Paths and Cycles in Graphs and Hypergraphs

by

Vincent Pfenninger

A thesis submitted to the University of Birmingham for the degree of Doctor of Philosophy in Pure Mathematics

Supervisor: Dr Allan Lo

Co-supervisor: Prof. Daniela Kühn

School of Mathematics

College of Engineering and Physical Sciences

University of Birmingham

18th November 2022

Abstract

In this thesis we present new results in graph and hypergraph theory all of which feature paths or cycles.

A *k-uniform tight cycle* $C_n^{(k)}$ is a k-uniform hypergraph on n vertices with a cyclic ordering of its vertices such that the edges are all k-sets of consecutive vertices in the ordering.

We consider a generalisation of Lehel's Conjecture, which states that every 2-edge-coloured complete graph can be partitioned into two cycles of distinct colour, to k-uniform hypergraphs and prove results in the 4- and 5-uniform case.

For a k-uniform hypergraph H, the *Ramsey number* $r(H)$ is the smallest integer N such that any 2-edge-colouring of the complete k-uniform hypergraph on N vertices contains a monochromatic copy of H. We determine the Ramsey number for 4-uniform tight cycles asymptotically in the case where the length of the cycle is divisible by 4, by showing that $r(C_{4n}^{(4)}) = (5 + o(1))n$.

We prove a resilience result for tight Hamiltonicity in random hypergraphs. More precisely, we show that for any $\gamma > 0$ and $k \geq 3$ asymptotically almost surely, every subgraph of the binomial random k-uniform hypergraph $G^{(k)}(n, n^{\gamma-1})$ in which all $(k-1)$-sets are contained in at least $(\frac{1}{2} + 2\gamma)pn$ edges has a tight Hamilton cycle.

A random graph model on a host graph H is said to be 1-*independent* if for every pair of vertex-disjoint subsets A, B of $E(H)$, the state of edges (absent or present) in A is independent of the state of edges in B. We show that $p = 4 - 2\sqrt{3}$ is the critical probability such that every 1-independent graph model on $\mathbb{Z}^2 \times K_n$ where each edge is present with probability at least p contains an infinite path.

DECLARATION

Sections 1.1 and 1.2 and Chapters 2 and 3 are based on joint work with Allan Lo [82, 83] (my contribution: 50% for each paper). Section 1.3 and Chapter 4 are based on joint work with Peter Allen and Olaf Parczyk [8] (my contribution: 33%). Section 1.4 and Chapter 5 are based on joint work with Victor Falgas-Ravry [47] (my contribution: 50%).

ACKNOWLEDGEMENTS

First of all I would like to thank my supervisor, Allan, for his guidance, encouragement and fruitful discussions on research throughout my PhD. I am also grateful for my other collaborators Olaf, Peter and Victor. It has been a pleasure working with you and I have learned a lot. Moreover, I would like to thank the combinatorics group at Birmingham and my office mates for such a friendly work environment. My PhD was made possible by generous funding provided by the Engineering and Physical Sciences Research Council (EPSRC) and by the College of Engineering and Physical Sciences at the University of Birmingham. Special thanks goes to my family, especially my parents, for being so supportive of my dreams to become a mathematician and to Amarja for always being there for me.

CONTENTS

CHAPTER 1

INTRODUCTION

Combinatorial problems have captured the interest of humanity for thousands of years. This is evidenced by the fact that one of the oldest known mathematical documents, the Rhind mathematical papyrus, which is more than 3500 years old, contains what can be interpreted as a combinatorial exercise [23]. The origin of graph theory, one of the most important areas of modern combinatorics, is much more recent and is commonly attributed to the work of Euler on the problem of the Bridges of Königsberg dating back to 1735 [118]. A graph G is defined to be a pair of sets $(V(G), E(G))$, where $V(G)$ is called the set of vertices of G and $E(G)$, the set of edges of G, is a subset of the set of 2-element subsets of $V(G)$. The objects of study in this thesis will be graphs and hypergraphs (a natural generalisation of graphs where each edge can contain any number of vertices instead of having to contain exactly two). In this thesis we will consider problems in two areas of combinatorics, extremal combinatorics and probabilistic combinatorics. These areas of combinatorics were both popularised by Paul Erdős and now form an integral part of modern combinatorics.

Extremal combinatorics concerns itself with how large or small a given parameter can be in a given set of discrete structures. One of the first such problems was considered by Mantel [87] in 1907 who showed that the maximal number of edges that a triangle-free graph on n vertices can have is $\left\lfloor \frac{n^2}{4} \right\rfloor$. One branch of extremal combinatorics that has seen extensive study is Ramsey theory which originated when Ramsey [100] proved what is now known as Ramsey's theorem. Ramsey's theorem states that for any integers $s, t \geq 1$ there exists an integer $R(s, t)$ such that any red-blue edge-colouring of the complete graph on $R(s, t)$ vertices contains a complete graph on s vertices which contains only red edges or a complete graph on t vertices which contains only blue edges. In general problems in Ramsey theory ask how big a certain structure has to be in order for it to be guaranteed to contain a specific ordered substructure. In Chapter 2 we consider a generalisation to hypergraphs of Lehel's conjecture which is closely related to Ramsey theory and in Chapter 3 we consider a problem in hypergraph Ramsey theory.

Combinatorics and probability theory have been closely related ever since the advent of probability theory in the seventeenth century. The relevance of combinatorics to probability theory is natural given the fact that one often has to enumerate combinatorial objects when calculating probabilities. Perhaps more surprising is the fact that probability theory can often be helpful in solving purely deterministic problems in combinatorics that do not at first seem to have anything to do with randomness. The use of such probabilistic techniques in combinatorics is called the probabilistic method and will feature at several points in the work presented in this thesis. Probabilistic combinatorics is about both the study of random discrete structures and about the probabilistic method. The most widely studied random structure in graph theory is the Erdős–Rényi random graph $G_{n,p}$ which is the graph on n vertices where each edge is present independently with probability p. The study of such random graphs was initiated by Erdős and Rényi [44] in 1959. In Chapter 4 we will consider a generalisation of the Erdős–Rényi random graph to hypergraphs and in Chapter 5 we will study different random graph models, where, in particular, we relax the condition that edges are added independently.

The chapters of this thesis will consider the following topics. In Chapter 2 we will examine a generalisation of monochromatic cycle partitioning to hypergraphs, Chapter 3 will be about the Ramsey number for 4-unifrom tight cycles, Chapter 4 will be about resilience for Hamiltonicity in random hypergraphs, and Chapter 5 will consider percolation in random graph models with a relaxed independence condition. We will now present our results and give a more detailed introduction for each of the relevant areas.

1.1 Monochromatic cycle partitioning

An r-*edge-colouring* of a graph (or hypergraph) is a colouring of its edges with r colours. A *monochromatic subgraph* of an r-edge-coloured graph is one in which all the edges have the same colour.

An old observation of Erdős and Rado states that every 2-edge-coloured complete graph contains a monochromatic spanning tree. Although this statement is easy to prove this led to further study of Ramsey type problems for large sparse structures. Gerencsér and Gyárfás [54] showed the tight result that any 2-edge-colouring of K_n contains a monochromatic path of length at least $\left\lfloor \frac{2n}{3} \right\rfloor$. They remarked that the easier result that any 2-edge-coloured K_n contains a monochromatic path of length at least $n/2$ can be shown by noting that the path of maximal length that consists of a red[1] path followed by a blue path is always spanning. In particular, every 2-edge-coloured complete graph on n vertices admits a partition of the vertex set into a red path and a blue path. Gyárfás [57] went on to show the stronger statement that in every 2-edge-coloured K_n the vertices can be covered by a red cycle and a blue cycle that share at most one vertex. Lehel conjectured that this statement remains true if one asks for the cycles to be disjoint. More precisely Lehel conjectured that every 2-edge-colouring of the complete graph on n vertices admits a partition of the vertex set into two monochromatic cycles of distinct colours, where the empty set, a single vertex and a single edge are considered to be *degenerate cycles* and are allowed as cycles for the partition. The conjecture was first stated in [11] where it was proved for special types of 2-edge-colourings of K_n. The conjecture was then proved for large n by Łuczak, Rödl and Szemerédi [85] using Szemerédi's Regularity Lemma. Allen [1] subsequently improved the bound on n by giving a different proof that avoids the use of the regularity lemma. Finally Bessy and Thomassé [22] proved Lehel's conjecture for all n

[1] We will always assume that the two colours used in a 2-edge-colouring are red and blue.

by giving a short and clever proof.

Similar problems for the complete bipartite graph $K_{n,n}$ have also been studied. Gyárfás and Lehel [59, 57] showed that any 2-edge-coloured $K_{n,n}$ contains a red path and a blue path that are disjoint and together cover all but possibly a single vertex unless the colouring is a split colouring. Here a *split colouring* is one such that each colour class consists of the disjoint union of two complete bipartite graphs. This was improved by Pokrovskiy [97] who showed that any 2-edge-colouring of $K_{n,n}$ that is not a split colouring admits a partition of the vertices into two monochromatic paths of distinct colours. Recently, Stein [113] further generalised this by showing that any such colouring admits a partition into a monochromatic cycle and a monochromatic path of distinct colours.

It is natural to ask if the condition for Lehel's conjecture that the graph be complete is necessary or if a minimum degree condition suffices. It was conjectured by Balogh, Barát, Gerbner, Gyárfás and Sárközy [20] that Lehel's conjecture can be strengthened in this way. They conjectured that any 2-edge-coloured graph G on n vertices with $\delta(G) > 3n/4$ can be partitioned into a red and a blue cycle. They gave a construction showing that their conjecture is best possible and proved an approximate version of it by showing that, for every $\varepsilon > 0$ and large n, under the stronger assumption that $\delta(G) > (3/4+\varepsilon)n$ all but εn of the vertices can be partitioned into a red and a blue cycle. This was improved by DeBiasio and Nelsen [38] who showed that under this stronger minimum degree condition a partition of all the vertices into a red and a blue cycle can be obtained. Finally Letzter [80] proved the full conjecture for large n. In a similar direction Pokrovskiy [98] conjectured that, for n large, any 2-edge-coloured graph G on n vertices with $\delta(G) > 2n/3$ can be partitioned into 3 monochromatic cycles and that if instead $\delta(G) > n/2$ then G can be partitioned into four monochromatic cycles. An approximate version of the first part of this conjecture was proved by Allen, Böttcher, Lang, Skokan and Stein [5] who showed that, for any $\varepsilon > 0$, if $\delta(G) \geq (2/3+\varepsilon)n$, then G can be partitioned into 3 monochromatic cycles. The second part of the conjecture however was disproved by Korándi, Lang, Letzter and Pokrovskiy [73] who showed that for sufficiently large n there exists a 2-edge-coloured graph G on n vertices with $\delta(G) \geq n/2 + \log n/(16 \log\log n)$ whose vertices cannot be partitioned into fewer than $\log n/(32 \log\log n)$ monochromatic cycles.

Similar problems have also been considered for colourings with a general number of colours. In particular, a lot of attention has been given to the problem of determining the number of monochromatic cycles that are needed to partition an r-edge-coloured complete graph. Erdős, Gyárfás and Pyber [45] proved that every r-edge-coloured complete graph can be partitioned into $O(r^2 \log r)$ monochromatic cycles and conjectured that r monochromatic cycles would suffice. Their result was improved by Gyárfás, Ruszinkó, Sárközy and Szemerédi [62] who showed that $O(r \log r)$ monochromatic cycles are enough. However, Pokrovskiy [97] disproved the conjecture by showing that for each $r \geq 3$ there exist infinitely many r-edge-coloured complete graphs which cannot be partitioned into r monochromatic cycles. Even so, in these counterexamples it is still possible to cover all but 1 of the vertices with r vertex-disjoint monochromatic cycles. This lead Pokrovskiy to propose a weaker version of the conjecture stating that each r-edge-coloured complete graph contains r vertex-disjoint monochromatic cycles that together cover all but at most c_r of the vertices, where c_r is a constant depending only on r. Pokrovskiy [98] subsequently proved that we can take $c_3 \leq 43000$ for large enough n. Minimum degree conditions have also been considered in this setting. It was shown by Korándi, Lang, Letzter and Pokrovskiy [73] that there exists a constant c such that any r-edge-coloured graph on n vertices with minimum degree at least $n/2 + cr \log n$ can be partitioned into $O(r^2)$ monochromatic cycles. They also provided a construction showing that this is essentially best possible.

Recently, generalisations of Lehel's conjecture to hypergraphs have also been considered. For any positive integer k, a *k-uniform hypergraph*, or *k-graph*, H is an ordered pair of sets $(V(H), E(H))$ such that $E(H) \subseteq \binom{V(H)}{k}$, where $\binom{S}{k}$ is the set of all subsets of S of size k. Let $K_n^{(k)}$ be the complete k-graph on n vertices.

In k-graphs there are several notions of cycle. For integers $1 \leq \ell < k < n$, a k-graph C on n vertices is called an *ℓ-cycle* if there is an ordering of its vertices $V(C) = \{v_0, \ldots, v_{n-1}\}$

such that $E(C) = \{\{v_{i(k-\ell)}, \dots, v_{i(k-\ell)+k-1}\} \colon 0 \leq i \leq n/(k-\ell) - 1\}$, where the indices are taken modulo n. That is, an ℓ-cycle is a k-graph with a cyclic ordering of its vertices such that its edges are sets of k consecutive vertices and consecutive edges share exactly ℓ vertices. (Note that a k-uniform ℓ-cycle on n vertices only exists if $k - \ell$ divides n.) A single edge or any set of fewer than k vertices is considered to be a *degenerate ℓ-cycle*. Further, 1-cycles and $(k-1)$-cycles are called *loose cycles* and *tight cycles*, respectively.

For loose cycles, Gyárfás and Sárközy [60] showed that every r-edge-coloured complete k-graph on n vertices can be partitioned into $c(k, r)$ monochromatic loose cycles. Sárközy [110] showed that, for n sufficiently large, $50kr \log(kr)$ loose cycles are enough. For tight cycles, Bustamante, Corsten, Frankl, Pokrovskiy and Skokan [27] showed that every r-edge-coloured complete k-graph can be partitioned into $C(k, r)$ monochromatic tight cycles. See [58] for a survey on other results about monochromatic cycle partitions and related problems.

In Chapter 2 we investigate monochromatic tight cycle partitions in 2-edge-coloured complete k-graphs on n vertices. When $k = 3$, Bustamante, Hàn and Stein [28] showed that there exist two vertex-disjoint monochromatic tight cycles of distinct colours covering all but at most $o(n)$ of the vertices. Recently, Garbe, Mycroft, Lang, Lo and Sanhueza-Matamala [53] proved that two monochromatic tight cycles are sufficient to cover all vertices. However, these cycles may not be of distinct colours. We show that for all $k \geq 3$, there are arbitrarily large 2-edge-coloured complete k-graphs that cannot be partitioned into two monochromatic tight cycles of distinct colours.

Proposition 1.1.1. *For all $k \geq 3$ and $m \geq k+1$, there exists a 2-edge-colouring of $K^{(k)}_{k(m+1)+1}$ that does not admit a partition into two tight cycles of distinct colours.*

It is interesting to note that, as was recently proved by Stein [113], every 2-edge-coloured $K^{(3)}_n$ admits a partition into two tight paths[1] of distinct colours.

It is natural to ask whether we can cover almost all vertices of a 2-edge-coloured complete k-graphs with two vertex-disjoint monochromatic tight cycles of distinct colours.

[1]A *k-uniform tight path* is a k-graph obtained from a k-uniform tight cycle by deleting a vertex.

The case when $k = 3$ is affirmed in [28]. Here, we show that this is true when $k = 4$.

Theorem 1.1.2. *For every $\varepsilon > 0$, there exists an integer n_1 such that, for all $n \geq n_1$, every 2-edge-coloured complete 4-graph on n vertices contains two vertex-disjoint monochromatic tight cycles of distinct colours covering all but at most εn of the vertices.*

When $k = 5$, we prove a weaker result that four monochromatic tight cycles are sufficient to cover almost all vertices.

Theorem 1.1.3. *For every $\varepsilon > 0$, there exists an integer n_1 such that, for all $n \geq n_1$, every 2-edge-coloured complete 5-graph on n vertices contains four vertex-disjoint monochromatic tight cycles covering all but at most εn of the vertices.*

1.2 Ramsey theory

The *Ramsey number* $r(H_1, \ldots, H_m)$ for k-graphs $H_1, \ldots, H_m$ is the smallest integer N such that any m-edge-colouring of the complete k-graph $K_N^{(k)}$ contains a monochromatic copy of H_i in the i-th colour for some $1 \leq i \leq m$. If $H_1, \ldots, H_m$ are all isomorphic to H then we let $r_m(H) = r(H_1, \ldots, H_m)$ and call it the *m-colour Ramsey number for H*. We also write $r(H)$ for $r_2(H)$ and simply call it the *Ramsey number for H*.

The Ramsey number for the complete graph has seen extensive study. Nevertheless, the important problem of determining the asymptotic behaviour remains widely open. The best known bounds are

$$(1 - o(1))\frac{\sqrt{2}n}{e}\sqrt{2}^n \leq r(K_n) \leq n^{-C\frac{\log n}{\log \log n}} 4^n$$

for a constant $C > 0$, where the lower bound is by Spencer [112] and the upper bound is by Conlon [31]. As this example shows, determining the asymptotic behaviour of the Ramsey number for certain graphs can be very hard. In general, we know even less about Ramsey numbers for hypergraphs. However, Ramsey numbers for cycles in graphs and hypergraphs, which we focus on here, are a bit better understood. In particular, cycles in graphs and hypergraphs have bounded degree and their Ramsey numbers are thus known to be linear in the number of their vertices. That bounded degree graphs have Ramsey numbers linear in their number of vertices was shown by Chvatál, Rödl, Szemerédi and Trotter [29]. The analogous result for 3-uniform hypergraphs was shown independently by Cooley, Fountoulakis, Kühn and Osthus [34] and by Nagle, Olsen, Rödl and Schacht [93]. Subsequently Cooley, Fountoulakis, Kühn and Osthus proved a generalisation to hypergraphs of any constant uniformity [35]. A shorter proof of this that gives better constants was later given by Conlon, Fox, and Sudakov [32] by avoiding the use of the hypergraph regularity lemma that was employed by previous proofs.

The Ramsey number for cycles in graphs was determined exactly in [26, 48, 108]. In particular, for $n \geq 5$, we have

$$r(C_n) = \begin{cases} \frac{3}{2}n - 1, & \text{if } n \text{ is even,} \\ 2n - 1, & \text{if } n \text{ is odd.} \end{cases}$$

Note that there is a dependence on the parity of the length of the cycle. For the m-colour Ramsey number, Jenssen and Skokan [68] proved that for $m \geq 2$ and any large enough odd integer n we have $r_m(C_n) = 2^{m-1}(n-1) + 1$.

Some Ramsey numbers for k-graphs related to cycles have also been studied. A k-uniform tight cycle $C_n^{(k)}$ is a k-graph on n vertices with a cyclic ordering of its vertices such that its edges are the sets of k consecutive vertices. The Ramsey number of the 3-uniform tight cycle on n vertices $C_n^{(3)}$ was determined asymptotically by Haxell, Łuczak, Peng, Rödl, Ruciński and Skokan, see [66, 67]. They showed that, for $i \in \{1, 2\}$,

$$r(C_{3n}^{(3)}) = (1 + o(1))4n \quad \text{and} \quad r(C_{3n+i}^{(3)}) = (1 + o(1))6n.$$

Note that just as for cycles in graphs the Ramsey number for 3-uniform tight cycles depends on the parity of the length of the cycle.

We define the *k-uniform tight path on n vertices* $P_n^{(k)}$ to be the k-graph obtained from $C_{n+1}^{(k)}$ by deleting a vertex. Using the bound on the Turán number for tight paths that was recently shown by Füredi, Jiang, Kostochka, Mubayi and Verstraëte [52], one can deduce that $r(P_n^{(k)}) \leq k(n-k+1)$ for any even $k \geq 2$.

The Ramsey number for loose cycles have also been studied. We denote by $LC_n^{(k)}$, where $n = \ell(k-1)$, the k-uniform loose cycle on n vertices, that is the k-graph with vertex set $\{v_1, \dots, v_n\}$ and edges $e_i = \{v_{1+i(k-1)}, \dots, v_{k+i(k-1)}\}$ for $0 \leq i \leq \ell - 1$, where indices are taken modulo n. Note that the k-uniform loose cycle on n vertices only exists if n is divisible by $k-1$. Haxell, Łuczak, Peng, Rödl, Ruciński, Simonovits, and Skokan [65] showed that $r(LC_n^{(3)}) = (1+o(1))5n/4$ for even n. This was generalised to all uniformities by Gyárfás, Sárközy and Szemerédi [61] who showed that, for $k \geq 3$ and n divisible by $k-1$,

$$r(LC_n^{(k)}) = (1+o(1))\frac{2k-1}{2k-2}n.$$

Recently, the exact values of Ramsey numbers for loose cycles have been determined in various cases, see [111] for more details.

Another problem of interest in this area is determining the Ramsey number of a complete graph and a cycle. For graphs, Keevash, Long and Skokan [69] showed that there exists an absolute constant $C \geq 1$ such that

$$r(C_\ell, K_n) = (\ell-1)(n-1)+1 \text{ provided } \ell \geq \frac{C\log n}{\log\log n}.$$

Analogous problems for hypergraphs have also been considered. See [90, 92, 96] for the analogous problem with loose, tight and Berge cycles, respectively.

In this Chapter 3 we will consider the Ramsey number for tight cycles. We determine the Ramsey number for the 4-uniform tight cycle on n vertices $C_n^{(4)}$ asymptotically in the case where n is divisible by 4.

Theorem 1.2.1. *Let $\varepsilon > 0$. For n large enough we have $r(C_{4n}^{(4)}) \leq (5+\varepsilon)n$.*

It is easy to see that this is asymptotically tight.

Proposition 1.2.2. *For $n, k \geq 2$, we have that $r(C_{kn}^{(k)}) \geq (k+1)n - 1$.*

Proof. Let $N = (k+1)n - 2$. We show that there exists a red-blue edge-colouring of $K_N^{(k)}$ that does not contain a monochromatic copy of $C_{kn}^{(k)}$. We partition the vertex set of $K_N^{(k)}$ into two sets X and Y of sizes $n-1$ and $kn-1$, respectively. We colour every edge that intersects the set X red and every other edge blue. It is easy to see that this red-blue edge-colouring of $K_N^{(k)}$ does not even contain a monochromatic matching of size n and thus also cannot contain a monochromatic copy of $C_{kn}^{(k)}$. Indeed, there is no red matching of size n since every red edge must intersect X and $|X| = n - 1$. Moreover, there is no blue matching of size n since all blue edges are entirely contained in Y and $|Y| = kn - 1$. $\square$

It is clear that the proof of Proposition 1.2.2 also shows that $r(P_{4n+i}^{(4)}) \geq 5n - 1$ for $0 \leq i \leq 3$. Since $C_{4(n+1)}^{(4)}$ contains $P_{4n+i}^{(4)}$ for each $0 \leq i \leq 3$, Theorem 1.2.1 also determines the Ramsey number for the 4-uniform tight path asymptotically.

Corollary 1.2.3. *We have $r(P_n^{(4)}) = (5/4 + o(1))n$.*

1.3 Resilience for Hamiltonicity in random graphs

The study of Hamilton cycles in graphs is one of the oldest topics in graph theory. In extremal graph theory, Dirac [41] in 1952 proved that an n-vertex graph with minimum degree at least $\frac{n}{2}$ contains a Hamilton cycle. The graph $K_{n,n+1}$ shows that this result is tight. In random graph theory, Pósa [99] and Korshunov [74, 75] independently showed in the 1970s that Hamilton cycles first appear in the random graph $G(n,p)$ — that is, the n-vertex graph where edges are present independently with probability p — at a threshold $p = \Theta\left(\frac{\log n}{n}\right)$. Komlós and Szemerédi [72] showed that the sharp threshold for Hamiltonicity coincides with that for minimum degree 2, and Bollobás [24] strengthened this by showing a hitting time version: if edges are added one by one, the edge which causes minimum degree 2 will asymptotically almost surely[1] also cause Hamiltonicity.

[1] Asymptotically almost surely (a.a.s.) is with probability tending to 1 as n tends to infinity.

Combining these areas, Sudakov and Vu [114] introduced the term *resilience* (though the same concept appears earlier in work of Alon, Capalbo, Kohayakawa, Rödl, Ruciński and Szemerédi [9]). They proved that for each $\gamma > 0$, the random graph $\Gamma = G(n,p)$ is a.a.s. $\left(\frac{1}{2}+\gamma\right)$-resiliently Hamiltonian whenever $p = \omega(n^{-1}\log^4 n)$; that is, every subgraph of Γ with minimum degree at least $\left(\frac{1}{2}+\gamma\right)pn$ has a Hamilton cycle. This result is sharp in the minimum degree, for the same reason as Dirac's theorem, but the probability can be improved. This was done over a succession of papers: Lee and Sudakov [79] showed that p can be reduced to the threshold $\Omega(n^{-1}\log n)$ and recently Montgomery [91] and, independently, Nenadov, Steger and Trujić [95] showed the hitting time version of this result (for which one needs to be a little more careful with edge deletion: it is permitted to delete only a $\left(\frac{1}{2}-\gamma\right)$-fraction of the edges at any given vertex).

Hamilton cycles in hypergraphs have only much more recently been attacked. There are several natural notions of paths and cycles in hypergraphs: the one that will concern us here is that of tight paths and cycles in k-uniform hypergraphs. That is, we work with hypergraphs in which all edges have uniformity k. We say that a given linear ordering of some vertices is a *tight path* if each consecutive k-set of vertices forms an edge; a given cyclic ordering of some vertices with the same condition forms a tight cycle. The $k = 2$ case of this definition reduces to the usual paths and cycles in graphs. For brevity, in what follows we write *k-graph* for k-uniform hypergraph.

In terms of extremal results, there are again several reasonable questions — one should place some form of 'minimum degree' condition for tight Hamilton cycles, but this can take the form of insisting that every j-set of vertices is in sufficiently many edges, where one can choose j between 1 and $k-1$. This leads to several significantly different problems (and even more if one considers other notions of cycle). We refer the reader to the comprehensive survey of Kühn and Osthus [77] for details, and focus on the version of minimum degree we want to work with. This is the case $j = k-1$, sometimes called *codegree*. Here, the Hamiltonicity problem is resolved. Rödl, Ruciński and Szemerédi [103, 104], first for 3-uniform and then for general uniformity, showed that if n is sufficiently large, any

n-vertex k-graph with minimum codegree at least $\left(\frac{1}{2}+\gamma\right)n$ (i.e. every $(k-1)$-set is in at least that many edges) contains a tight Hamilton cycle. For 3-graphs, they [105] were also able to give the exact result for sufficiently large n (finding exactly what should replace the error term γn).

In random hypergraphs, Dudek and Frieze [42, 43] found for several different notions of 'cycle' the threshold for Hamiltonicity in the binomial random hypergraph $G^{(k)}(n,p)$, that is the n-vertex k-graph in which k-sets are edges independently with probability p. In particular, in [43] they showed by the second moment method that for $k=3$ the threshold is $\omega\left(n^{-1}\right)$, and for $k \geq 4$ the sharp threshold is at en^{-1}. Narayanan and Schacht [94] strengthened these results, in particular showing that en^{-1} is also the sharp threshold for $k=3$.

Combining these (and answering a question of Frieze [51]), we prove, in Chapter 4, the following corresponding codegree resilience statement.

Theorem 1.3.1. *Given any $\gamma > 0$ and $k \geq 3$, if $p \geq n^{-1+\gamma}$, we show that $\Gamma = G^{(k)}(n,p)$ a.a.s. satisfies the following. Let G be any n-vertex subgraph of Γ such that $\delta_{k-1}(G) \geq \left(\frac{1}{2}+2\gamma\right)pn$. Then G contains a tight Hamilton cycle.*

Observe that this theorem is sharp in the minimum degree requirement, but it is presumably not sharp in the probability. More precisely, when $p = \Omega(\log n/n)$ then a.a.s. in Γ there is an n-vertex subgraph G such that $\delta_{k-1}(G) \geq (1/2-\gamma)pn$ and G does not contain a tight Hamilton cycle. When $p = o(\log n/n)$, there a.a.s. are $(k-1)$-tuples in Γ that are not contained in any edges and, therefore, no G as required by the theorem exists. For this regime the resilience condition needs to be adjusted, perhaps as explained above for the hitting time results in graphs from [91, 95]. We certainly need $p \geq 2en^{-1}$ for any statement of this kind to be true, otherwise randomly deleting half of the edges from Γ would a.a.s. destroy the tight Hamiltonicity.

This is the first resilience statement for tight Hamilton cycles in sparse random hypergraphs to the best of our knowledge; however for Berge cycles, Clemens, Ehrenmüller and Person [30] proved a resilience statement which is both tight in the minimum degree and has only a polylogarithmic gap in the probability. For perfect matchings it was shown by Ferber and Hirschfeld [49] that the same codegree resilience as in Theorem 1.3.1 holds with $p = \Omega(\log n/n)$, which is significantly above the threshold for the appearance of perfect matchings, but optimal for the same reasons as discussed above. More generally, Ferber and Kwan [50] studied the transference of results for perfect matchings in dense hypergraphs into resilience statements in random hypergraphs.

It would be interesting to investigate this transference for other types of Hamilton cycles and other degree conditions. For example, in the case of 3-graphs Reiher, Rödl, Ruciński, Schacht, and Szemerédi [101] show that any n-vertex 3-graph with minimum vertex degree $(\frac{5}{9}+\gamma)\binom{n}{2}$ contains a tight Hamilton cycle. Can this be extended to a resilience statement in random 3-graphs? More precisely, can the condition $\delta_2(G) \geq (\frac{1}{2}+\gamma)pn$ in Theorem 1.3.1 for $k = 3$ be replaced by $\delta_1(G) \geq (\frac{5}{9}+\gamma)p\binom{n}{2}$? The bound on the minimum degree would again be sharp.

1.4 Percolation and 1-independent random graph models

Percolation theory lies at the interface of probability theory, statistical physics and combinatorics. Its object of study is, roughly speaking, the connectivity properties of random subgraphs of infinite connected graphs, and in particular the points at which these undergo drastic transitions such as the emergence of infinite components. Since its inception in Oxford in the late 1950s, percolation theory has become a rich field of study (see e.g. the monographs [25, 56, 88]). One of the cornerstones of the discipline is the Harris–Kesten Theorem [64, 70], which states that if each edge of the integer square lattice $\mathbb{Z}^2$ is open independently at random with probability p, then if $p \leq \frac{1}{2}$ almost surely all connected components of open edges are finite, while if $p > \frac{1}{2}$ almost surely there exists an infinite connected component of open edges. Thus 1/2 is what is known as the *critical probability for independent bond percolation* on $\mathbb{Z}^2$.

In general, given an infinite connected graph H, determining the critical probability for independent bond percolation on H is a hard problem, with the answer known exactly only in a handful of cases. There is thus great interest in methods for rigorously estimating such critical probabilities. One of the most powerful and effective techniques for doing just that was developed by Balister, Bollobás and Walters [16], and relies on comparing percolation processes with locally dependent bond percolation on $\mathbb{Z}^2$ (to be more precise: 1-independent bond percolation; see below for a definition). The method of Balister, Bollobás and Walters has proved influential, and has been widely applied to obtain the best rigorous confidence interval estimates for the value of the critical parameter in a wide range of models, see e.g. [12, 14, 16, 15, 17, 19, 21, 39, 40, 109, 102].

However, as noted by the authors of [16] and again by Balister and Bollobás [13] in 2012, locally dependent bond percolation is poorly understood. To quote from the latter work, "*[given that] 1-independent percolation models have become a key tool in establishing bounds on critical probabilities [...], it is perhaps surprising that some of the most basic questions about 1-independent models are open*". In particular, there is no known locally dependent analogue of the Harris–Kesten Theorem, nor even until now much of a sense of what the corresponding 1-independent critical probability ought to be. In Chapter 5, we contribute to the broader project initiated by Balister and Bollobás of addressing the gap in our knowledge about 1-independent bond percolation by making some first steps towards a 1-independent Harris–Kesten Theorem. To state our results and place them in their proper context, we first need to give some definitions.

Let $H = (V, E)$ be a graph. Given a probability measure μ on subsets of E, a *μ-random graph* $\mathbf{H}_\mu$ is a random spanning subgraph of H whose edge-set is chosen randomly from subsets of E according to the law given by μ. Each probability measure μ on subsets of E thus gives rise to a *random graph model* on the host graph H, and we use the two terms (probability measure μ on subsets of E/random graph model $\mathbf{H}_\mu$ on H) interchangeably. We will be interested in random graph models where the state (present/absent) of edges is dependent only on the states of nearby edges. Recall that the *graph distance* between two subsets $A, B \subseteq E$ is the length of the shortest path in H from an endpoint of an edge in A to an endpoint of an edge in B. So in particular if an edge in A shares a vertex with an edge in B, then the graph distance from A to B is zero, while if A and B are supported on disjoint vertex-sets then the graph distance from A to B is at least one.

Definition 1.4.1 (k-independence)**.** *A random graph model $\mathbf{H}_\mu$ on a host graph H is* k-independent *if whenever A, B are disjoint subsets of $E(H)$ such that the graph distance between A and B is at least k, the random variables $E(\mathbf{H}_\mu) \cap A$ and $E(\mathbf{H}_\mu) \cap B$ are independent. If $\mathbf{H}_\mu$ is k-independent, we say that the associated probability measure μ is a* k-independent measure*, or* k-ipm*, on H.*

Let $\mathcal{M}_{k,\geq p}(H)$ denote the collection of all k-independent measures μ on $E(H)$ in which each edge of H is included in $\mathbf{H}_\mu$ with probability at least p. We define $\mathcal{M}_{k,\leq p}(H)$ mutatis mutandis, and let $\mathcal{M}_{k,p}(H)$ denote $\mathcal{M}_{k,\geq p} \cap \mathcal{M}_{k,\leq p}$ — in other words $\mathcal{M}_{k,p}$ is the collection of all k-ipm μ on H in which each edge of H is included in $\mathbf{H}_\mu$ with probability exactly p.

Observe that a 0-independent measure μ is what is known as a *Bernoulli* or *product* measure on E: each edge in E is included in $\mathbf{H}_\mu$ at random independently of all the others. We refer to such measures as *independent measures*. The collection $\mathcal{M}_{0,p}(H)$ thus consists of a single measure, the *p-random measure*, in which each edge of H is included in the associated random graph with probability p, independently of all the other edges. When the host graph H is K_n, the complete graph on n vertices, this gives rise to the celebrated *Erdős–Rényi random graph model*, while when $H = \mathbb{Z}^2$ this is exactly the *independent bond percolation* model considered in the Harris–Kesten Theorem.

We will focus instead on $\mathcal{M}_{1,\geq p}(H)$ and $\mathcal{M}_{1,p}(H)$, whose probability measures allow for some local dependence between the edges. A simple and well-studied example of a model from $\mathcal{M}_{1,p}(H)$ is given by *site percolation*: build a random spanning subgraph $\mathbf{H}_\theta^{\text{site}}$ of H by assigning each vertex $v \in V(H)$ a state S_v independently at random, with $S_v = 1$ with probability θ and $S_v = 0$ otherwise, and including an edge $uv \in E(H)$ in $\mathbf{H}_\theta^{\text{site}}$ if and only if $S_u = S_v = 1$. Each edge of H is in this random graph with probability $p = \theta^2$, and the model is clearly 1-independent since 'randomness resides in the vertices', and so what happens inside two disjoint vertex sets is independent. More generally, any *state-based model* obtained by first assigning independent random states S_v to vertices $v \in V(H)$ and then adding an edge uv according to some deterministic or probabilistic rule depending only on the ordered pair (S_u, S_v) will give rise to a 1-ipm on H. State-based models are a generalisation of the probabilistic notion of a *two-block factor*, see [81] for details.

Given a 1-ipm μ on an infinite connected graph H, we say that μ *percolates* if $\mathbf{H}_\mu$ almost surely (i.e. with probability 1) contains an infinite connected component.

Definition 1.4.2. *Given an infinite connected graph H, we define the* 1-independent critical percolation probability *for H to be*

$$p_{1,c}(H) := \inf \{p \geq 0 : \ \forall \mu \in \mathcal{M}_{1,\geq p}(H), \ \mu \text{ percolates}\}.$$

Remark 1.4.3. *Given $\mu \in \mathcal{M}_{1,\geq p}(H)$ we can obtain a random graph $\mathbf{H}_\nu$ from $\mathbf{H}_\mu$ by deleting each edge uv of $\mathbf{H}_\mu$ independently at random with probability $1 - p/\left(\mathbb{P}[uv \in E(\mathbf{H}_\mu)]\right)$. Clearly $\mathbf{H}_\mu$ stochastically dominates (i.e. is a supergraph of) $\mathbf{H}_\nu$ and $\nu \in \mathcal{M}_{1,p}(H)$. Thus the definition of $p_{1,c}(H)$ above is unchanged if we replace $\mathcal{M}_{1,\geq p}(H)$ by $\mathcal{M}_{1,p}(H)$.*

Remark 1.4.4. *The probability $p_{1,c}(H)$ is in fact one of five natural critical probabilities for* 1*-independent percolation one could consider, all of which are distinct in general — see [36, Section 11.3, Corollary 50 and Question 53].*

Balister, Bollobás and Walters [16] devised a highly effective method for giving rigorous confidence interval results for critical parameters in percolation theory via comparison with 1-independent models on the square integer lattice $\mathbb{Z}^2$. Their method relies on estimating the probability of certain *finite, bounded* events (usually via Monte Carlo methods, whence the confidence intervals) and on bounds on the 1-independent critical probability $p_{1,c}(\mathbb{Z}^2)$. Work of Liggett, Schonman and Stacey [81] on stochastic domination of independent models by 1-independent models implied $p_{1,c}(\mathbb{Z}^2) < 1$. Balister, Bollobás and Walters [16, Theorem 2] obtained the effective upper bound $p_{1,c}(\mathbb{Z}^2) < 0.8639$ via a renormalisation argument and noted *"it would be of interest to give significantly better* *bounds for $p_{1,c}(\mathbb{Z}^2)$; unfortunately, we cannot even hazard a guess as to [its] value".* The question of determining $p_{1,c}(\mathbb{Z}^2)$ was raised again by Balister and Bollobás [13, Question 2], who noted the difficulty of the problem:

Problem 1.4.5 (1-independent Harris–Kesten problem)**.** *Determine $p_{1,c}(\mathbb{Z}^2)$.*

Very recently, Balister, Johnston, Savery and Scott [18] proved the new upper bound $p_{1,c}(\mathbb{Z}^2) \leq 0.8457$.

Balister and Bollobás [13] observed that a simple modification of site percolation due to Newman shows that $p_{1,c}(\mathbb{Z}^2) \geq (\theta_s)^2 + (1-\theta_s)^2$, where $\theta_s = \theta_s(\mathbb{Z}^2)$ is the critical probability for site percolation in $\mathbb{Z}^2$. Since it is known that $\theta_s \in [0.556, 0.679492]$ (see [116, 117]), this shows that $p_{1,c}(\mathbb{Z}^2) \geq 0.5062$. Non-rigorous simulation-based estimates $\theta_s \approx 0.597246$ [119] improve this to a non-rigorous lower bound of 0.5172. Recently, Day, Falgas-Ravry and Hancock gave significant improvements on these lower bounds. In [36, Theorem 7], they constructed measures based on an idea from the PhD thesis [46, Theorem 62] of Falgas-Ravry showing that for any $d \in \mathbb{N}$, $p_{1,c}(\mathbb{Z}^d) \geq 4 - 2\sqrt{3} = 0.5358\ldots$. They in fact showed $p_{1,c}(H) \geq 4 - 2\sqrt{3}$ for any host graph H satisfying what they call the *finite* 2*-percolation property* (see [36, Corollary 24]), a family which includes the graphs $\mathbb{Z}^2 \times K_n$ for any $n \in \mathbb{N}$. Further, the same authors gave a different construction [36, Theorem 8] showing that

$$p_{1,c}(\mathbb{Z}^2) \geq (\theta_s)^2 + \frac{1-\theta_s}{2}, \tag{1.4.1}$$

where $\theta_s = \theta_s(\mathbb{Z}^2)$ is the critical probability for site percolation in $\mathbb{Z}^2$. Using the aforementioned simulation-based estimates for θ_s, this gives a non-rigorous lower bound of 0.5549 on $p_{1,c}(\mathbb{Z}^2)$. Very recently, Balister, Johnston, Savery and Scott [18] proved the new rigorous lower bound $p_{1,c}(\mathbb{Z}^2) \geq \frac{1}{32}(35 - 3\sqrt{3}) = 0.555197\ldots$ and also gave a new non-rigorous simulation-based bound $p_{1,c}(\mathbb{Z}^2) \geq 0.5921$. All these lower bounds remain far apart from the upper bound of 0.8457 from [18], and, as noted in [16], part of the difficulty of Problem 1.4.5 has been the absence of a clear candidate conjecture to aim for.

In view of the difficulty of Problem 1.4.5, there has been interest in increasing our understanding of 1-independent models on other host graphs than $\mathbb{Z}^2$. Balister and Bollobás noted $p_{1,c}(\mathbb{Z}^d)$ is non-increasing in d and must therefore converge to a limit as $d \to \infty$. They showed this limit is at least $1/2$ and posed the following problem [13, Question 2]:

Problem 1.4.6 (Balister and Bollobás problem)**.** *Determine* $\lim_{d\to\infty} p_{1,c}(\mathbb{Z}^d)$.

By the construction of Day, Falgas-Ravry and Hancock mentioned above, this limit is in fact at least $4 - 2\sqrt{3}$; the best known upper bound is 0.5847 and was proved very recently by Balister, Johnston, Savery and Scott [18].

Balister and Bollobás have further studied 1-independent models on infinite trees, obtaining in this setting 1-independent analogues of classical results of Lyons [86] for independent bond percolation. Day, Falgas-Ravry and Hancock for their part gave a number of results on the connectivity of 1-independent random graphs on paths and complete graphs, and on the almost sure emergence of arbitrarily long paths in 1-independent models. More precisely, they introduced the *long paths critical probability* $p_{1,\mathrm{LP}}(H)$ of H, given by

$$p_{1,\mathrm{LP}}(H) := \inf\left\{p \in [0,1] : \ \forall \mu \in \mathcal{M}_{1,p}, \forall \ell \in \mathbb{N}, \ \mathbb{P}\left[\mathbf{H}_\mu \text{ contains a path of length } \ell\right] > 0\right\},$$

and showed $p_{1,\mathrm{LP}}(\mathbb{Z}) = 3/4$, $p_{1,\mathrm{LP}}(\mathbb{Z} \times K_2) = 2/3$. Since the sequence $p_{1,\mathrm{LP}}(\mathbb{Z} \times K_n)$ is non-increasing in n, it tends to a limit in $[0,1]$ as $n \to \infty$. Day, Falgas-Ravry and Hancock showed in [36, Theorem 12(v)] that this limit lies in the interval $[4 - 2\sqrt{3}, 5/9]$ and asked [36, Problem 54]:

Problem 1.4.7 (Day, Falgas–Ravry and Hancock)**.** *Determine* $\lim_{n\to\infty} p_{1,\mathrm{LP}}(\mathbb{Z} \times K_n)$.

In Chapter 5 we determine the limit of the 1-independent critical probability for percolation in $\mathbb{Z}^2 \times K_n$ as $n \to \infty$:

Theorem 1.4.8. *The following hold:*

(i) *If* $p > 4 - 2\sqrt{3}$ *is fixed, then there exists* $N \in \mathbb{N}$ *such that* $p_{1,c}\left(\mathbb{Z}^2 \times K_N\right) \leq p$.

(ii) *For every* $n \in \mathbb{N}$, $p_{1,c}\,(\mathbb{Z}^2 \times K_n) \geq 4 - 2\sqrt{3}$.

In particular, we have $\lim_{n\to\infty} p_{1,c}(\mathbb{Z}^2 \times K_n) = 4 - 2\sqrt{3} = 0.5358\ldots$.

As a corollary to the key result in our proof of Theorem 1.4.8, we also obtain a solution to the problem of Day, Falgas–Ravry and Hancock on long paths in 1-independent percolation, Problem 1.4.7 above:

Theorem 1.4.9. $\lim_{n\to\infty} p_{1,\mathrm{LP}}\,(\mathbb{Z} \times K_n) = 4 - 2\sqrt{3}$.

In fact, we are able to show the conclusions of Theorems 1.4.8 and 1.4.9 still hold if we replace the complete graph K_n by a suitable *pseudorandom graph*. Recall that the study of pseudorandom graphs originates in the ground-breaking work of Thomason [115]. We shall use the following notion of weak pseudorandomness (see Condition (3) in the survey of Krivelevich and Sudakov [76]):

Definition 1.4.10. *Let* $q = q(n)$ *be a sequence in* $[0, 1]$. *A sequence* $(G_n)_{n\in\mathbb{N}}$ *of* n*-vertex graphs is* weakly q-pseudorandom *if*

$$\max\left\{\left|e(G_n[U]) - q\frac{|U|^2}{2}\right| :\ U \subseteq V(G_n)\right\} = o(qn^2).$$

Note that if $(G_n)_{n\in\mathbb{N}}$ is a sequence of weakly q-pseudorandom graphs, then for any $U_1, U_2 \subseteq V(G_n)$ with $U_1 \cap U_2 = \varnothing$, we have

$$e(G_n[U_1, U_2]) = q\,|U_1|\,|U_2| + o(qn^2).$$

Theorem 1.4.11. *Let* $q = q(n)$ *satisfy* $nq(n) \gg \log n$. *Then for any sequence* $(G_n)_{n\in\mathbb{N}}$ *of* n*-vertex graphs which is weakly* q*-pseudorandom, we have* $\lim_{n\to\infty} p_{1,c}(\mathbb{Z}^2 \times G_n) = 4 - 2\sqrt{3}$.

Theorem 1.4.12. *Let* $q = q(n)$ *satisfy* $nq(n) \gg \log n$. *Then for any sequence* $(G_n)_{n\in\mathbb{N}}$ *of* n*-vertex graphs which is weakly* q*-pseudorandom, we have* $\lim_{n\to\infty} p_{1,\mathrm{LP}}(\mathbb{Z} \times G_n) = 4 - 2\sqrt{3}$.

We conjecture that the conclusion of Theorem 1.4.8 still holds if we replace the complete graph K_n by an n-dimensional hypercube.

Conjecture 1.4.13. $\lim_{n\to\infty} p_{1,c}(\mathbb{Z}^2 \times Q_n) = 4 - 2\sqrt{3}$.

Observe Conjecture 1.4.13 implies the answer to the problem of Balister and Bollobás, Problem 1.4.6 above, is $4 - 2\sqrt{3}$. In fact, we make the following bolder conjecture:

Conjecture 1.4.14 (1-independent percolation in high dimension)**.** *There exists $d \geq 3$ such that*

$$p_{1,c}(\mathbb{Z}^d) = 4 - 2\sqrt{3}.$$

Finally we prove some modest results on component evolution in 1-independent models on K_n and on pseudorandom graphs. The main point of these results is that 'the two-state measure minimises the size of the largest component', a heuristic which in turn guides our Conjecture 1.4.13. Here by the *two-state measure*, we mean the following variant of site percolation, due to Newman (see [89]):

Definition 1.4.15 (Two-state measure)**.** *Let H be a graph, and let $p \in [\frac{1}{2}, 1]$. The* two-state measure *$\mu_{2s,p} \in \mathcal{M}_{1,p}(H)$ is constructed as follows: assign to each vertex $v \in V(H)$ a state S_v independently and uniformly at random, with $S_v = 1$ with probability $\theta = \theta(p) = (1 + \sqrt{2p-1})/2$ and $S_v = 0$ otherwise. Then let $\mathbf{H}_{\mu_{2s,p}}$ be the random subgraph of H obtained by including an edge if and only if its endpoints are in the same state.*

Day, Falgas-Ravry and Hancock showed in [36, Theorem 16] that $\mu_{2s,p}$ minimises the probability of connected subgraphs over all 1-ipm $\mu \in \mathcal{M}_{1,p}(K_{2n})$. We show below that it also minimises the probability of having a component of size at least n. Explicitly, given a set of edges $F \subseteq E(H)$ in a graph H, we let $C_i(F)$ denote the i-th largest connected component in the associated subgraph $(V(H), F)$ of H. Then:

Proposition 1.4.16. *Set $p_{2n} = \frac{1}{2}\left(1 - \tan^2\left(\frac{\pi}{4n}\right)\right)$ and $H = K_{2n}$. Then for all $p \in [p_{2n}, 1]$,*

$$\min\Big\{\mathbb{P}\left[|C_1(\mathbf{H}_\mu)| > n\right] : \ \mu \in \mathcal{M}_{1,\geq p}(K_{2n})\Big\} = 1 - \binom{2n}{n}\left(\frac{1-p}{2}\right)^n.$$

Further, we show that the two-state measure also asymptotically minimises the likely size of a largest component in 1-independent models on pseudorandom graphs:

Theorem 1.4.17. *Let $r \in \mathbb{N}$, and let $p \in (\frac{1}{r+1}, \frac{1}{r}]$ be fixed. Let $(H_n)_{n\in\mathbb{N}}$ be a sequence of weakly q-pseudorandom graphs on n vertices with $q = q(n) \gg \log(n)/n$. Then the following hold for $H = H_n$:*

(i) *For every $\mu \in \mathcal{M}_{1,p}(H)$, with probability $1 - o(1)$ we have*

$$|C_1(\mathbf{H}_\mu)| \geq (1 - o(1)) \frac{1 + \sqrt{\frac{(r+1)p-1}{r}}}{r + 1} n.$$

(ii) *There exists $\mu \in \mathcal{M}_{1,p}(H)$ such that with probability $1 - o(1)$ the random graph $\mathbf{H}_\mu$ satisfies $|C_1(\mathbf{H}_\mu)| \leq (1 + o(1)) \frac{1+\sqrt{\frac{(r+1)p-1}{r}}}{r+1} n$.*

This leads us to the natural conjecture that the two-state measure asymptotically minimises the size of a largest component in 1-independent models on the hypercube Q_n:

Conjecture 1.4.18. *Let $p \in (\frac{1}{2}, 1]$ be fixed, and let $H = Q_n$. Then for all $\mu \in \mathcal{M}_{1,\geq p}(Q_n)$, with probability $1 - o(1)$ we have $|C_1(\mathbf{H}_\mu)| \geq \left(\frac{1+\sqrt{2p-1}}{2} - o(1)\right) 2^n$.*

We suspect a proof of this conjecture combined with the ideas presented in this chapter would yield a proof of Conjecture 1.4.13.

Overall, our results would lead us to speculate that the true value of $p_{1,c}(\mathbb{Z}^2)$ is probably a lot closer to the non-rigorous lower bound of 0.5921 than to the upper bound of 0.8457 (both obtained in [18]). However a rigorous proof of improved upper bounds on $p_{1,c}(\mathbb{Z}^2)$ remains elusive for the time being.

CHAPTER 2

TOWARDS LEHEL'S CONJECTURE FOR 4-UNIFORM TIGHT CYCLES

The main aim of this chapter is to prove Theorems 1.1.2 and 1.1.3. We first recall some definitions and then restate these theorems. Recall that an *r-edge-colouring of a k-graph H* is a colouring of the edges of H with r colours. Also recall that a *monochromatic subgraph* of an r-edge-coloured graph is one in which all the edges have the same colour. Moreover, recall that a *k-uniform tight cycle* is a k-graph with a cyclic ordering of its vertices such that its edges are precisely all k-sets of consecutive vertices in the ordering. We now restate Theorems 1.1.2 and 1.1.3.

Theorem 1.1.2. *For every $\varepsilon > 0$, there exists an integer n_1 such that, for all $n \geq n_1$, every 2-edge-coloured complete 4-graph on n vertices contains two vertex-disjoint monochromatic tight cycles of distinct colours covering all but at most εn of the vertices.*

Theorem 1.1.3. *For every $\varepsilon > 0$, there exists an integer n_1 such that, for all $n \geq n_1$, every 2-edge-coloured complete 5-graph on n vertices contains four vertex-disjoint monochromatic tight cycles covering all but at most εn of the vertices.*

To prove Theorems 1.1.2 and 1.1.3, we use the *connected matching method* that has often been credited to Łuczak [84]. We now present a sketch-proof for Theorem 1.1.2. Consider a 2-edge-coloured complete 4-graph $K_n^{(4)}$ on n vertices. We start by applying the Hypergraph Regularity Lemma to the 2-edge-coloured complete 4-graph $K_n^{(4)}$. More precisely the

Regular Slice Lemma of Allen, Böttcher, Cooley and Mycroft [2], see Lemma 2.3.3. We obtain a 2-edge-coloured reduced graph $\mathcal{R}$ that is almost complete. A monochromatic matching in a k-graph is a set of vertex-disjoint edges of the same colour. We say that it is tightly connected if, for any two edges f and f', there exists a sequence of edges $e_1, \dots, e_t$ of the same colour such that $e_1 = f$, $e_t = f'$ and $|e_i \cap e_{i+1}| = k - 1$ for all $i \in [t-1]$. Using Corollary 2.3.12, it suffices to find two vertex-disjoint monochromatic tightly connected matchings of distinct colours in the reduced graph $\mathcal{R}$. The main challenge is to identify the 'tightly connected components' (see Section 2.1 for the formal definition) in which we will find the matchings. To do so, we introduce the concept of 'blueprint', which is a 2-edge-coloured 2-graph with the same vertex set as $\mathcal{R}$. The key property is that connected components in the blueprint correspond to tightly connected components in $\mathcal{R}$.

The remainder of this chapter is organised as follows. In Section 2.1, we introduce some basic notation and definitions. In Section 2.2, we prove Proposition 1.1.1. In Section 2.3, we introduce the statements about hypergraph regularity and prove the crucial Corollary 2.3.12 that allows us to reduce our problem of finding cycles in the complete graph to one about finding tightly connected matchings in the reduced graph. In Section 2.4, we give the definition of blueprint and setup some useful results. In Sections 2.5 and 2.6, we prove Theorems 1.1.2 and 1.1.3, respectively. Finally, we make some concluding remarks in Section 2.7.

2.1 Preliminaries

If we say that a statement holds for $0 < a \ll b \leq 1$, then we mean that there exists a non-decreasing function $f\colon (0,1] \to (0,1]$ such that the statement holds for all $a, b \in (0,1]$ with $a \leq f(b)$. Similar expressions with more variables are defined analogously. If $1/n$ appears in one of these expressions, then we implicitly assume that n is a positive integer. We omit floors and ceilings whenever doing so does not affect the argument.

We often write $x_1 \dots x_j$ for the set $\{x_1, \dots, x_j\}$. Moreover, for each positive integer n, we let $[n] = \{1, \dots, n\}$. For a set S and a non-negative integer k we denote by $\binom{S}{k}$ the set of all subsets of S of size k.

A *k-graph* H is a pair of sets $(V(H), E(H))$ such that $E(H) \subseteq \binom{V(H)}{k}$. We let $v(H) = |V(H)|$ be the number of vertices of H. We abuse notation by identifying the k-graph H with its edge set $E(H)$. Hence by $|H|$ we mean the number of edges of H.

A 2-edge-coloured k-graph is a k-graph together with a colouring of its edges with the colours red and blue. For a 2-edge-coloured k-graph H we denote by H^{red} and H^{blue} the k-graph on $V(H)$ induced by the red and the blue edges of H, respectively. A subgraph of a 2-edge-coloured k-graph is called *monochromatic* if all its edges have the same colour.

Let H be a 2-edge-coloured k-graph. Two edges f and f' in H are *tightly connected* if there exists a sequence of edges $e_1, \dots, e_t$ such that $e_1 = f$, $e_t = f'$ and $|e_i \cap e_{i+1}| = k-1$ for all $i \in [t-1]$. A subgraph H' of H is *tightly connected* if every pair of edges in H' is tightly connected in H. A maximal tightly connected subgraph of H is called a *tight component* of H. Note that a tight component is a subgraph rather than a vertex subset as in the traditional graph case. In a 2-graph G, we simply call a tight component a *component* and a *spanning component* is one that covers all the vertices of G. A *tightly connected matching* in a k-graph H is a matching contained in a tight component of H. A *red tight component* and a *red tightly connected matching* are a tight component and a tightly connected matching in H^{red}, respectively. We define these terms similarly for *blue*.

Let H be a k-graph and $S, W \subseteq V(H)$. We denote by $H - W$ the k-graph with $V(H-W) = V(H) \setminus W$ and $E(H-W) = \{e \in E(H) \colon e \cap W = \varnothing\}$. We call $H - W$ the *k-graph obtained from H by deleting W*. Further we let $H[W] = H - (V(H) \setminus W)$. Let F be a k-graph or a set of k-element sets. We denote by $H - F$ the subgraph of H obtained by deleting the edges in F. We define $N_H(S, W)$ to be the set $\{e \in \binom{W}{k-|S|} \colon e \cup S \in H\}$ and we define $d_H(S, W)$ to be its cardinality. Further we write $N_H(S)$ and $d_H(S)$ for $N_H(S, V(H))$ and $d_H(S, V(H))$, respectively. If H is 2-edge-coloured, then we write $N_H^{\mathrm{red}}(S, W)$, $d_H^{\mathrm{red}}(S, W)$, $N_H^{\mathrm{blue}}(S, W)$, $d_H^{\mathrm{blue}}(S, W)$ for $N_{H^{\mathrm{red}}}(S, W)$, $d_{H^{\mathrm{red}}}(S, W)$,

$N_{H^{\text{blue}}}(S,W)$, $d_{H^{\text{blue}}}(S,W)$, respectively. The *link graph of H with respect to S*, denoted by H_S, is the $(k-|S|)$-graph satisfying $V(H_S) = V(H) \setminus S$ and $E(H_S) = N_H(S)$.

For $j \in [k-1]$, the *j-th shadow of H*, denoted by $\partial^j H$, is the $(k-j)$-graph with vertex set $V(\partial^j H) = V(H)$ and edge set

$$E(\partial^j H) = \left\{ e \in \binom{V(H)}{k-j} : e \subseteq f \text{ for some } f \in E(H) \right\}.$$

For the 1-st shadow of H, we also simply write ∂H instead of $\partial^1 H$.

For $\mu, \alpha > 0$, we say that a k-graph H on n vertices is *(μ,α)-dense* if, for each $i \in [k-1]$, we have $d_H(S) \geq \mu\binom{n}{k-i}$ for all but at most $\alpha\binom{n}{i}$ sets $S \in \binom{V(H)}{i}$ and $d_H(S) = 0$ for all other $S \in \binom{V(H)}{i}$.

Proposition 2.1.1. *Let $0 \leq \alpha, \mu \leq 1$ and let H be a (μ,α)-dense k-graph on n vertices. Then $|H| \geq (\mu-\alpha)\binom{n}{k}$. Moreover, if $\mu > 1/2$, then H is tightly connected.*

Proof. Note that

$$|H| = \frac{1}{k} \sum_{S \in \binom{V(H)}{k-1}} d_H(S) \geq \frac{1}{k}(1-\alpha)\binom{n}{k-1}\mu n \geq (\mu-\alpha)\binom{n}{k}.$$

Now suppose that $\mu > 1/2$. We show that H is tightly connected. Note that, for $S, S' \in \binom{V(H)}{k-1}$ with $d_H(S), d_H(S') > 0$, we have $d_H(S), d_H(S') \geq \mu n > n/2$ and thus

$$N_H(S) \cap N_H(S') \neq \varnothing.$$

Let $f = x_1 \dots x_k$ and $f' = y_1 \dots y_k$ be two edges of H. Inductively choose vertices $z_1, \dots, z_{k-1} \in V(H)$ such that

$$z_i \in N_H(z_1 \dots z_{i-1} x_{i+1} \dots x_k) \cap N_H(z_1 \dots z_{i-1} y_{i+1} \dots y_k)$$

for all $i \in [k-1]$. It follows that f and f' are tightly connected. □

The following proposition shows that any k-graph that has all but a small fraction of the possible edges contains a $(1-\varepsilon,\alpha)$-dense subgraph. The proof was inspired by the proof of Lemma 8.8 in [63]. A different generalisation of this lemma can also be found as Lemma 2.3 in [78].

Proposition 2.1.2. *Let $1/n \ll \alpha \ll 1/k \leq 1/2$. Let H be a k-graph on n vertices with $|H| \geq (1-\alpha)\binom{n}{k}$. Then there exists a subgraph H' of H such that $V(H')=V(H)$ and H' is $(1-2\alpha^{1/4k^2}, 2\alpha^{1/4k^2})$-dense.*

Proof. We call a set $S \subseteq V(H)$ with $|S| \in [k-1]$ *bad* if $d_H(S) < (1-\alpha^{1/2})\binom{n}{k-|S|}$. For $i \in [k-1]$, let $\mathcal{B}_i$ be the set of all bad i-sets. For each $i \in [k-1]$, we have

$$(1-\alpha)\binom{k}{i}\binom{n}{k} \leq \binom{k}{i}|H| = \sum_{S \in \binom{V(H)}{i}} d_H(S) \leq \binom{n}{i}\binom{n}{k-i} - \alpha^{1/2}\binom{n}{k-i}|\mathcal{B}_i|.$$

This implies

$$|\mathcal{B}_i| \leq \frac{1}{\alpha^{1/2}}\left(\binom{n}{i} - \frac{(1-\alpha)\binom{k}{i}\binom{n}{k}}{\binom{n}{k-i}}\right) \leq 2\alpha^{1/2}\binom{n}{i}.$$

Let $\beta = \alpha^{1/2k}$. For all $j \in \{k-1, k-2, \dots, 1\}$ in turn, we construct $\mathcal{A}_j \subseteq \binom{V(H)}{j}$ inductively as follows. We set $\mathcal{A}_{k-1} = \mathcal{B}_{k-1}$. Given $2 \leq j \leq k-1$ and $\mathcal{A}_j$, we define $\mathcal{A}_{j-1} \subseteq \binom{V(H)}{j-1}$ to be the set of all $X \in \binom{V(H)}{j-1}$ such that $X \in \mathcal{B}_{j-1}$ or $d_{\mathcal{A}_j}(X) \geq \beta^{1/2}n$.

Claim 2.1.3. *For all $i \in [k-1]$, $|\mathcal{A}_i| \leq \beta^i\binom{n}{i}$. Moreover, if $1 \leq i < j \leq k-1$ and a set $S \in \binom{V(H)}{i}$ satisfies $d_{\mathcal{A}_j}(S) \geq \beta^{1/2(j-i)}\binom{n}{j-i}$, then $S \in \mathcal{A}_i$.*

Proof of Claim. We first prove the first part by induction on $k-i$. For $i=k-1$, we have $|\mathcal{A}_{k-1}| = |\mathcal{B}_{k-1}| \leq 2\alpha^{1/2}\binom{n}{k-1} \leq \beta^{k-1}\binom{n}{k-1}$.

Now suppose $2 \leq i \leq k-1$ and $|\mathcal{A}_i| \leq \beta^i\binom{n}{i}$. By double counting tuples (X,w) with $X \in \mathcal{A}_{i-1} \setminus \mathcal{B}_{i-1}$ and $X \cup w \in \mathcal{A}_i$, we have $(|\mathcal{A}_{i-1}| - |\mathcal{B}_{i-1}|)\,\beta^{1/2}n \leq i\,|\mathcal{A}_i|$. Hence

$$\begin{aligned}|\mathcal{A}_{i-1}| &\leq \frac{i}{\beta^{1/2}n}|\mathcal{A}_i| + |\mathcal{B}_{i-1}| \leq \frac{i}{\beta^{1/2}n}\beta^i\binom{n}{i} + 2\alpha^{1/2}\binom{n}{i-1} \\ &= \beta^{i-1/2}\binom{n-1}{i-1} + 2\alpha^{1/2}\binom{n}{i-1} \leq \beta^{i-1}\binom{n}{i-1}.\end{aligned}$$

This proves the first part of the claim.

We now prove the second part of the claim. Fix $i \in [k-1]$. We proceed by induction on $j-i$. For $j = i+1$, the statement holds by the definition of $\mathcal{A}_i$. Now let $S \in \binom{V(H)}{i}$ and $j \geq i+2$ be such that $d_{\mathcal{A}_j}(S) \geq \beta^{1/2(j-i)}\binom{n}{j-i}$. If $S \in \mathcal{B}_i$, then $S \in \mathcal{A}_i$. Recall that if $T \in \binom{V(H)}{j-1} \setminus \mathcal{A}_{j-1}$, then $d_{\mathcal{A}_j}(T) < \beta^{1/2}n$. We have

$$\begin{aligned}\beta^{1/2(j-i)}\binom{n}{j-i} \leq d_{\mathcal{A}_j}(S) &\leq \sum_{\substack{T \in \mathcal{A}_{j-1}\\ S \subseteq T}} d_{\mathcal{A}_j}(T) + \sum_{\substack{T \in \binom{V(H)}{j-1} \setminus \mathcal{A}_{j-1}\\ S \subseteq T}} d_{\mathcal{A}_j}(T)\\ &\leq n d_{\mathcal{A}_{j-1}}(S) + \beta^{1/2} n d_{\binom{V(H)}{j-1} \setminus \mathcal{A}_{j-1}}(S)\\ &\leq n d_{\mathcal{A}_{j-1}}(S) + \beta^{1/2} n \binom{n}{j-i-1},\end{aligned}$$

and thus

$$d_{\mathcal{A}_{j-1}}(S) \geq \beta^{1/2(j-i-1)}\binom{n}{j-i-1}.$$

Hence by the induction hypothesis we have $S \in \mathcal{A}_i$. ■

For each $j \in [k-1]$, let F_j be the set of edges $e \in H$ for which there exists some $S \in \mathcal{A}_j$ with $S \subseteq e$. Let $F = \bigcup_{j \in [k-1]} F_j$ and let $H' = H - F$. We will show that H' is the desired k-graph. For $i \in [k-1]$, let $\mathcal{S}_i$ be the set of all $S \in \binom{V(H)}{i}$ such that $d_F(S) \geq \beta^{1/2k}\binom{n}{k-i}$.

