



RESEARCH ARTICLE

Cycle Partitions in Dense Regular Digraphs and Oriented Graphs

Allan Lo¹, Viresh Patel² and Mehmet Akif Yıldız³

¹School of Mathematics, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom;
E-mail: s.a.lo@bham.ac.uk.

²School of Mathematical Sciences, Queen Mary University of London, Mile End Road, London, E1 4NS, United Kingdom;
E-mail: viresh.patel@qmul.ac.uk.

³Korteweg-de Vries Instituut voor Wiskunde, Universiteit van Amsterdam, Science Park 107, Amsterdam, 1098XG, The Netherlands; E-mail: m.a.yildiz@uva.nl (corresponding author).

Received: 20 September 2023; **Revised:** 6 December 2024; **Accepted:** 20 December 2024

2020 Mathematics Subject Classification: *Primary* – 05C35; *Secondary* – 05C20, 05C45

Abstract

A conjecture of Jackson from 1981 states that every d -regular oriented graph on n vertices with $n \leq 4d + 1$ is Hamiltonian. We prove this conjecture for sufficiently large n . In fact we prove a more general result that for all $\alpha > 0$, there exists $n_0 = n_0(\alpha)$ such that every d -regular digraph on $n \geq n_0$ vertices with $d \geq \alpha n$ can be covered by at most $n/(d + 1)$ vertex-disjoint cycles, and moreover that if G is an oriented graph, then at most $n/(2d + 1)$ cycles suffice.

Contents

1	Introduction	2
1.1	Related work	3
1.1.1	Path cover	3
1.1.2	Extending perfect matchings	3
1.2	Extremal examples for Conjecture 1.1	4
2	Notation and preliminaries	4
2.1	Sketch of proof of Theorem 1.3	5
3	Robust expanders	6
4	Finding long cycles	8
5	Balanced path systems and path contraction	10
5.1	Proof of Theorem 1.3	11
6	Proof of Lemma 5.4	12
7	Proof of Lemma 6.1	15
7.1	Flows	17
7.2	Preliminaries	18
7.3	Proof of Lemma 6.1	20
8	Proof of Lemma 6.3	24

9 Conclusion	29
9.1 Path cover for (non-regular) graphs	29
9.2 Edge-disjoint cycles	29
9.3 Connectivity and regularity	30
References	30

1. Introduction

A *Hamilton cycle* in a (directed) graph is a (directed) cycle that visits every vertex. Hamilton cycles are one of the most intensely studied structures in graph theory and there are numerous results that establish (best-possible) conditions guaranteeing their existence. The seminal result of Dirac [7] states that every graph on $n \geq 3$ vertices with minimum degree at least $n/2$ is Hamiltonian. Ghouila-Houri [12] showed the corresponding version in directed graphs (*digraph* for short), that is, every digraph on $n \geq 3$ vertices with minimum semi-degree at least $n/2$ (i.e., every vertex has in- and outdegree at least $n/2$) is Hamiltonian. These bounds are tight by taking for instance the disjoint union of two cliques (a regular extremal example) or a slightly imbalanced complete bipartite (di)graph (an irregular extremal example). Recall that an *oriented graph* is a digraph that can have at most one edge between each pair of vertices (whereas a digraph can have up to two, one in each direction). For oriented graphs, a more recent result of Keevash, Kühn, and Osthus [17] established a (tight) minimum semi-degree threshold of $\lceil (3n - 4)/8 \rceil$ for Hamiltonicity. In contrast to graphs and digraphs, there are no *regular* extremal examples for this result. Jackson [16] conjectured in 1981 that regularity actually reduces the degree threshold significantly for oriented graphs:

Conjecture 1.1 (Jackson [16]). *For each $d > 2$, every d -regular oriented graph (i.e., every vertex has d in- and outneighbours) on $n \leq 4d + 1$ vertices has a Hamilton cycle.*

The disjoint union of two regular tournaments shows that Jackson's conjecture is best possible. This example works for $n \equiv 2 \pmod{4}$ (since regular tournaments require an odd number of vertices), but similar examples exist for $n \equiv 0, 1, 3 \pmod{4}$ (see Section 1.2).

We note that an approximate version of Jackson's conjecture was recently verified by the current authors in [27], that is, for every $\varepsilon > 0$, there exists $n_0(\varepsilon)$ such that every d -regular oriented graph on $n \geq n_0(\varepsilon)$ vertices with $d \geq (1/4 + \varepsilon)n$ is Hamiltonian. Here, we verify the exact version for large n .

Theorem 1.2. *There exists an integer n_0 such that every d -regular oriented graph on $n \geq n_0$ vertices with $n \leq 4d + 1$ has a Hamilton cycle.*

Generalizing questions about Hamilton cycles, one can consider the question of covering the vertices of a (di)graph with as few vertex-disjoint cycles as possible. Indeed, we prove Theorem 1.2 by showing a more general result about covering regular digraphs and oriented graphs with few vertex-disjoint cycles.

Theorem 1.3. *For all $\alpha > 0$, there exists $n_0 = n_0(\alpha)$ such that every d -regular digraph G on $n \geq n_0$ vertices with $d \geq \alpha n$ can be covered by at most $n/(d + 1)$ vertex-disjoint cycles. Moreover, if G is an oriented graph, then at most $n/(2d + 1)$ cycles suffice.*

This is best possible by considering the disjoint union of complete digraphs of order $d + 1$ for digraphs and the disjoint union of regular tournaments of order $2d + 1$ for oriented graphs. Notice that we have $n/(2d + 1) < 2$ when $n \leq 4d + 1$, so that Theorem 1.3 implies Theorem 1.2. We also point out that the proof of Theorem 1.3 in fact shows that each cycle is relatively long (of length at least $d/2$).

Theorem 1.3 generalizes the following result of Gruslys and Letzter [13] from regular graphs to regular digraphs and oriented graphs.

Theorem 1.4 (Gruslys and Letzter [13]). *For all $\alpha > 0$, there exists $n_0 = n_0(\alpha)$ such that every d -regular graph on $n \geq n_0$ vertices with $d \geq \alpha n$ can be covered by at most $n/(d + 1)$ vertex-disjoint cycles.*

Theorem 1.3 implies Theorem 1.4 by making every edge into a directed 2-cycle.

1.1. Related work

Theorem 1.3 also has connections with several well-studied problems in extremal graph theory: here we mention some of them.

1.1.1. Path cover

A weaker version of cycle cover is path cover. The *path cover number* of a (di)graph G , denoted by $\pi(G)$, is the minimum number of vertex-disjoint (directed) paths needed to cover $V(G)$. This was introduced by Ore [30], and he showed that $\pi(G) \leq n - \sigma_2(G)$ holds where $\sigma_2(G)$ denotes the minimum sum of the degrees over all non-adjacent vertices. Magnant and Martin [28] conjectured that *regularity* significantly reduces the upper bound for $\pi(G)$:

Conjecture 1.5 (Magnant and Martin [28]). *If G is a d -regular graph on n vertices, then $\pi(G) \leq n/(d+1)$.*

It is known that Conjecture 1.5 holds for small values of d (see [28] for $d \leq 5$ and see [9] for $d = 6$). Han [15] showed that, for dense graphs, it is enough to use $1 + n/(d+1)$ paths to cover almost all vertices. Also, Theorem 1.4 verifies Conjecture 1.5 in the dense case. It is worth noting that the Linear Arboricity Conjecture [2] implies Conjecture 1.5 for odd values of d , and gives $\pi(G) \leq 2n/(d+2)$ for general d (see [9] for a detailed discussion).

For digraphs, the classical result of Gallai and Milgram [11] states that $\pi(G)$ can be bounded above by the size of the maximum independent set (and Dilworth's [6] theorem says that equality holds for the special case of posets). As our Theorem 1.3 generalizes Theorem 1.4 from graphs to digraphs and oriented graphs, we believe the following stronger version of Conjecture 1.5 holds, which Theorem 1.3 establishes in the dense case.

Conjecture 1.6. *If G is a d -regular digraph on n vertices, then $\pi(G) \leq n/(d+1)$. Moreover, $\pi(G) \leq n/(2d+1)$ holds if G is oriented.*

Also, Conjecture 1.6 implies Conjecture 1.5 by making every edge into a directed 2-cycle.

1.1.2. Extending perfect matchings

Gruslys and Letzter [13], as well as proving Theorem 1.4, proved that every large d -regular bipartite graph G on n vertices with d linear in n can be covered by at most $n/2d$ vertex-disjoint paths. They mentioned that one should be able to replace paths by cycles. Indeed, as a corollary of Theorem 1.3, the result below shows that those cycles can be found in such a way that they even contain any prescribed perfect matching.

Corollary 1.7. *For all $\alpha > 0$, there exists $n_1 = n_1(\alpha)$ such that, for every d -regular bipartite graph on $n \geq n_1$ vertices with $d \geq \alpha n$, any perfect matching can be extended to vertex-disjoint cycles covering all vertices with at most $n/2d$ cycles.*

Proof of Corollary 1.7. Let $\alpha > 0$ and $n_1 = 2n_0(\alpha)$, where n_0 is the function given in Theorem 1.3. Let G be a d -regular bipartite graph on $n \geq n_1$ vertices with $d \geq \alpha n$ and vertex classes X and Y . Since G is bipartite and regular, n is even and $|X| = |Y| = n/2$. Let M be any perfect matching of G . Let $X = \{x_1, \dots, x_{n/2}\}$ and $Y = \{y_1, \dots, y_{n/2}\}$ be such that $x_i y_i \in E(M)$ for all i . Define the digraph H on X such that for any distinct $i, j \in [n/2]$, $x_i x_j \in E(H)$ if and only if $x_i y_j \in E(G)$. Note that H is $(d-1)$ -regular on $n/2$ vertices. By Theorem 1.3, H can be covered by at most $n/2d$ vertex-disjoint cycles. Note that a (directed) cycle $x_{i_1} x_{i_2} \dots x_{i_\ell}$ in H corresponds to a cycle $x_{i_1} y_{i_2} x_{i_2} y_{i_3} \dots x_{i_\ell} y_{i_1}$ in G . Therefore G can be covered by at most $n/2d$ vertex-disjoint cycles that contain all the edges of M . $\square$

Note that Corollary 1.7 is tight by considering the disjoint union of $n/2d$ many $K_{d,d}$'s. It also shows that d -regular bipartite graphs on n (sufficiently large) vertices with $d > n/4$ are examples of graphs in which every perfect matching can be extended into a Hamilton cycle. This property is called the PMH-property in [1]. Häggkvist [14] initiated the study of sufficient conditions for the PMH-property (using

the name F -Hamiltonian where F is a perfect matching) by showing $\sigma_2(G) \geq n + 1$ is sufficient. Las Vergnas [24] proved a similar condition for bipartite graphs, which also (almost) implies Corollary 1.7 in the case $d > n/4$, and Yang [35] gave minimum edge density conditions to guarantee the PMH property in graphs and bipartite graphs. In the sparse setting, as a special case of a conjecture of Ruskey and Savage [32], Fink [10] proved that the hypercube has the PMH-property.

1.2. Extremal examples for Conjecture 1.1

We end the introduction with the deferred examples from earlier. Note that we give a general overview of the paper at the end of the next section.

Recall that the disjoint union of two regular tournaments on $n/2$ vertices shows that Jackson's conjecture (Conjecture 1.1) is best possible. This example works when $n \equiv 2 \pmod{4}$ and similar examples for other values of n are constructed as follows. When $n \equiv 0 \pmod{4}$, we take the disjoint union of two complete graphs on $n/2$ vertices, remove a perfect matching, and orient the edges so that the resulting oriented graph is regular. When $n \equiv 3 \pmod{4}$, we take a disjoint union of a regular tournament on $(n-1)/2$ vertices and a regular orientation of the complete graph minus a perfect matching on $(n+1)/2$ vertices. When $n \equiv 1 \pmod{4}$, we take a regular orientation of a disjoint union of a clique minus a perfect matching on $(n-1)/2$ vertices and a clique minus a Hamilton cycle on $(n+1)/2$ vertices.

2. Notation and preliminaries

Throughout the paper, we use standard graph theory notation and terminology. For $k \in \mathbb{N}$, we sometimes denote the set $\{1, 2, \dots, k\}$ by $[k]$. For a digraph G , we denote its vertex set by $V(G)$ and its edge set $E(G)$, and sometimes write $|G| := |V(G)|$ and $e(G) := |E(G)|$. For $a, b \in V(G)$, we write ab for the directed edge from a to b . We write $H \subseteq G$ to mean H is a *subdigraph* of G (i.e., $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$). We sometimes think of $F \subseteq E(G)$ as a subdigraph of G with vertex set consisting of those vertices incident to edges in F and with edge set F . We write $G - F$ for the digraph obtained from G by deleting the edges in F . For $S \subseteq V(G)$, we write $G[S]$ for the subdigraph of G induced by S and $G - S$ for the digraph $G[V(G) \setminus S]$. For $A, B \subseteq V(G)$ not necessarily disjoint, we define $E_G(A, B) := \{ab \in E(G) : a \in A, b \in B\}$ and we write $e_G(A, B) := |E_G(A, B)|$ ¹. We often drop subscripts if these are clear from context. For two digraphs H_1 and H_2 , the union $H_1 \cup H_2$ is the digraph with vertex set $V(H_1) \cup V(H_2)$ and edge set $E(H_1) \cup E(H_2)$. We say an undirected graph G is bipartite with bipartition A, B if $V(G) = A \cup B$ and $E(G) \subseteq \{ab : a \in A, b \in B\}$.

For a digraph G and $v \in V(G)$, we denote the set of outneighbours and inneighbours of v by $N_G^+(v)$ and $N_G^-(v)$, respectively, and we write $d_G^+(v) := |N_G^+(v)|$ and $d_G^-(v) := |N_G^-(v)|$ for the out- and indegree of v , respectively. For $S \subseteq V(G)$ we write $d_G^-(v, S) := |N_G^-(v) \cap S|$ and $d_G^+(v, S) := |N_G^+(v) \cap S|$. We write $\delta^+(G)$ and $\delta^-(G)$, respectively, for the minimum out- and indegree of G and $\delta^0(G) := \min\{\delta^+(G), \delta^-(G)\}$ for the minimum semi-degree. Similarly, the maximum semi-degree $\Delta^0(G)$ of G is defined by $\Delta^0(G) := \max\{\Delta^+(G), \Delta^-(G)\}$ where $\Delta^+(G)$ and $\Delta^-(G)$ denote the maximum out- and maximum indegree of G , respectively. A digraph is called d -regular if each vertex has exactly d outneighbours and d inneighbours.

The notation above extends to undirected graphs in the obvious ways. In particular for an undirected graph G , we write $\Delta(G)$ and $\delta(G)$, respectively for the maximum degree and the minimum degree. A graph is called d -regular if each vertex has exactly d neighbours. For a vertex $v \in V(G)$ and subset $S \subseteq V(G)$, we write $d_G(v, S) := |N_G(v) \cap S|$.

A directed path Q in a digraph G is a subdigraph of G such that $V(Q) = \{v_1, \dots, v_k\}$ for some $k \in \mathbb{N}$ and $E(Q) = \{v_1v_2, v_2v_3, \dots, v_{k-1}v_k\}$. We denote such a directed path by its vertices in order (i.e., we

¹We DO NOT write $G[A, B]$ for the graph with vertex set $A \cup B$ and edge set $E_G(A, B)$, but instead for a bipartite undirected graph; see below.

write $Q = v_1 v_2 \cdots v_k$). A directed cycle in G is exactly the same except that it also includes the edge $v_k v_1$. Sometimes we identify paths with their edge sets.

A set of vertex-disjoint directed paths $\mathcal{Q} = \{Q_1, Q_2, \dots\}$ in a digraph G is called a *path system* in G . We interchangeably think of $\mathcal{Q}$ as a set of vertex-disjoint directed paths in G and as a subdigraph of G with vertex set $V(\mathcal{Q}) = \bigcup_i V(Q_i)$ and edge set $E(\mathcal{Q}) = \bigcup_i E(Q_i)$. We sometimes call this subdigraph the graph induced by $\mathcal{Q}$. A matching M in a digraph (or undirected graph) G is a set of edges $M \subseteq E(G)$ such that every vertex of G is incident to at most one edge in M . If $M = \{a_i b_i : i \in [m]\}$ is a matching in a digraph, then we write $V^+(M) := \{a_i : i \in [m]\}$ and $V^-(M) = \{b_i : i \in [m]\}$. A *1-factor* in a digraph G is a set of vertex-disjoint directed cycles whose union has the same vertices as G .

Throughout the paper, we will work with partitions of a vertex set V of the form $\{V_{ij} : i, j \in [k]\}$ for some $k \in \mathbb{N}$ (so a partition into k^2 parts). For each i , we write $V_{i*} := V_{i1} \cup \cdots \cup V_{ik}$ and $V_{*i} := V_{1i} \cup \cdots \cup V_{ki}$. Note that $\{V_{i*} : i \in [k]\}$ is a partition of V as is $\{V_{*i} : i \in [k]\}$. Note also that $V_{ij} = V_{i*} \cap V_{*j}$.

For two sets A and B , the *symmetric difference* of A and B is the set $A \Delta B := (A \setminus B) \cup (B \setminus A)$. For $x, y \in (0, 1]$, we often use the notation $x \ll y$ to mean that x is sufficiently small as a function of y [i.e., $x \leq f(y)$] for some implicitly given non-decreasing function $f : (0, 1] \rightarrow (0, 1]$. We implicitly assume all constants in such hierarchies are positive, and we omit floors and ceilings whenever this does not affect the argument.

We use the following non-standard notation. Let G be a directed graph and let $U, W \subseteq V(G)$ not necessarily disjoint. We define $G[U, W]$ to be the auxiliary undirected bipartite graph with bipartition U, W where, for each $u \in U$ and $w \in W$, uw is an (undirected) edge of $G[U, W]$ if and only if uw is a directed edge in $E(G)$. Note that for each vertex in $U \cap W$, there are two copies of the vertex in $G[U, W]$ which are viewed as distinct. So $G[U, W]$ has $|U| + |W|$ vertices and $e_G(U, W)$ edges.

2.1. Sketch of proof of Theorem 1.3

Theorem 1.3 is proved in several steps. One of the key ingredients is a structural result for directed graphs (Theorem 3.5) that we derive from a result of Kühn, Lo, Osthus and Staden [21] about partitioning dense regular undirected graphs into robust expanders. Robust expansion is a notion introduced and used by Kühn and Osthus together with several coauthors to obtain a number of breakthrough results on (di)graph decompositions and Hamiltonicity (see e.g., [23, 22, 21]). In Section 3, we give the necessary background on robust expansion before proving Theorem 3.5, which we informally describe below.

Informally, robust expanders are dense (di)graphs that are highly connected in a certain sense, and one of their key properties is that they are Hamiltonian under suitable (mild) degree conditions (see Theorem 4.2). Moreover, they are robust to small alterations (see Lemma 3.1). If we could show that every d -regular digraph (resp. oriented graph) can be partitioned into at most $n/(d+1)$ (resp. $n/(2d+1)$) robust expanders, it would be enough to prove Theorem 1.3. Such a partition does not exist in general, but our structural result, Theorem 3.5, gives us a starting point. Roughly, it says that for any d -regular n -vertex digraph G with d linear in n and n sufficiently large, there exist two vertex partitions $V(G) = V_{1*} \cup \cdots \cup V_{k*}$ and $V(G) = V_{*1} \cup \cdots \cup V_{*k}$ with $k \leq 1 + n/(d+1)$ such that for each $i \in [k]$, $G[V_{i*}, V_{*i}]$ is a bipartite robust expander and $|V_{i*}| \approx |V_{*i}|$. Note here that k , the number of parts in each partition, is at most one more than the number of cycles we desire (in the case of digraphs). Writing $V_{ij} = V_{i*} \cap V_{*j}$ for all $i, j \in [k]$, we can think of our two partitions as a single k^2 -partition $\mathcal{P} = \{V_{ij} : i, j \in [k]\}$ of $V(G)$.

