

RESEARCH ARTICLE

Cycle type in Hall–Paige: a proof of the Friedlander–Gordon–Tannenbaum conjecture

Alp Müyesser 

University College London, UK; E-mail: alp.muyesser@new.ox.ac.uk.

Received: 28 July 2023; Revised: 18 December 2025; Accepted: 18 December 2025

2020 Mathematics Subject Classification: Primary – 05B15, 20P05

Abstract

An orthomorphism of a finite group G is a bijection $\phi: G \rightarrow G$ such that $g \mapsto g^{-1}\phi(g)$ is also a bijection. In 1981, Friedlander, Gordon, and Tannenbaum conjectured that when G is abelian, for any $k \geq 2$ dividing $|G| - 1$, there exists an orthomorphism of G fixing the identity and permuting the remaining elements as products of disjoint k -cycles as long as the Sylow 2-subgroups of G are trivial or noncyclic. We prove this conjecture for all sufficiently large groups.

1. Introduction

An **orthomorphism** of a finite group G is a bijection $\phi: G \rightarrow G$ such that $g \mapsto g^{-1}\phi(g)$ is also bijective. Orthomorphisms have attracted much interest in recent years, not least due to their link with Latin squares. The multiplication tables of groups with orthomorphisms yield Latin squares with orthogonal mates, which in turn give useful constructions in design theory (see the book of Evans [15] for an overview of the area). A fundamental conjecture in the area is the Hall–Paige conjecture [22] which states that a group G admits an orthomorphism if and only if the product of all elements in the group (in any order) belongs to the commutator subgroup (this property is henceforth referred to as the **Hall–Paige condition**, and is equivalent to the Sylow 2-subgroups of G being trivial or noncyclic). For abelian groups, the Hall–Paige condition simply means that the sum of all elements in the group is the identity. The conjecture was confirmed by Wilcox [36], Evans [14], and Bray [8] in 2009. It is not too difficult to see that the Hall–Paige condition is necessary, but the fact that it is also sufficient is quite remarkable.

The proof of Wilcox, Evans and Bray has the disadvantage that it relies extensively on the classification of finite simple groups. Recently, two new proofs of the Hall–Paige conjecture have been found which do not rely on this classification, with the caveat that both proofs require the group to be sufficiently large. On the other hand, both proofs strengthen the original statement of the Hall–Paige conjecture in a distinct, novel direction. The first of these proofs is due to Eberhard, Manners, and Mrazović [12]. This proof uses tools from analytic number theory, and it yields a strikingly accurate asymptotic on the number of orthomorphisms for groups with the Hall–Paige condition. The second proof is due to the author and Pokrovskiy [30], and this proof has the advantage of finding orthomorphisms in random-like subsets of groups. This flexibility turns out to be quite fruitful as demonstrated by the numerous applications of the ‘random Hall–Paige conjecture’ given in [30].

The current paper is focused on a third way to strengthen the Hall–Paige conjecture, this time by asserting the existence of orthomorphisms with specific cycle types. Recall that the **cycle type** of a

permutation π encodes how many cycles of each length are present when π is written as a product of disjoint cycles. For example, orthomorphisms that consist of a single cycle come up naturally in Ringel’s resolution of the Heawood map colouring conjecture, which motivated Ringel to ask for a classification of all groups with such orthomorphisms (see [34, 17, 2, 33]). Several other problems of a similar flavour concerning ‘sequenceable groups’ were raised by numerous authors with the motivation to construct Latin squares with additional properties (see [32] and Section 1.1.2 in [30]). There are also motivations to study orthomorphisms with other cycle types. For example, orthomorphisms that are products of disjoint 6-cycles give constructions such as ‘cyclic’ Steiner triple systems [25].

A unifying conjecture in the area was given by Friedlander, Gordon, and Tannenbaum in 1981 [16].

Conjecture 1.1 (The Friedlander–Gordon–Tannenbaum (FGT) conjecture, 1981). *Let G be an abelian group of order n satisfying the Hall–Paige condition. Suppose for some integer $k \geq 2$ that k divides $n - 1$. Then, there exists an orthomorphism of G that fixes the identity element, and permutes the remaining elements as products of disjoint cycles of length k .*

The Hall–Paige conjecture is not very laborious to verify for abelian groups, and this was already achieved by Hall and Paige when they posed their conjecture. The FGT conjecture, on the other hand, has remained open for more than forty years. There are several partial results towards the FGT conjecture in the literature. Friedlander, Gordon, and Tannenbaum themselves confirmed their conjecture for groups of order at most 15, and abelian p -groups where $p \geq 3$ [16]. We refer the reader to [15] for a more detailed overview (see also [5, 6, 35] for results about the very related concept of complete mappings). We just remark that the $k = 3$ and the cyclic group case of the FGT conjecture is open, signifying the difficulty of the problem. In this paper, we resolve the FGT conjecture for sufficiently large groups.

Theorem 1.2. *The Friedlander–Gordon–Tannenbaum conjecture is true for all sufficiently large groups.*

We use methods from probabilistic combinatorics, so our proof needs large groups just to get concentration for some random variables with fairly simple distributions. We do not make this constant explicit to make the presentation neater.

We verify the FGT conjecture by developing fairly general methods which can potentially be used to shed light on many adjacent embedding problems with an algebraic flavour. As we explore in Section 7.1, our methods seem adaptable for the study of graceful and harmonious graph labellings [18]. We delve further into the proof in the next section, which serves as a skeleton for the paper and contains a bird’s eye view of the proof of Theorem 1.2.

We make three further remarks.

Remark 1.3. At the time the FGT conjecture was posed, the Hall–Paige conjecture was known to be true for abelian groups, but not in general, which perhaps explains why Conjecture 1.1 is concerned only with abelian groups. Given the present work, it seems reasonable to suspect that the FGT conjecture can be extended to nonabelian groups, perhaps even quasi-groups/Latin squares, which would generalise the famous Ryser–Brualdi–Stein conjecture. We discuss this further in the concluding remarks, Section 7.

Remark 1.4. Our proof of Theorem 1.2 actually gives much more, and can be used to give many other cycle types that can be realised via orthomorphisms. We discuss this further in Section 7.

Remark 1.5. A very related notion is that of a **complete mapping**, which is a permutation ϕ of a group G such that $g \rightarrow g\phi(g)$ is also bijective. A group admits a complete mapping if and only if it admits an orthomorphism, essentially because the map $g \rightarrow g^{-1}$ is a bijection. Therefore, the Hall–Paige conjecture is sometimes stated with respect to complete mappings instead of orthomorphisms. However, this equivalence does not hold when we make restrictions on the cycle type. For example, in an abelian group, there cannot be a complete mapping inducing any cycle of length 2, therefore the FGT conjecture does not hold when orthomorphisms are replaced with complete mappings (for a more detailed discussion of cycle types of complete mappings, see [5, 6]). However, some appropriate modification of the FGT conjecture likely holds for complete mappings as well, and we discuss this further in Section 7. We should also remark that, confusingly, orthomorphisms are called complete mappings in [16], but the convention in the current paper seems to be standard following the book of Evans [15].