Claim 2.1.4. *For $i \in [k-1]$, $|\mathcal{S}_i| \leq \beta^{1/2}\binom{n}{i}$.*

Proof of Claim. For $j \in [k-1]$, we have

$$|F_j| \leq |\mathcal{A}_j|\binom{n-j}{k-j} \overset{\text{Claim 2.1.3}}{\leq} \beta^j \binom{n}{j}\binom{n-j}{k-j} = \beta^j \binom{k}{j}\binom{n}{k}.$$

Thus

$$|F| \leq \sum_{j \in [k-1]} |F_j| \leq \sum_{j \in [k-1]} \beta^j \binom{k}{j}\binom{n}{k} \leq 2^k \beta \binom{n}{k}.$$

Now, for $i \in [k-1]$, we have

$$\frac{|\mathcal{S}_i|\,\beta^{1/2k}\binom{n}{k-i}}{\binom{k}{i}} \le |F| \le 2^k\beta\binom{n}{k}$$

and thus $|\mathcal{S}_i| \le \beta^{1/2}\binom{n}{i}$. ■

Consider $i \in [k-1]$. Note that $|\mathcal{S}_i \cup \mathcal{B}_i| \le 2\alpha^{1/4k^2}\binom{n}{i}$. Now let $S \in \binom{V(H)}{i} \setminus (\mathcal{S}_i \cup \mathcal{B}_i)$. As $S \notin \mathcal{B}_i$, we have $d_H(S) \ge (1-\alpha^{1/2})\binom{n}{k-i}$. As $S \notin \mathcal{S}_i$, we have

$$\begin{aligned} d_{H'}(S) &= d_H(S) - d_F(S) \ge d_H(S) - \beta^{1/2k}\binom{n}{k-i} \\ &\ge (1-\alpha^{1/2}-\beta^{1/2k})\binom{n}{k-i} \ge (1-2\alpha^{1/4k^2})\binom{n}{k-i}. \end{aligned}$$

Consider $X \in \binom{V(H)}{i}$ with $d_{H'}(X) \ne 0$. We want to show that $d_{H'}(X) \ge (1-2\alpha^{1/4k^2})\binom{n}{k-i}$. By the above, it suffices to show that $X \notin \mathcal{B}_i \cup \mathcal{S}_i$. Let $e \in H'$ with $X \subseteq e$. Since $e \notin F_i$, $X \notin \mathcal{A}_i$ and thus $X \notin \mathcal{B}_i$. It remains for us to show that $X \notin \mathcal{S}_i$. Assume the contrary that X is contained in more that $\beta^{1/2k}\binom{n}{k-i}$ edges of F. Let $\mathcal{Y} = N_F(X)$, so $|\mathcal{Y}| \ge \beta^{1/2k}\binom{n}{k-i}$. For each $Y \in \mathcal{Y}$, fix a set $A_Y \in \bigcup_{j\in[k-1]}\mathcal{A}_j$ such that $A_Y \subseteq X \cup Y$ and let $T_Y = X \cap A_Y$ and $S_Y = Y \setminus A_Y$. If $A_Y \subseteq X$, then $A_Y \subseteq e \in H'$, a contradiction. Hence $A_Y \cap Y \ne \varnothing$ for all $Y \in \mathcal{Y}$. For $Y \in \mathcal{Y}$, we have $X \cap Y = \varnothing$, and thus $|T_Y| \le |A_Y| - 1 \le k-2$. By an averaging argument, there exist $t \in \{0,1,\dots,k-2\}$, $T \in \binom{X}{t}$, $a \in [k-1]$, $S \in \binom{V(H)}{k-i-a+t}$ and $\widetilde{\mathcal{Y}} \subseteq \mathcal{Y}$ such that, for all $Y \in \widetilde{\mathcal{Y}}$, we have $T_Y = T$, $|A_Y| = a$, $S_Y = S$ and

$$\left|\widetilde{\mathcal{Y}}\right| \ge \frac{|\mathcal{Y}|}{2^i(k-1)\binom{n}{k-i-a+t}} \ge \beta^{1/2(k-1)}\binom{n}{a-t}.$$

Since $Y \setminus A_Y = S_Y = S$ and $|A_Y| = a$ for all $Y \in \widetilde{\mathcal{Y}}$, the A_Y are distinct for all $Y \in \widetilde{\mathcal{Y}}$. Recall that $T \subseteq A_Y \in \mathcal{A}_a$ for each $Y \in \widetilde{\mathcal{Y}}$. If $T = \varnothing$, then $t = 0$ and so $|\mathcal{A}_a| \ge \left|\widetilde{\mathcal{Y}}\right| > \beta^a\binom{n}{a}$ contradicting Claim 2.1.3. If $T \ne \varnothing$, then we have $d_{\mathcal{A}_a}(T) \ge \left|\widetilde{\mathcal{Y}}\right| \ge \beta^{1/2(k-1)}\binom{n}{a-t}$. Claim 2.1.3 implies that $T \in \mathcal{A}_t$. Since $T \subseteq X \subseteq e$, we have $e \in F_t$ contradicting the fact that $e \in H' = H - \bigcup_{j \in [k-1]} F_j$. $\square$

2.2 Extremal example

In this section, we prove Proposition 1.1.1, that is, we prove that, for $k \geq 3$, there exist arbitrarily large 2-edge-coloured complete k-graphs that do not admit a partition into two tight cycles of distinct colours.

A *k-uniform tight path* is a k-graph obtained by deleting a vertex from a tight cycle. First we need the following proposition.

Proposition 2.2.1. *Let $k \geq 3$, let P and C be a k-uniform tight path and tight cycle, respectively. We have the following.*

(i) *If X and Y partition $V(P)$ such that $|e \cap Y| \geq 2$ for all $e \in P$, then $2(|X| - (k-1)) \leq (k-2)|Y|$.*

(ii) *If X and Y partition $V(C)$ such that $|e \cap Y| \geq 2$ for all $e \in C$, then $2|X| \leq (k-2)|Y|$.*

Proof. We first prove (i). Let M be a matching of maximum size in P. Since each edge of P contains at least 2 vertices of Y,

$$|X| \leq |X \cap V(M)| + |V(P) \setminus V(M)| \leq (k-2)|M| + k - 1 \leq \frac{(k-2)|Y|}{2} + k - 1.$$

Now we prove (ii). Since $|e \cap Y| \geq 2$ and $|e \cap X| \leq k-2$ for each edge $e \in C$, we have

$$|X| = \frac{1}{k} \sum_{e \in C} |e \cap X| = \frac{1}{k} \sum_{e \in C} \frac{|e \cap X|}{|e \cap Y|} |e \cap Y| \leq \frac{1}{k} \sum_{e \in C} \frac{k-2}{2} |e \cap Y| = \frac{k-2}{2} |Y|.$$

$\square$

We are now ready to give our extremal example. Note that the case $k = 3$ of the extremal example is already given in [53]. Recall that, in a k-graph, we consider a single edge and any set of fewer than k vertices to be degenerate cycles.

Proof of Proposition 1.1.1. Let $k \geq 3$, $m \geq k+1$ and $n = k(m+1)+1$. Let X, Y and $\{z\}$ be three disjoint vertex sets of $K_n^{(k)}$ of sizes $(k-1)m+k-2$, $m+2$ and 1, respectively. We colour an edge e in $K_n^{(k)}$ red if $z \in e$ and $|e \cap Y| \geq 2$ or $z \notin e$ and $|e \cap Y| = 1$. Otherwise we colour it blue. Note that $K_n^{(k)} - z$ has the following 3 monochromatic tight components:

$$B_1 = \binom{X}{k}, \ B_2 = \left\{ e \in \binom{X \cup Y}{k} : \ |e \cap Y| \geq 2 \right\}, \ R = \left\{ e \in \binom{X \cup Y}{k} : \ |e \cap Y| = 1 \right\}.$$

Note that B_1 and B_2 are blue and R is red. Suppose for a contradiction that $K_n^{(k)}$ can be partitioned into a red tight cycle C_R and a blue tight cycle C_B.

First assume $z \in V(C_R)$. Since all the red edges containing z are in a red tight component disjoint from R, we have $|V(C_R)| \leq k$. Hence $|V(C_B)| = n - |V(C_R)| \geq n - k \geq km > k$ and $|V(C_B) \cap Y| = |Y \setminus V(C_R)| \geq m + 2 - (k-1) \geq 1$. So C_B is not degenerate and $C_B \subseteq B_2$. Any edge $e \in C_B$ contains at least 2 vertices in Y. By Proposition 2.2.1(ii), $2\,|V(C_B) \cap X| \leq (k-2)\,|V(C_B) \cap Y|$. It follows that

$$\begin{aligned} 2(k-1)m - 2 = 2(|X| - (k-1)) &\leq 2\,|V(C_B) \cap X| \\ &\leq (k-2)\,|V(C_B) \cap Y| \leq (k-2)\,|Y| = (k-2)(m+2). \end{aligned}$$

This implies that $m \leq 2$, a contradiction.

Hence, we may assume that $z \in V(C_B)$. This implies that $C_R \subseteq R$ or $|V(C_R)| \leq k-1$. Let $x_R = |V(C_R) \cap X|$, $y_R = |V(C_R) \cap Y|$, $x_B = |V(C_B) \cap X|$ and $y_B = |V(C_B) \cap Y|$. Let P_B be the tight path $C_B - z$. Clearly $|V(P_B) \cap X| = x_B$ and $|V(P_B) \cap Y| = y_B$. Since $C_R \subseteq R$ or $|V(C_R)| \leq k-1$,

$$y_R \leq \max\left\{ \left\lfloor \frac{|X|}{k-1} \right\rfloor, k-1 \right\} = m < |Y|. \tag{2.2.1}$$

Hence, $V(P_B) \cap Y \neq \varnothing$ and $|V(P_B)| \geq (n-1) - km \geq k$. We must have $P_B \subseteq B_2$. By

Proposition 2.2.1(i), we have that

$$2(x_B - (k-1)) \leq (k-2)y_B. \tag{2.2.2}$$

Thus

$$\begin{aligned}|V(P_B)| &= x_B + y_B \leq \frac{k}{2} y_B + k - 1 \leq \frac{k}{2}|Y| + k - 1 = \frac{k}{2}(m+2) + k - 1 \\ &\leq mk = n - 1 - k.\end{aligned}$$

This implies that $|V(C_R)| \geq k$. Hence $C_R \subseteq R$ and thus

$$x_R = (k-1)y_R. \tag{2.2.3}$$

Since $x_R + x_B = |X| = (k-1)m + k - 2$ and $y_R + y_B = |Y| = m + 2$, (2.2.2) implies

$$\begin{aligned}(k-2)(m+2-y_R) &\geq 2(|X| - x_R - (k-1)) \\ &= 2((k-1)m + k - 2 - (k-1)y_R - (k-1)),\end{aligned}$$

which implies $y_R \geq m - 1$. If $y_R = m - 1$, then (2.2.3) implies that $x_R = (k-1)(m-1)$ and thus $x_B = 2k - 3$ and $y_B = 3$. Let $P_B = v_1 \dots v_{2k}$. Either the edge $v_1 \dots v_k$ or the edge $v_{k+1} \dots v_{2k}$ contains at most one vertex of Y, a contradiction to $P_B \subseteq B_2$. Thus we may assume $y_R \geq m$ and since $y_R \leq m$ by (2.2.1), we have $y_R = m$. By (2.2.3), we have $x_R = (k-1)m$ and thus $x_B = k - 2$ and $y_B = 2$. Moreover, C_B is a copy of $K_{k+1}^{(k)}$ that has a blue edge containing z and at least two vertices of Y, a contradiction. □

2.3 Hypergraph regularity

In this section, we follow the notation of Allen, Böttcher, Cooley and Mycroft [2]. A *hypergraph* $\mathcal{H}$ is an ordered pair $(V(\mathcal{H}), E(\mathcal{H}))$, where $E(\mathcal{H}) \subseteq 2^{V(\mathcal{H})}$. Again, we identify the hypergraph $\mathcal{H}$ with its edge set $E(\mathcal{H})$. A subgraph $\mathcal{H}'$ of $\mathcal{H}$ is a hypergraph with $V(\mathcal{H}') \subseteq V(\mathcal{H})$ and $E(\mathcal{H}') \subseteq E(\mathcal{H})$. It is *spanning* if $V(\mathcal{H}') = V(\mathcal{H})$. For $U \subseteq V(\mathcal{H})$, we define $\mathcal{H}[U]$ to be the subgraph of $\mathcal{H}$ with $V(\mathcal{H}[U]) = U$ and $E(\mathcal{H}[U]) = \{e \in E(\mathcal{H}) \colon e \subseteq U\}$. We call $\mathcal{H}$ a *complex* if $\mathcal{H}$ is down-closed, that is if $e \in \mathcal{H}$ and $f \subseteq e$, then $f \in \mathcal{H}$. A k*-complex* is a complex with only edges of size at most k. We denote by $\mathcal{H}^{(i)}$ the spanning subgraph of $\mathcal{H}$ containing only the edges of size i. Let $\mathcal{P}$ be a partition of $V(\mathcal{H})$ into parts $V_1, \ldots, V_s$. Then we say that a set $S \subseteq V(\mathcal{H})$ is $\mathcal{P}$*-partite* if $|S \cap V_i| \leq 1$ for all $i \in [s]$. For $\mathcal{P}' = \{V_{i_1}, \ldots, V_{i_r}\} \subseteq \mathcal{P}$, we define the subgraph of $\mathcal{H}$ induced by $\mathcal{P}'$, denoted by $\mathcal{H}[\mathcal{P}']$ or $\mathcal{H}[V_{i_1}, \ldots, V_{i_r}]$, to be the subgraph of $\mathcal{H}[\bigcup \mathcal{P}']$ containing only the edges that are $\mathcal{P}'$-partite. The hypergraph $\mathcal{H}$ is $\mathcal{P}$-partite if all of its edges are $\mathcal{P}$-partite. In this case we call the parts of $\mathcal{P}$ the *vertex classes* of $\mathcal{H}$. We say that $\mathcal{H}$ is s*-partite* if it is $\mathcal{P}$-partite for some partition $\mathcal{P}$ of $V(\mathcal{H})$ into s parts. Let $\mathcal{H}$ be a $\mathcal{P}$-partite hypergraph. If X is a k-set of vertex classes of $\mathcal{H}$, then we write $\mathcal{H}_X$ for the k-partite subgraph of $\mathcal{H}^{(k)}$ induced by $\bigcup X$, whose vertex classes are the elements of X. Moreover, we denote by $\mathcal{H}_{X^<}$ the k-partite hypergraph with $V(\mathcal{H}_{X^<}) = \bigcup X$ and $E(\mathcal{H}_{X^<}) = \bigcup_{X' \subsetneq X} \mathcal{H}_{X'}$. In particular, if $\mathcal{H}$ is a complex, then $\mathcal{H}_{X^<}$ is a $(k-1)$-complex because X is a set of size k.

Let $i \geq 2$, and let $\mathcal{P}_i$ be a partition of a vertex set V into i parts. Let H_i and H_{i-1} be a $\mathcal{P}_i$-partite i-graph and a $\mathcal{P}_i$-partite $(i-1)$-graph on a common vertex set V, respectively. We say that a $\mathcal{P}_i$-partite i-set in V is *supported on* H_{i-1} if it induces a copy of the complete $(i-1)$-graph $K_i^{(i-1)}$ on i vertices in H_{i-1}. We denote by $K_i(H_{i-1})$ the $\mathcal{P}_i$-partite i-graph on V whose edges are all $\mathcal{P}_i$-partite i-sets contained in V which are supported on H_{i-1}. Now we define the *density of* H_i *with respect to* H_{i-1} to be

$$d(H_i \mid H_{i-1}) = \frac{|K_i(H_{i-1}) \cap H_i|}{|K_i(H_{i-1})|}$$

if $|K_i(H_{i-1})| > 0$ and $d(H_i \mid H_{i-1}) = 0$ if $|K_i(H_{i-1})| = 0$. So $d(H_i \mid H_{i-1})$ is the proportion of $\mathcal{P}_i$-partite copies of K_i^{i-1} in H_{i-1} which are also edges of H_i. More generally, if $\mathbf{Q} = (Q_1, Q_2, \ldots, Q_r)$ is a collection of r (not necessarily disjoint) subgraphs of H_{i-1}, we define $K_i(\mathbf{Q}) = \bigcup_{j=1}^r K_i(Q_j)$ and

$$d(H_i \mid \mathbf{Q}) = \frac{|K_i(\mathbf{Q}) \cap H_i|}{|K_i(\mathbf{Q})|}$$

if $|K_i(\mathbf{Q})| > 0$ and $d(H_i \mid \mathbf{Q}) = 0$ if $|K_i(\mathbf{Q})| = 0$. We say that H_i is (d_i, ε, r)-*regular with respect to* H_{i-1}, if we have $d(H_i \mid \mathbf{Q}) = d_i \pm \varepsilon$ for every r-set $\mathbf{Q}$ of subgraphs of H_{i-1} with $|K_i(\mathbf{Q})| > \varepsilon\, |K_i(H_{i-1})|$. We say that H_i is (ε, r)-*regular with respect to* H_{i-1} if there exists some d_i for which H_i is (d_i, ε, r)-regular with respect to H_{i-1}. Finally, given an i-graph G whose vertex set contains that of H_{i-1}, we say that G is (d_i, ε, r)-*regular with respect to* H_{i-1} if the i-partite subgraph of G induced by the vertex classes of H_{i-1} is (d_i, ε, r)-regular with respect to H_{i-1}. We refer to the density of this i-partite subgraph of G with respect to H_{i-1} as the *relative density of* G *with respect to* H_{i-1}.

Now let $s \geq k \geq 3$ and let $\mathcal{H}$ be an s-partite k-complex on vertex classes $V_1, \dots, V_s$. For any set $A \subseteq [s]$, we write V_A for $\bigcup_{i \in A} V_i$. Note that, if $e \in \mathcal{H}^{(i)}$ for some $2 \leq i \leq k$, then the vertices of e induce a copy of K_i^{i-1} in $\mathcal{H}^{(i-1)}$. Therefore, for any set $A \in \binom{[s]}{i}$, the density $d(\mathcal{H}^{(i)}[V_A] \mid \mathcal{H}^{(i-1)}[V_A])$ is the proportion of 'possible edges' of $\mathcal{H}^{(i)}[V_A]$, which are indeed edges. We say that $\mathcal{H}$ is $(d_k, \dots, d_2, \varepsilon_k, \varepsilon, r)$-*regular* if

(a) for any $2 \leq i \leq k-1$ and any $A \in \binom{[s]}{i}$, the induced subgraph $\mathcal{H}^{(i)}[V_A]$ is $(d_i, \varepsilon, 1)$-regular with respect to $\mathcal{H}^{(i-1)}[V_A]$, and

(b) for any $A \in \binom{[s]}{k}$, the induced subgraph $\mathcal{H}^{(k)}[V_A]$ is (d_k, ε_k, r)-regular with respect to $\mathcal{H}^{(k-1)}[V_A]$.

For a $(k-1)$-tuple $\mathbf{d} = (d_k, \dots, d_2)$, we write $(\mathbf{d}, \varepsilon_k, \varepsilon, r)$-regular to mean $(d_k, \dots, d_2, \varepsilon_k, \varepsilon, r)$-regular. We say that a $(k-1)$-complex $\mathcal{J}$ is (t_0, t_1, ε)-*equitable* if it has the following properties.

(a) $\mathcal{J}$ is $\mathcal{P}$-partite for some $\mathcal{P}$ which partitions $V(\mathcal{J})$ in to t parts, where $t_0 \leq t \leq t_1$, of equal size. We refer to $\mathcal{P}$ as the *ground partition* of $\mathcal{J}$, and to the parts of $\mathcal{P}$ as the *clusters* of $\mathcal{J}$.

(b) There exists a *density vector* $\mathbf{d} = (d_{k-1}, \dots, d_2)$ such that, for each $2 \le i \le k-1$, we have $d_i \ge 1/t_1$ and $1/d_i \in \mathbb{N}$, and $\mathcal{J}$ is $(\mathbf{d}, \varepsilon, \varepsilon, 1)$-regular.

For any k-set X of clusters of $\mathcal{J}$, we denote by $\hat{\mathcal{J}}_X$ the k-partite $(k-1)$-graph $(\mathcal{J}_{X^<})^{(k-1)}$ and call $\hat{\mathcal{J}}_X$ a *polyad.* Given a (t_0, t_1, ε)-equitable $(k-1)$-complex $\mathcal{J}$ and a k-graph G on $V(\mathcal{J})$, we say that G is (ε_k, r)*-regular with respect to a k-set X of clusters of $\mathcal{J}$* if there exists some d such that G is (d, ε_k, r)-regular with respect to the polyad $\hat{\mathcal{J}}_X$. Moreover, we write $d^*_{G,\mathcal{J}}(X)$ for the relative density of G with respect to $\hat{\mathcal{J}}_X$; we may drop either subscript if it is clear from context.

We can now give the crucial definition of a regular slice.

Definition 2.3.1 (Regular slice)**.** Given $\varepsilon, \varepsilon_k > 0, r, t_0, t_1 \in \mathbb{N}$, a graph G and a $(k-1)$-complex $\mathcal{J}$ on $V(G)$, we call $\mathcal{J}$ a $(t_0, t_1, \varepsilon, \varepsilon_k, r)$*-regular slice* for G if $\mathcal{J}$ is (t_0, t_1, ε)-equitable and G is (ε_k, r)-regular with respect to all but at most $\varepsilon_k \binom{t}{k}$ of the k-sets of clusters of $\mathcal{J}$, where t is the number of clusters of $\mathcal{J}$.

If we specify the density vector $\mathbf{d}$ and the number of clusters t of an equitable complex or a regular slice, then it is not necessary to specify t_0 and t_1 (since the only role of these is to bound $\mathbf{d}$ and t). In this situation we write that $\mathcal{J}$ is $(\cdot, \cdot, \varepsilon)$-equitable, or is a $(\cdot, \cdot, \varepsilon, \varepsilon_k, r)$-regular slice for G.

Given a regular slice $\mathcal{J}$ for a k-graph G, we define the d-reduced k-graph $\mathcal{R}^{\mathcal{J}}_d(G)$ as follows.

Definition 2.3.2 (The d-reduced k-graph)**.** Let $k \ge 3$. Let G be a k-graph and let $\mathcal{J}$ be a $(t_0, t_1, \varepsilon, \varepsilon_k, r)$-regular slice for G. Then, for $d > 0$, we define the *d-reduced k-graph* $\mathcal{R}^{\mathcal{J}}_d(G)$ to be the k-graph whose vertices are the clusters of $\mathcal{J}$ and whose edges are all k-sets X of clusters of $\mathcal{J}$ such that G is (ε_k, r)-regular with respect to X and $d^*(X) \ge d$.

We now state the version of the Regular Slice Lemma that we need, which is a special case of [2, Lemma 10].

Lemma 2.3.3 (Regular Slice Lemma [2, Lemma 10])**.** *Let $k \ge 3$. For all positive integers t_0 and s, positive ε_k and all functions $r\colon \mathbb{N} \to \mathbb{N}$ and $\varepsilon\colon \mathbb{N} \to (0,1]$, there are integers t_1*

and n_0 such that the following holds for all $n \geq n_0$ which are divisible by $t_1!$. Let K be a 2-edge-coloured complete k-graph on n vertices. Then there exists a $(k-1)$-complex $\mathcal{J}$ on $V(K)$ which is a $(t_0, t_1, \varepsilon(t_1), \varepsilon_k, r(t_1))$-regular slice for both K^{red} and K^{blue}.

Given a 2-edge-coloured complete k-graph H we want to apply the Regular Slice Lemma to H^{red} and H^{blue}. The following lemma shows that in this setting the union of the corresponding reduced graphs $\mathcal{R}_d^{\mathcal{J}}(H^{\mathrm{red}}) \cup \mathcal{R}_d^{\mathcal{J}}(H^{\mathrm{blue}})$ is almost complete.

Lemma 2.3.4 ([53, Lemma 8.5])**.** *Let $k \geq 3$. Let K be a 2-edge-coloured complete k-graph and let $\mathcal{J}$ be a $(\cdot, \cdot, \varepsilon, \varepsilon_k, r)$-regular slice for both K^{red} and K^{blue}. Let t be the number of clusters of $\mathcal{J}$. Then, provided that $d \leq 1/2$, we have $\left|\mathcal{R}_d^{\mathcal{J}}(K^{\mathrm{red}}) \cup \mathcal{R}_d^{\mathcal{J}}(K^{\mathrm{blue}})\right| \geq (1-2\varepsilon_k)\binom{t}{k}$.*

Proof. Since $\mathcal{J}$ is a $(\cdot, \cdot, \varepsilon, \varepsilon_k, r)$-regular slice for both K^{red} and K^{blue} there are at least $(1-2\varepsilon_k)\binom{t}{k}$ k-sets X of clusters of $\mathcal{J}$ such that both K^{red} and K^{blue} are (ε_k, r)-regular with respect to X. Let X be such a k-set. Since K^{red} and K^{blue} are complements of each other, we have $d^*_{K^{\mathrm{red}}}(X) + d^*_{K^{\mathrm{blue}}}(X) = 1$. Hence $d^*_{K^{\mathrm{red}}}(X) \geq 1/2$ or $d^*_{K^{\mathrm{blue}}}(X) \geq 1/2$ and thus, since $d \leq 1/2$, we have $X \in \mathcal{R}_d^{\mathcal{J}}(K^{\mathrm{red}}) \cup \mathcal{R}_d^{\mathcal{J}}(K^{\mathrm{blue}})$. □

Let H be a k-graph. A *fractional matching* in H is a function $\omega : E(H) \to [0,1]$ such that for all $v \in V(H)$, $\sum_{e \in H : v \in e} \omega(e) \leq 1$. The *weight* of the fractional matching is defined to be $\sum_{e \in H} \omega(e)$. A fractional matching is *tightly connected* if the subgraph induced by the edges e with $\omega(e) > 0$ is tightly connected in H. The following result from [2] converts a tightly connected fractional matching in the reduced graph into a tight cycle in the original graph.

Lemma 2.3.5 ([2, Lemma 13])**.** *Let k, r, n_0, t be positive integers, and let $\psi, \varepsilon, \varepsilon_k, d_k, \ldots, d_2$ be positive constants such that $1/d_i \in \mathbb{N}$ for each $2 \leq i \leq k-1$, and such that $1/n_0 \ll 1/t$,*

$$\frac{1}{n_0} \ll \frac{1}{r}, \varepsilon \ll \varepsilon_k, d_{k-1}, \ldots, d_2 \quad \text{and} \quad \varepsilon_k \ll \psi, d_k, \frac{1}{k}.$$

Then the following holds for all integers $n \geq n_0$. Let G be a k-graph on n vertices, and $\mathcal{J}$ be *a $(\cdot, \cdot, \varepsilon, \varepsilon_k, r)$-regular slice for G with t clusters and density vector $(d_{k-1}, \dots, d_2)$. Suppose that $\mathcal{R}^{\mathcal{J}}_{d_k}(G)$ contains a tightly connected fractional matching with weight μ. Then G contains a tight cycle of length ℓ for every $\ell \leq (1-\psi)k\mu n/t$ that is divisible by k.*

We use the following fact, lemma and proposition to prove Lemma 2.3.10 which is a stronger version of Lemma 2.3.5 that allows us to control the location of the tight cycle.

Fact 2.3.6 ([2, Fact 7]). *Suppose that $1/m_0 \ll \varepsilon \ll 1/t_1, 1/t_0, \beta, 1/k \leq 1/3$ and that $\mathcal{J}$ is a (t_0, t_1, ε)-equitable $(k-1)$-complex with density vector $(d_{k-1}, \dots, d_2)$ whose clusters each have size $m \geq m_0$. Let X be a set of k clusters of $\mathcal{J}$. Then*

$$\left| K_k((\mathcal{J}_{X^<})^{(k-1)}) \right| = (1 \pm \beta) m^k \prod_{i=2}^{k-1} d_i^{\binom{k}{i}}.$$

Lemma 2.3.7 (Regular Restriction Lemma [2, Lemma 28]). *Suppose integers k, m and reals $\alpha, \varepsilon, \varepsilon_k, d_k, \dots, d_2 > 0$ are such that*

$$\frac{1}{m} \ll \varepsilon \ll \varepsilon_k, d_{k-1}, \dots, d_2 \quad \text{and} \quad \varepsilon_k \ll \alpha, \frac{1}{k}.$$

For any $r, s \in \mathbb{N}$ and $d_k > 0$, set $\mathbf{d} = (d_k, \dots, d_2)$, and let $\mathcal{G}$ be an s-partite k-complex whose vertex classes $V_1, \dots, V_s$ each have size m and which is $(\mathbf{d}, \varepsilon_k, \varepsilon, r)$-regular. Choose any $V_i' \subseteq V_i$ with $|V_i'| \geq \alpha m$ for each $i \in [s]$. Then the induced subcomplex $\mathcal{G}[V_1' \cup \dots \cup V_s']$ is $(\mathbf{d}, \sqrt{\varepsilon_k}, \sqrt{\varepsilon}, r)$-regular.

The following proposition shows that a refinement of a regular slice is also a regular slice.

Proposition 2.3.8. *Let $1/m \ll \varepsilon \ll 1/N, 1/t_0, 1/t_1, 1/k \leq 1/3$. Let $\mathcal{J}$ be a (t_0, t_1, ε)-equitable $(k-1)$-complex with density vector $(d_{k-1}, \dots, d_2)$ and clusters $V_1, \dots, V_t$ each of size m. Let $V_{i,1}, \dots, V_{i,N}$ be an equipartition of V_i for each $i \in [t]$. Then there exists a $(Nt_0, Nt_1, \sqrt{\varepsilon})$-equitable $(k-1)$-complex $\widetilde{\mathcal{J}}$ with density vector $(d_{k-1}, \dots, d_2)$, ground partition $\{V_{i,j} \colon i \in [t], j \in [N]\}$ and $\widetilde{\mathcal{J}}[V_1, \dots, V_t] = \mathcal{J}$.*

Proof. We construct $\widetilde{\mathcal{J}}$ from $\mathcal{J}$ as follows. Let the ground partition of $\widetilde{\mathcal{J}}$ be $\{V_{i,j} \colon i \in [t], j \in [N]\}$. Starting with the edges of $\mathcal{J}$ we iteratively add additional edges at random as follows. For each $2 \le i \le k-1$, beginning with $i = 2$, we add each i-edge that contains two vertices that are in vertex classes with the same first index and is supported on the $(i-1)$-edges independently with probability d_i.

We now show that with high probability $\widetilde{\mathcal{J}}$ is the desired $(k-1)$-complex. Note that it suffices to show that with high probability $\widetilde{\mathcal{J}}$ is $(\mathbf{d}, \sqrt{\varepsilon}, \sqrt{\varepsilon}, 1)$-regular.

Let $\widetilde{\mathcal{J}}^{\le i} = \bigcup_{j \in [i]} \widetilde{\mathcal{J}}^{(j)}$ and $\mathbf{d}^{\le i} = (d_i, \dots, d_2)$. For $i \in [k-1]$, let B_i be the event that $\widetilde{\mathcal{J}}^{\le i}$ is not $(\mathbf{d}^{\le i}, \sqrt{\varepsilon}, \sqrt{\varepsilon}, 1)$-regular. Note that $B_1 = \varnothing$. Consider $2 \le i \le k-1$ and $A \in \binom{[t]\times[N]}{i}$. Let $B_{i,A}$ be the event that $\widetilde{\mathcal{J}}^{(i)}[V_A]$ is not $(d_i, \sqrt{\varepsilon}, 1)$-regular with respect to $\widetilde{\mathcal{J}}^{(i-1)}[V_A]$.

Claim 2.3.9. *For $i \in [k-1]$ and $A \in \binom{[t]\times[N]}{i}$, we have $\mathbb{P}\left[B_{i,A} \mid \overline{B_{i-1}}\right] = e^{-\Omega(m^i)}$ as $m \to \infty$.*

Proof of Claim. Assume $\overline{B_{i-1}}$ holds. Let $A = \{(r_j, s_j) \colon j \in [i]\}$. Define $\widetilde{A} = \{r_j \colon j \in [i]\}$. If the r_j are distinct, then the claim holds by Lemma 2.3.7 with $\mathcal{G} = \mathcal{J}[V_{\widetilde{A}}]$ and $\alpha = 1/N$. If not all the r_j are distinct, then $\left|K_i(\widetilde{\mathcal{J}}^{(i-1)}[V_A])\right| \ge \frac{1}{2}\left(\prod_{j=2}^{i-1} d_j^{\binom{i}{j}}\right)(m/N)^i$, by Fact 2.3.6. Thus for each subgraph Q of $\widetilde{\mathcal{J}}^{(i-1)}[V_A]$ such that $|K_i(Q)| > \sqrt{\varepsilon}\left|K_i(\widetilde{\mathcal{J}}^{(i-1)}[V_A])\right|$, a Chernoff bound implies that

$$
\begin{aligned}
&\mathbb{P}\left[d(\widetilde{\mathcal{J}}^{(i)}[V_A] \mid Q) \ne d_i \pm \sqrt{\varepsilon} \middle| \overline{B_{i-1}}\right] \\
=&\mathbb{P}\left[\left|\left|\widetilde{\mathcal{J}}^{(i)}[V_A] \cap K_i(Q)\right| - d_i\left|K_i(Q)\right|\right| > \frac{\sqrt{\varepsilon}}{d_i} d_i \left|K_i(Q)\right| \middle| \overline{B_{i-1}}\right] \\
\le&2\exp\left(-\frac{1}{3}\left(\frac{\sqrt{\varepsilon}}{d_i}\right)^2 d_i\left|K_i(Q)\right|\right) \le 2\exp\left(-\frac{1}{3}\frac{\varepsilon^{3/2}}{d_i}\left|K_i(\widetilde{\mathcal{J}}^{(i-1)}[V_A]\right|\right) \\
\le&2\exp\left(-\frac{1}{6}\frac{\varepsilon^{3/2}}{d_i}\left(\prod_{j=2}^{i-1} d_j^{\binom{i}{j}}\right)\left(\frac{m}{N}\right)^i\right) \le e^{-\Omega(m^i)}.
\end{aligned}
$$

Since there are at most $2^{(im)^{i-1}}$ choices for Q, the claim follows by a union bound. ∎

Note that if $\widetilde{\mathcal{J}}$ is not $(\mathbf{d}, \sqrt{\varepsilon}, \sqrt{\varepsilon}, 1)$-regular, then there exists some $i \in [k-1]$ and

$A \in \binom{[t]\times[N]}{i}$ such that $B_{i,A}$ holds. Further by choosing i minimal we can ensure that $\overline{B_{i-1}}$ holds. Thus, by a union bound and Claim 2.3.9, we have

$$\begin{aligned}\mathbb{P}\left[\widetilde{\mathcal{J}} \text{ is not } (\mathbf{d}, \sqrt{\varepsilon}, \sqrt{\varepsilon}, 1)\text{-regular}\right] &\leq \sum_{i=1}^{k-1} \sum_{A \in \binom{[t]\times[N]}{i}} \mathbb{P}\left[B_{i,A} \cap \overline{B_{i-1}}\right] \\ &\leq \sum_{i=1}^{k-1} \sum_{A \in \binom{[t]\times[N]}{i}} \mathbb{P}\left[B_{i,A} \mid \overline{B_{i-1}}\right] = o(1).\end{aligned}$$

□

The following lemma is a strengthening of Lemma 2.3.5. We believe the constant β and the corresponding condition could be removed if one were to go through the proof of Lemma 2.3.5 to prove a stronger result.

Lemma 2.3.10. *Let $1/n \ll 1/r, \varepsilon \ll \varepsilon_k, d_{k-1}, \dots, d_2$ and $\varepsilon_k \ll \varepsilon' \ll \psi, d_k, \beta, 1/k \leq 1/3$ and $1/n \ll 1/t$ such that t divides n and $1/d_i \in \mathbb{N}$ for all $2 \leq i \leq k-1$. Let G be a k-graph on n vertices and $\mathcal{J}$ be a $(\cdot, \cdot, \varepsilon, \varepsilon_k, r)$-regular slice for G. Further, let $\mathcal{J}$ have t clusters $V_1, \dots, V_t$ all of size n/t and density vector $\mathbf{d} = (d_{k-1}, \dots, d_2)$. Suppose that the reduced graph $\mathcal{R}^{\mathcal{J}}_{d_k}(G)$ contains a tightly connected fractional matching φ with weight μ. Assume that all edges with non-zero weight have weight at least β. For each $i \in [t]$, let $W_i \subseteq V_i$ be such that $|W_i| \geq ((1-3\varepsilon')\varphi(V_i) + \varepsilon')n/t$. Then $G\left[\bigcup_{i\in[t]} W_i\right]$ contains a tight cycle of length ℓ for each $\ell \leq (1-\psi)k\mu n/t$ that is divisible by k.*

We first explain the main ideas of the proof. We would like to find a regular slice for $G' = G[\bigcup_{i\in[t]} W_i]$ to then apply Lemma 2.3.5. The issue is that not all vertex classes in G' have the same size. To get around this we take a refinement of the original partition and use Proposition 2.3.8 to find a new regular slice with that ground partition. The reduced graph for this new regular slice will be a blow up of the original reduced graph. We can find a corresponding tightly connected matching in this new reduced graph. Then we simply apply Lemma 2.3.5.

Proof of Lemma 2.3.10. Let $m = n/t$ and $\widetilde{m} = \lfloor \varepsilon' m/2 \rfloor$. For each $i \in [t]$, let $\widetilde{V}_i \subseteq V_i$

such that $\widetilde{m} \mid \left|\widetilde{V}_i\right|$ and $\left|V_i \setminus \widetilde{V}_i\right| \leq \varepsilon' m/2$. By Lemma 2.3.7, $\mathcal{J}[\widetilde{V}_1, \dots, \widetilde{V}_t]$ is $(\cdot, \cdot, \sqrt{\varepsilon})$-equitable with density vector $(d_{k-1}, \dots, d_2)$. Let $N = \lfloor m/\widetilde{m} \rfloor$ and, for each $i \in [t]$, let $N_i = \lfloor ((1-3\varepsilon')\varphi(V_i) + \varepsilon')N \rfloor \leq \lfloor |W_i| / \widetilde{m} \rfloor$. For each $i \in [t]$, let $V_{i,1}, \dots, V_{i,N}$ be an equipartition of $\widetilde{V}_i$ such that $V_{i,1}, \dots, V_{i,N_i} \subseteq W_i$. Let $\widetilde{W} = \{V_{i,j} : i \in [t], j \in [N_i]\}$ and $\widetilde{t} = \left|\widetilde{W}\right|$. By Proposition 2.3.8, there exists a $(\cdot, \cdot, \varepsilon^{1/4})$-equitable $(k-1)$-complex $\mathcal{J}^*$ with density vector $(d_{k-1}, \dots, d_2)$ and ground partition $\{V_{i,j} : i \in [t], j \in [N]\}$ such that $\mathcal{J}[\widetilde{V}_1, \dots, \widetilde{V}_t] = \mathcal{J}^*[\widetilde{V}_1, \dots, \widetilde{V}_t]$. Let $\widetilde{\mathcal{J}} = \mathcal{J}^*_{\widetilde{W}}$, that is $\widetilde{\mathcal{J}}$ is the $(k-1)$-complex contained in $\mathcal{J}^*$ induced by the vertex classes in $\widetilde{W}$.

Let $\widetilde{G}$ be the subgraph of $G[\bigcup \widetilde{W}]$ obtained by removing all edges contained in k-tuples of density less than d_k and in irregular k-tuples. We show that $\widetilde{\mathcal{J}}$ is a regular slice for $\widetilde{G}$. Let X be a set of k clusters of $\widetilde{\mathcal{J}}$. If the k clusters in X are all contained in distinct clusters of $\mathcal{J}$ that form a regular k-tuple of density at least d_k, then let Y denote the k-set of these clusters. Note that $(G \cup \mathcal{J})[Y]$ is $((d, d_{k-1}, \dots, d_2), \varepsilon_k, \varepsilon, r)$-regular, for some $d \geq d_k - \varepsilon_k$, and thus, by Lemma 2.3.7, $(\widetilde{G} \cup \widetilde{\mathcal{J}})[X]$ is $((d, d_{k-1}, \dots, d_2), \sqrt{\varepsilon_k}, \sqrt{\varepsilon}, r)$-regular. Hence $\widetilde{G}$ is $(d, \sqrt{\varepsilon_k}, r)$-regular with respect to $(\widetilde{\mathcal{J}}_{X^<})^{(k-1)}$. Note that, for all other k-sets of clusters X, the k-partite subgraph of $\widetilde{G}$ induced by the clusters in X is empty. For these k-sets of clusters, $\widetilde{G}$ is $(0, \sqrt{\varepsilon_k}, r)$-regular with respect to the polyad $(\widetilde{\mathcal{J}}_{X^<})^{(k-1)}$. Thus $\widetilde{\mathcal{J}}$ is a $(\cdot, \cdot, \sqrt{\varepsilon_k}, \varepsilon^{1/4}, r)$-regular slice for $\widetilde{G}$.

Note that $\widetilde{\mathcal{R}} = \mathcal{R}^{\widetilde{\mathcal{J}}}_{d_k - 2\sqrt{\varepsilon_k}}(\widetilde{G})$ is a blow-up of $\mathcal{R}^{\mathcal{J}}_{d_k}(G)$. Consider the tightly connected fractional matching φ on $\mathcal{R}^{\mathcal{J}}_{d_k}(G)$ with weight μ. We construct a tightly connected matching on $\widetilde{\mathcal{R}}$ as follows. For each $e \in \mathcal{R}^{\mathcal{J}}_{d_k}(G)$, we will pick a matching M_e in $\widetilde{\mathcal{R}}$ of size $\widetilde{\varphi}(e) = \lfloor (1-3\varepsilon')\varphi(e)N \rfloor$. Note that, for each $i \in [t]$,

$$\sum_{e \ni V_i} \widetilde{\varphi}(e) \leq \lfloor ((1-3\varepsilon')\varphi(V_i) + \varepsilon')N \rfloor = N_i. \tag{2.3.1}$$

For each vertex V_i in $\mathcal{R}^{\mathcal{J}}_{d_k}(G)$ and each edge $e \in \mathcal{R}^{\mathcal{J}}_{d_k}(G)$ that contains V_i, we choose disjoint sets $I_{i,e} \subseteq [N_i]$ such that $|I_{i,e}| = \widetilde{\varphi}(e)$. This is possible by (2.3.1). Recall that $\widetilde{\mathcal{R}}$ is a blow-up of $\mathcal{R}^{\mathcal{J}}_{d_k}(G)$. For each edge $e = \{V_{i_1}, V_{i_2}, \dots, V_{i_k}\} \in \mathcal{R}^{\mathcal{J}}_{d_k}(G)$, the subgraph $\widetilde{\mathcal{R}}_e$

of $\widetilde{\mathcal{R}}$ induced by the set of edges $\{\{V_{i_1,j_1},\ldots,V_{i_k,j_k}\}\colon j_1\in I_{i_1,e},\ldots,j_k\in I_{i_k,e}\}$ is a balanced complete k-partite k-graph. Pick a perfect matching M_e in $\widetilde{\mathcal{R}}_e$. Let $M=\bigcup_{e\in\mathcal{R}^{\mathcal{J}}_{d_k}(G)} M_e$. Note that M is a matching of size

$$\begin{aligned}\sum_{e\in\mathcal{R}^{\mathcal{J}}_{d_k}(G)}\widetilde{\varphi}(e)&=\sum_{e\in\mathcal{R}^{\mathcal{J}}_{d_k}(G)}\lfloor(1-3\varepsilon')\varphi(e)N\rfloor\geq\sum_{\substack{e\in\mathcal{R}^{\mathcal{J}}_{d_k}(G)\\ \varphi(e)>0}}((1-3\varepsilon')\varphi(e)N-1)\\&\geq(1-3\varepsilon')\mu N-\mu/\beta=\left(1-3\varepsilon'-\frac{1}{N\beta}\right)\mu N\\&\geq(1-3\varepsilon'-\varepsilon'/\beta)\mu N\geq(1-\sqrt{\varepsilon'})\mu N\geq(1-2\sqrt{\varepsilon'})\mu\frac{m}{\widetilde{m}}.\end{aligned}$$

In the second inequality above we used the fact that since φ is a fractional matching with weight μ and all edges have weight at least β, there are at most μ/β edges of positive weight. Since $\widetilde{\mathcal{R}}$ is a blow-up of $\mathcal{R}^{\mathcal{J}}_{d_k}(G)$, M is tightly connected. We conclude by applying Lemma 2.3.5 with $k,r,n,\widetilde{t},\psi^2,\varepsilon^{1/4},\sqrt{\varepsilon_k},d_k-2\sqrt{\varepsilon_k},d_{k-1},\ldots,d_2,\widetilde{\mathcal{J}},\widetilde{G},\ell$ playing the roles of $k,r,n_0,t,\psi,\varepsilon,\varepsilon_k,d_k,\ldots,d_2,\mathcal{J},G,\ell$. □

For the next result, we need the following definition.

Definition 2.3.11. Let $\mu^s_k(\beta,\varepsilon,n)$ be the largest μ such that every 2-edge-coloured $(1-\varepsilon,\varepsilon)$-dense k-graph on n vertices contains a fractional matching with weight μ such that all edges with non-zero weight have weight at least β and lie in s monochromatic tight components. Let $\mu^s_k(\beta)=\liminf_{\varepsilon\to 0}\liminf_{n\to\infty}\mu^s_k(\beta,\varepsilon,n)/n$. Similarly, let $\mu^*_k(\beta,\varepsilon,n)$ be the largest μ such that every 2-edge-coloured $(1-\varepsilon,\varepsilon)$-dense k-graph on n vertices contains a fractional matching with weight μ such that all edges with non-zero weight have weight at least β and lie in one red and one blue tight component. Let $\mu^*_k(\beta)=\liminf_{\varepsilon\to 0}\liminf_{n\to\infty}\mu^*_k(\beta,\varepsilon,n)/n$.

The following is the crucial result that reduces finding cycles in the original graph to finding tightly connected matchings in the reduced graph.

Corollary 2.3.12. *Let $1/n\ll\eta,\beta,1/k,1/s$ with $k\geq 3$. Let K be a 2-edge-coloured complete k-graph on n vertices. Then the following hold.*

(i) K *contains* s *vertex-disjoint monochromatic tight cycles covering at least* $(\mu_k^s(\beta) - \eta)kn$ *vertices.*

(ii) K *contains two vertex-disjoint monochromatic tight cycles of distinct colours covering at least* $(\mu_k^*(\beta) - \eta)kn$ *vertices.*

(iii) K *contains a monochromatic tight cycle of length* ℓ *for any* $\ell \leq (\mu_k^1(\beta) - \eta)kn$ *divisible by* k.

Proof. We prove the first statement. The other two statements can be proved similarly (where for the third statement we additionally make use of the fact that Lemma 2.3.10 also allows us to control the length of the resulting cycle). Without loss of generality assume that $\eta \leq 1/3$. Let $d_k = 1/2$ and $1/t_0 \ll \varepsilon_k \ll \varepsilon' \ll \varepsilon \ll \eta, \beta, 1/k, 1/s$. Note that $\mu_k^s(\beta, \varepsilon, t) \geq (\mu_k^s(\beta) - \eta^2)t$ for all $t \geq t_0$. We choose functions $\widetilde{\varepsilon}(\cdot)$ and $r(\cdot)$ where $\widetilde{\varepsilon}(\cdot)$ approaches zero sufficiently quickly and $r(\cdot)$ increases sufficiently quickly such that for any integer $t^* \geq t_0$ and $d_2, \ldots, d_{k-1} \geq 1/t^*$ we may apply Lemma 2.3.10 with $\widetilde{\varepsilon}(t^*)$ and $r(t^*)$ playing the roles of ε and r, respectively. We apply Lemma 2.3.3 to obtain n_0 and t_1. Let $\widetilde{\varepsilon} = \widetilde{\varepsilon}(t_1)$ and $r = r(t_1)$. Let $n_1 \geq n_0$ be large enough such that for all $n \geq n_1$ and $d_2, \ldots, d_{k-1} \geq 1/t_1$ we may apply Lemma 2.3.10. Let $n_2 = n_1 + t_1!$. We show that the theorem holds for all $n \geq n_2$. Let K be a 2-edge-coloured complete k-graph on n vertices. Let $\widetilde{n} \leq n$ be the largest integer such that $t_1!$ divides $\widetilde{n}$. Let $\widetilde{K}$ be a complete subgraph of K on $\widetilde{n}$ vertices. Note that $\widetilde{n} \geq n_1$. By Lemma 2.3.3, there exists a $(t_0, t_1, \widetilde{\varepsilon}, \varepsilon_k, r)$-regular slice $\mathcal{J}$ for both $\widetilde{K}^{\text{red}}$ and $\widetilde{K}^{\text{blue}}$. Let t be the number of clusters of $\mathcal{J}$ and let $(d_{k-1}, \ldots, d_2)$ be the density vector of $\mathcal{J}$. Let $\widetilde{H} = \mathcal{R}_{d_k}^{\mathcal{J}}(\widetilde{K}^{\text{red}}) \cup \mathcal{R}_{d_k}^{\mathcal{J}}(\widetilde{K}^{\text{blue}})$ be a 2-edge-coloured k-graph such that $\mathcal{R}_{d_k}^{\mathcal{J}}(\widetilde{K}^{\text{red}}) \setminus \mathcal{R}_{d_k}^{\mathcal{J}}(\widetilde{K}^{\text{blue}}) \subseteq \widetilde{H}^{\text{red}}$ and $\mathcal{R}_{d_k}^{\mathcal{J}}(\widetilde{K}^{\text{blue}}) \setminus \mathcal{R}_{d_k}^{\mathcal{J}}(\widetilde{K}^{\text{red}}) \subseteq \widetilde{H}^{\text{blue}}$. By Lemma 2.3.4, we have $\left|\widetilde{H}\right| \geq (1 - 2\varepsilon_k)\binom{t}{k}$. By Proposition 2.1.2, there exists a $(1 - (2\varepsilon_k)^{1/(4k^2+1)}, (2\varepsilon_k)^{1/(4k^2+1)})$-dense subgraph $H \subseteq \widetilde{H}$ with $V(H) = V(\widetilde{H})$. Since $\varepsilon_k \ll \varepsilon$, H is $(1 - \varepsilon, \varepsilon)$-dense. Let φ be a fractional matching in H of weight $\mu = \mu_k^s(\beta, \varepsilon, t) \geq (\mu_k^s(\beta) - 2\eta^2)t$ such that all edges with non-zero weight have weight at least β and lie in s monochromatic tight components $K_1, \ldots, K_s$ of H. For each $j \in [s]$, we define a fractional matching φ_j in H by setting $\varphi_j(e) = \varphi(e)$ if $e \in K_i$ and $\varphi(e) = 0$ otherwise. For each $j \in [s]$, let μ_j be the weight of φ_j. It follows that $\sum_{j\in[s]} \mu_j = \mu$.

Let $V_1, \ldots, V_t$ be the clusters of $\mathcal{J}$. For each $i \in [t]$ and $j \in [s]$, we define

$$w_{i,j} = \max\{\sum_{\substack{e\in H\\ V_i\in e}} \varphi_j(e) - s\varepsilon', \varepsilon'\}.$$

For each $i \in [t]$, let $V_{i,1}, \ldots, V_{i,s}$ be disjoint subsets of V_i such that $|V_{i,j}| = \lceil w_{i,j}n/t\rceil$. By Lemma 2.3.10, there exist tight cycles $C_1, \ldots, C_s$ in K such that, for all $j \in [s]$, $|C_j| = (1-\eta^2)\mu_j k\tilde{n}/t$, $C_j \subseteq K\left[\bigcup_{i\in[t]} V_{i,j}\right]$ and C_j has the same colour as K_j. Hence $C_1, \ldots, C_s$ are vertex-disjoint and together cover

$$(1-\eta^2)\mu k\tilde{n}/t \geq (1-\eta^2)(\mu_k^s(\beta) - \eta^2)k\tilde{n} \geq (\mu_k^s(\beta) - \eta)kn$$

vertices of K. □

2.4 Blueprints

Let H be a 2-edge-coloured k-graph. We define what we call a *blueprint for H* which is an auxiliary graph that can be used as a guide when finding connected matchings in H. A form of the notion of blueprint for $k = 3$ already appeared in [67].

Definition 2.4.1. Let $\varepsilon > 0$, $k \geq 3$ and let H be a 2-edge-coloured k-graph on n vertices. We say that a 2-edge-coloured $(k-2)$-graph G with $V(G) \subseteq V(H)$ is an *ε-blueprint for H*, if

(BP1) for every edge $e \in G$, there exists a monochromatic tight component $H(e)$ in H such that $H(e)$ has the same colour as e and $d_{\partial H(e)}(e) \geq (1-\varepsilon)n$ and

(BP2) for $e, e' \in G$ of the same colour with $|e \cap e'| = k-3$, we have $H(e) = H(e')$.

We say that e *induces* $H(e)$ and write $R(e)$ or $B(e)$ instead of $H(e)$ if e is red or blue, respectively. We simply say that G is a *blueprint*, when H is clear from context and there exists $\varepsilon > 0$ such that G is an ε-blueprint for H. For $S \in \binom{V(H)}{k-3}$, all the red (blue) edges of a blueprint containing S induce the same red (blue) tight component, so we call that component the red (blue) tight component induced by S. Note that any subgraph of a blueprint is also a blueprint.

Example 2.4.2. *Let $k \geq 3$ and let n be a positive integer. Let A and B be disjoint vertex sets with $|A \cup B| = n$. Let $K^{(k)}(A, B)$ be the 2-edge-coloured complete k-graph with vertex set $A \cup B$ where an edge e is red if and only if $|e \cap A|$ is even (and blue otherwise). Let H be $K^{(k)}(A, B)$ and let G be $K^{(k-2)}(A, B)$ with colours reversed. If $\varepsilon \geq \frac{k-2}{n}$, then G is an ε-blueprint for H. Indeed, for an edge $e \in G$ we can set $H(e) = \{f \in H \colon |f \cap A| = |e \cap A| + 1\}$.*

The main aim of this section is to prove the following lemma that establishes the existence of blueprints for 2-edge-coloured $(1 - \varepsilon, \alpha)$-dense graphs.

Lemma 2.4.3. *Let $1/n \ll \varepsilon \leq \alpha \ll 1/k \leq 1/3$. Let H be a 2-edge-coloured $(1-\varepsilon, \alpha)$-dense k-graph on n vertices. Then there exists a $3\sqrt{\varepsilon}$-blueprint G_* for H with $V(G_*) = V(H)$ and $|G_*| \geq (1 - \alpha - 24k\sqrt{\varepsilon})\binom{n}{k-2}$. Moreover, if $k \geq 4$ and $\varepsilon \ll \alpha$, there exists a $(1 - \alpha^{1/(4(k-2)^2+1)}, \alpha^{1/(4(k-2)^2+1)})$-dense spanning subgraph G of G_*.*

We need a few simple preliminary results to prove Lemma 2.4.3. First we show that any 2-edge-coloured 2-graph with large minimum degree contains a large monochromatic connected subgraph.

Proposition 2.4.4. *Let $0 < \beta \leq 1/6$ and let F be a 2-edge-coloured 2-graph with $|V(F)| \leq n$ and $\delta(F) \geq (1 - \beta)n$. Then there exists a subgraph F' of F of order at least $(1 - \beta)n$ that contains a spanning monochromatic component and $\delta(F') \geq (1 - 2\beta)n$.*

Proof. Let F' be a subgraph of F of maximum order that contains a spanning monochromatic component. Assume without loss of generality that F' contains a spanning red component. Let $S = V(F')$ and $\overline{S} = V(F) \setminus V(F')$. Since $\delta(F) \geq (1 - \beta)n$, we have that $|S| \geq (1 - \beta)n/2$. Suppose, for a contradiction, that $|S| < (1 - \beta)n$. Note that all edges between S and $\overline{S}$ are blue. If $\delta(F) - |S| + 1 > \left|\overline{S}\right|/2$, then each pair of vertices in S has a common neighbour in $\overline{S}$ and so there is a blue component strictly containing S which contradicts the maximality of F'. Therefore

$$\delta(F) - |S| + 1 \le \left|\overline{S}\right|/2 = (|V(F)| - |S|)/2 \le (n - |S|)/2.$$

Hence

$$|S| \ge 2\delta(F) - n + 2 \ge 2(1-\beta)n - n + 2 = (1-2\beta)n + 2.$$

But now every pair of vertices in $\overline{S}$ has a common neighbour in S, since $\left|\overline{S}\right| \le |V(F)| - |S| \le 2\beta n$ and so

$$\delta(F) - \left|\overline{S}\right| + 1 \ge (1-\beta)n - 2\beta n + 1 = (1-3\beta)n + 1 > n/2.$$

Thus $\overline{S} \cup N_F(\overline{S})$ is spanned by a blue component. But since

$$\left|\overline{S} \cup N_F(\overline{S})\right| \ge \delta(F) \ge (1-\beta)n,$$

we have a contradiction. It is easy to see that $\delta(F') \ge (1-2\beta)n$. □

Proposition 2.4.5. *Let $1/n \ll \gamma \le 1/9$. Let F be a 2-graph with $|V(F)| \le n$ and $|E(F)| \ge (1-\gamma)\binom{n}{2}$. Then there exists a subgraph of F with minimum degree at least $(1-3\sqrt{\gamma})n$.*

Proof. Let $W = \{v \in V(F) \colon d(v) < (1-2\sqrt{\gamma})n\}$. We have that

$$(1-2\gamma)n^2 \le 2\,|E(F)| = \sum_{v \in V(F)} d(v) \le n^2 - 2\sqrt{\gamma}n\,|W|\,.$$

This implies that $|W| \le \sqrt{\gamma}n$. Let $F^* = F - W$. It follows that $\delta(F^*) \ge (1-2\sqrt{\gamma})n - |W| \ge (1-3\sqrt{\gamma})n$. □

Corollary 2.4.6. *Let $1/n \ll \varepsilon \le 1/324$. Let F be a 2-edge-coloured 2-graph with*

$|V(F)| \leq n$ *and* $|E(F)| \geq (1-\varepsilon)\binom{n}{2}$. *Then there exists a subgraph* F' *of* F *of order at least* $(1-3\sqrt{\varepsilon})n$ *that contains a spanning monochromatic component and* $\delta(F') \geq (1-6\sqrt{\varepsilon})n$.

Proof. By Proposition 2.4.5, there exists a subgraph F^* of F with $\delta(F^*) \geq (1-3\sqrt{\varepsilon})n$. We conclude by applying Proposition 2.4.4 with $F = F^*$ and $\beta = 3\sqrt{\varepsilon}$. □

2.4.1 Proof of Lemma 2.4.3

Now we show that for any $(1-\varepsilon, \alpha)$-dense 2-edge-coloured graph we can find a dense blueprint.

Proof of Lemma 2.4.3. Let $F = \partial^2 H$. Since H is $(1-\varepsilon, \alpha)$-dense,

$$E(F) = \left\{ e \in \binom{V(H)}{k-2} : d_H(e) > 0 \right\} = \left\{ e \in \binom{V(H)}{k-2} : d_H(e) \geq (1-\varepsilon)\binom{n}{2} \right\}$$

and

$$|E(F)| \geq (1-\alpha)\binom{n}{k-2}. \tag{2.4.1}$$

We now colour each edge e of F as follows. Note that the link graph H_e is a 2-graph. We induce a 2-edge-colouring on H_e by colouring the 2-edge $f \in H_e$ with the colour of the k-edge $e \cup f \in H$. By Corollary 2.4.6, there exists a monochromatic component in H_e of order at least $(1-3\sqrt{\varepsilon})n$. Let K_e be such a component chosen arbitrarily. We colour the edge e according to the colour of K_e. If e is red in F, then we define $R(e) \subseteq H$ to be the red tight component containing all the edges $e \cup f$ where $f \in K_e$. If e is blue in F, then we define $B(e)$ analogously.

In the next claim we show that, for each $S \in \binom{V(H)}{k-3}$, almost all edges in F of the same colour containing S induce the same monochromatic tight component in H.

Claim 2.4.7. *For each* $S \in \binom{V(H)}{k-3}$, *there exist* $\Gamma^{\text{red}}(S) \subseteq N_F^{\text{red}}(S)$ *and* $\Gamma^{\text{blue}}(S) \subseteq N_F^{\text{blue}}(S)$ *with* $\left|\Gamma^{\text{red}}(S)\right| \geq \left|N_F^{\text{red}}(S)\right| - 6\sqrt{\varepsilon}n$ *and* $\left|\Gamma^{\text{blue}}(S)\right| \geq \left|N_F^{\text{blue}}(S)\right| - 6\sqrt{\varepsilon}n$ *such that, for all* $y_1, y_2 \in \Gamma^{\text{red}}(S)$, $R(S \cup y_1) = R(S \cup y_2)$ *and, for all* $y_1', y_2' \in \Gamma^{\text{blue}}(S)$, $B(S \cup y_1') = B(S \cup y_2')$.

Proof of Claim. We only prove the statement for $N_F^{\text{red}}(S)$ as the proof of the statement for $N_F^{\text{blue}}(S)$ is analogous. Assume $\left|N_F^{\text{red}}(S)\right| > 6\sqrt{\varepsilon}n$ (or else we simply set $\Gamma^{\text{red}}(S) = \varnothing$). Let D be the directed graph with vertex set $N_F^{\text{red}}(S)$ and edge set

$$E(D) = \{y_1y_2 \colon y_1 \in V(K_{S\cup y_2})\}\,.$$

Note that, for $y_1y_2 \in E(D)$, there exists an edge in $R(S \cup y_2)$ containing $S \cup y_1y_2$. So if y_1y_2 is a double edge (that is, $y_1y_2, y_2y_1 \in E(D)$), then $R(S \cup y_1) = R(S \cup y_2)$. For $y \in N_F^{\text{red}}(S)$,

$$d_D^-(y) \geq \left|N_F^{\text{red}}(S) \cap V(K_{S\cup y})\right| \geq \left|N_F^{\text{red}}(S)\right| - 3\sqrt{\varepsilon}n,$$

since $|V(K_{S\cup y})| \geq (1 - 3\sqrt{\varepsilon})n$. Hence the number of double edges in D is at least

$$\left|N_F^{\text{red}}(S)\right| \left(\left|N_F^{\text{red}}(S)\right| - 3\sqrt{\varepsilon}n\right) - \frac{1}{2}\left|N_F^{\text{red}}(S)\right|^2 = \frac{1}{2}\left|N_F^{\text{red}}(S)\right| \left(\left|N_F^{\text{red}}(S)\right| - 6\sqrt{\varepsilon}n\right).$$

Thus there exists a vertex $y_0 \in N_F^{\text{red}}(S)$ that is incident to at least $\left|N_F^{\text{red}}(S)\right| - 6\sqrt{\varepsilon}n$ double edges. Let $\Gamma^{\text{red}}(S) = \{y_0\} \cup \{y \in N_F^{\text{red}}(S) \colon yy_0, y_0y \in E(D)\}$. Note that $\left|\Gamma^{\text{red}}(S)\right| \geq \left|N_F^{\text{red}}(S)\right| - 6\sqrt{\varepsilon}n$ and $R(S \cup y) = R(S \cup y_0)$ for all $y \in \Gamma^{\text{red}}(S)$. ■

Consider the multi-$(k-2)$-graph D^* with

$$E(D^*) = \left\{S \cup y \colon S \in \binom{V(H)}{k-3}, y \in \Gamma^{\text{red}}(S) \cup \Gamma^{\text{blue}}(S)\right\}.$$

Note that

$$\begin{aligned}|E(D^*)| &= \sum_{S\in\binom{V(H)}{k-3}} \left|\Gamma^{\text{red}}(S) \cup \Gamma^{\text{blue}}(S)\right| \geq \sum_{S\in\binom{V(H)}{k-3}} (d_F(S) - 12\sqrt{\varepsilon}n) \\ &\geq (k-2)\,|F| - 24k\sqrt{\varepsilon}\binom{n}{k-2}.\end{aligned}$$

Every edge in D^* has multiplicity at most $k-2$. So at least $|F| - 24k\sqrt{\varepsilon}\binom{n}{k-2}$ edges $e \in \binom{V(H)}{k-2}$ have multiplicity $k-2$ in D^*. Let G_* be the $(k-2)$-graph on $V(H)$ such that $e \in G_*$ if and only if e has multiplicity $k-2$ in D^*. So, by (2.4.1), $|G_*| \geq |F| - 24k\sqrt{\varepsilon}\binom{n}{k-2} \geq (1-\alpha-24k\sqrt{\varepsilon})\binom{n}{k-2}$.

We now show that G_* is a $3\sqrt{\varepsilon}$-blueprint for H. Consider any $e, e' \in G_*^{\mathrm{red}}$ with $|e \cap e'| = k-3$. Let $S = e \cap e'$, $y = e' \setminus S$ and $y' = e \setminus S$. Since $e, e' \in G_*^{\mathrm{red}}$, we have $y, y' \in \Gamma^{\mathrm{red}}(S)$ and so $R(e) = R(S \cup y) = R(S \cup y') = R(e')$. Further, for $e \in G_*^{\mathrm{red}}$, we have $d_{\partial R(e)}(e) \geq |V(K_e)| \geq (1-3\sqrt{\varepsilon})n$. Analogous statements hold for edges of G_*^{blue}.

If $k \geq 4$ and $\varepsilon \ll \alpha$, then $|G_*| \geq (1-4\alpha)\binom{n}{k-2}$ and thus by Proposition 2.1.2 there exists a subgraph $G \subseteq G_*$ such that G is $(1-\alpha^{1/(4(k-2)^2+1)}, \alpha^{1/(4(k-2)^2+1)})$-dense and $V(G) = V(G_*) = V(H)$. $\square$

2.4.2 Some lemmas about blueprints

Let H be a k-graph and G be a blueprint for H. We write $H(G)$ for $\bigcup_{e \in G} H(e)$. We write G^+ for the subgraph of $H(G)$ with edge set

$$E(G^+) = \{e \in H(G) \colon f \subseteq e \text{ for some } f \in G\},$$

that is, the subgraph of $H(G)$ obtained by deleting all edges that do not contain an edge of G. Note that this also defines $(G')^+$ for any subgraph G' of G as a subgraph of a blueprint for H is also a blueprint for H. Moreover, note that G^+ is a subgraph of H, not of G. For a red tight component R_* and a blue tight component B_* in H, we denote by R_*^{k-2} and B_*^{k-2} the sets of edges of G that induce R_* and B_*, respectively.

We prove some lemmas that we will use several times later on. Roughly speaking, the following lemma states that if S is a set of $k-4$ vertices of H contained in many edges of both R_*^{k-2} and B_*^{k-2}, then S is contained in an edge of R_* or B_*.

Lemma 2.4.8. *Let $1/n \ll \varepsilon \ll \alpha \ll 1$. Let H be a 2-edge-coloured $(1-\varepsilon, \alpha)$-dense k-graph on n vertices and G a $3\sqrt{\varepsilon}$-blueprint for H. Let R_* and B_* be a red and a blue tight* *component of H, respectively. Let $U \subseteq V(G)$ and $S \in \binom{U}{k-4}$ such that*

$$d_{R_*^{k-2}}(S,U), d_{B_*^{k-2}}(S,U) \geq \varepsilon^{1/4} n^2.$$

Then there exist $x, x', y, y' \in U$ such that $S \cup xx' \in R_*^{k-2}$, $S \cup yy' \in B_*^{k-2}$, $S \cup xx'y \in \partial R_*$, $S \cup yy'x \in \partial B_*$ and $S \cup xx'yy' \in H$. In particular, $(R_*^{k-2})^+[U] \cup (B_*^{k-2})^+[U] \neq \varnothing$.

Proof. Let $X_{R_*} = \{x \in U \colon d_{R_*^{k-2}}(S \cup x, U) \geq \varepsilon^{1/2} n\}$ and $X_{B_*} = \{x \in U \colon d_{B_*^{k-2}}(S \cup x, U) \geq \varepsilon^{1/2} n\}$. Note that

$$\varepsilon^{1/4} n^2 \leq d_{R_*^{k-2}}(S,U) = \frac{1}{2} \sum_{x \in U} d_{R_*^{k-2}}(S \cup x, U) \leq n \left| X_{R_*} \right| + \varepsilon^{1/2} n^2.$$

Thus $|X_{R_*}| \geq (\varepsilon^{1/4} - \varepsilon^{1/2}) n \geq \frac{1}{2} \varepsilon^{1/4} n$. Similarly, $|X_{B_*}| \geq \frac{1}{2} \varepsilon^{1/4} n$.

