilton cycle problem (though NP-complete in general) is solvable in polynomial time for tournaments. Chudnovsky, Scott and Seymour [22] solved a long-standing problem of Bang-Jensen and Thomassen [8] by showing that the linkedness problem is also solvable in polynomial time for tournaments. More precisely, for a given tournament on n vertices, one can determine in time polynomial in n whether it is k-linked and if yes, one can produce a corresponding set of k paths (also in polynomial time). Fortune, Hopcroft and Wyllie [34] showed that for general digraphs, the problem is NP-complete even for $k = 2$. We can use the result in [22] to obtain an algorithmic version of Theorem 3.1.2. More precisely, given a $C'k^2 \log k$-linked tournament on n vertices, one can find k edge-disjoint Hamilton cycles in time polynomial in n (where k is fixed). We discuss this in more detail in Section 3.9. Note that this immediately results in an algorithmic version of Theorem 3.1.1.

3.1.4 Related results and spanning regular subgraphs

Proposition 3.5.1 actually suggests that the 'bottleneck' to finding k edge-disjoint Hamilton cycles is the existence of a k-regular subdigraph: it states that if the connectivity of a tournament T is significantly lower than in Theorem 3.1.1, then T may not even contain a spanning k-regular subdigraph. There are other results which exhibit this phenomenon: if T is itself regular, then Kelly's conjecture from 1968 states that T itself has a Hamilton decomposition. Kelly's conjecture was proved very recently (for large tournaments) by Kühn and Osthus [61].

Erdős raised a 'probabilistic' version of Kelly's conjecture: for a tournament T, let $\delta^0(T)$ denote the minimum of the minimum out-degree and the minimum in-degree. He conjectured that for almost all tournaments T, the maximum number of edge-disjoint Hamilton cycles in T is exactly $\delta^0(T)$. In particular, this would imply that with high probability, $\delta^0(T)$ is also the degree of a densest spanning regular subdigraph in a random tournament T. This conjecture of Erdős was proved by Kühn and Osthus [59], based on the main result in [61].

It would be interesting to obtain further conditions which relate the degree of the densest spanning regular subdigraph of a tournament T to the number of edge-disjoint Hamilton cycles in T. For undirected graphs, one such conjecture was made in [55]: it states that for any graph G satisfying the conditions of Dirac's theorem, the number of edge-disjoint Hamilton cycles in G is exactly half the degree of a densest spanning even-regular subgraph of G. An approximate version of this conjecture was proved by Ferber, Krivelevich and Sudakov [33], see e.g. [55, 59] for some related results.

The methods used in the current chapter are quite different from those used e.g. in the papers mentioned in Section 3.1.4. A crucial ingredient is the construction of highly structured dominating sets (see Section 3.3 for an informal description). We believe that this approach will have further applications. Indeed, Kühn, Osthus and Townsend [62] have recently developed it to give an affirmative answer to the following question of Thomassen (see [82]): given any positive integers $k_1, \dots, k_t$, does there exist an integer $f(k_1, \dots, k_t)$ such that every strongly $f(k_1, \dots, k_t)$-connected tournament T admits a partition of its vertex set into vertex classes $V_1, \dots, V_t$ such that for all $1 \leq i \leq t$ the subtournament $T[V_i]$ is strongly k_i-connected? In fact [62] contains a stronger result, which has further applications to a problem on cycle factors.

3.1.5 Organization of the chapter

In the next section, we introduce the notation that will be used for the remainder of the chapter. In Section 3.3, we give an overview of the proof of Theorem 3.1.2. In Sections 3.4 and 3.5, we give the relatively short proofs of Theorem 3.1.3 and Proposition 3.5.1. In Section 3.6, we show that given a 'linked domination structure' (as introduced in the proof sketch), we can find a single Hamilton cycle (Lemma 3.6.7). In Section 3.7, we show that given several suitable linked domination structures, we can repeatedly apply Lemma 3.6.7 to find k edge-disjoint Hamilton cycles. In Section 3.8 we show that any highly linked tournament contains such suitable linked domination structures. Finally, Section 3.9 contains some concluding remarks.

3.2 Notation

The digraphs considered in this chapter do not have loops and we allow up to two edges between any pair of x, y of distinct vertices, at most one in each direction. A digraph is an *oriented graph* if there is at most one edge between any pair x, y of distinct vertices, i.e. if it does not contain a cycle of length two.

Given a digraph D, we write $V(D)$ for its vertex set, $E(D)$ for its edge set, $e(D) := |E(D)|$ for the number of its edges and $|D|$ for its *order*, i.e. for the number of its vertices. We write $H \subseteq D$ to mean that H is a subdigraph of D, i.e. $V(H) \subseteq V(D)$ and $E(H) \subseteq E(D)$. Given $X \subseteq V(D)$, we write $D - X$ for the digraph obtained from D by deleting all vertices in X, and $D[X]$ for the subdigraph of D induced by X. Given $F \subseteq E(D)$, we write $D - F$ for the digraph obtained from D by deleting all edges in F. We write $V(F)$ for the set of all endvertices of edges in F. If H is a subdigraph of D, we write $D - H$

for $D - E(H)$.

We write xy for an edge directed from x to y. Unless stated otherwise, when we refer to paths and cycles in digraphs, we mean directed paths and cycles, i.e. the edges on these paths and cycles are oriented consistently. Given a path $P = x \dots y$ from x to y and a vertex z outside P which sends an edge to x, we write zxP for the path obtained from P by appending the edge zx. The *length* of a path or cycle is the number of its edges. We call the terminal vertex of a path P the *head* of P and denote it by $h(P)$. Similarly, we call the initial vertex of a path P the *tail* of P and denote it by $t(P)$. The *interior* $\mathrm{Int}(P)$ of a path P is the subpath obtained by deleting $t(P)$ and $h(P)$. Thus $\mathrm{Int}(P) = \emptyset$ if P has length at most one. Two paths P and P' are *internally disjoint* if $P \neq P'$ and $V(\mathrm{Int}(P)) \cap V(\mathrm{Int}(P')) = \emptyset$. A *path system* $\mathcal{P}$ is a collection of vertex-disjoint paths. We write $V(\mathcal{P})$ for the set of all vertices lying on paths in $\mathcal{P}$ and $E(\mathcal{P})$ for the set of all edges lying on paths in $\mathcal{P}$. We write $h(\mathcal{P})$ for the set consisting of the heads of all paths in $\mathcal{P}$ and $t(\mathcal{P})$ for the set consisting of the tails of all paths in $\mathcal{P}$. If $v \in V(\mathcal{P})$, we write v^+ and v^- for the successor and predecessor of v on the path in $\mathcal{P}$ containing v. A path system $\mathcal{P}$ is a *path cover* of a directed graph D if every path in $\mathcal{P}$ lies in D and together the paths in $\mathcal{P}$ cover all the vertices of D. If $X \subseteq V(D)$ and $\mathcal{P}$ is a path cover of $D[X]$, we sometimes also say that $\mathcal{P}$ is a path cover of X.

If x is a vertex of a digraph D, then $N^+_D(x)$ denotes the *out-neighbourhood* of x, i.e. the set of all those vertices y for which $xy \in E(D)$. Similarly, $N^-_D(x)$ denotes the *in-neighbourhood* of x, i.e. the set of all those vertices y for which $yx \in E(D)$. We write $d^+_D(x) := |N^+_D(x)|$ for the *out-degree* of x and $d^-_D(x) := |N^-_D(x)|$ for its *in-degree.* We denote the *minimum out-degree* of D by $\delta^+(D) := \min\{d^+_D(x) : x \in V(D)\}$ and the *maximum out-degree* of D by $\Delta^+(D) := \max\{d^+_D(x) : x \in V(D)\}$. We define the *minimum in-degree* $\delta^-(D)$ and the *maximum in-degree* $\Delta^-(D)$ similarly. The *minimum degree* of D is defined by $\delta(D) := \min\{d^+_D(x) + d^-_D(x) : x \in V(D)\}$ and its *minimum semi-degree* by $\delta^0(D) := \min\{\delta^+(D), \delta^-(D)\}$. Whenever $X, Y \subseteq V(D)$ are disjoint, we write $e_D(X)$ for the number of edges of D having both endvertices in X, and $e_D(X, Y)$ for the number of edges of D with tail in X and head in Y. We write $N_D^+(X) := \bigcup_{x\in X} N_D^+(x)$ and define $N_D^-(X)$ similarly. In all these definitions we often omit the subscript D if the digraph D is clear from the context.

A digraph D is *strongly connected* if for all $x, y \in V(D)$, there is a directed path in D from x to y. Given $k \in \mathbb{N}$, we say a digraph is *strongly k-connected* if $|D| > k$ and for every $S \subseteq V(D)$ of size at most $k-1$, $D - S$ is strongly connected. We say a digraph D is *k-linked* if $|D| \geq 2k$ and whenever $x_1, \dots, x_k, y_1, \dots, y_k$ are $2k$ distinct vertices of D, there exist vertex-disjoint paths $P_1, \dots, P_k$ such that P_i is a path from x_i to y_i.

Given a digraph D and sets $X, Y \subseteq V(D)$, we say that X *in-dominates* Y if each vertex in Y is an in-neighbour of some vertex in X. Similarly, we say that X *out-dominates* Y if each vertex in Y is an out-neighbour of some vertex in X.

A tournament T is *transitive* if there exists an ordering $v_1, \dots, v_n$ of its vertices such that $v_iv_j \in E(T)$ if and only if $i < j$. In this case, we often say that v_1 is the *tail* of T and v_n is the *head* of T.

Given $k \in \mathbb{N}$, we write $[k] := \{1, \dots, k\}$. We write log for the binary logarithm and $\log^2 n := (\log n)^2$.

3.3 Sketch of the proof of Theorem 3.1.2

In this section, we give an outline of the proof of Theorem 3.1.2. An important idea is the notion of a 'covering edge'. Given a small (pre-determined) set S of vertices in a tournament T, this will mean that it will suffice to find a cycle covering all vertices of $T - S$. More precisely, let T be a tournament, let $x \in V(T)$, and suppose C is a cycle in T covering $T - x$. If $yz \in E(C)$ and $yx, xz \in E(T)$, then we can replace yz by yxz in C to turn C into a Hamilton cycle. We call yz a *covering edge* for x. More generally, if $S \subseteq V(T)$ and C is a cycle in T spanning $V(T) - S$ such that C contains a covering edge for each $x \in S$, then we can turn C into a Hamilton cycle by using all these covering edges. Note that this idea still works if C covers some part of S. On the other hand, note that S needs to be fixed at the beginning – this is different than in the recently popularized 'absorbing method' (see e.g. [53, 84]).

Another important tool will be the following consequence of the Gallai-Milgram theorem: suppose that G is an oriented graph on n vertices with $\delta(G) \geq n - \ell$. Then the vertices of G can be covered with ℓ vertex-disjoint paths. We use this as follows: suppose we are given a highly linked tournament T and have already found i edge-disjoint Hamilton cycles in T. Then the Gallai-Milgram theorem implies that we can cover the vertices of the remaining oriented graph by a set of $2i$ vertex-disjoint paths. Very roughly, the aim is to link together these paths using the high linkedness of the original tournament T.

To achieve this aim, we introduce and use the idea of 'transitive dominating sets'. Here a transitive out-dominating set A_ℓ has the following properties:

- A_ℓ out-dominates $V(T) \setminus A_\ell$, i.e. every vertex of $V(T) \setminus A_\ell$ receives an edge from A_ℓ.
- A_ℓ induces a transitive tournament in T.

Transitive in-dominating sets B_ℓ are defined similarly.

Now suppose that we have already found i edge-disjoint Hamilton cycles in a highly linked tournament T. Let T' be the oriented subgraph of T obtained by removing the edges of these Hamilton cycles. Suppose that we also have the following 'linked dominating structure' in T', which consists of:

- small disjoint transitive out-dominating sets $A_1, \ldots, A_t$, where $t := 2i + 1$;
- small disjoint transitive in-dominating sets $B_1, \ldots, B_t$;

- a set of short vertex-disjoint paths $P_1, \dots, P_t$, where each P_ℓ is a path from the head b_ℓ of B_ℓ to the tail a'_ℓ of A_ℓ.

Recall that the head of a transitive tournament is the vertex of out-degree zero and the tail is defined analogously. The paths P_ℓ are found at the outset of the proof, using the assumption that the original tournament T is highly linked. (Note that T' need not be highly linked.)

Let A^* denote the union of the A_i and let B^* denote the union of the B_i. Note that $\delta(T' - A^* \cup B^*) \geq |T' - (A^* \cup B^*)| - 1 - 2i = |T' - (A^* \cup B^*)| - t$. So the Gallai-Milgram theorem implies that we can cover the vertices of $T' - A^* \cup B^*$ with t vertex-disjoint paths $Q_1, \dots, Q_t$. Now we can link up successive paths using the above dominating sets as follows. The final vertex of Q_1 sends an edge to some vertex b in B_2 (since B_2 is in-dominating). Either b is equal to the head b_2 of B_2 or there is an edge in $T'[B_2]$ from b to b_2 (since $T'[B_2]$ is a transitive tournament). Now follow the path P_2 from b_2 to the tail a'_2 of A_2. Using the fact that $T'[A_2]$ is transitive and that A_2 is out-dominating, we can similarly find a path of length at most two from a'_2 to the initial vertex of Q_2. Continuing in this way, we can link up all the paths Q_ℓ and P_ℓ into a single cycle C which covers all vertices outside $A^* \cup B^*$ (and some of the vertices inside $A^* \cup B^*$). The idea is illustrated in Figure 3.1.

In our construction, we will ensure that the paths P_ℓ contain a set of covering edges for $A^* \cup B^*$. So C also contains covering edges for $A^* \cup B^*$, and so we can transform C into a Hamilton cycle as discussed earlier.

A major obstacle to the above strategy is that in order to guarantee the P_ℓ in $T' - A^* \cup B^*$, we would need the linkedness of T to be significantly larger than $|A^* \cup B^*|$ (and thus larger than $|A_\ell|$). However, there are many tournaments where any in- or out-dominating set contains $\Omega(\log n)$ vertices (consider a random tournament). This leads to a linkage requirement on T which depends on n (and not just on k, as required in Theorem 3.1.2).

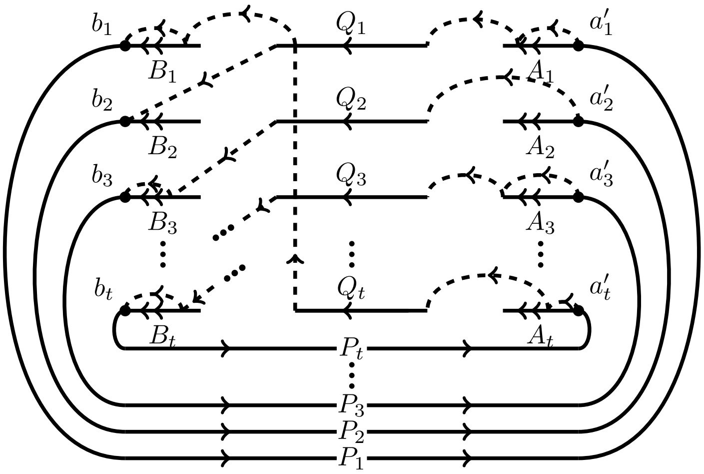

Figure 3.1: Illustrating the paths Q_i and P_i as well as the edges linking them up via the linked domination structure.

We overcome this problem by considering 'almost dominating sets': instead of out-dominating all vertices outside A_ℓ, the A_ℓ will out-dominate almost all vertices outside A_ℓ. (Analogous comments apply to the in-dominating sets B_ℓ.) This means that we have a small 'exceptional set' E of vertices which are not out-dominated by all of the A_ℓ. The problem with allowing an exceptional set is that if the tail of a path Q_ℓ in our cover is in the exceptional set E, we cannot extend it directly into the out-dominating set A_ℓ as in the above description. However, if we make sure that the A_ℓ include the vertices of smallest in-degree of T, we can deal with this issue. Indeed, in this case we can show that every vertex $v \in E$ has in-degree $d^-(v) > 2|E|$ say, so we can always extend the tail of a path out of the exceptional set if necessary (and then into an almost out-dominating set A_ℓ as before). Unfortunately, we may 'break' one of the paths P_ℓ in the process. However, if we are careful about the place where we break it and construct some 'spare' paths at the outset, it turns out that the above strategy can be made to work.

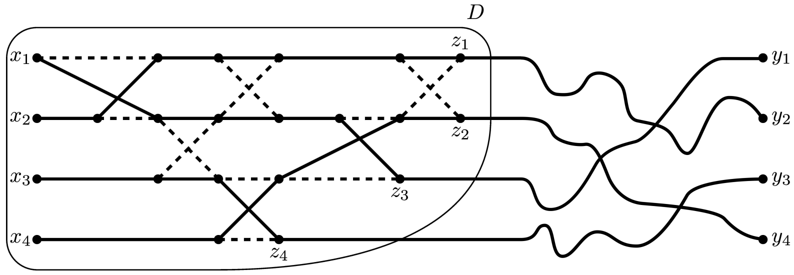

Figure 3.2: Illustrating our construction of a digraph D which corresponds to a sorting network for $k = 4$. D is used to link x_i to y_i. In the notation of the proof of Theorem 3.1.3, we have $\pi(3) = 1$.

3.4 Connectivity and linkedness in tournaments

In this section we give the proof of Theorem 3.1.3. We will also collect some simple properties of highly linked directed graphs which we will use later on. The proof of Theorem 3.1.3 is based on an important result of Ajtai, Komlós and Szemerédi [1, 2] on sorting networks. Roughly speaking, the proof idea of Theorem 3.1.3 is as follows. Suppose that we are given a highly connected tournament T and we want to link an ordered set X of k vertices to a set Y of the same size. Then we construct the equivalent of a sorting network D inside $T - Y$ with 'initial vertices' in X and 'final vertices' in a set Z. The high connectivity of T guarantees an 'unsorted' set of k ZY-paths which avoid the vertices in $D - Z$. One can then extend these paths via D to the appropriate vertices in X. In this way, we obtain paths linking the vertices in X to the appropriate ones in Y. An example is shown in Figure 3.2.

We now introduce the necessary background on non-adaptive sorting algorithms and sorting networks; see [52] for a more detailed treatment. In a sorting problem, we are given k registers $R_1, \dots, R_k$, and each register R_i is assigned a distinct element from $[k]$, which we call the *value* of R_i; thus there is some permutation π of $[k]$ such that value i has been assigned to register $R_{\pi(i)}$. Our task is to sort the values into their corresponding registers (so that value i is assigned to R_i) by making a sequence of comparisons: a comparison entails taking two registers and reassigning their values so that the higher value is assigned to the higher register and the lower value to the lower register. A non-adaptive sorting algorithm is a sequence of comparisons specified in advance such that for any initial assignment of k values to k registers, applying the prescribed sequence of comparisons results in every value being assigned to its corresponding register.

Ajtai, Komlós and Szemerédi [1, 2] proved, via the construction of sorting networks, that there exists an absolute constant C' and a non-adaptive sorting algorithm (for k registers and values) that requires $C'k \log k$ comparisons, and this is asymptotically best possible. It is known that we can take $C' := 3050$ [74] (results of this type are often stated in terms of the *depth* of a sorting network rather than the number of comparisons).

The next theorem is a consequence of the above. Before we can state it, we first need to introduce some notation. A comparison c, which is part of some non-adaptive sorting algorithm for k registers, will be denoted by $c = (s; t)$, where $1 \leq s < t \leq k$, to indicate that c is a comparison in which the values of registers R_s and R_t are compared (and sorted so the higher value is assigned to the higher register).

Theorem 3.4.1 *(see [1, 2, 74]) Let $C' := 3050$ and $k \in \mathbb{N}$ be such that $k \geq 2$. Then there exist $r \leq C'k \log k$ and a sequence of comparisons $c_1, \dots, c_r$ satisfying the following property: for any initial assignment of k values to k registers, applying the comparisons in sequence results in register R_i being assigned the value i for all $i \in [k]$.*

We now show how to obtain a structure within a highly connected tournament that simulates the function of a non-adaptive sorting algorithm. Each comparison in the sorting algorithm will be simulated by a 'switch', which we now define. An (a_1, a_2)-*switch* is a digraph D on 5 distinct vertices a_1, a_2, b, b_1, b_2, where either $E(D) = \{a_1b, bb_1, bb_2, a_2b_1, a_2b_2\}$ or $E(D) = \{a_2b, bb_1, bb_2, a_1b_1, a_1b_2\}$. We call b_1 and b_2 the *terminal vertices* of the (a_1, a_2)-switch. Note that for any permutation π of $\{1, 2\}$, there exist vertex-disjoint paths P_1, P_2 of D such that P_i joins a_i to $b_{\pi(i)}$ for $i = 1, 2$.

Proposition 3.4.2 *Let T be a tournament. Given distinct vertices $a_1, a_2 \in V(T)$, if $d^+_T(a_1), d^+_T(a_2) \geq 7$, then T contains an (a_1, a_2)-switch.*

Proof. We may choose disjoint sets $A_1 \subseteq N^+_T(a_1) \setminus \{a_2\}$ and $A_2 \subseteq N^+_T(a_2) \setminus \{a_1\}$ with $|A_1| = |A_2| = 3$. Consider the bipartite digraph H induced by T between A_1 and A_2. It is easy to check that there exists $b \in A_1 \cup A_2$ with $d^+_H(b) \geq 2$. Let b_1 and b_2 be two out-neighbours of b in H. Now the vertices a_1, a_2, b, b_1, b_2 with suitably chosen edges from T form an (a_1, a_2)-switch (with terminal vertices b_1 and b_2). □

Given $k \in \mathbb{N}$, we write S_k for the set of permutations of $[k]$ and id_k for the identity permutation of $[k]$. The following structural lemma for tournaments is at the heart of the proof of Theorem 3.1.3. It constructs the equivalent of a sorting network in a tournament of high minimum outdegree.

Lemma 3.4.3 *Let $C' := 3050$ and $k \in \mathbb{N}$ be such that $k \geq 2$. Let T be a tournament with $\delta^+(T) \geq (3C' + 5)k \log k$, and let $x_1, \ldots, x_k \in V(T)$ be distinct vertices. Then there exists a digraph $D \subseteq T$ and distinct vertices $z_1, \ldots, z_k \in V(D)$ with the following properties:*

(i) $x_1, \ldots, x_k \in V(D)$.

(ii) $|D| \leq (3C' + 1)k \log k$.

(iii) *For any $\pi \in S_k$, we can find vertex-disjoint paths $P_1, \ldots, P_k$ such that P_i joins $x_{\pi(i)}$ to z_i for all $i \in [k]$.*

Proof. Consider the sorting problem for k registers, and apply Theorem 3.4.1 to obtain a sequence $c_1, \ldots, c_r$ of $r \leq C'k \log k$ comparisons such that for any $\pi \in S_k$, if value i is initially assigned to register $R_{\pi(i)}$, then applying the comparisons $c_1, \ldots, c_r$ results in every value being assigned to its corresponding register. Given $\pi \in S_k$, we write $\pi_q \in S_k$ for the permutation such that after applying the first q comparisons $c_1, \ldots, c_q$, value i is assigned to register $R_{\pi_q(i)}$ for all i; thus $\pi_r = id_k$.

Let D_0 be the digraph with vertex set $\{x_1, \dots, x_k\}$ and empty edge set. We inductively construct digraphs $D_0 \subseteq D_1 \subseteq \cdots \subseteq D_r \subseteq T$ and for each D_q we maintain a set $Z_q = \{z_1^q, \dots, z_k^q\}$ of k distinct *final vertices* such that the following holds:

(a) $|D_q| = 3q + k$.

(b) Whenever $\pi \in S_k$ is a permutation, there exist vertex-disjoint paths $P_1^q, \dots, P_k^q$ in D_q such that P_i^q joins $x_{\pi(i)}$ to $z_{\pi_q(i)}^q$ for all $i \in [k]$.

Assuming the above statement holds for $q = 0, \dots, r$, then taking $D := D_r$ with $z_i := z_i^r$ for all $i \in [k]$ proves the lemma. Indeed $|D_r| = 3r + k \le 3C'k \log k + k \le (3C'+1)k \log k$ and $\pi_r = id_k$.

Having already defined D_0, let us describe the inductive step of our construction. Suppose that for some $q \in [r]$ we have constructed $D_{q-1} \subseteq T$ and a corresponding set $Z_{q-1} = \{z_1^{q-1}, \dots, z_k^{q-1}\}$ of final vertices. Let $s, t \in [k]$ with $s < t$ be such that $c_q = (s; t)$. Define the tournament $T' := T - (V(D_{q-1}) \setminus \{z_s^{q-1}, z_t^{q-1}\})$. Then T' has minimum out-degree at least

$$(3C' + 5)k \log k - |D_{q-1}| \ge (3C' + 5)k \log k - 3r - k \ge 5k \log k - k \ge 7,$$

and so in particular $d_{T'}^+(z_s^{q-1}), d_{T'}^+(z_t^{q-1}) \ge 7$. Thus we may apply Proposition 3.4.2 to obtain a (z_s^{q-1}, z_t^{q-1})-switch σ in T'. Write b_1, b_2 for the terminal vertices of σ. Now D_q is constructed from D_{q-1} by adding the vertices and edges of σ to D_{q-1}; note that z_s^{q-1} and z_t^{q-1} are precisely the common vertices of D_{q-1} and σ. We define the set $Z_q = \{z_1^q, \dots, z_k^q\}$ by setting $z_i^q := z_i^{q-1}$ for all $i \ne s, t$ and $z_s^q := b_1$ as well as $z_t^q := b_2$. Note that $z_1^q, \dots, z_k^q$ are distinct.

Finally we check that conditions (a) and (b) hold for D_q. Condition (a) holds since D_q has exactly 3 more vertices than D_{q-1}. For (b), by induction we may assume that there are vertex-disjoint paths $P_1^{q-1},\dots,P_k^{q-1}$ in D_{q-1} such that P_i^{q-1} joins $x_{\pi(i)}$ to $z^{q-1}_{\pi_{q-1}(i)}$ for all $i \in [k]$. Choose vertex-disjoint paths Q_s and Q_t in σ such that

- if c_q swaps values in registers R_s and R_t, then Q_s joins z_s^{q-1} to z_t^q and Q_t joins z_t^{q-1} to z_s^q;
- if c_q does not swap values in registers R_s and R_t, then Q_s joins z_s^{q-1} to z_s^q and Q_t joins z_t^{q-1} to z_t^q.

Now exactly two of the paths from $P_1^{q-1},\dots,P_k^{q-1}$ end at z_s^{q-1} and z_t^{q-1}, namely those indexed by $\pi_{q-1}^{-1}(s)$ and $\pi_{q-1}^{-1}(t)$. We extend these two paths using Q_s and Q_t, and leave all others unchanged to obtain paths $P_1^q,\dots,P_k^q$. It is straightforward to check that these paths are vertex-disjoint and that P_i joins $x_{\pi(i)}$ to $z^q_{\pi_q(i)}$ for all $i \in [k]$. □

It is now an easy step to prove Theorem 3.1.3. We will use the following directed version of Menger's Theorem.

Theorem 3.4.4 (Menger's Theorem) *Suppose D is a strongly k-connected digraph with $A, B \subseteq V(D)$ and $|A|, |B| \geq k$. Then there exist k vertex-disjoint paths in D each starting in A and ending in B.*

Proof of Theorem 3.1.3. Set $C' := 3050$ and $C := 3C' + 6 < 10^4$. We must show that, given a strongly $Ck\log k$-connected tournament T and distinct vertices $x_1,\dots,x_k$, $y_1,\dots,y_k \in V(T)$, we can find vertex-disjoint paths $R_1,\dots,R_k$ such that R_i joins x_i to y_i for all $i \in [k]$.

Let $X := \{x_1,\dots,x_k\}$, $Y := \{y_1,\dots,y_k\}$ and $T' := T - Y$. Note that T' is strongly $(3C'+5)k\log k$-connected, and in particular $\delta^+(T'') \geq (3C'+5)k\log k$. Thus we can apply Lemma 3.4.3 to T' and $x_1,\dots,x_k$ to obtain a digraph $D \subseteq T'$ and vertices $z_1,\dots,z_k \in V(D)$ satisfying properties (i)–(iii) of Lemma 3.4.3. Let $Z := \{z_1,\dots,z_k\}$. Since $|D| \leq (3C'+1)k\log k$, the tournament $T'' := T - (V(D) \setminus Z)$ is strongly k-connected. Therefore, by Theorem 3.4.4, there exist k vertex-disjoint paths, with each path starting in Z and ending in Y. For each $i \in [k]$, let us assume that $P_{\pi(i)}$ is the path that joins z_i to $y_{\pi(i)}$, where π is some permutation of $[k]$. By Lemma 3.4.3, we can find vertex-disjoint paths $Q_1, \dots, Q_k$ in D such that Q_i joins $x_{\pi(i)}$ to z_i. Then the path $R_i := Q_{\pi^{-1}(i)} P_{\pi^{-1}(i)}$ joins x_i to y_i and these paths are vertex-disjoint. $\square$

Batcher [9] (see also [52]) gave a construction of sorting networks which is asymptotically not optimal but which gives better values for small k. More precisely, it uses at most $2k \log^2 k$ comparisons for $k \geq 3$. If we use these as a building block in the proof of Lemma 3.4.3 instead of the asymptotically optimal ones leading to Theorem 3.4.1, we immediately obtain the following result which improves Theorem 3.1.3 for small values of k.

Theorem 3.4.5 *For all $k \in \mathbb{N}$ with $k \geq 3$, every strongly $12k \log^2 k$-connected tournament is k-linked.*

For $k = 2$, the best bound is obtained by a result of Bang-Jensen [5], who showed that every strongly 5-connected semi-complete digraph is 2-linked, which is best possible even for tournaments.

We will now collect some simple properties of highly linked directed graphs which we will use later on. The first two follow straightforwardly from the definition of linkedness.

Proposition 3.4.6 *Let $k \in \mathbb{N}$. Then a digraph D is k-linked if and only if $|D| \geq 2k$ and whenever $(x_1, y_1), \dots, (x_k, y_k)$ are ordered pairs of (not necessarily distinct) vertices of D, there exist internally disjoint paths $P_1, \dots, P_k$ such that P_i joins x_i to y_i.*

Proposition 3.4.7 *Let $k, \ell \in \mathbb{N}$ with $\ell < k$, and let D be a k-linked digraph. Let $X \subseteq V(D)$ and $F \subseteq E(D)$ be such that $|X| + 2|F| \leq 2\ell$. Then $D - X - F$ is $(k - \ell)$-linked.*

The next lemma shows that in a sufficiently highly linked digraph we can link given pairs of vertices by vertex-disjoint paths which together do not contain too many vertices.

Lemma 3.4.8 *Let $k, s \in \mathbb{N}$, and let D be a $2ks$-linked digraph. Let $(x_1, y_1), \dots, (x_k, y_k)$ be ordered pairs of (not necessarily distinct) vertices in D. Then there exist internally disjoint paths $P_1, \dots, P_k$ such that P_i joins x_i to y_i for all $i \in [k]$ and $|P_1 \cup \dots \cup P_k| \leq |D|/s$.*

Proof. By Proposition 3.4.6 there exist internally disjoint paths $P_1^1, \dots, P_k^{2s}$ such that P_i^j joins x_i to y_i for all $i \in [k]$ and all $j \in [2s]$. For any j, the interiors of $P_1^j, \dots, P_k^j$ contain at least $|P_1^j \cup \dots \cup P_k^j| - 2k$ vertices. So the disjointness of the paths implies that there is a $j \in [2s]$ with $|P_1^j \cup \dots \cup P_k^j| - 2k \leq |D|/2s$. The result now follows by setting $P_i := P_i^j$ and noting that $2k \leq |D|/2s$. □

3.5 Nearly extremal example

The aim of this section is to prove the following proposition, which shows that the bound on the connectivity in Theorem 3.1.1 is close to best possible.

Proposition 3.5.1 *Fix $n, k \in \mathbb{N}$ with $k \geq 2$ and $n > k^2 + k + 2$. There exists a strongly $\lfloor k^2/4 \rfloor$-connected tournament T of order n such that if $D \subseteq T$ is a spanning r-regular subdigraph, then $r \leq k$. In particular, T contains at most k edge-disjoint Hamilton cycles.*

It is easy to see that the above tournament T is also $\Omega(k^2)$-linked. This shows that the bound in Theorem 3.1.2 has to be at least quadratic in k.

Proof. Let $\ell \in \mathbb{N}$. We will first describe a tournament $T_\ell = (V_\ell, E_\ell)$ of order $2\ell+1$ which is strongly ℓ-connected. We then use T_ℓ as a building block to construct a tournament as desired in the proposition.

Let $V_\ell := \{v_0, \dots, v_{2\ell}\}$ and let E_ℓ consist of the edges $v_i v_{i+t}$ for all $i = 0, \dots, 2\ell$ and all $t \in [\ell]$, where indices are understood to be modulo $2\ell + 1$. One may think of T_ℓ as the tournament with vertices $v_0, \dots, v_{2\ell}$ placed in order, clockwise, around a circle, where the out-neighbours of each v_i are the ℓ closest vertices to v_i in the clockwise direction, and the in-neighbours are the ℓ closest vertices in the anticlockwise direction. Note that T_ℓ is regular. Note also that, for any distinct $x, y \in V_\ell$, we can find a path in T_ℓ from x to y by traversing vertices from x to y in clockwise order; this remains true even if we delete any $\ell - 1$ vertices from T_ℓ.

Next we construct a tournament $T_{m,\ell} = (V_{m,\ell}, E_{m,\ell})$ as follows. We take $V_{m,\ell}$ to be the disjoint union of sets $A_\ell := \{a_0, \dots, a_{2\ell}\}$, $B_\ell := \{b_0, \dots, b_{2\ell}\}$, and $C_m := \{c_1, \dots, c_m\}$. The edges of $T_{m,\ell}$ are defined as follows: $T_{m,\ell}[A_\ell]$ and $T_{m,\ell}[B_\ell]$ are isomorphic to T_ℓ (with the natural labelling of vertices), and $T[C_m]$ is a transitive tournament which respects the given order of the vertices in C_m (i.e. $c_i c_j$ is an edge if and only if $i < j$). Each vertex in A_ℓ is an in-neighbour of all vertices in C_m, and each vertex in B_ℓ is an out-neighbour of all vertices in C_m. Finally, a vertex $a_i \in A_\ell$ is an in-neighbour of a vertex $b_j \in B_\ell$ if and only if $i \neq j$. Note that $|T_{m,\ell}| = m + 4\ell + 2$.

Claim 1. *The tournament $T_{m,\ell}$ is strongly ℓ-connected.*

To see that $T_{m,\ell}$ is strongly ℓ-connected, we check that if $S \subseteq V_{m,\ell}$ with $|S| \leq \ell - 1$, then $T_{m,\ell} - S$ is strongly connected. Write A'_ℓ, B'_ℓ and C'_m respectively for $A_\ell \setminus S$, $B_\ell \setminus S$, and $C_m \setminus S$. Note that there is at least one edge of $T_{m,\ell} - S$ from B'_ℓ to A'_ℓ, which we may assume by symmetry to be $b_0 a_0$. Ordering the vertices of $T_{m,\ell}$ as $a_0, \dots, a_{2\ell}, c_1, \dots, c_m, b_1, \dots, b_{2\ell}, b_0$ and removing the vertices of S from this ordering gives a Hamilton cycle in $T_{m,\ell} - S$. Thus $T_{m,\ell} - S$ must be strongly connected. This completes the proof of Claim 1.

Claim 2. *Let $m, \ell \in \mathbb{N}$ be such that $m > \sqrt{4\ell}$. Then for every r-regular spanning subdigraph $D \subseteq T_{m,\ell}$ we have $r \leq \sqrt{4\ell}$.*

Suppose for a contradiction that $D \subseteq T_{m,\ell}$ is an r-regular spanning subdigraph with $r := \lfloor \sqrt{4\ell} \rfloor + 1 > \sqrt{4\ell}$. Since D is regular, we have $e_D(A_\ell, \bar{A}_\ell) = e_D(\bar{A}_\ell, A_\ell)$, where $\bar{A}_\ell := V(D) \setminus A_\ell$. Noting that $r \leq m$, consider the first r vertices $c_1, \dots, c_r$ of C_m. Since $N^-_D(c_i) \subseteq N^-_{T_{m,\ell}}(c_i) = A_\ell \cup \{c_1, \dots, c_{i-1}\}$ and $|N^-_D(c_i)| = r$, we have $|N^-_D(c_i) \cap A_\ell| \geq r - i + 1$, so that $e_D(A_\ell, \{c_i\}) \geq r - i + 1$. Thus

$$e_D(\bar{A}_\ell, A_\ell) = e_D(A_\ell, \bar{A}_\ell) \geq e(A_\ell, \{c_1, \ldots, c_r\}) \geq r + \cdots + 1 = \binom{r+1}{2}.$$

But $e_D(\bar{A}_\ell, A_\ell) \leq e_{T_{m,\ell}}(\bar{A}_\ell, A_\ell) = 2\ell+1$, so $\binom{r+1}{2} \leq 2\ell+1$. This is easily seen to contradict $r > \sqrt{4\ell}$ for all $\ell \in \mathbb{N}$. This completes the proof of Claim 2.

To prove the proposition, we set $\ell := \lfloor k^2/4 \rfloor$ and $m := n - 4\ell - 2$, and take T to be $T_{m,\ell}$. Thus $|T| = |T_{m,\ell}| = m + 4\ell + 2 = n$. By Claim 1, T is strongly $\lfloor k^2/4 \rfloor$-connected. Since $n > k^2+k+2 \geq 4\ell+\sqrt{4\ell}+2$, we have $m > \sqrt{4\ell}$, so Claim 2 implies that if $D \subseteq T = T_{m,\ell}$ is a spanning r-regular subdigraph, then $r \leq \sqrt{4\ell} \leq k$. $\square$

As mentioned in the introduction, Thomassen [89] observed that no lower bound on the strong edge-connectivity of a tournament can guarantee two edge-disjoint Hamilton cycles. (Recall that a digraph D is strongly k-edge-connected if $|D| \geq 2$ and if for every $S \subseteq E(D)$ of size at most $k-1$, $T - S$ is strongly connected.) Here, for completeness, we provide an explicit example for Thomassen's observation.

Let $T = (V, E)$ be a tournament where V is the disjoint union of three sets, A, B, and $C = \{x_1, x_2\}$, and where x is a distinguished vertex of B. We choose $T[A]$ and $T[B]$ to be any strongly k-edge-connected tournament and let $x_1x_2 \in E(T)$. All edges between A and C are directed from A to C; all edges between B and C are directed from C to B; and all edges between A and B are directed from A to B except edges between A and x, which are directed from x to A.

It is easy to check that T is strongly k-edge-connected and that all Hamilton cycles in T use the edge x_1x_2. Hence there are tournaments with arbitrarily high strong edge-connectivity but with no two edge-disjoint Hamilton cycles.

3.6 Finding a single Hamilton cycle in suitable oriented graphs

We first state two simple, well-known facts concerning the degree sequences of tournaments.

