ROBUST EXPANSION AND HAMILTONICITY

by

KATHERINE STADEN

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

School of Mathematics
College of Engineering and Physical Sciences
The University of Birmingham
January 2015

Abstract

This thesis contains four results in extremal graph theory relating to the recent notion of robust expansion, and the classical notion of Hamiltonicity. In Chapter 2 we prove that every sufficiently large 'robustly expanding' digraph which is dense and regular has an approximate Hamilton decomposition. This provides a common generalisation of several previous results and in turn was a crucial tool in Kühn and Osthus's proof that in fact these conditions guarantee a Hamilton decomposition, thereby proving a conjecture of Kelly from 1968 on regular tournaments.

In Chapters 3 and 4, we prove that every sufficiently large 3-connected D-regular graph on n vertices with $D \geq n/4$ contains a Hamilton cycle. This answers a problem of Bollobás and Häggkvist from the 1970s. Along the way, we prove a general result about the structure of dense regular graphs, and consider other applications of this.

Chapter 5 is devoted to a degree sequence analogue of the famous Pósa conjecture. Our main result is the following: if the ith largest degree in a sufficiently large graph G on n vertices is at least a little larger than $n/3+i$ for $i \leq n/3$, then G contains the square of a Hamilton cycle.

CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

To my supervisors, Deryk Osthus and Daniela Kühn: my most sincere gratitude for your patience, dedication and knowledge; and for introducing me to such a wonderful area of mathematics.

To my parents, to Luke and to Albert: thank you for your love and support.

To my friends in the maths department: for coffee breaks, light relief and much more.

CHAPTER 1

INTRODUCTION

The topic of this thesis is extremal graph theory. The objects of study are (undirected and directed) graphs, discrete structures which are used to model various real-world systems and with many applications in computer science and other areas. A graph is a collection of vertices, wherein pairs of vertices are joined by at most one edge. Attempting to understand, characterise and describe the behaviour of these simple structures has given rise to a great many deep and difficult questions.

The sorts of problems which define graph theory are numerous and varied. Given a class $\mathcal{G}$ of graphs and a property $\mathcal{P}$, we might ask *how many* graphs in $\mathcal{G}$ have $\mathcal{P}$? If we choose a member of $\mathcal{G}$ at random, according to some distribution, can we say that $\mathcal{P}$ holds with high probability? In contrast to considering the 'typical behaviour' of graphs in relation to a property, *extremal* questions consider the 'worst possible' case. How big or small must a graph invariant be to *guarantee* a particular property $\mathcal{P}$?

> Given a graph invariant μ, what is the least m such that every $G \in \mathcal{G}$ with $\mu(G) \geq m$ has property $\mathcal{P}$?

For example, how many edges guarantee the existence of a triangle? What minimum degree guarantees a Hamilton cycle? These particular questions were answered many decades ago, but continue to motivate many natural, fascinating and indeed surprising problems in modern combinatorics, some of which I will address in this thesis.

1.1 Sufficient conditions for Hamilton cycles

A *Hamilton cycle* is a cycle in which every vertex appears exactly once. It is a very natural object which has motivated a huge body of research since the birth of graph theory. The decision problem of finding a Hamilton cycle is NP-complete, and appears on Karp's original list of NP-complete problems [69]. Thus it is unlikely that there exists an efficient algorithm to solve it. Instead, one can focus on finding sufficient conditions which guarantee the existence of a Hamilton cycle. Ideally we want sufficient conditions which are easy to compute and which hold for a large class of graphs. Furthermore, we aim to find conditions which are *best possible*.

Now begins a brief survey of extremal results on Hamilton cycles and related problems. For a far more comprehensive exploration of this area, we direct the reader to e.g. [55] or [83].

1.1.1 Vertex degree conditions

Many of the most fundamental extremal results give, for various H, a bound on the number of edges of G in terms of the number of vertices that guarantee $H \subseteq G$. Indeed, for graphs G on n vertices, $\lfloor n^2/4 \rfloor$ edges guarantee a triangle (Mantel's theorem, 1907), while $\binom{n-1}{2}+2$ edges are required to guarantee a Hamilton cycle (due to Ore and, independently, Bondy; see e.g. [22]). In the case of spanning subgraphs (such as the Hamilton cycle), it makes more sense to consider conditions which guarantee many edges at every vertex. All of the results proved in this thesis include some *vertex degree* condition. The simplest such condition is a *minimum degree* condition: a lower bound on $\delta(G)$. Dirac [40] proved the following:

Theorem 1.1.1. *Let G be a graph on $n \geq 3$ vertices with $\delta(G) \geq n/2$. Then G contains a Hamilton cycle.*

One can immediately see that this result is tight by considering two equal-sized cliques which intersect at one vertex, or the almost balanced complete bipartite graph. Ore [97]

generalised this by giving a condition on the sums of degrees of non-adjacent vertices. Another more refined degree condition is a *degree sequence condition*, which gives a lower bound on the ith largest degree d_i. Pósa [100] proved that if $d_1 \leq \ldots \leq d_n$ is the degree sequence of a graph G where $d_i \geq i+1$ for all $i < (n-1)/2$ and if additionally $d_{\lceil n/2 \rceil} \geq \lceil n/2 \rceil$ when n is odd, then G contains a Hamilton cycle. This is significantly stronger than Dirac's theorem, as almost half the vertices of G can have degree less than $n/2$. Pósa's theorem was generalised by Chvátal [35], who characterised those degree sequences which guarantee a Hamilton cycle. Finally, Bondy and Chvátal [21] provided a generalisation of all of these results by proving that a graph is Hamiltonian if and only if its closure is Hamiltonian. (The *closure* $cl(G)$ of a graph G is obtained by exhaustively adding an edge between pairs of non-adjacent vertices whose degree sum is at least $|G|$.)

1.1.2 Cycles of different lengths

We now consider the problem of finding cycles of many different lengths. A graph is *weakly pancyclic* if it contains a cycle of every length from its *girth* (length of shortest cycle) to the *circumference* (longest cycle). Brandt [25] showed that every non-bipartite graph G on n vertices with more than $\lfloor (n-1)^2/4 + 1 \rfloor$ edges is weakly pancyclic. He conjectured that actually $(n-1)(n-3)/4+4$ edges should suffice. This was nearly solved by Bollobás and Thomason [16]. If G has the stronger property of containing every cycle of length $3 \leq \ell \leq n$, we say it is *pancyclic*. Bondy [17] generalised Dirac's theorem in this direction by showing that any graph $G \neq K_{n/2,n/2}$ on $n \geq 3$ vertices with $\delta(G) \geq n/2$ is pancyclic. He proposed a striking metaconjecture [18] which would generalise his result: almost any non-trivial condition on a graph which implies Hamiltonicity in fact implies pancyclicity (apart from maybe a simple family of exceptions).

What about the problem of finding a 2-*factor*, a spanning collection of vertex-disjoint cycles (of which the Hamilton cycle is one example)? A conjecture of El-Zahar [41] from the 80s was that, for an n-vertex graph G, if $\delta(G) \geq \lceil k_1/2 \rceil + \ldots + \lceil k_\ell/2 \rceil$, then G contains the vertex-disjoint union of cycles $C_{k_1} \cup \ldots \cup C_{k_\ell}$, where $k_1 + \ldots + k_\ell = n$. In

this direction, Aigner and Brandt [2] showed that minimum degree $(2n-1)/3$ guarantees that G contains any graph on at most n vertices and maximum degree 2. This is a special case of a conjecture of Bollobás, Eldridge [14] and Catlin [29]: if G_1 and G_2 are n-vertex graphs such that $(\Delta(G_1)+1)(\Delta(G_2)+1) < n+1$, then G_1 and G_2 are edge-disjoint subgraphs of the complete graph on n vertices. We prove a degree sequence version of the result of Aigner and Brandt in Chapter 5, which is an approximate generalisation. El-Zahar's conjecture was eventually proved in 1999 for large n in the PhD thesis of Abbasi [1].

1.1.3 Digraphs

Extra complexity becomes apparent when one turns to digraphs (where we seek cycles whose edges are consistently oriented). Here, one analogue of a minimum degree condition is a *minimum semidegree condition*, which stipulates that every vertex in G has in- and outdegree at least some k (which we write as $\delta^0(G) \geq k$). Ghouila-Houri [53] proved that every strongly connected digraph on n vertices has a Hamilton cycle if the sum of in- and outdegrees at every vertex is at least n. (Here, a graph G is *strongly connected* if for every pair x, y of vertices, there is a (directed) path in G from x to y, and from y to x.) In particular, $\delta^0(G) \geq n/2$ is sufficient. Ore-type analogues were obtained by Woodall [115] and Meyniel [92].

The situation changes when one considers oriented graphs (digraphs in which 2-cycles are forbidden).

Theorem 1.1.2. (Keevash, Kühn and Osthus) [70] *There exists $n_0 \in \mathbb{N}$ such that every oriented graph G on $n \geq n_0$ vertices with $\delta^0(G) \geq (3n-4)/8$ contains a Hamilton cycle, and this bound is best possible.*

Strong connectivity is necessary for the presence of a Hamilton cycle in a digraph; and in fact Camion [28] proved that it is sufficient in a tournament (a *tournament* is an orientation of a complete graph). Thomassen [111] asked if a 'stronger' connectivity

condition actually guarantees more in a tournament, i.e. the existence of many edge-disjoint Hamilton cycles. This was verified by Kühn, Lapinskas, Osthus and Patel [74], whose bounds were later improved by Pokrovskiy [99]. Observe that, even though no degree condition is assumed, such a (weak) condition is implied by high connectivity.

In another direction, suppose we wish to find not only a consistently oriented Hamilton cycle in a digraph, but also all other orientations of a Hamilton cycle. A best possible minimum semidegree condition for this was provided by DeBiasio, Kühn, Molla, Osthus and Taylor [39]. Interestingly, the cycle whose orientations alternate requires a higher semidegree bound than any other cycle.

1.1.4 Hypergraphs

There has been much interest in Hamilton cycles in *hypergraphs*, which are generalisations of graphs. Here, edges are not necessarily pairs, but may consist of larger sets of vertices. Now there are notions of degree for sets of vertices, not just singletons; and edges may intersect in more than one vertex, giving rise to different notions of 'cycle'.

Suppose that H is a hypergraph on n vertices in which every edge is a set of exactly k vertices (*k-uniform*), and we are interested in finding an *ℓ-cycle*, in which consecutive edges overlap in exactly ℓ vertices. Let us also sensibly generalise the notion of degree – given $S \subseteq V(H)$, we let $d_H(S)$ be number of edges of H containing S as a subset. Then define $\delta_t(H)$ to be the minimum $d_H(S)$ taken over all S with $|S| = t$.

One can considerably broaden the questions we have posed for graphs by asking them in the more general hypergraph setting. For example, what is the least $\delta_t(H)$ that guarantees a Hamilton ℓ-cycle in H?

The case $(k, \ell, t) = (2, 1, 1)$ is Dirac's theorem. Recently, the case $(k, \ell, k-1)$ was solved asymptotically in a series of papers by various authors [77, 102, 103]. However, even the conjectured bound [101] for the very natural case $(k, k-1, 1)$ remains unproven.

To say more would be outside the scope of this thesis; we refer the interested reader to [101].

1.2 Hamilton decompositions

Given a graph or digraph G, we now wish to find not one but several edge-disjoint Hamilton cycles. Suppose we actually want a *Hamilton decomposition* – a collection of edge-disjoint Hamilton cycles which together contain every edge of G.

The study of Hamilton decompositions began over one hundred years ago when Walecki (see [8]) proved that the complete graph K_n on n vertices has a Hamilton decomposition if and only if n is odd. Tillson [112] showed that complete digraphs on $n \neq 4, 6$ vertices have a Hamilton decomposition. Until recently, little else was known for more general classes of (di)graphs. Observe that being regular is a necessary condition for a digraph to have a Hamilton decomposition (for a graph, being regular of even degree is necessary). In 1968, Kelly (see e.g. [93]) conjectured that, in fact, for tournaments, this is also sufficient: every regular tournament has a Hamilton decomposition. This was proved in 2013 by Kühn and Osthus [81], for large graphs.

Theorem 1.2.1. (Kühn and Osthus) [81] *There exists $n_0 \in \mathbb{N}$ such that for all regular tournaments T on $n \geq n_0$ vertices, T has a Hamilton decomposition.*

In fact their result was much more general. They showed that a particular structural property of digraphs (which is possessed by regular tournaments) guarantees a Hamilton decomposition. This property is *robust expansion*, which turns out to have a close connection with Hamiltonicity.

1.3 Robust expansion

Expansion is a familiar concept in graph theory. Roughly speaking, a graph is an *expander* if each set of vertices has neighbourhood larger than itself. *Robust expansion* asks for something stronger. Let G be a digraph on n vertices and suppose $S \subseteq V(G)$. The *ν-robust outneighbourhood* $RN^+_{\nu,G}(S)$ of $S \subseteq V(G)$ is the set of all vertices in G with at

least νn inneighbours in S. We say that G is a *robust* (ν, τ)*-outexpander* if

$$|RN^+_{\nu,G}(S)| \geq |S| + \nu n \quad \text{for all} \quad S \subseteq V(G) \quad \text{with} \quad \tau n \leq |S| \leq (1-\tau)n.$$

There is an analogous notion of *robust expansion* for graphs. Why is this a useful notion to study? Firstly, it is a property shared by many widely-studied classes of *dense* graphs, i.e. those in which each vertex is adjacent to some positive proportion of the total number of vertices. Robustly expanding (di)graphs include oriented graphs G with minimum degree at least slightly larger than $3|G|/8$, quasirandom (di)graphs (e.g. (n, d, λ)-graphs for appropriate values of these parameters) and dense random regular graphs. In particular, every regular tournament is a robust outexpander. Secondly, robust expansion has been used as an essential concept in the recent solution of several longstanding conjectures, including an implicit use in Kühn, Mycroft and Osthus's resolution of Sumner's universal tournament conjecture [78, 79].

Chapter 2 is dedicated to the proof of the following theorem, which states that every sufficiently large regular robust outexpander has an approximate Hamilton decomposition.

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

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

(ii) G is a robust (ν, τ)-outexpander.

Then G contains at least $(1-\eta)r$ edge-disjoint Hamilton cycles. Moreover, such a set of Hamilton cycles can be found in time polynomial in n.

This was a crucial tool in Kühn and Osthus's proof that the same hypotheses in fact guarantee a Hamilton decomposition. As noted above, this proves Kelly's conjecture for large graphs. A version of our result for regular tournaments was proved by Kühn, Osthus and Treglown [86].

Theorem 1.1.2 told us that minimum semidegree almost $3n/8$ was sufficient to guarantee a Hamilton cycle in an oriented graph on n vertices. But the result of Kühn and Osthus implies that, for regular oriented graphs, degree just a little larger than this in fact guarantees a Hamilton decomposition. However, in the case of regular oriented graphs, Jackson [64] conjectures that degree just less than $n/4$ should suffice for the existence of a single Hamilton cycle.

One consequence of Kühn and Osthus's result is that every large even D-regular graph G on n vertices with degree D at least slightly larger than $n/2$ has a Hamilton decomposition. Together with Csaba, Lo and Treglown [37], they recently showed that actually $D \geq \lfloor n/2 \rfloor$ suffices, thus answering a question of Nash-Williams [94, 95].

This result is particularly striking – the threshold at which a single Hamilton cycle appears [96] matches the threshold which guarantees a decomposition. It improves upon work in [34, 62].

1.4 Below the threshold

Having considered one way of extending Dirac's theorem by looking for not one but many Hamilton cycles, we now turn to a different extension. It is natural to ask the following: how might one decrease the degree bound in Dirac's theorem at the expense of introducing some extra conditions? Alternatively, is there some additional barrier to Hamiltonicity just below the degree threshold?

This type of question has been asked many times in different extremal contexts. Recall our initial question: given a property $\mathcal{P}$, a graph invariant μ and a class $\mathcal{G}$ of graphs, what is the least m such that every $G \in \mathcal{G}$ with $\mu(G) \geq m$ has property $\mathcal{P}$? Having answered this, we can delve deeper: is there a property $\mathcal{P}'$ and an $m' < m$ such that every $G \in \mathcal{G}$ with $\mu(G) \geq m'$ and property $\mathcal{P}'$ also has property $\mathcal{P}$? In other words, just below the threshold of m, is there some reason (i.e. the absence of $\mathcal{P}'$) that prevents $\mathcal{P}$?

We give an example to illustrate this. Let $\mathcal{P}$ be the property of containing a triangle,

let μ be minimum degree δ, and let $\mathcal{G}$ be the class of all graphs on n vertices. Mantel's theorem states that every $G \in \mathcal{G}$ with $\mu(G) > n/2$ has property $\mathcal{P}$. Now let $\mathcal{P}'$ be the property of having no bipartition. Andrásfai, Erdős and Sós [10] proved that every $G \in \mathcal{G}$ with $\mu(G) > 2n/5$ and property $\mathcal{P}'$ also has property $\mathcal{P}$.

1.5 Hamilton cycles in dense regular graphs

We can now return to the case when $\mathcal{P}$ is the property of containing a Hamilton cycle. The well-known extremal examples for Dirac's theorem of two equal-sized cliques which intersect in one vertex and the almost balanced complete bipartite graph suggest what $\mathcal{P}'$ must be: the property of being regular and having a sufficiently strong connectivity requirement.

In this spirit, Szekeres (see [63]) asked for which $D = D(n)$ does every 2-connected regular graph G on n vertices with degree at least D contain a Hamilton cycle. This question was answered by Jackson [63] who showed that $D \geq n/3$ suffices. So in fact the degree threshold decreases dramatically with the additional assumptions of 2-connectivity and regularity. Bollobás and Häggkvist [13, 58] had, earlier and independently, posed a striking and natural conjecture which is a generalisation to t-connected graphs: every t-connected D-regular graph on n vertices with $D \geq n/(t+1)$ contains a Hamilton cycle. An example of Jung [68] and independently Jackson, Li and Zhu [66] showed this to be false for $t \geq 4$; but if true for $t = 3$ then this would be best possible. In Chapter 4, we prove this completely for large n, thereby verifying the only remaining case of Bollobás and Häggkvist's conjecture. It is indeed surprising that the relationship between D and t should abruptly end at $t = 4$.

Theorem C. *There exists $n_0 \in \mathbb{N}$ such that every 3-connected D-regular graph on $n \geq n_0$ vertices with $D \geq n/4$ contains a Hamilton cycle.*

1.6 The structure of dense regular graphs

The most vital tool in our proof of Theorem C is Theorem B, a structural result for dense regular graphs, which is the subject of Chapter 3. We have seen that the class of robust expanders is very rich indeed. Despite this, there are many dense regular graphs which do not have the property. How can one harness the powerful properties of robust expansion in a general dense regular graph G? Roughly speaking, Theorem B states that such a G has a remarkably simple structure – it can be partitioned into a small number of robust components, each of which has strong expansion properties. We do not explicitly state Theorem B here, as it requires a number of technical definitions.

1.7 Powers of Hamilton cycles

There is a long history of embedding spanning structures in graphs. This, unsurprisingly, is considerably harder than finding a non-spanning structure. We have discussed the case when the spanning structure H is a Hamilton cycle. There are many results for other H. For example, Hajnal and Szemerédi [59] proved that every graph G on $n \in r\mathbb{N}$ vertices with $\delta(G) \geq (r-1)n/r$ contains a perfect K_r-packing; that is, a collection of vertex-disjoint cliques on r vertices. This result is best possible.

The rth *power* H^r of a graph H is obtained from H by adding additional edges between every pair of vertices at distance at most r. (The adjacency matrix of H^r is obtained by normalising the rth power of the adjacency matrix of H.) In the 1960s, Pósa [43] conjectured that, in a graph on n vertices, minimum degree $2n/3$ guarantees C_n^2, the square of a Hamilton cycle. This was strengthened by Seymour [105], who suggested that, for all $r \geq 1$, minimum degree $rn/(r+1)$ guarantees the rth power of a Hamilton cycle C_n^r. This degree bound would be best possible. Observe that, when $n \in r\mathbb{N}$, C_n^r contains a perfect K_{r+1}-packing. So the Pósa-Seymour conjecture is strictly stronger than the Hajnal-Szemerédi theorem. After intensive study in the 1990s (see

e.g. [46, 47, 48, 49, 50]), Komlós, Sárközy and Szemerédi [73] were able to prove the conjecture for large n.

In Chapter 5, we consider a degree sequence analogue of Pósa's conjecture. Can we still find the square of a Hamilton cycle in a graph in which many of the vertices have degree significantly lower than $2n/3$? Our main result is the following:

Theorem D. *Given any $\eta > 0$ there exists an $n_0 \in \mathbb{N}$ such that the following holds. If G is a graph on $n \geq n_0$ vertices whose degree sequence $d_1 \leq \cdots \leq d_n$ satisfies*

$$d_i \geq n/3 + i + \eta n \quad \textit{for all } i \leq n/3,$$

then G contains the square of a Hamilton cycle.

Note that Theorem D allows for almost $n/3$ vertices in G to have degree substantially smaller than $2n/3$. However, it does not quite imply Pósa's conjecture for large graphs due to the term ηn. Up to this term, examples show that the result is best possible.

Due to space considerations, we are only able to present a sketch of the proof of Theorem D. The full proof appears in [107].

1.8 Tools

We now briefly describe some of the tools used to prove the results in this thesis. Each chapter contains the definitions and precise statements of the tools used within.

One tool used either explicitly or implicitly in every chapter is Szemerédi's Regularity lemma [108]. Proved in the 1970s, it states that, in every sufficiently large graph, the vertex set can be partitioned into a constant number of parts of roughly equal size, so that the edge distribution between any pair exhibits random-like behaviour (the pairs are *ε-regular*). So it reduces the problem of embedding into deterministic structures to the easier problem of embedding into random-like objects. The lemma has been used in the proofs of many results in many areas of combinatorics, from its original appearance as

an auxiliary lemma in Szemerédi's theorem on arithmetic progressions [109] to a great many extremal and Ramsey-Turán problems.

Given an input parameter ε (which can be thought of an an approximation error), the original proof gives that the number of parts has tower-type dependence on ε. Unfortunately, as proved by Gowers [56], one cannot do much better than this. So the lemma only draws useful conclusions for graphs which are very large indeed.

There have been some attempts in recent years to reprove results which were originally proved using the Regularity lemma, sometimes to obtain proofs applicable to smaller graphs. Levitt, Sárkozy and Szeméredi [87] found a proof of Pósa's conjecture for graphs on $n \geq n_0$ vertices which avoids the Regularity lemma. Châu, DeBiasio and Kierstead [31] were able to find an explicit bound for n_0 by proving that $n_0 \geq 2 \times 10^8$ suffices. This number is large, but is nonetheless tiny compared to the regularity-sized bound given in the original proof by Komlós, Sárközy and Szemerédi [73].

Their proof uses the *Connecting-Absorbing* technique, first introduced by Rödl, Ruciński and Szemerédi [104]. Roughly speaking, absorption allows one to turn an almost spanning structure into a spanning structure. In [81], a complicated version of absorbing was employed by Kühn and Osthus to strengthen Theorem A and prove Kelly's conjecture. Theorem D is proved using an amalgam of the Connecting-Absorbing and Regularity methods.

The proof of Theorem D also uses the so-called Blow-up lemma of Komlós, Sárközy and Szemerédi [72]. This states that, with regard to embedding spanning graphs of bounded degree, ε-regular pairs behave like complete bipartite graphs.

Probabilistic arguments have been a cornerstone of combinatorial proofs since they were first employed by Erdős to prove a lower bound for the Ramsey number $R(s, s)$ [42]. We only employ basic probabilistic tools in this thesis, e.g. to find subgraphs which inherit the properties of their host graph.

The proofs of all four of our main results are very long. We develop tailored machinery to solve each of the problems we consider. However, we hope that some of our

methods (e.g. the main result of Chapter 3) will be of independent interest and will find applications elsewhere.

1.9 Organisation of the thesis

The main content of this thesis lies in Chapters 2–5. The first three of these chapters contain the proofs of Theorems A–C respectively; while in Chapter 5, we sketch the proof of Theorem D. Each chapter is self-contained, and the relevant notation is stated near the beginning of the chapter. (Much of the notation in Chapter 4 is carried over from Chapter 3, so it will not be restated, but the reader is referred back to the appropriate sections from Chapter 3.)

Chapter 2 is based on joint work [98] with Osthus, Chapters 3,4 on joint work [75, 76] with Kühn, Lo and Osthus, and Chapter 5 on joint work [107] with Treglown.

CHAPTER 2

APPROXIMATE HAMILTON DECOMPOSITIONS OF ROBUSTLY EXPANDING REGULAR DIGRAPHS

2.1 Introduction

A Hamilton decomposition of a graph or digraph G is a set of edge-disjoint Hamilton cycles which together cover all the edges of G. The first result in the area was proved by Walecki in 1892, who showed that a complete graph K_n has a Hamilton decomposition if and only if n is odd (see e.g. [91], [8], [9]). Tillson [112] solved the analogous problem for complete digraphs in 1980. Though the area is rich in beautiful conjectures, until recently there were few general results.

Starting with a result of Frieze and Krivelevich [51], a very successful recent direction of research has been to find 'approximate' Hamilton decompositions, i.e. a set of edge-disjoint Hamilton cycles which cover almost all the edges of the given (di)graph. The result in [51] concerns dense quasirandom graphs and digraphs. Hypergraph versions of this result were proved by Frieze, Krivelevich and Loh [52] as well as Bal and Frieze [11]. Also, Kühn, Osthus and Treglown [86] proved an approximate version of Kelly's conjecture. This long-standing conjecture (see [93]) states that every regular tournament has a Hamilton decomposition. In fact, the result in [86] is much more general, namely it states that every regular oriented graph on n vertices whose in- and outdegree is slightly

larger than $3n/8$ has an approximate Hamilton decomposition. Here an oriented graph is a digraph with at most one edge between each pair of vertices (whereas a digraph may have one edge in each direction between a pair of vertices).

The main result of this chapter is in turn a far reaching generalisation of the result in [86]. Instead of a degree condition, it involves an expansion condition that has recently been shown to have a close connection with Hamiltonicity. This notion was introduced by Kühn, Osthus and Treglown in [85]. The condition states that for every set S which is not too small and not too large, its 'robust' outneighbourhood is at least a little larger than S itself. More precisely, suppose that G is a digraph of order n and $S \subseteq V(G)$. The *ν-robust outneighbourhood* $RN^+_{\nu,G}(S)$ of S is the set of vertices with at least νn inneighbours in S. We say that G is a *robust (ν, τ)-outexpander* if

$$|RN^+_{\nu,G}(S)| \geq |S| + \nu n \quad \text{for all} \quad S \subseteq V(G) \quad \text{with} \quad \tau n \leq |S| \leq (1-\tau)n.$$

Our main result states that every sufficiently large robustly outexpanding regular digraph has an approximate Hamilton decomposition.

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

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

(ii) *G is a robust (ν, τ)-outexpander.*

Then G contains at least $(1-\eta)r$ edge-disjoint Hamilton cycles. Moreover, such a set of Hamilton cycles can be found in time polynomial in n.

As observed in Lemma 12.1 of [81], every oriented graph whose in- and outdegrees are all at least slightly larger than $3n/8$ is a robust outexpander, so this does generalise the main result of [86]. Moreover, it turns out that one can relax condition (i) to the requirement that G is 'almost regular'. This is due to the fact (observed in [81]) that

every almost regular robustly expanding digraph contains a spanning regular digraph of similar degree.

Corollary 2.1.1. *For every $\alpha > 0$ there exists $\tau > 0$ such that for all $\nu, \eta > 0$ there exist $n_0 = n_0(\alpha, \nu, \tau, \eta)$ and $\gamma = \gamma(\alpha, \nu, \tau, \eta) > 0$ for which the following holds. Suppose that*

(i) G is a digraph on $n \geq n_0$ vertices with $(\alpha - \gamma)n \leq d_G^{\pm}(x) \leq (\alpha + \gamma)n$ for every x in G;

(ii) G is a robust (ν, τ)-outexpander.

Then G contains at least $(\alpha - \eta)n$ edge-disjoint Hamilton cycles. Moreover, this set of Hamilton cycles can be found in time polynomial in n.

The result in [86] extends to almost regular oriented graphs in the same way, but is inherently non-algorithmic (see Section 2.2). Since, for dense digraphs, the condition of being a robust outexpander is much weaker than that of being quasirandom, Corollary 2.1.1 is much more general than the result in [51] mentioned earlier. Moreover, it is best possible in the sense that, for an almost regular digraph, an approximate Hamilton decomposition is obviously the best one can hope for.

Theorem A is used as an essential tool by Kühn and Osthus in [81] to prove the following result, which (under the same conditions) guarantees not only an approximate decomposition, but a Hamilton decomposition.

Theorem 2.1.2. *For every $\alpha > 0$ there exists $\tau > 0$ such that for every $\nu > 0$ there exists $n_0 = n_0(\alpha, \nu, \tau)$ for which the following holds. Suppose that*

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

(ii) *G is a robust (ν, τ)-outexpander.*

Then G has a Hamilton decomposition. Moreover, this decomposition can be found in time polynomial in n.

So as a special case, Theorem 2.1.2 implies that Kelly's conjecture holds for all sufficiently large regular tournaments. It also implies a conjecture of Erdős on packing Hamilton cycles in random tournaments (see [110]). However, it turns out that the notion of robust (out)expansion extends far beyond the class of tournaments and many further applications of Theorem A are explored by Kühn and Osthus in [82]. For example, the notion of robust expansion can be extended to undirected graphs in a natural way and one can deduce a version of Theorem 2.1.2 for undirected graphs. In [82] this in turn is used to prove an approximate version of a conjecture of Nash-Williams on Hamilton decompositions of dense regular graphs. Random regular graphs of linear degree as well as (n, d, λ)-graphs (for appropriate values of these parameters) are further examples of robustly expanding graphs. In combination with a result of Gutin and Yeo [57], Theorem 2.1.2 can also be used to solve a problem of Glover and Punnen [54] as well as Alon, Gutin and Krivelevich [6] on TSP tour domination (see [81] for details). For this application, it is crucial that the Hamilton decomposition can be found in polynomial time.

Roughly speaking, the argument leading to Theorem 2.1.2 uses Theorem A in the following way: let G be a robustly expanding digraph. The first step is to remove a 'robustly decomposable' spanning regular digraph H from G to obtain G'. H will be sparse compared to G and will have the property that it has a Hamilton decomposition even if we add the edges of a digraph H', which is very sparse compared to H and also regular (on the same vertex set) but otherwise arbitrary. Now G' is still a robust outexpander, so one can apply Theorem A to G' obtain an approximate Hamilton decomposition of G'. Let H' denote the set of edges not contained in any of the Hamilton cycles of this approximate decomposition of G'. Then the fact that H is robustly decomposable implies that $H \cup H'$ has a Hamilton decomposition. Together with the approximate decomposition of G', this yields a Hamilton decomposition of the entire digraph G. Note that the above approach means that for Theorem 2.1.2 to be algorithmic, one needs Theorem A to be algorithmic too.

This chapter is organised as follows. In the next section, we give a brief outline of the argument. We then collect the necessary tools in Section 2.3 (which is mostly concerned with Szemerédi's Regularity lemma) and Section 2.4 (which mainly collects properties of robust outexpanders). We then prove Theorem A in Section 2.5. In Section 2.6, we deduce Corollary 2.1.1 from Theorem A.

2.2 Sketch of the proof of Theorem A

Roughly speaking, the strategy of the proof of Theorem A is the following. Suppose that a digraph G satisfies the conditions of Theorem A. First remove the edges of a carefully chosen spanning sparse subdigraph H from G and let G' consist of the remaining edges of G. Next, find an approximate decomposition of G' into edge-disjoint 1-factors F_i (where a 1-factor is a spanning union of vertex-disjoint cycles). Finally, the aim is to transform each F_i into a Hamilton cycle by removing some of its edges and adding some edges of H. One immediate obstacle to a naïve implementation of this approach is that the F_i might consist of many cycles, so turning each of them into a Hamilton cycle might require more edges from H than one can afford. In [51, 86], this was overcome (loosely speaking) by choosing the 1-factors F_i randomly. It turns out that this has the advantage that the F_i will have few cycles, i.e. they are already close to being Hamilton cycles. One disadvantage is that this approach is inherently non-algorithmic (and does not seem derandomisable).

A second problem is how to make sure that H contains the edges that are required to transform each F_i into a Hamilton cycle. We overcome this by choosing H and the 1-factors F_i according to the vertex partition of G obtained from Szemerédi's Regularity lemma. More precisely, we apply the Regularity lemma to partition G into clusters $V_1, \dots, V_L$ of vertices such that almost all ordered pairs of clusters induce a pseudorandom subdigraph of G, together with a small (but typically troublesome) exceptional set V_0. For some small constant β, we define the 'reduced multidigraph' $R(\beta)$ whose vertices are the clusters V_j with (multiple) edges from V_j to V_k if the corresponding subdigraph of G

is pseudorandom and dense. Here the number of edges from V_j to V_k is proportional to the density of $G[V_j, V_k]$. So each edge of $R(\beta)$ corresponds to a bipartite pseudorandom digraph between the corresponding pair of clusters in G (where all these pseudorandom digraphs have the same density β). $R(\beta)$ inherits many of the properties of G, in particular it is an almost regular robust outexpander with large minimum semidegree.

The next step is to use the Max-Flow-Min-Cut Theorem to find a spanning regular subdigraph of $R(\beta)$ which contains almost all edges of $R(\beta)$. We can now (arbitrarily) partition this regular subdigraph into a collection of edge-disjoint 1-factors F_i of $R(\beta)$ (see Section 2.5.1). Each of the F_i corresponds to a vertex-disjoint collection of 'blown-up' cycles which spans most of $V(G)$. We will denote each of these collections by G_i and call G_i the ith slice of G. Note that the G_i are all edge-disjoint.

Roughly speaking, the aim is to add a small number of edges (which do not lie in any of the other slices) to each G_i to transform G_i into a regular digraph which has an approximate Hamilton decomposition. Together, these approximate Hamilton decompositions of the slices then yield an approximate Hamilton decomposition of G. In Section 2.5.2, we put aside three sparse subdigraphs H_0, H_1, H_2 which we will use to add the required edges to each G_i. So together, H_0, H_1 and H_2 play the role of the digraph H mentioned earlier.

So far we have ignored the exceptional vertices, but to obtain a regular spanning subdigraph we need to incorporate them into each slice G_i. For convenience, we call any exceptional vertex $x \in V_0$ and each edge incident with V_0 'red'. In Sections 2.5.6 and 2.5.7, we will add red edges to each G_i in such a way that the resulting slice G_i is almost regular and only a small part of each cluster is incident to any red edges. Some of these edges come from H_1 and the others will be edges of G which are not contained in any of the H_j or any of the G_i constructed so far.

Together with these red edges, each G_i is now an almost regular digraph consisting mainly of a union of blown-up cycles. On the other hand, G_i may not even be connected. But to guarantee many edge-disjoint Hamilton cycles in G_i, we clearly need to have suf-

ficiently many edge-disjoint paths between these blown-up cycles. For this, we define an ordering of the cycles $D_1, \ldots, D_\ell$ of F_i and specify 'bridge vertices' $x_{i,j}$ (one for each successive pair of cycles) so that $x_{i,j}$ has many inneighbours in D_j and many outneighbours in D_{j+1}. We find the edges incident to these $x_{i,j}$ within H_0 (see Section 2.5.5).

We would now like to find a spanning regular subdigraph in each G_i whose degree is almost as large as that of G_i. Trivially, this regular subdigraph would then have a decomposition into 1-factors. However, as we have little control over the red edges added so far, they may prevent us from finding a regular subdigraph (see Section 2.5.8 for a discussion and an example). For this reason, we add extra (red) edges to G_i from H_2 to balance out the existing red edges. In this way, we can ensure that for each cluster V of a blown-up cycle D, the number of edges leaving V in G_i equals the number of edges entering its successor V^+ on D. This is achieved in Sections 2.5.8 and 2.5.9, by considering an auxiliary reduced digraph R^* which also turns out to be a robust outexpander (the latter property is crucial here).

As indicated above, in Section 2.5.10, we can now find a spanning κ-regular subdigraph G_i^* of each G_i (for a suitable κ). We now decompose each G_i^* into 1-factors $f_{i,1}, \ldots, f_{i,\kappa}$. Our aim is to transform each $f_{i,j}$ into a Hamilton cycle by adding and removing a few edges. The edges we add will be taken from a very sparse digraph $H_{3,i}$ which we removed from G_i earlier (so $H_{3,i}$ can also be viewed as a union of blown-up cycles). The key point of the proof is that we can achieve this transformation by using a very small number of edges from $H_{3,i}$ for each $f_{i,j}$. The reason for this is that we can guarantee that the red edges added in the course of the proof are 'localised' within each G_i, i.e. on each blown-up cycle of each F_i there are long intervals of clusters which are not incident to any red edges. This means that for each 1-factor $f_{i,j}$, its subdigraph induced by any such interval I consists of long paths. If some of these paths lie on different cycles of $f_{i,j}$, we can merge these into a single cycle by adding and removing edges of $H_{3,i}$ which are induced by just a single pair of consecutive clusters on I. Crucially, this enables us to use the bipartite subdigraphs of $H_{3,i}$ induced by other pairs of consecutive clusters on I to transform other

1-factors $f_{i,j'}$ of the slice G_i. Repeating this process until we have merged all cycles of $f_{i,j}$ into a single cycle eventually transforms the $f_{i,j}$ into κ edge-disjoint Hamilton cycles, as required (see Lemma 2.4.5 and Section 2.5.11).

An approach based on the Regularity lemma was already used in [86]. However, as mentioned earlier, the argument there relied on a random choice of the 1-factors, which did not translate into an algorithm. This problem is overcome by the above 'localisation' idea, which automatically produces 1-factors which are 'well behaved' with respect to red edges in the sense described above. However, this 'localisation property' is quite difficult to achieve and relies on additional ideas such as a refinement of the original regularity partition and a special 'unwinding' of blown-up cycles (see Section 2.5.3 and Lemma 2.4.4).

2.3 Notation and the Diregularity lemma

2.3.1 Notation

Throughout we will omit floors and ceilings where the argument is unaffected. The constants in the hierarchies used to state our results are chosen from right to left. For example, if we claim that a result holds whenever $0 < 1/n \ll a \ll b \ll c \leq 1$ (where n is the order of the graph or digraph), then there are non-decreasing functions $f : (0,1] \to (0,1]$, $g : (0,1] \to (0,1]$ and $h : (0,1] \to (0,1]$ such that the result holds for all $0 < a, b, c \leq 1$ and all $n \in \mathbb{N}$ with $b \leq f(c)$, $a \leq g(b)$ and $1/n \leq h(a)$. Hierarchies with more constants are defined in a similar way. Note that $a \ll b$ implies that we may assume in the proof that e.g. $a < b$ or $a < b^2$. We write $a = b \pm \varepsilon$ for $a \in [b - \varepsilon, b + \varepsilon]$.

For an undirected graph G containing a vertex x we write $N_G(x)$ for the neighbourhood of x and $d_G(x)$ for its degree. For a digraph G we write xy for the edge directed from x to y and write $N_G^+(x)$ for the *outneighbourhood*, the set of vertices receiving an edge from x, and write $d_G^+(x) := |N_G^+(x)|$ for the *outdegree* of x. We define the *inneighbourhood* $N_G^-(x)$

and *indegree* $d_G^-(x)$ similarly. For a collection of vertices $U \subseteq V(G)$ we write $d_G^+(U)$ for the total number of edges sent out by the vertices in U. We define $d_G^-(U)$ analogously. We will omit the G subscript in the above and in similar situations elsewhere if this is unambiguous. Denote the minimum outdegree by $\delta^+(G)$ and the minimum indegree by $\delta^-(G)$. Let the *minimum semidegree* $\delta^0(G)$ be the minimum of $\delta^+(G)$ and $\delta^-(G)$. Denote the maximum outdegree by $\Delta^+(G)$ and define $\Delta^-(G)$ and analogously. Let $\Delta^0(G)$ denote the maximum of $\Delta^+(G)$ and $\Delta^-(G)$. If G is a multidigraph then neighbourhoods are multisets. For any positive integer r, an *r-regular* digraph on n vertices is such that every vertex has exactly r outneighbours and r inneighbours. A *1-factor* of a multidigraph G is a 1-regular spanning digraph; that is, a collection of vertex-disjoint cycles that together contain all the vertices of G.

If G is a multidigraph and $U \subseteq V(G)$, we write $G[U]$ for the sub-multidigraph of G *induced* by U. That is, the digraph with vertex set U and edge set obtained from $E(G)$ by including only those edges with both endpoints contained in U. If $G[U]$ has empty edge set, we say that U is an *isolated* subset of G. If G is a digraph and $U \subseteq V(G)$ we write $G \setminus U$ for the digraph with vertex set $V(G) \setminus U$ and edge set obtained from $E(G)$ by deleting all edges incident to a vertex of U.

Given a digraph R and a positive integer r, the *r-fold blow-up* $r \otimes R$ of R is the digraph obtained from R by replacing every vertex x of R by r vertices and replacing every edge xy of R by the complete bipartite graph $K_{r,r}$ between the two sets of r vertices corresponding to x and y such that all the edges of $K_{r,r}$ are oriented towards the r vertices corresponding to y. We say that any edge in this $K_{r,r}$ is *contained in the blow-up of xy*. Now consider the case when $V_1, \ldots, V_k$ is a partition of some set V of vertices and R is a digraph whose vertices are $V_1, \ldots, V_k$. If R is a directed cycle, say $R = C = V_1 \ldots V_k$, and G is a digraph with $V(G) \subseteq V = V_1 \cup \ldots \cup V_k$, we say that *(the edges of) G wind(s) around C* if, for every edge xy of G, there exists an index j such that $x \in V_j$ and $y \in V_{j+1}$ (where indices are taken modulo k).

2.3.2 A Chernoff bound and its derandomisation

In the proof of Claims 2.5.3 and 2.5.4, we will use the following standard Chernoff type bound (see e.g. Corollary 2.3 in [67] and Theorem 2.2 in [106]).

Proposition 2.3.1. *Suppose X has binomial distribution and $0 < a < 1$. Then*

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

To obtain an algorithmic version of Theorem A, we need to 'derandomise' our applications of Proposition 2.3.1. This can be done via the well known 'method of conditional probabilities.' The following result of Srivastav and Stangier (Theorem 2.10 in [106]) provides a convenient way to apply this method. It implies that any construction based on a polynomial number of applications of Proposition 2.3.1 can be derandomised to provide a polynomial time algorithm.

Theorem 2.3.2. *[106] Let $X_1, \ldots, X_N$ be independent $0/1$ random variables, where $\mathbb{P}(X_j = 1) = p$ and $\mathbb{P}(X_j = 0) = 1 - p$ for some rational $0 \leq p \leq 1$. Suppose that $1 \leq i \leq m$ and let $w_{ij} \in \{0, 1\}$. Let $\phi_i := \sum_{j=1}^{N} w_{ij} X_j$ and fix β_i with $0 < \beta_i < 1$. Let E_i^+ denote the event that $\phi_i \geq (1+\beta_i)\mathbb{E}[\phi_i]$ and let E_i^- denote the event that $\phi_i \leq (1-\beta_i)\mathbb{E}[\phi_i]$. Let $E_i \in \{E_i^+, E_i^-\}$. Suppose further that*

$$\sum_{i=1}^{m} e^{-\beta_i^2 \mathbb{E}(\phi_i)/3} \leq 1/2.$$

Then

$$\mathbb{P}\left(\bigcap_{i=1}^{m} E_i\right) \geq 1/2$$

and a vector $x \in \bigcap_{i=1}^{m} E_i$ can be constructed in time $O(mN^2 \log(mN))$.

In general, it will usually be clear that the proofs can be translated into polynomial time algorithms. We do not prove an explicit bound on the time needed to find the set of edge-disjoint Hamilton cycles guaranteed by Theorem A, apart from the fact that the

time is polynomial in n.

2.3.3 The Diregularity lemma

We will use the directed version of Szemerédi's Regularity lemma. To state it we need some definitions. We write $d_G(A, B)$ for the *density* $\frac{e_G(A,B)}{|A||B|}$ of an undirected bipartite graph G with vertex classes A and B. Given $\varepsilon > 0$ we say that G is *ε-regular* if every $X \subseteq A$ and $Y \subseteq B$ with $|X| \geq \varepsilon|A|$ and $|Y| \geq \varepsilon|B|$ satisfy $|d(A, B) - d(X, Y)| \leq \varepsilon$. Given $\varepsilon, d \in (0, 1)$ we say that G is (ε, d)-regular if G is ε-regular and $d_G(A, B) = d \pm \varepsilon$. We say that G is *(ε, d)-superregular* if both of the following hold:

- G is (ε, d)-regular;
- $d(a) = (d \pm \varepsilon)|B|$, $d(b) = (d \pm \varepsilon)|A|$ for all $a \in A$, $b \in B$.

Given disjoint vertex sets A and B in a digraph G, write $(A, B)_G$ for the oriented bipartite subgraph of G whose vertex classes are A and B and whose edges are all those from A to B in G. We say that $(A, B)_G$ has any of the regularity properties above if the requirements hold for the underlying undirected bipartite graph of $(A, B)_G$.

The Diregularity lemma is a variant of the Regularity lemma for digraphs due to Alon and Shapira [7]. We will use the degree form which can be derived from the standard version in the same manner as the undirected degree form. The proof of the Diregularity lemma itself is similar to the undirected version.

Lemma 2.3.3. (Degree form of the Diregularity lemma) *For every $\varepsilon \in (0, 1)$ and every integer M' there are integers M and n_0 such that if G is a digraph on $n \geq n_0$ vertices and $d \in [0, 1]$, then there is a partition of the vertex set of G into $V_0, \ldots, V_L$ and a spanning subdigraph G' of G such that the following holds:*

- *$M' \leq L \leq M$;*
- *$|V_0| \leq \varepsilon n$;*

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

We call $V_1, \ldots, V_L$ *clusters*, V_0 the *exceptional set* and the vertices in V_0 *exceptional vertices*. We refer to G' as the *pure digraph*. The last condition of the lemma says that all pairs of clusters are ε-regular in both directions (but possibly with different densities). The *reduced digraph* R of G with parameters ε, d and M' is the digraph whose vertices are $V_1, \ldots, V_L$ and in which V_iV_j is an edge precisely when $(V_i, V_j)_{G'}$ is ε-regular and has density at least d. For each edge V_iV_j of G we write d_{ij} for the density of $(V_i, V_j)_{G'}$. Suppose $0 < 1/M' \ll \varepsilon \ll \beta \ll d \ll 1$. The *reduced multidigraph* $R(\beta)$ of G with parameters $\varepsilon, \beta, d, M'$ is obtained from R by setting $V(R(\beta)) := V(R)$ and adding $\lfloor d_{ij}/\beta \rfloor$ directed edges from V_i to V_j whenever $V_iV_j \in E(R)$. These digraphs inherit some of the key properties of G, as the next few results show (which are variants of well known observations, see e.g. Lemma 11 in [86] for the next result).

Lemma 2.3.4. *Let* $0 < 1/n_0 \ll 1/M' \ll \varepsilon \ll \beta \ll d \leq d' \ll c_1 \leq c_2 < 1$ *and let* G *be a digraph of order* $n \geq n_0$ *with* $\delta^0(G) \geq c_1 n$ *and* $\Delta^0(G) \leq c_2 n$. *Apply Lemma 2.3.3 with parameters* ε, d *and* M' *to obtain a pure digraph* G' *and a reduced digraph* R *of* G *and let* R' *denote the subdigraph of* R *whose edges correspond to pairs of density at least* d'. *Let* $R(\beta)$ *denote the reduced multidigraph of* G *with parameters* ε, β, d *and* M' *and let* $R'(\beta)$ *be the multidigraph obtained from* $R(\beta)$ *by including only those edges which also correspond to an edge of* R'. *Let* $L := |R| = |R(\beta)|$. *Then*

(i) $\delta^0(R') \geq (c_1 - 3d')L$.

(ii) $\delta^0(R'(\beta)) \geq (c_1 - 4d')\dfrac{L}{\beta}$ *and* $\Delta^0(R'(\beta)) \leq (c_2 + 2\varepsilon)\dfrac{L}{\beta}$.

Proof. To prove (i), we consider the weighted digraph R'_w obtained from R' by giving each edge V_iV_j of R' weight d_{ij}. Given a cluster V_i, we write $w^+(V_i)$ for the sum of the weights of all edges sent out by V_i in R'_w. We define $w^-(V_i)$ similarly and write $w^0(R'_w)$ for the minimum of $\min\{w^+(V_i), w^-(V_i)\}$ over all clusters V_i. Note that $\delta^0(R') \geq w^0(R'_w)$. Moreover, Lemma 2.3.3 implies that $d^{\pm}_{G'\setminus V_0}(x) > (c_1 - 2d)n$ for all $x \in V(G' \setminus V_0)$. Thus each $V_i \in V(R')$ satisfies

$$(c_1 - 2d)nm \leq e_{G'}(V_i, V(G') \setminus V_0) \leq m^2 w^+(V_i) + (d'm^2)L$$

and so $w^+(V_i) \geq (c_1 - 2d - d')L \geq (c_1 - 3d')L$. Arguing in the same way for inweights gives us $\delta^0(R') \geq w^0(R'_w) \geq (c_1 - 3d')L$. We can deduce the first part of (ii) by noting that

$$d^+_{R'(\beta)}(V_i) = \sum_{V_j \in N^+_{R'}(V_i)} \lfloor d_{ij}/\beta \rfloor \geq w^+(V_i)/\beta - L > (c_1 - 4d')\frac{L}{\beta}.$$

Similar arguments can be used to show the remaining bounds. □

Note that the previous lemma implies that, when G is dense, R and R' are certainly spanning.

Lemma 2.3.5. *Let M', n_0 be positive integers and let $\varepsilon, d, \nu, \tau$ be positive constants such that $1/n_0 \ll 1/M' \ll \varepsilon \ll d \leq d' \leq \nu \leq \tau < 1$ and $d' \leq \nu/20$. Let G be a digraph on $n \geq n_0$ vertices such that G is a robust (ν, τ)-outexpander. Let R be the reduced digraph of G with parameters ε, d and M' with clusters of size m and let R' be the subdigraph of R whose edges correspond to pairs of density at least d'. Then R' is a robust $(\nu/4, 3\tau)$-outexpander.*

Proof. Let G' denote the pure digraph, $L := |V(R)|$, and $V_1, \ldots, V_L$ be the clusters of G, and V_0 the exceptional set. Let $m := |V_1| = \ldots = |V_L|$. Suppose $S \subseteq V(R')$ has $3\tau L \leq |S| \leq (1-3\tau)L$. Let S_G denote the set of vertices which is the union of all clusters in S. So $S_G \subseteq V(G)$ and $2\tau n \leq |S_G| \leq (1 - 2\tau)n$. For every $x \in RN^+_{\nu,G}(S_G)$ we have

that $|N^-_{G'}(x) \cap S_G| \geq |N^-_G(x) \cap S_G| - (d+\varepsilon)n \geq (\nu - d - \varepsilon)n \geq \nu n/2$. This implies that

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

and every vertex $x \in RN^+_{\nu/2,G'}(S_G)$ has at least $\nu n/2$ inneighbours in S_G. Suppose, for a contradiction, that $|RN^+_{\nu/4,R'}(S)| < |S| + \nu L/4$. Let RN' denote the union of all vertices in clusters in $RN^+_{\nu/4,R'}(S)$ and let $T := RN^+_{\nu/2,G'}(S_G) \setminus RN'$; then $|T| \geq \nu n/4$.

Note that by definition, for all V outside $RN^+_{\nu/4,R'}(S)$, there exists a collection $\mathcal{V}$ of at least $|S| - \nu L/4$ clusters $U \in S$ so that there is no edge from U to V in R'. So by assumption such a $\mathcal{V}$ exists for any V which has non-empty intersection with T.

We say that a vertex $x \in V$ is *bad* if it has indegree at least $2d'm$ in at least $\sqrt{\varepsilon}L$ of the clusters in $\mathcal{V}$. The final property of Lemma 2.3.3 implies that there are at most εm vertices in V that have indegree at least $2d'm$ in some fixed cluster of $\mathcal{V}$. So by double counting the number of such vertex-cluster pairs, we see that any cluster contains at most $\sqrt{\varepsilon}m$ bad vertices.

Say that a cluster V is *significant* if $|V \cap T| \geq \varepsilon^{1/3}m$. Then there are at least $\nu L/5$ significant clusters and we write $V' := V \cap T$. Consider any $x \in V'$, where V is significant. We say that a cluster U in S is *rich for* x if x has at least $\nu m/10$ inneighbours in U. Since x has at least $\nu n/2$ inneighbours in S_G, there are at least $\nu L/3$ clusters in S which are rich for x. So there are at least $\nu L/12 \geq \sqrt{\varepsilon}L$ clusters in $\mathcal{V}$ which are rich for x. Since $d' \leq \nu/20$, this means that every x in V' is bad. Thus V contains at least $\varepsilon^{1/3}m$ bad vertices, a contradiction. □

The following simple observation is well known, the version given here is proved as Proposition 4.3(i) and (iii) in [81].

Proposition 2.3.6. *Suppose that* $0 < 1/m \ll \varepsilon \leq d' \leq d \ll 1$. *Let* G *be a bipartite graph with vertex classes* A *and* B *of size* m. *Suppose that* G' *is obtained from* G *by removing at most* $d'm$ *vertices from each vertex class and at most* $d'm$ *edges incident to each vertex from* G.

(i) If G is (ε, d)-regular then G' is $(2\sqrt{d'}, d)$-regular.

(ii) If G is (ε, d)-superregular then G' is $(2\sqrt{d'}, d)$-superregular.

The following well known observation (similar to Proposition 4.2 in [81]) states that in an ε-regular bipartite graph almost all vertices have the expected degree and almost all pairs of vertices have the expected codegree (i.e. the expected number of common neighbours). Its proof follows immediately from the definition of regularity.

Proposition 2.3.7. *Suppose that $0 < \varepsilon \leq d \ll 1$. Let G be an (ε, d)-regular bipartite graph with vertex classes A and B of size m. Then the following conditions hold.*

- *All but at most $2\varepsilon m$ vertices in A have degree $(d \pm 2\varepsilon)m$.*
- *All but at most $4\varepsilon m^2$ pairs $a \neq a'$ of distinct vertices in A satisfy $|N(a) \cap N(a')| = (d^2 \pm 2\varepsilon)m$.*
- *The vertices in B satisfy the analogues of these statements.*

The following is (a special case of) Lemma 4.9 in [81].

Lemma 2.3.8. *Suppose that $1/m \ll \varepsilon, d \ll 1/C \leq 1$. Let $G = (U, V)$ be a bipartite graph with vertex classes U and V of size m. Suppose that all but at most εm vertices in V have degree at least $(1-\varepsilon)dm$ and for all pairs of distinct vertices in V the number of common neighbours is at most Cd^2m. Suppose also that there are at most εm^2 pairs of distinct vertices in V that have at least $(1+\varepsilon)d^2m$ common neighbours. Then G is $(\varepsilon^{1/6}, d)$-regular.*

The next result (similar to Lemma 4.10(iii) and (iv) in [81]) shows that we can partition an ε-(super)regular pair into edge-disjoint ε'-(super)regular spanning subgraphs.

Lemma 2.3.9. *Let K be a positive integer and let $0 < 1/m \ll \varepsilon \ll \gamma_1, \ldots, \gamma_K \ll 1$ such that $\gamma_1 + \ldots + \gamma_K \leq d \leq 1$.*

(i) *If G is an (ε, d)-regular bipartite graph with vertex classes X, Y of size m, then it contains K edge-disjoint spanning subgraphs $J_1, \ldots, J_K$ such that for each $1 \leq k \leq K$ we have that J_k is $(\varepsilon^{1/12}, \gamma_k)$-regular. Moreover, if $x \in X$ satisfies $d_G(x) = (d \pm \varepsilon)m$, then $d_{J_k}(x) = (\gamma_k \pm \varepsilon^{1/12})m$ for each $1 \leq k \leq K$.*

(ii) *If G is an (ε, d)-superregular bipartite graph with vertex classes of size m, then it contains K edge-disjoint spanning subgraphs $J_1, \ldots, J_K$ such that for each $1 \leq k \leq K$ we have that J_k is $(\varepsilon^{1/12}, \gamma_k)$-superregular.*

Moreover, the spanning subgraphs can be found in time polynomial in m.

Proof. We only prove (i). The proof of (ii) is similar. Suppose that G is (ε, d)-regular with vertex classes U and V of size m. Assign each edge of G to J_k with probability $p_k := \gamma_k/d$ independently from all other edges. So the probability that an edge is assigned to none of $J_1, \ldots, J_K$ is $1 - (\gamma_1 + \ldots + \gamma_K)/d \geq 0$.

Consider any vertex $v \in V$ with $d_G(v) = (d \pm 2\varepsilon)m$. Then for each $1 \leq k \leq K$, the expected degree of v in J_k is $p_k(d \pm 2\varepsilon)m = (1 \pm \sqrt{\varepsilon}/2)\gamma_k m$. So Proposition 2.3.1 implies that

$$\begin{aligned}\mathbb{P}\left(d_{J_k}(v) \neq (1 \pm \sqrt{\varepsilon})\gamma_k m\right) &\leq \mathbb{P}\left(|d_{J_k}(v) - \mathbb{E}(d_{J_k}(v))| \geq \frac{\sqrt{\varepsilon}}{3}\mathbb{E}(d_{J_k}(v))\right) \\ &\leq 2e^{-\varepsilon\mathbb{E}(d_{J_k}(v))/27} \leq 2e^{-\varepsilon\gamma_k m/28}.\end{aligned}$$

For the remainder of the proof, we let the *codegree* $d_G(x, x')$ of a pair x, x' of vertices in G be the number of common neighbours of x and x'. Consider any pair $v, v' \in V$ of distinct vertices with codegree $d_G(v, v') = (d^2 \pm 2\varepsilon)m$. Then the expected codegree of v, v' in J_k is $\mathbb{E}(d_{J_k}(v, v')) = (p_k)^2(d^2 \pm 2\varepsilon)m = (1 \pm \sqrt{\varepsilon}/2)(\gamma_k)^2 m$. So Proposition 2.3.1

implies that

$$\begin{aligned}
& \mathbb{P}\left(d_{J_k}(v,v') \neq (1 \pm \sqrt{\varepsilon})(\gamma_k)^2 m\right) \\
\leq\ & \mathbb{P}\left(|d_{J_k}(v,v') - \mathbb{E}(d_{J_k}(v,v'))| \geq \frac{\sqrt{\varepsilon}}{3}\mathbb{E}(d_{J_k}(v,v'))\right) \\
\leq\ & 2e^{-\varepsilon\mathbb{E}(d_{J_k}(v,v'))/27} \leq 2e^{-\varepsilon(\gamma_k)^2 m/28}.
\end{aligned}$$

Similarly, consider any pair $x \neq x'$ of vertices in G (with no restriction on the codegree $d_G(x,x')$). If $d_G(x,x') \leq 3(\gamma_k)^2 m/2d^2$, then clearly $d_{J_k}(x,x') \leq 3(\gamma_k)^2 m/2d^2$. So suppose that $d_G(x,x') \geq 3(\gamma_k)^2 m/2d^2$. Then $3(\gamma_k)^4 m/2d^4 \leq \mathbb{E}(d_{J_k}(x,x')) \leq (p_k)^2 m = (\gamma_k)^2 m/d^2$ and so

$$\begin{aligned}
\mathbb{P}\left(d_{J_k}(x,x') \geq \frac{3(\gamma_k)^2 m}{2d^2}\right) &\leq \mathbb{P}\left(d_{J_k}(x,x') \geq \frac{3}{2}\mathbb{E}(d_{J_k}(x,x'))\right) \\
&\leq 2e^{-\mathbb{E}(d_{J_k}(x,x'))/12} \leq 2e^{-(\gamma_k)^4 m/8d^4}.
\end{aligned}$$

Proposition 2.3.7 implies that V contains at most $2\varepsilon m$ vertices whose degree in G is not $(d \pm 2\varepsilon)m$ as well as at most $4\varepsilon m^2$ pairs of distinct vertices whose codegree in G is not $(d^2 \pm 2\varepsilon)m$. Note that certainly $\gamma_k \geq 1/m^{1/8}$ so $K/m^{1/8} \leq \gamma_1 + \ldots + \gamma_K \leq d$ so $K \leq dm^{1/8}$. Thus a union bound implies that with probability at least

$$\begin{aligned}
& 1 - \sum_{1 \leq k \leq K}\left(me^{-\varepsilon\gamma_k m/28} + m^2 e^{-\varepsilon(\gamma_k)^2 m/28} + m^2 e^{-(\gamma_k)^4 m/8d^4}\right) \\
\geq\ & 1 - 2dm^{1/8}\left(me^{-\varepsilon m^{7/8}/28} + m^2 e^{-\varepsilon m^{3/4}/28} + m^2 e^{-m^{1/2}/8d^4}\right) \geq 1/2
\end{aligned}$$

all of the following properties are satisfied for each $1 \leq k \leq K$:

- All but at most $2\varepsilon m$ vertices $v \in V$ satisfy $d_{J_k}(v) = (1 \pm \sqrt{\varepsilon})\gamma_k m$.
- All but at most $4\varepsilon m^2$ pairs $v \neq v'$ of vertices in V satisfy $d_{J_k}(v,v') = |N_{J_k}(v) \cap N_{J_k}(v')| \leq (1 + \sqrt{\varepsilon})(\gamma_k)^2 m$.
- All pairs $v \neq v'$ of vertices in V satisfy $d_{J_k}(v,v') \leq 3(\gamma_k)^2 m/2d^2$.

Then Lemma 2.3.8 (applied with $\sqrt{\varepsilon}, \gamma_k, 3/2d^2$ playing the roles of ε, d, C) implies that we can ensure that J_k is $(\varepsilon^{1/12}, \gamma_k)$-regular for each $1 \leq k \leq K$. $\square$

The proof of Theorem A begins by decomposing our digraph into 'blown-up' 1-factors and we will need the following well known and easy fact that allows us to extract almost spanning blown-up 1-factors in which pairs are superregular.

Lemma 2.3.10. *Let $0 < \varepsilon \leq \gamma \leq 1 \leq m$ and let D be a digraph with vertex clusters $V_1, \ldots, V_k$ each of size m such that $(V_j, V_{j+1})_D$ is (ε, γ)-regular for $1 \leq j \leq k$, where $V_{k+1} := V_1$. Then there exists a subdigraph D' of D with vertex clusters $V'_1, \ldots, V'_k$ where $V'_j \subseteq V_j$, $|V'_j| = (1 - 2\varepsilon)m$ and $(V'_j, V'_{j+1})_{D'}$ is $(4\varepsilon, \gamma)$-superregular for $1 \leq j \leq k$, where $V'_{k+1} := V'_1$.*

Proof. For each $1 \leq j \leq k$, each V_i contains at most $2\varepsilon m$ vertices whose outdegree or indegree in D is either at most $(\gamma - 2\varepsilon)m$ or at least $(\gamma + 2\varepsilon)m$. Deleting exactly $2\varepsilon m$ vertices including these from each cluster gives us D'. $\square$

We will use the following crude version of the fact that every ε-regular pair contains a subgraph of given maximum degree Δ whose average degree is close to Δ, which is Lemma 13 in [86].

Lemma 2.3.11. *Suppose that $0 < 1/m \ll \varepsilon', \varepsilon \ll d_0 \leq d_1 \ll 1$ and that (A, B) is an (ε, d_1)-regular pair with m vertices in each class. Then (A, B) contains a subgraph H whose maximum degree is at most $d_0 m$ and whose average degree is at least $d_0 m/8$.*

The proof proceeds by greedily removing matchings and so H can be found in polynomial time. Part (ii) of the following observation is proved as Lemma 5.3 in [81]; (i) is immediate from the definition.

Lemma 2.3.12. *Let $r \geq 3$ and let G be a robust (ν, τ)-outexpander with $0 < 3\nu \leq \tau < 1$. Let G' be the r-fold blow-up of G. Then*

(i) $\delta^0(G') = r\delta^0(G)$.

(ii) *G' is a robust $(\nu^3, 2\tau)$-outexpander.*

2.3.4 Uniform refinements

We will also need to partition each vertex cluster into equal parts in such a way that the in- and outneighbourhood of each vertex restricted to each part is roughly the size we expect it to be. This is very similar to Lemma 4.7 in [81]. To state the result, we need the following definitions. Let G be a digraph and let $\mathcal{P}$ be a partition of $V(G)$ into an exceptional set V_0 and clusters of equal size. Suppose that $\mathcal{P}'$ is another partition of $V(G)$ into an exceptional set V_0' and clusters of equal size. We say that $\mathcal{P}'$ is an *ℓ-refinement of* $\mathcal{P}$ if $V_0 = V_0'$ and if the clusters in $\mathcal{P}'$ are obtained by partitioning each cluster in $\mathcal{P}$ into ℓ subclusters of equal size. (So if $\mathcal{P}$ contains k clusters then $\mathcal{P}'$ contains $k\ell$ clusters.) $\mathcal{P}'$ is an *ε-uniform ℓ-refinement of $\mathcal{P}$ with respect to G* if it is an ℓ-refinement of $\mathcal{P}$ which satisfies the following condition:

(URef) Whenever x is a vertex of G, V is a cluster in $\mathcal{P}$ and $|N_G^+(x) \cap V| \geq \varepsilon|V|$ then $|N_G^+(x) \cap V'| = (1 \pm \varepsilon)|N_G^+(x) \cap V|/\ell$ for each cluster $V' \in \mathcal{P}'$ with $V' \subseteq V$. The inneighbourhoods of the vertices of G satisfy an analogous condition.

Let $\mathcal{G}$ be a collection of digraphs on the same vertex set. If $\mathcal{P}'$ is a refinement of a partition $\mathcal{P}$ with respect to G for all $G \in \mathcal{G}$ then we say that it is a refinement *with respect to $\mathcal{G}$*.

The next lemma is a generalisation of Lemma 4.7 in [81]. Its proof is very similar but we include it here for completeness. The proof proceeds by considering a random partition of V^* (which can be derandomised by Theorem 2.3.2).

Lemma 2.3.13. *Suppose that $0 < 1/m \ll 1/k, \varepsilon \ll \varepsilon', d, 1/\ell, 1/t \leq 1$ and $m/\ell \in \mathbb{N}$. Suppose that $\mathcal{G}$ is a collection of t digraphs on the same set V^* of $n \leq 2km$ vertices and that $\mathcal{P}$ is a partition of V^* into an exceptional set V_0 and k clusters of size m. Then there exists an ε-uniform ℓ-refinement of $\mathcal{P}$ with respect to $\mathcal{G}$. Moreover, any ε-uniform ℓ-refinement $\mathcal{P}'$ of $\mathcal{P}$ automatically satisfies the following conditions for all $G \in \mathcal{G}$:*

(i) Suppose that V, W are clusters in $\mathcal{P}$ and V', W' are clusters in $\mathcal{P}'$ with $V' \subseteq V$ and $W' \subseteq W$. If $G[V, W]$ is (ε, d)-superregular then $G[V', W']$ is (ε', d)-superregular.

(ii) Suppose that V, W are clusters in $\mathcal{P}$ and V', W' are clusters in $\mathcal{P}'$ with $V' \subseteq V$ and $W' \subseteq W$. If $G[V,W]$ is (ε, d)-regular then $G[V',W']$ is (ε', d)-regular.

Proof. To prove the existence of an ε-uniform ℓ-refinement of $\mathcal{P}$, let $\mathcal{P}^*$ be a partition obtained by splitting each cluster $V \in \mathcal{P}$ uniformly at random into ℓ subclusters. More precisely, the probability that a vertex $x \in V$ is assigned to the ith subcluster is $1/\ell$, independently of all other vertices. Consider a fixed graph G in $\mathcal{G}$, a fixed vertex x of G and a cluster $V \in \mathcal{P}$ with $d^+ := |N_G^+(x) \cap V| \geq \varepsilon m$. Given a cluster $V' \in \mathcal{P}^*$ with $V' \subseteq V$, we say that (x, G) is *out-bad for* V' if the outdegree of x into V' in G is not $(1 \pm \varepsilon/2)d^+/\ell$. We say that x is *out-bad for* V' if (x, G) is out-bad for V' for some $G \in \mathcal{G}$. Then Proposition 2.3.1 implies that the probability that x is out-bad for V' is at most $2te^{-\varepsilon^2 d^+/3\cdot 4\ell} \leq 2te^{-\varepsilon^4 m}$. Since $\mathcal{P}^*$ contains $k\ell \leq n$ clusters, the probability that the common vertex set of the graphs in $\mathcal{G}$ contains some vertex which is out-bad for at least one cluster $V' \in \mathcal{P}^*$ is at most $tn^2 e^{-\varepsilon^4 m} < 1/8$. We argue analogously for the inneighbourhoods of the vertices in G (by considering 'in-bad' vertices).

We now say that a cluster V' of $\mathcal{P}^*$ is *good* if $|V'| = (1 \pm \varepsilon^2/2)m/\ell$. A similar argument as above shows that the probability that $\mathcal{P}^*$ has a cluster which is not good is at most $1/4$. So with probability at least $1/2$, all clusters of $\mathcal{P}^*$ are good, and no vertices are out-bad or in-bad.

Now obtain $\mathcal{P}'$ from $\mathcal{P}^*$ as follows: for each cluster V of $\mathcal{P}$, equalize the sizes of the corresponding ℓ subclusters in $\mathcal{P}$ by moving at most $\varepsilon^2 m/2\ell$ vertices from one subcluster to another. So whenever G is a graph in $\mathcal{G}$, x is a vertex of G, V is a cluster in $\mathcal{P}$ and $|N_G^+(x) \cap V| \geq \varepsilon |V|$, it follows that we have

$$|N_G^+(x) \cap V'| = (1 \pm \varepsilon/2)|N_G^+(x) \cap V|/\ell \pm \varepsilon^2 m/2\ell$$

for each cluster $V' \in \mathcal{P}'$ with $V' \subseteq V$. The inneighbourhoods of the vertices of G satisfy an analogous condition. So (URef) holds and so $\mathcal{P}'$ is an ε-uniform ℓ-refinement of $\mathcal{P}$.

To prove (i), suppose that $\mathcal{P}'$ is any ε-uniform ℓ-refinement of $\mathcal{P}$, that $G \in \mathcal{G}$, and

that $G[V,W]$ is (ε,d)-superregular (where V and W are clusters in $\mathcal{P}$). Let V' and W' be clusters in $\mathcal{P}'$ with $V' \subseteq V$ and $W' \subseteq W$. Then $G[V',W']$ is $\varepsilon\ell$-regular and thus ε'-regular. Consider any $x \in V'$ and let $d^+ := |N_G^+(x) \cap W|$. Thus $d^+ = (d \pm \varepsilon)m$ since $G[V,W]$ is (ε,d)-superregular. Together with the ε-uniformity of $\mathcal{P}'$ this implies that $|N_G^+(x) \cap W'| = (1 \pm \varepsilon)d^+/\ell = (d \pm \varepsilon')m/\ell$. The inneighbourhoods in V' of the vertices in W' satisfy the analogous property. Thus $G[V',W']$ is (ε',d)-superregular.

The proof of (ii) is almost identical. □

Let $\varepsilon > 0$ and let $\mathcal{P}$ be a partition of $V(G)$ into an exceptional set V_0 and clusters of size m. Let $\mathcal{P}'$ be another partition of $V(G)$ into an exceptional set V_0' and clusters of size m' where $m \geq m'$ and $|m - m'| \leq 2\varepsilon m'$. We say that $\mathcal{P}$ and $\mathcal{P}'$ are *ε-close* if $|V_0 \cap V_0'| \geq (1-\varepsilon)|V_0'|$ and if for each cluster U in $\mathcal{P}'$ there is a cluster V in $\mathcal{P}$ such that $|U \cap V| \geq (1-\varepsilon)m'$. In this case we say that U and V are *associated*. Note that V is unique when $\varepsilon < 1/2$. Suppose that R is a multidigraph whose vertices are the clusters of $\mathcal{P}$. Let R' be the multidigraph obtained from R by relabelling V by V' for each $V \in \mathcal{P}$ associated with $V' \in \mathcal{P}'$. So R' has vertex set consisting precisely of the clusters of $\mathcal{P}'$. Moreover, for each edge E from U to V in R, there is a unique edge E' from U' to V' in R' which is *associated* with E. The following lemma states that refinements of ε-close partitions are still ε'-close with a slightly bigger parameter ε'.

Lemma 2.3.14. *Suppose that $0 < 1/m \ll 1/k, \varepsilon_1, \varepsilon_2 \ll \varepsilon', d, 1/\ell \leq 1$ and that $m/\ell \in \mathbb{N}$. Suppose that G is a digraph on $n \leq 2km$ vertices and that $\mathcal{P}$ is a partition of $V(G)$ into an exceptional set V_0 and k clusters of size m. Let $\mathcal{P}'$ be an ε_1-uniform ℓ-refinement of $\mathcal{P}$. Suppose that $\mathcal{R}$ is another partition of $V(G)$ into an exceptional set V_0' and clusters of size m' that is ε_2-close to $\mathcal{P}$. Then, in time polynomial in m, one can find an ε'-uniform ℓ-refinement $\mathcal{R}'$ of $\mathcal{R}$ which is ε'-close to $\mathcal{P}'$.*

Proof. Let U be a cluster of $\mathcal{P}$ and let V be the cluster of $\mathcal{R}$ associated with U. Then, for each U' in $\mathcal{P}'$ such that $U' \subseteq U$ we have that $|U' \cap V| \geq m'/\ell - \varepsilon_2 m'$, so we can pick a subset V' of $U' \cap V$ of size exactly $(1-\varepsilon_2\ell)m'/\ell$. There are now exactly $\varepsilon_2\ell m'$

vertices of V which do not lie in any subcluster V'. Distribute these among the V' so that every subcluster has equal size m'/ℓ. Together with V_0', these subclusters form the partition $\mathcal{R}'$. Clearly U' and V' are associated clusters of $\mathcal{P}'$ and $\mathcal{R}'$ respectively and $|U' \cap V'| \geq (1-\varepsilon')m'/\ell$. It is easy to see that $\mathcal{R}'$ has the required properties. $\square$

Observe that if $\varepsilon_1 \leq \varepsilon_2$ then any ε_1-uniform refinement is also an ε_2-uniform refinement, and two ε_1-close partitions are also ε_2-close.

Let $\mathcal{P}_2$ denote the partition obtained by taking an ε-uniform ℓ_1-refinement $\mathcal{P}_1$ of a partition $\mathcal{P}$ and then taking an ε-uniform ℓ_2-refinement of $\mathcal{P}_1$. Then $\mathcal{P}_2$ is a 3ε-uniform $\ell_2\ell_1$-refinement of $\mathcal{P}$. Indeed, whenever x is a vertex of G, V is a cluster in $\mathcal{P}$ and $|N_G^+(x) \cap V| \geq \varepsilon|V|$, then for each cluster $V' \in \mathcal{P}_2$ with $V' \subseteq V$, we have

$$|N_G^+(x) \cap V'| = (1 \pm \varepsilon)^2 |N_G^+(x) \cap V|/\ell_2\ell_1 = (1 \pm 3\varepsilon)|N_G^+(x) \cap V|/\ell_2\ell_1, \tag{2.3.1}$$

and similarly for the inneighbourhoods.

2.4 Tools for finding subgraphs, 1-factors and Hamilton cycles

2.4.1 Almost regular spanning subgraphs

The following result (which is proved as Lemma 5.2 in [81]) shows that in a robust outexpander, we can guarantee a spanning subdigraph with a given degree sequence (as long as the required degrees are not too large and do not deviate too much from each other). If x is a vertex of a multidigraph Q, we write $d_Q^+(x)$ for the number of edges in Q whose initial vertex is x and $d_Q^-(x)$ for the number of edges in Q whose final vertex is x.

Lemma 2.4.1. *Let $q \in \mathbb{N}$. Suppose that $0 < 1/n \ll \varepsilon \ll \nu \leq \tau \ll \alpha < 1$ and that $1/n \ll \rho \leq q\nu^2/3$. Let G be a digraph on n vertices with $\delta^0(G) \geq \alpha n$ which is a robust (ν, τ)-outexpander. Suppose that Q is a multidigraph on $V(G)$ such that*

whenever $xy \in E(G)$ *then* Q *contains at least* q *edges from* x *to* y*. For every vertex* x *of* G*, let* $n_x^+, n_x^- \in \mathbb{N}$ *be such that* $(1-\varepsilon)\rho n \leq n_x^+, n_x^- \leq (1+\varepsilon)\rho n$ *and such that* $\sum_{x \in V(G)} n_x^+ = \sum_{x \in V(G)} n_x^-$*. Then* Q *contains a spanning submultidigraph* Q' *such that* $d_{Q'}^+(x) = n_x^+$ *and* $d_{Q'}^-(x) = n_x^-$ *for every* $x \in V(G) = V(Q)$*.*

The next result (Lemma 16 in [86]) is an analogue of the previous one where we consider superregular pairs instead of robust outexpanders. In both cases, the proof is algorithmic (as it is based on the Max-Flow-Min-Cut Theorem).

Lemma 2.4.2. *Let* $0 < 1/n \ll \varepsilon \ll \beta \ll \alpha' \ll \alpha \ll 1$*. Suppose that* $G = (A, B)$ *is an* $(\varepsilon, \beta + \varepsilon)$*-superregular pair where* $|A| = |B| = n$*. Define* $\kappa := (1-\alpha)\beta n$*. Suppose we have a non-negative integer* $m_a^+ \leq \alpha'\beta n$ *associated with each* $a \in A$ *and a non-negative integer* $m_b^- \leq \alpha'\beta n$ *associated with each* $b \in B$ *such that* $\sum_{a \in A} m_a^+ = \sum_{b \in B} m_b^-$*. Then* G *contains a spanning subgraph* H *in which* $d_H(a) = n_a^+ := \kappa - m_a^+$ *for any* $a \in A$ *and* $d_H(b) = n_b^- := \kappa - m_b^-$ *for any* $b \in B$*.*

2.4.2 Decomposing regular digraphs into 1-factors

Petersen proved that every regular undirected graph can be decomposed into 1-factors. The corresponding result for directed graphs is well known; for completeness we include the proof (which is algorithmic as perfect matchings can be found in polynomial time).