For each $x \in X_{R_*}$, let

$$\begin{aligned} Y_x &= \{y \in X_{B_*} \colon S \cup yy' \in B_*^{k-2} \text{ and } S \cup xyy' \in \partial B_* \text{ for some } y' \in U\} \\ &= \bigcup_{y' \in U} N_{B_*^{k-2}}(S \cup y') \cap N_{\partial B_*}(S \cup xy'). \end{aligned}$$

For each $y \in X_{B_*}$, there exists $y' \in U$ with $S \cup yy' \in B_*^{k-2}$. By (BP1), $d_{\partial B_*}(S \cup yy', X_{R_*}) \geq |X_{R_*}| - 3\sqrt{\varepsilon} n$. Hence each $y \in X_{B_*}$ is contained in at least $|X_{R_*}| - 3\sqrt{\varepsilon} n$ of the sets Y_x. By averaging, there exists an $x \in X_{R_*}$ such that

$$|Y_x| \geq \frac{(|X_{R_*}| - 3\sqrt{\varepsilon} n) \left| X_{B_*} \right|}{2 \left| X_{R_*} \right|} \geq \frac{1}{4} \left| X_{B_*} \right| \geq \frac{1}{8} \varepsilon^{1/4} n.$$

Fix such an $x \in X_{R_*}$. For each $y \in Y_x$, choose a vertex $y' \in U$ such that $S \cup yy' \in B_*^{k-2}$ and $S \cup xyy' \in \partial B_*$. Let $X = N_{R_*^{k-2}}(S \cup x, U)$, so $|X| \geq \varepsilon^{1/2} n$, since $x \in X_{R_*}$. For each $y \in Y_x$, since H is $(1-\varepsilon, \alpha)$-dense, there are at least $|X| - \varepsilon n$ vertices $x' \in X$ such that $S \cup xx'yy' \in H$. Thus, by averaging, there exists a vertex $x' \in X$ and a set $\widetilde{Y}_x \subseteq Y_x$ with

$$\left| \widetilde{Y}_x \right| \geq \frac{(|X| - \varepsilon n) \left| Y_x \right|}{2 \left| X \right|} \geq \frac{1}{4} \left| Y_x \right| \geq \frac{1}{32} \varepsilon^{1/4} n$$

such that $S \cup xx'yy' \in H$ for all $y \in \widetilde{Y}_x$. Fix such an $x' \in X$. Since $S \cup xx' \in R_*^{k-2}$, we have that

$$\left|N_{\partial R_*}(S \cup xx') \cap \widetilde{Y}_x\right| \geq \left|\widetilde{Y}_x\right| - 3\sqrt{\varepsilon}n \geq \left(\frac{1}{32}\varepsilon^{1/4} - 3\sqrt{\varepsilon}\right)n > 0.$$

Choose $y \in N_{\partial R_*}(S \cup xx') \cap \widetilde{Y}_x$. We have $S \cup xx' \in R_*^{k-2}$, $S \cup yy' \in B_*^{k-2}$, $S \cup xx'y \in \partial R_*$, $S \cup xyy' \in \partial B_*$ and $S \cup xx'yy' \in H$ as required. $\square$

The following lemma shows that if we have a vertex set $T \in \binom{V(G)}{k-3}$ such that $d_G^{\text{red}}(T)$ and $d_G^{\text{blue}}(T)$ are both large, then T is contained a lot of sets in $\partial R \cap \partial B$, where R and B are the red and blue tight components induced by the red and blue edges incident to T, respectively.

Lemma 2.4.9. *Let $1/n \ll \varepsilon \ll 1$, $k \geq 3$ and $\delta > 5\sqrt{\varepsilon}$. Let H be a 2-edge-coloured k-graph on n vertices and G a $3\sqrt{\varepsilon}$-blueprint for H. Let $T \in \binom{V(H)}{k-3}$. Let $S^{\text{blue}} \subseteq N_G^{\text{blue}}(T)$ and $S^{\text{red}} \subseteq N_G^{\text{red}}(T)$ be such that $|S^{\text{blue}}|, |S^{\text{red}}| \geq \delta n$. Then there exists a vertex $y \in S^{\text{blue}}$ such that, for*

$$\Gamma_y^{\text{red}} = \{x \in S^{\text{red}} \colon T \cup xy \in \partial R(T \cup x) \cap \partial B(T \cup y)\},$$

we have $\left|\Gamma_y^{\text{red}}\right| \geq (\delta - 6\sqrt{\varepsilon})n$. Moreover, if $\delta \geq \varepsilon^{1/9}$, then $\left|\Gamma_y^{\text{red}}\right| \geq (1 - \varepsilon^{1/4})\left|S^{\text{red}}\right|$. The same statements hold when the colours are reversed.

Proof. Let $m_{\text{blue}} = |S^{\text{blue}}|$ and $m_{\text{red}} = |S^{\text{red}}|$. If $\delta < \varepsilon^{1/9}$, then we may assume that $m_{\text{blue}} = m_{\text{red}} = \lceil \delta n \rceil$ by deleting vertices in S^{blue} and S^{red} if necessary. Let D be the bipartite directed graph with vertex classes S^{blue} and S^{red} such that, for each $y \in S^{\text{blue}}$ and $x \in S^{\text{red}}$, we have $N_D^+(y) = N_{\partial B}(T \cup y) \cap S^{\text{red}}$ and $N_D^+(x) = N_{\partial R}(T \cup x) \cap S^{\text{blue}}$. Since G is a $3\sqrt{\varepsilon}$-blueprint for H, we have that

$$\begin{aligned}|E(D)| &\geq m_{\text{blue}}(m_{\text{red}} - 3\sqrt{\varepsilon}n) + m_{\text{red}}(m_{\text{blue}} - 3\sqrt{\varepsilon}n) \\ &= 2m_{\text{blue}}m_{\text{red}} - 3\sqrt{\varepsilon}n(m_{\text{blue}} + m_{\text{red}}).\end{aligned}$$

Thus the number of double edges in D is at least $m_{\text{blue}}m_{\text{red}} - 3\sqrt{\varepsilon}n(m_{\text{blue}} + m_{\text{red}})$. For each $y \in S^{\text{blue}}$, let $\Gamma_y = \{x \in S^{\text{red}} \colon xy, yx \in D\}$. Hence there is some vertex $y \in S^{\text{blue}}$ such that

$$|\Gamma_y| \geq m_{\text{red}} - 3\sqrt{\varepsilon}n\left(\frac{m_{\text{blue}} + m_{\text{red}}}{m_{\text{blue}}}\right) \geq \begin{cases} (\delta - 6\sqrt{\varepsilon})n, & \text{if } \delta < \varepsilon^{1/9}; \\ m_{\text{red}}(1 - \varepsilon^{1/4}), & \text{otherwise.} \end{cases}$$

Note that if $xy, yx \in D$ with $x \in S^{\text{red}}$ and $y \in S^{\text{blue}}$, then $T \cup xy \in \partial R(T \cup x) \cap \partial B(T \cup y)$. Hence $\Gamma_y \subseteq \Gamma_y^{\text{red}}$ and thus the lemma follows. □

Roughly speaking, in the next lemma we consider the following situation. Let R be a red tight complement in H, G be a blueprint for H and $R_G \subseteq G^{\text{red}}$ be such that $H(R_G) \subseteq R$. We pick a maximal matching in R_G^+ and let U be the remaining vertices of H not in this matching, so $R_G^+[U]$ is empty. Then the lemma implies that the number of monochromatic tight components in U is less than what we would expect. In particular, if $k = 4$, then the edges in $G[U]$ induce only two monochromatic tight components in H.

Lemma 2.4.10. *Let $k \geq 4$ and $1/n \ll \varepsilon \ll \alpha, \delta \ll \eta \ll 1$. Let H be a $(1-\varepsilon, \alpha)$-dense k-graph and G a $3\sqrt{\varepsilon}$-blueprint for H. Let R be a red tight component in H. Let $R_G \subseteq G^{\text{red}}$ be such that $H(R_G) \subseteq R$. Let $U \subseteq V(H)$ be such that $|U| \geq \eta n/2$ and $R_G^+[U] = \varnothing$. Let $S \in \binom{U}{k-4}$ be such that the link graph G_S of G satisfies $G_S^{\text{red}}[U] \subseteq (R_G)_S$ and $\delta(G_S[U]) \geq |U| - \delta n$. Then there exists a subgraph J_S of $G_S[U]$ such that $|J_S| \geq |G_S[U]| - 7\delta^{1/4}n^2$ and $H(S \cup e) = H(S \cup e')$ for all $e, e' \in J_S$ of the same colour. In particular, if $k = 4$, then the edges in J induce only one red and one blue tight component in H. The same statement holds when the colours are reversed.*

Proof. Set $J_S^{\text{red}} = G_S^{\text{red}}[U]$. Note that for $e, e' \in J_S^{\text{red}}$, we have $e, e' \in (R_G)_S$ and thus $H(S \cup e) = H(S \cup e') = R$ since $H(R_G) \subseteq R$. Therefore to prove the lemma, it suffices to prove that there exists $J_S^{\text{blue}} \subseteq G_S^{\text{blue}}[U]$ such that $\left|J_S^{\text{red}}\right| + \left|J_S^{\text{blue}}\right| \geq |G_S[U]| - 7\delta^{1/4}n^2$ and $H(S \cup e) = H(S \cup e')$ for all $e, e' \in J_S^{\text{blue}}$.

For simplicity we assume $k = 4$ and $S = \varnothing$. It is easy to see that an analogous argument works in the general case. Thus for the rest of the proof, we omit the subscript S.

Let $K = G[U]$. If $\left|K^{\text{blue}}\right| < 2\delta^{1/2}n^2$, then we are done by setting $J^{\text{blue}} = \varnothing$ as

$$\left|J^{\text{red}}\right| = \left|K^{\text{red}}\right| = |K| - \left|K^{\text{blue}}\right| \geq |K| - 2\delta^{1/2}n^2 \geq |K| - 7\delta^{1/4}n^2.$$

Now assume $\left|K^{\text{blue}}\right| \geq 2\delta^{1/2}n^2$. Let $X = \{x \in V(K) \colon d_K^{\text{blue}}(x) \geq \delta n\}$. We have that

$$2\delta^{1/2}n^2 \leq \left|K^{\text{blue}}\right| \leq \sum_{x\in U} d_K^{\text{blue}}(x) \leq n\,|X| + \delta n^2.$$

Thus $|X| \geq \delta^{1/2}n$. Let D be the digraph with vertex set X such that, for each $x \in X$,

$$N_D^+(x) = N_K^{\text{blue}}(x, X) \cup \{x' \in N_K^{\text{red}}(x, X) \colon xx'y \in \partial R \cap \partial B(xy) \text{ for some } y \in N_K^{\text{blue}}(x)\}.$$

We now bound $\delta^+(D)$ as follows. If $d_K^{\text{red}}(x, X) \geq \delta n$, then by applying Lemma 2.4.9 (with $x, N_G^{\text{blue}}(x, U), N_G^{\text{red}}(x, X), \delta$ playing the roles of $T, S^{\text{blue}}, S^{\text{red}}, \delta$), we deduce that

$$\begin{aligned}&\left|\{x' \in N_K^{\text{red}}(x, X) \colon xx'y \in \partial R(xx') \cap \partial B(xy) \text{ for some } y \in N_K^{\text{blue}}(x)\}\right|\\ &\geq (1 - \varepsilon^{1/4})d_K^{\text{red}}(x, X).\end{aligned}$$

Recall that $R = R(xx')$ for all $x' \in N_K^{\text{red}}(x, X)$, $|X| \geq \delta^{1/2}n$ and $\varepsilon \ll \delta$. Hence

$$\begin{aligned}d_D^+(x) &\geq d_K^{\text{blue}}(x, X) + (1 - \varepsilon^{1/4})d_K^{\text{red}}(x, X) \geq (1 - \varepsilon^{1/4})(d_K^{\text{blue}}(x, X) + d_K^{\text{red}}(x, X))\\ &= (1 - \varepsilon^{1/4})d_K(x, X) \geq (1 - \varepsilon^{1/4})(|X| - \delta n) \geq (1 - 2\delta^{1/2})\,|X|\,.\end{aligned}$$

On the other hand, if $d_K^{\text{red}}(x, X) < \delta n$, then

$$d_D^+(x) \geq d_K^{\text{blue}}(x, X) \geq |X| - \delta n - d_K^{\text{red}}(x, X) \geq |X| - 2\delta n \geq (1 - 2\delta^{1/2})\,|X|\,.$$

Therefore, we have $\delta^+(D) \geq (1 - 2\delta^{1/2})\,|X|$ and so $|E(D)| \geq (1 - 2\delta^{1/2})\,|X|^2 \geq 2(1 -$

$2\delta^{1/2})\binom{|X|}{2}$. Let F be the graph with vertex set X in which xx' forms an edge if and only if it forms a double edge in D. Note that $|F| \geq (1-4\delta^{1/2})\binom{|X|}{2}$. By Proposition 2.4.5, there exists a subgraph F^* of F with $\delta(F^*) \geq (1-6\delta^{1/4})\,|X|$. Clearly, F^* is connected.

Let $J^{\mathrm{blue}} = \{xx' \in K^{\mathrm{blue}} \colon x \in V(F^*)\}$. We have

$$\begin{aligned}\left|J^{\mathrm{red}} \cup J^{\mathrm{blue}}\right| &\geq |K| - \sum_{x' \in U \setminus X} d_K^{\mathrm{blue}}(x') - |X \setminus V(F^*)|\, n \\ &\geq |K| - \delta n^2 - 6\delta^{1/4} n^2 \geq |G[U]| - 7\delta^{1/4} n^2.\end{aligned}$$

We now show that $B(x_1z_1) = B(x_2z_2)$ for all $x_1z_1, x_2z_2 \in J^{\mathrm{blue}}$. Since F^* is connected and $d_{J^{\mathrm{blue}}}(x) > 0$ for all $x \in V(F^*)$, it suffices to consider the case when $x_1x_2 \in F^*$. If $x_1x_2 \in K^{\mathrm{blue}}$, then $x_1z_1, x_1x_2, x_2z_2 \in G^{\mathrm{blue}}$ and so $B(x_1z_1) = B(x_1x_2) = B(x_2z_2)$, since G is a blueprint. Now assume that $x_1x_2 \in K^{\mathrm{red}}$. Since $x_1x_2 \in F^* \subseteq F$, there are $y_1 \in N_K^{\mathrm{blue}}(x_1)$ and $y_2 \in N_K^{\mathrm{blue}}(x_2)$ such that $x_1x_2y_1 \in \partial R \cap \partial B(x_1y_1)$ and $x_1x_2y_2 \in \partial R \cap \partial B(x_2y_2)$. Let $u \in N_H(x_1x_2y_1) \cap N_H(x_1x_2y_2) \cap U$. Since $R_G^+[U] = \varnothing$, we have $x_1x_2y_1u, x_1x_2y_2u \in H^{\mathrm{blue}}$. Hence, $B(x_1y_1) = B(x_2y_2)$. Moreover, since $x_1y_1, x_1z_1, x_2y_2, x_2z_2 \in G^{\mathrm{blue}}$, we have $B(x_1z_1) = B(x_1y_1) = B(x_2y_2) = B(x_2z_2)$ as required. □

2.5 Monochromatic connected matchings in $K_n^{(4)}$

In this section, we prove that every almost complete red-blue edge-coloured 4-graph H contains a red and a blue tightly connected matching that are vertex-disjoint and together cover almost all vertices of H.

Lemma 2.5.1. *Let $1/n \ll \varepsilon \ll \alpha \ll \eta < 1$. Let H be a 2-edge-coloured $(1-\varepsilon, \alpha)$-dense 4-graph on n vertices. Then H contains two vertex-disjoint monochromatic tightly connected matchings of distinct colours such that their union covers all but at most $3\eta n$ of the vertices of H.*

Note that this implies $\mu_4^*(1, \varepsilon, n) \geq (1-3\eta)n/4$ for $1/n \ll \varepsilon \ll \eta < 1$. Hence $\mu_4^*(1) \geq 1/4$. Therefore, together with Corollary 2.3.12, Lemma 2.5.1 implies Theorem 1.1.2.

To prove Lemma 2.5.1 we first need the following lemma which chooses the initial tight components in H in which we find our tightly connected matchings.

Lemma 2.5.2. *Let $1/n \ll \varepsilon \ll \alpha \ll \eta < 1$. Let H be a 2-edge-coloured $(1-\varepsilon, \alpha)$-dense 4-graph on n vertices. Suppose that H does not contain two vertex-disjoint monochromatic tightly connected matchings of distinct colours such that their union covers all but at most $3\eta n$ of the vertices of H. Then, there exists a red tight component R in H, a blue tight component B in H, a $3\sqrt{\varepsilon}$-blueprint G for H and a matching M_0 in $R \cup B$ such that the following hold, where $W_0 = V(G) \setminus V(M_0)$.*

(i) $\delta(G) \geq (1 - \alpha^{1/30})n$.

(ii) *$R(e) = R$ and $B(e') = B$ for all edges $e \in G^{\mathrm{red}}[V(M_0^{\mathrm{red}}) \cup W_0]$ and all edges $e' \in G^{\mathrm{blue}}[V(M_0^{\mathrm{blue}}) \cup W_0]$.*

(iii) $M_0 \subseteq (G^{\mathrm{red}})^+ \cup (G^{\mathrm{blue}})^+$.

(iv) *$(G^{\mathrm{red}})^+[W_0] \cup (G^{\mathrm{blue}})^+[W_0]$ is empty.*

Proof. By Lemma 2.4.3, there exists a $3\sqrt{\varepsilon}$-blueprint G_0 for H with $V(G_0) = V(H)$ and $|G_0| \geq (1 - \alpha - 96\sqrt{\varepsilon})\binom{n}{2} \geq (1 - 4\alpha)\binom{n}{2}$. By Corollary 2.4.6, there exists a subgraph G_1 of G_0 of order at least $(1 - 6\sqrt{\alpha})n$ that contains a spanning monochromatic component and $\delta(G_1) \geq (1 - 12\sqrt{\alpha})n$. Note that that G_1 is also a $3\sqrt{\varepsilon}$-blueprint for H.

We assume without loss of generality that G_1 contains a spanning red component. Since G_1 is a blueprint, all the red edges in G_1 induce the same red tight component R in H. Let $R^+ = (G_1^{\mathrm{red}})^+ \subseteq R$. Let M be a matching in R^+ of maximum size. Let $U = V(G_1) \setminus V(M)$.

Thus $|U| \geq \eta n$ (or else $|V(M)| \geq |V(G_1)| - |U| \geq (1-2\eta)n$, a contradiction). Moreover, $R^+[U] = \varnothing$. Since $\delta(G_1) \geq (1 - 12\sqrt{\alpha})n$, we have $\delta(G_1[U]) \geq |U| - \alpha^{1/3}n$. Hence, by Lemma 2.4.10 (with $4, U, \varnothing, \alpha^{1/3}$ playing the roles of k, U, S, δ), there exists a subgraph J of $G_1[U]$ such that $|J| \geq |G_1[U]| - 2\alpha^{1/13}n^2$, such that $H(e) = H(e')$ for all $e, e' \in J$ of the same colour. Let $G_2 = (G_1 - G_1^{\mathrm{blue}}[U]) \cup J$ and $B = B(e)$ for $e \in J^{\mathrm{blue}}$. Note that

$|G_2| \geq (1-\alpha^{1/14})\binom{n}{2}$. By Proposition 2.4.5, there exists a subgraph G of G_2 such that $\delta(G) \geq (1-\alpha^{1/30})n$, so (i) holds.

Let $W = V(G) \setminus V(M)$. Next, we show that (ii) and (iii) hold but with M, W instead of M_0, W_0. Note that $M^{\text{blue}} = \varnothing$, so (iii) holds by our construction. Since $G^{\text{red}} \subseteq G_1^{\text{red}}$ and G_1^{red} is connected and a blueprint, $R(e) = R$ for all $e \in G^{\text{red}}$. Note that $G^{\text{blue}}[V(M^{\text{blue}}) \cup W] = G^{\text{blue}} - V(M) \subseteq G_2^{\text{blue}}[U] = J^{\text{blue}}$, so $B(e) = B$ for all $e \in G^{\text{blue}}[V(M^{\text{blue}}) \cup W]$. Hence (ii) holds. We now add vertex-disjoint edges of $(G^{\text{red}})^+[W] \cup (G^{\text{blue}})^+[W]$ to M and call the resulting matching M_0. We deduce that M_0 satisfies (ii)–(iv). □

We now prove Lemma 2.5.1.

Proof of Lemma 2.5.1. Suppose the contrary that H does not contain two vertex-disjoint monochromatic tightly connected matchings of distinct colours such that their union covers all but at most $3\eta n$ of the vertices of H. We call this the initial assumption. Apply Lemma 2.5.2 and obtain a red tight component R, a blue tight component B in H, a $3\sqrt{\varepsilon}$-blueprint G for H and a matching M_0 in $R \cup B$ satisfying Lemma 2.5.2(i)–(iv).

We now fix G, R and B. We use the following notation for the rest of the proof. For a matching M in $R \cup B$, we set

$$\begin{aligned} W &= W(M) = V(G) \setminus V(M), \\ W_{\text{red}} &= W_{\text{red}}(M) = \{w \in W \colon d^{\text{blue}}_{G[W]}(w) \leq 8\sqrt{\varepsilon} n\}, \\ W_{\text{blue}} &= W_{\text{blue}}(M) = \{w \in W \colon d^{\text{red}}_{G[W]}(w) \leq 8\sqrt{\varepsilon} n\}. \end{aligned}$$

Note that $|W| \geq \eta n$ by the initial assumption. Without loss of generality, $|W_{\text{blue}}(M_0)| \leq |W_{\text{red}}(M_0)|$.

We define $\mathcal{M}$ be the set of matchings M in $R \cup B$ such that

(i′) $R(e) = R$ and $B(e') = B$ for all edges $e \in G^{\text{red}}[W]$ and $e' \in G^{\text{blue}}[V(M^{\text{blue}}) \cup W]$,

(ii′) $M^{\text{blue}} \subseteq (G^{\text{blue}})^+$, and

(iii′) $(G^{\text{red}})^+[W] \cup (G^{\text{blue}})^+[W]$ is empty.

Note that (i′) and (ii′) are weaker statements of those in Lemma 2.5.2(ii) and (iii), so $M_0 \in \mathcal{M}$. Let $\mathcal{M}'$ be the set of $M \in \mathcal{M}$ also satisfying

(iv′) $|W_{\text{blue}}| \leq |W_{\text{red}}|$.

Observe that $M_0 \in \mathcal{M}'$, so $\mathcal{M}'$ is nonempty.

Let $\gamma = 10\alpha^{1/30}$. We now show that, for all $M \in \mathcal{M}$, W_{red} and W_{blue} partition W, and moreover one of them is small.

Claim 2.5.3. *Let $M \in \mathcal{M}$. The following holds.*

(a) *For all $w \in W$, either $d^{\text{red}}_{G[W]}(w) \leq 7\sqrt{\varepsilon}n$ or $d^{\text{blue}}_{G[W]}(w) \leq 7\sqrt{\varepsilon}n$.*

(b) *W_{red} and W_{blue} partition W.*

(c) *Either $|W_{\text{blue}}| \leq \gamma n$ or $|W_{\text{red}}| \leq \gamma n$.*

In particular, if $M \in \mathcal{M}'$, then $|W_{\text{blue}}| \leq \gamma n$.

Proof of Claim. Suppose that there exists a vertex $w \in W$ that satisfies $d^{\text{red}}_{G[W]}(w)$, $d^{\text{blue}}_{G[W]}(w) > 7\sqrt{\varepsilon}n$. By Lemma 2.4.9 (with $7\sqrt{\varepsilon}, w, N^{\text{red}}_{G[W]}(w), N^{\text{blue}}_{G[W]}(w)$ playing the roles of $\delta, T, S^{\text{red}}, S^{\text{blue}}$), there exist $x \in N^{\text{red}}_{G[W]}(w)$ and $y \in N^{\text{blue}}_{G[W]}(w)$ such that $wxy \in \partial R \cap \partial B$. In particular, $d_H(wxy) \neq 0$ and thus $d_H(wxy) \geq (1-\varepsilon)n$, which implies that there exists a vertex $w' \in W$ such that $ww'xy \in H$. Note that $ww'xy \in (G^{\text{red}})^+[W] \cup (G^{\text{blue}})^+[W]$ contradicting (iii′). Hence, $\min\{d^{\text{red}}_{G[W]}(w), d^{\text{blue}}_{G[W]}(w)\} \leq 7\sqrt{\varepsilon}n$. Since Lemma 2.5.2(i) implies that $\delta(G[W]) \geq |W| - \alpha^{1/30}n > 16\sqrt{\varepsilon}n$, we deduce that (a) and (b) hold.

Recall that $|W| \geq \eta n > 2\gamma n$. So one of W_{red} and W_{blue} has size greater than γn. Suppose both are (that is, (c) is false). Since $\delta(G) \geq (1-\alpha^{1/30})n = (1-\gamma/10)n$, we have that there are at least

$$|W_{\text{blue}}| \left(|W_{\text{red}}| - \gamma n/10 - 8\sqrt{\varepsilon}n\right) \geq |W_{\text{blue}}| \left(|W_{\text{red}}| - \gamma n/5\right) > 3\,|W_{\text{red}}|\,|W_{\text{blue}}|\,/4$$

blue edges between W_{blue} and W_{red} and similarly there are at least $3\,|W_{\text{red}}|\,|W_{\text{blue}}|\,/4$ red edges between W_{blue} and W_{red}. Thus $e(W_{\text{red}}, W_{\text{blue}}) > |W_{\text{red}}|\,|W_{\text{blue}}|$, a contradiction. ■

Let $M_* \in \mathcal{M}'$ be such that $(|M_*|, |M_*^{\mathrm{red}}|)$ is lexicographically maximum. We write $W^*, W^*_{\mathrm{red}}, W^*_{\mathrm{blue}}$ for $W(M_*), W_{\mathrm{red}}(M_*), W_{\mathrm{blue}}(M_*)$, respectively.

The next claim shows that almost all 4-edges in $H[W^*]$ are blue and they form a tight component. Indeed, this follows from the fact that almost all edges in $G[W^*]$ are red and thus almost all triples in W^* are in ∂R.

Claim 2.5.4. *There exists a blue tight component B' in H such that the number of triples $xyz \in \binom{W^*_{\mathrm{red}}}{3} \cap \partial B'$ with $d_{B'}(xyz, W^*_{\mathrm{red}}) \geq |W^*_{\mathrm{red}}| - \varepsilon n$ is at least $(1 - \alpha^{1/31}) \left|\binom{W^*_{\mathrm{red}}}{3}\right|$.*

Proof of Claim. Let $\mathcal{T}$ be the set of triples $xyz \in \binom{W^*_{\mathrm{red}}}{3} \cap \partial R$ such that $xy \in G^{\mathrm{red}}$. Note that, for any $x \in W^*_{\mathrm{red}}$, $y \in N_G^{\mathrm{red}}(x, W^*_{\mathrm{red}})$ and $z \in N_{\partial R}(xy, W^*_{\mathrm{red}})$, we have $xyz \in \mathcal{T}$. Thus

$$\begin{aligned}|\mathcal{T}| &\geq \frac{1}{3!}|W^*_{\mathrm{red}}| \left(|W^*_{\mathrm{red}}| - \alpha^{1/30}n - 8\sqrt{\varepsilon n}\right)\left(|W^*_{\mathrm{red}}| - 3\sqrt{\varepsilon n}\right) \\ &\geq \frac{|W^*_{\mathrm{red}}|^3}{3!}\left(1 - \frac{2\alpha^{1/30}n}{|W^*_{\mathrm{red}}|}\right) \geq \left(1 - \alpha^{1/31}\right)\left|\binom{W^*_{\mathrm{red}}}{3}\right|,\end{aligned}$$

as $|W^*_{\mathrm{red}}| \geq \eta n/2$. By (iii′), we have that if $xyz \in \mathcal{T}$ and $w \in N_H(xyz, W^*_{\mathrm{red}})$, then $wxyz \in H^{\mathrm{blue}}$. For $xyz \in \mathcal{T}$, let $B(xyz)$ be the maximal blue tight component containing all the edges $xyzw$, where $w \in N_H(xyz, W^*_{\mathrm{red}})$. We say that xyz generates the blue tight component $B(xyz)$. It suffices to show that all $xyz \in \mathcal{T}$ generate the same blue tight component. First we show that triples that share two vertices generate the same blue tight component. Note that, for $xyz_1, xyz_2 \in \mathcal{T}$, we have $d_H(xyz_1, W^*_{\mathrm{red}}), d_H(xyz_2, W^*_{\mathrm{red}}) \geq |W^*_{\mathrm{red}}| - \varepsilon n > |W^*_{\mathrm{red}}|/2$ and thus there exists $w \in N_H(xyz_1) \cap N_H(xyz_2) \cap W^*_{\mathrm{red}}$. Since the edges $wxyz_1$ and $wxyz_2$ are blue, it follows that $B(xyz_1) = B(xyz_2)$.

Now let $x_1y_1z_1, x_2y_2z_2 \in \mathcal{T}$, where $x_1y_1, x_2y_2 \in G^{\mathrm{red}}$. Let $w_1 \in N_{\partial R}(x_1y_1) \cap N_{\partial R}(x_2y_2) \cap N_{G^{\mathrm{red}}}(x_1) \cap N_{G^{\mathrm{red}}}(x_2) \cap W^*_{\mathrm{red}}$ and $w_2 \in N_{\partial R}(x_1w_1) \cap N_{\partial R}(x_2w_1) \cap W^*_{\mathrm{red}}$. It follows that $x_1y_1w_1, x_1w_1w_2, x_2w_1w_2, x_2y_2w_1 \in \mathcal{T}$. Hence $B(x_1y_1z_1) = B(x_1y_1w_1) = B(x_1w_1w_2) = B(x_2w_1w_2) = B(x_2y_2w_1) = B(x_2y_2z_2)$. Let B' be the unique blue tight component generated by all triples $xyz \in \mathcal{T}$. ■

The previous claim and a greedy argument imply that there is a matching $M_*^{B'}$ in $B'[W^*]$

that covers all but ηn of the vertices in W^*. Thus we may assume that $\left|M_*^{\text{blue}}\right| \geq \eta n/4$, otherwise $\left|V(M_*^{\text{red}} \cup M_*^{B'})\right| \geq n - 3\eta n$, which is a contradiction to the initial assumption. To complete the proof, we will show that in fact $B' = B$, implying M_*^{red} and $M_*^{\text{blue}} \cup M_*^{B'}$ are tightly connected matchings, a contradiction to the initial assumption.

We now pick a special edge $e^* \in M_*^{\text{blue}}$. Its special property that we desire is stated in Claim 2.5.6.

Claim 2.5.5. *There exist an edge $e^* = v_1^* v_2^* v_3^* v_4^* \in M_*^{\text{blue}}$ and distinct vertices $w_1, \dots, w_4$, $w_1', \dots, w_4' \in W_{\text{red}}^*$ such that, for each $j \in [4]$,*

(a) *all the red edges of G incident to v_j^* induce R, or*

(b) *$v_j^* w_j \in G^{\text{blue}}$ and $v_j^* w_j w_j' \in \partial R \cap \partial B$.*

Proof of Claim. For each edge $e \in M_*^{\text{blue}}$, let $v_1^e, v_2^e, v_3^e, v_4^e$ be an enumeration of its vertices. It is easy to see that there exists $M_1^{\text{blue}} \subseteq M_*^{\text{blue}}$ with $\left|M_1^{\text{blue}}\right| = \left|M_*^{\text{blue}}\right|/16$ such that for each $j \in [4]$ we have that either

(a′) for all $e \in M_1^{\text{blue}}$, there is a red edge in G between v_j^e and W_{red}^*, or

(b′) for all $e \in M_1^{\text{blue}}$, all edges in G between v_j^e and W_{red}^* are blue.

Let J_1 be the set of $j \in [4]$ such that (a′) holds and $J_2 = [4] \setminus J_1$. Since each vertex in W_{red}^* is incident to a red edge of G that induces R and G is a blueprint for H, we have that, for all $e \in M_1^{\text{blue}}$ and all $j \in J_1$, all the red edges incident to v_j^e induce R. For every $j \in J_2$, we have that

$$\left|G^{\text{blue}}[\left\{v_j^e \colon e \in M_1^{\text{blue}}\right\}, W_{\text{red}}^*]\right| \geq \left|M_1^{\text{blue}}\right| \left(|W_{\text{red}}^*| - \alpha^{1/30} n\right) \geq \left(1 - \alpha^{1/31}\right) \left|M_1^{\text{blue}}\right| |W_{\text{red}}^*| \,.$$

Thus there exists $w_j \in W_{\text{red}}^*$ such that $w_j v_j^e$ is blue for at least $\left|M_1^{\text{blue}}\right|(1 - \alpha^{1/32})$ of the vertices v_j^e, with $e \in M_1^{\text{blue}}$. It is easy to see that we can choose the w_j to be distinct. Hence there exist distinct vertices $w_1, w_2, w_3, w_4 \in W_{\text{red}}^*$ and $M_2^{\text{blue}} \subseteq M_1^{\text{blue}}$ with $\left|M_2^{\text{blue}}\right| = \left|M_1^{\text{blue}}\right|/2 \geq \eta n/128$ such that for all $j \in J_2$ and all $e \in M_2^{\text{blue}}$ we have that $w_j v_j^e \in G^{\text{blue}}$.

For $j \in J_2$, let $V_j = \{v_j^e \colon e \in M_2^{\mathrm{blue}}\}$ and note that $d_G^{\mathrm{blue}}(w_j, V_j) = \left|M_2^{\mathrm{blue}}\right| \geq \eta n/128$ and $d_G^{\mathrm{red}}(w_j, W^*_{\mathrm{red}}) \geq \eta n/2$. For each $j \in J_2$, we apply Lemma 2.4.9 with colours reversed and $w_j, V_j, \widetilde{W}^*_{\mathrm{red}}$ playing the roles of $T, S^{\mathrm{blue}}, S^{\mathrm{red}}$ where $\widetilde{W}^*_{\mathrm{red}}$ denotes W^*_{red} with all previously chosen vertices removed. Thus, we find distinct $w'_j \in W^*_{\mathrm{red}} \setminus \{w_1, w_2, w_3, w_4\}$ and $M_3^{\mathrm{blue}} \subseteq M_2^{\mathrm{blue}}$ with $\left|M_3^{\mathrm{blue}}\right| = \left|M_2^{\mathrm{blue}}\right|/2$ such that, for all $j \in J_2$ and all $e \in M_3^{\mathrm{blue}}$, we have that $v_j^e w_j \in G^{\mathrm{blue}}$ and $v_j^e w_j w'_j \in \partial R \cap \partial B$. We complete the proof by choosing $e^* = v_1^* v_2^* v_3^* v_4^* \in M_3^{\mathrm{blue}}$ and a distinct vertex $w'_j \in W^*_{\mathrm{red}}$ for each $j \in J_1$. ■

Let $W' = W^*_{\mathrm{red}} \setminus \{w_1, \ldots, w_4, w'_1, \ldots, w'_4\}$.

Claim 2.5.6. *The graph $B[e^* \cup W']$ does not contain two vertex-disjoint edges each of which contains an edge of G^{blue} and $R[e^* \cup W']$ is empty. In particular, there do not exist two vertex-disjoint edges f_1 and f_2 in $(R \cup B)[e^* \cup W']$ each containing an edge of G^{blue}.*

Proof of Claim. First suppose there exist two vertex-disjoint edges $f_1, f_2 \in B[e^* \cup W']$ each of which contains an edge of G^{blue}. By the maximality of $|M_*|$, both f_1 and f_2 must intersect e^*. For simplicity, we only consider the case that $e^* \setminus (f_1 \cup f_2) = \{v_1^*\}$ (the other cases can be proved similarly). By Claim 2.5.5, we have that all red edges of G incident to v_1^* induce R or $v_1^* w_1 \in G^{\mathrm{blue}}$ and $v_1^* w_1 w'_1 \in \partial R \cap \partial B$.

First suppose that $v_1^* w_1 \in G^{\mathrm{blue}}$ and $v_1^* w_1 w'_1 \in \partial R \cap \partial B$. Let $w''_1 \in N_H(v_1^* w_1 w'_1, W^* \setminus (f_1 \cup f_2))$ and $f_3 = v_1^* w_1 w'_1 w''_1$. Let $M' = (M_* \setminus \{e^*\}) \cup \{f_1, f_2, f_3\}$. Note that $W(M') \subseteq W^*$. Since $|W| \geq \eta n \geq 3\gamma n$ and $|W^*_{\mathrm{blue}}| \leq \gamma n$ by Claim 2.5.3, we deduce that M' satisfies (iv′). Hence $M' \in \mathcal{M}'$ contradicting the maximality of $|M_*|$.

Now assume that all the red edges of G incident to v_1^* induce R. Let M be a matching in $R \cup B$ containing $(M_* \setminus \{e^*\}) \cup \{f_1, f_2\}$ satisfying (ii′) and (iii′). We now show that $M \in \mathcal{M}'$, which then contradicts the maximality of $|M_*|$. Recall that $v_1^* \in e^* \in M_*^{\mathrm{blue}}$, so

$$W \subseteq (W^* \setminus (f_1 \cup f_2)) \cup \{v_1^*\} \text{ and } V(M^{\mathrm{blue}}) \cup W \subseteq V(M_*^{\mathrm{blue}}) \cup W^*. \tag{2.5.1}$$

Together with our assumption on v_1^*, M satisfies (i′). Hence $M \in \mathcal{M}$. For all $w \in W \cap W^*_{\mathrm{red}}$,

$$d^{\mathrm{blue}}_{G[W]}(w) \overset{(2.5.1)}{\leq} d^{\mathrm{blue}}_{G[W^*]}(w) + |v_1^*| \overset{\text{Claim 2.5.3(a)}}{\leq} 7\sqrt{\varepsilon}n + 1 \leq 8\sqrt{\varepsilon}n.$$

and a similar inequality holds for all $w \in W \cap W^*_{\mathrm{blue}}$. This implies that $W_{\mathrm{blue}} \subseteq W^*_{\mathrm{blue}} \cup \{v^*\}$. Since $|W| \geq \eta n \geq 3\gamma n$ and $|W^*_{\mathrm{blue}}| \leq \gamma n$ by Claim 2.5.3, we deduce that M satisfies (iv′). Hence, $M \in \mathcal{M}'$ as required, a contradiction.

Therefore, $B[e^* \cup W']$ does not contain two vertex-disjoint edges each of which contains an edge of G^{blue}. If $R[e^* \cup W']$ contains an edge f, then a similar argument holds with f replacing $\{f_1, f_2\}$. Note that if $|M| = |M_*|$, then we obtain a contradiction by showing that $\left|M^{\mathrm{red}}_*\right| < \left|M^{\mathrm{red}}\right|$. ■

Since $e^* \in M^{\mathrm{blue}}_* \subseteq (G^{\mathrm{blue}})^+$, we may assume without loss of generality that $v_1^* v_2^* \in G^{\mathrm{blue}}$. The following claim shows that one of the vertices v_1^* and v_2^* has small blue degree in G to W' (and thus it has large red degree to W').

Claim 2.5.7. *We have $d^{\mathrm{blue}}_G(v_1^*, W') \leq 3\gamma n$ or $d^{\mathrm{blue}}_G(v_2^*, W') \leq 3\gamma n$.*

Proof of Claim. Suppose to the contrary that we have $d^{\mathrm{blue}}_G(v_1^*, W'), d^{\mathrm{blue}}_G(v_2^*, W') > 3\gamma n$. By Claim 2.5.6, it suffices to show that we can find two vertex-disjoint edges f_1 and f_2 in $(R \cup B)[e^* \cup W']$ each containing an edge of G^{blue}. It is easy to see that we can greedily choose vertices $x \in N^{\mathrm{blue}}_G(v_1^*, W')$, $x' \in N^{\mathrm{red}}_G(x, W') \cap N_{\partial B}(v_1^* x, W')$ and $x'' \in N_{\partial R}(xx', W') \cap N_H(v_1^* x x', W')$. Set $f_1 = v_1^* x x' x''$. By our construction, $v_1^* x x' \in \partial B$ and $x x' x'' \in \partial R$ implying $f_1 \in (R \cup B)[e^* \cup W']$. Similarly there exists an edge $f_2 = v_2^* y y' y'' \in (R \cup B)[e^* \cup W']$ disjoint from f_1 with $y, y', y'' \in W'$. ■

Without loss of generality assume $d^{\mathrm{blue}}_G(v_1^*, W') \leq 3\gamma n$ and so $d^{\mathrm{red}}_G(v_1^*, W') \geq |W'| - \alpha^{1/31} n$. Let $w \in N_{\partial B}(v_1^* v_2^*) \cap N^{\mathrm{red}}_G(v_1^*) \cap W'$, $w' \in N^{\mathrm{red}}_G(w) \cap N_{\partial R}(v_1^* w) \cap N_H(v_1^* v_2^* w) \cap W'$ and $w'' \in N_H(v_1^* w w', W')$. (We can find these vertices greedily one by one.) By Claim 2.5.4, we may further assume that $w w' w'' \in \partial B'$. By construction, we have that $v_1^* w w' \in \partial R$ and thus Claim 2.5.6 implies that both $v_1^* v_2^* w w'$ and $v_1^* w w' w''$ are blue. Since $v_1^* v_2^* w \in \partial B$, we deduce that $v_1^* v_2^* w w', v_1^* w w' w'' \in B$ and so $w w' w'' \in \partial B$ implying that $\partial B \cap \partial B' \neq \varnothing$. Therefore $B = B'$ as required. $\square$

2.6 Monochromatic connected matchings in $K_n^{(5)}$

The aim of this section is to prove the following lemma which shows that 2-edge-coloured dense 5-graphs can be almost partitioned into four monochromatic tightly connected matchings.

Lemma 2.6.1. *Let $1/n \ll \varepsilon \ll \alpha \ll \eta < 1$. Let H be a 2-edge-coloured $(1-\varepsilon, \alpha)$-dense 5-graph on n vertices. Then H contains four vertex-disjoint monochromatic tightly connected matchings such that their union covers all but at most $3\eta n$ of the vertices of H.*

Note that this implies $\mu_5^4(1, \varepsilon, n) \geq (1 - 3\eta)n/5$ for $1/n \ll \varepsilon \ll \eta < 1$. Hence $\mu_5^4(1) \geq 1/5$. Together with Corollary 2.3.12, Lemma 2.6.1 implies Theorem 1.1.3.

We use the following notation throughout this section. Let H be a 2-edge-coloured 5-graph and let G be a blueprint for H. Given a red tight component $R \subseteq H$, we write R^3 for the edges of G that induce R. We use analogous notation for blue tight components.

Let H be a 2-edge-coloured dense 5-graph. We first apply Lemma 2.4.3 to H to get a blueprint G for H. Since G is 2-edge-coloured dense 3-graph, we can apply Lemma 2.4.3 again to G to obtain a blueprint for G, which is a 2-coloured 1-graph. The following lemma summarises the structural information about H that we obtain in this way.

Lemma 2.6.2. *Let $1/n \ll \varepsilon \ll \alpha \ll 1$. Let H be a 2-edge-coloured $(1-\varepsilon, \alpha)$-dense 5-graph on n vertices. Then there exists a 3-graph G with $V(G) = V(H)$, two disjoint subsets V^{red} and V^{blue} of $V(H)$, a red tight component $R \subseteq H$ and a blue tight component $B \subseteq H$ such that the following properties hold.*

(i) *G is a $(1-\alpha^{1/37}, \alpha^{1/37})$-dense $3\sqrt{\varepsilon}$-blueprint for H.*

(ii) $\left|V(H) \setminus (V^{\text{red}} \cup V^{\text{blue}})\right| \leq \alpha^{1/75} n$.

(iii) $d_{\partial R^3}(v) \geq (1-\alpha^{1/75})n$ *for all* $v \in V^{\text{red}}$.

(iv) $d_{\partial B^3}(v) \geq (1-\alpha^{1/75})n$ *for all* $v \in V^{\text{blue}}$.

Proof. By Lemma 2.4.3, there exists a $(1-\alpha^{1/37}, \alpha^{1/37})$-dense $3\sqrt{\varepsilon}$-blueprint G for H with $V(G) = V(H)$. We apply Lemma 2.4.3 to G and obtain a $\alpha^{1/75}$-blueprint J for G with $|J| \geq (1-\alpha^{1/75})n$. Note that, as a blueprint for a 3-graph, J is a 1-graph. Hence each edge of J contains precisely one vertex. By the definition of a blueprint all the red edges of J induce the same red tight component R_G of G. Let $V^{\text{red}} = \bigcup J^{\text{red}}$. Since R_G is a red tight component of G all its edges induce the same red tight component R of H. Define V^{blue} and B analogously. $\square$

Two edges f and f' in H are *loosely connected* if there exists a sequence of edges $e_1, \ldots, e_t$ such that $e_1 = f$, $e_t = f'$ and $|e_i \cap e_{i+1}| \geq 1$ for all $i \in [t-1]$. A subgraph H' of H is *loosely connected* if every pair of edges in H' is loosely connected. A maximal loosely connected subgraph of H is called a *loose component* of H.

We now prove Lemma 2.6.1. The proof works by first finding a maximal matching in $R \cup B$, where R and B are the components given by Lemma 2.6.2, and then finding maximal connected matchings in the remaining vertices.

Proof of Lemma 2.6.1. Assume, for a contradiction, that such matchings do not exist. We call this the initial assumption. Apply Lemma 2.6.2 and obtain $V^{\text{red}}, V^{\text{blue}}, G, R^3, R, B^3, B$ and let $V^* = V^{\text{red}} \cup V^{\text{blue}}$. Since there are only few vertices in $V(H) \setminus V^*$ we ignore these vertices from the start and construct our matchings in $H[V^*]$.

We begin by choosing a matching $M \subseteq (R \cup B)[V^*]$ of maximum size. Let $U = V^* \setminus V(M)$. Note that we have $R[U] = B[U] = \varnothing$ and $|U| \geq \eta n$ by the initial assumption. Let $U^{\text{red}} = U \cap V^{\text{red}}$ and $U^{\text{blue}} = U \cap V^{\text{blue}}$. The following claim shows that if U^{red} and U^{blue} are both large, then $G[U]$ must contain many edges in R^3 or many edges in B^3.

Claim 2.6.3. *If* $\left|U^{\text{red}}\right|, \left|U^{\text{blue}}\right| \geq \alpha^{1/309}n$, *then we have that* $\max\{|R^3[U]|, |B^3[U]|\} \geq \frac{1}{2}\left|U^{\text{red}}\right|\left|U^{\text{blue}}\right||U| - 3\alpha^{1/155}n^3$.

Proof of Claim. Define a bipartite graph K_0 with vertex classes U^{red} and U^{blue} such that $x \in U^{\text{red}}$ and $y \in U^{\text{blue}}$ are joined by an edge if and only if $xy \in \partial R^3 \cap \partial B^3$. Recall that $d_{\partial R^3}(x) \geq (1-\alpha^{1/75})n$ and $d_{\partial B^3}(y) \geq (1-\alpha^{1/75})n$ for all $x \in U^{\text{red}}$ and $y \in U^{\text{blue}}$. Hence

$$|K_0| \geq \left|U^{\text{blue}}\right| \left|U^{\text{red}}\right| - \alpha^{1/75}n^2.$$

Since G is $(1-\alpha^{1/37}, \alpha^{1/37})$-dense, we have $d_G(xy, U) \geq |U| - \alpha^{1/37}n$ for $xy \in K_0$. We now colour the edges of K_0 such that $xy \in K_0$ is red if $d_{R^3}(xy, U) \geq |U| - 2\alpha^{1/76}n$ and blue if $d_{B^3}(xy, U) \geq |U| - 2\alpha^{1/76}n$. Since $K_0 \subseteq \partial R^3 \cap \partial B^3$, if $xyz \in G$ with $xy \in K_0$, then $xyz \in R^3 \cup B^3$. Hence it suffices to show that almost all edges of K_0 are of the same colour. Indeed, if we have that at least $\left|U^{\text{red}}\right| \left|U^{\text{blue}}\right| - 3\alpha^{1/154}n^2$ edges of K_0 are red, then we have

$$\left|R^3[U]\right| \geq \frac{1}{2}(\left|U^{\text{red}}\right| \left|U^{\text{blue}}\right| - 3\alpha^{1/154}n^2)(|U| - 2\alpha^{1/76}n) \geq \frac{1}{2}\left|U^{\text{red}}\right| \left|U^{\text{blue}}\right| |U| - 3\alpha^{1/155}n^3.$$

We show that each edge $xy \in K_0$ is coloured either red or blue. It suffices to show that either $d_{R^3}(xy, U) < \alpha^{1/76}n$ or $d_{B^3}(xy, U) < \alpha^{1/76}n$. Indeed if $d_{R^3}(xy, U), d_{B^3}(xy, U) \geq \alpha^{1/76}n$, then by Lemma 2.4.9, there exists $u, u' \in U$ such that $xyu \in R^3$, $xyu' \in B^3$ and $xyuu' \in \partial R \cap \partial B$. For any $u'' \in N_H(xyuu', U)$, we would have $xyuu'u'' \in R[U] \cup B[U]$, a contradiction to the maximality of M. Moreover, by Lemma 2.4.8, we have that $\min\{d^{\text{red}}_{K_0}(u), d^{\text{blue}}_{K_0}(u)\} \leq \alpha^{1/76}n$ for all $u \in U$.

Let K_1 be the graph obtained from K_0 by, for each $u \in U$, deleting all red edges incident to u if $d^{\text{red}}_{K}(u) \leq \alpha^{1/76}n$ and all blue edges incident to u if $d^{\text{blue}}_{K}(u) \leq \alpha^{1/76}n$. Note that $|K_1| \geq \left|U^{\text{red}}\right| \left|U^{\text{blue}}\right| - \alpha^{1/77}n^2$ and that, in K_1, each vertex is incident to only edges of one colour. It is not too hard to see that by deleting at most $2\alpha^{1/154}n^2$ additional edges, we can obtain a subgraph K_2 of K_1 for which each vertex has degree 0 or large degree. More precisely, for all $u \in U^{\text{red}}$,

$$d_{K_2}(u) \geq \left|U^{\text{blue}}\right| - 3\alpha^{1/308}n \text{ or } d_{K_2}(u) = 0$$

and, for all $u \in U^{\text{blue}}$,

$$d_{K_2}(u) \geq \left|U^{\text{red}}\right| - 3\alpha^{1/308}n \text{ or } d_{K_2}(u) = 0.$$

Since each vertex is incident to only edges of one colour and any two vertices in U^{red} that have non-zero degree have a common neighbour this implies that all edges in K_2 are of the same colour. Since $|K_2| \geq \left|U^{\text{red}}\right|\left|U^{\text{blue}}\right| - 3\alpha^{1/154}n^2$, this concludes the proof. ■

The following claim shows that there is a red tight component R_* and a blue tight component B_* of H such that almost all the edges in $G[U]$ induce one of these components.

Claim 2.6.4. *Let $\gamma = \alpha^{1/1110}$. There exists a red tight component R_* and a blue tight component B_* of H such that*

(i) $|R_*^3[U]| \geq \left|G^{\text{red}}[U]\right| - 8\gamma^{1/5}n^3$ *and* $|B_*^3[U]| \geq \left|G^{\text{blue}}[U]\right| - 8\gamma^{1/5}n^3$,

(ii) $|(R_*^3 \cup B_*^3)[U]| \geq (1-\gamma^{1/6})\binom{|U|}{3}$ *and*

(iii) $R_* = R$ *or* $B_* = B$.

Proof of Claim. First we show that, for each $u \in U$, there exists $J_u \subseteq G_u[U]$, where G_u is the link graph of G at u, such that $|J_u| \geq |G_u[U]| - \alpha^{1/14}n^2$ and $R(e \cup u) = R(e' \cup u)$ for $e, e' \in J_u^{\text{red}}$ and $B(e \cup u) = B(e' \cup u)$ for $e, e' \in J_u^{\text{blue}}$.

To show this fix $u \in U$. Without loss of generality assume that $u \in U^{\text{red}}$. By Lemma 2.6.2, $d_{\partial R^3}(u, U) \geq |U| - \alpha^{1/75}n$. Let $U_* = N_{\partial R^3}(u, U)$. Clearly, $|U_*| \geq \eta n/2$ and $G_u^{\text{red}}[U_*] \subseteq R_u^3$. Moreover, for all $x \in U_*$, we have $d_G(ux) > 0$ and thus, since G is $(1-\alpha^{1/37}, \alpha^{1/37})$-dense, $d_G(ux) \geq (1-\alpha^{1/37})n$. It follows that $\delta(G_u[U_*]) \geq |U_*| - \alpha^{1/37}n$. Thus by applying Lemma 2.4.10 with $R^3, u, U_*, \alpha^{1/37}$ playing the roles of R_G, S, U, δ, there exists $J_u \subseteq G_u[U_*] \subseteq G_u[U]$ such that

$$|J_u| \geq |G_u[U_*]| - 7\alpha^{1/148}n^2 \geq |G_u[U]| - \alpha^{1/75}n^2 - 7\alpha^{1/148}n^2 \geq |G_u[U]| - \alpha^{1/149}n^2$$

and $H(u \cup e) = H(u \cup e')$ for $e, e' \in J_u$ of the same colour.

Now consider the auxiliary multi-3-graph $D = \bigcup_{u \in U} \{e \cup u : e \in J_u\}$. Note that

$$|D| = \sum_{u \in U} |J_u| \geq \sum_{u \in U} \left(|G_u[U]| - \alpha^{1/149} n^2 \right) \geq 3\, |G[U]| - \alpha^{1/149} n^3.$$

Let F be the subgraph of $G[U]$ for which $e \in F$ if and only if e is an edge of multiplicity 3 in D. Since G is $(1-\alpha^{1/37}, \alpha^{1/37})$-dense, Proposition 2.1.1 implies that $|G| \geq (1-2\alpha^{1/37})\binom{n}{3}$. Hence

$$\begin{aligned} |G[U]| &\geq \binom{|U|}{3} - 2\alpha^{1/37}\binom{n}{3} \geq \binom{|U|}{3} - 2\alpha^{1/37}\binom{|U|/\eta}{3} \\ &\geq \binom{|U|}{3} - \frac{4\alpha^{1/37}}{\eta^3}\binom{|U|}{3} \geq (1-\alpha^{1/38})\binom{|U|}{3}. \end{aligned}$$

Therefore $|F| \geq |G[U]| - \alpha^{1/149} n^3 \geq (1 - \alpha^{1/150})\binom{|U|}{3}$. Recall that $\gamma = \alpha^{1/1110}$. By Propositions 2.1.1 and 2.1.2, there exists a $(1-\gamma^{1/5}, \gamma^{1/5})$-dense subgraph $\widetilde{F} \subseteq F$ with $V(\widetilde{F}) = V(F) = U$ and, by Proposition 2.1.1, $\left|\widetilde{F}\right| \geq (1-2\gamma^{1/5})\binom{|U|}{3}$. Hence $\left|\widetilde{F}^{\text{red}}\right| \geq \left|G^{\text{red}}[U]\right| - 2\gamma^{1/5} n^3$. Let $S^{\text{red}} = \{x \in U : d_{\widetilde{F}^{\text{red}}}(x) \geq 6\gamma^{1/5} n^2\}$. Let F_0^{red} be the subgraph of $\widetilde{F}^{\text{red}}$ consisting of all edges that contain a vertex in S^{red}. Note that $\left|F_0^{\text{red}}\right| \geq \left|\widetilde{F}^{\text{red}}\right| - 6\gamma^{1/5} n^3 \geq \left|G^{\text{red}}[U]\right| - 8\gamma^{1/5} n^3$.

We claim that all the edges in F_0^{red} induce the same red tight component R_* in H. Let $e, e' \in \widetilde{F}^{\text{red}}$ with $u \in e \cap e'$. Note that $e \setminus u, e' \setminus u \in J_u^{\text{red}}$ and so $R(e) = R(e')$. Hence edges in the same loose component of $\widetilde{F}^{\text{red}}$ induce the same red tight component in H. In particular, since $F_0^{\text{red}} \subseteq \widetilde{F}^{\text{red}}$, for $u \in S^{\text{red}}$, all the edges in $N_{F_0^{\text{red}}}(u)$ induce the same red tight component $R(u)$ of H.

Let $u, v \in S^{\text{red}}$. We want to show that $R(u) = R(v)$. We may assume that u and v are in distinct loose components L and L' of $\widetilde{F}^{\text{red}}$, respectively. In particular, any edge of $\widetilde{F}$ that intersects both $V(L)$ and $V(L')$ is in $\widetilde{F}^{\text{blue}}$. If $u, v \in V^{\text{red}}$, then $d_{\partial R^3}(u), d_{\partial R^3}(v) \geq (1-\alpha^{1/75})n$ implying $R(u) = R = R(v)$. Thus we may assume that one of u and v is in V^{blue}, say $v \in V^{\text{blue}}$. Let $\Gamma_L(u) = \{u' \in V(L) : d_L(uu') \geq \gamma^{1/5} n\}$ and $\Gamma_{L'}(v) = \{v' \in V(L') : d_{L'}(vv') \geq \gamma^{1/5} n\}$. It is easy to see that $|\Gamma_L(u)|, |\Gamma_{L'}(v)| \geq 5\gamma^{1/5} n$.

Let D' be the bipartite directed graph with parts $\Gamma_L(u)$ and $\Gamma_{L'}(v)$ such that, for $u' \in \Gamma_L(u)$,

$$N^+_{D'}(u') = \{v' \in \Gamma_{L'}(v) \colon uu'v' \in \widetilde{F}^{\text{blue}} \text{ and } uu'u''v' \in \partial R(uu'u'') \cap \partial B(uu'v') \\ \text{and } uu'u''vv' \in H \text{ for some } u'' \in N_L(uu')\},$$

and, for $v' \in \Gamma_{L'}(v)$,

$$N^+_{D'}(v') = \{u' \in \Gamma_L(u) \colon vv'u' \in \widetilde{F}^{\text{blue}} \text{ and } u'v \in \partial B^3 \text{ and } vv'v''u' \in \partial B \cap \partial R(vv'v'') \\ \text{and } vv'v''uu' \in H \text{ for some } v'' \in N_{L'}(vv')\}.$$

By Lemma 2.4.9, the fact that $\widetilde{F}$ is $(1-\gamma^{1/5}, \gamma^{1/5})$-dense and the fact that H is $(1-\varepsilon, \alpha)$-dense, we have, for $u' \in \Gamma_L(u)$,

$$d^+_{D'}(u') \geq |\Gamma_{L'}(v)| - \gamma^{1/5}n - \varepsilon^{1/4}n - \varepsilon n > |\Gamma_{L'}(v)|/2.$$

Similarly, also using the fact that $d_{\partial B^3}(v) \geq (1-\alpha^{1/75})n$, we have, for $v' \in \Gamma_{L'}(v)$,

$$d^+_{D'}(v') \geq |\Gamma_L(u)| - \gamma^{1/5}n - \alpha^{1/75}n - \varepsilon^{1/4}n - \varepsilon n > |\Gamma_L(u)|/2.$$

It follows that D' contains a double edge $u'v'$, where $u' \in \Gamma_L(u)$ and $v' \in \Gamma_{L'}(v)$. Let $u'' \in N_L(uu')$ and $v'' \in N_{L'}(vv')$ be the vertices that are guaranteed to exist by the definition of D'. Since $u'v \in \partial B^3$, we have that $vv'u \in B^3$ and thus also $uu'v' \in B^3$. As $B[U] = \varnothing$, we have $vv'v''uu', uu'u''vv' \in H^{\text{red}}$ and thus $R(uu'u'') = R(vv'v'')$. Hence $R(u) = R(v)$. We define F_0^{blue} and B_* in an analogous way. This proves (i).

Note that (ii) follows from (i) using the facts $|U| \geq \eta n$ and $|G[U]| \geq (1-\alpha^{1/38})\binom{|U|}{3}$, which were noted earlier in this proof.

We will now prove (iii). We distinguish between two cases.

Case 1: $\left|U^{\text{red}}\right|, \left|U^{\text{blue}}\right| \geq \gamma^{1/13}n$.

By Claim 2.6.3, we have $\max\{|R^3[U]|, |B^3[U]|\} \geq \frac{1}{2}\left|U^{\text{red}}\right|\left|U^{\text{blue}}\right||U| - 3\alpha^{1/155}n^3$. Since

$\frac{1}{2}\left|U^{\text{red}}\right|\left|U^{\text{blue}}\right||U| - 3\alpha^{1/155}n^3 \geq \frac{1}{2}\gamma^{2/13}\eta n^3 - 3\alpha^{1/155}n^3 \geq 2\gamma^{1/6}n^3$, we have $R^3_* \cap R^3 \neq \varnothing$ or $B^3_* \cap B^3 \neq \varnothing$ and thus $R_* = R$ or $B_* = B$.

Case 2: $\left|U^{\text{blue}}\right| \leq \gamma^{1/13}n$ or $\left|U^{\text{red}}\right| \leq \gamma^{1/13}n$.
Say $\left|U^{\text{blue}}\right| \leq \gamma^{1/13}n$. Then $\left|U^{\text{red}}\right| = |U| - \left|U^{\text{blue}}\right| \geq |U| - \gamma^{1/13}n$. Let $Q^3 = \{T \in \binom{U}{3} : \binom{T}{2} \cap \partial R^3 \neq \varnothing\}$. Since $d_{\partial R}(u, U) \geq |U| - \alpha^{1/75}n$ for $u \in U^{\text{red}}$, there can be at most $\left|U^{\text{red}}\right| \alpha^{2/75}n^2$ triples that intersect U^{red} and are not in Q^3. Hence

$$\begin{aligned}\left|Q^3\right| &\geq \binom{|U|}{3} - \left|U^{\text{blue}}\right|^3 - \left|U^{\text{red}}\right|\alpha^{2/75}n^2 \\ &\geq \binom{|U|}{3} - \gamma^{3/13}n^3 - \alpha^{2/75}n^3 \geq \binom{|U|}{3} - 2\gamma^{1/5}n^3.\end{aligned}$$

Note that $|R^3[U]| \geq \left|Q \cap G^{\text{red}}[U]\right| \geq \left|G^{\text{red}}[U]\right| - 2\gamma^{1/5}n^3$. Therefore, we have $R_* = R$. ■

We define $R_\diamond = R \cup R_*$ and $B_\diamond = B \cup B_*$. Note that, by Claim 2.6.4(iii), $R_\diamond \cup B_\diamond$ is the union of at most three monochromatic tight components. Let $M_\diamond$ be a maximal matching in $(R_\diamond \cup B_\diamond)[V^*]$ containing M. Let $W = V^* \setminus V(M_\diamond)$. Since $M \subseteq M_\diamond$, we have $W \subseteq U$. By the initial assumption, we have $|W| \geq \eta n$. Note that $(R_* \cup B_*)[W] = \varnothing$ and, since $W \subseteq U$, $(R_* \cup B_*)[W] \geq \binom{|W|}{3} - \gamma^{1/6}n^3$. The following claim shows that almost all the edges in $G[W]$ are of the same colour.

Claim 2.6.5. *We have* $|R^3_*[W]| \geq \binom{|W|}{3} - \gamma^{1/9}n^3$ *or* $|B^3_*[W]| \geq \binom{|W|}{3} - \gamma^{1/9}n^3$.

Proof of Claim. Let $G_* = R^3_* \cup B^3_*$. We define

$$\begin{aligned}W_{\text{red}} &= \{u \in W : d_{G_*}(u, W) \geq 2\alpha n^2 \text{ and } d_{B^3_*}(u, W) < \alpha n^2\}, \\ W_{\text{blue}} &= \{u \in W : d_{G_*}(u, W) \geq 2\alpha n^2 \text{ and } d_{R^3_*}(u, W) < \alpha n^2\}, \\ W_0 &= \{u \in W : d_{G_*}(u, W) < 2\alpha n^2\}.\end{aligned}$$

Since $(R_* \cup B_*)[W] = \varnothing$, by Lemma 2.4.8, $W_{\text{red}}, W_{\text{blue}}$ and W_0 partition W. Let J be the subgraph of $G_*[W]$ obtained by deleting all red edges containing a vertex in $W_{\text{blue}} \cup W_0$ and all blue edges containing a vertex in $W_{\text{red}} \cup W_0$. Note that $|J| \geq |G_*[W]| - 2\alpha n^3 \geq$

$(1-\gamma^{1/7})\binom{|W|}{3}$ and $J \subseteq \binom{W_{\mathrm{red}}}{3} \dot\cup \binom{W_{\mathrm{blue}}}{3}$. Hence

$$(1-\gamma^{1/7})\binom{|W|}{3} \le \binom{|W_{\mathrm{red}}|}{3} + \binom{|W_{\mathrm{blue}}|}{3}. \tag{2.6.1}$$

Suppose that $|W_{\mathrm{red}}|, |W_{\mathrm{blue}}| \le (1-\alpha^{1/8})|W|$. By (2.6.1), we may assume without loss of generality assume that $|W_{\mathrm{red}}| \ge |W|/2$. Noting that $x \mapsto x^3 + (|W|-x)^3$ is an increasing function for $x \ge |W|/2$ we have

$$\begin{aligned}\binom{|W_{\mathrm{red}}|}{3} + \binom{|W_{\mathrm{blue}}|}{3} &\le \frac{1}{6}\left(|W_{\mathrm{red}}|^3 + |W_{\mathrm{blue}}|^3\right) \le \frac{1}{6}\left(|W_{\mathrm{red}}|^3 + (|W| - |W_{\mathrm{red}}|)^3\right) \\ &\le ((1-\gamma^{1/8})^3 + \gamma^{3/8})\frac{|W|^3}{6} < (1-\gamma^{1/7})\binom{|W|}{3},\end{aligned}$$

a contradiction to (2.6.1).

Hence at least one of W_{red} and W_{blue} has size at least $(1-\gamma^{1/8})|W|$. Without loss of generality assume $|W_{\mathrm{red}}| \ge (1-\gamma^{1/8})|W|$. Note that any edge of J contained in W_{red} is in R^3_*, hence

$$\left|R^3_*[W]\right| \ge |J| - \left|W \setminus W^{\mathrm{red}}\right| n^2 \ge \binom{|W|}{3} - \gamma^{1/9} n^3.$$

This proves the claim. ■

Now assume without loss of generality that $|R^3_*[W]| \ge \binom{|W|}{3} - \gamma^{1/9}n^3$. Note that almost all edges in $H[W]$ are blue (otherwise there would have to be an edge in $R_*[W]$, which would contradict the maximality of M). More precisely, we have

$$\left|H^{\mathrm{blue}}[W]\right| \ge \frac{3!}{5!}|R_*[W]|(|W| - 3\sqrt{\varepsilon}n)(|W| - \varepsilon n) \ge (1-\gamma^{1/10})\binom{|W|}{5}.$$

By Propositions 2.1.1 and 2.1.2, there exists a $(1-\gamma^{1/1010}, \gamma^{1/1010})$-dense tightly connected subgraph $\widetilde{H}^{\mathrm{blue}}$ of $H^{\mathrm{blue}}[W]$ with $V(\widetilde{H}^{\mathrm{blue}}) = W$ and $\left|\widetilde{H}^{\mathrm{blue}}\right| \ge (1-2\gamma^{1/1010})\binom{|W|}{5}$. By an easy greedy argument, there exists a matching M' in $\widetilde{H}^{\mathrm{blue}}$ that covers all but at most ηn of the vertices in W. The matching $M' \cup M_\diamond$ covers all but at most $3\eta n$ of the vertices of H. This contradicts the initial assumption. $\square$

2.7 Concluding remarks

For $k \geq 3$, let $f(k)$ be the minimum integer m such that, for all large 2-edge-coloured complete k-graphs, there exists m vertex-disjoint monochromatic tight cycles covering almost all vertices. Note that $f(k)$ is well defined by [27] but the bound is very large. It is easy to see that $f(k) \geq 2$ for all $k \geq 3$. Indeed, consider the k-graph $H = K^{(k)}(A, B)$ given in Example 2.4.2 with $|A| = \frac{3k-1}{3k}n$. Note that $H[A]$ is a red tight component. Moreover, note that any tight cycle contained in a monochromatic tight component other than $H[A]$ covers at most about a third of the vertices of H and any tight cycle in $H[A]$ leaves all $\frac{n}{3k}$ vertices in B uncovered. Hence no monochromatic tight cycle covers almost all vertices in H. We have $f(3) = 2$ by [28]. Theorems 1.1.2 and 1.1.3 imply $f(4) = 2$ and $f(5) \leq 4$, respectively. In general, we believe that $f(k) = 2$ for all k. However, we believe that new ideas may be needed as indicated by again considering the k-graph $H = K^{(k)}(A, B)$ with $|A| = \frac{3k-1}{3k}n$ (as above). If H contains two vertex-disjoint monochromatic tight cycles of distinct colour covering almost all vertices, then one of the two cycles must lie entirely in the red tight component $H[A]$. However, this tight component is not induced by any edge in the blueprint of H (which is $K^{(k-2)}(A, B)$ with colours swapped). Thus we ask the weaker question of whether one can bound $f(k)$ by some suitable function of k.