If the partition $\mathcal{P}$ described above is balanced, meaning that $|V_{i*}| = |V_{*i}|$ for all i , then it turns out that G can be covered by k vertex-disjoint cycles.² So to prove Theorem 1.3, we would like to ensure our partition is balanced, and when $k = 1 + \lfloor n/(d+1) \rfloor$, we would like to slightly improve on the number of cycles. For the latter, in Section 4, we define an auxiliary graph $S(\mathcal{P})$ for the partition $\mathcal{P}$, which has vertex set $[k]$ and $ij \in E(S(\mathcal{P}))$ if and only if $V_{ij} \cup V_{ji} \neq \emptyset$. For each connected component $I \subseteq [k]$ of $S(\mathcal{P})$, Lemma 4.1 shows how to find a cycle whose vertices are exactly $\bigcup_{i \in I} V_{i*} \cup V_{*i}$

²This follows essentially from Lemma 4.1 although we do not show it explicitly in the paper.

(provided $\mathcal{P}$ is balanced). These cycles are necessarily disjoint for different connected components of $S(\mathcal{P})$, so we obtain a cycle partition where the number of cycles is exactly the number of connected components of $S(\mathcal{P})$. For digraphs, if $S(\mathcal{P})$ has at least one edge (so has at most $k - 1 \leq n/(d + 1)$ connected components), then Theorem 1.3 follows. Similarly, for oriented graphs, one can show (using Proposition 6.5) that at least one edge in $S(\mathcal{P})$ is enough to prove Theorem 1.3 for oriented graphs. Thus, it is enough if our partition $\mathcal{P}$ is balanced and $S(\mathcal{P})$ has at least one edge.

To ensure our partition is balanced (i.e., that $|V_{i*}| = |V_{*i}|$ for all i), note first that, as mentioned earlier, Theorem 3.5 already guarantees that $|V_{i*}| \approx |V_{*i}|$ for all i . The idea will be to find and contract a suitable collection of vertex-disjoint paths $\mathcal{Q}$. Here suitable simply means that the numbers of edges of $\mathcal{Q}$ between the various parts of $\mathcal{P}$ are related in the right way, in which case we call $\mathcal{Q}$ a $\mathcal{P}$ -balanced path system (introduced in Section 5). If we can find such a $\mathcal{P}$ -balanced path system with few edges, then contracting the paths will result in a graph with an adjusted partition $\mathcal{P}'$ that is balanced and has similar properties as $\mathcal{P}$ (Proposition 5.3), and any cycle partition in the contracted graph can be lifted to a cycle partition in the original graph (Propositions 5.1). If the path system $\mathcal{Q}$ satisfies some further properties then we can also guarantee that $S(\mathcal{P}')$ has at least one edge and this will be enough to prove Theorem 1.3. Lemma 5.4 says that we can always find a $\mathcal{P}$ -balanced path system with the required properties, and so we reduce the task of proving Theorem 1.3 to the task of proving Lemma 5.4.

In Section 6 we reduce the proof of Lemma 5.4 to two lemmas, namely Lemmas 6.1 and 6.3. Lemma 6.1 allows us to find a $\mathcal{P}$ -balanced path system with the desired properties under most circumstances, while Lemma 6.3 (together with Proposition 6.2) allows us to find such a path system in the remaining ‘extremal’ circumstances when G is close to the disjoint union of cliques. Lemma 6.1 is proved in Section 7 using a flow argument and Lemma 6.3 is proved in Section 8. These last two lemmas are the most technical parts of the paper, so we defer their sketch of proofs to their respective sections.

3. Robust expanders

We first define robust expansion for graphs. Let $0 < \nu \leq \tau < 1$ and let G be a graph on n vertices. For $S \subseteq V(G)$, the ν -robust neighbourhood $RN_{\nu,G}(S)$ is the set of all those vertices with at least νn neighbours in S . We say that G is a robust (ν, τ) -expander if every $S \subseteq V(G)$ with $\tau n \leq |S| \leq (1 - \tau)n$ satisfies $|RN_{\nu,G}(S)| \geq |S| + \nu n$. In fact we will mainly be concerned with bipartite robust expanders. Let G be a bipartite graph with bipartition A, B . We say that G is a *bipartite robust (ν, τ) -expander with bipartition A, B* if every $S \subseteq A$ with $\tau|A| \leq |S| \leq (1 - \tau)|A|$ satisfies $|RN_{\nu,G}(S) \cap B| \geq |S| + \nu n$. Note that the order of A and B matters here. Our first lemma says that bipartite robust expansion is robust to small alterations; the lemma can easily be derived from the definition of robust expansion.

Lemma 3.1 [33, Lemma 3.4.9]. *Let $0 < 1/n \ll \nu \leq \tau \ll 1$ with $\nu \leq 1/2$. Let G be a bipartite graph with $U \subseteq V(G)$. Suppose $G[U]$ is a bipartite robust (ν, τ) -expander on n vertices with bipartition A, B and that $A', B' \subseteq V(G)$ are sets satisfying $|A \Delta A'| + |B \Delta B'| \leq \nu|A|/4$. Then $G[A' \cup B']$ is a bipartite robust $(\nu/2, 2\tau)$ -expander with bipartition A', B' .*

Next we give a structural result due to Kühn, Lo, Osthus and Staden [21] which states that any regular graph of linear minimum degree has a vertex partition into a small number of parts where each part induces a robust expander or a bipartite robust expander. In fact we state the special case of this result for bipartite graphs, which is all we require.

Theorem 3.2 (Bipartite special case of [21, Theorem 3.1]). *For all $\alpha, \tau > 0$ and every non-decreasing function $f : (0, 1) \rightarrow (0, 1)$, there exists n_0 such that the following holds. For all d -regular bipartite graphs G on $2n \geq n_0$ vertices with bipartition A, B and $d \geq \alpha n$, there exist ρ, ν with*

$$1/n_0 \leq \rho \leq \nu \leq \tau; \quad \rho \leq f(\nu); \quad \text{and} \quad 1/n_0 \leq f(\rho)$$

such that there is a partition of $V(G)$ into sets $A_1, \dots, A_k, B_1, \dots, B_k$ with the following properties for all $i \in [k]$:

- (i) $G[A_i \cup B_i]$ is a bipartite robust (ν, τ) -expander with bipartition A_i, B_i ;
- (ii) for all $x \in A_i \cup B_i$ and $j \in [k]$, $d(x, A_i \cup B_i) \geq d(x, A_j \cup B_j)$, so in particular, $\delta(G[A_i \cup B_i]) \geq d/k$;
- (iii) $||A_i| - |B_i|| \leq 2\rho n$;
- (iv) all but at most $2\rho n$ vertices $x \in A_i \cup B_i$ satisfy $d(x, A_i \cup B_i) \geq d - 2\rho n$;
- (v) $k \leq n/(d - 2\rho n)$;
- (vi) $A_1, \dots, A_k$ is a partition of A and $B_1, \dots, B_k$ is a partition of B .

Remark 3.3. Note that (i)–(iv) follow directly from the statement of [21, Theorem 3.1]. While (v) and (vi) cannot be deduced immediately from the statement of [21, Theorem 3.1], they follow immediately from the proof.³

We now define robust expansion for digraphs. Let $0 < \nu \leq \tau < 1$ and let G be a digraph on n vertices. For $S \subseteq V(G)$, the ν -robust outneighbourhood $RN_{\nu, G}^+(S)$ is the set of all those vertices with at least νn inneighbours in S . We say that G is a robust (ν, τ) -outexpander if every $S \subseteq V(G)$ with $\tau n \leq |S| \leq (1 - \tau)n$ satisfies $|RN_{\nu, G}^+(S)| \geq |S| + \nu n$.

Proposition 3.4. Let $0 < 1/m \ll \nu \leq \tau \ll 1$ with $\nu \leq 1/2$. Suppose G is a bipartite graph with bipartition A, B , where $A = \{a_1, \dots, a_m\}$ and $B = \{b_1, \dots, b_m\}$. If G is a bipartite robust (ν, τ) -expander with bipartition A, B , then the digraph H with vertex set A and edge set $E(H) = \{a_i a_j : a_i b_j \in E(G), i \neq j\}$ is a robust $(\nu/2, 2\tau)$ -outexpander.

The proposition above follows immediately from the definitions; we crudely replace (ν, τ) with $(\nu/2, 2\tau)$ to account for the loss of any edges of the form $a_i b_i$.

We now state and prove a structure lemma for regular digraphs that is derived from Theorem 3.2. This will be one of the key ingredients in the proof of Theorem 1.3. It says that any dense regular digraph G has two vertex partitions $V(G) = V_{1*} \cup \dots \cup V_{k*}$ and $V(G) = V_{*1} \cup \dots \cup V_{*k}$ with k relatively small such that for each i , the (undirected) bipartite graph $G[V_{i*}, V_{*i}]$ is a bipartite robust expander (with bipartition V_{i*}, V_{*i}). Various other degree and size conditions relating to the partition are also given. Note that in the theorem below, we actually give a partition $\{V_{ij} : i, j \in [k]\}$ and recall that for each $i \in [k]$, we write $V_{i*} := V_{i1} \cup \dots \cup V_{ik}$ and $V_{*i} := V_{1i} \cup \dots \cup V_{ki}$. Thus $V_{ij} = V_{i*} \cap V_{*j}$.

Theorem 3.5. For all $\alpha, \tau > 0$ and every non-decreasing function $f : (0, 1) \rightarrow (0, 1)$, there exists n_0 such that the following holds. For all d -regular digraphs G on $n \geq n_0$ vertices with $d \geq \alpha n$, there exist ρ, ν with

$$1/n_0 \leq \rho \leq \nu \leq \tau; \quad \rho \leq f(\nu); \quad \text{and} \quad 1/n_0 \leq f(\rho)$$

such that there is a partition $\mathcal{P} = \{V_{ij} : i, j \in [k]\}$ of $V(G)$ satisfying, for all $i, j \in [k]$,

- (i) $G[V_{i*}, V_{*i}]$ is a bipartite robust (ν, τ) -expander with bipartition V_{i*}, V_{*i} ;
- (ii) for all $x \in V_{ij}$ and $i', j' \in [k]$, $d^+(x, V_{*i}) \geq d^+(x, V_{*i'})$ and $d^-(x, V_{j*}) \geq d^-(x, V_{j'*})$, so in particular, $\delta(G[V_{i*}, V_{*i}]) \geq d/k$;
- (iii) $||V_{i*}| - |V_{*i}|| \leq 2\rho n$;
- (iv) all but at most $2\rho n$ vertices $x \in V_{ij}$ satisfy $d^+(x, V_{*i}), d^-(x, V_{j*}) \geq d - 2\rho n$;
- (v) $k \leq n/(d - 2\rho n)$.

Proof. We simply apply Theorem 3.2 to the obvious bipartite graph obtained from the directed graph G , with the natural correspondence in parameters, as follows.

³The idea of the proof is to successively refine partitions of $V(G) = A \cup B$ as follows. Assume we have obtained a partition $\mathcal{U} = \{U_1, \dots, U_r\}$ of $V(G)$ such that there are few edges leaving or entering U for every $U \in \mathcal{U}$. If for some $U \in \mathcal{U}$ we have that $G[U]$ is not a bipartite robust expander (with bipartition $U \cap A, U \cap B$), then there is some $S \subseteq U \cap A$ whose robust neighbourhood $R \subseteq U \cap B$ is not much larger than S . Writing $U' = S \cup R$ and $U'' = U \setminus U'$, we let $\mathcal{U}' = \mathcal{U} \setminus \{U\} \cup \{U', U''\}$. It is not too hard to show (using the fact that G is regular) that, as with $\mathcal{U}$, there are not many edges entering or leaving each $U \in \mathcal{U}'$. We continue refining the partition in this way until we obtain a partition $\mathcal{U}^*$ where every $U \in \mathcal{U}^*$ satisfies that $G[U]$ is a bipartite robust expander with bipartition $U \cap A, U \cap B$. The process of refining the partition must eventually stop because each $U \in \mathcal{U}$ cannot be much smaller than d (the degree of G) since not many edges enter or leave U . This essentially shows (i), (iii), (v) and (vi), while (ii) and (iv) are obtained by making slight adjustments to the final partition.

For any d -regular digraph G on n vertices, let H_G be the d -regular bipartite undirected graph on $2n$ vertices defined as follows. Let $V(H_G) = A \cup B$, where A and B are disjoint copies of $V(G)$ and for $a \in A$ and $b \in B$ let $ab \in E(H_G)$ if and only if $ab \in E(G)$. Now applying Theorem 3.2 with H_G playing the role of G , we obtain the following statement.

For all $\alpha, \tau > 0$ and every non-decreasing function $f : (0, 1) \rightarrow (0, 1)$, there exists n_0 such that the following holds. For all d -regular digraphs G on $n \geq n_0$ vertices and $d \geq \alpha n$, taking H_G to be the corresponding bipartite d -regular graph on $2n \geq n_0$ vertices there exist ρ, ν with

$$1/n_0 \leq \rho \leq \nu \leq \tau; \quad \rho \leq f(\nu); \quad \text{and} \quad 1/n_0 \leq f(\rho)$$

such that there is a partition of $V(H_G)$ into sets $A_1, \dots, A_k, B_1, \dots, B_k$ with the following properties for all $i \in [k]$:

- (i') $H_G[A_i \cup B_i]$ is a bipartite robust (ν, τ) -expander with bipartition A_i, B_i ;
- (ii') for all $x \in A_i \cup B_i$ and $j \in [k]$, $d_{H_G}(x, A_i \cup B_i) \geq d_{H_G}(x, A_j \cup B_j)$, so in particular, $\delta(H_G[A_i \cup B_i]) \geq d/k$;
- (iii') $||A_i| - |B_i|| \leq 2\rho n$;
- (iv') all but at most $2\rho n$ vertices $x \in A_i \cup B_i$ satisfy $d_{H_G}(x, A_i \cup B_i) \geq d - 2\rho n$;
- (v') $k \leq n/(d - 2\rho n)$;
- (vi') $A_1, \dots, A_k$ is a partition of A and $B_1, \dots, B_k$ is a partition of B .

For $i, j \in [k]$, define $V_{ij} = A_i \cap B_j$ (where we think of A_i and B_j as sets of vertices of the digraph G). First note that, by (vi'), $\mathcal{P} = \{V_{ij} : i, j \in [k]\}$ is a partition of $V(G)$, and $A_i = V_{i*}$ and $B_i = V_{*i}$. Now the natural correspondence between H_G and G means that (i')–(v') imply (i)–(v), respectively. $\square$

4. Finding long cycles

Theorem 3.5 from the previous section shows that every (dense) regular digraph has a vertex partition with some useful properties. In this section, we show how properties (i) and (ii) from Theorem 3.5 together with a simple balancing condition on the partition allow us to construct few long cycles that can cover all the vertices of several parts in the partition.

Let G be a digraph on V and let $\mathcal{P} = \{V_{ij} : i, j \in [k]\}$ be a partition of V where we allow some parts to be empty. We say that $\mathcal{P}$ is *balanced* if $|V_{i*}| = |V_{*i}|$ for all $i \in [k]$. For $i, j \in [k]$, let G_{ij} be the subdigraph of G on $V_{i*} \cup V_{*j}$ with edges from V_{i*} to V_{*j} , that is, $E(G_{ij}) = E_G(V_{i*}, V_{*j}) = E(G) \cap (V_{i*} \times V_{*j})$.

Define $S(\mathcal{P})$ to be the graph (without loops) on $[k]$ such that for all $i, j \in [k]$ with $i \neq j$, $ij \in E(S(\mathcal{P}))$ if and only if $V_{ij} \cup V_{ji} \neq \emptyset$. The connected components of $S(\mathcal{P})$ will determine which parts of $\mathcal{P}$ can be covered by one long cycle. We remark that, by definition, $\bigcup_{i \in I} (V_{i*} \cup V_{*i})$ and $\bigcup_{j \in J} (V_{j*} \cup V_{*j})$ are disjoint for two distinct connected components I and J of $S(\mathcal{P})$. The aim of this section is to prove the following lemma.

Lemma 4.1. *Let $1/m \ll \nu \leq \tau \ll \alpha < 1$. Let G be a digraph with a balanced vertex partition $\mathcal{P} = \{V_{ij} : i, j \in [k]\}$. Let I be a connected component in $S(\mathcal{P})$. Suppose that, for all $i \in I$, $G[V_{i*}, V_{*i}]$ is a bipartite robust (ν, τ) -expander with bipartition V_{i*}, V_{*i} such that $|V_{i*}| \geq m$ and $\delta(G[V_{i*}, V_{*i}]) \geq \alpha|V_{i*}|$. Then there exists a cycle C in G with $V(C) = \bigcup_{i \in I} (V_{i*} \cup V_{*i})$.*

The lemma above will be used as follows. Suppose we have a dense d -regular digraph G with a vertex partition $\mathcal{P}$ such as that given by Theorem 3.5 but with the additional property that $\mathcal{P}$ is balanced. Then Lemma 4.1 applied to each connected component of $S(\mathcal{P})$ gives us a collection of s vertex-disjoint cycles that cover $V(G)$, where s is the number of connected components in $S(\mathcal{P})$. So in later sections we will be interested in obtaining balanced partitions $\mathcal{P}$ where the number of connected components of $S(\mathcal{P})$ is ‘small’.

We need the following theorem, which states that a robust outexpander with linear minimum degree contains a (directed) Hamilton cycle.

Theorem 4.2 ([23]; see also [25]). Let $1/n \ll \nu \leq \tau \ll \gamma < 1$. Let G be a robust (ν, τ) -outexpander on n vertices with $\delta^0(G) \geq \gamma n$. Then G contains a Hamilton cycle.

The next lemma shows that, under the conditions of Lemma 4.1, $V_{i*} \cup V_{*i}$ can be covered either by vertex-disjoint paths from $V_{i*} \setminus V_{ii}$ to $V_{*i} \setminus V_{ii}$ or by a cycle, and it is used inductively to prove Lemma 4.1.

Lemma 4.3. Let $1/m \ll \nu \leq \tau \ll \alpha < 1$. Let G be a digraph with a vertex partition $\mathcal{P} = \{V_{ij} : i, j \in [k]\}$. Let $i \in [k]$. Suppose that $G[V_{i*}, V_{*i}]$ is a bipartite robust (ν, τ) -expander with bipartition V_{i*}, V_{*i} such that $|V_{i*}| = |V_{*i}| = m$ and $\delta(G[V_{i*}, V_{*i}]) \geq \alpha m$. Let $\phi : V_{i*} \setminus V_{ii} \rightarrow V_{*i} \setminus V_{ii}$ be a bijection. Then there exists a path system $\mathcal{Q}$ in G_{ii} such that $\mathcal{Q} \cup \{\phi(v)v : v \in V_{i*} \setminus V_{ii}\}$ is a cycle with vertex set $V_{i*} \cup V_{*i}$. In the special case that $V_{i*} = V_{*i} = V_{ii}$, we have that ϕ is an empty function and $\mathcal{Q}$ is a cycle (rather than a path system) with vertex set V_{ii} .