Proposition 2.4.3. *Any* r*-regular multidigraph* G *contains* r *edge-disjoint 1-factors.*

Proof. Define an undirected bipartite graph J with two vertex classes A and B, each of which is a copy of $V(G)$, with an edge from $a \in A$ to $b \in B$ for each edge from a to b in G. J is r-regular so, by Hall's Theorem [60], contains a perfect matching M_1. Then $J \setminus M_1$ is $(r-1)$-regular so contains a perfect matching M_2. Repeating this procedure we can decompose J into r perfect matchings, each of which corresponds to a 1-factor in G. $\square$

2.4.3 Unwinding cycles

At two points in the proof, we will partition a blown-up cycle into several longer, thinner blown-up cycles on subclusters of the original clusters. The following section describes how this process is implemented and describes a special approximate decomposition to be used in Section 2.5.3.

Suppose that $D = p \otimes C_n$ is a p-fold blow-up of a cycle C_n of length n. Let $X_1, \ldots, X_n$ be the vertex classes of D. We call any edge-disjoint collection $C^1, \ldots, C^{p'}$ of p' Hamilton cycles of D a *p'-unwinding* of D. The following lemma guarantees a $(p-1)$-unwinding in which, for each C^d and each i, the ith vertices of two distinct classes X_j and $X_{j'}$ have distance at least p on C^d.

Lemma 2.4.4. *Suppose that $p > 2$ is a prime, suppose $n \in \mathbb{N}$ and let $D = p \otimes C_n$ be a p-fold blow-up of a cycle C_n of length n. Denote the vertex classes of D by $X_1, \ldots, X_n$ where, for all j with $1 \leq j \leq n$, we have $X_j = \{x_j^1, \ldots, x_j^p\}$. Then D contains a $(p-1)$-unwinding $C^1, \ldots, C^{p-1}$ such that for every $1 \leq d \leq p-1$ and every $1 \leq i \leq p$,*

(i) if p is coprime to n, then the vertices $x_1^i, \ldots, x_n^i$ have pairwise distance at least p on C^d;

(ii) if p is not coprime to n, then the vertices $x_1^i, \ldots, x_{n-2}^i$ have pairwise distance at least p on C^d.

Proof. We first prove (i). Let $\{a\}$ denote the residue of a modulo p and $[b]$ the residue of b modulo n where we adopt the convention that $\{\ell p\} := p$ and $[\ell n] := n$ for any $\ell \in \mathbb{N}$. For $1 \leq d \leq p-1$ define the modular arithmetic progression

$$P(d) := (\{1\}, \{1+d\}, \ldots, \{1+(np-1)d\})$$

in $\mathbb{Z}_p$. For each $1 \leq k \leq np$ and $1 \leq d \leq p-1$, define the edge

$$e_k^d := x_{[k]}^{P(d)_k} x_{[k+1]}^{P(d)_{k+1}},$$

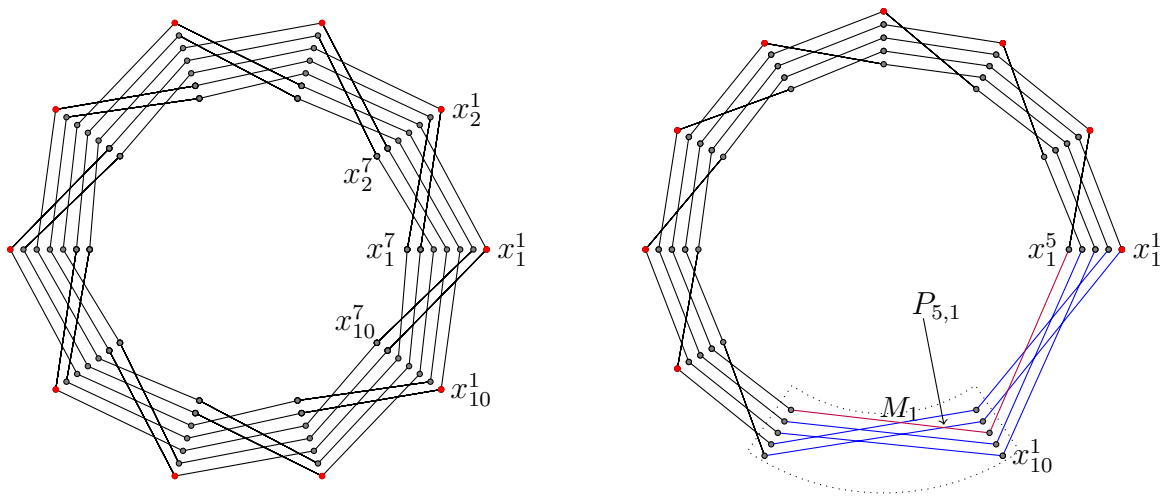

Figure 2.1: Illustrating Lemma 2.4.4(i) with $n = 10, p = 7, d = 2$ and Lemma 2.4.4(ii) with $n = 10, p = 5, d = 1$.

where $P(d)_k$ denotes the kth term of $P(d)$ and $P(d)_{np+1} := P(d)_1$. We define C^d to be the digraph with vertex set $V(D)$ and edges $e_1^d, \ldots, e_{np}^d$ (see Figure 2.1). Note that C^d is clearly a closed walk in D.

Claim A. *For each $1 \leq d, d' \leq p-1$ and $1 \leq k, k' \leq np$ the following hold:*

(a) $P(d)$ is periodic with period p.

(b) Suppose $P(d)_k = P(d')_{k'}$ and $P(d)_{k+1} = P(d')_{k'+1}$. Then $d = d'$.

We first show that the claim implies (i). First note that (a) and the fact that n is coprime to p imply that every vertex is visited exactly once in the closed walk C^d, so C^d must in fact be a Hamilton cycle. Now suppose $e_k^d = e_{k'}^{d'}$. Then (b) implies that also $d = d'$. Thus no two C^d share an edge; thus $C^1, \ldots, C^{p-1}$ is a collection of edge-disjoint Hamilton cycles. (a) implies that, on each C^d, the distance between x_ℓ^i and $x_{\ell'}^i$ is a multiple of p for any $1 \leq \ell, \ell' \leq n$. Therefore $C^1, \ldots, C^{p-1}$ have the required property.

Proof of Claim A. To prove (a) of the claim, note that $P(d)_k = P(d)_{k'}$ if and only if $1 + kd \equiv 1 + k'd \mod p$ if and only if $k \equiv k' \mod p$ since d is coprime to p. To

prove (b), note that $P(d)_k = P(d')_{k'}$ and $P(d)_{k+1} = P(d')_{k'+1}$ imply that

$$1 + kd \equiv 1 + k'd' \mod p \tag{2.4.1}$$

$$1 + (k+1)d \equiv 1 + (k'+1)d' \mod p. \tag{2.4.2}$$

Subtracting (2.4.1) from (2.4.2) gives $d \equiv d' \mod p$; but $1 \leq d, d' \leq p-1$ so $d = d'$. This proves the claim and completes the proof of (i). ■

We now prove (ii). So suppose instead that n and p are not coprime. Then $n' := n-2$ is coprime to p since $p > 2$. The idea is to use paths derived from the cycles defined above for the first n' clusters and extend them into Hamilton cycles via the remaining clusters. To this end, form an auxiliary blown-up cycle $\tilde{D}$ from D by identifying x^i_j with $x^i_{j'}$ whenever $1 \leq i \leq p$ and $j, j' \in \{n-1, n, 1\}$ and call this vertex x^i_1 in $\tilde{D}$. Now remove any resulting loops from $\tilde{D}$. So $\tilde{D} = p \otimes C_{n-2}$. Next, apply (i) to $\tilde{D}$ to obtain $\tilde{C}^1, \ldots, \tilde{C}^{p-1}$. Now, for each $1 \leq d \leq p-1$, obtain $E_1(C^d)$ from $E(\tilde{C}^d)$ by replacing any edge $x^i_{n-2}x^{i'}_1$ by $x^i_{n-2}x^{i'}_{n-1}$. Note that, in D, $E_1(C^1), \ldots, E_1(C^{p-1})$ is an edge-disjoint collection of $p-1$ paths each of length n'.

Claim B. *The collections $E_1(C^1), \ldots, E_1(C^{p-1})$ of paths can be extended into $p-1$ edge-disjoint Hamilton cycles $C^1, \ldots, C^{p-1}$ respectively such that C^d is a subdivision of $\tilde{C}^d$ for each $1 \leq d \leq p-1$.*

Proof. For each $1 \leq d \leq p-1$ we will need to find a collection of edge-disjoint paths from x^i_{n-1} to x^i_1 for $1 \leq i \leq p$ to extend $E_1(C^d)$ to C^d. Moreover, these collections must be pairwise edge-disjoint. By Hall's Theorem, we can find $p-1$ edge-disjoint perfect matchings $M_1, \ldots, M_{p-1}$ in the complete bipartite subgraph (X_{n-1}, X_n) of D. For each $1 \leq d \leq p-1$ and $1 \leq i \leq p$, define

$$P_{i,d} = x^i_{n-1}x^{i'}_n x^i_1$$

whenever $x^i_{n-1}x^{i'}_n$ is an edge in M_d. Since the M_d are edge-disjoint matchings, the

$P_{i,d}$ are edge-disjoint paths with the required property. Thus, for each $1 \leq d \leq p-1$, defining

$$E(C^d) := E_1(C^d) \cup \bigcup_{1 \leq i \leq p} P_{i,d}$$

gives $p-1$ edge-disjoint Hamilton cycles $C^1, \ldots, C^{p-1}$. This proves the claim. ■

Note that since C^d is a subdivision of $\tilde{C}^d$, the distance between any two vertices in C^d is at least the distance in $\tilde{C}^d$. This immediately gives the required property, and completes the proof of (ii). □

2.4.4 Merging 1-factors in blown-up cycles

In Section 2.5.10 we will have found an approximate decomposition of a robustly expanding digraph into 1-factors. The following lemma will use the special structure of the 1-factors to merge their cycles into a single Hamilton cycle. It is a special case of Lemma 6.5 in [81], which in turn is based on an idea in [33]. As noted in [81], the cycle guaranteed by the lemma can be found in polynomial time. Roughly speaking, the lemma asserts that if we have a 1-regular digraph F where most of the edges wind around a 'blown-up' cycle $C = V_1 \ldots V_k$, then under certain circumstances we can turn F into a (single) cycle by replacing a few edges of F by edges from a digraph G whose edges all wind around C.

Lemma 2.4.5. *Let $0 < 1/m \ll \varepsilon \ll d < 1$. Let $V_1, \ldots, V_k$ be pairwise disjoint clusters, each of size m and let $C = V_1 \ldots V_k$ be a directed cycle on these clusters. Let $J \subseteq E(C)$. Let G be a digraph on $V_1 \cup \cdots \cup V_k$ such that $G[V_i, V_{i+1}]$ is (ε, d)-superregular for every $V_iV_{i+1} \in J$. Suppose that F is a 1-regular digraph with $V_1 \cup \cdots \cup V_k \subseteq V(F)$ such that the following properties hold:*

(i) *For each edge $V_iV_{i+1} \in J$ the digraph $F[V_i, V_{i+1}]$ is a perfect matching.*

(ii) *For each cycle D in F there is some edge $V_iV_{i+1} \in J$ such that D contains a vertex in V_i.*

(iii) *Whenever* $V_iV_{i+1}, V_jV_{j+1} \in J$ *are such that* J *avoids all edges in the segment* $V_{i+1}CV_j$ *of* C *from* V_{i+1} *to* V_j*, then* F *contains a path* P_{ij} *joining some vertex* $u_{i+1} \in V_{i+1}$ *to some vertex* $u'_j \in V_j$ *such that* P_{ij} *winds around* C*.*

Then we can obtain a cycle on $V(F)$ *from* F *by replacing* $F[V_i, V_{i+1}]$ *with a suitable perfect matching in* $G[V_i, V_{i+1}]$ *for each edge* $V_iV_{i+1} \in J$*.*

It will also be convenient to use the following result from [85], which guarantees a Hamilton cycle in a robustly expanding digraph. The proof of Lemma 2.4.5 actually consists of repeated applications of Theorem 2.4.6 to a suitable auxiliary digraph. The proof of Theorem 2.4.6 can be made algorithmic but this is not needed here as we only apply it to a 'reduced' digraph, obtained from the Regularity lemma.

Theorem 2.4.6. *Let* n_0 *be a positive integer and* α, ν, τ *be positive constants such that* $1/n_0 \ll \nu \leq \tau \ll \alpha < 1$*. Let* G *be a digraph on* $n \geq n_0$ *vertices with* $\delta^0(G) \geq \alpha n$ *which is a robust* (ν, τ)*-outexpander. Then* G *contains a Hamilton cycle.*

2.5 The proof of Theorem A

2.5.1 Applying the Diregularity lemma

We choose τ so that $\tau \ll \alpha$. Without loss of generality we may assume that $\nu \ll \tau$ as any robust (ν, τ)-outexpander is also a robust (ν', τ)-outexpander for any $\nu' \leq \nu$. We may also assume that $0 < \eta \ll \nu$ as a collection of $(1-\eta')r$ edge-disjoint Hamilton cycles certainly contains a collection of $(1-\eta)r$ edge-disjoint Hamilton cycles if $\eta' \leq \eta$. Define further constants satisfying

$$\begin{aligned} 0 < 1/n_0 \ll 1/M \ll 1/M' \ll \tilde{\varepsilon} \ll \varepsilon \ll \varepsilon' \ll \xi \ll 1/p \\ \ll \beta \ll d \ll 1/s \ll \gamma \ll d' \ll \eta \ll \nu \ll \tau \ll \alpha, \end{aligned} \tag{2.5.1}$$

where $s \in \mathbb{N}$ is even and p is a prime.

Let G be a digraph of order $n \geq n_0$ such that G is an r-regular robust (ν, τ)-outexpander with $r \geq \alpha n$. Define $\tilde{\alpha}$ by $r = \tilde{\alpha} n$. Apply the Diregularity lemma (stated as Lemma 2.3.3) to G with parameters $\tilde{\varepsilon}^{12}, d, M'$ to obtain clusters $\tilde{V}_1, \ldots, \tilde{V}_{\tilde{L}}$ of size $\tilde{m}$, an exceptional set V_0, a pure digraph G' and a reduced digraph $\tilde{R}$. So $|\tilde{R}| = \tilde{L}$ and $M' \leq \tilde{L} \leq M$. We denote the above partition of G by $\tilde{\mathcal{P}}$ and call the $\tilde{V}_j$ the *clusters* of $\tilde{\mathcal{P}}$, frequently referred to as *base primary clusters* (to distinguish them from other types of cluster defined later on). Let $\tilde{R}'$ be the spanning subdigraph of $\tilde{R}$ whose edges correspond to pairs of density at least d'. So $\tilde{V}_i\tilde{V}_j$ is an edge of $\tilde{R}'$ if $(\tilde{V}_i, \tilde{V}_j)_{G'}$ has density at least d'.

When $\tilde{E}$ is an edge of $\tilde{R}$ from $\tilde{V}_i$ to $\tilde{V}_j$ we write $G'(\tilde{E})$ for the subdigraph $(\tilde{V}_i, \tilde{V}_j)_{G'}$ and d_{ij} for the density of this pair. Then by Lemma 2.3.3, $G'(\tilde{E})$ is $(\tilde{\varepsilon}^{12}, d_{ij})$-regular. Let $\tilde{R}(\beta)$ denote the reduced multidigraph of G (obtained from $\tilde{R}$) with parameters $\tilde{\varepsilon}^{12}, \beta, d$ and M'. Let $\tilde{R}'(\beta)$ be the multidigraph obtained from $\tilde{R}(\beta)$ by including only those edges which also correspond to an edge of $\tilde{R}'$. Roughly speaking, our aim is to find an approximate decomposition of $\tilde{R}(\beta)$ into edge-disjoint 1-factors $\tilde{F}$, and then find an approximate Hamilton decomposition of a subdigraph of G consisting mainly of edges that correspond to a pair in $\tilde{F}$.

For each edge $\tilde{E}$ of $\tilde{R}$, apply Lemma 2.3.9(i) to $G'(\tilde{E})$ with parameters $K := \lfloor d_{ij}/\beta \rfloor$ and $\gamma_k := \beta$ for each $1 \leq k \leq K$ to obtain K edge-disjoint $(\tilde{\varepsilon}, \beta)$-regular subdigraphs. We associate each of these with a unique edge E from $\tilde{V}_i$ to $\tilde{V}_j$ of $\tilde{R}(\beta)$ and call the corresponding digraph $G'(E)$.

Let A be a cluster of $\tilde{R}$ and let $E(A)$ denote the set of edges incident to A in $\tilde{R}(\beta)$. For an edge E in $E(A)$ and $x \in A$, we say that the pair (x, E) is *good* if

- A is the initial cluster of E and $d^+_{G'(E)}(x) = (\beta \pm 2\tilde{\varepsilon})\tilde{m}$; or
- A is the final cluster of E and $d^-_{G'(E)}(x) = (\beta \pm 2\tilde{\varepsilon})\tilde{m}$

and say it is *bad* otherwise (recall that $\tilde{m}$ is the cluster size). We say that x is *good* if x forms a bad pair with at most $\xi|E(A)|$ edges in $E(A)$. Note that for a fixed edge E in $E(A)$, at most $\tilde{\varepsilon}\tilde{m}$ vertices $x \in A$ are bad. So by double counting the number of bad

pairs, it is easy to see that the number of bad vertices in A is at most $\tilde{\varepsilon}\tilde{m}/\xi$.

We remove every bad vertex from its cluster as well as possibly some more arbitrary vertices so that exactly $\tilde{\varepsilon}\tilde{m}/\xi$ vertices have been removed from each cluster. We then remove at most $2sp$ further vertices from each cluster in order to guarantee that the cluster size is divisible by $2sp$. We still denote the cluster size by $\tilde{m}$ and still call the clusters *base primary*. Each vertex removed here is added to the exceptional set V_0, which we now call the *core exceptional set*. So

$$|V_0| \leq (\tilde{\varepsilon}^{12} + \tilde{\varepsilon}/\xi)n + 2sp\tilde{L} \overset{(2.5.1)}{\leq} \sqrt{\tilde{\varepsilon}}n/2. \tag{2.5.2}$$

We still denote the partition of $V(G)$ into V_0 and these clusters by $\tilde{\mathcal{P}}$. Note that for each edge E of $\tilde{R}(\beta)$, the digraph $G'(E)$ is still $(\sqrt{\tilde{\varepsilon}}, \beta)$-regular by Proposition 2.3.6(i) (at most $\tilde{\varepsilon}\tilde{m}/4$ vertices were removed from each cluster). Lemma 2.3.4 implies that

$$\begin{aligned} \delta^0(\tilde{R}') \geq (\tilde{\alpha} - 3d')\tilde{L} \quad &\text{and} \quad \delta^0(\tilde{R}'(\beta)) \geq (\tilde{\alpha} - 4d')\frac{\tilde{L}}{\beta}, \\ \delta^0(\tilde{R}(\beta)) \geq (\tilde{\alpha} - 4d)\frac{\tilde{L}}{\beta} \quad &\text{and} \quad \Delta^0(\tilde{R}(\beta)) \leq (\tilde{\alpha} + 2\tilde{\varepsilon}^{12})\frac{\tilde{L}}{\beta}. \end{aligned} \tag{2.5.3}$$

By Lemma 2.3.5, $\tilde{R}'$ is a robust $(\nu/4, 3\tau)$-outexpander. Note that it is a subdigraph of $\tilde{R}'(\beta) \subseteq \tilde{R}(\beta)$ and that all of its edges have multiplicity at least $q := d'/\beta$ in $\tilde{R}'(\beta)$. Let

$$\tilde{r} := (\tilde{\alpha} - \gamma)\tilde{L}/\beta. \tag{2.5.4}$$

For each cluster U, let $n_U^{\pm} := d^{\pm}_{\tilde{R}(\beta)}(U) - \tilde{r}$ and let $\rho := \gamma/\beta$, so $\rho \leq q\nu^2/3$. Note that

$$(1 - \frac{4d}{\gamma})\rho\tilde{L} = (\gamma - 4d)\frac{\tilde{L}}{\beta} \leq n_U^{\pm} \leq (\gamma + 2\tilde{\varepsilon}^{12})\frac{\tilde{L}}{\beta} = (1 + \frac{2\tilde{\varepsilon}^{12}}{\gamma})\rho\tilde{L}.$$

So we can apply Lemma 2.4.1 to $(G, Q) := (\tilde{R}', \tilde{R}'(\beta))$ to obtain a sub-multidigraph W of $\tilde{R}'(\beta)$ (and hence of $\tilde{R}(\beta)$) such that the in- and outdegrees of each cluster U are exactly $n_U^{\pm}$. So $\tilde{R}(\beta) \setminus W$ is a spanning $\tilde{r}$-regular sub-multidigraph of $\tilde{R}(\beta)$. Apply

Proposition 2.4.3 to decompose $\tilde{R}(\beta) \setminus W$ into $\tilde{r}$ 1-factors $\tilde{F}_1, \ldots, \tilde{F}_{\tilde{r}}$ of $\tilde{R}(\beta)$. So each $\tilde{F}_t$ corresponds to a collection of blown-up cycles spanning $V(G) \setminus V_0$. Note that this step would not work if we only considered $\tilde{R}$ and $\tilde{R}(\beta)$ and tried to apply Lemma 2.4.1 to find W in $\tilde{R}(\beta)$ directly.

2.5.2 Thin auxiliary digraphs H

We now define edge-disjoint subdigraphs H_0^+, H_0^-, H_1^+, H_1^- and H_2 of G, which are sparse 'shadows' of the reduced multidigraph. They act as reservoirs of well-distributed edges which will be used at various stages in the proof. The role of $H_0^\pm$ is to connect blown-up cycles (in Section 2.5.5) to ensure that our final merging procedure does indeed yield Hamilton cycles. In Section 2.5.6 edges will be taken from $H_1^\pm$ to connect the vertices in the special exceptional sets $V_{0,i}$ (defined later) to the non-exceptional vertices in each slice G_i (defined in Section 2.5.3). H_2 will be used to construct 'balancing edges' which will be introduced in Section 2.5.8. We choose these subdigraphs already at this point because if we remove them later then this might destroy the superregularity of the pairs in the G_i.

We obtain $H_0^+, H_0^-, H_1^+, H_1^-, H_2$ as follows. Each has vertex set $V(G)$ and initially contains no edges. Then, for each edge E of $\tilde{R}(\beta)$, $G'(E)$ is a $(\sqrt{\tilde{\varepsilon}}, \beta)$-regular pair and we can apply Lemma 2.3.9(i) to $G'(E)$ with $\gamma_1 := \beta_1$ where

$$\beta_1 := (1 - 5\gamma)\beta \tag{2.5.5}$$

and $\gamma_2 := \ldots = \gamma_6 := \gamma\beta$, to obtain six edge-disjoint pairs $J_1, \ldots, J_6$, where J_k is $(\tilde{\varepsilon}^{1/24}, \gamma_k)$-regular, and we call these digraphs $G^*(E)$, $H_0^+(E)$, $H_0^-(E)$, $H_1^+(E)$, $H_1^-(E)$ and $H_2(E)$ respectively. We denote the union of $H(E)$ over all edges E of $\tilde{R}(\beta)$ by H. We will only use the weaker bounds that the 'remaining' subdigraph $G^*(E)$ of $G'(E)$ is $(\varepsilon/8, \beta_1)$-regular and for each $H = H_0^+, H_0^-, H_1^+, H_1^-, H_2$ we have that $H(E)$ is $(\varepsilon, \gamma\beta)$-regular. Moreover, Lemma 2.3.9(i) implies that if E is an edge from A to B and if $x \in A$

and $y \in B$ are good for E, then

$$d^+_{H(E)}(x), d^-_{H(E)}(y) = (\gamma\beta \pm 2\varepsilon)\tilde{m}. \tag{2.5.6}$$

Note also that V_0 is isolated in each H. We now derive some further properties of these digraphs which we will need later. Firstly, we have the following property for H_0^+ and H_0^-:

(H0) Suppose that $\tilde{A}\tilde{B}$ is an edge of $\tilde{R}$. Then for at least $(1-\varepsilon')|\tilde{A}|$ of the vertices $x \in \tilde{A}$ and $(1-\varepsilon')|\tilde{B}|$ of the vertices $y \in \tilde{B}$ we have

$$|N^+_{H_0^+}(x) \cap \tilde{B}| \geq \gamma d\tilde{m}/2 \quad \text{and} \quad |N^-_{H_0^-}(y) \cap \tilde{A}| \geq \gamma d\tilde{m}/2.$$

To see this, note first that every edge E of $\tilde{R}$ has multiplicity at least d/β in $\tilde{R}(\beta)$. Let $E_1, \ldots, E_\ell$ be the edges of $\tilde{R}(\beta)$ corresponding to E. So $d/\beta \leq \ell \leq 1/\beta$. Recall that $H_0^+(E_i)$ is $(\varepsilon, \gamma\beta)$-regular. Let A' be the set of all vertices $x \in \tilde{A}$ such that x has outdegree at least $(\gamma\beta - 2\varepsilon)\tilde{m}$ in each of $H_0^+(E_1), \ldots, H_0^+(E_\ell)$. Then $|A'| \geq (1-\ell\varepsilon)\tilde{m} \geq (1-\varepsilon')\tilde{m}$. Moreover, for all $x \in A'$, we have

$$|N^+_{H_0^+}(x) \cap \tilde{B}| \geq \ell(\gamma\beta - 2\varepsilon)\tilde{m} \geq \frac{d}{\beta}\frac{\gamma\beta}{2}\tilde{m} = \frac{\gamma d\tilde{m}}{2}.$$

The proof of the second inequality is similar.

We also have the following property of H_1^+ and H_1^-:

(H1) For all $x \in V(G) \setminus V_0$, we have $\gamma\tilde{\alpha}n/3 \leq d^\pm_{H_1^+}(x), d^\pm_{H_1^-}(x) \leq 2\gamma\tilde{\alpha}n$.

Recall the definitions of 'good' and 'bad' from Section 2.5.1:

Let A be a cluster of $\tilde{R}$ and let $E(A)$ denote the set of edges incident to A in $\tilde{R}(\beta)$. For an edge E in $E(A)$ and $x \in A$, we say that the pair (x, E) is *good* if

- A is the initial cluster of E and $d^+_{G'(E)}(x) = (\beta \pm 2\tilde{\varepsilon})\tilde{m}$; or

- A is the final cluster of E and $d^-_{G'(E)}(x) = (\beta \pm 2\tilde{\varepsilon})\tilde{m}$

and say it is *bad* otherwise (recall that $\tilde{m}$ is the cluster size). We say that x is *good* if x forms a bad pair with at most $\xi|E(A)|$ edges in $E(A)$.

Now, (H1) follows from the fact that V_0 contains all the bad vertices (in the sense of Section 2.5.1). Indeed, since any vertex $x \in V(G) \setminus V_0$ is good we have

$$d^+_{H_1^+}(x) \overset{(2.5.6)}{\geq} \delta^+(\tilde{R}(\beta))(1-\xi)(\gamma\beta - 2\varepsilon)\tilde{m} \overset{(2.5.3)}{\geq} \frac{\tilde{\alpha}\tilde{L}}{2\beta}\gamma\beta\tilde{m} \geq \gamma\tilde{\alpha}n/3.$$

The other bounds in (H1) follow similarly.

2.5.3 Unwinding cycles

For each $1 \leq t \leq \tilde{r}$ we now apply Lemma 2.3.10 to each cycle in $\tilde{F}_t$ to remove vertices from each cluster, so that they now have size exactly $(1-\varepsilon/4)\tilde{m}$ and such that each edge E of $\tilde{F}_t$ corresponds to an $(\varepsilon/2, \beta_1)$-superregular pair $G^*(E)$. By removing at most $2sp$ further vertices from each cluster we obtain clusters of size m such that $2sp \mid m$. We call these *adapted primary clusters* or *adapted primary* (t)*-clusters* if we wish to emphasise the dependence on t, and say that each such cluster is *associated* with the base primary cluster from which it was formed. Since $2sp \leq \varepsilon\tilde{m}/4$ it is easy to see that now each edge E of $\tilde{F}_t$ corresponds to an (ε, β_1)-superregular pair $G^*(E)$. Note that

$$\frac{1}{m} \leq \frac{2\tilde{L}}{n} \leq \frac{2M}{n_0} \ll \frac{1}{\tilde{L}} \quad \text{and} \quad (1-\varepsilon)n \overset{(2.5.2)}{\leq} m\tilde{L} \leq \tilde{m}\tilde{L} \leq n. \tag{2.5.7}$$

Let $\tilde{V}^{\text{spec}}_{0,t}$ denote the set of all those vertices in G which were removed from the clusters in this step. We call them the *special exceptional vertices* (for the original slice t). So $|\tilde{V}^{\text{spec}}_{0,t}| \leq \varepsilon n/4 + 2sp\tilde{L} \leq \varepsilon n/2$. Let $\tilde{V}_{0,t} = V_0 \cup \tilde{V}^{\text{spec}}_{0,t}$. Then

$$|\tilde{V}_{0,t}| \overset{(2.5.2)}{\leq} \frac{2\varepsilon n}{3}. \tag{2.5.8}$$

We denote the collection of the adapted primary (t)-clusters together with the exceptional set $\tilde{V}_{0,t}$ by $\mathcal{P}(t)$. Note that $\mathcal{P}(t)$ and $\tilde{\mathcal{P}}$ are $2\varepsilon/3$-close for every $1 \leq t \leq \tilde{r}$ (recall that this notion was defined before Lemma 2.3.14).

For each cycle C in a given 1-factorisation, we would like to ensure that the outneighbourhood of an exceptional vertex is well-distributed on each cycle, in the sense that each cluster V of C contains only a small fraction of the neighbours of any exceptional vertex. Currently, we cannot guarantee this. But we will be able to achieve this property by considering a refinement of the partition $\mathcal{P}(t)$ for each t. As associated clusters in each $\mathcal{P}(t)$ only differ slightly from one another, we can find this refinement in such a way that the subclusters are also similar by ensuring that all such refinements are close to a refinement of $\tilde{\mathcal{P}}$.

Let $\mathcal{G} = \{G, H_0^+, H_0^-, H_1^+, H_1^-, H_2\}$. We now apply Lemma 2.3.13 to our base primary clusters and exceptional set V_0 to obtain an $\tilde{\varepsilon}$-uniform s-refinement $\mathcal{P}'_s$ of $\tilde{\mathcal{P}}$ with respect to $\mathcal{G}$, and we call the resulting subclusters *base s-clusters*. So we have $L_s := s\tilde{L}$ base s-clusters. Apply Lemma 2.3.13 to $\mathcal{P}'_s$ to obtain an $\tilde{\varepsilon}$-uniform p-refinement $\mathcal{P}'_p$ of $\mathcal{P}'_s$ with respect to $\mathcal{G}$. Let

$$L_p := pL_s = sp\tilde{L}. \tag{2.5.9}$$

We call the L_p subclusters obtained from an s-cluster *base p-clusters*. By the remark before (2.3.1), $\mathcal{P}'_p$ is also a $3\tilde{\varepsilon}$-uniform sp-refinement of $\tilde{\mathcal{P}}$. Finally apply Lemma 2.3.13 to $\mathcal{P}'_p$ to obtain an $\tilde{\varepsilon}$-uniform 2-refinement $\mathcal{P}'_{2p}$ of $\mathcal{P}'_p$ with respect to $\mathcal{G}$. The argument before (2.3.1) implies that $\mathcal{P}'_{2p}$ is a $4\tilde{\varepsilon}$-uniform $2p$-refinement of $\mathcal{P}'_s$ and a $5\tilde{\varepsilon}$-uniform $2sp$-refinement of $\tilde{\mathcal{P}}$. We call the subclusters obtained from an s-cluster *base $2p$-clusters*.

Define constants $\varepsilon_s, \varepsilon_p, \varepsilon_{2p}$ such that $\varepsilon \ll \varepsilon_s \ll \varepsilon_p \ll \varepsilon_{2p} \ll \varepsilon'$. Now do the following for each t with $1 \leq t \leq \tilde{r}$. Apply Lemma 2.3.14 to $\tilde{\mathcal{P}}$ to obtain an ε_s-uniform s-refinement $\mathcal{P}_s(t)$ of $\mathcal{P}(t)$ that is ε_s-close to $\mathcal{P}'_s$. Next apply Lemma 2.3.14 to $\mathcal{P}'_s$ to obtain an ε_p-uniform p-refinement $\mathcal{P}_p(t)$ of $\mathcal{P}_s(t)$ that is ε_p-close to $\mathcal{P}'_p$. By the observation at the end of Section 2.3.4, $\mathcal{P}_p(t)$ is also an ε'-uniform sp-refinement of $\mathcal{P}(t)$. Finally apply Lemma 2.3.13 to $\mathcal{P}'_p$ to obtain an ε_{2p}-uniform 2-refinement $\mathcal{P}_{2p}(t)$ of $\mathcal{P}_p(t)$ that is ε_{2p}-

close to $\mathcal{P}'_{2p}$. Again, $\mathcal{P}_{2p}(t)$ is an ε'-uniform $2p$-refinement of $\mathcal{P}_s(t)$ and an ε'-uniform $2sp$-refinement of $\mathcal{P}(t)$. For $j = s, p, 2p$ we call the clusters of $\mathcal{P}_j(t)$ the *(adapted) j-clusters* or *j-(t)-clusters* if we wish to emphasise the dependence on t. For each such j we have that $\mathcal{P}_j(t)$ is an ε'-uniform refinement of $\mathcal{P}(t)$ that is ε'-close to $\mathcal{P}'_j$, so each adapted j-cluster in $\mathcal{P}_j(t)$ is associated with a unique base j-cluster in $\mathcal{P}'_j$. Write

$$m_s := m/s \quad \text{and} \quad m_p := m/sp \tag{2.5.10}$$

for the respective sizes of the s- and p-clusters (which are the same for all t though the clusters themselves are different). Note that

$$m_p \leq \frac{n}{L_p} \leq 2m_p. \tag{2.5.11}$$

By a slight abuse of notation we can consider $\tilde{R}$ and $\tilde{R}(\beta)$ as digraphs on either base or adapted (t)-clusters, depending on the context. For each $0 \leq t \leq \tilde{r}$, we now define corresponding reduced digraphs for the refinements defined above, where, for convenience, $\mathcal{P}_j(0) := \mathcal{P}'_j$.

Let $R_s = s \otimes \tilde{R}$ be the s-fold blow-up of $\tilde{R}$, where for a vertex W of $\tilde{R}$ (which is an adapted primary (t)-cluster if $t \geq 1$), the corresponding vertices in R_s are the subclusters of W in $\mathcal{P}_s(t)$. Define $R_s(\beta) = s \otimes \tilde{R}(\beta)$ analogously. Also let $R_p = p \otimes R_s$, where for a vertex U of R_s the corresponding vertices in R_p are the subclusters of U in $\mathcal{P}_p(t)$. So the vertices of R_p are precisely the p-clusters and also $R_p = sp \otimes \tilde{R}$. Define $R_p(\beta) = p \otimes R_s(\beta)$ $= sp \otimes \tilde{R}(\beta)$ analogously. Note that apart from the fact that the clusters which form their vertex sets are slightly different for different values of t, these digraphs are the same, so there is no need for any dependence on t in the notation.

Suppose that $\tilde{E}$ is an edge of $\tilde{R}(\beta)$ from $\tilde{U}$ to $\tilde{W}$ and that U is an s-cluster which is a subcluster of $\tilde{U}$ and W is an s-cluster which a subcluster of $\tilde{W}$. Note that there is a unique edge E in $R_s(\beta)$ from U to W corresponding to $\tilde{E}$. Thus to each edge E of $R_s(\beta)$

we can associate the digraph

$$G^*(E) := G^*(\tilde{E})[U, W]. \tag{2.5.12}$$

We make a similar association for each edge F of $R_p(\beta)$ by defining $G^*(F)$ analogously.

We now use the 1-factors $\tilde{F}_t$ to define edge-disjoint 1-factors F'_j in the reduced digraph $R_s(\beta)$ and then use the F'_j to find edge-disjoint 1-factors F_i in $R_p(\beta)$. Note that each cycle C of $\tilde{F}_t$ corresponds to an s-fold blow-up C' of C in $R_s(\beta)$. So for each cycle C in $\tilde{F}_t$ we can apply Lemma 2.4.4 to obtain an $(s-1)$-unwinding $C_1, \ldots, C_{s-1}$ of C'. Here we do not need the special properties of the $(s-1)$-unwinding which are guaranteed by Lemma 2.4.4; in fact any unwinding yielding edge-disjoint Hamilton cycles will do. So $\tilde{F}_t$ corresponds to a set of $(s-1)$ 1-factors F'_j (with $(t-1)(s-1)+1 \leq j \leq t(s-1)$) of $R_s(\beta)$. We say that such an F'_j has *original factor type* t (and that t is the *original type* of j). Note that for each cluster W of $\tilde{R}$, there are s clusters of F'_j which are subclusters of W. Moreover, all of these lie on the same cycle of F'_j. Let

$$r_s := (s-1)\tilde{r}. \tag{2.5.13}$$

Then altogether this gives us a set of r_s edge-disjoint 1-factors $F'_1, \ldots, F'_{r_s}$ of $R_s(\beta)$.

Consider any cycle C_ℓ of a 1-factor F'_j obtained from a cycle C of $\tilde{F}_t$ as above. Let K be the length of C; so C_ℓ has length Ks. We say that an s-cluster lying on C_ℓ is *clean for* F'_j if it belongs to the last K clusters of C_ℓ (where for each cycle we pick a consistent ordering of its vertices in advance). Note that $K \geq 2$ and so C_ℓ has at least two clean s-clusters. Moreover, for each adapted primary cluster W, exactly one subcluster of W in $\mathcal{P}_s(t)$ is clean for F'_j. Note that for different 1-factors F'_j, the set of clean clusters will usually be different.

It turns out that we actually need a stronger property than the one described above, namely we need that $(\star)$ below holds. (This will enable us to ensure that, in the digraphs G_i that we consider later, only a few clusters will contain vertices sending or receiving

an edge from the exceptional set and these will be sufficiently far apart.) For this, we use our refinement $\mathcal{P}_p(t)$ of each s-cluster into p subclusters and unwind the cycles in the above 1-factorisation again.

For every $V \in \mathcal{P}_s(t)$, let $V^1, \dots, V^p$ be the p-clusters contained in V. Note that the collection of all V^k over all s-clusters V contained in an adapted primary cluster W are precisely the p-clusters refining W. For each cycle $D = V_1 \dots V_K$ in F'_j (where this is the same ordering we specified above) let D' be the p-fold blow-up of D whose vertex classes are the p-clusters V_ℓ^k contained in $V_1, \dots, V_K$. Apply Lemma 2.4.4 to D' to find a $(p-1)$-unwinding $D_1, \dots, D_{p-1}$ of D' with V_ℓ^k playing the role of x_ℓ^k. We have the following property:

($\star$) For each $1 \leq d \leq p-1$ and $1 \leq k \leq p$, the p-clusters $V_1^k, \dots, V_{K-2}^k$ have pairwise distance at least p on D_d.

Note that ($\star$) holds only for the p-clusters in $V_1, \dots, V_{K-2}$ and not necessarily V_{K-1} or V_K. This is the reason for introducing the clean s-clusters: recall that V_{K-1} and V_K are clean. This will mean that we will never introduce any edges between their vertices and the exceptional set (see (b) in Section 2.5.4).

Moreover, for all $1 \leq \ell \leq K$, we have from Lemma 2.4.4 that $V_\ell^1, \dots, V_\ell^p$ lie on the same cycle D_d. Additionally, their successors on D_d all belong to a single adapted primary cluster $V_{\ell+1}$. Also F'_j gives rise to a set of $(p-1)$ edge-disjoint 1-factors F_i (with $(j-1)(p-1)+1 \leq i \leq j(p-1)$) of $R_p(\beta)$. We say that such an F_i has *intermediate factor type* j and *original factor type* t where t is the original type of F'_j. For each i, write $V_{0,i}^{\text{spec}} := \tilde{V}_{0,t}^{\text{spec}}$ for the special exceptional set associated with F_i, where t is the original factor type of F_i. Note that for every i, every vertex in G is contained either in a p-cluster of F_i, in V_0 or in $V_{0,i}^{\text{spec}}$. Note also that for each adapted primary cluster W of $\tilde{R}$, there are sp clusters of F_i which are subclusters of W. Also let

$$r_p := (p-1)r_s. \tag{2.5.14}$$

Then altogether this gives us a set of r_p edge-disjoint 1-factors $F_1, \dots, F_{r_p}$ of $R_p(\beta)$. Note that for each t, there are exactly $(s-1)(p-1)$ of the F_i which have original factor type t. Furthermore,

$$r_p \overset{(2.5.4)}{\leq} \frac{\tilde{\alpha} sp\tilde{L}}{\beta} \quad \text{and so} \quad r_p \overset{(2.5.9)}{\leq} \frac{\tilde{\alpha} L_p}{\beta}; \quad \text{also} \quad 1/m \leq 1/m_p \overset{(2.5.7)}{\ll} 1/r_p. \tag{2.5.15}$$

For each edge $E \in E(F_i)$ from A to B, let $G_i(E) := G^*(E)$, where $G^*(E)$ was defined just after (2.5.12). Let G_i denote the union of the digraphs $G_i(E)$ over all E with $E \in E(F_i)$ and call it the ith *slice*. Clearly $G_1, \dots, G_{r_p}$ are edge-disjoint. Given $E \in E(F_i)$ for some $1 \leq i \leq r_p$, let $\tilde{E} \in E(\tilde{F}_t)$ be the unique edge such that E is in the blow-up of $\tilde{E}$, where F_i has original factor type t. As noted directly before (2.5.7), $G^*(\tilde{E})$ is (ε, β_1)-superregular and hence $G_i(E) = G^*(E)$ is (ε', β_1)-superregular by Lemma 2.3.13(i).

Recall that since $V_{0,i}^{\text{spec}}$ is different for each i, the vertex set of a p-cluster will be slightly different in G_i and $G_{i'}$ when F_i and $F_{i'}$ have different original factor types. Note that if U is a base $2p$-cluster (of size $m_p/2$), and $U_{(t)}$ is the associated $2p$-(t)-cluster, then

$$|U \cap U_{(t)}| \geq (1-\varepsilon')m_p/2. \tag{2.5.16}$$

as the corresponding partitions are ε'-close. (On the other hand, $\bigcap_{1 \leq t \leq \tilde{r}} U_{(t)}$ may be empty.) The same statements hold for s- and p-clusters. When adding edges incident to exceptional vertices in Section 2.5.6 we need to be careful about distinguishing between base $2p$-clusters and the $2p$-(t)-clusters which are actually contained in the clusters of our slices.

2.5.4 Red clusters and edges

The aim of this section is to lay some groundwork for Sections 2.5.5, 2.5.6 and 2.5.7 by specifying the properties that the edges between the exceptional vertices and the rest of $V(G)$ need to satisfy. In Section 2.5.5 our aim is to remove a bounded number of *bridge*

vertices $V_{0,i}^{\text{bridge}}$ from each $G_i \setminus (V_0 \cup V_{0,i}^{\text{spec}})$ and change their neighbourhoods in such a way that the blown-up cycles in G_i are connected via bridge vertices. Some additional vertices will also be removed and added to $V_{0,i}^{\text{spec}}$ to keep the cluster sizes equal. In Section 2.5.6 we will add edges to G_i which are incident to V_0. In Section 2.5.7 we will do the same for $V_{0,i}^{\text{spec}}$. We will then let

$$V_{0,i} := V_0 \cup V_{0,i}^{\text{spec}} \cup V_{0,i}^{\text{bridge}}.$$

$V_{0,i}$ is then the exceptional set for the slice G_i: each vertex will lie either in a cluster of G_i or in $V_{0,i}$. Any edge incident to a vertex in $V_{0,i}$ and any vertex in a cluster of G_i incident to such an edge will be called *i-red* (or *red* if this is unambiguous). Roughly speaking, when adding red edges to G_i, we will need to ensure that G_i is a spanning almost-regular digraph, that no non-exceptional vertex has large i-red degree and that the set of red vertices is small and well-distributed.

To achieve this, for each i we will only add red edges incident to some carefully selected $2p$-clusters and then apply property $(\star)$ (which we derived from Lemma 2.4.4). More precisely, for fixed i, let $j = j(i)$ and $t = t(i)$ respectively be the intermediate and original factor types of G_i. For each s-cluster U of $\mathcal{P}_s(t)$, let $U_1, \dots, U_p$ denote the p-clusters of $\mathcal{P}_p(t)$ which are subclusters of U. For $1 \leq \ell \leq p$, let $U(\ell)$ and $U(\ell + p)$ be the $2p$-clusters contained in U_ℓ. In G_i, we will add red edges between $V_{0,i}$ and $U(k)$ only if

(a) $t \equiv k \mod 2p$ and

(b) U is not a clean cluster in F'_j.

We call such a $2p$-cluster $U(k)$ *i-red* and we call a p-cluster *i-red* if it contains an i-red $2p$-cluster (or simply *red* if this is unambiguous). Note that (a) implies that every s-cluster U which is not clean contains exactly one red $2p$-cluster (and thus exactly one red p-cluster). Moreover, recall that any adapted primary cluster contains exactly one clean s-cluster; thus it contains exactly $s - 1$ red p-clusters.

All red vertices will be contained in red $2p$-clusters, but note that we do not require every red cluster to contain a red vertex. Let

$$\kappa := (1-\gamma)\beta_1 m_p. \tag{2.5.17}$$

We would like to find exactly κ edge-disjoint Hamilton cycles in each of the G_i. For this, we will first need to add edges so that G_i satisfies the following for all i with $1 \le i \le r_p$:

(Red0) There exists a sequence $D_1 x_1 D_2 x_2 \dots x_{\ell-1} D_\ell x_\ell D_1$ with the following properties:

- Each D_j is a cycle of F_i and every cycle of F_i appears at least once in the sequence;
- $V_{0,i}^{\text{bridge}} := \{x_1, \dots, x_\ell\}$ and each x_j has exactly κ outneighbours in D_{j+1} and exactly κ inneighbours in D_j.

(Red1) $d^{\pm}_{G_i}(x) = \kappa$ for all $x \in V_{0,i}$;

(Red2) $V_{0,i}$ is an independent set in G_i;

(Red3) $|N^{\pm}_{G_i}(y) \cap V_{0,i}| \le \sqrt{\xi}\beta_1 m_p$ for all $y \notin V_{0,i}$;

(Red4) For every red p-cluster V, all red edges of G_i are incident to a single $2p$-cluster contained in V. In particular, $|N^{\pm}_{G_i}(V_{0,i}) \cap V| \le m_p/2$ for all clusters $V \in R_p$;

(Red5) If V, V' are red p-clusters on a cycle C of F_i, then they have distance at least p on C;

(Red6) If a p-cluster V contains the final vertex of a red edge in G_i, then it contains no initial vertices of red edges in G_i, and vice versa;

(Red7) $G_1, \dots, G_{r_p}$ are edge-disjoint and $G_i(E)$ is $(2\varepsilon', \beta_1)$-superregular for all $E \in E(F_i)$.

Roughly speaking, given a 1-factor f of G_i, (Red0) and (Red1) will ensure that f has a path between any pair of successive cycles D_j of F_i. (Red2)–(Red6) imply that the red

edges are well-distributed. This will be crucial when applying Lemma 2.4.5 to transform f into a Hamilton cycle in Section 2.5.11.

Suppose that V is a red p-cluster. Since only one of the two $2p$-clusters contained in V is red, it follows that (Red4) will automatically be satisfied for V if we add red edges according to (a) and (b). It is easy to see that this will also satisfy (Red5). Indeed, recall that every non-clean s-cluster contains exactly one red p-cluster. Moreover, if U_ℓ and $U_{\ell'}$ are non-clean s-clusters then the p-cluster U_ℓ^k is red if and only if $U_{\ell'}^k$ is red. Suppose a cycle C in F_i was obtained by unwinding the blow-up of C' in F'_j; then the last two s-clusters in C' (using the same ordering as in Section 2.5.3) are clean and hence contain no red p-clusters by (b). So $(\star)$ implies that the red clusters on C will have distance at least p apart.

Let $F = r_p/4p$. For each k with $1 \leq k \leq 2p$, note that the number of all digraphs G_i whose original type t satisfies (a) is $2F$. For each k, consider an ordering of all these graphs.

We will now fix which of the (red) $2p$-clusters will receive red edges and which of them will send out red edges. For each $i = 1, \ldots, r_p$, let $t = t(i)$ be the original type of G_i and let k with $1 \leq k \leq 2p$ satisfy (a). Suppose that G_i is the fth graph with original type t (so $1 \leq f \leq 2F$). For each adapted primary cluster $W \in \mathcal{P}(t)$, let $W_1, \ldots, W_s$ denote the set of all s-clusters contained in W. Recall that exactly one of the W_j is clean. Now choose a set S_W^+ of s-clusters from $W_1, \ldots, W_{s/2}$ so that none of the s-clusters in S_W^+ is clean and so that $|S_W^+| = s/2 - 1$ (recall that s is even). Let I_W^+ denote the set of indices of the s-clusters in S_W^+. Similarly choose S_W^- from $W_{s/2+1}, \ldots, W_s$ with $|S_W^-| = s/2 - 1$ which avoids the clean cluster and let I_W^- be the corresponding set of indices.

For each s-cluster W_ℓ contained in W, let $W_\ell(k)$ denote the kth $2p$-cluster contained in W_ℓ, where k is defined as in the previous paragraph. We call $W_\ell(k)$ *in-red* (for i) if

- $1 \leq f \leq F$ and $\ell \in I_W^+$, or $F < f \leq 2F$ and $\ell \in I_W^-$.

We call $W_\ell(k)$ *out-red* (for i) if

- $1 \leq f \leq F$ and $\ell \in I_W^-$, or $F < f \leq 2F$ and $\ell \in I_W^+$.

If a p-cluster V contains an in-red $2p$-cluster, we say that V is *in-red* (and similarly for *out-red* clusters). So the number of in-red p-clusters in each adapted primary cluster W is exactly

$$|I_W^-| = \frac{s}{2} - 1 \tag{2.5.18}$$

and similarly for out-red clusters.

2.5.5 Connecting blown-up cycles

In the final section of the proof we will successively find 1-factors in each G_i and then turn each of these into a Hamilton cycle. As mentioned earlier, (Red0) will be used to ensure that each 1-factor f of G_i has a path connecting any pair of consecutive cycles of F_i, which will make it possible to merge the cycles of f into a Hamilton cycle. In this section we will modify G_i so that (Red0) holds.

We will join cycles by choosing *bridge vertices* $x_{i,j}$ in $V(G) \setminus (V_0 \cup V_{0,i}^{\text{spec}})$ whose neighbourhoods will be chosen from the sparse digraphs $H_0^{\pm}$ defined in Section 2.5.2. In what follows, we write A_j^- for the predecessor of the p-cluster A_j in F_i.

Claim 2.5.1. *There is a sequence $A_1B_1A_2B_2 \dots A_{\tilde{L}}B_{\tilde{L}}A_1$ of p-clusters in R_p such that, for each $1 \leq j, j' \leq \tilde{L}$ the following hold:*

(i) Let $\tilde{A}_j$ and $\tilde{B}_j$ be the adapted primary clusters containing A_j and B_j respectively. Then there is an edge from $\tilde{A}_j$ to $\tilde{B}_j$ in $\tilde{R}$.

(ii) A_j^- is out-red and B_j is in-red;

(iii) B_j and A_{j+1} lie in the same adapted primary cluster (where $A_{\tilde{L}+1} := A_1$).

(iv) Every adapted cluster contains exactly one A_j and exactly one $B_{j'}$.

(v) All the A_j and $B_{j'}$ are distinct.

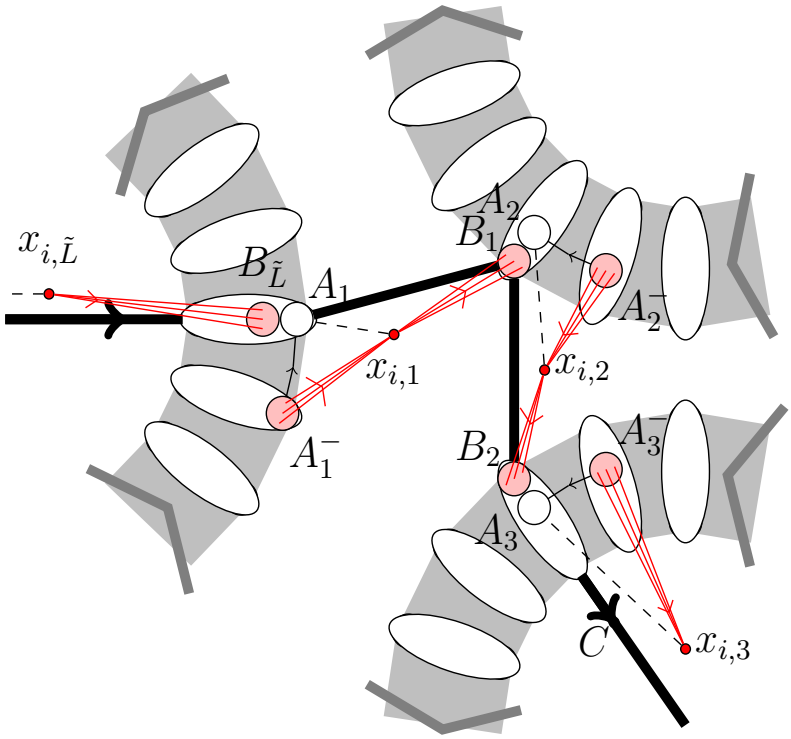

Figure 2.2: Bridge vertices $x_{i,1}, x_{i,2}, x_{i,3}$ chosen from p-clusters A_1, A_2, A_3 respectively.

Proof. Observe that, by Lemma 2.3.5, $\tilde{R}$ is a robust $(\nu/4, 3\tau)$-outexpander, so Theorem 2.4.6 implies that $\tilde{R}$ contains a Hamilton cycle $C = \tilde{A}_1 \dots \tilde{A}_{\tilde{L}}$. We will choose A_j in $\tilde{A}_j$ and B_j in $\tilde{A}_{j+1}$. This will automatically satisfy (i), (iii) and (iv) and ensures that they will be distinct, except possibly $A_j = B_{j-1}$. Now recall that, by (2.5.18), each adapted primary cluster contains exactly $s/2-1$ in-red and $s/2-1$ out-red p-clusters. Moreover, as noted after $(\star)$, they all lie on the same cycle in F_i and p-clusters directly preceding those in $\tilde{A}_j$ on F_i all lie in the same adapted primary cluster, which we call $\tilde{A}_j^-$. Thus we can always choose an in-red B_{j-1} in $\tilde{A}_j$ and an out-red $A_j^- \in \tilde{A}_j^-$ whose successor A_j on F_i lies in $\tilde{A}_j$, proving (ii). Moreover, we have $s/2-1 > 1$ choices for A_j so we may assume that A_j and B_{j-1} are distinct. This proves (v) and thus the claim. ■

We will choose the bridge vertices in the sets A_j. The next claim guarantees many candidates for these bridge vertices whose neighbourhoods have the required properties.

Claim 2.5.2. *For each i with $1 \leq i \leq r_p$, whenever a p-cluster $A_j \in V(F_i)$ is joined by an edge in R_p to a p-cluster $B_j \in V(F_i)$ the following holds. Let $(A_j^-)'$ and B_j' be $2p$-clusters contained in A_j^- and B_j respectively. Then A_j contains at least $m_p/2$ vertices x such that both $|N^-_{H_0^-}(x) \cap (A_j^-)'| > \kappa$ and $|N^+_{H_0^+}(x) \cap B_j'| > \kappa$.*

We say that a vertex x as in Claim 2.5.2 is (i,j)-*useful*.

Proof of Claim 2.5.2. Note that (H0) and (URef) imply that, for at least $3m_p/4$ of the vertices $x \in A_j$, we have

$$|N^+_{H_0^+}(x) \cap B_j'| \geq \gamma d m_p/5 \overset{(2.5.17)}{>} \kappa.$$

As F_i (and thus R_p) contains the edge $A_j^- A_j$, for at least $3m_p/4$ of the vertices $x \in A_j$ we similarly have

$$|N^-_{H_0^-}(x) \cap (A_j^-)'| > \kappa.$$

So at least $m_p/2$ the vertices in A_j satisfy both inequalities, which proves the claim. ■

Now we choose the set $V_{0,i}^{\text{bridge}}$ satisfying (Red0). For each $1 \leq i \leq r_p$, consider the sequence guaranteed by Claim 2.5.1 and for each $1 \leq j \leq \tilde{L}$, let $(A_j^-)^{\text{red}}$ and B_j^{red} be the unique red $2p$-clusters contained in A_j^- and B_j respectively. So $(A_j^-)^{\text{red}}$ is out-red and B_j^{red} is in-red. For each $1 \leq j \leq \tilde{L}$, apply Claim 2.5.2 to the pair (A_j, B_j) with $(A_j^-)^{\text{red}}, B_j^{\text{red}}$ playing the roles of $(A_j^-)', B_j'$ respectively, to obtain a vertex $x_{i,j} \in A_j$ which is (i,j)-useful and which is distinct from all vertices chosen so far. Note that the latter is possible since Claim 2.5.1(v) implies that for each i, we only choose one vertex from A_j. So altogether, we choose at most r_p vertices from each A_j, which is at most $m_p/3$ by (2.5.15). In each G_i, remove each $x_{i,j}$ from A_j and denote the collection of all $x_{i,j}$ with $1 \leq j \leq \tilde{L}$ by $V_{0,i}^{\text{bridge}}$. This process is illustrated in Figure 2.2.

Now, for each $1 \leq i \leq r_p$, there are exactly $sp - 1$ of the p-clusters in each adapted

primary (t)-cluster of G_i from which no vertices have been removed in this step, where t is the original type of F_i. We still need to ensure that p-clusters of G_i have equal size, so we choose a further $(sp-1)r_p\tilde{L} \leq \varepsilon n/3$ distinct vertices such that exactly one is removed from each untouched p-cluster in each G_i. Each such vertex is moved from its cluster into $V_{0,i}^{\text{spec}}$. The final inequality in (2.5.15) implies that we can assume that each vertex x is moved into $V_{0,i}^{\text{spec}}$ for at most one $1 \leq i \leq r_p$ in this step. Now adapted primary (t)-clusters become *adapted primary* $[i]$*-clusters* and (adapted) s-, p- and $2p$-(t)-clusters become *(adapted)* s-, p- and $2p$-$[i]$-*clusters* respectively (or s-, p-, $2p$-*clusters* if this is unambiguous). This will not overlap with previous notation as from now on we never refer to (t)-clusters and only ever refer to base and $[i]$-clusters. (2.5.16) implies that if U is a base $2p$-cluster and $U_{[i]}$ is the associated $2p$-$[i]$-cluster, then

$$|U \cap U_{[i]}| \geq (1-\varepsilon')m_p/2 - 1. \tag{2.5.19}$$

We still refer to the cluster sizes m, m_s and m_p in the same way since each one has only lost at most sp vertices (which does not affect any calculations). The $2p$-clusters may no longer have exactly the same size, but this also does not affect any of the calculations. We call $V_{0,i}^{\text{bridge}}$ the set of *bridge vertices* and say that every edge incident to a bridge vertex is i-red. We now have that

$$|V_{0,i}| \overset{(2.5.8)}{\leq} \varepsilon n. \tag{2.5.20}$$

Since in G_i, we removed exactly one vertex from each p-cluster, we still have

$$|N^+_{H_0^+}(x_{i,j}) \cap B_j^{\text{red}}| \geq \kappa \quad \text{and} \quad |N^-_{H_0^-}(x_{i,j}) \cap (A_j^-)^{\text{red}}| \geq \kappa.$$

Since the $x_{i,j}$ are all distinct, it follows that for each $x_{i,j}$, we can choose κ of these outedges from H_0^+ and add them to G_i. Similarly, we can choose κ of these inedges from H_0^- and add them to G_i, whilst also removing every other edge incident to $x_{i,j}$ in G_i. So (Red1)

is satisfied for $V_{0,i}^{\text{bridge}}$.

It is now easy to verify (Red0). For each $1 \leq i \leq r_p$, consider the sequence given by Claim 2.5.1. Let D_j be the cycle of F_i containing the adapted p-cluster A_j for each $1 \leq j \leq \tilde{L}$ and let $x_j := x_{i,j}$ be the bridge vertex which was removed from A_j in G_i. Note that each cycle of F_i appears several times in the sequence. We claim that $D_1 x_1 D_2 x_2 \dots x_{\tilde{L}} D_\ell x_{\tilde{L}} D_1$ is a sequence satisfying (Red0). The first property is immediate from Claim 2.5.1(iv). Each x_j has inneighbourhood contained in A_j^- which is in D_j since A_j is, and its outneighbourhood is contained in B_j which lies in the same adapted cluster as A_{j+1} and thus in the cycle D_{j+1}. Therefore the second property is also satisfied.

(Red2) follows since the in- and outedges incident to bridge vertices were chosen from edge-disjoint subdigraphs H_0^- and H_0^+ respectively. Furthermore, by Claim 2.5.1(v), any $y \notin V_{0,i}$ is incident to at most one i-red edge so (Red3) holds. In each red p-cluster V, red edges were only added to the unique red $2p$-cluster W contained in V so (Red4) is satisfied. (Red5) is satisfied by the comments after the statement of (Red7). Moreover every out-red p-cluster only sends out red edges and every in-red p-cluster only receives red edges so (Red6) holds. The edge-disjointness in (Red7) is immediate from the construction. Finally, note that any vertex in $V(G) \setminus V_{0,i}$ lost at most one inneighbour and one outneighbour in G_i, so for each edge E of F_i, $G_i(E)$ is certainly still $(2\varepsilon', \beta_1)$-superregular. Therefore (Red0) and (Red2)–(Red7) are all satisfied. Note that (Red1) holds for all vertices in $V_{0,i}^{\text{bridge}}$. The aim of the next two sections is to maintain these properties while also achieving (Red1) for all vertices in $V_{0,i}$.

2.5.6 Incorporating the core exceptional set V_0

Note that so far, G_i contains no edges with initial or final vertex in $V_0 \cup V_{0,i}^{\text{spec}}$. In this section and the next we will add edges incident to these vertices into the G_i. Recall that we call such edges and any incident vertices *i-red* or *red* if this is unambiguous. Throughout both sections we will refer to (a) and (b) in Section 2.5.4. To achieve (Red1), we consider the core exceptional set V_0 and the special exceptional set $V_{0,i}^{\text{spec}}$ separately. In this section

we consider the core exceptional set. Roughly speaking, the set of edges between V_0 and $G_i \setminus V_0$ will consist of a random subdigraph of G induced by V_0 and the red $2p$-clusters of G_i. The following claim guarantees the existence of suitable edge-disjoint random subdigraphs. Recall from Section 2.5.4 that $F = r_p/4p$.

Claim 2.5.3. *Let X be a base $2p$-cluster which is a subcluster of a base primary cluster W. Then in G we can find F edge-disjoint bipartite graphs $E_1^+(X), \dots, E_F^+(X)$ with all edges oriented from V_0 to X so that for all $1 \le f \le F$ the following hold:*

(i) For all $x \in V_0$ we have $d^+_{E_f^+(X)}(x) \ge \frac{1-\varepsilon}{2spF}\left(|N_G^+(x) \cap W| - 5\tilde{\varepsilon}\tilde{m}\right)$.

(ii) For all $y \in X$ we have $d^-_{E_f^+(X)}(y) < \sqrt{\xi}\beta_1 m_p/2$.

We can also find $E_1^-(X), \dots, E_F^-(X)$ satisfying analogous properties for the in-neighbourhoods.

Proof. Let $E^+(X)$ denote the digraph induced by the set of edges from V_0 to X in G. Now consider a random partition of the edges of $E^+(X)$ into F parts $E_f^+(X)$. More precisely, assign each edge of $E^+(X)$ to $E_f^+(X)$ with probability $1/F$, independently of all other edges. There are several cases to consider. Say that $x \in V_0$ is *prolific* if $|N_G^+(x) \cap W| > 5\tilde{\varepsilon}\tilde{m}$. Say that V_0 is *large* if $|V_0| \ge \sqrt{\xi}\beta_1 m_p/2$ and *small* otherwise. Every $x \in V_0$ which is not prolific satisfies the condition in (i) with probability 1, and the inequality in (ii) is satisfied with probability 1 if V_0 is small. Suppose that x is prolific. Then since $\mathcal{P}'_{2p}$ is a $5\tilde{\varepsilon}$-uniform $2sp$-refinement of $\tilde{\mathcal{P}}$, (URef) implies that $d^+_{E^+(X)}(x) \ge \frac{1-5\tilde{\varepsilon}}{2sp}|N_G^+(x) \cap W|$.

Then for each $1 \le f \le F$, each prolific $x \in V_0$ and each $y \in X$,

$$\mathbb{E}\left(d^+_{E_f^+(X)}(x)\right) \ge \frac{1-5\tilde{\varepsilon}}{2spF}|N_G^+(x) \cap W| \quad \text{and} \quad \mathbb{E}\left(d^-_{E_f^+(X)}(y)\right) \le \frac{|V_0|}{F}.$$

By Proposition 2.3.1 (with $a := \varepsilon/2$) we have that, for fixed f and prolific

$x \in V_0$,

$$\mathbb{P}\left(d^+_{E^+_f(X)}(x) \leq \frac{1-\varepsilon}{2spF}|N^+_G(x) \cap W|\right) \leq \exp\left(-\frac{5\varepsilon^2\tilde{\varepsilon}\tilde{m}(1-5\tilde{\varepsilon})}{24spF}\right) \overset{(2.5.15)}{\leq} \exp\left(-\frac{\tilde{\varepsilon}^2\beta n}{s^2p\tilde{L}^2}\right) \overset{(2.5.1)}{\leq} e^{-\sqrt{n}}$$

and $|V_0|F \leq n^2$. So taking a union bound over all f and all $x \in V_0$ we see that the probability that (i) fails for some f in this partition is at most $n^2e^{-\sqrt{n}}$. Similarly, for large V_0, fixed f and $y \in X$, Proposition 2.3.1 implies that

$$\mathbb{P}\left(d^-_{E^+_f(X)}(y) > \frac{2|V_0|}{F}\right) \leq \exp\left(-\frac{\sqrt{\xi}\beta_1 m_p}{6F}\right) = \exp\left(-\frac{2\sqrt{\xi}\beta_1 m}{3sr_p}\right) \overset{(2.5.1),(2.5.15)}{\leq} e^{-\sqrt{n}}. \tag{2.5.21}$$

Note that $n \leq 2L_pm_p$ by (2.5.11) and

$$r_p \overset{(2.5.14)}{\geq} sp\tilde{r}/2 \overset{(2.5.4)}{\geq} sp\tilde{\alpha}\tilde{L}/3 \overset{(2.5.9)}{=} \tilde{\alpha}L_p/3.$$

Thus

$$\frac{2|V_0|}{F} \overset{(2.5.2)}{\leq} 4\sqrt{\tilde{\varepsilon}}n\frac{p}{r_p} \leq \frac{24\sqrt{\tilde{\varepsilon}}pm_p}{\tilde{\alpha}} \overset{(2.5.1)}{<} \frac{\sqrt{\xi}\beta_1 m_p}{2}. \tag{2.5.22}$$

Furthermore, $|X|F \leq n^2$, so (2.5.21) and (2.5.22) imply that the probability that (ii) fails for this partition is at most $n^2e^{-\sqrt{n}}$. Therefore the partition satisfies both (i) and (ii) with probability $1 - 2n^2e^{-\sqrt{n}} \geq 1/2$. This proves the claim. ■

For each k with $1 \leq k \leq 2p$, recall from the end of Section 2.5.4 that the number of all graphs G_i whose original type t satisfies (a) is $2F$. For each k, consider the ordering of all these digraphs as chosen in Section 2.5.4 and suppose that G_i is the fth digraph with original type t. We now define the edges of G_i between V_0 and $V(G) \setminus V_{0,i}$. For each base s-cluster W_ℓ, let $W_\ell(k)$ denote the kth base $2p$-cluster contained in W_ℓ. Apply Claim 2.5.3 to obtain F bipartite digraphs $E^+_f(W_\ell(k))$ and F bipartite digraphs $E^-_f(W_\ell(k))$ for each $W_\ell(k)$. Now let W'_ℓ denote the s-$[i]$-cluster associated with W_ℓ and $W_\ell(k)'$ denote the

$2p$-$[i]$-cluster associated with $W_\ell(k)$. Let $E_f^+(W_\ell(k)')$ be the subdigraph of $E_f^+(W_\ell(k))$ consisting of all edges whose final vertex lies in $W_\ell(k)'$ and let $E_f^-(W_\ell(k))')$ be the subdigraph of $E_f^-(W_\ell(k))$ consisting of all edges whose initial vertex lies in $W_\ell(k)'$. Then, by (2.5.19), for all $x \in V(G)$ we have

$$d^+_{E_f^+(W_\ell(k)')}(x) \geq d^+_{E_f^+(W_\ell(k))}(x) - \varepsilon' m_p/2 - 1. \tag{2.5.23}$$

An analogous statement is true for the indegrees in E_f^-. Recall that k with $1 \leq k \leq 2p$ is defined by the fact that G_i has original type t and $t \equiv k \mod 2p$, and that I_W^+ and I_W^- are the indices of the non-clean s-clusters in W defined at the end of Section 2.5.4. If $1 \leq f \leq F$, we add the following edges to G_i:

- all edges lying in the digraphs $E_f^+(W_\ell(k)')$ with $\ell \in I_W^+$;
- all edges lying in the digraphs $E_f^-(W_\ell(k)')$ with $\ell \in I_W^-$.

If $F < f \leq 2F$, we add the following edges to G_i:

- all edges lying in the digraphs $E_{f-F}^+(W_\ell(k)')$ with $\ell \in I_W^-$;
- all edges lying in the digraphs $E_{f-F}^-(W_\ell(k)')$ with $\ell \in I_W^+$.

Note this implies that all edges from $G_i \setminus V_0$ to V_0 have initial vertex in an out-red cluster and similarly for the in-red clusters. Moreover, the sets of edges assigned to G_i and $G_{i'}$ are disjoint for $i \neq i'$. Indeed, this follows from the fact that, for $j \neq k$, $E_f^\pm(W_\ell(j))$ and $E_f^\pm(W_\ell(k))$ are clearly edge-disjoint; that for $f \neq f'$, $E_f^\pm(W_\ell(k)')$ and $E_{f'}^\pm(W_\ell(k)')$ are also edge-disjoint; and that each $E_f^\pm(W_\ell(k))$ is used for at most one of the G_i.

Therefore Claim 2.5.3, (2.5.23) and (2.5.18) imply that for all $x \in V_0$, we have that

$$d^+_{G_i}(x) \geq (s/2 - 1) \sum_{W \in \tilde{\mathcal{P}}} \left(\frac{1-\varepsilon}{2spF} (|N_G^+(x) \cap W| - 5\tilde{\varepsilon}\tilde{m}) - \frac{\varepsilon' m_p}{2} - 1 \right).$$

Note also that

$$2spF = \frac{sr_p}{2} \overset{(2.5.15)}{\leq} \frac{s}{2} \frac{\tilde{\alpha} L_p}{\beta} \overset{(2.5.11)}{\leq} \frac{s}{2} \frac{\tilde{\alpha} n}{\beta m_p}. \tag{2.5.24}$$

So

$$\begin{aligned} d^+_{G_i}(x) &\geq (s/2-1)\frac{(1-\varepsilon')}{2spF}(\tilde{\alpha}n - |V_0| - 2\varepsilon' n) \\ &\overset{(2.5.2),(2.5.24)}{\geq} (1-4\varepsilon')\beta m_p \overset{(2.5.5)}{\geq} \beta_1 m_p \overset{(2.5.17)}{\geq} \kappa, \end{aligned} \tag{2.5.25}$$

and we have an analogue for indegrees. So we can delete edges from each $x \in V_0$ so that $d^{\pm}_{G_i}(x) = \kappa$ in each slice and hence (Red1) holds for all vertices in V_0.

2.5.7 Incorporating the special exceptional set $V^{\text{spec}}_{0,i}$

We now prove a claim which will be used to achieve (Red1) for the set $V^{\text{spec}}_{0,i}$ of special exceptional vertices. Before this, we first need to derive a further property (H1′) of $H^{\pm}_1$ from (H1).

Write S^+_i for the collection of vertices contained in the out-red $2p$-$[i]$-clusters and define S^-_i analogously. Note that each of $S^{\pm}_i$ consists of the vertices in exactly $s/2-1$ of the $2p$-$[i]$-clusters in each adapted s-$[i]$-cluster.

For every $1 \leq k \leq 2p$ and every base s-cluster $U \in \mathcal{P}'_s$, let $U(k)$ be the kth base $2p$-cluster of U, and write $H^+_{1,k}$ for the spanning subdigraph of H^+_1 consisting of all edges whose final vertex lies in $\bigcup_{U\in\mathcal{P}'_s} U(k)$. Also define $H^-_{1,k}$ to be the spanning subdigraph of H^-_1 consisting of all edges whose initial vertex lies in $\bigcup_{U\in\mathcal{P}'_s} U(k)$. We have the following property of $H^{\pm}_1$:

(H1′) For all $x \in V(G)\setminus V_0$, whenever i has original type t and k satisfies $1 \leq k \leq 2p$ and

(a) we have that

$$\frac{\gamma\tilde{\alpha}n}{20p} \leq |N^+_{H^+_{1,k}}(x) \cap S^-_i| \quad , \quad |N^-_{H^-_{1,k}}(x) \cap S^+_i| \leq \frac{\gamma\tilde{\alpha}n}{p}.$$

To prove (H1′), note that since $\mathcal{P}'_{2p}$ was a $5\tilde{\varepsilon}$-uniform $2sp$-refinement of $\tilde{\mathcal{P}}$, (URef) implies

that, for each $x \in V(G) \setminus V_0$, each $1 \leq k \leq 2p$ and $U \in \mathcal{P}'_s$,

$$|N^+_{H^+_{1,k}}(x) \cap U(k)| \geq \frac{(1-5\tilde{\varepsilon})}{2sp}\left(|N^+_{H^+_1}(x) \cap \tilde{U}| - 5\tilde{\varepsilon}\tilde{m}\right)$$

where $\tilde{U}$ is the base primary cluster containing U. If $U(k)_i$ is the $2p$-$[i]$-cluster associated with $U(k)$, (2.5.19) implies that

$$|N^+_{H^+_{1,k}}(x) \cap U(k)_i| \geq |N^+_{H^+_{1,k}}(x) \cap U(k)| - \varepsilon' m_p/2 - 1.$$

But whenever i has original type t and k satisfies (a), S_i^- contains all the vertices from exactly $s/2-1$ of the $2p$-$[i]$-clusters $U(k)_i$ contained in each adapted $[i]$-cluster $\tilde{U}_i$ associated with $\tilde{U}$, so

$$\begin{aligned}|N^+_{H^+_{1,k}}(x) \cap S_i^- \cap \tilde{U}_i| &\geq (s/2-1)\left(\frac{1-5\tilde{\varepsilon}}{2sp}(|N^+_{H^+_1}(x) \cap \tilde{U}| - 5\tilde{\varepsilon}\tilde{m}) - \frac{\varepsilon' m_p}{2} - 1\right)\\ &\geq |N^+_{H^+_1}(x) \cap \tilde{U}|/6p - \varepsilon' s m_p.\end{aligned}$$

Therefore, summing over all base primary clusters $\tilde{U}$ and recalling that V_0 is an isolated set in H_1^+ we have that

$$|N^+_{H^+_{1,k}}(x) \cap S_i^-| \geq \frac{d^+_{H^+_1}(x)}{6p} - \varepsilon' s m_p \tilde{L} \overset{(\mathrm{H1})}{\geq} \frac{\gamma\tilde{\alpha}n}{20p}.$$

The other bounds in (H1′) follow similarly.

Claim 2.5.4. *For each i with $1 \leq i \leq r_p$, there are subdigraphs Q_i^+ of H_1^+ and Q_i^- of H_1^- each consisting of edges between $V^{\mathrm{spec}}_{0,i}$ and $V(G) \setminus V_{0,i}$ so that*

(i) for all $x \in V^{\mathrm{spec}}_{0,i}$ we have $|N^+_{Q_i^+}(x) \cap S_i^-|, |N^-_{Q_i^-}(x) \cap S_i^+| \geq \kappa$.

(ii) For all $y \in V(G) \setminus V_{0,i}$ we have $d^+_{Q_i^-}(y), d^-_{Q_i^+}(y) \leq \sqrt{\xi}\beta_1 m_p/3$.

(iii) all the $Q_i^\pm$ are pairwise edge-disjoint.

Proof. For each vertex x in $V(G) \setminus V_0$, we let $T(x) := \{i : x \in V_{0,i}^{\text{spec}}\}$. Recall that $x \in V_{0,i}^{\text{spec}}$ if and only if

(A) $x \in \tilde{V}_{0,t}^{\text{spec}}$ and i has original type t; or

(B) x was removed to compensate for the removal of a bridge vertex.

Note that x can satisfy both (A) and (B). Suppose that x satisfies (A). Let $\mathcal{L}_x = \{t : x \in \tilde{V}_{0,t}^{\text{spec}}\}$. Note $x \notin V_0$. So x is good in the sense of Section 2.5.1, and hence $|\mathcal{L}_x| \leq \xi\tilde{L}/\beta$. As observed before (2.5.20), any $x \in V(G) \setminus V_0$ is in at most one set $V_{0,i}^{\text{spec}}$ due to (B). Therefore

$$|T(x)| \leq |\mathcal{L}_x|(s-1)(p-1) + 1 \leq \xi\tilde{L}sp/\beta \overset{(2.5.9)}{=} \xi L_p/\beta.$$

For each $1 \leq i \leq r_p$ and each $1 \leq k \leq 2p$ we define digraphs $Q_{i,k}^+$ as follows. For each k, we randomly assign each edge of $H_{1,k}^+$ whose initial vertex is x to one of the digraphs $Q_{i,k}^+$ with $i \in T(x)$ with probability $q := \beta/\xi L_p$ (independently of all other edges, and each edge is assigned to at most one of the $Q_{i,k}^+$). The sum of the probabilities is at most 1. Note that V_0 is an isolated set in $H_{1,k}^{\pm}$. Now define $Q_i^+ := Q_{i,k}^+$ where i has original type t and k satisfies (a). Then (iii) certainly holds, and for all $x \in V_{0,i}^{\text{spec}}$, we have

$$\mathbb{E}\left(|N_{Q_i^+}^+(x) \cap S_i^-|\right) = \frac{\beta|N_{H_{1,k}^+}^+(x) \cap S_i^-|}{\xi L_p} \overset{(\text{H1}')}{\geq} \frac{\gamma\tilde{\alpha}\beta n}{20p\xi L_p} \overset{(2.5.11)}{\geq} 2\beta m_p. \tag{2.5.26}$$

Proposition 2.3.1 implies that, for fixed $1 \leq i \leq r_p$ and fixed $x \in V_{0,i}^{\text{spec}}$,

$$\mathbb{P}\left(|N_{Q_i^+}^+(x) \cap S_i^-| < \beta_1 m_p\right) \leq \exp\left(-\frac{\beta m_p}{6}\right) \overset{(2.5.9),(2.5.11)}{\leq} \exp\left(-\frac{\beta n}{12sp\tilde{L}}\right) \leq e^{-\sqrt{n}}.$$

So a union bound implies that the probability that there exist i and x not satisfying this inequality is at most $n^2 e^{-\sqrt{n}} < 1/4$. (i) now follows since $\kappa \leq \beta_1 m_p$ by (2.5.17).