CHAPTER 3

THE RAMSEY NUMBER FOR 4-UNIFORM TIGHT CYCLES

Our aim in this chapter is to prove Theorem 1.2.1. We recall some definitions and then restate the theorem. Recall that for a k-graph H, we define the *Ramsey number of H*, denoted by $r(H)$, to be the least positive integer N such that any 2-edge-coloured complete graph on N vertices contains a monochromatic copy of H. Also recall that the *k-uniform tight cycle $C_n^{(k)}$* is defined to be the k-graph on n vertices with a cyclic ordering of its vertices such that its edges are the k-sets of consecutive vertices in the ordering. We now restate Theorem 1.2.1.

Theorem 1.2.1. *Let $\varepsilon > 0$. For n large enough we have $r(C_{4n}^{(4)}) \leq (5 + \varepsilon)n$.*

The remainder of this chapter is organised as follows. In the next section we give a sketch of the proof of Theorem 1.2.1. In Section 3.2, we introduce basic notation and definitions. In Section 3.3, we define blow-ups and prove some basic propositions about them. In Section 2.4, we define blueprints, state a result about their existence and prove some basic results about how they interact with blow-ups. In Section 3.5, we prove that an almost complete 2-edge-coloured 4-graph contains a monochromatic tightly connected fractional matching with large weight. In Section 3.6, we show how to use this to prove Theorem 1.2.1.

3.1 Sketch of the proof of Theorem 1.2.1

We now sketch the proof of Theorem 1.2.1. We use a hypergraph version of the connected matching method of Łuczak [84] as follows. We consider a red-blue edge-colouring of $K_N^{(4)}$ for $N = (5/4 + \varepsilon)n$. We begin by applying the Hypergraph Regularity Lemma. More precisely, we use the Regular Slice Lemma of Allen, Böttcher, Cooley and Mycroft [2]. This gives us a reduced graph $\mathcal{R}$, which is a red-blue edge-coloured almost complete 4-graph on $(5/4 + \varepsilon)n'$ vertices. To prove Theorem 1.2.1, it now suffices to find a *monochromatic tightly connected matching* of size $n'/4$ in $\mathcal{R}$. A monochromatic tightly connected matching is a monochromatic matching M such that for any two edges $f, f' \in M$, there exists a tight walk[1] in $\mathcal{R}$ of the same colour as M connecting f and f'. This reduction of our problem to finding a monochromatic tightly connected matching in the reduced graph is formalised in Corollary 2.3.12.

Let γ be a constant such that $0 < \gamma \ll \varepsilon$ and let M be a maximal monochromatic tightly connected matching in $\mathcal{R}$. Suppose that M has size less than $n'/4$ and is red. We show that we can find a monochromatic tightly connected matching of size at least $|M| + \gamma n'$. By iterating this we get our desired result. We actually find a *fractional matching* instead. Note that by taking a blow-up of $\mathcal{R}$ we can then convert it back into an integral matching. For simplicity, let us further assume that $\mathcal{R}$ has only one red and one blue tight component (see Section 3.2 for the definition). Then any monochromatic matching is tightly connected. Consider an edge $f \in M$ and a vertex w not covered by M. Observe that if all the edges in $\mathcal{R}[f \cup \{w\}]$ are red, then we get a larger red fractional matching (by giving weight 1/4 to each of the five edges in $\mathcal{R}[f \cup \{w\}]$). Thus for almost all the edges $f \in M$ there is a blue edge f' such that $|f \cap f'| = 3$. This gives us a blue matching M' of almost the same size as M. Note that the set of leftover vertices $W = V(\mathcal{R}) \setminus V(M \cup M')$ has size at least $\varepsilon n'$. By the maximality of M, any edge in $\mathcal{R}[W]$ must be blue. So we can extend M' by adding a matching in W to get the desired

[1] A *tight walk* in a k-graph is a sequence of edges $e_1, \ldots, e_t$ such that $|e_i \cap e_{i+1}| = k - 1$ for all $1 \le i \le t - 1$.

matching of size at least $|M| + \gamma n'$.

However, $\mathcal{R}$ may contain many monochromatic tight components (instead of just two). Hence we need to choose monochromatic tight components carefully. To do this we use a novel auxiliary graph called the blueprint which we will also use in Chapter 2. The blueprint is a graph with the key property that monochromatic connected components in it correspond to monochromatic tight components in the 4-graph we are considering. Since the blueprint is red-blue edge-coloured and almost complete, it contains an almost-spanning monochromatic tree. Using the key property of blueprints this shows that $\mathcal{R}$ contains a large monochromatic tight component.

3.2 Preliminaries

In this chapter we reuse definitions from Chapter 2. If H is a k-graph and $\mathcal{P} = \{V_1, \dots V_s\}$ a partition of $V(H)$, then we call an edge $e \in H$ *$\mathcal{P}$-partite* if $|V_i \cap e| \leq 1$ for all $i \in [s]$. If all edges of H are $\mathcal{P}$-partite, then we call H *$\mathcal{P}$-partite.* We call H s-partite if H is $\mathcal{P}$-partite for some partition $\mathcal{P}$ of $V(H)$ into s sets. For vertex-disjoint k-graphs H_1 and H_2, we define the k-graph $H_1 \cup H_2 = (V(H_1) \cup V(H_2), E(H_1) \cup E(H_2))$.

A *fractional matching* in a k-graph H is a function $\varphi\colon E(H) \to [0,1]$ such that for every $v \in V(H)$, $\sum_{e \in E(H) : v \in e} \varphi(e) \leq 1$. For each $e \in E(H)$ we call $\varphi(e)$ the *weight of* e. The *weight of* φ is $\sum_{e \in E(H)} \varphi(e)$. For a positive integer r, a $1/r$*-fractional matching* φ in a k-graph H is a fractional matching such that each edge has weight in $\{0, \frac{1}{r}, \frac{2}{r}, \dots, \frac{r-1}{r}, 1\}$, that is $\{\varphi(e)\colon e \in E(H)\} \subseteq \{0, \frac{1}{r}, \frac{2}{r}, \dots, \frac{r-1}{r}, 1\}$. For vertex-disjoint k-graphs H_1 and H_2 and fractional matchings φ_1 and φ_2 in H_1 and H_2, respectively, we define the fractional matching $\varphi_1 + \varphi_2\colon E(H_1 \cup H_2) \to [0,1]$ in $H_1 \cup H_2$ by setting $(\varphi_1 + \varphi_2)(e) = \varphi_1(e)$ if $e \in H_1$ and $(\varphi_1 + \varphi_2)(e) = \varphi_2(e)$ if $e \in H_2$. For a k-graph H, a subgraph H' of H and a fractional matching φ in H', we define the *completion of* φ *with respect to* H, denoted φ^H, to be the fractional matching $\varphi^H\colon E(H) \to [0,1]$ such that $\varphi^H(e) = \varphi(e)$ if $e \in H'$ and $\varphi(e) = 0$ otherwise. For a matching M in a k-graph H, we define the *fractional matching* *induced by the matching* M to be the fractional matching $\varphi\colon E(M) \to [0,1]$ with $\varphi(e) = 1$ for all $e \in M$. A *tightly connected fractional matching* in a k-graph H is a fractional matching $\varphi\colon E(H') \to [0,1]$ where H' is a tight component of H.

A *red tight component*, a *red tightly connected matching* and a *red tightly connected fractional matching* in a 2-edge-coloured k-graph H are a tight component, a tightly connected matching and a tightly connected fractional matching, respectively, in H^{red}. We define these terms analogously for blue. A *monochromatic tight component* in H is a red or a blue tight component in H and similarly for the other terms.

3.3 Blow-ups

We will later need blow-ups to convert fractional matchings to integral ones. So we define blow-ups here and show some basic facts.

Definition 3.3.1. Given a k-graph H we say that H_* is a blow-up of H if there exists a partition $\mathcal{P} = \{V_x \colon x \in V(H)\}$ of $V(H_*)$ such that $H_* = \bigcup_{x_1\dots x_k \in H} K_{V_{x_1},\dots,V_{x_k}}$, where $K_{V_{x_1},\dots,V_{x_k}}$ is the complete k-partite k-graph with vertex classes $V_{x_1}, \dots, V_{x_k}$. Moreover, if H is 2-edge-coloured, then we have

$$H_*^{\mathrm{red}} = \bigcup_{x_1\dots x_k \in H^{\mathrm{red}}} K_{V_{x_1},\dots,V_{x_k}} \text{ and } H_*^{\mathrm{blue}} = \bigcup_{x_1\dots x_k \in H^{\mathrm{blue}}} K_{V_{x_1},\dots,V_{x_k}}.$$

If $|V_x| = r$ for all $x \in V(H)$, then we call H_* an *r-blow-up* of H. If $e_* = y_1 \dots y_k \in H_*$, then we let $f_{e_*} = x_1 \dots x_k$ be the unique edge in H such that $y_i \in V_{x_i}$ for all $i \in [k]$.

Recall that a k-graph H on n vertices is called (μ, α)-*dense* if, for each $i \in [k-1]$, we have $d_H(S) \geq \mu\binom{n}{k-i}$ for all but at most $\alpha\binom{n}{i}$ sets $S \in \binom{V(H)}{i}$ and $d_H(S) = 0$ for all other sets $S \in \binom{V(H)}{i}$. The following proposition shows that the r-blow-up of a $(1-\varepsilon, \alpha)$-dense k-graph is $(1-2\varepsilon, 2\alpha)$-dense.

Proposition 3.3.2. *Let* $1/n \ll \varepsilon, \alpha, 1/r, 1/k$. *Let* H *be a* $(1-\varepsilon, \alpha)$*-dense* k*-graph on* n *vertices and let* H_* *be an* r*-blow-up of* H. *Then* H_* *is* $(1-2\varepsilon, 2\alpha)$*-dense.*

Proof. Let $\mathcal{P} = \{V_x : x \in V(H)\}$ be the partition of $V(H_*)$. Let $i \in [k-1]$ and $S \in \binom{V(H_*)}{i}$ be such that S is $\mathcal{P}$-partite. Let $S' = x_1 \dots x_i \in \binom{V(H)}{i}$ be such that $S \in K_{V_{x_1},\dots,V_{x_i}}$. Suppose $d_H(S') \geq (1-\varepsilon)\binom{n}{k-i}$. Then $d_{H_*}(S) \geq (1-\varepsilon)\binom{n}{k-i}r^{k-i} \geq (1-2\varepsilon)\binom{nr}{k-i}$, since $\binom{nr}{k-i} = (1+o(1))\binom{n}{k-i}r^{k-i}$ as $n \to \infty$. The number of sets $S \in \binom{V(H_*)}{i}$ for which this is true is at least $r^i(1-\alpha)\binom{n}{i} \geq (1-2\alpha)\binom{nr}{i}$. For all other sets $S \in \binom{V(H_*)}{i}$, we have $d_{H_*}(S) = 0$. Hence H_* is $(1-2\varepsilon, 2\alpha)$-dense. □

The following proposition shows how to turn a matching in a blow-up of a k-graph H into a fractional matching in H and vice versa.

Proposition 3.3.3. *Let $1/N \ll \varepsilon \ll 1/r$ and $k \geq 2$. Let H be an edge-coloured k-graph on N vertices and let H_* be an r-blow-up of H. Then H_* contains a monochromatic tightly connected matching M of size m if and only if H contains a monochromatic tightly connected $1/r$-fractional matching of weight m/r of the same colour.*

Proof. Let $\mathcal{P} = \{V_x : x \in V(H)\}$ be the partition of $V(H_*)$. Let F_* be the monochromatic tight component of H_* that contains M. There exists a monochromatic tight component F of H such that

$$F_* = \bigcup_{x_1 \dots x_k \in F} K_{V_{x_1},\dots,V_{x_k}}.$$

We define the fractional matching $\varphi\colon F \to [0,1]$ as follows. For each edge $e = x_1 \dots x_k \in F$, we set

$$\varphi(e) = \frac{\left|M \cap K_{V_{x_1},\dots,V_{x_k}}\right|}{r}.$$

For each $x \in V(H)$,

$$\sum_{e \in H\colon x \in e} \varphi(e) = \frac{1}{r}\left|\{f \in M\colon f \cap V_x \neq \varnothing\}\right| \leq 1,$$

since M is a matching and $|V_x| = r$. Hence φ is a monochromatic tightly connected $1/r$-fractional matching. We conclude by noting that φ has weight m/r.

The other direction is proved similarly. □

3.4 Blueprints and blow-ups

We use the notion of blueprint introduced in Section 2.4, which allows us to track monochromatic tight components. In this section we show that blueprints work well together with blow-ups.

Proposition 3.4.1. *Let $1/n \ll \varepsilon \ll 1/k \leq 1/4$ and let $r \geq 2$ be an integer. Let H be a 2-edge-coloured k-graph on n vertices, let G be an ε-blueprint for H, and H_* an r-blow-up of H with vertex partition $\mathcal{P} = \{V_x : x \in V(H)\}$. Let $G_* = G_*(H, H_*, G)$ be the r-blow-up of G with vertex partition $\mathcal{P}' = \{V_x : x \in V(G)\}$. Then G_* is an ε-blueprint for H_*. Moreover, for $e_* = y_1 \dots y_{k-2} \in G_*$ and $f_{e_*} = x_1 \dots x_{k-2} \in G$ where $y_i \in V_{x_i}$ for all $i \in [k-2]$, we have $H_*(e_*) = \bigcup_{z_1 \dots z_k \in H(f_{e_*})} K_{V_{z_1}, \dots, V_{z_k}}$, that is $H_*(e_*)$ is the r-blow-up of $H(f_{e_*})$ in H_*.*

Proof. For $e_* = y_1 \dots y_{k-2} \in G_*$, we let $H_*(e_*)$ be the blow-up of $H(f_{e_*})$ with respect to $\mathcal{P}$, that is $H_*(e_*) = \bigcup_{z_1 \dots z_k \in H(f_{e_*})} K_{V_{z_1}, \dots, V_{z_k}}$. Since $H(f_{e_*})$ is a monochromatic tight component in H, $H_*(e_*)$ is indeed a monochromatic tight component in H_* as required. Moreover, since f_{e_*} has the same colour as e_*, $H(e_*)$ has the same colour as e_*.

Let $e_* \in G_*$. We show that $d_{\partial H_*(e_*)}(e_*) \geq (1-\varepsilon)nr$. Since $H_*(e_*)$ is the blow-up of $H(f_{e_*})$ with respect to $\mathcal{P}$, $\partial H_*(e_*)$ is the blow-up of $\partial H(f_{e_*})$ with respect to $\mathcal{P}$. It follows that $d_{\partial H_*(e_*)}(e_*) = r d_{\partial H(f_{e_*})}(f_{e_*}) \geq (1-\varepsilon)nr$.

Now let $e_*, e'_* \in G_*$ of the same colour with $|e_* \cap e'_*| = k-3$. We show that $H_*(e_*) = H_*(e'_*)$. We have $|f_{e_*} \cap f_{e'_*}| = k-3$ and f_{e_*} and $f_{e'_*}$ have the same colour. Thus since G is a blueprint for H, we have $H(f_{e_*}) = H(f_{e'_*})$. Thus, by definition, $H_*(e_*) = H_*(e'_*)$. □

The blueprint of a 2-edge-coloured 4-graph is a 2-edge-coloured graph. We use the following proposition to show that the blow-up of such a blueprint retains the properties of having large minimum degree and of having a spanning red component.

Proposition 3.4.2. *Let $1/n \ll \beta, r$. Let G be a 2-edge-coloured 2-graph with $\delta(G) \geq (1-\beta)n$ and let G_* be the r-blow-up of G with vertex partition $\mathcal{P} = \{V_x : x \in V(G)\}$. Then*

$\delta(G_*) \geq (1-\beta)nr$. Further, if G contains a spanning red component, then G_* contains a spanning red component. The same statement holds with the colours reversed.

Proof. Let $v \in V(G_*)$. There exists $x \in V(G)$ such that $v \in V_x$. We have $N_{G_*}(v) = \bigcup_{y \in N_G(x)} V_y$ and thus $d_{G_*}(v) \geq \delta(G)r \geq (1-\beta)nr$. Now assume that G contains a spanning red component. We show that G_* contains a spanning red component. Let $u, v \in V(G_*)$. There exist $x, y \in V(G)$ such that $u \in V_x$ and $v \in V_y$. Let $z \in V(G) \setminus \{x, y\}$. Since G contains a spanning red component, there exist walks $xx_1 \dots x_k z$ and $yy_1 \dots y_\ell z$ in G^{red}. Choose vertices $u_i \in V_{x_i}$, $v_j \in V_{y_i}$, and $v_z \in V_z$ for $i \in [k]$ and $j \in [\ell]$. Note that $uu_1 \dots u_k z v_\ell \dots v_1 v$ is a walk in G_*^{red}. Hence G_* contains a spanning red component. $\square$

We also reuse Lemma 2.4.3 and Corollary 2.4.6 from Chapter 2 to show that a 2-edge-coloured almost complete 4-graph has a blueprint with large minimum degree that contains a spanning monochromatic component.

3.5 Finding monochromatic tightly connected matchings

Our goal in this section is to prove the following lemma which is the main ingredient in the proof of Theorem 1.2.1.

Lemma 3.5.1. *Let $1/n \ll \varepsilon \ll c \ll \eta$. Let $N = (5/4 + 3\eta)n$. Let H be a 2-edge-coloured $(1-\varepsilon, \varepsilon)$-dense 4-graph on N vertices. Then there exists a monochromatic tightly connected fractional matching in H with weight at least $n/4$ and all weights at least c.*

By using Proposition 3.3.3, proving Lemma 3.5.1 is reduced to showing that, for some $\varepsilon \ll 1/s \ll \eta$, an s-blow-up H_* of H contains a monochromatic tightly connected matching of size at least $v(H_*)/5$. We will prove this as follows. First we find a monochromatic tightly connected matching in H of size δn for some $1/s \ll \delta \ll \eta$. We then iteratively take blow-ups of H that contain monochromatic tightly connected matchings that cover a larger and larger proportion of the vertices. We prove this by showing that as long as our current matching M in a blow-up H_* of H is not yet large enough, we can find a fractional monochromatic matching of weight $|M| + \gamma v(H_*)$ (where $1/s \ll \gamma \ll \delta$). This is the main ingredient in the proof of Lemma 3.5.1 and is formalised in Lemma 3.5.3. We then convert this fractional matching into an integral matching by taking another blow-up (see Proposition 3.5.4).

Since H is only almost complete, its blueprint will also only be almost complete. To overcome some difficulties arising from this, we mostly work with edges of H that work well with respect to its blueprint. We call these edges *good edges* and define them as follows.

Definition 3.5.2 (Good edges, good sets of edges, good fractional matchings, J^+)**.** Let H be a 2-edge-coloured 4-graph and G a blueprint for H. We call an edge $f \in H$ *good for* (H, G) if

(G1) $f \subseteq V(G)$,

(G2) $G[f] \cong K_4$ and

(G3) there exists $z \in f$ such that $xyz \in \partial H(xy)$ for every $xy \in \binom{f \setminus \{z\}}{2}$.

We call a set of edges $F \subseteq H$ *good for* (H, G), if every edge $f \in F$ is good for (H, G). For a subgraph J of H, a fractional matching $\varphi \colon E(J) \to [0, 1]$ is called *good for* (H, G) if $\{e \in E(J) \colon \varphi(e) > 0\}$ is good for (H, G). If H and G are clear from context, then we simply call such edges, sets of edges and fractional matchings *good*. For a subgraph J of H, we let J^+ be the subgraph of J that contains only the edges of J that are good for (H, G). Note that this is different from the notion of G^+ for a blueprint G that we introduced in Section 2.4.2. We do not use this latter notion in this chapter.

Intuitively, by using only good edges we can ignore some of the problems that arise from the fact that H and G are only almost complete. The purpose of (G3) is to allow us to deduce, in some situations, that the edge f is in one of the monochromatic tight components induced by the edges of $G[f]$. For example, if $f = x_1x_2x_3x_4$ is a blue edge in H and $x_1x_2, x_3x_4 \in G^{\text{blue}}$, then (G3) implies that $f \in B(x_1x_2) \cup B(x_3x_4)$.

The following lemma is the main ingredient in the proof of Lemma 3.5.1. It states that if we have a monochromatic matching that is not large enough, then we can find a larger one.

Lemma 3.5.3. *Let $r = \binom{9}{4}!$. Let $1/n \ll \varepsilon \ll \gamma \ll \delta \ll \eta \ll 1$. Let $N = (5/4 + 3\eta)n$. Let H be a 2-edge-coloured $(1-\varepsilon, \varepsilon)$-dense 4-graph on N vertices that does not contain a monochromatic tightly connected matching of size at least $n/4$. Let G be an ε-blueprint for H with $\delta(G) \geq (1-\varepsilon)N$. Suppose H contains a red tight component R satisfying $H(e) = R$ for every $e \in G^{\mathrm{red}}$. Let M be a good matching in H with $3\delta n \leq |M| < n/4$ such that one of the following holds.*

(H1) *$M \subseteq R$ or*

(H2) *M is contained in a blue tight component B of H.*

Then H contains a good $1/r$-fractional matching in R or in a blue tight component of H of weight at least $|M| + \gamma n$. The same statement holds with colours reversed.

3.5.1 Proof of Lemma 3.5.1 assuming Lemma 3.5.3

Before proving Lemma 3.5.3, we show how to prove Lemma 3.5.1 using Lemma 3.5.3. To do this we need a few other small results.

The following proposition shows that we can turn a good $1/r$-fractional matching into a good integral matching by taking an r-blow-up.

Proposition 3.5.4. *Let $n, r \geq 2$ be integers and $\mu \geq 0$. Let H be 2-edge-coloured 4-graph on n vertices. Let H_* be an r-blow-up of H with vertex partition $\mathcal{P} = \{V_x : x \in V(H)\}$. Let G be a blueprint for H and let G_* be the corresponding blueprint for H_* as defined in Proposition 3.4.1. Let F be a monochromatic tight component of H and let $F_* = \bigcup_{x_1 \dots x_4 \in F} K_{V_{x_1}, \dots, V_{x_4}}$ be the corresponding monochromatic tight component of H_*. Let φ be a $1/r$-fractional matching in F with weight μ that is good for (H, G). Then there exists an (integral) matching in F_* of size μr that is good for (H_*, G_*).*

Proof. For each vertex $x \in V(H)$ and each edge $e \in F$ containing x, choose disjoint sets $U_{x,e} \subseteq V_x$ such that $|U_{x,e}| = r\varphi(e)$. This is possible since φ is a $1/r$-fractional matching and $|V_x| = r$ for each $x \in V(H)$. For each edge $e = x_1 \dots x_4 \in F$, let M_e be a perfect matching in $K_{U_{x_1,e},\dots,U_{x_4,e}}$. Clearly, the M_e are disjoint. Let $M = \bigcup_{e \in F} M_e$. Note that $|M| = \sum_{e \in F} r\varphi(e) = \mu r$. It is easy to see that since φ is a fractional matching in F that is good for (H, G), M is a matching in F_* that is good for (H_*, G_*). $\square$

The following proposition shows that (in a strong sense) most edges are good in our usual setting of having a $(1-\varepsilon, \varepsilon)$-dense 2-edge-coloured 4-graph H and a blueprint G for it with large minimum degree. Recall that H^+ is the subgraph of H that contains only the edges of H that are good for (H, G).

Proposition 3.5.5. *Let $1/N \ll \varepsilon \ll \gamma$. Let H be a $(1-\varepsilon, \varepsilon)$-dense 2-edge-coloured 4-graph on N vertices, let G be an ε-blueprint for H with $\delta(G) \geq (1-\varepsilon)N$ and $W \subseteq V(G)$ a set of size at least γN. Then $\delta_1(H^+[W]) \geq \binom{|W|}{3} - 2\varepsilon N^3$. Moreover, $H^+[W]$ contains a matching of size at least $\frac{|W|}{4} - \gamma N$.*

Proof. Fix $v \in V(H)$. Choose vertices

$$
\begin{aligned}
&z_1 \in N_G(v),\\
&z_2 \in N_G(v) \cap N_G(z_1) \cap N_{\partial H(vz_1)}(vz_1) \text{ and}\\
&z_3 \in N_G(v) \cap N_G(z_1) \cap N_G(z_2) \cap \bigcap_{xy \in \binom{vz_1z_2}{2}} N_{\partial H(xy)}(xy) \cap N_H(vz_1z_2).
\end{aligned}
$$

Note that $vz_1z_2z_3$ is good. Since G is an ε-blueprint for H with $\delta(G) \geq (1-\varepsilon)N$ and H is $(1-\varepsilon, \varepsilon)$-dense, the number of choices for z_1, z_2 and z_3 are at least $(1-\varepsilon)N$, $(1-3\varepsilon)N$ and $(1-7\varepsilon)N$, respectively. Hence

$$\delta_1(H^+) \geq \frac{(1-11\varepsilon)N^3}{3!} \geq \binom{N}{3} - 2\varepsilon N^3.$$

It follows that $\delta_1(H^+[W]) \geq \binom{|W|}{3} - 2\varepsilon N^3$. By a greedy argument, $H^+[W]$ contains a matching of size at least $\frac{|W|}{4} - \gamma N$. $\square$

The following proposition shows that in our usual setting the following holds. Given two sets of vertices T_1 and T_2 (with $|T_i| \in \{2, 3, 4\}$ and satisfying some simple conditions) there exist vertices z_1, z_2, z_3 such that all the edges in $H[T_i \cup z_1z_2z_3]$ are good. It is also shown that we can choose these vertices z_1, z_2 and z_3 in any not too small set of vertices. We use this to find tight connections of good edges between small sets of vertices.

Proposition 3.5.6. *Let $1/N \ll \varepsilon \ll \gamma$. Let H be a $(1-\varepsilon, \varepsilon)$-dense 2-edge-coloured 4-graph on N vertices, let G be an ε-blueprint for H with $\delta(G) \geq (1-\varepsilon)N$. Let $W \subseteq V(G)$ be a set of size at least γN. Let $T_1, T_2 \subseteq V(G)$ be sets such that, for $i \in [2]$,*

(a) $2 \leq |T_i| \leq 4$,

(b) $T_i \in H^+$ *if* $|T_i| = 4$,

(c) $G[T_i] \cong K_{|T_i|}$,

(d) $N_H(S) \neq \varnothing$ *for all* $S \in \binom{T_i}{3}$.

Then there exist vertices $z_1, z_2, z_3 \in W$ such that, for $i \in [2]$,

(i) $H[T_i \cup z_1z_2z_3] \cong K^{(4)}_{|T_i|+3}$,

(ii) $G[T_i \cup z_1z_2z_3] \cong K_{|T_i|+3}$,

(iii) $xyz_1 \in \partial H(xy)$ *for all* $xy \in \binom{T_i}{2}$,

(iv) $xyz_2 \in \partial H(xy)$ *for all* $xy \in \binom{T_i \cup z_1}{2}$,

(v) $xyz_3 \in \partial H(xy)$ *for all* $xy \in \binom{T_i \cup z_1z_2}{2}$.

In particular, $H^+[T_i \cup z_1z_2z_3] \cong K^{(4)}_{|T_i|+3}$ *for* $i \in [2]$.

Proof. Choose vertices

$$z_1 \in W \cap \bigcap_{S \in \binom{T_1}{3} \cup \binom{T_2}{3}} N_H(S) \cap \bigcap_{x \in T_1 \cup T_2} N_G(x) \cap \bigcap_{xy \in \binom{T_1}{2} \cup \binom{T_2}{2}} N_{\partial H(xy)}(xy),$$

$$z_2 \in W \cap \bigcap_{S \in \binom{T_1 \cup z_1}{3} \cup \binom{T_2 \cup z_1}{3}} N_H(S) \cap \bigcap_{x \in T_1 \cup T_2 \cup z_1} N_G(x) \cap \bigcap_{xy \in \binom{T_1 \cup z_1}{2} \cup \binom{T_2 \cup z_2}{2}} N_{\partial H(xy)}(xy) \text{ and}$$

$$z_3 \in W \cap \bigcap_{S \in \binom{T_1 \cup z_1 z_2}{3} \cup \binom{T_2 \cup z_1 z_2}{3}} N_H(S) \cap \bigcap_{x \in T_1 \cup T_2 \cup z_1 z_2} N_G(x)$$
$$\cap \bigcap_{xy \in \binom{T_1 \cup z_1 z_2}{2} \cup \binom{T_2 \cup z_1 z_3}{2}} N_{\partial H(xy)}(xy),$$

noting that these vertices exist since H is $(1-\varepsilon, \varepsilon)$-dense,[1] G is an ε-blueprint for H with $\delta(G) \geq (1-\varepsilon)N$ and $|W| \geq \gamma N$. □

The following corollary states that in our usual setting the good edges are tightly connected in any not too small induced subgraph of H.

Corollary 3.5.7. *Let $1/N \ll \varepsilon \ll \gamma$. Let H be a $(1-\varepsilon, \varepsilon)$-dense 2-edge-coloured 4-graph on N vertices. Let G be an ε-blueprint for H with $\delta(G) \geq (1-\varepsilon)N$. Let $W \subseteq V(G)$ be a set of size at least γN. Then $H^+[W]$ is tightly connected.*

Proof. Let $f_1, f_2 \in H^+[W]$. By Proposition 3.5.6, there exist vertices $z_1, z_2, z_3 \in W$ such that $H^+[f_1 \cup z_1 z_2 z_3] \cong H^+[f_2 \cup z_1 z_2 z_3] \cong K_7^{(4)}$. It follows that f_1 and f_2 are in the same tight component of H^+. □

The next lemma allows us to find blue tight components with useful properties. Recall that given a 2-edge-coloured 4-graph H, a blueprint G for H and a blue tight component B of H, we denote by B^2 the set of edges $e \in G^{\text{blue}}$ such that $B(e) = B$.

Lemma 3.5.8. *Let $1/n \ll \varepsilon \ll \gamma \ll \eta < 1$. Let $N = (5/4 + 3\eta)n$. Let H be a 2-edge-coloured $(1-\varepsilon, \varepsilon)$-dense 4-graph on N vertices and let G be an ε-blueprint for H with $\delta(G) \geq (1-\varepsilon)N$. Suppose H contains a red tight component R satisfying $H(e) = R$ for every $e \in G^{\text{red}}$. Then for each $W \subseteq V(G)$ with $|W| \geq \gamma n$ such that $R^+[W] =$*

[1] We use that for $x, y, z \in V(H)$ with $xyz \in \partial H$, we have $d_H(xyz) > 0$ and thus $d_H(xyz) \geq (1-\varepsilon)N$.

$\varnothing$, there exists a blue tight component B_W of H such that the following holds. Let $\mathcal{T}_W = \{T \in \binom{W}{3} \colon G[T] \cong K_3, G^{\text{red}}[T] \neq \varnothing, N_H(T) \neq \varnothing\}$. For $T \in \mathcal{T}_W$, let $\Gamma_W(T) = W \cap \bigcap_{x \in T} N_G(x) \cap \bigcap_{xy \in \binom{T}{2}} N_{\partial H(xy)}(xy) \cap N_H(T)$.

(B1) *For any $T \in \mathcal{T}_W$, we have $\Gamma_W(T) \neq \varnothing$ and $T \cup w \in B_W^+$ for all $w \in \Gamma_W(T)$. In particular, $\mathcal{T}_W \subseteq \partial B_W$.*

(B2) *For each $e \in G^{\text{blue}}[W]$, we have $B(e) = B_W$, that is, $G^{\text{blue}}[W] \subseteq B_W^2$.*

Moreover, if $W_1, W_2 \subseteq V(G)$ satisfy $|W_1 \cap W_2| \geq \gamma n$ and $R^+[W_1] = R^+[W_2] = \varnothing$, then $B_{W_1} = B_{W_2}$.

Proof. For $T \in \mathcal{T}_W$, since G is an ε-blueprint for H with $\delta(G) \geq (1-\varepsilon)N$ and H is $(1-\varepsilon, \varepsilon)$-dense, we have

$$|\Gamma_W(T)| \geq |W| - 7\varepsilon N \geq |W| - 14\varepsilon n. \tag{3.5.1}$$

In particular, for any $T \in \mathcal{T}_W$, $\Gamma_W(T) \neq \varnothing$ (since $|W| \geq \gamma n$ and $\varepsilon \ll \gamma$).

Note that if $T \in \mathcal{T}_W$ and $w \in \Gamma_W(T)$, then $T \cup w \in H^{\text{blue}}$ (or else $T \cup w \in R^+$ as $xyw \in \partial R$ for some $xy \in G^{\text{red}}[T]$ contradicting $R^+[W] = \varnothing$). For $T \in \mathcal{T}_W$, let B_W^T be the blue tight component of H containing all the edges $T \cup w$ where $w \in \Gamma_W(T)$. Note that, in particular, $T \in \partial B_W^T$ for every $T \in \mathcal{T}_W$. Moreover,

$$B(e) = B_W^T \text{ for all } T \in \mathcal{T}_W \text{ and } e \in G^{\text{blue}}[T]. \tag{3.5.2}$$

Claim 3.5.9. *There exists a blue tight component B_W such that $B_W^T = B(e) = B_W$ for any $e \in G^{\text{blue}}[W]$ and any $T \in \mathcal{T}_W$.*

Proof of Claim. First assume that $\mathcal{T}_W = \varnothing$. This implies that $G^{\text{red}}[W] = \varnothing$. So $G^{\text{blue}}[W]$ is connected and thus, since G is a blueprint, $B(e_1) = B(e_2)$ for any $e_1, e_2 \in G^{\text{blue}}[W]$. So we may set $B_W = B(e)$ for all $e \in G^{\text{blue}}[W]$. It follows that (B1) and (B2) both hold.

Now assume that $\mathcal{T}_W \neq \varnothing$. First we show that

$$B_W^{T_1} = B_W^{T_2} \text{ for any } T_1, T_2 \in \mathcal{T}_W \text{ with } |T_1 \cap T_2| \geq 2. \tag{3.5.3}$$

Let $T_1, T_2 \in \mathcal{T}_W$ with $|T_1 \cap T_2| = 2$. By (3.5.1), there exists $w \in \Gamma_W(T_1) \cap \Gamma_W(T_2)$. We have $T_1 \cup w \in B_W^{T_1}$ and $T_2 \cup w \in B_W^{T_2}$. Since $|(T_1 \cup w) \cap (T_2 \cup w)| = 3$ and $B_W^{T_1}$ and $B_W^{T_2}$ are blue tight components, we have $B_W^{T_1} = B_W^{T_2}$.

Now we show that (3.5.3) actually holds for any $T_1, T_2 \in \mathcal{T}_W$, that is

$$B_W^{T_1} = B_W^{T_2} \text{ for any } T_1, T_2 \in \mathcal{T}_W. \tag{3.5.4}$$

Let $T_1, T_2 \in \mathcal{T}_W$. Say $T_1 = x_1x_2x_3$ and $T_2 = y_1y_2y_3$, where $x_1x_2 \in G^{\text{red}}$ and $y_1y_2 \in G^{\text{red}}$. By Proposition 3.5.6, there exist vertices $z_1, z_2 \in W$ such that

$$H[T_i \cup z_1z_2] \cong K_5^{(4)} \text{ and } G[T_i \cup z_1z_2] \cong K_5 \text{ for } i \in [2].$$

Note that $x_1x_2z_1, y_1y_2z_1 \in \mathcal{T}_W$. If x_1z_1 and y_1z_1 are both in G^{red}, then $x_1z_1z_2, y_1z_1z_2 \in \mathcal{T}_W$ and thus by (3.5.3), we have $B_W^{T_1} = B_W^{x_1x_2z_1} = B_W^{x_1z_1z_2} = B_W^{y_1z_1z_2} = B_W^{y_1y_2z_1} = B_W^{T_2}$. If x_1z_1 and y_1z_1 are both in G^{blue}, then by (3.5.2) and the fact that G is a blueprint, we have $B_W^{T_1} = B_W^{x_1x_2z_1} = B(x_1z_1) = B(y_1z_1) = B_W^{y_1y_2z_1} = B_W^{T_2}$. Now assume that exactly one of x_1z_1 and y_1z_1 is in G^{red}, say $x_1z_1 \in G^{\text{red}}$ and $y_1z_1 \in G^{\text{blue}}$. Note that $x_1z_1z_2 \in \mathcal{T}_W$. If $z_1z_2 \in G^{\text{red}}$, then $x_1z_1z_2, y_1z_1z_2 \in \mathcal{T}_W$ and so by (3.5.2), we have $B_W^{T_1} = B_W^{x_1x_2z_1} = B_W^{x_1z_1z_2} = B_W^{y_1z_1z_2} = B_W^{y_1y_2z_1} = B_W^{T_2}$. If $z_1z_2 \in G^{\text{blue}}$, then by (3.5.2), (3.5.3) and the fact that G is a blueprint, we have $B_W^{T_1} = B_W^{x_1x_2z_1} = B_W^{x_1z_1z_2} = B(z_1z_2) = B(y_1z_1) = B_W^{y_1y_2z_1} = B_W^{T_2}$.

Now we show that

$$B_W^{T} = B(e) \text{ for any } T \in \mathcal{T}_W \text{ and any } e \in G^{\text{blue}}[W]. \tag{3.5.5}$$

Let $T \in \mathcal{T}_W$ with $e_1 = x_1x_2 \in G^{\text{red}}[T]$ and $e_2 = y_1y_2 \in G^{\text{blue}}[W]$. By Proposition 3.5.6, there exist vertices $z_1, z_2 \in W$ such that

$$e_i \cup z_1z_2 \in H^+ \text{ for } i \in [2].$$

If $y_1z_1 \in G^{\text{red}}$, then $y_1y_2z_1 \in \mathcal{T}_W$ and by (3.5.2) and (3.5.4), we have $B_W^T = B_W^{y_1y_2z_1} = B(e_2)$. Now assume $y_1z_1 \in G^{\text{blue}}$. If $x_1z_1 \in G^{\text{blue}}$, then by (3.5.2), (3.5.4) and the fact that G is a blueprint, we have $B_W^T = B_W^{x_1x_2z_1} = B(x_1z_1) = B(y_1z_1) = B(e_2)$. Next assume $x_1z_1 \in G^{\text{red}}$. If $z_1z_2 \in G^{\text{red}}$, then by (3.5.2), (3.5.4) and the fact that G is a blueprint, we have $B_W^T = B_W^{y_1z_1z_2} = B(y_1z_1) = B(e_2)$. If $z_1z_2 \in G^{\text{blue}}$, then by (3.5.2), (3.5.4) and the fact that G is a blueprint, we have $B_W^T = B_W^{x_1z_1z_2} = B(z_1z_2) = B(y_1z_1) = B(e_2)$.

Since $\mathcal{T}_W \neq \varnothing$, (3.5.5) implies that

$$B(e_1) = B(e_2) \text{ for any } e_1, e_2 \in G^{\text{blue}}[W]. \tag{3.5.6}$$

It follows from (3.5.4), (3.5.5) and (3.5.6) that we may set $B_W = B_W^T = B(e)$ for all $T \in \mathcal{T}_W$ and all $e \in G^{\text{blue}}[W]$. It follows that (B1) and (B2) hold. ■

Now we show the final statement of the lemma. Let $W_1, W_2 \subseteq V(G)$ with $|W_1 \cap W_2| \geq \gamma n$ and $R^+[W_1] = R^+[W_2] = \varnothing$. Greedily choose vertices $x_1, x_2, x_3 \in W_1 \cap W_2$ such that $G[x_1x_2x_3] \cong K_3$ and $x_1x_2x_3 \in \partial H(x_1x_2)$ (this is possible since G is an ε-blueprint with $\Delta(\overline{G}) \leq 2\varepsilon n$ and $1/n \ll \varepsilon \ll \gamma$). Note that since $x_1x_2x_3 \in \partial H(x_1x_2)$, we have $N_H(x_1x_2x_3) \neq \varnothing$. If $x_1x_2 \in G^{\text{red}}$, then $x_1x_2x_3 \in \mathcal{T}_{W_2} \cap \mathcal{T}_{W_2} \subseteq \partial B_{W_1} \cap \partial B_{W_2}$ and thus $B_{W_1} = B_{W_2}$. If $x_1x_2 \in G^{\text{blue}}$, then $B_{W_1} = B(x_1x_2) = B_{W_2}$. □

We are now ready to prove Lemma 3.5.1 assuming Lemma 3.5.3.

Proof of Lemma 3.5.1. Suppose for a contradiction that there does not exist a tightly connected fractional matching in H with weight at least $n/4$ and all weights at least c. Let $r = \binom{9}{4}!$. Choose new constants ε_0, γ and δ such that $1/n \ll \varepsilon \ll \varepsilon_0 \ll c \ll \gamma \ll \delta \ll \eta$. By Lemma 2.4.3 and Corollary 2.4.6, there exists an ε_0-blueprint for H with $\delta(G) \geq (1 - \varepsilon_0)N$ that contains a spanning monochromatic component. Without loss of generality assume that G contains a spanning red component and let R be the unique red tight component of H such that $H(e) = R$ for every edge $e \in G^{\mathrm{red}}$.

Claim 3.5.10. *There exists a good matching M of size at least $3\delta n$ in H that is contained in R or in a blue tight component of H.*

Proof of Claim. Let M be a maximum good matching in R and let $W = V(G) \setminus V(M)$. It follows that $R^+[W] = \varnothing$. Moreover, we may assume that $|M| < 3\delta n$ (or else we are done). Thus $|W| \geq |V(G)| - 12\delta n \geq N - 13\delta n$. Let $B = B_W$ be the blue tight component that exists by Lemma 3.5.8.

Case A: $G^{\mathrm{red}}[W]$ contains a matching of size at least $3\delta n$. Let $t = 3\delta n$ and let $\{u_i v_i \colon i \in [t]\}$ be a matching in $G^{\mathrm{red}}[W]$. Let $\mathcal{T}_W$ and Γ_W be defined as in Lemma 3.5.8. Since G is an ε_0-blueprint for H with $\delta(G) \geq (1-\varepsilon_0)N$, there exist distinct vertices $w_i \in W$, one for each $i \in [t]$, such that $u_i v_i w_i \in \mathcal{T}_W$. Since H is $(1-\varepsilon, \varepsilon)$-dense, there exist disjoint vertices $w_i' \in \Gamma_W(u_i v_i w_i)$, one for each $i \in [t]$. By Lemma 3.5.8 (B1), $u_i v_i w_i w_i' \in B^+$ for each $i \in [t]$. It follows that $\{u_i v_i w_i w_i' \colon i \in [t]\}$ is a good matching of size $3\delta n$ in B, as required.

Case B: $G^{\mathrm{red}}[W]$ does not contain a matching of size at least $3\delta n$. It follows that there exists a set $W' \subseteq W$ of size at least $|W| - 6\delta n \geq N - 19\delta n$ such that $G[W'] \subseteq G^{\mathrm{blue}}$. By Lemma 3.5.8 (B2), $G[W'] \subseteq B^2$. Let M' be a maximum matching in $B^+[W']$. We may assume that $|M'| < 3\delta n$ (or else we are done). Let $W'' = W' \setminus V(M')$ and so $|W''| \geq N - 31\delta n$. Note that by the maximality of M' and $G[W'] \subseteq B^2$, we have that $H^+[W''] \subseteq H^{\mathrm{red}}$. By Corollary 3.5.7, there exists a red tight component R_* of H such that $H^+[W''] = R_*^+[W'']$. Thus by Proposition 3.5.5, $R_*^+[W'']$ contains a matching of size at least

$$\frac{|W''|}{4} - \delta N \geq \frac{N - 31\delta n}{4} - 2\delta n \geq \frac{n}{4},$$

a contradiction. ■

Let $0 \leq L \leq 1/\gamma$ be the largest integer such that the following holds. Let H_* be an

r^L-blow-up of H. Let $\mathcal{P} = \{V_x : x \in V(H)\}$ be the partition of $V(H_*)$. Let G_* be the ε-blueprint for H_* as defined in Proposition 3.4.1. Let R_* be the red tight component in H_* that is the r^L-blow-up of R. Then H_* contains a matching M_* in R_* or in a blue tight component of H_* with $|M_*| \geq r^L(3\delta n + L\gamma n)$ that is good for (H_*, G_*).

By Claim 3.5.10, we have that L is well-defined. Moreover, if $L \geq \frac{1}{4\gamma}$, then by Proposition 3.3.3, we are done. Hence we may assume that $L < \frac{1}{4\gamma}$. Let $n_* = nr^L$ and $N_* = Nr^L = (5/4 + 3\eta)n_*$. Since H is $(1-\varepsilon, \varepsilon)$-dense, Proposition 3.3.2 implies that H_* is $(1-2\varepsilon, 2\varepsilon)$-dense and hence also $(1-\varepsilon_0, \varepsilon_0)$-dense. By Proposition 3.4.1 and Proposition 3.4.2, G_* is an ε_0-blueprint for H_* with $\delta(G_*) \geq (1-\varepsilon_0)nr^L = (1-\varepsilon_0)n_*$ such that $H_*(e) = R_*$ for all $e \in G_*^{\mathrm{red}}$. We may further assume that H_* does not contain a monochromatic tightly connected matching of size at least $n_*/4$ (or else we are done by Proposition 3.3.3). We apply Lemma 3.5.3 with $r, n_*, \varepsilon_0, \gamma, \delta, \eta, N_*, H_*, G_*, R_*, M_*$ playing the roles of $r, n, \varepsilon, \gamma, \delta, \eta, N, H, G, R, M$. We deduce that H_* contains a $1/r$-fractional matching in R_* or in a blue tight component of H_* that is good for (H_*, G_*) of weight at least $|M_*| + \gamma n_*$. Let H_{**} be an r-blow-up of H_* and note that H_{**} is an r^{L+1}-blow-up of H. Let R_{**} be the red tight component of H_{**} that is the r-blow-up of R_* and thus is the r^{L+1}-blow-up of R. By Proposition 3.4.1, $H_{**}(e) = R_{**}$ for all $e \in G_{**}^{\mathrm{red}}$. By Proposition 3.5.4, H_{**} contains a matching in R_{**} or in a blue tight component of H_{**} that is good for (H_{**}, G_{**}) of size at least

$$r(|M_*| + \gamma n_*) \geq r(r^L(3\delta n + L\gamma n) + \gamma r^L n) = r^{L+1}(3\delta n + (L+1)\gamma n).$$

This is a contradiction to the maximality of L. □

3.5.2 Sketch of the proof of Lemma 3.5.3 and suitable pairs

Before proceeding with the proof of Lemma 3.5.3, we give a sketch of the proof. Recall that our aim is given a monochromatic tightly connected matching M, to find a larger monochromatic tightly connected fractional matching. We split the proof into two cases depending on whether (H1) or (H2) holds. We only sketch the proof of the case where (H1) holds, that is $M \subseteq R$ (the other case is similar). We may assume that M is a matching of maximum size in R. Moreover, for simplicity, we assume that

$$\begin{aligned}&H \text{ and } G \text{ are complete, } V(G) = V(H) \text{ and for every } e \in G \text{ and } v \in V(H) \setminus e,\\ &\text{we have } e \cup v \in \partial H(e),\end{aligned} \tag{A}$$

the last of which is an idealised version of (BP1).

Let $W = V(H) \setminus V(M)$ be the vertices of H not covered by M. By Lemma 3.5.8, there exists a blue tight component B in H such that $\binom{W}{3} \subseteq \partial B$ and every $e \in G^{\text{blue}}[W]$ induces B. We will find our desired fractional matching in R or B.

Our first step is to find a subset $U' \subseteq W$ and two matchings $M' = \{f'(u) : u \in U'\} \subseteq M \subseteq R$ and $M_1^* = \{f_u^* : u \in U'\} \subseteq B$ such that $|U'| \approx |M|$, and

$$|f'(u) \cap f_u^*| = 3 \text{ and } f_u^* \setminus f'(u) = u \text{ for all } u \in U'.$$

If no such U' exists, then we can find a small matching $M'' \subseteq M$ and for each $f \in M''$ a disjoint 4-set $W_f \subseteq W$ such that, for each $f \in M''$ there exists a fractional matching φ_f in $R[f \cup W_f]$ of weight at least $\frac{r+1}{r}$ (where r is some absolute constant). By starting with M and replacing each edge $f \in M''$ with φ_f we get a larger fractional matching. (See Claim 3.5.16 for the details.)

In our second step we then extend the matching M_1^* to a larger fractional matching in B completing the proof.

We will use the following fact which allows us to find a fractional matching in $R[f \cup W_f]$.

Fact 3.5.11. *Let $k, s \geq 2$. Let H be a k-graph and let $F \subseteq E(H)$ be a nonempty set such that $\bigcap F = \varnothing$ and $|F| = s$. Then there exists a $\frac{1}{s-1}$-fractional matching in H with weight $\frac{s}{s-1}$.*

Proof. Let $\varphi \colon E(H) \to [0,1]$ be defined by $\varphi(e) = \frac{1}{s-1}$ for each $e \in F$ and $\varphi(e) = 0$ otherwise. Since $\bigcap F = \varnothing$ each vertex of H is contained in at most $s-1$ edges of F. It follows that for every $v \in V(H)$, we have $\sum_{e \in E(H)\colon v \in e} \varphi(e) \leq 1$. Thus φ is a $\frac{1}{s-1}$-fractional matching in H. $\square$

In the actual proof of Lemma 3.5.3, H and G will only be almost complete and so do not satisfy (A). In particular, we may not have $e \cup v \in \partial H(e)$ for some $e \in G$ and $v \in V(H) \setminus e$. To overcome the difficulties that arise from this, we introduce the following notion of *suitable pairs*. For a suitable pair (f, W), it is useful to think of f as an edge of some good matching and W as a subset of the vertices not covered by that matching. Then (SP1) to (SP6) are the properties that we would have in the idealised case where (A) holds that are necessary for our proof.

Definition 3.5.12 (Suitable pairs)**.** Let H be a 2-edge-coloured 4-graph and G a blueprint for H. Let $f \in H$ be a good edge and $W \subseteq V(G) \setminus f$. We call (f, W) a *suitable pair for* (H, G) if the following properties hold, where $s = |W|$.

(SP1) $H[f \cup W] \cong K_{s+4}^{(4)}$.

(SP2) $G[f \cup W] \cong K_{s+4}$.

(SP3) If $xy \in \binom{f}{2}$ and $z \in W$, then $xyz \in \partial H(xy)$.

(SP4) If $xy \in \binom{W}{2}$ and $z \in f$, then $xyz \in \partial H(xy)$.

(SP5) If $x \in f$ and $yz \in \binom{W}{2}$, then $xyz \in \partial H(xy)$.

(SP6) If $xyz \in \binom{W}{3}$, then $xyz \in \partial H(xy)$.

If H and G are clear from context, we simply call (f, W) a suitable pair. Note that if (f, W) is a suitable pair and $W' \subseteq W$, then (f, W') is a suitable pair. Moreover, if (f, W) is a suitable pair, then any edge in $H[f \cup W]$ is good. Also if $e \in R^2[W]$, then $f' \in R^+$ for any edge $f' \in H[f \cup W]$ with $e \subseteq f'$.

We use the following lemma to find suitable pairs. The main idea is that if we choose uniformly at random an edge f from a good matching M and a subset W_f of constant size from $V(G) \setminus V(M)$, then (f, W_f) is likely to be a suitable pair.

Lemma 3.5.13. *Let $1/N \ll \varepsilon \ll \gamma \ll \delta \ll 1/s \leq 1$. Let H be a $(1-\varepsilon, \varepsilon)$-dense 2-edge-coloured 4-graph on N vertices and let G be an ε-blueprint for H with $\delta(G) \geq (1-\varepsilon)N$. Let M be a good matching in H of size at least δN and let $W \subseteq V(G) \setminus V(M)$ with $|W| \geq \delta N$. Then there exist a matching $M' \subseteq M$ with $|M'| \geq \gamma N$ and disjoint sets $W_f \in \binom{W}{s}$ for each $f \in M'$ such that (f, W_f) is a suitable pair for each $f \in M'$.*

Proof. Let $m = 2\gamma N$. We independently choose $\{f_i : i \in [m]\}$ uniformly at random among all subsets of M of size m and $\{W_{f_i} : i \in [m]\}$ uniformly at random among all sets of m disjoint sets in $\binom{W}{s}$. Note that for each $i \in [m]$, f_i is distributed uniformly in M and W_{f_i} is independent from f_i and distributed uniformly in $\binom{W}{s}$. For each $i \in [m]$, let A_i be the event that (f_i, W_{f_i}) is a suitable pair. Note that it suffices to show that $\mathbb{P}[A_i] \geq 1/2$ for each $i \in [m]$ since then for $M' = \{f_i : i \in [m] \text{ such that } A_i \text{ holds}\}$, we have $\mathbb{E}[|M'|] \geq m/2 \geq \gamma N$.

Fix $i \in [m]$. For $j \in \{1, 2\}$, let $A_{i,j}$ be the event that (f_i, W_{f_i}) satisfies (SPj). For $j \in \{3, \dots, 6\}$, we modify the statements slightly for the following probability calculation. For $j \in \{3, \dots, 6\}$, let $A_{i,j}$ be the event that (f_i, W_{f_i}) satisfies (SPj) but only for those pairs xy such that $xy \in G$. Note that $A_i \subseteq \bigcap_{j \in [6]} A_{i,j}$.

To prove the lemma, it suffices to show that $\mathbb{P}[A_{i,j}] \geq 11/12$ for each $j \in [6]$.

To bound $\mathbb{P}[A_{i,1}]$, we fix f_i and count the number of sets $W_i \in \binom{W}{s}$ such that (f_i, W_i) satisfies (SP1). Note that if we iteratively choose

$$w_j \in W \cap \bigcap_{S \in \binom{f_i \cup w_1 \dots w_{j-1}}{3}} N_H(S)$$

for each $j \in [s]$, then $(f_i, \{w_1, \dots, w_s\})$ satisfies (SP1). Since $d_H(S) \geq (1-\varepsilon)N$ for each $S \in \binom{V(H)}{3}$ with $N_H(S) \neq \varnothing$, the number of choices for w_j is at least $|W| - \binom{s+3}{3}\varepsilon N$. Hence the number of sets $W_i \in \binom{W}{s}$ such that (f_i, W_i) satisfies (SP1) is at least

$$\frac{(|W| - \binom{s+3}{3}\varepsilon N)^s}{s!}.$$

It follows that

$$\mathbb{P}[A_{i,1}] \geq \frac{(|W| - \binom{s+3}{3}\varepsilon N)^s}{\binom{|W|}{s}s!} \geq \left(1 - \binom{s+3}{3}\sqrt{\varepsilon}\right)^s \geq 1 - s\binom{s+3}{3}\sqrt{\varepsilon} \geq \frac{11}{12},$$

where in the second inequality we used that $|W| \geq \delta N$. By a similar argument, since $W \subseteq V(G)$ and $\delta(G) \geq (1-\varepsilon)N$, we have

$$\mathbb{P}[A_{i,2}] \geq \frac{(|W| - (s+3)\varepsilon N)^s}{\binom{|W|}{s}s!} \geq 1 - s(s+3)\sqrt{\varepsilon} \geq \frac{11}{12}.$$

To bound $\mathbb{P}[A_{i,3}]$, we fix f_i and recall that since $W_{f_i} = \{w_1, \ldots, w_s\}$ is chosen uniformly at random in $\binom{W}{s}$, each w_j is distributed uniformly in W. Since G is an ε-blueprint, (BP1) implies that $d_{\partial H(xy)}(xy) \geq (1-\varepsilon)N$ for every $xy \in G$. Hence for $xy \in f_i$ and $j \in [s]$, we have that $\mathbb{P}[w_i \notin N_{\partial H(xy)}(xy)] \leq \varepsilon N / |W| \leq \sqrt{\varepsilon}$. A union bound implies that

$$\mathbb{P}[A_{i,3}] \geq 1 - 6s\sqrt{\varepsilon} \geq \frac{11}{12}.$$

To bound $\mathbb{P}[A_{i,4}]$ we fix W_{f_i} and recall that f_i is distributed uniformly in M. Since G is an ε-blueprint, for each $xy \in G[W_{f_i}]$, we have $d_{\partial H(xy)}(xy) \geq (1-\varepsilon)N$. Hence there are at most $4\binom{s}{2}\varepsilon N$ elements of M for which (f_i, W_{f_i}) does not satisfy (SP4). It follows that

$$\mathbb{P}[A_{i,4}] \geq \frac{|M| - 4\binom{s}{2}\varepsilon N}{|M|} \geq 1 - 4\binom{s}{2}\sqrt{\varepsilon} \geq \frac{11}{12},$$

as $|M| \geq \delta N$.

To bound $\mathbb{P}[A_{i,5}]$, we fix f_i and let $W_{f_i} = \{w_1, \ldots, w_s\}$. For each distinct $j, j' \in [s]$, $w_j w_{j'}$ is distributed uniformly in $\binom{W}{2}$. Since G is an ε-blueprint, for each $xy \in G$, we have $d_{\partial H(xy)}(xy) \geq (1-\varepsilon)N$. Hence, for $x \in f_i$ and distinct $j, j' \in [s]$, we have

$$\mathbb{P}[xw_j \notin G \text{ or } w_{j'} \notin N_{\partial H(xw_j)}(xw_j)] \leq \frac{|W|\varepsilon N}{2\binom{|W|}{2}}.$$

A union bound implies that

$$\mathbb{P}[A_{i,5}] \geq 1 - 4s^2 \frac{|W| \, \varepsilon N}{2\binom{|W|}{2}} \geq 1 - 4s^2\sqrt{\varepsilon} \geq \frac{11}{12}.$$

Similarly,

$$\mathbb{P}[A_{i,6}] \geq 1 - s^3 \frac{|W| \, (|W|-1)\varepsilon N}{3!\binom{|W|}{3}} \geq 1 - s^3\sqrt{\varepsilon} \geq \frac{11}{12}.$$

□

The following proposition analyses a specific pattern that we will encounter a few times in our proof.

Proposition 3.5.14. *Let H be a 2-edge-coloured 4-graph and let G be a blueprint for H. Let R be a red tight component of H. Let $f \in R^+$ and let (f, W) be a suitable pair such that $|W| = 3$ and there is an edge $e \in R^2[W]$. Let $F = \{f' \in H^{\mathrm{red}}[f \cup W] \colon e \subseteq f'\}$. Suppose $\bigcap R[f \cup W] \neq \varnothing$. Then there exists $x \in f \cap \bigcap F$. In particular, for any edge $f' \in H[f \cup W]$ with $e \subseteq f'$ and $x \notin f'$, we have $f' \in H^{\mathrm{blue}}$. The same statement holds with colours reversed.*

Proof. Suppose for a contradiction that $f \cap \bigcap F = \varnothing$. Let $f' \in F$ and let $z \in f' \cap f$. By (SP4), $e \cup z \in \partial H(e) = \partial R$. Since $e \cup z \subseteq f'$ and $f' \in H^{\mathrm{red}}$, we have $f' \in R$. Thus $F \subseteq R$. It follows that $\bigcap R[f \cup W] = \varnothing$, a contradiction. □

3.5.3 Proof of Lemma 3.5.3 assuming (H1)

We prove Lemma 3.5.3 for the case that (H1) holds, that is, $M \subseteq R$.

Proof of Lemma 3.5.3 assuming (H1). Assume for a contradiction that H does not contain a good $1/r$-fractional matching in R or in a blue tight component of H of weight at least $|M| + \gamma n$. Note that $|V(G)| \geq (1-\varepsilon)N \geq N - 2\varepsilon n \geq (5/4 + 2\eta)n$. We will construct our fractional matching in $H[V(G)]$ ignoring the small number of vertices in $V(H) \setminus V(G)$.

It suffices to assume that M is a maximum good matching in R (that is a maximum matching in R^+) and that among all such matchings M contains the smallest number of edges f such that

$$G^{\text{blue}}[f] \neq \varnothing \text{ and } B(e) \neq B_W \text{ for all } e \in G^{\text{blue}}[f], \tag{3.5.7}$$

where $W = W(M) = V(G)\backslash V(M)$ and B_W is as in Lemma 3.5.8. Here B_W is defined since $R^+[W] = \varnothing$ by the maximality of M and $|W| \geq (5/4 + 2\eta)n - 4|M| \geq (1/4 + 2\eta)n \geq \gamma n$ as $|M| < n/4$. Let $B = B_W$.

Claim 3.5.15. *Let $f \in M$ and $e \in G^{\text{blue}}[W]$ such that $G^{\text{blue}}[f] \neq \varnothing$ and (f, e) is a suitable pair. Then f contains an edge $e' \in G^{\text{blue}}[f]$ with $B(e') = B$.*

Proof of Claim. Suppose for a contradiction that $B(e') \neq B$ for all $e' \in G^{\text{blue}}[f]$. Let $e' = x_1y_1 \in G^{\text{blue}}[f]$ and $e = x_2y_2$. By Lemma 3.5.8, $B(e) = B$. Since G is a blueprint and $B(e') \neq B$, we have $x_1x_2 \in G^{\text{red}}$. By (SP5), $x_1x_2y_2 \in \partial R$. If $f_* = x_1x_2y_1y_2 \in H^{\text{red}}$, then $f_* \in R^+$ and thus the matching $M^* = (M \setminus \{f\}) \cup \{f_*\}$ is good and has one less edge satisfying (3.5.7) than M (since $B_{W(M^*)} = B_{W(M)}$ by Lemma 3.5.8), a contradiction. Hence $f_* \in H^{\text{blue}}$. By (SP4), $x_1x_2y_2 \in \partial H(x_2y_2) = \partial B$ and by (SP3), $x_1y_1x_2 \in \partial H(x_1y_1) = \partial B(e')$. It follows that $B(e') = B$, a contradiction. ■

Let $U \subseteq W$ be a set of maximum size such that for each $u \in U$, there exists a distinct edge $f(u) \in M$ so that $(f(u), u)$ is a suitable pair and $H^{\text{red}}[f(u) \cup u] \cong K_5^{(4)}$. If $|U| \geq 4\gamma n$, then we are done since the $1/r$-fractional matching $\varphi\colon R \to [0, 1]$ with $\varphi(e) = 1$ for $e \in M \setminus \bigcup_{u \in U} f(u)$, $\varphi(e) = 1/4$ for each edge $e \in \bigcup_{u \in U} H^{\text{red}}[f(u) \cup u]$ and $\varphi(e) = 0$ for all other edges $e \in R$ is a good $1/r$-fractional matching in R of weight at least $|M| + \gamma n$. Now assume that $|U| < 4\gamma n$. Let $W' = W \setminus U$ and $M' = M \setminus \bigcup_{u \in U} f(u)$. Note that $|W'| \geq (5/4 + 2\eta)n - 4|M| - 4\gamma n \geq (1/4 + \eta)n$ and $2\delta n \leq |M| - 4\gamma n \leq |M'| \leq n/4$.

By the maximality of U, we have that

$$H^{\text{blue}}[f \cup w] \neq \varnothing \tag{3.5.8}$$

for every suitable pair $(f, w) \in M' \times W'$. Let $U' \subseteq W'$ be a set of maximum size such that there exists for each $u \in U'$, a distinct edge $f'(u) \in M'$ such that $(f'(u), u)$ is a suitable pair and $B[f'(u) \cup u] \neq \varnothing$. Let

$$W'' = W' \setminus U' \text{ and } M'' = M' \setminus \bigcup_{u \in U'} f'(u).$$

Note that $|W''| \geq |W'| - |U'| \geq (1/4 + \eta)n - |M'| \geq \eta n \geq \eta N/2$. Let $\gamma \ll \delta_0 \ll \delta$.

Claim 3.5.16. *We have $|U'| \geq |M'| - \delta_0 n$.*

Proof of Claim. Suppose not. We have $|M''| \geq \delta_0 n \geq \delta_0 N/2$. By the maximality of U', we have

$$B[f \cup w] = \varnothing \tag{3.5.9}$$

for every suitable pair $(f, w) \in M'' \times W''$. By Lemma 3.5.13, there exists $M^* \subseteq M''$ with $|M^*| = r\gamma n$ and disjoint sets $W_f \in \binom{W''}{3}$ for each $f \in M^*$ such that (f, W_f) is a suitable pair for each $f \in M^*$. Let φ_0 be the fractional matching induced by the matching $M \setminus M^*$. It suffices to show that, for every $f \in M^*$, there exists a $1/r$-fractional matching φ_f in $R[f \cup W_f]$ of weight at least $\frac{r+1}{r}$. Indeed, the completion of $\varphi_0 + \sum_{f \in M^*} \varphi_f$ with respect to R is a good $1/r$-fractional matching in R of weight at least $|M \setminus M^*| + \frac{r+1}{r}|M^*| \geq |M| + \gamma n$ giving us a contradiction.

Consider any $f = x_1x_2x_3x_4 \in M^*$. By Fact 3.5.11 and since $r = \binom{9}{4}!$, we may assume that $\bigcap R[f \cup W_f] \neq \varnothing$. We distinguish between several cases.

Case A: $\boldsymbol{G^{\mathrm{red}}[W_f] \neq \varnothing}$. Let $W_f = uvw$ with $e = uv \in G^{\mathrm{red}}[W_f]$. Recall that (f, W_f) is a suitable pair. We apply Proposition 3.5.14 with r, H, G, R, f, W_f, e playing the roles of r, H, G, R_*, f, W, e. So there exists $x \in f$ such that

$$f' \in H^{\mathrm{blue}} \text{ for any edge } f' \in H[f \cup W_f] \text{ with } e \subseteq f' \text{ and } x \notin f'. \tag{3.5.10}$$

By (3.5.8) and (3.5.9), there exists an edge $f_* \in H^{\mathrm{blue}}[f \cup u] \setminus B$. Since $f \in R$, we have $u \in f_*$. Let $yz \subseteq (f_* \cap f) \setminus x$. By (3.5.10), $yuvw, yzuv \in H^{\mathrm{blue}}$. By Lemma 3.5.8 (B1),

$uvw \in \partial B$. Hence $yuvw, yzuv, f_* \in B$, a contradiction to $f_* \notin B$.

Case B: $G^{\mathrm{red}}[W_f] = \varnothing$. By Lemma 3.5.8 (B2), we have $G[W_f] \subseteq G^{\mathrm{blue}} \subseteq B^2$. Let $uv \in B^2[W_f]$. We distinguish between the following cases. It is easy to see that these cases exhaust all possibilities.

Case B.1: $\left|G^{\mathrm{blue}}[f]\right| \geq 3$ and $G^{\mathrm{blue}}[f] \ncong K_{1,3}$. Note that $G^{\mathrm{blue}}[f]$ is connected, hence Claim 3.5.15 and the fact that G is a blueprint imply that $G^{\mathrm{blue}}[f] \subseteq B^2$. By (3.5.8), there exists an edge $f_* \in H^{\mathrm{blue}}[f \cup u]$. Observe that f_* contains an edge $xy \in G^{\mathrm{blue}}[f]$ and $u \in f_*$. By (SP3), we have $xyu \in \partial H(xy) = \partial B$. Hence $f_* \in B$, a contradiction to (3.5.9).

