Embedding problems in graphs and hypergraphs

by

Andrew Clark Treglown

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

School of Mathematics
The University of Birmingham
September 2010

Abstract

The first part of this thesis concerns perfect matchings and their generalisations. We determine the minimum vertex degree that ensures a perfect matching in a 3-uniform hypergraph, thereby answering a question of Hàn, Person and Schacht.

We say that a graph G has a perfect H-packing (also called an H-factor) if there exists a set of disjoint copies of H in G which together cover all the vertices of G. Given a graph H, we determine, asymptotically, the Ore-type degree condition which ensures that a graph G has a perfect H-packing.

The second part of the thesis concerns Hamilton cycles in directed graphs. We give a condition on the degree sequences of a digraph G that ensures G is Hamiltonian. This gives an approximate solution to a problem of Nash-Williams concerning a digraph analogue of Chvátal's theorem.

We also show that every sufficiently large regular tournament can almost completely be decomposed into edge-disjoint Hamilton cycles. More precisely, for each $\eta > 0$ every regular tournament G of sufficiently large order n contains at least $(1/2 - \eta)n$ edge-disjoint Hamilton cycles. This gives an approximate solution to a conjecture of Kelly from 1968.

Acknowledgements

I would like to thank Daniela Kühn and Deryk Osthus for the guidance and help I have received from them in the writing of this thesis. You have both been excellent PhD supervisors and collaborators.

Thanks must also go to the other combinatorialists who have come and gone whilst I have been at Birmingham; Olly Cooley, Luke Kelly, Fiachra Knox and Richard Mycroft. As well as having many useful conversations about maths we've sung campfire songs, spent freezing winter evenings at football matches and even cooked on national television together!

Many people have helped shape my academic life, from the teachers at Wyvern College, Salisbury to those who lecturered me at the University of Birmingham. Particular thanks should go to Brian Philp who, whilst I was an undergraduate, always had time to answer my questions.

Finally the biggest thanks goes to my family, especially my mum, dad and sister for their continued support.

Contents

Chapter 1
Introduction

A natural question is to establish conditions that ensure a graph G contains some spanning subgraph F. For example F could be a Hamilton cycle or a perfect matching. Of course, it is desirable to fully characterise those graphs G which contain a spanning copy of a given graph F. For example, Tutte's theorem [93] characterises those graphs with a perfect matching. However, for some graphs F (for example Hamilton cycles) it is unlikely that such a characterisation exists. Indeed, for many graphs F (including Hamilton cycles) the decision problem of whether a graph G contains F is NP-complete. Thus, it is of interest to find sufficient conditions.

1.1 Generalisations of perfect matchings

1.1.1 Perfect H-packings

Perhaps the simplest parameter of a graph G to consider is the minimum degree $\delta(G)$ of G. Dirac [23] showed that any graph G on $n \geq 3$ vertices has a Hamilton cycle provided that $\delta(G) \geq n/2$. So when n is even this implies that G contains a perfect matching (and it is easy to see that this bound is tight). Chapter 3 of this thesis is concerned with the case when F is composed of many copies of a small graph H, i.e. when F is a perfect H-packing. More precisely, a perfect H-packing in G consists of vertex-disjoint copies of H in G covering all the vertices of G. So if $H = K_2$, a perfect H-packing in G is simply a perfect matching in G. Hajnal and Szemerédi [33] established the bound on the minimum degree of a graph G which guarantees that G contains a perfect K_r-packing. Given *any* graph H, Kühn and Osthus [57] determined, up to an additive constant, the bound on the minimum degree of a graph G that ensures a perfect H-packing in G.

It is also of interest to consider other types of degree conditions that force a perfect H-packing in a graph G. Ore's theorem [73] generalises Dirac's theorem. This result states that a graph G of order $n \geq 3$ contains a Hamilton cycle if $d(x) + d(y) \geq n$ for all non-adjacent $x \neq y \in V(G)$. We refer to such conditions on the sum of the degrees of non-adjacent vertices of a graph as Ore-type degree conditions. A result of Kierstead and Kostochka [45] implies a 'best possible' Ore-type degree condition which guarantees the existence of a perfect K_r-packing in a graph G. In Chapter 3 we asymptotically determine the Ore-type degree condition that ensures a perfect H-packing in a graph G for *any* graph H (see Theorem 3.2). Thus, this provides an Ore-type analogue of the result of Kühn and Osthus mentioned above.

Notice that the Ore-type degree condition which forces a Hamilton cycle in a graph is 'twice the minimum degree condition' in Dirac's theorem. Further the corresponding bound in the aforementioned result of Kierstead and Kostochka is again 'twice the minimum degree condition' in the Hajnal-Szemerédi theorem. Thus, one may imagine that we have a similar phenomenon for our result concerning perfect H-packings. However, perhaps surprisingly, this is not the case. Indeed, for some graphs H the Ore-type degree condition which ensures a perfect H-packing in a graph G involves the so-called colour extension number of H. This parameter is not relevant, however, in the corresponding minimum degree condition.

The Erdős-Stone theorem gives a condition on the number of edges in a graph G which forces a copy of some fixed graph H in G. Clearly a necessary condition for the existence of a perfect H-packing in a graph G is the property that for all $x \in V(G)$ there exists a copy of H in G containing x. In Section 3.1.4 we characterise, up to an error term, the minimum and Ore-type degree conditions that ensure a copy of a graph H in G containing a given $x \in V(G)$. In some sense the bound in this latter result is the 'reason' why the Ore-type degree condition which guarantees a perfect H-packing in a graph G is not twice the minimum degree condition in the result of Kühn and Osthus. This will be discussed in more depth in Section 3.1.4.

1.1.2 Matchings in r-uniform hypergraphs

As mentioned earlier, a theorem of Tutte [93] characterises all those graphs that contain a perfect matching. In contrast, a result of Garey and Johnson [29] implies that the decision problem whether an r-uniform hypergraph contains a perfect matching is NP-complete for $r \geq 3$. So again it is natural to seek simple sufficient conditions that ensure a perfect matching. Given an r-uniform hypergraph H and distinct vertices $v_1, \dots, v_\ell \in V(H)$ (where $1 \leq \ell \leq r-1$) we define $d_H(v_1, \dots, v_\ell)$ to be the number of edges containing each of $v_1, \dots, v_\ell$. The minimum ℓ-degree $\delta_\ell(H)$ of H is the minimum of $d_H(v_1, \dots, v_\ell)$ over all ℓ-element sets of vertices in H. We refer to $\delta_1(H)$ as the minimum vertex degree of H, and $\delta_{r-1}(H)$ as the minimum codegree of H.

In recent years there has been significant progress on this problem. Indeed, following on from work in [54, 79], Rödl, Ruciński and Szemerédi [80] characterised the minimum codegree that ensures a perfect matching in an r-uniform hypergraph. However, much less is known about minimum vertex degree conditions for perfect matchings in r-uniform hypergraphs H. Hàn, Person and Schacht [34] gave conditions on $\delta_1(H)$ that ensure a perfect matching in the case when $r \geq 4$. These bounds were subsequently lowered by Markström and Ruciński [65]. This result, however, is believed to be far from tight. In the case when $r = 3$, Hàn, Person and Schacht [34] asymptotically determined the minimum vertex degree that ensures a perfect matching. In Chapter 4 we determine this threshold exactly.

It is also natural to ask for conditions that ensure a matching of given size d in an r-uniform hypergraph H. In the case when d is small compared to the order of H, Bollobás, Daykin and Erdős [11] determined the minimum vertex degree that forces a matching of size d in an r-uniform hypergraph H. In Chapter 4 we extend this result to all possible values of d in the case when H is 3-uniform.

1.2 Hamilton cycles in directed graphs

1.2.1 Degree sequences forcing Hamilton cycles

Dirac's theorem is best possible in the sense that a lower minimum degree condition does not force Hamiltonicity. However, it is of interest to strengthen Dirac's theorem by finding conditions on a graph G of order n which ensure Hamiltonicity but which allow some vertices to have degree much less than $n/2$. Pósa [75] gave such a condition on the so-called degree sequence of a graph: Suppose that the degrees of a graph G of even order n are $d_1 \leq \cdots \leq d_n$. If $n \geq 4$ and $d_i \geq i+1$ for all $i < n/2$ then G contains a Hamilton cycle. So the condition considers graphs G for which nearly half the vertices may have degree much less than $n/2$. Chvátal [19] generalised this result by characterising all those degree sequences that ensure the existence of a Hamilton cycle in a graph.

Finding analogous results for directed graphs (digraphs) has proved to be much more difficult. (Throughout this thesis the digraphs we consider do not have loops and we allow at most one edge in each direction between any pair of vertices.) Ghouila-Houri [30] proved an analogue of Dirac's theorem for digraphs. In Chapter 5 we consider a conjecture of Nash-Williams which, if true, provides a digraph analogue of Chvátal's theorem. Indeed, the conjecture would imply a complete characterisation of all those digraph degree sequences which force Hamiltonicity. No progress has been made on Nash-Williams' conjecture so far. However, we will prove an approximate version of this conjecture for sufficiently large digraphs (Theorem 5.2). In order to prove this result we will prove a stronger result which ensures a Hamilton cycle in 'robustly expanding digraphs' of linear degree (see Theorem 5.13).

An oriented graph is a digraph which can be obtained from an undirected graph by orienting its edges. Thomassen [88] raised the question of an analogue of Dirac's theorem for oriented graphs. Proving a conjecture of Häggkvist [31], Keevash, Kühn and Osthus [41] determined the bound on the minimum semidegree of an oriented graph G which forces G to contain a Hamilton cycle (for sufficiently large oriented graphs). As indicated earlier, for undirected graphs Pósa's theorem is much stronger than Dirac's theorem. It is natural to seek a result which strengthens Häggkvist's conjecture in the same way. Interestingly though, in Section 5.4 we show that no such analogue of Pósa's theorem exists.

1.2.2 Powers of Hamilton cycles and related problems

A well-studied generalisation of the notion of a Hamilton cycle is that of the rth power of a Hamilton cycle. (The rth power of a Hamilton cycle C is obtained from C by adding an edge between every pair of vertices of distance at most r on C.) Seymour [84] gave a conjectural bound on the minimum degree of a graph G that forces G to contain the rth power of a Hamilton cycle. This conjecture was verified for large graphs by Komlós, Sárközy and Szemerédi [50]. Seymour's conjecture extends a conjecture of Pósa (see [25]) who proposed the bound in the case of the square of a Hamilton cycle (that is, when $r = 2$).

The notion of the rth power of a Hamilton cycle also makes sense in the digraph setting: In this case the rth power of a Hamilton cycle C is the digraph obtained from C by adding a directed edge from x to y if there is a path of length at most r from x to y on C. In Section 6.1 we give a conjecture on the minimum semidegree of an oriented graph G which ensures that G contains the square of a Hamilton cycle. We also show that, if true, the conjecture would be best possible.

Notice that in the case when $r + 1$ divides $|G|$, a necessary condition for a graph G to contain the rth power of a Hamilton cycle is that G contains a perfect K_{r+1}-packing. In fact, the Hajnal-Szemerédi theorem together with the result of Komlós, Sárközy and Szemerédi show that the minimum degree bound which forces a perfect K_{r+1}-packing is the same as the minimum degree bound which forces the rth power of a Hamilton cycle.

Similarly when 3 divides $|G|$, a necessary condition for an oriented graph G to contain the square of a Hamilton cycle is that G contains a perfect packing of transitive triangles. In Section 6.2 we give a conjecture on the minimum semidegree of an oriented graph G which ensures that G contains a perfect packing of transitive triangles. Perhaps surprisingly, this bound is lower than the bound given in our conjecture concerning the square of a Hamilton cycle.

1.2.3 Decomposing oriented graphs into Hamilton cycles

Another variant of the Hamilton cycle problem which has received much attention is the problem of whether a graph or digraph G has a Hamilton decomposition. That is, whether the edge set of G can be decomposed into a collection of edge-disjoint Hamilton cycles. The problem originates from 1892 when Walecki showed that K_n has a Hamilton decomposition precisely when n is odd.

A regular tournament is an orientation of a complete graph such that every vertex has equal in- and outdegree. In 1968 Kelly (see e.g. [8, 13, 67]) conjectured that every regular tournament has a Hamilton decomposition. Despite receiving much attention this problem remains open. However, in Chapter 7 we prove an approximate version of Kelly's conjecture (Theorem 7.2) which roughly states that all sufficiently large regular tournaments G can be 'almost' decomposed into edge-disjoint Hamilton cycles (i.e. all but $o(|G|^2)$ edges of G lie in a collection of edge-disjoint Hamilton cycles).

Instead of proving our approximate version of Kelly's conjecture directly, we prove a much stronger result (Theorem 7.3). Indeed, we give a condition on the minimum semidegree of an 'almost regular' oriented graph G that ensures the edge set of G can be almost decomposed into edge-disjoint Hamilton cycles. (Here, by 'almost regular' we mean every vertex has roughly the same in- and outdegree.)

In 1982 Thomassen [89] posed a weaker version of Kelly's conjecture: If G is a regular tournament on $2k+1$ vertices and A is any set of at most $k-1$ edges of G, then $G-A$ has a Hamilton cycle. Using our result (Theorem 5.13) concerning Hamilton cycles in robustly expanding digraphs we prove this conjecture in the case when G is large. The content of Chapters 3, 4, 5 and 7 is based on joint work [60, 63, 61, 62] with Kühn and Osthus.

1.3 Szemerédi's Regularity lemma

Szemerédi's Regularity lemma [86] allows us to approximate sufficiently large and dense graphs by a 'random-like' graph. The Blow-up lemma of Komlós, Sárközy and Szemerédi [49] provides a way of embedding spanning subgraphs H of bounded degree into such random-like graphs. Thus these results are essential tools in proving several of the theorems given in this thesis. Indeed, Alon and Shapira [3] established a variant of the Regularity lemma for digraphs. This will be exploited in the proof of Theorems 5.2 and 7.3. We will not use the Blow-up lemma directly in the proof of Theorem 5.2. However, we will use a result (Lemma 5.9) from [41] whose proof uses a version of the Blow-up lemma due to Csaba [20]. (We do not, however, use the Blow-up lemma in the proof of Theorem 7.3.) The proof of Theorem 3.2 given in Chapter 3 uses the 'standard' version of the Blow-up lemma (Lemma 2.6). In Chapter 2 we draw together all the information we require concerning the Regularity lemma and the Blow-up lemma.

1.4 Notation and preliminaries

If G is a graph or digraph we write $V(G)$ to denote the set of vertices of G and $E(G)$ the set of its edges. Furthermore $e(G)$ denotes the number of edges in G and $|G|$ the order of G.

Given a graph G and a vertex $x \in V(G)$ we denote by $d_G(x)$ the degree of x in G, $\delta(G)$ the minimum degree of G and $\Delta(G)$ the maximum degree of G. The chromatic number of G is denoted by $\chi(G)$.

Given a graph G and disjoint $A, B \subseteq V(G)$, an *A-B edge* is an edge of G with one endvertex in A and the other in B. The number of these edges is denoted by $e_G(A, B)$ or $e(A, B)$ if this is unambiguous. We write $(A, B)_G$ for the bipartite subgraph of G with vertex classes A and B whose edges are precisely the A-B edges in G. Similarly, given a digraph G and disjoint $A, B \subseteq V(G)$, we write $e_G(A, B)$ for the number of all those edges which are directed from some vertex in A to some vertex in B. We also write $(A, B)_G$ for the oriented bipartite subgraph of G with vertex classes A and B whose edges are precisely the edges from A to B in G.

Given two vertices x and y of a digraph G, we write xy for the edge directed from x to y. We denote by $N^+_G(x)$ and $N^-_G(x)$ the out- and the inneighbourhood of x and by $d^+_G(x)$ and $d^-_G(x)$ its out- and indegree. We will write $N^+(x)$ for example, if this is unambiguous.

Given $S \subseteq V(G)$, we write $N^+_G(S)$ for the union of $N^+_G(x)$ for all $x \in S$ and define $N^-_G(S)$ analogously. The *minimum semidegree* $\delta^0(G)$ of G is the minimum of its minimum outdegree $\delta^+(G)$ and its minimum indegree $\delta^-(G)$. The maximum of the maximum outdegree $\Delta^+(G)$ and the maximum indegree $\Delta^-(G)$ is denoted by $\Delta^0(G)$.

Throughout this thesis we omit floors and ceilings whenever this does not affect the argument.

Chapter 2
The Regularity lemma and the Blow-up lemma

2.1 The Regularity lemma for graphs

Szemerédi's Regularity lemma [86] has proved to be an incredibly powerful and useful tool in graph theory as well as in Ramsey theory, combinatorial number theory and other areas of mathematics and theoretical computer science. Indeed, the result was initially proved by Szemerédi in order to prove a conjecture of Erdős and Turán [26] that sequences of integers of positive upper density must contain long arithmetic progressions.

The lemma essentially says that large dense graphs can be approximated by a random-like graph. The strength of this result will be useful for the proof of the results concerning packings in graphs given in Chapter 3. As mentioned in Section 1.3, there is a version of the Regularity lemma for digraphs due to Alon and Shapira [3], which will be used in the proof of Theorems 5.2 and 7.3.

Before stating the Regularity lemma we first need to introduce some more notation and definitions. The *density* of a bipartite graph G with vertex classes A and B is defined to be

$$d_G(A,B) := \frac{e_G(A,B)}{|A||B|}.$$

We will write $d(A,B)$ if this is unambiguous. Given any $\varepsilon, \varepsilon' > 0$, we say that G is $[\varepsilon, \varepsilon']$-*regular* if for all sets $X \subseteq A$ and $Y \subseteq B$ with $|X| \geq \varepsilon|A|$ and $|Y| \geq \varepsilon|B|$ we have

$|d(A,B) - d(X,Y)| < \varepsilon'$. In the case when $\varepsilon = \varepsilon'$ we say that G is *ε-regular* (we also say that $(A,B)_G$ is an *ε-regular pair*). One can think of an ε-regular pair as a bipartite graph which has its edges distributed in a fairly uniform way. Further, the smaller ε is, the 'more uniform' the pair is.

The notion of a super-regular pair is similar to that of a regular pair. However, here we require a lower bound on the degrees of the vertices in such a pair. Indeed, given a bipartite graph G with vertex classes A and B and given any $\varepsilon > 0$ and $d \in [0,1)$ we say that G is *(ε, d)-super-regular* if all sets $X \subseteq A$ and $Y \subseteq B$ with $|X| \geq \varepsilon|A|$ and $|Y| \geq \varepsilon|B|$ satisfy $d(X,Y) > d$ and, furthermore, if $d_G(a) > d|B|$ for all $a \in A$ and $d_G(b) > d|A|$ for all $b \in B$. (In Chapter 7 it will be more convenient to use a slight variant of this definition.) The next fact states that every regular pair has an almost spanning subgraph which is super-regular.

Fact 2.1 *If (A,B) is an ε-regular pair with density d (where $0 < \varepsilon < 1/3$), then there exists $A' \subseteq A$ and $B' \subseteq B$ with $|A'| \geq (1-\varepsilon)|A|$ and $|B'| \geq (1-\varepsilon)|B|$, such that (A', B') is a $(2\varepsilon, d - 3\varepsilon)$-super-regular pair.*

Szemerédi's Regularity lemma states that we can partition the vertices of any large graph into a bounded number of 'clusters' so that most of the pairs of clusters induce ε-regular pairs.

Lemma 2.2 (Szemerédi [86]) *For every $\varepsilon > 0$ and each integer ℓ_0 there is an $M = M(\varepsilon, \ell_0)$ such that if G is any graph on at least M vertices then there exists a partition of $V(G)$ into $V_0, V_1, \dots, V_\ell$ such that the following holds:*

- $\ell_0 \leq \ell \leq M$,
- $|V_0| \leq \varepsilon|G|$,
- $|V_1| = \dots = |V_\ell| =: L$,
- *for all but $\varepsilon\ell^2$ pairs $1 \leq i < j \leq \ell$ the graph $(V_i, V_j)_G$ is ε-regular.*

In this thesis we will use the following degree form of Szemerédi's Regularity lemma which can easily be derived from Lemma 2.2.

Lemma 2.3 (Degree form of the Regularity lemma) *For every $\varepsilon > 0$ and each integer ℓ_0 there is an $M = M(\varepsilon, \ell_0)$ such that if G is any graph on at least M vertices and $d \in [0, 1)$, then there exists a partition of $V(G)$ into $\ell + 1$ classes $V_0, V_1, ..., V_\ell$, and a spanning subgraph $G' \subseteq G$ with the following properties:*

- $\ell_0 \leq \ell \leq M$, $|V_0| \leq \varepsilon|G|$, $|V_1| = \cdots = |V_\ell| =: L$,
- $d_{G'}(v) > d_G(v) - (d + \varepsilon)|G|$ *for all* $v \in V(G)$,
- $e(G'[V_i]) = 0$ *for all* $i \geq 1$,
- *for all* $1 \leq i < j \leq \ell$ *the graph* $(V_i, V_j)_{G'}$ *is* ε*-regular and has density either* 0 *or greater than* d.

The sets $V_1, \ldots, V_\ell$ are called *clusters*, V_0 is called the *exceptional set* and the vertices in V_0 *exceptional vertices*. We refer to G' as the *pure graph of* G. Clearly, we may assume that $(V_i, V_j)_G$ is not ε-regular or has density at most d whenever $(V_i, V_j)_{G'}$ contains no edges (for all $1 \leq i < j \leq \ell$). The *reduced graph* R *of* G is the graph whose vertices are $V_1, \ldots, V_\ell$ and in which V_i is adjacent to V_j whenever $(V_i, V_j)_{G'}$ is ε-regular and has density greater than d. The reduced graph R of a graph G inherits certain properties of G. For example, the next fact states that R 'almost inherits' the minimum degree of G.

Fact 2.4 *Suppose R is the reduced graph of G with parameters ε and d. If $0 < 2\varepsilon \leq d \leq c/2$ and $\delta(G) \geq c|G|$ then $\delta(R) \geq (c - 2d)|R|$.*

It is often useful to consider the reduced graph R of a graph G when seeking some given substructure in G. This is illustrated by the following Embedding lemma. The proof is based on a simple greedy argument, see e.g. Lemma 7.5.2 in [22] or Theorem 2.1 in [52] for details.

Lemma 2.5 (Embedding lemma) *Let H be an r-partite graph with vertex classes $X_1, \ldots, X_r$ and let ε, d, n_0 be constants such that $0 < 1/n_0 \ll \varepsilon \ll d, 1/|H|$. Let G be an r-partite graph with vertex classes $V_1, \ldots, V_r$ of size at least n_0 such that $(V_i, V_j)_G$ is ε-regular and has density at least d whenever H contains an edge between X_i and X_j (for all $1 \leq i < j \leq r$). Then G contains a copy of H such that $X_i \subseteq V_i$.*

Here (and later on) we write $0 < a_1 \ll a_2 \ll a_3 \leq 1$ to mean that we can choose the constants a_1, a_2, a_3 from right to left. More precisely, there are increasing functions f and g such that, given a_3, whenever we choose some $a_2 \leq f(a_3)$ and $a_1 \leq g(a_2)$, all calculations needed in the proof of Lemma 2.5 are valid.

Let H be an r-partite graph and suppose we have applied Lemma 2.3 with parameters ε and d to G to obtain clusters of size L such that $0 < 1/L \ll \varepsilon \ll d, 1/|H|$. Then the Embedding lemma tells us that if we have found a copy of K_r in R then we can find a copy of H in G.

The Embedding lemma cannot be used by itself to find spanning subgraphs of a graph G. However, the Blow-up lemma of Komlós, Sárközy and Szemerédi [49] states that one can even find a spanning subgraph H in G provided that H has bounded maximum degree and the bipartite pairs forming G are super-regular.

Lemma 2.6 (Blow-up lemma) *Given a graph R with $V(R) = \{1, \dots, r\}$ and $d, \Delta > 0$, there is a constant $\varepsilon_0 = \varepsilon_0(d, \Delta, r) > 0$ such that the following holds. Given $L_1, \dots, L_r \in \mathbb{N}$ and $0 < \varepsilon \leq \varepsilon_0$, let R^* be the graph obtained from R by replacing each vertex $i \in V(R)$ with a set V_i of L_i new vertices and joining all vertices in V_i to all vertices in V_j precisely when $ij \in E(R)$. Let G be a spanning subgraph of R^* such that for every $ij \in E(R)$ the bipartite graph $(V_i, V_j)_G$ is (ε, d)-super-regular. Then G contains a copy of every subgraph H of R^* with $\Delta(H) \leq \Delta$.*

2.2 The Regularity lemma for digraphs

In the proof of Theorems 5.2 and 7.3 we will use the directed version of Szemerédi's Regularity lemma. Before we state it we need to define what we mean by an ε-regular pair in a digraph. Recall that given disjoint vertex sets A and B in a digraph G, we write $(A, B)_G$ for the oriented bipartite subgraph of G whose vertex classes are A and B and whose edges are all the edges from A to B in G. We say $(A, B)_G$ is *$[\varepsilon, \varepsilon']$-regular and has density d'* if this holds for the underlying undirected bipartite graph of $(A, B)_G$. (Note that the ordering of the pair $(A, B)_G$ is important here.) In the case when $\varepsilon = \varepsilon'$ we say that $(A, B)_G$ *is*

ε*-regular and has density* d'. Similarly, given $d \in [0,1)$ we say $(A,B)_G$ is (ε,d)*-super-regular* if this holds for the underlying undirected bipartite graph. The Diregularity lemma is a variant of the Regularity lemma for digraphs due to Alon and Shapira [3]. Its proof is similar to the undirected version. We will use the degree form of the Diregularity lemma which is derived (see for example [95]) from the standard version in the same manner as the undirected degree form.

Lemma 2.7 (Degree form of the Diregularity lemma) *For every* $\varepsilon \in (0,1)$ *and every integer* M' *there are integers* M *and* n_0 *such that if* G *is a digraph on* $n \geq n_0$ *vertices and* $d \in [0,1)$ *is any real number, then there is a partition of the vertex set of* G *into* $V_0, V_1, \ldots, V_L$ *and a spanning subdigraph* G' *of* G *such that the following holds:*

- $M' \leq L \leq M$,
- $|V_0| \leq \varepsilon n$,
- $|V_1| = \cdots = |V_L| =: m$,
- $d^+_{G'}(x) > d^+_G(x) - (d+\varepsilon)n$ *for all vertices* $x \in V(G)$,
- $d^-_{G'}(x) > d^-_G(x) - (d+\varepsilon)n$ *for all vertices* $x \in V(G)$,
- *for all* $i = 1, \ldots, L$ *the digraph* $G'[V_i]$ *is empty,*
- *for all* $1 \leq i,j \leq L$ *with* $i \neq j$ *the pair* $(V_i,V_j)_{G'}$ *is* ε*-regular and has density either* 0 *or density at least* d.

As in the graph case, we call $V_1, \ldots, V_L$ *clusters*, V_0 the *exceptional set* and the vertices in V_0 *exceptional vertices*. We refer to G' as the *pure digraph*. The last condition of the lemma says that all pairs of clusters are ε-regular in both directions (but possibly with different densities). The *reduced digraph* R *of* G *with parameters* ε, d *and* M' is the digraph whose vertices are $V_1, \ldots, V_L$ and in which V_iV_j is an edge precisely when $(V_i,V_j)_{G'}$ is ε-regular and has density at least d.

Chapter 3
An Ore-type theorem for perfect packings in graphs

3.1 Introduction

3.1.1 Perfect packings in graphs of large minimum degree

Given two graphs H and G, an *H-packing in G* is a collection of vertex-disjoint copies of H in G. An H-packing is called *perfect* if it covers all the vertices of G. In this case one also says that G contains an *H-factor*. H-packings are generalisations of graph matchings (which correspond to the case when H is a single edge).

In the case when H is an edge, Tutte's theorem characterises those graphs which have a perfect H-packing. However, for other connected graphs H no characterisation is known. Furthermore, Hell and Kirkpatrick [35] showed that the decision problem whether a graph G has a perfect H-packing is NP-complete precisely when H has a component consisting of at least 3 vertices. It is natural therefore to ask for simple sufficient conditions which ensure the existence of a perfect H-packing. One such result is a theorem of Hajnal and Szemerédi [33] which states that a graph G whose order n is divisible by r has a perfect K_r-packing provided that $\delta(G) \geq (1 - 1/r)n$. It is easy to see that the minimum degree condition here is best possible. So for $H = K_r$, the parameter which governs the existence of a perfect H-packing in a graph G of large minimum degree is $\chi(H) = r$.

Kühn and Osthus [56, 57] showed that for any graph H either the so-called critical chromatic number or the chromatic number of H is the relevant parameter. Here the *critical chromatic number* $\chi_{cr}(H)$ of a graph H is defined as

$$\chi_{cr}(H) := (\chi(H) - 1)\frac{|H|}{|H| - \sigma(H)},$$

where $\sigma(H)$ denotes the size of the smallest possible colour class in any $\chi(H)$-colouring of H. When considering H-packings we will only consider graphs H which contain at least one edge (without mentioning this explicitly), so $\chi_{cr}(H)$ is well defined. Note that $\chi(H) - 1 < \chi_{cr}(H) \leq \chi(H)$ for all graphs H, and $\chi_{cr}(H) = \chi(H)$ precisely when every $\chi(H)$-colouring of H has colour classes of equal size. The characterisation of when $\chi(H)$ or $\chi_{cr}(H)$ is the relevant parameter depends on the so-called highest common factor of H, which is defined as follows.

We say that a colouring of H is *optimal* if it uses exactly $\chi(H) =: r$ colours. Given an optimal colouring c of H, let $x_1 \leq x_2 \leq \dots \leq x_r$ denote the sizes of the colour classes of c. We write $\mathcal{D}(c) := \{x_{i+1} - x_i \mid i = 1, \dots, r-1\}$, and let $\mathcal{D}(H)$ denote the union of all the sets $\mathcal{D}(c)$ taken over all optimal colourings c of H. We denote by $\mathrm{hcf}_\chi(H)$ the highest common factor of all integers in $\mathcal{D}(H)$. If $\mathcal{D}(H) = \{0\}$ then we define $\mathrm{hcf}_\chi(H) := \infty$. We write $\mathrm{hcf}_c(H)$ for the highest common factor of all the orders of components of H. For non-bipartite graphs H we say that $\mathrm{hcf}(H) = 1$ if $\mathrm{hcf}_\chi(H) = 1$. If $\chi(H) = 2$ then we say $\mathrm{hcf}(H) = 1$ if $\mathrm{hcf}_c(H) = 1$ and $\mathrm{hcf}_\chi(H) \leq 2$. (See [57] for some examples.) Put

$$\chi^*(H) := \begin{cases} \chi_{cr}(H) & \text{if } \mathrm{hcf}(H) = 1; \\ \chi(H) & \text{otherwise.} \end{cases}$$

Also let $\delta(H, n)$ denote the smallest integer k such that every graph G whose order n is divisible by $|H|$ and with $\delta(G) \geq k$ contains a perfect H-packing.

Theorem 3.1 (Kühn and Osthus [57]) *For every graph H there exists a constant $C =$*

$C(H)$ such that

$$\left(1 - \frac{1}{\chi^*(H)}\right) n - 1 \le \delta(H,n) \le \left(1 - \frac{1}{\chi^*(H)}\right) n + C.$$

Theorem 3.1 improved previous bounds by Alon and Yuster [5], who showed that $\delta(H,n) \le (1 - 1/\chi(H))n + o(n)$, and by Komlós, Sárközy and Szemerédi [51], who replaced the $o(n)$-term by a constant depending only on H. Further related results are discussed in the surveys [46, 47, 52, 58, 97].

3.1.2 Ore-type degree conditions for perfect packings

Of course, one can also consider other types of degree conditions that ensure a perfect H-packing in a graph G. One natural such condition is an Ore-type degree condition requiring a lower bound on the sum of the degrees of non-adjacent vertices of G. (The name comes from Ore's theorem [73], which states that a graph G of order $n \ge 3$ contains a Hamilton cycle if $d(x) + d(y) \ge n$ for all non-adjacent $x \ne y \in V(G)$.)

A result of Kierstead and Kostochka [45] on equitable colourings implies that a graph G whose order n is divisible by r and with $d(x) + d(y) \ge 2(1 - 1/r)n - 1$ for all non-adjacent $x \ne y \in V(G)$ contains a perfect K_r-packing. Note that this is a strengthening of the Hajnal-Szemerédi theorem. Kawarabayashi [40] asked for the Ore-type condition which guarantees a K_4^--packing in a graph G covering a given number of vertices of G. (Here K_4^- denotes the graph obtained from K_4 by removing an edge.) Similarly it is natural to seek an Ore-type analogue of Theorem 3.1. This will be the main result of this chapter (but with an $o(n)$-error term). Perhaps surprisingly, the Ore-type condition needed is not 'twice the minimum degree condition'. For some graphs H it depends on the so-called colour extension number of H, which we will define now. Roughly speaking, this is a measure of how many extra colours we need to properly colour H if we try to build this colouring by extending an $(r-2)$-colouring of a neighbourhood of a vertex of H.

More precisely, suppose that H is a graph with $\chi(H) =: r$ which contains a vertex x for which the subgraph $H[N(x)]$ induced by the neighbourhood of x is $(r-2)$-colourable.

Given such a vertex $x \in V(H)$, let m_x denote the smallest integer for which there exists an $(r-2)$-colouring of $H[N(x)]$ that can be extended to an $(r+m_x)$-colouring of H. The *colour extension number* $CE(H)$ *of* H is defined as

$$CE(H) := \min\{m_x \mid x \in V(H) \text{ with } \chi(H[N(x)]) \leq r-2\}.$$

If $\chi(H[N(x)]) = r-1$ for all $x \in V(H)$ we define $CE(H) := \infty$. So every bipartite graph H without isolated vertices has $CE(H) = \infty$. All other bipartite graphs H have $CE(H) = 0$. In general, $1 \leq CE(H) < \infty$ if for any optimal colouring of H and any $v \in V(H)$, $N(v)$ lies in exactly $r-1$ colour classes of H, but there exists a vertex $x \in V(H)$ such that $\chi(H[N(x)]) \leq r-2$. Note that in this case $CE(H) \leq r-2$. (Indeed, we can colour $H - N(x)$ with r different colours to obtain a $(2r-2)$-colouring of H.)

In order to help the readers to familiarise themselves with the notion of the colour extension number we now give a number of examples. $\chi(K_4^-) = 3$ and $\chi(K_4^-[N(x)]) = 2$ for every vertex x of K_4^-. Thus $CE(K_4^-) = \infty$. Next consider the graph $F^\diamond$ obtained from the complete 3-partite graph $K_{2,2,2}$ by removing an edge xy of $K_{2,2,2}$ and adding a new vertex z which is adjacent to x and y only. Then $\chi(F^\diamond) = 3$, $\chi(F^\diamond[N(w)]) = 2$ for every vertex $w \neq z$ in $F^\diamond$ and $\chi(F^\diamond[N(z)]) = 1$. Note that in any 3-colouring of $F^\diamond$, x and y are coloured differently. So if we 1-colour $N(z) = \{x, y\}$, this colouring can be extended to a 4-colouring of $F^\diamond$ but not a 3-colouring. Thus $CE(F^\diamond) = 1$.

For each $k \geq 1$ and $r \geq k+2$ we now give an example of a family of graphs $H^\diamond$ with $CE(H^\diamond) = k$ and $\chi(H^\diamond) = r$. Consider a complete r-partite graph whose vertex classes $V_1, \dots, V_r$ have size $> k$. Let $H^\diamond$ be obtained from this graph by deleting the edges of k vertex-disjoint copies $K^1, \dots, K^k$ of K_{k+1} which lie in $V_1 \cup \dots \cup V_{k+1}$, and by adding a new vertex x which is adjacent to the $k(k+1)$ vertices lying in these copies of K_{k+1} as well as to all the vertices in $V_{k+2}, \dots, V_{r-1}$ (see Figure 3.1). Note that $\chi(H^\diamond) = r$. Furthermore, any vertex $y \in V_1 \cup \dots \cup V_r$ lies in a copy of K_r in $H^\diamond$. So $\chi(H^\diamond[N(y)]) = r-1$. However, the subgraph $D := H^\diamond[N(x) \cap V_1 \cap \dots \cap V_{k+1}]$ has a k-colouring c'_x with colour classes $V(K^1), \dots, V(K^k)$ and it is easy to check that this is the only k-colouring of D (and so in

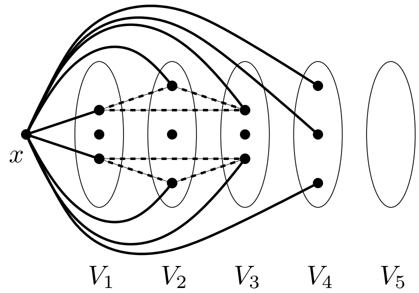

Figure 3.1: The graph $H^\diamond$ in the case when $k = 2$, $r = 5$ and when each V_i has size 3. The dashed lines indicate the deleted edges.

particular $\chi(D) = k$). Thus $\chi(H^\diamond[N(x)]) = r-2$ and the only $(r-2)$-colouring of $H^\diamond[N(x)]$ is the one which agrees with c'_x on D and colours each of $V_{k+2}, \dots, V_{r-1}$ with a new colour. Let c_x denote this colouring. When extending c_x to a proper colouring of $H^\diamond$ we cannot reuse the $r-2$ colours used in c_x since every $y \in V(H^\diamond) \setminus N(x)$ is adjacent to a vertex in each colour class of c_x. As $\chi(H^\diamond - N(x)) = r - (r-k-2) = k+2$ this means that we require $r+k$ colours in total to extend c_x to a proper colouring of $H^\diamond$. Thus $CE(H^\diamond) = k$.

Let

$$\chi_{\text{Ore}}(H) := \begin{cases} \chi(H) & \text{if } \text{hcf}(H) \neq 1 \text{ or } CE(H) = \infty; \\ \max\left\{\chi_{cr}(H), \chi(H) - \frac{2}{CE(H)+2}\right\} & \text{otherwise.} \end{cases}$$

Recall that $CE(K_4^-) = \infty$ and $CE(F^\diamond) = 1$, where $F^\diamond$ was defined above. So $\chi_{\text{Ore}}(K_4^-) = \chi(K_4^-) = 3$. Any 3-colouring of $F^\diamond$ has one colour class of size 3 and two colour classes of size 2. So $\text{hcf}(F^\diamond) = 1$ and thus $\chi_{\text{Ore}}(F^\diamond) = \max\{\chi_{cr}(F^\diamond), 3-2/3\} = \max\{14/5, 7/3\} = 14/5$.

Note that if $\text{hcf}(H) = 1$ and $CE(H) = 0$ then $\chi_{\text{Ore}}(H) = \chi_{cr}(H)$ (an odd cycle of length at least 5 provides an example of such a graph H). On the other hand, one can choose the sizes of the vertex classes V_i in the preceding example $H^\diamond$ so that $\chi_{\text{Ore}}(H^\diamond)$ lies strictly between $\chi_{cr}(H^\diamond)$ and $\chi(H^\diamond)$. (For instance, take k large, $|V_1| = k+1$, $|V_2| = 2k$ and $|V_i| = 2k+1$ for all $i \geq 3$. Then $\chi_{cr}(H^\diamond)$ is close to $\chi(H^\diamond) - 1/2$, $\text{hcf}(H^\diamond) = 1$ and so $\chi_{\text{Ore}}(H^\diamond) = \chi(H^\diamond) - 2/(k+2)$.)

Given a graph H, let $\delta_{\text{Ore}}(H, n)$ be the smallest integer k such that every graph G

whose order n is divisible by $|H|$ and with $d(x) + d(y) \geq k$ for all non-adjacent $x \neq y \in V(G)$ contains a perfect H-packing. Roughly speaking, our next result states that when considering an Ore-type degree condition, for any graph H, $\chi_{\text{Ore}}(H)$ is the relevant parameter which governs the existence of a perfect H-packing. In particular, it implies that we do not have a 'dichotomy' involving only $\chi(H)$ and $\chi_{cr}(H)$ as in Theorem 3.1.

Theorem 3.2 *For every graph H and each $\eta > 0$ there exists a constant $C = C(H)$ and an integer $n_0 = n_0(H, \eta)$ such that if $n \geq n_0$ then*

$$2\left(1 - \frac{1}{\chi_{\text{Ore}}(H)}\right) n - C \leq \delta_{\text{Ore}}(H, n) \leq 2\left(1 - \frac{1}{\chi_{\text{Ore}}(H)} + \eta\right) n.$$

So for example, Theorem 3.2 implies that $\lim_{n\to\infty} \delta_{\text{Ore}}(K_4^-, n)/n = 4/3$ and $\lim_{n\to\infty} \delta_{\text{Ore}}(F^\diamond, n)/n = 9/7$.

The upper bound in Theorem 3.2 follows from Lemmas 3.11 and 3.12 in Section 3.3, which in turn are proved in Sections 3.3 and 3.5. The lower bound is proved in Section 3.2. For every graph H there are infinitely many values of n for which we can take $C = 2$ in Theorem 3.2. In fact, if $\text{hcf}(H) \neq 1$ or $CE(H) = \infty$ then $C = 2$ suffices for all n divisible by $|H|$. In general $C \leq 2|H|^4$ (see Section 3.2). It would be interesting to know whether one can replace the error term ηn by a constant depending only on H.

3.1.3 Almost perfect packings

The critical chromatic number was first introduced by Komlós [48], who showed that it is the relevant parameter when considering 'almost' perfect H-packings.

Theorem 3.3 (Komlós [48]) *For every graph H and each $\gamma > 0$ there exists an integer $n_0 = n_0(\gamma, H)$ such that every graph G of order $n \geq n_0$ and minimum degree at least $(1 - 1/\chi_{cr}(H))n$ contains an H-packing which covers all but at most γn vertices of G.*

It is easy to see that the bound on the minimum degree in Theorem 3.3 is best possible. In the proof of Theorem 3.2 we will use the following result which provides an Ore-type analogue of Theorem 3.3. Again, the critical chromatic number is the relevant parameter for any graph H. In particular, this means that Theorem 3.4 is a generalisation of Theorem 3.3. The proof of Theorem 3.4 is almost identical to that of Theorem 3.3, details can be found in [92].

Theorem 3.4 *For every graph H and each $\eta > 0$ there exists an integer $n_0 = n_0(H, \eta)$ such that if G is a graph on $n \geq n_0$ vertices and*

$$d(x) + d(y) \geq 2\left(1 - \frac{1}{\chi_{cr}(H)}\right) n$$

for all non-adjacent $x \neq y \in V(G)$ then G has an H-packing covering all but at most ηn vertices.

Shokoufandeh and Zhao [85] showed that in Theorem 3.3 the bound on the number of uncovered vertices can be reduced to a constant depending only on H. We conjectured in [60] that this should also be the case for Theorem 3.4.

3.1.4 Copies of H covering a given vertex

In the proof of Theorem 3.2 it will be useful to determine the Ore-type degree condition which guarantees a copy of H covering a given vertex of G. Let $\delta'_{\mathrm{Ore}}(H, n)$ denote the smallest integer k such that whenever w is a vertex of a graph G of order n with $d(x)+d(y) \geq k$ for all non-adjacent $x \neq y \in V(G)$ then G contains a copy of H covering w. Define

$$\chi'_{\mathrm{Ore}}(H) := \begin{cases} \chi(H) & \text{if } CE(H) = \infty; \\ \chi(H) - \frac{2}{CE(H)+2} & \text{otherwise.} \end{cases}$$

Theorem 3.5 *For every graph H and every $\eta > 0$ there exists an integer $n_0 = n_0(H, \eta)$ and a constant $C = C(H)$ such that if $n \geq n_0$ then*

$$2\left(1 - \frac{1}{\chi'_{\mathrm{Ore}}(H)}\right) n - C \leq \delta'_{\mathrm{Ore}}(H, n) \leq 2\left(1 - \frac{1}{\chi'_{\mathrm{Ore}}(H)} + \eta\right) n.$$

Theorem 3.5 is proved in Section 3.3. As in the case of perfect H-packings, the Ore-type degree condition in Theorem 3.5 does not quite match the bound needed for the corresponding minimum degree version. Indeed, let $\delta'(H,n)$ denote the smallest integer k such that whenever w is a vertex of a graph G of order n with $\delta(G) \geq k$ then G contains a copy of H covering w. Together with the Erdős-Stone theorem the next result implies that asymptotically $\delta'(H,n)$ is the same as the minimum degree needed to force any copy of H in a graph of order n.

Proposition 3.6 *For every graph H and every $\eta > 0$ there exists an integer $n_0 = n_0(H,\eta)$ such that if $n \geq n_0$ then*

$$\left(1 - \frac{1}{\chi(H)-1}\right) n - 1 \leq \delta'(H,n) \leq \left(1 - \frac{1}{\chi(H)-1} + \eta\right) n.$$

Proof. Let $r := \chi(H)$. The lower bound on $\delta'(H,n)$ follows by considering a complete $(r-1)$-partite graph G whose vertex classes are as equal as possible. We now prove the upper bound on $\delta'(H,n)$. Let G be a sufficiently large graph of order n such that $\delta(G) \geq \left(1 - \frac{1}{r-1} + \eta\right) n$. Let x be any vertex of G. We have to find a copy of H in G which contains x.

Choose additional constants ε, d, η_1 and α such that

$$0 < \varepsilon \ll d \ll \eta_1 \ll \alpha \ll \eta$$

and let $\ell_0 := 1/\varepsilon$. Apply Lemma 2.3 with parameters ε, d, ℓ_0 to G to obtain clusters $V_1, \dots, V_\ell$ of size L, an exceptional set V_0, a pure graph G' and a reduced graph R. Fact 2.4 implies that

$$\delta(R) \geq \left(1 - \frac{1}{r-1} + \frac{\eta}{2}\right) |R|. \tag{3.1}$$

By adding the vertices of one cluster to V_0 if necessary (and deleting this cluster from R) we may assume that $x \in V_0$. (So now $|V_0| \leq 2\varepsilon n$.)

Let $t \in \mathbb{N}$ be sufficiently large and let F denote the complete r-partite graph with one vertex class of size one and $r-1$ vertex classes of size t. Now $\chi_{cr}(F) = (r-1)\frac{|F|}{|F|-1} = r-1+\frac{1}{t}$. Thus, we may assume that t was chosen so that

$$1 - \frac{1}{r-1} + \frac{\eta}{2} > 1 - \frac{1}{\chi_{cr}(F)}.$$

In fact, we will need to assume that t was chosen so that

$$1-\frac{1}{r-1}+\frac{\eta}{2}>1-\frac{1}{r-1}+\frac{2}{(t+1)(r-1)}\geq 1-\frac{t-1}{|F|}. \tag{3.2}$$

So by the choice of t and by (3.1) we have that

$$\delta(R)\geq\left(1-\frac{1}{\chi_{cr}(F)}\right)|R|.$$

Since ε was chosen to be sufficiently small, $|R|\geq\ell_0$ is sufficiently large so that we can apply Komlós' theorem (Theorem 3.3) to obtain an F-packing $\mathcal{F}$ in R covering all but at most $\eta_1|R|$ vertices in R. We remove all clusters in R that are not covered by this F-packing, and put all the vertices lying in such clusters into V_0. (So now $|V_0|\leq 2\eta_1 n$.)

We say that x *is adjacent to a cluster* $V_i\in V(R)$ if x is adjacent to at least αL vertices of V_i in G. We let $d_R(x)$ denote the number of clusters $V_i\in V(R)$ that x is adjacent to. Now

$$\left(1-\frac{1}{r-1}+\eta\right)n\leq d_G(x)\leq d_R(x)L+(|R|-d_R(x))\alpha L+|V_0|\leq d_R(x)L+2\alpha n$$

and so

$$d_R(x)\geq\left(1-\frac{1}{r-1}+\frac{\eta}{2}\right)|R|. \tag{3.3}$$

We say a copy $F'\in\mathcal{F}$ of F is *useful for* x if x is adjacent to $r-1$ clusters belonging to different vertex classes of F'. Notice that if we have a useful copy F' of F in $\mathcal{F}$ then we can apply the Embedding lemma (Lemma 2.5) to obtain our desired copy of H in G which contains x. Indeed, in this case x could play the role of any vertex $y\in V(H)$. The vertices in $N_H(y)$ would be embedded into the aforementioned $r-1$ clusters of F' that x is adjacent to, and $H-N_H(y)$ would be embedded into the clusters of F'. Thus, it suffices to find a useful copy of F in $\mathcal{F}$.