2. Main theorem and overview of the proof

2.1. Definitions of key auxiliary graphs and hypergraphs

It is customary in combinatorics to rephrase statements such as Conjecture 1.1 in terms of finding perfect matchings in hypergraphs, or finding rainbow structures in edge-coloured graphs, and we follow this tradition in the current paper.

Given a group G of order n , we denote by \vec{K}_G the edge-coloured directed graph defined as follows. $V(\vec{K}_G) := G$, and $E(\vec{K}_G) := \{(a, b) \in G \times G : a \neq b\}$, and the *colour* of an edge (a, b) is the group element ab^{-1} . Given subsets $V, C \subseteq G$, by $\vec{K}_G[V; C]$ we denote the subgraph of \vec{K}_G obtained by keeping only the vertices in V , and the directed edges with colours in C . Occasionally, the following related definition will also be useful. Given multiple subsets $V_1, V_2, \dots, V_k \subseteq G$, we denote by $\vec{K}_G[V_1, V_2, \dots, V_k]$ the edge-coloured directed graph with vertex set $V_1 \sqcup V_2 \sqcup \dots \sqcup V_k$ (\sqcup indicates that we are taking a disjoint union) and edge set consisting of edges of the form $e = (v, w) \in V_i \times V_{i+1}$ (with colour vw^{-1}) for some $i \in \{1, 2, \dots, k\}$ (where $k + 1 = 1$). By $\vec{K}_G[V_1, V_2, \dots, V_k; C]$, we denote the same graph obtained by keeping only edges whose colour is in C .

Recall that a subgraph of an edge-coloured graph is called **rainbow** if all edges have distinct colours. Given $V, C \subseteq G$, let $\mathcal{H}_k[V; C]$ be the $2k$ -uniform hypergraph on the vertex set $V \sqcup C$ where $v \sqcup c$ is an edge whenever $v \subseteq V$ induces a rainbow directed cycle of length k in \vec{K}_G with the colour set of the cycle being precisely c . \mathcal{H}_k denotes $\mathcal{H}_k[G; G]$. Sometimes we overload the terms vertex and colour by referring to elements of $V(\mathcal{H}_k[V; C])$ which come from V as **vertices** and those which come from C as **colours**. The following observation is quite critical.

Observation 2.1. Assuming the ambient group G is abelian, if c is the colour set of an edge in \mathcal{H}_k , or the colour set of some directed rainbow cycle in \vec{K}_G (of any length), then the sum of all the elements of c must equal 0, that is, c is a zero-sum set.

Proof. As in a directed cycle each vertex has one in-edge and one out-edge, when we take a sum of all the colours of a cycle in \vec{K}_G , each vertex appears twice, once positive, and once negative. The statement follows. \square

Given graphs H and G , we say that G contains an **H -factor** if there exists a collection of copies of H in G that partition the vertex set of G . For example, a K_2 -factor in a graph is a perfect matching. \vec{P}_k denotes a directed path of length k (meaning with k edges). \vec{C}_k denotes a directed cycle of length k (meaning with k vertices and k edges). The following proposition follows from all the definitions presented thus far.

Proposition 2.2. Let G be a finite abelian group and let k be an integer with $k \geq 2$. The following are equivalent.

- G admits an orthomorphism fixing the identity and permuting the remaining elements as products of disjoint k -cycles.
- $\vec{K}_G[G \setminus \{0\}; G \setminus \{0\}]$ contains a rainbow \vec{C}_k -factor.
- $\mathcal{H}_k[G \setminus \{0\}; G \setminus \{0\}]$ has a perfect matching.

We invite the reader to verify the above proposition. Thanks to Proposition 2.2, we can phrase our main result in the language of hypergraph matchings in the next subsection.

2.2. Main theorem and its proof modulo key lemmas

The $k = 2$ case of Conjecture 1.1 is proven implicitly in [16], where the authors give orthomorphisms of odd-order cyclic groups which are products of disjoint transpositions (see also [15] for a proof of the $k = 2$ case). The case of $k > \log^{10} n$ (the ‘high-girth case’), on the other hand, can be resolved by using some tools from [30]. In fact, the $k = n$ case (the Hamilton cycle case) was implicitly solved in [30] already, and it turns out the method is general enough to handle cycles of length at least polylogarithmic in n . We give the details for this in Section 6. We remark that for the methods of [30], this polylogarithmic

lower bound on the cycle length is a hard barrier, essentially because any sorting network (see [4, 1]) must have depth at least $\log n$ (we discuss this in more detail later on). Hence, our main theorem is concerned with the $3 \leq k \leq \log^{10} n$ case of the FGT conjecture.

We recall the following conventions before stating our main theorem. Recall that a **p -random** subset of set S is one obtained by sampling each element of S independently with probability p . Similarly, we say a collection of random sets $R_1, \dots, R_k \subseteq S$ is **disjoint p -random** if each element of S belongs to each R_i with probability p , and to none of the R_i with probability $1 - pk$, and these decisions are made independently for each element of S . We reserve the letter n for the size of the ambient group throughout the paper. When we say that an event holds ‘with high probability’, we mean that the probability of the event approaches 1 as n tends to infinity.

Theorem 2.3 (Main theorem). *There exists an absolute constant $\varepsilon_{2.3} > 0$ such that the following holds. Let G be an abelian group of order n , let $p \geq n^{-\varepsilon_{2.3}}$, and suppose k is some integer such that $3 \leq k \leq \log^{10} n$. Let $R_1, R_2 \subseteq G$ be p -random subsets, sampled independently. Then, the following holds with high probability.*

Let $V, C \subseteq G$ be equal-sized subsets with $|V \Delta R_1|, |C \Delta R_2| \leq n^{3/4}$. Suppose k divides $|V|$ (and thus, $|C|$), and suppose $\sum C = 0$ and $0 \notin C$. Then, $\mathcal{H}_k[V; C]$ has a perfect matching.

Theorem 2.3 turns into a deterministic statement when applied with $p = 1$. This statement, when n is sufficiently large, implies Conjecture 1.1 (when $3 \leq k \leq \log^{10} n$, the ‘low-girth case’) by setting $V = C = G \setminus \{0\}$. Theorem 2.3 can thus be interpreted as a randomised version of Conjecture 1.1. As far as our proof method is concerned, it does not take extra work to prove Theorem 2.3 compared to Conjecture 1.1. Theorem 2.3 also has further applications. Using its full strength, one can find orthomorphisms with other cycle types, see Section 7 for more details.