Proposition 3.6.1 *Let T be a tournament on n vertices. Then T contains at least one vertex of in-degree at most $n/2$, and at least one vertex of out-degree at most $n/2$.*

Proposition 3.6.2 *Let T be a tournament on n vertices and let $d \geq 0$. Then T has at most $2d+1$ vertices of in-degree at most d, and at most $2d+1$ vertices of out-degree at most d.*

We will also use the following well-known result due to Gallai and Milgram (see for example [18]). (The *independence number* of a digraph T is the maximal size of a set $X \subseteq V(T)$ such that $T[X]$ contains no edges.)

Theorem 3.6.3 *Let T be a digraph with independence number at most k. Then T has a path cover consisting of at most k paths.*

The following corollary is an immediate consequence of Theorem 3.6.3.

Corollary 3.6.4 *Let T be an oriented graph on n vertices with $\delta(T) \geq n-k$. Then T has a path cover consisting of at most k paths.*

Given a digraph T, we define a *covering edge* for a vertex v to be an edge xy of T such that $xv, vy \in E(T)$. We call xv and vy the *activating edges* of xy. Note that if xy is a covering edge for v and C is a cycle in T containing xy but not v, we can form a new cycle C' with $V(C') = V(C) \cup \{v\}$ by replacing xy with xvy in C. We will see in Section 3.8 that covering edges are easy to find in strongly 2-connected tournaments.

Recall that, given a path system $\mathcal{P}$, we write $h(\mathcal{P})$ for the set of heads of paths in $\mathcal{P}$ and $t(\mathcal{P})$ for the set of tails of paths in $\mathcal{P}$. If $v \in V(\mathcal{P})$, we write v^+ and v^- respectively for the successor and predecessor of v on the path in $\mathcal{P}$ containing v.

The following lemma allows us to take a path cover $\mathcal{P}$ of a digraph and modify it into a path cover $\mathcal{P}'$ with no heads in some "bad" set I, without adding any heads or tails in $I \cup J$ for some other "bad" set J. Moreover, we can do this without losing any edges in some "good" set $F \subseteq E(\mathcal{P})$, and without altering too many paths in $\mathcal{P}$. In our applications, F will consist of covering edges. We require that every vertex in I has high out-degree.

Lemma 3.6.5 *Let T be a digraph. Let $I, J \subseteq V(T)$ be disjoint. Let $\mathcal{P} = \mathcal{P}_1 \dot\cup \mathcal{P}_2$ be a path cover of T satisfying $h(\mathcal{P}_2) \cap I = \emptyset$. Let $F \subseteq E(\mathcal{P})$. Suppose $d^+(v) > 3(|I| + |J|) + 2|F|$ for all $v \in I$. Then there exists a path cover $\mathcal{P}'$ of T satisfying the following properties:*

(i) $h(\mathcal{P}') \cap I = \emptyset$.

(ii) $h(\mathcal{P}') \cap J = h(\mathcal{P}) \cap J$.

(iii) $t(\mathcal{P}') \cap (I \cup J) = t(\mathcal{P}) \cap (I \cup J)$.

(iv) $F \subseteq E(\mathcal{P}')$.

(v) $|\mathcal{P}'| \leq |\mathcal{P}| + |\mathcal{P}_1|$.

(vi) $|\mathcal{P}' \cap \mathcal{P}_2| \geq |\mathcal{P}_2| - |\mathcal{P}_1|$.

If in addition $d^+(v) > 3(|I| + |J|) + 2|F| + |V(\mathcal{P}_2)|$ for all $v \in I$, then we may strengthen (vi) to $\mathcal{P}_2 \subseteq \mathcal{P}'$.

Proof. We will use the degree condition on the vertices in I in the hypothesis to repeatedly extend paths with heads in I out of I, breaking other paths in $\mathcal{P}$ as a result. We must ensure that we do not create new paths with endpoints in $I \cup J$ in the process.

Let $r := |\mathcal{P}_1|$ and $\mathcal{P}^0 := \mathcal{P}$. We shall find path covers $\mathcal{P}^1, \dots, \mathcal{P}^r$ of T such that the following properties hold for all $0 \leq i \leq r$:

(P1) $|h(\mathcal{P}^i) \cap I| \leq r - i$.

(P2) $h(\mathcal{P}^i) \cap J = h(\mathcal{P}) \cap J$.

(P3) $t(\mathcal{P}^i) \cap (I \cup J) = t(\mathcal{P}) \cap (I \cup J)$.

(P4) $F \subseteq E(\mathcal{P}^i)$.

(P5) $|\mathcal{P}^i| \leq |\mathcal{P}| + i$.

(P6) $|\mathcal{P}^i \cap \mathcal{P}_2| \geq |\mathcal{P}_2| - i$.

If this is possible, we may then take $\mathcal{P}' := \mathcal{P}^r$.

By hypothesis, $\mathcal{P}^0$ satisfies (P1)–(P6). So suppose we have found $\mathcal{P}^0, \dots, \mathcal{P}^{i-1}$ for some $i \in [r]$. We then form $\mathcal{P}^i$ as follows. If $|h(\mathcal{P}^{i-1}) \cap I| \leq r - i$, we simply let $\mathcal{P}^i := \mathcal{P}^{i-1}$. Otherwise, let $P \in \mathcal{P}^{i-1}$ be a path with head $v \in I$. We will form $\mathcal{P}^i$ by extending the head v of P and breaking the path in $\mathcal{P}^{i-1}$ which P now intersects into two subpaths. Define

$$X := \{x \in V(T) : \{x^+, x, x^-\} \cap (I \cup J) \neq \emptyset\}.$$

We have

$$d^+(v) > 3(|I| + |J|) + 2|F| \geq |X| + |V(F)| \geq |X \cup V(F)|,$$

and so there exists $w \in N^+(v) \setminus (X \cup V(F))$. Let Q be the path in $\mathcal{P}^{i-1}$ containing w (note that we may have $Q = P$). Split Q into (at most) two paths and an isolated vertex by removing any of the edges w^-w, ww^+ that exist, and let $\mathcal{P}^*$ be the set of paths obtained from $\mathcal{P}^{i-1}$ in this way. Let P^* be the path in $\mathcal{P}^*$ containing v. (Note that $P^* = P$ unless $w \in V(P)$.) We then form $\mathcal{P}^i$ by replacing P^* by P^*vw in $\mathcal{P}^*$.

First suppose $w \in \mathrm{Int}(Q)$. Then $\mathcal{P}^i$ is a path cover of T such that

$$h(\mathcal{P}^i) = (h(\mathcal{P}^{i-1}) \setminus \{v\}) \cup \{w, w^-\} \qquad \text{and} \qquad t(\mathcal{P}^i) = t(\mathcal{P}^i) \cup \{w^+\}.$$

Since $w \notin X$, we have $w, w^- \notin I$ and hence

$$|h(\mathcal{P}^i) \cap I| = |h(\mathcal{P}^{i-1}) \cap I| - 1 \leq r - i.$$

Thus (P1) holds. Similarly,

$$h(\mathcal{P}^i) \cap J = h(\mathcal{P}^{i-1}) \cap J = h(\mathcal{P}) \cap J,$$
$$t(\mathcal{P}^i) \cap (I \cup J) = t(\mathcal{P}^{i-1}) \cap (I \cup J) = t(\mathcal{P}) \cap (I \cup J),$$

and so (P2) and (P3) hold. By similar arguments, (P1)–(P3) also hold if w is an endpoint of Q. Since $w \notin V(F)$ and $F \subseteq E(\mathcal{P}^{i-1})$ we have $F \subseteq E(\mathcal{P}^i)$ and (P4) holds. (P5) holds too since $|\mathcal{P}^i| \leq |\mathcal{P}^{i-1}| + 1$. Finally, we have altered at most two paths in $\mathcal{P}^{i-1}$. One of these had its head in I, so we have altered at most one path in $\mathcal{P}^{i-1} \cap \mathcal{P}_2$. Thus (P6) holds.

If in addition we have

$$d^+(v) > 3(|I| + |J|) + 2|F| + |V(\mathcal{P}_2)|,$$

then we may use almost exactly the same argument to prove the strengthened version of the result. Instead of choosing $w \in N^+(v) \setminus (X \cup V(F))$, we may choose $w \in N^+(v) \setminus (X \cup V(F) \cup V(\mathcal{P}_2))$. We also strengthen (P6) to the requirement that $\mathcal{P}_2 \subseteq \mathcal{P}^i$. The strengthened (P6) must hold in each step since we now have that $w \notin V(\mathcal{P}_2)$. $\square$

The following analogue of Lemma 3.6.5 for tails can be obtained by reversing the orientation of each edge of T.

Lemma 3.6.6 *Let T be a digraph. Let $I, J \subseteq V(T)$ be disjoint. Let $\mathcal{P} = \mathcal{P}_1 \dot\cup \mathcal{P}_2$ be a path cover of T satisfying $t(\mathcal{P}_2) \cap I = \emptyset$. Let $F \subseteq E(\mathcal{P})$. Suppose $d^-(v) > 3(|I| + |J|) + 2|F|$ for all $v \in I$. Then there exists a path cover $\mathcal{P}'$ of T satisfying the following properties:*

(i) $t(\mathcal{P}') \cap I = \emptyset$.

(ii) $t(\mathcal{P}') \cap J = t(\mathcal{P}) \cap J$.

(iii) $h(\mathcal{P}') \cap (I \cup J) = h(\mathcal{P}) \cap (I \cup J)$.

(iv) $F \subseteq E(\mathcal{P}')$.

(v) $|\mathcal{P}'| \leq |\mathcal{P}| + |\mathcal{P}_1|$.

(vi) $|\mathcal{P}' \cap \mathcal{P}_2| \geq |\mathcal{P}_2| - |\mathcal{P}_1|$.

If in addition $d^-(v) > 3(|I| + |J|) + 2|F| + |V(\mathcal{P}_2)|$ for all $v \in I$, then we may strengthen (vi) to $\mathcal{P}_2 \subseteq \mathcal{P}'$.

The following lemma is the main building block of the proof of Theorem 3.1.2. It will be applied repeatedly to find the required edge-disjoint Hamilton cycles. Roughly speaking, the lemma guarantees a Hamilton cycle provided that we have well-chosen disjoint (almost) dominating sets A_i and B_i which are linked by short paths containing covering edges for all vertices in these dominating sets. (This is the linked dominating structure described in Sections 3.1 and 3.3.) An additional assumption is that we have not removed too many edges of our tournament T already. In general, the statement and proof roughly follow the sketch in Section 3.3, with the addition of a set $X \subseteq V(T)$.

The role of X is as follows. The sets A_i and B_i in the lemma dominate only almost all vertices of T, so we have some small exceptional sets E_A and E_B of vertices which are not dominated. We will use Lemmas 3.6.5 and 3.6.6 to extend a certain path system out of these exceptional sets E_A and E_B. For this we need that the vertices in $E_A \cup E_B$ have relatively high in- and out-degree. But T may have vertices which do not satisfy this degree condition. When we apply Lemma 3.6.7, these problematic vertices will be the elements of X.

Lemma 3.6.7 *Let $C := 10^6$, $k \geq 20$, $t := 164k$, and $c := \lceil \log 50t + 1 \rceil$. Suppose that T is an oriented graph of order n satisfying $\delta(T) > n - 4k$ and $\delta^0(T) \geq Ck^2$. Suppose moreover that T contains disjoint sets of vertices $A_1, \dots, A_t$, $B_1, \dots, B_t$ and X, a matching F, and vertex-disjoint paths $P_1, \dots, P_t$ such that the following conditions hold, where $A^* := A_1 \cup \dots \cup A_t$ and $B^* := B_1 \cup \dots \cup B_t$:*

(i) $2 \leq |A_i| \leq c$ *for all* $i \in [t]$. *Moreover, $T[A_i]$ is a transitive tournament whose head has out-degree at least $n/3$ in T.*

(ii) There exists a set $E_A \subseteq V(T) \setminus (A^* \cup B^*)$, such that each A_i out-dominates $V(T) \setminus (A^* \cup B^* \cup E_A)$. Moreover, $|E_A| \leq d^-/40$, where $d^- := \min\{d^-_T(v) : v \in E_A \setminus X\}$.

(iii) $2 \leq |B_i| \leq c$ *for all* $i \in [t]$. *Moreover, $T[B_i]$ is a transitive tournament whose tail has in-degree at least $n/3$ in T.*

(iv) There exists a set $E_B \subseteq V(T) \setminus (A^* \cup B^*)$, such that each B_i in-dominates $V(T) \setminus (A^* \cup B^* \cup E_B)$. Moreover, $|E_B| \leq d^+/40$, where $d^+ := \min\{d^+_T(v) : v \in E_B \setminus X\}$.

(v) For all $i \in [t]$, P_i is a path from the head of $T[B_i]$ to the tail of $T[A_i]$ which is internally disjoint from $A^* \cup B^*$. Moreover, $|P_1 \cup \dots \cup P_t| \leq n/20$.

(vi) $F \subseteq E(P_1 \cup \dots \cup P_t)$ and $V(F) \cap (A^* \cup B^*) = \emptyset$. Moreover, $F = \{e_v : v \in A^* \cup B^*\}$, where e_v is a covering edge for v and $e_v \neq e_{v'}$ whenever $v \neq v'$. In particular, $|F| = |A^* \cup B^*| \leq 2ct$.

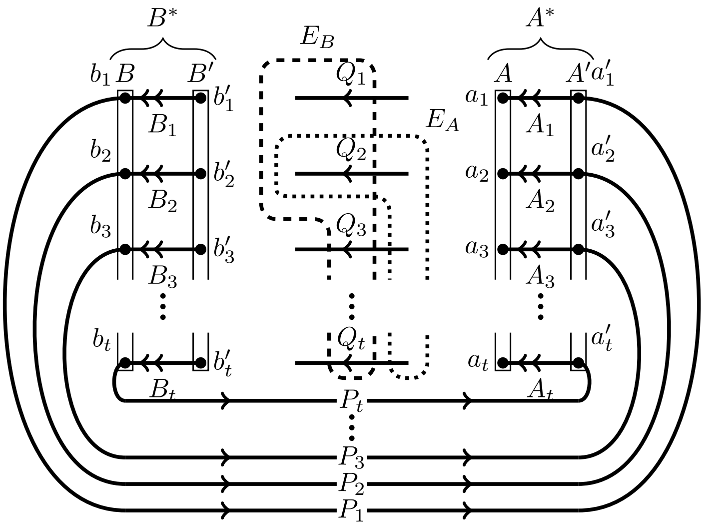

Figure 3.3: Our linked domination structure and path cover at the beginning of the proof of Lemma 3.7.2.

(vii) We have $X \subseteq V(P_1 \cup \cdots \cup P_t)$, $X \cap (A^* \cup B^*) = \emptyset$ and $|X| \leq 2kt$.

Then T contains a Hamilton cycle.

Proof. Without loss of generality, suppose that $d^- \leq d^+$. (Otherwise, reverse the orientation of every edge in T.) Write a_i for the head of $T[A_i]$ and a'_i for its tail. Similarly, write b_i for the head of $T[B_i]$ and b'_i for its tail. Let

$$A := \{a_1, \dots, a_t\}, \quad A' := \{a'_1, \dots, a'_t\}, \quad B := \{b_1, \dots, b_t\} \text{ and } B' := \{b'_1, \dots, b'_t\}.$$

Thus the sets A, A', B, B' are disjoint, and by condition (v) the paths P_i join B to A'. Let

$$N := V(T) \setminus (A^* \cup B^*), \quad T' := T[N \cup A' \cup B], \quad \text{and} \quad \mathcal{P}_2 := \{P_1, \dots, P_t\}.$$

By Corollary 3.6.4, there exists a path cover $\mathcal{P}_1$ of $N \setminus V(\mathcal{P}_2)$ with $|\mathcal{P}_1| \leq 4k$. Then $\mathcal{Q}_1 := \mathcal{P}_1 \dot\cup \mathcal{P}_2$ is a path cover of T'. The situation is illustrated in Figure 3.3.

Claim. *There exists an oriented graph T'' with $T' \subseteq T'' \subseteq T[V(T') \cup A \cup B']$ and a path cover $\mathcal{Q}$ of T'' such that the following properties hold:*

(Q1) $F \subseteq E(\mathcal{Q})$.

(Q2) $t(\mathcal{Q}) \cap E_A = \emptyset$.

(Q3) $h(\mathcal{Q}) \cap E_B = \emptyset$.

(Q4) $|\mathcal{Q} \cap \mathcal{P}_2| \geq |\mathcal{Q}_1| - 20k$.

(Q5) If a_i *or* b'_i *is in* $V(\mathcal{Q})$*, then* $P_i \notin \mathcal{Q}$.

(Q6) $|\mathcal{Q}| \leq |\mathcal{Q}_1| + 124k$.

(Q7) No paths in $\mathcal{Q} \setminus \mathcal{P}_2$ *have endpoints in* $A^* \cup B^*$.

We will prove the claim by applying Lemmas 3.6.5 and 3.6.6 repeatedly to improve our current path cover. More precisely, we will construct path covers $\mathcal{Q}_2, \ldots, \mathcal{Q}_6$ such that eventually $\mathcal{Q}_6$ satisfies (Q1)–(Q7). So we can take $\mathcal{Q} := \mathcal{Q}_6$.

In order to be able to apply Lemmas 3.6.5 and 3.6.6, we must first bound the degrees of the vertices in T' from below. For all $v \in V(T')$, we have

$$d^+_{T'}(v) \geq d^+_T(v) - |A^* \cup B^*| \overset{\text{(i),(iii)}}{\geq} d^+_T(v) - 2ct \geq d^+_T(v) - \frac{\delta^0(T)}{5} \geq \frac{4}{5} d^+_T(v). \tag{3.6.8}$$

Similarly,

$$d^-_{T'}(v) \geq \frac{4}{5} d^-_T(v) \tag{3.6.9}$$

for all $v \in V(T')$.

We will first extend the tails of paths in $\mathcal{Q}_1$ out of E_A. We do this by applying Lemma 3.6.6 to T' and $\mathcal{Q}_1 = \mathcal{P}_1 \dot\cup \mathcal{P}_2$ with $I := E_A \setminus X$, $J := X \cup A' \cup B$ to form a new path cover $\mathcal{Q}_2$ of T' which will satisfy (Q1) and (Q2). By conditions (ii) and (v), no paths in $\mathcal{P}_2$ have endpoints in I. By condition (vi), $F \subseteq E(\mathcal{Q}_1)$. Moreover,

$$\begin{aligned} 3(|I|+|J|)+2|F| &\leq 3|E_A|+3|X|+3|A'|+3|B|+2|F| \\ &\overset{(ii),(vii),(vi)}{\leq} \frac{3}{40}d^- + 6kt + 6t + 4ct < \frac{4}{5}d^-. \end{aligned} \tag{3.6.10}$$

In the final inequality we used the fact that $d^- \geq \delta^0(T) \geq Ck^2$. Thus for all $v \in I$ we have

$$d^-_{T'}(v) \overset{(3.6.9)}{\geq} \frac{4}{5}d^-_T(v) \overset{(ii)}{\geq} \frac{4}{5}d^- \overset{(3.6.10)}{>} 3(|I|+|J|)+2|F|.$$

Thus the requirements of Lemma 3.6.6 are satisfied, and we can apply the lemma to obtain a path cover $\mathcal{Q}_2$ of T'.

Lemma 3.6.6(iv) implies that $\mathcal{Q}_2$ satisfies (Q1). Moreover, Lemma 3.6.6(v),(vi) imply that

$$\begin{aligned} |\mathcal{Q}_2| \leq |\mathcal{Q}_1| + 4k \qquad \text{as well as} \qquad & |\mathcal{Q}_2 \cap \mathcal{P}_2| \geq |\mathcal{P}_2| - 4k \geq |\mathcal{Q}_1| - 8k, \\ \text{and thus} \qquad & |\mathcal{Q}_2 \setminus \mathcal{P}_2| \leq 12k, \end{aligned} \tag{3.6.11}$$

where we have used that $|\mathcal{Q}_1| = |\mathcal{P}_1| + |\mathcal{P}_2| \leq |\mathcal{P}_2| + 4k$ for the second inequality above. Recall from condition (vii) that $X \subseteq V(\mathcal{P}_2)$ and $X \cap (A^* \cup B^*) = \emptyset$. Thus no paths in $\mathcal{Q}_1$ have endpoints in X. Moreover, since $t(\mathcal{P}_2) = B$ and $h(\mathcal{P}_2) = A'$, no paths in $\mathcal{Q}_1$ have tails in A' or heads in B. Together with Lemma 3.6.6(i)–(iii) this implies that $\mathcal{Q}_2$ satisfies (Q2) and

(a_1) $t(\mathcal{Q}_2) \cap A' = h(\mathcal{Q}_2) \cap B = \emptyset$.

(a_2) $h(\mathcal{Q}_2) \cap X = \emptyset$.

We will now extend the heads of paths in $\mathcal{Q}_2$ out of E_B. We do this by applying Lemma 3.6.5 to T', $(\mathcal{Q}_2 \setminus \mathcal{P}_2)\dot{\cup}(\mathcal{Q}_2 \cap \mathcal{P}_2)$ with $I := E_B \setminus X$, $J := (E_A \setminus E_B) \cup X \cup A' \cup B$ to form a new path cover $\mathcal{Q}_3$ of T' which will satisfy (Q1)–(Q4). As before, no paths in $\mathcal{P}_2 \supseteq \mathcal{Q}_2 \cap \mathcal{P}_2$ have endpoints in I, and $F \subseteq E(\mathcal{Q}_2)$ by (Q1) for $\mathcal{Q}_2$. Moreover, similarly as in (3.6.10) we obtain

$$\begin{aligned}3(|I| + |J|) + 2|F| &\leq 3|E_B| + 3|E_A| + 3|X| + 3|A'| + 3|B| + 2|F| \\ &\leq \frac{3}{40}d^+ + \frac{3}{40}d^- + 6kt + 6t + 4ct < \frac{4}{5}d^+.\end{aligned}$$

(In the final inequality we used our assumption that $d^- \leq d^+$.) Together with (3.6.8) this implies that $d^+_{T'}(v) \geq 4d^+/5 > 3(|I| + |J|) + 2|F|$ for all $v \in I$. Thus the requirements of Lemma 3.6.5 are satisfied, and we can apply the lemma to obtain a path cover $\mathcal{Q}_3$ of T'.

By Lemma 3.6.5(iv), $\mathcal{Q}_3$ satisfies (Q1). Lemma 3.6.5(v) implies that

$$|\mathcal{Q}_3| \leq |\mathcal{Q}_2| + |\mathcal{Q}_2 \setminus \mathcal{P}_2| \stackrel{(3.6.11)}{\leq} |\mathcal{Q}_2| + 12k \stackrel{(3.6.11)}{\leq} |\mathcal{Q}_1| + 16k. \tag{3.6.12}$$

Similarly, Lemma 3.6.5(vi) implies that

$$|\mathcal{Q}_3 \cap \mathcal{P}_2| \geq |\mathcal{Q}_2 \cap \mathcal{P}_2| - |\mathcal{Q}_2 \setminus \mathcal{P}_2| \stackrel{(3.6.11)}{\geq} |\mathcal{Q}_1| - 20k. \tag{3.6.13}$$

So $\mathcal{Q}_3$ satisfies (Q4). Lemma 3.6.5(iii) and (Q2) for $\mathcal{Q}_2$ together imply that $\mathcal{Q}_3$ satisfies (Q2). Moreover, (a_2) and Lemma 3.6.5(i),(ii) together imply that no path in $\mathcal{Q}_3$ has its head in $(E_B \setminus X) \cup X \supseteq E_B$ and so $\mathcal{Q}_3$ satisfies (Q3). Finally, (a_1) and Lemma 3.6.5(ii),(iii) together imply that

(b_1) no paths in $\mathcal{Q}_3$ have tails in A' or heads in B.

We will now extend the paths in $\mathcal{Q}_3 \setminus \mathcal{P}_2$ so that their endpoints lie in $A \cup B'$ rather than $A' \cup B$. More precisely, if $P \in \mathcal{Q}_3 \setminus \mathcal{P}_2$ has head $a'_i \in A'$, then we replace P by Pa'_ia_i

(recall that $a'_i a_i \in E(T)$ by condition (i) and $a_i \in A \subseteq V(T) \setminus V(\mathcal{Q}_3)$ by the definition of N). If $P \in \mathcal{Q}_3 \setminus \mathcal{P}_2$ has tail $b_i \in B$, we replace P by $b'_i b_i P$ (recall that $b'_i b_i \in E(T)$ by condition (iii) and $b'_i \in B' \subseteq V(T) \setminus V(\mathcal{Q}_3)$). Let $\mathcal{Q}_4$ be the path system thus obtained from $\mathcal{Q}_3$. Let $T'' := T[V(\mathcal{Q}_4)]$. Then

$$T' \subseteq T'' \subseteq T[V(T') \cup A \cup B'].$$

and $\mathcal{Q}_4$ is a path cover of T'' satisfying (Q1)–(Q4) and such that

$$|\mathcal{Q}_4| = |\mathcal{Q}_3| \qquad \text{and} \qquad \mathcal{Q}_4 \cap \mathcal{P}_2 = \mathcal{Q}_3 \cap \mathcal{P}_2. \tag{3.6.14}$$

Moreover, $h(\mathcal{Q}_4 \setminus \mathcal{P}_2) \cap A' = \emptyset$ and $t(\mathcal{Q}_4 \setminus \mathcal{P}_2) \cap B = \emptyset$. Together with ($b_1$) this implies that

(c_1) no paths in $\mathcal{Q}_4 \setminus \mathcal{P}_2$ have endpoints in $A' \cup B$.

Moreover, by construction of $\mathcal{Q}_4$, every vertex $a_i \in V(\mathcal{Q}_4) \cap A$ is a head of some path $P \in \mathcal{Q}_4 \setminus \mathcal{P}_2$ and this path P also contains a'_i (so in particular $P_i \notin \mathcal{Q}_4 \cap \mathcal{P}_2$). Similarly, every vertex in $b'_i \in V(\mathcal{Q}_4) \cap B'$ is a tail of some path $P \in \mathcal{Q}_4 \setminus \mathcal{P}_2$ and this path P also contains b_i (in particular $P_i \notin \mathcal{Q}_4 \cap \mathcal{P}_2$). Thus (Q5) as well as the following assertion hold:

(c_2) no paths in $\mathcal{Q}_4$ have heads in B' or tails in A.

We will now extend the tails of paths in $\mathcal{Q}_4 \setminus \mathcal{P}_2$ out of $A^* \cup B^*$. We do this by applying the strengthened form of Lemma 3.6.6 to T'', $(\mathcal{Q}_4 \setminus \mathcal{P}_2) \dot\cup (\mathcal{Q}_4 \cap \mathcal{P}_2)$ with $I := B'$, $J := E_A \cup E_B \cup A' \cup A \cup B$ to form a new path cover $\mathcal{Q}_5$ of T'' which still satisfies (Q1)–(Q5), and such that no path in $\mathcal{Q}_5 \setminus \mathcal{P}_2$ has endpoints in $A' \cup B' \cup B$. Clearly no paths in $\mathcal{P}_2 \supseteq \mathcal{Q}_4 \cap \mathcal{P}_2$ have tails in I, and $F \subseteq E(\mathcal{Q}_4)$ by (Q1). By condition (iii) we have $d^-_T(v) \geq n/3$ for all $v \in I$. Together with (3.6.9) this implies that $d^-_{T''}(v) \geq d^-_{T'}(v) \geq n/4$

for all $v \in I$. Note also that $|V(\mathcal{P}_2)| \le n/20$ by condition (v). So similarly as in (3.6.10), it follows that

$$\begin{aligned}
&3(|I|+|J|)+2|F|+|V(\mathcal{Q}_4\cap\mathcal{P}_2)| \\
&\qquad \le 3(|A'|+|A|+|B'|+|B|+|E_A|+|E_B|)+2|F|+|V(\mathcal{P}_2)| \\
&\qquad \le 12t+\frac{3}{20}d^+ + 4ct+\frac{n}{20} < \frac{n}{4} \le d^-_{T''}(v)
\end{aligned}$$

for all $v \in I$. Thus the requirements of the strengthened form of Lemma 3.6.6 are satisfied, and we can apply the lemma to obtain a path cover $\mathcal{Q}_5$ of T'' such that $\mathcal{Q}_5\cap\mathcal{P}_2 \supseteq \mathcal{Q}_4\cap\mathcal{P}_2$. Note that Lemma 3.6.6(ii),(iii) imply that the endpoints of $\mathcal{Q}_5 \setminus (\mathcal{P}_2\cap\mathcal{Q}_4)$ in J are the same as those of $\mathcal{Q}_4\setminus\mathcal{P}_2$. Together with ($c_1$) this implies that no paths in $\mathcal{Q}_5\setminus(\mathcal{P}_2\cap\mathcal{Q}_4)$ have endpoints in $A'\cup B$. In particular, this means that $\mathcal{Q}_5\cap\mathcal{P}_2 = \mathcal{Q}_4\cap\mathcal{P}_2$ and so

(d_1) no paths in $\mathcal{Q}_5\setminus\mathcal{P}_2$ have endpoints in $A'\cup B$.

Thus (Q5) for $\mathcal{Q}_4$ implies that $\mathcal{Q}_5$ satisfies (Q5) as well. Lemma 3.6.6(ii)–(iv), (vi) (strengthened) and (Q1)–(Q4) for $\mathcal{Q}_4$ together imply that $\mathcal{Q}_5$ satisfies (Q1)–(Q4). Moreover, Lemma 3.6.6(v) implies that

$$\begin{aligned}
|\mathcal{Q}_5| &\le |\mathcal{Q}_4|+|\mathcal{Q}_4\setminus\mathcal{P}_2| \overset{(3.6.14)}{=} |\mathcal{Q}_3|+|\mathcal{Q}_3\setminus\mathcal{P}_2| = 2|\mathcal{Q}_3|-|\mathcal{Q}_3\cap\mathcal{P}_2| \\
&\overset{(3.6.12),(3.6.13)}{\le} |\mathcal{Q}_1|+52k. \qquad (3.6.15)
\end{aligned}$$

By Lemma 3.6.6(i),(ii) and (c_2), we can also strengthen (d_1) to

(d_2) no paths in $\mathcal{Q}_5\setminus\mathcal{P}_2$ have endpoints in $A'\cup B'\cup B$ and no paths in $\mathcal{Q}_5$ have tails in A.

Finally, we will extend the heads of paths in $\mathcal{Q}_5\setminus\mathcal{P}_2$ out of $A^*\cup B^*$. We do this by applying the strengthened form of Lemma 3.6.5 to T'', $(\mathcal{Q}_5\setminus\mathcal{P}_2)\dot\cup(\mathcal{Q}_5\cap\mathcal{P}_2)$ with $I := A$,

$J := E_A \cup E_B \cup A' \cup B' \cup B$ to form a new path cover $\mathcal{Q}_6$ of T'' which will satisfy (Q1)–(Q7). Clearly no paths in $\mathcal{P}_2 \supseteq \mathcal{Q}_5 \cap \mathcal{P}_2$ have heads in I, and $F \subseteq E(\mathcal{Q}_5)$ by (Q1). Similarly as before, condition (i) and (3.6.8) together imply that

$$3(|I| + |J|) + 2|F| + |V(\mathcal{Q}_5) \cap \mathcal{P}_2| < \frac{n}{4} \le d^+_{T''}(v)$$

for all $v \in I$. Thus the requirements of the strengthened form of Lemma 3.6.5 are satisfied, and we can apply the lemma to obtain a path cover $\mathcal{Q}_6$ of T'' such that $\mathcal{Q}_6 \cap \mathcal{P}_2 = \mathcal{Q}_5 \cap \mathcal{P}_2$. (The fact that we have equality follows using a similar argument as in (d_1) above.)

Thus (Q5) for $\mathcal{Q}_5$ implies that $\mathcal{Q}_6$ satisfies (Q5) as well. Lemma 3.6.5(ii)–(iv), (vi) (strengthened) and (Q1)–(Q4) for $\mathcal{Q}_5$ together imply that $\mathcal{Q}_6$ satisfies (Q1)–(Q4). Also, by Lemma 3.6.5(v) we have

$$|\mathcal{Q}_6| \le |\mathcal{Q}_5| + |\mathcal{Q}_5 \setminus \mathcal{P}_2| = 2|\mathcal{Q}_5| - |\mathcal{Q}_5 \cap \mathcal{P}_2| \overset{(\mathrm{Q4}),(3.6.15)}{\le} |\mathcal{Q}_1| + 124k.$$

So (Q6) holds. Moreover, by Lemma 3.6.5(i)–(iii), (d_2) and the fact that $\mathcal{Q}_6 \cap \mathcal{P}_2 = \mathcal{Q}_5 \cap \mathcal{P}_2$, no paths in $\mathcal{Q}_6 \setminus \mathcal{P}_2$ have endpoints in $A' \cup A \cup B' \cup B$. Since no vertex in $(A^* \cup B^*) \setminus (A' \cup A \cup B' \cup B)$ lies in $V(T'') = V(\mathcal{Q}_6)$, this in turn implies (Q7). So the path system $\mathcal{Q} := \mathcal{Q}_6$ is as required in the claim.

We will now use the fact that each A_i and each B_i is an almost dominating set in order to extend the paths in $\mathcal{Q} \setminus \mathcal{P}_2$ into those A_i and B_i which contain the endpoints of paths in $\mathcal{Q} \cap \mathcal{P}_2$. We then use the paths in $\mathcal{Q} \cap \mathcal{P}_2$ to join these extended paths into a long cycle C covering (at least) N, and with $F \subseteq E(C)$. Finally, we will deploy whatever covering edges we need from F in order to absorb any vertices in $A^* \cup B^*$ not already covered into C.

Let $\mathcal{R} := \mathcal{Q} \setminus \mathcal{P}_2$ and $\mathcal{S} := \mathcal{Q} \cap \mathcal{P}_2$. In order to carry out the steps above, we would like to have $|\mathcal{R}| = |\mathcal{S}|$ to avoid having any paths in $\mathcal{S}$ left over. So we first split the paths in $\mathcal{R}$ until we have exactly $|\mathcal{S}|$ of them. In this process, we wish to preserve (Q1)–(Q3), (Q5) and (Q7). To show that this can be done, first note that by (Q4) and (Q6), we have

$$|\mathcal{R}| = |\mathcal{Q} \setminus \mathcal{P}_2| \leq 144k = t - 20k \leq |\mathcal{Q}_1| - 20k \leq |\mathcal{Q} \cap \mathcal{P}_2| = |\mathcal{S}|.$$

The number of edges in $\mathcal{R}$ which are incident to vertices in $E_A \cup E_B \cup A^* \cup B^*$, or which belong to F, is bounded above by

$$2(|E_A| + |E_B| + |A^*| + |B^*|) + |F| \leq \frac{d^+}{10} + 6ct \leq \frac{n}{4}.$$

On the other hand,

$$\begin{aligned}|E(\mathcal{R})| &= |V(\mathcal{R})| - |\mathcal{R}| \geq (n - |A^* \cup B^*| - |V(\mathcal{P}_2)|) - 144k \\ &\geq n - 2ct - \frac{n}{20} - 144k \geq \frac{n}{2}.\end{aligned}$$

Hence

$$|E(\mathcal{R})| - 2(|E_A| + |E_B| + |A^*| + |B^*|) - |F| \geq \frac{n}{4} > t \geq |\mathcal{S}|.$$

We may therefore form a path cover $\mathcal{R}'$ of $T[V(\mathcal{R})]$ with $|\mathcal{R}'| = |\mathcal{S}|$ by greedily removing edges of paths in $\mathcal{R}$ which are neither incident to $A^* \cup B^* \cup E_A \cup E_B$ nor elements of F. Then $\mathcal{R}' \cup \mathcal{S}$ satisfies (Q1)–(Q3), (Q5) and (Q7).

Next, we extend the paths in $\mathcal{R}'$ into $A^* \cup B^*$ and join them with the paths in $\mathcal{S}$ to form a long cycle C. By relabeling the P_i if necessary, we may assume that $\mathcal{S} = \{P_1, \ldots, P_\ell\}$. Let $R_1, \ldots, R_\ell$ denote the paths in $\mathcal{R}'$ and for each $j \in [\ell]$ let x_j be the tail of R_j and y_j the head of R_j. Recall from (Q2) and (Q7) that $x_j \notin A^* \cup B^* \cup E_A$. Hence by condition (ii) there exists $x'_j \in A_{j-1}$ with $x'_j x_j \in E(T)$, where the indices are understood to be modulo ℓ. Similarly $y_j \notin A^* \cup B^* \cup E_B$ by (Q3) and (Q7), so by condition (iv) there exists $y'_j \in B_j$ with $y_j y'_j \in E(T)$. Let $R'_j := x'_j x_j R_j y_j y'_j$. If $x'_j \neq a'_{j-1}$, then we extend R'_j by adding the edge $a'_{j-1}x'_j$. Similarly, if $y'_j \neq b_j$ we extend R'_j by adding the edge y'_jb_j. In all cases, we still denote the resulting path from a'_{j-1} to b_j by R'_j.

Recall that P_j is a path from b_j to a'_j for all $j \in [\ell]$. Moreover, we have $x'_j, y'_j \notin V(\mathcal{Q} \setminus \mathcal{P}_2) = V(\mathcal{R}')$ for all $j \in [\ell]$. (Indeed, if $x'_j \neq a_j$ this follows since for the oriented graph T'' defined in the claim we have $V(T'') \cap A_i \subseteq \{a_i, a'_i\}$. If $x'_j = a_j$, this follows since $P_j \in \mathcal{Q}$ and so (Q5) implies that $a_j \notin V(\mathcal{Q})$. The argument for y'_j is similar.) Thus $R'_1, \dots, R'_\ell$ are pairwise vertex-disjoint and internally disjoint from the paths in $\mathcal{S}$. So we can define a cycle C by

$$C := R'_1P_1R'_2P_2 \dots P_{\ell-1}R'_\ell P_\ell.$$

Note that $N \subseteq V(C)$ since $\mathcal{R}' \cup \mathcal{S}$ is a path cover of T'', and $F \subseteq E(C)$ by (Q1). Recall from condition (vi) that F consists of covering edges e_v for all $v \in A^* \cup B^*$ and that these e_v are pairwise distinct. Thus each e_v lies on C and so neither of the two activating edges of e_v can lie on C. Writing $e_v = x_vy_v$, it follows from these observations that we may form a new cycle C' by replacing x_vy_v by x_vvy_v in C for all $v \in (A^* \cup B^*) \setminus V(C)$. Then C' is a Hamilton cycle of T, as desired. $\square$

3.7 Finding many edge-disjoint Hamilton cycles in a good tournament

In the proof of Theorem 3.1.2, we will find the edge-disjoint Hamilton cycles in a given highly-linked tournament by repeatedly applying Lemma 3.6.7. In each application, we will need to set up all the dominating sets and paths required by Lemma 3.6.7. The following definition encapsulates this idea. (Recall that $\mathrm{Int}(P)$ denotes the interior of a path P.)