Proof. Let H be the digraph on V_{i*} obtained from G_{ii} by identifying v with $\phi(v)$ for all $v \in V_{i*} \setminus V_{ii}$ and deleting any loops. Note that $\delta^0(H) \geq \delta(G[V_{i*}, V_{*i}]) - 1 \geq \alpha m/2$ and H is a robust $(\nu/2, 2\tau)$ -outexpander by Proposition 3.4 applied to the undirected bipartite graph $G[V_{i*}, V_{*i}]$. By Theorem 4.2, H has a Hamilton cycle C . Note that if $V_{i*} = V_{ii} = V_{*i}$, then H coincides with $G[V_{ii}]$, proving the lemma in this case. If not, let $V_{i*} \setminus V_{ii} = \{x_1, \dots, x_\ell\}$ be such that C visits $x_1, \dots, x_\ell$ in that order. Thus, C can be decomposed into paths $P_1, \dots, P_\ell$ such that P_j is a path from x_j to x_{j+1} , where we take $x_{\ell+1} = x_1$. Let Q_j be obtained from P_j by replacing x_{j+1} with $\phi(x_{j+1})$. Note that Q_j is a path in G_{ii} from x_j to $\phi(x_{j+1})$. The result follows by setting $\mathcal{Q} = \{Q_1, \dots, Q_\ell\}$. $\square$

We now prove Lemma 4.1. For a connected component I in $S(\mathcal{P})$, the idea is to apply Lemma 4.3 to each $i \in I$. Some care is needed to ensure that the union of path systems forms only one cycle.

Proof of Lemma 4.1. If I consists of a single vertex say i , then the result follows by Lemma 4.3 (since in that case $V_{i*} = V_{*i} = V_{ii}$). Now assume $|I| \geq 2$, so $V_{i*} \cup V_{*i} \neq V_{ii}$ for all $i \in I$. Without loss of generality, let $I = [\ell]$ and order the indices of I such that for each $j \in [\ell - 1]$, there is a $j' \in [\ell] \setminus [j]$ with $j j' \in E(S(\mathcal{P}))$. This can be achieved since I is connected, (e.g., by taking a reverse breadth-first search ordering). Note that $V_{ij} = V_{ji} = \emptyset$ for all $i \in [\ell]$ and $j \notin [\ell]$.

For $i \in [\ell]$, let $W_i^+ = \bigcup_{i' \in [i], j' \notin [i]} V_{i'j'} = \bigcup_{i' \in [i]} V_{i'*} \cup \bigcup_{i', j' \in [i]} V_{i'j'}$ and $W_i^- = \bigcup_{i' \notin [i], j' \in [i]} V_{i'j'} = \bigcup_{j' \in [i]} V_{*j'} \setminus \bigcup_{i', j' \in [i]} V_{i'j'}$. Since $\mathcal{P}$ is balanced, $|W_i^+| = |W_i^-|$ for each $i \in [\ell]$. Also, by our ordering of the indices, we have that $W_i^+, W_i^- \neq \emptyset$ for $i \in [\ell - 1]$ and that $W_\ell^+ = W_\ell^- = \emptyset$.

Let $\mathcal{Q}_0 = W_0^+ = W_0^- = \emptyset$ and suppose for some $i \in [\ell]$, we have already found a path system $\mathcal{Q}_{i-1}$ such that $V(\mathcal{Q}_{i-1}) = \bigcup_{i' \in [i-1]} (V_{i'*} \cup V_{*i'})$ and $\mathcal{Q}_{i-1}$ consists of precisely $|W_{i-1}^+|$ paths from W_{i-1}^+ to W_{i-1}^- . We now construct $\mathcal{Q}_i$ as follows. Let $\phi : V_{i*} \setminus V_{ii} \rightarrow V_{*i} \setminus V_{ii}$ be any bijection such that if there is a path in $\mathcal{Q}_{i-1}$ from $v_+ \in V_{i*} \setminus V_{ii}$ to $v_- \in V_{*i} \setminus V_{ii}$, then $\phi(v_-) = v_+$. Apply Lemma 4.3 and obtain a path system $\mathcal{Q}'_i$ in G_{ii} such that $\mathcal{Q}'_i \cup \{\phi(v)v : v \in V_{i*} \setminus V_{ii}\}$ is a cycle with vertex set $V_{i*} \cup V_{*i}$. We set $\mathcal{Q}_i = \mathcal{Q}_{i-1} \cup \mathcal{Q}'_i$.

Suppose $i = \ell$. Note that $W_{\ell-1}^- = V_{\ell*} \setminus V_{\ell\ell}$ and $W_{\ell-1}^+ = V_{*\ell} \setminus V_{\ell\ell}$. Hence $\mathcal{Q}_{\ell-1}$ consists of paths from $\phi(v) \in W_{\ell-1}^+$ to $v \in W_{\ell-1}^-$, we have $V(\mathcal{Q}_{\ell-1}) = \bigcup_{i' \in [\ell-1]} (V_{i'*} \cup V_{*i'}) = \bigcup_{i' \in [\ell]} (V_{i'*} \cup V_{*i'}) \setminus V_{\ell\ell}$. Since $\mathcal{Q}'_\ell \cup \{\phi(v)v : v \in V_{\ell*} \setminus V_{\ell\ell}\}$ is a cycle with vertex set $V_{\ell*} \cup V_{*\ell}$, we deduce that $\mathcal{Q}_\ell = \mathcal{Q}'_\ell \cup \mathcal{Q}_{\ell-1}$ is a cycle with vertex set $\bigcup_{i' \in [\ell]} (V_{i'*} \cup V_{*i'})$ as required.

Suppose $i \in [\ell - 1]$. Note that $V(\mathcal{Q}_i) = \bigcup_{i' \in [i]} (V_{i'*} \cup V_{*i'})$. It remains to check that $\mathcal{Q}_i$ is a path system consisting of precisely $|W_i^+|$ paths from W_i^+ to W_i^- . Note first that paths in $\mathcal{Q}'_i \subseteq G_{ii}$ start in $V_{i*} \setminus V_{ii}$ and end in $V_{*i} \setminus V_{ii}$ and have all internal vertices in V_{ii} . On the other hand a path in $\mathcal{Q}_{i-1}$ can only intersect V_{*i} at its start point and V_{i*} at its end point and it avoids V_{ii} . Therefore in $\mathcal{Q}_i = \mathcal{Q}_{i-1} \cup \mathcal{Q}'_i$, all indegrees and outdegrees are at most 1. A vertex has indegree zero in $\mathcal{Q}_{i-1} \cup \mathcal{Q}'_i$ if and only if it is a start point of some path in $\mathcal{Q}_{i-1} \cup \mathcal{Q}'_i$ but not an endpoint of any path in $\mathcal{Q}_{i-1} \cup \mathcal{Q}'_i$, (i.e., the set of vertices of indegree zero is

$$(W_{i-1}^+ \cup (V_{i*} \setminus V_{ii})) \setminus (W_{i-1}^- \cup (V_{*i} \setminus V_{ii})) = (W_{i-1}^+ \cup V_{i*}) \setminus V_{*i} = W_i^+$$

and similarly the set of vertices of outdegree zero is W_i^- . Finally it remains to check that there are no cycles in $\mathcal{Q}_{i-1} \cup \mathcal{Q}'_i$. By our choice of ϕ , if there is a cycle, it spans all the vertices of $V(\mathcal{Q}_{i-1} \cup \mathcal{Q}'_i)$, but this cannot happen because vertices in $W_i^+ \neq \emptyset$ have in-degree zero. $\square$

5. Balanced path systems and path contraction

As mentioned in the previous section, in order to apply Lemma 4.1 to obtain a suitable cycle partition of a dense regular digraph G , we will require a vertex partition $\mathcal{P}$ of G such as that given by Theorem 3.5 but with the additional properties that $\mathcal{P}$ is balanced and $S(\mathcal{P})$ has few connected components. In this section, we state a result (Lemma 5.4) that allows us to adjust a partition $\mathcal{P}$ to have these additional properties. In particular, if $\mathcal{P}$ is not balanced, Lemma 5.4 guarantees us a so-called $\mathcal{P}$ -balanced path system in G whose ‘contraction’ makes $\mathcal{P}$ balanced in a suitable way. At the end of the section we show how Lemma 5.4 implies our main result, Theorem 1.3.

Let $\mathcal{P} = \{V_{ij} : i, j \in [k]\}$ be a partition of a vertex set V and let G be a digraph on V . Recall the definition of G_{ij} and of $\mathcal{P}$ being balanced at the beginning of Section 4. Write $G_{i*} = \bigcup_{j \neq i} G_{ij}$ and $G_{*j} = \bigcup_{i \neq j} G_{ij}$. Let $\mathcal{B}(G, \mathcal{P}) = G - \bigcup_{i \in [k]} G_{ii}$. Note that $\mathcal{B}(G, \mathcal{P}) = \bigcup_{i \in [k]} G_{i*} = \bigcup_{j \in [k]} G_{*j}$. We say that a digraph H on V (where H is usually a path system in G) is $\mathcal{P}$ -balanced if, for all $i \in [k]$,

$$|V_{i*}| - |V_{*i}| = e_H(V_{i*}, V) - e_H(V, V_{*i}) = \sum_{j \in [k]} e(H_{ij}) - \sum_{j \in [k]} e(H_{ji}) = e(H_{i*}) - e(H_{*i}).$$

The above gives three equivalent conditions for H to be $\mathcal{P}$ -balanced. Note that if H is regular, then it is $\mathcal{P}$ -balanced for any $\mathcal{P}$.

Let Q be a path in G from $v_+ \in V_{i_+j_+}$ to $v_- \in V_{i_-j_-}$. We define the Q -contracted subgraph G' of G to be $G \setminus V(Q)$ together with a new vertex w such that $N_{G'}^+(w) = N_G^+(v_-) \setminus V(Q)$ and $N_{G'}^-(w) = N_G^-(v_+) \setminus V(Q)$. We call $\mathcal{P}' = \{V'_{ij} : i, j \in [k]\}$ the Q -contracted partition of $\mathcal{P}$, where

$$V'_{ij} = \begin{cases} (V_{ij} \setminus V(Q)) \cup \{w\} & \text{if } (i, j) = (i_-, j_+); \\ V_{ij} \setminus V(Q) & \text{otherwise.} \end{cases} \quad (5.1)$$

Let $\mathcal{Q}$ be a path system in G . The $\mathcal{Q}$ -contracted subgraph of G (and $\mathcal{Q}$ -contracted partition of $\mathcal{P}$) is obtained by successively contracting each $Q \in \mathcal{Q}$ for G (and $\mathcal{P}$, respectively).

The following two propositions follow from the definition of $\mathcal{Q}$ -contraction. In fact the definitions are chosen precisely so that these propositions hold.

Proposition 5.1. *Let G be a digraph and $\mathcal{Q}$ be a path system in G . Suppose that the $\mathcal{Q}$ -contracted subgraph of G contains a 1-factor with ℓ cycles. Then G also contains a 1-factor containing $\mathcal{Q}$ with ℓ cycles. (The new 1-factor is simply the old 1-factor with the paths in $\mathcal{Q}$ uncontracted.)*

Proposition 5.2. *Let G be a digraph on n vertices and let $\mathcal{P} = \{V_{ij} : i, j \in [k]\}$ be a partition of $V(G)$. Let $\mathcal{Q}$ be a path system in G such that $e(\mathcal{Q}) < |V_{i*}|, |V_{*i}|$ for all $i \in [k]$. Let G' and $\mathcal{P}' = \{V'_{ij} : i, j \in [k]\}$ be the $\mathcal{Q}$ -contracted subgraph of G and $\mathcal{Q}$ -contracted partition of $\mathcal{P}$, respectively. Then for all $i \in [k]$, $\delta(G'[V'_{i*}, V'_{*i}]) \geq \delta(G[V_{i*}, V_{*i}]) - 2e(\mathcal{Q})$.⁴*

Proposition 5.3. *Let $\mathcal{P} = \{V_{ij} : i, j \in [k]\}$ be a partition of a vertex set V and $\mathcal{Q}$ be a $\mathcal{P}$ -balanced path system on V . Then the $\mathcal{Q}$ -contracted partition of $\mathcal{P}$ is balanced.*

Proof. Let $Q \in \mathcal{Q}$ be a path from $v_+ \in V_{i_+j_+}$ to $v_- \in V_{i_-j_-}$. For each $i, j \in [k]$, let $q_{ij} = e(Q_{ij})$ (i.e., the number of edges of Q from V_{i*} to V_{*j} .) Let $\mathcal{P}' = \{V'_{ij} : i, j \in [k]\}$ be the $\mathcal{Q}$ -contracted partition of $\mathcal{P}$. Consider $i \in [k]$. Note that

⁴Note that, in the definition of contraction, the new vertices created are placed in exactly the right parts to ensure the degree condition in the proposition.

$$|V'_{i*}| = |V_{i*}| - |V_{i*} \cap V(Q)| + \mathbb{1}(i = i_-) = |V_{i*}| - \sum_{j \in [k]} q_{ij},$$

$$|V'_{*i}| = |V_{*i}| - |V_{*i} \cap V(Q)| + \mathbb{1}(i = j_+) = |V_{*i}| - \sum_{j \in [k]} q_{ji}.$$

Therefore, $|V'_{i*}| - |V'_{*i}| = |V_{i*}| - |V_{*i}| - \sum_{j \in [k]} (q_{ij} - q_{ji}) = |V_{i*}| - |V_{*i}| - (e(Q_{i*}) - e(Q_{*i}))$. A similar statement holds for the Q -contracted partition $\mathcal{P}^* = \{V_{ij}^* : i, j \in [k]\}$ of $\mathcal{P}$, and the result follows as Q is $\mathcal{P}$ -balanced (i.e., for each $i \in [k]$, $|V_{i*}^*| - |V_{*i}^*| = |V_{i*}| - |V_{*i}| - (e(Q_{i*}) - e(Q_{*i})) = 0$). $\square$

The next lemma says that one can adjust the vertex partition $\mathcal{P}$ found in Theorem 3.5 by a small amount, and find a $\mathcal{P}$ -balanced path system Q such that $S(\mathcal{P}^*)$ has few connected components, where $\mathcal{P}^*$ is the Q -contracted partition of $\mathcal{P}$. It is exactly what we need to prove Theorem 1.3, as we shall see. In Section 6, we break this lemma down into two further lemmas.

Lemma 5.4. *Let $1/n \ll \gamma \ll \alpha, 1/k$. Let G be a d -regular digraph on n vertices with $d \geq \alpha n$ and $\mathcal{P} = \{V_{ij} : i, j \in [k]\}$ be a partition of $V(G)$ such that, for all $i, j \in [k]$,*

- (i) $\delta(G[V_{i*}, V_{*i}]) \geq d/k$;
- (ii) $||V_{i*}| - |V_{*i}|| \leq \gamma n$;
- (iii) $|V_{i*}|, |V_{*i}| \geq d - \gamma n$;
- (iv) *all but at most γn vertices $x \in V_{ij}$ satisfy $d^+(x, V_{*i}), d^-(x, V_{j*}) \geq d - \gamma n$.*

Then there exists a partition $\mathcal{P}' = \{V'_{ij} : i, j \in [k]\}$ of $V(G)$ and a $\mathcal{P}'$ -balanced path system Q in G such that

- (a) *for all $i, j \in [k]$, $|V_{ij} \Delta V'_{ij}| \leq \gamma^{1/2} n$;*
- (b) *for all $i \in [k]$, $\delta(G[V'_{i*}, V'_{*i}]) \geq (d/k) - \gamma^{1/2} n$;*
- (c) $e(Q) \leq \gamma^{1/2} n$;
- (d) $S(\mathcal{P}^*)$ *has at most $n/(qd + 1)$ connected components, where $\mathcal{P}^*$ is the Q -contracted partition of $\mathcal{P}'$, and we set $q = 2$ if G is an oriented graph and $q = 1$ otherwise.*

5.1. Proof of Theorem 1.3

We now prove Theorem 1.3 assuming Lemma 5.4.

Proof of Theorem 1.3. Choose a non-decreasing function $g : (0, 1) \rightarrow (0, 1)$ such that the requirements of Lemmas 4.1 (with $m = \alpha n/2$) and 5.4 are satisfied whenever $n, \gamma, \nu, \tau, \alpha$ satisfy

$$1/n \ll_g \gamma \ll_g \nu \leq \tau \ll_g \alpha,$$

where we write $a \ll_g b$ to mean $a \leq g(b)$. Set $f(x) := \min(x, g(x^2/4))$. Fix $\tau \leq f(\alpha)$ and apply Theorem 3.5 to obtain n_0 such that for any d -regular digraph G on $n \geq n_0$ vertices with $d \geq \alpha n$, there exists γ and ν such that

$$1/n \ll_f \gamma \ll_f \nu \leq \tau \ll_f \alpha$$

and a partition $\mathcal{P} = \{V_{ij} : i, j \in [k]\}$ of $V(G)$ satisfying, for all $i, j \in [k]$,

- (a₁) $G[V_{i*}, V_{*i}]$ is a bipartite robust (ν, τ) -expander with partition V_{i*}, V_{*i} ;
- (a₂) for all $x \in V_{ij}$ and $i', j' \in [k]$, $d^+(x, V_{*i}) \geq d^+(x, V_{*i'})$ and $d^-(x, V_{j*}) \geq d^-(x, V_{j' *})$, so in particular, $\delta(G[V_{i*}, V_{*i}]) \geq d/k$;
- (a₃) $||V_{i*}| - |V_{*i}|| \leq \gamma n$;
- (a₄) all but at most γn vertices $x \in V_{ij}$ satisfy $d^+(x, V_{*i}), d^-(x, V_{j*}) \geq d - \gamma n$;
- (a₅) $k \leq n/(d - \gamma n)$.

Note that (a₅) in particular implies that $\gamma \ll 1/k$. Furthermore, (a₂) and (a₄) together with $\gamma \ll 1/k$ imply that

(a₆) for all $i \in [k]$, $|V_{i*}|, |V_{*i}| \geq d - \gamma n$.

Let $q = 2$ if G is oriented, and $q = 1$ otherwise. Apply Lemma 5.4 and obtain a partition $\mathcal{P}' = \{V'_{ij} : i, j \in [k]\}$ of $V(G)$ and a $\mathcal{P}'$ -balanced path system $\mathcal{Q}$ in G such that

(b₁) for all $i, j \in [k]$, $|V_{ij} \Delta V'_{ij}| \leq \gamma^{1/2}n$;

(b₂) for all $i \in [k]$, $\delta(G[V'_{i*}, V'_{*i}]) \geq (d/k) - \gamma^{1/2}n$;

(b₃) $e(\mathcal{Q}) \leq \gamma^{1/2}n$;

(b₄) $S(\mathcal{P}^*)$ has at most $n/(qd+1)$ connected components, where $\mathcal{P}^*$ is the $\mathcal{Q}$ -contracted partition of $\mathcal{P}'$.

Let $\mathcal{P}^* = \{V^*_{ij} : i, j \in [k]\}$ and let G^* be the $\mathcal{Q}$ -contracted subgraph of G . We will check that, for all $i, j \in [k]$,

(c₁) $\mathcal{P}^*$ is balanced;

(c₂) $|V_{ij} \Delta V^*_{ij}| \leq 5\gamma^{1/2}n$;

(c₃) $|V^*_{i*}|, |V^*_{*i}| \geq d - 6k\gamma^{1/2}n \geq d/2$;

(c₄) $G^*[V^*_{i*}, V^*_{*i}]$ is a bipartite robust $(\nu/2, 2\tau)$ -expander with bipartition V^*_{i*}, V^*_{*i} ;

(c₅) $\delta(G^*[V^*_{i*}, V^*_{*i}]) \geq \alpha^2|V^*_{i*}|/2$.

