SUBSTRUCTURES IN LARGE GRAPHS

by

AMELIA MAY TAYLOR

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

School of Mathematics
College of Engineering and Physical Sciences
University of Birmingham
August 2016

Abstract

The first problem we address concerns Hamilton cycles. Suppose G is a large digraph in which every vertex has in- and outdegree at least $|G|/2$. We show that G contains every orientation of a Hamilton cycle except, possibly, the antidirected one. The antidirected case was settled by DeBiasio and Molla. Our result is best possible and improves on an approximate result by Häggkvist and Thomason.

We then investigate the random greedy F-free process which was initially studied by Erdős, Suen and Winkler and by Spencer. This process greedily adds edges without creating a copy of F, terminating in a maximal F-free graph. We provide an upper bound on the number of hyperedges at the end of this process for a large class of hypergraphs.

The remainder of this thesis focuses on F-decompositions, i.e., whether the edge set of a graph can be partitioned into copies of F. We obtain the best known bounds on the minimum degree which ensures a K_r-decomposition of an r-partite graph, with applications to Latin squares. Lastly, we find exact bounds on the minimum degree for a large graph to have a C_{2k}-decomposition where $k \neq 3$. In both cases, we assume necessary divisibility conditions are satisfied.

ACKNOWLEDGEMENTS

First and foremost, I am indebted to my supervisors Daniela Kühn and Deryk Osthus for their encouragement, guidance and expertise.

Thank you to everyone I have met over the years at the University of Birmingham. You have made my time here truly special. Finally, thanks are due to my family for their constant support.

CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Extremal graph theory

What conditions guarantee that a graph contains a triangle? When we can be sure that a graph contains a perfect matching? These questions are typical of those asked in extremal graph theory. Indeed, Mantel [59] showed that having more than $|G|^2/4$ edges suffices for a graph G to contain a triangle. In a similar spirit, Tutte [79] described all graphs which have a perfect matching. Extremal results demonstrate how global parameters such as the total number of edges or the chromatic number of a graph can have considerable influence on its local structure. Take Turán's theorem [78], for example, which determines the maximum number of edges in any graph which contains no clique of size r.

Often, the subgraph of interest will be a Hamilton cycle, that is, a cycle which visits every vertex of the graph exactly once. The problem of finding a Hamilton cycle is exactly that faced in the famous Travelling Salesman Problem which has long fascinated mathematicians. Imagine a salesman has been given a list of cities. He must visit each city exactly once before returning to his starting point. Clearly he wants to minimise the time spent travelling, so the question we are asked is: can we find a Hamilton cycle of minimum length? Problems of this type are faced daily by those working in logistics, transport and telecommunications.

Karp [45] showed that the problem of finding a Hamilton cycle in a graph is NP-complete and so it is unlikely that we can find a complete classification of those graphs that are Hamiltonian. Instead, we seek sufficient conditions that will ensure a graph contains a Hamilton cycle. These conditions often involve the minimum degree or the degree sequence of the graph.

A classical result is Dirac's theorem [26] which states that if G is a graph on $n \geq 3$ vertices with minimum degree $\delta(G) \geq n/2$, then G contains a Hamilton cycle. This result is best possible in that there are graphs with minimum degree $\lceil n/2 \rceil - 1$ which do not contain a Hamilton cycle. Indeed, if n is even, consider the graph consisting of two disjoint cliques each on $n/2$ vertices and if n is odd consider the complete bipartite graph with vertex classes of size $(n-1)/2$ and $(n+1)/2$.

It is also natural to consider conditions on the degree sequence of a graph. We define the *degree sequence* of G to be the sequence $d_1, d_2, \dots, d_n$, which lists the degrees of the vertices in G such that $d_1 \leq d_2 \leq \dots \leq d_n$. In 1962, Pósa showed that if $d_i \geq i+1$ for all $i < (n-1)/2$ and, if n is odd, $d_{\lceil n/2 \rceil} \geq \lceil n/2 \rceil$, then G contains a Hamilton cycle. This result is much stronger than Dirac's theorem since we allow the graph to contain vertices with degree much smaller than $n/2$. Chvátal [22] generalised Pósa's theorem further still by describing those degree sequences which ensure that a graph is Hamiltonian. His result states that if G is a graph on n at least three vertices and $d_i \geq i+1$ or $d_{n-i} \geq n-i$ for all $i < n/2$, then G has a Hamilton cycle. What is more, if we fix n and $1 \leq r < n/2$, we can find a graph G with degree sequence $d_1, \dots, d_n$ which satisfies this condition, apart from at $i = r$ in which case $d_r = r$ and $d_{n-r} = n-r-1$, such that G does not contain a Hamilton cycle. This shows that Chvátal's theorem is best possible.

Hamiltonicity has also been intensively studied in the digraph setting. Many results involve the minimum semidegree $\delta^0(G)$ of a digraph G, the minimum of all the in- and outdegrees of the vertices in G. For instance, Ghouila-Houri [37] proved an analogue of Dirac's theorem for digraphs which guarantees that any digraph of minimum semidegree $\delta^0(G) \geq n/2$ contains a consistently oriented Hamilton cycle. (By *consistently oriented* we mean that all edges of the cycle are oriented in the same direction.) To see that Ghouila-

Houri's result is best possible, take the extremal graphs for Dirac's theorem and orient the edges in both directions. Whether an analogue of Chvátal's theorem, conjectured by Nash-Williams [63] in 1975, holds for digraphs remains an open problem.

An oriented graph is a special type of digraph which can be obtained by orienting the edges of a graph. So, whilst in a digraph we allow two edges of opposite orientations between a pair of vertices, in an oriented graph at most one edge is allowed between any pair of vertices. Keevash, Kühn and Osthus [47] proved a version of Dirac's theorem for oriented graphs. Here the minimum semidegree threshold turns out to be $\delta^0(G) \geq (3n-4)/8$.

A notion that has proved very useful in the search for Hamilton cycles is that of robust expansion, first introduced by Kühn, Osthus and Treglown in [57]. Roughly speaking, a digraph is a robust outexpander if every set of vertices of reasonable size has an out-neighbourhood at least a little larger than itself and this property should hold, even if we delete a small proportion of edges from the graph. Kühn, Osthus and Treglown showed that if a sufficiently large digraph is a robust expander then a linear minimum semidegree condition $\delta^0(G) \geq \eta n$ guarantees a consistently oriented Hamilton cycle. Their proof uses Szemerédi's regularity lemma [73]. This powerful tool allows us to approximate any large graph by a random one and is particularly useful in embedding problems.

So far, we have always assumed that in a digraph, the edges of a Hamilton cycle should be oriented consistently around the cycle. But it is natural to seek minimum semidegree conditions which guarantee a Hamilton cycle whose edges have any prescribed orientation. Perhaps we desire a Hamilton cycle whose edges are oriented alternately forwards and backwards around the cycle (we call such a cycle *antidirected*). In Chapter 2, we provide an exact bound on the minimum semidegree for a (sufficiently large) digraph to contain any given orientation of a Hamilton cycle.

1.2 Probabilistic graph theory

We are surrounded by a vast collection of networks: social networks, transport infrastructures and the internet to name but a few. Random graphs have attracted significant attention because of their potential to model these large networks. There is also a keen interest in generating random graphs for the purpose of algorithm testing. Running an algorithm on a random graph allows us to analyse how well the algorithm performs on average.

Another motivation for the study of random graphs is the following. Sometimes it can be difficult to find a graph satisfying a certain property $\mathcal{P}$ and this is when probabilistic methods come into their own. Instead of trying to design the required graph, we construct one at random. If we are able to show that this random graph satisfies $\mathcal{P}$ with positive probability, we have proved the existence of a graph with property $\mathcal{P}$ without ever finding it explicitly. A surprising result obtained using probabilistic techniques, due to Erdős [31], proves the existence of graphs with large girth (i.e. graphs containing no short cycles) and large chromatic number.

Ramsey theory, the search for structure in large graphs, provided Erdős with the initial motivation for developing probabilistic techniques. Ramsey's theorem [69] tells us that, given any sufficiently large graph, we are guaranteed to find a large complete graph or a large independent set. The Ramsey number $R(s,t)$ is the smallest positive integer n such that any graph on n vertices contains a clique of size s or an independent set of size t. Ramsey numbers are notoriously difficult to calculate and, as a result, very few are known. In 1947, Erdős [30] considered a random two-colouring to give a lower bound on the diagonal Ramsey number $R(k,k)$.

The binomial random graph $G_{n,p}$ is the probability space consisting of all graphs G with n vertices and an edge between each pair of vertices independently with probability p. For example, we could construct the random graph $G_{n,1/2}$ by tossing a coin for each pair of vertices in turn and drawing an edge if the coin shows heads. Random graphs have been studied extensively and questions asked include:

(i) for what values of p do we expect $G_{n,p}$ to be connected and

(ii) how large does p have to be before $G_{n,p}$ will almost certainly contain a Hamilton cycle?

The first of these was answered by Erdős and Rényi [32]. Bollobás [15] and Ajtai, Komlós and Szemerédi [1] solved the second. The $G_{n,p}$ model continues to captivate mathematicians and forms the basis of a huge body of research.

Random graph processes are used to gain an insight into how the random graph develops over time. We start by assigning a birthtime which is uniformly distributed in $[0, 1]$ to each edge of the complete graph on n vertices. Initially the graph is empty and we gradually increase p, adding in new edges as they are born. At time p in this process, the graph is $G_{n,p}$. This process is well understood but the analysis becomes more complicated if we add extra rules. For example, we can produce a graph with bounded maximum degree by adding the condition that an edge can only be added if it does not create a vertex of degree greater than d. This process was studied by Ruciński and Wormald [70].

We are particularly interested in how local constraints can influence the global evolution of a random process. Studying these processes allows us to obtain probabilistic analogues of classical extremal problems. Recall Mantel's theorem which says that any graph which does not contain a triangle has at most $|G|^2/4$ edges. The complete bipartite graph $K_{n/2,n/2}$ attains this bound. We can also study a random graph process which creates maximal triangle-free graphs. At each step of the triangle-free process, we only add in the new edge if it does not create a triangle. By counting the average number of edges at the end of this process, we obtain a lower bound on the number of edges permitted in a triangle-free graph. This greedy process falls significantly short of $n^2/4$ edges, so studying random processes can be thought of as analysing an average case.

The triangle-free process was suggested as a means to study the off-diagonal Ramsey number $R(3, k)$ and was first investigated by Erdős, Suen and Winkler [33] and Spencer [72]. Bohman and Keevash [13] and Fiz Pontiveros, Griffiths and Morris [35] studied this process using the differential equation method introduced by Wormald [84]. Independently, they obtained a lower bound of $R(3,k) \geq (1/4 - o(1))k^2/\log k$, improving on previous results in [50] and [11]. The power of random techniques is highlighted by the fact that the best explicit construction (i.e. the largest known concrete example of a graph with no triangles and no independent set of size k) gives a lower bound of only $\Omega(k^{3/2})$, see [2]. The natural variant of the triangle-free process, the F-free process where F is any fixed graph, is discussed further in Chapter 3.

It is logical to start asking similar questions of hypergraphs. A k-uniform hypergraph is made up of hyperedges, each of which contains exactly k vertices (so a 2-uniform hypergraph is a graph). We can define a random k-uniform hypergraph $H_{n,p}$ in exactly the same way as $G_{n,p}$ and we can now consider random hypergraph processes. However, much less is known about these processes for hypergraphs than graphs. We investigate the F-free hypergraph process in Chapter 3.

1.3 Graph decompositions

Given graphs F and G, is it possible to cover the edges of G completely using edge-disjoint copies of F? If the answer to this question is yes, we say that G has an F-decomposition. One of the first results of this kind was proved by Kirkman [51] in 1847. He showed that the complete graph on n vertices can be decomposed into triangles if and only if $n \equiv 1, 3 \mod 6$. In order for a graph G to have a triangle decomposition, it is clearly necessary that the number of edges in G must be divisible by three and that every vertex in G must have even degree (these conditions are guaranteed for K_n precisely when $n \equiv 1, 3 \mod 6$). We say that a graph which satisfies these edge and degree divisibility conditions is K_3-divisible. But these conditions alone are not sufficient; there are graphs which are K_3-divisible but which do not have a K_3-decomposition, take G to be a cycle of length six for example.

In 1850, Kirkman [52] set the following puzzle in the Lady's and Gentleman's Diary:

> *Fifteen schoolgirls must go for a walk three abreast each day for seven days. Find an arrangement so that no pair of girls must walk in the same row as each other more than once.*

With a little thought, we can translate this problem into a graph setting. Each girl becomes a vertex and we add all edges between. Any row of three girls defines a triangle so, on any day, the arrangement defines five disjoint triangles in this graph. An answer to the puzzle gives a triangle decomposition of the graph since each edge (or pair of girls) appears in a triangle (or row) exactly once. In fact, what we have just seen is an example of a Steiner triple system. A Steiner triple system is a family of triples $\mathcal{S} \subseteq \{1, \ldots, n\}$ such that every pair $\{i, j\} \subseteq \{1, \ldots, n\}$ lies in exactly one set $S \in \mathcal{S}$. This system is none other than a K_3-decomposition of K_n. So an equivalent statement of Kirkman's theorem would be: Steiner triple systems exist if and only if $n \equiv 1, 3 \mod 6$. A famous example of a Steiner triple system is the Fano plane.

In a similar fashion, Steiner systems can be defined for larger sets. In general, a Steiner system $S(t, k, n)$ is a family of k-sets $\mathcal{S} \subseteq \{1, \ldots, n\}$ such that every t-set is contained in exactly one $S \in \mathcal{S}$. A key objective in design theory is to determine for which values of t, k and n Steiner systems exist. When $t = 2$, Steiner systems $S(2, k, n)$ correspond directly to K_k-decompositions of K_n and, for higher values of t, Steiner systems relate to hypergraph decompositions, see Keevash [46]. This means that decompositions are particularly prevalent in design theory.

But let us return once again to 1847 when Kirkman determined exactly which cliques have triangle decompositions. It would be more than 100 years before anyone generalised Kirkman's result and the person in question was Wilson [82]. Wilson proved an analogue for arbitrary F-decompositions of large cliques. He showed that any sufficiently large clique which satisfies the necessary divisibility conditions can be decomposed into copies of F. But when G is not a clique, deciding whether G is F-decomposable is a very difficult problem. In fact, it is NP-complete when F has a connected component with at least three edges, see [27]. For this reason, we seek sufficient conditions which guarantee an F-decomposition and these often focus on the minimum degree. For triangles, Nash-

Williams [62] conjectured that every K_3-divisible graph G with minimum degree at least $3|G|/4$ has a K_3-decomposition. Recently, there has been much progress in the study of F-decompositions. For example, Barber, Kühn, Lo and Osthus [7], showed how to turn an approximate F-decomposition (one which covers almost all of the edges in G) into a perfect one. This reduces the problem of finding a decomposition to instead bounding the so-called fractional decomposition threshold. Currently, the best known minimum degree bound for triangles is $0.9|G|$ (see [28]), still some way away from Nash-Williams' conjectured bound.

We can even extend the Hamiltonicity problem discussed in Section 1.1 into a decomposition setting. Here, we are interested in whether a graph G has a Hamilton-decomposition, that is, whether we can partition the edges of G into edge-disjoint Hamilton cycles. In 1892, Walecki showed that every clique on an odd number of vertices has a Hamilton-decomposition (see [4], for example). Tillson [77] considered the directed analogue, determining when a complete digraph has a Hamilton-decomposition. Kelly conjectured in 1962 that every regular tournament (an orientation of the complete graph K_n) should also have a decomposition into Hamilton cycles. This was recently verified for large n by Kühn and Osthus [55]. A surprising application of their result is to the Asymmetric Travelling Salesman Problem (a weighted directed version of the problem discussed in Section 1.1).

In this thesis, we will investigate two distinct decomposition problems. The first of which is explored in Chapter 4 and concerns clique decompositions of graphs in a multipartite setting. For instance, we bound the minimum degree for a tripartite graph to have a decomposition into triangles. The direct correspondence between such decompositions and Latin squares makes these results particularly meaningful. Latin squares are $n \times n$ grids which are filled with entries from $\{1, \ldots, n\}$ in such a way that each number appears exactly once in each row and column. They were notably investigated by Euler. These grids appear in many branches of mathematics, studied not only for their own sake (they form the basis of the popular Sudoku puzzle), but because of their applications to experiment design, group theory and error-correcting codes. In Chapter 4, we address the question: given a partially completed Latin square, when are we able to fill in the rest of the boxes?

Finally, in Chapter 5 we investigate C_{2k}-decompositions, that is, decompositions of graphs into cycles of even length. We determine exact minimum degree bounds for a graph G (which is large and satisfies the necessary divisibility conditions) to have such a decomposition for all lengths apart from six.

Chapter 2 is based on work with DeBiasio, Kühn, Molla and Osthus [24]. Chapter 3 is based on work with Kühn and Osthus [56]. Chapter 4 is based on work with Barber, Kühn, Lo and Osthus [8]. Chapter 5 is based on [75].

CHAPTER 2

ARBITRARY ORIENTATIONS OF HAMILTON CYCLES IN DIGRAPHS

2.1 Introduction

A classical result on Hamilton cycles is Dirac's theorem [26] which states that if G is a graph on $n \geq 3$ vertices with minimum degree $\delta(G) \geq n/2$, then G contains a Hamilton cycle. Ghouila-Houri [37] proved an analogue of Dirac's theorem for digraphs which guarantees that any digraph of minimum semidegree at least $n/2$ contains a consistently oriented Hamilton cycle (where the minimum semidegree $\delta^0(G)$ of a digraph G is the minimum of all the in- and outdegrees of the vertices in G). In [47], Keevash, Kühn and Osthus proved a version of this theorem for oriented graphs. Here the minimum semidegree threshold turns out to be $\delta^0(G) \geq (3n-4)/8$. (In a digraph we allow two edges of opposite orientations between a pair or vertices, in an oriented graph at most one edge is allowed between any pair of vertices.)

Instead of asking for consistently oriented Hamilton cycles in an oriented graph or digraph, it is natural to consider different orientations of a Hamilton cycle. For example, Thomason [76] showed that every sufficiently large strongly connected tournament contains every orientation of a Hamilton cycle. Häggkvist and Thomason [42] proved an approximate version of Ghouila-Houri's theorem for arbitrary orientations of Hamilton cycles. They showed that a minimum semidegree of $n/2+n^{5/6}$ ensures the existence of an arbitrary orientation of a Hamilton cycle in a digraph. This improved a result of Grant [39] for antidirected Hamilton cycles. The exact threshold in the antidirected case was obtained by DeBiasio and Molla [25], here the threshold is $\delta^0(G) \geq n/2 + 1$, i.e., larger than in Ghouila-Houri's theorem. In Figure 2.1, we give two digraphs G on $2m$ vertices which satisfy $\delta^0(G) = m$ and have no antidirected Hamilton cycle, showing that this bound is best possible. (The first of these examples is already due to Cai [18].)

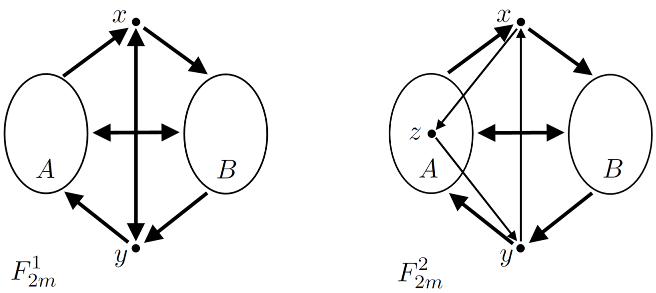

Figure 2.1: In digraphs F^1_{2m} and F^2_{2m}, A and B are independent sets of size $m-1$ and bold arrows indicate that all possible edges are present in the directions shown.

Theorem 2.1.1 (DeBiasio & Molla, [25]). *There exists an integer m_0 such that the following hold for all $m \geq m_0$. Let G be a digraph on $2m$ vertices. If $\delta^0(G) \geq m$, then G contains an antidirected Hamilton cycle, unless G is isomorphic to F^1_{2m} or F^2_{2m}. In particular, if $\delta^0(G) \geq m+1$, then G contains an antidirected Hamilton cycle.*

In this thesis, we settle the problem by completely determining the exact threshold for arbitrary orientations. We show that a minimum semidegree of $n/2$ suffices if the Hamilton cycle is not antidirected. This bound is best possible by the extremal examples for Ghouila-Houri's theorem, i.e., if n is even, the digraph consisting of two disjoint complete digraphs on $n/2$ vertices and, if n is odd, the complete bipartite digraph with vertex classes of size $(n-1)/2$ and $(n+1)/2$.

Theorem 2.1.2. *There exists an integer n_0 such that the following holds. Let G be a digraph on $n \geq n_0$ vertices with $\delta^0(G) \geq n/2$. If C is any orientation of a cycle on n vertices which is not antidirected, then G contains a copy of C.*

Kelly [49] proved an approximate version of Theorem 2.1.2 for oriented graphs. He showed that the semidegree threshold for an arbitrary orientation of a Hamilton cycle in an oriented graph is $3n/8 + o(n)$. It would be interesting to obtain an exact version of this result.

2.2 Proof sketch

The proof of Theorem 2.1.2 utilizes the notion of robust expansion which has been very useful in several settings recently. Roughly speaking, a digraph G is a robust outexpander if every vertex set S of reasonable size has an outneighbourhood which is at least a little larger than S itself, even if we delete a small proportion of the edges of G. A formal definition of robust outexpansion is given in Section 2.4. In Lemma 2.4.4, we observe that any graph satisfying the conditions of Theorem 2.1.2 must be a robust outexpander or have a large set which does not expand, in which case we say that G is ε-extremal. Theorem 2.1.2 was verified for the case when G is a robust outexpander by Taylor in [74] based on the approach of Kelly [49]. This allows us to restrict our attention to the ε-extremal case. We introduce three refinements of the notion of ε-extremality: ST-extremal, AB-extremal and $ABST$-extremal. These are illustrated in Figure 2.2, the arrows indicate that G is almost complete in the directions shown. In each of these cases, we have that $|A| \sim |B|$ and $|S| \sim |T|$. If G is ST-extremal, then the sets A and B are almost empty and so G is close to the digraph consisting of two disjoint complete digraphs on $n/2$ vertices. If G is AB-extremal, then the sets S and T are almost empty and so in this case G is close to the complete bipartite digraph with vertex classes of size $n/2$ (thus both digraphs in Figure 2.1 are AB-extremal). Within each of these cases, we further subdivide the proof depending on how many changes of direction the desired Hamilton cycle has. Note that in the directed setting the set of extremal structures is much less restricted than in the undirected setting (in the undirected case, it is well known that the extremal graphs are close to the complete bipartite graph $K_{n/2,n/2}$ or two disjoint cliques on $n/2$ vertices).

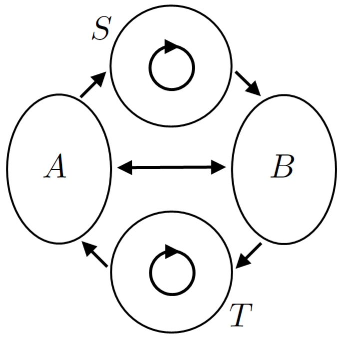

Figure 2.2: An $ABST$-extremal graph. When G is AB-extremal, the sets S and T are almost empty and when G is ST-extremal the sets A and B are almost empty.

The main difficulty in each of the cases is covering the exceptional vertices, i.e., those vertices with low in- or outdegree in the vertex classes where we would expect most of their neighbours to lie. When G is AB-extremal, we also consider the vertices in $S \cup T$ to be exceptional and, when G is ST-extremal, we consider the vertices in $A \cup B$ to be exceptional. In each case we find a short path P in G which covers all of these exceptional vertices. When the cycle C is close to being consistently oriented, we cover these exceptional vertices by short consistently oriented paths and when C has many changes of direction, we will map sink or source vertices in C to these exceptional vertices (here a sink vertex is a vertex of indegree two and a source vertex is a vertex of outdegree two).

An additional difficulty is that in the AB- and $ABST$-extremal cases we must ensure that the path P leaves a balanced number of vertices in A and B uncovered. Once we have found P in G, the remaining vertices of G (i.e., those not covered by P) induce a balanced almost complete bipartite digraph and one can easily embed the remainder of C using a bipartite version of Dirac's theorem. When G is ST-extremal, our aim will be to split the cycle C into two paths P_S and P_T and embed P_S into the digraph $G[S]$ and P_T into $G[T]$. So a further complication in this case is that we need to link together P_S and P_T as well as covering all vertices in $A \cup B$.

This chapter is organised as follows. Sections 2.3 and 2.4 introduce the notation and tools which will be used throughout this chapter. In Section 2.4.3 we describe the structure of an ε-extremal digraph and formally define what it means to be ST-, AB- or $ABST$-extremal. The remaining sections prove Theorem 2.1.2 in each of these three cases: we consider the ST-extremal case in Section 2.5, the AB-extremal case in Section 2.6 and the $ABST$-extremal case in Section 2.7.

2.3 Notation

Let G be a digraph on n vertices. We will write $xy \in E(G)$ to indicate that G contains an edge oriented from x to y. If G is a digraph and $x \in V(G)$, we will write $N^+_G(x)$ for the *outneighbourhood* of x and $N^-_G(x)$ for the *inneighbourhood* of x. We define $d^+_G(x) := |N^+_G(x)|$ and $d^-_G(x) := |N^-_G(x)|$. We will write, for example, $d^\pm_G(x) \geq a$ to mean $d^+_G(x), d^-_G(x) \geq a$. We sometimes omit the subscript G if this is unambiguous. We let $\delta^0(G) := \min\{d^+(x), d^-(x) : x \in V(G)\}$. If $A \subseteq V(G)$, we let $d^+_A(x) := |N^+_G(x) \cap A|$ and define $d^-_A(x)$ and $d^\pm_A(x)$ similarly. We say that $x \in V(G)$ is a *sink vertex* if $d^+(x) = 0$ and a *source vertex* if $d^-(x) = 0$.

Let $A, B \subseteq V(G)$ and $xy \in E(G)$. If $x \in A$ and $y \in B$ we say that xy is an AB-*edge*. We write $E(A, B)$ for the set of all AB-edges and we write $E(A)$ for $E(A, A)$. We let $e(A, B) := |E(A, B)|$ and $e(A) := |E(A)|$. We write $G[A, B]$ for the digraph with vertex set $A \cup B$ and edge set $E(A, B) \cup E(B, A)$ and we write $G[A]$ for the digraph with vertex set A and edge set $E(A)$. We say that a path $P = x_1x_2 \dots x_q$ is an AB-*path* if $x_1 \in A$ and $x_q \in B$. If $x_1, x_q \in A$, we say that P is an A-*path*. If $A \subseteq V(P)$, we say that P *covers* A. If $\mathcal{P}$ is a collection of paths, we write $V(\mathcal{P})$ for $\bigcup_{P \in \mathcal{P}} V(P)$.

Let $P = x_1x_2 \dots x_q$ be a path. The *length* of P is the number of its edges. Given sets $X_1, \dots, X_q \subseteq V(G)$, we say that P has *form* $X_1X_2 \dots X_q$ if $x_i \in X_i$ for $i = 1, 2, \dots, q$. We will use the following abbreviation

$$(X)^k := \underbrace{XX \dots X}_{k \text{ times}}.$$

We will say that P is a *forward* path of the form $X_1X_2\dots X_q$ if P has form $X_1X_2\dots X_q$ and $x_ix_{i+1} \in E(P)$ for all $i = 1, 2, \dots, q-1$. Similarly, P is a *backward* path of the form $X_1X_2\dots X_q$ if P has form $X_1X_2\dots X_q$ and $x_{i+1}x_i \in E(P)$ for all $i = 1, 2, \dots, q-1$.

A digraph G is *oriented* if it is an orientation of a simple graph (i.e., if there are no $x, y \in V(G)$ such that $xy, yx \in E(G)$). Suppose that $C = (u_1u_2\dots u_n)$ is an oriented cycle. We let $\sigma(C)$ denote the number of sink vertices in C. We will write $(u_iu_{i+1}\dots u_j)$ or (u_iCu_j) to denote the subpath of C from u_i to u_j. In particular, (u_iu_{i+1}) may represent the edge u_iu_{i+1} or $u_{i+1}u_i$. Given edges $e = (u_i, u_{i+1})$ and $f = (u_j, u_{j+1})$, we write (eCf) for the path (u_iCu_{j+1}). We say that an edge (u_iu_{i+1}) is a *forward edge* if $(u_iu_{i+1}) = u_iu_{i+1}$ and a *backward edge* if $(u_iu_{i+1}) = u_{i+1}u_i$. We say that a cycle is *consistently oriented* if all of its edges are oriented in the same direction (forward or backward). We define a consistently oriented subpath P of C in the same way. We say that P is *forward* if it consists of only forward edges and *backward* if it consists of only backward edges. A collection of subpaths of C is *consistent* if they are all forward paths or if they are all backward paths. We say that a path or cycle is *antidirected* if it contains no consistently oriented subpath of length two.

Given C as above, we define $d_C(u_i, u_j)$ to be the length of the path (u_iCu_j) (so, for example, $d_C(u_1, u_n) = n-1$ and $d_C(u_n, u_1) = 1$). For a subpath $P = (u_iu_{i+1}\dots u_k)$ of C, we call u_i the *initial* vertex of P and u_k the *final* vertex. We write $(u_jP) := (u_ju_{j+1}\dots u_k)$ and $(Pu_j) := (u_iu_{i+1}\dots u_j)$. If P_1 and P_2 are subpaths of C, we define $d_C(P_1, P_2) := d_C(v_1, v_2)$, where v_i is the initial vertex P_i. In particular, we will use this definition when one or both of P_1, P_2 are edges. Suppose $P_1, P_2, \dots, P_k$ are internally disjoint subpaths of C such that the final vertex of P_i is the initial vertex of P_{i+1} for $i = 1, \dots, k-1$. Let x denote the initial vertex of P_1 and y denote the final vertex of P_k. If $x \neq y$, we write $(P_1P_2\dots P_k)$ for the subpath of C from x to y. If $x = y$, we sometimes write $C = (P_1P_2\dots P_k)$.

Throughout this thesis we will use hierarchies, for example $1/n \ll a \ll b < 1$, where constants are chosen from right to left. The notation $a \ll b$ means that there exists an increasing function f for which the result holds whenever $a \leq f(b)$. In order to simplify the presentation, we will not determine these functions explicitly.

2.4 Tools

2.4.1 Hamilton cycles in dense graphs and digraphs

We will use the following standard results concerning Hamilton paths and cycles. Theorem 2.4.1 is a bipartite version of Dirac's theorem. Proposition 2.4.2 is a simple consequence of Dirac's theorem and this bipartite version.

Theorem 2.4.1 (Moon & Moser, [61])**.** *Let $G = (A, B)$ be a bipartite graph with $|A| = |B| = n$. If $\delta(G) \geq n/2 + 1$, then G contains a Hamilton cycle.*

Proposition 2.4.2. (i) *Let G be a digraph on n vertices with $\delta^0(G) \geq 7n/8$. Let $x, y \in V(G)$ be distinct. Then G contains a Hamilton path of any orientation between x and y.*

(ii) *Let $m \geq 10$ and $G = (A, B)$ be a bipartite digraph with $|A| = m + 1$ and $|B| = m$. Suppose that $\delta^0(G) \geq (7m + 2)/8$. Let $x, y \in A$. Then G contains a Hamilton path of any orientation between x and y.*

Proof. To prove (i), we define an undirected graph G' on the vertex set $V(G)$ where $uv \in E(G')$ if and only if $uv, vu \in E(G)$. Let G'' be the graph obtained from G' by contracting the vertices x and y to a single vertex x' with $N_{G''}(x') := N_{G'}(x) \cap N_{G'}(y)$. Note that

$$\delta(G'') \geq (n - 1)/2 = |G''|/2.$$

Hence G'' has a Hamilton cycle by Dirac's theorem. This corresponds to a Hamilton path of any orientation between x and y in G.

For (ii), we proceed in the same way, using Theorem 2.4.1 instead of Dirac's theorem.

□

2.4.2 Robust expanders

Let $0 < \nu \leq \tau < 1$, let G be a digraph on n vertices and let $S \subseteq V(G)$. The *ν-robust outneighbourhood* $RN^+_{\nu,G}(S)$ of S is the set of all those vertices $x \in V(G)$ 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 $\tau n < |S| < (1-\tau)n$.

Recall from Section 2.1 that Kelly [49] showed that any sufficiently large oriented graph with minimum semidegree at least $(3/8+\alpha)n$ contains any orientation of a Hamilton cycle. It is not hard to show that any such oriented graph is a robust outexpander (see [55]). In fact, in [49], Kelly observed that his arguments carry over to robustly expanding digraphs of linear degree. Taylor [74] has verified that this is indeed the case, proving the following result.

Theorem 2.4.3 ([74])**.** *Suppose $1/n \ll \nu \leq \tau \ll \eta < 1$. Let G be a digraph on n vertices with $\delta^0(G) \geq \eta n$ and suppose G is a robust (ν, τ)-outexpander. If C is any orientation of a cycle on n vertices, then G contains a copy of C.*

2.4.3 Structure

Let $\varepsilon > 0$ and G be a digraph on n vertices. We say that G is *ε-extremal* if there is a partition A, B, S, T of its vertices into sets of sizes a, b, s, t such that $|a - b|, |s - t| \leq 1$ and $e(A \cup S, A \cup T) < \varepsilon n^2$.

The following lemma describes the structure of a graph which satisfies the conditions of Theorem 2.1.2.

Lemma 2.4.4. *Suppose $0 < 1/n \ll \nu \ll \tau, \varepsilon < 1$ and let G be a digraph on n vertices with $\delta^0(G) \geq n/2$. Then G satisfies one of the following:*

(i) *G is ε-extremal;*

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

Proof. Suppose that G is not a robust (ν, τ)-outexpander. Then there is a set $X \subseteq V(G)$ with $\tau n \leq |X| \leq (1-\tau)n$ and $|RN^+_{\nu,G}(X)| < |X| + \nu n$. Define $RN^+ := RN^+_{\nu,G}(X)$. We consider the following cases:

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

Note that any vertex in $\overline{RN^+}$ has fewer than νn inneighbours in X so $e(X, \overline{RN^+}) < \nu n^2$. Together with the fact that $\delta^0(G) \geq n/2$, this implies

$$|X|n/2 \leq e(X, RN^+) + e(X, \overline{RN^+}) \leq |X||RN^+| + \nu n^2 \leq |X|(|RN^+| + \nu n/\tau).$$

So $|RN^+| \geq (1/2 - \nu/\tau)n \geq |X| + \nu n$, which gives a contradiction.

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

For any $v \in V(G)$ we note that $d^-_X(v) \geq \nu n$. Hence $|RN^+| = |G| \geq |X| + \nu n$, a contradiction.

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

Suppose that $|RN^+| < (1/2 - 3\nu)n$. Since $\delta^0(G) \geq n/2$, each vertex in X has more than $3\nu n$ outneighbours in $\overline{RN^+}$. Thus, there is a vertex $v \notin RN^+$ with more than $3\nu n|X|/n > \nu n$ inneighbours in X, which is a contradiction. Therefore,

$$(1/2 - 3\nu)n \leq |RN^+| < |X| + \nu n < (1/2 + 2\nu)n. \tag{2.1}$$

Write $A_0 := X \setminus RN^+$, $B_0 := RN^+ \setminus X$, $S_0 := X \cap RN^+$ and $T_0 := \overline{X} \cap \overline{RN^+}$. Let a_0, b_0, s_0, t_0, respectively, denote their sizes. Note that $|X| = a_0 + s_0$, $|RN^+| = b_0 + s_0$ and $a_0 + b_0 + s_0 + t_0 = n$. It follows from (2.1) and the conditions of Case 3 that

$$(1/2 - \sqrt{\nu})n \leq a_0 + s_0, b_0 + t_0, b_0 + s_0, a_0 + t_0 \leq (1/2 + \sqrt{\nu})n$$

and so $|a_0 - b_0|, |s_0 - t_0| \leq 2\sqrt{\nu}n$. Note that

$$e(A_0 \cup S_0, A_0 \cup T_0) = e(X, \overline{RN^+}) < \nu n^2.$$

By moving at most $\sqrt{\nu}n$ vertices between the sets A_0 and B_0 and $\sqrt{\nu}n$ between the sets S_0 and T_0, we obtain new sets A, B, S, T of sizes a, b, s, t satisfying $|a-b|, |s-t| \leq 1$ and $e(A \cup S, A \cup T) \leq \varepsilon n^2$. So G is ε-extremal. □

2.4.4 Refining the notion of ε-extremality

Let $n \in \mathbb{N}$ and $\varepsilon, \varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \eta_1, \eta_2, \tau$ be positive constants satisfying

$$1/n \ll \varepsilon \ll \varepsilon_1 \ll \varepsilon_2 \ll \eta_1 \ll \tau \ll \varepsilon_3 \ll \varepsilon_4 \ll \eta_2 \ll 1.$$

We now introduce three refinements of ε-extremality. (The constants ε_2 and ε_4 do not appear in these definitions but will be used at a later stage in the proof so we include them here for clarity.) Let G be a digraph on n vertices.

Firstly, we say that G is *ST-extremal* if there is a partition A, B, S, T of $V(G)$ into sets of sizes a, b, s, t such that:

(P1) $a \leq b$, $s \leq t$;

(P2) $\lfloor n/2 \rfloor - \varepsilon_3 n \leq s, t \leq \lceil n/2 \rceil + \varepsilon_3 n$;

(P3) $\delta^0(G[S]), \delta^0(G[T]) \geq \eta_2 n$;

(P4) $d^{\pm}_S(x) \geq n/2 - \varepsilon_3 n$ for all but at most $\varepsilon_3 n$ vertices $x \in S$;

(P5) $d^{\pm}_T(x) \geq n/2 - \varepsilon_3 n$ for all but at most $\varepsilon_3 n$ vertices $x \in T$;

(P6) $a + b \leq \varepsilon_3 n$;

(P7) $d^-_T(x), d^+_S(x) > n/2 - 3\eta_2 n$ and $d^-_S(x), d^+_T(x) \leq 3\eta_2 n$ for all $x \in A$;

(P8) $d^-_S(x), d^+_T(x) > n/2 - 3\eta_2 n$ and $d^-_T(x), d^+_S(x) \leq 3\eta_2 n$ for all $x \in B$.

Secondly, we say that G is *AB-extremal* if there is a partition A, B, S, T of $V(G)$ into sets of sizes a, b, s, t such that:

(Q1) $a \le b$, $s \le t$;

(Q2) $\lfloor n/2 \rfloor - \varepsilon_3 n \le a, b \le \lceil n/2 \rceil + \varepsilon_3 n$;

(Q3) $\delta^0(G[A,B]) \ge n/50$;

(Q4) $d^\pm_B(x) \ge n/2 - \varepsilon_3 n$ for all but at most $\varepsilon_3 n$ vertices $x \in A$;

(Q5) $d^\pm_A(x) \ge n/2 - \varepsilon_3 n$ for all but at most $\varepsilon_3 n$ vertices $x \in B$;

(Q6) $s + t \le \varepsilon_3 n$;

(Q7) $d^-_A(x), d^+_B(x) \ge n/50$ for all $x \in S$;

(Q8) $d^-_B(x), d^+_A(x) \ge n/50$ for all $x \in T$;

(Q9) if $a < b$, $d^\pm_B(x) < n/20$ for all $x \in B$; $d^-_B(x) < n/20$ for all $x \in S$ and $d^+_B(x) < n/20$ for all $x \in T$.

Thirdly, we say that G is *ABST-extremal* if there is a partition A, B, S, T of $V(G)$ into sets of sizes a, b, s, t such that:

(R1) $a \le b$, $s \le t$;

(R2) $a, b, s, t \ge \tau n$;

(R3) $|a - b|, |s - t| \le \varepsilon_1 n$;

(R4) $\delta^0(G[A,B]) \ge \eta_1 n$;

(R5) $d^+_{B\cup S}(x), d^-_{A\cup S}(x) \ge \eta_1 n$ for all $x \in S$;

(R6) $d^+_{A\cup T}(x), d^-_{B\cup T}(x) \ge \eta_1 n$ for all $x \in T$;

(R7) $d^\pm_B(x) \ge b - \varepsilon^{1/3} n$ for all but at most $\varepsilon_1 n$ vertices $x \in A$;

(R8) $d^\pm_A(x) \ge a - \varepsilon^{1/3} n$ for all but at most $\varepsilon_1 n$ vertices $x \in B$;

(R9) $d^+_{B\cup S}(x) \geq b+s-\varepsilon^{1/3}n$ and $d^-_{A\cup S}(x) \geq a+s-\varepsilon^{1/3}n$ for all but at most $\varepsilon_1 n$ vertices $x \in S$;

(R10) $d^+_{A\cup T}(x) \geq a+t-\varepsilon^{1/3}n$ and $d^-_{B\cup T}(x) \geq b+t-\varepsilon^{1/3}n$ for all but at most $\varepsilon_1 n$ vertices $x \in T$.

Proposition 2.4.5. *Suppose*

$$1/n \ll \varepsilon \ll \varepsilon_1 \ll \eta_1 \ll \tau \ll \varepsilon_3 \ll \eta_2 \ll 1$$

and G is an ε-extremal digraph on n vertices with $\delta^0(G) \geq n/2$. Then there is a partition of $V(G)$ into sets A, B, S, T of sizes a, b, s, t satisfying (P2)–(P8), (Q2)–(Q9) *or* (R2)–(R10). *Moreover, if A, B, S, T satisfies* (Q2)–(Q9), *we also have that $a \leq b$.*

Proof. Consider a partition A_0, B_0, S_0, T_0 of $V(G)$ into sets of sizes a_0, b_0, s_0, t_0 such that $|a_0 - b_0|, |s_0 - t_0| \leq 1$ and $e(A_0 \cup S_0, A_0 \cup T_0) < \varepsilon n^2$. Define

$$\begin{aligned}
X_1 &:= \{x \in A_0 \cup S_0 : d^+_{B_0\cup S_0}(x) < n/2 - \sqrt{\varepsilon}n\},\\
X_2 &:= \{x \in A_0 \cup T_0 : d^-_{B_0\cup T_0}(x) < n/2 - \sqrt{\varepsilon}n\},\\
X_3 &:= \{x \in B_0 \cup T_0 : d^+_{A_0\cup T_0}(x) < n/2 - \sqrt{\varepsilon}n\},\\
X_4 &:= \{x \in B_0 \cup S_0 : d^-_{A_0\cup S_0}(x) < n/2 - \sqrt{\varepsilon}n\}
\end{aligned}$$

and let $X := \bigcup_{i=1}^4 X_i$. We now compute an upper bound for $|X|$. Each vertex $x \in X_1$ has $d^+_{A_0\cup T_0}(x) > \sqrt{\varepsilon}n$, so $|X_1| \leq \varepsilon n^2/\sqrt{\varepsilon}n = \sqrt{\varepsilon}n$. Also, each vertex $x \in X_2$ has $d^-_{A_0\cup S_0}(x) > \sqrt{\varepsilon}n$, so $|X_2| \leq \sqrt{\varepsilon}n$. Observe that

$$\begin{aligned}
|A_0 \cup T_0|n/2 - \varepsilon n^2 &\leq e(B_0 \cup T_0, A_0 \cup T_0)\\
&\leq (n/2 - \sqrt{\varepsilon}n)|X_3| + |A_0 \cup T_0|(|B_0 \cup T_0| - |X_3|)
\end{aligned}$$

which gives

$$|X_3|(|A_0 \cup T_0| - n/2 + \sqrt{\varepsilon}n) \leq |A_0 \cup T_0|(|B_0 \cup T_0| - n/2) + \varepsilon n^2 \leq 2\varepsilon n^2.$$

So $|X_3| \leq 2\varepsilon n^2/(\sqrt{\varepsilon}n/2) = 4\sqrt{\varepsilon}n$. Similarly, we find that $|X_4| \leq 4\sqrt{\varepsilon}n$. Therefore, $|X| \leq 10\sqrt{\varepsilon}n$.

Case 1: $a_0, b_0 < 2\tau n$.

Let $Z := X \cup A_0 \cup B_0$. Choose disjoint $Z_1, Z_2 \subseteq Z$ so that $d^{\pm}_{S_0}(x) \geq 2\eta_2 n$ for all $x \in Z_1$ and $d^{\pm}_{T_0}(x) \geq 2\eta_2 n$ for all $x \in Z_2$ and $|Z_1 \cup Z_2|$ is maximal. Let $S := (S_0 \setminus X) \cup Z_1$ and $T := (T_0 \setminus X) \cup Z_2$. The vertices in $Z \setminus (Z_1 \cup Z_2)$ can be partitioned into two sets A and B so that $d^+_S(x), d^-_T(x) \geq n/2 - 3\eta_2 n$ for all $x \in A$ and $d^-_S(x), d^+_T(x) \geq n/2 - 3\eta_2 n$ for all $x \in B$. The partition A, B, S, T satisfies (P2)–(P8).

Case 2: $s_0, t_0 < 2\tau n$.

Partition X into four sets Z_1, Z_2, Z_3, Z_4 so that $d^{\pm}_{B_0}(x) \geq n/5$ for all $x \in Z_1$; $d^{\pm}_{A_0}(x) \geq n/5$ for all $x \in Z_2$; $d^+_{B_0}(x), d^-_{A_0}(x) \geq n/5$ for all $x \in Z_3$ and $d^-_{B_0}(x), d^+_{A_0}(x) \geq n/5$ for all $x \in Z_4$. Then set $A_1 := (A_0 \setminus X) \cup Z_1$, $B_1 := (B_0 \setminus X) \cup Z_2$.

Assume, without loss of generality, that $|A_1| \leq |B_1|$. To ensure that the vertices in B satisfy (Q9), choose disjoint sets $B', B'' \subseteq B_1$ so that $|B' \cup B''|$ is maximal subject to: $|B' \cup B''| \leq |B_1| - |A_1|$, $d^+_{B_1}(x) \geq n/20$ for all $x \in B'$ and $d^-_{B_1}(x) \geq n/20$ for all $x \in B''$. Set $B := B_1 \setminus (B' \cup B'')$, $S_1 := (S_0 \setminus X) \cup Z_3 \cup B'$ and $T_1 := (T_0 \setminus X) \cup Z_4 \cup B''$. To ensure that the vertices in $S \cup T$ satisfy (Q9), choose sets $S' \subseteq S_1, T' \subseteq T_1$ which are maximal subject to: $|S'| + |T'| \leq |B| - |A_1|$, $d^{\pm}_B(x) \geq n/20$ for all $x \in S'$ and $d^{\pm}_B(x) \geq n/20$ for all $x \in T'$. We define $A := A_1 \cup S' \cup T'$, $S := S_1 \setminus S'$ and $T := T_1 \setminus T'$. Then $a \leq b$ and (Q2)–(Q9) hold.

Case 3: $a_0, b_0, s_0, t_0 \geq 2\tau n - 1$.

The case conditions imply $a_0, b_0, s_0, t_0 < n/2 - \tau n$. Then, since $\delta^0(G) \geq n/2$, each vertex must have at least $2\eta_1 n$ inneighbours in at least two of the sets A_0, B_0, S_0, T_0. The same holds when we consider outneighbours instead. So we can partition the vertices in X

into sets Z_1, Z_2, Z_3, Z_4 so that: $d^{\pm}_{B_0}(x) \geq 2\eta_1 n$ for all $x \in Z_1$; $d^{\pm}_{A_0}(x) \geq 2\eta_1 n$ for all $x \in Z_2$; $d^{+}_{B_0 \cup S_0}(x), d^{-}_{A_0 \cup S_0}(x) \geq 2\eta_1 n$ for all $x \in Z_3$ and $d^{+}_{A_0 \cup T_0}(x), d^{-}_{B_0 \cup T_0}(x) \geq 2\eta_1 n$ for all $x \in Z_4$. Let $A := (A_0 \setminus X) \cup Z_1$, $B := (B_0 \setminus X) \cup Z_2$, $S := (S_0 \setminus X) \cup Z_3$ and $T := (T_0 \setminus X) \cup Z_4$. This partition satisfies (R2)–(R10). □

The above result implies that to prove Theorem 2.1.2 for ε-extremal graphs it will suffice to consider only graphs which are ST-extremal, AB-extremal or $ABST$-extremal. Indeed, to see that we may assume that $a \leq b$ and $s \leq t$, suppose that G is ε-extremal. Then G has a partition satisfying (P2)–(P8), (Q2)–(Q9) or (R2)–(R10) by Proposition 2.4.5. Note that relabelling the sets of the partition (A, B, S, T) by (B, A, T, S) if necessary allows us to assume that $a \leq b$. If $s \leq t$, then we are done. If $s > t$, reverse the orientation of every edge in G to obtain the new graph G'. Relabel the sets (A, B, S, T) by (A, B, T, S). Under this new labelling, the graph G' satisfies all of the original properties as well as $a \leq b$ and $s \leq t$. Obtain C' from the cycle C by reversing the orientation of every edge in C. The problem of finding a copy of C in G is equivalent to finding a copy of C' in G'.

2.5 G is ST-extremal

The aim of this section is to prove the following lemma which settles Theorem 2.1.2 in the case when G is ST-extremal.

Lemma 2.5.1. *Suppose that $1/n \ll \varepsilon_3 \ll \varepsilon_4 \ll \eta_2 \ll 1$. Let G be a digraph on n vertices such that $\delta^0(G) \geq n/2$ and G is ST-extremal. If C is any orientation of a cycle on n vertices, then G contains a copy of C.*

We will split the proof of Lemma 2.5.1 into two cases based on how close the cycle C is to being consistently oriented. Recall that $\sigma(C)$ denotes the number of sink vertices in C. Observe that in any oriented cycle, the number of sink vertices is equal to the number of source vertices.

2.5.1 C has many sink vertices, $\sigma(C) \geq \varepsilon_4 n$

The rough strategy in this case is as follows. We would like to embed half of the cycle C into $G[S]$ and half into $G[T]$, making use of the fact that these graphs are nearly complete. At this stage, we also suitably assign the vertices in $A \cup B$ to $G[S]$ or $G[T]$. We will partition C into two disjoint paths, P_S and P_T, each containing at least $\sigma(C)/8$ sink vertices, which will be embedded into $G[S]$ and $G[T]$. The main challenge we will face is finding appropriate edges to connect the two halves of the embedding.

Lemma 2.5.2. *Suppose that $1/n \ll \varepsilon_3 \ll \varepsilon_4 \ll \eta_2 \ll 1$. Let G be a digraph on n vertices with $\delta^0(G) \geq n/2$. Suppose A, B, S, T is a partition of $V(G)$ satisfying* (P1)–(P8). *Let C be an oriented cycle on n vertices with $\sigma(C) \geq \varepsilon_4 n$. Then there exists a partition S^*, T^* of the vertices of G and internally disjoint paths R_1, R_2, P_S, P_T such that $C = (P_S R_1 P_T R_2)$ and the following hold:*

(i) *$S \subseteq S^*$ and $T \subseteq T^*$;*

(ii) *$|P_T| = |T^*|$;*

(iii) *P_S and P_T each contain at least $\varepsilon_4 n/8$ sink vertices;*

(iv) *$|R_i| \leq 3$ and G contains disjoint copies R_i^G of R_i such that R_1^G is an ST-path, R_2^G is a TS-path and all interior vertices of R_i^G lie in S^*.*

In the proof of Lemma 2.5.2 we will need the following proposition.

Proposition 2.5.3. *Suppose that $1/n \ll \varepsilon_3 \ll \varepsilon_4 \ll \eta \ll 1$. Let G be a digraph on n vertices with $\delta^0(G) \geq n/2$. Suppose A, B, S, T is a partition of $V(G)$ satisfying* (P1)–(P8).

(i) *If $a = b \in \{0, 1\}$ then there are two disjoint edges between S and T of any given direction.*

(ii) *If $A = \emptyset$ then there are two disjoint TS-edges.*

(iii) *If $a = 1$ and $b \geq 2$ then there are two disjoint TS-edges.*

(iv) *There are two disjoint edges in* $E(S, T \cup A) \cup E(T, S \cup B)$.

Proof. Let

$$S' := \{x \in S : N^+_A(x), N^-_B(x) = \emptyset\} \text{ and } T' := \{x \in T : N^+_B(x), N^-_A(x) = \emptyset\}.$$

First we prove (i). If $a = b \in \{0, 1\}$ then it follows from (P7), (P8) that $|S'|, |T'| \geq n/4$. Since $s \leq t$, it is either the case that $s \leq (n-1)/2-b$ or $s = t = n/2-b$. If $s \leq (n-1)/2-b$ choose any $x \neq y \in S'$. Both x and y have at least $\lceil n/2 - ((n-1)/2 - b - 1 + b)\rceil = 2$ inneighbours and outneighbours in T, so we find the desired edges. Otherwise $s = t = n/2 - b$ and each vertex in S' must have at least one inneighbour and at least one outneighbour in T and each vertex in T' must have at least one inneighbour and at least one outneighbour in S. It is now easy to check that (i) holds. Indeed, König's theorem gives the two required disjoint edges provided they have the same direction. Using this, it is also easy to find two edges in opposite directions.

We now prove (ii). Suppose that $A = \emptyset$. We have already seen that the result holds when $B = \emptyset$. So assume that $b \geq 1$. Since $s \leq (n-b)/2$, each vertex in S must have at least $b/2+1$ inneighbours in $T \cup B$. Assume for contradiction that there are no two disjoint TS-edges. Then all but at most one vertex in S must have at least $b/2$ inneighbours in B. So $e(B, S) \geq bn/8$ which implies that there is a vertex $v \in B$ with $d^+_S(v) \geq n/8$. But this contradicts (P8). So there must be two disjoint TS-edges.

For (iii), suppose that $a = 1$ and $b \geq 2$. Since $s \leq (n-b-1)/2$, each vertex in S must have at least $(b+1)/2$ inneighbours in $T \cup B$. Assume that there are no two disjoint TS-edges. Then all but at most one vertex in S have at least $(b-1)/2$ inneighbours in B. So $e(B, S) \geq nb/12$ which implies that there is a vertex $v \in B$ with $d^+_S(v) \geq n/12$ which contradicts (P8). Hence (iii) holds.

For (iv), we observe that $\min\{s+b, t+a\} \leq (n-1)/2$ or $s+b = t+a = n/2$. If $s + b \leq (n-1)/2$ then each vertex in S has at least two outneighbours in $T \cup A$, giving the desired edges. A similar argument works if $t + a \leq (n-1)/2$. If $s + b = t + a = n/2$

then each vertex in S has at least one outneighbour in $T \cup A$ and each vertex in T has at least one outneighbour in $S \cup B$. It is easy to see that there must be two disjoint edges in $E(S, T \cup A) \cup E(T, S \cup B)$. □

Proof of Lemma 2.5.2. Observe that C must have a subpath P_1 of length $n/3$ containing at least $\varepsilon_4 n/3$ sink vertices. Let $v \in P_1$ be a sink vertex such that the subpaths $(P_1 v)$ and $(v P_1)$ of P_1 each contain at least $\varepsilon_4 n/7$ sink vertices. Write $C = (v_1 v_2 \dots v_n)$ where $v_1 := v$ and write $k' := n - t$.

Case 1: $a \leq 1$

If $a = b$, set $S^* := S \cup A \cup B$, $T^* := T$, $R_1 := (v_{k'} v_{k'+1})$ and $R_2 := (v_n v_1) = v_n v_1$. By Proposition 2.5.3(i), G contains a pair of disjoint edges between S and T of any given orientation. So we can map $v_n v_1$ to a TS-edge and $(v_{k'} v_{k'+1})$ to an edge between S and T of the correct orientation such that the two edges are disjoint.

Suppose now that $b \geq a + 1$. By Proposition 2.5.3(ii)–(iii), we can find two disjoint TS-edges e_1 and e_2. If $v_{k'}$ is not a source vertex, set $S^* := S \cup A \cup B$, $T^* := T$, $R_1 := (v_{k'-1} v_{k'} v_{k'+1})$ and $R_2 := v_n v_1$. Map $v_n v_1$ to e_1. If $v_{k'+1} v_{k'} \in E(C)$, map R_1 to a path of the form SST which uses e_2. Otherwise, since $v_{k'}$ is not a source vertex, R_1 is a forward path. Using (P8), we find a forward path of the form SBT for R_1^G.

So let us suppose that $v_{k'}$ is a source vertex. Let $b_1 \in B$ and set $S^* := S \cup A \cup B \setminus \{b_1\}$ and $T^* := T \cup \{b_1\}$. Let $R_1 := (v_{k'-1} v_{k'}) = v_{k'} v_{k'-1}$ and $R_2 := v_n v_1$. We know that $v_n v_1, v_{k'} v_{k'-1} \in E(C)$, so we can map these edges to e_1 and e_2.

In each of the above, we define P_S and P_T to be the paths, which are internally disjoint from R_1 and R_2, such that $C = (P_S R_1 P_T R_2)$. Note that (i)–(iv) are satisfied.

Case 2: $a \geq 2$

Apply Proposition 2.5.3(iv) to find two disjoint edges $e_1, e_2 \in E(S, T \cup A) \cup E(T, S \cup B)$. Choose any distinct $x, y \in A \cup B$ such that x and y are disjoint from e_1 and e_2.

First let us suppose that $v_{k'}$ is a sink vertex. If $e_1, e_2 \in E(S, A) \cup E(T, S \cup B)$, set $S^* := S \cup A \cup B$, $T^* := T$, $R_1 := (v_{k'-1} v_{k'} v_{k'+1})$ and $R_2 := (v_n v_1 v_2)$. If $e_1 \in E(T, S \cup B)$, use

(P3) and (P8) to find a path of the form $S(S\cup B)T$ which uses e_1 for R_1^G. If $e_1 \in E(S,A)$, we use (P7) to find a path of the form SAT using e_1 for R_1^G. In the same way, we find a copy R_2^G of R_2. If exactly one of e_i, e_2 say, lies in $E(S,T)$, set $S^* := (S \cup A \cup B) \setminus \{x\}$, $T^* := T \cup \{x\}$, $R_1 := (v_{k'-1}v_{k'}v_{k'+1})$ and $R_2 := (v_1v_2)$. Then v_2v_1 can be mapped to e_2 and we use e_1 to find a copy R_1^G of R_1 as before. If both $e_1, e_2 \in E(S,T)$, set $S^* := (S \cup A \cup B) \setminus \{x,y\}$, $T^* := T \cup \{x,y\}$, $R_1 := (v_{k'-1}v_{k'})$ and $R_2 := (v_1v_2)$. Then map v_2v_1 and $v_{k'-1}v_{k'}$ to the edges e_1 and e_2.

Suppose now that $(v_{k'-1}v_{k'}v_{k'+1})$ is a consistently oriented path. If $e_2 \notin E(S,T)$, let $S^* := S \cup A \cup B$, $T^* := T$, $R_1 := (v_{k'-1}v_{k'}v_{k'+1})$ and $R_2 := (v_nv_1v_2)$ and, if $e_2 \in E(S,T)$, let $S^* := (S \cup A \cup B) \setminus \{x\}$, $T^* := T \cup \{x\}$, $R_1 := (v_{k'-1}v_{k'}v_{k'+1})$ and $R_2 := (v_1v_2)$. Then use the edge e_2 to find a copy R_2^G of R_2 as above. We use (P7) or (P8) to map R_1 to a backward path of the form SAT or a forward path of the form SBT as appropriate.

We let P_S and P_T be paths which are internally disjoint from R_1 and R_2 such that $C = (P_SR_1P_TR_2)$. Then (i)–(iv) are satisfied.

It remains to consider the case when $v_{k'}$ is a source vertex. We now consider the vertex $v_{k'-1}$ instead of $v_{k'}$. Note that C cannot contain two adjacent source vertices, so either $v_{k'-1}$ is a sink vertex or $(v_{k'-2}v_{k'-1}v_{k'})$ is a backward path. We proceed as previously. Note that when we define the path P_T it will have one additional vertex and so we must allocate an additional vertex from $A\cup B$ to T^*, we are able to do this since $a+b > 3$. □

Apply Lemma 2.5.2 to G and C to obtain internally disjoint subpaths R_1, R_2, P_S and P_T of C as well as a partition S^*, T^* of $V(G)$. Let R_i^G be copies of R_i in G satisfying the properties of the lemma. Write R' for the set of interior vertices of the R_i^G. Define $G_S := G[S^* \setminus R']$ and $G_T := G[T^*]$. Let x_T and x_S be the images of the final vertices of R_1 and R_2 and let y_S and y_T be the images of the initial vertices of R_1 and R_2, respectively. Also, let $V_S := S^* \cap (A \cup B)$ and $V_T := T^* \cap (A \cup B)$.

The following proposition allows us to embed copies of P_S and P_T in G_S and G_T. The idea is to greedily find a short path which will contain all of the vertices in V_S and V_T and any vertices of "low degree". We then use that the remaining graph is nearly complete to complete the embedding.

Proposition 2.5.4. *Let G_S, P_S, P_T, x_S, y_S, x_T and y_T be as defined above.*

(i) *There is a copy of P_S in G_S such that the initial vertex of P_S is mapped to x_S and the final vertex is mapped to y_S.*

(ii) *There is a copy of P_T in G_T such that the initial vertex of P_T is mapped to x_T and the final vertex is mapped to y_T.*

Proof. We prove (i), the proof of (ii) is identical. Write $P_S = (u_1 u_2 \ldots u_k)$. An averaging argument shows that there exists a subpath P of P_S of order at most $\varepsilon_4 n$ containing at least $\sqrt{\varepsilon_3} n$ sink vertices.

Let $X := \{x \in S : d^+_S(x) < n/2 - \varepsilon_3 n \text{ or } d^-_S(x) < n/2 - \varepsilon_3 n\}$. By (P4), $|X| \leq \varepsilon_3 n$ and so, using (P3), we see that every vertex $x \in X$ is adjacent to at least $\eta_2 n/2$ vertices in $S \setminus X$. So we can assume that $x_S, y_S \in S \setminus X$ since otherwise we can embed the second and penultimate vertices on P_S to vertices in $S \setminus X$ and consider these vertices instead.

Let u'_1 be the initial vertex of P and u'_k be the final vertex. Define $m_1 := d_{P_S}(u_1, u'_1)+1$ and $m_2 := d_{P_S}(u'_k, u_k) + 1$. Suppose first that $m_1, m_2 > \eta_2^2 n$. We greedily find a copy P^G of P in G_S which covers all vertices in $V_S \cup X$ such that u'_1 and u'_k are mapped to vertices $s_1, s_2 \in S \setminus X$. This is possible since any two vertices in X can be joined by a path of length at most three of any given orientation, by (P3) and (P4), and we can use each vertex in V_S as the image of a sink or source vertex of P. Partition $(V(G_S) \setminus V(P^G)) \cup \{s_1, s_2\}$, arbitrarily, into two sets L_1 and L_2 of size m_1 and m_2 respectively so that $s_1, x_S \in L_1$ and $s_2, y_S \in L_2$. Consider the graphs $G_i := G_S[L_i]$ for $i = 1, 2$. Then (P4) implies that $\delta(G_i) \geq m_i - \varepsilon_3 n - \varepsilon_4 n \geq 7m_i/8$. Applying Proposition 2.4.2(i), we find suitably oriented Hamilton paths from s_1 to x_S in G_1 and s_2 to y_S in G_2 which, when combined with P, form a copy of P_S in G_S (with endvertices x_S and y_S).

It remains to consider the case when $m_1 < \eta_2^2 n$ or $m_2 < \eta_2^2 n$. Suppose that the former holds (the latter is similar). Let P' be the subpath of P_S between u_1 and u'_k. So $P \subseteq P'$. Similarly as before, we first greedily find a copy of P' in G_S which covers all vertices of

$X \cup V_S$ and then extend this to an embedding of P_S. □

Proposition 2.5.4 allows us to find copies of P_S and P_T in G_S and G_T with the desired endvertices. Combining these with R_1^G and R_2^G found in Lemma 2.5.2, we obtain a copy of C in G. This proves Lemma 2.5.1 when $\sigma(C) \geq \varepsilon_4 n$.

2.5.2 C has few sink vertices, $\sigma(C) < \varepsilon_4 n$

Our approach will closely follow the argument when C had many sink vertices. The main difference will be how we cover the exceptional vertices. We will call a consistently oriented subpath of C which has length 20 a *long run*. If C contains few sink vertices, it must contain many of these long runs. So, whereas previously we used sink and source vertices, we will now use long runs to cover the vertices in $A \cup B$.

Proposition 2.5.5. *Suppose that $1/n \ll \varepsilon \ll 1$ and $n/4 \leq k \leq 3n/4$. Let C be an oriented cycle with $\sigma(C) < \varepsilon n$. Then we can write C as $(u_1 u_2 \dots u_n)$ such that there exist:*

(i) *Long runs P_1, P_2 such that P_1 is a forward path and $d_C(P_1, P_2) = k$,*

(ii) *Long runs P_1', P_2', P_3', P_4' such that $d_C(P_i', P_{i+1}') = \lfloor n/4 \rfloor$ for $i = 1, 2, 3$.*

Proof. Let P be a subpath of C of length $n/8$. Let $\mathcal{Q}$ be a consistent collection of vertex disjoint long runs in P of maximum size. Then $|\mathcal{Q}| \geq 2\varepsilon n$, with room to spare. We can write C as $(u_1 u_2 \dots u_n)$ so that the long runs in $\mathcal{Q}$ are forward paths.

Suppose that (i) does not hold. For each $Q_i \in \mathcal{Q}$, let Q_i' be the path of length 20 such that $d_C(Q_i, Q_i') = k$. Since Q_i' is not a long run, Q_i' must contain at least one sink or source vertex. The paths Q_i' are disjoint so, in total, C must contain at least $|\mathcal{Q}|/2 \geq \varepsilon n > \sigma(C)$ sink vertices, a contradiction. Hence (i) holds.

We call a collection of four disjoint long runs P_1, P_2, P_3, P_4 *good* if $P_1 \in \mathcal{Q}$ and $d_C(P_i, P_{i+1}) = \lfloor n/4 \rfloor$ for all $i = 1, 2, 3$. Suppose C does not contain a good collection of long runs. In particular, this means that each long run in $\mathcal{Q}$ does not lie in a good collection. For each path $Q_i \in \mathcal{Q}$, let $Q_{i,1}, Q_{i,2}, Q_{i,3}$ be subpaths of C of length 20 such that $d_C(Q_i, Q_{i,j}) = j\lfloor n/4 \rfloor$. Since $\{Q_i, Q_{i,1}, Q_{i,2}, Q_{i,3}\}$ does not form a good collection, at least one of the $Q_{i,j}$ must contain a sink or source vertex. The paths $Q_{i,j}$ where $Q_i \in \mathcal{Q}$ and $j = 1, 2, 3$ are disjoint so, in total, C must contain at least $|\mathcal{Q}|/2 \geq \varepsilon n > \sigma(C)$ sink vertices, which is a contradiction. This proves (ii). □

The following proposition finds a collection of edges oriented in an atypical direction for an ε-extremal graph. We will use these edges to find consistently oriented S- and T-paths covering all of the vertices in $A \cup B$. This proposition will be used again in Section 2.7.1, where it allows us to correct an imbalance in the sizes of A and B.

Proposition 2.5.6. *Let G be a digraph on n vertices with $\delta^0(G) \geq n/2$. Let $d \geq 0$ and suppose A, B, S, T is a partition of $V(G)$ into sets of size a, b, s, t with $t \geq s \geq d + 2$ and $b = a + d$. Then G contains a collection M of $d + 1$ edges in $E(T, S \cup B) \cup E(B, S)$ satisfying the following. The endvertices of M outside B are distinct and each vertex in B is the endvertex of at most one TB-edge and at most one BS-edge in M. Moreover, if $e(T, S) > 0$, then M contains a TS-edge.*

Proof. Let $k := t - s$. We define a bipartite graph G' with vertex classes $S' := S \cup B$ and $T' := T \cup B$ together with all edges xy such that $x \in S', y \in T'$ and $yx \in E(T, S \cup B) \cup E(B, S)$. We claim that G' has a matching of size $d + 2$. To prove the claim, suppose that G' has a vertex cover X of size $|X| < d + 2$. Then $|X \cap S'| < (d - k)/2 + 1$ or $|X \cap T'| < (d + k)/2 + 1$. Suppose that the former holds and consider any vertex $t_1 \in T \setminus X$. Since $\delta^+(G) \geq n/2$ and $a + t = (n - d + k)/2$, t_1 has at least $(d - k)/2 + 1$ outneighbours in S'. But these vertices cannot all be covered by X. So we must have that $|X \cap T'| < (d + k)/2 + 1$. Consider any vertex $s_1 \in S \setminus X$. Now $\delta^-(G) \geq n/2$ and $a + s = (n - d - k)/2$, so s_1 must have at least $(d + k)/2 + 1$ inneighbours in T'. But not all of these vertices can be covered by X. Hence, any vertex cover of G' must have size at least $d + 2$ and so König's theorem implies that G' has a matching of size $d + 2$.

If $e(T, S) > 0$, either the matching contains a TS-edge, or we can choose any TS-edge

e and at least d of the edges in the matching will be disjoint from e. This corresponds to a set of $d+1$ edges in $E(T, S \cup B) \cup E(B, S)$ in G with the required properties. $\square$

We define a *good path system* $\mathcal{P}$ to be a collection of disjoint S- and T-paths such that each path $P \in \mathcal{P}$ is consistently oriented, has length at most six and covers at least one vertex in $A \cup B$. Each good path system $\mathcal{P}$ gives rise to a modified partition $A_{\mathcal{P}}, B_{\mathcal{P}}, S_{\mathcal{P}}, T_{\mathcal{P}}$ of the vertices of G (we allow $A_{\mathcal{P}}$, $B_{\mathcal{P}}$ to be empty) as follows. Let $\mathrm{Int}_S(\mathcal{P})$ be the set of all interior vertices on the S-paths in $\mathcal{P}$ and $\mathrm{Int}_T(\mathcal{P})$ be the set of all interior vertices on the T-paths. We set $A_{\mathcal{P}} := A \setminus V(\mathcal{P})$, $B_{\mathcal{P}} := B \setminus V(\mathcal{P})$, $S_{\mathcal{P}} := (S \cup \mathrm{Int}_S(\mathcal{P})) \setminus \mathrm{Int}_T(\mathcal{P})$ and $T_{\mathcal{P}} := (T \cup \mathrm{Int}_T(\mathcal{P})) \setminus \mathrm{Int}_S(\mathcal{P})$ and say that $A_{\mathcal{P}}, B_{\mathcal{P}}, S_{\mathcal{P}}, T_{\mathcal{P}}$ is the *$\mathcal{P}$-partition* of $V(G)$.