If a copy $F' \in \mathcal{F}$ of F is not useful then x is adjacent to at most $|F| - t - 1$ clusters in F'. However,

$$|\mathcal{F}|(|F| - t - 1) = \left(1 - \frac{t-1}{|F|}\right)|R| < d_R(x)$$

by (3.2) and (3.3). Thus we must have a useful copy of F in $\mathcal{F}$, as required. $\square$

We will not use Proposition 3.6 in the proof of Theorem 3.2, however it does help to explain the difference between Theorems 3.1 and 3.2. Indeed, Theorem 3.3 and Proposition 3.6 show that the minimum degree which ensures an almost perfect H-packing is larger than the minimum degree which guarantees a copy of H covering any given vertex. In contrast, Theorems 3.4 and 3.5 imply that for some H this is not true in the Ore-type case. So it is natural that $\delta_{\text{Ore}}(H,n)$ involves this property explicitly (since the property that every vertex is contained in a copy of H is clearly necessary to ensure a perfect H-packing). In fact, this is the only real difference to the expression for $\delta(H,n)$ in Theorem 3.1: note that we have $\chi_{\text{Ore}}(H) = \max\{\chi^*(H), \chi'_{\text{Ore}}(H)\}$ and thus Theorems 3.1, 3.2 and 3.5 imply that

$$\delta_{\text{Ore}}(H,n) = \max\left\{2\delta(H,n), \delta'_{\text{Ore}}(H,n)\right\} + o(n).$$

3.1.5 Forcing a single copy of H

In view of Theorem 3.5, one might also wonder what Ore-type degree condition ensures at least one copy of H (i.e. we do not require every vertex to lie in a copy of H). It is easy to see that if G is of order n then the condition is similar to the condition on the minimum degree.

Proposition 3.7 *For every graph H and every $\eta > 0$ there exists an integer $n_0 = n_0(H,\eta)$ such that if $n \geq n_0$ and G is a graph on n vertices which satisfies*

$$d(x) + d(y) \geq 2\left(1 - \frac{1}{\chi(H) - 1} + \eta\right)n$$

for all non-adjacent $x \neq y \in V(G)$, then G contains a copy of H.

Proposition 3.7 immediately follows from the Erdős-Stone theorem and the following observation (which we expect to be known, but we were unable to find a reference):

Proposition 3.8 *Let G be a graph with $d(x) + d(y) \geq 2k$ for all non-adjacent $x \neq y \in V(G)$. Then G has average degree at least k.*

To prove Proposition 3.8, let A be the set of vertices in G whose degree is less than k and let B be the set of remaining vertices. Let $\overline{G}$ denote the complement of G and let F denote the bipartite subgraph of $\overline{G}$ induced by A and B. Hall's theorem implies that F has a matching covering all of A (Hall's condition can be verified by noting that for all $X \subseteq A$ the number of edges in F between X and the neighbourhood of X is at least $|X|(n-k-1)$ and at most $|N(X)|(n-k-1)$). Now apply the Ore-type degree condition to all pairs of vertices of G which are contained in this matching.

3.2 Extremal examples

Let us now prove the lower bound in Theorem 3.2. The next proposition deals with the case when $CE(H) = \infty$.

Proposition 3.9 *Let H be a graph with $CE(H) = \infty$. Let $n \geq |H|$. Then there exists a graph G of order n with*

$$d(x) + d(y) \geq 2\left(1 - \frac{1}{\chi(H)}\right)n - 2$$

for all non-adjacent $x \neq y \in V(G)$ containing a vertex that does not belong to a copy of H. (In particular, G has no perfect H-packing.)

Proof. Let $r := \chi(H)$. Consider the complete r-partite graph of order n whose vertex classes $V_1', V_2', V_3, \dots, V_r$ have sizes as equal as possible, where $|V_1'| \leq |V_2'| \leq |V_3| \leq \dots \leq |V_r|$. Note that $n - |V_1'| - |V_2'| \geq n - 2n/r$.

Let G be obtained from this graph by moving all but one vertex, w say, from V_1' to V_2', by making the set $V_2 \supseteq V_2'$ thus obtained from V_2' into a clique and by deleting all the edges between w and the vertices in V_2.

Any vertex $y \in V_3 \cup \dots \cup V_r$ satisfies $d(y) \geq n - \lceil \frac{n}{r} \rceil \geq (1 - 1/\chi(H))n - 1$. Thus $d(y_1) + d(y_2) \geq 2(1 - 1/\chi(H))n - 2$ for all non-adjacent $y_1 \neq y_2 \in V(G) \backslash (\{w\} \cup V_2)$. Moreover, $d(w) = n - |V_1'| - |V_2'| \geq n - 2n/r$ and for any $z \in V_2$ we have $d(z) = n - 2$. So $d(w) + d(z) \geq 2(1 - 1/\chi(H))n - 2$. Hence G satisfies our Ore-type degree condition.

The neighbourhood of w in G induces an $(r-2)$-partite subgraph of G. Therefore, since $\chi(H[N(x)]) = r - 1$ for all $x \in V(H)$, w cannot play the role of any vertex in H. So G does not contain a copy of H covering w. $\square$

The following proposition will be used for the case when H is non-bipartite and $CE(H) < \infty$.

Proposition 3.10 *Let H be a graph with $r := \chi(H) \geq 3$ for which $m := CE(H) < \infty$. Then there are infinitely many graphs G whose order n is divisible by $|H|$ and such that*

$$d(x) + d(y) \geq 2\left(1 - \frac{1}{r - \frac{2}{m+2}}\right) n - 1$$

for all non-adjacent $x \neq y \in V(G)$ containing a vertex that does not belong to a copy of H. (In particular, G has no perfect H-packing.)

Proof. Let $t \in \mathbb{N}$ be such that $((m+2)r-2)(r-2)$ divides t. Define $s := 2|H|/((m+2)r-2)$. Let G' be the complete $(r + m - 1)$-partite graph with one vertex class V_1 of size $st - 1$, m vertex classes $V_2, \dots, V_{m+1}$ of size st and $r - 2$ vertex classes $V_{m+2}, \dots, V_{r+m-1}$ of size $\frac{|H|t-(m+1)st}{r-2}$. Let G be obtained from G' by adding a vertex w to G' such that w is adjacent to precisely those vertices in $V_{m+2} \cup \dots \cup V_{r+m-1}$. So $|G| = |H|t$.

Any $y \in V_1 \cup \dots \cup V_{m+1}$ satisfies

$$d(y) + d(w) \geq 2|H|t - (m + 2)st - 1 = 2\left(1 - \frac{m + 2}{(m + 2)r - 2}\right)|G| - 1.$$

Furthermore, given any $y_1 \neq y_2 \in V_i$ for some $m+2 \leq i \leq r+m-1$, we have

$$\begin{aligned} d(y_1)+d(y_2) &= 2|H|t - 2\left(\frac{|H|t-(m+1)st}{r-2}\right) = 2|G| - \frac{2}{r-2}\left(1-\frac{2(m+1)}{(m+2)r-2}\right)|G| \\ &= 2|G| - \frac{2}{r-2}\frac{(m+2)(r-2)}{(m+2)r-2}|G| = 2\left(1-\frac{m+2}{(m+2)r-2}\right)|G|. \end{aligned}$$

Since $d(y)+d(y') \geq d(y)+d(w)$ for any $y \neq y' \in V_i$ with $1 \leq i \leq m+1$ this implies that G satisfies our Ore-type degree condition.

Suppose that w belongs to some copy H_w of H in G. Since $\chi(G) = m+r-1$, an optimal colouring of G induces an $(m+r-1)$-colouring of H_w and an $(r-2)$-colouring of $G[N(w)]$. But then w must be playing the role of a vertex $x \in V(H)$ such that $\chi(H[N(x)]) \leq r-2$, contradicting the definition of $m = CE(H)$. $\square$

We will now use Propositions 3.9 and 3.10 to prove the lower bound of Theorem 3.2.

Proof of Theorem 3.2 (lower bound). In the case when $\mathrm{hcf}(H) \neq 1$ the lower bound follows from the lower bound in Theorem 3.1. Proposition 3.9 settles the case when $CE(H) = \infty$. So we may assume that $\mathrm{hcf}(H) = 1$ and $CE(H) < \infty$. In this case, the lower bound in Theorem 3.1 also implies that

$$\delta_{\mathrm{Ore}}(H,n) \geq 2(1-1/\chi_{cr}(H))n - 2 \tag{3.4}$$

(for any graph H). Suppose first that H is bipartite. Since $CE(H) < \infty$ this means that H must have an isolated vertex and so $CE(H) = 0$. Thus $\chi_{\mathrm{Ore}}(H) = \chi_{cr}(H)$ and so we are done by (3.4).

So suppose next that $\chi(H) \geq 3$. In this case the proof of Proposition 3.10 implies the lower bound whenever n is divisible by $((m+2)r-2)(r-2)|H|$. To deduce the lower bound for any $n \geq ((m+2)r-2)(r-2)|H|$ which is divisible by $|H|$ we proceed as follows. Let n' be the largest integer such that $n' \leq n$ and n' is divisible by $((m+2)r-2)(r-2)|H|$. Construct a graph G of order n' as in the proof of Proposition 3.10. Then add $n-n' < ((m+2)r-2)(r-2)|H|$ new vertices to V_1 so that these vertices have the same neighbourhoods as the original vertices in V_1. Then $|G| = n$ and by the same argument as in Proposition 3.10, G does not contain a perfect H-packing. Moreover, it is easy to check that $d(x)+d(y) \geq 2(1-1/(r-2/(m+2)))n - 2|H|^4$ for all non-adjacent $x \neq y \in V(G)$.

$\square$

3.3 Some useful results

In Section 3.2 we proved the lower bound on $\delta_{\mathrm{Ore}}(H,n)$ in Theorem 3.2. The following two results together imply the upper bound.

Lemma 3.11 *Let H be a graph and let $\eta > 0$. There exists an integer $n_0 = n_0(H,\eta)$ such that if G is a graph whose order $n \geq n_0$ is divisible by $|H|$ and*

$$d(x) + d(y) \geq 2\left(1 - \frac{1}{\chi(H)} + \eta\right) n$$

for all non-adjacent $x \neq y \in V(G)$ then G contains a perfect H-packing.

Lemma 3.12 *Let $\eta > 0$ and suppose that H is a graph such that $\mathrm{hcf}(H) = 1$ and $CE(H) < \infty$. There exists an integer $n_0 = n_0(H,\eta)$ such that if G is a graph whose order $n \geq n_0$ is divisible by $|H|$ and*

$$d(x) + d(y) \geq \max\left\{2\left(1 - \frac{1}{\chi(H) - \frac{2}{CE(H)+2}} + \eta\right) n, 2\left(1 - \frac{1}{\chi_{cr}(H)} + \eta\right) n\right\} \quad (3.5)$$

for all non-adjacent $x \neq y \in V(G)$ then G contains a perfect H-packing.

Note that Lemma 3.11 implies the upper bound on $\delta(H,n)$ by Alon and Yuster (which we mentioned in Section 3.1). We now deduce Lemma 3.11 from Lemma 3.12.

Proof of Lemma 3.11. Let $h := |H|$ and $r := \chi(H)$. Given any $k \geq 2$, define H^* to be the complete $(r+1)$-partite graph with one vertex class of size 1, one vertex class of size $hk-1$ and $r-1$ vertex classes of size hk. Let H' be obtained from H^* by removing an edge between some vertex y in a vertex class of size hk and the vertex in the singleton vertex class. So $\chi(H') = r+1$, $|H'| = hkr$ and $\chi(H'[N(y)]) = r-1$. Moreover, $CE(H') = 0$ since $N(y)$

lies in $r-1$ vertex classes of H'. It is easy to see that H' contains a perfect H-packing and that $\mathrm{hcf}(H')=1$. So $\chi_{\mathrm{Ore}}(H')=\chi_{cr}(H')=(\chi(H')-1)\frac{|H'|}{|H'|-\sigma(H')}=r\frac{|H'|}{|H'|-1}$. In particular, we can choose k sufficiently large to guarantee that $1/\chi_{cr}(H')\geq 1/\chi(H)-\eta/4$.

Consider any graph G as in Lemma 3.11. Choose $a\leq kr$ such that $n-ah$ is divisible by $|H'|=hkr$. Apply Proposition 3.7 to obtain a disjoint copies of H in G. Remove these a copies of H from G to obtain a graph G' whose order is divisible by $|H'|$ and which satisfies

$$d_{G'}(x_1)+d_{G'}(x_2)\geq 2\left(1-\frac{1}{\chi(H)}+\frac{\eta}{2}\right)|G'|\geq 2\left(1-\frac{1}{\chi_{cr}(H')}+\frac{\eta}{4}\right)|G'|$$

for all non-adjacent $x_1\neq x_2\in V(G')$. Apply Lemma 3.12 to find a perfect H'-packing in G'. In particular, this induces a perfect H-packing in G'. Thus, together with all those copies of H in $G-G'$ we have chosen before, we obtain a perfect H-packing in G. $\square$

Thus to prove Theorem 3.2 it remains to prove Lemma 3.12, which we will do in Section 3.5. In order to deal with the 'exceptional' vertices in the proof of Lemma 3.12 we use the following result which implies that every vertex w of a graph G as in Lemma 3.12 is contained in a copy of H. We prove Lemma 3.13 in Section 3.4.

Lemma 3.13 *Let H be a graph such that $m:=CE(H)<\infty$. Let $r:=\chi(H)$ and $\eta>0$. There exists an integer $n_0=n_0(\eta,H)$ such that whenever G is a graph on $n\geq n_0$ vertices with*

$$d(x)+d(y)\geq 2\left(1-\frac{1}{r-\frac{2}{m+2}}+\eta\right)n \tag{3.6}$$

for all non-adjacent $x\neq y\in V(G)$ then every vertex of G lies in a copy of H in G.

The above results also imply Theorem 3.5:

Proof of Theorem 3.5. The lower bound in the case when $CE(H)=\infty$ follows from Proposition 3.9. If $CE(H)<\infty$ and $\chi(H)\geq 3$ then Proposition 3.10 gives the lower bound for infinitely many values of n and as in the proof of the lower bound in Theorem 3.2 it can be used to derive the lower bound for any n. If $CE(H)<\infty$ and $\chi(H)=2$ then

$CE(H)=0$ and so the lower bound is trivial. The upper bound follows from Lemmas 3.11 and 3.13. $\square$

Fact 2.4 states that the minimum degree of a graph G is almost inherited by its reduced graph. We now prove an analogue of this for an Ore-type degree condition. This will be useful in the proof of Lemmas 3.12 and 3.13.

Lemma 3.14 *Given a constant c, let G be a graph such that $d_G(x)+d_G(y)\ge c|G|$ for all non-adjacent $x\neq y\in V(G)$. Suppose we have applied Lemma 2.3 with parameters ε and d to G. Let R be the corresponding reduced graph. Then $d_R(V_i)+d_R(V_j)>(c-2d-4\varepsilon)|R|$ for all non-adjacent $V_i\neq V_j\in V(R)$.*

Proof. Let $V_1,\dots,V_\ell$ denote the clusters obtained from Lemma 2.3. Let $L:=|V_1|=\dots=|V_\ell|$, let V_0 denote the exceptional set and let G' be the pure graph. Set $G'':=G'-V_0$. Consider any pair V_iV_j of clusters which does not form an edge in R. Pick $x\in V_i$ and $y\in V_j$ such that $xy\notin E(G)$. So $d_G(x)+d_G(y)\ge c|G|$ and thus $d_{G''}(x)+d_{G''}(y)>(c-2d-4\varepsilon)|G|$. However, by definition of G'', each cluster containing a neighbour of x in G'' must be a neighbour of V_i in R and the analogue holds for the clusters containing the neighbours of y. Thus $d_R(V_i)+d_R(V_j)\ge(d_{G''}(x)+d_{G''}(y))/L\ge(c-2d-4\varepsilon)|R|$, as required. $\square$

In our proof of Lemma 3.12 we will also use the following result, Lemma 12 from [57]. It gives a sufficient condition on the sizes of the vertex classes of a complete $\chi(H)$-partite graph G which ensures that G has a perfect H-packing. Lemma 3.15 is the point where the assumption that $\mathrm{hcf}(H)=1$ is crucial – it is false for graphs with $\mathrm{hcf}(H)\neq 1$.

Lemma 3.15 *Let H be a graph with $\mathrm{hcf}(H)=1$. Put $r:=\chi(H)$ and $\gamma:=(r-1)\sigma(H)/(|H|-\sigma(H))$. Let $0<\beta_1\ll\lambda_1\ll\gamma,1-\gamma,1/|H|$ be positive constants. Suppose that G is a complete r-partite graph with vertex classes $U_1,\dots,U_r$ such that $|G|\gg|H|$ is divisible by $|H|$, $(1-\lambda_1^{1/10})|U_r|\le\gamma|U_i|\le(1-\lambda_1)|U_r|$ for all $i<r$ and such that $||U_i|-|U_j||\le\beta_1|G|$ whenever $1\le i<j<r$. Then G contains a perfect H-packing.*

3.4 Proof of Lemma 3.13

Let H be as in the statement of the lemma and let G be a graph of sufficiently large order n which satisfies (3.6). Recall that $r = \chi(H)$ and $m = CE(H)$. Let x be any vertex of G. We have to find a copy of H in G which contains x. Suppose first that $r = 2$. Then H must have an isolated vertex v (since $CE(H) < \infty$). So we can apply Proposition 3.7 to find a copy of $H - v$ in $G - x$ and thus a copy of H in G (where x plays the role of v).

So suppose that $r \geq 3$. Choose additional constants ε, d and α such that

$$0 < \varepsilon \ll d \ll \alpha \ll \eta$$

and let $\ell_0 := 1/\varepsilon$. Apply the Regularity lemma with parameters ε, d, ℓ_0 to G to obtain clusters $V_1, \dots, V_\ell$ of size L, an exceptional set V_0, a pure graph G' and a reduced graph R. Let

$$k := (m+2)r - 2.$$

Lemma 3.14 implies that

$$d_R(V_i) + d_R(V_j) \geq 2\left(1 - \frac{1}{r - \frac{2}{m+2}} + \frac{\eta}{2}\right)|R| = 2\left(1 - \frac{m+2}{k} + \frac{\eta}{2}\right)|R| \tag{3.7}$$

for all $V_i \neq V_j \in V(R)$ with $V_iV_j \notin E(R)$. By adding the vertices of one cluster to V_0 if necessary (and deleting this cluster from R) we may assume that $x \in V_0$. (So now $|V_0| \leq 2\varepsilon n$.) We say that x *is adjacent to a cluster* $V_i \in V(R)$ if x is adjacent to at least αL vertices of V_i in G. We denote by S the set of clusters $V_i \in V(R)$ that x is adjacent to, and define $s := |S|/|R|$. Also, we write $\overline{S} := V(R) \setminus S$. Note that

$$d_G(x) \leq |S|L + |\overline{S}|\alpha L + |V_0| \leq (s + \alpha + 2\varepsilon)n \leq (s + 2\alpha)n \tag{3.8}$$

and so

$$s \geq \frac{\delta(G)}{n} - 2\alpha \overset{(3.6)}{\geq} \left(1 - \frac{2}{r - \frac{2}{m+2}} + 2\eta\right) - 2\alpha \geq 1 - \frac{2(m+2)}{k} + \eta. \tag{3.9}$$

In particular $s > 0$ since $r \geq 3$. Our aim now is to find either a copy K'_r of K_r in R containing $r-1$ clusters adjacent to x (i.e. $|V(K'_r) \cap S| \geq r-1$), or a copy K'_{r+m} of K_{r+m} in R containing $r-2$ clusters adjacent to x. In both cases we could apply the Embedding lemma (Lemma 2.5) to find the desired copy H_x of H in G. Indeed, in the case where we find K'_{r+m} we could use x to play the role of a vertex $y \in V(H)$ for which there exists an $(r-2)$-colouring of $H[N(y)]$ that can be extended to an $(r+m)$-colouring of H. The neighbourhood $N_H(y)$ of y would be embedded into the clusters belonging to $V(K'_{r+m}) \cap S$ and $H - N_H(y)$ would be embedded into the clusters belonging to $V(K'_{r+m})$ (so here we use the fact that $CE(H) = m$). In the case where we find K'_r, x can play the role of any vertex of H. Given some optimal colouring of H, the vertices of H which have a different colour than x are embedded into the clusters in $V(K'_r) \cap S$ (so we only use that $\chi(H) = r$ in this case).

Let C be the set of clusters $U \in S$ with $d_R(U) < (1 - (m+2)/k + \eta/2)|R|$. By (3.7), C induces a clique. So we may assume that $|C| < r$, since otherwise we have our copy K'_r of K_r. Suppose now that for some $1 \leq i \leq r-1$ we have already found i clusters $U_1, \ldots, U_i \in S \setminus C$ such that $U_1, \ldots, U_i$ form a copy K'_i of K_i in R. Then

$$|\bigcap_{1 \leq j \leq i} N_R(U_j)| \geq -(i-1)|R| + \sum_{j=1}^{i} d_R(U_j) \geq \left(1 - \frac{i(m+2)}{k} + \eta/2\right)|R|. \tag{3.10}$$

Case 1. $1 - s \leq (2m+2)/k$

In this case, we will find a copy of K_r which contains at least $r-1$ vertices in S. Suppose that $i \leq r-2$ and we have found $U_1, \ldots, U_i$ as above. Then $1 - i(m+2)/k \geq (2m+2)/k$ and so (3.10) implies that the common neighbourhood $N_R(K'_i)$ of K'_i satisfies $|N_R(K'_i)| \geq (1 - s + \eta/2)\,|R|$. So we can choose $U_{i+1} \in S \setminus C$ to extend K'_i into a copy of K_{i+1} in $R[S \setminus C]$ (we can avoid C when choosing U_{i+1} since $|C| < r \ll \eta|R|$). If $i = r-1$, then $1 - \frac{i(m+2)}{k} = \frac{m}{k} \geq 0$. So $|N_R(K'_i)| \geq \eta|R|/2$ and we can extend $K'_i = K'_{r-1}$ into the desired copy of K_r using an arbitrary vertex of R.

Case 2. $1 - s \geq (2m+2)/k$

In this case, we will either find a copy of K_r which contains at least $r-1$ vertices in S or find a copy of K_{r+m} which contains at least $r-2$ vertices in S. Suppose that $i \le r-3$ and we have found $U_1, \dots, U_i$ as described before Case 1 which form a copy K'_i of K_i in $R[S \setminus C]$. Note that

$$1 - \frac{i(m+2)}{k} \ge \frac{k-(r-3)(m+2)}{k} = \frac{3(m+2)-2}{k} \ge \frac{2(m+2)}{k} \overset{(3.9)}{\ge} 1-s.$$

Thus (3.10) implies that we can choose a cluster $U_{i+1} \in S \setminus C$ which forms a K_{i+1} together with K'_i. This shows that we can find a copy K'_{r-2} of K_{r-2} which lies in $R[S \setminus C]$. Note that (3.10) also implies that the common neighbourhood $N_R(K'_{r-2})$ of K'_{r-2} satisfies

$$|N_R(K'_{r-2})| \ge \left(1 - \frac{(r-2)(m+2)}{k} + \frac{\eta}{2}\right)|R| = \left(\frac{2(m+1)}{k} + \frac{\eta}{2}\right)|R|. \tag{3.11}$$

Now we aim to extend K'_{r-2} into a copy K'_{r+m} of K_{r+m}. We will aim to find the additional vertices in $\overline{S}$. Suppose for some $0 \le i \le m+1$ we have found i clusters $W_1, \dots, W_i \in \overline{S}$ which together with K'_{r-2} form a copy K'_{r-2+i} of K_{r-2+i} in R. We will need a lower bound on $d_R(W_j)$ for all $j = 1, \dots, i$. To derive this, note that the definition of S implies that W_j contains a vertex y which is not adjacent to x in G. So (3.6) and (3.8) and the inequality in Case 2 imply that

$$d_G(y) \ge \left(2\left(1 - \frac{m+2}{k} + \eta\right) - s - 2\alpha\right)n \ge \left(1 - \frac{2}{k} + \eta\right)n$$

and so $d_{G'}(y) \ge (1 - 2/k + \eta/2)n$. But each cluster containing a neighbour of y in G' must be a neighbour of W_j in R. Hence

$$d_R(W_j) \ge \frac{d_{G'}(y) - |V_0|}{L} \ge \left(1 - \frac{2}{k}\right)|R|. \tag{3.12}$$

So the common neighbourhood $N_R(K'_{r-2+i})$ of K'_{r-2+i} satisfies

$$|N_R(K'_{r-2+i})| \ge |N_R(K'_{r-2})| - i|R| + \sum_{j=1}^{i} d_R(W_j) \overset{(3.11),(3.12)}{\ge} \left(\frac{2(m+1)}{k} - i\frac{2}{k} + \frac{\eta}{2}\right)|R| \ge \frac{\eta|R|}{2}. \tag{3.13}$$

So we can choose a vertex $W_{i+1} \in V(R) \setminus C$ that is a common neighbour of the clusters in K'_{r-2+i}. Suppose that $W_{i+1} \in S$. Then together with K'_{r-2} this forms a copy K'_{r-1} of K_{r-1} in $R[S \setminus C]$. Now (3.10) implies that $|N_R(K'_{r-1})| \geq (m/k + \eta/2)\,|R|$ and so we can extend K'_{r-1} to a copy of K_r with at least $r-1$ vertices in S. So we may assume that $W_{i+1} \in \overline{S}$. Continuing in this way, we obtain a copy of K_{r+m} having $r-2$ clusters in S, as required.

3.5 Proof of Lemma 3.12

3.5.1 Preliminaries and an outline of the proof

Let H, G and $\eta > 0$ be as in Lemma 3.12 and let $r := \chi(H)$. Choose $t \in \mathbb{N}$ such that $t|H|(r-1) \geq 4r/\eta$. Let $z_1 := t(r-1)\sigma(H)$ and $z := t(|H| - \sigma(H))$. Put $\gamma := z_1/z$. Note that $0 < \gamma < 1$ since $\mathrm{hcf}(H) = 1$. Define B^* to be the complete r-partite graph with one vertex class of size z_1 and $r-1$ vertex classes of size z. Then B^* has a perfect H-packing and $\eta|B^*|/4 \geq r$. Moreover,

$$\chi_{cr}(B^*) = \chi_{cr}(H) = (r-1)\frac{|H|}{|H| - \sigma(H)} = r - 1 + \frac{(r-1)\sigma(H)}{|H| - \sigma(H)} = r - 1 + \gamma. \tag{3.14}$$

Choose $s \in \mathbb{N}$ and a new constant λ such that $0 < \lambda \ll \eta, \gamma, 1-\gamma$ as well as $s_1 := \gamma(1+\lambda)s \in \mathbb{N}$ and $s_1 \leq s$. Let B' denote the complete r-partite graph with one vertex class of size s_1 and $r-1$ vertex classes of size s. Thus,

$$\chi_{cr}(B') = (r-1)\frac{|B'|}{|B'| - s_1} = r - 1 + \gamma(1+\lambda). \tag{3.15}$$

Note that the proportion $\gamma(1+\lambda)$ of the size of the smallest vertex class of B' compared to the size of one of the larger classes is slightly larger than the corresponding proportion γ associated with B^*. We can therefore choose s and λ in such a way that B' has a perfect B^*-packing, and thus a perfect H-packing. (Indeed, the perfect B^*-packing would consist of 'most' but not all of the copies of B^* having their smallest vertex class lying in the smallest vertex class of B'.)

We now give an outline for the proof of Lemma 3.12. We first apply the Regularity lemma to G to obtain a reduced graph R. Since R almost inherits the Ore-type condition on G we may apply Theorem 3.4 to find an almost perfect B'-packing of R. We then remove all clusters from R that are not covered by this B'-packing and add the vertices in these clusters to the exceptional set V_0.

For each exceptional vertex $x \in V_0$, we apply Lemma 3.13 to find a copy of H in G containing x, and remove the vertices in this copy from G. Thus some vertices in clusters in R will be removed from G. The copies of H will be chosen to be disjoint for different exceptional vertices.

Our aim is to apply the Blow-up lemma to each copy B'_i of B' in the B'-packing of R in order to find an H-packing in G which covers all the vertices belonging to (the modified) clusters in B'_i. Then all these H-packings together with all those copies of H chosen for the exceptional vertices would form a perfect H-packing in G. However, to do this, we need that the complete r-partite graph F_i^* whose jth vertex class is the union of all the clusters in the jth vertex class of B'_i has a perfect H-packing. Lemma 3.15 gives a condition which guarantees this.

To apply Lemma 3.15 we need that $|F_i^*|$ is divisible by $|H|$. We will remove a bounded number of further copies of H from G to ensure this (see Section 3.5.4). Furthermore, we require that F_i^* has $r-1$ vertex classes of roughly the same size, u say, and that its other vertex class is a little larger than γu. But this condition will be satisfied automatically by the choice of the sizes of the vertex classes in B'. In fact, this is the reason why we chose a B'-packing in R rather than a B^*-packing. The above strategy is based on that in [56]. However, there are additional difficulties.

3.5.2 Applying the Regularity lemma and modifying the reduced graph

We define further constants satisfying

$$0 < \varepsilon \ll d \ll \eta_1 \ll \beta \ll \alpha \ll \lambda \ll \eta, \gamma, 1-\gamma.$$

We also choose η_1 so that

$$\eta_1 \ll \frac{1}{|B'|}.$$

Throughout the proof we assume that the order n of our graph G is sufficiently large for our calculations to hold. Apply the Regularity lemma with parameters ε, d and $\ell_0 := 1/\varepsilon$ to obtain clusters $V_1, \dots, V_\ell$ of size L, an exceptional set V_0, a pure graph G' and a reduced graph R. Let $m := CE(H)$. By Lemma 3.14 we have that

$$d_R(V_{j_1}) + d_R(V_{j_2}) \geq \max\left\{2\left(1 - \frac{1}{r - \frac{2}{m+2}} + \frac{\eta}{2}\right)|R|, 2\left(1 - \frac{1}{\chi_{cr}(H)} + \frac{\eta}{2}\right)|R|\right\}$$

for all $V_{j_1} \neq V_{j_2} \in V(R)$ with $V_{j_1}V_{j_2} \notin E(R)$. Together with (3.14) and (3.15) this implies that

$$d_R(V_{j_1}) + d_R(V_{j_2}) \geq 2\left(1 - \frac{1}{\chi_{cr}(B')}\right)|R|$$

for all $V_{j_1} \neq V_{j_2} \in V(R)$ with $V_{j_1}V_{j_2} \notin E(R)$. So we can apply Theorem 3.4 to R to obtain a B'-packing covering all but at most $\eta_1|R|$ vertices. We denote the copies of B' in this packing by $B'_1, \dots, B'_{\ell'}$. We delete all the clusters not contained in some B'_i from R and add all vertices lying in these clusters to V_0. So $|V_0| \leq \varepsilon n + \eta_1 n \leq 2\eta_1 n$. We now refer to R as this modified reduced graph. We still have that

$$d_R(V_{j_1}) + d_R(V_{j_2}) \geq \max\left\{2\left(1 - \frac{1}{r - \frac{2}{m+2}} + \frac{\eta}{4}\right)|R|, 2\left(1 - \frac{1}{\chi_{cr}(H)} + \frac{\eta}{4}\right)|R|\right\} \tag{3.16}$$

for all $V_{j_1} \neq V_{j_2} \in V(R)$ with $V_{j_1}V_{j_2} \notin E(R)$. Recall that by definition of B', each B'_i contains a perfect B^*-packing. Fix such a B^*-packing for each $i = 1, \dots, \ell'$. The union of all these B^*-packings gives us a perfect B^*-packing $\mathcal{B}^*$ in R.

Given any B'_i, it is easy to check that we can replace each cluster $V_j \in V(B'_i)$ with a subcluster of size $L' := (1 - \varepsilon|B'|)L$ such that for each edge $V_{j_1}V_{j_2}$ of B'_i the chosen subclusters of V_{j_1} and V_{j_2} form a $(2\varepsilon, d/2)$-super-regular pair in G'. (Indeed, this is just a generalisation of Fact 2.1.) We do this for each $i = 1, \dots, \ell'$ and add all the vertices not belonging to our chosen subclusters to V_0. We now refer to these subclusters as the clusters of R. Then for every edge $V_{j_1}V_{j_2}$ of R the pair $(V_{j_1}, V_{j_2})_{G'}$ is still 2ε-regular and has density more than $d/2$. Moreover,

$$|V_0| \leq 2\eta_1 n + \varepsilon|B'|n \leq 3\eta_1 n. \tag{3.17}$$

We now partition each cluster V_j into a red part V_j^{red} and a blue part V_j^{blue} where $\big||V_j^{red}| - |V_j^{blue}|\big| \leq \varepsilon L'$ and $\big||N_G(x) \cap V_j^{red}| - |N_G(x) \cap V_j^{blue}|\big| \leq \varepsilon L'$ for all $x \in V(G)$. (Consider a random partition to see that there are V_j^{red} and V_j^{blue} with these properties.) Together all these partitions of the clusters yield a partition of $V(G) - V_0$ into a set V^{red} of red vertices and a set V^{blue} of blue vertices. In Section 3.5.3 we will choose certain copies of H in G to cover the exceptional vertices in V_0, but each of these copies will avoid the red vertices. All the vertices contained in these copies of H will be removed from the clusters they belong to. However, for every edge $V_{j_1}V_{j_2}$ of B_i' the modified bipartite subgraph of G' whose vertex classes are the remainders of V_{j_1} and V_{j_2} will still be $(5\varepsilon, d/5)$-super-regular since it still contains all vertices in $V_{j_1}^{red} \cup V_{j_2}^{red}$. Furthermore, all edges in R will still correspond to 5ε-regular pairs of density more than $d/5$. After Section 3.5.3 we will only remove a bounded number of further vertices from the clusters, which will not affect the super-regularity significantly.

3.5.3 Incorporating the exceptional vertices

In this section we cover all the exceptional vertices with vertex-disjoint copies of H. Let G^{blue} denote the induced subgraph of G with vertex set $V^{blue} \cup V_0$. The definition of V^{blue}, (3.5) and (3.17) together imply that

$$d_{G^{blue}}(x) + d_{G^{blue}}(y) \geq \max\left\{2\left(1 - \frac{1}{r - \frac{2}{m+2}} + \frac{\eta}{2}\right)|G^{blue}|, 2\left(1 - \frac{1}{\chi_{cr}(H)} + \frac{\eta}{2}\right)|G^{blue}|\right\}$$

for all non-adjacent $x \neq y \in V(G^{blue})$. Let $v_1, \ldots, v_{|V_0|}$ be an enumeration of the exceptional vertices. Lemma 3.13 gives us a copy H_{v_1} of H in G^{blue} covering v_1. Delete the vertices of H_{v_1} from G^{blue} and apply the lemma again to find a copy H_{v_2} of H covering v_2. We would like to continue this way. However, for later purposes it is convenient to be able to assume that from each cluster we only delete a small proportion of vertices during this process. So before choosing the copy H_{v_j} for v_j (say), we call B'_i *bad* if it contains a cluster meeting the copies $H_{v_1}, \dots, H_{v_{j-1}}$ that we have chosen before in at least $\beta L'$ vertices. So at most $|V_0||H|/(\beta L') \leq 3\eta_1|H|n/(\beta L') \leq \eta\ell'/10$ of the B'_i are bad. We delete all the vertices belonging to clusters in bad B'_i from G^{blue}. Since there are at most $\eta n/10 \leq \eta|G^{blue}|/4$ such vertices, we can still apply Lemma 3.13 to find H_{v_j}. Thus we can cover all the exceptional vertices. We remove all the vertices lying in the copies $H_{v_1}, \dots, H_{v_{|V_0|}}$ of H from the clusters they belong to (and from G).

3.5.4 Making the blow-up of each $B \in \mathcal{B}^*$ divisible by $|H|$

Given a subgraph $S \subseteq R$ we write $V_G(S)$ for the set of all those vertices of G that belong to a cluster in S. Our aim now is to find, for each B'_i in our B'-packing in R, an H-packing in G covering all the vertices in $V_G(B'_i)$. Thus, taking the union of these H-packings and the copies of H containing the vertices in V_0, we will obtain a perfect H-packing in G. If we can ensure that the complete r-partite graph whose jth vertex class is the union of all clusters in the jth vertex class of B'_i has a perfect H-packing, then by the Blow-up lemma the subgraph of G' corresponding to B'_i will have a perfect H-packing. By Lemma 3.15 the former will turn out to be the case provided that $|H|$ divides $|V_G(B'_i)|$. So our next aim is to remove a bounded number of copies of H from G to ensure that $|V_G(B'_i)|$ is divisible by $|H|$ for all $i = 1, \dots, \ell'$. This in turn will be achieved by ensuring that $|H|$ divides $|V_G(B)|$ for all $B \in \mathcal{B}^*$.

Consider the auxiliary graph F whose vertices are the elements of $\mathcal{B}^*$ where $B_1, B_2 \in \mathcal{B}^*$ are adjacent in F if R contains a copy of K_r with one vertex in B_1 and $r-1$ vertices in B_2 or vice versa.

Suppose first that F is connected. Consider a spanning tree T of F with root $B_0 \in \mathcal{B}^*$, say. If B_1, $B_2 \in \mathcal{B}^*$ are adjacent in F then by the Embedding lemma G contains a copy of H with one vertex in $V_G(B_1)$ and all the other vertices in $V_G(B_2)$, or vice versa. (To see this, let K'_r be a copy of K_r in R with one vertex $V \in V_R(B_1)$ and all other vertices in $V_R(B_2)$. Choose any $V' \in V_R(B_2)$ which is adjacent to all of $V(K'_r) \setminus \{V\}$. Then our copy of H will have one vertex, v say, in V. All other vertices of H lying in the same colour class as v will be embedded into V' and all the remaining vertices of H will be embedded into $V(K'_r)\setminus\{V\}$.) In fact, we can choose $|H|-1$ disjoint such copies of H. So by removing at most $|H|-1$ such copies of H we can ensure $|V_G(B_1)|$ is divisible by $|H|$.

We can use this observation to 'shift the remainders mod $|H|$' along T to achieve that $|H|$ divides $|V_G(B)|$ for all $B \in \mathcal{B}^*$ as follows. Let j_{max} be the largest distance of some $B \in \mathcal{B}^*$ from B_0 in T. Then for all $B \in \mathcal{B}^*$ of distance j_{max} from B_0 we can remove copies of H as indicated above to ensure that $|H|$ divides $|V_G(B)|$. We can repeat this for all those $B \in \mathcal{B}^*$ of distance $j_{max}-1$ from B_0 etc. until $|V_G(B)|$ is divisible by $|H|$ for all $B \in \mathcal{B}^*$. (This follows as $\sum_{B\in\mathcal{B}^*}|V_G(B)|$ is divisible by $|H|$ since $|G|$ is divisible by $|H|$.)

So we may assume that F is not connected. Let $\mathcal{C}$ denote the set of all components of F. Given $C \in \mathcal{C}$, we denote by $V_R(C) \subseteq V(R)$ the set of all those clusters which belong to some $B \in \mathcal{B}^*$ with $B \in C$. We write $V_G(C) \subseteq V(G)$ for the union of all the clusters in $V_R(C)$. We will show that we can remove a bounded number of copies of H from G to achieve that $|V_G(C)|$ is divisible by $|H|$ for all $C \in \mathcal{C}$. As in the case when F is connected, we can then 'shift the remainders mod $|H|$' along a spanning tree of each component to make $|V_G(B)|$ divisible by $|H|$ for all $B \in \mathcal{B}^*$.

In the case when $r=2$ this is straightforward. Indeed, in this case H contains an isolated vertex (since $CE(H) < \infty$). So given any $C \in \mathcal{C}$ we can apply the Embedding lemma to find $|H|-1$ vertex-disjoint copies of H in G such that one vertex (playing the role of the isolated vertex) lies in $V_G(C)$ and the other vertices lie in $V_G(C')$ for some $C' \in \mathcal{C}\setminus\{C\}$. By removing a suitable number of such copies we can ensure that $|H|$ divides $|V_G(C)|$. Since in the above argument we can choose any $C' \in \mathcal{C}\setminus\{C\}$ to contain the remaining vertices of our copy of H (and since $|G|$ is divisible by $|H|$) we can apply this argument repeatedly to make $|V_G(C'')|$ divisible by $|H|$ for all $C'' \in \mathcal{C}$.

So now we consider the case when $r \geq 3$. We need the following claim.

Claim 3.16 *Let $C_1, C_2 \in \mathcal{C}$ and let $V \in V_R(C_2)$. Then*

$$|N_R(V) \cap V_R(C_1)| < \left(1 - \frac{1}{r-1+\gamma}\right)|V_R(C_1)|.$$

Proof. Suppose not. Then there exists some $B \in \mathcal{B}^*$ such that $B \in C_1$ and

$$|N_R(V) \cap B| \geq \left(1 - \frac{1}{r-1+\gamma}\right)|B| = |B| - \frac{(r-1)z + z_1}{r-1+z_1/z} = |B| - z.$$

Hence V has a neighbour in at least $r-1$ vertex classes of B. So R contains a copy of K_r with one vertex, namely V, in a copy $B_0 \in \mathcal{B}^*$ and $r-1$ vertices in B. So B and B_0 are adjacent in F. But they lie in different components of F, a contradiction. $\square$

We now show that we can remove a bounded number of copies of H from G to make $|V_G(C)|$ divisible by $|H|$ for *some* $C \in \mathcal{C}$. (In particular, if F consists of exactly two components C and C' this also ensures that $|V_G(C')|$ is divisible by $|H|$.)

Claim 3.17 *There exists a component $C \in \mathcal{C}$ with $|V_R(C)| \leq |R|/2$ for which we can ensure that $|H|$ divides $|V_G(C)|$ by removing at most $|H|-1$ copies of H from G.*

Proof. To prove the claim we will distinguish two cases.

Case 1. *There exists a component $C_1 \in \mathcal{C}$ with $|V_R(C_1)| \leq |R|/2$ and such that there is a cluster $V_1 \in V_R(C_1)$ with $d_R(V_1) \geq (1 - 1/\chi_{cr}(H) + \eta/4)|R|$.*

Recall that K_{r+1}^- is a K_{r+1} with one edge removed. We call the two non-adjacent vertices of K_{r+1}^- *small*. We say that a copy K' of K_{r+1}^- in R is *good* if either (i) $V(K') \cap V_R(C_1)$ consists of a small vertex of K' or (ii) $V(K') \setminus V_R(C_1)$ consists of a small vertex of K'. Once we have found a good K', we can use the Embedding lemma to find at most $|H|-1$ vertex-disjoint copies of H in G such that their removal from G ensures that $|V_G(C_1)|$ is divisible by $|H|$, as desired. (In case (i) precisely one vertex in each of these copies of H lies in $V_G(C_1)$ while in case (ii) precisely $|H|-1$ vertices in each of these copies of H lies in $V_G(C_1)$.) So it suffices to find a good copy of K_{r+1}^-.

Let S denote the set of neighbours of V_1 outside $V_R(C_1)$ in R. Let K be the set of vertices $V \in S$ with $d_R(V) < (1 - 1/\chi_{cr}(H) + \eta/4)|R|$. By (3.16), K induces a clique in R. If $|K| \geq r$, then we have a found a good copy of K_{r+1}^- (consisting of V_1 and r vertices of K). So we may assume that $|K| < r$.

Since $r \geq 3$ we have that $d_R(V_1) \geq (1/2 + \eta/4)|R|$. So $|S \setminus K| \geq \eta|R|/4 - r > 0$. Thus we can choose $V_2 \in S \setminus K$. By (3.14) the number of common neighbours of V_1 and V_2 in R is at least

$$\left(1 - \frac{2}{r-1+\gamma} + \frac{\eta}{4}\right)|R|. \tag{3.18}$$

We first consider the case when at least $(1 - \frac{2}{r-1+\gamma} + \frac{\eta}{4})|V(R) \setminus V_R(C_1)|$ common neighbours of V_1 and V_2 lie outside $V_R(C_1)$. We claim that we can find $V_3, \dots, V_r \in S \setminus K$ which form a K_r with V_1 and V_2. Suppose that we have found $V_3, \dots, V_i$ where $2 \leq i \leq r-1$. Note that Claim 3.16 and the definition of S imply that for $j \geq 2$ the number of neighbours of V_j outside $V_R(C_1)$ is at least $(1 - 1/(r-1+\gamma))|V(R) \setminus V_R(C_1)|$. Together with (3.18), this implies that the common neighbourhood of $V_1, \dots, V_i$ outside $V_R(C_1)$ has size at least

$$\left(1 - \frac{i}{r-1+\gamma} + \frac{\eta}{4}\right)|V(R) \setminus V_R(C_1)| \geq \frac{\eta}{4}|V(R) \setminus V_R(C_1)| > r > |K|. \tag{3.19}$$

This shows that we can find V_{i+1} and more generally $V_3, \dots, V_r$ as required. A similar calculation as in (3.19), shows that the common neighbourhood of $V_2, \dots, V_r$ outside $V_R(C_1)$ is non-empty and so contains some vertex V_{r+1} say. Together with $V_1, \dots, V_r$, V_{r+1} forms a good copy of K^-_{r+1}.

Now consider the case when at least $(1 - \frac{2}{r-1+\gamma} + \frac{\eta}{4})|V_R(C_1)|$ common neighbours of V_1 and V_2 lie inside $V_R(C_1)$. Since $\eta|V_R(C_1)|/4 \geq \eta|B^*|/4 \geq r$ we can argue as in the previous case. Indeed, this time we choose $V_3, \dots, V_r$ inside $V_R(C_1)$ to obtain a copy of K_r in R with one vertex, namely V_2, outside $V_R(C_1)$. We also choose a vertex V_{r+1} inside $V_R(C_1)$ that is adjacent to V_1, $V_3, \dots, V_r$. Again, $V_1, \dots, V_{r+1}$ form a good copy of K^-_{r+1}.

Case 2. *Every component $C \in \mathcal{C}$ with $|V_R(C)| \leq |R|/2$ is such that $d_R(V) < (1 - 1/\chi_{cr}(H) + \eta/4)|R|$ for all $V \in V_R(C)$.*

Together with (3.16) this implies that $V_1V_2 \in E(R)$ for all $V_1 \in V_R(C_1)$, $V_2 \in V_R(C_2)$ where $C_1, C_2 \in \mathcal{C}$ are such that $|V_R(C_1)|, |V_R(C_2)| \leq |R|/2$. But this means that there is only one component $C' \in \mathcal{C}$ with $|V_R(C')| \leq |R|/2$. So F consists of precisely two components C' and C'' where $V_R(C')$ forms a clique in R and $|V_R(C'')| > |R|/2$.

We first consider the case when $r = 3$. Note that R contains an edge between $V_R(C')$

and $V_R(C'')$. Indeed, if not then for any $V' \in V_R(C')$ and $V'' \in V_R(C'')$ by (3.16) we have that $d_R(V') + d_R(V'') \geq 2(1 - 1/\chi_{cr}(H) + \eta/4)|R| > |R|$ and so there must be an edge from V' to $V_R(C'')$ or from V'' to $V_R(C')$, a contradiction.

So since $|V_R(C')| \geq |B^*| \geq r + m$ we have a copy K'_{r+m} of K_{r+m} in $V_R(C')$ such that there is a cluster $V'' \in V_R(C'')$ adjacent to one of the clusters, V' say, of K'_{r+m}. Using the definition of m and the Embedding lemma we can find at most $|H| - 1$ copies of H in G each containing precisely one vertex in $V_G(C'')$ such that their removal ensures that $|H|$ divides $|V_R(C')|$ and thus also $|V_R(C'')|$. (Indeed, by definition of m there exists a vertex y of H such that $\chi(H[N(y)]) = r - 2 = 1$ and such that some 1-colouring of $N(y)$ can be extended to an $(r + m)$-colouring of H. So in our copies of H the vertex y will lie in V'', $N(y)$ will lie in V' and the remaining vertices of H will lie in $V(K'_{r+m})$.)