Note that (c₁) holds by Proposition 5.3. Consider $i, j \in [k]$. Note that

$$|V^*_{ij} \Delta V_{ij}| \leq |V^*_{ij} \Delta V'_{ij}| + |V'_{ij} \Delta V_{ij}| \leq 2|V(\mathcal{Q})| + |V'_{ij} \Delta V_{ij}| \stackrel{(b_1), (b_3)}{\leq} 4\gamma^{1/2}n + \gamma^{1/2}n = 5\gamma^{1/2}n$$

implying (c₂). Hence (c₃) follows from (a₆) and (c₂) as $\gamma \ll \alpha, 1/k$. By (a₁), (c₂), (c₃), and $\gamma \ll \nu$, Lemma 3.1 (applied to the bipartite graph $G^*[V^*_{i*}, V^*_{*i}] \cup G[V_{i*}, V_{*i}]$ with $V_{i*}, V_{*i}, V^*_{i*}, V^*_{*i}$ playing the roles of A, B, A', B' respectively) implies that (c₄) holds. By (a₆), (b₁), and (b₃), we have $e(\mathcal{Q}) \leq \gamma^{1/2}n < d - \gamma n - k\gamma^{1/2}n \leq |V'_{i*}|, |V'_{*i}|$ as $\gamma \ll \alpha, 1/k$. Hence, by Proposition 5.2,

$$\delta(G^*[V^*_{i*}, V^*_{*i}]) \geq \delta(G[V'_{i*}, V'_{*i}]) - 2e(\mathcal{Q}) \stackrel{(b_2), (b_3)}{\geq} (d/k) - 3\gamma^{1/2}n \geq \alpha^2|V^*_{i*}|/2,$$

where the last inequality holds as $\gamma \ll \alpha, 1/k$ and $k(\alpha - \gamma) \leq 1$ by (a₅). Hence (c₅) holds.

Apply Lemma 4.1 (with $G^*, \mathcal{P}^*, d/2, \nu/2, 2\tau, \alpha^2/2$ playing the roles of $G, \mathcal{P}, m, \nu, \tau, \alpha$, respectively) to each connected component I in $S(\mathcal{P}^*)$. Thus, by (b₄), G^* can be covered with at most $n/(qd+1)$ vertex-disjoint cycles (where we have a cycle on $\cup_{i \in I} (V_{i*} \cup V_{*i})$ for each component I of $S(\mathcal{P}^*)$ so that each cycle has length at least $d/2$ by (c₃)). By Proposition 5.1, G can be covered with at most $n/(qd+1)$ vertex-disjoint cycles (each of length at least $d/2$). $\square$

6. Proof of Lemma 5.4

In the last section we reduced the task of proving Theorem 1.3 to the task of proving Lemma 5.4. In this section we reduce this further to the task of proving Lemmas 6.1 and 6.3. We start by introducing the notion of a *non-trivial path system*, which will help us control the number of connected components of $S(\mathcal{P}^*)$ (as required in the conclusion of Lemma 5.4).

Recall the definitions at the beginning of Sections 4 and 5. Let $\mathcal{P} = \{V_{ij} : i, j \in [k]\}$ be a partition of a vertex set V . We call a path system $\mathcal{Q}$ on V *non-trivial* if there exists $i \neq j$ such that $V_{ij} \setminus V(\mathcal{Q}) \neq \emptyset$ or $\mathcal{Q}$ contains a path from V_{*j} to V_{i*} . Otherwise, we call $\mathcal{Q}$ a *trivial* path system. Note that if $\mathcal{Q}$ is non-trivial, then the $\mathcal{Q}$ -contracted partition $\mathcal{P}'$ of $\mathcal{P}$ has the property that $S(\mathcal{P}')$ has at least one edge so that its number of connected components is strictly less than its number of vertices.

Lemma 6.1 below shows that, under the same hypothesis as Lemma 5.4, there exists a $\mathcal{P}$ -balanced path system with few edges, and moreover that this path system is non-trivial if further assumptions are

made. We shall see at the end of the section that the existence of such a non-trivial $\mathcal{P}$ -balanced path system is enough to prove Lemma 5.4. In fact, a $\mathcal{P}$ -balanced path system alone (without the condition of being non-trivial) is enough to prove Lemma 5.4 except in the extremal cases when $d \approx n/k$ if G is digraph and when $d \approx n/2k$ if G is an oriented graph.

Lemma 6.1. *Let $1/n \ll \gamma \ll 1/k, \alpha$. Let G^0 be a d -regular digraph on n vertices with $d \geq \alpha n$ and $\mathcal{P} = \{V_{ij} : i, j \in [k]\}$ be a partition of $V(G^0)$ such that, for all $i \in [k]$,*

- (i) $\delta(G^0[V_{i*}, V_{*i}]) \geq d/k$;
- (ii) $||V_{i*}| - |V_{*i}|| \leq \gamma n$;
- (iii) $|V_{i*}|, |V_{*i}| \geq d - \gamma n$.

Then G^0 contains a $\mathcal{P}$ -balanced path system $\mathcal{Q}$ such that $e(\mathcal{Q}) \leq k^2 \gamma n$. Moreover, if one of the following holds

- (m1) $\sum_{i,j \in [k]: i \neq j} |V_{ij}| > k^2 \gamma n$, or
- (m2) *there exists $v_0 \in V_{i_0 j_0}$ for some $i_0 \neq j_0$ such that $d^+(v_0, V_{*i_0}) - d^-(v_0, V_{i_0*}) \geq 100k^{12} \gamma^{1/3} d$,*

then we may assume that $\mathcal{Q}$ is non-trivial.

In case Lemma 5.4 gives a $\mathcal{P}$ -balanced but trivial path system (i.e., when both (m1) and (m2) fail), we use the next proposition to slightly modify $\mathcal{P}$ and Lemma 6.3 to find the desired non-trivial, $\mathcal{P}$ -balanced path system.

Proposition 6.2. *Let $1/n \ll \gamma \ll 1/k, \alpha$. Let G^0 be a d -regular digraph on n vertices with $d \geq \alpha n$ and $\mathcal{P} = \{V_{ij} : i, j \in [k]\}$ be a partition of $V(G^0)$ such that, for all distinct $i, j \in [k]$,*

- (i) $\delta(G^0[V_{i*}, V_{*i}]) \geq d/k$;
- (ii) $\sum_{i', j' \in [k]: i' \neq j'} |V_{i' j'}| \leq \gamma n$;
- (iii) *for all $v \in V_{ij}$, we have $d^+(v, V_{*i}) - d^-(v, V_{i*}) \leq \gamma n$.*

Then there exists a partition $\mathcal{P}' = \{V'_{ii} : i \in [k]\}$ of $V(G^0)$ such that, for all $i \in [k]$,

- (i') $|V_{ii} \Delta V'_{ii}| \leq \gamma n$;
- (ii') $\delta^0(G^0[V'_{ii}]) \geq d/k - \gamma n$.

Proof. For each $i \in [k]$, let $V'_{ii} = V_{i*}$. Clearly, $\mathcal{P}' = \{V'_{ii} : i \in [k]\}$ is a partition of $V(G^0)$. Note that (ii) implies (i'). For $v \in V_{ij}$ (possibly $i = j$), note that

$$d^+(v, V'_{ii}) = d^+(v, V_{i*}) \stackrel{(ii)}{\geq} d^+(v, V_{ii}) \stackrel{(i)}{\geq} d^+(v, V_{*i}) - \gamma n \geq d/k - \gamma n.$$

If $i \neq j$, similarly, we have

$$d^-(v, V'_{ii}) = d^-(v, V_{i*}) \stackrel{(iii)}{\geq} d^+(v, V_{*i}) - \gamma n \geq d/k - \gamma n.$$

Also, for all $v \in V_{ii}$, (i) implies $d^-(v, V'_{ii}) = d^-(v, V_{i*}) \geq d/k$. Hence (ii') holds. $\square$

Lemma 6.3. *Let $d, k \in \mathbb{N}$ be such that $k \geq 2$ and $d > 165k^5$. Let G be a d -regular digraph on n vertices and $\mathcal{P} = \{V_{ii} : i \in [k]\}$ be a partition of $V(G)$ such that, for all $i \in [k]$,*

- (i) $|V_{ii}| \geq d/2$;
- (ii) $n < (2d + 1)k$, and $n < (d + 1)k$ if G is not oriented;
- (iii) *for all $v \in V_{ii}$, $d^+(v, V_{ii}) + d^-(v, V_{ii}) \geq d/k$.*

Then G contains a non-trivial $\mathcal{P}$ -balanced path system $\mathcal{Q}$ with $e(\mathcal{Q}) \leq k$.

Before we can prove Lemma 5.4, we need a good lower bound in oriented graphs for $|\bigcup_{i \in I} (V_{i*} \cup V_{*i})|$ for any connected component I in $S(\mathcal{P})$; see Proposition 6.5. For this, we use the following result on the minimum semi-degree threshold for an oriented graph to be a robust outexpander. Recall that the definition of robust outexpander is given immediately after Remark 3.3.

Lemma 6.4 [22, Lemma 13.1]. Let $1/n \ll \nu \ll \tau \leq \varepsilon/2 \leq 1$. Then every oriented graph G on n vertices with $\delta^0(G) \geq (3/8 + \varepsilon)n$ is a robust (ν, τ) -outexpander.

Proposition 6.5. Let $1/n \ll \gamma \ll \alpha, 1/k$. Let G be an oriented graph on n vertices with $\Delta^0(G) \leq d$ where $d \geq \alpha n$. Let $\mathcal{P} = \{V_{ij} : i, j \in [k]\}$ be a partition of $V(G)$ such that $|V_{i*}| \geq d/2$ for all $i \in [k]$. Suppose that, for all $i, j \in [k]$, all but at most γn vertices $x \in V_{ij}$ satisfy $d^+(x, V_{*i}), d^-(x, V_{j*}) \geq d - \gamma n$. Let $I \subseteq S(\mathcal{P})$ be a connected component. Then

$$\left| \bigcup_{i \in I} (V_{i*} \cup V_{*i}) \right| \geq \begin{cases} 2(d - \gamma n) & \text{if } |I| = 1, \\ 9(d - 5\gamma n)/4 & \text{otherwise.} \end{cases}$$

Proof. Without loss of generality, let $I = [\ell]$. For each $i \in [\ell]$, we can find vertices $v_i \in V_{i*}$ such that $d^+(v_i, V_{*i}) \geq d - \gamma n$ as $\gamma \ll \alpha, 1/k$, which shows in particular that $|V_{*i}| \geq d - \gamma n$. Hence, if $\ell \geq 3$,

$$\left| \bigcup_{i \in I} (V_{i*} \cup V_{*i}) \right| \geq \sum_{i \in [\ell]} |V_{*i}| \geq \ell(d - \gamma n) \geq 3(d - \gamma n) \geq 9(d - 5\gamma n)/4.$$

If $\ell = 1$, then $V_{11} = V_{1*} = V_{*1}$. There exists a vertex $v \in V_{11}$ such that $d^+(v, V_{11}), d^-(v, V_{11}) \geq d - \gamma n$. Since G is an oriented graph, we have $|V_{11}| \geq d^+(v, V_{11}) + d^-(v, V_{11}) \geq 2(d - \gamma n)$.

If $\ell = 2$, then $\bigcup_{i \in [2]} V_{i*} = \bigcup_{i \in [2]} V_{*i}$ and $V_{ij} = V_{ji} = \emptyset$ for all $(i, j) \in [2] \times ([k] \setminus [2])$. Let $V_I = \bigcup_{i \in [2]} V_{i*} = \bigcup_{i \in [2]} V_{*i}$. There exists a vertex subset $W \subseteq V_I$ such that

$$|W| \geq |V_I| - 4\gamma n \geq 2(d - \gamma n) - 4\gamma n \geq \alpha n,$$

and such that, writing $W_{ij} = W \cap V_{ij}$ for $i, j \in [2]$, we have for all $i, j \in [2]$ that

$$d^+(x, W_{*i}), d^-(y, W_{j*}) \geq d - 5\gamma n \text{ for all } x \in W_{i*} \text{ and } y \in W_{*j}. \quad (6.1)$$

Hence $\delta^0(G[W]) \geq d - 5\gamma n$.

Let τ be a constant with $\gamma \ll \tau \ll \alpha$. Next we show that $G[W]$ is not a robust $(\gamma^{1/3}, \tau)$ -outexpander. For $i \in [2]$, (6.1) implies that $|W_{i*}| \geq d - 5\gamma n \geq \tau n \geq \tau|W|$ and so $\tau|W| \leq |W_{i*}| \leq (1 - \tau)|W|$. Since $W_{1*} \cup W_{2*} = W_{*1} \cup W_{*2}$, we may assume without loss of generality that $|W_{1*}| \geq |W_{*1}|$. Recall that $\Delta^0(G) \leq d$. By (6.1), each vertex in $|W_{1*}|$ has at most $5\gamma n$ outneighbours in W_{*2} , so

$$|\text{RN}_{\gamma^{1/3}}^+(W_{1*}) \cap W_{*2}| \leq \frac{e_G(W_{1*}, W_{*2})}{\gamma^{1/3}|W|} \leq \frac{5\gamma n^2}{\alpha \gamma^{1/3} n} < \gamma^{1/3} \alpha n \leq \gamma^{1/3} |W|,$$

where the penultimate inequality holds as $\gamma \ll \alpha$. This implies that

$$|\text{RN}_{\gamma^{1/3}}^+(W_{1*})| \leq |W_{*1}| + |\text{RN}_{\gamma^{1/3}}^+(W_{1*}) \cap W_{*2}| < |W_{1*}| + \gamma^{1/3} |W|.$$

Hence $G[W]$ is not a robust $(\gamma^{1/3}, \tau)$ -outexpander as claimed. Lemma 6.4 with $\varepsilon = 5/72$ implies $\delta^0(G[W]) \leq (3/8 + \varepsilon)|W| = 4|W|/9$. Then $4|W|/9 \geq d - 5\gamma n$, and the result follows. $\square$

We now prove Lemma 5.4 assuming Lemmas 6.1 and 6.3.

Proof of Lemma 5.4. Let G and $\mathcal{P} = \{V_{ij} : i, j \in [k]\}$ be as in the statement of Lemma 5.4. We apply Lemma 6.1 (with $G^0 = G$) and obtain a $\mathcal{P}$ -balanced path system $\mathcal{Q}^0$ such that $e(\mathcal{Q}^0) \leq k^2 \gamma n$. Let $\mathcal{P}^0$ be the $\mathcal{Q}^0$ -contracted partition of $\mathcal{P}$. Let k^* be the smallest integer larger than $n/(qd + 1)$, that is, $k^* = \lfloor n/(qd + 1) \rfloor + 1$. If $S(\mathcal{P}^0)$ has at most $n/(qd + 1)$ connected components, then we are done by setting $\mathcal{P}' = \mathcal{P}$ and $\mathcal{Q} = \mathcal{Q}^0$ since $e(\mathcal{Q}^0) \leq k^2 \gamma n \leq \gamma^{1/2} n$. Hence we may assume that $S(\mathcal{P}^0)$ has at least k^* connected components.

Claim 6.6. We have $k^* = k$ and $S(\mathcal{P}^0)$ is an empty graph on $[k^*]$.

Proof of claim. Since $1/n \ll \gamma \ll \alpha$ and $d \geq \alpha n$, we have that

$$\frac{n}{q(d-\gamma n)} - \frac{n}{qd+1} \leq \frac{1}{10} \text{ and so } \left\lfloor \frac{n}{q(d-\gamma n)} \right\rfloor \leq \left\lfloor \frac{n}{qd+1} \right\rfloor + 1 = k^*. \quad (6.2)$$

First suppose that G is not oriented, so $q = 1$. Recall (iii) that $|V_{i*}| \geq d - \gamma n$ for all $i \in [k]$, so $n \geq k(d - \gamma n)$. Together with (6.2), we have $k \leq \lfloor n/(d - \gamma n) \rfloor \leq k^*$. On the other hand, since $S(\mathcal{P}^0)$ has at least k^* connected components, we have $k = |V(S(\mathcal{P}^0))| \geq k^*$. Thus $k = k^*$. Furthermore, $S(\mathcal{P}^0)$ is an empty graph on $[k^*]$ (or else $S(\mathcal{P}^0)$ would have fewer than k^* connected components).

Next, suppose that G is oriented, so $q = 2$. Let $I_1, \dots, I_{a+b}$ be connected components of $S(\mathcal{P}^0)$ such that $|I_j| = 1$ if and only if $j \in [a]$. Note that $a + b \geq k^*$ and $\{\bigcup_{i \in I_j} (V_{i*} \cup V_{*i}) : j \in [a + b]\}$ is a partition of $V(G)$. Recall (iv) that for all $i, j \in [k]$, all but at most γn vertices $x \in V_{ij}$ satisfy $d^+(x, V_{*i}), d^-(x, V_{j*}) \geq d - \gamma n$. Also, we have that $|V_{i*}| \geq d - \gamma n \geq d/2$ for all $i \in [k]$ as $\gamma \ll \alpha$. By Proposition 6.5, we have

$$n = \sum_{j \in [a+b]} \left| \bigcup_{i \in I_j} (V_{i*} \cup V_{*i}) \right| \geq 2(d - \gamma n)a + \frac{9(d - 5\gamma n)}{4}b = 2(d - \gamma n)(a + b) + \frac{d - 37\gamma n}{4}b.$$

Hence (6.2) implies that

$$\frac{1}{10} \geq \frac{n}{2(d - \gamma n)} - \frac{n}{2d + 1} > \frac{n}{2(d - \gamma n)} - k^* \geq (a + b - k^*) + \frac{b(d - 37\gamma n)}{8(d - \gamma n)} > (a + b - k^*) + \frac{b}{9},$$

where the last inequality holds as $\gamma \ll \alpha$. Recalling that $a + b \geq k^*$, this shows that $a + b = k$ and $b = 0$, and so $a = k^*$, proving the claim. $\square$

Since $S(\mathcal{P}^0)$ is an empty graph on $[k^*]$, we deduce that $\mathcal{Q}^0$ is trivial as defined at the start of the section. By the moreover statement of Lemma 6.1, we have

- (m1) $\sum_{i,j \in [k]: i \neq j} |V_{ij}| \leq k^2 \gamma n \leq \gamma^{1/4} n$ and
- (m2) for all distinct $i, j \in [k]$ and $v \in V_{ij}$, we have $d^+(v, V_{*i}) - d^-(v, V_{i*}) \leq 100k^{12}\gamma^{1/3}d \leq \gamma^{1/4}n$.

By Proposition 6.2 (with $\gamma^{1/4}$ playing the role of γ), there exists a partition $\mathcal{P}' = \{V'_{ii} : i \in [k]\}$ of $V(G)$ (and we take $V'_{ij} = \emptyset$ for $i \neq j$) such that, for all $i \in [k]$,

- (i') $|V_{ii} \Delta V'_{ii}| \leq \gamma^{1/4}n$;
- (ii') $\delta^0(G[V'_{ii}]) \geq (d/k) - \gamma^{1/4}n$.