Lemma 2.5.7. *Suppose that $1/n \ll \varepsilon_3 \ll \varepsilon_4 \ll \eta_2 \ll 1$. Let G be a digraph on n vertices with $\delta^0(G) \geq n/2$. Suppose A, B, S, T is a partition of $V(G)$ satisfying* (P1)–(P8). *Let C be a cycle on n vertices with $\sigma(C) < \varepsilon_4 n$. Then there exists t^* such that one of the following holds:*

- *There exist internally disjoint paths P_S, P_T, R_1, R_2 such that:*
  - (i) $C = (P_S R_1 P_T R_2)$;
  - (ii) $|P_T| = t^*$;
  - (iii) *R_1 and R_2 are paths of length two and G contains disjoint copies R_i^G of R_i whose interior vertices lie in $V(G) \setminus T$. Moreover, R_1^G is an ST-path and R_2^G is a TS-path.*
- *There exist internally disjoint paths $P_S, P_S', P_T, P_T', R_1, R_2, R_3, R_4$ such that:*
  - (i) $C = (P_S R_1 P_T R_2 P_S' R_3 P_T' R_4)$;
  - (ii) $|P_T| + |P_T'| = t^*$ *and* $|P_S|, |P_S'|, |P_T|, |P_T'| \geq n/8$;
  - (iii) *R_1, R_2, R_3, R_4 are paths of length two and G contains disjoint copies R_i^G of R_i whose interior vertices lie in $V(G) \setminus T$. Moreover, R_1^G and R_3^G are ST-paths and R_2^G and R_4^G are TS-paths.*

Furthermore, G has a good path system $\mathcal{P}$ such that the paths in $\mathcal{P}$ are disjoint from each R_i^G, $\mathcal{P}$ covers $(A \cup B) \setminus \bigcup V(R_i^G)$ and the $\mathcal{P}$-partition $A_\mathcal{P}, B_\mathcal{P}, S_\mathcal{P}, T_\mathcal{P}$ of $V(G)$ satisfies $|T_\mathcal{P}| = t^*$.

Proof. Let $d := b - a$ and $k := t - s$.

We first obtain a good path system $\mathcal{P}_0$ covering $A \cup B$ as follows. Apply Proposition 2.5.6 to obtain a collection M_0 of $d + 1$ edges as described in the proposition. Let $M \subseteq M_0$ of size d such that M contains a TS-edge if $d \geq 1$ and $e(T, S) > 0$. We use each edge $e \in M$ together with properties (P3), (P5) and (P8) to cover one vertex in B by a consistently oriented path of length at most six as follows. If $e \in E(T, B)$ and e is disjoint from all other edges in M, find a consistently oriented path of the form TBT using e. If $e \in E(B, S)$ and e is disjoint from all other edges in M, find a consistently oriented path of the form SBS using e. If $e \in E(T, S)$, we note that (P3), (P5) and (P8) allows us to find a consistently oriented path of length three between any vertex in B and any vertex in T. So we can find a consistently oriented path of the form $SB(T)^3S$ which uses e. Finally, if $e \in E(T, B)$ and shares an endvertex with another edge $e' \in M \cap E(B, S)$ we find a consistently oriented path of the form $SB(T)^3BS$ using e and e'. This path uses two edges in M but covers two vertices in B. Since we have many choices for each such path, we can choose them to be disjoint, so M allows us to find a good path system $\mathcal{P}_1$ covering d vertices in B.

Label the vertices in A by $a_1, a_2, \dots, a_a$ and the remaining vertices in B by $b_1, b_2, \dots, b_a$. We now use (P6)–(P8) to find a consistently oriented S- or T-path L_i covering each pair a_i, b_i. If $1 \leq i \leq \lceil (4a + k)/8 \rceil$, cover the pair a_i, b_i by a path of the form $SBTAS$. If $\lceil (4a+k)/8 \rceil < i \leq a$ cover the pair a_i, b_i by a path of the form $TASBT$. Let $\mathcal{P}_2 := \bigcup_{i=1}^a L_i$.

We are able to choose all of these paths so that they are disjoint and thus obtain a good path system $\mathcal{P}_0 := \mathcal{P}_1 \cup \mathcal{P}_2$ covering $A \cup B$. Let $A_{\mathcal{P}_0}, B_{\mathcal{P}_0}, S_{\mathcal{P}_0}, T_{\mathcal{P}_0}$ be the $\mathcal{P}_0$-partition of $V(G)$ and let $t' := |T_{\mathcal{P}_0}|$, $s' := |S_{\mathcal{P}_0}|$.

By Proposition 2.5.5(i), we can enumerate the vertices of C so that there are long runs P_1, P_2 such that P_1 is a forward path and $d_C(P_1, P_2) = t'$. We will find consistently oriented ST- and TS-paths for R_1^G and R_2^G which depend on the orientation of P_2. The paths R_1 and R_2 will be consistently oriented subpaths of P_1 and P_2 respectively, whose position will be chosen later.

Case 1: $b \geq a + 2$.

Suppose first that P_2 is a backward path. If $\mathcal{P}_1$ contains a path of the form $SB(T)^3BS$, let b_0 and b_0' be the two vertices in B on this path. Otherwise, let b_0 and b_0' be arbitrary vertices in B which are covered by $\mathcal{P}_1$. Use (P8) to find a forward path for R_1^G which is of the form $S\{b_0\}T$. We also find a backward path of the form $T\{b_0'\}S$ for R_2^G. We choose the paths R_1^G and R_2^G to be disjoint from all paths in $\mathcal{P}_0$ which do not contain b_0 or b_0'.

Suppose now that P_2 is a forward path. If $a \geq 1$, consider the path $L_1 \in \mathcal{P}_2$ covering $a_1 \in A$ and $b_1 \in B$. Find forward paths of the form $S\{b_1\}T$ for R_1^G and $T\{a_1\}S$ for R_2^G, using (P7) and (P8), which are disjoint from all paths in $\mathcal{P}_0 \setminus \{L_1\}$. Finally, we consider the case when $a = 0$. Recall that $e(T, S) > 0$ by Proposition 2.5.3(ii) and so M contains a TS-edge. Hence there is a path P' in $\mathcal{P}_1$ of the form $SB(T)^3S$, covering a vertex $b_0 \in B$ and an edge $t_1s_1 \in E(T, S)$, say. We use (P3) and (P8) to find forward paths of the form $S\{b_0\}T$ for R_1^G and $\{t_1\}\{s_1\}S$ for R_2^G which are disjoint from all paths in $\mathcal{P}_0 \setminus \{P'\}$.

Obtain the good path system $\mathcal{P}$ from $\mathcal{P}_0$ by removing all paths meeting R_1^G or R_2^G. Let $A_\mathcal{P}, B_\mathcal{P}, S_\mathcal{P}, T_\mathcal{P}$ be the $\mathcal{P}$-partition of $V(G)$ and $t^* := |T_\mathcal{P}|$. The only vertices which could have moved to obtain $T_\mathcal{P}$ from $T_{\mathcal{P}_0}$ are interior vertices on the paths in $\mathcal{P}_0 \setminus \mathcal{P}$, so $|t^* - t'| \leq 2 \cdot 5 = 10$. Thus we can choose R_1 and R_2 to be subpaths of length two of P_1 and P_2 so that $|P_T| = t^*$, where P_S and P_T are defined by $C = (P_S R_1 P_T R_2)$.

Case 2: $b \leq a + 1$.

Case 2.1: $a \leq 1$.

If $a = b$, by Proposition 2.5.3(i) we can find disjoint $e_1, e_2 \in E(S, T)$ and disjoint $e_3 \in E(S, T)$, $e_4 \in E(T, S)$. Note that $\mathcal{P}_0 = \mathcal{P}_2$, since $a = b$, so we may assume that all paths in $\mathcal{P}_0$ are disjoint from e_1, e_2, e_3, e_4. If P_2 is a forward path, find a forward path of the form SST for R_1^G using e_3 and a forward path of the form TSS for R_2^G using e_4. If P_2 is a backward path, find a forward path of the form SST for R_1^G using e_1 and a backward path of the form TSS for R_2^G using e_2. In both cases, we choose R_1^G and R_2^G to be disjoint from all paths in $\mathcal{P}_0$.

If $b = a + 1$, note that there exist $e_1 \in E(S, T)$ and $e_2 \in E(T, S)$. (To see this, use that $\delta^0(G) \geq n/2$ and the fact that (P7) and (P8) imply that $|\{x \in S : N_A^+(x), N_B^-(x) = \emptyset\}| \geq n/4$.) We may assume that all paths in $\mathcal{P}_2$ are disjoint from e_1, e_2. Let $b_0 \in B$ be the vertex covered by the single path in $\mathcal{P}_1$. Find a forward path of the form $S\{b_0\}T$ for R_1^G, using (P8). Find a consistently oriented path of the form TSS for R_2^G which uses e_1 if P_2 is a backward path and e_2 if P_2 is a forward path. Choose the paths R_1^G and R_2^G to be disjoint from the paths in $\mathcal{P}_0 \setminus \mathcal{P}_1 = \mathcal{P}_2$.

In both cases, we obtain the good path system $\mathcal{P}$ from $\mathcal{P}_0$ by removing at most one path which meets R_1^G or R_2^G. Let $A_\mathcal{P}, B_\mathcal{P}, S_\mathcal{P}, T_\mathcal{P}$ be the $\mathcal{P}$-partition of $V(G)$ and let $t^* := |T_\mathcal{P}|$. The only vertices which could have moved to obtain $T_\mathcal{P}$ from $T_{\mathcal{P}_0}$ are interior vertices on the path in $\mathcal{P}_0 \setminus \mathcal{P}$ if $\mathcal{P}_0 \neq \mathcal{P}$, so $|t^* - t'| \leq 5$. So we can choose subpaths R_i of P_i so that $|P_T| = t^*$, where P_S and P_T are defined by $C = (P_S R_1 P_T R_2)$.

Case 2.2: $2 \leq a \leq k$.

If P_2 is a forward path, consider $a_1 \in A$ and $b_1 \in B$ which were covered by the path $L_1 \in \mathcal{P}_0$. Use (P7) and (P8) to find forward paths, disjoint from all paths in $\mathcal{P}_0 \setminus \{L_1\}$, of the form $S\{b_1\}T$ and $T\{a_1\}S$ for R_1^G and R_2^G respectively.

Suppose now that P_2 is a backward path. We claim that G contains $2 - d$ disjoint ST-edges. Indeed, suppose not. Then $d_T^+(x) \leq 1 - d$ for all but at most one vertex in S. Note that $b + s = (n - k + d)/2$, so $d_{A \cup T}^+(x) \geq (k - d)/2 + 1$ for all $x \in S$. So

$$e(S, A) \geq (s - 1)((k - d)/2 + 1 - (1 - d)) = (s - 1)(k + d)/2 \geq nk/8 \geq na/8.$$

Hence, there is a vertex $x \in A$ with $d_S^-(x) \geq n/8$, contradicting (P7). Let $E = \{e_i : 1 \leq i \leq 2 - d\}$ be a set of $2 - d$ disjoint ST-edges. We may assume that $\mathcal{P}_2$ is disjoint from E.

If $a = b$, use (P3) to find a forward path of the form SST using e_1 for R_1^G and a backward path of the form TSS using e_2 for R_2. If $b = a + 1$, let $b_0 \in B$ be the vertex covered by the single path in $\mathcal{P}_1$. Use (P3) and (P8) to find a forward path of the form $S\{b_0\}T$ for R_1^G and a backward path of the form TSS using e_1 for R_2^G. We choose the paths R_1^G and R_2^G to be disjoint from all paths in $\mathcal{P}_2$.

In both cases, we obtain the good path system $\mathcal{P}$ from $\mathcal{P}_0$ by removing at most one path which meets R_1^G or R_2^G. Let $A_{\mathcal{P}}, B_{\mathcal{P}}, S_{\mathcal{P}}, T_{\mathcal{P}}$ be the $\mathcal{P}$-partition of $V(G)$ and $t^* := |T_{\mathcal{P}}|$. The only vertices which could have moved to obtain $T_{\mathcal{P}}$ from $T_{\mathcal{P}_0}$ are interior vertices on the path in $\mathcal{P}_0 \setminus \mathcal{P}$ if $\mathcal{P}_0 \neq \mathcal{P}$, so $|t^* - t'| \leq 5$. Thus we can choose R_1 and R_2 to be subpaths of length two of P_1 and P_2 so that $|P_T| = t^*$, where P_S and P_T are defined by $C = (P_S R_1 P_T R_2)$.

Case 2.3: $a \geq 2, k$.

We note that

$$\begin{aligned} t' - s' &= |(T \cup \mathrm{Int}_T(\mathcal{P}_0)) \setminus \mathrm{Int}_S(\mathcal{P}_0)| - |(S \cup \mathrm{Int}_S(\mathcal{P}_0)) \setminus \mathrm{Int}_T(\mathcal{P}_0)| \\ &= |(T \cup \mathrm{Int}_T(\mathcal{P}_2)) \setminus \mathrm{Int}_S(\mathcal{P}_2)| - |(S \cup \mathrm{Int}_S(\mathcal{P}_2)) \setminus \mathrm{Int}_T(\mathcal{P}_2)| + c \\ &= (t + 3a - 4\lceil (4a+k)/8 \rceil) - (s + 4\lceil (4a+k)/8 \rceil - a) + c \\ &= 4a + k - 8\lceil (4a+k)/8 \rceil + c \end{aligned}$$

where $-7 \leq c \leq 1$ is a constant representing the contribution of interior vertices on the path in $\mathcal{P}_1$ if $b = a + 1$ and $c = 0$ if $b = a$. In particular, this implies that $|t' - s'| \leq 15$ and

$$(n - 15)/2 \leq s', t' \leq (n + 15)/2.$$

Apply Proposition 2.5.5(ii) to find long runs P_1', P_2', P_3', P_4' such that $d_C(P_i', P_{i+1}') = \lfloor n/4 \rfloor$ for $i = 1, 2, 3$. Let x_i be the initial vertex of each P_i'. If $\{P_i', P_{i+2}'\}$ is consistent for some $i \in \{1, 2\}$, consider $a_1 \in A$, $b_1 \in B$ which which were covered by the path $L_1 \in \mathcal{P}_0$. If P_i', P_{i+2}' are both forward paths, let R_1^G and R_2^G be forward paths of the form $S\{b_1\}T$ and $T\{a_1\}S$ respectively. If P_i', P_{i+2}' are both backward paths, let R_1^G and R_2^G be backward paths of the form $S\{a_1\}T$ and $T\{b_1\}S$ respectively. Choose the paths R_1^G and R_2^G to be disjoint from the paths in $\mathcal{P} := \mathcal{P}_0 \setminus \{L_1\}$. Let $A_\mathcal{P}, B_\mathcal{P}, S_\mathcal{P}, T_\mathcal{P}$ be the $\mathcal{P}$-partition of $V(G)$ and let $t^* = |T_\mathcal{P}|$. The only vertices which could have been added or removed to obtain $T_\mathcal{P}$ from $T_{\mathcal{P}_0}$ are interior vertices on L_1 so $(n-15)/2 - 3 \leq t^* \leq (n+15)/2 + 3$. Then we can choose R_1 and R_2 to be subpaths of length two of P'_i and P'_{i+2} so that $|P_T| = t^*$, where P_S, P_T are defined so that $C = (P_S R_1 P_T R_2)$.

So let us assume that $\{P'_i, P'_{i+2}\}$ is not consistent for $i = 1, 2$. We may assume that the paths P'_1 and P'_4 are both forward paths, by relabelling if necessary, and we illustrate the situation in Figure 2.3.

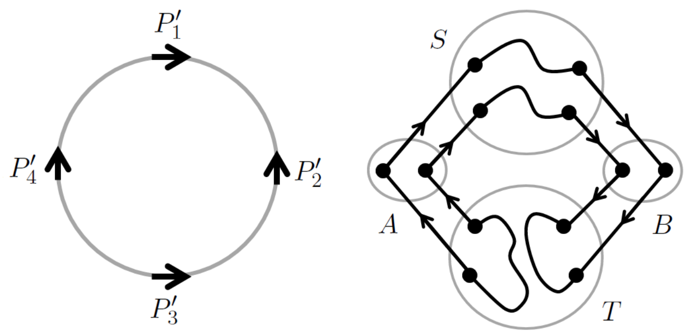

Figure 2.3: A good collection of long runs.

Consider the vertices $a_i \in A$ and $b_i \in B$ covered by the paths $L_i \in \mathcal{P}_0$ for $i = 1, 2$. Let $\mathcal{P} := \mathcal{P}_0 \setminus \{L_1, L_2\}$ and let $A_\mathcal{P}, B_\mathcal{P}, S_\mathcal{P}, T_\mathcal{P}$ be the $\mathcal{P}$-partition of $V(G)$. Let $t^* := |T_\mathcal{P}|$. The only vertices which can have been added or removed to obtain $T_\mathcal{P}$ from $T_{\mathcal{P}_0}$ are interior vertices on the paths L_1 and L_2, so $(n-15)/2 - 6 \leq t^* \leq (n+15)/2 + 6$. Find a forward path of the form $S\{b_1\}T$ for R_1^G. Then find backward paths of the form $T\{b_2\}S$ and $S\{a_1\}T$ for R_2^G and R_3^G respectively. Finally, find a forward path of the form $T\{a_2\}S$ for R_4^G. We can choose the paths R_i^G to be disjoint from all paths in $\mathcal{P}$. Since P'_1 and P'_2 are of length 20 we are able to find subpaths R_1, R_2, R_3, R_4 of P'_1, P'_2, P'_3, P'_4 so that $|P_T| + |P'_T| = t^*$, where P_S, P'_S, P_T, P'_T are defined so that $C = (P_S R_1 P_T R_2 P'_S R_3 P'_T R_4)$.

□

In order to prove Lemma 2.5.1 in the case when $\sigma(C) < \varepsilon_4 n$, we first apply Lemma 2.5.7 to G. We now proceed similarly as in the case when C has many sink vertices (see

Proposition 2.5.4) and so we only provide a sketch of the argument. We first observe that any subpath of the cycle of length $100\varepsilon_4 n$ must contain at least

$$\lfloor 100\varepsilon_4 n/21 \rfloor - 2\varepsilon_4 n > 2\varepsilon_3 n \geq a + b \geq |\mathcal{P}| \tag{2.2}$$

disjoint long runs. Let s_1 be the image of the initial vertex of P_S. Let P_S^* be the subpath of P_S formed by the first $100\varepsilon_4 n$ edges of P_S. We can cover all S-paths in $\mathcal{P}$ and all vertices $x \in S$ which satisfy $d_S^+(x) < n/2 - \varepsilon_3 n$ or $d_S^-(x) < n/2 - \varepsilon_3 n$ greedily by a path in G starting from s_1 which is isomorphic to P_S^*. Note that (2.2) ensures that P_S^* contains $|\mathcal{P}|$ disjoint long runs. So we can map the S-paths in $\mathcal{P}$ to subpaths of these long runs. Let P_S'' be the path formed by removing from P_S all edges in P_S^*.

If Lemma 2.5.7(i) holds and thus P_S is the only path to be embedded in $G[S]$, we apply Proposition 2.4.2(i) to find a copy of P_S'' in $G[S]$, with the desired endvertices. If Lemma 2.5.7(ii) holds, we must find copies of both P_S and P_S' in $G[S]$. So we split the graph into two subgraphs of the appropriate size before applying Proposition 2.4.2(i) to each. We do the same to find copies of P_T (or P_T and P_T') in $G[T]$. Thus, we obtain a copy of C in G. This completes the proof of Lemma 2.5.1.

2.6 G is AB-extremal

The aim of this section is to prove the following lemma which shows that Theorem 2.1.2 is satisfied when G is AB-extremal. Recall that an AB-extremal graph closely resembles a complete bipartite graph. We will proceed as follows. First we will find a short path which covers all of the exceptional vertices (the vertices in $S \cup T$). It is important that this path leaves a balanced number of vertices uncovered in A and B. We will then apply Proposition 2.4.2 to the remaining, almost complete, balanced bipartite graph to embed the remainder of the cycle.

Lemma 2.6.1. *Suppose that $1/n \ll \varepsilon_3 \ll 1$. Let G be a digraph on n vertices with*

$\delta^0(G) \geq n/2$ and assume that G is AB-extremal. If C is any orientation of a cycle on n vertices which is not antidirected, then G contains a copy of C.

If $b > a$, the next lemma implies that $E(B \cup T, B)$ contains a matching of size $b-a+2$. We can use $b-a$ of these edges to pass between vertices in B whilst avoiding A allowing us to correct the imbalance in the sizes of A and B.

Proposition 2.6.2. *Suppose $1/n \ll \varepsilon_3 \ll 1$. Let G be a digraph on n vertices with $\delta^0(G) \geq n/2$. Suppose A, B, S, T is a partition of $V(G)$ satisfying (Q1)–(Q9) and $b = a+d$ for some $d > 0$. Then there is a matching of size $d+2$ in $E(B \cup T, B)$.*

Proof. Consider a maximal matching M in $E(B \cup T, B)$ and suppose that $|M| \leq d+1$. Since $a + s \leq (n-d)/2$, each vertex in B has at least $d/2$ inneighbours in $B \cup T$. In particular, since M was maximal, each vertex in $B \setminus V(M)$ has at least $d/2$ inneighbours in $V(M)$. Then there is a $v \in V(M) \subseteq B \cup T$ with

$$d^+_B(v) \geq \frac{(b-2|M|)}{2|M|}\frac{d}{2} \geq \frac{n}{20},$$

contradicting (Q9). Therefore $|M| \geq d+2$. □

We say that P is an *exceptional cover of* G if $P \subseteq G$ is a copy of a subpath of C and

(EC1) P covers $S \cup T$;

(EC2) both endvertices of P are in A;

(EC3) $|A \setminus V(P)| + 1 = |B \setminus V(P)|$.

We will use the following notation when describing the form of a path. If $X, Y \in \{A, B\}$ then we write $X * Y$ for any path which alternates between A and B whose initial vertex lies in X and final vertex lies in Y. For example, $A * A(ST)^2$ indicates any path of the form $ABAB \dots ASTST$.

Suppose P is of the form $Z_1 Z_2 \dots Z_m$, where $Z_i \in \{A, B, S, T\}$. Let $Z_{i_1}, Z_{i_2}, \dots, Z_{i_j}$ be the appearances of A and B, where $i_j < i_{j+1}$. If $Z_{i_j} = A = Z_{i_{j+1}}$, we say that $Z_{i_{j+1}}$ is a *repeated* A. We define a *repeated* B similarly. Let $\mathrm{rep}(A)$ and $\mathrm{rep}(B)$ be the numbers of repeated As and repeated Bs, respectively. Suppose that P has both endvertices in A and P uses $\ell + \mathrm{rep}(B)$ vertices from B. Then P will use $\ell + \mathrm{rep}(A) + 1$ vertices from A (we add one because both endvertices of P lie in A). So we have that

$$|B \setminus V(P)| - |A \setminus V(P)| = b - a - \mathrm{rep}(B) + \mathrm{rep}(A) + 1. \tag{2.3}$$

Given a set of edges $M \subseteq E(G)$ we define the graph $G_M \subseteq G$ whose vertex set is $V(G)$ and whose edge set is $E(A, B \cup S) \cup E(B, A \cup T) \cup E(T, A) \cup E(S, B) \cup M \subseteq E(G)$. Informally, in addition to the edges of M, G_M has edges between two vertex classes when the bipartite graph they induce in G is dense.

We will again split our argument into two cases depending on the number of sink vertices in C.

2.6.1 Finding an exceptional cover when C has few sink vertices, $\sigma(C) < \varepsilon_4 n$

It is relatively easy to find an exceptional cover when C has few sink vertices by observing that C must contain many disjoint consistently oriented paths of length three. We can use these consistently oriented paths to cover the vertices in $S \cup T$ by forward paths of the form ASB or BTA, for example.

Proposition 2.6.3. *Suppose* $1/n \ll \varepsilon_3 \ll \varepsilon_4 \ll 1$*. Let* G *be a digraph on* n *vertices with* $\delta^0(G) \geq n/2$*. Suppose* A, B, S, T *is a partition of* $V(G)$ *satisfying* (Q1)–(Q9)*. If* $\sigma(C) < \varepsilon_4 n$*, then there is an exceptional cover of* G *of length at most* $21\varepsilon_4 n$*.*

Proof. Let $d := b - a$. Let P be any subpath of C of length $20\varepsilon_4 n$. Let $\mathcal{Q}$ be a maximum consistent collection of disjoint paths of length three in P, such that $d_C(Q, Q') \geq 7$ for all distinct $Q, Q' \in \mathcal{Q}$. Then

$$|\mathcal{Q}| \geq (\lfloor 20\varepsilon_4 n/7 \rfloor - 2\varepsilon_4 n)/2 > 4\varepsilon_3 n > d + s + t.$$

If necessary, reverse the order of all vertices in C so that the paths in $\mathcal{Q}$ are forward paths. Apply Proposition 2.6.2 to find a matching $M \subseteq E(B \cup T, B)$ of size d and write $M = \{e_1, \ldots, e_m, f_{m+1}, \ldots, f_d\}$, where $e_i \in E(B)$ and $f_i \in E(T, B)$. Map the initial vertex of P to any vertex in A. We will greedily find a copy of P in G_M which covers M and $S \cup T$ as follows.

Note that, by (Q8), we can cover each edge $f_i \in M$ by a forward path of the form BTB. By (Q7), each of the vertices in S can be covered by a forward path of the form ASB. Similarly, (Q8) allows us to find a forward path of the form BTA covering each vertex in T. Moreover, note that (Q2)–(Q5) allow us to find a path of length three of any orientation between any pair of vertices $x \in A$ and $y \in B$ using only edges from $E(A, B) \cup E(B, A)$. So we can find a copy of P which covers every edge in M (first the e_i and then the f_i) and every vertex in $(S \cup T) \setminus V(M)$ by a copy of a path in $\mathcal{Q}$ and which has the form

$$(A * BB)^m (A * BTB)^{d-m} (A * ASB)^s (A * BT)^{t-d+m} A * X,$$

where $X \in \{A, B\}$. We may assume that $X = A$ by extending the path P by one vertex if necessary. Let P^G denote this copy of P in G.

Now (EC1) and (EC2) hold. It remains to check (EC3). Observe that P^G contains no repeated As and exactly d repeated Bs, these occur in the subpath of P^G of the form $(A * BB)^m (A * BTB)^{d-m}$. By (2.3), we see that

$$|B \setminus V(P^G)| - |A \setminus V(P^G)| = 1,$$

so (EC3) is satisfied. Hence P^G forms an exceptional cover. □

2.6.2 Finding an exceptional cover when C has many sink vertices, $\sigma(C) \geq \varepsilon_4 n$

When C is far from being consistently oriented, we use sink and source vertices to cover the vertices in $S \cup T$. A natural approach would be to try to cover the vertices in $S \cup T$ by paths of the form ASA and BTB whose central vertex is a sink or by paths of the form ATA and BSB whose central vertex is a source. In essence, this is what we will do, but there are some technical issues we will need to address. The most obvious is that each time we cover a vertex in S or T by a path of one of the above forms, we will introduce a repeated A or a repeated B, so we will need to cover the exceptional vertices in a "balanced" way.

Let P be a subpath of C and let m be the number of sink vertices in P. Suppose that P_1, P_2, P_3 is a partition of P into internally disjoint paths such that $P = (P_1P_2P_3)$. We say that P_1, P_2, P_3 is a *useful tripartition of* P if there exist $\mathcal{Q}_i \subseteq V(P_i)$ such that:

- P_1 and P_2 have even length;
- $|\mathcal{Q}_i| \geq \lfloor m/12 \rfloor$ for $i = 1, 2, 3$;
- all vertices in $\mathcal{Q}_1 \cup \mathcal{Q}_3$ are sink vertices and are an even distance apart;
- all vertices in $\mathcal{Q}_2$ are source vertices and are an even distance apart.

Note that a useful tripartition always exists. We say that $\mathcal{Q}_1, \mathcal{Q}_2, \mathcal{Q}_3$ are *sink/source/sink sets* for the tripartition P_1, P_2, P_3. We say that a subpath $L \subseteq P_2$ is a *link* if L has even length and, if, writing x for the initial vertex and y for the final vertex of L, the paths (P_2x) and (yP_2) each contain at least $|\mathcal{Q}_2|/3$ elements of $\mathcal{Q}_2$.

Proposition 2.6.4. *Let $1/n \ll \varepsilon \ll \eta \ll \tau \leq 1$. Let G be a digraph on n vertices and let A, B, S, T be a partition of $V(G)$. Let S_A, S_B be disjoint subsets of S and T_A, T_B be disjoint subsets of T. Let $a := |A|$, $b := |B|$, $s_A := |S_A|$, $s_B := |S_B|$, $t_A := |T_A|$, $t_B := |T_B|$ and let $a_1 \in A$. Suppose that:*

(i) *$a, b \geq \tau n$;*

(ii) *$s_A, s_B, t_A, t_B \leq \varepsilon n$;*

(iii) *$\delta^0(G[A,B]) \geq \eta n$;*

(iv) *$d^{\pm}_B(x) \geq b - \varepsilon n$ for all but at most εn vertices $x \in A$;*

(v) *$d^{\pm}_A(x) \geq a - \varepsilon n$ for all but at most εn vertices $x \in B$;*

(vi) *$d^-_A(x) \geq \eta n$ for all $x \in S_A$, $d^+_B(x) \geq \eta n$ for all $x \in S_B$, $d^+_A(x) \geq \eta n$ for all $x \in T_A$ and $d^-_B(x) \geq \eta n$ for all $x \in T_B$.*

Suppose that P is a path of length at most $\eta^2 n$ which contains at least $200\varepsilon n$ sink vertices. Let P_1, P_2, P_3 be a useful tripartition of P with sink/source/sink sets $\mathcal{Q}_1, \mathcal{Q}_2, \mathcal{Q}_3$. Let $L \subseteq P_2$ be a link. Suppose that $G \setminus (S_A \cup S_B \cup T_A \cup T_B)$ contains a copy L^G of L which is an AB-path if $d_C(P, L)$ is even and a BA-path otherwise. Let r_A be the number of repeated As in L^G and r_B be the number of repeated Bs in L^G. Let G' be the graph with vertex set $V(G)$ and edges

$$E(A, B \cup S_A) \cup E(B, A \cup T_B) \cup E(T_A, A) \cup E(S_B, B) \cup E(L^G).$$

Then G' contains a copy P^G of P such that:

- *$L^G \subseteq P^G$;*
- *P^G covers S_A, S_B, T_A, T_B;*
- *a_1 is the initial vertex of P^G;*
- *The final vertex of P^G lies in B if P has even length and A if P has odd length;*
- *P^G has $s_A + t_A + r_A$ repeated As and $s_B + t_B + r_B$ repeated Bs.*

Proof. We may assume, without loss of generality, that the initial vertex of P lies in $\mathcal{Q}_1$. If not, let x be the first vertex on P lying in $\mathcal{Q}_1$ and greedily embed the initial segment (Px) of P starting at a_1 using edges in $E(A, B) \cup E(B, A)$. Let a_1' be the image of x. We can then use symmetry to relabel the sets A, B, S_A, S_B, T_A, T_B, if necessary, to assume that $a'_1 \in A$.

We will use (vi) to find a copy of P which covers the vertices in $S_A \cup T_B$ by sink vertices in $\mathcal{Q}_1 \cup \mathcal{Q}_3$ and the vertices in $S_B \cup T_A$ by source vertices in $\mathcal{Q}_2$. We will use that $|\mathcal{Q}_i| \geq 15\varepsilon n$ for all i and also that (iii)–(v) together imply that G' contains a path of length three of any orientation between any pair of vertices in $x \in A$ and $y \in B$. Consider any $q_1 \in \mathcal{Q}_1$ and $q_2 \in \mathcal{Q}_2$. The order in which we cover the vertices will depend on whether $d_C(q_1, q_2)$ is even or odd (note that the parity of $d_C(q_1, q_2)$ does not depend on the choice of q_1 and q_2).

Suppose first that $d_C(q_1, q_2)$ is even. We find a copy of P in G' as follows. Map the initial vertex of P to a_1. Then greedily cover all vertices in T_B so that they are the images of sink vertices in $\mathcal{Q}_1$ using a path P_1^G which is isomorphic to P_1 and has the form $(A{*}BT_BB)^{t_B}A{*}A$. Let x_L be the initial vertex of L and y_L be the final vertex. Let x_L^G and y_L^G be the images of x_L and y_L in L^G. Cover all vertices in S_B so that they are the images of source vertices in $\mathcal{Q}_2$ using a path isomorphic to (P_2x_L) which starts from the final vertex of P_1^G and ends at x_L^G. This path has the form $(A * BS_BB)^{s_B}A * X$, where $X := A$ if $d_C(P, L)$ is even and $X := B$ if $d_C(P, L)$ is odd. Now use the path L_G. Next cover all vertices in T_A so that they are the images of source vertices in $\mathcal{Q}_2$ using a path isomorphic to (y_LP_2) whose initial vertex is y_L^G. This path has the form $Y * A(B * AT_AA)^{t_A}B * B$, where $Y := B$ if $d_C(P, L)$ is even and $Y := A$ if $d_C(P, L)$ is odd. Let P_2^G denote the copy of P_2 obtained in this way. Finally, starting from the final vertex of P_2^G, find a copy of P_3 which covers all vertices in S_A by sink vertices in $\mathcal{Q}_3$ and has the form $(B{*}AS_AA)^{s_A}B{*}B$ if P (and thus also P_3) has even length and $(B * AS_AA)^{s_A}B * A$ if P (and thus also P_3) has odd length. If $d_C(q_1, q_2)$ is odd, we find a copy of P which covers T_B, T_A, $V(L^G)$, S_B, S_A (in this order) in the same way. Observe that P^G has $s_A + t_A + r_A$ repeated As and $s_B + t_B + r_B$ repeated Bs, as required. □

We are now in a position to find an exceptional cover. The proof splits into a number of cases and we will require the assumption that C is not antidirected. We will need a matching found using Proposition 2.6.2 and a careful assignment of the remaining vertices in $S \cup T$ to sets S_A, S_B, T_A and T_B to ensure that the path found by Proposition 2.6.4 leaves a balanced number of vertices in A and B uncovered.

Lemma 2.6.5. *Suppose $1/n \ll \varepsilon_3 \ll \varepsilon_4 \ll 1$. Let G be a digraph on n vertices with $\delta^0(G) \geq n/2$. Suppose A, B, S, T is a partition of $V(G)$ satisfying* (Q1)–(Q9). *If C is an oriented cycle on n vertices, C is not antidirected and $\sigma(C) \geq \varepsilon_4 n$, then there is an exceptional cover P of G of length at most $2\varepsilon_4 n$.*

Proof. Let $d := b - a$, $k := t - s$ and $r := s + t$. Since $\sigma(C) \geq \varepsilon_4 n$, we can use an averaging argument to guarantee a subpath Q' of C of length at most $\varepsilon_4 n$ such that Q' contains at least $2\sqrt{\varepsilon_3} n$ sink vertices. Let Q be an initial subpath of Q' which has odd length and contains $\sqrt{\varepsilon_3} n$ sink vertices.

Case 1: *$a < b$ or $s < t$.*

We will find disjoint sets of vertices S_A, S_B, T_A, T_B, of sizes s_A, s_B, t_A, t_B respectively, and a matching $M' = E \cup E'$ (where E and E' are disjoint) such that the following hold:

(E1) $S_A \cup S_B = S$ and $T_A \cup T_B = T \setminus V(E')$;

(E2) $E \subseteq E(B)$, $|E| \leq d$;

(E3) $E' \subseteq E(B \cup T, B) \cup E(A, A \cup T)$ and $1 \leq |E'| \leq 2$;

(E4) If $p := |E' \cap E(B)| - |E' \cap E(A)|$, then $s_A + t_A + d = s_B + t_B + p + |E|$.

We find sets satisfying (E1)–(E4) as follows. Suppose first that n is odd. Note that we can find a matching $M \subseteq E(B \cup T, B)$ of size $d + 1$. Indeed, if $a < b$ then M exists by Proposition 2.6.2 and if $a = b$, and so $s < t$, we use that $a + s < n/2$ and $\delta^0(G) \geq n/2$ to find M of size $d+1 = 1$. Fix one edge $e \in M$ and let $E' := \{e\}$. There are $r' := r - |V(E') \cap T|$ vertices in $S \cup T$ which are not covered by E'. Set $d' := \min\{r', d - p\}$ and let $E \subseteq (M \setminus E') \cap E(B)$ have size $d - p - d'$.

Suppose that n is even. If $a < b$, by Proposition 2.6.2, we find a matching M of size $d+2$ in $E(B \cup T, B)$. Fix two edges $e_1, e_2 \in M$ and let $E' := \{e_1, e_2\}$. Choose r', d' and E as above.

If n is even and $a = b$, then $a + s = b + s = (n-k)/2 \leq n/2 - 1$. So $d^+_{A \cup T}(x) \geq k/2$ for each $x \in A$ and $d^-_{B \cup T}(x) \geq k/2$ for each $x \in B$. Either we can find a matching M of size two in $E(B \cup T, B) \cup E(A, A \cup T)$ or $t = s + 2$ and there is a vertex $v \in T$ such that $A \subseteq N^-(v)$ and $B \subseteq N^+(v)$. In the latter case, move v to S to get a new partition satisfying (Q1)–(Q9) and the conditions of Case 2. So we will assume that the former holds. Let $E' := M$, $E := \emptyset$, $r' := r - |V(E') \cap T|$ and $d' := -p$.

In each of the above cases, note that $d' \equiv r' \mod 2$ and $|d'| \leq r'$. So we can choose disjoint subsets S_A, S_B, T_A, T_B satisfying (E1) such that $s_A + t_A = (r' - d')/2$ and $s_B + t_B = (r' + d')/2$. Then (E4) is also satisfied.

We construct an exceptional cover as follows. Let L_1 denote the oriented path of length two whose second vertex is a sink and let L_2 denote the oriented path of length two whose second vertex is a source. For each $e \in E'$, we find a copy $L(e)$ of L_1 or L_2 covering e. If $e \in E(A)$ let $L(e)$ be a copy of L_1 of the form AAB, if $e \in E(B)$ let $L(e)$ be a copy of L_1 of the form ABB, if $e \in E(A, T)$ let $L(e)$ be a copy of L_1 of the form ATB and if $e \in E(T, B)$ let $L(e)$ be a copy of L_2 of the form ATB. Note that for each $e \in E'$, the orientation of $L(e)$ is the same regardless of whether it is traversed from its initial vertex to final vertex or vice versa. This means that we can embed it either as an AB-path or a BA-path.

Let a_1 be any vertex in A and let $e_1 \in E'$. Let r_A and r_B be the number of repeated As and Bs, respectively, in $L(e_1)$. So $r_A = 1$ if and only if $e_1 \in E(A)$, otherwise $r_A = 0$. Also, $r_B = 1$ if and only if $e_1 \in E(B)$, otherwise $r_B = 0$. Consider a useful tripartition P_1, P_2, P_3 of Q. Let $L \subseteq P_2$ be a link which is isomorphic to $L(e_1)$. Let x denote the final vertex of Q. Using Proposition 2.6.4 (with $2\varepsilon_3, \varepsilon_4, 1/4$ playing the roles of ε, η, τ), we find a copy Q^G of Q covering S_A, S_B, T_A, T_B whose initial vertex is a_1. Moreover, $L(e_1) \subseteq Q^G \subseteq G_{\{e_1\}} \subseteq G_M$, the final vertex x^G of Q^G lies in A, Q^G has $s_A + t_A + r_A$

repeated As and $s_B+t_B+r_B$ repeated Bs. If $|E'|=2$, let $e_2 \in E' \setminus \{e_1\}$. Let $Q'' := (xQ')$. Let y be the second source vertex in Q'' if $e_2 \in E(T,B)$ and the second sink vertex in Q'' otherwise. Let y^- be the vertex preceding y on C, let y^+ be the vertex following y on C and let $q := d_C(x, y^-)$. Find a path in G whose initial vertex is x^G which is isomorphic to $(Q''y^-)$ and is of the form $A*A$ if q is even and $A*B$ if q is odd, such that the final vertex of this path is an endvertex of $L(e_2)$. Then use the path $L(e_2)$ itself. Let $Z := B$ if q is even and $Z := A$ if q is odd. Finally, extend the path to cover all edges in E using a path of the form $Z*B(A*ABB)^{|E|}A$ which is isomorphic to an initial segment of (y^+Q''). Let P denote the resulting extended subpath of C, so $Q \subseteq P \subseteq Q'$. Let P^G be the copy of P in G_M.

Note that (EC1) and (EC2) hold. Each repeated A in P^G is either a repeated A in Q^G or it occurs when P^G uses $L(e_2)$ in the case when $e_2 \in E(A)$. Similarly, each repeated B in P^G is either a repeated B in Q^G or it occurs when P^G uses $L(e_2)$ in the case when $e_2 \in E(B)$ or when P^G uses an edge in E. Substituting into (2.3) and recalling (E4) gives

$$\begin{aligned}|B \setminus V(P^G)| - |A \setminus V(P^G)| =& b - a - (s_B + t_B + |E| + |E' \cap E(B)|)\\ &+ (s_A + t_A + |E' \cap E(A)|) + 1\\ =& d - (s_B + t_B + |E|) - p + (s_A + t_A) + 1 = 1.\end{aligned}$$

So (EC3) is satisfied and P^G is an exceptional cover.

Case 2: *$a = b$ and $s = t$.*

If $s = t = 0$ then any path consisting of one vertex in A is an exceptional cover. So we will assume that $s, t \geq 1$. We say that C is *close to antidirected* if it contains an antidirected subpath of length $500\varepsilon_3 n$.

Case 2.1: *C is close to antidirected.*

If there is an edge $e \in E(T,B) \cup E(B,S) \cup E(S,A) \cup E(A,T)$ then we are able to find an exceptional cover in the graph $G_{\{e\}}$. We illustrate how to do this when $e = t_1 b_1 \in E(T,B)$, the other cases are similar. Since C is close to but not antidirected, it follows that C

contains a path P of length $500\varepsilon_3 n$ which is antidirected except for the initial two edges which are oriented consistently. Let $s_1 \in S$. If the initial edge of P is a forward edge, let P' be the subpath of P consisting of the first three edges of P and find a copy $(P')^G$ of P' in G of the form $A\{s_1\}BA$. If the initial edge of P is a backward edge, let P' consist of the first two edges of P and let $(P')^G$ be a backward path of the form $B\{s_1\}A$. Let P'' be the subpath of P formed by removing from P all edges in P'. Let $x^G \in A$ be the final vertex of $(P')^G$. Set $S_A := S \setminus \{s_1\}$, $T_B := T \setminus \{t_1\}$ and $S_B, T_A := \emptyset$. Let P_1, P_2, P_3 be a useful tripartition of P''. As in Case 1, let L_2 denote the oriented path of length two whose second vertex is a source. Let $L \subseteq P_2$ be a link which is isomorphic to L_2 and map L to a path L^G of the form BTA which uses the edge t_1b_1. We use Proposition 2.6.4 to find a copy $(P'')^G$ of P'' which uses L^G, covers $S_A \cup T_B$ and whose initial vertex is mapped to x^G. Moreover, the final vertex of P'' is mapped to $A \cup B$ and $(P'')^G$ has $s_A = s - 1$ repeated As and $t_B = t - 1$ repeated Bs. Let P^G be the path $(P')^G \cup (P'')^G$. Then P^G satisfies (EC1) and we may assume that (EC2) holds, by adding a vertex in A as a new initial vertex and/or final vertex if necessary. The repeated As and Bs in P^G are precisely the repeated As and Bs in $(P'')^G$. Therefore, (2.3) implies that (EC3) holds and P^G forms an exceptional cover.

Let us suppose then that $E(T, B) \cup E(B, S) \cup E(S, A) \cup E(A, T)$ is empty. If $S = \{s_1\}, T = \{t_1\}$ then, since $\delta^0(G) \geq n/2$, G must contain the edge s_1t_1 and edges a_1s_1, b_1t_1 for some $a_1 \in A, b_1 \in B$. Since C is not antidirected but has many sink vertices we may assume that C contains a subpath $P = (uvxyz)$ where $uv, vx, yx \in E(C)$. We use the edges a_1s_1, s_1t_1, b_1t_1, as well as an additional AB- or BA-edge, to find a copy P^G of P in G of the form $ASTBA$. The path P^G forms an exceptional cover.

If $s = t = 2$ and $e(S) = e(T) = 2$, we find an exceptional cover as follows. Write $S = \{s_1, s_2\}$, $T = \{t_1, t_2\}$. We have that $s_is_j, t_it_j \in E(G)$ for all $i \neq j$. Note that C is not antidirected, so C must contain a path of length six which is antidirected except for its initial two edges which are consistently oriented. Suppose first that the initial two edges of P are forward edges. Let $a_1 \in A$ be an inneighbour of s_1. Note that s_2 has an inneighbour in T, without loss of generality t_1. Let $b_1 \in B$ be an inneighbour of t_2 and $a_2 \in A$ be an outneighbour of b_1. We find a copy P^G of P which has the form $ASSTTBA$ and uses the edges $a_1s_1, s_1s_2, t_1s_2, t_1t_2, b_1t_2, b_1a_2$, in this order. If the initial two edges of P are backward, we instead find a path of the form $ATTSSBA$. Note that in both cases, P^G satisfies (EC1) and (EC2). P^G has no repeated As and Bs and (2.3) implies that (EC3) holds. So P^G forms an exceptional cover.

So let us assume that $s, t \geq 2$ and, additionally, $e(S) + e(T) < 4$ if $s = 2$. There must exist two disjoint edges $e_1 = t_1s_1$, $e_2 = s_2t_2$ where $s_1, s_2 \in S$ and $t_1, t_2 \in T$ (since $\delta^0(G) \geq n/2$ and $E(T, B) \cup E(B, S) \cup E(S, A) \cup E(A, T) = \emptyset$). We use these edges to find an exceptional cover as follows. We let $S_A := S \setminus \{s_1, s_2\}$, $T_B := T \setminus \{t_1, t_2\}$, $s_A := |S_A|$ and $t_B := |T_B|$. We use e_1 and e_2 to find an antidirected path P^G which starts with a backward edge and is of the form

$$A\{t_1\}\{s_1\}A(B * AS_AA)^{s_A}B * B\{s_2\}\{t_2\}B(A * BT_BB)^{s_B}A.$$

The length of P^G is less than $500\varepsilon_3 n$. So, as C is close to antidirected, C must contain a subpath isomorphic to P^G. We claim that P^G is an exceptional cover. Clearly, P^G satisfies (EC1) and (EC2). For (EC3), note that P^G contains an equal number of repeated As and repeated Bs. Then (2.3) implies that $|B \cap V(P^G)| = |A \cap V(P^G)| + 1$.

Case 2.2: *C is far from antidirected.*

Recall that Q is a subpath of C of length at most $\varepsilon_4 n$ containing at least $\sqrt{\varepsilon_3}n$ sink vertices. Let $\mathcal{Q}$ be a maximum collection of sink vertices in Q such that all vertices in $\mathcal{Q}$ are an even distance apart, then $|\mathcal{Q}| \geq \sqrt{\varepsilon_3}n/2$. Partition the path Q into 11 internally disjoint subpaths so that $Q = (P_1P_1'P_2P_2' \dots P_5P_5'P_6)$ and each subpath contains at least $300\varepsilon_3 n$ elements of $\mathcal{Q}$. Note that each P_i' has length greater than $500\varepsilon_3 n$ and so is not antidirected, that is, each P_i' must contain a consistently oriented subpath P_i'' of length two. At least three of the P_i'' must form a consistent set. Thus there must exist $i < j$ such that $d_C(P_i'', P_j'')$ is even and $\{P_i'', P_j''\}$ is consistent. We may assume, without loss of generality, that P_i'', P_j'' are forward paths and that the second vertex of P_i is in $\mathcal{Q}$. Let P be the subpath of Q whose initial vertex is the initial vertex of P_i and whose final vertex is the final vertex of P_j''.

We will find an exceptional cover isomorphic to P as follows. Choose $s_1 \in S$ and $t_1 \in T$ arbitrarily. Set $S_A := S \setminus \{s_1\}$ and $T_B := T \setminus \{t_1\}$. Map the initial vertex of P to A. We find a copy of P which maps each vertex in S_A to a sink vertex in P_i and each vertex in T_B to a sink vertex in P_j. If $d_C(P_i, P_i'')$ is even, P_i'' is mapped to a path L' of the form $A\{s_1\}B$ and P_j'' is mapped to a path L'' of the form $B\{t_1\}A$. If $d_C(P_i, P_i'')$ is odd, P_i'' is mapped to a path L' of the form $B\{t_1\}A$ and P_j'' is mapped to a path L'' of the form $A\{s_1\}B$. Thus, if $d_C(P_i, P_i'')$ is even, we obtain a copy P^G which starts with a path of the form $A(B * AS_AA)^{s_A}B * A$, then uses L' and continues with a path of the form $B * B(A * BT_BB)^{t_B}A * B$. Finally, the path uses L''. The case when $d_C(P_i, P_i'')$ is odd is similar. (EC1) holds and we may assume that (EC2) holds by adding one vertex to P if necessary. Note that P^G contains an equal number of repeated As and Bs, so (2.3) implies that (EC3) holds and P^G is an exceptional cover. □

2.6.3 Finding a copy of C

Proposition 2.6.3 and Lemma 2.6.5 allow us to find a short exceptional cover for any cycle which is not antidirected. We complete the proof of Lemma 2.6.1 by extending this path to cover the small number of vertices of low degree remaining in A and B and then applying Proposition 2.4.2.

Proof of Lemma 2.6.1. Let P be an exceptional cover of G of length at most $21\varepsilon_4 n$, guaranteed by Proposition 2.6.3 or Lemma 2.6.5. Let

$$X := \{v \in A : d_B^+(v) < n/2 - \varepsilon_3 n \text{ or } d_B^-(v) < n/2 - \varepsilon_3 n\} \text{ and}$$
$$Y := \{v \in B : d_A^+(v) < n/2 - \varepsilon_3 n \text{ or } d_A^-(v) < n/2 - \varepsilon_3 n\}.$$

(Q4) and (Q5) together imply that $|X \cup Y| \leq 2\varepsilon_3 n$. Together with (Q3), this allows us to cover the vertices in $X \cup Y$ by any orientation of a path of length at most $\varepsilon_4 n$. So we can extend P to cover the remaining vertices in $X \cup Y$ (by a path which alternates between A and B). Let P' denote this extended path. Thus $|P'| \leq 22\varepsilon_4 n$. Let x and y be the endvertices of P'. We may assume that $x, y \in A \setminus X$. Let $A' := (A \setminus V(P')) \cup \{x, y\}$ and $B' := B \setminus V(P')$ and consider $G' := G[A', B']$. Note that $|A'| = |B'| + 1$ by (EC3) and

$$\delta^0(G') \geq n/2 - \varepsilon_3 n - 22\varepsilon_4 n \geq (7|B'| + 2)/8.$$

Thus, by Proposition 2.4.2(ii), G' has a Hamilton path of any orientation between x and y in G. We combine this path with P', to obtain a copy of C. □

2.7 G is $ABST$-extremal

In this section we prove that Theorem 2.1.2 holds for all $ABST$-extremal graphs. When G is $ABST$-extremal, the sets A, B, S and T are all of significant size; $G[S]$ and $G[T]$ look like cliques and $G[A, B]$ resembles a complete bipartite graph. The proof will combine ideas from Sections 2.5 and 2.6.

Lemma 2.7.1. *Suppose that $1/n \ll \varepsilon \ll \varepsilon_1 \ll \eta_1 \ll \tau \ll 1$. Let G be a digraph on n vertices with $\delta^0(G) \geq n/2$ and assume that G is $ABST$-extremal. If C is any orientation of a cycle on n vertices which is not antidirected, then G contains a copy of C.*

We will again split the proof into two cases, depending on how many changes of direction C contains. In both cases, the first step is to find an exceptional cover (defined in Section 2.6) which uses only a small number of vertices from $A \cup B$.

2.7.1 Finding an exceptional cover when C has few sink vertices, $\sigma(C) < \varepsilon_2 n$

The following lemma allows us to find an exceptional cover when C is close to being consistently oriented. The two main components of the exceptional cover are a path $P_S \subseteq G[S]$ covering most of the vertices in S and another path $P_T \subseteq G[T]$ covering most of the vertices in T. We are able to find P_S and P_T because $G[S]$ and $G[T]$ are almost complete. A shorter path follows which uses long runs (recall that a long run is a consistently oriented path of length 20) and a small number of vertices from $A \cup B$ to cover any remaining vertices in $S \cup T$. We use edges found by Proposition 2.5.6 to control the number of repeated As and Bs on this path.

Lemma 2.7.2. *Suppose* $1/n \ll \varepsilon \ll \varepsilon_1 \ll \varepsilon_2 \ll \eta_1 \ll \tau \ll 1$. *Let* G *be a digraph on* n *vertices with* $\delta^0(G) \geq n/2$. *Suppose* A, B, S, T *is a partition of* $V(G)$ *satisfying* (R1)–(R10). *Let* C *be an oriented cycle on* n *vertices. If* $\sigma(C) < \varepsilon_2 n$, *then* G *has an exceptional cover* P *such that* $|V(P) \cap (A \cup B)| \leq 2\eta_1^2 n$.

Proof. Let $s^* := s - \lceil \varepsilon_2 n \rceil$ and $d := b - a$. Define $S' \subseteq S$ to consist of all vertices $x \in S$ with $d^+_{B \cup S}(x) \geq b + s - \varepsilon^{1/3} n$ and $d^-_{A \cup S}(x) \geq a + s - \varepsilon^{1/3} n$. Define $T' \subseteq T$ similarly. Note that $|S \setminus S'|, |T \setminus T'| \leq \varepsilon_1 n$ by (R9) and (R10).

We may assume that the vertices of C are labelled so that the number of forward edges is at least the number of backward edges. Let $Q \subseteq C$ be a forward path of length two, this exists since $\sigma(C) < \varepsilon_2 n$. If C is not consistently oriented, we may assume that Q is immediately followed by a backward edge. Define $e_1, e_2, e_3 \in E(C)$ such that $d_C(e_1, Q) = s^*$, $d_C(Q, e_2) = s^* + 1$, $d_C(Q, e_3) = 2$. Let $P_0 := (e_1 C e_2)$.

If at least one of e_1, e_2 is a forward edge, define paths P_T and P_S of order s^* so that $P_0 = (e_1 P_T Q P_S e_2)$. In this case, map Q to a path Q^G in G of the form $T'AS'$. If e_1 and e_2 are both backward edges, our choice of Q implies that e_3 is also a backward edge. Let P_T and P_S be defined so that $P_0 = (e_1 P_T Q e_3 P_S e_2)$. So $|P_T| = s^*$ and $|P_S| = s^* - 1$. In this case, map $(Q e_3)$ to a path Q^G of the form $T'ABS'$.

Let $p_T := |P_T|$ and $p_S := |P_S|$. Our aim is to find a copy P_0^G of P_0 which maps P_S to

$G[S]$ and P_T to $G[T]$. We will find P_0^G of the form F as given in Table 2.1. Let M be a set

e_1 e_2	forward forward	forward backward	backward forward	backward backward
F	$B(T)^{p_T}A(S)^{p_S}B$	$B(T)^{p_T}A(S)^{p_S}A$	$A(T)^{p_T}A(S)^{p_S}B$	$A(T)^{p_T}AB(S)^{p_S}A$

Table 2.1: Proof of Lemma 2.7.2: P_0^G has form F.

of $d+1$ edges in $E(T, B \cup S) \cup E(B, S)$ guaranteed by Proposition 2.5.6. We also define a subset M' of M which we will use to extend P_0^G to an exceptional cover. If e_1, e_2 are both forward edges, choose $M' \subseteq M$ of size d. Otherwise let $M' := M$. Let $d' := |M'|$. Let M_1' be the set of all edges in M' which are disjoint from all other edges in M' and let $d_1' := |M_1'|$. So $M' \setminus M_1'$ consists of $(d' - d_1')/2 =: d_2'$ disjoint consistently oriented paths of the form TBS.

We now fix copies e_1^G and e_2^G of e_1 and e_2. If e_1 is a forward edge, let e_1^G be a BT'-edge, otherwise let e_1^G be a $T'A$-edge. If e_2 is a forward edge, let e_2^G be a $S'B$-edge, otherwise let e_2^G be an AS'-edge. Let t_1 be the endpoint of e_1^G in T', s_2 be the endpoint of e_2^G in S' and let $t_2 \in T'$ and $s_1 \in S'$ be the endpoints of Q^G. Let v be the final vertex of e_2^G and let $X \in \{A, B\}$ be such that $v \in X$.

We now use (R5), (R6), (R9) and (R10) to find a collection $\mathcal{P}$ of at most $3\varepsilon_1 n + 1$ disjoint, consistently oriented paths which cover the edges in M' and the vertices in $S \setminus S'$ and $T \setminus T'$. $\mathcal{P}$ uses each edge $e \in M_1'$ in a forward path P_e of the form $B(S \cup T)^j B$ for some $1 \leq j \leq 4$ and $\mathcal{P}$ uses each path in $M' \setminus M_1'$ in a forward path of the form $BT^jBS^{j'}B$ for some $1 \leq j, j' \leq 4$. The remaining vertices in $S \setminus S'$, $T \setminus T'$ are covered by forward paths in $\mathcal{P}$ of the form $A(S)^jB$ or $B(T)^jA$, for some $1 \leq j \leq 3$.

Let $S'' \subseteq S \setminus (V(\mathcal{P}) \cup \{s_1, s_2\})$ and $T'' \subseteq T \setminus (V(\mathcal{P}) \cup \{t_1, t_2\})$ be sets of size at most $2\varepsilon_2 n$ so that $|S''| + p_S = |S \setminus V(\mathcal{P})|$ and $|T''| + p_T = |T \setminus V(\mathcal{P})|$. Note that $S'' \subseteq S'$ and $T'' \subseteq T'$. So we can cover the vertices in S'' by forward paths of the form ASB and we can cover the vertices in T'' by forward paths of the form BTA. Let $\mathcal{P}'$ be a collection of disjoint paths thus obtained. Let P_1 be the subpath of order $\eta_1^2 n$ following P_0 on C. Note that P_1 contains at least $\sqrt{\varepsilon_2} n$ disjoint long runs. Each path in $\mathcal{P} \cup \mathcal{P}'$ will be contained in the image of such a long run. (Each forward path in $\mathcal{P} \cup \mathcal{P}'$ might be traversed by P_1^G in a forward or backward direction, for example, a forward path of the form $BT^jBS^{j'}B$ could appear in P_1^G as a forward path of the form $BT^jBS^{j'}B$ or a backward path of the form $BS^{j'}BT^jB$.) So we can find a copy P_1^G of P_1 starting from v which uses $\mathcal{P} \cup \mathcal{P}'$ and has the form

$$X * AX_1X_2 \dots X_{d'_1}Y_1Y_2 \dots Y_{d'_2}Z_1Z_2 \dots Z_\ell B * Y$$

for some $\ell \geq 0$ and $Y \in \{A, B\}$, where $X_i \in \{B(S \cup T)^jB * A : 1 \leq j \leq 4\}$, $Y_i \in \{B(S \cup T)^jB(S \cup T)^{j'}B * A : 1 \leq j, j' \leq 4\}$ and

$$Z_i \in \{BA(S \cup T)^jB * A, B(S \cup T)^jA * A : 1 \leq j \leq 3\}.$$

Let S^* be the set of uncovered vertices in S together with the vertices s_1, s_2 and let T^* be the set of uncovered vertices in T together with t_1 and t_2. Write $G_S := G[S^*]$ and $G_T := G[T^*]$. Now $\delta^0(G_T) \geq t - \sqrt{\varepsilon_2}n \geq 7|G_T|/8$ and so G_T has a Hamilton path from t_1 to t_2 which is isomorphic to P_T, by Proposition 2.4.2(i). Similarly, we find a path isomorphic to P_S from s_1 to s_2 in G_S. Altogether, this gives us the desired copy P_0^G of P_0 in G. Let $P^G := P_0^G P_1^G$.

We now check that P^G forms an exceptional cover. Clearly (EC1) holds and we may assume that P^G has both endvertices in A (by extending the path if necessary) so that (EC2) is also satisfied. For (EC3), observe that P_1^G contains exactly $d'_1 + 2d'_2 = d'$ repeated Bs, these occur in the subpath of the form $X_1X_2 \dots X_{d'_1}Y_1Y_2 \dots Y_{d'_2}$ covering the edges in M'. If e_1 and e_2 are both forward edges, then, consulting Table 2.1, we see that P_0^G has no repeated As and that there are no other repeated As or Bs in P^G. Recall that in this case $d' = d$, so (2.3) gives $|B \setminus V(P^G)| - |A \setminus V(P^G)| = d - d' + 1 = 1$. If at least one of e_1, e_2 is a backward edge, using Table 2.1, we see that there is one repeated A in P_0^G and there are no other repeated As or Bs in P^G. In this case, we have $d' = d + 1$, so (2.3) gives $|B \setminus V(P^G)| - |A \setminus V(P^G)| = d - d' + 1 + 1 = 1$. Hence P^G satisfies (EC3) and forms an exceptional cover. Furthermore, $|V(P^G) \cap (A \cup B)| \leq 2\eta_1^2 n$. □

2.7.2 Finding an exceptional cover when C has many sink vertices, $\sigma(C) \geq \varepsilon_2 n$

In Lemma 2.7.4, we find an exceptional cover when C contains many sink vertices. The proof will use the following result which allows us to find short AB- and BA-paths of even length. We will say that an AB- or BA-path P in G is *useful* if it has no repeated As or Bs and uses an odd number of vertices from $S \cup T$.

Proposition 2.7.3. *Suppose* $1/n \ll \varepsilon \ll \varepsilon_1 \ll \eta_1 \ll \tau \ll 1$. *Let* G *be a digraph on* n *vertices with* $\delta^0(G) \geq n/2$. *Suppose* A, B, S, T *is a partition of* $V(G)$ *satisfying* (R1)–(R10). *Let* L_1 *and* L_2 *be oriented paths of length eight. Then* G *contains disjoint copies* L_1^G *and* L_2^G *of* L_1 *and* L_2 *such that each* L_i^G *is a useful path. Furthermore, we can specify whether* L_i^G *is an* AB*-path or a* BA*-path.*

Proof. Define $S' \subseteq S$ to be the set consisting of all vertices $x \in S$ with $d_S^{\pm}(x) \geq \eta_1 n/2$. Define $T' \subseteq T$ similarly. Note that $|S \setminus S'|, |T \setminus T'| \leq \varepsilon_1 n$ by (R9) and (R10). We claim that G contains disjoint edges $e, f \in E(B \cup T, S') \cup E(A \cup S, T')$. Indeed, if $a + s < n/2$ it is easy to find disjoint $e, f \in E(B \cup T, S')$, since $\delta^0(G) \geq n/2$. Otherwise, we must have $a + s = b + t = n/2$ and so each vertex in S' has at least one inneighbour in $B \cup T$ and each vertex in T' has at least one inneighbour in $A \cup S$. Let G' be the bipartite digraph with vertex classes $A \cup S$ and $B \cup T$ and all edges in $E(B \cup T, S') \cup E(A \cup S, T')$. The claim follows from applying König's theorem to the underlying undirected graph of G'.

We demonstrate how to find a copy L_1^G of L_1 in G which is an AB-path. The argument when L_1^G is a BA-path is very similar. L_1^G will have the form $A * B(T)^i(S)^j(T)^k A * B$ or $A * A(T)^i(S)^j(T)^k B * B$, for some $i, j, k \geq 0$ such that $i + j + k$ is odd. Note then that L_1^G will have no repeated As or Bs.

First suppose that L_1 is not antidirected, so L_1 has a consistently oriented subpath L' of length two. We will find a copy of L_1, using (R9)–(R10) to map L' to a forward path of the form ASB or BTA or a backward path of the form BSA or ATB. More precisely, if L' is a forward path, let L_1^G be a path of the form $A * ASB * B$ if $d_C(L_1, L')$ is even and a path of the form $A * BTA * B$ if $d_C(L_1, L')$ is odd. If L' is backward, let L_1^G be a path of the form $A * ATB * B$ if $d_C(L_1, L')$ is even and a path of the form $A * BSA * B$ if $d_C(L_1, L')$ is odd.

Suppose now that L_1 is antidirected. We will find a copy L_1^G of L_1 which contains e. If $e \in E(B, S')$, we use (R9) and the definition of S' to find a copy of L_1 of the following form. If the initial edge of L_1 is a forward edge, we find L_1^G of the form $A(S)^3B * B$. If the initial edge is a backward edge, we find L_1^G of the form $AB(S)^3A * B$. If $e \in E(A, T')$ we will use (R10) and the definition of T' to find a copy of L_1 of the following form. If the initial edge of L_1 is a forward edge, we find L_1^G of the form $A(T)^3B * B$. If the initial edge is a backward edge, we find L_1^G of the form $AB(T)^3A * B$.

If L_1 is antidirected and $e \in E(T, S')$, we will use (R4), (R6), (R9), (R10) and the definition of S' to find a copy of L_1 containing e. If the initial edge of L_1 is a forward edge, find L_1^G of the form $AB(S)^2(T)^{2h-1}A * B$, where $1 \leq h \leq 2$. If the initial edge is a backward edge, find L_1^G of the form $A(T)^{2h-1}(S)^2B * B$, where $1 \leq h \leq 2$. Finally, we consider the case when $e \in E(S, T')$. If the initial edge of L_1 is a forward edge, we find L_1^G of the form $AB(S)^{2h-1}(T)^2A * B$, where $1 \leq h \leq 2$. If the initial edge of L_1 is a backward edge, we find L_1^G of the form $A(T)^2(S)^{2h-1}B * B$, where $1 \leq h \leq 2$.

We find a copy L_2^G of L_2 (which is disjoint from L_1^G) in the same way, using the edge f if L_2 is an antidirected path. □

As when there were few sink vertices, we will map long paths to $G[S]$ and $G[T]$. It will require considerable work to choose these paths so that G contains edges which can be used to link these paths together and so that we are able to cover the remaining vertices in $S \cup T$ using sink and source vertices in a "balanced" way. In many ways, the proof is similar to the proof of Lemma 2.6.5. In particular, we will use Proposition 2.6.4 to map sink and source vertices to some vertices in $S \cup T$.

Lemma 2.7.4. *Suppose* $1/n \ll \varepsilon \ll \varepsilon_1 \ll \varepsilon_2 \ll \eta_1 \ll \tau \ll 1$. *Let* G *be a digraph on* n *vertices with* $\delta^0(G) \geq n/2$. *Suppose* A, B, S, T *is a partition of* $V(G)$ *satisfying* (R1)–(R10). *Let* C *be an oriented cycle on* n *vertices which is not antidirected. If* $\sigma(C) \geq \varepsilon_2 n$, *then* G *has an exceptional cover* P *such that* $|V(P) \cap (A \cup B)| \leq 5\varepsilon_2 n$.

Proof. Let $d := b - a$. Define $S' \subseteq S$ to be the set consisting of all vertices $x \in S$ with $d^{\pm}_S(x) \geq \eta_1 n/2$ and define $T' \subseteq T$ similarly. Let $S'' := S \setminus S'$ and $T'' := T \setminus T'$. Note that $|S''|, |T''| \leq \varepsilon_1 n$ by (R9) and (R10). By (R5), all vertices $x \in S''$ satisfy $d^-_A(x) \geq \eta_1 n/2$ or $d^+_B(x) \geq \eta_1 n/2$ and, by (R6), all $x \in T''$ satisfy $d^+_A(x) \geq \eta_1 n/2$ or $d^-_B(x) \geq \eta_1 n/2$. In our proof below, we will find disjoint sets $S_A, S_B \subseteq S$ and $T_A, T_B \subseteq T$ of suitable size such that

$$d^-_A(x) \geq \eta_1 n/2 \text{ for all } x \in S_A \text{ and } d^+_B(x) \geq \eta_1 n/2 \text{ for all } x \in S_B; \tag{2.4}$$

$$d^-_B(x) \geq \eta_1 n/2 \text{ for all } x \in T_B \text{ and } d^+_A(x) \geq \eta_1 n/2 \text{ for all } x \in T_A. \tag{2.5}$$

Note that (R9) implies that all but at most $\varepsilon_1 n$ vertices from S could be added to S_A or S_B and satisfy the conditions of (2.4). Similarly, (R10) implies that all but at most $\varepsilon_1 n$ vertices in T are potential candidates for adding to T_A or T_B so as to satisfy (2.5). We will write $s_A := |S_A|$, $s_B := |S_B|$, $t_A := |T_A|$ and $t_B := |T_B|$.

Let $s^* := s - \lceil \sqrt{\varepsilon_1} n \rceil$ and let $\ell := 2\lceil \varepsilon_2 n \rceil - 1$. If C contains an antidirected subpath of length ℓ, let Q_2 denote such a path. We may assume that the initial edge of Q_2 is a forward edge by reordering the vertices of C if necessary. Otherwise, choose Q_2 to be any subpath of C of length ℓ such that Q_2 contains at least $\varepsilon_1^{1/3} n$ sink vertices and the second vertex of Q_2 is a sink. Let Q_1 be the subpath of C of length ℓ such that $d_C(Q_1, Q_2) = 2s^* + \ell$. Note that if Q_1 is antidirected then Q_2 must also be antidirected. Let e_1, e_2 be the final two edges of Q_1 and let f_1, f_2 be the initial two edges of Q_2 (where the edges are listed in the order they appear in Q_1 and Q_2, i.e., $(e_1 e_2) \subseteq Q_1$ and $(f_1 f_2) \subseteq Q_2$). Note that f_1 is a forward edge and f_2 is a backward edge.

Let Q' be the subpath of C of length 14 such that $d_C(Q', Q_2) = s^*$. If Q' is antidirected, let Q be the subpath of Q' of length 13 whose initial edge is a forward edge. Otherwise let $Q \subseteq Q'$ be a consistently oriented path of length two. We will consider the three cases stated below.

Case 1: *Q_1 and Q_2 are antidirected. Moreover, $\{e_2, f_1\}$ is consistent if and only if n is* *even.*

We will assume that the initial edge of Q is a forward edge, the case when Q is a backward path of length two is very similar. We will find a copy Q^G of Q which is a $T'S'$-path. If Q is a forward path of length two, map Q to a forward path Q^G of the form $T'AS'$. If Q is antidirected, we find a copy Q^G of Q as follows. Let Q'' be the subpath of Q of length eight such that $d_C(Q, Q'') = 3$. Recall that a path in G is useful if it has no repeated As or Bs and uses an odd number of vertices from $S \cup T$. Using Proposition 2.7.3, we find a copy $(Q'')^G$ of Q'' in G which is a useful AB-path. We find Q^G which starts with a path of the form $T'ABA$, uses $(Q'')^G$ and then ends with a path of the form BAS'. Let q_S and q_T be the numbers of interior vertices of Q^G in S and T, respectively.