Now suppose that $r \geq 4$. We claim that there exists $V'' \in V_R(C'')$ which sends at least r edges to $V_R(C')$ in R. Suppose not. Then no $V \in V_R(C'')$ is joined to all of $V_R(C')$. Together with the definition of C' and (3.16) this implies that $d_R(V) \geq (1 - 1/\chi_{cr}(H) + \eta/4)|R|$. But then $|V_R(C')| < |R|/\chi_{cr}(H)$ since otherwise V is joined to $\eta|R|/4 \geq r$ vertices in $V_R(C')$. By assumption there are less than $r|V_R(C'')| < r|R|$ edges between $V_R(C')$ and $V_R(C'')$ in R. Moreover, by (3.16) and since $|V_R(C')| < |R|/\chi_{cr}(H)$ every cluster in $V_R(C')$ sends at least $(1 - 3/\chi_{cr}(H) + \eta/4)|R| > \eta|R|/4$ edges to $V_R(C'')$. So $\eta|R||V_R(C')|/4 < r|R|$. But $|V_R(C')| \geq |B^*| \geq 4r/\eta$ by definition of B^* and so $\eta|R||V_R(C')|/4 \geq r|R|$, a contradiction. So indeed there exists a vertex $V'' \in V_R(C'')$ sending at least r edges to $V_R(C')$. As before, we can remove at most $|H| - 1$ copies of H from G to ensure that $|H|$ divides both $|V_R(C')|$ and $|V_R(C'')|$. □

Claim 3.18 *We can make $|V_G(B)|$ divisible by $|H|$ for all $B \in \mathcal{B}^*$ by removing at most $|\mathcal{B}^*||H|$ copies of H from G.*

Proof. Our first aim is to take out some copies of H in G to achieve that $|V_G(C)|$ is divisible by $|H|$ for each $C \in \mathcal{C}$. We apply Claim 3.17 to remove at most $|H| - 1$ copies of H from G to ensure that $|V_G(C_1)|$ is divisible by $|H|$ for some component $C_1 \in \mathcal{C}$ with $|V_R(C_1)| \leq |R|/2$. Next we consider the graphs $F_1 := F - V(C_1)$ and $R_1 := R - V_R(C_1)$

instead of F and R. Claim 3.16 and (3.16) together imply that

$$d_{R_1}(V_{j_1}) + d_{R_1}(V_{j_2}) \geq 2\left(1 - \frac{1}{r-1+\gamma} + \frac{\eta}{4}\right)|R_1|$$

for all $V_{j_1} \neq V_{j_2} \in V(R_1)$ with $V_{j_1}V_{j_2} \notin E(R_1)$. Now suppose that $|\mathcal{C}| \geq 3$. Then similarly as in the proof of Claim 3.17 we can find a component $C_2 \in \mathcal{C}$ with $|V_R(C_2)| \leq |R_1|/2$ and such that by removing at most $|H|-1$ copies of H from G we ensure that $|H|$ divides $|V_G(C_2)|$. As $|G|$ was divisible by $|H|$ we can continue in this fashion to achieve that $|V_G(C)|$ is divisible by $|H|$ for each $C \in \mathcal{C}$.

During this process we have to take out at most $(|\mathcal{C}|-1)(|H|-1)$ copies of H in G. Now consider each $C \in \mathcal{C}$ separately. By proceeding as in the connected case for each C and taking out at most $(|C|-1)(|H|-1)$ further copies of H in each case, we can make $|V_G(B)|$ divisible by $|H|$ for all $B \in \mathcal{B}^*$. Hence, in total we have taken out at most $(|\mathcal{C}|-1)(|H|-1)+(|\mathcal{B}^*|-|\mathcal{C}|)(|H|-1) \leq |\mathcal{B}^*||H|$ copies of H. (Note that $|\mathcal{B}^*||H|$ is also an upper bound on the number of copies of H removed from G in the case when $r=2$.) □

3.5.5 Applying the Blow-up lemma

We now consider all the copies $B'_1, \dots, B'_{\ell'}$ of B' in the B'-packing of R, where the vertices of R are the modified clusters (i.e. they do not contain the vertices contained in the copies of H removed in Sections 3.5.3 and 3.5.4). For each $i \leq \ell'$ let G'_i denote the r-partite subgraph of G' whose jth vertex class is the union of all the clusters lying in the jth vertex class of B'_i (for $j = 1, \dots, r$). In Section 3.5.4 we made $|G'_i| = |V_G(B'_i)|$ divisible by $|H|$ for each i. Moreover, in Section 3.5.3 we removed at most $\beta L'$ vertices from each cluster. In Section 3.5.4 we removed only a bounded number of further vertices. So altogether we removed at most $2\beta L'$ vertices from each cluster. Since $\beta \ll \lambda \ll \gamma, 1-\gamma$ we may apply Lemma 3.15 to conclude that the complete r-partite graph whose vertex classes are the same as the vertex classes of G'_i has a perfect H-packing.

We observed at the end of Section 3.5.2 that the choice of those copies of H removed in Section 3.5.3 ensures that all the bipartite subgraphs corresponding to edges of B'_i are still $(5\varepsilon, d/5)$-super-regular. In Section 3.5.4 we only removed a bounded number of further vertices from each cluster. So after Section 3.5.4 the bipartite subgraphs of G'_i are still $(6\varepsilon, d/6)$-super-regular. Hence, for each $i = 1, \ldots, \ell'$, we may apply the Blow-up lemma to find a perfect H-packing in G'_i. All these H-packings together with the copies of H chosen previously form a perfect H-packing in G, as desired.

Chapter 4
Matchings in 3-uniform hypergraphs

4.1 Introduction

A *perfect matching* in a hypergraph H is a collection of vertex-disjoint edges of H which cover the vertex set $V(H)$ of H. A theorem of Tutte [93] gives a characterisation of all those graphs which contain a perfect matching. On the other hand, the decision problem whether an r-uniform hypergraph contains a perfect matching is NP-complete for $r \geq 3$. (See, for example, [39] for complexity results in the area.) It is natural therefore to seek simple sufficient conditions that ensure a perfect matching in an r-uniform hypergraph.

Given an r-uniform hypergraph H and distinct vertices $v_1, \dots, v_\ell \in V(H)$ (where $1 \leq \ell \leq r-1$) we define $d_H(v_1, \dots, v_\ell)$ to be the number of edges containing each of $v_1, \dots, v_\ell$. The *minimum ℓ-degree* $\delta_\ell(H)$ of H is the minimum of $d_H(v_1, \dots, v_\ell)$ over all ℓ-element sets of vertices in H. Of these parameters the two most natural to consider are the *minimum vertex degree* $\delta_1(H)$ and the *minimum collective degree* or *minimum codegree* $\delta_{r-1}(H)$. Rödl, Ruciński and Szemerédi [80] determined the minimum codegree that ensures a perfect matching in an r-uniform hypergraph. This improved bounds given in [54, 79]. An r-partite version was proved by Aharoni, Georgakopoulos and Sprüssel [1].

Much less is known about minimum vertex degree conditions for perfect matchings in r-uniform hypergraphs H. Hàn, Person and Schacht [34] showed that the threshold in the case when $r = 3$ is $(1 + o(1))\frac{5}{9}\binom{|H|}{2}$. This improved an earlier bound given by Daykin and Häggkvist [21]. In this chapter we determine the threshold exactly, which answers a question from [34].

Theorem 4.1 *There exists an $n_0 \in \mathbb{N}$ such that the following holds. Suppose that H is a 3-uniform hypergraph whose order $n \geq n_0$ is divisible by 3. If*

$$\delta_1(H) > \binom{n-1}{2} - \binom{2n/3}{2}$$

then H has a perfect matching.

While finalising this thesis we learned from [78] that the same result was also announced recently by Szemerédi. The following example shows that the result is best possible: let H^* be the 3-uniform hypergraph whose vertex set is partitioned into two vertex classes V and W of sizes $2n/3 + 1$ and $n/3 - 1$ respectively and whose edge set consists precisely of all those edges with at least one endpoint in W. Then H^* does not have a perfect matching and $\delta_1(H) = \binom{n-1}{2} - \binom{2n/3}{2}$.

The example generalises in the obvious way to r-uniform hypergraphs. This leads to the following conjecture, which is implicit in several papers (see e.g. [34, 58]). Partial results were proved by Hàn, Person and Schacht [34] as well as Markström and Ruciński [65].

Conjecture 4.2 *For each integer $r \geq 3$ there exists an integer $n_0 = n_0(r)$ such that the following holds. Suppose that H is an r-uniform hypergraph whose order $n \geq n_0$ is divisible by r. If*

$$\delta_1(H) > \binom{n-1}{r-1} - \binom{(r-1)n/r}{r-1},$$

then H has a perfect matching.

It is also natural to ask about the minimum (vertex) degree which guarantees a matching of given size d. Bollobás, Daykin and Erdős [11] solved this problem for the case when d is small compared to the order of H. We state the 3-uniform case of their result here. The above hypergraph H^* with W of size $d - 1$ shows that the minimum degree bound is best possible.

Theorem 4.3 (Bollobás, Daykin and Erdős [11]) *Let $d \in \mathbb{N}$. If H is a 3-uniform hypergraph on $n > 54(d+1)$ vertices and*

$$\delta_1(H) > \binom{n-1}{2} - \binom{n-d}{2}$$

then H contains a matching of size at least d.

In this chapter we extend this result to the entire range of d. Note that Theorem 4.4 generalises Theorem 4.1, so it suffices to prove Theorem 4.4.

Theorem 4.4 *There exists an $n_0 \in \mathbb{N}$ such that the following holds. Suppose that H is a 3-uniform hypergraph on $n \geq n_0$ vertices, that $n/3 \geq d \in \mathbb{N}$ and that*

$$\delta_1(H) > \binom{n-1}{2} - \binom{n-d}{2}.$$

Then H contains a matching of size at least d.

It would be interesting to obtain analogous results (i.e. minimum degree conditions which guarantee a matching of size d) for r-uniform hypergraphs and for r-partite hypergraphs (some bounds are given in [21]).

The situation for ℓ-degrees where $1 < \ell < r-1$ is also still open. Pikhurko [74] showed that if $\ell \geq r/2$ and H is an r-uniform hypergraph whose order n is divisible by r then H has a perfect matching provided that $\delta_\ell(H) \geq (1/2 + o(1))\binom{n}{r-\ell}$. This result is best possible up to the $o(1)$-term. In [34], Hàn, Person and Schacht provided conditions on $\delta_\ell(H)$ that ensure a perfect matching in the case when $\ell < r/2$. These bounds were subsequently lowered by Markström and Ruciński [65]. See [78] for further results concerning perfect matchings in hypergraphs.

4.2 Notation

Given a hypergraph H and subsets V_1, V_2, V_3 of its vertex set $V(H)$, we say that an edge $v_1v_2v_3$ is of *type* $V_1V_2V_3$ if $v_1 \in V_1$, $v_2 \in V_2$ and $v_3 \in V_3$.

Let $d \le n/3$ and let V, W be a partition of a set of n vertices such that $|W| = d$. Define $H_{n,d}(V, W)$ to be the hypergraph with vertex set $V \cup W$ consisting of all those edges which have type VVW or VWW. Thus $H_{n,d}(V, W)$ has a matching of size d,

$$\delta_1(H_{n,d}(V, W)) = \binom{n-1}{2} - \binom{n-d-1}{2}$$

and $H_{n,d}(V, W)$ is very close to the extremal hypergraph which shows that the degree condition in Theorem 4.4 is best possible. V and W are the *vertex classes of* $H_{n,d}(V, W)$.

Given $\varepsilon > 0$, a 3-uniform hypergraph H on n vertices and a partition V, W of $V(H)$ with $|W| = d$, we say that H is *ε-close to* $H_{n,d}(V, W)$ if

$$|E(H_{n,d}(V, W)) \setminus E(H)| \le \varepsilon n^3.$$

In this case we also call V and W *vertex classes of* H. (So H does not have unique vertex classes.) We say that H is *ε-close to* $H_{n,d}$ if there is a partition V, W of $V(H)$ such that $|W| = d$ and H is ε-close to $H_{n,d}(V, W)$.

Given a vertex v of a 3-uniform hypergraph H, we write $N_H(v)$ for the *neighbourhood of* v, i.e. the set of all those (unordered) tuples of vertices which form an edge together with v. Given two disjoint sets $A, B \subseteq V(H)$, we define the *link graph* $L_v(A, B)$ *of* v *with respect to* A, B to be the bipartite graph whose vertex classes are A and B and in which $a \in A$ is joined to $b \in B$ if and only if $ab \in N_H(v)$. Similarly, given a set $A \subseteq V(H)$, we define the *link graph* $L_v(A)$ *of* v *with respect to* A to be the graph whose vertex set is A and in which $a, a' \in A$ are joined if and only if $aa' \in N_H(v)$. Also, given disjoint sets $A, B, C, D, E \subseteq V(H)$, we write $L_v(ABCD)$ for $L_v(A, B) \cup L_v(B, C) \cup L_v(C, D)$. We define $L_v(ABCDE)$ similarly. If M is a matching in H and E, F are two edges in M with $v \notin E, F$, we write $L_v(EF)$ for $L_v(V(E), V(F))$. If $E_1, \dots, E_5$ are matching edges avoiding v, we define $L_v(E_1 \dots E_4)$ and $L_v(E_1 \dots E_5)$ similarly. If $e = uw$ is an edge in the link graph of v, then we write ve for the edge vuw of H. A matching in H of size d is called a *d-matching.*

Given a set M and $k \ge 2$, we write $\binom{M}{k}$ for the set of all k-element subsets of M. Given sets M and M', we write MM' for the set of all pairs mm' with $m \in M$ and $m' \in M'$. Given two graphs G and G', we write $G \cong G'$ if they are isomorphic. A bipartite graph is called *balanced* if its vertex classes have equal size.

4.3 Preliminaries and outline of proof

Our approach towards Theorem 4.4 follows the so-called *stability approach*: we prove an approximate version of the desired result which states that the minimum degree condition implies that either (i) H contains a d-matching or (ii) H is 'close' to the extremal hypergraph. The latter implies that H is 'close' to the hypergraph $H_{n,d}$ defined in the previous section. This extremal situation (ii) is then dealt with separately. We do this in Section 4.4, where we prove Lemma 4.7. The proof of Lemma 4.7 makes use of Theorem 4.3.

The non-extremal case is proved in Section 4.5. As mentioned earlier, an approximate version of Theorem 4.1 was proved in [34]. However, we need to proceed somewhat differently as the argument in [34] fails to guarantee the 'closeness' of H to the extremal hypergraph in case (ii). (But we do use the same general approach and a number of ideas from [34].)

We begin by considering a matching M of maximum size and suppose that $|M| < d$. We then carry out a sequence of steps, where in each step we show that we can either find a larger matching (and thus obtain a contradiction), or show that H is successively 'closer' to $H_{n,d}$. Amongst others, the following fact from [34] will be used to achieve this (see Figure 4.1 for the definitions of $B_{033}, B_{023}, B_{113}$).

Fact 4.5 *Let B be a balanced bipartite graph on 6 vertices.*

- *If $e(B) \geq 7$ then B contains a perfect matching.*
- *If $e(B) = 6$ then either B contains a perfect matching or $B \cong B_{033}$.*
- *If $e(B) = 5$ then either B contains a perfect matching or $B \cong B_{023}, B_{113}$.*

We call the vertices of degree 3 in B_{113} the *base vertices* of B_{113} and the edge between them the *base edge* of B_{113}.

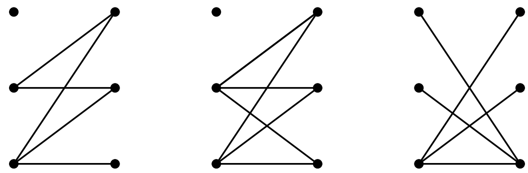

Figure 4.1: The graphs B with $e(B) \geq 5$ and no perfect matching

To see how the above fact can be used, suppose for example that x_1, x_2 and x_3 are unmatched vertices, that E and F are edges in M and that the link graphs $L_{x_i}(EF)$ are identical (call this graph B). The minimum degree condition implies that, for almost all unmatched vertices x, we have $e(L_x(EF)) \geq 5$. So let us assume this holds for x_1, x_2, x_3. If B contains a perfect matching, it is easy to see that we can transform M into a (larger) matching which also covers the x_i. If $B = B_{113}$, we can use this to prove that we are 'closer' to $H_{n,d}$. In particular, note that if $H = H_{n,d}$, then in the above example we have $B = B_{113}$. If $B \cong B_{023}, B_{033}$, we need to consider link graphs involving more than 2 edges from M in order to gain further information.

To find a matching which is larger than M, we will often need several vertices whose link graphs with respect to some set of matching edges are identical (as in the above example). We can usually achieve this with a simple application of the pigeonhole principle. But for this to work, we need to be able to assume that the number of vertices not covered by M is fairly large. This may not be true if e.g. we are seeking a perfect matching. To overcome this problem, we apply the 'absorbing method' which was first introduced in [80]. The method (as used in [34]) guarantees the existence of a small matching M^* which can 'absorb' any (very) small set of leftover vertices V' into a matching covering all of $V' \cup V(M^*)$. (The existence of M^* is shown using a probabilistic argument.) So if we are seeking e.g. a perfect matching, it suffices to prove the existence of an almost perfect one outside M^*. In particular, we can always assume that the set of vertices not covered by M is reasonably large, as otherwise we are done by the following lemma.

Lemma 4.6 (Hàn, Person and Schacht [34]) *Given any $\gamma > 0$ there exists an integer*

$n_0 = n_0(\gamma)$ such that the following holds. Suppose that H is a 3-uniform hypergraph on $n \geq n_0$ vertices such that $\delta_1(H) \geq (1/2 + 2\gamma)\binom{n}{2}$. Then there is a matching M^* in H of size $|M^*| \leq \gamma^3 n/3$ such that for every set $V' \subseteq V(H)\backslash V(M^*)$ with $\gamma^6 n \geq |V'| \in 3\mathbb{Z}$ there is a matching in H covering precisely the vertices in $V(M^*) \cup V'$.

4.4 Extremal case

The aim of this section is to show that hypergraphs which satisfy the degree condition in Theorem 4.4 and are close to $H_{n,d}$ contain a d-matching.

Lemma 4.7 *There exist $\varepsilon > 0$ and $n_0 \in \mathbb{N}$ such that the following holds. Suppose that H is a 3-uniform hypergraph on $n \geq n_0$ vertices and $d \leq n/3$ is an integer. If*

- *$\delta_1(H) > \binom{n-1}{2} - \binom{n-d}{2}$ and*
- *H is ε-close to $H_{n,d}$,*

then H contains a d-matching.

We will first prove the lemma in the case when H is not only close to $H_{n,d}$, but when for every vertex v most of the edges of $H_{n,d}$ incident to v also lie in H. More precisely, given $\alpha > 0$ and a 3-uniform hypergraph H on the same vertex set $V(H)$ as $H_{n,d}$, we say that a vertex $v \in V(H)$ is *α-bad* if $|N_{H_{n,d}}(v)\backslash N_H(v)| > \alpha n^2$. Otherwise we say that v is *α-good.* So if v is α-good then all but at most αn^2 of the edges incident to v in $H_{n,d}$ also lie in H. We will now show that if $d \geq n/150$ then any such H contains a d-matching.

Lemma 4.8 *Let $0 < \alpha < 10^{-6}$ and let $n, d \in \mathbb{N}$ be such that $n/150 \leq d \leq n/3$. Suppose that H is a 3-uniform hypergraph on the same vertex set as $H_{n,d}$ and every vertex of H is α-good. Then H contains a d-matching.*

Proof. Let V and W denote the vertex classes of $H_{n,d}$ of sizes $n-d$ and d respectively. Consider the largest matching M in H which consists entirely of edges of type VVW. Let V' denote the set of vertices in V uncovered by M. Define W' similarly. For a contradiction we assume that $|M| < d$. First note that $|M| \geq n/4$. Indeed, to see this consider any vertex $w \in W'$. Since w is α-good but $N_H(w) \cap \binom{V'}{2} = \emptyset$, it follows that $|V'| \leq 2\sqrt{\alpha}n$. Thus $|M| = |V \setminus V'|/2 \geq (n - d - 2\sqrt{\alpha}n)/2 \geq n/4$.

Consider $v_1, v_2 \in V'$ and $w \in W'$ where $v_1 \neq v_2$. Given a pair e_1e_2 of distinct matching edges from M, we say that e_1e_2 is *good for v_1v_2w* if there are all possible edges e in H which take the following form: e has type VVW and contains one vertex from $\{v_1, v_2, w\}$, one vertex from e_1 and one vertex from e_2. Note that if e_1e_2 is good for v_1v_2w then H has a 3-matching which consists of edges of type VVW and contains precisely the vertices in e_1, e_2 and $\{v_1, v_2, w\}$. So if such a pair e_1e_2 exists, we obtain a matching in H that is larger than M, yielding a contradiction.

Since $|M| \geq n/4$ we have at least $\binom{n/4}{2} > n^2/40$ pairs of distinct matching edges $e_1, e_2 \in M$. Since v_1, v_2 and w are α-good there are at most $3\alpha n^2 < n^2/40$ such pairs e_1e_2 that are not good for v_1v_2w. So one such pair must be good for v_1v_2w, a contradiction.

□

We now use Lemma 4.8 to prove Lemma 4.7. Our strategy is to obtain a 'small' matching M in H that covers all 'bad' vertices in H. We will construct M in stages so as to ensure that $H - V(M)$ satisfies the hypothesis of Lemma 4.8. Thus we obtain a $(d - |M|)$-matching M' of $H - V(M)$, and hence a d-matching $M \cup M'$ of H.

Proof of Lemma 4.7. Let $0 < 1/n_0 \ll \varepsilon \ll \varepsilon' \ll \varepsilon'' \ll \varepsilon''' \ll 1$. By Theorem 4.3 we may assume that $d \geq n/100$. Suppose that H is as in the statement of the lemma and let V and W denote the vertex classes of H of sizes $n - d$ and d respectively. Since H is ε-close to $H_{n,d}$, all but at most $3\sqrt{\varepsilon}n$ vertices in H are $\sqrt{\varepsilon}$-good. Let V^{bad} denote the set of $\sqrt{\varepsilon}$-bad vertices in V. Define W^{bad} similarly. So $|V^{bad}|, |W^{bad}| \leq 3\sqrt{\varepsilon}n$.

Define $c := |W^{bad}|$, $V_1 := V \cup W^{bad}$ and $W_1 := W \backslash W^{bad}$. Thus $a := |V_1| = n - d + c$ and $b := |W_1| = d - c$. Moreover,

$$\delta_1(H[V_1]) \geq \delta_1(H) - \binom{b}{2} - (a-1)b > \binom{n-1}{2} - \binom{n-d}{2} - \binom{b}{2} - (a-1)b.$$

But $\binom{n-1}{2} = \binom{a-1}{2} + (a-1)b + \binom{b}{2}$ and so

$$\delta_1(H[V_1]) > \binom{a-1}{2} - \binom{n-d}{2} = \binom{a-1}{2} - \binom{a-c}{2}.$$

Since $c \leq 3\sqrt{\varepsilon}n$ we can apply Theorem 4.3 to obtain a matching M_1 of size c in $H[V_1]$.

Let $H_1 := H - V(M_1)$ and $V_2 := V_1 \backslash V(M_1)$. (Note that if $W^{bad} = \emptyset$ then $H_1 = H$.) So H_1 has vertex classes V_2 and W_1 where $|V_2| = a - 3c$. Since H is ε-close to $H_{n,d}(V,W)$ and $3c \leq 9\sqrt{\varepsilon}n \ll \varepsilon' n$ we have that H_1 is ε'-close to $H_{|H_1|,b}(V_2, W_1)$. By definition of W_1 all vertices in W_1 are ε'-good in H_1. Furthermore, if a vertex $v \in V(H_1)$ is ε'-bad in H_1 then $v \in V_2$ and $v \in V^{bad} \cup W^{bad}$. Let V_2^{bad} denote the set of such vertices. So $|V_2^{bad}| \leq 3\sqrt{\varepsilon}n$. If $V_2^{bad} = \emptyset$ then we can apply Lemma 4.8 to obtain a b-matching M_2 in H_1. We thus obtain a matching $M_1 \cup M_2$ of size $b + c = d$ in H . So we may assume that $V_2^{bad} \neq \emptyset$.

We say that a vertex $v \in V_2^{bad}$ is *useful* if there are at least $\varepsilon' n^2$ pairs of vertices $v'w \in V_2 W_1$ such that $vv'w$ is an edge in H_1. Clearly we can greedily select a matching M_2 in H_1 such that $m_2 := |M_2| \leq |V_2^{bad}|$ where M_2 covers all useful vertices and consists entirely of edges of type $V_2V_2W_1$. Let $H_2 := H_1 - V(M_2)$, $V_3 := V_2 \backslash V(M_2)$ and $W_2 := W_1 \backslash V(M_2)$. Then $|V_3| = |V_2| - 2m_2 = a - 3c - 2m_2$ and $|W_2| = b - m_2$. Note that

$$\begin{aligned}\delta_1(H) &> \binom{n-1}{2} - \binom{n-d}{2} \geq (1-\varepsilon)\left(1 - \left(1 - \frac{d}{n}\right)^2\right)\frac{n^2}{2} \\ &= (1-\varepsilon)\left(\frac{2d}{n} - \frac{d^2}{n^2}\right)\frac{n^2}{2} = (1-\varepsilon)d\left(n - \frac{d}{2}\right).\end{aligned} \tag{4.1}$$

Consider any vertex $v \in V_2^{bad} \backslash V(M_2)$. Since v is not useful, it must lie in more than

$$\begin{aligned}\delta_1(H) - n|V(H) \setminus V(H_2)| - \varepsilon' n^2 - \binom{|W_2|}{2} &\overset{(4.1)}{\geq} (1-\varepsilon)d\left(n - \frac{d}{2}\right) - \varepsilon' n^2 - \varepsilon' n^2 - \frac{d^2}{2} \\ &\geq d(n-d) - \varepsilon dn - 2\varepsilon' n^2 \geq \frac{2dn}{3} - 3\varepsilon' n^2 \geq 2\varepsilon' n^2\end{aligned}$$

edges of $H_2[V_3]$. Since $|V_2^{bad}| \leq 3\sqrt{\varepsilon}n$ we can greedily select a matching M_3 in $H_2[V_3]$ of size $m_3 := |M_3| \leq |V_2^{bad}|$ which covers all the vertices in H_2 which lie in V_2^{bad}.

Let $H_3 := H_2 - V(M_3)$ and $V_4 := V_3 \backslash V(M_3)$. So H_3 has vertex classes V_4 and W_2 where $|V_4| = |V_3| - 3m_3 = a - 3c - 2m_2 - 3m_3$. Recall that every vertex in $V(H_1) \setminus V_2^{bad}$ is ε'-good in H_1. Since $V_2^{bad} \subseteq V(M_2 \cup M_3)$ and $|H_1| - |H_3| = 3(|M_2| + |M_3|) \ll \varepsilon' n$, it follows that every vertex of H_3 is ε''-good. So certainly for every vertex $w \in W_2$ there are at least $|V_4||W_2|/2$ pairs $vw' \in V_4W_2$ such that vww' is an edge in H_3. Thus we can greedily find a matching M_4 of size m_3 such that each edge in M_4 has type $V_4W_2W_2$.

Let $H_4 := H_3 - V(M_4)$, $V_5 := V_4 \backslash V(M_4)$ and $W_3 := W_2 \backslash V(M_4)$. So H_4 has vertex classes V_5 and W_3 of sizes $|V_5| = |V_4| - m_3 = a - 3c - 2m_2 - 4m_3 = n - d - 2c - 2m_2 - 4m_3$ and $|W_3| = |W_2| - 2m_3 = b - m_2 - 2m_3 = d - c - m_2 - 2m_3$. Moreover, every vertex of H_4 is ε'''-good. Thus we can apply Lemma 4.8 to H_4 to obtain a $|W_3|$-matching M_5 in H_4. But then $M_1 \cup M_2 \cup M_3 \cup M_4 \cup M_5$ is a matching of size $c + m_2 + m_3 + m_3 + |W_3| = d$ in H, as desired. $\square$

We remark that the only point in the proof of Theorem 4.4 where we need the full strength of the minimum degree condition is when we apply Theorem 4.3 to find the matching M_1 in the proof of Lemma 4.7.

4.5 Proof of Theorem 4.4

4.5.1 Preliminaries

We first define constants satisfying

$$0 < 1/n_0 \ll 1/C \ll \gamma'' \ll \gamma' \ll \gamma \ll \varepsilon' \ll \varepsilon \ll \eta' \ll \eta \ll \alpha' \ll \alpha \ll \rho' \ll \rho \ll \tau \ll 1. \tag{4.2}$$

Let H be a 3-uniform hypergraph on $n \geq n_0$ vertices such that

$$\delta_1(H) > \binom{n-1}{2} - \binom{n-d}{2} \geq (1-\gamma')d(n-d/2), \tag{4.3}$$

where d is an integer such that $1 \leq d \leq n/3$. (Note that the second inequality in (4.3) follows from the same argument as (4.1).) We wish to find a d-matching in H. Note that

Theorem 4.3 covers the case when $d \leq n/100$. So we may assume that $n/100 \leq d \leq n/3$.

Suppose $d \geq n/3 - \tau n$. Since $\tau \ll 1$, (4.3) gives us that $\delta_1(H) \geq (1/2 + 2\gamma'')\binom{n}{2}$. So by Lemma 4.6 there is a matching M^* in H of size $|M^*| \leq (\gamma'')^3 n/3$ such that for every set $V' \subseteq V(H)\backslash V(M^*)$ with $(\gamma'')^6 n \geq |V'| \in 3\mathbb{Z}$ there is a matching in H covering precisely the vertices in $V(M^*) \cup V'$. If $n/100 \leq d < n/3 - \tau n$ we set $M^* := \emptyset$.

In both cases we define $H' := H - V(M^*)$. (So $H' = H$ if $n/100 \leq d < n/3 - \tau n$.) Thus

$$\delta_1(H') \geq \delta_1(H) - \gamma' n^2. \tag{4.4}$$

Let M be the largest matching in H'. Clearly we may assume that $|M| < d$. Theorem 4.3 implies that

$$n/200 \leq |M| < d. \tag{4.5}$$

Let $V_M := V(M)$ and $V_0 := V(H')\backslash V_M$. So $|V_0| \leq n - |V_M|$. If $n/100 \leq d < n/3 - \tau n$ then $|V_0| > n - 3d > 3\tau n$. Suppose $d \geq n/3 - \tau n$. If $|V_0| \leq (\gamma'')^6 n$, then by definition of M^*, there is a matching M' in H containing all but at most two vertices from $V(M^*) \cup V_0$. But then $M \cup M'$ is a matching in H of size $\lfloor n/3 \rfloor \geq d$, as desired. So in both cases we may assume that

$$(\gamma'')^6 n \leq |V_0| \leq n - |V_M|. \tag{4.6}$$

4.5.2 Finding structure in the link graphs

In this section we show that 'most' of our link graphs $L_v(EF)$ with $v \in V_0$ and $EF \in \binom{M}{2}$ are copies of B_{113} (recall that B_{113} was defined after Fact 4.5).

Claim 4.9 *There does not exist $v_1v_2v_3 \in \binom{V_0}{3}$ and $EF \in \binom{M}{2}$ such that*

- *$L_{v_1}(EF) = L_{v_2}(EF) = L_{v_3}(EF)$ and*
- *$L_{v_1}(EF)$ contains a perfect matching.*

Proof. The proof is identical to the proof of Fact 17 in [34]. We include it here for completeness. Let $E = \{x_1, x_2, x_3\}$ and $F = \{y_1, y_2, y_3\}$ and suppose x_1y_1, x_2y_2 and x_3y_3 is a perfect matching in $L_{v_1}(EF)$. Since these edges lie in $L_{v_i}(EF)$ for each $1 \leq i \leq 3$ the edges $v_1x_1y_1$, $v_2x_2y_2$ and $v_3x_3y_3$ lie in H'. Replacing E and F in M with these edges we obtain a larger matching in H', a contradiction. $\square$

We will now use Claim 4.9 to show that only a constant number of vertices $v \in V_0$ have 'many' link graphs $L_v(EF)$ containing perfect matchings.

Claim 4.10 *Let V_0' denote the set of all those vertices $v \in V_0$ for which there are at least εn^2 pairs $EF \in \binom{M}{2}$ such that $L_v(EF)$ contains a perfect matching. Then $|V_0'| \leq C$.*

Proof. Let G be the bipartite graph with vertex classes V_0' and $\binom{M}{2}$ where $\{v, EF\}$ is an edge in G precisely when $L_v(EF)$ contains a perfect matching. So G contains at least $|V_0'|\varepsilon n^2$ edges. If $|V_0'| \geq C$ then there is a pair $EF \in \binom{M}{2}$ such that $d_G(EF) \geq C\varepsilon \geq 3 \cdot 2^9$ (since $1/C \ll \varepsilon$). Since there are 2^9 labelled bipartite graphs with vertex classes E and F, there are 3 vertices $v_1, v_2, v_3 \in V_0'$ such that $L_{v_1}(EF) = L_{v_2}(EF) = L_{v_3}(EF)$ and $L_{v_1}(EF)$ contains a perfect matching. This contradicts Claim 4.9, as required. $\square$

Claim 4.11 *Let V_0'' denote the set of all those vertices $v \in V_0$ for which there are at least εn^2 pairs $EF \in \binom{M}{2}$ such that $L_v(EF) \cong B_{023}, B_{033}$. Then $|V_0''| \leq C$.*

Proof. Suppose for a contradiction that $|V_0''| > C$. Given any $v \in V_0''$, define an auxiliary oriented graph G_v as follows: The vertex set of G_v is M and given $EF \in \binom{M}{2}$ there is an edge directed from E to F precisely when $L_v(EF) \cong B_{023}, B_{033}$ where E is the vertex class that contains the isolated vertex in $L_v(EF)$. Since $v \in V_0''$, we have that $e(G_v) \geq \varepsilon n^2$.

We call a path $E_1 \dots E_5$ of length 4 in G_v *suitable* if its (directed) edges are E_1E_2, E_3E_2, E_3E_4 and E_5E_4. Our first aim is to find at least $\varepsilon' n^5$ suitable paths in G_v. Choose a partition V_1, V_2 of $V(G_v)$ such that $e_{G_v}(V_1, V_2) \geq e(G_v)/5 \geq \varepsilon n^2/5$. (To see the existence of such a partition, consider the expected number of edges from V_1 to V_2 in a random partition of $V(G_v)$.) Let G_v' denote the undirected bipartite graph with vertex classes V_1 and V_2 whose edges are all those edges in G_v that are oriented from V_1 to V_2. Since $e(G'_v) \geq \varepsilon n^2/5$, G'_v contains a subgraph G''_v with $\delta(G''_v) \geq d(G'_v)/2 \geq \varepsilon n/5$. Thus we can greedily find at least

$$\frac{1}{2} \cdot \frac{\varepsilon n}{5} \left(\frac{\varepsilon n}{5} - 1 \right) \cdots \left(\frac{\varepsilon n}{5} - 4 \right) \geq \varepsilon' n^5$$

paths of length 4 in G''_v whose endpoints both lie in V_1. By definition of G''_v, each of these paths corresponds to a suitable path in G_v.

Consider a suitable path $E_1 \dots E_5$ in G_v. So $L_v(E_2E_3), L_v(E_3E_4) \cong B_{023}, B_{033}$ with the isolated vertex in both graphs lying in E_3. Choose edges e_1 of $L_v(E_2E_3)$ and e_2 of $L_v(E_3E_4)$ such that e_1 and e_2 are disjoint. Since $L_v(E_1E_2) \cong B_{023}, B_{033}$ and E_1 contains the isolated vertex in this graph, there is a 2-matching $\{e_3, e_4\}$ in $L_v(E_1E_2)$ that is disjoint from e_1. Similarly since $L_v(E_4E_5) \cong B_{023}, B_{033}$ and E_5 contains the isolated vertex in this graph, there is a 2-matching $\{e_5, e_6\}$ in $L_v(E_4E_5)$ that is disjoint from e_2. Hence $L_v(E_1E_2E_3E_4E_5)$ contains a 6-matching $\{e_1, e_2, e_3, e_4, e_5, e_6\}$.

Let G be the bipartite graph with vertex classes V''_0 and the set $(M)^5$ of all ordered 5-tuples of elements of M where $\{v, E_1E_2E_3E_4E_5\}$ is an edge in G precisely when $E_1 \dots E_5$ is a suitable path in G_v. So G contains at least $|V''_0|\varepsilon' n^5$ edges.

Since $|V''_0| > C$ there exists $E_1E_2E_3E_4E_5 \in (M)^5$ such that $d_G(E_1E_2E_3E_4E_5) \geq C\varepsilon' \geq 6 \cdot 2^{36}$. Further, there are at most 2^{36} distinct graphs in the collection of all those graphs $L_v(E_1E_2E_3E_4E_5)$ for which $v \in N_G(E_1E_2E_3E_4E_5)$. Thus there are 6 vertices $v_1, \dots, v_6 \in V''_0$ such that $v_1, \dots, v_6 \in N_G(E_1E_2E_3E_4E_5)$ and $L_{v_1}(E_1E_2E_3E_4E_5) = \cdots = L_{v_6}(E_1E_2E_3E_4E_5)$. Let $\{x_1y_1, \dots, x_6y_6\}$ be a 6-matching in $L_{v_1}(E_1E_2E_3E_4E_5)$. So $\{v_1x_1y_1, \dots, v_6x_6y_6\}$ is a 6-matching in H'. Replacing the edges $E_1, \dots, E_5$ in M with $\{v_1x_1y_1, \dots, v_6x_6y_6\}$ we obtain a larger matching, a contradiction. □

Claim 4.12 *Let V'''_0 denote the set of all those vertices $v \in V_0$ which fail to satisfy*

$$e(L_v(V_0, V_M)) \leq (1 + \sqrt{\gamma'})|V_0||M|. \tag{4.7}$$

Then $|V'''_0| \leq C$.

Proof. Suppose for a contradiction that $|V_0'''| > C \geq 2/\gamma'$. Given an edge E in M, we say that E is *good for* $v \in V_0'''$ if at least two vertices in E have degree at least 3 in $L_v(E, V_0)$. For every $v \in V_0'''$, there are at least $\gamma'|M|$ edges in M which are good for v. (To see this, suppose there are fewer edges which are good for v. Then

$$\begin{aligned} e(L_v(V_0, V_M)) &< (1-\gamma')|M|(4+|V_0|) + \gamma'|M| \cdot 3|V_0| \\ &\leq |M||V_0| \left((1-\gamma')(1+\gamma') + 3\gamma'\right) \leq (1+\sqrt{\gamma'})|V_0||M|, \end{aligned}$$

a contradiction to the fact that $v \in V_0'''$.) This in turn implies that there are $v_1, v_2 \in V_0'''$ and an edge E in M which is good for both v_1 and v_2. Then the definition of 'good' implies that are disjoint edges $e_1 \in L_{v_1}(E, V_0)$ and $e_2 \in L_{v_2}(E, V_0)$ which do not contain v_1 or v_2. Now we can enlarge M by removing E and adding v_1e_1 and v_2e_2. This contradiction to the maximality of M proves the claim. □

Claim 4.13 *Every vertex $v \in V_0 \backslash V_0'''$ satisfies*

$$e(L_v(V_M)) \geq (5-\gamma)\binom{|M|}{2}.$$

Proof. Suppose $v \in V_0 \backslash V_0'''$. Then as $e(L_v(V_0)) = 0$

$$\begin{aligned} e(L_v(V_M)) &\overset{(4.4)}{\geq} \delta_1(H) - e(L_v(V_0, V_M)) - \gamma' n^2 \\ &\overset{(4.3),(4.7)}{\geq} (1-\gamma')d(n-d/2) - \left(1+\sqrt{\gamma'}\right)|V_0||M| - \gamma' n^2. \end{aligned}$$

Now note that the function $d(n-d/2)$ is increasing in d for $d \leq n/3$. So

$$\begin{aligned} e(L_v(V_M)) &\geq (1-\gamma')|M|\left(n - \frac{|M|}{2}\right) - \left(1+\sqrt{\gamma'}\right)(n-3|M|)|M| - \gamma' n^2 \\ &\geq \left(n|M| - \frac{|M|^2}{2} - \gamma' n|M|\right) - \left(n|M| - 3|M|^2 + \sqrt{\gamma'}n|M|\right) - \gamma' n^2 \\ &\overset{(4.5)}{\geq} \frac{5|M|^2}{2} - 400\sqrt{\gamma'}|M|^2 \geq (5-\gamma)\binom{|M|}{2}, \end{aligned}$$

which completes the proof of the claim. □

Claim 4.14 *Let V_0'''' denote the set of all those vertices $v \in V_0 \backslash V_0'''$ for which there are at least ηn^2 pairs $EF \in \binom{M}{2}$ such that $L_v(EF)$ contains at most 4 edges. Then $|V_0''''| \leq 2C$.*

Proof. Suppose for a contradiction that $|V_0''''| > 2C$. Let $v \in V_0''''$. At most $3|M|$ edges $e = vv_1v_2$ in H containing v are such that v_1 and v_2 lie in the same edge $E \in M$. Thus Claim 4.13 implies that

$$\sum_{EF \in \binom{M}{2}} e(L_v(EF)) \geq (5-\gamma)\binom{|M|}{2} - 3|M| \geq 5\binom{|M|}{2} - \gamma n^2. \tag{4.8}$$

Let c denote the number of pairs $EF \in \binom{M}{2}$ such that $L_v(EF)$ contains at most 4 edges. Then $c \geq \eta n^2$ and so (4.8) implies that there are at least $\eta' n^2$ pairs $EF \in \binom{M}{2}$ such that $L_v(EF)$ contains at least 6 edges. Indeed, suppose that this is not the case. Then

$$\begin{aligned}\sum_{EF \in \binom{M}{2}} e(L_v(EF)) &\leq 4c + 9\eta' n^2 + 5\left[\binom{|M|}{2} - c\right] = 5\binom{|M|}{2} - c + 9\eta' n^2 \\ &< 5\binom{|M|}{2} - \gamma n^2\end{aligned}$$

since $\gamma \ll \eta' \ll \eta$. This contradicts (4.8), as desired.

Recall from Fact 4.5 that a balanced bipartite graph B on 6 vertices that contains at least 6 edges either has a perfect matching or $B \cong B_{033}$. Thus, given any $v \in V_0''''$ there are at least $r \geq \eta' n^2/2 \geq \varepsilon n^2$ pairs $E_1F_1, \dots, E_rF_r \in \binom{M}{2}$ such that either

- $L_v(E_iF_i)$ contains a perfect matching for all $1 \leq i \leq r$ or,
- $L_v(E_iF_i) \cong B_{033}$ for all $1 \leq i \leq r$.

So since $|V_0''''| > 2C$ one of the following holds:

(α_1) There are more than C vertices $v \in V_0''''$ for which there are at least εn^2 pairs $EF \in \binom{M}{2}$ such that $L_v(EF)$ contains a perfect matching.

(α_2) There are more than C vertices $v \in V_0''''$ for which there are at least εn^2 pairs $EF \in \binom{M}{2}$ such that $L_v(EF) \cong B_{033}$.

In either case we get a contradiction: (α_1) contradicts Claim 4.10 and (α_2) contradicts Claim 4.11. $\square$

Recall from Fact 4.5 that if B is a balanced bipartite graph on 6 vertices with $e(B) = 5$ then either B contains a perfect matching or $B \cong B_{023}, B_{113}$. If $e(B) \geq 6$ then either B contains a perfect matching or $B \cong B_{033}$. Thus Claims 4.10, 4.11, 4.12 and 4.14 together imply that all vertices $v \in V_0 \setminus (V_0' \cup V_0'' \cup V_0''' \cup V_0'''')$ satisfy

(β) $L_v(EF) \cong B_{113}$ for at least $\binom{|M|}{2} - 2\varepsilon n^2 - \eta n^2 \geq (1-\alpha')\binom{|M|}{2}$ pairs $EF \in \binom{M}{2}$.

Let $V_0^* := V_0 \setminus (V_0' \cup V_0'' \cup V_0''' \cup V_0'''')$. Thus

$$|V_0 \setminus V_0^*| \leq 5C.$$

Moreover, each $v \in V_0^*$ satisfies

$$e(L_v(V_M)) \leq 5(1-\alpha')\binom{|M|}{2} + 9\alpha'\binom{|M|}{2} + 3|M| \leq 5(1+\alpha')\binom{|M|}{2}. \tag{4.9}$$

Here the term $3|M|$ accounts for the edges which have both endpoints in the same matching edge of M.

We can now show that M has almost the required size. This will be used in Section 4.5.3 to prove that H is close to $H_{n,d}$.

Claim 4.15 $|M| > d - \alpha n$.

Proof. Assume for a contradiction that $|M| \leq d - \alpha n$. Consider any $v \in V_0^*$. Then

$$d_{H'}(v) \overset{(4.3),(4.4)}{\geq} (1-\gamma')d(n-d/2) - \gamma' n^2 \geq d(n-d/2) - 2\gamma' n^2. \tag{4.10}$$

Also $e(L_v(V_0)) = 0$ since M is maximal. Thus

$$\begin{aligned} d_{H'}(v) &= e(L_v(V_M)) + e(L_v(V_0, V_M)) \overset{(4.7),(4.9)}{\leq} 5(1+\alpha')\binom{|M|}{2} + (1+\sqrt{\gamma'})|V_0||M| \\ &\leq 5(1+\alpha')\binom{|M|}{2} + \left(|M|(n-3|M|) + \sqrt{\gamma'}n^2\right) \\ &\leq |M|(n-|M|/2) + \sqrt{\alpha'}n^2 < (d-\alpha n)(n-d/2+\alpha n/2) + \sqrt{\alpha'}n^2 \\ &< d(n-d/2) - 2\gamma' n^2, \end{aligned}$$

a contradiction to (4.10), as desired. (In the third line we again used that the function $d(n-d/2)$ is increasing in d for $d \leq n/3$.) $\square$

In the next sequence of claims, we will show that there are vertices $v_1, \dots, v_{10} \in V_0^*$ whose link graphs $L_{v_i}(V_M)$ are very similar to each other (see Claim 4.19 for the precise statement).

Claim 4.16 *Suppose $v_1, \dots, v_{10} \in V_0^*$ are distinct vertices such that for some $EF \in \binom{M}{2}$, $L_{v_1}(EF), \dots, L_{v_{10}}(EF) \cong B_{113}$. Then $L_{v_1}(EF) = \cdots = L_{v_{10}}(EF)$.*

Proof. We suppose for a contradiction that the claim does not hold. Since there are 9 labelled bipartite graphs with vertex classes E and F which are isomorphic to B_{113}, two of the $L_{v_i}(EF)$ must be the same. So we may assume that $L_{v_1}(EF) = L_{v_2}(EF)$ but $L_{v_1}(EF) \neq L_{v_3}(EF)$. Let $E = \{x_1, x_2, x_3\}$ and $F = \{y_1, y_2, y_3\}$. Suppose $E(L_{v_1}(EF)) = E(L_{v_2}(EF)) = \{x_1y_1, x_1y_2, x_1y_3, x_2y_1, x_3y_1\}$. (So x_1y_1 is the base edge of $L_{v_1}(EF)$ and $L_{v_2}(EF)$ as defined after Fact 4.5.) Since $L_{v_1}(EF) \neq L_{v_3}(EF)$ there is an edge $e \in L_{v_3}(EF) \backslash L_{v_1}(EF)$. We may assume $e = x_3y_3$. Replacing E and F with $v_1x_1y_2, v_2x_2y_1$ and $v_3x_3y_3$ in M we obtain a larger matching, a contradiction. $\square$

Choose distinct $v_1, \dots, v_{10} \in V_0^*$ which will be fixed throughout the remainder of the proof.

Claim 4.17 *There is a set $\mathcal{E}$ of at least $(1-\alpha)|M|$ matching edges $E \in M$ such that for each $E \in \mathcal{E}$ there are at least $(1-\alpha)|M|$ edges $F \in M$ for which*

$$L_{v_1}(EF) = \cdots = L_{v_{10}}(EF) \cong B_{113}.$$

Proof. By (β) and Claim 4.16 there are at least $(1-10\alpha')\binom{|M|}{2}$ pairs $EF \in \binom{M}{2}$ such that $L_{v_1}(EF) = \cdots = L_{v_{10}}(EF) \cong B_{113}$. This in turn immediately implies the claim. $\square$

Claim 4.18 *For every $E \in \mathcal{E}$ there is a set $\mathcal{F}_E$ of at least $(1-2\alpha)|M|$ edges in M such that*

(δ_1) $L_{v_1}(EF) = \cdots = L_{v_{10}}(EF) \cong B_{113}$ for each $F \in \mathcal{F}_E$ and

(δ_2) in each of the $L_{v_1}(EF)$ with $F \in \mathcal{F}_E$ the same vertex x plays the role of the base vertex in E.