Finally, apply Lemma 6.3 (with $\mathcal{P}'$ playing the role of $\mathcal{P}$) by noting that property (i) of Lemma 6.3 holds by property (iii) of Lemma 5.4 together with (m1) and (i') above, and properties (ii) and (iii) of Lemma 6.3 hold by Claim 6.6 and (ii') above, respectively. Thus, in G , we obtain a non-trivial $\mathcal{P}'$ -balanced path system $\mathcal{Q}$ with $e(\mathcal{Q}) \leq k \leq \gamma^{1/2}n$. Also, if $\mathcal{P}^*$ is the $\mathcal{Q}$ -contracted partition of $\mathcal{P}'$, then by Claim 6.6, $S(\mathcal{P}^*)$ has at most $k - 1 = k^* - 1 \leq n/(qd + 1)$ connected components since $\mathcal{Q}$ is non-trivial. This gives (c) and (d) in Lemma 5.4, while (m1) and (i') imply (a), and (ii') implies (b). $\square$

7. Proof of Lemma 6.1

Recall that in Lemma 6.1, we are given a d -regular digraph G^0 on n vertices together with a partition $\mathcal{P} = \{V_{ij} : i, j \in [k]\}$ of $V(G^0)$ satisfying certain size and degree conditions, and we must find a $\mathcal{P}$ -balanced path system $\mathcal{Q}$ with few edges. Moreover, we require that $\mathcal{Q}$ is non-trivial under certain circumstances. We begin with a sketch of our proof approach for this lemma (ignoring the moreover part for now). Recall the definitions at the start of Sections 4 and 5.

Let V^+ and V^- be two disjoint copies of $V(G^0)$. Let $V_{1*}^+, \dots, V_{k*}^+$ to be the partition of V^+ corresponding to $V_{1*}, \dots, V_{k*}$ in $V(G^0)$. Similarly, let $V_{*1}^-, \dots, V_{*k}^-$ to be the partition of V^- corresponding to

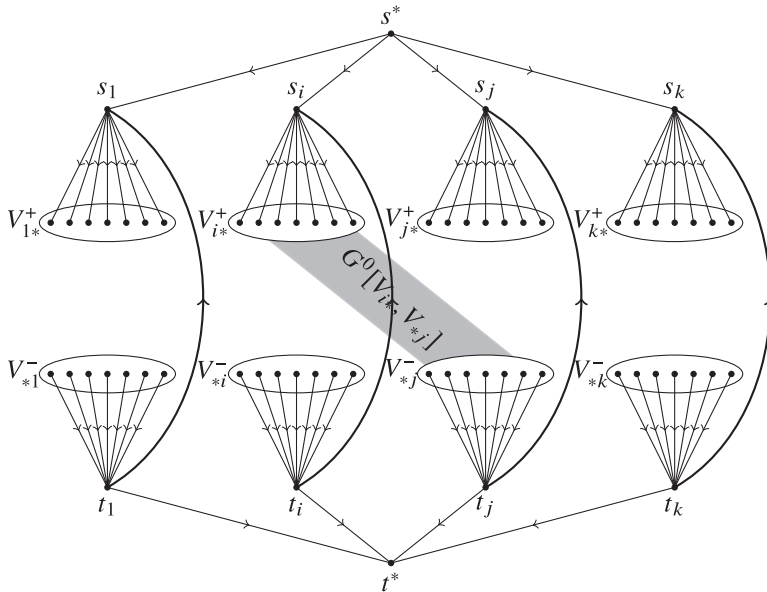


Figure 1. The network $\mathcal{F}^*$.

$V_{*1}, \dots, V_{*k}$ in $V(G^0)$. Let H be the bipartite digraph with bipartition V^+, V^- such that for each $x \in V_{i*}^+$ and $y \in V_{j*}^-$, we have $xy \in E(H)$ if and only if $xy \in E(G^0)$ and $i \neq j$. So there is a bijection between $E(H)$ and $E(\mathcal{B}(G^0, \mathcal{P}))$.

A naive approach is to find, for each $i, j \in [k]$ with $i \neq j$, a matching M_{ij} in $H[V_{i*}^+, V_{j*}^-]$ with $e(M_{ij}) = e(G_{ij}^0)/d$. If Q_{ij} is the subgraph in G^0 corresponding to M_{ij} , then $Q = \bigcup_{i,j \in [k]} Q_{ij}$ is $\mathcal{P}$ -balanced by construction since G^0 is d -regular. It is relatively easy to guarantee Q has few edges by first passing from G^0 to a suitable subdigraph G (see Proposition 7.6). However, there are three problems with this approach: (i) $\Delta^0(Q)$ might be greater than one (meaning that Q is not a path system), (ii) $e(G_{ij}^0)/d$ may not be an integer, and (iii) Q may contain cycles, even if $\Delta^0(Q) \leq 1$.

To overcome (i) and (ii), we consider a suitable flow problem by converting H to a network $\mathcal{F}^* = (F^*, w, s^*, t^*)$ as follows (the formal definition of a network and flow are stated in Section 7.1). Starting with the graph H (with all edges of H having capacity 1), we add new vertices $s^*, s_1, \dots, s_k, t^*, t_1, \dots, t_k$, where s^* and t^* are the source and sink respectively and where the s_i and t_j are viewed as ‘local’ sources and sinks, respectively. For each $i, j \in [k]$ and each $x^+ \in V_{i*}^+, y^- \in V_{j*}^-$, we add the edges $s_i x^+$ and $y^- t_j$, each of capacity 1. For each $i \in [k]$, we add the edge $s^* s_i$ of capacity $\max\{|V_{i*}^+| - |V_{i*}^-|, 0\}$, the edge $t_i t^*$ of capacity $\max\{|V_{i*}^-| - |V_{i*}^+|, 0\}$ and the edge $t_i s_i$ of infinite capacity. This gives the network $\mathcal{F}^*$; see Figure 1. Consider a maximum integer flow f for $\mathcal{F}^*$. Define Q to be the subdigraph of $\mathcal{B}(G^0, \mathcal{P})$ such that $xy \in E(Q)$ if and only if $x^+ y^-$ has a flow of one in f , where x^+ and y^- are the corresponding copies of x and y in V^+ and V^- , respectively. Note that $\Delta^0(Q) \leq 1$ by construction. Also it is easy to check that Q is $\mathcal{P}$ -balanced provided all edges at s^* (and hence also at t^*) are saturated by f .

To guarantee this latter condition on f , it turns out that it is enough to find a fractional flow f^* such that f_{ij}^* , the total amount of flow in f^* through the edges in $H[V_{i*}, V_{j*}]$, is roughly $e(G_{ij}^0)/d$ for all $i \neq j$. Such a fractional flow is almost a maximum flow, and we can use the max-flow min-cut theorem to convert this fractional flow into an integer flow that saturates the edges at s^* (and t^*); see Claim 7.10. This also addresses (ii) above.

To deal with (iii), suppose we could find a matching M^0 of G^0 such that $e(M^0[V_{i*}, V_{j*}]) \geq e(G_{ij}^0)/d$ for all $i, j \in [k]$. Then defining H (and $\mathcal{F}^*$) using M^0 instead of G^0 , we could apply the same argument

as before to obtain a $\mathcal{P}$ -balanced subgraph $Q \subseteq M^0$ (which is therefore necessarily a path system as required). However, it is not always possible to find such a matching, so instead, for each i, j , we find a large ‘extendable matching’ in G_{ij}^0 which consists of a matching $M_{ij} \subseteq G_{ij}^0$ together with sets of vertices $X_{ij}^+ \subseteq V_{i*}$ and $X_{ij}^- \subseteq V_{*j}$ where each vertex $v^+ \in X_{ij}^+$ has high outdegree in G_{ij}^0 and each vertex $v^- \in X_{ij}^-$ has high indegree in G_{ij}^0 , and where $\bigcup_{i,j} M_{ij}$ is a matching (see Lemma 7.8). Here the very large degrees of vertices in $X_{ij} = X_{ij}^+ \cup X_{ij}^-$ allow us to greedily extend M_{ij} into a suitably large path system⁵ (see Proposition 7.9). These path systems would be disjoint (as required) if the X_{ij} are disjoint, but a priori the X_{ij} will not be disjoint. Therefore, we must modify the X_{ij} so that each vertex is assigned to at most one X_{ij}^+ and at most one X_{ij}^- . It turns out that this assignment problem is not too hard to incorporate into the flow problem described earlier (we simply add an edge of capacity 1 from $v^+ \in V^+$ to t_j if $v^+ \in X_{ij}^+$ and similarly if $v^- \in X_{ij}^-$). Solving the flow problem gives us the *disjoint* ‘extendable matchings’ we seek (meaning the X_{ij} ’s are disjoint), which can greedily be extended to give the desired path system.

7.1. Flows

We recall some common definitions and facts about flow networks. We note that flows are only used in the proof of Lemma 6.1 (and not in any preliminary results).

A *flow network* is a tuple $\mathcal{F} = (F, w, S, T)$, where F is a digraph, $w: E(F) \rightarrow \mathbb{R}_{\geq 0}$ is the *capacity* function, and $S \subset V$ is a set of *sources* (i.e., each $s \in S$ only has outedges incident to it) and $T \subset V$ is a set of *sinks* (i.e., each $t \in T$ only has inedges incident to it). A *flow* for the flow network $\mathcal{F}$ is a function $f: E(F) \rightarrow \mathbb{R}_{\geq 0}$ such that, for all $e \in E$, we have $f(e) \leq w(e)$ and, for all $v \in V \setminus (S \cup T)$, we have $\sum_{u \in N_F^-(v)} f(uv) = \sum_{u \in N_F^+(v)} f(vu)$. We define the *value* of f as $\text{val}(f) := \sum_{s \in S} \sum_{v \in N_F^+(s)} f(sv) = \sum_{t \in T} \sum_{u \in N_F^-(t)} f(ut)$. A *maximum flow* in a given flow network is a flow f that maximises $\text{val}(f)$. We say that f is an *integer flow* if $f(e)$ is an integer for all $e \in E$, and to emphasize the contrast, we sometimes refer to a flow that is not necessarily an integer flow as a *fractional flow*. For any set of edges $E' \subseteq E$, we write $f(E') := \sum_{e \in E'} f(e)$. If $S = \{s\}$ and $T = \{t\}$, then we simply write (F, w, s, t) .

There are variants of the above notions that we will use. In particular, as well as having edge capacities, a flow network can also have vertex capacities (which restrict the amount of flow that can pass through a vertex) so that now $w: E(F) \cup V(F) \rightarrow \mathbb{R}^+$, and for each $v \in V(F) \setminus (S \cup T)$ a flow is defined as before with the added restriction that $f(v) \leq w(v)$, where $f(v) := \sum_{u \in N_F^-(v)} f(uv) = \sum_{u \in N_F^+(v)} f(vu)$. One can easily reduce this to the situation of just edge capacities by replacing each vertex v with a directed edge v^-v^+ of capacity $w(v)$ and each directed edge uv (or vw) with uv^- (or v^+w , respectively) where new edges inherit the capacities of their old counterparts. We say that a flow f *saturates* an edge e (or a vertex v) if $f(e) = w(e)$ (or $f(v) = w(v)$, respectively).

Let (F, w, s, t) be a flow network. For a partition (U, W) of $V(F)$ with $s \in U$, $t \in W$ the edge set $E_F(U, W)$ is called a *cut*. (Recall that $E_F(U, W)$ is the set of edges in F from U to W .) The *capacity* of a cut $E_F(U, W)$ is the sum of the capacities of its edges, (i.e., $w(E_F(U, W)) := \sum_{e \in E_F(U, W)} w(e)$.) A *minimum cut* of the given flow network is a cut of minimum capacity. We make use of the following well-known theorem.

Theorem 7.1 (Max-flow min-cut [4]). *Let (F, w, s, t) be a flow network.*

- (i) *If f is a flow and $E_F(U, W)$ is a cut then $\text{val}(f) \leq w(E_F(U, W))$.*
- (ii) *We have that f is a maximum flow and $E_F(U, W)$ is a minimum cut if and only if $\text{val}(f) = w(E_F(U, W))$ and in that case f saturates every edge in $E_F(U, W)$.*
- (iii) *If all capacities are non-negative integers, [i.e., $w(e) \in \mathbb{Z}_{\geq 0}$ for all $e \in E(F)$], then there exists a maximum flow f that is an integer flow.*

We also use the following whose proof is omitted as it is straightforward.

⁵When doing the greedy extension we obtain a path system (with paths of length at most 3) rather than a matching because we allow the start (resp. end) point of an edge in M_{ij} to be in X_{ij}^- (resp. X_{ij}^+).

Proposition 7.2. Let $\mathcal{F} = (F, w, s, t)$ be a flow network and f a flow for $\mathcal{F}$. Then, for any edge $e \in E(F)$, there exists a flow f' for $\mathcal{F}$ such that $f'(e) = 0$ and $\text{val}(f') \geq \text{val}(f) - f(e)$.

7.2. Preliminaries

We will require Vizing's theorem for multigraphs in the proof of Lemma 7.8. Let H be an (undirected) multigraph (without loops). The *multiplicity* $\mu(H)$ of H is the maximum number of edges between two vertices of H , and, as usual, $\Delta(H)$ is the maximum degree of H . A *proper k -edge-colouring* of H is an assignment of k colours to the edges of H such that incident edges receive different colours.

Theorem 7.3 ([34]; see e.g., [5]). Any multigraph H has a proper k -edge colouring with $k = \Delta(H) + \mu(H)$ colours. In particular, by taking the largest colour class, there is a matching in H of size at least $e(H)/(\Delta(H) + \mu(H))$.

We will also require the following simple proposition about decomposing acyclic digraphs into paths.

Proposition 7.4 [26, Proposition 2.6]. Let G be an acyclic digraph. Then the edges of G can be partitioned into $\sum_{v \in V(G)} |d^+(v) - d^-(v)|/2$ directed paths.

We will need the following property of regular digraphs; its proof is a simple exercise by considering $\sum_{v \in V_{i*}} d^+(v)$ and $\sum_{v \in V_{*i}} d^-(v)$.

Proposition 7.5 [27, Proposition 3.2]. Let G be a d -regular digraph, $k \in \mathbb{N}$, and $\mathcal{P} = \{V_{ij} : i, j \in [k]\}$ be a partition of $V(G)$. Then, for all $i \in [k]$, we have $d(|V_{i*}| - |V_{*i}|) = e(G_{i*}) - e(G_{*i})$.

The next proposition shows that we can consider a simpler subgraph G in $\mathcal{B}(G^0, \mathcal{P})$.

Proposition 7.6. Let G^0 be a digraph, $k, d \in \mathbb{N}$ and $\mathcal{P} = \{V_{ij} : i, j \in [k]\}$ be a partition of $V(G^0)$ such that for all $i \in [k]$, $e(G_{i*}^0) - e(G_{*i}^0) = d(|V_{i*}| - |V_{*i}|)$. Then, by reordering $[k]$ if necessary, there exists a subdigraph G of $\mathcal{B}(G^0, \mathcal{P})$ such that

$$e(G) \leq d(k-1) \sum_{i \in [k]} ||V_{i*}| - |V_{*i}||/2,$$

$$e(G_{i*}) - e(G_{*i}) = d(|V_{i*}| - |V_{*i}|) \text{ for all } i \in [k],$$

and for all $i, j \in [k]$ with $i \leq j$, $e(G_{ji}) = 0$.

Proof. Define an auxiliary multidigraph H^0 on $[k]$ (with loops) such that there are precisely $e(G_{ij}^0)$ many ij edges in H^0 . There is a natural bijection between $E(G^0)$ and $E(H^0)$. Let H be an acyclic subdigraph of H^0 obtained by successively removing all edges of a directed cycle where we treat loops also as cycles. Note that, for all $i \in [k]$,

$$d_H^+(i) - d_H^-(i) = d_{H^0}^+(i) - d_{H^0}^-(i) = e(G_{i*}^0) - e(G_{*i}^0) = d(|V_{i*}| - |V_{*i}|). \quad (7.1)$$

Since H is acyclic, by relabelling if necessary, we may assume that there are no edges ji with $i \leq j$. By Proposition 7.4, $E(H)$ can be decomposed into $\sum_{i \in [k]} |d_H^+(i) - d_H^-(i)|/2$ directed paths. Since each path can have length at most $k-1$ (as H has k vertices), we have

$$e(H) \leq (k-1) \sum_{i \in [k]} |d_H^+(i) - d_H^-(i)|/2 \stackrel{(7.1)}{=} d(k-1) \sum_{i \in [k]} ||V_{i*}| - |V_{*i}||/2.$$

The result follows by setting G to be the subdigraph of G^0 corresponding to H . □

We need the following lemma from our previous work [27]. It states that given a set of matchings of low total maximum degree, one can select a relatively large number of the edges from each matching so that the union of selected edges is also a matching.

Lemma 7.7 [27, Lemma 4.2]. Let $k, \ell \in \mathbb{N}$ and $M_1, M_2, \dots, M_\ell$ be matchings with $\Delta(\bigcup_{i \in [\ell]} M_i) \leq k$. Suppose $e(M_i) > 8k^6 \ln \ell$ for all $i \in [\ell]$. Then, there exists a matching $M \subseteq \bigcup_{i \in [\ell]} M_i$ with $|M \cap M_i| \geq e(M_i)/2k^2$ for all $i \in [\ell]$.

Any graph (or digraph) either has a large matching or has all its edges incident to a small set of vertices (so these vertices have relatively large degree). The following lemma allows us to interpolate between these extremes and moreover does it simultaneously for all G_{ij} . This corresponds to the ‘extendable matchings’ described in the sketch of proof.

Lemma 7.8. Let G be a digraph, $k, d \in \mathbb{N}$, and $\mathcal{P} = \{V_{ij} : i, j \in [k]\}$ be a partition of $V(G)$. Let $\theta \in (0, 1)$ with $\theta d \geq 2$. For all $i, j \in [k]$ and $v \in V(G)$, let $X_{ij}^+ = \{v \in V(G) : d_{G_{ij}}^+(v) \geq \theta d\}$ and define X_{ij}^- similarly. Then, for each $i, j \in [k]$, there exists a matching M_{ij} of G_{ij} such that

- (1) $e(G_{ij}) \leq \sum_{x^+ \in X_{ij}^+} d_{G_{ij}}^+(x^+) + \sum_{x^- \in X_{ij}^-} d_{G_{ij}}^-(x^-) + (6k^2 \theta d)e(M_{ij}) + 50k^{10} \theta d$;
- (2) for all $uv \in M_{ij}$, $u \notin X_{ij}^+$ and $v \notin X_{ij}^-$;
- (3) $\bigcup_{i,j \in [k]} M_{ij}$ is a matching.

Proof. Consider $i, j \in [k]$. Let H_{ij} be the multigraph obtained from G_{ij} by deleting all the edges uv with $u \in X_{ij}^+$ or $v \in X_{ij}^-$, and by making all the edges undirected. Note that we have $\Delta(H_{ij}) + \mu(H_{ij}) \leq 2\theta d + 2$ and

$$e(H_{ij}) \geq e(G_{ij}) - \left(\sum_{x^+ \in X_{ij}^+} d_{G_{ij}}^+(x^+) + \sum_{x^- \in X_{ij}^-} d_{G_{ij}}^-(x^-) \right). \quad (7.2)$$

Then, by Vizing’s theorem for multigraphs (Theorem 7.3), there exists a matching M_{ij}^H in H_{ij} of size at least $e(H_{ij})/(2\theta d + 2) \geq e(H_{ij})/3\theta d$. Let M_{ij}^0 be the corresponding matching in G_{ij} . If $e(M_{ij}^0) \leq 16k^{10}$, then we set M_{ij}^0 to be empty. Thus, together with (7.2) we have

$$e(G_{ij}) \leq \left(\sum_{x^+ \in X_{ij}^+} d_{G_{ij}}^+(x^+) + \sum_{x^- \in X_{ij}^-} d_{G_{ij}}^-(x^-) \right) + (3\theta d)e(M_{ij}^0) + 50k^{10} \theta d.$$

Observe that $\Delta(\bigcup_{i,j \in [k]} M_{ij}^0) \leq k$. Apply Lemma 7.7 for nonempty matchings M_{ij}^0 (with $\ell \leq k^2$) to obtain $M_{ij} \subseteq M_{ij}^0$ (set $M_{ij} = \emptyset$ if $M_{ij}^0 = \emptyset$) such that $\bigcup_{i,j \in [k]} M_{ij}$ is a matching and, for all $i, j \in [k]$, $e(M_{ij}) \geq e(M_{ij}^0)/2k^2$. The result follows. $\square$

Recall that for any directed matching M , $V^+(M)$ and $V^-(M)$ are the sets of starting and ending vertices of the directed edges in M , respectively. Formally, $V^+(M) = \{v \in V(M) : vw \in E(M) \text{ for some } w \in V(M)\}$ and similarly for $V^-(M)$.