If n is even, let $e := e_2$ and, if n is odd, let $e := e_1$. In both cases, let $f := f_1$. The assumptions of this case imply that e and f are both forward edges. Let $P := (Q_1CQ_2)$ and let P_T and P_S be subpaths of C which are internally disjoint from e, f and Q and are such that $(eCf) = (eP_TQP_Sf)$. Our plan is to find a copy of P_T in $G[T]$ and a copy of P_S in $G[S]$. Let $p_T := |P_T|$ and $p_S := |P_S|$. If Q is a consistently oriented path we have that $q_S, q_T = 0$ and $p_S + p_T = d_C(e, f) - 1$. If Q is antidirected, then $q_S + q_T$ is odd and $p_S + p_T = d_C(e, f) - 12$. So in both cases we observe that

$$p_S + p_T + q_S + q_T \equiv d_C(e, f) - 1 \equiv n \mod 2. \tag{2.6}$$

Choose S_A, S_B, T_A, T_B to satisfy (2.4) and (2.5) so that $S'' \setminus V(Q^G) \subseteq S_A \cup S_B$, $T'' \setminus V(Q^G) \subseteq T_A \cup T_B$, $s = s_A + s_B + p_S + q_S$, $t = t_A + t_B + p_T + q_T$ and $s_A + t_A + d = s_B + t_B$. To see that this can be done, first note that the choice of s^* implies that $s - p_S - q_S \geq \sqrt{\varepsilon_1}n/2 > |S''| + d$ and $t - p_T - q_T \geq \sqrt{\varepsilon_1}n/2 > |T''| + d$. Let $r := s + t - (p_S + p_T + q_S + q_T)$. So r is the number of vertices in $S \cup T$ which will not be covered by the copies of P_T, P_S or Q. Then (2.6) implies that

$$r \equiv s + t - n \equiv d \mod 2.$$

Thus we can choose the required subsets S_A, S_B, T_A, T_B so that $s_A + t_A = (r-d)/2$ and $s_B + t_B = (r+d)/2$. Note that (R3) and the choice of s^* also imply that $s_A + s_B, t_A + t_B \leq 2\sqrt{\varepsilon_1}n$.

Recall that Q_1 is antidirected. So we can find a path $(Q_1e)^G$ isomorphic to (Q_1e) which covers the vertices in T_A by source vertices and the vertices in T_B by sink vertices. We choose this path to have the form

$$X * A(BAT_AA * A)^{t_A}(BT_BB * A)^{t_B}B * BT',$$

where $X \in \{A, B\}$. Observe that $(Q_1e)^G$ has t_A repeated As and t_B repeated Bs. Find a path Q_2^G isomorphic to Q_2 of the form

$$S'B * A(BAS_AA * A)^{s_A}(BS_BB * A)^{s_B}B * B$$

which covers all vertices in S_A by sink vertices and all vertices in S_B by source vertices. Q_2^G has s_A repeated As and s_B repeated Bs. So far, we have been working under the assumption that Q starts with a forward edge. If Q is a backward path, the main difference is that we let $e := e_1$ if n is even and let $e := e_2$ if n is odd. We let $f := f_2$ so that e and f are both backward edges and we map Q to a backward path Q^G of the form $T'BS'$. Then (2.6) holds and we can proceed similarly as in the case when Q is a forward path.

We find copies of P_T in $G[T']$ and P_S in $G[S']$ as follows. Greedily embed the first $\sqrt{\varepsilon_1}n$ vertices of P_T to cover all uncovered vertices $x \in T'$ with $d_T^+(x) \leq t - \varepsilon^{1/3}n$ or $d_T^-(x) \leq t - \varepsilon^{1/3}n$. Note that, by (R10), there are at most $\varepsilon_1 n$ such vertices. Write $P_T' \subseteq P_T$ for the subpath still to be embedded and let t_1 and t_2 be the images of its endvertices in T. Let T^* denote the sets of so far uncovered vertices in T together with t_1 and t_2 and define $G_T := G[T^*]$. We have that $\delta^0(G_T) \geq t - \varepsilon^{1/3}n - 3\sqrt{\varepsilon_1}n \geq 7|G_T|/8$, using (R2), and so we can apply Proposition 2.4.2(i) to find a copy of P_T' in G_T with the desired endpoints. In the same way, we find a copy of P_S in $G[S']$. Together with Q^G, $(Q_1e)^G$ and Q_2^G, this gives a copy P^G of P in G such that $|V(P^G) \cap (A \cup B)| \leq 5\varepsilon_2 n$.

The path P^G satisfies (EC1) and we may assume that (EC2) holds, by extending the path by one or two vertices, if necessary, so that both of its endvertices lie in A. Let us now verify (EC3). All repeated As and Bs in P^G are repeated As and Bs in the paths $(Q_1e)^G$ and Q_2^G. So in total, P^G has $s_A + t_A$ repeated As and $s_B + t_B$ repeated Bs. Then (2.3) gives that P^G satisfies

$$|B \setminus V(P^G)| - |A \setminus V(P^G)| = d - (s_B + t_B) + (s_A + t_A) + 1 = 1.$$

So (EC3) is satisfied and P^G is an exceptional cover.

Case 2: *There exists $e \in \{e_1, e_2\}$ and $f \in \{f_1, f_2\}$ such that $\{e, f\}$ is consistent and $n - d_C(e, f)$ is even.*

Let v be the final vertex of f. Recall the definitions of a useful tripartition and a link from Section 2.6. Consider a useful tripartition P_1, P_2, P_3 of (vQ_2) and let $\mathcal{Q}_1, \mathcal{Q}_2, \mathcal{Q}_3$ be sink/source/sink sets. Let $L \subseteq P_2$ be a link of length eight such that $d_C(v, L)$ is even. If Q is a consistently oriented path, use Proposition 2.7.3 to find a copy L^G of L which is a useful BA-path if e is forward and a useful AB-path if e is backward. Map Q to a path Q^G of the form $T'AS'$ if Q is a forward path and $T'BS'$ if Q is a backward path. If Q is antidirected, let Q'' be the subpath of Q of length eight such that $d_C(Q, Q'') = 3$. Using Proposition 2.7.3, we find disjoint copies $(Q'')^G$ of Q'' and L^G of L in G such that $(Q'')^G$ is a useful AB-path and L^G is as described above. We find Q^G which starts with a path of the form $T'ABA$, uses $(Q'')^G$ and then ends with a path of the form BAS'. Let q_S be the number of interior vertices of Q^G and L^G in S and let q_T be the number of interior vertices of Q^G and L^G in T. Note that in all cases, Q^G is a $T'S'$-path with no repeated As or Bs.

Let $P := (eCQ_2)$ and let $P_0 := (eCf)$. Define subpaths P_T and P_S of C which are internally disjoint from Q, e, f and are such that $P_0 = (eP_TQP_Sf)$. Let $p_T := |P_T|$ and $p_S := |P_S|$. Our aim will be to find a copy P_0^G of P_0 which uses Q^G and maps P_T to $G[T]$ and P_S to $G[S]$. P_0^G will have the form F given in Table 2.2. We fix edges e^G and f^G

for e and f. If e is a forward edge, then choose e^G to be a BT'-edge and f^G to be an $S'B$-edge. If e is a backward edge, let e^G be a $T'A$-edge and f^G be an AS'-edge. We also define a constant d' in Table 2.2 which will be used to ensure that the final assignment is balanced. So, if r_A and r_B are the numbers of repeated As and Bs in P_0^G respectively, we

Initial edge of Q e	forward forward	forward backward	backward forward	backward backward
F	$BT^{p_T}\mathcal{A}S^{p_S}B$	$AT^{p_T}\mathcal{A}S^{p_S}A$	$BT^{p_T}BS^{p_S}B$	$AT^{p_T}BS^{p_S}A$
d'	d	$d+2$	$d-2$	d

Table 2.2: Proof of Lemma 2.7.4, Cases 2 and 3: P_0^G has form F, where $\mathcal{A}$ denotes an A-path with no repeated As or Bs.

will have $r_A - r_B = d' - d$.

Note that

$$p_T + p_S + q_T + q_S \equiv d_C(e,f) \equiv n \mod 2. \tag{2.7}$$

The number of vertices in $S \cup T$ which will not be covered by P_0^G or L^G is equal to $r := s + t - (p_T + p_S + q_T + q_S)$ and (2.7) implies that

$$r \equiv s + t - n \equiv d \equiv d' \mod 2.$$

Also note that the choice of s^* implies that $s - p_S - q_S \geq \sqrt{\varepsilon_1}n/2 > |S''| + d'$ and $t - p_T - q_T \geq \sqrt{\varepsilon_1}n/2 > |T''| + d'$. Thus we can choose sets S_A, S_B, T_A, T_B satisfying (2.4) and (2.5) so that $S'' \setminus V(Q^G \cup L^G) \subseteq S_A \cup S_B$, $T'' \setminus V(Q^G \cup L^G) \subseteq T_A \cup T_B$, $s = s_A + s_B + p_S + q_S$, $t = t_A + t_B + p_T + q_T$ and $s_A + t_A + d' = s_B + t_B$. (R3) and the choice of s^* imply that $s_A + s_B, t_A + t_B \leq 2\sqrt{\varepsilon_1}n$. Recall that v denotes the final vertex of f and let v^G be the image of v in G. If $v^G \in A$ (i.e., if e is backward), let $v' := v$ and $(v')^G := v^G$. If $v^G \in B$, let v' denote the successor of v on C. If $vv' \in E(C)$, map v' to an outneighbour of v^G in A and, if $v'v \in E(C)$, map v' to an inneighbour of v^G in A. Let $(v')^G$ be the image of v'. Then we can apply Proposition 2.6.4, with $2\sqrt{\varepsilon_1}, \eta_1/2, \tau/2, (v')^G$ playing the roles of $\varepsilon, \eta, \tau, a_1$, to find a copy $(v'Q_2)^G$ of $(v'Q_2)$ which starts at $(v')^G$, covers S_A, S_B, T_A, T_B and contains L^G. Note that we make use of (2.4) and (2.5) here. We obtain a copy $(vQ_2)^G$ of (vQ_2) (by combining $v^G(v')^G$ with $(v'Q_2)^G$ if $v' \neq v$) which has $s_A + t_A$ repeated As and $s_B + t_B$ repeated Bs.

We find copies of P_T in $G[T]$ and P_S in $G[S]$ as in Case 1. Combining these paths with $(vQ_2)^G$, e^G, Q^G and f^G, we obtain a copy P^G of P in G such that $|V(P^G) \cap (A \cup B)| \leq 3\varepsilon_2 n$. The path P^G satisfies (EC1) and we may assume that (EC2) holds, by extending the path if necessary to have both endvertices in A. All repeated As and Bs in P^G occur as repeated As and Bs in the paths P_0^G and $(vQ_2)^G$ so we can use (2.3) to see that

$$|B \setminus V(P^G)| - |A \setminus V(P^G)| = d - (s_B + t_B) + (d' - d) + (s_A + t_A) + 1 = 1.$$

Therefore, (EC3) is satisfied and P^G is an exceptional cover.

Case 3: *The assumptions of Cases* 1 *and* 2 *do not hold.*

Recall that f_1 is a forward edge and f_2 is a backward edge. Since Case 2 does not hold, this implies that e_2 is a forward edge if n is even (otherwise $e := e_2$ and $f := f_2$ would satisfy the conditions of Case 2) and e_2 is a backward edge if n is odd (otherwise $e := e_2$ and $f := f_1$ would satisfy the conditions of Case 2). In particular, since Case 1 does not hold, this in turn implies that Q_1 is not antidirected. We claim that $Q_1 \setminus \{e_2\}$ is not antidirected. Suppose not. Then it must be the case that $\{e_1, e_2\}$ is consistent. If e_1 and e_2 are forward edges (and so n is even), then $e := e_1$ and $f := f_1$ satisfy the conditions of Case 2. If e_1 and e_2 are both backward edges (and so n is odd), then $e := e_1$ and $f := f_2$ satisfy the conditions of Case 2. Therefore, $Q_1 \setminus \{e_2\}$ is not antidirected and must contain a consistently oriented path Q_1' of length two.

Let $e := e_2$. If n is even, let $f := f_1$ and, if n is odd, let $f := f_2$. In both cases, we have that $\{e, f\}$ is consistent. Let $P := (Q_1' C Q_2)$ and $P_0 := (ePf)$. Let P_T and P_S be subpaths of C defined such that $P_0 = (eP_T Q P_S f)$. Set $p_T := |P_T|$ and $p_S := |P_S|$. Our aim is to find a copy P_0^G which is of the form given in Table 2.2. We also define a constant d' as in Table 2.2. So if r_A and r_B are the numbers of repeated As and Bs in P_0^G respectively, then again $r_A - r_B = d' - d$.

Let v be the final vertex of f. Consider a tripartition P_1, P_2, P_3 of (vQ_2) and a link $L \subseteq P_2$ of length eight such that $d_C(v, L)$ is even. Proceed exactly as in Case 2 to find copies Q^G and L^G of Q and L. Use (R4), (R9) and (R10) to fix a copy $(Q_1'Ce)^G$ of $(Q_1'Ce)$ which is disjoint from Q^G and L^G and is of the form given in Table 2.3. Note that the

Q_1' $d_C(Q_1', e)$	forward odd	forward even	backward odd	backward even
Form of $(Q_1'Ce)^G$ if e is forward	$BTA * BT'$	$ASB * BT'$	$BSA * BT'$	$ATB * BT'$
Form of $(Q_1'Ce)^G$ if e is backward	$ASB * AT'$	$BTA * AT'$	$ATB * AT'$	$BSA * AT'$

Table 2.3: Form of $(Q_1'Ce)^G$ in Case 3.

interior of $(Q_1'Ce)^G$ uses exactly one vertex from $S \cup T$ and $(Q_1'Ce)^G$ has no repeated As or Bs. Write $(Q_1')^G$ for the image of Q_1'. We also fix an edge f^G for the image of f which is disjoint from Q^G, L^G and $(Q_1'Ce)^G$ and is an $S'B$-edge if e is forward and an AS'-edge if e is backward. Let q_S be the number of interior vertices of Q^G, L^G and $(Q_1')^G$ in S and let q_T be the number of interior vertices of Q^G, L^G and $(Q_1')^G$ in T.

Note that $p_S + p_T + q_S + q_T \equiv d_C(e, f) - 1 \equiv n \mod 2$. Using the same reasoning as in Case 2, we find sets S_A, S_B, T_A, T_B satisfying (2.4) and (2.5) such that $S'' \setminus V(Q^G \cup L^G \cup (Q_1')^G) \subseteq S_A \cup S_B$, $T'' \setminus V(Q^G \cup L^G \cup (Q_1')^G) \subseteq T_A \cup T_B$, $s = s_A + s_B + p_S + q_S$, $t = t_A + t_B + p_T + q_T$ and $s_A + t_A + d' = s_B + t_B$. (R3) and the choice of s^* imply that $s_A, t_A, s_B, t_B \leq 2\sqrt{\varepsilon_1}n$. Recall that v denotes the final vertex of f. Similarly as in Case 2, we now use Proposition 2.6.4 to find a copy $(vQ_2)^G$ of (vQ_2) which covers S_A, S_B, T_A, T_B, contains L^G and has $s_A + t_A$ repeated As and $s_B + t_B$ repeated Bs.

We find copies of P_T in $G[T]$ and P_S in $G[S]$ as in Case 1. Together with $(Q_1'Ce)^G$, Q^G, f^G and $(vQ_2)^G$, these paths give a copy P^G of P in G such that $|V(P^G) \cap (A \cup B)| \leq 5\varepsilon_2 n$. The path P^G satisfies (EC1) and we may assume that (EC2) holds, by extending the path so that both endvertices lie in A if necessary. All repeated As and Bs in P^G occur as repeated As and Bs in the paths P_0^G and $(vQ_2)^G$, so we can use (2.3) to see that

$$|B \setminus V(P^G)| - |A \setminus V(P^G)| = d - (s_B - t_B) - (d - d') + (s_A + t_A) + 1 = 1.$$

So (EC3) is satisfied and P^G is an exceptional cover. □

2.7.3 Finding a copy of C

As we did in the AB-extremal case, we will now use an exceptional cover to find a copy of C in G.

Proof of Lemma 2.7.1. Apply Lemma 2.7.2 or Lemma 2.7.4 to find an exceptional cover P of G which uses at most $2\eta_1^2 n$ vertices from $A \cup B$. Let P' be the path of length $\sqrt{\varepsilon_1} n$ following P on C. Extend P by a path isomorphic to P', using this path to cover all $x \in A$ such that $d_B^+(x) \leq b - \varepsilon^{1/3} n$ or $d_B^-(x) \leq b - \varepsilon^{1/3} n$ and all $x \in B$ such that $d_A^+(x) \leq a - \varepsilon^{1/3} n$ or $d_A^-(x) \leq a - \varepsilon^{1/3} n$, using only edges in $E(A, B) \cup E(B, A)$. Let P^* denote the resulting extended path.

We may assume that both endvertices a_1, a_2 of P^* are in A and also that $d_B^\pm(a_i) \geq b - \varepsilon^{1/3} n$ (by extending the path if necessary). Let A^*, B^* denote those vertices in A and B which have not already been covered by P^* together with a_1 and a_2 and let $G^* := G[A^*, B^*]$. We have that $|A^*| = |B^*| + 1$ and $\delta^0(G^*) \geq a - 3\eta_1^2 n \geq (7|B^*| + 2)/8$. Then G^* has a Hamilton path of any orientation with the desired endpoints by Proposition 2.4.2(ii). Together with P^*, this gives a copy of C in G. □

CHAPTER 3

ON THE RANDOM GREEDY F-FREE HYPERGRAPH PROCESS

3.1 Introduction

3.1.1 Results

Fix a k-uniform hypergraph F. In this thesis, we study the following random greedy process, which constructs a maximal F-free k-uniform hypergraph. Assign a birthtime which is uniformly distributed in $[0, 1]$ to each hyperedge of the complete k-uniform hypergraph K_n^k on n vertices. Start with the empty hypergraph on n vertices at time $p = 0$. Increase p and each time that a new hyperedge is born, add it to the hypergraph provided that it does not create a copy of F (edges with equal birthtime are added in any order). Denote the resulting hypergraph at time p by $R_{n,p}$.

The random greedy graph process (i.e. the case when $k = 2$) has been studied for many graphs. The initial motivation (see for example [33]) was to study the Ramsey number $R(3, t)$. Indeed, the best current lower bounds on $R(3, t)$ were obtained via the study of the triangle-free process ([13], [35]). Osthus and Taraz [64] gave an upper bound on the number of edges in the graph $R_{n,1}$ when F is strictly 2-balanced (this condition is defined formally on the next page but, roughly speaking, guarantees that F does not contain particularly dense subgraphs). They showed that a.a.s. $R_{n,1}$ has maximum degree $O(n^{1-(|F|-2)/(e(F)-1)}(\log n)^{1/(\Delta(F)-1)})$. (Here a.a.s. stands for 'asymptotically almost surely', i.e. for the property that an event occurs with probability tending to one as n tends to infinity.) Results for the cases when $F = C_4$ and $F = K_4$ were obtained independently by Bollobás and Riordan [16]. Bohman and Keevash [12] showed that a.a.s. $R_{n,1}$ has minimum degree $\Omega(n^{1-(|F|-2)/(e(F)-1)}(\log n)^{1/(e(F)-1)})$ whenever F is strictly 2-balanced and conjectured that this gives the correct order of magnitude. Improved upper bounds have been obtained for some graphs. For instance, the number of edges has been determined asymptotically when F is a cycle ([11], [13], [35], [65], [80]) and when $F = K_4$ ([81], [83]). Picollelli [66] determined asymptotically the number of edges when F is a diamond, i.e. the graph obtained by removing one edge from K_4. Note that this graph is not strictly 2-balanced.

Much less is known about the process when F is a k-uniform hypergraph and $k \geq 3$. The only known upper bound is due to Bohman, Mubayi and Picollelli [14], who studied the F-free process when F is a k-uniform generalisation of a graph triangle (with an application to certain Ramsey numbers). In this thesis, we obtain a generalisation of the upper bound in [64] to strictly k-balanced hypergraphs. Here we say that a k-uniform hypergraph F is *strictly k-balanced* if $|F| \geq k+1$ and for all proper subgraphs $F' \subsetneq F$ with $|F'| \geq k+1$ we have

$$\frac{e(F)-1}{|F|-k} > \frac{e(F')-1}{|F'|-k}.$$

We also need the following definition. Given a hypergraph H and $i \in \mathbb{N}$, we define the *maximum i-degree* of H by

$$\Delta_i(H) := \max\{d_H(U) : U \subseteq V(H), |U| = i\},$$

where $d_H(U)$ is the number of hyperedges in H containing U. For any k-uniform hypergraph, the *maximum co-degree,* $\Delta_{k-1}(H)$ refers to the maximum $(k-1)$-degree.

Theorem 3.1.1. *Let $k \in \mathbb{N}$ be such that $k \geq 2$. Let F be a strictly k-balanced k-uniform hypergraph which has v vertices and $h \geq v-k+1$ hyperedges. Suppose $\Delta_{k-1}(F) \geq 2$.*

Then there exists a constant c such that a.a.s.

$$\Delta_{k-1}(R_{n,1}) < t \qquad \textit{where} \qquad t := cn^{1-\frac{v-k}{h-1}}(\log n)^{\frac{3}{\Delta_{k-1}(F)-1}-\frac{1}{h-1}}. \tag{3.1}$$

In particular, a.a.s. $R_{n,1}$ has at most tn^{k-1} hyperedges.

Note that Theorem 3.1.1 applies, for example, to all k-uniform cliques K_v^k on $v \geq k+1$ vertices and more generally to all balanced complete ℓ-partite k-uniform hypergraphs with $\ell \geq k$ and more than k vertices.

Bennett and Bohman [10] studied a random greedy independent set algorithm in certain quasi-random hypergraphs. This algorithm finds a maximal independent set by choosing vertices uniformly at random and adding them to the existing set provided they do not create a hyperedge. Note that we can define an $e(F)$-regular hypergraph H whose set of vertices is $E(K_n^k)$ and whose hyperedges correspond to all copies of F in K_n^k. In this case, the random greedy independent set process on H is exactly the F-free process. Their result can be applied in the context of the F-free process to show that if F is a strictly k-balanced k-uniform hypergraph and every vertex of F lies in at least two hyperedges, then a.a.s. $R_{n,1}$ has $\Omega(n^{k-(|F|-k)/(e(F)-1)}(\log n)^{1/(e(F)-1)})$ hyperedges. Up to logarithmic factors, this matches the upper bound given in Theorem 3.1.1.

3.1.2 Open questions

There are many natural open questions related to the random greedy F-free process. First, we discuss bounds on the number of edges in $R_{n,1}$ when F is an ℓ-cycle. Theorem 3.1.1 applies in the case when F is a k-uniform tight cycle. However, there are other natural notions of a hypergraph cycle: Given $\ell \in \mathbb{N}$ with $\ell < k$, we say that a k-uniform hypergraph $C_{\ell,h}$ is an *ℓ-cycle* of length h if there is a cyclic ordering of its vertices $x_1, \dots, x_{h(k-\ell)}$ and a corresponding ordering on its hyperedges $e_0, \dots, e_{h-1}$ such that $e_i = \{x_{i(k-\ell)+1}, \dots, x_{i(k-\ell)+k}\}$. So consecutive hyperedges on the cycle intersect in exactly ℓ vertices. The case when $\ell = k-1$ corresponds to $C_{\ell,h}$ being a tight cycle of length h. It is easy to check that all ℓ-cycles are strictly k-balanced, but only tight cycles satisfy the co-degree condition in Theorem 3.1.1. In the case when $\ell \geq k/2$, ℓ-cycles meet the conditions in [10]. We conjecture that the bound on the number of hyperedges in [10] is of the correct magnitude for any ℓ.

Conjecture 3.1.2. *Let $\ell, k \in \mathbb{N}$ be such that $k \geq 2$ and $k > \ell$ and let $F := C_{\ell,h}$ be the ℓ-cycle of length h. Then a.a.s. $R_{n,1}$ has $\Theta(n^{\frac{h\ell}{h-1}}(\log n)^{\frac{1}{h-1}})$ hyperedges.*

One motivation for Conjecture 3.1.2 is that $p = n^{h\ell/(h-1)-k}(\log n)^{1/(h-1)}$ is the threshold for the property that every hyperedge in $H_{n,p}$ lies in an ℓ-cycle of length h.

Another open problem would be to generalise Theorem 3.1.1 by finding an upper bound on the number of steps in the random greedy independent set process studied in [10].

The random greedy independent set process can also be applied to study arithmetic progression free sets. Suppose $k, n \in \mathbb{N}$. The kAP-free process generates a subset I of $\mathbb{Z}_n$ which does not contain an arithmetic progression of length k as follows. The elements of $\mathbb{Z}_n$ are ordered uniformly at random. Each is then, in turn, added to the set I if it does not create a k term arithmetic progression. So this is another instance of the random greedy independent set algorithm, this time on the hypergraph with vertex set $\mathbb{Z}_n$ whose hyperedges are all arithmetic progressions of length k. When n is prime, Bennett and Bohman [10] showed that a.a.s. the kAP-free process generates a kAP-free set I of size $\Omega(n^{(k-2)/(k-1)}(\log n)^{1/(k-1)})$. It would be interesting to obtain a corresponding upper bound on I. (Note that an upper bound on the number of steps in the random greedy independent set process would imply an upper bound for the kAP-free process.)

3.1.3 Sketch of the argument

Rather than studying the random greedy process itself, we are able to prove Theorem 3.1.1 by obtaining precise information about the random binomial hypergraph $H_{n,p}$. (This idea was first used in [64].) More precisely, write $H_{n,p}$ for the random binomial k-uniform hypergraph on n vertices with hyperedge probability p, i.e., each hyperedge is included in $H_{n,p}$ with probability p, independently of all other hyperedges. We write $H^-_{n,p}$ for the hypergraph formed by removing all (hyperedges in) copies of F from $H_{n,p}$. Note that $H_{n,p}$ can also be viewed as the random hypergraph consisting of all hyperedges with birthtime at most p. Thus, for all $p \in [0,1]$ we have

$$H^-_{n,p} \subseteq R_{n,p} \subseteq R_{n,1}.$$

We will always assume that K^k_n, $H_{n,p}$, $H^-_{n,p}$ and $R_{n,p}$ use the vertex set $[n]$.

In Section 3.2, we collect some large deviation inequalities. The proof of Theorem 3.1.1 is given in Section 3.3, the strategy is as follows. We first identify the largest point p where we can still use $H_{n,p}$ to approximate the behaviour of $H^-_{n,p}$ (i.e. for this p, only a small proportion of edges of $H_{n,p}$ lie in a copy of F). Now let U be a set of $k-1$ vertices in F such that $d_F(U) = \Delta_{k-1}(F)$. Let $\hat{F}$ be the subgraph of F obtained by deleting all those hyperedges which contain U. Let t be as in (3.1). Suppose for a contradiction that there exists a $(k-1)$-set V of degree t in $R_{n,1}$ and let T be the neighbourhood of V in $R_{n,1}$. We will show that in this case we would almost certainly find a copy α of $\hat{F}$ in $H^-_{n,p}[T \cup V]$ which maps U to V. Since $H^-_{n,p} \subseteq R_{n,1}$, α would also be a copy of $\hat{F}$ in $R_{n,1}[T \cup V]$ which maps U to V. But this actually yields a copy of F in $R_{n,1}$, a contradiction. So a.a.s. $\Delta_{k-1}(R_{n,1}) < t$. It is perhaps surprising that for our analysis the order of hyperedges added after this critical point p is irrelevant.

3.2 Tools

Let $\mathcal{S}$ be a collection of subsets of $E(K^k_n)$. For each $\alpha \in \mathcal{S}$, let I_α denote the indicator variable which equals one if all hyperedges in α lie in $H_{n,p}$ and zero otherwise. Set

$$X := \sum_{\alpha \in \mathcal{S}} I_\alpha \qquad \text{and} \qquad \mu := \mathbb{E}[X].$$

Let Y be the size of a largest hyperedge-disjoint collection of elements of $\mathcal{S}$ in $H_{n,p}$ (i.e. the maximum size of a set $\mathcal{S}' \subseteq \mathcal{S}$ such that $I_\alpha = 1$ for all $\alpha \in \mathcal{S}'$ and $\alpha \cap \alpha' = \emptyset$ for all distinct $\alpha, \alpha' \in \mathcal{S}'$). Erdős and Tetali [34] proved the following upper tail bound on Y.

Theorem 3.2.1. *[34]. For every $a \in \mathbb{N}$, we have $\mathbb{P}[Y \geq a] \leq (e\mu/a)^a$.*

We also require a lower tail bound on Y. For all $\alpha, \alpha' \in \mathcal{S}$ with $\alpha \neq \alpha'$, we write $\alpha \sim \alpha'$ if $\alpha \cap \alpha' \neq \emptyset$. Define

$$\Delta := \sum_{\alpha' \sim \alpha} \mathbb{E}[I_\alpha I_{\alpha'}],$$

where the sum is over all ordered pairs $\alpha' \sim \alpha$ in $\mathcal{S}$. Also, let

$$\eta := \max_{\alpha \in \mathcal{S}} \mathbb{E}[I_\alpha] \qquad \text{and} \qquad \nu := \max_{\alpha \in \mathcal{S}} \sum_{\alpha' \in \mathcal{S} : \alpha' \sim \alpha} \mathbb{E}[I_{\alpha'}].$$

The following bound follows from Lemma 4.2 in Chapter 8 and Theorem A.15 in [3], see [64].

Theorem 3.2.2. *Let $\varepsilon > 0$. Then $\mathbb{P}[Y \leq (1-\varepsilon)\mu] \leq e^{(1-\varepsilon)\mu\nu + \frac{\Delta}{2(1-\eta)} - \frac{\varepsilon^2\mu}{2}}$.*

3.3 Proof of Theorem 3.1.1

3.3.1 Basic parameters

Let F be a strictly k-balanced k-uniform hypergraph which has v vertices, h hyperedges and $d := \Delta_{k-1}(F) \geq 2$. Choose positive constants c_1, c_2 satisfying

$$1/c_1 \ll 1/c_2 \ll 1/v, 1/h.$$

Given functions f and g, we will write $f = \tilde{O}(g)$ if there exists a constant c such that $f(n) \leq (\log n)^c g(n)$ for all sufficiently large n.

Set

$$p := \frac{1}{c_2(n^{v-k}\log n)^{1/(h-1)}} \qquad \text{and} \qquad t := c_1 np(\log n)^{3/(d-1)}.$$

Here p is chosen to be as large as possible subject to the constraint that a.a.s. only a small proportion of the hyperedges of $H_{n,p}$ lie in a copy of F. For each $k+1 \leq i \leq v$, we define

$$h_i := \max\{e(F') : F' \subsetneq F, |F'| = i\}.$$

Since F is strictly k-balanced, we have

$$\frac{h-1}{v-k} > \frac{h_i - 1}{i-k}.$$

So for each $k+1 \leq i \leq v$ we can define a positive constant

$$\delta_i := i - k - (h_i - 1)\frac{v-k}{h-1} > 0. \tag{3.2}$$

Let

$$\delta := \min\{\delta_i : k+1 \leq i \leq v\}.$$

We will often use that for $k+1 \leq i \leq v$

$$n^{v-i}p^{h-h_i} \leq n^{v-i-\frac{v-k}{h-1}(h-h_i)} \overset{(3.2)}{=} n^{v-i-\frac{v-k}{h-1}(h-1-\frac{i-k-\delta_i}{v-k}(h-1))} = n^{-\delta_i} \leq n^{-\delta}. \tag{3.3}$$

Note that this bounds the expected number of extensions of a fixed subgraph of F on i vertices into copies of F in $H_{n,p}$.

3.3.2 Many copies of F containing a fixed hyperedge

For a given hyperedge $f \in E(K_n^k)$, an (r, f)-*cluster* is a collection $F_1, F_2, \ldots, F_r$ of r copies of F such that each F_i contains f and for each $1 < i \leq r$, there exists $f_i \in E(F_i)$ such that $f_i \notin E(F_j)$ for any $j < i$. Define $\mathcal{A}$ to be the event that $H_{n,p}$ has no $(\log n, f)$-cluster for any hyperedge f. We will bound the probability of $\mathcal{A}^c$, i.e., the probability that $H_{n,p}$ has a $(\log n, f)$-cluster for some hyperedge f.

Lemma 3.3.1. *We have* $\mathbb{P}[\mathcal{A}^c] \leq n^{-k}$.

Proof. Fix some $f \in E(K_n^k)$. Write $Z_{r,f}$ for the number of (r,f)-clusters in $H_{n,p}$, so $Z_{1,f}$ counts copies of F which contain the hyperedge f. There are h hyperedges in F which could be mapped to f, so

$$\mathbb{E}[Z_{1,f}] \leq hn^{v-k}p^h \leq e^{-2k}$$

with room to spare. Let $r < \log n$ and consider a fixed (r,f)-cluster C in $H_{n,p}$. Let Z_C be the number of $(1,f)$-clusters in $H_{n,p}$ which contain at least one hyperedge which does not lie in C, so each of these $(1,f)$-clusters together with C forms an $(r+1,f)$-cluster. Suppose that α is a $(1,f)$-cluster sharing $k+1 \leq i \leq v$ vertices with C. The set of hyperedges shared by α and C forms a proper subgraph of F on i vertices, so α and C can have at most h_i common hyperedges. This allows us to estimate $\mathbb{E}[Z_C]$ as

$$\mathbb{E}[Z_C] \leq hn^{v-k}p^{h-1} + \sum_{i=k+1}^{v} v^i (rv)^{i-k} n^{v-i} p^{h-h_i} \overset{(3.3)}{\leq} e^{-3k} + \tilde{O}(n^{-\delta}) \leq e^{-2k}.$$

If we sum over all (r,f)-clusters in K_n^k, we find that

$$\mathbb{E}[Z_{r+1,f}] \leq \mathbb{E}[Z_{r,f}]e^{-2k} \leq e^{-2(r+1)k}$$

and hence $\mathbb{E}[Z_{\log n,f}] \leq n^{-2k}$. By summing over all $f \in E(K_n^k)$, we obtain

$$\mathbb{P}[\mathcal{A}^c] \leq \binom{n}{k} n^{-2k} \leq n^{-k},$$

as required. □

3.3.3 Estimating the number of extensions of a fixed set

Recall that $d = \Delta_{k-1}(F)$. Let $U = \{u_1, u_2, \ldots, u_{k-1}\} \subseteq V(F)$ be such that $d_F(U) = d$. Let $N_F(U)$ denote the neighbourhood of U in F, i.e. $N_F(U) := \{x \in V(F) : U \cup \{x\} \in E(F)\}$. Define $\hat{F} \subseteq F$ which has vertex set $V(F)$ and all hyperedges $f \in E(F)$ such that

$|f \cap U| \le k-2$. Fix $T \subseteq [n]$ of size t and an ordered sequence $V = (v_1, v_2, \ldots, v_{k-1})$ of distinct vertices, where $v_i \in [n] \setminus T$ for each $1 \le i \le k-1$. Given a hypergraph $H \subseteq K_n^k$, let $\mathcal{S}(H) = \mathcal{S}(H, T, V)$ be the set of all copies of $\hat{F}$ in H such that the following hold:

- for each $1 \le i \le k-1$, u_i is mapped to v_i;
- $N_F(U)$ is mapped into T and
- $V(F) \setminus N_F(U)$ is mapped into $[n] \setminus T$.

We let $X := |\mathcal{S}(H_{n,p})|$ and $X^- := |\mathcal{S}(H_{n,p}^-)|$. Note that $X^- \le X$ since $H_{n,p}^- \subseteq H_{n,p}$.

Note that if $T \subseteq N_{R_{n,1}}(V)$, then $\mathcal{S}(R_{n,1}) = \emptyset$, as otherwise we could find a copy of F in $R_{n,1}$. Since $H_{n,p}^- \subseteq R_{n,1}$, it follows that $X^- = 0$. So, in order to prove Theorem 3.1.1, it will suffice to prove that a.a.s. we have $X^- > 0$ for any choice of T, V.

Lemma 3.3.2. *Given $T \subseteq [n]$ of size t and an ordered sequence $V = (v_1, v_2, \ldots, v_{k-1})$ of distinct vertices, where $v_i \in [n] \setminus T$ for each $1 \le i \le k-1$, define X^- as above. Then*

$$\mathbb{P}[(X^- = 0) \cap \mathcal{A}] \le 2n^{-2t}.$$

Proof. Write $\mathcal{S} := \mathcal{S}(K_n^k)$. Note that

$$\begin{aligned}\mu_1 := \mathbb{E}[X] &\ge \binom{t}{d}\binom{n-t-k+1}{v-d-k+1}p^{h-d} \ge \frac{tt^{d-1}n^{v-d-k+1}p^{h-d}}{d^d v^v} \\ &= \frac{tc_1^{d-1}n^{v-k}p^{h-1}(\log n)^3}{d^d v^v} = \frac{c_1^{d-1}}{d^d v^v c_2^{h-1}} t(\log n)^2 \ge 24h^2 t(\log n)^2. \end{aligned} \tag{3.4}$$

Let $\mathcal{S}'(H_{n,p})$ be a hyperedge-disjoint collection of elements of $\mathcal{S}(H_{n,p})$ of maximum size and let $Y_1 := |\mathcal{S}'(H_{n,p})|$. In order to apply Theorem 3.2.2, we will estimate ν, Δ and η.

First we estimate ν. Define

$$\nu^* := \max_{\alpha \in \mathcal{S}} \sum_{\alpha' \in \mathcal{S} : \alpha' \sim \alpha} \mathbb{E}[I_{\alpha'} \mid I_\alpha = 1]$$

and note that $\nu \le \nu^*$. We count the expected number of elements $\alpha' \in \mathcal{S}(H_{n,p}) \setminus \{\alpha\}$ sharing at least one hyperedge with some fixed element $\alpha \in \mathcal{S}$. Note that α and α' must share at least two vertices outside V by the definition of $\hat{F}$. We let $k+1 \leq i+j \leq v$ denote the number of shared vertices, where i is the number of vertices shared in T. Consider any $\alpha' \in \mathcal{S} \setminus \{\alpha\}$ sharing $i+j$ vertices with α. Let K be the hypergraph on $i+j$ vertices formed by the set of hyperedges shared by α and α'. Let K' be the hypergraph on $i+j$ vertices obtained from K by adding all hyperedges of the form $V \cup x$ for each of the i vertices $x \in T$ shared by α and α'. Since $K' \subsetneq F$, $e(K') \leq h_{i+j}$ and so α and α' can have at most $h_{i+j} - i$ common hyperedges. Then

$$\begin{aligned}\nu \leq \nu^* &\leq \sum_{i+j=k+1}^{v} v^{i+j} t^{d-i} n^{v-d-j} p^{h-d-(h_{i+j}-i)} \\ &= \sum_{i+j=k+1}^{v} v^{i+j} (c_1 (\log n)^{\frac{3}{d-1}})^{d-i} n^{v-(i+j)} p^{h-h_{i+j}} \overset{(3.3)}{=} \tilde{O}(n^{-\delta}) = o(1).\end{aligned}$$

Since Δ counts the expected number of ordered pairs of elements in $\mathcal{S}(H_{n,p})$ which share at least one hyperedge, we have

$$\Delta \leq \mu_1 \nu^* = o(\mu_1).$$

Finally, the probability of a fixed element in $\mathcal{S}$ being present in $H_{n,p}$ is given by

$$\eta = p^{h-d} = o(1).$$

So we can apply Theorem 3.2.2 to see that

$$\mathbb{P}[Y_1 \leq \mu_1/2] \leq e^{-\mu_1/10} \overset{(3.4)}{\leq} n^{-2t}. \tag{3.5}$$

We define a *couple* (α, F') to be the union of an element $\alpha \in \mathcal{S}'(H_{n,p})$ and a copy F' of F in $H_{n,p}$ which share at least one hyperedge. Note that deleting F' from $H_{n,p}$ to form $H^-_{n,p}$ will destroy α.

We define an auxiliary graph G as follows. For each element of $\mathcal{S}'(H_{n,p})$ which lies in a couple, choose one. These couples form the vertices of G. Draw an edge between two vertices in G if the corresponding couples share a hyperedge. We will use that

$$|G| \leq (\Delta(G) + 1)\alpha(G), \tag{3.6}$$

where $\alpha(G)$ denotes the size of the largest independent set in G. We will use this inequality (which holds for all graphs) to bound the number of vertices in G and show that with sufficiently high probability $|G| < Y_1$. (This in turn implies that at least one element of $\mathcal{S}'(H_{n,p})$ will remain in $H^-_{n,p}$, i.e. $X^- > 0$.)

First, we bound $\alpha(G)$. Let X_2 be the number of couples in $H_{n,p}$. We estimate $\mu_2 := \mathbb{E}[X_2]$, breaking the sum into parts depending on the number i of vertices shared by α and F' in each couple (α, F'). For $k+1 \leq i \leq v$, we use that α and F' intersect in a proper subgraph of F (this is true even when $i = v$) and thus can have at most h_i common hyperedges. The first term in our bound on μ_2 corresponds to those couples (α, F') where α and F' share exactly one hyperedge:

$$\begin{aligned} \mu_2 = \mathbb{E}[X_2] &\leq \mu_1 h^2 n^{v-k} p^{h-1} + \sum_{i=k+1}^{v} \mu_1 v^i n^{v-i} p^{h-h_i} \\ &\overset{(3.3)}{\leq} \mu_1 h^2 n^{v-k} p^{h-1} + O(\mu_1 n^{-\delta}) \leq \mu_1/(12e^2 h^2 \log n). \end{aligned} \tag{3.7}$$

Let Y_2 be the size of a largest hyperedge-disjoint collection of couples in $H_{n,p}$. We note that $\alpha(G) \leq Y_2$ and use Theorem 3.2.1 to bound Y_2:

$$\begin{aligned} \mathbb{P}\left[\alpha(G) \geq \mu_1/(12h^2 \log n)\right] \leq \mathbb{P}\left[Y_2 \geq \mu_1/(12h^2 \log n)\right] &\leq \left(\frac{e\mu_2 12h^2 \log n}{\mu_1}\right)^{\mu_1/(12h^2 \log n)} \\ &\overset{(3.7)}{\leq} e^{-\mu_1/(12h^2 \log n)} \overset{(3.4)}{\leq} n^{-2t}. \end{aligned} \tag{3.8}$$

We now bound $\Delta(G)$. Assume that $\mathcal{A}$ holds, that is, $H_{n,p}$ does not contain a $(\log n, f)$-cluster for any hyperedge f. Fix some hyperedge $f \in E(H_{n,p})$. Let $\mathcal{F}$ be a collection of couples (α_i, F_i) such that $f \in E((\alpha_i, F_i))$ for each i and $\alpha_i \neq \alpha_j$ if $i \neq j$. Suppose, for contradiction, that $|\mathcal{F}| \geq h \log n + 1$. For each couple (α_i, F_i) in $\mathcal{F}$, let e_i be a hyperedge shared by α_i and F_i. The α_i are hyperedge-disjoint by the definition of $\mathcal{S}'(H_{n,p})$, so $f \in E(F_i)$ for all but at most one couple $(\alpha_i, F_i) \in \mathcal{F}$ where $f \in E(\alpha_i)$. If $\mathcal{F}$ contains such a couple, delete it from $\mathcal{F}$. Then, starting with $i = 1$, if (α_i, F_i) has not already been deleted, delete from $\mathcal{F}$ any couples (α_j, F_j) with $j > i$ such that e_j lies in (α_i, F_i). Do this for each i in turn. Since the α_i are hyperedge-disjoint, at each step we delete at most $h-1$ couples from $\mathcal{F}$. So a collection $\mathcal{F}' \subseteq \mathcal{F}$ of at least $\log n$ couples remains. Note that for any $i < j$ such that $(\alpha_i, F_i), (\alpha_j, F_j) \in \mathcal{F}'$, we have $e_j \in E(F_j)$ but $e_j \notin E(F_i)$. But then, the set of all F_i such that $(\alpha_i, F_i) \in \mathcal{F}'$ contains a $(\log n, f)$-cluster in $H_{n,p}$ which is a contradiction to $\mathcal{A}$. Thus the assumption that $|\mathcal{F}| \geq h \log n + 1$ was incorrect. Therefore, $|\mathcal{F}| < h \log n + 1$. Since every couple has fewer than $2h$ hyperedges, we must have

$$\Delta(G) < 2h^2 \log n. \tag{3.9}$$

So, if $\mathcal{A}$ holds, if $\alpha(G) < \mu_1/(12h^2 \log n)$ and if $|Y_1| \geq \mu_1/2$, then

$$|G| \overset{(3.6),(3.9)}{\leq} (2h^2 \log n + 1)\mu_1/(12h^2 \log n) \leq \mu_1/4 < |Y_1|.$$

Thus,

$$\begin{aligned}\mathbb{P}[(X^- = 0) \cap \mathcal{A}] &= \mathbb{P}[(|G| = Y_1) \cap \mathcal{A}] \\ &\leq \mathbb{P}[Y_1 \leq \mu_1/2] + \mathbb{P}[\alpha(G) \geq \mu_1/(12h^2 \log n)] \overset{(3.5),(3.8)}{\leq} 2n^{-2t},\end{aligned}$$

as desired. □

3.3.4 Combining the bounds

We now use Lemmas 3.3.1 and 3.3.2 to prove Theorem 3.1.1.

Proof of Theorem 3.1.1. Define $\mathcal{B}$ to be the event that there exist $T \subseteq [n]$ of size t and an ordered sequence $V = (v_1, v_2, \ldots, v_{k-1})$ of distinct vertices such that $v_i \in [n] \setminus T$

for each $1 \le i \le k-1$ and $X^- = 0$. As remarked before Lemma 3.3.2, $\Delta_{k-1}(R_{n,1}) \ge t$ implies $\mathcal{B}$. So we can apply Lemmas 3.3.1 and 3.3.2 to see that

$$\mathbb{P}[\Delta_{k-1}(R_{n,1}) \ge t] \le \mathbb{P}[\mathcal{B}] \le \mathbb{P}[\mathcal{A}^c] + \mathbb{P}[\mathcal{A} \cap \mathcal{B}] \le n^{-k} + n^{t+k-1}(2n^{-2t}) = o(1).$$

This completes the proof of Theorem 3.1.1. □

CHAPTER 4

CLIQUE DECOMPOSITIONS OF MULTIPARTITE GRAPHS AND COMPLETION OF LATIN SQUARES

4.1 Introduction

A K_r-*decomposition* of a graph G is a partition of its edge set $E(G)$ into cliques of order r. If G has a K_r-decomposition, then certainly $e(G)$ is divisible by $\binom{r}{2}$ and the degree of every vertex is divisible by $r-1$. A classical result of Kirkman [51] asserts that, when $r = 3$, these two conditions ensure that K_n has a triangle decomposition (i.e. Steiner triple systems exist). This was generalized to arbitrary r (for large n) by Wilson [82] and to hypergraphs by Keevash [46]. Recently, there has been much progress in extending this from decompositions of complete host graphs to decompositions of graphs which are allowed to be far from complete. In this chapter, we investigate this question in the r-partite setting. This is of particular interest as it implies results on the completion of partial Latin squares and more generally partial mutually orthogonal Latin squares.

4.1.1 Clique decompositions of r-partite graphs

Our main result (Theorem 4.1.1) states that if G is (i) balanced r-partite, (ii) satisfies the necessary divisibility conditions and (iii) its minimum degree is at least a little larger than the minimum degree which guarantees an approximate decomposition into r-cliques, then G in fact has a decomposition into r-cliques. (Here an approximate decomposition is a set of edge-disjoint copies of K_r which cover almost all edges of G.) To state this result precisely, we need the following definitions.

We say that a graph or multigraph G on $(V_1, \dots, V_r)$ is K_r-*divisible* if G is r-partite with vertex classes $V_1, \dots, V_r$ and for all $1 \leq j_1, j_2 \leq r$ and every $v \in V(G) \setminus (V_{j_1} \cup V_{j_2})$,

$$d(v, V_{j_1}) = d(v, V_{j_2}).$$

Note that in this case, for all $1 \leq j_1, j_2, j_3, j_4 \leq r$ with $j_1 \neq j_2$, $j_3 \neq j_4$, we automatically have $e(V_{j_1}, V_{j_2}) = e(V_{j_3}, V_{j_4})$. In particular, $e(G)$ is divisible by $e(K_r) = \binom{r}{2}$.

Let G be an r-partite graph on $(V_1, \dots, V_r)$ with $|V_1| = \dots = |V_r| = n$. Let

$$\hat{\delta}(G) := \min\{d(v, V_j) : 1 \leq j \leq r,\ v \in V(G) \setminus V_j\}.$$

An η-*approximate* K_r-decomposition of G is a set of edge-disjoint copies of K_r covering all but at most ηn^2 edges of G. We define $\hat{\delta}^{\eta}_{K_r}(n)$ to be the infimum over all δ such that every K_r-divisible graph G on $(V_1, \dots, V_r)$ with $|V_1| = \dots = |V_r| = n$ and $\hat{\delta}(G) \geq \delta n$ has an η-approximate K_r-decomposition. Let $\hat{\delta}^{\eta}_{K_r} := \limsup_{n \to \infty} \hat{\delta}^{\eta}_{K_r}(n)$. So if $\varepsilon > 0$ and G is sufficiently large, K_r-divisible and $\hat{\delta}(G) > (\hat{\delta}^{\eta}_{K_r} + \varepsilon)n$, then G has an η-approximate K_r-decomposition. Note that it is important here that G is K_r-divisible. Take, for example, the complete r-partite graph with vertex classes of size n and remove $\lceil \eta n \rceil$ edge-disjoint perfect matchings between one pair of vertex classes. The resulting graph G satisfies $\hat{\delta}(G) = n - \lceil \eta n \rceil$, yet has no η-approximate K_r-decomposition whenever $r \geq 3$.

Theorem 4.1.1. *For every $r \geq 3$ and every $\varepsilon > 0$ there exists an $n_0 \in \mathbb{N}$ and an $\eta > 0$ such that the following holds for all $n \geq n_0$. Suppose G is a K_r-divisible graph on $(V_1, \dots, V_r)$ with $|V_1| = \dots = |V_r| = n$. If $\hat{\delta}(G) \geq (\hat{\delta}^{\eta}_{K_r} + \varepsilon)n$, then G has a K_r-decomposition.*

By a result of Haxell and Rödl [43], the existence of an approximate decomposition follows from that of a fractional decomposition. So together with very recent results of Bowditch and Dukes [17] as well as Montgomery [60] on fractional decompositions into triangles and cliques respectively, Theorem 4.1.1 implies the following explicit bounds. We discuss this derivation in Section 4.1.3.

Theorem 4.1.2. *For every $r \geq 3$ and every $\varepsilon > 0$ there exists an $n_0 \in \mathbb{N}$ such that the following holds for all $n \geq n_0$. Suppose G is a K_r-divisible graph on $(V_1, \dots, V_r)$ with $|V_1| = \dots = |V_r| = n$.*

(i) If $r = 3$ and $\hat{\delta}(G) \geq \left(\frac{101}{104} + \varepsilon\right) n$, then G has a K_3-decomposition.

(ii) If $r \geq 4$ and $\hat{\delta}(G) \geq \left(1 - \frac{1}{10^6 r^3} + \varepsilon\right) n$, then G has a K_r-decomposition.

If G is the complete r-partite graph, this corresponds to a theorem of Chowla, Erdős and Straus [21]. A bound of $(1 - 1/(10^{16} r^{29}))n$ was claimed by Gustavsson [40]. The following conjecture seems natural (and is implicit in [40]).

Conjecture 4.1.3. *For every $r \geq 3$ there exists an $n_0 \in \mathbb{N}$ such that the following holds for all $n \geq n_0$. Suppose G is a K_r-divisible graph on $(V_1, \dots, V_r)$ with $|V_1| = \dots = |V_r| = n$. If $\hat{\delta}(G) \geq (1 - 1/(r+1))n$, then G has a K_r-decomposition.*

A construction which matches the lower bound in Conjecture 4.1.3 is described in Section 4.3.1 (this construction also gives a similar lower bound on $\hat{\delta}^{\eta}_{K_r}$). In the non-partite setting, the triangle case is a long-standing conjecture by Nash-Williams [62] that every graph G on n vertices with minimum degree at least $3n/4$ has a triangle decomposition (subject to divisibility conditions). Barber, Kühn, Lo and Osthus [7] recently reduced its asymptotic version to proving an approximate or fractional version. Corresponding results on fractional triangle decompositions were proved by Yuster [86], Dukes [29], Garaschuk [36] and Dross [28].

More generally [7] also gives results for arbitrary graphs, and corresponding fractional decomposition results have been obtained by Yuster [86], Dukes [29] as well as

Barber, Kühn, Lo, Montgomery and Osthus [6]. Further results on F-decompositions of non-partite graphs (leading on from [7]) have been obtained by Glock, Kühn, Lo, Montgomery and Osthus [38]. Amongst others, for any bipartite graph F, they asymptotically determine the minimum degree threshold which guarantees an F-decomposition.

4.1.2 Mutually orthogonal Latin squares and K_r-decompositions of r-partite graphs

A *Latin square* $\mathcal{T}$ of order n is an $n \times n$ grid of cells, each containing a symbol from $[n]$, such that no symbol appears twice in any row or column. It is easy to see that $\mathcal{T}$ corresponds to a K_3-decomposition of the complete tripartite graph $K_{n,n,n}$ with vertex classes consisting of the rows, columns and symbols.

Now suppose that we have a partial Latin square; that is, a partially filled in grid of cells satisfying the conditions defining a Latin square. When can it be completed to a Latin square? This natural question has received much attention. For example, a classical theorem of Smetaniuk [71] as well as Anderson and Hilton [5] states that this is always possible if at most $n-1$ entries have been made (this bound is best possible). The case $r=3$ of Conjecture 4.1.3 implies that, provided we have used each row, column and symbol at most $n/4$ times, it should also still be possible to complete a partial Latin square. This was conjectured by Daykin and Häggkvist [23]. (Note that this conjecture corresponds to the special case of Conjecture 4.1.3 when $r=3$ and the condition of G being K_r-divisible is replaced by that of G being obtained from $K_{n,n,n}$ by deleting edge-disjoint triangles.)

More generally, we say that two Latin squares R (red) and B (blue) drawn in the same $n \times n$ grid of cells are *orthogonal* if no blue symbol appears twice next to the same red symbol. In the same way that a Latin square corresponds to a K_3-decomposition of $K_{n,n,n}$, a pair of orthogonal Latin squares corresponds to a K_4-decomposition of $K_{n,n,n,n}$ where the vertex classes are rows, columns, red symbols and blue symbols. More generally, there is a natural bijection between sequences of $r-2$ *mutually orthogonal* Latin squares (where every pair from the sequence are orthogonal) and K_r-decompositions of complete r-partite graphs with vertex classes of equal size. Sequences of mutually orthogonal Latin squares are also known as *transversal designs*. Theorem 4.1.2 can be used to show the following (see Section 4.3.2 for details).

Theorem 4.1.4. *For every $r \geq 3$ and every $\varepsilon > 0$ there exists an $n_0 \in \mathbb{N}$ such that the following holds for all $n \geq n_0$. Let*

$$c_r := \begin{cases} \frac{3}{104} & \text{if } r = 3, \\ \frac{9}{10^7 r^3} & \text{if } r \geq 4. \end{cases}$$

Let $\mathcal{T}_1, \dots, \mathcal{T}_{r-2}$ be a sequence of mutually orthogonal partial $n \times n$ Latin squares (drawn in the same $n \times n$ grid). Suppose that each row and column of the grid contains at most $(c_r - \varepsilon)n$ non-empty cells and each coloured symbol is used at most $(c_r - \varepsilon)n$ times. Then $\mathcal{T}_1, \dots, \mathcal{T}_{r-2}$ can be completed to a sequence of mutually orthogonal Latin squares.

The best previous bound for the triangle case $r = 3$ is due to Bartlett [9], who obtained a minimum degree bound of $(1 - 10^{-4})n$. This improved an earlier bound of Chetwynd and Häggkvist [20] as well as the one claimed by Gustavsson [40].

4.1.3 Fractional and approximate decompositions

A *fractional K_r-decomposition* of a graph G is a non-negative weighting of the copies of K_r in G such that the total weight of all the copies of K_r containing any fixed edge of G is exactly 1. Fractional decompositions are of particular interest to us because of the following result of Haxell and Rödl, of which we state only a very special case.

Theorem 4.1.5 (Haxell and Rödl [43])**.** *For every $r \geq 3$ and every $\eta > 0$ there exists an $n_0 \in \mathbb{N}$ such that the following holds. Let G be a graph on $n \geq n_0$ vertices that has a fractional K_r-decomposition. Then G has an η-approximate K_r-decomposition.*

We define $\hat{\delta}^*_{K_r}(n)$ to be the infimum over all δ such that every K_r-divisible graph G on $(V_1, \dots, V_r)$ with $|V_1| = \dots = |V_r| = n$ and $\hat{\delta}(G) \geq \delta n$ has a fractional K_r-decomposition. Let $\hat{\delta}^*_{K_r} := \limsup_{n\to\infty} \hat{\delta}^*_{K_r}(n)$. Theorem 4.1.5 implies that, for every $\eta > 0$, $\hat{\delta}^\eta_{K_r} \leq \hat{\delta}^*_{K_r}$. Together with Theorem 4.1.1, this yields the following.

Corollary 4.1.6. *For every $r \geq 3$ and every $\varepsilon > 0$ there exists an $n_0 \in \mathbb{N}$ such that the following holds for all $n \geq n_0$. Suppose G is a K_r-divisible graph on $(V_1, \dots, V_r)$ with $|V_1| = \dots = |V_r| = n$. If $\hat{\delta}(G) \geq (\hat{\delta}^*_{K_r} + \varepsilon)n$, then G has a K_r-decomposition.*

In particular, to prove Conjecture 4.1.3 asymptotically, it suffices to show that $\hat{\delta}^*_{K_r} \leq 1 - 1/(r + 1)$. For triangles, the best bound on the 'fractional decomposition threshold' is due to Bowditch and Dukes [17].

Theorem 4.1.7 (Bowditch and Dukes [17])**.** $\hat{\delta}^*_{K_3} \leq \frac{101}{104}$.

For arbitrary cliques, Montgomery obtained the following bound. Somewhat weaker bounds (obtained by different methods) are also proved in [17].

Theorem 4.1.8 (Montgomery [60])**.** *For every $r \geq 3$, $\hat{\delta}^*_{K_r} \leq 1 - \frac{1}{10^6 r^3}$.*

Note that together with Corollary 4.1.6, these results immediately imply Theorem 4.1.2.

This chapter is organised as follows. In Section 4.2 we introduce some notation and tools which will be used throughout this chapter. In Section 4.3 we give extremal constructions which support the bounds in Conjecture 4.1.3 and we provide a proof of Theorem 4.1.4. Section 4.4 outlines the proof of Theorem 4.1.1 and guides the reader through the remaining sections in this chapter.

4.2 Notation and tools

Let G be a graph and let $\mathcal{P} = \{U^1, \dots, U^k\}$ be a partition of $V(G)$. We write $G[U^1]$ for the subgraph of G induced by U^1 and $G[U^1, U^2]$ for the bipartite subgraph of G induced by the vertex classes U^1 and U^2. We will also sometimes write $G[U^1, U^1]$ for $G[U^1]$. We write $G[\mathcal{P}] := G[U^1, \dots, U^k]$ for the k-partite subgraph of G induced by the partition $\mathcal{P}$. We write $U^{<i}$ for $U^1 \cup \cdots \cup U^{i-1}$. We say the partition $\mathcal{P}$ is *equitable* if its parts differ in size by at most one. Given a set $U \subseteq V(G)$, we write $\mathcal{P}[U]$ for the restriction of $\mathcal{P}$ to U.

Let G be a graph and let $U, V \subseteq V(G)$. We write $N_G(U, V) := \{v \in V : xv \in E(G) \text{ for all } x \in U\}$ and $d_G(U, V) := |N_G(U, V)|$. For $v \in V(G)$, we write $N_G(v, V)$ for $N_G(\{v\}, V)$ and $d_G(v, V)$ for $d_G(\{v\}, V)$. If U and V are disjoint, we let $e_G(U, V) := e(G[U, V])$.

Let G and H be graphs. We write $G - H$ for the graph with vertex set $V(G)$ and edge set $E(G) \setminus E(H)$. We write $G \setminus H$ for the subgraph of G induced by the vertex set $V(G) \setminus V(H)$. We call a vertex-disjoint collection of copies of H in G an *H-matching*. If the H-matching covers all vertices in G, we say that it is *perfect*.

Throughout this chapter, we consider a partition $V_1, \dots, V_r$ of a vertex set V such that $|V_j| = n$ for all $1 \leq j \leq r$. Given a set $U \subseteq V$, we write

$$U_j := U \cap V_j.$$

We write $K_r(k)$ for the complete r-partite graph with vertex classes of size k. We say that an r-partite graph G on $(V_1, \dots, V_r)$ is *balanced* if $|V_1| = \cdots = |V_r|$.

Let $m, n, N \in \mathbb{N}$ with $m, n < N$. The *hypergeometric distribution* with parameters N, n and m is the distribution of the random variable X defined as follows. Let S be a random subset of $\{1, 2, \dots, N\}$ of size n and let $X := |S \cap \{1, 2, \dots, m\}|$. We will frequently use the following bounds, which are simple forms of Hoeffding's inequality.

Lemma 4.2.1 (see [44, Remark 2.5 and Theorem 2.10]). *Let $X \sim B(n, p)$ or let X have a hypergeometric distribution with parameters N, n, m. Then $\mathbb{P}(|X - \mathbb{E}(X)| \geq t) \leq 2e^{-2t^2/n}$.*

Lemma 4.2.2 (see [44, Corollary 2.3 and Theorem 2.10]). *Suppose that X has binomial or hypergeometric distribution and $0 < a < 3/2$. Then $\mathbb{P}(|X - \mathbb{E}(X)| \geq a\mathbb{E}(X)) \leq 2e^{-a^2\mathbb{E}(X)/3}$.*

4.3 Extremal graphs and completion of Latin squares

4.3.1 Extremal graphs

The following proposition shows that the minimum degree bound conjectured in Conjecture 4.1.3 would be best possible. It also provides a lower bound on the approximate decomposition threshold $\hat{\delta}^{\eta}_{K_r}$ (and thus on the fractional decomposition threshold $\hat{\delta}^{*}_{K_r}$).

Proposition 4.3.1. *Let $r \in \mathbb{N}$ with $r \geq 3$ and let $\eta > 0$. For infinitely many n, there exists a K_r-divisible graph G on $(V_1, \ldots, V_r)$ with $|V_1| = \cdots = |V_r| = n$ and $\hat{\delta}(G) = \lceil (1 - 1/(r+1))n \rceil - 1$ which does not have a K_r-decomposition. Moreover, $\hat{\delta}^{\eta}_{K_r} \geq 1 - 1/(r+1) - \eta$.*

Proof. Let $m \in \mathbb{N}$ with $1/m \ll \eta$ and let $n := (r-1)m$. Let $\{U^1, \ldots, U^{r-1}\}$ be a partition of $V_1 \cup \cdots \cup V_r$ such that, for each $1 \leq i \leq r-1$ and each $1 \leq j \leq r$, $U^i_j = U^i \cap V_j$ has size m.

Let G_0 be the intersection of the complete r-partite graph on $(V_1, \ldots, V_r)$ and the complete $(r-1)$-partite graph on $(U^1, \ldots, U^{r-1})$. For each $1 \leq q \leq m$ and each $1 \leq i \leq r-1$, let H^i_q be a graph formed by starting with the empty graph on U^i and including a q-regular bipartite graph with vertex classes $(U^i_{j_1}, U^i_{j_2})$ for each $1 \leq j_1 < j_2 \leq r$. Let $H_q := H^1_q \cup \cdots \cup H^{r-1}_q$ and let $G_q := G_0 \cup H_q$. Observe that G_q is regular, K_r-divisible and

$$\hat{\delta}(G_q) = (r-2)m + q.$$

Now G_0 is $(r-1)$-partite, so every copy of K_r in G_q contains at least one edge of H_q. Therefore, any collection of edge-disjoint copies of K_r in G will leave at least

$$\begin{aligned} \ell(G_q) := e(G_q) - e(H_q)\binom{r}{2} &= \left((r-2)m + q - \binom{r}{2}q\right)\binom{r}{2}n \\ &= (m - (r+1)q/2)(r-2)\binom{r}{2}n \end{aligned}$$

edges of G_q uncovered. Let $q_0 := \lceil 2m/(r+1) \rceil - 1$. Then $\ell(G_{q_0}) > 0$, so G_{q_0} does not have a K_r-decomposition. Also,

$$\hat{\delta}(G_{q_0}) = (r-2)m + \lceil 2m/(r+1)\rceil - 1 = \lceil (1-1/(r+1))n\rceil - 1.$$

Now let $q_\eta := \lceil 2m/(r+1) - \eta n\rceil$. We have $\hat{\delta}(G_{q_\eta}) \geq (1-1/(r+1)-\eta)n$ and

$$\begin{aligned}\ell(G_{q_\eta}) &\geq (m - (2m/(r+1) - \eta n + 1)(r+1)/2)(r-2)\binom{r}{2}n \\ &= (\eta n - 1)(r+1)(r-2)r(r-1)n/4 \geq 6(\eta n - 1)n > \eta n^2.\end{aligned}$$

Thus, $\hat{\delta}^{\eta}_{K_r} \geq 1 - 1/(r+1) - \eta$. □

4.3.2 Completion of mutually orthogonal Latin squares

In this section, we give a proof of Theorem 4.1.4. Note that better bounds on the fractional decomposition threshold would immediately lead to better bounds on c_r. For any r-partite graph H on $(V_1, \dots, V_r)$, we let $\overline{H}$ denote the r-partite complement of H on $(V_1, \dots, V_r)$.

Proof of Theorem 4.1.4. By making ε smaller if necessary, we may assume that $\varepsilon \ll 1$. Let $n_0 \in \mathbb{N}$ be such that $1/n_0 \ll \varepsilon, 1/r$. Use $\mathcal{T}_1, \dots, \mathcal{T}_{r-2}$ to construct a balanced r-partite graph G with vertex classes $V_j = [n]$ for $1 \leq j \leq r$ as follows. For each $1 \leq i, j, k \leq n$ and each $1 \leq m \leq r-2$, if in $\mathcal{T}_m$ the cell (i,j) contains the symbol k, include a K_3 on the vertices $i \in V_{r-1}$, $j \in V_r$ and $k \in V_m$. (If the cell (i,j) is filled in different $\mathcal{T}_m$, this leads to multiple edges between $i \in V_{r-1}$ and $j \in V_r$, which we disregard.) For each $1 \leq i, j, k, k' \leq n$ and each $1 \leq m < m' \leq r-2$ such that the cell (i,j) contains symbol k in $\mathcal{T}_m$ and symbol k' in $\mathcal{T}_{m'}$, add an edge between the vertices $k \in V_m$ and $k' \in V_{m'}$.

If $r = 3$, then G is an edge-disjoint union of copies of K_3, so G is K_3-divisible. Then $\overline{G}$ is also K_3-divisible and $\hat{\delta}(\overline{G}) \geq (101/104 + \varepsilon)n$. So we can apply Theorem 4.1.2 to find a K_3-decomposition of $\overline{G}$ which we can then use to complete $\mathcal{T}_1$ to a Latin square.

Suppose now that $r \geq 4$. Observe that G consists of an edge-disjoint union of cliques

$H_1, \dots, H_q$ such that, for each $1 \le i \le q$, H_i contains an edge of the form xy where $x \in V_{r-1}$ and $y \in V_r$. We have $q \le (c_r - \varepsilon)n^2$. We now show that we can extend G to a graph which has a K_r-decomposition. We will do this by greedily extending each H_i in turn to a copy H_i' of K_r. Suppose that $1 \le p \le q$ and we have already found edge-disjoint $H_1', \dots, H_{p-1}'$. Given $v \in V(G)$, let $s(v, p-1)$ be the number of graphs in $\{H_1', \dots, H_{p-1}'\} \cup \{H_p, \dots, H_q\}$ which contain v. Suppose that $s(v, p-1) \le 10(c_r - \varepsilon^2)n/9$ for all $v \in V(G)$. For each $1 \le j \le r$, let $B_j := \{v \in V_j : s(v, p-1) \ge 10(c_r - \varepsilon)n/9\}$. We have

$$|B_j| \le \frac{q}{10(c_r - \varepsilon)n/9} \le \frac{9n}{10}. \tag{4.1}$$

Let $G_{p-1} := G \cup \bigcup_{i=1}^{p-1} (H_i' - H_i)$. Note that

$$\hat{\delta}(\overline{G}_{p-1}) \ge (1 - 10(c_r - \varepsilon^2)/9)n. \tag{4.2}$$

We will extend H_p to a copy of K_r as follows. Let $\{j_1, \dots, j_m\} = \{j : 1 \le j \le r \text{ and } V(H_p) \cap V_j = \emptyset\}$. For each j_i in turn, starting with j_1, choose one vertex x_{j_i} from the set $N_{\overline{G}_{p-1}}(V(H_p) \cup \{x_{j_1}, \dots, x_{j_{i-1}}\}, V_{j_i} \setminus B_{j_i})$. This is possible since (4.1) and (4.2) imply

$$d_{\overline{G}_{p-1}}(V(H_p) \cup \{x_{j_1}, \dots, x_{j_{i-1}}\}, V_{j_i} \setminus B_{j_i}) \ge (1/10 - (r-1)10(c_r - \varepsilon^2)/9)n > 0.$$

Let H_p' be the copy of K_r with vertex set $V(H_p) \cup \{x_j : 1 \le j \le r \text{ and } V(H_p) \cap V_j = \emptyset\}$. By construction, for every $v \in V(G)$, the number $s(v, p)$ of graphs in $\{H_1', \dots, H_p'\} \cup \{H_{p+1}, \dots, H_q\}$ which contain v satisfies $s(v, p) \le 10(c_r - \varepsilon^2)n/9$.

Continue in this way to find edge-disjoint $H_1', \dots, H_q'$ such that $s(v, q) \le 10(c_r - \varepsilon^2)n/9$. Let $G_q := \bigcup_{1 \le i \le q} H_i'$. We have $\hat{\delta}(\overline{G}_q) \ge (1 - 10(c_r - \varepsilon^2)/9)n = (1 - 1/10^6 r^3 + 10\varepsilon^2/9)n$ and, since G_q is an edge-disjoint union of copies of K_r, we know that $\overline{G}_q$ is K_r-divisible. So we can apply Theorem 4.1.2 to find a K_r-decomposition $\mathcal{F}$ of $\overline{G}_q$. Note that $\mathcal{F}' := \mathcal{F} \cup \bigcup_{1 \le i \le q} H_i'$ is a K_r-decomposition of the complete r-partite graph. Since $H_i \subseteq H_i'$

for each $1 \leq i \leq q$, we can use $\mathcal{F}'$ to complete $\mathcal{T}_1, \ldots, \mathcal{T}_{r-2}$ to a sequence of mutually orthogonal Latin squares. □

4.4 Proof sketch

Our proof of Theorem 4.1.1 builds on the proof of the main results of [7], but requires significant new ideas. In particular, the r-partite setting involves a stronger notion of divisibility (the non-partite setting simply requires that $r-1$ divides the degree of each vertex of G and that $\binom{r}{2}$ divides $e(G)$) and we have to work much harder to preserve it during our proof. This necessitates a delicate 'balancing' argument (see Section 4.10). In addition, we use a new construction for our absorbers, which allows us to obtain the best possible version of Theorem 4.1.1. (The construction of [7] would only achieve $1-1/3(r-1)$ in place of $1-1/(r+1)$.)

The idea behind the proof is as follows. We are assuming that we have access to a black box approximate decomposition result: given a K_r-divisible graph G on vertex classes of size n with $\hat{\delta}(G) \geq (\hat{\delta}^{\eta}_{K_r} + \varepsilon)n$ we can obtain an approximate K_r-decomposition that leaves only ηn^2 edges uncovered. We would like to obtain an exact decomposition by 'absorbing' this small remainder. By an *absorber* for a K_r-divisible graph H we mean a graph A_H such that both A_H and $A_H \cup H$ have a K_r-decomposition. For any fixed H we can construct an absorber A_H. But there are far too many possibilities for the remainder H to allow us to reserve individual absorbers for each in advance.