Definition 3.7.1 *We say that a tournament T is (C, k, t, c)-*good *if it contains vertex sets $A_1^1, \dots, A_k^t$, $B_1^1, \dots, B_k^t$, $E_{A,1}, \dots, E_{A,k}$, $E_{B,1}, \dots, E_{B,k}$, edge sets $F_1, \dots, F_k$, and paths*

$P_1^1, \dots, P_k^t$ *such that the following statements hold, where* $A_i^* := A_i^1 \cup \dots \cup A_i^t$, $A^* := A_1^* \cup \dots \cup A_k^*$, $B_i^* := B_i^1 \cup \dots \cup B_i^t$, *and* $B^* := B_1^* \cup \dots \cup B_k^*$*:*

(G1) The sets $A_1^1, \dots, A_k^t$ are disjoint and $2 \le |A_i^\ell| \le c$ for all $i \in [k]$ and $\ell \in [t]$. Moreover, each $T[A_i^\ell]$ is a transitive tournament whose head has out-degree at least $2n/5$ in T. Write $A := \{h(T[A_i^\ell]) : i \in [k], \ell \in [t]\}$.

(G2) The sets $B_1^1, \dots, B_k^t$ are disjoint from each other and from A^*, and $2 \le |B_i^\ell| \le c$ for all $i \in [k]$ and $\ell \in [t]$. Moreover, each $T[B_i^\ell]$ is a transitive tournament whose tail has in-degree at least $2n/5$ in T. Write $B' := \{t(T[B_i^\ell]) : i \in [k], \ell \in [t]\}$.

(G3) Write $d_- := \min\{d^-(v) : v \in V(T) \setminus (A \cup B')\}$. Each A_i^ℓ out-dominates $V(T) \setminus (A^* \cup B^* \cup E_{A,i})$. Moreover, $|E_{A,i}| \le d_-/50$ and $E_{A,i} \cap (A_i^* \cup B_i^*) = \emptyset$ for all $i \in [k]$.

(G4) Write $d_+ := \min\{d^+(v) : v \in V(T) \setminus (A \cup B')\}$. Each B_i^ℓ in-dominates $V(T) \setminus (A^* \cup B^* \cup E_{B,i})$. Moreover, $|E_{B,i}| \le d_+/50$ and $E_{B,i} \cap (A_i^* \cup B_i^*) = \emptyset$ for all $i \in [k]$.

(G5) Each P_i^ℓ is a path from the head of $T[B_i^\ell]$ to the tail of $T[A_i^\ell]$. For each $i \in [k]$, the paths $P_i^1, \dots, P_i^t$ are vertex-disjoint and $|P_1^1 \cup \dots \cup P_k^t| \le n/20$. For all $i \neq j$ and all $\ell, m \in [t]$, P_i^ℓ and P_j^m are edge-disjoint and

$$V(\mathrm{Int}(P_i^\ell)) \cap (A^* \cup B^*) \subseteq (A \cup B') \setminus (A_i^* \cup B_i^*).$$

(G6) $F_i \subseteq E(P_i^t)$ and $(A \cup B') \setminus (A_i^* \cup B_i^*) \subseteq V(P_i^t)$ for all $i \in [k]$.

(G7) The set $F_1 \cup \dots \cup F_k$ is a matching in $T - (A^* \cup B^*)$. For all $i \in [k]$ we have $F_i = \{e_v : v \in A_i^* \cup B_i^*\}$, where e_v is a covering edge for v and $e_v \neq e_{v'}$ whenever $v \neq v'$. Moreover, for each $i \in [k]$, let F_i^{act} be the set of activating edges corresponding to the covering edges in F_i. Then $F_i^{\mathrm{act}} \cap E(P_j^\ell) = \emptyset$ for all $i, j \in [k]$ and all $\ell \in [t]$.

(G8) We have $\delta^0(T) \ge Ck^2 \log k$.

For convenience, we collect the various disjointness conditions of Definition 3.7.1 into a single statement.

(G9)
- The sets $A_1^1, \dots, A_k^t$, $B_1^1, \dots, B_k^t$ are disjoint.
- $(E_{A,i} \cup E_{B,i}) \cap (A_i^* \cup B_i^*) = \emptyset$ for all $i \in [k]$.
- $F_1 \cup \cdots \cup F_k$ is a matching in $T - (A^* \cup B^*)$.
- For each $i \in [k]$, the paths $P_i^1, \dots, P_i^t$ are vertex-disjoint.
- For all $i \neq j$ and all $\ell, m \in [t]$, P_i^ℓ and P_j^m are edge-disjoint and $V(\mathrm{Int}(P_i^\ell)) \cap (A^* \cup B^*) \subseteq (A \cup B') \setminus (A_i^* \cup B_i^*)$. In particular, $P_i^1, \dots, P_i^t$ are internally disjoint from $A_i^* \cup B_i^*$.

The next lemma shows that for suitable parameters C, $t = t(k)$ and $c = c(k)$, every (C, k, t, c)-good tournament contains k edge-disjoint Hamilton cycles. In the next section we then show that there exists a constant $C' > 0$ such that any $C'k^2 \log k$-linked tournament is (C, k, t, c)-good (see Lemma 3.8.9). These two results together immediately imply Theorem 3.1.2.

As mentioned at the beginning of this section, in order to prove Lemma 3.7.2 we will apply Lemma 3.6.7 k times. In the notation for Definition 3.7.1, our convention is that the sets with subscript i will be used in the ith application of Lemma 3.6.7 to find the ith Hamilton cycle.

Lemma 3.7.2 *Let $C := 10^7$, $k \geq 20$, $t := 164k$, $c := \lceil \log 50t + 1 \rceil$. Then any (C, k, t, c)-good tournament contains k edge-disjoint Hamilton cycles.*

Proof. Let T be a (C, k, t, c)-good tournament, and let $n := |T|$. Let $A_1^1, \dots, A_k^t$, $B_1^1, \dots, B_k^t$, $E_{A,1}, \dots, E_{A,k}$, $E_{B,1}, \dots, E_{B,k}$, $F_1, \dots, F_k$, $P_1^1, \dots, P_k^t$, d_- and d_+ be as in Definition 3.7.1. (Note that this also implicitly defines sets $A_1^*, \dots, A_k^*$, A^*, A, $B_1^*, \dots, B_k^*$, B^*, B', and $F_1^{\mathrm{act}}, \dots, F_k^{\mathrm{act}}$ as in Definition 3.7.1.) Our aim is to apply Lemma 3.6.7

repeatedly to find k edge-disjoint Hamilton cycles. So suppose that for some $i \in [k]$ we have already found edge-disjoint Hamilton cycles $C_1, \dots, C_{i-1}$ such that the following conditions hold:

(a) $C_1, \dots, C_{i-1}$ are edge-disjoint from $T[A_j^\ell], T[B_j^\ell]$ and P_j^ℓ for all $i \le j \le k$ and all $\ell \in [t]$.

(b) $E(C_1 \cup \dots \cup C_{i-1}) \cap F_j^{\text{act}} = \emptyset$ for all $i \le j \le k$.

Intuitively, these conditions guarantee that none of the edges we will need in order to find $C_i, \dots, C_k$ are contained in $C_1, \dots, C_{i-1}$. We have to show that $T - C_1 - \dots - C_{i-1}$ contains a Hamilton cycle C_i which satisfies (a) and (b) (with i replaced by $i+1$).

Define

$$\begin{aligned}
T_i &:= T - \left(\bigcup_{j<i} C_j \cup \bigcup_{j>i} F_j^{\text{act}} \right) - \bigcup_{j>i,\ \ell \in [t]} (P_j^\ell \cup T[A_j^\ell] \cup T[B_j^\ell]), \\
E'_{A,i} &:= E_{A,i} \cup \left(\left(\bigcup_{j<i} N^+_{C_j}(A_i^*) \cup \bigcup_{j>i,\ \ell\in[t]} N^+_{P_j^\ell}(A_i^*) \cup A^* \cup B^* \right) \setminus (A_i^* \cup B_i^*) \right), \\
E'_{B,i} &:= E_{B,i} \cup \left(\left(\bigcup_{j<i} N^-_{C_j}(B_i^*) \cup \bigcup_{j>i,\ \ell\in[t]} N^-_{P_j^\ell}(B_i^*) \cup A^* \cup B^* \right) \setminus (A_i^* \cup B_i^*) \right), \\
X_i &:= (A \cup B') \setminus (A_i^* \cup B_i^*).
\end{aligned}$$

Then it suffices to find a Hamilton cycle C_i of T_i. We will do so by applying Lemma 3.6.7 to T_i, $A_i^1, \dots, A_i^t$, $B_i^1, \dots, B_i^t$, $P_i^1, \dots, P_i^t$, $E'_{A,i}$, $E'_{B,i}$, F_i and X_i. It therefore suffices to verify that the conditions of Lemma 3.6.7 hold.

We claim that for each $v \in V(T_i)$, we have

$$d^+_{T_i}(v) \ge d^+_T(v) - (i-1) - (k-i) - 1 - c > d^+_T(v) - 2k. \tag{3.7.3}$$

Indeed, it is immediate that $d^+_{C_1\cup\dots\cup C_{i-1}}(v) = i-1$. Since by (G9) for each $j > i$ the paths $P_j^1, \dots, P_j^t$ are vertex-disjoint, v is covered by at most $k-i$ of the paths $P_{i+1}^1, \dots, P_k^t$ and hence $d^+_{P_{i+1}^1\cup\dots\cup P_k^t}(v) \leq k-i$. Recall from (G7) that $F_1 \cup \dots \cup F_k$ consists of one covering edge e_v for each $v \in A^* \cup B^*$. Moreover, by (G9) the set $F_1 \cup \dots \cup F_k$ is a matching in $T - (A^* \cup B^*)$ and $A_1^1, \dots, A_k^t, B_1^1, \dots, B_k^t$ are all disjoint. Thus the digraph with edge set $F_1^{\rm act} \cup \dots \cup F_k^{\rm act}$ is a disjoint union of directed paths of length two and therefore has maximum out-degree one. Finally, since $A_1^1, \dots, A_k^t, B_1^1, \dots, B_k^t$ are disjoint, v belongs to at most one of $T[A_1^1], \dots, T[A_k^t], T[B_1^1], \dots, T[B_k^t]$. Moreover, $\Delta^+(T[A_j^\ell]), \Delta^+(T[B_j^\ell]) \leq c$ for all $j > i$ and all $\ell \in [t]$ by (G1) and (G2). So (3.7.3) follows. Similarly, we have

$$d^-_{T_i}(v) > d^-_T(v) - 2k. \tag{3.7.4}$$

In particular, $\delta(T_i) > n - 4k$, as required by Lemma 3.6.7.

We have $\delta^0(T) > Ck^2$ by (G8), and hence $\delta^0(T_i) > 10^6k^2$ as required by Lemma 3.6.7. The disjointness conditions of Lemma 3.6.7 are satisfied by (G9) and the definition of X_i. Since $V(T_i) = V(T)$, it is immediate that $A_i^1, \dots, A_i^t, B_i^1, \dots, B_i^t, X_i \subseteq V(T_i)$. We claim that $P_i^1, \dots, P_i^t \subseteq T_i$. Indeed, by (a) and (G5), each P_i^ℓ is edge-disjoint from $C_1 \cup \dots \cup C_{i-1}$ and from P_j^m for all $j > i$ and all $m \in [t]$. By (G7), each P_i^ℓ is edge-disjoint from $F_1^{\rm act} \cup \dots \cup F_k^{\rm act}$. Moreover, by (G5), each P_i^ℓ is edge-disjoint from $T[A_j^m] \cup T[B_j^m]$ for all $j > i$ and all $m \in [t]$. Altogether this implies that $P_i^1, \dots, P_i^t \subseteq T_i$. We have $F_i \subseteq E(P_i^t) \subseteq E(T_i)$ by (G6). It therefore suffices to prove that conditions (i)–(vii) of Lemma 3.6.7 hold.

Condition (v) follows from (G5). Condition (vi) follows from (G6) and (G7). (Note that (G7) implies that $F_i^{\rm act} \cap F_j^{\rm act} = \emptyset$ for all $i \neq j$. So (G7), (b) and the definition of T_i imply that $F_i^{\rm act} \subseteq T_i$.) By (G6) we have $X_i \subseteq V(P_i^t)$ and by (G1) and (G2) we have $|X_i| \leq |A \cup B'| = 2kt$, so condition (vii) holds too.

It therefore remains to verify conditions (i)–(iv). We first check (i). We have $2 \leq |A_i^\ell| \leq c$ by (G1). Moreover, we claim that $T_i[A_i^\ell] = T[A_i^\ell]$ for all $\ell \in [t]$. Indeed, to see this, note that $C_1, \dots, C_{i-1}$ are edge-disjoint from $T[A_i^\ell]$ by (a); by (G9) for all $j > i$ and all $m \in [t]$ each path P_j^m and each $T[A_j^m]$, $T[B_j^m]$ is edge-disjoint from $T[A_i^\ell]$; by (G7) all edges in F_j^{act} for $j > i$ are incident to a vertex in $A_j^* \cup B_j^*$, and hence by (G9) none of these edges belongs to $T[A_i^\ell]$. Thus $T_i[A_i^\ell] = T[A_i^\ell]$ is a transitive tournament by (G1). Finally, by (G1) the head of each $T[A_i^\ell]$ has out-degree at least $2n/5$ in T, and so by (3.7.3) out-degree at least $n/3$ in T_i. Hence condition (i) of Lemma 3.6.7 is satisfied. A similar argument shows that condition (iii) of Lemma 3.6.7 is also satisfied.

We will next verify that condition (ii) of Lemma 3.6.7 holds too. (G9) and the definition of $E'_{A,i}$ together imply that $E'_{A,i} \cap (A_i^* \cup B_i^*) = \emptyset$. By (G3), each A_i^ℓ out-dominates $V(T) \setminus (A^* \cup B^* \cup E_{A,i})$ in T, and hence out-dominates $V(T_i) \setminus (A^* \cup B^* \cup E_{A,i} \cup N^+_{T-T_i}(A_i^*))$ in T_i. However, it follows from (G9) that for all $j > i$ and all $\ell, m \in [t]$, no edge in F_j^{act} has an endpoint in A_i^ℓ and that $A_i^\ell \cap A_j^m = A_i^\ell \cap B_j^m = \emptyset$. Hence by (G9) we have that

$$N^+_{T-T_i}(A_i^*) = \bigcup_{j<i} N^+_{C_j}(A_i^*) \cup \bigcup_{j>i,\ \ell \in [t]} N^+_{P_j^\ell}(A_i^*).$$

It therefore follows from the definitions of $E'_{A,i}$ and T_i that A_i^ℓ out-dominates $V(T_i) \setminus (A_i^* \cup B_i^* \cup E'_{A,i})$ in T_i for all $\ell \in [t]$.

So in order to check that condition (ii) of Lemma 3.6.7 holds, it remains only to bound $|E'_{A,i}|$ from above. To do this, first note that by (G9), each vertex in A_i^* is contained in at most $k - i$ of the paths $P_{i+1}^1, \dots, P_k^t$. Moreover, $|E_{A,i}| \leq d_-/50$ by (G3). It therefore follows from the definition of $E'_{A,i}$, (G1) and (G2) that

$$
\begin{aligned}
|E'_{A,i}| &\le |E_{A,i}| + \left|\bigcup_{j<i} N^+_{C_i}(A^*_i)\right| + \left|\bigcup_{j>i,\ \ell\in[t]} N^+_{P^\ell_j}(A^*_i)\right| + |A^*| + |B^*| \\
&\le \frac{d_-}{50} + (i-1)|A^*_i| + (k-i)|A^*_i| + 2kct \le \frac{d_-}{50} + kct + 2kct \le \frac{d_-}{45}.
\end{aligned}
$$

The last inequality follows since $d_- \ge \delta^0(T) \ge Ck^2 \log k$ by (G8). Since $E'_{A,i}$ is disjoint from $A^*_i \cup B^*_i$, we have $E'_{A,i} \setminus X_i = E'_{A,i} \setminus (A \cup B')$. Hence for all $v \in E'_{A,i} \setminus X_i$ we have

$$
d^-_{T_i}(v) \overset{(3.7.4)}{\ge} d^-_T(v) - 2k \overset{(\mathrm{G3})}{\ge} d_- - 2k \ge \frac{19}{20} d_-
$$

and so

$$
|E'_{A,i}| \le \frac{d_-}{45} \le \frac{1}{40} \min\{d^-_{T_i}(v) : v \in E'_{A,i} \setminus X_i\}.
$$

This shows that condition (ii) of Lemma 3.6.7 is satisfied. The argument that (iv) holds is similar. We may therefore apply Lemma 3.6.7 to find a Hamilton cycle C_i in T_i as desired. □

3.8 Highly-linked tournaments are good

The aim of this section is to prove that any sufficiently highly-linked tournament is (C, k, t, c)-good. We first show that it is very easy to find covering edges for any given vertex – we will use the following lemma to find matchings $F_1, \dots, F_k$ consisting of covering edges as in Definition 3.7.1.

Lemma 3.8.1 *Suppose that T is a strongly 2-connected tournament, and $v \in V(T)$. Then there exists a covering edge for v.*

Proof. Since T is strongly connected and $|T| > 1$, we have $N^+(v), N^-(v) \neq \emptyset$. Since $T - v$ is strongly connected, there is an edge xy from $N^-(v)$ to $N^+(v)$. But then $xv, vy \in$

$E(T)$, so xy is a covering edge for v, as desired. $\square$

The next lemma will be used to obtain paths $P_1^1, \dots, P_k^t$ as in Definition 3.7.1. Recall that we require $F_i \subseteq E(P_i^t)$ and $(A \cup B') \setminus (A_i^* \cup B_i^*) \subseteq V(P_i^t)$ for all $i \in [k]$. We will ensure the latter requirement by first covering $(A \cup B') \setminus (A_i^* \cup B_i^*)$ with few paths and then linking these paths together – hence the form of the lemma.

Lemma 3.8.2 *Let $s \in \mathbb{N}$, and let T be a digraph. Let $x_1, \dots, x_k$, $y_1, \dots, y_k$ be distinct vertices of T, and let $\mathcal{Q}_1, \dots, \mathcal{Q}_k$ be (possibly empty) path systems in $T - \{x_1, \dots, x_k, y_1, \dots, y_k\}$ with $E(\mathcal{Q}_i) \cap E(\mathcal{Q}_j) = \emptyset$ whenever $i \neq j$. Write*

$$m := k + \sum_{i=1}^{k} |\mathcal{Q}_i| + \left| \bigcup_{i=1}^{k} V(\mathcal{Q}_i) \right|, \tag{3.8.3}$$

and suppose that T is $2sm$-linked. Then there exist edge-disjoint paths $P_1, \dots, P_k \subseteq T$ satisfying the following properties:

(i) P_i is a path from x_i to y_i for all $i \in [k]$.

(ii) $Q \subseteq P_i$ for all $Q \in \mathcal{Q}_i$ and all $i \in [k]$.

(iii) $V(P_i) \cap V(P_j) \subseteq V(\mathcal{Q}_i) \cap V(\mathcal{Q}_j)$ for all $i \neq j$.

(iv) $|P_1 \cup \dots \cup P_k| \leq |T|/s + |V(\mathcal{Q}_1) \cup \dots \cup V(\mathcal{Q}_k)|$.

Proof. For all $i \in [k]$, let $a_i^1 \dots b_i^1, \dots, a_i^{t_i} \dots b_i^{t_i}$ denote the paths in $\mathcal{Q}_i$. Let $F \subseteq E(T)$ denote the set of all those edges which form a path of length one in $\mathcal{Q}_1 \cup \dots \cup \mathcal{Q}_k$. Let

$$T' := T\left[\left(V(T) \setminus \bigcup_{i=1}^{k} V(\mathcal{Q}_i)\right) \cup \bigcup_{i=1}^{k} \bigcup_{j=1}^{t_i} \{a_i^j, b_i^j\}\right] - F.$$

Note that $E(T') \cap (E(\mathcal{Q}_1) \cup \cdots \cup E(\mathcal{Q}_k)) = \emptyset$. Define sets $X_1, \ldots, X_k$ of ordered pairs of vertices of T' by

$$X_i := \begin{cases} \{(x_i, a_i^1), (b_i^1, a_i^2), \ldots, (b_i^{t_i - 1}, a_i^{t_i}), (b_i^{t_i}, y_i)\}, & \text{if } \mathcal{Q}_i \neq \emptyset, \\ \{(x_i, y_i)\} & \text{if } \mathcal{Q}_i = \emptyset, \end{cases}$$

and let $X := X_1 \cup \cdots \cup X_k$. Let $\ell := 2sm - 2s|X|$. Since $|V(T) \setminus V(T')| + |F| \leq |V(\mathcal{Q}_1) \cup \cdots \cup V(\mathcal{Q}_k)|$ and $|X| = k + \sum_{i=1}^{k} |\mathcal{Q}_i|$, it follows that

$$2\ell = 4s(m - |X|) \overset{(3.8.3)}{=} 4s \left| \bigcup_{i=1}^{k} V(\mathcal{Q}_i) \right| \geq |V(T) \setminus V(T')| + 2|F|.$$

Thus by Proposition 3.4.7, T' is $2s|X|$-linked. We may therefore apply Lemma 3.4.8 to X in order to obtain, for each $i \in [k]$, a path system $\mathcal{P}_i$ whose paths link the pairs in X_i and such that whenever $i \neq j$, we have $E(\mathcal{P}_i) \cap E(\mathcal{P}_j) = \emptyset$ and $V(\mathcal{P}_i) \cap V(\mathcal{P}_j)$ consists of exactly the vertices that lie in a pair in both X_i and X_j. Let P_i be the path obtained from the union of all paths in $\mathcal{P}_i$ and all paths in $\mathcal{Q}_i$. Then $P_1, \ldots, P_k$ are edge-disjoint paths satisfying (i)–(iv). $\square$

The next lemma shows that given a vertex v in a tournament T, we can find a small transitive subtournament whose head is v and which out-dominates almost all vertices of T.

Lemma 3.8.4 *Let T be a tournament on n vertices, let $v \in V(T)$, and suppose that $c \in \mathbb{N}$ satisfies $2 \leq c \leq \log d^-(v) - 1$. Then there exist disjoint sets $A, E \subseteq V(T)$ such that the following properties hold:*

(i) $2 \leq |A| \leq c$ and $T[A]$ is a transitive tournament with head v.

(ii) A out-dominates $V(T) \setminus (A \cup E)$.

(iii) $|E| \leq (1/2)^{c-1} d^-(v)$.

The fact that the bound in (iii) depends on $d^-(v)$ is crucial: for instance, we can apply Lemma 3.8.4 with v being the vertex of lowest in-degree. Then (iii) implies that the 'exceptional set' $|E|$ is much smaller than $d^-(v) \leq d^-(w)$ for any $w \in E$. So while w is not dominated by A directly, it is dominated by many vertices outside E. This will make it possible to cover E by paths whose endpoints lie outside E. (More formally, the lemma is used to ensure (G3), which in turn is used for (Q2) in the proof of Lemma 3.6.7).

Proof. Let $v_1 := v$. We will find A by repeatedly choosing vertices $v_1, \dots, v_i$ such that the size of their common in-neighbourhood (i.e. the intersection of their individual in-neighbourhoods) is minimised at each step. More precisely, let $A_1 := \{v_1\}$. Suppose that for some $i < c$ we have already found a set $A_i = \{v_1, \dots, v_i\}$ such that $T[A_i]$ is a transitive tournament with head v_1, and such that the common in-neighbourhood E_i of $v_1, \dots, v_i$ satisfies

$$|E_i| \leq \frac{1}{2^{i-1}} d^-(v).$$

Note that these conditions are satisfied for $i = 1$. Moreover, note that E_i is the set of all those vertices in $T - A_i$ which are not out-dominated by A_i. If $|E_i| < 4$, then we have

$$|E_i| < 4 = \frac{1}{2^{\log d^-(v) - 2}} d^-(v) \leq \frac{1}{2^{c-1}} d^-(v), \tag{3.8.5}$$

and so A_i satisfies (i)–(iii). (Note that $|A_i| \geq 2$ since the assumptions imply that $d^-(v) \geq 8$.) Thus in this case we can take $A := A_i$ and $E := E_i$.

So suppose next that $|E_i| \geq 4$. In this case we will extend A_i to A_{i+1} by adding a suitable vertex v_{i+1}. By Proposition 3.6.1, E_i contains a vertex v_{i+1} of in-degree at most $|E_i|/2$ in $T[E_i]$. Let $A_{i+1} := \{v_1, \dots, v_{i+1}\}$ and let E_{i+1} be the common in-neighbourhood of $v_1, \dots, v_{i+1}$. Then $T[A_{i+1}]$ is a transitive tournament with head v_1 and

$$|E_{i+1}| \leq \frac{1}{2}|E_i| \leq \frac{1}{2^i} d^-(v).$$

By repeating this construction, either we will find $|E_i| < 4$ for some $i < c$ (and therefore take $A := A_i$ and $E := E_i$) or we will obtain sets A_c and E_c satisfying (i)–(iii). $\square$

We will also need the following analogue of Lemma 3.8.4 for in-dominating sets. It immediately follows from Lemma 3.8.4 by reversing the orientations of all edges.

Lemma 3.8.6 *Let T be a tournament on n vertices, let $v \in V(T)$, and suppose that $c \in \mathbb{N}$ satisfies $2 \leq c \leq \log d^+(v) - 1$. Then there exist disjoint sets $B, E \subseteq V(T)$ such that the following properties hold:*

(i) $2 \leq |B| \leq c$ *and* $T[B]$ *is a transitive tournament with tail* v.

(ii) B *in-dominates* $V(T) \setminus (B \cup E)$.

(iii) $|E| \leq (1/2)^{c-1} d^+(v)$.

We will now apply Lemma 3.8.4 repeatedly to obtain many pairwise disjoint small almost-out-dominating sets. We will also prove an analogue for in-dominating sets. These lemmas will be used in order to obtain sets $A_1^1, \dots, A_k^t$, $B_1^1, \dots, B_k^t$, $E_{A,1}, \dots, E_{A,k}$ and $E_{B,1}, \dots, E_{B,k}$ as in Definition 3.7.1.

Lemma 3.8.7 *Let T be a tournament on n vertices, $U \subseteq V(T)$ and $c \in \mathbb{N}$ with $c \geq 2$. Suppose that $\delta^-(T) \geq 2^{c+1} + c|U|$. Then there exist families $\{A_v : v \in U\}$ and $\{E_v : v \in U\}$ of subsets of $V(T)$ such that the following properties hold:*

(i) A_v *out-dominates* $V(T) \setminus (E_v \cup \bigcup_{u \in U} A_u)$ *for all* $v \in U$.

(ii) $T[A_v]$ *is a transitive tournament with head* v *for all* $v \in U$.

(iii) $|E_v| \leq (1/2)^{c-1} d^-(v)$ *for all* $v \in U$.

(iv) $2 \leq |A_v| \leq c$ *for all* $v \in U$.

(v) $A_u \cap E_v = \emptyset$ *for all* $u, v \in U$.

(vi) $A_u \cap A_v = \emptyset$ *for all* $u \neq v$.

Proof. We repeatedly apply Lemma 3.8.4. Suppose that for some $U' \subseteq U$ with $U' \neq U$ we have already found $\{A_u : u \in U'\}$ and $\{E'_u : u \in U'\}$ satisfying (ii)–(vi) (with U' playing the role of U and E'_u playing the role of E_u) such that

(a) A_v out-dominates $V(T) \setminus (\bigcup_{u \in U'} A_u \cup E'_v \cup U)$ for all $v \in U'$;

(b) $(\bigcup_{u \in U'} A_u) \cap U = U'$.

Pick $v \in U \setminus U'$. Our aim is to apply Lemma 3.8.4 to v and

$$T' := T - \left(\bigcup_{u \in U'} A_u \cup (U \setminus \{v\}) \right).$$

Note that $v \in V(T')$ by (b). Moreover,

$$d^-_{T'}(v) \geq \delta^-(T') \overset{\text{(iv)}}{\geq} \delta^-(T) - c|U'| - |U \setminus U'| \geq \delta^-(T) - c|U| \geq 2^{c+1},$$

where the final inequality holds by hypothesis, and so $c \leq \log d^-_{T'}(v) - 1$. Hence we can apply Lemma 3.8.4 to obtain disjoint sets $A_v, E_v \subseteq V(T')$ as described there. For all $u \in U'$, let $E_u := E'_u \setminus A_v$. Then the collections $\{A_u : u \in U' \cup \{v\}\}$ and $\{E_u : u \in U' \cup \{v\}\}$ satisfy (v) and (vi) (with $U' \cup \{v\}$ playing the role of U). Moreover, (b) holds too (with $U' \cup \{v\}$ playing the role of U'). Conditions (i)–(iii) of Lemma 3.8.4 imply that (a) holds (with $U' \cup \{v\}$, E_u playing the roles of U', E'_u) and that (ii)–(iv) hold (with $U' \cup \{v\}$ playing the role of U).

We continue in this way to obtain sets $\{A_u : u \in U\}$ and $\{E_u : u \in U\}$ which satisfy (ii)–(vi) as well as (a) (with U, E_u playing the roles of U', E'_u). But (a) implies (i) since $\bigcup_{u\in U} A_u \cup U = \bigcup_{u\in U} A_u$ (as $u \in A_u$ by (ii)). $\square$

The next lemma is an analogue of Lemma 3.8.7 for in-dominating sets. The proof is similar to that of Lemma 3.8.7.

Lemma 3.8.8 *Let T be a tournament on n vertices, $U \subseteq V(T)$ and $c \in \mathbb{N}$ with $c \geq 2$. Suppose that $\delta^+(T) \geq 2^{c+1}+c|U|$. Then there exist families $\{B_v : v \in U\}$ and $\{E_v : v \in U\}$ of subsets of $V(T)$ such that the following properties hold:*

(i) B_v *in-dominates* $V(T) \setminus (E_v \cup \bigcup_{u\in U} B_u)$ *for all* $v \in U$.

(ii) $T[B_v]$ *is a transitive tournament with tail* v *for all* $v \in U$.

(iii) $|E_v| \leq (1/2)^{c-1} d^+(v)$ *for all* $v \in U$.

(iv) $2 \leq |B_v| \leq c$ *for all* $v \in U$.

(v) $B_u \cap E_v = \emptyset$ *for all* $u, v \in U$.

(vi) $B_u \cap B_v = \emptyset$ *for all* $u \neq v$.

We will now combine the previous results in order to prove that any sufficiently highly-linked tournament is (C, k, t, c)-good. Note that Lemmas 3.7.2 and 3.8.9 together imply Theorem 3.1.2.

Lemma 3.8.9 *Let $C := 10^7$, $k \geq 20$, $t := 164k$ and $c := \lceil \log 50t + 1\rceil$. Then any $Ck^2 \log k$-linked tournament is (C, k, t, c)-good.*

Proof. Let T be a $Ck^2 \log k$-linked tournament, and let $n := |T|$. Note in particular that $\delta^0(T) \geq Ck^2 \log k$ by Proposition 3.4.6, so (G8) is satisfied. We have to choose $A^1_1, \dots, A^t_k$,

$B_1^1, \dots, B_k^t$, $E_{A,1}, \dots, E_{A,k}$, $E_{B,1}, \dots, E_{B,k}$, $F_1, \dots, F_k$ and $P_1^1, \dots, P_k^t$ satisfying (G1)–(G7) of Definition 3.7.1.

Construct a set $A \subseteq V(T)$ by greedily choosing kt vertices of least possible in-degree in T, and likewise construct a set $B' \subseteq V(T)$ by greedily choosing kt vertices of least possible out-degree in T. Note that by choosing the vertices in A and B' suitably, we may assume that $A \cap B' = \emptyset$. (Since $n \geq \delta^0(T) \geq 2kt$, this is indeed possible.) Define

$$d_- := \min\{d^-(v) : v \in V(T) \setminus (A \cup B')\},$$
$$d_+ := \min\{d^+(v) : v \in V(T) \setminus (A \cup B')\}.$$

Note that $d^-(a) \leq d_-$ for all $a \in A$ and $d^+(b) \leq d_+$ for all $b \in B'$.

Our first aim is to choose the sets $A_1^1, \dots, A_k^t$ using Lemma 3.8.7. Partition A arbitrarily into sets $A_1, \dots, A_k$ of size t, and write $A_i =: \{a_i^1, \dots, a_i^t\}$. Since $|B'| = kt \leq \delta^0(T)/2$, we have

$$2^{c+1} + c|A| \leq 400t + ckt \leq \frac{C}{2}k^2 \log k \leq \delta^-(T) - |B'| \leq \delta^-(T - B').$$

Thus we can apply Lemma 3.8.7 to $T - B'$, A and c in order to obtain almost out-dominating sets $A_i^\ell \ni a_i^\ell$ and corresponding exceptional sets $E_{A,i}^\ell$ as in the statement of Lemma 3.8.7 (for all $i \in [k]$ and all $\ell \in [t]$). Write $A_i^* := A_i^1 \cup \dots \cup A_i^t$ and $A^* := A_1^* \cup \dots \cup A_k^*$.

Let us now verify (G1). By Lemma 3.8.7(ii), (iv) and (vi), each $T[A_i^\ell]$ is a transitive tournament with head a_i^ℓ, $2 \leq |A_i^\ell| \leq c$, and the sets $A_1^1, \dots, A_k^t$ are all disjoint. In particular, $A = \{h(A_i^\ell) : i \in [k], \ell \in [t]\}$. We claim in addition that $d^+(a_i^\ell) \geq 2n/5$. Indeed, Proposition 3.6.2 implies that T has at most $4n/5 + 1$ vertices of out-degree at most $2n/5$, and hence at least $n/5 - 1$ vertices of out-degree at least $2n/5$. Moreover,

$$|A| = kt \leq \frac{Ck^2 \log k}{5} - 1 \leq \frac{n}{5} - 1.$$

So since the vertices of A were chosen to have minimal in-degree in T, it follows that $d^+(a_i^\ell) \geq 2n/5$ for all $i \in [k]$ and all $\ell \in [t]$. Thus (G1) holds.

We will next apply Lemma 3.8.8 in order to obtain the sets $B_1^1, \dots, B_k^t$. To do this, we first partition B' arbitrarily into sets $B'_1, \dots, B'_k$ of size t, and write $B'_i =: \{b_i'^1, \dots, b_i'^t\}$. Since $|A^*| \leq ktc \leq \delta^0(T)/2$, we have

$$2^{c+1} + c|B| \leq 400t + ckt \leq \frac{C}{2}k^2 \log k \leq \delta^+(T) - |A^*| \leq \delta^+(T - A^*).$$

Thus we can apply Lemma 3.8.8 to $T - A^*$, B' and c in order to obtain almost in-dominating sets $B_i^\ell \ni b_i'^\ell$ and corresponding exceptional sets $E_{B,i}^\ell$ as in the statement of Lemma 3.8.8 (for all $i \in [k]$ and all $\ell \in [t]$). Write $B_i^* := B_i^1 \cup \dots \cup B_i^t$ and $B^* := B_1^* \cup \dots \cup B_k^*$. Similarly as before one can show that (G2) holds.

We now define the exceptional sets $E_{A,i}$ and $E_{B,i}$. For all $i \in [k]$, let

$$E_{A,i} := (E_{A,i}^1 \cup \dots \cup E_{A,i}^t) \setminus B^* \quad \text{and} \quad E_{B,i} := (E_{B,i}^1 \cup \dots \cup E_{B,i}^t).$$

Recall from Lemmas 3.8.7(v) and 3.8.8(v) that $E_{A,i}^\ell \cap A^* = \emptyset$ and $E_{B,i}^\ell \cap (A^* \cup B^*) = \emptyset$ for all $i \in [k]$ and all $\ell \in [t]$. Thus $E_{A,i} \cap (A_i^* \cup B_i^*) = \emptyset$ and $E_{B,i} \cap (A_i^* \cup B_i^*) = \emptyset$ for all $i \in [k]$. By Lemma 3.8.7(i), each A_i^ℓ out-dominates $V(T) \setminus (A^* \cup B^* \cup E_{A,i})$. Lemma 3.8.7(iii) and the fact that $a_i^\ell \in A$ together imply that

$$|E_{A,i}| \leq \sum_{\ell=1}^{t} |E_{A,i}^\ell| \leq \sum_{\ell=1}^{t} \frac{1}{2^{c-1}} d^-(a_i^\ell) \leq \frac{t}{2^{c-1}} d_- \leq \frac{d_-}{50}, \tag{3.8.10}$$

so (G3) holds. Similarly, by Lemma 3.8.8(i), each B_i^ℓ in-dominates $V(T) \setminus (A^* \cup B^* \cup E_{B,i})$, and as in (3.8.10) one can show that $|E_{B,i}| \leq d_+/50$. Thus (G4) holds.

We now use Lemma 3.8.1 in order to define the sets $F_1, \dots, F_k$ of covering edges. Recall from (G7) that we require $F_1 \cup \dots \cup F_k$ to be a matching in $T - (A^* \cup B^*)$. Suppose that for some (possibly empty) subset $V' \subsetneq A^* \cup B^*$ we have defined a set $\{e_v : v \in V'\}$ of independent edges in $T - (A^* \cup B^*)$ such that e_v is a covering edge for v and $e_v \neq e_{v'}$ whenever $v \neq v'$. Pick any vertex $v \in (A^* \cup B^*) \setminus V'$. We will next define e_v. Let T' be the tournament obtained from T by deleting $(A^* \cup B^*) \setminus \{v\}$ as well as the endvertices of the covering edges $e_{v'}$ for all $v' \in V'$. Then

$$|V(T) \setminus V(T')| \leq |A^* \cup B^*| + 2|A^* \cup B^*| \leq 3ktc \leq \frac{C}{2}k^2 \log k,$$

so by Proposition 3.4.7, T' is still $(Ck^2 \log k/2)$-linked and hence strongly 2-connected. We may therefore apply Lemma 3.8.1 to find a covering edge e_v for v in T'. Continue in this way until we have chosen e_v for each $v \in A^* \cup B^*$ and let $F_i := \{e_v : v \in A_i^* \cup B_i^*\}$. Then the first part of (G7) holds.

It remains to choose the paths $P_1^1, \dots, P_k^t$. Recall from (G6) that we need to ensure that $(A \cup B') \setminus (A_i^* \cup B_i^*) \subseteq V(P_i^t)$ for all $i \in [k]$. We could achieve this by incorporating each of these vertices using the high linkedness of T. However, since $|A \cup B'| = 2kt$, a direct application of linkedness would require T to be $\Theta(k^3)$-linked. For each $i \in [k]$, we will therefore first choose a path cover $\mathcal{Q}_i$ of $T[(A \cup B') \setminus (A_i^* \cup B_i^*)]$ consisting of few paths and then use Lemma 3.8.2 (and thereby the high linkedness of T) to incorporate these paths into P_i^t. This has the advantage that we will only need T to be $\Theta(k^2 \log k)$-linked.

Let us first choose the path covers $\mathcal{Q}_i$ of $T[(A \cup B') \setminus (A_i^* \cup B_i^*)]$. Suppose that for some $j \in [k]$ we have already found path systems $\mathcal{Q}_1, \dots, \mathcal{Q}_{j-1}$ such that, for each $i < j$, $\mathcal{Q}_i$ is a path cover of $T[(A \cup B') \setminus (A_i^* \cup B_i^*)]$ with $|\mathcal{Q}_i| \leq 2k$, and such that for all $i < i' < j$ the paths in $\mathcal{Q}_i$ are edge-disjoint from paths in $\mathcal{Q}_{i'}$. To choose $\mathcal{Q}_j$, apply Corollary 3.6.4 to the oriented graph T'' obtained from $T[(A \cup B') \setminus (A_j^* \cup B_j^*)]$ by deleting the edges of all the paths in $\mathcal{Q}_1, \dots, \mathcal{Q}_{j-1}$. Since $\delta(T'') \geq |T''| - 1 - 2(j-1) \geq |T''| - 2k$, Corollary 3.6.4 ensures that $|\mathcal{Q}_j| \leq 2k$.