For (ii), note that for *any* vertex $y \in V(G)$ we have

$$\mathbb{E}\left(d^-_{Q_i^+}(y)\right) \leq q|V^{\text{spec}}_{0,i}| \overset{(2.5.20)}{\leq} \frac{\beta}{\xi L_p}\varepsilon n \overset{(2.5.11)}{\leq} \frac{2\varepsilon}{\xi}\beta m_p \leq \sqrt{\xi}\beta m_p/4. \tag{2.5.27}$$

Proposition 2.3.1 shows (as in Claim 2.5.3) that the probability that the condition in (ii) fails for some i and some $y \in V(G)$ is at most 1/4. So there is a choice of $Q_1^+, \ldots, Q_{r_p}^+$ so that all the conditions hold, and similarly for $Q_1^-, \ldots, Q_{r_p}^-$, as required. ■

It is now easy to obtain the edges of G_i between $V^{\text{spec}}_{0,i}$ and $V(G) \setminus V_{0,i}$. Apply Claim 2.5.4 to find edge-disjoint digraphs $Q_i^\pm$ for each $1 \leq i \leq r_p$. Recall that $S_i^\pm \subseteq V(G) \setminus V_{0,i}$ and so (Red6) will follow if we add i-red edges with initial vertex in S_i^+ or final vertex in S_i^-. So for each $x \in V_{0,i}$ we add exactly κ edges in Q_i^+ going from x to S_i^- and exactly κ edges in Q_i^- going to x from S_i^+.

We have now incorporated $V_{0,i}$ into each G_i. It remains to verify that (Red0)–(Red7) hold. Recall that we partially verified these properties for the red vertices incident to bridge vertices at the end of Section 2.5.5. In particular, (Red0) was achieved in Section 2.5.5 and the edges we have added here do not affect it. The previous paragraph shows that (Red1) holds for all vertices in $V^{\text{spec}}_{0,i}$. Since we already verified it for the bridge vertices $V^{\text{bridge}}_{0,i}$ in Section 2.5.5 and for V_0 in Section 2.5.6, it now holds for all vertices in $V_{0,i}$. Clearly, our construction satisfies (Red2). (Red3) follows from Claims 2.5.3(ii) and 2.5.4(ii) and the fact that each non-exceptional vertex is incident to at most one bridge vertex in each slice. Recall that, in Section 2.5.4, we showed how (Red4) and (Red5) follow from (a) and (b) of the construction. (Red6) follows from the fact that in constructions including V_0 and $V^{\text{spec}}_{0,i}$ and $V^{\text{bridge}}_{0,i}$, the outedges from $V_{0,i}$ always went to in-red clusters and the inedges to $V_{0,i}$ came from out-red clusters. (Red7) follows immediately from the edge-disjointness of the digraphs in Claims 2.5.3 and 2.5.4 and the observation in the final paragraph of Section 2.5.5.

Note that Theorem 2.3.2 implies that the proofs of Claims 2.5.3 and 2.5.4 can be

'derandomised' and so red edges satisfying (Red0)–(Red7) can be found in polynomial time.

2.5.8 Finding shadow balancing sequences

We have now incorporated all the exceptional vertices to form r_p edge-disjoint slices G_i of G, together containing almost all edges, such that each slice is a spanning almost-regular subdigraph of G. The main aim of this section is to add further red edges to each slice G_i so that the number of red edges sent out by vertices in each cluster V equals the number received by its successor V^+ on the cycle of F_i containing V.

This 'balancing property' is necessary for the following reason. Suppose that V is out-red and suppose that we have a 1-factor f containing a red edge sent out to $V_{0,i}$ by a vertex $x \in V$. If V^+ is not red, any edge of f to V^+ must have its initial vertex in V. So $f[V, V^+]$ must be a perfect matching, which is impossible since there can be no edge in f from x to a vertex in V^+. Note that the absence of red edges incident to V^- does not give rise to the above problem. But we observe a similar problem for U, U^- when U is in-red. So the above 'balancing property' is certainly necessary to obtain even a single 1-factor. We will see in Section 2.5.10 that, combined with our other properties, it is also sufficient.

We will add 'balancing edges' between non-exceptional vertices to achieve the above property while also ensuring that no vertex is incident to many red edges. As indicated above, it will turn out to be sufficient to only add such edges to either the predecessor or successor of existing red clusters. By the end of Section 2.5.9 our new red clusters will consist of consecutive pairs, well-spaced around each blown-up cycle.

We will first find 'shadow balancing edges' in the reduced digraph between suitable cluster pairs. For this, we will use the fact that R_p is a robust outexpander. Then we will choose the required number of edges from the sparse pre-reserved subdigraph H_2 induced on these pairs. When doing this, we need to be careful to maintain (Red6) with p replaced by $p - 1$.

Given G_i, we denote the set of red p-clusters by T (so we suppress the dependence on i here). Let T_{in} denote the set of in-red clusters and define T_{out} similarly, so $T = T_{\text{in}} \cup T_{\text{out}}$. For a set $S \subseteq T$ of p-clusters, we let S^- denote the predecessors of S on T and define S^+ similarly.

Now, for each $1 \leq i \leq r_p$ and each p-cluster V, let

$$s_i^{\pm}(V) := \sum_{y \in V} |N_{G_i}^{\pm}(y) \cap V_{0,i}| \tag{2.5.28}$$

be the number of red edges entering/leaving V. So $s_i^+(V) \neq 0$ only if $V \in T_{\text{out}}$ and $s_i^-(V) \neq 0$ only if $V \in T_{\text{in}}$. Note that (Red1) implies that

$$\sum_{V \in R_p} s_i^+(V) = \sum_{V \in R_p} s_i^-(V). \tag{2.5.29}$$

Let

$$b := \frac{\xi^{1/6}\beta_1 m_p^2}{L_p} \quad \text{and} \quad c := \xi^{1/5}\beta_1 m_p^2. \tag{2.5.30}$$

A *balancing sequence* B_i with respect to G_i is a spanning subdigraph of H_2 with the following properties:

(B1) $d_{B_i}^{\pm}(y) \leq 8\xi^{1/6}\beta_1 m_p$ for every $y \notin V_{0,i}$;

(B2) We have the following degree conditions:

$$d_{B_i}^+(V) = \begin{cases} s_i^-(V^+) + c & \text{if } V \in T_{\text{in}}^- \\ c & \text{if } V \in T_{\text{out}} \\ 0 & \text{otherwise,} \end{cases}$$

$$d_{B_i}^-(V) = \begin{cases} c & \text{if } V \in T_{\text{in}} \\ s_i^+(V^-) + c & \text{if } V \in T_{\text{out}}^+ \\ 0 & \text{otherwise.} \end{cases}$$

We will use so called 'shadow balancing sequences' as a framework to find balancing

sequences. For this, define an auxiliary digraph R^* with $V(R^*) = T$ as follows. Let

$$N^+_{R^*}(V) = \begin{cases} \left(N^+_{R_p}(V^-) \cap T_{\text{in}}\right) \cup \left(N^+_{R_p}(V^-) \cap T^+_{\text{out}}\right)^- & \text{if } V \in T_{\text{in}} \\ \left(N^+_{R_p}(V) \cap T_{\text{in}}\right) \cup \left(N^+_{R_p}(V) \cap T^+_{\text{out}}\right)^- & \text{if } V \in T_{\text{out}}. \end{cases} \tag{2.5.31}$$

This definition reflects the fact that red edges entering $V \in T_{\text{in}}$ will be balanced by edges leaving V^- (and entering either T_{in} or the successor W^+ of some $W \in T_{\text{out}}$). Similarly an edge leaving $V \in T_{\text{out}}$ will be balanced by an edge entering V^+. Note that R^* depends on i. If we need to emphasise this, we write R^*_i.

Define a *shadow balancing sequence* B'_i to be a multidigraph with vertex set $V(R^*)$ whose edges are copies of edges of R^* as follows. Let

$$n^+_V := \begin{cases} s^-_i(V) + c & \text{if } V \in T_{\text{in}} \\ c & \text{if } V \in T_{\text{out}} \end{cases} \quad \text{and} \quad n^-_V := \begin{cases} c & \text{if } V \in T_{\text{in}} \\ s^+_i(V) + c & \text{if } V \in T_{\text{out}}. \end{cases}$$

Then B'_i has the following properties:

(B1′) no edge of R^* appears more than b times in B'_i.

(B2′) For every $V \in V(R^*)$, we have $d^+_{B'_i}(V) = n^+_V$ and $d^-_{B'_i}(V) = n^-_V$.

Note that (2.5.29) implies that

$$\sum_{V \in R^*} n^+_V = \sum_{V \in R^*} n^-_V. \tag{2.5.32}$$

To find these shadow balancing sequences, we will need that R^* is a robust outexpander with sufficiently large minimum semidegree.

Claim 2.5.5. *Let $\nu' = \nu^3/64$. Then*

(i) R^* is a robust $(\nu', 12\tau)$-outexpander.

(ii) $\delta^0(R^*) \geq \tilde{\alpha}|R^*|/4$.

Proof. To prove part (i) of the claim, we will use the fact that an $(s/2-1)$-fold blow-up of a robust $(\nu/4, 3\tau)$-outexpander is a $(\nu', 6\tau)$-robust outexpander (see Lemma 2.3.12). Let $R_p^{\text{in}} = R_p[T_{\text{in}}]$ and $R_p^{\text{out}} = R_p[T_{\text{out}}^+]$. Since every adapted primary cluster contains exactly $s/2-1$ out-red p-clusters, it follows that R_p^{in} is an $(s/2-1)$-fold blow-up of $\tilde{R}$. So it is a robust $(\nu', 6\tau)$-outexpander. Similarly, R_p^{out} is a robust $(\nu', 6\tau)$-outexpander.

Consider any $S \subseteq T_{\text{in}}$ with $6\tau|T_{\text{in}}| \leq |S| \leq (1-6\tau)|T_{\text{in}}|$. Note that T_{in} and T_{out} are disjoint (see e.g. (Red6)). So T_{in} and $(T_{\text{out}}^+)^-$ are disjoint and hence (2.5.31) implies that

$$|RN^+_{\nu',R^*}(S)| = |RN^+_{\nu',R_p}(S^-) \cap T_{\text{in}}| + |(RN^+_{\nu',R_p}(S^-) \cap T^+_{\text{out}})^-|. \tag{2.5.33}$$

Now let S^-_{in} be obtained from S^- by replacing each p-cluster $V \in S^-$ by an arbitrary (but distinct) p-cluster $V_{\text{in}} \in T_{\text{in}}$ which lies in the same adapted primary cluster as S^-. Note this is possible as $S \subseteq T_{\text{in}}$ implies that S (and thus S^-) contains at most $s/2-1$ of the p-clusters from each adapted s-cluster. Note that in R_p, each cluster receives an edge from V_{in} if and only if it receives an edge from V. So (2.5.31) implies that

$$\begin{aligned}|RN^+_{\nu',R_p}(S^-) \cap T_{\text{in}}| &= |RN^+_{\nu',R_p}(S^-_{\text{in}}) \cap T_{\text{in}}| = |RN^+_{\nu',R_p^{\text{in}}}(S^-_{\text{in}})| \\ &\geq |S^-_{\text{in}}| + \nu'|R_p^{\text{in}}| = |S| + \nu'|R^*|/2.\end{aligned}$$

Similarly, let S^-_{out} be obtained from S^- by replacing each p-cluster $V \in S^-$ by an arbitrary (but distinct) cluster $V_{\text{out}} \in T^+_{\text{out}}$ which lies in the same adapted s-cluster

as V. Then we have

$$\begin{aligned}|(RN^+_{\nu',R_p}(S^-)\cap T^+_{\text{out}})^-| &= |RN^+_{\nu',R_p}(S^-)\cap T^+_{\text{out}}|\\ &= |RN^+_{\nu',R_p}(S^-_{\text{out}})\cap T^+_{\text{out}}| = |RN^+_{\nu',R_p^{\text{out}}}(S^-_{\text{out}})|\\ &\geq |S^-_{\text{out}}| + \nu'|R_p^{\text{out}}| = |S| + \nu'|R^*|/2.\end{aligned}$$

So altogether, we have $|RN^+_{\nu',R^*}(S)| \geq 2|S| + \nu'|R^*|$.

Now suppose that $S \subseteq T_{\text{out}}$ with $6\tau|T_{\text{out}}| \leq |S| \leq (1-6\tau)|T_{\text{out}}|$. Similarly as above, (2.5.31) implies that

$$\begin{aligned}|RN^+_{\nu',R^*}(S)| &= |RN^+_{\nu',R_p}(S)\cap T_{\text{in}}| + |(RN^+_{\nu',R_p}(S)\cap T^+_{\text{out}})^-|\\ &\geq 2|S| + \nu'|R^*|.\end{aligned} \tag{2.5.34}$$

Now consider any $S \subseteq V(R^*)$ with $6\tau|R^*| \leq |S| \leq (1-6\tau)|R^*|$. Then either $|S\cap T_{\text{in}}| \geq |S|/2$ or $|S\cap T_{\text{out}}| \geq |S|/2$. In either case, we get $|RN^+_{\nu',R^*}(S)| \geq |S| + \nu'|R^*|$. This proves part (i) of the claim.

To prove part (ii), suppose that $V \in T_{\text{in}}$. Note that R_p^{in} satisfies $\delta^0(R_p^{\text{in}}) \geq \tilde{\alpha}|R_p^{\text{in}}|/2$ by Lemma 2.3.12(i). Choose any $V^-_{\text{in}} \in T_{\text{in}}$ which lies in the same adapted primary cluster as V^-. Then, similarly as observed above, V^-_{in} has the same outneighbours within the set T_{in} as V^- (both in the digraph R_p). So the degree bound follows for V. The case when $V \in T_{\text{out}}$ is similar. ■

It is now easy to find shadow balancing sequences B'_i satisfying (B1′) and (B2′). Indeed, note that $c \leq n_V^{\pm} \leq c + \sqrt{\xi}\beta_1 m_p^2$ by (Red3). In particular, (2.5.30) implies that $n_V^+ = c\left(1 \pm \xi^{3/10}\right)$ and similarly for n_V^-. Let R' be obtained from R^* by replacing each of the edges of R^* by b copies of this edge and let $n' := |R^*| = (s-2)\tilde{L}$. We will apply Lemma 2.4.1 as follows:

	R^*	R'	n'	b	$\xi^{3/10}$	ν'	c/n'
playing the role of	G	Q	n	q	ε	ν	ρ

Then

$$\rho := \frac{c}{n'} \overset{(2.5.30)}{=} \frac{\xi^{1/5}\beta_1 m_p^2}{(s-2)\tilde{L}} \overset{(2.5.1)}{\leq} \frac{\xi^{1/6}\beta_1 m_p^2 \nu'^2}{3sp\tilde{L}} \overset{(2.5.9),(2.5.30)}{=} \frac{b\nu'^2}{3}$$

as required by Lemma 2.4.1, and we obtain a spanning subdigraph B_i' of R' with $d^{\pm}_{B_i'}(V) = n_V^{\pm}$ for each $V \in V(R') = V(R^*)$.

2.5.9 Adding balancing sequences

Note that for each edge E' of R_i^*, there is a unique edge E of R_p (from a p-cluster A to a p-cluster B) which corresponds to E'. More precisely, (2.5.31) shows that if $E' = VW \in E(R_i^*)$ then

$$E = \begin{cases} V^-W & \text{if } V \in T_{\text{in}}, W \in T_{\text{in}} \\ V^-W^+ & \text{if } V \in T_{\text{in}}, W \in T_{\text{out}} \\ VW & \text{if } V \in T_{\text{out}}, W \in T_{\text{in}} \\ VW^+ & \text{if } V \in T_{\text{out}}, W \in T_{\text{out}}. \end{cases} \tag{2.5.35}$$

(As before, V^- denotes the predecessor of V on F_i.) So for each edge of B_i', we can choose the corresponding edge of R_p. For each i and each edge E of R_p, let $c_i(E)$ denote the number of times that the edge E is chosen due to B_i'. So $c_i(E) \leq b$ by (B1′). If we now replace the chosen edges E of R_p with $c_i(E)$ edges in $H_2(E)$, this will give the required balancing sequence B_i. However, we need to be careful to ensure that we can do this for every i with $1 \leq i \leq r_p$ so that all edges are disjoint. We also wish to maintain (Red4) and (Red6).

We now need to consider the dependence on i again, as clusters in different slices are not quite the same. Given a base p-cluster A in R_p, let $A[i]$ be the associated p-$[i]$-cluster. Each p-$[i]$-cluster $A[i]$ contains at most one red $2p$-$[i]$-cluster by (Red4). If there is such a subcluster, denote it by $A^*[i]$. If there is no such subcluster, let $A^*[i]$ be an arbitrary subcluster of $A[i]$. We will only add balancing edges incident to $A^*[i]$. Let A^* be the base $2p$-cluster associated with $A^*[i]$. Suppose that E is an edge of R_p from A to B.

Let $\tilde{E} \in E(\tilde{R}(\beta))$ be one of the edges whose blow-up contains E; then $H_2(\tilde{E})$ is $(\varepsilon, \gamma\beta)$-regular as observed in Section 2.5.2. Write $H_2(E^*)$ for the subdigraph of $H_2(\tilde{E})$ induced on (A^*, B^*); then by Lemma 2.3.13(i) we have that $H_2(E^*)$ is $(\varepsilon', \gamma\beta)$-regular.

Write $H_2(E^*[i])$ for the subdigraph of $H_2(E^*)$ induced on $(A^*[i], B^*[i])$. Whenever E is chosen due to B'_i, we will add balancing edges to G_i from $H_2(E^*[i])$. By (2.5.19) we have that, for all i with $1 \leq i \leq r_p$, $H_2(E^*[i])$ is a subdigraph of $H_2(E^*)$ obtained by removing at most $\varepsilon' m_p/2 + 1$ vertices from each vertex class.

Now for each i in succession we aim to apply Lemma 2.3.11 to find a set $C_i(E)$ of $c_i(E)$ edges in $H_2(E^*)$, and remove the edges of $C_i(E)$ from further consideration. Suppose we have found $C_1(E), \ldots, C_{i-1}(E)$ in $H_2(E^*)$. Suppose further that each of these has maximum degree at most $d_0 m_p$ and that the edges are from A^* to B^*. We now wish to find $C_i(E)$.

Let $H_2^{i-1}(E^*)$ denote the subdigraph of $H_2(E^*)$ obtained by removing the edges of $C_1(E), \ldots, C_{i-1}(E)$ and removing any vertex not present in $H_2(E^*[i])$. So $H_2^{i-1}(E^*)$ is also a subdigraph of $H_2(E^*[i])$. By (2.5.19), the number of vertices in each vertex class of $H_2^{i-1}(E^*)$ is at most $\varepsilon' m_p/2 + 1$ less than that in $H_2(E^*)$. Moreover, at most $r_p d_0 m_p$ edges have been removed from each vertex.

We need the following short claim.

Claim 2.5.6. *Let $d_0 := 8b/m_p^2$ where b is defined in* (2.5.30)*. Suppose that H is a subdigraph of $H_2(E^*)$ obtained by removing at most $\varepsilon' m_p/2 + 1$ vertices from each of A^* and B^* and at most $r_p d_0 m_p$ edges at every vertex. Then H is $(\xi^{1/15}, \gamma\beta)$-regular.*

Proof. Note first that

$$d_0 = \frac{8b}{m_p^2} = \frac{8\xi^{1/6}\beta_1}{L_p}. \tag{2.5.36}$$

So

$$2r_p d_0 \overset{(2.5.15)}{\leq} \frac{16\xi^{1/6}\beta_1}{L_p}\frac{\tilde{\alpha}L_p}{\beta} \leq 16\xi^{1/6}\tilde{\alpha} \leq \xi^{1/7}.$$

Also $\varepsilon' \ll \xi^{1/7}$. So Proposition 2.3.6(i) with $\xi^{1/7}$ playing the role of d' implies the claim. ■

Claim 2.5.6 implies that $H_2^{i-1}(E^*)$ is $(\xi^{1/15}, \gamma\beta)$-regular. So we can apply Lemma 2.3.11 to find $C_i(E)$, with a maximum degree of at most $8c_i(E)/m_p \leq 8b/m_p = d_0 m_p$. We continue inductively until we have found $C_1(E), \ldots, C_{r_p}(E)$.

Now let B_i be the union of all $C_i(E)$ over all edges E of R_p. Note that the B_i are edge-disjoint by construction. To verify (B1), note that for all $y \in V(G) \setminus V_{0,i}$,

$$d^{\pm}_{B_i}(y) \leq L_p d_0 m_p \overset{(2.5.36)}{=} 8\xi^{1/6}\beta_1 m_p,$$

as required. (2.5.35) implies that the clusters that send out shadow balancing edges are precisely $T^-_{\text{in}} \cup T_{\text{out}}$ and the clusters that receive shadow balancing edges are precisely $T_{\text{in}} \cup T^+_{\text{out}}$. Suppose that $V \in T^-_{\text{in}}$. Then we have that

$$d^+_{B_i}(V) \overset{(2.5.35)}{=} d^+_{B'_i}(V^+) \overset{(\text{B2}')}{=} n^+_{V^+} = s^-_i(V^+) + c$$

so (B2) holds in this case. The other cases follow similarly. Therefore B_i satisfies (B1) and (B2). Note that only vertices in a single $2p$-subcluster of each p-cluster (which is the red subcluster if one of them is red) are incident to a balancing edge.

For each $1 \leq i \leq r_p$ we add the edges of B_i to G_i. So now $E(G_i)$ consists of edges from each cluster to its unique successor on F_i together with the i-red edges incident to $V_{0,i}$ and the balancing edges B_i.

2.5.10 Almost decomposing into 1-factors

Our aim now is to use Lemma 2.4.2 to find a κ-regular spanning subdigraph of each G_i. For this, the 'balancing property' achieved in Section 2.5.9 will be crucial.

Before this, for each i, we first remove a subdigraph $H_{3,i}$ of G_i, which will be needed in Section 2.5.11. We do this as follows. For each edge E of F_i, recall that $G_i(E)$ is $(2\varepsilon', \beta_1)$-

superregular by (Red7). Apply Lemma 2.3.9(ii) to $G_i(E)$ with parameters $K := 2$ and $\gamma_1 := \gamma^2\beta_1, \gamma_2 := \beta_2$ where

$$\beta_2 := (1-\gamma^2)\beta_1 \tag{2.5.37}$$

to obtain two edge-disjoint subdigraphs of $G_i(E)$: a $(2\varepsilon'^{1/12}, \gamma^2\beta_1)$-superregular digraph $H_{3,i}(E)$ and a $(2\varepsilon'^{1/12}, \beta_2)$-superregular 'remainder' subdigraph which we still denote by $G_i(E)$. We let $H_{3,i}$ have vertex set $V(G)$ and edge set given by the union of $H_{3,i}(E)$ over all edges E of F_i.

We now continue with finding a κ-regular spanning subdigraph of each G_i. Denote the collection of i-red edges incident to $V_{0,i}$ by $\mathcal{T}_i$. For each $1 \le i \le r_p$ we call the edges in $\mathcal{T}_i \cup B_i$ and any p-cluster containing a vertex incident to such an edge *i-red* or *red* (so balancing edges are also regarded as red now). Write $d_i^{\pm}(x) := d_{\mathcal{T}_i}^{\pm}(x) + d_{B_i}^{\pm}(x)$ for each $x \in V(G_i)$ and define $d_i^{\pm}(V) = \sum_{x\in V} d_i^{\pm}(x)$ for $V \in V(F_i)$. So by (2.5.28) we have that, for each $V \in V(F_i)$,

$$d_i^{\pm}(V) = s_i^{\pm}(V) + d_{B_i}^{\pm}(V). \tag{2.5.38}$$

For each $1 \le i \le r_p$ we now have the following properties:

(Red0′) There exists a sequence $D_1x_1D_2x_2\dots x_{\ell-1}D_\ell x_\ell D_1$ with the following properties:

- Each D_j is a cycle of F_i and every cycle of F_i appears at least once in the sequence;
- $V_{0,i}^{\text{bridge}} = \{x_1, \dots, x_\ell\}$ and each x_j has exactly κ outneighbours in D_{j+1} and exactly κ inneighbours in D_j;

(Red1′) $d_i^{\pm}(x) = \kappa$ for each $x \in V_{0,i}$;

(Red2′) $V_{0,i}$ is an independent set in G_i;

(Red3′) $d_i^{\pm}(y) \le \xi^{1/7}\beta_2 m_p$ for each $y \in G_i \setminus V_{0,i}$;

(Red4′) For every red cluster $V \in R_p$, all i-red edges are incident to a single $2p$-cluster

contained in V. In particular, at most $m_p/2$ vertices in V are incident to an i-red edge;

(Red5′) In F_i any out-red p-cluster V is preceded by $p-3$ p-clusters which are neither out-red nor in-red, and is succeeded by an in-red p-cluster. Any in-red p-cluster V is succeeded by $p-3$ p-clusters which are neither out-red nor in-red, and is preceded by an out-red p-cluster;

(Red6′) Each p-cluster is either out-red, in-red or contains no vertices incident to a red edge;

(Red7′) $G_1, \ldots, G_{r_p}$ are edge-disjoint and $G_i(E)$ is $(2\varepsilon'^{1/12}, \beta_2)$-superregular for all $E \in E(F_i)$;

(B2″) $d_i^+(V) = d_i^-(V^+)$ for all p-clusters $V \in V(F_i)$.

(Red0′), (Red1′) and (Red2′) follow immediately from (Red0), (Red1) and (Red2) respectively. (Red3′) follows from summing the degrees given by (Red3) and (B1) and using (2.5.37). (Red4′) is a consequence of (Red4) and our choice of edges in Section 2.5.9. (Red5′) follows from (Red5) and (B2): indeed, the (red) clusters in $T = T_{\text{in}} \cup T_{\text{out}}$ are separated by exactly $p-1$ non-red clusters by (Red5), and by (B2), the only other red clusters are precisely those in $T_{\text{in}}^- \cup T_{\text{out}}^+$. (Red6′) and edge-disjointness in (Red7′) follow from (Red6) and edge-disjointness in (Red7), as well as the construction of B_i in Sections 2.5.8 and 2.5.9. The second part of (Red7′) was verified directly after (2.5.37). (B2″) is a direct consequence of (B2) and (2.5.38). So for example, if $V \in T_{\text{out}}$ then

$$d_i^+(V) = s_i^+(V) + c = d_{B_i}^-(V^+) = d_i^-(V^+).$$

Consider any edge E from V to V^+ in F_i. We wish to find a subdigraph $G_i(E)^*$ of $G_i(E)$ such that, together with the red edges incident to V and V^+, every vertex in V has outdegree κ and every vertex in V^+ has indegree κ. The union of these subdigraphs over all edges $E \in E(F_i)$, together with the red edges $B_i \cup \mathcal{T}_i$, will form a κ-regular spanning subdigraph G_i^* of G_i. (Recall that κ was defined in (2.5.17).)

Given any $x \in V$, let $m_x^+ = d_i^+(x)$ and given any $y \in V^+$, let $m_y^- = d_i^-(y)$. By (Red3′) we have that $m_x^+, m_y^- \leq \xi^{1/7}\beta_2 m_p$ and by (B2″) we have that

$$\sum_{x \in V} m_x^+ = \sum_{y \in V^+} m_y^-.$$

Let $\hat{\varepsilon} := 2\varepsilon'^{1/12}$ and $\hat{\beta} := \beta_2 - \hat{\varepsilon}$. So (Red7′) implies that $G_i(E)$ is $(\hat{\varepsilon}, \hat{\beta} + \hat{\varepsilon})$-superregular for every $E \in E(F_i)$. Let

$$\hat{\alpha} := 1 - \frac{(1-\gamma)\beta_1}{\beta_2 - \hat{\varepsilon}}.$$

So $\kappa = (1-\hat{\alpha})\hat{\beta} m_p$, and it is easy to see that $\gamma/2 \leq \hat{\alpha} \leq 2\gamma$, so that $\hat{\beta} \ll \hat{\alpha} \ll 1$. Thus we can apply Lemma 2.4.2 to $G_i(E)$ with $\hat{\varepsilon}$ playing the role of ε, $\hat{\beta}$ playing the role of β and $\hat{\alpha}$ playing the role of α. Then we obtain a spanning subdigraph $G_i(E)^*$ of $G_i(E)$ in which each $x \in V$ has outdegree $\kappa - m_x^+$ and each $y \in V^+$ has indegree $\kappa - m_y^-$. Then

$$G_i^* := \bigcup_{E \in E(F_i)} G_i(E)^* \cup B_i \cup \mathcal{T}_i$$

is a κ-regular spanning subdigraph of G_i as required. Moreover $G_1^*, \ldots, G_{r_p}^*$ are edge-disjoint subdigraphs of G by (Red7′). Now apply Proposition 2.4.3 to each G_i^* to obtain κ edge-disjoint 1-factors $f_{i,1}, \ldots, f_{i,\kappa}$ of each G_i.

2.5.11 Merging 1-factors into Hamilton cycles

The final step is to use edges disjoint from our collection of 1-factors to merge cycles such that each 1-factor is transformed into a Hamilton cycle. Then we will have found an approximate decomposition into edge-disjoint Hamilton cycles. The argument will be exactly the same for each G_i. So since we will work within a fixed G_i, we will label the κ factors obtained from G_i as $f_1, \ldots, f_\kappa$. We wish to use Lemma 2.4.5 and edges from our pre-reserved digraph $H_{3,i}$ to merge the cycles in each f_j.

We will use the fact that each f_j has a structure closely related to that of F_i (which,

recall, is a 1-factor of $R_p(\beta)$). We say that a non-red cluster is *black* and we say that an edge of F_i is *black* if both the initial cluster and final cluster are black. So for all black edges VV^+ in F_i we have that $f_j[V, V^+]$ is a perfect matching for each f_j, since in G_i every edge from a vertex in V goes to a vertex in V^+. (Red5′) implies that every pair $U_{\text{out}}U_{\text{in}}$ of consecutive red clusters on any cycle of F_i is followed by $p-3$ consecutive black clusters. Denote the path of length $p-4$ from the first of these black clusters to the last by I_U, so every edge in I_U is black. So we can choose $p-4$ disjoint sets of edges $J_1, \ldots, J_{p-4}$ of F_i so that for each pair of consecutive red clusters $U_{\text{out}}U_{\text{in}}$, J_q contains exactly one edge of I_U. So each J_q consists of exactly $|T| = |T_{\text{in}}| + |T_{\text{out}}| = (s-2)\tilde{L}$ edges of F_i and has non-empty intersection with any cycle of F_i.

The idea is to apply Lemma 2.4.5 repeatedly to transform each of the f_j into a Hamilton cycle. Each time $H_{3,i}$ will play the role of G, and each J_q will play the role of J roughly κ/p times. If $\mathcal{E}$ is a set of edges in F_i, we write $H_{3,i}(\mathcal{E}) := \bigcup_{E\in\mathcal{E}} H_{3,i}(E)$.

We now describe the merging procedure for f_1. Denote the cycles of F_i by $D_1, \ldots, D_\ell$. Let K_1 be the 1-regular digraph consisting of all cycles of f_1 which contain a vertex in a cluster of D_1. Now apply Lemma 2.4.5 as follows: D_1 plays the role of C, $J_1 \cap E(D_1)$ plays the role of J, K_1 plays the role of F and $H_{3,i}(J_1)$ plays the role of G.

Condition (i) in Lemma 2.4.5 is clearly satisfied since every edge of J_1 is black. To verify condition (ii), let D be any cycle of K_1. We claim that D contains a vertex x from a black cluster B. To see this, suppose that D contains a vertex y which lies in an in-red cluster. Then the next vertex of D lies in a black cluster. Similarly, if y lies in an out-red cluster, then the vertex preceding y on D lies in a black cluster, which proves the claim. Now let I_U be the black interval containing B; then there is a path in D (containing x) which contains at least one vertex from each cluster in I_U. But $J_1 \cap E(D_1)$ contains an edge of I_U, as required.

To verify (iii), let VV^+ and WW^+ be edges of $J_1 \cap E(D_1)$ such that J_1 avoids all edges in the segment V^+D_1W. Then there is exactly one pair of successive red clusters $U_{\text{out}}U_{\text{in}}$ in this segment. So for each $v_a \in V^+$ there is a path P_a in f_1 from v_a to a distinct

vertex u_a^{out} in U_{out} which winds around D_1. Similarly, for each $u_{a'}^{\text{in}} \in U_{\text{in}}$ there is a path $P'_{a'}$ in f_1 from $u_{a'}$ to a distinct vertex $w_{a'} \in W$ which winds around D_1. But by (Red4′), for at least half of the vertices $u_a^{\text{out}} \in U_{\text{out}}$, there is an edge in f_1 to some $u_{a'}^{\text{in}} \in U_{\text{in}}$. So f_1 contains at least one path $v_a P_a u_a^{\text{out}} u_{a'}^{\text{in}} P'_{a'} w_{a'}$ from $v_a \in V^+$ to $w_{a'} \in W$ which winds around D_1, as required.

So we can find a matching M_1 in $H_{3,i}(J_1)$ and a cycle C_1 with $V(C_1) = V(K_1)$ and $E(C_1) \subseteq K_1 \cup M_1$. We replace the 1-regular subdigraph K_1 of f_1 by C_1. We call the resulting 1-factor $f_1(1)$ and we denote $H_{3,i} \setminus M_1$ by $H^2_{3,i}$. Note that all cycles of f_1 which contained a vertex in D_1 have now been merged into a single cycle of $f_1(1)$.

For $2 \leq k \leq \ell$ we define $f_1(k)$ inductively as follows. Let K_k be the 1-regular digraph consisting of all cycles of $f_1(k-1)$ which contain a vertex in a cluster of D_k. Now let D_k play the role of C, $J_1 \cap E(D_k)$ play the role of J, K_k play the role of F and $H_{3,i}(J_1)$ play the role of G. Note that the k choices $J_1 \cap E(D_{k'})$ with $1 \leq k' \leq k$ playing the role of J so far are pairwise vertex-disjoint. Exactly as above, the conditions (i)-(iii) are satisfied and we can apply Lemma 2.4.5 to obtain a 1-factor $f_1(k)$ in which all cycles containing a vertex in D_k have been merged. Moreover if two vertices x and y lie on a common cycle of $f_1(k-1)$ they lie on a common cycle of $f_1(k)$. We repeat this for all $1 \leq k \leq \ell$ to obtain $f'_1 := f_1(\ell)$. We will see below that f'_1 is a Hamilton cycle.

We now aim to carry out a similar procedure for $f_2, \ldots, f_\kappa$ to obtain $f'_2, \ldots, f'_\kappa$. The approach will be to use J_1 for $f_1, \ldots, f_{\kappa'}$ where $\kappa' := \kappa/(p-4)$ and more generally to use J_q for $f_{(q-1)\kappa'+1}, \ldots, f_{q\kappa'}$. Note that, to obtain f'_1, we removed exactly one perfect matching from each $H_{3,i}(E)$ for each edge E of J_1. To reuse J_1 we need only check that, at each step and for each edge E of J_1, the remainder of the sparse digraph $H_{3,i}(E)$ satisfies the conditions required of G in Lemma 2.4.5. For this, let $H^t_{3,i}(J_q)$ denote a subdigraph of $H_{3,i}(J_q)$ obtained by removing t arbitrary perfect matchings from $H_{3,i}(E)$ for each $E \in J_q$.

Claim 2.5.7. *Let κ' be defined as above and let $\varepsilon^* := 2\sqrt{\beta_1/p}$. Then $H^{\kappa'}_{3,i}(E)$ is*

$(\varepsilon^*, \gamma^2\beta_1)$*-superregular whenever* E *is an edge in* J_q*, where* $1 \le q \le p-4$.

Proof. To see this, it suffices to consider a single edge $E = XY$ in J_1. Write $H := H^{\kappa'}_{3,i}(E)$. Then, since at each stage we removed a perfect matching, in total we removed κ' edges incident to each vertex in $X \cup Y$, which is at most $\beta_1 m_p/p$ by (2.5.17). Since $H_{3,i}(E)$ is $(2\varepsilon'^{1/12}, \gamma^2\beta_1)$-superregular (see directly after (2.5.37)), we can apply Proposition 2.3.6(ii) with $H_{3,i}(E)$ playing the role of G, H playing the role of G' and $d' := \beta_1/p$ to find that H is $(\varepsilon^*, \gamma^2\beta_1)$-superregular. Note that $\varepsilon^* \ll \gamma^2\beta_1$ by (2.5.1). This proves the claim. ■

Suppose that we have constructed $f'_1, \dots, f'_t$ with $t < \kappa'$ in the same way as f'_1. Then we will have used t perfect matchings in $H_{3,i}(E)$ for each $E \in J_1$. Let $H^t_{3,i}(J_1)$ denote the subdigraph of $H_{3,i}(J_1)$ consisting of the remaining edges. Then Claim 2.5.7 implies that $H^t_{3,i}(J_1)$ can still play the role of G in Lemma 2.4.5. So we can construct f'_{t+1} in the same way as f'_1. Thus we can obtain $f'_1, \dots, f'_{\kappa'}$ as described above.

Now for each $2 \le q \le p-4$ and each $1 \le t \le \kappa'$ we can use J_q to obtain $f'_{(q-1)\kappa'+t}$ from $f_{(q-1)\kappa'+t}$ in exactly the same way (except that we use edges from $H_{3,i}(J_q)$ and so $J_q \cap E(D_k)$ now plays the role of J for $1 \le k \le \ell$).

More precisely, write $f_{(q-1)\kappa'+t}(0) := f_{(q-1)\kappa'+t}$ and $H^0_{3,i}(J_q) := H_{3,i}(J_q)$. For each $1 \le j \le \ell$ let $K_{(q-1)\kappa'+t}$ be the 1-regular digraph consisting of all cycles of $f_{(q-1)\kappa'+t}(j-1)$ which contain a vertex in a cluster of D_j. Apply Lemma 2.4.5 with D_j playing the role of C, $J_q \cap E(D_j)$ playing the role of J, K_1 playing the role of F and $H^{t-1}_{3,i}(J_q)$ playing the role of G to obtain $f_{(q-1)\kappa'+t}(j)$. By Claim 2.5.7, $H^t_{3,i}(J_q)$ satisfies the conditions required of G in the lemma. Exactly as for f_1 above, $f_{(q-1)\kappa'+t}(j)$ is a 1-factor in which all cycles containing a vertex in D_j have been merged, and if two vertices lie on a common cycle of $f_{(q-1)\kappa'+t}(j-1)$ they also lie on a common cycle of $f_{(q-1)\kappa'+t}(j)$. Write $f'_{(q-1)\kappa'+t} := f_{(q-1)\kappa'+t}(\ell)$. Now let $H^j_{3,i}(J_q)$ denote the remainder of $H^{j-1}_{3,i}(J_q)$ after these ℓ applications of the lemma.

We have now obtained $f'_1, \dots, f'_\kappa$. They are clearly edge-disjoint 1-factors. We claim that f'_j is a Hamilton cycle for each $1 \le j \le \kappa$. Indeed, suppose not. It suffices to

consider f_1'. Let C and C' be cycles in f_1' where C contains a vertex x in some cycle D of F_i and C' contains a vertex x' in some cycle D' of F_i. Recall that, by our construction, for all cycles D_k in F_i, every vertex in (a cluster of) D_k is contained in a single cycle in f_1'. Consider the sequence given by (Red0′) as a cyclic sequence and pick an interval

$$D_g x_g D_{g+1} x_{g+1} \dots x_{g'-1} D_{g'} x_{g'}$$

such that $D = D_g$ and $D' = D_{g'}$. By (Red0′) and (Red1′), the inneighbour of x_g in f_1' is contained in D, so $x_g \in V(C)$. But similarly the outneighbour of x_g in f_1' is contained in D_{g+1}, so all vertices lying in a cluster of D_{g+1} are contained in $V(C)$ and thus $x_{g+1} \in V(C)$. Continuing along the subsequence we conclude that every vertex lying in a cluster of D' lies on C. So x' lies on both C' and C; so since f_1' is a 1-factor we must have $C = C'$. Thus f_1' is a Hamilton cycle, and the same holds for $f_2', \dots, f_\kappa'$.

Finally, we can bound the total number of Hamilton cycles as follows. Note that

$$\kappa \overset{(2.5.5),(2.5.10),(2.5.17)}{=} (1-\gamma)(1-5\gamma)\beta\frac{m}{sp}.$$
$$r_p \overset{(2.5.4),(2.5.13),(2.5.14)}{=} (s-1)(p-1)(\tilde{\alpha}-\gamma)\frac{\tilde{L}}{\beta} \geq (1-\sqrt{\gamma})sp\frac{\tilde{\alpha}\tilde{L}}{\beta}.$$

So altogether, after repeating the procedure for every $1 \leq i \leq r_p$, we have found

$$\begin{aligned} r_p\kappa &\geq (1-\gamma)(1-5\gamma)(1-\sqrt{\gamma})\tilde{\alpha}\tilde{L}\tilde{m} \\ &\overset{(2.5.7)}{\geq} (1-\sqrt{\gamma})^3(1-\varepsilon)\tilde{\alpha}n \\ &\overset{(2.5.1)}{\geq} (1-\eta)r \end{aligned}$$

edge-disjoint Hamilton cycles, as required. This completes the proof of Theorem A.

2.6 The proof of Corollary 2.1.1

We now use Theorem A and Lemma 2.4.1 to prove Corollary 2.1.1.

Proof. As in the proof of Theorem A, we may assume without loss of generality that $0 < \eta \ll \nu \ll \tau \ll \alpha$. Choose n_0 and γ so that $0 < 1/n_0 \ll \gamma \ll \eta$. Suppose that G is a digraph on $n \geq n_0$ vertices satisfying (i) and (ii). Let

$$n_x^{\pm} := d_G^{\pm}(x) - (\alpha - \sqrt{\gamma})n$$

for each $x \in V(G)$. We apply Lemma 2.4.1 to G with $\rho = \varepsilon = \sqrt{\gamma}$ and with $Q = G$ (so $q = 1$) to obtain a subdigraph H of G such that $\tilde{G} := G \setminus H$ is an $(\alpha - \sqrt{\gamma})n$-regular digraph on n vertices. Note that for all $x \in V(G)$ we have $d_{\tilde{G}}^-(x) \geq d_G^-(x) - (\sqrt{\gamma} - \gamma)n \geq d_G^-(x) - \nu n/2$. So for all sets S of vertices,

$$RN^+_{\nu/2,\tilde{G}}(S) \supseteq RN^+_{\nu,G}(S).$$

Thus $\tilde{G}$ is a robust $(\nu/2, \tau)$-outexpander. Therefore we can apply Theorem A to $\tilde{G}$ with parameter $\eta' := \eta/2\alpha$ to find $(1-\eta')(\alpha - \sqrt{\gamma})n > (\alpha - \eta)n$ edge-disjoint Hamilton cycles in $\tilde{G}$ and hence in G. $\square$

CHAPTER 3

THE ROBUST COMPONENT STRUCTURE OF DENSE REGULAR GRAPHS AND APPLICATIONS

3.1 Introduction

3.1.1 The robust component structure of dense regular graphs

The main result of this chapter states that any dense regular graph G is the vertex-disjoint union of a bounded number of 'robust components'. Each such component has a strong expansion property that is highly 'resilient' and almost all edges of G lie inside these robust components. In other words, the result implies that the large scale structure of dense regular graphs is remarkably simple. This can be applied e.g. to Hamiltonicity problems in dense regular graphs. Note that the structural information obtained in this way is quite different from that given by Szemerédi's Regularity lemma.

The crucial notion in our partition is that of robust expansion. This is a structural property which has close connections to Hamiltonicity. Recall the following definitions. Given a graph G on n vertices, $S \subseteq V(G)$ and $0 < \nu \leq \tau < 1$, we define the *ν-robust neighbourhood* $RN_{\nu,G}(S)$ of S to be the set of all those vertices of G with at least νn neighbours in S. We say G is a *robust (ν, τ)-expander* if, for every $S \subseteq V(G)$ with $\tau n \leq |S| \leq (1-\tau)n$, we have that $|RN_{\nu,G}(S)| \geq |S| + \nu n$.

There is an analogous notion of robust outexpansion for digraphs. This was first introduced in [86] and has been instrumental in proving several longstanding conjectures. For example, Kühn and Osthus [81] recently settled a conjecture of Kelly from 1968 (for large tournaments) by showing that every sufficiently large dense regular robust outexpander has a Hamilton decomposition. This was discussed in more detail in Chapter 2. Another example is the recent proof [78, 79] of Sumner's universal tournament conjecture from 1971.

The main result of the current chapter is Theorem B. It allows us to harness the useful consequences of robust expansion even if the graph itself is not a robust expander. For this, we introduce the additional notion of 'bipartite robust expanders'. Let G be a bipartite graph with vertex classes A and B. Then clearly G is not a robust expander as the larger class does not expand. However, we can obtain a bipartite analogue of robust expansion by only considering sets $S \subseteq A$ with $\tau|A| \leq |S| \leq (1-\tau)|A|$. This notion extends in a natural way to graphs which are 'close to bipartite'.

Roughly speaking, our main result (Theorem B) implies the following.

(†) For all $r \in \mathbb{N}$ and all $\varepsilon > 0$, any sufficiently large D-regular graph on n vertices with $D \geq (\frac{1}{r+1} + \varepsilon)n$ has a vertex partition into at most r robust expander components and bipartite robust expander components, so that the number of edges between these is $o(n^2)$.

We give a formal statement of this in Section 3.3. In Section 3.5 we obtain a generalisation to almost regular graphs. (Here, G is 'almost regular' if $\Delta(G) - \delta(G) = o(n)$.)

In the special case of dense vertex-transitive graphs (which are always regular), Christofides, Hladký and Máthé [32] introduced a partition into 'iron connected components'. (Iron connectivity is closely related to robust expansion.) They applied this to resolve the dense case of a question of Lovász [90] on Hamilton paths (and cycles) in vertex-transitive graphs. It would be very interesting to obtain a similar partition result for further classes of graphs. In particular, it might be possible to generalise Theorem B to sparser graphs.

In this chapter, we apply Theorem B to give an approximate solution to a long-standing conjecture on Hamilton cycles in regular graphs (Theorem 3.1.2) as well as an asymptotically optimal result on the circumference of dense regular graphs of given connectivity (Theorem 3.1.4). We are also confident that our robust partition result will have applications to other problems.

Chapter 4 will be devoted to the proof of Theorem C, the exact version of Theorem 3.1.2.

3.1.2 An application to Hamilton cycles in regular graphs

Consider the classical result of Dirac that every graph on $n \geq 3$ vertices with minimum degree at least $n/2$ contains a Hamilton cycle. Suppose we wish to strengthen this by reducing the degree threshold at the expense of introducing some other condition(s). The two extremal examples for Dirac's theorem (i.e. the disjoint union of two cliques and the almost balanced complete bipartite graph) make it natural to consider regular graphs with some connectivity property, see e.g. the recent survey of Li [88] and handbook article of Bondy [20].

In particular, Szekeres (see [63]) asked for which D every 2-connected D-regular graph G on n vertices is Hamiltonian. Jackson [63] showed that $D \geq n/3$ suffices. This improved earlier results of Nash-Williams [94], Erdős and Hobbs [44] and Bollobás and Hobbs [15]. Hilbig [61] improved the degree condition to $n/3 - 1$, unless G is the Petersen graph or another exceptional graph. As discussed later on in this section, this bound is best possible.

Bollobás [13] as well as Häggkvist (see [63]) independently made the natural and far more general conjecture that any t-connected regular graph on n vertices with degree at least $n/(t+1)$ is Hamiltonian. However, the following counterexample (see Figure 3.1(i)), due to Jung [68] and independently Jackson, Li and Zhu [66], disproves this conjecture for $t > 3$.

For m divisible by four, construct G as follows. Let C_1, C_2 be two disjoint copies

of K_{m+1} and let A, B be two disjoint independent sets of orders $m, m-1$ respectively. Add every edge between A and B. Add a set of $m/2$ independent edges from each of C_1 and C_2 to A so that together these edges form a matching of size m. Delete $m/4$ independent edges in each of C_1, C_2 so that G is m-regular. Then G has $4m+1$ vertices and is $m/2$-connected. However G is not Hamiltonian since $G \setminus A$ has $|A|+1$ components (in other words, G is not 1-tough).

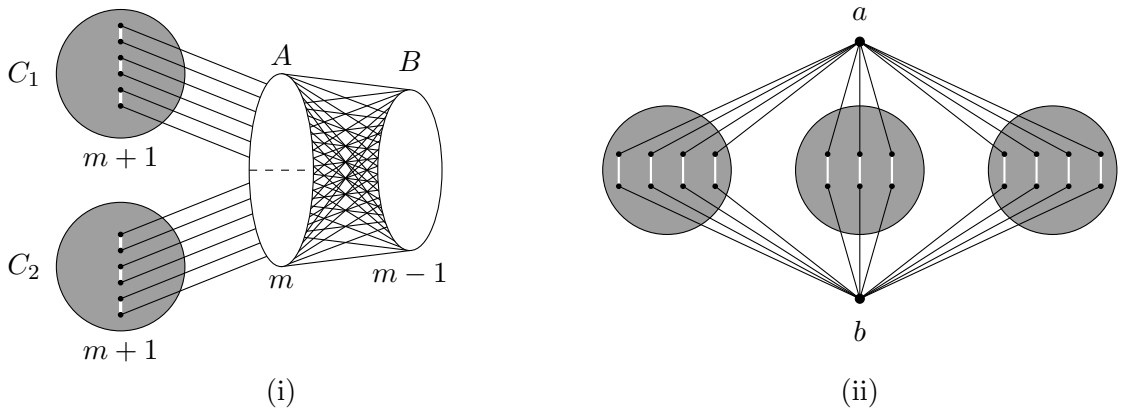

Figure 3.1: Extremal examples for Conjecture 3.1.1.

Jackson, Li and Zhu [66] believe that the conjecture of Bollobás and Häggkvist is true in the remaining open case when $t = 3$.

Conjecture 3.1.1. *Let G be a 3-connected D-regular graph on $n \geq 13$ vertices such that $D \geq n/4$. Then G contains a Hamilton cycle.*

The 3-regular graph obtained from the Petersen graph by replacing one vertex with a triangle shows that the conjecture does not hold for $n = 12$. The graph in Figure 3.1(i) is extremal and the bound on D is tight.

As mentioned earlier, there exist non-Hamiltonian 2-connected regular graphs on n vertices with degree close to $n/3$ (see Figure 3.1(ii)). Indeed, we can construct such a graph G as follows. Start with three disjoint cliques on $3m$ vertices each. In the ith clique choose disjoint sets A_i and B_i with $|A_i| = |B_i|$ and $|A_1| = |A_3| = m$ and $|A_2| = m-1$.

Remove a perfect matching between A_i and B_i for each i. Add two new vertices a and b, where a is connected to all vertices in the sets A_i and b is connected to all vertices in all the sets B_i. Then G is a $(3m-1)$-regular 2-connected graph on $n = 9m+2$ vertices. However, G is not Hamiltonian because $G \setminus \{a, b\}$ has three components. Therefore none of the conditions – degree, order or connectivity – of Conjecture 3.1.1 can be relaxed.

There have been several partial results in the direction of Conjecture 3.1.1. Fan [45] and Jung [68] independently showed that every 3-connected D-regular graph contains a cycle of length at least $3D$, or a Hamilton cycle. Li and Zhu [89] proved Conjecture 3.1.1 in the case when $D \geq 7n/22$ and Broersma, van den Heuvel, Jackson and Veldman [26] proved it for $D \geq 2(n+7)/7$. In [66] it is proved that, if G satisfies the conditions of the conjecture, any longest cycle in G is dominating provided that n is not too small. (Here, a subgraph H of a graph G is *dominating* if $G \setminus V(H)$ is an independent set.) By considering robust partitions, we are able to prove an approximate version of the conjecture.

Theorem 3.1.2. *For all $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that every 3-connected D-regular graph on $n \geq n_0$ vertices with $D \geq (1/4+\varepsilon)n$ is Hamiltonian.*

In fact, if D is at least a little larger than $n/5$ but G is not Hamiltonian we also determine the approximate structure of G (see Theorem 3.7.11). In Chapter 4, we use this to prove the exact version of Conjecture 3.1.1 for large n. Moreover, the proof in Chapter 4 does not supersede the results established in the current chapter, but rather uses them as an essential tool.

There are also natural analogues of the above results and questions for directed graphs. Here, a D-regular directed graph is such that every vertex has both in- and outdegree equal to D. An *oriented graph* is a digraph without 2-cycles.

Conjecture 3.1.3.

(a) For each $D > 2$, every D-regular oriented graph G on n vertices with $D \geq (n-1)/4$ is Hamiltonian.

(b) Every strongly 2-connected D-regular digraph on n vertices with $D \geq n/3$ is Hamiltonian.

(c) For each $D > 2$, every D-regular strongly 2-connected oriented graph G on n vertices with $D \geq n/6$ is Hamiltonian.

(a) was conjectured by Jackson [64], (b) and (c) were raised in [80], which also contains a more detailed discussion of these conjectures.

3.1.3 An application to the circumference of regular graphs

More generally, we also consider the circumference of dense regular graphs of given connectivity. Bondy [19] conjectured that, for $r \geq 3$, every sufficiently large 2-connected D-regular graph G on n vertices with $D \geq n/r$ has circumference $c(G) \geq 2n/(r-1)$. (Here the circumference $c(G)$ of G is the length of the longest cycle in G.) This was confirmed by Wei [114], who proved the conjecture for all n and in fact showed that $c(G) \geq 2n/(r-1) + 2(r-3)/(r-1)$, which is best possible. We are able to extend this (asymptotically) to t-connected dense regular graphs.

Theorem 3.1.4. *Let $t, r \in \mathbb{N}$. For all $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that the following holds. Whenever G is a t-connected D-regular graph on $n \geq n_0$ vertices where $D \geq (1/r + \varepsilon)n$, the circumference of G is at least* $\min\{t/(r-1), 1-\varepsilon\}n$.

This is asymptotically best possible. Indeed, in Proposition 3.8.1 we show that, for every $t, r \in \mathbb{N}$, there are infinitely many n such that there exists a graph G on n vertices which is $((n-t)/(r-1)-1)$-regular and t-connected with $c(G) \leq tn/(r-1) + t$. Moreover, as discussed above, the first extremal example in Figure 3.1 shows that in general $\min\{t/(r-1), 1-\varepsilon\}n$ cannot be replaced by $\min\{t/(r-1), 1\}n$.

Theorem 3.1.4 shows that the conjecture of Bollobás and Häggkvist is in fact close to being true after all – any t-connected regular graph with degree slightly higher than $n/(t+1)$ contains an almost spanning cycle.

3.1.4 An application to bipartite regular graphs

One can consider similar questions about dense regular bipartite graphs. Häggkvist [58] conjectured that every 2-connected D-regular bipartite graph on n vertices with $D \geq n/6$ is Hamiltonian. If true, this result would be best possible. Indeed, it was essentially verified by Jackson and Li [65] who proved it in the case when $D \geq (n+38)/6$. Recently, Li [88] conjectured a bipartite analogue of Conjecture 3.1.1, i.e. that every 3-connected D-regular bipartite graph on n vertices with $D \geq n/8$ is Hamiltonian.

Restricting to bipartite graphs strengthens the structural information implied by our main result Theorem B considerably. So it seems likely that one can use our partition result to make progress towards these and other related conjectures.

One might ask if a bipartite analogue of the conjecture of Bollobás and Häggkvist holds, i.e. whether every t-connected D-regular bipartite graph on n vertices with $D \geq n/2(t+1)$ contains a Hamilton cycle. However, as in the general case, it turns out that this is false for $t > 3$. Indeed, for each $t \geq 2$ and infinitely many $D \in \mathbb{N}$, Proposition 3.8.2 guarantees a D-regular bipartite graph G on $8D + 2$ vertices that is t-connected and contains no Hamilton cycle. (This observation generalises one from [88], which considered the case when $t = 3$.)

As in the general case, one may also consider the circumference of dense regular bipartite graphs. Indeed, the argument for Theorem 3.1.4 yields the following bipartite analogue. Again, it is asymptotically best possible (see Proposition 3.8.2(i)).

Theorem 3.1.5. *Let $t, r \in \mathbb{N}$, where r is even. For all $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that the following holds. Whenever G is a t-connected D-regular bipartite graph on $n \geq n_0$ vertices where $D \geq (1/r+\varepsilon)n$, the circumference of G is at least* $\min\{2tn/(r-2), n\} - \varepsilon n$.

3.1.5 Organisation of the chapter and sketch proof of Theorem 3.1.2

This chapter is organised as follows. In the remainder of this section we sketch the proof of Theorem 3.1.2. Section 4.4 lists some notation which will be used throughout. In Section 3.3 we state our robust partition result (Theorem B) which formalises (†). We prove it in Section 3.4, which also contains a sketch of the argument. In Section 3.5 we derive a version of Theorem B for almost regular graphs. In Section 3.6 we show how to find suitable path systems covering the robust components obtained from Theorem B. These tools are then used in Section 3.7 to prove Theorem 3.1.2 and used in Section 3.8 to prove Theorems 3.1.4 and 3.1.5.

In order to see how our partition result Theorem B may be applied, we now briefly outline the argument used to prove Theorem 3.1.2.

Let $\varepsilon > 0$ and suppose that G is a 3-connected D-regular graph on n vertices, where $D \geq (1/4 + \varepsilon)n$. Now (†) gives us a robust partition $\mathcal{V}$ of G containing exactly k robust expander components and ℓ bipartite robust expander components, where $k + \ell \leq 3$.

However, Theorem B actually gives the stronger bound that $k + 2\ell \leq 3$, so there are only five possible choices of (k, ℓ) (see Proposition 3.3.1). Assume for simplicity that $\mathcal{V}$ consists of three robust expander components G_1, G_2, G_3. So $(k, \ell) = (3, 0)$. The result of [86] mentioned in Section 3.1.1 implies that G_i contains a Hamilton cycle for $i = 1, 2, 3$. In fact, it can be used to show (see Corollary 3.6.8) that G_i is Hamilton p-linked for each bounded p. (Here a graph G is *Hamilton p-linked* if, whenever $x_1, y_1, \ldots, x_p, y_p$ are distinct vertices, there exist vertex-disjoint paths $P_1, \ldots, P_p$ such that P_j connects x_j to y_j, and such that together the paths $P_1, \ldots, P_p$ cover all vertices of G.) This means that the problem of finding a Hamilton cycle in G can be reduced to finding only a suitable set of *external edges*, where an edge is external if it has endpoints in different G_i. We use the assumption of 3-connectivity to find these external edges (in Section 3.7).

The cases where $\ell > 0$ are more difficult since a bipartite graph does not contain a Hamilton cycle if it is not balanced. So as well as suitable external edges, we need to find

some 'balancing edges' incident to the bipartite robust expander component. (Note that if $\ell > 0$ we must have $\ell = 1$ and $k \leq 1$.) Suppose for example that we have $k = \ell = 1$ and that we have a bipartite robust expander component with vertex classes A, B where $|A| = |B| + 1$, as well as a robust expander component X and an edge e joining A to X and an edge f joining B to X, where e and f are disjoint (so e and f are external edges). Then one possible set of balancing edges consists e.g. of two further external edges incident to A. Another example would be one edge inside A. These balancing edges are guaranteed by our assumption that G is regular. We construct them in Section 3.7.

3.2 Notation

For $A \subseteq V(G)$, complements are always taken within the entire graph G, so that $\overline{A} := V(G) \setminus A$. Given $A \subseteq V(G)$, we write $N(A) := \bigcup_{a \in A} N(a)$. For $x \in V(G)$ and $A \subseteq V(G)$ we write $d_A(x)$ for the number of edges xy with $y \in A$. For $A, B \subseteq V(G)$, we write $e(A, B)$ for the number of edges with exactly one endpoint in A and one endpoint in B. (Note that A, B are not necessarily disjoint.) Define $e'(A, B) := e(A, B) + e(A \cap B)$. So $e'(A, B) = \sum_{a \in A} d_B(a) = \sum_{b \in B} d_A(b)$ and if A, B are disjoint then $e'(A, B) = e(A, B)$. For a digraph G, we write $\delta^0(G)$ for the minimum of its minimum indegree and minimum outdegree.

For distinct $x, y \in V(G)$ and a path P with endpoints x and y, we sometimes write $P = xPy$ to emphasise this. Given disjoint subsets A, B of $V(G)$, we say that P is an *AB-path* if P has one endpoint in A and one endpoint in B. We call a vertex-disjoint collection of paths a *path system*. We will often think of a path system $\mathcal{P}$ as a graph with edge set $\bigcup_{P \in \mathcal{P}} E(P)$, so that e.g. $V(\mathcal{P})$ is the union of the vertex sets of each path in $\mathcal{P}$.

3.3 Robust partitions of regular graphs

In this section we list the definitions which are required to state the main result of this chapter. For a graph G on n vertices, $0 < \nu < 1$ and $S \subseteq V(G)$, recall that the *ν-robust neighbourhood* $RN_{\nu,G}(S)$ of S to be the set of all those vertices with at least νn neighbours in S. Also, recall that for $0 < \nu \leq \tau < 1$ we say that G is a *robust (ν, τ)-expander* if, for all sets S of vertices satisfying $\tau n \leq |S| \leq (1-\tau)n$, we have that $|RN_{\nu,G}(S)| \geq |S| + \nu n$. For $S \subseteq X \subseteq V(G)$ we write $RN_{\nu,X}(S) := RN_{\nu,G[X]}(S)$.

We now introduce the concept of 'bipartite robust expansion'. Let $0 < \nu \leq \tau < 1$. Suppose that H is a (not necessarily bipartite) graph on n vertices and that A, B is a partition of $V(H)$. We say that H 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 $|RN_{\nu,H}(S) \cap B| \geq |S| + \nu n$. Note that the order of A and B matters here. We do not mention the bipartition if it is clear from the context.

Note that for $0 < \nu' \leq \nu \leq \tau \leq \tau' < 1$, any robust (ν, τ)-expander is also a robust (ν', τ')-expander (and the analogue holds in the bipartite case).

Given $0 < \rho < 1$, we say that $U \subseteq V(G)$ is a *ρ-component* of a graph G on n vertices if $|U| \geq \sqrt{\rho} n$ and $e_G(U, \overline{U}) \leq \rho n^2$. Note that a ρ-component is not necessarily connected. Let $0 < \rho \leq \nu \leq 1$. Let G be a graph containing a ρ-component U and let $S \subseteq U$. We say that S is *ν-expanding in U* if $|RN_{\nu,U}(S)| \geq |S| + \nu|U|$, and *non-$\nu$-expanding* otherwise. So if $G[U]$ is a robust (ν, τ)-expander for some τ, then all $S \subseteq U$ satisfying $\tau|U| \leq |S| \leq (1-\tau)|U|$ are ν-expanding in U.

Recall that $\overline{U_1} = V(G) \setminus U_1$ and similarly for U_2. Suppose that G is a graph on n vertices and that $U \subseteq V(G)$. We say that U is *ρ-close to bipartite (with bipartition U_1, U_2)* if

(C1) U is the union of two disjoint sets U_1 and U_2 with $|U_1|, |U_2| \geq \sqrt{\rho} n$;

(C2) $\Big||U_1| - |U_2|\Big| \leq \rho n$;

(C3) $e(U_1, \overline{U_2}) + e(U_2, \overline{U_1}) \leq \rho n^2$.

So U is close to bipartite if a balanced bipartite graph can be obtained from U by removing a small number of vertices and edges.

Note that (C1) and (C3) together imply that U is a ρ-component. Suppose that G is a graph on n vertices and that $U \subseteq V(G)$. Let $0 < \rho \leq \nu \leq \tau < 1$. We say that U is a *(ρ, ν, τ)-robust expander component of G* if

(E1) U is a ρ-component;

(E2) $G[U]$ is a robust (ν, τ)-expander.

We say that U is a *bipartite (ρ, ν, τ)-robust expander component (with bipartition A, B) of G* if

(B1) U is ρ-close to bipartite with bipartition A, B;

(B2) $G[U]$ is a bipartite robust (ν, τ)-expander with bipartition A, B.

We say that U is a *(ρ, ν, τ)-robust component* if it is either a (ρ, ν, τ)-robust expander component or a bipartite (ρ, ν, τ)-robust expander component.

Our main result states that any sufficiently dense regular graph has a partition into robust components. Let $k, \ell, D \in \mathbb{N}$ and $0 < \rho \leq \nu \leq \tau < 1$. Given a D-regular graph G on n vertices, we say that $\mathcal{V}$ is a *robust partition of G with parameters ρ, ν, τ, k, ℓ* if the following conditions hold.

(D1) $\mathcal{V} = \{V_1, \ldots, V_k, W_1, \ldots, W_\ell\}$ is a partition of $V(G)$;

(D2) for all $1 \leq i \leq k$, V_i is a (ρ, ν, τ)-robust expander component of G;

(D3) for all $1 \leq j \leq \ell$, there exists a partition A_j, B_j of W_j such that W_j is a bipartite (ρ, ν, τ)-robust expander component with respect to A_j, B_j;

(D4) for all $X, X' \in \mathcal{V}$ and all $x \in X$, we have $d_X(x) \geq d_{X'}(x)$. In particular, $d_X(x) \geq D/m$, where $m := k + \ell$;

(D5) for all $1 \leq j \leq \ell$ we have $d_{B_j}(u) \geq d_{A_j}(u)$ for all $u \in A_j$ and $d_{A_j}(v) \geq d_{B_j}(v)$ for all $v \in B_j$; in particular, $\delta(G[A_j, B_j]) \geq D/2m$;

(D6) $k + 2\ell \leq \lfloor (1+\rho^{1/3})n/D \rfloor$;

(D7) for all $X \in \mathcal{V}$, all but at most ρn vertices $x \in X$ satisfy $d_X(x) \geq D - \rho n$.

As we shall see, (D6) can be derived from (D1)–(D5) but it is useful to state it explicitly. Our main result is the following theorem, which we prove in the next section. As mentioned in the introduction to the chapter, we can use Theorem B to derive a version for almost regular graphs (see Section 3.5).

Theorem B. *For all $\alpha, \tau > 0$ and every non-decreasing function $f : (0,1) \to (0,1)$, there exists $n_0 \in \mathbb{N}$ such that the following holds. For all D-regular graphs G on $n \geq n_0$ vertices where $D \geq \alpha n$, there exist ρ, ν with $1/n_0 \leq \rho \leq \nu \leq \tau$; $\rho \leq f(\nu)$ and $1/n_0 \leq f(\rho)$, and $k, \ell \in \mathbb{N}$ such that G has a robust partition $\mathcal{V}$ with parameters ρ, ν, τ, k, ℓ.*

The technical statement is necessary in order to apply Theorem B e.g. to prove Theorem 3.1.2. One must ensure that the robust partition parameters ρ, ν, τ are sufficiently small compared to the degree parameter ε, but also 'well-spaced' enough.

When the degree of G is large, (D6) implies that there are only a small number of possible choices for k and ℓ.

Proposition 3.3.1. *Let $n, D \in \mathbb{N}$ and suppose that $0 < 1/n \ll \rho \ll \nu \ll \tau \ll 1/r < 1$ and $\rho^{1/3} \leq \varepsilon/2$. Let G be a D-regular graph on n vertices where $D \geq (1/r + \varepsilon)n$ and let $\mathcal{V}$ be a robust partition of G with parameters ρ, ν, τ, k, ℓ. Then $k + 2\ell \leq r - 1$ and so $\ell \leq \lfloor (r-1)/2 \rfloor$ and $k \leq r - 1 - 2\ell$. In particular,*

(i) if $r = 4$ then $(k, \ell) \in \mathcal{S}$, where $\mathcal{S} := \{(1,0), (2,0), (3,0), (0,1), (1,1)\}$;

(ii) if $r = 5$ then $(k, \ell) \in \mathcal{S} \cup \{(4,0), (2,1), (0,2)\}$.

Proof. It suffices to show that $k+2\ell \leq r-1$. By (D6) and our assumption that $\rho^{1/3} \leq \varepsilon/2$ we have

$$k + 2\ell \leq \left\lfloor \frac{1+\varepsilon/2}{1/r+\varepsilon} \right\rfloor = \left\lfloor \frac{r + r\varepsilon/2}{1 + r\varepsilon} \right\rfloor = r - 1,$$

as required. □

We will prove Theorem 3.1.2 separately for each of these cases in Proposition 3.3.1(i). Proposition 3.3.1 is the only point of the proof where we need the full strength of the degree condition $D \geq (1/4 + \varepsilon)n$. (Within each case, $D \geq \varepsilon n$ will do.) Furthermore, Proposition 3.3.1(ii) implies that a $\lceil n/4 \rceil$-regular graph could have any of the structures specified by (i), as well as $(k, \ell) = (4, 0), (2, 1), (0, 2)$. Note also that Figure 3.1(i) has $(k, \ell) = (2, 1)$ and Figure 3.1(ii) has $(k, \ell) = (3, 0)$.

3.4 The proof of Theorem B

We begin by giving a brief sketch of the argument.

3.4.1 Sketch proof of Theorem B

The basic proof strategy is to successively refine an appropriate partition of G. So let G be a D-regular graph on n vertices, where D is linear in n. Suppose that G is not a (bipartite) robust expander. Then $V(G)$ contains a set S such that N is not much larger than S, where $N := RN_{\nu,G}(S)$ for appropriate ν. Consider a minimal S with this property. Since G is regular, N cannot be significantly smaller than S. One can use this to show that there are very few edges between $S \cup N$ and $X := V(G) \setminus (S \cup N)$. Moreover, one can show that S and N are either almost identical or almost disjoint. In the former case, $G[S \cup N]$ is close to a robust expander and in the latter $G[S \cup N]$ is close to a bipartite robust expander. So in both cases, $S \cup N$ is close to a (bipartite) robust expander component. Similarly, if X is non-empty, it is either a (bipartite) robust expander component or we can partition it further along the above lines. In this way, we eventually arrive at the desired partition.

3.4.2 Preliminary observations

We will often use the following simple observation about ρ-components.

Lemma 3.4.1. *Let $n, D \in \mathbb{N}$ and $\rho, \rho', \gamma > 0$ such that $\rho \leq \rho'$ and $\gamma \geq \rho + \rho'$. Let G be a D-regular graph on n vertices and let U be a ρ-component of G. Then*

(i) $|U| \geq D - \sqrt{\rho}n$;

(ii) if W, W' is a partition of U and W is a ρ'-component of G, then $e(W', \overline{W'}) \leq \gamma n^2$;

(iii) if $D \geq 2\sqrt{\rho'}n$, then U is a ρ'-component of G.

Let $X \subseteq V(G)$ have bipartition X_1, X_2 such that X is ρ-close to bipartite with bipartition X_1, X_2. Then

(iv) $|X_1|, |X_2| \geq D - 2\sqrt{\rho}n$;

(v) if $D \geq 3\sqrt{\rho'}n$, then X is ρ'-close to bipartite.

Proof. To prove (i), note that

$$|U|D = \sum_{x \in U} d_G(x) = 2e_G(U) + e_G(U, \overline{U}) \leq |U|^2 + \rho n^2.$$

So $|U| \geq D - \rho n^2/|U| \geq D - \sqrt{\rho}n$, as required. To see (ii), note that

$$e(W', \overline{W'}) = e(W', W) + e(W', \overline{U}) \leq e(\overline{W}, W) + e(U, \overline{U}) \leq (\rho + \rho')n^2 \leq \gamma n^2.$$

To see (iii), note first that $e(U, \overline{U}) \leq \rho n^2 \leq \rho' n^2$. Furthermore, (i) implies that $|U| \geq D - \sqrt{\rho}n \geq \sqrt{\rho'}n$.

We now prove (iv). Since $e'(X_1, \overline{X_2}) \leq 2e(X_1, \overline{X_2}) \leq 2\rho n^2$ and since G is D-regular, we have that

$$|X_1|D - 2\rho n^2 \leq e'(X_1, V(G)) - e'(X_1, \overline{X_2}) = e'(X_1, X_2) \leq |X_1||X_2|. \tag{3.4.1}$$

So $|X_2| \geq D - 2\rho n^2/|X_1| \geq D - 2\sqrt{\rho}n$. A similar argument shows that $|X_1| \geq D - 2\sqrt{\rho}n$. Finally, (v) holds since (C2) and (C3) are immediate, and (C1) follows from (iv). $\square$

The following lemma implies that, for any regular graph G and any $S \subseteq V(G)$ that is not too small, the robust neighbourhood of S cannot be significantly smaller than S itself.

Lemma 3.4.2. *Let $n, D \in \mathbb{N}$ and suppose that $0 < 1/n \ll \rho \ll \nu \ll \tau, \alpha < 1$. Let U be a ρ-component of a D-regular graph G on n vertices, where $D \geq \alpha n$. Let $S \subseteq U$ satisfy $|S| \geq \tau|U|$. Write $N := RN_{\nu,U}(S)$ and let $Y := S \setminus N$ and $W := V(G) \setminus (S \cup N)$. Then*

(i) $e(S, Y) \leq \nu n^2$ and $e(S, W) \leq 2\nu n^2$;

(ii) $|N| \geq |S| - \sqrt{\nu} n$;

(iii) $|N| \geq D - \sqrt{\nu} n$.

Proof. To prove (i), note that $e(S, Y) = e_{G[U]}(S, Y) \leq |Y|\nu|U| \leq \nu n^2$. Moreover, $e'(S, \overline{N} \cap U) = \sum_{x \in \overline{N} \cap U} d_S(x) \leq \nu|U|^2 \leq \nu n^2$. Since U is a ρ-component of G, we have that $e(U, \overline{U}) \leq \rho n^2$. Hence

$$e'(S, \overline{N}) = e'(S, \overline{N} \cap U) + e(S, \overline{U}) \leq (\nu + \rho)n^2 \leq 2\nu n^2. \tag{3.4.2}$$

This proves (i) as $e(S, W) \leq e'(S, \overline{N})$.

We now prove (ii). Certainly $e'(S, N) \leq \sum_{x \in N} d(x) = D|N|$. Similarly

$$e'(S, N) = D|S| - e'(S, \overline{N}) \overset{(3.4.2)}{\geq} D|S| - 2\nu n^2. \tag{3.4.3}$$

Then $|N| \geq |S| - 2\nu n^2/D \geq |S| - \sqrt{\nu} n$, which proves (ii). Finally, we prove (iii). Lemma 3.4.1(i) implies that $|U| \geq D - \sqrt{\rho} n$, so

$$|S| \geq \tau|U| \geq \tau D/2. \tag{3.4.4}$$

Moreover, (3.4.3) implies that $|S||N| \geq e'(S, N) \geq D|S| - 2\nu n^2$ and hence

$$|N| \geq D - \frac{2\nu}{|S|} n^2 \overset{(3.4.4)}{\geq} D - \frac{4\nu}{\tau D} n^2 \geq D - \sqrt{\nu} n,$$

as required. $\square$

The next lemma gives some sufficient conditions for U to be close to bipartite when G is a regular graph and $U \subseteq V(G)$.

Lemma 3.4.3. *Let $n, D \in \mathbb{N}$ and suppose that $0 < 1/n \ll \gamma' \leq \gamma \ll \alpha < 1$ where $\gamma' \leq \gamma^{7/6}$. Suppose that G is a D-regular graph on n vertices where $D \geq \alpha n$. Let Y, Z be disjoint subsets of $V(G)$ such that*

(i) $|Y| \geq \gamma n$;

(ii) $\Big||Y| - |Z|\Big| \leq \gamma n$;

(iii) $e(Y, \overline{Z}) \leq \gamma' n^2$.

Then $Y \cup Z$ is $\gamma^{1/3}$-close to bipartite with bipartition Y, Z.

Proof. First note that (C2) certainly holds with $\gamma^{1/3}$ playing the role of ρ. Since $e'(Y, \overline{Z}) \leq 2e(Y, \overline{Z}) \leq 2\gamma' n^2$ and G is D-regular, we have that

$$|Y|D - 2\gamma' n^2 \leq e'(Y, V(G)) - e'(Y, \overline{Z}) = e'(Y, Z) \leq |Y||Z|. \tag{3.4.5}$$

So $|Z| \geq D - 2\gamma' n^2/|Y| \geq 2\gamma^{1/6} n$ and $|Y| \geq |Z| - \gamma n \geq \gamma^{1/6} n$. Thus (C1) holds. We also have that

$$\begin{aligned} e(Z, \overline{Y}) &\leq e'(Z, \overline{Y}) = |Z|D - e'(Y, Z) \overset{(3.4.5)}{\leq} (|Z| - |Y|)D + 2\gamma' n^2 \\ &\overset{\text{(ii)}}{\leq} D\gamma n + 2\gamma' n^2 \leq 3\gamma n^2. \end{aligned}$$

So $e(Y, \overline{Z}) + e(Z, \overline{Y}) \leq 4\gamma n^2 \leq \gamma^{1/3} n^2$ and therefore (C3) holds. $\square$

We now show that if a regular graph G contains a non-expanding set S whose intersection with its robust neighbourhood is small, then G contains an induced subgraph which is close to bipartite.