To bridge the gap between the output of the approximate result and the capabilities of our absorbers, we use an iterative absorption approach (see also [7] and [53]). Our guiding principle is that, since we have no control on the remainder if we apply the approximate decomposition result all in one go, we should apply it more carefully. More precisely, we begin by partitioning $V(G)$ at random into a large number of parts $U^1, \ldots, U^k$. Since k is large, $G[U^1, \ldots, U^k]$ still has high minimum degree, and, since the partition is random, each $G[U^i]$ also has high minimum degree. We first reserve a sparse and well structured subgraph J of $G[U^1, \dots, U^k]$, then we obtain an approximate decomposition of $G[U^1, \dots, U^k] - J$ leaving a sparse remainder H. We then use a small number of edges from the $G[U^i]$ to cover all edges of $H \cup J$ by copies of K_r. Let G' be the subgraph of G consisting of those edges not yet used in the approximate decomposition. Then all edges of G' lie in some $G'[U^i]$, and each $G'[U^i]$ has high minimum degree, so we can repeat this argument on each $G'[U^i]$. Suppose that we can iterate in this way until we obtain a partition $W_1 \cup \dots \cup W_m$ of $V(G)$ such that each W_i has size at most some constant M and all edges of G have been used in the approximate decomposition except for those contained entirely within some W_i. Then the remainder is a vertex-disjoint union of graphs $H_1, \dots, H_m$, with each H_i contained within W_i. At this point we have already achieved that the total leftover $H_1 \cup \dots \cup H_m$ has only $O(n)$ edges. More importantly, the set of all possibilities for the graphs H_i has size at most $2^{M^2} m = O(n)$, which is a small enough number that we are able to reserve special purpose absorbers for each of them in advance (i.e. right at the start of the proof).

The above sketch passes over one genuine difficulty. Recall that $H \subseteq G[U^1, \dots, U^k]$ denotes the sparse remainder obtained from the approximate decomposition, which we aim to 'clean up' using a well structured graph J set aside at the beginning of the proof, i.e. we aim to cover all edges of $H \cup J$ with copies of K_r by using a few additional edges from the $G[U^i]$. So consider any vertex $v \in U^1_1$ (recall that $U^i_j = U^i \cap V_j$). In order to cover the edges in $H \cup J$ between v and U^2, we would like to find a perfect K_{r-1}-matching in $N(v) \cap U^2$. However, for this to work, the number of neighbours of v inside each of $U^2_2, \dots, U^2_r$ must be the same, and the analogue must hold with U^2 replaced by any of $U^3, \dots, U^k$. (This is in contrast to [7], where one only needs that the number of leftover edges between v and any of the parts U^i is divisible by r, which is much easier to achieve.) We ensure this balancedness condition by constructing a 'balancing graph' which can be used to transfer a surplus of edges or degrees from one part to another. This 'balancing graph' will be the main ingredient of J. Another difficulty is that whenever we apply the approximate decomposition result, we need to ensure that the graph is K_r-divisible. This means that we need to 'preprocess' the graph at each step of the iteration.

The rest of this chapter is organised as follows. In Section 4.5, we present general purpose embedding lemmas that allow us to find a wide range of desirable structures within our graph. In Section 4.6, we detail the construction of our absorbers. In Section 4.7, we prove some basic properties of random subgraphs and partitions. In Section 4.8, we show how we can assume that our approximate decomposition result produces a remainder with low maximum degree rather than simply a small number of edges. In Section 4.9, we clean up the edges in the remainder using a few additional edges from inside each part of the current partition. However, we assume in this section that our remainder is balanced in the sense described above. In Section 4.10, we describe the balancing operation which ensures that we can make this assumption. Finally, in Section 4.11 we put everything together to prove Theorem 4.1.1.

4.5 Embedding lemmas

Let G be an r-partite graph on $(V_1, \dots, V_r)$ and let $\mathcal{P} = \{U^1, U^2, \dots, U^k\}$ be a partition of $V(G)$. Recall that $U^i_j := U^i \cap V_j$ for each $1 \leq i \leq k$ and each $1 \leq j \leq r$. We say that a graph (or multigraph) H is *$\mathcal{P}$-labelled* if:

(a) every vertex of H is labelled by one of: $\{v\}$ for some $v \in V(G)$; U^i_j for some $1 \leq i \leq k$, $1 \leq j \leq r$ or V_j for some $1 \leq j \leq r$;

(b) the vertices labelled by singletons (called *root vertices*) form an independent set in H, and each $v \in V(G)$ appears as a label $\{v\}$ at most once;

(c) for each $1 \leq j \leq r$, the set of vertices $v \in V(H)$ such that v is labelled L for some $L \subseteq V_j$ forms an independent set in H.

Any vertex which is not a root vertex is called a *free vertex*.

Let H be a $\mathcal{P}$-labelled graph and let H' be a copy of H in G. We say that H' is *compatible with its labelling* if each vertex of H gets mapped to a vertex in its label.

Given a graph H and $U \subseteq V(H)$ with $e(H[U]) = 0$, we define the *degeneracy of H rooted at U* to be the least d for which there is an ordering $v_1, \ldots, v_b$ of the vertices of H such that

- there is an a such that $U = \{v_1, \ldots, v_a\}$ (the ordering of U is unimportant);
- for $a < j \leq b$, v_j is adjacent to at most d of the v_i with $1 \leq i < j$.

The *degeneracy of a $\mathcal{P}$-labelled graph H* is the degeneracy of H rooted at U, where U is the set of root vertices of H.

In the proof of Lemma 4.10.9, we use the following special case of Lemma 5.1 from [7] to find copies of labelled graphs inside a graph G, provided their degeneracy is small. Moreover, this lemma allows us to assume that the subgraph of G used to embed these graphs has low maximum degree.

Lemma 4.5.1. *Let $1/n \ll \eta \ll \varepsilon, 1/d, 1/b \leq 1$ and let G be a graph on n vertices. Suppose that:*

(i) *for each $S \subseteq V(G)$ with $|S| \leq d$, $d_G(S, V(G)) \geq \varepsilon n$.*

Let $m \leq \eta n^2$ and let $H_1, \ldots, H_m$ be labelled graphs such that, for every $1 \leq i \leq m$, every vertex of H_i is labelled $\{v\}$ for some $v \in V(G)$ or labelled by $V(G)$ and that property (b) *above holds for H_i. Moreover, suppose that:*

(ii) *for each $1 \leq i \leq m$, $|H_i| \leq b$;*

(iii) *for each $1 \leq i \leq m$, the degeneracy of H_i (rooted at the set of vertices labelled by singletons) is at most d;*

(iv) *for each $v \in V(G)$, there are at most ηn graphs H_i with some vertex labelled $\{v\}$.*

Then there exist edge-disjoint embeddings $\phi(H_1), \ldots, \phi(H_m)$ of $H_1, \ldots, H_m$ compatible with their labellings such that the subgraph $H := \bigcup_{i=1}^{m} \phi(H_i)$ of G satisfies $\Delta(H) \leq \varepsilon n$.

We will also use the following partite version of the lemma to find copies of $\mathcal{P}$-labelled graphs in an r-partite graph G. We omit the proof since it is very similar to the proof of Lemma 5.1 in [7] (for details, see Appendix A).

Lemma 4.5.2. *Let $1/n \ll \eta \ll \varepsilon, 1/d, 1/b, 1/k, 1/r \leq 1$ and let G be an r-partite graph on $(V_1, \ldots, V_r)$ where $|V_1| = \cdots = |V_r| = n$. Let $\mathcal{P} = \{U^1, \ldots, U^k\}$ be a k-partition of $V(G)$. Suppose that:*

(i) *for each $1 \leq i \leq k$ and each $1 \leq j \leq r$, if $S \subseteq V(G) \setminus V_j$ with $|S| \leq d$ then $d_G(S, U^i_j) \geq \varepsilon |U^i_j|$.*

Let $m \leq \eta n^2$ and let $H_1, \ldots, H_m$ be $\mathcal{P}$-labelled graphs such that the following hold:

(ii) *for each $1 \leq i \leq m$, $|H_i| \leq b$;*

(iii) *for each $1 \leq i \leq m$, the degeneracy of H_i is at most d;*

(iv) *for each $v \in V(G)$, there are at most ηn graphs H_i with some vertex labelled $\{v\}$.*

Then there exist edge-disjoint embeddings $\phi(H_1), \ldots, \phi(H_m)$ of $H_1, \ldots, H_m$ in G which are compatible with their labellings such that $H := \bigcup_{1 \leq i \leq m} \phi(H_i)$ satisfies $\Delta(H) \leq \varepsilon n$.

4.6 Absorbers

Let H be any r-partite graph on the vertex set $V = (V_1, \ldots, V_r)$. An *absorber for H* is a graph A such that both A and $A \cup H$ have K_r-decompositions.

Our aim is to find an absorber for each small K_r-divisible graph H on V. The construction develops ideas in [7]. In particular, we will build the absorber in stages using transformers, introduced below, to move between K_r-divisible graphs.

Let H and H' be vertex-disjoint graphs. An *$(H, H')_r$-transformer* is a graph T which is edge-disjoint from H and H' and is such that both $T \cup H$ and $T \cup H'$ have K_r-decompositions. Note that if H' has a K_r-decomposition, then $T \cup H'$ is an absorber for H. So the idea is that we can use a transformer to transform a given H into a new graph H', then into H'' and so on, until finally we arrive at a graph which has a K_r-decomposition.

Let $V = (V_1, \dots, V_r)$. Throughout this section, given two r-partite graphs H and H' on V, we say that H' is a *partition-respecting copy* of H if there is an isomorphism $f : H \to H'$ such that $f(v) \in V_j$ for every vertex $v \in V(H) \cap V_j$.

Given r-partite graphs H and H' on V, we say that H' is *obtained from H by identifying vertices* if there exists a sequence of r-partite graphs $H_0, \dots, H_s$ on V such that $H_0 = H$, $H_s = H'$ and the following holds. For each $0 \le i < s$, there exists $1 \le j_i \le r$ and vertices $x_i, y_i \in V(H_i) \cap V_{j_i}$ satisfying the following:

(i) $N_{H_i}(x_i) \cap N_{H_i}(y_i) = \emptyset$.

(ii) H_{i+1} is the graph which has vertex set $V(H_i) \setminus \{y_i\}$ and edge set $E(H_i \setminus \{y_i\}) \cup \{vx_i : vy_i \in E(H_i)\}$ (i.e., H_{i+1} is obtained from H_i by identifying the vertices x_i and y_i).

Condition (i) ensures that the identifications do not produce multiple edges. Note that if H and H' are r-partite graphs on V and H' is a partition-respecting copy of a graph obtained from H by identifying vertices then there exists a graph homomorphism $\phi : H \to H'$ that is edge-bijective and maps vertices in V_j to vertices in V_j for each $1 \le j \le r$.

In the following lemma, we find a transformer between a pair of K_r-divisible graphs H and H' whenever H' can be obtained from H by identifying vertices.

Lemma 4.6.1. *Let $r \ge 3$ and $1/n \ll \eta \ll 1/s \ll \varepsilon, 1/b, 1/r \le 1$. Let G be an r-partite graph on $V = (V_1, \dots, V_r)$ with $|V_1| = \dots = |V_r| = n$. Suppose that $\hat{\delta}(G) \ge (1 - 1/(r+1) + \varepsilon)n$. Let H and H' be vertex-disjoint K_r-divisible graphs on V with $|H| \le b$. Suppose further that H' is a partition-respecting copy of a graph obtained from H by identifying vertices. Let $B \subseteq V$ be a set of at most ηn vertices. Then G contains an $(H, H')_r$-transformer T such that $V(T) \cap B \subseteq V(H \cup H')$ and $|T| \le s^2$.*

In our proof of Lemma 4.6.1, we will use the following multipartite asymptotic version of the Hajnal–Szemerédi theorem.

Theorem 4.6.2 ([48] and [58]). *Let $r \geq 2$ and let $1/n \ll \varepsilon, 1/r$. Suppose that G is an r-partite graph on $(V_1, \dots, V_r)$ with $|V_1| = \cdots = |V_r| = n$ and $\hat{\delta}(G) \geq (1 - 1/r + \varepsilon)n$. Then G contains a perfect K_r-matching.*

Proof of Lemma 4.6.1. Let $\phi : H \to H'$ be a graph homomorphism from H to H' that is edge-bijective and maps vertices in V_j to V_j for each $1 \leq j \leq r$.

Let T be any graph defined as follows:

(a) For each $xy \in E(H)$, $Z^{xy} := \{z_j^{xy} : 1 \leq j \leq r \text{ and } x, y \notin V_j\}$ is a set of $r-2$ vertices. For each $x \in V(H)$, let $Z^x := \bigcup_{y \in N_H(x)} Z^{xy}$.

(b) For each $x \in V(H)$, S^x is a set of $(r-1)s$ vertices.

(c) For all distinct $e, e' \in E(H)$ and all distinct $x, x' \in V(H)$, the sets Z^e, $Z^{e'}$, S^x, $S^{x'}$ and $V(H \cup H')$ are disjoint.

(d) $V(T) := V(H) \cup V(H') \cup \bigcup_{e \in E(H)} Z^e \cup \bigcup_{x \in V(H)} S^x$.

(e) $E_H := \{xz : x \in V(H) \text{ and } z \in Z^x\}$.

(f) $E_{H'} := \{\phi(x)z : x \in V(H) \text{ and } z \in Z^x\}$.

(g) $E_Z := \{wz : e \in E(H) \text{ and } w, z \in Z^e\}$.

(h) $E_S := \{xv : x \in V(H) \text{ and } v \in S^x\}$.

(i) $E'_S := \{\phi(x)v : x \in V(H) \text{ and } v \in S^x\}$.

(j) For each $x \in V(H)$, F_1^x is a perfect K_{r-1}-matching on $S^x \cup Z^x$.

(k) For each $x \in V(H)$, F_2^x is a perfect K_{r-1}-matching on S^x.

(l) For each $x \in V(H)$, F_1^x and F_2^x are edge-disjoint.

(m) For each $x \in V(H)$, Z^x is independent in F_1^x.

(n) $E(T) := E_H \cup E_{H'} \cup E_Z \cup E_S \cup E'_S \cup \bigcup_{x \in V(H)} E(F_1^x \cup F_2^x)$.

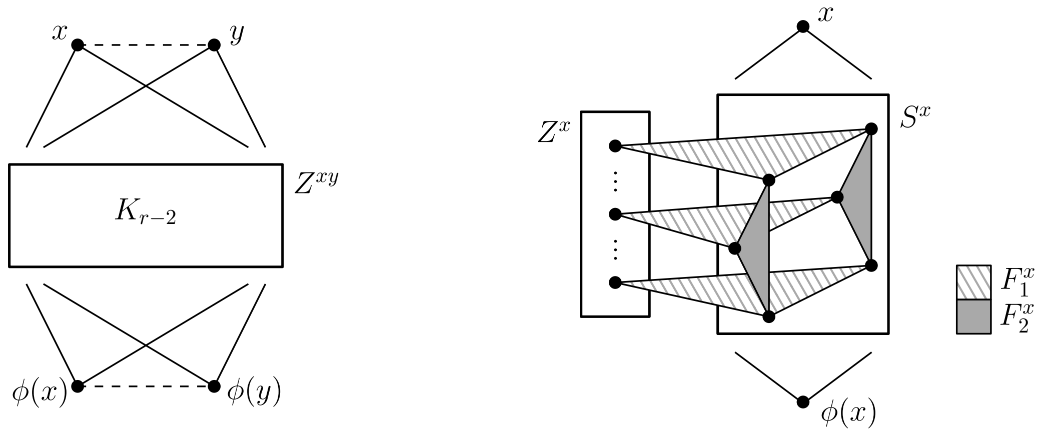

Figure 4.1: Left: Subgraph of T_1 associated with $xy \in E(H)$. Right: Subgraph of T_2 associated with $x \in V(H)$ in the case when $r = 4$.

Then

$$|T| = |H| + |H'| + \sum_{e \in E(H)} |Z^e| + \sum_{x \in V(H)} |S^x| = |H| + |H'| + (r-2)e(H) + (r-1)s|H| \leq s^2.$$

Let T_1 be the subgraph of T with edge set $E_H \cup E_{H'} \cup E_Z$ and let $T_2 := T - T_1$. So $E(T_2) = E_S \cup E'_S \cup \bigcup_{x \in V(H)} E(F_1^x \cup F_2^x)$. In what follows, we will often identify certain subsets of the edge set of T with the subgraphs of T consisting of these edges. For example, we will write $E_S[\{x\}, S^x]$ for the subgraph of T consisting of all the edges in E_S between x and S^x. Note that there are several possibilities for T as we have several choices for the perfect K_{r-1}-matchings in (j) and (k).

Lemma 4.6.1 will follow from Claims 1 and 2 below.

Claim 1: *If T satisfies* (a)–(n)*, then T is an $(H, H')_r$-transformer.*

Proof of Claim 1. Note that $H \cup E_H \cup E_Z$ can be decomposed into $e(H)$ copies of K_r, where each copy of K_r has vertex set $\{x, y\} \cup Z^{xy}$ for some edge $xy \in E(H)$. Similarly, $H' \cup E_{H'} \cup E_Z$ can be decomposed into $e(H)$ copies of K_r.

For each $x \in V(H)$, note that $(E_{H'} \cup E'_S)[\{\phi(x)\}, S^x \cup Z^x] \cup F_1^x$ and $E_S[\{x\}, S^x] \cup F_2^x$

are edge-disjoint and have K_r-decompositions. Since

$$T_2 \cup E_{H'} = \bigcup_{x \in V(H)} \big((E_{H'} \cup E'_S)[\{\phi(x)\}, S^x \cup Z^x] \cup F_1^x\big) \cup \bigcup_{x \in V(H)} \big(E_S[\{x\}, S^x] \cup F_2^x\big),$$

it follows that $T_2 \cup E_{H'}$ has a K_r-decomposition. Similarly, for each vertex $x \in V(H)$, $(E_H \cup E_S)[\{x\}, S^x \cup Z^x] \cup F_1^x$ and $E'_S[\{\phi(x)\}, S^x] \cup F_2^x$ are edge-disjoint and have K_r-decompositions, so $T_2 \cup E_H$ has a K_r-decomposition.

To summarise, $H \cup E_H \cup E_Z$, $H' \cup E_{H'} \cup E_Z$, $T_2 \cup E_H$ and $T_2 \cup E_{H'}$ all have K_r-decompositions. Therefore, $T \cup H = (H \cup E_H \cup E_Z) \cup (T_2 \cup E_{H'})$ has a K_r-decomposition, as does $T \cup H' = (H' \cup E_{H'} \cup E_Z) \cup (T_2 \cup E_H)$. Hence T is an $(H, H')_r$-transformer.

Claim 2: *G contains a graph T satisfying* (a)–(n) *such that $V(T) \cap B \subseteq V(H \cup H')$.*

Proof of Claim 2. We begin by finding a copy of T_1 in G. It will be useful to note that, for any graph T which satisfies (a)–(n), T_1 is r-partite with vertex classes $(V(H \cup H') \cap V_j) \cup \{z_j^{xy} : xy \in E(H) \text{ and } x, y \notin V_j\}$ where $1 \leq j \leq r$. Also, $T[V(H \cup H')]$ is empty and every vertex $z \in V(T_1) \setminus V(H \cup H')$ satisfies

$$d_{T_1}(z) = 2 + (r - 3) + 2 = r + 1. \tag{4.3}$$

So T_1 has degeneracy $r+1$ rooted at $V(H \cup H')$. Since $\hat{\delta}(G) \geq (1 - 1/(r+1) + \varepsilon/2)n + |B|$, we can find a copy of T_1 in G such that $V(T_1) \cap B \subseteq V(H \cup H')$.

We now show that, after fixing T_1, we can extend T_1 to T by finding a copy of T_2. Consider any ordering $x_1, \dots, x_{|H|}$ on the vertices of H. Suppose we have already chosen $S^{x_1}, \dots, S^{x_{q-1}}$, $F_1^{x_1}, \dots, F_1^{x_{q-1}}$ and $F_2^{x_1}, \dots, F_2^{x_{q-1}}$ and we are currently embedding S^{x_q}. Let $B' := B \cup V(T_1) \cup \bigcup_{i=1}^{q-1} S^{x_i}$; that is, B' is the set of vertices that are unavailable for S^{x_q}, either because they have been used previously or they lie in B. Note that $|B'| \leq |T| + |B| \leq 2\eta n$. We will choose suitable vertices for S^{x_q} in the common neighbourhood of x_q and $\phi(x_q)$.

To simplify notation, we write $x := x_q$ and assume that $x \in V_1$ (the argument is identical in the other cases). Choose a set $V' \subseteq (N_G(x) \cap N_G(\phi(x))) \setminus B'$ which is maximal subject to $|V'_2| = \cdots = |V'_r|$ (recall that $V'_j = V' \cap V_j$). Note that for each $2 \leq j \leq r$, we have

$$|V'_j| \geq (1 - 1/(r+1) + \varepsilon)n - (1/(r+1) - \varepsilon)n - |B'| \geq (1 - 2/(r+1))n.$$

Let $n' := |V'_2|$. For every $2 \leq j \leq r$ and every $v \in V(G) \setminus V_j$, we have

$$d_G(v, V'_j) \geq n' - (1/(r+1) - \varepsilon)n \geq (1 - 1/(r-1) + \varepsilon)n'. \tag{4.4}$$

Roughly speaking, we will choose S^x as a random subset of V'. For each $2 \leq j \leq r$, choose each vertex of V'_j independently with probability $p := (1 + \varepsilon/8)s/n'$ and let S'_j be the set of chosen vertices. Note that, for each j, $\mathbb{E}(|S'_j|) = n'p = (1+\varepsilon/8)s$. We can apply Lemma 4.2.2 to see that

$$\begin{aligned}\mathbb{P}(||S'_j| - (1 + \varepsilon/8)s| \geq \varepsilon s/8) &\leq \mathbb{P}(||S'_j| - (1 + \varepsilon/8)s| \geq \varepsilon \mathbb{E}(|S'_j|)/10) \\ &\leq 2e^{-\varepsilon^2 s/300} \leq 1/4(r-1). \end{aligned}\tag{4.5}$$

Given a vertex $v \in V(G)$ and $2 \leq j \leq r$ such that $v \notin V_j$, note that

$$\mathbb{E}(d_G(v, S'_j)) \overset{(4.4)}{\geq} (1 - 1/(r-1) + \varepsilon)n'p > (1 - 1/(r-1) + \varepsilon)s.$$

We will say that a vertex $v \in V(G)$ is *bad* if there exists $2 \leq j \leq r$ such that $v \notin V_j$ and $d_G(v, S'_j) < (1 - 1/(r-1) + 3\varepsilon/4)s$, that is, the degree of v in S'_j is lower than expected. We can again apply Lemma 4.2.2 to see that

$$\begin{aligned}\mathbb{P}(d_G(v, S'_j) \leq (1 - 1/(r-1) + 3\varepsilon/4)s) &\leq \mathbb{P}(|d_G(v, S'_j) - \mathbb{E}(d_G(v, S'_j))| \geq \varepsilon s/4) \\ &\leq \mathbb{P}(|d_G(v, S'_j) - \mathbb{E}(d_G(v, S'_j))| \geq \varepsilon \mathbb{E}(d_G(v, S'_j))/10) \\ &\leq 2e^{-\varepsilon^2 s/600}.\end{aligned}$$

So $\mathbb{P}(v \text{ is bad}) \leq 2(r-1)e^{-\varepsilon^2 s/600} \leq e^{-s^{1/2}}$. Let $S' := \bigcup_{j=2}^r S'_j$. We say that the set S' is *bad* if $S' \cup Z^x$ contains a bad vertex. We have

$$\begin{aligned}
\mathbb{P}(S' \text{ is bad}) &\leq \sum_{v \in V'} \mathbb{P}(v \in S' \text{ and } v \text{ is bad}) + \sum_{v \in Z^x} \mathbb{P}(v \text{ is bad}) \\
&= \sum_{v \in V'} \mathbb{P}(v \in S')\mathbb{P}(v \text{ is bad}) + \sum_{v \in Z^x} \mathbb{P}(v \text{ is bad}) \\
&\leq (n'p + (b-1)(r-2))e^{-s^{1/2}} \leq 2se^{-s^{1/2}} \leq 1/4. \qquad (4.6)
\end{aligned}$$

We apply (4.5) and (4.6) to see that with probability at least 1/2, the set S' chosen in this way is not bad and, for each $2 \leq j \leq r$, we have $s \leq |S'_j| \leq (1+\varepsilon/4)s$. Choose one such set S'. Delete at most $\varepsilon s/4$ vertices from each S'_j to obtain sets S^x_j satisfying $|S^x_2| = \cdots = |S^x_r| = s$. Let $S^x := \bigcup_{j=2}^r S^x_j$. Since S' was not bad, for each $2 \leq j \leq r$ and each vertex $v \in (S^x \cup Z^x) \setminus V_j$,

$$d_G(v, S^x_j) \geq (1 - 1/(r-1) + 3\varepsilon/4)s - \varepsilon s/4 = (1 - 1/(r-1) + \varepsilon/2)s. \qquad (4.7)$$

We now show that we can find F^x_1 and F^x_2 satisfying (j)–(m). Let $G^x := G[Z^x \cup S^x] - G[Z^x]$. Note that G^x is a balanced $(r-1)$-partite graph with vertex classes of size n_x where $s \leq n_x \leq s + (r-2)(b-1)/(r-1) < s + b$. Using (4.7), we see that

$$\hat{\delta}(G^x) \geq (1 - 1/(r-1) + \varepsilon/2)s \geq (1 - 1/(r-1) + \varepsilon/3)n_x.$$

So, using Theorem 4.6.2, we can find a perfect K_{r-1}-matching F^x_1 in G^x. Finally, let $G' := G - F^x_1$ and use (4.7) to see that

$$\hat{\delta}(G'[S^x]) \geq (1 - 1/(r-1) + \varepsilon/3)s.$$

So we can again apply Theorem 4.6.2, to find a perfect K_{r-1}-matching F^x_2 in $G'[S^x]$. In this way, we find a copy of T satisfying (a)–(n) such that $V(T) \cap B \subseteq V(H \cup H')$. $\square$

We now construct our absorber by combining several suitable transformers.

Let H be an r-partite multigraph on $V = (V_1, \dots, V_r)$ and let $xy \in E(H)$. A *K_r-expansion of xy* is defined as follows. Consider a copy F_{xy} of K_r on vertex set $\{u_1, \dots, u_r\}$ such that $u_j \in V_j \setminus V(H)$ for all $1 \leq j \leq r$. Let j_1, j_2 be such that $x \in V_{j_1}$ and $y \in V_{j_2}$. Delete xy from H and $u_{j_1}u_{j_2}$ from F_{xy} and add edges joining x to u_{j_2} and joining y to u_{j_1}. Let $H_{\rm exp}$ be the graph obtained by K_r-expanding every edge of H, where the F_{xy} are chosen to be vertex-disjoint for different edges $xy \in E(H)$.

Fact 4.6.3. *Suppose that the graph H' is obtained from a graph H by K_r-expanding the edge $xy \in E(H)$ as above. Then the graph obtained from H' by identifying x and u_{j_1} is H with a copy of K_r attached to x.*

Let $h \in \mathbb{N}$. We define a graph M_h as follows. Take a copy of K_r on V (consisting of one vertex in each V_j) and replace each edge by h multiedges. Let M denote the resulting multigraph. Let $M_h := M_{\rm exp}$ be the graph obtained by K_r-expanding every edge of M. We have $|M_h| = r + hr\binom{r}{2}$. Note that M_h has degeneracy $r-1$. To see this, list all vertices in $V(M)$ (in any order) followed by the vertices in $V(M_h \setminus M)$ (in any order).

We will now apply Lemma 4.6.1 twice in order to find an $(H, M_h)_r$-transformer in G.

Lemma 4.6.4. *Let $r \geq 3$ and $1/n \ll \eta \ll 1/s \ll \varepsilon, 1/b, 1/r \leq 1$. Let G be an r-partite graph on $V = (V_1, \dots, V_r)$ with $|V_1| = \dots = |V_r| = n$. Suppose that $\hat{\delta}(G) \geq (1 - 1/(r+1) + \varepsilon)n$. Let H be a K_r-divisible graph on V with $|H| \leq b$. Let $h := e(H)/\binom{r}{2}$. Let M'_h be a partition-respecting copy of M_h on V which is vertex-disjoint from H. Let $B \subseteq V$ be a set of at most ηn vertices. Then G contains an $(H, M'_h)_r$-transformer T such that $V(T) \cap B \subseteq V(H \cup M'_h)$ and $|T| \leq 3s^2$.*

Proof. We construct a graph $H_{\rm att}$ as follows. Start with the graph H. For each edge of H, arbitrarily choose one of it endpoints x and attach a copy of K_r (found in $G \setminus ((V(H \cup M'_h) \cup B) \setminus \{x\})$) to x. The copies of K_r should be chosen to be vertex-disjoint outside $V(H)$. Write $H_{\rm att}$ for the resulting graph. Let $H'_{\rm exp}$ be a partition-respecting copy of $H_{\rm exp}$ in $G \setminus (V(H_{\rm att} \cup M'_h) \cup B)$. Note that we are able to find these graphs since both have degeneracy $r-1$ and $\hat{\delta}(G) \geq (1-1/(r+1)+\varepsilon)n$.

By Fact 4.6.3, H_{att} is a partition-respecting copy of a graph obtained from H'_{exp} by identifying vertices, and this is also the case for M'_h. To see the latter, for each $1 \leq j \leq r$, identify all vertices of H'_{exp} lying in V_j. (We are able to do this since these vertices are non-adjacent with disjoint neighbourhoods.)

Apply Lemma 4.6.1 to find an $(H'_{\mathrm{exp}}, H_{\mathrm{att}})_r$-transformer T' in $G - M'_h$ such that $V(T') \cap B \subseteq V(H)$ and $|T'| \leq s^2$. Then apply Lemma 4.6.1 again to find an $(H'_{\mathrm{exp}}, M'_h)_r$-transformer T'' in $G - (H_{\mathrm{att}} \cup T')$ such that $V(T'') \cap B \subseteq V(M'_h)$ and $|T''| \leq s^2$.

Let $T := T' \cup T'' \cup H'_{\mathrm{exp}} \cup (H_{\mathrm{att}} - H)$. Then T is edge-disjoint from $H \cup M'_h$. Note that

$$T \cup H = (T' \cup H_{\mathrm{att}}) \cup (T'' \cup H'_{\mathrm{exp}}) \quad \text{and}$$
$$T \cup M'_h = (T' \cup H'_{\mathrm{exp}}) \cup (T'' \cup M'_h) \cup (H_{\mathrm{att}} - H),$$

both of which have K_r-decompositions. Therefore T is an $(H, M'_h)_r$-transformer. Moreover, $|T| \leq 3s^2$. Finally, observe that $V(T) \cap B = V(T' \cup T'' \cup H_{\mathrm{att}}) \cap B \subseteq V(H \cup M'_h)$. □

We now have all of the necessary tools to find an absorber for H in G.

Lemma 4.6.5. *Let $r \geq 3$ and let $1/n \ll \eta \ll 1/s \ll \varepsilon, 1/b, 1/r \leq 1$. Let G be an r-partite graph on $V = (V_1, \dots, V_r)$ with $|V_1| = \dots = |V_r| = n$. Suppose that $\hat{\delta}(G) \geq (1-1/(r+1)+\varepsilon)n$. Let H be a K_r-divisible graph on V with $|H| \leq b$. Let $B \subseteq V$ be a set of at most ηn vertices. Then G contains an absorber A for H such that $V(A) \cap B \subseteq V(H)$ and $|A| \leq s^3$.*

Proof. Let $h := e(H)/\binom{r}{2}$. Let $G' := G \setminus (V(H) \cup B)$. Write hK_r for the graph consisting of h vertex-disjoint copies of K_r. Since $\hat{\delta}(G') \geq (1-1/(r+1)+\varepsilon/2)n$, we can choose vertex-disjoint (partition-respecting) copies of M_h and hK_r in G' (and call these M_h and hK_r again). Use Lemma 4.6.4 to find an $(H, M_h)_r$-transformer T' in $G - hK_r$ such that $V(T') \cap B \subseteq V(H)$ and $|T'| \leq 3s^2$. Apply Lemma 4.6.4 again to find an

$(hK_r, M_h)_r$-transformer T'' in $G - (H \cup T')$ which avoids B and satisfies $|T''| \leq 3s^2$. It is easy to see that $T := T' \cup T'' \cup M_h$ is an $(H, hK_r)_r$-transformer.

Let $A := T \cup hK_r$. Note that both A and $A \cup H = (T \cup H) \cup hK_r$ have K_r-decompositions. So A is an absorber for H. Moreover, $V(A) \cap B \subseteq V(T') \cap B \subseteq V(H)$ and $|A| \leq s^3$. □

4.6.1 Absorbing sets

Let $\mathcal{H}$ be a collection of graphs on the vertex set $V = (V_1, \ldots, V_r)$. We say that $\mathcal{A}$ is an *absorbing set for* $\mathcal{H}$ if $\mathcal{A}$ is a collection of edge-disjoint graphs and, for every $H \in \mathcal{H}$ and every K_r-divisible subgraph $H' \subseteq H$, there is a distinct $A_{H'} \in \mathcal{A}$ such that $A_{H'}$ is an absorber for H'.

Lemma 4.6.6. *Let $r \geq 3$ and $1/n \ll \eta \ll \varepsilon, 1/b, 1/r \leq 1$. Let G be an r-partite graph on $V = (V_1, \ldots, V_r)$ with $|V_1| = \cdots = |V_r| = n$. Suppose that $\hat{\delta}(G) \geq (1 - 1/(r+1) + \varepsilon)n$. Let $m \leq \eta n^2$ and let $\mathcal{H}$ be a collection of m edge-disjoint graphs on $V = (V_1, \ldots, V_r)$ such that each vertex $v \in V$ appears in at most ηn of the elements of $\mathcal{H}$ and $|H| \leq b$ for each $H \in \mathcal{H}$. Then G contains an absorbing set $\mathcal{A}$ for $\mathcal{H}$ such that $\Delta(\bigcup \mathcal{A}) \leq \varepsilon n$.*

We repeatedly use Lemma 4.6.5 and aim to avoid any vertices which have been used too often.

Proof. Enumerate the K_r-divisible subgraphs of all $H \in \mathcal{H}$ as $H_1, \ldots, H_{m'}$. Note that each $H \in \mathcal{H}$ can have at most $2^{e(H)} \leq 2^{\binom{b}{2}}$ K_r-divisible subgraphs so $m' \leq 2^{\binom{b}{2}} \eta n^2$. For each $v \in V(G)$ and each $0 \leq j \leq m$, let $s(v, j)$ be the number of indices $1 \leq i \leq j$ such that $v \in V(H_i)$. Note that $s(v, j) \leq 2^{\binom{b}{2}} \eta n$.

Let $s \in \mathbb{N}$ be such that $\eta \ll 1/s \ll \varepsilon, 1/b, 1/r$. Suppose that we have already found absorbers $A_1, \ldots, A_{j-1}$ for $H_1, \ldots, H_{j-1}$ respectively such that $|A_i| \leq s^3$, for all $1 \leq i \leq j-1$, and, for every $v \in V(G)$,

$$d_{G_{j-1}}(v) \leq \eta^{1/2} n + (s(v, j-1) + 1)s^3, \tag{4.8}$$

where $G_{j-1} := \bigcup_{1\le i\le j-1} A_i$. We show that we can find an absorber A_j for H_j in $G - G_{j-1}$ which satisfies (4.8) with j replacing $j-1$.

Let $B := \{v \in V(G) : d_{G_{j-1}}(v) \ge \eta^{1/2}n\}$. We have

$$|B| \le \frac{2e(G_{j-1})}{\eta^{1/2}n} \le \frac{2m'\binom{s^3}{2}}{\eta^{1/2}n} \le \frac{2^{\binom{b}{2}+1}\eta n^2 s^6}{\eta^{1/2}n} \le \eta^{1/3}n.$$

We have

$$\begin{aligned}\hat{\delta}(G - G_{j-1}) &\overset{(4.8)}{\ge} (1 - 1/(r+1) + \varepsilon)n - \eta^{1/2}n - (s(v, j-1) + 1)s^3 \\ &\ge (1 - 1/(r+1) + \varepsilon)n - \eta^{1/2}n - (2^{\binom{b}{2}}\eta n + 1)s^3 > (1 - 1/(r+1) + \varepsilon/2)n.\end{aligned}$$

So we can apply Lemma 4.6.5 (with $\varepsilon/2$, $\eta^{1/3}$, $G - G_{j-1}$ and H_j playing the roles of ε, η, G and H) to find an absorber A_j for H_j in $G - G_{j-1}$ such that $V(A_j) \cap B \subseteq V(H_j)$ and $|A_j| \le s^3$.

We now check that (4.8) holds with j replacing $j-1$. If $v \in V(G) \setminus B$, this is clear. Suppose then that $v \in B$. If $v \in V(A_j)$, then $v \in V(H_j)$ and $s(v, j) = s(v, j-1) + 1$. So in all cases,

$$d_{G_j}(v) \le \eta^{1/2}n + (s(v, j) + 1)s^3.$$

Continue in this way until we have found an absorber A_i for each H_i. Then $\mathcal{A} := \{A_i : 1 \le i \le m'\}$ is an absorbing set. Using (4.8),

$$\Delta\left(\bigcup \mathcal{A}\right) = \Delta(G_{m'}) \le \eta^{1/2}n + (2^{\binom{b}{2}}\eta n + 1)s^3 \le \varepsilon n,$$

as required. □

4.7 Partitions and random subgraphs

In this section we consider a sequence $\mathcal{P}_1, \dots, \mathcal{P}_\ell$ of successively finer partitions which will underlie our iterative absorption process. We will also construct corresponding sparse quasirandom subgraphs R_i which will be used to 'smooth out' the leftover from the approximate decomposition in each step of the process.

Let G be an r-partite graph on $(V_1, \dots, V_r)$. An (α, k, δ)-*partition* for G on $(V_1, \dots, V_r)$ is a partition $\mathcal{P} = \{U^1, \dots, U^k\}$ of $V(G)$ such that the following hold:

(Pa1) for each $1 \le j \le r$, $\{U^i_j : 1 \le i \le k\}$ is an equitable partition of V_j (recall that $U^i_j = U^i \cap V_j$);

(Pa2) for each $1 \le i \le k$, $|U^i_1| = \dots = |U^i_r|$;

(Pa3) for each $v \in V(G)$, each $1 \le i \le k$ and each $1 \le j \le r$,

$$|d_G(v, U^i_j) - d_G(v, V_j)/k| < \alpha |U^i_j|;$$

(Pa4) for each $1 \le i \le k$, each $1 \le j \le r$ and each $v \notin V_j$, $d_G(v, U^i_j) \ge \delta |U^i_j|$.

We say that $\mathcal{P} = \{U^1, \dots, U^k\}$ is a k-*partition* if it satisfies (Pa1) and (Pa2).

The following proposition guarantees a $(n^{-1/3}/2, k, \delta - n^{-1/3}/2)$-partition of any sufficiently large balanced r-partite graph G with $\hat{\delta}(G) \ge \delta n$. To prove this result, it suffices to consider an equitable partition $U^1_j, U^2_j, \dots, U^k_j$ of V_j chosen uniformly at random (with $|U^1_j| \le \dots \le |U^k_j|$).

Proposition 4.7.1. *Let $k, r \in \mathbb{N}$. There exists n_0 such that if $n \ge n_0$ and G is any r-partite graph on $(V_1, \dots, V_r)$ with $|V_1| = \dots = |V_r| = n$ and $\hat{\delta}(G) \ge \delta n$, then G has a $(\nu, k, \delta - \nu)$-partition, where $\nu := n^{-1/3}/2$.*

We say that $\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_\ell$ is an (α, k, δ, m)-*partition sequence* for G on $(V_1, \dots, V_r)$ if, writing $\mathcal{P}_0 := \{V(G)\}$, (S1) for each $1 \leq i \leq \ell$, $\mathcal{P}_i$ refines $\mathcal{P}_{i-1}$;

(S2) for each $1 \leq i \leq \ell$ and each $W \in \mathcal{P}_{i-1}$, $\mathcal{P}_i[W]$ is an (α, k, δ)-partition for $G[W]$;

(S3) for each $1 \leq i \leq \ell$, all $1 \leq j_1, j_2, j_3 \leq r$ with $j_1 \neq j_2, j_3$, each $W \in \mathcal{P}_{i-1}$, each $U \in \mathcal{P}_i[W]$ and each $v \in W_{j_1}$,

$$|d_G(v, U_{j_2}) - d_G(v, U_{j_3})| < \alpha |U_{j_1}|;$$

(S4) for each $U \in \mathcal{P}_\ell$ and each $1 \leq j \leq r$, $|U_j| = m$ or $m - 1$.

Note that (S2) and (Pa2) together imply that $|U_{j_1}| = |U_{j_2}|$ for each $1 \leq i \leq \ell$, each $U \in \mathcal{P}_i$ and all $1 \leq j_1, j_2 \leq r$.

By successive applications of Proposition 4.7.1, we immediately obtain the following result which guarantees the existence of a suitable partition sequence (for details see Appendix A).

Lemma 4.7.2. *Let $k, r \in \mathbb{N}$ with $k \geq 2$ and let $0 < \alpha < 1$. There exists m_0 such that, for all $m' \geq m_0$, any K_r-divisible graph G on $(V_1, \ldots, V_r)$ with $|V_1| = \cdots = |V_r| = n \geq km'$ and $\hat{\delta}(G) \geq \delta n$ has an $(\alpha, k, \delta - \alpha, m)$-partition sequence for some $m' \leq m \leq km'$.*

Suppose that we are given a k-partition $\mathcal{P}$ of G. The following proposition finds a quasirandom spanning subgraph R of G so that each vertex in R has roughly the expected number of neighbours in each set $U \in \mathcal{P}$. The proof is an easy application of Lemma 4.2.1.

Proposition 4.7.3. *Let $1/n \ll \alpha, \rho, 1/k, 1/r \leq 1$. Let G be an r-partite graph on $(V_1, \ldots, V_r)$ with $|V_1| = \cdots = |V_r| = n$. Suppose that $\mathcal{P}$ is a k-partition for G. Let $\mathcal{S}$ be a collection of at most n^2 subsets of $V(G)$. Then there exists $R \subseteq G[\mathcal{P}]$ such that for all $1 \leq j \leq r$, all distinct $x, y \in V(G)$, all $U \in \mathcal{P}$ and all $S \in \mathcal{S}$:*

- $|d_R(x, U_j) - \rho d_{G[\mathcal{P}]}(x, U_j)| < \alpha |U_j|$;

- $|d_R(\{x,y\},U_j) - \rho^2 d_{G[\mathcal{P}]}(\{x,y\},U_j)| < \alpha|U_j|$;
- $|d_G(y, N_R(x,U_j)) - \rho d_G(y, N_{G[\mathcal{P}]}(x,U_j))| < \alpha|U_j|$;
- $|d_R(y,S_j) - \rho d_{G[\mathcal{P}]}(y,S_j)| < \alpha n$.

We need to reserve some quasirandom subgraphs R_i of G at the start of our proof, whilst the graph G is still almost balanced with respect to the partition sequence. We will add the edges of R_i back after finding an approximate decomposition of $G[\mathcal{P}_i]$ in order to assume the leftover from this approximate decomposition is quasirandom. The next lemma gives us suitable subgraphs for R_i.

Lemma 4.7.4. *Let $1/m \ll \alpha \ll \rho, 1/k, 1/r \leq 1$. Let G be an r-partite graph on $(V_1, \ldots, V_r)$ with $|V_1| = \cdots = |V_r|$. Suppose that $\mathcal{P}_1, \ldots, \mathcal{P}_\ell$ is a $(1,k,0,m)$-partition sequence for G. Let $\mathcal{P}_0 := \{V(G)\}$ and, for each $0 \leq q \leq \ell$, let $G_q := G[\mathcal{P}_q]$. Then there exists a sequence of graphs $R_1, \ldots, R_\ell$ such that $R_q \subseteq G_q - G_{q-1}$ for each q and the following holds. For all $1 \leq q \leq \ell$, all $1 \leq j \leq r$, all $W \in \mathcal{P}_{q-1}$, all distinct $x, y \in W$ and all $U \in \mathcal{P}_q[W]$:*

(i) $|d_{R_q}(x,U_j) - \rho d_{G_q}(x,U_j)| < \alpha|U_j|$;

(ii) $|d_{R_q}(\{x,y\},U_j) - \rho^2 d_{G_q}(\{x,y\},U_j)| < \alpha|U_j|$;

(iii) $d_{G'_{q+1}}(y, N_{R_q}(x,U_j)) \geq \rho d_{G_{q+1}}(y, N_{G_q}(x,U_j)) - 3\rho^2|U_j|$, *where* $G'_{q+1} := G_{q+1} - R_{q+1}$ *if* $q \leq \ell - 1$, $G'_{\ell+1} := G$ *and* $G_{\ell+1} := G$.

Proof. For $1 \leq q \leq \ell$, we say that the sequence of graphs $R_1, \ldots, R_q$ is *good* if $R_i \subseteq G_i - G_{i-1}$ and for all $1 \leq i \leq q$, all $1 \leq j \leq r$, all $W \in \mathcal{P}_{i-1}$, all distinct $x, y \in W$ and all $U \in \mathcal{P}_i[W]$:

(a) (i) and (ii) hold (with q replaced by i);

(b) $|d_{G_{i+1}}(y, N_{R_i}(x,U_j)) - \rho d_{G_{i+1}}(y, N_{G_i}(x,U_j))| < \alpha|U_j|$;

(c) if $i \leq q-1$, $d_{R_{i+1}}(y, N_{R_i}(x,U_j)) < \rho d_{G_{i+1}}(y, N_{R_i}(x,U_j)) + \alpha|U_j|$.

Suppose $1 \le q \le \ell$ and we have found a good sequence of graphs $R_1, \dots, R_{q-1}$. We will find R_q such that $R_1, \dots, R_q$ is good. Let $W \in \mathcal{P}_{q-1}$, let $\mathcal{S}_1$ be the empty set and, if $q \ge 2$, let $W' \in \mathcal{P}_{q-2}$ be such that $W \subseteq W'$ and let $\mathcal{S}_q := \{N_{R_{q-1}}(x, W) : x \in W'\}$. Apply Proposition 4.7.3 (with $|W|/r$, $G_{q+1}[W]$, $\mathcal{P}_q[W]$ and $\mathcal{S}_q$ playing the roles of n, G, $\mathcal{P}$ and $\mathcal{S}$) to find $R_W \subseteq G_{q+1}[W][\mathcal{P}_q[W]] = G_q[W]$ such that:

$$\begin{aligned} |d_{R_W}(x, U_j) - \rho d_{G_q}(x, U_j)| &< \alpha|U_j|, \\ |d_{R_W}(\{x, y\}, U_j) - \rho^2 d_{G_q}(\{x, y\}, U_j)| &< \alpha|U_j|, \\ |d_{G_{q+1}}(y, N_{R_W}(x, U_j)) - \rho d_{G_{q+1}}(y, N_{G_q}(x, U_j))| &< \alpha|U_j|, \\ |d_{R_W}(y, S_j) - \rho d_{G_q}(y, S_j)| &< \alpha|W_j|, \end{aligned} \tag{4.9}$$

for all $1 \le j \le r$, all distinct $x, y \in W$, all $U \in \mathcal{P}_q[W]$ and all $S \in \mathcal{S}_q$. Set $R_q := \bigcup_{W \in \mathcal{P}_{q-1}} R_W$. It is clear that $R_1, \dots, R_q$ satisfy (a) and (b). We now check that (c) holds when $1 \le i = q-1$. Let $1 \le j \le r$, $W \in \mathcal{P}_{q-2}$, $x, y \in W$ be distinct and $U \in \mathcal{P}_{q-1}[W]$. If $y \notin U$, then $d_{R_q}(y, U_j) = 0$ and so (c) holds. If $y \in U$, then $d_{R_q}(y, N_{R_{q-1}}(x, U)) = d_{R_U}(y, N_{R_{q-1}}(x, U))$ and (c) follows by replacing W and S by U and $N_{R_{q-1}}(x, U)$ in property (4.9). So $R_1, \dots, R_q$ is good.

So G contains a good sequence of graphs $R_1, \dots, R_\ell$. We will now check that this sequence also satisfies (iii). If $q = \ell$, this follows immediately from (b). Let $1 \le q < \ell$, $1 \le j \le r$, $W \in \mathcal{P}_{q-1}$, $x, y \in W$ be distinct and $U \in \mathcal{P}_q[W]$. We have

$$\begin{aligned} d_{R_{q+1}}(y, N_{R_q}(x, U_j)) &\overset{\text{(c)}}{<} \rho d_{G_{q+1}}(y, N_{R_q}(x, U_j)) + \alpha|U_j| \\ &\overset{\text{(b)}}{<} \rho^2 d_{G_{q+1}}(y, N_{G_q}(x, U_j)) + (\alpha\rho + \alpha)|U_j| \le 2\rho^2|U_j|. \end{aligned}$$

Therefore,

$$\begin{aligned} d_{G'_{q+1}}(y, N_{R_q}(x, U_j)) &= d_{G_{q+1}}(y, N_{R_q}(x, U_j)) - d_{R_{q+1}}(y, N_{R_q}(x, U_j)) \\ &\overset{\text{(b)}}{\ge} \rho d_{G_{q+1}}(y, N_{G_q}(x, U_j)) - 3\rho^2|U_j|. \end{aligned}$$

So $R_1, \dots, R_\ell$ satisfy (i)–(iii). □

We apply Lemma 4.7.4 when $\mathcal{P}_1, \dots, \mathcal{P}_\ell$ is an $(\alpha, k, 1-1/r+\varepsilon, m)$-partition sequence for G to obtain the following result. For details of the proof, see Appendix A.

Corollary 4.7.5. *Let $1/m \ll \alpha \ll \rho, 1/k \ll \varepsilon, 1/r \leq 1$. Let G be a K_r-divisible graph on $(V_1, \dots, V_r)$ with $|V_1| = \dots = |V_r|$. Suppose that $\mathcal{P}_1, \dots, \mathcal{P}_\ell$ is an $(\alpha, k, 1-1/r+\varepsilon, m)$-partition sequence for G. Let $\mathcal{P}_0 := \{V(G)\}$ and $G_q := G[\mathcal{P}_q]$ for $0 \leq q \leq \ell$. There exists a sequence of graphs $R_1, \dots, R_\ell$ such that $R_q \subseteq G_q - G_{q-1}$ for each $1 \leq q \leq \ell$ and the following holds. For all $1 \leq q \leq \ell$, all $1 \leq j, j' \leq r$, all $W \in \mathcal{P}_{q-1}$, all distinct $x, y \in W$ and all $U, U' \in \mathcal{P}_q[W]$:*

(i) $d_{R_q}(x, U_j) < \rho d_{G_q}(x, U_j) + \alpha |U_j|$;

(ii) $d_{R_q}(\{x, y\}, U_j) < (\rho^2 + \alpha)|U_j|$;

(iii) *if* $x \notin U \cup U' \cup V_j \cup V_{j'}$, $|d_{R_q}(x, U_j) - d_{R_q}(x, U'_{j'})| < 3\alpha |U_j|$;

(iv) *if* $x \notin U$, $y \in U$ *and* $x, y \notin V_j$, *then*

$$d_{G'_{q+1}}(y, N_{R_q}(x, U_j)) \geq \rho(1 - 1/(r-1)) d_{G_q}(x, U_j) + \rho^{5/4} |U_j|,$$

where $G'_{q+1} := G_{q+1} - R_{q+1}$ *if* $q \leq \ell - 1$ *and* $G'_{\ell+1} := G$.

4.8 A remainder of low maximum degree

The aim of this section is to prove the following lemma which lets us assume that the remainder of G after finding an η-approximate decomposition has small maximum degree.

Lemma 4.8.1. *Let $1/n \ll \alpha \ll \eta \ll \gamma \ll \varepsilon < 1/r < 1$. Let G be an r-partite graph on $(V_1, \dots, V_r)$ with $|V_1| = \dots = |V_r| = n$ and $\hat{\delta}(G) \geq (\hat{\delta}^{\eta}_{K_r} + \varepsilon)n$. Suppose also that, for all $1 \leq j_1, j_2 \leq r$ and every $v \notin V_{j_1} \cup V_{j_2}$,*

$$|d_G(v, V_{j_1}) - d_G(v, V_{j_2})| < \alpha n. \tag{4.10}$$

Then there exists $H \subseteq G$ such that $G - H$ has a K_r-decomposition and $\Delta(H) \leq \gamma n$.

Our strategy for the proof of Lemma 4.8.1 is as follows. We first remove a sparse random subgraph H of G and then choose an η-approximate K_r-decomposition of $G-H$. Now consider the remainder R obtained from G by deleting all edges in the copies of K_r in this decomposition. Suppose that v is a vertex whose degree in R is too high. Our aim will be to find a K_{r-1}-matching in a sparse random subgraph whose vertex set is the neighbourhood of v in G. Each vertex in this random subgraph sees, on average, at most $\rho d_G(v)/(r-1) \ll (1 - 1/(r-1) + \varepsilon)d_G(v)/(r-1)$ vertices in each other part, so Theorem 4.6.2 alone is of no use. But Theorem 4.6.2 can be combined with the Regularity lemma in order to find the desired K_r-matching.

4.8.1 Regularity

In this section, we introduce a version of the Regularity lemma which we will use to prove Lemma 4.8.1.

Let G be a bipartite graph on (A, B). For non-empty sets $X \subseteq A$, $Y \subseteq B$, we define the *density of* $G[X, Y]$ to be $d_G(X, Y) := e_G(X, Y)/|X||Y|$. Let $\varepsilon > 0$. We say that G is *ε-regular* if for all sets $X \subseteq A$ and $Y \subseteq B$ with $|X| \geq \varepsilon|A|$ and $|Y| \geq \varepsilon|B|$ we have

$$|d_G(A, B) - d_G(X, Y)| < \varepsilon.$$

The following simple result follows immediately from this definition.

Proposition 4.8.2. *Suppose that $0 < \varepsilon \leq \alpha \leq 1/2$. Let G be a bipartite graph on (A, B). Suppose that G is ε-regular with density d. If $A' \subseteq A, B' \subseteq B$ with $|A'| \geq \alpha|A|$ and $|B'| \geq \alpha|B|$ then $G[A', B']$ is ε/α-regular and has density greater than $d - \varepsilon$.*

Proposition 4.8.2 shows that regularity is robust, that is, it is not destroyed by deleting a small number of vertices. The next observation allows us to delete a small number of edges at each vertex and still maintain regularity. The proof again follows from the definition.

Proposition 4.8.3. *Let $n \in \mathbb{N}$ and let $0 < \gamma \ll \varepsilon \leq 1$. Let G be a bipartite graph on (A, B) with $|A| = |B| = n$. Suppose that G is ε-regular with density d. Let $H \subseteq G$ with $\Delta(H) \leq \gamma n$ and let $G' := G - H$. Then G' is 2ε-regular and has density greater than $d - \varepsilon/2$.*

The following proposition takes a graph G on $(V_1, \dots, V_r)$ where each pair of vertex classes induces an ε-regular pair and allows us to find a K_r-matching covering most of the vertices in G. Part (i) follows from Proposition 4.8.2 and the definition of regularity. For (ii), apply (i) repeatedly until only $\lceil \varepsilon^{1/r} n \rceil$ vertices remain uncovered in each V_j.

Proposition 4.8.4. *Let $1/n \ll \varepsilon \ll d, 1/r \leq 1$. Let G be an r-partite graph on $(V_1, \dots, V_r)$ with $|V_1| = \dots = |V_r| = n$. Suppose that, for all $1 \leq j_1 < j_2 \leq r$, the graph $G[V_{j_1}, V_{j_2}]$ is ε-regular with density at least d.*

(i) *For each $1 \leq j \leq r$, let $W_j \subseteq V_j$ with $|W_j| = \lceil \varepsilon^{1/r} n \rceil$. Then $G[W_1, \dots, W_r]$ contains a copy of K_r.*

(ii) *The graph G contains a K_r-matching which covers all but at most $2r\varepsilon^{1/r} n$ vertices of G.*

We will use a version of Szemerédi's Regularity lemma [73] stated for r-partite graphs. It is proved in the same way as the non-partite degree version.

Lemma 4.8.5 (Degree form of the r-partite Regularity lemma)**.** *Let $0 < \varepsilon < 1$ and $k_0, r \in \mathbb{N}$. Then there is an $N = N(\varepsilon, k_0, r)$ such that the following holds for every $0 \leq d < 1$ and for every r-partite graph G on $(V_1, \dots, V_r)$ with $|V_1| = \dots = |V_r| = n \geq N$. There exists a partition $\mathcal{P} = \{U^0, \dots, U^k\}$ of $V(G)$, $m \in \mathbb{N}$ and a spanning subgraph G' of G satisfying the following:*

(i) *$k_0 \leq k \leq N$;*

(ii) *for each* $1 \leq j \leq r$, $|U_j^0| \leq \varepsilon n$;

(iii) *for each* $1 \leq i \leq k$ *and each* $1 \leq j \leq r$, $|U_j^i| = m$;

(iv) *for each* $1 \leq j \leq r$ *and each* $v \in V(G)$, $d_{G'}(v, V_j) > d_G(v, V_j) - (d + \varepsilon)n$;

(v) *for all but at most* εk^2 *pairs* $U_{j_1}^{i_1}, U_{j_2}^{i_2}$ *where* $1 \leq i_1, i_2 \leq k$ *and* $1 \leq j_1 < j_2 \leq r$, *the graph* $G'[U_{j_1}^{i_1}, U_{j_2}^{i_2}]$ *is* ε*-regular and has density either* 0 *or* $> d$.

We define the *reduced graph* R as follows. The vertex set of R is the set of clusters $\{U_j^i : 1 \leq i \leq k \text{ and } 1 \leq j \leq r\}$. For each $U, U' \in V(R)$, UU' is an edge of R if the subgraph $G'[U, U']$ is ε-regular and has density greater than d. Note that R is a balanced r-partite graph with vertex classes $W_j := \{U_j^i : 1 \leq i \leq k\}$ for $1 \leq j \leq r$. The following simple proposition relates the minimum degree of G and the minimum degree of R.

Proposition 4.8.6. *Suppose that* $0 < 2\varepsilon \leq d \leq c/2$. *Let* G *be an* r*-partite graph on* $(V_1, \dots, V_r)$ *with* $|V_1| = \cdots = |V_r| = n$ *and* $\hat{\delta}(G) \geq cn$. *Suppose that* G *has a partition* $\mathcal{P} = \{U^0, \dots, U^k\}$ *and a subgraph* $G' \subseteq G$ *as given by Lemma 4.8.5. Let* R *be the reduced graph of* G. *Then* $\hat{\delta}(R) \geq (c - 2d)k$.

Proof. Let $W_j := \{U_j^i : 1 \leq i \leq k\}$, where $1 \leq j \leq r$, be the vertex classes of R. Let $1 \leq j_1, j_2 \leq r$ be such that $j_1 \neq j_2$. Consider any $U \in W_{j_1}$ and let $x \in U$. We observe that x has neighbours in at least $(d_{G'}(x, V_{j_2}) - |U_{j_2}^0|)/m$ different clusters in W_{j_2} in G'. By Lemma 4.8.5(v) and the definition of R, U is a neighbour of each of these clusters in R. So we have

$$d_R(U, W_{j_2}) \geq (d_{G'}(x, V_{j_2}) - |U_{j_2}^0|)/m \geq (d_{G'}(x, V_{j_2}) - \varepsilon n)/m.$$

From Lemma 4.8.5(iv), we also have that

$$d_{G'}(x, V_{j_2}) > d_G(x, V_{j_2}) - (d + \varepsilon)n \geq (c - (d + \varepsilon))n.$$

Combining these inequalities, we obtain that

$$d_R(U, W_{j_2}) \geq (c - (d + 2\varepsilon))n/m \geq (c - 2d)k$$

and hence $\hat{\delta}(R) \geq (c - 2d)k$. □

4.8.2 Degree reduction

At the beginning of our proof of Lemma 4.8.1, we will reserve a random subgraph H of G. Proposition 4.8.8 below ensures that we can partition the neighbourhood of each vertex so that H induces ε-regular graphs between these parts. In our proof of Proposition 4.8.8, we will use the following well-known result for which we omit the proof.

Proposition 4.8.7. *Let $1/n \ll \varepsilon \ll d, \rho \leq 1$. Let G be a bipartite graph on (A, B) with $|A| = |B| = n$. Suppose that G is ε-regular with density at least d. Let H be a graph formed by taking each edge of G independently with probability ρ. Then, with probability at least $1 - 1/n^2$, H is 4ε-regular with density at least $\rho d/2$.*

Proposition 4.8.8. *Let $1/n \ll \alpha \ll 1/N \ll 1/k_0 \leq \varepsilon^* \ll d \ll \rho < \varepsilon, 1/r < 1$. Let G be an r-partite graph on $(V_1, \dots, V_r)$ with $|V_1| = \dots = |V_r| = n$ and $\hat{\delta}(G) \geq (1 - 1/r + \varepsilon)n$. Suppose that for all $1 \leq j_1, j_2 \leq r$ and every $v \notin V_{j_1} \cup V_{j_2}$, $|d_G(v, V_{j_1}) - d_G(v, V_{j_2})| < \alpha n$. Then there exists $H \subseteq G$ satisfying the following properties:*

(i) *For each $1 \leq j \leq r$ and each $v \in V(G)$, $|d_H(v, V_j) - \rho d_G(v, V_j)| < \alpha n$. In particular, for any $1 \leq j_1, j_2 \leq r$ such that $v \notin V_{j_1} \cup V_{j_2}$, $|d_H(v, V_{j_1}) - d_H(v, V_{j_2})| < 3\alpha n$.*

(ii) *For each vertex $v \in V(G)$, there exists a partition $\mathcal{P}(v) = \{U^0(v), \dots, U^{k_v}(v)\}$ of $N_G(v)$ and $m_v \in \mathbb{N}$ such that:*

- *$k_0 \leq k_v \leq N$;*
- *for each $1 \leq j \leq r$, $|U^0_j(v)| \leq \varepsilon^* n$;*

- *for each $1 \le i \le k_v$ and each $1 \le j \le r$ such that $v \notin V_j$, $|U_j^i(v)| = m_v$;*
- *for each $1 \le i \le k_v$ and all $1 \le j_1 < j_2 \le r$ such that $v \notin V_{j_1} \cup V_{j_2}$, the graph $H[U_{j_1}^i(v), U_{j_2}^i(v)]$ is ε^*-regular with density greater than d.*

Roughly speaking, (ii) says that for each $v \in V(G)$ the reduced graph of $H[N_G(v)]$ has a perfect K_{r-1}-matching.

Proof. Let H be the graph formed by taking each edge of G independently with probability ρ. For each $1 \le j \le r$ and each $v \in V(G)$, Lemma 4.2.1 gives

$$\mathbb{P}(|d_H(v, V_j) - \rho d_H(v, V_j)| \ge \alpha n) \le 2e^{-2\alpha^2 n} < 1/rn^2.$$

So the probability that there exist $1 \le j \le r$ and $v \in V(G)$ such that $|d_H(v, V_j) - \rho d_G(v, V_j)| \ge \alpha n$ is at most $rn/rn^2 = 1/n$. Let $1 \le j_1, j_2 \le r$. Note that if $v \notin V_{j_1} \cup V_{j_2}$ and $|d_H(v, V_j) - \rho d_G(v, V_j)| < \alpha n$ for $j = j_1, j_2$, then

$$|d_H(v, V_{j_1}) - d_H(v, V_{j_2})| < |\rho d_G(v, V_{j_1}) - \rho d_G(v, V_{j_2})| + 2\alpha n < 3\alpha n.$$

So H satisfies (i) with probability at least $1 - 1/n$.

We will now show that H satisfies (ii) with probability at least $1/2$. We find partitions of the neighbourhood of each vertex $v \in V(G)$ as follows. To simplify notation, we will assume that $v \in V_1$ (the argument is identical for the other cases). For all $2 \le j_1, j_2 \le r$, we have $|d_G(v, V_{j_1}) - d_G(v, V_{j_2})| < \alpha n$. So, there exists n_v and, for each $2 \le j \le r$, a subset $V_j(v) \subseteq N_G(v, V_j)$ such that $|V_j(v)| > d_G(v, V_j) - \alpha n$ and

$$|V_j(v)| = n_v \ge \hat{\delta}(G) \ge (1 - 1/r)n.$$

Let G_v denote the balanced $(r-1)$-partite graph $G[V_2(v), \ldots, V_r(v)]$. Note that

$$\hat{\delta}(G_v) \ge n_v - \frac{n}{r} + \varepsilon n \ge \left(1 - \frac{1}{r-1} + \varepsilon\right) n_v. \tag{4.11}$$

Apply Lemma 4.8.5 (with $\varepsilon^*/4$, $2d/\rho$, k_0 and G_v playing the roles of ε, d, k_0 and G) to find a partition $\mathcal{Q}(v) = \{W^0(v), \ldots, W^{k_v}(v)\}$ of $V(G_v)$ satisfying properties (i)–(v) of Lemma 4.8.5. Let $m_v := |W_2^1(v)|$. Let R_v denote the reduced graph corresponding to this partition. Proposition 4.8.6 together with (4.11) implies that

$$\hat{\delta}(R_v) \geq (1 - 1/(r-1) + \varepsilon/2)k_v.$$

So we can use Theorem 4.6.2 to find a perfect K_{r-1}-matching M_v in R_v. Let $U^0(v) := W^0(v) \cup (N_G(v) \setminus V(G_v))$. Note that for each $2 \leq j \leq r$, $|U_j^0| < |W_j^0| + \alpha n \leq \varepsilon^* n$. Let $\mathcal{P}(v) := \{U^0(v), \ldots, U^{k_v}(v)\}$ be a partition of $N_G(v)$ which is chosen such that, for each $1 \leq i \leq k_v$, $\{U_j^i(v) : 2 \leq j \leq r\}$ induces a copy of K_{r-1} in M_v. By the definition of R_v, for each $1 \leq i \leq k_v$ and all $2 \leq j_1 < j_2 \leq r$, the graph $G[U_{j_1}^i(v), U_{j_2}^i(v)]$ is $\varepsilon^*/4$-regular with density greater than $2d/\rho$.

Fix $1 \leq i \leq k_v$ and $2 \leq j_1 < j_2 \leq r$. Proposition 4.8.7 (with m_v, $\varepsilon^*/4$, $2d/\rho$, $G[U_{j_1}^i(v), U_{j_2}^i(v)]$ and $H[U_{j_1}^i(v), U_{j_2}^i(v)]$ playing the roles of n, ε, d, G and H) gives that $H[U_{j_1}^i(v), U_{j_2}^i(v)]$ is ε^*-regular and has density greater than d with probability at least $1 - 1/m_v^2$.

We require the graph $H[U_{j_1}^i(v), U_{j_2}^i(v)]$ to be ε^*-regular with density greater than d for every edge $U_{j_1}^i(v)U_{j_2}^i(v) \in E(M_v)$. There are k_v choices for i and, for each i, there are $\binom{r-1}{2}$ choices for j_1 and j_2. So the probability that, for fixed $v \in V(G)$, there exists an edge $U_{j_1}^i(v)U_{j_2}^i(v) \in E(M_v)$ which fails to be ε^*-regular with density greater than d is at most

$$k_v r^2 \frac{1}{m_v^2} < \frac{1}{2rn}.$$

We multiply this probability by rn for each of the rn choices of v to see that H satisfies property (ii) with probability at least $1 - rn/2rn = 1/2$. Hence, the graph H satisfies both (i) and (ii) with probability at least $1/2 - 1/n > 0$. So we can choose such a graph H. □

In order to find an η-approximate K_r-decomposition in a graph G, we would like to use the definition of $\hat{\delta}^{\eta}_{K_r}$ which requires G to be K_r-divisible. The next proposition shows that, provided that $d_G(v, V_{j_1})$ is close to $d_G(v, V_{j_2})$ for all $1 \leq j_1, j_2 \leq r$ and $v \notin V_{j_1} \cup V_{j_2}$, G can be made K_r-divisible by removing only a small number of edges.

Proposition 4.8.9. *Let $1/n \ll \alpha \ll \gamma \ll 1/r < 1$. Let G be an r-partite graph on $(V_1, \dots, V_r)$ with $|V_1| = \cdots = |V_r| = n$ and $\hat{\delta}(G) \geq (1/2 + 2\gamma/r)n$. Suppose that, for all $1 \leq j_1, j_2 \leq r$ and every $v \in V(G) \setminus (V_{j_1} \cup V_{j_2})$, $|d_G(v, V_{j_1}) - d_G(v, V_{j_2})| < \alpha n$. Then there exists $H \subseteq G$ such that $G - H$ is K_r-divisible and $\Delta(H) \leq \gamma n$.*

To prove Proposition 4.8.9, we require the following result whose proof is based on the Max-Flow-Min-Cut theorem.

Proposition 4.8.10. *Suppose that $1/n \ll \alpha \ll \xi \ll 1$. Let G be a bipartite graph on (A, B) with $|A| = |B| = n$. Suppose that $\delta(G) \geq (1/2 + 4\xi)n$. For every vertex $v \in V(G)$, let $n_v \in \mathbb{N}$ be such that $(\xi - \alpha)n \leq n_v \leq (\xi + \alpha)n$ and such that $\sum_{a \in A} n_a = \sum_{b \in B} n_b$. Then G contains a spanning graph G' such that $d_{G'}(v) = n_v$ for every $v \in V(G)$.*

Proof. We will use the Max-Flow-Min-Cut theorem. Orient every edge of G towards B and give each edge capacity one. Add a source vertex s^* which is attached to every vertex $a \in A$ by an edge of capacity n_a. Add a sink vertex t^* which is attached to every vertex in $b \in B$ by an edge of capacity n_b. Let $c_0 := \sum_{a \in A} n_a = \sum_{b \in B} n_b$. Note that an integer-valued c_0-flow corresponds to the desired spanning graph G' in G. So, by the Max-Flow-Min-Cut theorem, it suffices to show that every cut has capacity at least c_0.

Consider a minimal cut C. Let $S \subseteq A$ be the set of all vertices $a \in A$ for which $s^*a \notin C$ and let $T \subseteq B$ be the set of all $b \in B$ for which $bt^* \notin C$. Let $S' := A \setminus S$ and $T' := B \setminus T$. Then C has capacity

$$c := \sum_{s \in S'} n_s + e_G(S, T) + \sum_{t \in T'} n_t.$$

First suppose that $|S| \geq (1/2 - 2\xi)n$. In this case, since $\delta(G) \geq (1/2 + 4\xi)n$, each vertex in T receives at least $2\xi n$ edges from S. So

$$c \geq \sum_{t\in T'} n_t + 2|T|\xi n \geq \sum_{t\in T'} n_t + |T|(\xi+\alpha)n \geq c_0.$$

A similar argument works if $|T| \geq (1/2-2\xi)n$. Suppose then that $|S|, |T| < (1/2-2\xi)n$. Then $|S'|, |T'| > (1/2+2\xi)n$ and

$$c \geq \sum_{s\in S'} n_s + \sum_{t\in T'} n_t \geq (|S'|+|T'|)(\xi-\alpha)n > (n+4\xi n)(\xi-\alpha)n \geq (\xi+\alpha)n^2 \geq c_0,$$

as required. □

We now use Proposition 4.8.10 to prove Proposition 4.8.9.