We will now choose $P_1^1, \dots, P_k^t$. For each $i \in [k]$ and each $\ell \in [t]$, let $a_i'^\ell$ denote the tail of $T[A_i^\ell]$ and b_i^ℓ the head of $T[B_i^\ell]$. Let

$$A' := \{a_i'^\ell : i \in [k],\ \ell \in [t]\} \quad \text{and} \quad B := \{b_i^\ell : i \in [k],\ \ell \in [t]\}.$$

For all $i \in [k]$ and all $\ell \in [t-1]$ let $\mathcal{Q}_i^\ell := \emptyset$. For all $i \in [k]$ let $\mathcal{Q}_i^t$ be the path system consisting of all the edges in F_i (each viewed as a path of length one) and all the paths in $\mathcal{Q}_i$. Let $T''' := T - ((A^* \cup B^*) \setminus (A \cup A' \cup B \cup B'))$. Our aim is to apply Lemma 3.8.2 with $s := 30$ to T''', the vertices $b_1^1, \dots, b_k^t$, $a_1'^1, \dots, a_k'^t$, and the path systems $\mathcal{Q}_1^1, \dots, \mathcal{Q}_k^t$. To verify that T''' is sufficiently highly linked, let m be as defined in (3.8.3) and note that

$$\begin{aligned} m &= kt + 3\sum_{i=1}^{k} |F_i| + \sum_{i=1}^{k} |\mathcal{Q}_i| + \left|\bigcup_{i=1}^{k} V(\mathcal{Q}_i)\right| \le kt + 6ckt + 2k^2 + |A \cup B'| \\ &\le 5kt + 6ckt \le \frac{C}{70} k^2 \log k. \end{aligned}$$

Together with the fact that $|T| - |T'''| \le 2ckt$ and Proposition 3.4.7 this implies that T''' is $2 \cdot 30m$-linked. So we can indeed apply Lemma 3.8.2 to find edge-disjoint paths P_i^ℓ in T''' (for all $i \in [k]$ and all $\ell \in [t]$) satisfying the following properties:

(i) P_i^ℓ is a path from b_i^ℓ to $a_i'^\ell$.

(ii) $Q \subseteq P_i^\ell$ for all $Q \in \mathcal{Q}_i^\ell$.

(iii) $V(P_i^\ell) \cap V(P_j^m) \subseteq V(\mathcal{Q}_i^\ell) \cap V(\mathcal{Q}_j^m)$ for all $(i, \ell) \neq (j, m)$.

(iv) We have that

$$\begin{aligned} |P_1^1 \cup \dots \cup P_k^t| &\le \frac{n}{30} + 2\sum_{i=1}^{k} |F_i| + \left|\bigcup_{i=1}^{k} V(\mathcal{Q}_i)\right| = \frac{n}{30} + 2|A^* \cup B^*| + |A \cup B'| \\ &\le \frac{n}{30} + 4ckt + 2kt \le \frac{n}{20}. \end{aligned}$$

Condition (ii) implies that $F_i \subseteq P_i^t$ and $(A \cup B') \setminus (A_i^* \cup B_i^*) = V(\mathcal{Q}_i) \subseteq V(\mathcal{Q}_i^t) \subseteq V(P_i^t)$ for all $i \in [k]$. Thus (G6) holds.

We now prove that (G5) holds. From (iii) and the fact that that $V(\mathcal{Q}_i^\ell) \cap V(\mathcal{Q}_i^m) = \emptyset$ for all $i \in [k]$, $\ell \neq m$, it follows that $P_i^1, \dots, P_i^t$ are vertex-disjoint for all $i \in [k]$. Together with (i) and (iv) this implies that in order to check (G5), it remains to show that

$$V(\mathrm{Int}(P_i^\ell)) \cap (A^* \cup B^*) \subseteq (A \cup B') \setminus (A_i^* \cup B_i^*) \qquad \text{for all } i \in [k], \ell \in [t]. \tag{3.8.11}$$

Clearly,

$$\begin{aligned} V(P_i^\ell) \cap (A^* \cup B^*) &\subseteq V(T''') \cap (A^* \cup B^*) \\ &= A \cup A' \cup B \cup B' \qquad \text{for all } i \in [k], \ell \in [t]. \end{aligned} \tag{3.8.12}$$

By definition, we have $(A' \cup B) \cap V(\mathcal{Q}_j^m) = \emptyset$ for all $j \in [k], m \in [t]$. It therefore follows from (iii) that each vertex in $A' \cup B$ may appear in at most one path P_j^m. However, by (i) each vertex in $A' \cup B$ is an endpoint of P_j^m for some $j \in [k], m \in [t]$. Hence

$$V(\mathrm{Int}(P_i^\ell)) \cap (A' \cup B) = \emptyset \qquad \text{for all } i \in [k], \ell \in [t]. \tag{3.8.13}$$

Fix $i \in [k], \ell \in [t]$ and take $j \in [k] \setminus \{i\}$. We have $(A \cup B') \cap (A_i^* \cup B_i^*) \cap V(\mathcal{Q}_i^\ell) = \emptyset$, and by (G6) we have $(A \cup B') \cap (A_i^* \cup B_i^*) \subseteq (A \cup B') \setminus (A_j^* \cup B_j^*) \subseteq V(P_j^t)$. Applying (iii) to P_i^ℓ and P_j^t, it therefore follows that

$$V(P_i^\ell) \cap (A \cup B') \cap (A_i^* \cup B_i^*) = \emptyset \qquad \text{for all } i \in [k], \ell \in [t]. \tag{3.8.14}$$

(3.8.12)–(3.8.14) now imply (3.8.11). Thus (G5) holds.

So it remains to check that the last part of (G7) holds too, i.e. that $F_i^{\mathrm{act}} \cap E(P_j^\ell) = \emptyset$ for all $i, j \in [k]$ and all $\ell \in [t]$. Consider any covering edge $e_v = x_v y_v \in F_i$. Then (G6) implies that x_v and y_v are contained in P_i^t. Moreover, (iii) implies that $V(P_i^t) \cap V(P_j^\ell) \subseteq V(\mathcal{Q}_i^t) \cap V(\mathcal{Q}_j^\ell) \subseteq A \cup B'$ whenever $(i,t) \neq (j,\ell)$. Since $x_v, y_v \notin A \cup B'$, this shows that $x_v v, v y_v \notin E(P_j^\ell)$ whenever $(i,t) \neq (j,\ell)$. But since $e_v \in E(P_i^t)$, we also have $x_v v, v y_v \notin E(P_i^t)$. This completes the proof that T is (C,k,t,c)-good. $\square$

3.9 Concluding remarks

3.9.1 Eliminating the logarithmic factor

A natural approach to improve the bound in Theorem 3.1.2 would be to reduce the parameter c, i.e. to consider smaller 'almost dominating' sets. In particular, if we could choose c independent of k, then we would obtain the (conjectured) optimal bound of $\Theta(k^2)$ for the linkedness. The obstacle to this in our argument is given by (3.8.10), which requires that c has a logarithmic dependence on k.

3.9.2 Algorithmic aspects

As remarked in the introduction, the proof of Theorem 3.1.2 is algorithmic. Indeed, when we apply the assumption of high linkedness to find appropriate paths in the proof of Lemma 3.8.9 (via Lemma 3.8.2), we can make use of the main result of [22] that these can be found in polynomial time. Moreover, the proof of the Gallai-Milgram theorem (Theorem 3.6.3) is also algorithmic (see [18]). These are the only tools we need in the proof, and the proof itself immediately translates into a polynomial time algorithm.

Chapter 4

Optimal covers with Hamilton cycles in random graphs

4.1 Introduction

Given graphs H and G, an H-decomposition of G is a set of edge-disjoint copies of H in G which cover all edges of G. The study of such decompositions forms an important area of Combinatorics but it is notoriously difficult. Often an H-decomposition does not exist (or it may be out of reach of current methods). In this case, the natural approach is to study the packing and covering versions of the problem. Here an *H-packing* is a set of edge-disjoint copies of H in G and an *H-covering* is a set of (not necessarily edge-disjoint) copies of H covering all the edges of G. An H-packing is *optimal* if it has the largest possible size and an H-covering is *optimal* if it has the smallest possible size. The two problems of finding (nearly) optimal packings and coverings may be viewed as 'dual' to each other.

By far the most famous problem of this kind is the Erdős-Hanani problem on packing and covering a complete r-uniform hypergraph with k-cliques, which was solved by Rödl [83]. In this case, it turns out that the (asymptotic) covering and packing versions of the problem are trivially equivalent and the solutions have approximately the same value.

Packings of Hamilton cycles in random graphs $G_{n,p}$ were first studied by Bollobás and Frieze [16]. (Here $G_{n,p}$ denotes the binomial random graph on n vertices with edge probability p.) Recently, the problem of finding optimal packings of edge-disjoint Hamilton cycles in a random graph has received a large amount of attention, leading to its complete solution in a series of papers by several authors (see below for more details on the history of the problem). The size of a packing of Hamilton cycles in a graph G is obviously at most $\lfloor \delta(G)/2 \rfloor$, and this trivial bound turns out to be tight in the case of $G_{n,p}$ for *any* p.

The covering version of the problem was first investigated by Glebov, Krivelevich and Szabó [39]. Note that the trivial bound on the size an optimal covering of a graph G with Hamilton cycles is $\lceil \Delta(G)/2 \rceil$. They showed that for $p \geq n^{-1+\varepsilon}$, this bound is a.a.s. approximately tight, i.e. in this range, a.a.s. the edges of $G_{n,p}$ can be covered with $(1+o(1))\Delta(G_{n,p})/2$ Hamilton cycles. Here we say that a property A holds a.a.s. (asymptotically almost surely), if the probability that A holds tends to 1 as n tends to infinity.

The authors of [39] also conjectured that their approximate bound could be extended to any $p = \omega(\log n/n)$. We are able to go further and prove the corresponding exact bound, unless p tends to 0 or 1 rather quickly.

Theorem 4.1.1 *Suppose that $G \sim G_{n,p}$, where $\frac{\log^{117} n}{n} \leq p \leq 1 - n^{-1/8}$. Then a.a.s. the edges of G can be covered by $\lceil \Delta(G)/2 \rceil$ Hamilton cycles.*

Note that the exact bound fails when p is sufficiently large. Indeed, let $n \geq 5$ be odd and take $p = 1 - n^{-2}$. Then with $\Omega(1)$ probability, $G \sim G_{n,p}$ is the complete graph with one edge uv removed. We claim that in this case, G cannot be covered by $(n-1)/2$ Hamilton cycles. Suppose such a cover exists. Then exactly one edge is contained in more than one Hamilton cycle in the cover. But u and v both have odd degrees, and hence are both incident to an edge contained in more than one Hamilton cycle. Since $uv \notin E(G)$, these edges must be distinct and we have a contradiction.

Note also that even though our result does not hold for $p > 1 - n^{-1/8}$, it still implies the conjecture of [39] in this range. Indeed, if $G \sim G_{n,p}$ with $p > 1 - n^{-1/8}$, we may simply partition G into two edge-disjoint graphs uniformly at random and apply Theorem 4.1.1 to each one to a.a.s. cover G with $(1 + o(1))n/2$ Hamilton cycles.

Unlike the situation with the Erdős-Hanani problem, the packing and covering problems are not equivalent in the case of Hamilton cycles. However, they do turn out to be closely related, so we now summarize the known results leading to the solution of the packing problem for Hamilton cycles in random graphs. Here 'exact' refers to a bound of $\lfloor \delta(G_{n,p})/2 \rfloor$, and ε is a positive constant.

authors	range of p	
Ajtai, Komlós & Szemerédi [3]	$\delta(G_{n,p}) = 2$	exact
Bollobás & Frieze [16]	$\delta(G_{n,p})$ bounded	exact
Frieze & Krivelevich [36]	p constant	approx.
Frieze & Krivelevich [37]	$p = \frac{(1+o(1)) \log n}{n}$	exact
Knox, Kühn & Osthus [50]	$p \geq \frac{C \log n}{n}$, C large	approx.
Ben-Shimon, Krivelevich & Sudakov [13]	$\frac{(1+o(1)) \log n}{n} \leq p \leq \frac{1.02 \log n}{n}$	exact
Knox, Kühn & Osthus [51]	$\frac{\log^{50} n}{n} \leq p \leq 1 - n^{-1/5}$	exact
Krivelevich & Samotij [54]	$\frac{\log n}{n} \leq p \leq n^{-1+\varepsilon}$	exact
Kühn & Osthus [59]	$p \geq 2/3$	exact

In particular, the results in [16, 51, 54, 59] (of which [51, 54] cover the main range) together show that for any p, a.a.s. the size of an optimal packing of Hamilton cycles in $G_{n,p}$ is $\lfloor \delta(G_{n,p})/2 \rfloor$. This confirms a conjecture of Frieze and Krivelevich [37] (a stronger conjecture was made in [36]).

The result in [59] is based on a recent result of Kühn and Osthus [61] which guarantees the existence of a Hamilton decomposition in every regular 'robustly expanding' digraph.

The main application of the latter was the proof (for large tournaments) of a conjecture of Kelly that every regular tournament has a Hamilton decomposition. But as discussed in [61, 59], the result in [61] also has a number of further applications to packings of Hamilton cycles in dense graphs and (quasi-)random graphs.

Recall that the above results imply an optimal packing result for any p. However, for the covering version, we need p to be large enough to ensure the existence of at least one Hamilton cycle before we can find any covering at all. This is the reason for the restriction $p = \omega(\log n/n)$ in the conjecture of Glebov, Krivelevich and Szabó [39] mentioned above. However, they asked the intriguing question whether this might extend to p which is closer to the threshold $\log n/n$ for the appearance of a Hamilton cycle in a random graph. In fact, it would be interesting to know whether a 'hitting time' result holds. For this, consider the well-known 'evolutionary' random graph process $G_{n,t}$: Let $G_{n,0}$ be the empty graph on n vertices. Consider a random ordering of the edges of K_n. Let $G_{n,t}$ be obtained from $G_{n,t-1}$ by adding the tth edge in the ordering. Given a property $\mathcal{P}$, let $t(\mathcal{P})$ denote the *hitting time* of $\mathcal{P}$, i.e. the smallest t so that $G_{n,t}$ has $\mathcal{P}$.

Question 4.1.2 *Let $\mathcal{C}$ denote the property that an optimal covering of a graph G with Hamilton cycles has size $\lceil \Delta(G)/2 \rceil$. Let $\mathcal{H}$ denote the property that a graph G has a Hamilton cycle. Is it true that a.a.s. $t(\mathcal{C}) = t(\mathcal{H})$?*

Note that $\mathcal{C}$ is not monotone. In fact, it is not even the case that for all $t > t(\mathcal{C})$, $G_{n,t}$ a.a.s. has $\mathcal{C}$. Taking $n \geq 5$ odd and $t = \binom{n}{2} - 1$, $G_{n,t}$ is the complete graph with one edge removed – which, as noted above, may not be covered by $(n-1)/2$ Hamilton cycles. It would be interesting to determine (approximately) the ranges of t such that a.a.s. $G_{n,t}$ has $\mathcal{C}$.

The approximate covering result of Glebov, Krivelevich and Szabó [39] uses the approximate packing result in [50] as a tool. More precisely, their proof applies the result in [50] to obtain an almost optimal packing. Then the strategy is to add a comparatively small number of Hamilton cycles which cover the remaining edges. Instead, our proof of Theorem 4.1.1 is based on the main technical lemma (Lemma 47) of the exact packing result in [51]. This is stated as Lemma 4.4.1 in the current chapter and (roughly) states the following: Suppose we are given a regular graph H which is close to being pseudorandom and a pseudorandom graph G_1, where G_1 is allowed to be surprisingly sparse compared to H. Then we can find a set of edge-disjoint Hamilton cycles in $G_1 \cup H$ covering all edges of H. Our proof involves several successive applications of this result, where we eventually cover all edges of $G_{n,p}$. In addition, our proof crucially relies on the fact that in the range of p we consider, there is a small but significant gap between the degree of the unique vertex x_0 of maximum degree and the other vertex degrees (and the same holds for the vertex of minimum degree). This means that for all vertices $x \neq x_0$, we can afford to cover a few edges incident to x more than once. The analogous observation for the minimum degree was exploited in [51] as well.

The result in [39] also holds for quasi-random graphs of edge density at least $n^{-1+\varepsilon}$, provided that they have an almost optimal packing of Hamilton cycles. It would be interesting to obtain such results for sparser quasi-random graphs too. In fact, the result in [51] does apply in a quasi-random setting (see Theorem 48 in [51]), but the assumptions are quite restrictive and it is not clear to which extent they can be used to prove results for (n, d, λ)-graphs, say. Note that even if the assumptions of [51] could be weakened, our results would still not immediately generalise to (n, d, λ)-graphs.

This chapter is organized as follows: In the next section, we collect several results and definitions regarding pseudorandom graphs, mainly from [51]. In Section 4.3, we apply Tutte's Theorem to give results which enable us to add a small number of edges to certain almost-regular graphs in order to turn them into regular graphs (without increasing the maximum degree). Finally, in Section 4.4 we put together all these tools to prove Theorem 4.1.1.

4.2 Pseudorandom graphs

The purpose of this section is to collect all the properties of $G_{n,p}$ that we need for our proof of Theorem 4.1.1. Throughout the rest of the chapter, we always assume that n is sufficiently large for our estimates to hold. In particular, some of our lemmas only hold for sufficiently large n, but we do not state this explicitly. We write log for the natural logarithm and $\log^a n$ for $(\log n)^a$. Given functions $f, g : \mathbb{N} \to \mathbb{R}$, we write $f = \omega(g)$ if $f/g \to \infty$ as $n \to \infty$. We denote the average degree of a graph G by $d(G)$.

We will need the following Chernoff bound (see e.g. Theorem 2.1 in [46]).

Lemma 4.2.1 *Suppose that $X \sim Bin(n,p)$. For any $0 < a < 1$ we have*

$$\mathbb{P}(X \leq (1-a)\mathbb{E}X) \leq e^{-\frac{a^2}{3}\mathbb{E}X}.$$

The following notion was first introduced by Thomason [88].

Definition 4.2.2 *Let $p, \beta \geq 0$ with $p \leq 1$. A graph G is (p,β)-jumbled if for all non-empty $S \subseteq V(G)$ we have*

$$\left| e_G(S) - p\binom{|S|}{2} \right| < \beta|S|.$$

We will also use the following immediate consequence of Definition 4.2.2. Suppose that G is a (p,β)-jumbled graph and $X, Y \subseteq V(G)$ are disjoint. Then

$$|e(X,Y) - p|X||Y|| \leq 2\beta(|X| + |Y|). \tag{4.2.3}$$

To see this, note that $e(X,Y) = e(X \cup Y) - e(X) - e(Y)$. Now (4.2.3) follows from Definition 4.2.2 by applying the triangle inequality.

The following notion was introduced in [51].

Definition 4.2.4 *Let G be a graph on n vertices. For a set $T \subseteq V(G)$, let $\overline{d}_G(T) := \frac{1}{|T|}\sum_{t\in T} d_G(t)$ be the average degree of the vertices of T in G. Then G is* strongly 2-jumping *if for all non-empty $T \subseteq V(G)$ we have*

$$\overline{d}_G(T) \geq \delta(G) + \min\{|T| - 1, \log^2 n\}.$$

Note that a strongly 2-jumping graph G is '2-jumping', i.e. it has a unique vertex of minimum degree and all other vertices have degree at least $\delta(G) + 2$.

The next definition collects (most of) the pseudorandomness properties that we need.

Definition 4.2.5 *A graph G on n vertices is p-*pseudorandom *if all of the following hold:*

(P1) G is $(p, 2\sqrt{np(1-p)})$-jumbled.

(P2) For any disjoint $S, T \subseteq V(G)$,

(i) if $\left(\frac{1}{|S|} + \frac{1}{|T|}\right)\frac{\log n}{p} \geq \frac{7}{2}$, then $e_G(S,T) \leq 2(|S| + |T|)\log n$,

(ii) if $\left(\frac{1}{|S|} + \frac{1}{|T|}\right)\frac{\log n}{p} \leq \frac{7}{2}$, then $e_G(S,T) \leq 7|S||T|p$.

(P3) For any $S \subseteq V(G)$,

(i) if $\frac{\log n}{|S|p} \geq \frac{7}{4}$, then $e(S) \leq 2|S|\log n$,

(ii) if $\frac{\log n}{|S|p} \leq \frac{7}{4}$, then $e(S) \leq \frac{7}{2}|S|^2 p$.

(P4) We have $np - 2\sqrt{np\log n} \leq \delta(G) \leq np - 200\sqrt{np(1-p)}$.

(P5) We have $\Delta(G) \leq np + 2\sqrt{np\log n}$.

(P6) G is strongly 2-jumping.

The following definition is essentially the same, except that some of the bounds are more restrictive.

Definition 4.2.6 *A graph G on n vertices is* strongly p-pseudorandom *if all of the following hold:*

(SP1) G is $(p, \frac{3}{2}\sqrt{np(1-p)})$-jumbled.

(SP2) For any disjoint $S, T \subseteq V(G)$,

(i) if $\left(\frac{1}{|S|} + \frac{1}{|T|}\right) \frac{\log n}{p} \geq \frac{7}{2}$, then $e_G(S,T) \leq \frac{3}{2}(|S| + |T|) \log n$,

(ii) if $\left(\frac{1}{|S|} + \frac{1}{|T|}\right) \frac{\log n}{p} \leq \frac{7}{2}$, then $e_G(S,T) \leq 6|S||T|p$.

(SP3) For any $S \subseteq V(G)$,

(i) if $\frac{\log n}{|S|p} \geq \frac{7}{4}$, then $e(S) \leq \frac{3}{2}|S| \log n$,

(ii) if $\frac{\log n}{|S|p} \leq \frac{7}{4}$, then $e(S) \leq 3|S|^2 p$.

(SP4) We have $np - 2\sqrt{np \log n} \leq \delta(G) \leq np - 200\sqrt{np(1-p)}$.

(SP5) We have $\Delta(G) \leq np + \frac{15}{8}\sqrt{np \log n}$.

(SP6) G is strongly 2-jumping.

The following lemma is an immediate consequence of Lemmas 9–11, 13 and 14 from [51].

Lemma 4.2.7 *Let $G \sim G_{n,p}$, where $48^2 \log^7 n/n \leq p \leq 1 - 36 \log^{\frac{7}{2}} n/\sqrt{n}$. Then G is strongly p-pseudorandom with probability at least $1 - 11/\log n$.*

The next observation shows that if we add a few edges at some vertex x_0 of a strongly pseudorandom graph such that none of these edges is incident to the unique vertex of minimum degree, then we obtain a graph which is still pseudorandom.

Lemma 4.2.8 *Suppose that G is a strongly p-pseudorandom graph with $p, 1-p = \omega(1/n)$. Let y_1 be the (unique) vertex of minimum degree in G and let $x_0 \neq y_1$ be any other vertex.*

Let F be a collection of edges of K_n not contained in G which are incident to x_0 but not to y_1 and such that $|F| \leq \sqrt{np \log n}/8$. Then the graph $G+F$ is p-pseudorandom.

Proof. Let $G' := G+F$. Clearly, (SP4) and (SP6) are not affected by adding the edges of F, so G' satisfies (P4) and (P6). The bound on $|F|$ together with (SP5) immediately imply that G' satisfies (P5).

We now show that G' satisfies (P1). Indeed, for any $S \subseteq V(G')$, (SP1) implies that

$$\begin{aligned} \left| e_{G'}(S) - p\binom{|S|}{2} \right| &\leq |e_{G'}(S) - e_G(S)| + \left| e_G(S) - p\binom{|S|}{2} \right| \\ &\leq |S| + \frac{3}{2}\sqrt{np(1-p)}|S| \leq 2\sqrt{np(1-p)}|S|. \end{aligned}$$

To check (P2), suppose that $S, T \subseteq V(G')$ are disjoint. Without loss of generality we may assume that $|S| \leq |T|$. First suppose $\left(\frac{1}{|S|} + \frac{1}{|T|}\right) \frac{\log n}{p} \geq \frac{7}{2}$. Then (i) of (SP2) implies that

$$e_{G'}(S,T) \leq e_G(S,T) + |T| \leq \frac{3}{2}\left(|S| + |T|\right)\log n + |T| \leq 2\left(|S| + |T|\right)\log n,$$

as required. Now suppose that $\left(\frac{1}{|S|} + \frac{1}{|T|}\right) \frac{\log n}{p} \leq \frac{7}{2}$. Then (ii) of (SP2) implies that

$$e_{G'}(S,T) \leq e_G(S,T) + |T| \leq |T|\left(6p|S| + 1\right) \leq 7|S||T|p.$$

So (ii) of (P2) holds. The proof that (P3) holds is essentially the same. $\square$

We say that a graph G on n vertices is *u-downjumping* if it has a unique vertex x_0 of maximum degree, and $d(x_0) \geq d(x) + u$ for all $x \neq x_0$. The following result follows from Lemma 17 in [51] by considering complements. The latter lemma in turn follows easily from Theorem 3.15 in [14].

Lemma 4.2.9 *Let $G \sim G_{n,p}$ with $p, 1-p = \omega\left(\log n/n\right)$. Then a.a.s. G is $5\frac{\sqrt{np(1-p)}}{\log n}$-*

downjumping.

The next result is intuitively obvious, but due to possible correlations between vertex degrees, it does merit some justification.

Lemma 4.2.10 *Suppose that* $\log^2 n/n < p' \leq p \leq 1 - \log^2 n/n$, *that* $p' \leq 1/2$ *and that* $G \sim G_{n,p}$. *Let* H *be a random subgraph of* G *obtained by including each edge of* G *into* H *with probability* p'/p. *Then a.a.s.* G *contains a unique vertex* x_0 *of maximum degree and* x_0 *does not have minimum degree in* H.

Proof. Fix any $\varepsilon > 0$. Let A be the event that G contains a unique vertex x_0 of maximum degree and that $d_H(x_0) = \delta(H)$. Let $f := np' - \sqrt{np' \log\log n}$. Let B be the event that $\delta(H) \leq f$. Note that $H \sim G_{n,p'}$. So Corollary 3.13 of [15] implies that $\mathbb{P}(\overline{B}) \leq \varepsilon$. Let C be the event that G contains a unique vertex x_0 of maximum degree and that $d_H(x_0) \leq f$ and note that $A \cap B \subseteq C$. Note also that $\mathbb{P}(A) \leq \mathbb{P}(A \cap B) + \mathbb{P}(\overline{B}) \leq \mathbb{P}(C) + \varepsilon$. We say that a graph F on n vertices is *typical* if $\Delta(F) \geq np$ and there is a unique vertex of degree $\Delta(F)$. Now let D be the event that G is typical. Then Corollary 3.13 of [15] and Lemma 4.2.9 together imply that $\mathbb{P}(\overline{D}) \leq \varepsilon$. For any fixed graph F on n vertices, let E_F denote the event that $G = F$. Then $\mathbb{P}(C) \leq \varepsilon + \sum_{F\colon F \text{ typical}} \mathbb{P}(C \mid E_F)\mathbb{P}(E_F)$. Suppose that E_F holds, where F is typical. Let $N := d_G(x_0)$ (note that E_F determines N and x_0). Whether the event C holds is now determined by a sequence of N Bernoulli trials, each with success probability p'/p. So let $X \sim \mathrm{Bin}(N, p'/p)$. Then $\mathbb{E}(X) = N(p'/p) \geq p'n$, which implies that $f \leq \mathbb{E}(X)(1 - \sqrt{\log\log n/\mathbb{E}(X)})$. Then an application of Lemma 4.2.1 gives us

$$\mathbb{P}(C \mid E_F) = \mathbb{P}(X \leq f) \leq e^{-\log\log n/3} \leq \varepsilon.$$

So $\mathbb{P}(C) \leq 2\varepsilon$, which in turn implies that $\mathbb{P}(A) \leq 3\varepsilon$. Since ε was arbitrary, this implies the result. □

Hefetz, Krivelevich and Szabó [44] proved a criterion for Hamiltonicity which requires only a rather weak quasirandomness notion. We will use a special case of their Theorem 1.2 in [44]. In that theorem, given a set S of vertices in a graph G, we let $N(S)$ denote the external neighbourhood of S, i.e. the set of all those vertices $x \notin S$ for which there is some vertex $y \in S$ with $xy \in E(G)$. Also, we say that G is *Hamilton-connected* if for any pair x, y of distinct vertices there is a Hamilton path with endpoints x and y.

Theorem 4.2.11 *Suppose that G is a graph on n vertices which satisfies the following:*

(HP1) For every $S \subseteq V(G)$ with $|S| \leq n/\sqrt{\log n}$, we have $|N(S)| \geq 20|S|$.

(HP2) G contains at least one edge between any two disjoint subsets $A, B \subseteq V(G)$ with $|A|, |B| \geq n/\log n$.

Then G is Hamilton-connected.

Theorem 4.2.12 *Let $G \sim G_{n,p}$ with $\log^8 n/n \leq p \leq 1 - n^{-1/3}$, and let x_0 be a vertex of maximum degree in G. Then a.a.s. $G - x_0$ is Hamilton-connected.*

Proof. It suffices to check that $G - x_0$ satisfies (HP1) and (HP2). For p in the above range, these properties are well known to hold a.a.s. for G with room to spare and so also hold for $G - x_0$. For completeness we point out explicit references. To check (HP1), first note that Lemma 4.2.7 implies that G is p-pseudorandom. So Corollary 37 of [51] applied with $A_x := N_G(x) \setminus \{x_0\}$ now implies that (HP1) holds. (HP2) is a special case of Theorem 2.11 in [15] – the latter guarantees a.a.s. the existence of many edges between A and B. $\square$

4.3 Extending graphs into regular graphs

The aim of this section is to show that whenever H is a graph which satisfies certain conditions and G is a p-pseudorandom graph on the same vertex set which is edge-disjoint from H, then G contains a spanning subgraph H' whose degree sequence complements that of H, i.e. such that $H \cup H'$ is $\Delta(H)$-regular. The conditions on H that we need are the following:

- H has even maximum degree.
- H is $\sqrt{np}$-downjumping.
- H satisfies $\Delta(H) - \delta(H) \leq (np \log n)^{5/7}$.

In order to show this we will use Tutte's f-factor theorem, for which we need to introduce the following notation. Given a graph $G = (V, E)$ and a function $f : V \to \mathbb{N} \cup \{0\}$, an *$f$-factor* of G is a subgraph G' of G such that $d_{G'}(v) = f(v)$ for all $v \in V$. Our approach will then be to set $f(v) := \Delta(H) - d_H(v)$ and attempt to find an f-factor in the pseudorandom graph G. The following result of Tutte [92, 93] gives a necessary and sufficient condition for a graph to contain an f-factor.

Theorem 4.3.1 *A graph $G = (V, E)$ has an f-factor if and only if for every two disjoint subsets $X, Y \subseteq V$, there are at most*

$$\sum_{x \in X} f(x) + \sum_{y \in Y} (d(y) - f(y)) - e(X, Y)$$

connected components K of $G - X - Y$ such that

$$\sum_{x \in K} f(x) + e(K, Y)$$

is odd.

When applying this result, we will often bound the number of components K of $G-X-Y$ for which $\sum_{x \in K} f(x) + e(K, Y)$ is odd by the total number of components of $G - X - Y$.

The next lemma (which is a special case of Lemma 20 in [51]) implies that there are at most $|X|+|Y|$ such components.

Lemma 4.3.2 *Let $G=(V,E)$ be a p-pseudorandom graph on n vertices with $pn \geq \log n$. Then for any nonempty $B \subseteq V$, the number of components of $G[V \setminus B]$ is at most $|B|$. In particular, G is connected.*

The following lemma guarantees an f-factor in a pseudorandom graph, as long as $\sum_{v\in V} f(v)$ is even, $f(v)$ is not too large and for all but at most one vertex $f(v)$ is not too small either. (Clearly, the requirement that $\sum_{v\in V} f(v)$ is even is necessary.)

Lemma 4.3.3 *Let $G=(V,E)$ be a p-pseudorandom graph on n vertices with $pn \geq \log^{21} n$, and let $f : V \to \mathbb{N}\cup\{0\}$ be a function such that $\sum_{v\in V} f(v)$ is even. Suppose that G contains a vertex x_0 such that $f(x_0)$ is even and such that*

$$f(x_0) \leq (np\log n)^{\frac{5}{7}} \quad \textit{and} \quad \sqrt{np} \leq f(v) \leq (np\log n)^{\frac{5}{7}} \;\; \textit{for all} \;\; v \in V\setminus\{x_0\}.$$

Then G has an f-factor.

Proof. Given two disjoint sets $X, Y \subseteq V$, we define $\alpha_f(X,Y)$ to be the number of connected components K of $G-X-Y$ such that

$$\sum_{x\in K} f(x) + e(K,Y)$$

is odd. We also define

$$\beta_f(X,Y) := \sum_{x\in X} f(x) + \sum_{y\in Y} (d(y)-f(y)) - e(X,Y).$$

By Theorem 4.3.1, it then suffices to prove that $\alpha_f(X,Y) \leq \beta_f(X,Y)$.

We will first show that $\alpha_f(X,Y) \leq |X| + |Y|$. If either X or Y is nonempty, this follows immediately from Lemma 4.3.2. If both X and Y are empty, then we must show that $\alpha_f(\emptyset,\emptyset) = 0$. But this holds since G is connected by Lemma 4.3.2, and $\sum_{x\in V} f(x)$ is even by hypothesis. Hence $\alpha_f(X,Y) \leq |X| + |Y|$ in all cases.

Hence if

$$\beta_f(X,Y) \geq |X| + |Y| \tag{4.3.4}$$

holds, then we have $\alpha_f(X,Y) \leq \beta_f(X,Y)$ and we are done. If $X = Y = \emptyset$, (4.3.4) holds. So it remains to consider the following cases.

Case 1. $|X| = 1$.

Let x denote the unique vertex in X. Suppose first that $Y = \emptyset$. In this case Lemma 4.3.2 implies that $G - x = G - X - Y$ is connected. If $x = x_0$ then $\sum_{v\in V\setminus\{x\}} f(v) = \sum_{v\in V} f(v) - f(x)$ is even. Thus $\alpha_f(X,Y) = 0$ and so $\beta_f(X,Y) \geq \alpha_f(X,Y)$, as desired. If $x \neq x_0$ then $\beta_f(X,Y) = f(x) \geq \sqrt{np} \geq 1 \geq \alpha_f(X,Y)$, as desired.

Thus we may assume that $Y \neq \emptyset$. Then

$$\begin{aligned}
\beta_f(X,Y) &\geq \sum_{y\in Y}(d(y) - f(y)) - |X||Y| \\
&\overset{(P4)}{\geq} \left(np - 2\sqrt{np\log n} - (np\log n)^{\frac{5}{7}}\right)|Y| - |Y| \\
&\geq \frac{np}{2}|Y| \geq |X| + |Y|
\end{aligned}$$

and so (4.3.4) holds.

Case 2. $|X| > 1$ and $|Y| \leq \frac{1}{4}|X|(np)^{-\frac{3}{14}}\log^{-\frac{5}{7}} n$.

Since $\sum_{y\in Y} d(y) \geq e(X,Y)$ it follows that in this case we have

$$\begin{aligned}\beta_f(X,Y) &\geq \sum_{x\in X} f(x) - \sum_{y\in Y} f(y) \geq (|X|-1)\sqrt{np} - |Y|(np\log n)^{\frac{5}{7}} \\ &\geq \frac{\sqrt{np}}{2}|X| - \frac{\sqrt{np}}{4}|X| \geq 2|X| \geq |X|+|Y|,\end{aligned}$$

and so (4.3.4) holds.

Case 3. $1 < |X| \leq \frac{n}{2}$ and $|Y| > \frac{1}{4}|X|(np)^{-\frac{3}{14}}\log^{-\frac{5}{7}} n$.

It follows by (P1) and (4.2.3) that

$$e(X,Y) \leq p|X||Y| + 4\sqrt{np}(|X|+|Y|).$$

Thus

$$\begin{aligned}\beta_f(X,Y) - \alpha_f(X,Y) &\geq \sum_{y\in Y}(d(y)-f(y)) - e(X,Y) - |X| - |Y| \\ &\overset{(P4)}{\geq} \left(np - 2\sqrt{np\log n} - (np\log n)^{\frac{5}{7}}\right)|Y| - p|X||Y| \\ &\qquad\qquad -5\sqrt{np}(|X|+|Y|) \\ &\geq \left(p(n-|X|) - 2(np\log n)^{\frac{5}{7}}\right)|Y| - 5\sqrt{np}|X| \qquad (4.3.5) \\ &\geq \left(\frac{np}{2} - 2(np\log n)^{\frac{5}{7}}\right)|Y| - 5\sqrt{np}|X| \\ &\geq \frac{1}{4}\left(\frac{(np)^{\frac{11}{14}}}{2\log^{\frac{5}{7}} n} - 22\sqrt{np}\right)|X| \geq 0,\end{aligned}$$

as desired.

Case 4. $|X| > \frac{n}{2}$ and $|Y| > \frac{1}{4}|X|(np)^{-\frac{3}{14}}\log^{-\frac{5}{7}} n$.

In this case we have

$$n - |X| \geq |Y| \geq \frac{|X|}{4(np)^{\frac{3}{14}} \log^{\frac{5}{7}} n} \geq \frac{n^{\frac{11}{14}}}{8p^{\frac{3}{14}} \log^{\frac{5}{7}} n}.$$

But as in the previous case, one can show that (4.3.5) still holds and so

$$\begin{aligned}
\beta_f(X,Y) - \alpha_f(X,Y) &\geq \left(p(n - |X|) - 2(np\log n)^{\frac{5}{7}}\right)|Y| - 5\sqrt{np}|X| \\
&\geq \left(\frac{(np)^{\frac{11}{14}}}{8\log^{\frac{5}{7}} n} - 2(np\log n)^{\frac{5}{7}}\right)|Y| - 5\sqrt{np}|X| \\
&\geq \frac{(np)^{\frac{11}{14}}}{9\log^{\frac{5}{7}} n}|Y| - 5\sqrt{np}|X| \\
&\geq \left(\frac{(np)^{\frac{4}{7}}}{36\log^{\frac{10}{7}} n} - 5\sqrt{np}\right)|X| \geq 0,
\end{aligned}$$

as desired.

This completes the proof of the lemma. $\square$

Corollary 4.3.6 *Let G be a p-pseudorandom graph on n vertices, where $pn \geq \log^{21} n$. Suppose that H is a graph on $V(G)$ which satisfies the following conditions:*

- *H is $\sqrt{np}$-downjumping.*
- *If x_0 is the unique vertex of maximum degree in H then $H - x_0$ and $G - x_0$ are edge-disjoint.*
- *$\Delta(H)$ is even.*
- *$\Delta(H) - \delta(H) \leq (np\log n)^{\frac{5}{7}}$.*

Then there exists a $\Delta(H)$-regular graph H' such that $H \subseteq H' \subseteq G \cup H$.