Lemma 3.4.4. *Let $n, D \in \mathbb{N}$ and suppose that $0 < 1/n \ll \rho \ll \nu \ll \tau, \alpha < 1$. Suppose that G is a D-regular graph on n vertices where $D \geq \alpha n$. Let $U \subseteq V(G)$ be a ρ-component of G. Suppose that $S \subseteq U$ is non-ν-expanding in U and $|S| \geq \tau|U|$. Let $N := RN_{\nu,U}(S)$, $Y := S \setminus N$ and $Z := N \setminus S$. Then*

(i) $\Big||Y| - |Z|\Big| \leq \sqrt{\nu}n$;

(ii) if also $|Y| > \sqrt{\nu}n$, then $Y \cup Z$ is $\nu^{1/6}$-close to bipartite with bipartition Y, Z.

Proof. Let $X := S \cap N$. So $S = X \cup Y$ and $N = X \cup Z$. Since S is non-ν-expanding in U, we have that $|N| < |S| + \nu|U|$. By Lemma 3.4.2(ii) we have that

$$|S| - \sqrt{\nu}n \leq |N| < |S| + \nu|U| \leq |S| + \sqrt{\nu}n,$$

which proves (i). To prove (ii), let $W := \overline{S \cup N} = \overline{X \cup Y \cup Z}$. Note that Lemma 3.4.2(i) implies that

$$e(Y, \overline{Z}) = e(Y, S \cup W) \leq e(S, Y) + e(S, W) \leq 3\nu n^2. \tag{3.4.6}$$

Set $\gamma' := 3\nu$ and $\gamma := \sqrt{\nu}$. Then $\gamma' \leq \gamma^{5/6}$. So we can apply Lemma 3.4.3 to see that $Y \cup Z$ is $\nu^{1/6}$-close to bipartite with bipartition Y, Z. □

The next proposition formalises the fact that, if a graph G contains a subset U that is close to bipartite; we may add or remove any small set of vertices so that it is still close to bipartite (with slightly weaker parameters).

Proposition 3.4.5. *Let $n, D \in \mathbb{N}$, $0 < 1/n \ll \rho_1, \rho_2 \ll \alpha < 1$ and let $\rho \geq \rho_1 + 2\rho_2$. Suppose that G is a D-regular graph on n vertices where $D \geq \alpha n$ and let $U \subseteq V(G)$ be such that U is ρ_1-close to bipartite, with bipartition A, B. Suppose that $A', B' \subseteq V(G)$ are disjoint and $|A \triangle A'| + |B \triangle B'| \leq \rho_2 n$. Let $U' := A' \cup B'$. Then U' is ρ-close to bipartite with bipartition A', B'.*

Proof. We need to check that (C1)–(C3) hold. Lemma 3.4.1(iv) implies that

$$|A'| \geq |A| - \rho_2 n \geq D - (2\sqrt{\rho_1} + \rho_2)n \geq \sqrt{\rho}n,$$

and similarly for $|B'|$. So (C1) holds. Also

$$\Big||A'| - |B'|\Big| \leq |A' \triangle A| + \Big||A| - |B|\Big| + |B \triangle B'| \leq (\rho_1 + \rho_2)n \leq \rho n,$$

so (C2) holds. Moreover,

$$\begin{aligned} e(A', \overline{B'}) + e(B', \overline{A'}) &\leq e(A, \overline{B}) + e(B, \overline{A}) + 2(|A' \triangle A| + |B' \triangle B|)n \\ &\leq (\rho_1 + 2\rho_2)n^2 \leq \rho n^2. \end{aligned}$$

So (C3) holds, as required. □

3.4.3 Properties of non-expanding subsets

In this subsection we prove that a ρ-component is either a robust expander component, a bipartite robust expander component, or the union of two ρ'-components (where $\rho \ll \rho'$). This forms the core of the proof of Theorem B.

For this, we first show that if U is a ρ-component in a regular graph G such that $G[U]$ is not a robust expander, then either U is close to bipartite, or U can be decomposed into two ρ'-components. To prove this, we consider a non-expanding set S and its robust neighbourhood N. We use our previous results to show that either $S \cup N$ and its complement in U are both ρ'-components or U is ρ'-close to bipartite.

Lemma 3.4.6. *Let $n \in \mathbb{N}$, suppose that $0 < 1/n \ll \rho \ll \nu \ll \rho' \ll \tau \ll \alpha < 1$ and let $D \geq \alpha n$ be a natural number. Let U be a ρ-component of a D-regular graph G on n vertices. Suppose that $G[U]$ is not a robust (ν, τ)-expander. Then at least one of the following hold:*

(i) U has a partition U_1, U_2 such that each of U_1, U_2 is a ρ'-component of G;

(ii) U is ρ'-close to bipartite.

Proof. Since $G[U]$ is not a robust (ν, τ)-expander, there exists $S \subseteq U$ with

$$\tau|U| \leq |S| \leq (1-\tau)|U| \tag{3.4.7}$$

and $|RN_{\nu,U}(S)| < |S| + \nu|U|$. Let $N := RN_{\nu,U}(S)$, $X := S \cap N$, $Y := S \setminus N$, $Z := N \setminus S$ and $W := V(G) \setminus (S \cup N)$. We consider two cases, depending on the size of Y.

Case 1. $|Y| \leq \sqrt{\nu}n$.

In this case, we will show that (i) holds. Let $U_1 := S \cup N = S \cup Z$ so that $\overline{U_1} = W$. Then Lemma 3.4.2(iii) implies that $|U_1| \geq |N| \geq D - \sqrt{\nu}n \geq \sqrt{\rho'}n$. By Lemma 3.4.4(i), we have

$$|Z| \leq |Y| + \sqrt{\nu}n \leq 2\sqrt{\nu}n \tag{3.4.8}$$

$$\leq \tau D/4 \leq \tau|U|/2, \tag{3.4.9}$$

where the last inequality holds since $|U| \geq D - \sqrt{\rho}n$ by Lemma 3.4.1(i).

Now Lemma 3.4.2(i) implies that

$$e(U_1, \overline{U_1}) = e(S, W) + e(Z, W) \leq 2\nu n^2 + |Z|n \overset{(3.4.8)}{\leq} 3\sqrt{\nu}n^2.$$

So U_1 is a $3\sqrt{\nu}$-component of G. Moreover, (3.4.9) and (3.4.7) together imply that

$$|U_1| = |S| + |Z| \leq (1 - \tau/2)\,|U|.$$

Let $U_2 := U \setminus U_1$. Then $|U_2| \geq \tau|U|/2 \geq \sqrt{\rho'}n$. Since U is a ρ-component, U_1 is a $3\sqrt{\nu}$-component, and $\rho + 3\sqrt{\nu} \leq \rho'$, we can apply Lemma 3.4.1(ii) with $U_1, U_2, \rho, 3\sqrt{\nu}, \rho'$ playing the roles of $W, W', \rho, \rho', \gamma$ respectively to see that $e(U_2, \overline{U_2}) \leq \rho' n^2$. Thus U_2 is a ρ'-component of G and so (i) holds.

Case 2. $|Y| > \sqrt{\nu}n$.

Let $U_1 := Y \cup Z = S \triangle N$. Lemma 3.4.4(ii) implies that U_1 is $\nu^{1/6}$-close to bipartite with

bipartition Y, Z. Therefore (C1) and (C3) imply that U_1 is a $\nu^{1/6}$-component. Moreover, $|U_1| \geq 2(D - 2\nu^{1/12}n) \geq 2(D - 2\sqrt{\rho'}n)$ by Lemma 3.4.1(iv). Let $U_2 := U \setminus U_1$. Now Lemma 3.4.1(ii) with $U_1, U_2, \rho, \nu^{1/6}, (\rho'/3)^2$ playing the roles of $W, W', \rho, \rho', \gamma$ implies that $e(U_2, \overline{U_2}) \leq (\rho'/3)^2 n^2$. If $|U_2| \geq \rho' n/3$ then U_2 is a $(\rho'/3)^2$-component. So Lemma 3.4.1(i) implies that $|U_2| \geq D - \rho' n/3$ and thus U_2 is actually a ρ'-component of G. So (i) holds in this case.

Thus we may assume that $|U_2| < \rho' n/3$. Let $Y' := Y \cup U_2$ and $Z' := Z$. Then Y', Z' are disjoint subsets whose union is U. Note that $|Y' \triangle Y| + |Z' \triangle Z| = |U_2| < \rho' n/3$. Now Proposition 3.4.5 with $U_1, U, Y, Z, Y', Z', \nu^{1/6}, \rho'/3, \rho'$ playing the roles of $U, U', A, B, A', B', \rho_1, \rho_2, \rho$ implies that U is ρ'-close to bipartite with bipartition Y', Z'. So (ii) holds. □

The following lemma is a bipartite analogue of Lemma 3.4.6. It states that if G is a regular graph and $U \subseteq V(G)$ such that U is close to bipartite and $G[U]$ is not a bipartite robust expander, then U can be decomposed into two components. The proof is similar to that of Lemma 3.4.6 – we find the partition by considering a non-expanding set and its robust neighbourhood.

Lemma 3.4.7. *Let $n, D \in \mathbb{N}$ and suppose that $0 < 1/n \ll \rho \ll \nu \ll \rho' \ll \tau \ll \alpha < 1$. Let G be a D-regular graph on n vertices where $D \geq \alpha n$. Suppose that $U \subseteq V(G)$ is such that U is ρ-close to bipartite with bipartition A, B and $G[U]$ is not a bipartite robust (ν, τ)-expander with bipartition A, B. Then there is a partition U_1, U_2 of U such that U_1, U_2 are ρ'-components.*

Proof. Since U is ρ-close to bipartite with bipartition A, B, Lemma 3.4.1(iv) and (C2) imply that

$$|A|, |B| \geq D - 2\sqrt{\rho}n \geq D/2, \tag{3.4.10}$$

$$\Big||A| - |B|\Big| \leq \rho n. \tag{3.4.11}$$

Since $G[U]$ is not a bipartite robust (ν, τ)-expander with bipartition A, B, there exists $S \subseteq A$ with

$$\tau|A| \leq |S| \leq (1-\tau)|A| \tag{3.4.12}$$

and such that

$$|N \cap B| < |S| + \nu|U|, \tag{3.4.13}$$

where $N := RN_{\nu,U}(S)$. We claim that

$$|N \cap A| \leq \sqrt{\rho}n. \tag{3.4.14}$$

To see this, note that (C3) implies that

$$\rho n^2 \geq e(A) = \frac{1}{2}\sum_{x \in A} d_A(x) \geq \frac{1}{2}\sum_{x \in N \cap A} d_S(x) \geq \frac{1}{2}\nu|U||N \cap A|.$$

So

$$|N \cap A| \leq 2\rho n^2/\nu|U| \overset{(3.4.10)}{\leq} 2\rho n^2/\nu D \leq \sqrt{\rho}n, \tag{3.4.15}$$

proving the claim. Therefore

$$|S| + \nu n \geq |S| + \nu|U| \overset{(3.4.13)}{>} |N \cap B| = |N| - |N \cap A| \overset{(3.4.14)}{\geq} |N| - \sqrt{\rho}n.$$

Let $\nu_0 := 2\nu/\alpha$. Then

$$|RN_{\nu_0,U}(S)| \leq |N| \leq |S| + (\nu + \sqrt{\rho})n < |S| + 2\nu n \overset{(3.4.10)}{\leq} |S| + \nu_0|U|. \tag{3.4.16}$$

That is, S is non-ν_0-expanding in U. Let $X := N \cap S$, $Y := S \setminus N$, $Z := N \setminus S$ as in Lemma 3.4.4.

Note that $X \subseteq S \subseteq A$ and $X \subseteq N$, so (3.4.14) implies that $|X| \leq \sqrt{\rho}n$. Moreover,

$$|Y| = |S| - |X| \geq |S| - \sqrt{\rho}n \overset{(3.4.12)}{\geq} \tau|A| - \sqrt{\rho}n \overset{(3.4.10)}{\geq} \tau D/3 > \sqrt{\nu_0}n. \tag{3.4.17}$$

Let $U_1 := Y \cup Z$ and $U_2 := U \setminus U_1$. Let $\nu' := \nu_0^{1/6}$. Now (3.4.11) and (3.4.12) imply that $|S| \geq \tau|U|/3$. Then (3.4.16), (3.4.17) and Lemma 3.4.4(ii) with $\rho, \nu_0, \tau/3$ playing the roles of ρ, ν, τ imply that U_1 is ν'-close to bipartite and hence a ν'-component, as required. Now Lemma 3.4.1(ii) with U, U_1, U_2, ν', ρ' playing the roles of U, W, W', ρ', γ implies that $e(U_2, \overline{U_2}) \leq \rho' n^2$. Note that

$$\begin{aligned} |U_2| &\geq |U| - |S| - |N| \overset{(3.4.16)}{\geq} |U| - 2|S| - 2\nu n/\alpha \\ &\overset{(3.4.12)}{\geq} |B| - |A| + 2\tau|A| - 2\sqrt{\nu} n \overset{(3.4.10),(3.4.11)}{\geq} \tau|A| \geq \sqrt{\rho'}|U|. \end{aligned}$$

Therefore U_2 is a ρ'-component. □

3.4.4 Adjusting partitions

The results of this subsection will be needed to ensure (D4), (D5) and (D7) in the proof of Theorem B.

The next two lemmas state that (bipartite) robust expanders are indeed robust, in the sense that the expansion property cannot be destroyed by adding or removing a small number of vertices.

Lemma 3.4.8. *Let $0 < \nu \ll \tau \ll 1$. Suppose that G is a graph and $U, U' \subseteq V(G)$ are such that $G[U]$ is a robust (ν, τ)-expander and $|U \triangle U'| \leq \nu|U|/4$. Then $G[U']$ is a robust $(\nu/2, 2\tau)$-expander.*

Proof. Let $S \subseteq U'$ be such that $2\tau|U'| \leq |S| \leq (1-2\tau)|U'|$. Then $\tau|U| \leq |S \cap U| \leq (1-\tau)|U|$. Observe that $RN_{\nu/2,U'}(S) \supseteq RN_{\nu,U}(S \cap U) \cap U'$. Since $G[U]$ is a robust (ν, τ)-expander, we have that

$$\begin{aligned} |RN_{\nu/2,U'}(S)| &\geq |RN_{\nu,U}(S \cap U)| - |U \setminus U'| \geq |S \cap U| + \nu|U| - |U \setminus U'| \\ &\geq |S| + \nu|U| - |U \triangle U'| \geq |S| + 3\nu|U|/4 \geq |S| + \nu|U'|/2, \end{aligned}$$

as required. □

Lemma 3.4.9. *Let $0 < \nu \ll \tau \ll 1$. Suppose that $U \subseteq V(G)$ and that $G[U]$ is a bipartite robust (ν, τ)-expander with bipartition A, B. Let $W, A', B' \subseteq V(G)$ be such that $|W| \leq \nu|A|/4$; A' and B' are disjoint; and $|A \triangle A'| + |B \triangle B'| \leq \nu|A|/4$. Then*

(i) $G[U \setminus W]$ is a bipartite robust $(\nu/2, 2\tau)$-expander with bipartition $A \setminus W, B \setminus W$;

(ii) $G[A' \cup B']$ is a bipartite robust $(\nu/2, 2\tau)$-expander with bipartition A', B'.

Proof. To prove (i), let $S \subseteq A \setminus W$ be such that $2\tau|A \setminus W| \leq |S| \leq (1 - 2\tau)|A \setminus W|$. Then $\tau|A| \leq |S| \leq (1 - \tau)|A|$. Observe that $RN_{\nu/2, B \setminus W}(S) \supseteq RN_{\nu, B}(S) \setminus W$. So

$$|RN_{\nu/2, B \setminus W}(S)| \geq |RN_{\nu, B}(S)| - |W| \geq |S| + 3\nu|U|/4 \geq |S| + \nu|U \setminus W|/2,$$

as required.

To prove (ii), let $S \subseteq A'$ be such that $2\tau|A'| \leq |S| \leq (1 - 2\tau)|A'|$. Then $\tau|A| \leq |S \cap A| \leq (1 - \tau)|A|$. Observe that $RN_{\nu/2, B'}(S) \supseteq RN_{\nu, B}(S \cap A) \cap B'$. Therefore

$$\begin{aligned} |RN_{\nu/2, B'}(S)| &\geq |RN_{\nu, B}(S \cap A)| - |B \setminus B'| \geq |S \cap A| + \nu|A \cup B| - |B \setminus B'| \\ &\geq |S| + \nu|A \cup B| - |B \setminus B'| - |A' \setminus A| \geq |S| + \nu|A' \cup B'|/2, \end{aligned}$$

as required. □

We now extend Lemma 3.4.9 by showing that, after adding and removing a small number of vertices, a bipartite robust component is still a bipartite robust component, with slightly weaker parameters.

Lemma 3.4.10. *Let $0 < 1/n \ll \rho \leq \gamma \ll \nu \ll \tau \ll \alpha < 1$ and suppose that G is a D-regular graph on n vertices where $D \geq \alpha n$.*

(i) Suppose that $A \cup B$ is a bipartite (ρ, ν, τ)-robust expander component of G with bipartition A, B. Let $A', B' \subseteq V(G)$ be disjoint such that $|A \triangle A'| + |B \triangle B'| \leq \gamma n$. Then $A' \cup B'$ is a bipartite $(3\gamma, \nu/2, 2\tau)$-robust expander component of G with bipartition A', B'.

(ii) Suppose that U is a bipartite (ρ, ν, τ)-robust expander component of G. Let $U' \subseteq V(G)$ be such that $|U \triangle U'| \leq \gamma n$. Then U' is a bipartite $(3\gamma, \nu/2, 2\tau)$-robust expander component of G.

Proof. We first prove (i). Observe that (B1) for $A \cup B$ and Lemma 3.4.1(iv) imply that $|A| \geq D - 2\sqrt{\rho}n \geq D/2$. Proposition 3.4.5 with $\rho, \gamma, 3\gamma$ playing the roles of ρ_1, ρ_2, ρ implies that $A' \cup B'$ is 3γ-close to bipartite with bipartition A', B'. So (B1) holds. Now Lemma 3.4.9(ii) implies that $G[A' \cup B']$ is a bipartite robust $(\nu/2, 2\tau)$-expander with bipartition A', B', so (B2) holds. This completes the proof of (i). It is easy to see that (ii) follows from (i). □

In any ρ-component, almost all vertices have very few neighbours outside the component. In particular, most vertices have more neighbours within their own component than in any other. The following lemma allows us to move a small number of vertices in a partition into ρ-components so that this property holds for all vertices.

Lemma 3.4.11. *Let $m, n, D \in \mathbb{N}$ and $0 < 1/n \ll \rho \ll \alpha, 1/m \leq 1$. Let G be a D-regular graph on n vertices where $D \geq \alpha n$. Suppose that $\mathcal{U} := \{U_1, \ldots, U_m\}$ is a partition of $V(G)$ such that U_i is a ρ-component for each $1 \leq i \leq m$. Then G has a vertex partition $\mathcal{V} := \{V_1, \ldots, V_m\}$ such that*

(i) $|U_i \triangle V_i| \leq \rho^{1/3} n$;

(ii) V_i is a $\rho^{1/3}$-component for each $1 \leq i \leq m$;

(iii) if $x \in V_i$ then $d_{V_i}(x) \geq d_{V_j}(x)$ for all $1 \leq i, j \leq m$. In particular, $d_V(x) \geq D/m$ for all $x \in V$ and all $V \in \mathcal{V}$;

(iv) for all but at most $\rho^{1/3} n$ vertices $x \in V_i$ we have $d_{V_i}(x) \geq D - 2\sqrt{\rho}n$.

Proof. First note that the second part of (iii) follows from the first. For each $1 \leq i \leq m$, let X_i be the collection of vertices $y \in U_i$ with $d_{\overline{U_i}}(y) \geq \sqrt{\rho}n$. Since U_i is a ρ-component,

we have $|X_i| \leq \sqrt{\rho}n$. Let $W_i := U_i \setminus X_i$. Then each $x \in W_i$ satisfies

$$d_{W_i}(x) = d(x) - d_{\overline{U_i} \cup X_i}(x) \geq d(x) - \sqrt{\rho}n - |X_i| \geq d(x) - 2\sqrt{\rho}n. \tag{3.4.18}$$

Let $X := \bigcup_{1 \leq i \leq m} X_i$. Among all partitions $X'_1, \ldots, X'_m$ of X, choose one such that $\sum_{1 \leq i \leq m} e(V_i, \overline{V_i})$ is minimal, where $V_i := W_i \cup X'_i$. It is easy to see that $d_{V_i}(x) \geq d_{V_j}(x)$ for all $x \in X'_i$ and all $1 \leq i, j \leq m$. So (iii) holds for all $x \in X'_i$ and $i \leq m$. Moreover, if $x \in W_i$, then (3.4.18) implies that $d_{V_i}(x) \geq d_{W_i}(x) \geq d(x) - 2\sqrt{\rho}n \geq d(x)/2$. So (iii) also holds for each vertex in W_i. Furthermore, by minimality,

$$\sum_{1 \leq i \leq m} e(V_i, \overline{V_i}) \leq \sum_{1 \leq i \leq m} e(U_i, \overline{U_i}) \leq \rho m n^2 \leq \rho^{1/3} n^2$$

and hence each V_i is a $\rho^{1/3}$-component, so (ii) holds.

Note that $U_i \cap V_i \supseteq W_i$, so

$$|U_i \triangle V_i| \leq \sum_{1 \leq i \leq m} |X'_i| = |X| \leq m\sqrt{\rho}n \leq \rho^{1/3} n, \tag{3.4.19}$$

which proves (i). Finally, (3.4.18) and the fact that $|V_i \setminus W_i| \leq |X| \leq \rho^{1/3} n$ by (3.4.19) together imply (iv). □

The next lemma shows that, in a bipartite robust expander component, we can adjust the bipartition slightly so that any vertex has at least as many neighbours in the opposite class as within its own class. The resulting graph will still be a bipartite robust expander component. The proof is very similar to that of Lemma 3.4.11.

Lemma 3.4.12. *Let $0 < 1/n \ll \rho \ll \nu \ll \tau \ll \alpha < 1$ and let G be a D-regular graph on n vertices where $D \geq \alpha n$. Suppose that U is a bipartite (ρ, ν, τ)-robust component of G. Then there exists a bipartition A, B of U such that*

(i) U is a bipartite $(3\sqrt{\rho}, \nu/2, 2\tau)$-robust component with partition A, B;

(ii) $d_B(u) \geq d_A(u)$ for all $u \in A$, and $d_A(v) \geq d_B(v)$ for all $v \in B$;

Proof. Let X, Y be a bipartition of U such that U is a bipartite (ρ, ν, τ)-robust expander component of G with respect to X, Y. Let X_0 be the collection of vertices $x \in X$ with $d_{\overline{Y}}(x) \geq 2\sqrt{\rho}n$. Let Y_0 be the collection of vertices $y \in Y$ with $d_{\overline{X}}(y) \geq 2\sqrt{\rho}n$. Then (B1) implies that

$$\rho n^2 \geq e(X, \overline{Y}) + e(Y, \overline{X}) \geq \frac{1}{2}\left(\sum_{x \in X} d_{\overline{Y}}(x) + \sum_{y \in Y} d_{\overline{X}}(y)\right) \geq \frac{1}{2}\left(\sum_{x \in X_0} d_{\overline{Y}}(x) + \sum_{y \in Y_0} d_{\overline{X}}(y)\right)$$
$$\geq (|X_0| + |Y_0|)\sqrt{\rho}n,$$

and so $|X_0| + |Y_0| \leq \sqrt{\rho}n$. Let $X' := X \setminus X_0$ and $Y' := Y \setminus Y_0$. So for all $x \in X'$, $d_{Y'}(x) \geq d(x) - d_{\overline{Y}}(x) - |Y_0| \geq D - 3\sqrt{\rho}n$. An analogous statement holds for all vertices $y \in Y'$.

Among all partitions A_0, B_0 of $X_0 \cup Y_0$, choose one such that $e(A, \overline{B}) + e(B, \overline{A})$ is minimal, where $A := X' \cup A_0$ and $B := Y' \cup B_0$. We claim that A, B is the required partition.

Indeed, our choice of A_0, B_0 implies that (ii) holds for all $u \in A_0$ and all $v \in B_0$. Moreover, if $u \in X'$, then $d_B(u) \geq d_{Y'}(u) \geq D - 3\sqrt{\rho}n \geq d_U(u)/2$. So $d_B(u) \geq d_A(u)$ for all $u \in X'$ and similarly $d_A(v) \geq d_B(v)$ for all $v \in Y'$. This completes the proof of (ii).

To prove (i), note that $A \cap X \supseteq X'$ and $B \cap Y \supseteq Y'$, so $|A \triangle X| + |B \triangle Y| \leq |X_0| + |Y_0| \leq \sqrt{\rho}n$. So Lemma 3.4.10(i) with $\rho, \sqrt{\rho}, \nu, \tau, X, Y, A, B$ playing the roles of $\rho, \gamma, \nu, \tau, A, B, A', B'$ implies that U is a bipartite $(3\sqrt{\rho}, \nu/2, 2\tau)$-robust component with bipartition A, B. This completes the proof of (i). □

3.4.5 Proof of Theorem B

We are now ready to prove Theorem B – that every sufficiently large dense regular graph has a robust partition. The first part of the proof is an iteration of Lemmas 3.4.6 and 3.4.7 – we begin with the trivial partition of $V(G)$ and successively refine it by applying Lemma 3.4.7 to those components which are close to bipartite and Lemma 3.4.6 to the

others, until we obtain a partition into robust components. We then use Lemma 3.4.11 to adjust the partition slightly and Lemma 3.4.12 to achieve an appropriate bipartition of the bipartite robust expander components.

Proof of Theorem B. Let $t := 3\lceil 2/\alpha \rceil$. Define further constants satisfying

$$0 < 1/n_0 \ll \rho_1 \ll \nu_1 \ll \rho_2 \ll \nu_2 \ll \ldots \ll \rho_t \ll \nu_t \ll \tau' \ll \alpha, \tau$$

so that $1/n_0 \leq f(\rho_1)$ and $3^{3/2}\rho_i^{1/6} \leq f(\nu_i/4)$ for all $1 \leq i \leq t$. We first prove the following claim.

Claim. *There is some $1 \leq i < t$ and a partition $\mathcal{U}$ of $V(G)$ such that U is a (ρ_i, ν_i, τ')-robust component for each $U \in \mathcal{U}$.*

Proof. To see this, let $\mathcal{U}_1 := \{V(G)\}$. Note that $V(G)$ is certainly a ρ_1-component of G, and $|\mathcal{U}_1| = 1$. Suppose, for some i with $1 \leq i < t$, we have inductively defined a partition $\mathcal{U}_i$ of $V(G)$ such that U is a ρ_i-component for each $U \in \mathcal{U}_i$ and $2|\mathcal{U}_i| + |\mathcal{W}_i| \geq i + 1$, where $\mathcal{W}_i$ is the collection of all those $U \in \mathcal{U}_i$ which are ρ_i-close to bipartite. If each $U \in \mathcal{U}_i$ is a (ρ_i, ν_i, τ')-robust component, then we are done by setting $\mathcal{U} := \mathcal{U}_i$. Otherwise, we obtain $\mathcal{U}_{i+1}$ from $\mathcal{U}_i$ as follows.

There is some $U \in \mathcal{U}_i$ which is not a (ρ_i, ν_i, τ')-robust component. If $U \in \mathcal{W}_i$, then apply Lemma 3.4.7 with $\rho_i, \nu_i, \rho_{i+1}, \tau'$ playing the roles of ρ, ν, ρ', τ to obtain a partition U_1, U_2 of U such that U_1, U_2 are ρ_{i+1}-components. Let $\mathcal{U}_{i+1} := (\mathcal{U}_i \setminus \{U\}) \cup \{U_1, U_2\}$. Lemma 3.4.1(v) implies that $\mathcal{W}_i \setminus \{U\} \subseteq \mathcal{W}_{i+1}$, where $\mathcal{W}_{i+1}$ is the collection of all those $U \in \mathcal{U}_{i+1}$ which are ρ_{i+1}-close to bipartite. Thus $|\mathcal{U}_{i+1}| = |\mathcal{U}_i| + 1$ and $|\mathcal{W}_{i+1}| \geq |\mathcal{W}_i| - 1$.

So suppose next that $U \in \mathcal{U}_i \setminus \mathcal{W}_i$. Apply Lemma 3.4.6 with $\rho_i, \nu_i, \rho_{i+1}, \tau'$ playing the roles of ρ, ν, ρ', τ. If Lemma 3.4.6(i) holds, then U has a partition U_1, U_2 such that U_1, U_2 are ρ_{i+1}-components. As before, we let $\mathcal{U}_{i+1} := (\mathcal{U}_i \setminus \{U\}) \cup \{U_1, U_2\}$. So $|\mathcal{U}_{i+1}| = |\mathcal{U}_i| + 1$ and $|\mathcal{W}_{i+1}| \geq |\mathcal{W}_i|$. Otherwise, Lemma 3.4.6(ii) holds. Then U is ρ_{i+1}-close to bipartite. We let $\mathcal{U}_{i+1} := \mathcal{U}_i$. Then $|\mathcal{U}_{i+1}| = |\mathcal{U}_i|$

and $|\mathcal{W}_{i+1}| \geq |\mathcal{W}_i| + 1$.

Note that in each case we have $2|\mathcal{U}_{i+1}|+|\mathcal{W}_{i+1}| \geq i+2$. Moreover, Lemma 3.4.1(iii) implies that each $W \in \mathcal{U}_i \setminus \{U\}$ is a ρ_{i+1}-component. Therefore each $W \in \mathcal{U}_{i+1}$ is a ρ_{i+1}-component.

It remains to show that this process must stop before we define $\mathcal{U}_t$. Suppose not, i.e. suppose we have defined $\mathcal{U}_t$. Since each $W \in \mathcal{U}_t$ is a ρ_t-component, Lemma 3.4.1(i) implies that $|W| \geq (\alpha - \sqrt{\rho_t})n$ for all $W \in \mathcal{U}_t$. Moreover, $|\mathcal{U}_t| > t/3$ since $3|\mathcal{U}_t| \geq 2|\mathcal{U}_t| + |\mathcal{W}_t| \geq t+1$. Altogether, this implies that

$$|V(G)| \geq \frac{t}{3}(\alpha - \sqrt{\rho_t})n \geq \frac{2}{\alpha}(\alpha - \sqrt{\rho_t})n > n,$$

a contradiction. This completes the proof of the claim. ■

Set $\rho' := \rho_i$, $\nu' := \nu_i$, $\rho := 3^{3/2}\rho'^{1/6}$ and $\nu := \nu'/4$. So

$$\rho = 3^{3/2}\rho'^{1/6} \leq f(\nu'/4) = f(\nu) \quad \text{and} \quad 1/n_0 \leq f(\rho_1) \leq f(\rho) \tag{3.4.20}$$

and every $U \in \mathcal{U}$ is a (ρ', ν', τ')-robust component of G. So there exist $k, \ell \in \mathbb{N}$ such that $\mathcal{U} = \{U_1, \ldots, U_k, Z_1, \ldots, Z_\ell\}$, where U_i is a (ρ', ν', τ')-robust expander component for all $1 \leq i \leq k$, and Z_j is a bipartite (ρ', ν', τ')-robust expander component for all $1 \leq j \leq \ell$. Let $m := k+\ell$. Note that for each $1 \leq i \leq k$, we have $|U_i| \geq D - \sqrt{\rho'}n$ (by Lemma 3.4.1(i) and since U_i is a ρ'-component). Moreover, for each $1 \leq j \leq \ell$, $|Z_j| \geq 2(D - 2\sqrt{\rho'}n)$ by Lemma 3.4.1(iv). Thus

$$n = \sum_{1 \leq i \leq k} |U_i| + \sum_{1 \leq j \leq \ell} |Z_j| \geq (D - 2\sqrt{\rho'}n)(k + 2\ell)$$

and so

$$k + 2\ell \leq \left\lfloor \frac{n}{D - 2\sqrt{\rho'}n} \right\rfloor \leq \left\lfloor (1 + \rho'^{1/3})\frac{n}{D} \right\rfloor. \tag{3.4.21}$$

In particular, $1/m \geq \alpha/2$.

To achieve (D4), we apply Lemma 3.4.11 with ρ' playing the role of ρ to $\mathcal{U}$ to obtain a new partition $\mathcal{V} = \{V_1, \dots, V_k, W_1, \dots, W_\ell\}$ of $V(G)$ satisfying (i)–(iv), so in particular

$$|U_i \triangle V_i|, |Z_j \triangle W_j| \leq \rho'^{1/3} n \tag{3.4.22}$$

for all $1 \leq i \leq k$ and all $1 \leq j \leq \ell$. We claim that $\mathcal{V}$ satisfies (D1)–(D7).

Now (D1) certainly holds, (D4) follows from Lemma 3.4.11(iii) and (D7) follows from Lemma 3.4.11(iv). To prove (D2), note that V_i is a $\rho'^{1/3}$-component by Lemma 3.4.11(ii) and $|V_i| \geq D/2 \geq \sqrt{\rho} n$. Thus V_i is a ρ-component, i.e. (E1) holds. Now, by (3.4.22) and Lemma 3.4.8 with ν', τ', U_i, V_i playing the roles of ν, τ, U, U', we have that $G[V_i]$ is a robust $(\nu'/2, 2\tau')$-expander and thus also a robust (ν, τ)-expander. So (E2) holds, proving (D2).

To check (D3), recall that Z_j is a bipartite (ρ', ν', τ')-robust expander component. Then (3.4.22) and Lemma 3.4.10(ii) applied with $\rho', \rho'^{1/3}, \nu', \tau', Z_j, W_j$ playing the roles of $\rho, \gamma, \nu, \tau, U, U'$ imply that W_j is a bipartite $(3\rho'^{1/3}, \nu'/2, 2\tau')$-robust expander component. Now for each $1 \leq j \leq \ell$, apply Lemma 3.4.12 to W_j with $3\rho'^{1/3}, \nu'/2, 2\tau'$ playing the roles of ρ, ν, τ to obtain a bipartition A_j, B_j of W_j satisfying (i) and (ii). Lemma 3.4.12(i) implies that W_j is a bipartite (ρ, ν, τ)-robust expander component with bipartition A_j, B_j. So (D3) holds. Lemma 3.4.12(ii) implies that (D5) holds. Finally, (D6) follows from (3.4.21). □

3.5 Extending Theorem B to almost regular graphs

In this section, we prove an extension of Theorem B which states that every dense almost regular graph has a robust partition. We first extend the definition of a robust partition to graphs which may not be regular. Let $k, \ell, D \in \mathbb{N}$ and $0 < \rho \leq \nu \leq \tau < 1$. Given a graph G on n vertices, we say that $\mathcal{V}$ is a *robust partition of G with parameters ρ, ν, τ, k, ℓ* if (D1)–(D7) hold with $\delta(G)$ playing the role of D. Note that, for D-regular graphs, this

coincides with the definition given in Section 3.3.

Theorem 3.5.1. *For all $\alpha, \tau > 0$ and every non-decreasing function $f : (0,1) \to (0,1)$, there exist $n_0 \in \mathbb{N}$ and $\gamma > 0$ such that the following holds. For all graphs G on $n \geq n_0$ vertices with $\alpha n \leq \delta(G) \leq \Delta(G) \leq \delta(G) + \gamma n$, there exist ρ, ν with $1/n_0, \gamma \leq \rho \leq \nu \leq \tau$; $\rho \leq f(\nu)$ and $1/n_0 \leq f(\rho)$, and $k, \ell \in \mathbb{N}$ such that G has a robust partition $\mathcal{V}$ with parameters ρ, ν, τ, k, ℓ.*

The proof proceeds by taking two copies of G and adding a small number of edges between them to obtain a regular graph G', whose degree is only slightly higher than $\Delta(G)$. We apply Theorem B to obtain a robust partition $\mathcal{V}$ of G'. The construction of G' implies that every robust component in $\mathcal{V}$ lies entirely in one copy of G. So there is a partition of $\mathcal{V}$ into two parts, one of which must be a robust partition of G. It seems highly likely that one can prove Theorem 3.5.1 directly by adapting the proof of Theorem B, although we did not attempt this.

In order to construct G' from G, we need some preliminaries. We say that a non-decreasing sequence $(d_i)_{1 \leq i \leq n}$ of positive integers is *bipartite graphic* if there exists a bipartite graph G with vertex classes A and B with $|A| = |B| = n$ such that the ith vertex of each of A and B has degree d_i. The following theorem of Alon, Ben-Shimon and Krivelevich [4] gives a sufficient condition for a sequence to be bipartite graphic. (Note that their original statement was different, but the two forms are equivalent, as observed in [27].)

Theorem 3.5.2. *Suppose that $(d_i)_{1 \leq i \leq n}$ is a non-decreasing sequence of positive integers. Then $(d_i)_{1 \leq i \leq n}$ is bipartite graphic if $nd_1 \geq (d_1 + d_n)^2/4$.*

We also need the following result (Lemma 3.8 from [82]).

Lemma 3.5.3. *Suppose that $0 < \nu \leq \tau \leq \varepsilon < 1$ are such that $\varepsilon \geq 2\nu/\tau$. Let G be a graph on n vertices with minimum degree $\delta(G) \geq (1/2 + \varepsilon)n$. Then G is a robust (ν, τ)-expander.*

We are now able to deduce Theorem 3.5.1 from Theorem B.

Proof that Theorem B implies Theorem 3.5.1. Define $f' : (0,1) \to (0,1)$ by $f'(x) := \min\{f(x)/4, \alpha x/2\}$ and let $\tau' := \min\{\tau, \alpha^2/20\}$. Apply Theorem B with α, τ', f' playing the roles of α, τ, f to obtain $n_0 \in \mathbb{N}$. Let $\gamma := 1/4n_0$. Let G be a graph on $n \geq n_0$ vertices with $\alpha n \leq \delta(G) \leq \Delta(G) \leq \delta(G) + \gamma n$. Let $D := \delta(G)$. Order the vertices $v_1, \ldots, v_n$ of G in order of increasing degree.

Obtain a graph G'' from G as follows. We let $W_1 := \{w_1, \ldots, w_n\}$ and $W_2 := \{x_1, \ldots, x_n\}$ be disjoint sets of vertices and let G'' have vertex set $W_1 \cup W_2$. We add the edges $w_i w_j$ and $x_i x_j$ whenever $v_i v_j \in E(G)$.

Choose a constant β such that $\gamma = \beta(1-\beta)$ and $\gamma \leq \beta \leq 2\gamma$. Let $d_i := D + \beta n - d_G(v_{n+1-i})$. Then $(d_i)_{1 \leq i \leq n}$ is a non-decreasing sequence and $(\beta - \gamma)n \leq d_1 \leq d_n \leq \beta n$. Observe that if $(d_i)_{1 \leq i \leq n}$ is bipartite graphic, then we can add edges to G'' between W_1 and W_2 to obtain a $(D + \beta n)$-regular graph G'. Since $(d_1 + d_n)^2/4 \leq \beta^2 n^2 = (\beta - \gamma)n^2 \leq n d_1$, Theorem 3.5.2 implies that such a G' exists. Note that

$$\Delta(G'[W_1, W_2]) = d_n \leq \beta n. \tag{3.5.1}$$

Theorem B applied to G' implies that there exist ρ', ν with $1/n_0 \leq \rho' \leq \nu \leq \tau'$; $\rho' \leq f'(\nu)$ and $1/n_0 \leq f'(\rho')$, and $k', \ell' \in \mathbb{N}$ such that G' has a robust partition $\mathcal{V}$ with parameters $\rho', \nu, \tau', k', \ell'$. Note that $\beta \leq 2\gamma = 1/2n_0 \leq \nu/2$.

Claim. *Let $U \in \mathcal{V}$ be arbitrary. Then U is contained entirely within one of W_1, W_2.*

Proof. Let $U_i := U \cap W_i$ for $i = 1, 2$. Assume, for a contradiction, that $U_1, U_2 \neq \emptyset$. Then

$$\begin{aligned} |U_i| \geq \delta(G'[U_i]) &\overset{\text{(D4),(3.5.1)}}{\geq} \frac{D}{k' + \ell'} - \beta n \overset{\text{(D6)}}{\geq} \frac{D}{2(1+\rho^{1/3})n/D} - \beta n \\ &\geq (\alpha^2/4 - \beta)n \geq \alpha^2 n/5. \end{aligned} \tag{3.5.2}$$

In particular this implies that $\tau'|U| \leq |U_i| \leq (1-\tau')|U|$. The fact that $\beta \leq \nu/2$ and (3.5.1) imply that $RN_{\nu,U}(U_i) \subseteq U_i$. Then U cannot be a robust expander component. So U is a bipartite robust expander component, with bipartition A, B, say. Let $A_i := A \cap U_i$ for $i = 1, 2$ and define B_i analogously. Similarly as in (3.5.2), using (D5) instead of (D4), one can show that $|A_i|, |B_i| \geq (\alpha^2/8 - \beta)n \geq \alpha^2 n/10$. In particular, $\tau'|A| \leq |A_i| \leq (1-\tau')|A|$. Without loss of generality, suppose that $|A_1| - |B_1| \geq |A_2| - |B_2|$. Then (C2) implies that $|A_1| - |B_1| \geq -\rho' n$ and so $|RN_{\nu,U}(A_1) \cap B| \leq |B_1| \leq |A_1| + \rho' n < |A_1| + \nu|U|$, a contradiction. (Here we used the fact that $|U| > \alpha n/2$ and $\rho' \leq f'(\nu)$.) This completes the proof of the claim. ■

So there is a partition $\mathcal{V}_1, \mathcal{V}_2$ of $\mathcal{V}$ such that $U \subseteq W_i$ for all $U \in \mathcal{V}_i$. For $i = 1, 2$, let k_i be the number of robust expander components and ℓ_i the number of bipartite robust expander components in $\mathcal{V}_i$. Let $\rho := 4\rho'$. We claim that, for at least one of $i = 1, 2$, we have that $\mathcal{V}_i$ is a robust partition of G with parameters $\rho, \nu, \tau', k_i, \ell_i$. Suppose that, for both $i = 1, 2$, we have $k_i + 2\ell_i > \lfloor (1+\rho^{1/3})n/D \rfloor$. Then

$$k' + 2\ell' \geq 2\left\lfloor \frac{(1+\rho^{1/3})n}{D} \right\rfloor + 2 > \left\lfloor \frac{2(1+\rho^{1/3})n}{D} \right\rfloor \geq \left\lfloor \frac{2(1+\rho'^{1/3})n}{D+\beta n} \right\rfloor,$$

contradicting (D6) for $\mathcal{V}$. So without loss of generality, we have that $\mathcal{V}_1$ satisfies (D6). It is easy to check that the remaining properties (D1)–(D5) and (D7) are also satisfied by $\mathcal{V}_1$. Therefore $\mathcal{V}_1$ is a robust partition of G with parameters $\rho, \nu, \tau', k_1, \ell_1$ and hence also with parameters $\rho, \nu, \tau, k_1, \ell_1$. □

3.6 How to obtain a long cycle given a robust partition

The main result of this section is Lemma 3.6.2 which implies that, given a suitable set $\mathcal{P}$ of paths joining up the robust components of a robust partition, one can extend $\mathcal{P}$ into

a Hamilton cycle. Actually, in the proof of Theorem 3.1.4 we will need to consider the more general notion of a weak robust subpartition, defined below.

3.6.1 Definitions and the main statement

Let $k, \ell \in \mathbb{N}$ and $0 < \rho \leq \nu \leq \tau \leq \eta < 1$. Given a graph G on n vertices, we say that $\mathcal{U}$ is a *weak robust subpartition in G with parameters $\rho, \nu, \tau, \eta, k, \ell$* if the following conditions hold.

(D1′) $\mathcal{U} = \{U_1, \ldots, U_k, Z_1, \ldots, Z_\ell\}$ is a collection of disjoint subsets of $V(G)$;

(D2′) for all $1 \leq i \leq k$, U_i is a (ρ, ν, τ)-robust expander component of G;

(D3′) for all $1 \leq j \leq \ell$, there exists a partition A_j, B_j of Z_j such that Z_j is a bipartite (ρ, ν, τ)-robust expander component with respect to A_j, B_j;

(D4′) $\delta(G[X]) \geq \eta n$ for all $X \in \mathcal{U}$;

(D5′) for all $1 \leq j \leq \ell$, we have $\delta(G[A_j, B_j]) \geq \eta n/2$.

A weak robust subpartition $\mathcal{U}$ is weaker than a robust partition in the sense that a non-regular graph can have a weak robust (sub)partition, $\mathcal{U}$ need not involve the entire graph, and we can make small adjustments to the partition while still maintaining (D1′)–(D5′) with slightly worse parameters. This is formalised by the following statement.

Proposition 3.6.1. *Let $k, \ell, D \in \mathbb{N}$ and suppose that $0 < 1/n \ll \rho \leq \nu \leq \tau \leq \eta \leq \alpha^2/2 < 1$.*

(i) Suppose that G is a D-regular graph on n vertices where $D \geq \alpha n$. Let $\mathcal{V}$ be a robust partition of G with parameters ρ, ν, τ, k, ℓ. Then $\mathcal{V}$ is a weak robust subpartition in G with parameters $\rho, \nu, \tau, \eta, k, \ell$.

(ii) Suppose that H is a graph and $\mathcal{U}$ is a weak robust subpartition in H with parameters $\rho, \nu, \tau, \eta, k, \ell$. Let $\mathcal{U}' \subseteq \mathcal{U}$ be non-empty. Then $\mathcal{U}'$ is a weak robust subpartition in H with parameters $\rho, \nu, \tau, \eta, k', \ell'$ for some $k' \leq k$ and $\ell' \leq \ell$.

Proof. We only prove (i) since (ii) is clear. Note that properties (D1′)–(D3′) are immediate. Note (D6) implies that $k + 2\ell \leq \lfloor (1 + \rho^{1/3})/\alpha \rfloor \leq 2/\alpha$. So $D/(k + \ell) \geq \alpha^2 n/2$. Together with (D4) and (D5) this shows that (D4′) and (D5′) hold. This completes the proof. □

For a path system $\mathcal{P}$, we say that a vertex x is an *endpoint* of $\mathcal{P}$ if x is an endpoint of some path in $\mathcal{P}$. Define the *internal vertices* of $\mathcal{P}$ similarly. If every endpoint of a path system $\mathcal{P}$ lies in some $U \subseteq V(G)$, we say that $\mathcal{P}$ is *U-anchored*. When $\mathcal{U}$ is a collection of disjoint subsets of $V(G)$, we say that $\mathcal{P}$ is *$\mathcal{U}$-anchored* if it is $\bigcup_{U \in \mathcal{U}} U$-anchored. Given a path P in G, we say that P' is an *extension* of P if P' is a path which contains P as a subpath. An *Euler tour* in a (multi)graph is a closed walk that visits every vertex and uses each edge exactly once.

Given a graph G with $U \subseteq V(G)$ and a path system $\mathcal{P}$ in G, we write $\mathrm{End}_{\mathcal{P}}(U)$ and $\mathrm{Int}_{\mathcal{P}}(U)$ for, respectively, the number of endpoints/internal vertices of $\mathcal{P}$ which lie in U. Given disjoint sets $A, B \subseteq V(G)$, we say that $\mathcal{P}$ is *(A, B)-balanced* if

- $\mathrm{End}_{\mathcal{P}}(A) = \mathrm{End}_{\mathcal{P}}(B) > 0$; and
- $|A| - \mathrm{Int}_{\mathcal{P}}(A) = |B| - \mathrm{Int}_{\mathcal{P}}(B)$.

Suppose that G is a graph and $\mathcal{U}$ is a collection of disjoint subsets of $V(G)$. Let $\mathcal{P}$ be a $\mathcal{U}$-anchored path system in G (so all endpoints of the paths in $\mathcal{P}$ lie in $\bigcup_{U \in \mathcal{U}} U$). We define the *reduced multigraph $R_{\mathcal{U}}(\mathcal{P})$ of $\mathcal{P}$ with respect to $\mathcal{U}$* to be the multigraph with vertex set $\mathcal{U}$ in which each edge between U and U' corresponds to a path in $\mathcal{P}$ with one endpoint in U and one endpoint in U'. So $R_{\mathcal{U}}(\mathcal{P})$ might contain loops.

Let $k, \ell \in \mathbb{N}$, let $0 < \rho \leq \nu \leq \tau \leq \eta < 1$ and let $0 < \gamma < 1$. Suppose that G is a graph on n vertices with a weak robust subpartition $\mathcal{U} = \{U_1, \ldots, U_k, Z_1, \ldots, Z_\ell\}$ with parameters $\rho, \nu, \tau, \eta, k, \ell$, so that the bipartition of Z_j specified by (D3′) is A_j, B_j. We say that $\mathcal{P}$ is a *$\mathcal{U}$-tour with parameter γ* if

(T1) $\mathcal{P}$ is a $\mathcal{U}$-anchored path system;

(T2) $R_{\mathcal{U}}(\mathcal{P})$ has an Euler tour;

(T3) for all $U \in \mathcal{U}$ we have $|V(\mathcal{P}) \cap U| \leq \gamma n$;

(T4) for all $1 \leq j \leq \ell$, $\mathcal{P}$ is (A_j, B_j)-balanced.

We will often think of $R_{\mathcal{U}}(\mathcal{P})$ as a walk rather than a multigraph. So in particular, we will often say that '$R_{\mathcal{U}}(\mathcal{P})$ is an Euler tour'. The aim of this section is to prove the following lemma, stating that every graph with a weak robust subpartition $\mathcal{U}$ and a $\mathcal{U}$-tour contains a cycle which covers every vertex within the components of $\mathcal{U}$.

Lemma 3.6.2. *Let $k, \ell, n \in \mathbb{N}$ and suppose that $0 < 1/n \ll \rho, \gamma \ll \nu \leq \tau \ll \eta < 1$. Suppose that G is a graph on n vertices and that $\mathcal{U}$ is a weak robust subpartition in G with parameters $\rho, \nu, \tau, \eta, k, \ell$. Suppose further that G contains a $\mathcal{U}$-tour $\mathcal{P}$ with parameter γ. Then there is a cycle in G which contains $\mathcal{P}$ and every vertex in $\bigcup_{U \in \mathcal{U}} U$.*

Since by Proposition 3.6.1(i) every robust partition is also a weak robust subpartition, Lemma 3.6.2 immediately implies the following result which will be used in the proof of Theorem 3.1.2 while for the proof of Theorem 3.1.4 we will need Lemma 3.6.2 itself.

Corollary 3.6.3. *Let $k, \ell, n, D \in \mathbb{N}$ and suppose that $0 < 1/n \ll \rho, \gamma \ll \nu \leq \tau \ll \alpha < 1$. Suppose that G is a D-regular graph on n vertices where $D \geq \alpha n$, with a robust partition $\mathcal{V}$ with parameters ρ, ν, τ, k, ℓ. Suppose further that G contains a $\mathcal{V}$-tour with parameter γ. Then G contains a Hamilton cycle.*

The remainder of this section is devoted to the proof of Lemma 3.6.2.

3.6.2 Spanning path systems in robust expanders

In this subsection, we prove Corollary 3.6.8, which states that when p is not too large, every robust expander G is Hamilton p-linked, i.e. given distinct vertices $y_1, y_1', \ldots, y_p, y_p'$, there exist p vertex-disjoint paths joining y_i to y_i' for all $i \leq p$ such that together these paths cover all the vertices of G. This, combined with a bipartite analogue in the next

subsection, will be the main tool in proving Lemma 3.6.2: the y_i and y_i' will be suitable endpoints of the paths in the $\mathcal{U}$-tour $\mathcal{P}$.

We now recall an analogue of robust expansion for digraphs (see Section 2.1). Let $0 < \nu \leq \tau < 1$. Given any digraph G on n vertices and $S \subseteq V(G)$, the *ν-robust outneighbourhood* $RN^+_{\nu,G}(S)$ of S is the set of all those vertices of G which have at least νn inneighbours in S. G is called a *robust (ν,τ)-outexpander* if $|RN^+_{\nu,G}(S)| \geq |S| + \nu n$ for all $S \subseteq V(G)$ with $\tau n \leq |S| \leq (1-\tau)n$.

The next lemma is a directed analogue of Lemma 3.4.8. Its proof follows immediately from the definition.

Lemma 3.6.4. *Let $0 < \nu \ll \tau \ll 1$. Suppose that G is a digraph and $U \subseteq W \subseteq V(G)$ are such that $G[U]$ is a robust (ν,τ)-outexpander and $|U \setminus W| \leq \nu|U|/2$. Then $G[W]$ is a robust $(\nu/2, 2\tau)$-outexpander.*

The next lemma shows that the diameter of a robust outexpander is small.

Lemma 3.6.5. *Let $n \in \mathbb{N}$ and $0 < 1/n \ll \nu \ll \tau \ll \eta \leq 1$. Suppose that G is a robust (ν,τ)-outexpander on n vertices with $\delta^0(G) \geq \eta n$. Then, given any distinct vertices $x, y \in V(G)$, there exists a path P in G from x to y such that $|V(P)| \leq 1/\nu$.*

Proof. Let X_i be the set of vertices v for which there is a directed walk from x to v in G of length at most i. So $X_0 = \{x\}$ and $X_1 = N^+(x) \cup \{x\}$. So $|X_1| \geq \eta n$. Note that $RN^+_{\nu,G}(X_i) \subseteq X_{i+1}$. Therefore, if $|X_i| \leq (1-\tau)n$, then $|X_{i+1}| \geq |RN^+_{\nu,G}(X_i)| \geq |X_i| + \nu n$. So certainly for $i' := \lfloor 1/\nu \rfloor - 2$ we have that $|X_{i'}| \geq (1-\tau)n$. But since $\delta^0(G) \geq \eta n \geq \tau n$ we have that $X_{i'+1} = V(G)$. In particular, this implies that for any $y \neq x$ there is a path P of length at most $1/\nu - 1$ between x and y in G. Therefore $|V(P)| \leq 1/\nu$. $\square$

We will need the following result of Kühn, Osthus and Treglown [86], which states that a robust outexpander whose minimum degree is not too small contains a (directed) Hamilton cycle.

Theorem 3.6.6 ([86]). *Let $n \in \mathbb{N}$ and suppose that $0 < 1/n \ll \nu \leq \tau \ll \eta < 1$. Let G be a robust (ν, τ)-outexpander on n vertices with $\delta^0(G) \geq \eta n$. Then G contains a Hamilton cycle.*

We say that a digraph G is *p-ordered Hamilton* if, given $x_1, \ldots, x_p \in V(G)$, G contains a Hamilton cycle which traverses $x_1, \ldots, x_p$ in this order.

Corollary 3.6.7. *Let $n, p \in \mathbb{N}$ and suppose that $0 < 1/n \ll \nu \ll \tau \ll \eta < 1$ and $p \leq \nu^3 n$. Let G be a robust (ν, τ)-outexpander on n vertices with $\delta^0(G) \geq \eta n$. Then G is p-ordered Hamilton.*

Proof. Let $x_1, \ldots, x_p \in V(G)$. We claim that we can find a path P in G joining x_1, x_p which traverses $x_1, \ldots, x_p$ in this order and such that $|V(P)| \leq \nu n/2$. To see this, suppose for some $i \leq p-1$ we have found a path P_i joining x_1, x_i with $|V(P_i)| \leq 2i/\nu$ which traverses $x_1, \ldots, x_i$ in this order and such that $x_{i+1}, \ldots, x_p$ do not lie in P_i. Let $G_i := G \setminus ((V(P_i) \setminus \{x_i\}) \cup \{x_{i+2}, \ldots, x_p\})$. Note that $n - |V(G_i)| \leq 2p/\nu \leq \nu n/2$. So Lemma 3.6.4 implies that G_i is a robust $(\nu/2, 2\tau)$-outexpander. Apply Lemma 3.6.5 with G_i, x_i, x_{i+1} playing the roles of G, x, y to obtain a path, which, when appended to P_i, gives a path P_{i+1} joining x_1, x_{i+1} which traverses $x_1, \ldots, x_{i+1}$ in this order such that $x_{i+2}, \ldots, x_p$ do not lie in P_{i+1} and $|V(P_{i+1})| \leq |V(P_i)| + 2/\nu \leq 2(i+1)/\nu$. Set $P := P_p$. This proves the claim.

Let G' be the graph obtained from $G \setminus V(P)$ by adding a new vertex z such that $N^-_{G'}(z) := N^-_{G \setminus V(P)}(x_1)$ and $N^+_{G'}(z) := N^+_{G \setminus V(P)}(x_p)$. Then $\delta^0(G') \geq \delta^0(G) - \nu n/2 \geq \eta |G'|/2$ and G' is a robust $(\nu/2, 2\tau)$-outexpander. Therefore we can apply Theorem 3.6.6 to find a directed Hamilton cycle in G'. This corresponds to a Hamilton cycle in G which traverses $x_1, \ldots, x_p$ in this order. $\square$

The following corollary states that robust (out)expanders are Hamilton p-linked provided that p is not too large.

Corollary 3.6.8. *Let $n, p \in \mathbb{N}$ and suppose that $0 < 1/n \ll \nu \ll \tau \ll \eta < 1$ and let $p \leq \nu^4 n$.*

(i) Let G be a robust (ν, τ)-outexpander on n vertices with $\delta^0(G) \geq \eta n$. Then G is Hamilton p-linked.

(ii) Let H be a robust (ν, τ)-expander on n vertices with $\delta(H) \geq \eta n$. Then H is Hamilton p-linked.

Proof. To prove (i), let $y_1, \ldots, y_p, y'_1, \ldots, y'_p \in V(G)$. Obtain G^* from G as follows. For each $1 \leq i \leq p$ (where indices are considered modulo p), replace the pair y_{i+1}, y'_i with a new vertex z_i such that $N^+_{G^*}(z_i) := N^+_G(y_{i+1})$ and $N^-_{G^*}(z_i) := N^-_G(y'_i)$. Then it is easy to see that G^* is a robust $(\nu/2, 2\tau)$-outexpander. Corollary 3.6.7 implies that G^* contains a Hamilton cycle which traverses $z_1, \ldots, z_p$ in this order. This corresponds to a collection $P_1, \ldots, P_p$ of vertex-disjoint paths such that P_i joins y_i to y'_i and all the P_i together cover $V(G)$, proving (i).

To prove (ii), let G be the digraph obtained from H by replacing each edge xy with directed edges $\overrightarrow{xy}$ and $\overrightarrow{yx}$. Then G is a robust (ν, τ)-outexpander with $\delta^0(G) \geq \eta n$. Now (i) implies that G is Hamilton p-linked. For each $xy \in E(H)$, any path system in G uses at most one of $\overrightarrow{xy}, \overrightarrow{yx}$. So H is Hamilton p-linked. □

3.6.3 Spanning path systems in bipartite robust expanders

Given $p \in \mathbb{N}$ and a bipartite graph G with vertex classes A, B, we say that G is (A, B)-*Hamilton p-linked* if, given any $Y := \{y_1, y'_1, y_2, y'_2, \ldots, y_p, y'_p\} \subseteq V(G)$ with $|Y \cap A| = |Y \cap B| = p$, we can find a set of vertex-disjoint paths joining y_i to y'_i in G such that together these paths cover all the vertices of G. Note that if G is (A, B)-Hamilton p-linked then it is balanced. In this subsection we show that, for p not too large, G is (A, B)-Hamilton p-linked when G is a balanced bipartite robust expander.

Given a balanced bipartite graph G with vertex classes A, B which contains a perfect matching M, we denote by G^* the *M-auxiliary digraph of G* obtained from G as follows. Let G^* have vertex set B. For each $v \in B$, we let v' be the unique vertex of A such that $vv' \in M$. Then, for all $x, v \in B$, we let $\overrightarrow{vx} \in E(G^*)$ if and only if $x \in N_G(v') \setminus \{v\}$. Note

that the order of A and B matters here.

Lemma 3.6.9. *Let $n \in \mathbb{N}$ and $0 < 1/n \ll \nu \ll \tau \ll \eta < 1$. Let G be a balanced bipartite graph with vertex classes A, B so that $|A| = |B| = n$ and $\delta(G) \geq \eta n$. Suppose further that G is a bipartite robust (ν, τ)-expander (with bipartition A, B). Then*

(i) G contains a perfect matching M;

(ii) the M-auxiliary digraph G^* of G is a robust (ν, τ)-outexpander with minimum degree at least $\eta n/2$.

Proof. Observe that (i) follows immediately from Hall's Theorem. Indeed, by Hall's Theorem, it suffices to show that whenever S is a proper subset of A, we have that $|N_G(S)| \geq |S|$. Suppose first that $|S| \leq \tau n$. Then $|N_G(S)| \geq \delta(G) \geq \eta n \geq \tau n \geq |S|$. Suppose instead that $\tau n \leq |S| \leq (1-\tau)n$. Then $|N_G(S)| \geq |RN_{\nu,G}(S)| \geq |S| + \nu n \geq |S|$. Finally, suppose that $|S| \geq (1-\tau)n$. Then $N_G(S) = B$ since $\tau \leq \eta$ and $\delta(G) \geq \eta n$. So certainly $|N_G(S)| \geq |S|$ in this case. Therefore G contains a perfect matching $M := \{xx' : x \in B, x' \in A\}$. To prove (ii), note that $\delta^0(G^*) \geq \delta(G) - 1 \geq \eta n/2$. Consider any $S \subseteq B$ with $\tau n \leq |S| \leq (1-\tau)n$. Let $S_A := \{x' : x \in S\}$ and note that $RN^+_{\nu,G^*}(S) \supseteq RN_{\nu,G}(S_A)$. Thus

$$|RN^+_{\nu,G^*}(S)| \geq |RN_{\nu,G}(S_A)| \geq |S_A| + \nu|V(G)| \geq |S| + \nu|V(G^*)|,$$

and therefore G^* is a robust (ν, τ)-outexpander, proving (ii). □

We now prove an analogue of Lemma 3.6.5 for bipartite robust expanders.

Lemma 3.6.10. *Let $n \in \mathbb{N}$ and $0 < 1/n \ll \nu \ll \tau \ll \eta < 1$. Suppose that G is a bipartite graph on n vertices with vertex classes A, B, where $\Big||A| - |B|\Big| \leq \nu^2 n$. Suppose further that $\delta(G) \geq \eta n$ and G is a bipartite robust (ν, τ)-expander (with bipartition A, B). Then, given any distinct vertices $x, y \in V(G)$ there exists a path P between x and y in G such that $|V(P)| \leq 4/\nu$.*

Proof. Consider each $u \in \{x, y\}$. If $u \in B$, let u' be a neighbour of u which lies in A. If $u \in A$, let $u' := u$. Make these choices so that x', y' are distinct. So $\{x', y'\} \subseteq A$. Remove at most $\Big||A| - |B|\Big| \leq \nu|A|/4$ vertices from $A \cup B$ to obtain $A' \subseteq A$ and $B' \subseteq B$ such that $|A'| = |B'|$ and $\{x', y'\} \subseteq A'$. Lemma 3.4.9(i) implies that $G' := G[A', B']$ is a bipartite robust $(\nu/2, 2\tau)$-expander and that $\delta(G') \geq \eta n'/2$ where $n' := |V(G')|$.

Let X_i be the set of vertices $v \in A'$ of distance at most $2i$ to x' in G'. Now Lemma 3.6.9(i) implies that G' contains a perfect matching M. So for all $i \geq 0$ we have $X_{i+1} \supseteq \{a \in A' : ab \in M, b \in N_{G'}(X_i)\}$. Thus $|X_1| \geq \eta n'/2$ and whenever $i \geq 1$ and $|X_i| < (1 - \tau)n$ then

$$|X_{i+1}| \geq |N_{G'}(X_i)| \geq |RN_{\nu/2,G'}(X_i)| \geq |X_i| + \nu n'/2.$$

So certainly for $i' := \lfloor 2/\nu \rfloor - 4$ we have that $|X_{i'}| \geq (1 - \tau)n'$. But since $\delta(G') \geq \eta n'/2 \geq \tau n'$ we have that $X_{i'+1} = A'$. In particular, this implies that there is a path of length at most $4/\nu - 5$ between x' and y' in G' and hence a path P with $|V(P)| \leq 4/\nu$ between x and y in G. □

The following is a bipartite analogue of Corollary 3.6.8. To prove it, we iterate Lemma 3.6.10 to find short paths between a small number of pairs of vertices. Then the graph obtained by deleting these paths is still a bipartite robust expander.

Lemma 3.6.11. *Let $n, p \in \mathbb{N}$, $0 < 1/n \ll \nu \ll \tau \ll \eta \leq 1$ and $p \leq \nu^4 n$. Suppose that G is a bipartite graph vertex classes A, B, so that $|A| = |B| = n$. Suppose further that G is a bipartite robust (ν, τ)-expander with $\delta(G) \geq \eta n$. Then G is (A, B)-Hamilton p-linked.*

Proof. Let $Y := \{y_1, y_1', y_2, y_2', \ldots, y_p, y_p'\}$ be a collection of distinct vertices in G such that $|Y \cap A| = |Y \cap B|$. For each $1 \leq i, j \leq p$, let $W_i := \{y_i, y_i'\}$ and let $W_{\geq j} := \bigcup_{j \leq i \leq p} W_i$. Suppose, for some $0 \leq \ell \leq p - 2$, we have already obtained vertex-disjoint paths $R_1, \ldots, R_\ell$, where for each $1 \leq i \leq \ell$, R_i has endpoints y_i, y_i' and $|V(R_i)| \leq 8/\nu$.

We obtain $R_{\ell+1}$ as follows. Let

$$G_\ell := G \setminus (V(R_1) \cup \ldots \cup V(R_{\ell-1}) \cup W_{\geq \ell+1})$$

and let $n_\ell := |V(G_\ell)|$. Note that

$$|V(G) \setminus V(G_\ell)| = \sum_{1 \leq i \leq \ell} |V(R_i)| + |W_{\geq \ell+1}| \leq 8p/\nu \leq \nu^2 n. \tag{3.6.1}$$

Let $A_\ell := A \cap V(G_\ell)$ and define B_ℓ analogously. Then Lemma 3.4.9(i) implies that G_ℓ is a bipartite robust $(\nu/2, 2\tau)$-expander with bipartition A_ℓ, B_ℓ, and $\delta(G_\ell) \geq \eta n_\ell/4$. Moreover $\Big||A_\ell| - |B_\ell|\Big| \leq |V(G) \setminus V(G_\ell)| \leq \nu^2 n \leq \nu^2 n_\ell$. Therefore we can apply Lemma 3.6.10 with $A_\ell, B_\ell, \nu/2, 2\tau, \eta/4$ playing the roles of A, B, ν, τ, η to see that G_ℓ contains a path $R_{\ell+1}$ between $y_{\ell+1}$ and $y'_{\ell+1}$ such that $|V(R_{\ell+1})| \leq 8/\nu$.

Therefore we can obtain vertex-disjoint paths $R_1, \ldots, R_{p-1}$ in $G \setminus \{y_p, y'_p\}$ such that $|V(R_i)| \leq 8/\nu$ and R_i joins y_i, y'_i for all $1 \leq i \leq p-1$. To obtain R_p, we now consider three cases depending on the classes in which y_p, y'_p lie. Let $V^* := \bigcup_{1 \leq i \leq p-1} V(R_i)$.

Case 1. *$y_p \in A$ and $y'_p \in B$.*

Using our assumption that $|Y \cap A| = |Y \cap B|$, it is easy to see that $|V^* \cap A| = |V^* \cap B|$. Let $G' := G \setminus \big(V^* \cup \{y_p, y'_p\}\big)$. Also let $A' := A \cap V(G')$ and define B' analogously. Then $|A'| = |B'| =: n'$. As above, G' is a bipartite robust $(\nu/2, 2\tau)$-expander with respect to A', B', and $\delta(G') \geq \eta(n'+1)/2$. Therefore G' contains a perfect matching M' by Lemma 3.6.9(i). Let $M'' := M' \cup \{y_p y'_p\}$. Then M'' is a perfect matching in the graph G^- obtained from $G \setminus V^*$ by adding the edge $y_p y'_p$ if necessary. Note that $|G^-| = 2(n'+1)$ and $\delta(G^-) \geq \eta(n'+1)/2$. Let G'' be the M''-auxiliary digraph of G^-. Then Lemma 3.6.9(ii) implies that G'' is a robust $(\nu/2, 2\tau)$-outexpander with minimum degree at least $\eta(n'+1)/4$. By Theorem 3.6.6, G'' contains a Hamilton cycle C. Then C corresponds to a Hamilton path R_p in $G \setminus V^*$ which joins y_p and y'_p. Thus $R_1, \ldots, R_p$ are vertex-disjoint from each other, join y_i to y'_i, and together cover all the vertices of G. So

G is (A, B)-Hamilton p-linked.

Case 2. $y_p, y'_p \in A$.

So it is easy to see that $|V^* \cap A| = |V^* \cap B| - 1$. Choose a neighbour z_p of y'_p in B which does not lie in V^*. Now delete y'_p from G and proceed as above with z_p playing the role of y'_p.

Case 3. $y_p, y'_p \in B$.

This is analogous to Case 2. □

3.6.4 Proof of Lemma 3.6.2

We are now ready to prove Lemma 3.6.2. Given a robust subpartition $\mathcal{U}$ in G and a $\mathcal{U}$-tour $\mathcal{P}$, we apply Corollary 3.6.8 within each robust expander component U of $\mathcal{U}$, with the endpoints of $\mathcal{P}$ which lie in U suitably ordered. Similarly, we apply Lemma 3.6.11 within each bipartite robust expander component Z of $\mathcal{U}$. In this way, we obtain a set $\mathcal{R}$ of 'joining paths'. Then together the paths in $\mathcal{P} \cup \mathcal{R}$ form a cycle containing every vertex of $\bigcup_{U \in \mathcal{U}} U$.

Proof of Lemma 3.6.2. Note that if $\nu' \leq \nu$, then any (bipartite) robust (ν, τ)-expander is also a (bipartite) robust (ν', τ)-expander. So without loss of generality, we may assume that $\nu \ll \tau$. Write $\mathcal{U} := \{U_1, \ldots, U_k, Z_1, \ldots, Z_\ell\}$ so that (D1′)–(D5′) are satisfied. Let $\mathcal{P}$ be a $\mathcal{U}$-tour with parameter γ, let $q := |\mathcal{P}|$ and $R := R_{\mathcal{U}}(\mathcal{P})$. So for each path $P \in \mathcal{P}$ there is a unique edge e_P in R. Without loss of generality, $e_{P_1} \ldots e_{P_q}$ is the Euler tour guaranteed by (T2). This corresponds to an ordering $P_1, \ldots, P_q$ of the paths in $\mathcal{P}$. Direct the edges of R so that $e_{P_1} \ldots e_{P_q}$ is a directed tour. Direct the edges of (the paths in) $\mathcal{P}$ correspondingly, so that for all $1 \leq s \leq q$, if e_{P_s} has startpoint U and endpoint W, then P_s is a directed path from some vertex $x_s^- \in U$ to some vertex $x_s^+ \in W$. We thus obtain an ordering $x_1^+, x_2^-, x_2^+, \ldots, x_q^-, x_q^+, x_1^-$ of the endpoints of $\mathcal{P}$. Note that for each $1 \leq i \leq q$, x_i^+, x_{i+1}^- lie in the same $X \in \mathcal{U}$, where the indices are considered modulo q.

Fix some $U \in \mathcal{U}$. Let $p := \mathrm{End}_{\mathcal{P}}(U)/2$. Thus $p \in \mathbb{N}$. Then there exists a subsequence $i_1, \ldots, i_p$ of $1, \ldots, q$ such that

$$K := (x^+_{i_1}, x^-_{i_1+1}, x^+_{i_2}, x^-_{i_2+1}, \ldots, x^+_{i_p}, x^-_{i_p+1})$$

is the subsequence of ordered endpoints of $\mathcal{P}$ which lie in U (where $x^-_{q+1} := x^-_1$). Let I be the (unordered) collection of internal vertices of $\mathcal{P}$ which lie in U. Let $U' := U \setminus I$. Note that each element of K lies in U'. Now (D4′) implies that $\delta(G[U]) \geq \eta n$. Furthermore, (T3) implies that $\mathrm{End}_{\mathcal{P}}(U) + \mathrm{Int}_{\mathcal{P}}(U) \leq \gamma n$. So

$$|U'| \overset{\text{(T3)}}{\geq} |U| - \gamma n \geq (\eta - \gamma)n \geq \eta n/2 \tag{3.6.2}$$

and hence

$$p = \mathrm{End}_{\mathcal{P}}(U)/2 \leq \gamma n/2 \leq \gamma |U'|/\eta \leq \sqrt{\gamma}|U'|; \tag{3.6.3}$$

$$\text{and} \quad |I| = \mathrm{Int}_{\mathcal{P}}(U) \leq 2\gamma n \leq 4\gamma |U'|/\eta \leq \nu |U'|/10. \tag{3.6.4}$$

Suppose first that $U = U_i$ for some $1 \leq i \leq k$. Then U is a (ρ, ν, τ)-robust expander component. By (3.6.4), we may apply Lemma 3.4.8 with $U, U \setminus I$ playing the roles of U, U' to see that $G[U']$ is a robust $(\nu/2, 2\tau)$-expander and $\delta(G[U']) \geq \eta n/2$. By (3.6.3) and Corollary 3.6.8, $G[U']$ is Hamilton p-linked. So there is an ordered collection $\mathcal{R}_U$ of p vertex-disjoint paths in $G[U']$ spanning U' such that the jth path in $\mathcal{R}_U$ joins $x^+_{i_j}$ and $x^-_{i_j+1}$.

Suppose instead that $U = Z_i$ for some $1 \leq i \leq \ell$. Then there exists a bipartition A, B of U such that U is a bipartite (ρ, ν, τ)-robust expander component with bipartition A, B. Let $A' := A \setminus I$ and $B' := B \setminus I$. So A', B' is a bipartition of U'. Recall from (T4) that $\mathcal{P}$ is (A, B)-balanced. Thus $|A'| = |B'|$ and $\mathrm{End}_{\mathcal{P}}(A') = \mathrm{End}_{\mathcal{P}}(B') > 0$.

Let $n' := |A'|$. Note that (D5′) implies that $\delta(G[A, B]) \geq \eta n/2$. By (B1) and (C2) we have that $\Big||A| - |B|\Big| \leq \rho n$ and hence (3.6.2) implies that $|A| \geq 2|U'|/5$. Now

(3.6.4) implies that $|I| \leq \nu|U'|/10 \leq \nu|A|/4$. So we may apply Lemma 3.4.9(i) to see that $G[U']$ is a bipartite robust $(\nu/2, 2\tau)$-expander with bipartition A', B', and that $\delta(G[A', B']) \geq \eta n'/4$. By (3.6.3) and Lemma 3.6.11, H is (A', B')-Hamilton p-linked. So there is an ordered collection $\mathcal{R}_U$ of p vertex-disjoint paths in H spanning U' such that the jth path in $\mathcal{R}_U$ joins $x^+_{i_j}$ and $x^-_{i_j+1}$.

Proceed in this way for each $U \in \mathcal{U}$ and let $\mathcal{R} := \bigcup_{U \in \mathcal{U}} \mathcal{R}_U$. Then for each $1 \leq i \leq q$, there exists exactly one path R_i in $\mathcal{R}$ which joins x^+_i and x^-_{i+1} (with indices modulo q). Let

$$C := x^-_1 P_1 x^+_1 R_1 x^-_2 P_2 x^+_2 \dots x^-_p P_p x^+_p R_p x^-_1.$$

Then C is a cycle in G which covers $\bigcup_{U \in \mathcal{U}} U$. □

3.7 The proof of Theorem 3.1.2

Our aim is to prove Theorem 3.1.2, i.e. that every sufficiently large 3-connected D-regular graph G on n vertices with $D \geq (1/4+\varepsilon)n$ contains a Hamilton cycle. By Theorem B and Proposition 3.3.1(i), G has a robust partition $\mathcal{V}$ such that (k, ℓ) takes one of five values. By Corollary 3.6.3, to find a Hamilton cycle it suffices to find a $\mathcal{V}$-tour. We achieve this for each case. In the first subsection we consider the case $\ell = 0$ (so $1 \leq k \leq 3$), i.e. when G is a union of robust expander components. Then in Subsection 3.7.2 we prove some lemmas which are useful for the case when $\ell \geq 1$. Finally in Subsections 3.7.3 and 3.7.4 we consider the cases $(k, \ell) = (0, 1), (1, 1)$ respectively.

3.7.1 Finding $\mathcal{V}$-tours in a 3-connected graph with at most three robust expander components

The main result of this section guarantees a $\mathcal{V}$-tour in a 3-connected graph G which has a robust partition $\mathcal{V}$ into at most three robust expander components.

To prove this, we use the fact that there is a matching of size 3 between any set A

of vertices with $3 \leq |A| \leq |G| - 3$, and its complement $\overline{A}$ in G. Applying this (possibly more than once) and using a case analysis implies that we can find a path system $\mathcal{P}$ such that $R_{\mathcal{V}}(\mathcal{P})$ has an Euler tour with at most 4 edges. Since $\mathcal{V}$ contains no bipartite robust components, $\mathcal{P}$ is a suitable $\mathcal{V}$-tour.

Lemma 3.7.1. *Let $D, n \in \mathbb{N}$, let $0 < 1/n \ll \rho \ll \nu \ll \tau \ll \alpha < 1$ and let $D \geq \alpha n$. Suppose that G is a D-regular 3-connected graph on n vertices and that $\mathcal{V}$ is a robust partition of G with parameters $\rho, \nu, \tau, k, 0$ where $k \leq 3$. Then G contains a $\mathcal{V}$-tour with parameter $4/n$.*

We will use the following proposition which is an immediate consequence of Menger's Theorem.

Proposition 3.7.2. *Let $k \in \mathbb{N}$ and let G be a k-connected graph. Suppose that A is a subset of G with $|A|, |\overline{A}| \geq k$. Then there is a matching of size k between A and $\overline{A}$.*

Lemma 3.7.1 is an immediate corollary of the following lemma. To see this, note that (T4) is vacuous here.

Lemma 3.7.3. *Let G be a 3-connected graph and let $\mathcal{V}$ be a partition of $V(G)$ into at most three parts, where $|V| \geq 3$ for each $V \in \mathcal{V}$. Then G contains a path system $\mathcal{P}$ such that*

(i) $e(\mathcal{P}) \leq 4$ and $\mathcal{P} \subseteq \bigcup_{V \in \mathcal{V}} G[V, \overline{V}]$;

(ii) $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour;

(iii) for each $V \in \mathcal{V}$, if c_i is the number of vertices in V with degree i in $\mathcal{P}$ (for $i = 1, 2$), then $c_1 + 2c_2 \in \{2, 4\}$ and $c_2 \leq 1$.