Proof of Proposition 4.8.9. For each $v \in V(G)$, let

$$m_v := \min\{d_G(v, V_j) : 1 \leq j \leq r \text{ with } v \notin V_j\}.$$

For each $1 \leq j \leq r$ and each $v \notin V_j$, let $a_{v,j} := d_G(v, V_j) - m_v$. Note that,

$$0 \leq a_{v,j} < \alpha n. \tag{4.12}$$

For each $1 \leq j \leq r$, let $N_j := \sum_{v\in V_j} m_v$. We have, for any $1 \leq j_1, j_2 \leq r$,

$$\begin{aligned} |N_{j_1} - N_{j_2}| &= \left| \sum_{v\in V_{j_1}} (d_G(v, V_{j_2}) - a_{v,j_2}) - \sum_{v\in V_{j_2}} (d_G(v, V_{j_1}) - a_{v,j_1}) \right| \\ &= \left| \sum_{v\in V_{j_1}} a_{v,j_2} - \sum_{v\in V_{j_2}} a_{v,j_1} \right| \overset{(4.12)}{<} \alpha n^2. \end{aligned} \tag{4.13}$$

Let $N := \min\{N_j : 1 \leq j \leq r\}$ and, for each $1 \leq j \leq r$, let $M_j := N_j - N$. Note that (4.13) implies $0 \leq M_j < \alpha n^2$. For each $1 \leq j \leq r$ and each $v \in V_j$, choose $p_v \in \mathbb{N}$ to be as equal as possible such that $\sum_{v\in V_j} p_v = M_j$. Then

$$0 \le p_v < \alpha n + 1. \tag{4.14}$$

Let $\xi := \gamma/2r$. For each $1 \le j \le r$ and each $v \notin V_j$, let

$$n_{v,j} := \lceil \xi n \rceil + a_{v,j} + p_v.$$

Using (4.12) and (4.14), we see that,

$$\xi n \le n_{v,j} \le (\xi + 3\alpha)n. \tag{4.15}$$

We will consider each pair $1 \le j_1 < j_2 \le r$ separately and choose $H_{j_1,j_2} = H[V_{j_1}, V_{j_2}]$. Fix $1 \le j_1 < j_2 \le r$ and observe that,

$$\begin{aligned}\sum_{v\in V_{j_1}} n_{v,j_2} &= \sum_{v\in V_{j_1}} (\lceil \xi n \rceil + a_{v,j_2} + p_v) = \lceil \xi n \rceil n + \sum_{v\in V_{j_1}} a_{v,j_2} + M_{j_1} \\ &= \lceil \xi n \rceil n + M_{j_1} + \sum_{v\in V_{j_1}} (d_G(v, V_{j_2}) - m_v) = \lceil \xi n \rceil n + M_{j_1} + e_G(V_{j_1}, V_{j_2}) - N_{j_1} \\ &= \lceil \xi n \rceil n - N + e_G(V_{j_1}, V_{j_2}) = \sum_{v\in V_{j_2}} n_{v,j_1}.\end{aligned}$$

Let $G_{j_1,j_2} := G[V_{j_1}, V_{j_2}]$ and note that $\delta(G_{j_1,j_2}) \ge (1/2 + 4\xi)n$. Apply Proposition 4.8.10 (with 3α, ξ, G_{j_1,j_2}, V_{j_1} and V_{j_2} playing the roles of α, ξ, G, A and B) to find $H_{j_1,j_2} \subseteq G_{j_1,j_2}$ such that $d_{H_{j_1,j_2}}(v) = n_{v,j_2}$ for every $v \in V_{j_1}$ and $d_{H_{j_1,j_2}}(v) = n_{v,j_1}$ for every $v \in V_{j_2}$.

Let $H := \bigcup_{1\le j_1<j_2\le r} H_{j_1,j_2}$. By (4.15), we have $\Delta(H) \le 2r\xi n = \gamma n$. For any $1 \le j \le r$ and any $v \notin V_j$, we have

$$\begin{aligned}d_{G-H}(v, V_j) &= d_G(v, V_j) - d_H(v, V_j) = d_G(v, V_j) - n_{v,j} \\ &= d_G(v, V_j) - \lceil \xi n \rceil - d_G(v, V_j) + m_v - p_v = m_v - p_v - \lceil \xi n \rceil.\end{aligned}$$

So $G - H$ is K_r-divisible. $\square$

We now have all the necessary tools to prove Lemma 4.8.1. This lemma finds an approximate K_r-decomposition which covers all but at most γn edges at any vertex.

Proof of Lemma 4.8.1. The lemma trivially holds if $r = 2$, so we may assume that $r \geq 3$. In particular, by Proposition 4.3.1, $\hat{\delta}(G) \geq (1-1/(r+1)+\varepsilon/2)n$. Choose constants N, k_0, ε^*, d and ρ satisfying

$$\eta \ll 1/N \ll 1/k_0 \leq \varepsilon^* \ll d \ll \rho \ll \gamma.$$

Apply Proposition 4.8.8 to find a subgraph $H_1 \subseteq G$ satisfying properties (i)–(ii).

Let $G_1 := G - H_1$. Using (4.10) and that H_1 satisfies Proposition 4.8.8(i), for all $1 \leq j_1, j_2 \leq r$ and each $v \notin V_{j_1} \cup V_{j_2}$,

$$\begin{aligned}|d_{G_1}(v, V_{j_1}) - d_{G_1}(v, V_{j_2})| &\leq |d_G(v, V_{j_1}) - d_G(v, V_{j_2})| + |d_{H_1}(v, V_{j_1}) - d_{H_1}(v, V_{j_2})| \\ &< \alpha n + 3\alpha n = 4\alpha n.\end{aligned}$$

Note also that $\hat{\delta}(G_1) \geq 3n/4$. So we can apply Proposition 4.8.9 (with G_1, 4α and $\gamma/2$ playing the roles of G, α and γ) to obtain $H_2 \subseteq G_1$ such that $G_1 - H_2$ is K_r-divisible and $\Delta(H_2) \leq \gamma n/2$. Then $\hat{\delta}(G_1 - H_2) \geq (\hat{\delta}^\eta_{K_r} + \varepsilon/2)n$, so we can find an η-approximate K_r-decomposition $\mathcal{F}$ of $G_1 - H_2$.

Let $G_2 := G_1 - H_2 - \bigcup \mathcal{F}$ be the graph consisting of all the remaining edges in $G_1 - H_2$. Let

$$B := \{v \in V(G) : d_{G_2}(v) > \eta^{1/2} n\}.$$

Note that

$$|B| \leq 2e(G_2)/\eta^{1/2} n \leq 2\eta^{1/2} n. \tag{4.16}$$

Let $\mathcal{F}_1 := \{F \in \mathcal{F} : F \cap B = \emptyset\}$ and let $G_3 := G - \bigcup \mathcal{F}_1$. If $v \in B$, then $N_{G_3}(v) = N_G(v)$. Suppose that $v \notin B$. For any $u \in B$, at most one copy of K_r in $\mathcal{F} \setminus \mathcal{F}_1$ can contain both

u and v. So there can be at most $(r-1)|B|$ edges in $\bigcup(\mathcal{F} \setminus \mathcal{F}_1)$ that are incident to v and so

$$\begin{aligned} d_{G_3}(v) &\leq d_{H_1}(v) + d_{H_2}(v) + d_{G_2}(v) + (r-1)|B| \\ &\leq (r-1)(\rho+\alpha)n + \gamma n/2 + \eta^{1/2}n + 2(r-1)\eta^{1/2}n \leq \gamma n. \end{aligned} \tag{4.17}$$

Label the vertices of $B = \{v_1, v_2, \dots, v_{|B|}\}$. We will use copies of K_r to cover most of the edges at each vertex v_i in turn. We do this by finding a K_{r-1}-matching M_i in $H_1[N_{G_3}(v_i)] = H_1[N_G(v_i)]$ in turn for each i. Suppose that we are currently considering $v := v_i$ and let $\mathcal{M} := \bigcup_{1 \leq j < i} M_j$. To simplify notation, we will assume that $v \in V_1$ (the proof in the other cases is identical).

Let $\mathcal{P}(v) = \{U^0(v), \dots, U^{k_v}(v)\}$ be a partition of $N_G(v)$ satisfying Proposition 4.8.8(ii). We can choose a partition $\mathcal{Q}(v) = \{W^0(v), \dots, W^{k_v}(v)\}$ of $N_G(v)$ and $m'_v \geq m_v - |B|$ such that, for each $1 \leq i \leq k_v$:

- $W^i(v) \subseteq U^i(v)$;
- $W^i(v) \cap B = \emptyset$;
- for each $2 \leq j \leq r$, $|W^i_j(v)| = m'_v$.

Note that, using (4.16), $|W^0(v)| \leq |U^0(v)| + |B|k_v r \leq r(\varepsilon^* n + 2\eta^{1/2}nk_v) \leq 2\varepsilon^* rn$.

By Proposition 4.8.8(ii), for each $1 \leq i \leq k_v$ and all $2 \leq j_1 < j_2 \leq r$, the graph $H_1[U^i_{j_1}(v), U^i_{j_2}(v)]$ is ε^*-regular with density greater than d. So Proposition 4.8.2 implies that $H_1[W^i_{j_1}(v), W^i_{j_2}(v)]$ is $2\varepsilon^*$-regular with density greater than $d/2$. Let $H'_1 := H_1 - \mathcal{M}$. Using (4.16), we have $\Delta(\mathcal{M}[W^i_{j_1}(v), W^i_{j_2}(v)]) \leq |B| \leq \eta^{1/3}m'_v$. So we can apply Proposition 4.8.3 (with m'_v, $\eta^{1/3}$ and $2\varepsilon^*$ playing the roles of n, γ and ε) to see that $H'_1[W^i_{j_1}(v), W^i_{j_2}(v)]$ is $4\varepsilon^*$-regular with density greater than $d/3$.

We use Proposition 4.8.4 (with m'_v, $4\varepsilon^*$, $d/3$ and $r-1$ playing the roles of n, ε, d and r) to find a K_{r-1}-matching covering all but at most $2(r-1)(4\varepsilon^*)^{1/(r-1)}m'_v$ vertices in $H'_1[W^i(v)]$ for each $1 \leq i \leq k_v$. Write M_i for the union of these K_{r-1}-matchings over

$1 \leq i \leq k_v$. Note that M_i covers all but at most

$$|W^0(v)| + 2(r-1)(4\varepsilon^*)^{1/(r-1)}m'_v k_v \leq 2\varepsilon^* rn + 2(r-1)(4\varepsilon^*)^{1/(r-1)}n \leq \gamma n \tag{4.18}$$

vertices in $N_G(v)$.

Continue to find edge-disjoint $M_1, \dots, M_{|B|}$. For each $1 \leq i \leq |B|$, $M'_i := \{v_i \cup K : K \in M_i\}$ is an edge-disjoint collection of copies of K_r in G_3 covering all but at most γn edges at v_i in G. Write $\mathcal{M}' := \bigcup_{1\leq i\leq |B|} M'_i$ and let $H := G_3 - \bigcup \mathcal{M}' = G - \bigcup(\mathcal{F}_1 \cup \mathcal{M}')$. Then $G - H = \bigcup(\mathcal{F}_1 \cup \mathcal{M}')$ has a K_r-decomposition and $\Delta(H) \leq \gamma n$, by (4.17) and (4.18). □

4.9 Covering a pseudorandom remainder between vertex classes

After applying Lemma 4.8.1, we are left with a graph H such that $H[\mathcal{P}]$ has low maximum degree. We will add a suitable quasirandom graph R to H to be able to assume that the remainder $H' = R \cup H$ is actually quasirandom. The results in this section will allow us to cover any remaining edges in $H'[\mathcal{P}]$ using only a small number of edges from $H' - H'[\mathcal{P}]$. This is done by finding, for each $x \in V(G)$, suitable vertex-disjoint copies of K_{r-1} inside $H' - H'[\mathcal{P}]$ such that each copy of K_{r-1} forms a copy of K_r together with the edges incident to x in $H'[\mathcal{P}]$.

Lemma 4.9.1. *Let $r \geq 2$ and $1/n \ll 1/k, 1/r, \rho \leq 1$. Let G be an r-partite graph on $(V_1, \dots, V_r)$ with $|V_1| = \dots = |V_r| = n$. Let $q \leq krn$ and let $W^1, \dots, W^q \subseteq V(G)$. Suppose that:*

(i) *for each $1 \leq i \leq q$, there exists $1 \leq j_i \leq r$ and $n_i \in \mathbb{N}$ such that, for each $1 \leq j \leq r$, $|W^i_j| = 0$ if $j = j_i$ and $|W^i_j| = n_i$ otherwise;*

(ii) *for each $1 \leq i \leq q$, $\hat{\delta}(G[W^i]) \geq (1 - 1/(r-1))n_i + 9kr^2\rho^{3/2}n$;*

(iii) *for all* $1 \leq i_1 < i_2 \leq q$, $|W^{i_1} \cap W^{i_2}| \leq 2r\rho^2 n$*;*

(iv) *each* $v \in V(G)$ *is contained in at most* $2k\rho n$ *of the* W^i*.*

Then there exist edge-disjoint $T_1, \dots, T_q$ *in* G *such that each* T_i *is a perfect* K_{r-1}*-matching in* $G[W^i]$*.*

The W^i in Lemma 4.9.1 will play the role of vertex neighbourhoods later on. The proof of Lemma 4.9.1 is similar to that of Lemma 10.7 in [7], we include it here for completeness. We will use the following result.

Proposition 4.9.2 (Jain, see [68, Lemma 8])**.** *Let* $X_1, \dots, X_n$ *be Bernoulli random variables such that, for any* $1 \leq s \leq n$ *and any* $x_1, \dots, x_{s-1} \in \{0, 1\}$,

$$\mathbb{P}(X_s = 1 \mid X_1 = x_1, \dots, X_{s-1} = x_{s-1}) \leq p.$$

Let $X = \sum_{s=1}^{n} X_i$ *and let* $B \sim B(n, p)$*. Then* $\mathbb{P}(X \geq a) \leq \mathbb{P}(B \geq a)$ *for any* $a \geq 0$*.*

Proof of Lemma 4.9.1. Set $t := \lceil 8kr\rho^{3/2}n \rceil$. Let $G_i := G[W^i]$ for $1 \leq i \leq q$. Suppose we have already found $T_1, \dots T_{s-1}$ for some $1 \leq s \leq q$. We find T_s as follows.

Let $H_{s-1} := \bigcup_{i=1}^{s-1} T_i$ and $G'_s := G_s - H_{s-1}[W^s]$. If $\Delta(H_{s-1}[W^s]) > (r-2)\rho^{3/2}n$, let $T'_1, \dots, T'_t$ be empty graphs on W^s. Otherwise, (ii) implies

$$\hat{\delta}(G'_s) \geq \big(1 - \frac{1}{r-1}\big)n_s + 8kr^2\rho^{3/2}n \geq \big(1 - \frac{1}{r-1} + \rho^{3/2}\big)n_s + (r-2)(t-1)$$

and we can greedily find t edge-disjoint perfect K_{r-1}-matchings $T'_1, \dots, T'_t$ in G'_s using Theorem 4.6.2. In either case, pick $1 \leq i \leq t$ uniformly at random and set $T_s := T'_i$. It suffices to show that, with positive probability,

$$\Delta(H_{s-1}[W^s]) \leq (r-2)\rho^{3/2}n \qquad \text{for all } 1 \leq s \leq q.$$

Consider any $1 \leq i \leq q$ and any $w \in W^i$. For $1 \leq s \leq q$, let $Y_s^{i,w}$ be the indicator function of the event that T_s contains an edge incident to w in G_i. Let $X^{i,w} := \sum_{s=1}^{q} Y_s^{i,w}$.

Note $d_{H_q}(w, W^i) \leq (r-2)X^{i,w}$. So it suffices to show that, with positive probability, $X^{i,w} \leq \rho^{3/2}n$ for all $1 \leq i \leq q$ and all $w \in W^i$.

Fix $1 \leq i \leq q$ and $w \in W^i$. Let $J^{i,w}$ be the set of indices $s \neq i$ such that $w \in W^s$; (iv) implies $|J^{i,w}| < 2k\rho n$. If $s \notin J^{i,w} \cup \{i\}$, then $w \notin W^s$ and $Y_s^{i,w} = 0$. So

$$X^{i,w} \leq 1 + \sum_{s \in J^{i,w}} Y_s^{i,w}. \tag{4.19}$$

Let $s_1 < \cdots < s_{|J^{i,w}|}$ be an enumeration of $J^{i,w}$. For any $b \leq |J^{i,w}|$, note that

$$d_{G_{s_b}}(w, W^i) \leq |W^i \cap W^{s_b}| \overset{\text{(iii)}}{\leq} 2r\rho^2 n.$$

So at most $2r\rho^2 n$ of the subgraphs T'_j that we picked in G'_{s_b} contain an edge incident to w in G_i. Thus

$$\mathbb{P}(Y_{s_b}^{i,w} = 1 \mid Y_{s_1}^{i,w} = y_1, \ldots, Y_{s_{b-1}}^{i,w} = y_{b-1}) \leq 2r\rho^2 n/t \leq \rho^{1/2}/4k$$

for all $y_1, \ldots, y_{b-1} \in \{0,1\}$ and $1 \leq b \leq |J^{i,w}|$. Let $B \sim B(|J^{i,w}|, \rho^{1/2}/4k)$. Using Proposition 4.9.2, Lemma 4.2.1 and that $|J^{i,w}| \leq 2k\rho n$, we see that

$$\begin{aligned}\mathbb{P}(X^{i,w} > \rho^{3/2}n) &\overset{(4.19)}{\leq} \mathbb{P}(\sum_{s \in J^{i,w}} Y_s^{i,w} > 3\rho^{3/2}n/4) \leq \mathbb{P}(B > 3\rho^{3/2}n/4) \\ &\leq \mathbb{P}(|B - \mathbb{E}(B)| > \rho^{3/2}n/4) \leq 2e^{-\rho^2 n/16k}.\end{aligned}$$

There are at most $qrn \leq kr^2n^2$ pairs (i, w), so there is a choice of $T_1, \ldots, T_q$ such that $X^{i,w} \leq \rho^{3/2}n$ for all $1 \leq i \leq q$ and all $w \in W^i$. □

The following is an immediate consequence of Lemma 4.9.1.

Corollary 4.9.3. *Let $r \geq 2$ and $1/n \ll 1/k, 1/r, \rho \leq 1$. Let G be an r-partite graph on $(V_1, \ldots, V_r)$ with $|V_1| = \cdots = |V_r| = n$. Let $U, W \subseteq V(G)$ be disjoint with $|W_1| = \cdots = |W_r| \geq \lfloor n/k \rfloor$. Suppose the following hold:*

(i) *for all* $1 \leq j_1, j_2 \leq r$ *and all* $x \in U \setminus (V_{j_1} \cup V_{j_2})$, $d_G(x, W_{j_1}) = d_G(x, W_{j_2})$*;*

(ii) *for all* $1 \leq j \leq r$ *and all* $x \in U \setminus U_j$, $\hat{\delta}(G[N_G(x, W)]) \geq (1 - 1/(r-1))d_G(x, W_j) + 9kr\rho^{3/2}|W|$*;*

(iii) *for all distinct* $x, x' \in U$, $|N_G(x, W) \cap N_G(x', W)| \leq 2\rho^2|W|$*;*

(iv) *for all* $y \in W$, $d_G(y, U) \leq 2k\rho|W_1|$*.*

Then there exists $G_W \subseteq G[W]$ *such that* $G[U, W] \cup G_W$ *has a* K_r*-decomposition and* $\Delta(G_W) \leq 2kr\rho|W_1|$*.*

Proof. Let $q := |U|$ and let $u^1, \dots, u^q$ be an enumeration of U. For each $1 \leq i \leq q$, let $W^i := N_G(u^i, W)$. Note that $q \leq kr|W_1|$. Apply Lemma 4.9.1 (with $G[W]$ and $|W_1|$ playing the roles of G and n) to obtain edge-disjoint perfect K_{r-1}-matchings T_i in each $G[W^i]$. Let $G_W := \bigcup_{i=1}^q T_i$. Then $G[U, W] \cup G_W$ has a K_r-decomposition. For each $y \in W$, we use (iv) to see that $d_{G_W}(y) \leq (r-1)d_G(y, U) < 2kr\rho|W_1|$. □

If we are given a k-partition $\mathcal{P}$ of the r-partite graph G, we can apply Corollary 4.9.3 repeatedly with each $U \in \mathcal{P}$ playing the role of W to obtain the following result.

Corollary 4.9.4. *Let* $r \geq 2$ *and* $1/n \ll \rho \ll 1/k, 1/r \leq 1$*. Let* G *be an* r*-partite graph on* $(V_1, \dots, V_r)$ *with* $|V_1| = \cdots = |V_r| = n$*. Let* $\mathcal{P} = \{U^1, \dots, U^k\}$ *be a* k*-partition for* G*. Suppose that the following hold for all* $2 \leq i \leq k$*:*

(i) *for all* $1 \leq j_1, j_2 \leq r$ *and all* $x \in U^{<i} \setminus (V_{j_1} \cup V_{j_2})$, $d_G(x, U^i_{j_1}) = d_G(x, U^i_{j_2})$*;*

(ii) *for all* $1 \leq j \leq r$ *and all* $x \in U^{<i} \setminus V_j$, $\hat{\delta}(G[N_G(x, U^i)] \geq (1 - 1/(r-1))d_G(x, U^i_j) + 9kr\rho^{3/2}|U^i|$*;*

(iii) *for all distinct* $x, x' \in U^{<i}$, $|N_G(x, U^i) \cap N_G(x', U^i)| \leq 2\rho^2|U^i|$*;*

(iv) *for all* $y \in U^i$, $d_G(y, U^{<i}) \leq 2k\rho|U^i_1|$*.*

Then there exists $G_0 \subseteq G - G[\mathcal{P}]$ *such that* $G[\mathcal{P}] \cup G_0$ *has a* K_r*-decomposition and* $\Delta(G_0) \leq 3r\rho n$*.*

Proof. For each $2 \le i \le k$, let $G_i := G[U^{<i}, U^i] \cup G[U^i]$. Apply Corollary 4.9.3 to each G_i with $U^{<i}$, U^i playing the roles of U, W to obtain $G'_i \subseteq G[U^i]$ such that $G[U^{<i}, U^i] \cup G'_i$ has a K_r-decomposition and $\Delta(G'_i) \le 2kr\rho\lceil n/k \rceil \le 3r\rho n$. Let $G_0 := \bigcup_{i=2}^k G'_i$. Then $G[\mathcal{P}] \cup G_0$ has a K_r-decomposition and $\Delta(G_0) \le 3r\rho n$. □

4.10 Balancing graph

In our proof we will consider a sequence of successively finer partitions $\mathcal{P}_1, \dots, \mathcal{P}_\ell$ in turn. When considering $\mathcal{P}_i$, we will assume the leftover is a subgraph of $G - G[\mathcal{P}_{i-1}]$ and aim to use Lemma 4.8.1 and then Corollary 4.9.4 to find copies of K_r such that the leftover is now contained in $G - G[\mathcal{P}_i]$ (i.e. inside the smaller partition classes). However, to apply Corollary 4.9.4 we need the leftover to be balanced with respect to the partition classes. In this section we show how this can be achieved.

Let $\mathcal{P} = \{U^1, \dots, U^k\}$ be a k-partition of the vertex set $V = (V_1, \dots, V_r)$ with $|V_1| = \dots = |V_r| = n$. We say that a graph H on $(V_1, \dots, V_r)$ is *locally $\mathcal{P}$-balanced* if

$$d_H(v, U^i_{j_1}) = d_H(v, U^i_{j_2})$$

for all $1 \le i \le k$, all $1 \le j_1, j_2 \le r$ and all $v \in U^i \setminus (V_{j_1} \cup V_{j_2})$. Note that a graph which is locally $\mathcal{P}$-balanced is not necessarily K_r-divisible but that $H[U^i]$ is K_r-divisible for all $1 \le i \le k$.

Let $\gamma > 0$. A *$(\gamma, \mathcal{P})$-balancing graph* is a K_r-decomposable graph B on V such that the following holds. Let H be any K_r-divisible graph on V with:

(P1) $e(H \cap B) = 0$;

(P2) $|d_H(v, U^i_{j_1}) - d_H(v, U^i_{j_2})| < \gamma n$ for all $1 \le i \le k$, all $1 \le j_1, j_2 \le r$ and all $v \notin V_{j_1} \cup V_{j_2}$.

Then there exists $B' \subseteq B$ such that $B - B'$ has a K_r-decomposition and

$$d_{H \cup B'}(v, U^i_{j_1}) = d_{H \cup B'}(v, U^i_{j_2})$$

for all $2 \le i \le k$, all $1 \le j_1, j_2 \le r$ and all $v \in U^{<i} \setminus (V_{j_1} \cup V_{j_2})$.

Our aim in this section will be to prove Lemma 4.10.1 which finds a $(\gamma, \mathcal{P})$-balancing graph in a suitable graph G.

Lemma 4.10.1. *Let $1/n \ll \gamma \ll \gamma' \ll 1/k \ll \varepsilon \ll 1/r \le 1/3$. Let G be an r-partite graph on $(V_1, \dots, V_r)$ with $|V_1| = \dots = |V_r| = n$. Let $\mathcal{P} = \{U^1, \dots, U^k\}$ be a k-partition for G. Suppose $d_G(v, U^i_j) \ge (1 - 1/(r+1) + \varepsilon)|U^i_j|$ for all $1 \le i \le k$, all $1 \le j \le r$ and all $v \notin V_j$. Then there exists $B \subseteq G$ which is a $(\gamma, \mathcal{P})$-balancing graph such that B is locally $\mathcal{P}$-balanced and $\Delta(B) < \gamma' n$.*

The balancing graph B will be made up of two graphs: B_{edge}, an edge balancing graph (which balances the total number of edges between appropriate classes), and B_{deg}, a degree balancing graph (which balances individual vertex degrees). These are described in Sections 4.10.1 and 4.10.2 respectively.

4.10.1 Edge balancing

Let $\mathcal{P} = \{U^1, \dots, U^k\}$ be a k-partition of the vertex set $V = (V_1, \dots, V_r)$ with $|V_1| = \dots = |V_r| = n$. Let $\gamma > 0$. A $(\gamma, \mathcal{P})$-*edge balancing graph* is a K_r-decomposable graph B_{edge} on V such that the following holds. Let H be any K_r-divisible graph on V which is edge-disjoint from B_{edge} and satisfies (P2). Then there exists $B'_{\text{edge}} \subseteq B_{\text{edge}}$ such that $B_{\text{edge}} - B'_{\text{edge}}$ has a K_r-decomposition and

$$e_{H \cup B'_{\text{edge}}}(U^{i_1}_{j_1}, U^{i_2}_{j_2}) = e_{H \cup B'_{\text{edge}}}(U^{i_1}_{j_1}, U^{i_2}_{j_3})$$

for all $1 \le i_1 < i_2 \le k$ and all $1 \le j_1, j_2, j_3 \le r$ with $j_1 \ne j_2, j_3$.

In this section, we first construct and then find a $(\gamma, \mathcal{P})$-edge balancing graph in G.

For any multigraph G on W and any $e \in W^{(2)}$, let $m_G(e)$ be the multiplicity of the edge e in G. We say that a K_r-divisible multigraph G on $W = (W_1, \dots, W_r)$ is *irreducible* if G has no non-trivial K_r-divisible proper subgraphs; that is, for every $H \subsetneq G$ with $e(H) > 0$,

H is not K_r-divisible. It is easy to see that there are only finitely many irreducible K_r-divisible multigraphs on W. In particular, this implies the following proposition.

Proposition 4.10.2. *Let $r \in \mathbb{N}$ and let $W = (W_1, \dots, W_r)$. Then there exists $N = N(W)$ such that every irreducible K_r-divisible multigraph on W has edge multiplicity at most N.*

Let $\mathcal{P} = \{U^1, \dots, U^k\}$ be a partition of $V = (V_1, \dots, V_r)$. Take a copy K of $K_r(k)$ with vertex set $(W_1, \dots, W_r)$ where $W_j = \{w_j^1, \dots, w_j^k\}$ for each $1 \le j \le r$. For each $1 \le i \le k$, let $W^i := \{w_j^i : 1 \le j \le r\}$. Given a graph H on V, we define an *excess multigraph* $\mathrm{EM}(H)$ on the vertex set $V(K)$ as follows. Between each pair of vertices $w_{j_1}^{i_1}$, $w_{j_2}^{i_2}$ such that $w_{j_1}^{i_1} w_{j_2}^{i_2} \in E(K)$ there are exactly

$$e_H(U_{j_1}^{i_1}, U_{j_2}^{i_2}) - \min\{e_H(U_j^{i_1}, U_{j'}^{i_2}) : 1 \le j, j' \le r, j \neq j'\}$$

multiedges in $\mathrm{EM}(H)$.

Proposition 4.10.3. *Let $r \in \mathbb{N}$ with $r \ge 3$. Let $\mathcal{P} = \{U^1, \dots, U^k\}$ be a k-partition of the vertex set $V = (V_1, \dots, V_r)$ with $|V_1| = \dots = |V_r| = n$. Let H be any K_r-divisible graph on V satisfying* (P2). *Then the excess multigraph* $\mathrm{EM}(H)$ *has a decomposition into at most $3\gamma k^2 r^2 n^2$ irreducible K_r-divisible multigraphs.*

Proof. First, note that for any $1 \le i_1, i_2 \le k$, any $1 \le j_1, j_2, j_3 \le r$ with $j_1 \neq j_2, j_3$ and any $v \in U_{j_1}^{i_1}$, we have $|d_H(v, U_{j_2}^{i_2}) - d_H(v, U_{j_3}^{i_2})| < \gamma n$ by (P2). Therefore,

$$|e_H(U_{j_1}^{i_1}, U_{j_2}^{i_2}) - e_H(U_{j_1}^{i_1}, U_{j_3}^{i_2})| < \gamma n |U_{j_1}^{i_1}| < \gamma n^2. \quad (4.20)$$

We claim that, for all $w_{j_1}^{i_1} w_{j_2}^{i_2} \in E(K)$,

$$m_{\mathrm{EM}(H)}(w_{j_1}^{i_1} w_{j_2}^{i_2}) < 3\gamma n^2. \quad (4.21)$$

Let $1 \leq j'_1, j'_2 \leq r$ with $j'_1 \neq j'_2$. Let $1 \leq j \leq r$ with $j \neq j_1, j'_1$. Then

$$\begin{aligned}
|e_H(U^{i_1}_{j_1}, U^{i_2}_{j_2}) - e_H(U^{i_1}_{j'_1}, U^{i_2}_{j'_2})| &\leq |e_H(U^{i_1}_{j_1}, U^{i_2}_{j_2}) - e_H(U^{i_1}_{j_1}, U^{i_2}_{j})| \\
&\quad + |e_H(U^{i_1}_{j_1}, U^{i_2}_{j}) - e_H(U^{i_1}_{j'_1}, U^{i_2}_{j})| \\
&\quad + |e_H(U^{i_1}_{j'_1}, U^{i_2}_{j}) - e_H(U^{i_1}_{j'_1}, U^{i_2}_{j'_2})| \\
&\overset{(4.20)}{<} 3\gamma n^2.
\end{aligned}$$

So (4.21) holds.

We will now show that $\mathrm{EM}(H)$ is K_r-divisible. Consider any vertex $w^{i_1}_{j_1} \in V(\mathrm{EM}(H))$ and any $1 \leq j_2, j_3 \leq r$ such that $j_1 \neq j_2, j_3$. Note that, since H is K_r-divisible,

$$\begin{aligned}
d_{\mathrm{EM}(H)}(w^{i_1}_{j_1}, W_{j_2}) &= \sum_{i=1}^{k} m_{\mathrm{EM}(H)}(w^{i_1}_{j_1}, w^{i}_{j_2}) \\
&= e_H(U^{i_1}_{j_1}, V_{j_2}) - \sum_{i=1}^{k} \min\{e_H(U^{i_1}_{j}, U^{i}_{j'}) : 1 \leq j, j' \leq r, j \neq j'\} \\
&= e_H(U^{i_1}_{j_1}, V_{j_3}) - \sum_{i=1}^{k} \min\{e_H(U^{i_1}_{j}, U^{i}_{j'}) : 1 \leq j, j' \leq r, j \neq j'\} \\
&= \sum_{i=1}^{k} m_{\mathrm{EM}(H)}(w^{i_1}_{j_1}, w^{i}_{j_3}) = d_{\mathrm{EM}(H)}(w^{i_1}_{j_1}, W_{j_3}).
\end{aligned}$$

So $\mathrm{EM}(H)$ is K_r-divisible and therefore has a decomposition $\mathcal{F}$ into irreducible K_r-divisible multigraphs. By (4.21), there are at most $3\gamma n^2$ edges between any pair of vertices in $\mathrm{EM}(H)$, so $|\mathcal{F}| \leq (3\gamma n^2)e(K) < 3\gamma k^2 r^2 n^2$. □

Let $N = N(V(K))$ be the maximum multiplicity of an edge in any irreducible K_r-divisible multigraph on $V(K) = (W_1, \ldots, W_r)$ (N exists by Proposition 4.10.2). Label each vertex w^i_j of K by U^i_j. Let $K(N)$ be the labelled multigraph obtained from K by replacing each edge of K by N multiedges.

We now construct a $\mathcal{P}$-labelled graph which resembles the multigraph $K(N)$ (when we compare relative differences in the numbers of edges between vertices) and has lower degeneracy. Consider any edge $e = w^{i_1}_{j_1} w^{i_2}_{j_2} \in E(K(N))$. Let $\theta(e)$ be the graph obtained by the following procedure. Take a copy K_e of $K[W^{i_1}, W^{i_2}] - w^{i_1}_{j_1} w^{i_2}_{j_2}$ (K_e inherits the labelling of $K[W^{i_1}, W^{i_2}]$). Note that $K[W^{i_1}, W^{i_2}]$ is a copy of K_r if $i_1 = i_2$ and a copy of the graph obtained from $K_{r,r}$ by deleting a perfect matching otherwise. Join $w^{i_1}_{j_1}$ to the copy of $w^{i_2}_{j_2}$ in K_e and join $w^{i_2}_{j_2}$ to the copy of $w^{i_1}_{j_1}$ in K_e. Write $\theta(e)$ for the resulting $\mathcal{P}$-labelled graph (so the vertex set of $\theta(e)$ consists of $w^{i_1}_{j_1}$, $w^{i_2}_{j_2}$ as well as all the vertices in K_e). Choose the graphs K_e to be vertex-disjoint for all $e \in E(K(N))$. For any $K' \subseteq K(N)$, let $\theta(K') := \bigcup\{\theta(e) : e \in E(K')\}$.

To see that the labelling of $\theta(K(N))$ is actually a $\mathcal{P}$-labelling, note that for any U^i_j, the set of vertices labelled U^i_j forms an independent set in $\theta(K(N))$. Moreover, note that $\theta(K(N))$ has degeneracy $r-1$. To see this, list its vertices in the following order. First list all the original vertices of $V(K)$. These form an independent set in $\theta(K(N))$. Then list the remaining vertices of $\theta(K(N))$ in any order. Each of these vertices has degree $r-1$ in $\theta(K(N))$, so the degeneracy of $\theta(K(N))$ can be at most $r-1$.

Proposition 4.10.4. *Let $\mathcal{P} = \{U^1, \dots, U^k\}$ be a k-partition of the vertex set $V = (V_1, \dots, V_r)$ with $|V_1| = \dots = |V_r| = n$. Let $J = \phi(\theta(K(N)))$ be a copy of $\theta(K(N))$ on V which is compatible with its $\mathcal{P}$-labelling. Then the following hold:*

(i) *J is K_r-divisible and locally $\mathcal{P}$-balanced;*

(ii) *for any multigraph $H \subseteq K(N)$, any $1 \le i_1, i_2 \le k$ and any $1 \le j_1 < j_2 \le r$,*

$$e_{\phi(\theta(H))}(U^{i_1}_{j_1}, U^{i_2}_{j_2}) = e_H(W^{i_1}, W^{i_2}) + m_H(w^{i_1}_{j_1} w^{i_2}_{j_2}).$$

Proof. We first prove that J is K_r-divisible. Consider any $x \in V(\theta(K(N)))$. If $x = w^i_j \in V(K)$, then $d_J(\phi(x), V_{j_1}) = Nk$ for all $1 \le j_1 \le r$ with $j_1 \ne j$ (since for each edge $w^i_j w^{i_1}_{j_1} \in E(K)$, x has exactly N neighbours labelled $U^{i_1}_{j_1}$ in $\theta(K(N))$). If $x \notin V(K)$, x must appear in a copy of K_e in $\theta(e)$ for some edge $e \in E(K(N))$. In this case, $d_J(\phi(x), V_j) = 1$ for all $1 \le j \le r$ such that $\phi(x) \notin V_j$. So J is K_r-divisible.

To see that J is locally $\mathcal{P}$-balanced, consider any $x \in V(\theta(K(N)))$. If $x = w^i_j \in V(K)$, then $\phi(x) \in U^i_j$ and $d_J(\phi(x), U^i_{j_1}) = N$ for all $1 \le j_1 \le r$ with $j_1 \ne j$. Otherwise, x must appear in a copy of K_e in $\theta(e)$ for some edge $e = w^{i_1}_{j_1} w^{i_2}_{j_2} \in E(K(N))$. Let i, j be such that $\phi(x) \in U^i_j$ (so $i \in \{i_1, i_2\}$). If $i_1 \neq i_2$, then $d_J(\phi(x), U^i_{j'}) = 0$ for all $1 \leq j' \leq r$. If $i_1 = i_2$, then $d_J(\phi(x), U^i_{j'}) = 1$ for all $1 \leq j' \leq r$ with $j' \neq j$. So J is locally $\mathcal{P}$-balanced. Thus (i) holds.

We now prove (ii). Let $1 \leq i_1, i_2 \leq k$ and $1 \leq j_1 < j_2 \leq r$. Consider any edge $w^i_j w^{i'}_{j'} \in E(K(N))$. The $\mathcal{P}$-labelling of $\theta(K(N))$ gives

$$e_{\phi(\theta(w^i_j w^{i'}_{j'}))}(U^{i_1}_{j_1}, U^{i_2}_{j_2}) = \begin{cases} 0 & \text{if } \{i, i'\} \neq \{i_1, i_2\}, \\ 2 & \text{if } \{(i,j), (i',j')\} = \{(i_1, j_1), (i_2, j_2)\}, \\ 1 & \text{otherwise.} \end{cases} \tag{4.22}$$

Let $H \subseteq K(N)$. Then (ii) follows from applying (4.22) to each edge in H. $\square$

The following proposition allows us to use a copy of $\theta(K(N))$ to correct imbalances in the number of edges between parts $U^{i_1}_{j_1}$ and $U^{i_2}_{j_2}$ when $\mathrm{EM}(H)$ is an irreducible K_r-divisible multigraph.

Proposition 4.10.5. *Let $\mathcal{P} = \{U^1, \dots, U^k\}$ be a k-partition of the vertex set $V = (V_1, \dots, V_r)$ with $|V_1| = \cdots = |V_r| = n$. Let H be a graph on V such that $\mathrm{EM}(H) = I$ is an irreducible K_r-divisible multigraph. Let $J = \phi(\theta(K(N)))$ be a copy of $\theta(K(N))$ on V which is compatible with its $\mathcal{P}$-labelling and edge-disjoint from H. Then there exists $J' \subseteq J$ such that $J - J'$ is K_r-divisible and $H' := H \cup J'$ satisfies*

$$e_{H'}(U^{i_1}_{j_1}, U^{i_2}_{j_2}) = e_{H'}(U^{i_1}_{j_1}, U^{i_2}_{j_3}) \tag{4.23}$$

for all $1 \leq i_1 < i_2 \leq k$ and all $1 \leq j_1, j_2, j_3 \leq r$ with $j_1 \neq j_2, j_3$.

Proof. Recall that N denotes the maximum multiplicity of an edge in an irreducible K_r-divisible multigraph on $V(K)$. So we may view I as a subgraph of $K(N)$. Let

$J' := J - \phi(\theta(I))$. For all $1 \le i_1 < i_2 \le k$, let

$$p_{i_1,i_2} := \min\{e_H(U^{i_1}_{j_1}, U^{i_2}_{j_2}) : 1 \le j_1, j_2, \le r, j_1 \ne j_2\}.$$

Proposition 4.10.4 gives, for all $1 \le i_1 < i_2 \le k$ and all $1 \le j_1, j_2 \le r$ with $j_1 \ne j_2$,

$$\begin{aligned} e_{J'}(U^{i_1}_{j_1}, U^{i_2}_{j_2}) &= e_{\phi(\theta(K(N)))}(U^{i_1}_{j_1}, U^{i_2}_{j_2}) - e_{\phi(\theta(I))}(U^{i_1}_{j_1}, U^{i_2}_{j_2}) \\ &= e_{K(N)}(W^{i_1}, W^{i_2}) + N - (e_I(W^{i_1}, W^{i_2}) + m_I(w^{i_1}_{j_1} w^{i_2}_{j_2})) \\ &= e_{K(N)-I}(W^{i_1}, W^{i_2}) + N - m_I(w^{i_1}_{j_1} w^{i_2}_{j_2}). \end{aligned}$$

Recall that $I = \mathrm{EM}(H)$, so $e_H(U^{i_1}_{j_1}, U^{i_2}_{j_2}) = m_I(w^{i_1}_{j_1} w^{i_2}_{j_2}) + p_{i_1,i_2}$ and

$$e_{H'}(U^{i_1}_{j_1}, U^{i_2}_{j_2}) = e_H(U^{i_1}_{j_1}, U^{i_2}_{j_2}) + e_{J'}(U^{i_1}_{j_1}, U^{i_2}_{j_2}) = e_{K(N)-I}(W^{i_1}, W^{i_2}) + N + p_{i_1,i_2}.$$

Note that the right hand side is independent of j_1, j_2. Thus (4.23) holds. □

The following proposition describes a $(\gamma, \mathcal{P})$-edge balancing graph based on the construction in Propositions 4.10.4 and 4.10.5

Proposition 4.10.6. *Let $k, r \in \mathbb{N}$ with $r \ge 3$. Let $\mathcal{P} = \{U^1, \dots, U^k\}$ be a k-partition of the vertex set $V = (V_1, \dots, V_r)$ with $|V_1| = \dots = |V_r| = n$. Let $J_1, \dots, J_\ell$ be a collection of $\ell \ge 3\gamma k^2 r^2 n^2$ copies of $\theta(K(N))$ on V which are compatible with their labellings. Let $\{A_1, \dots, A_m\}$ be an absorbing set for $J_1, \dots, J_\ell$ on V. Suppose that $J_1, \dots, J_\ell, A_1, \dots, A_m$ are edge-disjoint. Then $B_{\mathrm{edge}} := \bigcup_{i=1}^{\ell} J_i \cup \bigcup_{i=1}^{m} A_i$ is a $(\gamma, \mathcal{P})$-edge balancing graph.*

Proof. Let H be any K_r-divisible graph on V which is edge-disjoint from B_{edge} and satisfies (P2). Apply Proposition 4.10.3 to find a decomposition of $\mathrm{EM}(H)$ into a collection $\mathcal{I} = \{I_1, \dots, I_{\ell'}\}$ of irreducible K_r-divisible multigraphs, where $\ell' \le 3\gamma k^2 r^2 n^2 \le \ell$. If $\ell' = 0$, let $B'_{\mathrm{edge}} \subseteq B_{\mathrm{edge}}$ be the empty graph. If $\ell' > 0$, we proceed as follows to find B'_{edge}. Let $H_1, \dots, H_{\ell'}$ be graphs on V which partition the edge set of H and satisfy $\mathrm{EM}(H_s) = I_s$ for each $1 \le s \le \ell'$. (To find such a partition, for each $1 \le s < \ell'$ form H_s

by taking one $U_{j_1}^{i_1}U_{j_2}^{i_2}$-edge from H for each edge $w_{j_1}^{i_1}w_{j_2}^{i_2}$ in I_s. Let $H_{\ell'}$ consist of all the remaining edges.)

Apply Proposition 4.10.5 for each $1 \leq s \leq \ell'$ with H_s and J_s playing the roles of H and J to find $J'_s \subseteq J_s$ such that $J_s - J'_s$ is K_r-divisible and $H'_s := H_s \cup J'_s$ satisfies

$$e_{H'_s}(U_{j_1}^{i_1}, U_{j_2}^{i_2}) = e_{H'_s}(U_{j_1}^{i_1}, U_{j_3}^{i_2}) \tag{4.24}$$

for all $1 \leq i_1 < i_2 \leq k$ and all $1 \leq j_1, j_2, j_3 \leq r$ with $j_1 \neq j_2, j_3$. Let $B'_{\text{edge}} := \bigcup_{s=1}^{\ell'} J'_s$. Then (4.24) implies that the graph $H' := H \cup B'_{\text{edge}} = \bigcup_{s=1}^{\ell'} H'_s$ satisfies

$$e_{H'}(U_{j_1}^{i_1}, U_{j_2}^{i_2}) = e_{H'}(U_{j_1}^{i_1}, U_{j_3}^{i_2})$$

for all $1 \leq i_1 < i_2 \leq k$ and all $1 \leq j_1, j_2, j_3 \leq r$ with $j_1 \neq j_2, j_3$.

We now check that B_{edge} and $B_{\text{edge}} - B'_{\text{edge}}$ are K_r-decomposable. Recall that every absorber A_i is K_r-decomposable. Also recall that, for every $1 \leq s \leq \ell$, J_s is K_r-divisible, by Proposition 4.10.4(i). Since $\{A_1, \dots, A_m\}$ is an absorbing set, it contains a distinct absorber for each J_s. So for each $1 \leq s \leq \ell$, there exists a distinct $1 \leq i_s \leq m$ such that $A_{i_s} \cup J_s$ has a K_r-decomposition. Therefore B_{edge} is K_r-decomposable. To see that $B_{\text{edge}} - B'_{\text{edge}}$ is K_r-decomposable, recall that for each $1 \leq s \leq \ell'$, $J_s - J'_s$ is a K_r-divisible subgraph of J_s. So for each $1 \leq s \leq \ell$, there exists a distinct $1 \leq j_s \leq m$ such that, if $s \leq \ell'$, $A_{j_s} \cup (J_s - J'_s)$ has a K_r-decomposition and, if $s > \ell'$, $A_{j_s} \cup J_s$ has a K_r-decomposition. So we can find a K_r-decomposition of

$$B_{\text{edge}} - B'_{\text{edge}} = \bigcup_{s=1}^{\ell'} (J_s - J'_s) \cup \bigcup_{s=\ell'+1}^{\ell} J_s \cup \bigcup_{s=1}^{m} A_m.$$

Therefore, B_{edge} is a $(\gamma, \mathcal{P})$-edge balancing graph. □

The next proposition finds a copy of this $(\gamma, \mathcal{P})$-edge balancing graph in G.

Proposition 4.10.7. *Let $1/n \ll \gamma \ll \gamma' \ll 1/k \ll \varepsilon \ll 1/r \leq 1/3$. Let G be an r-partite* *graph on* $(V_1, \dots, V_r)$ *with* $|V_1| = \cdots = |V_r| = n$*. Let* $\mathcal{P} = \{U^1, \dots, U^k\}$ *be a* k*-partition for* G*. Suppose that* $d_G(v, U^i_j) \geq (1 - 1/(r+1) + \varepsilon)|U^i_j|$ *for all* $1 \leq i \leq k$*, all* $1 \leq j \leq r$ *and all* $v \notin V_j$*. Then there exists a* $(\gamma, \mathcal{P})$*-edge balancing graph* $B_{\mathrm{edge}} \subseteq G$ *such that* B_{edge} *is locally* $\mathcal{P}$*-balanced and* $\Delta(B_{\mathrm{edge}}) < \gamma' n$*.*

Proof. Let γ_1 be such that $\gamma \ll \gamma_1 \ll \gamma'$. Recall that $\theta(K(N))$ is a $\mathcal{P}$-labelled graph with degeneracy $r-1$ and all vertices of $\theta(K(N))$ are free vertices. Also,

$$|\theta(K(N))| \leq |K| + 2re(K)N = kr + 2rk^2 \binom{r}{2} N \leq k^2 r^3 N.$$

Let $\ell := \lceil 3\gamma k^2 r^2 n^2 \rceil \leq \gamma^{1/2} n^2$. We can apply Lemma 4.5.2 (with $\gamma^{1/2}$, γ_1, $r-1$, $k^2 r^3 N$ and $\theta(K(N))$ playing the roles of η, ε, d, b and H_i) to find edge-disjoint copies $J_1, \dots, J_\ell$ of $\theta(K(N))$ in G which are compatible with their labellings and satisfy $\Delta(\bigcup_{i=1}^{\ell} J_i) \leq \gamma_1 n$.

Let $G' := G[\mathcal{P}] - \bigcup_{i=1}^{\ell} J_i$ and note that

$$\hat{\delta}(G') \geq (1 - 1/(r+1) + \varepsilon)n - \lceil n/k \rceil - \gamma_1 n \geq (1 - 1/(r+1) + \gamma')n.$$

Apply Lemma 4.6.6 (with γ_1, $\gamma'/2$, $k^2 r^3 N$ and G' playing the roles of η, ε, b and G) to find an absorbing set $\mathcal{A}$ for $J_1, \dots, J_\ell$ in G' such that $\Delta(\bigcup \mathcal{A}) \leq \gamma' n/2$.

Let $B_{\mathrm{edge}} := \bigcup_{i=1}^{\ell} J_i \cup \bigcup \mathcal{A}$. Then B_{edge} is a $(\gamma, \mathcal{P})$-edge balancing graph by Proposition 4.10.6. Also, $\Delta(B_{\mathrm{edge}}) < \gamma' n$. Note that for each $1 \leq i \leq k$, $B_{\mathrm{edge}}[U^i] = \bigcup_{s=1}^{\ell} J_s[U^i]$ (this is the reason for finding $\mathcal{A}$ in $G[\mathcal{P}]$). Moreover, each J_s is locally $\mathcal{P}$-balanced by Proposition 4.10.4(i). Therefore B_{edge} is also locally $\mathcal{P}$-balanced. □

4.10.2 Degree balancing

Let $\mathcal{P} = \{U^1, \dots, U^k\}$ be a k-partition of the vertex set $V = (V_1, \dots, V_r)$ with $|V_1| = \cdots = |V_r| = n$. Let $\gamma > 0$. A $(\gamma, \mathcal{P})$-*degree balancing graph* is a K_r-decomposable graph B_{deg} on V such that the following holds. Let H be any K_r-divisible graph on V satisfying:

(Q1) $e(H \cap B_{\mathrm{deg}}) = 0$;

(Q2) $e_H(U^{i_1}_{j_1}, U^{i_2}_{j_2}) = e_H(U^{i_1}_{j_1}, U^{i_2}_{j_3})$ for all $1 \leq i_1 < i_2 \leq k$ and all $1 \leq j_1, j_2, j_3 \leq r$ with $j_1 \neq j_2, j_3$;

(Q3) $|d_H(v, U^i_{j_2}) - d_H(v, U^i_{j_3})| < \gamma |U^i_{j_1}|$ for all $2 \leq i \leq k$, all $1 \leq j_1, j_2, j_3 \leq r$ with $j_1 \neq j_2, j_3$ and all $v \in U^{<i}_{j_1}$.

Then there exists $B'_{\text{deg}} \subseteq B_{\text{deg}}$ such that $B_{\text{deg}} - B'_{\text{deg}}$ has a K_r-decomposition and

$$d_{H \cup B'_{\text{deg}}}(v, U^i_{j_1}) = d_{H \cup B'_{\text{deg}}}(v, U^i_{j_2})$$

for all $2 \leq i \leq k$, all $1 \leq j_1, j_2 \leq r$ and all $v \in U^{<i} \setminus (V_{j_1} \cup V_{j_2})$.

We will build a degree balancing graph by combining smaller graphs which correct the degrees between two parts of the partition at a time. So, let us assume that the partition has only two parts, i.e., let $\mathcal{P} = \{U^1, U^2\}$ partition the vertex set $V = (V_1, \ldots, V_r)$. We begin by defining those graphs which will form the basic gadgets of the degree balancing graph. Let D_0 be a copy of $K_r(3)$ with vertex classes $\{w^i_j : 1 \leq i \leq 3\}$ for $1 \leq j \leq r$. For each $1 \leq i \leq 3$, let $W^i := \{w^i_j : 1 \leq j \leq r\}$. We define a labelling $L : V(D_0) \to \{U^1_j, U^2_j : 1 \leq j \leq r\}$ as follows:

$$L(w^i_j) = \begin{cases} U^1_j & \text{if } i = 1, 2, \\ U^2_j & \text{if } i = 3. \end{cases}$$

Suppose that x, y are distinct vertices in $U^1_{j_1}$ where $1 \leq j_1 \leq r$. Obtain the $\mathcal{P}$-labelled graph $D_{x,y}$ by taking the labelled copy of D_0 and changing the label of $w^1_{j_1}$ to $\{x\}$ and $w^2_{j_1}$ to $\{y\}$. Let $1 \leq j_2 \leq r$ be such that $j_2 \neq j_1$. Let $D^{j_2}_{x \to y}$ be the $\mathcal{P}$-labelled subgraph of $D_{x,y}$ which has as its vertex set

$$W^1 \cup \{w^2_{j_1}\} \cup (W^3 \setminus \{w^3_{j_1}\}),$$

contains all possible edges in $W^1 \setminus \{w^1_{j_1}\}$, all possible edges in $W^3 \setminus \{w^3_{j_1}\}$, all edges of the form $w^1_{j_1} w^3_j$ and $w^1_j w^2_{j_1}$ where $1 \leq j \leq r$ and $j \neq j_1, j_2$, as well as the edges $w^1_{j_1} w^1_{j_2}$ and $w^2_{j_1} w^3_{j_2}$. (Note that if we were to identify the vertices $w^1_{j_1}$ and $w^2_{j_1}$ we would obtain two copies of K_r which have only one vertex in common.)

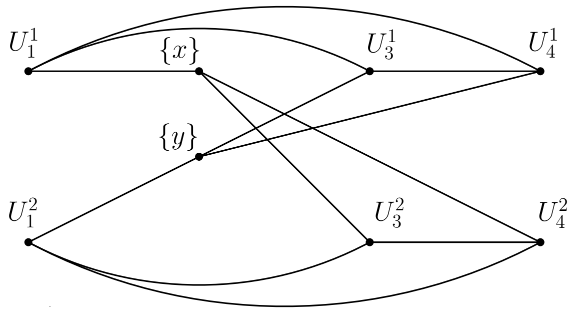

Figure 4.2: A copy of $D^1_{x\to y}$ when $r = 4$ and $x, y \in U^1_2$.

As in Section 4.10.1, we would like to reduce the degeneracy of $D_{x,y}$. The operation θ (which will be familiar from Section 4.10.1) replaces each edge of $D_{x,y}$ by a $\mathcal{P}$-labelled graph as follows. Consider any edge $e = w^{i_1}_{j_3} w^{i_2}_{j_4} \in E(D_{x,y})$. Take a labelled copy D_e of $D_0[W^{i_1}, W^{i_2}] - w^{i_1}_{j_3} w^{i_2}_{j_4}$ (D_e inherits the labelling of $D_0[W^{i_1}, W^{i_2}]$). Note that $D_0[W^{i_1}, W^{i_2}]$ is a copy of K_r if $i_1 = i_2$ and a copy of the graph obtained from $K_{r,r}$ by deleting a perfect matching otherwise. Join $w^{i_1}_{j_3}$ to the copy of $w^{i_2}_{j_4}$ in D_e and join $w^{i_2}_{j_4}$ to the copy of $w^{i_1}_{j_3}$ in D_e (so the vertex set of $\theta(e)$ consists of $w^{i_1}_{j_3}$, $w^{i_2}_{j_4}$ as well as all the vertices in D_e). Write $\theta(e)$ for the resulting $\mathcal{P}$-labelled graph. Choose the graphs D_e to be vertex-disjoint for all $e \in E(D_{x,y})$. For any $D' \subseteq D_{x,y}$, let $\theta(D') := \bigcup\{\theta(e) : e \in E(D')\}$. The graph $\theta(D_{x,y})$ has the following properties:

(θ1) $|\theta(D_{x,y})| \leq 3r + 2r3^2\binom{r}{2} \leq 10r^3$ (since we add at most $2re(K_r(3))$ new vertices to obtain $\theta(D_{x,y})$ from $D_{x,y}$);

(θ2) $\theta(D_{x,y})$ has degeneracy $r - 1$ (to see this, take the original vertices of $D_{x,y}$ first, followed by the remaining vertices in any order).

Suppose that H is a graph on V and $x, y \in U^1_{j_1}$. Suppose that $d_H(x, U^2_{j_2})$ is currently too large and $d_H(y, U^2_{j_2})$ is too small. The next proposition allows us to use copies of $\theta(D^{j_2}_{x\to y})$ to 'transfer' some of this surplus from x to y.

Proposition 4.10.8. *Let $\mathcal{P} = \{U^1, U^2\}$ be a partition of the vertex set $V = (V_1, \ldots, V_r)$. Let $1 \leq j_1, j_2 \leq r$ with $j_1 \neq j_2$ and suppose $x, y \in U^1_{j_1}$. Suppose that $D_1 = \phi(\theta(D_{x,y}))$ is a* *copy of $\theta(D_{x,y})$ on V which is compatible with its labelling. Let $D_2 := \phi(\theta(D^{j_2}_{x\to y})) \subseteq D_1$. Then the following hold:*

(i) *both D_1 and D_2 are K_r-divisible;*

(ii) *D_1 is locally $\mathcal{P}$-balanced;*

(iii) *for any $1 \leq j_3, j_4 \leq r$ with $j_4 \neq j_2$ and any $v \in U^1 \setminus (V_{j_3} \cup V_{j_4})$,*

$$d_{D_2}(v, U^2_{j_3}) - d_{D_2}(v, U^2_{j_4}) = \begin{cases} -1 & \textit{if } v = x \textit{ and } j_3 = j_2, \\ 1 & \textit{if } v = y \textit{ and } j_3 = j_2, \\ 0 & \textit{otherwise.} \end{cases}$$

Proof. First we show that (i) holds. Consider any $v \in V(\theta(D_{x,y}))$. If $v \in V(D_{x,y})$, then $d_{D_1}(\phi(v), V_j) = 3$ for all $1 \leq j \leq r$ such that $\phi(v) \notin V_j$. Otherwise, v appears in a copy of D_e for some edge $e \in E(D_{x,y})$ and $d_{D_1}(\phi(v), V_j) = 1$ for all $1 \leq j \leq r$ such that $\phi(v) \notin V_j$. So D_1 is K_r-divisible. For D_2, consider any $v \in V(\theta(D^{j_2}_{x\to y}))$. If $v \in V(D^{j_2}_{x\to y})$, then $d_{D_2}(\phi(v), V_j) = 1$ for all $1 \leq j \leq r$ with $\phi(v) \notin V_j$. Otherwise, v appears in a copy of D_e for some edge $e \in E(D^{j_2}_{x\to y})$ and $d_{D_2}(\phi(v), V_j) = 1$ for all $1 \leq j \leq r$ such that $\phi(v) \notin V_j$. So D_2 is K_r-divisible.

For (ii), consider any $v \in V(\theta(D_{x,y}))$. First suppose $v = w^i_j \in V(D_{x,y})$. If $i = 1, 2$, then $\phi(v) \in U^1_j$ and $d_{D_1}(\phi(v), U^1_{j'}) = 2$ for all $1 \leq j' \leq r$ with $j' \neq j$. If $i = 3$, then $\phi(v) \in U^2_j$ and $d_{D_1}(\phi(v), U^2_{j'}) = 1$ for all $1 \leq j' \leq r$ with $j' \neq j$. Otherwise, v must appear in a copy of D_e in $\theta(e)$ for some edge $e = w^{i_1}_{j_1} w^{i_2}_{j_2} \in E(D_{x,y})$. Let i, j be such that $\phi(v) \in U^i_j$. If $i_1, i_2 \in \{1, 2\}$ or if $i_1 = i_2 = 3$, then $d_{D_1}(\phi(v), U^i_{j'}) = 1$ for all $1 \leq j' \leq r$ with $j' \neq j$. Otherwise, $d_{D_1}(\phi(v), U^i_{j'}) = 0$ for all $1 \leq j' \leq r$. So D_1 is locally $\mathcal{P}$-balanced.

Property (iii) will follow from the $\mathcal{P}$-labelling of $\theta(D^{j_2}_{x\to y})$. Note that

$$d_{D_2}(x, U^2_{j'}) = \begin{cases} 0 & \text{if } j' \in \{j_1, j_2\}, \\ 1 & \text{otherwise} \end{cases} \qquad \text{and} \qquad d_{D_2}(y, U^2_{j'}) = \begin{cases} 1 & \text{if } j' = j_2, \\ 0 & \text{otherwise.} \end{cases}$$

The only other edges ab in D_2 of the form U^1U^2 are those which appear in the image of D_e for some $e = w^i_j w^3_{j'} \in E(D^{j_2}_{x\to y})$ with $i = 1, 2$. Note that such e must be incident to x or y and that a and b are new vertices, i.e., $a, b \notin V(D^{j_2}_{x\to y})$. But for any $v \in \phi(D_e) \cap U^1$, we have $d_{D_2}(v, U^2_{j'}) = 1$ for every $1 \le j' \le r$ such that $\phi(v) \notin V_{j'}$. It follows that (iii) holds. □

In what follows, given a collection $\mathcal{D}$ of graphs and an embedding $\phi(D)$ for each $D \in \mathcal{D}$, we write $\phi(\mathcal{D}) := \{\phi(D) : D \in \mathcal{D}\}$.

Lemma 4.10.9. *Let $1/n \ll \gamma \ll \gamma' \le 1/r \le 1/3$. Let $V = (V_1, \dots, V_r)$ with $|V_1| = \cdots = |V_r| = n$. Let $\mathcal{P} = \{U^1, U^2\}$ be a 2-partition of V. Let $1 \le j_1 \le r$. Then there exists $\mathcal{D} \subseteq \{\theta(D^j_{x\to y}) : x, y \in U^1_{j_1}, x \ne y, 1 \le j \le r, j \ne j_1\}$ such that the following hold.*

(i) $|\mathcal{D}| \le \gamma' n^2$.

(ii) *Each vertex $v \in V$ is a root vertex in at most $\gamma' n$ elements of $\mathcal{D}$.*

(iii) *Suppose that, for each $D \in \mathcal{D}$, $\phi(D)$ is a copy of D on V which is compatible with its labelling. Suppose further that $\phi(D)$ and $\phi(D')$ are edge-disjoint for all distinct $D, D' \in \mathcal{D}$. Let H be any r-partite graph on V which is edge-disjoint from $\bigcup \phi(\mathcal{D})$ and satisfies* (Q2) *and* (Q3). *Then there exists $\mathcal{D}' \subseteq \mathcal{D}$ such that $H' := H \cup \bigcup \phi(\mathcal{D}')$ satisfies the following. For all $v \in U^1_{j_1}$, and all $1 \le j_2, j_3 \le r$ such that $j_1 \ne j_2, j_3$,*

$$d_{H'}(v, U^2_{j_2}) = d_{H'}(v, U^2_{j_3})$$

and for all $1 \le j_2, j_3 \le r$ and all $v \in U^1 \setminus (V_{j_1} \cup V_{j_2} \cup V_{j_3})$,

$$d_{H'}(v, U^2_{j_2}) - d_{H'}(v, U^2_{j_3}) = d_H(v, U^2_{j_2}) - d_H(v, U^2_{j_3}).$$

In particular, H' satisfies (Q2) *and* (Q3).

Proof. Let $p := \gamma'/4(r-1)$ and $m := |U^1_{j_1}|$. Define an auxiliary graph R on $U^1_{j_1}$ such that $\Delta(R) < 2pm$ and

$$|N_R(S)| \geq p^2m/2 \tag{4.25}$$

for all $S \subseteq U^1_{j_1}$ with $|S| \leq 2$. It is easy to find such a graph R; indeed, a random graph with edge probability p has these properties with high probability.

Let

$$\mathcal{D} := \{\theta(D^j_{x\to y}), \theta(D^j_{y\to x}) : xy \in E(R), 1 \leq j \leq r, j \neq j_1\}.$$

Each vertex of V appears as x or y in some $\theta(D^j_{x\to y})$ in $\mathcal{D}$ at most $2(r-1)\Delta(R) < 4(r-1)pm = \gamma' m$ times. In particular, this implies $|\mathcal{D}| \leq \gamma' m^2$. So $\mathcal{D}$ satisfies (i) and (ii).

We now show that $\mathcal{D}$ satisfies (iii). Suppose that, for each $D \in \mathcal{D}$, $\phi(D)$ is a copy of D on V which is compatible with its labelling. Suppose further that $\phi(D)$ and $\phi(D')$ are edge-disjoint for all distinct $D, D' \in \mathcal{D}$. Let H be any r-partite graph on V which is edge-disjoint from $\bigcup \phi(\mathcal{D})$ and satisfies (Q2) and (Q3).

Let $j_{\min} := \min\{j : 1 \leq j \leq r, j \neq j_1\}$. For each $v \in U^1_{j_1}$ and each $j_{\min} < j \leq r$ such that $j \neq j_1$, let

$$f(v,j) := d_H(v, U^2_j) - d_H(v, U^2_{j_{\min}}). \tag{4.26}$$

By (Q3),

$$|f(v,j)| < \gamma(m+1) < 2\gamma m. \tag{4.27}$$

Let $U^+(j)$ be a multiset such that each $v \in U^1_{j_1}$ appears precisely $\max\{f(v,j), 0\}$ times. Let $U^-(j)$ be a multiset such that each $v \in U^1_{j_1}$ appears precisely $\max\{-f(v,j), 0\}$ times. Property (Q2) implies that $|U^+(j)| = |U^-(j)|$, so there is a bijection $g_j : U^+(j) \to U^-(j)$.

For each copy u' of u in $U^+(j)$, let $P_{u'}$ be a path of length two whose vertices are labelled, in order,

$$\{u\}, U^1_{j_1}, \{g_j(u')\}.$$

So $P_{u'}$ has degeneracy two. Let $\mathcal{S}_j := \{P_{u'} : u' \in U^+(j)\}$. It follows from (4.27) that each vertex is used as a root vertex at most $2\gamma m$ times in $\mathcal{S}_j$ and $|\mathcal{S}_j| \leq 2\gamma m^2$. Using

(4.25), we can apply Lemma 4.5.1 (with m, 2, 3, 2γ, $p^2/2$ and R playing the roles of n, d, b, η, ε and G) to find a set of edge-disjoint copies $\mathcal{T}_j$ of the paths in $\mathcal{S}_j$ in R which are compatible with their labellings. (Note that we do not require the paths in $\mathcal{T}_j$ to be edge-disjoint from the paths in $\mathcal{T}_{j'}$ for $j \neq j'$.) We will view the paths in $\mathcal{T}_j$ as directed paths whose initial vertex lies in $U^+(j)$ and whose final vertex lies in $U^-(j)$.

For each $j_{\min} < j \leq r$ such that $j \neq j_1$, let $\mathcal{D}_j := \{\theta(D^j_{x\to y}) : \overrightarrow{xy} \in E(\bigcup \mathcal{T}_j)\}$. Let

$$\mathcal{D}' := \bigcup_{\substack{j_{\min}<j\leq r \\ j\neq j_1}} \mathcal{D}_j \subseteq \mathcal{D}.$$

It remains to show that $H' := H \cup \bigcup \phi(\mathcal{D}')$ satisfies (iii). For each $j_{\min} < j \leq r$ such that $j \neq j_1$, let $H_j := \bigcup \phi(\mathcal{D}_j)$. Consider any vertex $v \in U^1_{j_1}$ and let $j_{\min} < j_2 \leq r$ be such that $j_2 \neq j_1$. Now v will be the initial vertex in exactly $a := \max\{f(v, j_2), 0\}$ paths and the final vertex in exactly $b := \max\{-f(v, j_2), 0\} = a - f(v, j_2)$ paths in $\mathcal{T}_{j_2}$. Let c be the number of paths in $\mathcal{T}_{j_2}$ for which v is an internal vertex. By definition, H_{j_2} contains $a + c$ graphs $\phi(D)$ where D is of the form $\theta(D^{j_2}_{v\to y})$ for some $y \in U^1_{j_1}$. Also, H_{j_2} contains $b + c$ graphs $\phi(D)$ where D of the form $\theta(D^{j_2}_{x\to v})$ for some $x \in U^1_{j_1}$. Proposition 4.10.8(iii) then implies that

$$d_{H_{j_2}}(v, U^2_{j_2}) - d_{H_{j_2}}(v, U^2_{j_{\min}}) = (b + c) - (a + c) = -f(v, j_2). \tag{4.28}$$

For any $j_{\min} < j_3 \leq r$ such that $j_3 \neq j_1, j_2$, Proposition 4.10.8(iii) implies that

$$d_{H_{j_3}}(v, U^2_{j_2}) - d_{H_{j_3}}(v, U^2_{j_{\min}}) = 0. \tag{4.29}$$

Equations (4.28) and (4.29) imply that

$$d_{\bigcup \phi(\mathcal{D}')}(v, U^2_{j_2}) - d_{\bigcup \phi(\mathcal{D}')}(v, U^2_{j_{\min}}) = d_{H_{j_2}}(v, U^2_{j_2}) - d_{H_{j_2}}(v, U^2_{j_{\min}}) = -f(v, j_2),$$

which together with (4.26) gives

$$d_{H'}(v,U^2_{j_2}) - d_{H'}(v,U^2_{j_{\min}}) = d_H(v,U^2_{j_2}) - d_H(v,U^2_{j_{\min}}) - f(v,j_2) = 0. \tag{4.30}$$

Thus, for all $v \in U^1_{j_1}$ and all $1 \le j_2, j_3 \le r$ such that $j_1 \neq j_2, j_3$,

$$d_{H'}(v,U^2_{j_2}) = d_{H'}(v,U^2_{j_{\min}}) = d_{H'}(v,U^2_{j_3}).$$

Finally, consider any $1 \le j_2, j_3 \le r$ and any $v \in U^1 \setminus (V_{j_1} \cup V_{j_2} \cup V_{j_3})$. Proposition 4.10.8(iii) implies that

$$d_{\bigcup \phi(\mathcal{D}')}(v,U^2_{j_2}) - d_{\bigcup \phi(\mathcal{D}')}(v,U^2_{j_3}) = 0,$$

so

$$d_{H'}(v,U^2_{j_2}) - d_{H'}(v,U^2_{j_3}) = d_H(v,U^2_{j_2}) - d_H(v,U^2_{j_3}). \tag{4.31}$$

That H' satisfies (Q2) and (Q3) follows immediately from (4.30) and (4.31). □

Let $\mathcal{P} = \{U^1, U^2\}$ partition the vertex set $V = (V_1, \dots, V_r)$ with $|V_1| = \cdots = |V_r| = n$. We say that a collection $\mathcal{D}$ of $\mathcal{P}$-labelled graphs is a (γ,γ')-*degree balancing set* for the pair (U^1, U^2) if the following properties hold. Suppose that, for each $D \in \mathcal{D}$, $\phi(D)$ is a copy of D on V which is compatible with its labelling. Suppose further that $\phi(D)$ and $\phi(D')$ are edge-disjoint for all distinct $D, D' \in \mathcal{D}$.