Proof. Since $E \in \mathcal{E}$ there is a set $\mathcal{F}'_E$ of at least $(1-\alpha)|M|$ edges in M such that $L_{v_1}(EF) = \cdots = L_{v_{10}}(EF) \cong B_{113}$ for each $F \in \mathcal{F}'_E$. Let $\mathcal{F}_E := \mathcal{F}'_E \cap \mathcal{E}$. Then $|\mathcal{F}_E| \geq (1-2\alpha)|M|$ and for each $F \in \mathcal{F}_E$ there are at least $(1-\alpha)|M|$ edges $F' \in M$ for which $L_{v_1}(FF') = \cdots = L_{v_5}(FF') \cong B_{113}$.

We claim that $\mathcal{F}_E$ satisfies the claim. Certainly $\mathcal{F}_E$ satisfies (δ_1). Suppose for a contradiction that there are $F_1, F_2 \in \mathcal{F}_E$ such that the vertex $x_1 \in E$ that plays the role of a base vertex in $L_{v_1}(EF_1)$ is different from the vertex $x_2 \in E$ that plays the role of a base vertex in $L_{v_1}(EF_2)$. Let $F' \in M$ be such that $L_{v_1}(F_2F') = \cdots = L_{v_5}(F_2F') \cong B_{113}$, and $F' \neq E, F_1$.

Since $L_{v_1}(EF_1) \cong B_{113}$ and $x_1 \neq x_2$, there exists a 2-matching $\{e_1, e_2\}$ in $L_{v_1}(EF_1)$ that is disjoint from x_2. Similarly since $L_{v_1}(F_2F') \cong B_{113}$ there exists a 2-matching $\{e_3, e_4\}$ in $L_{v_1}(F_2F')$. Since $x_2 \in E$ is a base vertex in $L_{v_1}(EF_2)$, there is an edge e_5 from x_2 to the vertex in F_2 that is uncovered by $\{e_3, e_4\}$. So $\{e_1, e_2, e_3, e_4, e_5\}$ is a 5-matching in $L_{v_1}(F_1EF_2F')$. We have chosen F_1, F_2 and F' so that $L_{v_1}(F_1EF_2F') = L_{v_2}(F_1EF_2F') = \cdots = L_{v_5}(F_1EF_2F')$. Thus $M' := \{v_1e_1, v_2e_2, v_3e_3, v_4e_4, v_5e_5\}$ is a 5-matching in H' that contains only vertices from $E \cup F' \cup F_1 \cup F_2 \cup \{v_1, v_2, v_3, v_4, v_5\}$. Replacing E, F', F_1 and F_2 in M with the edges in M' yields a larger matching, a contradiction. $\square$

Given $E \in \mathcal{E}$, we call the unique vertex $x \in V(E)$ satisfying (δ_2) a *bottom vertex.* If $y \in E$ is such that $y \neq x$ then we say that y is a *top vertex.* So each $E \in \mathcal{E}$ contains one bottom vertex and two top vertices whereas none of the at most $\alpha|M|$ edges in $M \setminus \mathcal{E}$ contains a top or bottom vertex.

Claim 4.19 *There are at least $(1-6\alpha)|M|^2/2$ pairs $EF \in \binom{M}{2}$ such that*

(ε_1) $L_{v_1}(EF) = \cdots = L_{v_{10}}(EF) \cong B_{113}$;

(ε_2) both E and F contain a bottom vertex w and z respectively;

(ε_3) wz is the base edge of $L_{v_1}(EF)$.

Proof. Consider the directed graph G whose vertex set is M and in which there is a directed edge from E to F if $E \in \mathcal{E}$ and $F \in \mathcal{F}_E$. Claims 4.17 and 4.18 together imply that G has at least $(1-3\alpha)|M|^2$ edges and thus at least $(1-6\alpha)|M|^2/2$ pairs EF of vertices in G must be joined by a double edge. But each such pair EF satisfies the claim. $\square$

4.5.3 Showing that H is $\sqrt{\rho}$-close to $H_{n,d}$

We have now collected all the information we need for showing that H is close to $H_{n,d}(V,W)$, where W will be constructed from the set of bottom vertices in M. More precisely, let W' denote the set of all the bottom vertices. So Claims 4.15 and 4.17 together imply that

$$d - 2\alpha n \leq (1-\alpha)|M| \leq |\mathcal{E}| = |W'| \leq |M| \leq d. \tag{4.11}$$

Let V' denote the set of all the top vertices in H. Thus

$$2d - 4\alpha n \leq 2(1-\alpha)|M| \leq |V'| = 2|W'| \leq 2d. \tag{4.12}$$

Choose a partition V, W of $V(H)$ such that $|W| = d$, $W' \subseteq W$, $V' \subseteq V$. Note that since (4.11) implies that $|W \setminus W'| \leq 2\alpha n$, all but at most $2\alpha n$ vertices of V_0 lie in V. Our aim is to show that H is $\sqrt{\rho}$-close to $H_{n,d}(V,W)$. Note that showing this proves Theorem 4.4 as we can apply Lemma 4.7 since we chose $\rho \ll 1$ in (4.2).

Claim 4.20 *H does not contain an edge of type $V'V_0V_0$.*

Proof. Suppose that the claim is false and let $v'vv_0$ be an edge of H with $v' \in V'$ and $v, v_0 \in V_0$. Let $E \in \mathcal{E}$ be the matching edge containing v'. Take any $F \in \mathcal{F}_E$. Take any 2 vertices from $v_1, \dots, v_{10}$ which are not equal to v_0 or v, call them x and y. Since v' is a top vertex of E, it follows that $L_x(EF)$ contains a 2-matching e_1, e_2 avoiding v'. Note that this is also a 2-matching in $L_y(EF)$. Now we can enlarge M by removing E, F and adding $v'vv_0$, xe_1 and ye_2. This contradicts the maximality of M and proves the claim. □

Claim 4.21

- *H contains at least* $(1-\rho')|W'||V'||V_0|$ *edges of type* $W'V'V_0$.
- *H contains at least* $(1-\rho')|V_0|\binom{|W'|}{2}$ *edges of type* $W'W'V_0$.
- *H contains at most* $\rho'|V_0|\binom{|V'|}{2}$ *edges of type* $V'V'V_0$.

Proof. To see the first part of the claim, consider any $v \in V_0^*$ and any pair w', v' with $w' \in W'$ and $v' \in V'$. Both w' and v' could lie in the same matching edge from M, but there are at most $3|M|$ such pairs. Also, w' and v' could lie in a pair E, F of matching edges from M for which either $L_v(EF) \ncong B_{113}$ or which does not satisfy (ε_1)–(ε_3) in Claim 4.19. But (β) and Claim 4.19 together imply that there are at most $\sqrt{\alpha}n^2$ such pairs E, F. So suppose next that w' and v' lie in a pair E, F satisfying $L_v(EF) \cong B_{113}$ and (ε_1)–(ε_3). Then $L_v(EF), L_{v_1}(EF), \dots, L_{v_9}(EF) \cong B_{113}$ and so $L_v(EF) = L_{v_1}(EF) = \dots = L_{v_9}(EF)$ by Claim 4.16. Conditions (ε_2) and (ε_3) now imply that $w'v' \in E(L_v(W', V'))$. So

$$e(L_v(V', W')) \geq |V'||W'| - 2\sqrt{\alpha}n^2 \geq (1-\rho'/2)|V'||W'|.$$

Summing over all vertices $v \in V_0^*$ and using that $|V_0 \setminus V_0^*| \leq 5C$ implies the first part of the claim. The remaining parts of the claim can be proved similarly. □

Claim 4.22 *H contains at least* $|W'|\binom{|V_0|}{2} - \rho n^3$ *edges of type* $W'V_0V_0$.

Proof. Consider any $v \in V_0$. By Claim 4.20 there are no edges in $L_v(V(H))$ with one endpoint in V' and the other in V_0. By (4.11) there are at most $3\alpha|M|n \leq 3\alpha n^2$ edges in

$L_v(V(H))$ with one endpoint in $V_M \backslash (V' \cup W')$ and the other in V_0. Furthermore, $L_v(V_0)$ contains no edges. Thus,

$$\begin{aligned}
e(L_v(W', V_0)) &\geq \delta_1(H') - e(L_v(V_M)) - 3\alpha n^2 \\
&\overset{(4.3),(4.4),(4.9)}{\geq} (1-\gamma')d\left(n - \frac{d}{2}\right) - \gamma' n^2 - 5(1+\alpha')\binom{|M|}{2} - 3\alpha n^2 \\
&\overset{(4.5)}{\geq} (1-\gamma')|M|\left(n - \frac{|M|}{2}\right) - (5+\sqrt{\alpha})\frac{|M|^2}{2} \\
&\geq |M|(n - 3|M|) - \sqrt{\alpha}|M|n \geq |W'||V_0| - \rho' n^2.
\end{aligned}$$

As earlier, here we use the fact that the function $d(n-d/2)$ is increasing in d for $d \leq n/3$. Summing over all vertices $v \in V_0^*$ and using the fact that $|V_0 \setminus V_0^*| \leq 5C$ now proves the claim. $\square$

Claim 4.23

- *H contains at least $(1-\rho)|W'|\binom{|V'|}{2}$ edges of type $W'V'V'$.*
- *H contains at least $(1-\rho)|V'|\binom{|W'|}{2}$ edges of type $W'W'V'$.*

Proof. First note that the last part of Claim 4.21 implies that all but at most $2\sqrt{\rho'}n$ vertices $x \in V'$ lie in at most $\sqrt{\rho'}|V'||V_0|$ edges of type $V'V'V_0$. Call such vertices x *useful.* Consider any useful x. Then $x \in E'$ for some $E' \in \mathcal{E} \subseteq M$. Further, since x is a top vertex in E', certainly there exists an edge $F' \in M$ such that $L_{v_1}(E'F') = L_{v_2}(E'F') \cong B_{113}$, where x is not a base vertex in $L_{v_1}(E'F')$. So $L_{v_1}(E'F')$ contains a 2-matching $\{e_1, e_2\}$ which avoids x.

Consider any pair $EF \in \binom{M \backslash \{E', F'\}}{2}$ satisfying (ε_1)–(ε_3). We claim that $L_x(EF) \subseteq L_{v_1}(EF)$. Indeed, if not then there exist disjoint edges e_3, e_4 and e_5 such that $e_3 \in E(L_x(EF))$ and $e_4, e_5 \in E(L_{v_1}(EF))$. Since $L_{v_1}(E'F') = L_{v_2}(E'F')$ and since EF satisfies (ε_1) we have that $v_1e_1, v_2e_2, xe_3, v_3e_4$ and v_4e_5 are edges in H'. Replacing E, F, E', F' with $v_1e_1, v_2e_2, xe_3, v_3e_4$ and v_4e_5 in M yields a larger matching in H', a contradiction. So indeed $L_x(EF) \subseteq L_{v_1}(EF)$.

There are at least $(1-6\alpha)|M|^2/2 - 2|M| \geq (1-7\alpha)|M|^2/2$ pairs $EF \in \binom{M\setminus\{E',F'\}}{2}$ satisfying (ε_1)–(ε_3). We claim that at most $\rho^2|M|^2/2$ of these pairs EF are such that $L_x(EF)$ contains fewer than 5 edges. Indeed, suppose not. Since for such EF, $L_x(EF) \subseteq L_{v_1}(EF) \cong B_{113}$, the number of edges of H which contain x and have no endpoint outside V_M is at most

$$4\cdot\rho^2|M|^2/2 + 5\cdot(1-7\alpha-\rho^2)|M|^2/2 + 9\cdot 7\alpha|M|^2/2 + 3|M| \leq (5+30\alpha-\rho^2)|M|^2/2.$$

Here the third term accounts for edges between pairs not satisfying (ε_1)–(ε_3) and the final term for edges with 2 vertices in the same matching edge from M. Let us now bound the number of edges containing x which have an endpoint outside V_M. There are at most $|W'|(n-3|M|) \leq |M|(n-3|M|)$ such edges having an endpoint in W' and at most $\sqrt{\alpha}n^2$ such edges having an endpoint outside $V'\cup W'\cup V_0$. Since H has no edge of type $V'V_0V_0$ by Claim 4.20, the only other such edges consist of x, one vertex in V' and one vertex in V_0. But since x is useful the number of such edges is at most $\sqrt{\rho'}|V'||V_0|$. Thus in total there are at most $|M|(n-3|M|) + 2\sqrt{\rho'}n^2$ edges which contain x and have an endpoint outside V_M. So the degree of x in H is at most

$$\begin{aligned}(5+30\alpha-\rho^2)|M|^2/2 + |M|(n-3|M|) + 2\sqrt{\rho'}n^2 &\leq |M|(n-|M|/2) - \rho^3n^2\\ &\leq d(n-d/2) - \rho^3n^2 \overset{(4.5),(4.3)}{<} \delta_1(H),\end{aligned}$$

a contradiction. Thus there are at least $(1-7\alpha-\rho^2)|M|^2/2$ pairs $EF \in \binom{M\setminus\{E',F'\}}{2}$ satisfying (ε_1)–(ε_3) such that $L_x(EF) = L_{v_1}(EF) \cong B_{113}$. Let $\mathcal{P}$ denote the set of such pairs.

Now consider any pair w', v' with $w' \in W'$ and $v' \in V'\setminus\{x\}$. Both w', v' could lie in the same matching edge from M, but there are at most $3|M|$ such pairs. Also, w', v' could lie in a pair E, F of matching edges which does not belong to $\mathcal{P}$. But there at most $5\rho^2|M|^2$ such pairs w', v'. So suppose next that w', v' lies in a pair E, F belonging to $\mathcal{P}$. Since $L_x(EF) = L_{v_1}(EF) \cong B_{113}$ and EF satisfies (ε_2) and (ε_3) it follows that $w'v' \in E(L_x(EF))$. Thus $e(L_x(W',V')) \geq (1-6\rho^2)|W'||V'|$. Summing over all useful vertices $x \in V'$ proves the first part of the claim. The second part follows similarly (the only change is that we consider a pair $w_1', w_2' \in W'$ in the final paragraph). $\square$

Claims 4.21–4.23 together with (4.11) and (4.12) now show that H contains all but at most $\sqrt{\rho}n^3$ edges of type WVV and WWV and thus H is $\sqrt{\rho}$-close to $H_{n,d}(V, W)$. Hence H contains a perfect matching by Lemma 4.7.

Remark. One can also obtain Theorem 4.4 by proving the result only in the case when $d = \lfloor n/3 \rfloor$. Indeed, suppose that H is as in the theorem. Let $a := \lfloor (n - 3d)/2 \rfloor$. Obtain a new 3-uniform hypergraph H' from H by adding a new vertices to H such that each of these vertices forms an edge with all pairs of vertices in H'. It is not hard to check that $\delta_1(H') > \binom{|H'|-1}{2} - \binom{|H'|-\lfloor |H'|/3 \rfloor}{2}$ and so H' has a matching M' of size $\lfloor |H'|/3 \rfloor$. One can then show that M' contains at least d edges from H, as desired. (We thank Peter Allen for suggesting this trick.)

However, the proof of Theorem 4.4 is only slightly simpler in the case when $d = \lfloor n/3 \rfloor$ (we do not need Claims 4.20–4.22 in this case) and to show that the above trick works, one requires some extra calculations.

Chapter 5
Hamiltonian degree sequences in digraphs

5.1 Introduction

Since it is unlikely that there is a characterisation of all those graphs which contain a Hamilton cycle it is natural to ask for sufficient conditions which ensure Hamiltonicity. One of the most general of these is Chvátal's theorem [19] that characterises all those degree sequences which ensure the existence of a Hamilton cycle in a graph: Suppose that the degrees of the graph are $d_1 \leq \cdots \leq d_n$. If $n \geq 3$ and $d_i \geq i+1$ or $d_{n-i} \geq n-i$ for all $i < n/2$ then G is Hamiltonian. This condition on the degree sequence is best possible in the sense that for any degree sequence violating this condition there is a corresponding graph with no Hamilton cycle. More precisely, if $d_1 \leq \cdots \leq d_n$ is a graphical degree sequence (i.e. there exists a graph with this degree sequence) then there exists a non-Hamiltonian graph G whose degree sequence $d'_1 \leq \cdots \leq d'_n$ is such that $d'_i \geq d_i$ for all $1 \leq i \leq n$.

A special case of Chvátal's theorem is Dirac's theorem, which states that every graph with $n \geq 3$ vertices and minimum degree at least $n/2$ has a Hamilton cycle. An analogue of Dirac's theorem for digraphs was proved by Ghouila-Houri [30]. Nash-Williams [72] raised the question of a digraph analogue of Chvátal's theorem quite soon after the latter was proved.

For a digraph G it is natural to consider both its outdegree sequence $d^+_1, \ldots, d^+_n$ and its indegree sequence $d_1^-, \dots, d_n^-$. Throughout this chapter we take the convention that $d_1^+ \leq \dots \leq d_n^+$ and $d_1^- \leq \dots \leq d_n^-$ without mentioning this explicitly. Note that the terms d_i^+ and d_i^- do not necessarily correspond to the degree of the same vertex of G.

Conjecture 5.1 (Nash-Williams [72]) *Suppose that G is a strongly connected digraph on $n \geq 3$ vertices such that for all $i < n/2$*

(i) $d_i^+ \geq i+1$ or $d_{n-i}^- \geq n-i$,

(ii) $d_i^- \geq i+1$ or $d_{n-i}^+ \geq n-i$.

Then G contains a Hamilton cycle.

No progress has been made on this conjecture so far (see also [8]). It is even an open problem whether the conditions imply the existence of a cycle through any pair of given vertices (see [10]).

As discussed in Section 5.2, one cannot omit the condition that G is strongly connected. At first sight one might also try to replace the degree condition in Chvátal's theorem by

- $d_i^+ \geq i+1$ or $d_{n-i}^+ \geq n-i$,
- $d_i^- \geq i+1$ or $d_{n-i}^- \geq n-i$.

However, Bermond and Thomassen [10] observed that the latter conditions do not guarantee Hamiltonicity. Indeed, consider the digraph obtained from the complete digraph K on $n-2 \geq 4$ vertices by adding two new vertices v and w which both send an edge to every vertex in K and receive an edge from one fixed vertex $u \in K$.

The following example shows that the degree condition in Conjecture 5.1 would be best possible in the sense that for all $n \geq 3$ and all $k < n/2$ there is a non-Hamiltonian strongly connected digraph G on n vertices which satisfies the degree condition except that $d_k^+, d_k^- \geq k+1$ are replaced by $d_k^+, d_k^- \geq k$ in the kth pair of conditions. To see this, take an independent set I of size $k < n/2$ and a complete digraph K of order $n-k$. Pick a set X of k vertices of K and add all possible edges (in both directions) between I and X. The digraph G thus obtained is strongly connected, not Hamiltonian and

$$\underbrace{k,\dots,k}_{k \text{ times}},\underbrace{n-1-k,\dots,n-1-k}_{n-2k \text{ times}},\underbrace{n-1,\dots,n-1}_{k \text{ times}}$$

is both the out- and indegree sequence of G. A more detailed discussion of extremal examples is given in Section 5.2.

In this chapter we prove the following approximate version of Conjecture 5.1 for large digraphs.

Theorem 5.2 *For every $\eta > 0$ there exists an integer $n_0 = n_0(\eta)$ such that the following holds. Suppose G is a digraph on $n \geq n_0$ vertices such that for all $i < n/2$*

- *$d_i^+ \geq i + \eta n$ or $d_{n-i-\eta n}^- \geq n - i$,*
- *$d_i^- \geq i + \eta n$ or $d_{n-i-\eta n}^+ \geq n - i$.*

Then G contains a Hamilton cycle.

Instead of proving Theorem 5.2 directly, we will prove the existence of a Hamilton cycle in a digraph satisfying a certain expansion property (Theorem 5.13). We defer the precise statement to Section 5.6.

The following weakening of Conjecture 5.1 was posed earlier by Nash-Williams [68, 69]. It would yield a digraph analogue of Pósa's theorem which states that a graph G on $n \geq 3$ vertices has a Hamilton cycle if its degree sequence $d_1, \dots, d_n$ satisfies $d_i \geq i + 1$ for all $i < (n-1)/2$ and if additionally $d_{\lceil n/2 \rceil} \geq \lceil n/2 \rceil$ when n is odd [75]. Note that this is much stronger than Dirac's theorem but is a special case of Chvátal's theorem.

Conjecture 5.3 (Nash-Williams [68, 69]) *Let G be a digraph on $n \geq 3$ vertices such that $d_i^+, d_i^- \geq i + 1$ for all $i < (n-1)/2$ and such that additionally $d_{\lceil n/2 \rceil}^+, d_{\lceil n/2 \rceil}^- \geq \lceil n/2 \rceil$ when n is odd. Then G contains a Hamilton cycle.*

The previous example shows that the degree condition would be best possible in the same sense as described there. The assumption of strong connectivity is not necessary in Conjecture 5.3, as it follows from the degree conditions. The following approximate version of Conjecture 5.3 is an immediate consequence of Theorem 5.2.

Corollary 5.4 *For every $\eta > 0$ there exists an integer $n_0 = n_0(\eta)$ such that every digraph G on $n \geq n_0$ vertices with $d_i^+, d_i^- \geq i + \eta n$ for all $i < n/2$ contains a Hamilton cycle.*

In Section 5.4 we give a construction which shows that for oriented graphs there is no analogue of Pósa's theorem.

It will turn out that the conditions of Theorem 5.2 even guarantee the digraph G to be *pancyclic*, i.e. G contains a cycle of length t for all $t = 2, \dots, n$.

Corollary 5.5 *For every $\eta > 0$ there exists an integer $n_0 = n_0(\eta)$ such that the following holds. Suppose G is a digraph on $n \geq n_0$ vertices such that for all $i < n/2$*

- $d_i^+ \geq i + \eta n$ *or* $d_{n-i-\eta n}^- \geq n - i$,
- $d_i^- \geq i + \eta n$ *or* $d_{n-i-\eta n}^+ \geq n - i$.

Then G is pancyclic.

Thomassen [87] proved an Ore-type condition which implies that every digraph with minimum in- and outdegree $> n/2$ is pancyclic. (The complete bipartite digraph whose vertex class sizes are as equal as possible shows that the latter bound is best possible.) Alon and Gutin [2] observed that one can use Ghouila-Houri's theorem to show that every digraph G with minimum in- and outdegree $> n/2$ is even vertex-pancyclic. Here a digraph G is called *vertex-pancyclic* if every vertex of G lies on a cycle of length t for all $t = 2, \dots, n$. In Proposition 5.7 we show that one cannot replace pancyclicity by vertex-pancyclicity in Corollary 5.5. Minimum degree conditions for (vertex-) pancyclicity of oriented graphs are discussed in [44].

This chapter is organised as follows. We first give a more detailed discussion of extremal examples for Conjecture 5.1. In Section 5.3 we then deduce Corollary 5.5 from Theorem 5.2 and show that one cannot replace pancyclicity by vertex-pancyclicity. The proof of Theorem 5.2 uses the Regularity lemma for digraphs (Lemma 2.7) which was introduced in Section 2.2. The proof of Theorem 5.2 is included in Section 5.6. It relies on a result (Lemma 5.9) of Keevash, Kühn and Osthus [41] which was used to prove an analogue of Dirac's theorem for oriented graphs. A related result was proved earlier in [43].

It is a natural question to ask whether the 'error terms' in Theorem 5.2 and Corollary 5.4 can be eliminated using an 'extremal case' or 'stability' analysis. However, this seems quite difficult as there are many different types of digraphs which come close to violating the conditions in Conjectures 5.1 and 5.3 (this is different e.g. to the situation in [41]). As a step in this direction, recently it was shown in [16] that the degrees in the first parts of the conditions in Theorem 5.2 can be capped at $n/2$, i.e. the conditions can be replaced by

- $d_i^+ \geq \min\{i + \eta n, n/2\}$ or $d_{n-i-\eta n}^- \geq n - i$,
- $d_i^- \geq \min\{i + \eta n, n/2\}$ or $d_{n-i-\eta n}^+ \geq n - i$.

The proof of this result is considerably more difficult than that of Theorem 5.2. A (parallel) algorithmic version of Chvátal's theorem for undirected graphs was recently considered in [83] and for directed graphs in [17].

5.2 Extremal examples for Conjecture 5.1 and a weaker conjecture

The example given in Section 5.1 does not quite imply that Conjecture 5.1 would be best possible, as for some k it violates both (i) and (ii) for $i = k$. Here is a slightly more complicated example which only violates one of the conditions for $i = k$ (unless n is odd and $k = \lfloor n/2 \rfloor$).

Suppose $n \geq 5$ and $1 \leq k < n/2$. Let K and K' be complete digraphs on $k - 1$ and $n - k - 2$ vertices respectively. Let G be the digraph on n vertices obtained from the disjoint union of K and K' as follows. Add all possible edges from K' to K (but no edges from K to K') and add new vertices u and v to the digraph such that there are all possible edges from K' to u and v and all possible edges from u and v to K. Finally, add a vertex w that sends and receives edges from all other vertices of G (see Figure 5.1). Thus G is strongly connected, not Hamiltonian and has outdegree sequence

$$\underbrace{k-1, \ldots, k-1}_{k-1 \text{ times}}, k, k, \underbrace{n-1, \ldots, n-1}_{n-k-1 \text{ times}}$$

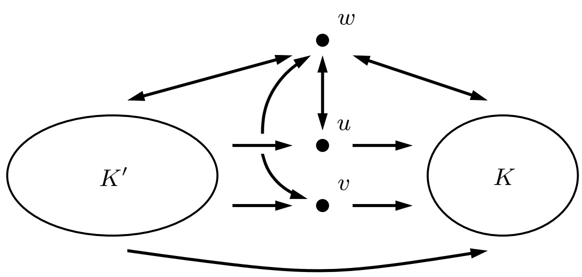

Figure 5.1: An extremal example for Conjecture 5.1

and indegree sequence

$$\underbrace{n-k-2,\dots,n-k-2}_{n-k-2 \text{ times}}, n-k-1, n-k-1, \underbrace{n-1,\dots,n-1}_{k \text{ times}}.$$

Suppose that either n is even or, if n is odd, we have that $k < \lfloor n/2 \rfloor$. One can check that G then satisfies the conditions in Conjecture 5.1 except that $d_k^+ = k$ and $d_{n-k}^- = n-k-1$. (When checking the conditions, it is convenient to note that our assumptions on k and n imply $n-k-1 \geq \lceil n/2 \rceil$. Hence there are at least $\lceil n/2 \rceil$ vertices of outdegree $n-1$ and so (ii) holds for all $i < n/2$.) If n is odd and $k = \lfloor n/2 \rfloor$ then conditions (i) and (ii) both fail for $i = k$. We do not know whether a similar construction as above also exists for this case. It would also be interesting to find an analogous construction as above for Conjecture 5.3.

Here is also an example which shows that the assumption of strong connectivity in Conjecture 5.1 cannot be omitted. Let $n \geq 4$ be even. Let K and K' be two disjoint copies of a complete digraph on $n/2$ vertices. Obtain a digraph G from K and K' by adding all possible edges from K to K' (but none from K' to K). It is easy to see that G is neither Hamiltonian, nor strongly connected, but satisfies the condition on the degree sequences given in Conjecture 5.1.

As it stands, the additional connectivity assumption means that Conjecture 5.1 does not seem to be a precise digraph analogue of Chvátal's theorem: in such an analogue, we would ask for a complete characterisation of *all* digraph degree sequences which force Hamiltonicity. However, it turns out that it makes sense to replace the strong connectivity assumption with an additional degree condition (condition (iii) below). If true, the following conjecture would provide the desired characterisation.

Conjecture 5.6 (Kühn, Osthus and Treglown [61]) *Suppose that G is a digraph on $n \geq 3$ vertices such that for all $i < n/2$*

(i) $d_i^+ \geq i+1$ or $d_{n-i}^- \geq n-i$,

(ii) $d_i^- \geq i+1$ or $d_{n-i}^+ \geq n-i$,

and such that (iii) *$d_{n/2}^+ \geq n/2$ or $d_{n/2}^- \geq n/2$ if n is even. Then G contains a Hamilton cycle.*

Conjecture 5.6 would actually follow from Conjecture 5.1. To see this, it of course suffices to check that the conditions in Conjecture 5.6 imply strong connectivity. This in turn is easy to verify, as the degree conditions imply that for any vertex set S with $|S| \leq n/2$ we have $|N^-(S) \cup S| > |S|$ and $|N^+(S) \cup S| > |S|$. (We need (iii) to obtain this assertion precisely for those S with $|S| = n/2$.)

It remains to check that Conjecture 5.6 would indeed characterise all digraph degree sequences which force a Hamilton cycle. Unless n is odd and $k = \lfloor n/2 \rfloor$, the construction at the beginning of the section already gives non-Hamiltonian graphs which satisfy all the degree conditions (including (iii)) except (i) for $i = k$. To cover the case when n is odd and $k = \lfloor n/2 \rfloor$, let G be the digraph obtained from two disjoint cliques K and K' of orders $\lceil n/2 \rceil$ and $\lfloor n/2 \rfloor$ by adding all edges from K to K'. If $i = k = \lfloor n/2 \rfloor$ then G satisfies (ii) (because $d_{n-k}^+ = n-1$) but not (i). For all other i, both conditions are satisfied. Finally, the example immediately preceding Conjecture 5.6 gives a graph on an even number n of vertices which satisfies (i) and (ii) for all $i < n/2$ but does not satisfy (iii).

Nash-Williams observed that Conjecture 5.1 would imply Chvátal's theorem. (Indeed, given an undirected graph G satisfying the degree condition in Chvátal's theorem, obtain a digraph by replacing each undirected edge with a pair of directed edges, one in each direction. This satisfies the degree condition in Conjecture 5.1. It is also strongly connected, as it is easy to see that G must be connected.) A disadvantage of Conjecture 5.6 is that it would not imply Chvátal's theorem in the same way: consider a graph G which is obtained from $K_{n/2,n/2}$ by removing a perfect matching and adding a spanning cycle in one of the two vertex classes. The degree sequence of this G satisfies the conditions of Chvátal's theorem. However, the digraph obtained by doubling the edges of G does not satisfy (iii) in Conjecture 5.6.

5.3 The proof of Corollary 5.5

We begin this section with a proof of Corollary 5.5.

Proof of Corollary 5.5. Our first aim is to prove the existence of a vertex $x \in V(G)$ such that $d^+(x)+d^-(x) \geq n$. Such a vertex exists if there is an index j with $d_j^+ + d_{n-j}^- \geq n$. Indeed, at least $n-j+1$ vertices of G have outdegree at least d_j^+ and at least $j+1$ vertices have indegree at least d_{n-j}^-. Thus there will be a vertex x with $d^+(x) \geq d_j^+$ and $d^-(x) \geq d_{n-j}^-$.

To prove the existence of such an index j, suppose first that there is an i with $2 \leq i < n/2$ and such that $d_{i-1}^+ \geq i$ but $d_i^+ = i$. Then $d_{n-i}^- \geq n-i$ and so $d_i^+ + d_{n-i}^- \geq n$ as required. The same argument works if there is an i with $2 \leq i < n/2$ and such that $d_{i-1}^- \geq i$ but $d_i^- = i$. Suppose next that $d_1^+ \leq 1$. Then $d_{n-1}^- \geq n-1$ and so $d_1^+ = 1$. Thus we can take $j := 1$. Again, the same argument works if $d_1^- \leq 1$. Thus we may assume that $d_{\lceil n/2 \rceil -1}^+, d_{\lceil n/2 \rceil -1}^- \geq \lceil n/2 \rceil$. But in this case we can take $j := \lfloor n/2 \rfloor$.

Now let x be a vertex with $d^+(x) + d^-(x) \geq n$, set $G' := G - x$ and $n' := |G'|$. Let $d_{1,G'}^+, \dots, d_{n',G'}^+$ and $d_{1,G'}^-, \dots, d_{n',G'}^-$ denote the out- and the indegree sequences of G'. Given some $i \leq n'$ and $s > 0$, if $d_i^+ \geq s$ then at least $n+1-i$ vertices in G have outdegree at least s. Thus at least $n-i = n'+1-i$ vertices in G' have outdegree at least $s-1$ and so $d_{i,G'}^+ \geq s-1$. Thus for all $i < n/2$ the degree sequences of G' satisfy

- $d_{i,G'}^+ \geq i + \eta n - 1$ or $d_{n-i-\eta n,G'}^- \geq n-i-1$,
- $d_{i,G'}^- \geq i + \eta n - 1$ or $d_{n-i-\eta n,G'}^+ \geq n-i-1$

and so

- $d_{i,G'}^+ \geq i + \eta n'/2$ or $d_{n'-i-\eta n'/2,G'}^- \geq n'-i$,
- $d_{i,G'}^- \geq i + \eta n'/2$ or $d_{n'-i-\eta n'/2,G'}^+ \geq n'-i$.

Hence we can apply Theorem 5.2 with η replaced by $\eta/2$ to obtain a Hamilton cycle $C = x_1 \dots x_{n'}$ in G'. We now apply the same trick as in [2] to obtain a cycle (through x) in G of the desired length, t say (where $2 \leq t \leq n$): Since $d^+_G(x) + d^-_G(x) \geq n > n'$ there exists an i such that $x_i \in N^+_G(x)$ and $x_{i+t-2} \in N^-_G(x)$ (where we take the indices modulo n'). But then $xx_i \dots x_{i+t-2}x$ is the required cycle of length t. □

Note that the proof of Corollary 5.5 shows that if Conjecture 5.1 holds and G is a strongly 2-connected digraph with

- $d^+_i \geq i+2$ or $d^-_{n-i-1} \geq n-i$,
- $d^-_i \geq i+2$ or $d^+_{n-i-1} \geq n-i$

for all $i < n/2$ then G is pancyclic.

The next result implies that we cannot replace pancyclicity with vertex-pancyclicity in Corollary 5.5.

Proposition 5.7 *Given any $k \geq 3$ there are $\eta = \eta(k) > 0$ and $n_0 = n_0(k)$ such that for every $n \geq n_0$ there exists a digraph G on n vertices with $d^+_i, d^-_i \geq i + \eta n$ for all $i < n/2$, but such that some vertex of G does not lie on a cycle of length less than k.*

Proof. Let $\eta := 1/(k3^k)$ and suppose that n is sufficiently large. Let G be the digraph obtained from the disjoint union of $k-2$ independent sets $V_1, \dots, V_{k-2}$ with $|V_i| = 3^i \lceil \eta n \rceil$ and a complete digraph K on $n-1-|V_1 \cup \dots \cup V_{k-2}|$ vertices as follows. Add a new vertex x which sends an edge to all vertices in V_1 and receives an edge from all vertices in K. Add all possible edges from V_i to V_{i+1} (but no edges from V_{i+1} to V_i) for each $i \leq k-3$. Finally, add all possible edges going from vertices in K to other vertices and add all edges from V_{k-2} to K. Then $d^-_i \geq |K| \geq 2n/3$ and $d^+_i \geq i + \eta n$ for all $i < n/2$ with room to spare. However, if C is a cycle containing x then the inneighbour of x on C must lie in K. But the shortest path from x to K has length $k-1$ and so $|C| \geq k$, as required. □

5.4 Degree sequences for Hamilton cycles in oriented graphs

In Section 5.1 we mentioned Ghouila-Houri's theorem which gives a bound on the minimum semidegree of a digraph G guaranteeing a Hamilton cycle. A natural question raised by Thomassen [88] is that of determining the minimum semidegree which ensures a Hamilton cycle in an oriented graph. Häggkvist [31] conjectured that every oriented graph G of order $n \geq 3$ with $\delta^0(G) \geq (3n-4)/8$ contains a Hamilton cycle. The bound on the minimum semidegree would be best possible. Keevash, Kühn and Osthus [41] confirmed this conjecture for sufficiently large oriented graphs.

Pósa's theorem implies the existence of a Hamilton cycle in a graph G even if G contains a significant number of vertices of degree much less than $n/2$, i.e. of degree much less than the minimum degree required to force a Hamilton cycle. In particular, Pósa's theorem is much stronger than Dirac's theorem. In the same sense, Conjecture 5.3 would be much stronger than Ghouila-Houri's theorem. The following proposition implies that we cannot strengthen Häggkvist's conjecture in this way: there are non-Hamiltonian oriented graphs which contain just a bounded number of vertices whose semidegree is (only slightly) smaller than $3n/8$. To state this proposition we need to introduce the notion of dominating sequences: Given sequences $x_1, \dots, x_n$ and $y_1, \dots, y_n$ of numbers we say that $y_1, \dots, y_n$ *dominates* $x_1, \dots, x_n$ if $x_i \leq y_i$ for all $1 \leq i \leq n$.

Proposition 5.8 *For every $0 < \alpha < 3/8$, there is an integer $c = c(\alpha)$ and infinitely many oriented graphs G whose in- and outdegree sequences both dominate*

$$\underbrace{\alpha|G|, \dots, \alpha|G|}_{c \text{ times}}, 3|G|/8, \dots, 3|G|/8$$

but such that G does not contain a Hamilton cycle.

Proof. Define $c := 4t$ where $t \in \mathbb{N}$ is chosen such that $3 - 1/t > 8\alpha$. Let n be sufficiently large and such that $8t$ divides n and define vertex sets A, B, C, D and E of sizes $n/4, n/8, n/8-1, n/4+1$ and $n/4$ respectively.

Let G be the oriented graph obtained from the disjoint union of A, B, C, D and E by defining the following edges: G contains all possible edges from A to B, B to C, C to D, A to C, B to D and D to A. E sends out all possible edges to A and B and receives all possible edges from C and D. B and C both induce tournaments that are as regular as possible (see Figure 5.2). So certainly $d^+_G(x), d^-_G(x) \geq 3n/8$ for all $x \in B \cup C \cup E$. Furthermore,

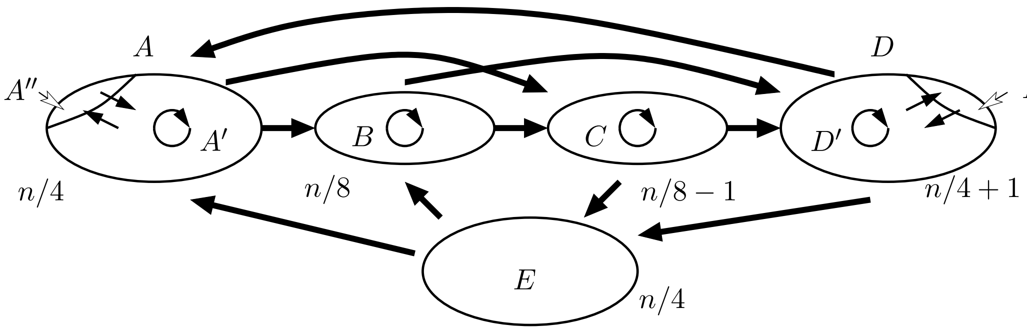

Figure 5.2: The oriented graph G in Proposition 5.8

currently, $d^+_G(a) = n/4 - 1$, $d^-_G(a) = n/2 + 1$, $d^+_G(d) = n/2$ and $d^-_G(d) = n/4 - 1$ for all $a \in A$ and all $d \in D$.

Partition A into A' and A'' where $|A''| = c$ and thus $|A'| = n/4 - c$. Write $A' =: \{x_1, x_2, \ldots, x_{n/8-c/2}, y_1, y_2, \ldots, y_{n/8-c/2}\}$ and $A'' =: \{z_1, \ldots, z_{2t}, w_1, \ldots, w_{2t}\}$. Let A' induce a tournament that is as regular as possible. In particular, every vertex in A' sends out at least $n/8 - c/2 - 1$ edges to other vertices in A'. We define the edges between A' and A'' as follows: Add the edges $x_i z_j, y_i w_j$ to G for all $1 \leq i \leq n/8 - c/2$ and $1 \leq j \leq 2t$. Note that we can partition both $\{x_1, \ldots, x_{n/8-c/2}\}$ and $\{y_1, \ldots, y_{n/8-c/2}\}$ into t sets of size $s := n/(2c) - 2$. For each $0 \leq i \leq t-1$ add all possible edges from $\{x_{si+1}, \ldots, x_{s(i+1)}\}$ to $\{w_{2i+1}, w_{2i+2}\}$ and from $\{y_{si+1}, \ldots, y_{s(i+1)}\}$ to $\{z_{2i+1}, z_{2i+2}\}$. If $a' \in A'$ and $a'' \in A''$ are such that the edge $a'a''$ has not been included into G so far then add the edge $a''a'$ to G. Thus, $d^+_G(a') \geq (n/4 - 1) + (n/8 - c/2 - 1) + c/2 + 2 = 3n/8$ for all $a' \in A'$ and

$$d^+_G(a'') \geq (n/4 - 1) + (n/8 - c/2 - s) = 3n/8 - c/2 - n/(2c) + 1 \geq \alpha n$$

for all $a'' \in A''$.

Partitioning D into D' and D'' (where $|D''| = c$) and defining edges inside D in a similar fashion to those inside A, we can ensure that $d_G^-(d') \geq 3n/8$ for all $d' \in D'$ and $d_G^-(d'') \geq \alpha n$ for all $d'' \in D''$. So indeed G has the desired degree sequences.

E is an independent set, so if G contains a Hamilton cycle H then the inneighbour of each vertex in E on H must lie in $C \cup D$ while its outneighbour lies in $A \cup B$. So H contains at least $|E| = n/4$ disjoint edges going from $A \cup B$ to $C \cup D$. However, all such edges in G have at least one endvertex in $B \cup C$. So there are at most $|B| + |C| = n/4 - 1 < |E|$ such disjoint edges in G. Thus G does not contain a Hamilton cycle (in fact, G does not contain a 1-factor). □

5.5 Expansion and robustness in digraphs

Given $0 < \nu \leq \tau < 1$, we call a digraph G a (ν, τ)-*outexpander* if $|N^+(S)| \geq |S| + \nu|G|$ for all $S \subseteq V(G)$ with $\tau|G| < |S| < (1 - \tau)|G|$. The main tool in the proof of Theorem 5.2 is the following result from [41].

Lemma 5.9 *Let M', n_0 be positive integers and let $\varepsilon, d, \eta, \nu, \tau$ be positive constants such that $1/n_0 \ll 1/M' \ll \varepsilon \ll d \ll \nu \leq \tau \ll \eta < 1$. Let G be an oriented graph on $n \geq n_0$ vertices such that $\delta^0(G) \geq 2\eta n$. Let R be the reduced digraph of G with parameters ε, d and M'. Suppose that there exists a spanning oriented subgraph R^* of R with $\delta^0(R^*) \geq \eta|R^*|$ which is a (ν, τ)-outexpander. Then G contains a Hamilton cycle.*

Our next aim is to show that any digraph G as in Theorem 5.2 is an outexpander. In fact, we will show that even the 'robust outneighbourhood' of any set $S \subseteq V(G)$ of reasonable size is significantly larger than S. More precisely, let $0 < \nu \leq \tau < 1$. Given any digraph G and $S \subseteq V(G)$, the ν-*robust outneighbourhood* $RN^+_{\nu,G}(S)$ *of* S is the set of all those vertices x of G which have at least $\nu|G|$ inneighbours in S. G is called a *robust* (ν, τ)-*outexpander* if $|RN^+_{\nu,G}(S)| \geq |S| + \nu|G|$ for all $S \subseteq V(G)$ with $\tau|G| < |S| < (1 - \tau)|G|$.

Lemma 5.10 *Let n_0 be a positive integer and τ, η be positive constants such that $1/n_0 \ll \tau \ll \eta < 1$. Let G be a digraph on $n \geq n_0$ vertices with*

(i) $d_i^+ \geq i + \eta n$ or $d^-_{n-i-\eta n} \geq n - i$,

(ii) $d^-_i \geq i + \eta n$ *or* $d^+_{n-i-\eta n} \geq n - i$

for all $i < n/2$*. Then* $\delta^0(G) \geq \eta n$ *and* G *is a robust* (τ^2, τ)*-outexpander.*

Proof. Clearly, if $d^+_1 \geq 1 + \eta n$ then $\delta^+(G) \geq \eta n$. If $d^+_1 < 1 + \eta n$ then (i) implies that $d^-_{n-1-\eta n} \geq n - 1$. Thus G has at least $\eta n + 1$ vertices of indegree $n - 1$ and so $\delta^+(G) \geq \eta n$. It follows similarly that $\delta^-(G) \geq \eta n$.

Consider any non-empty set $S \subseteq V(G)$ with $\tau n < |S| < (1-\tau)n$ and $|S| \neq n/2 + \lfloor \tau n \rfloor$. Let us first deal with the case when $d^+_{|S|-\lfloor \tau n \rfloor} \geq |S| - \lfloor \tau n \rfloor + \eta n \geq |S| + \eta n/2$. Then S contains a set X of $\lfloor \tau n \rfloor$ vertices, each having outdegree at least $|S| + \eta n/2$. Let Y be the set of all those vertices of G that have at least $\tau^2 n$ inneighbours in X. Then

$$|X|(|S| + \eta n/2) \leq |Y||X| + (n - |Y|)\tau^2 n \leq |Y||X| + \tau^2 n^2$$

and so $|RN^+_{\tau^2,G}(S)| \geq |Y| \geq |S| + 2\tau^2 n$.

So suppose next that $d^+_{|S|-\lfloor \tau n \rfloor} < |S| - \lfloor \tau n \rfloor + \eta n$. Since $\delta^-(G) \geq \eta n$ we may assume that $|S| \leq (1 - \eta + \tau^2)n < n - 1 - \eta n + \lfloor \tau n \rfloor$ (otherwise $RN^+_{\tau^2,G}(S) = V(G)$ and we are done). Thus

$$d^-_{n-|S|+\lfloor \tau n \rfloor - \eta n} \geq n - |S| + \lfloor \tau n \rfloor \geq n - |S| + \tau^2 n$$

by (i) and (ii). (Here we use that $|S| \neq n/2 + \lfloor \tau n \rfloor$.)

So G contains at least $|S| - \lfloor \tau n \rfloor + \eta n \geq |S| + \eta n/2$ vertices x of indegree at least $n - |S| + \tau^2 n$. If $|RN^+_{\tau^2,G}(S)| < |S| + 2\tau^2 n$ then $V(G) \setminus RN^+_{\tau^2,G}(S)$ contains such a vertex x. But then x has at least $\tau^2 n$ neighbours in S, i.e. $x \in RN^+_{\tau^2,G}(S)$, a contradiction.

If $|S| = n/2 + \lfloor \tau n \rfloor$ then considering the outneighbourhood of a subset of S of size $|S| - 1$ shows that $|RN^+_{\tau^2,G}(S)| \geq |S| - 1 + 2\tau^2 n \geq |S| + \tau^2 n$. □

The next result implies that the property of a digraph G being a robust outexpander is 'inherited' by the reduced digraph of G. For this (and for Lemma 5.12) we need that G is a robust outexpander, rather than just an outexpander.

Lemma 5.11 *Let* M', n_0 *be positive integers and let* $\varepsilon, d, \eta, \nu, \tau$ *be positive constants such that* $1/n_0 \ll \varepsilon \ll d \ll \nu, \tau, \eta < 1$ *and such that* $M' \ll n_0$*. Let* G *be a digraph on* $n \geq n_0$

vertices with $\delta^0(G) \geq \eta n$ and such that G is a robust (ν, τ)-outexpander. Let R be the reduced digraph of G with parameters ε, d and M'. Then $\delta^0(R) \geq \eta|R|/2$ and R is a robust $(\nu/2, 2\tau)$-outexpander.