2.2.1. Comparison with the random Hall–Paige conjecture

It is worth discussing how Theorem 2.3 is different from the main result in [30] (see Theorem 3.20 in the current paper) and the applications given therein. The main result of [30], or the so-called random Hall–Paige conjecture, is concerned with finding perfect matchings in hypergraphs whose vertices are group elements, and edges are given by triples (a, b, c) where $a + b + c = 0$. In comparison, the hypergraph in Theorem 2.3, for example when $k = 3$, is defined by 6-tuples (a, b, c, d, e, f) where $a = d - e$, $b = e - f$, $c = f - d$ (note this implies in particular that $a + b + c = 0$). With extra ideas, it is sometimes possible to glue together simpler equations to deduce information about hypergraphs with more complicated structure. For example, the random Hall–Paige conjecture guarantees existence of perfect rainbow matchings in random subsets of \vec{K}_G . Collecting multiple disjoint random sets, and finding rainbow matchings between them (using disjoint colour sets), we can build long rainbow path forests in \vec{K}_G . However, when the lengths of the rainbow paths approach k , we need to ‘stitch’ the endpoints of these rainbow paths together in order to obtain k -cycles. It turns out that one can set aside some structure in the beginning which allows us to stitch arbitrary endpoints together, using a specified set of vertices and colours (see Lemma 6.18 from [30]). The catch is that for the stitching, we need to use paths of length at least $\log n$. This barrier comes from, roughly speaking, the stitching process being equivalent to running a sorting algorithm. Due to limitations on how many comparisons any sorting algorithm has to make, there unfortunately is no room for improvement for this part of the argument, and therefore we need novel ideas.

In fact, it turns out that when k is small (say $k = 3$), the stitching statement we require (corresponding to Lemma 6.1 when $t = 2$) simply becomes false, in the sense that conditions coming from simply summing elements together, as in the Hall–Paige conjecture, are not sufficient to characterise the existence of the desired spanning structure. This is closely related to obstructions which arise in the toroidal n -queens problem for which Bowtell and Keevash recently made a break-through [7]. We won’t give much context for this problem here, mentioning only that the problem is about finding perfect matchings in hypergraphs where edges are of the form (a, b, c, d) where $c = a + b$ and $d = a - b$, similar in spirit to the FGT conjecture. In order to characterise when such perfect matchings arise, even

in cyclic groups, one needs a condition about the sum of *squares* in the group. This is in stark contrast with both the Hall–Paige and FGT conjectures, where the obstructions can be expressed using only the additive structure of the group in question.

2.2.2. Proof ideas

As outlined before, the low-girth case of the FGT conjecture is highly sensitive, and several novel ideas are required compared to the random Hall–Paige conjecture from [30]. To illustrate just one specific challenge we need to overcome, consider the following question: For which k -subsets $C \subseteq G$ does \mathcal{H}_k have an edge with C as its colour set? Equivalently, which colour sets induce directed rainbow k -cycles in \vec{K}_G ? This is highly relevant as a necessary condition for solving the FGT conjecture is coming up with a *partition* of $G \setminus \{0\}$ into such colour sets. It is not hard to see that we need that $\sum C = 0$ (see Observation 2.1) and that $0 \notin C$. Further thought reveals that to find a directed rainbow k -cycle (and not just a closed k -walk) in \vec{K}_G with colour set C , we need to be able to order C as (c_1, \dots, c_k) such that $c_1, c_1 + c_2, c_1 + \dots + c_k$ are all distinct. We call (c_1, \dots, c_k) a **cycle-candidate** if this property holds. Can a zero-sum subset not containing 0 be ordered to be a cycle-candidate? It turns out that even for cyclic groups of prime order, this is an open problem, posed initially by Ronald Graham in 1971. Surprisingly little is known about this problem (see Problem 10 from [20], see also [3, 11, 10, 9]). For example, the problem is already open for $k = 13$ and cyclic groups of prime order.

Due to difficulties surrounding Graham’s problem, our understanding of the structure of sequences yielding cycle-candidates is rather limited. A central aspect of our approach (which we believe translates well to adjacent problems, see Section 7.1) capitalises on the fact that there is a rich and well-behaved subset of cycle-candidates, namely, those which can be constructed by gluing together short, carefully curated families of subsequences (see for example the definition of *dissociable* in Section 3.2, as well as Lemma 3.27). A key tension in the proof is that the definitions of these building blocks need to be strong enough so that when they are combined, they produce cycle-candidates, but also weak enough that they exist in abundance throughout G . The latter need arises because we will have to construct *absorbers* for these building blocks from the beginning (see Lemma 5.19), in order to have some flexibility in how we use them towards the end of the proof.

Absorption is a general framework in probabilistic combinatorics designed to turn almost spanning structures into fully spanning structures (see [29, 7, 27] for some recent breakthroughs using this framework). The method has been developed immensely in the past decade, making it difficult to pinpoint exactly what the absorption method is. However, the common denominator in all implementations of the method is relying on the existence of small scale structures (gadgets) with various properties in order to build a large scale, flexible structure. A relevant question is which small scale structures can even be found in the host structure. Lemma 3.19 is a key tool we use in the current paper which gives a general method to produce *substructure patterns* which exist in abundance throughout host structures which are algebraically defined. We refer the reader to Figure 1 for examples of such patterns we work with in the current paper.

We emphasise that this result is quite flexible (in particular, it is possible to extend Definition 3.10 with further sufficient conditions if a future problem requires a richer class of substructures). We refer the reader to Section 7.1 for a discussion about a potential connection with the study of graceful and harmonious labellings.

2.2.3. Proof of Theorem 2.3

For the rest of this section, we focus on Theorem 2.3, which is concerned with the ‘low-girth’ case of the FGT conjecture. The key lemma used to prove Theorem 2.3 is the following, which states the existence of an ‘absorber for zero-sum subsets’. Roughly speaking, this lemma states that random subsets contain ‘absorbers’ which have the ability to combine with any small enough set to produce matchings (we say that the small set is ‘absorbed’), provided that this small set satisfies some straightforward necessary conditions.

Lemma 2.4 (Zero-sum absorption). *There exist absolute constants $\varepsilon = \varepsilon_{2.4} > 0$ and $K = K_{2.4} \geq 1$ with $\varepsilon K \leq 10^{-10}$ such that the following holds. Let $3 \leq k \leq \log^{10} n$, $p \geq n^{-\varepsilon}$. Let $R_1, R_2 \subseteq G$ be p -random subsets, sampled independently. Let $m \in k \cdot \mathbb{N}$ with $m \leq (p/k \log n)^K n$. Then, the following holds with high probability.*

Let $U \subseteq G$ with $|U| \leq n^{4/5}$. Then, there exist $V \subseteq R_1 \setminus U$ and $C \subseteq R_2 \setminus U$ with the following property. For any $V' \subseteq G \setminus V$ and $C' \subseteq G \setminus C$ with $|V'| = |C'| = m$, $\sum C' = 0$, $0 \notin C'$, we have that $\vec{K}_G[V \cup V'; C \cup C']$ has a rainbow \vec{C}_k -factor, or equivalently, $\mathcal{H}_k[V \cup V'; C \cup C']$ has a perfect matching.