Proof. Define $f(v) := \Delta(H) - d_H(v)$ for all $v \in V(G)$. Then

$$\sum_{v \in V} f(v) = n\Delta(H) - \sum_{v \in V} d_H(v),$$

which is even. Moreover $f(x_0) = 0$ and our assumptions on H imply that

$$\sqrt{np} \leq f(v) \leq \Delta(H) - \delta(H) \leq (np \log n)^{\frac{5}{7}}$$

for all $v \in V \setminus \{x_0\}$. We may therefore apply Lemma 4.3.3 to find an f-factor G' in G. Then $H' := H \cup G'$ is a $\Delta(H)$-regular graph as desired. $\square$

4.4 Proof of Theorem 4.1.1

The main tool for our proof of Theorem 4.1.1 is the following result from [51, Lemma 47]. Roughly speaking, it asserts that given a regular graph H_0 which is contained in a pseudorandom graph G and given a pseudorandom subgraph G_0 of G which is allowed to be quite sparse compared to H_0, we can find a set of edge-disjoint Hamilton cycles in $H_0 \cup G_0$ which cover all edges of H_0. For technical reasons, instead of a single pseudorandom graph G_0, in its proof we actually need to consider a union of several edge-disjoint pseudorandom graphs $G_1, \ldots, G_{2m+1}$, where m is close to $\log n$.

Lemma 4.4.1 *Suppose that $p_0 \geq \frac{\log^{14} n}{n}$ and $p_1 \geq \frac{(np_0)^{\frac{3}{4}} \log^{\frac{5}{2}} n}{n}$. Let $m := \frac{\log(n^2 p_1)}{\log \log n}$, and for all $i \in [2m+1]$ set $p_i := p_1$ if i is odd, and $p_i := 10^{10} p_1$ if i is even. Let G be a p_0-pseudorandom graph on n vertices. Suppose that $G_1, \ldots, G_{2m+1}$ are pairwise edge-disjoint spanning subgraphs of G such that each G_i is p_i-pseudorandom. Moreover, for all $i \in [2m+1]$, let H_i be an even-regular spanning subgraph of G_i with $\delta(G_i) - 1 \leq d(H_i) \leq \delta(G_i)$. Suppose that H_0 is an even-regular spanning subgraph of G which is edge-disjoint from $\bigcup_{i=1}^{2m+1} H_i$. Then there exists a collection $\mathcal{HC}$ of edge-disjoint Hamilton cycles such* *that the union* $HC := \bigcup \mathcal{HC}$ *of all these Hamilton cycles satisfies* $H_0 \subseteq HC \subseteq \bigcup_{i=0}^{2m+1} H_i$.

The following lemma is a special case of Lemma 22(ii) of [51]. Given p_i-pseudorandom graphs G_i as in Lemma 4.4.1, it allows us to find the even-regular spanning subgraphs H_i required by Lemma 4.4.1.

Lemma 4.4.2 *Let* G *be a* p*-pseudorandom graph on* n *vertices such that* $p, 1-p = \omega\left(\log^2 n/n\right)$. *Then* G *has an even-regular spanning subgraph* H *with* $\delta(G) - 1 \leq d(H) \leq \delta(G)$.

The next lemma ensures that $G \sim G_{n,p}$ contains a collection of Hamilton cycles which cover all edges of G except for some edges at the vertex x_0 of maximum degree and such that every edge at x_0 is covered at most once. Theorem 4.1.1 will then be an easy consequence of this lemma and Theorem 4.2.12.

Lemma 4.4.3 *Let* $G \sim G_{n,p}$, *where* $\frac{\log^{117} n}{n} \leq p \leq 1 - n^{-\frac{1}{8}}$. *Then a.a.s.* G *has a unique vertex* x_0 *of degree* $\Delta(G)$ *and there exist a collection* $\mathcal{HC}$ *of Hamilton cycles in* G *and a collection* F *of edges incident to* x_0 *such that*

(i) *every edge of* $G - F$ *is covered by some Hamilton cycle in* $\mathcal{HC}$;

(ii) *no edge in* F *is covered by a Hamilton cycle in* $\mathcal{HC}$;

(iii) *no edge incident to* x_0 *is covered by more than one Hamilton cycle in* $\mathcal{HC}$.

Note that in Lemma 4.4.3, we have $|\mathcal{HC}| = (\Delta(G) - |F|)/2$.

The strategy of our proof of Lemma 4.4.3 is as follows. We split $G \sim G_{n,p}$ into three edge-disjoint random graphs G_1, G_2 and R such that the density of G_1 is almost p and both G_2 and R are much sparser. It turns out we may assume that the vertex x_0 of maximum degree in G also has maximum degree in G_1. We then apply Corollary 4.3.6 in order to extend G_1 into a $\Delta(G_1)$-regular graph by using some edges of R. Next we apply

Lemma 4.4.1 in order to cover this regular graph with edge-disjoint Hamilton cycles, using some edges of G_2.

Let H_2 be the subgraph of $R \cup G_2$ which is not covered by these Hamilton cycles. Again, we can make sure that x_0 is still the vertex of maximum degree in H_2. We now apply Corollary 4.3.6 again in order to extend H_2 into a $\Delta(H_2)$-regular graph H_2' by using edges of a random subgraph R' of G_1 (i.e. edges which we have already covered by Hamilton cycles). Finally, we would like to apply Lemma 4.4.1 in order to cover this regular graph by edge-disjoint Hamilton cycles, using edges of another sparse random subgraph G' of G_1. However, this means that in the last step we might use edges of G' at x_0, i.e. edges which have already been covered with edge-disjoint Hamilton cycles. Clearly, this would violate condition (iii) of the lemma.

We overcome this problem as follows: at the beginning, we delete all those edges at x_0 from G_1 which lie in G', and then we regularize and cover the graph H_1 thus obtained from G_1 as before, instead of G_1 itself. However, we have to ensure that x_0 is still the vertex of maximum degree in H_1. This forces us to make G' quite sparse: the average degree of G' needs to be significantly smaller than the gap between $d_G(x_0) = \Delta(G)$ and the degree of the next vertex, i.e. significantly smaller than $\sqrt{np(1-p)}/\log n$. Unfortunately it turns out that such a choice would make G' too sparse to apply Lemma 4.4.1 in order to cover H_2. Thus the above two 'iterations' are not sufficient to prove the lemma (where each iteration consists of an application of Corollary 4.3.6 to regularize and then an application of Lemma 4.4.1 to cover). But with three iterations, the above approach can be made to work.

Proof of Lemma 4.4.3. Lemmas 4.2.7 and 4.2.9 imply that a.a.s. G satisfies the following two conditions:

(a) G is p-pseudorandom.

(b) G is $5u$-downjumping, where $u := \frac{\sqrt{np(1-p)}}{\log n}$.

Note that

$$(np)^{\frac{27}{64}} \log^{\frac{259}{32}} n = \frac{\sqrt{np(1-p)}}{\log n} \cdot \frac{\log^{\frac{291}{32}} n}{(np)^{\frac{5}{64}} \sqrt{1-p}} \leq \frac{u}{2}. \tag{4.4.4}$$

Indeed, to see the last inequality note that either $1 - p \geq 1/2$ and $(np)^{\frac{5}{64}} \geq \log^{\frac{292}{32}} n$ or $(np)^{\frac{5}{64}} \geq (n/2)^{\frac{5}{64}}$ and $\sqrt{1-p} \geq n^{-\frac{1}{16}}$. So here we use the bounds on p in the lemma. Define

$$\begin{aligned}
p_2 &:= \frac{(np)^{\frac{3}{4}} \log^{\frac{7}{2}} n}{n} \geq \frac{\log^{91} n}{n}, \\
p_3 &:= \frac{(np_2)^{\frac{3}{4}} \log^{\frac{7}{2}} n}{n} = \frac{(np)^{\frac{9}{16}} \log^{\frac{49}{8}} n}{n} \geq \frac{\log^{71} n}{n}, \\
p_3' &:= 1600 p_3, \\
p_4 &:= \frac{(np_3)^{\frac{3}{4}} \log^{\frac{7}{2}} n}{n} = \frac{(np)^{\frac{27}{64}} \log^{\frac{259}{32}} n}{n} \geq \frac{\log^{57} n}{n}, \\
p_1 &:= p - 2p_2 - p_3, \\
m_i &:= \frac{\log(n^2 p_i)}{\log\log n} \quad \text{for all} \quad 2 \leq i \leq 4, \\
p_{(i,j)} &:= \begin{cases} \frac{p_i}{(10^{10}+1)m_i+1} & \text{if } 2 \leq i \leq 4 \text{ and if } j \in [2m_i+1] \text{ is odd,} \\ \frac{10^{10} p_i}{(10^{10}+1)m_i+1} & \text{if } 2 \leq i \leq 4 \text{ and if } j \in [2m_i+1] \text{ is even.} \end{cases}
\end{aligned}$$

Now form random subgraphs of G as follows. First partition G into edge-disjoint random graphs G_1, G_2, G_3 and R_2 such that $G_i \sim G_{n,p_i}$ for $i = 1, 2, 3$ and $R_2 \sim G_{n,p_2}$. (This can be done by randomly including each edge e of G into precisely one of G_1, G_2, G_3 and R_2, where the probability that e is included into G_i is p_i/p and the probability that e is included into R_2 is p_2/p, independently of all other edges of G.) We then choose edge-disjoint random subgraphs R_2', R_4 and G_4 of G_1 with $R_2' \sim G_{n,p_2}$, $R_4 \sim G_{n,p_4}$, and $G_4 \sim G_{n,p_4}$. (Since $p_1 \geq p_2 + 2p_4$ this can be done similarly to before.) Next we choose a random subgraph G_3' of G_2 such that $G_3' \sim G_{n,p_3'}$. To summarize, we thus have the following containments, where $\dot\cup$ denotes the edge-disjoint union of graphs:

$$G = G_1 \mathbin{\dot\cup} G_2 \mathbin{\dot\cup} G_3 \mathbin{\dot\cup} R_2 \quad \text{and} \quad G_1 \supseteq R_2' \mathbin{\dot\cup} R_4 \mathbin{\dot\cup} G_4 \quad \text{and} \quad G_2 \supseteq G_3'.$$

Finally, for each $i \in \{2, 3, 4\}$, we partition G_i into edge-disjoint random subgraphs $G_{(i,1)}, \dots, G_{(i,2m_i+1)}$ with $G_{(i,j)} \sim G_{n,p_{(i,j)}}$. Lemma 4.2.7 and a union bound implies that a.a.s. the following conditions hold:

(c) G_i is p_i-pseudorandom for all $i = 1, \dots, 4$.

(d) $G_{(i,j)}$ is $p_{(i,j)}$-pseudorandom for all $i = 2, 3, 4$ and all $j \in [2m_i + 1]$.

(e) R_2 and R_2' are p_2-pseudorandom, and R_4 is p_4-pseudorandom.

(f) $R_2 \cup G_2 \cup R_2' \cup G_3$ is strongly $(3p_2 + p_3)$-pseudorandom and $G_3' \cup G_3 \cup R_4 \cup G_4$ is strongly $(p_3' + p_3 + 2p_4)$-pseudorandom.

Since $R_2 \cup G_2 \cup R_2' \cup G_3 \sim G_{n,3p_2+p_3}$ and $G_3' \cup G_3 \cup R_4 \cup G_4 \sim G_{n,p_3'+p_3+2p_4}$, Lemma 4.2.10 implies that a.a.s. the following condition holds:

(g) Let x_0 be the unique vertex of maximum degree of G. Then x_0 is not the vertex of minimum degree in $R_2 \cup G_2 \cup R_2' \cup G_3$ or $G_3' \cup G_3 \cup R_4 \cup G_4$.

It follows that a.a.s. conditions (a)–(g) are all satisfied; in the remainder of the proof we will thus assume that they are. We can apply Lemma 4.4.2 for each $i = 2, 3, 4$ and each $j \in [2m_i + 1]$ to obtain an even-regular spanning subgraph $H_{(i,j)}$ of $G_{(i,j)}$ with $\delta(G_{(i,j)}) - 1 \leq d(H_{(i,j)}) \leq \delta(G_{(i,j)})$.

As indicated earlier, our strategy consists of the following three iterations. The purpose of the first iteration is to cover all the edges of G_1. To do this, we will apply Corollary 4.3.6 in order to extend G_1 into a regular graph H_1', using some edges of R_2. (Actually we will first set aside a set F_1 of edges of G_1 at x_0, but this will still leave x_0 the vertex of maximum degree in $H_1 := G_1 - F_1$. In particular, F_1 will contain the set F^* of all edges of G_4 at x_0.) We will then apply Lemma 4.4.1 to cover H_1' with edge-disjoint Hamilton cycles, using some edges of G_2.

The purpose of the second iteration is to cover all the edges of $G_2 \cup R_2$ not already covered in the first iteration – we denote this remainder by H_2. It turns out that x_0 will still be the vertex of maximum degree in H_2. If $\Delta(H_2)$ is odd, then we will add one edge from $F_1 \setminus F^*$ to H_2 to obtain a graph H_2' of even maximum degree. Otherwise, we simply let $H_2' := H_2$. We extend H_2' into a regular graph H_2'' using Corollary 4.3.6 and some edges of R_2', then cover H_2'' with edge-disjoint Hamilton cycles using Lemma 4.4.1 and some edges of G_3.

The purpose of the third iteration is to cover all the edges of G_3 not already covered in the second iteration – we denote this remainder by H_3. We first add some (so far unused) edges from $F_1 \setminus F^*$ to H_3 in order to make x_0 the unique vertex of maximum degree. Let H_3' denote the resulting graph. We then extend H_3' into a regular graph H_3'' using Corollary 4.3.6 and some edges of R_4, and finally cover H_3'' with edge-disjoint Hamilton cycles using Lemma 4.4.1 and some edges of G_4.

It is in this iteration that we make use of G_3', for technical reasons. It turns out that $G_3 \cup G_4 \cup R_4$ is so sparse that adding the required edges from $F_1 \setminus F^*$ may destroy its pseudorandomness, rendering it unsuitable as a choice of G in Lemma 4.4.1. Since the only role of G in Lemma 4.4.1 is that of a 'container' for the other graphs, this issue is easy to solve by adding a slightly denser random graph to $G_3 \cup G_4 \cup R_4$, namely G_3'.

Note that we did not use any edges of R_2' at x_0 when turning H_2' into H_2'' since x_0 is a vertex of maximum degree in H_2'. Similarly, we did not use any edges of R_4 at x_0 when turning H_3' into H_3''. Moreover, F^* was the set of all edges of G_4 at x_0 and no edge in F^* was covered in the first two iterations. Altogether this means that we do not cover any edge at x_0 more than once.

Note that in the second and third iterations, the graphs R_2' and R_4 we use for regularising consist of edges we have already covered. In the second iteration, this turns out to be a convenient way of controlling the difference between the maximum and minimum degree of H_3 (which might have been about $\Delta(G)-\delta(G)$ if we had used uncovered edges). In the third iteration, there are simply no more uncovered edges available.

After outlining our strategy, let us now return to the actual proof. We claim that x_0 is the unique vertex of maximum degree in G_1 and that G_1 is $4u$-downjumping. Indeed, for all $x \neq x_0$ we have

$$\begin{aligned}
d_{G_1}(x) &= d_G(x) - d_{G_2\cup G_3\cup R_2}(x) \overset{\text{(b)}}{\leq} d_G(x_0) - 5u - d_{G_2\cup G_3\cup R_2}(x)\\
&= d_{G_1}(x_0) + d_{G_2\cup G_3\cup R_2}(x_0) - 5u - d_{G_2\cup G_3\cup R_2}(x)\\
&\leq d_{G_1}(x_0) + \Delta(G_2) + \Delta(G_3) + \Delta(R_2) - 5u - \delta(G_2) - \delta(G_3) - \delta(R_2)\\
&\leq d_{G_1}(x_0) - \left(5u - 12\sqrt{np_2\log n}\right),
\end{aligned}$$

where the last inequality follows from the facts that both G_2 and R_2 are p_2-pseudorandom, G_3 is p_3-pseudorandom, $p_3 \leq p_2$ as well as from (P4) and (P5). But

$$\sqrt{np_2\log n} = (np)^{\frac{3}{8}}\log^{\frac{9}{4}} n \overset{(4.4.4)}{\leq} \frac{u}{2}\cdot(np)^{-\frac{3}{64}} \leq \frac{u}{\log n}. \tag{4.4.5}$$

Altogether this shows that $d_{G_1}(x) \leq d_{G_1}(x_0) - 4u$ for all $x \neq x_0$. Thus G_1 is $4u$-downjumping and x_0 is the unique vertex of maximum degree in G_1, as desired. Note that

$$\Delta(G_4) \leq 2np_4 = 2(np)^{\frac{27}{64}}\log^{\frac{259}{32}} n \overset{(4.4.4)}{\leq} u. \tag{4.4.6}$$

Let F^* be the set of all edges of G_4 which are incident to x_0. Thus $|F^*| \leq u$ by (4.4.6). Choose a set F_1 of edges incident to x_0 in G_1 such that $F^* \subseteq F_1$,

$$3u - 1 \leq |F_1| \leq 3u, \tag{4.4.7}$$

and such that $\Delta(G_1 - F_1)$ is even. Note that we used (4.4.6) and thus the full strength of (4.4.4) (in the sense that it would no longer hold if we replace 117 by 116 in the lower bound on p stated in Lemma 4.4.3) in order to be able to guarantee that $F^* \subseteq F_1$. So this is the point where we need the bounds on p in the lemma. Let $H_1 := G_1 - F_1$. Thus H_1 is still u-downjumping.

Our next aim is to apply Corollary 4.3.6 in order to extend H_1 into a $\Delta(H_1)$-regular graph H_1', using some of the edges of R_2. So we need to check that the conditions in Corollary 4.3.6 are satisfied. But since G_1 is p_1-pseudorandom we have

$$\begin{aligned}\Delta(H_1) - \delta(H_1) \leq \Delta(G_1) - \delta(G_1) &\overset{\text{(P4),(P5)}}{\leq} 4\sqrt{np_1 \log n} \\ &\leq 4\sqrt{np \log n} = 4(np_2)^{\frac{2}{3}} \log^{-\frac{11}{6}} n \leq (np_2 \log n)^{\frac{5}{7}}.\end{aligned} \tag{4.4.8}$$

Moreover $p_2 \geq \log^{21} n/n$ and H_1 is u-downjumping and so $\sqrt{np_2}$-downjumping by (4.4.5). Since R_2 is p_2-pseudorandom we may therefore apply Corollary 4.3.6 to find a regular graph H_1' of degree $\Delta(H_1)$ with $H_1 \subseteq H_1' \subseteq H_1 \cup R_2$.

Next, we wish to apply Lemma 4.4.1 in order to cover H_1' with edge-disjoint Hamilton cycles. Note that for every $1 \leq j \leq 2m_2 + 1$

$$np_{(2,j)} \geq \frac{np_2}{(10^{10}+1)m_2 + 1} \geq \frac{(np)^{\frac{3}{4}} \log^{\frac{7}{2}} n \log\log n}{10^{11} \log n} \geq (np)^{\frac{3}{4}} \log^{\frac{5}{2}} n. \tag{4.4.9}$$

So we can apply Lemma 4.4.1 with G, H_1', $G_{(2,1)}, \ldots, G_{(2,2m_2+1)}$ and $H_{(2,1)}, \ldots, H_{(2,2m_2+1)}$ playing the roles of G, H_0, $G_1, \ldots, G_{2m+1}$ and $H_1, \ldots, H_{2m+1}$ to obtain a collection $\mathcal{HC}_1$

of edge-disjoint Hamilton cycles such that the union $HC_1 := \bigcup \mathcal{HC}_1$ of these Hamilton cycles satisfies

$$H'_1 \subseteq HC_1 \subseteq H'_1 \cup \bigcup_{j=1}^{2m_2+1} H_{(2,j)} \subseteq H'_1 \cup G_2.$$

Write $H_2 := (G_2 \cup R_2) \setminus E(HC_1)$ for the uncovered remainder of $G_2 \cup R_2$. Note that

(HC1) no edge of G incident to x_0 is covered more than once in $\mathcal{HC}_1$;

(HC1$'$) HC_1 contains no edges from F_1.

Our next aim is to extend H_2 into a regular graph H'_2 using some of the edges of R'_2. We will then use some of the edges of G_3 in order to find edge-disjoint Hamilton cycles which cover H'_2. Note that

$$d_{H_2}(x) = d_{H_1}(x) + d_{R_2 \cup G_2}(x) - 2|\mathcal{HC}_1| \tag{4.4.10}$$

for all $x \in V(G)$. Together with the fact that H_1 is u-downjumping this implies that for all $x \neq x_0$ we have

$$\begin{aligned} d_{H_2}(x_0) - d_{H_2}(x) &= (d_{H_1}(x_0) - d_{H_1}(x)) + (d_{R_2 \cup G_2}(x_0) - d_{R_2 \cup G_2}(x)) \\ &\geq u - (\Delta(R_2) + \Delta(G_2) - (\delta(R_2) + \delta(G_2))) \\ &\geq u - 8\sqrt{np_2 \log n} \overset{(4.4.5)}{\geq} \sqrt{np_2}. \end{aligned}$$

(For the second inequality we used the fact that both R_2 and G_2 are p_2-pseudorandom together with (P4) and (P5).) Thus x_0 is the unique vertex of maximum degree in H_2 and H_2 is $\sqrt{np_2}$-downjumping. If $\Delta(H_2)$ is odd, let H'_2 be obtained from H_2 by adding some edge from $F_1 \setminus F^*$. Condition (g) ensures that we can choose this edge in such a way that it is not incident to the unique vertex of minimum degree in the $(3p_2 + p_3)$-pseudorandom graph $R_2 \cup G_2 \cup R'_2 \cup G_3$. Let F'_1 be the set consisting of this edge. If $\Delta(H_2)$ is even, let $H_2' := H_2$ and $F_1' := \emptyset$. In both cases, let $F_2 := F_1 \setminus F_1'$ and note that H_2' is still $\sqrt{np_2}$-downjumping. Moreover,

$$
\begin{aligned}
\Delta(H_2') - \delta(H_2') &\leq \Delta(H_2) - \delta(H_2) + 1 \\
&\overset{(4.4.10)}{\leq} \Delta(H_1) + \Delta(G_2) + \Delta(R_2) - \delta(H_1) - \delta(G_2) - \delta(R_2) + 1 \\
&\leq \Delta(G_1) + \Delta(G_2) + \Delta(R_2) - \delta(G_1) - \delta(G_2) - \delta(R_2) + 1 \\
&\leq 4\sqrt{np_1 \log n} + 8\sqrt{np_2 \log n} + 1 \leq 5\sqrt{np \log n} \\
&\leq (np_2 \log n)^{\frac{5}{7}}.
\end{aligned}
$$

(For the fourth inequality we used the facts that G_1 is p_1-pseudorandom and both R_2 and G_2 are p_2-pseudorandom together with (P4) and (P5). The final inequality follows similarly to (4.4.8).) Furthermore, note that $E(H_2') \cap E(R_2') \subseteq F_1'$ and so $H_2' - x_0$ and $R_2' - x_0$ are edge-disjoint. Thus we may apply Corollary 4.3.6 to find a regular graph H_2'' of degree $\Delta(H_2')$ with $H_2' \subseteq H_2'' \subseteq H_2' \cup R_2'$. Since x_0 is of maximum degree in H_2', we have the following:

$$\text{No edge from } R_2' \text{ incident to } x_0 \text{ was added to } H_2' \text{ in order to obtain } H_2''. \tag{4.4.11}$$

Let $G_2^* := (R_2 \cup G_2 \cup R_2' \cup G_3) + F_1'$. Our choice of F_1' and condition (f) together ensure that we can apply Lemma 4.2.8 with $R_2 \cup G_2 \cup R_2' \cup G_3$ and F_1' playing the roles of G and F to see that G_2^* is $(3p_2 + p_3)$-pseudorandom. Note that for every $1 \leq j \leq 2m_3 + 1$

$$np_{(3,j)} \geq (4np_2)^{\frac{3}{4}} \log^{\frac{5}{2}} n \geq (n(3p_2 + p_3))^{\frac{3}{4}} \log^{\frac{5}{2}} n,$$

where the first inequality follows similarly to (4.4.9). Hence we may apply Lemma 4.4.1 with G_2^*, H_2'', $G_{(3,1)}, \dots, G_{(3,2m_3+1)}$ and $H_{(3,1)}, \dots, H_{(3,2m_3+1)}$ playing the roles of G, H_0, $G_1, \dots, G_{2m+1}$ and $H_1, \dots, H_{2m+1}$ to obtain a collection $\mathcal{HC}_2$ of edge-disjoint Hamilton cycles such that the union $HC_2 := \bigcup \mathcal{HC}_2$ of these Hamilton cycles satisfies

$$H_2'' \subseteq HC_2 \subseteq H_2'' \cup \bigcup_{j=1}^{2m_3+1} H_{(3,j)} \subseteq H_2'' \cup G_3.$$

We now have the following properties:

(HC2) no edge of G incident to x_0 is covered more than once in $\mathcal{HC}_1 \cup \mathcal{HC}_2$;

(HC2$'$) $HC_1 \cup HC_2$ contains no edges from F_2;

(HC2$''$) $\mathcal{HC}_1 \cup \mathcal{HC}_2$ covers all edges in $(G_1 - F_2) \cup G_2 \cup R_2$.

Indeed, to see (HC2), first note that (4.4.11) implies that all edges incident to x_0 in HC_2 are contained in $H_2' \cup G_3$ and thus in $(H_2 + F_1') \cup G_3$, which is edge-disjoint from HC_1. Now (HC2) follows from (HC1) together with the fact that the Hamilton cycles in $\mathcal{HC}_2$ are pairwise edge-disjoint.

Write $H_3 := G_3 \setminus E(HC_2)$ for the subgraph of G_3 which is not covered by the Hamilton cycles in $\mathcal{HC}_2$. Our final aim is to extend H_3 into a regular graph H_3' using some of the edges of R_4. We will then use the edges of G_4 in order to find edge-disjoint Hamilton cycles which cover H_3' (and thus the edges of G_3 not covered so far). Note that for all $x \in V(G)$

$$d_{H_3}(x) = d(H_2'') + d_{G_3}(x) - 2|\mathcal{HC}_2|.$$

Together with the fact that G_3 is p_3-pseudorandom this implies that

$$\Delta(H_3) - \delta(H_3) = \Delta(G_3) - \delta(G_3) \overset{\text{(P4),(P5)}}{\leq} 4\sqrt{np_3 \log n}. \tag{4.4.12}$$

Thus we can add a set $F_2' \subseteq F_2 \setminus F^*$ of edges at x_0 to H_3 to ensure that x_0 is the unique vertex of maximum degree in the graph H_3' thus obtained from H_3, that H_3' is

$\sqrt{np_4}$-downjumping, $\Delta(H_3')$ is even and such that

$$|F_2'| \leq 4\sqrt{np_3 \log n} + \sqrt{np_4} + 1 \leq 5\sqrt{np_3 \log n} \leq \sqrt{np_2 \log n} \overset{(4.4.5)}{\leq} \frac{u}{\log n}. \tag{4.4.13}$$

Note that $|F_2 \setminus F^*| = |F_1 \setminus (F_1' \cup F^*)| \geq 2u - 2$ by (4.4.7) and since $|F^*| \leq u$ by (4.4.6). So we can indeed choose such a set F_2'. Moreover, condition (g) ensures that we can choose F_2' in such a way that it contains no edge which is incident to the unique vertex of minimum degree in the $(p_3' + p_3 + 2p_4)$-pseudorandom graph $G_3' \cup G_3 \cup R_4 \cup G_4$. Let $F_3 := F_2 \setminus F_2'$ and note that

$$\begin{aligned}\Delta(H_3') - \delta(H_3') &\leq \Delta(H_3) - \delta(H_3) + \sqrt{np_4} + 1 \overset{(4.4.12)}{\leq} 5\sqrt{np_3 \log n} = 5(np_4)^{\frac{2}{3}} \log^{-\frac{11}{6}} n \\ &\leq (np_4 \log n)^{\frac{5}{7}}.\end{aligned}$$

Furthermore, $E(H_3') \cap E(R_4) \subseteq F_2'$ and so $H_3' - x_0$ and $R_4 - x_0$ are edge-disjoint. Since also $p_4 \geq \log^{21} n/n$, we may apply Corollary 4.3.6 to obtain a regular graph H_3'' of degree $\Delta(H_3')$ such that $H_3' \subseteq H_3'' \subseteq H_3' \cup R_4$. Note that since x_0 is of maximum degree in H_3', we have the following:

No edge from R_4 incident to x_0 was added to H_3' in order to obtain H_3''. (4.4.14)

Let $G_3^* := (G_3' \cup G_3 \cup R_4 \cup G_4) + F_2'$. Since $|F_2'| \leq 5\sqrt{np_3 \log n} = \sqrt{np_3' \log n}/8$ by (4.4.13), we may apply Lemma 4.2.8 with $G_3' \cup G_3 \cup R_4 \cup G_4$ and F_2' playing the roles of G and F to see that G_3^* is $(p_3' + p_3 + 2p_4)$-pseudorandom.

Note that for every $1 \leq j \leq 2m_4 + 1$

$$np_{(4,j)} \geq (4np_3')^{\frac{3}{4}} \log^{\frac{5}{2}} n \geq (n(p_3' + p_3 + 2p_4))^{\frac{3}{4}} \log^{\frac{5}{2}} n,$$

where the first inequality follows similarly to (4.4.9). Recall that F^* denotes the set of all those edges of G_4 which are incident to x_0. Since $F_2' \cap F^* = \emptyset$, H_3'' and G_4 are edge-disjoint (and so $H_3'', H_{(4,1)}, \dots, H_{(4,2m_4+1)}$ are pairwise edge-disjoint). Thus we can apply Lemma 4.4.1 with G_3^*, H_3'', $G_{(4,1)}, \dots, G_{(4,2m_4+1)}$ and $H_{(4,1)}, \dots, H_{(4,2m_4+1)}$ playing the roles of G, H_0, $G_1, \dots, G_{2m+1}$ and $H_1, \dots, H_{2m+1}$ to obtain a collection $\mathcal{HC}_3$ of edge-disjoint Hamilton cycles such that the union $HC_3 := \bigcup \mathcal{HC}_3$ of these Hamilton cycles satisfies

$$H_3'' \subseteq HC_3 \subseteq H_3'' \cup \bigcup_{j=1}^{2m_4+1} H_{(4,j)} \subseteq H_3'' \cup G_4.$$

We claim that no edge of G incident to x_0 is covered more than once in $\mathcal{HC} := \mathcal{HC}_1 \cup \mathcal{HC}_2 \cup \mathcal{HC}_3$. Indeed, (HC2) implies that this was the case for $\mathcal{HC}_1 \cup \mathcal{HC}_2$. Moreover, recall that the Hamilton cycles in $\mathcal{HC}_3$ are pairwise edge-disjoint. In addition, (4.4.14) implies that all edges incident to x_0 in HC_3 are contained in

$$H_3' + F^* = H_3 + F_2' + F^* \subseteq H_3 + F_2.$$

So (HC2$'$) implies that none of these edges lies in $HC_1 \cup HC_2$, which proves the claim.

Note that (HC2$''$) and the definition of $\mathcal{HC}_3$ together imply that $\mathcal{HC}$ covers all edges of $G - F_3$. Let $F \subseteq F_3$ be the set of uncovered edges. Then F and $\mathcal{HC}$ are as required in the lemma. □

We remark that for the final application of Lemma 4.4.1 in the proof of Lemma 4.4.3 it would have been enough to consider $G_3 \cup R_4 \cup G_4$ instead of $G_3' \cup G_3 \cup R_4 \cup G_4$ (since H_3'' and all the $G_{(4,j)}$ are contained in $(G_3 \cup R_4 \cup G_4) + F_2'$). However, we would not have been able to apply Lemma 4.2.8 in this case since $|F_2'| > \sqrt{np_3 \log n}/8$. Introducing G_3' ensures that the conditions of Lemma 4.2.8 are satisfied (and this is the only purpose of G_3').

We can now combine Theorem 4.2.12 and Lemma 4.4.3 in order to prove Theorem 4.1.1.

Proof of Theorem 4.1.1. Lemma 4.4.3 implies that a.a.s. G contains a collection $\mathcal{HC}$ of Hamilton cycles and a collection F of edges incident to the unique vertex x_0 of maximum degree such that no edge of G incident to x_0 is contained in more than one Hamilton cycle in $\mathcal{HC}$ and such that the Hamilton cycles in $\mathcal{HC}$ cover precisely the edges of $G - F$. Moreover, by Theorem 4.2.12, a.a.s. $G - x_0$ is Hamilton-connected.

If $|F|$ is odd, we add one edge of $G - F$ incident to x_0 to F. We still denote the resulting set of edges by F. Let $r := |F|/2$ and $e_1e'_1, \dots, e_re'_r$ be pairs of edges such that F is the union of all these $2r$ edges. Since $G - x_0$ is Hamilton-connected, for each $1 \leq i \leq r$ there exists a Hamilton cycle C_i of G containing both e_i and e'_i. Then $\mathcal{HC} \cup \{C_1, \dots, C_r\}$ is a collection of $\lceil \Delta(G)/2 \rceil$ Hamilton cycles covering G, as desired. $\square$

Using further iterations in the proof of Lemma 4.4.3, one could reduce the exponent 117 in Lemma 4.4.3 (and thus in Theorem 4.1.1). One further iteration would lead to an exponent of 60, while the effect of yet further iterations quickly becomes insignificant.

Chapter 5

On-line Ramsey numbers of paths and cycles

5.1 Introduction

Ramsey's theorem [81] states that for all $k \in \mathbb{N}$, there exists $t \in \mathbb{N}$ such that any red-blue edge colouring of a clique K_t contains a monochromatic clique of order k. We call the least such t the k^{th} *Ramsey number*, and denote it by $r(k)$. Ramsey numbers and their generalisations have been a fundamentally important area of study in combinatorics for many years. Particularly well-studied are Ramsey numbers for graphs. Here the *Ramsey number* of two graphs G and H, denoted by $r(G, H)$, is the least t such that any red-blue edge colouring of K_t contains a red copy of G or a blue copy of H. See e.g. [80] for a survey of known Ramsey numbers.

An important generalisation of Ramsey numbers, first defined by Erdős, Faudree, Rousseau and Schelp [31], is as follows. Let G and H be two graphs. We say that a graph K has the (G, H)-*Ramsey property* if any red-blue edge colouring of K must contain either a red copy of G or a blue copy of H. Then the *size Ramsey number* $\hat{r}(G, H)$ is given by the minimum number of edges of any graph with the (G, H)-Ramsey property.

In this chapter, we consider the following related generalisation defined independently by Beck [12] and Kurek and Ruciński [64]. Let G and H be two graphs. Consider a game played on the edge set of the infinite clique $K_{\mathbb{N}}$ with two players, Builder and Painter. In each round of the game, Builder chooses an edge and Painter colours it red or blue. Builder wins by creating either a red copy of G or a blue copy of H, and wishes to do so in as few rounds as possible. Painter wishes to delay Builder for as many rounds as possible. (Note that Painter may not delay Builder indefinitely – for example, Builder may simply choose every edge of $K_{r(G,H)}$.) The *on-line Ramsey number* $\tilde{r}(G,H)$ is the minimum number of rounds it takes Builder to win, assuming that both Builder and Painter play optimally. We call this game the $\tilde{r}(G,H)$*-game*, and write $\tilde{r}(G) = \tilde{r}(G,G)$. Note that $\tilde{r}(G,H) \geq e(G) + e(H) - 1$ for all graphs G and H, as Painter may simply colour the first $e(G) - 1$ edges red and all subsequent edges blue. It is also clear that $\tilde{r}(G,H) \leq \hat{r}(G,H)$.

On-line Ramsey theory has been well-studied. The best known bounds for $\tilde{r}(K_t)$ are given by

$$\frac{r(t)}{2} \leq \tilde{r}(K_t) \leq t^{-c\frac{\log t}{\log\log t}} 4^t,$$

where c is a positive constant. The lower bound is due to Alon (see [12]), and the upper bound is due to Conlon [24]. Note that these bounds are similar to the best known bounds for classical Ramsey numbers $r(K_t)$, although Conlon also proves in [24] that

$$\tilde{r}(t) \leq C^{-t}\binom{r(t)}{2}$$

for some constant $C > 1$ and infinitely many values of t. For general graphs G, the best known lower bound for $\tilde{r}(G)$ is given by Grytczuk, Kierstead and Prałat [41].

Theorem 5.1.1 *For graphs G, we have $\tilde{r}(G) \geq \beta(G)(\Delta(G) - 1)/2 + e(G)$, where $\beta(G)$*

denotes the vertex cover number of G*.*

Various general strategies for Builder and Painter have also been studied. For example, consider the following strategy for Builder in the $\tilde{r}(G, H)$-game. Builder chooses a large but finite set of vertices in $K_{\mathbb{N}}$, say a set of size $n \in \mathbb{N}$, with $n \geq r(G, H)$. Then Builder chooses the edges of the induced K_n in a uniformly random order, allowing Painter to colour each edge as they wish, until the game ends. This strategy was analysed for the $\tilde{r}(K_3)$-game by Friedgut, Kohayakawa, Rödl, Ruciński and Tetali [35], and for the more general $\tilde{r}(G)$-game by Marciniszyn, Spöhel and Steger [67, 66].

Finally, it is interesting to consider the results of possible restrictions to Builder's strategy. For example, Grytczuk, Hałuszczak and Kierstead [40] proved (among other things) that if $\chi(G) \leq k$, then Builder can win the $\tilde{r}(G)$-game without uncovering a graph with chromatic number greater than k. Kierstead and Konjevod [49] consider similar questions for a generalisation of the $\tilde{r}(G, H)$-game to hypergraphs.

Given the known bounds on $\tilde{r}(K_t)$, it is not surprising that determining on-line Ramsey numbers exactly has proved even more difficult than determining classical Ramsey numbers exactly, and very few results are known. A significant amount of effort has been focused on the special case where G and H are paths. Grytczuk, Kierstead and Prałat [41] and Prałat [77, 78] have determined $\tilde{r}(P_{k+1}, P_{\ell+1})$ exactly when $\max\{k, \ell\} \leq 8$. In addition, Beck [10] has proved that the size Ramsey number $\hat{r}(P_{k+1})$ is at most linear in k. The following general bounds are the best known, and were proved in [41].

Theorem 5.1.2 *For all* $k, \ell \in \mathbb{N}$*, we have* $k + \ell - 1 \leq \tilde{r}(P_{k+1}, P_{\ell+1}) \leq 2k + 2\ell - 3$*.*

In general, it seems difficult to bound on-line Ramsey numbers $\tilde{r}(G, H)$ below. One of the major difficulties in doing so is the variety of possible strategies for Builder. We present a strategy for Painter which mitigates this problem somewhat.