Proof. Suppose first that $|\mathcal{V}| = 1$. Let $\mathcal{P}$ consist of a single arbitrary edge. So (i) and (iii) are clear. Then $R_{\mathcal{V}}(\mathcal{P})$ is a loop, so (ii) holds.

Suppose instead that $|\mathcal{V}| = 2$ and write $\mathcal{V} := \{V, W\}$. Then Proposition 3.7.2 implies that G contains a matching $\mathcal{P}$ of size two between V and W. So (i) holds. In this case,

$R_{\mathcal{V}}(\mathcal{P})$ consists of exactly two VW-edges, so (ii) holds. Moreover, for each $V \in \mathcal{V}$ we have $(c_1, c_2) = (2, 0)$, implying (iii).

Suppose finally that $|\mathcal{V}| = 3$ and write $\mathcal{V} := \{V_1, V_2, V_3\}$. We write M_{ij} for a matching between V_i and V_j. Given a path system $\mathcal{P}$ in G, we write c_i^j for the number of vertices in V_j with degree i in $\mathcal{P}$. Proposition 3.7.2 implies that there is a matching of size three between V_1 and $V_2 \cup V_3$. Without loss of generality, choose M_{12} such that $|M_{12}| = 2$. By Proposition 3.7.2 there is a matching of size three between V_3 and $V_1 \cup V_2$. Therefore there exist vertex-disjoint M_{13}, M_{23} such that $|M_{13}| + |M_{23}| = 3$. Throughout the remainder of the proof, we will let u_1, v_1, w_1, x_1 be distinct vertices in V_1 and we will label vertices in other classes similarly.

Case 1. $|M_{13}| = 3$.

If M_{13} contains two edges e, e' that are vertex-disjoint from M_{12}, then we let $\mathcal{P}$ have edge-set $\{e, e'\} \cup M_{12}$. So (i) holds. Note that $R_{\mathcal{V}}(\mathcal{P})$ consists of precisely two V_1V_2-edges and two V_1V_3-edges. Therefore (ii) holds. Moreover, $(c_1^1, c_2^1) = (4, 0)$ and $(c_1^j, c_2^j) = (2, 0)$ for $j = 2, 3$, implying (iii).

Otherwise, M_{13} contains exactly two edges that share endpoints with edges in M_{12}. Without loss of generality, let $M_{13} := \{u_1u_3, v_1v_3, w_1w_3\}$ and $M_{12} := \{u_1u_2, v_1v_2\}$. In this case, let $\mathcal{P} := \{u_1u_2, v_2v_1v_3, w_1w_3\}$. (i) is immediate, and $R_{\mathcal{V}}(\mathcal{P}) \cong C_3$ so (ii) holds. Moreover, $(c_1^1, c_2^1) = (2, 1)$ and $(c_1^j, c_2^j) = (2, 0)$ for $j = 2, 3$, implying (iii).

Case 2. *Without loss of generality,* $|M_{13}| = 2$ *and* $|M_{23}| = 1$.

Let v_2v_3 be the edge in M_{23}. Since $|M_{12}| = |M_{13}| = 2$ we can pick edges $w_1w_2 \in M_{12}$ and $x_1x_3 \in M_{13}$ so that $w_2 \neq v_2$ and $x_1 \neq w_1$. But M_{13} and M_{23} are vertex-disjoint, so $x_3 \neq v_3$. In this case, we let $\mathcal{P} := \{w_1w_2, v_2v_3, x_3x_1\}$. (i) is immediate, and $R_{\mathcal{V}}(\mathcal{P}) \cong C_3$ so (ii) holds. Moreover, $(c_1^j, c_2^j) = (2, 0)$ for all $V \in \mathcal{V}$, implying (iii). This completes the proof of the case $|\mathcal{V}| = 3$. $\square$

3.7.2 Finding an (A, B)-balanced path system in a bipartite robust expander

In Section 3.6 we showed that, given a robust partition $\mathcal{V}$, 'the balancing property' (T4) was sufficient to extend a $\mathcal{V}$-tour into a Hamilton cycle. In this section we prove some lemmas which will be useful in finding a path system which satisfies (T4).

Suppose that W is a bipartite robust component with bipartition A, B, where $|A| \geq |B|$. We will show that if a path system $\mathcal{P}$ satisfies a particular condition (3.7.1) on $e_{\mathcal{P}}(A), e_{\mathcal{P}}(B), e_{\mathcal{P}}(W, \overline{W})$, we can add AB-edges to $\mathcal{P}$ to obtain an (A, B)-balanced path system. So to find a path system $\mathcal{P}'$ which satisfies (T4), it suffices to find $\mathcal{P}$ which satisfies (3.7.1) for all bipartite robust components.

We begin by observing the following crucial fact.

Proposition 3.7.4. *Let G be a D-regular graph with vertex partition A, B, V. Then*

(i) $2(e(A) - e(B)) + e(A, V) - e(B, V) = (|A| - |B|)D$.

In particular,

(ii) $2e(A) + e(A, V) \geq (|A| - |B|)D$;

(iii) if $V = \emptyset$ then $2(e(A) - e(B)) = (|A| - |B|)D$.

Proof. It suffices to prove (i) since (ii) and (iii) are then immediate. We have that

$$\sum_{x \in A} d_B(x) = e(A, B) = \sum_{y \in B} d_A(y).$$

Moreover, by counting degrees,

$$2e(A) + e(A, V) = \sum_{x \in A} (D - d_B(x)) = D|A| - \sum_{x \in A} d_B(x),$$

and similarly for B. So $2e(A) - 2e(B) + e(A, V) - e(B, V) = D(|A| - |B|)$, as desired. □

The following proposition is an immediate consequence of Vizing's Theorem on edge-colourings.

Proposition 3.7.5. *Let H be a graph with $\Delta(H) \leq \Delta$. Then H contains a matching of size $\lceil e(H)/(\Delta+1) \rceil$.*

Given a graph G, a collection $\mathcal{U}$ of disjoint subsets of $V(G)$ and a $\mathcal{U}$-anchored path system $\mathcal{P}$ in G, we say that a path system $\mathcal{P}'$ is a *$\mathcal{U}$-extension of $\mathcal{P}$* if

- every edge which lies in a path of $\mathcal{P}'$ but not a path of $\mathcal{P}$ lies in $\bigcup_{U \in \mathcal{U}} G[U]$;
- for every $P' \in \mathcal{P}'$ there is at most one $P \in \mathcal{P}$ such that $P \subseteq P'$.

If $U \subseteq V(G)$ we will write *U-extension* for $\{U\}$-extension. The next lemma shows that a $\mathcal{U}$-extension $\mathcal{P}'$ of $\mathcal{P}$ 'behaves similarly' to $\mathcal{P}$ in the reduced multigraph $R_{\mathcal{U}}$, and also that $R_{\mathcal{U}}$ is not affected by considering a slightly different partition.

Lemma 3.7.6. *Let $\mathcal{U}$ be a collection of disjoint vertex-subsets of a graph G and let $\mathcal{P}$ be a $\mathcal{U}$-anchored path system in G.*

(i) Suppose that $\mathcal{P}'$ is a $\mathcal{U}$-extension of $\mathcal{P}$. Then $\mathcal{P}'$ is a $\mathcal{U}$-anchored path system.

(ii) Suppose that $\mathcal{P}'$ is an $\mathcal{X}$-extension of $\mathcal{P}$ for some $\mathcal{X} \subseteq \mathcal{U}$. Then $\mathcal{P}'$ is a $\mathcal{U}$-extension of $\mathcal{P}$.

(iii) Suppose that $\mathcal{P}'$ is a $\mathcal{U}$-extension of $\mathcal{P}$. Then $R_{\mathcal{U}}(\mathcal{P}')$ is an Euler tour if and only if $R_{\mathcal{U}}(\mathcal{P})$ is an Euler tour.

(iv) Suppose that $\mathcal{U} := \{U_1, \ldots, U_t\}$, $\mathcal{X} := \{X_1, \ldots, X_t\}$, $X_i \subseteq U_i$ for all $1 \leq i \leq t$, and $\mathcal{P}$ is $\mathcal{X}$-anchored. Then $R_{\mathcal{X}}(\mathcal{P}) \cong R_{\mathcal{U}}(\mathcal{P})$.

Proof. Note that (i), (ii) and (iv) are immediate. To prove (iii), let $\mathcal{R}$ be the subset of $\mathcal{P}'$ such that every $R \in \mathcal{R}$ contains some $P_R \in \mathcal{P}$. So $|\mathcal{R}| = |\mathcal{P}|$. Observe that P_R has endpoints in $U, U' \in \mathcal{U}$ if and only if R has endpoints in U, U'. So $R_{\mathcal{U}}(\mathcal{R}) \cong R_{\mathcal{U}}(\mathcal{P})$. Let $\mathcal{Q} := \mathcal{P}' \setminus \mathcal{R}$. Then every edge in a path in $\mathcal{Q}$ lies in $\bigcup_{U \in \mathcal{U}} G[U]$. So $R_{\mathcal{U}}(\mathcal{Q})$ consists entirely of loops. Therefore $R_{\mathcal{U}}(\mathcal{P}') = R_{\mathcal{U}}(\mathcal{R}) \cup R_{\mathcal{U}}(\mathcal{Q})$ is an Euler tour if and only if $R_{\mathcal{U}}(\mathcal{R})$ is, i.e. if and only if $R_{\mathcal{U}}(\mathcal{P})$ is. This proves (iii). $\square$

Suppose that $A, B \subseteq V(G)$ are disjoint. The following lemma gives a sufficient condition which ensures that a path system $\mathcal{P}$ can be extended into an (A, B)-balanced path system which does not cover too much of $A \cup B$. Whenever we wish to find a balanced path system we will find a collection of paths which satisfy this condition.

The idea is that, given an (A, B)-balanced path system $\mathcal{P}'$, removing every AB-edge gives a path system $\mathcal{P}$ which satisfies (3.7.1) below. The lemma proceeds in the opposite direction: one can add AB-edges to such a $\mathcal{P}$ to recover $\mathcal{P}'$.

Lemma 3.7.7. *Let $n \in \mathbb{N}$ and $0 < 1/n \ll \rho < 1$ and suppose that G is a graph on n vertices. Let $U \subseteq V(G)$ have bipartition A, B where $\Big||A| - |B|\Big| \leq \rho n$ and $\delta(G[A, B]) > 9\rho n$. Let $\mathcal{P}$ be a path system in G such that $|V(\mathcal{P}) \cap U| \leq \rho n$,*

$$2e_{\mathcal{P}}(A) - 2e_{\mathcal{P}}(B) + e_{\mathcal{P}}(A, \overline{U}) - e_{\mathcal{P}}(B, \overline{U}) = 2(|A| - |B|) \tag{3.7.1}$$

and $\mathcal{P}$ has at least one endpoint in U. Then G contains a path system $\mathcal{P}'$ such that

(α) $\mathcal{P}'$ is a U-extension of $\mathcal{P}$;

(β) $\mathcal{P}'$ is (A, B)-balanced;

(γ) $|V(\mathcal{P}') \cap U| \leq 9\rho n$.

Proof. Without loss of generality, suppose that $|A| \geq |B|$. Let $A_0 \subseteq A$ and $B_0 \subseteq B$ be minimal such that $V(\mathcal{P}) \cap U \subseteq A_0 \cup B_0$ and

$$|A_0| - |B_0| = |A| - |B|. \tag{3.7.2}$$

Note that

$$|A_0| + |B_0| = |A| - |B| + 2|B_0| \leq \Big||A| - |B|\Big| + 2|V(\mathcal{P}) \cap U| \leq 3\rho n. \tag{3.7.3}$$

For each $u \in A_0$, find a set N_u of $2 - d_{\mathcal{P}}(u)$ neighbours of u in $B \setminus B_0$. For each $v \in B_0$, find a set N_v of $2 - d_{\mathcal{P}}(v)$ neighbours of v in $A \setminus A_0$. Choose these sets to be disjoint and

such that $(N_u \cup N_v) \cap V(\mathcal{P}) = \emptyset$. This is possible since for each $u \in A$ and $v \in B$ we have $d_B(u), d_A(v) > 3(|A_0| + |B_0|)$. Obtain $\mathcal{P}'$ from $\mathcal{P}$ by adding the edges xx' to (the paths in) $\mathcal{P}$ for each $x \in A_0 \cup B_0$ and for each $x' \in N_x$. It is clear that $\mathcal{P}'$ is a U-extension of $\mathcal{P}$, so (α) holds.

Note that the set of internal vertices of $\mathcal{P}'$ which lie in U is precisely $A_0 \cup B_0$. Then $\mathrm{Int}_{\mathcal{P}'}(A) - \mathrm{Int}_{\mathcal{P}'}(B) = |A_0| - |B_0| = |A| - |B|$ by (3.7.2). So to show (β), it is enough to check that $\mathrm{End}_{\mathcal{P}'}(A) = \mathrm{End}_{\mathcal{P}'}(B)$ and that this value is non-zero. Since

$$\sum_{u \in A} d_{\mathcal{P}'}(u) = 2e_{\mathcal{P}'}(A) + e_{\mathcal{P}'}(A, B) + e_{\mathcal{P}'}(A, \overline{U}) = 2e_{\mathcal{P}}(A) + e_{\mathcal{P}'}(A, B) + e_{\mathcal{P}}(A, \overline{U}),$$

and similarly for B, we have that

$$\begin{aligned} \sum_{u \in A} d_{\mathcal{P}'}(u) - \sum_{v \in B} d_{\mathcal{P}'}(v) &= 2e_{\mathcal{P}}(A) - 2e_{\mathcal{P}}(B) + e_{\mathcal{P}}(A, \overline{U}) - e_{\mathcal{P}}(B, \overline{U}) \\ &\overset{(3.7.1)}{=} 2(|A| - |B|). \end{aligned} \tag{3.7.4}$$

By construction, $\sum_{u \in A} d_{\mathcal{P}'}(u) = \mathrm{End}_{\mathcal{P}'}(A) + 2|A_0|$, and similarly for B. So, by (3.7.4),

$$\mathrm{End}_{\mathcal{P}'}(A) - \mathrm{End}_{\mathcal{P}'}(B) = 2(|A| - |B|) - 2(|A_0| - |B_0|) \overset{(3.7.2)}{=} 0. \tag{3.7.5}$$

Recall that $\mathcal{P}$ has at least one endpoint x lying in U. Then $|N_x| = 1$ and the vertex in N_x is an endpoint of a path in $\mathcal{P}'$. So $\mathrm{End}_{\mathcal{P}'}(A) = \mathrm{End}_{\mathcal{P}'}(B)$ is non-zero, proving (β).

Finally, note that every vertex in $V(\mathcal{P}') \cap U$ which does not lie in $A_0 \cup B_0$ is a neighbour of some $x \in A_0 \cup B_0$ in $\mathcal{P}'$. So (3.7.3) implies that

$$|V(\mathcal{P}') \cap U| \leq |A_0 \cup B_0| + |N_{\mathcal{P}'}(A_0 \cup B_0)| \leq 3(|A_0| + |B_0|) \leq 9\rho n,$$

proving (γ). □

The next lemma is an iteration of Lemma 3.7.7. We will use it to successively extend a path system into one that is (A, B)-balanced for all appropriate A, B.

Lemma 3.7.8. *Let $n, k, \ell \in \mathbb{N}$ and $0 < 1/n \ll \rho \ll \nu \ll \tau \ll \eta < 1$. Let G be a graph on n vertices and suppose that $\mathcal{U} := \{U_1, \dots, U_k, W_1, \dots, W_\ell\}$ is a weak robust subpartition in G with parameters $\rho, \nu, \tau, \eta, k, \ell$. For each $1 \le j \le \ell$, let A_j, B_j be the bipartition of W_j specified by* (D3′)*. Let $\mathcal{P}$ be a $\mathcal{U}$-anchored path system such that for each $1 \le j \le \ell$,*

$$2e_{\mathcal{P}}(A_j) - 2e_{\mathcal{P}}(B_j) + e_{\mathcal{P}}(A_j, \overline{W_j}) - e_{\mathcal{P}}(B_j, \overline{W_j}) = 2(|A_j| - |B_j|). \tag{3.7.6}$$

Suppose further that $|V(\mathcal{P}) \cap U| \le \rho n$ for all $U \in \mathcal{U}$, and that $R_{\mathcal{U}}(\mathcal{P})$ is an Euler tour. Then G contains a $\mathcal{U}$-extension $\mathcal{P}'$ of $\mathcal{P}$ that is a $\mathcal{U}$-tour with parameter 9ρ.

Proof. Let $\mathcal{P}_0 := \mathcal{P}$. Suppose that for some $0 \le i < \ell$, we have already defined a path system $\mathcal{P}_i$ such that

(α_i) $\mathcal{P}_i$ is a $\{W_1, \dots, W_i\}$-extension of $\mathcal{P}$;

(β_i) for all $1 \le j \le i$, $\mathcal{P}_i$ is (A_j, B_j)-balanced;

(γ_i) for all $1 \le j \le i$, $|V(\mathcal{P}_i) \cap W_j| \le 9\rho n$.

Now we obtain $\mathcal{P}_{i+1}$ from $\mathcal{P}_i$ as follows. Note that (D3′) implies that (B1) and (C2) hold and hence that $\Big||A_{i+1}| - |B_{i+1}|\Big| \le \rho n$. Moreover, by (D5′) we have that $\delta(G[A_j, B_j]) \ge \eta n/2 > 9\rho n$. Also ($\alpha_i$) implies that $|V(\mathcal{P}_i) \cap W_{i+1}| = |V(\mathcal{P}) \cap W_{i+1}| \le \rho n$ and that (3.7.6) still holds with $i+1$ and $\mathcal{P}_i$ playing the roles of j and $\mathcal{P}$. Finally, $R_{\mathcal{U}}(\mathcal{P})$ is a non-empty Euler tour, so $\mathcal{P}$ contains at least one endpoint in W_{i+1}. Thus $\mathcal{P}_i$ contains at least one endpoint in W_{i+1} by (α_i). Therefore we can apply Lemma 3.7.7 with $W_{i+1}, A_{i+1}, B_{i+1}, \mathcal{P}_i, \rho$ playing the roles of $U, A, B, \mathcal{P}, \rho$. We thus obtain a path system $\mathcal{P}_{i+1}$ satisfying Lemma 3.7.7(α)–(γ). Now (α) and (α_i) imply that (α_{i+1}) holds. We obtain (β_{i+1}) and (γ_{i+1}) in a similar way.

Therefore we can obtain $\mathcal{P}' := \mathcal{P}_\ell$ that satisfies (α_ℓ)–(γ_ℓ). Now (α_ℓ) and Lemma 3.7.6(ii) imply that $\mathcal{P}'$ is a $\mathcal{U}$-extension of $\mathcal{P}$. It remains to show that (T1)–(T4) hold for $\mathcal{P}'$ with 9ρ playing the role of γ. Indeed, (T1) follows from Lemma 3.7.6(i) and the fact that $\mathcal{P}'$ is a $\mathcal{U}$-extension of $\mathcal{P}$. Since $\mathcal{R}_{\mathcal{U}}(\mathcal{P})$ is an Euler tour, Lemma 3.7.6(iii) implies that $\mathcal{R}_{\mathcal{U}}(\mathcal{P}')$

is an Euler tour, and hence (T2) holds. We have $|V(\mathcal{P}') \cap W_j| \leq 9\rho n$ for all $1 \leq j \leq \ell$ by (γ_ℓ). Moreover, by (α_ℓ) we have that $|V(\mathcal{P}') \cap U_j| = |V(\mathcal{P}) \cap U_j| \leq \rho n$ for all $1 \leq j \leq k$. So (T3) holds. Finally, (T4) is immediate from (β_ℓ). □

3.7.3 Finding a $\mathcal{V}$-tour in a regular bipartite robust expander

We now consider the case when G has a robust partition with $(k, \ell) = (0, 1)$, i.e. G is a regular bipartite robust expander. By Corollary 3.6.3, in order to find a Hamilton cycle in G it suffices to find a $\{V(G)\}$-tour with an appropriate parameter. This is guaranteed by the following lemma.

Lemma 3.7.9. *Let $D, n \in \mathbb{N}$ and let $0 < 1/n \ll \rho \ll \nu \ll \tau \ll \alpha < 1$. Let G be a D-regular graph on n vertices where $D \geq \alpha n$. Suppose that G has a robust partition $\mathcal{V}$ with parameters $\rho, \nu, \tau, 0, 1$. Then G contains a $\mathcal{V}$-tour with parameter 18ρ.*

Proof. Note (D3) implies that there exists a bipartition A, B of $V(G)$ such that G is a bipartite (ρ, ν, τ)-expander with bipartition A, B. By (D5) we have that $\delta(G[A, B]) \geq D/2$. Therefore

$$\Delta(G[A]), \Delta(G[B]) \leq D/2. \tag{3.7.7}$$

Moreover, (B1) (which follows from (D3)) implies that G is ρ-close to bipartite with bipartition A, B. So (C2) holds, i.e.

$$\Big||A| - |B|\Big| \leq \rho n. \tag{3.7.8}$$

Suppose first that $|A| = |B|$. Then let $\mathcal{P}$ consist of exactly one AB-edge. Note that $R_{\mathcal{V}}(\mathcal{P})$ is a loop and that $\mathcal{P}$ is (A, B)-balanced. All of (T1)–(T4) hold.

Without loss of generality, assume that $|A| > |B|$. Proposition 3.7.4(iii) implies that

$$e(A) \geq e(A) - e(B) = (|A| - |B|)D/2. \tag{3.7.9}$$

Proposition 3.7.5 implies that $G[A]$ contains a matching of size

$$\left\lceil \frac{e(A)}{\Delta(A)+1} \right\rceil \overset{(3.7.7),(3.7.9)}{\geq} \left\lceil \frac{(|A|-|B|)D/2}{D/2+1} \right\rceil = |A|-|B| - \left\lfloor \frac{|A|-|B|}{D/2+1} \right\rfloor$$
$$\overset{(3.7.8)}{\geq} |A|-|B| - \lfloor 2\rho/\alpha \rfloor = |A|-|B|.$$

So we can choose a matching M of size $|A|-|B|$ in $G[A]$.

Now Proposition 3.6.1(i) implies that $\mathcal{V}$ is a weak robust subpartition in G with parameters $\rho, \nu, \tau, \alpha^2/2, 0, 1$. Certainly M is $\mathcal{V}$-anchored and

$$2e_M(A) - 2e_M(B) + e_M(A, \overline{V(G)}) - e_M(B, \overline{V(G)}) = 2e_M(A) = 2(|A|-|B|).$$

We also have that $|V(M)| = 2(|A|-|B|) \leq 2\rho n$. Moreover, M is non-empty since $|A|-|B| > 0$. Thus $R_{\mathcal{V}}(M)$ is a non-empty collection of loops and hence a non-empty Euler tour. Therefore we can apply Lemma 3.7.8 with $\mathcal{V}, 0, 1, V(G), A, B, M, 2\rho, \alpha^2/2$ playing the roles of $\mathcal{U}, k, \ell, W_j, A_j, B_j, \mathcal{P}, \rho, \eta$ to obtain a path system $\mathcal{P}$ which is a $\mathcal{V}$-tour with parameter 18ρ. □

3.7.4 Finding a $\mathcal{V}$-tour when there is exactly one component of each type

We would like to find a Hamilton cycle when G is the union of a robust expander component V and a bipartite robust expander component W. By Corollary 3.6.3, it is sufficient to find a $\mathcal{V}$-tour for this robust partition $\mathcal{V}$. This is guaranteed by the following lemma.

Lemma 3.7.10. *Let $n, D \in \mathbb{N}$, $0 < 1/n \ll \rho \ll \nu \ll \tau \ll \alpha < 1$ and let $D \geq \alpha n$. Suppose that G is a 3-connected D-regular graph on n vertices and that $\mathcal{V}$ is a robust partition of G with parameters $\rho, \nu, \tau, 1, 1$. Then G contains a $\mathcal{V}$-tour with parameter 36ρ.*

Let V, W be as above and let A, B be a bipartition of W such that W is a bipartite robust expander with respect to A, B. Suppose that $|A| \geq |B|$. To prove Lemma 3.7.10,

our aim is to find a path system $\mathcal{P}$ to which we can apply Lemma 3.7.8 and hence obtain a $\mathcal{V}$-tour. Roughly speaking, $\mathcal{P}$ will consist of the union of two matchings, M_A in $G[A]$ and $M_{A,V}$ in $G[A,V]$ which together have the right size to 'balance' W.

Proof of Lemma 3.7.10. Let $\mathcal{V} := \{V, W\}$, where V is a (ρ,ν,τ)-robust expander component and W has bipartition A, B so that W is a bipartite (ρ,ν,τ)-robust expander component with respect to A, B. So (B1) and (C2) imply that

$$\Big||A| - |B|\Big| \leq \rho n. \tag{3.7.10}$$

Moreover, (D4) implies that $\delta(G[V]), \delta(G[W]) \geq D/2$ and therefore

$$D/2 \geq \Delta(G[W,V]) \geq \Delta(G[A,V]). \tag{3.7.11}$$

By (D5) we have

$$\Delta(G[A]) \leq D/2. \tag{3.7.12}$$

Claim 1. *It suffices to find a path system $\mathcal{P}$ in G such that the following hold:*

(i) $2e_{\mathcal{P}}(A) - 2e_{\mathcal{P}}(B) + e_{\mathcal{P}}(A,V) - e_{\mathcal{P}}(B,V) = 2(|A| - |B|)$*;*

(ii) $e(\mathcal{P}) \leq 2\rho n$*;*

(iii) $\mathcal{P}$ *has at least one* VW*-path.*

Proof. Note that Proposition 3.6.1(i) implies that $\mathcal{V}$ is a weak robust subpartition in G with parameters $\rho, \nu, \tau, \alpha^2/2, 1, 1$. Clearly, $\mathcal{P}$ is a $\mathcal{V}$-anchored path system. Observe that (D5) implies that $\delta(G[A,B]) \geq D/4$. Let p be the number of VW-paths in $\mathcal{P}$. Then $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour if and only if p is positive and even. By (iii) we have $p > 0$. Now (i) implies that

$$e_{\mathcal{P}}(W,V) = e_{\mathcal{P}}(A,V) + e_{\mathcal{P}}(B,V) = 2(|A| - |B|) - 2e_{\mathcal{P}}(A) + 2e_{\mathcal{P}}(B) + 2e_{\mathcal{P}}(B,V)$$

is even. Note that any $P \in \mathcal{P}$ contains an odd number of VW-edges if P is a VW-

path, and an even number otherwise. Therefore p is even and so $R_{\mathcal{V}}(\mathcal{P})$ is a non-empty Euler tour. Finally, for each $X \in \mathcal{V}$ we have $|V(\mathcal{P}) \cap X| \leq 2e(\mathcal{P}) \leq 4\rho n$ by (ii). Therefore we can apply Lemma 3.7.8 with $\mathcal{V}, 1, 1, W, A, B, \mathcal{P}, 4\rho, \alpha^2/2$ playing the roles of $\mathcal{U}, k, \ell, W_j, A_j, B_j, \mathcal{P}, \rho, \eta$ to find a $\mathcal{V}$-extension $\mathcal{P}'$ of $\mathcal{P}$ that is a $\mathcal{V}$-tour with parameter 36ρ, proving the claim. ■

So it remains to find a path system $\mathcal{P}$ as in Claim 1. Suppose first that $|A| = |B|$. Since G is 3-connected, Proposition 3.7.2 implies that $G[V, W]$ contains a matching of size three. We only consider the case when $G[A, V]$ contains a matching $M_{A,V}$ of size two. (The case when this holds for $G[B, V]$ is similar.) Now Proposition 3.7.4(i) implies that

$$2e(B) + e(B, V) = 2e(A) + e(A, V) \geq 2.$$

If $e(B) \geq 1$, let $\mathcal{P} := M_{A,V} \cup \{e\}$, where e is an edge in $G[B]$. Otherwise, $e(B) = 0$ and hence $e(B, V) \geq 2$. In this case we let $\mathcal{P}$ consist of two vertex-disjoint edges $e \in G[A, V]$ and $e' \in G[B, V]$. In both cases, (i)–(iii) clearly hold for $\mathcal{P}$ and we are done.

Without loss of generality, assume that $|A| > |B|$. Proposition 3.7.4(ii) implies that

$$2e(A) + e(A, V) \geq (|A| - |B|)D. \tag{3.7.13}$$

Suppose first that $e(A) < D/5$. Then (3.7.13) implies that $e(A, V) \geq (|A| - |B|)D - 2D/5$. Now Proposition 3.7.5 implies that $G[A, V]$ contains a matching of size at least

$$\begin{aligned} \left\lceil \frac{e(A,V)}{\Delta(G[A,V])+1} \right\rceil &\overset{(3.7.11)}{\geq} \left\lceil \frac{(|A|-|B|)D - 2D/5}{D/2+1} \right\rceil \\ &= 2(|A|-|B|) - \left\lfloor \frac{2(|A|-|B|) + 2D/5}{D/2+1} \right\rfloor \\ &\overset{(3.7.10)}{\geq} 2(|A|-|B|) - \left\lfloor \frac{D/2}{D/2+1} \right\rfloor = 2(|A|-|B|). \end{aligned} \tag{3.7.14}$$

Let $\mathcal{P}$ be a matching of size $2(|A| - |B|)$ in $G[A, V]$. Then $\mathcal{P}$ satisfies (i)–(iii) (indeed, (ii) follows from (3.7.10)).

Therefore we can assume that $e(A) \geq D/5$. Let

$$\ell := \min\left\{\left\lceil \frac{e(A)}{D/2+1} \right\rceil, |A| - |B|\right\}. \tag{3.7.15}$$

Note that $\ell \geq 1$. Clearly $G[A]$ contains a matching of size ℓ by Proposition 3.7.5 and (3.7.12). We now consider two cases, depending on the value of ℓ.

Case 1. $\ell = |A| - |B|$.

Let M be a matching of size ℓ in $G[A]$. Since G is 3-connected, Proposition 3.7.2 implies that $G[V, W]$ contains a matching of size three. Suppose first that $G[A, V]$ contains a matching $M_{A,V}$ of size two. Write $V(M_{A,V}) \cap A := \{u, u'\}$. If uu' is an edge in M, delete it to obtain M'. Otherwise delete an arbitrary edge from M to obtain M'. Let $\mathcal{P} := M' \cup M_{A,V}$. Then $\mathcal{P}$ is a path system satisfying (i). Also (ii) follows from (3.7.10). Moreover, u lies in a VW-path in $\mathcal{P}$, so (iii) holds.

So suppose that $G[A, V]$ does not contain a matching of size two. Then $G[B, V]$ contains a matching $M_{B,V}$ of size two. Moreover, there is at most one vertex in $A \cup V$ such that every edge in $G[A, V]$ is incident to this vertex. Therefore (3.7.11) implies that $e(A, V) \leq \Delta(G[A, V]) \leq D/2$. So

$$e(A) - |M| \overset{(3.7.13)}{\geq} (|A| - |B|)D/2 - D/4 - |M| \geq D/4 - 1 > 0,$$

where the penultimate inequality follows from the fact that $|M| = |A| - |B| > 0$. So we can find an edge e in $G[A]$ that is not contained in M. Let $\mathcal{P} := M_{B,V} \cup M \cup \{e\}$. Then $\mathcal{P}$ is a path system satisfying (i)–(iii). This completes the proof of Case 1.

Case 2. $\ell < |A| - |B|$ *and so* $\ell = \lceil e(A)/(D/2+1) \rceil$.

> **Claim 2.** *Suppose that $G[A]$ contains no matching of size $\ell + 1$. Then $G[A]$ contains a matching M^- of size $\ell - 1$ and a path $P := xyz$ which is vertex-disjoint from M^-.*
>
> *Proof.* Suppose first that $\Delta(G[A]) \leq D/8 - 1$. Then Proposition 3.7.5 implies that

$G[A]$ contains a matching of size

$$\left\lceil \frac{e(A)}{D/8} \right\rceil = \left\lceil \frac{e(A)}{D/3} + \frac{5e(A)}{D} \right\rceil \geq \left\lceil \frac{e(A)}{D/3} + 1 \right\rceil \geq \ell + 1, \tag{3.7.16}$$

a contradiction. So $\Delta(G[A]) > D/8 - 1 > 2\ell$ by (3.7.10) and (3.7.15). Recall that $G[A]$ contains a matching M of size ℓ. Since M must be maximal, there is some $y \in V(M)$ such that $d_A(y) > 2\ell$. Let $x \in A$ be a neighbour of y such that $x \notin V(M)$. Let z be the neighbour of y in M. Let $M^- := M \setminus \{yz\}$ and $P := xyz$.

■

Proposition 3.7.5 implies that $G[A, V]$ contains a matching of size

$$\begin{aligned}
\left\lceil \frac{e(A,V)}{\Delta(G[A,V])+1} \right\rceil &\overset{(3.7.11)}{\geq} \left\lceil \frac{e(A,V)}{D/2+1} \right\rceil + 2\left\lceil \frac{e(A)}{D/2+1} \right\rceil - 2\ell \\
&\geq \left\lceil \frac{2e(A)+e(A,V)}{D/2+1} \right\rceil - 2\ell \\
&\overset{(3.7.13)}{\geq} \left\lceil \frac{(|A|-|B|)D}{D/2+1} \right\rceil - 2\ell \geq 2(|A|-|B|-\ell),
\end{aligned}$$

where the final inequality follows in a similar way to (3.7.14). So we can choose a matching $M_{A,V}$ in $G[A, V]$ of size $2(|A| - |B| - \ell) > 0$.

Let E be any collection of ℓ edges in $G[A]$ and let $H := E \cup M_{A,V}$. Then

$$2e_H(A) - 2e_H(B) + e_H(A,V) - e_H(B,V) = 2|E| + |M_{A,V}| = 2(|A| - |B|). \tag{3.7.17}$$

Moreover,

$$e(H) = |M_{A,V}| + |E| = 2(|A| - |B|) - \ell \overset{(3.7.10)}{\leq} 2\rho n. \tag{3.7.18}$$

Suppose that $G[A]$ contains a matching M of size $\ell+1$. Then $\mathcal{P}^+ := M \cup M_{A,V}$ is a path system. If $\mathcal{P}^+$ contains a VW-path then obtain $\mathcal{P}$ from $\mathcal{P}^+$ by deleting an arbitrary edge of M. Otherwise there is an edge e in M which is incident to some edge in $M_{A,V}$. Let $\mathcal{P} := \mathcal{P}^+ \setminus \{e\}$. Then at least one endpoint of e is an endpoint of a VW-path in $\mathcal{P}$. In both cases, (iii) holds. Also (i) and (ii) hold by (3.7.17) and (3.7.18).

Therefore we may assume that $G[A]$ contains no matching of size $\ell+1$. Let $M^-, P = xyz$ be as guaranteed by Claim 2. Then $M_1 := M^- \cup \{xy\}$ and $M_2 := M^- \cup \{yz\}$ are both matchings of size ℓ in $G[A]$. For $i = 1, 2$, let $\mathcal{P}_i := M_i \cup M_{A,V}$. These are both path systems. Now (3.7.17) and (3.7.18) imply that both of $\mathcal{P}_1$ and $\mathcal{P}_2$ satisfy (i) and (ii). If, for some $i = 1, 2$, $\mathcal{P}_i$ also satisfies (iii) then we are done by setting $\mathcal{P} := \mathcal{P}_i$, so suppose not. Then for each $i = 1, 2$ there exists $M_i' \subseteq M_i$ such that $V(M_i') = V(M_{A,V}) \cap A$. In particular, this implies that $M_1', M_2' \subseteq M^-$. Pick any edge $e \in M_1'$ and let $\mathcal{P} := P \cup (M^- \setminus \{e\}) \cup M_{A,V}$. Then both endpoints of e are endpoints of a VW-path in $\mathcal{P}$, so (3.7.17) and (3.7.18) imply that $\mathcal{P}$ satisfies (i)–(iii). $\square$

3.7.5 The proof of Theorem 3.1.2

As already indicated at the beginning of the section, Theorem 3.1.2 now follows easily. Indeed, recall that we have a robust partition $\mathcal{V}$ with only five possible values of (k, ℓ). But Lemmas 3.7.1, 3.7.9 and 3.7.10 guarantee a $\mathcal{V}$-tour in each of these cases. Now Corollary 3.6.3 implies that G contains a Hamilton cycle.

Actually, we even prove the following stronger stability result of which Theorem 3.1.2 is an immediate consequence: if the degree of G is close to $n/4$ and G is not Hamiltonian, then G is either close to the union of four cliques, or the union of two complete bipartite graphs, or the first extremal example discussed in Subsection 3.1.2. This result will be very important in our proof of Theorem C in Chapter 4.

Theorem 3.7.11. *For every $\varepsilon, \tau > 0$ with $2\tau^{1/3} \leq \varepsilon$ and every non-decreasing function $g : (0,1) \to (0,1)$, there exists $n_0 \in \mathbb{N}$ such that the following holds. For all 3-connected D-regular graphs G on $n \geq n_0$ vertices where $D \geq (1/5+\varepsilon)n$, at least one of the following holds:*

(i) G has a Hamilton cycle;

(ii) $D < (1/4 + \varepsilon)n$ and there exist ρ, ν with $1/n_0 \leq \rho \leq \nu \leq \tau$; $1/n_0 \leq g(\rho)$; $\rho \leq g(\nu)$, and $(k, \ell) \in \{(4,0), (2,1), (0,2)\}$ such that G has a robust partition $\mathcal{V}$ with

parameters ρ, ν, τ, k, ℓ.

Proof. Let $\alpha := 1/5 + \varepsilon$. Choose a non-decreasing function $f : (0,1) \to (0,1)$ with $f(x) \leq \min\{x, g(x)\}$ for all $x \in (0,1)$ such that the requirements of Proposition 3.3.1 (applied with $r := 5$), Corollary 3.6.3 and Lemmas 3.7.1, 3.7.9 and 3.7.10 (each applied with τ' playing the role of τ) are satisfied whenever $n, \rho, \gamma, \nu, \tau'$ satisfy

$$\begin{aligned} 1/n \leq f(\rho), f(\gamma); \quad \rho \leq f(\nu), \varepsilon^3/8; \quad \gamma \leq f(\nu); \\ \nu \leq f(\tau'); \quad \tau' \leq f(\varepsilon), f(1/5), \tau \end{aligned} \tag{3.7.19}$$

(and so $\tau' \leq f(\alpha)$). Choose τ', τ'' such that $0 < \tau' \leq f(\varepsilon), f(1/5), \tau$ and let $\tau'' := f(\tau')$. Apply Theorem B with $f/36, \alpha, \tau''$ playing the roles of f, α, τ to obtain an integer n_0. Let G be a 3-connected D-regular graph on $n \geq n_0$ vertices where $D \geq \alpha n$. Theorem B now guarantees ρ, ν, k, ℓ with $1/n_0 \leq \rho \leq \nu \leq \tau''$, $1/n_0 \leq f(\rho)$ and $36\rho \leq f(\nu)$ such that G has a robust partition $\mathcal{V}$ with parameters $\rho, \nu, \tau'', k, \ell$ (and thus also a robust partition with parameters $\rho, \nu, \tau', k, \ell$).

Let $\gamma := 36\rho$. Note that $n, \rho, \gamma, \nu, \tau'$ satisfy (3.7.19). So we can apply Proposition 3.3.1(ii) with $\tau', 5$ playing the roles of τ, r to see that (k, ℓ) is equal to (a) $(k, 0)$ for $1 \leq k \leq 3$; (b) $(0,1)$; (c) $(1,1)$; or (d) $(4,0), (2,1), (0,2)$. Apply Lemmas 3.7.1, 3.7.9 and 3.7.10 (with τ' playing the role of τ) in the cases (a), (b), (c) respectively to obtain a $\mathcal{V}$-tour of G with parameter $36\rho = \gamma$. Then Corollary 3.6.3 (with τ' playing the role of τ) implies that G contains a Hamilton cycle so we are in case (i). If instead (d) holds, Proposition 3.3.1(i) implies that $D < (1/4 + \varepsilon)n$. Since $f \leq g$ and $\mathcal{V}$ is a robust partition with parameters ρ, ν, τ, k, ℓ (as $\tau' \leq \tau$) we are in case (ii). □

Proof of Theorem 3.1.2. Let $\varepsilon > 0$. Choose a positive constant τ such that $2\tau^{1/3} \leq \varepsilon$. Apply Theorem 3.7.11 (with $g(x) = x$, say) to obtain an integer n_0. Let G be a 3-connected D-regular graph on $n \geq n_0$ vertices with $D \geq (1/4 + \varepsilon)n$. Then Theorem 3.7.11 implies that G has a Hamilton cycle. □

3.8 The proofs of Theorems 3.1.4 and 3.1.5

We first show that Theorem 3.1.4 is approximately best possible.

Proposition 3.8.1. *Let $t, r \in \mathbb{N}$ be such that $r \geq 2$. Then there are infinitely many $n \in \mathbb{N}$ for which there exists a t-connected D-regular graph G on n vertices with $D := (n-t)/(r-1) - 1$ and circumference $c(G) \leq tn/(r-1) + t$.*

One can easily modify the construction to obtain a t-connected D-regular graph G with the same bound on $c(G)$ for smaller values of D (e.g. $D = n/r$).

Proof. We may suppose that $t \leq r-1$. Pick any $k \in \mathbb{N}$ with $k \geq 2t$. Let

$$n := (r-1)(2k(r-1)+1) + t \quad \text{and} \quad D := \frac{n-t}{r-1} - 1 = 2k(r-1).$$

Construct a graph G on n vertices as follows. Let $X, U_1, \ldots, U_{r-1}$ be a partition of $V(G)$, where $|X| = t$ and the $|U_i| = D+1$. Add all edges within the U_i. So $G[U_i]$ is D-regular. Let M_i be a matching in $G[U_i]$ with $|V(M_i)| = tD/(r-1)$. Note that M_i exists since $tD/(r-1)$ is even, and at most D since $t \leq r-1$. Add exactly one edge from each $y \in V(M_i)$ to X so that each $x \in X$ receives exactly $D/(r-1)$ edges from $V(M_i)$. Remove M_i from G.

Therefore G is t-connected (with vertex cut-set X) and D-regular. But any cycle in G traverses at most t of the U_i, so

$$c(G) \leq t|U_i| + |X| \leq tn/(r-1) + |X| = tn/(r-1) + t,$$

as required. □

The first part of the following proposition shows that the bound on the circumference in Theorem 3.1.5 is close to best possible. The second part of the proposition is a bipartite analogue of the extremal example in Figure 3.1(i).

Proposition 3.8.2.

(i) Let $t, r \in \mathbb{N}$ be such that $r \geq 4$ is even and $t \geq 2$. Then there are infinitely many $n \in \mathbb{N}$ for which there exists a t-connected D-regular bipartite graph G on n vertices with $D := (n-2)/(r-2)$ and circumference $c(G) \leq 2tn/(r-2) + t$;

(ii) For every $t \in \mathbb{N}$ with $t \geq 2$, there are infinitely many $D \in \mathbb{N}$ such that there exists a bipartite graph on $8D + 2$ vertices which is D-regular and t-connected but does not contain a Hamilton cycle.

Proof. To prove (i), Consider any $k \in \mathbb{N}$ with $k \geq 2t$. Let

$$n := \frac{k(r-2)^2}{2} + 2 \quad \text{and} \quad D := \frac{n-2}{r-2} = \frac{k(r-2)}{2}.$$

Let V be a set of n vertices and let $\{x, y\}, A_1, \ldots, A_{r/2-1}, B_1, \ldots, B_{r/2-1}$ be a partition of V, where $|A_i| = |B_i| = D$. Construct a D-regular graph G with $V(G) = V$ as follows. For each $1 \leq i \leq r/2 - 1$, add all edges between A_i and B_i. For all $1 \leq i \leq r/2 - 1$, choose a matching $M_i \in G[A_i, B_i]$ of size k.

For each $i \geq 2$, partition $B_i \cap V(M_i)$ into a set B_i' of size $k - t + 2$ and a set B_i'' of size $t - 2$. Add an edge between x and every vertex in B_i'. Add an edge between y and every vertex in $A_i \cap V(M_i)$. Remove M_i from G.

Choose $U \subseteq A_1$ such that $|U| = t - 2$ and $U \cap V(M_1) = \emptyset$. For each $u \in U$, let $N_u \subseteq B_1 \setminus V(M_1)$ be a collection of $r/2 - 2$ distinct neighbours of u. Choose the N_u to be disjoint. (This is possible since, for each $u \in U$, we have $d_{B_1 \setminus V(M_1)}(u) = D - k = k(r/2 - 2) \geq (t-2)(r/2-2)$.) For each $u \in U$ and all $i \geq 2$, add a single edge from u to a vertex of B_i'' such that every $b \in B_i''$ has exactly one neighbour in U. For each $u \in U$, remove every edge between u and every vertex in N_u. Add an edge between x and every vertex in $(B_1 \cap V(M_1)) \cup \bigcup_{u \in U} N_u$. Add an edge between y and every vertex in $A_1 \cap V(M_1)$. Remove M_1 from G. This completes the construction of G (see Figure 3.2 for an illustration of the case when $r = 10, t = 3$ (and with $U = \{a\}$)).

Note that G is bipartite and D-regular. It is not hard to see that $U \cup \{x, y\}$ is a vertex cut-set of minimal size. So G is t-connected. Let P be a path in G from $v_i \in (A_i \cup B_i) \setminus U$

to $v_j \in (A_j \cup B_j) \setminus U$, where $i \neq j$. Then there is a vertex $w \in U \cup \{x, y\}$ between v_i and v_j on P. Therefore any cycle C in G traverses at most t of the $A_i \cup B_i$, so $c(G) \leq 2tD + t \leq 2tn/(r-2) + t$, proving (i).

To obtain (ii), let $r := 10$ and consider G as in (i). Then $G \setminus (A_1 \cup \{x, y\})$ has $|B_1| + 3 > |A_1| + 2$ components. So G is not 1-tough so does not contain a Hamilton cycle. $\square$

One can easily modify the construction to obtain a t-connected D-regular graph G with the same bound on $c(G)$ for smaller values of D.

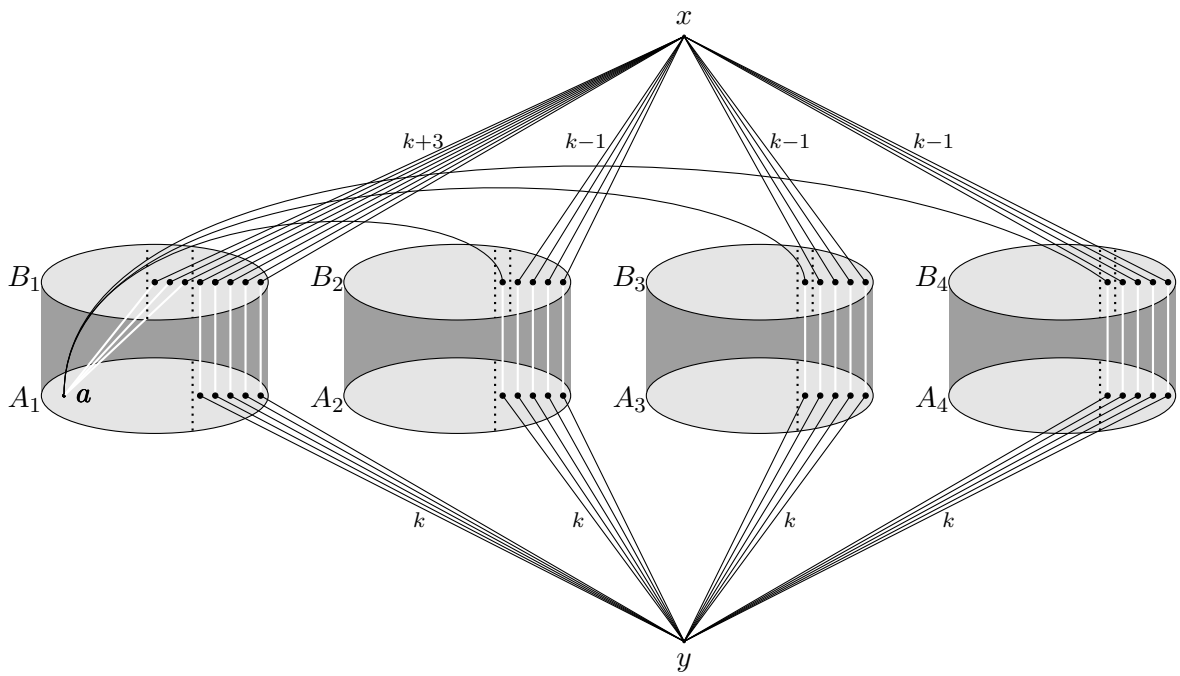

Figure 3.2: The graph G in Proposition 3.8.2(i) in the case when $r = 10$ and $t = 3$.

The proof of Theorem 3.1.4 uses robust partitions as the main tool (Theorem B). We show that, in a t-connected graph G with a robust partition, we can find a cycle that contains every vertex in the t largest robust components of G (or at least almost all the vertices in the case of bipartite robust components). When G has degree slightly larger than n/r, its robust partition contains at most $r-1$ components. So the t largest components together contain at least $tn/(r-1)$ vertices, as required.

We let C_1 denote a loop and C_2 a double edge. The following result shows that, given

any t-connected graph G and any collection $\mathcal{U}$ of t disjoint subsets of $V(G)$, we can find a path system $\mathcal{P}$ such that $R_{\mathcal{U}}(\mathcal{P}) \cong C_t$.

Proposition 3.8.3. *Let $t \in \mathbb{N}$, let G be a t-connected graph and let $\mathcal{U} := \{U_1, \ldots, U_t\}$ be a collection of disjoint vertex-subsets of G with $|U_i| \geq 2t$ for each $1 \leq i \leq t$. Then there exists a $\mathcal{U}$-anchored path system $\mathcal{P}$ in G such that $R_{\mathcal{U}}(\mathcal{P}) \cong C_t$.*

Proof. For each i, let $\mathcal{U}_i := \{U_1, \ldots, U_i\}$. Let P be a non-trivial path in G with both endpoints in U_1 and let $\mathcal{P}_1 := \{P\}$. Thus $R_{\mathcal{U}_1}(\mathcal{P}) \cong C_1$. Now suppose, for some $i < t$, we have obtained a $\mathcal{U}_i$-anchored path system $\mathcal{P}_i$ in G such that $R_{\mathcal{U}_i}(\mathcal{P}_i) \cong C_i$. Without loss of generality, we may assume that this cycle is $U_1U_2 \ldots U_i$. So $\mathcal{P}_i$ consists of i paths $P_1, \ldots, P_i$ where P_j has endpoints $x_j \in U_j, y_{j+1} \in U_{j+1}$ (with indices modulo i).

Suppose that there is some path $P_j \in \mathcal{P}_i$ with $|V(P_j) \cap U_{i+1}| \geq 2$. Let $u, v \in V(P_j) \cap U_{i+1}$ be distinct such that u is closer than v to x_j on P_j. Let $\mathcal{P}_{i+1}$ be the path system obtained from $\mathcal{P}_i$ be replacing P_j with the paths x_jP_ju, vP_jy_{j+1}.

So we may assume that $|V(\mathcal{P}_i) \cap U_{i+1}| = \sum_{1 \leq j \leq i} |V(P_j) \cap U_{i+1}| \leq i$. Let $U'_{i+1} := U_{i+1} \setminus V(\mathcal{P}_i)$. Note that $|U'_{i+1}| \geq 2t - i > t$. By Menger's Theorem, there exists a path system $\mathcal{R}$ consisting of $i+1$ paths which join $V(\mathcal{P}_i)$ to U'_{i+1} and have no internal vertices in $V(\mathcal{P}_i)$. By the pigeonhole principle, there exist $j \leq i$ and distinct paths $xRy, x'R'y' \in \mathcal{R}$ such that $x, x' \in V(P_j)$. Without loss of generality, x is closer to x_j on P_j than x'. Obtain $\mathcal{P}_{i+1}$ from $\mathcal{P}_i$ by replacing P_j with x_jP_jxRy, $y'R'x'P_jy_{j+1}$.

In both cases, $\mathcal{P}_{i+1}$ is a $\mathcal{U}_{i+1}$-anchored path system, and

$$R_{\mathcal{U}_{i+1}}(\mathcal{P}_{i+1}) = U_1 \ldots U_jU_{i+1}U_{j+1} \ldots U_i$$

is a cycle with vertex set $\mathcal{U}_{i+1}$. The path system $\mathcal{P}_t$ obtained in this way is as required in the proposition. $\square$

Now we show that, if $R_{\mathcal{U}}(\mathcal{P})$ is an Euler tour, we can discard suitable subpaths of each $P \in \mathcal{P}$ to ensure that $|V(\mathcal{P}) \cap U|$ is small for each $U \in \mathcal{U}$.

Proposition 3.8.4. *Let $\mathcal{U}$ be a collection of disjoint non-empty vertex-subsets of a graph G and let $\mathcal{P}$ be a $\mathcal{U}$-anchored path system in G containing t paths such that $R_{\mathcal{U}}(\mathcal{P})$ is an Euler tour. Then there exists a $\mathcal{U}$-anchored path system $\mathcal{P}'$ in G such that $R_{\mathcal{U}}(\mathcal{P}')$ is an Euler tour, and for each $U \in \mathcal{U}$ we have that $|V(\mathcal{P}') \cap U| \leq 2t$.*

Proof. Let $s := |\mathcal{U}|$. Clearly, the proposition holds if $s = 1$. So we may assume that $s \geq 2$ and that no $P \in \mathcal{P}$ has both endpoints in the same $X \in \mathcal{U}$ (otherwise we could remove P from $\mathcal{P}$). Fix a path $P \in \mathcal{P}$ with endpoints $u \in U, v \in V$ where $U, V \in \mathcal{U}$ are distinct. We will define a sequence of path systems $\mathcal{R}_\ell$ with $E(\mathcal{R}_\ell) \subseteq E(P)$ as follows. Let $\mathcal{R}_0 := \{P\}$. Suppose, for some $0 \leq \ell < s$, we have already defined a path system $\mathcal{R}_\ell$ such that

(α_ℓ) $\mathcal{R}_\ell$ is $\mathcal{U}$-anchored;

(β_ℓ) if $\ell \geq 1$ then $E(\mathcal{R}_\ell) \subseteq E(\mathcal{R}_{\ell-1})$;

(γ_ℓ) $E(R_{\mathcal{U}}(\mathcal{R}_\ell))$ forms a walk from U to V;

(δ_ℓ) for at least ℓ of the X in $\mathcal{U}$, $|X \cap V(\mathcal{R}_\ell)| \leq 2$.

Now we obtain $\mathcal{R}_{\ell+1}$ from $\mathcal{R}_\ell$ as follows. We are done if there are at least $\ell + 1$ sets X in $\mathcal{U}$ such that $|X \cap V(\mathcal{R}_\ell)| \leq 2$, so suppose not. Let $W \in \mathcal{U}$ be such that $|W \cap V(\mathcal{R}_\ell)| \geq 3$. By ($\gamma_\ell$), there exists an integer $p \geq 1$ such that $R_{\mathcal{U}}(\mathcal{R}_\ell)$ equals the walk $U_1U_2 \ldots U_{p+1}$ from $U_1 := U$ to $U_{p+1} := V$. So $\mathcal{R}_\ell$ consists of p paths $R_1, \ldots, R_p$ such that R_j has endpoints $x_j \in U_j$ and $y_{j+1} \in U_{j+1}$. Choose $j \leq j'$ such that $W \cap V(R_j)$ and $W \cap V(R_{j'})$ are both non-empty, and $j' - j$ is maximal with this property. Let $w \in W$ be the vertex on R_j which is closest to x_j and let $w' \in W$ be the vertex on $R_{j'}$ which is closest to $y_{j'+1}$. Let $\mathcal{R}_{\ell+1} := \{R_1, \ldots, R_{j-1}, x_jR_jw, w'R_{j'}y_{j'+1}, R_{j'+1}, \ldots, R_p\}$. Certainly $\mathcal{R}_{\ell+1}$ satisfies ($\beta_{\ell+1}$) and ($\delta_{\ell+1}$) from the construction. ($\alpha_{\ell+1}$) follows from (α_ℓ). Since w, w' lie in the same set in $\mathcal{U}$, ($\gamma_{\ell+1}$) holds by (γ_ℓ).

Therefore we can obtain $\mathcal{P}_P := \mathcal{R}_s$ that satisfies (α_s)–(δ_s). We can obtain $\mathcal{P}_P$ independently for each $P \in \mathcal{P}$. Since the P are vertex-disjoint and (β_s) implies that

$E(\mathcal{P}_P) \subseteq E(P)$, it follows that $\mathcal{P}' := \bigcup_{P\in\mathcal{P}} \mathcal{P}_P$ is a path system. Moreover $\mathcal{P}'$ is certainly $\mathcal{U}$-anchored by (α_s). We write $R := R_{\mathcal{U}}(\mathcal{P})$ and $R' := R_{\mathcal{U}}(\mathcal{P}')$. Note (γ_s) implies that one can obtain R' from R by replacing each edge UV of R with a walk joining U, V. Since R is an Euler tour we therefore have that R' is an Euler tour. Moreover, (δ_s) implies that for each $X \in \mathcal{U}$ we have $|V(\mathcal{P}') \cap X| = \sum_{P\in\mathcal{P}} |V(\mathcal{P}_P) \cap X| \leq 2t$ as required. □

In the following proposition, we show that, given a weak robust subpartition $\mathcal{U}$ in a t-connected graph G, we can adjust $\mathcal{U}$ slightly so that G contains a path system $\mathcal{P}$ which is a $\mathcal{U}$-tour. For this, we simply apply Propositions 3.8.3 and 3.8.4 to obtain a suitable $\mathcal{U}$-anchored path system and remove a small number of vertices from each bipartite robust component.

Proposition 3.8.5. *Let $t, n \in \mathbb{N}$ and let $0 < 1/n \ll \rho \ll \nu \ll \tau \ll \eta, 1/t \leq 1$. Suppose that G is a regular t-connected graph on n vertices. Let $\mathcal{U}$ be a weak robust subpartition in G with parameters $\rho, \nu, \tau, \eta, k, \ell$ where $k + \ell \leq t$. Then*

(i) G has a weak robust subpartition $\mathcal{X}$ with parameters $6\rho, \nu/2, 2\tau, \eta/2, k, \ell$;

(ii) $|\bigcup_{X\in\mathcal{X}} X| \geq |\bigcup_{U\in\mathcal{U}} U| - 2\rho\ell n$;

(iii) G contains an $\mathcal{X}$-tour with parameter 54ρ.

Proof. Write $\mathcal{U} = \{U_1, \ldots, U_k, Z_1, \ldots, Z_\ell\}$ satisfying (D1′)–(D5′). Apply Proposition 3.8.3 to $\mathcal{U}$ with $t' := k + \ell$ playing the role of t to obtain a $\mathcal{U}$-anchored path system $\mathcal{P}^*$ such that $R_{\mathcal{U}}(\mathcal{P}^*) \cong C_{t'}$. Since $\mathcal{P}^*$ contains at most t paths, we may apply Proposition 3.8.4 to $\mathcal{P}^*$ to obtain a $\mathcal{U}$-anchored path system $\mathcal{P}$ such that $R_{\mathcal{U}}(\mathcal{P})$ is an Euler tour and $|V(\mathcal{P}) \cap U| \leq 2t$ for all $U \in \mathcal{U}$.

Consider any $1 \leq j \leq \ell$. Let A_j, B_j be the bipartition of Z_j guaranteed by (D3′). So Z_j is a bipartite (ρ, ν, τ)-robust expander component with respect to A_j, B_j. Moreover,

$$2e_{\mathcal{P}}(A_j) + e_{\mathcal{P}}(A_j, \overline{Z_j}) \leq \sum_{x\in Z_j} d_{\mathcal{P}}(Z_j) \leq 2|V(\mathcal{P}) \cap Z_j| \leq 4t.$$

A similar inequality holds for B_j. Now $\Big||A_j| - |B_j|\Big| \leq \rho n$ by (D3′), (B1) and (C2). Therefore we can remove at most $\rho n + 4t \leq 2\rho n$ vertices from $Z_j \setminus V(\mathcal{P})$ to obtain $A'_j \subseteq A_j$, $B'_j \subseteq B_j$ and $Z'_j := A'_j \cup B'_j$ such that

$$2e_{\mathcal{P}}(A'_j) - 2e_{\mathcal{P}}(B'_j) + e_{\mathcal{P}}(A'_j, \overline{Z'_j}) - e_{\mathcal{P}}(B'_j, \overline{Z'_j}) = 2(|A'_j| - |B'_j|). \tag{3.8.1}$$

To see this, it suffices to check that $e_{\mathcal{P}}(A'_j, \overline{Z'_j}) - e_{\mathcal{P}}(B'_j, \overline{Z'_j})$ (and thus the left-hand side of (3.8.1)) is even. To verify the latter note that, modulo two,

$$e_{\mathcal{P}}(Z'_j, \overline{Z'_j}) \equiv \mathrm{End}_{\mathcal{P}}(Z'_j) = d_{R_{\mathcal{U}}(\mathcal{P})}(Z'_j).$$

So

$$e_{\mathcal{P}}(A'_j, \overline{Z'_j}) - e_{\mathcal{P}}(B'_j, \overline{Z'_j}) = e_{\mathcal{P}}(Z'_j, \overline{Z'_j}) - 2e_{\mathcal{P}}(B'_j, \overline{Z'_j}) \equiv d_{R_{\mathcal{U}}(\mathcal{P})}(Z'_j) - 2e_{\mathcal{P}}(B'_j, \overline{Z'_j}) \equiv 0$$

where the final congruence follows since $R_{\mathcal{U}}(\mathcal{P})$ is an Euler tour. Therefore (3.8.1) can be satisfied.

Let $\mathcal{X} := \{U_1, \ldots, U_k, Z'_1, \ldots, Z'_\ell\}$. Clearly (ii) holds. To prove (i), first note that for each $1 \leq j \leq \ell$, we have $|A'_j \triangle A_j| + |B'_j \triangle B_j| = |Z_j \setminus Z'_j| \leq 2\rho n$. Then Lemma 3.4.10(i) implies that Z'_j is a bipartite $(6\rho, \nu/2, 2\tau)$-robust expander component of G with bipartition A'_j, B'_j. So (D3′) holds. The remaining properties (D1′), (D2′), (D4′) and (D5′) are clear.

Finally, by (3.8.1) and the properties of $\mathcal{P}$ stated above, we can apply Lemma 3.7.8 with $\mathcal{X}, \mathcal{P}, 6\rho, \nu/2, 2\tau, \eta/2, k, \ell$ playing the roles of $\mathcal{U}, \mathcal{P}, \rho, \nu, \tau, \eta, k, \ell$ to obtain an $\mathcal{X}$-extension $\mathcal{P}'$ of $\mathcal{P}$ in G that is an $\mathcal{X}$-tour with parameter 54ρ. This proves (iii). □

We are now able to prove Theorem 3.1.4.

Proof of Theorem 3.1.4. Let $\alpha := 1/r + \varepsilon$ and $\eta := 1/2r^2 \leq \alpha^2/2$. Choose a non-decreasing function $f : (0,1) \to (0,1)$ with $f(x) \leq x$ for all $x \in (0,1)$ such that the requirements of Propositions 3.3.1, 3.6.1 and 3.8.5 as well as Lemma 3.6.2 are satisfied

whenever $n, \rho, \gamma, \nu, \tau$ satisfy the following:

$$1/n \leq f(\rho); \quad \rho \leq f(\nu), \varepsilon^3/8; \quad \nu \leq f(\tau); \quad \tau \leq f(\eta), f(1/t), f(1/r); \tag{3.8.2}$$

as well as $1/n \leq f(\gamma)$ and $\gamma \leq f(\nu)$. Choose τ, τ' so that

$$0 < \tau' \leq \tau \leq \frac{1}{2r^2}, \frac{\varepsilon}{2t}, \frac{\varepsilon^3}{8}, \frac{f(1/t)}{2}, \frac{f(\eta/2)}{2}, f(1/r) \quad \text{and } \tau' \leq f(\tau). \tag{3.8.3}$$

Choose a non-decreasing function $f' : (0,1) \to (0,1)$ such that $54f'(x) \leq f(x/2)$ for all $x \in (0,1)$. Apply Theorem B with f', α, τ' playing the roles of f, α, τ to obtain an integer n_0. Let G be a t-connected D-regular graph on $n \geq n_0$ vertices where $D \geq \alpha n$. Theorem B now guarantees ρ, ν, k', ℓ' with

$$1/n_0 \leq \rho \leq \nu \leq \tau', \quad 1/n_0 \leq f'(\rho) \quad \text{and} \quad \rho \leq f'(\nu) \tag{3.8.4}$$

such that G has a robust partition $\mathcal{V}$ with parameters $\rho, \nu, \tau', k', \ell'$ (and thus also with parameters $\rho, \nu, \tau, k', \ell'$). Note that (3.8.3) and (3.8.4) together imply that (3.8.2) holds. Moreover,

$$2\rho \leq 1/r^2 \quad \text{and} \quad 2\rho t \leq \varepsilon. \tag{3.8.5}$$

Claim. *There are integers k, ℓ with $k + \ell \leq t$ such that G has a weak robust subpartition $\mathcal{U}$ with parameters $\rho, \nu, \tau, \eta, k, \ell$ where*

$$\sum_{U \in \mathcal{U}} |U| \geq \min\left\{\frac{t}{r-1} + \frac{\ell}{r^2}, 1\right\} n. \tag{3.8.6}$$

Proof. Recall that $\mathcal{V}$ is a robust partition in G with parameters $\rho, \nu, \tau, k', \ell'$. Let $m := k' + \ell'$. Suppose first that $m \leq t$. Since by Proposition 3.6.1(i), $\mathcal{V}$ is a weak robust subpartition in G with parameters $\rho, \nu, \tau, \eta, k', \ell'$ we can take $\mathcal{U} := \mathcal{V}$ (and so $k = k'$ and $\ell = \ell'$). Therefore we may assume that $t \leq m - 1$. Order the

members of $\mathcal{V}$ as $X_1, \ldots, X_m$ so that $|X_1| \geq \ldots \geq |X_m|$. Let $\mathcal{U} := \{X_1, \ldots, X_t\}$. Now by Proposition 3.6.1(i) and (ii) there exist integers k, ℓ so that $k + \ell = t$ and $\mathcal{U}$ is a weak robust subpartition in G with parameters $\rho, \nu, \tau, \eta, k, \ell$.

By averaging, we have that $\sum_{U \in \mathcal{U}} |U| \geq tn/m$. Note also that $m + \ell \leq m + \ell' = k' + 2\ell' \leq r - 1$ where the last inequality follows from Proposition 3.3.1. Therefore

$$\sum_{U \in \mathcal{U}} |U| \geq \frac{tn}{m} \geq \frac{tn}{r-1-\ell} = \frac{tn}{r-1}\left(1 + \frac{\ell}{r-1-\ell}\right) \geq \frac{tn}{r-1} + \frac{\ell n}{r^2},$$

proving the claim. ■

Apply Proposition 3.8.5 to $G, \mathcal{U}$ to obtain a weak robust subpartition $\mathcal{X}$ with parameters $6\rho, \nu/2, 2\tau, \eta/2, k, \ell$ in G and an $\mathcal{X}$-anchored path system $\mathcal{P}$ such that $\sum_{X \in \mathcal{X}} |X| \geq \sum_{U \in \mathcal{U}} |U| - 2\rho\ell n$ and $\mathcal{P}$ is an $\mathcal{X}$-tour with parameter $\gamma := 54\rho$. Now (3.8.3) and (3.8.4) imply that

$$1/n \leq f(6\rho), f(\gamma); \quad 6\rho, \gamma \leq f(\nu/2); \quad \nu/2 \leq f(2\tau); \quad 2\tau \leq f(\eta/2). \tag{3.8.7}$$

Then Lemma 3.6.2 with $\mathcal{X}, \mathcal{P}, 6\rho, \gamma, \nu/2, 2\tau, \eta/2$ playing the roles of $\mathcal{U}, \mathcal{P}, \rho, \gamma, \nu, \tau, \eta$ implies that there is a cycle C in G which contains every vertex in $\bigcup_{X \in \mathcal{X}} X$. So

$$\begin{aligned} |V(C)| &\geq \sum_{U \in \mathcal{U}} |U| - 2\rho\ell n \overset{(3.8.6)}{\geq} \min\left\{\frac{t}{r-1} + \frac{\ell}{r^2} - 2\rho\ell, 1 - 2\rho\ell\right\} n \\ &\overset{(3.8.5)}{\geq} \min\left\{\frac{t}{r-1}, 1 - \varepsilon\right\} n, \end{aligned}$$

as required. □

Proof of Theorem 3.1.5 (Sketch). The proof is almost the same as that of Theorem 3.1.4. We proceed similarly as we did there to obtain a robust partition $\mathcal{V}$ with parameters $\rho, \nu, \tau', k', \ell'$. Since G is bipartite, it is easy to check that $k' = 0$. Thus $\ell' \leq \lfloor (r-1)/2 \rfloor = (r-2)/2$ by Proposition 3.3.1. Instead of the claim in the proof of Theorem 3.1.4, we now show that there exists an integer $\ell \leq t$ such that G has a weak robust subpartition

$\mathcal{U}$ with parameters $\rho, \nu, \tau, \eta, 0, \ell$ where $\sum_{U \in \mathcal{U}} |U| \geq \min\{2tn/(r-2), n\}$. (Using that $\ell' \leq (r-2)/2$, this follows as in the claim.) The remainder of the proof is now similar to that of Theorem 3.1.4. $\square$

CHAPTER 4

SOLUTION TO A PROBLEM OF BOLLOBÁS AND HÄGGKVIST ON HAMILTON CYCLES IN REGULAR GRAPHS

4.1 Introduction

In this chapter we give an exact solution to a longstanding conjecture on Hamilton cycles in regular graphs, posed independently by Bollobás and Häggkvist:

Theorem C. *There exists $n_0 \in \mathbb{N}$ such that every 3-connected D-regular graph on $n \geq n_0$ vertices with $D \geq n/4$ is Hamiltonian.*

The content of this chapter leads on from Chapter 3; the background to and motivation for the problem were discussed in detail in Section 3.1.2, so we do not repeat this here.

In Chapter 3, we proved Theorem 3.1.2, an approximate version of Theorem C, namely that for all $\varepsilon > 0$, whenever n is sufficiently large, any 3-connected D-regular graph on n vertices with $D \geq (1/4 + \varepsilon)n$ is Hamiltonian. In fact we proved a stronger result (Theorem 3.7.11), which we will use in this chapter to prove Theorem C. Recall that the major tool in proving Theorems 3.1.2 and 3.7.11 was a structural decomposition result (Theorem B) which holds for any dense regular graph: it gives a partition into (bipartite) robust expanders with few edges between these (see Section 4.3). We proved further applications of this partition result in Chapter 3.

This chapter is organised as follows. In Section 4.2, we discuss the extremal examples which show that Theorem C is best possible. Section 4.3 contains a sketch of the proof of Theorem C. The proof of Theorem C is split into three cases, and these are considered in Sections 4.5–4.7 respectively. Finally, we derive Theorem C in Section 4.8.

4.2 The extremal examples

In this section we show that Theorem C is best possible in the sense that neither the degree condition nor the connectivity condition can be reduced. An example of Jung [68] and Jackson, Li and Zhu [66] shows that the degree condition cannot be reduced for graphs with $n \equiv 1 \mod 8$ vertices; for completeness we extend this to all possible n in the following proposition. An illustration of their example may be found in Figure 4.1(i).

Proposition 4.2.1. *Let $n \geq 5$ and let D be the largest integer such that $D \leq \lceil n/4 \rceil - 1$ and nD is even. Then there is an $(\lfloor n/8 \rfloor - 1)$-connected D-regular graph G_n on n vertices which does not contain a Hamilton cycle.*

Proof. Recall that a D-regular graph on n vertices exists if and only if $n \geq D+1$ and nD is even. For each $n \geq 5$, we define a graph G_n on n vertices as follows. Let V_1, V_2, A, B be disjoint independent sets where $|A| = D$, $|B| = D - 1$, and the other classes have sizes according to the table below. Let A_1, A_2 be a partition of A so that $\Big|D/2 - |A_1|\Big|$ is minimal subject to the parity conditions below being satisfied:

n	D	$\lvert V_1\rvert$	$\lvert V_2\rvert$	$\lvert A_1\rvert$	$\lvert A_2\rvert$
$8k+1$	$2k$	$2k+1$	$2k+1$	even	even
$8k+2$	$2k$	$2k+2$	$2k+1$	even	even
$8k+3$	$2k$	$2k+2$	$2k+2$	even	even
$8k+4$	$2k$	$2k+3$	$2k+2$	even	even
$8k+5$	$2k$	$2k+3$	$2k+3$	even	even
$8k+6$	$2k+1$	$2k+3$	$2k+2$	odd	even
$8k+7$	$2k$	$2k+4$	$2k+4$	even	even
$8k+8$	$2k+1$	$2k+4$	$2k+3$	even	odd

Note that $|V_i| \geq D+1$ for $i = 1, 2$. Add every edge between A and B. First consider the cases when $D = 2k$. Then $|A_i|$ is even for $i = 1, 2$. For each $i = 1, 2$, add edges so that $G_n[V_i]$ is D-regular. Let M_i be a matching of size $|A_i|/2$ in $G_n[V_i]$ and remove it. Let $V_i' := V(M_i)$. So $|V_i'| = |A_i|$. Add a perfect matching between V_i' and A_i.

Now consider the case when $D = 2k+1$. Then, by our choice of A_i and V_i we have that $|A_i| \equiv |V_i| \mod 2$. Fix $V_i' \subseteq V_i$ with $|V_i'| := |A_i|$. Define the edge set of $G_n[V_i]$ so that for all $x \in V_i'$ we have $d_{V_i}(x) = D-1$ and for all $y \in V_i \setminus V_i'$ we have $d_{V_i}(y) = D$. Add a perfect matching between V_i' and A_i.

Then G_n has n vertices, is D-regular and has connectivity $\min\{|A_1|, |A_2|\} \geq \lfloor n/8 \rfloor - 1$. Moreover, G_n does not contain a Hamilton cycle because it is not 1-tough ($G_n \setminus A$ contains more than $|A|$ components). $\square$

There also exist non-Hamiltonian 2-connected regular graphs on n vertices with degree close to $n/3$ (see Figure 4.1(ii)). Indeed, we can construct such a graph G as follows. Start with three disjoint cliques on $3k$ vertices each. In the ith clique choose disjoint sets A_i and B_i with $|A_i| = |B_i|$ and $|A_1| = |A_3| = k$ and $|A_2| = k-1$. Remove a perfect matching between A_i and B_i for each i. Add two new vertices a and b, where a is connected to all vertices in the sets A_i and b is connected to all vertices in all the sets

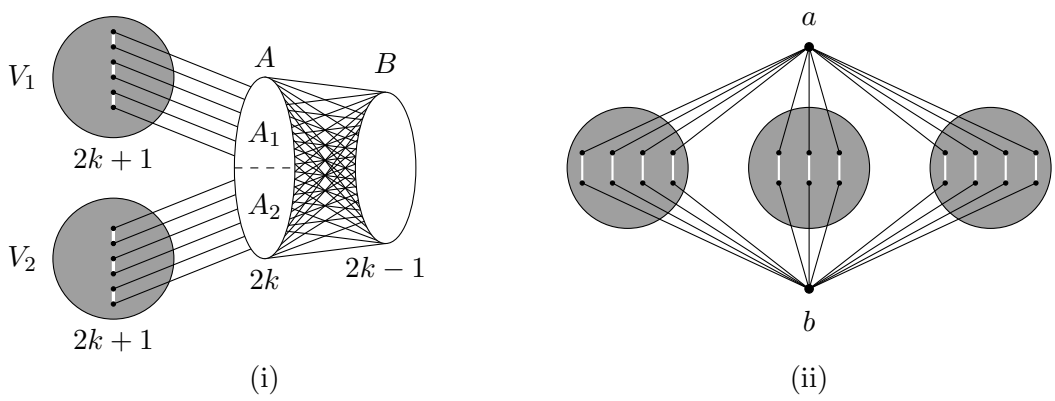

Figure 4.1: Extremal examples for Theorem C. (i) is an illustration for the case $n = 8k+1$. Here, each V_i is a clique of order $2k+1$ with a matching of size k removed.

B_i. Then G is a $(3k-1)$-regular 2-connected graph on $n = 9k+2$ vertices. However, G is not Hamiltonian because $G \setminus \{a, b\}$ has three components. One can construct similar examples for all $n \in \mathbb{N}$.

Altogether this shows that none of the conditions — degree or connectivity — of Theorem C can be relaxed.

4.3 Sketch of the proof

4.3.1 Robust partitions of dense regular graphs

The main tool in our proof is Theorem B, a structural result on dense regular graphs that was the main result of Chapter 3. Roughly speaking, this allows us to partition the vertex set of such a graph G into a small number of 'robust components', each of which has strong expansion properties and sends few edges to the rest of the graph.

Theorem B roughly says the following:

(♣) For all $r \in \mathbb{N}$ and $\varepsilon > 0$ and n sufficiently large, every D-regular graph G on n vertices with $D \geq (\frac{1}{r+1} + \varepsilon)n$ has a robust partition with parameters k, ℓ, where $k + 2\ell \leq r$.

In particular, the number of edges between robust components is $o(n^2)$.

4.3.2 Finding a Hamilton cycle using a robust partition

Now suppose that G is a D-regular graph on n vertices with $D \geq n/4$, where n is sufficiently large. Then (♣) applied with $r = 4$ implies that G has a robust partition $\mathcal{V}$ with parameters k, ℓ, where $k+2\ell \leq 4$. This gives eight possible structures, parametrised by $(k, \ell) \in S_{\leq 3} \cup S_4$, where

$$S_{\leq 3} := \{(1,0),(2,0),(3,0),(0,1),(1,1)\} \quad \text{and } S_4 := \{(4,0),(0,2),(2,1)\}.$$

Note that the extremal example in Figure 4.1(i) corresponds to the case $(2,1)$ and the one in (ii) corresponds to the case $(3,0)$. Also note that when $D \geq (1/4+\varepsilon)n$, we have $k+2\ell \leq 3$ and so $(k,\ell) \in S_{\leq 3}$. In Chapter 3, we proved that if G is 3-connected and has a robust partition $\mathcal{V}$ with parameters k, ℓ where $(k,\ell) \in S_{\leq 3}$, then G is Hamiltonian. In particular, this implies an approximate version of Theorem C. The proof proceeded by considering each possible structure separately. Therefore, to prove Theorem C, it remains to show that if G is 3-connected and has a robust partition $\mathcal{V}$ with parameters k, ℓ where $(k,\ell) \in S_4$, then G is Hamiltonian (see Theorem 3.7.11). So the current chapter uses the result of Chapter 3 as an essential ingredient. Again, we consider each structure separately in Sections 4.5, 4.6 and 4.7 respectively.