Case B.2: $\left|G^{\mathrm{blue}}[f]\right| \geq 2$ and $\bigcap G^{\mathrm{blue}}[f] \neq \varnothing$. Without loss of generality assume that $x_1x_2, x_1x_3 \in G^{\mathrm{blue}}[f]$ and $x_2x_3, x_2x_4, x_3x_4 \in G^{\mathrm{red}}$. By Claim 3.5.15 and the fact that G is a blueprint, we have that $G^{\mathrm{blue}}[f] \subseteq B^2$. By (3.5.8) and (3.5.9), there exists an edge $f_* \in H^{\mathrm{blue}}[f \cup u] \setminus B$. Observe that $f_* = x_2x_3x_4u$ since, by (SP3), $x_1x_2u, x_1x_3u \in \partial B$. Let $e_{f,1} = x_2x_3uv$, $e_{f,2} = x_2x_4uv$, $e_{f,3} = x_3x_4uv$. Note that, for all $i \in [3]$, we have $e_{f,i} \cap f \in R^2$ and $e_{f,i} \cap W_f \in B^2$ and thus, by (SP3) and (SP4), we have $e_{f,i} \in R \cup B$. Since $f_* \notin B$, we have $e_{f,i} \in R$ for all $i \in [3]$. We are done since $\{f, e_{f,1}, e_{f,2}, e_{f,3}\} \subseteq R[f \cup W_f]$ has an empty intersection, a contradiction.

Case B.3: $G^{\mathrm{red}}[f]$ contains a copy of C_4. Without loss of generality assume that $x_1x_2, x_2x_3, x_3x_4, x_1x_4 \in G^{\mathrm{red}}[f]$. By (SP4), we have $uvx_j \in \partial B$ for all $j \in [4]$. By (SP3), we have $x_1x_2u, x_2x_3u, x_3x_4u, x_1x_4u \in \partial R$ for all $i \in [4]$.

If x_1x_2uv and x_3x_4uv are red, then $F = \{f, x_1x_2uv, x_3x_4uv\} \subseteq R$ has an empty intersection, a contradiction. So we may assume that x_1x_2uv is blue and thus in B. Similarly, by considering $\{f, x_1x_4uv, x_2x_3uv\}$, we may assume that $x_1x_4uv \in B$. By (3.5.8) and (3.5.9), there exists an edge $f_* \in H^{\mathrm{blue}}[f \cup u] \setminus B$. Since $x_1x_2uv, x_1x_4uv \in B$, we have $f_* = x_2x_3x_4u$ and $x_2x_3uv, x_2x_4uv, x_3x_4uv \in R$. Thus we obtain a contradiction as $\{f, x_2x_3uv, x_2x_4uv, x_3x_4uv\} \subseteq R$ has an empty intersection. ■

For the remainder of the proof, our aim is to find a good $1/r$-fractional matching in B of weight at least $|M| + \gamma n$. For each $u \in U'$, choose an edge $f_u^* \in B[f'(u) \cup u]$

which exists by the definition of U'. Since $(f'(u), u)$ is a suitable pair, f_u^* is good. Let $M_1^* = \{f_u^* \colon u \in U'\}$ and note that M_1^* is a good matching in B. Note

$$|M_1^*| = |U'| \geq |M'| - \delta_0 n \geq |M| - 2\delta_0 n.$$

Let $M_2^* \subseteq B^+[V(G) \setminus V(M_1^*)]$ be a maximum matching. If $|M_1^*| + |M_2^*| \geq |M| + \gamma n$, then we are done. Thus we may assume $|M_1^*| + |M_2^*| < |M| + \gamma n$, so $|M_2^*| \leq 3\delta_0 n$. Let

$$\begin{aligned} U'' &= \{u \in U' \colon (f'(u) \cup u) \cap V(M_2^*) = \varnothing\}, \\ M_0 &= \bigcup_{u \in U''} f'(u) \text{ and} \\ W_0 &= W'' \setminus V(M_2^*) = W'. \end{aligned}$$

We have

$$\begin{aligned} &|U''| \geq |U'| - 4\,|M_2^*| \geq |M| - 14\delta_0 n \geq 2\delta n \geq \delta N, \\ &|M_0| = |U''| \geq \delta N \text{ and} \\ &|W_0| \geq |W''| - 4\,|M_2^*| \geq \eta n/2 \geq \eta N/4. \end{aligned}$$

By Lemma 3.5.13, there exist a subset $U_0 \subseteq U''$ corresponding to the matching $\{f'(u) \colon u \in U_0\} \subseteq M_0$ of size $3r\delta_0 n$ and disjoint sets $W_u \in \binom{W_0}{4}$ for each $u \in U_0$ such that $(f'(u), W_u)$ is a suitable pair for each $u \in U_0$.

We now construct a good $1/r$-fractional matching $\varphi \colon B \to [0,1]$ in B as follows. Let φ_0 be the fractional matching induced by the matching $(M_1^* \setminus \{f_u^* \colon u \in U_0\}) \cup M_2^*$. Suppose that, for each $u \in U_0$, there exists a good $1/r$-fractional matching φ_u in $B[f'(u) \cup u \cup W_u]$ of weight at least $\frac{r+1}{r}$. Then the completion of $\varphi_0 + \sum_{u \in U_0} \varphi_u$ with respect to B is a good $1/r$-fractional matching in B of weight at least $|M_1^*| + |M_2^*| + |U_0|/r \geq |M| - 2\delta_0 n + 3\delta_0 n \geq |M| + \gamma n$. Thus it suffices to show that, for each $u \in U_0$, there exists a good $1/r$-fractional matching φ_u in $B[f'(u) \cup u \cup W_u]$ of weight at least $\frac{r+1}{r}$. Note that $B[f'(u) \cup W_u] \cup \{f_u^*\} \subseteq B^+$. By Fact 3.5.11, it suffices to show that $\bigcap (B[f'(u) \cup W_u] \cup \{f_u^*\}) = \varnothing$.

Consider any $u \in U_0$. Let

$$f'(u) = yz_1z_2z_3 \in R,\ f_u^* = z_1z_2z_3u \in B \text{ and } W_u = w_1w_2w_3w_4.$$

By the maximality of M_2^*, we have $w_1w_2w_3w_4 \notin B$. Hence (B1) implies that $G^{\text{red}}[W_u] = \varnothing$. Thus by Lemma 3.5.8 (B2), we have $G[W_u] \subseteq B^2$. In particular, $w_1w_2 \in B^2$ and thus (SP4) and (SP6) imply $yw_1w_2, w_1w_2w_3 \in \partial B$. By the maximality of M_2^* and the maximality of M, we have that $yw_1w_2w_3, w_1w_2w_3w_4 \in H^{\text{red}} \setminus R$.[1]

We distinguish between two cases.

Case A: At least two of yz_1, yz_2, yz_3 are in G^{red}. Without loss of generality assume that $yz_1, yz_2 \in G^{\text{red}}$. By (SP3), we have $yz_1w_1, yz_2w_1 \in \partial R$. Since $yw_1w_2w_3 \in H^{\text{red}} \setminus R$ and $yw_1w_2 \in \partial B$, we have $yz_1w_1w_2, yz_2w_1w_2 \in B$. Thus we are done since $\{z_1z_2z_3u, yz_1w_1w_2, yz_2w_1w_2\} \subseteq B$ has an empty intersection.

Case B: At least two of yz_1, yz_2, yz_3 are in G^{blue}. Without loss of generality assume that $yz_1, yz_2 \in G^{\text{blue}}$, so $G^{\text{blue}}[yz_1z_2z_3]$ is connected. Since $w_1w_2 \in G^{\text{blue}}[W]$ and $(yz_1z_2z_3, w_1w_2)$ is a suitable pair, Claim 3.5.15 implies that $B(yz_1) = B(yz_2) = B$ and thus $yz_1w_1, yz_2w_1 \in \partial B$ by (SP3). If $yz_1w_1w_2, yz_2w_1w_2 \in H^{\text{blue}}$, then $yz_1w_1w_2, yz_2w_1w_2 \in B$ (since $w_1w_2y \in \partial B$ by (SP4)). Note that $\{z_1z_2z_3u, yz_1w_1w_2, yz_2w_1w_2\} \subseteq B$ has an empty intersection and thus we are done.

Hence, we may assume without loss of generality that $yz_1w_1w_2$ is red. Since $yw_1w_2w_3 \in H^{\text{red}} \setminus R$ and $yz_1w_1 \in \partial B$, we have $yz_1z_2w_1, yz_1z_3w_1 \in B$. If $yz_2w_1w_2$ is red, then $yz_2z_3w_1 \in B$ (since $yw_1w_2w_3 \in H^{\text{red}} \setminus R$ and $yz_2w_1 \in \partial B$). Thus we are done since $\{z_1z_2z_3u, yz_1z_2w_1, yz_1z_3w_1, yz_2z_3w_1\} \subseteq B$ has an empty intersection. If $yz_2w_1w_2$ is blue, then we have $yz_2w_1w_2 \in B$ since $yw_1w_2 \in \partial B$. Thus we are done since $\{z_1z_2z_3u, yz_1z_3w_1, yz_2w_1w_2\} \subseteq B$ has an empty intersection.

This completes the proof. $\square$

[1] Since $yz_1z_2z_3 \in R$ and R is a red tight component, this implies that we have $yz_{i_1}z_{i_2}w_{i_3} \in H^{\text{blue}}$ or $yz_{i_1}w_{i_2}w_{i_3} \in H^{\text{blue}}$ for all distinct $i_1, i_2, i_3 \in [4]$.

3.5.4 Proof of Lemma 3.5.3 assuming (H2)

We now prove the remaining case of Lemma 3.5.3, that is when M is contained in a blue tight component B of H. Note that the proof is similar to the proof for the case where we assume (H1).

Proof of Lemma 3.5.3 assuming (H2). Assume for a contradiction that H does not contain a good $1/r$-fractional matching in R or in a blue tight component of H of weight at least $|M| + \gamma n$. Note that $|V(G)| \geq (1-\varepsilon)N \geq N - 2\varepsilon n \geq (5/4 + 2\eta)n$. We will construct all our good fractional matching in $H[V(G)]$ ignoring the small number of vertices in $V(H) \setminus V(G)$.

It suffices to assume that M is a maximum good matching in B, that is a maximum matching in B^+. Let $W = V(G) \setminus V(M)$. Note that $B^+[W] = \varnothing$.

Claim 3.5.17. *If $f \in M$ is an edge such that $G^{\mathrm{blue}}[f]$ contains a triangle or a matching of size* 2*, then $G^{\mathrm{blue}}[f]$ contains an edge $e \in B^2$. Moreover, if $\left|G^{\mathrm{blue}}[f]\right| \geq 4$, then $G^{\mathrm{blue}}[f] \subseteq B^2$.*

Proof of Claim. Let f be such an edge in M. Since M is a good matching, there exists $z \in f$ such that $xyz \in \partial H(xy)$ for every $xy \in \binom{f \setminus \{z\}}{2}$. Observe that there exists $e \in \binom{f \setminus \{z\}}{2} \cap G^{\mathrm{blue}}$. Hence $e \cup z \in \partial B(e)$. Since $f \in B$, we have $B(e) = B$, that is, $e \in B^2$.

If $\left|G^{\mathrm{blue}}[f]\right| \geq 4$, then $G^{\mathrm{blue}}[f]$ contains a triangle or a matching of size 2 and thus by the previous argument $G^{\mathrm{blue}}[f]$ contains an edge $e \in B^2$. Moreover, $G^{\mathrm{blue}}[f]$ is connected and thus, since G is a blueprint, we have $G^{\mathrm{blue}}[f] \subseteq B^2$. ■

Let $U \subseteq W$ be a set of maximum size such that for each $u \in U$ there exists a distinct edge $f(u) \in M$ for which $(f(u), u)$ is a suitable pair and $H^{\mathrm{blue}}[f(u) \cup u] \cong K_5^{(4)}$. If $|U| \geq 4\gamma n$, then we are done since the $1/r$-fractional matching $\varphi\colon B \to [0,1]$ with $\varphi(e) = 1$ for $e \in M \setminus \bigcup_{u \in U} f(u)$, $\varphi(e) = 1/4$ for each edge $e \in \bigcup_{u \in U} H^{\mathrm{blue}}[f(u) \cup u]$ and $\varphi(e) = 0$ for all other edges $e \in B$ is a good $1/r$-fractional matching in B and has weight at least $|M| + \gamma n$. Now assume that $|U| < 4\gamma n$.

Let $W' = W \setminus U$ and $M' = M \setminus \bigcup_{u \in U} f(u)$. Note that

$$|W'| \geq (5/4 + 2\eta)n - 4|M| - 4\gamma n \geq (1/4 + \eta)n \tag{3.5.11}$$

and $2\delta n \leq |M| - 4\gamma n \leq |M'| \leq n/4$. By the maximality of U, we have

$$H^{\mathrm{red}}[f \cup w] \neq \varnothing \tag{3.5.12}$$

for every suitable pair $(f, w) \in M' \times W'$.

We distinguish between two cases.

Case 1: $G^{\mathrm{red}}[W'] = \varnothing$ and $B^2[W'] \neq \varnothing$. Recall that $|W'| \geq (1/4 + \eta)n$ and $\delta(G) \geq (1 - \varepsilon)N$. So $G[W']$ is connected. Hence $G[W'] \subseteq B^2$. We start by proving the following claim.

Claim 3.5.18. *There exists a red tight component R_* of H such that $H^+[W'] \subseteq R_*^+$.*

Proof of Claim. By the maximality of M and $G[W'] \subseteq B^2$, we have $H^+[W'] \subseteq H^{\mathrm{red}}$.[1] By Corollary 3.5.7, $H^+[W']$ is tightly connected. Hence there exists a red tight component R_* of H such that $H^+[W'] \subseteq R_*^+$. ■

Let $U' \subseteq W'$ be a set of maximum size such that for each $u \in U'$, there exists a distinct edge $f'(u) \in M'$ so that $(f'(u), u)$ is a suitable pair and $R_*[f'(u) \cup u] \neq \varnothing$. Let $W'' = W' \setminus U'$ and $M'' = M' \setminus \bigcup_{u \in U'} f'(u)$. Note that

$$|W''| \geq |W'| - |U'| \geq (1/4 + \eta)n - |M'| \geq \eta n.$$

Let δ_0 be a new constant such that $\gamma \ll \delta_0 \ll \delta$.

Claim 3.5.19. *We have $|U'| \geq |M'| - \delta_0 n \geq |M| - 2\delta_0 n$.*

[1]Suppose there was an edge $f \in H^+[W'] \cap H^{\mathrm{blue}}$. Since $f \in H^+$, by (G3), there exists $z \in f$ such that $xyz \in \partial H(xy)$ for all $xy \in \binom{f \setminus \{z\}}{2}$. Let $xy \in \binom{f \setminus \{z\}}{2}$. Since $G[W'] \subseteq B^2$, we have $xy \in B^2$ and thus $xyz \in \partial B$. Hence $f \in B$, a contradiction to the maximality of M.

Proof of Claim. Suppose not. We have $|M''| \geq \delta_0 n \geq \delta_0 N/2$. By the maximality of U', we have

$$R_*[f \cup w] = \varnothing \tag{3.5.13}$$

for every suitable pair $(f, w) \in M'' \times W''$. By Lemma 3.5.13, there exists $M^* \subseteq M''$ with $|M^*| = r\gamma n$ and disjoint sets $W_f \in \binom{W''}{4}$ for each $f \in M^*$ such that (f, W_f) is a suitable pair for each $f \in M^*$. Let φ_0 be the fractional matching induced by the matching $M \setminus M^*$. It suffices to show that, for every $f \in M^*$, there exists a good $1/r$-fractional matching φ_f in $B[f \cup W_f]$ of weight at least $\frac{r+1}{r}$. Indeed the completion of $\varphi_0 + \sum_{f \in M^*} \varphi_f$ with respect to B is a good $1/r$-fractional matching in B of weight at least $|M \setminus M^*| + \frac{r+1}{r}|M^*| \geq |M| + \gamma n$ giving us a contradiction.

Consider any $f = x_1x_2x_3x_4 \in M^*$. By Fact 3.5.11, $r = \binom{9}{4}!$ and $B[f \cup W_f] \subseteq B^+$, we may assume that $\bigcap B[f \cup W_f] \neq \varnothing$.

Let $W_f = w_1w_2w_3w_4 \in H^+[W''] \subseteq R_*^+$. Let $W_f^* = w_1w_2w_3$ and $e = w_1w_2$. Note that (f, W_f^*) is a suitable pair and $e \in B^2$ since $G[W'] \subseteq B^2$. We apply Proposition 3.5.14 with colours reversed and r, H, G, B, f, W_f^*, e playing the roles of r, H, G, R_*, f, W, e. So there exists $x \in f$ such that

$$f' \in H^{\text{red}} \text{ for any edge } f' \in H[f \cup W_f^*] \text{ with } e = w_1w_2 \subseteq f' \text{ and } x \notin f'. \tag{3.5.14}$$

By (3.5.12) and (3.5.13), there exists an edge $f_* \in H^{\text{red}}[f \cup w_1] \setminus R_*$. Since $f \in B$, we have $w_1 \in f_*$. Let $yz = (f_* \cap f) \setminus x$. By (3.5.14), $yw_1w_2w_3, yzw_1w_2 \in H^{\text{red}}$. Since $w_1w_2w_3w_4 \in H^+[W''] \subseteq R_*^+$, $w_1w_2w_3 = W_f^* \in \partial R_*$. Hence $yw_1w_2w_3, yzw_1w_2, f_* \in R_*$, a contradiction to $f_* \notin R_*$. ■

We now find a matching in R_*^+ as follows. For each $u \in U'$, choose an edge $f_u^* \in R_*[f'(u) \cup u]$ and note that, since $(f'(u), u)$ is a suitable pair, f_u^* is good. Let $M_1^* = \{f_u^* \colon u \in U'\}$, so M_1^* is a matching in $R_*^+[V(H) \setminus W'']$. By Claim 3.5.18, $H^+[W''] \subseteq R_*^+$. By Proposition 3.5.5, $R_*^+[W'']$ contains a matching M_2^* of size at least $\frac{|W''|}{4} - \gamma N \geq$

$\frac{|W''|}{4} - 2\gamma n$.

Thus $M_1^* \cup M_2^*$ is a matching in R_*^+ of size

$$\begin{aligned} |M_1^*| + |M_2^*| &\geq |U'| + \frac{|W''|}{4} - 2\gamma n = \frac{1}{4}(3\,|U'| + |W'|) - 2\gamma n \\ &\overset{\text{Claim 3.5.19, (3.5.11)}}{\geq} \frac{1}{4}(3\,|M| - 6\delta_0 n + (5/4 + 2\eta)n - 4\,|M| - 4\gamma n) - 2\gamma n \\ &\geq \frac{1}{4}((5/4+\eta)n - |M|) \geq \frac{n}{4}, \end{aligned}$$

where the last inequality holds as $|M| < n/4$. Hence H contains a good monochromatic tightly connected matching of size at least $n/4$, a contradiction.

Case 2: $G^{\text{red}}[W'] \neq \varnothing$ or $B^2[W'] = \varnothing$. Recall that by the maximality of M, we have $B^+[W'] = \varnothing$.

Claim 3.5.20. *Let $y_1y_2y_3y_4 \in H^+[W']$ with $y_1y_2 \in B^2$ and $y_1y_2y_3 \in \partial B$. Then $y_1y_2y_3y_4 \in R^+$.*

Proof of Claim. If $B^2[W'] = \varnothing$, then this is vacuously true. Hence we may assume that $G^{\text{red}}[W'] \neq \varnothing$. Suppose to the contrary, that $y_1y_2y_3y_4 \notin R^+$. Since $B^+[W'] = \varnothing$, we have $y_1y_2y_3y_4 \in H^{\text{red}} \setminus R$. Let $x_1x_2 \in G^{\text{red}}[W']$. By Proposition 3.5.6, there exist vertices $z_1, z_2, z_3 \in W'$ such that $H^+[y_1y_2y_3y_4z_1z_2z_3] \cong K_7^{(4)}$, $H^+[x_1x_2z_1z_2z_3] \cong K_5^{(4)}$, $y_1y_2z_1 \in \partial H(y_1y_2)$, $y_1z_1z_2 \in \partial H(y_1z_1)$, $z_1z_2z_3 \in \partial H(z_1z_2)$, $x_1z_1z_2 \in \partial H(x_1z_1)$ and $x_1x_2z_1 \in \partial H(x_1x_2)$.

Since $y_1y_2 \in B^2$, we have $y_1y_2z_1 \in \partial B$. Since $B^+[W'] = \varnothing$ and $y_1y_2y_3y_4 \in H^{\text{red}} \setminus R$, we have $y_1y_2z_1z_2, y_1y_2y_3z_1 \in H^{\text{red}} \setminus R$. This implies that $y_1z_1 \in G^{\text{blue}}$ (or else $y_1y_2z_1z_2 \in R$) and so $y_1z_1 \in B^2$. Thus $y_1z_1z_2 \in \partial B$ and $y_1z_1z_2z_3 \in H^{\text{red}} \setminus R$ (or else $B^+[W'] \neq \varnothing$). Similarly, we deduce that $z_1z_2 \in G^{\text{blue}}$ and so $z_1z_2 \in B^2$. Thus $z_1z_2z_3 \in \partial B$ and since $B^+[W'] = \varnothing$, $x_1z_1z_2z_3 \in H^{\text{red}} \setminus R$. It follows that $x_1z_1 \in G^{\text{blue}}$ and so $x_1z_1 \in B^2$. Thus $x_1z_1z_2 \in \partial B$ and since $B^+[W'] = \varnothing$, $x_1x_2z_1z_2 \in H^{\text{red}}$. Since $x_1x_2 \in G^{\text{red}}$, we have $x_1x_2z_1 \in \partial R$ and thus $y_1y_2y_3y_4 \in R^+$, a contradiction. ■

Let $U' \subseteq W'$ be a set of maximum size such that for each $u \in U'$, there exists a distinct edge $f'(u) \in M'$ for which $(f'(u), u)$ is a suitable pair and $R[f'(u) \cup u] \neq \varnothing$. Let $W'' = W' \setminus U'$ and $M'' = M' \setminus \bigcup_{u \in U'} f'(u)$. Note that $|W''| = |W'| - |U'| \geq (1/4 + \eta)n - |M'| \geq \eta n \geq \eta N/2$. Let δ_0 be a new constant such that $\gamma \ll \delta_0 \ll \delta$.

Claim 3.5.21. *We have $|U'| \geq |M'| - \delta_0 n$.*

Proof of Claim. Suppose not. We have $|M''| \geq \delta_0 n \geq \delta_0 N/2$. By the maximality of U', we have

$$R[f \cup w] = \varnothing \tag{3.5.15}$$

for every suitable pair $(f, w) \in M'' \times W''$. By Lemma 3.5.13, there exists $M^* \subseteq M''$ with $|M^*| = r\gamma n$ and disjoint sets $W_f \in \binom{W''}{4}$ for each $f \in M^*$ such that (f, W_f) is a suitable pair for each $f \in M^*$.

Let φ_0 be the fractional matching induced by the matching $M \setminus M^*$. Suppose that, for every $f \in M^*$, there exists good a $1/r$-fractional matching φ_f in $B[f \cup W_f]$ of weight at least $\frac{r+1}{r}$. Then the completion of $\varphi_0 + \sum_{f \in M^*} \varphi_f$ with respect to B is a good $1/r$-fractional matching in B of weight at least $|M \setminus M^*| + \frac{r+1}{r}|M^*| \geq |M| + \gamma n$. Thus it suffices to show that, for every $f \in M^*$, there exists a $1/r$-fractional matching φ_f in $B[f \cup W_f]$ of weight at least $\frac{r+1}{r}$.

Consider any $f = x_1x_2x_3x_4 \in M^*$. By Fact 3.5.11, $B[f \cup W_f] \subseteq B^+$ and $r = \binom{9}{4}!$, it suffices to show that $\bigcap B[f \cup W_f] = \varnothing$. Let $uv \in G[W_f]$. We distinguish between several cases.

Case A: $\left|G^{\text{red}}[f]\right| \geq 3$ with $G^{\text{red}}[f] \ncong K_{1,3}$ or $G^{\text{red}}[f]$ is a matching of size 2. By (3.5.12), there exists an edge $f_* \in H^{\text{red}}[f \cup u]$. Observe that f_* contains an edge $xy \in G^{\text{red}}[f]$ and $u \in f_*$. By (SP3), we have $xyu \in \partial H(xy) = \partial R$. Hence $f_* \in R$, a contradiction to (3.5.15).

Case B: $uv \in G^{\text{red}}$, $\left|G^{\text{red}}[f]\right| \geq 2$ and $\bigcap G^{\text{red}}[f] \neq \varnothing$. Without loss of generality assume that $x_1x_2, x_1x_3 \in G^{\text{red}}[f]$ and $x_2x_3, x_2x_4, x_3x_4 \in G^{\text{blue}}$. By Claim 3.5.17, we have

$B(x_2x_3) = B(x_2x_4) = B(x_3x_4) = B$. By (3.5.12) and (3.5.15), there exists an edge $f_* \in H^{\text{red}}[f \cup u] \setminus R$. Observe that $f_* = x_2x_3x_4u$ since, by (SP3), $x_1x_2u, x_1x_3u \in \partial R$. Let $e_{f,1} = x_2x_3uv$, $e_{f,2} = x_2x_4uv$, $e_{f,3} = x_3x_4uv$. Note that, for all $i \in [3]$, we have $e_{f,i} \cap f \in B^2$ and $e_{f,i} \cap W_f = uv \in R^2$ and thus, by (SP3) and (SP4), we have $e_{f,i} \in R \cup B$. Since $f_* \notin R$, we have $e_{f,i} \in B$ for all $i \in [3]$. We are done since $\{f, e_{f,1}, e_{f,2}, e_{f,3}\} \subseteq B$ has an empty intersection.

Case C: $\boldsymbol{uv \in G^{\text{red}}}$ and $\boldsymbol{G^{\text{blue}}[f]}$ contains a copy of $\boldsymbol{C_4}$. We assume without loss of generality that $x_1x_2, x_2x_3, x_3x_4, x_1x_4 \in G^{\text{blue}}[f]$. By Claim 3.5.17, $x_1x_2, x_2x_3, x_3x_4, x_1x_4 \in B^2$. By (SP4), we have $uvx_i \in \partial R$ for all $i \in [4]$. By (SP3), we have $x_1x_2u, x_2x_3u, x_3x_4u$, $x_1x_4u \in \partial B$. If x_1x_2uv and x_3x_4uv are blue, then both are in B^+ and together with f they form a set $F \subseteq B[f \cup W_f]$ with $\bigcap F = \varnothing$. So we may assume that x_1x_2uv is red and thus in R. Similarly, we may assume that $x_1x_4uv \in R$. By (3.5.12) and (3.5.15), there exists an edge $f_* \in H^{\text{red}}[f \cup u] \setminus R$. Since $x_1x_2uv, x_1x_4uv \in R$, we have $f_* = x_2x_3x_4u$ and $x_2x_3uv, x_2x_4uv, x_3x_4uv \in B$. Thus we are done since $F = \{f, x_2x_3uv, x_2x_4uv, x_3x_4uv\}$ has an empty intersection.

Case D: $\boldsymbol{uv \in B^2}$. Let $W_f^* = uvw \subseteq W_f$. Note that since (f, W_f) is a suitable pair and $W_f^* \subseteq W_f$, we have that (f, W_f^*) is a suitable pair. Suppose for a contradiction, that $\bigcap B[f \cup W_f] \neq \varnothing$. We apply Proposition 3.5.14 with colours reversed and r, H, G, B, f, W_f^*, e playing the roles of r, H, G, R_*, f, W, e. We have that there exists $x \in f$ such that

$$f' \in H^{\text{red}} \text{ for any edge } f' \in H[f \cup W_f^*] \text{ with } e = uv \subseteq f' \text{ and } x \notin f'. \tag{3.5.16}$$

By (3.5.12) and (3.5.15), there exists an edge $f_* \in H^{\text{red}}[f \cup u] \setminus R$. Since $f \in B$, we have $u \in f_*$. Let $yz \subseteq (f_* \cap f) \setminus x$. By (3.5.16), $yuvw, yzuv \in H^{\text{red}}$. Let $w' \in W_f \setminus uvw$. Since $uv \in B^2$, we have $uvw \in \partial B$. Since (f, W_f) is a suitable pair, $uvww' \in H^+$. By Claim 3.5.20, we have $uvww' \in R^+$ and thus $uvw \in \partial R$. Hence $yuvw, yzuv, f_* \in R$, a contradiction to $f_* \notin R$.

Case E: $\boldsymbol{uv \in G^{\text{blue}} \setminus B^2}$. If $\left|G^{\text{red}}[f]\right| \geq 3$ with $G^{\text{red}}[f] \ncong K_{1,3}$ or $G^{\text{red}}[f]$ is a matching of size 2, then we are in Case A. Hence we may assume that $G^{\text{blue}}[f]$ contains a triangle. We assume without loss of generality that $x_1x_2, x_2x_3, x_1x_3 \in G^{\text{blue}}$. By Claim 3.5.17, we have $G^{\text{blue}}[f] \subseteq B^2$. By (SP3), we have $x_1x_2u, x_2x_3u, x_1x_3u \in \partial B$. Let $B_* = B(uv) \neq B$. By (SP4), we have $uvx_i \in \partial B_*$ for all $i \in [3]$. Hence, since $B_* \neq B$, we have $E = \{x_1x_2uv, x_2x_3uv, x_1x_3uv\} \subseteq H^{\text{red}}$. Since $uv \in B_*^2$ and $B_* \neq B$, we have $x_iu \in G^{\text{red}}$ for all $i \in [3]$. By (SP5), we have $x_iuv \in \partial R$ for all $i \in [3]$. Hence $E \subseteq R$. It follows that $x_1x_2u, x_2x_3u, x_1x_3u \in \partial R$. This contradicts the fact that $H^{\text{red}}[f \cup u] \setminus R \neq \varnothing$ which holds by (3.5.12) and (3.5.15). ■

For the remainder of the proof, our aim is to find a good $1/r$-fractional matching in R of weight at least $|M| + \gamma n$. For each $u \in U'$, choose an edge $f_u^* \in R[f'(u) \cup u]$. Since $(f'(u), u)$ is a suitable pair, f_u^* is good. Let $M_1^* = \{f_u^* \colon u \in U'\}$, so M_1^* is a good matching in R. Note

$$|M_1^*| = |U'| \geq |M'| - \delta_0 n \geq |M| - 2\delta_0 n.$$

Let $M_2^* \subseteq R^+$ be a maximum matching vertex-disjoint from M_1^*. If $|M_1^*|+|M_2^*| \geq |M|+\gamma n$, then we are done. Thus we may assume $|M_1^*| + |M_2^*| < |M| + \gamma n$, so $|M_2^*| \leq 3\delta_0 n$. Let $U'' = \{u \in U' \colon (f'(u) \cup u) \cap V(M_2^*) = \varnothing\}$. We have $|U''| \geq |U'| - 4\,|M_2^*| \geq |M| - 14\delta_0 n \geq 2\delta n \geq \delta N$. Let $M_0 = \bigcup_{u \in U''} f'(u)$ and note that $|M_0| = |U''| \geq \delta N$. Recall that $W'' = W' \setminus U'$ and $|W''| \geq \eta N/2$ and let $W_0 = W'' \setminus V(M_2^*)$ and note that $|W_0| \geq |W''| - 4\,|M_2^*| \geq \eta N/4$. By Lemma 3.5.13, there exist a subset $U_0 \subseteq U''$ corresponding to the matching $\bigcup_{u \in U_0} f'(u) \subseteq M_0$ of size $3r\delta_0 n$ and disjoint sets $W_u \in \binom{W_0}{4}$ for each $u \in U_0$ such that $(f'(u), W_u)$ is a suitable pair for each $u \in U_0$.

We now construct a good $1/r$-fractional matching $\varphi \colon R \to [0,1]$ in R as follows. Let φ_0 be the fractional matching induced by the matching $(M_1^* \setminus \{f_u^* \colon u \in U_0\}) \cup M_2^*$. Suppose that, for each $u \in U_0$, there exists a good $1/r$-fractional matching φ_u in $R[f'(u) \cup u \cup W_u]$ of weight at least $\frac{r+1}{r}$. Then the completion of $\varphi_0 + \sum_{u \in U_0} \varphi_u$ with respect to R is a good $1/r$-fractional matching in R of weight at least $|M_1^*|+|M_2^*|+|U_0|/r \geq |M| - 2\delta_0 n + 3\delta_0 n \geq |M| + \gamma n$. Thus it suffices to show that, for each $u \in U_0$, there exists a good $1/r$-fractional matching φ_u in

$R[f'(u) \cup u \cup W_u]$ of weight at least $\frac{r+1}{r}$.

Consider any $u \in U_0$. Note that f_u^* is good and since $(f'(u), W_u)$ is a suitable pair, any edge in $H[f'(u) \cup W_u]$ is good. By Fact 3.5.11, it suffices to show that $\bigcap(R[f'(u) \cup W_u] \cup \{f_u^*\}) = \varnothing$.

Let

$$f'(u) = yz_1z_2z_3 \in B,\ f_u^* = z_1z_2z_3u \in R \text{ and } W_u = w_1w_2w_3w_4.$$

By the maximality of M, we have $w_1w_2w_3w_4 \notin B$. By the maximality of M_2^*, we have $w_1w_2w_3w_4 \notin R$. Hence Claim 3.5.20 and (SP6) imply that $B^2[W_u] = \varnothing$. It follows that the following cases exhaust all possibilities.

Case A: $G^{\text{red}}[W_u] \neq \varnothing$. We assume without loss of generality that $w_1w_2 \in G^{\text{red}}$. By (SP6) and (SP4), $w_1w_2w_3, w_1w_2y \in \partial R$. Since $w_1w_2w_3w_4 \notin R \cup B$, we have $w_1w_2w_3w_4 \in H^{\text{blue}} \setminus B$. By the maximality of M_2^*, we have $yw_1w_2w_3 \in H^{\text{blue}} \setminus B$. We now consider the colours of the edges yz_i for $i \in [3]$.

Case A.1: At least two edges in $\{yz_i\colon i \in [3]\}$ are in B^2. We assume without loss of generality that $yz_1, yz_2 \in B^2$. By (SP3), $yz_1w_1, yz_2w_1 \in \partial B$. Since $yw_1w_2w_3 \in H^{\text{blue}} \setminus B$, we have $yz_1w_1w_2, yz_2w_1w_2 \in H^{\text{red}}$. Since $w_1w_2y \in \partial R$, we have $\{z_1z_2z_3u,\ yz_1w_1w_2, yz_2w_1w_2\} \subseteq R$. Moreover, this set has an empty intersection and so we are done.

Case A.2: At least two edges in $\{yz_i\colon i \in [3]\}$ are in G^{red}. We assume without loss of generality that $yz_1, yz_2 \in G^{\text{red}}$. By (SP3), we have $yz_1w_1, yz_2w_1 \in \partial R$. We distinguish between the following three subcases.

If $yz_1w_1w_2, yz_2w_1w_2$ are both red, then, $yz_1w_1w_2, yz_2w_1w_2 \in R$ as $yw_1w_2 \in \partial R$. We are done since $\{z_1z_2z_3u, yz_1w_1w_2, yz_2w_1w_2\} \subseteq R$ has an empty intersection.

If $yz_1w_1w_2$ is blue and $yz_2w_1w_2$ is red, then since $yz_2w_1 \in \partial R$ and $yw_1w_2w_3 \in H^{\text{blue}} \setminus B$, we have $yz_2w_1w_2 \in R$ and $yz_1w_1w_2 \in H^{\text{blue}} \setminus B$. From $yz_1z_2z_3 \in B$ and $yz_1w_1w_2 \in H^{\text{blue}} \setminus B$, it follows that $yz_1z_3w_1 \in H^{\text{red}}$. Since $yz_1w_1 \in \partial R$, we have $yz_1z_3w_1 \in R$. We are done since $\{z_1z_2z_3u, yz_1z_3w_1, yz_2w_1w_2\} \subseteq R$ has an empty intersection.

Hence we may assume that $yz_1w_1w_2, yz_2w_1w_2$ are both blue. Since $yw_1w_2w_3 \in H^{\text{blue}} \setminus B$

and $yz_1z_2z_3 \in B$, we have $yz_1z_2w_1, yz_1z_3w_1, yz_2z_3w_1 \in H^{\text{red}}$. From $yz_1w_1, yz_2w_1 \in \partial R$, it follows that $yz_1z_2w_1, yz_1z_3w_1, yz_2z_3w_1 \in R$. We are done since $\{z_1z_2z_3u, yz_1z_2w_1, yz_1z_3w_1, yz_2z_3w_1\} \subseteq R$ has an empty intersection.

Case A.3: At least two edges in $\{yz_i : i \in [3]\}$ are in $G^{\text{blue}} \setminus B^2$. We assume without loss of generality that $yz_1, yz_2 \in G^{\text{blue}} \setminus B^2$. Let $B_* = B(yz_1) = B(yz_2)$ and note that $B_* \neq B$. We have $z_1z_2, z_1z_3, z_2z_3 \in G^{\text{red}}$ (or else Claim 3.5.17 implies $yz_1, yz_2 \in B^2$). Since $yz_1w_1, yz_2w_1 \in \partial B_*$ by (SP3) and $yz_1z_2z_3 \in B \neq B_*$, we have $yz_1z_2w_1, yz_1z_3w_1, yz_2z_3w_1 \in H^{\text{red}}$. Since $z_1z_2w_1, z_1z_3w_1, z_2z_3w_1 \in \partial R$ by (SP3), we have $yz_1z_2w_1, yz_1z_3w_1, yz_2z_3w_1 \in R$. We are done since $\{z_1z_2z_3u, yz_1z_2w_1, yz_1z_3w_1, yz_2z_3w_1\} \subseteq R$ has an empty intersection.

Case B: $G[W_u] \subseteq G^{\text{blue}} \setminus B^2$. Since G is a blueprint all the edges in $G[W_u]$ induce the same blue tight component $B_* \neq B$ of H. By (SP4) and (SP6), yw_1w_2, yw_1w_3, yw_2w_3, $w_1w_2w_3 \in \partial B_*$.

Case B.1: At least one edge in $\{yz_i : i \in [3]\}$ is in B^2. We assume without loss of generality that $yz_1 \in B^2$. By (SP4), $yz_1w_1 \in \partial B$. Note that $yw_1 \in G^{\text{red}}$ (else $B_* = B$ since G is a blueprint). By (SP5), $yw_1w_2 \in \partial R$ and the maximality of M_2^* implies $yw_1w_2w_3 \in B_*$. Since $B \neq B_*$, $yz_1w_1 \in \partial B$ and $yw_1w_2 \in \partial R$, we have $yz_1w_1w_2 \in R$.

If $yz_2w_1w_2$ is red, then we have $yz_2w_1w_2 \in R$ as $yw_1w_2 \in \partial R$. Moreover, $\{z_1z_2z_3u, yz_1w_1w_2, yz_2w_1w_2\} \subseteq R$ has an empty intersection. Hence we may assume that $yz_2w_1w_2$ is blue. We have $yz_2w_1w_2 \in B_*$ since $yw_1w_2w_3 \in B_*$. It follows that $yz_2z_3w_1 \in H^{\text{red}}$ (else $B = B_*$).

Now if $yz_2 \in G^{\text{red}}$, then $yz_2w_1 \in \partial R$ by (SP3) and thus $yz_2z_3w_1 \in R$. We are done since $\{z_1z_2z_3u, yz_1w_1w_2, yz_2z_3w_1\} \subseteq R$ has an empty intersection. Hence we may assume that $yz_2 \in G^{\text{blue}}$. Since $yz_1 \in B^2$ and G is a blueprint, we have $yz_2 \in B^2$. By (SP3), we have $yz_2w_1 \in \partial B$. Since $B_* \neq B$, $yz_2w_1 \in \partial B$ and $yw_1w_2 \in \partial R$, we have $yz_2w_1w_2 \in R$. We are done since $\{z_1z_2z_3u, yz_1w_1w_2, yz_2w_1w_2\} \subseteq R$ has an empty intersection.

Case B.2: At least one of the edges z_1z_2, z_1z_3, z_2z_3 is in B^2. We assume without loss of generality that $z_1z_2 \in B^2$. We may assume that $yz_1, yz_2 \in G^{\text{red}}$ (else we are in Case B.1). By (SP4), we have $yz_1w_1, yz_2w_1 \in \partial R$. Let $F_1 = \{yz_1w_1w_2, yz_1z_3w_1\}$ and

$F_2 = \{yz_2w_1w_2, yz_2z_4w_1\}$. We claim that each of F_1 and F_2 contains a red edge. Suppose not, and assume without loss of generality that $F_1 \subseteq H^{\text{blue}}$. Since $yz_1z_2z_3 = f'(u) \in B$ and $yw_1w_2 \in \partial B_*$, we have $B_* = B$, a contradiction. Let $f_1 \in F_1$ and $f_2 \in F_2$ be red edges. Since $yz_1w_1, yz_2w_1 \in \partial R$, we have $f_1, f_2 \in R$. We are done since $\{f_u^*, f_1, f_2\} \subseteq R$ has an empty intersection.

Case B.3: $\boldsymbol{f}$ contains no edges of $\boldsymbol{B^2}$. Since f contains no edges of B^2, Claim 3.5.17 implies that $G^{\text{blue}}[f]$ does not contain a triangle. Thus we may choose edges $e_{12} \in G^{\text{red}}[yz_1z_2], e_{13} \in G^{\text{red}}[yz_1z_3]$ and $e_{23} \in G^{\text{red}}[yz_2z_3]$. Let $F_{12} = \{yz_1z_2w_1, e_{12} \cup w_1w_2\}$, $F_{13} = \{yz_1z_3w_1, e_{13} \cup w_1w_3\}$ and $F_{23} = \{yz_2z_3w_2, e_{23} \cup w_2w_3\}$. Suppose that each of F_{12}, F_{13} and F_{23} contains a red edge f_{12}, f_{13} and f_{23}, respectively. By (SP3), we have that $e_{12} \cup w_1, e_{13} \cup w_1, e_{23} \cup w_2 \in \partial R$ and thus $F = \{f_u^*, f_{12}, f_{13}, f_{23}\} \subseteq R$. We are done since F has an empty intersection. Hence we may assume that one of F_{12}, F_{13} and F_{23} contains only blue edges. We assume without loss of generality that F_{12} contains only blue edges. That is $yz_1z_2w_1$ and $e_{12} \cup w_1w_2$ are blue. Note that these two edges are in B since $yz_1z_2z_3 = f'(u) \in B$. By (SP4), we have $z_1w_1w_2, z_2w_1w_2 \in \partial B_*$. Hence $yz_1z_2w_1, e_{12} \cup w_1w_2 \in B_*$. This contradicts $B_* \neq B$.

This completes the proof. □

3.6 Proof of Theorem 1.2.1

Definition 3.6.1. Let $\mu_k^s(\beta, \varepsilon, n)$ be the largest μ such that every 2-edge-coloured $(1-\varepsilon, \varepsilon)$-dense k-graph on n vertices contains a factional matching with weight μ such that all edges with non-zero weight have weight at least β and lie in s monochromatic tight components. Let $\mu_k^s(\beta) = \liminf_{\varepsilon \to 0} \liminf_{n \to \infty} \mu_k^s(\beta, \varepsilon, n)/n$.

We will also reuse the crucial result Corollary 2.3.12 from Chapter 2 that reduces finding cycles in the original graph to finding tightly connected matchings in the reduced graph.

We are now ready to prove Theorem 1.2.1.

Proof of Theorem 1.2.1. Let $1/n \ll c \ll \eta \ll \varepsilon$. Let K be a 2-edge-coloured complete 4-graph on $N = (5+\varepsilon)n$ vertices. We show that K contains a monochromatic tight cycle of length $4n$. Note that Lemma 3.5.1 implies that $\mu_4^1(c) \geq 1/5 - \eta$. Applying Corollary 2.3.12 with $N, \eta, c, 4, K$ playing the roles of n, η, β, k, K we obtain that K contains a monochromatic tight cycle of length ℓ for any $\ell \leq (\mu_4^1(c) - \eta)4N$ divisible by 4. Since

$$(\mu_4^1(c) - \eta)4N \geq (1/5 - 2\eta)(5+\varepsilon)4n \geq 4n,$$

we have that K contains a monochromatic tight cycle of length $4n$. $\square$

3.7 Concluding remarks

Here we determined the Ramsey number for 4-uniform tight cycles asymptotically in the case where the length of the tight cycle is divisible by 4. The cases where the length of the tight cycle is not divisible by 4 are still open. The general conjecture for the Ramsey numbers of tight cycles is as follows.

Conjecture 3.7.1 (Haxell, Łuczak, Peng, Rödl, Ruciński, Skokan [67]). *Let $k \geq 2$, $0 \leq i \leq k-1$ and $d = \gcd(k,i)$. Then $r(C_{kn+i}^{(k)}) = (1+o(1))\frac{d+1}{d}kn$.*

The lower bound is given by the following extremal example.

Proposition 3.7.2. *Let $n \geq 1, k \geq 2$ and $0 \leq i \leq k-1$. Then $r(C_{kn+i}^{(k)}) \geq \frac{d+1}{d}kn - 2$ where $d = \gcd(k,i)$.*

Proof. Let $N = \frac{d+1}{d}kn - 2$ and consider the following red-blue edge-colouring of $K_N^{(k)}$. Partition the vertex set of $K_N^{(k)}$ into two sets X and Y such that $|X| = \frac{k}{d}n - 1$ and $|Y| = kn - 1$. Colour each edge that has an even number of vertices in X red and all other edges blue. Note that each monochromatic tight component in this red-blue edge-colouring of $K_N^{(k)}$ consists of all edges e such that $|e \cap X| = r_1$ and $|e \cap Y| = r_2$ for some pair of nonnegative integers (r_1, r_2) with $r_1 + r_2 = k$. We claim that this red-blue edge-colouring of $K_N^{(k)}$ does not contain a monochromatic copy of $C_{kn+i}^{(k)}$. Suppose for a contradiction that there is a monochromatic copy C of $C_{kn+i}^{(k)}$. Let (r_1, r_2) be the pair of nonnegative integers that correspond to the monochromatic tight component that contains C. Note that if $r_1 = 0$, then $V(C) \subseteq Y$. But since $|V(C)| = kn + i > kn - 1 = |Y|$, this is impossible. Hence $r_1 \geq 1$.

First suppose that $i = 0$. Then $d = k$ and so $|X| = n - 1$. By double counting the pairs (v, e) such that $v \in V(C) \cap X$ and $v \in e \in C$, we have $k\,|V(C) \cap X| = r_1 kn$. Since $n - 1 = |X| \geq |V(C) \cap X| = r_1 n \geq n$, we have a contradiction.

Now suppose that $1 \leq i \leq k - 1$. By the same double counting argument as above, we have $k\,|V(C) \cap X| = r_1(kn + i)$. Hence $k \mid r_1(kn + i)$ and thus $k \mid r_1 i$. Thus $r_1 i$ is a common multiple of i and k. It follows that $ik = \gcd(k, i)\mathrm{lcm}(k, i) \leq d r_1 i$ and so $r_1 \geq \frac{k}{d}$. Now we have $|X| \geq |X \cap V(C)| = \frac{r_1}{k}(kn + i) \geq r_1 n \geq \frac{k}{d} n > |X|$, a contradiction. $\square$

For 4-uniform tight cycles, Conjecture 3.7.1 implies that $r(C_{4n+1}^{(4)}) = (1 + o(1))8n = r(C_{4n+3}^{(4)})$ and $r(C_{4n+2}^{(4)}) = (1 + o(1))6n$. In order to prove these remaining cases, finding a large monochromatic tightly connected fractional matching in the reduced graph is no longer sufficient. Indeed, if the corresponding monochromatic tight component in the original graph is a complete 4-partite 4-graph, then it only contains tight cycles of length divisible by 4. A natural approach to overcome this problem is to additionally require that the chosen monochromatic tight component in the reduced graph contains a copy of $C_5^{(4)}$ (or a subgraph homomorphic to $C_5^{(4)}$). One of the difficulties with this approach is that we can no longer just choose a maximum matching in a monochromatic tight component, as these matchings can now be arbitrarily large (as long as we cannot also find a subgraph homomorphic to $C_5^{(4)}$). Thus a lot of the arguments we used in our proof do no longer apply in this setting. Nevertheless we hope that some of our methods will be useful for further research on this conjecture.

CHAPTER 4

RESILIENCE FOR TIGHT HAMILTONICITY

Due to space constraints for this thesis, we omit several more technical proofs of results in this chapter. For any proofs thus omitted we refer the reader to the full paper this chapter is based on [8]. Our aim in this chapter is to prove Theorem 1.3.1. We recall some definitions before stating the theorem again.

Recall that $G^{(k)}(n,p)$ is the *binomial random k-graph*, that is the k-graph on n vertices for which each k-set of vertices forms an edge independently with probability p. Recall that we say an event holds *asymptotically almost surely* (a.a.s. for short) if its probability tends to 1 as n tends to infinity. Finally, recall that for a k-graph G, we denote by $\delta_{k-1}(G)$ the *minimum codegree of* G, that is $\delta_{k-1}(G) = \min_{S \in \binom{V(G)}{k-1}} d_G(S)$, where, for $S \subseteq V(G)$, $d_G(S)$ is the number of edges of G that contain S.

We now recall the statement of Theorem 1.3.1.

Theorem 1.3.1. *Given any $\gamma > 0$ and $k \geq 3$, if $p \geq n^{-1+\gamma}$, we show that $\Gamma = G^{(k)}(n,p)$ a.a.s. satisfies the following. Let G be any n-vertex subgraph of Γ such that $\delta_{k-1}(G) \geq \left(\frac{1}{2} + 2\gamma\right)pn$. Then G contains a tight Hamilton cycle.*

The remainder of this chapter is organised as follows. We first give the ideas of the proof of Theorem 1.3.1. Then we gather the tools we need for the proof of Theorem 1.3.1 (including a special version of hypergraph regularity and some useful properties of the random hypergraph) in Section 4.2 and finally prove the theorem in Section 4.3.

4.1 Ideas of the proof

Our proof strategy for Theorem 1.3.1 uses the *reservoir method*, which was previously used in [4] and [7], in a similar way to the use we will make here, to give polynomial-time algorithms that find tight Hamilton cycles in Γ itself for broadly similar values of p. Very briefly, the reservoir method is as follows.

In a first step, we identify a *reservoir set* R, which contains a small (but bounded away from 0) fraction of the vertices of G. We construct a *reservoir path* P_{res}, which is a tight path that contains all the vertices of R and in addition for any subset R' of R, there is a tight path with the same ends as P_{res} whose vertex set is $V(P_{\mathrm{res}}) \setminus R'$.

In a second step, we extend P_{res} to an almost-spanning tight path P_{almost}. In the final step we re-use some vertices of R to extend P_{almost} further to a structure which is 'almost' a tight Hamilton cycle, except that some vertices R' of R are used twice. Finally we apply the reservoir property of P_{res} to obtain the desired tight Hamilton cycle.

In [7], in the random hypergraph, there are two main tools needed to put this plan into action. First, for any given ordered $(k-1)$-tuple $\mathbf{x}$ of vertices and set S of 'unused' vertices which is not too small, there will be lots of ways to start a tight path from $\mathbf{x}$ and continuing with vertices of S. Second, for any given pair of ordered $(k-1)$-tuples $\mathbf{x}$ and $\mathbf{y}$, and any given set S of unused vertices which is not too small, it is possible to find a tight path from $\mathbf{x}$ to $\mathbf{y}$ in S.[1]

Neither of these statements is true in the resilience setting. Instead, we make use of hypergraph regularity to help us. In Section 4.2 we state our main tools, and prove some of them. We first introduce spike paths, which we need to construct our reservoir structure (much as in [7]).

We give the notational setup for hypergraph regularity, and state a sparse, strengthened version of the Strong Hypergraph Regularity Lemma, Lemma 4.2.4, which may be of independent interest. We show that the output of this Regularity Lemma is, for k-graphs

[1] To be accurate, these statements will be true for all the sets S that actually appear in the proof, by a careful revealing argument; they are not true for every S.

with our minimum degree condition, a structure which is robustly *tightly linked*: this is a version of connectivity appropriate for tight paths.

We show that the random hypergraph has certain nice properties: in particular, once one removes a small fraction of $(k-1)$-tuples, for any remaining $(k-1)$-tuple $\mathbf{x}$ and set S which is reasonably small (it cannot contain more than $n/2$ vertices) there are lots of ways to start constructing a tight path from $\mathbf{x}$ *avoiding* S (Lemma 4.2.9), and if we do so for a sufficiently large (but independent of n) number of steps, we reach a positive fraction of all $(k-1)$-tuples. This statement (Lemma 4.2.10) is one of the key points in our proof: most of the time, we can expand in a few steps from any given $(k-1)$-tuple to a positive density of $(k-1)$-tuples (and a similar statement holds for spike paths).

Using Lemma 4.2.10, regularity and tight linkedness, we can prove a Connecting Lemma (Lemma 4.2.16) which states that for any reasonably small set S and most pairs $\mathbf{x}$ and $\mathbf{y}$ of $(k-1)$-tuples, there is a short tight path from $\mathbf{x}$ to $\mathbf{y}$ which avoids S.

These tools are enough to prove a Reservoir Lemma 4.2.21, which (much as in [7]) constructs P_{res} mentioned above. However again at this point difficulties arise. In the random hypergraph of [7], the vertices outside P_{res} have no particular structure. In our setting, P_{res} interacts in some rather unpredictable way with the existing structure provided by the Regularity Lemma. To deal with this, we use LP-duality in Lemma 4.2.19 to find a fractional matching which will tell us how many vertices we should use in each part of our regular partition in order to obtain P_{almost}. We also at this point run into the difficulty that we can only guarantee expansion from the minimum degree when we are avoiding less than $n/2$ vertices, yet P_{almost} is supposed to cover almost all of the vertices; it is here that we need the 'strengthened' property of our Regularity Lemma.

4.2 Tools

In this section we present the tools needed to prove Theorem 1.3.1.

4.2.1 Spike paths

To build our reservoir structure we need spike paths, which are the following variant of a tight path that changes orientation every $(k-1)$ steps. We will only consider spike paths with a number of vertices divisible by $k-1$.

Definition 4.2.1 (Spike path)**.** *In an k-uniform hypergraph, a spike path with t spikes consists of a sequence of t pairwise disjoint $(k-1)$-tuples $\mathbf{a}_1, \dots, \mathbf{a}_t$, where $\mathbf{a}_i = (a_{i,1}, \dots, a_{i,k-1})$ for all i, with the property, that the edges $\{a_{i,k-j}, \dots, a_{i,1}, a_{i+1,1}, \dots, a_{i+1,j}\}$ are present for all $i = 1, \dots, t-1$ and $j = 1, \dots, k-1$. We call $\mathbf{a}_i$ the ith spike.*

4.2.2 Notation for k-multicomplexes

In this section we explain some notation for k-multicomplexes most of which is needed for our version of Hypergraph Regularity Lemma (Lemma 4.2.4).

A *k-complex* is a hypergraph H all of whose edges have size at most k, which is down-closed, i.e. if $e \in E(H)$ and $e' \subseteq e$ then $e' \in E(H)$. The *layers* of a k-complex are, for each $0 \leq i \leq k$, the i-uniform hypergraph $H^{(i)}$ on the same vertex set, where $E\left(H^{(i)}\right) = \{e \in E(H) : |e| = i\}$.

A *k-multicomplex* is, informally, a k-complex in which multiple edges of any size between 2 and k are permitted, together with a map *boundary* ∂ identifying the $(i-1)$-edges which support a given i-edge. Formally, a k-multicomplex H consists of a *vertex set* $V(H)$, together with a set of *edges* $E(H)$, a *vertices* map $\mathrm{vertices} : E \to \mathcal{P}(V)$ such that $\mathrm{vertices}(e)$ is a set of size between 0 and k for each $e \in E(H)$, and a *boundary* map $\partial : E \setminus \{\emptyset\} \to \mathcal{P}(E)$ such that ∂e contains exactly one edge whose vertices are $\mathrm{vertices}(e) \setminus \{v\}$ for each $v \in \mathrm{vertices}(e)$, and no other edges. We further insist on the following consistency condition: if $2 \leq i \leq k$, and S is a set of i edges each with $i-1$ vertices, such that $\left|\bigcup_{f \in S} \partial f\right| > \binom{i}{i-2}$, then there are no edges $e \in H$ such that $\partial e = S$. We say that the *uniformity* of an edge e is $|\mathrm{vertices}(e)|$, and we may write that e is an edge *on the set* $\mathrm{vertices}(e)$, or that e is *a* $|\mathrm{vertices}(e)|$*-edge*. We will also say, given a set S

consisting of i edges of uniformity $(i-1)$, that e is *supported* on S if $\partial e = S$.

Note that the boundary of a 1-edge is necessarily $\{\emptyset\}$, and that 'down-closure' is forced by the condition of the boundary map. To better understand the consistency condition, consider the following. If e is an edge of H with at least two vertices, and x and y are distinct vertices of e, let e_x and e_y be the edges in ∂e whose vertices do not contain respectively x and y. There is an edge e_{xy} in ∂e_x, and an edge e_{yx} in ∂e_y, on vertices$(e) \setminus \{x, y\}$. The consistency condition is equivalent to insisting that for any e, x and y we have $e_{yx} = e_{xy}$.

Observe that a k-complex is a k-multicomplex, where the vertices of each edge are simply its members as a set, and the boundary map is the usual boundary $\partial e = \Big\{e \setminus \{v\} : v \in e\Big\}$ (which is in this case the only possible boundary map for the given vertices map). However in general, for a given ground set, edge set and vertices map, there may be several different boundary maps which fit the definition of k-multicomplex; these return different multicomplexes. The idea here is that we will need to think of a given edge (say with vertices $\{1, 2, 3\}$) as containing specific edges with vertices $\{1, 2\}$, $\{1, 3\}$ and $\{2, 3\}$, and the map ∂ tells us which edges these are. We should stress that it is possible to have a k-multicomplex in which there are two different edges which have the same boundary and vertices, and indeed the multicomplexes we consider in this paper will have this property for edges of uniformity two and above (though for us a 1-edge will always be the unique 1-edge on a given vertex).

Given a vector $\mathbf{d} = (d_2, \ldots, d_k)$ where $1/d_i \in \mathbb{N}$ for each i, we call a k-multicomplex H $\mathbf{d}$-*equitable* if there is exactly one 1-edge on each vertex, and furthermore for any $2 \le i \le k$ and i-set X of vertices the following holds. Whenever S is a collection of i edges of uniformity $i-1$ in H, one on the vertices $X \setminus \{x\}$ for each $x \in X$, if the union $\bigcup_{f \in S} \partial f$ has exactly $\binom{i}{i-2}$ edges then the number of i-edges in H supported on S is exactly $1/d_i$. We refer to $\mathbf{d}$ as the *density vector* of the multicomplex.

Finally, we need a notion of connectedness for multicomplexes.

Definition 4.2.2 (tight link, tightly linked)**.** *Given a k-multicomplex $\mathcal{R}$, and two* $(k-1)$-

edges u, v of $\mathcal{R}$, let $\mathbf{u}$ be u together with an ordering $(u_1, \dots, u_{k-1})$ of its vertices, and similarly let $\mathbf{v}$ be v together with an ordering $(v_1, \dots, v_{k-1})$ of its vertices. A tight link *from* $\mathbf{u}$ *to* $\mathbf{v}$ *in* $\mathcal{R}$ *is the following collection of (not necessarily distinct) vertices and edges of* $\mathcal{R}$.

For each $1 \le j \le k-1$, there is a vertex w_j. There are k-edges $e_{1,u}$ and $e_{1,v}$ of $\mathcal{R}$, where $e_{1,u}$ is on vertices $\{u_1, \dots, u_{k-1}, w_1\}$ and $u \in \partial e_{1,u}$, and $e_{1,v}$ is on vertices $\{v_1, \dots, v_{k-1}, w_1\}$ and $v \in \partial e_{1,v}$. For each $2 \le j \le k-1$, there are k-edges $e_{j,u}$ and $e_{j,v}$ of $\mathcal{R}$, where $e_{j,u}$ is on vertices $\{u_j, \dots, u_{k-1}, w_1, \dots, w_j\}$ and $\partial e_{j-1,u} \cap \partial e_{j,u} \neq \emptyset$, and $e_{j,v}$ is on vertices $\{v_j, \dots, v_{k-1}, w_1, \dots, w_j\}$ and $\partial e_{j-1,v} \cap \partial e_{j,v} \neq \emptyset$. Finally $\partial e_{k-1,u} \cap \partial e_{k-1,v} \neq \emptyset$.

We say that a k-multicomplex $\mathcal{R}$ is tightly linked *if for any two $(k-1)$-edges in $\mathcal{R}$, and any orderings of their vertices, $\mathbf{u}$ and $\mathbf{v}$, there is a tight link from $\mathbf{u}$ to $\mathbf{v}$ in $\mathcal{R}$.*

The precise sequence of vertices and edges is not critical (it is simply a particular structure we can easily construct). However it will be convenient to note that the k-edges of a tight link are in fact a spike path with three spikes. Note that there is $\ell \in \mathbb{N}$ and a permutation ϱ on $[k-1]$ such that for any $\mathbf{u}$ and $\mathbf{v}$, if there is a tight link from $\mathbf{u}$ to $\mathbf{v}$ then there is a homomorphism from the ℓ-vertex tight path to $\mathcal{R}$, using only the k-edges of the tight link, where the first $k-1$ vertices of the tight path are sent to $\mathbf{u}$ in order and the last $k-1$ vertices to the vertices of $\mathbf{v}$ in the order ϱ.

4.2.3 Sparse hypergraph regularity

We need a strengthened version of the Strong Hypergraph Regularity Lemma for sparse hypergraphs. The Strong Hypergraph Regularity Lemma was first proved by Rödl and Skokan [107] and Gowers [55]; we use a version due to Rödl and Schacht [106], from which we deduce a strengthened version by a standard method. We then use a weak sparse regularity lemma of Conlon, Fox and Zhao [33] to transfer this strengthened version to a sparse version, following [6].

In order to state our regularity lemma, we need quite a few definitions. These are either standard definitions for the dense ($p = 1$) case, or the natural sparse versions of the same, as taken from [3]. Note that definitions here differ from the definitions for hypergraph regularity in Chapter 2 as we need a different version of hypergraph regularity in this chapter.

Let $\mathcal{P}$ partition a vertex set V into parts $V_1, \ldots, V_s$. We say that a subset $S \subseteq V$ is *$\mathcal{P}$-partite* if $|S \cap V_i| \leq 1$ for every $i \in [s]$ and the *index* of a $\mathcal{P}$-partite set $S \subseteq V$ is $i(S) := \{i \in [s] : |S \cap V_i| = 1\}$. For any $A \subseteq [s]$ we write V_A for $\bigcup_{i \in A} V_i$. Similarly, we say that a hypergraph H is *$\mathcal{P}$-partite* if all of its edges are $\mathcal{P}$-partite. In this case we refer to the parts of $\mathcal{P}$ as the *vertex classes* of H. Moreover, we say that a hypergraph H is *s-partite* if there is some partition $\mathcal{P}$ of $V(H)$ into s parts for which H is $\mathcal{P}$-partite.

Let $i \geq 2$, let H_i be any i-partite i-graph, and let H_{i-1} be any i-partite $(i-1)$-graph, on a common vertex set V partitioned into i common vertex classes. We denote by $K_i(H_{i-1})$ the i-partite i-graph on V whose edges are all i-sets in V which are supported on H_{i-1} (i.e. induce a copy of the complete $(i-1)$-graph K_i^{i-1} on i vertices in H_{i-1}). Given $p \in (0, 1]$, the *p-density of H_i with respect to H_{i-1}* is then defined to be

$$d_p(H_i|H_{i-1}) := \frac{|K_i(H_{i-1}) \cap H_i|}{p|K_i(H_{i-1})|}$$

if $|K_i(H_{i-1})| > 0$. For convenience we take $d_p(H_i|H_{i-1}) := 0$ if $|K_i(H_{i-1})| = 0$, and we assume H_1 is the complete 1-graph on V, whose edge set is V. So $d_p(H_i|H_{i-1})$ is the proportion of copies of K_i^{i-1} in H_{i-1} which are also edges of H_i, scaled by p. When H_{i-1} is clear from the context, we simply refer to $d_p(H_i|H_{i-1})$ as the *relative p-density of H_i*. We say that H_i is *(d_i, ε, p)-regular with respect to H_{i-1}* if we have $d_p(H_i|H'_{i-1}) = d_i \pm \varepsilon$ for every subgraph H'_{i-1} of H_{i-1} such that $|K_i(H'_{i-1})| > \varepsilon |K_i(H_{i-1})|$. Given an i-graph G whose vertex set contains that of H_{i-1}, we say that G is *(d_i, ε, p)-regular with respect to H_{i-1}* if the i-partite subgraph of G induced by the vertex classes of H_{i-1} is (d_i, ε, p)-regular with respect to H_{i-1}. Finally, we say G is *(ε, p)-regular with respect to H_{i-1}* if there exists d_i such that G is (d_i, ε, p)-regular with respect to H_{i-1}. Similarly as before, when H_{i-1} is clear from the context, we refer to the relative density of this i-partite subgraph of G with respect to H_{i-1} as the *relative p-density of G*.

Now let H be an s-partite k-complex on vertex classes $V_1, \dots, V_s$, where $s \geq k \geq 3$. Recall that, since H is a complex, if $e \in H$ and $e' \subseteq e$ then $e' \in H$. So if $e \in H^{(i)}$ for some $2 \leq i \leq k$, then the vertices of e induce a copy of K_i^{i-1} in $H^{(i-1)}$. We say that H is $(d_k, \dots, d_2, \varepsilon_k, \varepsilon, p)$*-regular* if

(a) for any $2 \leq i \leq k-1$ and any $A \in \binom{[s]}{i}$, the induced subgraph $H^{(i)}[V_A]$ is $(d_i, \varepsilon, 1)$-regular with respect to $H^{(i-1)}[V_A]$, and

(b) for any $A \in \binom{[s]}{k}$, the induced subgraph $H^{(k)}[V_A]$ is (d_k, ε_k, p)-regular with respect to $H^{(k-1)}[V_A]$.

So each constant d_i approximates the relative density of each subgraph $H^{(i)}[V_A]$ for $A \in \binom{[s]}{i}$. For a $(k-1)$-tuple $\mathbf{d} = (d_k, \dots, d_2)$ we write $(\mathbf{d}, \varepsilon_k, \varepsilon, p)$-regular to mean $(d_k, \dots, d_2, \varepsilon_k, \varepsilon, p)$-regular.

The definition of a $(\mathbf{d}, \varepsilon_k, \varepsilon, p)$-regular complex H is the 'right' generalisation of an ε-regular pair (X, Y) in dense graphs to sparse hypergraphs. The Szemerédi Regularity Lemma states that there is a partition of the vertices of any graph into boundedly many parts such that most pairs of parts are regular; now our aim is to define a generalisation of 'partition' in order to say that we can partition any k-uniform hypergraph G such that most k-sets lie in regular complexes. As one can guess from the phrasing, the k-layer of each complex will consist of (all) edges of G supported by the complex. The lower layers will be in the 'partition', and we now set up the notation to define this.