Definition 5.1.3 *Let $\mathcal{F}$ be a family of graphs. We define the* $\mathcal{F}$-blocking strategy *for Painter as follows. Write R_i for the graph consisting of all uncovered red edges immediately before the ith move of the game, and write e_i for the ith edge chosen by Builder. Then Painter colours e_i red if $R_i + e_i$ is $\mathcal{F}$-free, and blue otherwise. (Recall that a graph is* $\mathcal{F}$-free *if it contains no graph in $\mathcal{F}$ as a subgraph.)*

In an $\tilde{r}(G, H)$-game, it is natural to consider $\mathcal{F}$-blocking strategies with $G \in \mathcal{F}$. For example, if $\mathcal{F} = \{G\}$, then the $\mathcal{F}$-blocking strategy for Painter consists of colouring every edge red unless doing so would cause Painter to lose the game. If Painter is using an $\mathcal{F}$-blocking strategy, one clear strategy for Builder would be to construct a red $\mathcal{F}$-free graph, then use it to force a blue copy of H in $e(H)$ moves. We will show that this is effectively Builder's only strategy (see Proposition 5.3.3), and thus to bound $\tilde{r}(G, H)$ below it suffices to prove that no small red $\mathcal{F}$-free graph can be used to force a blue copy of H. We use this technique to derive some lower bounds for on-line Ramsey numbers of the form $\tilde{r}(P_{k+1}, H)$, taking $\mathcal{F} = \{P_{k+1}\} \cup \{C_i : i \geq 3\}$.

Theorem 5.1.4 *Let $k, \ell, d \in \mathbb{N}$ with $k \geq 2$. Let H be a graph with $e(H) = \ell$ and $\Delta(H) = d$. Then*

$$\tilde{r}(P_{k+1}, H) \geq \begin{cases} (2d+1)\ell/(2d) & \text{if } k = 2, \\ (5d+4)\ell/(5d) & \text{if } k = 3, \\ (d+1)\ell/d & \text{if } k \geq 4. \end{cases}$$

Moreover, if H is connected and $k \geq 3$, then

$$\tilde{r}(P_{k+1}, H) \geq \begin{cases} \ell + 2\left\lceil (2\ell+1)/d \right\rceil/5 & \text{if } k = 3, \\ (d+1)\ell/d + \min\{k/2 - 1, \ell/d\} - 1 & \text{if } k \geq 4. \end{cases}$$

Taking $H = P_{\ell+1}$, we determine $\tilde{r}(P_3, P_{\ell+1})$ exactly for all $\ell \geq 2$. Furthermore, we obtain strong bounds on $\tilde{r}(P_{k+1}, P_{\ell+1})$ for all $\ell \geq k \geq 3$.

Theorem 5.1.5 *For all $\ell \geq k \geq 2$, we have*

$$\begin{aligned} \tilde{r}(P_{k+1}, P_{\ell+1}) &= \lceil 5\ell/4 \rceil && \text{if } k = 2, \\ (7\ell+2)/5 \leq \tilde{r}(P_{k+1}, P_{\ell+1}) &\leq (7\ell+52)/5 && \text{if } k = 3, \\ (3\ell+k)/2 - 2 \leq \tilde{r}(P_{k+1}, P_{\ell+1}) &\leq 2\ell + 2k - 3 && \text{if } k \geq 4. \end{aligned}$$

Here our lower bound for $k \geq 4$ follows from Theorem 5.1.4, and our upper bound is taken from Theorem 5.1.2. Note that our lower bound for $k \geq 4$ substantially improves Theorem 5.1.2 unless k is very close to ℓ. Our proof of the upper bound for $k = 3$ is complicated, and in the interest of clarity we have chosen not to fully optimise the bound. We do not believe that it could be made tight without substantial additional work, however.

Motivated by Theorem 5.1.5 and the known values of $\tilde{r}(P_{k+1}, P_{\ell+1})$, we make the following conjecture.

Conjecture 5.1.6 *For all $\ell \geq k$, we have*

$$\tilde{r}(P_{k+1}, P_{\ell+1}) = \begin{cases} \ell & \text{if } k = 1, \\ \lceil 5\ell/4 \rceil & \text{if } k = 2, \\ \lceil (7\ell+2)/5 \rceil & \text{if } k = 3, \\ \lceil 3\ell/2 \rceil + k - 3 & \text{if } k \geq 4. \end{cases}$$

In particular, we have $\tilde{r}(P_{k+1}) = \lceil 5k/2 \rceil - 3$ for $k \geq 3$.

Note that Conjecture 5.1.6 would imply Conjecture 4.1 of [77]. The conjecture is trivially true for $k = 1$. Theorem 5.1.5 implies that it is true for $k = 2$ and that it is true up to an additive error in the upper bound for $k = 3$. It also implies that an approximate lower bound holds when $k = o(\ell)$ as $\ell \to \infty$. Finally, the conjecture is already known when $\max\{k, \ell\} \leq 8$ or when $k = \ell = 9$.

We also determine $\tilde{r}(P_3, C_\ell)$ exactly for all ℓ.

Theorem 5.1.7 *For all $\ell \geq 3$, we have*

$$\tilde{r}(P_3, C_\ell) = \begin{cases} \ell + 2 & \text{if } \ell \leq 4, \\ \lceil 5\ell/4 \rceil & \text{if } \ell \geq 5. \end{cases}$$

Note that $\tilde{r}(P_3, C_\ell) = \tilde{r}(P_3, P_{\ell+1})$ for $\ell \geq 5$. This is somewhat surprising, as $e(C_\ell) = e(P_{\ell+1})$ but it seems much harder for Builder to close a blue cycle than to extend a blue path. This result gives rise to the following natural question.

Question 5.1.8 *For what graphs G and integers ℓ do we have $\tilde{r}(G, C_\ell) = \tilde{r}(G, P_{\ell+1})$?*

Further, we give bounds on $\tilde{r}(C_4, P_{\ell+1})$.

Theorem 5.1.9 *For $\ell \geq 3$, we have $2\ell \leq \tilde{r}(C_4, P_{\ell+1}) \leq 4\ell - 4$. Moreover, $\tilde{r}(C_4, P_4) = 8$.*

Many of the lower bounds in Theorems 5.1.5 and 5.1.7 follow from Theorem 5.1.4, and all of them follow from analysing $\mathcal{F}$-blocking strategies. In particular, we obtain tight lower bounds on $\tilde{r}(P_3, P_{\ell+1})$ and $\tilde{r}(P_3, C_\ell)$ in this way, as well as a lower bound on $\tilde{r}(P_4, P_{\ell+1})$ which matches Conjecture 5.1.6. We are therefore motivated to ask the following question.

Question 5.1.10 *For which graphs G and H does there exist a family $\mathcal{F}$ of graphs such that the $\mathcal{F}$-blocking strategy is optimal for Painter in the $\tilde{r}(G, H)$-game?*

The chapter is laid out as follows. In Section 5.3, we prove Theorem 5.1.4. In Section 5.4, we prove Theorem 5.1.5 for $k = 2$ (see Theorem 5.4.3). We use a similar argument to prove Theorem 5.1.7 in Section 5.5 (see Proposition 5.5.2 and Theorem 5.5.3). In Section 5.6 we prove Theorem 5.1.5 for $k = 3$. Finally, in Section 5.7 we prove Theorem 5.1.9.

5.2 Notation and conventions

We write $\mathbb{N}$ for the set $\{1, 2, \dots\}$ of natural numbers, and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.

Suppose $P = v_1 \dots v_k$ and $Q = w_1 \dots w_\ell$ are paths. If $i < j$, we write $v_i P v_j$ (or $v_j P v_i$) for the subpath $v_i v_{i+1} \dots v_j$ of P. We also write PQ for the concatenation of P and Q. For example, if $i < j$ and $i' < j'$ then $uv_iPv_jyw_{i'}Qw_{j'}$ denotes the path $uv_iv_{i+1} \dots v_jyw_{i'}w_{i'+1} \dots w_{j'}$.

In the context of an $\tilde{r}(G, H)$-game, an *uncovered edge* is an edge of $K_\mathbb{N}$ that has previously been chosen by Builder, and a *new vertex* is a vertex in $K_\mathbb{N}$ not incident to any uncovered edge.

Many of our lemmas say that in an $\tilde{r}(G, H)$-game, given a finite coloured graph $X \subseteq K_\mathbb{N}$, Builder can force Painter to construct a coloured graph $Y \subseteq K_\mathbb{N}$ satisfying some desired property. We will often apply such a lemma to a finite coloured graph $X' \supsetneq X$, and in these cases we will implicitly require $V(Y) \cap V(X') \subseteq V(X)$. (Intuitively, when Builder chooses a new vertex while constructing Y, it should be new with respect to X' rather than X.) This is formally valid, since we may apply the lemma to an $\tilde{r}(G, H)$-game on the board $K_\mathbb{N} - (V(X') \setminus V(X))$ and have Builder choose the corresponding edges in $K_\mathbb{N}$.

For technical convenience, we allow Builder to "waste" a round in the $\tilde{r}(G, H)$-game by choosing an uncovered edge. Clearly this change does not affect the duration of an optimally-played game.

5.3 General lower bounds

Our aim is to bound $\tilde{r}(G,H)$ below for graphs G and H. In this section, Painter will always use an $\mathcal{F}$-blocking strategy for some family $\mathcal{F}$ of graphs with $G \in \mathcal{F}$. Hence, as we shall demonstrate in Proposition 5.3.3 below, Builder's strategy boils down to choosing a red graph with which to force a blue copy of H.

Definition 5.3.1 *Let $\mathcal{F}$ be a family of graphs and let $R \subseteq K_{\mathbb{N}}$ be an $\mathcal{F}$-free graph. We say that an edge $e \in K_{\mathbb{N}} - R$ is $(R,\mathcal{F})$-*forceable *iff $R + e$ is not $\mathcal{F}$-free. We say a graph H is $(R,\mathcal{F})$-*forceable *iff there exists $H' \subseteq K_{\mathbb{N}} - R$ with $H' \simeq H$ such that every edge $e \in E(H')$ is $(R,\mathcal{F})$-forceable. We call H' an $(R,\mathcal{F})$-*forced copy *of H. If R and $\mathcal{F}$ are clear from context, we will omit '$(R,\mathcal{F})$-'.*

Definition 5.3.2 *Let $\mathcal{F}$ be a family of graphs and let H be a graph. We say a graph $R \subseteq K_{\mathbb{N}}$ is an $\mathcal{F}$-*scaffolding *for H iff the following properties hold.*

(i) R is $\mathcal{F}$-free.

(ii) H is $(R,\mathcal{F})$-forceable.

(iii) R contains no isolated vertices.

Note that (iii) is simply a convenience – any isolated vertices in R have no bearing on Builder's ability to use R to force a copy of H, so we disregard them.

Proposition 5.3.3 *Let G and H be graphs. Let $\mathcal{F}$ be a family of graphs with $G \in \mathcal{F}$. Suppose every $\mathcal{F}$-scaffolding for H has at least m edges. Then $\tilde{r}(G,H) \geq m + e(H)$.*

Proof. Consider an $\tilde{r}(G,H)$-game in which Painter uses an $\mathcal{F}$-blocking strategy. Further suppose Builder wins by claiming edges $e_1, \ldots, e_r$. Since Builder choosing an edge which Painter colours blue has no effect on Painter's subsequent choices, without loss of generality we may assume that there exists i such that Painter colours $e_1, \ldots, e_i$ red and $e_{i+1}, \ldots, e_r$ blue. Let $R \subseteq K_{\mathbb{N}}$ be the subgraph with edge set $\{e_1, \ldots, e_i\}$, and let $B \subseteq K_{\mathbb{N}}$ be the subgraph with edge set $\{e_{i+1}, \ldots, e_r\}$. Thus R is the uncovered red graph and B is the uncovered blue graph.

We will show that R is an $\mathcal{F}$-scaffolding for H. First note that R is $\mathcal{F}$-free by Painter's strategy, and R has no isolated vertices by definition. Moreover, since $G \in \mathcal{F}$ and Builder wins, there exists $H' \subseteq B$ with $H' \simeq H$. So $e(B) \geq e(H)$. Moreover, by Painter's strategy all edges in B must be $(R, \mathcal{F})$-forceable, so H is $(R, \mathcal{F})$-forceable. Hence R is an $\mathcal{F}$-scaffolding for H, so $e(R) \geq m$. Therefore, Builder wins in $r \geq e(R) + e(B) \geq m + e(H)$ rounds. $\square$

Therefore, to bound $\tilde{r}(G, H)$ below, it suffices to bound the number of edges in an $\mathcal{F}$-scaffolding for H below for some family $\mathcal{F}$ of graphs with $G \in \mathcal{F}$. To prove Theorem 5.1.4, we set $G = P_{k+1}$ and $\mathcal{F} = \{P_{k+1}\} \cup \{C_i : i \geq 3\}$. Thus an $\mathcal{F}$-free graph is a forest whose components have diameter less than k. Lemma 5.3.7 gives a lower bound on the number of edges in an $\mathcal{F}$-scaffolding for H. Theorem 5.1.4 then follows immediately from Lemma 5.3.7 and Proposition 5.3.3.

Note that replacing $\mathcal{F}$ by $\{P_{k+1}\}$ and attempting a similar proof yields a worse lower bound in some cases. For example, taking $H = P_{2k+1}$ with $k \geq 3$, if Painter follows the $\{P_{k+1}\}$-blocking strategy then Builder can win in $3k$ moves by first constructing a red C_k.

We will see in the proof of Lemma 5.3.7 that if R is a red $\mathcal{F}$-free graph with no isolated vertices, and $X \subseteq V(R)$ is the set of endpoints of P_k's whose vertices lie in R, then Builder may force at most $\Delta(H)(|R| + |X|)$ edges of H using R. It will therefore be very useful to bound $|R| + |X|$ above in terms of $e(R)$, first in the special case where R is a tree (see Lemma 5.3.4) and then in general (see Lemma 5.3.5).

Lemma 5.3.4 *Let $k, m \in \mathbb{N}$ with $k \geq 2$. Let R be a P_{k+1}-free tree with m edges. Let X*

be the set of endpoints of P_k's whose vertices lie in R. Then

$$|R| + |X| \leq \max\{2m - k + 4, 2m\}.$$

Moreover, if $X \neq \emptyset$, then $|R| + |X| \leq 2m - k + 4$.

Proof. If $k = 2$, then $R = K_2$ and so $|R| + |X| = 2m + 2$ and we are done. If $X = \emptyset$, then $|R| + |X| = |R| = m + 1 \leq 2m$ and we are done. We may therefore assume that $k \geq 3$ and $X \neq \emptyset$.

We claim that if $x \in X$, then x is a leaf of R. Indeed, let P be a P_k with one endpoint equal to x. Let $y \in V(P)$ be the neighbour of x in P, and suppose $xz \in E(R)$ for some $z \neq y$. Then either $z \in V(P)$ and $xzPx$ is a cycle in R, or $z \notin V(P)$ and Pxz is a P_{k+1} in R – both are contradictions. Hence if $x \in X$, then x is a leaf. But since $X \neq \emptyset$, R contains a P_k and hence at least $k - 2$ vertices of degree greater than 1. Hence

$$|R| + |X| \leq |R| + |R| - (k - 2) = 2m - k + 4,$$

and the proposition follows. □

Lemma 5.3.5 *Let $k, m \in \mathbb{N}$ with $k \geq 2$. Let R be a P_{k+1}-free forest with m edges and no isolated vertices. Let X be the set of all endpoints of P_k's whose vertices lie in R. Then*

$$|R| + |X| \leq \begin{cases} 4m & \text{if } k = 2, \\ 5m/2 & \text{if } k = 3, \\ 2m - q(k-4) & \text{if } k \geq 4, \end{cases}$$

where q is the number of components of R containing a P_k. Moreover, if $k \geq 4$ and there exists an edge e such that $R + e$ contains a P_{k+1}, then $|R| + |X| \leq 2m - k + 4$.

Proof. Let $R_1, \ldots, R_r$ be the components of R. Let $m_i = e(R_i)$ and $X_i = X \cap V(R_i)$ for all $1 \leq i \leq r$. If $k = 2$, by Lemma 5.3.4 we have

$$|R| + |X| = \sum_{i=1}^{r}(|R_i| + |X_i|) \leq \sum_{i=1}^{r}(2m_i + 2) = 2(m + r) \leq 4m.$$

Similarly if $k \geq 4$, suppose without loss of generality that $R_1, \ldots, R_q$ are the components of R containing P_k's. Then by Lemma 5.3.4 we have

$$|R| + |X| \leq \sum_{i=1}^{q}(2m_i - k + 4) + \sum_{i=q+1}^{r} 2m_i = 2m - q(k - 4). \tag{5.3.6}$$

Suppose $k = 3$. Without loss of generality, let $R_1, \ldots, R_{r'}$ be those components of R which consist of a single edge. (Note that we may have $r' = 0$.) Then $m = r' + \sum_{i=r'+1}^{r} m_i$ and $r - r' \leq m/2$. Then by Lemma 5.3.4 we have

$$\begin{aligned}
|R| + |X| &= \sum_{i=1}^{r'}(|R_i| + |X_i|) + \sum_{i=r'+1}^{r}(|R_i| + |X_i|) \\
&\leq 2r' + \sum_{i=r'+1}^{r}(2m_i + 1) \\
&= 2m + r - r' \leq 5m/2
\end{aligned}$$

and so the result follows.

Finally, suppose $k \geq 4$ and there exists an edge e such that $R + e$ contains a P_{k+1}. If $X \neq \emptyset$, then $q \geq 1$ and so $|R| + |X| \leq 2m - k + 4$ by (5.3.6). Hence we may assume that $X = \emptyset$, and so e is an edge between two vertices of R. It follows that R contains two vertex-disjoint paths of combined length at least $k - 1$, and hence that

$$|R| + |X| = |R| = m + r \leq m + (m - k + 3) < 2m - k + 4,$$

as desired. The first inequality follows since all edges in a given path must lie in the same component of R. $\square$

Lemma 5.3.7 *Let $k, \ell, d \in \mathbb{N}$ with $k \geq 2$. Let H be a graph with $e(H) = \ell$ and $\Delta(H) = d$. Let $\mathcal{F} = \{P_{k+1}\} \cup \{C_i : i \geq 3\}$. Suppose R is an $\mathcal{F}$-scaffolding for H. Then, we have*

$$e(R) \geq \begin{cases} \ell/(2d) & \text{if } k = 2, \\ 4\ell/(5d) & \text{if } k = 3, \\ \ell/d & \text{if } k \geq 4. \end{cases}$$

Moreover, if H is connected then

$$e(R) \geq \begin{cases} \frac{2}{5}\left\lceil \frac{2\ell+1}{d} \right\rceil & \text{if } k = 3, \\ \min\left\{\frac{\ell}{d} + \frac{k-4}{2}, \frac{2\ell}{d} - 1\right\} & \text{if } k \geq 4. \end{cases}$$

Proof. Let $m = e(R)$. Note that R is a P_{k+1}-free forest with m edges and no isolated vertices. Let X be the set of endpoints of P_k's whose vertices lie in R and let $Y = V(R) \backslash X$.

We first claim that any $(R, \mathcal{F})$-forceable edge is either incident to X or internal to Y. Suppose not. Then there exist $y \in Y$ and $z \notin V(R)$ such that yz is a forceable edge. Let $F \in \mathcal{F}$ be such that $F \subseteq R+e$. Note that $e \in E(F)$, since R is $\mathcal{F}$-free. Since $d_{R+e}(z) = 1$, we have $F = P_{k+1}$. But then y is an endpoint of a P_k in R, contradicting $y \in Y$.

Let H' be a forced copy of H. Then H' contains at most $d|X|$ edges incident to X, and at most $d|Y|/2$ edges internal to Y. All edges of H' are forceable, so it follows that

$$\ell = e(H') \leq d|X| + \frac{d|Y|}{2} = \frac{d(|R| + |X|)}{2}. \tag{5.3.8}$$

Then (5.3.8) and the first case of Lemma 5.3.5 imply the lemma holds when $k = 2$ or H is not connected.

Now suppose H is connected and $k \geq 4$. If there exists an edge e such that $R+e$ contains a P_{k+1}, then $|R|+|X| \leq 2m-k+4$ by Lemma 5.3.5. Hence, (5.3.8) implies that $m \geq \frac{\ell}{d} + \frac{k-4}{2}$. Therefore, we may assume that no such edge exists, and in particular that $X = \emptyset$. This implies that R is a $\{C_i : i \geq 3\}$-scaffolding for H. Since no edge between components of R is $(R, \{C_i : i \geq 3\})$-forceable, and H is connected, we may assume that R is connected and therefore a tree. Hence, $|R| = m+1$. Moreover, (5.3.8) implies that $m \geq 2\ell/d - 1$. Therefore

$$m \geq \min\left\{\frac{\ell}{d} + \frac{k-4}{2}, \frac{2\ell}{d} - 1\right\}$$

in all cases, as required.

Finally, suppose H is connected and $k = 3$. First suppose $X = \emptyset$, so that R is a matching. Note that $m \geq 2$, or there would be no $(R, \mathcal{F})$-forceable edges. If $\ell \leq 2d$, it follows that

$$m \geq 2 = \frac{2}{5}\left(\frac{4d+d}{d}\right) \geq \frac{2}{5}\left(\frac{2\ell}{d} + 1\right) \geq \frac{2}{5}\left\lceil\frac{2\ell+1}{d}\right\rceil,$$

as desired. Otherwise, if $\ell > 2d$, we have

$$m = \frac{|R|}{2} \overset{(5.3.8)}{\geq} \frac{\ell}{d} = \frac{2}{5}\left(\frac{4\ell+\ell}{2d}\right) > \frac{2}{5}\left(\frac{2\ell}{d} + 1\right) \geq \frac{2}{5}\left\lceil\frac{2\ell+1}{d}\right\rceil,$$

as desired.

We may therefore assume that $X \neq \emptyset$. Moreover, $Y \neq \emptyset$. (Indeed, since R is a P_4-free forest only leaves of R can be elements of X. Since $X \neq \emptyset$, R contains a P_3 and hence a non-leaf.) Since H is connected, H' either contains an edge between X and Y, consists entirely of edges incident to X, or consists entirely of edges internal to Y. We will show that in all three cases, we have

$$\ell \leq \frac{d(|R| + |X|) - 1}{2}. \tag{5.3.9}$$

If there is an edge of H' between X and Y, then there are at most $(d|Y| - 1)/2$ edges internal to Y and so (5.3.9) holds by a calculation similar to that of (5.3.8). If H' consists entirely of edges incident to X, then H' contains at most $d|X|$ edges. Since $Y \neq \emptyset$, we have $d|X| < d(|R| + |X|)/2$ and so (5.3.9) holds. Finally, suppose all edges of H' are internal to Y. Then

$$\ell = e(H') \leq \frac{d|Y|}{2} \leq \frac{d(|R| - 1)}{2},$$

where the last inequality follows since $X \neq \emptyset$, and so again (5.3.9) holds. Hence (5.3.9) holds in all cases.

It now follows from (5.3.9) that $|R| + |X| \geq \lceil (2\ell + 1)/d \rceil$, and so Lemma 5.3.5 implies that $m \geq (2/5) \cdot \lceil (2\ell + 1)/d \rceil$ as required. □

Theorem 5.1.4 now follows immediately from Proposition 5.3.3 and Lemma 5.3.7.

5.4 Determining $\tilde{r}(P_3, P_{\ell+1})$ for $\ell \geq 2$

Theorem 5.1.4 implies that $\tilde{r}(P_3, P_{\ell+1}) \geq \lceil 5\ell/4 \rceil$ for $\ell \geq 2$. To bound $\tilde{r}(P_3, P_{\ell+1})$ above, we shall present a strategy for Builder. In the discussion that follows, we assume for clarity that Painter will never voluntarily lose the $\tilde{r}(P_3, P_{\ell+1})$-game.

Builder will use the threat of a red P_3 to force a blue $P_{\ell+1}$. First, Builder will use Lemma 5.4.1 to construct a blue path P with one endpoint incident to a red edge. Builder will then use a procedure outlined in Lemma 5.4.2 to efficiently extend P until it has length between $\ell - 4$ and ℓ. Finally, Builder will carefully extend P into a blue $P_{\ell+1}$, yielding a tight upper bound for $\tilde{r}(P_3, P_{\ell+1})$ (see Theorem 5.4.3).

Lemma 5.4.1 *Let $q \in \mathbb{N}$ with $q \geq 5$. Builder can force one of the following structures independent of Painter's choices:*

(i) a red P_3 in at most $q-1$ rounds.

(ii) a blue P_q in $q-1$ rounds.

(iii) a blue P_t with one endpoint incident to a red edge in t rounds for some $4 \leq t \leq q-1$.

Proof. Builder first chooses an arbitrary vertex x_1, then proceeds as follows. Suppose that Builder has already obtained a blue path $x_1 \dots x_i$ in $i-1$ rounds for some $1 \leq i < q$. Builder then chooses the edge $x_i x_{i+1}$, where x_{i+1} is a new vertex. If Painter colours $x_i x_{i+1}$ blue, we have obtained a blue path $x_1 \dots x_{i+1}$ in i rounds, and so if $i+1 < q$ we may repeat the process. If Painter colours all such edges blue, we will obtain a blue path $x_1 \dots x_q$ in $q-1$ rounds and achieve (ii). Suppose instead that for some $1 \leq i \leq q-1$, within i rounds we obtain a path $x_1 \dots x_{i+1}$ such that $x_1 \dots x_i$ is blue and $x_i x_{i+1}$ is red. If $i \geq 4$ then we have achieved (iii), so suppose in addition $i \leq 3$.

First suppose $i \in \{1, 2\}$. In this case, Builder chooses the two edges $x_i v$ and $v x_{i+1}$ where v is a new vertex. If $i = 1$, Builder also chooses the edge $x_{i+1} w$ where w is a new vertex. If Painter colours $x_i v$, $v x_{i+1}$ or $x_{i+1} w$ red, then $x_{i+1} x_i v$, $v x_{i+1} x_i$ or $x_i x_{i+1} w$ respectively is a red P_3 and we have achieved (i). Otherwise, we have achieved (iii). Indeed, if $i = 1$ then $x_1 v x_2 w$ is a blue P_4 constructed in 4 rounds with x_1 incident to the red edge $x_1 x_2$, and if $i = 2$ then $x_1 x_2 v x_3$ is a blue P_4 constructed in 4 rounds with x_3 incident to the red edge $x_3 x_2$.

Finally, suppose $i = 3$. Then Builder chooses the edge $x_4 x_1$. If Painter colours the edge red, then $x_3 x_4 x_1$ is a red P_3 and we have achieved (i), so suppose Painter colours the edge blue. Then $x_4 x_1 x_2 x_3$ is a blue P_4 constructed in 4 rounds with x_3 incident to the red edge $x_3 x_4$, so we have achieved (iii). $\square$

Lemma 5.4.2 *Let $\ell \in \mathbb{N}$ with $\ell \geq 4$. Builder can force one of the following structures independent of Painter's choices:*

(i) a red P_3 in at most $5\ell/4 - 1$ rounds.

(ii) a blue $P_{\ell+1}$ in at most $5\ell/4 - 1$ rounds.

(iii) a blue P_t with one endpoint incident to a red edge in at most $5t/4 - 1$ rounds for some $\ell - 3 \leq t \leq \ell$.

Proof. Throughout the proof, we assume for clarity that Painter will always avoid (i) and (ii) if possible. By Lemma 5.4.1 (taking $q = \ell + 1$) we may assume that Builder has constructed a blue P_t, say $v_1 \dots v_t$, which satisfies

($*$) $v_1 \dots v_t$ has one endpoint incident to a red edge $v_1 u$, and Builder constructed $v_1 \dots v_t$ in at most $5t/4 - 1$ rounds. Moreover, $4 \leq t \leq \ell$.

Note that $t \leq 5t/4 - 1$ since $t \geq 4$.

If $t \geq \ell - 3$, then we have achieved (iii). Hence, we may assume that $4 \leq t < \ell - 3$. Without loss of generality, let $v_1 u$ be a red edge as in ($*$). Builder will extend $v_1 \dots v_t$ as follows. We apply Lemma 5.4.1 with $q = \ell - t + 1 \geq 5$ on a set of new vertices. We split into cases depending on Painter's choice.

Case 1: Builder obtains a red P_3 in at most $\ell - t$ rounds, as in Lemma 5.4.1(i).

In this case, Builder has spent at most $5t/4 - 1 + \ell - t \leq 5\ell/4 - 2$ rounds in total since $t \leq \ell - 4$, and so we have achieved (i).

Case 2: Builder obtains a blue path $w_1 \dots w_{\ell-t+1}$ in $\ell - t$ rounds, as in Lemma 5.4.1(ii).

In this case, Builder has again spent at most $5\ell/4 - 2$ rounds in total. Builder now chooses the edge $w_1 v_1$. If Painter colours it red, then $w_1 v_1 u$ is a red P_3 and we have achieved (i). If Painter colours it blue, then $w_{\ell-t+1} \dots w_1 v_1 \dots v_t$ is a blue $P_{\ell+1}$ and we have achieved (ii).

Case 3: Builder obtains a blue path $w_1 \dots w_{t'}$ and a red edge w_1x in at most t' rounds for some $4 \le t' \le \ell - t$, as in Lemma 5.4.1(iii).

In this case, Builder has spent at most

$$\frac{5t}{4} - 1 + t' = \frac{5t}{4} + \frac{5t'}{4} - \frac{t'}{4} - 1 \le \frac{5(t+t')}{4} - 2 \le \frac{5\ell}{4} - 2$$

rounds in total. Builder now chooses the edge v_tw_1. If Painter colours it red, then v_tw_1x is a red P_3 and we have achieved (i). If Painter colours it blue, then $v_1 \dots v_tw_1 \dots w_{t'}$ is a blue $P_{t+t'}$ with v_1 incident to the red edge v_1u. Moreover, this $P_{t+t'}$ satisfies $(*)$ with $t + t' > t$. Hence by iterating the argument above, the result follows. $\square$

Theorem 5.4.3 *For all $\ell \ge 2$, $\tilde{r}(P_3, P_{\ell+1}) = \lceil 5\ell/4 \rceil$.*

Proof. Theorem 5.1.4 implies that $\tilde{r}(P_3, P_{\ell+1}) \ge \lceil 5\ell/4 \rceil$. It therefore suffices to prove that Builder can win the $\tilde{r}(P_3, P_{\ell+1})$-game within $\lceil 5\ell/4 \rceil$ rounds. First note that $\tilde{r}(P_3, P_3) = 3$ and $\tilde{r}(P_3, P_4) = 4$, as shown by Grytczuk, Kierstead and Prałat [41] and Prałat [78] respectively, so we may assume $\ell \ge 4$. Applying Lemma 5.4.2, either Builder obtains a blue path $v_1 \dots v_{t+1}$ and a red edge v_1u in at most $5(t+1)/4 - 1$ rounds for some $\ell - 3 \le t + 1 \le \ell$ or we are done. Write

$$r(t) = \left\lceil \frac{5\ell}{4} \right\rceil - \left(\left\lfloor \frac{5(t+1)}{4} \right\rfloor - 1 \right) = \left\lceil \frac{\ell}{4} \right\rceil - \left\lfloor \frac{t+1}{4} \right\rfloor + (\ell - t),$$

and note that Builder has at least $r(t)$ rounds left to construct either a red P_3 or a blue $P_{\ell+1}$. We now split into cases depending on the precise value of t.

Case 1: $t = \ell - 1$, so that $r(t) = 1$.

Builder chooses the edge v_0v_1, where v_0 is a new vertex. If Painter colours it red, then v_0v_1u is a red P_3 and we are done. Otherwise, $v_0v_1 \dots v_\ell$ is a blue $P_{\ell+1}$ and we are done.

Case 2: $t = \ell - 2$, so that $r(t) \ge 3$.

Builder chooses the edge $v_{\ell-1}x$, where x is a new vertex. If Painter colours it blue, then we are in Case 1 with an extra round to spare. If Painter colours it red, Builder chooses the edges $v_{\ell-1}w$ and wx, where w is a new vertex. If Painter colours either edge red then $xv_{\ell-1}w$ or $wxv_{\ell-1}$ respectively is a red P_3 and we are done. Otherwise, $v_1 \dots v_{\ell-1}wx$ is a blue $P_{\ell+1}$ and we are done.

Case 3: $t = \ell - 3$, so that $r(t) \geq 4$.

Builder chooses the edge $v_{\ell-2}x$, where x is a new vertex. If Painter colours it blue, then we are in Case 2. If Painter colours it red, Builder chooses the edges $v_{\ell-2}w$, wx and xy, where w and y are new vertices. If Painter colours any of these edges red then $xv_{\ell-2}w$, $wxv_{\ell-2}$ or $v_{\ell-2}xy$ respectively is a red P_3 and we are done. Otherwise, $v_1 \dots v_{\ell-2}wxy$ is a blue $P_{\ell+1}$ and we are done.

Case 4: $t = \ell - 4$, so that $r(t) \geq 5$.

Builder chooses the edge $v_{\ell-3}x$, where x is a new vertex. If Painter colours it blue, then we are in Case 3. If Painter colours it red, Builder chooses the edges v_0v_1, $v_{\ell-3}w$, wx and xy, where v_0, w and y are new vertices. If Painter colours any of these edges red then v_0v_1u, $xv_{\ell-3}w$, $wxv_{\ell-3}$ or $v_{\ell-3}xy$ respectively is a red P_3 and we are done. Otherwise, $v_0v_1 \dots v_{\ell-3}wxy$ is a blue $P_{\ell+1}$ and we are done. □

5.5 Determining $\tilde{r}(P_3, C_\ell)$ for $\ell \geq 3$

Our aim is to prove Theorem 5.1.7, i.e. to determine $\tilde{r}(P_3, C_\ell)$ for all $\ell \geq 3$. As a warmup, we first determine $\tilde{r}(P_3, C_3)$ and $\tilde{r}(P_3, C_4)$. Note that Theorem 5.1.4 implies that $\tilde{r}(P_3, C_3) \geq 5\ell/4$ for all $\ell \geq 3$, but this lower bound is too weak when $\ell \leq 4$. Instead, we consider the $\{C_\ell\}$-blocking strategy for Painter in an $\tilde{r}(C_\ell, P_3)$-game.

Proposition 5.5.1 *For all $\ell \geq 3$, we have $\tilde{r}(P_3, C_\ell) \geq \ell + 2$.*

Proof. We consider the $\{C_\ell\}$-blocking strategy for Painter in the $\tilde{r}(C_\ell, P_3)$-game. Let R be an edge-minimal $\{C_\ell\}$-scaffolding for P_3. Then R must contain two distinct P_ℓ's, so $e(R) \geq \ell$. The result therefore follows from Proposition 5.3.3. $\square$

The upper bounds are both relatively straightforward.

Proposition 5.5.2 *We have $\tilde{r}(P_3, C_3) = 5$ and $\tilde{r}(P_3, C_4) = 6$.*

Proof. By Proposition 5.5.1, we have $\tilde{r}(P_3, C_3) \geq 5$ and $\tilde{r}(P_3, C_4) \geq 6$. It is easy to show that $r(P_3, C_4) = 4$ (see e.g. Radziszowski [80]), so we also have $\tilde{r}(P_3, C_4) \leq \binom{4}{2} = 6$ as Builder may simply choose the edges of a K_4. It therefore suffices to prove that Builder can win the $\tilde{r}(P_3, C_3)$-game in 5 rounds.

Take new vertices u, v, w, x, y and z. Builder first chooses the edges uv, uw and ux. If Painter colours more than one of these edges red, then we have obtained a red P_3 and we are done.

Suppose Painter colours uv, uw and ux blue. Then Builder chooses the edges vw and wx. If Painter colours either edge blue, then $vwuv$ or $wxuw$ respectively is a blue C_3 and we are done. If Painter colours both edges red, then vwx is a red P_3 and we are done.

Finally, suppose Painter colours (without loss of generality) uv red, but uw and ux blue. Then Builder chooses the edge xy. If Painter colours xy red, Builder chooses the edge wx, yielding either a red P_3 (namely wxy), or a blue C_3, $wxuw$, and we are done. If Painter colours xy blue, Builder chooses the edge yu, yielding either a red P_3 (namely yuv) or a blue C_3 (namely $uxyu$), and we are done. $\square$

We now determine $\tilde{r}(P_3, C_\ell)$ for $\ell \geq 5$. As in Section 5.4, Builder's strategy will be to build up a long blue path using Lemma 5.4.2. Builder will then carefully close this path into a blue C_ℓ.

Theorem 5.5.3 *For all $\ell \geq 5$, $\tilde{r}(P_3, C_\ell) = \lceil 5\ell/4 \rceil$.*

Proof. Theorem 5.1.4 implies that $\tilde{r}(P_3, C_\ell) \geq \lceil 5\ell/4 \rceil$. It therefore suffices to prove that Builder can win the $\tilde{r}(P_3, C_\ell)$-game within $\lceil 5\ell/4 \rceil$ rounds. By Lemma 5.4.2, Builder can force one of the following structures independent of Painter's choices:

(i) a red P_3 in at most $5(\ell-1)/4 - 1$ rounds.

(ii) a blue P_ℓ in at most $5(\ell-1)/4 - 1$ rounds.

(iii) a blue P_t with one endpoint incident to a red edge in at most $5t/4 - 1$ rounds for some $\ell - 4 \leq t \leq \ell - 1$.

If Painter chooses (i), then we are done. Suppose Painter chooses (ii), so that Builder has at least

$$\left\lceil \frac{5\ell}{4} \right\rceil - \left(\frac{5(\ell-1)}{4} - 1 \right) = \left\lceil \frac{5\ell}{4} \right\rceil - \frac{5\ell}{4} + \frac{9}{4} > 2$$

rounds to construct a red P_3 or a blue C_ℓ. (Thus Builder has at least 3 rounds.) Let $v_1 \dots v_\ell$ be the corresponding blue path. Then Builder chooses the edges $v_\ell v_1$, $v_1 v_3$ and $v_\ell v_2$. If Painter colours $v_\ell v_1$ blue then $v_1 \dots v_\ell v_1$ is a blue C_ℓ and we are done. If Painter colours $v_\ell v_1$ red and $v_1 v_3$ or $v_\ell v_2$ red, then $v_\ell v_1 v_3$ or $v_1 v_\ell v_2$ respectively is a red P_3 and we are done. Finally, if Painter colours both $v_1 v_3$ and $v_\ell v_2$ blue, then $v_1 v_3 v_4 \dots v_\ell v_2 v_1$ is a blue C_ℓ and we are done.

Finally, suppose Painter chooses (iii). Let $v_1 \dots v_t$ be the corresponding blue path and let $v_1 u$ be a red edge. Write

$$r(t) = \left\lceil \frac{5\ell}{4} \right\rceil - \left(\left\lfloor \frac{5t}{4} \right\rfloor - 1 \right) = \left\lceil \frac{\ell}{4} \right\rceil - \left\lfloor \frac{t}{4} \right\rfloor + \ell - t + 1,$$

so that Builder has at least $r(t)$ rounds left to construct either a red P_3 or a blue C_ℓ. We split into cases depending on the precise value of t.

Case 1: $t = \ell - 1$, so that $r(t) \geq 3$.

Builder first chooses the edge $v_{\ell-1}w$, where w is a new vertex. If Painter colours $v_{\ell-1}w$ blue, then Builder chooses the edge wv_1. If Painter colours wv_1 red then wv_1u is a red P_3, and if Painter colours wv_1 blue then $v_1v_2\dots v_{\ell-1}wv_1$ is a blue C_ℓ. Now suppose Painter colours $v_{\ell-1}w$ red instead. Then Builder chooses the edges $v_{\ell-1}x$ and xv_1, where x is a new vertex. If Painter colours either edge red, then $wv_{\ell-1}x$ or xv_1u respectively is a red P_3 and we are done. Otherwise, $v_1\dots v_{\ell-1}xv_1$ is a blue C_ℓ and we are done.