In each case we adopt the following strategy. Let $\mathcal{V}$ be a robust partition of G with parameters k, ℓ. Kühn, Osthus and Treglown [86] proved that every large robust expander H with linear minimum degree contains a Hamilton cycle. This can be strengthened (see Corollary 3.6.8) to show that one can cover all the vertices of a robust expander with a set of paths with prescribed endvertices. More precisely, one can show that each robust expander component $G[V_i]$ is Hamilton p-linked for each small p and all $1 \leq i \leq k$. (Here a graph H is *Hamilton p-linked* if, whenever $X := \{x_1, y_1, \ldots, x_p, y_p\}$ is a collection of distinct vertices, there exist vertex-disjoint paths $P_1, \ldots, P_p$ such that P_j connects x_j to y_j, and such that together the paths $P_1, \ldots, P_p$ cover all vertices of H.) Balanced bipartite robust expanders have the same property, provided X is distributed equally

between the bipartition classes. This means that we can hope to reduce the problem of finding a Hamilton cycle in G to finding a suitable set of *external edges* E_{ext}, where an edge is external if it has endpoints in different members of $\mathcal{V}$. We then apply the Hamilton p-linked property to each robust component to join up the external edges into a Hamilton cycle. The assumption of 3-connectivity is crucial for finding E_{ext}.

However, several problems arise. When $(k, \ell) = (4, 0)$, we have four robust components and only the assumption of 3-connectivity, which makes it difficult to find a suitable set E_{ext} joining all four components directly. However, we can appeal to the dominating cycle result in [66] mentioned in the introduction to Chapter 3, giving us a fairly short argument for this case. Note that the condition that $D \geq n/4$ is essential in this case — 3-connectivity on its own is not sufficient.

Now suppose that $\ell \geq 1$, i.e. $\mathcal{V}$ contains a bipartite robust expander component. These cases are challenging since a bipartite graph does not contain a Hamilton cycle if it is not balanced. So as well as a suitable set E_{ext}, we need to find a set E_{bal} of *balancing edges* incident to the bipartite robust expander component. Suppose for example that $(k, \ell) = (0, 2)$ and G consists of two bipartite robust expander components W_1, W_2 such that W_i has vertex classes A_i, B_i where $|A_1| = |B_1|$ and $|A_2| = |B_2| + 1$. Then we could choose E_{bal} to be a single edge with both endpoints in A_2. A second example would be $E_{\text{bal}} = \{a_1a_2, b_1a_2'\}$ where $a_1 \in A_1$, $b_1 \in B_1$ and $a_2, a_2' \in A_2$ are distinct. (Note that these are also external edges and in this case we can actually take $E_{\text{ext}} \cup E_{\text{bal}} = \{a_1a_2, b_1a_2'\}$.) Observe that we need at least $\Big||A_1| - |B_1|\Big| + \Big||A_2| - |B_2|\Big|$ balancing edges.

Our robust partition guarantees that the vertex classes of any bipartite robust expander component differ by at most $o(n)$, so we must potentially find a similar number of balancing edges. This must be done in such a way that $\mathcal{P} := E_{\text{ext}} \cup E_{\text{bal}}$ can be extended into a Hamilton cycle. So in particular $\mathcal{P}$ must be a collection of vertex-disjoint paths. We use the Hamilton p-linkedness of the (bipartite) robust expander components to find these edges which extend $\mathcal{P}$ into a Hamilton cycle. Consider the second example above, with $\mathcal{P} = \{a_1a_2, b_1a_2'\}$. Choose a neighbour b_2 of a_2 in B_2 and let $\mathcal{P}' := \{a_1a_2b_2, b_1a_2'\}$.

Then the Hamilton 1-linkedness of W_1, W_2 implies that we can find a path P_1 with endpoints a_1, b_1 which spans W_1, and a path P_2 with endpoints a_2', b_2 which spans $W_2 \setminus \{a_2\}$. Then the edges of $P_1, P_2, \mathcal{P}'$ together form a Hamilton cycle.

It turns out that the condition that $D \geq n/4$ is crucial in the case when $(k, \ell) = (2, 1)$ (see Section 4.2) but its full strength is not required in the case when $(k, \ell) = (0, 2)$. A sketch of the proof in each of the three cases can be found at the beginning of Sections 4.5, 4.6 and 4.7 respectively.

4.4 Notation and tools

4.4.1 General notation

All notation is as in Chapter 3. For general notation, we refer the reader to Section 3.2. We will also need the following basic concepts:

If S, T are sets of vertices which are not necessarily disjoint and may not be subsets of $V(G)$, we write $e_G(S)$ for the number of edges of G with both endpoints in S, and $e_G(S, T)$ for the number of ST-edges of G, i.e. for the number of all edges with one endpoint in S and the other endpoint in T. We also set $G[S] := G[S \cap V(G)]$. Moreover, when S, T are disjoint, we write $G[S, T]$ for the bipartite graph with vertex classes $S \cap V(G)$, $T \cap V(G)$ whose edge set consists of all the ST-edges of G. We omit the subscript G whenever the graph G is clear from the context.

Given disjoint subsets X, Y of $V(G)$, we say that P is an *XY-path* if P has one endpoint in X and one endpoint in Y. We call a vertex-disjoint collection of non-trivial paths a *path system*. We will often think of a path system $\mathcal{P}$ as a graph with edge set $\bigcup_{P \in \mathcal{P}} E(P)$, so that e.g. $V(\mathcal{P})$ is the union of the vertex sets of each path in $\mathcal{P}$, and $e_{\mathcal{P}}(X)$ denotes the number of edges on the paths in $\mathcal{P}$ having both endpoints in X. By slightly abusing notation, given two vertex sets S and T and a path system $\mathcal{P}$, we write $\mathcal{P}[S]$ for the graph obtained from $\mathcal{P}[S]$ by deleting isolated vertices and define $\mathcal{P}[S, T]$

similarly. We say that a vertex x is an *endpoint* of $\mathcal{P}$ if x is an endpoint of some path in $\mathcal{P}$. An *Euler tour* in a (multi)graph is a closed walk that uses each edge exactly once.

We write $\mathbb{N}$ for the set of positive integers and write $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. $\mathbb{R}_{\geq 0}$ denotes the set of non-negative reals.

Given $0 < \varepsilon < 1$ and $x \in \mathbb{R}$, we write $\lceil x \rceil_\varepsilon := \lceil x - \varepsilon \rceil$.

In addition, we will need the following specialised notation that was defined in various parts of Chapter 3. For convenience we list below the sections in which the relevant definitions are stated.

Section 3.3: *(ν, τ)-robust expander*, *bipartite (ν, τ)-robust expander (with bipartition A, B)*, *ρ-component*, *ρ-close to bipartite (with bipartition U_1, U_2)*, (C1)–(C3), *(ρ, ν, τ)-robust expander component*, (E1), (E2), *bipartite (ρ, ν, τ)-robust expander component (with bipartition A, B)*, (B1), (B2), *(ρ, ν, τ)-robust component*, *robust partition of G with parameters ρ, ν, τ, k, ℓ*, (D1)–(D7), the statement of Theorem B.

Section 3.6: *weak robust subpartition with parameters $\rho, \nu, \tau, \eta, k, \ell$*, *reduced multigraph $R_{\mathcal{V}}(\mathcal{P})$ of $\mathcal{P}$ with respect to $\mathcal{V}$*, *$\mathcal{V}$-tour with parameter γ*, (T1)–(T4).

4.5 (4,0): Four robust expander components

The aim of this section is to prove the following lemma.

Lemma 4.5.1. *Let $D, n \in \mathbb{N}$ and $0 < 1/n \ll \rho \ll \nu \ll \tau \ll 1$. Suppose that G is a 3-connected D-regular graph on n vertices with $D \geq n/4$. Suppose further that G has a robust partition $\mathcal{V}$ with parameters $\rho, \nu, \tau, 4, 0$. Then G contains a $\mathcal{V}$-tour with parameter $33/n$.*

We will find a $\mathcal{V}$-tour $\mathcal{P}$ as follows. Let $\mathcal{V} := \{V_1, \ldots, V_4\}$. Suppose that there are $1 \leq i < j \leq 4$ such that $G[V_i, V_j]$ contains a large matching M. We can use 3-connectivity with the tripartition $\mathcal{V}' := \mathcal{V} \cup \{V_i \cup V_j\} \setminus \{V_i, V_j\}$ to obtain a path system $\mathcal{P}'$ such that $R_{\mathcal{V}'}(\mathcal{P}')$ is a $\mathcal{V}'$-tour. Then $\mathcal{P}'$ together with some suitable edges of M will form a $\mathcal{V}$-tour.

Suppose instead that for all $1 \leq i < j \leq 4$, every matching in $G[V_i, V_j]$ is small. In this case, we appeal to the result of Jackson, Li and Zhu [66] mentioned in the introduction to Chapter 3: any longest cycle in G is dominating. Thus C visits all the V_i. Moreover, since there are very few edges between the V_i it follows that most of the edges of C lie within some V_i. If we remove all such edges, what remains is a $\mathcal{V}$-tour.

Before proceeding, we make a small remark. The result of [66] allows us to avoid a potentially intricate case analysis in the case when every matching between components is small, but this could conceivably be done 'by hand'. So it seems likely that Lemma 4.5.1 could be proved without appealing to [66].

Let $\mathcal{V}'$ be a partition of $V(G)$ into three parts such that $\mathcal{V}$ is a refinement of $\mathcal{V}'$. Then, by Lemma 3.7.3, we can easily find a collection of paths $\mathcal{P}'$ such that $R_{\mathcal{V}'}(\mathcal{P}')$ is an Euler tour. The following result will enable us to 'extend' $\mathcal{P}'$ into $\mathcal{P}$ such that $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour.

Proposition 4.5.2. *Let $\mathcal{U}$ be a partition of $V(G)$. Let $U, V \in \mathcal{U}$ and let $\mathcal{U}' := \mathcal{U} \cup \{U \cup V\} \setminus \{U, V\}$. Suppose that G contains a path system $\mathcal{P}'$ such that $R_{\mathcal{U}'}(\mathcal{P}')$ is an Euler tour. Suppose further that $G[U, V]$ contains a matching M of size at least $|V(\mathcal{P}') \cap (U \cup V)| + 2$. Then G contains a path system $\mathcal{P}$ with $E(\mathcal{P}) \supseteq E(\mathcal{P}')$ such that $R_{\mathcal{U}}(\mathcal{P})$ is an Euler tour and $|V(\mathcal{P}) \cap X| \leq |V(\mathcal{P}') \cap X| + 2$ for all $X \in \mathcal{U}$.*

Proof. Note that there are at least two edges e, e' of M which are vertex-disjoint from $\mathcal{P}'$. Let $R' := R_{\mathcal{U}}(\mathcal{P}')$ and $R'' := R_{\mathcal{U}'}(\mathcal{P}')$. We have that $d_{R'}(U) + d_{R'}(V) = d_{R''}(U \cup V)$ is even since R'' is an Euler tour. Moreover, $d_{R'}(X) = d_{R''}(X)$ for all $X \in \mathcal{U}' \cap \mathcal{U}$.

If both $d_{R'}(U)$ and $d_{R'}(V)$ are odd, let $\mathcal{P} := \mathcal{P}' \cup \{e\}$. Otherwise, both $d_{R'}(U)$ and $d_{R'}(V)$ are even (but one could be zero). In this case, let $\mathcal{P} := \mathcal{P}' \cup \{e, e'\}$. It is straightforward to check that in both cases $R_{\mathcal{U}}(\mathcal{P})$ is an Euler tour. $\square$

A subgraph H of a graph G is said to be *dominating* if $G \setminus V(H)$ is an independent set. In our proof of Lemma 4.5.1 we will use the following theorem of Jackson, Li and Zhu.

Theorem 4.5.3. *[66] Let G be a 3-connected D-regular graph on n vertices with $D \geq n/4$. Then any longest cycle in C is dominating.*

Proof of Lemma 4.5.1. Let C be a longest cycle in G. Then Theorem 4.5.3 implies that C is dominating. We consider two cases according to the number of edges in C between classes of $\mathcal{V}$.

Case 1. $e_C(U,V) \geq 12$ *for some distinct* $U, V \in \mathcal{V}$.

Since C is a cycle we have that $\Delta(C[U,V]) \leq 2$. König's theorem implies that $C[U,V]$ has a proper edge-colouring with at most two colours, and thus $C[U,V]$ contains a matching of size at least $e_C(U,V)/2 \geq 6$.

Let $\mathcal{V}' := \mathcal{V} \cup \{U \cup V\} \setminus \{U,V\}$. So $\mathcal{V}'$ is a tripartition of $V(G)$, and certainly $|V| \geq 3$ for each $V \in \mathcal{V}'$. Apply Lemma 3.7.3 to obtain a path system $\mathcal{P}'$ in G such that (i)–(iii) hold. Then $R_{\mathcal{V}'}(\mathcal{P}')$ is an Euler tour and (iii) implies that $|V(\mathcal{P}') \cap X| \leq 4$ for all $X \in \mathcal{V}'$.

Now Proposition 4.5.2 with $\mathcal{V}, \mathcal{V}'$ playing the roles of $\mathcal{U}, \mathcal{U}'$ implies that G contains a path system $\mathcal{P}$ such that $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour, and $|V(\mathcal{P}) \cap X| \leq 6$ for all $X \in \mathcal{V}$. So $\mathcal{P}$ is a $\mathcal{V}$-tour with $6/n$ playing the role of γ.

Case 2. $e_C(U,V) \leq 11$ *for all distinct* $U, V \in \mathcal{V}$.

Let $\mathcal{P}$ be the collection of disjoint paths with edge set $E(C) \setminus \bigcup_{V \in \mathcal{V}} E(C[V])$. For each $V \in \mathcal{V}$, let $\mathcal{P}_V := \bigcup_{U \in \mathcal{V} \setminus \{V\}} \mathcal{P}[U,V]$. Then

$$e(\mathcal{P}_V) = \sum_{U \in \mathcal{V} \setminus \{V\}} e_C(U,V) \leq 33. \tag{4.5.1}$$

Suppose that $|V(C) \cap V| < D - 2\rho^{1/3}n$. Let $X := V \setminus V(C)$. So X is an independent set in G. Moreover, (D7) implies that, for all but at most ρn vertices in $x \in V$, we have $d_V(x) \geq D - \rho n$. In particular, $|V| \geq D - \rho n$ and so $|X| \geq \rho^{1/3}n$. Thus there is some $x \in X$ such that $d_V(x) \geq D - \rho n$. Therefore x has a neighbour in X, a contradiction.

Thus $|V(C) \cap V| \geq D - 2\rho^{1/3}n$ for all $V \in \mathcal{V}$. But

$$2|V(C) \cap V| = \sum_{v \in V} d_C(v) = 2e_C(V) + e(\mathcal{P}_V)$$

and hence

$$e_C(V) = |V(C) \cap V| - \frac{1}{2}e(\mathcal{P}_V) \geq D - 2\rho^{1/3}n - 33/2 > 0.$$

Thus $E(C[V]) \neq \emptyset$ for all $V \in \mathcal{V}$. It is straightforward to check that this implies that $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour. Finally, note that, for each $V \in \mathcal{V}$, (4.5.1) implies that we have $|V(\mathcal{P}) \cap V| \leq e(\mathcal{P}_V) \leq 33$. So $\mathcal{P}$ is a $\mathcal{V}$-tour with parameter $33/n$. □

4.6 (0,2): Two bipartite robust expander components

The aim of this section is to prove the following lemma.

Lemma 4.6.1. *Let $D, n \in \mathbb{N}$, let $0 < 1/n \ll \rho \ll \nu \ll \tau \ll \alpha < 1$ and let $D \geq \alpha n$. Suppose that G is a 3-connected D-regular graph on n vertices and that $\mathcal{V}$ is a robust partition of G with parameters $\rho, \nu, \tau, 0, 2$. Then G contains a $\mathcal{V}$-tour with parameter $\rho^{1/3}$.*

We first give a brief outline of the argument.

4.6.1 Sketch of the proof of Lemma 4.6.1

Let $\mathcal{V} := \{W_1, W_2\}$ be as above and let A_i, B_i be a bipartition of W_i such that $G[W_i]$ is a bipartite robust expander component with bipartition A_i, B_i, where $|A_i| \geq |B_i|$. To prove Lemma 4.6.1, our aim is to find a 'balancing' path system $\mathcal{P}$ to which we can apply Lemma 3.7.8 and hence obtain a $\mathcal{V}$-tour. In other words, the path system has to 'compensate for' the differences in the sizes of the vertex classes A_i and B_i and has to 'join up' W_1 and W_2.

One could try to first find a path system which balances W_1, and then add additional edges so that W_2 is also balanced; however these additional edges may cause W_1 to become unbalanced. So one must find a path system $\mathcal{P}$ which simultaneously balances both components.

This is not too difficult if both A_1 and A_2 contain sufficiently large matchings M_1 and M_2 (see Lemma 4.6.5). In this case, the 3-connectivity of G guarantees a matching of size two connecting W_1 and W_2, to which we can add suitable edges from M_1 and M_2 to obtain $\mathcal{P}$.

So suppose that this is not the case. Then (see Lemmas 4.6.4 and 4.6.12) we show that we can choose $C_i \in \{A_i, B_i\}$ for each $i = 1, 2$ such that König's theorem on edge-colourings guarantees the following: $G[C_1], G[C_2], G[W_1, A_2]$ contain matchings $M_1, M_2, M_{1,2}$ respectively, such that the union $\mathcal{R}$ of these matchings balances both W_1 and W_2. However, two problems can arise: $\mathcal{R}$ may not connect W_1 and W_2 (it could contain no W_1W_2-path) and it may contain cycles.

Therefore the bulk of the proof of Lemma 4.6.1 is devoted to choosing M_1, M_2 and $M_{1,2}$ carefully to avoid these problems. Observe that since we use König's theorem to find matchings, we can actually find much larger matchings in $H \subseteq G$ when $\Delta(H)$ is small, and choosing a 'good' matching is easier. So most of the difficulty in the proof arises from the presence of vertices of high degree.

4.6.2 Balanced subgraphs with respect to a partition

Consider a graph G with vertex partition $\mathcal{V} := \{W_1, W_2\}$, where W_i has bipartition A_i, B_i for $i = 1, 2$. Write $\mathcal{V}^*$ for the ordered partition (A_1, B_1, A_2, B_2). Given $D \in \mathbb{N}$, we say that G is *D-balanced (with respect to $\mathcal{V}^*$)* if both of the following hold.

$$2e(A_1) - 2e(B_1) + e(A_1, W_2) - e(B_1, W_2) = D(|A_1| - |B_1|); \tag{4.6.1}$$

$$2e(A_2) - 2e(B_2) + e(A_2, W_1) - e(B_2, W_1) = D(|A_2| - |B_2|).$$

Proposition 3.7.4(i) easily implies that any D-regular graph with arbitrary ordered partition $\mathcal{V}^*$ is D-balanced.

Proposition 4.6.2. *Suppose that G is a D-regular graph and let A_1, B_1, A_2, B_2 be a partition of $V(G)$. Then G is D-balanced with respect to (A_1, B_1, A_2, B_2).* □

The next proposition shows that, to prove Lemma 4.6.1, it suffices to find a path system $\mathcal{P}$ which is 2-balanced with respect to $\mathcal{V}^*$, contains a W_1W_2-path, and does not have many edges.

Proposition 4.6.3. *Let $n, D \in \mathbb{N}$ and $0 < 1/n \ll \rho \leq \gamma \ll \nu \ll \tau \ll \alpha < 1$. Let G be a D-regular graph on n vertices with $D \geq \alpha n$. Suppose further that G has a robust partition $\mathcal{V} := \{W_1, W_2\}$ with parameters $\rho, \nu, \tau, 0, 2$. For each $i = 1, 2$, let A_i, B_i be the bipartition of W_i guaranteed by* (D3)*. Let $\mathcal{P}$ be a 2-balanced path system with respect to (A_1, B_1, A_2, B_2) in G. Suppose that $e(\mathcal{P}) \leq \gamma n$ and that $\mathcal{P}$ contains at least one W_1W_2-path. Then G contains a $\mathcal{V}$-tour with parameter 18γ.*

Proof. Let p be the number of W_1W_2-paths in $\mathcal{P}$. Any W_1W_2-path in $\mathcal{P}$ contains an odd number of W_1W_2-edges. Since $\mathcal{P}$ is 2-balanced with respect to (A_1, B_1, A_2, B_2), we have that $e_{\mathcal{P}}(W_1, W_2) = e_{\mathcal{P}}(A_1, W_2) - e_{\mathcal{P}}(B_1, W_2) + 2e_{\mathcal{P}}(B_1, W_2)$ is even. Hence p is even. Since $p > 0$, we have that $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour.

The hypothesis $e(\mathcal{P}) \leq \gamma n$ implies that $|V(\mathcal{P}) \cap V| \leq 2\gamma n$ for all $V \in \mathcal{V}$. Proposition 3.6.1 implies that $\mathcal{V}$ is a weak robust partition with parameters $2\gamma, \nu, \tau, \alpha^2/2, 0, 2$. Thus we can apply Lemma 3.7.8 with $\mathcal{V}, 0, 2, W_j, A_j, B_j, \mathcal{P}, 2\gamma$ playing the roles of $\mathcal{U}, k, \ell, W_j, A_j, B_j, \mathcal{P}, \rho$ to find a $\mathcal{V}$-tour $\mathcal{P}'$ with parameter 18γ. □

The next lemma shows that we can find a D-balanced subgraph of G which only contains edges in some of the parts of G. (Recall the definition of $\lceil \cdot \rceil_\varepsilon$ from the end of Subsection 4.4.)

Lemma 4.6.4. *Let $D \in \mathbb{N}$ be such that $D \geq 20$. Let G be a graph and let $\mathcal{V}^* := (A_1, B_1, A_2, B_2)$ be an ordered partition of $V(G)$ with $0 \leq |A_i| - |B_i| \leq D/2$ for $i = 1, 2$.*

Suppose that $e_G(A_1, B_2) \leq e_G(B_1, A_2)$ and $\Delta(G[A_i]) \leq D/2$ for $i = 1, 2$. Suppose further that G is D-balanced with respect to $\mathcal{V}^*$. Then one of the following holds:

(i) for $i = 1, 2$, $G[A_i]$ contains a matching M_i of size $|A_i| - |B_i| \leq \lceil e_G(A_i)/5 \rceil_{1/4}$;

(ii) there exists a spanning subgraph G' of G which is D-balanced with respect to $\mathcal{V}^*$ and $E(G') \subseteq E(G[C_1]) \cup E(G[C_2]) \cup E(G[A_1 \cup B_1, A_2])$, where $C_1 \in \{A_1, B_1\}$ and $C_2 \in \{A_2, B_2\}$.

Proof. Observe that the graph obtained by removing $E(G[A_i, B_i])$ from G for $i = 1, 2$ is D-balanced. So we may assume that $E(G[A_i, B_i]) = \emptyset$ for $i = 1, 2$. Consider each of the pairs

$$\{G[A_1], G[B_1]\}, \{G[A_2], G[B_2]\}, \{G[A_1, A_2], G[B_1, B_2]\}, \{G[A_1, B_2], G[B_1, A_2]\}$$

of induced subgraphs. For each such pair $\{J, J'\}$, remove $\min\{e_G(J), e_G(J')\}$ arbitrary edges from each of J, J' in G. Let H be the subgraph obtained from G in this way. Then H is D-balanced and for each pair $\{J, J'\}$, we have that $E(H[V(J)]) = \emptyset$ whenever $e_G(J) \leq e_G(J')$ (and vice versa). In particular, $e_H(A_1, B_2) = 0$. Suppose that we cannot take $G' := H$ so that (ii) holds. Then $H \subseteq G[C_1] \cup G[C_2] \cup G[B_1, A_2 \cup B_2]$ for some $C_1 \in \{A_1, B_1\}$ and $C_2 \in \{A_2, B_2\}$ with $e_H(B_1, B_2) \geq 1$. So $e_H(A_1, A_2) = 0$. Let $v_i := D(|A_i| - |B_i|) \geq 0$. Since H is D-balanced we have that $2e_H(A_1) - 2e_H(B_1) - e_H(B_1, A_2 \cup B_2) = v_1 \geq 0$. In particular, $e_H(A_1) \geq e_H(B_1)$. So $e_H(B_1) = 0$. Let $t := e_H(B_1, A_2)$. Thus

$$\begin{aligned} &2e_H(A_1) \geq v_1 + t + 1 \quad \text{and similarly} \\ &2e_H(A_2) \geq v_2 - t + 1. \end{aligned} \tag{4.6.2}$$

Suppose first that $t \geq v_2$. Then $2e_H(A_1) \geq v_1 + v_2 + 1$. Since G is D-balanced, summing the two equations in (4.6.1) implies that $v_1 + v_2$ is even. Let $H_{B_1A_2}$ consist of v_2 arbitrary edges in $H[B_1, A_2]$ and let H_{A_1} consist of $(v_1 + v_2)/2$ arbitrary edges in $H[A_1]$.

In this case, we let $G' := H_{A_1} \cup H_{B_1A_2}$. So (ii) holds.

Suppose instead that $t < v_2$. First consider the case when $t = 0$. Then (4.6.2) implies that $2e_G(A_i) \geq 2e_H(A_i) \geq v_i + 1$ for $i = 1, 2$. Since $\Delta(G[A_i]) \leq D/2$, Vizing's theorem implies that $G[A_i]$ contains a matching M_i of size

$$\left\lceil \frac{e_G(A_i)}{D/2+1} \right\rceil \geq \left\lceil \frac{D(|A_i| - |B_i|)/2}{D/2+1} \right\rceil \geq |A_i| - |B_i| - \lfloor D/(D+2) \rfloor = |A_i| - |B_i|.$$

Note that the right hand side is at most $\lceil e(A_i)/5 \rceil_{1/4}$. So (i) holds.

Therefore we may assume that $t > 0$. Recall that $v_1 \equiv v_2 \mod 2$. We will choose $H_{B_1A_2} \subseteq H[B_1, A_2]$ and $H_{A_i} \subseteq H[A_i]$ for $i = 1, 2$ by arbitrarily choosing edges according to the relative parities of v_1 and t, such that the following hold:

- if $v_1 + t$ is even then choose $e(H_{B_1A_2}) = t$, $2e(H_{A_1}) = v_1 + t$, $2e(H_{A_2}) = v_2 - t$;
- if v_1+t is odd then choose $e(H_{B_1A_2}) = t-1$, $2e(H_{A_1}) = v_1+t-1$, $2e(H_{A_2}) = v_2-t+1$.

These choices are possible by (4.6.2). We let $G' := H_{A_1} \cup H_{A_2} \cup H_{B_1A_2}$. Observe that G' is D-balanced. So (ii) holds. □

Observe that the subgraph $M_1 \cup M_2$ of G guaranteed by Lemma 4.6.4(i) is a 2-balanced path system. The next lemma shows that, when G is 3-connected, one can modify such a path system into one which also contains paths between $A_1 \cup B_1$ and $A_2 \cup B_2$.

Lemma 4.6.5. *Let $n, D \in \mathbb{N}$ and $0 < 1/n \ll \gamma \ll 1$. Let G be a 3-connected D-regular graph on n vertices. Let W_1, W_2 be a partition of $V(G)$ and let A_i, B_i be a partition of W_i for $i = 1, 2$, where $|A_i| \geq |B_i|$. Suppose that there exist matchings M_1, M_2 in $G[A_1], G[A_2]$ respectively so that $|A_i| - |B_i| = e(M_i) \leq \lceil e(A_i)/5 \rceil_{1/4}$ and $e(M_i) \leq \gamma n$ for $i = 1, 2$. Then G contains a path system $\mathcal{P}$ which is 2-balanced with respect to (A_1, B_1, A_2, B_2) and contains a W_1W_2-path, and $e(\mathcal{P}) \leq 3\gamma n$.*

Proof. Proposition 4.6.2 implies that G is D-balanced with respect to (A_1, B_1, A_2, B_2). Suppose that there exist edges $e \in E(G[A_1, A_2])$ and $e' \in E(G[B_1, B_2])$. Then we can

take $\mathcal{P} := M_1 \cup M_2 \cup \{e, e'\}$. We are similarly done if there exist edges $f \in E(G[A_1, B_2])$ and $f' \in E(G[B_1, A_2])$. If either of these two hold then we say that G contains a *balanced matching*. So we may assume that G does not contain a balanced matching. The 3-connectivity of G implies that there is a matching N of size at least three in $G[W_1, W_2]$. Since G does not contain a balanced matching, $e_N(C_1, C_2) \geq 2$ for some $C_i \in \{A_i, B_i\}$. So we can choose a matching N' of size two in $G[C_1, C_2]$. Let D_i be such that $\{C_i, D_i\} := \{A_i, B_i\}$. Note that $e_G(D_1, D_2) = 0$ or G would contain a balanced matching. Without loss of generality, we may assume that $e(M_1) \leq e(M_2)$.

Case 1. $e(M_2) > 0$.

Note that $1 \leq e(M_2) \leq e_G(A_2)/5 + 3/4$. Thus $e_G(A_2) - e(M_2) \geq 4e_G(A_2)/5 - 3/4 > 0$. So we can always choose an edge $e_2 \in E(G[A_2]) \setminus E(M_2)$. If possible, let f_2 be the edge of M_2 spanned by $V(N') \cap A_2$. If there is no such edge, let f_2 be an arbitrary edge in M_2. Let

$$M_2' := \begin{cases} M_2 \setminus \{f_2\} & \text{if } C_2 = A_2 \\ M_2 \cup \{e_2\} & \text{if } C_2 = B_2. \end{cases}$$

Case 1.a. $e(M_1) > 0$.

Define e_1, f_1 and hence M_1' analogously to e_2, f_2, M_2'. It is straightforward to check that $\mathcal{P} := N' \cup M_1' \cup M_2'$ is as required in the lemma.

Case 1.b. $e(M_1) = 0$.

We have $|A_1| = |B_1|$. Without loss of generality we may suppose that $C_1 = A_1$ or we can swap A_1, B_1. So $e_G(A_1, W_2) \geq e_N(C_1, C_2) \geq 2$. Since G is D-balanced and $e_G(B_1, C_2) = e_G(B_1, W_2)$, this in turn implies that $2e_G(B_1) + e_G(B_1, C_2) \geq 2$. If $e_G(B_1) > 0$ let $e \in E(G[B_1])$ be arbitrary and define $\mathcal{P} := N' \cup M_2' \cup \{e\}$. Otherwise, there exists $e_{12} \in E(G[B_1, C_2])$. Let $e_{12}' \in E(N')$ be vertex-disjoint from e_{12}. If possible, let $f_2' \in E(M_2)$ be the edge spanning the endpoints of e_{12}, e_{12}' which lie in A_2; otherwise, let $f_2' \in E(M_2)$ be arbitrary. If $C_2 = A_2$, let $\mathcal{P} := M_2 \cup \{e_{12}, e_{12}'\} \setminus \{f_2'\}$. If $C_2 = B_2$, let

$\mathcal{P} := M_2 \cup \{e_{12}, e'_{12}\}$. It is straightforward to check that in all cases $\mathcal{P}$ is as required in the lemma.

Case 2. $e(M_2) = 0$.

So $e(M_1) = 0$ and $|A_i| = |B_i|$ for $i = 1, 2$. Without loss of generality, we may assume that $C_i := A_i$ (and hence $D_i := B_i$). Write $\{i, j\} = \{1, 2\}$. Since G is D-balanced we have that

$$2e_G(A_i) - 2e_G(B_i) + e_G(A_i, A_j) + e_G(A_i, B_j) - e_G(B_i, A_j) = 0.$$

So $2e_G(B_i) + e_G(B_i, A_j) \geq e_N(A_1, A_2) \geq 2$. Therefore either $e_G(B_i) > 0$ or $e_G(B_i, A_j) > 0$ (or both). So for $i = 1, 2$, either we can find $e_i \in E(G[B_i])$ or $e_{ij} \in E(G[B_i, A_j])$ (or both). Note that not both $e_G(B_1, A_2), e_G(A_1, B_2)$ can be positive since G does not contain a balanced matching.

Suppose that $e_G(B_1), e_G(B_2) > 0$. Let $\mathcal{P} := N' \cup \{e_1, e_2\}$, as required. Otherwise we may assume without loss of generality that $e_G(B_1) > 0$ and $e_G(B_2, A_1) > 0$. Let $e'_{12} \in N'$ be vertex-disjoint from e_{21}. Let $\mathcal{P} := \{e_1, e'_{12}, e_{21}\}$. It is straightforward to check that in both cases $\mathcal{P}$ is as required in the lemma. □

4.6.3 Tools for finding matchings

Given any bipartite graph G, König's theorem on edge-colourings guarantees that we can find a matching of size at least $\lceil e(G)/\Delta(G) \rceil$. The following lemma shows that, given any matching M in G, we can find a matching M' of at least this size such that $V(M) \subseteq V(M')$.

Lemma 4.6.6. *Let G be a bipartite graph with vertex classes V, W such that $\Delta(G) \leq \Delta$. Let M be a matching in G with $e(M) \leq \lceil e(G)/\Delta \rceil$. Then there exists a matching M' in G such that $e(M') = \lceil e(G)/\Delta \rceil$ and $V(M) \subseteq V(M')$.*

Proof. Let M' be a matching in G such that $V(M) \subseteq V(M')$ and $e(M') \leq \lceil e(G)/\Delta \rceil$ is

maximal with this property. Suppose that $e(M') < \lceil e(G)/\Delta \rceil$. Since, by König's theorem on edge-colourings, G contains a matching of size $\lceil e(G)/\Delta \rceil$, this means that M' is not a maximum matching. So, by Berge's lemma, G contains an augmenting path P for M', i.e. a path with endpoints not in $V(M')$ which alternates between edges in $E(M')$ and edges outside of $E(M')$. But then $P \setminus E(M')$ is a matching contradicting the maximality of $e(M')$. $\square$

We now show that given a bipartite graph $G = (U, Z)$ and any partition V, W of Z, we can find a large matching in G which has the 'right' density in each of $G[U, V]$ and $G[U, W]$.

Lemma 4.6.7. *Let G be a bipartite graph with vertex classes $U, V \cup W$, where V, W are disjoint. Suppose that $\Delta(G) \leq \Delta$. Let b_V, b_W be non-negative integers such that $b_V + b_W \leq \lceil e(G)/\Delta \rceil$, $b_V \leq \lceil e_G(U, V)/\Delta \rceil$ and $b_W \leq \lceil e_G(U, W)/\Delta \rceil$. Then G contains a matching M such that $e_M(U, V) = b_V$ and $e_M(U, W) = b_W$.*

Proof. By increasing b_V, b_W if necessary, we may assume that $b_V + b_W = \lceil e(G)/\Delta \rceil$. Note that either $b_V = \lceil e_G(U, V)/\Delta \rceil$, or $b_W = \lceil e_G(U, W)/\Delta \rceil$, or both. Suppose without loss of generality that $b_V = \lceil e_G(U, V)/\Delta \rceil$. Choose a matching M' in G of size $\lceil e(G)/\Delta \rceil$. Let $m_V := e_{M'}(U, V)$ and let $m_W := e_{M'}(U, W)$. Let $k := b_V - m_V$. Then

$$m_W = \lceil e(G)/\Delta \rceil - m_V = b_V + b_W - m_V = b_W + k.$$

If $k = 0$ we are done, so suppose first that $k > 0$. Apply Lemma 4.6.6 to obtain a matching J_V in $G[U, V]$ such that $e(J_V) = b_V$ and $V(J_V) \supseteq V(M'[U, V])$. So $|(V(J_V) \setminus V(M'[U, V])) \cap U| = k$. Thus we can choose a submatching J_W of $M'[U, W]$ of size $m_W - k = b_W$ that is vertex-disjoint from J_V. Let $M := J_V \cup J_W$.

Otherwise, $k < 0$. Apply Lemma 4.6.6 to obtain a matching J_W in $G[U, W]$ such that $e(J_W) = b_W$ and $V(J_W) \supseteq V(M'[U, W])$. As above, we can choose a submatching J_V of $M'[U, V]$ of size b_V that is vertex-disjoint from J_W. Let $M := J_V \cup J_W$. $\square$

4.6.4 Acyclic unions of matchings

The next lemma shows that, in a graph with low maximum degree, we can find a large matching that does not completely span a given set of vertices.

Proposition 4.6.8. *Let $0 < 1/\Delta \ll \eta \ll 1$. Let G be a graph with $\Delta(G) \leq \eta\Delta$ and suppose that $e(G) \geq 2\eta\Delta$. Suppose that $K \subseteq V(G)$. Then there exists a matching M in G such that $e(M) = \lceil e(G)/\Delta \rceil$ and $M[K]$ is not a perfect matching.*

Proof. By Vizing's theorem, G contains a matching M' of size

$$\left\lceil \frac{e(G)}{\Delta(G)+1} \right\rceil \geq \left\lceil \frac{e(G)}{3\eta\Delta/2} \right\rceil \geq \left\lceil \frac{e(G)}{\Delta} \right\rceil + 1.$$

Delete edges so that M' has size $\lceil e(G)/\Delta \rceil + 1$. If M' contains an edge with both endpoints in K, remove this edge to obtain M. Otherwise, obtain M from M' by removing an arbitrary edge. □

Proposition 4.6.8 and the following observation will be used to guarantee that, given a matching M in $G[W_1, A_2]$, we can find a suitable matching N in $G[A_2]$ such that the path system $M \cup N$ contains a W_1A_2-path.

Fact 4.6.9. *Let G be a graph with vertex partition U, V and let M be a non-empty matching between U and V. Let $K := V(M) \cap V$ and let M' be a matching in $G[V]$ such that $M'[K]$ is not a perfect matching. Then $M \cup M'$ is a path system containing a UV-path.*

Given a graph G with low maximum degree, vertex partition U, V and a non-empty matching M in $G[U, V]$, the next lemma shows that we can find matchings in $G[U], G[V]$ which extend M into a path system $\mathcal{P}$ containing a UV-path.

Lemma 4.6.10. *Let $0 < 1/\Delta \ll \eta \ll 1$. Let G be a graph with partition U, V and suppose that $\Delta(G) \leq \eta\Delta$. Let M be a matching between U and V. Suppose further that $e_G(U) \leq e_G(V) \leq \eta\Delta^2$. Then there exist matchings M_U, M_V in $G[U], G[V]$ respectively such that*

(i) $\mathcal{P} := M \cup M_U \cup M_V$ *is a path system;*

(ii) $e(M_U) \leq \lceil e_G(U)/\Delta \rceil$ *with equality if* $e_G(U) \geq \sqrt{\eta}\Delta$*; and* $e(M_V) \leq \lceil e_G(V)/\Delta \rceil$ *with equality if* $e_G(V) \geq \sqrt{\eta}\Delta$*;*

(iii) if $M \neq \emptyset$*, then* $\mathcal{P}$ *contains a* UV*-path.*

Proof. If $M = \emptyset$ then Vizing's theorem implies that we can find matchings M_U, M_V of size $\lceil e_G(U)/\Delta \rceil$, $\lceil e_G(V)/\Delta \rceil$ respectively. Then (i)–(iii) hold. So we may assume that $M \neq \emptyset$. If $e_G(U) \leq e_G(V) < \sqrt{\eta}\Delta$, then we are done by taking $M_U, M_V := \emptyset$. Suppose instead that $e_G(U) < \sqrt{\eta}\Delta \leq e_G(V)$. Apply Proposition 4.6.8 with $G[V], V(M) \cap V$ playing the roles of G, K to obtain a matching M_V in $G[V]$ such that $e(M_V) = \lceil e_G(V)/\Delta \rceil$ and $M_V[V(M) \cap V]$ is not a perfect matching. Fact 4.6.9 implies that we are done by taking $M_U = \emptyset$.

Therefore we may assume that $\sqrt{\eta}\Delta \leq e_G(U) \leq e_G(V)$. Apply Proposition 4.6.8 with $G[U], V(M) \cap U$ playing the roles of G, K to obtain a matching M_U in $G[U]$ of size $\lceil e_G(U)/\Delta \rceil$ such that $M_U[V(M) \cap U]$ is not a perfect matching. Let $\mathcal{P}_U$ be the path system with edge set $E(M) \cup E(M_U)$. So Fact 4.6.9 implies that $\mathcal{P}_U$ contains at least one UV-path P. Let $u_0 \in U$ and $v_0 \in V$ be the endpoints of P. Let Y be the set of all those vertices in V which are endpoints of a VV-path in $\mathcal{P}_U$. Now

$$|Y| \leq 2e(M_U) = 2\lceil e_G(U)/\Delta \rceil \leq 2\lceil e_G(V)/\Delta \rceil. \tag{4.6.3}$$

Obtain G' from $G[V]$ by removing every edge incident with $Y \cup \{v_0\}$. So

$$e(G') \geq e_G(V) - \eta\Delta(|Y| + 1) \overset{(4.6.3)}{\geq} (1 - 4\sqrt{\eta})e_G(V) \geq e_G(V)/2.$$

So G' contains a matching of size

$$\lceil e(G')/(\eta\Delta + 1) \rceil \geq \lceil e(G')/2\eta\Delta \rceil \geq \lceil e_G(V)/4\eta\Delta \rceil \geq \lceil e_G(V)/\Delta \rceil.$$

Let M_V be an arbitrary submatching of this matching of size $\lceil e_G(V)/\Delta \rceil$. Let $\mathcal{P} := M \cup M_U \cup M_V$.

Clearly (ii) holds. Observe that $\mathcal{P}$ has a UV-path, namely P. Hence (iii) holds. To show (i), it is enough to show that $\mathcal{P}$ is acyclic. Suppose not and let C be a cycle in $\mathcal{P}$. Now C contains at least one edge $e \in E(M_V)$. Then both endpoints of this edge belong to Y, and hence $e \notin E(G')$, a contradiction. $\square$

The following is a version of Lemma 4.6.10 for sparse graphs which may have a small number of vertices with high degree.

Lemma 4.6.11. *Let $0 < 1/\Delta \ll \rho \ll 1$. Let G be a graph with vertex partition U, V and suppose that $\Delta(G[U]), \Delta(G[V]) \leq \Delta$. Let M be a matching between U and V such that $e(M) \leq \rho\Delta$. Suppose further that $e_G(U), e_G(V) \leq \rho\Delta^2$. Then, for any integers $0 \leq a_U \leq \lceil e_G(U)/\Delta \rceil_{1/4}$ and $0 \leq a_V \leq \lceil e_G(V)/\Delta \rceil_{1/4}$, G contains a path system $\mathcal{P}$ such that*

(i) $\mathcal{P}[U, V] = M$ and both of $\mathcal{P}[U]$,$\mathcal{P}[V]$ are matchings;

(ii) $e_{\mathcal{P}}(U) = a_U$, $e_{\mathcal{P}}(V) = a_V$;

(iii) if $M \neq \emptyset$, then $\mathcal{P}$ contains a UV-path.

Proof. By removing edges in $G[U]$ and $G[V]$ we may assume without loss of generality that $a_U = \lceil e_G(U)/\Delta \rceil_{1/4}$ and $a_V = \lceil e_G(V)/\Delta \rceil_{1/4}$. Choose η with $\rho \ll \eta \ll 1$. Let $U' := \{u \in U : d_U(u) \geq \eta\Delta\}$ and define V' analogously. Then $2e_G(U) \geq \sum_{u \in U'} d_U(u) \geq |U'|\eta\Delta$ and similarly for V', so

$$|U'|, |V'| \leq \sqrt{\rho}\Delta. \tag{4.6.4}$$

Let $U_0 := U \setminus U'$ and $V_0 := V \setminus V'$. Let H be the graph with vertex set $V(G)$ and edge set $E(G[U_0]) \cup E(G[V_0]) \cup M$. So $E_H(U) = E_G(U_0)$ and $E_H(V) = E_G(V_0)$. Moreover, $\Delta(H) \leq 2\eta\Delta$. Note that

$$e_G(U_0) \geq e_G(U) - \Delta|U'| \quad \text{and} \quad e_G(V_0) \geq e_G(V) - \Delta|V'|. \tag{4.6.5}$$

Assume without loss of generality that $e_G(U_0) \leq e_G(V_0)$. Apply Lemma 4.6.10 with $H, M, U, V, 2\eta$ playing the roles of G, M, U, V, η to obtain matchings M_{U_0}, M_{V_0} in $H[U_0] = G[U_0], H[V_0] = G[V_0]$ respectively such that $\mathcal{P}_0 := M \cup M_{U_0} \cup M_{V_0}$ is a path system satisfying Lemma 4.6.10(i)–(iii). So $\mathcal{P}_0$ contains a UV-path if $M \neq \emptyset$. Moreover, $e(M_{U_0}) \leq \lceil e_G(U_0)/\Delta \rceil$ with equality if $e_G(U_0) \geq \sqrt{2\eta}\Delta$, and $e(M_{V_0}) \leq \lceil e_G(V_0)/\Delta \rceil$ with equality if $e_G(V_0) \geq \sqrt{2\eta}\Delta$. Thus

$$|V(\mathcal{P}_0)| \leq 2e(\mathcal{P}_0) \leq 2\left(e(M) + \lceil e_G(U)/\Delta \rceil + \lceil e_G(V)/\Delta \rceil\right) \leq \sqrt{\rho}\Delta. \tag{4.6.6}$$

For every $u \in U'$ and $v \in V'$ we have that

$$d_{U_0 \setminus V(\mathcal{P}_0)}(u), d_{V_0 \setminus V(\mathcal{P}_0)}(v) \overset{(4.6.6)}{\geq} \eta\Delta/2 \overset{(4.6.4)}{>} |U'|, |V'|.$$

So for each $u \in U'$, we may choose a distinct neighbour $w_u \in U_0 \setminus V(\mathcal{P}_0)$ of u. Let $M_{U'} := \{uw_u : u \in U'\} \subseteq G[U', U_0 \setminus V(\mathcal{P}_0)]$. Define a matching $M_{V'}$ in $G[V', V_0 \setminus V(\mathcal{P}_0)]$ (which covers V') similarly.

Let $\mathcal{P} := \mathcal{P}_0 \cup M_{U'} \cup M_{V'}$. Note that $\mathcal{P}$ is a path system since $\mathcal{P}_0$ is. Certainly $\mathcal{P}[U, V] = \mathcal{P}_0[U, V] = M$, so (i) holds. Suppose that $e_G(U_0) \geq \sqrt{2\eta}\Delta$. Then

$$\begin{aligned} e_{\mathcal{P}}(U) = e(M_{U_0}) + e(M_{U'}) &= \lceil e_G(U_0)/\Delta \rceil + |U'| \overset{(4.6.5)}{\geq} \lceil e_G(U)/\Delta - |U'| \rceil + |U'| \\ &= \lceil e_G(U)/\Delta \rceil \geq \lceil e_G(U)/\Delta \rceil_{1/4}. \end{aligned}$$

Suppose instead that $e_G(U_0) < \sqrt{2\eta}\Delta$. Then

$$e_{\mathcal{P}}(U) \geq |U'| \overset{(4.6.5)}{\geq} \lceil e_G(U)/\Delta - \sqrt{2\eta} \rceil \geq \lceil e_G(U)/\Delta \rceil_{1/4}$$

since $\sqrt{2\eta} < 1/4$. Analogous statements are true for $e_{\mathcal{P}}(V)$. So by removing edges in $e_{\mathcal{P}}(U), e_{\mathcal{P}}(V)$ if necessary, we may assume that (ii) holds. Note that $\mathcal{P}$ has a UV-path if $\mathcal{P}_0$ does (there is a one-to-one correspondence between the UV-paths in $\mathcal{P}$ and the

UV-paths in $\mathcal{P}_0$). $\square$

4.6.5 Rounding

Given a small collection of reals which sum to an integer, the following lemma shows that we can suitably round these reals so that their sum is unchanged. Lemmas 4.6.7 and 4.6.11 together enable us to find three matchings, one in each of $G[W_1], G[W_2]$ and $G[W_1, W_2]$, each of which is not too large, such that their union is a path system $\mathcal{P}$. Lemma 4.6.12 will allow us to choose the size of each matching correctly, so that $\mathcal{P}$ is 2-balanced.

Lemma 4.6.12. *Let $0 < \varepsilon < 1/2$. Let $a_1, a_2, b, c \in \mathbb{R}$ with $b, c \geq 0$ and let $x_1, x_2 \in \mathbb{N}_0$. Suppose that*

$$2a_1 + b - c = 2x_1 \quad \textit{and} \quad 2a_2 + b + c = 2x_2.$$

Then there exist integers a_1', a_2', b', c' such that

$$2a_1' + b' - c' = 2x_1 \quad \textit{and} \quad 2a_2' + b' + c' = 2x_2,$$

where $0 \leq b' \leq \lceil b \rceil$, $0 \leq c' \leq \lceil c \rceil$, $b' + c' \leq \lceil b + c \rceil$; and for $i = 1, 2$, $|a_i'| \leq \lceil |a_i| \rceil_\varepsilon$; and finally $a_i' \geq 0$ if and only if $a_i \geq 0$.

Proof. Note that

$$\lfloor 2a_1 \rfloor + \lceil b - c \rceil = 2x_1 \quad \text{and} \quad \lfloor 2a_2 \rfloor + \lceil b + c \rceil = 2x_2. \tag{4.6.7}$$

In particular, either $\lfloor 2a_1 \rfloor$, $\lceil b - c \rceil$ are both odd, or both even. The same is true for the pair $\lfloor 2a_2 \rfloor, \lceil b + c \rceil$. Let $A_i := \lfloor 2a_i \rfloor / 2$ for $i = 1, 2$. Let also

$$B := \frac{\lceil b + c \rceil + \lceil b - c \rceil}{2} \quad \text{and} \quad C := \frac{\lceil b + c \rceil - \lceil b - c \rceil}{2}.$$

Observe that $\{A_1, A_2, B, C\} \subseteq \mathbb{Z} \cup (\mathbb{Z} + 1/2)$. Let $i \in \{1, 2\}$. Suppose first that $a_i \geq 0$

(and so $A_i \geq 0$). If $a_i - \lfloor a_i \rfloor \leq \varepsilon$ then $2\lceil a_i \rceil_\varepsilon = 2\lfloor a_i \rfloor = \lfloor 2a_i \rfloor = 2A_i$. If $a_i - \lfloor a_i \rfloor > \varepsilon$ then $2\lceil a_i \rceil_\varepsilon = 2\lceil a_i \rceil \geq \lfloor 2a_i \rfloor = 2A_i$. Therefore $\lceil A_i \rceil \leq \lceil a_i \rceil_\varepsilon$. Suppose now that $a_i < 0$ (and so $A_i < 0$). If $a_i - \lfloor a_i \rfloor < 1 - \varepsilon$ then $2\lfloor a_i + \varepsilon \rfloor = 2\lfloor a_i \rfloor \leq \lfloor 2a_i \rfloor = 2A_i$. If $a_i - \lfloor a_i \rfloor \geq 1 - \varepsilon$ then $2\lfloor a_i + \varepsilon \rfloor = 2\lfloor a_i \rfloor + 2 = \lfloor 2a_i \rfloor + 1 = 2A_i + 1$ since $1 - \varepsilon \geq 1/2$. Since $-\lceil -a_i \rceil_\varepsilon = \lfloor a_i + \varepsilon \rfloor$, this shows that $-\lceil -a_i \rceil_\varepsilon \leq \lceil A_i \rceil$. Altogether this implies that

$$\begin{aligned} |A_i| &\leq \lceil |a_i| \rceil_\varepsilon \quad \text{when } A_i \in \mathbb{Z}, \quad \text{and} \\ |A_i + 1/2| &\leq \lceil |a_i| \rceil_\varepsilon \quad \text{when } A_i \in \mathbb{Z} + 1/2. \end{aligned} \tag{4.6.8}$$

We also have that

$$B + C = \lceil b + c \rceil \quad \text{and} \quad B - C = \lceil b - c \rceil. \tag{4.6.9}$$

Note that

$$\begin{aligned} &\lceil 2b \rceil = \lceil b + c + b - c \rceil \leq 2B \leq \lceil b + c + (b - c) \rceil + 1 = \lceil 2b \rceil + 1 \leq 2\lceil b \rceil + 1; \\ &\lceil 2c \rceil - 1 = \lceil b + c - (b - c) \rceil - 1 \leq 2C \leq \lceil b + c - (b - c) \rceil = \lceil 2c \rceil \leq 2\lceil c \rceil. \end{aligned} \tag{4.6.10}$$

It is straightforward to check that these equations (together with the definition of C) imply the following:

$$\begin{aligned} 0 \leq B \leq \lceil b \rceil &\quad \text{when } B \in \mathbb{Z} \\ 0 \leq B - 1/2 \leq \lceil b \rceil &\quad \text{when } B \in \mathbb{Z} + 1/2 \\ 0 \leq C \leq \lceil c \rceil &\quad \text{when } C \in \mathbb{Z} \\ 0 \leq C - 1/2 < C + 1/2 \leq \lceil c \rceil &\quad \text{when } C \in \mathbb{Z} + 1/2. \end{aligned} \tag{4.6.11}$$

Finally, note that (4.6.7) and (4.6.9) together imply that

$$2A_1 + B - C = 2x_1 \quad \text{and} \quad 2A_2 + B + C = 2x_2. \tag{4.6.12}$$

We choose a_1', a_2', b', c' as follows:

	a_1'	a_2'	b'	c'	
(i)	A_1	A_2	B	C	if $\lceil b+c \rceil$, $\lceil b-c \rceil$ both even;
(ii)	$A_1 + 1/2$	A_2	$B - 1/2$	$C + 1/2$	if $\lceil b+c \rceil$ even, $\lceil b-c \rceil$ odd;
(iii)	A_1	$A_2 + 1/2$	$B - 1/2$	$C - 1/2$	if $\lceil b+c \rceil$ odd, $\lceil b-c \rceil$ even;
(iv)	$A_1 + 1/2$	$A_2 + 1/2$	$B - 1$	C	if $b > 0$ and $\lceil b+c \rceil$, $\lceil b-c \rceil$ both odd;
(v)	$A_1 - 1/2$	$A_2 + 1/2$	B	$C - 1$	if $b = 0$ and $\lceil b+c \rceil$, $\lceil b-c \rceil$ both odd.

By the definition of A_i we have for each $i = 1, 2$ that $a_i' \geq 0$ if and only if $a_i \geq 0$. Then $\{a_1', a_2', b', c'\} \subseteq \mathbb{Z}$ and (4.6.12) implies that

$$2a_1' + b' - c' = 2x_1 \quad \text{and} \quad 2a_2' + b' + c' = 2x_2.$$

Moreover, $b' + c' \leq B + C = \lceil b + c \rceil$. We claim that $0 \leq b' \leq \lceil b \rceil$ and $0 \leq c' \leq \lceil c \rceil$ and $|a_i'| \leq \lceil |a_i| \rceil_\varepsilon$ for $i = 1, 2$ respectively in all cases (i)–(v). To see this, suppose first that we are in case (iv). Since $b > 0$, (4.6.10) implies that $B \geq \lceil 2b \rceil / 2 > 0$, so, since $B \in \mathbb{Z}$, $B - 1 \geq 0$ in this case.

Suppose now that we are in case (v). Then $\lceil c \rceil, \lceil -c \rceil = -\lfloor c \rfloor$ are both odd. Therefore $\lceil c \rceil, \lfloor c \rfloor$ are both odd so $\lceil c \rceil = \lfloor c \rfloor = c$. So $c \in \mathbb{N}_0$ is odd, $B = 0$ and $C = c$. Thus $C - 1 \geq 0$. Moreover $c = 2A_1 - 2x_1$, so $2A_1$ is odd and positive, which implies that $A_1 - 1/2 \geq 0$. Then (4.6.8) implies that $|A_1 - 1/2| \leq \lceil |a_i| \rceil_\varepsilon$.

In all cases (i)–(v), these last deductions together with (4.6.8)–(4.6.11) complete the proof of the lemma. □

4.6.6 Proof of Lemma 4.6.1

Before we can prove Lemma 4.6.1, we need one more preliminary result which guarantees a path system $\mathcal{P}$ that can balance out the vertex class sizes of the bipartite graphs

induced by the W_i. If $e_{\mathcal{P}}(W_1, W_2) = 0$, then we will use 3-connectivity (via Lemma 4.6.5) to modify $\mathcal{P}$ into a balanced path system which also links up the W_i.

Lemma 4.6.13. *Let $0 < 1/n \ll \rho \ll \nu \ll \tau \ll \alpha < 1$ and let G be a D-regular graph on n vertices with $D \geq \alpha n$. Suppose that G has a robust partition $\mathcal{V} := \{W_1, W_2\}$ with parameters $\rho, \nu, \tau, 0, 2$. For each $i = 1, 2$, let A_i, B_i be the bipartition of W_i guaranteed by* (D3)*, and suppose that $|A_i| \geq |B_i|$. Then*

(i) G contains a path system $\mathcal{P}$ which is 2-balanced with respect to (A_1, B_1, A_2, B_2) such that $e(\mathcal{P}) \leq \sqrt{\rho} n$;

(ii) if $e_{\mathcal{P}}(W_1, W_2) > 0$ then $\mathcal{P}$ contains a W_1W_2-path;

(iii) for $i = 1, 2$, $\mathcal{P}[W_i]$ consists either of a matching in $G[A_i]$ of size at most $\lceil e_G(A_i)/5 \rceil_{1/4}$, or a matching in $G[B_i]$ of size at most $\lceil e_G(B_i)/5 \rceil_{1/4}$.

Proof. Write $\mathcal{V}^* := (A_1, B_1, A_2, B_2)$. Let $\Delta := D/2$ and note that

$$\Delta(G[A_i]), \Delta(G[B_i]), \Delta(G[W_1, W_2]) \leq \Delta$$

for $i = 1, 2$ by (D4) and (D5). Without loss of generality, we may suppose that $e_G(A_1, B_2) \leq e_G(B_1, A_2)$. Note that G is D-balanced with respect to $\mathcal{V}^*$ by Proposition 4.6.2. Apply Lemma 4.6.4 to G. Suppose that Lemma 4.6.4(i) holds. Then $G[A_i]$ contains a matching M_i of size $|A_i| - |B_i| \leq \lceil e_G(A_i)/5 \rceil_{1/4}$ for $i = 1, 2$. Set $\mathcal{P} := M_1 \cup M_2$. So (iii) holds, (D3) and (C2) imply that (i) holds, and (ii) is vacuous.

So we may assume that Lemma 4.6.4(ii) holds. Let H be a spanning subgraph of G which is D-balanced with respect to $\mathcal{V}^*$ such that $E(H) \subseteq E(G[C_1]) \cup E(G[C_2]) \cup E(G[W_1, A_2])$ for some $C_1 \in \{A_1, B_1\}$ and $C_2 \in \{A_2, B_2\}$. Observe that

$$e(H) \leq \sum_{i=1,2} \left(e_G(A_i, \overline{B_i}) + e_G(B_i, \overline{A_i}) \right) \overset{\text{(D3),(C3)}}{\leq} 2\rho n^2. \tag{4.6.13}$$

For each $H' \subseteq H$ and $i = 1, 2$, define

$$f_i(H') = e_{H'}(A_i) - e_{H'}(B_i). \tag{4.6.14}$$

Now (4.6.1) implies that, for any $t \in \mathbb{N}_0$, H' is t-balanced if

$$2f_i(H') + e_{H'}(A_i, W_j) - e_{H'}(B_i, W_j) = t(|A_i| - |B_i|) \tag{4.6.15}$$

for $\{i, j\} = \{1, 2\}$. Observe that $e_H(C_i) = e_H(W_i) = |f_i(H)|$. For $i = 1, 2$, let

$$a_i := f_i(H)/\Delta. \tag{4.6.16}$$

Then the D-balancedness of H and (4.6.15) imply that

$$\begin{aligned} 2a_1 + \frac{e_H(A_1, A_2)}{\Delta} - \frac{e_H(B_1, A_2)}{\Delta} &= 2(|A_1| - |B_1|) \\ \text{and} \quad 2a_2 + \frac{e_H(A_1, A_2)}{\Delta} + \frac{e_H(B_1, A_2)}{\Delta} &= 2(|A_2| - |B_2|). \end{aligned}$$

Apply Lemma 4.6.12 with $a_1, a_2, e_H(A_1, A_2)/\Delta, e_H(B_1, A_2)/\Delta, |A_1| - |B_1|, |A_2| - |B_2|, 1/4$ playing the roles of $a_1, a_2, b, c, x_1, x_2, \varepsilon$ to obtain integers a_1', a_2', b', c' with

$$|a_i'| \leq \lceil |a_i| \rceil_{1/4} = \lceil e_H(C_i)/\Delta \rceil_{1/4} \quad \text{for } i = 1, 2; \tag{4.6.17}$$

$$a_i' \geq 0 \quad \text{if and only if} \quad a_i \geq 0; \tag{4.6.18}$$

$0 \leq b' \leq \lceil e_H(A_1, A_2)/\Delta \rceil$; $0 \leq c' \leq \lceil e_H(B_1, A_2)/\Delta \rceil$ and

$$b' + c' \leq \lceil e_H(W_1, A_2)/\Delta \rceil; \tag{4.6.19}$$

$$2a_1' + b' - c' = 2(|A_1| - |B_1|) \quad \text{and} \quad 2a_2' + b' + c' = 2(|A_2| - |B_2|). \tag{4.6.20}$$

Apply Lemma 4.6.7 with $H[W_2, W_1], W_2, A_1, B_1$ playing the roles of G, U, V, W to obtain a matching M in $H[W_2, W_1]$ such that

$$e_M(A_1, A_2) = e_M(A_1, W_2) = b', \quad e_M(B_1, A_2) = e_M(B_1, W_2) = c' \tag{4.6.21}$$

$$\text{and} \quad e_M(W_1, B_2) = 0.$$

Then (4.6.13) and (4.6.19) imply that $e(M) = b' + c' \leq \lceil e(H)/\Delta \rceil \leq \sqrt{\rho}\Delta$. By (4.6.13) and (4.6.17), we can apply Lemma 4.6.11 to H with $\sqrt{\rho}, M, \Delta, W_1, W_2, |a_1'|, |a_2'|$ playing the roles of $\rho, M, \Delta, U, V, a_U, a_V$ to obtain a path system $\mathcal{P}$ such that

$$\mathcal{P}[W_1, W_2] = M; \tag{4.6.22}$$

$$e_{\mathcal{P}}(W_i) = e_{\mathcal{P}}(C_i) = |a_i'| \quad \text{for } i = 1, 2; \tag{4.6.23}$$

$\mathcal{P}[C_i]$ is a matching for $i = 1, 2$, and if $M \neq \emptyset$, then $\mathcal{P}$ contains a W_1W_2-path. So (ii) holds. (Note that (4.6.23) follows from the fact that $H[W_i] = H[C_i]$.) Moreover, (4.6.17) and (4.6.23) imply that the matching $\mathcal{P}[C_i]$ has size at most $\lceil e_H(C_i)/\Delta \rceil_{1/4} \leq \lceil e_G(C_i)/\Delta \rceil_{1/4} \leq \lceil e_G(C_i)/5 \rceil_{1/4}$. So (iii) holds. Equations (4.6.14), (4.6.16), (4.6.18) and (4.6.23) imply that

$$f_i(\mathcal{P}) = a_i'. \tag{4.6.24}$$

Furthermore, by (4.6.21) and (4.6.22) we have

$$e_{\mathcal{P}}(A_1, W_2) - e_{\mathcal{P}}(B_1, W_2) = b' - c' \quad \text{and} \quad e_{\mathcal{P}}(W_1, A_2) - e_{\mathcal{P}}(W_1, B_2) = b' + c'.$$

Together with (4.6.15), (4.6.20) and (4.6.24), this implies that $\mathcal{P}$ is 2-balanced with respect to $\mathcal{V}^*$. Finally,

$$e(\mathcal{P}) = |a_1'| + |a_2'| + b' + c' \overset{(4.6.17),(4.6.19)}{\leq} e(H)/\Delta + 3 \overset{(4.6.13)}{\leq} \sqrt{\rho}n,$$

as required. □

Proof of Lemma 4.6.1. Let $\mathcal{V} := \{W_1, W_2\}$ and for $i = 1, 2$, let A_i, B_i be the partition of W_i guaranteed by (D3). Without loss of generality, we may suppose that $|A_i| \geq |B_i|$. Apply Lemma 4.6.13 to obtain a path system $\mathcal{P}$ which is 2-balanced with respect to (A_1, B_1, A_2, B_2) such that $e(\mathcal{P}) \leq \sqrt{\rho}n$.

Suppose first that $e_{\mathcal{P}}(W_1, W_2) > 0$. Then $\mathcal{P}$ contains a W_1W_2-path by Lemma 4.6.13(ii). So we are done by Proposition 4.6.3. Therefore we may assume that $e_{\mathcal{P}}(W_1, W_2) = 0$. Lemma 4.6.13(iii) implies that, for each $i = 1, 2$, at least one of $\mathcal{P}[A_i], \mathcal{P}[B_i]$ is empty, and the other is a matching of size at most $\lceil e_G(B_i)/5 \rceil_{1/4}, \lceil e_G(A_i)/5 \rceil_{1/4}$ respectively. The 2-balancedness of $\mathcal{P}$ implies that $e_{\mathcal{P}}(A_i) - e_{\mathcal{P}}(B_i) = |A_i| - |B_i| \geq 0$. So $\mathcal{P} = M_1 \cup M_2$ for some matchings $M_i \subseteq G[A_i]$. Apply Lemma 4.6.5 to obtain a path system $\mathcal{P}'$ which is 2-balanced with respect to (A_1, B_1, A_2, B_2) and contains a W_1W_2-path, and $e(\mathcal{P}) \leq 3\sqrt{\rho}n$. Again, we are done by Proposition 4.6.3. □

4.7 (2,1) : Two robust expander components and one bipartite robust expander component

The aim of this section is to prove the following lemma.

Lemma 4.7.1. *Let $0 < 1/n \ll \rho \ll \nu \ll \tau \ll 1$. Let G be a 3-connected D-regular graph on n vertices where $D \geq n/4$. Let $\mathcal{X}$ be a robust partition of G with parameters $\rho, \nu, \tau, 2, 1$. Then G contains a Hamilton cycle.*

This — the final case — is the longest and most difficult. This is perhaps unsurprising given that the extremal example in Figure 4.1(i) has precisely this structure. Moreover, the presence of a bipartite robust expander component means that the path system we find to join the robust components needs to be balanced with respect to the bipartite component – the regularity of G is essential to achieve this. On the other hand, since we have to join up three components, the 3-connectivity of G is essential too. The main challenge is to find a path system which satisfies both requirements simultaneously, i.e.

one that is both balanced and joins up the three components. We need to invoke the degree bound $D \geq n/4$ for this. We begin by giving a brief outline of the argument.

4.7.1 Sketch of the proof of Lemma 4.7.1

Let $\mathcal{X}' := \{V_1', V_2', W'\}$, where $G[V_i']$ is a robust expander component for $i = 1, 2$, and $G[W']$ is a bipartite robust expander component with bipartition A', B', where $|A'| \geq |B'|$. One can hope to use the regularity of G to find a path system $\mathcal{P}'$ consisting of a matching in A', together with a matching from A' to $U' := V_1' \cup V_2'$, which balances (the sizes of the vertex classes A', B' of) $G[W']$. However, $\mathcal{P}'$ may not connect W' to each of V_1' and V_2' in the right way. We could for example have that $e_{\mathcal{P}'}(W', V_1') = 0$ or that $e_{\mathcal{P}'}(W', V_1')$ is odd. In both cases, $\mathcal{P}'$ requires modification. But if one adds an edge to $\mathcal{P}'$ between one of the V_i' and W', then $\mathcal{P}'$ will no longer balance $G[W']$, meaning that $\mathcal{P}'$ must be further adapted.

It turns out that it is better to begin with a small path system $\mathcal{P}_0$ for which $R_{\mathcal{X}'}(\mathcal{P}_0)$ has an Euler tour, but which does not necessarily balance $G[W']$. If $\mathcal{P}_0$ also balances $G[W']$ then we are done. So suppose not. We then attempt to balance $\mathcal{P}_0$ by adding edges of $G[W']$ to $\mathcal{P}_0$. When such an attempt fails, we will slightly modify $\mathcal{P}_0$ using the additional structural information about G that this failure implies. We then add edges of $G[W']$ to the modified path system.

To find $\mathcal{P}_0$ which corresponds to an Euler tour, one could simply use Lemma 3.7.3. However, since the proof of the lemma uses the 3-connectivity of G, we have insufficient control on the structure of $\mathcal{P}_0$ (i.e. it may not be possible to extend it into a balancing path system). Instead, we will construct $\mathcal{P}_0$ by first finding a large matching M in $G[A', \overline{W'}]$. Typically this matching will be obtained using König's theorem on edge-colourings, so $e(M) \geq e_G(A', \overline{W'})/\Delta(G[A', \overline{W'}])$. Since $\mathcal{X}'$ is a robust partition, (D4) implies that $\Delta(G[A', \overline{W'}]) \leq 2D/3$. This would give $e(M) \geq 3e_G(A', \overline{W'})/2D$, which is insufficient for our purposes. To improve on this, we alter the partition $\mathcal{X}'$ very slightly to obtain a weak robust partition $\mathcal{V} = \{V_1, V_2, W\}$ so that $\Delta(G[W, \overline{W}]) \leq D/2$ (where $G[V_1]$ and $G[V_2]$

are robust expander components and $G[W]$ is a bipartite robust expander component with bipartition A, B, where $|A| \geq |B|$). By Lemma 3.6.2 it is *still* sufficient to find a $\mathcal{V}$-tour using the approach outlined above (see Lemma 4.7.3 and Subsection 4.7.3 for the statement and proof of this reduction). Now the matching in $G[A, \overline{W}]$ which will be used to construct the initial path system $\mathcal{P}_0$ has size at least $2e_G(A, \overline{W})/D$.

We prove Lemma 4.7.1 separately in each of the following four cases:

- $|A| - |B| \geq 2$ and $e_G(A, \overline{W})$ is at least a little larger than $3D/2$ (Subsection 4.7.5);
- $|A| - |B| \geq 2$ and $e_G(A, \overline{W})$ is at most a little larger than $3D/2$ (Subsection 4.7.6);
- $|A| - |B| = 1$ (Subsection 4.7.7);
- $|A| = |B|$ (Subsection 4.7.8).

The reason for these distinctions will be discussed at the end of Subsection 4.7.4. The full strength of the minimum degree bound $D \geq n/4$ is only used in the last two cases.

4.7.2 Notation

Throughout the remainder of the chapter, whenever we say that a graph G has vertex partition $\mathcal{V} = \{V_1, V_2, W := A \cup B\}$, we assume that $V(G)$ has a partition into parts V_1, V_2, W, each of size at least $|V(G)|/100 \geq 100$, that A and B are disjoint and $|A| \geq |B|$. Moreover, we will say that G has a weak robust partition $\mathcal{V} = \{V_1, V_2, W := A \cup B\}$ (for some given parameters) if $\mathcal{V}$ satisfies the above properties and is a weak robust partition of G such that $G[V_1], G[V_2]$ are two robust expander components and $G[W]$ is a bipartite robust expander component, and the bipartition of W as specified by (D3′) is A, B. We will use a similar notation when $\mathcal{V}$ is a robust partition of G.

Given $0 < \varepsilon < 1$ and $\Delta > 0$, consider any graph G with vertex partition U, A, B such that $\Delta(G[A]), \Delta(G[A, U]) \leq \Delta$. We say that

$$\mathrm{char}_{\Delta,\varepsilon}(G) := (\ell, m) \tag{4.7.1}$$

when $\ell := \lceil e_G(A)/\Delta \rceil_\varepsilon$ and m is the largest even integer less than or equal to $\lceil e_G(A,U)/\Delta \rceil_\varepsilon$. (Recall the definition of $\lceil \cdot \rceil_\varepsilon$ from the end of Subsection 4.4.) We think of 'char' as being short for 'character'. The character of G encodes what sort of $\mathcal{V}$-tour $\mathcal{P}$ we can hope to find. Typically, when G has character (ℓ, m), a $\mathcal{V}$-tour will closely resemble the union of a matching of size ℓ in $G[A]$, and a matching of size m is $G[A,U]$. (Recall that, in a $\mathcal{V}$-tour $\mathcal{P}$, we have that $e_{\mathcal{P}}(W,U)$ is even.)

Given any path system $\mathcal{P}$ in G, we write

$$\mathrm{bal}_{AB}(\mathcal{P}) := e_{\mathcal{P}}(A) - e_{\mathcal{P}}(B) + (e_{\mathcal{P}}(A,U) - e_{\mathcal{P}}(B,U))/2. \tag{4.7.2}$$

When $\mathcal{V} = \{V_1, V_2, W := A \cup B\}$ is a vertex partition of G, we take $U := V_1 \cup V_2$ in the definitions of $\mathrm{char}_{\Delta,\varepsilon}$ and bal_{AB}.

Given $0 < \varepsilon < 1$, $\Delta > 0$ and a graph G with partition $\mathcal{V} = \{V_1, V_2, W := A \cup B\}$ and $\mathrm{char}_{\Delta,\varepsilon}(G) = (\ell, m)$, we will find a path system satisfying the following properties:

(P1) $e(\mathcal{P}) \leq \ell + m + 6$;

(P2) $\mathrm{bal}_{AB}(\mathcal{P}) = |A| - |B|$;

(P3) $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour.

4.7.3 Preliminaries and a reduction

In this subsection we show that, in order to prove Lemma 4.7.1, it is sufficient to prove Lemma 4.7.3 below. We then state some tools which will be used in the next subsections to do so. The following observation provides us with a convenient check for a path system $\mathcal{P}$ to be such that $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour.

Fact 4.7.2. *Let G be a graph with vertex partition $\mathcal{V}$ into three parts. Then, for a path system $\mathcal{P}$ in G,* (P3) *is equivalent to the following. For each $X \in \mathcal{V}$, $e_{\mathcal{P}}(X, \overline{X})$ is even and there exists $X' \in \mathcal{V} \setminus \{X\}$ such that $\mathcal{P}$ contains an XX'-path.*

The remainder of Section 4.7 is devoted to the proof of the following lemma, which states that G contains a path system satisfying (P1)–(P3) (when the partition $\mathcal{V}$ and the parameters involved are suitably defined).

Lemma 4.7.3. *Let $n, D \in \mathbb{N}$ and $\ell, m \in \mathbb{N}_0$. Let $0 < 1/n \ll \rho \ll \nu \ll \tau \ll \varepsilon \ll 1$. Let G be a 3-connected D-regular graph on n vertices where $D \geq n/4$. Suppose that G has a weak robust partition $\mathcal{V} = \{V_1, V_2, W := A \cup B\}$ with parameters $\rho, \nu, \tau, 1/16, 2, 1$ such that $|V_1|, |V_2| \geq D/2$. Suppose further that $\Delta(G[A, V_1 \cup V_2]) \leq D/2$, $d_{V_i}(x_i) \geq d_{V_j}(x_i)$ for all $x_i \in V_i$ and all $\{i, j\} = \{1, 2\}$, and $d_A(a) \leq d_B(a)$ for all $a \in A$. Let $\mathrm{char}_{D/2,\varepsilon}(G) = (\ell, m)$. Then G contains a path system $\mathcal{P}$ satisfying* (P1)–(P3).

The following proposition gives bounds on ℓ and m when $\mathrm{char}_{\Delta,\varepsilon}(G) = (\ell, m)$.

Proposition 4.7.4. *Let $n, D \in \mathbb{N}$ and $\ell, m \in \mathbb{N}_0$. Let $0 < 1/n \ll \rho \ll \nu \ll \tau \ll \varepsilon, \eta \ll 1$ and suppose $D \geq n/4$. Let G be a graph on n vertices with weak robust partition $\mathcal{V} = \{V_1, V_2, W := A \cup B\}$ with parameters $\rho, \nu, \tau, \eta, 2, 1$. Suppose further that $\Delta(G[A]), \Delta(G[A, V_1 \cup V_2]) \leq D/2$ and that $\mathrm{char}_{D/2,\varepsilon}(G) = (\ell, m)$. Then $\ell, m \leq 12\rho n$.*

Proof. (D3′) implies that $G[W]$ is ρ-close to bipartite with bipartition A, B. So $e_G(A) + e_G(A, V_1 \cup V_2) \leq \rho n^2$. Thus $\ell = \lceil 2e_G(A)/D \rceil_\varepsilon \leq 3\rho n^2/D \leq 12\rho n$. An almost identical calculation gives the same bound for m. □

We now show that, to prove Lemma 4.7.1, it suffices to prove Lemma 4.7.3.

Proof of Lemma 4.7.1 *(assuming Lemma* 4.7.3*).* Choose ε with $\tau \ll \varepsilon \ll 1$. Let $\mathcal{X} = \{U_1, U_2, W' := A' \cup B'\}$ be a robust partition of G with parameters $\rho, \nu, \tau, 2, 1$, where $G[U_1], G[U_2]$ are (ρ, ν, τ)-robust expander components and $G[W']$ is a bipartite (ρ, ν, τ)-robust expander component with bipartition A', B' as guaranteed by (D3). We will alter $\mathcal{X}$ slightly so that it is a weak robust partition and that additionally the degree conditions of Lemma 4.7.3 hold.

> **Claim.** *There exists a weak robust partition $\mathcal{V} = \{V_1, V_2, W := A \cup B\}$ of G with parameters $\rho^{1/3}, \nu/2, 2\tau, 1/16, 2, 1$ such that $|V_1|, |V_2| \geq D/2$, $\Delta(G[A, V_1 \cup V_2]) \leq$*

$D/2$, $d_{V_i}(x_i) \geq d_{V_j}(x_i)$ *for all* $x_i \in V_i$ *and* $\{i,j\} = \{1,2\}$, *and* $d_A(a) \leq d_B(a)$ *for all* $a \in A$.