(a) Each $D \in \mathcal{D}$ has degeneracy at most $r-1$ and $|D| \le 10r^3$.

(b) $|\mathcal{D}| \le \gamma' n^2$.

(c) Each vertex $v \in V$ is a root vertex in at most $\gamma' n$ elements of $\mathcal{D}$.

(d) For each $D \in \mathcal{D}$, $\phi(D)$ is K_r-divisible and locally $\mathcal{P}$-balanced.

(e) Let H be any r-partite graph on V which is edge-disjoint from $\bigcup \phi(\mathcal{D})$ and satisfies (Q2) and (Q3). Then, for each $D \in \mathcal{D}$, there exists $D' \subseteq D$ such that $\phi(D')$ is K_r-divisible and, if $\mathcal{D}' := \{D' : D \in \mathcal{D}\}$ and $H' := H \cup \bigcup \phi(\mathcal{D}')$, then

$$d_{H'}(v, U^2_{j_1}) = d_{H'}(v, U^2_{j_2})$$

for all $1 \leq j_1, j_2 \leq r$ and all $v \in U^1 \setminus (V_{j_1} \cup V_{j_2})$.

The following result describes a (γ, γ')-degree balancing set based on the gadgets constructed so far.

Proposition 4.10.10. *Let $1/n \ll \gamma \ll \gamma' \leq 1/r \leq 1/3$. Let $V = (V_1, \dots, V_r)$ with $|V_1| = \dots = |V_r| = n$. Let $\mathcal{P} = \{U^1, U^2\}$ be a 2-partition for V. Then (U^1, U^2) has a (γ, γ')-degree balancing set.*

Proof. Apply Lemma 4.10.9 for each $1 \leq j_1 \leq r$ with γ'/r playing the role of γ' to find sets $\mathcal{D}_{j_1} \subseteq \{\theta(D^j_{x \to y}) : x, y \in U^1_{j_1}, x \neq y, 1 \leq j \leq r, j \neq j_1\}$ satisfying the properties (i)–(iii). Let $\mathcal{D}$ consist of one copy of $\theta(D_{x,y})$ for each $\theta(D^j_{x \to y})$ in $\bigcup_{j=1}^r \mathcal{D}_j$. We claim that $\mathcal{D}$ is a (γ, γ')-degree balancing set. Note that each $\theta(D_{x,y})$ satisfies $|\theta(D_{x,y})| \leq 10r^3$ and has degeneracy at most $r-1$ by (θ1) and (θ2), so (a) holds. For each $1 \leq j \leq r$, $|\mathcal{D}_j| \leq \gamma' n^2/r$, so (b) holds. Also, each vertex $v \in V$ is used as a root vertex in at most $\gamma' n/r$ elements of each $\mathcal{D}_j$. Since $\theta(D_{x,y})$ and $\theta(D^j_{x \to y})$ have the same set of root vertices, (c) holds. Property (d) follows from Proposition 4.10.8(i) and (ii).

It remains to show that (e) is satisfied. Suppose that, for each $D \in \mathcal{D}$, $\phi(D)$ is a copy of D on V which is compatible with its labelling. Suppose further that $\phi(D)$ and $\phi(D')$ are edge-disjoint for all distinct $D, D' \in \mathcal{D}$. Let H be any r-partite graph on V which is edge-disjoint from $\bigcup \phi(\mathcal{D})$ and satisfies (Q2) and (Q3). Using property (iii) of $\mathcal{D}_1$ in Lemma 4.10.9, we can find $\mathcal{D}'_1 \subseteq \mathcal{D}_1$ such that $H_1 := H \cup \bigcup \phi(\mathcal{D}'_1)$ satisfies (Q2), (Q3) and

$$d_{H_1}(v, U^2_{j_1}) = d_{H_1}(v, U^2_{j_2})$$

for all $v \in U^1_1$ and all $2 \leq j_1, j_2 \leq r$. We can then find $\mathcal{D}'_2 \subseteq \mathcal{D}_2$ such that $H_2 := H_1 \cup \bigcup \phi(\mathcal{D}'_2)$ satisfies (Q2), (Q3) and

$$d_{H_2}(v, U^2_{j_1}) = d_{H_2}(v, U^2_{j_2})$$

for all $v \in U^1_j$ where $j = 1, 2$ and all $1 \leq j_1, j_2 \leq r$ with $j \neq j_1, j_2$. Continuing in this way, we eventually find $\mathcal{D}'_r \subseteq \mathcal{D}_r$ such that $H_r := H_{r-1} \cup \bigcup \phi(\mathcal{D}'_{r-1})$ satisfies

$$d_{H_r}(v, U^2_{j_1}) = d_{H_r}(v, U^2_{j_2}) \tag{4.32}$$

for all $1 \leq j_1, j_2 \leq r$ and all $v \in U^1 \setminus (V_{j_1} \cup V_{j_2})$.

For each $D \in \mathcal{D}_j$, if $D \in \mathcal{D}'_j$, then let $D' := D$; otherwise let D' be the empty graph. Let $\mathcal{D}' := \{D' : D \in \bigcup_{j=1}^r \mathcal{D}_j\}$. For each $D' \in \mathcal{D}'$, D' is either empty or of the form $\theta(D^j_{x \to y})$, so $\phi(D')$ is K_r-divisible by Proposition 4.10.8(i). By (4.32), $\mathcal{D}'$ satisfies (e). So $\mathcal{D}$ satisfies (a)–(e) and is a (γ, γ')-degree balancing set for (U^1, U^2). □

The following result finds copies of the degree balancing sets described in the previous proposition.

Proposition 4.10.11. *Let $1/n \ll \gamma \ll \gamma' \ll 1/k \ll \varepsilon \ll 1/r \leq 1/3$. Let G be an r-partite graph on $(V_1, \dots, V_r)$ with $|V_1| = \dots = |V_r| = n$. Let $\mathcal{P} = \{U^1, \dots, U^k\}$ be a k-partition for G. Suppose that $d_G(v, U^i_j) \geq (1 - 1/(r+1) + \varepsilon)|U^i_j|$ for all $1 \leq i \leq k$, all $1 \leq j \leq r$ and all $v \notin V_j$. Then there exists a $(\gamma, \mathcal{P})$-degree balancing graph $B_{\rm deg} \subseteq G$ such that $B_{\rm deg}$ is locally $\mathcal{P}$-balanced and $\Delta(B_{\rm deg}) < \gamma' n$.*

Proof. Choose γ_1, γ_2 such that $\gamma \ll \gamma_1 \ll \gamma_2 \ll \gamma'$. Proposition 4.10.10 describes a (γ, γ_1^2)-degree balancing set $\mathcal{D}_{i_1,i_2}$ for each pair (U^{i_1}, U^{i_2}) with $1 \leq i_1 < i_2 \leq k$. Let $\mathcal{D} := \bigcup_{1 \leq i_1 < i_2 \leq k} \mathcal{D}_{i_1,i_2}$. We have $|\mathcal{D}| \leq k^2\gamma_1^2 n^2 \leq \gamma_1 n^2$ and each vertex is used as a root vertex in at most $k^2\gamma_1^2 n \leq \gamma_1 n$ elements of $\mathcal{D}$. By (a), we can apply Lemma 4.5.2 (with γ_1, γ_2, $r-1$ and $10r^3$ playing the roles of η, ε, d and b) to find edge-disjoint copies $\phi(D)$ of each $D \in \mathcal{D}$ in G which are compatible with their labellings and satisfy $\Delta(\bigcup \phi(\mathcal{D})) \leq \gamma_2 n$.

Let $G' := G[\mathcal{P}] - \bigcup \phi(\mathcal{D})$ and note that

$$\hat{\delta}(G') \geq (1 - 1/(r+1) + \varepsilon)n - \lceil n/k \rceil - \gamma_2 n \geq (1 - 1/(r+1) + \gamma')n.$$

Apply Lemma 4.6.6 (with γ_2, $\gamma'/2$, $10r^3$ and G' playing the roles of η, ε, b and G) to find an absorbing set $\mathcal{A}$ for $\phi(\mathcal{D})$ in G' such that $\Delta(\bigcup \mathcal{A}) \leq \gamma' n/2$.

Let $B_{\text{deg}} := \bigcup \phi(\mathcal{D}) \cup \bigcup \mathcal{A}$. Then, $\Delta(B_{\text{deg}}) < \gamma' n$. For all $1 \leq i_1 < i_2 \leq k$, $\mathcal{D}_{i_1,i_2}$ is a degree balancing set so $\bigcup \phi(\mathcal{D}_{i_1,i_2})$ is locally $\mathcal{P}$-balanced by (d). Since $B_{\text{deg}}[U^i] = \bigcup \phi(\mathcal{D})[U^i]$ for each $1 \leq i \leq k$, the graph B_{deg} must also be locally $\mathcal{P}$-balanced.

We now check that B_{deg} is a $(\gamma, \mathcal{P})$-degree balancing graph. Let H be any K_r-divisible graph on V satisfying (Q1)–(Q3). Consider any $1 \leq i_1 < i_2 \leq k$. Note that $H[U^{i_1} \cup U^{i_2}]$ satisfies (Q1)–(Q3). Since $\mathcal{D}_{i_1,i_2}$ is a (γ, γ')-degree balancing set for (U^{i_1}, U^{i_2}), there exist $D' \subseteq D$ for each $D \in \mathcal{D}_{i_1,i_2}$ such that $\phi(D')$ is K_r-divisible and, if $\mathcal{D}'_{i_1,i_2} := \{D' : D \in \mathcal{D}_{i_1,i_2}\}$ and $H'_{i_1,i_2} := H \cup \bigcup \phi(\mathcal{D}'_{i_1,i_2})$, then

$$d_{H'_{i_1,i_2}}(v, U^{i_2}_{j_1}) = d_{H'_{i_1,i_2}}(v, U^{i_2}_{j_2})$$

for all $1 \leq j_1, j_2 \leq r$ and all $v \in U^{i_1} \setminus (V_{j_1} \cup V_{j_2})$. Let $B'_{\text{deg}} := \bigcup_{1 \leq i_1 < i_2 \leq k} \phi(\mathcal{D}'_{i_1,i_2})$ and let $H' := H \cup B'_{\text{deg}}$. Note that $V(\bigcup \phi(\mathcal{D}'_{i_1,i_2})) \subseteq U^{i_1} \cup U^{i_2}$ for all $1 \leq i_1 < i_2 \leq k$. So we have $d_{H'}(v, U^i_{j_1}) = d_{H'}(v, U^i_{j_2})$ for all $2 \leq i \leq k$, all $1 \leq j_1, j_2 \leq r$ and all $v \in U^{<i} \setminus (V_{j_1} \cup V_{j_2})$.

It remains to show that B_{deg} and $B_{\text{deg}} - B'_{\text{deg}}$ both have K_r-decompositions. Recall that $\mathcal{A}$ is an absorbing set for $\phi(\mathcal{D})$. So, for any K_r-divisible subgraph D^* of any graph in $\phi(\mathcal{D})$, $\mathcal{A}$ contains an absorber for D^*. Also, A is K_r-decomposable for each $A \in \mathcal{A}$. Since $\phi(D)$ is K_r-divisible for each $D \in \mathcal{D}$ by (d), we see that B_{deg} has a K_r-decomposition. Note that, for each $D \in \mathcal{D}_{i_1,i_2}$, $\phi(D')$ is K_r-divisible by (e) and hence $\phi(D) - \phi(D')$ is also K_r-divisible. So

$$B_{\text{deg}} - B'_{\text{deg}} = \bigcup \mathcal{A} \cup \bigcup_{D \in \mathcal{D}} (\phi(D) - \phi(D'))$$

has a K_r-decomposition. Therefore, B_{deg} is a $(\gamma, \mathcal{P})$-degree balancing graph. □

4.10.3 Finding the balancing graph

Finally, we combine the edge balancing graph and degree balancing graph from Propositions 4.10.7 and 4.10.11 respectively to find a $(\gamma, \mathcal{P})$-balancing graph in G.

Proof of Lemma 4.10.1. Choose constants γ_1 and γ_2 such that $\gamma \ll \gamma_1 \ll \gamma_2 \ll \gamma'$. First apply Proposition 4.10.7 to find a $(\gamma, \mathcal{P})$-edge balancing graph $B_{\text{edge}} \subseteq G$ such that B_{edge} is locally $\mathcal{P}$-balanced and $\Delta(B_{\text{edge}}) < \gamma_1 n$. Now $G' := G - B_{\text{edge}}$ satisfies $d_{G'}(v, U_j^i) \geq (1 - 1/(r+1) + \varepsilon/2)|U_j^i|$ for all $v \notin V_j$, so we can apply Proposition 4.10.11 to find a $(\gamma_2, \mathcal{P})$-degree balancing graph $B_{\text{deg}} \subseteq G'$ such that B_{deg} is locally $\mathcal{P}$-balanced and $\Delta(B_{\text{deg}}) < \gamma' n/2$. Let $B := B_{\text{edge}} \cup B_{\text{deg}}$. Then $\Delta(B) < \gamma' n$ and B is locally $\mathcal{P}$-balanced. Also, since both B_{edge} and B_{deg} are K_r-decomposable, B is K_r-decomposable.

We now show that B is a $(\gamma, \mathcal{P})$-balancing graph. Let H be any K_r-divisible graph on V satisfying (P1) and (P2). Since B_{edge} is a $(\gamma, \mathcal{P})$-edge balancing graph, there exists $B'_{\text{edge}} \subseteq B_{\text{edge}}$ such that $B_{\text{edge}} - B'_{\text{edge}}$ has a K_r-decomposition and $H_1 := H \cup B'_{\text{edge}}$ satisfies

$$e_{H_1}(U_{j_1}^{i_1}, U_{j_2}^{i_2}) = e_{H_1}(U_{j_1}^{i_1}, U_{j_3}^{i_2})$$

for all $1 \leq i_1 < i_2 \leq k$ and all $1 \leq j_1, j_2, j_3 \leq r$ with $j_1 \neq j_2, j_3$.

Note that H_1 is K_r-divisible. Also

$$|d_{H_1}(v, U_{j_2}^i) - d_{H_1}(v, U_{j_3}^i)| \leq |d_H(v, U_{j_2}^i) - d_H(v, U_{j_3}^i)| + \Delta(B_{\text{edge}}) < \gamma n + \gamma_1 n \leq \gamma_2 |U_{j_1}^i|$$

for all $2 \leq i \leq k$, all $1 \leq j_1, j_2, j_3 \leq r$ with $j_1 \neq j_2, j_3$ and all $v \in U_{j_1}^{<i}$. So H_1 satisfies (Q1)–(Q3) with H_1 and γ_2 replacing H and γ. Now, B_{deg} is a $(\gamma_2, \mathcal{P})$-degree balancing graph so there exists $B'_{\text{deg}} \subseteq B_{\text{deg}}$ such that $B_{\text{deg}} - B'_{\text{deg}}$ has a K_r-decomposition and $H_2 := H_1 \cup B'_{\text{deg}}$ satisfies

$$d_{H_2}(v, U_{j_1}^i) = d_{H_2}(v, U_{j_2}^i)$$

for all $2 \leq i \leq k$, all $1 \leq j_1, j_2 \leq r$ and all $v \in U^{<i} \setminus (V_{j_1} \cup V_{j_2})$.

Let $B' := B'_{\text{edge}} \cup B'_{\text{deg}}$. Then $B - B' = (B_{\text{edge}} - B'_{\text{edge}}) \cup (B_{\text{deg}} - B'_{\text{deg}})$ has a K_r-

decomposition. Note that $H \cup B' = H_2$. So B is a $(\gamma, \mathcal{P})$-balancing graph. □

4.11 Proof of Theorem 4.1.1

In this section, we prove our main result, Theorem 4.1.1. The idea is to take a suitable partition $\mathcal{P}$ of $V(G)$, cover all edges in $G[\mathcal{P}]$ by edge-disjoint copies of K_r and then absorb all remaining edges using an absorber which we set aside at the start of the process. However, for the final step to work, we need that the classes of $\mathcal{P}$ have bounded size. A key step towards this is the following lemma which, for a partition $\mathcal{P}$ into a bounded number of parts, finds an approximate K_r-decomposition which covers all edges of $G[\mathcal{P}]$. We then iterate this lemma inductively to get a similar lemma where the parts have bounded size (see Lemma 4.11.2).

Lemma 4.11.1. *Let $1/n \ll \alpha \ll \eta \ll \rho \ll 1/k \ll \varepsilon \ll 1/r \leq 1/3$. Let G be a K_r-divisible graph on $(V_1, \ldots, V_r)$ with $|V_1| = \cdots = |V_r| = n$. Let $\mathcal{P}$ be a k-partition for G. For each $x \in V(G)$, each $U \in \mathcal{P}$ and each $1 \leq j \leq r$, let $0 \leq d_{x,U_j} \leq |U_j|$. Let $G_0 \subseteq G - G[\mathcal{P}]$, $G_1 := G - G_0$ and $R \subseteq G[\mathcal{P}]$. Suppose the following hold for all $U, U' \in \mathcal{P}$ and all $1 \leq j, j_1, j_2 \leq r$ such that $j \neq j_1, j_2$:*

(a) *for all $x \in U_j$, $|d_G(x, U_{j_1}) - d_G(x, U_{j_2})| < \alpha|U_j|$;*

(b) *for all $x \notin V_j$, $d_{G_1}(x, U_j) \geq (\hat{\delta}^{\eta}_{K_r} + \varepsilon)|U_j|$;*

(c) *for all $x \in V(G)$, $d_R(x, U_j) < \rho d_{x,U_j} + \alpha|U_j|$;*

(d) *for all distinct $x, y \in V(G)$, $d_R(\{x, y\}, U_j) < (\rho^2 + \alpha)|U_j|$;*

(e) *for all $x \notin U \cup U' \cup V_{j_1} \cup V_{j_2}$, $|d_R(x, U_{j_1}) - d_R(x, U'_{j_2})| < 3\alpha|U_{j_1}|$;*

(f) *for all $x \notin U$ and all $y \in U$ such that $x, y \notin V_j$,*

$$d_{G_1}(y, N_R(x, U_j)) \geq \rho(1 - 1/(r-1))d_{x,U_j} + \rho^{5/4}|U_j|.$$

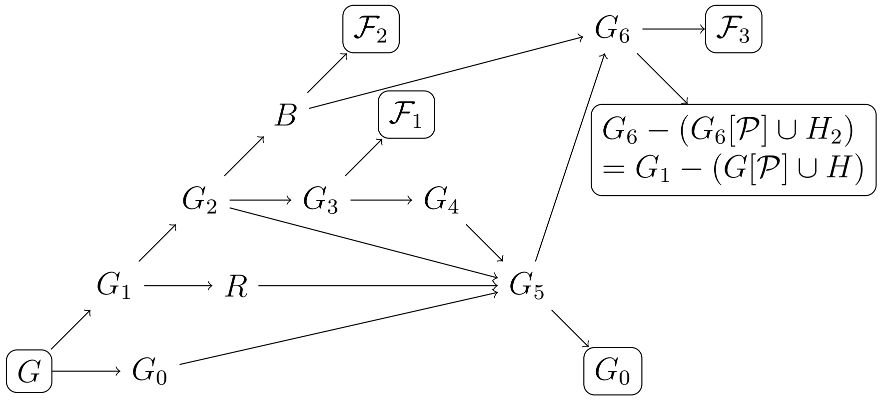

Figure 4.3: Outline for the proof of Lemma 4.11.1.

Then there is a subgraph $H \subseteq G_1 - G[\mathcal{P}]$ such that $G[\mathcal{P}] \cup H$ has a K_r-decomposition and $\Delta(H) \leq 4r\rho n$.

To prove Lemma 4.11.1, we apply Lemma 4.8.1 to cover almost all the edges of $G[\mathcal{P}]$. We then balance the leftover using Lemma 4.10.1. The remaining edges in $G[\mathcal{P}]$ can then be covered using Corollary 4.9.4. The graph R in Lemma 4.11.1 forms the main part of the graph G in Corollary 4.9.4. Conditions (c)–(f) ensure that R is 'quasirandom'.

Proof. Write $\mathcal{P} = \{U^1, \dots, U^k\}$. Let $G_2 := G_1 - R = G - G_0 - R$. Note that Proposition 4.3.1 together with (b) and (c) implies that for any $1 \leq i \leq k$, any $1 \leq j \leq r$ and any $x \notin V_j$,

$$d_{G_2}(x, U^i_j) \geq (\hat{\delta}^{\eta}_{K_r} + \varepsilon - 2\rho)|U^i_j| \geq (1 - 1/(r+1) + \varepsilon/2)|U^i_j|.$$

Choose constants γ_1, γ_2 such that $\eta \ll \gamma_1 \ll \gamma_2 \ll \rho$. Apply Lemma 4.10.1 (with γ_1, γ_2, $\varepsilon/2$, k, G_2, $\mathcal{P}$ playing the roles of γ, γ', ε, k, G, $\mathcal{P}$) to find a $(\gamma_1, \mathcal{P})$-balancing graph $B \subseteq G_2$ such that

$$\Delta(B) < \gamma_2 n \tag{4.33}$$

and B is locally $\mathcal{P}$-balanced. As B is also K_r-decomposable, for all $1 \leq j_1, j_2 \leq r$ and all $x \notin V_{j_1} \cup V_{j_2}$,

$$d_{B[\mathcal{P}]}(x, V_{j_1}) = d_{B[\mathcal{P}]}(x, V_{j_2}). \tag{4.34}$$

Let $G_3 := G_2[\mathcal{P}] - B = G[\mathcal{P}] - R - B$. Then (b), (c) and (4.33) give

$$\hat{\delta}(G_3) \geq (\hat{\delta}^{\eta}_{K_r} + \varepsilon)n - \lceil n/k \rceil - 2\rho n - \gamma_2 n \geq (\hat{\delta}^{\eta}_{K_r} + \varepsilon/2)n.$$

Consider any $1 \leq j_1, j_2 \leq r$ and any $x \notin V_{j_1} \cup V_{j_2}$. Using (a), (e) and (4.34), we have

$$\begin{aligned}|d_{G_3}(x, V_{j_1}) - d_{G_3}(x, V_{j_2})| &\leq |d_{G[\mathcal{P}]}(x, V_{j_1}) - d_{G[\mathcal{P}]}(x, V_{j_2})| + |d_R(x, V_{j_1}) - d_R(x, V_{j_2})| \\ &< \alpha n + 3\alpha n = 4\alpha n.\end{aligned}$$

So we can apply Lemma 4.8.1 (with 4α, η, $\gamma_1/2$, $\varepsilon/2$, G_3 playing the roles of α, η, γ, ε, G) to find $G_4 \subseteq G_3$ such that $G_3 - G_4$ has a K_r-decomposition $\mathcal{F}_1$ and

$$\Delta(G_4) \leq \gamma_1 n/2. \tag{4.35}$$

The graphs G, $G_3 - G_4$ and B are all K_r-divisible (and $G_3 - G_4$ and B are edge-disjoint), so

$$G_5 := G - (G_3 - G_4) - B = (G - G[\mathcal{P}] - B) \cup G_4 \cup R$$

must also be K_r-divisible. Note that $e(G_5 \cap B) = 0$ and $G_5[\mathcal{P}] = G_4 \cup R$. Consider any $1 \leq i \leq k$, any $1 \leq j_1, j_2 \leq r$ and any $x \notin V_{j_1} \cup V_{j_2}$. If $x \notin U^i$, (4.35) and (e) give

$$\begin{aligned}|d_{G_5}(x, U^i_{j_1}) - d_{G_5}(x, U^i_{j_2})| &= |d_{G_4 \cup R}(x, U^i_{j_1}) - d_{G_4 \cup R}(x, U^i_{j_2})| \\ &\leq \Delta(G_4) + |d_R(x, U^i_{j_1}) - d_R(x, U^i_{j_2})| < (\gamma_1/2 + 3\alpha)n < \gamma_1 n.\end{aligned}$$

If $x \in U^i$, then we use (a), that B is locally $\mathcal{P}$-balanced and that $G_4, R \subseteq G[\mathcal{P}]$ to see that

$$\begin{aligned}|d_{G_5}(x, U^i_{j_1}) - d_{G_5}(x, U^i_{j_2})| &\leq |d_G(x, U^i_{j_1}) - d_G(x, U^i_{j_2})| + |d_B(x, U^i_{j_1}) - d_B(x, U^i_{j_2})| \\ &< \alpha n \leq \gamma_1 n.\end{aligned}$$

So (P1) and (P2) in Section 4.10 hold with G_5 and γ_1 replacing H and γ. Since B is a $(\gamma_1, \mathcal{P})$-balancing graph, there exists $B' \subseteq B$ such that $B - B'$ has a K_r-decomposition $\mathcal{F}_2$ and, for all $2 \le i \le k$, all $1 \le j_1, j_2 \le r$ and all $x \in U^{<i} \setminus (V_{j_1} \cup V_{j_2})$,

$$d_{G_5 \cup B'}(x, U^i_{j_1}) = d_{G_5 \cup B'}(x, U^i_{j_2}). \tag{4.36}$$

Write $H_1 := \bigcup_{i=1}^k (B - B')[U^i]$ and let

$$G_6 := G_5 \cup B' - G_0 = (G - G[\mathcal{P}] - G_0 - B) \cup R \cup G_4 \cup B'.$$

Note that

$$G_6[\mathcal{P}] = R \cup G_4 \cup B'[\mathcal{P}] = G_5[\mathcal{P}] \cup B'[\mathcal{P}]. \tag{4.37}$$

We now check conditions (i)–(iv) of Corollary 4.9.4 (with G_6 playing the role of G). Since $G_0 \subseteq G - G[\mathcal{P}]$, (i) follows immediately from (4.36). For (ii), suppose that $2 \le i \le k$ and $x \in U^{<i}$. For any $1 \le j \le r$, using (c), (4.35) and (4.33), we have

$$\begin{aligned} d_{G_6}(x, U^i_j) &\overset{(4.37)}{\le} d_R(x, U^i_j) + \Delta(G_4) + \Delta(B) < \rho d_{x,U^i_j} + \alpha|U^i_j| + \gamma_1 n/2 + \gamma_2 n \\ &\le \rho d_{x,U^i_j} + 2\gamma_2 n. \end{aligned} \tag{4.38}$$

Consider any $y \in N_{G_6}(x, U^i)$. Note that $G_6[U^i] = G_1[U^i] - (B - B')[U^i]$. So, for any $1 \le j \le r$ such that $x, y \notin V_j$, we have

$$\begin{aligned} d_{G_6}(y, N_{G_6}(x, U^i_j)) \ge d_{G_6}(y, N_R(x, U^i_j)) &\ge d_{G_1}(y, N_R(x, U^i_j)) - \Delta(B) \\ &\overset{(f),(4.33)}{\ge} (1 - 1/(r-1))\rho d_{x,U^i_j} + \rho^{5/4}|U^i_j| - \gamma_2 n \\ &\overset{(4.38)}{\ge} (1 - 1/(r-1)) d_{G_6}(x, U^i_j) + \rho^{5/4}|U^i_j| - 3\gamma_2 n \\ &> (1 - 1/(r-1)) d_{G_6}(x, U^i_j) + 9kr\rho^{3/2}|U^i|. \end{aligned}$$

So (ii) holds.

To see that G_6 satisfies property (iii) of Corollary 4.9.4, note that for all $2 \leq i \leq k$ and all distinct $x, x' \in U^{<i}$, (d), (4.33), (4.35) and (4.37) imply that

$$|N_{G_6}(x, U^i) \cap N_{G_6}(x', U^i)| \leq d_R(\{x, x'\}, U^i) + \Delta(G_4) + \Delta(B) \\ < (\rho^2 + \alpha)|U^i| + \gamma_1 n/2 + \gamma_2 n \leq 2\rho^2 |U^i|.$$

Finally, by (c), (4.33), (4.35) and (4.37), for any $y \in U^i$, we have that

$$d_{G_6}(y, U^{<i}) \leq \Delta(R) + \Delta(G_4) + \Delta(B) \leq 3\rho n/2 \leq 2k\rho |U_1^i|,$$

and (iv) holds. Hence we can apply Corollary 4.9.4 to G_6 to find a subgraph $H_2 \subseteq G_6 - G_6[\mathcal{P}]$ such that $G_6[\mathcal{P}] \cup H_2$ has a K_r-decomposition $\mathcal{F}_3$ and $\Delta(H_2) \leq 3r\rho n$. Set $H := H_1 \cup H_2 \subseteq G_1 - G[\mathcal{P}]$. We have $\Delta(H) \leq \Delta(H_1) + \Delta(H_2) \leq \Delta(B) + \Delta(H_2) \leq 4r\rho n$. Now,

$$\begin{aligned} G[\mathcal{P}] \cup H &= G_2[\mathcal{P}] \cup R \cup H = G_3 \cup R \cup H \cup B[\mathcal{P}] \\ &= \bigcup \mathcal{F}_1 \cup G_4 \cup R \cup H \cup B[\mathcal{P}] = \bigcup \mathcal{F}_1 \cup G_5[\mathcal{P}] \cup H_1 \cup H_2 \cup B[\mathcal{P}] \\ &= \bigcup (\mathcal{F}_1 \cup \mathcal{F}_2) \cup G_5[\mathcal{P}] \cup H_2 \cup B'[\mathcal{P}] \overset{(4.37)}{=} \bigcup (\mathcal{F}_1 \cup \mathcal{F}_2) \cup G_6[\mathcal{P}] \cup H_2 \\ &= \bigcup (\mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3). \end{aligned}$$

So $G[\mathcal{P}] \cup H$ has a K_r-decomposition $\mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3$. □

We now iterate Lemma 4.11.1, applying it to each partition $\mathcal{P}_i$ in a partition sequence $\mathcal{P}_1, \dots, \mathcal{P}_\ell$ for G. This allows us to cover all of the edges in $G[\mathcal{P}_\ell]$ by edge-disjoint copies of K_r, leaving only a small remainder in $\bigcup_{U \in \mathcal{P}_\ell} G[U]$.

Lemma 4.11.2. *Let $1/m \ll \alpha \ll \eta \ll \rho \ll 1/k \ll \varepsilon \ll 1/r \leq 1/3$. Let G be a K_r-divisible graph on $(V_1, \dots, V_r)$ with $|V_1| = \dots = |V_r| = n$. Let $\mathcal{P}_1, \dots, \mathcal{P}_\ell$ be a $(1, k, \hat{\delta}^\eta_{K_r} + \varepsilon/2, m)$-partition sequence for G. For each $1 \leq q \leq \ell$, each $1 \leq j \leq r$, each $U \in \mathcal{P}_q$ and each $x \in V(G)$, let $0 \leq d_{x,U_j} \leq |U_j|$ be given. Let $\mathcal{P}_0 := \{V(G)\}$ and, for each $0 \leq q \leq \ell$,* *let $G_q := G[\mathcal{P}_q]$. Let $R_1, \dots, R_\ell$ be a sequence of graphs such that $R_q \subseteq G_q - G_{q-1}$ for each q. Suppose the following hold for all $1 \le q \le \ell$, all $1 \le j, j_1, j_2 \le r$ such that $j \neq j_1, j_2$, all $W \in \mathcal{P}_{q-1}$, all distinct $x, y \in W$ and all $U, U' \in \mathcal{P}_q[W]$:*

(i) *if $q \ge 2$, $\mathcal{P}_q[W]$ is a $(1, k, \hat{\delta}^\eta_{K_r} + \varepsilon)$-partition for $G[W]$;*

(ii) *if $x \in U_j$, $|d_G(x, U_{j_1}) - d_G(x, U_{j_2})| < \alpha |U_j|$;*

(iii) *$d_{R_q}(x, U_j) < \rho d_{x,U_j} + \alpha |U_j|$;*

(iv) *$d_{R_q}(\{x, y\}, U_j) < (\rho^2 + \alpha)|U_j|$;*

(v) *if $x \notin U \cup U' \cup V_{j_1} \cup V_{j_2}$, $|d_{R_q}(x, U_{j_1}) - d_{R_q}(x, U'_{j_2})| < 3\alpha |U_{j_1}|$;*

(vi) *if $x \notin U$, $y \in U$ and $x, y \notin V_j$, then*

$$d_{G'_{q+1}}(y, N_{R_q}(x, U_j)) \ge \rho(1 - 1/(r-1)) d_{x,U_j} + \rho^{5/4} |U_j|$$

where $G'_{q+1} := G_{q+1} - R_{q+1}$ if $q \le \ell - 1$ and $G'_{\ell+1} := G$.

Then there is a subgraph $H \subseteq \bigcup_{U \in \mathcal{P}_\ell} G[U]$ such that $G - H$ has a K_r-decomposition.

Proof. We will use induction on ℓ. If $\ell = 1$, apply Lemma 4.11.1 (with $\varepsilon/2$, $\mathcal{P}_1$, R_1 and the empty graph playing the roles of ε, $\mathcal{P}$, R and G_0) to find $H' \subseteq G - G[\mathcal{P}_1]$ such that $G[\mathcal{P}_1] \cup H'$ has a K_r-decomposition. Letting $H := G - G[\mathcal{P}_1] - H' \subseteq \bigcup_{U \in \mathcal{P}_\ell} G[U]$, shows the result holds for $\ell = 1$.

Suppose then that $\ell \ge 2$ and the result holds for all smaller ℓ. Note that for each $1 \le j \le r$, each $x \notin V_j$ and each $U \in \mathcal{P}_1$, $d_{G[\mathcal{P}_2] - R_2}(x, U_j) \ge (\hat{\delta}^\eta_{K_r} + \varepsilon/3)|U_j|$, since R_2 satisfies (iii) and $\mathcal{P}_1, \dots, \mathcal{P}_\ell$ is a $(1, k, \hat{\delta}^\eta_{K_r} + \varepsilon/2, m)$-partition sequence for G. So we may apply Lemma 4.11.1 (with $\varepsilon/3$, $\mathcal{P}_1$, R_1, G and $(G - G[\mathcal{P}_2]) \cup R_2$ playing the roles of ε, $\mathcal{P}$, R, G and G_0) to find $H' \subseteq G[\mathcal{P}_2] - (G[\mathcal{P}_1] \cup R_2)$ such that $G[\mathcal{P}_1] \cup H'$ has a K_r-decomposition $\mathcal{F}_1$ and $\Delta(H') \le 4r\rho n$. Let $G^* := G - G[\mathcal{P}_1] - H' = G - \bigcup \mathcal{F}_1$, so G^* is K_r-divisible. Observe that $G^* = \bigcup_{U \in \mathcal{P}_1} G^*[U]$, so $G^*[U]$ is K_r-divisible for each $U \in \mathcal{P}_1$.

Consider any $U \in \mathcal{P}_1$. We check that

$$G^*[U], \mathcal{P}_2[U], \dots, \mathcal{P}_\ell[U], R_2[U], \dots, R_\ell[U]$$

satisfy the conditions of Lemma 4.11.2. Since $\Delta(H') \leq 4r\rho n \leq \varepsilon n/4k^2$, $\mathcal{P}_2[U]$ is a $(1, k, \hat{\delta}^\eta_{K_r} + \varepsilon/2)$-partition for $G^*[U]$. For any $3 \leq q \leq \ell$ and any $W \in \mathcal{P}_{q-1}$, $G^*[W] = G[W]$ since $H' \subseteq G[\mathcal{P}_2]$. So (i) holds and $\mathcal{P}_2[U], \dots, \mathcal{P}_\ell[U]$ is a $(1, k, \hat{\delta}^\eta_{K_r} + \varepsilon/2, m)$-partition sequence for $G^*[U]$. For (ii), note that for any $2 \leq q \leq \ell$, any $1 \leq j \leq r$, any $U' \in \mathcal{P}_q[U]$ and any $x \in U'$, $d_{G^*}(x, U'_j) = d_G(x, U'_j)$. Conditions (iii)–(v) are automatically satisfied. To see that (vi) holds, note that for any $2 \leq q \leq \ell$ and any $U' \in \mathcal{P}_q[U]$, $G^*_{q+1}[U'] = G_{q+1}[U']$ since $H' \subseteq G[\mathcal{P}_2]$. So we can apply the induction hypothesis to $G^*[U], \mathcal{P}_2[U], \dots, \mathcal{P}_\ell[U], R_2[U], \dots, R_\ell[U]$ to obtain a subgraph $H_U \subseteq \bigcup_{U' \in \mathcal{P}_\ell[U]} G^*[U']$ such that $G^*[U] - H_U$ has a K_r-decomposition $\mathcal{F}_U$. Set $H := \bigcup_{U \in \mathcal{P}_1} H_U$. Then, $H \subseteq \bigcup_{U \in \mathcal{P}_\ell} G[U]$ and $G - H$ has a K_r-decomposition $\mathcal{F}_1 \cup \bigcup_{U \in \mathcal{P}_1} \mathcal{F}_U$. □

We are now ready to prove Theorem 4.1.1.

Proof of Theorem 4.1.1. Let $n_0 \in \mathbb{N}$ and $\eta > 0$ be such that $1/n_0 \ll \eta \ll \varepsilon$ and choose additional constants η_1, m', α, ρ and k such that

$$1/n_0 \ll \eta_1 \ll 1/m' \ll \alpha \ll \eta \ll \rho \ll 1/k \ll \varepsilon.$$

Let G be any K_r-divisible graph on $(V_1, \dots, V_r)$ with $|V_1| = \dots = |V_r| = n \geq n_0$ and $\hat{\delta}(G) \geq (\hat{\delta}^\eta_{K_r} + \varepsilon)n$. Apply Lemma 4.7.2 to find an $(\alpha, k, \hat{\delta}^\eta_{K_r} + \varepsilon - \alpha, m)$-partition sequence $\mathcal{P}_1, \dots, \mathcal{P}_\ell$ for G where $m' \leq m \leq km'$. So in particular, by (S3), for each $1 \leq q \leq \ell$, all $1 \leq j_1, j_2, j_3 \leq r$ with $j_1 \neq j_2, j_3$, each $U \in \mathcal{P}_q$ and each $x \in U_{j_1}$,

$$|d_G(x, U_{j_2}) - d_G(x, U_{j_3})| < \alpha |U_{j_1}|. \tag{4.39}$$

Let $\mathcal{P}_0 := \{V(G)\}$ and $G_q := G[\mathcal{P}_q]$ for $0 \leq q \leq \ell$. Note that $\hat{\delta}^\eta_{K_r} + \varepsilon - \alpha \geq 1 - 1/r + \varepsilon$ (with room to spare) by Proposition 4.3.1. So we can apply Corollary 4.7.5 to find a sequence of graphs $R_1, \dots, R_\ell$ such that $R_q \subseteq G_q - G_{q-1}$ for each $1 \le q \le \ell$ and the following holds. For all $1 \le q \le \ell$, all $1 \le j, j' \le r$, all $W \in \mathcal{P}_{q-1}$, all distinct $x, y \in W$ and all $U, U' \in \mathcal{P}_q[W]$:

(a) $d_{R_q}(x, U_j) < \rho d_{G_q}(x, U_j) + \alpha|U_j|$;

(b) $d_{R_q}(\{x, y\}, U_j) < (\rho^2 + \alpha)|U_j|$;

(c) $|d_{R_q}(x, U_j) - d_{R_q}(x, U'_{j'})| < 3\alpha|U_j|$ if $x \notin U \cup U' \cup V_j \cup V_{j'}$;

(d) $d_{G'_{q+1}}(y, N_{R_q}(x, U_j)) \ge \rho(1 - 1/(r-1))d_{G_q}(x, U_j) + \rho^{5/4}|U_j|$ if $x \notin U, y \in U$ and $x, y \notin V_j$, where $G'_{q+1} := G_{q+1} - R_{q+1}$ if $q \le \ell - 1$ and $G'_{\ell+1} := G$.

Let $\mathcal{H} := \{G[U] : U \in \mathcal{P}_\ell\}$. Each $H \in \mathcal{H}$ satisfies $|H| \le rm$. Note that

$$\hat{\delta}(G[\mathcal{P}_1] - R_1) \ge (\hat{\delta}^{\eta}_{K_r} + \varepsilon)n - \lceil n/k \rceil - 2\rho n > (1 - 1/(r+1) + \varepsilon/2)n.$$

So we can apply Lemma 4.6.6 (with η_1, α, rm and $G[\mathcal{P}_1] - R_1$ playing the roles of η, ε, b and G) to find an absorbing set $\mathcal{A}$ for $\mathcal{H}$ inside $G[\mathcal{P}_1] - R_1$ such that $A^* := \bigcup \mathcal{A}$ satisfies $\Delta(A^*) \le \alpha n$.

Let $G^* := G - A^*$. Note that both G and A^* are K_r-divisible, so G^* is K_r-divisible. Since $\Delta(A^*) \le \alpha n$ and $A^* \subseteq G[\mathcal{P}_1]$, $\mathcal{P}_1, \dots, \mathcal{P}_\ell$ is an $(1, k, \hat{\delta}^{\eta}_{K_r} + \varepsilon/2, m)$-partition sequence for G^*. For each $1 \le q \le \ell$, each $1 \le j \le r$, each $U \in \mathcal{P}_q$ and each $x \in V(G)$, set $d_{x,U_j} := d_{G_q}(x, U_j)$. Using (4.39), (a)–(d) and that $A^* \subseteq G[\mathcal{P}_1]$, we see that G^*, the partition sequence $\mathcal{P}_1, \dots, \mathcal{P}_\ell$ and the sequence of graphs $R_1, \dots, R_\ell$ satisfy properties (i)–(vi) of Lemma 4.11.2 (with $\varepsilon - \alpha$ playing the role of ε). So we may apply Lemma 4.11.2 to find $H \subseteq \bigcup_{U \in \mathcal{P}_\ell} G^*[U]$ such that $G^* - H$ has a K_r-decomposition $\mathcal{F}_1$.

Note that H is a K_r-divisible subgraph of $\bigcup_{U \in \mathcal{P}_\ell} G[U]$, so for each $U \in \mathcal{P}_\ell$, $H[U] \subseteq G[U]$ is K_r-divisible. Since $\mathcal{A}$ is an absorbing set for $\mathcal{H}$, it contains a distinct absorber for each $H[U]$. So $H \cup A^*$ has a K_r-decomposition $\mathcal{F}_2$. Thus $G = (G^* - H) \cup (H \cup A^*)$ has a K_r-decomposition $\mathcal{F}_1 \cup \mathcal{F}_2$. □

CHAPTER 5

ON THE EXACT DECOMPOSITION THRESHOLD FOR EVEN CYCLES

5.1 Introduction

Let F and G be graphs. We say that G has an F*-decomposition* (or is F*-decomposable*) if its edge set can be partitioned into copies of F. One of the first results in the study of graph decompositions was due to Kirkman [51] who gave conditions for a clique to have a K_3-decomposition. His result was generalised by Wilson [82] who determined when large cliques have F-decompositions for arbitrary F. When G is not a clique, the problem becomes more challenging and the corresponding decision problem is NP-complete [27].

Clearly, every graph which has an F-decomposition must satisfy certain vertex degree and edge divisibility conditions. There have been many recent developments bounding the F-decomposition threshold, that is, the minimum degree which ensures an F-decomposition in any large graph satisfying the necessary divisibility conditions. General results on the F-decomposition threshold establishing a close connection to its fractional counterpart are obtained in [7] and [38]. Moreover, [7] determines the asymptotic decomposition threshold for even cycles and [38] generalises this to arbitrary bipartite graphs. The results in [7] and [38] can be combined with bounds for the fractional version of this problem in [6] and [28] to obtain good explicit bounds on the F-decomposition threshold. Corresponding results for the multipartite setting (with applications to the completion of partially filled Latin squares) were considered in [8], [17] and [60]. The only known exact minimum degree bound (prior to Theorem 5.1.2) was obtained by Yuster [85] who studied the case when F is a tree.

From here on, we restrict our attention to the case when F is a cycle. We say that G is C_k*-divisible* if $e(G)$ is divisible by k and every vertex of G has even degree. Note that any graph which has a C_k-decomposition is necessarily C_k-divisible. For each $k \in \mathbb{N}$ with $k \geq 2$, let us define

$$\delta_k := \begin{cases} 2/3 & \text{if } k = 2, \\ 1/2 & \text{if } k \geq 3. \end{cases}$$

Barber, Kühn, Lo and Osthus [7] proved asymptotically best possible minimum degree bounds for a graph to have a C_{2k}-decomposition.

Theorem 5.1.1 ([7]). *Let $k \in \mathbb{N}$ with $k \geq 2$. For each $\varepsilon > 0$, there is an n_0 such that every C_{2k}-divisible graph G on $n \geq n_0$ vertices with $\delta(G) \geq (\delta_k + \varepsilon)n$ has a C_{2k}-decomposition.*

In this thesis we remove the linear error term from Theorem 5.1.1 to obtain best possible minimum degree bounds for cycles of all even lengths except length six. We structure the proof into extremal cases where we construct the decompositions directly and non-extremal cases where the iterative absorption approach of [7] and [38] remains effective. In Proposition 5.1.4, we give constructions which show that our bounds are best possible.

Theorem 5.1.2. *Let $k \in \mathbb{N}$ with $k = 2$ or $k \geq 4$. There is an n_0 such that every C_{2k}-divisible graph G on $n \geq n_0$ vertices with*

$$\delta(G) \geq \begin{cases} 2n/3 - 1 & \textit{if } k = 2, \\ n/2 & \textit{if } k \geq 4 \end{cases}$$

has a C_{2k}-decomposition.

It is an open problem to determine the exact minimum degree guaranteeing a C_6-

decomposition, this is discussed in more detail in Section 5.8.

Along the way to proving Theorem 5.1.2, we also obtain a bipartite version of Theorem 5.1.1 which is stated as Theorem 5.1.3 below. If G is a bipartite graph with vertex classes A and B, we introduce the following variant on the minimum degree. Given $0 \leq \delta \leq 1$, we will write $\delta_{\mathrm{bip}}(G) \geq \delta$ if, for each $v \in A$, $d_G(v) \geq \delta|B|$ and for each $v \in B$, $d_G(v) \geq \delta|A|$. This definition is convenient when the bipartite graph is not balanced. Cavenagh [19] already studied C_4-decompositions and proved a bound of $\delta_{\mathrm{bip}}(G) \geq 95/96$ ensures a C_4-decomposition. Theorem 5.1.3 is asymptotically best possible, see Proposition 5.1.5.

Theorem 5.1.3. *Let $k \in \mathbb{N}$ with $k \geq 2$. For each $\varepsilon > 0$, there is an n_0 such that every C_{2k}-divisible bipartite graph $G = (A, B)$ with $n_0 \leq |A| \leq |B| \leq 2|A|$ and $\delta_{\mathrm{bip}}(G) \geq \delta_k + \varepsilon$ has a C_{2k}-decomposition.*

5.1.1 Extremal graphs

In this section we provide extremal constructions which show that Theorem 5.1.2 is best possible and Theorem 5.1.3 is asymptotically so.

Proposition 5.1.4. (i) *There are infinitely many C_4-divisible graphs G with $\delta(G) \geq 2|G|/3 - 2$ and no C_4-decomposition.*

(ii) *Let $k \in \mathbb{N}$, $k \geq 2$. There are infinitely many C_{2k}-divisible graphs G with $\delta(G) \geq |G|/2 - 1$ and no C_{2k}-decomposition.*

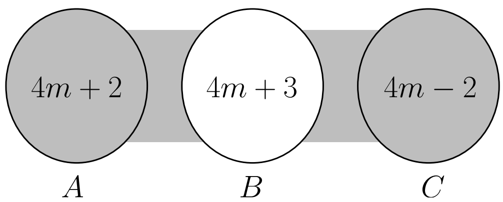

Figure 5.1: The extremal graph for C_4, Proposition 5.1.4(i). All possible edges are present in the shaded regions.

Proof. We begin with (i). Let $m \in \mathbb{N}$ and let A, B, C be disjoint sets of vertices of sizes $4m+2, 4m+3, 4m-2$ respectively. Form a graph G which has vertex set $A \cup B \cup C$. The edge set of G is such that A and C form cliques and G contains all possible edges between $A \cup C$ and B. For each $v \in V(G)$, $d(v) \in \{8m+4, 8m\}$, so every vertex has even degree and $\delta(G) = 8m = 2|G|/3 - 2$. We also have

$$e(G) = \binom{4m+2}{2} + 8m(4m+3) + \binom{4m-2}{2} = 4(12m^2 + 5m + 1).$$

So G is C_4-divisible. Any copy of C_4 in G must use an even number of edges from $G[A]$. But $e(A) = \binom{4m+2}{2} = (2m+1)(4m+1)$ is odd. Hence, G does not have a C_4-decomposition.

For (ii), let n be such that $n \equiv 2k+1 \mod 4k$ and let G be the union of two vertex-disjoint copies of K_n. Every vertex in G has degree $n-1 = |G|/2-1$ which is even and $2k$ divides $e(G) = n(n-1)$. So G is C_{2k}-divisible. But G does not have a C_{2k}-decomposition since $2k$ does not divide $\binom{n}{2}$. □

Proposition 5.1.5. (i) *There are infinitely many C_4-divisible bipartite graphs $G = (A, B)$ with $|A| = |B|$, $\delta(G) \geq 2|A|/3 - 2$ and no C_4-decomposition.*

(ii) *Let $k \in \mathbb{N}$, $k \geq 2$. There are infinitely many C_{2k}-divisible bipartite graphs $G = (A, B)$ with $|A| = |B|$, $\delta(G) \geq |A|/2 - 1$ and no C_{2k}-decomposition.*

Proof. First, we prove (i). Let $m \in \mathbb{N}$. Start with independent sets $V_1, \dots, V_6$ each of size $2m+1$ and add all edges between V_i and V_{i+1} for each $1 \leq i \leq 6$ (consider indices modulo 6). Remove one copy of C_6 between V_5 and V_6 and let G denote the resulting graph. Then G is bipartite with vertex classes $A := V_1 \cup V_3 \cup V_5$ and $B := V_2 \cup V_4 \cup V_6$ of size $6m+3$. The degree of each vertex in G is either $4m+2$ or $4m$, both of which are even, and $\delta(G) = 4m = 2|A|/3 - 2$. The number of edges in G is $6(2m+1)^2 - 6 = 24m(m+1)$. So G is C_4-divisible. But G does not have a C_4-decomposition. To see this, note that any copy of C_4 in G must use an even number of edges between V_1 and V_2 but $e_G(V_1, V_2) = (2m+1)^2$ is odd.

Now we consider (ii). For each $n \in N$, let $K^-_{n,n}$ denote the graph formed by removing a perfect matching from $K_{n,n}$. Suppose first that k is even. Choose $m \in \mathbb{N}$ such that $m \equiv k+1 \mod 2k$. Let G be the vertex-disjoint union of two copies of $K^-_{m,m}$. Then G is a balanced bipartite graph with vertex classes of size $2m$. Each vertex in G has degree $m-1 \equiv k \mod 2k$ which is even and

$$e(G) = 2(m-1)m \equiv 2k(k+1) \equiv 0 \mod 2k.$$

So G is C_{2k}-divisible. But G does not have a C_{2k}-decomposition because

$$e(K^-_{m,m}) = (m-1)m \equiv k(k+1) \equiv k \mod 2k.$$

Now we consider k odd. Choose $m \in \mathbb{N}$ such that $4m \equiv k-1 \mod 2k$ (i.e., choose $m \equiv (k-1)/4 \mod 2k$ if $k \equiv 1 \mod 4$ and $m \equiv (3k-1)/4 \mod 2k$ if $k \equiv 3 \mod 4$). Let G be the vertex-disjoint union of $K^-_{2m+1,2m+1}$ and $K_{2m,2m}$, so that G is a balanced bipartite graph with vertex classes of size $4m+1$. Note that each vertex in G has degree $2m$ which is even and, since

$$e(G) = 2m(4m+1) \equiv 2mk \equiv 0 \mod 2k,$$

G is C_{2k}-divisible. However, $2k$ does not divide

$$e(K^-_{2m+1,2m+1}) - e(K_{2m,2m}) = 2m,$$

so $K^-_{2m+1,2m+1}$ and $K_{2m,2m}$ (and hence also G) are not C_{2k}-decomposable. □

5.1.2 Outline of the proof

Our argument is based on an iterative absorption approach. This method was introduced in [53] and further developed in the context of F-decompositions in [7] and [38]. In our setting, the idea of iterative absorption is as follows. Let U be a subset of $V(G)$ of constant size and let $U_0 \supseteq U_1 \supseteq \cdots \supseteq U_\ell$ be a decreasing sequence of sets of vertices with $U_\ell := U$. We use an iterative argument to cover almost all edges of G by copies of C_{2k}. Here, it is to our advantage that C_{2k} is bipartite since we can always greedily find an approximate decomposition of G using the Erdős-Stone theorem (this is not true for F-decompositions in general). At the end of the i^{th} iteration, we are left with a diminishing subgraph $H_i \subseteq G[U_i]$ until, eventually, all that remains is a small leftover $H \subseteq G[U]$. But we have prepared for H by removing an "absorber" at the start of the process, a subgraph A of G with the property that $A \cup H$ has a C_{2k}-decomposition. This absorber must be able to deal with all possible leftover graphs in $G[U]$, but this is feasible since U only has constant size. Thus we obtain a C_{2k}-decomposition of G. So the proof of Theorem 5.1.1 using iterative absorption relies on two parts:

1. G contains an absorber and
2. we can cover all edges in $G - G[U]$.

When we relax the minimum degree condition on G to prove Theorem 5.1.2, one or both of these properties can become considerably more challenging to attain.

When the cycle has length at least eight, we need to show that a minimum degree of $|G|/2$ suffices to find a C_{2k}-decomposition. If G satisfies a certain expansion property this guarantees many disjoint paths between any pair of vertices, which enables us to show that (1) and (2) still hold. If G is not an expander, then G has one of two well-defined extremal structures. Either G resembles a complete bipartite graph or the disjoint union of two cliques. In either case, we can construct C_{2k}-decompositions directly. We first deal with any edges or vertices which are unusual in some way to leave behind disjoint graphs or bipartite graphs which have high minimum degree. These can be decomposed using the existing Theorem 5.1.1 or the bipartite version, Theorem 5.1.3, (which is proved in Section 5.6).

Cycles of length four are treated separately since in this case we require a higher minimum degree, namely $\delta(G) \geq 2|G|/3 - 1$. In fact, this minimum degree is sufficient (with room to spare) for (2) and it is only finding an absorber which causes any difficulty. We are able to show that any graph which does not contain an absorber will, as in the previous case, have a well-defined structure and we find a C_4-decomposition directly.

This chapter is organised as follows. In Section 5.2, we introduce the notation which will be used throughout. We construct absorbers in Section 5.3. We prove Theorem 5.1.2 for $k = 2$ in Section 5.4 and for $k \geq 4$ in Section 5.5 (see Table 5.1 for a guide). As mentioned above, these proofs rely on decomposition results when the host graph G is bipartite (see Theorem 5.1.3) and when G is an expander (see Theorem 5.5.2). These results are proved in Sections 5.6 and 5.7 respectively.

	$\mathbf{C_4}$	$\mathbf{C_{8+}}$
non-extremal	Lemma 5.4.1	Theorem 5.5.2
extremal	Lemma 5.4.2	Lemmas 5.5.3 and 5.5.7

Table 5.1: Components in the proof of Theorem 5.1.2.

5.2 Notation and tools

Let G be a graph and let $\mathcal{P} = \{U_1, \ldots, U_k\}$ be a partition of $V(G)$. We write $G[U_1]$ for the subgraph of G induced by U_1 and $G[U_1, U_2]$ for the bipartite subgraph of G induced by the vertex classes U_1 and U_2. We write $G[\mathcal{P}] := G[U_1, \ldots, U_k]$ for the k-partite subgraph of G induced by the partition $\mathcal{P}$. We say the partition $\mathcal{P}$ is *equitable* if its parts differ in size by at most one.

Let $U, V \subseteq V(G)$. We write $E_G(U) := E(G[U])$ and $e_G(U) := e(G[U])$. If U and V are disjoint, we let $E_G(U, V) := E(G[U, V])$ and $e_G(U, V) := e(G[U, V])$. For any $v \in V(G)$, $N_G(v, U) := N_G(v) \cap U$ and $d_G(v, U) := |N_G(v, U)|$. Let H be a graph. We write $G - H$ for the graph with vertex set $V(G)$ and edge set $E(G) \setminus E(H)$. We write $G \setminus H$ for the subgraph of G induced by the vertex set $V(G) \setminus V(H)$. (Note that, in general, $G - H \neq G \setminus H$.)

Let F and G be graphs and let $\eta > 0$. We say that a collection $\mathcal{F}$ of edge-disjoint copies of F in G is an *η-approximate F-decomposition of G* if $e(G - \bigcup \mathcal{F}) \leq \eta |G|^2$. In this chapter, the graph F will always be bipartite, so we can greedily apply the Erdős-Stone theorem to find an η-approximate F-decomposition of any large graph G. We say that G is *2-divisible* if every vertex in G has even degree.

5.3 Absorbers

As described earlier, the main idea in the proof of the non-extremal cases of Theorem 5.1.2 is to cover as many edges of G as possible with copies of C_{2k} using an iterative approach. Then, as long as only a small number of edges remain, we can "absorb" these using a special graph which was reserved at the start of the process. Let H and H' be vertex-disjoint graphs. The graph A is an *F-absorber* for H if both A and $A \cup H$ have F-decompositions. An *$(H, H')_F$-transformer* is a graph T which is edge-disjoint from H and H' and is such that both $T \cup H$ and $T \cup H'$ have F-decompositions. Note that if H' has an F-decomposition, then $T \cup H'$ is an F-absorber for H. So we can use transformers to build an absorber.

The following fact follows directly from H being Eulerian.

Fact 5.3.1. *Let H be any connected 2-divisible graph and let C be a cycle of length $e(H)$. There is a graph homomorphism ϕ from C to H that is edge-bijective.*

We will make use of the following graphs. For any $i, k \in \mathbb{N}$, define $L(i, k)$ to be the graph consisting of i copies of C_{2k} with exactly one common vertex. For any graph H, we say that H^{con} is a *C_{2k}-connector* for H if:

- $H \cup H^{\text{con}}$ is connected and
- H^{con} has a C_{2k}-decomposition.

The following simple procedure finds a C_{2k}-connector for H. Suppose H is not connected and choose vertices u and v which lie in separate components of H. Form a copy of C_{2k} containing these vertices by adding two edge-disjoint paths of length k between u and v. If the resulting graph H_1 is not connected, repeat this process on H_1. Eventually, a connected graph H' is obtained with $|H'| \leq (2k-1)|H|$. The graph $H^{\mathrm{con}} := H' - H$ is a C_{2k}-connector for H.

5.3.1 Absorbers for long cycles

The following simple transformer construction suits our purpose. Let H be a connected 2-divisible graph and let $C = u_1 u_2 \dots u_h$ be a cycle of length $h := e(H)$ which is vertex-disjoint from H. Let ϕ be a graph homomorphism from C to H that is edge-bijective. For each $1 \leq i \leq h$, let P_i be a path of length k from u_i to $\phi(u_i)$ and let Q_i be a path of length $k-1$ from u_{i+1} to $\phi(u_i)$ (we consider indices modulo h). Suppose that the paths P_i, Q_i are internally disjoint and that they are edge-disjoint from H and C. Note that for each $1 \leq i \leq h$, $u_i u_{i+1} \cup P_i \cup Q_i$ and $\phi(u_i u_{i+1}) \cup P_{i+1} \cup Q_i$ form copies of C_{2k}. So $T := \bigcup_{i=1}^{h} (P_i \cup Q_i)$ is a $(C, H)_{C_{2k}}$-transformer and $|T| = 2ke(H)$.

Lemma 5.3.2. *Let $k \in \mathbb{N}$, $k \geq 4$ and $1/n \ll 1/m' \ll 1/m \ll 1/k$. Let G be a graph on n vertices and let $U \subseteq V(G)$ with $|U| = m$. Suppose that between any pair of vertices $x, y \in V(G)$ there are at least m' internally disjoint paths of length $k-1$. Then G contains a C_{2k}-divisible subgraph A^* such that $|A^*| \leq 2^{m^2}$ and if H is any C_{2k}-divisible graph on U that is edge-disjoint from A^* then $A^* \cup H$ has a C_{2k}-decomposition.*

Proof. Let $H_1, \dots, H_p$ be an enumeration of all possible C_{2k}-divisible graphs on U (note that $p \leq 2^{\binom{m}{2}}$). We will find an absorber for each H_i. For each $1 \leq i \leq p$, find an edge-disjoint C_{2k}-connector $H_i^{\mathrm{con}} \subseteq G - G[U]$ using the procedure outlined above. Each $H_i' := H_i \cup H_i^{\mathrm{con}}$ is C_{2k}-divisible and $|H_i'| \leq (2k-1)m$.

For each $1 \leq i \leq p$, let $h_i := e(H_i')$, let C^i be a cycle of length h_i and let J_i be a copy of the graph $L(h_i/2k, k)$, defined at the beginning of this section. Find copies of

C^i and J_i in G which are vertex-disjoint from each other and from the graphs H_i'. Find a $(H_i', C^i)_{C_{2k}}$-transformer T_i and a $(C^i, J_i)_{C_{2k}}$-transformer T_i' in G (such that T_i and T_i' are edge-disjoint and avoid all edges fixed so far). It is easy to find these transformers using the construction described above since G contains many internally disjoint paths of length $k-1$ (and hence k also) between any pair of vertices. Then $T_i \cup C^i \cup T_i' \cup J_i$ is an absorber for H_i'. Hence $A_i := H_i^{\text{con}} \cup T_i \cup C^i \cup T_i' \cup J_i$ is an absorber for H_i. Letting $A^* := \bigcup_{i=1}^{p} A_i$ and noting $|A^*| \le 4kh_ip \le 2^{m^2}$ completes the proof. □

5.3.2 C_4-absorbers

For cycles of length four, we will require the following alternative construction of a transformer. This is exactly the construction given in [7] and it is illustrated in Figure 5.2.

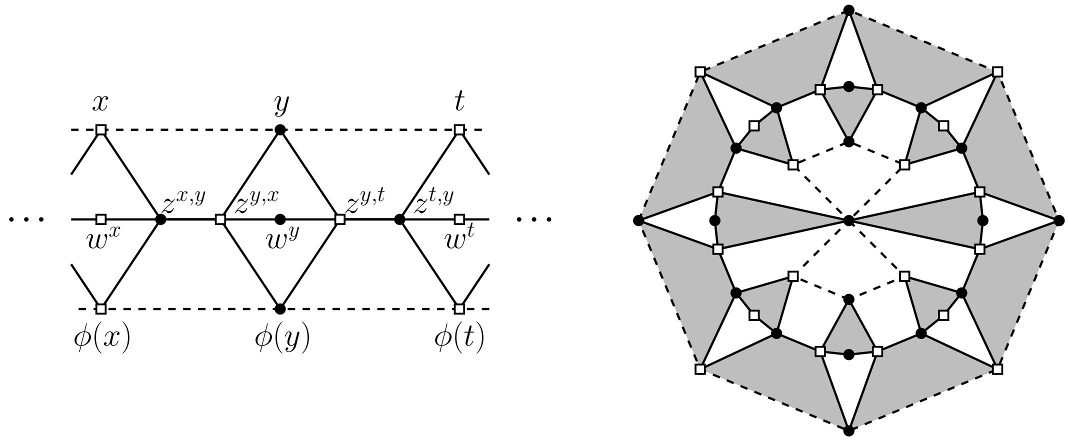

Figure 5.2: The transformer construction for cycles of length four (left) and a $(C_8, L(2,2))_{C_4}$-transformer (right). The square/round vertices give a bipartition of the transformer which is used by Lemma 5.3.3.

Let H be a connected, C_4-divisible graph and let C be a cycle of length $e(H)$. Suppose H and C are vertex-disjoint. Let ϕ be an edge-bijective graph homomorphism from C to H. For each $xy \in E(C)$, choose a set of vertices $Z^{xy} := \{z^{x,y}, z^{y,x}\}$ and, for each $x \in V(C)$, choose a vertex w^x. Choose the vertices so that $V(H)$, $V(C)$, Z^e, $Z^{e'}$, $\{w^x\}$ and $\{w^{x'}\}$ are disjoint for all distinct $e, e' \in E(C)$ and all distinct $x, x' \in V(C)$. Let

- $E_1 := \{xz^{x,y}, yz^{y,x} : xy \in E(C)\}$;

- $E_2 := \{z^{x,y}z^{y,x} : xy \in E(C)\}$;
- $E_3 := \{\phi(x)z^{x,y}, \phi(y)z^{y,x} : xy \in E(C)\}$;
- $E_4 := \{w^x z^{x,y} : xy \in E(C)\}$.

The transformer T has $V(T) := V(H) \cup V(C) \cup \bigcup_{e \in E(C)} Z^e \cup \bigcup_{x \in V(C)} \{w^x\}$ and $E(T) := \bigcup_{i=1}^4 E_i$. Note that $|T| \leq 5|C| = 5e(H)$. To see that T is a $(C,H)_{C_4}$-transformer, it remains to verify that both $C \cup T$ and $H \cup T$ have C_4-decompositions (the details are given in Section 8 of [7]).

5.3.3 Finding absorbers in a bipartite setting

We must also be able to find absorbers when the host graph G is bipartite.

Lemma 5.3.3. *Let $k \in \mathbb{N}$, $k \geq 2$ and $1/n \ll 1/m' \ll 1/m \ll 1/k$. Let $G = (A,B)$ be a bipartite graph with $|A|, |B| \geq n$ and let $U \subseteq V(G)$ with $|U| = m$. Suppose that for each $v \in A$, $d(v) \geq \delta_k|B| + m'$ and, for each $v \in B$, $d(v) \geq \delta_k|A| + m'$. Then G contains a C_{2k}-divisible subgraph A^* such that $|A^*| \leq 2^{m^2}$ and if H is any C_{2k}-divisible graph on U that is edge-disjoint from A^* then $A^* \cup H$ has a C_{2k}-decomposition.*

The proof is very similar to that of Lemma 5.3.2 so we omit the details and restrict ourselves to the following outline. For $k \geq 3$ we find transformers using the construction given in Section 5.3.1 and for C_4 we use the construction described in Section 5.3.2. The following observations allow us to find absorbers:

- Given a connected, 2-divisible graph H and a vertex-disjoint cycle C of length $e(H)$ on (A,B), there is a bipartition of the $(C,H)_{C_{2k}}$-transformer which respects the bipartitions of $V(H)$ and $V(C)$ (with a suitable choice of the graph homomorphism ϕ). An example for cycles of length four is given in Figure 5.2.
- For $k \geq 3$, $(C,H)_{C_{2k}}$-transformers are constructed from a collection of internally-disjoint paths of length k or $k-1$ between vertices in C and H. Any pair of vertices in

A has at least $2m'$ common neighbours in B since, for any $v \in A$, $d_G(v) \geq |B|/2+m'$. Similarly, any pair of vertices in B has at least $2m'$ common neighbours in A. So we can find the transformers greedily.

- List the vertices of the $(C,H)_{C_4}$-transformer described in Section 5.3.2 so that they appear in the following order: $V(C \cup H)$, $\bigcup_{e \in E(C)} Z^e$, $\bigcup_{x \in V(C)} \{w^x\}$. Each vertex in the transformer has at most three of its neighbours appearing before itself in this list. For any $v \in A$, $d_G(v) \geq 2|B|/3 + m'$, so any three vertices in A have at least $3m'$ common neighbours in B. The same is true with the roles of A and B reversed. So we can greedily embed the vertices of the transformer in this order.

5.4 Cycles of length four

5.4.1 Case distinction

For cycles of length four, the εn term in Theorem 5.1.1 is required only to find the absorber in the proof. We show that a minimum degree of $2n/3 - 1$ suffices by observing that any such graph either contains an absorber or has one of two extremal structures pictured in Figure 5.3 (both of which have C_4-decompositions).

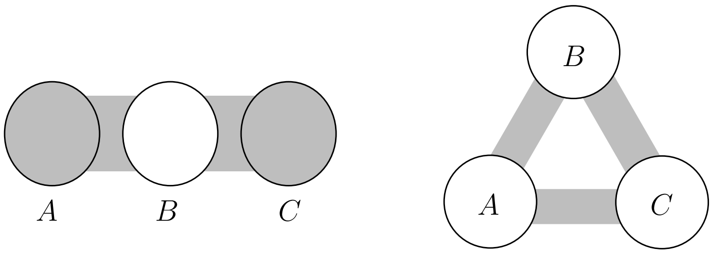

Figure 5.3: If G is extremal and $\delta(G) \geq 2n/3 - 1$, then G resembles the graph on the left (type 1) or the right (type 2). Here $|A|, |B|, |C| \sim n/3$ and shaded areas are dense.

We say that a graph G on n vertices is *m-extremal* if there exist disjoint sets $S, T \subseteq V(G)$ such that $|S|, |T| \geq n/3 - m$ which satisfy one of the following:

- $e(S,T) = 0$; (*Type 1*)

- $e(S) = e(T) = 0$. (*Type 2*)

If G is not close to being m-extremal, Lemma 5.4.1 finds a C_4-decomposition.

Lemma 5.4.1. *Let $n, m_1, m_2 \in \mathbb{N}$ with $1/n \ll 1/m_1 \ll 1/m_2 \ll 1$. Let G be a C_4-divisible graph on n vertices with $\delta(G) \geq 2n/3 - 1$. Suppose that for every spanning subgraph G' of G such that $\delta(G') \geq 2n/3 - m_2$, G' is not m_1-extremal. Then G has a C_4-decomposition.*

If Lemma 5.4.1 does not apply, then G has a subgraph G' which is m_1-extremal and has $\delta(G') \geq 2n/3 - m_2 \geq 2n/3 - m_1$. In this case, we use the following result.

Lemma 5.4.2. *Let $n, m \in \mathbb{N}$ with $1/n \ll 1/m \ll 1$. Let G be a C_4-divisible graph on n vertices with $\delta(G) \geq 2n/3 - 1$. Suppose that there exists a spanning subgraph G' of G such that $\delta(G') \geq 2n/3 - m$ and G' is m-extremal of* (i) *type 1 or* (ii) *type 2. Then G has a C_4-decomposition.*

So, together, Lemmas 5.4.1 and 5.4.2 imply Theorem 5.1.2 when $k = 2$.

5.4.2 G is not extremal

In this section we prove Lemma 5.4.1, which finds a C_4-decomposition of G whenever G is not extremal. Let G be a graph on n vertices. A (δ, μ, m)-*vortex* in G (as defined in [38]) is a sequence $U_0 \supseteq U_1 \supseteq \cdots \supseteq U_\ell$ such that:

- $U_0 = V(G)$;
- $|U_i| = \lfloor \mu |U_{i-1}| \rfloor$, for all $1 \leq i \leq \ell$, and $|U_\ell| = m$;
- $d_G(x, U_i) \geq \delta |U_i|$, for all $1 \leq i \leq \ell$ and all $x \in U_{i-1}$.

We use Lemma 4.3 from [38] to find a vortex in G.