Case 2: $t = \ell - 2$, so that $r(t) \geq 4$.

Builder first chooses the edge $v_{\ell-2}w$, where w is a new vertex. If Painter colours $v_{\ell-2}w$ blue then we are in Case 1, so suppose Painter colours $v_{\ell-2}w$ red. Builder then chooses the edges $v_{\ell-2}x$, xw and wv_1, where x is a new vertex. If Painter colours any of these edges red, then $wv_{\ell-2}x$, $xwv_{\ell-2}$ or $v_{\ell-2}wv_1$ respectively is a red P_3 and we are done. Otherwise, $v_1v_2\dots v_{\ell-2}xwv_1$ is a blue C_ℓ and we are done.

Case 3: $t = \ell - 3$, so that $r(t) \geq 5$.

Builder first chooses the edge $v_{\ell-3}w$, where w is a new vertex. If Painter colours $v_{\ell-3}w$ blue then we are in Case 2, so suppose Painter colours $v_{\ell-3}w$ red. Builder then chooses the edges $v_{\ell-3}x$, xw, wy and yv_1, where x and y are new vertices. If Painter colours any of these edges red, then $wv_{\ell-3}x$, $xwv_{\ell-3}$, $v_{\ell-3}wy$ or yv_1u respectively is a red P_3 and we are done. Otherwise, $v_1v_2\dots v_{\ell-3}xwyv_1$ is a blue C_ℓ and we are done.

Case 4: $t = \ell - 4$, so that $r(t) \geq 6$.

Builder first chooses two edges wx and xy, where w, x and y are new vertices. If Painter colours both edges red, wxy is a red P_3 and we are done. Now suppose that Painter colours one edge blue and one red, say wx red and xy blue. Then Builder chooses the edges $v_{\ell-4}w$, wz, zx and yv_1, where z is a new vertex. If Painter colours any of these edges red, then $v_{\ell-4}wx$, xwz, zxw or yv_1u respectively is a red P_3 and we are done. Otherwise, $v_1v_2\dots v_{\ell-4}wzxyv_1$ is a blue C_ℓ and we are done.

We may therefore assume that Painter colours both wx and xy blue. Builder now chooses the edge $v_{\ell-4}w$. If Painter colours $v_{\ell-4}w$ blue, we are in Case 1 (taking our path to be $v_1v_2\dots v_{\ell-4}wxy$), so suppose Painter colours $v_{\ell-4}w$ red. Then Builder chooses the edges $v_{\ell-4}z$, zw and yv_1, where z is a new vertex. If Painter colours any of these edges red, then $wv_{\ell-4}z$, $zwv_{\ell-4}$ or yv_1u respectively is a red P_3 and we are done. Otherwise, $v_1v_2\dots v_{\ell-4}zwxyv_1$ is a blue C_ℓ and we are done. □

5.6 Bounding $\tilde{r}(P_4, P_{\ell+1})$ for $\ell \geq 3$

Theorem 5.1.4 implies that $\tilde{r}(P_4, P_{\ell+1}) \geq (7\ell + 2)/5$ for $\ell \geq 3$. In order to prove Theorem 5.1.5 for the case when $k = 3$, it therefore suffices to bound $\tilde{r}(P_4, P_{\ell+1})$ above, which we do in Theorem 5.6.24. In the following discussion we take on the role of Builder, and we will assume for clarity that Painter will not voluntarily lose the game by creating a red P_4. Finally, note that throughout this section the variable R will be used to refer to a path, not a scaffolding.

We will employ the following strategy to construct a blue $P_{\ell+1}$. We will obtain two (initially trivial) vertex-disjoint blue paths Q and R, repeatedly extend them, and then join them together to form a blue $P_{\ell+1}$ when they are sufficiently long. Here Q is distinct from R in that we require one of Q's endpoints to be incident to a red edge bc disjoint from $V(R)$. Some of our methods for extending a blue path require this property, and others destroy it. Thus at each stage we will extend either Q or R depending on which of our extension methods Painter allows us to use.

We will use the following lemma to join Q and R together (and sometimes to extend Q).

Lemma 5.6.1 *Let Q be a (possibly trivial) blue path with endpoints a and b, where b is incident to a red edge bc. Let R be a (possibly trivial) blue path vertex-disjoint from*

$V(Q) \cup \{c\}$. *Then Builder can force Painter to construct one of the following while uncovering at most* 2 *edges:*

(i) a blue path Q' *of length* $e(Q) + e(R) + 1$ *with one endpoint incident to a red edge.*

(ii) a red P_4.

Proof. First suppose that R is non-trivial, and let x and y be the endpoints of R. Moreover, suppose that either $a = c$ or Q is trivial, so that both endpoints of Q are incident to bc. Builder chooses the edges bx and cy. If Painter colours both edges red, then $xbcy$ is a red P_4. Hence, without loss of generality, we may assume that Painter colours bx blue. Then $Q' := aQbxRy$ is a blue path of length $e(Q) + e(R) + 1$, where a is incident to the red edge bc.

Now suppose that Q is non-trivial and $a \neq c$. Builder chooses the edge ax. If Painter colours ax blue, then $bQaxRy$ is a blue path of length $e(Q) + e(R) + 1$ with endpoint b incident to the red edge bc. So we may assume that Painter colours ax red. Builder then chooses the edge bx. If Painter colours bx red, then $cbxa$ is a red P_4. Otherwise $Q' := aQbxRy$ is a blue path of length $e(Q) + e(R) + 1$ where a is incident to the red edge ax.

Finally, suppose R is trivial with endpoint x. Let y be a new vertex. Then the argument above implies the lemma on replacing xRy with x throughout. □

The arguments that follow are by necessity somewhat technical. The reader may therefore find the following intuition useful.

(i) For every seven edges we uncover, we will extend either Q or R by five blue edges.

(ii) When we join Q and R, $e(Q) + e(R) + 1$ should not be too much greater than ℓ.

It is clear that following the above principles will yield a bound of the form $\tilde{r}(P_4, P_{\ell+1}) \leq 7\ell/5 + C$ for some constant C. We will violate (i) in the first and last phases of Builder's strategy, but this introduces only constant overhead.

Before we can apply Lemma 5.6.1 to join Q and R and obtain a blue $P_{\ell+1}$, we must extend them until $e(Q) + e(R) + 1 \geq \ell$. Each time we extend Q and R, we require two independent edges of the same colour. (Naturally, we can obtain these by choosing three independent edges.) If these edges are blue, we may extend Q efficiently using Lemma 5.6.5 (see Section 5.6.1). If they are red, we may extend either Q or R efficiently using Lemma 5.6.18 (see Section 5.6.2). Note that the latter case is significantly harder. We then apply Lemmas 5.6.5 and 5.6.18 repeatedly to prove Theorem 5.6.24 (see Section 5.6.3).

In our figures throughout the section, we shall represent blue edges with solid lines and red edges with dotted lines.

5.6.1 Extending Q using two independent blue edges e and f.

Throughout this subsection, e and f will be two independent blue edges vertex-disjoint from Q and R. We will prove that we can use these two edges to efficiently extend Q – see Lemma 5.6.5. We first define a special type of path which will be important to the extension process.

Definition 5.6.2 *We say that a path $xySz$ is* of type A *if xy is a red edge and S is a non-trivial blue path with endpoints y and z.*

Note that the above definition requires $x \notin V(S)$. For the remainder of the section, if we refer to a path $xySz$ of type A, we shall take it as read that x, y, z and S are as in Definition 5.6.2.

We now sketch the proof of Lemma 5.6.5. By greedily extending the blue edge e into a path, Builder can obtain either a long blue path or a path of type A (see Lemma 5.6.3). If Builder obtains a long blue path P, then we can simply join P and Q together using Lemma 5.6.1. Suppose instead Builder obtains a path $xySz$ of type A. Then we use

Lemma 5.6.4 to efficiently join S and Q together. In either case, the resulting blue path Q' also has an endpoint incident to a red edge, so Q' retains the defining property of Q.

We first prove that Builder can obtain either a long blue path or a path of type A by greedily extending e.

Lemma 5.6.3 *Let $m \in \mathbb{N}$ and let e be a blue edge. Then Builder can force Painter to construct one of the following:*

(i) a path $xySz$ of type A with $e(S) = t$ while uncovering t edges for some $1 \leq t < m$.

(ii) a blue path of length m while uncovering $m - 1$ edges.

Proof. Let S_1 be the blue path formed by e. Builder proceeds to extend S_1 greedily until either Builder has constructed a blue path of length m or Painter has coloured an edge red.

Indeed, suppose S_i is a blue path of length i for some $1 \leq i \leq m-1$ with endpoints y and z, and that Builder has uncovered $i - 1$ edges in forming S_i from S_1. Then Builder chooses the edge xy, where x is a new vertex. If Painter colours xy red then xyS_iz is a path of type A with $e(S_i) = i$, where $1 \leq i < m$. Moreover, Builder has uncovered i edges in constructing it, and so we have achieved (i). If instead Painter colours xy blue, then $S_{i+1} := xyS_iz$ is a blue path of length $i + 1$ and Builder has uncovered i edges in constructing it.

By repeating this process, Builder must either obtain a path of type A as in (i) or a blue path S_m of length m as in (ii). □

We now prove that Builder can use a path of type A to efficiently extend Q. Recall that we were given two independent blue edges, e and f, and that we have already used e to construct a path of type A.

Lemma 5.6.4 *Suppose Q is a non-trivial blue path with endpoints a and b, where b is incident to a red edge bc. Suppose $xySz$ is a path of type A which is vertex-disjoint from*

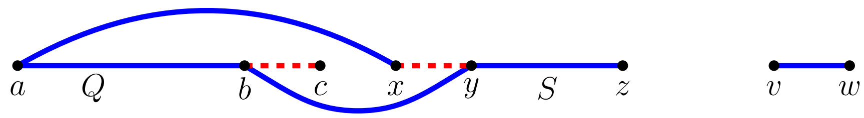

Figure 5.1: Extending Q using a path of type A as in Lemma 5.6.4(i).

$V(Q) \cup \{c\}$. *Further suppose that* $f = vw$ *is a blue edge vertex-disjoint from* $V(Q) \cup V(xySz) \cup \{c\}$. *Then Builder can force Painter to construct one of the following:*

(i) a blue path Q' *of length* $e(Q) + e(S) + 2$ *with one endpoint* b' *incident to a red edge* $b'c'$ *while uncovering* 2 *edges. Moreover,* f *is vertex-disjoint from* $V(Q') \cup \{c'\}$.

(ii) a blue path Q' *of length* $e(Q)+e(S)+4$ *with one endpoint incident to a red edge* $b'c'$ *while uncovering* 4 *edges. (Note that* f *need not be vertex-disjoint from* $V(Q') \cup \{c'\}$.*)*

(iii) a red P_4 *while uncovering at most* 4 *edges.*

Proof.

Builder chooses the edge ax. First suppose Painter colours ax blue. Builder then chooses the edge by. If Painter colours the edge by red, then $cbyx$ is a red P_3 and we have achieved (iii). Suppose not. Then $Q' := xaQbySz$ (see Figure 5.1) is a blue path of length $e(Q) + e(S) + 2$, where x is incident to the red edge xy, and we have achieved (i).

Now suppose Painter instead colours ax red. Builder then chooses the edges av, wy and xb. If Painter colours any of these edges red, then $yxav$, $wyxa$ or $yxbc$ respectively is a red P_4 and we have achieved (iii). Suppose not. Then $Q' := xbQavwySz$ (see Figure 5.2) is a blue path of length $e(Q) + e(S) + 4$, where x is incident to the red edge xy, and we have achieved (ii). □

We now consolidate Lemmas 5.6.3 and 5.6.4 into a single lemma which says that given two independent blue edges, Builder can efficiently extend Q. In applying Lemma 5.6.5, we will take m to be $\ell - e(Q) - e(R) - 1$. Thus if we can extend Q by at least m edges, then we can join Q and R to obtain a blue $P_{\ell+1}$ immediately afterwards.

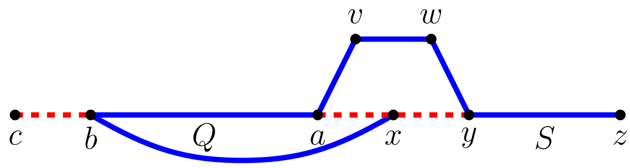

Figure 5.2: Extending Q using a path of type A and an blue independent edge vw as in Lemma 5.6.4(ii).

Lemma 5.6.5 *Let $m \in \mathbb{N}$. Suppose Q is a non-trivial blue path with endpoints a and b, where b is incident to a red edge bc. Suppose e and f are two independent blue edges which are vertex-disjoint from $V(Q) \cup \{c\}$. Then Builder can force Painter to construct one of the following:*

(i) a blue path Q' with $e(Q') = e(Q) + \ell'$ for some $3 \leq \ell' \leq m+3$ such that Q' has an endpoint b' incident to a red edge $b'c'$. A total of ℓ' edges are uncovered in the process. Moreover, if $\ell' < 5 \leq m$, then f is vertex-disjoint from $V(Q') \cup \{c'\}$.

(ii) a red P_4 while uncovering at most $m+3$ edges.

Proof. We apply Lemma 5.6.3 to e and m, and split into cases depending on Painter's choice.

Case 1: As in Lemma 5.6.3(i), we obtain a path $xySz$ of type A with $e(S) = t$ for some $1 \leq t < m$ which is vertex-disjoint from $V(f) \cup V(Q) \cup \{c\}$, while uncovering t edges.

We apply Lemma 5.6.4 to Q, $xySz$ and f. First suppose that as in Lemma 5.6.4(i), we obtain a blue path Q' of length $e(Q)+t+2$ with one endpoint incident to a red edge while preserving f's independence. In total we have uncovered $t+2$ edges. Hence Q' satisfies (i) on setting $\ell' = t+2$.

Now suppose that as in Lemma 5.6.4(ii), we obtain a blue path Q' of length $e(Q)+t+4$ with one endpoint incident to a red edge. We have uncovered $t+4$ edges in total. Hence setting $\ell' = t+4$, we have achieved (i) with $\ell' \geq 5$.

Finally, suppose that as in Lemma 5.6.4(iii) we obtain a red P_4. Then we have uncovered at most $t+4 \leq m+3$ edges in total and so we have achieved (ii).

Case 2: As in Lemma 5.6.3(ii), we obtain a blue path S of length m which is vertex-disjoint from $V(Q) \cup \{c\}$ while uncovering $m-1$ edges.

We apply Lemma 5.6.1 to Q and S to construct either a blue path Q' of length $e(Q)+m+1$ with one endpoint incident to a red edge or a red P_4 while uncovering at most 2 additional edges. We have uncovered at most $m+1$ edges in total. Hence in the former case we have achieved (i), and in the latter case we have achieved (ii). □

5.6.2 Extending Q and R using two red edges e and f.

In this subsection, our aim is to extend Q or R efficiently when given two independent red edges e and f – see Lemma 5.6.18. As in Section 5.6.1, it will be convenient to define some special paths that we will use in the extension process. These paths can be viewed as analogues of paths of type A.

Definition 5.6.6 *A path $vwxyz$ is* of type B *if vw and yz are red edges, and wx and xy are blue edges.*

Definition 5.6.7 *A path $T_1 \dots T_k$ is* of type C *if the following statements hold:*

(C1) k is odd and $k \geq 3$.

(C2) T_1 is either a blue edge or a path of the form $x_1y_1z_1$, where $z_1 \in V(T_2)$ and y_1z_1 is red (and x_1y_1 may be red or blue).

(C3) T_k is either a blue edge or a path of the form $x_ky_kz_k$, where $x_k \in V(T_{k-1})$ and x_ky_k is red (and y_kz_k may be red or blue).

(C4) $T_2, T_4, \dots, T_{k-1}$ are blue paths. Exactly one of these paths has length 1 and the rest have length 2.

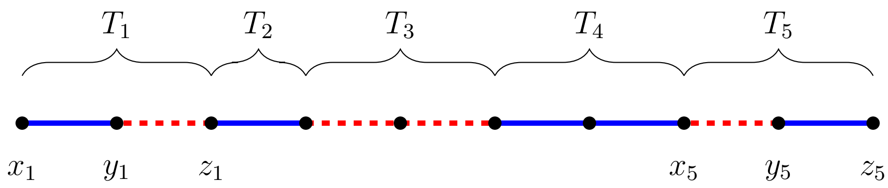

Figure 5.3: A complete path $T_1 \dots T_5$ of type C.

(C5) $T_3, T_5, \dots, T_{k-2}$ *are all red* P_3*'s.*

We say $T_1 \dots T_k$ *is* incomplete *if* T_1 *or* T_k *is a red* P_3*. Otherwise, we say* $T_1 \dots T_k$ *is* complete.

For the remainder of the section, if we refer to a path $vwxyz$ of type B or a path $T_1 \dots T_k$ of type C, we shall take it as read that v, w, x, y, z and $T_1, \dots, T_k$ are as in Definitions 5.6.6 and 5.6.7 respectively. Note that paths of type C are well-defined with respect to direction of traversal – if $v_1 \dots v_p$ is a path of type C, then so is $v_p \dots v_1$. See Figure 5.3 for an example of a path of type C.

We now sketch the proof of Lemma 5.6.18. Let e and f be two independent red edges. Using these edges, Builder can force either a path of type B or a path of type C using Lemma 5.6.8. If Builder obtains a path $vwxyz$ of type B, they will apply Lemma 5.6.9 to efficiently extend Q using $vwxyz$.

Suppose instead Builder obtains a path $T_1 \dots T_k$ of type C. Then we run into a problem – $T_1 \dots T_k$ is not complete, and only a complete path of type C may be used to efficiently extend R (see Lemma 5.6.13). Builder will therefore use Corollary 5.6.12 to extend $T_1 \dots T_k$ into a path $T'_1 \dots T'_{k'}$ of type C which is either complete or arbitrarily long. Builder then uses Lemma 5.6.13 to extend R using $T'_1 \dots T'_{k'}$. If $T'_1 \dots T'_{k'}$ is complete, this extension is efficient; otherwise, Builder wins the game immediately afterwards by joining Q and the resulting blue path. Thus an incomplete path of type C is used to extend R at most once over the course of the game, adding only constantly many rounds to the game's length.

We first prove that given two independent red edges Builder can force either a path of type B or a path of type C.

Lemma 5.6.8 *Given two independent red edges e and f, Builder can force Painter to construct one of the following:*

(i) a path of type B while uncovering 2 *edges;*

(ii) an incomplete path $T_1T_2T_3$ of type C and length 5 *while uncovering* 3 *edges;*

(iii) a red P_4 while uncovering 2 *edges.*

Proof. Write $e = uv$ and $f = xy$. Builder chooses the edges vw and wx, where w is a new vertex. If Painter colours both edges red, then $uvwx$ is a red P_4 and we have achieved (iii). Suppose without loss of generality that Painter colours vw blue. If Painter also colours wx blue, then $uvwxy$ is a path of type B and we have achieved (i). If instead Painter colours wx red, then Builder chooses the edge tu. However Painter colours tu, $tuvwxy$ is now a path of type C and length 5, taking $T_1 = tuv$, $T_2 = vw$ and $T_3 = wxy$. Moreover, T_3 is a red P_3, so $T_1T_2T_3$ is incomplete and we have achieved (ii). $\square$

We next prove that Builder can use a path of type B to efficiently extend Q.

Lemma 5.6.9 *Suppose Q is a non-trivial blue path with endpoints a and b, where b is incident to a red edge bc. Suppose $vwxyz$ is a path of type B vertex-disjoint from $V(Q) \cup \{c\}$. Then, by uncovering at most* 3 *edges, Builder can force Painter to construct one of the following:*

(i) a blue path Q' of length $e(Q) + 5$ with one endpoint b' incident to a red edge $b'c'$.

(ii) a red P_4.

Proof. Builder chooses the edges bv, vy and wz. If Painter colours any of these edges red, then $cbvw$, $wvyz$ or $vwzy$ respectively is a red P_4 and we have achieved (ii).

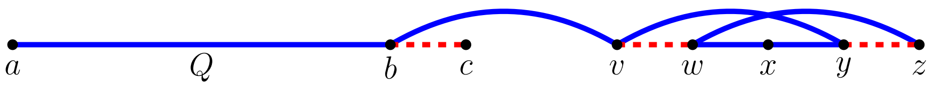

Figure 5.4: Extending Q using a path of type B as in Lemma 5.6.9.

Otherwise, $aQbvyxwz$ is a blue path of length $e(Q) + 5$, where z is incident to the red edge zy (see Figure 5.4), and we have achieved (i). □

We now focus on paths of type C. We first note the following simple property of such paths, which follows immediately from their definition (Definition 5.6.7).

Proposition 5.6.10 *Suppose $T_1 \dots T_k$ is a path of type C. Then*

$$e(T_1 \dots T_k) = 2k - 5 + e(T_1) + e(T_k).$$

Let $T_1 \dots T_k$ be an incomplete path of type C. We first prove an ancillary lemma, which says that Builder can always extend an incomplete path of type C into a slightly longer path of type C.

Lemma 5.6.11 *Suppose $T_1 \dots T_k$ is an incomplete path of type C and length ℓ. Then Builder can force Painter to do one of the following:*

(i) for some $i \in \{3, 4\}$, extend $T_1 \dots T_k$ to a path $T'_1 \dots T'_{k+2}$ of type C and length $\ell + i$ while uncovering i edges.

(ii) construct a red P_4 while uncovering at most 4 edges.

Proof. Suppose without loss of generality that $T_k = x_k y_k z_k$ is a red P_3, where $x_k \in V(T_{k-1})$. Set $T'_i = T_i$ for $i \leq k$. Then Builder chooses two edges uv and vw, where u, v and w are new vertices.

First suppose Painter colours both edges blue. Then Builder chooses the edge $z_k u$. If Painter colours $z_k u$ red, then $x_k y_k z_k u$ is a red P_4 and we have achieved (ii). If Painter colours $z_k u$ blue, then set $T'_{k+1} = z_k uv$ and $T'_{k+2} = vw$. Thus, $T'_1 \ldots T'_{k+2}$ is a path of type C and length $\ell + 3$, and we have achieved (i).

Now suppose that Painter colours both uv and vw red. Then Builder chooses the edges $z_k t$ and tu, where t is a new vertex. If Painter colours one of these edges red, then $x_k y_k z_k t$ or $tuvw$ is a red P_4, respectively, and we have achieved (ii). If Painter colours both $z_k t$ and tu blue, then set $T'_{k+1} = z_k tu$ and $T'_{k+2} = uvw$. Thus, $T'_1 \ldots T'_{k+2}$ is a path of type C and length $\ell + 4$, and we have achieved (i).

Finally, suppose without loss of generality that Painter colours uv blue and vw red. Then Builder chooses the edges $z_k u$ and wx, where x is a new vertex. If Painter colours $z_k u$ red, then $x_k y_k z_k u$ is a red P_4 and we have achieved (ii). If Painter colours $z_k u$ blue, then set $T'_{k+1} = z_k uv$ and $T'_{k+2} = vwx$. Thus $T'_1 \ldots T'_{k+2}$ is a path of type C of length $\ell + 4$, however Painter colours wx, and we have achieved (i). $\square$

By applying Lemma 5.6.11 repeatedly, Builder can extend the path $T_1 T_2 T_3$ of type C given by Lemma 5.6.8 into either a complete path of type C or an arbitrarily long incomplete path of type C. Recall from Proposition 5.6.10 that a path $T_1 \ldots T_k$ of type C has length at most $2k - 1$.

Corollary 5.6.12 *Let $k_0 \geq 5$ be an odd integer. Suppose $T_1 T_2 T_3$ is an incomplete path of type C and length 5. Then Builder can force Painter to do one of the following:*

(i) for some $k, \ell \in \mathbb{N}$, extend $T_1 T_2 T_3$ to a complete path $T'_1 \ldots T'_k$ of type C and length ℓ such that $5 \leq k \leq k_0$, while uncovering $\ell - 5$ edges.

(ii) for some $\ell \in \mathbb{N}$, extend $T_1 T_2 T_3$ to an incomplete path $T'_1 \ldots T'_{k_0}$ of type C and length ℓ while uncovering $\ell - 5$ edges.

(iii) construct a red P_4 while uncovering at most $2k_0 - 6$ edges.

We next prove that Builder can extend R using a path of type C.

Lemma 5.6.13 *Suppose $T_1 \dots T_k$ is a path of type C with $k \geq 5$ and $e(T_1 \dots T_k) = \ell$. Suppose R is a (possibly trivial) blue path which is vertex-disjoint from $T_1 \dots T_k$. Then Builder can force Painter to construct one of the following:*

(i) a blue path R' of length $e(R) + (5k-7)/2$ while uncovering $3(k-1)/2$ edges. This case can only occur if $T_1 \dots T_k$ is incomplete.

(ii) a blue path R' of length $e(R) + \ell'$ while uncovering at most $7\ell'/5 - \ell$ edges for some $1 \leq \ell' \leq 5(k-1)/2$. This case can only occur if $T_1 \dots T_k$ is complete.

(iii) a red P_4 while uncovering at most $3(k-1)/2$ edges.

Proof. Let a and b be the endpoints of R. (If R is trivial, then let $a = b$.) For $i \in \{3, 5, \dots, k-2\}$, write $T_i = x_i y_i z_i$ where $x_i \in V(T_{i-1})$ and $z_i \in V(T_{i+1})$. Thus $x_i y_i z_i$ is a red P_3 for each $i \in \{3, 5, \dots, k-2\}$. Builder chooses the set

$$F_1 = \{x_3 a, b z_3, x_5 c_1, c_1 z_5, x_7 c_2, c_2 z_7, \dots, x_{k-2} c_{\frac{k-5}{2}}, c_{\frac{k-5}{2}} z_{k-2}\}$$

of edges, where $c_1, \dots, c_{\frac{k-5}{2}}$ are new vertices. Note that

$$|F_1| = 2 + 2 \cdot \frac{k-5}{2} = k - 3 < \frac{3(k-1)}{2}. \tag{5.6.14}$$

If Painter colours an edge in F_1 red, say $x_i w$ or $w z_i$ for some integer i and some vertex w, then $z_i y_i x_i w$ or $w z_i y_i x_i$ respectively is a red P_4. So in this case we have achieved (iii).

Now suppose Painter colours all edges in F_1 blue. Then we have obtained a blue path

$$S_1 = T_2 x_3 a R b z_3 T_4 x_5 c_1 z_5 T_6 x_7 c_2 z_7 \dots T_{k-3} x_{k-2} c_{\frac{k-5}{2}} z_{k-2} T_{k-1}.$$

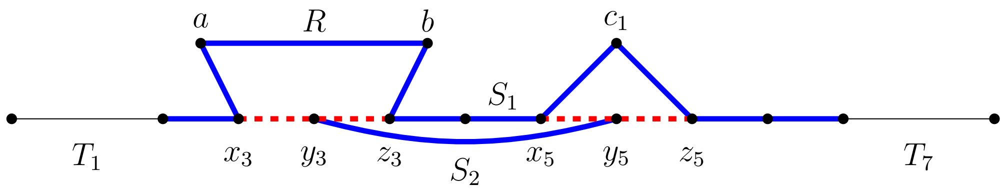

Figure 5.5: Structure of S_1 and S_2 in Lemma 5.6.13 for a path $T_1 \dots T_7$ of type C.

Note that S_1 has length

$$\begin{aligned} e(S_1) &= e(T_2) + e(T_4) + \cdots + e(T_{k-1}) + |F_1| + e(R) \\ &= \left(2 \cdot \frac{k-3}{2} + 1\right) + (k-3) + e(R) = e(R) + 2k - 5, \end{aligned} \tag{5.6.15}$$

where the second equality follows from (5.6.14).

Builder now chooses the set

$$F_2 = \{y_3y_5, y_5y_7, \dots, y_{k-4}y_{k-2}\}$$

of edges. Note that $|F_2| = (k-5)/2$, so by (5.6.14) we have uncovered

$$|F_1| + |F_2| = k - 3 + \frac{k-5}{2} = \frac{3k-11}{2} \tag{5.6.16}$$

edges in total so far. If Painter colours an edge in F_2 red, say y_iy_{i+2} for some $i \in \{3, 5, \dots, k-4\}$, then $z_iy_iy_{i+2}x_{i+2}$ is a red P_4. So in this case we have achieved (iii). Suppose Painter colours all edges in F_2 blue. Then we have obtained a blue path

$$S_2 = y_{k-2}y_{k-4} \dots y_5y_3.$$

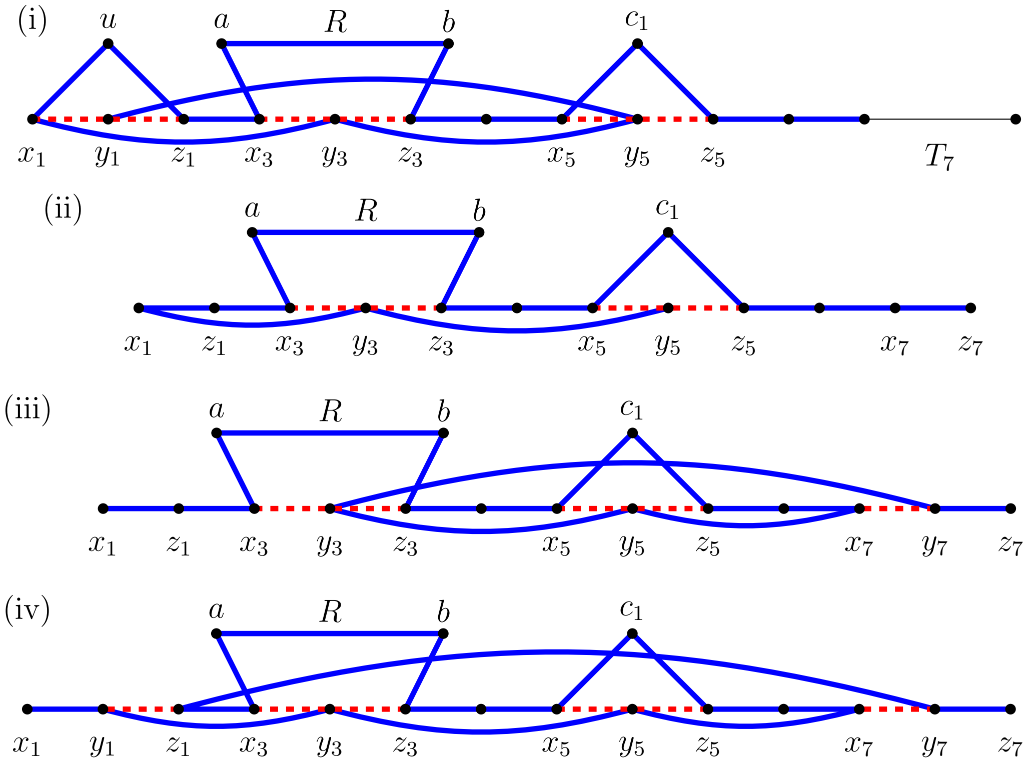

Figure 5.6: Extending a blue path R with a path $T_1 \dots T_7$ as in cases 1 through 4 (respectively) of Lemma 5.6.13.

Note that S_2 has length $|F_2| = (k-5)/2$. Moreover, S_1 and S_2 are vertex-disjoint (see Figure 5.5) and by (5.6.15) we have

$$e(S_1) + e(S_2) = e(R) + 2k - 5 + \frac{k-5}{2} = e(R) + \frac{5(k-3)}{2}. \tag{5.6.17}$$

Our aim is now to join S_1 and S_2 together to form R'. The way in which we do this depends on the structure of T_1 and T_k.

Case 1: $T_1 \dots T_k$ is incomplete.

Without loss of generality we may assume that T_1 is a red P_3, say $x_1y_1z_1$ with $z_1 \in V(T_2)$. Builder chooses the edges y_1y_{k-2}, y_3x_1, x_1u and uz_1, where u is a new vertex. In total, Builder has uncovered $|F_1| + |F_2| + 4 = 3(k-1)/2$ edges by (5.6.16). If Painter colours any of the edges red, then $x_1y_1y_{k-2}z_{k-2}$, $y_3x_1y_1z_1$, $z_1y_1x_1u$ or $uz_1y_1x_1$ is a red P_4, respectively, and we have achieved (iii). Suppose Painter colours them all blue. Then $R' := y_1y_{k-2}S_2y_3x_1uz_1S_1$ is a blue path of length $e(S_1) + e(S_2) + 4 = e(R) + (5k - 7)/2$ by (5.6.17) (see Figure 5.6(i)) and hence we have achieved (i).

Case 2: $T_1 \dots T_k$ is complete and each of T_1 and T_k is a blue edge.

Write $T_1 = x_1z_1$ and $T_k = x_kz_k$ with $z_1 \in V(T_2)$ and $x_k \in V(T_{k-1})$. First suppose that $k \geq 7$. Builder chooses the edges y_3x_1 and $y_{k-2}x_1$. In total, Builder has uncovered $|F_1| + |F_2| + 2 = (3k - 7)/2$ edges by (5.6.16). If Painter colours both edges red, then $x_3y_3x_1y_{k-2}$ is a red P_4 and we have achieved (iii). Suppose Painter colours x_1y_3 blue. Then $R' := S_2y_3x_1z_1S_1x_kz_k$ is a blue path of length $e(S_1) + e(S_2) + 3 = e(R) + (5k - 9)/2$ by (5.6.17) (see Figure 5.6(ii)). Writing

$$\ell' := e(R') - e(R) = \frac{5k - 9}{2},$$

Builder has uncovered

$$\frac{3k - 7}{2} < \frac{7}{5} \cdot \frac{5k - 9}{2} - (2k - 3) = \frac{7\ell'}{5} - \ell$$

edges in total, where the last equality follows from Proposition 5.6.10. Hence we have achieved (ii). If instead Painter colours x_1y_{k-2} blue, the same argument shows we have achieved (ii) on replacing S_2y_3 by S_2y_{k-2}. So if $k \geq 7$, we are done.

If instead $k = 5$, Builder chooses the edges y_3x_1 and ux_1, where u is a new vertex. If Painter colours both edges red, then $ux_1y_3z_3$ is a red P_4 and we have achieved (iii). Suppose instead Painter colours wx_1 blue for some $w \in \{u, y_3\}$. Then $R' := wx_1z_1S_1x_5z_5$ is a blue path of length $e(S_1) + e(S_2) + 3$ (as $e(S_2) = 0$) and Builder has uncovered $|F_1| + |F_2| + 2$ edges. Thus we have achieved (ii) as above.

Case 3: $T_1 \dots T_k$ is complete and exactly one of T_1 and T_k is a blue edge.

Without loss of generality we may assume that T_1 is a blue edge. Let $T_1 = x_1 z_1$ with $z_1 \in V(T_2)$, and let $T_k = x_k y_k z_k$ with $x_k \in V(T_{k-1})$. Note that $x_k y_k$ is red and $y_k z_k$ is blue. Builder chooses the edges $x_k y_{k-2}$ and $y_3 y_k$. In total, Builder has uncovered $|F_1| + |F_2| + 2 = (3k-7)/2$ edges by (5.6.16). If Painter colours either $x_k y_{k-2}$ or $y_3 y_k$ red, then $y_k x_k y_{k-2} x_{k-2}$ or $x_3 y_3 y_k x_k$ is a red P_4 respectively, and we have achieved (iii). Suppose Painter instead colours both edges blue. Then $R' := x_1 z_1 S_1 x_k y_{k-2} S_2 y_3 y_k z_k$ is a blue path of length $e(S_1) + e(S_2) + 4 = e(R) + (5k-7)/2$ by (5.6.17) (see Figure 5.6(iii)). Writing

$$\ell' := e(R') - e(R) = \frac{5k-7}{2},$$

Builder has uncovered

$$\frac{3k-7}{2} < \frac{7}{5} \cdot \frac{5k-7}{2} - (2k-2) = \frac{7\ell'}{5} - \ell$$

edges in total, where the last equality follows from Proposition 5.6.10. Hence we have achieved (ii).

Case 4: $T_1 \dots T_k$ is complete and neither T_1 nor T_k is a blue edge.

Let $T_1 = x_1 y_1 z_1$ and $T_k = x_k y_k z_k$ where $z_1 \in V(T_2)$ and $x_k \in V(T_{k-1})$. Thus $x_1 y_1$ and $y_k z_k$ are blue, and $y_1 z_1$ and $x_k y_k$ are red. Then Builder chooses the edges $y_k z_1$, $x_k y_{k-2}$, and $y_3 y_1$. In total, Builder has uncovered $|F_1| + |F_2| + 3 = (3k-5)/2$ edges by (5.6.16). If Painter colours one of these edges red, then $x_k y_k z_1 y_1$, $y_k x_k y_{k-2} x_{k-2}$ or $z_3 y_3 y_1 z_1$ respectively is a red P_4 and we have achieved (iii). Suppose Painter colours them all blue. Then $R' := z_k y_k z_1 S_1 x_k y_{k-2} S_2 y_3 y_1 x_1$ is a blue path (see Figure 5.6(iv)) of length $e(S_1) + e(S_2) + 5 = e(R) + 5(k-1)/2$ by (5.6.17). Writing

$$\ell' := e(R') - e(R) = \frac{5k-5}{2},$$

Builder has uncovered

$$\frac{3k-5}{2} = \frac{7}{5} \cdot \frac{5k-5}{2} - (2k-1) = \frac{7\ell'}{5} - \ell$$

edges in total, where the last equality follows from Proposition 5.6.10. We have achieved case (ii). $\square$

Finally, we consolidate Lemmas 5.6.8, 5.6.9 and 5.6.13 and Corollary 5.6.12 into a single lemma which says that given two independent red edges, Builder can extend either Q or R. As with Lemma 5.6.5, in applying Lemma 5.6.18 we will take m to be $\ell - e(Q) - e(R) - 1$.

Lemma 5.6.18 *Let $m \geq 9$ be an integer. Let Q and R be blue paths and let e and f be two red edges. Suppose that Q is non-trivial and has an endpoint b incident to a red edge bc. Further suppose that $V(Q) \cup \{c\}$, R, e and f are pairwise vertex-disjoint. Then Builder can force Painter to construct one of the following:*

(i) a blue path Q' with one endpoint b' incident to a red edge $b'c'$ such that $e(Q') = e(Q)+5$, while uncovering 5 edges. Moreover, R is vertex-disjoint from $V(Q') \cup \{c'\}$.

(ii) a blue path R' such that $e(R') = e(R) + \ell'$ for some $1 \leq \ell' \leq m+5$ while uncovering at most $7\ell'/5 - 2$ edges. Moreover, R' is vertex-disjoint from $V(Q) \cup \{c\}$.

(iii) a blue path R' such that $e(R') \geq e(R) + m$ while uncovering at most $7m/5 + 6$ edges. Moreover, R' is vertex-disjoint from $V(Q) \cup \{c\}$.

(iv) a red P_4 while uncovering at most $7m/5 + 6$ edges.

Proof. We first apply Lemma 5.6.8 to e and f. If as in Lemma 5.6.8(iii) we obtain a red P_4 while uncovering 2 edges, then we have achieved (iv). Suppose we do not. Then we split into cases depending on Painter's choice.

Case 1: We obtain a path $vwxyz$ of type B while uncovering 2 edges, as in Lemma 5.6.8(i). Moreover, $vwxyz$ is vertex-disjoint from $V(Q) \cup \{c\}$ and R.