Proof. Let G' denote the pure digraph, $L := |R|$, let $V_1, \dots, V_L$ be the clusters of G (i.e. the vertices of R) and V_0 the exceptional set. Let $m := |V_1| = \dots = |V_L|$. Then

$$\delta^0(R) \geq (\delta^0(G') - |V_0|)/m \geq (\delta^0(G) - (d + 2\varepsilon)n)/m \geq \eta L/2.$$

Consider any $S \subseteq V(R)$ with $2\tau L \leq |S| \leq (1 - 2\tau)L$. Let S' be the union of all the clusters belonging to S. Then $\tau n \leq |S'| \leq (1 - 2\tau)n$. Since $|N^-_{G'}(x) \cap S'| \geq |N^-_G(x) \cap S'| - (d + \varepsilon)n \geq \nu n/2$ for every $x \in RN^+_{\nu,G}(S')$ this implies that

$$|RN^+_{\nu/2,G'}(S')| \geq |RN^+_{\nu,G}(S')| \geq |S'| + \nu n \geq |S|m + \nu mL.$$

However, in G' every vertex $x \in RN^+_{\nu/2,G'}(S') \setminus V_0$ receives edges from vertices in at least $|N^-_{G'}(x) \cap S'|/m \geq (\nu n/2)/m \geq \nu L/2$ clusters $V_i \in S$. Thus by the final property of the partition in Lemma 2.7 the cluster V_j containing x is an outneighbour of each such V_i (in R). Hence $V_j \in RN^+_{\nu/2,R}(S)$. This in turn implies that

$$|RN^+_{\nu/2,R}(S)| \geq (|RN^+_{\nu/2,G'}(S')| - |V_0|)/m \geq |S| + \nu L/2,$$

as required. $\square$

The strategy of the proof of Theorem 5.2 is as follows. By Lemma 5.10 our given digraph G is a robust outexpander and by Lemma 5.11 this also holds for the reduced digraph R of G. The next result gives us a spanning oriented subgraph R^* of R which is still an outexpander. The somewhat technical property concerning the subdigraph $H \subseteq R$ in Lemma 5.12 will be used to guarantee an oriented subgraph G^* of G which has linear minimum semidegree and such that R^* is a reduced digraph of G^*. (G^* will be obtained from the spanning subgraph of the pure digraph G' which corresponds to R^* by modifying the neighbourhoods of a small number of vertices.) Finally, we will apply Lemma 5.9

with R^* playing the role of both R and R^* and G^* playing the role of G to find a Hamilton cycle in G^* and thus in G.

Lemma 5.12 *Given positive constants $\nu \le \tau \le \eta$, there exists a positive integer n_0 such that the following holds. Let R be a digraph on $n \ge n_0$ vertices which is a robust (ν, τ)-outexpander. Let H be a spanning subdigraph of R with $\delta^0(H) \ge \eta n$. Then R has a spanning oriented subgraph R^* which is a robust $(\nu/12, \tau)$-outexpander and such that $\delta^0(R^* \cap H) \ge \eta n/4$.*

Proof. Consider a random spanning oriented subgraph R^* of R obtained by deleting one of the edges xy, yx (each with probability 1/2) for every pair $x, y \in V(R)$ for which $xy, yx \in E(R)$, independently from all other such pairs. Given a vertex x of R, we write $N_R^{\pm}(x)$ for the set of all those vertices of R which are both out- and inneighbours of x and define $N_H^{\pm}(x)$ similarly. Let $H^* := H \cap R^*$. Clearly, $d_{H^*}^+(x), d_{H^*}^-(x) \ge \eta n/4$ if $|N_H^{\pm}(x)| \le 3\eta n/4$. So suppose that $|N_H^{\pm}(x)| \ge 3\eta n/4$. Let $X := |N_H^{\pm}(x) \cap N_{H^*}^+(x)|$. Then $\mathbb{E}X \ge 3\eta n/8$ and so a standard Chernoff estimate (see e.g. [4, Cor. A.14]) implies that

$$\mathbb{P}(d_{H^*}^+(x) < \eta n/4) \le \mathbb{P}(X < \eta n/4) \le \mathbb{P}(X < 2\mathbb{E}X/3) < 2\mathrm{e}^{-c\mathbb{E}X} \le 2\mathrm{e}^{-3c\eta n/8},$$

where c is an absolute constant (i.e. it does not depend on ν, τ or η). Similarly it follows that $\mathbb{P}(d_{H^*}^-(x) < \eta n/4) \le 2\mathrm{e}^{-3c\eta n/8}$.

Consider any set $S \subseteq V(R^*) = V(R)$. Let $ERN_{\nu/3,R}^+(S) := RN_{\nu/3,R}^+(S) \setminus S$ and define $ERN_{\nu/12,R^*}^+(S)$ similarly. We say that S is *good* if all but at most $\nu n/6$ vertices in $ERN_{\nu/3,R}^+(S)$ are contained in $ERN_{\nu/12,R^*}^+(S)$. Our next aim is to show that

$$\mathbb{P}(S \text{ is not good}) \le \mathrm{e}^{-n}. \tag{5.1}$$

To prove (5.1), write $ERN_R^{\pm}(S)$ for the set of all those vertices $x \in ERN_{\nu/3,R}^+(S)$ for which $|N_R^{\pm}(x) \cap S| \ge \nu n/4$. Note that every vertex in $ERN_{\nu/3,R}^+(S) \setminus ERN_R^{\pm}(S)$ will automatically lie in $ERN_{\nu/12,R^*}^+(S)$. We say that a vertex $x \in ERN_R^{\pm}(S)$ *fails* if $x \notin ERN_{\nu/12,R^*}^+(S)$. The expected size of $N_{R^*}^-(x) \cap N_R^{\pm}(x) \cap S$ is at least $\nu n/8$. So as before, a Chernoff estimate gives

$$\mathbb{P}(x \text{ fails}) \leq \mathbb{P}(|N^-_{R^*}(x) \cap N^\pm_R(x) \cap S| < \nu n/12) \leq 2\mathrm{e}^{-c\nu n/8} =: p.$$

Let Y be the number of all those vertices $x \in ERN^\pm_R(S)$ which fail. Then $\mathbb{E}Y \leq p|ERN^\pm_R(S)| \leq pn$. Note that the failure of distinct vertices is independent (which is the reason we are only considering vertices in the external neighbourhood of S). So we can apply the following Chernoff estimate (see e.g. [4, Theorem A.12]): If $C \geq \mathrm{e}^2$ we have

$$\mathbb{P}(Y \geq C\mathbb{E}Y) \leq \mathrm{e}^{(C - C\ln C)\mathbb{E}Y} \leq \mathrm{e}^{-C(\ln C)\mathbb{E}Y/2}.$$

Setting $C := \nu n/(6\mathbb{E}Y) \geq \nu/(6p)$ this gives

$$\begin{aligned}\mathbb{P}(S \text{ is not good}) = \mathbb{P}(Y > \nu n/6) = \mathbb{P}(Y > C\mathbb{E}Y) &\leq \mathrm{e}^{-C(\ln C)\mathbb{E}Y/2} = \mathrm{e}^{-\nu n(\ln C)/12} \\ &\leq \mathrm{e}^{-n}.\end{aligned}$$

(The last inequality follows since $p \ll \nu$ if n is sufficiently large.) This completes the proof of (5.1).

Since $4n\mathrm{e}^{-3c\eta n/8} + 2^n\mathrm{e}^{-n} < 1$ (if n is sufficiently large) this implies that there is an outcome for R^* such that $\delta^0(R^* \cap H) \geq \eta n/4$ and such that every set $S \subseteq V(R)$ is good. We will now show that the latter property implies that such an R^* is a robust $(\nu/12, \tau)$-outexpander. So consider any set $S \subseteq V(R)$ with $\tau n < |S| < (1-\tau)n$. Let $EN := ERN^+_{\nu,R}(S)$ and $N := RN^+_{\nu,R}(S) \cap S$. So $EN \cup N = RN^+_{\nu,R}(S)$. Since S is good and $EN \subseteq ERN^+_{\nu/3,R}(S)$ all but at most $\nu n/6$ vertices in EN are contained in $ERN^+_{\nu/12,R^*}(S) \subseteq RN^+_{\nu/12,R^*}(S)$.

Now consider any partition of S into S_1 and S_2 such that every vertex $x \in N$ satisfies $|N^-_R(x) \cap S_i| \geq \nu n/3$ for $i = 1, 2$. (The existence of such a partition follows by considering a random partition.) Then $S_1 \cap N \subseteq ERN^+_{\nu/3,R}(S_2)$. But since S_2 is good this implies that all but at most $\nu n/6$ vertices in $S_1 \cap N$ are contained in $ERN^+_{\nu/12,R^*}(S_2) \subseteq RN^+_{\nu/12,R^*}(S)$. Similarly, since S_1 is good, all but at most $\nu n/6$ vertices in $S_2 \cap N$ are contained in

$ERN^+_{\nu/12,R^*}(S_1) \subseteq RN^+_{\nu/12,R^*}(S)$. Altogether this shows that

$$|RN^+_{\nu/12,R^*}(S)| \geq |EN \cup (S_1 \cap N) \cup (S_2 \cap N)| - \frac{3\nu n}{6} = |RN^+_{\nu,R}(S)| - \frac{\nu n}{2} \geq |S| + \frac{\nu n}{2},$$

as required. □

5.6 Proof of Theorem 5.2

As indicated in Section 5.1, instead of proving Theorem 5.2 directly, we will prove the following stronger result. It immediately implies Theorem 5.2 since by Lemma 5.10 any digraph G as in Theorem 5.2 is a robust outexpander and satisfies $\delta^0(G) \geq \eta n$.

Theorem 5.13 *Let n_0 be a positive integer and ν, τ, η be positive constants such that $1/n_0 \ll \nu \leq \tau \ll \eta < 1$. Let G be a digraph on $n \geq n_0$ vertices with $\delta^0(G) \geq \eta n$ which is a robust (ν, τ)-outexpander. Then G contains a Hamilton cycle.*

Proof. Pick a positive integer M' and additional constants ε, d such that $1/n_0 \ll 1/M' \ll \varepsilon \ll d \ll \nu$. Apply the Regularity lemma (Lemma 2.7) with parameters ε, d and M' to G to obtain clusters $V_1, \dots, V_L$, an exceptional set V_0 and a pure digraph G'. Then $\delta^0(G') \geq (\eta - (d+\varepsilon))n$ by Lemma 2.7. Let R be the reduced digraph of G with parameters ε, d and M'. Lemma 5.11 implies that $\delta^0(R) \geq \eta L/2$ and that R is a robust $(\nu/2, 2\tau)$-outexpander.

Let H be the spanning subdigraph of R in which V_iV_j is an edge if $V_iV_j \in E(R)$ and the density $d_{G'}(V_i, V_j)$ of the oriented subgraph $(V_i, V_j)_{G'}$ of G' is at least $\eta/4$. We will now give a lower bound on $\delta^+(H)$. So consider any cluster V_i and let $m := |V_i|$. Writing $e_{G'}(V_i, V(G) \setminus V_0)$ for the number of all edges from V_i to $V(G) \setminus V_0$ in G', we have

$$\sum_{V_j \in N^+_R(V_i)} d_{G'}(V_i, V_j) m^2 = e_{G'}(V_i, V(G) \setminus V_0) \geq \delta^0(G')m - |V_0|m \geq (\eta - 2d)nm.$$

It is easy to see that this implies that there are at least $\eta L/4$ outneighbours V_j of V_i in R such that $d_{G'}(V_i, V_j) \geq \eta/4$. But each such V_j is an outneighbour of V_i in H and so

$\delta^+(H) \geq \eta L/4$. It follows similarly that $\delta^-(H) \geq \eta L/4$. We now apply Lemma 5.12 to find a spanning oriented subgraph R^* of R which is a (robust) $(\nu/24, 2\tau)$-outexpander and such that $\delta^0(R^* \cap H) \geq \eta L/16$. Let $H^* := H \cap R^*$.

Our next aim is to modify the pure digraph G' into a spanning oriented subgraph of G having minimum semidegree at least $\eta^2 n/100$. Let G^* be the spanning subgraph of G' which corresponds to R^*. So G^* is obtained from G' by deleting all those edges xy that join some cluster V_i to some cluster V_j with $V_iV_j \in E(R) \setminus E(R^*)$. Note that $G^* - V_0$ is an oriented graph. However, some vertices of $G^* - V_0$ may have small degrees. We will show that there are only a few such vertices and we will add them to V_0 in order to achieve that the out- and indegrees of all the vertices outside V_0 are large. So consider any cluster V_i. For any cluster $V_j \in N^+_{H^*}(V_i)$ at most εm vertices in V_i have less than $(d_{G'}(V_i, V_j) - \varepsilon)m \geq \eta m/5$ outneighbours in V_j (in the digraph G'). Call all these vertices of V_i *useless for* V_j. Thus on average any vertex of V_i is useless for at most $\varepsilon |N^+_{H^*}(V_i)|$ clusters $V_j \in N^+_{H^*}(V_i)$. This implies that at most $\sqrt{\varepsilon} m$ vertices in V_i are useless for more than $\sqrt{\varepsilon}|N^+_{H^*}(V_i)|$ clusters $V_j \in N^+_{H^*}(V_i)$. Let $U_i^+ \subseteq V_i$ be a set of size $\sqrt{\varepsilon} m$ which consists of all these vertices and some extra vertices from V_i if necessary. Similarly, we can choose a set $U_i^- \subseteq V_i \setminus U_i^+$ of size $\sqrt{\varepsilon} m$ such that for every vertex $x \in V_i \setminus U_i^-$ there are at most $\sqrt{\varepsilon}|N^-_{H^*}(V_i)|$ clusters $V_j \in N^-_{H^*}(V_i)$ such that x has less than $\eta m/5$ inneighbours in V_j. For each $i = 1, \dots, L$ remove all the vertices in $U_i^+ \cup U_i^-$ and add them to V_0. We still denote the subclusters obtained in this way by $V_1, \dots, V_L$ and the exceptional set by V_0. Thus we now have that $|V_0| \leq 3\sqrt{\varepsilon} n$. Moreover,

$$\delta^0(G^* - V_0) \geq \frac{\eta m}{5}(1 - \sqrt{\varepsilon})\delta^0(H^*) - |V_0| \geq \frac{\eta m}{5}\frac{\eta L}{17} - 3\sqrt{\varepsilon} n \geq \frac{\eta^2 n}{100}.$$

We now modify G^* by altering the neighbours of the exceptional vertices: For every $x \in V_0$ we select a set of $\eta n/2$ outneighbours of x in G and a set of $\eta n/2$ inneighbours such that these two sets are disjoint and add the edges between x and the selected neighbours to G^*. We still denote the oriented graph thus obtained from G^* by G^*. Then $\delta^0(G^*) \geq \eta^2 n/100$. Since the partition $V_0, V_1, \dots, V_L$ of $V(G^*)$ is as described in the Regularity lemma (Lemma 2.7) with parameters $3\sqrt{\varepsilon}$, $d - \varepsilon$ and M' (where G^* plays the role of G' and G) we can say that R^* is a reduced digraph of G^* with these parameters. Thus we may apply Lemma 5.9 with R^* playing the role of both R and R^* and G^* playing the role of G to find a Hamilton cycle in G^* and thus in G. $\square$

Theorem 5.13 is used as a tool in [53] to prove an approximate version of Sumner's universal tournament conjecture. The result also has an application to a conjecture of Thomassen on tournaments which will be discussed in Chapter 7.

The notion of robust expansion can also be defined in the graph setting: Let $0 < \nu \leq \tau < 1$. Given a graph G and $S \subseteq V(G)$, the *ν-robust neighbourhood $RN_{\nu,G}(S)$ of S* is the set of all those vertices x of G which have at least $\nu|G|$ neighbours in S. G is called a *robust (ν, τ)-expander* if $|RN_{\nu,G}(S)| \geq |S| + \nu|G|$ for all $S \subseteq V(G)$ with $\tau|G| < |S| < (1-\tau)|G|$. We obtain the following immediate corollary of Theorem 5.13.

Corollary 5.14 *Let n_0 be a positive integer and ν, τ, η be positive constants such that $1/n_0 \ll \nu \leq \tau \ll \eta < 1$. Let G be a graph on $n \geq n_0$ vertices with $\delta(G) \geq \eta n$ which is a robust (ν, τ)-expander. Then G contains a Hamilton cycle.*

Chapter 6
Some embedding problems for oriented graphs

6.1 Powers of Hamilton cycles

A generalisation of the notion of a Hamilton cycle is that of the rth power of a Hamilton cycle. Indeed, the rth power of a Hamilton cycle C is obtained from C by adding an edge between every pair of vertices of distance at most r on C. Seymour [84] conjectured the following strengthening of Dirac's theorem.

Conjecture 6.1 (Seymour [84]) *Let G be a graph on n vertices. If $\delta(G) \geq \frac{r}{r+1}n$ then G contains the rth power of a Hamilton cycle.*

Pósa (see [25]) had earlier proposed the conjecture in the case of the square of a Hamilton cycle (that is, when $r = 2$). Komlós, Sárközy and Szemerédi [50] proved Conjecture 6.1 for sufficiently large graphs.

The notion of the rth power of a Hamilton cycle also makes sense in the digraph setting: In this case the rth power of a Hamilton cycle C is the digraph obtained from C by adding a directed edge from x to y if there is a path of length at most r from x to y on C. Bollobás and Häggkvist [12] proved that given any $\varepsilon > 0$ and any $r \in \mathbb{N}$, all sufficiently large tournaments T on n vertices with $\delta^0(T) \geq (1/4 + \varepsilon)n$ contain the rth power of a Hamilton cycle.

One would expect that the minimum semidegree threshold that ensures a digraph contains the rth power of a Hamilton cycle is the 'same' as the condition in Conjecture 6.1. But it is far less clear at first sight what to expect in the oriented case. We propose the following oriented graph analogue of Pósa's conjecture.

Conjecture 6.2 *Suppose G is an oriented graph on n vertices such that $\delta^0(G) \geq 5n/12$. Then G contains the square of a Hamilton cycle.*

The following proposition shows that, if true, Conjecture 6.2 is 'best possible'.

Proposition 6.3 *Let $n \in \mathbb{N}$ be divisible by 12. Then there is an oriented graph G on n vertices with $\delta^0(G) = 5n/12 - 1$ which does not contain the square of a Hamilton cycle.*

Proof. Let G denote the oriented graph on n vertices whose vertex set consists of the sets A, B, C, D and E where $|A| = n/6+1$, $|B| = n/6-1$, $|C| = n/3$ and $|D| = |E| = n/6$. The edge set of G is obtained as follows: Add all possible edges from $A \cup B$ to C, from C to $D \cup E$, from D to $A \cup B$ and from E to $A \cup D$. Let B, C and D all induce tournaments that are as regular as possible (so $\delta^0(G[B]) = \delta^0(G[D]) = n/12 - 1$ and $\delta^0(G[C]) = n/6 - 1$). We add edges between A and B in such a way that every vertex in A sends and receives at least $n/12 - 1$ edges to and from B, and every vertex in B sends and receives at least $n/12$ edges to and from A. Similarly, we add edges between B and E in such a way that every vertex in B sends and receives $n/12$ edges to and from E, and every vertex in E sends and receives at least $n/12 - 1$ edges to and from B. A and E are both independent sets (see Figure 6.1). So $\delta^0(G) = 5n/12 - 1$.

Assume that G contains the square of a Hamilton cycle F. Since $|B| < |E|$, showing that F must visit B between any two visits of E would yield a contradiction. Thus, consider any vertex $e \in E$. Its predecessor c_1 on F lies in $B \cup C$, so without loss of generality we may assume that $c_1 \in C$. The predecessor c_2 of c_1 on F must lie in $N^-(e) \cap N^-(c_1) \subseteq B \cup C$. So without loss of generality we may assume that $c_2 \in C$. The predecessor c_3 of c_2 on F lies in $A \cup B \cup C$. Again we are done if $c_3 \in B$, so we assume that $c_3 \in A \cup C$. Since F visits all the vertices of G we must eventually arrive at a predecessor $a \in A$ whose successor

c on F lies in C. But now the predecessor of a on F must lie in $N^-(c) \cap N^-(a) \subseteq B$, as required. □

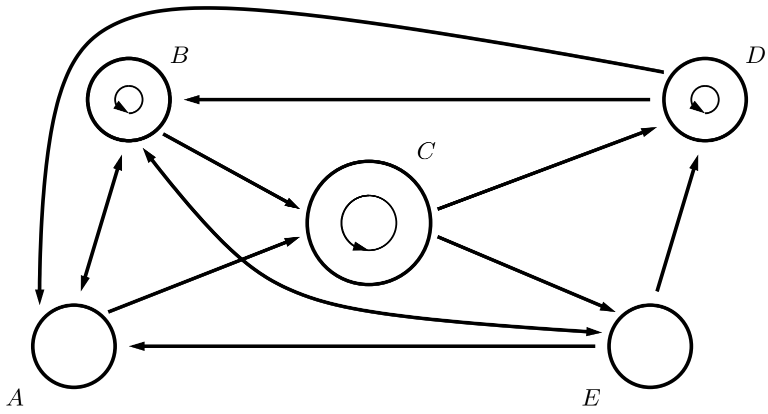

Figure 6.1: The oriented graph G from Proposition 6.3

6.2 Transitive triangle packings

In Chapter 3 we considered perfect H-packings in graphs G. It is also natural to consider the case when H and G are oriented graphs. As discussed earlier, perfect H-packings in graphs have been widely studied. However, far less is known in the oriented graph case. Keevash and Sudakov [42] showed that any oriented graph G on n vertices with $\delta^0(G) \geq (1/2-o(1))n$ contains a packing of cyclic triangles covering all but at most 3 vertices.

It is natural to ask for the minimum semidegree of an oriented graph which ensures a perfect packing of transitive triangles T_3. Note that if 3 divides $|G|$ then a necessary condition for an oriented graph G to contain a square of a Hamilton cycle is that G contains a perfect packing of transitive triangles. Let $\delta(G)$ denote the minimum degree of an oriented graph G (that is, the minimum number of edges incident to a vertex in G). The following proposition from [96] implies that a minimum semidegree as in Conjecture 6.2 ensures a perfect T_3-packing.

Proposition 6.4 (Yuster [96]) *Suppose G is an oriented graph whose order n is divisible by 3. If $\delta(G) \geq 5n/6$ then G contains a perfect T_3-packing.*

Proposition 6.4 is best possible in the sense that there are oriented graphs G whose order n is divisible by 3 and where $\delta(G) = (5n-3)/6$ but which do not contain a perfect T_3-packing. (Indeed, consider the oriented graph G on $6m+3$ vertices consisting of 3 vertex sets A, B and C where $|A| = |B| = m+1$ and $|C| = 4m+1$, and such that C induces a tournament, A sends out all possible edges to B, B sends out all possible edges to C and C sends out all possible edges to A. Then G does not contain a perfect T_3-packing since every copy of T_3 in G has at most one vertex in $A \cup B$.) We believe however that, in terms of minimum semidegree, one can improve on the bound given in Proposition 6.4.

Conjecture 6.5 *Suppose G is an oriented graph whose order n is divisible by 3. If $\delta^0(G) \geq 7n/18$ then G contains a perfect T_3-packing.*

If true, Conjecture 6.5 would characterise the minimum semidegree which ensures an oriented graph has a perfect T_3-packing.

Proposition 6.6 *Let $n \in \mathbb{N}$ be divisible by 18. Then there is an oriented graph G on n vertices with $\delta^0(G) = 7n/18 - 1$ which does not contain a perfect T_3-packing.*

Proof. Let G denote the oriented graph on n vertices whose vertex set consists of the sets A, B, C and D where $|A| = 2n/9+1$, $|B| = |C| = 2n/9$ and $|D| = n/3-1$ and whose edge set is obtained as follows: Add all possible edges from A to B, from B to C and from C to A. Let D induce a regular tournament. Partition D into two sets D' and D'' of sizes $n/6$ and $n/6-1$ respectively. Add all possible edges from D' to $B \cup C$, from A to D', from D'' to A and from $B \cup C$ to D'' (see Figure 6.2). It is easy to see that $\delta^0(G) = 7n/18-1$. Note that G does not have a perfect T_3-packing since every copy of T_3 in G must have at least one vertex in D. □

6.3 Packing transitive tournaments

Let T_k denote the transitive tournament on k vertices. In light of Conjecture 6.5 we ask the following question.

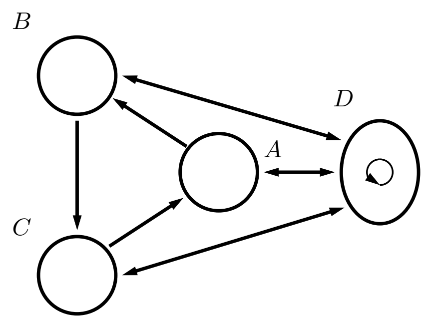

Figure 6.2: The oriented graph G from Proposition 6.6

Question 1 *What minimum semidegree condition ensures that an oriented graph contains a perfect T_k-packing?*

Recall that in the oriented graph G given in Proposition 6.6 the vertex set $A \cup B \cup C$ induces an oriented graph which does not contain a copy of T_3. This is the 'reason' why G does not contain a perfect T_3-packing. It would be of interest to establish whether the extremal examples, in terms of perfect T_k-packings, take a similar form. Thus, Question 1 is closely linked to the following question.

Question 2 *What minimum semidegree condition ensures that an oriented graph contains a copy of T_k?*

Valadkhan [94] has investigated this problem with respect to density conditions. It is easy to see that an oriented graph G on n vertices with $\delta^0(G) > n/3$ contains a copy of T_3 (and the blow-up of a cyclic triangle shows that this bound is best possible).

6.4 Perfect packings and Ramsey numbers

The *oriented tiling Ramsey number* $\overrightarrow{TR}(k)$ *of* k is the smallest integer n divisible by k such that any orientation of the complete graph K_n contains a perfect T_k-packing. Erdős (see [76]) proved the existence of these numbers. The following simple result gives a bound on the minimum degree which ensures an oriented graph G contains a perfect T_k-packing.

Proposition 6.7 *Suppose G is an oriented graph whose order n is divisible by k and such that $\delta(G) \geq (1 - \frac{1}{\overrightarrow{TR}(k)})n$. Then G contains a perfect T_k-packing.*

Sketch proof. Let $m := \overrightarrow{TR}(k)$. Consider the case when m divides n. By disregarding the orientations of the edges of G we obtain a graph G^* on n vertices with $\delta(G^*) \geq (1 - \frac{1}{m})n$. The Hajnal-Szemerédi theorem [33] implies that G^* has a perfect K_m-packing. By definition of m this implies that G has a perfect T_k-packing. If n is not divisible by m, we remove a number of vertex-disjoint copies of T_k from G until m divides $|G|$. We then proceed as before. □

Note that $\overrightarrow{TR}(3) = 6$ so Proposition 6.7 implies Proposition 6.4. In view of Proposition 6.7 it is natural to seek good upper bounds on $\overrightarrow{TR}(k)$. The *oriented Ramsey number* $\overrightarrow{R}(k)$ *of* k is the smallest integer n such that any orientation of K_n contains a copy of T_k. The following proposition gives an upper bound on $\overrightarrow{TR}(k)$ in terms of oriented Ramsey numbers.

Proposition 6.8 *Given any $k \in \mathbb{N}$, $\overrightarrow{TR}(k) \leq \overrightarrow{R}(2k-1) + (2k-1)\overrightarrow{R}(k)$.*

Proof. We use the same trick as Caro used in [15]. Let n be the largest integer divisible by k such that $n \leq \overrightarrow{R}(2k-1) + (2k-1)\overrightarrow{R}(k)$ and ℓ the largest integer divisible by k which satisfies $\ell \leq \overrightarrow{R}(k)$. Consider any orientation $\overrightarrow{K}$ of K_n. By definition of n, $\overrightarrow{K}$ contains ℓ vertex-disjoint copies of T_{2k-1}. We can cover all but ℓ of the remaining vertices of $\overrightarrow{K}$ with vertex-disjoint copies of T_k. Each of the ℓ uncovered vertices x are paired off with one of our copies T'_{2k-1} of T_{2k-1}. Since x either sends out at least k edges to T'_{2k-1} in $\overrightarrow{K}$ or receives at least k edges from T'_{2k-1} in $\overrightarrow{K}$, we have that the oriented subgraph of $\overrightarrow{K}$ induced by $V(T'_{2k-1}) \cup \{x\}$ contains a perfect T_k-packing. Thus $\overrightarrow{K}$ contains a perfect T_k-packing. □

The numbers $\overrightarrow{R}(k)$ are known for $k \leq 6$ (see [77, 81]). Sanchez-Flores [82] showed that $\overrightarrow{R}(7) \leq 54$ which by an induction argument implies that $\overrightarrow{R}(k) \leq 54 \cdot 2^{k-7}$ for $k \geq 7$ (this is the best known general upper bound on oriented Ramsey numbers). Note also that $\overrightarrow{R}(k) \leq R(k)$ where $R(k)$ denotes the Ramsey number of k.

Chapter 7
Hamilton decompositions of regular tournaments

7.1 Introduction

7.1.1 Kelly's conjecture

A Hamilton decomposition of a graph or digraph G is a set of edge-disjoint Hamilton cycles which together cover all the edges of G. The topic has a long history but some of the main questions remain open. In 1892, Walecki showed that the edge set of the complete graph K_n on n vertices has a Hamilton decomposition if n is odd (see e.g. [6, 64] for the construction). If n is even, then n is not a factor of $\binom{n}{2}$, so clearly K_n does not have such a decomposition. Walecki's result implies that a complete digraph G on n vertices has a Hamilton decomposition if n is odd. More generally, Tillson [91] proved that a complete digraph G on n vertices has a Hamilton decomposition if and only if $n \neq 4, 6$.

A tournament is an orientation of a complete graph. We say that a tournament is *regular* if every vertex has equal in- and outdegree. Thus regular tournaments contain an odd number n of vertices and each vertex has in- and outdegree $(n-1)/2$. The following beautiful conjecture of Kelly (see e.g. [8, 13, 67]), which has attracted much attention, states that every regular tournament has a Hamilton decomposition:

Conjecture 7.1 (Kelly) *Every regular tournament on n vertices can be decomposed into $(n-1)/2$ edge-disjoint Hamilton cycles.*

In this chapter we prove an approximate version of Kelly's conjecture.

Theorem 7.2 *For every $\eta > 0$ there exists an integer n_0 so that every regular tournament on $n \geq n_0$ vertices contains at least $(1/2 - \eta)n$ edge-disjoint Hamilton cycles.*

Most of the previous partial results towards Kelly's conjecture have been obtained by giving bounds on the minimum semidegree of an oriented graph which guarantees a Hamilton cycle. This approach was first used by Jackson [37], who showed that every regular tournament on at least 5 vertices contains a Hamilton cycle and a Hamilton path which are edge-disjoint. Zhang [98] then showed that every such tournament contains two edge-disjoint Hamilton cycles. Improved bounds on the value of $\delta^0(G)$ which forces a Hamilton cycle were then found by Thomassen [89], Häggkvist [31], Häggkvist and Thomason [32] as well as Kelly, Kühn and Osthus [43]. Finally, Keevash, Kühn and Osthus [41] showed that every sufficiently large oriented graph G on n vertices with $\delta^0(G) \geq (3n-4)/8$ contains a Hamilton cycle. This bound on $\delta^0(G)$ is best possible and confirmed a conjecture of Häggkvist [31]. Note that this result implies that every sufficiently large regular tournament on n vertices contains at least $n/8$ edge-disjoint Hamilton cycles. This was the best bound so far towards Kelly's conjecture. Kelly's conjecture has also been verified for $n \leq 9$ by Alspach (see the survey [10]).

We do not prove Theorem 7.2 directly, rather we prove the following stronger result. (We say that an oriented graph G on n vertices is $(\alpha \pm \eta)n$*-regular* if $\delta^0(G) \geq (\alpha - \eta)n$ and $\Delta^0(G) \leq (\alpha + \eta)n$.)

Theorem 7.3 *For every $\gamma > 0$ there exist $n_0 = n_0(\gamma)$ and $\eta = \eta(\gamma) > 0$ such that the following holds. Suppose that G is an $(\alpha \pm \eta)n$-regular oriented graph on $n \geq n_0$ vertices where $3/8 + \gamma \leq \alpha < 1/2$. Then G contains at least $(\alpha - \gamma)n$ edge-disjoint Hamilton cycles.*

We will prove Theorem 7.3 only for the case when $\alpha = 3/8 + \gamma$ since the general result follows immediately from this.

Theorem 7.3 is best possible in the sense that there are almost regular oriented graphs whose semidegrees are all close to $3n/8$ but which do not contain a Hamilton cycle. These were first found by Häggkvist [31]. However, we believe that if one requires G to be completely regular, then one can actually obtain a Hamilton decomposition of G. Note this would be a significant generalisation of Kelly's conjecture.

Conjecture 7.4 (Kühn, Osthus and Treglown [62]) *For every $\gamma > 0$ there exists $n_0 = n_0(\gamma)$ such that for all $n \geq n_0$ and all $r \geq (3/8 + \gamma)n$ each r-regular oriented graph on n vertices has a decomposition into Hamilton cycles.*

At present we do not even have any examples to rule out the possibility that one can reduce the constant $3/8$ in the above conjecture:

Question 3 *Is there a constant $c < 3/8$ such that for every sufficiently large n every cn-regular oriented graph G on n vertices has a Hamilton decomposition or at least a set of edge-disjoint Hamilton cycles covering almost all edges of G?*

It is clear that we cannot take $c < 1/4$ since there are non-Hamiltonian k-regular oriented graphs on n vertices with $k = n/4 - 1/2$ (consider a union of two regular tournaments).

7.1.2 Related results and problems

Jackson [37] introduced the following bipartite version of Kelly's conjecture (both versions are also discussed e.g. in the Handbook article by Bondy [13]). A *bipartite tournament* is an orientation of a complete bipartite graph.

Conjecture 7.5 (Jackson [37]) *Every regular bipartite tournament has a Hamilton decomposition.*

An undirected version of Conjecture 7.5 was proved independently by Auerbach and Laskar [7], as well as Hetyei [36].

Kelly's conjecture has been generalised in several directions. For instance, given an oriented graph G, define its *excess* by

$$\mathrm{ex}(G) := \sum_{v \in V(G)} \max\{d^+(v) - d^-(v), 0\},$$

where $d^+(v)$ denotes the number of outneighbours of the vertex v, and $d^-(v)$ the number of its inneighbours. Pullman (see e.g. Conjecture 8.25 in [13]) conjectured that if G is an oriented graph such that $d^+(v) + d^-(v) = d$ for all vertices v of G, where d is odd, then G has a decomposition into $\mathrm{ex}(G)$ directed paths. To see that this would imply Kelly's conjecture, let G be the oriented graph obtained from a regular tournament by deleting a vertex. Another generalisation was made by Bang-Jensen and Yeo [9], who conjectured that every k-edge-connected tournament has a decomposition into k spanning strong digraphs.

In [89], Thomassen also formulated the following weakening of Kelly's conjecture.

Conjecture 7.6 (Thomassen [89]) *If G is a regular tournament on $2k+1$ vertices and A is any set of at most $k-1$ edges of G, then $G - A$ has a Hamilton cycle.*

([89] also contains the related conjecture that for any $\ell \geq 2$, there is an $f(\ell)$ so that every strongly $f(\ell)$-connected tournament contains ℓ edge-disjoint Hamilton cycles.) Recall that in Section 5.6 we proved a result on the existence of Hamilton cycles in 'robust expander digraphs' (Theorem 5.13). In Section 7.6 we use Theorem 5.13 to prove Conjecture 7.6 for large tournaments.

Further support for Kelly's conjecture was also provided by Thomassen [90], who showed that the edges of every regular tournament on n vertices can be covered by $12n$ Hamilton cycles. In [59] Kühn and Osthus observed that one can use Theorem 7.2 to reduce this to $(1/2 + o(1))n$ Hamilton cycles. A discussion of further recent results about Hamilton cycles in directed graphs can be found in the survey [59].

It seems likely that the techniques developed in this chapter will also be useful in solving further problems. In fact, Christofides, Kühn and Osthus [18] used similar ideas to prove approximate versions of the following two long-standing conjectures of Nash-Williams [70, 71]:

Conjecture 7.7 (Nash-Williams [70]) *Let G be a $2d$-regular graph on at most $4d+1$ vertices, where $d \geq 1$. Then G has a Hamilton decomposition.*

Conjecture 7.8 (Nash-Williams [71]) *Let G be a graph on n vertices with minimum degree at least $n/2$. Then G contains $n/8 + o(n)$ edge-disjoint Hamilton cycles.*

(Actually, Nash-Williams initially formulated Conjecture 7.8 with the term $n/8$ replaced by $n/4$, but Babai found a counterexample to this.)

Another related problem was raised by Erdős (see [89]), who asked whether almost all tournaments G have at least $\delta^0(G)$ edge-disjoint Hamilton cycles. Note that an affirmative answer would not directly imply that Kelly's conjecture holds for almost all regular tournaments, which would of course be an interesting result in itself. There are also a number of corresponding questions for random undirected graphs (see e.g. [28]).

After giving an outline of the argument in the next section, we will give some useful results related to the Regularity lemma in Section 7.3. Section 7.4 contains statements and proofs of several auxiliary results, mostly on (almost) 1-factors in (almost) regular oriented graphs. The proof of Theorem 7.3 is given in Section 7.5.

7.2 Sketch of the proof of Theorem 7.3

Let $\gamma > 0$ and $\alpha := 3/8 + \gamma$. Suppose we are given an αn-regular oriented graph G on n vertices and our aim is to 'almost' decompose it into Hamilton cycles. One possible approach might be the following: first remove a spanning regular oriented subgraph H whose degree ηn satisfies $\eta \ll 1$. Let G' be the remaining oriented subgraph of G. Now consider a decomposition of G' into 1-factors $F_1, \dots, F_r$ (which clearly exists). Next, try to transform each F_i into a Hamilton cycle by removing some of its edges and adding some suitable edges of H. This is of course impossible if many of the F_i consist of many cycles. However, an auxiliary result of Frieze and Krivelevich in [28] implies that we can 'almost' decompose G' so that each 1-factor F_i consists of only a few cycles.

If H were a 'quasi-random' oriented graph, then (as in [28]) one could use it to successively 'merge' the cycles of each F_i into Hamilton cycles using a 'rotation-extension' argument: delete an edge of a cycle C of F_i to obtain a path P from a to b, say. If there is an edge of H from b to another cycle C' of F_i, then extend P to include the vertices of C' (and similarly for a). Continue until there is no such edge. Then (in H) the current endvertices of the path P have many neighbours on P. One can use this together with the quasi-randomness of H to transform P into a cycle with the same vertices as P. Now repeat this, until we have merged all the cycles into a single (Hamilton) cycle. Of course, one has to be careful to maintain the quasi-randomness of H in carrying out this 'rotation-extension' process for the successive F_i (the fact that F_i contains only few cycles is important for this).

The main problem is that G need not contain such a spanning 'quasi-random' subgraph H. So instead, in Section 7.5.1 we use Szemerédi's regularity lemma to decompose G into quasi-random subgraphs. We then choose both our 1-factors F_i and the graph H according to the structure of this decomposition. More precisely, we apply the directed version of Szemerédi's regularity lemma (Lemma 2.7) to obtain a partition of the vertices of G into a bounded number of clusters V_i so that almost all of the bipartite subgraphs spanned by ordered pairs of clusters are quasi-random. This then yields a reduced digraph R, whose vertices correspond to the clusters, with an edge from one cluster U to another cluster W if the edges from U to W in G form a quasi-random graph. (Note that R need not be oriented.) We view R as a weighted digraph whose edge weights are the densities of the corresponding ordered pair of clusters. We then obtain an unweighted multidigraph R_m from R as follows: given an edge e of R joining a cluster U to W, replace it with $K = K(e)$ copies of e, where K is approximately proportional to the density of the ordered pair (U, W). It is not hard to show that R_m is approximately regular (see Lemma 7.11). If R_m were regular, then it would have a decomposition into 1-factors, but this assumption may not be true. However, we can show that R_m can 'almost' be decomposed into 'almost' 1-factors. In other words, there exist edge-disjoint collections $\mathcal{F}_1, \dots, \mathcal{F}_r$ of vertex-disjoint cycles in R_m such that each $\mathcal{F}_i$ covers almost all of the clusters in R_m (see Lemma 7.17).

Now we choose edge-disjoint oriented spanning subgraphs $C_1, \dots, C_r$ of G so that each C_i corresponds to $\mathcal{F}_i$. For this, consider an edge e of R from U to W and suppose for example that $\mathcal{F}_1$, $\mathcal{F}_2$ and $\mathcal{F}_8$ are the only $\mathcal{F}_i$ containing copies of e in R_m. Then for each edge of G from U to W in turn, we assign it to one of C_1, C_2 and C_8 with equal probability. Then with high probability, each C_i consists of bipartite quasi-random oriented graphs which together form a disjoint union of 'blown-up' cycles. Moreover, we can arrange that all the vertices have degree close to βm (here m is the cluster size and β a small parameter which does not depend on i). We now remove a small proportion of the edges from G (and thus from each C_i) to form oriented subgraphs $H_1^+, H_1^-, H_2, H_{3,i}, H_4, H_{5,i}$ of G, where

$1 \leq i \leq r$. Ideally, we would like to show that each C_i can almost be decomposed into Hamilton cycles. Since the C_i are edge-disjoint, this would yield the required result.

One obvious obstacle is that the C_i need not be spanning subgraphs of G (because of the exceptional set V_0 returned by the regularity lemma and because the $\mathcal{F}_i$ are not spanning). So in Section 7.5.2 we add suitable edges between C_i and the leftover vertices to form edge-disjoint oriented spanning subgraphs G_i of G where every vertex has degree close to βm. (The edges of H_1^- and H_1^+ are used in this step.) But the distribution of the edges added in this step may be somewhat 'unbalanced', with some vertices of C_i sending out or receiving too many of them. In fact, as discussed at the beginning of Section 7.5.4, we cannot even guarantee that G_i has a single 1-factor. We overcome this new difficulty by adding carefully chosen further edges (from H_2 this time) to each G_i which compensate the above imbalances.

Once these edges have been added, in Section 7.5.5 we can use the max-flow min-cut theorem to almost decompose each G_i into 1-factors $F_{i,j}$. (This is one of the points where we use the fact that the C_i consist of quasi-random graphs which form a union of blown-up cycles.) Moreover, (i) the number of cycles in each of these 1-factors is not too large and (ii) most of the cycles inherit the structure of $\mathcal{F}_i$. More precisely, (ii) means that most vertices u of C_i have the following property: let U be the cluster containing u and let U^+ be the successor of U in $\mathcal{F}_i$. Then the successor u^+ of u in $F_{i,j}$ lies in U^+.

In Section 7.5.6 we can use (i) and (ii) to merge the cycles of each $F_{i,j}$ into a 1-factor $F'_{i,j}$ consisting only of a bounded number of cycles – for each cycle $\mathcal{C}$ of $\mathcal{F}_i$, all the vertices of G_i which lie in clusters of $\mathcal{C}$ will lie in the same cycle of $F'_{i,j}$. We will apply a rotation-extension argument for this, where the additional edges (i.e. those not in $F_{i,j}$) come from $H_{3,i}$. Finally, in Section 7.5.7 we will use the fact that R_m contains many short paths to merge each $F'_{i,j}$ into a single Hamilton cycle. The additional edges will come from H_4 and $H_{5,i}$ this time.

7.3 Notation and some results related to the Diregularity lemma

7.3.1 Notation

Given a multidigraph G, we denote by $N^+_G(x)$ the *multiset* of vertices where a vertex $y \in V(G)$ appears k times in $N^+_G(x)$ if G contains precisely k edges from x to y. Again, we have an analogous definition for $N^-_G(x)$. We will write $N^+(x)$ for example, if this is unambiguous. Given a vertex x of a multidigraph G, we write $d^+_G(x) := |N^+(x)|$ for the outdegree of x, $d^-_G(x) := |N^-(x)|$ for its indegree and $d(x) := d^+(x) + d^-(x)$ for its degree. The maximum of the maximum outdegree $\Delta^+(G)$ and the maximum indegree $\Delta^-(G)$ is denoted by $\Delta^0(G)$. The *minimum semidegree* $\delta^0(G)$ of G is the minimum of its minimum outdegree $\delta^+(G)$ and its minimum indegree $\delta^-(G)$. Throughout this chapter we will use $d^\pm_G(x)$, $\delta^\pm(G)$ and $N^\pm_G(x)$ as 'shorthand' notation. For example, $\delta^\pm(G) \geq \delta^\pm(H)/2$ is read as $\delta^+(G) \geq \delta^+(H)/2$ and $\delta^-(G) \geq \delta^-(H)/2$.

A *1-factor* of a multidigraph G is a collection of vertex-disjoint cycles in G which together cover all the vertices of G. Given $A, B \subseteq V(G)$, we write $e_G(A, B)$ to denote the number of edges in G with startpoint in A and endpoint in B. Given a multiset X and a set Y we define $X \cap Y$ to be the multiset where x appears as an element precisely k times in $X \cap Y$ if $x \in X$, $x \in Y$ and x appears precisely k times in X. We write $a = b \pm \varepsilon$ for $a \in [b - \varepsilon, b + \varepsilon]$.

7.3.2 A Chernoff bound

We will often use the following Chernoff bound for binomial and hypergeometric distributions (see e.g. [38, Corollary 2.3 and Theorem 2.10]). Recall that the binomial random variable with parameters (n, p) is the sum of n independent Bernoulli variables, each taking value 1 with probability p or 0 with probability $1-p$. The hypergeometric random variable X with parameters (n, m, k) is defined as follows. We let N be a set of size n, fix $S \subseteq N$ of size $|S| = m$, pick a uniformly random $T \subseteq N$ of size $|T| = k$, then define $X = |T \cap S|$.

Note that $\mathbb{E}X = km/n$.

Proposition 7.9 *Suppose X has binomial or hypergeometric distribution and $0 < a < 3/2$. Then $\mathbb{P}(|X - \mathbb{E}X| \geq a\mathbb{E}X) \leq 2e^{-\frac{a^2}{3}\mathbb{E}X}$.*

7.3.3 The Diregularity lemma

In the proof of Theorem 7.3 we will use the directed version of Szemerédi's Regularity lemma (Lemma 2.7). To prove Theorem 7.3 it will be more convenient to use the following definition of super-regularity (which is different to the definition used earlier in this thesis): Given $\varepsilon > 0$ and $d \in [0, 1)$ we say that G is (ε, d)-*super-regular* if all sets $X \subseteq A$ and $Y \subseteq B$ with $|X| \geq \varepsilon|A|$ and $|Y| \geq \varepsilon|B|$ satisfy $d(X, Y) = d \pm \varepsilon$ and, furthermore, if $d_G(a) = (d \pm \varepsilon)|B|$ for all $a \in A$ and $d_G(b) = (d \pm \varepsilon)|A|$ for all $b \in B$.

The next result shows that we can partition the set of edges of an ε-(super)-regular pair into edge-disjoint subgraphs such that each of them is still (super)-regular.

Lemma 7.10 *Let $0 < \varepsilon \ll d_0 \ll 1$ and suppose $K \geq 1$. Then there exists an integer $m_0 = m_0(\varepsilon, d_0, K)$ such that for all $d \geq d_0$ the following holds.*

(i) *Suppose that $G = (A, B)$ is an ε-regular pair of density d where $|A| = |B| = m \geq m_0$. Then there are $\lfloor K \rfloor$ edge-disjoint spanning subgraphs $S_1, \dots, S_{\lfloor K \rfloor}$ of G such that each S_i is $[\varepsilon, 4\varepsilon/K]$-regular of density $(d \pm 2\varepsilon)/K$.*

(ii) *If $K = 2$ and $G = (A, B)$ is (ε, d)-super-regular with $|A| = |B| = m \geq m_0$. then there are two edge-disjoint spanning subgraphs S_1 and S_2 of G such that each S_i is $(2\varepsilon, d/2)$-super-regular.*

Proof. We first prove (i). Suppose we have chosen m_0 sufficiently large. Initially set $E(S_i) = \emptyset$ for each $i = 1, \dots, \lfloor K \rfloor$. We consider each edge of G in turn and add it to each $E(S_i)$ with probability $1/K$, independently of all other edges of G. So the probability that xy is added to none of the S_i is $1 - \lfloor K \rfloor/K$. Moreover, $\mathbb{E}(e(S_i)) = e(G)/K = dm^2/K$.

Given $X \subseteq A$ and $Y \subseteq B$ with $|X|, |Y| \geq \varepsilon m$ we have that $|d_G(X, Y) - d| < \varepsilon$. Thus

$$\frac{1}{K}(d - \varepsilon)|X||Y| < \mathbb{E}(e_{S_i}(X, Y)) < \frac{1}{K}(d + \varepsilon)|X||Y|$$

for each i. Proposition 7.9 for the binomial distribution implies that with high probability $(d-2\varepsilon)|X||Y|/K < e_{S_i}(X,Y) < (d+2\varepsilon)|X||Y|/K$ for each $i \leq \lfloor K \rfloor$ and every $X \subseteq A$ and $Y \subseteq B$ with $|X|, |Y| \geq \varepsilon m$. Such S_i are as required in (i).