Proposition 7.9. Let G be a digraph, $k \in \mathbb{N}$, and $\mathcal{P} = \{V_{ij} : i, j \in [k]\}$ be a partition of $V(G)$. Let $W \subseteq V(G)$. Suppose that, for each $i, j \in [k]$ with $i \neq j$, there exist $Y_{ij}^+ \subseteq V_{i*}$, $Y_{ij}^- \subseteq V_{*j}$ and $M_{ij} \subseteq G_{ij}$ such that

- (a) $M = \bigcup_{i,j \in [k]} M_{ij}$ is a matching;
- (b) both of $\{Y_{ij}^+, V^+(M_{ij}) : i, j \in [k]\}$ and $\{Y_{ij}^-, V^-(M_{ij}) : i, j \in [k]\}$ are sets of disjoint sets;
- (c) for all $y^+ \in Y_{ij}^+$ and $y^- \in Y_{ij}^-$, $d_{G_{ij}}^+(y^+) \geq 2 \sum_{i,j \in [k]} (|Y_{ij}^+| + |Y_{ij}^-| + e(M_{ij})) + |W| + 1$.

Then $\mathcal{B}(G, \mathcal{P})$ contains a path system $\mathcal{Q}$ such that the following hold:

- (i) $e(\mathcal{Q}_{ij}) = |Y_{ij}^+| + |Y_{ij}^-| + e(M_{ij})$ for all $i, j \in [k]$, where $\mathcal{Q}_{ij} = \mathcal{Q} \cap E(G_{ij})$;
- (ii) $V(\mathcal{Q}) \cap W \subseteq \bigcup_{i,j \in [k]} (Y_{ij}^+ \cup Y_{ij}^- \cup V(M_{ij}))$;

(iii) if $y^+ \in Y_{ij}^+ \setminus V^-(M)$, then there exists $u \in V_{*j} \setminus W$ such that the single edge y^+u is a path in $\mathcal{Q}$, and a similar statement for $y^- \in Y_{ij}^- \setminus V^+(M)$ holds.

Proof. Start by setting $\mathcal{Q} = M$ and then for each $y \in Y_{ij}^+$ (and $y \in Y_{ij}^-$), greedily add an edge yv (and vy , respectively) in $G_{ij} \setminus W$ such that v avoids all current vertices in $\mathcal{Q}$ and all vertices in $\bigcup_{i,j} (Y_{ij}^+ \cup Y_{ij}^-)$, which is possible by (c). It is clear that we always maintain a path system, and that (i)–(iii) hold by construction. $\square$

7.3. Proof of Lemma 6.1

We now prove Lemma 6.1.

Proof of Lemma 6.1. We split the proof into several steps.

Step 1: Defining G . By Proposition 7.5, $e(G_{i*}^0) - e(G_{*i}^0) = d(|V_{i*}| - |V_{*i}|)$ for all $i \in [k]$. Apply Proposition 7.6 and, without loss of generality, obtain a subdigraph G of $\mathcal{B}(G^0, \mathcal{P})$ such that

$$e(G) \leq (k-1)d \sum_{i \in [k]} \frac{||V_{i*}| - |V_{*i}||}{2} \stackrel{(ii)}{\leq} k^2 \gamma dn. \quad (7.3)$$

$$e(G_{i*}) - e(G_{*i}) = d(|V_{i*}| - |V_{*i}|) \text{ for all } i \in [k], \quad (7.4)$$

and, for all $i, j \in [k]$ with $i \leq j$, $e(G_{ji}) = 0$.

Step 2: Finding X_{ij}^+ , X_{ij}^- and M_{ij} . For all $i, j \in [k]$, let

$$X_{ij}^+ = \{v \in V(G) : d_{G_{ij}}^+(v) \geq \gamma^{1/3} d\} \text{ and } X_{ij}^- = \{v \in V(G) : d_{G_{ij}}^-(v) \geq \gamma^{1/3} d\}. \quad (7.5)$$

Observe that $X_{ij}^+ \subseteq V_{i*}$ and $X_{ij}^- \subseteq V_{*j}$ for all $i, j \in [k]$. Note that $\gamma^{1/3} d \geq \gamma^{1/3} \alpha n \geq 2$ as $1/n \ll \gamma \ll \alpha$. Apply Lemma 7.8 to G (with $\theta = \gamma^{1/3}$) and obtain a matching M_{ij} of G_{ij} for $i, j \in [k]$, such that

- (i') $e(G_{ij}) \leq \sum_{x^+ \in X_{ij}^+} d_{G_{ij}}^+(x^+) + \sum_{x^- \in X_{ij}^-} d_{G_{ij}}^-(x^-) + (6k^2 \gamma^{1/3} d)e(M_{ij}) + 50k^{10} \gamma^{1/3} d$;
- (ii') for all $uv \in M_{ij}$, $u \notin X_{ij}^+$ and $v \notin X_{ij}^-$;
- (iii') $\bigcup_{i,j \in [k]} M_{ij}$ is a matching.

By deleting vertices in $X_{ij}^+ \cup X_{ij}^-$ and edges of M_{ij} if necessary, we may assume that the RHS of (i') is bounded above by $e(G_{ij}) + d$. Hence

$$\begin{aligned} \gamma^{1/3} d(|X_{ij}^+| + |X_{ij}^-| + e(M_{ij})) &\leq \sum_{x^+ \in X_{ij}^+} d_{G_{ij}}^+(x^+) + \sum_{x^- \in X_{ij}^-} d_{G_{ij}}^-(x^-) + (6k^2 \gamma^{1/3} d)e(M_{ij}) + 50k^{10} \gamma^{1/3} d \\ &\leq e(G_{ij}) + d \stackrel{(7.3)}{\leq} k^2 \gamma dn + d \leq \gamma^{2/3} k^{-2} d^2 / 3, \end{aligned}$$

where the last inequality holds as $d \geq \alpha n$ and $1/n \ll \gamma \ll \alpha, 1/k$. Therefore,

$$(iv') \sum_{i,j \in [k]} (|X_{ij}^+| + |X_{ij}^-| + e(M_{ij})) \leq \gamma^{1/3} d / 3.$$

Step 3: Defining flow networks. For $i, j \in [k]$, let

$$M = \bigcup_{i,j \in [k]} M_{ij}, \quad X_{i*}^+ = \bigcup_{j \in [k]} (X_{ij}^+ \cup V^+(M_{ij})), \text{ and } X_{*j}^- = \bigcup_{i \in [k]} (X_{ij}^- \cup V^-(M_{ij})).$$

Note that $X_{i*}^+ \subseteq V_{i*}$ and $X_{*j}^- \subseteq V_{*j}$. Let $X^+ = \bigcup_{i \in [k]} X_{i*}^+$ and $X^- = \bigcup_{j \in [k]} X_{*j}^-$.

We now define a flow network $\mathcal{F} = (F, w, S, T)$ with multiple sources and sinks as follows. Let $S = \{s_i : i \in [k]\}$ and $T = \{t_i : i \in [k]\}$. Let $V(F) = S \cup T \cup X^+ \cup X^-$. Here we treat X^+ and X^- as

disjoint (i.e., both X^+ and X^- contain a distinct copy of v for any vertex $v \in X^+ \cap X^-$). For $v \in V(G)$, we write v^+ (and v^-) for the copy of v belonging to X^+ (and X^- , respectively). We define $E(F)$ as the union of edge-disjoint paths as follows: for each $i, j \in [k]$ with $i < j$,

- for each $x^+ \in X_{ij}^+$, we add the directed path $s_i x^+ t_j$;
- for each $x^- \in X_{ij}^-$, we add the directed path $s_i x^- t_j$;
- for each $e = uv \in M_{ij}$, we add the directed path $s_i u^+ v^- t_j$;
- every edge and vertex has capacity 1 (i.e., $w(e) = 1$ for all $e \in E(F)$ and $w(v) = 1$ for all $v \in V(F) \setminus (S \cup T) = X^+ \cup X^-$).

We modify the flow network slightly if (m1) or (m2) holds. Let

$$F_0 = \begin{cases} F - \{v_0^+\} - \{s_{i_0} v_0^-\} - E_F(X^+, v_0^-) & \text{if } v_0 \in V_{i_0 j_0} \text{ exists satisfying (m2),} \\ F & \text{otherwise.} \end{cases}$$

We further remove v_0^- from F_0 if v_0^- has no inneighbour in F_0 . Write w_0 for the capacities on F_0 inherited from F and define $\mathcal{F}_0 = (F_0, w_0, S, T)$. By the definition of $\mathcal{F}_0$, note for later that

$$\text{if there is any flow through } v_0^-, \text{ then it must be via some edge } s_i v_0^- \text{ with } i \neq i_0. \quad (7.6)$$

Given a flow f on $\mathcal{F}$ or $\mathcal{F}_0$, for each $i, j \in [k]$, let f_{ij} be the sum of flow over all edges from $s_i \cup X_{i*}^+$ to $t_j \cup X_{*j}^-$, that is,

$$f_{ij} = \sum_{e \in E_F(s_i \cup X_{i*}^+, t_j \cup X_{*j}^-)} f(e). \quad (7.7)$$

Note that the total amount of flow going out of s_i (all going to X^+) is $\sum_{j \in [k]} f_{ij}$, and the total amount of flow into t_j is $\sum_{i \in [k]} f_{ij}$. We now reduce the lemma to the following claim about fractional flows.

Claim 7.10. To prove the lemma, it suffices to find a fractional flow f on $\mathcal{F}_0$ such that

$$\sum_{i, j \in [k]} \max\{e(G_{ij})/d - f_{ij}, 0\} < 1. \quad (7.8)$$

Proof of claim. Assume f is as given in the claim and we wish to prove the lemma. We may assume that $f_{ij} \leq e(G_{ij})/d$ for all $i, j \in [k]$.⁶ We define a new flow network $\mathcal{F}^* = (F^*, w^*, s^*, t^*)$ with a single source s^* and sink t^* as follows. We obtain $\mathcal{F}^*$ from $\mathcal{F}_0$ by adding the new vertices s^* and t^* to $V(F_0)$ and, for each $i \in [k]$,

- adding the edge $s^* s_i$ of capacity $\max\{|V_{i*}| - |V_{*i}|, 0\}$;
- adding the edge $t_i t^*$ of capacity $\max\{|V_{*i}| - |V_{i*}|, 0\}$;
- adding an edge $t_i s_i$ of infinite capacity;
- giving the vertices in $S \cup T$ infinite capacity.

We will use f to find an integer flow in $\mathcal{F}^*$ of value $\sum_{i \in [k]} ||V_{i*}| - |V_{*i}||/2$. Then, by looking at the vertices through which there is non-zero flow, we will define the sets Y_{ij}^+ , Y_{ij}^- , and M_{ij} satisfying the properties for Proposition 7.9, which will give the desired path system $\mathcal{Q}$.

First we find a fractional flow on $\mathcal{F}^*$ as follows. Let $\mathcal{F}^+ = (F^+, w^+, s^*, t^*)$ be obtained from $\mathcal{F}^*$ by adding, for each $i, j \in [k]$, an edge $s_i t_j$ with edge capacity 1. Define a flow f^+ on $\mathcal{F}^+$ such that, for all $i, j \in [k]$,

⁶Otherwise we could suitably decrease flow along paths P with positive flow from any source to any sink that uses an edge from $E_F(s_i \cup X_{i*}^+, t_j \cup X_{*j}^-)$, and this can be done independently for each $i, j \in [k]$, since the $E_F(s_i \cup X_{i*}^+, t_j \cup X_{*j}^-)$ are disjoint over all $i, j \in [k]$.

- $f^+(e) = f(e)$ for all $e \in E(F_0)$;
- $f^+(s^*s_i) = \max\{|V_{i*}| - |V_{*i}|, 0\}$;
- $f^+(t_jt^*) = \max\{|V_{*j}| - |V_{j*}|, 0\}$;
- $f^+(t_is_i) = \min\{e(G_{i*})/d, e(G_{*i})/d\}$;
- $f^+(s_it_j) = e(G_{ij})/d - f_{ij} \geq 0$.

One can check that f^+ is indeed a fractional flow⁷ on $\mathcal{F}^+$ with value $\sum_{i \in [k]} ||V_{i*}| - |V_{*i}||/2$. By applying Proposition 7.2 to reduce the flow in s_it_j to 0 for all $i, j \in [k]$, we obtain a fractional flow on $\mathcal{F}^*$ of value

$$r \geq \frac{1}{2} \sum_{i \in [k]} ||V_{i*}| - |V_{*i}|| - \sum_{i,j \in [k]} f^+(s_it_j) = \frac{1}{2} \sum_{i \in [k]} ||V_{i*}| - |V_{*i}|| - \sum_{i,j \in [k]} \left(\frac{e(G_{ij})}{d} - f_{ij} \right) \\ \stackrel{(7.8)}{>} \frac{1}{2} \left(\sum_{i \in [k]} ||V_{i*}| - |V_{*i}|| \right) - 1.$$

Note that all edge capacities are integral, so Theorem 7.1 implies there exists an integral flow f^* on $\mathcal{F}^*$ of value at least $\lceil r \rceil = \sum_{i \in [k]} ||V_{i*}| - |V_{*i}||/2$. Moreover, since the cuts $E_{F^*}(s^*, V(F^*) \setminus \{s^*\})$ and $E_{F^*}(V(F^*) \setminus \{t^*\}, t^*)$ have capacity $\sum_{i \in [k]} ||V_{i*}| - |V_{*i}||/2$, Theorem 7.1 implies that f^* saturates all the edges s^*s_i and t_jt^* for all $i, j \in [k]$.

We now define $Y_{ij}^+ \subseteq X_{ij}^+$, $Y_{ij}^- \subseteq X_{ij}^-$ and $M_{ij}^* \subseteq M_{ij}$ as follows. Initially, set $Y_{ij}^+ = Y_{ij}^- = M_{ij} = \emptyset$ for all $i, j \in [k]$. For each $x^+ \in X^+$ with $f^*(x^+) = 1$, since f^* is an integral flow, there exist unique $i, j \in [k]$ such that either $f^*(s_ix^+) = f^*(x^+t_j) = 1$ or $f^*(s_ix^+) = f^*(x^+z) = f^*(zt_j) = 1$ for some (unique) $z \in X^-$. In the former case we add x^+ to Y_{ij}^+ while we add x^+z into M_{ij}^* for the latter. In a similar way, for $x^- \in X^-$ with $f^*(x^-) = 1$, we either add x^- to the set Y_{ij}^- or add an edge ending at x^- into M_{ij}^* . Note that

$$|Y_{ij}^+| + |Y_{ij}^-| + e(M_{ij}^*) = \sum_{e \in E_{F^*}(s_i \cup X_{i*}^+, t_j \cup X_{*j}^-)} f^*(e) =: f_{ij}^*. \quad (7.9)$$

Clearly $M^* := \bigcup_{i,j \in [k]} M_{ij}^* \subseteq M$ is a matching by (iii'). Both of $\{Y_{ij}^+, V^+(M_{ij}^*) : i, j \in [k]\}$ and $\{Y_{ij}^-, V^-(M_{ij}^*) : i, j \in [k]\}$ are sets of disjoint sets as each $x \in X^+ \cup X^-$ has vertex capacity of one. Set

$$W = \begin{cases} \bigcup_{i,j \in [k]: i \neq j} V_{ij} & \text{if } \sum_{i,j \in [k]: i \neq j} |V_{ij}| \leq k^2 \gamma n, \\ \emptyset & \text{otherwise.} \end{cases}$$

For all $y^+ \in Y_{ij}^+ \subseteq X_{ij}^+$, we have

$$d_{G_{ij}}^+(y^+) \stackrel{(7.5)}{\geq} \gamma^{1/3} d \stackrel{(iv')}{\geq} 2 \sum_{i,j \in [k]} \left(|X_{ij}^+| + |X_{ij}^-| + e(M_{ij}^*) \right) + \gamma^{1/3} d/3 \\ \geq 2 \sum_{i,j \in [k]} \left(|Y_{ij}^+| + |Y_{ij}^-| + e(M_{ij}^*) \right) + |W| + 1$$

since $k^2 \gamma n + 1 \leq 2k^2 \gamma n \leq \gamma^{1/3} d/3$ as $\gamma \ll \alpha \leq d/n$, and a similar inequality holds for all $y^- \in Y_{ij}^-$. We apply Proposition 7.9 and obtain a path system $\mathcal{Q}$ of G such that

- (i'') $e(\mathcal{Q}_{ij}) = |Y_{ij}^+| + |Y_{ij}^-| + e(M_{ij}^*) \stackrel{(7.9)}{=} f_{ij}^*$ for all $i, j \in [k]$;
- (ii'') $V(\mathcal{Q}) \cap W \subseteq \bigcup_{i,j \in [k]} (Y_{ij}^+ \cup Y_{ij}^- \cup V(M_{ij}^*))$;

⁷Since f is a fractional flow on $\mathcal{F}_0$, it suffices to examine only s_i and t_i for $i \in [k]$. For s_i , the amount of flow going out of s_i is $\sum_{j \in [k]} f_{ij}$ (from s_i to X^+) and $\sum_{j \in [k]} (e(G_{ij})/d - f_{ij})$ (from s_i to T), and the flow going into s_i is $\min\{e(G_{i*})/d, e(G_{*i})/d\}$ (from t_i to s_i) and $\max\{|V_{i*}| - |V_{*i}|, 0\}$ (from s^* to s_i). By noting that (7.4) gives $e(G_{i*})/d - e(G_{*i})/d = |V_{i*}| - |V_{*i}|$, the total contribution is 0. For t_i , the calculation is similar.

(iii'') if $y^+ \in Y_{ij}^+ \setminus V^-(M^*)$, then there exists $u \in V_{*j} \setminus W$ such that the single edge y^+u is a path in $\mathcal{Q}$, and a similar statement for $y^- \in Y_{ij}^- \setminus V^+(M^*)$ holds.