Lemma 5.4.3 ([38])**.** *Let $0 \leq \delta \leq 1$. For all $0 < \mu < 1$, there exists an $m_0 = m_0(\mu)$ such that for all $m' \geq m_0$ the following holds. Whenever G is a graph on $n \geq m'$ vertices with $\delta(G) \geq \delta n$, then G has a $(\delta - \mu, \mu, m)$-vortex for some $\lfloor \mu m' \rfloor \leq m \leq m'$.*

The following result (taken from the more general statement for F-decompositions, Lemma 5.1 in [38]) finds an approximate C_4-decomposition of G leaving only a very small (and very restricted) leftover H.

Lemma 5.4.4 ([38])**.** *Let $1/m \ll \mu$. Let G be a C_4-divisible graph with $\delta(G) \geq (1/2 + 3\mu)|G|$ and let $U_0 \supseteq U_1 \supseteq \cdots \supseteq U_\ell$ be a $(1/2 + 4\mu, \mu, m)$-vortex in G. Then there exists $H \subseteq G[U_\ell]$ such that $G - H$ is C_4-decomposable.*

We must prove the following lemma which reserves an absorber that can be used to deal with this leftover graph H.

Lemma 5.4.5. *Let $n, m_1, m_2, m_3 \in \mathbb{N}$ with $1/n \ll 1/m_1 \ll 1/m_2 \ll 1/m_3 \ll 1$. Let G be a graph on n vertices with $\delta(G) \geq 2n/3 - m_2$. Suppose that G is not m_1-extremal. Let $U \subseteq V(G)$ with $|U| = m_3$. Then G contains a C_4-divisible subgraph A^* with $|A^*| \leq 2^{m_3^2}$ such that for any C_4-divisible graph H on U that is edge-disjoint from A^*, the graph $A^* \cup H$ has a C_4-decomposition.*

Lemma 5.4.1 follows directly from these results.

Proof of Lemma 5.4.1. (Assuming Lemma 5.4.5.) Let $m_3 \in \mathbb{N}$ and μ be such that

$$1/n \ll 1/m_1 \ll 1/m_2 \ll 1/m_3 \ll \mu \ll 1.$$

Apply Lemma 5.4.3 to G to find a $(2/3 - 2\mu, \mu, m_3)$-vortex $U_0 \supseteq U_1 \supseteq \cdots \supseteq U_\ell$ in G. Define $\ell_0 := \lceil \log_\mu(m_2/n) \rceil + 1$. We have

$$\mu^2 m_2 - 2 \leq \mu^{\ell_0} n - 2 \leq |U_{\ell_0}| \leq \mu^{\ell_0} n \leq \mu m_2.$$

Let $G' := G - G[U_{\ell_0}]$. We have $\delta(G') \geq 2n/3 - 1 - |U_{\ell_0}| \geq 2n/3 - m_2$, so G' is not m_1-extremal. Apply Lemma 5.4.5 to the graph G' with U_ℓ playing the role of U to find $A^* \subseteq G'$ as in the lemma. We have $\Delta(A^*) \leq |A^*| \leq 2^{m_3^2} \leq |U_{\ell_0}|/10$, so $U_0 \supseteq U_1 \supseteq \cdots \supseteq U_\ell$ is a $(1/2 + 4\mu, \mu, m_3)$-vortex in $G^* := G - A^*$. Then apply Lemma 5.4.4 to G^* to find

$H \subseteq G^*[U_\ell]$ such that $G^* - H$ has a C_4-decomposition. Observing that $A^* \cup H$ has a C_4-decomposition (by Lemma 5.4.5) completes the proof. □

To prove Lemma 5.4.5, we will find a C_4-absorber for each possible C_4-divisible graph on U. We will use the transformer construction which was given in Section 5.3.2 and embed the vertices of the transformer in the order: $V(H \cup C), \bigcup_{e \in E(C)} Z^e, \bigcup_{x \in V(C)} \{w^x\}$. The difficulty arises when we try to embed the vertices in $\bigcup_{e \in E(C)} Z^e$ since, unlike in [7], we can no longer guarantee that any set of three vertices will have a common neighbour.

We will say that the edge $v_x v_y$ *transforms* xy to $\phi(x)\phi(y)$ if $v_x \in N(x) \cap N(\phi(x))$ and $v_y \in N(y) \cap N(\phi(y))$. Suppose we are transforming the edge xy to $\phi(x)\phi(y)$. We are able to do this if there is an edge between $N(x) \cap N(\phi(x))$ and $N(y) \cap N(\phi(y))$. These "transforming" edges are related to the vertices in $\bigcup_{e \in E(C)} Z^e$. That is, for each $xy \in E(C)$, the edge $z^{x,y} z^{y,x}$ transforms xy to $\phi(x)\phi(y)$. This suggests that we will be able to find an absorber as long as there do not exist $X, Y \subseteq V(G)$ with $|X|, |Y| \sim n/3$ and $e(X, Y) = 0$ (note that X and Y are not necessarily disjoint, unlike in the definition of m-extremal).

Proof of Lemma 5.4.5. Let $H_1, \dots, H_p$ be an enumeration of all possible C_4-divisible graphs on U and note that $p \leq 2^{\binom{m_3}{2}}$. For each $1 \leq i \leq p$, find an edge-disjoint C_4-connector $H_i^{\text{con}} \subseteq G - G[U]$ (using the procedure given in Section 5.3). Each $H_i' := H_i \cup H_i^{\text{con}}$ is C_4-divisible and $|H_i'| \leq 3m_3$. Let $h_i := e(H_i')$, let C^i be a cycle of length h_i and let J_i be a copy of $L(h_i/4, 2)$. Our strategy is as follows. Suppose that $G \setminus \bigcup_{i=1}^p H_i'$ contains vertex-disjoint copies of C^i and J_i such that we are able to find edge-disjoint $(C^i, H_i')_{C_4}$- and $(C^i, J_i)_{C_4}$-transformers T_i and T_i'. Then we can combine these to obtain a C_4-absorber A_i for H_i as in the proof of Lemma 5.3.2 (more precisely, letting $A_i := H_i^{\text{con}} \cup T_i \cup C^i \cup T_i' \cup J_i$). We use the following claim.

Claim: *There exist vertex-disjoint copies of $C^1, \dots, C^p, J_1, \dots, J_p$ in $G \setminus \bigcup_{i=1}^p H_i'$ such that the following holds. Let $W \subseteq V(G)$ with $|W| \leq m_2$. For any $1 \leq i \leq p$, any $xy \in E(C^i)$ and any $\phi(x)\phi(y) \in E(G)$ there is an edge $v_x v_y \in E(G \setminus W)$ which transforms* *xy to $\phi(x)\phi(y)$.*

We consider two cases.

Case 1: *For all sets $X, Y \subseteq V(G)$ with $|X|, |Y| \geq n/3 - 3m_2$, $e_G(X,Y) > 0$.*

Find vertex-disjoint copies of $C^1, \ldots, C^p, J_1, \ldots, J_p$ (anywhere) in $G \setminus \bigcup_{i=1}^p H_i'$. Consider any $W \subseteq V(G)$ with $|W| \leq m_2$, any $xy, \phi(x)\phi(y) \in E(G)$. Let $X := (N_G(x) \cap N_G(\phi(x))) \setminus W$ and $Y := (N_G(y) \cap N_G(\phi(y))) \setminus W$. Note that

$$|X|, |Y| \geq 2\delta(G) - n - |W| \geq n/3 - 3m_2,$$

so $e_G(X,Y) > 0$. Any edge $v_x v_y \in E_G(X,Y)$ transforms xy to $\phi(x)\phi(y)$.

Case 2: *There exist $X, Y \subseteq V(G)$ with $|X|, |Y| \geq n/3 - 3m_2$ such that $e_G(X,Y) = 0$.*

Since G is not m_1-extremal, $X \cap Y \neq \emptyset$. Let $v \in X \cap Y$ and note that $N_G(v) \subseteq V(G) \setminus (X \cup Y)$. So $|X \cup Y| \leq n/3 + m_2$ and

$$|X \cap Y| \geq 2(n/3 - 3m_2) - (n/3 + m_2) = n/3 - 7m_2.$$

Let $X' \subseteq X \cap Y$ of size $\lfloor n/3 \rfloor - 7m_2$. Note that $e_G(X') = 0$.

Let $m := m_1/10$. For each $i \in \{m, n/3 - \sqrt{m}\}$, let $U_i := \{v : v \in V(G) \setminus X', d_G(v, X') \leq i\}$. We have

$$|X'|(2n/3 - m_2) \leq e_G(X', V(G) \setminus X') \leq |U_i| i + (n - |X'| - |U_i|)|X'|$$

which yields

$$|U_i| \leq \frac{|X'|(n/3 - |X'| + m_2)}{|X'| - i} \leq \frac{|X'|(8m_2 + 1)}{|X'| - i}.$$

Thus, we have $|U_m| \leq 9m_2$ and $|U_{n/3-\sqrt{m}}| \leq n/100$. Set $X'' := X' \cup U_m$, $Y' := V(G) \setminus X''$ and $Y'' := Y' \setminus U_{n/3-\sqrt{m}}$. Note that:

(i) for every $v \in X''$, $d_G(v, Y') \geq 2n/3 - 2m$;

(ii) for every $v \in Y'$, $d_G(v, X'') \geq m$ and $d_G(v, Y') \geq 2n/3 - m_2 - |X''|$;

(iii) for every $v \in Y''$, $d_G(v, X'') \geq n/3 - \sqrt{m}$;

(iv) $n/3 - 8m_2 \leq |X''| \leq n/3 + 2m_2$ and $2n/3 - 2m_2 \leq |Y'| \leq 2n/3 + 8m_2$.

Find vertex-disjoint copies of $C^1, \ldots, C^p, J_1, \ldots, J_p$ in $G \setminus \bigcup_{i=1}^p H_i'$ such that each cycle $C^i \subseteq G[X'', Y'']$. Consider any $W \subseteq V(G)$ with $|W| \leq m_2$, any $1 \leq i \leq p$, any $xy \in E(C^i)$ and any $\phi(x)\phi(y) \in E(G)$. We will assume, without loss of generality, that $x \in X''$ and $y \in Y''$.

Suppose first that $\phi(x), \phi(y) \in X''$. Note that (i) and (ii) imply

$$\begin{aligned}|N_G(y, Y') \cap N_G(\phi(y), Y')| &\geq (2n/3 - m_2 - |X''|) + (2n/3 - 2m) - |Y'| \\ &= n/3 - 2m - m_2 \geq n/3 - 3m.\end{aligned}$$

Choose v_y to be any vertex in $(N_G(y, Y') \cap N_G(\phi(y), Y')) \setminus W$. By (ii) and (iv), v_y has at least $2n/3 - m_2 - |X''| > n/4$ neighbours in Y'. Since

$$|N_G(x, Y') \cap N_G(\phi(x), Y')| \overset{\text{(i)}}{\geq} 2(2n/3 - 2m) - |Y'| \overset{\text{(iv)}}{\geq} |Y'| - 5m,$$

v_y has many neighbours in $(N_G(x, Y') \cap N_G(\phi(x), Y')) \setminus W$, choose any one of these for v_x.

Now suppose that $\phi(x), \phi(y) \in Y'$. It follows from (ii)–(iv) that

$$\begin{aligned}|N_G(y, X'') \cap N_G(\phi(y), X'')| &\geq (n/3 - \sqrt{m}) + m - (n/3 + 2m_2) \\ &= m - \sqrt{m} - 2m_2 \geq m/2.\end{aligned}$$

Choose any vertex from $(N_G(y, X'') \cap N_G(\phi(y), X'')) \setminus W$ for v_y. This vertex is adjacent to all but at most $3m$ vertices in Y', by (i) and (iv). Use (i) and (ii) to see that $|N_G(x, Y') \cap N_G(\phi(x), Y')| \geq n/3 - 3m$. Thus v_y must have many neighbours in $(N_G(x, Y') \cap N_G(\phi(x), Y')) \setminus W$. Choose any suitable vertex for v_x.

A similar argument deals with the case when $\phi(x) \in X''$ and $\phi(y) \in Y'$. We use that

$$|N_G(y, X'') \cap N_G(\phi(y), X'')| \geq m/2 \text{ and } |N_G(x, Y') \cap N_G(\phi(x), Y')| \geq |Y'| - 5m$$

to find suitable vertices $v_y \in (N_G(y, X'') \cap N_G(\phi(y), X'')) \setminus W$ and $v_x \in (N_G(x, Y') \cap N_G(\phi(x), Y')) \setminus W$.

Finally, suppose that $\phi(x) \in Y'$ and $\phi(y) \in X''$. We again use (i) and (ii) to see that

$$|N_G(x, Y') \cap N_G(\phi(x), Y')|, |N_G(y, Y') \cap N_G(\phi(y), Y')| \geq n/3 - 3m.$$

Let

$$Y_x := (N_G(x, Y') \cap N_G(\phi(x), Y')) \setminus W \text{ and}$$
$$Y_y := (N_G(y, Y') \cap N_G(\phi(y), Y')) \setminus W,$$

so $|Y_x|, |Y_y| \geq n/3 - 4m$. If $e_G(Y_x, Y_y) > 0$ choose any $v_x v_y \in E_G(Y_x, Y_y)$. Suppose then that $e_G(Y_x, Y_y) = 0$. Note that $Y_x \cap Y_y \neq \emptyset$, else G is m_1-extremal of type 1. So, as previously, we can let $v \in Y_x \cap Y_y$ and note that $N_G(v) \subseteq V(G) \setminus (Y_x \cup Y_y)$. So $|Y_x \cup Y_y| \leq n/3 + m_2$ and

$$|Y_x \cap Y_y| \geq 2(n/3 - 4m) - (n/3 + m_2) \geq n/3 - 9m.$$

But then G is m_1-extremal of type 2 (take $S := X'$ and $T := Y_x \cap Y_y$) which is a contradiction. This completes the proof of the claim.

We now explain how to use the claim to find, for each $1 \leq i \leq p$, a $(C^i, H_i')_{C_4}$-transformer (and $(C^i, J_i)_{C_4}$-transformers are found in exactly the same way). We will use the construction described in Section 5.3.2. Let ϕ be an edge-bijective graph homomorphism from C^i to H_i. For each edge $xy \in E(C^i)$, use the claim (with W set to be all vertices which have been used at any point previously in the construction) to find an edge which transforms $xy \in E(C^i)$ to $\phi(x)\phi(y)$ and thus obtain suitable embeddings for the vertices in $\bigcup_{e \in E(C^i)} Z^e$. It is then an easy task to greedily embed remaining vertices of the transformer (the vertices of the form w^x for some $x \in V(C^i)$), since each vertex of this type has at most two neighbours previously embedded. Continuing in this way, we find edge-disjoint absorbers A_i for each H_i such that $|A_i| \leq m_3^3$. Let $A^* := \bigcup_{i=1}^{p} A_i$ and note that $|A^*| \leq pm_3^3 \leq 2^{m_3^2}$. □

5.4.3 Type 1 extremal

In this section, we will prove Lemma 5.4.2 for graphs which are type 1 extremal. The next result takes any graph G which is type 1 extremal and partitions its vertices into sets A, B and C so that each vertex has many neighbours in two of the parts.

Proposition 5.4.6. *Let $n, m \in \mathbb{N}$ such that $1/n \ll 1/m \ll 1$. Let G be a graph on n vertices with $\delta(G) \geq 2n/3 - m$. Suppose G is m-extremal of type 1. Then there exists a partition A, B, C of $V(G)$ satisfying:*

(P1) *for all $v \in A$, $d_G(v, A), d_G(v, B) \geq 5n/18$;*

(P2) *for all $v \in C$, $d_G(v, B), d_G(v, C) \geq 5n/18$;*

(P3) *for all but at most $3m$ vertices $v \in A$, $d_G(v, A), d_G(v, B) \geq n/3 - 6m$;*

(P4) *for all but at most $3m$ vertices $v \in C$, $d_G(v, B), d_G(v, C) \geq n/3 - 6m$;*

(P5) *for all $v \in B$, $d_G(v, A), d_G(v, C) \geq n/50$;*

(P6) *for all but at most $50m$ vertices $v \in B$, $d_G(v, A), d_G(v, C) \geq 5n/18$;*

(P7) *$n/3 - 5m \leq |A|, |B|, |C| \leq n/3 + 3m$.*

Proof. Since G is m-extremal of type 1, there exist disjoint sets $A_1, C_1 \subseteq V(G)$ such that $|A_1|, |C_1| = \lceil n/3 \rceil - m$ and $e_G(A_1, C_1) = 0$. Let $B_1 := V(G) \setminus (A_1 \cup C_1)$. Since

$\delta(G) \geq 2n/3 - m$, for all $v \in A_1$, $d_G(v, A_1) \geq n/3 - 3m$ and $d_G(v, B_1) \geq n/3$. Likewise, for all $v \in C_1$, $d_G(v, C_1) \geq n/3 - 3m$ and $d_G(v, B_1) \geq n/3$.

Let B_C consist of all vertices v in B_1 such that $d_G(v, A_1) < n/50$. By considering $e_G(A_1, B_1)$, we obtain the following bound.

$$|A_1|n/3 \leq |B_C|n/50 + (n/3 + 2m - |B_C|)|A_1|$$

which gives

$$|B_C| \leq \frac{2m|A_1|}{|A_1| - n/50} \leq \frac{2m|A_1|}{9|A_1|/10} \leq 3m.$$

Similarly, defining B_A to consist of all vertices v in B_1 such that $d_G(v, C_1) < n/50$, we get $|B_A| \leq 3m$. Note that $B_A \cap B_C = \emptyset$. In exactly the same way, we can show that for all but at most

$$2 \cdot \frac{2m|A_1|}{|A_1| - 5n/18} \leq 50m$$

vertices $v \in B$, $d_G(v, A), d_G(v, C) \geq 5n/18$. Set $A := A_1 \cup B_A$, $C := C_1 \cup B_C$ and $B := B_1 \setminus (B_A \cup B_C)$. Properties (P1)–(P7) are satisfied. □

The next result refines this partition and covers all atypical edges by copies of C_4 to leave a dense graph with a well-defined structure.

Proposition 5.4.7. *Let $n, m \in \mathbb{N}$ such that $1/n \ll 1/m \ll 1$. Let G be a C_4-divisible graph on n vertices with $\delta(G) \geq 2n/3 - 1$. Suppose that there exists a spanning subgraph G' of G such that $\delta(G') \geq 2n/3 - m$ and G' is m-extremal of type 1. Then there exists $G'' \subseteq G$ and a partition A, B, C of $V(G'')$ satisfying:*

(Q1) *$e_{G''}(A)$ and $e_{G''}(C)$ are even;*

(Q2) *$G'' \subseteq G[A] \cup G[C] \cup G[B, A \cup C]$ and $G - G''$ has a C_4-decomposition;*

(Q3) *for all $v \in A$, $d_{G''}(v, A), d_{G''}(v, B) \geq n/4$;*

(Q4) *for all $v \in B$, $d_{G''}(v, A), d_{G''}(v, C) \geq n/4$;*

(Q5) *for all* $v \in C$, $d_{G''}(v, B), d_{G''}(v, C) \geq n/4$*;*

(Q6) $n/3 - 55m \leq |A|, |B|, |C| \leq n/3 + 3m$*.*

Note that we do not require G'' to be spanning.

Proof. First apply Proposition 5.4.6 to G' to find a partition A, B, C of $V(G)$ satisfying (P1)–(P7). Suppose that $e_G(A, C) + e_G(B) = 0$. It is clear that taking G'' as G with the partition A, B, C will satisfy (Q2)–(Q6). We must check (Q1). Since $N_G(x) \subseteq A \cup B$ for all $x \in A$ and so on,

$$|A| + |B| - 1, |A| + |C|, |B| + |C| - 1 \geq \delta(G) \tag{5.1}$$

which implies that $2n = 2(|A|+|B|+|C|) \geq 3\delta(G)+2$ and $\delta(G) \leq (2n-2)/3$. Note that $n \not\equiv 0 \mod 3$, otherwise $\delta(G) \geq 2n/3$ since $2n/3 - 1$ is odd and G is 2-divisible. We can show that $n \not\equiv 2 \mod 3$ either, else $\delta(G) \geq \lceil 2n/3 \rceil - 1 = (2n-1)/3$. Thus $n = 3N + 1$ for some $N \in \mathbb{N}$ and $\delta(G) = 2N$. The inequalities in (5.1) must be satisfied with equality, else $|A| + |B| + |C| > n$. Hence $|A| = |C| = N$ and $|B| = N + 1$; the graphs $G[A]$, $G[B, A \cup C]$ and $G[C]$ are complete and G is $2N$-regular. If $e_G(A) = e_G(C) = \binom{N}{2}$ is odd, it is easy to check that $N \equiv 2, 3 \mod 4$. But then $e(G) = N(3N + 1)$ is not divisible by four which contradicts G being C_4-divisible. Hence (Q1) is also satisfied.

Let us assume then that $e_G(A, C)+e_G(B) > 0$. Our first step will be to cover all edges inside B and between A and C using copies of C_4. We begin by reducing the maximum degree in $G[A, C] \cup G[B]$. Choose any edge $xy \in E_G(A, C) \cup E_G(B)$, we will protect this edge for the time being since we might need it later on. Let $G_0 := (G[A, C]\cup G[B])-\{xy\}$. Let η be chosen such that $1/m \ll \eta \ll 1$. The Erdős-Stone theorem allows us to greedily remove copies of C_4 from G_0 until at most ηn^2 edges remain. Let $\mathcal{F}_0$ denote this collection of edge-disjoint copies of C_4 and let $G_1 := G_0 - \bigcup \mathcal{F}_0$ with $e(G_1) \leq \eta n^2$.

We say that a vertex v is *bad* if $d_{G_1}(v) \geq \eta^{1/2}n$. Note that G contains at most $2\eta n^2/(\eta^{1/2}n) = 2\eta^{1/2}n$ bad vertices. Let $B' \subseteq B$ consist of all the vertices $v \in B$ such that $d_G(v, A) < 5n/18$ or $d_G(v, C) < 5n/18$. Then $|B'| \leq 50m$ by (P6). For each bad vertex v, let $S_v \subseteq N_{G_1}(v)$ be a set of vertices of maximal size such that $|S_v|$ is even, no vertex in S_v is bad and $S_v \cap B' = \emptyset$. Note that each vertex appears in at most $2\eta^{1/2}n$ sets S_v. Pair up the vertices in each S_v arbitrarily. Our aim is to find a path of length two between each pair in $G_2 := G - (G[A, C] \cup G[B] \cup \{xy\})$. In total we have to find at most $\eta n^2/2$ paths. Note that each pair in S_v has at least $n/9$ common neighbours in G_2 (for S_v where $v \in B$, it is important that $S_v \subseteq B \setminus B'$). This allows us to greedily embed the paths so that each vertex is used at most $\eta^{1/3}n/3$ times. Write $\mathcal{F}_1$ for the edge-disjoint collection of copies of C_4 formed by taking $\bigcup G[v \cup S_v]$ together with these paths. Let $G_3 := G - \bigcup(\mathcal{F}_0 \cup \mathcal{F}_1)$. We have:

(a) for all $v \in V(G_3)$, $d_{G_3}(v) \geq d_{G_2}(v) - \eta^{1/3}n$;

(b) $\Delta(G_3[B]), \Delta(G_3[A, C]) \leq \eta^{1/3}n$;

(c) $1 = |\{xy\}| \leq e_{G_3}(A, C) + e_{G_3}(B) \leq \eta n^2 + 1$.

We make the following observation

$$e_{G_3}(A, C) + e_{G_3}(B) \equiv e_{G_3}(A) + e_{G_3}(C) \mod 2. \tag{5.2}$$

To see (5.2), note that G_3 is C_4-divisible since it was obtained by removing edge-disjoint copies of C_4 from G. In particular, this means that G_3 is 2-divisible and so $e_{G_3}(A \cup C, B)$ is even. Since $e(G_3)$ is also even, the result follows.

We use (5.2) to cover all remaining edges in $E_{G_3}(A, C) \cup E_{G_3}(B)$, at the same time ensuring we leave an even number of edges behind in each of A and C. If $e_{G_3}(C)$ is odd, then assign one edge from $E_{G_3}(A, C) \cup E_{G_3}(B)$ to C (we use (c) to ensure that this edge exists) and the remainder to A. Otherwise, assign all edges from $E_{G_3}(A, C) \cup E_{G_3}(B)$ to A. Find a copy of C_4 covering each $e \in E_{G_3}(A, C) \cup E_{G_3}(B)$ of the following form (here we say that a cycle has the form $X_1X_2X_3X_4$ to indicate that the cycle visits vertices in X_1, X_2, X_3 and X_4 in this order):

- $BBXX$, if $e \in E_G(B)$ and e is assigned to $X \in \{A, C\}$;

- $CAAB$, if $e \in E_G(A, C)$ and e is assigned to A;
- $ACCB$, if $e \in E_G(A, C)$ and e is assigned to C.

We first check that it is possible to find cycles of these forms without using any vertex too often. The ordering of each cycle above is suggestive of the order in which its vertices should be embedded (for cycles of the form $BBXX$, choose the first vertex in X to satisfy (P3) or (P4) in G, i.e., not one of the exceptional $3m$ vertices). Properties (P1)–(P7) together with (a) ensure that there are at least $n/100$ suitable candidates in G_3 for each vertex which is not an endpoint of the fixed edge e. In total we must find at most $\eta n^2 + 1$ cycles and each vertex appears in the fixed edge e for at most $\eta^{1/3}n$ of these cycles, by (b) and (c). So it is possible to embed cycles of the required forms so that each vertex is used at most $2\eta^{1/3}n$ times. Let $\mathcal{F}_2$ denote the collection of cycles thus obtained and let $G_4 := G_3 - \bigcup \mathcal{F}_2$. For each $v \in V(G_4)$, we have

$$d_{G_4}(v) \geq d_{G_2}(v) - 5\eta^{1/3}n. \tag{5.3}$$

We now check that removing these cycles has the desired effect. Observe that any edge which is assigned to A forms a C_4 which uses one edge from $E_{G_3}(A)$ and no edges from $E_{G_3}(C)$. The same statement holds with A and C swapped. If $e_{G_3}(C)$ is odd, deleting the cycles in $\mathcal{F}_3$ will remove one edge from $E_{G_3}(C)$ leaving $e_{G_4}(C)$ even. If $e_{G_3}(C)$ is even, no edges were assigned to C so $e_{G_4}(C)$ remains even. To see that $e_{G_4}(A)$ will also be even, we note that (5.2) implies that the number of edges assigned to A was congruent to $e_{G_3}(A)$ mod 2.

Lastly, we cover all edges incident to vertices in B' (so that we can ignore B'). Take each vertex $v \in B'$ and pair its neighbours up arbitrarily. Find a path of length two between each pair in $G_4[A \cup C, B]$ (each such path will form a copy of C_4 which covers two edges incident at v). By (P1), (P2) and (5.3), any pair of vertices in $A \cup C$ has at least $n/10$ common neighbours in B and, in total, we are required to find at most $|B'|n/2 \leq 25mn$ paths. So we can find a collection $\mathcal{F}_3$ of edge-disjoint copies of C_4 which covers all edges incident at B' and uses each vertex in $V(G) \setminus B'$ at most ηn times. Let $B'' := B \setminus B'$ and let $G'' := (G_4 - \bigcup \mathcal{F}_3) \setminus B'$. It is easy to check that G'' with the partition A, B'', C satisfies (Q1)–(Q6). □

Proposition 5.4.7 takes us most of the way towards proving Lemma 5.4.2 for graphs of type 1. All that remains is to show that the graphs $G''[A]$, $G''[C]$, $G''[A, B]$ and $G''[B, C]$ can be made to be C_4-divisible and then to decompose these using Theorems 5.1.1 and 5.1.3.

Proof of Lemma 5.4.2(i). Apply Proposition 5.4.7 to G to find $G'' \subseteq G$ and a partition A, B, C of $V(G'')$ satisfying properties (Q1)–(Q6). We begin by making the graphs $G''[A]$ and $G''[C]$ C_4-divisible. Towards this aim, let $A' \subseteq A$ consist of all vertices $v \in A$ such that $d_{G''}(v, A)$ is odd. Clearly, $|A'|$ is even. Pair up the vertices in A' arbitrarily. For each pair a_1, a_2, find a copy of C_4 of the form a_1Aa_2B in G''. Note that on removing a copy of C_4 of this form, a_1 and a_2 will both have even degree in A and the degree of the third vertex in A is reduced by two so its parity will not be changed. Do the same for the vertices in C (finding cycles of the form c_1Cc_2B). Note that in total we must find at most $n/2$ copies of C_4. Properties (Q3), (Q5) and (Q6) imply that each pair has at least $n/10$ common neighbours in the required vertex classes, so we can avoid using any vertex more than 20 times. Write $\mathcal{F}_1$ for this collection of copies of C_4 and let $G_1 := G'' - \bigcup \mathcal{F}_1$. Now every vertex in $G_1[A]$ and $G_1[C]$ has even degree.

We also require the number of edges in $G_1[A]$ and in $G_1[C]$ to be divisible by four. We know already that the number of edges will be even (from (Q1) and the fact that $\mathcal{F}_1$ uses an even number of edges from both $G''[A]$ and $G''[C]$). Say that $e_{G_1}(A) \equiv 2 \mod 4$. We can fix this by removing a graph F consisting of three edge-disjoint copies of C_4 which take the following form: $a_1Aa_2B, a_2Aa_3B, a_1Aa_3B$ where $a_1, a_2, a_3 \in A$. Note that $F[A]$ is a copy of C_6, so removing F does not cause the degree of any vertex in $G_1[A]$ to become odd. We can remove a similar graph if $e_{G_1}(C)$ not divisible by four. We obtain a graph

G_2 such that $G_2[A]$ and $G_2[C]$ are C_4-divisible. It follows from (Q3), (Q5) and (Q6) that

$$\delta(G_2[A]), \delta(G_2[C]) \geq n/4 - 50 \geq (2/3 + 1/100)|A|, (2/3 + 1/100)|C|.$$

So we can apply Theorem 5.1.1 to find C_4-decompositions $\mathcal{F}_A$ and $\mathcal{F}_C$ of $G_2[A]$ and $G_2[C]$, respectively. Let $G_3 := G_2 - \bigcup(\mathcal{F}_A \cup \mathcal{F}_C)$.

We will now make the bipartite graphs $G_3[A, B]$ and $G_3[B, C]$ C_4-divisible. Note that for any $v \in A \cup C$, $d_{G_3}(v, B)$ is necessarily even. Let $B' \subseteq B$ consist of all vertices $v \in B$ such that $d_{G_3}(v, A)$ (and hence $d_{G_3}(v, C)$) is odd. Since $e_{G_3}(A, B)$ is even, $|B'|$ must also be even. Pair up the vertices in B' arbitrarily. For each pair b_1, b_2, find a copy of C_4 of the form Ab_1Cb_2. On removing these copies from G_3, we see that b_1 and b_2 now have even degree in A and in C. Properties (Q4) and (Q6) ensure that there are at least $n/10$ suitable candidates at each step of the embedding. Since there are fewer than $n/2$ pairs, we can choose these copies of C_4 so that no vertex is used more than 10 times. If, after removing these copies, the number of edges between A and B is not divisible by four then it must be congruent to $2 \mod 4$. We can correct this by removing three further edge-disjoint copies of C_4 of the form: b_1Ab_2C, b_2Ab_3C, b_1Ab_3C where b_1, b_2, b_3 are distinct vertices in B. Note that removing these copies of C_4 removes $6 \equiv 2 \mod 4$ edges between A and B but will not change the parity of $d(b_i, A)$ for any $i \in \{1, 2, 3\}$. Write $\mathcal{F}_2$ for the copies of C_4 removed in this step and let $G_4 := G_3 - \bigcup \mathcal{F}_2$. We now have C_4-divisible bipartite graphs $G_4[A, B]$ and $G_4[B, C]$ and $d_{G_4}(v, B) \geq n/4 - 100$ for all $v \in A \cup C$. Recall (Q6), which implies

$$\delta_{\text{bip}}(G_4[A, B]), \delta_{\text{bip}}(G_4[B, C]) \geq 2/3 + 1/100.$$

So we can use Theorem 5.1.3 to find a C_4-decomposition of G_4. Thus we have found a C_4-decomposition of G. □

5.4.4 Type 2 extremal

In this section, we prove Lemma 5.4.2 for graphs which are type 2 extremal. We begin by showing that graphs of this type closely resemble a balanced tripartite graph with high minimum degree.

Proposition 5.4.8. *Let $n, m \in \mathbb{N}$ such that $1/n \ll 1/m \ll 1$. Let G be a C_4-divisible graph on n vertices. Suppose that there exists a spanning subgraph G' of G such that $\delta(G') \geq 2n/3 - m$ and G' is m-extremal of type 2. Then there exists $G'' \subseteq G$ and a partition A, B, C of $V(G'')$ satisfying:*

(R1) *$|A|$, $|B|$ and $|C|$ are even;*

(R2) *$n/3 - 50m \leq |A|, |B|, |C| \leq n/3 + 2m$;*

(R3) *$G - G''$ has a C_4-decomposition;*

(R4) *for each $X \in \{A, B, C\}$ and each $v \in V(G'') \setminus X$, we have $d_{G''}(v, X) \geq n/4$.*

Again, G'' is not necessarily spanning.

Proof. Since G' is m-extremal of type 2, there exist disjoint sets $A_1, B_1 \subseteq V(G)$ such that $|A_1|, |B_1| = \lceil n/3 \rceil - m$ and $e_{G'}(A_1) = e_{G'}(B_1) = 0$. Let $C_1 := V(G) \setminus (A_1 \cup B_1)$. For all $v \in A_1$, $d_G(v, B_1) \geq n/3 - 3m$ and $d_G(v, C_1) \geq n/3 - 1$ since $\delta(G') \geq 2n/3 - m$. Likewise, for all $v \in B_1$, $d_G(v, A_1) \geq n/3 - 3m$ and $d_G(v, C_1) \geq n/3 - 1$.

Let $C_{1,A}$ consist of all vertices $v \in C_1$ such that $d_G(v, A_1) < 5n/18$. By considering $e_{G'}(A_1, C_1)$, we obtain the following bound.

$$|A_1|(n/3 - 1) \leq |C_{1,A}|5n/18 + (n/3 + 2m - |C_{1,A}|)|A_1|$$

which gives

$$|C_{1,A}| \leq \frac{(2m+1)|A_1|}{|A_1| - 5n/18} \leq \frac{(2m+1)|A_1|}{|A_1|/12} \leq 25m.$$

Similarly, defining $C_{1,B}$ to consist of all vertices v in C_1 such that $d_G(v, B_1) < 5n/18$, we get $|C_{1,B}| \leq 25m$. Choose at most one further vertex from each of A_1, B_1 and

$C_1 \setminus (C_{1,A} \cup C_{1,B})$ so that $|A_1|$, $|B_1|$ and $|C_1 \setminus (C_{1,A} \cup C_{1,B})|$ are made even by their removal. Let U be the set which is formed by adding these vertices to $C_{1,A} \cup C_{1,B}$. Then $|U| \leq 50m + 3$.

Since any pair of vertices in G has at least $n/4$ common neighbours, we can easily find a collection of edge-disjoint copies of C_4 which covers all edges incident at U and uses each vertex in $V(G) \setminus U$ at most m^2 times. Write $\mathcal{F}$ for this collection of copies of C_4. Let $G'' := (G - \bigcup \mathcal{F}) \setminus U$. Together with the partition $A := A_1 \setminus U$, $B := B_1 \setminus U$ and $C := C_1 \setminus U$, this graph satisfies (R1)–(R4). □

We now complete the proof of Lemma 5.4.2. The idea is to cover all atypical edges to leave behind a tripartite graph with vertex classes A, B, C and high minimum degree. A little more work produces a graph such that each pair of vertex classes induces a C_4-divisible bipartite graph which we can decompose using Theorem 5.1.3.

Proof of Lemma 5.4.2(ii). Apply Proposition 5.4.8 to find $G_1 \subseteq G$ and a partition A, B, C of $V(G_1)$ satisfying (R1)–(R4). The next step is to cover the edges in $G_1' := G_1[A] \cup G_1[B] \cup G_1[C]$ using copies of C_4. Let ε be such that $1/m \ll \varepsilon \ll 1$. Using the Erdős-Stone theorem, we may assume that $e(G_1') \leq \varepsilon n^2$ (by greedily removing copies of C_4 if necessary). Let $U \subseteq V(G_1')$ consist of all vertices v such that $d_{G_1'}(v) \geq \varepsilon^{1/2} n$. It is clear that $|U| \leq 2\varepsilon^{1/2} n$. For each $v \in U$, let $S_v \subseteq N_{G'}(v) \setminus U$ be as large as possible such that $|S_v|$ is even. For each $v \in U$, arbitrarily pair up the vertices in S_v and find edge-disjoint paths of length two in $G_1 - G_1'$ which join the pairs (to form copies of C_4 together with v). Properties (R2) and (R4) allow us to do this in such a way that each vertex is used at most $3\varepsilon^{1/2} n$ times. Denote the set of edge-disjoint copies of C_4 found in this step by $\mathcal{F}_1$. Let $G_2 := G_1 - \bigcup \mathcal{F}_1$. For each $X \in \{A, B, C\}$ and each $v \notin X$,

$$d_{G_2}(v, X) \geq n/4 - 6\varepsilon^{1/2} n \quad \text{and} \tag{5.4}$$

$$\Delta(G_2[A]), \Delta(G_2[B]), \Delta(G_2[C]) \leq \max\{\varepsilon^{1/2} n, |U| + 1\} \leq 3\varepsilon^{1/2} n. \tag{5.5}$$

Now cover each remaining edge in $G_2' := G_2[A] \cup G_2[B] \cup G_2[C]$ by a copy of C_4 using a path of length three in $G_2 - G_2'$ between its endvertices. We require at most εn^2 such paths and each vertex is an endvertex of at most $3\varepsilon^{1/2}n$ paths, by (5.5). There are at least $n/10$ possibilities to embed each vertex by (5.4) and (R2), so we are able to find these paths so that each vertex is used at most $\varepsilon^{1/3}n/3$ times. Remove these copies of C_4 and write G_3 for the resulting graph. Note that A, B, C are independent sets in G_3 and, for each $X \in \{A, B, C\}$ and each $v \notin X$,

$$d_{G_3}(v, X) \geq n/4 - \varepsilon^{1/3}n. \tag{5.6}$$

In this final step, we ensure that each pair of vertex classes induces a C_4-divisible graph. Since G_3 is 2-divisible, $e_{G_3}(A, B)$ must be even. So there is an even number of vertices $v \in A$ such that $d_{G_3}(v, B)$ is odd (note that such v will necessarily also have $d_{G_3}(v, C)$ odd since G_3 is 2-divisible). Pair these odd vertices up arbitrarily and, for each pair a_1, a_2, remove one copy of C_4 of the form a_1Ba_2C (this changes the parities of $d_{G_3}(a_1, B)$ and $d_{G_3}(a_2, B)$). Each pair has many common neighbours in B and C by (5.6), so we can do this in such a way that each vertex is used at most ten times. Do the same for the vertices in B and C to obtain a graph G_4 such that each bipartite graph induced by a pair from $\{A, B, C\}$ is C_4-divisible (that the number of edges in these graphs is divisible by four follows from 2-divisibility and (R1)). Each of these bipartite graphs has minimum degree at least $n/4 - 2\varepsilon^{1/3}n$ and (R4) implies

$$\delta_{\text{bip}}(G_4[A, B]), \delta_{\text{bip}}(G_4[A, C]), \delta_{\text{bip}}(G_4[B, C]) \geq 2/3 + \varepsilon.$$

So we may apply Theorem 5.1.3 to find C_4-decompositions of $G_4[A, B]$, $G_4[A, C]$ and $G_4[B, C]$. This completes our C_4-decomposition of G. □

5.5 Even cycles of length at least eight

The aim of this section is to prove Theorem 5.1.2 for even cycles of length at least eight. We will again split our argument into extremal and non-extremal cases. When G is not extremal, it will satisfy an expansion property which we now describe. Let G be a graph on n vertices. We define the *robust neighbourhood* of a set $S \subseteq V(G)$ to be the set of vertices $R_{\nu,G}(S) := \{v \in V(G) : d_G(v,S) \geq \nu n\}$. We say that a set $S \subseteq V(G)$ is *ν-expanding in G* if $|R_{\nu,G}(S)| \geq n/2 + \nu n$. We say that G is a *ν-expander* if for every $x \in V(G)$, $N_G(x)$ is ν-expanding. Note that every ν-expander G satisfies $\delta(G) \geq \nu n$.

Any graph which is not a ν-expander falls into one of two classes of extremal graph. We say that a graph G on n vertices is *ε-close to $K_{n/2} \cup K_{n/2}$* if there exists $S \subseteq V(G)$ such that $|S| = \lfloor n/2 \rfloor$ and $e(S, \overline{S}) \leq \varepsilon n^2$. We say that G is *ε-close to bipartite* if there exists $S \subseteq V(G)$ such that $|S| = \lfloor n/2 \rfloor$ and $e(S) \leq \varepsilon n^2$. The following is a weak form of Lemma 26 in [54].

Proposition 5.5.1 ([54])**.** *Let $1/n \ll \nu \ll \varepsilon < 1$. Let G be a graph on n vertices with $\delta(G) \geq n/2$. Then one of the following holds:*

(i) *G is a ν-expander;*

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

(iii) *G is ε-close to bipartite.*

The following result, which will be proved in Section 5.7, is a version of Theorem 5.1.1 which relies on ν-expansion (instead of solely the minimum degree). This result finds a C_{2k}-decomposition of G when G is a ν-expander and $k \geq 4$.

Theorem 5.5.2. *Let $k \in \mathbb{N}$, $k \geq 4$ and $1/n \ll \nu, 1/k$. Let G be a C_{2k}-divisible ν-expander on n vertices. If $k = 4$, assume further that $\delta(G) \geq n/2$. Then G has a C_{2k}-decomposition.*

Given Theorem 5.5.2, it remains to find decompositions of graphs which are close to $K_{n/2} \cup K_{n/2}$ or close to bipartite. This is achieved in the current section. Theorem 5.1.2

for $k \geq 4$ will then follow directly from Proposition 5.5.1, Theorem 5.5.2, Lemma 5.5.3 and Lemma 5.5.7.

5.5.1 G is close to $K_{n/2} \cup K_{n/2}$

The next result finds a C_{2k}-decomposition when G is close to $K_{n/2} \cup K_{n/2}$. The idea of the proof is to exploit the fact that G resembles two disjoint cliques: first dealing with any unusual edges or exceptional vertices and then using Theorem 5.1.1 to decompose the (almost) cliques.

Lemma 5.5.3. *Let $k \in \mathbb{N}$, $k \geq 4$ and $1/n \ll \varepsilon \ll 1$. Suppose that G is a C_{2k}-divisible graph on n-vertices and $\delta(G) \geq n/2$. Suppose further that G is ε-close to $K_{n/2} \cup K_{n/2}$. Then G has a C_{2k}-decomposition.*

We will prove Lemma 5.5.3 in stages.

Proposition 5.5.4. *Let $1/n \ll \varepsilon \ll 1$. Suppose that G is a graph on n vertices with $\delta(G) \geq n/2$ which is ε-close to $K_{n/2} \cup K_{n/2}$. Then there exists a partition A, B of $V(G)$ such that:*

(S1) $\delta(G[A]), \delta(G[B]) \geq n/5$;

(S2) *for all but at most $2\sqrt{\varepsilon}n$ vertices $v \in A$, $d_G(v, A) \geq n/2 - 2\sqrt{\varepsilon}n$;*

(S3) *for all but at most $2\sqrt{\varepsilon}n$ vertices $v \in B$, $d_G(v, B) \geq n/2 - 2\sqrt{\varepsilon}n$;*

(S4) $n/2 - 4\varepsilon n \leq |A| \leq |B| \leq n/2 + 4\varepsilon n$.

Proof. Let $S \subseteq V(G)$ such that $|S| = \lfloor n/2 \rfloor$ and $e(S, \overline{S}) \leq \varepsilon n^2$. Let $T := \overline{S}$. For each $p \in \{11n/50, n/2 - \sqrt{\varepsilon}n\}$, let $S_p := \{v \in S : d_G(v, S) \leq p\}$ and define T_p similarly. We have $|S_p|, |T_p| \leq \frac{\varepsilon n^2}{n/2-p}$, so that

$$|S_{11n/50}|, |T_{11n/50}| \leq 25\varepsilon n/7 \;\text{ and }\; |S_{n/2-\sqrt{\varepsilon}n}|, |T_{n/2-\sqrt{\varepsilon}n}| \leq \sqrt{\varepsilon}n.$$

Let $S' := (S \setminus S_{11n/50}) \cup T_{11n/50}$. Setting A to be the smallest of S' and $\overline{S'}$ and setting $B := \overline{A}$ gives the desired partition. □

Before we begin decomposing G, we must reserve some edges between A and B using the following simple proposition. These edges will be used at a later stage to ensure that the graphs on A and B are C_{2k}-divisible.

Proposition 5.5.5. *Let $k \in \mathbb{N}$ and $1/n \ll \varepsilon \ll 1/k$. Suppose that G is a graph on n vertices with $\delta(G) \geq n/2$ and A, B is a partition of $V(G)$ satisfying* (S1)–(S4). *Then there exist $4k$ distinct edges $e_1, \dots e_{2k}, f_1, \dots f_{2k} \in E_G(A, B)$ such that, for each $1 \leq i \leq 2k$, e_i and f_i are vertex-disjoint and $d_G(a_i, A) \geq n/2 - 2\sqrt{\varepsilon}n$ where $a_i := V(e_i) \cap A$.*

Proof. If $|A| < n/2$, each vertex in A has at least two neighbours in B so the result is clear. So we assume that $|A| = |B| = n/2$, in which case $\delta(G[A, B]) \geq 1$. Suppose that the proposition is false and let $\ell < 2k$ be maximal such that G contains edges $e_1, \dots e_\ell, f_1, \dots f_\ell \in E_G(A, B)$ such that, for each $1 \leq i \leq \ell$, e_i and f_i are vertex-disjoint and $d_G(a_i, A) \geq n/2 - 2\sqrt{\varepsilon}n$ where $a_i := V(e_i) \cap A$.

Let $U := \bigcup_{i=1}^{\ell} V(e_i \cup f_i)$, $A' := A \setminus U$ and $B' := B \setminus U$. Choose any vertex $a \in A'$ such that $d_G(a, A) \geq n/2 - 2\sqrt{\varepsilon}n$ and let $b \in N_G(a, B)$. Let $a' \in A' \setminus \{a\}$ and $b' \in B' \setminus \{b\}$. If $a'b'' \in E(G)$ for some $b'' \neq b$, we can take $e_{\ell+1} := ab$ and $f_{\ell+1} = a'b''$, contradicting the maximality of ℓ. Since $d_G(a', B) \geq 1$, we must have $a'b \in E(G)$. Similarly, $ab' \in E(G)$. But then taking $e_{\ell+1} := ab'$ and $f_{\ell+1} = a'b$ gives a contradiction. □

The next result covers the remaining edges between A and B.

Proposition 5.5.6. *Let $k \in \mathbb{N}$, $k \geq 4$ and $1/n \ll \eta \ll \varepsilon \ll 1/k$. Suppose that G is a graph on n vertices and A, B is a partition of $V(G)$ satisfying* (S1)–(S4). *Suppose that $e_G(A, B)$ is even. Then there exists a C_{2k}-decomposable graph $H \subseteq G$ such that $G[A, B] \subseteq H$ and $\Delta(H[A]), \Delta(H[B]) \leq \eta n$.*

Proof. Let $G' := G[A, B]$. Use the Erdős-Stone theorem to greedily find an η^4-approximate C_{2k}-decomposition $\mathcal{F}$ of G' and let $H_0 := \bigcup \mathcal{F}$. Let $X_A := \{v \in A :$

$d_{G'-H_0}(v) \geq \eta^2 n\}$ and note that $|X_A| \leq \eta^4 n^2/(\eta^2 n) = \eta^2 n$. For each vertex $x \in X_A$, pair up the vertices in $N_{G'-H_0}(x)$ arbitrarily, leaving at most one vertex unpaired. Find edge-disjoint paths of length $2k-2$ in $G[B]$ between each of the pairs (to obtain copies of C_{2k} which cover all but at most one of the edges incident at x in $G'-H_0$). Properties (S1), (S3) and (S4) allow us to find these paths so that each vertex appears as an interior vertex on at most $\eta^3 n$ of the paths. Let H_A denote the C_{2k}-decomposable graph thus obtained and repeat the process for the set of vertices $X_B := \{v \in B : d_{G'-H_0-H_B}(v) \geq \eta^2 n\}$, obtaining a C_{2k}-decomposable graph H_B which covers all but at most one edge incident at each $x \in X_B$. Now $H' := H_0 \cup H_A \cup H_B$ is C_{2k}-decomposable, $\Delta(G[A,B] - H') \leq \eta^2 n$ and

$$\Delta(H'[A]), \Delta(H'[B]) \leq 2\eta^3 n + \eta^2 n \leq 2\eta^2 n.$$

Since $e_G(A,B)$ and $e_{H'}(A,B)$ are even, so is $e_{G-H'}(A,B)$. Pair up the edges in $E_{G-H'}(A,B)$ arbitrarily and complete each to a copy of C_{2k} as follows. If the two edges share an endvertex, in A say, find a path of length $2k-2$ between their endpoints in B as above. If the edges are disjoint, find paths of length $k-1 \geq 3$ between their endpoints in A and in B. Again, properties (S1)–(S4) allow us to find these paths so that they are edge-disjoint and each vertex appears as an interior vertex on at most $\eta^3 n$ of the paths. Let H'' denote the C_{2k}-decomposable graph obtained in this way. We have ensured that

$$\Delta(H''[A]), \Delta(H''[B]) \leq \Delta(G[A,B] - H') + 2\eta^3 n \leq 2\eta^2 n.$$

Finally, let $H := H' \cup H''$. This graph is C_{2k}-decomposable,

$$\Delta(H[A]), \Delta(H[B]) \leq 4\eta^2 n \leq \eta n$$

and $G[A,B] \subseteq H$. □

We combine the previous results to find a C_{2k}-decomposition when G is ε-close to $K_{n/2} \cup K_{n/2}$.

Proof of Lemma 5.5.3. Choose a constant η such that $1/n \ll \eta \ll \varepsilon$. Apply Proposition 5.5.4 to obtain a partition A, B of G satisfying (S1)–(S4). Then apply Proposition 5.5.5 to reserve edges $\mathcal{E} := \{e_1, \dots, e_{2k}, f_1, \dots, f_{2k}\}$. Let $G' := G - \bigcup \mathcal{E}$ and note that G' with the partition A, B still satisfies (S1)–(S4) of Proposition 5.5.4. Since G is 2-divisible, $e_G(A, B)$ is even, so $e_{G'}(A, B) = e_G(A, B) - 4k$ is also even. So we can apply Proposition 5.5.6 to find $H \subseteq G'$ which has a C_{2k}-decomposition $\mathcal{F}_1$ such that $G'[A, B] \subseteq H$ and $\Delta(H[A]), \Delta(H[B]) \leq \eta n$.

We have covered all edges in $G[A, B]$ apart from those in $\mathcal{E}$, which we will use to ensure that $(G - H)[A]$ and $(G - H)[B]$ are C_{2k}-divisible. To this end, let $0 \leq r \leq 2k - 1$ be chosen such that $e_{G-H}(A) \equiv r \mod 2k$. We will find $2k$ copies of C_{2k}, each containing a pair e_i, f_i, as follows. For each $1 \leq i \leq 2k - r$, find a path of length 2 between the endpoints of e_i and f_i in $(G - H)[A]$ and a path of length $2k - 4$ between the endpoints of e_i and f_i in $(G - H)[B]$. For each $2k - r < i \leq 2k$, find a path of length 3 between the endpoints of e_i and f_i in $(G - H)[A]$ and a path of length $2k - 5$ between the endpoints of e_i and f_i in $(G - H)[B]$. (The property $d_G(a_i, A) \geq n/2 - 2\sqrt{\varepsilon} n$ where $a_i := e_i \cap A$ is needed for finding the paths of length 2.) We can choose these paths to be edge-disjoint. Let $\mathcal{F}_2$ denote the copies of C_{2k} thus obtained and let $G'' := G - H - \bigcup \mathcal{F}_2$. We make the following important observation: $G''[A]$ and $G''[B]$ are C_{2k}-divisible. That these graphs are 2-divisible is clear (they were obtained by removing edge-disjoint copies of C_{2k} from a 2-divisible graph G). To see that $e_{G''}(A)$ is divisible by $2k$, note that

$$e_{G''}(A) = e_{G-H}(A) - 2(2k - r) - 3r \equiv r - 4k + 2r - 3r \equiv 0 \mod 2k$$

(and $e_{G''}(B)$ is also divisible by $2k$).

Finally, note that

$$\Delta((G - G'')[A]), \Delta((G - G'')[B]) \leq 2\eta n$$

and recall (S1)–(S4). Let $X := \{x : d_{G''}(x) < n/2 - 3\sqrt{\varepsilon} n\}$. Then $|X| \leq 4\sqrt{\varepsilon} n$ and we can easily cover all edges incident at vertices in X using a collection $\mathcal{F}_3$ of edge-disjoint copies of C_{2k} such that no vertex in $V(G'') \setminus X$ is used more than $\varepsilon^{1/3}n$ times. Let $G''' := (G'' - \bigcup \mathcal{F}_3) \setminus X$. Now

$$\delta(G''') \geq n/2 - 3\varepsilon^{1/3}n \geq (2/3 + \varepsilon) \cdot \max\{|A \setminus X|, |B \setminus X|\}.$$

We then find C_{2k}-decompositions $\mathcal{F}_4$ and $\mathcal{F}_5$ of $G'''[A]$ and $G'''[B]$ respectively, using Theorem 5.1.1. Then $\bigcup_{i=1}^{5} \mathcal{F}_i$ gives a C_{2k}-decomposition of G. □

5.5.2 G is close to bipartite

We now consider the case when G is close to bipartite. We will process the graph, covering any unusual edges or exceptional vertices with copies of C_{2k} until we really are left with a dense bipartite graph. This we can decompose using Theorem 5.1.3.

Lemma 5.5.7. *Let $k \in \mathbb{N}$, $k \geq 4$ and $1/n \ll \varepsilon \ll 1$. Suppose that G is a C_{2k}-divisible graph on n-vertices and $\delta(G) \geq n/2$. Suppose further that G is ε-close to bipartite. Then G has a C_{2k}-decomposition.*

The following proposition partitions the vertices of G into an "almost bipartite" graph with high minimum degree.

Proposition 5.5.8. *Let $k \in \mathbb{N}$, $k \geq 4$ and $1/n \ll \varepsilon \ll 1$. Suppose that G is a C_{2k}-divisible graph on n-vertices and $\delta(G) \geq n/2$. Suppose further that G is ε-close to bipartite. Then there exists $G' \subseteq G$ and a partition A, B of $V(G')$ such that the following hold:*

(T1) $\delta(G'[A, B]) \geq n/3$;

(T2) $G - G'$ *has a C_{2k}-decomposition;*

(T3) $|A|, |B| = n/2 \pm 6\sqrt{\varepsilon}n$;

(T4) $e(G'[A]) + e(G'[B]) < \varepsilon n^2$.

Note that G' is not necessarily spanning.

Proof. Let $S \subseteq V(G)$ be such that $|S| = \lfloor n/2 \rfloor$ and $e_G(S) \leq \varepsilon n^2$. Let $T := \overline{S}$ and consider the bipartite graph $G_0 := G[S,T]$. We want to transform G_0 into a bipartite graph whose minimum degree is as high as possible. We first modify the bipartition S, T to obtain a new bipartition S', T'. Let

$$X := \{x : d_{G_0}(x) < n/2 - \sqrt{\varepsilon} n\}.$$

It is easy to see that $|X| \leq 5\sqrt{\varepsilon} n$. Let

$$X_S := \{x \in X : d_G(x, S) < 5n/12\} \text{ and let } X_T := X \setminus X_S.$$

Let $S' := (S \setminus X) \cup X_S$ and let $T' := \overline{S'} = (T \setminus X) \cup X_T$. It is useful to note that:

(i) for any $x \in S'$, $d_G(x, T') \geq n/13$ and, if $x \in S' \setminus X$, $d_G(x, T') \geq n/2 - 6\sqrt{\varepsilon} n$;

(ii) for any $x \in T'$, $d_G(x, S') \geq 5n/13$ and, if $x \in T' \setminus X$, $d_G(x, S') \geq n/2 - 6\sqrt{\varepsilon} n$.

Let $X_0 := \{x \in X : d_G(x, S) \text{ and } d_G(x, T) < 5n/12\}$. Since $X_0 \subseteq X_S$, the vertices in X_0 have all been assigned to S' but they do not naturally belong to either side of the partition so we will cover all edges incident at these vertices in the next step.

Choose any vertex $x \in X_0$. Suppose that $d_G(x, S')$ is odd. Note that $e_G(S', T')$ is even (since G is 2-divisible). This means that $e_G(S') + e_G(T')$ is also even (recall that the number of edges in G is divisible by $2k$). In particular, there must be an edge $uv \in E(S') \cup E(T')$ which is not incident at x. Let $y \in N_G(x, S') \setminus (X \cup \{u, v\})$. We now find a copy of C_{2k} which uses both xy and uv. If $u, v \in S'$, note that $|N_G(y) \cap N_G(u)| \geq n/15$ by (i), so we can find a path of length two from u to y. We also find a path of length $2k-4 \geq 4$ between x and v. At each stage, we can choose from at least $n/20$ vertices. This gives a copy of C_{2k}. We proceed in a similar way when $u, v \in T'$. We may now assume that $d_G(x, S')$ is even. Pair up the neighbours of x arbitrarily and find edge-disjoint paths of length $2k-2$ between each pair in $G[S', T']$ (to obtain edge-disjoint copies of

C_{2k}). Remove all copies of C_{2k} obtained in this way from G. Repeat this process for the remaining vertices in X_0 and write $\mathcal{F}_1$ for the collection of copies of C_{2k} thus obtained. Let $G_1 := G - \bigcup \mathcal{F}_1$. At the end of this process, we may assume that each vertex in $V(G) \setminus X_0$ has been used in at most $\varepsilon^{1/3}n/2$ copies of C_{2k}.

Let $A := S' \setminus X_0$, $B := T'$. Observe that $|A|, |B| = n/2 \pm 6\sqrt{\varepsilon}n$ and

$$\delta(G_1[A,B]) \geq 5n/13 - \varepsilon^{1/3}n \geq n/3.$$

Using the Erdős-Stone theorem, we greedily find an ε-approximate C_{2k}-decomposition $\mathcal{F}_2$ of $G_1[A] \cup G_1[B]$. Letting $G' := G - \bigcup(\mathcal{F}_1 \cup \mathcal{F}_2)$ completes the proof. □

We use this proposition to prove Lemma 5.5.7.

Proof of Lemma 5.5.7. Apply Proposition 5.5.8 to find G', A, B satisfying (T1)–(T4). Let $\mathcal{F}_1$ be a C_{2k}-decomposition of $G - G'$. Let $A' := \{x \in A : d_{G'}(x, A) \geq \sqrt{\varepsilon}n\}$ and define B' similarly. Note that $|A' \cup B'| \leq 2\sqrt{\varepsilon}n$ by (T4). Take each vertex $x \in A'$ in turn and split $N_{G'}(x, A)$ into pairs (leaving at most one vertex). Use (T1) and (T3) to find a path of length $2k-2$ between each pair in $G'[A, B]$ to obtain a copy of C_{2k} together with x. Carry out this process for the remaining edges at each remaining vertex in A'. Do the same for the vertices in B'. We may carry out this process so that each vertex appears in at most $\varepsilon^{1/3}n$ of the paths. Write $\mathcal{F}_2$ for the collection of copies of C_{2k} obtained in this way and let $G_1 := G' - \bigcup \mathcal{F}_2$. We have $\Delta(G_1[A]), \Delta(G_1[B]) < \varepsilon^{1/3}n$ and

$$\delta(G_1[A, B]) \geq n/3 - 2\varepsilon^{1/3}n. \tag{5.7}$$

We now cover the remaining edges in $E_{G_1}(A) \cup E_{G_1}(B)$. There are an even number of these so we can pair them up arbitrarily. We use (5.7) to find paths of even length at least two between any two vertices in the same class and paths of odd length at least three between any two vertices in different classes. At each step we have a choice of at least $n/10$ vertices so we are able to find edge-disjoint copies of C_{2k} (by finding paths of suitable length between the endpoints of each pair of edges) so that each pair of edges is covered and no vertex appears in more than $2\varepsilon^{1/3}n$ of the cycles. Write $\mathcal{F}_3$ for the collection of copies of C_{2k} obtained in this step. The graph $G_2 := G_1 - \bigcup \mathcal{F}_3$ is C_{2k}-divisible and bipartite with vertex classes A and B of size $n/2 \pm 6\sqrt{\varepsilon}n$. Furthermore, $\delta(G_2) \geq n/3 - 6\varepsilon^{1/3}n$ so $\delta_{\text{bip}}(G_2) \geq 1/2 + \varepsilon$. Thus G_2 has a C_{2k}-decomposition $\mathcal{F}_4$ by Theorem 5.1.3. Together, $\bigcup_{i=1}^{4} \mathcal{F}_i$ gives a C_{2k}-decomposition of G. □

5.6 Decompositions of bipartite graphs

In this section we prove Theorem 5.1.3, the bipartite version of Theorem 5.1.1. Theorem 5.1.3 finds a C_{2k}-decomposition of G when G is bipartite and has high minimum degree. We used this result to prove Theorem 5.1.2 earlier on. The proof closely follows the iterative absorption argument of [38], thus we omit some of the details.

We require the following definition, a bipartite version of the vortices considered in Section 5.4. Let $G = (A, B)$ be a bipartite graph. A (δ, μ, m)*-vortex respecting* (A, B) in G is a sequence $U_0 \supseteq U_1 \supseteq \cdots \supseteq U_\ell$ such that

- $U_0 = V(G)$;
- $|U_i \cap X| = \lfloor \mu |U_{i-1} \cap X| \rfloor$ for all $1 \leq i \leq \ell$ and each $X \in \{A, B\}$, and $|U_\ell| = m$;
- $d_G(x, U_i \cap X) \geq \delta |U_i \cap X|$, for all $1 \leq i \leq \ell$, each $X \in \{A, B\}$ and all $x \in U_{i-1} \setminus X$.

The following observation guarantees a vortex in G. It is proved by repeatedly applying the Chernoff bound given by Lemma 4.2.1 (for more details, see Appendix B.1).

Lemma 5.6.1. *Let* $0 \leq \delta \leq 1$ *and* $1/m' \ll \mu < 1$. *Suppose that* $G = (A, B)$ *is a bipartite graph with* $m' \leq |A| \leq |B| \leq 2|A|$ *and* $\delta_{\text{bip}} \geq \delta$. *Then* G *has a* $(\delta - \mu, \mu, m)$*-vortex respecting* (A, B) *for some* $2\lfloor \mu m' \rfloor \leq m \leq m'$.

The idea is to use the following result to cover almost all of the edges in G leaving only a small (very restricted) remainder which can be dealt with using the absorbers given by Lemma 5.3.3.

Lemma 5.6.2. *Let $k \in \mathbb{N}$, $k \geq 2$ and let $1/m \ll \mu \ll 1/k$. Let $G = (A, B)$ be a bipartite 2-divisible graph with $n \leq |A|, |B| \leq 2n$ and $\delta_{\text{bip}} \geq 1/2 + 3\mu$. Let $U_0 \supseteq U_1 \supseteq \cdots \supseteq U_\ell$ be a $(1/2 + 4\mu, \mu, m)$-vortex respecting (A, B) in G. Then there exists $H_\ell \subseteq G[U_\ell]$ such that $G - H_\ell$ is C_{2k}-decomposable.*

We will prove Lemma 5.6.2 in Section 5.6.1. Theorem 5.1.3 then follows directly from these results.

Proof of Theorem 5.1.3. (Assuming Lemma 5.6.2.) Let $m, n_0 \in \mathbb{N}$ and μ be such that

$$1/n_0 \ll 1/m \ll \mu \ll \varepsilon, 1/k.$$

Apply Lemma 5.6.1 to find $U_0 \supseteq U_1 \supseteq \cdots \supseteq U_\ell$, a $(\delta_k + \varepsilon/2, \mu, m)$-vortex respecting (A, B) in G.

Let $G_1 := G - G[U_1]$. We have $\delta_{\text{bip}}(G_1) \geq \delta_k + \varepsilon/2$. Apply Lemma 5.3.3 to G_1 with U_ℓ playing the role of U to find an absorber $A^* \subseteq G_1$ as in the lemma. We have $\Delta(A^*) \leq |A^*| \leq 2^{m^2}$, so $U_0 \supseteq U_1 \supseteq \cdots \supseteq U_\ell$ is a $(\delta_k + 4\mu, \mu, m)$-vortex respecting (A, B) in $G^* := G - A^*$ and $\delta_{\text{bip}}(G^*) \geq 1/2 + 3\mu$. Then apply Lemma 5.6.2 to G^* to find $H_\ell \subseteq G^*[U_\ell]$ such that $G^* - H_\ell$ has a C_{2k}-decomposition. Observing that $A^* \cup H_\ell$ has a C_{2k}-decomposition (by Lemma 5.3.3) completes the proof. $\square$

5.6.1 Proving Lemma 5.6.2

First, we state some useful results. We require the following simple proposition on decompositions of cliques. It is a special case of Wilson's Theorem and is proved very easily (see [38], for example).

Proposition 5.6.3. *Let p be prime. Then for every $k \in \mathbb{N}$, K_{p^k} has a K_p-decomposition.*

We use the next result to find an approximate C_{2k}-decomposition of G and maintain some control over the number of edges incident at any vertex in a given set X.

Lemma 5.6.4. *Let $k \in \mathbb{N}$, $k \geq 2$ and $1/n \ll \eta \ll \varepsilon, 1/k$. Suppose that $G = (A, B)$ is a bipartite graph with $n \leq |A|, |B| \leq 4n$ and $\delta_{\text{bip}}(G) \geq 1/2 + \varepsilon$. Let $X \subseteq V(G)$ of size at most $\eta^{1/2}n$. Then there exists $H \subseteq G$ such that $G - H$ is C_{2k}-decomposable, $Y := \{x \in V(G) : d_H(x) > \eta n\}$ has size at most ηn and $X \cap Y = \emptyset$.*

Proof. The first step is to cover the edges in $G[X]$ by edge-disjoint copies of C_{2k}. That is, for each edge $xy \in E_G(X)$, find a path of length $2k-1$ between the x and y in $G - G[X]$ (x and y lie in different vertex classes so any path between them is necessarily odd). In total we must find at most ηn^2 paths. Since $\delta_{\text{bip}}(G - G[X]) \geq 1/2 + 3\varepsilon/4$, we may choose these paths to be edge-disjoint and use each vertex at most $\eta^{1/3}n$ times. These paths, together with $E_G(X)$ give an edge-disjoint collection $\mathcal{F}_X$ of copies of C_{2k} with $\Delta(\bigcup \mathcal{F}_X) \leq 2\eta^{1/3}n$.

Consider the graph $G' := G \setminus \bigcup \mathcal{F}_X$. Our next step is to cover all but at most one of the remaining edges incident at each vertex in X. For each $x \in X$, pair up the vertices in $N_{G'}(x)$, leaving at most one vertex. Note that both vertices in any pair lie in the same vertex class. Since $\delta_{\text{bip}}(G' \setminus X) \geq 1/2 + \varepsilon/2$, we can find edge-disjoint paths of (even) length $2k-2$ between each pair in $G' \setminus X$. Each path combines with two edges incident at X to form a copy of C_{2k}. Thus we obtain a collection $\mathcal{F}'_X$ of edge-disjoint copies of C_{2k} which, together with $\mathcal{F}_X$, cover all but at most one edge incident at each $x \in X$.

Let $H' := G - \bigcup(\mathcal{F}_X \cup \mathcal{F}'_X)$ and note that $d_{H'}(x) \leq 1$ for all $x \in X$. Use the Erdős-Stone theorem to greedily find an η^3-approximate C_{2k}-decomposition of H' which we will denote by $\mathcal{F}$. Let $H := H' - \bigcup \mathcal{F}$ and note that $G - H$ has a C_{2k}-decomposition given by $\mathcal{F}_X \cup \mathcal{F}'_X \cup \mathcal{F}$. If $Y := \{x \in V(G) : d_H(x) > \eta n\}$, then $|Y| \leq 2e(H)/(\eta n) \leq \eta n$ and $X \cap Y = \emptyset$. □

We use Lemma 5.6.4 to prove the following result which finds a C_{2k}-decomposition of G so that every vertex has low degree in the remainder.

Lemma 5.6.5. *Let $k \in \mathbb{N}$, $k \geq 2$ and $1/n \ll \varepsilon, 1/k$. Let $G = (A, B)$ be a bipartite graph with $n \leq |A|, |B| \leq 3n$ and $\delta_{\text{bip}}(G) \geq 1/2 + \varepsilon$. Then G has an approximate C_{2k}-decomposition $\mathcal{F}$ such that $\Delta(G - \bigcup \mathcal{F}) \leq \varepsilon n$.*

Proof. Choose $s, t \in \mathbb{N}$ and $\eta > 0$ such that

$$1/n \ll \eta \ll 1/s \ll 1/t \ll \varepsilon, 1/k$$

and K_s has a K_t-decomposition (s and t exist by Proposition 5.6.3). Let $\mathcal{P} = \{V_1, \dots, V_s\}$ be a partition of $V(G)$ satisfying the following for all $1 \leq i \leq s$ and each $X \in \{A, B\}$:

(i) $|V_i \cap X| = \lfloor |X|/s \rfloor$ or $\lceil |X|/s \rceil$;

(ii) $d_G(x, V_i \cap X) \geq (1/2 + 2\varepsilon/3)|V_i \cap X|$ for all $x \in V(G) \setminus X$.

To see $\mathcal{P}$ exists, consider random equitable partitions $V_1^A, \dots, V_s^A$ of A and $V_1^B, \dots, V_s^B$ of B and let $V_i := V_i^A \cup V_i^B$. Lemma 4.2.1 implies that this partition satisfies (ii) with probability at least 3/4.