Fix $k \geq 3$, and let $\mathcal{P}$ partition a vertex set V into parts $V_1, \dots, V_t$. For any $A \subseteq [t]$, we denote by $\mathrm{Cross}_A(\mathcal{P})$ the collection of $\mathcal{P}$-partite subsets $S \subseteq V$ of index $i(S) = A$. Likewise, we denote by $\mathrm{Cross}_j(\mathcal{P})$ the union of Cross_A for each $A \in \binom{[t]}{j}$, so $\mathrm{Cross}_j(\mathcal{P})$ contains all $\mathcal{P}$-partite subsets $S \subseteq V$ of size j. When $\mathcal{P}$ is clear from the context, we write simply Cross_A and Cross_j. For each $2 \leq j \leq k-1$ and $A \in \binom{[t]}{j}$ let $\mathcal{P}_A$ be a partition of Cross_A. For consistency of notation we also define the trivial partitions $\mathcal{P}_{\{s\}} := \{V_s\}$

for $s \in [t]$ and $\mathcal{P}_\emptyset := \{\emptyset\}$. Let $\mathcal{P}^*$ consist of the partitions $\mathcal{P}_A$ for each $A \in \binom{[t]}{j}$ and each $0 \le j \le k-1$. We say that $\mathcal{P}^*$ is a $(k-1)$-*family of partitions on* V if whenever $S, T \in \mathrm{Cross}_A$ lie in the same part of $\mathcal{P}_A$ and $B \subseteq A$, then $S \cap \bigcup_{j \in B} V_j$ and $T \cap \bigcup_{j \in B} V_j$ lie in the same part of $\mathcal{P}_B$. In other words, given $A \in \binom{[t]}{j}$, if we specify one part of each $\mathcal{P}_B$ with $B \in \binom{A}{j-1}$, then we obtain a subset of Cross_A consisting of all $S \in \mathrm{Cross}_A$ whose $(j-1)$-subsets are in the specified parts. We say that this subset of Cross_A is the subset *supported* by the specified parts of $\mathcal{P}_B$. In general, we say that a j-set e is *supported* by a collection S, with $|S| = j$, of $(j-1)$-graphs if exactly one $(j-1)$-subset of e is in each member of S, and we say a set of j-edges E is supported by S if each edge of E is supported by S.

Thus the partitions $\mathcal{P}_B$ give a natural partition of Cross_A, and we are saying that $\mathcal{P}_A$ must refine it.

We refer to the parts of each member of $\mathcal{P}^*$ as *cells*. Also, we refer to $\mathcal{P}$ as the *ground partition* of $\mathcal{P}^*$, and the parts of $\mathcal{P}$ (i.e. the vertex classes V_i) as the *clusters* of $\mathcal{P}^*$. For each $0 \le j \le k-1$ let $\mathcal{P}^{(j)}$ denote the partition of Cross_j formed by the parts (which we call j-*cells*) of each of the partitions $\mathcal{P}_A$ with $A \in \binom{[t]}{j}$ (so in particular $\mathcal{P}^{(1)} = \mathcal{P}$).

Observe that a $(k-1)$-family of partitions $\mathcal{P}^*$ naturally form the edges of a k-multicomplex, whose vertex set is the (set of parts of the) ground partition, whose edges of uniformity $j \le k-1$ are the j-cells, with the vertices map identifying the j parts of the ground partition which contain a given j-cell, and where the boundary operator ∂e identifies the $(|e|-1)$-cells supporting e. So far we have described a $(k-1)$-multicomplex; we extend this to a k-complex by adding, for each set S of k edges of uniformity $k-1$ which can be a boundary (i.e. which is such that $\left|\bigcup_{f \in S} \partial f\right| = \binom{k}{k-2}$) one edge of uniformity k whose boundary is S. When we refer to the *multicomplex of the family of partitions* $\mathcal{P}^*$ we mean this multicomplex. Note that we have defined the word 'support' both in terms of multicomplexes and in terms of a family of partitions: but these definitions are consistent, i.e. that a given j-cell is supported by some $(j-1)$-cells means the same thing whether one reads 'support' in terms of the family of partitions or its multicomplex.

For any $0 \le j \le k-1$, any $A \in \binom{[t]}{j}$ and any $Q' \in \mathrm{Cross}_A$, let $C_{Q'}$ denote the cell of $\mathcal{P}_A$ which contains Q'. Then the fact that $\mathcal{P}^*$ is a family of partitions implies that for any $Q \in \mathrm{Cross}_k$ the union $\mathcal{J}(Q) \coloneqq \bigcup_{Q' \subsetneq Q} C_{Q'}$ of cells containing subsets of Q is a k-partite $(k-1)$-complex. We say that the $(k-1)$-family of partitions $\mathcal{P}^*$ is (t_0, t_1, ε)-*equitable* if

(a) $\mathcal{P}$ partitions V into t clusters of equal size, where $t_0 \le t \le t_1$,

(b) for each $2 \le j \le k-1$, $\mathcal{P}^{(j)}$ partitions Cross_j into at most t_1 cells,

(c) there exists $\mathbf{d} = (d_{k-1}, \dots, d_2)$ such that for each $2 \le j \le k-1$ we have $d_j \ge 1/t_1$ and $1/d_j \in \mathbb{N}$, and for every $Q \in \mathrm{Cross}_k$ the k-partite $(k-1)$-complex $\mathcal{J}(Q)$ is $(\mathbf{d}, \varepsilon, \varepsilon, 1)$-regular.

Note that conditions 4.2.3 and 4.2.3 imply that $\mathcal{J}(Q)$ is a $(1, t_1, \varepsilon)$-equitable $(k-1)$-complex (with the same density vector $\mathbf{d}$) for any $Q \in \mathrm{Cross}_k$.

Next, for any $\mathcal{P}$-partite set Q with $2 \le |Q| \le k$, define $\hat{P}(Q; \mathcal{P}^*)$ to be the $|Q|$-partite $(|Q|-1)$-graph on $V_{i(Q)}$ with edge set $\bigcup_{Q' \in \binom{Q}{|Q|-1}} C_{Q'}$. We refer to $\hat{P}(Q; \mathcal{P}^*)$ as a $|Q|$-*polyad*; when the family of partitions $\mathcal{P}^*$ is clear from the context, we write simply $\hat{P}(Q)$ rather than $\hat{P}(Q; \mathcal{P}^*)$. Note that the condition for $\mathcal{P}^*$ to be a $(k-1)$-family of partitions can then be rephrased as saying that if $2 \le |Q| \le k-1$ then the cell C_Q is supported on $\hat{P}(Q)$, and in the multicomplex corresponding to $\mathcal{P}^*$ we have edges corresponding to the cells of each uniformity from 1 to $k-1$ inclusive, together with edges corresponding to the k-polyads supported by $\mathcal{P}^*$. As shown in [3, Claim 32], if $\mathcal{P}^*$ is (t_0, t_1, ε)-equitable for sufficiently small ε, then for any $2 \le j \le k-1$ and any $Q \in \mathrm{Cross}_j$ the number of j-cells of $\mathcal{P}^*$ supported on $\hat{P}(Q)$ is precisely equal to $1/d_j$. More specifically, if $\left(d_j^{-1} - 1\right)(d_j + \varepsilon) < 1$, and $\left(d_j^{-1} + 1\right)(d_j - \varepsilon) > 1$, then by definition necessarily there are exactly d_j^{-1} cells supported; it suffices to choose $\varepsilon \ll d_j^2$ to ensure these two inequalities. In other words, the multicomplex corresponding to $\mathcal{P}^*$ is $\mathbf{d}$-equitable.

Now let G be a k-graph on V, and let $\mathcal{P}^*$ be a $(k-1)$-family of partitions on V. Let $Q \in \mathrm{Cross}_k$, so the polyad $\hat{P}(Q)$ is a k-partite $(k-1)$-graph. We say that G is (ε_k, p)-*regular with respect to* $\mathcal{P}^*$ if there are at most $\varepsilon_k \binom{|V|}{k}$ sets $Q \in \mathrm{Cross}_k$ for which G

is not (ε_k, p)-regular with respect to the polyad $\hat{P}(Q)$. That is, at most an ε_k-proportion of subsets of V of size k yield polyads with respect to which G is not regular (though some subsets of V of size k do not yield any polyad due to not being members of Cross_k).

At this point we have the setup to state the Strong Hypergraph Regularity Lemma, which states that for any k-uniform hypergraph G there is a $(k-1)$-family of partitions $\mathcal{P}^*$, which is (t_0, t_1, ε)-equitable for some t_1 independent of $|V(G)|$, such that G is regular with respect to $\mathcal{P}^*$. However we need a stronger version, which is not standard (the dense graph version, called the Strengthened Regularity Lemma, is due to Alon, Fischer, Krivelevich and Szegedy [10], and it is folklore that the hypergraph version we now state should exist). To that end, given two families of partitions $\mathcal{P}^*$ and $\mathcal{Q}^*$ on the same vertex set, we say that $\mathcal{P}^*$ *refines* $\mathcal{Q}^*$ if every cell of $\mathcal{P}^*$ is a subset of some cell of $\mathcal{Q}^*$.

Definition 4.2.3. *Given a k-uniform hypergraph G, we call a pair of families of partitions $(\mathcal{P}_c^*, \mathcal{P}_f^*)$ on $V(G)$ a $(t_0, t_1, t_2, \varepsilon_k, \varepsilon, f_k, f, p)$-strengthened pair for G if the following are true.*

(S1) *$\mathcal{P}_f^*$ refines $\mathcal{P}_c^*$.*

(S2) *$\mathcal{P}_c^*$ is (t_0, t_1, ε)-equitable.*

(S3) *G is (ε_k, p)-regular with respect to $\mathcal{P}_c^*$.*

(S4) *$\mathcal{P}_f^*$ is (t_0, t_2, f)-equitable.*

(S5) *G is (f_k, p)-regular with respect to $\mathcal{P}_f^*$.*

(S6) *For all but at most $\varepsilon_k^2\binom{|V(G)|}{k}$ elements Q of $\mathrm{Cross}_k(\mathcal{P}_c)$, we have $d_p\Big(G\Big|\hat{\mathcal{P}}(Q, \mathcal{P}_c^*)\Big) = d_p\Big(G\Big|\hat{\mathcal{P}}(Q, \mathcal{P}_f^*)\Big) \pm \varepsilon_k$.*

We refer to $\mathcal{P}_c^*$ as the *coarse partition* and $\mathcal{P}_f^*$ as the *fine partition.* Slightly extending the usual definition, we say a k-polyad $\hat{P}(Q; \mathcal{P}_c^*)$ is *irregular* (with respect to G) if any one of the following three things occurs:

(i) G is not (ε_k, p)-regular with respect to $\hat{P}(Q; \mathcal{P}_c^*)$,

(ii) for more than an ε_k-fraction of the k-sets Q' supported on $\hat{P}(Q;\mathcal{P}^*_c)$, G is not $\big(f_k, p\big)$-regular with respect to $\hat{P}(Q';\mathcal{P}^*_f)$, or

(iii) for more than an ε_k-fraction of the k-sets Q' supported on $\hat{P}(Q;\mathcal{P}^*_c)$, we have $d_p\Big(G\Big|\hat{P}(Q';\mathcal{P}^*_f)\Big) \neq d_p\Big(G\Big|\hat{P}(Q;\mathcal{P}^*_c)\Big) \pm \varepsilon_k$.

If a polyad of $\mathcal{P}^*_c$ is not irregular, we say it is *regular.*

We will always choose f_k such that $f_k \leq \varepsilon_k^2$, and ε small enough that every k-polyad supports very close to the same number of k-edges. Under this assumption, it is straightforward to check that at most a $4\varepsilon_k$-fraction of polyads in $\mathcal{P}^*_c$ are irregular (the proof of this can be found in the Appendix of [8]).

We need one more definition. Given any (not necessarily distinct) subsets $E_1, \dots, E_k$ in $\binom{[n]}{k-1}$, we say a k-set $S \subseteq [n]$ is *rainbow* for the E_i if there is an injective labelling of the $(k-1)$-subsets of S with the numbers $1, \dots, k$ such that the $(k-1)$-subset labelled i is in E_i. We write $K_k(E_1, \dots, E_k)$ for the set of rainbow k-sets in $[n]$. We say that a graph G on $[n]$ is (η, p)-*upper regular* if the following holds. For any $E_1, \dots, E_k$, we have

$$\Big|E(G) \cap K_k(E_1, \dots, E_k)\Big| \leq p\Big|K_k(E_1, \dots, E_k)\Big| + p\eta n^k \,.$$

Finally, we are in a position to state our strengthened sparse version of the Strong Hypergraph Regularity Lemma. Informally, what this states is that we can find $\mathcal{P}^*_c$ and $\mathcal{P}^*_f$ which are simultaneously a strengthened pair for s edge-disjoint graphs, for any (fixed) regularity ε_k of $\mathcal{P}^*_c$, where ε and f can be as small as desired depending on the number of parts in $\mathcal{P}^*_c$ and $\mathcal{P}^*_f$ respectively, and furthermore the regularity f_k of $\mathcal{P}^*_f$ can depend arbitrarily on the number of parts of $\mathcal{P}^*_c$.

Lemma 4.2.4 (Strengthened Sparse Strong Hypergraph Regularity Lemma [8, Lemma 5])**.** *Given integers $k \geq 2$ and t_0 and s, real $\varepsilon_k > 0$ and functions $\varepsilon, f_k, f : \mathbb{N} \to (0,1]$, there exists a real $\eta > 0$ and integers T and n_0 such that the following holds for all $n \geq n_0$ with $T!|n$. Let V be a vertex set of size n, suppose that $G_1, \dots, G_s$ are k-uniform hypergraphs on V, and suppose $\mathcal{Q}^*$ is a family of partitions on V which is $(1, t_0, \eta)$-equitable.*

Suppose furthermore that for each $1 \le i \le s$ there is a real $p_i \in (0,1]$ such that G_i is (η, p_i)-upper regular. Then there are integers t_1, t_2 with $t_0 \le t_1 \le t_2 \le T$, and families of partitions $\mathcal{P}_c^*$ and $\mathcal{P}_f^*$, both refining $\mathcal{Q}^*$, such that for each $1 \le i \le s$, the pair $(\mathcal{P}_c^*, \mathcal{P}_f^*)$ is a $\Big(t_0, t_1, t_2, \varepsilon_k, \varepsilon(t_1), f_k(t_1), f(t_2), p_i\Big)$-strengthened pair for G_i.

In this thesis, we omit the proof of this lemma (for the proof see [8, Lemma 5]). Note that the case $k = 2$ will not be used here; and in this setting the 'families of partitions' are simply vertex set partitions and the functions ε and f play no rôle.

Given a $(t_0, t_1, t_2, \varepsilon_k, \varepsilon, f_k, f, p)$-strengthened pair $(\mathcal{P}_c^*, \mathcal{P}_f^*)$ for G, recall that $\mathcal{P}_c^*$ has the structure of a multicomplex. We denote by $\mathcal{R}_{\varepsilon_k}(G; \mathcal{P}_c^*, \mathcal{P}_f^*)$ the *ε_k-reduced multicomplex of G with respect to $(\mathcal{P}_c^*, \mathcal{P}_f^*)$*, which is the (unique) maximal submulticomplex of $\mathcal{P}_c^*$ which has the following properties.

(RG1) Every k-edge of $\mathcal{R}_{\varepsilon_k}(G; \mathcal{P}_c^*, \mathcal{P}_f^*)$ is regular.

(RG2) For each $1 \le i \le k-1$, each i-edge of $\mathcal{R}_{\varepsilon_k}(G; \mathcal{P}_c^*, \mathcal{P}_f^*)$ is in the boundary of at least

$$\Big(1 - 2^{i+2}\varepsilon_k^{1/k}\Big) t \prod_{j=2}^{i+1} d_j^{-\binom{i}{j-1}} \quad \text{if } i < k-1, \text{ and}$$
$$\Big(1 - 2^{k+1}\varepsilon_k^{1/k}\Big) t \prod_{j=2}^{k-1} d_j^{-\binom{k-1}{j-1}} \quad \text{if } i = k-1$$

$(i+1)$-edges of $R_{\varepsilon_k}(G; \mathcal{P}_c^*, \mathcal{P}_f^*)$.

The existence and uniqueness of the reduced multicomplex are trivial: we obtain it by simply iteratively removing from the multicomplex $\mathcal{P}_c^*$ edges which either fail one of (RG1) or (RG2), or from whose boundary we removed edges (so that they are no longer supported and cannot be in the multicomplex). It is easy, but not quite trivial, to show that most of the vertices of $\mathcal{R}_{\varepsilon_k}(G; \mathcal{P}_c^*, \mathcal{P}_f^*)$ (i.e. the parts of $\mathcal{P}_c$) are also 1-edges of $\mathcal{R}_{\varepsilon_k}(G; \mathcal{P}_c^*, \mathcal{P}_f^*)$. Now given $d > 0$, we let $\mathcal{R}_{\varepsilon_k, d}(G; \mathcal{P}_c^*, \mathcal{P}_f^*)$ be the (unique) submulticomplex of $\mathcal{R}_{\varepsilon_k}(G; \mathcal{P}_c^*, \mathcal{P}_f^*)$ obtained by removing all k-edges corresponding to polyads whose relative p-density is less than d. We call $\mathcal{R}_{\varepsilon_k, d}(G; \mathcal{P}_c^*, \mathcal{P}_f^*)$ the *(ε_k, d)-reduced multicomplex of G with respect to $(\mathcal{P}_c^*, \mathcal{P}_f^*)$*.

The following lemma shows some useful properties about the (ε_k, d)-reduced multicomplex of G in our usual setting. In this thesis, we omit its proof (for the proof see [8, Lemma 6]).

Lemma 4.2.5 ([8, Lemma 6]). *Given $k \in \mathbb{N}$ and $d > 0$ suppose that $t_0 \in \mathbb{N}$ is sufficiently large. Given any constants $\delta, \varepsilon_k, \nu > 0$, any function $\varepsilon : \mathbb{N} \to (0, 1]$ which tends to zero sufficiently fast, any $t_1, t_2 \in \mathbb{N}$, any $0 < f_k \leq \varepsilon_k^2$ and any $f > 0$, there exists $\eta > 0$ such that the following holds for any sufficiently large n and any $p > 0$. Suppose G is an n-vertex hypergraph which is (η, p)-upper regular and every $(k-1)$-set in $V(G)$ is contained in at least δpn edges. Suppose that $(\mathcal{P}_c^*, \mathcal{P}_f^*)$ is a $(t_0, t_1, t_2, \varepsilon_k, \varepsilon(t_1), f_k, f, p)$-strengthened pair for G.*

Let $\mathcal{R} = \mathcal{R}_{\varepsilon_k,d}(G; \mathcal{P}_c^*, \mathcal{P}_f^*)$ be the (ε_k, d)-reduced multicomplex of G, and suppose that $\mathcal{P}_c^*$ has t clusters and density vector $\mathbf{d} = (d_{k-1}, \ldots, d_2)$. Then $\mathcal{R}$ contains at least $\left(1 - 4\varepsilon_k^{1/k}\right)t$ 1-edges, and every $(k-1)$-edge of $\mathcal{R}$ is contained in at least

$$\left(\delta - 2d - 2^{k+2}\varepsilon_k^{1/k}\right)t \cdot \prod_{i=2}^{k-1} d_i^{-\binom{k-1}{i-1}}$$

k-edges of $\mathcal{R}$.

Finally, if $\delta > \frac{1}{2} + 2d + 2^{k+2}\varepsilon_k^{1/k} + \nu$, then any induced subcomplex of $\mathcal{R}$ on at least $(1-\nu)t$ 1-edges is tightly linked.

If we do not remove too many vertices from the 1-cells we still have a regular complex with slightly different parameters.

Lemma 4.2.6 (Regular Restriction Lemma [3, Lemma 28]). *For all integers $k \geq 2$ and constants $\alpha, d_0 > 0$, there exists $\varepsilon > 0$ such that the following holds. Let $\mathbf{d} = (d_{k-1}, \ldots, d_2)$ be a vector of real numbers with $d_i \geq d_0$ for each $2 \leq i \leq k-1$, and let G be a k-partite $(k-1)$-complex with parts $V_1, \ldots, V_k$ of size $m \geq \varepsilon^{-1}$ which is $(\mathbf{d}, \varepsilon, \varepsilon, 1)$-regular. Choose any $V_i' \subseteq V_i$ of size at least αm for $i = 1, \ldots, k$. Then the induced subcomplex $G[V_1', \ldots, V_k']$ is $(\mathbf{d}, \sqrt{\varepsilon}, 1)$-regular.*

4.2.4 Properties of the random hypergraph

We use the following standard versions of the Chernoff bound.

Theorem 4.2.7. *Let X be a random variable with distribution* $\mathrm{Bin}(n,p)$*. Then for any $\varepsilon > 0$ we have*

$$\mathbb{P}(X \geq pn + \varepsilon n) \leq \exp(-D(p+\varepsilon||p)n) \quad \textit{and} \quad \mathbb{P}(X \leq pn - \varepsilon n) \leq \exp(-D(p-\varepsilon||p)n)\,,$$

where $D(x||y) = x\log\left(\frac{x}{y}\right) + (1-x)\log\left(\frac{1-x}{1-y}\right)$ is the Kullback-Leibler divergence between two Bernoulli-distributed random variables with parameters x and y, respectively. From this it follows

$$\mathbb{P}(|X - pn| > \varepsilon pn) < 2\exp\left(-\tfrac{\varepsilon^2 pn}{3}\right) \quad \textit{for any } \varepsilon \leq \tfrac{3}{2}$$

and if $t \geq 6pn$ we have

$$Pr(X \geq pn + t) < \exp(-t)\,.$$

Lemma 4.2.8. *Given $\eta > 0$, $k \in \mathbb{N}$ there exists C such that if $p \geq \frac{C}{n}$, then $\Gamma = G^{(k)}(n,p)$, and all its subgraphs, are a.a.s. (η,p)-upper regular.*

Proof. Observe that if $\Gamma = G^{(k)}(n,p)$ is (η,p)-upper-regular, then automatically all its subgraphs are also. We assume without loss of generality that $\eta < 1$, and set $C = 18k\eta^{-3}$.

Given any $E_1,\dots,E_k \subseteq \binom{[n]}{k}$, we aim to estimate the probability that $E_1,\dots,E_k$ witness the failure of $G^{(k)}(n,p)$ to be (η,p)-upper regular. The expected number of edges of $G^{(k)}(n,p)$ which appear on the sets $K_k(E_1,\dots,E_k)$ is $p\Big|K_k(E_1,\dots,E_k)\Big|$, and the distribution is binomial, so we may apply the Chernoff bound.

If $\Big|K_k(E_1,\dots,E_k)\Big| \leq \frac{1}{6}\eta n^k$, then failure to be (η,p)-upper regular means that the number of k-edges appearing on $K_k(E_1,\dots,E_k)$ is at least seven times the expected number; by the Chernoff bound the probability of this event is less than $\exp(-p\eta n^k) < \exp(-kn^{k-1})$.

If $\Big|K_k(E_1,\dots,E_k)\Big| \geq \frac{1}{6}\eta n^k$, then, by the Chernoff bound, the probability that more than $(1+\eta)p\Big|K_k(E_1,\dots,E_k)\Big|$ edges appear is at most

$$\exp\Big(-\frac{\eta^2 p\Big|K_k(E_1,\dots,E_k)\Big|}{3}\Big)\le\exp\Big(-\frac{\eta^2\frac{C}{n}\eta n^k}{18}\Big)=\exp\Big(-kn^{k-1}\Big)\,.$$

Since there are at most $2^{\binom{n}{k-1}}$ choices for each E_i, by the union bound the probability that $G^{(k)}(n,p)$ is not (η,p)-upper regular is at most

$$2^{k\binom{n}{k-1}}\exp\Big(-kn^{k-1}\Big)$$

which tends to zero as n tends to infinity. $\square$

Given a set $S\subseteq V(\Gamma)$, we say a $(k-1)$-set x is $(\varepsilon,p,1)$-*good for* S if we have

$$\Big|\Big\{s\in S:x\cup\{s\}\in E(\Gamma)\Big\}\Big|=p|S|\pm\varepsilon pn\,.$$

For each $\ell\ge 2$, we say inductively that a $(k-1)$-set x is (ε,p,ℓ)-*good for* S if it is $(\varepsilon,p,\ell-1)$-good for S and there are at most εpn edges of Γ which contain x and in addition contain a set which is not $(\varepsilon,p,\ell-1)$-good for S.

Lemma 4.2.9. *Given $\varepsilon>0$, $k\in\mathbb{N}$ there exists C such that if $p\ge\frac{C\log n}{n}$, then $\Gamma=G^{(k)}(n,p)$ a.a.s. has the following property. For each set $S\subseteq V(\Gamma)$, and each $1\le\ell\le\frac{1}{C}\log\log n$, there are at most $o(n)$ $(k-1)$-sets in $V(\Gamma)$ outside S which are not (ε,p,ℓ)-good for S.*

Proof. Given S, we first estimate the number of $(k-1)$-sets x which are outside S and not $(\varepsilon,p,1)$-good for S.

If $|S|<\frac{1}{6}\varepsilon n$, then failure of a given x to be $(\varepsilon,p,1)$-good for S means x forms an edge with at least $7p|S|$ vertices in S, the probability of which is by the Chernoff bound at most $\exp(-\varepsilon pn)$, which for large enough C is smaller than n^{-k}. If on the other hand $|S|\ge\frac{1}{6}pn$, then the probability that x does not form an edge with $(1\pm\varepsilon)p|S|$ vertices of S is at most $2\exp\Big(\frac{-\varepsilon^2p|S|}{3}\Big)<n^{-k}$ for large enough C. We see that in either case, the probability that x

is not $(\varepsilon, p, 1)$-good for S is at most n^{-k}. Now if x and x' are two different $(k-1)$-sets outside S, then the events of x and of x' being not $(\varepsilon, p, 1)$-good for S are independent, so again using the Chernoff bound we can estimate the likelihood of many sets being bad for S. The expected number of bad sets for S is at most $n^{k-1} \cdot n^{-k} = n^{-1}$. Therefore, for any $t \geq 1$, we can bound the probability that there are t or more bad $(k-1)$-sets for S by

$$\exp\left(-D(n^{-k} + tn^{1-k} || n^{-k})n^{k-1}\right) \leq \exp\left(-\frac{t \log n}{2}\right).$$

In particular, taking $t = 4n/\log n$ and using the union bound, the probability that there exists a set S for which more than $4n/\log n$ $(k-1)$-sets are not $(\varepsilon, p, 1)$-good is at most 2^{-n}. Suppose that Γ is such that this good event occurs, and in addition that every $(k-1)$-set of vertices of Γ is contained in at most $2pn$ edges of Γ.

Let $K = 2k\varepsilon^{-1}$. Now given S and $\ell \geq 1$, we claim that the number of $(k-1)$-sets outside S which are not (ε, p, ℓ)-good for S is at most $4n \cdot K^{\ell-1}/\log n$. We prove this by induction on ℓ; the base case $\ell = 1$ is the assumption on Γ. Suppose $\ell \geq 2$, and that the number of $(k-1)$-sets outside S which are not $(\varepsilon, p, \ell-1)$-good for S is at most $4n \cdot K^{\ell-2}/\log n$. For each $(k-1)$-set x outside S which is not $(\varepsilon, p, \ell-1)$-good for S, we assign to each $(k-1)$-set y such that $x \cup y$ is an edge of Γ one unit of badness. Observe that the total number of units of badness assigned is at most $(k-1) \cdot 2pn \cdot 4n \cdot K^{\ell-2}/\log n$. On the other hand, a set y which is $(\varepsilon, p, \ell-1)$-good for S can only fail to be (ε, p, ℓ)-good for S if it is assigned at least εpn units of badness. It follows that the total number of such sets is at most $2(k-1)\varepsilon^{-1} \cdot 4nK^{\ell-2}/\log n$, and so the number of $(k-1)$-sets outside S which are not (ε, p, ℓ)-good for S is at most

$$2(k-1)\varepsilon^{-1} \cdot 4nK^{\ell-2}/\log n + 4n \cdot K^{\ell-2}/\log n \leq 4n \cdot K^{\ell-1}/\log n\,,$$

as desired. In particular, this formula is in $o(n)$ for all $1 \leq \ell \leq \frac{1}{C}\log\log n$ with C large enough. $\square$

For a given $(k-1)$-tuple, we will find many paths starting from there. To get expansion we need to ensure that they have many different end-tuples.

Lemma 4.2.10. *For $\gamma > 0$, $k \geq 3$, any fixed integer $\ell > \frac{k-1}{\gamma} + k - 1$, and any $\mu > 0$ a.a.s. in $\Gamma = G^{(k)}(n,p)$ with $p = n^{-1+\gamma}$ the following holds. For any $(k-1)$-tuple $\mathbf{x}$ in $V(\Gamma)$ and a set $\mathcal{P}$ of at least $(\mu pn)^{\ell}$ tight paths in Γ with $\ell + (k-1)$ vertices and rooted at $\mathbf{x}$, the number of end $(k-1)$-tuples of the paths in $\mathcal{P}$ is at least $\frac{\mu^{2\ell}}{8(2\ell)!}n^{k-1}$. Moreover, when $(k-1)|\ell$, the same holds for spike paths rooted at $\mathbf{x}$.*

To prove Lemma 4.2.10 we need a concentration result of Kim and Vu [71]. We first give some definitions and then state the result. Let m be a positive integer and H be a hypergraph with $|V(H)| = m$ and each edge has at most r vertices. Let $p \in [0,1]$ and let $X_i, i \in V(H)$ be independent random variables with $\mathbb{P}[X_i = 1] = p$ and $\mathbb{P}[X_i = 0] = 1 - p$. We define the random variable

$$Y_H = \sum_{f \in E(H)} \prod_{i \in f} X_i.$$

For each subset $A \subseteq V(H)$, we define the A-truncated subgraph $H(A)$ of H to be the subgraph of H with $V(H(A)) = V(H) \setminus A$ and $E(H(A)) = \{f \subseteq V(H(A)) \colon f \cup A \in E(H)\}$. Hence

$$Y_{H(A)} = \sum_{\substack{f \in E(H) \\ A \subseteq f}} \prod_{i \in f \setminus A} X_i.$$

Now, for $0 \leq i \leq r$, we set $\mathcal{E}_i(H) = \max_{A \subseteq V(H), |A| = i} \mathbb{E}[Y_{H(A)}]$. Note that $\mathcal{E}_0(H) = \mathbb{E}[Y_H]$. Finally, we let $\mathcal{E}(H) = \max_{0 \leq i \leq r} \mathcal{E}_i(H)$ and $\mathcal{E}'(H) = \max_{1 \leq i \leq r} \mathcal{E}_i(H)$.

Theorem 4.2.11 (Kim-Vu polynomial concentration [71])**.** *In this setting we have*

$$\mathbb{P}[|Y_H - \mathbb{E}(Y_H)| > a_r(\mathcal{E}(H)\mathcal{E}'(H))^{1/2}\lambda^r)] = O(\exp(-\lambda + (r-1)\log m))$$

for any $\lambda > 1$ and $a_r = 8^r r!^{1/2}$.

Moreover, we will need the following definitions. Let $k \geq 3$, $\ell \geq k-1$, and $\gamma > 0$.

Roughly speaking, we define D_ℓ to be the k-graph obtained from two vertex-disjoint tight paths on $\ell + k - 1$ vertices by identifying the end $(k-1)$-tuples and let $\mathcal{D}_\ell$ be the set of hypergraphs obtained from D_ℓ by additionally identifying some (or none) of the not yet identified vertices from the first tight path with such vertices from the second without completely collapsing it into a tight path. More precisely, we let $U = \{u_1, \dots, u_{\ell+k-1}\}$ and $W = \{w_1, \dots, w_{\ell+k-1}\}$ be two sets of vertices that are disjoint except that $\mathbf{x} = (u_1, \dots, u_{k-1}) = (w_1, \dots, w_{k-1})$ and $\mathbf{y} = (u_{\ell+k-1}, \dots, u_{\ell+1}) = (w_{\ell+k-1}, \dots, w_{\ell+1})$. Then D_ℓ is the hypergraph with vertex set $U \cup W$ and edge set

$$\{\{u_i, \dots, u_{i+k-1}\} \colon i \in [\ell]\} \cup \{\{w_i, \dots, w_{i+k-1}\} \colon i \in [\ell]\}.$$

For $0 \leq j \leq \ell - (k-1)$, we denote by $\mathcal{D}_\ell^j$ the graphs obtained from D_ℓ by taking sets $I_1, I_2 \subseteq \{k, \dots, \ell\}$ each of size j and a bijection $\sigma \colon I_1 \to I_2$ and identifying u_i with $w_{\sigma(i)}$ for all $i \in I_1$, where, if $j = \ell - (k-1)$, then we do not allow σ to be the identity (since that would collapse D_ℓ into a tight path). We say that such a graph $F \in \mathcal{D}_\ell^j$ is rooted at $\mathbf{x}$ and call the vertices in $\mathbf{y}$ the end-vertices of F. Moreover, for any edge $\{u_i, \dots, u_{i+k-1}\}$ in $F \in \mathcal{D}_\ell^j$ as above we call u_i the *first vertex* and u_{i+k-1} the *last vertex* of the edge and say that the edge *starts* in u_i and *ends* in u_{i+k-1} (and analogously for edges contained in W). Finally, we let $\mathcal{D}_\ell = \bigcup_{0 \leq j \leq \ell-(k-1)} \mathcal{D}_\ell^j$.

We now prove the following lemma which we will use to prove Lemma 4.2.10.

Lemma 4.2.12. *For $\gamma > 0$, $k \geq 3$, and any fixed integer $\ell > \frac{k-1}{\gamma} + k - 1$ a.a.s. in $\Gamma = G^{(k)}(n,p)$ with $p = n^{-1+\gamma}$ the following holds. For all $(k-1)$-tuples $\mathbf{x}$ in $V(\Gamma)$, the number of copies of elements of $\mathcal{D}_\ell$ in Γ that are rooted at $\mathbf{x}$ is at most $2p^{2\ell}n^{2\ell-(k-1)}$.*

Proof. Fix a $(k-1)$-tuple $\mathbf{x}$ in $V(\Gamma)$ and an integer $\ell > \frac{k-1}{\gamma} + k - 1$ and let n be large enough. We call the vertices in $\mathbf{x}$ *rooted* and any other vertices *unrooted*. Let $F \in \mathcal{D}_\ell$ and consider the complete k-graph $K_n^{(k)}$ on n vertices. We define a hypergraph H_F as follows.

Let $V(H_F) = E(K_n^{(k)})$ and let

$$E(H_F) = \left\{ \mathcal{F} \in \binom{E(K_n^{(k)})}{e(F)} : \mathcal{F} \text{ spans a copy of } F \text{ in } K_n^{(k)} \text{ rooted at } \mathbf{x} \right\}.$$

Note that, since $e(F) \leq 2\ell$, each edge in H_F has size at most 2ℓ. For each $e \in V(H_F)$ let X_e be the random variable for which $X_e = 1$ if e is an edge of Γ and $X_e = 0$ otherwise. Note that $\mathbb{P}[X_e = 1] = p$. It is easy to see that with these definitions Y_{H_F} is the number of copies of F in Γ rooted at $\mathbf{x}$. Since $e(D_\ell) = 2\ell$, $v(D_\ell) = 2\ell$, and $k-1$ vertices are rooted, we have $\binom{n}{2\ell-(k-1)} p^{2\ell} \leq \mathbb{E}[Y_{H_{D_\ell}}] \leq p^{2\ell} n^{2\ell-(k-1)}$, in particular $\mathbb{E}[Y_{H_{D_\ell}}] = \Theta(p^{2\ell} n^{2\ell-(k-1)})$.

Claim 4.2.13. *For $F \in \mathcal{D}_\ell \setminus \{D_\ell\}$, we have*

$$\mathbb{E}\left[Y_{H_F}\right] = o\left(\mathbb{E}\left[Y_{H_{D_\ell}}\right]\right).$$

Proof of Claim. We split the proof into two cases depending on the integer j for which we have $F \in \mathcal{D}_\ell^j$.

First suppose that $F \in \mathcal{D}_\ell^j$ for some $j \in \{1, \ldots, \ell - 2(k-1)\}$. Note that $v(F) = 2\ell - j$. We claim that $e(F) \geq 2\ell - j$. This can be seen as follows. Recall that F is obtained from D_ℓ by identifying j additional vertices from the first tight path in D_ℓ with vertices from the second. This leaves $\ell - (k-1) - j \geq k-1$ unidentified vertices in the first tight path. In addition to the ℓ edges in the second path, F contains an edge ending in each of the unidentified vertices and one more additional edge starting with each of the last $k-1$ unidentified vertices (these edges cannot end in an unidentified vertex, so there is no double counting). Thus

$$e(F) \geq \ell + \ell - (k-1) - j + (k-1) = 2\ell - j.$$

Hence, since $k-1$ vertices are rooted,

$$\mathbb{E}[Y_{H_F}] \leq p^{2\ell-j} n^{2\ell-j-(k-1)} = p^{2\ell} n^{2\ell-(k-1)} (pn)^{-j} = p^{2\ell} n^{2\ell-(k-1)} n^{-j\gamma} = o\left(\mathbb{E}\left[Y_{H_{D_\ell}}\right]\right).$$

Now suppose that $F \in \mathcal{D}_\ell^j$ for some $j \in \{\ell - 2(k-1)+1, \dots, \ell-(k-1)\}$. As in the previous case, in addition to the ℓ edges in the second path, F contains an edge ending in each of the $\ell-(k-1)-j$ unidentified vertices. Thus $e(F) \geq 2\ell-(k-1)-j$. Hence, since $v(F) = 2\ell - j$, $k-1$ vertices are rooted, and $j > \ell - 2(k-1)$, we have

$$\begin{aligned}\mathbb{E}[Y_{H_F}] &\leq p^{2\ell-j-(k-1)} n^{2\ell-j-(k-1)} = p^{2\ell} n^{2\ell-(k-1)} (pn)^{-j} p^{-(k-1)} \\ &= p^{2\ell} n^{2\ell-(k-1)} n^{-j\gamma+(k-1)-\gamma(k-1)} < p^{2\ell} n^{2\ell-(k-1)} n^{-\gamma(\ell-2(k-1))+(k-1)-\gamma(k-1)} \\ &= p^{2\ell} n^{2\ell-(k-1)} n^{-\gamma(\ell-(k-1))+(k-1)} = o\left(\mathbb{E}\left[Y_{H_{D_\ell}}\right]\right),\end{aligned}$$

since $\ell > \frac{k-1}{\gamma} + k - 1$. ■

Combining the claim with our bound on $\mathbb{E}[Y_{H_{D_\ell}}]$ we obtain

$$\sum_{F \in \mathcal{D}_\ell} \mathbb{E}[Y_{H_F}] \leq \frac{3}{2} p^{2\ell} n^{2\ell-(k-1)}. \tag{4.2.1}$$

Next we show that, for all $F \in \mathcal{D}_\ell$, the random variable Y_{H_F} is concentrated around its expectation.

Claim 4.2.14. *For all $F \in \mathcal{D}_\ell$, we have*

$$\mathbb{P}\left[|Y_{H_F} - \mathbb{E}[Y_{H_F}]| > \frac{p^{2\ell} n^{2\ell-(k-1)}}{2|\mathcal{D}_\ell|}\right] = O\left(\exp(-n^{\gamma/(6\ell)})\right).$$

Proof of Claim. Let $F \in \mathcal{D}_\ell$. We first show that $\mathcal{E}'(H_F) \leq n^{2\ell-(k-1)} p^{2\ell} n^{-\gamma}$. To that end let $1 \leq i \leq 2\ell$ and $A \subseteq V(H_F) = E(K_n^{(k)})$ with $|A| = i$. Note that it suffices to show that $\mathbb{E}[Y_{H_F(A)}] \leq n^{2\ell-(k-1)} p^{2\ell} n^{-\gamma}$. We let v_A be the number of vertices covered by A, that is, $v_A = |\bigcup A|$ and we let r_A be the number of rooted vertices not covered by A, that is, $r_A = |\mathbf{x} \setminus \bigcup A|$. Moreover, we call edges in A *covered edges*, edges not in A *uncovered edges*, vertices in $\bigcup A$ *covered vertices* and vertices not in $\bigcup A$ *uncovered vertices*. Note that $Y_{H_F(A)}$ is the number of copies of F in $\Gamma + A$ that are rooted at $\mathbf{x}$ and contain A. So if A is not contained in (the edge set of) any copy of F rooted at $\mathbf{x}$, then $Y_{H_F(A)} = 0$ and we are done. Now assume that A is contained in a copy of F rooted at $\mathbf{x}$. We consider two cases.

First suppose that the number of uncovered unrooted vertices in any copy of F rooted at $\mathbf{x}$ and containing A is at most $\ell+k-2$, that is, $v(F)-v_A-r_A \leq \ell+k-2$. Note that there are at least $v(F)-v_A-r_A$ uncovered edges, that is, $e(F)-i \geq v(F)-v_A-r_A$ as for each uncovered unrooted vertex, there is at least one uncovered edge ending in that vertex. Thus

$$\begin{aligned}\mathbb{E}[Y_{H_F(A)}] &\leq n^{v(F)-v_A-r_A}p^{e(F)-i} \leq (np)^{v(F)-v_A-r_A} \leq n^{\gamma(\ell+k-2)} = n^{\gamma\ell+\gamma(k-2)} \\ &\leq n^{2\gamma\ell-(k-1)-\gamma(k-1)+\gamma(k-2)} = n^{2\gamma\ell-(k-1)-\gamma} = n^{2\ell-(k-1)}p^{2\ell}n^{-\gamma},\end{aligned}$$

where the last inequality follows from the fact that $\ell > \frac{k-1}{\gamma}+k-1$.

Now we consider the case where the number of uncovered unrooted vertices in any copy of F rooted at $\mathbf{x}$ and containing A is at least $\ell+k-1$, that is, $v(F)-v_A-r_A \geq \ell+k-1$. Consider such a copy of F. Recall that F is obtained from two tight paths by identifying vertices. Note that in this case there are at least $k-1$ vertices in each of the tight paths that are unrooted, uncovered, not on the other tight path, and not part of the $k-1$ end-vertices of F. Let B be the set of the last $k-1$ such vertices on the first tight path. We now show that F has at least $v(F)-v_A-r_A+k-1$ uncovered edges, that is, $e(F)-i \geq v(F)-v_A-r_A+k-1$. Note that for each uncovered unrooted vertex there is at least one uncovered edge ending in that vertex. That gives us $v(F)-v_A-r_A$ uncovered edges. We show that for each vertex in B there is at least one more uncovered edge that we have not yet counted. For each vertex $u \in B$, consider the uncovered edge starting in u. If it ends in a covered vertex, then it is one more uncovered edge that we have not yet counted. If it ends in an uncovered vertex w, then the corresponding edge on the other tight path is also uncovered. Thus there are two uncovered edges ending in w and we have only counted one of them. Since there are $k-1$ vertices in B, it follows that there are at least $v(F)-v_A-r_A+k-1$ uncovered edges in F. Thus for any copy of F containing A, we have

$$2\ell - 1 \geq e(F) - i \geq v(F) - v_A - r_A + k - 1.$$

Hence

$$\begin{aligned}
\mathbb{E}[Y_{H_F(A)}] &\leq n^{v(F)-v_A-r_A} p^{e(F)-i} \leq n^{e(F)-i-(k-1)} p^{e(F)-i} \\
&\leq (np)^{e(F)-i} n^{-(k-1)} \leq (np)^{2\ell-1} n^{-(k-1)} = n^{2\ell-(k-1)} p^{2\ell} n^{-\gamma}.
\end{aligned}$$

We have shown that $\mathcal{E}'(H_F) \leq n^{2\ell-(k-1)} p^{2\ell} n^{-\gamma}$. Now note that, since $\mathbb{E}[Y_{H_F}] \leq \mathbb{E}[Y_{H_{D_\ell}}] \leq n^{2\ell-(k-1)} p^{2\ell}$, we have $\mathcal{E}(H_F) = \max\{\mathcal{E}'(H_F), \mathbb{E}[Y_{H_F}]\} \leq n^{2\ell-(k-1)} p^{2\ell}$. It follows that

$$(\mathcal{E}(H_F)\mathcal{E}'(H_F))^{1/2} \leq n^{2\ell-(k-1)} p^{2\ell} n^{-\gamma/2}.$$

Therefore, with

$$\lambda_F = \left(\frac{p^{2\ell} n^{2\ell-(k-1)}}{2|\mathcal{D}_\ell| a_{2\ell} (\mathcal{E}(H_F)\mathcal{E}'(H_F))^{1/2}} \right)^{1/(2\ell)} \geq \left(\frac{n^{\gamma/2}}{2|\mathcal{D}_\ell| a_{2\ell}} \right)^{1/(2\ell)} \geq n^{\gamma/(5\ell)},$$

we have, by Theorem 4.2.11,

$$\begin{aligned}
\mathbb{P}\left[|Y_{H_F} - \mathbb{E}[Y_{H_F}]| > \frac{p^{2\ell} n^{2\ell-(k-1)}}{2|\mathcal{D}_\ell|} \right] &= \mathbb{P}[|Y_{H_F} - \mathbb{E}[Y_{H_F}]| > a_{2\ell} (\mathcal{E}(H_F)\mathcal{E}'(H_F))^{1/2} \lambda_F^{2\ell}] \\
&= O\left(\exp\left(-\lambda_F + (2\ell - 1) \log \binom{n}{k} \right) \right) \\
&= O\left(\exp(-n^{\gamma/(5\ell)} + (2\ell-1) k \log n) \right) = O\left(\exp(-n^{\gamma/(6\ell)}) \right).
\end{aligned}$$

■

Now let $Z_{\mathbf{x}}$ be the number of copies of elements of $\mathcal{D}_\ell$ in Γ rooted at $\mathbf{x}$. Note that

$Z_{\mathbf{x}} = \sum_{F \in \mathcal{D}_\ell} Y_{H_F}$. We have

$$\begin{aligned}
\mathbb{P}[Z_{\mathbf{x}} > 2p^{2\ell}n^{2\ell-(k-1)}] &\leq \mathbb{P}\left[\left|\sum_{F \in \mathcal{D}_\ell} (Y_{H_F} - \mathbb{E}[Y_{H_F}])\right| > 2p^{2\ell}n^{2\ell-(k-1)} - \sum_{F \in \mathcal{D}_\ell} \mathbb{E}[Y_{H_F}]\right] \\
&\overset{(4.2.1)}{\leq} \mathbb{P}\left[\sum_{F \in \mathcal{D}_\ell} |Y_{H_F} - \mathbb{E}[Y_{H_F}]| > \frac{1}{2}p^{2\ell}n^{2\ell-(k-1)}\right] \\
&\leq \sum_{F \in \mathcal{D}_\ell} \mathbb{P}\left[|Y_{H_F} - \mathbb{E}[Y_{H_F}]| > \frac{p^{2\ell}n^{2\ell-(k-1)}}{2|\mathcal{D}_\ell|}\right] \\
&= O\left(|\mathcal{D}_\ell| \exp(n^{-\gamma/(6\ell)})\right),
\end{aligned}$$

where the last inequality follows from the union bound together with the fact that if the average of the values $|Y_{H_F} - \mathbb{E}[Y_{H_F}]|$ is at least $\frac{p^{2\ell}n^{2\ell-(k-1)}}{2|\mathcal{D}_\ell|}$, then at least one of the values has to be that large. Finally, the result follows by the union bound over all $(k-1)$-tuples $\mathbf{x}$ in $V(\Gamma)$. $\square$

We are now ready to prove Lemma 4.2.10.

Proof of Lemma 4.2.10. Let $\mathbf{x}$ be a $(k-1)$-tuple in $V(\Gamma)$ and $\mathcal{P}$ be a set of at least $(\mu p n)^\ell$ tight paths in Γ with $\ell + (k-1)$ vertices and rooted at $\mathbf{x}$. Let Q be the set of end-tuples we reach from $\mathbf{x}$ with paths in $\mathcal{P}$. For each $\mathbf{q} \in Q$, let $\mathcal{P}_{\mathbf{q}}$ be those paths in $\mathcal{P}$ that end in $\mathbf{q}$. Note that, for $\mathbf{q} \in Q$ and distinct elements $P, P' \in \mathcal{P}_{\mathbf{q}}$, we have $P \cup P' \in \mathcal{D}_\ell$. Thus for each $\mathbf{q} \in Q$, there are at least $\frac{1}{(2\ell)!}\binom{|\mathcal{P}_{\mathbf{q}}|}{2}$ copies of elements of $\mathcal{D}_\ell$ in Γ rooted at $\mathbf{x}$ and ending in $\mathbf{q}$ (we divide by $(2\ell)!$ since there are at most $(2\ell)!$ ways the union of two paths in $\mathcal{P}$ could result in the same copy of an element of $\mathcal{D}_\ell$). Hence the number of copies of elements of $\mathcal{D}_\ell$ in Γ rooted at $\mathbf{x}$ is at least

$$\sum_{\mathbf{q} \in Q} \frac{1}{(2\ell)!}\binom{|\mathcal{P}_{\mathbf{q}}|}{2} \geq \frac{|Q|}{(2\ell)!}\binom{\frac{(\mu p n)^\ell}{|Q|}}{2} \geq \frac{(\mu p n)^{2\ell}}{4(2\ell)!|Q|},$$

where the penultimate inequality follows by Jensen's inequality. Thus, by Lemma 4.2.12, we have a.a.s.

$$2p^{2\ell}n^{2\ell-(k-1)} \geq \frac{(\mu p n)^{2\ell}}{4(2\ell)!|Q|}$$

and thus

$$|Q| \geq \frac{\mu^{2\ell}}{8(2\ell)!} n^{k-1}.$$

Moreover, an analogous argument shows the result for spike paths. $\square$

Together with the definition of tuples that are (ε, p, ℓ)-good for S, Lemma 4.2.10 implies the following.

Corollary 4.2.15. *For any $\gamma > 0$ and any $0 < \varepsilon \leq \frac{1}{4}\gamma$, and integers s, $k \geq 3$, and $\ell > \frac{k-1}{\gamma} + k - 1$, there exists $\nu > 0$ such that in $\Gamma = G^{(k)}(n, p)$ a.a.s. the following holds when $p = n^{-1+\gamma}$. Let $G \subseteq \Gamma$ satisfy $\delta_{k-1}(G) \geq \left(\frac{1}{2} + \gamma\right)pn$. Let $S, S' \subseteq V(\Gamma)$ be sets with $|S| \leq \frac{1}{2}n$ and $|S'| \leq s$. Let $\mathbf{x}$ be a $(k-1)$-tuple, which is (ε, p, ℓ)-good for S. Then there are at least νn^{k-1} different $(k-1)$-tuples $\mathbf{y}$, such that there exists a tight path in G of length ℓ with ends $\mathbf{x}$ and $\mathbf{y}$ and no vertices of the path in $S \cup S'$ except for possibly some of the vertices in $\mathbf{x}$. Moreover, when $(k-1)|\ell$, the same holds for spike paths in G of length ℓ.*

Proof. We only prove the statement for tight paths as it is easy to see that the proof can be adapted for spike paths. We set $\mu = \frac{1}{4}\gamma$, and $\nu = \frac{\mu^{2\ell}}{8(2\ell)!}$. Suppose that the good event of Lemma 4.2.10, with input γ, k, and μ, holds for $\Gamma = G^{(k)}(n, p)$.

Given G and $\mathbf{x}$ as in the lemma statement, let $\mathbf{x} = \left(x_1, \ldots, x_{k-1}\right)$. We construct tight paths $x_1 \ldots x_{\ell+k-1}$ rooted at $\mathbf{x}$ by choosing vertices $x_k, \ldots, x_{\ell+k-1}$ one by one as follows. For each $k \leq i \leq \ell + k - 1$, we choose x_i such that $x_i \notin S \cup S' \cup \{x_1, \ldots, x_{i-1}\}$ and $\{x_{i-k+1}, \ldots, x_i\} \in E(G)$. If $i < \ell + k - 1$, we insist in addition that $\{x_{i-k+2}, \ldots, x_i\}$ is $(\varepsilon, p, \ell - (i - k + 1))$-good for S. Since $\mathbf{x}$ is (ε, p, ℓ)-good for S, for each $k \leq i \leq \ell + k - 1$, the number of choices for each x_i, such that $\{x_{i-k+1}, \ldots, x_i\}$ is an edge of G, $x_i \notin S \cup S' \cup \{x_1, \ldots, x_{i-1}\}$, and $\{x_{i-k+2}, \ldots, x_i\}$ is $(\varepsilon, p, \ell - (i - k + 1))$-good, is at least

$$\left(\tfrac{1}{2} + \gamma\right)pn - (p\,|S| + \varepsilon pn) - s - \ell - (k-1) - \varepsilon pn \geq (\gamma - 2\varepsilon)pn - s - \ell - (k-1) \geq \tfrac{1}{4}\gamma pn\,.$$

Let $\mathcal{P}$ be the set of tight paths constructed in this way; then we have $|\mathcal{P}| \geq \left(\frac{1}{4}\gamma pn\right)^{\ell} =$

$(\mu pn)^\ell$. Since the good event of Lemma 4.2.10 holds, the number of end $(k-1)$-tuples of these paths is at least $\frac{\mu^{2\ell}}{8(2\ell)!}n^{k-1} = \nu n^{k-1}$, as desired. $\square$

4.2.5 Connecting lemma

The next lemma will enable us to connect two $(k-1)$-tuples, which are (ε', p, ℓ)-good for some set S, by a path of length at most ℓ avoiding S.

Lemma 4.2.16 ([8, Lemma 20]). *Given $k \geq 3$, and $\gamma > 0$, there exists an integer ℓ such that for any integer s the following holds. For any $d, \eta > 0$, any $0 < \varepsilon' \leq \frac{1}{4}\gamma$, any integer t_0, any small enough $\nu, \varepsilon_k > 0$, any functions $\varepsilon, f, f_k : \mathbb{N} \to (0, 1]$ which tend to zero sufficiently fast, and any large enough $t_1, t_2 \in \mathbb{N}$, the following holds a.a.s. in $\Gamma = G^{(k)}(n, p)$ with $p \geq n^{-1+\gamma}$. Suppose $G \subseteq \Gamma$ is an n-vertex k-graph with $\delta_{k-1}(G) \geq \left(\frac{1}{2} + \gamma\right)pn$, that $(\mathcal{P}_c^*, \mathcal{P}_f^*)$ is a $(t_0, t_1, t_2, \varepsilon_k, \varepsilon(t_1), f_k(t_1), f(t_2), p)$-strengthened pair for G, and that t is the number of 1-cells in $\mathcal{P}_c^*$. Let $\mathcal{R}' \subseteq \mathcal{R} = \mathcal{R}_{\varepsilon_k, d}(G; \mathcal{P}_c^*, \mathcal{P}_f^*)$ be an induced subcomplex of the (ε_k, d)-reduced multicomplex of G on at least $(1-\nu)t$ 1-edges and assume that it is tightly linked. Further, let $S \subseteq V(G)$ be such that $|S| \leq \frac{1}{2}n$ and it intersects all 1-cells of $\mathcal{R}'$ in at most an $(1-\eta)$-fraction. Then for any two $(k-1)$-tuples $\mathbf{x}$ and $\mathbf{y}$, which are (ε', p, ℓ)-good for S, and any set S' of size at most s, there exists a tight path of length ℓ with ends $\mathbf{x}$ and $\mathbf{y}$.*

In this thesis, we omit the proof of Lemma 4.2.16 (for the proof see [8, Lemma 20]). In [8] we prove this lemma by using Lemma 4.2.17, which allows us to connect a fraction of any good $(k-1)$-cell to a fraction of an adjacent good $(k-1)$-cell, where adjacency is with respect to regular polyads.

Lemma 4.2.17 ([8, Lemma 21]). *Given $k \geq 3$ and $\gamma > 0$, there exists an integer ℓ such that for any integer s the following holds. For any $d, \eta, \nu > 0$, any $t_0 \in \mathbb{N}$, any small enough $\varepsilon_k > 0$, any functions $\varepsilon, f : \mathbb{N} \to (0, 1]$ which tend to zero sufficiently fast, any integers $t_2 \geq t_1 \geq t_0$, and any small enough $f_k > 0$, the following holds a.a.s. in $\Gamma = G^{(k)}(n, p)$ with $p \geq n^{-1+\gamma}$. Suppose $G \subseteq \Gamma$ is an n-vertex k-graph, that $(\mathcal{P}_c^*, \mathcal{P}_f^*)$ is* *a $(t_0, t_1, t_2, \varepsilon_k, \varepsilon(t_1), f_k, f(t_2), p)$-strengthened pair for G, let $H := \hat{P}(Q; \mathcal{P}_c^*)$ be a regular polyad in the reduced complex $\mathcal{R}_{\varepsilon_k,d}(G; \mathcal{P}_c^*, \mathcal{P}_f^*)$, $V_1, \ldots, V_k$ its underlying 1-cells, and $S \subseteq V(G)$ is a set intersecting each of these in at most an $(1-\eta)$-fraction. Further, let E_1 and E_k be the two $(k-1)$-cells of H missing V_1 and V_k, respectively.*

Let the tuples in E_1 and E_k be ordered according to $V_1, \ldots, V_k$. Then there is $\overline{E}_k \subseteq E_k$ with $|\overline{E}_k| \geq (1-\nu)|E_k|$, such that for any $\mathbf{x} \in \overline{E}_k$ and any set S' of at most s vertices there is a tight path from $\mathbf{x}$ to $\mathbf{y}$ of length ℓ with internal vertices not in $S \cup S'$ for a $(1-\nu)$-fraction of the tuples $\mathbf{y} \in E_1$.

With this lemma and Corollary 4.2.15 it is straightforward to prove Lemma 4.2.16. The idea of the proof is to first expand from the tuples $\mathbf{x}$ and $\mathbf{y}$ using Corollary 4.2.15 and then connect two of the many ends that we found with Lemma 4.2.17 by following a tight link given by Lemma 4.2.5.

In this thesis, we also omit the proof of Lemma 4.2.17 (for the full details see [8, Lemma 21]) but here we briefly explain the idea for the proof by sketching an easier version. Suppose that $k = 2$ (i.e. we are dealing with graphs, not hypergraphs) and rather than having two clusters which are adjacent in the reduced graph, we have a path of $\ell + 1$ clusters $V_1, \ldots, V_\ell, V_{\ell+1}$ in the reduced graph. We want to show that for most vertices $x \in V_1$, there is a path from x to y for most $y \in V_{\ell+1}$. To begin with, we look at the fine parts within V_ℓ. We discard those fine parts which do not form $(f_2, \frac{1}{2}d, p)$-regular pairs with most fine parts in $V_{\ell+1}$; by definition of the reduced graph, there are few such, and we let X_ℓ be the remaining subset of V_ℓ. Next, for each $i = \ell - 1, \ldots, 1$ we discard from V_i those vertices with fewer than $(d - \varepsilon_2)p|X_{i+1}|$ neighbours in X_{i+1} to obtain X_i. Again, by regularity we discard few vertices at each step, so X_1 is most of V_1. Now if we choose any $x \in X_1$, we claim there is a path from x to y for most $y \in V_{\ell+1}$.

To see this, note that there are many paths which start at x and go out to X_ℓ: we can construct these paths greedily starting from x, and we have at least $\frac{1}{2}dp|V_i|$ choices in each X_i. By Lemma 4.2.10 and choice of ℓ, there are linearly many different endvertices of these paths in X_ℓ. We call this the *coarse expansion.* However the number of these endvertices will be much *smaller* than $\varepsilon_2|X_\ell|$, so we cannot use the coarse regularity to say anything about the set of endvertices. This is where we need the fine partition: we can ensure the fine regularity constant f_2 is so small that the number of endvertices is much larger than $f_2|X_\ell|$. By averaging, there is a fine part Z contained in X_ℓ which contains a set R_0 of endvertices, where $|R_0| \geq f_2|Z|$. Now Z forms a $(f_2, \frac{1}{2}d, p)$-regular pair with most fine parts in $V_{\ell+1}$. For any such fine part Z', by $(f_2, \frac{1}{2}d, p)$-regularity, the set $\overline{R_1}$ of vertices in Z' which we cannot reach, i.e. which do not send an edge to R_0, is of size at most $f_2|Z'|$. In other words, we have found, for most fine parts Z' in $V_{\ell+1}$, a path from x to most vertices of Z'; that is the desired paths to most vertices of $V_{\ell+1}$. We call this second step the *fine expansion*.

It is fairly easy to see that this strategy still works with sets S and S' to avoid. It is also not very hard to modify it to work with one regular pair rather than a path of regular pairs: we split off a small fraction of each cluster to use for the coarse expansion (and we do not reuse this part for the fine expansion). What is not, however, so easy is to make this argument work for k-graphs for $k \geq 3$. The coarse expansion step works much as described above, but the fine expansion requires more care. If we are given $k = 3$ and a regular polyad on parts (X, Y, Z), and a significant fraction of the XY 2-cell are marked as end-tuples of tight paths from some given $\mathbf{x}$, then we cannot necessarily conclude that almost all pairs in the YZ 2-cell are end-tuples of tight paths from $\mathbf{x}$. We can only conclude this for those pairs whose vertex in Y is also in many marked pairs. However this does then imply that most vertices of Z are in YZ pairs which form an edge with a marked pair, and taking another step, using another regular polyad (Y, Z, W), we can finally argue that most ZW pairs are end-tuples of tight paths from $\mathbf{x}$; so the ZW pairs play the same role as Z' in the argument sketched above. For higher uniformity, we generalise this argument; in uniformity k, we need $k-1$ steps.

The following lemma deals with the fine expansion mentioned above. It is also used in proving Lemma 4.2.21. In this thesis, we omit the proof of this lemma (for the proof see [8, Lemma 25]).

Lemma 4.2.18 ([8, Lemma 25]). *Given $k \geq 3$ and δ, $d_k > 0$, for all sufficiently small $f'_k > 0$ we have: given $d_0 > 0$, for all sufficiently small $f' > 0$ and all sufficiently large m the following holds.*

Given a set V of vertices, suppose that we have a ground partition $\mathcal{P} = \{X_0, \dots, X_{2k-3}\}$ with $|X_i| = m$ for each i, and for each $2 \leq i \leq k-1$ a $\mathcal{P}$-partite i-graph G_i on V such that for each $Y \subseteq \{0, \dots, 2k-3\}$ the graph $G_i\Big[\prod_{y \in Y} X_y\Big]$ is $(d_i, f', 1)$-regular with respect to G_{i-1} (where we assume $E(G_1) = V$). Furthermore suppose that we have a $\mathcal{P}$-partite k-graph G_k, such that $G_k[X_j, \dots, X_{j+k-1}]$ is (d_k, f'_k, p)-regular with respect to G_{k-1} for each $0 \leq j \leq k-2$. Suppose that $d_i \geq d_0$ for each $2 \leq i \leq k-1$. Suppose that for each $2 \leq i \leq k$, all the edges of G_i are supported by G_{i-1}.

Suppose that we are given a set $R_0 \subseteq G_{k-1}[X_0, \dots, X_{k-2}]$ of size at least

$$\delta m^{k-1} \prod_{\ell=2}^{k-1} d_\ell^{\binom{k-1}{\ell}} .$$

Let $R_{k-1} \subseteq G_{k-1}[X_{k-1}, \dots, X_{2k-3}]$ be those $(k-1)$-edges which are the end-tuples of some tight path in G_k with one vertex in each of $X_0, \dots, X_{2k-3}$ and whose start $(k-1)$-tuple is in R_0.

Then we have $|R_{k-1}| \geq (1-\delta) m^{k-1} \prod_{\ell=2}^{k-1} d_\ell^{\binom{k-1}{\ell}}$.

4.2.6 Fractional matchings

While the clusters of a regular partition are all the same size, and are still about the same size after we remove the reservoir set, the reservoir path may intersect the clusters in very different amounts. When we extend the reservoir path to an almost-spanning path, this means we need to use different numbers of vertices in the different clusters. To guide the construction of the almost-spanning path, the following lemma returns a fractional matching in the cluster graph such that the total weight on each cluster is at most the fraction of vertices still to use in that cluster, and the total weight of the fractional matching is very close to $\frac{1}{k}$ times the fraction of vertices in total still to use.

Lemma 4.2.19. *Let H be an m-vertex k-complex, and let $w : V(H) \to [0,1]$ be a weight function. Given $\varepsilon > 0$, suppose that H has at least $(1-\varepsilon)m$ edges of size 1, and that for each $1 \le i \le k-2$, each i-edge of H is contained in at least $(1-\varepsilon)m$ edges of size $i+1$. Finally suppose that each $(k-1)$-edge of H is contained in at least $\left(\frac{1}{2}+\gamma\right)m$ edges of size k, and suppose $\sum_{v\in V(H)} w(v) \ge (1-\gamma)m$. Then there is a weight function $w^* : E\left(H^{(k)}\right) \to [0,1]$ such that for each $v \in V(H)$ we have $\sum_{e\ni v} w^*(e) \le w(v)$ and $\sum_{e\in E\left(H^{(k)}\right)} w^*(e) \ge \left(\sum_{v\in V(H)} w(v) - \varepsilon m\right) \cdot \frac{1}{k}$.*

Proof. Consider the linear program

$$\begin{aligned} &\text{maximise} \quad \sum_{e\in E(H^{(k)})} w^*(e) \\ &\text{subject to} \quad \sum_{e\ni v} w^*(e) \le w(v) \text{ for each } v \in V(H) \text{ and } w^*(e) \ge 0\,. \end{aligned}$$

The dual program has variables $y : V(H) \to [0,1]$ such that for each $e \in E\left(H^{(k)}\right)$ we have $\sum_{v\in e} y(v) \ge 1$, where we minimise $\sum_{v\in V(H)} y(v)w(v)$. Suppose that y is a feasible solution to the dual program.

We order $V(H)$ according to decreasing y. We find a k-edge of H as follows. We take the last v_1 such that $\{v_1\}$ is a 1-edge of H. Then for each $2 \le i \le k$ in succession, we choose the last vertex v_i such that $\{v_1,\dots,v_i\}$ is an i-edge of H.

For each $1 \le i \le k-1$, since by construction $\{v_1,\dots,v_{i-1}\}$ is an $(i-1)$-edge of H, there are at most εm choices of v_i which do not give an i-edge of H, and in particular v_i will be at or after position $(1-\varepsilon)m$ in the order. Finally since $\{v_1,\dots,v_{k-1}\}$ is a $(k-1)$-edge of H, necessarily v_k will be at position at or after $\left(\frac{1}{2}+\gamma\right)m$ in the order.

Suppose that the vertex v of H at position $(1-\varepsilon)m$ in the order satisfies $y(v) = a$, and let $y(v_k) = b$. Then we have $(k-1)a + b \ge \sum_{i=1}^{k} y(v_i) \ge 1$, where the second inequality is since y is feasible for the dual program. On the other hand, let α denote the sum of $w(u)$ over vertices u equal to or earlier in the order than v_k, and let β denote the sum of $w(u)$ over vertices u after v_k but not after v (where v is at position $(1-\varepsilon)m$ in the order). Then we have $\sum_{v\in V(H)} w(v)y(v) \ge \alpha b + \beta a$.

We view this as an optimisation problem: given $0 \le a \le b \le 1$ such that $(k-1)a+b \ge 1$, minimise $\alpha b + \beta a$. Trivially we can assume the minimum occurs for $(k-1)a+b = 1$, and since $k \ge 2$ and $\alpha > \beta$, the unique minimum occurs when $a = b = \frac{1}{k}$.