Consider $i \in [k]$. Recall that $N_{F^*}^-(s_i) = \{s^*, t_i\}$, $N_{F^*}^+(s_i) \subseteq X_{i*}^+ \cup X^-$, and each vertex in X^+ has capacity one. We have

$$\begin{aligned} f^*(s^*s_i) + f^*(t_is_i) &= \sum_{u \in N_{F^*}^-(s_i)} f^*(us_i) = \sum_{v \in N_{F^*}^+(s_i)} f^*(s_iv) \\ &= \sum_{v \in X_{i*}^+ \cup X^-} f^*(s_iv) = \sum_{j \in [k]} \sum_{e \in E_{F^*}(s_i \cup X_{i*}^+, t_j \cup X_{*j}^-)} f^*(e) = \sum_{j \in [k]} f_{ij}^{*(i'')} \sum_{j \in [k]} e(\mathcal{Q}_{ij}) \end{aligned}$$

and similarly we have $f^*(t_it^*) + f^*(t_is_i) = \sum_{j \in [k]} e(\mathcal{Q}_{ji})$. Since f^* saturates the edges s^*s_i and t_it^* , we deduce that,

$$\begin{aligned} \sum_{j \in [k]} e(\mathcal{Q}_{ij}) - \sum_{j \in [k]} e(\mathcal{Q}_{ji}) &= f^*(s^*s_i) - f^*(t_it^*) = w^*(s^*s_i) - w^*(t_it^*) \\ &= \max\{|V_{i*}| - |V_{*i}|, 0\} - \max\{|V_{*i}| - |V_{i*}|, 0\} = |V_{i*}| - |V_{*i}| \end{aligned}$$

implying that $\mathcal{Q}$ is $\mathcal{P}$ -balanced, as required. By applying Proposition 7.6 (with $d = 1$ and $G^0 = \mathcal{Q}$), we may further assume that

$$e(\mathcal{Q}) \leq (k-1) \sum_{i \in [k]} \|V_{i*}| - |V_{*i}|\|/2 \stackrel{(ii)}{\leq} k^2 \gamma n. \quad (7.10)$$

We now check the moreover statement of the Lemma 6.1. If $\sum_{i,j \in [k]: i \neq j} |V_{ij}| > k^2 \gamma n$ (that is, (m1) holds), then (7.10) implies that $\bigcup_{i,j \in [k]: i \neq j} V_{ij} \setminus V(\mathcal{Q}) \neq \emptyset$, so $\mathcal{Q}$ is non-trivial. Suppose that $\sum_{i,j \in [k]: i \neq j} |V_{ij}| \leq k^2 \gamma n$ and there exists $v_0 \in V_{i_0j_0}$ satisfying (m2). If $v_0 \notin V(\mathcal{Q})$, then $\mathcal{Q}$ is non-trivial. So suppose that $v_0 \in V(\mathcal{Q})$. Since $v_0 \in W$, by (ii'') we must have $v_0 \in Y_{ij}^+ \cup Y_{ij}^- \cup V(M_{ij}^*)$ for some $i, j \in [k]$. This means that there is a flow through v_0^+ as $v_0^+ \notin V(F^*)$. By (7.6), $v_0 \in Y_{i_0j_0}^-$ for some $i \neq i_0$. Also, $v_0^+ \notin V(F^*)$ implies that $v_0 \notin V^+(M^*)$. By (iii''), we deduce that there exists $u \in V_{i*} \setminus W = V_{ii}$ such that uv_0 is a path in $\mathcal{Q}$. Note that uv_0 is a path from $V_{ii} \subseteq V_{*i}$ to $V_{i_0j_0} \subseteq V_{i_0*}$. Hence $\mathcal{Q}$ is non-trivial, as required. $\square$

Step 4: Define a distribution function on $X^+ \cup X^-$. For each $i \in [k]$, $x^+ \in X_{i*}^+$ and $x^- \in X_{*i}^-$, $j \in [k] \setminus \{i\}$, we set

$$p_j(x^+) = \max\{d_{G_{ij}}^+(x^+)/d, 6k^2\gamma^{1/3}\} \quad \text{and} \quad p_j(x^-) = \max\{d_{G_{ji}}^-(x^-)/d, 6k^2\gamma^{1/3}\}.$$

Given $x^+ \in X_{i*}^+$, we ‘view’ $p_j(x^+)$ to be flow through x^+ from s_i to t_j , and similarly for $p_j(x^-)$.

Claim 7.11. For each $i \in [k]$ and $x^+ \in X_{i*}^+$,

$$\sum_{j \in [k] \setminus \{i\}} p_j(x^+) \leq 1 - \frac{d_{G^0}^+(x, V_{*i})}{d} + 6k^3\gamma^{1/3} < 1$$

and a similar statement holds for each $j \in [k]$ and $x^- \in X_{*j}^-$.

Proof of claim. We will only consider the case when $x^+ \in X_{i*}^+$ as the other case when $x^- \in X_{*j}^-$ can be proved similarly. Since G is a subgraph of G^0 , we have

$$\sum_{j \in [k] \setminus \{i\}} d_{G_{ij}}^+(x^+) \leq \sum_{j \in [k] \setminus \{i\}} d_{G_{ij}}^+(x^+) = d - d_{G^0}^+(x, V_{*i}).$$

Hence

$$\begin{aligned} \sum_{j \in [k] \setminus \{i\}} p_j(x^+) &\leq \sum_{j \in [k] \setminus \{i\}} \left(\frac{d_{G_{ij}}^+(x^+)}{d} + 6k^2\gamma^{1/3} \right) \leq \frac{\sum_{j \in [k] \setminus \{i\}} d_{G_{ij}}^+(x^+)}{d} + 6k^3\gamma^{1/3} \\ &\leq 1 - \frac{d_{G_0}^+(x, V_{*i})}{d} + 6k^3\gamma^{1/3} < 1, \end{aligned}$$

where the last inequality holds as $d_{G_0}^+(x, V_{*i}) \geq d/k$ by (i) and $\gamma \ll 1/k$. $\square$

Step 5: Defining fractional flow on $\mathcal{F}_0$. We first define the fractional flow f on $\mathcal{F}$ as follows. For each $i, j \in [k]$, we do the following

- for each $x^+ \in X_{ij}^+ \subseteq V_{i*}$, add a flow through $s_i x^+ t_j$ of value $p_j(x^+)$;
- for each $x^- \in X_{ij}^- \subseteq V_{*j}$, add a flow through $s_i x^- t_j$ of value $p_i(x^-)$;
- for each edge $e = uv \in M_{ij}$, add a flow through $s_i u^+ v^- t_j$ of value $6k^2\gamma^{1/3}$.

Our construction and Claim 7.11 imply that each vertex $v \in X^+ \cup X^-$ has a flow of value at most 1 through it. Thus f is indeed a fractional flow on $\mathcal{F}$. For each $i, j \in [k]$,

$$\begin{aligned} df_{ij} &\stackrel{(7.7)}{=} d \left(\sum_{x^+ \in X_{ij}^+} p_j(x^+) + \sum_{x^- \in X_{ij}^-} p_i(x^-) + \sum_{uv \in M_{ij}} 6k^2\gamma^{1/3} \right) \\ &\geq \sum_{x^+ \in X_{ij}^+} d_{G_{ij}}^+(x^+) + \sum_{x^- \in X_{ij}^-} d_{G_{ij}}^-(x^-) + (6k^2\gamma^{1/3}d)e(M_{ij}) \stackrel{(i')}{\geq} e(G_{ij}) - 50k^{10}\gamma^{1/3}d. \end{aligned}$$

Therefore,

$$\sum_{i,j \in [k]} \max\{e(G_{ij})/d - f_{ij}, 0\} \leq 50k^{12}\gamma^{1/3} < 1.$$

If there is no v_0 satisfying (m2), then $\mathcal{F}_0 = \mathcal{F}$ and we are done by Claim 7.10. If such a v_0 exists, then recall that F_0 is obtained from F removing the vertex v_0^+ , the edge $s_{i_0}v_0^-$ and edges in $E_F(X^+, v_0^-)$. By repeated application of Proposition 7.2, we obtain a flow in $\mathcal{F}_0$ whose value is lower than that of f by at most

$$\begin{aligned} &\sum_{j \in [k] \setminus \{i_0\}} p_j(v_0^+) + p_{i_0}(v_0^-) + 2 \cdot 6k^2\gamma^{1/3} \\ &\stackrel{\text{Claim 7.11}}{\leq} 1 - \frac{d_{G_0}^+(v_0, V_{*i_0})}{d} + 6k^3\gamma^{1/3} + \left(\frac{d_{G_{i_0j_0}}^-(v_0)}{d} + 6k^2\gamma^{1/3} \right) + 12k^2\gamma^{1/3} \\ &\leq 1 - \frac{d_{G_0}^+(v_0, V_{*i_0})}{d} + \frac{d_{G_0}^-(v_0, V_{i_0*})}{d} + 24k^3\gamma^{1/3} \\ &\stackrel{(m2)}{\leq} 1 - 100k^{12}\gamma^{1/3} + 24k^3\gamma^{1/3} < 1 - 50k^{12}\gamma^{1/3}. \end{aligned}$$

Therefore, $\sum_{i,j \in [k]} \max\{e(G_{ij})/d - f_{ij}, 0\}$ is still less than 1. Hence we are done by Claim 7.10. $\square$

8. Proof of Lemma 6.3

We begin with an outline of the proof of the Lemma 6.3. Let G and $\mathcal{P}$ be the digraph and vertex partition as in the statement of Lemma 6.3. Suppose, contrary to the lemma, that there is no non-trivial $\mathcal{P}$ -balanced path system.

Consider distinct $i, j \in [k]$. Note that if there exist two vertex-disjoint edges $e_1 \in E(G_{ij})$ and $e_2 \in E(G_{ji})$, then $\{e_1, e_2\}$ forms a non-trivial $\mathcal{P}$ -balanced path system. Thus, if $e(G_{ij}), e(G_{ji}) \geq 3$, then we may assume that there exists a vertex $w_{ij} \in V_{ii} \cup V_{jj}$ that is contained in all edges of $G_{ij} \cup G_{ji}$. Let H be the oriented graph on $[k]$ where $ij \in E(H)$ if $e(G_{ij}), e(G_{ji}) \geq 3$ and $w_{ij} \in V_{jj}$; we give weight $w(ij) := |E(G_{ij} \cup G_{ji})|$ to ij .⁸ It turns out that the underlying undirected graph of H is acyclic (see Claim 8.1) and that total weight of edges in H is almost equal to $|\mathcal{B}(G, \mathcal{P})|$ (see (8.6)). Set $w(i) = \sum_{j \in [k]} (w(ij) - w(ji))$.

We focus on the V_{ii} that are relatively small: let $c_i = qd + 1 - |V_{ii}|$ (where we set $q = 2$ if G is an oriented graph and $q = 1$ otherwise) and after relabelling indices assume $|\{i \in [k] : c_i \geq 0\}| = [k_0]$. We will lower and upper bound $\sum_{i \in [k_0]} w(i)$. For the lower bound, we note that if $c_i > 0$ then every vertex of V_{ii} has at least one edge from outside V_{ii} , and so we can lower bound $w(i)$ in terms of c_i for each $i \in [k_0]$ (see (8.12)). For the upper bound, we note that $w(ij)$ can be upper bounded using (iii) (see (8.2)). Then, as H is acyclic (so has few edges), we are able to find a good upper bound for $\sum_{i \in [k_0]} w(i)$ by considering connected components (i.e., trees) in $H[[k_0]]$. Combining the lower and upper bounds, we obtain $\sum_{i \in [k_0]} c_i \leq k - k_0 - 1$ (see (8.15)) whereas (ii) implies $\sum_{i \in [k_0]} c_i > k - k_0$, a contradiction.

Recall that for a digraph G and $A, B \subseteq V(G)$ not necessarily disjoint, we write $E_G(A, B) := \{ab \in E(G) : a \in A, b \in B\}$ and $e_G(A, B) := |E_G(A, B)|$. We write $E_G(A)$ and $e_G(A)$ for $E_G(A, A)$ and $e_G(A, A)$, respectively.

Proof of Lemma 6.3. Assume G and $\mathcal{P}$ satisfy (i)–(iii) in the statement of Lemma 6.3. Let $q = 2$ if G is an oriented graph and $q = 1$ otherwise, so (ii) says that $n < (qd + 1)k$. We write V_i instead of V_{ii} for $i \in [k]$, so that, for all distinct $i, j \in [k]$, G_{ij} becomes the bipartite digraph with vertex classes V_i and V_j and edges given by the edges of G from V_i to V_j . In this context (where $V_{ij} = \emptyset$ for all distinct $i, j \in [k]$), $\mathcal{Q}$ is a $\mathcal{P}$ -balanced path system if, for each $i \in [k]$, the number of edges of $\mathcal{Q}$ going into V_i is equal to the number of edges of $\mathcal{Q}$ leaving V_i . Also, a $\mathcal{P}$ -balanced path system is non-trivial if it has at least one path whose endpoints lie in distinct parts of the partition $\mathcal{P} = \{V_i : i \in [k]\}$. Now suppose for a contradiction that G does not contain any non-trivial $\mathcal{P}$ -balanced path system with at most k edges.

Claim 8.1. Let H_0 be a digraph on $[k]$ such that $ij \in E(H_0)$ if $e(G_{ij}) \geq 3$ and $i \neq j$. Then H_0 does not have any directed cycle of length at least 3.

Proof of claim. Suppose to the contrary that C is such a cycle in H_0 . For each $ij \in E(C)$, we pick an edge $e_{ij} \in E(G_{ij})$ with $ij \in E(C)$ and call the resulting subdigraph G' . Note that G' is $\mathcal{P}$ -balanced and has at most k edges; moreover it is either a cycle or a path system. If G' is a cycle, then pick an $ij \in E(C)$, and find $e'_{ij} \in E(G_{ij})$ with $e'_{ij} \neq e_{ij}$. Write $G'' = (G' - \{e_{ij}\}) \cup \{e'_{ij}\}$ if G' is a cycle and $G'' = G'$ otherwise. Note that G'' is a $\mathcal{P}$ -balanced path system. If G'' is a single path, then there exists $t \in [k]$ such that the endpoints of G'' lies in V_t . Then, pick an $i'j' \in E(C)$ with $i', j' \neq t$, and find $e'_{i'j'} \in E(G_{i'j'})$ with $e'_{i'j'} \neq e_{i'j'}, e_{i'j'}$. Write $G''' = (G'' - \{e_{i'j'}\}) \cup \{e'_{i'j'}\}$ if G'' is a single path and $G''' = G''$ otherwise. Note that G''' consists of at least two vertex-disjoint paths and that for any path in G''' , the endpoints lie in distinct parts of the partition $\mathcal{P} = \{V_i : i \in [k]\}$, so G''' is a non-trivial $\mathcal{P}$ -balanced path system with at most k edges, a contradiction. $\square$

Claim 8.2. For all distinct $i, j \in [k]$, if $e(G_{ij}) \leq 2$, then $e(G_{ji}) \leq 2k^2$.

Proof of claim. Suppose to the contrary and without loss of generality that $e(G_{21}) \leq 2$ and $e(G_{12}) \geq 2k^2 + 1$. Let H_0 be as defined in Claim 8.1, so $12 \in E(H_0)$ and $21 \notin E(H_0)$. Let A be the set of $i \in [k]$ such that H_0 contains a directed path from 1 to i starting with the edge 12, where we a priori allow $1 \in A$ (i.e., resulting from a closed path). Let $B = [k] \setminus A$. It is clear from the definition that $2 \in A$, and

⁸Note that it is convenient to introduce the weight $w(ij)$ (and later $w(i)$) for the sketch of proof; these are not used in the actual proof.

by Claim 8.1 and since $21 \notin E(H_0)$ we have $1 \in B$. Note that by the definition of A , there are no edges in $E_{H_0}(A, B)$ so that $e(G_{ij}) \leq 2$ for all $ij \in A \times B$. Together with the regularity of G , we have

$$2k^2 \geq \sum_{ij \in A \times B} e(G_{ij}) = \sum_{ij \in B \times A} e(G_{ij}) \geq e(G_{12}) \geq 2k^2 + 1,$$

which is a contradiction. $\square$

Let H^* be the graph on $[k]$ such that $ij \in E(H^*)$ if $e(G_{ij}), e(G_{ji}) \geq 3$ (i.e., H^* is the graph obtained from H_0 by deleting all the edges which do not lie in a directed 2-cycle and by making each directed 2-cycle into an undirected edge). By Claim 8.1, H^* is acyclic. Consider distinct $i, j \in [k]$. Note that there do not exist disjoint edges e_1, e_2 with $e_1 \in E(G_{ij})$ and $e_2 \in E(G_{ji})$ because otherwise $\{e_1, e_2\}$ would form a non-trivial $\mathcal{P}$ -balanced path system. Thus if $ij \in E(H^*)$ (that is, $e(G_{ij}), e(G_{ji}) \geq 3$), then there exists a vertex $w_{ij} \in V_i \cup V_j$ that is contained in all edges of $G_{ij} \cup G_{ji}$. So, we have

$$e(G_{ij}) + e(G_{ji}) = d_{G_{ij} \cup G_{ji}}^+(w_{ij}) + d_{G_{ij} \cup G_{ji}}^-(w_{ij}) \stackrel{\text{(iii)}}{\leq} (2 - 1/k)d.$$

If $ij \notin E(H^*)$, then $\min\{e(G_{ij}), e(G_{ji})\} \leq 2$ and by Claim 8.2, we obtain

$$e(G_{ij}) + e(G_{ji}) \leq 2k^2 + 2 \leq (2 - 1/k)d \quad (8.1)$$

as $k \geq 2$ and $d > 165k^5$. Hence we have

$$e(G_{ij}) + e(G_{ji}) \leq (2 - 1/k)d \text{ for all distinct } i, j \in [k]. \quad (8.2)$$

Recall that $q = 2$ if G is an oriented graph, and $q = 1$ otherwise. Note that, if $|V_i| \leq qd + 1$ for some $i \in [k]$, then

$$\begin{aligned} 2kd &\stackrel{(8.2)}{\geq} \sum_{j \in [k] \setminus \{i\}} e(G_{ij}) = d|V_i| - e(G_{ii}) \geq d|V_i| - \frac{|V_i|(|V_i| - 1)}{q} \\ &= \frac{qd + 1 - |V_i|}{q} |V_i| \stackrel{(i)}{\geq} \frac{d(qd + 1 - |V_i|)}{2q} \geq \frac{d(qd + 1 - |V_i|)}{4}. \end{aligned}$$

Hence, for all $i \in [k]$,

$$|V_i| \geq qd + 1 - 8k. \quad (8.3)$$

We now orient H^* to obtain an oriented graph H on $[k]$ as follows. Recall that if $ij \in E(H^*)$, then there exists a vertex $w_{ij} \in V_i \cup V_j$ contained in all edges of $E(G_{ij}) \cup E(G_{ji})$. We orient from i to j if $w_{ij} \in V_j$.

Let W be the set of w_{ij} and $W_i = W \cap V_i$ for all $i \in [k]$. Let $V'_i = V_i \setminus W_i$. Note that

$$|W_i| \leq k - 1 \text{ and so } |V'_i| = |V_i| - |W_i| \stackrel{(8.3)}{\geq} qd - 9k. \quad (8.4)$$

Recall that $\mathcal{B}(G, \mathcal{P}) = \bigcup_{i, j \in [k]: i \neq j} G_{ij}$. Let $G^* = \bigcup_{i, j \in [k]} G_{ij}^*$ where G_{ij}^* is the subdigraph of $\mathcal{B}(G, \mathcal{P})$ consisting of the edges (in both directions) between V'_i and w_{ij} if $ij \in E(H)$ and $G_{ij}^* = \emptyset$ otherwise.⁹ Note that G_{ij}^* is a subdigraph of $G_{ij} \cup G_{ji}$ whenever $ij \in E(H)$ and $G_{ij}^* = \emptyset$ otherwise. Therefore,

$$e(G_{ij}^*) \begin{cases} \leq e(G_{ij} \cup G_{ji}) & \text{if } ij \in E(H); \\ = 0 & \text{if } ij \notin E(H). \end{cases} \quad (8.5)$$

⁹i.e. $G_{ij}^* = G_{ij}[V'_i \cup w_{ij}] \cup G_{ji}[V'_i \cup w_{ij}]$ if $ij \in E(H)$ and $G_{ij}^* = \emptyset$ otherwise.