Since $|V_i| \leq \varepsilon n/2$ for all $1 \leq i \leq s$, it suffices to show that $G[\mathcal{P}]$ has an approximate C_{2k}-decomposition $\mathcal{F}$ such that $\Delta(G[\mathcal{P}] - \bigcup \mathcal{F}) \leq \varepsilon n/2$. Let $\{T_1, \dots, T_\ell\}$ be a K_t-decomposition of K_s, where $V(K_s) = \{1, \dots, s\}$. For each $1 \leq i \leq \ell$, define $G_i := \bigcup_{jk \in E(T_i)} G[V_j, V_k]$, so the G_i decompose $G[\mathcal{P}]$. For each $1 \leq i \leq \ell$, each $X \in \{A, B\}$ and all $x \in V(G_i) \setminus X$, we have

$$d_{G_i}(x) \geq (t-1)(1/2 + 2\varepsilon/3)\lfloor |X|/s \rfloor \geq (1/2 + \varepsilon/2)t\lceil |X|/s \rceil \geq (1/2 + \varepsilon/2)|V(G_i) \cap X|,$$

by (i) and (ii). So $\delta_{\text{bip}}(G_i) \geq 1/2 + \varepsilon/2$. We also note that

$$n' := t\lfloor n/s \rfloor \leq |V(G_i) \cap A|, |V(G_i) \cap B| \leq t\lceil 3n/s \rceil \leq 4n'.$$

Let $X_1 := \emptyset$. For each $1 \leq i \leq \ell$ in turn, apply Lemma 5.6.4 (with G_i, $\varepsilon/2$ and $X_i \cap V(G_i)$ playing the roles of G, ε and X) to find $H_i \subseteq G_i$ such that $G_i - H_i$ is C_{2k}-

decomposable, $d_{H_i}(x) \leq \eta n'$ for all $x \in X_i$ and $|Y_i| \leq \eta n'$, where $Y_i := \{x \in V(G_i) : d_{H_i}(x) > \eta n'\}$. Let $X_{i+1} := X_i \cup Y_i$. Note that, for all $1 \leq i \leq \ell$, $|X_i| \leq s^2\eta n' \leq \eta^{1/2}n'$ so we can indeed use Lemma 5.6.4. Let $H := \bigcup_{i=1}^{\ell} H_i$ and consider any $x \in V(G)$. We know that

$$d_H(x) \leq \ell\eta n' + 4n' \leq (s^2\eta + 4)tn/s \leq \varepsilon n/2,$$

since $d_{H_i}(x) \leq \eta n'$ for all but at most one $1 \leq i \leq \ell$. □

The following proposition takes a subset R of $V(G)$ and covers all the edges in a sparse subgraph H of $G[\overline{R}]$ using copies of C_{2k} without using any vertex too many times. It is an analogue of Proposition 5.10 in [38] and the proof is identical, so we omit the details.

Proposition 5.6.6. *Let $k \in \mathbb{N}$, $k \geq 2$ and $1/n \ll \gamma \ll \mu, 1/k$. Let $G = (A, B)$ be a bipartite graph with $n \leq |A|, |B| \leq 5n$. Let $V(G) = L \mathbin{\dot\cup} R$ such that $|R \cap X| \geq \mu n$ and $d_G(x, R \cap X) \geq (1/2 + \mu)|R \cap X|$ for each $X \in \{A, B\}$ and all $x \in V(G) \setminus X$. Let H be any subgraph of $G[L]$ such that $\Delta(H) \leq \gamma n$. Then there exists $J \subseteq G$ such that $J[L]$ is empty, $J \cup H$ is C_{2k}-decomposable and $\Delta(J) \leq \mu^2 n$.*

We now use each of the results obtained so far to prove Lemma 5.6.7. This lemma forms the basis of the induction proof of Lemma 5.6.2.

Lemma 5.6.7. *Let $k \in \mathbb{N}$, $k \geq 2$ and $1/n \ll \mu \ll 1/k$. Let $G = (A, B)$ be a bipartite graph with $n \leq |A|, |B| \leq 3n$. Let $U \subseteq V(G)$ with $|U \cap A| = \lfloor \mu|A| \rfloor$ and $|U \cap B| = \lfloor \mu|B| \rfloor$. Suppose $\delta_{\mathrm{bip}}(G) \geq 1/2 + 2\mu$ and $d_G(x, U \cap X) \geq (1/2 + \mu)|U \cap X|$ for each $X \in \{A, B\}$ and all $x \in V(G) \setminus X$. Then, if $2 \mid d_G(x)$ for all $x \in V(G) \setminus U$, there exists a collection $\mathcal{F}$ of edge-disjoint copies of C_{2k} such that every edge in $G - G[U]$ is covered and $\Delta(\bigcup \mathcal{F}[U]) \leq \mu^3|U|$.*

Proof. Choose constants γ, ξ such that $1/n \ll \gamma \ll \xi \ll \mu \ll 1/k$. Let $W := V(G) \setminus U$, $m := \lceil \xi^{-1} \rceil$ and $M := \binom{m+1}{2}$. Let $V_1, \dots, V_M$ be a partition of U such that for all $1 \leq i \leq M$, each $X \in \{A, B\}$ and all $x \in V(G) \setminus X$:

1. $d_G(x, V_i \cap X) \geq (1/2 + \mu/2)|V_i \cap X|$;

2. $|V_i \cap X| = \lfloor |U \cap X|/M \rfloor$ or $\lceil |U \cap X|/M \rceil$.

To see that such a partition exists, consider random equipartitions $V_1^A, \dots, V_M^A$ of $U \cap A$ and $V_1^B, \dots, V_M^B$ of $U \cap B$. Let $V_i := V_i^A \cap V_i^B$. Lemma 4.2.1 implies that this partition satisfies (1) with probability at least 3/4.

Let $W_1, \dots, W_m$ be a partition of W such that $W_1 \cap A, \dots, W_m \cap A$ and $W_1 \cap B, \dots, W_m \cap B$ are equipartitions of $W \cap A$ and $W \cap B$ respectively. Let $G_W^1, \dots, G_W^M$ be an enumeration of the M graphs of the form $G[W_i]$ or $G[W_i, W_j]$. Note $G[W] = \bigcup_{i=1}^M G_W^i$ and, for all $1 \le i \le M$,

$$|V(G_W^i) \cap A|, |V(G_W^i) \cap B| \le 2(3n/m + 1) \le 7\xi n \tag{5.8}$$

For each $1 \le i \le M$, let $R_i := G[V_i, V(G_W^i)]$. Let $R := \bigcup_{i=1}^M R_i$. For each $v \in V_i$ we see that $d_R(v) \le 7\xi n$ by (5.8) and for each $v \in W$, we have $d_R(v) \le m((3n\mu/M) + 1) \le 7\xi n$. Thus $\Delta(R) \le 7\xi n$.

Let $G' := G - (G[U] \cup R)$. Since $|U \cap A| = \lfloor \mu |A| \rfloor$, $|U \cap B| = \lfloor \mu |B| \rfloor$ and $\Delta(R) \le 7\xi n$, we note that $\delta_{\text{bip}}(G') \ge 1/2 + \mu/2$. So, by Lemma 5.6.5 (with γ playing the role of ε), G' has an approximate C_{2k}-decomposition $\mathcal{F}_1$ such that $H := G' - \bigcup \mathcal{F}_1$ satisfies $\Delta(H) \le \gamma n$.

We now use R and Proposition 5.6.6 to cover the edges in $H[W]$. For each $1 \le i \le M$, let $H_i := H[W] \cap G_W^i$ (so $H[W] = \bigcup H_i$) and $G_i := G[V_i] \cup R_i \cup H_i$. Observe that G_i is a bipartite graph and $V(G_i) = V_i \cup V(G_W^i)$. Let us check that G_i satisfies the conditions of Proposition 5.6.6 (with G_i, $\sqrt{\gamma}$, ξ^2 and V_i playing the roles of G, γ, μ and R). Let $n_i := \min\{|V(G_i) \cap A|, |V(G_i) \cap B|\}$, then

$$n_i \le |V(G_i) \cap A|, |V(G_i) \cap B| \le 4n_i.$$

Note that

$$n_i \le |V(G_i) \cap A| = |V_i \cap A| + |V(G_W^i) \cap A| \overset{(5.8)}{\le} 3\mu n/M + 7\xi n \le 8\xi n$$

which gives $n \geq n_i/8\xi$. We use this to see that

$$|V_i \cap A|, |V_i \cap B| \geq \mu n/2M \geq \mu\xi^2 n/2 \geq \xi^2 n_i.$$

Also $\Delta(H_i) \leq \gamma n \leq \sqrt{\gamma} n_i$ and (1) implies that $d_{G_i}(x, V_i \cap X) \geq (1/2+\xi^2)|V_i \cap X|$ for each $X \in \{A, B\}$ and all $x \in V(G_i) \setminus X$. So we may apply Proposition 5.6.6 to find $J_i \subseteq G_i$ such that $J_i[V(G_i) \setminus V_i]$ is empty, $J_i \cup H_i$ is C_{2k}-decomposable and $\Delta(J_i) \leq \xi^4 n_i$. Let $J := \bigcup_{i=1}^{M} J_i$. Then $J \cup H[W]$ has a C_{2k}-decomposition $\mathcal{F}_2$ and $\Delta(J) \leq \xi n$.

We must now cover the remaining edges in $H[U, W] \cup R$. Let $G'' := G - \bigcup(\mathcal{F}_1 \cup \mathcal{F}_2)$. Note that $G''[W]$ is empty and

$$\Delta(G'') \leq \Delta(H) + \Delta(R) \leq \gamma n + 7\xi n \leq 8\xi n.$$

Since $\Delta(J) \leq \xi n$, $\delta_{\rm bip}(G''[U]) \geq 1/2+\mu/2$. For each $w \in W$, $d_{G''}(w)$ is even, so we can pair up the vertices in $N_{G''}(w)$ arbitrarily and let P denote the list of pairs of all neighbours of W. Each vertex in U appears in at most $\Delta(G'') \leq \sqrt{\xi}|U|$ of the pairs in P and $|P| \leq \Delta(G'')3n \leq \sqrt{\xi}|U|^2$. The vertices in each pair lie in the same vertex class so we can find paths of (even) length $2k-2$ between each pair so that these paths are edge-disjoint and no vertex is used more than $\mu^3|U|/4$ times. We obtain a collection $\mathcal{F}_3$ of edge-disjoint copies of C_{2k} which cover the edges of $G'' - G''[W]$ such that $\Delta(\bigcup \mathcal{F}_3) \leq \mu^3|U|/2$. Let $\mathcal{F} := \mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3$. Then

$$\Delta\Big(\bigcup \mathcal{F}[U]\Big) \leq \Delta(J) + \Delta\Big(\bigcup \mathcal{F}_3\Big) \leq \mu^3|U|$$

and $\mathcal{F}$ covers every edge of $G - G[U]$. □

Finally, we use Lemma 5.6.7 and induction to prove Lemma 5.6.2.

Proof of Lemma 5.6.2. If $\ell = 0$, we can set $H_\ell := G$, so we assume $\ell \geq 1$. We begin by observing that for any $0 \leq i \leq \ell$, we have $2\mu^i n/3 \leq \mu^i n - 1/(1-\mu) \leq |U_i \cap A|, |U_i \cap B| \leq 2\mu^i n$. The lemma will follow from the following statement which we will prove by induction on ℓ.

> *Let $G = (A, B)$ be a 2-divisible bipartite graph with $\delta_{\text{bip}} \geq 1/2 + 3\mu$ and $|A| \leq |B| \leq 3|A|$. Let $U_1 \subseteq V(G)$ with $|U_1 \cap A| = \lfloor \mu|A| \rfloor$ and $|U_1 \cap B| = \lfloor \mu|B| \rfloor$. Suppose that $d_G(x, U_1 \cap X) \geq (1/2 + 7\mu/2)|U_1 \cap X|$ for each $X \in \{A, B\}$ and all $x \in V(G) \setminus X$. Let $U_1 \supseteq \cdots \supseteq U_\ell$ be a $(1/2 + 4\mu, \mu, m)$-vortex respecting $(U_1 \cap A, U_1 \cap B)$ in $G[U_1]$ such that $|U_i \cap B| \leq 3|U_i \cap A|$, for each $1 \leq i \leq \ell$. Then there exists $H_\ell \subseteq G[U_\ell]$ such that $G - H_\ell$ is C_{2k}-decomposable.*

If $\ell = 1$, the statement follows directly from Lemma 5.6.7 applied to G and U_1. Assume then that $\ell \geq 2$ and the statement holds for $\ell - 1$. Let $G' := G - G[U_2]$ and note that $\delta_{\text{bip}}(G') \geq 1/2 + 2\mu$ and $d_{G'}(x, U_1 \cap X) \geq (1/2 + \mu)|U_1 \cap X|$ for each $X \in \{A, B\}$ and all $x \in V(G) \setminus X$. Furthermore, for all $x \in V(G') \setminus U_1$, $d_{G'}(x) = d_G(x)$ so $2 \mid d_{G'}(x)$. Apply Lemma 5.6.7 to find an edge-disjoint collection $\mathcal{F}$ of copies of C_{2k} covering all edges in $G' - G[U_1]$ such that

$$\Delta\big(\bigcup \mathcal{F}[U_1]\big) \leq \mu^3|U_1| \leq 5\mu^2|U_2 \cap A|.$$

Let $G'' := G[U_1] - \bigcup \mathcal{F}$. Then G'' is a 2-divisible bipartite graph with $\delta_{\text{bip}}(G'') \geq 1/2 + 3\mu$. For each $X \in \{A, B\}$, $|U_2 \cap X| = \lfloor \mu|U_1 \cap X| \rfloor$ and, for any $x \in V(G'') \setminus X$,

$$d_{G''}(x, U_2 \cap X) \geq (1/2 + 4\mu)|U_2 \cap X| - \Delta\big(\bigcup \mathcal{F}[U_1]\big) \geq (1/2 + 7\mu/2)|U_2 \cap X|.$$

Since $G''[U_2] = G[U_2]$, $U_2 \supseteq \cdots \supseteq U_\ell$ is a $(1/2+4\mu, \mu, m)$-vortex respecting $(U_2 \cap A, U_2 \cap B)$ in $G''[U_2]$. Hence, by induction, there exists a subgraph H_ℓ of $G[U_\ell]$ such that $G'' - H_\ell$ has a C_{2k}-decomposition $\mathcal{F}'$. Together $\mathcal{F} \cup \mathcal{F}'$ is a C_{2k}-decomposition of $G - H_\ell$. □

5.7 Decompositions of expanders

The purpose of this section is to prove Theorem 5.5.2 which finds a C_{2k}-decomposition of any C_{2k}-divisible ν-expander G when $k \geq 4$. The significance of G being a ν-expander (defined in Section 5.5) is that there are many internally disjoint paths between any pair of vertices in G. We can use these paths to construct copies of C_{2k} and to find absorbers and this allows us to use the arguments of [38] with only slight modification. We will make use of the fact that ν-expansion is a robust property in the sense that the graph remains a $\nu/2$-expander when we remove a sparse subgraph.

5.7.1 Finding paths

The next result can be used to find many internally disjoint paths with predetermined endpoints without using any vertex too often.

Proposition 5.7.1. *Let $k \in \mathbb{N}$, $k \geq 4$ and $1/n \ll \gamma \ll \nu, 1/k$. Let G be a ν-expander on n vertices and let $P = \{(x_1, y_1), \ldots, (x_m, y_m)\}$ be a collection of $m \leq \gamma n^2$ pairs of distinct vertices of G. Suppose that each vertex appears in at most γn pairs in P. Then G contains a collection of edge-disjoint paths $\mathcal{P} = \{P^1, \ldots, P^m\}$ such that, for each $1 \leq i \leq m$, P^i is a path of length k from x_i to y_i. Furthermore, $\Delta(\bigcup \mathcal{P}) \leq \gamma^{1/3} n$.*

Proof. Let $1 \leq j \leq m$ and suppose we have already found paths $P^1, \ldots, P^{j-1}$ such that each vertex in $V(G)$ appears as an internal vertex in at most $2\sqrt{\gamma}n$ of the paths. Let B be the set of all vertices which appear as an internal vertex in at least $\sqrt{\gamma}n$ paths in $P^1, \ldots, P^{j-1}$. Note that

$$|B| \leq m(k-1)/(\sqrt{\gamma}n) \leq \nu^2 n.$$

Let $G_j := G - \bigcup_{i=1}^{j-1} P^i$. Note that $\Delta(\bigcup_{i=1}^{j-1} P^i) \leq 4\sqrt{\gamma}n + \gamma n$ so G_j is a $\nu/2$-expander (which implies $\delta(G_j) \geq \nu n/2$). We find a path P^j between x_j and y_j in G_j whose interior vertices avoid B as follows. Since $\nu n/2 \geq |B| + k$, we can embed a path of length $k-4$ starting at x_j greedily. Let x'_j denote its endpoint. In order to find a path of length four between x'_j and y_j it suffices to note that

$$|R_{\nu/2,G_j}(N_{G_j}(x'_j)) \cap R_{\nu/2,G_j}(N_{G_j}(y_j))| \geq \nu n \geq |B| + k.$$

Continuing in this way, we obtain edge-disjoint paths $P^1, \ldots, P^m$ of length k such that no vertex is used as an internal vertex more than $2\sqrt{\gamma}n$ times. Thus $\Delta(\bigcup \mathcal{P}) \leq 4\sqrt{\gamma}n + \gamma n \leq$

$\gamma^{1/3}n$. □

5.7.2 Expander vortices

We now introduce a further variant of the vortex, this time for expanders, where we replace the minimum degree condition with an expansion property instead. Let G be a graph on n vertices. A (ν, μ, m)-*expander vortex* in G is a sequence $U_0 \supseteq U_1 \supseteq \cdots \supseteq U_\ell$ such that

- $U_0 = V(G)$;
- $|U_i| = \lfloor \mu |U_{i-1}| \rfloor$, for all $1 \le i \le \ell$, and $|U_\ell| = m$;
- $N_G(x, U_i)$ is ν-expanding in $G[U_i]$, for all $1 \le i \le \ell$ and all $x \in U_{i-1}$.

Proposition 5.7.2. *Let $0 \le \nu \le 1$ and $1/n \ll \mu < 1$. Suppose that G is a ν-expander on n vertices. Then there exists $U \subseteq V(G)$ of size $\lfloor \mu n \rfloor$ such that, for every $x \in V(G)$, $N_G(x, U)$ is $(\nu - n^{-1/3})$-expanding in $G[U]$.*

Proof. Let U be a random subset of $V(G)$ of size $\lfloor \mu n \rfloor$. Fix $x \in V(G)$. Lemma 4.2.1 gives

$$\mathbb{P}(|R_{\nu,G}(N_G(x)) \cap U| < (1/2 + \nu - n^{-1/3})|U|) \le 2e^{-2n^{-2/3}|U|^2/n} \le 2e^{-\mu^2 n^{1/3}} \le 1/n^3.$$

Consider any $y \in R_{\nu,G}(N_G(x))$. Again by Lemma 4.2.1,

$$\mathbb{P}(d_G(y, N_G(x, U)) < (\nu - n^{-1/3})|U|) \le 2e^{-2n^{-2/3}|U|^2/n} \le 2e^{-\mu^2 n^{1/3}} \le 1/n^3.$$

By summing over all choices of x and y, we see that with probability at least $1 - 2/n$ the set U chosen in this way satisfies:

1. $|R_{\nu,G}(N_G(x)) \cap U| \ge (1/2 + \nu - n^{-1/3})|U|$, for all $x \in V(G)$ and
2. $d_G(y, N_G(x, U)) \ge (\nu - n^{-1/3})|U|$, for all $x \in V(G)$ and all $y \in R_{\nu,G}(N_G(x))$.

For any $x \in V(G)$, we have

$$|R_{\nu-n^{-1/3},G[U]}(N_G(x,U))| \overset{(2)}{\geq} |R_{\nu,G}(N_G(x)) \cap U| \overset{(1)}{\geq} (1/2+\nu-n^{-1/3})|U|$$

so U is the required set. □

We use the following result to find an expander vortex in G.

Lemma 5.7.3. *Let $0 \leq \nu \leq 1$ and $1/m' \ll \mu < 1$. Suppose that G is a ν-expander on $n \geq m'$ vertices. Then G has a $(\nu-\mu,\mu,m)$-expander vortex for some $\lfloor \mu m' \rfloor \leq m \leq m'$.*

This result follows from repeated applications of Proposition 5.7.2 (see Appendix B.2 for more details).

5.7.3 Covering most of the edges

In this section we decompose almost all of the graph G into cycles except for a very restricted remainder using the following result. This is exactly the technique we used in Section 5.6, so again we omit some details.

Lemma 5.7.4. *Let $k \in \mathbb{N}$, $k \geq 3$ and $1/m \ll \nu, 1/k$. Let G be a 2-divisible 4ν-expander and let $U_0 \supseteq U_1 \supseteq \cdots \supseteq U_\ell$ be a $(5\nu,\nu,m)$-expander vortex in G. Then there exists $H_\ell \subseteq G[U_\ell]$ such that $G-H_\ell$ is C_{2k}-decomposable.*

We require some preliminary results. The first of which finds an approximate C_{2k}-decomposition of G whilst maintaining control over the number of edges incident at all vertices in a given set X.

Lemma 5.7.5. *Let $k \in \mathbb{N}$, $k \geq 3$ and $1/n \ll \eta \ll \nu, 1/k$. Suppose that G is a ν-expander on n vertices and that $X \subseteq V(G)$ of size at most $\eta^{1/2}n$. Then there exists $H \subseteq G$ such that $G-H$ is C_{2k}-decomposable, $Y := \{x \in V(G) : d_H(x) > \eta n\}$ has size at most ηn and $X \cap Y = \emptyset$.*

Proof. We begin by finding edge-disjoint copies of C_{2k} which cover all the edges in $G[X]$. To this end, let $P_X := \{(x, y) : xy \in E_G(X)\}$. Since $|X| \leq \eta^{1/2}n$, $G - G[X]$ is a $3\nu/4$-expander and we may apply Proposition 5.7.1 (with P_X, $\eta^{1/2}$, $G - G[X]$, $2k-1$ and $3\nu/4$ playing the roles of P, γ, G, k and ν) to find a collection $\mathcal{P}_X$ of edge-disjoint paths of length $2k-1$ between the endpoints of each edge in $E_G(X)$ such that $\Delta(\bigcup \mathcal{P}_X) \leq \eta^{1/6}n$. Thus we obtain a collection $\mathcal{F}_X$ of edge-disjoint copies of C_{2k} which cover all of the edges in $G[X]$ such that $\Delta(\bigcup \mathcal{F}_X) \leq 2\eta^{1/6}n$. Let $G' := G \setminus \bigcup \mathcal{F}_X$.

Our next step is to cover all but at most one of the remaining edges incident at each vertex in X. For each $x \in X$, pair up the vertices in $N_{G'}(x)$, leaving at most one vertex. Let P'_X denote the list of pairs for all $x \in X$. Note that $G' \setminus X$ is a $\nu/2$-expander. Then, as previously, apply Proposition 5.7.1 (with P'_X, $\eta^{1/2}$, $G' \setminus X$, $2k-2$ and $\nu/2$ playing the roles of P, γ, G, k and ν) to find a collection $\mathcal{P}'_X$ of edge-disjoint paths of length $2k-2$ in $G' \setminus X$ between each pair in P'_X. These paths combine with edges incident at X to form a collection $\mathcal{F}'_X$ of edge-disjoint copies of C_{2k} which, together with $\mathcal{F}_X$, cover all but at most one edge incident at each $x \in X$.

Finally, let $H' := G - \bigcup(\mathcal{F}_X \cup \mathcal{F}'_X)$. Use the Erdős-Stone theorem to greedily find an η^3-approximate C_{2k}-decomposition of H' which we will denote by $\mathcal{F}$. Let $H := H' - \bigcup \mathcal{F}$ and note that $G - H$ has a C_{2k}-decomposition given by $\mathcal{F}_X \cup \mathcal{F}'_X \cup \mathcal{F}$. If $Y := \{x \in V(G) : d_H(x) > \eta n\}$, then $|Y| \leq 2e(H)/(\eta n) \leq \eta n$. Since $d_H(x) \leq 1$ for all $x \in X$, $X \cap Y = \emptyset$. □

We use Lemma 5.7.5 to prove the following result which finds a C_{2k}-decomposition of G so that every vertex has low degree in the remainder.

Lemma 5.7.6. *Let $k \in \mathbb{N}$, $k \geq 3$ and $1/n \ll \nu, 1/k$. Let G be a ν-expander on n vertices. Then G has an approximate C_{2k}-decomposition $\mathcal{F}$ such that $\Delta(G - \bigcup \mathcal{F}) \leq \nu n$.*

Proof. Choose $s, t \in \mathbb{N}$ and $\eta > 0$ such that

$$1/n \ll \eta \ll 1/s \ll 1/t \ll \nu, 1/k$$

and K_s has a K_t-decomposition (s and t exist by Proposition 5.6.3). Let $\mathcal{P} = \{V_1, \dots, V_s\}$ be an equipartition of $V(G)$ satisfying the following for all $1 \le i \le s$:

(i) $d_G(y, N_G(x, V_i)) \ge (\nu - \eta)|V_i|$ for all $x \in V(G)$ and all $y \in R_{\nu,G}(N_G(x))$;

(ii) $|R_{\nu,G}(N_G(x)) \cap V_i| \ge (1/2 + \nu - \eta)|V_i|$ for all $x \in V(G)$.

To see that such a partition exists, consider a random equipartition of $V(G)$ into s parts and apply Lemma 4.2.1 to see that this partition satisfies (i)–(ii) with probability at least 3/4. It will suffice to show that $G[\mathcal{P}]$ has an approximate C_{2k}-decomposition $\mathcal{F}$ such that $\Delta(G[\mathcal{P}] - \bigcup \mathcal{F}) \le \nu n/2$ (since $|V_i| \le \nu n/2$ for all $1 \le i \le s$).

Consider $\{T_1, \dots, T_\ell\}$, a K_t-decomposition of K_s, where $V(K_s) = \{1, \dots, s\}$. For each $1 \le i \le \ell$, define $G_i := \bigcup_{jk \in E(T_i)} G[V_j, V_k]$, so the G_i decompose $G[\mathcal{P}]$. Consider any $x \in V(G_i)$ and any $y \in R_{\nu,G}(N_G(x)) \cap V(G_i)$. We have

$$\begin{aligned} d_{G_i}(y, N_{G_i}(x)) = \sum_{\substack{V_j \subseteq V(G_i) \\ x,y \notin V_j}} d_G(y, N_G(x, V_j)) &\overset{\text{(i)}}{\ge} (t-2)(v - \eta)\lfloor n/s \rfloor \\ &\ge (\nu/2)t\lceil n/s \rceil \ge (\nu/2)|G_i|. \end{aligned} \tag{5.9}$$

So

$$\begin{aligned} |R_{\nu/2,G_i}(N_{G_i}(x))| \overset{(5.9)}{\ge} |R_{\nu,G}(N_G(x)) \cap V(G_i)| &\overset{\text{(ii)}}{\ge} t(1/2 + \nu - \eta)\lfloor n/s \rfloor \\ &\ge (1/2 + \nu/2)t\lceil n/s \rceil \ge (1/2 + \nu/2)|G_i|. \end{aligned}$$

Thus G_i is a $\nu/2$-expander for each $1 \le i \le \ell$.

Let $X_1 := \emptyset$. For each $1 \le i \le \ell$ in turn, apply Lemma 5.7.5 (with G_i, $\nu/2$ and $X_i \cap V(G_i)$ playing the roles of G, ν and X) to find $H_i \subseteq G_i$ such that $G_i - H_i$ is C_{2k}-decomposable, $d_{H_i}(x) \le \eta|G_i|$ for all $x \in X_i$ and $|Y_i| \le \eta|G_i|$, where $Y_i := \{x \in V(G_i) : d_{H_i}(x) > \eta|G_i|\}$. Let $X_{i+1} := X_i \cup Y_i$. Note that, for all $1 \le i \le \ell$, $|X_i| \le s^2\eta t\lceil n/s \rceil \le \eta^{1/2}t\lfloor n/s \rfloor$, so we can indeed use Lemma 5.7.5. Let $H := \bigcup_{i=1}^{\ell} H_i$ and consider any

$x \in V(G)$. We know that

$$d_H(x) \leq \ell\eta t\lceil n/s\rceil + t\lceil n/s\rceil \leq 2s\eta tn + 2tn/s \leq \nu n/2,$$

since $d_{H_i}(x) \leq \eta t\lceil n/s\rceil$ for all but at most one $1 \leq i \leq \ell$. □

The following proposition, an analogue of Proposition 5.6.6, takes a subset R of G and covers all the edges in a sparse subgraph H which have no endpoint in this set R. It is proved by mimicking the proof of Proposition 5.10 in [38] (see Appendix B.2 for more details).

Proposition 5.7.7. *Let $k \in \mathbb{N}$, $k \geq 3$ and $1/n \ll \gamma \ll \mu, 1/k$. Let G be a graph on n vertices and let $V(G) = L \,\dot\cup\, R$ such that $|R| \geq \mu n$ and $N_G(x, R)$ is μ-expanding in $G[R]$ for all $x \in V(G)$. Let H be any subgraph of $G[L]$ such that $\Delta(H) \leq \gamma n$. Then there exists a subgraph A of G such that $A[L]$ is empty, $A \cup H$ is C_{2k}-decomposable and $\Delta(A) \leq \gamma^{1/3}|R|$.*

We will obtain Lemma 5.7.4 from the following result by induction. The proof of Lemma 5.7.8 very closely resembles that of Lemma 5.6.7 (and uses Lemma 5.7.6, Proposition 5.7.7 and Proposition 5.7.1, in this order). We omit the details here and refer the reader instead to Appendix B.2.

Lemma 5.7.8. *Let $k \in \mathbb{N}$, $k \geq 3$ and $1/n \ll \nu, 1/k$. Let G be a 3ν-expander on n vertices and $U \subseteq V(G)$ with $|U| = \lfloor \nu n\rfloor$. Suppose that $N_G(x, U)$ is ν-expanding in $G[U]$ for all $x \in V(G)$. Then, if $2 \mid d_G(x)$ for all $x \in V(G) \setminus U$, there exists a collection $\mathcal{F}$ of edge-disjoint copies of C_{2k} such that every edge in $G - G[U]$ is covered and $\Delta(\bigcup \mathcal{F}[U]) \leq \nu^2|U|/4$.*

Finally, we use Lemma 5.7.8 and induction to prove Lemma 5.7.4.

Proof of Lemma 5.7.4. If $\ell = 0$, we can set $H_\ell := G$, so we assume $\ell \geq 1$. We will prove the following statement (which implies Lemma 5.7.4) by induction on ℓ.

Let G be a 2-divisible 4ν-expander and let $U_1 \subseteq V(G)$ of size $\lfloor \nu|G| \rfloor$ such that $N_G(x)$ is $9\nu/2$-expanding in $G[U_1]$ for all $x \in V(G)$. Let $U_1 \supseteq \cdots \supseteq U_\ell$ be a $(5\nu, \nu, m)$-expander vortex in $G[U_1]$. Then there exists $H_\ell \subseteq G[U_\ell]$ such that $G - H_\ell$ is C_{2k}-decomposable.

If $\ell = 1$, the statement follows directly from Lemma 5.7.8 applied to G and U_1. Assume then that $\ell \geq 2$ and the claim holds for $\ell - 1$. Let $G' := G - G[U_2]$ and note that G' is a 3ν-expander and $N_{G'}(x)$ is ν-expanding in $G'[U_1]$ for all $x \in V(G)$. Furthermore, for all $x \in V(G') \setminus U_1$, $d_{G'}(x) = d_G(x)$ so $2 \mid d_{G'}(x)$. Apply Lemma 5.7.8 to find a collection $\mathcal{F}$ of edge-disjoint copies of C_{2k} covering all edges in $G' - G[U_1]$ such that $\Delta(\bigcup \mathcal{F}[U_1]) \leq \nu^2|U_1|/4$. Let $G'' := G[U_1] - \bigcup \mathcal{F}$. Then G'' is 2-divisible and G'' is a 4ν-expander and $U_2 \subseteq V(G'')$ with $|U_2| = \lfloor \nu|G''| \rfloor$. Moreover, for any $x \in V(G'')$, $N_{G''}(x)$ is $9\nu/2$-expanding in $G[U_2]$. Since $G''[U_2] = G[U_2]$, $U_2 \supseteq \cdots \supseteq U_\ell$ is a $(5\nu, \nu, m)$-expander vortex in $G[U_2]$. Hence, by induction, there exists $H_\ell \subseteq G[U_\ell]$ such that $G'' - H_\ell$ has a C_{2k}-decomposition $\mathcal{F}'$. Together $\mathcal{F} \cup \mathcal{F}'$ is a C_{2k}-decomposition of $G - H_\ell$. □

Finally, we prove the main result in this section, Theorem 5.5.2.

Proof of Theorem 5.5.2. Let $m, m' \in \mathbb{N}$ and μ be such that

$$1/n \ll 1/m' \ll 1/m \ll \mu \ll \nu, 1/k.$$

Let $\nu' := \nu/7$. Apply Lemma 5.7.3 to G to find a $(6\nu', \nu', m)$-expander vortex $U_0 \supseteq U_1 \supseteq \cdots \supseteq U_\ell$ in G. Let $G_1 := G - G[U_1]$. Since $|U_1| \leq \nu' n$, G_1 is a $\nu/2$-expander (and if $k = 4$, $\delta(G_1) \geq n/2 - \nu' n$) which implies that between any two vertices in G_1, there are at least m' internally disjoint paths of length $k - 1$. Apply Lemma 5.3.2 to the graph G_1 with U_ℓ playing the role of U to find $A^* \subseteq G_1$ as in the lemma. Let $G^* := G - A^*$ and note that G^* is C_{2k}-divisible. We have $\Delta(A^*) \leq |A^*| \leq 2^{m^2}$, so G^* is a $4\nu'$-expander and $U_0 \supseteq U_1 \supseteq \cdots \supseteq U_\ell$ is a $(5\nu', \nu', m)$-vortex in G^*. Then apply Lemma 5.7.4 to G^* to find $H_\ell \subseteq G^*[U_\ell]$ such that $G^* - H_\ell$ has a C_{2k}-decomposition. Observing that $A^* \cup H_\ell$ has a C_{2k}-decomposition (by Lemma 5.3.2) completes the proof. □

5.8 Concluding remarks

In Theorem 5.1.2, we have found exact minimum degree bounds for a graph to have a decomposition into cycles of all even lengths apart from six. For cycles of length six, the best bound is given by Theorem 5.1.1 and remains at $(1/2+\varepsilon)|G|$ which is asymptotically best possible. We conjecture that the bound should also be $|G|/2$ in this case but were unable to prove this using the methods of Section 5.5. The primary reason for this was that we were unable to construct a C_6-absorber which could be found in a ν-expander. The transformer construction given in Section 5.3.1 works well for longer cycles since these transformers can be constructed using paths of length at least three between the fixed vertices. But, when the cycle is shorter, we do not have enough flexibility when choosing the intermediate vertices. This means that we were only able to prove the expander decomposition result, Theorem 5.5.2, for cycles of length at least eight. There are also places in the proofs of Lemmas 5.5.3 and 5.5.7 where we require the cycle to have length at least eight, though it is likely that these arguments could be adapted for C_6-decompositions if required.

APPENDIX A

SUPPLEMENTARY DETAILS FOR CHAPTER 4

Some results in Chapter 4 were very similar to existing results in [7] and their proofs follow in a similar manner, requiring only minor adaptations. For this reason, we omitted the details in the main body of this thesis but we include proofs here for completeness.

First we prove Lemma 4.5.2 which finds copies of $\mathcal{P}$-labelled graphs in an r-partite graph G.

Proof of Lemma 4.5.2. For each $v \in V(G)$ and each $0 \leq j \leq m$, let $s(v, j)$ be the number of indices $1 \leq i \leq j$ such that some vertex of H_i is labelled $\{v\}$. Note that (iv) implies that $s(v, j) \leq \eta n$.

Suppose that we have already found copies $\phi(H_1), \ldots \phi(H_{j-1})$ of $H_1, \ldots, H_{j-1}$ such that, for every $v \in V(G)$,

$$d_{G_{j-1}}(v) \leq \eta^{1/2} n + (s(v, j-1) + 1)b, \tag{A.1}$$

where $G_{j-1} := \bigcup_{1 \leq i \leq j-1} \phi(H_i)$. We show that we can find a copy of H_j in $G - G_{j-1}$ which satisfies (A.1) with j replacing $j-1$.

Let $B := \{v \in V(G) : d_{G_{j-1}}(v) \geq \eta^{1/2} n\}$. We have

$$|B| \leq \frac{2e(G_{j-1})}{\eta^{1/2} n} \leq \frac{2mdb}{\eta^{1/2} n} \leq \frac{2\eta dbn^2}{\eta^{1/2} n} \leq 2\eta^{1/2} dbn.$$

By (iii), we can order the vertices of H_j so that root vertices precede free vertices and each free vertex is preceded by at most d of its neighbours. We will embed the vertices in this order. Suppose that we are currently embedding the vertex x. If x is a root vertex, embed it at its assigned vertex. This is possible since we have not yet embedded any neighbour of x.

Suppose that x is a free vertex labelled $V \subseteq V_i$. Let U denote the image of the neighbours of x in H_j which have already been embedded. Note that $|U| \leq d$ and, by the definition of a $\mathcal{P}$-labelling, $U \cap V_i = \emptyset$. Then (i) implies that $d_G(U, V) \geq \varepsilon|V|$. We have

$$\begin{aligned} d_{G-G_{j-1}}(U, V) &\geq d_G(U, V) - \sum_{u \in U} d_{G_{j-1}}(u, V) \overset{(A.1)}{\geq} \varepsilon|V| - d(\eta^{1/2}n + (\eta n + 1)b) \\ &> |B| + |H_j|. \end{aligned}$$

So we can map x to a suitable vertex in $V \setminus B$.

Suppose that we have embedded all vertices of H_j. We now check that (A.1) holds with j replacing $j-1$. If $v \in V(G) \setminus B$, this is clear. Suppose then that $v \in B$. If v was used in the embedding of H_j, v must be the image of a root vertex and $s(v, j) = s(v, j-1) + 1$. So in all cases,

$$d_{G_j}(v) \leq \eta^{1/2}n + (s(v, j) + 1)b.$$

Continue in this way until all the H_i have been embedded. Using (A.1),

$$\Delta(H) = \Delta(G_m) \leq \eta^{1/2}n + (\eta n + 1)b \leq \varepsilon n,$$

as required. □

Next we find a partition sequence in G, proving Lemma 4.7.2. This result follows from repeated applications of Proposition 4.7.1.

Proof of Lemma 4.7.2. Choose m_0 such that $1/m_0 \ll 1/k, 1/r, \alpha$ and let $m' \geq m_0$. Let $\ell := \lfloor \log_k(n/m') \rfloor$. Define $\mathcal{P}_0, \dots, \mathcal{P}_\ell$ as follows. Let $\mathcal{P}_0 := \{V(G)\}$. For each $j \in \mathbb{N}$, let

$$a_j := n^{-1/3} + (n/k)^{-1/3} + \cdots + (n/k^{j-1})^{-1/3}.$$

Suppose that, for some $1 \le p \le \ell$, we have already chosen $\mathcal{P}_0, \dots, \mathcal{P}_{p-1}$ such that, for each $1 \le i \le p-1$ and each $W \in \mathcal{P}_{i-1}$, $\mathcal{P}_i[W]$ is an $(a_i, k, \delta - a_i)$-partition for $G[W]$ and for all $1 \le j_1, j_2, j_3 \le r$ with $j_1 \ne j_2, j_3$, each $U \in \mathcal{P}_i[W]$ and each $v \in W_{j_1}$,

$$|d_G(v, U_{j_2}) - d_G(v, U_{j_3})| < 2a_i|U_{j_1}|. \tag{A.2}$$

Let $\beta := \lfloor n/k^{p-1} \rfloor^{-1/3}/2$. For each $W \in \mathcal{P}_{p-1}$, $|W_1| = \cdots = |W_r| \ge n/k^{p-1} - 1 \ge n/k^{\ell-1} - 1 \ge m_0$. Also, $\hat{\delta}(G[W]) \ge (\delta - a_{p-1})|W_1|$. So we can choose a $(\beta, k, \delta - \beta - a_{p-1})$-partition $\mathcal{P}_W$ for $G[W]$, using Proposition 4.7.1. Note that $\beta + a_{p-1} < a_p$ so $\mathcal{P}_W$ is an $(a_p, k, \delta - a_p)$-partition. Let $\mathcal{P}_p := \bigcup_{W \in \mathcal{P}_{p-1}} \mathcal{P}_W$.

Consider any $W \in \mathcal{P}_{p-1}$, any $U \in \mathcal{P}_p[W]$, any $1 \le j_1, j_2, j_3 \le r$ with $j_1 \ne j_2, j_3$ and any $v \in W_{j_1}$. We use that $\mathcal{P}_W$ is a $(\beta, k, 0)$-partition for $G[W]$ together with (A.2) to see that

$$\begin{aligned}
|d_G(v, U_{j_2}) - d_G(v, U_{j_3})| &< |d_G(v, U_{j_2}) - d_G(v, W_{j_2})/k| + |d_G(v, U_{j_3}) - d_G(v, W_{j_3})/k| \\
&\quad + |d_G(v, W_{j_2})/k - d_G(v, W_{j_3})/k| \\
&< \beta|U_{j_2}| + \beta|U_{j_3}| + 2a_{p-1}|W_{j_1}|/k \\
&\le 2(\beta + a_{p-1})|U_{j_1}| + 2a_{p-1} \le 2a_p|U_{j_1}|.
\end{aligned}$$

Finally, we note that

$$a_\ell = (n/k^{\ell-1})^{-1/3} \sum_{i=0}^{\ell-1} k^{-i/3} \le \frac{(n/k^{\ell-1})^{-1/3}}{1 - k^{-1/3}} \le \frac{m_0^{-1/3}}{1 - 2^{-1/3}} \le \frac{\alpha}{2}.$$

This completes the proof with $m = \lceil n/k^\ell \rceil$. □

Finally, we prove Corollary 4.7.5. The proof is simply a case of verifying that the sequence of graphs $R_1, \dots, R_\ell$ obtained by Lemma 4.7.4 have the required properties.

Proof of Corollary 4.7.5. Apply Lemma 4.7.4 to G to find a sequence of graphs $R_1, \ldots, R_\ell$, such that $R_q \subseteq G_q - G_{q-1}$ for each $1 \le q \le \ell$, which satisfies the following. For all $1 \le q \le \ell$, all $1 \le j \le r$, all $W \in \mathcal{P}_{q-1}$, all distinct $x, y \in W$ and all $U \in \mathcal{P}_q[W]$,

$$|d_{R_q}(x, U_j) - \rho d_{G_q}(x, U_j)| < \alpha |U_j| \qquad \text{(A.3)}$$

$$|d_{R_q}(\{x, y\}, U_j) - \rho^2 d_{G_q}(\{x, y\}, U_j)| < \alpha |U_j| \qquad \text{(A.4)}$$

$$d_{G'_{q+1}}(y, N_{R_q}(x, U_j)) \ge \rho d_{G_{q+1}}(y, N_{G_q}(x, U_j)) - 3\rho^2 |U_j| \qquad \text{(A.5)}$$

where $G_{\ell+1} := G$. So properties (i) and (ii) hold.

We now show that (iii) is satisfied. Fix $1 \le q \le \ell$, $1 \le j, j' \le r$, $W \in \mathcal{P}_{q-1}$, $U, U' \in \mathcal{P}_q[W]$ and $x \in W \setminus (U \cup U' \cup V_j \cup V_{j'})$. Since $\mathcal{P}_q[W]$ is an $(\alpha, k, 1 - 1/r + \varepsilon)$-partition for $G[W]$,

$$|d_{G_q}(x, U_j) - d_{G_q}(x, U'_j)| < \alpha |U_j| + \alpha |U'_j|.$$

We use that $\mathcal{P}_1, \ldots, \mathcal{P}_\ell$ is a partition sequence to see that

$$|d_{G_q}(x, U'_j) - d_{G_q}(x, U'_{j'})| < \alpha |U'_j|.$$

Together these give

$$|d_{G_q}(x, U_j) - d_{G_q}(x, U'_{j'})| < 3\alpha(|U_j| + 1). \qquad \text{(A.6)}$$

We use (A.6) together with (A.3) to see that

$$\begin{aligned} |d_{R_q}(x, U_j) - d_{R_q}(x, U'_{j'})| &\le |d_{R_q}(x, U_j) - \rho d_{G_q}(x, U_j)| + |d_{R_q}(x, U'_{j'}) - \rho d_{G_q}(x, U'_{j'})| \\ &\quad + \rho |d_{G_q}(x, U_j) - d_{G_q}(x, U'_{j'})| \\ &< \alpha |U_j| + \alpha |U'_{j'}| + 3\alpha\rho(|U_j| + 1) \le 3\alpha |U_j|. \end{aligned}$$

So (iii) holds.

For (iv), fix $1 \le q \le \ell$, $1 \le j \le r$, $W \in \mathcal{P}_{q-1}$, $U \in \mathcal{P}_q[W]$, $x \in W \setminus U$ and $y \in U$.

Suppose $x, y \notin V_j$. Since $\mathcal{P}_q[W]$ is an $(\alpha, k, 1-1/r+\varepsilon)$-partition for $G[W]$,

$$d_{G_{q+1}}(y, U_j) \geq d_G(y, U_j) - |U_j|/k - 1 \geq (1 - 1/r + \varepsilon/2)|U_j| \quad \text{and} \tag{A.7}$$

$$d_{G_q}(x, U_j) \geq (1 - 1/r)|U_j|. \tag{A.8}$$

So

$$\begin{aligned} d_{G_{q+1}}(y, N_{G_q}(x, U_j)) &\geq d_{G_q}(x, U_j) + d_{G_{q+1}}(y, U_j) - |U_j| \overset{(A.7)}{\geq} d_{G_q}(x, U_j) - |U_j|/r + \varepsilon|U_j|/2 \\ &\overset{(A.8)}{\geq} (1 - 1/(r-1))d_{G_q}(x, U_j) + \varepsilon|U_j|/2. \end{aligned} \tag{A.9}$$

Thus

$$\begin{aligned} d_{G'_{q+1}}(y, N_{R_q}(x, U_j)) &\overset{(A.5)}{\geq} \rho d_{G_{q+1}}(y, N_{G_q}(x, U_j)) - 3\rho^2|U_j| \\ &\overset{(A.9)}{\geq} \rho\big[(1 - 1/(r-1))d_{G_q}(x, U_j) + \varepsilon|U_j|/2\big] - 3\rho^2|U_j| \\ &\geq \rho(1 - 1/(r-1))d_{G_q}(x, U_j) + \rho^{5/4}|U_j| \end{aligned}$$

and (iv) holds. □

APPENDIX B

SUPPLEMENTARY DETAILS FOR CHAPTER 5

We omitted some proofs from Chapter 5 in order to make the argument more concise and draw attention to new results. These proofs are obtained by slight modifications to proofs of similar results in [38] and we provide full details here for completeness.

B.1 Decompositions of bipartite graphs

This section supports Section 5.6. We will prove Lemma 5.6.1 which finds a vortex in a bipartite graph G. We use the following proposition which is a simple application of Lemma 4.2.1.

Proposition B.1.1. *Let* $0 \leq \delta \leq 1$ *and* $1/n \ll \mu < 1$. *Suppose that* $G = (A, B)$ *is a bipartite graph with* $n = |A| \leq |B| \leq 3n$ *and* $\delta_{\mathrm{bip}}(G) \geq \delta$. *Then there exists* $U \subseteq V(G)$ *such that* $|U \cap X| = \lfloor \mu |X| \rfloor$ *and* $d_G(x, U \cap X) \geq (\delta - n^{-1/3}) \lfloor \mu |X| \rfloor$ *for each* $X \in \{A, B\}$ *and every* $x \in V(G) \setminus X$.

Proof of Lemma 5.6.1. For each $X \in \{A, B\}$ and each $i \in \mathbb{N}$, define $n_0(X) := |X|$ and $n_i(X) := \lfloor \mu n_{i-1}(X) \rfloor$. Let $a_0 := 0$ and $a_i := |A|^{-1/3} \sum_{j=1}^{i} \mu^{-(j-1)/3}$. Let $n_i := n_i(A) + n_i(B)$. Observe that

$$\mu^i |X| - 1/(1-\mu) \leq n_i(X) \leq \mu^i |X| \tag{B.1}$$

and so

$$n_i(A) \leq n_i(B) \leq 3(\mu^i|B|/2 - 2) \leq 3(\mu^i|A| - 2) \leq 3n_i(A).$$

Define $\ell := 1 + \max\{i \geq 0 : n_i \geq m'\}$, $m := n_\ell$ and note that $2\lfloor \mu m' \rfloor \leq m \leq m'$.

Let $1 \leq i \leq \ell$. Suppose that we have already found $U_0, \dots, U_{i-1}$, a $(\delta - 2a_{i-1}, \mu, n_{i-1})$-vortex respecting (A, B) in G. By Proposition B.1.1, there exists $U_i \subseteq U_{i-1}$ such that, for each $X \in \{A, B\}$, $|U_i \cap X| = n_i(X)$ and $d_G(x, U_i \cap X) \geq (\delta - 2a_{i-1} - n_{i-1}(A)^{-1/3})n_i(X)$ for every $x \in U_{i-1} \setminus X$. Since

$$2(a_i - a_{i-1}) = 2(|A|\mu^{i-1})^{-1/3} \overset{(B.1)}{\geq} 2(n_{i-1}(A) + 2)^{-1/3} \geq n_{i-1}(A)^{-1/3},$$

$U_0, \dots, U_i$ is a $(\delta - 2a_i, \mu, n_i)$-vortex respecting (A, B) in G. Eventually we find $U_0, \dots, U_\ell$, a $(\delta - 2a_\ell, \mu, m)$-vortex respecting (A, B) in G.

Finally, observe that $\mu^{\ell-1}|A| \geq n_{\ell-1}(A) \geq n_{\ell-1}/4 \geq m'/4$. So

$$a_\ell = |A|^{-1/3} \frac{\mu^{-\ell/3} - 1}{\mu^{-1/3} - 1} \leq \frac{(\mu^{\ell-1}|A|)^{-1/3}}{1 - \mu^{1/3}} \leq \frac{(m'/4)^{-1/3}}{1 - \mu^{1/3}} \leq \mu/2$$

and the result follows. □

B.2 Decompositions of expanders

In Section 5.7 we found C_{2k}-decompositions of expanders. This section contains some additional details. We begin by proving Lemma 5.7.3 which finds an expander vortex in G.

Proof of Lemma 5.7.3. Define $n_0 := n$ and, for each $i \in \mathbb{N}$, $n_i := \lfloor \mu n_{i-1} \rfloor$. Then

$$\mu^i n - 1/(1 - \mu) \leq n_i \leq \mu^i n.$$

If we let $\ell := 1 + \max\{i \geq 0 : n_i \geq m'\}$ and $m := n_\ell$, then $\lfloor \mu m' \rfloor \leq m \leq m'$. Let $a_0 := 0$

and, for each $i \in \mathbb{N}$, let $a_i := n^{-1/3} \sum_{j=1}^{i} \mu^{-(j-1)/3}$.

Fix $1 \le i \le \ell$ and suppose that we have already found a $(\nu - 2a_{i-1}, \mu, n_{i-1})$-expander vortex $U_0, \ldots, U_{i-1}$ in G. By Proposition 5.7.2, there exists $U_i \subseteq U_{i-1}$ of size n_i such that $N_G(x, U_i)$ is $(\nu - 2a_{i-1} - n_{i-1}^{-1/3})$-expanding in $G[U_i]$ for every $x \in U_{i-1}$. Thus, $U_0, \ldots, U_i$ is a $(\nu - 2a_i, \mu, n_i)$-expander vortex in G. All that remains is to note that $U_0, \ldots, U_\ell$ is a $(\nu - 2a_\ell, \mu, m)$-expander vortex in G and

$$a_\ell = n^{-1/3} \frac{\mu^{-\ell/3} - 1}{\mu^{-1/3} - 1} \le \frac{(\mu^{\ell-1} n)^{-1/3}}{1 - \mu^{1/3}} \le \frac{m'^{-1/3}}{1 - \mu^{1/3}} \le \mu/2$$

since $\mu^{\ell-1} n \ge n_{\ell-1} \ge m'$. Thus $U_0, \ldots, U_\ell$ is a $(\nu - \mu, \mu, m)$-expander vortex as required.

□

We now prove Proposition 5.7.7 which is used to cover all edges in a sparse subgraph H of G which have no endvertex in a set R. Its proof is very similar to that of Proposition 5.10 in [38].

Proof of Proposition 5.7.7. Enumerate the edges of $E(H)$: $e_1, \ldots, e_m$. For each edge e_i in turn, we will find a copy F_i of C_{2k} which contains e_i such that $V(F_i) \cap L = V(e_i)$. The graphs $F_1, \ldots, F_m$ must be edge-disjoint.

Fix $1 \le j \le m$ and suppose that we have already found $F_1, \ldots, F_{j-1}$. Let $G_{j-1} := \bigcup_{i=1}^{j-1} F_i$ and suppose that $\Delta(G_{j-1}) \le \sqrt{\gamma} n + 2$. Let $X := \{x \in V(G) : d_{G_{j-1}}(x) > \sqrt{\gamma} n\}$ and note that $X \cap L = \emptyset$, since $d_{G_{j-1}}(x) \le 2\Delta(H) \le \sqrt{\gamma} n$ for all $x \in L$. We have

$$|X| \sqrt{\gamma} n \le 2e(G_{j-1}) \le 4ke(H) \le 2k\gamma n^2,$$

giving $|X| \le 2k\sqrt{\gamma} n \le \gamma^{1/3}|R|$. Let $G' := (G - G_{j-1})[(R \setminus X) \cup V(e_j)]$. Recall that, for any $x \in V(G')$, $N_G(x, R)$ is μ-expanding in $G[R]$, i.e., $|R_{\mu, G[R]}(N_G(x, R))| \ge (1/2 + \mu)|R|$. Since $\Delta(G_{j-1}), |X| \le \gamma^{1/3}|R|$, we have

$$|R_{\mu/2, G'}(N_{G'}(x))| \ge (1/2 + \mu/2)(|R| + 2) \ge (1/2 + \mu/2)|G'|.$$

Thus G' is a $\mu/2$-expander. This allows us to find a copy F_j of C_{2k} in G' that contains e_j. Moreover, F_j avoids X so $\Delta(G_j) \leq \sqrt{\gamma}n + 2$. Letting $A := \bigcup_{i=1}^{m}(F_i - e_i)$ completes the proof. □

Finally, we prove Lemma 5.7.8 (the proof uses the same ideas as the corresponding bipartite result Lemma 5.6.7).

Proof of Lemma 5.7.8. Choose constants γ, ξ such that $1/n \ll \gamma \ll \xi \ll \nu, 1/k$. Let $W := V(G) \setminus U$, $m := \lceil \xi^{-1} \rceil$ and $M := \binom{m+1}{2}$. Let $V_1, \dots, V_M$ be an equipartition of U such that for all $x \in V(G)$ and all $1 \leq i \leq M$,

$$N_G(x, V_i) \text{ is } \nu/2\text{-expanding in } G[V_i]. \tag{B.2}$$

To see that such a partition exists, consider a random equipartition of U into M parts. Apply Lemma 4.2.1 to see that such a partition satisfies (B.2) with probability at least 3/4.

Let $W_1, \dots, W_m$ be an equipartition of W and let $G_W^1, \dots, G_W^M$ be an enumeration of the M graphs of the form $G[W_i]$ or $G[W_i, W_j]$. Note $G[W] = \bigcup_{i=1}^{M} G_W^i$ and

$$|G_W^i| \leq 2(|W|/m + 1) \leq 2\xi n$$

for all $1 \leq i \leq M$. For each $1 \leq i \leq M$, let $R_i := G[V_i, V(G_W^i)]$. Let $R := \bigcup_{i=1}^{M} R_i$. Note that $\Delta(R) \leq \max\{2\xi n, 2|U|m/M\} \leq 4\xi n$.

Let $G' := G - (G[U] \cup R)$. Since $|U| = \lfloor \nu n \rfloor$ and $\Delta(R) \leq 2\xi n$, we note that G' is a ν-expander. So, by Lemma 5.7.6, G' has an approximate C_{2k}-decomposition $\mathcal{F}_1$ such that $H := G' - \bigcup \mathcal{F}_1$ satisfies $\Delta(H) \leq \gamma n$.

We now use R and Proposition 5.7.7 to cover the edges in $H[W]$. For each $1 \leq i \leq M$, let $H_i := H[W] \cap G_W^i$ (so $H[W] = \bigcup H_i$) and $G_i := G[V_i] \cup R_i \cup H_i$. Note that $V(G_i) = V_i \cup V(G_W^i)$ and thus $\nu\xi^2 n/10 \leq |V_i| \leq |G_i| \leq 3\xi n$, implying that $|V_i| \geq \xi^2 |G_i|$. Also $\Delta(H_i) \leq \gamma n \leq \sqrt{\gamma}|G_i|$ and (B.2) implies that $N_{G_i}(x, V_i)$ is ξ^2-expanding in $G[V_i]$

for all $x \in V(G_i)$. So we may apply Proposition 5.7.7 (with G_i, $\sqrt{\gamma}$, ξ^2 and V_i playing the roles of G, γ, μ and R) to find $A_i \subseteq G_i$ such that $A_i[V(G_i) \setminus V_i]$ is empty, $A_i \cup H_i$ is C_{2k}-decomposable and $\Delta(A_i) \leq \xi^4 |V_i|$. Let $A := \bigcup_{i=1}^M A_i$. So $A \cup H[W]$ has a C_{2k}-decomposition $\mathcal{F}_2$ and $\Delta(A) \leq \xi n$.

We must now cover the remaining edges in $H[U, W] \cup R'$. Let $G'' := G - \bigcup(\mathcal{F}_1 \cup \mathcal{F}_2)$. Note that $G''[W]$ is empty and $\Delta(G'') \leq \Delta(H) + \Delta(R) \leq \gamma n + 4\xi n \leq 5\xi n$. Since $\Delta(A) \leq \xi n$, $G''[U]$ is a $\nu/2$-expander. For each $w \in W$, $d_{G''}(w)$ is even, so we can pair up the vertices in $N_{G''}(w)$ arbitrarily and let P denote the list of pairs of all neighbours of W. Each vertex in U appears in at most $\Delta(G'') \leq 5\xi n \leq \sqrt{\xi}|U|$ of the pairs in P and

$$|P| \leq \Delta(G'')n \leq 5\xi n^2 \leq \sqrt{\xi}|U|^2.$$

Then we can apply Proposition 5.7.1 to $G''[U]$ (with $|U|$, $\sqrt{\xi}$, $\nu/2$, $2k-2$ playing the roles of n, γ, ν and k) to find a collection $\mathcal{F}_3$ of edge-disjoint copies of C_{2k} which cover the edges of $G'' - G''[W]$ such that $\Delta(\mathcal{F}_3) \leq \nu^3|U|$. Let $\mathcal{F} := \mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3$. Then $\mathcal{F}$ covers every edge of $G - G[U]$ and $\Delta(\mathcal{F}[U]) \leq \Delta(A) + \bigcup \mathcal{F}[U] \leq \nu^2|U|/4$. $\square$

INDEX

LIST OF REFERENCES

[1] M. Ajtai, J. Komlós and E. Szemerédi, The first occurrence of Hamilton cycles in random graphs, *Ann. Disc. Math.* **27** (1985), 173–178.

[2] N. Alon, Explicit Ramsey graphs and orthonormal labellings, *Electronic J. Combin.* **1** (1994), R12.

[3] N. Alon and J.H. Spencer, *The Probabilistic Method*, Wiley-Interscience, New York (1992).

[4] B. Alspach, The wonderful Walecki construction, *Bull. Inst. Combin. Appl.* **52** (2008), 7–20.

[5] L.D. Anderson and A.J.W. Hilton, Thanks Evans!, *Proc. London Math. Soc.* **47** (1983), 507–522.

[6] B. Barber, D. Kühn, A. Lo, R. Montgomery and D. Osthus, Fractional clique decompositions of dense graphs and hypergraphs, *arXiv:1507.04985*, 2015.

[7] B. Barber, D. Kühn, A. Lo and D. Osthus, Edge-decompositions of graphs with high minimum degree, *Advances in Mathematics* **288** (2016), 337–385.

[8] B. Barber, D. Kühn, A. Lo, D. Osthus and A. Taylor, Clique decompositions of multipartite graphs and completion of Latin squares, *arXiv:1603.01043*, 2016.

[9] P. Bartlett, Completions of ε-dense partial Latin squares, *J. Combin. Designs* **21** (2013), 447–463.

[10] P. Bennett and T. Bohman, A note on the random greedy independent set algorithm, *Random Structures Algorithms* **49(3)** (2016), 479–502.

[11] T. Bohman, The triangle-free process, *Advances in Math.* **221** (2009), 1653–1677.

[12] T. Bohman and P. Keevash, The early evolution of the H-free process, *Invent. Math.* **181** (2010), 291–336.

[13] T. Bohman and P. Keevash, Dynamic concentration of the triangle-free process, *arXiv:1302.5963*, 2013.

[14] T. Bohman, D. Mubayi and M. Picollelli, The independent neighborhoods process, *Israel J. Math.* **214(1)** (2016), 333–357.

[15] B. Bollobás, The evolution of sparse graphs, *Graph theory and combin.* Academic Press, London (1984), 35–57.

[16] B. Bollobás and O. Riordan, Constrained graph processes, *Electronic J. Combin.* **7** (2000), R18.

[17] F. Bowditch and P. Dukes, Fractional triangle decompositions of dense 3-partite graphs, *arXiv:1510.08998*, 2015.

[18] M.C. Cai, A counterexample to a conjecture of Grant, *Discrete Math.* **44**(1) (1983), 111.

[19] N. Cavenagh, Decomposing dense bipartite graphs into 4-cycles, *Electron. J. Combin.* **22(1)** (2015) #P1.50.

[20] A.G. Chetwynd and R. Häggkvist, Completing partial Latin squares where each row, column and symbol is used at most cn times, *Reports Dept. Mathematics*, Univ. Stockholm (1985).

[21] S. Chowla, P. Erdős and E. Straus, On the maximal number of pairwise orthogonal Latin squares of a given order, *Canadian J. Math.* **12** (1960), 204–208.

[22] V. Chvátal, On Hamilton's ideals, *J. Combin. Theory B.* **12** (1972), 163–168.

[23] D.E. Daykin and R. Häggkvist, Completion of sparse partial latin squares, Graph theory and combinatorics (Cambridge, 1983), pages 127–132, Academic Press, London, 1984.

[24] L. DeBiasio, D. Kühn, T. Molla, D. Osthus and A. Taylor, Arbitrary orientations of Hamilton cycles in digraphs, *SIAM J. Discrete Math.* **29** (2015), 1553–1584.

[25] L. DeBiasio and T. Molla, Semi-degree threshold for anti-directed Hamilton cycles, *Electronic J. Combin.* **22**(4) (2015), 4–34.

[26] G.A. Dirac, Some theorems on abstract graphs, *Proc. London Math. Soc.* **2** (1952), 69–81.

[27] D. Dor and M. Tarsi, Graph decomposition is NP-complete: a complete proof of Holyer's conjecture, *SIAM J. Comput.* **26** (1997), 1166–1187.

[28] F. Dross, Fractional triangle decompositions in graphs with large minimum degree, *SIAM J. Discrete Math.* **30** (2016), 36–42.

[29] P. Dukes, Rational decomposition of dense hypergraphs and some related eigenvalue estimates, *Linear Algebra Appl.* **436(9)** (2012), 3736–3746. Corrigendum: *Linear Algebra Appl.* **467** (2015), 267–269.

[30] P. Erdős, Some remarks on the theory of graphs, *Bull. Amer. Math. Soc.* **53** (1947), 292–294.

[31] P. Erdős, Graph theory and probability, *Canad. J. Math.* **11** (1959), 34–38.

[32] P. Erdős and A. Rényi, On random graphs. I., *Publicationes Mathematicae* **6** (1959), 290–297.

[33] P. Erdős, S. Suen and P. Winkler, On the size of a random maximal graph, *Random Structures Algorithms* **6** (1995), 309–318.

[34] P. Erdős and P. Tetali, Representation of integers as the sum of k terms, *Random Structures Algorithms* **1** (1990), 245–261.

[35] G. Fiz Pontiveros, S. Griffiths and R. Morris, The triangle-free process and $R(3,k)$, Mem. Amer. Math. Soc., to appear.

[36] K. Garaschuk, *Linear methods for rational triangle decompositions.* PhD thesis, Univ. Victoria, https://dspace.library.uvic.ca//handle/1828/5665, 2014.

[37] A. Ghouila-Houri, Une condition suffisante d'existence d'un circuit Hamiltonien, *C.R. Acad. Sci. Paris* **25** (1960), 495–497.

[38] S. Glock, D. Kühn, A. Lo, R. Montgomery and D. Osthus, On the decomposition threshold of a given graph, *arXiv:1603.04724*, 2016.

[39] D. Grant. Antidirected Hamilton cycles in digraphs, *Ars Combinatoria* **10** (1980), 205–209.

[40] T. Gustavsson, *Decompositions of large graphs and digraphs with high minimum degree.* PhD thesis, Univ. Stockholm (1991).

[41] R. Häggkvist and A. Thomason, Oriented Hamilton cycles in oriented graphs, *Combinatorics, Geometry and Probability*, Cambridge University Press (1997), 339–353.

[42] R. Häggkvist and A. Thomason, Oriented Hamilton cycles in digraphs, *J. Graph Theory* **20** (1995), 471–479.

[43] P.E. Haxell and V. Rödl, Integer and fractional packings in dense graphs, *Combinatorica* **21(1)** (2001), 13–38.

[44] S. Janson, T. Łuczak and A. Ruciński, *Random graphs.* Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[45] R. Karp, Reducibility among combinatorial problems, *Complexity of computer computations* (1972), 85–103.

[46] P. Keevash, The existence of designs, *arXiv:1401.3665*, 2014.

[47] P. Keevash, D. Kühn and D. Osthus, An exact minimum degree condition for Hamilton cycles in oriented graphs, *J. London Math. Soc.* **79** (2009), 144–166.

[48] P. Keevash and R. Mycroft, A geometric theory for hypergraph matching, *Memoirs American Math. Soc.* **233** (2014), number 1098.

[49] L. Kelly, Arbitrary orientations of Hamilton cycles in oriented graphs, *Electronic J. Combin.* **18** (2011).

[50] J.H. Kim, The Ramsey number $R(3,t)$ has order of magnitude $t^2/\log t$, *Random Structures Algorithms* **7** (1995), 173–207.

[51] T.P. Kirkman, On a problem in combinatorics, *Cambridge Dublin Mathematical Journal* **2** (1847), 191–204.

[52] T.P. Kirkman, *Query VI.* Lady's and Gentleman's Diary, page 48, 1850.

[53] D. Kühn, F. Knox and D. Osthus, Edge-disjoint Hamilton cycles in random graphs, *Random Structures Algorithms* **46** (2015), 397–445.

[54] D. Kühn, J. Lapinskas and D. Osthus, Optimal packings of Hamilton cycles in graphs of high minimum degree, *Combin. Probab. Comput.* **22** (2013), 394–416.

[55] D. Kühn and D. Osthus. Hamilton decompositions of regular expanders: A proof of Kelly's conjecture for large tournaments, *Advances in Math.* **237** (2013), 62–146.

[56] D. Kühn, D. Osthus and A. Taylor, On the random greedy F-free hypergraph process, *SIAM J. Discrete Math.* **30** (2016), 1343–1350.

[57] D. Kühn, D. Osthus and A. Treglown, Hamilton Degree Sequences in Digraphs, *J. Combinatorial Theory B* **100** (2012), 367–380.

[58] A. Lo and K. Markström, A multipartite Hajnal-Szemerédi theorem for graphs and hypergraphs, *Comb. Probab. Comput.* **22** (2013), 97–111.

[59] W. Mantel, Problem 28, *Wiskundige Opgaven* (1907), 10:60–61.

[60] R. Montgomery, Fractional clique decompositions of dense partite graphs, *arXiv:1603.01039*, 2016.

[61] J. Moon and L. Moser, On Hamiltonian bipartite graphs, *Israel J. Math.* **1** (1963), 163–165.

[62] C.St.J.A. Nash-Williams, An unsolved problem concerning decomposition of graphs into triangles. In *Combinatorial Theory and its Applications III*, pages 1179–1183. North Holland, 1970.

[63] C.St.J.A. Nash-Williams, Hamiltonian circuits, *Studies in Math.* **12** (1975), 301–360.

[64] D. Osthus and A. Taraz, Random maximal H-free graphs, *Random Structures Algorithms* **18** (2001), 61–82.

[65] M. Picollelli, The final size of the C_ℓ-free process, *SIAM J. Discrete Math.* **28(3)** (2014), 1276–1305.

[66] M. Picollelli, The diamond-free process, *Random Structures Algorithms* **45** (2014), 513-551.

[67] L. Pósa, A theorem concerning hamiltonian lines, *Magyar Tud. Akad. Mat. Fiz. Oszt. Kozl.* **7** (1962), 255–226.

[68] R. Raman, The power of collision: randomized parallel algorithms for chaining and integer sorting. In *Foundations of software technology and theoretical computer science (Bangalore, 1990)*, volume 472 of *Lecture Notes in Comput. Sci.*, pages 161–175. Springer, Berlin, 1990.

[69] F. Ramsey, On a problem of formal logic, *Proc. London Math. Soc.* **30** (1930), 264–286.

[70] A. Ruciński and N. Wormald, Random graph processes with degree restrictions, *Combin. Probab. Comput.* **1** (1992), 169–180.

[71] B. Smetianuk, A new construction on latin squares I. A proof of the Evans conjecture, *Ars Combinatorica* **11** (1981), 155–172.

[72] J.H. Spencer, Maximal triangle-free graphs and Ramsey $R(3,t)$, unpublished manuscript (1995).

[73] E. Szemerédi, Regular Partitions of Graphs, *Problémes Combinatoires et Théorie des Graphes Colloques Internationaux CNRS* **260** (1978), 399–401.

[74] A. Taylor, The regularity method for graphs and digraphs, MSci thesis, School of Mathematics, University of Birmingham, *arXiv:1406.6531*, 2013.

[75] A. Taylor, On the exact decomposition threshold for even cycles, *arXiv:1607.06315*, 2016.

[76] A. Thomason, Paths and cycles in tournaments, *Proc. American Math. Soc.* **296** (1986), 167–180.

[77] T.W. Tillson, A Hamiltonian decomposition of K^*_{2m}, $2m \geq 8$, *J. Combin. Theory* **29** (1980), 68–74.

[78] P. Turán, On an extremal problem in graph theory, *Matematikai és Fizikai Lapok* **48** (1941), 436–452.

[79] W. Tutte, The factorisation of linear graphs, *J. London Math. Soc.* **22** (1947), 107–111.

[80] L. Warnke, The C_ℓ-free process, *Random Structures Algorithms* **44** (2014), 490–526.

[81] L. Warnke, When does the K_4-free process stop?, *Random Structures Algorithms* **44** (2014), 355–397.

[82] R.M. Wilson, Decompositions of complete graphs into subgraphs isomorphic to a given graph. In *Proceedings of the Fifth British Combinatorial Conference (Univ. Aberdeen, Aberdeen, 1975)*, pages 647–659. Congressus Numerantium, No. XV, Utilitas Math., Winnipeg, Man., 1976.

[83] G. Wolfovitz, The K_4-free process, *arXiv:1008.4044*, 2010.

[84] N.C. Wormald, Differential equations for random processes and random graphs, *Ann. Appl. Probab.* **5** (1995), 1217–1235.

[85] R. Yuster, Packing and decomposition of graphs with trees, *J. Combin. Theory Series B* **78** (2000), 123–140.

[86] R. Yuster, Asymptotically optimal K_k-packings of dense graphs via fractional K_k-decompositions, *J. Combin. Theory Ser. B* **95(1)** (2005), 1–11.