To finish, we need a version of Rödl nibble that works with regular hypergraphs plus a few ‘junk vertices’, which correspond to the leftover in the smaller random sets after the absorber is removed. The following lemma encapsulates this idea.

Lemma 2.5. *There exists an absolute constant $\varepsilon_{2.5} > 0$ such that the following holds. Let G be an abelian group of order n . Let $3 \leq k \leq \log^{10} n$, and let $p \geq n^{-\varepsilon_{2.5}}$. Let R_1 and R_2 be p -random subsets of G , sampled independently. The following holds with high probability.*

For any $V_D, C_D \subseteq G$ disjoint with R_1 and R_2 (respectively) with $n^{99/1000} \leq |V_D| = |C_D| \leq \varepsilon_{2.5} p^3 n / k^{100}$, $\mathcal{H}_k[R_1 \cup V_D; R_2 \cup C_D]$ contains a matching covering all but at most $n^{1-1/10^8}$ vertices.

The proof of Lemma 2.5 comes down to establishing certain pseudorandomness properties of \mathcal{H}_k . Checking pseudorandomness in hypergraphs is notoriously tricky, for example see [23] for a useful criterion for dense hypergraphs. Unfortunately, \mathcal{H}_k is quite sparse, and potentially has large uniformity, so [23] is not immediately useful in our set-up. For this reason, we have to put a fair bit of care into the proof of Lemma 2.5.

We can now give the proof of our main theorem, assuming these two lemmas. We remark that often in our proofs, we have random subsets $R' \subseteq R$ where R itself is a random subset of the group G . When we say that R' is a q -random subset, we always mean that R' is a q -random subset of the group G , and not of R .

Proof of Theorem 2.3. Pick a value of $\varepsilon_{2.3}$ such that $0 < \varepsilon_{2.3} \ll \varepsilon_{2.4}, \varepsilon_{2.5}, 1/K_{2.4}$. For each $i \in \{1, 2\}$, partition R_i into $R_i^{(1)}$ and $R_i^{(2)}$ which are disjoint p_1 -random and p_2 -random sets respectively, where $p_1 = (1/1000)\varepsilon_{2.5} p^4 / k^{100}$ (and $p_2 = p - p_1$). We have that $p_1 \geq n^{-\varepsilon_{2.4}}$ and $p_2 \geq n^{\varepsilon_{2.5}}$ if $\varepsilon_{2.3}$ is small enough. Select some $m \in k \cdot \mathbb{N}$ such that $10n^{1-1/10^8} \leq m \leq (p_1/k \log n)^{K_{2.4}n}$ (there exist such values of m as $\varepsilon_{2.4} K_{2.4} \leq 10^{-10}$). With high probability, Lemma 2.4 holds with $(R_1^{(1)}, R_2^{(1)}, p_1, m)$ in place of (R_1, R_2, p, m) and Lemma 2.5 holds with $(R_1^{(2)}, R_2^{(2)}, p_2)$ in place of (R_1, R_2, p) . Also with high probability, the size of each random set is at most $\sqrt{n} \log n$ away from its expectation (by Chernoff’s bound, see Lemma 3.1). With high probability, all of these properties hold simultaneously.

Now, fix random sets having all these properties and let V and C be given as in the statement of the theorem. Set $U := (R_1 \setminus V) \cup (R_2 \setminus C)$ noting $|U| \leq 2n^{3/4}$. Apply Lemma 2.4 to find absorbing subsets $V_A \subseteq R_1^{(1)} \setminus U \subseteq V$ and $C_A \subseteq R_2^{(1)} \setminus U \subseteq C$ which can combine with m -sized vertex-sets and m -sized zero-sum colour-sets to produce perfect matchings. Note this implies in particular that $|V_A| = |C_A|$.

Set $V_D := (R_1^{(1)} \cap V) \setminus V_A$ and $C_D := (R_2^{(1)} \cap C) \setminus C_A$, noting $|V_D|, |C_D| \leq 3p_1 n$ (as this bounds from above even the sizes of $R_1^{(1)}$ and $R_2^{(1)}$) and also note that $3p_1 n \leq \varepsilon_{2.5} p^4 n / k^{100}$ (using the definition of p_1 and that $p_2 \geq p/2$). Note also that $||V_D| - |C_D|| \leq 10n^{3/4}$ so V_D and C_D are nearly equal sized¹. By an application of Lemma 2.5, we may deduce that $\mathcal{H}_k[V_D \cup (R_1^{(2)} \cap V); C_D \cup (R_2^{(2)} \cap C)]$ has a matching M_1 covering all but at most $2n^{1-1/10^8}$ vertices². Unmatch some edges of M_1 so that the number of leftover vertices $V' := V \setminus V(M_1)$ and colours $C' := C \setminus V(M_1)$ are both equal to m (this is possible as

¹This holds because $|V_A| = |C_A|$, V, R_1 and C, R_2 have small symmetric difference by assumption of the theorem, and we assume the random sets deviate not so much from their expected size

²To see this, first delete $\leq 10n^{3/4}$ elements from C_D or V_D so that $|V_D| = |C_D|$, add a minimal number of dummy elements to both V_D and C_D so that their size is at least $n^{99/1000}$ (but still at most $\varepsilon_{2.5} p^3 n / k^{100}$), and then we may find directly by Lemma 2.5 a matching inside $\mathcal{H}_k[V_D \cup (R_1^{(2)} \cap V); C_D \cup (R_2^{(2)} \cap C)]$ leaving at most $n^{1-1/10^8}$ uncovered vertices. All edges of this matching exist also in $\mathcal{H}_k[V_D \cup (R_1^{(2)} \cap V); C_D \cup (R_2^{(2)} \cap C)]$, with the exception of those edges incident on the ($\leq 10n^{3/4}$) vertices we deleted from V_D or V_C to balance the sets, the ($\leq n^{99/1000}$) dummy vertices added to V_D and C_D and the edges

k divides $|V|, |C|, m$ and $|V_A| = |C_A|$, and m is at least twice as large as the number of leftover vertices). Note that the matching M_1 guarantees that all colours in $C \setminus C'$ admit a partition C_1, C_2, \dots where each C_i is the colour set of a rainbow cycle in \vec{K}_G , meaning that $\sum C_i = 0$ for each i (see Observation 2.1). So we must have $\sum C \setminus C' = 0$ also. As $\sum C = 0$ by assumption, this implies that $\sum C' = 0$, so we can invoke the property coming from Lemma 2.4. This means that V' and C' combine with V_A and C_A to produce a matching, M_2 . $M_1 \cup M_2$ is then the desired perfect matching. \square

2.3. Organisation of the rest of the paper

We collect some preliminary tools in Section 3. We have already broken up the task of proving Theorem 2.3 into proving Lemma 2.5 and Lemma 2.4. The former is proved in Section 4 and the latter is proved in Section 5. In Section 6, we show how the high-girth case of the FGT conjecture can be derived from results from [30], as promised earlier on in this section. In Section 7, we discuss some directions for future research.