Thus we have $\sum_{v \in V(H)} w(v)y(v) \ge (\alpha+\beta) \cdot \frac{1}{k}$ for any feasible solution y to the dual program, so the value of the dual program is at least $(\alpha+\beta) \cdot \frac{1}{k}$. By the Duality Theorem for linear programming, the value of the primal program is the same. Finally since $w(v) \in [0,1]$ we have $\sum_{v \in V(H)} w(v) \le \alpha + \beta + \varepsilon m$, and the lemma follows. $\square$

4.2.7 Reservoir path

Definition 4.2.20 (Reservoir path)**.** *A* reservoir path P_{res} *with a* reservoir set $R \subsetneq V(P_{\text{res}})$ *is an k-uniform hypergraph with two $(k-1)$-tuples* $\mathbf{v}$ *and* $\mathbf{w}$*, such that for* any $R' \subseteq R$, P_{res} *contains a tight path with the vertex set $V(P_{\text{res}}) \setminus R'$ and end-tuples* $\mathbf{v}$ *and* $\mathbf{w}$*.*

Lemma 4.2.21 (Reservoir Lemma [8, Lemma 24])**.** *Given $k \ge 3$, $\gamma > 0$, and $\ell' \in \mathbb{N}$, there exist an integer c, such that for $0 < \varepsilon' \le \frac{1}{4}\gamma$, $0 < d \le \frac{1}{8}\gamma$, large enough t_0, small enough $\nu, \varepsilon_k > 0$, any functions $\varepsilon, f_k, f : \mathbb{N} \to (0,1]$ which tend to zero sufficiently fast and any integers $t_2 \ge t_1 \ge t_0$ the following holds a.a.s. in $\Gamma = G^{(k)}(n,p)$ with $p \ge n^{-1+\gamma}$. Suppose $G \subseteq \Gamma$ with $\delta_{k-1}(G) \ge (\frac{1}{2}+\gamma)pn$, that $(\mathcal{P}_c^*, \mathcal{P}_f^*)$ is a $\big(t_0, t_1, t_2, \varepsilon_k, \varepsilon(t_1), f_k(t_1), f(t_2), p\big)$-strengthened pair for G, and that t is the number of 1-cells in $\mathcal{P}_c^*$. Let $\mathcal{R} = \mathcal{R}_{\varepsilon_k,d}(G; \mathcal{P}_c^*, \mathcal{P}_f^*)$ be the (ε_k, d)-reduced multicomplex of G and let S be the union of the 1-cells that are not in $\mathcal{R}$. Then given $R \subseteq V(G)$ with $|R| \le \nu n$ there exists a reservoir path P_{res} in G with reservoir set R and ends* $\mathbf{v}$ *and* $\mathbf{w}$*, such that* $\mathbf{v}, \mathbf{w}$ *are (ε', p, ℓ')-good for $S \cup V(P_{\text{res}})$ and $|V(P_{\text{res}})| \le c|R|$.*

We briefly sketch the proof of this lemma but omit the full proof for the purpose of this thesis (for the proof see [8, Lemma 24]). We fix G and let $R \subseteq V(G)$ be a set of size $r = |R| \le \nu n$. For every $u \in R$ we need a reservoir path P_u with reservoir set $\{u\}$ on a constant number of vertices with end-tuples $\mathbf{v}_u$ and $\mathbf{w}_u$; this is a tight path with end-tuples $\mathbf{v}_u$ and $\mathbf{w}_u$ and vertex set $V(P_u)$ such that there also exists a tight path with the same end-tuples and vertex set $V(P_u) \setminus \{u\}$.

To build P_{res} we begin with an arbitrary $(k-1)$-tuple $\mathbf{v} = \mathbf{w}$, which is a reservoir path with an empty reservoir set, and call this P_0. Assume we have built a reservoir path P_{i-1} with reservoir set $R' \subseteq R$ of size $i-1$ and end-tuples $\mathbf{v}_{i-1}$ and $\mathbf{w}_{i-1}$ such that $V(P_{i-1})$ does not intersect $R \setminus R'$. Then, for some $u \in R \setminus R'$, we construct a reservoir path P_u with end-tuples $\mathbf{v}_u$ and $\mathbf{w}_u$ that is disjoint from P_{i-1}. If i is odd we connect $\mathbf{w}_{i-1}$ to $\mathbf{v}_u$ by a tight path (using Lemma 4.2.16) and let $\mathbf{w}_i = \mathbf{w}_u$ and $\mathbf{v}_i = \mathbf{v}$. If $i-1$ is even we connect $\mathbf{v}_{i-1}$ to $\mathbf{w}_u$ by a tight path and let $\mathbf{v}_i = \mathbf{v}_u$ and $\mathbf{w}_i = \mathbf{w}$. In both cases we obtain a reservoir path P_i with reservoir set $R' \cup \{u\}$, end-tuples $\mathbf{v}_i$ and $\mathbf{w}_i$, and continue. By alternating between the endpoints we ensure that the end-tuples are always (ε', p, ℓ')-good for $V(P'_{\text{res}})$.

Finally, let us sketch how we construct P_u, a picture of which (for $k = 5$) is in Figure 4.1. We begin by finding a $(2k-1)$-vertex tight path with u its central vertex; this gives the spikes $\mathbf{u}$ and $\mathbf{x}_1$ in the figure. We look at all the ways to fill in the upper and lower spike paths in the figure. Using Lemma 4.2.10, we see that from each we can get to a positive density of end-tuples. In particular, we can get to a positive density of each of two vertex-disjoint coarse $(k-1)$-cells in the regular partition, and two applications of Lemma 4.2.18 gives us the tuple $\mathbf{v}$ connecting the paths, completing the spikes. We then use Lemma 4.2.16 repeatedly to create the paths between pairs of spikes. The only point where we need to be a bit careful is to ensure that when creating the upper and lower spike paths, and when connecting them, we do not reuse vertices; for this purpose we randomly split the vertex set into three parts and use one for each of the upper spike path, the lower spike path, and the connection.

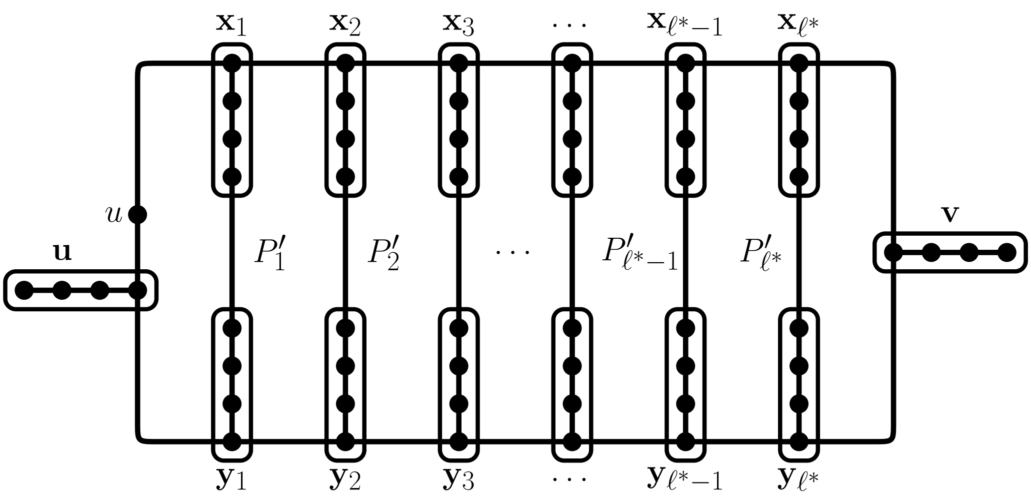

Figure 4.1: Reservoir structure in the case $k = 5$ with $\ell^* = \ell/(k-1)$ for one vertex u with two tight paths that both have end-tuples $\mathbf{u}$ and $\mathbf{v}$, where one is using all vertices and the other all but u.

4.3 Proof of Theorem 1.3.1

Proof. Given $\gamma > 0$ and $k \geq 3$, let $\ell_{\text{L4.2.17}} \geq \frac{k-1}{\gamma} + k$ be returned by Lemma 4.2.17 for input k, 0, and γ. Similarly, let $\ell_{\text{L4.2.16}}$ be given by Lemma 4.2.16 with input k and $\frac{1}{2}\gamma$. Let $\nu_{\text{C4.2.15}}$ be returned by Corollary 4.2.15 for input γ, $\varepsilon = \frac{1}{4}\gamma$, $s = k$, k, and $\ell_{\text{L4.2.17}}$ and let $\nu_{\text{L4.2.17}} = \frac{1}{4}\nu_{\text{C4.2.15}}$. Let c be the integer returned by Lemma 4.2.21 for input γ, k and $\ell_{\text{L4.2.17}}$ and then let $\frac{1}{8}\gamma \geq d > 0$. Let $t_0 \geq k!\nu_{\text{C4.2.15}}^{-2}$ be sufficiently large for Lemma 4.2.21 with input as above, $\varepsilon' = \frac{1}{4}\gamma$ and d, for Lemma 4.2.16 with input as above and $\eta = \frac{1}{2}$, $\varepsilon' = \frac{1}{8}\gamma$, and $s = 3\ell_{\text{L4.2.16}} + 3k$, and for Lemma 4.2.5 with input k and d.

We then choose $\nu_{\text{res}} < \frac{\gamma}{8c}$ such that $2\nu_{\text{res}}$ is sufficiently small for Lemma 4.2.21 with the given input and $\nu_{\text{L4.2.16}} > 0$ is small enough for Lemma 4.2.16 with the given input. We let $\eta_{\text{L4.2.17}} = 10^{-6}\ell_{\text{L4.2.16}}^{-1}\nu_{\text{res}}\nu_{\text{L4.2.16}}$. Next we choose $\varepsilon_k \leq 10^{-6}k^{-k}\nu_{\text{C4.2.15}}^k\nu_{\text{res}}^k\eta_{\text{L4.2.17}}^k$ small enough for Lemma 4.2.17 with input as above and $s = \ell_{\text{L4.2.17}}$, d, $\eta_{\text{L4.2.17}}$, $\nu_{\text{L4.2.17}}$, and $\varepsilon' = \frac{1}{8}\gamma$, for Lemma 4.2.21 with input as above, and also such that $2\nu_{\text{res}}^{-k}\varepsilon_k$ is small enough for Lemma 4.2.16 with input as above.

We choose functions $\varepsilon, f_k, f : \mathbb{N} \to (0,1]$ such that $\sqrt{\varepsilon}$, $2\nu_{\mathrm{res}}^{-k} f_k$ and $\sqrt{f}$ are all smaller than ε_k, small enough for each of Lemmas 4.2.16, 4.2.17 and 4.2.21 with the above inputs, and that for each t, both $\varepsilon(t)$ and $f(t)$ are small enough for Lemma 4.2.6 with input k, $\alpha = \frac{1}{2}\nu_{\mathrm{res}}$ and $d_0 = \frac{1}{2t}$. Let $\eta_{\mathrm{L4.2.4}}$ and $T_{\mathrm{L4.2.4}}$ be returned by Lemma 4.2.4 for input k, t_0, $s = 1$, ε_k, ε, f_k, f.

Given n, let $p \geq n^{-1+\gamma}$. Let $\tilde{L}$ be a set of at most $T_{\mathrm{L4.2.4}}! - 1$ vertices in $[n]$ such that $n - |\tilde{L}|$ is divisible by $T_{\mathrm{L4.2.4}}!$. Suppose that $\tilde{\Gamma} = G^{(k)}(n,p)$ and its induced subgraph $\Gamma = \tilde{\Gamma} - \tilde{L}$ are in the good events of Corollary 4.2.15, Lemmas 4.2.17 and 4.2.21 with inputs as above and Lemma 4.2.9 with input $\frac{1}{4}\gamma\nu_{\mathrm{res}}$ and k. Suppose that Γ and all its subgraphs are $(\eta_{\mathrm{L4.2.4}}, p)$-upper regular, which by Lemma 4.2.8 holds a.a.s. In addition, suppose that Γ satisfies the following: if R is a set of vertices chosen independently with probability ν_{res} from $V(\Gamma)$, then a.a.s. $\Gamma[R]$ is in the good event of Lemma 4.2.16 with input as above. Note that this last event occurs a.a.s. for the following reason: if we first choose R randomly then expose the edges of Γ, a.a.s. we obtain a set R of size $\left(1 \pm \frac{1}{2}\right)\nu_{\mathrm{res}} n$, and given this Lemma 4.2.16 states that a.a.s. $\Gamma[R]$ will be in the good event. Thus the probability of obtaining a pair (R, Γ) such that $\Gamma[R]$ is not in the good event of Lemma 4.2.16 is $o(1)$, and it follows that, for any $\iota > 0$, the probability of choosing Γ such that $\Big($ $\Gamma[R]$ has probability at least ι of not being in the good event of Lemma 4.2.16 $\Big)$, is $o(1)$.

Given $\tilde{G} \subseteq \Gamma$ with $\delta_{k-1}(\tilde{G}) \geq \left(\frac{1}{2} + 2\gamma\right)pn$, we remove $\tilde{L}$ to obtain an induced subgraph G with $T_{\mathrm{L4.2.4}}! | v(G)$. Observe that $\delta_{k-1}(G) \geq \left(\frac{1}{2} + \gamma\right)pn$. We apply Lemma 4.2.4 to G, with input as above, to obtain a $\big(t_0, t_1, t_2, \varepsilon_k, \varepsilon(t_1), f_k(t_1), f(t_2), p\big)$-strengthened pair $(\mathcal{P}_c^*, \mathcal{P}_f^*)$, where $t_0 \leq t_1 \leq t_2 \leq T_{\mathrm{L4.2.4}}$. Let t be the number of clusters of $\mathcal{P}_c$; by definition we have $t_0 \leq t \leq t_1$. Applying Lemma 4.2.5, we see that the (ε_k, d)-reduced multicomplex $\mathcal{R}$ of G, with respect to this strengthened pair, has at least $\left(1 - 4\varepsilon_k^{1/k}\right)t$ 1-edges, every $(k-1)$-edge of $\mathcal{R}$ is contained in at least

$$\left(\tfrac{1}{2} + \gamma - 2d - 2^{k+2}\varepsilon_k^{1/k}\right)t \prod_{i=2}^{k-1} d_i^{-\binom{k-1}{i-1}} \geq \left(\tfrac{1}{2} + \tfrac{1}{2}\gamma\right)t \prod_{i=2}^{k-1} d_i^{-\binom{k-1}{i-1}}$$

k-edges, and every induced subcomplex of $\mathcal{R}$ on at least $\Big(1-\gamma+2d+2^{k+2}\varepsilon_k^{1/k}\Big)t < \Big(1-\frac{1}{2}\gamma\Big)t$ vertices is tightly linked.

We choose a subset R of $[n]$ by selecting vertices uniformly at random with probability ν_{res}. A.a.s. we have $|R| = \Big(1+o(1)\Big)\nu_{\mathrm{res}}n$. By Chernoff's inequality and the union bound, a.a.s. for each V which is a part of either $\mathcal{P}_c$ or $\mathcal{P}_f$, we have $|V \cap R| = \Big(1 \pm o(1)\Big)\nu_{\mathrm{res}}|V|$. Furthermore, for each S which is the neighbourhood in $\tilde{G}$ or in Γ of some $(k-1)$-set of vertices, we have $|S \cap R| = \Big(1 \pm o(1)\Big)\nu_{\mathrm{res}}|S|$ (recall that any such set S has size at least $\frac{1}{2}pn \geq n^{\gamma/2}$). Finally, by our assumption on Γ, we have a.a.s. that $\Gamma[R]$ is in the good event of Lemma 4.2.16 with input as above. Suppose that R is such that all of these likely events occur.

We apply Lemma 4.2.21, with inputs as above, to find a reservoir path P_{res} in G with reservoir set R whose ends are $\mathbf{v}_{\mathrm{res}}$ and $\mathbf{w}_{\mathrm{res}}$, such that $\mathbf{v}_{\mathrm{res}}$ and $\mathbf{w}_{\mathrm{res}}$ are both $(\frac{1}{4}\gamma, p, \ell_{\mathrm{L4.2.17}})$-good for $S \cup V(P_{\mathrm{res}})$, where S is the union of all 1-cells not in $\mathcal{R}$, and such that $\Big|V(P_{\mathrm{res}})\Big| \leq c|R| \leq \frac{1}{8}\gamma n$.

We now aim to extend P_{res}, from its end $\mathbf{w}_{\mathrm{res}}$, to a path P_{almost} covering almost all vertices of G. To begin with, let $\mathcal{R}'$ denote the complex on $V(\mathcal{R})$ obtained by letting e' be an edge of $\mathcal{R}'$ whenever there is an edge e of $\mathcal{R}$ such that vertices$(e) = e'$. Thus the 1-edges of $\mathcal{R}$ and $\mathcal{R}'$ are identical, and it follows inductively from the definition of an (ε_k, d)-reduced multicomplex that for each $1 \leq i \leq k-2$, each i-edge of $\mathcal{R}'$ is contained in at least $\Big(1 - 2^{i+2}\varepsilon_k^{1/k}\Big)t$ $(i+1)$-edges, and each $(k-1)$-edge of $\mathcal{R}'$ is contained in at least $\Big(\frac{1}{2} + \frac{1}{2}\gamma\Big)t$ k-edges. We define a weight function ω on $V(\mathcal{R}')$ as follows. Given a cluster $V_i \in V(\mathcal{R}')$, if $|V_i \setminus V(P_{\mathrm{res}})| < 2\eta_{\mathrm{L4.2.17}}\frac{n}{t}$, we set $\omega(V_i) = 0$. Otherwise, we set

$$\omega(V_i) = \frac{|V_i \setminus V(P_{\mathrm{res}})| - 2\eta_{\mathrm{L4.2.17}}\frac{n}{t}}{(1-\nu_{\mathrm{res}})\frac{n}{t}}\,.$$

Note that since $|V_i \setminus V(P_{\mathrm{res}})| \leq |V_i \setminus R| \leq \Big(1+o(1)\Big)(1-\nu_{\mathrm{res}})\frac{n}{t}$, this weight function takes values in $[0,1]$. Furthermore, we have

$$\sum_{V_i \in V(\mathcal{R}')} \omega(V_i) = \frac{n - |V(P_{\text{res}})| - 2\eta_{\text{L4.2.17}} n}{(1-\nu_{\text{res}})\frac{n}{t}} > \left(1 - \tfrac{1}{2}\gamma\right) t \,.$$

This is the required setup to apply Lemma 4.2.19, with input $2^{k+2}\varepsilon_k^{1/k}$ and $\frac{1}{2}\gamma$. The result is a weight function $\omega^* : E(\mathcal{R}'^{(k)}) \to [0,1]$ such that for each $V_i \in V(\mathcal{R}')$ we have $\sum_{e \ni V_i} \omega^*(e) \le \omega(V_i)$, and

$$\begin{aligned}\sum_{e \in \mathcal{R}'^{(k)}} \omega^*(e) &\ge \tfrac{1}{k}\Big(\sum_{V_i \in V(\mathcal{R}')} \omega(V_i) - 2^{k+2}\varepsilon_k^{1/k} t \Big) > \tfrac{1}{k} \cdot \frac{n - |V(P_{\text{res}})| - 2\eta_{\text{L4.2.17}} n - 2^{k+3}\varepsilon_k^{1/k} n}{(1-\nu_{\text{res}})\frac{n}{t}} \\ &> \tfrac{1}{k} \cdot \frac{n - |V(P_{\text{res}})| - 3\eta_{\text{L4.2.17}} n}{(1-\nu_{\text{res}})\frac{n}{t}} \,. \end{aligned} \tag{4.3.1}$$

Recall that $\mathcal{R}$ is tightly linked, even if an arbitrary set of $\frac{1}{2}\gamma t$ vertices is removed. If a cluster of $\mathcal{P}_c$ has ω-weight zero, then it contains at least $\frac{n}{2t}$ vertices of P_{res}, so there are at most $\frac{2c\nu_{\text{res}} n \cdot 2t}{n} = 4c\nu_{\text{res}} t \le \frac{1}{2}\gamma t$ clusters with ω-weight zero. In particular, the submulticomplex of $\mathcal{R}$ induced by removing clusters of ω-weight zero is tightly linked.

We next construct a path P_{almost} extending P_{res} from $\mathbf{w}_{\text{res}}$ as follows. Recall that $\mathbf{w}_{\text{res}}$ is $\left(\frac{1}{4}\gamma, p, \ell_{\text{L4.2.17}}\right)$-good for $S \cup V(P_{\text{res}})$. To begin with, we use Corollary 4.2.15 to obtain a collection of $(k-1)$-tuples, of size at least $\nu_{\text{C4.2.15}} n^{k-1}$, each of which is the end-tuple of a path of length $\ell_{\text{L4.2.17}}$ starting at $\mathbf{w}_{\text{res}}$ whose vertices, other than those in $\mathbf{w}_{\text{res}}$, are disjoint from $V(P_{\text{res}})$. Note that all these tuples are by construction outside $V(P_{\text{res}})$ and so also outside R. By definition of a strengthened pair and (ε_k, d)-reduced multicomplex, and choice of t_0 and ε_k, at least half of these end-tuples are contained in $(k-1)$-cells of $\mathcal{P}_c^*$ which are in $\mathcal{R}$. In particular, by averaging there is a $(k-1)$-cell of $\mathcal{R}$, with clusters in a given order, f_0, such that at least a $\frac{1}{2}\nu_{\text{C4.2.15}}$-fraction of these end-tuples are in f_0 in the given order. Let $P_0 = P_{\text{res}}$, and let Q_0 denote the set of $(k-1)$-tuples in f_0 which are ends of paths of length $\ell_{\text{L4.2.17}}$ starting from $\mathbf{w}_0 := \mathbf{w}_{\text{res}}$ whose vertices outside $\mathbf{w}_0$ are disjoint from P_0.

We order arbitrarily the k-edges of $\mathcal{R}'$ with positive ω^*-weight, and for each j, let g_j be a k-edge of $\mathcal{R}$ whose vertices are the same as the jth k-edge of $\mathcal{R}'$; we let $\omega^*(g_j)$ be given by ω^* at the jth edge of $\mathcal{R}'$. We now create a sequence $e_1, \dots$ of k-edges of $\mathcal{R}$ as follows. To begin with, we choose a tight link in $\mathcal{R}$ from f_0 to a $(k-1)$-tuple in g_1 using only clusters of positive weight, and we let the first edges $e_1, \dots$ be the edges of a homomorphic copy of a minimum length tight path following this tight link. We then repeat g_1 in the sequence

$$\left\lceil \frac{k(1-\nu_{\mathrm{res}})\frac{n}{t} \cdot \omega^*(g_1)}{\ell_{\text{L4.2.17}}} \right\rceil$$

times, follow a tight link to g_2, and so on. When we follow a tight link, we always do so such that the edges $e_1, \dots$ form a homomorphic copy of a tight path in $\mathcal{R}$, using only vertices whose weight according to ω is positive; this is possible since the vertices of each g_j have weight at least $\omega^*(g_j) > 0$, and since the positive-weight induced submulticomplex of $\mathcal{R}$ is tightly linked. Note that the number of repetitions of g_1 fixes the ordered $(k-1)$-cell in the boundary of g_1 from which we follow a tight link to g_2, and so on.

Since $\mathcal{R}$ is a bounded size multicomplex — it contains in total at most $t_1^k \cdot t_1^{\binom{k}{2}} \dots t_1^{\binom{k}{k-1}} \leq t_1^{2^k}$ edges of size k — the total number of edges e_i used in following tight links is at most $4k^3 \cdot t_1^{2^{k+1}}$.

We now use the following procedure repeatedly for $i \geq 1$. We are given P_{i-1} which is a path from $\mathbf{v}_{\mathrm{res}}$ to $\mathbf{w}_{i-1}$, an ordered $(k-1)$-cell f_{i-1} of $\mathcal{R}$ (which is contained in e_{i-1} and also in e_i), and a set Q_{i-1} of $(k-1)$-tuples in f_{i-1} which are ends of paths of length $\ell_{\text{L4.2.17}}$ from $\mathbf{w}_{i-1}$ whose vertices outside $\mathbf{w}_{i-1}$ are disjoint from P_{i-1}. We suppose Q_{i-1} contains at least a $2\nu_{\text{L4.2.17}}$-fraction of the $(k-1)$-tuples in f_{i-1}. We let f_i be the $(k-1)$-cell in the boundary of e_i on the last $k-1$ clusters of e_i, with the order inherited from e_i.

By Lemma 4.2.17, with input as above, and $S = V(P_{i-1})$, and choice of $\nu_{\text{L4.2.17}}$ there is a tuple $\mathbf{w}_i$ in Q_{i-1} such that the following holds. Let P_i denote the extension of P_{i-1} to $\mathbf{w}_i$ by adding a path of length $\ell_{\text{L4.2.17}}$ witnessing $\mathbf{w}_i \in Q_{i-1}$; let S' be the vertices $V(P_i) \setminus V(P_{i-1})$. There is a set Q_i of $(1-\nu_{\text{L4.2.17}})$-fraction of the tuples of f_i, each of which is the end of a path of length $\ell_{\text{L4.2.17}}$ from $\mathbf{w}_i$, whose vertices outside $\mathbf{w}_i$ are disjoint from $V(P_{i-1})$ and from S'. Note that this is the setup required to iterate the application of Lemma 4.2.17, provided that we ensure that at no stage does $S = V(P_{i-1})$ intersect any cluster of e_i in more than a $(1 - \eta_{\mathrm{L4.2.17}})$-fraction. This is guaranteed for the following reason. Given a cluster V_j of $\mathcal{P}_c$, if $\omega(V_j) = 0$ then V_j is not a vertex of any e_i. If on the other hand $\omega(V_j) > 0$, then the total number of vertices used in V_j is at most

$$\ell_{\mathrm{L4.2.17}} \cdot 4k^3 \cdot t_1^{2^{k+1}} + 2\ell_{\mathrm{L4.2.17}} \cdot \binom{t_1}{k-1} + \tfrac{\ell_{\mathrm{L4.2.17}}}{k} \cdot \frac{k(1-\nu_{\mathrm{res}})\frac{n}{t}}{\ell_{\mathrm{L4.2.17}}} \cdot \sum_{g_i \ni V_j} \omega^*(g_i)\,,$$

where the first term counts vertices used in following tight links, the second accounts for the rounding up in the weighting at each edge g_i containing V_j and the (at most) one vertex per g_i extra since the tight path may use one more vertex in some clusters than others (since $\frac{\ell_{\mathrm{L4.2.17}}}{k}$ may not be an integer). Note that these first two terms are bounded from above by a constant. Since $\sum_{g_i \ni V_j} \omega^*(g_i) \le \omega(V_j)$, we see that the number of vertices used in V_j is at most

$$O(1)+\tfrac{\ell_{\mathrm{L4.2.17}}}{k}\cdot\frac{k(1-\nu_{\mathrm{res}})\frac{n}{t}}{\ell_{\mathrm{L4.2.17}}}\cdot\omega(V_j) = O(1)+|V_j\setminus V(P_{\mathrm{res}})|-2\eta_{\mathrm{L4.2.17}}\tfrac{n}{t} \le \Big|V_j\setminus V(P_{\mathrm{res}})\Big|-\eta_{\mathrm{L4.2.17}}\tfrac{n}{t}\,,$$

and in particular at all times at least $\eta_{\mathrm{L4.2.17}}\frac{n}{t}$ vertices remain in V_j. We let P_{almost} denote the final tight path from $\mathbf{v}_{\mathrm{res}}$ to $\mathbf{w}_{\mathrm{alm}}$ obtained by this procedure.

Observe that, just counting repetitions of the g_i, the total number of vertices $\Big|V(P_{\mathrm{almost}})\setminus V(P_{\mathrm{res}})\Big|$ is at least

$$\begin{aligned} \ell_{\mathrm{L4.2.17}} \cdot \tfrac{k(1-\nu_{\mathrm{res}})\frac{n}{t}}{\ell_{\mathrm{L4.2.17}}} \cdot \sum_{e\in\mathcal{R}'^{(k)}} \omega^*(e) &> k(1-\nu_{\mathrm{res}})\tfrac{n}{t} \cdot \tfrac{1}{k} \cdot \frac{n - |V(P_{\mathrm{res}})| - 3\eta_{\mathrm{L4.2.17}}n}{(1-\nu_{\mathrm{res}})\frac{n}{t}} \\ &= n - |V(P_{\mathrm{res}})| - 3\eta_{\mathrm{L4.2.17}}n\,. \end{aligned} \tag{4.3.2}$$

It follows that $n - \Big|V(P_{\mathrm{almost}})\Big| \le 3\eta_{\mathrm{L4.2.17}}n$. Let $L = \Big(V(G)\setminus V(P_{\mathrm{almost}})\Big)\cup\tilde{L}$. Recall that $\tilde{L}$ is the set of at most $T_{\mathrm{L4.2.4}}! - 1$ vertices we removed from $\tilde{G}$ in order to guarantee the required divisibility condition.

Our final task is to extend P_{almost}, re-using some vertices of R, to cover the vertices of L

and connect the ends. Critically, observe that $|L|$ is much smaller than $|R|$, and that by assumption on Γ and R, the good event of Lemma 4.2.16 holds for $\Gamma[R]$, for the input given at the start of the proof. Recall that $G[R]$ has minimum codegree at least $\left(\frac{1}{2}+\frac{1}{2}\gamma\right)p|R|$. Let $\mathcal{P}^*_{cr}$ and $\mathcal{P}^*_{fr}$ denote the families of partitions obtained from $\mathcal{P}^*_c$ and $\mathcal{P}^*_f$ respectively by reducing each cell to only those elements contained in R. By Lemma 4.2.6 and choice of ε_k, ε, f_k and f, $(\mathcal{P}^*_{cr}, \mathcal{P}^*_{fr})$ is a $(t_0, t_1, t_2, 2\nu_{\text{res}}^{-k}\varepsilon_k, \sqrt{\varepsilon(t_1)}, 2\nu_{\text{res}}^{-k}f_k, \sqrt{f(t_2)}, p)$-strengthened pair for $G[R]$. Let $\mathcal{R}_r$ denote the multicomplex obtained from $\mathcal{R}$ by replacing the cells of $\mathcal{P}^*_c$ with those of $\mathcal{P}^*_{cr}$. Note that $\mathcal{R}_r$ is still the (ε_k, d)-reduced multicomplex for this strengthened pair, so it is contained in the $(2\nu_{\text{res}}^{-k}\varepsilon_k, d)$-reduced multicomplex.

Let $S_{-1}=\emptyset$. We now construct for $i=0,1,\dots$ two disjoint tight paths $P_{v,i}$ and $P_{w,i}$ and $S_i = V(P_{v,i}) \cup V(P_{v,i})$, where one end of $P_{v,i}$ is $\overleftarrow{\mathbf{v}_{\text{res}}}$ and the other, $\mathbf{v}_i$, is $(\frac{1}{4}\gamma\nu_{\text{res}}, p, \ell_{\text{L4.2.16}})$-good for S_{i-1}, and $P_{v,i}$ contains i vertices of L and all other vertices, except those of $\mathbf{v}_{\text{res}}$, are in R. Similarly one end of $P_{w,i}$ is $\mathbf{w}_{\text{alm}}$ and the other, $\mathbf{w}_i$, is $(\frac{1}{4}\gamma\nu_{\text{res}}, p, \ell_{\text{L4.2.16}})$-good for S_{i-1}, and $P_{w,i}$ contains i vertices of L, not in $P_{v,i}$, and all other vertices, except those of $\mathbf{w}_{\text{alm}}$, are in R. We do this as follows. To begin with, we find a tight path $P_{v,0}$ of length $k-1$, one of whose end tuples is $\overleftarrow{\mathbf{v}_{\text{res}}}$ and the other of which, $\mathbf{v}_0$, is contained in R. Recall that every $(k-1)$-set in $V(G)$ is contained in at least $\left(\frac{1}{2}+\frac{1}{2}\gamma\right)p|R|$ edges of size k with the extra vertex in R, so in particular we can greedily build the required path of length $k-1$. We construct $P_{w,0}$ from $\mathbf{w}_{\text{alm}}$ to $\mathbf{w}_0$ similarly. Observe that, by definition, both $\mathbf{v}_0$ and $\mathbf{w}_0$ are $(\frac{1}{4}\gamma\nu_{\text{res}}, p, \ell_{\text{L4.2.16}})$-good for S_{-1}.

Now suppose $i \geq 1$ and that we have constructed tight paths $P_{v,i-1}$ and $P_{w,i-1}$ as above, whose ends $\mathbf{v}_{i-1}$ and $\mathbf{w}_{i-1}$ are both $(\frac{1}{4}\gamma\nu_{\text{res}}, p, \ell_{\text{L4.2.16}})$-good for S_{i-2}, and we have $|S_{i-1}| \leq 4(i-1)\ell_{\text{L4.2.16}}$. We first extend $P_{v,i-1}$ to a path $P_{v,i}$ as follows. We choose any $u \in L \setminus S_{i-1}$ and vertices $v_1,\dots,v_{k-2}$ from $R \setminus S_{i-1}$ such that the tuple $(u, v_1, \dots, v_{k-2})$ is $(\frac{1}{4}\gamma\nu_{\text{res}}, p, \ell_{\text{L4.2.16}}+k-1)$-good for S_{i-1}. This step always succeeds, as $o(n)$ of these tuples are not $(\frac{1}{4}\gamma\nu_{\text{res}}, p, \ell_{\text{L4.2.16}}+k-1)$-good for S_{i-1}, by the good event of Lemma 4.2.9 assumed above. Then we can easily choose additional vertices $v_{k-1}, u_1, \dots, u_{k-1}$ from $R \setminus S_{i-1}$ such that for $j=1,\dots,k$ there is a k-edge $\{u_j,\dots,u_{k-1},u,v_1,\dots,v_{j-1}\}$ and the tuples $\mathbf{v}_i = (v_1, \dots, v_{k-1})$ and $\mathbf{u} = (u_1, \dots, u_{k-1})$ are $(\frac{1}{4}\gamma\nu_{\text{res}}, p, \ell_{\text{L4.2.16}})$-good for S_{i-1}. For example, there are at least $(\frac{1}{2} + \frac{1}{4}\gamma)p|R|$ edges $\{u, v_1, \dots, v_{k-1}\}$ in G with $v_{k-1} \in R$, of which at most $p|S_{i-1}| + \frac{1}{4}\gamma\nu_{\text{res}}pn \le \frac{1}{4}p|R|$ have $v_{k-1} \in S_{i-1}$ and at most $\frac{1}{4}\gamma\nu_{\text{res}}pn$ are such that $(v_1, \dots, v_{k-1})$ is not $(\frac{1}{4}\gamma\nu_{\text{res}}, p, \ell_{\text{L4.2.16}} + k - 2)$-good for S_{i-1}.

Next, with Lemma 4.2.16, we connect $\mathbf{v}_{i-1}$ to $\mathbf{u}$ with a tight path of length $\ell_{\text{L4.2.16}}$ and internal vertices not in S_{i-1}. Note that here we added the set S' containing the vertices $V(P_{v,i-1}) \setminus V(P_{v,i-2})$, $V(P_{w,i-1}) \setminus V(P_{w,i-2})$, and $\{u, v_1, \dots, v_{k-1}\}$ and that $|S'| \le 2\ell_{\text{L4.2.16}} + 2k$. To see that the conditions of Lemma 4.2.16 are satisfied, recall that $|S_{i-1}| \le 4(i-1)\ell_{\text{L4.2.16}} \le 6\eta_{\text{L4.2.17}}\ell_{\text{L4.2.16}}n$. By the choice of $\eta_{\text{L4.2.17}}$, this is at most $\frac{1}{4}|R|$ and there can bet at most $\nu_{\text{L4.2.16}}t$ 1-cells in $\mathcal{R}_r$ which intersect S_{i-1} in at least a $\frac{1}{2}$-fraction. We then let $P_{v,i}$ be the path obtained by concatenating $P_{v,i-1}$, the path from $\mathbf{v}_{i-1}$ to $\mathbf{u}$, and the path from $\mathbf{u}$ via u to $\mathbf{v}_i$.

If there remain uncovered vertices in L, we repeat the same procedure to extend $P_{w,i-1}$ to $P_{w,i}$, where in the last step we also add the vertices from $V(P_{v,i}) \setminus V(P_{v,i-1})$ to S' and get $|S'| \le 3\ell_{\text{L4.2.16}} + 3k$. Note that afterwards with $S_i = V(P_{v,i}) \cup V(P_{w,i})$ we have $|S_i| \le 4i\ell_{\text{L4.2.16}}$ and all conditions of $P_{v,i}$ and $P_{w,i}$ needed for the next iterations are satisfied. We stop this procedure as soon as all vertices of L are used; we let P_v denote the final $P_{v,i}$ with end tuple $\mathbf{v} = \mathbf{v}_i$, and P_w denote either $P_{w,i}$ or $P_{w,i-1}$ (depending on whether $|L|$ is even or odd, respectively) with end tuple $\mathbf{w}$ either $\mathbf{w}_i$ or $\mathbf{w}_{i-1}$, respectively. Finally we make a last use of Lemma 4.2.16 to find a tight path in R whose interior vertices are disjoint from $V(P_v) \cup V(P_w)$, and whose ends are $\overleftarrow{\mathbf{v}}$ and $\mathbf{w}$. This is possible for the same reasons as above. Concatenating these three tight paths, we obtain a tight path P_{cover} whose end tuples are $\mathbf{w}_{\text{alm}}$ and $\mathbf{v}_{\text{res}}$, such that $L \subseteq V(P_{\text{cover}})$, and such that all interior vertices of P_{cover} are contained in $L \cup R$.

Let R' denote the set of vertices $V(P_{\text{cover}}) \cap R$. By the reservoir property of P_{res}, there is a tight path P^*_{res} whose end tuples are identical to P_{res} and whose vertex set is $V(P_{\text{res}}) \setminus R'$. We replace P_{res} with P^*_{res} in P_{almost} to obtain a tight path P^*_{almost} whose end tuples are identical to those of P_{almost} and whose vertex set is $V(P_{\text{almost}}) \setminus R' = V(\tilde{G}) \setminus (L \cup R')$.

Concatenating P^*_{almost} and P_{cover}, we obtain the desired tight Hamilton cycle in $\tilde{G}$. $\square$

CHAPTER 5

1-INDEPENDENT PERCOLATION ON $\mathbb{Z}^2 \times K_N$

The main aim of this chapter is to prove Theorems 1.4.8, 1.4.9, 1.4.11 and 1.4.12. We recall the following definitions and then restate the theorems. Let $H = (V, E)$ be a graph. Given a probability measure μ on subsets of E, a *μ-random graph* $\mathbf{H}_\mu$ is a random spanning subgraph of H whose edge-set is chosen randomly from subsets of E according to the law given by μ. A random graph model $\mathbf{H}_\mu$ is 1-*independent* if, for vertex-disjoint subsets $A, B \subseteq E(H)$, the random variables $\mathbf{H}_\mu \cap A$ and $\mathbf{H}_\mu \cap B$ are independent. In this case, we call the associated probability measure μ a 1-*independent measure (*1-*imp)* on H. For $p \in [0,1]$, let $\mathcal{M}_{1,p}$ and $\mathcal{M}_{1,\geq p}$, be the sets of 1-independent measures on H for which each edge of H is present in $\mathbf{H}_\mu$ with probability exactly p and with probability at least p, respectively. Given a 1-ipm μ on an infinite connected graph H, we say that μ *percolates* if $\mathbf{H}_\mu$ almost surely (i.e. with probability 1) contains an infinite connected component. Given an infinite connected graph H, we define the 1-*independent critical percolation probability* for H to be $p_{1,c}(H) := \inf \{p \geq 0 : \ \forall \mu \in \mathcal{M}_{1,\geq p}(H), \ \mu \textit{ percolates}\}$. Moreover, we define the *long paths critical probability* $p_{1,\mathrm{LP}}(H)$ of H, to be given by

$$p_{1,\mathrm{LP}}(H) := \inf \{p \in [0,1] : \ \forall \mu \in \mathcal{M}_{1,p}, \forall \ell \in \mathbb{N}, \ \mathbb{P}[\mathbf{H}_\mu \text{ contains a path of length } \ell] > 0\}.$$

In this chapter we will prove Theorems 1.4.8 and 1.4.9.

Theorem 1.4.8. *The following hold:*

(i) *If* $p > 4 - 2\sqrt{3}$ *is fixed, then there exists* $N \in \mathbb{N}$ *such that* $p_{1,c}\left(\mathbb{Z}^2 \times K_N\right) \leq p$.

(ii) *For every* $n \in \mathbb{N}$, $p_{1,c}\left(\mathbb{Z}^2 \times K_n\right) \geq 4 - 2\sqrt{3}$.

In particular, we have $\lim_{n\to\infty} p_{1,c}(\mathbb{Z}^2 \times K_n) = 4 - 2\sqrt{3} = 0.5358\ldots$.

Theorem 1.4.9. $\lim_{n\to\infty} p_{1,\mathrm{LP}}\left(\mathbb{Z} \times K_n\right) = 4 - 2\sqrt{3}$.

We also prove Theorems 1.4.11 and 1.4.12 which are stronger versions of these theorems where K_n is replaced by a sequence of weakly pseudorandom graphs. Recall that, for a sequence $q = q(n)$ in $[0,1]$, a sequence of n-vertex graphs $(G_n)_{n\in\mathbb{N}}$ is *weakly* q*-pseudorandom* if

$$\max\left\{\left|e(G_n[U]) - q\frac{|U|^2}{2}\right| :\ U \subseteq V(G_n)\right\} = o(qn^2).$$

Theorem 1.4.11. *Let* $q = q(n)$ *satisfy* $nq(n) \gg \log n$*. Then for any sequence* $(G_n)_{n\in\mathbb{N}}$ *of* n*-vertex graphs which is weakly* q*-pseudorandom, we have* $\lim_{n\to\infty} p_{1,c}(\mathbb{Z}^2 \times G_n) = 4 - 2\sqrt{3}$.

Theorem 1.4.12. *Let* $q = q(n)$ *satisfy* $nq(n) \gg \log n$*. Then for any sequence* $(G_n)_{n\in\mathbb{N}}$ *of* n*-vertex graphs which is weakly* q*-pseudorandom, we have* $\lim_{n\to\infty} p_{1,\mathrm{LP}}(\mathbb{Z} \times G_n) = 4 - 2\sqrt{3}$.

The remainder of this chapter is organised as follows. In the next section we explain some basic notation. The key step in the proof of our main results in this chapter, Theorem 5.2.1, is proved in Section 5.2; it establishes that $p = 4 - 2\sqrt{3}$ is the threshold for ensuring there is a high probability in any 1-independent model of finding a path between the largest components in two disjoint copies of K_n joined by a matching. The argument in a sense captures 'what makes the $4 - 2\sqrt{3}$ measure of [36, 46] tick'. We then use Theorem 5.2.1 in Section 5.3 to prove Theorems 1.4.8, 1.4.9, 1.4.11 and 1.4.12. Our component evolution results, Proposition 1.4.16 and Theorem 1.4.17 are proved in Section 5.4.

5.1 Notation

Given $n \in \mathbb{N}$ we write $[n]$ for the discrete interval $\{1, 2, \ldots, n\}$. We write $S^{(2)}$ for the collection of all unordered pairs from a set S. We use standard graph-theoretic notation. Given a graph H, we use $V = V(H)$ and $E = E(H)$ to refer to its vertex-set and edge-set respectively, and write $e(H)$ for the size of $E(H)$. Given $X \subseteq V$, we write $H[X]$ for the subgraph of H induced by X, i.e. the graph $(X, E(H) \cap X^{(2)})$. For disjoint subsets X, Y of V we also write $H[X, Y]$ for the bipartite subgraph of H induced by $X \sqcup Y$, that is the graph $(X \cup Y, \{xy \in E(H) : \ x \in X, y \in Y\})$. We denote by K_n the complete graph on n vertices, $K_n = ([n], [n]^{(2)})$.

The *Cartesian product* of two graphs G_1 and G_2 is the graph $G_1 \times G_2$ with $V(G_1 \times G_2) = \{(v_1, v_2) : \ v_1 \in V(G_1), v_2 \in V(G_2)\}$ and $E(G_1 \times G_2)$ consisting of all pairs $\{(u_1, u_2), (v_1, v_2)\}$ with either $u_1 = v_1 \in V(G_1)$ and $u_2v_2 \in E(G_2)$ or $u_1v_1 \in E(G_1)$ and $u_2 = v_2 \in V(G_2)$. In particular if $G_1 = K_2$, i.e. a single edge, then $G_1 \times G_2$ is the *bunkbed graph* of G_2 consisting of two disjoint copies of G_2, the *left copy* $\{1\} \times G_2$ and the *right copy* $\{2\} \times G_2$, together with a perfect matching joining each vertex $(1, v)$ in the left copy to its image $(2, v)$ in the right copy.

Figure 5.1: The Cartesian product $K_2 \times K_3$

Finally we use the standard Landau notation for asymptotic behaviour: given functions $f, g : \ \mathbb{N} \to \mathbb{R}$, we write $f = O(g)$ if $|f(n)| \leq C|g(n)|$ for some $C > 0$ and all n sufficiently large, and $f = o(g)$ if $\lim_{n\to\infty} |f(n)/g(n)| = 0$. We use $f = \Omega(g)$ and $f = \omega(g)$ to denote $g = O(f)$ and $g = o(f)$, respectively. We also sometimes use $f \ll g$ and $f \gg g$ as a shorthand for $f = o(g)$ and $f = \omega(g)$, respectively. Given a sequence of events $(E_n)_{n\in\mathbb{N}}$ in some probability space, we say that E_n occurs *with high probability (whp)* if $\mathbb{P}[E_n] = 1 - o(1)$.

5.2 When left meets right: joining the largest components on either side of $K_2 \times G_n$

Let $(G_n)_{n\in\mathbb{N}}$ be a sequence of weakly q-pseudorandom n-vertex graphs where $qn \gg \log n$. Consider the Cartesian product $H = K_2 \times G_n$. Given $\mu \in \mathcal{M}_{1,p}(H)$, let 'Left meets Right' denote the event that the μ-random graph $\mathbf{H}_\mu$ contains a connected component containing both strictly more than half of the vertices in $\{1\} \times [n]$ and strictly more than half of the vertices in $\{2\} \times [n]$. Our main result in this section is showing that the event 'Left meets Right' undergoes a sharp transition at $p = 4 - 2\sqrt{3}$, in the sense that for $p \leq 4 - 2\sqrt{3}$ it is possible to construct 1-independent measures $\mu \in \mathcal{M}_{1,p}(H)$ such that whp the event 'Left meets Right' does not occur, while for $p > 4 - 2\sqrt{3}$ it occurs whp regardless of the choice of μ.

Theorem 5.2.1. (i) *Let $p > 4 - 2\sqrt{3}$ be fixed. Then for every $\mu \in \mathcal{M}_{1,p}(H)$,*

$$\mathbb{P}\left[\text{Left meets Right}\right] = 1 - o(1).$$

(ii) *Let $\frac{1}{2} < p \leq 4 - 2\sqrt{3}$ be fixed. Then there exists $\mu \in \mathcal{M}_{1,\geq p}(H)$ such that*

$$\mathbb{P}\left[\text{Left meets Right}\right] = o(1).$$

For $p \in (\frac{1}{2}, 1]$, let $\theta = \theta(p)$ be given by

$$\theta(p) := \frac{1 + \sqrt{2p-1}}{2}.$$

The quantity θ will play an important role in the proof of both parts of Theorem 5.2.1. Observe that $\theta \in [p, 1]$ and satisfies

$$\theta^2 + (1-\theta)^2 = p \qquad \text{and} \qquad 2\theta(1-\theta) = 1 - p.$$

The latter of these relations and the resolution of the quadratic inequality $p^2 - 8p + 4 \geq 0$

for $p \in [0,1]$ can be used to show

$$\theta\sqrt{p} \leq 1 - p \qquad \text{if and only if} \qquad p \leq 4 - 2\sqrt{3}. \tag{5.2.1}$$

Our proofs will also make extensive use of the following Chernoff bound: given a binomial random variable $X \sim \mathrm{Binom}(N,p)$ and $\varepsilon \in (0,1)$, we have

$$\mathbb{P}\left[|X - Np| \geq \varepsilon Np\right] \leq 2e^{-\frac{\varepsilon^2 Np}{3}}. \tag{5.2.2}$$

5.2.1 Lower bound construction: proof of Theorem 5.2.1(ii)

For each $1/2 < p \leq 4 - 2\sqrt{3}$, we construct a state-based measure $\mu_F \in \mathcal{M}_{\geq p}(K_2 \times G_n)$, based on the ideas behind constructions in [36, 46]. Assume without loss of generality that $V(G_n) = [n]$. We randomly assign to each vertex $(i,v) \in [2] \times [n]$ a state $S_v \in \{0, 1, \star\}$, independently of all the other vertices, with

(a) $S_{(1,v)} = 1$ with probability θ and $S_{(1,v)} = 0$ otherwise;

(b) $S_{(2,v)} = 0$ with probability $\sqrt{p}$ and $S_{(2,v)} = \star$ otherwise.

We then include edges of $H = K_2 \times G_n$ in our random subgraph $\mathbf{H}_{\mu_F}$ according to the following rules:

(i) an edge $\{(1,u),(1,v)\}$ is included if $S_{(1,u)} = S_{(1,v)}$;

(ii) an edge $\{(2,u),(2,v)\}$ is included if $S_{(2,u)} = S_{(2,v)} = 0$;

(iii) an edge $\{(1,v),(2,v)\}$ is included if $S_{(2,v)} = \star$ or if $S_{(1,v)} = S_{(2,v)} = 0$.

See Fig. 5.2 for an illustration of the construction. Since μ_F is state-based, it is clearly a 1-ipm. Our state distributions (a)–(b) imply that every edge in the left copy of G_n is open (included in our random graph) with probability $\theta^2 + (1-\theta)^2 = p$ (by the edge-rule (i) above), and that every edge in the right copy of G_n is open with probability $(\sqrt{p})^2 = p$

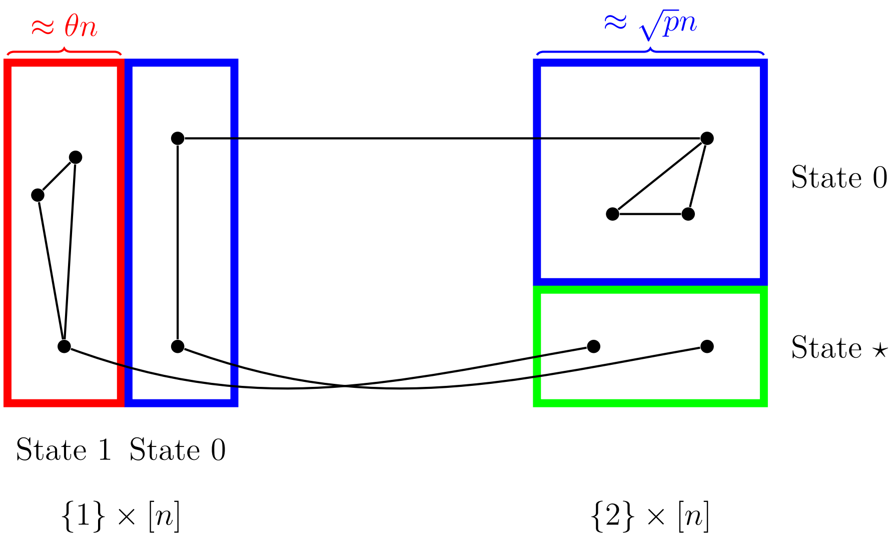

Figure 5.2: The Lower bound construction

(by the edge-rule (ii) above). On the other hand, (by the edge-rule (iii) above) an edge $\{(1,v),(2,v)\}$ from the left copy to the right copy is closed if and only if $S_{(1,v)} = 1$ and $S_{(2,v)} = 0$, which by (5.2.1) occurs with probability $\theta\sqrt{p} \leq 1-p$ provided $p \leq 4-2\sqrt{3}$. Thus $\mu_F \in \mathcal{M}_{1,\geq p}(K_2 \times G_n)$ as claimed.

All that remains to show is that for this measure the event 'Left meets Right' occurs with probability $o(1)$ in the random graph $\mathbf{H}_{\mu_F}$. Observe that the construction of μ_F ensures there is no path in $\mathbf{H}_{\mu_F}$ from the vertices in $\{1\} \times [n]$ in state 1 to the vertices in $\{2\} \times [n]$ in state 0. Indeed the only edges of $\mathbf{H}_{\mu_F}$ in which the endpoints are in different states are those edges containing a vertex $(2,v)$ in state $S_{(2,v)} = \star$. Since by construction vertices in state $\star$ have degree exactly one in $\mathbf{H}_{\mu_F}$, it follows that there is no component of $\mathbf{H}_{\mu_F}$ containing both vertices in state 1 and vertices in state 0.

Since the expected number of vertices in $\{1\} \times [n]$ in state 1 is $\theta n > pn$ and the expected number of vertices in $\{2\} \times [n]$ in state 0 is $\sqrt{p}n > pn$, and since states are assigned independently, it follows from (5.2.2) that for all fixed p with $1/2 < p \leq 4-2\sqrt{3}$, with probability $1-o(1)$ there is no connected component in $\mathbf{H}_{\mu_F}$ containing at least half of the vertices of both $\{1\} \times [n]$ and $\{2\} \times [n]$. Thus 'Left meets Right' occurs with probability $o(1)$ for $\mathbf{H}_{\mu_F}$, as claimed. $\square$

5.2.2 Upper bound: proof of Theorem 5.2.1(i)

Suppose $p > 4 - 2\sqrt{3}$ is fixed. We shall show that for n sufficiently large this implies that for any $\mu \in \mathcal{M}_{1,p}(H)$, whp 'Left meets Right' occurs. Our strategy for doing this is as follows: first of all we show in Lemma 5.2.6 that, for each $i \in [2]$, in any fixed tripartition $\sqcup_{j=1}^{3} V_j$ of $\{i\} \times [n]$, whp each of the parts V_j contains roughly the expected number of edges of $\mathbf{H}_\mu$, i.e. $(p + o(1))\, e(H[V_j])$. This immediately implies that whp there is a component C_L of $\mathbf{H}_\mu$ containing strictly more than half of the vertices of $\{1\} \times [n]$, and another component C_R containing at least half of the vertices of $\{2\} \times [n]$.

If these two components C_L and C_R are not the same, then we colour vertices of $[2] \times [n]$ Green if they lie in a small component of $\mathbf{H}_\mu[\{i\} \times [n]]$ for some $i \in [2]$, and otherwise Red if they are part of C_L and Blue if not (so in particular vertices in C_R are coloured Blue). This gives rise to a partition of $[n]$ into 9 sets $V_{c,c'}$, corresponding to the possible ordered colour pairs assigned to the vertex pairs $((1, v), (2, v))$, $v \in [n]$. Since whp at least $(p - o(1))n$ of the n edges from $\{1\} \times [n]$ to $\{2\} \times [n]$ are present in $\mathbf{H}_\mu$, we can combine the probabilistic information from Lemma 5.2.6 to show that whp the relative sizes of the $V_{c,c'}$ almost satisfy a certain system $\mathcal{S} = \mathcal{S}(p)$ of inequalities (5.2.7) to (5.2.10) (or more precisely that we can extract from the $|V_{c,c'}|/n$ a solution to $\mathcal{S}(p_\star)$ for some $p_\star$ a little smaller than p). For $p > 4 - 2\sqrt{3}$ and n sufficiently large, we are able to show this leads to a contradiction (Lemma 5.2.7). Having outlined our proof strategy, we now fill in the details. We shall use the following path-decomposition theorem due to Dean and Kouider.

Theorem 5.2.2 (Dean and Kouider [37])**.** *Let G be an n-vertex graph. Then there exists a set $\mathcal{P}$ of edge-disjoint paths in G such that $|\mathcal{P}| \leq \frac{2n}{3}$ and $\bigcup_{P \in \mathcal{P}} E(P) = E(G)$.*

Recall that a *matching* in a graph is a set of vertex-disjoint edges.

Corollary 5.2.3. *Let $\varepsilon > 0$ and let G be an n-vertex graph with $e(G) \geq 2n/\varepsilon$. Then there exists a set $\mathcal{M}$ of edge-disjoint matchings in G such that*

(M1) $|\mathcal{M}| \leq 2n$,

(M2) $|E(G) \setminus \bigcup_{M \in \mathcal{M}} M| \leq 2\varepsilon e(G)$, *and*

(M3) $|M| \geq \frac{\varepsilon e(G)}{2n}$ *for every* $M \in \mathcal{M}$.

Proof. By Theorem 5.2.2, there exists a set $\mathcal{P}$ of edge-disjoint paths in G such that $|\mathcal{P}| \leq \frac{2n}{3}$ and $E(G) = \bigcup_{P \in \mathcal{P}} E(P)$. Let $\mathcal{P}_{\text{short}} = \{P \in \mathcal{P} \colon e(P) \leq 2\varepsilon \frac{e(G)}{n}\}$. Let $\mathcal{M}$ be the set of matchings obtained by decomposing each path in $\mathcal{P} \setminus \mathcal{P}_{\text{short}}$ into two matchings. We have $|\mathcal{M}| \leq 2|\mathcal{P}| \leq 2n$. Moreover, each $M \in \mathcal{M}$ satisfies $|M| \geq \lfloor \frac{\varepsilon e(G)}{n} \rfloor \geq \frac{\varepsilon e(G)}{2n}$. Finally, $|E(G) \setminus \bigcup_{M \in \mathcal{M}} E(M)| \leq \frac{2n}{3} \cdot 2\varepsilon \frac{e(G)}{n} \leq 2\varepsilon e(G)$. $\square$

Matchings are useful in a 1-independent context since the states of their edges (present or absent) are independent. This is shown formally in the following proposition.

Proposition 5.2.4. *Let* G *be an* n*-vertex graph,* $p \in [0,1]$*, and* $\mu \in \mathcal{M}_{1,p}(G)$*. Let* M *be a matching in* G *and for each* $e \in M$*, let* $X_e = 1$ *if* $e \in E(\mathbf{G}_\mu)$ *and* $X_e = 0$ *otherwise. Then* $\{X_e \colon e \in M\}$ *is a set of independent random variables.*

Proof. Let $M' = \{e_1, \dots, e_t\} \subseteq M$ and for each $i \in [t]$, let $x_i \in \{0,1\}$. It suffices to show that

$$\mathbb{P}\left[\bigcap_{i \in [t]} \{X_{e_i} = x_i\}\right] = \prod_{i \in [t]} \mathbb{P}\left[X_{e_i} = x_i\right].$$

This follows immediately from the fact that for each $2 \leq j \leq t$,

$$\mathbb{P}\left[\bigcap_{i \in [j]} \{X_{e_i} = x_i\}\right] = \mathbb{P}\left[\bigcap_{i \in [j-1]} \{X_{e_i} = x_i\}\right] \cdot \mathbb{P}\left[X_{e_j} = x_j\right],$$

where we have used that μ is a 1-independent probability measure and thus the random variables $\{e_1, \dots, e_{j-1}\} \cap E(\mathbf{G}_\mu)$ and $\{e_j\} \cap E(\mathbf{G}_\mu)$ are independent. $\square$

We can thus combine Corollary 5.2.3 with a Chernoff bound to show the number of edges in a 1-independent model is concentrated around its mean.

Lemma 5.2.5. *Let* $\varepsilon > 0$ *and* $p \in (0,1]$*. Let* G *be an* n*-vertex graph with* $e(G) \geq 2n/\varepsilon$

and let $\mu \in \mathcal{M}_{1,p}(G)$. Then

$$\mathbb{P}\left[e(\mathbf{G}_\mu) \leq (1-3\varepsilon)pe(G)\right] \leq 4n \exp\left(-\frac{\varepsilon^3 pe(G)}{6n}\right).$$

Proof. We apply Corollary 5.2.3 to obtain a set $\mathcal{M}$ of edge-disjoint matchings in G such that properties (M1) to (M3) hold. For every $M \in \mathcal{M}$, we have $|M| \geq \frac{\varepsilon e(G)}{2n}$. Thus by (5.2.2) and 1-independence,

$$\mathbb{P}\left[e(\mathbf{G}_\mu \cap M) \leq (1-\varepsilon)p\,|M|\right] \leq 2\exp\left(-\frac{\varepsilon^3 pe(G)}{6n}\right).$$

By a union bound, we have

$$\begin{aligned}\mathbb{P}\left[e(\mathbf{G}_\mu \cap M) \geq (1-\varepsilon)p\,|M| \text{ for all } M \in \mathcal{M}\right] &\geq 1 - 2|M| \exp\left(-\frac{\varepsilon^3 pe(G)}{6n}\right) \\ &\geq 1 - 4n\exp\left(-\frac{\varepsilon^3 pe(G)}{6n}\right).\end{aligned}$$

Thus with probability at least $1 - 4n\exp\left(-\frac{\varepsilon^3 pe(G)}{6n}\right)$ we have

$$e(\mathbf{G}_\mu) \geq \sum_{M \in \mathcal{M}} (1-\varepsilon)p\,|M| \geq (1-\varepsilon)p(1-2\varepsilon)e(G) \geq (1-3\varepsilon)pe(G).$$

This completes the proof. □

Lemma 5.2.6. *Let $p \in (\frac{1}{2}, 1]$, and let $\varepsilon = \varepsilon(p) > 0$ be fixed and sufficiently small. Let G be an n-vertex graph satisfying*

$$\left|e(G[U]) - q\frac{|U|^2}{2}\right| \leq \frac{\varepsilon^2}{4}qn^2 \tag{5.2.3}$$

for all $U \subseteq V(G)$, where $q(n) \gg \frac{\log n}{n}$. Consider a fixed tripartition $V(G) = V_1 \sqcup V_2 \sqcup V_3$. Then for every $\mu \in \mathcal{M}_{1,p}(G)$, the following hold whp:

(P1) $e(\mathbf{G}_\mu[V_i]) \geq pq\frac{|V_i|^2}{2} - \varepsilon qn^2$ *for every $i \in [3]$.*

(P2) $e(\mathbf{G}_\mu[V_i, V_j]) \geq pq\,|V_i|\,|V_j| - \varepsilon qn^2$ *for all $1 \leq i < j \leq 3$.*

(P3) *For every* $i \in [3]$ *with* $|V_i| \geq \varepsilon^{1/4}n$, $\mathbf{G}_\mu[V_i]$ *contains a unique largest connected component* C_i *of order at least* $(\theta - \varepsilon^{1/4})\,|V_i|$.

(P4) *For all* $1 \leq i < j \leq 3$ *with* $|V_i|, |V_j| \geq \varepsilon^{1/4}n$, *there exists a path from* C_i *to* C_j *in* $\mathbf{G}_\mu[V_i, V_j]$.

(P5) *There is a unique largest connected component* C *in* $\mathbf{G}_\mu$ *such that* $|C| \geq (\theta - 3\varepsilon^{1/4})n$ *and for each* $i \in [3]$ *with* $|V_i| \geq \varepsilon^{1/4}n$, $C_i \subseteq C$.

Proof. We first show that (P1) holds whp. Fix $i \in [3]$. If $|V_i| \leq \sqrt{\varepsilon}n$, then (P1) trivially holds. Hence we assume that $|V_i| \geq \sqrt{\varepsilon}n$. By our pseudorandomness assumption (5.2.3) on G we have $e(G[V_i]) \geq q\frac{|V_i|^2}{2} - \frac{\varepsilon}{2}qn^2$ (which for n sufficiently large is greater than $\frac{2n}{\varepsilon}$ so that we can apply Lemma 5.2.5). Thus we have

$$\begin{aligned}
\mathbb{P}\left[e(\mathbf{G}_\mu[V_i]) \leq pq\frac{|V_i|^2}{2} - \varepsilon qn^2\right] &\leq \mathbb{P}\left[e(\mathbf{G}_\mu[V_i]) \leq pe(G[V_i]) - \frac{\varepsilon}{2}qn^2\right] \\
&\leq \mathbb{P}\left[e(\mathbf{G}_\mu[V_i]) \leq \left(1 - \frac{\varepsilon}{3}\right)pe(G[V_i])\right] \\
&\leq 4n\exp\left(-\Omega\left(\frac{e(G[V_i])}{n}\right)\right) = 4n\exp\left(-\Omega(qn)\right) = o(1),
\end{aligned}$$

where the inequality in the third line follows from Lemma 5.2.5. So (P1) holds whp.

Next we show that (P2) holds whp. Fix $1 \leq i < j \leq 3$. If $|V_i| \leq \varepsilon n$ or $|V_j| \leq \varepsilon n$, then (P2) trivially holds. Hence we may assume that $|V_i|, |V_j| \geq \varepsilon n$. By (5.2.3) applied three times (to V_i, V_j and $V_i \cup V_j$), we have $e(G[V_i, V_j]) \geq q\,|V_i|\,|V_j| - 3\frac{\varepsilon^2}{4}qn^2$. In particular, $e(G[V_i, V_j]) \geq \frac{\varepsilon^2}{4}qn^2$, which for n sufficiently large is greater than $\frac{2n}{\varepsilon}$. We now apply Lemma 5.2.5 to show that (P2) holds whp. We have

$$\begin{aligned}
&\mathbb{P}\left[e(\mathbf{G}_\mu[V_i, V_j]) \leq pq\,|V_i|\,|V_j| - \varepsilon qn^2\right] \leq \mathbb{P}\left[e(\mathbf{G}_\mu[V_i, V_j]) \leq pe(G[V_i, V_j]) - \frac{\varepsilon}{2}qn^2\right] \\
&\leq \mathbb{P}\left[e(\mathbf{G}_\mu[V_i, V_j]) \leq \left(1 - \frac{\varepsilon}{3}\right)pe(G[V_i, V_j])\right] \leq 4n\exp\left(-\Omega\left(\frac{e(G[V_i, V_j])}{n}\right)\right) \\
&= 4n\exp\left(-\Omega(qn)\right) = o(1).
\end{aligned}$$

So (P2) holds whp.

Now we show that (P1) implies (P3). Assume that (P1) holds. Fix $i \in [3]$ and assume that $|V_i| \geq \varepsilon^{1/4}n$. Let $C \subseteq V_i$ be a largest connected component in $\mathbf{G}_\mu[V_i]$ and suppose for a contradiction that $|C| \leq (\theta - \varepsilon^{1/4})\,|V_i|$.

If $|C| \leq \frac{|V_i|}{2}$, then there is a partition of V_i into at most 4 sets, each of size at most $\frac{|V_i|}{2}$, such that every connected component of $\mathbf{G}_\mu[V_i]$ is entirely contained in one of the sets of the partition. Indeed, such a partition can be obtained by starting with the partition where every connected component of $\mathbf{G}_\mu[V_i]$ forms its own part and then as long as there are two parts of size at most $\frac{|V_i|}{4}$ merging them into a single part. Since for any quadruple (x_1, x_2, x_3, x_4) with $\frac{1}{2} \geq x_i \geq 0$ and $\sum_i x_i = 1$ we have $\sum_i (x_i)^2 \leq \frac{1}{2}$, it follows from (P1) and (5.2.3) that

$$pq\frac{|V_i|^2}{2} - \varepsilon q n^2 \leq e(\mathbf{G}_\mu[V_i]) \leq q\frac{|V_i|^2}{4} + \varepsilon^2 q n^2.$$

Rearranging terms, this gives

$$(p - \frac{1}{2})q\frac{\varepsilon^{1/2}n^2}{2} \leq (p - \frac{1}{2})q\frac{|V_i|^2}{2} \leq q(\varepsilon + \varepsilon^2)n^2,$$

which is a contradiction for ε chosen sufficiently small. Thus we may assume $|C| \geq \frac{|V_i|}{2}$. Now by (P1) and (5.2.3) again, we have

$$\begin{aligned} pq\frac{|V_i|^2}{2} - \varepsilon q n^2 &\leq e(\mathbf{G}_\mu[V_i]) \leq e(\mathbf{G}_\mu[C]) + e(\mathbf{G}_\mu[V_i \setminus C]) \\ &\leq q\frac{|C|^2}{2} + q\frac{(|V_i| - |C|)^2}{2} + \frac{\varepsilon^2}{2}qn^2. \end{aligned}$$

Dividing by $q\frac{|V_i|^2}{2}$ and using $|V_i| \geq \varepsilon^{1/4}n$, we deduce that

$$p - 3\sqrt{\varepsilon} \leq \left(\frac{|C|}{|V_i|}\right)^2 + \left(1 - \frac{|C|}{|V_i|}\right)^2. \tag{5.2.4}$$

Since $x \mapsto x^2 + (1-x)^2$ is an increasing function in the interval $[\frac{1}{2}, 1]$, $\frac{1}{2}|V_i| \leq |C| \leq$

$(\theta - \varepsilon^{1/4})\,|V_i|$, and $\theta^2 + (1-\theta)^2 = p$, we have

$$\begin{aligned}\left(\frac{|C|}{|V_i|}\right)^2 + \left(1 - \frac{|C|}{|V_i|}\right)^2 &\leq (\theta - \varepsilon^{1/4})^2 + (1 - \theta + \varepsilon^{1/4})^2 \\ &= \theta^2 + (1-\theta)^2 - 2\varepsilon^{1/4}(2\theta - 1) + 2\sqrt{\varepsilon} \leq p - 4\sqrt{\varepsilon},\end{aligned}$$

contradicting (5.2.4). Hence $|C| \geq (\theta - \varepsilon^{1/4})\,|V_i|$. Note that since $\theta - \varepsilon^{1/4} > 1/2$ (for $\varepsilon = \varepsilon(p)$ chosen sufficiently small), C is the unique largest component in $\mathbf{G}_\mu[V_i]$. So (P3) holds whp.