The proof of (ii) is similar. Indeed, as in (i) one can show that with high probability any $X \subseteq A$ and $Y \subseteq B$ with $|X|, |Y| \geq \varepsilon m$ satisfy $d_{S_i}(X,Y) = d/2 \pm 2\varepsilon$ (for $i = 1, 2$). Moreover, each vertex $a \in A$ satisfies $\mathbb{E}(d_{S_i}(a)) = d_G(a)/2 = (d \pm \varepsilon)m/2$ (for $i = 1, 2$) and similarly for the vertices in B. So again Proposition 7.9 for the binomial distribution implies that with high probability $d_{S_i}(a) = (d/2 \pm 2\varepsilon)m$ for all $a \in A$ and $d_{S_i}(b) = (d/2 \pm 2\varepsilon)m$ for all $b \in B$. Altogether this shows that with high probability both S_1 and S_2 are $(2\varepsilon, d/2)$-super-regular. $\square$

Suppose $0 < 1/M' \ll \varepsilon \ll \beta \ll d \ll 1$ and let G be a digraph. Let R and G' denote the reduced digraph and pure digraph respectively, obtained by applying Lemma 2.7 to G with parameters ε, d and M'. For each edge V_iV_j of R we write $d_{i,j}$ for the density of $(V_i, V_j)_{G'}$. (So $d_{i,j} \geq d$.) The *reduced multidigraph* R_m of G with parameters ε, β, d and M' is obtained from R by setting $V(R_m) := V(R)$ and adding $\lfloor d_{i,j}/\beta \rfloor$ directed edges from V_i to V_j whenever $V_iV_j \in E(R)$.

We will always consider the reduced multidigraph R_m of a digraph G whose order is sufficiently large in order to apply Lemma 7.10 to any pair $(V_i, V_j)_{G'}$ of clusters with $V_iV_j \in E(R)$. Let $K := d_{i,j}/\beta$ and $S_{i,j,1}, \dots, S_{i,j,\lfloor K \rfloor}$ be the spanning subgraphs of $(V_i, V_j)_{G'}$ obtained from Lemma 7.10. (So each $S_{i,j,k}$ is ε-regular of density $\beta \pm \varepsilon$.) Let $(V_iV_j)_1, \dots, (V_iV_j)_{\lfloor K \rfloor}$ denote the directed edges from V_i to V_j in R_m. We associate each $(V_iV_j)_k$ with the edges in $S_{i,j,k}$.

Lemma 7.11 *Let $0 < 1/M' \ll \varepsilon \ll \beta \ll d \ll c_1 \leq c_2 < 1$ and let G be a digraph of sufficiently large order n with $\delta^0(G) \geq c_1 n$ and $\Delta^0(G) \leq c_2 n$. Apply Lemma 2.7 with parameters ε, d and M' to obtain a pure digraph G' and a reduced digraph R of G. Let R_m denote the reduced multidigraph of G with parameters ε, β, d and M'. Then*

$$\delta^0(R_m) > (c_1 - 3d)\frac{|R_m|}{\beta} \text{ and } \Delta^0(R_m) < (c_2 + 2\varepsilon)\frac{|R_m|}{\beta}.$$

Note the corresponding upper bound would not hold if we considered R instead of R_m here.

Proof. Given any $V_i, V_j \in V(R)$, let $d_{i,j}$ denote the density of $(V_i, V_j)_{G'}$. Then

$$(c_1 - 2d)|R| \le \frac{(c_1 - 2d)nm}{m^2} \le \frac{\sum_{v\in V_i}\left(d^+_{G'}(v) - |V_0|\right)}{m^2} \le \sum_{V_j \in V(R)} d_{i,j} \tag{7.1}$$

by Lemma 2.7. Thus

$$\begin{aligned} d^+_{R_m}(V_i) &= \sum_{V_j\in V(R_m)} \left\lfloor \frac{d_{i,j}}{\beta} \right\rfloor \ge \frac{1}{\beta} \sum_{V_j \in V(R)} d_{i,j} - |R_m| \overset{(7.1)}{\ge} (c_1 - 2d - \beta)\frac{|R_m|}{\beta} \\ &> (c_1 - 3d)\frac{|R_m|}{\beta}. \end{aligned}$$

So indeed $\delta^+(R_m) > (c_1 - 3d)|R_m|/\beta$. Similar arguments can be used to show that $\delta^-(R_m) > (c_1 - 3d)|R_m|/\beta$ and $\Delta^0(R_m) < (c_2 + 2\varepsilon)|R_m|/\beta$. $\square$

We will also need the well-known fact that for any cycle C of the reduced multigraph R_m we can delete a small number of vertices from the clusters in C in order to ensure that each edge of C corresponds to a super-regular pair. We include a proof for completeness.

Lemma 7.12 *Let $C = V_{j_1} \dots V_{j_s}$ be a cycle in the reduced multigraph R_m as in Lemma 7.11. For each $t = 1, \dots, s$ let $(V_{j_t}V_{j_{t+1}})_{k_t}$ denote the edge of C which joins V_{j_t} to $V_{j_{t+1}}$ (where $V_{j_{s+1}} := V_{j_1}$). Then we can choose subclusters $V'_{j_t} \subseteq V_{j_t}$ of size $m' := (1 - 4\varepsilon)m$ such that $(V'_{j_t}, V'_{j_{t+1}})_{S_{j_t,j_{t+1},k_t}}$ is $(10\varepsilon, \beta)$-super-regular (for each $t = 1, \dots, s$).*

Proof. Recall that for each $t = 1, \dots, s$ the digraph S_{j_t,j_{t+1},k_t} corresponding to the edge $(V_{j_t}V_{j_{t+1}})_{k_t}$ of C is ε-regular and has density $\beta \pm \varepsilon$. So V_{j_t} contains at most $2\varepsilon m$ vertices whose outdegree in S_{j_t,j_{t+1},k_t} is either at most $(\beta - 2\varepsilon)m$ or at least $(\beta + 2\varepsilon)m$. Similarly, there are at most $2\varepsilon m$ vertices in V_{j_t} whose indegree in $S_{j_{t-1},j_t,k_{t-1}}$ is either at most $(\beta - 2\varepsilon)m$ or at least $(\beta + 2\varepsilon)m$. Let V'_{j_t} be a set of size m' obtained from V_{j_t} by deleting all these vertices (and some additional vertices if necessary). It is easy to check that $V'_{j_1}, \dots, V'_{j_t}$ are subclusters as required. $\square$

Finally, we will use the following crude version of the fact that every $[\varepsilon, \varepsilon']$-regular pair contains a subgraph of given maximum degree Δ whose average degree is close to Δ.

Lemma 7.13 *Suppose that* $0 < 1/n \ll \varepsilon', \varepsilon \ll d_0 \leq d_1 \ll 1$ *and that* (A, B) *is an* $[\varepsilon, \varepsilon']$*-regular pair of density* d_1 *with* n *vertices in each class. Then* (A, B) *contains a subgraph* H *whose maximum degree is at most* $d_0 n$ *and whose average degree is at least* $d_0 n/8$*.*

Proof. Let $A'' \subseteq A$ be the set of vertices of degree at least $2d_1 n$ and define B'' similarly. Then $|A''|, |B''| \leq \varepsilon n$. Let $A' := A \setminus A''$ and $B' := B \setminus B''$. Then (A', B') is still $[2\varepsilon, 2\varepsilon']$-regular of density at least $d_1/2$. Now consider a spanning subgraph H of (A', B') which is obtained from (A', B') by including each edge with probability $d_0/3d_1$. So the expected degree of every vertex is at most $2d_0 n/3$ and the expected number of edges of H is at least $d_0(n - \varepsilon n)^2/6$. Now apply the Chernoff bound on the binomial distribution in Proposition 7.9 to each of the vertex degrees and to the total number of edges in H to see that with high probability H has the desired properties. □

7.4 Useful results

7.4.1 1-factors in multidigraphs

Our main aim in this subsection is to show that the reduced multidigraph R_m contains a collection of 'almost' 1-factors which together cover almost all the edges of R_m (see Lemma 7.17). To prove this we will need the following collection of results.

Lemma 7.14 *Let* $0 < 1/n \ll 1/M' \ll \varepsilon \ll d \ll d' \ll \gamma < 1$*. Suppose that* G *is an oriented graph of order* n *with* $\delta^0(G) \geq (\alpha + \gamma)n$ *for some* $\alpha > 0$*. Let* R *denote the reduced digraph of* G *with parameters* ε, d *and* M' *obtained by applying Lemma 2.7. Then there exists a spanning oriented subgraph* R'_o *of* R *whose edges correspond to pairs of density at least* d' *and*

$$\delta^0(R'_o) \geq (\alpha + \gamma/2)|R'_o|.$$

Proof. Applying Lemma 2.7 to G with parameters ε, d and M' we obtain clusters $V_1, \dots, V_L$ of size m, an exceptional set V_0 and a pure digraph G'. Let R' denote the spanning subdigraph of R whose edge set consists of precisely those edges corresponding to pairs of density at least d'.

Let G'' denote the subgraph of G' 'induced' by the edges of R'. More precisely, let G'' be the subgraph of G' with vertex set $V(G)\backslash V_0$ and with edge set consisting precisely of those edges in G' which correspond to an edge in R'. Notice that given any $V_i \in V(R')$

$$\sum_{j\neq i} e_{G''}(V_i, V_j) \geq \left(\sum_{j\neq i} e_{G'}(V_i, V_j)\right) - d'm^2L. \tag{7.2}$$

We will now obtain R'_o from R' by deleting edges randomly as follows. Given an unordered pair of clusters V_i, V_j of R' we delete the edge V_iV_j (if it exists) with probability

$$\frac{e_{G''}(V_j, V_i)}{e_{G''}(V_i, V_j) + e_{G''}(V_j, V_i)}. \tag{7.3}$$

Otherwise we delete V_jV_i (if it exists). In the case when $V_iV_j, V_jV_i \notin E(R')$ then we interpret (7.3) as 0. Thus if at most one of V_iV_j and V_jV_i is an edge then with probability 1 we do not delete an edge. We repeat this for all unordered pairs of clusters V_i, V_j of R'. Thus given any $V_i \in V(R'_0)$,

$$\begin{aligned}
\mathbb{E}(d^+_{R'_0}(V_i)) &= \sum_{j\neq i} \frac{e_{G''}(V_i, V_j)}{e_{G''}(V_i, V_j) + e_{G''}(V_j, V_i)} \geq \sum_{j\neq i} \frac{e_{G''}(V_i, V_j)}{|V_i||V_j|} \\
&\overset{(7.2)}{\geq} \left(\sum_{j\neq i} \frac{e_{G'}(V_i, V_j)}{m^2}\right) - d'L \geq \frac{L}{mn}\left(\sum_{j\neq i} e_{G'}(V_i, V_j)\right) - d'L \\
&\geq \frac{L}{mn}\sum_{x\in V_i}\left(d^+_{G'}(x) - |V_0|\right) - d'L \geq \frac{L}{mn}(\alpha + 4\gamma/5)nm - d'L = (\alpha + 3\gamma/4)L.
\end{aligned}$$

Similarly $\mathbb{E}(d^-_{R'_0}(V_i)) \geq (\alpha+3\gamma/4)L$. Applying, for example, a Simple Concentration Bound (see [66]), since $L \geq M'$ and M' is sufficiently large we have that, with probability > 0, $\delta^0(R'_o) \geq (\alpha + \gamma/2)L$, as desired. $\square$

The following is an immediate consequence of Lemma 12 in [43].

Lemma 7.15 *Let R be an oriented graph on L vertices with $\delta^0(R) \geq (3/8 + \gamma)L$ for some*

$0 < \gamma \ll 1$. *If $X \subseteq V(R)$ is non-empty such that $|X| \leq (1-\gamma)L$ then $|N^+(X)| \geq |X| + \gamma L$.*

Lemma 7.16 *Let R be an oriented graph on L vertices with $\delta^0(R) \geq (3/8+\gamma)L$ for some $0 < \gamma \ll 1$. Given any distinct vertices $x, y \in V(R)$ there exists a directed path of length at most $1/\gamma$ from x to y in R.*

Proof. Let X_i be the set of vertices v for which there is a directed walk from x to v in R of length at most i. So $X_0 = \{x\}$ and $X_1 = N^+(x) \cup \{x\}$. If $|X_i| \leq (1-\gamma)L$ then Lemma 7.15 implies that $|X_{i+1}| \geq |N^+(X_i)| \geq |X_i| + \gamma L$. So certainly for $i' := \lfloor 1/\gamma \rfloor - 1$ we have that $|X_{i'}| \geq (1-\gamma)L$. But since $\delta^0(R) \geq (3/8+\gamma)L$ we have that $X_{i'+1} = V(R)$. In particular this implies that for any $y \neq x$ there is a directed path of length at most $1/\gamma$ from x to y in R. $\square$

Lemma 7.17 *Let $0 < 1/n \ll 1/M' \ll \varepsilon \ll \beta \ll \eta \ll d \ll c' \ll c \ll d' \ll \gamma \ll 1$ and define $\alpha := 3/8+\gamma$. Suppose that G is an $(\alpha n \pm \eta n)$-regular oriented graph of order n. Let R_m denote the reduced multidigraph of G with parameters ε, β, d and M' obtained by applying Lemma 2.7. Let $r := (\alpha - c)|R_m|/\beta$. Then there exist edge-disjoint collections $\mathcal{F}_1, \dots, \mathcal{F}_r$ of vertex-disjoint cycles in R_m such that each $\mathcal{F}_i$ covers all but at most $c|R_m|$ of the clusters in R_m.*

Proof. Let $L := |R_m|$. Lemma 7.11 implies that

$$\delta^0(R_m) \geq (\alpha - 4d)\frac{L}{\beta} \quad \text{and} \quad \Delta^0(R_m) \leq (\alpha + 2\eta)\frac{L}{\beta}. \tag{7.4}$$

Let R denote the reduced digraph R of G with parameters ε, d and M'. Let R'_o denote the oriented spanning subgraph of R obtained by applying Lemma 7.14 with parameter d'. So $\delta^0(R'_o) \geq (3/8+\gamma/4)L$.

First we find a set of clusters $X \subseteq V(R)$ with the following properties:

- $|X| = cL$,
- $|N^{\pm}_{R_m}(V_i) \cap X| = (\alpha \pm 5d)\frac{cL}{\beta}$ for all $V_i \in V(R_m)$,
- $|N^{\pm}_{R'_o}(V_i) \cap X| \geq (3/8+\gamma/5)cL$ for all $V_i \in V(R'_o)$.

We obtain X by choosing a set of cL clusters uniformly at random. Then each cluster V_i satisfies

$$\mathbb{E}(|N^{\pm}_{R_m}(V_i) \cap X|) = c|N^{\pm}_{R_m}(V_i)| \overset{(7.4)}{=} c(\alpha \pm 4d)\frac{L}{\beta}$$

and

$$\mathbb{E}(|N^{\pm}_{R'_o}(V_i) \cap X|) \geq (3/8 + \gamma/4)cL.$$

Proposition 7.9 for the hypergeometric distribution now implies that with nonzero probability X satisfies our desired conditions. (Recall that $N^{+}_{R_m}(V_i)$ is a multiset. Formally Proposition 7.9 does not apply to multisets. However, for each $j = 1, \dots, 1/\beta$ we can apply Proposition 7.9 to the set of all those clusters which appear at least j times in $N^{+}_{R_m}(V_i)$, and similarly for $N^{-}_{R_m}(V_i)$.)

Note that

$$d^{\pm}_{R_m \setminus X}(V_i) = (\alpha - \alpha c \pm 5d)\,\frac{L}{\beta}$$

for each $V_i \in V(R_m \setminus X)$. We now add a small number of *temporary edges* to $R_m \setminus X$ in order to turn it into an r'-regular multidigraph where $r' := (\alpha - \alpha c + 5d)\frac{L}{\beta}$. We do this as follows. As long as $R_m \setminus X$ is not r'-regular there exist $V_i, V_j \in V(R_m \setminus X)$ such that V_i has outdegree less than r' and V_j has indegree less than r'. In this case we add an edge from V_i to V_j. (Note we may have $i = j$, in which case we add a loop.)

We decompose the edge set of $R_m \setminus X$ into r' 1-factors $\mathcal{F}'_1, \dots, \mathcal{F}'_{r'}$. (To see that we can do this, consider the bipartite multigraph H where both vertex classes A, B consist of a copy of $V(R_m \setminus X)$ and we have s edges between $a \in A$ and $b \in B$ if there are precisely s edges from a to b in $R_m \setminus X$, including the temporary edges. Then H is regular and so has a perfect matching. This corresponds to a 1-factor $\mathcal{F}'_1$. Now remove the edges of $\mathcal{F}'_1$ from H and continue to find $\mathcal{F}'_2, \dots, \mathcal{F}'_{r'}$ in the same way.) Since at each cluster we added at most $20d\frac{L}{\beta}$ temporary edges, all but at most $20\sqrt{d}\frac{L}{\beta}$ of the $\mathcal{F}'_i$ contain at most $\sqrt{d}L$ temporary edges. By relabeling if necessary we may assume that $\mathcal{F}'_1, \dots, \mathcal{F}'_r$ are such 1-factors. We now remove the temporary edges from each of these 1-factors, though we still refer to the digraphs obtained in this way as $\mathcal{F}'_1, \dots, \mathcal{F}'_r$. So each $\mathcal{F}'_i$ spans $R_m \setminus X$ and consists of cycles and at most $\sqrt{d}L$ paths.

Our aim is to use the clusters in X to piece up these paths into cycles in order to obtain edge-disjoint directed subgraphs $\mathcal{F}_1, \dots, \mathcal{F}_r$ of R_m where each $\mathcal{F}_i$ is a collection of vertex-disjoint cycles and $\mathcal{F}'_i \subseteq \mathcal{F}_i$.

Let $P'_1, \dots, P'_\ell$ denote all the paths lying in one of $\mathcal{F}'_1, \dots, \mathcal{F}'_r$ (so $\ell \leq \sqrt{d}Lr \leq \sqrt{d}L^2/\beta$). Our next task is to find edge-disjoint paths and cycles $P_1, \dots, P_\ell$ of length at most $10/\gamma$ in R_m with the following properties.

(i) If P'_j consists of a single cluster $V_{j'} \in V(R)$ then P_j is a cycle consisting of at most $8/\gamma$ clusters in X as well as $V_{j'}$.

(ii) If P'_j is a path of length ≥ 1 then P_j is a path whose startpoint is the endpoint of P'_j. Similarly the endpoint of P_j is the startpoint of P'_j.

(iii) If P'_j is a path of length ≥ 1 then the internal clusters in the path P_j lie in X.

(iv) If P'_{j_1} and P'_{j_2} lie in the same $\mathcal{F}'_i$ then P_{j_1} and P_{j_2} are vertex-disjoint.

So conditions (i)–(iii) imply that $P'_j \cup P_j$ is a directed cycle for each $1 \leq j \leq \ell$. Assuming we have found such paths and cycles $P_1, \dots, P_\ell$, we define $\mathcal{F}_1, \dots, \mathcal{F}_r$ as follows. Suppose $P'_{j_1}, \dots, P'_{j_t}$ are the paths in $\mathcal{F}'_i$. Then we obtain $\mathcal{F}_i$ from $\mathcal{F}'_i$ by adding the paths and cycles $P_{j_1}, \dots, P_{j_t}$ to $\mathcal{F}'_i$. Condition (iv) ensures that the $\mathcal{F}_i$ are indeed collections of vertex-disjoint cycles.

It remains to show the existence of $P_1, \dots, P_\ell$. Suppose that for some $j \leq \ell$ we have already found $P_1, \dots, P_{j-1}$ and now need to define P_j. Consider P'_j and suppose it lies in $\mathcal{F}'_i$. Let V_a denote the startpoint of P'_j and V_b its endpoint.

We call an edge $(V_{i_1}V_{i_2})_k$ in R_m *full* if it has been used in one of $P_1, \dots, P_{j-1}$. Otherwise we call $(V_{i_1}V_{i_2})_k$ *free*. We have at most $\frac{10}{\gamma}(j-1) \leq \frac{10}{\gamma}\sqrt{d}\frac{L^2}{\beta} \ll c'\frac{L^2}{\beta}$ full edges in R_m. So at most $2\sqrt{c'}L$ clusters in X have more than $\sqrt{c'}L/\beta$ full edges incident to them in R_m. Let X_0 denote the set of such clusters in X and set $X' := X \backslash X_0$. So $|X'| \geq |X| - 2\sqrt{c'}L$.

Let $P'_{j_1}, \dots, P'_{j_t}$ denote the paths which lie in $\mathcal{F}'_i$ (so $t \leq \sqrt{d}L$). At most $\frac{10}{\gamma}\sqrt{d}L$ clusters in X' lie in the paths and cycles $P_{j_1}, \dots, P_{j_t}$ already defined. Let X_1 denote the set of such clusters in X' and set $X'' := X' \backslash X_1$. So $|X''| \geq |X| - 2\sqrt{c'}L - \frac{10}{\gamma}\sqrt{d}L \geq |X| - 3\sqrt{c'}L$. Let $R_{X''} := R'_o[X'']$. Thus $\delta^0(R_{X''}) \geq (3/8 + \gamma/5)cL - 3\sqrt{c'}L \geq (3/8 + \gamma/6)|X''|$ since $c' \ll c \ll \gamma$.

Every edge VV^+ in $R_{X''}$ corresponds to an ε-regular pair of density at least d'. Thus there are at least $\lfloor d'/\beta \rfloor$ edges in R_m associated with VV^+. We say such an edge VV^+ in $R_{X''}$ is *full* if all the edges in R_m associated with VV^+ are full. Since $X'' \subseteq X'$ each cluster $V \in X''$ is incident to at most $\sqrt{c'}L/\beta$ full edges in R_m. Thus given any cluster $V \in V(R_{X''})$, there are at most $\sqrt{c'}\frac{L}{\beta}/\lfloor d'/\beta \rfloor \leq 2\frac{\sqrt{c'}}{d'}L$ full edges in $R_{X''}$ incident to V. We remove all full edges from $R_{X''}$. So we now have that

$$\delta^0(R_{X''}) \geq (3/8 + \gamma/6)|X''| - 2\frac{\sqrt{c'}}{d'}L \geq (3/8 + \gamma/7)|X''|. \tag{7.5}$$

Since $|N^-_{R_m}(V_a) \cap X| \geq (\alpha - 5d)cL/\beta$ and $|X''| \geq |X| - 3\sqrt{c'}L$ we have that $|N^-_{R_m}(V_a) \cap X''| \geq (\alpha - 5d)cL/\beta - 3\sqrt{c'}L/\beta \geq (3/8 + \gamma/2)cL/\beta$. There are at most $20\sqrt{d}L/\beta$ full edges in R_m incident to V_a. Since $(3/8 + \gamma/2)cL/\beta - 20\sqrt{d}L/\beta \gg 1$ we can still choose a suitable cluster V_{a^-} in $N^-_{R_m}(V_a) \cap X''$ which will play the role of the inneighbour of V_a on P_j. Let $(V_{a^-}V_a)_{k_a}$ denote the corresponding free edge in R_m which will be used in P_j. A similar argument shows that we can find a cluster $V_{b^+} \neq V_{a^-}$ to play the role of the outneighbour of V_b on P_j. So $V_{b^+} \in X''$ and there is a free edge $(V_bV_{b^+})_{k_b}$ in R_m.

Using (7.5) to apply Lemma 7.16 to $R_{X''}$ we see that there exists a directed path of length at most $7/\gamma$ from V_{b^+} to V_{a^-} in $R_{X''}$. By definition of $R_{X''}$ this path corresponds to a directed path P^*_j from V_{b^+} to V_{a^-} in R_m which consists of free edges and which avoids clusters lying on the paths $P_{j_1}, \dots, P_{j_t}$. We take P_j to be the directed path or cycle $P^*_j \cup \{(V_{a^-}V_a)_{k_a}, (V_bV_{b^+})_{k_b}\}$. $\square$

7.4.2 Spanning subgraphs of super-regular pairs

Frieze and Krivelevich [28] showed that every (ε, β)-super-regular pair Γ contains a regular subgraph Γ' whose density is almost the same as that of Γ. The following lemma is an extension of this, where we can require Γ' to have a given degree sequence, as long as this degree sequence is almost regular.

Lemma 7.18 *Let $0 < 1/m \ll \varepsilon \ll \beta \ll \alpha' \ll \alpha \ll 1$. Suppose that $\Gamma = (U, V)$ is an $(\varepsilon, \beta+\varepsilon)$-super-regular pair where $|U| = |V| = m$. Define $\tau := (1-\alpha)\beta m$. Suppose we have a non-negative integer $x_i \leq \alpha'\beta m$ associated with each $u_i \in U$ and a non-negative integer $y_i \leq \alpha'\beta m$ associated with each $v_i \in V$ such that $\sum_{u_i \in U} x_i = \sum_{v_i \in V} y_i$. Then Γ contains a spanning subgraph Γ' in which $c_i := \tau - x_i$ is the degree of $u_i \in U$ and $d_i := \tau - y_i$ is the degree of $v_i \in V$.*

Proof. We first obtain a directed network N from Γ by adding a source s and a sink t. We add an edge su_i of capacity c_i for each $u_i \in U$ and an edge $v_i t$ of capacity d_i for each $v_i \in V$. We give all the edges in Γ capacity 1 and direct them from U to V.

Our aim is to show that the capacity of any cut is at least $\sum_{u_i \in U} c_i = \sum_{v_i \in V} d_i$. By the max-flow min-cut theorem this would imply that N admits a flow of value $\sum_{u_i \in U} c_i$, which by construction of N implies the existence of our desired subgraph Γ'.

So consider any (s,t)-cut $(S, \bar{S})$ where $S = \{s\} \cup S_1 \cup S_2$ with $S_1 \subseteq U$ and $S_2 \subseteq V$. Let $\bar{S}_1 := U \backslash S_1$ and $\bar{S}_2 := V \backslash S_2$. The capacity of this cut is

$$\sum_{u_i \in \bar{S}_1} c_i + \sum_{v_i \in S_2} d_i + e(S_1, \bar{S}_2)$$

and so our aim is to show that

$$e(S_1, \bar{S}_2) \geq \sum_{u_i \in S_1} c_i - \sum_{v_i \in S_2} d_i. \tag{7.6}$$

Now

$$\sum_{u_i \in S_1} c_i - \sum_{v_i \in S_2} d_i \leq |S_1|(1-\alpha)\beta m - |S_2|(1-\alpha-\alpha')\beta m \tag{7.7}$$

and similarly

$$\sum_{u_i \in S_1} c_i - \sum_{v_i \in S_2} d_i = \sum_{v_i \in \bar{S}_2} d_i - \sum_{u_i \in \bar{S}_1} c_i \leq |\bar{S}_2|(1-\alpha)\beta m - |\bar{S}_1|(1-\alpha-\alpha')\beta m. \tag{7.8}$$

By (7.7) we may assume that $|S_1| \geq (1-2\alpha')|S_2|$. (Since otherwise $\sum_{u_i \in S_1} c_i - \sum_{v_i \in S_2} d_i < 0$ and thus (7.6) is satisfied.) Similarly by (7.8) we may assume that $|\bar{S}_2| \geq (1-2\alpha')|\bar{S}_1|$. Let $\alpha^* := \alpha'/\alpha$. We now consider several cases.

Case 1. $|S_1|, |\bar{S}_2| \geq \varepsilon m$ and $|S_1| \geq (1+\alpha^*)|S_2|$.

Since Γ is $(\varepsilon, \beta+\varepsilon)$-super-regular we have that

$$\begin{aligned} e(S_1, \bar{S}_2) &\geq \beta|S_1|(m - |S_2|) \geq \beta m(|S_1| - |S_2|) \\ &= \big(|S_1|(1-\alpha)\beta m - |S_2|(1-\alpha-\alpha')\beta m\big) + \alpha\beta m|S_1| - (\alpha+\alpha')\beta m|S_2| \\ &\geq |S_1|(1-\alpha)\beta m - |S_2|(1-\alpha-\alpha')\beta m. \end{aligned}$$

(The last inequality follows since $\alpha|S_1| \geq (\alpha+\alpha')|S_2|$.) Together with (7.7) this implies (7.6).

Case 2. $|S_1|, |\bar{S}_2| \geq \varepsilon m$, $|S_1| < (1+\alpha^*)|S_2|$ and $|S_2| \leq (1-\alpha^*)m$.

Again since Γ is $(\varepsilon, \beta+\varepsilon)$-super-regular we have that

$$e(S_1, \bar{S}_2) \geq \beta|S_1|(m - |S_2|) = \beta|S_1||\bar{S}_2|. \tag{7.9}$$

As before, to prove (7.6) we will show that

$$e(S_1, \bar{S}_2) \geq |S_1|(1-\alpha)\beta m - |S_2|(1-\alpha-\alpha')\beta m.$$

Thus by (7.9) it suffices to show that $\alpha m|S_1| - |S_1||S_2| + (1-\alpha-\alpha')m|S_2| \geq 0$. We know that $|S_2|(1-\alpha-\alpha') \geq |S_1|(1-\alpha-\alpha^*)$ since $(1+\alpha^*)|S_2| > |S_1|$. Hence, $\alpha|S_1| - |S_1|(1-\alpha^*) + |S_2|(1-\alpha-\alpha') \geq 0$. So $\alpha m|S_1| - |S_1||S_2| + (1-\alpha-\alpha')m|S_2| \geq 0$ as $|S_2| \leq (1-\alpha^*)m$. So indeed (7.6) is satisfied.

Case 3. $|S_1|, |\bar{S}_2| \geq \varepsilon m$, $|S_1| < (1+\alpha^*)|S_2|$ and $|S_2| > (1-\alpha^*)m$.

By (7.8) in order to prove (7.6) it suffices to show that

$$e(S_1, \bar{S}_2) \geq |\bar{S}_2|(1-\alpha)\beta m - |\bar{S}_1|(1-\alpha-\alpha')\beta m.$$

Since (7.9) also holds in this case, this means that it suffices to show that $\alpha|\bar{S}_2|m - |\bar{S}_1||\bar{S}_2| + (1-\alpha-\alpha')|\bar{S}_1|m \geq 0$. Since $|S_1| \geq (1-2\alpha')|S_2|$ and $|S_2| > (1-\alpha^*)m$ we have that $|S_1| > (1-\alpha)m$. Thus $\alpha|\bar{S}_2|m \geq |\bar{S}_1||\bar{S}_2|$ and so indeed (7.6) holds.

Case 4. $|S_1| < \varepsilon m$ and $|\bar{S}_2| \geq \varepsilon m$.

Since $|S_1| \geq (1-2\alpha')|S_2|$ we have that $|S_2| \leq 2\varepsilon m$. Hence,

$$e(S_1, \bar{S}_2) \geq \beta m|S_1| - |S_1||S_2| \geq (\beta - 2\varepsilon)m|S_1| \geq (1-\alpha)\beta m|S_1|$$

and so by (7.7) we see that (7.6) is satisfied, as desired.

Case 5. $|S_1| \geq \varepsilon m$ and $|\bar{S}_2| < \varepsilon m$.

Similarly as in Case 4 it follows that $e(S_1, \bar{S}_2) \geq (1-\alpha)\beta m|\bar{S}_2|$ and so by (7.8) we see that (7.6) is satisfied, as desired.

Note that we have considered all possible cases since we cannot have that $|S_1|, |\bar{S}_2| < \varepsilon m$. Indeed, if $|S_1|, |\bar{S}_2| < \varepsilon m$ then $|S_2| \geq (1-\varepsilon)m$ and as $|S_1| \geq (1-2\alpha')|S_2|$ this implies $|S_1| \geq (1-2\alpha')(1-\varepsilon)m$, a contradiction. □

7.4.3 Special 1-factors in graphs and digraphs

It is easy to see that every regular oriented graph G contains a 1-factor. The following result states that if G is also dense, then (i) we can guarantee a 1-factor with few cycles. Such 1-factors have the advantage that we can transform them into a Hamilton cycle by adding/deleting a comparatively small number of edges. (ii) implies that even if G contains a sparse 'bad' subgraph H, then there will be a 1-factor which does not contain 'too many' edges of H.

Lemma 7.19 *Let $0 < \theta_1, \theta_2, \theta_3 < 1/2$ and $\theta_1/\theta_3 \ll \theta_2$. Let G be a ρ-regular oriented graph whose order n is sufficiently large and where $\rho := \theta_3 n$. Suppose $A_1, \dots, A_{5n}$ are sets of vertices in G with $a_i := |A_i| \geq n^{1/2}$. Let H be an oriented subgraph of G such that $d^{\pm}_H(x) \leq \theta_1 n$ for all $x \in A_i$ (for each i). Then G has a 1-factor F such that* *(i) F contains at most $n/(\log n)^{1/5}$ cycles;*

(ii) For each i, at most $\theta_2 a_i$ edges of $H \cap F$ are incident to A_i.

To prove this result we will use ideas similar to those used by Frieze and Krivelevich [28]. In particular, we will use the following bounds on the number of perfect matchings in a bipartite graph.

Theorem 7.20 *Suppose that B is a bipartite graph whose vertex classes have size n and $d_1, \dots, d_n$ are the degrees of the vertices in one of these vertex classes. Let $\mu(B)$ denote the number of perfect matchings in B. Then*

$$\mu(B) \leq \prod_{k=1}^{n} (d_k!)^{1/d_k}.$$

Furthermore, if B is ρ-regular then

$$\mu(B) \geq \left(\frac{\rho}{n}\right)^n n!.$$

The upper bound in Theorem 7.20 was proved by Brégman [14]. The lower bound is a consequence of the Van der Waerden conjecture which was proved independently by Egorychev [24] and Falikman [27].

We will deduce (i) from the following result in [55], which in turn is similar to Lemma 2 in [28].

Lemma 7.21 *For all $\theta \leq 1$ there exists $n_0 = n_0(\theta)$ such that the following holds. Let B be a θn-regular bipartite graph whose vertex classes U and W satisfy $|U| = |W| =: n \geq n_0$. Let M_1 be any perfect matching from U to W which is disjoint from B. Let M_2 be a perfect matching chosen uniformly at random from the set of all perfect matchings in B. Let $F = M_1 \cup M_2$ be the resulting 2-factor. Then the probability that F contains more than $n/(\log n)^{1/5}$ cycles is at most e^{-n}.*

Proof of Lemma 7.19. Consider the ρ-regular bipartite graph B whose vertex classes V_1, V_2 are copies of $V(G)$ and where $x \in V_1$ is joined to $y \in V_2$ if xy is a directed edge in G.

Note that every perfect matching in B corresponds to a 1-factor of G and vice versa. Let $\mu(B)$ denote the number of perfect matchings of B. Then

$$\mu(B) \geq \left(\frac{\rho}{n}\right)^n n! \geq \left(\frac{\rho}{n}\right)^n \left(\frac{n}{e}\right)^n = \left(\frac{\rho}{e}\right)^n \tag{7.10}$$

by Theorem 7.20. Here we have also used Stirling's formula which implies that for sufficiently large m,

$$\left(\frac{m}{e}\right)^m \leq m! \leq \left(\frac{m}{e}\right)^{m+1}. \tag{7.11}$$

We now count the number $\mu_i(G)$ of 1-factors of G which contain more than $\theta_2 a_i$ edges of H which are incident to A_i. Note that

$$\mu_i(G) \leq \binom{2a_i}{\theta_2 a_i} (\theta_1 n)^{\theta_2 a_i} (\rho!)^{(n-\theta_2 a_i)/\rho}. \tag{7.12}$$

Indeed, the term $\binom{2a_i}{\theta_2 a_i}(\theta_1 n)^{\theta_2 a_i}$ in (7.12) gives an upper bound for the number of ways we can choose $\theta_2 a_i$ edges from H which are incident to A_i such that no two of these edges have the same startpoint and no two of these edges have the same endpoint. The term $(\rho!)^{(n-\theta_2 a_i)/\rho}$ in (7.12) uses the upper bound in Theorem 7.20 to give a bound on the number of 1-factors in G containing $\theta_2 a_i$ fixed edges. Now

$$(\rho!)^{(n-\theta_2 a_i)/\rho} \overset{(7.11)}{\leq} \left(\frac{\rho}{e}\right)^{(1+1/\rho)(n-\theta_2 a_i)} \leq \left(\frac{\rho}{e}\right)^{n-\theta_2 a_i + 1/\theta_3} \tag{7.13}$$

since $\rho = \theta_3 n$ and

$$\left(\frac{e}{\rho}\right)^{\theta_2 a_i - 1/\theta_3} \leq \left(\frac{2e}{\theta_3 n}\right)^{\theta_2 a_i} \tag{7.14}$$

since $a_i \geq n^{1/2}$. Furthermore,

$$\binom{2a_i}{\theta_2 a_i} \leq \frac{(2a_i)^{\theta_2 a_i}}{(\theta_2 a_i)!} \overset{(7.11)}{\leq} \left(\frac{2e}{\theta_2}\right)^{\theta_2 a_i}. \tag{7.15}$$

So by (7.12) we have that

$$\mu_i(G) \overset{(7.13),(7.15)}{\leq} \left(\frac{2e}{\theta_2}\right)^{\theta_2 a_i} (\theta_1 n)^{\theta_2 a_i} \left(\frac{\rho}{e}\right)^{n-\theta_2 a_i+1/\theta_3}$$
$$\overset{(7.14)}{\leq} \left(\frac{2e}{\theta_2}\theta_1 n \frac{2e}{\theta_3 n}\right)^{\theta_2 a_i} \left(\frac{\rho}{e}\right)^n \overset{(7.10)}{\leq} \left(\frac{4e^2\theta_1}{\theta_2\theta_3}\right)^{\theta_2 a_i} \mu(B) \ll \frac{\mu(B)}{5n}$$

since $\theta_1/\theta_3 \ll \theta_2$, $a_i \geq n^{1/2}$ and n is sufficiently large.

Now we apply Lemma 7.21 to B where M_1 is the identity matching (i.e. every vertex in V_1 is matched to its copy in V_2). Then a cycle of length 2ℓ in $M_1 \cup M_2$ corresponds to a cycle of length ℓ in G. So, since n is sufficiently large, the number of 1-factors of G containing more than $n/(\log n)^{1/5}$ cycles is at most $e^{-n}\mu(B)$. So there exists a 1-factor F of G which satisfies (i) and (ii). $\square$

7.4.4 Rotation-Extension lemma

The following lemma will be a useful tool when transforming 1-factors into Hamilton cycles. Given such a 1-factor F, we will obtain a path P by cutting up and connecting several cycles in F (as described in the proof sketch in Section 7.2). We will then apply the lemma to obtain a cycle C containing precisely the vertices of P.

Lemma 7.22 *Let $0 < 1/m \ll \varepsilon \ll \gamma < 1$. Let G be an oriented graph on $n \geq 2m$ vertices. Suppose that U and V are disjoint subsets of $V(G)$ of size m with the following property:*

$$\text{If } S \subseteq U,\ T \subseteq V \text{ are such that } |S|, |T| \geq \varepsilon m \text{ then } e_G(S,T) \geq \gamma |S||T|/2. \tag{7.16}$$

Suppose that $P = u_1 \dots u_k$ is a directed path in G where $u_1 \in V$ and $u_k \in U$. Let X denote the set of inneighbours u_i of u_1 which lie on P so that $u_i \in U$ and $u_{i+1} \in V$. Similarly let Y denote the set of outneighbours u_i of u_k which lie on P so that $u_i \in V$ and $u_{i-1} \in U$. Suppose that $|X|, |Y| \geq \gamma m$. Then there exists a cycle C in G containing precisely the vertices of P such that $|E(C) \backslash E(P)| \leq 5$. Furthermore, $E(P) \backslash E(C)$ consists of edges from X to X^+ and edges from Y^- to Y. (Here X^+ is the set of successors of vertices in X on P and Y^- is the set of predecessors of vertices in Y on P.)

Proof. Clearly we may assume that $u_ku_1 \notin E(G)$. Let X_1 denote the set of the first $\gamma m/2$ vertices in X along P and X_2 the set of the last $\gamma m/2$ vertices in X along P. We define Y_1 and Y_2 analogously. So $X_1, X_2 \subseteq U$ and $Y_1, Y_2 \subseteq V$. We have two cases to consider.

Case 1. All the vertices in X_1 precede those in Y_2 along P.

Partition $X_1 = X_{11} \cup X_{12}$ where X_{11} denotes the set of the first $\gamma m/4$ vertices in X_1 along P. We partition Y_2 into Y_{21} and Y_{22} analogously. Let X_{12}^+ denote the set of successors on P of the vertices in X_{12} and Y_{21}^- the set of predecessors of the vertices in Y_{21}. So $X_{12}^+ \subseteq V$ and $Y_{21}^- \subseteq U$. Further define

- $X'_{11} := \{u_i \mid u_{i-1} \in X_{11}$ and $\exists$ edge from u_{i-1} to $X_{12}^+\}$ and
- $Y'_{22} := \{u_i \mid u_{i+1} \in Y_{22}$ and $\exists$ edge from Y_{21}^- to $u_{i+1}\}$.

So $X'_{11} \subseteq V$ and $Y'_{22} \subseteq U$.

From (7.16) it follows that $|X'_{11}| \geq \frac{(\gamma/2)(\gamma m/4)|X_{12}^+|}{|X_{12}^+|} \geq \varepsilon m$ and similarly $|Y'_{22}| \geq \varepsilon m$. Since $X'_{11} \subseteq V$ and $Y'_{22} \subseteq U$, by (7.16) G contains an edge $u_{i'}u_i$ from Y'_{22} to X'_{11}. Since $u_i \in X'_{11}$, by definition of X'_{11} it follows that G contains an edge $u_{i-1}u_j$ for some $u_j \in X_{12}^+$. Likewise, since $u_{i'} \in Y'_{22}$, there is an edge $u_{j'}u_{i'+1}$ for some $u_{j'} \in Y_{21}^-$. Furthermore, $u_{j-1}u_1$ and $u_ku_{j'+1}$ are edges of G by definition of X_{12}^+ and Y_{21}^-. It is easy to check that the cycle

$$C = u_1 \dots u_{i-1}u_ju_{j+1} \dots u_{j'}u_{i'+1}u_{i'+2} \dots u_ku_{j'+1}u_{j'+2} \dots u_{i'}u_iu_{i+1} \dots u_{j-1}u_1$$

has the required properties (see Figure 7.1). For example, $E(P)\backslash E(C)$ consists of the edges $u_{i-1}u_i$, $u_{j-1}u_j$, $u_{j'}u_{j'+1}$ and $u_{i'}u_{i'+1}$. The former two edges go from X to X^+ and the latter two from Y^- to Y.

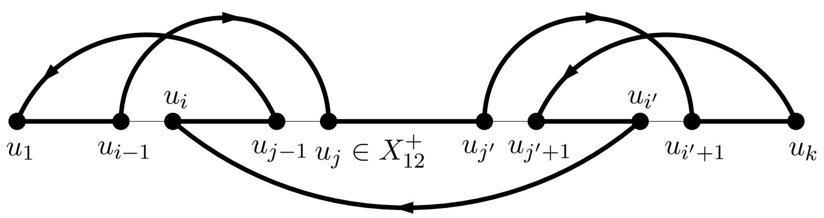

Figure 7.1: The cycle C from Case 1

Case 2. All the vertices in Y_1 precede those in X_2 along P.

Let Y_1^- be the predecessors of the vertices in Y_1 and X_2^+ the successors of the vertices in X_2 on P. So $|Y_1^-| = |X_2^+| = \gamma m/2$ and $Y_1^- \subseteq U$ and $X_2^+ \subseteq V$. Thus by (7.16) there exists an edge $u_i u_j \in E(G)$ from Y_1^- to X_2^+. Again, it is easy to check that the cycle

$$C = u_1 \dots u_i u_j u_{j+1} \dots u_k u_{i+1} u_{i+2} \dots u_{j-1} u_1$$

has the desired properties. □

7.4.5 Shifted walks

Suppose R is a digraph and F is a collection of vertex-disjoint cycles with $V(F) \subseteq V(R)$. A *closed shifted walk W in R with respect to F* is a walk in $R \cup F$ of the form

$$W = c_1^+ C_1 c_1 c_2^+ C_2 c_2 \dots c_{s-1}^+ C_{s-1} c_{s-1} c_s^+ C_s c_s c_1^+,$$

where

- $\{C_1, \dots, C_s\}$ is the set of all cycles in F;
- c_i lies on C_i and c_i^+ is the successor of c_i on C_i for each $1 \leq i \leq s$;
- $c_i c_{i+1}^+$ is an edge of R (here $c_{s+1}^+ := c_1^+$).

Note that the cycles $C_1, \dots, C_s$ are not necessarily distinct. If a cycle C_i in F appears exactly t times in W we say that C_i is *traversed t times*. Note that a closed shifted walk W has the property that for every cycle C of F, every vertex of C is visited the same number of times by W. The next lemma will be used in Section 7.5.7 to combine cycles of G which correspond to different cycles of F into a single (Hamilton) cycle. Shifted walks were introduced in [43], where they were used for a similar purpose.

Lemma 7.23 *Let $0 < 1/n \ll 1/M' \ll \varepsilon \ll d \ll c \ll d' \ll \gamma \ll 1$. Suppose that G is an oriented graph of order n with $\delta^0(G) \geq (3/8 + \gamma)n$. Let R denote the reduced digraph* *of G with parameters ε, d and M' obtained by applying Lemma 2.7 and set $L := |R|$. Let R'_o denote the spanning oriented subgraph of R obtained by applying Lemma 7.14 to R with parameter d'. Suppose F is a collection of vertex-disjoint cycles with $V(F) \subseteq V(R'_o)$ and $|V(F)| \geq (1-c)L$. Then R'_o contains a closed shifted walk with respect to F so that each cycle C in F is traversed at most L/γ times.*

Proof. From Lemma 7.14 we know that $\delta^0(R'_o) \geq (3/8 + \gamma/2)L$. Let $R'_F := R'_o[V(F)]$. So $\delta^0(R'_F) \geq (3/8+\gamma/2)L - cL \geq (3/8+\gamma/3)L$. Arguing in a similar fashion to the proof of Corollary 15 in [43] we obtain a closed shifted walk W in R'_F with respect to F which traverses each cycle in F at most $|R'_F|/\gamma \leq L/\gamma$ times. Since $R'_F \subseteq R'_o$, W is also a closed shifted walk in R'_o with respect to F, as desired. □

7.5 Proof of Theorem 7.3

7.5.1 Applying the Diregularity lemma

Without loss of generality we may assume that $0 < \gamma \ll 1$. Define further constants satisfying

$$0 < 1/M' \ll \varepsilon \ll \beta \ll \eta \ll d \ll c \ll c' \ll \eta_1 \ll \eta_2 \ll \eta_3 \ll \eta_4 \ll \eta_5 \ll d' \ll \eta' \ll \gamma. \tag{7.17}$$

Define $\alpha := 3/8 + \gamma$. Let G be an oriented graph of order $n \gg M'$ such that G is $(\alpha n \pm \eta n)$-regular. Apply the Diregularity lemma (Lemma 2.7) to G with parameters ε, d and M' to obtain clusters $V_1, \dots, V_L$ of size m, an exceptional set V_0, a pure digraph G' and a reduced digraph R (so $L = |R|$). Let R'_o be the oriented spanning subdigraph of R obtained by applying Lemma 7.14 to R with parameter d'. So if V_iV_j is an edge of R'_o then $(V_i, V_j)_{G'}$ has density at least d'.

Let R_m denote the reduced multidigraph of G with parameters ε, β, d and M'. For each edge V_iV_j of R let $d_{i,j}$ denote the density of the ε-regular pair $(V_i, V_j)_{G'}$. Recall that each edge $(V_iV_j)_k \in E(R_m)$ is associated with the kth spanning subgraph $S_{i,j,k}$ of $(V_i, V_j)_{G'}$

obtained by applying Lemma 7.10 with parameters $\varepsilon, d_{i,j}$ and $K := d_{i,j}/\beta$. Each $S_{i,j,k}$ is ε-regular with density $\beta \pm \varepsilon$. Lemma 7.11 implies that

$$\delta^0(R_m) \geq (\alpha - 4d)\frac{L}{\beta} \quad \text{and} \quad \Delta^0(R_m) \leq (\alpha + 2\eta)\frac{L}{\beta}. \tag{7.18}$$

Apply Lemma 7.17 to R_m in order to obtain

$$r := (\alpha - \eta')L/\beta \tag{7.19}$$

edge-disjoint collections $\mathcal{F}_1, \dots, \mathcal{F}_r$ of vertex-disjoint cycles in R_m such that each $\mathcal{F}_i$ contains all but at most cL of the clusters in R_m. Let $V_{0,i}$ denote the set of all those vertices in G which do not lie in clusters covered by $\mathcal{F}_i$. So $V_0 \subseteq V_{0,i}$ for all $1 \leq i \leq r$ and $|V_{0,i}| \leq |V_0| + cLm \leq (\varepsilon + c)n$. We now apply Lemma 7.12 to each cycle in $\mathcal{F}_i$ to obtain subclusters of size $m' := (1 - 4\varepsilon)m$ such that the edges of $\mathcal{F}_i$ now correspond to $(10\varepsilon, \beta)$-super-regular pairs. By removing one extra vertex from each cluster if necessary we may assume that m' is even. All vertices not belonging to the chosen subclusters of $\mathcal{F}_i$ are added to $V_{0,i}$. So now

$$|V_{0,i}| \leq 2cn. \tag{7.20}$$

We refer to the chosen subclusters as the clusters of $\mathcal{F}_i$ and still denote these clusters by $V_1, \dots, V_L$. (This is a slight abuse of notation since the clusters of $\mathcal{F}_i$ might be different from those of $\mathcal{F}_{i'}$.) Thus an edge $(V_{j_1}V_{j_2})_k$ in $\mathcal{F}_i$ corresponds to the $(10\varepsilon, \beta)$-super-regular pair $S'_{j_1,j_2,k} := (V_{j_1}, V_{j_2})_{S_{j_1,j_2,k}}$.

Let C_i denote the oriented subgraph of G whose vertices are all those vertices belonging to clusters in $\mathcal{F}_i$ such that for each $(V_{j_1}V_{j_2})_k \in E(\mathcal{F}_i)$ the edges between V_{j_1} and V_{j_2} are precisely all the edges in $S'_{j_1,j_2,k}$. Clearly $C_1, \dots, C_r$ are edge-disjoint.

We now define 'random' edge-disjoint oriented subgraphs H_1^+, H_1^-, H_2, $H_{3,i}$, H_4 and $H_{5,i}$ of G (for each $i = 1, \dots, r$). H_1^+ and H_1^- will be used in Section 7.5.2 to incorporate the exceptional vertices in $V_{0,i}$ into C_i. H_2 will be used to choose the skeleton walks in

Section 7.5.4. The $H_{3,i}$ will be used in Section 7.5.6 to merge certain cycles. H_4 and the $H_{5,i}$ will be used in Section 7.5.7 to find our almost decomposition into Hamilton cycles. We will choose these subgraphs to satisfy the following properties:

Properties of H_1^+ and H_1^-.

- H_1^+ is a spanning oriented subgraph of G.
- For all $x \in V(H_1^+)$, $\eta_1 n \leq d^{\pm}_{H_1^+}(x) \leq 2\eta_1 n$.
- For all $x \in V(H_1^+)$ and each $1 \leq i \leq r$, $|N^{\pm}_{H_1^+}(x) \cap V_{0,i}| \leq 5\eta_1 |V_{0,i}|$.
- H_1^- satisfies analogous properties.

Properties of H_2.

- The vertex set of H_2 consists of precisely all those vertices of G which lie in a cluster of R (i.e. $V(H_2) = V(G) \setminus V_0$).
- For each edge $(V_{j_1}V_{j_2})_k$ of R_m, H_2 contains a spanning oriented subgraph of $S_{j_1,j_2,k}$ which forms an ε-regular pair of density at least $\eta_2\beta$.
- All edges of H_2 belong to one of these ε-regular pairs.
- For all $x \in V(H_2)$, $d^{\pm}_{H_2}(x) \leq 2\eta_2 n$.

Properties of each $H_{3,i}$.

- The vertex set of $H_{3,i}$ consists of precisely all those vertices of G which lie in a cluster of $\mathcal{F}_i$ (i.e. $V(H_{3,i}) = V(G) \setminus V_{0,i}$).
- For each edge $(V_{j_1}V_{j_2})_k$ of $\mathcal{F}_i$, $H_{3,i}$ contains a spanning oriented subgraph of $S'_{j_1,j_2,k}$ which forms a $(\sqrt{\varepsilon}/2, 2\eta_3\beta)$-super-regular pair.
- All edges in $H_{3,i}$ belong to one of these pairs.
- Let H_3 denote the union of all the oriented graphs $H_{3,i}$. The last two properties together with (7.19) imply that $d^{\pm}_{H_3}(x) \leq 3\eta_3 n$ for all $x \in V(H_3)$.

Properties of H_4.

- The vertex set of H_4 consists of precisely all those vertices of G which lie in a cluster of R'_o (i.e. $V(H_4) = V(G) \setminus V_0$).
- For each edge $V_{j_1}V_{j_2}$ of R'_o, $(V_{j_1}, V_{j_2})_{H_4}$ is ε-regular of density at least $\eta_4 d'$.
- All edges in H_4 belong to one of these ε-regular pairs.
- For all $x \in V(H_4)$, $d^{\pm}_{H_4}(x) \leq 2\eta_4 n$.

Properties of each $H_{5,i}$.

- The vertex set of $H_{5,i}$ consists of precisely all those vertices of G which lie in a cluster of $\mathcal{F}_i$.
- For each edge $(V_{j_1}V_{j_2})_k$ of $\mathcal{F}_i$, $H_{5,i}$ contains a spanning oriented subgraph of $S'_{j_1,j_2,k}$ which forms a $(\sqrt{\varepsilon}/2, 2\eta_5\beta)$-super-regular pair.
- All edges in $H_{5,i}$ belong to one of these pairs.
- Let H_5 denote the union of all the oriented graphs $H_{5,i}$. The last two properties together with (7.19) imply that $d^{\pm}_{H_5}(x) \leq 3\eta_5 n$ for all $x \in V(H_5)$.

Properties of each $S'_{i,j,k}$.

- For each edge $(V_{j_1}V_{j_2})_k$ of $\mathcal{F}_i$ the oriented subgraph obtained from $S'_{j_1,j_2,k}$ by removing all the edges in $H_1^+, H_1^-, H_2, \dots, H_5$ is $(\varepsilon^{1/3}, \beta_1)$-super-regular for some β_1 with

$$(1-\eta')\beta \leq \beta_1 \leq \beta. \tag{7.21}$$

The existence of H_1^+, H_1^-, H_2, $H_{3,i}$, H_4 and $H_{5,i}$ can be shown by considering suitable random subgraphs of G and applying the Chernoff bound in Proposition 7.9. For example, to show that H_1^+ exists, consider a random subgraph of G which is obtained by including each edge of G with probability $4\eta_1$. Similarly, to define H_2 choose every edge in $S_{j_1,j_2,k}$ with probability $3\eta_2/2$ (for all $S_{j_1,j_2,k}$) and argue as in the proof of Lemma 7.10. Note that since H_4 only consists of edges between pairs of clusters V_{j_1}, V_{j_2} which form an edge in R'_o, the densities of oriented subgraphs obtained from the $S'_{j_1,j_2,k}$ by deleting all the edges in $H_1^+, H_1^-, H_2, \dots, H_5$ will not be close enough to each other. Indeed, if $V_{j_1}V_{j_2} \notin E(R'_o)$, then the corresponding density will be larger. However, for such pairs we can delete approximately a further η_4-proportion of the edges to ensure this property holds. Again, the deletion is done by considering a random subgraph obtained by deleting edges with probability η_4.

We now remove the edges in $H_1^+, H_1^-, H_2, \dots, H_5$ from each C_i. We still refer to the subgraphs of C_i and $S'_{j_1,j_2,k}$ thus obtained as C_i and $S'_{j_1,j_2,k}$.

7.5.2 Incorporating $V_{0,i}$ into C_i

Our ultimate aim is to use each of the C_i as a 'framework' to piece together roughly $\beta_1 m'$ Hamilton cycles in G. In this section we will incorporate the vertices in $V_{0,i}$, together with some edges incident to these vertices, into C_i. For each $i = 1, \dots, r$, let G_i denote the oriented spanning subgraph of G obtained from C_i by adding the vertices of $V_{0,i}$. So initially G_i contains no edges with a start- or endpoint in $V_{0,i}$. We now wish to add edges to G_i so that

(i) $d^{\pm}_{G_i}(x) \geq (1 - \sqrt{c})\beta_1 m'$ where x has neighbours only in C_i, for all $x \in V_{0,i}$;

(ii) $|N^{\pm}_{G_i}(y) \cap V_{0,i}| \leq \sqrt{c}\beta_1 m'$ for all $y \in V(C_i)$;

(iii) $G_1, \dots, G_r$ are edge-disjoint.

For each $x \in V(G)$ we define $\mathcal{L}_x := \{i \mid x \in V_{0,i}\}$ and let $L_x := |\mathcal{L}_x|$. Let

$$B' := \left\{ x \in V(G) \mid L_x \geq \frac{\eta_1 n}{2\beta_1 m'} \right\}.$$

We now consider the vertices in B' and $V(G) \backslash B'$ separately.

First consider any $x \in V(G) \setminus B'$. Let $p := 2\beta_1 m'/\eta_1 n$ and consider each edge e sent out by x in H_1^+. With probability $L_x p \leq 1$ we will assign e to exactly one of the G_i with $i \in \mathcal{L}_x$. More precisely, for each $i \in \mathcal{L}_x$ we assign e to G_i with probability p. So the probability e

is not assigned to any of the G_i is $1 - L_x p \geq 0$. We randomly distribute the edges of H_1^- received by x in an analogous way amongst all the G_i with $i \in \mathcal{L}_x$.

We proceed similarly for all the vertices in $V(G) \setminus B'$, with the random choices being independent for different such vertices. Since H_1^+ and H_1^- are edge-disjoint from each other and from all the C_i, the oriented graphs obtained from $G_1, \dots, G_r$ in this way will still be edge-disjoint. Moreover, $\mathbb{E}(d^{\pm}_{G_i}(x)) \geq \eta_1 np$ and $\mathbb{E}(d^{\pm}_{G_i[V_{0,i}]}(x)) \leq |V_{0,i}|p \leq 2cnp$ for every $x \in V(G) \setminus B'$ and each $i \in \mathcal{L}_x$. Thus

$$\mathbb{E}(|N^{\pm}_{G_i}(x) \cap V(C_i)|) \geq (\eta_1 - 2c)np \geq \beta_1 m'. \tag{7.22}$$

Let $B_i := V_{0,i} \cap B'$ and $\bar{B}_i := V_{0,i} \backslash B'$. Since $|N^{\pm}_{H_1^+ \cup H_1^-}(y) \cap V_{0,i}| \leq 10\eta_1 |V_{0,i}|$ for every $y \in V(C_i)$ (by definition of H_1^+ and H_1^-) we have that

$$\mathbb{E}(|N^{\pm}_{G_i}(y) \cap \bar{B}_i|) \leq 10\eta_1 |V_{0,i}|p \overset{(7.20)}{\leq} 40c\beta_1 m'. \tag{7.23}$$

Applying the Chernoff bound in Proposition 7.9 (for the binomial distribution) for each i and summing up the error probabilities for all i we see that with nonzero probability the following properties hold:

- (7.22) implies that $|N^{\pm}_{G_i}(x) \cap V(C_i)| \geq (1 - \sqrt{c})\beta_1 m'$ for every $x \in \bar{B}_i$.
- (7.23) implies that $|N^{\pm}_{G_i}(y) \cap \bar{B}_i| \leq \sqrt{c}\beta_1 m'/2$ for every $y \in V(C_i)$.

For each i we delete all the edges with both endpoints in $V_{0,i}$ from G_i.

Having dealt with the vertices in $V(G) \setminus B'$, let us now consider any $x \in B'$. We call each edge of G with startpoint x *free* if it does not lie in any of C_i, $H_1^+, H_1^-, H_2, \dots, H_5$ (for all $i = 1, \dots, r$) and if the endpoint is not in B'. Note that

$$|B'|\frac{\eta_1 n}{2\beta_1 m'} \leq \sum_{i=1}^{r} |V_{0,i}| \overset{(7.20)}{\leq} 2crn \overset{(7.19)}{\leq} cn\frac{L}{\beta},$$

and so $|B'| \leq \frac{2cn}{\eta_1}$. So the number of free edges sent out by x is at least

$$
\begin{aligned}
&(\alpha-\eta)n-(\beta_1+\varepsilon^{1/3})m'(r-L_x)-4\eta_1 n-2\eta_2 n-3\eta_3 n-2\eta_4 n-3\eta_5 n-|B'| \\
&\overset{(7.19)}{\geq} (\alpha-\eta)n-(\beta+\varepsilon^{1/3})m'(\alpha-\eta')\frac{L}{\beta}+L_x\beta_1 m'-4\eta_5 n-\frac{2cn}{\eta_1} \\
&\overset{(7.17)}{\geq} (\alpha-\eta)n-\left(\frac{\alpha\varepsilon^{1/3}n}{\beta}+\alpha n\right)+\frac{\eta' n}{2}+L_x\beta_1 m'-5\eta_5 n \overset{(7.17)}{\geq} L_x\beta_1 m'.
\end{aligned}
$$

We consider $L_x\beta_1 m'$ of these free edges sent out by x and distribute them randomly amongst all the G_i with $i \in \mathcal{L}_x$. More precisely, each such edge is assigned to G_i with probability $1/L_x$ (for each $i \in \mathcal{L}_x$). So for each $i \in \mathcal{L}_x$,

$$
\mathbb{E}(d^+_{G_i}(x)) = \beta_1 m' \tag{7.24}
$$

and

$$
\mathbb{E}(d^+_{G_i[V_{0,i}]}(x)) \leq |V_{0,i}|\frac{1}{L_x} \overset{(7.20)}{\leq} 2cn\left(\frac{2\beta_1 m'}{\eta_1 n}\right) = \frac{4c\beta_1 m'}{\eta_1} \ll \sqrt{c}\beta_1 m'/4. \tag{7.25}
$$

We can introduce an analogous definition of a free edge at x but for edges whose endpoint is x. As above we randomly distribute $L_x\beta_1 m'$ such edges amongst all the G_i with $i \in \mathcal{L}_x$. Thus for each $i \in \mathcal{L}_x$,

$$
\mathbb{E}(d^-_{G_i}(x)) = \beta_1 m' \quad \text{and} \quad \mathbb{E}(d^-_{G_i[V_{0,i}]}(x)) \ll \sqrt{c}\beta_1 m'/4. \tag{7.26}
$$

We proceed similarly for all vertices in B', with the random choices being independent for different vertices $x \in B'$. (Note that every edge of G is free with respect to at most one vertex in B'.) Then using the lower bound on L_x for all $x \in B'$ we have

$$
\mathbb{E}(|N^{\pm}_{G_i}(y) \cap B_i|) \leq |V_{0,i}|\frac{2\beta_1 m'}{\eta_1 n} \overset{(7.20)}{\leq} \sqrt{c}\beta_1 m'/4 \tag{7.27}
$$

for each $i = 1, \dots, r$ and all $y \in V(C_i)$. As before, applying the Chernoff type bound in Proposition 7.9 for each i and summing up the failure probabilities over all i shows that with nonzero probability the following properties hold:

- (7.24)–(7.26) imply that $|N^{\pm}_{G_i}(x) \cap V(C_i)| \geq (1-\sqrt{c})\beta_1 m'$ for each $x \in B_i$.
- (7.27) implies that $|N^{\pm}_{G_i}(y) \cap B_i| \leq \sqrt{c}\beta_1 m'/2$ for each $y \in V(C_i)$.

Together with the properties of G_i established after choosing the edges at the vertices in $V(G) \setminus B'$ it follows that $|N^{\pm}_{G_i}(x) \cap V(C_i)| \geq (1-\sqrt{c})\beta_1 m'$ for every $x \in V_{0,i}$ and $|N^{\pm}_{G_i}(y) \cap V_{0,i}| \leq \sqrt{c}\beta_1 m'$ for every $y \in V(C_i)$. Furthermore, $G_1, \dots, G_r$ are still edge-disjoint since when dealing with the vertices in B' we only added free edges. By discarding any edges assigned to G_i which lie entirely in $V_{0,i}$ we can ensure that (i) holds. So altogether (i)–(iii) are satisfied, as desired.

7.5.3 Randomly splitting the G_i

As mentioned in the previous section we will use each of the G_i to piece together roughly $\beta_1 m'$ Hamilton cycles of G. We will achieve this by firstly adding some more special edges to each G_i (see Section 7.5.4) and then almost decomposing each G_i into 1-factors. However, in order to use these 1-factors to create Hamilton cycles we will need to ensure that no 1-factor contains a 2-path with start- and endpoint in $V_{0,i}$, and midpoint in C_i. Unfortunately G_i might contain such paths. To avoid them, we will 'randomly split' each G_i.

We start by considering a random partition of each $V \in V(\mathcal{F}_i)$. Using the Chernoff bound in Proposition 7.9 for the hypergeometric distribution one can show that there exists a partition of V into subclusters V' and V'' so that the following conditions hold:

- $|V'|, |V''| = m'/2$ for each $V \in V(\mathcal{F}_i)$.
- $|N^{\pm}_{G_i}(x) \cap \mathcal{V}'| \geq (1/2-\sqrt{c})\beta_1 m'$ and $|N^{\pm}_{G_i}(x) \cap \mathcal{V}''| \geq (1/2-\sqrt{c})\beta_1 m'$ for each $x \in V_{0,i}$. (Here $\mathcal{V}' := \bigcup_{V \in V(\mathcal{F}_i)} V'$ and $\mathcal{V}'' := \bigcup_{V \in V(\mathcal{F}_i)} V''$.)

Recall that each edge $(V_{j_1}V_{j_2})_k \in E(\mathcal{F}_i)$ corresponds to the $(\varepsilon^{1/3}, \beta_1)$-super-regular pair $S'_{j_1,j_2,k}$. Let $\beta_2 := \beta_1/2$. So

$$(1/2-\eta')\beta \overset{(7.21)}{\leq} \beta_2 \overset{(7.21)}{\leq} \beta/2. \tag{7.28}$$

Apply Lemma 7.10(ii) to obtain a partition $E'_{j_1,j_2,k}, E''_{j_1,j_2,k}$ of the edge set of $S'_{j_1,j_2,k}$ so that the following condition holds:

- The edges of $E'_{j_1,j_2,k}$ and $E''_{j_1,j_2,k}$ both induce an $(\varepsilon^{1/4}, \beta_2)$-super-regular pair which spans $S'_{j_1,j_2,k}$.

We now partition G_i into two oriented spanning subgraphs G'_i and G''_i as follows.

- The edge set of G'_i is the union of all $E'_{j_1,j_2,k}$ (over all edges $(V_{j_1}V_{j_2})_k$ of $\mathcal{F}_i$) together with all the edges in G_i from $V_{0,i}$ to $\mathcal{V}'$, and all edges in G_i from $\mathcal{V}''$ to $V_{0,i}$.
- The edge set of G''_i is the union of all $E''_{j_1,j_2,k}$ (over all edges $(V_{j_1}V_{j_2})_k$ of $\mathcal{F}_i$) together with all the edges in G_i from $V_{0,i}$ to $\mathcal{V}''$, and all edges in G_i from $\mathcal{V}'$ to $V_{0,i}$.

Note that neither G'_i nor G''_i contains the type of 2-paths we wish to avoid. For each $i = 1, \dots, r$ we use Lemma 7.10(ii) to partition the edge set of each $H_{3,i}$ to obtain edge-disjoint oriented spanning subgraphs $H'_{3,i}$ and $H''_{3,i}$ so that the following condition holds:

- For each edge $(V_{j_1}V_{j_2})_k$ in $\mathcal{F}_i$, both $H'_{3,i}$ and $H''_{3,i}$ contain a spanning oriented subgraph of $S'_{j_1,j_2,k}$ which is $(\sqrt{\varepsilon}, \eta_3\beta)$-super-regular. Moreover, all edges in $H'_{3,i}$ and $H''_{3,i}$ belong to one of these pairs.

Similarly we partition the edge set of each $H_{5,i}$ to obtain edge-disjoint oriented spanning subgraphs $H'_{5,i}$ and $H''_{5,i}$ so that the following condition holds:

- For each edge $(V_{j_1}V_{j_2})_k$ in $\mathcal{F}_i$, both $H'_{5,i}$ and $H''_{5,i}$ contain a spanning oriented subgraph of $S'_{j_1,j_2,k}$ which is $(\sqrt{\varepsilon}, \eta_5\beta)$-super-regular. Moreover, all edges in $H'_{5,i}$ and $H''_{5,i}$ belong to one of these pairs.

We pair $H'_{3,i}$ and $H'_{5,i}$ with G'_i and pair $H''_{3,i}$ and $H''_{5,i}$ with G''_i. We now have $2r$ edge-disjoint oriented subgraphs of G, namely $G'_1, G''_1, \dots, G'_r, G''_r$. To simplify notation, we relabel these oriented graphs as $G_1, \dots, G_{r'}$ where

$$r' := 2r \overset{(7.19)}{=} 2(\alpha - \eta')L/\beta. \tag{7.29}$$

We similarly relabel the oriented graphs $H'_{3,1}, H''_{3,1}, \dots, H'_{3,r}, H''_{3,r}$ as $H_{3,1}, \dots, H_{3,r'}$ and relabel $H'_{5,1}, H''_{5,1}, \dots, H'_{5,r}, H''_{5,r}$ as $H_{5,1}, \dots, H_{5,r'}$ in such a way that $H_{3,i}$ and $H_{5,i}$ are the oriented graphs which we paired with G_i. For each i we still use the notation $\mathcal{F}_i$, C_i and $V_{0,i}$ in the usual way. Now (i) from Section 7.5.2 becomes

(i′) $d^{\pm}_{G_i}(x) \geq (1/2 - \sqrt{c})\beta_1 m'$ where x has neighbours only in C_i, for all $x \in V_{0,i}$,

while (ii) and (iii) remain valid.

7.5.4 Adding skeleton walks to the G_i

Note that all vertices (including the vertices of $V_{0,i}$) in each G_i now have in- and outdegree close to $\beta_2 m'$. In Section 7.5.5 our aim is to find a τ-regular oriented subgraph of G_i, where

$$\tau := (1 - \eta')\beta_2 m'. \tag{7.30}$$

However, this may not be possible: suppose for instance that $V_{0,i}$ consists of a single vertex x, $\mathcal{F}_i$ consists of 2 cycles C and C' and that all outneighbours of x lie on C and all inneighbours lie on C'. Then G_i does not even contain a 1-factor. A similar problem arises if for example $V_{0,i}$ consists of a single vertex x, $\mathcal{F}_i$ consists of a single cycle $C = V_1 \dots V_t$, all outneighbours of x lie in the cluster V_2 and all inneighbours in the cluster V_8. Note that in both situations, the edges between $V_{0,i}$ and C_i are not 'well-distributed' or 'balanced'. To overcome this problem, we add further edges to C_i which will 'balance out' the edges between C_i and $V_{0,i}$ which we added previously. These edges will be part of the skeleton walks which we define below. To motivate the definition of the skeleton walks it may be helpful to consider the second example above: Suppose that we add an edge e from V_1 to V_9. Then G_i now has a 1-factor. In general, we cannot find such an edge, but it will turn out that we can find a collection of a bounded number of edges fulfilling the same purpose.

A *skeleton walk* S in G with respect to G_i is a collection of distinct edges x_1x_2, $x_2^-x_3$, $\dots$, $x_{z-1}^-x_z$ and $x_z^-x_1$ of G with the following properties:

- $x_1 \in V_{0,i}$ and all vertices in $V(S)\backslash\{x_1\}$ lie in C_i.

- Given some $2 \le j \le z$, let $V \in V(\mathcal{F}_i)$ denote the cluster in $\mathcal{F}_i$ containing x_j and let C denote the cycle in $\mathcal{F}_i$ containing V. Then $x_j^- \in V^-$, where V^- is the predecessor of V on C.

The edges $x_2^- x_3, \dots, x_{z-1}^- x_z$ are referred to as the *internal edges* of S. We define z to be the *length* of S.

Note that whenever $\mathcal{S}$ is a union of edge-disjoint skeleton walks and V is a cluster in $\mathcal{F}_i$, then number of edges in $\mathcal{S}$ whose endpoint is in V is the same as the number of edges in $\mathcal{S}$ whose startpoint is in V^-. As indicated above, this 'balanced' property will be crucial when finding a τ-regular oriented subgraph of G_i in Section 7.5.5.

The internal edges of each skeleton walk S with respect to G_i will lie in the 'random' graph H_2 chosen in Section 7.5.1. More precisely, each of these edges will lie in a 'slice' $H_{2,i}$ of H_2 assigned to G_i. We will now partition H_2 into these 'slices' $H_{2,1}, \dots, H_{2,r'}$. To do this, recall that any edge $(V_{j_1} V_{j_2})_k$ in R_m corresponds to an ε-regular pair of density at least $\eta_2 \beta$ in H_2. Here V_{j_1} and V_{j_2} are viewed as clusters in R_m, so $|V_{j_1}| = |V_{j_2}| = m$. Apply Lemma 7.10(i) to each such pair of clusters to find edge-disjoint oriented subgraphs $H_{2,1}, \dots, H_{2,r'}$ of H_2 so that for each $H_{2,i}$ all the edges $(V_{j_1} V_{j_2})_k$ in R_m correspond to $[\varepsilon, 6\beta\varepsilon/L]$-regular pairs with density at least $(\eta_2\beta - 2\varepsilon)\beta/L \ge \eta_2\beta^2/2L$ in $H_{2,i}$.

Recall that by (i′) in Section 7.5.3 each vertex $x \in V_{0,i}$ has at least $(1/2 - \sqrt{c})\beta_1 m' \ge \tau$ outneighbours in C_i and at least $(1/2 - \sqrt{c})\beta_1 m'$ inneighbours in C_i. We pair τ of these outneighbours x^+ with distinct inneighbours x^-. For each of these τ pairs x^+, x^- we wish to find a skeleton walk with respect to G_i whose start edge is xx^+ and whose end edge is $x^- x$. We denote the union of these τ pairs $xx^+, x^- x$ of edges over all $x \in V_{0,i}$ by $\mathcal{T}_i$.

In Section 7.5.3 we partitioned each cluster $V \in V(\mathcal{F}_i)$ into subclusters V' and V''. We next show how to choose the skeleton walks for all those G_i for which each edge in G_i with startpoint in $V_{0,i}$ has its endpoint in $\mathcal{V}'$ (and so each edge in G_i with endpoint in $V_{0,i}$ has startpoint in $\mathcal{V}''$). The other case is similar, one only has to interchange $\mathcal{V}'$ and $\mathcal{V}''$.

Claim 7.24 *We can find a set $\mathcal{S}_i$ of $\tau|V_{0,i}|$ skeleton walks of length at most $20/\gamma$ with respect to G_i, one for each pair of edges in $\mathcal{T}_i$, such that $\mathcal{S}_i$ has the following properties:*

(i) *For each skeleton walk in $\mathcal{S}_i$, its internal edges all lie in $H_{2,i}$ and all these edges have their startpoint in $\mathcal{V}''$ and endpoint in $\mathcal{V}'$.*

(ii) *Any two of the skeleton walks in $\mathcal{S}_i$ are edge-disjoint.*

(iii) *Every $y \in V(C_i)$ is incident to at most $c^{1/5}\beta_2 m'$ edges belonging to the skeleton walks in $\mathcal{S}_i$.*

Note that $|\mathcal{S}_i| = |\mathcal{T}_i| = \tau|V_{0,i}| \leq 2c\beta_2 m'n$ by (7.20) and (7.30). To find $\mathcal{S}_i$, we will first find so-called shadow skeleton walks (here the internal edges are edges of R_m instead of G). More precisely, a *shadow skeleton walk* S' with respect to G_i is a collection of two edges x_1x_2, $x_z^- x_1$ of G and $z-2$ edges $(X_2^- X_3)_{k_2}$, $(X_3^- X_4)_{k_3}, \dots$, $(X_{z-1}^- X_z)_{k_{z-1}}$ of R_m with the following properties:

- x_1x_2, $x_z^- x_1$ is a pair in $\mathcal{T}_i$.
- $x_2 \in X_2$, $x_z^- \in X_z^-$ and each X_j is a vertex of a cycle in $\mathcal{F}_i$ and X_j^- is the predecessor of X_j on that cycle.

We refer to the edges $(X_3^- X_4)_{k_3}, \dots$, $(X_{z-2}^- X_{z-1})_{k_{z-2}}$ as the *internal edges* of S', $(X_2^- X_3)_{k_2}$ as the *internal start edge* and $(X_{z-1}^- X_z)_{k_{z-1}}$ as the *internal end edge* of S'. The *length* of S' is z.

Note that in the second condition we slightly abused the notation: as X_j is a cluster in R_m, it only corresponds to a cluster in $\mathcal{F}_i$ (which has size m' and is a subcluster of the one in R_m). However, in order to simplify our exposition, we will use the same notation for a cluster in R_m as for the cluster in $\mathcal{F}_i$ corresponding to it.

Given a collection $\mathcal{S}'$ of shadow skeleton walks (with respect to G_i) we say an edge of R_m is *bad* if it is used at least

$$B := c^{1/4}\beta^2(m')^2/L$$

times as an internal edge in $\mathcal{S}'$. We say an edge from V to U in R_m is $(V,+)$-*bad* if it is used at least B times as an internal start edge in the shadow skeleton walks of $\mathcal{S}'$. An edge from W to V in R_m is $(V,-)$-*bad* if it is used at least B times as an internal end edge in the shadow skeleton walks of $\mathcal{S}'$. We say an edge in R_m is *very bad* if it is used at least $10B$ times as an edge in $\mathcal{S}'$.

To prove Claim 7.24 we will first prove the following result.

Claim 7.25 *We can find a collection $\mathcal{S}'_i$ of $\tau|V_{0,i}|$ shadow skeleton walks with respect to G_i, one for each of pair in $\mathcal{T}_i$, and each of length at most $20/\gamma$, such that no edge in R_m is very bad.*

In order to find the internal edges of our desired shadow skeleton walks in Claim 7.25 we will have to find certain collections of edges in a special oriented subgraph of R. One can view these as 'skeletons' of the shifted walks defined in Section 7.4.5.

Suppose R' is a digraph and F a collection of vertex-disjoint cycles with $V(F) = V(R')$ (note F doesn't have to lie in R' here). Suppose $V, W \in V(R')$. A *V-W skeleton walk S in R' with respect to F of length* $k+1$ is a collection of edges

$$VV_1, V_1^-V_2, V_2^-V_3, \dots, V_{k-1}^-V_k \text{ and } V_k^-W$$

in R' with the following properties:

- If V_i belongs to the cluster C on F then V_i^- denotes the predecessor of V on C;
- $V \notin \{V_1, V_2, \dots, V_k, V_1^-, V_2^-, \dots, V_k^-\}$.

We say a V-W skeleton walk S in *strict* if either S has length 1 or
$W \notin \{V_1, V_2, \dots, V_k, V_1^-, V_2^-, \dots, V_k^-\}$. Note that if $V = W$ then a V-W skeleton walk S must be strict.

Claim 7.26 *Suppose that $R^\diamond$ is an oriented subgraph of R where $V(R^\diamond) = L' \geq (1-\gamma)L$ and F is a collection of vertex-disjoint cycles such that $V(F) = V(R^\diamond)$. Let $\mathcal{B}$ be a set of at most $\gamma L'/8$ vertices in $R^\diamond$. Suppose that for all $V \in V(R^\diamond)\backslash\mathcal{B}$,*

$$d^{\pm}_{R^\diamond}(V) \geq (3/8+\gamma)L'. \tag{7.31}$$

Given any $V, W \in V(R^\diamond)\backslash\mathcal{B}$ there exists a strict V-W skeleton walk S in $R^\diamond$ with respect to F of length at most $2/\gamma$ such that no edge of S is incident to a vertex in $\mathcal{B}$.

Proof. We first consider the case when $V \neq W$. Let $R^\diamond\backslash\mathcal{B}$ denote the oriented subgraph of $R^\diamond$ induced by $V(R^\diamond)\backslash\mathcal{B}$. By (7.31) we have that

$$\delta^0(R^\diamond\backslash\mathcal{B}) \geq (3/8+\gamma)L' - |\mathcal{B}| \geq (3/8+\gamma/2)L''$$

where $L'' := |R^\diamond\backslash\mathcal{B}| = L' - |\mathcal{B}| \geq (1-\gamma/8)L'$. Thus by Lemma 7.15 we have that given any non-empty $X \subseteq V(R^\diamond)\backslash\mathcal{B}$ and $|X| \leq (1-\gamma/2)L''$ then

$$|N^+_{R^\diamond\backslash\mathcal{B}}(X)| \geq |X| + \gamma L''/2. \tag{7.32}$$

Let X_i denote the set of vertices $V' \in V(R^\diamond)\backslash\mathcal{B}$ where $V' \neq V$ for which there is a V-V' skeleton walk S' of length at most i for which no edge in S' is incident to a vertex in $\mathcal{B}$, and for which W doesn't play the role of one of the V_j^- in S' (i.e. W does not play the role of the startpoint of any of the edges in S').

So $X_1 = N^+_{R^\diamond\backslash\mathcal{B}}(V)$ and hence $|X_1| \geq (3/8+\gamma/2)L''$. Let X_i^- denote the set of those vertices which are predecessors of the vertices in X_i on the cycles from F but which do not lie in $\mathcal{B}\cup\{V,W\}$. Thus $|X_i^-| \geq |X_i| - |\mathcal{B}| - 2$. Note that

$$X_{i+1} = \left(N^+(X_i^-) \cup X_i\right) \setminus (\mathcal{B}\cup\{V\}). \tag{7.33}$$

Suppose that $|X_i| \leq (1-\gamma/2)L''$. If $|X_i^-| > 0$ then since $|X_i^-| \leq |X_i| \leq (1-\gamma/2)L''$ by (7.32) we have that

$$|N^+(X_i^-)| \geq |X_i^-| + \gamma L''/2 \geq |X_i| - |\mathcal{B}| - 2 + \gamma L''/2 \geq |X_i| + \gamma L''/4 + 1. \tag{7.34}$$

Since $|X_1| \gg \gamma L'$ we will have that $|X_i| \gg \gamma L'$ for all i and so $|X_i^-| \geq |X_i| - |\mathcal{B}| - 2 > 0$. Thus (7.34) holds for all $i \geq 1$ such that $|X_i| \leq (1-\gamma/2)L''$. So for such i, (7.33) and (7.34) imply that

$$|X_{i+1}| \geq |X_i| + \gamma L''/4 + 1 - |\mathcal{B}| - 1 \geq |X_i| + \gamma L''/8.$$

Since $|X_1| \geq (3/8 + \gamma/2)L''$, for $i^* := \lfloor 1/\gamma \rfloor - 1$ we must have that $|X_{i^*}| \geq (1 - \gamma/2)L''$. Thus $|X^-_{i^*}| \geq (1-\gamma/2)L'' - |\mathcal{B}| - 2 \geq (1-\gamma)L''$. Since W has at least $(3/8+\gamma/2)L'' \gg \gamma L''$ inneighbours in $R^\diamond \backslash \mathcal{B}$, $W \in N^+(X^-_{i^*})$. So there exits a V-W skeleton walk S in $R^\diamond$ of length at most $1/\gamma$ which is disjoint from $\mathcal{B}$ and for which W only appears as the endpoint of an edge in S. If we restrict S to all those edges up to and including the first edge on S containing W then we see that this forms our desired strict V-W skeleton walk.

In the case when $V = W$ we choose any $W' \in V(R^\diamond)\backslash\mathcal{B}$ such that $W' \neq V$. As above we can choose strict V-W' and strict W'-V skeleton walks S_1 and S_2 of length at most $1/\gamma$ which are both disjoint from $\mathcal{B}$. $S_1 \cup S_2$ gives us our desired strict V-W skeleton walk. □

Proof of Claim 7.25. Suppose that we have already found $\ell < \tau|V_{0,i}|$ of our desired shadow skeleton walks for G_i. Let xx^+, x^-x be a pair in $\mathcal{T}_i$ for which we have yet to define a shadow skeleton walk. We will now find such a shadow skeleton walk S'. Suppose $x^+ \in V^+$ and $x^- \in W^-$, where $V^+, W^- \in V(\mathcal{F}_i)$. Let V denote the predecessor of V^+ in $\mathcal{F}_i$ and W the successor of W^- in $\mathcal{F}_i$.

Our first aim is to find a strict V-W skeleton walk in R which will be used to define the internal edges of S'. Recall that R'_o is the oriented spanning subgraph of R obtained by applying Lemma 7.14 to R with parameter d'. From Lemma 7.14 we know that $\delta^0(R'_o) \geq (3/8 + \gamma/3)L$. Let $R'_{o,i}$ denote the oriented subgraph of R'_o induced by the clusters of $\mathcal{F}_i$. Since $\mathcal{F}_i$ contains all but at most cL of the clusters of R'_o we have that $|R'_{o,i}| \geq (1-c)L$ and

$$\delta^0(R'_{o,i}) \geq (3/8 + \gamma/3)L - cL \geq (3/8 + \gamma/4)L.$$

Given any edge $V_{a_1}V_{a_2} \in V(R'_{o,i})$ there are at least $\lfloor d'/\beta \rfloor$ edges $(V_{a_1}V_{a_2})_k$ in R_m associated with it. By definition of G_i (condition (ii) in Section 7.5.2), each $y \in V(C_i)$ has at most $\sqrt{c}\beta_1 m'$ inneighbours in $V_{0,i}$ in G_i. So the number of $(V,+)$-bad edges is at most

$$\frac{\sqrt{c}\beta_1(m')^2}{B} = \frac{\sqrt{c}\beta_1(m')^2}{c^{1/4}\beta^2(m')^2/L} = \frac{c^{1/4}\beta_1 L}{\beta^2} \overset{(7.21)}{\leq} \frac{c^{1/4}L}{\beta}. \tag{7.35}$$

We remove an edge VV_{a_1} from $R'_{o,i}$ if all edges $(VV_{a_1})_k$ in R_m associated with VV_{a_1} are $(V,+)$-bad. By (7.35) we are removing at most

$$\left(c^{1/4}L/\beta\right) / \left(\lfloor d'/\beta \rfloor\right) \leq \frac{2c^{1/4}L}{d'}$$

edges sent out by V in $R'_{o,i}$. Thus,

$$d^+_{R'_{o,i}}(V) \geq (3/8+\gamma/4)L - \frac{2c^{1/4}L}{d'} \geq (3/8+\gamma/5)L. \tag{7.36}$$

We remove an edge $V_{a_1}W$ from $R'_{o,i}$ if all edges $(V_{a_1}W)_k$ in R_m associated with $V_{a_1}W$ are $(W,-)$-bad. A similar argument as above shows that

$$d^-_{R'_{o,i}}(W) \geq (3/8+\gamma/5)L. \tag{7.37}$$

Further we now have that for all $V_{a_1} \in V(R'_{o,i})\backslash\{V,W\}$,

$$d^\pm_{R'_{o,i}}(V_{a_1}) \geq (3/8+\gamma/4)L - 2$$

and

$$d^-_{R'_{o,i}}(V), d^+_{R'_{o,i}}(W) \geq (3/8+\gamma/4)L - 1. \tag{7.38}$$

Since each of the ℓ shadow skeleton walks already defined have length at most $20/\gamma$, the number of bad edges in R_m is at most

$$\frac{20\tau|V_{0,i}|/\gamma}{B} \overset{(7.20),(7.30)}{\leq} \frac{40\beta_2 m' cn}{\gamma c^{1/4}\beta^2(m')^2/L} \leq \frac{45c^{3/4}\beta_2 L^2}{\gamma\beta^2} \overset{(7.28)}{\leq} \frac{45c^{3/4}L^2}{\gamma\beta}.$$

We say a cluster V_{a_1} in $R'_{o,i}$ is *bad* if at least $\gamma d' L/(40\beta)$ edges incident to V_{a_1} in R_m are bad. Thus the number of bad vertices in $R'_{o,i}$ is at most

$$\frac{90c^{3/4}L^2/(\gamma\beta)}{\gamma d'L/(40\beta)} = \frac{3600c^{3/4}L}{\gamma^2 d'} \leq \gamma^{10}L.$$

Let $\mathcal{B}$ denote the set of all bad vertices in $R'_{o,i}$. So $|\mathcal{B}| \leq \gamma^{10}L$. Given an edge $V_{a_1}V_{a_2}$ that is disjoint from the clusters V and W we remove it from $R'_{o,i}$ if all edges $(V_{a_1}V_{a_2})_k$ in R_m associated with $V_{a_1}V_{a_2}$ are bad. Thus if $V_{a_1} \notin \mathcal{B}$ we have removed at most

$$\frac{\gamma d'L/(40\beta)}{\lfloor d'/\beta \rfloor} \leq \gamma L/20$$

edges incident to V_{a_1} in $R'_{o,i}$. Hence, we have that for all $V_{a_1} \in V(R'_{o,i})\backslash(\mathcal{B} \cup \{V, W\})$,

$$d^{\pm}_{R'_{o,i}}(V_{a_1}) \geq (3/8 + \gamma/4)L - 2 - \gamma L/20 \geq (3/8 + \gamma/6)L. \tag{7.39}$$

So (7.36), (7.37), (7.38) and (7.39) imply that

$$d^{\pm}_{R'_{o,i}}(V_{a_1}) \geq (3/8 + \gamma/6)L$$

for all $V_{a_1} \in V(R'_{o,i})\backslash\mathcal{B}$.

Thus we can apply Claim 7.26 to obtain a strict V-W skeleton walk S in $R'_{o,i}$ with respect to $\mathcal{F}_i$ of length at most $12/\gamma$ that avoids $\mathcal{B}$. Suppose S consists of the edges

$$VV_1, V_1^- V_2, V_2^- V_3, \ldots, V_{s-1}^- V_s \text{ and } V_s^- W.$$

Then by definition of $R'_{o,i}$ there is an edge $(VV_1)_{k_1}$ in R_m that is not $(V,+)$-bad. Similarly there is an edge $(V_s^- W)_{k_{s+1}}$ in R_m that is not $(W,-)$-bad. Further, given any $2 \leq s' \leq s$, there exists an edge $(V_{s'-1}^- V_{s'})_{k_{s'}}$ in R_m which is not bad. (Note that this follows from the definition of $R'_{o,i}$ and since $V_{s'-1}^- V_{s'}$ is disjoint from V and W.) The edges

$$(VV_1)_{k_1}, (V_1^- V_2)_{k_2}, (V_2^- V_3)_{k_3}, \ldots, (V_{s-1}^- V_s)_{k_s} \text{ and } (V_s^- W)_{k_{s+1}}$$

together with xx^+, x^-x yield our desired shadow skeleton walk. We repeat this process until we have our collection $\mathcal{S}'_i$ of skeleton shadow walks. By construction no edge in R_m plays the role of an internal start edge in $\mathcal{S}'_i$ more than B times, the role of an internal end edge more than B times, and the role of an internal edge more than B times. So no edge in R_m is very bad, as desired. $\square$

We now use Claim 7.25 to prove Claim 7.24.

Proof of Claim 7.24. We apply Claim 7.25 to obtain a collection $\mathcal{S}'_i$ of shadow skeleton walks. We will replace each edge of R_m in these shadow skeleton walks with a distinct edge of $H_{2,i}$ to obtain our desired collection $\mathcal{S}_i$ of skeleton walks.

Recall that each edge $(VW)_k$ in R_m corresponds to an $[\varepsilon, 6\varepsilon\beta/L]$-regular pair of density at least $\eta_2\beta^2/2L$ in $H_{2,i}$. Thus in $H_{2,i}$ the edges from V'' to W' induce a $[3\varepsilon, 12\varepsilon\beta/L]$-regular pair of density $d_1 \geq \eta_2\beta^2/3L$. (Here V', V'' and W', W'' are the partitions of V and W chosen in Section 7.5.3.) Let $d_0 := 80B/(m'/2)^2$ and note that $d_0 \leq d_1$. So we can now apply Lemma 7.13 to $(V'', W')_{H_{2,i}}$ to obtain a subgraph $H'_{2,i}[V'', W']$ with maximum degree at most $d_0m'/2$ and at least $d_0(m'/2)^2/8 = 10B$ edges. We do this for all those edges in R_m which are used in a shadow skeleton walk in $\mathcal{S}'_i$.

Since no edge in R_m is very bad, for each $S' \in \mathcal{S}'_i$ we can replace an edge $(VW)_k$ in S' with a distinct edge e from V'' to W' lying in $H'_{2,i}[V'', W']$. Thus we obtain a collection $\mathcal{S}_i$ of skeleton walks which satisfy properties (i) and (ii) of Claim 7.24. Note that by the construction of $\mathcal{S}_i$ every vertex $y \in V(C_i)$ is incident to at most $d_0m'L/\beta \ll c^{1/5}\beta_2m'/2$ edges which play the role of an internal edge in a skeleton walk in $\mathcal{S}_i$. Condition (ii) in Section 7.5.2 implies that y is incident to at most $2\sqrt{c}\beta_1m'$ edges in $\mathcal{T}_i$. So in total y is incident to at most $c^{1/5}\beta_2m'/2 + 2\sqrt{c}\beta_1m' \leq c^{1/5}\beta_2m'$ edges of the skeleton walks in $\mathcal{S}_i$. Hence (iii) and thus the entire claim is satisfied. $\square$

We now add the edges of the skeleton walks in $\mathcal{S}_i$ to G_i. Moreover, for each $x \in V_{0,i}$ we delete all those edges at x which do not lie in a skeleton walk in $\mathcal{S}_i$.

7.5.5 Almost decomposing the G_i into 1-factors

Our aim in this section is to find a suitable collection of 1-factors in each G_i which together cover almost all the edges of G_i. In order to do this, we first choose a τ-regular spanning oriented subgraph G_i^* of G_i and then apply Lemma 7.19 to G_i^*.

We will refer to all those edges in G_i which lie in a skeleton walk in $\mathcal{S}_i$ as *red*, and all other edges in G_i as *white*. Given $V \in V(\mathcal{F}_i)$ and $x \in V$, we denote by $N_w^+(x)$ the set of all those vertices which receive a white edge from x in G_i. Similarly we denote by $N_w^-(x)$ the set of all those vertices which send out a white edge to x in G_i. So $N_w^+(x) \subseteq V^+$ and $N_w^-(x) \subseteq V^-$, where V^+ and V^- are the successor and the predecessor of V in $\mathcal{F}_i$. Note that G_i has the following properties:

(α_1) $d^{\pm}_{G_i}(x) = \tau$ for each $x \in V_{0,i}$. Moreover, x does not have any in- or outneighbours in $V_{0,i}$.

(α_2) Every path in G_i consisting of two red edges has its midpoint in $V_{0,i}$.

(α_3) For each $(V_jV_j^+)_k \in E(\mathcal{F}_i)$ the white edges in G_i from V_j to V_j^+ induce a $(\varepsilon^{1/4}, \beta_2)$-super-regular pair $(V_j, V_j^+)_{G_i}$.

(α_4) Every vertex $u \in V(C_i)$ receives at most $c^{1/5}\beta_2 m'$ red edges and sends out at most $c^{1/5}\beta_2 m'$ red edges in G_i.

(α_5) In total, the vertices in G_i lying in a cluster $V_j \in V(\mathcal{F}_i)$ send out the same number of red edges as the vertices in V_j^+ receive.

In order to find our τ-regular spanning oriented subgraph of G_i, consider any edge $(V_jV_j^+)_k \in E(\mathcal{F}_i)$. Given any $u_\ell \in V_j$, let x_ℓ denote the number of red edges sent out by u_ℓ in G_i. Similarly given any $v_\ell \in V_j^+$, let y_ℓ denote the number of red edges received by v_ℓ in G_i. By (α_4) we have that $x_\ell, y_\ell \le c^{1/5}\beta_2 m'$ and by (α_5) we have that

$$\sum_{u_\ell \in V_j} x_\ell = \sum_{v_\ell \in V_j^+} y_\ell .$$

Thus we can apply Lemma 7.18 to obtain an oriented spanning subgraph of $(V_j, V_j^+)_{G_i}$ in which each u_ℓ has outdegree $\tau - x_\ell$ and each v_ℓ has indegree $\tau - y_\ell$. We apply Lemma 7.18 to each $(V_j V_j^+)_k \in E(\mathcal{F}_i)$. The union of all these oriented subgraphs together with the red edges in G_i clearly yield a τ-regular oriented subgraph G_i^* of G_i, as desired.

We will use the following claim to almost decompose G_i^* into 1-factors with certain useful properties.

Claim 7.27 *Let G^* be a spanning ρ-regular oriented subgraph of G_i where $\rho \geq \eta'\beta_2 m'$. Then G^* contains a 1-factor F^* with the following properties:*

(i) *F^* contains at most $n/(\log n)^{1/5}$ cycles.*

(ii) *For each $V_j \in V(\mathcal{F}_i)$, F^* contains at most $c'm'$ red edges incident to vertices in V_j.*

(iii) *Let F^*_{red} denote the set of vertices which are incident to a red edge in F^*. Then $|F^*_{red} \cap N^{\pm}_{H_{3,i}}(x)| \leq 2c'\eta_3\beta m'$ for each $x \in V(C_i)$.*

(iv) *$|F^*_{red} \cap N^{\pm}_{w}(x)| \leq 2c'\beta_2 m'$ for each $x \in V(C_i)$.*

Proof. A direct application of Lemma 7.19 to G^* proves the claim. Indeed, we apply the lemma with $\theta_1 = (c^{1/5}\beta_2 m')/n$, $\theta_2 = c'$, $\theta_3 = \rho/n \geq (\eta'\beta_2 m')/n$ and with the oriented spanning subgraph of G^* whose edge set consists precisely of the red edges in G^* playing the role of H. Furthermore, the clusters in $V(\mathcal{F}_i)$ together with the sets $N^{\pm}_{w}(x)$ and $N^{\pm}_{H_{3,i}}(x)$ (for each $x \in V(C_i)$) play the role of the A_j. □

Repeatedly applying Claim 7.27 we obtain edge-disjoint 1-factors $F_{i,1}, \dots, F_{i,\psi}$ of G_i satisfying conditions (i)–(iv) of the claim, where

$$\psi := (1 - 2\eta')\beta_2 m'. \tag{7.40}$$

Our aim is now to transform each of the $F_{i,j}$ into a Hamilton cycle using the edges of $H_{3,i}$, H_4 and $H_{5,i}$.

7.5.6 Merging the cycles in $F_{i,j}$ into a bounded number of cycles

Let $D_1, \dots, D_\xi$ denote the cycles in $\mathcal{F}_i$ and define $V_G(D_k)$ to be the set of vertices in G_i which lie in clusters in the cycle D_k. In this subsection, for each i and j we will merge the cycles in $F_{i,j}$ to obtain a 1-factor $F'_{i,j}$ consisting of at most ξ cycles.

Recall from Section 7.5.5 that we call the edges of G_i which lie on a skeleton walk in $\mathcal{S}_i$ red and the non-red edges of G_i white. We call the edges of the 'random' oriented graph $H_{3,i}$ defined in Section 7.5.1 *green*. (Recall that $H_{3,i}$ was modified in Section 7.5.3.) We will use the edges from $H_{3,i}$ to obtain 1-factors $F'_{i,1}, \dots, F'_{i,\psi}$ for each G_i with the following properties:

(β_1) If $i \neq i'$ or $j \neq j'$ then $F'_{i,j}$ and $F'_{i',j'}$ are edge-disjoint.

(β_2) For each $V \in V(\mathcal{F}_i)$ all $x \in V$ which send out a white edge in $F_{i,j}$ lie on the same cycle C in $F'_{i,j}$.

(β_3) $|E(F'_{i,j}) \backslash E(F_{i,j})| \leq 6n/(\log n)^{1/5}$ for all i and j. Moreover, $E(F'_{i,j}) \backslash E(F_{i,j})$ consists of green and white edges only.

(β_4) For every edge in $F_{i,j}$ both endvertices lie on the same cycle in $F'_{i,j}$.

(β_5) All the red edges in $F_{i,j}$ still lie in $F'_{i,j}$.

Before showing the existence of 1-factors satisfying (β_1)–(β_5), we will derive two further properties (β_6) and (β_7) from them which we will use in the next subsection. So suppose that $F'_{i,j}$ is a 1-factor satisfying the above conditions. Consider any cluster $V \in V(\mathcal{F}_i)$. Claim 7.27(ii) implies that $F_{i,j}$ contains at most $c'm'$ red edges with startpoint in V. So the cycle C in $F'_{i,j}$ which contains all vertices $x \in V$ sending out a white edge in $F_{i,j}$ must contain at least $(1-c')m'$ such vertices x. In particular there are at least $(1-c')m' > c'm'$ vertices $y \in V^+$ which lie on C. So some of these vertices y send out a white edge in $F_{i,j}$. But by (β_2) this means that C contains all those vertices $y \in V^+$ which send out a white edge in $F_{i,j}$. Repeating this argument shows that C contains all vertices in $V(D_k)$ which send out a white edge in $F_{i,j}$ (here D_k is the cycle on $\mathcal{F}_i$ that contains V). Furthermore, by property (β_4), C contains all vertices in $V(D_k)$ which receive a white edge in $F_{i,j}$. By property (α_2) in Section 7.5.5 no vertex of C_i is both a startpoint of a red edge in G_i and an endpoint of a red edge in G_i. So this implies that all vertices in $V_G(D_k)$ lie on C. Thus if we obtain 1-factors $F'_{i,1}, \dots, F'_{i,\psi}$ satisfying (β_1)–(β_5) then the following conditions also hold:

(β_6) For each $j = 1, \dots, \psi$ and each $k = 1, \dots, \xi$ all the vertices in $V_G(D_k)$ lie on the same cycle in $F'_{i,j}$.

(β_7) For each $V \in V(\mathcal{F}_i)$ and each $j = 1, \dots, \psi$ at most $c'm'$ vertices in V lie on a red edge in $F'_{i,j}$.

(Condition (β_7) follows from Claim 7.27(ii) and the 'moreover' part of (β_3).)

For every i, we will define the 1-factors $F'_{i,1}, \dots, F'_{i,\psi}$ sequentially. Initially, we let $F'_{i,j} = F_{i,j}$. So the $F'_{i,j}$ satisfy all conditions except (β_2). Next, we describe how to modify $F'_{i,1}$ so that it also satisfies (β_2).

Recall from Section 7.5.3 that for each edge $(VV^+)_k$ of $\mathcal{F}_i$ the pair $(V, V^+)_{H_{3,i}}$ is $(\sqrt{\varepsilon}, \eta_3\beta)$-super-regular and thus $\delta^{\pm}(H_{3,i}) \geq (\eta_3\beta - \sqrt{\varepsilon})m' \geq \eta_3\beta m'/2$. Furthermore, whenever $V \in V(\mathcal{F}_i)$ and $x \in V$, the outneighbourhood of x in $H_{3,i}$ lies in V^+ and the inneighbourhood of x in $H_{3,i}$ lies in V^-. Let $H'_{3,i}$ denote the oriented spanning subgraph of $H_{3,i}$ whose edge set consists of those edges xy of $H_{3,i}$ for which x is not a startpoint of a red edge in our current 1-factor $F'_{i,1}$ and y is not an endpoint of a red edge in $F'_{i,1}$. Consider a white edge xy in $F'_{i,1}$. Claim 7.27(iii) implies that x sends out most $2c'\eta_3\beta m'$ green edges xz in $H_{3,i}$ which do not lie in $H'_{3,i}$. So $d^+_{H'_{3,i}}(x) \geq (1/2 - 2c')\eta_3\beta m'$. Similarly, $d^-_{H'_{3,i}}(y) \geq (1/2 - 2c')\eta_3\beta m'$. (However, if uv is a red edge in $F'_{i,1}$ then $d^+_{H'_{3,i}}(u) = d^-_{H'_{3,i}}(v) = 0$.) Thus we have the following properties of $H_{3,i}$ and $H'_{3,i}$:

(γ_1) For each $V \in V(\mathcal{F}_i)$ all the edges in $H_{3,i}$ sent out by vertices in V go to V^+.

(γ_2) If xy is a white edge in $F'_{i,1}$ then $d^+_{H'_{3,i}}(x), d^-_{H'_{3,i}}(y) \geq \eta_3\beta m'/3$.

(γ_3) Consider any $V \in V(\mathcal{F}_i)$. Let $S \subseteq V$ and $T \subseteq V^+$ be such that $|S|, |T| \geq \sqrt{\varepsilon}m'$. Then $e_{H_{3,i}}(S, T) \geq \eta_3\beta|S||T|/2$.

If $F'_{i,1}$ does not satisfy (β_2), then it contains cycles $C \neq C^*$ such that there is a cluster $V \in V(\mathcal{F}_i)$ and white edges xy on C and x^*y^* on C^* with $x, x^* \in V$ and $y, y^* \in V^+$.

We have 3 cases to consider. Firstly, we may have a green edge $xz \in E(H'_{3,i})$ such that z lies on a cycle $C' \neq C$ in $F'_{i,1}$. Then $z \in V^+$ and z is the endpoint of a white edge in $F'_{i,1}$ (by (γ_1) and the definition of $H'_{3,i}$). Secondly, there may be a green edge $wy^* \in E(H'_{3,i})$ such that w lies on a cycle $C' \neq C^*$ in $F'_{i,1}$. So here $w \in V$ is the startpoint of a white edge in $F'_{i,1}$. If neither of these cases hold, then $N^+_{H'_{3,i}}(x)$ lies on C and $N^-_{H'_{3,i}}(y^*)$ lies on C^*. Since $d^+_{H'_{3,i}}(x), d^-_{H'_{3,i}}(y^*) \geq \eta_3\beta m'/3$ by (γ_2), we can use (γ_3) to find a green edge $x'y'$ from $N^-_{H'_{3,i}}(y^*)$ to $N^+_{H'_{3,i}}(x)$. Then $x' \in V$, $y' \in V^+$, x' is the startpoint of a white edge in $F'_{i,1}$ and y' is the endpoint of a white edge in $F'_{i,1}$.

We will only consider the first of these 3 cases. The other cases can be dealt with analogously: In the second case w plays the role of x and y^* plays the role of z. In the third case x' plays the role of x and y' plays the role of z.

So let us assume that the first case holds, i.e. there is a green edge $xz \in E(H'_{3,i})$ such that z lies on a cycle $C' \neq C$ in $F'_{i,1}$ and z lies on a white edge wz on C'. Let P denote the directed path $(C \cup C' \cup \{xz\}) \setminus \{xy, wz\}$ from $y \in V^+$ to $w \in V$. Suppose that the endpoint w of P lies on a green edge $wv \in E(H'_{3,i})$ such that v lies outside P. Then $v \in V^+$ is the endpoint of a white edge uv lying on the cycle C'' in $F'_{i,1}$ which contains v. We extend P by replacing P and C'' with $(P \cup C'' \cup \{wv\}) \setminus \{uv\}$. We make similar extensions if the startpoint y of P has an inneighbour in $H'_{3,i}$ outside P. We repeat this 'extension' procedure as long as we can. Let P denote the path obtained in this way, say P joins $a \in V^+$ to $b \in V$. Note that a must be the endpoint of a white edge in $F'_{i,1}$ and b the startpoint of a white edge in $F'_{i,1}$.

We will now apply a 'rotation' procedure to close P into a cycle. By (γ_2) a has at least $\eta_3\beta m'/3$ inneighbours in $H'_{3,i}$ and b has at least $\eta_3\beta m'/3$ outneighbours in $H'_{3,i}$ and all these in- and outneighbours lie on P since we could not extend P any further. Let $X := N^-_{H'_{3,i}}(a)$ and $Y := N^+_{H'_{3,i}}(b)$. So $|X|, |Y| \geq \eta_3\beta m'/3$ and $X \subseteq V$ and $Y \subseteq V^+$ by (γ_1). Moreover, whenever $c \in X$ and c^+ is the successor of c on P, then either cc^+ was a white edge in $F'_{i,1}$ or $cc^+ \in E(H'_{3,i})$. Thus in both cases $c^+ \in V^+$. So the set X^+ of successors in P of all the vertices in X lies in V^+ and no vertex in X sends out a red edge in P. Similarly one can show that the set Y^- of predecessors in P of all the vertices in Y lies in V and no vertex in Y receives a red edge in P. Together with (γ_3) this shows that we can apply Lemma 7.22 with $P \cup H_{3,i}$ playing the role of G and V^+ playing the role of V and V playing the role of U to obtain a cycle $\hat{C}$ containing precisely the vertices of P such that $|E(\hat{C})\backslash E(P)| \leq 5$, $E(\hat{C})\backslash E(P) \subseteq E(H_{3,i})$ and such that $E(P)\backslash E(\hat{C})$ consists of edges from X to X^+ and edges from Y^- to Y. Thus $E(P)\backslash E(\hat{C})$ contains no red edges. Replacing P with $\hat{C}$ gives us a 1-factor (which we still call $F'_{i,j}$) with fewer cycles. Also note that if the number of cycles is reduced by ℓ, then we use at most $\ell + 5 \leq 6\ell$ edges in $H_{3,i}$ to achieve this. So $F'_{i,j}$ still satisfies all requirements with the possible exception of (β_2). If it still does not satisfy (β_2), we will repeatedly apply this 'rotation-extension' procedure until the current 1-factor $F'_{i,1}$ also satisfies (β_2). However, we need to be careful since we do not want to use edges of $H_{3,i}$ several times in this process. Simply deleting the edges we use may not work as (γ_2) might fail later on (when we will repeat the above process for $F'_{i,j}$ with $j > 1$).

So each time we modify $F'_{i,1}$, we also modify $H_{3,i}$ as follows. All the edges from $H_{3,i}$ which are used in $F'_{i,1}$ are removed from $H_{3,i}$. All the edges which are removed from $F'_{i,1}$ in the rotation-extension procedure are added to $H_{3,i}$. (Note that by (β_5) we never add red edges to $H_{3,i}$.) When we refer to $H_{3,i}$, we always mean the 'current' version of $H_{3,i}$, not the original one. Furthermore, at every step we still refer to an edge of $H_{3,i}$ as green, even if initially the edge did not lie in $H_{3,i}$. Similarly at every step we refer to the non-red edges of our current 1-factor as white, even if initially they belonged to $H_{3,i}$.

Note that if we added a green edge xz into $F'_{i,1}$, then x lost an outneighbour in $H_{3,i}$, namely z. However, (β_5) implies that we also moved some (white) edge xy of $F'_{i,1}$ to $H_{3,i}$, where y lies in the same cluster $V^+ \in V(\mathcal{F}_i)$ as z (here $x \in V$). So we still have that $\delta^+(H_{3,i}) \geq \eta_3\beta m'/3$. Similarly, at any stage $\delta^-(H_{3,i}) \geq \eta_3\beta m'/3$. When $H_{3,i}$ is modified, then $H'_{3,i}$ is modified accordingly. This will occur if we add some white edges to $H_{3,i}$ whose start or endpoint lies on a red edge in $F'_{i,1}$. However, Claim 7.27(iv) implies that at any stage we still have

$$d^+_{H'_{3,i}}(x), d^-_{H'_{3,i}}(y) \geq (1/2 - 2c')\eta_3\beta m' - 2c'\beta_2 m' \geq \eta_3\beta m'/3.$$

Also note that by (β_3), the modified version of $H_{3,i}$ still satisfies

$$e_{H_{3,i}}(S,T) \geq (\eta_3\beta - \sqrt{\varepsilon})|S||T| - 6n/(\log n)^{1/5} \geq \eta_3\beta|S||T|/2. \tag{7.41}$$

So $H_{3,i}$ and $H'_{3,i}$ will satisfy (γ_1)–(γ_3) throughout and thus the above argument still works. So after at most $n/(\log n)^{1/5}$ steps $F'_{i,1}$ will also satisfy (β_2).

Suppose that for some $1 < j \leq \psi$ we have found 1-factors $F'_{i,1}, \dots, F'_{i,j-1}$ satisfying (β_1)–(β_5). We can now carry out the rotation-extension procedure for $F'_{i,j}$ in the same way as for $F'_{i,1}$ until $F'_{i,j}$ also satisfies (β_2). In the construction of $F'_{i,j}$, we do not use the original $H_{3,i}$, but the modified version obtained in the construction of $F'_{i,j-1}$. We then introduce the oriented spanning subgraph $H'_{3,i}$ of $H_{3,i}$ similarly as before (but with respect to the current 1-factor $F'_{i,j}$). Then all the above bounds on these graphs still hold, except that in the middle expression of (7.41) we multiply the term $6n/(\log n)^{1/5}$ by j to account for the total number of edges removed from $H_{3,i}$ so far. But this does not affect the next inequality. So eventually, all the $F'_{i,j}$ will satisfy (β_1)–(β_5).

7.5.7 Merging the cycles in $F'_{i,j}$ to obtain Hamilton cycles

Our final aim is to piece together the cycles in $F'_{i,j}$, for each i and j, to obtain edge-disjoint Hamilton cycles of G. Since we have ψ 1-factors $F'_{i,1}, \dots, F'_{i,\psi}$ for each G_i, in total we will find

$$\begin{aligned}\psi r' &\overset{(7.29),(7.40)}{=} (1-2\eta')\beta_2 m' 2(\alpha - \eta')L/\beta \overset{(7.28)}{\geq} 2(1-2\eta')(\alpha-\eta')(1/2-\eta')m'L \\ &\overset{(7.17)}{\geq} (\alpha - \gamma)n\end{aligned}$$

edge-disjoint Hamilton cycles of G, as desired.

Recall that R'_o was defined in Section 7.5.1. Given any i, apply Lemma 7.23 to obtain a closed shifted walk

$$W_i = U_1^+ D_1' U_1 U_2^+ D_2' U_2 \dots U_{s-1}^+ D_{s-1}' U_{s-1} U_s^+ D_s' U_s U_1^+$$

in R_o' with respect to $\mathcal{F}_i$ such that each cycle in $\mathcal{F}_i$ is traversed at most $2L/\gamma$ times. So $\{D_1', \dots, D_s'\}$ is the set of all cycles in $\mathcal{F}_i$, U_k^+ is the successor of U_k on D_k' and $U_k U_{k+1}^+ \in E(R_o')$ for each $k = 1, \dots, s$ (where $U_{s+1} := U_1$). Moreover,

$$s \leq 2L^2/\gamma. \tag{7.42}$$

For each 1-factor $F_{i,j}'$ we will now use the edges of H_4 and $H_{5,i}$ to obtain a Hamilton cycle $C_{i,j}$ with the following properties:

(i) If $i \neq i'$ or $j \neq j'$ then $C_{i,j}$ and $C_{i',j'}$ are edge-disjoint.

(ii) $E(C_{i,j})$ consists of edges from $F_{i,j}'$, H_4 and $H_{5,i}$ only.

(iii) There are at most $2L^2/\gamma$ edges from H_4 lying in $C_{i,j}$.

(iv) There are at most $2L^2/\gamma + 5$ edges from $H_{5,i}$ lying in $C_{i,j}$.

For each j, we will use W_i to 'guide' us how to merge the cycles in $F_{i,j}'$ into the Hamilton cycle $C_{i,j}$. Suppose that we have already defined $\ell < \psi r'$ of the Hamilton cycles $C_{i',j'}$ satisfying (i)–(iv), but have yet to define $C_{i,j}$. We remove all those edges which have been used in these ℓ Hamilton cycles from both H_4 and $H_{5,i}$.

For each $V \in V(\mathcal{F}_i)$, we denote by V_w the subcluster of V containing all those vertices which do not lie on a red edge in $F_{i,j}'$. We refer to V_w as the *white subcluster of* V. Thus $|V_w| \geq (1 - c')m'$ by property (β_7) in Section 7.5.6. Note that the outneighbours of the vertices in V_w on $F_{i,j}'$ all lie in V^+ while their inneighbours lie in V^-. For each $k = 1, \dots, s$ we will denote the white subcluster of a cluster U_k by $U_{k,w}$. We use similar notation for U_k^+ and U_k^-.

Consider any $UV \in E(R_o')$. Recall that U and V are viewed as clusters of size m in R_o', but when considering $\mathcal{F}_i$ we are in fact considering subclusters of U and V of size m'.

When viewed as clusters in R'_o, UV initially corresponded to an ε-regular pair of density at least $\eta_4 d'$ in H_4. Thus when viewed as clusters in $\mathcal{F}_i$, UV initially corresponded to a 2ε-regular pair of density at least $\eta_4 d'/2$ in H_4. Moreover, initially the edges from U_w to V_w in H_4 induce a 3ε-regular pair of density at least $\eta_4 d'/3$. However, we have removed all the edges lying in the ℓ Hamilton cycles $C_{i',j'}$ which we have defined already. Property (iii) implies that we have removed at most $2L^2\ell/\gamma \leq 2L^2 n/\gamma$ edges from H_4. Thus we have the following property:

(δ_1) Given any $UV \in E(R'_o)$, let $S \subseteq U_w$, $T \subseteq V_w$ be such that $|S|, |T| \geq 3\varepsilon m'$. Then $e_{H_4}(S,T) \geq \eta_4 d'|S||T|/4$.

When constructing $C_{i,j}$ we will remove at most $2L^2/\gamma$ more edges from H_4. But since (δ_1) is far from being tight, it will hold throughout the argument below. Similarly, the initial definition of $H_{5,i}$ (c.f. Section 7.5.3) and (iv) together imply the following property:

(δ_2) Consider any edge $VV^+ \in E(\mathcal{F}_i)$. Let $S \subseteq V$ and $T \subseteq V^+$ be such that $|S|, |T| \geq \sqrt{\varepsilon} m'$. Then $e_{H_{5,i}}(S,T) \geq \eta_5 \beta |S||T|/2$.

We now construct $C_{i,j}$ from $F'_{i,j}$. Condition (β_6) in Section 7.5.6 implies that, for each $k = 1, \dots, s$, every vertex in $V_G(D'_k)$ lies on the same cycle, C'_k say, in $F'_{i,j}$. Let $x_1 \in U_{1,w}$ be such that x_1 has at least $\eta_4 d'|U^+_{2,w}|/4 \geq \eta_4 d' m'/5$ outneighbours in H_4 which lie in $U^+_{2,w}$. By (δ_1) all but at most $3\varepsilon m'$ vertices in $U_{1,w}$ have this property. Note that the outneighbour in $F'_{i,j}$ of any such vertex lies in U^+_1. However, by (δ_2) all but at most $\sqrt{\varepsilon} m'$ vertices in U^+_1 have at least $\eta_5 \beta |U_{1,w}|/2 \geq \eta_5 \beta m'/3$ inneighbours in $H_{5,i}$ which lie in $U_{1,w}$. Thus we can choose x_1 with the additional property that its outneighbour $y_1 \in U^+_1$ in $F'_{i,j}$ has at least $\eta_5 \beta m'/3$ inneighbours in $H_{5,i}$ which lie in $U_{1,w}$.

Let P denote the directed path $C'_1 - x_1 y_1$ from y_1 to x_1. We now have two cases to consider.

Case 1. $C'_1 \neq C'_2$.

Note that x_1 has at least $\eta_4 d' m'/5 - c' m' \gg \eta_4 d' m'/6$ outneighbours $y'_2 \in U^+_{2,w}$ in H_4 such that the inneighbour of y'_2 in $F'_{i,j}$ lies in $U_{2,w}$. However, by (δ_1) all but at most $3\varepsilon m'$ vertices in $U_{2,w}$ have at least $\eta_4 d'm'/5$ outneighbours in H_4 which lie in $U^+_{3,w}$. Thus we can choose an outneighbour $y'_2 \in U^+_{2,w}$ of x_1 in H_4 such that the inneighbour x'_2 of y'_2 in $F'_{i,j}$ lies in $U_{2,w}$ and x'_2 has at least $\eta_4 d'm'/5$ outneighbours in H_4 which lie in $U^+_{3,w}$. We extend P by replacing it with $(P \cup C'_2 \cup \{x_1 y'_2\}) \backslash \{x'_2 y'_2\}$.

Case 2. $C'_1 = C'_2$.

In this case the vertices in $V_G(D'_2)$ already lie on P. We will use the following claim to modify P.

Claim 7.28 *There is a vertex $y_2 \in U^+_{2,w}$ such that:*

- $x_1 y_2 \in E(H_4)$.
- *The predecessor x_2 of y_2 on P lies in $U_{2,w}$.*
- *There is an edge $x_2 y'_2$ in $H_{5,i}$ such that $y'_2 \in U^+_{2,w}$ and y_2 precedes y'_2 on P (but need not be its immediate predecessor).*
- *The predecessor x'_2 of y'_2 on P lies in $U_{2,w}$.*
- *x'_2 has at least $\eta_4 d'm'/5$ outneighbours in H_4 which lie in $U^+_{3,w}$.*

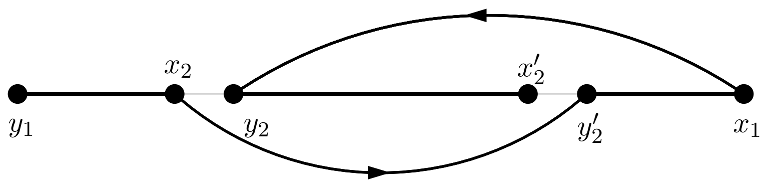

Figure 7.2: The modified path P in Case 2

Proof. Since x_1 has at least $\eta_4 d'm'/5$ outneighbours in H_4 which lie in $U^+_{2,w}$, at least $\eta_4 d'm'/5 - c'm' - 3\varepsilon m' \geq \eta_4 d'm'/6$ of these outneighbours y are such that the predecessor x of y on P lies in $U_{2,w}$ and at least $\eta_4 d'm'/5$ outneighbours of x in H_4 lie in $U^+_{3,w}$. This follows since all such vertices y have their predecessor on P lying in U_2 (since $y \in U^+_{2,w}$), since $|U_{2,w}| \geq (1 - c')m'$ and since by (δ_1) all but at most $3\varepsilon m'$ vertices in $U_{2,w}$ have at least $\eta_4 d'm'/5$ outneighbours in $U^+_{3,w}$. Let Y_2 denote the set of all such vertices y, and let X_2 denote the set of all such vertices x. So $|X_2| = |Y_2| \geq \eta_4 d'm'/6$, $X_2 \subseteq U_{2,w}$, $Y_2 \subseteq U^+_{2,w} \cap N^+_{H_4}(x_1)$. Let X_2^* denote the set of the first $\eta_4 d'm'/12$ vertices in X_2 on P and Y_2^* the set of the last $\eta_4 d'm'/12$ vertices in Y_2 on P. Then (δ_2) implies the existence of an edge x_2y_2' from X_2^* to Y_2^* in $H_{5,i}$. Then the successor y_2 of x_2 on P satisfies the claim. $\square$

Let x_2, y_2, x_2' and y_2' be as in Claim 7.28. We modify P by replacing P with

$$(P \cup \{x_1y_2, x_2y_2'\}) \backslash \{x_2y_2, x_2'y_2'\}$$

(see Figure 7.2).

In either of the above cases we obtain a path P from y_1 to some vertex $x_2' \in U_{2,w}$ which has at least $\eta_4 d'm'/5$ outneighbours in H_4 lying in $U^+_{3,w}$. We can repeat the above process: If $C_3' \neq C_1', C_2'$ then we extend P as in Case 1. If $C_3' = C_1'$ or $C_3' = C_2'$ then we modify P as in Case 2. In both cases we obtain a new path P which starts in y_1 and ends in some $x_3' \in U_{3,w}$ that has at least $\eta_4 d'm'/5$ outneighbours in H_4 lying in $U^+_{4,w}$. We can continue this process, for each C_k' in turn, until we obtain a path P which contains all the vertices in $C_1', \dots, C_s'$ (and thus all the vertices in G), starts in y_1 and ends in some $x_s' \in U_{s,w}$ having at least $\eta_4 d'm'/5$ outneighbours in H_4 which lie in $U^+_{1,w}$.

Claim 7.29 *There is a vertex $y_1' \in U_1^+ \setminus \{y_1\}$ such that:*

- $x_s'y_1' \in E(H_4)$.
- *The predecessor x_1' of y_1' on P lies in $U_{1,w}$.*
- *There is an edge $x_1'y_1''$ in $H_{5,i}$ such that $y_1'' \in U^+_{1,w}$ and y_1' precedes y_1'' on P.*
- *The predecessor x_1'' of y_1'' on P lies in $U_{1,w}$.*
- *x_1'' has at least $\eta_5 \beta m'/3$ outneighbours in $H_{5,i}$ which lie in $U^+_{1,w}$.*

Proof. The proof is almost identical to that of Claim 7.28 except that we apply (δ_2) to ensure that x_1'' has at least $\eta_5 \beta m'/3$ outneighbours in $H_{5,i}$ which lie in $U^+_{1,w}$. $\square$

Let x_1', y_1', x_1'' and y_1'' be as in Claim 7.29. We modify P by replacing it with the path

$$(P \cup \{x_s' y_1', x_1' y_1''\}) \backslash \{x_1' y_1', x_1'' y_1''\}$$

from y_1 to x_1''. So P is a Hamilton path in G which is edge-disjoint from the ℓ Hamilton cycles $C_{i',j'}$ already defined. In each of the s steps in our construction of P we have added at most one edge from each of H_4 and $H_{5,i}$. So by (7.42) P contains at most $2L^2/\gamma$ edges from H_4 and at most $2L^2/\gamma$ edges from $H_{5,i}$. All other edges of P lie in $F'_{i,j}$. Recall that y_1 has at least $\eta_5 \beta m'/3$ inneighbours in $H_{5,i}$ which lie in $U_{1,w}$ and x_1'' has at least $\eta_5 \beta m'/3$ outneighbours in $H_{5,i}$ which lie in $U_{1,w}^+$. Thus we can apply Lemma 7.22 to $P \cup H_{5,i}$ with U_1^+ playing the role of V and U_1 playing the role of U to obtain a Hamilton cycle $C_{i,j}$ in G where $|E(C_{i,j}) \backslash E(P)| \leq 5$. By construction, $C_{i,j}$ satisfies (i)–(iv). Thus we can indeed find $(\alpha - \gamma)n$ Hamilton cycles in G, as desired.

7.6 Proof of Conjecture 7.6 for large tournaments

In this section we prove Conjecture 7.6 for sufficiently large regular tournaments. The following observation of Keevash and Sudakov [42] will be useful for this.

Proposition 7.30 *Let $0 < c < 10^{-4}$ and let G be an oriented graph on n vertices such that $\delta^0(G) \geq (1/2 - c)n$. Then for any (not necessarily disjoint) $S, T \subseteq V(G)$ of size at least $(1/2 - c)n$ there are at least $n^2/60$ directed edges from S to T.*

We now show that Theorem 5.13 implies Conjecture 7.6 for sufficiently large regular tournaments.

Theorem 7.31 *There exists an integer n_0 such that the following holds. Given any regular tournament G on $n \geq n_0$ vertices and a set A of less than $(n-1)/2$ edges of G, then $G - A$ contains a Hamilton cycle.*

Proof. Let $0 < \nu \ll \tau \ll \eta \ll 1$. It is not difficult to show that G is a robust (ν, τ)-outexpander. Indeed, if $S \subseteq V(G)$ and $(1/2 + \tau)n < |S| < (1 - \tau)n$ then $RN^+_{\nu,G}(S) = V(G)$.

If $\tau n < |S| < (1/2 - \tau)n$ then it is easy to see that $|RN^+_{\nu,G}(S)| \geq (1-\tau)n/2 \geq |S| + \nu n$. So consider the case when $(1/2 - \tau)n \leq |S| \leq (1/2 + \tau)n$. Suppose $|RN^+_{\nu,G}(S)| < |S| + \nu n \leq (1/2 + 2\tau)n$. Then by Proposition 7.30 there are at least $n^2/60$ directed edges from S to $V(G)\backslash RN^+_{\nu,G}(S)$. By definition each vertex $x \in V(G)\backslash RN^+_{\nu,G}(S)$ has less than νn inneighbours in S, a contradiction. So $|RN^+_{\nu,G}(S)| \geq |S| + \nu n$ as desired.

Since $|A| < (n-1)/2$ and n is sufficiently large, $G - A$ must be a robust $(\nu/2, \tau)$-outexpander. Thus if $\delta^0(G - A) \geq \eta n$ then by Theorem 5.13, $G - A$ contains a Hamilton cycle.

If $\delta^0(G - A) < \eta n$ then there exists precisely one vertex $x \in V(G - A)$ such that either $d^+_{G-A}(x) < \eta n$ or $d^-_{G-A}(x) < \eta n$. Without loss of generality we may assume that $d^+_{G-A}(x) < \eta n$. Note that $d^+_{G-A}(x) \geq 1$ and let $y \in N^+_{G-A}(x)$. Let G' be the digraph obtained from $G - A$ by removing x and y from $G - A$ and adding a new vertex z so that $N^+_{G'}(z) := N^+_{G-A}(y)$ and $N^-_{G'}(z) := N^-_{G-A}(x)$. So $\delta^0(G') \geq \eta n - 2 \geq \eta n/2$ and G' is a robust $(\nu/3, 2\tau)$-outexpander. Thus by Theorem 5.13 G' contains a Hamilton cycle which corresponds to one in G. $\square$

Bibliography

[1] R. Aharoni, A. Georgakopoulos and P. Sprüssel, Perfect matchings in r-partite r-graphs, *European J. Combin.* **30** (2009), 39–42.

[2] N. Alon and G. Gutin, Properly colored Hamilton cycles in edge colored complete graphs, *Random Structures and Algorithms* **11** (1997), 179–186.

[3] N. Alon and A. Shapira, Testing subgraphs in directed graphs, *Journal of Computer and System Sciences* **69** (2004), 354–382.

[4] N. Alon and J. Spencer, *The Probabilistic Method* (2nd edition), Wiley-Interscience 2000.

[5] N. Alon and R. Yuster, H-factors in dense graphs, *J. Combin. Theory B* **66** (1996), 269–282.

[6] B. Alspach, J.-C. Bermond and D. Sotteau, Decompositions into cycles. I. Hamilton decompositions, *Cycles and rays (Montreal, PQ, 1987)*, Kluwer Acad. Publ., Dordrecht, 1990, 9–18.

[7] B. Auerbach and R. Laskar, On decompositions of r-partite graphs into edge-disjoint Hamiltonian circuits, *Discrete Math.* **14** (1976), 265–268.

[8] J. Bang-Jensen and G. Gutin, *Digraphs: Theory, Algorithms and Applications*, Springer 2000.

[9] J. Bang-Jensen, A. Yeo, Decomposing k-arc-strong tournaments into strong spanning subdigraphs, *Combinatorica* **24** (2004), 331–349.

[10] J.C. Bermond and C. Thomassen, Cycles in digraphs – a survey, *J. Graph Theory* **5** (1981), 1–43.

[11] B. Bollobás, D.E. Daykin and P. Erdős, Sets of independent edges of a hypergraph, *Quart. J. Math. Oxford Ser.* **27** (1976), 25–32.

[12] B. Bollobás and R. Häggkvist, Powers of Hamilton cycles in tournaments, *J. Combin. Theory B* **50** (1990), 309–318.

[13] A. Bondy, Basic graph theory: paths and circuits, in *Handbook of Combinatorics*, Vol. 1, Elsevier, Amsterdam (1995), 3–110.

[14] L.M. Brégman, Some properties of nonnegative matrices and their permanents, *Sov. Math. Dokl.* **14** (1973), 945–949.

[15] Y. Caro, Decompositions of large combinatorial structures, *Arch. Math* **52** (1989), 289–297.

[16] D. Christofides, P. Keevash, D. Kühn and D. Osthus, A semi-exact degree condition for Hamilton cycles in digraphs, *SIAM J. Discrete Math.* **24** (2010), 709–756.

[17] D. Christofides, P. Keevash, D. Kühn and D. Osthus, Finding Hamilton cycles in robustly expanding digraphs, preprint.

[18] D. Christofides, D. Kühn and D. Osthus, Edge disjoint Hamilton cycles in graphs, preprint.

[19] V. Chvátal, On Hamilton's ideals, *J. Combin. Theory B* **12** (1972), 163–168.

[20] B. Csaba, On the Bollobás–Eldridge conjecture for bipartite graphs, *Combin. Probab. Comput.* **16**, (2007), 661–691.

[21] D.E. Daykin and R. Häggkvist, Degrees giving independent edges in a hypergraph, *Bull. Austral. Math. Soc.* **23** (1981), 103–109.

[22] R. Diestel, *Graph Theory*, Graduate texts in Mathematics 173 (3rd Edition), Springer-Verlag 2005.

[23] G.A. Dirac, Some theorems on abstract graphs, *Proc. London. Math. Soc.* **2** (1952), 69–81.

[24] G.P. Egorychev, The solution of the Van der Waerden problem for permanents, *Dokl. Akad. Nauk SSSR* **258** (1981), 1041–1044.

[25] P. Erdős, Problem 9, in: M. Fieldler (Ed.), *Theory of Graphs and its Applications*, Czech. Acad. Sci. Publ., Prague, 1964, p. 159.

[26] P. Erdős and P. Turán, On some sequences of integers, *J. London Math. Soc.* **11** (1936) 261–264.

[27] D.I. Falikman, A proof of Van der Waerden problem for permanents of a doubly stochastic matrix, *Mat. Zametki* **29** (1981), 931–938.

[28] A. Frieze and M. Krivelevich, On packing Hamilton cycles in ε-regular graphs, *J. Combin. Theory B* **94** (2005), 159–172.

[29] M.R. Garey, D.S. Johnson, Computers and intractability, Freeman, 1979.

[30] A. Ghouila-Houri, Une condition suffisante d'existence d'un circuit hamiltonien, *C.R. Acad. Sci. Paris* **25** (1960), 495–497.

[31] R. Häggkvist, Hamilton cycles in oriented graphs, *Combin. Probab. Comput.* **2** (1993), 25–32.

[32] R. Häggkvist and A. Thomason, Oriented Hamilton cycles in oriented graphs, in *Combinatorics, Geometry and Probability*, Cambridge University Press 1997, 339–353.

[33] A. Hajnal and E. Szemerédi, Proof of a conjecture of Erdős, *Combinatorial Theory and its Applications vol. II* **4** (1970), 601–623.

[34] H. Hàn, Y. Person and M. Schacht, On perfect matchings in uniform hypergraphs with large minimum vertex degree, *SIAM J. Discrete Math.* **23** (2009), 732–748.

[35] P. Hell and D.G. Kirkpatrick, On the complexity of general graph factor problems, *SIAM J. Computing* **12** (1983), 601–609.

[36] G. Hetyei, On Hamiltonian circuits and 1-factors of the regular complete n-partite graphs, *Acta Acad. Pedagog, Civitate Press Ser.* **19** (1975), 5–10.

[37] B. Jackson, Long paths and cycles in oriented graphs, *J. Graph Theory* **5** (1981), 145–157.

[38] S. Janson, T. Łuczak and A. Rucinski, *Random Graphs*, Wiley, 2000.

[39] M. Karpiński, A Ruciński and E. Szymańska, Computational complexity of the perfect matching problem in hypergraphs with subcritical density, preprint.

[40] K. Kawarabayashi, K_4^--factor in a graph, *J. Graph Theory* **39** (2002), 111–128.

[41] P. Keevash, D. Kühn and D. Osthus, An exact minimum degree condition for Hamilton cycles in oriented graphs, *J. London Math. Soc.* **79** (2009), 144–166.

[42] P. Keevash and B. Sudakov, Triangle packings and 1-factors in oriented graphs, *J. Combin. Theory B* **99** (2009), 709–727.

[43] L. Kelly, D. Kühn and D. Osthus, A Dirac-type result on Hamilton cycles in oriented graphs, *Combin. Prob. Comput.* **17** (2008), 689–709.

[44] L. Kelly, D. Kühn and D. Osthus, Cycles of given length in oriented graphs, *J. Combin. Theory B* **100** (2010), 251–264.

[45] H.A. Kierstead and A.V. Kostochka, An Ore-type Theorem on Equitable Coloring, *J. Combin. Theory B* **98** (2008), 226–234.

[46] H.A. Kierstead, A.V. Kostochka and G. Yu, Extremal graph packing problems: Ore-type versus Dirac-type, in *Surveys in Combinatorics* (S. Huczynska, J.D. Mitchell and C.M. Roney-Dougal eds.), *London Math. Soc. Lecture Notes* **365**, 113–135, Cambridge University Press, 2009.

[47] J. Komlós, The Blow-up Lemma, *Combinatorics, Probability and Computing* **8** (1999), 161–176.

[48] J. Komlós, Tiling Turán Theorems, *Combinatorica* **20** (2000), 203–218.

[49] J. Komlós, G.N. Sárközy and E. Szemerédi, Blow-up Lemma, *Combinatorica* **17** (1997), 109–123.

[50] J. Komlós, G.N. Sárközy and E. Szemerédi, Proof of the Seymour conjecture for large graphs, *Annals of Combinatorics* **2** (1998), 43–60.

[51] J. Komlós, G.N. Sárközy and E. Szemerédi, Proof of the Alon-Yuster conjecture, *Discrete Math.* **235** (2001), 255–269.

[52] J. Komlós and M. Simonovits, Szemerédi's Regularity Lemma and its applications in graph theory, *Combinatorics: Paul Erdős is eighty, Vol.2* (1996), 295–352.

[53] D. Kühn, R. Mycroft and D. Osthus, An approximate version of Sumner's universal tournament conjecture, preprint.

[54] D. Kühn and D. Osthus, Matchings in hypergraphs of large minimum degree, *J. Graph Theory* **51** (2006), 269–280.

[55] D. Kühn and D. Osthus, Multicoloured Hamilton cycles and perfect matchings in pseudo-random graphs, *SIAM J. Discrete Mathematics* **20** (2006), 273–286.

[56] D. Kühn and D. Osthus, Critical chromatic number and the complexity of perfect packings in graphs, *17th ACM-SIAM Symposium on Discrete Algorithms* (SODA 2006), 851–859.

[57] D. Kühn and D. Osthus, The minimum degree threshold for perfect graph packings, *Combinatorica* **29** (2009), 65–107.

[58] D. Kühn and D. Osthus, Embedding large subgraphs into dense graphs, in *Surveys in Combinatorics* (S. Huczynska, J.D. Mitchell and C.M. Roney-Dougal eds.), *London Math. Soc. Lecture Notes* **365**, 137–167, Cambridge University Press, 2009.

[59] D. Kühn and D. Osthus, A survey on Hamilton cycles in directed graphs, preprint.

[60] D. Kühn, D. Osthus and A. Treglown, An Ore-type theorem for perfect packings in graphs, *SIAM J. Discrete Math.* **23** (2009), 1335–1355.

[61] D. Kühn, D. Osthus and A. Treglown, Hamiltonian degree sequences in digraphs, *J. Combin. Theory B* **100** (2010), 367–380.

[62] D. Kühn, D. Osthus and A. Treglown, Hamilton decompositions of regular tournaments, *Proc. Lond. Math. Soc.* **101** (2010), 303–335.

[63] D. Kühn, D. Osthus and A. Treglown, Matchings in 3-uniform hypergraphs, preprint.

[64] E. Lucas, *Récréations Mathématiques*, Vol. 2, Gautheir-Villars, 1892.

[65] K. Markström and A. Ruciński, Perfect matchings and Hamilton cycles in hypergraphs with large degrees, preprint.

[66] M. Molloy and B. Reed, *Graph colouring and the probabilistic method*, Springer-Verlag, Berlin-New York, 2002.

[67] J.W. Moon, *Topics on tournaments*, Holt, Rinehart and Winston, New York, 1968.

[68] C.St.J.A. Nash-Williams, Problem 47, *Proceedings of Colloq. Tihany 1966*, Academic Press 1968, p. 366.

[69] C.St.J.A. Nash-Williams, Hamilton circuits in graphs and digraphs, *The many facets of graph theory*, Springer Verlag Lecture Notes 110, Springer Verlag 1968, 237–243.

[70] C.St.J.A. Nash-Williams, Hamiltonian arcs and circuits, in *Recent Trends in Graph Theory (Proc. Conf., New York, 1970)*, Lecture Notes in Mathematics, Vol. 186, Springer, Berlin (1971), 197–210.

[71] C.St.J.A. Nash-Williams, Edge-disjoint Hamiltonian circuits in graphs with vertices of large valency, in *1971 Studies in Pure Mathematics (Presented to Richard Rado)*, Academic Press, London, 1971, 157–183.

[72] C.St.J.A. Nash-Williams, Hamiltonian circuits, *Studies in Math.* **12** (1975), 301–360.

[73] O. Ore, Note on Hamilton circuits, *Amer. Math. Monthly* **67** (1960), 55.

[74] O. Pikhurko, Perfect matchings and K_4^3-tilings in hypergraphs of large codegree, *Graphs Combin.* **24** (2008), 391–404.

[75] L. Pósa, A theorem concerning Hamiltonian lines, *Magyar Tud. Akad. Mat. Fiz. Oszt. Kozl.* **7** (1962), 225–226.

[76] K.B. Reid, Three problems on tournaments, in *Graph Theory and its Applications: East and West, Annals of the New York Academy of Sciences* **576** (1989), 466–473.

[77] K.B. Reid and E.T. Parker, Disproof of a conjecture of Erdős and Moser on tournaments, *J. Combin. Theory* **9** (1970), 225–238.

[78] V. Rödl and A. Ruciński, Dirac-type questions for hypergraphs – a survey (or more problems for Endre to solve), *An Irregular Mind (Szemerédi is 70)*, Bolyai Soc. Math. Studies **21** (2010), 1–30.

[79] V. Rödl, A. Ruciński and E. Szemerédi, Perfect matchings in uniform hypergraphs with large minimum degree, *European J. Combin.* **27** (2006), 1333–1349.

[80] V. Rödl, A. Ruciński and E. Szemerédi, Perfect matchings in large uniform hypergraphs with large minimum collective degree, *J. Combin. Theory Ser. A* **116** (2009), 613–636.

[81] A. Sanchez-Flores, On tournaments and their largest transitive subtournaments, *Graphs Combin.* **10** (1994), 367–376.

[82] A. Sanchez-Flores, On tournaments free of large transitive subtournaments, *Graphs Combin.* **14** (1998), 181–200.

[83] G.N. Sárközy, A fast parallel algorithm for finding Hamiltonian cycles in dense graphs, *Discrete Math.* **309** (2009), 1611–1622.

[84] P. Seymour, Problem section, in *Combinatorics: Proceedings of the British Combinatorial Conference 1973* (T.P. McDonough and V.C. Mavron eds.), 201–202, Cambridge University Press, 1974.

[85] A. Shoukoufandeh and Y. Zhao, Proof of a conjecture of Komlós, *Random Struct. Alg.* **23** (2003), 180–205.

[86] E. Szemerédi, Regular partitions of graphs, *Problémes Combinatoires et Théorie des Graphes Colloques Internationaux CNRS* **260** (1978), 399–401.

[87] C. Thomassen, An Ore-type condition implying a digraph to be pancyclic, *Discrete Math.* **19** (1977), 85–92.

[88] C. Thomassen, Long cycles in digraphs with constraints on the degrees, in *Surveys in Combinatorics* (B. Bollobás ed.), *London Math. Soc. Lecture Notes* **38**, 211–228, Cambridge University Press, 1979.

[89] C. Thomassen, Edge-disjoint Hamiltonian paths and cycles in tournaments, *Proc. London Math. Soc.* **45** (1982), 151–168.

[90] C. Thomassen, Hamilton circuits in regular tournaments, in *Cycles and rays*, vol. 115 of *North-Holland Math. Studies* (1985), 159–162.

[91] T.W. Tillson, A Hamiltonian decomposition of K^*_{2m}, $2m \geq 8$, *J. Combin. Theory B* **29** (1980), 68–74.

[92] A. Treglown, The Regularity Lemma and applications to packings in graphs, MSci thesis, School of Mathematics, University of Birmingham, 2007.

[93] W.T. Tutte, The factorisation of linear graphs, *J. London Math. Soc* **22** (1947), 107–111.

[94] P. Valadkhan, *Master's thesis, Simon Fraser University, 2009.*

[95] A. Young, Extremal problems for directed graphs, PhD thesis, School of Mathematics, University of Birmingham, 2007.

[96] R. Yuster, Tiling transitive tournaments and their blow-ups, *Order* **20** (2003), 121–133.

[97] R. Yuster, Combinatorial and computational aspects of graph packing and graph decomposition, *Computer Science Review* **1** (2007), 12–26.

[98] C.Q. Zhang, Every regular tournament has two arc-disjoint Hamilton cycles, J. Qufu Normal College, Special Issue Oper. Research (1980), 70–81.