3. Preliminaries

3.1. Probabilistic tools

3.1.1. Concentration inequalities

We need the following two basic concentration inequalities. We will refer to the following as Chernoff’s bound.

Lemma 3.1 (Chernoff bound). *Let $X := \sum_{i=1}^m X_i$ where $(X_i)_{i \in [m]}$ is a sequence of independent indicator random variables with $\mathbb{P}(X_i = 1) = p_i$. Let $\mathbb{E}[X] = \mu$. Then, for any $0 < \gamma < 1$, we have that $\mathbb{P}(|X - \mu| \geq \gamma\mu) \leq 2e^{-\mu\gamma^2/3}$.*

We use the following corollary of Chernoff’s bound often: that if R is a p -random subset of an n -element set, then with high probability we have that $|pn - |R|| \leq \log n\sqrt{n}$.

Sometimes the random variables we consider have slight dependencies. In this case, we rely on Azuma’s inequality which we now cite. Given a product probability space $\Omega = \prod_{i \in [n]} \Omega_i$, a random variable $X : \Omega \rightarrow \mathbb{R}$ is called **C -Lipschitz** if $|X(\omega) - X(\omega')| \leq C$ whenever ω and ω' differ in at most 1-coordinate.

Lemma 3.2 (Azuma’s inequality). *Let X be a C -Lipschitz random variable on a product probability space with n coordinates. Then, for any $t > 0$,*

$$\mathbb{P}(|X - \mathbb{E}(X)| > t) \leq 2e^{-\frac{t^2}{nC^2}}.$$

3.1.2. Nibble-type lemmas

We say that a r -partite r -uniform hypergraph H is (γ, p, n, k) -**regular** if every part has $(1 \pm \gamma)n$ vertices and every vertex has degree $(1 \pm \gamma)pn^k$. For a 3-uniform 3-partite hypergraph H , vertices u, v and a subset $U \subseteq V(H)$, we define the **pair degree** of (u, v) into U as the number of vertices z in U which are in the neighbourhood of both u and v , that is, the number of vertices z in U such that there exists $y, w \in V(H)$ such that $\{u, z, y\}$ and $\{v, z, w\}$ are both edges of H . A 3-uniform 3-partite hypergraph H is (γ, p, n) -**typical** if it is $(\gamma, p, n, 1)$ -**regular** and every pair of vertices u, v coming from the same part has pair degree $(1 \pm \gamma)p^2n$ into every other part of H . A hypergraph is **linear** if its maximum co-degree at most 1.

The following nibble-type result due to Ehard, Glock, and Joos is convenient to use for our application here.

incident on some vertex or colour of U which has size at most $2n^{3/4}$. Deleting such edges from the matching uncovers at most $12kn^{3/4} + 2n^{999/1000} \ll n^{1-1/10^8}$ vertices as $k \leq \log^{10}(n)$ and n is large, and this yields a matching of the desired form.

Theorem 3.3 [13]. *Suppose $\delta \in (0, 1)$ and $r \in \mathbb{N}$ with $r \geq 2$, and let $\varepsilon := \delta/50r^2$. Then there exists Δ_0 such that for all $\Delta \geq \Delta_0$, the following holds. Let \mathcal{H} be an r -uniform hypergraph with $\Delta(\mathcal{H}) \leq \Delta$ and $\Delta^c(\mathcal{H}) \leq \Delta^{1-\delta}$ as well as $e(\mathcal{H}) \leq \exp(\Delta^{\varepsilon^2})$. Suppose that \mathcal{W} is a set of at most $\exp(\Delta^{\varepsilon^2})$ weight functions on $E(\mathcal{H})$. Then, there exists a matching \mathcal{M} in \mathcal{H} such that $\omega(\mathcal{M}) = (1 \pm \Delta^{-\varepsilon})\omega(E(\mathcal{H}))/\Delta$ for all $\omega \in \mathcal{W}$ with $\omega(E(\mathcal{H})) \geq \max_{e \in E(\mathcal{H})} \omega(e)\Delta^{1+\delta}$.*

Applying Theorem 3.3 with a single uniform weight function, we obtain the following.

Corollary 3.4. *Let n be sufficiently large and let $p \geq n^{-1/10000}$.*

1. *Let \mathcal{H} be a 6-uniform 6-partite hypergraph on n vertices which is $(n^{-0.01}, p, n, 2)$ -regular with maximum co-degree at most $10n$. Then, \mathcal{H} has a matching covering all but $n^{1-10^{-5}}$ vertices.*
2. *For any $\gamma \geq 0$, every (γ, δ, n) -regular linear tripartite hypergraph has a matching covering all but at most $n^{1-1/500} + 3\gamma n$ vertices.*

The below lemma allows us to incorporate some nonrandom vertices/colours into the nibble process. It unfortunately does not directly imply Lemma 2.5, but it will be an important ingredient in its proof. It was proved in [30, Lemma 6.20].

Lemma 3.5. *There exists $C = C_{3.5} \geq 10$ sufficiently large so that the following holds. Let $1/n \ll \gamma$, and let $1 \geq a, b, c \geq 1/\log^C n$. Set $m := \max\{an, bn, cn\}$ and let $\zeta \in [0, 1]$ be such that $1/m \leq \zeta^C/C$ and $\zeta \leq \min\{a, b, c\}/100$. Suppose $\ell \geq m - m^{1-\gamma}$ and setting $(x, y, z) := (\ell - an, \ell - bn, \ell - cn)$, suppose that $x + y \leq cn/2$, $x + z \leq bn/2$ and $y + z \leq an/2$. Let $A, B, C \subseteq G$ be a, b, c -random subsets of G respectively, sampled with A and B disjoint, and C independent of A, B . Then, with probability at least $1 - 1/n^{2.5}$ the following holds.*

Let $A', B', C' \subseteq G$ with $|B \setminus B'|, |A \setminus A'|, |C \setminus C'| \leq n^{1-\gamma}$, $(1 - \zeta)|C'| = |A'| = |B'| = \ell$. Then, there is a perfect directed C' -matching in $K_G^\pm[A', B'; C']$.

3.2. Group theoretic tools

Given a sequence $\vec{c} = (c_1, c_2, \dots, c_k)$ of group elements, and another group element v , we define the following sequences:

- $P_{out}(v, \vec{c}) := (v, v - c_1, v - c_1 - c_2, \dots, v - c_1 - c_2 - \dots - c_k)$
- $P_{in}(v, \vec{c}) := (v, v + c_1, v + c_1 + c_2, \dots, v + c_1 + c_2 + \dots + c_k)$

Observe that in \vec{K}_G , $P_{out}(v, \vec{c})$ denotes the vertex sequence obtained by starting a walk from v , and following the out-edges given by the sequence \vec{c} . $P_{in}(v, \vec{c})$ is analogous, except it follows the in-edges.