We apply Lemma 5.6.9 to Q and $vwxyz$. Hence we have uncovered at most 5 edges in total. If we obtain a red P_4, then we have achieved (iv). Suppose instead we obtain a blue path Q' of length $q+5$ with one endpoint b' incident to a red edge $b'c'$, where $V(Q') \cup \{c'\}$ is vertex-disjoint from R. Then we have achieved (i).

Case 2: We obtain an incomplete path $T_1T_2T_3$ of type C and length 5 while uncovering 3 edges, as in Lemma 5.6.8(ii). Moreover, $T_1T_2T_3$ is vertex-disjoint from $V(Q) \cup \{c\}$ and R.

Let k_0 be the least odd number such that $k_0 \geq (2m+7)/5$. Since $5k_0 < (2m+7)+5\cdot 2$, and both $5k_0$ and $2m+17$ are odd integers, we have $k_0 \leq 2m/5+3$. Moreover, $k_0 \geq (2m+7)/5 \geq 5$ since $m \geq 9$. We apply Corollary 5.6.12 to $T_1T_2T_3$ and k_0. If we obtain a red P_4 while uncovering at most $2k_0 - 6$ additional edges, then we have achieved (iv). Suppose we do not. Then we split into subcases depending on Painter's choice.

Case 2a: For some $k, \ell \in \mathbb{N}$, we obtain a complete path $T'_1 \dots T'_k$ of type C and length ℓ such that $5 \leq k \leq k_0$ while uncovering $\ell - 5$ additional edges, as in Corollary 5.6.12(i). Moreover, $T'_1 \dots T'_k$ is vertex-disjoint from $V(Q) \cup \{c\}$ and R.

We now apply Lemma 5.6.13 to $T'_1 \dots T'_k$ and R. Suppose we obtain a blue path R' with length $e(R) + \ell'$, where

$$\ell' \leq \frac{5(k-1)}{2} \leq \frac{5(k_0-1)}{2} \leq \frac{5}{2} \cdot \left(\frac{2m}{5} + 2\right) = m+5,$$

while uncovering at most $7\ell'/5 - \ell$ edges as in Lemma 5.6.13(ii). Note that R' is vertex-disjoint from $V(Q) \cup \{c\}$. In total we have uncovered at most $3 + (\ell - 5) + (7\ell'/5 - \ell) = 7\ell'/5 - 2$ edges, so we have achieved (i).

Suppose instead we obtain a red P_4 while uncovering at most $3(k-1)/2$ edges as in Lemma 5.6.13(iii). Note that $\ell \leq 2k_0 - 1$ by Proposition 5.6.10. In total we have therefore uncovered at most

$$3 + (\ell - 5) + \frac{3(k_0 - 1)}{2} \leq \frac{7k_0 - 9}{2} \leq \frac{7}{2} \cdot \left(\frac{2m}{5} + 3 \right) - \frac{9}{2} = \frac{7m}{5} + 6 \qquad (5.6.19)$$

edges, and thus we have achieved (iv).

Case 2b: For some $\ell \in \mathbb{N}$, we obtain an incomplete path $T'_1 \ldots T'_{k_0}$ of type C and length ℓ while uncovering $\ell - 5$ additional edges, as in Corollary 5.6.12(ii). Moreover, $T'_1 \ldots T'_{k_0}$ is vertex-disjoint from $V(Q) \cup \{c\}$ and R.

We apply Lemma 5.6.13 to $T'_1 \ldots T'_{k_0}$ and R. Whatever the outcome, we uncover at most $3(k_0 - 1)/2$ edges. We have therefore uncovered at most $7m/5 + 6$ edges in total, as in (5.6.19). If we obtain a red P_4 as in Lemma 5.6.13(iii), then we have achieved (iv). Hence we may assume that we obtain a blue path R' of length

$$e(R) + \frac{5k_0 - 7}{2} \geq e(R) + \frac{5}{2} \cdot \frac{2m + 7}{5} - \frac{7}{2} = e(R) + m,$$

as in Lemma 5.6.13(i). (The inequality follows from the definition of k_0.) We have therefore achieved (iii). $\square$

5.6.3 An upper bound on $\tilde{r}(P_4, P_{\ell+1})$ for $\ell \geq 3$

We now use Lemmas 5.6.1, 5.6.5 and 5.6.18 to bound $\tilde{r}(P_4, P_{\ell+1})$ above in Theorem 5.6.24. Together with Theorem 5.1.4, this will imply the $k = 3$ case of Theorem 5.1.5.

Recall that Builder's strategy is to extend blue paths Q and R using independent edges. For the remainder of the section, we denote the graph Builder has uncovered by G. In order to keep track of the lengths of Q and R and the number of independent edges available, we introduce the following notation.

Definition 5.6.20 *Given $q, r, n_{\text{blue}}, n_{\text{red}} \in \mathbb{N}_0$, we say that a graph G contains a* $(q, r, n_{\text{blue}}, n_{\text{red}})$-structure *if it satisfies the following properties:*

(P1) G contains a (possibly trivial) blue path Q of length q with one endpoint b incident to a red edge bc.

(P2) G contains a (possibly trivial) blue path R of length r.

(P3) G contains a set F of independent edges containing n_{blue} blue edges and n_{red} red edges.

(P4) $V(Q) \cup \{c\}$, R and F are pairwise vertex-disjoint.

This notation substantially simplifies the statements of Lemmas 5.6.1, 5.6.5 and 5.6.18. The corresponding statements are as follows.

Corollary 5.6.21 *Let $q, r, n_{\text{red}}, n_{\text{blue}} \in \mathbb{N}_0$. Suppose G is a graph containing a $(q, r, n_{\text{blue}}, n_{\text{red}})$-structure. Then Builder can force Painter to construct a graph $G' \supseteq G$ with $e(G') \leq e(G) + 2$ such that G' contains a $(q + r + 1, 0, n_{\text{blue}}, n_{\text{red}})$-structure or a red P_4.*

Corollary 5.6.22 *Let $m, q, r, n_{\text{red}} \in \mathbb{N}_0$ with $q, m \geq 1$. Suppose G is a graph containing a $(q, r, 2, n_{\text{red}})$-structure. Then Builder can force Painter to construct a graph $G' \supseteq G$ such that one of the following holds:*

(i) G' contains a $(q + \ell', r, n_{\text{blue}}, n_{\text{red}})$-structure and $e(G') = e(G) + \ell'$ for some $3 \leq \ell' \leq m + 3$ and some $n_{\text{blue}} \in \mathbb{N}_0$. Moreover, if $3 \leq \ell' < 5 \leq m$, then we may take $n_{\text{blue}} = 1$.

(ii) G' contains a red P_4 and $e(G') \leq e(G) + m + 3$.

Corollary 5.6.23 *Let $m, q, r, n_{\text{blue}} \in \mathbb{N}_0$ with $q \geq 1$ and $m \geq 9$. Suppose G is a graph containing a $(q, r, n_{\text{blue}}, 2)$-structure. Then Builder can force Painter to construct a graph $G' \supseteq G$ such that one of the following holds:*

(i) $e(G') = e(G) + 5$ *and G' contains a $(q + 5, r, n_{\text{blue}}, 0)$-structure.*

(ii) There exists $1 \leq \ell' \leq m + 5$ such that $e(G') \leq e(G) + 7\ell'/5 - 2$ and G' contains a $(q, r + \ell', n_{\text{blue}}, 0)$-structure.

(iii) $e(G') \leq e(G) + 7m/5 + 6$ *and G' contains a $(q, r + m, n_{\text{blue}}, 0)$-structure.*

(iv) $e(G') \leq e(G) + 7m/5 + 6$ *and G' contains a red P_4.*

Theorem 5.6.24 *For all $\ell \in \mathbb{N}$, we have $\tilde{r}(P_4, P_{\ell+1}) \leq (7\ell + 52)/5$.*

Proof. Our aim is to show that Builder can construct a graph G with $e(G) \leq (7\ell+52)/5$ containing a red P_4 or a blue $P_{\ell+1}$.

We first obtain an initial blue path Q with one endpoint incident to a red edge. We claim that either Builder can construct a path $xySz$ of type A with $e(S) < \ell$, while uncovering at most $(7e(S) + 4)/5$ edges, or we are done. We proceed as follows. Builder chooses an edge $e = uv$. First suppose Painter colours uv blue. Then apply Lemma 5.6.3 to uv, taking $m = \ell$. If we find a blue $P_{\ell+1}$ while uncovering $\ell - 1$ additional edges, then since we have uncovered ℓ edges in total we are done. Suppose instead we find a path $xySz$ of type A with $e(S) < \ell$, while uncovering $e(S)$ additional edges in the process. Then in total Builder has uncovered $e(S) + 1 < (7e(S) + 4)/5$ edges, as desired.

Suppose instead Painter colours uv red. Then Builder chooses the edge vx, where x is a new vertex. If Painter colours vx blue, then uvx is a path of type A constructed while uncovering $2 < (7 + 4)/5$ edges in total. If Painter colours vx red, then Builder chooses the edges tu, uw and wx, where t and w are new vertices. If Painter colours any of these edges red, then $tuvx$, $xvuw$ or $wxvu$ respectively is a red P_4 and we are done.

Otherwise, $tuwxv$ is a path of type A (taking $S = tuwx$), constructed while uncovering $5 = (7 \cdot 3 + 4)/5$ edges in total. Therefore, we may assume that Builder has constructed a path $xySz$ of type A with $e(S) < \ell$ while uncovering at most $(7e(S) + 4)/5$ edges as claimed.

Let G_0 be the graph consisting of all edges uncovered so far. Thus G_0 contains a $(q_0, 0, 0, 0)$-structure for some $1 \leq q_0 < \ell$, and $e(G_0) \leq (7q_0 + 4)/5$. Suppose that for some $i \geq 0$, Builder has already constructed a graph G_i such that there exist $q_i, r_i, n_{\text{blue},i}, n_{\text{red},i} \in \mathbb{N}_0$ satisfying the following properties:

(G1) $G_i \subseteq K_{\mathbb{N}}$ is the graph of all uncovered edges.

(G2) G_i contains a $(q_i, r_i, n_{\text{blue},i}, n_{\text{red},i})$-structure, and $q_i > 0$.

(G3) $q_i + r_i \leq \ell + 4$.

(G4) $n_{\text{red},i}, n_{\text{blue},i} \leq 1$.

(G5) $e(G_i) \leq (7(q_i + r_i) + 4)/5 + n_{\text{blue},i} + n_{\text{red},i}$.

Note that (G1)–(G5) hold for $i = 0$. We are going to show that Builder can force a graph $G_{i+1} \supseteq G_i$ such that one of the following holds:

(a) G_{i+1} contains a red P_4 or a blue $P_{\ell+1}$ and $e(G_{i+1}) \leq (7\ell + 52)/5$.

(b) there exist $q_{i+1}, r_{i+1}, n_{\text{blue},i+1}, n_{\text{red},i+1} \in \mathbb{N}_0$ such that $q_{i+1} + r_{i+1} > q_i + r_i$ and G_{i+1}, $q_{i+1}, r_{i+1}, n_{\text{blue},i+1}$ and $n_{\text{red},i+1}$ together satisfy (G1)–(G5).

If (a) holds, we are done. If (b) holds, then Builder can repeat the algorithm to obtain G_{i+2}. We then simply repeat the process until it terminates, which must happen by (G3) (since $q_{i+1} + r_{i+1} > q_i + r_i$ whenever these quantities are defined). It therefore remains only to prove that forcing such a graph is possible.

Let $m = \ell - q_i - r_i - 1$. We split into cases depending on the values of $q_i, r_i, n_{\text{blue},i}$ and $n_{\text{red},i}$.

Case 1: $q_i + r_i \geq \ell - 1$.

In this case, we may simply join our two blue paths together to achieve (a). Apply Corollary 5.6.21 to G_i. Builder obtains a graph $G_{i+1} \supseteq G_i$ with

$$e(G_{i+1}) = e(G_i) + 2 \overset{\text{(G5)}}{\leq} \frac{7(q_i + r_i) + 4}{5} + n_{\text{blue},i} + n_{\text{red},i} + 2 \overset{\text{(G3),(G4)}}{\leq} \frac{7\ell + 52}{5}.$$

Moreover, G' contains a red P_4 or a blue $P_{\ell+1}$, so we have achieved (a).

Case 2: $\ell - 9 \leq q_i + r_i \leq \ell - 2$, so that $1 \leq m \leq 8$.

In this case, it is more efficient to naively extend our paths to the right combined length and join them than it is to apply our normal extension methods and potentially end up with paths longer than we need. Builder will force a red P_4 or a blue $P_{\ell+1}$ as follows. Apply Corollary 5.6.21 to G_i to obtain a graph $G' \supseteq G_i$ with $e(G') = e(G_i) + 2$. Note that G' contains a red P_4 or a $(q_i + r_i + 1, 0, n_{\text{blue},i}, n_{\text{red},i})$-structure. By repeating the process at most m additional times, Builder obtains a graph $G'' \supseteq G' \supseteq G_i$, where

$$\begin{aligned} e(G'') &\leq e(G) + 2m + 2 \overset{\text{(G5)}}{\leq} \frac{7(q_i + r_i) + 4}{5} + n_{\text{blue},i} + n_{\text{red},i} + 2m + 2 \\ &\overset{\text{(G4)}}{\leq} \frac{7(\ell - m - 1) + 4}{5} + 2 + 2m + 2 = \frac{7\ell}{5} + \frac{3m + 17}{5} \leq \frac{7\ell + 41}{5}, \end{aligned}$$

such that G'' contains a red P_4 or a $(q_i + r_i + m + 1, 0, n_{\text{blue},i}, n_{\text{red},i})$-structure (which contains a blue $P_{\ell+1}$). Thus we have achieved (a).

Case 3: $q_i + r_i \leq \ell - 10$, so that $m \geq 9$.

In this case, we will extend our paths efficiently using Corollaries 5.6.22 and 5.6.23. By choosing at most $3 - n_{\text{blue},i} - n_{\text{red},i}$ additional independent edges (on new vertices), Builder obtains a graph $G_i' \supseteq G_i$ containing a $(q_i, r_i, n'_{\text{blue}}, n'_{\text{red}})$-structure such that $n'_{\text{blue}} + n'_{\text{red}} \leq 3$, either $n'_{\text{blue}} = 2$ or $n'_{\text{red}} = 2$, and

$$e(G'_i) \overset{\text{(G5)}}{\leq} \frac{7(q_i + r_i) + 4}{5} + n'_{\text{blue}} + n'_{\text{red}}. \tag{5.6.25}$$

We split into subcases depending on the values of n'_{blue} and n'_{red}.

Case 3a: $n'_{\text{blue}} = 2$ and $n'_{\text{red}} \leq 1$.

We apply Corollary 5.6.22 to G'_i, obtaining a graph $G' \supseteq G'_i$. First suppose Corollary 5.6.22(i) holds, so that there exists some $3 \leq \ell' \leq m + 3$ such that G' contains a $(q_i + \ell', r_i, n''_{\text{blue}}, n'_{\text{red}})$-structure and $e(G') = e(G'_i) + \ell'$. Set $G_{i+1} = G'$, $q_{i+1} = q_i + \ell'$, $r_{i+1} = r_i$ and $n_{\text{red},i+1} = n'_{\text{red}}$. Set $n_{\text{blue},i+1} = 0$ if $\ell' \geq 5$ and $n_{\text{blue},i+1} = 1$ otherwise. Clearly $q_{i+1} + r_{i+1} > q_i + r_i$, and (G1) and (G4) are satisfied. Recall from Corollary 5.6.22(i) that if $\ell' < 5 \leq m$ then we may take $n''_{\text{blue}} = 1$, so (G2) is satisfied. We have $q_{i+1} + r_{i+1} \leq q_i + m + 3 + r_i = \ell + 2$, so (G3) is satisfied. If $3 \leq \ell' \leq 4$, we have

$$\begin{aligned} e(G') = e(G'_i) + \ell' &\overset{(5.6.25)}{\leq} \frac{7(q_i + r_i) + 4}{5} + 2 + n'_{\text{red}} + \ell' \\ &= \frac{7(q_i + r_i + \ell') + 4}{5} - \frac{2\ell'}{5} + 2 + n'_{\text{red}} \\ &\leq \frac{7(q_{i+1} + r_{i+1}) + 4}{5} + 1 + n'_{\text{red}} \\ &= \frac{7(q_{i+1} + r_{i+1}) + 4}{5} + n_{\text{blue},i+1} + n_{\text{red},i+1}. \end{aligned} \tag{5.6.26}$$

If instead $\ell' \geq 5$, then by a calculation similar to the above, we have

$$\begin{aligned} e(G') &\overset{(5.6.25)}{\leq} \frac{7(q_i + r_i) + 4}{5} + 2 + n'_{\text{red}} + \ell' \leq \frac{7(q_{i+1} + r_{i+1}) + 4}{5} + n'_{\text{red}} \\ &= \frac{7(q_{i+1} + r_{i+1}) + 4}{5} + n_{\text{blue},i+1} + n_{\text{red},i+1}. \end{aligned} \tag{5.6.27}$$

Thus, by (5.6.26) and (5.6.27), (G5) is satisfied. We have therefore achieved (b).

Suppose instead that Corollary 5.6.22(ii) holds, so that G' contains a red P_4 and

$e(G') \leq e(G'_i) + m + 3$. Then we have

$$\begin{aligned} e(G') &\overset{(5.6.25)}{\leq} \frac{7(q_i + r_i) + 4}{5} + 2 + n'_{\text{red}} + m + 3 \\ &\leq \frac{2(q_i + r_i) + 4}{5} + \ell + 5 \leq \frac{7\ell + 9}{5}, \end{aligned}$$

where the final inequality follows since $q_i + r_i \leq \ell - 10$. We have therefore achieved (a).

Case 3b: $n'_{\text{red}} = 2$ and $n'_{\text{blue}} \leq 1$.

We apply Corollary 5.6.23 to G'_i, obtaining a graph $G' \supseteq G'_i$. Suppose Corollary 5.6.23(i) or (ii) holds. In either case, it follows that there exist q' and r' such that G' contains a $(q', r', n'_{\text{blue}}, 0)$-structure and

$$1 \leq q' + r' - (q_i + r_i) \leq m + 5.$$

Write $\ell' = q' + r' - (q_i + r_i)$. Set $G_{i+1} = G'$, $q_{i+1} = q'$, $r_{i+1} = r'$, $n_{\text{blue},i+1} = n'_{\text{blue}}$ and $n_{\text{red},i+1} = 0$. Clearly (G1)–(G4) are satisfied, and $q_{i+1} + r_{i+1} > q_i + r_i$. Moreover, we have

$$\begin{aligned} e(G_{i+1}) \leq e(G'_i) + \frac{7\ell'}{5} - 2 &\overset{(5.6.25)}{\leq} \frac{7(q_i + r_i + \ell') + 4}{5} + n'_{\text{blue}} \\ &= \frac{7(q_{i+1} + r_{i+1}) + 4}{5} + n_{\text{blue},i+1} + n_{\text{red},i+1}, \end{aligned}$$

so (G5) is satisfied. We have therefore achieved (b).

Now suppose Corollary 5.6.23(iii) holds, so that G' contains a $(q_i, r_i + m, n'_{\text{blue}}, 0)$-structure and $e(G') \leq e(G'_i) + 7m/5 + 6$. We apply Corollary 5.6.21 to G', obtaining a graph G'' such that

$$\begin{aligned} e(G'') &= e(G') + 2 \leq e(G'_i) + \frac{7m}{5} + 8 \\ &\overset{(5.6.25)}{\leq} \frac{7(q_i + r_i + m) + 4}{5} + n'_{\text{blue}} + 10 \leq \frac{7\ell + 52}{5}. \end{aligned}$$

Moreover, G'' contains a red P_4 or an $(\ell, 0, n'_{\text{blue}}, 0)$-structure (which contains a blue $P_{\ell+1}$). We have therefore achieved (a).

Finally suppose Corollary 5.6.23(iv) holds, so that G' contains a red P_4 and $e(G') \leq e(G'_i) + 7m/5 + 6$. Then we have

$$e(G') \leq e(G'_i) + \frac{7m}{5} + 6 \overset{(5.6.25)}{\leq} \frac{7(q_i + r_i + m) + 4}{5} + n'_{\text{blue}} + 8 \leq \frac{7\ell + 42}{5}.$$

We have therefore achieved (a). This completes the proof of the theorem. □

5.7 Bounding $\tilde{r}(C_4, P_{\ell+1})$ for $\ell \geq 3$

Our aim is to prove Theorem 5.1.9, i.e. to bound $\tilde{r}(C_4, P_{\ell+1})$ for all $\ell \geq 3$. The lower bound is proved by considering a $\{C_4\}$-blocking strategy for Painter.

Proposition 5.7.1 *Let $k \in \mathbb{N}$ with $k \geq 3$. Let H be a connected graph. Then $\tilde{r}(C_k, H) \geq |H| + e(H) - 1$.*

Proof. We consider the $\{C_k\}$-blocking strategy for Painter in the $\tilde{r}(C_k, H)$-game. Let R be a $\{C_k\}$-scaffolding for H with $e(R)$ minimal. Note that each $(R, \{C_k\})$-forceable edge must lie entirely in a component of R. Since H is connected, R is connected and $|R| \geq |H|$. Hence, $e(R) \geq |H| - 1$. We are done by Proposition 5.3.3. □

Next, we prove that $\tilde{r}(C_4, P_4) = 8$. Note that a more detailed analysis of the $\{C_4\}$-blocking strategy for Painter is needed in order to obtain a better lower bound.

Proposition 5.7.2 $\tilde{r}(C_4, P_4) = 8$.

Proof. First, we consider the $\{C_4\}$-blocking strategy for Painter in the $\tilde{r}(C_4, P_4)$-game. Let R be an edge-minimal $\{C_4\}$-scaffolding for P_4. Then R must contain three distinct P_4's, so $e(R) \geq 5$ as R is C_4-free. Proposition 5.3.3 implies that $\tilde{r}(C_4, P_4) \geq 8$.

It therefore suffices to prove that Builder can win the $\tilde{r}(C_4, P_4)$-game within 8 rounds. Builder first chooses the edges $uv_1, \ldots, uv_4$ for distinct vertices $u, v_1, \ldots, v_4$. Without loss of generality we may assume that there exists an integer j such that Painter colours the edges uv_i blue if $i \leq j$, and red otherwise.

Suppose $j \geq 2$. Then Builder chooses four edges v_1w, v_2w, v_1w' and v_2w', where w and w' are new vertices. If Painter colours all edges red, then $v_1wv_2w'v_1$ is a red C_4. If Painter colours one of the edges blue say v_2w, then v_1uv_2w is a blue P_4.

Suppose $j \leq 1$. Then Builder chooses edges v_1v_2 and v_1v_3. If Painter colours both edges red, then $uv_2v_1v_3u$ is a red C_4. Suppose that Painter colours both edges blue. Builder then chooses the edges v_2v_4 and v_3v_4. If Painter colours both v_2v_4 and v_3v_4 red, then $uv_2v_4v_3u$ is a red C_4. Otherwise, $v_3v_1v_2v_4$ or $v_2v_1v_3v_4$ is a blue P_4. Therefore we may assume that v_1v_2 is blue and v_1v_3 is red. Further suppose that $j = 1$ and so uv_1 is blue. Then Builder chooses the edges v_2v_3 and v_2v_4. If Painter colours one of them blue, then $uv_1v_2v_3$ or $uv_1v_2v_4$ is a blue P_4. Otherwise $uv_3v_2v_4u$ is a red C_4. Finally, suppose that $j = 0$. Builder chooses the edges v_2v_3 and v_3v_4. If Painter colours one of them red, then $uv_1v_3v_2u$ or $uv_1v_3v_4u$ is a red C_4. Otherwise $v_1v_2v_3v_4$ is a blue P_4. $\square$

We now prove Theorem 5.1.9.

Proof. [Proof of Theorem 5.1.9] The lower bound follows from Proposition 5.7.1 and $\tilde{r}(C_4, P_4) = 8$ by Proposition 5.7.2. To prove the theorem, it is enough to show that $\tilde{r}(C_4, P_{\ell+1}) \leq 4\ell - 4$ for all $\ell \geq 3$. We proceed by induction on ℓ. By Proposition 5.7.2, this is true for $\ell = 3$. Suppose instead that $\ell \geq 4$ and Builder first spends at most $4\ell - 8$ rounds forcing Painter to construct a red C_4 or a blue $P_\ell = v_1 \ldots v_\ell$. (This is possible by the induction hypothesis.) We may assume that the latter holds or else we are done. Then Builder chooses four edges v_1x, $v_\ell x$, v_1y and $v_\ell y$, where x and y are new vertices. If Painter colours all edges red, then $v_1xv_\ell yv_1$ is a red C_4. If Painter colours one of the edges blue, say $v_\ell x$, then $v_1 \ldots v_\ell x$ is a blue $P_{\ell+1}$. In total, Builder has chosen at most

$4\ell - 4$ edges and the proposition follows. $\square$

List of References

[1] M. Ajtai, J. Komlós, and E. Szemerédi. An $O(n \log n)$ sorting network. *Proc. 15th Ann. ACM Symp. on Theory of Computing*, pages 1–19, 1983.

[2] M. Ajtai, J. Komlós, and E. Szemerédi. Sorting in $c \log n$ parallel steps. *Combinatorica*, 3:1–19, 1983.

[3] M. Ajtai, J. Komlós, and E. Szemerédi. The first occurrence of Hamilton cycles in random graphs. *Ann. Discrete Math.*, 27:173–178, 1985.

[4] T. Akiyama, T. Nishizeki, and N. Saito. NP-completeness of the Hamiltonian cycle problem for bipartite graphs. *J. Information Processing*, 3:73–76, 1980.

[5] J. Bang-Jensen. On the 2-linkage problem for semicomplete digraphs. *Annals of Discrete Math.*, 41:23–38, 1989.

[6] J. Bang-Jensen. Problems and conjectures concerning connectivity, paths, trees and cycles in tournament-like digraphs. *Discrete Math.*, 309:5655–5667, 2009.

[7] J. Bang-Jensen and G. Gutin. *Digraphs: Theory, Algorithms and Applications.* Springer, New York, 2nd edition, 2008.

[8] J. Bang-Jensen and C. Thomassen. A polynomial algorithm for the 2-path problem for semicomplete digraphs. *SIAM J. Discrete Mathematics*, 5:366–376, 1992.

[9] K. E. Batcher. Sorting networks and their applications. *Proc. AFIPS Spring Joint Computer Conference*, 32:307–314, 1968.

[10] J. Beck. On size Ramsey number of paths, trees and cycles I. *J. Graph Theory*, 7:115–130, 1983.

[11] J. Beck. There is no fast method for finding monochromatic complete subgraphs. *J. Combin. Theory B*, 34:58–64, 1983.

[12] J. Beck. Achievement games and the probabilistic method. In *Combinatorics, Paul Erdős is Eighty*, volume 1, pages 51–78. Bolyai Soc. Math. Studies, 1993.

[13] S. Ben-Shimon, M. Krivelevich, and B. Sudakov. On the resilience of Hamiltonicity and optimal packing of Hamilton cycles in random graphs. *SIAM J. Discrete Mathematics*, 25:1176–1193, 2011.

[14] B. Bollobás. The evolution of sparse graphs. In *Graph Theory and Combinatorics*, pages 35–37. Academic Press, London, 1984.

[15] B. Bollobás. *Random Graphs*. Academic Press, London, 1985.

[16] B. Bollobás and A. Frieze. On matchings and Hamiltonian cycles in random graphs. *North-Holland Math. Stud.*, 118:23–46, 1985.

[17] B. Bollobás and A. Thomason. Highly linked graphs. *Combinatorica*, 16:313–320, 1996.

[18] J. A. Bondy and U. S. R. Murty. *Graph Theory*, volume 244 of *Graduate Texts in Mathematics*. Springer, New York, 2008.

[19] S. A. Burr. Determining generalised Ramsey numbers is NP-hard. *Ars. Combin.*, 17:21–25, 1984.

[20] P. Camion. Chemins et circuits hamiltoniens des graphes complets. *C. R. Acad. Sci. Paris*, 249:2151–2152, 1959.

[21] D. Christofides, D. Kühn, and D. Osthus. Edge-disjoint Hamilton cycles in graphs. *J. Combin. Theory B*, 102:1035–1060, 2012.

[22] M. Chudnovsky, A. Scott, and P. Seymour. Disjoint paths in tournaments. Preprint.

[23] D. Conlon. A new upper bound for diagonal Ramsey numbers. *Ann. of Math.*, 170:941–960, 2009.

[24] D. Conlon. On-line Ramsey numbers. *SIAM J. Discrete Mathematics*, 23:1954–1963, 2009.

[25] B. Csaba, D. Kühn, A. Lo, D. Osthus, and A. Treglown. Proof of the 1-factorization and Hamilton decomposition conjectures II: the bipartite case. Preprint.

[26] B. Csaba, D. Kühn, A. Lo, D. Osthus, and A. Treglown. Proof of the 1-factorization and Hamilton decomposition conjectures III: approximate decompositions. Preprint.

[27] J. Cyman, T. Dzido, J. Lapinskas, and A. Lo. On-line Ramsey numbers of paths and cycles. Preprint.

[28] G. A. Dirac. Some theorems on abstract graphs. *Proc. Lond. Math. Soc.*, 2:69–81, 1952.

[29] P. Erdős. On some new inequalities concerning extremal properties of graphs. *Theory of Graphs*, Proc. Coll. Tihany:77–81, 1966.

[30] P. Erdős. Some recent results on extremal problems in graph theory. *Theory of Graphs*, Intern. Symp. Rome:118–123, 1966.

[31] P. Erdős, R. J. Faudree, C. C. Rousseau, and R. H. Schelp. The size Ramsey number. *Period. Math. Hungar.*, 9:145–161, 1978.

[32] P. Erdős and A. Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl*, 5:17–61, 1960.

[33] A. Ferber, M. Krivelevich, and B. Sudakov. Counting and packing Hamilton cycles in dense graphs and oriented graphs. Preprint.

[34] S. Fortune, J. E. Hopcroft, and J. Wyllie. The directed subgraph homeomorphism problem. *J. Theoret. Comput. Sci.*, 10:111–121, 1980.

[35] E. Friedgut, Y. Kohayakawa, V. Rödl, A. Ruciński, and P. Tetali. Ramsey games against a one-armed bandit. *Combin. Probab. Comput.*, 12:513–545, 2003.

[36] A. Frieze and M. Krivelevich. On packing Hamilton cycles in ε-regular graphs. *J. Combin. Theory B*, 94:159–172, 2005.

[37] A. Frieze and M. Krivelevich. On two Hamilton cycle problems in random graphs. *Israel J. Math.*, 166:221–234, 2008.

[38] E. N. Gilbert. Random graphs. *Ann. Math. Stat.*, 30:1141–1144, 1959.

[39] R. Glebov, M. Krivelevich, and T. Szabó. On covering expander graphs by Hamilton cycles. *Random Structures & Algorithms*, 44:183–200, 2014.

[40] J. Grytczuk, M. Hałuszczak, and H. Kierstead. On-line Ramsey theory. *Electron. J. Combin.*, 11:R57, 2004.

[41] J. Grytczuk, H. Kierstead, and P. Prałat. On-line Ramsey number for paths and stars. *Discrete Math. Theor. Comput. Sci.*, 10:63–74, 2008.

[42] S. G. Hartke, R. Martin, and T. Seacrest. Relating minimum degree and the existence of a k-factor. Research manuscript.

[43] S. G. Hartke and T. Seacrest. Random partitions and edge-disjoint Hamilton cycles. *J. Combin. Theory B*, 103.6:742–766, 2013.

[44] D. Hefetz, M. Krivelevich, and T. Szabó. Hamilton cycles in highly connected and expanding graphs. *Combinatorica*, 29:547–568, 2009.

[45] D. Hefetz, D. Kühn, J. Lapinskas, and D. Osthus. Optimal covers with Hamilton cycles in random graphs. *Combinatorica*, to appear.

[46] S. Janson, T. Łuczak, and A. Ruciński. *Random Graphs.* Wiley, New York, 2000.

[47] R. M. Karp. Reducibility among combinatorial problems. In *Complexity of Computer Computations*, pages 85–103. Springer, New York, 1972.

[48] P. Katerinis. Minimum degree of a graph and the existence of k-factors. *Proc. Indian Acad. Sci. Math. Sci.*, 94:123–127, 1985.

[49] H. Kierstead and G. Konjevod. Coloring number and on-line Ramsey theory for graphs and hypergraphs. *Combinatorica*, 29:49–64, 2009.

[50] F. Knox, D. Kühn, and D. Osthus. Approximate Hamilton decompositions of random graphs. *Random Structures & Algorithms*, 40:133–149, 2012.

[51] F. Knox, D. Kühn, and D. Osthus. Edge-disjoint Hamilton cycles in random graphs. *Random Structures & Algorithms*, to appear.

[52] D. E. Knuth. *The Art of Computer Programming*, volume 3. Addison-Wesley, Boston, 1997.

[53] M. Krivelevich. Triangle factors in random graphs. *Combin. Probab. Comput.*, 6:337–347, 1997.

[54] M. Krivelevich and W. Samotij. Optimal packings of Hamilton cycles in sparse random graphs. *SIAM J. Discrete Mathematics*, 26:964–982, 2012.

[55] D. Kühn, J. Lapinskas, and D. Osthus. Optimal packings of Hamilton cycles in graphs of high minimum degree. *Combin. Probab. Comput.*, 22:394–416, 2013.

[56] D. Kühn, J. Lapinskas, D. Osthus, and V. Patel. Proof of a conjecture of Thomassen on Hamilton cycles in highly connected tournaments. *Proc. Lond. Math. Soc.*, to appear.

[57] D. Kühn, A. Lo, and D. Osthus. Proof of the 1-factorization and Hamilton decomposition conjectures IV: exceptional systems for the two cliques case. Preprint.

[58] D. Kühn, A. Lo, D. Osthus, and A. Treglown. Proof of the 1-factorization and Hamilton decomposition conjectures I: the two cliques case. Preprint.

[59] D. Kühn and D. Osthus. Hamilton decompositions of regular expanders: applications. Preprint.

[60] D. Kühn and D. Osthus. A survey on Hamilton cycles in directed graphs. *European J. Combin.*, 33:750–766, 2012.

[61] D. Kühn and D. Osthus. Hamilton decompositions of regular expanders: a proof of Kelly's conjecture for large tournaments. *Adv. Math.*, 237:62–146, 2013.

[62] D. Kühn, D. Osthus, and T. Townsend. Proof of a tournament partition conjecture and an application to 1-factors with prescribed cycle lengths. *Combinatorica*, to appear.

[63] D. Kühn, D. Osthus, and A. Treglown. Hamiltonian degree sequences in digraphs. *J. Combin. Theory B*, 100:367–380, 2010.

[64] A. Kurek and A. Ruciński. Two variants of the size Ramsey number. *Discuss. Math. Graph Theory*, 25:141–149, 2005.

[65] W. Mantel. Vraagstuk XXVIII. *Wiskundige Opgaven*, 10:60–61, 1907.

[66] M. Marciniszyn, R. Spöhel, and A. Steger. Online Ramsey games in random graphs. *Combin. Probab. Comput.*, 18:271–300, 2009.

[67] M. Marciniszyn, R. Spöhel, and A. Steger. Upper bounds for online Ramsey games in random graphs. *Combin. Probab. Comput.*, 18:259–270, 2009.

[68] R. Matai, S. P. Singh, and M. L. Mittal. Traveling salesman problem: an overview of problem, theory and applications. In *Traveling salesman problem, theory and applications*, pages 1–24. InTech, Rijeka, 2010.

[69] K. Menger. Zur allgemeinen kurventheorie. *Fund. Math.*, 10:96–115, 1927.

[70] J. W. Moon. On subtournaments of a tournament. *Canad. Math. Bull.*, 9:297–301, 1966.

[71] C. St. J. A. Nash-Williams. Hamiltonian lines in graphs whose vertices have sufficiently large valencies. In *Combinatorial theory and its applications III (proc. Colloq. Balatonfüred)*, pages 813–819. North-Holland, Amsterdam, 1970.

[72] C. St. J. A. Nash-Williams. Edge-disjoint Hamiltonian circuits in graphs with vertices of large valency. In *Studies in pure mathematics (presented to Richard Rado)*, pages 157–183. Academic Press, London, 1971.

[73] C. St. J. A. Nash-Williams. Hamiltonian arcs and circuits. In *Recent Trends in Graph Theory (Proc. Conf. New York, 1970)*, pages 197–210. Springer, Berlin, 1971.

[74] M. S. Paterson. Improved sorting networks with $O(\log n)$ depth. *Algorithmica*, 5:75–92, 1990.

[75] A. Pokrovskiy. Edge disjoint hamiltonian cycles in highly connected tournaments. Preprint.

[76] A. Pokrovskiy. Highly linked tournaments. Preprint.

[77] P. Prałat. A note on small on-line Ramsey numbers for paths and their generalisations. *Australas. J. Combin.*, 40:27–36, 2008.

[78] P. Prałat. A note on off-diagonal small on-line Ramsey numbers for paths. *Ars. Combin.*, 107:295–306, 2012.

[79] R. Rado. Studien zur kombinatorik. *Math. Zeit.*, 36:242–280, 1933.

[80] S. Radziszowski. Small Ramsey numbers. *Electron. J. Combin.*, page DS1, 2014.

[81] F. Ramsey. On a problem of formal logic. *Proc. Lond. Math. Soc.*, 30:264–286, 1930.

[82] K. B. Reid. Three problems on tournaments. In *Graph Theory and its Applications: East and West*, volume 576, pages 466–473. Ann. New York Acad. Sci., 1989.

[83] V. Rödl. On a packing and covering problem. *European J. Combin.*, 6:69–78, 1985.

[84] V. Rödl, A. Ruciński, and E. Szemerédi. A Dirac-type theorem for 3-uniform hypergraphs. *Combin. Probab. Comput.*, 15:229–251, 2006.

[85] M. Simonovits. A method for solving extremal problems in graph theory. *Theory of Graphs*, Proc. Coll. Tihany:279–319, 1966.

[86] J. Spencer. Ramsey's theorem—a new lower bound. *J. Combin. Theory A*, 18:108–115, 1975.

[87] R. Thomas and P. Wollan. An improved extremal function for graph linkages. *European J. Combin.*, 26:309–324, 2005.

[88] A. Thomason. Pseudo-random graphs. *Ann. Discrete Math.*, 33:307–331, 1987.

[89] C. Thomassen. Edge-disjoint Hamiltonian paths and cycles in tournaments. *Proc. Lond. Math. Soc.*, 45:151–168, 1982.

[90] C. Thomassen. Connectivity in tournaments. In *Graph Theory and Combinatorics, A volume in honour of Paul Erdős*, pages 305–313. Academic Press, London, 1984.

[91] C. Thomassen. Note on highly connected non-2-linked digraphs. *Combinatorica*, 11:393–395, 1991.

[92] W. T. Tutte. The factors of graphs. *Canadian J. Math.*, 4:314–328, 1952.

[93] W. T. Tutte. A short proof of the factor theorem for finite graphs. *Canad. J. Math.*, 4:314–328, 1952.

[94] B. L. van der Waerden. Beweis einer baudetschen vermutung. *Nieuw Arch. Wisk*, 15:212–216, 1927.