Proof. For $i = 1, 2$, let X_i be the collection of vertices $x \in U_i$ with $d_{\overline{U_i}}(x) > \rho n$. Then (D7) implies that $|X_i| \leq \rho n$. Let $Y_i := U_i \setminus X_i$. Then each $y \in Y_i$ satisfies

$$d_{Y_i}(y) = d(y) - d_{\overline{U_i} \cup X_i}(y) \geq d(y) - \rho n - |X_i| \geq d(y) - 2\rho n. \tag{4.7.3}$$

Let A_0 be the collection of vertices $a \in A'$ such that $d_{\overline{B'}}(a) \geq \sqrt{\rho} n$. Let $A_1 := A' \setminus A_0$. Define B_0, B_1 analogously. By (D3), $G[W']$ is ρ-close to bipartite with bipartition A', B'. Therefore (C3) holds, from which one can easily derive that $|A_0|, |B_0| \leq 2\sqrt{\rho} n$. Similarly as in (4.7.3), for each $a \in A_1$ and $b \in B_1$ we have

$$d_{B_1}(a) \geq d(a) - 3\sqrt{\rho} n \quad \text{and} \quad d_{A_1}(b) \geq d(b) - 3\sqrt{\rho} n. \tag{4.7.4}$$

Let $V_0 := X_1 \cup X_2 \cup A_0 \cup B_0$. Then

$$|V_0| \leq 5\sqrt{\rho} n. \tag{4.7.5}$$

Among all partitions X_1', X_2', A_0', B_0' of V_0, choose one such that $e(A \cup B, V_1 \cup V_2)$ is minimised; and subject to $e(A \cup B, V_1 \cup V_2)$ being minimal we have that $e(V_1, V_2) + e(A) + e(B)$ is minimal, where $V_i := Y_i \cup X_i'$, $A := A_1 \cup A_0'$ and $B := B_1 \cup B_0'$. It is easy to see that $d_{A \cup B}(w) \geq d_{V_1 \cup V_2}(w)$ for all $w \in A_0' \cup B_0'$; $d_{V_1 \cup V_2}(v) \geq d_{A \cup B}(v)$ for all $v \in X_1' \cup X_2'$; $d_{V_i}(v_i) \geq d_{V_j}(v_i)$ for all $v_i \in X_i'$ and $\{i,j\} = \{1,2\}$; $d_A(a) \leq d_B(a)$ for all $a \in A_0'$; and $d_B(b) \leq d_A(b)$ for all $b \in B_0'$. If $v_i \in Y_i$, then (4.7.3) implies that $d_{V_i}(v_i) \geq d_{Y_i}(v_i) \geq d(v_i) - 2\rho n \geq d(v_i)/2$. So $d_{V_i}(v_i) \geq d_{A \cup B}(v_i), d_{V_j}(v_i)$ for $\{i,j\} = \{1,2\}$. Similarly, (4.7.4) implies that, for all $w \in A_1 \cup B_1$ we have $d_{A \cup B}(w) \geq d_{V_1 \cup V_2}(w)$; for all $a \in A_1$ we have $d_A(a) \leq d_B(a)$ and for all $b \in B_1$ we have $d_B(b) \leq d_A(b)$. Observe that (4.7.3), (4.7.4) imply that $|V_i| \geq D - 2\rho n$ and $|A|, |B| \geq D - 3\sqrt{\rho} n$ respectively. It remains to prove

that $\mathcal{V} := \{V_1, V_2, W := A \cup B\}$ is a weak robust partition with parameters $\rho^{1/3}, \nu/2, 2\tau, 1/16, 2, 1$. Property (D1′) is clear. By relabelling if necessary, we may assume that $|A| \geq |B|$. We now prove (D2′). Observe that

$$e(V_i, \overline{V_i}) \leq e(U_i, \overline{U_i}) + D|X_i| + D|X_i'| \leq (\rho + 6\sqrt{\rho})n^2 \leq \rho^{1/3}n^2.$$

Therefore each V_i is a $\rho^{1/3}$-robust component of G. Note also that

$$|V_i \triangle U_i| \leq |V_0| \overset{(4.7.5)}{\leq} 5\sqrt{\rho}n \leq \nu|U_i|/2.$$

Lemma 3.4.8 implies that $G[V_i]$ is a $(\nu/2, 2\tau)$-robust expander. Therefore $G[V_i]$ is a $(\rho^{1/3}, \nu/2, 2\tau)$-robust expander component for $i = 1, 2$, so (D2′) holds. To prove (D3′), note that $|A \triangle A'| + |B \triangle B'| \leq 2|V_0| \leq \rho^{1/3}n/3$ where the final inequality follows from (4.7.5). Now Lemma 3.4.10 implies that $G[A \cup B]$ is a bipartite $(\rho^{1/3}, \nu/2, 2\tau)$-robust expander component of G with bipartition A, B. Thus (D3′) holds. Finally, (D4′) and (D5′) are clear from the degree conditions we have already obtained. ■

Given the partition $\mathcal{V}$ of $V(G)$, let ℓ, m satisfy $\mathrm{char}_{D/2,\varepsilon}(G) = (\ell, m)$. Let $\mathcal{P}$ be a path system in G guaranteed by Lemma 4.7.3, i.e. $\mathcal{P}$ satisfies (P1)–(P3). Note that $\mathcal{V}$ is also a weak robust partition with parameters $\rho^{1/3}, \nu/2, 2\tau, \varepsilon, 2, 1$. So (P1) and Proposition 4.7.4 with $\rho^{1/3}, \varepsilon$ playing the roles of ρ, η imply that $e(\mathcal{P}) \leq 25\rho^{1/3}n$. Then, for each $X \in \mathcal{V}$ we have that $|V(\mathcal{P}) \cap X| \leq |V(\mathcal{P})| \leq 2e(\mathcal{P}) \leq 50\rho^{1/3}n \leq \rho^{1/4}n/9$. So Lemma 3.7.8 applied with $2, 1, W, A, B, \mathcal{P}, \rho^{1/4}/9$ playing the roles of $k, \ell, W_j, A_j, B_j, \mathcal{P}, \rho$ implies that G contains a path system $\mathcal{P}'$ that is a $\mathcal{V}$-tour with parameter $\rho^{1/4}$. Now Lemma 3.6.2 with $\mathcal{P}', \rho^{1/3}, \rho^{1/4}, \nu/2, 2\tau, 1/16, 2, 1$ playing the roles of $\mathcal{P}, \rho, \gamma, \nu, \tau, \eta, k, \ell$ implies that G contains a cycle whose vertex set includes every vertex in $\bigcup_{V \in \mathcal{V}} V$, i.e. a Hamilton cycle. □

4.7.4 Tools

In this section we gather some useful tools which will be used repeatedly in the sections to come. We will often use the following lower bounds for $e_G(A), e_G(A,U)$ implied by $\mathrm{char}_{\Delta,\varepsilon}(G)$.

Proposition 4.7.5. *Let $\Delta, \Delta' \in \mathbb{N}$ and $\ell, m \in \mathbb{N}_0$. Let $\Delta'/\Delta \leq \varepsilon < 1$. Suppose that G is a graph with vertex partition U, A, B such that $\Delta(G[A]), \Delta(G[A,U]) \leq \Delta$ and $\mathrm{char}_{\Delta,\varepsilon}(G) = (\ell, m)$. Then $e_G(A) \geq (\ell-1)\Delta + \Delta'$ and $e_G(A,U) \geq (m-1)\Delta + \Delta'$.*

Proof. We have that $\ell = \lceil e_G(A)/\Delta \rceil_\varepsilon = \lceil e_G(A)/\Delta - \varepsilon \rceil$ so $\ell - 1 < e_G(A)/\Delta - \varepsilon \leq (e_G(A) - \Delta')/\Delta$, as required. A near identical calculation proves the second assertion. □

The path system we require will contain edges in $G[A]$ and $G[V_1 \cup V_2, A]$, and will 'roughly look like' a matching within each of these subgraphs. The following lemma allows us to find a structure which in turn contains a large matching even if certain vertices need to be avoided.

Lemma 4.7.6. *Let $\Delta, \Delta' \in \mathbb{N}$ and $\ell \in \mathbb{N}_0$ be such that $\ell/\Delta', \Delta'/\Delta, 1/\Delta' \ll 1$. Let G be a graph with $\Delta(G) \leq \Delta$, and let $e(G) \geq (\ell-1)\Delta + \Delta'$. Then G contains one of the following:*

(i) a matching M of size $\ell + 1$ and $uv \in E(G)$ with $u \notin V(M)$;

(ii) ℓ vertices each with degree at least Δ'.

Moreover, if $\ell \geq 1$ and $e(G) \geq \ell\Delta + 1$; or $\ell = 0$ and $e(G) \geq 2$, then (i) *holds.*

Proof. We will use induction on ℓ in order to show that either (i) or (ii) holds. The cases $\ell = 0, 1$ are trivial. Suppose now that $\ell \geq 2$. Suppose first that $\Delta(G) \leq \Delta'$. Then, by Vizing's theorem, $E(G)$ can be properly coloured with at most $\Delta' + 1$ colours. Therefore G contains a matching of size

$$\left\lceil \frac{e(G)}{\Delta' + 1} \right\rceil \geq \left\lceil \frac{(\ell-1)\Delta + \Delta'}{\Delta' + 1} \right\rceil \geq \ell + 2.$$

So (i) holds. Thus we may assume that there exists $x \in V(G)$ with $d(x) \geq \Delta'$. Let $G^- := G \setminus \{x\}$. Then $e(G^-) \geq e(G) - \Delta \geq (\ell - 2)\Delta + \Delta'$. By induction, $e(G^-)$ contains either a matching M^- of size ℓ and $uv \in E(G^-)$ with $u \notin V(M^-)$, or $\ell - 1$ vertices of degree at least Δ'. In the first case, choose $y \in N(x) \setminus V(M^-)$ with $y \neq u$ and let $M := M^- \cup \{xy\}$. Then (i) holds. In the second case, x is our ℓth vertex of degree at least Δ' in G, so (ii) holds.

For the moreover part, suppose now that $\ell \geq 1$ and $e(G) \geq \ell\Delta + 1$. Suppose that (i) does not hold. Let $x_1, \ldots, x_\ell$ be ℓ distinct vertices of degree at least Δ'. Then $e(G \setminus \{x_1, \ldots, x_\ell\}) \geq e(G) - \Delta\ell \geq 1$. So G contains an edge e which is not incident to $\{x_1, \ldots, x_\ell\}$. We obtain a contradiction by considering $\{e, x_1z_1\} \cup \{x_1y_1, \ldots, x_\ell y_\ell\}$, where $z_1 \in N(x_1)$ avoids e and for $1 \leq i \leq \ell$ the vertices $y_i \in N(x_i)$ are distinct, and avoid e, z_1 and $x_1, \ldots, x_\ell$.

Finally, if $\ell = 0$, then any two edges of G satisfy (i). □

Given an even matching M in $G[A, V_1 \cup V_2]$ and a lower bound on $e_G(A)$, we would like to extend M into a path system $\mathcal{P}$ using edges from $G[A]$ so that $\mathrm{bal}_{AB}(\mathcal{P})$ is large. Lemma 4.7.6 gives us two useful structures in $G[A]$ from which we can choose suitable edges to add to M to form $\mathcal{P}$. The following proposition does this in the case when Lemma 4.7.6(i) holds.

Proposition 4.7.7. *Let G be a graph with vertex partition X, Y. Suppose that $G[Y]$ contains a matching M' of size $\ell + 1$ and an edge uv with $u \notin V(M')$. Let M be a non-empty even matching of size m in $G[X, Y]$. Then G contains a path system $\mathcal{P}$ such that*

(i) $\mathcal{P}[X, Y] = M$ and $\mathcal{P} \subseteq M \cup M' \cup \{uv\}$;

(ii) $e_{\mathcal{P}}(Y) = \ell + 1$;

(iii) $\mathcal{P}$ contains at least two XY-paths.

Proof. We will extend M by adding edges from $M' \cup \{uv\}$, so (i) automatically holds. Note that any path system $\mathcal{P}$ obtained in this way contains an even number of XY-paths. So it suffices to find such a $\mathcal{P}$ with at least one XY-path. If $M \cup M'$ contains an XY-path, then we are done by setting $\mathcal{P} := M \cup M'$. So suppose not. Then $M'[V(M) \cap Y]$ is a perfect matching M''. If $v \in V(M'')$, let f be the edge of M'' containing v. Otherwise, let $f \in E(M'')$ be arbitrary. We take $\mathcal{P} := M \cup M' \cup \{uv\} \setminus \{f\}$. Now both of the two edges in M which are incident to f lie in distinct XY-paths of $\mathcal{P}$, so (iii) holds. Clearly (ii) holds too. $\square$

Following on from the previous proposition, we now consider how to extend M into $\mathcal{P}$ when instead Lemma 4.7.6(ii) holds in $G[A]$.

Proposition 4.7.8. *Let $\Delta' \in \mathbb{N}$ and let $\ell, m, r \in \mathbb{N}_0$ with $\Delta' \geq 3\ell + m$. Let G be a graph with vertex partition X, Y and let M be a matching in $G[X, Y]$ of size m. Let $\{x_1, \ldots, x_\ell\} \subseteq Y$ such that $d_Y(x_i) \geq \Delta'$ and $|\{x_1, \ldots, x_\ell\} \setminus V(M)| \geq r$. Then there exists a path system $\mathcal{P} \subseteq G[X, Y] \cup G[Y]$ such that $e_{\mathcal{P}}(Y) = \ell + r$, $\mathcal{P}[X, Y] = M$ and every edge of M lies in a distinct XY-path in $\mathcal{P}$.*

Proof. Since $\Delta' \geq 3\ell + m$, $G[Y]$ contains a collection of ℓ vertex-disjoint paths $P_1, \ldots, P_\ell$ of length two with midpoints $x_1, \ldots, x_\ell$ respectively, such that $V(P_i) \cap V(M) \subseteq \{x_i\}$. For each $x_i \in V(M)$, delete one arbitrary edge from P_i. Let $\mathcal{P}$ consist of M together with $P_1, \ldots, P_\ell$. Then $\mathcal{P}$ is a path system, and every edge of M lies in a distinct XY-path. Moreover, $e_{\mathcal{P}}(Y) \geq 2\ell - (\ell - r) = \ell + r$. Delete additional edges from $\mathcal{P}[Y]$ if necessary. $\square$

Proposition 4.7.9. *Let $0 < \varepsilon < 1/3$. Let $a, b \in \mathbb{R}_{\geq 0}$ and let $x \in \mathbb{N}_0$. Suppose that $2a + b \geq 2x$. Let $a' := \lceil a \rceil_\varepsilon$ and let b' be the largest even integer of size at most $\lceil b \rceil_\varepsilon$. Then $a', b' \geq 0$ and $2a' + b' \geq 2x$.*

Proof. Note that

$$2\lceil a \rceil_\varepsilon + \lceil b \rceil_\varepsilon = 2\lceil a - \varepsilon \rceil + \lceil b - \varepsilon \rceil \geq \lceil 2a - 2\varepsilon + b - \varepsilon \rceil \geq \lceil 2x - 3\varepsilon \rceil \geq 2x.$$

This implies the proposition. □

Proposition 4.7.10. *Let $D \in \mathbb{N}$ and let $0 < \varepsilon < 1/3$. Let G be a D-regular graph and let U, A, B be a partition of $V(G)$ where $|A| \geq |B|$. Suppose that $\Delta(G[A,U]), \Delta(G[A]) \leq D/2$ and that $\mathrm{char}_{D/2,\varepsilon}(G) = (\ell, m)$. Then $\ell, m \geq 0$ and $\ell + m/2 \geq |A| - |B|$.*

Proof. Proposition 3.7.4(ii) implies that $4e(A)/D + 2e(A,U)/D \geq 2(|A| - |B|)$. Apply Proposition 4.7.9 with $2e(A)/D, 2e(A,U)/D, |A|-|B|$ playing the roles of a, b, x to obtain a', b'. Note that $a' = \ell$ and $b' = m$. □

We will first prove Lemma 4.7.3 in the case when $|A| - |B| \geq 2$. This constraint arises for the following reason. We will show that we can find a path system $\mathcal{P}$ such that $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour, but $\mathcal{P}$ is 'overbalanced'. More precisely, $\mathrm{bal}_{AB}(\mathcal{P}) = \ell + m/2$, which is at least as large as $|A| - |B|$ by Proposition 4.7.10. We would like to remove edges from $\mathcal{P}$ so that (P2) holds, and $R_{\mathcal{V}}(\mathcal{P})$ is still an Euler tour. However, there exist path systems $\mathcal{P}_0$ such that $\mathrm{bal}_{AB}(\mathcal{P}_0) = 2$, $R_{\mathcal{V}}(\mathcal{P}_0)$ is an Euler tour, but any $\mathcal{P}_0'$ with $E(\mathcal{P}_0') \subsetneqq E(\mathcal{P}_0)$ is such that $R_{\mathcal{V}}(\mathcal{P}_0')$ is not an Euler tour. (For example, a matching of size two in $G[V_1, A]$ together with a matching of size two in $G[V_2, A]$, such that these edges are all vertex-disjoint.) So, if $|A| - |B| < 2$, we cannot guarantee, simply by removing edges, that we will ever be able to find $\mathcal{P}'$ with $\mathrm{bal}_{AB}(\mathcal{P}') = |A| - |B|$ without violating (P3).

We will split the case when $|A| - |B| \geq 2$ further into the subcases $m \geq 4$ and $m \leq 2$, i.e. when $e_G(A, V_1 \cup V_2)$ is at least a little larger than $3D/2$, and when it is not. We will call these the *dense* and *sparse* cases respectively.

4.7.5 The proof of Lemma 4.7.3 in the case when $|A| - |B| \geq 2$ and $m \geq 4$

This subsection concerns the dense case when $m \geq 4$, i.e. when $e_G(A, V_1 \cup V_2)$ is at least slightly larger than $3D/2$. Now $G[A, V_1 \cup V_2]$ contains a matching M of size m. We will add edges to M to obtain a path system $\mathcal{P}$ which satisfies (P1)–(P3). If $M[A, V_i]$ is an

even non-empty matching for both $i = 1, 2$, then M satisfies (P3). In every other case we must modify M by adding and/or subtracting edges. We do this separately depending on the relative values of $e_M(A, V_1)$ and $e_M(A, V_2)$. We thus obtain a path system $\mathcal{P}_0$ which satisfies (P1) and (P3). Then we obtain $\mathcal{P}$ by adding edges to $\mathcal{P}_0$ from $G[A]$ so that (P2) is also satisfied. We must pay attention to the way in which these sets of edges interact to ensure that $\mathcal{P}$ still satisfies (P3).

We begin with the subcase when $e_M(V_1, A), e_M(V_2, A)$ are both even and positive.

Lemma 4.7.11. *Let $\Delta, \Delta' \in \mathbb{N}$, $\ell \in \mathbb{N}_0$ and $m \in 2\mathbb{N}$ with $\Delta'/\Delta, m/\Delta', \ell/\Delta' \ll 1$. Let G be a graph with vertex partition $\mathcal{V} = \{V_1, V_2, W := A \cup B\}$. Let M be a matching in $G[V_1 \cup V_2, A]$ of size m, and let $M_i := M[V_i, A]$ and $m_i := e(M_i)$. Suppose that $\{m_1, m_2\} \subseteq 2\mathbb{N}$. Let $e(A) \geq (\ell - 1)\Delta + \Delta'$ and $\Delta(G[A]) \leq \Delta$. Then G contains a path system $\mathcal{P}$ such that $\mathcal{P} \subseteq G[A] \cup G[A, V_1 \cup V_2]$, $\mathcal{P}[A, V_1 \cup V_2] = M$, $e(\mathcal{P}) = \ell + m$, $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour and $\mathrm{bal}_{AB}(\mathcal{P}) = \ell + m/2$. Moreover, $\mathcal{P}$ contains at least one V_iA-path for each $i = 1, 2$.*

Proof. We will find $\mathcal{P}$ by adding suitable edges of $G[A]$ to M such that $\mathcal{P}$ contains at least one V_iA-path for each $i = 1, 2$. Then by Fact 4.7.2 we have that $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour. Apply Lemma 4.7.6 to $G[A]$. Suppose first that Lemma 4.7.6(i) holds. Let M' be a matching of size $\ell + 1$ in $G[A]$ and let $uv \in E(G[A])$ be such that $u \notin V(M')$. Then

$$\mathrm{bal}_{AB}(M \cup M') = \ell + m/2 + 1 \quad \text{and} \quad e(M \cup M') = \ell + m + 1. \tag{4.7.6}$$

If $M \cup M'$ contains a V_iA-path for both $i = 1, 2$ we are done by setting $\mathcal{P} := M \cup M' \setminus \{e\}$ where $e \in M'$ is arbitrary. Suppose now that $M \cup M'$ contains a V_1A-path but no V_2A-path. Then $V(M_2) \cap A \subseteq V(M')$. Choose $e_2 \in E(M')$ with an endpoint in $V(M_2)$. Then $\mathcal{P} := M \cup M' \setminus \{e_2\}$ contains a V_iA-path for both $i = 1, 2$, and (4.7.6) implies that $\mathrm{bal}_{AB}(\mathcal{P}) = \ell + m/2$ and $e(\mathcal{P}) = \ell + m$, as required. The case when $M \cup M'$ contains a V_2A-path but no V_1A-path is identical.

So we may assume that $M \cup M'$ contains no V_iA-path for both $i = 1, 2$. Suppose

that there is $a_1a_2 \in E(M')$ with $a_i \in V(M_i)$. Then $\mathcal{P} := M \cup M' \setminus \{a_1a_2\}$ contains a V_iA-path with endpoint a_i for $i = 1, 2$. Moreover, (4.7.6) implies that $\mathcal{P}$ satisfies the other conditions. Therefore we may assume that $M_i' := M'[V(M_i) \cap A]$ is a (non-empty) perfect matching for $i = 1, 2$. Choose $f_i \in E(M_i')$ for $i = 1, 2$ such that $v \in V(f_1) \cup V(f_2)$ if possible. We set $\mathcal{P} := M \cup M' \cup \{uv\} \setminus \{f_1, f_2\}$. Note that every vertex in $V(f_i) \setminus \{v\}$ is the endpoint of a V_iA-path in $\mathcal{P}$. Then (4.7.6) implies that $\mathrm{bal}_{AB}(\mathcal{P}) = \mathrm{bal}_{AB}(M \cup M') + 1 - 2 = \ell + m/2$ and $e(\mathcal{P}) = \ell + m$, as required.

Suppose instead that Lemma 4.7.6(ii) holds and let $x_1, \dots, x_\ell$ be ℓ distinct vertices in A with $d_A(x_i) \geq \Delta'$ for all $1 \leq i \leq \ell$. Apply Proposition 4.7.8 with $G \setminus B, V_1 \cup V_2, A, M, x_i, 0$ playing the roles of G, X, Y, M, x_i, r to obtain a path system $\mathcal{P} \subseteq G[A] \cup G[A, V_1 \cup V_2]$ with $e_{\mathcal{P}}(A) = \ell$, $\mathcal{P}[A, V_1 \cup V_2] = M$ and such that every edge in M lies in a distinct AV_i-path in $\mathcal{P}$ for some $i \in \{1, 2\}$. Therefore $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour, $e(\mathcal{P}) = \ell + m$, and since $V(\mathcal{P}) \cap B = \emptyset$ we have that $\mathrm{bal}_{AB}(\mathcal{P}) = \ell + m/2$. $\square$

We now consider the case when $e_M(V_1, A), e_M(V_2, A)$ are both odd and at least three.

Lemma 4.7.12. *Let $\Delta, \Delta' \in \mathbb{N}$, $\ell \in \mathbb{N}_0$ and $m \in 2\mathbb{N}$ with $\Delta'/\Delta, m/\Delta', \ell/\Delta' \ll 1$. Let G be a graph with vertex partition $\mathcal{V} = \{V_1, V_2, W := A \cup B\}$. Let $m < e_G(V_1 \cup V_2, A)$, $e_G(A) \geq (\ell - 1)\Delta + \Delta'$ and $\Delta(G[A]) \leq \Delta$. Let M be a matching in $G[V_1 \cup V_2, A]$ of size m, and let $M_i := M[V_i, A]$, $m_i := e(M_i)$. Suppose $\{m_1, m_2\} \subseteq 2\mathbb{N}+1$. Then G contains a path system $\mathcal{P}$ such that $e(\mathcal{P}) \leq \ell + m$, $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour and $\mathrm{bal}_{AB}(\mathcal{P}) = \ell + m/2$.*

Proof. We will find $\mathcal{P}$ such that $e_{\mathcal{P}}(V_i, A) = e_{\mathcal{P}}(V_i, W)$ is even for $i = 1, 2$, $e_{\mathcal{P}}(V_1, V_2) = 0$ and such that for each $X \in \mathcal{V}$, there exists $X' \in \mathcal{V} \setminus \{X\}$ such that $\mathcal{P}$ contains an XX'-path. Then by Fact 4.7.2 we have that $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour.

Let us first suppose that $\ell = 0$. Since $m < e_G(V_1 \cup V_2, A)$, there exists an edge $e^+ \in G[V_1 \cup V_2, A] \setminus E(M)$. Suppose, without loss of generality, that $e^+ \in G[V_1, A]$. Let e^- be an arbitrary edge in M_2. Let $\mathcal{P} := M \cup \{e^+\} \setminus \{e^-\}$. Then $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour and $\mathrm{bal}_{AB}(\mathcal{P}) = (m_1 + 1)/2 + (m_2 - 1)/2 = m/2$, as required.

Therefore we assume that $\ell \geq 1$. Apply Lemma 4.7.6 to $G[A]$. Suppose first that

Lemma 4.7.6(i) holds. So $G[A]$ contains a matching M' of size $\ell+1$. Note that it suffices to find $e_i \in M_i$ for $i = 1, 2$ such that $M \cup M' \setminus \{e_1, e_2\}$ contains a V_iA-path for $i = 1, 2$. Then it is straightforward to check that we are done by setting $\mathcal{P} := M \cup M' \setminus \{e_1, e_2\}$.

We say that $xy \in E(G[A])$ is a *connecting edge* if $x \in V(M_1)$ and $y \in V(M_2)$. Suppose that M' contains no connecting edge. So $M \cup M'$ contains no V_1V_2-paths. But an even number of edges in M_i lie in V_iV_i-paths of $M \cup M'$. Since m_i is odd, there must be a V_iA-path P_i in $M \cup M'$ for $i = 1, 2$. We are done by choosing $e_i \in E(M_i) \setminus E(P_i)$ arbitrarily.

Therefore we may assume that there exists a connecting edge $a_1a_2 \in M'$, with $a_i \in V(M_i)$. Suppose that there exists a second connecting edge $a_1'a_2' \in M'$, with $a_i' \in V(M_i)$. Then we are done by choosing $e_1 \in M_1$ with endpoint a_1 and $e_2 \in M_2$ with endpoint a_2'. Therefore we may suppose that a_1a_2 is the only connecting edge in G. Let P be the V_1V_2-path containing a_1a_2. Let $\mathcal{P}' := (M \cup M') \setminus \{E(P)\}$. Then, for each $i = 1, 2$, either $\mathcal{P}'$ contains a V_iA-path $P_{i,A}$, or a V_iV_i-path $P_{i,i}$. In the first case, let e_i be an arbitrary edge of M_i that does not lie in $P_{i,A}$. In the second case, let $e_i \in E(P_{i,i}) \cap E(M_i)$ be arbitrary.

Suppose instead that Lemma 4.7.6(ii) holds in $G[A]$ and let $x_1, \ldots, x_\ell$ be ℓ distinct vertices in A with $d_A(x_i) \geq \Delta'$ for all $1 \leq i \leq \ell$. Since $\ell \geq 1$, we can choose $e_1 \in M_1$ and $e_2 \in M_2$ so that $\{x_1, \ldots, x_\ell\} \not\subseteq V(M \setminus \{e_1, e_2\})$. Apply Proposition 4.7.8 with $G \setminus B, V_1 \cup V_2, A, M \setminus \{e_1, e_2\}, x_i, 1$ playing the roles of G, X, Y, M, x_i, r to obtain a path system $\mathcal{P} \subseteq G[A] \cup G[A, V_1 \cup V_2]$ such that $e_{\mathcal{P}}(A) = \ell + 1$, $\mathcal{P}[A, V_1 \cup V_2] = M \setminus \{e_1, e_2\}$, and every edge in $M \setminus \{e_1, e_2\}$ lies in a distinct AV_i-path in $\mathcal{P}$ for some $i \in \{1, 2\}$. Then $e(\mathcal{P}) = \ell + m - 1$ and $\mathrm{bal}_{AB}(\mathcal{P}) = \ell + 1 + (m-2)/2 = \ell + m/2$. Since $\mathcal{P}[A, V_i]$ is an even matching for $i = 1, 2$ and $\mathcal{P}[V_1, V_2]$ is empty, we have that $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour and we are done. $\square$

We now consider the case when $e_M(V_2, A)$ is odd and at least three, and $e_M(V_1, A) = 1$.

Lemma 4.7.13. *Let $\Delta, \Delta' \in \mathbb{N}$, $\ell \in \mathbb{N}_0$ and $m \in 2\mathbb{N}$ with $\Delta'/\Delta, m/\Delta', \ell/\Delta' \ll 1$. Let G be a 3-connected graph with vertex partition $\mathcal{V} = \{V_1, V_2, W := A \cup B\}$. Let*

$e_G(A) \geq (\ell-1)\Delta+\Delta'$ *and* $\Delta(G[A]) \leq \Delta$. *Let* M_2 *be a matching in* $G[V_2, A]$ *of size* $m-1$ *where* $3 \leq m-1 < e_G(V_2, A)$ *and let* $e_1 \in G[V_1, A]$ *be an edge not incident to* M_2. *Then* G *contains a path system* $\mathcal{P}$ *such that* $e(\mathcal{P}) \leq \ell + m + 2$, $R_{\mathcal{V}}(\mathcal{P})$ *is an Euler tour and* $\mathrm{bal}_{AB}(\mathcal{P}) = \ell + m/2$.

Proof. We will find a path system $\mathcal{P}$ such that, for each $X \in \mathcal{V}$, $e_{\mathcal{P}}(X, \overline{X})$ is even and there exists $X' \in \mathcal{V} \setminus \{X\}$ such that $\mathcal{P}$ contains an XX'-path. Then by Fact 4.7.2, $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour. We will choose $\mathcal{P}$ such that $\mathcal{P}[V_1 \cup V_2, W]$ is obtained from $M_2 \cup \{e_1\}$ by adding/removing at most one edge. Since G is 3-connected, G contains an edge v_1v with $v_1 \in V_1$ and $v \in V_2 \cup A \cup B$ such that vv_1 and e_1 are vertex-disjoint. We consider cases depending on the location of v.

Case 1. $v \in A$.

If possible, let e_2 be the edge of M_2 incident to v; otherwise, let e_2 be an arbitrary edge of M_2. Then we are done by applying Lemma 4.7.11 with $M_2 \cup \{e_1, v_1v\} \setminus \{e_2\}$ playing the role of M.

Case 2. $v \in V_2$.

If possible, choose $e_2 \in E(M_2)$ whose endpoint $v_2 \in V_2$ satisfies $v_2 = v$, otherwise let $e_2 \in E(M_2)$ be arbitrary. Set $V_1' := V_1 \cup \{v, v_2\}$ and $V_2' := V_2 \setminus \{v, v_2\}$. Observe that $e_{M_2 \cup \{e_1\}}(A, V_i') \in 2\mathbb{N}$ for $i = 1, 2$. Let $\mathcal{V}' := \{V_1', V_2', W\}$. Apply Lemma 4.7.11 with $G \setminus \{v_1\}, V_1', V_2', A, B, M_2 \cup \{e_1\}$ playing the roles of G, V_1, V_2, A, B, M to obtain a path system $\mathcal{P}'$ such that $\mathcal{P}' \subseteq G[A] \cup G[A, V_1' \cup V_2']$, $\mathcal{P}'[A, V_1' \cup V_2'] = M_2 \cup \{e_1\}$, $e(\mathcal{P}') = \ell + m$, $R_{\mathcal{V}'}(\mathcal{P}')$ is an Euler tour and $\mathrm{bal}_{AB}(\mathcal{P}') = \ell + m/2$. Moreover, $\mathcal{P}'$ contains at least one $V_i'A$-path for each $i = 1, 2$. Let P_i be such a path.

Let $\mathcal{P} := \mathcal{P}' \cup \{vv_1\}$. Then $e(\mathcal{P}) = \ell + m + 1$ and $\mathrm{bal}_{AB}(\mathcal{P}) = \ell + m/2$. Moreover, each of $e_{\mathcal{P}}(V_1, \overline{V_1}) = e_{\mathcal{P}'}(V_1', \overline{V_1'}) = 2$, $e_{\mathcal{P}}(V_2, \overline{V_2}) = e_{\mathcal{P}'}(V_2', \overline{V_2'}) + 2$ and $e_{\mathcal{P}}(W, \overline{W}) = e_{\mathcal{P}'}(W, \overline{W})$ is even. Now P_2 is a V_2A-path in $\mathcal{P}$. Similarly, if P_1 avoids e_2, then P_1 is a V_1A-path in $\mathcal{P}$. If P_1 contains e_2 and $v_2 = v$, then v_1vP_1 is a V_1A-path in $\mathcal{P}$. If $v_2 \neq v$ then v_1v is a V_1V_2-path in $\mathcal{P}$. Therefore, by Fact 4.7.2, $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour, as required.

Case 3. $v \in B$.

Apply Lemma 4.7.6 to $G[A]$. Suppose first that Lemma 4.7.6(i) holds. Let M' be a matching of size $\ell + 1$ in $G[A]$ and let $uw \in E(G[A])$ with $u \notin V(M')$. Apply Proposition 4.7.7 with $G \setminus B, V_1 \cup V_2, A, M_2 \cup \{e_1\}, M', u, w$ playing the roles of G, X, Y, M, M', u, v to obtain a path system $\mathcal{P}_0$ such that $\mathcal{P}_0[V_1 \cup V_2, A] = M_2 \cup \{e_1\}$; $e_{\mathcal{P}_0}(A) = \ell + 1$; and $\mathcal{P}_0$ contains at least two $(V_1 \cup V_2)A$-paths. But $\mathcal{P}_0$ contains at most one V_1A-path, and hence at least one V_2A-path P. Now Proposition 4.7.7(i) implies that $e_P(V_2, A) = 1$. So we can choose $e \in E(\mathcal{P}_0[V_2, A]) \setminus E(P)$. Let $\mathcal{P} := \mathcal{P}_0 \cup \{v_1v\} \setminus \{e\}$. Then $e_{\mathcal{P}}(X, \overline{X})$ is even for all $X \in \{V_1, V_2, W\}$ and $\mathcal{P}$ contains a V_1B-path and a V_2A-path. Moreover, $\mathrm{bal}_{AB}(\mathcal{P}) = e_{\mathcal{P}_0}(A) + e_{\mathcal{P}_0}(A, V_1 \cup V_2)/2 - 1 = \ell + m/2$, as required.

Suppose instead that Lemma 4.7.6(ii) holds. Then $G[A]$ contains ℓ distinct vertices $x_1, \ldots, x_\ell$ such that $d_A(x_i) \geq \Delta'$ for all $1 \leq i \leq \ell$. Choose $e \in E(G[V_2, A]) \setminus E(M_2)$. If $\ell = 0$ then $\mathcal{P} := M_2 \cup \{e_1, v_1v, e\}$ is as required. Suppose now that $\ell = 1$. Let $w_1, y_1 \in N_A(x_1) \setminus V(M_2 \cup \{e_1\})$ be distinct. Suppose that $x_1 \notin V(e_1)$. If possible, choose e_2 to be the edge of M_2 that contains x_1; otherwise, let e_2 be an arbitrary edge of M_2. In this case we let $\mathcal{P} := M_2 \cup \{e_1, v_1v, w_1x_1y_1\} \setminus \{e_2\}$. Suppose now that $x_1 \in V(e_1)$. In this case we let $\mathcal{P} := M_2 \cup \{e_1, v_1v, e\} \cup \{x_1y_1\}$. In all cases, we have that $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour, $e(\mathcal{P}) \leq \ell + m + 2$ and $\mathrm{bal}_{AB}(\mathcal{P}) = m/2 + 1$, as required.

Suppose finally that $\ell \geq 2$. Then we can choose $e_2 \in M_2$ so that $\{x_1, \ldots, x_\ell\} \not\subseteq V(M_2 \cup \{e_1\} \setminus \{e_2\})$. Apply Proposition 4.7.8 with $G \setminus B, V_1 \cup V_2, A, M_2 \cup \{e_1\} \setminus \{e_2\}, x_i, 1$ playing the roles of G, X, Y, M, x_i, r to obtain a path system $\mathcal{P}_0$ in $G[A] \cup G[A, V_1 \cup V_2]$ such that $e_{\mathcal{P}_0}(A) = \ell + 1$, $\mathcal{P}_0[A, V_1 \cup V_2] = M_2 \cup \{e_1\} \setminus \{e_2\}$, and every edge in $M_2 \cup \{e_1\} \setminus \{e_2\}$ lies in a distinct AV_i-path in $\mathcal{P}_0$ for some $i \in \{1, 2\}$. Let $\mathcal{P} := \mathcal{P}_0 \cup \{v_1v\}$. Then $e(\mathcal{P}) = \ell + m + 1$ and

$$\mathrm{bal}_{AB}(\mathcal{P}) = e_{\mathcal{P}_0}(A) + e_{\mathcal{P}_0}(A, V_1 \cup V_2)/2 - 1/2 = \ell + 1 + (m-1)/2 - 1/2 = \ell + m/2.$$

Note finally that $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour by Fact 4.7.2. $\square$

We are now ready to prove a more general version of Lemmas 4.7.11–4.7.13 in which $G[A, V_1 \cup V_2]$ contains an arbitrary even matching of size at least four.

Lemma 4.7.14. *Let $\Delta, \Delta' \in \mathbb{N}$, $\ell \in \mathbb{N}_0$ and $m \in 2\mathbb{N}$ with $\Delta'/\Delta, m/\Delta', \ell/\Delta' \ll 1$ and $m \geq 4$. Let $\Delta'/\Delta < \varepsilon < 1/3$. Let G be a 3-connected graph with vertex partition $\mathcal{V} = \{V_1, V_2, W := A \cup B\}$. Suppose that $\Delta(G[A]), \Delta(G[A, V_1 \cup V_2]) \leq \Delta$ and $\mathrm{char}_{\Delta,\varepsilon}(G) = (\ell, m)$. Then G contains a path system $\mathcal{P}$ such that $e(\mathcal{P}) \leq \ell + m + 4$, $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour and $\mathrm{bal}_{AB}(\mathcal{P}) = \ell + m/2$.*

Proof. Write $U := V_1 \cup V_2$. Proposition 4.7.5 implies that

$$e_G(A) \geq (\ell - 1)\Delta + \Delta' \quad \text{and} \quad e_G(A, U) \geq (m - 1)\Delta + \Delta'. \tag{4.7.7}$$

Recall also that $m \leq \lceil e_G(A, U)/\Delta \rceil$ and m is even. Choose non-negative integers b_1, b_2 such that $b_i \leq \lceil e_G(A, V_i)/\Delta \rceil$ for $i = 1, 2$ and $b_1 + b_2 = m$. Apply Lemma 4.6.7 with $G[A, U], A, V_1, V_2$ playing the roles of G, U, V, W to obtain a matching M in $G[A, U]$ such that $e_M(A, V_i) = b_i$ for $i = 1, 2$. Without loss of generality we assume that $b_1 \leq b_2$. Suppose first that b_1, b_2 are both even and positive. Then we are done by applying Lemma 4.7.11. If b_1, b_2 are both odd and at least three, then we are done by applying Lemma 4.7.12. Suppose that $b_1 = 1$. Then $\lceil e_G(A, V_2)/\Delta \rceil \geq b_2 = m - 1$ so $m - 1 < e_G(A, V_2)$. Therefore we can apply Lemma 4.7.13 with M playing the role of $M_2 \cup \{e_1\}$. So we can assume that $b_1 = 0$, and hence that $M \subseteq G[A, V_2]$. Suppose that $e_G(A, V_1) > 0$. Then there is an edge $e \in E(G[A, V_1])$ and $m - 1$ edges in M which are not incident with e. We are similarly done by applying Lemma 4.7.13. The only remaining case is when $e_G(A, V_1) = 0$. Now (4.7.7) implies that

$$e_G(A, V_2) \geq (m - 1)\Delta + \Delta'. \tag{4.7.8}$$

Since G is 3-connected, $G[V_1, \overline{V_1}]$ contains a matching of size three. So $G[V_1, V_2 \cup B]$ contains a matching of size three. Then at least one of $G[V_1, V_2]$, $G[V_1, B]$ contains a

matching of size two.

Case 1. *$G[V_1, V_2]$ contains a matching M^* of size two.*

Choose two distinct edges $e_2, e_2' \in E(M)$ such that $|V(M^*) \cap \{v_2, v_2'\}|$ is as large as possible, where v_2, v_2' are the endvertices of e_2, e_2' in V_2. Set $V_1' := V_1 \cup \{v_2, v_2'\}$ and $V_2' := V_2 \setminus \{v_2, v_2'\}$. Observe that $e_M(A, V_i') \in 2\mathbb{N}$ for $i = 1, 2$ since $m \geq 4$. Let $\mathcal{V}' := \{V_1', V_2', W\}$. Apply Lemma 4.7.11 with G, V_1', V_2', A, B, M playing the roles of G, V_1, V_2, A, B, M to obtain a path system $\mathcal{P}'$ such that $\mathcal{P}' \subseteq G[A] \cup G[A, V_1' \cup V_2']$, $\mathcal{P}'[A, V_1' \cup V_2'] = M$, $e(\mathcal{P}') = \ell + m$, $R_{\mathcal{V}'}(\mathcal{P}')$ is an Euler tour and $\mathrm{bal}_{AB}(\mathcal{P}') = \ell + m/2$. Moreover, $\mathcal{P}'$ contains at least one $V_i'A$-path for each $i = 1, 2$. Let P_i be such a path. Then P_1 contains either e_2 or e_2'. Without loss of generality we may assume that P_1 contains e_2.

Let $\mathcal{P} := \mathcal{P}' \cup M^*$. Then $e(\mathcal{P}) = \ell + m + 2$ and $\mathrm{bal}_{AB}(\mathcal{P}) = \ell + m/2$. Moreover, each of $e_{\mathcal{P}}(V_1, \overline{V_1}) = e_{\mathcal{P}'}(V_1', \overline{V_1'}) = 2$, $e_{\mathcal{P}}(V_2, \overline{V_2}) = e_{\mathcal{P}'}(V_2', \overline{V_2'}) + 4$ and $e_{\mathcal{P}}(W, \overline{W}) = e_{\mathcal{P}'}(W, \overline{W})$ is even. Now P_2 is an V_2A-path in $\mathcal{P}$. If M^* contains an edge e which avoids both v_2, v_2' (and thus is vertex-disjoint from all edges in M), then e is a V_1V_2-path in $\mathcal{P}$. If there is no such edge e, then M^* contains an edge e' whose endvertex in V_2 is v_2. Then $e' \cup P_1$ is a V_1A-path in $\mathcal{P}$. Therefore, by Fact 4.7.2, $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour, as required.

Case 2. *$G[V_1, B]$ contains a matching M^* of size two.*

Apply Lemma 4.7.6 to $G[A]$. Suppose first that Lemma 4.7.6(i) holds. Then $G[A]$ contains a matching M' of size $\ell + 1$ and an edge uv with $u \notin V(M')$. Apply Proposition 4.7.7 with $G \setminus B, V_1 \cup V_2, A, M, M', u, v$ playing the roles of G, X, Y, M, M', u, v to obtain a path system $\mathcal{P}_0$ such that $\mathcal{P}_0[V_1 \cup V_2, A] = M$; $\mathcal{P}_0 \subseteq M \cup M' \cup \{uv\}$; $e_{\mathcal{P}_0}(A) = \ell + 1$; and $\mathcal{P}_0$ contains at least two V_2A-paths. Let $\mathcal{P} := \mathcal{P}_0 \cup M^*$. Then $\mathcal{P}$ contains at least two V_2A-paths and two V_1B-paths (namely the edges of M^*), so $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour. Moreover $\mathrm{bal}_{AB}(\mathcal{P}) = \ell + m/2$ and $e(\mathcal{P}) = \ell + m + 3$, as required.

Suppose now that Lemma 4.7.6(ii) holds in $G[A]$. Assume first that $\ell \geq 2$. Let $x_1, \ldots, x_\ell$ be ℓ distinct vertices in A such that $d_A(x_i) \geq \Delta'$ for $1 \leq i \leq \ell$. Since $m \geq 4$,

we can choose distinct $e_1, e_2 \in M$ such that $|\{x_1, \dots, x_\ell\} \setminus V(M \setminus \{e_1, e_2\})| \geq 2$. Then Proposition 4.7.8 applied with $G \setminus B, V_1 \cup V_2, A, M \setminus \{e_1, e_2\}, x_i, 2$ playing the roles of G, X, Y, M, x_i, r implies that there is a path system $\mathcal{P}' \subseteq G[A] \cup G[A, V_1 \cup V_2]$ such that $e_{\mathcal{P}'}(A) = \ell + 2$, $\mathcal{P}'[A, V_1 \cup V_2] = M \setminus \{e_1, e_2\}$, and such that every edge of $M \setminus \{e_1, e_2\}$ lies in a distinct AV_2-path. Let $\mathcal{P} := \mathcal{P}' \cup M^*$. Then $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour, $e(\mathcal{P}) = \ell + m + 2$, and

$$\mathrm{bal}_{AB}(\mathcal{P}) = e_{\mathcal{P}'}(A) + e_{\mathcal{P}'}(A, V_1 \cup V_2)/2 - 1 = \ell + 2 + (m-2)/2 - 1 = \ell + m/2.$$

Finally we consider the case when $\ell \leq 1$. Lemma 4.7.6 applied to $G[A, V_1 \cup V_2]$ and (4.7.8) imply that $G[A, V_1 \cup V_2]$ contains a matching M' of size m together with a matching M^+ of size two which is edge-disjoint from M', such that both edges in M^+ contain a vertex outside of $V(M')$. Since $e_G(A, V_1) = 0$ by our assumption, we have $M' \cup M^+ \subseteq G[A, V_2]$. Suppose first that $\ell = 0$. In this case we let $\mathcal{P} := M' \cup M^+ \cup M^*$. It is clear that $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour, $e(\mathcal{P}) = m + 4$ and $\mathrm{bal}_{AB}(\mathcal{P}) = m/2$, as required. The final case is when $\ell = 1$. Choose $e \in M^+$ and $e' \in M'$ such that $|V(e) \cap \{x_1\}| + |V(e') \cap \{x_1\}|$ is maximal. So $\mathcal{P}' := M' \cup M^+ \setminus \{e, e'\}$ is a matching of size $m - 1$ together with an extra edge, and $x_1 \notin V(\mathcal{P}')$. In particular, $\mathcal{P}'$ contains a V_2A-path P_2. Since $m/\Delta' \ll 1$, we can choose distinct vertices w_1, y_1 in $N_A(x_1) \setminus V(\mathcal{P}')$. Let $\mathcal{P} := \mathcal{P}' \cup M^* \cup \{w_1 x_1 y_1\}$. Then P_2 is a V_2A-path in $\mathcal{P}$ and each edge of M^* is a V_1B-path in $\mathcal{P}$. So Fact 4.7.2 implies that $\mathcal{P}$ is an Euler tour. Moreover, $\mathrm{bal}_{AB}(\mathcal{P}) = m/2 + 1$, and $e(\mathcal{P}) = m + 4$, as required. $\square$

The proof of Lemma 4.7.3 in the 'dense' case is now just a short step away.

Proof of Lemma 4.7.3 *in the case when* $|A| - |B| \geq 2$ *and* $m \geq 4$. Let $\Delta := D/2$. Observe that $d_A(a) \leq d_B(a)$ for all $a \in A$ implies that $\Delta(G[A]) \leq \Delta$. Proposition 4.7.10 implies that $\ell + m/2 \geq |A| - |B|$. Choose non-negative integers $\ell' \leq \ell$ and $m' \leq m$ such that m' is even, $\ell' + m'/2 = |A| - |B|$ and $m' \geq 4$. This is possible since $|A| - |B| \geq 2$. Let $\Delta' := \nu n$. Proposition 4.7.4 implies that $\ell', m' \leq 12\rho n$. Then $\Delta'/\Delta \ll 1$, $m'/\Delta' \ll 1$, $\ell'/\Delta' \ll 1$, $\Delta'/\Delta < \varepsilon$. Apply Lemma 4.7.14 with ℓ', m' playing the roles of ℓ, m to obtain

a path system $\mathcal{P}$ such that $e(\mathcal{P}) \leq \ell' + m' + 4 \leq \ell + m + 4$, $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour, and $\mathrm{bal}(\mathcal{P}) = \ell' + m'/2 = |A| - |B|$. So (P1)–(P3) hold. $\square$

4.7.6 The proof of Lemma 4.7.3 in the case when $|A| - |B| \geq 2$ and $m \leq 2$

We now deal with the sparse case, i.e. when the largest even matching we can guarantee between A and $V_1 \cup V_2$ has size at most two. For this, we need to introduce some notation which will be used in all of the remaining cases.

More notation and tools

The remaining cases are quite delicate and we are forced to now introduce some further notation, which we will attempt to motivate.

Given a graph G containing a path system $\mathcal{P}$, and $A \subseteq V(G)$, we write

$$F_{\mathcal{P}}(A) := (a_1, a_2) \tag{4.7.9}$$

when a_i is the number of vertices in A of degree i in $\mathcal{P}$ for $i = 1, 2$. Note that, if $e_{\mathcal{P}}(A) = 0$, then

$$e_{\mathcal{P}}(A, \overline{A}) = a_1 + 2a_2. \tag{4.7.10}$$

Before defining a 'basic connector', we give some motivation for this concept. Let $\mathcal{P}$ be a path system such that $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour. Let $\mathcal{P}_0$ be be a minimal subgraph of $\mathcal{P}$ such that $R_{\mathcal{V}}(\mathcal{P}_0)$ is an Euler tour. We will call such a path system $\mathcal{P}_0$ a 'basic connector'. So a basic connector satisfies (P1) and (P3), but not necessarily (P2). It is not hard to see that $-2 \leq \mathrm{bal}_{AB}(\mathcal{P}_0) \leq 2$ and $e(\mathcal{P}_0) \leq 4$ (see Proposition 4.7.15). So in the case when $|A| - |B| \geq 2$, we can find an 'overbalanced' path system $\mathcal{P}$ (with $\mathrm{bal}_{AB}(\mathcal{P}) \geq |A| - |B|$) and simply remove edges to obtain a $\mathcal{V}$-tour. We did this when $m \geq 4$ in the previous subsection. This extra condition guaranteed the presence of a

large matching in $G[A, U]$ which we used to suitably connect the components.

In this section, however, we have $m \leq 2$ (recall that m is even). So $G[A, U]$ may not contain a large matching, and so connecting the components may be difficult. Therefore we use a basic connector as the foundation of our $\mathcal{V}$-tour.

The final two subsections concern the case when $|A| - |B| \leq 1$. Now, as well as satisfying (P1) and (P3), any basic connector $\mathcal{P}_0$ is very close to being balanced; in fact $|\mathrm{bal}_{AB}(\mathcal{P}_0) - (|A| - |B|)| \leq 3$. So here we find the basic connector $\mathcal{P}_0$ in G which is closest to what we want, and carefully modify it.

Formally, we say that $\mathcal{P}$ is a *basic connector* (for $\mathcal{V} = \{V_1, V_2, W := A \cup B\}$) if

(BC1) $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour;

(BC2) $e(\mathcal{P}) \leq 4$ and $|\mathrm{bal}_{AB}(\mathcal{P})| \leq 2$;

(BC3) $e_{\mathcal{P}}(A \cup B) = 0$;

(BC4) if $F_{\mathcal{P}}(A) = (a_1, a_2)$ then $\mathrm{bal}_{AB}(\mathcal{P}) \in \{a_1 + 2a_2 - 2, a_1 + 2a_2 - 1\}$ and $a_2 \leq 1$.

It can be shown that (BC1)–(BC3) imply (BC4) (cf. the proof of Proposition 4.7.15). Observe (BC3) implies that if $\mathcal{P}$ is a basic connector, then

$$2\mathrm{bal}_{AB}(\mathcal{P}) = e_{\mathcal{P}}(A, V_1 \cup V_2) - e_{\mathcal{P}}(B, V_1 \cup V_2) = a_1 + 2a_2 - e_{\mathcal{P}}(B, V_1 \cup V_2). \tag{4.7.11}$$

Roughly speaking, the existence of a basic connector $\mathcal{P}$ follows from 3-connectivity. We would like to modify/extend $\mathcal{P}$ into a path system $\mathcal{P}'$ which balances the sizes of A, B, i.e. for which $\mathrm{bal}_{AB}(\mathcal{P}') = |A| - |B|$. The following notion will be very useful for this. Given a graph G, disjoint $A_1, A_2 \subseteq V(G)$ and $t \in \mathbb{N}_0$, we say that

$$\mathrm{acc}(G; A_1, A_2) \geq t$$

if G contains a path system $\mathcal{P}$ such that

(A1) $e(\mathcal{P}) = t$;

(A2) $d_{\mathcal{P}}(x_2) = 0$ for each $x_2 \in A_2$;

(A3) $d_{\mathcal{P}}(x_1) \leq 1$ for each $x_1 \in A_1$, and no path of $\mathcal{P}$ has both endpoints in A_1.

We say that such a $\mathcal{P}$ *accommodates* A_1, A_2, where 'acc' is short for 'accomodating'.

In a typical application of this notion, we have already constructed a path system $\mathcal{P}_0$. We let A_1 be the set of all those vertices in A which have degree one in $\mathcal{P}_0$ and A_2 be the set of all those vertices in A which have degree two in $\mathcal{P}_0$. Then, if $\mathrm{acc}(G[A]; A_1, A_2) \geq t$, we can find a path system $\mathcal{P}$ in $G[A]$ with t edges such that $\mathcal{P}_0 \cup \mathcal{P}$ is also a path system.

We now collect some tools which will be used to prove Lemma 4.7.3 in the case when $|A| - |B| \geq 2$ and $m \leq 2$. The next proposition uses Lemma 3.7.3 to show that G contains a basic connector.

Proposition 4.7.15. *Let G be a 3-connected graph with vertex partition $\mathcal{V} = \{V_1, V_2, W := A \cup B\}$. Then G contains a basic connector $\mathcal{P}$.*

Proof. Apply Lemma 3.7.3 to G and $\mathcal{V}$ to obtain a path system $\mathcal{P}$ satisfying the conditions (i)–(iii). We claim that $\mathcal{P}$ is a basic connector. Write $F_{\mathcal{P}}(A) = (a_1, a_2)$ and $F_{\mathcal{P}}(B) = (b_1, b_2)$. In particular, (iii) implies that

$$a_1 + b_1 + 2(a_2 + b_2) \in \{2, 4\} \tag{4.7.12}$$

and $a_2 + b_2 \leq 1$. Note that (BC1) and (BC3) are immediate from (ii) and (i) respectively. Moreover, (i) implies $e_{\mathcal{P}}(A \cup B) = 0$. So $e_{\mathcal{P}}(A, V_1 \cup V_2) = a_1 + 2a_2$ and $e_{\mathcal{P}}(B, V_1 \cup V_2) = b_1 + 2b_2$. So (4.7.12) implies that

$$2\mathrm{bal}_{AB}(\mathcal{P}) = a_1 + 2a_2 - b_1 - 2b_2 \in \{2a_1 + 4a_2 - 4, 2a_1 + 4a_2 - 2\}$$

and $|2\mathrm{bal}_{AB}(\mathcal{P})| \leq 4$, so (BC2) and (BC4) hold. □

By Proposition 4.7.15, we can find a basic connector $\mathcal{P}_0$ in G, which may not satisfy (P2). Our aim now is to find a suitable path system $\mathcal{P}_A$ in $G[A]$ so that $\mathcal{P}_0 \cup \mathcal{P}_A$ satisfies

(P1)–(P3). Let A_i be the collection of all those vertices of A with degree i in $\mathcal{P}_0$. The next result shows that it suffices to show that $\mathrm{acc}(G[A]; A_1, A_2) \geq |A| - |B| - \mathrm{bal}_{AB}(\mathcal{P}_0)$.

Proposition 4.7.16. *Let G be a graph with vertex partition $\mathcal{V} = \{V_1, V_2, W := A \cup B\}$. Let $\mathcal{P}_0$ be a basic connector in G and for $i = 1, 2$ let A_i be the collection of all those vertices of A with degree i in $\mathcal{P}_0$. Then, for any integer $0 \leq t \leq \mathrm{acc}(G[A]; A_1, A_2)$, we have that G contains a path system $\mathcal{P}$ such that $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour, $\mathrm{bal}_{AB}(\mathcal{P}) = \mathrm{bal}_{AB}(\mathcal{P}_0) + t$ and $e(\mathcal{P}) \leq t + 4$.*

Proof. Let $\mathcal{P}_A$ be a path system in $G[A]$ which accommodates A_1, A_2 such that $e(\mathcal{P}_A) = t$. Let $\mathcal{P} := \mathcal{P}_0 \cup \mathcal{P}_A$. Properties (A2) and (A3) imply that $\mathcal{P}$ is a path system. It is straightforward to check that (BC1) implies that $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour. Moreover, $\mathrm{bal}_{AB}(\mathcal{P}) = \mathrm{bal}_{AB}(\mathcal{P}_0) + e(\mathcal{P}_A)$, as required. Finally, (BC2) gives the required bound on $e(\mathcal{P})$. □

Building a basic connector from a matching

The next lemma shows that in the case when $G[A, V_1 \cup V_2]$ contains a matching of size at least three, we can obtain a basic connector with additional useful properties.

Lemma 4.7.17. *Let G be a 3-connected graph with vertex partition $\mathcal{V} = \{V_1, V_2, W := A \cup B\}$. Suppose that $G[A, V_1 \cup V_2]$ contains a matching M of size three. Then one of the following holds:*

(i) G contains a basic connector $\mathcal{P}$ with $\mathrm{bal}_{AB}(\mathcal{P}) \geq 1$, and if $F_{\mathcal{P}}(A) = (a_1, a_2)$, then $a_1 \geq 2$;

(ii) $e_G(A, V_i) = 0$ for some $i \in \{1, 2\}$, and for each $a \in A$, G contains matchings $M_{a,A}, M_{a,B}$ in $G[A \setminus \{a\}, V_j], G[B, V_i]$ respectively, where $j \in \{1, 2\} \setminus \{i\}$, each of which has size two. In particular, $\mathcal{P}_a := M_{a,A} \cup M_{a,B}$ is a basic connector with $\mathrm{bal}_{AB}(\mathcal{P}_a) = 0$, $a \notin V(\mathcal{P}_a)$ and $F_{\mathcal{P}}(A) = (2, 0)$.

Proof. Without loss of generality we may assume that $e_M(A, V_2) \geq e_M(A, V_1)$. Suppose first that $e_G(A, V_1) > 0$. We claim that $G[A, V_1 \cup V_2]$ contains a matching M' of size three such that $e_{M'}(A, V_1) = 1$ and $e_{M'}(A, V_2) = 2$. To see this, we may assume that we cannot set $M' := M$, so $M \subseteq G[A, V_2]$. Let $e_1 \in E(G[A, V_1])$. Then $V(e_1) \cap V(M) \subseteq A$. If possible, let e' be the edge of M incident to e_1, otherwise let $e' \in E(M)$ be arbitrary. Let $M' := M \cup \{e_1\} \setminus \{e'\}$, proving the claim.

Since G is 3-connected, there exists $e \in E(G[V_1, \overline{V_1}])$ that is not incident with the unique edge $e_1 \in M'[A, V_1]$. Let x be the endpoint of e that does not lie in V_1. If $x \in V_2$ then we can choose $e_2 \in M'[A, V_2]$ which is not incident with e and then $\mathcal{P} := \{e, e_1, e_2\}$ is a path system with $\mathrm{bal}_{AB}(\mathcal{P}) = 1$ and $F_{\mathcal{P}}(A) = (2, 0)$. It is easy to check that $\mathcal{P}$ is a basic connector, so (i) holds. If $x \in A \cup B$ then similarly $\mathcal{P} := M' \cup \{e\}$ satisfies (i).

Suppose now that $e_G(A, V_1) = 0$. Thus $e_M(A, V_2) = 3$. Since G is 3-connected, there is a matching M' of size three in $G[V_1, \overline{V_1}]$. Let $E(M') = \{e_1, e_2, e_3\}$ and let x_1, x_2, x_3 respectively be the endpoints of e_1, e_2, e_3 which do not lie in V_1. Note that $\{x_1, x_2, x_3\} \subseteq B \cup V_2$. Suppose first that $|V(M') \cap B| \leq 1$. Without loss of generality we assume that $\{x_1, x_2\} \subseteq V_2$. Let $e, e' \in E(M)$ be such that $\{x_1, x_2\} \not\subseteq V(\{e, e'\})$. Then $\mathcal{P} := \{e, e', e_1, e_2\}$ is such that $\mathrm{bal}_{AB}(\mathcal{P}) = 1$ and $F_{\mathcal{P}}(A) = (2, 0)$. Moreover, $\mathcal{P}$ is a basic connector, so (i) holds. So without loss of generality we may assume that $|V(M') \cap B| \geq 2$ and $\{x_1, x_2\} \subseteq B$. Given an arbitrary $a \in A$, choose $e, e' \in E(M)$ such that $a \notin V(\{e, e'\})$. Let $M_{a,A} := \{e, e'\}$ and $M_{a,B} := \{e_1, e_2\}$. So (ii) holds. $\square$

We now show how this result implies that, whenever $G[A, V_1 \cup V_2]$ contains a matching of size two, we are again able to find a basic connector with additional useful properties (though not as useful as those in Lemma 4.7.17).

Lemma 4.7.18. *Let G be a 3-connected graph with vertex partition $\mathcal{V} = \{V_1, V_2, W := A \cup B\}$. Let M be a matching in $G[A, V_1 \cup V_2]$ of size two. Then G contains a basic connector $\mathcal{P}$ with $\mathrm{bal}_{AB}(\mathcal{P}) \geq 0$, and if $F_{\mathcal{P}}(A) = (a_1, a_2)$, then $a_1 \geq 1$.*

Proof. Write $U := V_1 \cup V_2$. Since G is 3-connected, $G[A \cup B, U]$ contains a matching M'

of size three. We claim that $M \cup M'$ contains a matching M^* of size three such that at least two of the edges in M^* lie in $G[A, U]$. To see this, assume that $e_{M'}(A, U) \leq 1$ (or we could take $M^* := M'$). Assume further that there is no edge $e \in E(M')$ without an endpoint in $V(M)$ (or we could take $M^* := M \cup \{e\}$). Then, if we write $M := \{au, a'u'\}$ where $a, a' \in A$ and $u, u' \in U$, we have that M' consists of distinct edges $e_u, e_{u'}, e$ incident with u, u' and $\{a, a'\}$ respectively. Suppose that $a \in V(e)$. Then $e \in E(G[A, U])$ and so $e_u, e_{u'} \in E(G[B, U])$. Moreover, neither e nor e_u is incident with $a'u'$. We can set $M^* := \{a'u', e, e_u\}$. If instead $a' \in V(e)$, then we can set $M^* := \{au, e, e_{u'}\}$. This proves the claim.

If $M^* \subseteq G[A, U]$, we are done by Lemma 4.7.17. Otherwise, let bu be the unique edge in $M^*[B, U]$ with $u \in U$ and $b \in B$. Let $A' := A \cup \{b\}$ and $B' := B \setminus \{b\}$. Apply Lemma 4.7.17 with G, M^*, A', B' playing the roles of G, M, A, B. Suppose first that (i) holds. Then G contains a basic connector $\mathcal{P}$ with $\mathrm{bal}_{A'B'}(\mathcal{P}) \geq 1$. But $\mathrm{bal}_{AB}(\mathcal{P}) = \mathrm{bal}_{A'B'}(\mathcal{P}) - d_{\mathcal{P}}(b)$ if $b \in V(\mathcal{P})$ and $\mathrm{bal}_{AB}(\mathcal{P}) = \mathrm{bal}_{A'B'}(\mathcal{P})$ otherwise. If $d_{\mathcal{P}}(b) = 1$ then $\mathrm{bal}_{AB}(\mathcal{P}) \geq 0$, as required. Suppose that $d_{\mathcal{P}}(b) = 2$. Write $F_{\mathcal{P}}(A') := (a'_1, a'_2)$. Thus $a'_2 = 1$ by (BC4). Moreover, Lemma 4.7.17(i) implies that $a'_1 \geq 2$. Now $a'_1 + 2a'_2 \leq \mathrm{bal}_{A'B'}(\mathcal{P}) + 2 \leq 4$ by (BC2) and (BC4), so $(a'_1, a'_2) = (2, 1)$ and $\mathrm{bal}_{A'B'}(\mathcal{P}) = 2$. Then $\mathrm{bal}_{AB}(\mathcal{P}) \geq 0$, as required. Let $F_{\mathcal{P}}(A) =: (a_1, a_2)$. As above, $(a_1, a_2) \in \{(a'_1 - 1, a'_2), (a'_1, a'_2 - 1), (a'_1, a'_2)\}$. So $a_1 \geq a'_1 - 1 \geq 1$ by Lemma 4.7.17(i). Suppose instead that Lemma 4.7.17(ii) holds. The 'in particular' part implies that G contains a basic connector $\mathcal{P}_b$ with $\mathrm{bal}_{A'B'}(\mathcal{P}_b) = 0$, $F_{\mathcal{P}_b}(A) = (2, 0)$ and $b \notin V(\mathcal{P}_b)$. Then $\mathrm{bal}_{AB}(\mathcal{P}_b) = \mathrm{bal}_{A'B'}(\mathcal{P}_b)$, and $F_{\mathcal{P}_b}(A) = F_{\mathcal{P}_b}(A')$ as required. □

Accommodating path systems

The following proposition gives a lower bound for $\mathrm{acc}(G; A_1, A_2)$ whenever G contains several vertices of degree much larger than $|A_1| + |A_2|$ (i.e. when Lemma 4.7.6(ii) holds in G).

Proposition 4.7.19. *Let $\Delta' \in \mathbb{N}$ and let $\ell, a_1, a_2 \in \mathbb{N}_0$ be such that $\Delta' \geq 3\ell + a_1 + a_2$.*

Let G be a graph and let X be a collection of ℓ vertices in G such that $d_G(x) \geq \Delta'$ for all $x \in X$. Then for all disjoint $A_1, A_2 \subseteq V(G)$ with $|A_i| = a_i$ for $i = 1, 2$, we have

$$\mathrm{acc}(G; A_1, A_2) \geq 2\ell - |X \cap A_1| - 2|X \cap A_2|.$$

Proof. Write $X := \{x_1, \ldots, x_\ell\}$. Since $\Delta' \geq 3\ell + a_1 + a_2$ we can choose distinct vertices $w_1, \ldots, w_\ell, y_1, \ldots, y_\ell$ such that $\{w_i, y_i\} \subseteq N(x_i) \setminus (A_1 \cup A_2 \cup X)$. For each $1 \leq i \leq \ell$, define

$$P_i := \begin{cases} x_i y_i & \text{if } x_i \in A_1; \\ \emptyset & \text{if } x_i \in A_2; \\ w_i x_i y_i & \text{otherwise.} \end{cases} \tag{4.7.13}$$

Then $\mathcal{P} := \bigcup_{1 \leq i \leq \ell} P_i$ is a path system which accommodates A_1, A_2. Clearly

$$\mathrm{acc}(G; A_1, A_2) \geq e(\mathcal{P}) = 2\ell - |X \cap A_1| - 2|X \cap A_2|, \tag{4.7.14}$$

as required. □

The following proposition shows that, if A contains a collection X of vertices of high degree and G contains a basic connector $\mathcal{P}_0$ which does not interact too much with X, then we can extend $\mathcal{P}_0$ such that it still induces an Euler tour but $\mathrm{bal}_{AB}(\mathcal{P}_0)$ has increased.

Proposition 4.7.20. *Let $\Delta' \in \mathbb{N}$ and let $\ell, r \in \mathbb{N}_0$ be such that $\Delta' \geq 3\ell + 4$. Let G be a graph with vertex partition $\mathcal{V} = \{V_1, V_2, W := A \cup B\}$ and let $\mathcal{P}_0$ be a basic connector in G. For $i = 1, 2$, let A_i be the collection of all those vertices in A with degree i in $\mathcal{P}_0$. Let $X := \{x_1, \ldots, x_\ell\} \subseteq A$ where $d_A(x_i) \geq \Delta'$ for all $1 \leq i \leq \ell$. Suppose that $X \cap A_2 = \emptyset$ and $|X \setminus A_1| \geq r$. Then G contains a path system $\mathcal{P}$ such that $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour, $\mathrm{bal}_{AB}(\mathcal{P}) = \mathrm{bal}_{AB}(\mathcal{P}_0) + \ell + r$ and $e(\mathcal{P}) \leq \ell + r + 4$.*

Proof. Write $F_{\mathcal{P}_0}(A) := (a_1, a_2)$. So $|A_i| = a_i$ and hence $a_1 + a_2 = |V(\mathcal{P}_0) \cap A| \leq 4$ by

(BC2) and (BC3). Therefore we can apply Proposition 4.7.19 to see that

$$\mathrm{acc}(G[A]; A_1, A_2) \geq 2\ell - |X \cap A_1| - 2|X \cap A_2| \geq 2\ell - (\ell - r) = \ell + r.$$

Then Proposition 4.7.16 implies that there exists a path system $\mathcal{P}$ as required. □

The following lemma gives lower bounds for $\mathrm{acc}(G[A]; A_1, A_2)$. Together with Proposition 4.7.16, this will enable us to see 'how far' we can extend a basic connector. We show that $\mathrm{acc}(G[A]; A_1, A_2)$ is 'sufficiently large' unless we are in one of two special cases.

Lemma 4.7.21. *Let $k \in \{0, 1\}$, $\Delta, \Delta', \ell \in \mathbb{N}$ be such that $\ell + k \geq 2$. Suppose that $\Delta'/\Delta, \ell/\Delta' \ll 1$. Let G be a graph with vertex partition U, A and suppose that $e_G(A) \geq (\ell - 1)\Delta + \Delta'$ and $\Delta(G[A]), \Delta(G[A, U]) \leq \Delta$. Let $a_1, a_2 \in \mathbb{N}_0$ with $a_1 \geq k$ and $\Delta' \geq 3\ell + a_1 + a_2$. Let $A_1, A_2 \subseteq A$ be disjoint such that $|A_i| = a_i$ for $i = 1, 2$. Then one of the following holds.*

(I) $\mathrm{acc}(G[A]; A_1, A_2) \geq \ell - a_1 - 2a_2 + k + 2$*;*

(II) $k = 1$*,* $(a_1, a_2) = (1, 0)$ *and* $\mathrm{acc}(G[A]; A_1, A_2) \geq \ell + 1$*;*

(III) $k = 1$*,* $1 \leq \ell, a_1 + a_2 \leq 2$*,* $e_G(A) \leq \ell\Delta$ *and* $\mathrm{acc}(G[A]; A_1, A_2) \geq \ell - a_2$*. Moreover, let* $X := \{x \in A : d_A(x) \geq \Delta'\}$*. Then* $|X| = \ell$ *and all edges of* $G[A]$ *are incident with* X*.*

Proof. Apply Lemma 4.7.6 to $G[A]$. Suppose first that (i) holds. Let M be a matching in $G[A]$ of size $\ell + 1$ and let $uv \in E(G[A])$ be such that $u \notin V(M)$. Obtain M' from M by deleting all those edges with both endpoints in A_1 or at least one endpoint in A_2. Then M' accommodates A_1, A_2 by construction, so

$$\mathrm{acc}(G[A]; A_1, A_2) \geq e(M') \geq \ell + 1 - \lfloor a_1/2 \rfloor - a_2. \tag{4.7.15}$$

If $\lceil a_1/2 \rceil + a_2 \geq k + 1$, then (4.7.15) implies that (I) holds.

So suppose instead that $\lceil a_1/2 \rceil + a_2 \leq k$. First consider the case $k = 0$. Then $\lceil a_1/2 \rceil + a_2 = 0$ and hence $(a_1, a_2) = (0, 0)$. Now $A_1 = A_2 = \emptyset$, so $M \cup \{uv\}$ is a path system which accommodates A_1, A_2, and $e(M \cup \{uv\}) = \ell + 2$, so (I) holds.

Now consider the case $k = 1$. We have $\lceil a_1/2 \rceil + a_2 \leq 1$. But $a_1 \geq k \geq 1$ so $(a_1, a_2) = (1, 0)$. Observe that $\mathrm{acc}(G[A]; A_1, A_2) \geq \ell + 1$ by (4.7.15). So (II) holds.

Suppose now that Lemma 4.7.6(i) does not hold in $G[A]$. Since $\ell \geq 1$, we have $e_G(A) \leq \ell\Delta$ by the final assertion in Lemma 4.7.6. Let $X := \{x \in A : d_A(x) \geq \Delta'\}$. Then $|X| \geq \ell$. Since Lemma 4.7.6(i) does not hold, we must have that $|X| = \ell$ and that all edges of $G[A]$ are incident with X.

Apply Proposition 4.7.19 to see that

$$\begin{aligned} \mathrm{acc}(G[A]; A_1, A_2) &\geq 2\ell - |X \cap A_1| - 2|X \cap A_2| \geq 2\ell - \min\{a_1, \ell - a_2\} - 2a_2 \\ &= \ell - a_1 - 2a_2 + \max\{\ell, a_1 + a_2\} \geq \ell - a_2. \end{aligned} \tag{4.7.16}$$

In particular, if $\max\{\ell, a_1 + a_2\} \geq k + 2$, (4.7.16) implies that (I) holds. So we may suppose that $\max\{\ell, a_1 + a_2\} \leq k + 1$. Recall that $k + \ell \geq 2$ and $a_1 \geq k$ in the hypothesis. Hence, we have $k = 1$ and so $1 \leq \ell, a_1 + a_2 \leq 2$. So (III) holds. □

We are now ready to prove Lemma 4.7.3 in the case when $|A| - |B| \geq 2$ and $m \leq 2$. Roughly speaking, the approach is as follows. Proposition 4.7.15 implies that G contains a basic connector $\mathcal{P}_0$. When $m = 2$, Lemmas 4.7.17 and 4.7.18 allow us to assume that $\mathrm{bal}_{AB}(\mathcal{P}_0)$ is non-negative. We would like to extend $\mathcal{P}_0$ to a path system $\mathcal{P}$ in such a way that $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour and $\mathrm{bal}_{AB}(\mathcal{P}) = \ell + m/2 \geq |A| - |B|$. Proposition 4.7.16 implies that, in order to do this, it suffices to find a path system $\mathcal{P}_A$ in $G[A]$ which accommodates A_1, A_2 (where A_i is the collection of all those vertices in A with degree i in $\mathcal{P}_0$) and has enough edges. Now Lemma 4.7.21 implies that we can do this unless $m = 2$, ℓ is small and $(|A_1|, |A_2|)$ takes one of a small number of special values. Some additional arguments are required in these cases.

Proof of Lemma 4.7.3 *in the case when* $|A| - |B| \geq 2$ *and* $m \leq 2$. Let $k := m/2$. Since

$m \in 2\mathbb{N}_0$ we have $k \in \{0,1\}$. Let $\Delta := D/2$, $\Delta' := \nu n$ and $U := V_1 \cup V_2$. Proposition 4.7.10 implies that

$$\ell + k \geq |A| - |B| \geq 2. \tag{4.7.17}$$

Proposition 4.7.4 implies that $\ell, m \leq 12\rho n$. Then $\Delta'/\Delta, \ell/\Delta', m/\Delta' \ll 1$, $\Delta'/\Delta \ll \varepsilon$. Proposition 4.7.5 implies that

$$e_G(A) \geq (\ell - 1)\Delta + \Delta' \quad \text{and} \quad e_G(A, U) \geq (m-1)\Delta + \Delta'. \tag{4.7.18}$$

By Proposition 4.7.15, G contains a basic connector $\mathcal{P}_0$. Further assume that $\mathrm{bal}_{AB}(\mathcal{P}_0)$ is maximal, and given $\mathrm{bal}_{AB}(\mathcal{P}_0)$, a_1 is maximal where $F_{\mathcal{P}_0}(A) := (a_1, a_2)$. Let

$$t := |A| - |B| - \mathrm{bal}_{AB}(\mathcal{P}_0).$$

Then (BC2) implies that $t \geq 0$. In fact we may assume that $t \geq 1$ as otherwise $\mathcal{P}_0$ satisfies (P1)–(P3). For $i = 1, 2$ let A_i be the set of all those vertices in A which have degree i in $\mathcal{P}_0$. So $|A_i| = a_i$. Proposition 4.7.16 implies that, to prove Lemma 4.7.3, it suffices to show that

$$\mathrm{acc}(G[A]; A_1, A_2) \geq t.$$

(To check (P1), note that (BC2) and (4.7.17) imply $t \leq |A| - |B| + 2 \leq \ell + k + 2 \leq \ell + m + 2$.)

Claim A.

(i) *Suppose that* $k = 1$. *Then* $\mathrm{bal}_{AB}(\mathcal{P}_0) \geq 0$, *and if* $\mathrm{bal}_{AB}(\mathcal{P}_0) = 0$ *then* $a_1 \geq 1$.

(ii) $a_1 \geq k$.

Proof. To prove Claim A(i), note that if $k = 1$ (and so $m = 2$), then (4.7.18) and Lemma 4.7.6 imply that $G[A, U]$ contains a matching of size two. Together with Lemma 4.7.18 and our choice of $\mathcal{P}_0$ this in turn implies Claim A(i). Claim A(ii) clearly holds if $k = 0$, so assume $k = 1$. If $\mathrm{bal}_{AB}(\mathcal{P}_0) = 2$, then $a_1 \geq 1$ by (BC4). Together with Claim A(i) this shows that we may assume that $\mathrm{bal}_{AB}(\mathcal{P}_0) = 1$.