Hence, we deduce that

$$\begin{aligned} e(\mathcal{B}(G, \mathcal{P})) - e(G^*) &\leq \sum_{ij \notin E(H^*)} (e(G_{ij}) + e(G_{ji})) + \sum_{i, j \in [k]: i \neq j} |W_i||W_j| \\ &\stackrel{(8.1), (8.4)}{\leq} \binom{k}{2} (2k^2 + 2) + k(k-1)^3 < 3k^4. \end{aligned} \quad (8.6)$$

For $i \in [k]$, let $\bar{V}_i = V(G) \setminus V_i$. Since all edges in G^* contain a vertex in W and $G^*[W] \cup \bigcup_{i \in [k]} G_{ii}^*$ is empty, we have for all $i \in [k]$,

$$\sum_{j \in [k]} e(G_{ij}^*) = \sum_{j \in [k], ij \in E(H)} \left(e_{G^*}(V'_i, w_{ij}) + e_{G^*}(w_{ij}, V'_i) \right) = e_{G^*}(V'_i, \bar{V}_i) + e_{G^*}(\bar{V}_i, V'_i) \quad (8.7)$$

and, for all $j \in [k]$,

$$\sum_{i \in [k]} e(G_{ij}^*) = \sum_{i \in [k], ij \in E(H)} \left(e_{G^*}(V'_i, w_{ij}) + e_{G^*}(w_{ij}, V'_i) \right) = e_{G^*}(\bar{V}_j, W_j) + e_{G^*}(W_j, \bar{V}_j). \quad (8.8)$$

We also need the following inequality

$$e_G(V'_i, V_i) = e_G(V'_i) + \sum_{w \in W_i} d_G^-(w, V'_i) \leq e_G(V'_i) + \sum_{w \in W_i} d_G^-(w, V_i). \quad (8.9)$$

Now, let $k_0 = |\{i \in [k] : |V_i| \leq qd + 1\}|$. Note that $k_0 > 0$ since $n < (qd + 1)k$ by (ii). Without loss of generality,

$$|V_i| \leq qd + 1 \text{ if and only if } i \in [k_0]. \quad (8.10)$$

For $i \in [k_0]$,

$$\begin{aligned} d|V'_i| - e_G(V'_i) &\geq d|V'_i| - \frac{|V'_i|(|V'_i| - 1)}{q} = \frac{|V'_i|}{q} (qd + 1 - |V'_i|) = \frac{|V'_i|}{q} (qd + 1 - |V_i|) + \frac{|V'_i||W_i|}{q} \\ &\stackrel{(8.4)}{\geq} \frac{qd - 9k}{q} (qd + 1 - |V_i|) + \frac{|V'_i||W_i|}{q} \\ &= d(qd + 1 - |V_i|) - \frac{9k(qd + 1 - |V_i|)}{q} + \frac{|V'_i||W_i|}{q} \\ &\stackrel{(8.3)}{\geq} d(qd + 1 - |V_i|) - 72k^2 + \frac{|V'_i||W_i|}{q}. \end{aligned} \quad (8.11)$$

Then,

$$\begin{aligned} e_G(V'_i, \bar{V}_i) &= \sum_{v \in V'_i} d_G^+(v, \bar{V}_i) = \sum_{v \in V'_i} (d - d_G^+(v, V_i)) = d|V'_i| - e_G(V'_i, V_i) \\ &\stackrel{(8.9)}{\geq} d|V'_i| - e_G(V'_i) - \sum_{w \in W_i} d_G^-(w, V_i) = d|V'_i| - e_G(V'_i) - \sum_{w \in W_i} (d - d_G^-(w, \bar{V}_i)) \\ &= d|V'_i| - e_G(V'_i) - d|W_i| + e_G(\bar{V}_i, W_i) \\ &\stackrel{(8.11)}{\geq} d(qd + 1 - |V_i|) - 72k^2 - \frac{|W_i|}{q} (qd - |V'_i|) + e_G(\bar{V}_i, W_i) \\ &\stackrel{(8.4)}{\geq} d(qd + 1 - |V_i|) - 81k^2 + e_G(\bar{V}_i, W_i) \end{aligned}$$

and a similar inequality holds for $e_G(\overline{V}_i, V'_i)$. Hence, for $i \in [k_0]$,

$$e_G(V'_i, \overline{V}_i) + e_G(\overline{V}_i, V'_i) \geq 2d(qd + 1 - |V_i|) - 162k^2 + e_G(\overline{V}_i, W_i) + e_G(W_i, \overline{V}_i). \quad (8.12)$$

Let $\mathcal{I}$ be the set of all connected components I in $H^*[[k_0]]$ with $\sum_{i \in I}(qd + 1 - |V_i|) > 0$.¹⁰ Such $\mathcal{I}$ is non empty as $n < (qd + 1)k$ by (ii). We have that

$$\sum_{i \in [k_0]} (qd + 1 - |V_i|) = \sum_{I \in \mathcal{I}} \sum_{i \in I} (qd + 1 - |V_i|). \quad (8.13)$$

Recall that H^* is acyclic and H is an orientation of H^* . Note that for each $I \in \mathcal{I}$, $H[I]$ is an oriented tree and $E_H(I, [k] \setminus [k_0]) = E_H(I, [k]) \setminus E_H(I)$. Therefore, for each $I \in \mathcal{I}$,

$$\begin{aligned} (2 - 1/k)d \cdot e_H(I, [k] \setminus [k_0]) &= (2 - 1/k)d \cdot |E_H(I, [k]) \setminus E_H(I)| \\ &\stackrel{(8.2)}{\geq} \sum_{ij \in E(H): i \in I, j \in [k] \setminus I} e(G_{ij} \cup G_{ji}) \stackrel{(8.5)}{\geq} \sum_{ij \in E(H): i \in I, j \in [k] \setminus I} e(G_{ij}^*) \\ &= \sum_{ij \in E(H): i \in I, j \in [k]} e(G_{ij}^*) - \sum_{ij \in E(H): i, j \in I, i \neq j} e(G_{ij}^*) \\ &\stackrel{(8.5)}{\geq} \sum_{ij \in I \times [k]} e(G_{ij}^*) - \sum_{ij \in [k] \times I} e(G_{ij}^*) \\ &\stackrel{(8.7), (8.8)}{=} \sum_{i \in I} \left(e_{G^*}(V'_i, \overline{V}_i) + e_{G^*}(\overline{V}_i, V'_i) \right) - \sum_{j \in I} \left(e_{G^*}(\overline{V}_j, W_j) + e_{G^*}(W_j, \overline{V}_j) \right) \\ &\stackrel{(8.6)}{\geq} \sum_{i \in I} \left(e_G(V'_i, \overline{V}_i) + e_G(\overline{V}_i, V'_i) \right) - \sum_{j \in I} \left(e_G(\overline{V}_j, W_j) + e_G(W_j, \overline{V}_j) \right) - 3k^4 \\ &= \sum_{i \in I} \left(\left(e_G(V'_i, \overline{V}_i) + e_G(\overline{V}_i, V'_i) \right) - \left(e_G(\overline{V}_i, W_i) + e_G(W_i, \overline{V}_i) \right) \right) - 3k^4 \\ &\stackrel{(8.12)}{\geq} 2d \sum_{i \in I} (qd + 1 - |V_i|) - 162|I|k^2 - 3k^4 \geq 2d \sum_{i \in I} (qd + 1 - |V_i|) - 165k^4. \end{aligned}$$

After rearranging, for each $I \in \mathcal{I}$, we have

$$\begin{aligned} e_H(I, [k] \setminus [k_0]) &\geq \frac{2}{2 - 1/k} \sum_{i \in I} (qd + 1 - |V_i|) - \frac{165k^4}{(2 - 1/k)d} \\ &= \sum_{i \in I} (qd + 1 - |V_i|) + \frac{1}{2k - 1} \sum_{i \in I} (qd + 1 - |V_i|) - \frac{165k^5}{(2k - 1)d} \\ &\geq \sum_{i \in I} (qd + 1 - |V_i|) + \frac{1}{2k - 1} - \frac{165k^5}{(2k - 1)d} > \sum_{i \in I} (qd + 1 - |V_i|) \end{aligned}$$

¹⁰Here we identify the connected component with its vertex set.

as $\sum_{i \in I} (qd + 1 - |V_i|) \geq 1$ and $d > 165k^5$. Hence, we have

$$\begin{aligned} \sum_{I \in \mathcal{I}} e_H(I, [k] \setminus [k_0]) &\geq \sum_{I \in \mathcal{I}} \left(1 + \sum_{i \in I} (qd + 1 - |V_i|) \right) = |\mathcal{I}| + \sum_{I \in \mathcal{I}} \sum_{i \in I} (qd + 1 - |V_i|) \\ &\stackrel{(8.13)}{=} |\mathcal{I}| + \sum_{i \in [k_0]} (qd + 1 - |V_i|). \end{aligned} \quad (8.14)$$

Recall that H is an oriented forest and so $H[I]$ is an oriented tree for all $I \in \mathcal{I}$. Then,

$$\begin{aligned} \sum_{I \in \mathcal{I}} e_H(I, [k] \setminus [k_0]) &\leq e_H\left([k] \setminus [k_0] \cup \bigcup_{I \in \mathcal{I}} I\right) - \sum_{I \in \mathcal{I}} e_H(I) \\ &\leq \left(\left(k - k_0 + \sum_{I \in \mathcal{I}} |I| \right) - 1 \right) - \sum_{I \in \mathcal{I}} (|I| - 1) = k - k_0 + |\mathcal{I}| - 1. \end{aligned}$$

Together with (8.14), we have

$$\sum_{i \in [k_0]} (qd + 1 - |V_i|) \leq k - k_0 - 1. \quad (8.15)$$

Finally

$$\begin{aligned} n &= \sum_{i \in [k]} |V_i| \stackrel{(8.10)}{\geq} \sum_{i \in [k_0]} |V_i| + (qd + 2)(k - k_0) \\ &= (qd + 1)k + (k - k_0) - \sum_{i \in [k_0]} (qd + 1 - |V_i|) \stackrel{(8.15)}{\geq} (qd + 1)k + 1, \end{aligned}$$

a contradiction to property (ii) of Lemma 6.3. This completes the proof of the lemma. $\square$

9. Conclusion

9.1. Path cover for (non-regular) graphs

Magnant, Wang and Yuan [29] gave a stronger version of Conjecture 1.5. Recall that $\pi(G)$ is the minimum number of vertex-disjoint paths needed to cover G .

Conjecture 9.1 (Magnant, Wang and Yuan [29]). *If G is a graph on n vertices with $\Delta(G) = \Delta$ and $\delta(G) = \delta$, then $\pi(G) \leq \max\{n/(\delta + 1), (\Delta - \delta)n/(\Delta + \delta)\}$.*

The bound is tight by considering a disjoint union of $K_{\delta+1}$ or a disjoint union of $K_{\delta, \Delta}$. The conjecture holds if $\delta \leq 2$ [29] and when $\Delta \geq 2\delta$ [20]. Naturally, one can ask for the directed or oriented analogues.

9.2. Edge-disjoint cycles

In a weaker version of the problem that we have considered, one is interested in finding *edge-disjoint* cycles whose union covers all the vertices. As a generalization of Dirac's theorem, it was conjectured by Enomoto, Kaneko and Tuza [8] that if a graph G on n vertices has minimum degree at least n/k , then $V(G)$ can be covered by $k - 1$ edge-disjoint cycles. The case $k = 3$ was also proved in [8]. The conjecture was proved for 2-connected graphs [18], and has been completely resolved in [19]. Later, Balogh, Mousset and Skokan [3] obtained a stability result, showing that every graph on n vertices with minimum degree nearly n/k has a special structure if it does not have $k - 1$ edge-disjoint cycles covering

all vertices. One can ask the same question for digraphs as a generalization of Ghoulia-Houri's theorem [12], and Theorem 1.3 answers it affirmatively for regular digraphs:

Conjecture 9.2. *Let $k \in \mathbb{N}$ with $k \geq 2$. If G is a digraph on n vertices with minimum semi-degree at least n/k , then $V(G)$ can be covered by $k - 1$ edge-disjoint cycles.*

9.3. Connectivity and regularity

Jackson's conjecture states that imposing regularity on an oriented graph reduces the degree threshold for Hamiltonicity. One might hope that imposing connectivity on regular oriented graphs can reduce the degree threshold for Hamiltonicity. We refer the reader to [27, Sections 1 and 7] for history and conjectures on Hamiltonicity in regular (directed or oriented) graphs with given connectivity.

Similarly, in the setting of cycle partitions, one might hope that connectivity in addition to regularity might reduce the upper bound in Conjecture 1.5. In the sparse setting, Reed [31] proved that every 3-regular connected n -vertex graph can be covered by at most $\lceil n/9 \rceil$ vertex-disjoint paths, and conjectured that it suffices to use $\lceil n/10 \rceil$ paths if connectivity is replaced by 2-connectivity. Recall that Conjecture 1.5 gives the upper bound of $n/4$ for 3-regular graphs (that are not necessarily connected). Yu [36] recently verified Reed's conjecture and gave an example of a (2-connected) d -regular graph on n vertices which requires at least $\approx n/(d+4)$ paths for $d \geq 13$. It would be interesting to investigate the general relationship between the degree and connectivity of a regular (di)graph or oriented graph that guarantees a small number of vertex-disjoint cycles that cover all the vertices.

Competing interest. The authors have no competing interests to declare.

Financial support. A. Lo was partially supported by EPSRC, grant no. EP/V002279/1 and EP/V048287/1. V. Patel was supported by the Netherlands Organisation for Scientific Research (NWO) through the Gravitation Programme Networks (024.002.003). M.A. Yıldız was supported by a Marie Skłodowska-Curie Action from the EC (COFUND grant no. 945045) and by the NWO Gravitation project NETWORKS (grant no. 024.002.003).

Data availability statement. There are no additional data beyond that contained within the main manuscript.

Ethical standards. The research meets all ethical guidelines, including adherence to the legal requirements of the study country.

Author contributions. All authors contributed equally to the paper.

References

- [1] M. Abreu, J.B. Gauci, D. Labbate, G. Mazzuocollo, and J. P. Zerafa, 'Extending perfect matchings to Hamiltonian cycles in line graphs', *Electron. J. Comb.* **28**(1) (2021), Paper P1.7.
- [2] J. Akiyama, G. Exoo, and F. Harary, 'Covering and packing in graphs IV: Linear arboricity', *Networks* **11**(1), (2006) 69–72.
- [3] J. Balogh, F. Mousset, and J. Skokan, 'Stability for vertex cycle covers', *Electron. J. Comb.* **24** (2017), Paper P3.56.
- [4] G.B. Dantzig and D.R. Fulkerson, 'On the Max Flow MinCut Theorem of Networks', in *Linear Inequalities and Related systems*, *Ann. Math. Studies* **38** (1956), 215–221.
- [5] R. Diestel, 'Graph Theory (4th Edition)', *Graduate Texts in Mathematics* 173, Springer (2012).
- [6] P. Dilworth, 'A decomposition theorem for partially ordered sets', *Ann. of Math.* **51**(1), (1950), 161–166.
- [7] G.A. Dirac, 'Some theorems on abstract graphs', *Proc. Lond. Math. Soc.* **2** (1952), 69—81.
- [8] H. Enomoto, A. Kaneko, and Z. Tuza, ' P_3 -factors and covering cycles in graphs of minimum degree $n/3$ ', *Combinatorics (Proc. Coll. Math. Soc. János Bolyai Eger, 1987)*, North Holland (1988), 213–220.
- [9] U. Feige and E. Fuchs, 'On the path partition number of 6-regular graphs', *J Graph Theory* **101**, 345–378.
- [10] J. Fink, 'Perfect matchings extend to Hamilton cycles in hypercubes', *J. Combin. Theory. Ser. B* **97** (2007) 1074–1076.
- [11] T. Gallai and A.N. Milgram, 'Verallgemeinerung eines graphentheoretischen satzes von Redei', *Acta Sci. Math.* **21**, (1960), 181–186.
- [12] A. Ghouila-Houri, 'Une condition suffisante d'existence d'un circuit hamiltonien', *C.R. Acad. Sci. Paris* **25** (1960), 495–497.
- [13] V. Gruslys and S. Letzter, 'Cycle partitions of regular graphs', *Comb. Probab. Comput.* **30**(4) (2021), 526–549.
- [14] R. Häggkvist, 'On F -Hamiltonian graphs', In *Graph Theory and Related Topics*, J.A. Bondy and U.S.R. Murty, (Academic Press, New York, 1979), 219–231.
- [15] J. Han, 'On vertex-disjoint paths in regular graphs', *Electron. J. Comb.* **25** (2018), Paper P2.12.

- [16] B. Jackson, 'Long paths and cycles in oriented graphs', *J Graph Theory* **5** (1981), 245–252.
- [17] P. Keevash, D. Kühn, and D. Osthus, 'An exact minimum degree condition for Hamilton cycles in oriented graphs', *J. Lond. Math. Soc.* **79** (2009), 144–166.
- [18] M. Kouider, 'Covering vertices by cycles', *J Graph Theory* **18** (1994) 757–776.
- [19] M. Kouider and Z. Lonc, 'Covering cycles and k-term degree sums', *Combinatorica*, **16** (1996), 407–412.
- [20] M. Kouider and M. Zamime, 'On the path partition of graphs', [arXiv:2212.12793](https://arxiv.org/abs/2212.12793) (2022).
- [21] D. Kühn, A. Lo, D. Osthus, and K. Staden, 'The robust component structure of dense regular graphs and applications', *Proc. Lond. Math. Soc.* **110**(1) (2015), 19–56.
- [22] D. Kühn and D. Osthus, 'Hamilton decompositions of regular expanders: a proof of Kelly's conjecture for large tournaments', *Adv. Math.* **237** (2013), 62–146.
- [23] D. Kühn, D. Osthus, and A. Treglown, 'Hamiltonian degree sequences in digraphs', *J. Combin. Theory. Ser. B* **100** (2010), 367–380.
- [24] M. Las Vergnas, 'Problèmes de couplages et problèmes hamiltoniens en Théorie des graphes', Ph.D. Thesis, University of Paris, (1972).
- [25] A. Lo and V. Patel, 'Hamilton cycles in sparse robustly expanding digraphs', *Electron. J. Comb.* **25**(3) (2018), P3.44
- [26] A. Lo., V. Patel, J. Skokan, and J. Talbot, 'Decomposing tournaments into paths', *Proc. Lond. Math. Soc.* **121**(2) (2020), 426–461.
- [27] A. Lo, V. Patel, and M. A. Yıldız, 'Hamilton cycles in dense regular digraphs and oriented graphs', *J. Combin. Theory. Ser. B* **164** (2024), 119–160.
- [28] C. Magnant and D. Martin, 'A note on the path cover number of regular graphs', *Australas. J. Comb.* **43** (2009), 211–217.
- [29] C. Magnant, H. Wang, and S. Yuan, 'Path partitions of almost regular graphs', *Australas. J. Comb.* **64** (2016), 334–340.
- [30] O. Ore, 'Arc coverings of graphs', *Annali di Matematica Pura ed Applicata* **55** (1961), 315–321.
- [31] B. Reed, 'Paths, stars and the number three', *Comb. Probab. Comput.* **3** (1996), 277–295.
- [32] F. Ruskey and C. D. Savage, 'Hamilton cycles that extend transposition matchings in Cayley graphs of S_n ', *SIAM J. Discrete Math.* **6**(1) (1993), 152–166.
- [33] K. Staden, 'Robust expansion and Hamiltonicity', PhD Thesis, University of Birmingham (2015).
- [34] V.G. Vizing, 'On an estimate of the chromatic class of a p-graph', *Diskret. Analiz.* **3** (1964), 25–30.
- [35] Z. Yang, 'On F -Hamiltonian graphs', *Discrete Math.* **196** (1999), 281–286.
- [36] G. Yu, 'Covering 2-connected 3-regular graphs with disjoint paths' *J Graph Theory* **88** (2018), 385–401.