We call a sequence of group elements $\vec{c} = (c_1, c_2, \dots, c_k)$ a **path-candidate** if the partial sums $\sum_{i \in [j]} c_i$ for each $j \in [k]$ (including $j = 0$) are all distinct. Equivalently, all nonempty partial sums (of consecutive elements) are nonzero. Observe that \vec{c} being a path-candidate simply means that for any vertex $v \in G$, the walks $P_{out}(v, \vec{c})$ and $P_{in}(v, \vec{c})$ both give paths in \vec{K}_G .

We call a sequence of group elements $\vec{c} = (c_1, c_2, \dots, c_k)$ a **cycle-candidate** if $(c_1, c_2, \dots, c_{k-1})$ is a path-candidate and $\sum_{i \in [k]} c_i = 0$. This means that $P_{out}(v, \vec{c})$ and $P_{in}(v, \vec{c})$ both give cycles (of length k) in \vec{K}_G .

We call a sequence of group elements **rainbow** if all coordinates are distinct. Notice that a necessary condition for solving the FGT conjecture is a partition of $G \setminus \{0\}$ into rainbow cycle-candidates, each of length k .

The following two definitions only come up in the cover-down strategy for $k \geq 10$.

We call a collection of length k sequences **dissociable** if for any two distinct sequences $\vec{c} = (c_1, c_2, \dots, c_k)$ and $\vec{b} = (b_1, b_2, \dots, b_k)$ and $j, j' \in [k]$, $\sum_{i \in [j]} c_i \neq \sum_{i \in [j']} b_i$. This means that $P_{out}(v, \vec{c})$ and $P_{out}(v, \vec{b})$ are disjoint except on v . We call such a collection **near-dissociable** if the previous property holds for each $j, j' \leq k - 1$ (or equivalently, the sequences obtained by removing the last element from each tuple gives a dissociable family). This means that the corresponding directed walks are disjoint except on the endpoints.

We call two length k sequences (c_1, c_2, \dots, c_k) and $\vec{b} = (b_1, b_2, \dots, b_k)$ **separable at distance d** if for all $j, j' \in [k]$, $\sum_{i \in [j]} c_i + \sum_{i \in [j']} b_i \notin \{-d, d\}$. This means that for v and w where $v - w = d$, $P_{out}(v, \vec{c})$ and $P_{in}(w, \vec{b})$ are disjoint (except potentially on v or w).

The following simple lemma is key to the gadget finding strategy presented in Section 3.2.1.

Lemma 3.6. *Let G be abelian of order n . Then either the map $x \rightarrow 2x$ or the map $x \rightarrow 3x$ has an image of size at least $n^{1/2}$.*

Proof. Denote by G_k the size of the image of the homomorphism $x \rightarrow k \cdot x$. Denote by $\ker(k) := \{x \in G : k \cdot x = 0\}$. Note that $\ker(2) \cap \ker(3) = \{0\}$. Also, $\ker(2) + \ker(3)$ is a subgroup of G of size $|\ker(2)| \cdot |\ker(3)|$, so $|\ker(2)| \cdot |\ker(3)| \leq n$, so one of $\ker(2)$ or $\ker(3)$ must be at most $n^{1/2}$. By the first isomorphism theorem, $G_k = |G|/|\ker(k)|$, implying the claim. \square

3.2.1. Finding gadgets

In this section, we refine some tools from Section 3.6 of [30] in order to prove Lemma 3.19, a versatile tool to find small substructures in \vec{K}_G .

By F_k , we denote the free abelian group on k generators, the free variables are denoted as v_1, \dots, v_k (recall that $F_k \cong \mathbb{Z}^k$). $G * F_k$ denotes the free product, and $(G * F_k)^{ab}$ denotes the abelianization of the free product (recall that the abelianization G^{ab} of a group G is defined by the property that any homomorphism $G \rightarrow H$ where H is abelian factors uniquely through G^{ab}). A **word** is simply an element of $(G * F_k)^{ab}$. As all groups G are abelian in this paper, $(G * F_k)^{ab} \cong G \times F_k$, where the latter denotes a direct product. However, the former perspective makes it clear that each word w can be represented as

$$w = z_1 \cdot v_1 + \dots + z_k \cdot v_k + g$$

where each v_i is a free variable, each z_i is a (nonzero) integer, and $g \in G$, and this representation is unique up to reordering the summands. So w can also be viewed as an affine linear form $G^k \rightarrow G$.

A word w is **constant** if $w \in G$, that is, w does not include any free variables. We say that z_i is the **coefficient** of v_i . We say that w is **linear in v_i** if $z_i \in \{1, -1\}$. We say that w is **linear** if each $z_i \in \{1, -1\}$, and w is not constant. That is, w is linear in each free variable, and there exists at least one free variable in w . The **length** of a word is the sum $1 + \sum |z_i|$.

A homomorphism $\pi : (G * F_k)^{ab} \rightarrow G$ is a **projection** if $\pi(g) = g$ for all $g \in G$. We show two basic properties of projections. We remind the reader that throughout, G is a finite abelian group of order n .

Lemma 3.7. *For each function $f : \{v_1, \dots, v_k\} \rightarrow G$, there is precisely one projection $\pi_f : (G * F_k)^{ab} \rightarrow G$ which agrees with f on $\{v_1, \dots, v_k\}$. In particular, there are precisely n^k projections $(G * F_k)^{ab} \rightarrow G$.*

Proof. By the universal property of free abelian groups, there is a unique homomorphism $g : F_k \rightarrow G$ which agrees with f on $\{v_1, \dots, v_k\}$. By the universal property of free products, there is a unique homomorphism $h : G * F_k \rightarrow G$ that agrees with g on F_k and with the identity homomorphism $G \rightarrow G$. As G is abelian, h can be written uniquely as $h = h' \circ p$ where $p : (G * F_k) \rightarrow (G * F_k)^{ab}$ is the quotient map, and $h' : (G * F_k)^{ab} \rightarrow G$ is a projection that agrees with f on $\{v_1, \dots, v_k\}$. This gives the desired one to one correspondence. \square

Lemma 3.8. *Let $w \in \{v_1, \dots, v_k\}$ be linear in some free variable v_i and let $g \in G$. Then there are exactly n^{k-1} projections $\pi : (G * F_k)^{ab} \rightarrow G$ having $\pi(w) = g$.*

Proof. Suppose that $i = k$, without loss of generality. By Lemma 3.7 there are exactly n^{k-1} projections $\pi : (G * F_{k-1})^{ab} \rightarrow G$. For each such π , we show that there is a unique projection π' that agrees with π and additionally has $\pi'(w) = g$. By linearity of w in v_k , the equation $w = g$ rearranges into $v_k = h$ for some $h \in (G * F_k)^{ab}$ and v_k does not appear in h . So, $\pi'(w) = g$ is equivalent to $\pi'(w) = \pi'(g)$ (as π' is a projection) which is equivalent to $\pi'(v_k) = \pi'(h)$ (as π is a homomorphism). As π' agrees

with π and $h \in (G * F_{k-1})^{\text{ab}}$, we have that $\pi'(v_k) = \pi(h) \in G$. Therefore, π' has that $\pi'(v_k) = \pi(h)$, and $\pi'(v_i) = \pi(v_i)$ for $1 \leq i < k$. By Lemma 3.7, there is a unique projection with this property. \square

The following is a simple consequence of the previous lemma.

Lemma 3.9. *Let $S \subseteq (G * F_k)^{\text{ab}}$ be a set of elements which are each linear in at least one variable, and let $U \subseteq G$. Then the number of projections $\pi : (G * F_k)^{\text{ab}} \rightarrow G$ for which $\pi(S)$ intersects U is $\leq |S||U|n^{k-1}$.*

Definition 3.10. Let $w, w' \in (G * F_k)^{\text{ab}}$. We say that w and w' are **separable** if any of the following hold.

- (a) $w' - w$ is linear in some free variable v_i . Note that this is equivalent to asking that there exists a free variable v with coefficient z in w and z' in w' and we have $|z - z'| = 1$.
- (b) The equation $w = w'$ rearranges into $g = 0$ for some nonzero group element $g \in G$.
- (c) The equation $w = w'$ rearranges into $3v_i - 2v_j = g$ for some group element $g \in G$ and distinct free variables v_i and v_j .

Definition 3.11. Let $S \subseteq G * F_k$. We say that a homomorphism $\phi : (G * F_k)^{\text{ab}} \rightarrow G$, **separates** S if for every separable $w, w' \in S$ we have $\phi(w) \neq \phi(w')$.

Lemma 3.12. *Let $n \geq 10^{100}$. Let $S \subseteq (G * F_k)^{\text{ab}}$ be a set of size ≤ 1000 . Then there are at most $|S|^2 n^{k-1/2}$ projections $\pi : (G * F_k)^{\text{ab}} \rightarrow G$ which do not separate S .*

Proof. Let w and w' be two separable words in S . We case on which of the conditions (a)/(b)/(c) makes w and w' separable, and count the projections which do not separate them in each case.

- (a) In this case, $\pi(w) = \pi(w')$ is equivalent to $\pi(w - w') = e$ (using that π is a homomorphism). By Lemma 3.8, there are n^{k-1} projections π satisfying this latter identity.
- (b) We can rearrange $\pi(w) = \pi(w')$ into $\pi(g) = \pi(e)$ using that π is a homomorphism. The latter implies that $g = e$ using that π is a projection, which is a contradiction. Hence there can be no projections π with $\pi(w) = \pi(w')$ in this case.
- (c) Similarly to the previous cases, we can rearrange $\pi(w) = \pi(w')$ into $3\pi(v_i) - 2\pi(v_j) = g$. Using Lemma 3.6, suppose first that $x \rightarrow 3x$ has at least $n^{1/2}$ images, and suppose π also satisfies $\pi(v_j) = g'$ for some $g' \in G$, so we have $3\pi(v_i) = g'' \in G$ (where $g'' = g + 2g'$). As $x \rightarrow 3x$ is a homomorphism (G is abelian) and has at least $n^{1/2}$ images, the preimage of g'' under the map $x \rightarrow 3x$ has size at most $n^{1/2}$ (each nonempty preimage must have the same size in a group homomorphism). This means $\pi(v_i)$ must live in a set $T_{g''}$ of size at most $n^{1/2}$ assuming that $\pi(w) = \pi(w')$ and $\pi(v_j) = g'$. Thus, if $\pi(w) = \pi(w')$, π must agree with one of $n^{k-1}n^{1/2}$ functions $f : \{v_1, \dots, v_k\} \rightarrow G$, meaning that there are at most $n^{k-1/2}$ such projections, using Lemma 3.7. A symmetric argument works when $x \rightarrow 2x$ has at least $n^{1/2}$ images.

As there are at most $\binom{|S|}{2}$ pairs of separable words in S , the desired bound follows. \square

Lemma 3.13. *Let $n \geq 10^{10}$, and let $S \subseteq (G * F_k)^{\text{ab}}$ be a set of at most 100 elements which are all linear in at least one variable. Then, there are projections $\pi_1, \dots, \pi_{n/10^5}$ which separate S and have $\pi_1(S), \dots, \pi_{n/10^5}(S)$ disjoint.*

Proof. Call a projection *good* if it separates S . Let π_1, \dots, π_t be a maximal collection of good projections with the sets $\pi_i(S)$ being pairwise disjoint. Set $T = \pi_1(S) \cup \dots \cup \pi_t(S)$ noting $|T| = |S|t$. For any good projection π , we must have $\pi(S) \cap T \neq \emptyset$ by maximality, so by Lemma 3.9, we have that there are at most $|S|(|S|t)n^{k-1} = 10^4 t n^{k-1}$ good projections. On the other hand, there are at most $|S|^2 n^{k-1/2} \leq n^k/2$ projections which are not good by Lemma 3.12, so there are at least $n^k/2$ good projections (there are n^k projections total). Combining, we have $n^k/2 \leq 10^4 t n^{k-1}$, meaning $t \geq n/(2 \cdot 10^4)$, as desired. \square

Combining the previous lemma with a standard application of Chernoff's bound, we obtain the following. The different random sets are present in the statement as some words will represent vertices, others will represent colours in our applications.

Lemma 3.14. *Let $p \geq n^{-1/700}$. Let R_1, R_2 be p -random subset of G , sampled independently. With high probability, the following holds.*

*Let $S \subseteq (G * F_k)^{ab}$ a set of ≤ 100 elements which are each linear in at least one variable. Suppose $k \leq 20$ and each word has length at most 20. Let $S_1 \cup S_2 = S$ be a partition of S . Let $U \subseteq G$ with $|U| \leq p^{100}n/10^7$. Then there is a projection $\pi : (G * F_k)^{ab} \rightarrow G$ which separates S , has $\pi(S) \cap U = \emptyset$ and $\pi(S_1) \subseteq R_1$ and $\pi(S_2) \subseteq R_2$.*

Proof. Note that there are at most $20 \cdot 40^{20} \cdot n$ words of length at most 20 in $(G * F_k)^{ab}$ for some $k \leq 20$, and thus at most $40^{2100}n^{100}$ distinct choices for S . Fix some choice of S . We will show that the statement holds for S with probability $\geq 1 - e^{-\Omega(\sqrt{n})}$, so the desired statement holds for all S simultaneously by a union bound.