Next we show that (P2) and (P3) together imply (P4). Assume that (P2) and (P3) hold. Fix $1 \leq i < j \leq 3$ and assume that $|V_i|, |V_j| \geq \varepsilon^{1/4} n$. Suppose for a contradiction that there is no path in $\mathbf{G}_\mu[V_i, V_j]$ from C_i to C_j. Let $A_i \subseteq V_i$ and $A_j \subseteq V_j$ be the sets of vertices which cannot be reached by a path in $\mathbf{G}_\mu[V_i, V_j]$ from C_j and C_i, respectively. Since there is no path from C_i to C_j, we must have $C_i \subseteq A_i$ and $C_j \subseteq A_j$. By (P2), by the definition of A_i and A_j, and by (5.2.3) (applied in A_i, A_j, $V_i \setminus A_i$, $V_j \setminus A_j$, $A_i \cup (V_j \setminus A_j)$ and $A_j \cup (V_i \setminus A_i)$), we have

$$\begin{aligned}pq\,|V_i|\,|V_j| - \varepsilon q n^2 &\leq e(\mathbf{G}_\mu[V_i, V_j]) \leq e(\mathbf{G}_\mu[A_i, V_j \setminus A_j]) + e(\mathbf{G}_\mu[V_i \setminus A_i, A_j]) \\ &\leq q\,|A_i|\,(|V_j| - |A_j|) + q\,|A_j|\,(|V_i| - |A_i|) + \frac{3\varepsilon^2}{2} q n^2.\end{aligned} \tag{5.2.5}$$

Let $x_i = \frac{|A_i|}{|V_i|}$ and $x_j = \frac{|A_j|}{|V_j|}$. By (P3), $x_i \geq \frac{|C_i|}{|V_i|} \geq \theta - \varepsilon^{1/4} \geq \frac{1}{2}$ and similarly $x_j \geq \frac{1}{2}$. From (5.2.5) we get by dividing by $q\,|V_i|\,|V_j|$ and using $|V_i|, |V_j| \geq \varepsilon^{1/4} n$, that

$$p - 2\sqrt{\varepsilon} \leq x_i(1 - x_j) + x_j(1 - x_i) = x_i + x_j - 2x_i x_j \leq \frac{1}{2}, \tag{5.2.6}$$

where the last inequality follows since $(x, y) \mapsto x + y - 2xy$ is non-increasing in both x and y for $x, y \geq \frac{1}{2}$. Note that (5.2.6) gives a contradiction for ε sufficiently small since $p > \frac{1}{2}$. So (P4) holds whp.

Finally, we observe that (P5) follows directly from (P3) and (P4). $\square$

Let $\mathcal{S}(p)$ denote the collection of 3×3 matrices A with non-negative entries $A_{ij} \geq 0$,

$i, j \in [3]$, satisfying the following inequalities:

$$A_{11} + A_{22} + p \leq \sum_{i,j} A_{ij} \leq 1, \tag{5.2.7}$$

$$\begin{aligned} A_{1j} &\geq \frac{1}{2}\sum_i A_{ij} \quad \forall j \in [3] \quad \text{and} \\ A_{i1} &\geq \frac{1}{2}\sum_j A_{ij} \quad \forall i \in [3], \end{aligned} \tag{5.2.8}$$

$$(A_{1j})^2 + (A_{2j})^2 \geq p\left(\sum_i A_{ij}\right)^2 \quad \forall j \in [3], \tag{5.2.9}$$

$$(A_{i1})^2 + (A_{i2})^2 \geq p\left(\sum_j A_{ij}\right)^2 \quad \forall i \in [3]. \tag{5.2.10}$$

The key step in our proof of Theorem 5.2.1 will be, assuming that 'Left meets Right' does not occur whp, to use Lemma 5.2.6 to exhibit a partition of $[n]$ into 9 parts whose relative sizes can be used to find a solution to $\mathcal{S}(p_\star)$, for some $p_\star$ such that $4 - 2\sqrt{3} < p_\star < p$. We will then be able to use the following lemma to derive a contradiction.

Lemma 5.2.7. *For $4 - 2\sqrt{3} < p \leq 1$, $\mathcal{S}(p) = \varnothing$.*

Proof. Suppose not and let $A \in \mathcal{S}(p)$. Note that the bound for $\sum_{i,j} A_{ij}$ in (5.2.7) implies

$$A_{11} + A_{22} \leq 1 - p. \tag{5.2.11}$$

By transpose-symmetry of $\mathcal{S}(p)$ and (5.2.7), we may assume without loss of generality that

$$w := A_{21} + A_{31} + A_{32} + A_{33} \geq \frac{p}{2}. \tag{5.2.12}$$

Note that if $\sum_j A_{3j} > \frac{A_{31}}{\theta}$, then, since $x \mapsto x^2 + (1-x)^2$ is an increasing function of x in the interval $[\frac{1}{2}, 1]$ and since $A_{31} \geq \frac{1}{2}\sum_j A_{3j}$ by (5.2.8),

$$\left(\frac{A_{31}}{\sum_j A_{3j}}\right)^2 + \left(\frac{A_{32}}{\sum_j A_{3j}}\right)^2 \leq \left(\frac{A_{31}}{\sum_j A_{3j}}\right)^2 + \left(1 - \frac{A_{31}}{\sum_j A_{3j}}\right)^2 < \theta^2 + (1-\theta)^2 = p,$$

contradicting (5.2.10). Hence

$$\sum_j A_{3j} \leq \frac{A_{31}}{\theta}. \tag{5.2.13}$$

By an analogous argument, we have $\sum_i A_{i1} \leq \frac{A_{11}}{\theta}$ and thus

$$A_{21} \leq A_{21} + A_{31} \leq \frac{1-\theta}{\theta} A_{11}. \tag{5.2.14}$$

Now, by (5.2.13) we have $w \leq A_{21} + \frac{A_{31}}{\theta}$. By (5.2.9), we have that

$$A_{31} \leq \frac{\sqrt{(A_{11})^2 + (A_{21})^2}}{\sqrt{p}} - A_{11} - A_{21}.$$

Substituting this expression into our upper bound on w, we get

$$w \leq -\frac{(1-\theta)A_{21}}{\theta} - \frac{A_{11}}{\theta} + \frac{\sqrt{(A_{11})^2 + (A_{21})^2}}{\theta\sqrt{p}}.$$

For A_{11} fixed, the continuous function $f_{A_{11}}(y) = -\frac{(1-\theta)y}{\theta} - \frac{A_{11}}{\theta} + \frac{\sqrt{(A_{11})^2+y^2}}{\theta\sqrt{p}}$ is convex in $(0, +\infty)$ as its derivative $f'_{A_{11}}(y) = -\frac{(1-\theta)}{\theta} + \frac{1}{\theta\sqrt{p}\sqrt{(A_{11}/y)^2+1}}$ is increasing in y in that interval. By (5.2.14), $0 \leq A_{21} \leq \frac{1-\theta}{\theta}A_{11}$, which together with the convexity of $f_{A_{11}}$ gives:

$$\begin{aligned}
w &\leq \max\left\{f_{A_{11}}(0),\ f_{A_{11}}\left(\frac{1-\theta}{\theta}A_{11}\right)\right\} \\
&\leq \max\left\{-\frac{A_{11}}{\theta} + \frac{A_{11}}{\theta\sqrt{p}},\ -\left(\frac{1-\theta}{\theta}\right)^2 A_{11} - \frac{A_{11}}{\theta} + A_{11}\frac{\sqrt{1+\left(\frac{1-\theta}{\theta}\right)^2}}{\theta\sqrt{p}}\right\} \\
&\leq \max\left\{\frac{A_{11}}{\theta}\left(\frac{1}{\sqrt{p}} - 1\right),\ \frac{A_{11}}{\theta}(1-\theta)\right\} \\
&\leq \max\left\{\frac{1-p}{\theta}\left(\frac{1}{\sqrt{p}} - 1\right),\ \frac{1-p}{\theta}(1-\theta)\right\},
\end{aligned}$$

where the last inequality follows from the upper bound (5.2.11) on A_{11}. We now claim that this contradicts (5.2.12), i.e. that

$$\max\left\{\frac{1-p}{\theta}\left(\frac{1}{\sqrt{p}}-1\right),\frac{1-p}{\theta}(1-\theta)\right\}<\frac{p}{2}.$$

Note that $p\mapsto\frac{1-p}{\theta}\left(\frac{1}{\sqrt{p}}-1\right)-\frac{p}{2}$ and $p\mapsto\frac{1-p}{\theta}(1-\theta)-\frac{p}{2}$ are both strictly decreasing functions (as θ is increasing in p). Hence to prove the claim above, it suffices to show that for $p=4-2\sqrt{3}$, we have $\frac{1-p}{\theta}\left(\frac{1}{\sqrt{p}}-1\right)\leq\frac{p}{2}$ and $\frac{1-p}{\theta}(1-\theta)\leq\frac{p}{2}$. Let $p=4-2\sqrt{3}$. Note that $(\sqrt{3}-1)^2=4-2\sqrt{3}$ and $(2-\sqrt{3})^2=7-4\sqrt{3}$. Hence $\sqrt{p}=\sqrt{3}-1$, $\sqrt{2p-1}=2-\sqrt{3}$, and $\theta=(3-\sqrt{3})/2$. Now it is easy to check that

$$\frac{1}{\sqrt{p}}-1=1-\theta=\frac{\theta}{(1-p)}\frac{p}{2}=\frac{\sqrt{3}-1}{2},$$

which completes the proof. $\square$

We are now ready to complete the proof of Theorem 5.2.1 (i).

Proof. Let $p>4-2\sqrt{3}$ be fixed. Let $\varepsilon=\varepsilon(p)>0$ be fixed and chosen sufficiently small. Let $p_\star=\frac{1}{2}\left(4-2\sqrt{3}+p\right)$. Finally, let n be sufficiently large so that for $G=G_n$ the pseudorandomness assumption (5.2.3) holds, and let $\mu\in\mathcal{M}_{1,p}(H)$, where $H=K_2\times G_n$.

For $i\in[2]$, let $\mathbf{G}^i_\mu=\mathbf{H}_\mu[\{i\}\times[n]]$. For $i,j\in[2]$ with $i\neq j$, let $\mathcal{E}_{ij}$ be the event that for any partition $(\{i\}\times V_1)\sqcup(\{i\}\times V_2)\sqcup(\{i\}\times V_3)$ of $\{i\}\times[n]$ such that $\{i\}\times V_1$ and $\{i\}\times V_2$ are each a union of components of order at least $\varepsilon^{1/4}n$ in $\mathbf{G}^i_\mu$, we have that $\mathbf{G}^j_\mu$ satisfies (P1) to (P5) of Lemma 5.2.6 with $\{j\}\times V_1$, $\{j\}\times V_2$, $\{j\}\times V_3$ playing the roles of V_1, V_2, V_3. Given $\mathbf{G}^i_\mu$ and ε fixed, the number of such partitions is at most $3^{\varepsilon^{-1/4}}=O(1)$. Hence Lemma 5.2.6 implies that $\mathcal{E}_{ij}$ holds whp.

Further, by 1-independence and (5.2.2), whp there are at least $(p-\varepsilon)n$ edges in the matching $\mathbf{H}_\mu[\{1\}\times[n],\{2\}\times[n]]$. Let $\mathcal{E}_{\text{good}}$ be the event that $\mathcal{E}_{12}$ and $\mathcal{E}_{21}$ both occur and that in addition $e(\mathbf{H}_\mu[\{1\}\times[n],\{2\}\times[n]])\geq(p-\varepsilon)n$. Then $\mathcal{E}_{\text{good}}$ holds whp. We claim that if $\mathcal{E}_{\text{good}}$ holds, then so does 'Left meets Right' (which implies the statement of the theorem).

Suppose for a contradiction that $\mathcal{E}_{\text{good}}$ holds but 'Left meets Right' does not. For $i \in [2]$, let C^i be the unique largest connected component in $\mathbf{G}^i_\mu$ (this exist by (P5)). Let $U_1 \sqcup U_2 \sqcup U_3 = [n]$ and $W_1 \sqcup W_2 \sqcup W_3 = [n]$ be such that the following hold.

(a) $\{1\} \times U_1$ is the union of C^1 and all connected components in $\mathbf{G}^1_\mu$ of order at least $\varepsilon^{1/4}n$ that can be reached from C^1 by a path in $\mathbf{H}_\mu$.

(b) $\{1\} \times U_2$ is the union of all other connected components in $\mathbf{G}^1_\mu$ of order at least $\varepsilon^{1/4}n$.

(c) $\{1\} \times U_3$ is the union of all connected components of order less than $\varepsilon^{1/4}n$ in $\mathbf{G}^1_\mu$.

(d) $\{2\} \times W_1$ is the union of all connected components in $\mathbf{G}^2_\mu$ of order at least $\varepsilon^{1/4}n$ that cannot be reached from C^1 by a path in $\mathbf{H}_\mu$.

(e) $\{2\} \times W_2$ is the union of all connected components in $\mathbf{G}^2_\mu$ of order at least $\varepsilon^{1/4}n$ that can be reached from C^1 by a path in $\mathbf{H}_\mu$.

(f) $\{2\} \times W_3$ is the union of all connected components in $\mathbf{G}^2_\mu$ of order less than $\varepsilon^{1/4}n$.

We can think of these partitions as giving us a 3-colouring of the vertices in $V(H)$: a vertex in $\{i\} \times [n]$ is coloured red if it belongs to a large component in $\mathbf{G}^i_\mu$ and can be reached from C^1 in $\mathbf{H}_\mu$, blue if it belongs to a large component in $\mathbf{G}^i_\mu$ and cannot be reached by C^1 in $\mathbf{H}_\mu$, and green if it belongs to a small component in $\mathbf{G}^i_\mu$. The key properties of this colouring are that the large components C^1 and C^2 in $\mathbf{G}^1_\mu$ and $\mathbf{G}^2_\mu$ are coloured red and blue respectively, that there are no edges from red vertices to blue vertices, and that the green vertices span few edges in $\mathbf{G}^i_\mu$, $i \in [2]$. Our 3-colouring of $V(H)$ gives rise to a partition of $[n]$ into 9 sets in a natural way, by considering the possible colour pairs for $((1, v), (2, v))$, $v \in [n]$. This partition is illustrated in Fig. 5.3.

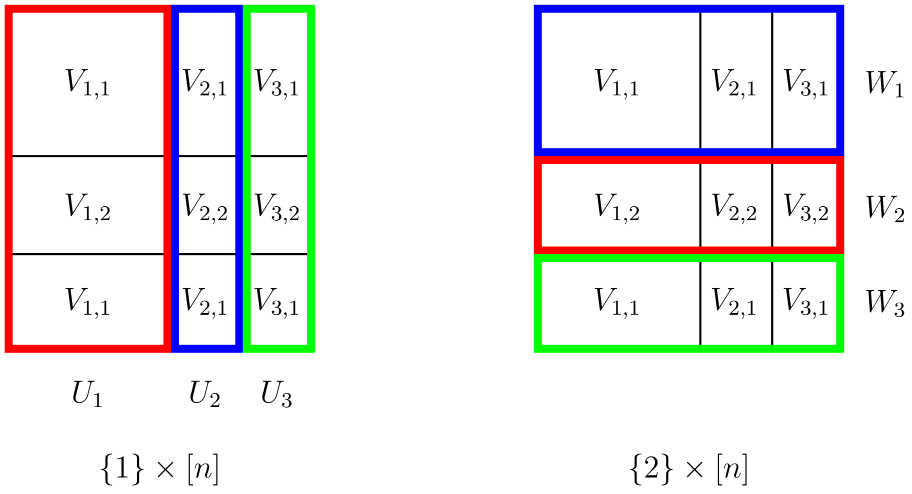

Figure 5.3: The partition of $V(H)$

We now investigate the relative sizes of this 9-partition. For $i, j \in [3]$, let $V_{ij} = U_i \cap W_j$. Since there is no path from C^1 to C^2 in $\mathbf{H}_\mu$, there are no edges present in the bipartite graphs $\mathbf{H}_\mu[\{1\} \times V_{11}, \{2\} \times V_{11}]$ and $\mathbf{H}_\mu[\{1\} \times V_{22}, \{2\} \times V_{22}]$. Since $\mathcal{E}_{\text{good}}$ holds, there are at least $(p - \varepsilon)n$ edges in $\mathbf{H}_\mu[\{1\} \times [n], \{2\} \times [n]]$ in total, which implies

$$|V_{11}| + |V_{22}| \leq (1 - p + \varepsilon)n. \tag{5.2.15}$$

Moreover, $\sum_{i,j} |V_{ij}| = n$. Hence

$$\sum_{i,j} |V_{ij}| - |V_{11}| - |V_{22}| \geq (p - \varepsilon)n. \tag{5.2.16}$$

For $j \in [3]$, if $|W_j| \geq \varepsilon^{1/4}n$, we have by (P3) and (P5) that there is a unique largest connected component C_j^1 in $\mathbf{G}_\mu^1[\{1\} \times W_j]$, and that this component satisfies $C_j^1 \subseteq C^1$ and $|C_j^1| \geq (\theta - \varepsilon^{1/4})|W_j|$, which for $\varepsilon = \varepsilon(p)$ chosen sufficiently small is greater than $\frac{1}{2}|W_j|$. Translating this in terms of our 9-partition, we have that for all $j \in [3]$ such that $\sum_i V_{ij} \geq \varepsilon^{1/4}n$

$$|V_{1j}| \geq \frac{1}{2} \sum_i |V_{ij}| \tag{5.2.17}$$

holds. By a symmetric argument, for every $i \in [3]$ such that $\sum_j V_{ij} \geq \varepsilon^{1/4} n$ we have

$$|V_{i1}| \geq \frac{1}{2} \sum_j |V_{ij}| \,. \tag{5.2.18}$$

Let $j \in [3]$. Note that $\mathbf{G}^1_\mu[U_3]$ contains only connected components of size at most $\varepsilon^{1/4} n$. These components can be covered by at most $\frac{2}{\varepsilon^{1/4}}$ sets, each of order at least $\frac{\varepsilon^{1/4} n}{2}$ and at most $\varepsilon^{1/4} n$. By (5.2.3) (which holds by our choice of n), each of these sets contains at most $q\frac{\varepsilon^{1/2} n^2}{2} + \frac{\varepsilon^2}{4} qn^2 < q\varepsilon^{1/2} n^2$ edges. Hence we have $e(\mathbf{G}^1_\mu[U_3]) \leq 2\varepsilon^{1/4} qn^2$. Since $V_{3j} \subseteq U_3$, we have $e(\mathbf{G}^1_\mu[V_{3j}]) \leq 2\varepsilon^{1/4} qn^2$. By (P1) and the pseudorandomness assumption (5.2.3), we have

$$\begin{aligned}
pq\frac{|W_j|^2}{2} - \varepsilon qn^2 &\leq e(\mathbf{G}^1_\mu[\{1\} \times W_j]) \\
&= e(\mathbf{G}^1_\mu[\{1\} \times V_{1j}]) + e(\mathbf{G}^1_\mu[\{1\} \times V_{2j}]) + e(\mathbf{G}^1_\mu[\{1\} \times V_{3j}]) \\
&\leq q\frac{|V_{1j}|^2}{2} + q\frac{|V_{2j}|^2}{2} + 2\varepsilon^{1/4} qn^2 + \frac{\varepsilon^2}{2} qn^2 < q\frac{|V_{1j}|^2}{2} + q\frac{|V_{2j}|^2}{2} + 3\varepsilon^{1/4} qn^2.
\end{aligned}$$

Hence, for every $j \in [3]$ and ε chosen sufficiently small,

$$|V_{1j}|^2 + |V_{2j}|^2 \geq p\left(\sum_i |V_{ij}|\right)^2 - 7\varepsilon^{1/4} n^2. \tag{5.2.19}$$

Similarly, for every $i \in [3]$,

$$|V_{i1}|^2 + |V_{i2}|^2 \geq p\left(\sum_j |V_{ij}|\right)^2 - 7\varepsilon^{1/4} n^2. \tag{5.2.20}$$

Let A be the 3×3 matrix with entries

$$A_{ij} = \begin{cases} \frac{|V_{ij}|}{n}, & \text{if } |V_{ij}| \geq \varepsilon^{1/9} n, \\ 0, & \text{otherwise.} \end{cases}$$

We claim that, provided $\varepsilon = \varepsilon(p)$ was chosen sufficiently small, $A \in \mathcal{S}(p_\star)$. Indeed, A

clearly has nonnegative entries summing up to at most 1, thus the second inequality of (5.2.7) is satisfied, while the first inequality (with $p_\star$ instead of p) follows from (5.2.16) and an appropriately small choice of ε (more specifically, we need $p_\star \le p - \varepsilon - 7\varepsilon^{1/9}$).

Next, consider $j \in [3]$. If $\sum_i |V_i| \ge \varepsilon^{1/4} n$, then by (5.2.17) we have $A_{1j} \ge \frac{1}{2}\sum_i A_{ij}$ (regardless of whether some of the V_{ij}, $i \in [3]$ have size less than $\varepsilon^{1/9} n$). Other the other hand if $\sum_i |V_i| < \varepsilon^{1/4} n$, then $A_{1j} = A_{2j} = A_{3j} = 0$. In either case, $A_{1j} \ge \frac{1}{2}\sum_i A_{ij}$ holds. By a symmetric argument we obtain that $A_{i1} \ge \frac{1}{2}\sum_j A_{ij}$ holds for every $i \in [3]$. Thus (5.2.8) is satisfied by A.

Finally, pick $j \in [3]$. If $|V_{i2}| \ge \varepsilon^{1/9} n$, then by (5.2.8) which we have just established and the definition of A_{i1}, we have $|V_{i1}| \ge \varepsilon^{1/9} n$ also. In this case (5.2.19) and an appropriately small choice of ε ensure that $(A_{1j})^2 + (A_{2j})^2 \ge p_\star \left(\sum_i A_{ij}\right)^2$. On the other hand, suppose $|V_{i2}| < \varepsilon^{1/9} n$. If $|V_{i1}| < \varepsilon^{1/9} n$, then by (5.2.8) the inequality $(A_{1j})^2 + (A_{2j})^2 \ge p_\star \left(\sum_i A_{ij}\right)^2$ holds trivially, since the right hand-side is zero. So suppose that $|V_{i1}| \ge \varepsilon^{1/9} n > |V_{i2}|$. Then (5.2.19), and $p > 1/2$ imply that

$$|V_{i1}|^2 > |V_{i1}|^2 - |V_{i2}| \left(2p\, |V_{i1}| - (1-p)\, |V_{i2}|\right) \ge p \left(|V_{i1}| + |V_{i3}|\right)^2 - 7\varepsilon^{1/4} n^2.$$

Together with an appropriately small choice of ε, this ensures $(A_{1j})^2 + (A_{2j})^2 \ge p_\star \left(\sum_i A_{ij}\right)^2$ again. Thus in every case (5.2.9) is satisfied by A (with $p_\star$ instead of p). A symmetric argument shows A satisfies (5.2.10) for $p_\star$ as well.

Thus $A \in \mathcal{S}(p_\star)$ as claimed. However, since $p_\star > 4 - 2\sqrt{3}$, Lemma 5.2.7 implies that $\mathcal{S}(p_\star) = \varnothing$, a contradiction. Thus the event $\mathcal{E}_{\text{good}}$, which holds whp, does imply the event 'Left meets Right', proving the theorem. □

5.3 Proof of Theorems 1.4.8, 1.4.9, 1.4.11 and 1.4.12

Our main theorems are all proved via a renormalisation argument combined with Theorem 5.2.1. Given two graphs G and H, we may view the Cartesian product $H \times G$ as a kind of 'augmented' version of H, and use any 1-independent random graph $(\mathbf{H} \times \mathbf{G})_\mu$ on $H \times G$ to construct a new 1-independent random graph $\mathbf{H}_\nu$ on H as follows: given an edge $uv \in E(H)$, we let uv be present in $\mathbf{H}_\nu$ if in the restriction of $(\mathbf{H} \times \mathbf{G})_\mu$ to $\{u, v\} \times V(G)$ there is a connected component containing strictly more than half of the vertices in each of $\{u\} \times V(G)$ and $\{v\} \times V(G)$.

That $\mathbf{H}_\nu$ is a 1-independent random graph follows immediately from the fact that $(\mathbf{H} \times \mathbf{G})_\mu$ was 1-independent: the states of edges inside vertex-disjoint edge-sets in $\mathbf{H}_\nu$ are determined by the states of edges inside vertex-disjoint edge sets in $(\mathbf{H} \times \mathbf{G})_\mu$. Further, any path in $\mathbf{H}_\nu$ can be 'lifted' up to a path in $(\mathbf{H} \times \mathbf{G})_\mu$ of equal or greater length: if uv, vw are present in $\mathbf{H}_\nu$, then there exist connected subgraphs C_{uv} and C_{vw} in $(\mathbf{H} \times \mathbf{G})_\mu$ with $C_{uv} \subseteq \{u, v\} \times V(G)$, $C_{vw} \subseteq \{v, w\} \times V(G)$, $C_{uv} \cap (\{u\} \times V(G))$ and $C_{vw} \cap (\{w\} \times V(G))$ both non-empty, and $C_{uv}, C_{v,w}$ both containing strictly more than half of the vertices in $\{v\} \times V(G)$ (and hence having non-empty intersection).

Now the likelihood of an edge uv being present in $\mathbf{H}_\nu$ is exactly the probability of the event corresponding to 'Left meets Right' occurring in the restriction of $(\mathbf{H} \times \mathbf{G})_\mu$ to the vertex-set $\{u, v\} \times V(G)$ (which induces a copy of $K_2 \times G$ in $H \times G$). Thus for $p > 4 - 2\sqrt{3}$ and a suitable choice of G, we can use Theorem 5.2.1(i) to ensure that each edge in the 1-independent random graph $\mathbf{H}_\nu$ is present with probability $1 - o(1)$. With such a high edge probability, we can then establish the almost sure existence of infinite components or long paths in $\mathbf{H}_\nu$ in a straightforward way — either by using results in the literature, or by direct argument.

On the other hand if $p \leq 4 - 2\sqrt{3}$, we can use ideas from the lower bound construction in the proof of Theorem 5.2.1(ii), which date back to [36, 46], in order to construct a 1-independent random subgraph $\mathbf{G}$ of $H \times K_n$ that fails to percolate (or, if $H = \mathbb{Z}$, that only contain paths of length $O(n)$). For the convenience of the reader, we sketch below how this works in the special case $H = \mathbb{Z}^2$.

Take $p = 4 - 2\sqrt{3}$, and set $\theta = (1 + \sqrt{2p - 1})/2$. Independently assign to each vertex $(x, y, z) \in \mathbb{Z}^2 \times V(K_n)$ a random state $S_{x,y,z} \in \{0, 1, \star\}$ as follows:

- if $\|(x,y)\|_\infty \cong 0 \mod 6$, set $S_{x,y,z} = 1$ with probability 1;
- if $\|(x,y)\|_\infty \cong 1 \mod 6$, set $S_{x,y,z} = 1$ with probability θ, and 0 otherwise;
- if $\|(x,y)\|_\infty \cong 2 \mod 6$, set $S_{x,y,z} = 0$ with probability $\sqrt{p}$, and $\star$ otherwise;
- if $\|(x,y)\|_\infty \cong 3 \mod 6$, set $S_{x,y,z} = 0$ with probability 1;
- if $\|(x,y)\|_\infty \cong 4 \mod 6$, set $S_{x,y,z} = 0$ with probability θ, and 1 otherwise;
- if $\|(x,y)\|_\infty \cong 5 \mod 6$, set $S_{x,y,z} = 1$ with probability $\sqrt{p}$, and $\star$ otherwise.

We now use these random states to build a 1-independent random graph $\mathbf{G}$ as follows. Given an edge$\{(x_1,y_1,z_1),(x_2,y_2,z_2)\}$ of $H \times K_n$, include it in $\mathbf{G}$ if one of the following holds:

- $S_{x_1,y_1,z_1} = S_{x_2,y_2,z_2} \neq \star$
- $\|(x_1,y_1)\|_\infty < \|(x_2,y_2)\|_\infty$ and $S_{x_2,y_2,z_2} = \star$.

Then the choice of probabilities for our random states ensure each edge is open[1] with probability at least $p = 4 - 2\sqrt{3}$, and our edge rules further imply that every connected component C in $\mathbf{G}$ meets at most four consecutive cylinders $\mathcal{C}_r := \{(x,y,z) : \ \|(x,y)\|_\infty = r\}$, $r \in \mathbb{Z}_{\geq 0}$ since, as is easily checked, a connected component in $\mathbf{G}$ cannot both contain a vertex assigned state 0 and a vertex assigned state 1 — we leave this as an exercise to the reader, and refer them to [36, Corollary 24] for a proof of this fact in a more general setting. In particular, we have that $\mathbf{G}$ does not percolate.

Having thus outlined our proof ideas, we now fill in the details. First we formalise our renormalisation argument with the following lemma.

Lemma 5.3.1 (Renormalisation lemma)**.** *Let H be a graph. Let $q = q(n)$ satisfy $nq(n) \gg \log n$, and let $(G_n)_{n\in\mathbb{N}}$ be a sequence of n-vertex graphs which is weakly q-pseudorandom. Then for every $\varepsilon > 0$ and every $p > 4 - 2\sqrt{3}$ fixed, there exists n_0 such that for all $n \geq n_0$,*

[1]Here we say that an edge is open if it is included in the random graph corresponding to the measure.

$G = G_n$ and $\mu \in \mathcal{M}_{1,\geq p}(H \times G)$ there exists $\nu \in \mathcal{M}_{1,\geq 1-\varepsilon}(H)$ and a coupling between $\mathbf{H}_\nu$ and $(\mathbf{H} \times \mathbf{G})_\mu$ such that there exists a path from u to v in $\mathbf{H}_\nu$ only if there exists a path from $\{u\} \times V(G)$ to $\{v\} \times V(G)$ in $(\mathbf{H} \times \mathbf{G})_\mu$.

Proof. Let $p > 4 - 2\sqrt{3}$ and $\varepsilon > 0$ be fixed. By Theorem 5.2.1(i), there exists $n_0 \in \mathbf{N}$ such that for all $n \geq n_0$ and all $\mu \in \mathcal{M}_{1,\geq p}(K_2 \times G_n)$, the μ-probability of the event 'Left meets Right' is at least $1 - \varepsilon$. For $n \geq n_0$, $G = G_n$ and $\mu \in \mathcal{M}_{1,\geq p}(H \times G)$, define a random graph model $\mathbf{H}_\nu$ from $(\mathbf{H} \times \mathbf{G})_\mu$ as follows: for each edge $uv \in E(H)$, we add uv to $\mathbf{H}_\nu$ if and only if there is a connected component in $(\mathbf{H} \times \mathbf{G})_\mu[\{u, v\} \times V(G_n)]$ containing strictly more than half of the vertices in $\{u\} \times V(G_n)$ and strictly more than half of the vertices $\{v\} \times V(G_n)$. The model $\mathbf{H}_\nu$ is clearly 1-independent, has edge-probability at least $1 - \varepsilon$, and has the property that any path in $\mathbf{H}_\nu$ can be lifted up to a path in $(\mathbf{H} \times \mathbf{G})_\mu$. This proves the Lemma. □

Recall that 2-neighbour bootstrap percolation on a graph G is a discrete-time process defined as follows. At time $t = 0$, an initial set of infected vertices $A = A_0$ is given. At every time $t \geq 1$, every vertex of G which has at least 2 neighbours in A_{t-1} becomes infected and is added to A_{t-1} to form A_t. We denote by $\overline{A}$ the set of all vertices of G which are eventually infected, $\overline{A} = \bigcup_{t \geq 0} A_t$. Following Day, Falgas-Ravry and Hancock [36], we say that a graph G has the *finite 2-percolation property* if for every finite set of initially infected vertices A, the set of eventually infected vertices $\overline{A}$ is finite. The content of [36][Corollary 24] is, informally, that the construction based on random-states we outline above 'works on all host graphs that have the finite 2-percolation property'.

Proof of Theorem 1.4.11. Let $H = \mathbb{Z}^2$. Pick $\varepsilon > 0$ such that $1 - \varepsilon > 0.8639$. Then by Lemma 5.3.1, for any $p > 4 - 2\sqrt{3}$, n sufficiently large and $G = G_n$, we can couple a random graph $(\mathbf{H} \times \mathbf{G})_\mu$, $\mu \in \mathcal{M}_{1,\geq p}(H)$ with a random graph $\mathbf{H}_\nu$, $\mu \in \mathcal{M}_{1,\geq 1-\varepsilon}(H)$ such that if $\mathbf{H}_\nu$ percolates then so does $(\mathbf{H} \times \mathbf{G})_\mu$. Since $p_{1,c}(H) < 0.8639$, as proved in [16, Theorem 2], it follows that $p_{1,c}(H \times G) \leq p$. Since $p > 4 - 2\sqrt{3}$ was arbitrary, we have the claimed upper bound $\lim_{n\to\infty} p_{1,c}(H \times G_n) \leq 4 - 2\sqrt{3}$. The lower bound

$\lim_{n\to\infty} p_{1,c}(H \times G_n) \geq 4 - 2\sqrt{3}$ follows from [36, Corollary 24] and the fact that $\mathbb{Z}^2 \times G_n$ is easily seen to have the finite 2-percolation property. Indeed, for any finite set of vertices A in $\mathbb{Z}^2 \times G_n$, there is some finite N such that $A \subseteq [N]^2 \times V(G_n)$. Now every vertex outside $[N]^2 \times V(G_n)$ has at most one neighbour in $[N]^2 \times V(G_n)$, and thus can never be infected by a 2-neighbour bootstrap percolation process started from A. □

Remark 5.3.2. *The proof above in fact works in a more general setting than $\mathbb{Z}^2$: suppose H has the finite 2-percolation property and satisfies $p_{1,c}(H) < 1$. Let $(G_n)_{n\in\mathbb{N}}$ be a sequence of weakly q-pseudorandom n-vertex graphs with $nq(n) \gg \log n$. Then $H \times G_n$ also has the finite 2-percolation property, and the proof above shows*

$$\lim_{n\to\infty} p_{1,c}(H \times G_n) = 4 - 2\sqrt{3}.$$

Examples of graphs with the finite 2-percolation property include many of the standard lattices studied in percolation theory, such as the honeycomb (hexagonal) lattice, the dice (rhombile) lattice or the tetrakis ('Union Jack') lattice.

Proof of Theorem 1.4.8. Since K_n is 1-pseudorandom, Theorem 1.4.8 is immediate from Theorem 1.4.11. □

Proof of Theorem 1.4.12. Let $H = \mathbb{Z}^2$. Pick $\varepsilon > 0$ such that $1 - \varepsilon > 3/4$. Then by Lemma 5.3.1, for any $p > 4 - 2\sqrt{3}$, n sufficiently large and $G = G_n$, we can couple a random graph $(\mathbf{H} \times \mathbf{G})_\mu$, $\mu \in \mathcal{M}_{1,\geq p}(H)$ with a random graph $\mathbf{H}_\nu$, $\mu \in \mathcal{M}_{1,\geq 1-\varepsilon}(H)$ such that if $\mathbf{H}_\nu$ contains a path of length ℓ then so does $(\mathbf{H} \times \mathbf{G})_\mu$. Since $p_{1,\mathrm{LP}}(H) = \frac{3}{4}$, as proved in [36, Theorem 11(i)][1] it follows that $p_{1,\mathrm{LP}}(H \times G) \leq p$. Since $p > 4 - 2\sqrt{3}$ was arbitrary, we have the claimed upper bound $\lim_{n\to\infty} p_{1,\mathrm{LP}}(H \times G_n) \leq 4 - 2\sqrt{3}$. The lower bound $\lim_{n\to\infty} p_{1,c}(H \times G_n) \geq 4 - 2\sqrt{3}$ was proved in [36, Theorem 12(v)] (with the same construction as we outlined at the beginning of this section, adapted mutatis mutandis to the setting $H = \mathbb{Z}$). □

[1]For the proof of this theorem, all we need is $p_{1,\mathrm{LP}}(H) < 1$, and thus the weaker bound $p_{1,\mathrm{LP}}(H) \leq 1 - 1/3e$ (which follows directly from an application of the Lovász local lemma) would suffice for our purposes here.

Proof of Theorem 1.4.9. Since K_n is 1-pseudorandom, Theorem 1.4.9 is immediate from Theorem 1.4.12. □

5.4 Component evolution in 1-independent models

Recall that the independence number $\alpha(G)$ of a graph G is the size of a largest independent (edge-free) subset of $V(G)$, and that a *perfect matching* in a graph G is a matching whose edges together cover all the vertices in $V(G)$. Moreover, a graph G is a *complete multipartite graph* if there exists a partition of $V(G)$ such that two vertices in $V(G)$ are joined by an edge in G if and only if they are contained in different parts of the partition. Finally, the *complement* G^c of a graph G is the graph on $V(G)$ whose edges are the non-edges of G, $G^c := (V(G), V(G)^{(2)} \setminus E(G))$.

Lemma 5.4.1. *If G is a complete multipartite graph on $2n$ vertices with independence number $\alpha(G) \leq n$, then G contains at least $n!$ perfect matchings.*

Proof. Let G be a complete multipartite graph on $2n$ vertices with the minimum number of perfect matchings subject to $\alpha(G) \leq n$. Let $V_1, V_2, \ldots, V_r$ denote the parts of G with $|V_1| \geq |V_2| \geq \cdots \geq |V_r|$. If $|V_{r-1}|+|V_r| \leq n$, then the graph G' obtained from G by deleting all edges in $G[V_{r-1}, V_r]$ satisfies $\alpha(G') \leq n$ and has at most as many perfect matchings as G. We may therefore assume that $|V_{r-1}| + |V_r| \geq n$, and thus in particular that $r \leq 3$. Consider a perfect matching M in G and let i be the number of edges in $E(G[V_1, V_2]) \cap M$. Clearly $|E(G[V_1, V_3]) \cap M| = |V_1| - i$ and $|E(G[V_2, V_3]) \cap M| = |V_2| - i = |V_3| - (|V_1| - i)$. From this we deduce that $i = \frac{1}{2}(|V_1| + |V_2| - |V_3|) = n - |V_3|$. Hence the number $\mathrm{PM}(G)$ of perfect matchings in G is:

$$\begin{aligned}\mathrm{PM}(G) &= \binom{|V_1|}{i}\binom{|V_2|}{i}\binom{|V_3|}{|V_1|-i} i!(|V_2|-i)!(|V_1|-i)! \\ &= \frac{|V_1|!\,|V_2|!\,|V_3|!}{(n-|V_1|)!(n-|V_2|)!(n-|V_3|)!}.\end{aligned}$$

(Here $\binom{|V_1|}{i}\binom{|V_2|}{i}i!$ counts the number of different ways of selecting i-sets of vertices from each of V_1 and V_2 and joining them by a perfect matching, while $\binom{|V_3|}{|V_1|-i}(|V_2|-i)!(|V_1|-i)!$ counts the number of ways of joining the vertices of V_3 by a perfect matching to the remaining vertices of $V_1 \cup V_2$.)

If $|V_3| > 0$, then let G' be the complete tripartite graph with parts of size $|V_1|, |V_2| + 1, |V_3| - 1$. Note that $\alpha(G') \leq n$. By the formula above , we have

$$\frac{\mathrm{PM}(G)}{\mathrm{PM}(G')} = \frac{|V_3|\,(n - |V_3| + 1)}{(|V_2| + 1)(n - |V_2|)} \geq 1,$$

since $|V_3|\,(n - |V_3| + 1) - (|V_2| + 1)(n - |V_2|) = (|V_2| - |V_3| + 1)(|V_2| + |V_3| - n) \geq 0$ (as $|V_2| \geq |V_3|$ and $|V_2| + |V_3| \geq n$). It follows that $\mathrm{PM}(G) \geq \mathrm{PM}(K_{n,n}) = n!$ as claimed. $\square$

Proof of Proposition 1.4.16. Let $H = K_{2n}$. For all $p \in [\frac{1}{2}, 1]$, we may construct the two-state measure $\mu_{2s,p} \in \mathcal{M}_{1,p}(H)$ which satisfies:

$$\mathbb{P}\left[|C_1(\mathbf{H}_{\mu_{2s,p}})| \leq n\right] = \mathbb{P}\left[|C_1(\mathbf{H}_{\mu_{2s,p}})| = n\right] = \binom{2n}{n}\theta^n(1-\theta)^n = \binom{2n}{n}\left(\frac{1-p}{2}\right)^n,$$

proving the upper bound in that range. For $p_{2n} \leq p \leq \frac{1}{2}$, we note that $\theta = \theta(p)$ is no longer a real number. However, as shown in [36, Section 7.1], we may take a 'complex limit' of the 2-state measure $\mu_{2s,p}$, and the conclusion above still holds.

For the lower bound, let $C_1, C_2, \ldots, C_r$ be the connected components of a μ-random subgraph $\mathbf{H}_\mu$ of K_{2n}. Let $\mathbf{G}$ denote the complete multipartite graph associated with the partition $\sqcup_i C_i$ of $V(K_{2n}) = [2n]$. Observe that $\mathbf{G}$ is a subgraph of the complement $\mathbf{H}_\mu^c$ of $\mathbf{H}_\mu$. If $|C_i| \leq n$ for all i, then $\alpha(\mathbf{G}) \leq n$, whence by Lemma 5.4.1 $\mathbf{G}$ contains at least $n!$ perfect matchings. In particular, $\mathbf{H}_\mu^c$ must contain at least $n!$ perfect matchings. By Markov's inequality, we thus have

$$\begin{aligned}
\mathbb{P}\left[|C_1(\mathbf{H}_\mu)| \leq n\right] &\leq \mathbb{P}\left[\mathbf{H}_\mu^c \text{ contains } \geq n! \text{ perfect matchings}\right] \\
&\leq \frac{1}{n!}\mathbb{E}\left[\#\{\text{perfect matchings in } \mathbf{H}_\mu^c\}\right] \\
&= \frac{1}{n!}\left(\frac{1}{n!}\prod_{i=0}^{n-1}\binom{2n-2i}{2}\right)(1-p)^n = \binom{2n}{n}\left(\frac{1-p}{2}\right)^n.
\end{aligned}$$

(Here $\left(\frac{1}{n!}\prod_{i=0}^{n-1}\binom{2n-2i}{2}\right)$ counts the number of perfect matchings in K_{2n} by selecting n vertex-disjoint edges sequentially one after the other, and dividing through by $n!$.) The lower bound follows. $\square$

Proof of Theorem 1.4.17. Let $p \in (\frac{1}{r+1}, \frac{1}{r}]$ be fixed. Fix $\varepsilon = \varepsilon(p) > 0$ sufficiently small. For n large enough, we have by the pseudorandomness assumption on H_n that for every $U \subseteq V(H_n)$, $e(H_n[U]) \leq q\frac{|U|^2}{2} + \varepsilon^2 pqn^2$. It then follows from Lemma 5.2.5 that whp

$$e(\mathbf{H}_\mu) \geq pq\frac{n^2}{2}(1 - 4\varepsilon^2), \tag{5.4.1}$$

which is strictly greater than $\frac{qn^2}{2(r+1)}$ for $\varepsilon = \varepsilon(p)$ chosen sufficiently small. Assume (5.4.1). We show this implies the claimed lower bound on the size of a largest component.

If $|C_1(\mathbf{H}_\mu)| \leq \frac{n}{r+1} - \varepsilon n$, then for ε sufficiently small there is a partition of $V(H)$ into at most $2(r+1)+1$ sets, each of which has size at most $\frac{n}{r+1} - \varepsilon n$, such that every connected component of $\mathbf{H}_\mu$ is wholly contained in one of the sets of the partition. Indeed, such a partition can be obtained by starting with a partition of $V(H)$ into the connected components of $\mathbf{H}_\mu$, and then as long as the partition contains two parts of size at most $\frac{1}{2}\left(\frac{n}{r+1} - \varepsilon n\right)$, choosing two such parts arbitrarily and merging them into a single part. Since for any $(2r+3)$-tuple $(x_1, \ldots, x_{2r+3})$ with $\frac{1}{r+1} - \varepsilon \geq x_i \geq 0$ and $\sum_i x_i = 1$ we have $\sum_i (x_i)^2 \leq (r+1)\left(\frac{1}{r+1} - \varepsilon\right)^2 + ((r+1)\varepsilon)^2$, we have by our pseudorandomness assumption that

$$e(\mathbf{H}_\mu) \leq \frac{q(r+1)}{2}\left(\frac{1}{r+1} - \varepsilon\right)^2 n^2 + \frac{q}{2}\left((r+1)\varepsilon\right)^2 n^2 + (2r+3)\varepsilon^2 pqn^2 < \frac{qn^2}{2(r+1)}$$

for ε sufficiently small, contradicting (5.4.1). Thus we may assume that $|C_1(\mathbf{H}_\mu)| > \frac{n}{r+1} - \varepsilon n$.

If $|C_1(\mathbf{H}_\mu)| \geq \frac{n}{r}$, then we have nothing to show. Finally if $\frac{n}{r+1} - \varepsilon n \leq |C_1(\mathbf{H}_\mu)| < \frac{n}{r}$, then $\mathbf{H}_\mu$ contains at least $r+1$ components. Let αn denote the size of a largest component, where $\frac{1}{r+1} - \varepsilon < \alpha < \frac{1}{r}$. Then

$$\left(r\alpha^2 + (1-r\alpha)^2\right) q\frac{n^2}{2} + (r+2)\varepsilon^2 pqn^2 \geq e(\mathbf{H}_\mu) \geq pq\frac{n^2}{2}(1-4\varepsilon^2).$$

Dividing through by $qn^2/2$, rearranging terms and using the fact ε is chosen sufficiently small, we get

$$r\alpha^2 + (1-r\alpha)^2 \geq p - \varepsilon.$$

Solving for α, we get that

$$\alpha \geq \frac{1 + \sqrt{\frac{(r+1)(p-\varepsilon)-1}{r}}}{r+1},$$

giving part (i).

For part (ii), consider the $r+1$-state measure in which each vertex is assigned state $r+1$ with probability $\frac{1-\sqrt{r((r+1)p-1)}}{r+1}$ and a uniform random state from the set $\{1, 2, \ldots, r\}$ otherwise, and in which an edge is open if and only if its vertices are in the same state. This is easily seen to be a 1-ipm with the requisite properties. $\square$

BIBLIOGRAPHY

[1] P. Allen. Covering two-edge-coloured complete graphs with two disjoint monochromatic cycles. *Combin. Probab. Comput.*, 17(4):471–486, 2008.

[2] P. Allen, J. Böttcher, O. Cooley, and R. Mycroft. Tight cycles and regular slices in dense hypergraphs. *J. Combin. Theory Ser. A*, 149:30–100, 2017.

[3] P. Allen, J. Böttcher, O. Cooley, and R. Mycroft. Tight cycles and regular slices in dense hypergraphs. *J. Combin. Theory Ser. A*, 149:30–100, 2017.

[4] P. Allen, J. Böttcher, Y. Kohayakawa, and Y. Person. Tight Hamilton cycles in random hypergraphs. *Random Structures Algorithms*, 46(3):446–465, 2015.

[5] P. Allen, J. Böttcher, R. Lang, J. Skokan, and M. Stein. Partitioning a 2-edge-coloured graph of minimum degree $2n/3 + o(n)$ into three monochromatic cycles. *arXiv e-prints*, arXiv:2204.00496, 2022.

[6] P. Allen, E. Davies, and J. Skokan. Regularity inheritance in hypergraphs. *arXiv e-prints*, arXiv:1901.05955, 2019.

[7] P. Allen, C. Koch, O. Parczyk, and Y. Person. Finding tight Hamilton cycles in random hypergraphs faster. *Combin. Probab. Comput*, pages 1–19, Sep 2020.

[8] P. Allen, O. Parczyk, and V. Pfenninger. Resilience for tight Hamiltonicity. *arXiv e-prints*, arXiv:2105.04513, 2021.

[9] N. Alon, M. Capalbo, Y. Kohayakawa, V. Rödl, A. Ruciński, and E. Szemerédi. Universality and tolerance (extended abstract). In *41st Annual Symposium on Foundations of Computer Science (Redondo Beach, CA, 2000)*, pages 14–21. IEEE Comput. Soc. Press, Los Alamitos, CA, 2000.

[10] N. Alon, E. Fischer, M. Krivelevich, and M. Szegedy. Efficient testing of large graphs. *Combinatorica*, 20(4):451–476, 2000.

[11] J. Ayel. Sur l'existence de deux cycles supplémentaires unicolores, disjoints et de couleurs différentes dans un graphe complet bicolore. *Theses, Université Joseph-Fourier - Grenoble I*, May 1979.

[12] A. Bálint, V. Beffara, and V. Tassion. On the critical value function in the divide and color model. *ALEA Lat. Am. J. Probab. Math. Stat.*, 10(2):653–666, 2013.

[13] P. Balister and B. Bollobás. Critical probabilities of 1-independent percolation models. *Combin. Probab. Comput.*, 21(1-2):11–22, 2012.

[14] P. Balister and B. Bollobás. Percolation in the k-nearest neighbor graph. In *Recent Results in Designs and Graphs: a Tribute to Lucia Gionfriddo*, volume 28 of *Quaderni di Matematica*, pages 83–100. 2013.

[15] P. Balister, B. Bollobas, A. Sarkar, and S. Kumar. Reliable density estimates for coverage and connectivity in thin strips of finite length. In *Proceedings of the 13th annual ACM international conference on Mobile computing and networking*, pages 75–86. ACM, 2007.

[16] P. Balister, B. Bollobás, and M. Walters. Continuum percolation with steps in the square or the disc. *Random Structures Algorithms*, 26(4):392–403, 2005.

[17] P. Balister, B. Bollobás, and M. Walters. Random transceiver networks. *Adv. in Appl. Probab.*, 41(2):323–343, 2009.

[18] P. Balister, T. Johnston, M. Savery, and A. Scott. Improved bounds for 1-independent percolation on $\mathbb{Z}^n$. *arXiv e-prints*, arXiv:2206.12335, 2022.

[19] N. Ball. Rigorous confidence intervals on critical thresholds in 3 dimensions. *J. Stat. Phys.*, 156(3):574–585, 2014.

[20] J. Balogh, J. Barát, D. Gerbner, A. Gyárfás, and G. N. Sárközy. Partitioning 2-edge-colored graphs by monochromatic paths and cycles. *Combinatorica*, 34(5):507–526, 2014.

[21] I. Benjamini and A. Stauffer. Perturbing the hexagonal circle packing: a percolation perspective. In *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, volume 49, pages 1141–1157. Institut Henri Poincaré, 2013.

[22] S. Bessy and S. Thomassé. Partitioning a graph into a cycle and an anticycle, a proof of lehel's conjecture. *J. Combin. Theory Ser. B*, 100(2):176 – 180, 2010.

[23] N. Biggs. The roots of combinatorics. *Historia Mathematica*, 6(2):109–136, 1979.

[24] B. Bollobás. The evolution of sparse graphs. In *Graph theory and combinatorics (Cambridge, 1983)*, pages 35–57. Academic Press, London, 1984.

[25] B. Bollobás and O. Riordan. *Percolation.* Cambridge University Press, New York, 2006.

[26] J. A. Bondy and P. Erdős. Ramsey numbers for cycles in graphs. *J. Combin. Theory Ser. B*, 14:46–54, 1973.

[27] S. Bustamante, J. Corsten, N. Frankl, A. Pokrovskiy, and J. Skokan. Partitioning edge-colored hypergraphs into few monochromatic tight cycles. *SIAM J. Discrete Math.*, 34(2):1460–1471, 2020.

[28] S. Bustamante, H. Hàn, and M. Stein. Almost partitioning 2-colored complete 3-uniform hypergraphs into two monochromatic tight or loose cycles. *J. Graph Theory*, 91(1):5–15, 2019.

[29] C. Chvatál, V. Rödl, E. Szemerédi, and W. T. Trotter, Jr. The Ramsey number of a graph with bounded maximum degree. *J. Combin. Theory Ser. B*, 34(3):239–243, 1983.

[30] D. Clemens, J. Ehrenmüller, and Y. Person. A Dirac-type theorem for Hamilton Berge cycles in random hypergraphs. In *Discrete mathematical days. Extended abstracts of the 10th "Jornadas de matemática discreta y algorítmica" (JMDA), Barcelona, Spain, July 6–8, 2016*, pages 181–186. Amsterdam: Elsevier, 2016.

[31] D. Conlon. A new upper bound for diagonal Ramsey numbers. *Ann. of Math. (2)*, 170(2):941–960, 2009.

[32] D. Conlon, J. Fox, and B. Sudakov. Ramsey numbers of sparse hypergraphs. *Random Structures Algorithms*, 35(1):1–14, 2009.

[33] D. Conlon, J. Fox, and Y. Zhao. A relative Szemerédi theorem. *Geom. Funct. Anal.*, 25(3):733–762, 2015.

[34] O. Cooley, N. Fountoulakis, D. Kühn, and D. Osthus. 3-uniform hypergraphs of bounded degree have linear Ramsey numbers. *J. Combin. Theory Ser. B*, 98(3):484–505, 2008.

[35] O. Cooley, N. Fountoulakis, D. Kühn, and D. Osthus. Embeddings and Ramsey numbers of sparse k-uniform hypergraphs. *Combinatorica*, 29(3):263–297, 2009.

[36] A. N. Day, V. Falgas-Ravry, and R. Hancock. Long paths and connectivity in 1-independent random graphs. *Random Structures Algorithms*, 57(4):1007–1049, 2020.

[37] N. Dean and M. Kouider. Gallai's conjecture for disconnected graphs. *Discrete Math.*, 213(1-3):43–54, 2000. Selected topics in discrete mathematics (Warsaw, 1996).

[38] L. DeBiasio and L. L. Nelsen. Monochromatic cycle partitions of graphs with large minimum degree. *J. Combin. Theory Ser. B*, 122:634–667, 2017.

[39] M. Deijfen, O. Häggström, and A. E. Holroyd. Percolation in invariant Poisson graphs with iid degrees. *Arkiv för matematik*, 50(1):41–58, 2012.

[40] M. Deijfen, A. E. Holroyd, and Y. Peres. Stable Poisson graphs in one dimension. *Electron. J. Probab.*, 16:no. 44, 1238–1253, 2011.

[41] G. A. Dirac. Some theorems on abstract graphs. *Proc. Lond. Math. Soc.*, 3(1):69–81, 1952.

[42] A. Dudek and A. Frieze. Loose Hamilton cycles in random uniform hypergraphs. *Electron. J. Combin.*, 18(1):P48, 2011.

[43] A. Dudek and A. Frieze. Tight Hamilton cycles in random uniform hypergraphs. *Random Structures Algorithms*, 42(3):374–385, 2013.

[44] P. Erdős and A. Rényi. On random graphs. I. *Publ. Math. Debrecen*, 6:290–297, 1959.

[45] P. Erdős, A. Gyárfás, and L. Pyber. Vertex coverings by monochromatic cycles and trees. *J. Combin. Theory Ser. B*, 51(1):90 – 95, 1991.

[46] V. Falgas-Ravry. *Thresholds in probabilistic and extremal combinatorics.* PhD thesis, University of London, 2012.

[47] V. Falgas-Ravry and V. Pfenninger. 1-independent percolation on $\mathbb{Z}^2 \times K_n$. *arXiv e-prints*, arXiv:2106.08674, 2021.

[48] R. J. Faudree and R. H. Schelp. All Ramsey numbers for cycles in graphs. *Discrete Math.*, 8:313–329, 1974.

[49] A. Ferber and L. Hirschfeld. Co-degrees resilience for perfect matchings in random hypergraphs. *Electron. J. Combin.*, 27:P1.40, 2020.

[50] A. Ferber and M. Kwan. Dirac-type theorems in random hypergraphs. *J. Combin. Theory Ser. B*, 155:318–357, 2022.

[51] A. Frieze. Hamilton Cycles in Random Graphs: a bibliography. *arXiv e-prints*, arXiv:1901.07139, 2019.

[52] Z. Füredi, T. Jiang, A. Kostochka, D. Mubayi, and J. Verstraëte. Tight paths in convex geometric hypergraphs. *Adv. Comb.*, pages Paper No. 1, 14, 2020.

[53] F. Garbe, R. Mycroft, R. Lang, A. Lo, and N. Sanhueza-Matamala. Partitioning 2-coloured complete 3-uniform hypergraphs into two monochromatic tight cycles. *In preparation.*

[54] L. Gerencsér and A. Gyárfás. On Ramsey-type problems. *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.*, 10:167–170, 1967.

[55] W. T. Gowers. Hypergraph regularity and the multidimensional Szemerédi theorem. *Ann. of Math.*, pages 897–946, 2007.

[56] G. Grimmett. *Percolation*, volume 321 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin, second edition, 1999.

[57] A. Gyárfás. Vertex coverings by monochromatic paths and cycles. *J. Graph Theory*, 7(1):131–135, 1983.

[58] A. Gyárfás. Vertex covers by monochromatic pieces — a survey of results and problems. *Discrete Math.*, 339(7):1970 – 1977, 2016.

[59] A. Gyárfás and J. Lehel. A Ramsey-type problem in directed and bipartite graphs. *Period. Math. Hungar.*, 3(3-4):299–304, 1973.

[60] A. Gyárfás and G. N. Sárközy. Monochromatic path and cycle partitions in hypergraphs. *Electron. J. Combin.*, 20(1):18, 2013.

[61] A. Gyárfás, G. N. Sárközy, and E. Szemerédi. The Ramsey number of diamond-matchings and loose cycles in hypergraphs. *Electron. J. Combin.*, 15(1):Research Paper 126, 14, 2008.

[62] A. Gyárfás, M. Ruszinkó, G. N. Sárközy, and E. Szemerédi. An improved bound for the monochromatic cycle partition number. *J. Combin. Theory Ser. B*, 96(6):855 – 873, 2006.

[63] J. Han, A. Lo, and N. Sanhueza-Matamala. Covering and tiling hypergraphs with tight cycles. *Combin. Probab. Comput.*, to appear.

[64] T. E. Harris. A lower bound for the critical probability in a certain percolation process. *Proc. Cambridge Philos. Soc.*, 56:13–20, 1960.

[65] P. E. Haxell, T. Ł uczak, Y. Peng, V. Rödl, A. Ruciński, M. Simonovits, and J. Skokan. The Ramsey number of hypergraph cycles. I. *J. Combin. Theory Ser. A*, 113(1):67–83, 2006.

[66] P. E. Haxell, T. Łuczak, Y. Peng, V. Rödl, A. Ruciński, and J. Skokan. The Ramsey number for hypergraph cycles II. *CDAM Research Report, LSE-CDAM-2007-04*, 2007.

[67] P. E. Haxell, T. Łuczak, Y. Peng, V. Rödl, A. Ruciński, and J. Skokan. The Ramsey number for 3-uniform tight hypergraph cycles. *Combin. Probab. Comput.*, 18(1-2):165–203, 2009.

[68] M. Jenssen and J. Skokan. Exact Ramsey numbers of odd cycles via nonlinear optimisation. *Adv. Math.*, 376:107444, 46, 2021.

[69] P. Keevash, E. Long, and J. Skokan. Cycle-complete Ramsey numbers. *Int. Math. Res. Not. IMRN*, (1):277–302, 2021.

[70] H. Kesten. The critical probability of bond percolation on the square lattice equals 1/2. *Comm. Math. Phys.*, 74(1):41–59, 1980.

[71] J. H. Kim and V. H. Vu. Concentration of multivariate polynomials and its applications. *Combinatorica*, 20(3):417–434, 2000.

[72] J. Komlós and E. Szemerédi. Limit distribution for the existence of Hamiltonian cycles in a random graph. *Discrete Math.*, 43(1):55–63, 1983.

[73] D. Korándi, R. Lang, S. Letzter, and A. Pokrovskiy. Minimum degree conditions for monochromatic cycle partitioning. *J. Combin. Theory Ser. B*, 146:96–123, 2021.

[74] A. D. Korshunov. Solution of a problem of Erdős and Renyi on Hamiltonian cycles in nonoriented graphs. *Sov. Math., Dokl.*, 17:760–764, 1976.

[75] A. D. Korshunov. Solution of a problem of P. Erdős and A. Renyi on Hamiltonian cycles in undirected graphs. *Metody Diskretn. Anal.*, 31:17–56, 1977.

[76] M. Krivelevich and B. Sudakov. Pseudo-random graphs. In *More sets, graphs and numbers*, pages 199–262. Springer, 2006.

[77] D. Kühn and D. Osthus. Hamilton cycles in graphs and hypergraphs: an extremal perspective. In *Proceedings of the International Congress of Mathematicians—Seoul 2014. Vol. IV*, pages 381–406. Kyung Moon Sa, Seoul, 2014.

[78] R. Lang and N. Sanhueza-Matamala. Minimum degree conditions for tight Hamilton cycles. *arXiv e-prints*, arXiv:2005.05291, 2020.

[79] C. Lee and B. Sudakov. Dirac's theorem for random graphs. *Random Structures Algorithms*, 41(3):293–305, 2012.

[80] S. Letzter. Monochromatic cycle partitions of 2-coloured graphs with minimum degree $3n/4$. *Electron. J. Combin.*, 26(1):Paper No. 1.19, 67, 2019.

[81] T. M. Liggett, R. H. Schonmann, and A. M. Stacey. Domination by product measures. *Ann. Probab.*, 25(1):71–95, 1997.

[82] A. Lo and V. Pfenninger. Towards Lehel's conjecture for 4-uniform tight cycles. *arXiv e-prints*, arXiv:2012.08875, 2020.

[83] A. Lo and V. Pfenninger. The Ramsey number for 4-uniform tight cycles. *arXiv e-prints*, arXiv:2111.05276, 2021.

[84] T. Łuczak. $R(C_n, C_n, C_n) \leq (4 + o(1))n$. *J. Combin. Theory Ser. B*, 75(2):174 – 187, 1999.

[85] T. Łuczak, V. Rödl, and E. Szemerédi. Partitioning two-coloured complete graphs into two monochromatic cycles. *Combin. Probab. Comput.*, 7(4):423–436, 1998.

[86] R. Lyons. Random walks and percolation on trees. *Ann. Probab.*, 18(3):931–958, 1990.

[87] W. Mantel. Problem 28. *Wiskundige Opgaven*, 10:60–61, 1907.

[88] R. Meester and R. Roy. *Continuum percolation*, volume 119 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1996.

[89] R. W. J. Meester. Uniqueness in percolation theory. *Statist. Neerlandica*, 48(3):237–252, 1994.

[90] A. Méroueh. The Ramsey number of loose cycles versus cliques. *J. Graph Theory*, 90(2):172–188, 2019.

[91] R. Montgomery. Hamiltonicity in random graphs is born resilient. *J. Combin. Theory Ser. B*, 139:316–341, 2019.

[92] D. Mubayi and V. Rödl. Hypergraph Ramsey numbers: tight cycles versus cliques. *Bull. Lond. Math. Soc.*, 48(1):127–134, 2016.

[93] B. Nagle, S. Olsen, V. Rödl, and M. Schacht. On the Ramsey number of sparse 3-graphs. *Graphs Combin.*, 24(3):205–228, 2008.

[94] B. Narayanan and M. Schacht. Sharp thresholds for nonlinear hamiltonian cycles in hypergraphs. *Random Structures Algorithms*, 57(1):244–255, 2020.

[95] R. Nenadov, A. Steger, and M. Trujić. Resilience of perfect matchings and hamiltonicity in random graph processes. *Random Structures Algorithms*, 54(4):797–819, 2019.

[96] J. Nie and J. Verstraëte. Ramsey numbers for nontrivial Berge cycles. *SIAM J. Discrete Math.*, 36(1):103–113, 2022.

[97] A. Pokrovskiy. Partitioning edge-coloured complete graphs into monochromatic cycles and paths. *J. Combin. Theory Ser. B*, 106:70 – 97, 2014.

[98] A. Pokrovskiy. Partitioning a graph into a cycle and a sparse graph. *arXiv e-prints*, arXiv:1607.03348, 2016.

[99] L. Pósa. Hamiltonian circuits in random graphs. *Discrete Math.*, 14(4):359–364, 1976.

[100] F. P. Ramsey. On a Problem of Formal Logic. *Proc. London Math. Soc. (2)*, 30(4):264–286, 1929.

[101] C. Reiher, V. Rödl, A. Ruciński, M. Schacht, and E. Szemerédi. Minimum vertex degree condition for tight hamiltonian cycles in 3-uniform hypergraphs. *Proc. Lond. Math. Soc.*, 119(2):409–439, 2019.

[102] O. Riordan and M. Walters. Rigorous confidence intervals for critical probabilities. *Phys. Rev. E*, 76:011110, Jul 2007.

[103] V. Rödl, A. Ruciński, and E. Szemerédi. A Dirac-type theorem for 3-uniform hypergraphs. *Combin. Probab. Comput.*, 15(1-2):229–251, 2006.

[104] V. Rödl, A. Ruciński, and E. Szemerédi. An approximate Dirac-type theorem for k-uniform hypergraphs. *Combinatorica*, 28(2):229–260, 2008.

[105] V. Rödl, A. Ruciński, and E. Szemerédi. Dirac-type conditions for Hamiltonian paths and cycles in 3-uniform hypergraphs. *Adv. Math.*, 227(3):1225–1299, 2011.

[106] V. Rödl and M. Schacht. Regular partitions of hypergraphs: regularity lemmas. *Combin. Probab. Comput.*, 16(6):833–885, 2007.

[107] V. Rödl and J. Skokan. Regularity lemma for k-uniform hypergraphs. *Random Structures Algorithms*, 25(1):1–42, 2004.

[108] V. Rosta. On a Ramsey-type problem of J. A. Bondy and P. Erdős. I, II. *J. Combin. Theory Ser. B*, 15:94–104; ibid. 15 (1973), 105–120, 1973.

[109] A. Sarkar and M. Haenggi. Percolation in the secrecy graph. *Discrete Appl. Math.*, 161(13-14):2120–2132, 2013.

[110] G. N. Sárközy. Improved monochromatic loose cycle partitions in hypergraphs. *Discrete Math.*, 334:52 – 62, 2014.

[111] M. Shahsiah. Ramsey numbers of 5-uniform loose cycles. *Graphs Combin.*, 38(1):Paper No. 5, 23, 2022.

[112] J. Spencer. Ramsey's theorem—a new lower bound. *J. Combinatorial Theory Ser. A*, 18:108–115, 1975.

[113] M. Stein. Monochromatic paths in 2-edge coloured graphs and hypergraphs. *arXiv e-prints*, arXiv:2204.12464, 2022.

[114] B. Sudakov and V. H. Vu. Local resilience of graphs. *Random Structures Algorithms*, 33(4):409–433, 2008.

[115] A. Thomason. Pseudo-random graphs. In M. Karoński, editor, *Proceedings of Random Graphs, Poznán 1985*, volume 33 of *Ann. Discrete Math.*, pages 307–331. North-Holland, 1987.

[116] J. van den Berg and A. Ermakov. A new lower bound for the critical probability of site percolation on the square lattice. *Random Structures Algorithms*, 8(3):199–212, 1996.

[117] J. C. Wierman. Substitution method critical probability bounds for the square lattice site percolation model. *Combin. Probab. Comput.*, 4(2):181–188, 1995.

[118] R. Wilson and J. J. Watkins, editors. *Combinatorics: ancient and modern.* Oxford University Press, Oxford, 2013.

[119] R. M. Ziff. Spanning probability in 2d percolation. *Phys. Rev. Lett.*, 69(18):2670, 1992.