By (BC4), we may further assume that $(a_1, a_2) = (0, 1)$. Then (4.7.11) implies that $e_{\mathcal{P}_0}(B, U) = 0$. But then $\mathcal{P}_0$ has no endpoints in $W = A \cup B$, contradicting (BC1). ■

Apply Lemma 4.7.21 with $G \backslash B, A, U, F_{\mathcal{P}_0}(A), \ell, k$ playing the roles of $G, A, U, (a_1, a_2), \ell, k$. Suppose first that (I) holds, so

$$\mathrm{acc}(G[A]; A_1, A_2) \geq \ell - a_1 - 2a_2 + k + 2 \overset{\text{(BC4),(4.7.17)}}{\geq} |A| - |B| - \mathrm{bal}_{AB}(\mathcal{P}_0) = t,$$

as required. Therefore we may assume that one of Lemma 4.7.21(II) or (III) holds. So $k = 1$ and therefore $\mathrm{bal}_{AB}(\mathcal{P}_0) \geq 0$ by Claim A(i). Suppose first that (II) holds. Then

$$\mathrm{acc}(G[A]; A_1, A_2) \geq \ell + 1 \overset{(4.7.17)}{\geq} |A| - |B| \geq t,$$

as required. Therefore we may assume that (III) holds. So $1 \leq \ell, a_1 + a_2 \leq 2$, $e_G(A) \leq \ell\Delta$ and $\mathrm{acc}(G[A]; A_1, A_2) \geq \ell - a_2$. Let $X := \{x \in A : d_A(x) \geq \Delta'\}$. Then Lemma 4.7.21(III) also implies that $|X| = \ell$ and all edges of $G[A]$ are incident with X.

We claim that we are done if $\mathrm{bal}_{AB}(\mathcal{P}_0) \neq a_2$. To see this, suppose first that $\mathrm{bal}_{AB}(\mathcal{P}_0) \leq a_2 - 1$. Since $\mathrm{bal}_{AB}(\mathcal{P}_0) \geq 0$ this implies that $a_2 = 1$ and $\mathrm{bal}_{AB}(\mathcal{P}_0) = 0$. But $a_1 \geq k \geq 1$ by Claim A(ii) and $a_1 + a_2 \leq 2$, so $a_1 = a_2 = 1$. This is a contradiction to (BC4). Suppose instead that $\mathrm{bal}_{AB}(\mathcal{P}_0) \geq a_2 + 1$. Then

$$\mathrm{acc}(G[A]; A_1, A_2) \geq \ell - a_2 \geq \ell + 1 - \mathrm{bal}_{AB}(\mathcal{P}_0) = \ell + 1 - (|A| - |B|) + t \overset{(4.7.17)}{\geq} t.$$

Therefore we may assume that $\mathrm{bal}_{AB}(\mathcal{P}_0) = a_2$. In particular, this together with (BC4) implies that $\mathrm{bal}_{AB}(\mathcal{P}_0) \in \{0, 1\}$. We claim that we can further assume that

$$\ell = |A| - |B| - 1. \tag{4.7.19}$$

Indeed, to see this, note that by (4.7.17), it suffices to show that we are done if $\ell \geq$

$|A| - |B|$. But in this case we have $\mathrm{acc}(G[A]; A_1, A_2) \geq \ell - a_2 \geq |A| - |B| - a_2 = t$, as required.

We will now distinguish two cases.

Case 1. *$G[A, U]$ contains a matching of size three.*

Recall that $\mathrm{bal}_{AB}(\mathcal{P}_0) \in \{0, 1\}$. So Lemma 4.7.17 and our choice of $\mathcal{P}_0$ imply that $a_1 \geq 2$. Since $a_1 + a_2 \leq 2$ we have that $(a_1, a_2) = (2, 0)$. Therefore $\mathrm{bal}_{AB}(\mathcal{P}_0) = a_2 = 0$. Now, by Lemma 4.7.17 and our choice of $\mathcal{P}_0$ we deduce that there is some $i \in \{1, 2\}$ such that for $j \in \{1, 2\} \setminus \{i\}$ and for each $a \in A$, there are matchings $M_{a,A}, M_{a,B}$ in $G[A \setminus \{a\}, V_i], G[B, V_j]$ respectively, each of which has size two. Moreover, $\mathcal{P}_a := M_{a,A} \cup M_{a,B}$ is a basic connector with $\mathrm{bal}_{AB}(\mathcal{P}_a) = 0$.

Let $x \in X$ be arbitrary. (Recall that $|X| = \ell \geq 1$.) Apply Proposition 4.7.20 with $\mathcal{P}_x, V(M_{x,A}) \cap A, \emptyset, X, \ell, 1$ playing the roles of $\mathcal{P}_0, A_1, A_2, X, \ell, r$ to obtain a path system $\mathcal{P}$ in G such that $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour, $\mathrm{bal}_{AB}(\mathcal{P}) = \mathrm{bal}_{AB}(\mathcal{P}_x) + \ell + 1 = |A| - |B|$ (using (4.7.19)), and $e(\mathcal{P}) \leq \ell + 5$. Thus, $\mathcal{P}$ satisfies (P1)–(P3).

Case 2. *$G[A, U]$ does not contain a matching of size three.*

Together with König's theorem on edge-colourings this implies that $e_G(A, U) \leq 2\Delta$.

Claim B. *$X \cap V(\mathcal{P}_0) = \emptyset$.*

Proof. Since $e_G(A, U) \leq 2\Delta$, Proposition 3.7.4(ii) implies that

$$e_G(A) \geq \Delta(|A| - |B|) - e_G(A, U)/2 \overset{(4.7.19)}{\geq} \ell\Delta.$$

In fact, equality holds since $e_G(A) \leq \ell\Delta$ by Lemma 4.7.21(III). Since all edges of $G[A]$ are incident with X and $|X| = \ell$ it follows that $d_A(x) = \Delta = D/2$ for all $x \in X$. For all $x \in X$, $d_U(x) = D - d_A(x) - d_B(x) \leq D - 2d_A(x) = D - 2\Delta = 0$. The claim follows by (BC3). ■

Recall that we assume that $t \geq 1$. Observe that, since $\mathrm{bal}_{AB}(\mathcal{P}_0) \in \{0, 1\}$, the definition of t and (4.7.19) imply that $1 \leq t \leq |A| - |B| = \ell + 1$. Choose an arbitrary $X' \subseteq X$

with $|X'| = t-1$. Apply Proposition 4.7.20 with $\mathcal{P}_0, X', t-1, 1$ playing the roles of $\mathcal{P}_0, X, \ell, r$ to obtain a path system $\mathcal{P}$ in G such that $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour, $\mathrm{bal}_{AB}(\mathcal{P}) = \mathrm{bal}_{AB}(\mathcal{P}_0) + t = |A| - |B|$, and $e(\mathcal{P}) \leq \ell + 5$. Thus, $\mathcal{P}$ satisfies (P1)–(P3). $\square$

4.7.7 The proof of Lemma 4.7.3 in the case when $|A| = |B| + 1$.

Note that the extremal example in Figure 4.1(i) satisfies the conditions of this case. Therefore the degree bound $D \geq n/4$ is essential here. We will follow a similar strategy as in Section 4.7.6. We first find a basic connector $\mathcal{P}_0$ and then modify it to obtain a path system $\mathcal{P}$ satisfying (P1)–(P3). To be more precise, $\mathcal{P}$ will satisfy $e(\mathcal{P}) \leq 6$ and $\mathrm{bal}_{AB}(\mathcal{P}) = 1$. Throughout this section, we will assume that the basic connector $\mathcal{P}_0$ is chosen so that $|\mathrm{bal}_{AB}(\mathcal{P}_0) - 1|$ is minimal. We will distinguish cases depending on the value of $\mathrm{bal}_{AB}(\mathcal{P}_0)$.

Let G be a D-regular graph with vertex partition A, B, U where $|A| = |B| + 1$. Then Proposition 3.7.4(i) implies that

$$2e_G(A) + e_G(A, U) = 2e_G(B) + e_G(B, U) + D. \tag{4.7.20}$$

We will need the following simple facts for the case when $|\mathrm{bal}_{AB}(\mathcal{P}_0)| = 2$.

Proposition 4.7.22. *Let G be a 3-connected graph with vertex partition $\mathcal{V} = \{V_1, V_2, W := A \cup B\}$. Then the following holds:*

(i) if $\mathcal{P}_0$ is a basic connector in G with $\mathrm{bal}_{AB}(\mathcal{P}_0) = 2$, then $V(\mathcal{P}_0) \cap B = \emptyset$ and $\mathcal{P}_0[A, V_i]$ is a matching of size two for each $i = 1, 2$. In particular, $\mathcal{P}_0[A, V_1 \cup V_2]$ contains a matching of size three.

(ii) if $e_G(B, U) > 0$ and G contains a basic connector $\mathcal{P}_0'$ with $\mathrm{bal}_{AB}(\mathcal{P}_0') = 2$, then G also contains a basic connector $\mathcal{P}_0$ with $\mathrm{bal}_{AB}(\mathcal{P}_0) = 1$;

(iii) if $e_G(A, U) > 0$ then G contains a basic connector $\mathcal{P}_0$ with $\mathrm{bal}_{AB}(\mathcal{P}_0) \geq -1$;

(iv) if $e_G(A,U), e_G(B,U) > 0$ then G contains a basic connector $\mathcal{P}_0$ with $|\mathrm{bal}_{AB}(\mathcal{P}_0)| \leq 1$.

Proof. (i) follows immediately from (BC1)–(BC4). To prove (ii), note that by (i), for both $i = 1, 2$ there are matchings $M_i \subseteq G[A, V_i]$ of size two such that $\mathcal{P}_0' = M_1 \cup M_2$. Let $e \in E(G[B, U])$ be arbitrary. Without loss of generality, suppose that $e \in E(G[B, V_1])$. If possible, let $e' \in E(M_1)$ be the edge incident with e; otherwise let $e' \in E(M_1)$ be arbitrary. Then $\mathcal{P}_0 := (\mathcal{P}_0' \cup \{e\}) \setminus \{e'\}$ is a basic connector with $\mathrm{bal}_{AB}(\mathcal{P}_0) = 1$, as required. (iii) and (iv) follow from Proposition 4.7.15 together with an argument similar to the one for (ii). □

The next lemma concerns the case when $G[A, V_1 \cup V_2]$ contains a matching of size three. This extra condition ensures the existence of a basic connector with useful properties of which we can take advantage.

Lemma 4.7.23. *Let $n, D \in \mathbb{N}$ be such that $D \geq n/4$ and $1/n \ll 1$. Let G be a 3-connected D-regular graph with vertex partition $\mathcal{V} = \{V_1, V_2, W := A \cup B\}$, where $|V_i| \geq D/2$ for $i = 1, 2$. Suppose that $|A| = |B| + 1$, that $\Delta(G[A, V_1 \cup V_2]) \leq D/2$ and that $G[A, V_1 \cup V_2]$ contains a matching of size three. Then G contains a path system $\mathcal{P}$ which satisfies* (P1)–(P3).

Proof. Let $U := V_1 \cup V_2$. Without loss of generality we may assume that $e_G(A, V_1) \leq e_G(A, V_2)$. We will obtain $\mathcal{P}$ by adding at most two edges to a basic connector $\mathcal{P}_0$. Therefore $e(\mathcal{P}) \leq 6$ so (P1) will hold. We may assume that there does not exist a basic connector $\mathcal{P}_0'$ with $\mathrm{bal}_{AB}(\mathcal{P}_0') = 1$ (otherwise we can take $\mathcal{P} := \mathcal{P}_0'$). Apply Lemma 4.7.17 to obtain a basic connector in G which satisfies (i) or (ii).

Case 1. *Lemma 4.7.17(i) holds.*

So G contains a basic connector $\mathcal{P}_0$ such that $\mathrm{bal}_{AB}(\mathcal{P}_0) \geq 1$ and, if $F_{\mathcal{P}_0}(A) = (a_1, a_2)$, then $a_1 \geq 2$. Thus $\mathrm{bal}_{AB}(\mathcal{P}_0) = 2$ by our assumption. Proposition 4.7.22(i) implies that $V(\mathcal{P}_0) \cap B = \emptyset$. Furthermore, Proposition 4.7.22(ii) implies that $e_G(B, U) = 0$.

Suppose that $e_G(B) \geq 1$. For arbitrary $e \in E(G[B])$ we have that $\mathcal{P} := \mathcal{P}_0 \cup \{e\}$ satisfies (P1)–(P3). So we may assume that $e_G(B) = 0$. So (4.7.20) implies that

$$2e_G(A) + e_G(A, U) = D. \tag{4.7.21}$$

Moreover, for each $b \in B$ we have that $N_G(b) \subseteq A$ and thus $|A| \geq D$. So $|B| \geq D - 1$ and since $D \geq n/4$ we have that $|U| \leq 2D + 1$. We will only prove the case when $|V_1| = D - s$ for some $s \in \mathbb{N}_0$. (The same argument also works for $|V_2| = D - s$.) Recall that $s \leq D/2$ by assumption. Then every vertex in V_1 has at least $s + 1$ neighbours in $\overline{V_1}$. Since $e_G(B, U) = 0$ and $e_G(A, V_1) \leq e_G(A, V_2)$ we have that

$$e_G(V_1, V_2) \geq e_G(V_1, \overline{V_1}) - e_G(A, V_1) \overset{(4.7.21)}{\geq} (s + 1)(D - s) - D/2 \geq D/2.$$

Suppose that $\mathcal{P}_0$ is a matching of size four in $G[A, U]$. Then, given any $e \in E(G[V_1, V_2])$, we can choose $e_i \in \mathcal{P}_0[A, V_i]$ such that e, e_1, e_2 is a matching of size three. Otherwise, Proposition 4.7.22(i) implies that $\mathcal{P}_0$ consists of vertex-disjoint paths u_1a_1, u_2a_2, v_1av_2, where $v_i, u_i \in V_i$ and $a, a_1, a_2 \in A$. Since $e_G(V_1, V_2) \geq 2$, we can pick $e \in E(G[V_1, V_2]) \setminus \{u_1u_2\}$. It is easy to see that we can similarly find $e_i \in E(\mathcal{P}_0[A, V_i])$ such that e, e_1, e_2 is a matching of size three. In both cases, $\mathcal{P} := \{e, e_1, e_2\}$ satisfies (P1)–(P3).

Case 2. *Lemma 4.7.17(ii) holds.*

Since $e_G(A, V_1) \leq e_G(A, V_2)$ this implies that $e_G(V_1, A) = 0$. Moreover, Lemma 4.7.17(ii) also implies that, for each $a \in A$, there are matchings $M_{a,A}, M_{a,B}$ in $G[A \setminus \{a\}, V_2], G[B, V_1]$ respectively, each of which has size two. In particular $e_G(B, U) \geq 2$. Suppose that $e_G(A) > 0$. Let $aa' \in E(G[A])$. Then $\mathcal{P} := M_{a,A} \cup M_{a,B} \cup \{aa'\}$ satisfies (P1)–(P3). So we may assume that $e_G(A) = 0$. Then (4.7.20) implies that $e_G(A, V_2) = e_G(A, U) \geq D + e_G(B, U) \geq D + 2$. The 'moreover' part of Lemma 4.7.6 with $G[A, V_2], D/2, 2$ playing the roles of G, Δ, ℓ implies that $G[A, V_2]$ contains a matching M_A of size three and an edge xy with $x \notin V(M_A)$. Let $a \in A$ be arbitrary. Then $\mathcal{P} := M_{a,B} \cup M_A \cup \{xy\}$ satisfies

(P1)–(P3). $\square$

The following proposition will be used to find edges in $G[A]$ which can be added to a basic connector $\mathcal{P}_0$ so that it is still a path system and $R_{\mathcal{V}}(\mathcal{P}_0)$ is still an Euler tour. For example, if $a \in A$ is such that $d_{\mathcal{P}_0}(a) = 2$, then we cannot add any edges in $G[A]$ which are incident with a. (Recall that the partition given in Lemma 4.7.3 satisfies $d_A(a) \leq d_B(a)$ for all $a \in A$.)

Proposition 4.7.24. *Let G be a D-regular graph with vertex partition A, B, U where $|A| = |B| + 1$. Let $a \in A$ be such that $d_A(a) \leq d_B(a)$. Then*

$$2e_G(A \setminus \{a\}) + e_G(A \setminus \{a\}, U) \geq e_G(B, U).$$

Proof. Note that

$$\begin{aligned} 2e_G(A \setminus \{a\}) + e_G(A \setminus \{a\}, U) &= 2e_G(A) + e_G(A, U) - 2d_A(a) - d_U(a) \\ &\geq 2e_G(A) + e_G(A, U) - d_A(a) - d_B(a) - d_U(a) \\ &= 2e_G(A) + e_G(A, U) - D \overset{(4.7.20)}{\geq} e_G(B, U), \end{aligned}$$

as required. $\square$

By Lemma 4.7.23, we may assume that $G[A, V_1 \cup V_2]$ contains no matching of size three. Then Proposition 4.7.22(i) allows us to assume that $\mathrm{bal}_{AB}(\mathcal{P}_0) \leq 0$ (or we are done). In the next lemma, we consider the case when $\mathrm{bal}_{AB}(\mathcal{P}_0) = 0$.

Lemma 4.7.25. *Let $D \in \mathbb{N}$. Let G be a 3-connected D-regular graph with vertex partition $\mathcal{V} = \{V_1, V_2, W := A \cup B\}$. Suppose that $|A| = |B| + 1$, $\Delta(G[A, V_1 \cup V_2]) \leq D/2$ and $d_A(a) \leq d_B(a)$ for all $a \in A$. Suppose further that $G[A, V_1 \cup V_2]$ does not contain a matching of size three. Let $\mathcal{P}_0$ be a basic connector in G with $\mathrm{bal}_{AB}(\mathcal{P}_0) = 0$. Then G contains a path system $\mathcal{P}$ which satisfies* (P1)–(P3).

Proof. Let $U := V_1 \cup V_2$. Since $G[A, U]$ does not contain a matching of size three, König's theorem on edge-colourings implies that

$$e_G(A, U) \leq D. \tag{4.7.22}$$

Property (BC4) implies that $a_1 + 2a_2 \in \{1, 2\}$ and so $F_{\mathcal{P}_0}(A) \in \{(2, 0), (1, 0), (0, 1)\}$. We will distinguish cases based on the value of $F_{\mathcal{P}_0}(A)$.

Case 1. $F_{\mathcal{P}_0}(A) = (2, 0)$.

Then (4.7.11) implies that $e_{\mathcal{P}_0}(A, U) = e_{\mathcal{P}_0}(B, U) = 2$. Since $\mathcal{P}_0$ is an Euler tour and $e(\mathcal{P}_0) \leq 4$ by (BC1) and (BC2), there are distinct vertices $a, a' \in A$, a collection of distinct vertices $X := \{u, u', v, v'\} \subseteq U$ with $|X \cap V_i| = 2$ for $i = 1, 2$ and $b, b' \in B$ which are not necessarily distinct, such that $\mathcal{P}_0 := \{au, a'u', bv, b'v'\}$.

Observe that we are done if there exists $e \in E(G[A]) \setminus \{aa'\}$ since then $\mathcal{P}_0 \cup \{e\}$ satisfies (P1)–(P3). So we may assume that $E(G[A]) \subseteq \{aa'\}$. Now

$$2 = e_{\mathcal{P}_0}(B, U) \leq e_G(B, U) \overset{(4.7.22)}{\leq} 2e_G(B) + e_G(B, U) + D - e_G(A, U) \overset{(4.7.20)}{=} 2e_G(A) \leq 2.$$

Therefore we have $e_G(B) = 0$, $e_G(A) = 1$, $e_G(A, U) = D$ and $e_G(B, U) = 2$, so $E(G[B, U]) = \{bv, b'v'\}$ and $E(G[A]) = \{aa'\}$.

We will assume that either $\{u, u'\} \subseteq V_1$ and $\{v, v'\} \subseteq V_2$; or $\{u, v\} \subseteq V_1$ and $\{u', v'\} \subseteq V_2$ since the other cases are similar.

Case 1.a. $\{u, u'\} \subseteq V_1$ *and* $\{v, v'\} \subseteq V_2$.

Suppose that $e_G(V_1, V_2) \neq 0$. Let $v_1v_2 \in E(G[V_1, V_2])$ with $v_i \in V_i$. Choose $e_1 \in \mathcal{P}_0[A, V_1 \setminus \{v_1\}]$ and $e_2 \in \mathcal{P}_0[B, V_2 \setminus \{v_2\}]$. Then $\mathcal{P} := \{e_1, e_2, v_1v_2, aa'\}$ satisfies (P1)–(P3). Suppose that $e_G(A, V_2) \neq 0$. Let $a''x_2 \in E(G[A, V_2])$ with $a'' \in A$ and $x_2 \in V_2$. Choose $e_2 \in \mathcal{P}_0[B, V_2 \setminus \{x_2\}]$. Then $\mathcal{P} := \{au, a'u', a''x_2, e_2\}$ satisfies (P1)–(P3). Therefore $e_G(A \cup V_1, V_2) = 0$. So $E(G[V_2, \overline{V_2}]) = \{bv, b'v'\}$, contradicting the 3-connectivity of G.

Case 1.b. $\{u, v\} \subseteq V_1$ *and* $\{u', v'\} \subseteq V_2$.

We may assume that $b = b'$ since otherwise $\mathcal{P} := \mathcal{P}_0 \cup \{aa'\}$ satisfies (P1)–(P3). Since $G[A, U]$ does not contain a matching of size three, every edge in $G[A, U]$ is incident with at least one of a, a', u, u'. Suppose that there exists $a'' \in A \setminus \{a, a'\}$ such that $ua'' \in E(G)$. Then $\mathcal{P} := \mathcal{P}_0 \cup \{ua'', aa'\} \setminus \{ua\}$ satisfies (P1)–(P3). A similar deduction can be made with u' playing the role of u. Therefore every edge in $G[A, U]$ is incident with a or a'. Since $e_G(A, U) = D$ we have $d_U(a), d_U(a') = D/2$.

Suppose that $e_G(V_1, V_2) \neq 0$. Let $v_1v_2 \in E(G[V_1, V_2])$ with $v_i \in V_i$. If $v_1 \neq u$ and $v_2 \neq u'$ then $\mathcal{P} := \{au, a'u', v_1v_2\}$ satisfies (P1)–(P3). Therefore we may suppose, without loss of generality, that $v_1 = u$. Suppose that $v_2 \neq u'$. Then $\mathcal{P} := \{a'u', v_1v_2, bv, aa'\}$ satisfies (P1)–(P3). Therefore we may suppose that $v_2 = u$. Thus $uu' \in E(G)$. Since $d_U(a) \geq D/2$, we can choose $w \in N_U(a) \setminus \{v, v', u, u'\}$. Suppose that $w \in V_1$. Then $\mathcal{P} := \{aw, uu', aa', bv'\}$ satisfies (P1)–(P3). If $w \in V_2$ then $\mathcal{P} := \{aw, uu', aa', bv\}$ satisfies (P1)–(P3).

Thus we may assume that $e_G(V_1, V_2) = 0$. Choose $Y_a \in \{V_1, V_2\}$ such that $d_{Y_a}(a) \geq D/4$. Note that there is always such a Y_a. Define $Y_{a'}$ analogously. Suppose that $Y_{a'} = V_1$. Choose $w' \in N_{V_1}(a') \setminus \{u, v\}$. Then $\mathcal{P} := \mathcal{P}_0 \cup \{a'w'\} \setminus \{bv\}$ satisfies (P1)–(P3). We can argue similarly if $Y_a = V_2$.

Therefore we may assume that $Y_{a'} = V_2$ and $Y_a = V_1$. Suppose that $d_{V_1}(a') \neq 0$. Let $w' \in N_{V_1}(a')$. Since $d_{V_1}(a) \geq D/4$, we can choose $w \in N_{V_1}(a) \setminus \{w'\}$. Then $\mathcal{P} := \mathcal{P}_0 \cup \{aw, a'w'\} \setminus \{au, bv\}$ satisfies (P1)–(P3). So $d_{V_1}(a') = 0$. Since every edge of $G[A, U]$ is incident with a or a', we have that every edge in $G[A, V_1]$ is incident with a. We have shown that every edge in $G[V_1, \overline{V_1}]$ is incident with a or b, contradicting the 3-connectivity of G.

Case 2. $F_{\mathcal{P}_0}(A) = (1, 0)$.

Then (4.7.11) implies that $e_G(B, U) \geq e_{\mathcal{P}_0}(B, U) = 1$. So (4.7.20) and (4.7.22) give $2e_G(A) = D + 2e_G(B) + e_G(B, U) - e_G(A, U) \geq 1$. Let $e \in E(G[A])$ be arbitrary. Then $\mathcal{P} := \mathcal{P}_0 \cup \{e\}$ satisfies (P1)–(P3).

Case 3. $F_{\mathcal{P}_0}(A) = (0, 1)$.

Now (4.7.11) implies that $e_{\mathcal{P}_0}(B,U) = e_{\mathcal{P}_0}(A,U) = 2$. So (BC2) implies that $e_{\mathcal{P}_0}(V_1,V_2) = 0$ and that there exist distinct $v_i, u_i \in V_i$ for $i = 1,2$, and $b, b' \in B$ and $a \in A$ such that $\mathcal{P}_0 = \{v_1b, v_2b', u_1au_2\}$. Proposition 4.7.24 implies that $2e_G(A\setminus\{a\})+e_G(A\setminus\{a\},U) \geq 2$. Suppose first that $e_G(A \setminus \{a\}) \geq 1$. Choose $e \in E(G[A \setminus \{a\}])$. Then $\mathcal{P} := \mathcal{P}_0 \cup \{e\}$ satisfies (P1)–(P3). Therefore we may assume that $e_G(A \setminus \{a\}, U) \geq 2$. Suppose there exists $e' \in E(G[A \setminus \{a\}, U \setminus \{u_1, u_2\}])$. Without loss of generality, suppose that e' has an endpoint in V_1. Then $\mathcal{P} := \mathcal{P}_0 \cup \{e'\} \setminus \{v_1b\}$ satisfies (P1)–(P3). Therefore we may assume that G contains an edge $a'u_1$ where $a' \in A \setminus \{a\}$. Let $\mathcal{P}_0' := \mathcal{P}_0 \cup \{a'u_1\} \setminus \{au_1\}$. Then $\mathcal{P}_0'$ is a basic connector with $\mathrm{bal}_{AB}(\mathcal{P}_0') = 0$ and $F_{\mathcal{P}_0'}(A) = (2,0)$. So we are in Case 1. □

The next lemma concerns the case when $\mathrm{bal}_{AB}(\mathcal{P}_0) = -1$.

Lemma 4.7.26. *Let $D \in \mathbb{N}$ where $D \geq 12$. Let G be a 3-connected D-regular graph with vertex partition $\mathcal{V} = \{V_1, V_2, W := A \cup B\}$. Suppose that $|A| = |B|+1$, $\Delta(G[A, V_1 \cup V_2] \leq D/2$ and $d_A(a) \leq d_B(a)$ for all $a \in A$. Let $\mathcal{P}_0$ be a basic connector in G such that $|\mathrm{bal}_{AB}(\mathcal{P}_0) - 1|$ is minimal. Suppose that $\mathrm{bal}_{AB}(\mathcal{P}_0) = -1$. Then G contains a path system $\mathcal{P}$ which satisfies* (P1)–(P3).

Proof. Let $U := V_1 \cup V_2$. Observe that $G[A,U]$ does not contain a matching of size two since otherwise Lemma 4.7.18 would imply that $\mathrm{bal}_{AB}(\mathcal{P}_0) \geq 0$. Therefore $e_G(A,U) \leq D/2$, and so (4.7.20) implies that

$$e_G(A) \geq D/4. \tag{4.7.23}$$

Write $F_{\mathcal{P}_0}(A) := (a_1, a_2)$. Then (BC4) implies that $a_1 + 2a_2 \in \{0,1\}$. So $(a_1,a_2) \in \{(0,0),(1,0)\}$. Suppose first that $(a_1,a_2) = (0,0)$. Then by (4.7.23), we can choose distinct $e, e' \in E(G[A])$. In this case $\mathcal{P} := \mathcal{P}_0 \cup \{e, e'\}$ satisfies (P1)–(P3).

Now suppose that $(a_1,a_2) = (1,0)$. Then (4.7.11) implies that

$$e_G(B,U) \geq e_{\mathcal{P}_0}(B,U) = 3. \tag{4.7.24}$$

Let au be the single edge in $\mathcal{P}_0[A,U]$, where $a \in A$ and $u \in U$. Note that any edge in $E(G[A \setminus \{a\}, U])$ is incident with u since $G[A,U]$ contains no matching of size two. So $e_G(A \setminus \{a\}, U) = d_{A\setminus\{a\}}(u)$. Thus Proposition 4.7.24 and (4.7.24) imply that

$$2e_G(A \setminus \{a\}) + d_{A\setminus\{a\}}(u) \geq 3. \tag{4.7.25}$$

Suppose first that $d_A(a) \leq 1$. In this case, (4.7.23) implies that $e_G(A\setminus\{a\}) \geq D/4-1 \geq 2$. Let $e, e' \in E(G[A \setminus \{a\}])$ be distinct. Then $\mathcal{P} := \mathcal{P}_0 \cup \{e, e'\}$ satisfies (P1)–(P3).

Now suppose that $d_A(a) \geq 2$. Let $a', a'' \in N_A(a)$ be distinct. Suppose that $e_G(A \setminus \{a\}) \neq 0$. Then we can choose $e \in E(G[A \setminus \{a\}])$, and $\mathcal{P} := \mathcal{P}_0 \cup \{aa', e\}$ satisfies (P1)–(P3). Suppose instead that $e_G(A \setminus \{a\}) = 0$. Then $d_{A\setminus\{a\}}(u) \geq 3$ by (4.7.25), so there exists $a^* \in A \setminus \{a, a', a''\}$ such that $ua^* \in E(G[A,U])$. We have that $\mathcal{P} := \mathcal{P}_0 \cup \{ua^*, a'aa''\} \setminus \{ua\}$ satisfies (P1)–(P3). □

We are now ready to combine the preceding lemmas to prove Lemma 4.7.3 fully in the case when $|A| = |B| + 1$.

Proof of Lemma 4.7.3 *in the case when* $|A| = |B| + 1$. Let $U := V_1 \cup V_2$. Suppose first that $G[A,U]$ contains a matching of size three. Then we are done by Lemma 4.7.23, so assume not. Proposition 4.7.15 implies that G contains a basic connector. Choose a basic connector $\mathcal{P}_0$ in G such that $|\mathrm{bal}_{AB}(\mathcal{P}_0) - 1|$ is minimal. Recall that (BC2) implies $|\mathrm{bal}_{AB}(\mathcal{P}_0)| \leq 2$. Since $G[A,U]$ does not contain a matching of size three, Proposition 4.7.22(i) implies that $\mathrm{bal}_{AB}(\mathcal{P}_0) \leq 1$. We may assume that $\mathrm{bal}_{AB}(\mathcal{P}_0) \leq 0$ or we are done. Lemmas 4.7.25 and 4.7.26 prove the lemma in the case when $\mathrm{bal}_{AB}(\mathcal{P}_0) = 0, -1$ respectively. So we may assume that $\mathrm{bal}_{AB}(\mathcal{P}_0) = -2$. Thus, by (4.7.11), we have $e_G(B,U) \geq 4$. Moreover, by Proposition 4.7.22(iii) we may assume that $e_G(A,U) = 0$. Now (4.7.20) implies $e_G(A) \geq D/2 + 2$. The 'moreover' part of Lemma 4.7.6 with $G[A], D/2, 1$ playing the roles of G, Δ, ℓ implies that $G[A]$ contains a matching M_A of size two and an edge aa' with $a \notin V(M_A)$. So $\mathcal{P} := \mathcal{P}_0 \cup M_A \cup \{aa'\}$ satisfies (P1)–(P3). □

4.7.8 The proof of Lemma 4.7.3 in the case when $|A| = |B|$

In this subsection we consider the only remaining case of Lemma 4.7.3: when the bipartite vertex classes A and B have equal size. Our aim is to find a path system $\mathcal{P}$ such that $R_{\mathcal{V}}(\mathcal{P})$ is an Euler tour, and $\mathrm{bal}_{AB}(\mathcal{P}) = 0$. As in the previous section, we will appropriately modify a basic connector guaranteed by Proposition 4.7.15. The degree bound $D \geq n/4$ is used again here.

Proof of Lemma 4.7.3 *in the case when* $|A| = |B|$. Let $U := V_1 \cup V_2$. Proposition 3.7.4(i) implies that

$$2e_G(A) + e_G(A, U) = 2e_G(B) + e_G(B, U). \tag{4.7.26}$$

Proposition 4.7.15 implies that G contains a basic connector. Choose a basic connector $\mathcal{P}_0$ in G such that $|\mathrm{bal}_{AB}(\mathcal{P}_0)|$ is minimal. Write $F_{\mathcal{P}_0}(A) := (a_1, a_2)$.

Suppose first that $e_G(B, U) = 0$. Then

$$2\mathrm{bal}_{AB}(\mathcal{P}_0) \overset{(4.7.11)}{=} a_1 + 2a_2 = e_{\mathcal{P}_0}(A, U) \leq e_G(A, U) \overset{(4.7.26)}{\leq} 2e_G(B).$$

(In particular, $\mathrm{bal}_{AB}(\mathcal{P}_0) \geq 0$.) Let $E' \subseteq E(G[B])$ be a collection of $\mathrm{bal}_{AB}(\mathcal{P}_0)$ distinct edges (so $|E'| \leq 2$ by (BC2)). Then $\mathcal{P} := \mathcal{P}_0 \cup E'$ satisfies (P1)–(P3). Thus we may assume that $e_G(B, U) \geq 1$ and a similar argument allows us to assume that $e_G(A, U) \geq 1$.

Together with the 3-connectivity of G, this implies that $G[W, U]$ contains a matching M of size two such that one edge is incident with A and one edge is incident with B. Proposition 4.7.22(iv) and our choice of $\mathcal{P}_0$ together imply that $|\mathrm{bal}_{AB}(\mathcal{P}_0)| \leq 1$. Without loss of generality we suppose that $\mathrm{bal}_{AB}(\mathcal{P}_0) = -1$ (otherwise $\mathrm{bal}_{AB}(\mathcal{P}_0) = 1$ and we could swap A and B, or $\mathrm{bal}_{AB}(\mathcal{P}_0) = 0$ and we are done by taking $\mathcal{P} := \mathcal{P}_0$). Then (BC4) implies that $(a_1, a_2) \in \{(0, 0), (1, 0)\}$. If $e_G(A) \geq 1$ then, for any $e \in E(G[A])$ we have that $\mathcal{P} := \mathcal{P}_0 \cup \{e\}$ satisfies (P1)–(P3). So we may assume that

$$e_G(A) = 0. \tag{4.7.27}$$

Claim 1. *$G[A,U]$ does not contain a matching of size two.*

Proof. Suppose not. We will show that if $G[A,U]$ contains a matching of size two, then the minimality of $|\mathrm{bal}_{AB}(\mathcal{P}_0)|$ will be contradicted. First consider the case when $(a_1,a_2)=(1,0)$. So $e_{\mathcal{P}_0}(A,U)=1$ and therefore $e_{\mathcal{P}_0}(B,U)=3$ by (4.7.11). But (BC2) implies that $e(\mathcal{P}_0)\leq 4$, so $e_{\mathcal{P}_0}(V_1,V_2)=0$. Now by (BC1) we have that $|V(\mathcal{P}_0)\cap V_i|=2$ for $i=1,2$, and $d_{\mathcal{P}_0}(v)=1$ for all $v\in V(\mathcal{P}_0)\cap V_i$. In particular, $e_{\mathcal{P}_0}(V_i,B)>0$ for both $i=1,2$. Let e be the single edge in $\mathcal{P}_0[A,U]$. Without loss of generality, we may assume that $G[A,U]$ contains an edge e' which is vertex-disjoint from e. (Otherwise, $G[A,U]$ contains a matching $av,a'v'$ such that $e=av'$. Then $\mathcal{P}_0':=\mathcal{P}_0\cup\{a'v'\}\setminus\{e\}$ is a basic connector with $\mathrm{bal}_{AB}(\mathcal{P}_0')=\mathrm{bal}_{AB}(\mathcal{P}_0)$ and $a'v'$ is the single edge in $\mathcal{P}_0'[A,U]$; and av is an edge which is vertex-disjoint from $a'v'$.) Suppose first that e' has an endpoint in V_1. If possible, choose $f\in E(\mathcal{P}_0[V_1,B])$ which is incident with e'; otherwise let $f\in E(\mathcal{P}_0[V_1,B])$ be arbitrary. Then $\mathcal{P}:=\mathcal{P}_0\cup\{e'\}\setminus\{f\}$ contradicts the minimality of $|\mathrm{bal}_{AB}(\mathcal{P}_0)|$. The case when e' has an endpoint in V_2 is similar.

Suppose now that $(a_1,a_2)=(0,0)$. Then $e_{\mathcal{P}_0}(A,U)=0$ and hence $e_{\mathcal{P}_0}(B,U)=2$. Moreover, $\mathcal{P}_0[B,U]$ is a matching e,e' since $\mathcal{P}_0$ is an Euler tour by (BC1). Now $d_{R_{\mathcal{V}}(\mathcal{P}_0)}(V_i)\geq 2$ for $i=1,2$, so $e_{\mathcal{P}_0}(V_1,V_2)\geq 1$. But (BC2) implies that $e(\mathcal{P}_0)\leq 4$, so $e_{\mathcal{P}_0}(V_1,V_2)\leq 2$. Suppose that $e_{\mathcal{P}_0}(V_1,V_2)=1$ and let $f\in E(\mathcal{P}_0[V_1,V_2])$. Then $\mathcal{P}_0=\{e,e',f\}$ is a matching of size three. Moreover $e_{\mathcal{P}_0}(B,V_i)=1$ for $i=1,2$. If there exists $e_A\in E(G[A,U]\setminus V(f))$ then we can replace one of e,e' by e_A to contradict the minimality of $|\mathrm{bal}_{AB}(\mathcal{P}_0)|$. Therefore there is a matching $\{e_A,e_A'\}\subseteq E(G[A,U])$ such that both e_A,e_A' are incident to $V(f)$. Then they are vertex-disjoint from $\{e,e'\}$, so $\mathcal{P}:=\{e,e',e_A,e_A'\}$ contradicts the minimality of $|\mathrm{bal}_{AB}(\mathcal{P}_0)|$. Suppose now that $e_{\mathcal{P}_0}(V_1,V_2)=2$. Then $\mathcal{P}_0[B,U]\subseteq G[B,V_i]$ for some $i=1,2$. Without loss of generality we assume that $i=2$. Suppose that there exists $e_A\in E(G[A,V_1])$. Choose $f\in E(\mathcal{P}_0[V_1,V_2])$ that is not incident to e_A. Choose $e_B\in E(\mathcal{P}_0[B,V_2])$ that is not incident to f. Then $\mathcal{P}:=\{e_A,f,e_B\}$

contradicts the minimality of $|\mathrm{bal}_{AB}(\mathcal{P}_0)|$. Therefore we may assume that there is a matching $M_A \subseteq G[A, V_2]$ of size two. There is at least one V_1V_2-path in $\mathcal{P}_0$ (which consists of a single edge f'). Choose $e \in M_A$ which is not incident to f'. If possible, let e_B be the edge of $\mathcal{P}_0[B, V_2]$ which is incident to e; otherwise let $e_B \in E(\mathcal{P}_0[B, V_2])$ be arbitrary. Then $\mathcal{P} := \mathcal{P}_0 \cup \{e\} \setminus \{e_B\}$ contradicts the minimality of $|\mathrm{bal}_{AB}(\mathcal{P}_0)|$. This completes the proof of the claim. ■

Therefore $e_G(A, U) \leq D/2$ since $\Delta(G[A, U)] \leq D/2$. So (4.7.26) and (4.7.27) together imply that

$$e_G(W, U) = e_G(B, U) - e_G(A, U) + 2e_G(A, U) \leq D. \tag{4.7.28}$$

Suppose first that $|A| = |B| = D - k$ for some $k \in \mathbb{N}$. Then (4.7.27) implies that, for all $a \in A$, we have $d_U(a) = D - d_A(a) - d_B(a) \geq D - |B| = k$. So $e_G(A, U) \geq k|A| = k(D - k) \geq D - 1$, a contradiction. So $|A| = |B| \geq D$ and hence $|U| = n - |A| - |B| \leq n - 2D \leq 2D$ since $D \geq n/4$.

Claim 2. *There exists a matching M' of size three in $G[V_1, V_2]$.*

Proof. To prove the claim, assume without loss of generality that $|V_1| \leq |V_2|$. Then there exists $s \in \mathbb{N}_0$ such that $|V_1| = D - s$. Recall from our assumption in Lemma 4.7.3 that $|V_1| \geq D/2$. Suppose first that $s \geq 2$. Then

$$\begin{aligned} e_G(V_1, V_2) &\geq D|V_1| - e_G(U, W) - 2\binom{|V_1|}{2} \overset{(4.7.28)}{\geq} |V_1|(D - |V_1| + 1) - D \\ &\geq \min\{D^2/4 - D/2, 2D - 6\} \geq D + 1. \end{aligned} \tag{4.7.29}$$

Recall that $d_{V_i}(x_i) \geq d_{V_j}(x_i)$ for all $x_i \in V_i$ and $\{i, j\} = \{1, 2\}$. So $\Delta(G[V_1, V_2]) \leq D/2$. Therefore we are done by König's theorem on edge-colourings.

Thus we may assume that $s \in \{0, 1\}$. Let $H := G[V_1, V_2]$. Suppose that H contains no matching of size three. By König's theorem on vertex covers, H contains a vertex cover $\{v_i, v_j\}$ where $v_i \in V_i$, $v_j \in V_j$ and i, j are not necessarily

distinct. So $e(H) \leq d_H(v_i) + d_H(v_j)$. Note that the complement $\overline{G}$ of G satisfies

$$\begin{aligned} e_{\overline{G}}(V_1) + e_{\overline{G}}(V_2) &\geq d_{\overline{G}[V_i]}(v_i) + d_{\overline{G}[V_j]}(v_j) - 1 = |V_i| - d_{V_i}(v_i) + |V_j| - d_{V_j}(v_j) - 3 \\ &\geq D - d_{V_i}(v_i) + D - d_{V_j}(v_j) - 5 \geq d_H(v_i) + d_H(v_j) - 5 \\ &\geq e(H) - 5. \end{aligned} \tag{4.7.30}$$

Therefore by counting the degrees in G of the vertices in U, we have that

$$\begin{aligned} e_G(U,W) &= \sum_{v \in V_1} d_G(v) + \sum_{v \in V_2} d_G(v) - 2e(H) - 2e_G(V_1) - 2e_G(V_2) \\ &= D(|V_1| + |V_2|) - 2e(H) \\ &\quad - 2\left(\binom{|V_1|}{2} - e_{\overline{G}}(V_1) + \binom{|V_2|}{2} - e_{\overline{G}}(V_2)\right) \\ &\overset{(4.7.30)}{\geq} D(|V_1| + |V_2|) - 10 - 2\binom{|V_1|}{2} - 2\binom{|V_2|}{2} \\ &= |V_1|(D - |V_1|) + |V_2|(D - |V_2|) + |V_1| + |V_2| - 10 \geq 2D - 14, \end{aligned}$$

a contradiction to (4.7.28). This proves the claim. ■

Recall that M is a matching of size two in $G[W,U]$ with one edge incident to A and one edge incident to B. Assume without loss of generality that $e_M(V_2,W) \geq e_M(V_1,W)$. There exists $e \in E(M')$ which is vertex-disjoint from M. Suppose first that $e_M(V_2,W) = 2$. Let $e' \in E(M') \setminus \{e\}$ be arbitrary. Then $\mathcal{P} := M \cup \{e, e'\}$ satisfies (P1)–(P3). Suppose instead that $e_M(V_2,W) = e_M(V_1,W) = 1$. Then $\mathcal{P} := M \cup \{e\}$ satisfies (P1)–(P3). This completes the proof of Lemma 4.7.3 in all cases. □

4.8 The proof of Theorem C

We are now ready to prove Theorem C. It is a consequence of Theorem 3.7.11 and Lemma 3.6.2, as well as Lemmas 4.5.1, 4.6.1 and 4.7.1.

Proof of Theorem C. Choose a non-decreasing function $g : (0,1) \to (0,1)$ with $g(x) \leq x$

for all $x \in (0,1)$ such that the requirements of Proposition 3.6.1 and Lemmas 3.6.2, 4.5.1, 4.6.1, 4.7.1 (each applied, where relevant, with $1/32, 1/4$ playing the roles of η, α) are satisfied whenever $n, \rho, \gamma, \nu, \tau$ satisfy

$$1/n \leq g(\rho), g(\gamma); \quad \rho, \gamma \leq g(\nu); \quad \nu \leq g(\tau); \quad \tau \leq g(1/32). \tag{4.8.1}$$

Choose τ, τ' so that

$$0 \leq \tau' \leq \tau \leq g(1/32), 40^{-3} \quad \text{and} \quad \tau' \leq g(\tau).$$

Define a function $g' : (0,1) \to (0,1)$ by $g'(x) = (g(x))^3$. Apply Theorem 3.7.11 with $g', \tau', 1/20$ playing the roles of g, τ, ε to obtain an integer n_0. Let G be a 3-connected D-regular graph on $n \geq n_0$ vertices where $D \geq n/4$. We may assume that Theorem 3.7.11(ii) holds or we are done. Thus there exist ρ, ν with $1/n_0 \leq \rho \leq \nu \leq \tau'$, $1/n_0 \leq g'(\rho)$ and $\rho \leq g'(\nu)$; and $(k, \ell) \in \{(4,0), (2,1), (0,2)\}$ such that G has a robust partition $\mathcal{V}$ with parameters $\rho, \nu, \tau', k, \ell$ (and thus also a robust partition with parameters ρ, ν, τ, k, ℓ).

Let $\gamma := \rho^{1/3}$. Note that $n, \rho, \gamma, \nu, \tau$ satisfy (4.8.1). Apply Lemmas 4.5.1, 4.6.1 in the cases when (k, ℓ) equals $(4,0), (0,2)$ respectively to obtain a $\mathcal{V}$-tour of G with parameter γ. Proposition 3.6.1 implies that $\mathcal{V}$ is a weak robust partition with parameters $\rho, \nu, \tau, 1/32, k, \ell$. Then Lemma 3.6.2 implies that G contains a Hamilton cycle. Apply Lemma 4.7.1 in the case when $(k, \ell) = (2,1)$ to obtain a Hamilton cycle in G. This completes the proof of the theorem. $\square$

CHAPTER 5

ON DEGREE SEQUENCES FORCING THE SQUARE OF A HAMILTON CYCLE

5.1 Introduction

One of the most fundamental results in extremal graph theory is Dirac's theorem [40] which states that every graph G on $n \geq 3$ vertices with minimum degree $\delta(G)$ at least $n/2$ contains a Hamilton cycle. It is easy to see that the minimum degree condition here is best possible. The *square* of a Hamilton cycle C is obtained from C by adding an edge between every pair of vertices of distance two on C. A famous conjecture of Pósa from 1962 (see [43]) provides an analogue of Dirac's theorem for the square of a Hamilton cycle.

Conjecture 5.1.1 (Pósa [43])**.** *Let G be a graph on n vertices. If $\delta(G) \geq 2n/3$, then G contains the square of a Hamilton cycle.*

Again, it is easy to see that the minimum degree condition in Pósa's conjecture cannot be lowered. To see this, consider the complete tripartite graph whose parts are almost the same size (so, when the number of vertices n is divisible by 3, this would be $K_{n/3-1,n/3,n/3+1}$). This graph does not even contain a perfect K_3-packing, so certainly does not contain the square of a Hamilton cycle.

The conjecture was intensively studied in the 1990s (see e.g. [46, 47, 48, 49, 50]),

culminating in its proof for large graphs G by Komlós, Sárközy and Szemerédi [71]. The proof applies Szemerédi's Regularity lemma and as such the graphs G considered are extremely large. More recently, the lower bound on the size of G in this result has been significantly lowered (see [31, 87]).

Although the minimum degree condition is best possible in Dirac's theorem, this does not necessarily mean that one cannot significantly strengthen this result. Indeed, Ore [97] showed that a graph G of order $n \geq 3$ contains a Hamilton cycle if $d(x)+d(y) \geq n$ for all non-adjacent $x \neq y \in V(G)$. The following result of Pósa [100] provides a *degree sequence* condition that ensures Hamiltonicity.

Theorem 5.1.2 (Pósa [100])**.** *Let G be a graph on $n \geq 3$ vertices with degree sequence $d_1 \leq \cdots \leq d_n$. If $d_i \geq i+1$ for all $i < (n-1)/2$ and if additionally $d_{\lceil n/2 \rceil} \geq \lceil n/2 \rceil$ when n is odd, then G contains a Hamilton cycle.*

Notice that Theorem 5.1.2 is significantly stronger than Dirac's theorem as it allows for almost half of the vertices of G to have degree less than $n/2$. A theorem of Chvátal [35] generalises Theorem 5.1.2 by characterising all those degree sequences which ensure the existence of a Hamilton cycle in a graph: Suppose that the degrees of a graph G are $d_1 \leq \cdots \leq d_n$. If $n \geq 3$ and $d_i \geq i+1$ or $d_{n-i} \geq n-i$ for all $i < n/2$ then G is Hamiltonian. Moreover, if $d_1 \leq \cdots \leq d_n$ is a degree sequence that does not satisfy this condition then there exists a non-Hamiltonian graph G whose degree sequence $d'_1 \leq \cdots \leq d'_n$ is such that $d'_i \geq d_i$ for all $1 \leq i \leq n$.

Recently there has been an interest in generalising Pósa's conjecture. An 'Ore-type' analogue of Pósa's conjecture has been proven for large graphs in [30, 38]. In [3], Allen, Böttcher and Hladký determined the minimum degree threshold that ensures a large graph contains a square cycle of a given length. The focus of this chapter is to investigate degree sequence conditions that guarantee a graph contains the square of a Hamilton cycle. This problem was raised in the arXiv version of [12]. The main result of this chapter is the following approximate degree sequence version of Pósa's conjecture.

Theorem D. *Given any $\eta > 0$ there exists an $n_0 \in \mathbb{N}$ such that the following holds. If G is a graph on $n \geq n_0$ vertices whose degree sequence $d_1 \leq \cdots \leq d_n$ satisfies*

$$d_i \geq n/3 + i + \eta n \quad \textit{for all} \quad i \leq n/3,$$

then G contains the square of a Hamilton cycle.

Note that Theorem D allows for almost $n/3$ vertices in G to have degree substantially smaller than $2n/3$. However, it does not quite imply Pósa's conjecture for large graphs due to the term ηn. An example from the arXiv version of [12] shows that the term ηn in Theorem D cannot be globally replaced by $o(\sqrt{n})$ for every $i \leq n/3$. So in this sense Theorem D is close to best possible. We suspect though that the degrees in Theorem D can be capped at $2n/3$.

Conjecture 5.1.3. *Given any $\eta > 0$ there exists an $n_0 \in \mathbb{N}$ such that the following holds. If G is a graph on $n \geq n_0$ vertices whose degree sequence $d_1 \leq \cdots \leq d_n$ satisfies*

$$d_i \geq \min\{n/3 + i + \eta n, 2n/3\} \quad \textit{for all} \quad i,$$

then G contains the square of a Hamilton cycle.

It would be extremely interesting to establish an approximate analogue of Chvátal's theorem for the square of a Hamilton cycle, i.e., to provide an approximate characterisation of those degree sequences which force the square of a Hamilton cycle.

A well-known result of Aigner and Brandt [2] and Alon and Fischer [5] states that if G is a graph on n vertices with minimum degree $\delta(G) \geq (2n-1)/3$ then G contains every graph H on n vertices with maximum degree $\Delta(H) \leq 2$. (A conjecture of El-Zahar [41], that was proven for large graphs by Abbasi [1], implies that for many graphs H with $\Delta(H) \leq 2$, the minimum degree condition here can be substantially lowered.) Since a square path on n vertices contains any such graph H, an immediate consequence of Theorem D is the following degree sequence result.

Corollary 5.1.4. *Given any $\eta > 0$ there exists an $n_0 \in \mathbb{N}$ such that the following holds. Suppose that H is a graph on $n \geq n_0$ vertices such that $\Delta(H) \leq 2$. If G is a graph on n vertices whose degree sequence $d_1 \leq \cdots \leq d_n$ satisfies*

$$d_i \geq n/3 + i + \eta n \quad \textit{for all} \quad i \leq n/3,$$

then G contains H.

The case when H is a triangle factor was proved in [113], and in fact this result is used as a tool in the proof of Theorem D.

The proof of Theorem D makes use of Szemerédi's Regularity lemma [108] and the Blow-up lemma [72]. In Section 5.2 we give a detailed sketch of the proof.

5.2 Overview of the proof

Over the last few decades a number of powerful techniques have been developed for embedding problems in graphs. The Blow-up lemma [72], in combination with the Regularity lemma [108], has been used to resolve a number of long-standing open problems, including Pósa's conjecture for large graphs [71]. More recently, the so-called Connecting-Absorbing method developed by Rödl, Ruciński and Szemerédi [102] has also proven to be highly effective in tackling such embedding problems.

Typically, both these approaches have been applied to graphs with 'large' minimum degree. Our graph G in Theorem D may have minimum degree $(1/3+o(1))n$. In particular, this is significantly smaller than the minimum degree threshold that forces the square of a Hamilton cycle in a graph (namely, $2n/3$). As we describe below, having vertices of relatively small degree makes the proof of Theorem D highly involved and rather delicate. Further, we also develop a number of new ideas in order to deal with these vertices of small degree.

5.2.1 An approximate version of Pósa's conjecture

In order to highlight some of the difficulties in the proof of Theorem D, we first give a sketch of a proof of an approximate version of Pósa's conjecture. This is based on the proof of Pósa's conjecture for large graphs given in [87].

Let $0 < \varepsilon \ll \gamma \ll \eta$. Suppose that G is a sufficiently large graph on n vertices with $\delta(G) \geq (2/3+\eta)n$. We wish to find the square of a Hamilton cycle in G. The proof splits into three main parts.

- **Step 1 (Absorbing path):** Find an 'absorbing' square path P_A in G such that $|P_A| \leq \gamma n$. P_A has the property that given *any* set $A \subseteq V(G) \setminus V(P_A)$ such that $|A| \leq 2\varepsilon n$, G contains a square path P with vertex set $V(P_A) \cup A$, where the first and last two vertices on P are the same as the first and last two vertices on P_A.

- **Step 2 (Reservoir set):** Let $G' := G \setminus V(P_A)$. Find a 'reservoir' set $\mathcal{R} \subseteq V(G')$ such that $|\mathcal{R}| \leq \varepsilon n$. $\mathcal{R}$ has the property that, given *arbitrary* disjoint ordered edges $ab, cd \in E(G)$, there are 'many' short square paths P in G so that: (i) The first two vertices on P are a, b respectively; (ii) The last two vertices on P are c, d respectively; (iii) $V(P) \setminus \{a, b, c, d\} \subseteq \mathcal{R}$.

- **Step 3 (Almost tiling with square paths):** Let $G'' := G' \setminus \mathcal{R}$. Find a collection $\mathcal{P}$ of a bounded number of vertex-disjoint square paths in G'' that together cover all but εn of the vertices in G''.

Assuming that $\delta(G) \geq (2/3 + \eta)n$, the proof of each of these three steps is not too involved. (Note though that the proof in [87] is more technical since there $\delta(G) \geq 2n/3$.)

After completing Steps 1–3, it is straightforward to find the square of a Hamilton cycle in G. Indeed, suppose ab is the last edge on a square path P_1 from $\mathcal{P}$ and cd is the first edge on a square path P_2 from $\mathcal{P}$. Then Step 2 implies that we can 'go through' $\mathcal{R}$ to join P_1 and P_2 into a single square path in G. Repeating this process we can obtain a square cycle C in G that contains all the square paths from $\mathcal{P}$. Further, we may also

incorporate the absorbing square path P_A into C. C now covers almost all the vertices of G. We then use P_A to absorb all the vertices from $V(G) \setminus V(C)$ into C to obtain the square of a Hamilton cycle.

5.2.2 A degree sequence version of Pósa's conjecture

Suppose that G is a sufficiently large graph on n vertices as in the statement of Theorem D. A result of Treglown [113] guarantees that G contains a collection of $\lfloor n/3 \rfloor$ vertex-disjoint triangles. Further, this result together with a simple application of the Regularity lemma implies that G in fact contains a collection $\mathcal{P}$ of a bounded number of vertex-disjoint square paths that together cover almost all of the vertices in G. So we can indeed prove an analogue of Step 3 in this setting. In particular, if we could find a reservoir set $\mathcal{R}$ as above, then certainly we would be able to join together the square paths in $\mathcal{P}$ through $\mathcal{R}$, to obtain an almost spanning square cycle C in G.

Suppose that $ab, cd \in E(G)$ and we wish to find a square path P in G between ab and cd. If $d_G(a), d_G(b) < n/2$ then it may be the case that a and b have no common neighbours. Then it is clearly impossible to find such a square path P between ab and cd (since ab does not lie in a single square path!). The degree sequence condition on G is such that almost $n/6$ vertices in G may have degree less than $n/2$. Therefore we cannot hope to find a reservoir set precisely as in Step 2 above.

We overcome this significant problem as follows. We first show that G contains a reservoir set $\mathcal{R}$ that can *only* be used to find a square path between pairs of edges $ab, cd \in E(G)$ of *large degree* (namely, at least $(2/3 + \eta)n$). This turns out to be quite involved. In order to use $\mathcal{R}$ to join together the square paths $P \in \mathcal{P}$ into an almost spanning square cycle, we now require that the first and last two vertices on each such P have large degree.

To find such a collection of square paths $\mathcal{P}$ we first find a special collection $\mathcal{F}$ of so-called 'folded paths' in a reduced graph R of G. Roughly speaking, folded paths are a generalisation of the notion of a square path. Each such folded path $F \in \mathcal{F}$ will act

as a 'guide' for embedding one of the paths $P \in \mathcal{P}$ into G. More precisely, there is a homomorphism from a square path P into a folded path F. In particular, the structure of F will ensure that the first and last two vertices on P are 'mapped' to large degree vertices in G.

Given our new reservoir set $\mathcal{R}$ and collection of square paths $\mathcal{P}$, we again can obtain an almost spanning square cycle C in G. Further, if we could construct an absorbing square path P_A as in Step 1, we would be able to absorb the vertices in $V(G) \setminus V(C)$ to obtain the square of a Hamilton cycle. However, we were unable to construct such an absorbing square path, and do not believe there is a 'simple' way to construct one. (Though, one could construct such a square path P_A if one only requires P_A to absorb vertices of large degree.) Instead, our method now turns towards the Regularity-Blow-up approach.

Using what we have achieved thus far, we can obtain an almost spanning square cycle in the reduced graph R of G. In fact, we obtain a much richer structure Z_ℓ in R called a 'triangle cycle'. Z_ℓ is a special 6-regular graph on 3ℓ vertices that contains the square of a Hamilton cycle. In particular, Z_ℓ contains a collection of vertex-disjoint triangles T_ℓ that together cover all the vertices in Z_ℓ. We then show that G contains an almost spanning structure $\mathcal{C}$ that looks like the 'blow-up' of Z_ℓ. More precisely, if $V(Z_\ell) = \{1, \dots, 3\ell\}$ and $V_1, \dots, V_{3\ell}$ are the corresponding clusters in G, then

- $V(\mathcal{C}) = V_1 \cup \cdots \cup V_{3\ell}$;
- $\mathcal{C}[V_i, V_j]$ is ε-regular whenever $ij \in E(Z_\ell)$;
- If ij is an edge in a triangle $T \in T_\ell$ then $\mathcal{C}[V_i, V_j]$ is ε-superregular.

We call $\mathcal{C}$ a 'cycle structure'. The initial structure of $\mathcal{C}$ is such that it contains a spanning square cycle. However, since $\mathcal{C}$ is not necessarily spanning in G, this does not correspond to the square of a Hamilton cycle in G. We thus need to incorporate the 'exceptional vertices' of G into this cycle structure $\mathcal{C}$ in a balanced way so that at the end $\mathcal{C}$ (and hence G) contains the square of a Hamilton cycle. The rich structure of Z_ℓ and thus $\mathcal{C}$ is vital

for this. Again particular care is needed when incorporating exceptional vertices of small degree into our cycle structure. This part of the proof builds on ideas used in [23, 24].

Unfortunately, space considerations prevent us from presenting the proof of Theorem D in its entirety. All the details may be found in [107].

LIST OF REFERENCES

[1] S. Abbasi, Spanning cycles in dense graphs, PhD thesis, Rutgers University, 1999.

[2] M. Aigner and S. Brandt, Embedding arbitrary graphs of maximum degree two, *J. London Math. Soc.* **48** (1993), 39–51.

[3] P. Allen, J. Böttcher and J. Hladký, Filling the gap between Turán's theorem and Pósa's conjecture, *J. London Math. Soc.* **84** (2011), 269–302.

[4] N. Alon, S. Ben-Shimon and M. Krivelevich, A note on regular Ramsey graphs, *J. Graph Theory* **64** (2010), 244–249.

[5] N. Alon and E. Fischer, 2-factors in dense graphs, *Discrete Math.* **152** (1996), 13–23.

[6] N. Alon, G. Gutin and M. Krivelevich, Algorithms with large domination ratio, *J. of Algorithms* **50** (2004), 118–131.

[7] N. Alon and A. Shapira, Testing subgraphs in directed graphs, *J. of Computer and System Sciences* **69** (2004), 354–382.

[8] B. Alspach, The wonderful Walecki construction, *Bull. Inst. Combin. Appl.* **52** (2008), 7–20.

[9] B. Alspach, J.C. Bermond and D. Sotteau, Decompositions into cycles. I. Hamilton decompositions, *Cycles and rays (Montreal, PQ, 1987)*, Kluwer Acad. Publ., Dordrecht, 1990, 9–18.

[10] B. Andrásfai, P. Erdős and V.T. Sós, On the connection between chromatic number, maximal clique and minimal degree of a graph, *Disc. Math.* **8** (3) (1974), 205–218.

[11] D. Bal and A. Frieze, Packing tight Hamilton cycles in uniform hypergraphs, *SIAM J. Disc. Math.* **26** (2012), 435–451.

[12] J. Balogh, A.V. Kostochka and A. Treglown, On perfect packings in dense graphs, *Electron. J. Combin.* **20** (2013) (1) P#57, arXiv:1110.3490.

[13] B. Bollobás, *Extremal Graph Theory*, p167, Academic Press, 1978.

[14] B. Bollobás and S. E. Eldridge, Maximal matchings in graphs with given maximal and minimal degrees, *Cong. Num. XV* (1976), 165–168.

[15] B. Bollobás and A.M. Hobbs, Hamiltonian cycles in regular graphs, *Advances in Graph Theory* **3** (1978), 43–48.

[16] B. Bollobás and A. Thomason, Weakly pancyclic graphs, *J. Combin. Theory B* **77** (1999), 121–137.

[17] J.A. Bondy, Pancyclic graphs I, *J. Combin. Theory B* **11** (1971), 80–84.

[18] J.A. Bondy, Pancyclic graphs: recent results, infinite and finite sets, in Colloq. Math. Soc. János Bolyai, Keszthely, Hungary, 1973, 111–113.

[19] J.A. Bondy, Hamilton cycles in graphs and digraphs, *Congress Numerantium* **21** (1978), 3–28.

[20] J.A. Bondy, Basic graph theory: paths and circuits, *Handbook of Combinatorics, Vol. 1* (1995), 3–110.

[21] J.A. Bondy and V. Chvátal, A method in graph theory, *Disc. Math.* **15** (1976), 111–135.

[22] J.A. Bondy and U.S.R. Murty, Graph theory with applications, North-Holland, 1976.

[23] J. Böttcher, M. Schacht and A. Taraz, Spanning 3-colourable subgraphs of small bandwidth in dense graphs, *J. Combin. Theory B* **98** (2008), 752–777.

[24] J. Böttcher, M. Schacht and A. Taraz, Proof of the bandwidth conjecture of Bollobás and Komlós, *Math. Ann.* **343** (1) (2009), 175–205.

[25] S. Brandt, A sufficient condition for all short cycles, *Disc. Appl. Math.* **79** (1997), 63–66.

[26] H.J. Broersma, J. van den Heuvel, B. Jackson and H.J. Veldman, Hamiltonicity of regular 2-connected graphs, *J. Graph Theory* **22** (1996), 105–124.

[27] G. Cairns, S. Mendan and Y. Nikolayevsky, A sharp refinement of a result of Alon, Ben-Shimon and Krivelevich on bipartite graph vertex sequences, *Australas. J. Combin.* **60** (2014), 217–226.

[28] P. Camion, Chemins et circuits hamiltoniens des graphes complets, *C. R. Acad. Sci. Paris* **249** (1959), 2151–2152.

[29] P. A. Catlin, Embedding subgraphs and coloring graphs under extremal degree conditions, PhD thesis, Ohio State University, 1976.

[30] P. Châu, An Ore-type theorem on hamiltonian square cycles, *Graphs Combin.* **29** (2013), 795–834.

[31] P. Châu, L. DeBiasio and H.A. Kierstead, Pósa's Conjecture for graphs of order at least 2×10^8, *Random Structures and Algorithms* **39** (2011), 507–525.

[32] D. Christofides, J. Hladký and A. Máthé, Hamilton cycles in dense vertex-transitive graphs, *J. Combin. Theory B* **109** (2014), 34–72.

[33] D. Christofides, P. Keevash, D. Kühn and D. Osthus, A semi-exact degree condition for Hamilton cycles in digraphs, *SIAM J. Discrete Mathematics* **24** (2010), 709–756.

[34] D. Christofides, D. Kühn and D. Osthus, Edge-disjoint Hamilton cycles in graphs, *J. Combin. Theory. B* **102** (2012), 1035–1060.

[35] V. Chvátal, On Hamilton's ideals, *J. Combin. Theory B* **12** (1972), 163–168.

[36] V. Chvátal, V. Rödl, E. Szemerédi and W. R. Trotter Jr, The Ramsey number of a graph with bounded maximum degree. *J. Combin Theory B* **34** (3) (1983), 239–243.

[37] B. Csaba, D. Kühn, A. Lo, D. Osthus and A. Treglown, Proof of the 1-factorization and Hamilton decomposition conjectures, *Mem. Amer. Math. Soc.*, to appear.

[38] L. DeBiasio, S. Faizullah and I. Khan, Ore-degree threshold for the square of a Hamiltonian cycle, *Discrete Math. and Theoretical Comp. Science*, to appear.

[39] L. DeBiasio, D. Kühn, T. Molla and A. Taylor, Arbitrary orientations of Hamilton cycles in digraphs, preprint.

[40] G.A. Dirac, Some theorems on abstract graphs, *Proc. London. Math. Soc.* **2** (1952), 69–81.

[41] M.H. El-Zahar, On circuits in graphs, *Discrete Math.* **50** (1984), 227–230.

[42] P. Erdős, Some remarks on the theory of graphs, *Bull. Amer. Math. Soc.* **53** (1959), 292–294.

[43] P. Erdős, Problem 9, in: M. Fieldler (Ed.), *Theory of Graphs and its Applications*, Czech. Acad. Sci. Publ., Prague, 1964, p. 159.

[44] P. Erdős and A.M. Hobbs, A class of Hamiltonian regular graphs, *J. Combin. Theory B* **37** (1978), 129–135.

[45] G. Fan, Longest cycles in regular graphs, *J. Combin. Theory B* **39** (1985), 325–345.

[46] G. Fan, R. Häggkvist, The square of a Hamiltonian cycle, *SIAM J. Discrete Math.* **7** (1994), 203–212.

[47] G. Fan, H.A. Kierstead, The square of paths and cycles, *J. Combin. Theory B* **63** (1995), 55–64.

[48] G. Fan, H.A. Kierstead, Hamiltonian square-paths, *J. Combin. Theory B* **67** (1996), 167–182.

[49] G. Fan, H.A. Kierstead, Partitioning a graph into two square-cycles, *J. Graph Theory* **23** (1996), 241–256.

[50] R.J. Faudree, R.J. Gould, M.S. Jacobson and R.H. Schelp, On a problem of Paul Seymour, *Recent Advances in Graph Theory* (V. R. Kulli ed.), Vishwa International Publication (1991), 197–215.

[51] A. Frieze and M. Krivelevich, On packing Hamilton cycles in ε-regular graphs, *J. Combin. Theory B* **94** (2005), 159–172.

[52] A. Frieze, M. Krivelevich and P.-S. Loh, Packing tight Hamilton cycles in 3-uniform hypergraphs, *Random Structures Algorithms* **40** (2012), 269-300.

[53] A. Ghouila-Houri, Une condition suffisante d'existence d'un circuit hamiltonien, *C. R. Acad. Sci. Paris* **25** (1960), 495–497.

[54] F. Glover and A.P. Punnen, The travelling salesman problem: new solvable cases and linkages with the development of approximation algorithms, *J. of the Operational Research Society* **48** (1997), 502–510.

[55] R.J. Gould, Recent advances on the Hamiltonian problem: survey III, *Graphs and Combinatorics* **30** (1) (2014), 1–46.

[56] W.T. Gowers, Lower bounds of tower type for Szemerédi's uniformity lemma, *Geom. Func. Anal.* **7** (2) (1997), 322–337.

[57] G. Gutin and A. Yeo, TSP tour domination and Hamilton cycle decompositions of regular graphs, *Oper. Res. Letters* **28** (2001), 107–111.

[58] R. Häggkvist, Unsolved problems, *Proc. Fifth Hungarian Colloq. Comb.* (1976).

[59] A. Hajnal and E. Szemerédi, Proof of a conjecture of P. Erdős, in Combinatorial Theory and its Application, (P. Erdős, A. Rényi and V.T. Sós, Eds.), 601–623, North-Holland, London, 1970.

[60] P. Hall, On representatives of subsets, *J. London Math. Soc.* **10** (1935), 26–30.

[61] F. Hilbig, Kantenstrukturen in nichthamiltonischen Graphen, Ph.D. Thesis, Technical University Berlin 1986.

[62] B. Jackson, Edge-disjoint Hamilton cycles in regular graphs on large degree, *J. London Math. Soc.* **19** (1979), 13–16.

[63] B. Jackson, Hamilton cycles in regular 2-connected graphs, *J. Combin. Theory B* **29** (1980), 27–46.

[64] B. Jackson, Long paths and cycles in oriented graphs, *J. Graph Theory* **5** (1981), 245–252.

[65] B. Jackson and H. Li, Hamilton cycles in 2-connected regular bipartite graphs, *J. Combin. Theory B* **62** (1994), 236–258.

[66] B. Jackson, H. Li and Y. Zhu, Dominating cycles in regular 3-connected graphs, *Discrete Mathematics* **102** (1991), 163–176.

[67] S. Janson, T. Łuczak and A. Ruciński, *Random graphs*, Wiley-Interscience, 2000.

[68] H.A. Jung, Longest circuits in 3-connected graphs, *Finite and infinite sets, Vol I, II, Colloq. Math. Soc. János Bolyai* **37** (1984), 403–438.

[69] R. Karp, Reducibility among combinatorial problems, Complexity of computer computations, Plenum, New York, 1972, 85–103.

[70] P. Keevash, D. Kühn and D. Osthus, A exact minimum degree condition for Hamilton cycles in oriented graphs, *J. London Math. Soc* **79** (2009), 144–166.

[71] J. Komlós, G.N. Sárközy and E. Szemerédi, On the square of a Hamiltonian cycle in dense graphs, *Random Structures and Algorithms* **9** (1996), 193–211.

[72] J. Komlós, G. N. Sárközy and E. Szemerédi, Blow-up lemma, *Combinatorica* **17** (1997), 109–123.

[73] J. Komlós, G.N. Sárközy and E. Szemerédi, Proof of the Seymour conjecture for large graphs, *Annals of Combinatorics* **2** (1998), 43–60.

[74] D. Kühn, J. Lapinskas, D. Osthus and V. Patel, Proof of a conjecture of Thomassen on Hamilton cycles in highly connected tournaments, *Proc. London Math. Soc.* **109** (2014), 733–762.

[75] D. Kühn, A. Lo, D. Osthus and K. Staden, The robust component structure of dense regular graphs and applications, *Proc. London Math. Soc.* **110** (1) (2015), 19–56.

[76] D. Kühn, A. Lo, D. Osthus and K. Staden, Solution to a problem of Bollobás and Häggkvist on Hamilton cycles in regular graphs, preprint.

[77] D. Kühn, R. Mycroft and D. Osthus, Hamilton ℓ-cycles in uniform hypergraphs, *J. Combin. Theory A* **117** (2010), 910–927.

[78] D. Kühn, R. Mycroft and D. Osthus, An approximate version of Sumner's universal tournament conjecture, *J. Combin. Theory B* **101** (2011), 415–447.

[79] D. Kühn, R. Mycroft and D. Osthus, A proof of Sumner's universal tournament conjecture for large tournaments, *Proc. London Math. Soc.* **102** (2011), 731–766.

[80] D. Kühn and D. Osthus, A survey on Hamilton cycles in directed graphs, *European J. Combinatorics* **33** (2012), 750–766.

[81] D. Kühn and D. Osthus, Hamilton decompositions of regular expanders: a proof of Kelly's conjecture for large tournaments, *Advances in Mathematics* **237** (2013), 62–146.

[82] D. Kühn and D. Osthus, Hamilton decompositions of regular expanders: applications, *J. Combin. Theory B* **104** (2014), 1–27.

[83] D. Kühn and D. Osthus, Hamilton cycles in graphs and hypergraphs: an extremal perspective, *Proc. Int. Con. of Math. 2014*, Seoul, South Korea, **4** (2014), 381–406.

[84] D. Kühn, D. Osthus and A. Taraz, Large planar subgraphs in dense graphs, *J. Combin. Theory B* **95** (2005), 263–281.

[85] D. Kühn, D. Osthus and A. Treglown, Hamiltonian degree sequences in digraphs, *J. Combin. Theory B* **100** (2010), 367–380.

[86] D. Kühn, D. Osthus and A. Treglown, Hamilton decompositions of regular tournaments, *Proc. London Math. Soc.* **101** (2010), 303–335.

[87] I. Levitt, G.N. Sárközy and E. Szemerédi, How to avoid using the Regularity lemma: Pósa's conjecture revisited, *Discrete Math.* **310** (2010), 630–641.

[88] H. Li, Generalizations of Dirac's theorem in Hamiltonian graph theory – a survey, *Discrete Mathematics* **313** (2013), 2034–2053.

[89] H. Li and Y. Zhu, Hamilton cycles in regular 3-connected graphs, *Discrete Mathematics* **110** (1992), 229–249.

[90] L. Lovász, *Problem 11* in *Combinatorial Structures and their Applications*, Gordon and Breach Science Publishers, 1970.

[91] E. Lucas, *Récréations Mathématiques*, Vol. 2, Gautheir-Villars, 1892.

[92] M. Meyniel, Une condition suffisante d'existence d'un circuit Hamiltonien dans un graphe orienté, *J. Combin. Theory B* **14** (1973), 137–147.

[93] J.W. Moon, *Topics on tournaments*, Holt, Rinehart and Winston, New York, 1968.

[94] C.St.J.A. Nash-Williams, Hamiltonian arcs and circuits, *Recent trends in Graph Theory* **186** (1971), 197–210.

[95] C.St.J.A. Nash-Williams, Hamiltonian lines in graphs whose vertices have sufficiently large valencies, in *Combinatorial theory and its applications, III (Proc. Colloq., Balatonfüred, 1969)*, North-Holland, Amsterdam (1970) 813–819.

[96] C.St.J.A. Nash-Williams, Valnecy sequences which force graphs to have Hamiltonian circuits, University of Waterloo Research Report, Waterloo, Ontario, 1969.

[97] O. Ore, Note on Hamilton circuits, *Amer. Math. Monthly* **67** (1960), 55.

[98] D. Osthus and K. Staden, Approximate Hamilton decompositions of regular robustly expanding digraphs, *SIAM J. Disc. Math.* **27** (3), (2013), 1372–1409.

[99] A. Pokrovskiy, Edge disjoint Hamiltonian cycles in highly connected tournaments, arXiv:1406.7552

[100] L. Pósa, A theorem concerning Hamiltonian lines, *Magyar Tud. Akad. Mat. Fiz. Oszt. Kozl.* **7** (1962), 225–226.

[101] V. Rödl and A. Ruciński, Dirac-type questions for hypergraphs – a survey (or more problems for Endre to solve), An Irregular Mind, *Bolyai Soc. Math. Studies* **21** (2010), 561–590.

[102] V. Rödl, A. Ruciński and E. Szemerédi, A Dirac-type theorem for 3-uniform hypergraphs, *Combin. Probab. Comput.* **15** (2006), 229–251.

[103] V. Rödl, A. Ruciński and E. Szemerédi, An approximate Dirac-type theorem for k-uniform hypergraphs, *Combinatorica* **28** (2008), 229–260.

[104] V. Rödl, A. Ruciński and E. Szemerédi, Dirac-type conditions for hamiltonian paths and cycles in 3-uniform hypergraphs, *Advances in Mathematics* **227** (2011), 1225–1299.

[105] P. Seymour, Problem section, in *Combinatorics: Proceedings of the British Combinatorial Conference 1973* (T.P. McDonough and V.C. Mavron eds.), 201–202, Cambridge University Press, 1974.

[106] A. Srivastav and P. Stangier, Algorithmic Chernoff-Hoeffding inequalities in integer programming, *Random Structures Algorithms* **8** (1996), 27–58.

[107] K. Staden and A. Treglown, On degree sequences forcing the square of a Hamilton cycle, preprint.

[108] E. Szemerédi, Regular partitions of graphs, *Problémes Combinatoires et Théorie des Graphes Colloques Internationaux CNRS* **260** (1978), 399–401.

[109] E. Szemerédi, On sets of integers containing no k elements in arithmetic progression, *Acta Arith.* **27** (1975), 199–245.

[110] C. Thomassen, Long cycles in digraphs with constraints on the degrees, in *Surveys in Combinatorics* (B. Bollobás ed.), *London Math. Soc. Lecture Notes* **38** (1979), 211–228.

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

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

[113] A. Treglown, A degree sequence Hajnal-Szemerédi theorem, preprint.

[114] B. Wei, On the circumferences of regular 2-connected graphs, *J. Combin. Theory B* **75** (1999), 88–99.

[115] D. Woodall, Sufficient conditions for cycles in digraphs, *Proc. London Math. Soc.* **24** (1972), 739–755.