Let $\pi_1, \dots, \pi_{n/10^5}$ be the projections separating S with disjoint images as guaranteed by Lemma 3.13. For any of these projections, that $\pi(S_1) \subseteq R_1$ and $\pi(S_2) \subseteq R_2$ is an event with probability at least p^{100} , so the expected number π satisfying this property is $p^{100}n/10^5$, and further these events are mutually independent over the various π as they are determined by whether disjoint sets of vertices are sampled into R_1 (or R_2). Hence, that at least $p^{100}n/10^6$ of the π satisfy this property happens with probability at least $\geq 1 - e^{-\Omega(\sqrt{n})}$ by Chernoff's bound (using that $p \geq n^{-1/700}$).

Once this property is satisfied, any $U \subseteq G$ can meet at most $|U|$ distinct images of π (as the π are disjoint), so for any $|U|$ of size at most $p^{100}n/10^7$, we may find a π of the desired form, disjointly with U . □

We now package everything we have so far into a lemma (Lemma 3.19) that fits nicely with our application in the setting of \vec{K}_G .

Definition 3.15. Given a group G , a **pattern** P is a directed (simple) graph equipped with a vertex and edge labelling ϕ with the following properties.

1. ϕ maps vertices and edges to $(G * F_k)^{ab}$ for some positive integer k .
2. Each vertex gets a distinct label via ϕ (i.e., $\phi|_{V(P)}$ is injective, but distinct edges can potentially receive the same label)
3. If $\vec{e} \in E(P)$ is a directed edge from v to w for $v, w \in V(P)$, we have that $\phi(v) - \phi(w) = \phi(\vec{e})$.

In Figure 1 we have several examples of patterns. We can naturally view the edge-labels as colours, hence each pattern can also be viewed as an edge-coloured graph.

A **pairwise separable** subset S is a subset where any two distinct words w and w' are separable.

Definition 3.16. We call a pattern (P, ϕ) **well-distributed** if the following two conditions hold.

1. The subsets (viewed as sets, not multisets) $\{\phi(v) : v \in V(P)\}$ and $\{\phi(\vec{e}) : \vec{e} \in E(P)\}$ are both pairwise separable subsets of $G * F_k$. (In particular, a two identical words, potentially both constants, may appear as a label on two distinct edges.)
2. Each label is either a constant, or linear in at least one free variable.
3. There are at most 20 distinct free variables, and each label word has length at most 20.

Notice that we are not insisting that any $\phi(v)$ and $\phi(\vec{e})$ are separable for a vertex v and edge \vec{e} . This is because in our applications vertex sets and colour sets are sampled independently, hence we don't need any separability properties.

Definition 3.17. A **copy** of a well-distributed pattern (P, ϕ) is a subgraph S of $\vec{K}_G[V; C]$ such that there exists a projection $\pi : (G * F_k)^{ab} \rightarrow G$ (where k is the number of free variables used in ϕ) with the following properties.

1. π maps $\phi[V(P)]$ (the vertex labels) to $V(S) \subseteq V$ and $\phi[E(P)]$ (the edge labels) to C .
2. π separates $\phi[V(P)]$ and π separates $\phi[E(P)]$. In particular, π is injective when restricted to $\phi[V(P)]$.
3. The edge (v, w) is a directed edge of S if and only if there is a directed edge from the vertex with the label $\pi^{-1}(v)$ to the vertex with the label $\pi^{-1}(w)$ in P .

An **edge-coloured directed graph isomorphism** ψ between two edge-coloured simple directed graphs G_1, G_2 is a graph isomorphism mapping vertices of G_1 to vertices of G_2 and mapping edges (v, w) of G_1 to $(\psi(v), \psi(w))$ that preserves the direction of each edge and respects colours. This means that e_1 and e_2 of G_1 have the same colour if and only if $\psi(e_1)$ and $\psi(e_2)$ have the same colour.

Observation 3.18. Let S be a copy of P . Then, there is an edge-coloured directed graph isomorphism ψ between P and S . Furthermore, if x is the label of a vertex or colour of P , and x is a constant, $\psi(x) = x$.

Proof. The projection π witnessing that S is a copy naturally corresponds to a ψ with the desired properties, as π fixes elements of $(G * F_k)^{\text{ab}}$ which are constants by definition of a projection. \square

The following is a consequence of the definition of well-distributed, copy, and applying Lemma 3.14 to R_1 and R_2 .

Lemma 3.19. Let $p \geq n^{-1/700}$. Let R_1 and R_2 be p -random subsets of G , sampled independently. With high probability, the following holds.

Let P be a well-distributed pattern with $|V(P)| + |E(P)| \leq 100$. Let $U \subseteq G$ with $|U| \leq p^{100}n/10^7$. Let V' and C' be the set of labels of vertices and colours in P which are constants. Then, there is a copy of P in $\tilde{K}_G[(R_1 \setminus U) \cup V'; (R_2 \setminus U) \cup C']$.

Proof. With high probability, R_1 and R_2 satisfy the conclusion of Lemma 3.14. Fix such R_1 and R_2 .

Let P be a well-distributed pattern with $|V(P)| + |E(P)| \leq 100$ so we have that $\phi[V(P)] \sqcup \phi[C(P)] := S$ (\sqcup indicates a disjoint union) has size at most 100. We may safely ignore the words that are constants in $\phi[V(P)]$ or $\phi[C(P)]$ as then their presence in the target vertex/colour of $\tilde{K}_G[(R_1 \setminus U) \cup V'; (R_2 \setminus U) \cup C']$ is guaranteed. Hence, we may assume all words are linear in at least one free variable by the definition of well-distributed. We then apply the property from Lemma 3.14 (its hypotheses are satisfied by linearity in at least one variable and item 3. of well-distributed) with $S_1 = \phi[V(P)]$ and $S_2 = \phi[C(P)]$ and $S = \phi[V(P)] \sqcup \phi[C(P)]$. This allows us to deduce that there exists a projection π of S mapping vertices to $R_1 \setminus U$ and $R_2 \setminus U$ separating S_1 and S_2 . The image of this projection naturally corresponds to a copy by picking a subgraph keeping exactly the directed edges that were present in the pattern P (in order to ensure item 3. of a copy). \square

As an example application of the above result, we recommend the reader to inspect the proof of Lemma 5.2.

3.2.2. Partitioning into sets with fixed sum

In this subsection, we prove some lemmas designed to ‘cover-down’ part of the absorption strategy. See the proof overview for more context. For the reader interested in the $k = 3$ case, the $k = 3$ case of Lemma 3.25 is all that is required, and this case follows directly from Lemma 3.21 (without having to use Lemma 3.24). We cite the following three results from [30]. Recall that $H_G[X, Y, Z]$ denotes the 3-uniform hypergraph whose vertex set is a disjoint union of X, Y and Z , and $(x, y, z) \in X \times Y \times Z$ is an edge whenever $x + y + z = 0$.

The below is [30, Theorem 4.6] specialised to abelian groups.

Theorem 3.20. Let $p \geq n^{-1/10^{102}}$. Let G be an abelian group of order n . