Hamilton Cycles In Directed Graphs

by

Luke Kelly

A thesis submitted to
The University of Birmingham
for the degree of
Master of Philosophy (Qual)

School of Mathematics
The University of Birmingham
December 2007

Abstract

The main results of this thesis are the following. We show that for each $\alpha > 0$ every sufficiently large oriented graph G with $\delta^+(G), \delta^-(G) \geq 3|G|/8 + \alpha|G|$ contains a Hamilton cycle. This gives an approximate solution to a problem of Thomassen [**52**]. In fact, we prove the stronger result that G is still Hamiltonian if $\delta(G) + \delta^+(G) + \delta^-(G) \geq 3|G|/2 + \alpha|G|$. Up to the term $\alpha|G|$ this confirms a conjecture of Häggkvist [**30**]. This result is then used to derive a corresponding Ore type result and a new result on pancyclicity in oriented graphs.

Acknowledgements

I would like to thank Deryk Osthus and Daniela Kühn for all their assistance in the writing of this thesis, including, but certainly not limited to, their coauthorship of the main result of this thesis. Thanks is also due to Peter Keevash who pointed out a simpler argument for one of the main lemmas than that in an earlier manuscript.

Contents

Chapter 1

Introduction and Notation

1.1. Introduction

Hamiltonicity is one of the most important, and most studied, areas of graph theory. There are many papers published every year seeking more sufficient conditions for a graph to contain a Hamilton cycle, looking at the behaviour of Hamilton cycles in various models of random graphs and examining refinements of the idea of Hamiltonicity. This thesis consists of an introduction to the topic of Hamiltonicity within digraphs and a result recently proved with Deryk Osthus and Daniela Kühn [**37**] which provides an analogue of Dirac's famous theorem on Hamilton cycles in graphs for oriented graph. Extensions of this result, to an Ore type result and a pancyclicity result, are also discussed.

A simple graph $G = (V(G), E(G))$ is a set of *vertices*, $V(G)$ (or V if this is unambiguous), often taken to be $[n] := \{1, \ldots, n\}$, with a set of edges $E(G) \subseteq V^{(2)}$ (or E). The number of vertices in a graph is called its *order* and is often denoted $|G|$. The number of edges in a graph is denoted by $e(G)$. A *multigraph* is a graph in which edges are given a multiplicity. A *Hamilton cycle* is an ordering $x_1, \ldots, x_n$ of the vertices of a graph such that $x_i x_{i+1} \in E(G)$ for all i (counting modulo n). The *degree* $d(x)$ of a vertex $x \in V(G)$ is the number of vertices sharing an edge with x. The *minimum degree* $\delta(G)$ of a graph is the minimum of the degrees of the vertices of G. A fundamental result of Dirac states that a minimum degree of $|G|/2$ guarantees a Hamilton cycle in an undirected graph G on at least 3 vertices. Ore in 1960 gave a stronger sufficient condition: if the sum of the degrees of every pair of non-adjacent vertices is at least $|G|$, then the graph is Hamiltonian [**48**].

A *digraph* or *directed graph* is a multigraph in which all the edges are assigned a direction and there are no multiple edges of the same direction. I.e. we allow an edge in each direction between two vertices, but no other multiple edges are allowed. An *oriented graph* is a (simple) graph in which every edge is assigned a direction. Equivalently, an oriented graph is a digraph with no multiple edges.

There is an obvious analogue of a Hamilton cycle for digraphs. That is, an ordering $x_1, \ldots, x_n$ of the vertices of a digraph D such that $x_i x_{i+1}$ is a directed edge for all i. When discussing cycles and paths in digraphs we always mean that they are directed without mentioning this explicitly. The *minimum semi-degree* $\delta^0(G)$ of an oriented graph G (or of a digraph) is the minimum of its minimum outdegree $\delta^+(G)$ and its minimum indegree $\delta^-(G)$. There are corresponding versions of the famous theorems of Dirac and Ore for digraphs. Ghouila-Houri [**28**] proved in 1960 that every digraph D with minimum semi-degree at least $|D|/2$ contains a Hamilton cycle. Meyniel [**44**] showed that an analogue of Ore's theorem holds for digraphs, that is a digraph on at least 4 vertices is either Hamiltonian or the sum of the degrees of a pair of non-adjacent vertices is less than $2|D|-1$. All these bounds are best possible. See Theorems 2.1 and 2.2 for proofs of the results of Ghouila-Houri and Meyniel.

It is natural to ask for the (smallest) minimum semi-degree which guarantees a Hamilton cycle in an oriented graph G. This question was first raised by Thomassen [**51**], who [**53**] showed that a minimum semi-degree of $|G|/2-\sqrt{|G|/1000}$ suffices (see also [**52**]). Note that this degree requirement means that G is not far from being complete. Häggkvist [**30**] improved the bound further to $|G|/2-2^{-15}|G|$ and conjectured that the actual value lies close to $3|G|/8$. The best previously known bound is due to Häggkvist and Thomason [**31**], who showed that for each $\alpha>0$ every sufficiently large oriented graph G with minimum semi-degree at least $(5/12+\alpha)|G|$ has a Hamilton cycle. Our first result implies that the actual value is indeed close to $3|G|/8$.

THEOREM 1.1. *For every $\alpha>0$ there exists an integer $N=N(\alpha)$ such that every oriented graph G of order $|G|\geq N$ with $\delta^0(G)\geq(3/8+\alpha)|G|$ contains a Hamilton cycle.*

A construction of Häggkvist [**30**] shows that the bound in Theorem 1.1 is essentially best possible (see Proposition 5.1).

In fact, Häggkvist [**30**] formulated the following stronger conjecture. Given an oriented graph G, let $\delta(G)$ denote the minimum degree of G (i.e. the minimum number of edges incident to a vertex) and set $\delta^*(G):=\delta(G)+\delta^+(G)+\delta^-(G)$.

CONJECTURE 1.2 (Häggkvist [**30**]). *Every oriented graph G with $\delta^*(G)>(3n-3)/2$ has a Hamilton cycle.*

Our next result provides an approximate confirmation of this conjecture for large oriented graphs.

THEOREM 1.3. *For every $\alpha>0$ there exists an integer $N=N(\alpha)$ such that every oriented graph G of order $|G|\geq N$ with $\delta^*(G)\geq(3/2+\alpha)|G|$ contains a Hamilton cycle.*

Note that Theorem 1.1 is an immediate consequence of this. Once one has a Dirac-type result it is natural to ask if there is a corresponding Ore type result and indeed in this case there is. The proof for this is similar to that of Theorem 1.3 so we do not give the entire proof.

THEOREM 1.4. *For every $\alpha>0$ there exists an integer $N=N(\alpha)$ such that if G is an oriented graph on $n\geq N$ vertices with $d^+(u)+d^-(v)\geq 3n/4+\alpha n$ for all non-adjacent vertices $u,v\in V(G)$ then G contains a Hamilton cycle.*

We do though give a proof of the one important lemma which is different, along with a brief discussion, in Chapter 5.

Moreover, note that Theorem 1.1 immediately implies a partial result towards a classical conjecture of Kelly (see e.g. [**4**]), which states that every regular tournament on n vertices can be partitioned into $(n-1)/2$ edge-disjoint Hamilton cycles, with a *tournament* being a complete oriented graph T in which all vertices have indegree and outdegree $(|T|-1)/2$.

Corollary 1.5. *For every $\alpha > 0$ there exists an integer $N = N(\alpha)$ such that every regular tournament of order $n \geq N$ contains at least $(1/8 - \alpha)n$ edge-disjoint Hamilton cycles.*

Indeed, Corollary 1.5 follows from Theorem 1.1 by successively removing Hamilton cycles until the oriented graph G obtained from the tournament in this way has minimum semi-degree less than $(3/8 + \alpha)|G|$. The best previously known bound on the number of edge-disjoint Hamilton cycles in a regular tournament is the one which follows from the result of Häggkvist and Thomason [**31**] mentioned above. A related result of Frieze and Krivelevich [**27**] implies that almost every tournament contains a collection of edge-disjoint Hamilton cycles which covers almost all of its edges and that the same holds for almost all regular tournaments. There is a more detailed discussion of result on packing Hamilton cycles in digraphs in the survey in Chapter 2.

An oriented graph on n vertices is called pancyclic if it contains a cycle of length ℓ for every $3 \leq \ell \leq n$. With some extra work, the minimum semi-degree condition in Theorem 1.3 also gives the following pancyclicity result. Since our main result was proved, Keevash, Kühn and Osthus [**36**] have extended the methods to give an exact minimum semi-degree result, by proving the (best-possible) bound of $\delta^0(G) \geq (3n-4)/8$ for containing a Hamilton cycle. We extend this further to give the following exact pancyclicity result.

Theorem 1.6. *There exists a number n_0 so that any oriented graph G on $n \geq n_0$ vertices with minimum semi-degree $\delta^0(G) \geq \lceil (3n-4)/8 \rceil$ contains an ℓ-cycle through u for any vertex $u \in V(G)$ and $4 \leq \ell \leq n$.*

This improves upon the previous work of Darbinyan [**22**], who proved that a minimum semi-degree of $\lfloor n/2 \rfloor - 1 \geq 4$ implies pancyclicity.

This thesis is organised as follows. In the rest of this chapter we introduce further notation. In Chapter 2 we give a survey of some of the main results on Hamilton cycles in digraphs. In Chapters 3 and 4 we give an overview of results on Hamiltonian decompositions of digraphs and pancyclicity in digraphs as well as some results of our own on short cycles in oriented graphs. In Chapter 5 we prove our main result, Theorem 1.3, and discuss how to prove an Ore type result for oriented graphs. Finally in Chapter 6 we prove the pancyclicity result discussed above.

1.2. Further Notation and Terminology

Given two vertices x and y of an oriented graph G, we write xy for the edge directed from x to y. We will sometimes say that x and y are *adjacent* to mean $xy \in E(G)$. Given $A \subset V(G)$ we say that $x \in V(G) \backslash A$ *dominates* A if $xy \in E(G)$ for all $y \in A$, and similarly we say that x is dominated by A if $yx \in E(G)$ for all $y \in A$. We write $N^+_G(x) := \{y \in V(G) : xy \in E(G)\}$ for the outneighbourhood of a vertex x and $d^+_G(x) := |N^+_G(x)|$ for its outdegree. Similarly, we write $N^-_G(x)$ for the inneighbourhood of x and $d^-_G(x) := |N^-_G(x)|$ for its indegree. We write $N_G(x) := N^+_G(x) \cup N^-_G(x)$ for the neighbourhood of x and use $N^+(x)$ etc. whenever this is unambiguous.

We write $\Delta(G)$ for the maximum of $|N(x)|$ over all vertices $x \in G$. We write $\delta(G)$, $\delta^+(G)$ and $\delta^-(G)$ respectively for the minimum of $|N(x)|$,

$|N^+(x)|$ and $|N^-(x)|$ over all vertices $x \in G$. We will sometimes write δ, δ^+ and δ^- where the meaning is clear.

Given a set A of vertices of G, we write $N^+_G(A)$ for the set of all outneighbours of vertices in A. So $N^+_G(A)$ is the union of $N^+_G(a)$ over all $a \in A$. $N^-_G(A)$ is defined similarly. The oriented subgraph of $G(V,E)$ induced by A is $G[A] := (A, E \cap A^{(2)})$. We say that A is *independent* if $G[A]$ contains no edges. Given two vertices x, y of G, an *x-y path* is a directed path which joins x to y. Given two disjoint subsets A and B of vertices of G, an A-B edge is an edge ab where $a \in A$ and $b \in B$, the number of these edges is denoted by $e_G(A,B)$.

Recall that when referring to paths and cycles in oriented graphs we always mean that they are directed without mentioning this explicitly. Given two vertices x and y on a directed cycle C, we write xCy for the subpath of C from x to y. Similarly, given two vertices x and y on a directed path P such that x precedes y, we write xPy for the subpath of P from x to y. A *walk* in an oriented graph G is a sequence of (not necessarily distinct) vertices $v_1, v_2, \dots, v_\ell$ where $v_i v_{i+1}$ is an edge for all $1 \leq i < \ell$. The walk is *closed* if $v_1 = v_\ell$. We define things similarly for graphs and for directed graphs. The *underlying graph* of an oriented graph G is the graph obtained from G by ignoring the directions of its edges. By a *strongly connected* or *strong* oriented graph we mean one in which between every 2 vertices there exists a directed path, which is the obvious oriented graph version of a connected (undirected) graph. The *distance* $d_G(x,y)$ between 2 vertices $x, y \in V(G)$ is the length of the shortest path from x to y in G.

Given disjoint vertex sets A and B in a graph G, we write $(A,B)_G$ for the induced bipartite subgraph of G whose vertex classes are A and B. We write (A,B) where this is unambiguous. We call an orientation of a complete graph a *tournament* and an orientation of a complete bipartite graph a *bipartite tournament*. A digraph D is *d-regular* if all vertices have in- and outdegree d. D is *regular* if it is d-regular for some d. It is easy to see (e.g. by induction) that for every odd n there exists a regular tournament on n vertices. A *1-factor* of a digraph is a 1-regular spanning subdigraph, i.e. a covering of the digraph by pairwise-disjoint cycles. Note that a Hamilton cycle is a connected 1-factor.

Note that all these definitions apply equally well for digraphs and oriented graphs.

For two functions $f, g : \mathbb{N} \to \mathbb{R}$ we write $f(n) = o(g(n))$ to mean $f(x)/g(x) \to 0$ as $x \to \infty$.

Chapter 2

Survey

2.1. Classical Results And Degree Conditions

In this section we give several, mainly classical, degree conditions forcing a digraph to be Hamiltonian. We start with digraph analogues of the famous theorems of Dirac and Ore.

Theorem 2.1 (Ghouila-Houri, 1960 [**28**]). *If D is a digraph on n vertices with minimum semi-degree $\delta^0(D) \geq n/2$ then D is Hamiltonian.*

Proof. Suppose not. Let D be a counterexample and let C be a cycle of maximum length in D (C is not necessarily unique), and write ℓ for the length of the cycle C. It is easy to see that $\ell \geq \max(\delta^+(D), \delta^-(D)) \geq n/2$. Indeed, consider a maximal path in D. By the minimum semi-degree condition, this path has length at least $\max(\delta^+(D), \delta^-(D))$. Since the endvertex has no edges to vertices outside this path it is adjacent to at least $\delta^+(D)$ vertices inside the path. Hence the end vertex must be adjacent to a vertex in the path at a distance (along the path) of at least $\delta^+(D)$ from itself. Similarly considering the first vertex in this maximal path gives us a cycle of length at least $\delta^-(D)$.

Now let $P = u_0u_1u_2 \dots u_k$ be a maximal path in $D - V(C)$. Since $\ell + k + 1 \leq n$ and $\ell \geq n/2$ we have $k < n/2$. Define $S := N^-(u_0) \cap C$, $T := N^+(u_k) \cap C$. By the maximality of P, $N^-(u_0) \subseteq C \cup P$, $N^+(u_k) \subseteq C \cup P$. Hence

$$|S|, |T| \geq \delta^0(D) - k \geq n/2 - k > 0.$$

Also, the maximality of C implies that for all $s \in S$ and $t \in T$, $\text{dist}_C(s, t) > k + 1$. Otherwise we could replace the path from s to t inside C with P to create a longer cycle. Hence there exists a vertex $s \in S$ followed by at least $k + 1$ vertices not in S. These are forbidden from T, along with all vertices in C which succeed a vertex in S. Thus at least $|S| - 1 + k + 1 \geq n/2$ vertices of C are not in T. Hence $|C| \geq n - k$, but this contradicts our earlier observation that $1 + \ell + k \leq n$.

$\square$

Next we give Meyniel's theorem, originally proved in 1973, which provides an analogue of Ore's theorem for directed graphs. The proof we give is due to Bondy and Thomassen [**11**]. It proceeds by proving a more technical looking result for which Meyniel's theorem is an immediate corollary.

Theorem 2.2 (Meyniel, 1973 [**44**]). *A strong digraph D on n vertices contains a Hamilton cycle if $d(u) + d(v) \geq 2n - 1$ for every pair $u, v \in V(D)$ of non-adjacent vertices.*

Theorem 2.3 (Bondy and Thomassen, 1976 [**11**]). *Let D be a strong digraph on n vertices containing no Hamilton cycle and let $S := x_1x_2 \dots x_k$*

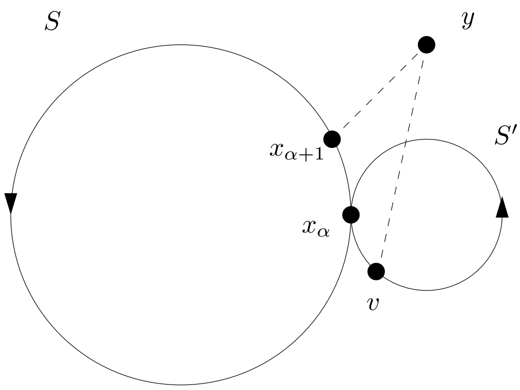


FIGURE 1. The setup for the case $\beta = 1$ of the proof of Theorem 2.3

be a maximal cycle in D. Then D contains a vertex v not in S and there exist numbers α, β ($1 \leq \alpha \leq k$, $1 \leq \beta < k$) such that

- $x_\alpha v \in E(D)$,
- *v is adjacent to no $x_{\alpha+i}$ with $1 \leq i \leq \beta$,*
- $d(v) + d(x_{\alpha+\beta}) \leq 2n - 1 - \beta$.

Proof. Suppose first that there is no path in D having only its start and finish in S. Then since D is strong and $V(S)$ is a proper subset of $V(D)$, D contains a cycle S' having precisely one vertex, say x_α, in common with S. Let v denote the successor of x_α on S'. If D contains a path of the form $x_{\alpha+1}yv$ or $vyx_{\alpha+1}$, where $y \in V \backslash V(S)$ then we contradict the assumption that D has no path having only its start and finish in S. So no such path exists and $e(y, \{x_{\alpha+1}, v\}) \leq 2$ for all such y. Further, we can assume that v is adjacent to no vertex of S other than x_α. Hence we obtain

$$\begin{aligned} d(v) + d(x_{\alpha+1}) &\leq 2 + 2(k-1) + \sum_{y \in V \backslash V(S)} (e(y, \{x_{\alpha+1}, v\}) + e(\{x_{\alpha+1}, v\}, y)) \\ &\leq 2k + 2(n-k-1) = 2n - 2, \end{aligned}$$

and the theorem is proved with $\beta = 1$.

Therefore we may assume that D contains a path having only its start and finish in S. Let $P := x_\alpha y_1 y_2 \dots y_s x_{\alpha+\gamma}$ be chosen such that γ is minimal. Note that since S is maximal, $\gamma > 1$. Put $v = y_1$.

(1) By the maximality of S we can't insert v into the path $x_{\alpha+\gamma}x_{\alpha+\gamma+1}\dots x_\alpha$. If there were more than $k-\gamma+2$ edges between v and this path then there would be an index i ($\alpha + \gamma \leq i < \alpha$) such that $x_i v, v x_{i+1} \in E(D)$, which would be a contradiction. Hence there are at most $k - \gamma + 2$ edges connecting this path to v.
(2) Furthermore, by the minimality of γ, v is not adjacent to any $x_{\alpha+i}$ with $1 \leq i < \gamma$, and D contains no path of the form $x_{\alpha+i}yv$ or $vyx_{\alpha+i}$ with $y \in V \backslash V(S)$ and $1 \leq i < \gamma$.

Now let β be defined as the largest integer i ($1 \leq i \leq \gamma$), such that D contains an $x_{\alpha+\gamma}x_\alpha$ path with vertex set $\{x_{\alpha+\gamma}, x_{\alpha+\gamma+1}, \dots, x_\alpha, x_{\alpha+1}, \dots, x_{\alpha+\beta-1}\}$. Let P' be an $x_{\alpha+\gamma}x_\alpha$ path with this vertex set. Note that it is possible that $\beta = 1$. Since $P \cup P'$ is a cycle, it follows from the maximality of S that $\beta < \gamma$. We cannot integrate $x_{\alpha+\beta}$ into this path so by the pigeonhole principle they

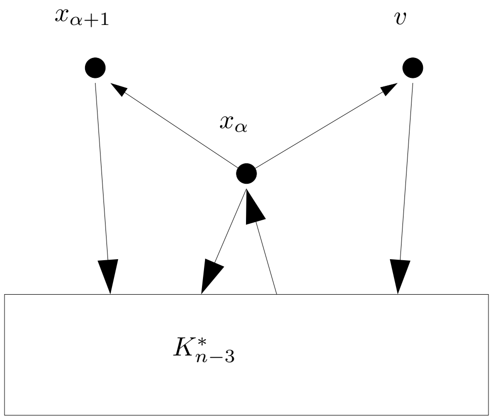


Figure 2. Extremal example for Theorem 2.2. In Theorem 2.3 we would have $\beta = 1$ and S as a Hamilton path in the K^*_{n-3} together with x_α and $x_{\alpha+1}$.

are joined by at most $k - \gamma + \beta + 1$ edges. Combining this with (1) and (2) we get

$$\begin{aligned} d(v)+d(x_{\alpha+\beta}) &\leq (k-\gamma+2)+(k-\gamma+\beta+1)+2(\gamma-\beta-1)+ \\ &+\sum_{y\in V\setminus V(S)} (e(y,\{x_{\alpha+1},v\})+e(\{x_{\alpha+1},v\},y)) \\ &\leq 2k-\beta+1+2(n-k-1) = 2n-1-\beta. \end{aligned}$$

This completes the proof. □

Taking the disjoint union of two complete digraphs of equal size shows that Theorem 2.1 is best possible. Figure 2.1 shows that Theorem 2.2 is also best possible.

The following important theorem of Moon states that for a tournament being strong is equivalent to having a Hamilton cycle.

Theorem 2.4 (Moon, 1966 [**45**]). *Let T be a strong tournament on $n \geq 3$ vertices. For every $x \in V(T)$ and every integer $k \in \{3, 4, \ldots, n\}$, there exists a cycle of length k through x in T. In particular, a tournament is Hamiltonian if and only if it is strong.*

Proof. Let $x \in V(T)$ and observe that since T is strong both $N^+(x)$ and $N^-(x)$ are non-empty. Moreover the bipartite subdigraph $(N^+(x), N^-(x))$ of T with vertex classes $N^+(x)$ and $N^-(x)$, is non-empty, so let yz be an edge of $(N^+(x), N^-(x))$. Then xyz is a 3-cycle through x as required. We now proceed by induction on k. Let $C = x_1x_2\ldots x_k$ be a cycle in T with $x_1 := x$ and $3 \leq k \leq n-1$. We show that T has a $(k+1)$-cycle through x.

If there is a vertex $y \in V(T)\setminus V(C)$ with an outneighbour and an inneighbour in C then it is easy to see that there exists an index i such that $x_iy, yx_{i+1} \in E(T)$. Thus the cycle $C' := x_1Cx_iyx_{i+1}Cx_k$ completes the proof, where we write x_1Cx_i for the subpath of C from x_1 to x_i. So we may assume that for all $y \in V(T)\setminus V(C)$ either $V(C) \subseteq N^+(y)$ or $V(C) \subseteq N^-(y)$. Denote by R the subset of $V(T)\setminus V(C)$ containing all vertices with inedges to all vertices in C. Similarly let S contain all vertices in $V(T)\setminus V(C)$ with outedges to all vertices in C. Then since T is strong, $R, S \neq \emptyset$ and (R, S) is not empty. Hence there exists $r \in R$, $s \in S$ with $rs \in E(T)$. Then the cycle $C' := x_1 r s x_3 C x_k$ is a $(k+1)$-cycle through x. □

This has the following corollary.

COROLLARY 2.5. *Every tournament T on n vertices with minimum semi-degree $\delta^0(T) \geq n/4$ is Hamiltonian.*

Proof. All we need to show is that a minimum semi-degree of $n/4$ implies that T is strong. Given a non-empty set $X \subset V(T)$, if every vertex $x \in X$ has $|N^+(x) \cap X| > (|X|-1)/2$ then $e(X) > \binom{|X|}{2}$, contradicting the fact that T is a tournament. Hence there exists a vertex $x \in X$ with $|N^+(x) \backslash X| \geq n/4 - |X|/2 + 1/2$ and so $|N^+(X)| \geq n/4 + |X|/2 + 1/2$. Similarly for outedges we get that for all non-empty $Y \subset V(T)$, $|N^-(Y)| \geq n/4 + |Y|/2 + 1/2$ – just reverse every edge in T to see this.

Then for any distinct vertices $u, v \in V(T)$ define $X_1 := N^+(u)$, $Y_1 := N^-(v)$ and $X_{i+1} := N^+(X_i)$, $Y_{i+1} := N^-(Y_i)$. By the minimum semi-degree condition $|X_1|, |Y_1| \geq n/4$ and induction gives that for $i \geq 2$, $|X_i|, |Y_i| \geq n/2 - 2^{-i-1}n + 1/2$. Indeed, for $i = 2$ we have

$$|X_2| \geq n/4 + |X_1|/2 + 1/2 \geq n/2 - n/8 + 1/2.$$

Now suppose that it holds for $i - 1$. Then by the above averaging argument we have that:

$$\begin{aligned} |X_i| \geq n/4 + |X_{i-1}|/2 + 1/2 &\geq n/4 + n/4 - 2^{-(i-1)-1}n/2 + 1/2 + 1/2 \\ &= n/2 - 2^{-i-1}n + 1. \end{aligned}$$

Hence there exists i with $|X_i| + |Y_i| \geq n + 1$ and so their intersection is non-empty. This gives us a uv path as desired. □

Taking the disjoint union of two regular tournaments T_1 and T_2 and adding all the edges from T_1 to T_2 shows that Corollary 2.5 is best possible.

Bollobás and Häggkvist showed that a slight increase in minimum semi-degree gives us not just a Hamilton cycle, but all small powers of a Hamilton cycle. Given a digraph D define the *kth power of D*, denoted $D^{(k)}$, to be the digraph with vertex set $V(D)$ and in which $uv \in E(D^{(k)})$ if and only if $\mathrm{dist}_D(u, v) \leq k$.

THEOREM 2.6 (Bollobás and Häggkvist, 1990 [**10**]). *For every $\varepsilon > 0$ and every natural number k there exists $N_0 = N_0(\varepsilon, k)$ such that the following holds. If T is a tournament on $n \geq N_0$ vertices with $\delta^0(T) \geq (1/4 + \varepsilon)n$, then T contains the kth power of a Hamilton cycle.*

There is a conjecture of Jackson (see e.g. [**4**]) which says that this minimum semi-degree condition implies the existence of a Hamilton cycle in any regular oriented graph.

CONJECTURE 2.7. *Every k-regular oriented graph of order at most $4k+1$, where $k \neq 2$, contains a Hamilton cycle.*

Taking the disjoint union of two regular tournaments shows that this is best possible.

For undirected graphs there are many degree conditions other than Ore's theorem on only some pairs of non-adjacent vertices that force a graph to be Hamiltonian. One of the best known examples of this is Fan's theorem [**23**] which says that if $\max\{d(x), d(y)\} \geq n/2$ for all pairs of vertices x, y at distance 2 from each other in a graph G on n vertices, then G is Hamiltonian. Similar theorems which impose degree conditions only on some pairs of vertices are rare for directed graphs. One such result that does exist is the following result of Bang-Jensen, Gutin and Li [**6**]. Recall that given a digraph D we say that a pair of vertices $\{x, y\} \subset V(D)$ is *dominated* if there exists a vertex $z \in V(D)$ such that $zx, zy \in E(D)$. We say that $\{x, y\} \subset V(D)$ is *dominating* if there exists a vertex $z \in V(D)$ such that $xz, yz \in E(D)$.

THEOREM 2.8 (Bang-Jensen, Gutin and Li, 1995 [**6**]). *Let D be a strong digraph on n vertices. Suppose that*

$$\min\{d^+(x) + d^-(y), d^-(x) + d^+(y)\} \geq n$$

for every pair of dominating or dominated non-adjacent vertices $\{x, y\}$. Then D contains a Hamilton cycle.

Note that this result neither implies nor is implied by Meyniel's theorem (Theorem 2.2).

2.2. Random Graphs

The study of random graphs has been an extremely productive area of research ever since Erdős and Rényi laid the foundations for the subject in 1959. Naturally random digraphs have also been extensively studied, and in particular many people have asked questions about Hamilton cycles in random digraphs.

We are interested not in whether some property definitely holds, but with the behaviour of 'typical' graphs, that is, in the probability that a property holds in a given probability space. In particular, many results try and show that the probability that some property P holds tends to 1 as the number of vertices of a graph given by a particular model tends to infinity. If this happens we say that P occurs *with high probability* (*w.h.p.*) or *almost surely*. In an arbitrary digraph on n vertices we can have n edges and a Hamilton cycle, or $(n-1)^2$ edges and no Hamilton cycle. Surprisingly, at least for those not accustomed to the study of random graphs, requiring only that we have a Hamilton cycle w.h.p. almost completely removes this huge window of uncertainty as we shall see in the rest of this section.

The following result provides an answer to perhaps the most obvious question one could ask about Hamilton cycles in random digraphs. That is, what happens in $D(n, p)$, the model of random digraphs consisting of digraphs with n vertices in which the edges are chosen independently and randomly with probability p. The answer was given through a rather surprising method using percolation by McDiarmid [**43**]. He showed that the probability that a random digraph $D \in D(n, p)$ is Hamiltonian is not smaller than the probability that $G(n, p)$ is Hamiltonian, where $G(n, p)$ is the model of random graphs consisting of graphs with n vertices in which the edges are chosen independently and randomly with probability p. This can then be combined with the following important theorem of Pósa to obtain a result for $D(n, p)$.

Theorem 2.9 (Pósa, 1976 [**8**]). *For a random graph $G \in G(n,p)$ with $p = (\log n + \log\log n + \omega(n))/n$ for some function $\omega(n) \to \infty$, G almost surely contains a Hamilton cycle.*

Theorem 2.10 (McDiarmid, 1981 [**43**]). *Let ε be constant and let $p := (1+\varepsilon)(\log n)/n$. Then*

$$\lim_{n\to\infty} \mathbb{P}(D \in D(n,p) \text{ is Hamiltonian}) = \begin{cases} 1 & \varepsilon > 0 \\ 0 & \varepsilon < 0 \end{cases}.$$

Frieze [**26**] later improved this result in two ways. Firstly he gave a much sharper threshold. Secondly he gave an explicit algorithm DHAM for finding Hamilton cycles in random digraphs.

Theorem 2.11 (Frieze, 1986 [**26**]). *Let $e_1, e_2, \dots, e_{n(n-1)}$ be a random permutation of the edges of the complete digraph on n vertices. Let $E_m := \{e_1, e_2, \dots, e_m\}$ and let D_m be the digraph on n vertices with edgeset E_m. Define $m^* := \min\{m : \delta^0(D_m) \geq 1\}$. Then*

$$\lim_{n\to\infty} \mathbb{P}(\textit{DHAM find a Hamilton cycle in } D_{m^*}) = 1.$$

There are many other families of random digraphs we might be interested in. When studying random graphs one finds, perhaps surprisingly, that there is commonly a sharp threshold below which a particular property does not hold w.h.p. and above which the property holds w.h.p. The result above shows that the very edge which increases the minimum semi-degree to 1 also makes the digraph Hamiltonian with high probability. With this behaviour in mind the following space of random digraphs is a natural one to study.

The random digraph $D_{k-\text{in},\ell-\text{out}}$ has vertex set $[n] = \{1, 2, \dots, n\}$ and each vertex $v \in [n]$ chooses a set $\text{in}(v)$ of k random edges directed to v and a set $\text{out}(v)$ of ℓ random edges directed from v. It is not normally important if we choose edges with replacement, i.e. if we forbid multiple edges, and the result we give below holds with or without it. Thus $D_{k-\text{in},\ell-\text{out}}$ usually has around $(k+\ell)n$ edges. This model of random digraphs was originally introduced by Fenner and Frieze and subsequently the question of the Hamiltonicity of such digraphs was studied by Cooper and Frieze (see [**17**] and [**18**]). In particular they proved, in the course of several papers, the following result.

Theorem 2.12 (Cooper and Frieze, 1999 [**17**]). *The digraph $D \in D_{2-in,2-out}$ is Hamiltonian with high probability.*

Note that this result is best possible as w.h.p. $D_{1-\text{in},2-\text{out}}$ contains 2 vertices of indegree 1 sharing a common inneighbour. The proof of this theorem is long and hard so we will not give the proof here. Instead we will note two of the main tools used in the proof, which are commonly used in other work in this area and which we use in the proof of our main result (Theorem 1.3) later.

Recall that a *1-factor* is a spanning collection of disjoint cycles. It turns out that the study of 1-factors in digraphs proves very helpful when looking for Hamilton cycles. It is often easier to find a 1-factor of a digraph, and then connect the cycles of the 1-factor up to form a Hamilton cycle, than it is to try and find a Hamilton cycle directly. This is an idea used in the proof of Theorem 2.12, another important result in this section (Theorem 2.15) and one of the results in the chapter on Hamiltonian decompositions (Chapter 3), as well as in the proof of our main result, Theorem 1.3.

The normal trick used when finding a 1-factor in an arbitrary digraph is to look for a perfect matching in a suitable (undirected) bipartite graph. Given a digraph $D = (V, E)$ define G to be the bipartite graph with vertex classes $V = \{1, 2, \ldots, n\}$ and $V' = \{1', 2', \ldots, n'\}$ where $ij' \in E(G)$ if and only if $ij \in E(D)$. Then crucially a perfect matching in G corresponds to a 1-factor of D. To find a perfect matching in G we need only check that Hall's condition holds, that is, for every subset $X \subset V(D)$, $|N_G(X)| = |N_D^+(X)| \geq |X|$.

Using this idea Cooper and Frieze proved that w.h.p. $D_{2-\text{in},2-\text{out}}$ contains a 1-factor. After this they give an algorithm which with high probability joins together the cycles of the 1-factor to create a Hamilton cycle. In showing that their algorithm succeeds w.h.p. they make heavy use of the following tail estimate of Chernoff (see e.g. [**3**, Cor. A.14]).

THEOREM 2.13. *Let $B(n, p)$ denote the binomial random variable with parameters n and p. Then for any $\varepsilon > 0$ we have*

$$\mathbb{P}(|B(n, p) - np| \geq \varepsilon n) < 2\exp(-2\varepsilon^2 n).$$

Compared to Chebychev's bound, which gives

$$\mathbb{P}(|B(n, p) - np| \geq \varepsilon n) \leq p^2/\varepsilon^2$$

for a binomial random variable, this is much stronger. Crucially it allows us to show that in a random graph a property that for a single vertex behaves like a binomial variable holds with high probability for every vertex. For example, the property of having almost expected vertex degree, holds w.h.p. *for all* vertices.

In a similar vein to the result on $D_{2-\text{in},2-\text{out}}$, Cooper, Frieze and Molloy [**19**] recently proved the following result on Hamilton cycles in random regular digraphs.

THEOREM 2.14 (Cooper, Frieze and Molloy, 2005 [**19**]). *Let $r \geq 2$ be a fixed constant and let $\Omega_{n,r}$ denote the set of digraphs with vertex set $[n]$ such that each vertex has indegree and outdegree r. Let D be chosen uniformly at random from $\Omega_{n,r}$. Then*

$$\lim_{n\to\infty} \mathbb{P}(D \text{ is Hamiltonian}) = \begin{cases} 0 & r = 2 \\ 1 & r \geq 3 \end{cases}.$$

As with almost every topic in graph theory, the use of probabilistic tools is vitally important in the study of Hamilton cycles. This is the case not just when looking at random digraphs. Probabilistic methods are an incredibly powerful tool when studying Hamilton cycles in any context. The following result of Alon [**3**] uses the probabilistic method to bound the number of Hamilton paths in a tournament, a problem, like many others, which it is not clear how to approach otherwise.

First we need some definitions. The *permanent* of an $n \times n$ matrix $A = (a_{i,j})$ is defined as

$$\mathrm{per}(A) := \sum_{\sigma \in S_n} \prod_{i=1}^{n} a_{i,\sigma(i)},$$

where the sum is taken over every permutation of $\{1, 2, \ldots, n\}$, i.e. over S_n. The definition is similar to that of the determinant of a matrix, the difference being that the signatures of the permutations are not taken into account when calculating the permanent. Unlike the determinant there is no pleasant geometric interpretation, crucially for us though it does have one very nice combinatorial interpretation. Given a digraph D denote by A_D the *adjacency matrix* of D. That is, let $A_D = (a_{i,j})$ be the matrix with rows and columns indexed by the vertices of D with $a_{i,j} = 1$ if $ij \in E(D)$ and $a_{i,j} = 0$ otherwise. A little thought shows that the permanent of A_D is precisely the number of 1-factors of D.

Theorem 2.15 (Alon, 1990 [**3**]). *Let $P(n)$ denote the maximum possible number of Hamilton paths in a tournament on n vertices. Then there exists a positive constant c such that for every n,*

$$\frac{n!}{2^{n-1}} \leq P(n) \leq cn^{3/2} \frac{n!}{2^{n-1}}.$$

Before we can prove this result we will need some extra notation and some other results. Let T be a tournament on n vertices and let A_T be its adjacency matrix. We will denote the number of 1-factors of T by $F(T)$ and the number of Hamilton cycles by $C(T)$. Clearly $C(T) \leq F(T)$. We will also, in a slight abuse of notation, write $P(T)$ for the number of Hamilton paths in a given tournament T and $C(n)$ for the maximum number of Hamilton cycles in any tournament on n vertices. ($F(n)$ is defined in the obvious way.) We will use the following solution of the Minc conjecture by Brégman to bound the permanent.

Theorem 2.16 (Brégman, 1973 [**12**]). *Let $A = (a_{i,j})$ be an $n \times n$ matrix with all $a_{i,j} \in \{0, 1\}$. Then the permanent of A satisfies*

$$\mathit{per}(A) \leq \prod_{1 \leq i \leq n} (r_i!)^{1/r_i},$$

where r_i is the number of non-zero entries in the ith row.

We will also need the following technical lemma (see e.g. [**3**]).

Lemma 2.17. *Define $g(x) = (x!)^{1/x}$. For every integer $S \geq n$, the maximum of the function $\prod_{i=1}^{n} g(x_i)$ subject to the constraints $\sum_{i=1}^{n} x_i = S$ and $x_i \geq 1$ are integers, is obtained if and only if the variables x_i are as equal as possible, i.e. iff each x_i is either $\lfloor S/n \rfloor$ or $\lceil S/n \rceil$.*

Proof. [of Theorem 2.15] Observe that the numbers r_i defined in the theorem above are precisely the outdegrees of the vertices of T. If at least one of these is 0, then clearly $C(T) = F(T) = 0$. Otherwise, by Theorem 2.16 and the technical lemma above, $F(T)$ is at most the value of the function $\prod_{i=1}^{n} (r_i!)^{1/r_i}$, where the integer values r_i satisfy $\sum_i r_i = \binom{n}{2}$ and are as equal as possible. Applying Stirling's formula gives, after some tedious calculation, that for every tournament T on n vertices

$$C(T) \leq F(T) \leq (1+o(1))\frac{\sqrt{\pi}}{\sqrt{2}\mathrm{e}}(n+1)^{3/2}\frac{(n-1)!}{2^n}. \tag{1}$$

Thus we have $C(n) \leq (1+o(1))(n+1)^{3/2}\frac{(n-1)!}{2^n}$.

Next we derive an upper bound on the number of Hamilton paths in a tournament in terms of the number of Hamilton cycles in a larger tournament. Given a tournament T let S be the tournament obtained from T by adding a new vertex y and by adding the edge xy to $E(S)$ with probability $1/2$ and the edge yx otherwise for every $x \in V(T)$. For every Hamilton path in T, the probability that it can be extended to a Hamilton cycle in S is precisely $1/4$. Thus the expected number of Hamilton cycles in S is $P(T)/4$, and hence there exists S for which $C(S) \geq P(T)/4$. Hence we have $P(n) \leq 4C(n+1)$.

Combining this with (1) we have

$$P(n) \leq 4C(n+1) \leq (1+o(1))\frac{\sqrt{\pi}}{\sqrt{2}\mathrm{e}}n^{3/2}\frac{n!}{2^{n-1}} = O\left(n^{3/2}\frac{n!}{2^{n-1}}\right).$$

The lower bound comes from a simple application of the linearity of expectation. For a permutation σ of $V(T)$ let X_σ be the indicator random variable for σ giving a Hamilton path, where we pick σ uniformly at random from all possible permutations of $V(T)$. Then the number of Hamilton paths in T is precisely $X := \sum X_\sigma$, where we sum over all possible permutations of $V(T)$, and

$$\mathbb{E}(X) = \sum \mathbb{E}(X_\sigma) = n!2^{-(n-1)}.$$

Thus some tournament has at least $\mathbb{E}(X)$ Hamilton paths and we are done. □

This result has recently been improved by Friedgut and Kahn [**24**].

THEOREM 2.18 (Friedgut and Kahn, 2004 [**24**]). *Let $C(n)$ denote the maximum number of Hamilton cycles in a tournament on n vertices. Then*

$$C(n) < O\left(n^{3/2-\alpha}\frac{(n-1)!}{2^n}\right),$$

where $\alpha = 0.2507\ldots$.

Recently Busch [**13**] has given the following best possible lower bound on the number of Hamilton paths in a strong tournament.

THEOREM 2.19 (Busch, 2005 [**13**]). *The minimum number of distinct Hamilton paths in a strong tournament on n vertices is $5^{(n-1)/3}$.*

2.3. Hamilton Cycles With Additional Constraints

In this section we discuss some results on Hamilton cycles where we impose additional constraints. Primarily this will be requiring that the Hamilton cycles either contain or avoid given edges or vertices.

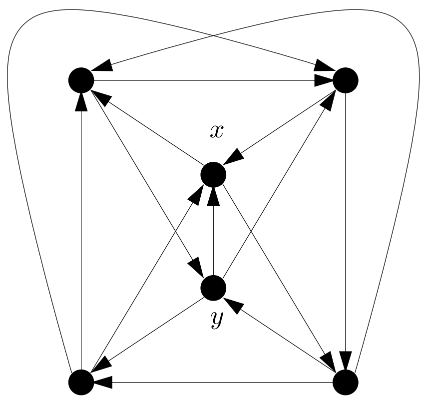


FIGURE 3. The exceptional tournament in Theorem 2.20.

We first give a classical result in this area. We say that a digraph D has an $[x, y]$-*Hamilton-path* for distinct vertices $x, y \in V(D)$ if it has a Hamilton path either from x to y or from y to x. The following theorem of Thomassen [**54**] gives a rather technical classification of tournaments containing an $[x, y]$-Hamilton path which has a nice corollary concerning such paths in strong tournaments. First we need some definitions. A maximal strong induced subgraph of a digraph D is called a *strong component*. We call a strong component of D *initial* if no other strong component sends an edge to it. Similarly a strong component is called *terminal* if it sends an end to no other strong component. See [**4**] for a fuller discussion of strong components.

THEOREM 2.20 (Thomassen, 1978 [**54**]). *Let T be a tournament on n vertices and let x_1, x_2 be distinct vertices of T. Then T has an $[x_1, x_2]$-Hamilton path if and only if none of the following holds.*

(a) *T is not strong and either neither x_1 nor x_2 belong to the initial strong component of T or neither x_1 nor x_2 belong to the terminal strong component.*
(b) *T is strong and for $i = 1, 2$, $T - x_i$ is not strong and x_{3-i} belongs to neither the initial nor the terminal strong component of $T - x_i$.*
(c) *T is isomorphic to Figure 2.3.*

This gives the following corollary, which is used in induction step of the proof of the theorem itself.

COROLLARY 2.21. *Let T be a strong tournament and let x, y, z be distinct vertices of T. Then T has a Hamilton path connecting two of the vertices in the set $\{x, y, z\}$.*

Proof. Suppose x, y, z are distinct vertices of a strong tournament T and that T has no Hamilton path connecting x and y. Then either (b) or (c) holds. The result is clear if T is isomorphic to Figure 2.3, so assume not. Then we have that (b) holds. If z belongs to the initial or the terminal strong component of $T - x$, there is a Hamilton path connecting x and z. If not, then $T - z$ is strong. Also as T is strong and $T - x$ is not strong we must have that x has an outneighbour in the initial strong component of $T - x$ and an inneighbour in the terminal strong component of $T - x$. But y belongs to neither the initial nor the terminal strong component of $T - x$

so we must have that $T - y$ is strong. But then T has a Hamilton path connecting z and y by Theorem 2.20. $\square$

The following result of Fraisse and Thomassen [**25**] provides a condition for the existence of Hamilton cycles in tournaments which avoid a given set of edges. We say that a digraph is *k-strong* if it is strong and if after removing any set of $k - 1$ vertices it is still strong.

THEOREM 2.22 (Fraisse and Thomassen, 1986 [**25**]). *For every k-strong tournament $T = (V, E)$ and every set $E' \subset E$ such that $|E'| \leq k - 1$, there is a Hamilton cycle in $T - E'$.*

Thomassen also proved the following result on Hamilton cycles in tournaments with restrictions on both edges included and excluded.

THEOREM 2.23 (Thomassen, 1983 [**56**]). *Let T be a tournament on n vertices and let k be a positive constant. Then for all k there exists $h(k)$ such that if T is $h(k)$-strong and $A_1 \subset E(T)$ and $A_2 \subset E(T) \backslash A_1$ are sets of at most k edges then $T - A_1$ contains a Hamilton cycle containing all edges in A_2.*

Kühn, Osthus and Young [**42**] recently gave a minimum degree condition that forces a sufficiently large digraph to contain a Hamilton cycle in which a number of vertices appear in a prescribed order. We say a digraph D is *k-ordered Hamiltonian* if for every sequence of k distinct vertices in D, $v_1, \ldots, v_k$, there exists a Hamilton cycle encountering $v_1, \ldots, v_k$ in this order.

THEOREM 2.24 (Kühn, Osthus and Young, 2007 [**42**]). *For every $k \geq 3$ there is an integer $n_0 = n_0(k)$ such that every digraph D on $n \geq n_0$ vertices with*

$$\delta^0(D) \geq \lceil (n + k)/2 \rceil - 1$$

is k-ordered Hamiltonian.

Note that the case $k = 1, 2$ is the theorem of Ghouila-Houri (Theorem 2.1), so this theorem generalises that result. They also proved a similar result on the existence of Hamilton cycles encountering given edges in the prescribed order.

2.4. Regularity and Hamiltonicity

In this section we will introduce the concept of ε-regularity and give a recent result by Frieze and Krivelevich on packing Hamilton cycles into ε-regular random digraphs. Essentially, a graph is ε-regular if it looks like a 'typical' random graph. More precisely, we define it as follows. The density of a bipartite graph $G = (A, B)$ with vertex classes A and B is defined to be

$$d_G(A, B) := \frac{e_G(A, B)}{|A|\,|B|}. \tag{2}$$

We often write $d(A, B)$ if this is unambiguous. Given $\varepsilon > 0$, we say that G is *ε-regular* if for all subsets $X \subseteq A$ and $Y \subseteq B$ with $|X| > \varepsilon\,|A|$ and $|Y| > \varepsilon\,|B|$ we have that $|d(X, Y) - d(A, B)| < \varepsilon$. A graph $G = (V, E)$ is ε-regular if this condition holds for all disjoint subsets; this is, $|d(X, Y) - d(A, B)| < \varepsilon$ for all disjoint $A, B \subset V$ with $|A|, |B| > \varepsilon|V|$.

So a bipartite graph is ε-regular if all sufficiently large pairs of subsets have almost the expected number of edges between them. This turns out to be the right way to capture the notion of 'looking random-like'. The following theorem (see e.g. [**3**]) shows that this simple property is equivalent to almost every property we would want from a 'random-like' graph. Equivalence here means that for all i, j if we are given c_i such that the statement relying on c_i is true then there exists $c_j = c_j(c_i)$ such that $c_j \to 0$ as $c_i \to 0$ and the statement relying on c_j is true for that value of c_j.

THEOREM 2.25. *Let G be a graph on n vertices with edge density p. Then the following properties are equivalent.*

- P1. *For every graph $H(s)$ on s vertices with t edges, the number of labelled copies of $H(s)$ in G is $(1 + c_1)n^s p^t$.*
- P2. *G contains at most $(p^4 + c_2)n^4$ labelled 4-cycles.*
- P3. $\sum_{x,x' \in V(G)} |N(x) \cap N(x')|^2 \leq (p^4 + c_2)n^4$.
- P4. $||N(x) \cap N(x')| - p^2 n| \leq c_3 n$ *for all but at most* $c_3 n^2$ *pairs of vertices* $x, x' \in V(G)$.
- P5. *For any 2 disjoint sets $A, B \subset V(G)$, $|e(A, B) - p|A||B|| \leq c_4 n^2$.*
- P6. *For any set $A \subset V(G)$, the number of edges inside A differs from $p|A|^2/2$ by at most $c_5 n^2$.*
- P7. *If $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of the adjacency matrix of G then* $\sum \lambda_i^4 \leq (p^4 + c_2)n^4$.

Furthermore, if G is regular (of degree pn) then these are all equivalent to:

- P8. *The second largest absolute value of the eigenvalues of the adjacency matrix of G is at most $c_6 n$.*

Note that Properties 2, 3 and 7 all count the same thing.

The fact that Property 5, which we have taken as the definition of ε-regularity, is strong enough to force almost the expected number of labelled copies of every subgraph is one of the interesting special cases of this result, and one of the reasons why ε-regularity is so interesting.

The full significance of the idea of ε-regularity comes from the enormous power of Szemerédi's famous Regularity lemma. Roughly speaking, this tells us that *any* large graph can be approximated by a union of ε-regular bipartite graphs. It has led to an astonishing variety of new results in the last twenty years in almost every area of graph theory. Alon and Shapira [**2**] proved a digraph version of the Regularity lemma, and we will use this heavily in our result. See Section 5.2 for a statement of the Diregularity lemma and a much longer discussion.

Clearly no digraph D contains more than $\delta^0(D)$ edge disjoint Hamilton cycles. It is conjectured that for random digraphs we achieve this bound with high probability. The corresponding problem for undirected graphs has been the subject of intensive research and has been settled for relatively sparse graphs [**9**] (at most $n \ln n$ edges in a graph of order n). The question appears to be very hard for the remaining cases. Erdős (see [**57**]) asked the following similar question on Hamilton cycles in tournaments.

QUESTION 2.26. *Do almost all tournaments contain $\delta^0(T)$ edge-disjoint Hamilton cycles?*

We now give one result which says dense random digraphs can almost be decomposed into edge-disjoint Hamilton cycles. Let $D(n, \alpha, \varepsilon)$ be the collection of digraphs on n vertices which are ε-regular and have minimum semi-degree αn, where an ε-regular digraph is defined in the obvious way. That is, given $\varepsilon > 0$ and a bipartite digraph D with vertex classes A and B, we say that D is *ε-regular* if for all subsets $X \subseteq A$ and $Y \subseteq B$ with $|X| > \varepsilon |A|$ and $|Y| > \varepsilon |B|$ we have that $|d(X, Y) - d(A, B)| < \varepsilon$ and $|d(Y, X) - d(B, A)| < \varepsilon$. I.e. we demand ε-regularity in *both* directions. A digraph $D = (V, E)$ with density $d := |E|/(n(n-1))$ is *ε-regular* if all $A, B \subset V$ with $|A|, |B| > \varepsilon|V|$ satisfy $|d - d(A, B)| < \varepsilon$ and $|d - d(B, A)| < \varepsilon$

Theorem 2.27 (Frieze and Krivelevich, 2003 [**27**]). *For all $\alpha > 0$ and for all sufficiently large $n \in \mathbb{N}$ there exists ε satisfying*

$$10 \left(\frac{\ln n}{n} \right)^{1/6} \leq \varepsilon < \alpha$$

such that if $D \in D(n, \alpha, \varepsilon)$ then D contains $(\alpha - 4\varepsilon^{1/2})n$ edge-disjoint Hamilton cycles with probability tending to 1 as n tends to infinity.

Corollary 2.28. *Let $0 < p < 1$ be constant. Then w.h.p. $D(n, p)$ contains $(p - o(1))n$ edge-disjoint Hamilton cycles.*

The corollary follows since using Chernoff type bounds it is relatively straightforward to show that w.h.p. a graph in $D(n, p)$ is ε-regular, i.e. w.h.p. $D(n, p) \in D(n, p, \varepsilon)$ for some constant ε.

Proof. [Sketch for Theorem 2.27] Let $\gamma := \varepsilon^{1/2}/2$. First choose a random subdigraph Γ of D with edge density γn and define $D_1 := D - \Gamma$. Straightforward use of Chernoff type bounds (Theorem 2.13) shows that w.h.p. Γ is dense and that there are almost the expected number of edges between any two disjoint subsets of $V(D)$ of size at least εn. We then extract a $2r$-regular subdigraph F from D_1, where $r := (\alpha - 6\gamma)n$. This is a fairly simple task, done by applying the max-flow min-cut theorem to a suitably defined bipartite graph.

For essentially the same reasons as discussed in the proof of Theorem 2.15, the number of perfect matching in a bipartite graph is equal to the permanent of its adjacency matrix. This fact and Brégman's proof of the Minc conjecture (Theorem 2.16) are used to show that any $2r$-regular digraph (with r not too small) contains many 1-factors, and hence one can show that it contains a 1-factor containing at most $10\varepsilon^{-1}(n/\ln n)^{1/2}$ cycles. Repeating this gives $t := r - \lfloor \varepsilon n \rfloor$ 1-factors $F_1, \ldots, F_t$ each not containing too many cycles.

Then for $i = 1, 2, \ldots, t$ we convert F_i into a Hamilton cycle H_i using the edges of the subgraph $D \backslash (H_1 \cup \ldots \cup H_{i-1} \cup F_i \cup \ldots \cup F_t)$. Note that this subgraph contains almost all the edges of Γ, and so we are able to use the fact that any two small sets, and thus the neighbourhoods of any two vertices, have an edge between them in Γ. The bound on the number of cycles in each 1-factor ensures that at every stage we have sufficient edges outside $H_1 \cup \ldots \cup H_{i-1} \cup F_i \cup \ldots \cup F_t$ to convert F_i into a Hamilton cycle. $\square$

This result also answers a long standing conjecture of Thomassen which in some sense generalises Kelly's conjecture (Conjecture 3.4). As with Corollary 2.28, this corollary follows from Theorem 2.27 by using Chernoff type bounds to show that w.h.p. a random tournament is ε-regular.

THEOREM 2.29 (Frieze and Krivelevich, 2003 [**27**]). *For any* $\varepsilon > 0$ *almost all tournaments of order* n *contain* $\lfloor (1/2-\varepsilon)n \rfloor$ *edge-disjoint Hamilton cycles.*

Chapter 3

Hamiltonian Decompositions

3.1. Results on Hamiltonian Decompositions

In this chapter we discuss Hamiltonian decompositions of digraphs and show how our result gives an improved partial result towards one of the most famous conjectures in tournament theory, Kelly's conjecture.

Given a digraph D, a *Hamiltonian decomposition* of D is a collection of edge-disjoint Hamilton cycles of D covering every edge of D. For undirected graphs it is a well-known and straightforward result that for $n \geq 1$ the complete graph K_{2n+1} is decomposable into edge-disjoint Hamilton cycles.

The question of the existence of Hamiltonian decompositions is harder for digraphs. We do though have the following result of Tillson showing that Hamiltonian decompositions exist for complete digraphs. The *complete digraph* of order n, K_n^*, is the unique digraph on n vertices such that for all $x, y \in V(K_n^*)$, $xy \in E(K_n^*)$.

Theorem 3.1 (Tillson, 1977 [**58**]). *The edges of K_n^* can be decomposed into Hamilton cycles if and only if $n \neq 4, 6$.*

The constructive proof uses Latin squares to give actual decompositions and can be found in Tillson's paper [**58**].

Whilst finding Hamiltonian decompositions is hard, if we just ask for a decomposition into edge-disjoint 1-factors the situation is easier. The following observation of Kotzig [**41**] says that we can always partition the edge set of a regular digraph into 1-factors.

Theorem 3.2 (Kotzig, 1969 [**41**]). *If $D = (V, E)$ is an r-regular digraph on n vertices, then the edge set of D can be partitioned into 1-factors.*

Proof. Define G to be the bipartite graph with vertex classes $V = \{1, 2, \ldots, n\}$ and $V' = \{1', 2', \ldots, n'\}$ where $ij' \in E(G)$ if and only if $ij \in E(D)$. Then G is r-regular and a perfect matching in G corresponds to a 1-factor of D. To check the existence of a perfect matching in G we need only check that Hall's condition holds, that is, for every subset $X \subset V$, $|N_G(X)| \geq |X|$.

So let $X \subset V$. By the regularity of G, the number of edges from X to $N(X)$ is exactly $r|X|$. Counting in the other direction, the regularity of G gives that the number of edges from $N(X)$ to X is at most $r|N(X)|$. Hence

$$r|X| = e(X, N(X)) \leq r|N(X)|,$$

as desired.

Remove this perfect matching from G to obtain an $(r-1)$-regular bipartite graph G'. By the same argument as before, G' contains a perfect matching. Repeating this we obtain r edge-disjoint perfect matchings in G, and hence r edge-disjoint 1-factors of D as required. $\square$

As in Theorem 2.12, it is sometimes desirable to restrict the number of cycles in a 1-factor, so as to make joining them together to form a Hamilton cycle easier. The following result of Frieze and Krivelevich was mentioned in the sketch of the proof of Theorem 2.27, but as it is of independent interest we state it seperately here.

Lemma 3.3 (Frieze and Krivelevich, 2005 [**27**]). *Let $\varepsilon > 0$ and let D be an r-regular digraph on n vertices, where $r \geq \varepsilon n$. Then D contains a 1-factor with at most $10\varepsilon^{-1}(n \ln n)^{1/2}$ cycles.*

3.2. Kelly's Conjecture

For tournaments we have the following famous conjecture of Kelly (see e.g. [**4**]).

Conjecture 3.4 (Kelly's Conjecture). *Every regular tournament on n vertices can be partitioned into $(n-1)/2$ edge-disjoint Hamilton cycles.*

There have been several partial results towards this conjecture. The first result which gave a number of Hamilton cycles dependent on n is due to Thomassen [**57**] who showed that every tournament contains at least $\lfloor \sqrt{n/1000} \rfloor$ edge-disjoint Hamilton cycles. The best previous result prior to that said that every regular tournament on at least 5 vertices contains 2 edge-disjoint Hamilton cycles (see e.g. [**4**]).

The first person to prove that there are a linear number of Hamilton cycles was Häggkvist [**30**], who proved that every regular tournament contains at least $2^{-18}n$ edge-disjoint Hamilton cycles. In work with Thomason [**31**] he later improved this bound to $n/12 - o(n)$, which was previously the best known bound. The following corollary of our main theorem (Theorem 1.3) improves these bounds.

Corollary 3.5. *For every $\alpha > 0$ there exists an integer $N = N(\alpha)$ such that every regular tournament T of order $n \geq N$ contains at least $(1/8 - \alpha)n$ edge-disjoint Hamilton cycles.*

Proof. A regular tournament has minimum semi-degree $(n-1)/2$ so we can apply Theorem 1.3 to obtain a Hamilton cycle H_1. Let $T_1 := T - H_1$. Then $\delta^0(T_1) = (n-1)/2 - 1$ and so we can find a Hamilton cycle H_2 and define $T_2 := T_1 - H_1$ as before. We can continue in this manner until T_i has minimum semi-degree less than $3n/8 + \alpha n$, i.e. we get at least $n/8 - \alpha n$ Hamilton cycles as desired. □

Note that whilst a full solution of Kelly's conjecture appears very hard, an improved partial result would be obtained by a solution of the conjecture of Jackson (Conjecture 2.7) stating that any regular oriented graph on n vertices with minimum semi-degree at least $n/4$ contains a Hamilton cycle.

Chapter 4

Pancyclicity

4.1. Pancylicity In Digraphs

An interesting way to strengthen the property of having a Hamilton cycle is to ask for the digraph to contain a cycle of every possible length. In this chapter we survey some of the results in this area. We also demonstrate how our result, in conjunction with partial results towards the Caccetta-Häggkvist conjecture (Conjecture 4.9), implies not just the presence of a Hamilton cycle in every sufficiently large oriented graph of minimum semi-degree $3n/8 + \alpha n$, but also a cycle of every possible length.

A digraph D of order n is *pancyclic* if it has cycles of all lengths $3, 4, \ldots, n$. We say that D is *vertex-pancyclic* if for every $v \in V(D)$ and any $k \in \{3, 4, \ldots, n\}$ there is a cycle of length k containing v. Furthermore, we say that D is *vertex-m-pancyclic* if D contains a cycle of length k containing v for all $m \leq k \leq n$ and each $v \in V(D)$.

Whilst this appears a much stronger notion than that of having a Hamilton cycle, it turns out that in some cases we require only slightly stronger conditions to force pancyclicity. The following example shows that an increase in minimum semi-degree of just 1 in the conditions of Theorem 2.1 gives us pancyclicity.

Theorem 4.1 (Alon and Gutin, 1997 [**1**]). *Every directed graph D on n vertices with minimum semi-degree $\delta^0(D) \geq n/2 + 1$ is vertex-2-pancyclic.*

Proof. Let $v \in V$ be an arbitrary vertex. The subdigraph $D - v$ has minimum semi-degree at least $n/2 = (n-1)/2 + 1/2$ so by Theorem 2.1 there is a Hamilton cycle $u_1 u_2 \ldots u_{n-1} u_1$ in $D - v$. If there is no cycle of length k through v then for no index i is $u_i \in N^+(v)$ and $u_{i+k-2} \in N^-(v)$, where the indices are modulo $n-1$. By summing over all values of i, we get that $|N^+(v)| + |N^-(v)| \leq n-1$, contradicting the minimum semi-degree condition. $\square$

Note that this theorem says that there are 2-cycles, which is something that is not required in our definition of pancyclicity. Similarly, adding just one to the Ore type condition in Meyniel's theorem (Theorem 2.2) again forces pancyclicity.

Theorem 4.2 (Thomassen, 1977 [**55**]). *Let D be a strong digraph on n vertices such that $d(x) + d(y) \geq 2n$ whenever x and y are nonadjacent. Then either D is pancyclic or n is even and D is isomorphic to the complete bipartite digraph on n vertices.*

Song [**50**] showed that for oriented graphs we can get a slightly stronger result.

Theorem 4.3 (Song, 1994 [**50**]). *Let D be an oriented graph on $n \geq 9$ vertices with minimum degree $n-2$. If*

$$xy \notin E(D) \Rightarrow d^+(x) + d^-(y) \geq n-3$$

then D is pancyclic.

There are many more nice results in this area. The following result of Häggkvist puts a bound not on the minimum degree but on the number of edges.

Theorem 4.4 (Häggkvist and Thomassen, 1976 [**32**]). *Every Hamiltonian digraph on n vertices and at least $n(n+1)/2 - 1$ edges is pancyclic.*

If we relax the conditions of pancyclicity and ask only for every short cycle we get a further set of interesting results. Erdős proved in 1963 a result of this type for undirected graphs, we state a directed analogue of his result for digraphs, due to Häggkvist and Thomassen.

Theorem 4.5 (Häggkvist and Thomassen, 1974 [**32**]). *Let $k \geq 2$ be an integer. Every strong digraph D with $n \geq (k-1)^2$ vertices and more than $n^2/2$ edges contains a cycle of length i for all $i \leq k$.*

The following simple observation says that if a tournament contains no triangles then it is acyclic, i.e. it contains no cycles of any length.

Proposition 4.6. *Every triangle-free tournament T is acyclic.*

Proof. Let $x_1x_2\dots x_k$ be a cycle of minimum length in T. By hypothesis $k \geq 4$. Since T is a tournament either $x_1x_3 \in E(T)$ or $x_3x_1 \in E(T)$. But in the first case $x_1x_3\dots x_k$ is a $(k-1)$-cycle, contradicting the minimality of k and in the second case $x_1x_2x_3$ is a triangle contradicting our assumption that T is triangle-free. □

Earlier when we stated Moon's theorem (Theorem 2.4) we actually gave a stronger version which he proved in 1968 stating that every strong tournament is not just Hamiltonian, but that for every vertex there is a cycle of every length containing it.

Theorem 4.7 (Moon, 1968 [**46**]). *Every strong tournament is vertex pancyclic.*

4.2. The Caccetta-Häggkvist Conjecture

To extend the main theorem of this thesis (Theorem 1.3) to show pancyclicity we will need to use a partial result towards the famous conjecture of Caccetta and Häggkvist on small cycles in digraphs. It is reasonable to expect that in a digraph with many edges there should be many cycles, and in particular, there should be some short cycles. The following theorem of Chvátal and Szemerédi shows that this intuition is correct.

Theorem 4.8 (Chvátal and Szemerédi, 1983 [**16**]). *Let D be a digraph on n vertices and let r be a positive constant such that the minimum outdegree of D is at least r. Then D contains a cycle of length at most $2n/(r+1)$.*

In 1978 Caccetta and Häggkvist [**14**] made the following striking conjecture (see e.g. [**47**]).

CONJECTURE 4.9 (Caccetta and Häggkvist, 1978 [**14**]). *Every digraph on n vertices with minimum outdegree at least r contains a cycle of length at most $\lceil n/r \rceil$.*

The following example shows that the result is best possible in the case $r = n/3$. It is not hard to extend it to a family of digraphs showing that the conjecture is best possible for all r.

PROPOSITION 4.10. *There exists an oriented graph G on n vertices with minimum outdegree $n/3 - 1$ such that G contains no triangle.*

Proof. Let the vertex set of G be $\{1, 2, \ldots, n\}$ and connect i to j if $1 \leq j - i \leq n/3 - 1$, counting modulo n. If $i_1 i_2 i_3$ is a triangle in G then by the pigeon hole principle one of $\{i_2 - i_1, i_3 - i_2, i_1 - i_3\}$ is at least $n/3$, where we again count modulo n. But this contradicts the definition of the edgeset of G. $\square$

Caccetta and Häggkvist proved the result themselves in the case $r = 2$ [**14**] and proofs also exist for the cases $r = 3$ (Hamidoune [**34**]) and $4 \leq r \leq 5$ (Hoáng and Reed [**35**]). Whilst the general case is unproven there are partial results in several different directions. One weakening which has led to a result is to insist that r is small compared to n.

THEOREM 4.11 (Shen, 1998 [**49**]). *Let r be a positive constant and let D be a digraph on $n \geq 2r^2 - 3r + 1$ vertices. If D has minimum outdegree r then it contains a cycle of length at most $\lceil n/r \rceil$.*

Allowing an additive constant error term in the size of the cycle has also been a productive avenue of research. The first result in this direction was due to Chvátal and Szemerédi [**16**], who gave an additive constant of 2500. The current best error term of 73 is due to Shen [**49**].

THEOREM 4.12 (Chvátal and Szemerédi, 1983 [**16**]). *Let D be a directed graph on n vertices and let r be a positive constant such that the minimum outdegree of D is at least r. Then D contains a cycle of length at most $\lceil n/r \rceil + 2500$.*

The third way one can weaken the conjecture is to increase the minimum outdegree demanded, and it is a partial result in this direction which we will need to prove that oriented graphs with minimum semi-degree $3n/8 + \alpha n$ are pancyclic. We will state results of this form in the case $r = n/3$, i.e. when we are trying to force a triangle in our digraph. This is by far the most famous form of this conjecture, and has been the subject of huge quantities of research over the last 25 years. One of the strongest results of this type, and the one we will use, is the following result of Shen.

THEOREM 4.13 (Shen, 1998 [**49**]). *Let $c = 3 - \sqrt{7} = 0.354\ldots$ and suppose that G is an oriented graph on n vertices with minimum outdegree at least cn. Then G contains a triangle.*

The following older (and weaker) result of Caccetta and Häggkvist is in a similar vein.

Theorem 4.14 (Caccetta and Häggkvist, 1978 [**14**]). *Let $c = (3-\sqrt{5})/2$ and suppose that G is an oriented graph on n vertices with minimum outdegree at least cn. Then G contains a triangle.*

Proof. Let $0 < c < 1$ and let G be an oriented graph on n vertices with minimum outdegree at least cn and suppose that G contains no triangles. We now show by induction that $c < (3-\sqrt{5})/2$.

By a simple averaging argument G contains a vertex v_0 with indegree at least cn. Let $A := N^-(v_0)$ and let $B := N^+(v_0)$. Then $|A|, |B| \geq cn$. Since G is an oriented graph we can assume that A, B and $\{v_0\}$ are all disjoint.

Let G' be the subdigraph of G induced by B. Since $|B| < n$ our induction hypothesis gives us a triangle in G', and thus in G, if $\delta^+(G') \geq c|B|$. Hence there exists a vertex $b_0 \in B$ with $|N^+_B(b_0)| < c|B|$. Let $W := N^+_G(b_0) \backslash B$ and note that

$$|W| > cn - c|B|.$$

If $v_0 \in W$ then G contains a 2-cycle, contradicting the fact that G is oriented. If $A \cap W \neq \emptyset$ then G contains a triangle and we are done. Therefore we can assume that the sets A, B, W and $\{v_0\}$ are all disjoint, and so

$$n \geq |A| + |B| + |W| + 1 > 2cn + (1-c)|B| + 1 > 3cn - c^2 n.$$

This implies that

$$c^2 - 3c + 1 > 0$$

and rearranging this gives $c > (3-\sqrt{5})/2$ as desired. □

In 1992 De Graaf, Schrijver and Seymour [**29**] considered a similar problem. Instead of asking for the minimum outdegree which forces a short cycle they asked for the minimum semi-degree which does this. The main result of this type is due to Shen and gives a stronger bound than is currently known for the ordinary version of the Caccetta-Häggkvist conjecture.

Theorem 4.15 (Shen, 1997 [**49**]). *If G is any oriented graph on n vertices with minimum semi-degree $\delta^0(G)/n \geq 0.3477\ldots$ then G contains a triangle.*

Recently Hamburger, Haxell and Kostochka [**33**] have used some results of Chudnovsky, Seymour and Sullivan [**15**] to slightly improve the bound given by Shen's method.

Theorem 4.16 (Hamburger, Haxell and Kostochka, 2007 [**33**]). *If G is any oriented graph on n vertices with minimum semi-degree $\delta^0(G) \geq 0.34564n$ then G contains a triangle.*

4.3. Short Cycles

To prove our pancyclicity result we need a result on short cycles in oriented graphs which, as it is of interest in its own right, we will state and discuss first. After seeing the triangle case of the Caccetta-Häggkvist conjecture (Conjecture 4.9) the following question is a natural one to ask.

Question 4.17. *What minimum semi-degree condition forces an oriented graph to contain a cycle of length k for a given constant k?*

We can provide an essentially best possible answer for $k \geq 4$, $k \not\equiv 0 \pmod 3$.

Theorem 4.18. *Let G be an oriented graph on n vertices and let $u, v \in V(G)$ be distinct vertices. Then the following holds.*

(P1) *If $\delta^0(G) \geq n/3 + 1$ then G contains a path of length exactly 4 from u to v.*

(P2) *If $\delta^0(G) \geq n/3 + 1 + (k-5)$ with $5 \leq k \leq n/7$ then G contains a path of length exactly k from u to v.*

If $u \in V(G)$ is any vertex then we have the following.

(C1) *If $\delta^0(G) \geq n/3 + 1$ then G contains a cycle of length 4 containing u.*

(C2) *If $\delta^0(G) \geq n/3 + 1$ then G contains a cycle of length 5 containing u.*

(C3) *If $\delta^0(G) \geq n/3 + (k-5)$ with $6 \leq k \leq n/7$ then G contains a cycle of length k containing u.*

Proof. We start with (P1). Assume that there is no path of length 4 from u to v. First note that given a walk $uu_1u_2u_3v$ of length 4 from u to v the only ways in which this will not be a path are if $u_1 = v$ or $u_3 = u$. Thus if we exclude these 2 possibilities then any such walk will in fact be a path.

Choose a set $X \subseteq N^+(u)\backslash\{v\}$ of size $n/3$ and define $Y := V(G)\backslash(X \cup N^+(X))$. A walk of length 4 from u to v through X which avoids u will necessarily be a path. For all vertices $x \in V(G)$ we have $|N(x)| \geq 2n/3 + 2$ and so the maximum size of an independent set is at most $n/3 - 2$. Writing X^+ for $N^+(X)$, we thus have $X \cap X^+ \neq \emptyset$.

Let $X^{++} := N^+(X^+)$. If $|X^{++}| > 2n/3$ then as $|N^-(v)\backslash\{u\}| \geq n/3$ their intersection is nonempty and we have the desired path. So assume not.

Assume $|X \cup X^+| \leq 2n/3$ and let $Z := \{z \in X \cap X^+ : N^-(z) \cap (X \cap X^+) = \emptyset\}$, i.e. Z is the set of vertices in $X \cap X^+$ which are not in $N^+(X \cap X^+)$. If $Z = X \cap X^+$ then $X \cap X^+$ is an independent set and

$$|N^+(X \cap X^+)\backslash X| \geq n/3 + 1 = n/3 + 1 - |X \cap X^+|/2 + |Z|/2.$$

Now suppose Z is a proper subset of $X \cap X^+$ and note that there are no edges from $X \cap X^-$ to Z. So

$$e((X \cap X^+)\backslash Z, X \cap X^+) \leq (|X \cap X^+| - |Z|)^2/2.$$

Hence there must exist a vertex $x \in (X \cap X^+)\backslash Z$ with $|N^+(x)\backslash X| \geq n/3 + 1 - |X \cap X^+|/2 + |Z|/2$ and hence

$$|X^+\backslash X| \geq |N^+(x)\backslash X| \geq n/3 + 1 - |X \cap X^+|/2 + |Z|/2. \tag{3}$$

Define $Y_B := \{y \in Y : N^-(y)\backslash Y = \emptyset\}$ and $Y_G := Y\backslash Y_B$. Y_B should be thought of as the set of 'bad' vertices – it contains vertices which are not in X^{++}.

If $Y_B = \emptyset$ then $|Y_G| = n - |X \cup X^+| > n/3$ by our assumption that $|X \cup X^+| \leq 2n/3$. But for any vertex $x \in X \cap X^+$ the minimum semidegree condition gives that $|N^+(x) \cap X^+| = |N^+(x)| \geq n/3 + 1$. Thus $|X^{++}| > 2n/3 + 1$, which is a contradiction.

So suppose that $Y_B \neq \emptyset$ and observe that

$$e(Y_B) \leq (|Y_B| - 1)|Y_B|/2, \quad e(Y_G, Y_B) \leq |Y_B||Y_G|, \quad e(V(G), Y_B) \geq |Y_B|n/3.$$

Combining these we get that

$$\begin{aligned} n/3 \le e(V(G), Y_B)/|Y_B| &\le |Y_B|/2 - 1/2 + |Y_G| \\ &= |Y|/2 - |Y_G|/2 - 1/2 + |Y_G|. \end{aligned}$$

Hence

$$\begin{aligned} &2n/3 \le |Y| + |Y_G| - 1 \\ (4) \qquad &\Rightarrow |Y_G| \ge 2n/3 - (n - |X \cup X^+|) + 1 = |X \cup X^+| - n/3 + 1. \end{aligned}$$

Observe that as there are no edges from X to Y, $Y_G \subseteq X^{++}$. Noting that $N^+(X^+)$ is at least the union of its intersections with X, $X^+\backslash X$ and Y we now have that

$$\begin{aligned} |X^{++}| &\ge |N^+(X \cap X^+) \cap (X^+\backslash X)| + |N^+(X \cap X^+) \cap X| + |Y_G| \\ &\ge |N^+(X \cap X^+) \cap (X^+\backslash X)| + (|X \cap X^+| - |Z|) + |Y_G| \\ &\overset{(3)}{\ge} (n/3 + 1/2 - |X \cap X^+|/2 + |Z|/2) + (|X \cap X^+| - |Z|) + |Y_G| \\ &\overset{(4)}{\ge} n/3 + 1/2 + |X \cap X^+|/2 - |Z|/2 + |X \cup X^+| - n/3 + 1 \\ &\ge |X \cap X^+|/2 - |Z|/2 + |X \cup X^+| + 3/2 \\ &\ge |X \cap X^+|/2 - |Z|/2 + |X| + |N^+(X \cap X^+)\backslash X| + 3/2 \\ &\overset{(3)}{\ge} |X \cap X^+|/2 - |Z|/2 + |X| + n/3 + 1 - |X \cap X^+|/2 + |Z|/2 + 3/2 \\ &= |X| + n/3 + 5/2 \\ &\ge 2n/3 + 5/2. \end{aligned}$$

Which contradicts our earlier assumption that $|X^{++}| \le 2n/3 + 1$.

Now suppose that $|X \cup X^+| > 2n/3$. Then $|Y| < n/3$ and so by the minimum semi-degree condition every vertex in Y has an inneighbour in $X \cup X^+$. But $e(X, Y) = 0$ by the definition of X^+ so every vertex in Y has an inneighbour in X^+ and $Y \subset X^{++}$.

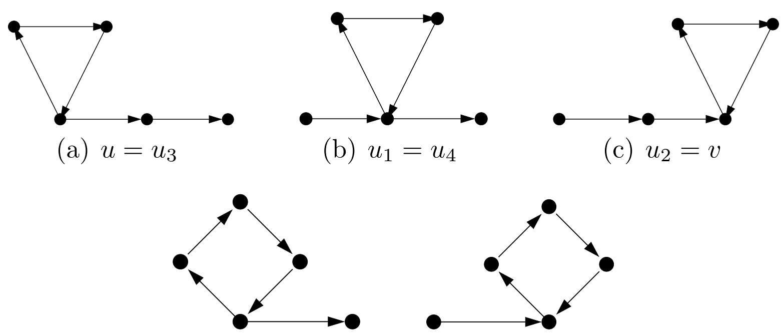
Also, $X\backslash X^+$ is an independent set, and so every vertex in $X\backslash X^+$ receives $\ge n/3 + 1$ edges from $(X^+ \cup Y)\backslash X$. But $|Y| < n/3$ so every vertex in X receives an edge from X^+. Thus $X\backslash X^+ \subset X^{++}$. Finally, any vertex in $X \cap X^+$ has at least $n/3 + 1$ outneighbours in X^+, and so $|X^{++} \cap X^+| \ge n/3 + 1$. (Recall that $X \cap X^+ \ne \emptyset$ and so we can indeed do this.)

Putting these observations together we get

$$\begin{aligned} |X^{++}| &\ge |X\backslash X^+| + |Y| + |X^{++} \cap X^+| \\ &\ge (|X| - |X \cap X^+|) + (n - |X \cup X^+|) + (n/3 + 1) \\ &\ge 4n/3 + 1 - |X^+| > 2n/3 + 1. \end{aligned}$$

Which is a contradiction. The final inequality holds since if $|X^+| \ge 2n/3 - 1$ then $|X^{++}| = n$, which is a contraction, and so we have shown that (P1) holds. Property (C1) holds by exactly the same argument; just set $u = v$.

Case (P2) is proved by a different method. We first prove it for $k = 5$ and then show how this implies the result in general. The maximum size of an independent set in G is $n - 2\delta^0(G) \le n/3 - 2$, so $N^+(u) \cap N^+(N^+(u)\backslash\{v\})$ has size at least 2. Thus there exist vertices $x \ne v$ and $x' \in N^+(u)\backslash\{v\}$ with



(a) $u = u_3$ (b) $u_1 = u_4$ (c) $u_2 = v$ (d) $u = u_4$ (e) $u_1 = v$

FIGURE 1. All possible ways in which a walk $uu_1u_2u_3u_4v$ of length 5 in an oriented graph could not be a path.

$ux, x'x \in E(G)$. (If not then there exists $y \in N^+(u)$ such that $N^+(u)\backslash\{y, v\}$ is an independent set of size $\geq n/3 - 1$, which is a contradiction.)

Note that the minimum semi-degree condition is impossible for $n < 9$ and for $n = 9$ it implies a regular tournament in which case the result is trivial. So we may assume that $n \geq 10$.

Define $X := N^+(x)\backslash\{v\}$. Then $|X| \geq n/3$. We now proceed as before, seeking a path of length 3 or 4 from x to v which avoids x' and u. Let $Y := N^-(v)\backslash\{u, x, x'\}$. By the minimum semi-degree condition, $|Y| \geq n/3 - 2$. We can assume the following, as otherwise we are done.

(i) $N^+(X) \cap Y = \emptyset$.
(ii) $(N^+(X) \cap N^-(Y))\backslash\{u\} = \emptyset$.
(iii) $X \cap N^-(Y) = \emptyset$.

Note that (i) and (iii) are equivalent. The exclusion of vertices from condition (ii) and the definitions of X and Y are necessary to prevent us getting a 'path' from u to v which self-intersects. We do not have to exclude u from X since $X \subseteq N^+(N^+(u))$ so $u \in X$ would give us a 2-cycle, contradicting the fact that G is an oriented graph. To see that we are done if one of (i)–(iii) fails let $uu_1u_2u_3u_4v$ be the walk from u to v which exists if one these conditions fails. Then either the negation of (i) gives

$$u_1 = x',\ u_2 = x,\ u_3 \in X,\ u_4 \in N^+(X) \cap Y \tag{5}$$

or the negation of (ii) gives

$$u_1 = x,\ u_2 \in X,\ u_3 \in (N^+(X) \cap N^-(Y))\backslash\{u\}, u_4 \in Y. \tag{6}$$

This walk will not be a path if any of the following occur: $u = u_3$, $u_1 = u_4$, $u_2 = v$, $u = u_4$ or $u_1 = v$ – see Figure 1 for an illustration. So in our situation we won't have a path if any of the following hold: $u_1 = v$, $u_2 = v$, $u_3 = u$ or $u_4 \in \{u, x, x'\}$. In (5) this is impossible because of the definition of x', the definition of x, the definition of X and the definition of Y respectively. Again in (6) this is impossible because of the definition of x, the definition of X, condition (ii) and the definition of Y respectively.

We now have to consider two separate cases.

CASE 1. *Either $X\backslash Y = \emptyset$ or $Y\backslash X = \emptyset$.*

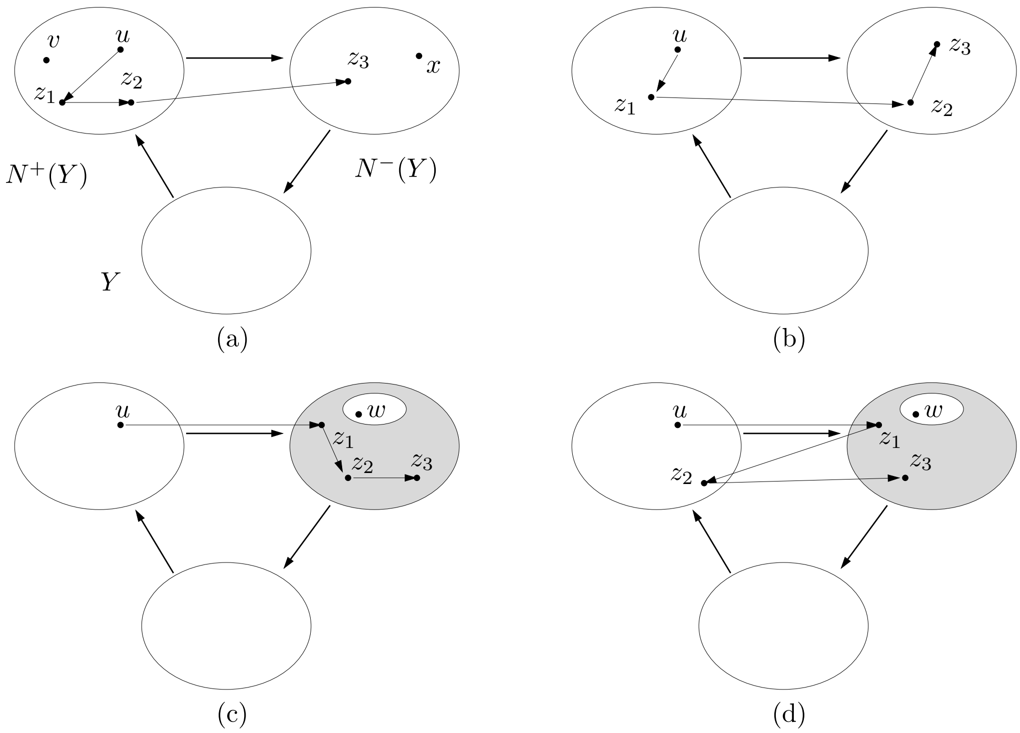


FIGURE 2. Diagram for Case 1 of the proof of (P2) in Theorem 4.18 when $Y \backslash X = \emptyset$.

If $X \backslash Y = \emptyset$ then $|X \cap Y| = |X| \geq n/3$, which is larger than the maximum size of an independent set. Thus $X \cap Y$ contains an edge $z_1 z_2$, but then $ux'xz_1z_2v$ is the desired path.

If $Y \backslash X = \emptyset$ then $|Y \cap X| = |Y| \geq n/3 - 2$, and we may assume that $Y = Y \cap X$ is an independent set. An independent set has size at most $n/3 - 2$ so we know that $|Y| = n/3 - 2$. Since any vertex $y \in Y \cap X$ has none of its $\geq 2n/3 + 2$ neighbours in Y, the minimum semi-degree condition gives us that $|N^+(y)|, |N^-(y)| = n/3 + 1$ and every vertex outside Y is a neighbour of y.

If there exists $y_1, y_2 \in Y$ with $N^+(y_1) \neq N^+(y_2)$ then there exist vertices z_1, z_2 with $z_1 \in N^+(y_1) \cap N^-(y_2)$ and $z_2 \in N^+(y_2) \cap N^-(y_1)$. Since $z_1 = u = z_2$ is impossible we may assume that $z_1 \neq u$. Then we have a path $uxy_1z_1y_2v$ of length 5 from u to v. Note that we can't have $z_1 = v$ as $y_1 \in N^+(v)$.

So we now have that $N^+(Y) \cap N^-(Y) = \emptyset$ and Y, $N^+(Y)$ and $N^-(Y)$ partition $V(G)$. Clearly $v \in N^+(Y)$ and $x \in N^-(Y)$. If $u \in N^-(Y)$ then any path of length 3 from Y to Y avoiding u and v completes the proof, and there are many such paths. Indeed, as no vertex in $N^+(Y)$ has an outneighbour in Y we have

$$e(N^+(Y), N^-(Y)) \geq \delta^0(G)|N^+(Y)| - e(N^+(Y)) \geq (n/3 + 1)^2/2.$$

So to get an edge from $N^+(Y)$ to $N^-(Y)$ avoiding u and v, and hence the desired path, we only need $(n/3+1)^2/2 > 2(n/3+1)$, which holds for $n \geq 10$, which we have assumed to be the case.

So we may assume that $u \notin N^-(Y)$. If u has an outneighbour $z_1 \in N^+(Y) \backslash \{v\}$ then one of the following must occur.

(a) Suppose there exists $z_2 \in N^+(Y)$, $z_2 \neq v$, with $z_1z_2 \in E(G)$. By the minimum semi-degree condition z_2 has an outneighbour $z_3 \in N^-(Y)$ and so for any $y \in Y$ we have $uz_1z_2z_3yv$ as a path of length 5 from u to v. Note that $z_3 \neq u$ since $u \notin N^-(Y)$.

(b) So assume not, i.e. $N^+(z_1) \cap N^+(Y) \subseteq \{v\}$. Then we have $|N^+(z_1) \cap N^-(Y)| \geq n/3$. By the minimum semi-degree condition $G[N^+(z_1) \cap N^-(Y)]$ is nonempty so pick an edge z_2z_3 in it. Then we again have the desired path.

So we can assume that u has no outneighbours in $N^+(Y)\backslash\{v\}$, and hence $|N^+(u) \cap N^-(Y)| \geq n/3$. Thus there exists at most one vertex, w say, in $N^-(Y)$ which is not in the outneighbourhood of u. We have one of the following cases.

(c) If there exists a path $z_1z_2z_3$ in $N^-(Y)\backslash\{w\}$ of length 2 then we are done.

(d) So assume no such path exists. Then in particular there exists a vertex $z_1 \in N^-(Y)\backslash\{w\}$ with no outneighbours in $N^-(Y)$ other than, possibly, w. Thus it has at least 2 outneighbours in $N^+(Y)$, and hence an outneighbour $z_2 \in N^+(Y)\backslash\{v\}$. By the minimum semi-degree condition z_2 has an outneighbour $z_3 \neq z_1$ in $N^-(Y)$. The path $uz_1z_2z_3yv$ completes the proof, for any vertex $y \in Y$.

This completes the analysis of Case 1.

Case 2. *$X\backslash Y \neq \emptyset$ and $Y\backslash X \neq \emptyset$.*

Suppose that all the vertices in $X\backslash Y$ have more than $(|X\backslash Y| - 1)/2$ outneighbours in X. Then $e(X\backslash Y, X) > \binom{|X\backslash Y|-1}{2}$, and so as $G[X]$ is an oriented graph at least one vertex must have an outneighbour in $X \cap Y$. But this gives us a path of length 3 from x to v and hence a contradiction.

Hence there exists a vertex in X with at most $(|X\backslash Y| - 1)/2$ outneighbours in X, and so at least $\delta^0(G) - (|X\backslash Y| - 1)/2$ outneighbours outside X. Thus

$$\begin{aligned} |N^+(X)\backslash X| &\geq n/3 + 1 - |X\backslash Y|/2 + 1/2 \\ &= n/3 + 3/2 - |X|/2 + |X \cap Y|/2. \end{aligned}$$

As $Y\backslash X \neq \emptyset$ we can use the same averaging argument as before to get that there exists a vertex in Y with at most $(|Y\backslash X| - 1)/2$ inneighbours in Y, and so at least $\delta^0(G) - (|Y\backslash X| - 1)/2$ inneighbours outside Y. Thus

$$|N^-(Y)\backslash Y| \geq n/3 + 3/2 - |Y|/2 + |X \cap Y|/2.$$

Combining these with the bounds on the sizes of $|X|$ and $|Y|$ we get

$$\begin{aligned} |X \cup N^+(X)| &\geq |X|/2 + n/3 + 3/2 + |X \cap Y|/2, \\ |Y \cup N^-(Y)| &\geq |Y|/2 + n/3 + 3/2 + |X \cap Y|/2. \end{aligned}$$

Finally, adding these together gives

$$\begin{aligned} |X \cup N^+(X)| + |Y \cup N^-(Y)| &\geq 2n/3 + 3 + |X|/2 + |Y|/2 + |X \cap Y| \\ &\geq n + 3/2 + |X \cap Y|. \end{aligned}$$

But $|X \cap N^-(Y)| + |N^+(X) \cap (Y \cup N^-(Y))| \leq 1$ by conditions (i)-(iii) so

$$|X \cup N^+(X)| + |Y \cup N^-(Y)| \leq n + |X \cap Y| + 1.$$

This contradiction completes the proof of case (P2) when $k = 5$.

To find a cycle of length 5 we do exactly the same thing, setting $u = v$.

To find paths of length $k \geq 6$ we first find a path $uu_1 \dots u_{k-5}$ in G of length $k-5$ starting at u, with $u_i \neq v$ for all i. Remove $u, u_1, \dots, u_{k-6}$ from $V(G)$ to form a new oriented graph H with $m := n - (k-5)$ vertices. A path of length 5 from u_{k-5} to v in H corresponds to a path of length k from u to v in G. But as $\delta^0(H) \geq m/3 + 1$ we can now apply (P2) to find the desired uv path.

To find a cycle of length $k \geq 6$ containing a given vertex u we first find a path $uu_1 \dots u_{k-5}$ of length $k-5$ starting at u. Remove $u_1, \dots u_{k-6}$ from $V(G)$ to form a new oriented graph H with $m := n - (k-6)$ vertices and minimum semi-degree $\delta^0(H) \geq m/3 + 1$. Now use (P2) to find a path of length 5 from u_{k-5} to u in H and thus obtain the desired cycle. Hence (C3) holds and the proof is complete. $\square$

Consider the blow-up of a triangle; the oriented graph on n vertices with 3 vertex classes V_0, V_1, V_2 of as equal size as possible, with all possible edges from V_i to V_{i+1}, counting modulo 3. It contains no cycles of length k where $k \geq 4$, $k \not\equiv 0 \pmod 3$, and so our result on minimum semi-degree forcing the existence of cycles of prescribed length is essentially best possible for all such k, where we regard k as fixed and n large.

For $k \equiv 0 \pmod 3$ the blow-up of a triangle contains many cycles of length k, but does not contain paths of length k between all pairs of vertices. Adding an extra vertex u with $N^+(u) = V_1$ and $N^-(u) = V_0$ gives an oriented graph with minimum semi-degree $\geq \lfloor n/3 \rfloor - 1$ with no cycle of length $k \equiv 0 \pmod 3$ containing u.

This example also gives a lower bound of $\lfloor n/3 \rfloor$ for the existence of paths of length exactly k ($k \geq 4$) between every pair of vertices, and so our result is essentially best possible here for all $k \geq 5$. Note that proving the existence of paths of length *at most* 4 between any vertices in an oriented graph on n vertices with minimum semi-degree $3n/8+1$ is very straightforward (a simple averaging argument suffices), the difficulty is in showing the existence of a path of prescribed length.

Chapter 5

Proof of our main theorem

5.1. Extremal Example

We now prove our main result, Theorem 1.3, and show how this can be extended to an Ore type result. In this section we give the extremal example. In Section 5.2 we discuss some machinery which we need, in particular the Diregularity Lemma. In Section 5.3 we prove our main technical lemmas and in Section 5.4 we prove the result. Finally in Section 5.5 we prove Theorem 1.4.

The following construction of Häggkvist [**30**] shows that Conjecture 1.2 is best possible for infinitely many values of $|G|$. We include it here for completeness.

Proposition 5.1. *There are infinitely many oriented graphs G with minimum semi-degree $(3|G|-5)/8$ which do not contain a 1-factor and thus do not contain a Hamilton cycle.*

Proof. Let $n := 4m+3$ for some odd $m \in \mathbb{N}$. Let G be the oriented graph obtained from the disjoint union of two regular tournaments A and C on m vertices, a set B of $m+2$ vertices and a set D of $m+1$ vertices by adding all edges from A to B, all edges from B to C, all edges from C to D as well as all edges from D to A. Finally, between B and D we add edges to obtain a bipartite tournament which is as regular as possible, i.e. the in- and outdegree of every vertex differ by at most 1. So in particular every vertex in B sends exactly $(m+1)/2$ edges to D (Figure 1).

It is easy to check that the minimum semi-degree of G is $(m-1)/2+(m+1) = (3n-5)/8$, as required. Since every path which joins two vertices in B has to pass through D, it follows that every cycle contains at least as many vertices from D as it contains from B. As $|B| > |D|$ this means that one cannot cover all the vertices of G by disjoint cycles, i.e. G does not contain a 1-factor. $\square$

5.2. The Diregularity Lemma and Other Tools

In this section we collect all the information we need about the Diregularity lemma and the Blow-up lemma. See [**40**] for a survey on the Regularity lemma and [**38**] for a survey on the Blow-up lemma. We start with some more notation. Recall that the density of a bipartite graph $G = (A, B)$ with vertex classes A and B is defined to be

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

We often write $d(A, B)$ if this is unambiguous. Given $\varepsilon > 0$, we say that G is *ε-regular* if for all subsets $X \subseteq A$ and $Y \subseteq B$ with $|X| > \varepsilon\,|A|$ and $|Y| > \varepsilon\,|B|$ we have that $|d(X, Y) - d(A, B)| < \varepsilon$. Given $d \in [0, 1]$ we

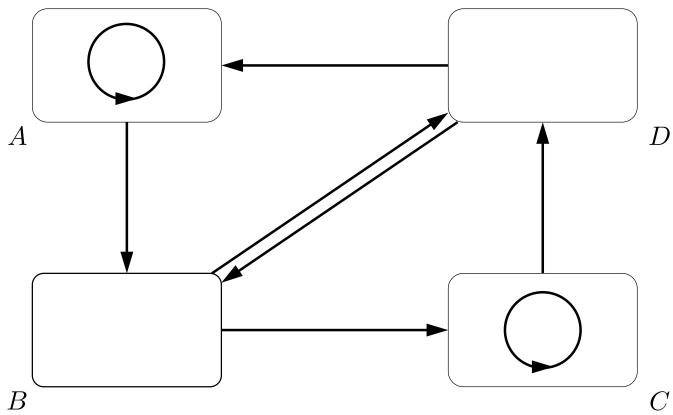


FIGURE 1. The oriented graph in the proof of Proposition 5.1.

say that G is (ε, d)-*super-regular* if it is ε-regular and furthermore $d_G(a) \geq (d-\varepsilon)|B|$ for all $a \in A$ and $d_G(b) \geq (d-\varepsilon)|A|$ for all $b \in B$. (This is a slight variation of the standard definition of (ε, d)-super-regularity where one requires $d_G(a) \geq d|B|$ and $d_G(b) \geq d|A|$.)

The Diregularity lemma is a version of the Regularity lemma for digraphs due to Alon and Shapira [**2**]. Its proof is quite similar to the undirected version. We will use the degree form of the Diregularity lemma which can be easily derived (see e.g. [**59**]) from the standard version, in exactly the same manner as the undirected degree form.

LEMMA 5.2 (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 vertices of G into $V_0, V_1, \ldots, V_k$ and a spanning subdigraph G' of G such that the following holds:*

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

$V_1, \ldots, V_k$ are called *clusters*, V_0 is called the *exceptional set* and the vertices in V_0 are called *exceptional vertices*. The last condition of the lemma says that all pairs of clusters are ε-regular in both directions (but possibly with different densities). We call the spanning digraph $G' \subseteq G$ given by the Diregularity lemma the *pure digraph*. Given clusters $V_1, \ldots, V_k$ and the pure digraph G', the *reduced digraph* R' is the digraph whose vertices are $V_1, \ldots, V_k$ and in which V_iV_j is an edge if and only if G' contains a V_i-V_j edge. Note that the latter holds if and only if the bipartite graph whose vertex classes are V_i and V_j and whose edges are all the V_i-V_j edges in G' is ε-regular and has density at least d. It turns out that R' inherits many properties of G, a fact that is crucial in our proof. However, R' is not necessarily oriented even if the original digraph G is, but the next lemma shows that by discarding edges with appropriate probabilities one can go over to a reduced oriented graph $R \subseteq R'$ which still inherits many of the properties of G.

LEMMA 5.3. *For every $\varepsilon \in (0,1)$ there exist integers $M' = M'(\varepsilon)$ and $n_0 = n_0(\varepsilon)$ such that the following holds. Let $d \in [0,1]$, let G be an oriented graph of order at least n_0 and let R' be the reduced digraph obtained by applying the Diregularity lemma to G with parameters ε, d and M'. Then R' has a spanning oriented subgraph R with*

(a) $\delta^+(R) \geq (\delta^+(G)/|G| - (3\varepsilon + d))\,|R|$,
(b) $\delta^-(R) \geq (\delta^-(G)/|G| - (3\varepsilon + d))\,|R|$,
(c) $\delta(R) \geq (\delta(G)/|G| - (3\varepsilon + 2d))\,|R|$.

Proof. Let us first show that every cluster V_i satisfies

$$|N_{R'}(V_i)|/|R'| \geq \delta(G)/|G| - (3\varepsilon + 2d). \tag{7}$$

To see this, consider any vertex $x \in V_i$. As G is an oriented graph, the Diregularity lemma implies that $|N_{G'}(x)| \geq \delta(G) - 2(d+\varepsilon)|G|$. On the other hand, $|N_{G'}(x)| \leq |N_{R'}(V_i)|m + |V_0| \leq |N_{R'}(V_i)||G|/|R'| + \varepsilon|G|$. Altogether this proves (7).

We first consider the case when

$$\delta^+(G)/|G| \geq 3\varepsilon + d \quad \text{and} \quad \delta^-(G)/|G| \geq 3\varepsilon + d. \tag{8}$$

Let R be the spanning oriented subgraph obtained from R' by deleting edges randomly as follows. For every unordered pair V_i, V_j of clusters we delete the edge V_iV_j (if it exists) with probability

$$\frac{e_{G'}(V_j, V_i)}{e_{G'}(V_i, V_j) + e_{G'}(V_j, V_i)}. \tag{9}$$

Otherwise we delete V_jV_i (if it exists). We interpret (9) as 0 if $V_iV_j, V_jV_i \notin E(R')$. So if R' contains at most one of the edges V_iV_j, V_jV_i then we do nothing. We do this for all unordered pairs of clusters independently and let X_i be the random variable which counts the number of outedges of the vertex $V_i \in R$. Then

$$\begin{aligned}
\mathbb{E}(X_i) &= \sum_{j \neq i} \frac{e_{G'}(V_i, V_j)}{e_{G'}(V_i, V_j) + e_{G'}(V_j, V_i)} \geq \sum_{j \neq i} \frac{e_{G'}(V_i, V_j)}{|V_i|\,|V_j|} \\
&\geq \frac{|R'|}{|G|\,|V_i|} \sum_{x \in V_i} (d^+_{G'}(x) - |V_0|) \geq (\delta^+(G')/|G| - \varepsilon)\,|R| \\
&\geq (\delta^+(G)/|G| - (2\varepsilon + d))\,|R| \overset{(8)}{\geq} \varepsilon|R|.
\end{aligned}$$

A Chernoff-type bound (see e.g. [**3**, Cor. A.14]) now implies that there exists a constant $c = c(\varepsilon)$ such that

$$\begin{aligned}
\mathbb{P}(X_i < (\delta^+(G)/|G| - (3\varepsilon + d))\,|R|) &\leq \mathbb{P}(|X_i - \mathbb{E}(X_i)| > \varepsilon\mathbb{E}(X_i)) \\
&\leq \mathrm{e}^{-c\mathbb{E}(X_i)} \leq \mathrm{e}^{-c\varepsilon|R|}.
\end{aligned}$$

Writing Y_i for the random variable which counts the number of inedges of the vertex V_i in R, it follows similarly that

$$\mathbb{P}(Y_i < (\delta^-(G)/|G| - (3\varepsilon + d))\,|R|) \leq \mathrm{e}^{-c\varepsilon|R|}.$$

As $2|R|\mathrm{e}^{-c\varepsilon|R|} < 1$ if M' is chosen to be sufficiently large compared to ε, this implies that there is some outcome R with $\delta^+(R) \geq (\delta^+(G)/|G| - (3\varepsilon + d))\,|R|$ and $\delta^-(R) \geq (\delta^-(G)/|G| - (3\varepsilon + d))\,|R|$. But $N_{R'}(V_i) = N_R(V_i)$ for every cluster V_i and so (7) implies that $\delta(R) \geq (\delta(G)/|G| - (3\varepsilon + 2d))|R|$. Altogether this shows that R is as required in the lemma.

If neither of the conditions in (8) hold, then (a) and (b) are trivial and one can obtain an oriented graph R which satisfies (c) from R' by arbitrarily deleting one edge from each double edge. If exactly one of the conditions in (8) holds, say the first, then (b) is trivial. To obtain an oriented graph R which satisfies (a) we consider the X_i as before, but ignore the Y_i. Again, $N_{R'}(V_i) = N_R(V_i)$ for every cluster V_i and so (c) is also satisfied. □

The oriented graph R given by Lemma 5.3 is called the *reduced oriented graph*. The spanning oriented subgraph G^* of the pure digraph G' obtained by deleting all the V_i-V_j edges whenever $V_iV_j \in E(R')\setminus E(R)$ is called the *pure oriented graph*. Given an oriented subgraph $S \subseteq R$, the *oriented subgraph of G^* corresponding to S* is the oriented subgraph obtained from G^* by deleting all those vertices that lie in clusters not belonging to S as well as deleting all the V_i-V_j edges for all pairs V_i, V_j with $V_iV_j \notin E(S)$.

In our proof of Theorem 1.3 we will also need the Blow-up lemma, in both the original form of Komlós, Sárközy and Szemerédi [**39**] and a recent strengthening due to Csaba [**20**]. Roughly speaking, they say that an r-partite graph formed by r clusters such that all the pairs of these clusters are dense and ε-regular behaves like a complete r-partite graph with respect to containing graphs H of bounded maximum degree as subgraphs.

Lemma 5.4 (Blow-up Lemma, Komlós, Sárközy and Szemerédi [**39**]). *Given a graph F on $[k]$ and positive numbers d, b, Δ, there exist positive numbers $\eta_0 = \eta_0(d, \Delta, k)$ and $\alpha = \alpha(d, \Delta, r, k) \leq 1/2$ such that the following holds for all positive numbers $\ell_1, \dots, \ell_k$ and all $0 < \eta \leq \eta_0$. Let F' be the graph obtained from F by replacing each vertex $i \in F$ with a set V_i of ℓ_i new vertices and joining all vertices in V_i to all vertices in V_j whenever ij is an edge of F. Let G' be a spanning subgraph of F' such that for every edge $ij \in F$ the graph $(V_i, V_j)_{G'}$ is (η, d)-super-regular. Then G' contains a copy of every subgraph H of F' with $\Delta(H) \leq \Delta$. Moreover, this copy of H in G' maps the vertices of H to the same sets V_i as the copy of H in F', i.e. if $h \in V(H)$ is mapped to V_i by the copy of H in F', then it is also mapped to V_i by the copy of H in G'.*

Furthermore, we can additionally require that for vertices $x \in H \subseteq R(L)$ lying in V_i their images in the copy of H in G are contained in (arbitrary) given sets $C_x \subseteq V_i$ provided that $|C_x| \geq bL$ for each such x and provided that in each V_i there are at most αL such vertices x.

Observe that in this version the pairs of clusters have to be super-regular and ε has to be sufficiently small compared to $1/r$ (and so in particular we cannot take $r = |R|$). We also need a stronger (and more technical) version due to Csaba [**20**], which allows us to take $r = |R|$ and does not demand super-regularity. The case when $\Delta = 3$ of this is implicit in [**21**].

In the statement of Lemma 5.5 and later on we write $0 < a_1 \ll a_2 \ll a_3$ 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 $a_2 \le f(a_3)$ and $a_1 \le g(a_2)$, all calculations needed in the proof of Lemma 5.5 are valid. Hierarchies with more constants are defined in the obvious way.

Lemma 5.5 (Blow-up Lemma, Csaba [**20**]). *For all integers Δ, K_1, K_2, K_3 and every positive constant c there exists an integer N such that whenever $\varepsilon, \varepsilon', \delta', d$ are positive constants with*

$$0 < \varepsilon \ll \varepsilon' \ll \delta' \ll d \ll 1/\Delta, 1/K_1, 1/K_2, 1/K_3, c$$

the following holds. Suppose that G^* is a graph of order $n \ge N$ and $V_0, \dots, V_k$ is a partition of $V(G^*)$ such that the bipartite graph $(V_i, V_j)_{G^*}$ is ε-regular with density either 0 or d for all $1 \le i < j \le k$. Let H be a graph on n vertices with $\Delta(H) \le \Delta$ and let $L_0 \cup L_1 \cup \dots \cup L_k$ be a partition of $V(H)$ with $|L_i| = |V_i| =: m$ for every $i = 1, \dots, k$. Furthermore, suppose that there exists a bijection $\phi : L_0 \to V_0$ and a set $I \subseteq V(H)$ of vertices at distance at least 4 from each other such that the following conditions hold:

(C1) $|L_0| = |V_0| \le K_1 dn$.
(C2) $L_0 \subseteq I$.
(C3) *L_i is independent for every $i = 1, \dots, k$.*
(C4) *$|N_H(L_0) \cap L_i| \le K_2 dm$ for every $i = 1, \dots, k$.*
(C5) *For each $i = 1, \dots, k$ there exists $D_i \subseteq I \cap L_i$ with $|D_i| = \delta' m$ and such that for $D := \bigcup_{i=1}^{k} D_i$ and all $1 \le i < j \le k$*

$$\big||N_H(D) \cap L_i| - |N_H(D) \cap L_j|\big| < \varepsilon m.$$

(C6) *If $xy \in E(H)$ and $x \in L_i, y \in L_j$ then $(V_i, V_j)_{G^*}$ is ε-regular with density d.*
(C7) *If $xy \in E(H)$ and $x \in L_0, y \in L_j$ then $|N_{G^*}(\phi(x)) \cap V_j| \ge cm$.*
(C8) *For each $i = 1, \dots, k$, given any $E_i \subseteq V_i$ with $|E_i| \le \varepsilon' m$ there exists a set $F_i \subseteq (L_i \cap (I \setminus D))$ and a bijection $\phi_i : E_i \to F_i$ such that $|N_{G^*}(v) \cap V_j| \ge (d - \varepsilon)m$ whenever $N_H(\phi_i(v)) \cap L_j \ne \emptyset$ (for all $v \in E_i$ and all $j = 1, \dots, k$).*
(C9) *Writing $F := \bigcup_{i=1}^{k} F_i$ we have that $|N_H(F) \cap L_i| \le K_3 \varepsilon' m$.*

Then G^* contains a copy of H such that the image of L_i is V_i for all $i = 1, \dots, k$ and the image of each $x \in L_0$ is $\phi(x) \in V_0$.

The additional properties of the copy of H in G^* are not included in the statement of the lemma in [**20**] but are stated explicitly in the proof.

Let us briefly motivate the conditions of the Blow-up lemma. The embedding of H into G guaranteed by the Blow-up lemma is found by a randomised algorithm which first embeds each vertex $x \in L_0$ to $\phi(x)$ and then successively embeds the remaining vertices of H. So the image of L_0 will be the exceptional set V_0. Condition (C1) requires that there are not too many exceptional vertices and (C2) ensures that we can embed the vertices in L_0 without affecting the neighbourhood of other such vertices. As L_i will be embedded into V_i we need to have (C3). Condition (C5) gives us a reasonably large set D of 'buffer vertices' which will be embedded last by the randomised algorithm. (C6) requires that edges between vertices of $H - L_0$ are embedded into ε-regular pairs of density d. (C7) ensures that the exceptional vertices have large degree in all 'neighbouring clusters'. (C8) and (C9) allow us to embed those vertices whose set of candidate images in G^* has grown very small at some point of the algorithm. Conditions (C6), (C8) and (C9) correspond to a substantial weakening of the super-regularity that the usual form of the Blow-up lemma requires, namely that whenever H contains an edge xy with and $x \in L_i, y \in L_j$ then $(V_i, V_j)_{G^*}$ is (ε, d)-super-regular.

We would like to apply the Blow-up lemma with G^* being obtained from the underlying graph of the pure oriented graph by adding the exceptional vertices. It will turn out that in order to satisfy (C8), it suffices to ensure that all the edges of a suitable 1-factor in the reduced oriented graph R correspond to (ε, d)-superregular pairs of clusters. A well-known simple fact (see the first part of the proof of Proposition 5.6) states that this can be ensured by removing a small proportion of vertices from each cluster V_i, and so (C8) will be satisfied. However, (C6) requires all the edges of R to correspond to ε-regular pairs of density precisely d and not just at least d. (As remarked by Csaba [**20**], it actually suffices that the densities are close to d in terms of ε.) The second part of the following proposition shows that this too does not pose a problem.

Proposition 5.6. *Let M', n_0, D be integers and let ε, d be positive constants such that $1/n_0 \ll 1/M' \ll \varepsilon \ll d \ll 1/D$. Let G be an oriented graph of order at least n_0. Let R be the reduced oriented graph and let G^* be the pure oriented graph obtained by successively applying first the Diregularity lemma with parameters ε, d and M' to G and then Lemma 5.3. Let S be an oriented subgraph of R with $\Delta(S) \leq D$. Let G' be the underlying graph of G^*. Then one can delete $2D\varepsilon|V_i|$ vertices from each cluster V_i to obtain subclusters $V_i' \subseteq V_i$ in such a way that G' contains a subgraph G'_S whose vertex set is the union of all the V_i' and such that*

- *$(V_i', V_j')_{G'_S}$ is $(\sqrt{\varepsilon}, d - 4D\varepsilon)$-superregular whenever $V_iV_j \in E(S)$,*
- *$(V_i', V_j')_{G'_S}$ is $\sqrt{\varepsilon}$-regular and has density $d - 4D\varepsilon$ whenever $V_iV_j \in E(R)$.*

Proof. Consider any cluster $V_i \in V(S)$ and any neighbour V_j of V_i in S. Recall that $m = |V_i|$. Let d_{ij} denote the density of the bipartite subgraph $(V_i, V_j)_{G'}$ of G' induced by V_i and V_j. So $d_{ij} \geq d$ and this bipartite graph is ε-regular. Thus there are at most $2\varepsilon m$ vertices $v \in V_i$ such that $||N_{G'}(v) \cap V_j| - d_{ij}m| > \varepsilon m$. So in total there are at most $2D\varepsilon m$ vertices $v \in V_i$ such that $||N_{G'}(v) \cap V_j| - d_{ij}m| > \varepsilon m$ for some neighbour V_j of V_i in S. Delete all these vertices as well as some more vertices if necessary to obtain a subcluster $V_i' \subseteq V_i$ of size $(1 - 2D\varepsilon)m =: m'$. Delete any $2D\varepsilon m$ vertices from each cluster $V_i \in V(R) \setminus V(S)$ to obtain a subcluster V_i'. It is easy to check that for each edge $V_iV_j \in E(R)$ the graph $(V_i', V_j')_{G'}$ is still 2ε-regular and that its density d'_{ij} satisfies

$$d' := d - 4D\varepsilon < d_{ij} - \varepsilon \leq d'_{ij} \leq d_{ij} + \varepsilon.$$

Moreover, whenever $V_iV_j \in E(S)$ and $v \in V_i'$ we have that

$$(d_{ij} - 4D\varepsilon)m' \le |N_{G'}(v) \cap V_j'| \le (d_{ij} + 4D\varepsilon)m'.$$

For every pair V_i, V_j of clusters with $V_iV_j \in E(S)$ we now consider a spanning random subgraph G_{ij}' of $(V_i', V_j')_{G'}$ which is obtained by choosing each edge of $(V_i', V_j')_{G'}$ with probability d'/d_{ij}', independently of the other edges. Consider any vertex $v \in V_i'$. Then the expected number of neighbours of v in V_j' (in the graph G_{ij}') is at least $(d_{ij} - 4D\varepsilon)d'm'/d_{ij}' \ge (1 - \sqrt{\varepsilon})d'm'$. So we can apply a Chernoff-type bound to see that there exists a constant $c = c(\varepsilon)$ such that

$$\mathbb{P}(|N_{G_{ij}'}(v) \cap V_j'| \le (d' - \sqrt{\varepsilon})m') \le \mathrm{e}^{-cd'm'}.$$

Similarly, whenever $X \subseteq V_i'$ and $Y \subseteq V_j'$ are sets of size at least $2\varepsilon m'$ the expected number of X-Y edges in G_{ij}' is $d_{G'}(X, Y)d'|X||Y|/d_{ij}'$. Since $(V_i', V_j')_{G'}$ is 2ε-regular this expected number lies between $(1 - \sqrt{\varepsilon})d'|X||Y|$ and $(1 + \sqrt{\varepsilon})d'|X||Y|$. So again we can use a Chernoff-type bound to see that

$$\mathbb{P}(|e_{G_{ij}'}(X, Y) - d'|X||Y|| > \sqrt{\varepsilon}|X||Y|) \le \mathrm{e}^{-cd'|X||Y|} \le \mathrm{e}^{-4cd'(\varepsilon m')^2}.$$

Moreover, with probability at least $1/(3m')$ the graph G_{ij}' has its expected density d' (see e.g. [**8**, p. 6]). Altogether this shows that with probability at least

$$1/(3m') - 2m'\mathrm{e}^{-cd'm'} - 2^{2m'}\mathrm{e}^{-4cd'(\varepsilon m')^2},$$

which is greater than 0 for sufficiently large m', we have that G_{ij}' is $(\sqrt{\varepsilon}, d')$-superregular and has density d'. Proceed similarly for every pair of clusters forming an edge of S. An analogous argument applied to a pair V_i, V_j of clusters with $V_iV_j \in E(R) \setminus E(S)$ shows that with non-zero probability the random subgraph G_{ij}' is $\sqrt{\varepsilon}$-regular and has density d'. Altogether this gives us the desired subgraph G_S' of G'. $\square$

5.2.1. Overview of the proof of Theorem 1.3. Let G be our given oriented graph. The rough idea of the proof is to apply the Diregularity lemma and Lemma 5.3 to obtain a reduced oriented graph R and a pure oriented graph G^*. The following result of Häggkvist implies that R contains a 1-factor.

THEOREM 5.7 (Häggkvist [**30**]). *Let R be an oriented graph with $\delta^*(R) > (3|R| - 3)/2$. Then R is strongly connected and contains a 1-factor.*

So one can apply the Blow-up lemma (together with Proposition 5.6) to find a 1-factor in $G^* - V_0 \subseteq G - V_0$. One now would like to glue the cycles of this 1-factor together and to incorporate the exceptional vertices to obtain a Hamilton cycle of G^* and thus of G. However, we were only able to find a method which incorporates a set of vertices whose size is small compared to the cluster size m. This is not necessarily the case for V_0. So we proceed as follows. We first choose a random partition of the vertex set of G into two sets A and $V(G) \setminus A$ having roughly equal size. We then apply the Diregularity lemma to $G - A$ in order to obtain clusters $V_1, \dots, V_k$ and an exceptional set V_0. We let m denote the size of these clusters and set $B := V_1 \cup \dots V_k$. By arguing as indicated above, we can find a Hamilton cycle C_B in $G[B]$. We then apply the Diregularity lemma to $G-B$, but with an ε which is small compared to $1/k$, to obtain clusters $V_1', \dots, V_\ell'$ and an exceptional set V_0'. Since the choice of our partition $A, V(G) \setminus A$ will imply that $\delta^*(G-B) \geq (3/2+\alpha/2)|G-B|$ we can again argue as before to obtain a cycle C_A which covers precisely the vertices in $A' := V_1' \cup \dots \cup V_\ell'$. Since we have chosen ε to be small compared to $1/k$, the set V_0' of exceptional vertices is now small enough to be incorporated into our first cycle C_B. (Actually, C_B is only determined at this point and not yet earlier on.) Moreover, by choosing C_B and C_A suitably we can ensure that they can be joined together into the desired Hamilton cycle of G.

5.3. Shifted Walks

In this section we will introduce the tools we need in order to glue certain cycles together and to incorporate the exceptional vertices. Let R^* be a digraph and let $\mathcal{C}$ be a collection of disjoint cycles in R^*. We call a closed walk W in R^* *balanced w.r.t.* $\mathcal{C}$ if

- for each cycle $C \in \mathcal{C}$ the walk W visits all the edges on C an equal number of times,
- W visits every vertex of R^*,
- every vertex not in any cycle from $\mathcal{C}$ is visited exactly once.

Let us now explain why balanced walks are helpful in order to incorporate the exceptional vertices. Suppose that $\mathcal{C}$ is a 1-factor of the reduced oriented graph R and that R^* is obtained from R by adding all the exceptional vertices $v \in V_0$ and adding an edge vV_i (where V_i is a cluster) whenever v sends edges to a significant proportion of the vertices in V_i, say we add vV_i whenever v sends at least cm edges to V_i. (Recall that m denotes the size of the clusters.) The edges in R^* of the form V_iv are defined in a similar way. Let G^c be the oriented graph obtained from the pure oriented graph G^* by making all the nonempty bipartite subgraphs between the clusters complete (and orienting all the edges between these clusters in the direction induced by R) and adding the vertices in V_0 as well as all the edges of G between V_0 and $V(G) \setminus V_0$. Suppose that W is a balanced closed walk in R^* which visits all the vertices lying on a cycle $C \in \mathcal{C}$ precisely $m_C \leq m$ times. Furthermore, suppose that $|V_0| \leq cm/2$ and that the vertices in V_0 have distance at least 3 from each other on W. Then by 'winding around' each cycle $C \in \mathcal{C}$ precisely $m - m_C$ times (at the point when W first visits C) we can obtain a Hamilton cycle in G^c. Indeed, the two conditions on V_0 ensure that the neighbours of each $v \in V_0$ on the Hamilton cycle can be chosen amongst the at least cm neighbours of v in the neighbouring clusters of v on W in such a way that they are distinct for different exceptional vertices. The idea then is to apply the Blow-up lemma to show that this Hamilton cycle corresponds to one in G. So our aim is to find such a balanced closed walk in R^*. However, as indicated in Section 5.2.1, the difficulties arising when trying to ensure that the exceptional vertices lie on this walk will force us to apply the above argument to the subgraphs induced by a random partition of our given oriented graph G.

Let us now go back to the case when R^* is an arbitrary digraph and $\mathcal{C}$ is a collection of disjoint cycles in R^*. Given vertices $a, b \in R^*$, a *shifted a-b* *walk* is a walk of the form

$$W = aa_1C_1b_1a_2C_2b_2 \dots a_tC_tb_tb$$

where $C_1, \dots, C_t$ are (not necessarily distinct) cycles from $\mathcal{C}$ and a_i is the successor of b_i on C_i for all $i \leq t$. (We might have $t = 0$. So an edge ab is a shifted a-b walk.) We call $C_1, \dots, C_t$ the cycles which are *traversed* by W. So even if the cycles $C_1, \dots, C_t$ are not distinct, we say that W traverses t cycles. Note that for every cycle $C \in \mathcal{C}$ the walk $W - \{a, b\}$ visits the vertices on C an equal number of times. Thus it will turn out that by joining the cycles from $\mathcal{C}$ suitably via shifted walks and incorporating those vertices of R^* not covered by the cycles from $\mathcal{C}$ we can obtain a balanced closed walk on R^*.

Our next lemma will be used to show that if R^* is oriented and $\delta^*(R^*) \geq (3/2 + \alpha)|R^*|$ then any two vertices of R^* can be joined by a shifted walk traversing only a small number of cycles from $\mathcal{C}$ (see Corollary 5.10). The lemma itself shows that the δ^* condition implies expansion, and this will give us the 'expansion with respect to shifted neighbourhoods' we need for the existence of shifted walks. The proof of Lemma 5.8 is similar to that of Theorem 5.7.

LEMMA 5.8. *Let R^* be an oriented graph on N vertices with $\delta^*(R^*) \geq (3/2+\alpha)N$ for some $\alpha > 0$. If $X \subseteq V(R^*)$ is nonempty and $|X| \leq (1-\alpha)N$ then $|N^+(X)| \geq |X| + \alpha N/2$.*

Proof. For simplicity, we write $\delta := \delta(R^*)$, $\delta^+ := \delta^+(R^*)$ and $\delta^- := \delta^-(R^*)$. Suppose the assertion is false, i.e. there exists $X \subseteq V(R^*)$ with $|X| \leq (1 - \alpha)N$ and

$$\big|N^+(X)\big| < |X| + \alpha N/2. \tag{10}$$

We consider the following partition of $V(R^*)$:

$A := X \cap N^+(X)$, $B := N^+(X) \setminus X$, $C := V(R^*) \setminus (X \cup N^+(X))$, $D := X \setminus N^+(X)$.

(10) gives us

$$|D| + \alpha N/2 > |B|. \tag{11}$$

Suppose $A \neq \emptyset$. Then by an averaging argument there exists $x \in A$ with $|N^+(x) \cap A| < |A|/2$. Hence $\delta^+ \leq |N^+(x)| < |B| + |A|/2$. Combining this with (11) we get

$$|A| + |B| + |D| \geq 2\delta^+ - \alpha N/2. \tag{12}$$

If $A = \emptyset$ then $N^+(X) = B$ and so (11) implies $|D| + \alpha N/2 \geq |B| \geq \delta^+$. Thus (12) again holds. Similarly, if $C \neq \emptyset$ then considering the inneighbourhood of a suitable vertex $x \in C$ gives

$$|B| + |C| + |D| \geq 2\delta^- - \alpha N/2. \tag{13}$$

If $C = \emptyset$ then the fact that $|X| \leq (1 - \alpha)N$ and (10) together imply that $D \neq \emptyset$. But then $N^-(D) \subseteq B$ and thus $|B| \geq \delta^-$. Together with (11) this shows that (13) holds in this case too.

If $D = \emptyset$ then trivially $|A| + |B| + |C| = N \geq \delta$. If not, then for any $x \in D$ we have $N(x) \cap D = \emptyset$ and hence

$$|A| + |B| + |C| \geq |N(x)| \geq \delta. \tag{14}$$

Combining (12), (13) and (14) gives

$$3|A| + 4|B| + 3|C| + 2|D| \geq 2\delta^- + 2\delta^+ + 2\delta - \alpha N = 2\delta^*(R^*) - \alpha N.$$

Finally, substituting (11) gives

$$3N + \alpha N/2 \geq 2\delta^*(R^*) - \alpha N \geq 3N + \alpha N,$$

which is a contradiction. □

As indicated before, we will now use Lemma 5.8 to prove the existence of shifted walks in R^* traversing only a small number of cycles from a given 1-factor of R^*. For this (and later on) the following fact will be useful.

FACT 5.9. *Let G be an oriented graph with $\delta^*(G) \geq (3/2 + \alpha)|G|$ for some constant $\alpha > 0$. Then $\delta^0(G) > \alpha|G|$.*

Proof. Suppose that $\delta^-(G) \leq \alpha|G|$. As G is oriented we have that $\delta^+(G) < |G|/2$ and so $\delta^*(G) < 3n/2 + \alpha|G|$, a contradiction. The proof for $\delta^+(G)$ is similar. □

COROLLARY 5.10. *Let R^* be an oriented graph on N vertices with $\delta^*(R^*) \geq (3/2 + \alpha)N$ for some $\alpha > 0$ and let $\mathcal{C}$ be a 1-factor in R^*. Then for any distinct $x, y \in V(R^*)$ there exists a shifted x-y walk traversing at most $2/\alpha$ cycles from $\mathcal{C}$.*

Proof. Let X_i be the set of vertices v for which there is a shifted x-v walk which traverses at most i cycles. So $X_0 = N^+(x) \neq \emptyset$ and $X_{i+1} = N^+(X_i^-) \cup X_i$, where X_i^- is the set of all predecessors of the vertices in X_i on the cycles from $\mathcal{C}$. Suppose that $|X_i| \leq (1 - \alpha)N$. Then Lemma 5.8 implies that

$$|X_{i+1}| \geq |N^+(X_i^-)| \geq |X_i^-| + \alpha N/2 = |X_i| + \alpha N/2.$$

So for $i^* := \lfloor 2/\alpha \rfloor - 1$, we must have $|X_{i^*}^-| = |X_{i^*}| \geq (1-\alpha)N$. But $|N^-(y)| \geq \delta^-(R^*) > \alpha N$ and so $N^-(y) \cap X_{i^*}^- \neq \emptyset$. In other words, $y \in N^+(X_{i^*}^-)$ and so there is a shifted x-y walk traversing at most $i^* + 1$ cycles. □

COROLLARY 5.11. *Let R^* be an oriented graph with $\delta^*(R^*) \geq (3/2 + \alpha)|R^*|$ for some $0 < \alpha \leq 1/6$ and let $\mathcal{C}$ be a 1-factor in R^*. Then R^* contains a closed walk which is balanced w.r.t. $\mathcal{C}$ and meets every vertex at most $|R^*|/\alpha$ times and traverses each edge lying on a cycle from $\mathcal{C}$ at least once.*

Proof. Let $C_1, \dots, C_s$ be an arbitrary ordering of the cycles in $\mathcal{C}$. For each cycle C_i pick a vertex $c_i \in C_i$. Denote by c_i^+ the successor of c_i on the cycle C_i. Corollary 5.10 implies that for all i there exists a shifted c_i-c_{i+1}^+ walk W_i traversing at most $2/\alpha$ cycles from $\mathcal{C}$, where $c_{s+1} := c_1$. Then the closed walk

$$W' := c_1^+ C_1 c_1 W_1 c_2^+ C_2 c_2 \dots W_{s-1} c_s^+ C_s c_s W_s c_1^+$$

is balanced w.r.t. $\mathcal{C}$ by the definition of shifted walks. Since each shifted walk W_i traverses at most $2/\alpha$ cycles of $\mathcal{C}$, the closed walk W meets each vertex at most $(|R^*|/3)(2/\alpha) + 1$ times. Let W denote the walk obtained from W' by 'winding around' each cycle $C \in \mathcal{C}$ once more. (That is, for each $C \in \mathcal{C}$ pick a vertex v on C and replace one of the occurences of v on W' by vCv.) Then W is still balanced w.r.t. $\mathcal{C}$, traverses each edge lying on a cycle from $\mathcal{C}$ at least once and visits each vertex of R^* at most $(|R^*|/3)(2/\alpha) + 2 \leq |R^*|/\alpha$ times as required. $\square$

5.4. Proof of Theorem 1.3

5.4.1. Partitioning G and applying the Diregularity lemma. Let G be an oriented graph on n vertices with $\delta^*(G) \geq (3/2+\alpha)n$ for some constant $\alpha > 0$. Clearly we may assume that $\alpha \ll 1$. Define positive constants ε, d and integers M'_A, M'_B such that

$$1/M'_A \ll 1/M'_B \ll \varepsilon \ll d \ll \alpha \ll 1.$$

Throughout this section, we will assume that n is sufficiently large compared to M'_A for our estimates to hold. Choose a subset $A \subseteq V(G)$ with $(1/2 - \varepsilon)n \leq |A| \leq (1/2+\varepsilon)n$ and such that every vertex $x \in G$ satisfies

$$\frac{d^+(x)}{n} - \frac{\alpha}{10} \leq \frac{|N^+(x) \cap A|}{|A|} \leq \frac{d^+(x)}{n} + \frac{\alpha}{10}$$

and such that $N^-(x) \cap A$ satisfies a similar condition. (The existence of such a set A can be shown by considering a random partition of $V(G)$.) Apply the Diregularity lemma (Lemma 5.2) with parameters ε^2, $d + 8\varepsilon^2$ and M'_B to $G - A$ to obtain a partition of the vertex set of $G - A$ into $k \geq M'_B$ clusters $V_1, \dots, V_k$ and an exceptional set V_0. Set $B := V_1 \cup \ldots \cup V_k$ and $m_B := |V_1| = \cdots = |V_k|$. Let R_B denote the reduced oriented graph obtained by an application of Lemma 5.3 and let G^*_B be the pure oriented graph. Since $\delta^+(G-A)/|G-A| \geq \delta^+(G)/n - \alpha/9$ by our choice of A, Lemma 5.3 implies that

$$\delta^+(R_B) \geq (\delta^+(G)/n - \alpha/8)|R_B|. \tag{15}$$

Similarly

$$\delta^-(R_B) \geq (\delta^-(G)/n - \alpha/8)|R_B| \tag{16}$$

and $\delta(R_B) \geq (\delta(G)/n - \alpha/4)|R_B|$. Altogether this implies that

$$\delta^*(R_B) \geq (3/2 + \alpha/2)|R_B|. \tag{17}$$

So Theorem 5.7 gives us a 1-factor $\mathcal{C}_B$ of R_B. We now apply Proposition 5.6 with $\mathcal{C}_B$ playing the role of S, ε^2 playing the role of ε and $d + 8\varepsilon^2$ playing the role of d. This shows that by adding at most $4\varepsilon^2 n$ further vertices to the exceptional set V_0 we may assume that each edge of R_B corresponds to an ε-regular pair of density d (in the underlying graph of G^*_B) and that each edge in the union $\bigcup_{C \in \mathcal{C}_B} C \subseteq R_B$ of all the cycles from $\mathcal{C}_B$ corresponds to an (ε, d)-superregular pair. (More formally, this means that we replace the clusters with the subclusters given by Proposition 5.6 and replace G^*_B with its oriented subgraph obtained by deleting all edges not corresponding to edges of the graph $G'_{\mathcal{C}_B}$ given by Proposition 5.6, i.e. the underlying graph of G^*_B will now be $G'_{\mathcal{C}_B}$.) Note that the new exceptional set now satisfies $|V_0| \leq \varepsilon n$.

Apply Corollary 5.11 with $R^* := R_B$ to find a closed walk W_B in R_B which is balanced w.r.t. $\mathcal{C}_B$, meets every cluster at most $2|R_B|/\alpha$ times and traverses all the edges lying on a cycle from $\mathcal{C}_B$ at least once.

Let G^c_B be the oriented graph obtained from G^*_B by adding all the V_i-V_j edges for all those pairs V_i, V_j of clusters with $V_iV_j \in E(R_B)$. Since $2|R_B|/\alpha \ll m_B$, we could make W_B into a Hamilton cycle of G^c_B by 'winding around' each cycle from $\mathcal{C}_B$ a suitable number of times. We could then apply the Blow-up lemma to show that this Hamilton cycle corresponds to one in G^*_B. However, as indicated in Section 5.2.1, we will argue slightly differently as it is not clear how to incorporate all the exceptional vertices by the above approach.

Set $\varepsilon_A := \varepsilon/|R_B|$. Apply the Diregularity lemma with parameters ε_A^2, $d+8\varepsilon_A^2$ and M'_A to $G[A\cup V_0]$ to obtain a partition of the vertex set of $G[A\cup V_0]$ into $\ell \geq M'_A$ clusters $V'_1, \dots, V'_\ell$ and an exceptional set V'_0. Let $A' := V'_1\cup\dots\cup V'_\ell$, let R_A denote the reduced oriented graph obtained from Lemma 5.3 and let G^*_A be the pure oriented graph. Similarly as in (17), Lemma 5.3 implies that $\delta^*(R_A) \geq (3/2+\alpha/2)|R_A|$ and so, as before, we can apply Theorem 5.7 to find a 1-factor $\mathcal{C}_A$ of R_A. Then as before, Proposition 5.6 implies that by adding at most $4\varepsilon_A^2 n$ further vertices to the exceptional set V'_0 we may assume that each edge of R_A corresponds to an ε_A-regular pair of density d and that each edge in the union $\bigcup_{C\in\mathcal{C}_A} C \subseteq R_A$ of all the cycles from $\mathcal{C}_A$ corresponds to an (ε_A, d)-superregular pair. So we now have that

$$|V'_0| \leq \varepsilon_A n = \varepsilon n/|R_B|. \tag{18}$$

Similarly as before, Corollary 5.11 gives us a closed walk W_A in R_A which is balanced w.r.t. $\mathcal{C}_A$, meets every cluster at most $2|R_A|/\alpha$ times and traverses all the edges lying on a cycle from $\mathcal{C}_A$ at least once.

5.4.2. Incorporating V'_0 into the walk W_B. Recall that the balanced closed walk W_B in R_B corresponds to a Hamilton cycle in G^c_B. Our next aim is to extend this walk to one which corresponds to a Hamilton cycle which also contains the vertices in V'_0. (The Blow-up lemma will imply that the latter Hamilton cycle corresponds to one in $G[B\cup V'_0]$.) We do this by extending W_B into a walk on a suitably defined digraph $R^*_B \supseteq R_B$ with vertex set $V(R_B)\cup V'_0$ in such a way that the new walk is balanced w.r.t. $\mathcal{C}_B$. R^*_B is obtained from the union of R_B and the set V'_0 by adding an edge vV_i between a vertex $v \in V'_0$ and a cluster $V_i \in V(R_B)$ whenever $\left|N^+_G(v)\cap V_i\right| > \alpha m_B/10$ and adding the edge V_iv whenever $\left|N^-_G(v)\cap V_i\right| > \alpha m_B/10$. Thus

$$|N^+_G(v)\cap B| \leq |N^+_{R^*_B}(v)|m_B + |R_B|\alpha m_B/10.$$

Hence

$$\begin{aligned}
|N^+_{R^*_B}(v)| &\geq |N^+_G(v)\cap B|/m_B - \alpha|R_B|/10 \geq |N^+_G(v)\cap B||R_B|/|B| - \alpha|R_B|/10\\
&\geq (|N^+_{G-A}(v)| - |V_0|)|R_B|/|G-A| - \alpha|R_B|/10\\
&\geq (\delta^+(G)/n - \alpha/2)|R_B| \geq \alpha|R_B|/2. \qquad (19)
\end{aligned}$$

(The penultimate inequality follows from the choice of A and the final one from Fact 5.9.) Similarly

$$|N^-_{R^*_B}(v)| \geq \alpha|R_B|/2.$$

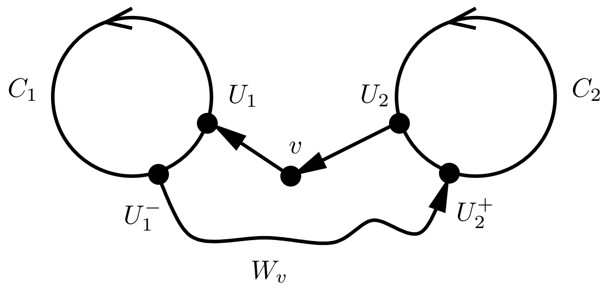


FIGURE 2. Incorporating the exceptional vertex v.

Given a vertex $v \in V_0'$ pick $U_1 \in N^+_{R_B^*}(v)$, $U_2 \in N^-_{R_B^*}(v) \setminus \{U_1\}$. Let C_1 and C_2 denote the cycles from $\mathcal{C}_B$ containing U_1 and U_2 respectively. Let U_1^- be the predecessor of U_1 on C_1, and U_2^+ be the successor of U_2 on C_2. (19) implies that we can ensure $U_1^- \neq U_2^+$. (However, we may have $C_1 = C_2$.) Corollary 5.10 gives us a shifted walk W_v from U_1^- to U_2^+ traversing at most $4/\alpha$ cycles of $\mathcal{C}_B$. To incorporate v into the walk W_B, recall that W_B traverses all those edges of R_B which lie on cycles from $\mathcal{C}_B$ at least once. Replace one of the occurences of $U_1^- U_1$ on W_B with the walk

$$W_v' := U_1^- W_v U_2^+ C_2 U_2 v U_1 C_1 U_1,$$

i.e. the walk that goes from U_1^- to U_2^+ along the shifted walk W_v, it then winds once around C_2 but stops in U_2, then it goes to v and further to U_1, and finally it winds around C_1. The walk obtained from W_B by including v in this way is still balanced w.r.t. $\mathcal{C}_B$, i.e. each vertex in R_B is visited the same number of times as every other vertex lying on the same cycle from $\mathcal{C}_B$. We add the extra loop around C_1 because when applying the Blow-up lemma we will need the vertices in V_0' to be at a distance of at least 4 from each other. Using this loop, this can be ensured as follows. After we have incorporated v into W_B we 'ban' all the 6 edges of (the new walk) W_B whose endvertices both have distance at most 3 from v. The extra loop ensures that every edge in each cycle from $\mathcal{C}$ has at least one occurence in W_B which is not banned. (Note that we do not have to add an extra loop around C_2 since if $C_2 \neq C_1$ then all the banned edges of C_2 lie on W_v' but each edge of C_2 also occurs on the original walk W_B.) Thus when incorporating the next exceptional vertex we can always pick an occurence of an edge which is not banned to be replaced by a longer walk. (When incorporating v we picked $U_1^- U_1$.) Repeating this argument, we can incorporate all the exceptional vertices in V_0' into W_B in such a way that all the vertices of V_0' have distance at least 4 on the new walk W_B.

Recall that G_B^c denotes the oriented graph obtained from the pure oriented graph G_B^* by adding all the V_i-V_j edges for all those pairs V_i, V_j of clusters with $V_i V_j \in E(R_B)$. Let $G^c_{B \cup V_0'}$ denote the graph obtained from G_B^c by adding all the V_0'-B edges of G as well as all the B-V_0' edges of G. Moreover, recall that the vertices in V_0' have distance at least 4 from each other on W_B and $|V_0'| \le \varepsilon n / |R_B| \ll \alpha m_B / 20$ by (18). As already observed at the beginning of Section 5.3, altogether this shows that by winding around each cycle from $\mathcal{C}_B$, one can obtain a Hamilton cycle $C^c_{B \cup V_0'}$ of $G^c_{B \cup V_0'}$ from the walk W_B, provided that W_B visits any cluster $V_i \in R_B$ at most m_B times.

To see that the latter condition holds, recall that before we incorporated the exceptional vertices in V_0' into W_B, each cluster was visited at most $2|R_B|/\alpha$ times. When incorporating an exceptional vertex we replaced an edge of W_B by a walk whose interior visits every cluster at most $4/\alpha + 2 \le 5/\alpha$ times. Thus the final walk W_B visits each cluster $V_i \in R_B$ at most

$$2|R_B|/\alpha + 5|V_0'|/\alpha \overset{(18)}{\le} 6\varepsilon n/(\alpha|R_B|) \le \sqrt{\varepsilon} m_B \tag{20}$$

times. Hence we have the desired Hamilton cycle $C^c_{B\cup V_0'}$ of $G^c_{B\cup V_0'}$. Note that (20) implies that we can choose $C^c_{B\cup V_0'}$ in such a way that for each cycle $C \in \mathcal{C}_B$ there is subpath P_C of $C^c_{B\cup V_0'}$ which winds around C at least

$$(1-\sqrt{\varepsilon})m_B \tag{21}$$

times in succession.

5.4.3. Applying the Blow-up lemma to find a Hamilton cycle in $G[B\cup V_0']$. Our next aim is to use the Blow-up lemma to show that $C^c_{B\cup V_0'}$ corresponds to a Hamilton cycle in $G[B\cup V_0']$. Recall that $k = |R_B|$ and that for each exceptional vertex $v \in V_0'$ the outneighbour U_1 of v on W_B is distinct from its inneighbour U_2 on W_B. We will apply the Blow-up lemma with H being the underlying graph of $C^c_{B\cup V_0'}$ and G^* being the graph obtained from the underlying graph of G^*_B by adding all the vertices $v \in V_0'$ and joining each such v to all the vertices in $N^+_G(v) \cap U_1$ as well as to all the vertices in $N^-_G(v) \cap U_2$. Recall that after applying the Diregularity lemma to obtain the clusters $V_1, \dots, V_k$ we used Proposition 5.6 to ensure that each edge of R_B corresponds to an ε-regular pair of density d (in the underlying graph of G^*_B and thus also in G^*) and that each edge of the union $\bigcup_{C\in\mathcal{C}_B} C \subseteq R_B$ of all the cycles from $\mathcal{C}_B$ corresponds to an (ε, d)-superregular pair.

V_0' will play the role of V_0 in the Blow-up lemma and we take $L_0, L_1, \dots, L_k$ to be the partition of H induced by $V_0', V_1, \dots, V_k$. $\phi : L_0 \to V_0'$ will be the obvious bijection (i.e. the identity). To define the set $I \subseteq V(H)$ of vertices of distance at least 4 from each other which is used in the Blow-up lemma, let P_C' be the subpath of H corresponding to P_C (for all $C \in \mathcal{C}_B$). For each $i = 1, \dots, k$, let $C_i \in \mathcal{C}_B$ denote the cycle containing V_i and let $J_i \subseteq L_i$ consist of all those vertices in $L_i \cap V(P'_{C_i})$ which have distance at least 4 from the endvertices of P'_{C_i}. Thus in the graph H each vertex $u \in J_i$ has one of its neighbours in the set L_i^- corresponding to the predecessor of V_i on C_i and its other neighbour in the set L_i^+ corresponding to the successor of V_i on C_i. Moreover, all the vertices in J_i have distance at least 4 from all the vertices in L_0 and (21) implies that $|J_i| \ge 9m_B/10$. It is easy to see that one can greedily choose a set $I_i \subseteq J_i$ of size $m_B/10$ such that the vertices in $\bigcup_{i=1}^k I_i$ have distance at least 4 from each other. We take $I := L_0 \cup \bigcup_{i=1}^k I_i$.

Let us now check conditions (C1)–(C9). (C1) holds with $K_1 := 1$ since $|L_0| = |V_0'| \le \varepsilon_A n = \varepsilon n/k \le d|H|$. (C2) holds by definition of I. (C3) holds since H is a Hamilton cycle in $G^c_{B\cup V_0'}$ (c.f. the definition of the graph $G^c_{B\cup V_0'}$). This also implies that for every edge $xy \in H$ with $x \in L_i, y \in L_j$ $(i, j \ge 1)$ we must have that $V_iV_j \in E(R_B)$. Thus (C6) holds as every edge of R_B corresponds to an ε-regular pair of clusters having density d. (C4) holds with $K_2 := 1$ because

$$|N_H(L_0) \cap L_i| \leq 2\,|L_0| = 2\,|V_0'| \overset{(18)}{\leq} 2\varepsilon n/|R_B| \leq 5\varepsilon m_B \leq dm_B.$$

For (C5) we need to find a set $D \subseteq I$ of buffer vertices. Pick any set $D_i \subseteq I_i$ with $|D_i| = \delta' m_B$ and let $D := \bigcup_{i=1}^{k} D_i$. Since $I_i \subseteq J_i$ we have that $|N_H(D) \cap L_j| = 2\delta' m_B$ for all $j = 1, \dots, k$. Hence

$$||N_H(D) \cap L_i| - |N_H(D) \cap L_j|| = 0$$

for all $1 \leq i < j \leq k$ and so (C5) holds. (C7) holds with $c := \alpha/10$ by our choice $U_1 \in N^+_{R_B^*}(v)$ and $U_2 \in N^-_{R_B^*}(v)$ of the neighbours of each vertex $v \in V_0'$ in the walk W_B (c.f. the definition of the graph R_B^*).

(C8) and (C9) are now the only conditions we need to check. Given a set $E_i \subseteq V_i$ of size at most $\varepsilon' m_B$, we wish to find $F_i \subseteq (L_i \cap (I \setminus D)) = I_i \setminus D$ and a bijection $\phi_i : E_i \to F_i$ such that every $v \in E_i$ has a large number of neighbours in every cluster V_j for which L_j contains a neighbour of $\phi_i(v)$. Pick any set $F_i \subseteq I_i \setminus D$ of size $|E_i|$. (This can be done since $|D \cap I_i| = \delta' m_B$ and so $|I_i \setminus D| \geq m_B/10 - \delta' m_B \gg \varepsilon' m_B$.) Let $\phi_i : E_i \to F_i$ be an arbitrary bijection. To see that (C8) holds with these choices, consider any vertex $v \in E_i \subseteq V_i$ and let j be such that L_j contains a neighbour of $\phi_i(v)$ in H. Since $\phi_i(v) \in F_i \subseteq I_i \subseteq J_i$, this means that V_j must be a neighbour of V_i on the cycle $C_i \in \mathcal{C}_B$ containing V_i. But this implies that $|N_{G^*}(v) \cap V_j| \geq (d - \varepsilon)m_B$ since each edge of the union $\bigcup_{C \in \mathcal{C}_B} C \subseteq R_B$ of all the cycles from $\mathcal{C}_B$ corresponds to an (ε, d)-superregular pair in G^*.

Finally, writing $F := \bigcup_{i=1}^{k} F_i$ we have

$$|N_H(F) \cap L_i| \leq 2\varepsilon' m_B$$

(since $F_j \subseteq J_j$ for each $j = 1, \dots, k$) and so (C9) is satisfied with $K_3 := 2$. Hence (C1)–(C9) hold and so we can apply the Blow-up lemma to obtain a Hamilton cycle in G^* such that the image of L_i is V_i for all $i = 1, \dots, k$ and the image of each $x \in L_0$ is $\phi(x) \in V_0$. (Recall that G^* was obtained from the underlying graph of G_B^* by adding all the vertices $v \in V_0'$ and joining each such v to all the vertices in $N_G^+(v) \cap U_1$ as well as to all the vertices in $N_G^-(v) \cap U_2$, where U_1 and U_2 are the neighbours of v on the walk W_B.) Using the fact that H was obtained from the (directed) Hamilton cycle $C^c_{B \cup V_0'}$ and since $U_1 \neq U_2$ for each $v \in V_0'$, it is easy to see that our Hamilton cycle in G^* corresponds to a (directed) Hamilton cycle C_B in $G[B \cup V_0']$.

5.4.4. Finding a Hamilton cycle in G. The last step of the proof is to find a Hamilton cycle in $G[A']$ which can be connected with C_B into a Hamilton cycle of G. Pick an arbitrary edge $v_1 v_2$ on C_B and add an extra vertex v^* to $G[A']$ with outneighbourhood $N_G^+(v_1) \cap A'$ and inneighbourhood $N_G^-(v_2) \cap A'$. A Hamilton cycle C_A in the digraph thus obtained from $G[A']$ can be extended to a Hamilton cycle of G by replacing v^* with $v_2 C_B v_1$. To find such a Hamilton cycle C_A, we can argue as before. This time, there is only one exceptional vertex, namely v^*, which we incorporate into the walk W_A. Note that by our choice of A and B the analogue of (19) is satisfied and so this can be done as before. We then use the Blow-up lemma to obtain the desired Hamilton cycle C_A corresponding to this walk.

5.5. Ore Type Condition

The proof of the Ore type result is essentially the same as that of the Dirac type result. The only place in the proof of the Diract type result where we used the full minimum semi-degree condition was in the proof of the expansion property, so we give a version of that lemma with Ore type conditions here. In the rest of the proof a weaker minimum degree condition suffices. This is guaranteed by the following proposition.

PROPOSITION 5.12. *Let G be an oriented graph such that for all distinct $x, y \in V(G)$ with $xy \notin E(G)$ we have $d^+(x) + d^-(y) \geq 3n/4$. Then $\delta^0(G) \geq n/8$.*

Proof. Suppose not. When without loss of generality assume that $\delta^+(G) \leq \delta^-(G)$. Indeed, if not then replace every edge $xy \in E(G)$ with yx. Let $u \in V(G)$ have $d^+(u) = \delta^0(G)$. Define $S = \{s \in V(G)\backslash\{u\} : us \notin E(G)\}$. Then $|S| \geq n - n/8 = 7n/8$ and moreover for all $s \in S$ we have $d^-(s) \geq 3n/4 - n/8 = 5n/8$. Hence

$$n^2/2 > e(G) \geq (5n/8)|S| \geq 35n^2/64,$$

which is a contradiction. □

The only further difference is the need for an analogue of Lemma 5.3, which guarantees that we can find an reduced oriented graph still approximately satisfying the original Ore type condition.

LEMMA 5.13. *For every $\varepsilon \in (0, 1)$ there exist integers $M' = M'(\varepsilon)$ and $n_0 = n_0(\varepsilon)$ such that the following holds. Let $d \in [0, 1]$, let G be an oriented graph of order at least n_0 and let R' be the reduced digraph obtained by applying the Diregularity lemma to G with parameters ε, d and M'. Suppose further that $\varepsilon + d \ll \alpha < 1$ and for all vertices $x, y \in V(G)$ with $xy \notin E(G)$, $d^+(x) + d^-(y) \geq \alpha |G|$. Then R' has a spanning oriented subgraph R with*

$$d^+(i) + d^-(j) \geq (\alpha - (5\varepsilon + 2d)) |G| \tag{22}$$

for all $ij \notin E(R)$.

Proof. Let us first show that all clusters V_i, V_j with $ij \notin E(R')$ satisfy

$$(|N^+_{R'}(i)| + |N^-_{R'}(j)|)/|R'| \geq \alpha - (4\varepsilon + 2d). \tag{23}$$

To see this observe that as $ij \notin E(R')$ these do not form an ε-regular pair of density at least d in the pure graph. Thus we can find vertices $x \in V_i$, $y \in V_j$ with $xy \notin E(G)$ then by our hypothesis and the Diregularity lemma $d^+_{G'}(x) + d^-_{G'}(y) \geq (\alpha - 2(\varepsilon + d)) |G|$. On the other hand,

$$\begin{aligned} |N^+_{G'}(x)| + |N^-_{G'}(y)| &\leq (|N^+_{R'}(i)| + |N^-_{R'}(j)|)m + 2|V_0| \\ &\leq (|N^+_{R'}(i)| + |N^-_{R'}(j)|)|G|/|R'| + 2\varepsilon|G|. \end{aligned}$$

Altogether this proves (23).

Let R be the spanning oriented subgraph obtained from R' by deleting edges randomly as follows. For every unordered pair V_i, V_j of clusters we delete the edge ij (if it exists) with probability

$$\frac{e_{G'}(V_j, V_i)}{e_{G'}(V_i, V_j) + e_{G'}(V_j, V_i)}. \tag{24}$$

Otherwise we delete ji (if it exists). We interpret (24) as 0 if $ij, ji \notin E(R')$. So if R' contains at most one of the edges ij, ji then we do nothing. We do this for all unordered pairs of clusters independently and let X_i be the random variable which counts the number of outedges of the vertex $i \in R$ and let Y_i count the number of inedges. Suppose $ij, ji \in E(R')$. Then

$$\begin{aligned}
\mathbb{E}(X_i + Y_j) &= \sum_{k \neq i} \frac{e_{G'}(V_i, V_k)}{e_{G'}(V_i, V_k) + e_{G'}(V_k, V_i)} + \sum_{k \neq j} \frac{e_{G'}(V_k, V_j)}{e_{G'}(V_j, V_k) + e_{G'}(V_k, V_j)} \\
&\geq \sum_{k \neq i} \frac{e_{G'}(V_i, V_j)}{|V_i|\,|V_j|} + \sum_{k \neq j} \frac{e_{G'}(V_k, V_j)}{|V_j|\,|V_k|} \\
&\geq \frac{|R'|}{|G|\,|V_i|} \left[\sum_{x \in V_i} (d^+_{G'}(x) - |V_0|) + \sum_{y \in V_j} (d^-_{G'}(y) - |V_0|) \right].
\end{aligned}$$

Now observe that as $ji \in E(R')$, the bipartite oriented graph $(V_j, V_i)_{G'}$ is ε-regular with density at least d. Thus it contains a matching M of size at least $(1-\varepsilon)|V_i|$. Since G is oriented, this gives us $|M|$ ordered pairs of vertices $(x, y) \in (V_i, V_j)$ with $xy \notin E(G)$ and hence $d^+_{G'}(x) + d^-_{G'}(y) \geq (\alpha - 2(\varepsilon + d))\,|G'|$. Thus

$$\begin{aligned}
\mathbb{E}(X_i + Y_j) &\geq \frac{|R'|}{|G|\,|V_i|} \left[(1-\varepsilon)(\alpha - 2(\varepsilon + d))\,|R|\,|V_i| - 2|V_0|\,|V_i| \right] \\
&\geq (\alpha - (5\varepsilon + 2d)\,|R| \geq \varepsilon |R|.
\end{aligned}$$

As in Lemma 5.3, a straightforward application of a Chernoff type bound shows that there exists $R \subseteq R'$ with this property, and combining this with (23) we see that the result holds. $\square$

The following expansion lemma is the analogue of Lemma 5.8 in the proof of the Ore type result.

LEMMA 5.14. *Let R^* be an oriented graph on N vertices with $d^+(u) + d^-(v) \geq 3N/4 + \alpha N$ for some $\alpha > 0$ and for vertices $u, v \in V(R^*)$ with $uv \notin E(R^*)$. If $X \subseteq V(R^*)$ is nonempty and $|X| \leq (1-\alpha)N$ then $|N^+(X)| \geq |X| + \alpha N/2$.*

Proof. Suppose the assertion is false, i.e. there exists nonempty $X \subseteq V(R^*)$ with $|X| \leq (1-\alpha)N$ and

$$\left|N^+(X)\right| < |X| + \alpha N/2. \tag{25}$$

We consider the following partition of $V(R^*)$:

$A := X \cap N^+(X)$, $B := N^+(X) \setminus X$, $C := V(R^*) \setminus (X \cup N^+(X))$, $D := X \setminus N^+(X)$.

(25) gives us

$$|D| + \alpha N/2 > |B|. \tag{26}$$

Suppose $|D| < \alpha N/10$. Then by (26) $|B| < 3\alpha N/4$ and so $A, C \neq \emptyset$. By an averaging argument there exists $a \in A$ with $|N^+(a) \cap A| \leq |A|/2$ and similarly there exists $c \in C$ with $|N^-(c) \cap C| \leq |C|/2$. By construction $ac \notin E(R^*)$ so $d^+(a) + d^-(c) \geq 3N/4 + \alpha N$ and so

$$3N/4 + \alpha N \leq d^+(a) + d^-(c) \leq (|A|/2 + |B|) + (|C|/2 + |B|).$$

Hence we have

$$N \geq |A| + |C| \geq 3N/2 + 2\alpha N - 4|B| \geq 3N/2 - \alpha N,$$

which is a contradiction. Thus for the rest of the proof we shall assume that $|D| > \alpha N/10$. Note that this is much stronger than we need, $|D| \geq 2$ would suffice.

For simplicity, we write $\delta := \delta(R^*)$, $\delta^+ := \delta^+(R^*)$ and $\delta^- := \delta^-(R^*)$. We write δ^+_D for the minimum outdegree in R^* of a vertex in D, that is, $\delta^+_D := \min_{d \in D}(d^+_{R^*}(d))$. We define δ^-_D, δ^+_A, etc. similarly.

Note that $R^*[D]$ contains no edges, and so for all $x \neq y \in D$ we have $d(x) + d(y) \geq 3N/2 + 2\alpha n$. Hence there is at most one vertex in D with degree less than $3N/4 + \alpha n$. In the case that such a vertex d_{bad} exists, define $D' := D \backslash \{d_{\text{bad}}\}$ and observe that $\delta_{D'} \geq 3N/4 + \alpha N$ and $D' \neq \emptyset$. If no such vertex exists then let $D' := D$ and again $\delta_{D'} \geq 3N/4 + \alpha N$.

Suppose $A \neq \emptyset$. Then by an averaging argument there exists $x \in A$ with $|N^+(x) \cap A| < |A|/2$. Hence $\delta^+_A \leq |N^+(x)| < |B| + |A|/2$. Combining this with (26) we get

$$|A| + |B| + |D| \geq 2\delta^+_A - \alpha N/2. \tag{27}$$

If $A = \emptyset$ then $N^+(X) = B$ and so (26) implies $|D| + \alpha N/2 \geq |B| \geq \delta^+_{D'}$. Thus we get

$$|A| + |B| + |D| \geq 2\delta^+_{D'} - \alpha N/2. \tag{28}$$

Similarly, if $C \neq \emptyset$ then considering the inneighbourhood of a suitable vertex $x \in C$ gives

$$|B| + |C| + |D| \geq 2\delta^-_C - \alpha N/2. \tag{29}$$

If $C = \emptyset$ then $N^-(D') \subseteq B$ and thus $|B| \geq \delta^-_{D'}$. Together with (26) this shows that

$$|B| + |C| + |D| \geq 2\delta^-_{D'} - \alpha N/2. \tag{30}$$

For any $x \in D'$ we have $N(x) \cap D = \emptyset$ and hence

$$|A| + |B| + |C| \geq |N(x)| \geq \delta_{D'}. \tag{31}$$

We now combine the above equations to generate a contradiction. We have four separate, but very similar, cases.

Case 1: $A \neq \emptyset, C \neq \emptyset$. First note that $e(A, C) = \emptyset$ so $\delta^+_A + \delta^-_C \geq 3N/4 + \alpha N$. Combining equations (27), (29) and (31) gives

$$\begin{aligned} 3|A| + 4|B| + 3|C| + 2|D| &\geq 2\delta^+_A + 2\delta^-_C + 2\delta_{D'} - \alpha N \\ &\geq 4(3N/4 + \alpha N) - \alpha N. \end{aligned}$$

Case 2: $A \neq \emptyset$, $C = \emptyset$. Since $e(A, D') = \emptyset$ we have $\delta^+_A + \delta^-_{D'} \geq 3N/4 + \alpha N$. Equations (27), (30) and (31) give

$$\begin{aligned} 3|A| + 4|B| + 3|C| + 2|D| &\geq 2\delta^+_A + 2\delta^-_{D'} + 2\delta_{D'} - \alpha N \\ &\geq 4(3N/4 + \alpha N) - \alpha N. \end{aligned}$$

Case 3: $A = \emptyset$, $C \neq \emptyset$. Since $e(D', C) = \emptyset$ we have $\delta^+_{D'} + \delta^-_C \geq 3N/4 + \alpha N$. Equations (28), (29) and (31) give

$$\begin{aligned} 3|A| + 4|B| + 3|C| + 2|D| &\geq 2\delta^+_{D'} + 2\delta^-_C + 2\delta_{D'} - \alpha N \\ &\geq 4(3N/4 + \alpha N) - \alpha N. \end{aligned}$$

Case 4: $A = \emptyset$, $C = \emptyset$. By the definition of D' and our earlier remarks we have $\delta^+_{D'} + \delta^-_{D'} \geq 3N/4 + \alpha N$. Equations (28), (30) and (31) give

$$\begin{aligned} 3|A| + 4|B| + 3|C| + 2|D| &\geq 2\delta^+_{D'} + 2\delta^-_{D'} + 2\delta_{D'} - \alpha N \\ &\geq 4(3N/4 + \alpha N) - \alpha N. \end{aligned}$$

Finally, substituting (26) gives

$$3N + \alpha N/2 \geq 3N + 3\alpha N,$$

which is a contradiction. $\square$

Chapter 6

Proof of theorem 1.6

With the results of Section 4.3 in mind, we are now in a position to prove Theorem 1.6. The proof that this result holds for 'long' cycles uses somewhat similar methods to those in [**36**], and we will use some results from that paper. Using the 'stability method' we will distinguish between a non-extremal case where our oriented graph has some form of expansion property, and an extremal case where the oriented graph is shown to be similar to that in Figure 5.1.

We have already proved the result for $4 \leq \ell \leq n/10^{10}$ in Theorem 4.18. Thus we can assume that $n/10^{10} \leq \ell < n$.

We will need the following slight extension of Lemma 5.3, due to Keevash, Kühn and Osthus [**36**].

Lemma 6.1. *For every $\varepsilon \in (0,1)$ and there exist numbers $M' = M'(\varepsilon)$ and $n_0 = n_0(\varepsilon)$ such that the following holds. Let $d \in [0,1]$ with $\varepsilon \leq d/2$, let G be an oriented graph of order $n \geq n_0$ and let R' be the reduced digraph with parameters (ε, d) obtained by applying the Diregularity Lemma to G with M' as the lower bound on the number of clusters. Then R' has a spanning oriented subgraph R such that*

(a) $\delta^+(R) \geq (\delta^+(G)/|G| - (3\varepsilon + d))|R|$,
(b) $\delta^-(R) \geq (\delta^-(G)/|G| - (3\varepsilon + d))|R|$,
(c) *for all disjoint sets $S, T \subset V(R)$ with $e_G(S^*, T^*) \geq 3dn^2$ we have $e_R(S,T) > d|R|^2$, where $S^* := \bigcup_{i \in S} V_i$ and $T^* := \bigcup_{i \in T} V_i$.*
(d) *for every set $S \subset V(R)$ with $e_G(S^*) \geq 3dn^2$ we have $e_R(S) > d|R|^2$, where $S^* := \bigcup_{i \in S} V_i$.*

Define a hierarchy of constants so that

$$1/n_0 \ll \varepsilon \ll d \ll c \ll \eta \ll 1.$$

Let G be an oriented graph on $n \geq n_0$ vertices with minimum semi-degree $\delta^0(G) \geq \lceil (3n-4)/8 \rceil$ and let $u \in V(G)$. Suppose that G contains no cycle of length ℓ containing u. Apply the Diregularity Lemma (Theorem 5.2) and Lemma 6.1 to G with parameters $(\varepsilon^2/3,\ d)$. This gives us a partition of $V(G)$ into $V_0, V_1, \ldots, V_k$ with $m := |V_1| = \ldots = |V_k|$ and a reduced oriented graph R. Lemma 6.1 gives us that

$$\delta^0(R) > (3/8 - 1/(2n) - d - \varepsilon^2)k > (3/8 - 2d)k. \tag{32}$$

Case 1. *$|N_R^+(S)| \geq |S| + 2ck$ for every $S \subset [k]$ with $k/3 < |S| < 2k/3$.*

In this case we use probabilistic methods to find a subdigraph G' of G with ℓ vertices and a new reduced oriented graph which still satisfies the conditions of Case 1, possibly with modified constants. Also, we can ensure that $u \in V(G')$. We can then use the following result from [**36**], which says that all such graphs contain a Hamilton cycle.

Lemma 6.2. *Let M', n_0 be positive numbers and let $\varepsilon, d, \eta, \nu, \tau$ be positive constants such that $1/n_0 \ll 1/M' \ll \varepsilon \ll d \ll \nu \leq \tau \ll 1$. 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 such that $|R'| \geq M'$. Suppose that there exists a spanning oriented subgraph R of R' with $\delta^0(G) \geq \eta|R|$ and such that $|N_R^+(S)| \geq |S| + \nu|S|$ for all sets $S \subseteq V(R)$ with $|S| < (1-\tau)|R|$. Then G contains a Hamilton cycle.*

The argument we use to find an appropriate subdigraph G' is similar to that in [**36**], and uses standard probabilistic techniques. Recall that there are k (non-exceptional) clusters, each with size m.

Claim 1.1. *Let m' satisfy $10^{-11}n/k < m' < m$ and $p := m'/m$. Then there exists a partition of $V(G)\setminus V_0$ into sets A and B which has the following properties:*

(a) $|A_i| = m'$, *where we write* $A_i := V_i \cap A$ *for every* $i \in [k]$;
(b) $|N_G^+(v) \cap A_i| = p|N_G^+(v) \cap V_i| \pm n^{2/3}$ *for every vertex* $v \in V(G)$; *and similarly for* $N_G^-(v)$;
(c) R *is the oriented reduced graph with parameters* $(\varepsilon^2/10^{11}, 3d/4)$ *corresponding to the partition* $A_1, \dots, A_k$ *of the vertex set of* $G[A]$;
(d) $\delta^0(G[A]) \geq (3/8 - \varepsilon)|A|$.

Proof. For each cluster V_i define a partition into A_i and B_i as follows. Let $\eta := n^{2/3}/(4|V_i|)$ and put $x \in V_i$ in A_i with probability $p + \eta$, independently of all other vertices. Then standard Chernoff type bounds give that the probability that $p|V_i| < |A_i| < p|V_i| + n^{2/3}/2$ does not occur is exponentially small in $|V_i|$. Further, they also give that the probability that any vertex $v \in A_i$ has outneighbourhood varying from $p|N_G^+(v) \cap V_i|$ by more than $n^{2/3}/2$ is exponentially small. Thus for sufficiently large n a partition exists satisfying both these conditions, and we can discard up to $n^{2/3}/2$ vertices from each A_i to obtain a partition satisfying (a) and (b).

To see (c) note that the definition of regularity implies that the pair (A_i, A_j) consisting of all the A_i-A_j edges in the pure oriented graph G^* is $\varepsilon^2/10^{11}$-regular and has density at least $3d/4$ whenever $ij \in E(R)$. On the other hand, (A_i, A_j) is empty whenever $ij \notin E(R)$ since $(V_i, V_j) \supset (A_i, A_j)$ is empty in this case. Property (d) follows immediately from (b). □

If $\ell \geq n - |V_0|$ then form G' by discarding $n - \ell$ arbitrary vertices from $V(G) \setminus \{u\}$. Otherwise apply the previous claim to G with $m' := \lfloor \ell/k \rfloor - 1$. Let $G' := G[A \cup V_0']$, where $V_0' \subseteq V(G) \setminus A$ is an arbitrary set of vertices containing u (if $u \notin A$) of size $\ell - |A|$. Then G' has exactly ℓ vertices and satisfies the conditions of Lemma 6.2 with $\tau = \eta = 1/3$ and $\nu = 2c$. Apply that result to obtain a Hamilton cycle in G' and thus a cycle of length ℓ through u in G.

Case 2. *There is a set $S \subset [k]$ with $k/3 < |S| < 2k/3$ and $|N_R^+(S)| < |S| + 2ck$.*

In this case we exploit the minimum semi-degree condition to demonstrate that G has roughly the same structure as the extremal graph. The proof proceeds in three steps.

(i) Show that the G has roughly the same structure as the extremal graph.
(ii) Show that if the cluster sizes and vertices satisfy certain conditions then using the Blow-up Lemma (Lemma 5.4) we have the desired cycle (Claim 2.3).
(iii) Use (ii) to obtain further structural refinements, eventually showing that G either contains a Hamilton cycle or contradicts the minimum semi-degree condition.

The difference between the proof here and the proof of the exact Hamiltonicity result in [**36**] is primarily in Step (ii), Claim 2.3. We have similar conditions here, but the stronger conclusion that we get a cycle of any length, not just a Hamilton cycle. Their proofs of the results needed for (iii) in the Hamiltonicity case implicitly require only that the conditions of (ii) are not satisfied, and so the proof of Step (iii) for us is implicit in their paper. Hence we will not give their proofs for either Step (i) or (iii). Instead we give a complete proof of the result in Step (ii) and refer the reader to [**36**] for all remaining details.

Let

$$A_R := S \cap N^+_R(S),\ B_R := N^+_R(S) \backslash S,\ C_R := [k] \backslash (S \cup N^+_R(S)),\ D_R := S \backslash N^+_R(S).$$

These sets will have similar properties as the sets A, B, C and D in the extremal example. Let $A := \bigcup_{i \in A_R} V_i$ and define B, C, D similarly. The following notation will prove useful. Let $P(1) := A$, $P(2) := B$, $P(3) := C$ and $P(4) := D$. When we refer to $P(i+1)$ or $P(i-1)$ we will always mean modulo 4. Define $P(i \oplus 1)$ by $P(1 \oplus 1) := P(1)$, $P(2 \oplus 1) := P(4)$, $P(3 \oplus 1) := P(3)$ and $P(4 \oplus 1) := P(2)$. This operation should be viewed with reference to the extremal graph as being the 'other' out-class of $P(i)$, and has the obvious inverse $P(1 \ominus 1) := P(1)$, $P(2 \ominus 1) := P(4)$, $P(3 \ominus 1) := P(3)$ and $P(4 \ominus 1) := P(2)$. Since we will show that G has a somewhat similar structure to the extremal graph it will be useful to define the following graph on $V(G)$. Let $F[(P(i), P(i+1)]$ contain all edges from $P(i)$ to $P(i+1)$, let $F[A]$ and $F[C]$ be tournaments which are as regular as possible. Finally let $F[B, D]$ be a bipartite tournament which is as regular as possible. We will show that G roughly looks like F, and hence contains a cycle of length ℓ. From now on we will not calculate explicit constants multiplying c, and just write $O(c)$. The constants implicit in the $O(*)$ notation will always be absolute.

We call a vertex $x \in P(i)$ *cyclic* if it has almost the same number of neighbours in $P(i-1)$ and $P(i+1)$ as a vertex in the corresponding vertex class in F. More precisely, call a vertex $x \in P(i)$ *cyclic* if $|N^+_G(x) \cap P(i+1)| \geq (1 - O(\sqrt{c})|P(i+1)|$ and $|N^-_G(x) \cap P(i-1)| \geq (1 - O(\sqrt{c})|P(i-1)|$, counting modulo 4. A vertex is *acceptable* if it has a significant outneighbourhood in one of its two 'out-classes' and one of its two 'in-classes', where these are understood with reference to F. More precisely, $x \in P(i)$ is *acceptable* if both the following hold.

- $|N^+_G(x) \cap P(i+1)| \geq (1/100 - O(\sqrt{c}))n$ or $|N^+_G(x) \cap P(i \oplus 1)| \geq (1/100 - O(\sqrt{c}))n$,

- $|N_G^-(x) \cap P(i-1)| \geq (1/100 - O(\sqrt{c}))n$ or $|N_G^-(x) \cap P(i \ominus 1)| \geq (1/100 - O(\sqrt{c}))n$.

An edge from $P(i)$ to $P(j)$ in G is acceptable if $P(j) = P(i+1)$ or $P(j) = P(i \oplus 1)$.

The next claim combines several results from [**36**] and shows that these sets have roughly the same structure as in F.

Claim 2.2 (Keevash, Kühn and Osthus, [**36**]). *The following hold for all i.*

(a) $|P(i)| = (1/4 \pm O(c))n$,
(b) $e(P(i), P(i+1)) > (1 - O(c))n^2/16$,
(c) $e(P(i), P(i \oplus 1)) > (1/2 - O(c))n^2/16$.

Furthermore, by reassigning vertices that are not cyclic to A, B, C or D we can arrange that every vertex of G is acceptable. We can also arrange that there are no vertices that are not cyclic but would become so if they were reassigned.

Note that these properties of A, B, C and D are invariant under the relabelling $A \leftrightarrow C$, $B \leftrightarrow D$. Thus we may assume that $|B| \geq |D|$.

Given a path $P := v_1 \dots v_k$ in G with $v_1, v_k \in P(i)$ we say we *contract* P to refer to the following process, which yields a new digraph H. Remove $v_1, \dots, v_k$ from G and add an extra vertex v^* to $P(i)$ with outneighbourhood $N^+(v_k)$ and inneighbourhood $N^-(v_1)$. The 'moreover' part of the next claim is not in the statement of the corresponding claim in [**36**]. That we are not seeking a Hamilton cycle allows us this modified condition and a simpler proof than would otherwise be the case.

Claim 2.3. *If $|B| = |D|$ and every vertex is acceptable then G has an ℓ-cycle containing u. Moreover, the assertion also holds if we allow one non-acceptable vertex $x \in A \cup C$.*

Proof. The idea is as follows. First we contract suitable paths to leave us with a digraph G_1 containing only cyclic vertices. Then we find suitable paths to contract to give a digraph G_2 with $|A| = |B| = |C| = |D|$. We can then apply the Blow-up Lemma to the underlying graph to find a cycle in G_2 which 'winds around' A, B, C, D. By our choice of the vertices in this cycle and the definition of our contractions this will correspond to the desired cycle in G. We will say that a 4-partite graph with vertex classes $(P(1), P(2), P(3), P(4))$ has *type* (p_1, p_2, p_3, p_4) if $|P(i)| = p_i + q$ and $p_i \in \mathbb{N}$ for all i and some q. Our initial condition on the sizes means that G has type $(p_1, 0, p_3, 0)$. The *type sum* is $p_1 + p_2 + p_3 + p_4$.

Firstly, move the non-acceptable vertex x (if it exists) to a vertex class in which it is acceptable, and readjust the $O(c)$ notation if necessary. This gives us type $(p_1, 0 \leq p_2 \leq 1, p_3, 0)$, possibly with new values for the p_i. Let $v_1, \dots, v_t$ be vertices which are acceptable but not cyclic. Claim 2.2 (a) and (b) give us that $t = O(\sqrt{c})n$, so we can pick cyclic neighbours v_i^+ and v_i^- of each v_i such that the edges $v_i v_i^+$ and $v_i^- v_i$ are acceptable and all these vertices are distinct. We want to contract $v_i^- v_i v_i^+$ so that we form a new graph in which all vertices are cyclic. We need to ensure that after contracting we are still of type $(p_1, 0 \leq p_2 \leq 1, p_3, 0)$ (although possibly with different p_i to above) and $p_1, p_3 = O(\sqrt{c})n$. For each v_i find a path P_i' of length at most 3 starting at v_i^+, ending at some cyclic vertex in the same cluster as v_i^- and 'winding around the clusters,' i.e. following the order $P(i)$, $P(i+1)$ etc. If v_i^+ and v_i^- are in the same cluster then the path P_i' is the empty path. Let $P_i := v_i^- v_i v_i^+ P_i'$ and note that we can choose the P_i to be disjoint.

Contract the paths P_i to form a new digraph G_1. Note that G_1 is not necessarily oriented. Every vertex in G_1 is cyclic by construction, possibly with a new constant in the $O(\sqrt{c})$ notation in the definition of a cyclic vertex. G_1 also has type $(p_1, 0 \leq p_2 \leq 1, p_3, 0)$ and $p_1, p_3 = O(\sqrt{c})n$.

Now suppose that $|A| < |C|$ and let $s := |C| - |A| = p_3 - p_1$. Greedily find a path P_C in G_1 which follows the pattern $CCDAB$ s times and then ends in C. I.e. find an edge between 2 cyclic vertices in C, extend around the clusters back to C and repeat until we have a path from C to C with s (cyclic) vertices from A, B and D and $2s+1$ vertices from C. We can do this as Claim 2.2 (a) and (c) imply that almost all unordered pairs of vertices in C are connected by an edge and $s = O(\sqrt{c})n$. Let G_2 be the digraph obtained by contracting P_C. Then in G_2 has type $(p_1, 0 \leq p_2 \leq 1, p_1, 0)$. If $|A| > |C|$ we can achieve type $(p_1, 0 \leq p_2 \leq 1, p_1, 0)$ in a similar way by contracting a path P_A from A to A following the pattern $AABCD$. Note that since $s = O(\sqrt{c})n$, all vertices of G_2 are still cyclic. Now suppose that in G_2 we have $|D| > |A|$. Let $s := |D| - |A| = -p_1$. This time we find a path P_D from D to D following the pattern $DBCDABDABC$ which contains $s+1$ more vertices from D than it contains from A, and similarly for C. Note that contracting P_D does not change $|B| - |D|$. Contracting P_D gives us a digraph (which we still call G_2) with type $(0, 0 \leq p_2 \leq 1, 0, 0)$ and all of whose vertices are still cyclic. The last case to consider is when we have $|D| < |A|$. In this case we can equalize the sets by contracting two paths P_A and P_C of appropriate length as above.

We now find and contract a short path in G_2 to form a new oriented graph G_3 with $|G_3| - n + \ell \equiv 0 \pmod 4$. Let $p := n - \ell \pmod 4$. This is (congruent to) the number of vertices we do not want in the cycle we will find in G_3. We now contract paths to ensure that G_3 has type sum p, and thus $|G_3| - n + \ell \equiv 0 \pmod 4$. Suppose G_3 has type $(0,0,0,0)$. If $p = 0$ we are done. If $p = 1$ use one path P_C and one path P_D as above to obtain type $(1,0,0,0)$. If $p = 2$ then a path P_D gives us type $(1,0,1,0)$ and finally if $p = 3$ a path P_C gives us type $(1,1,0,1)$. Now suppose G_3 has type $(0,1,0,0)$. If $p = 1$ we are done already. If $p = 2$ contract one path P_D and one path P_C to get type $(1,1,0,0)$. If $p = 3$ a path P_D gives us type $(1,1,1,0)$. Finally if $p = 4$ two paths P_D and one path P_C gives type $(2,1,1,0)$.

At most $O(\sqrt{c})n$ vertices in G_3 correspond to paths in G. Call these vertices and u *special vertices*. We now contract the special vertices. Let S_1 consist of the special vertices in A. Find a path from A to A that 'winds around' the 4 clusters of the oriented graph G_3 $|S_1|$ times and contains all vertices in S_1. As $|S_1| \leq O(\sqrt{c})n$ we can find such a path easily with a greedy algorithm. Contract this path and repeat for B, C and D to reduce the number of special vertices to at most 4 without otherwise affecting the structure of G_3. Let S consist of these remaining special vertices.

Let G_3' be the underlying graph corresponding to the set of edges oriented from $P(i)$ to $P(i+1)$, for $1 \leq i \leq 4$. Since all vertices of G_3 are cyclic and we chose $c \ll \eta \ll 1$, each pair $(P(i), P(i+1))$ is $(\eta, 1)$-super-regular in G_3'. Furthermore, G_3' contains no multiple edges. Let F' be the 4-partite graph with vertex classes $P(i)$ where the 4 bipartite graphs induced by $(P(i), P(i+1))$ are all complete. Define $\ell' := |F'| - n + \ell$ and note that it satisfies $\ell' \equiv 0 \pmod 4$ and $\ell'/4 \leq |D|$. Thus 'winding around' the 4 clusters $\ell'/4$ times we can find a cycle of length ℓ' in F' including all the special vertices. Note that we need $\ell < n$ here, since the one non-acceptable vertex means that we cannot ensure that G_3 has type $(0,0,0,0)$. Remove each special vertex $v_j \in S$ from this cycle to split the cycle into a series of disjoint paths $P_1 := v_1^+ P_1' v_2^-$, $P_2 := v_2^+ P_2' v_3^-$ etc. For each $v_j \in S \cap P(i)$ and every i pick sets $C_j^+ \subset N^+(v_j) \cap P(i+1)$ and $C_j^- \subset N^-(v_j) \cap P(i-1)$ of size $10^{-8}|G_3|$. We now apply Lemma 5.4 with $k = 4$, $\Delta = 2$, $b = 10^{-8}$ and the C_j^+ and C_j^- as the sets C_x (for $x \in \{v_1^+, v_2^-, v_2^+, \dots, v_1^-\}$) to embed the paths $P_1, \dots, P_{|S|}$. This gives us disjoint paths in $G_3' - S$ with endpoints in the C_j^+ and C_j^- and the sum of whose lengths is $|G_3| - n + \ell - 2|S|$. The 'moreover' part of Lemma 5.4 implies that we can assume that these paths continually 'wind around' A, B, C, D. The condition on the endpoints of the paths ensures that we can add in the special vertices to obtain a cycle C in G_3 of length $|G_3| - n + \ell$. As every vertex outside C in G_3 corresponds to a single vertex in G, the cycle C_ℓ in G corresponding to C has length ℓ and contains u.

□

Since we are done if we satisfy the conditions of Claim 2.3, assume that $|B| > |D|$. The argument in [**36**] reaches a similar point to us here, and proceeds by showing that either G contains a Hamilton cycle, or is even more like the extremal graph. More precisely, they show that G either satisfies certain structural conditions, which we state below, or the conditions of Claim 2.3 are satisfied. They do this by moving vertices between clusters to obtain $|B| = |D|$ whilst ensuring that all vertices are acceptable. The situation can arise though that $|B| = |D| + 1$ and the only vertex class that it is possible to move vertices in B to without stopping them being acceptable is D. In this case we can shift an arbitrary vertex in B to $A \cup C$ to satisfy the conditions of Claim 2.3.

Claim 2.4. *For each of the following properties, there are fewer than $|B| - |D|$ vertices with that property or the conditions of Claim 2.3 are satisfied.*

- $x \in A$ *and* $|N^-(x) \cap C| \geq (1/100 - O(\sqrt{c}))n$.
- $x \in A$ *and* $|N^-(x) \cap B| \geq (1/100 - O(\sqrt{c}))n$.
- $x \in C$ *and* $|N^+(x) \cap A| \geq (1/100 - O(\sqrt{c}))n$.
- $x \in C$ *and* $|N^+(x) \cap B| \geq (1/100 - O(\sqrt{c}))n$.

We now define a new class of vertices. We say that a vertex is *good* if it is acceptable and satisfies one of the following.

- $x \in A$ and $|N^-(x) \cap C|$, $|N^-(x) \cap B| \leq (1/100 + O(\sqrt{c}))n$.
- $x \in B$ and $|N^+(x) \cap A|$, $|N^+(x) \cap B| \leq (1/100 + O(\sqrt{c}))n$ and $|N^-(x) \cap B|$, $|N^-(x) \cap C| \leq (1/100 + O(\sqrt{c}))n$.

- $x \in C$ and $|N^+(x) \cap A|$, $|N^+(x) \cap B| \leq (1/100 + O(\sqrt{c}))n$.
- $x \in D$.

Note that cyclic vertices are not necessarily good.

Claim 2.5. *By reassigning at most $O(\sqrt{c})n$ vertices we can arrange that every vertex is good or the conditions of Claim 2.3 are satisfied.*

Let M be a maximal matching in $e(B, A) \cup e(B) \cup e(C, A) \cup e(C, B)$.

Claim 2.6. *$e(M) = 0$ and $|B| - |D| = 1$ or the conditions of Claim 2.3 are satisfied.*

If the conditions of Claim 2.3 are satisfied we are done, so assume not. Since $e(M) = 0$ we now have $e(B \cup C, A) = 0$. Since $e(A) < |A|^2/2$ there exists a vertex $a \in A$ with $d^-(a) \leq (|A| - 1)/2 + |D|$. Furthermore, we also now have that $e(C, B) = 0$ and $e(B) = 0$, and so there exist vertices $c \in C$ and $b \in B$ with $d^+(c) \leq (|C| - 1)/2 + |D|$ and $d(b) \leq |A| + |C| + |D|$. Since $|D| = |B| - 1$ we see that

$$(3n-4)/2 \leq d^-(a) + d^+(c) + d(b) \leq \frac{3}{2}(|A| + |C| + 2|D|) - 1 = \frac{3}{2}(n-1) - 1.$$

This contradiction completes the proof. $\square$

Bibliography

[1] N. Alon and G. Gutin, Properly colored Hamilton cycles in edge colored complete graphs. *Random Structures and Algorithms* **11** (1997), 179–186.

[2] N. Alon and A. Shapira, Testing subgraphs in directed graphs, in *STOC '03: Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing*, ACM Press 2003, 700–709.

[3] N. Alon and J. Spencer, *The Probabilistic Method* (2nd edition), Wiley-Interscience 2000.

[4] J. Bang-Jensen and G. Gutin, *Digraphs: Theory, Algorithms and Applications*, Springer 2000.

[5] J. Bang-Jensen and G. Gutin, Paths, trees and cycles in tournaments, *Surveys in Graph Theory* **115** (1996), 131–170.

[6] J. Bang-Jensen, G. Gutin and H. Li, Sufficient conditions for a digraph to be Hamiltonian, *Journal of Graph Theory* **22(2)** (1996), 181–187.

[7] B. Bollobás, *Modern Graph Theory*, Springer-Verlag 1998.

[8] B. Bollobás, *Random Graphs* (2nd edition), Cambridge University Press 2001.

[9] B. Bollobás and A.M. Frieze, On matchings and Hamilton cycles in random graphs, *Ann. Discrete Math.* **28** (1985), 23-46.

[10] B. Bollovás and R. Häggkvist, Powers of Hamilton cycles in tournaments, *Journal of Combinatorial Theory, Ser. B* **50(2)** (1990), 309–318.

[11] J.A. Bondy and C. Thomassen, A short proof of Meyniel's theorem, *Discrete Math.* **19(2)** (1977), 195–197.

[12] L.M. Bregman, Some properties of nonnegative matrices and their permanents, *Soviet Mathematics Doklady* **14** (1973), 945–949.

[13] A. Busch, A Note on the Number of Hamiltonian Paths in Strong Tournaments, *The Electronic Journal Of Combinatorics* **13(1)** (2006).

[14] L. Caccetta and R. Häggkvist, On minimal digraphs with given girth, in *Proceedings of the Ninth Southeastern Conference on Combinatorics, Graph Theory and Computing*, Congressus Numerantium 21 (1978), 181–187.

[15] M. Chudnovsky, P. Seymour and B. Sullivan, Cycles in dense graphs, *submitted* (2007).

[16] V. Chvátal and E. Szemerédi, Short cycles in directed graphs, *Journal of Combinatorial Theory Series B* **35(3)** (1983), 323–327.

[17] C. Cooper and A. Frieze, Hamilton cycles in random graphs and directed graphs, *Random Structures and Algorithms* **16(4)** (2000), 369–401.

[18] C. Cooper and A. Frieze, Hamilton cycles in a class of random directed graphs, *Journal of Combinatorial Theory Series B* **62(1)** (1991), 151–163.

[19] C. Cooper, A. Frieze, M. Molloy, Hamilton cycles in random regular digraphs, *Combinatorics, Computing and Probability* **3** (1994), 39–49 .

[20] B. Csaba, On the Bollobás - Eldridge conjecture for bipartite graphs, preprint 2006.

[21] B. Csaba, A. Shokoufandeh, and E. Szemerédi, Proof of a conjecture of Bollobás and Eldridge for graphs of maximum degree three, *Combinatorica* **23** (2003), 35–72.

[22] S. Darbinyan, Pancyclicity of digraphs with large semidegrees, *Akad. Nauk Armyan. SSR Dokl.* **80** (1985), 51–54.

[23] G. Fan, New sufficient conditions for cycles in graphs, *J. Combin. Theory B* **37(3)** (1984), 221–227.

[24] E. Friedgut and J. Kahn, On the number of Hamilton cycles in a tournament, *Combinatorics, Probability and Computing* **14** (2005), 769–781.

[25] P. Fraisse and C. Thomassen, Hamiltonian dicycles avoiding prescribed arcs in tournaments, *Graphs and Cominatorics* **3** (1986), 239–250 295–352.

[26] A. Frieze, An algorithm for finding Hamilton cycles in random directed graphs, *Journal of Algorithms* **(9)** (1988), 181–204.

[27] A. Frieze and M. Krivelevich, On packing Hamilton cycles in ε-regular graphs, *J. Combin. Theory B* **94** (2005), 159–172.
[28] A. Ghouila-Houri, Une condition suffisante d'existence d'un circuit hamiltonien, *C.R. Acad. Sci. Paris* **25** (1960), 495–497.
[29] M. de Graaf, A. Schrijver, and P. D. Seymour, Directed triangles in directed graphs, *Discrete Math.* **110** (1992), 279–282.
[30] R. Häggkvist, Hamilton cycles in oriented graphs, *Combin. Probab. Comput.* **2** (1993), 25–32.
[31] R. Häggkvist and A. Thomason, Oriented Hamilton cycles in oriented graphs, in *Combinatorics, Geometry and Probability*, Cambridge University Press 1997, 339–353.
[32] R. Häggkvist and C. Thomassen, On pancyclic digraphs, *Journal of Combinatorial Theory, Ser. B* **20(1)** (1976), 20–40.
[33] P. Hamburger, P. Haxell and A. Kostochka, On directed triangles in digraphs, *submitted* (2007).
[34] Y. Hamidoune, An application of connectivity theory in graphs to factorizations of elements in groups, *European Journal of Cominatorics* **2(4)** (1981), 349–355.
[35] C. Hoang and B. Reed, A note on short cycles in digraphs, *Discrete Math.* **66(1-2)** (1987), 103–107.
[36] P. Keevash, D. Kühn and D. Osthus, An exact minimum degree condition for Hamilton cycles in oriented graphs, preprint 2007.
[37] L. Kelly, D. Kühn and D. Osthus, A Dirac type result on Hamilton cycles in oriented graphs, *submitted* (2007).
[38] J. Komlós, The Blow-up lemma, *Combin. Probab. Comput.* **8** (1999), 161–176.
[39] J. Komlós, G. N. Sárközy, and E. Szemerédi, Blow-up lemma, *Combinatorica* **17** (1997), 109–123.
[40] J. Komlós and M. Simonovits, Szemerédi's Regularity Lemma and its applications in graph theory, *Bolyai Society Mathematical Studies 2, Combinatorics, Paul Erdős is Eighty (Vol. 2)* (D. Miklós, V. T. Sós and T. Szőnyi eds.), Budapest (1996), 295–352.
[41] A. Kotzig, The decomposition of a directed graph into quadratic factors, *Acta Fac. Rerum Natur. Univ Comenin. Math Publ.* **22** (1969), 27-29.
[42] D. Kühn, D. Osthus and A. Young, k-ordered Hamilton cycles in digraphs, preprint 2007.
[43] C.J.H. McDiarmid, General percolation and random graphs, *Advances In Applied Probability* **13** (1981), 40–60.
[44] M. Meyniel,, Une condition suffisante dexistence dun circuit hamiltonien dans un graphe orient, *Journal of Combinatorial Theory, Ser. B* **14** (1973), 137-147.
[45] J.W. Moon, On subtournaments of a tournament, *Canad. Math. Bull.* **9** (1966), 297–301.
[46] J.W. Moon *Topics on tournaments*, Holt, Rinehart and Winston 1968.
[47] M. Nathanson, The Caccetta-Häggkvist Conjecture and additive number theory, *ArXiv Mathematics e-prints* (2006).
[48] Ø. Ore, Note on Hamiltonian circuits, *American Mathematics Monthly* **67** (1960), 55.
[49] J. Shen, Directed triangles in digraphs, *Journal of Combinatorial Theory, Ser. B* **74(2)** (1998), 405-407.
[50] Z.M. Song, Pancyclic oriented graphs, *Journal of Combinatorial Theory, Ser. B* **18(5)** (1994), 461–468.
[51] C. Thomassen, Long cycles in digraphs with constraints on the degrees, in *Surveys in Combinatorics* (B. Bollobás ed.), *London Math. Soc. Lecture Notes* **38** (1979), 211–228, Cambridge University Press.
[52] C. Thomassen, Long cycles in digraphs. *Proc. London Math. Soc.* **42** (1981), 231–251.
[53] C. Thomassen, Edge-disjoint Hamiltonian paths and cycles in tournaments, *Proc. London Math. Soc.* **45** (1982), 151–168.
[54] C. Thomassen, Hamiltonian-connected tournaments, *Journal of Combinatorial Theory, Ser. B* **45(1)** (1982), 151–168.

[55] C. Thomassen, An Ore-type condition implying a digraph to be pancyclic, *Discrete Math.* **19(1)** (1977), 85–92.

[56] C. Thomassen, Connectivity in tournaments. In *Graph theory and combinatorics (Cambridge, 1983)*, 305–313.

[57] C. Thomassen, Edge-disjoint Hamiltonian paths and cycles in tournaments, *Proc London Math. Soc. 3* **45(1)** (1982), 151–168.

[58] T.W. Tillson, A Hamiltonian decomposition of K^*_{2m}, $2m \geq 8$, *Journal of Combinatorial Theory, Ser. B* **29(1)** (1980), 68–74.

[59] A. Young, Extremal problems for dense graphs and digraphs, PhD thesis, School of Mathematics, University of Birmingham 2007.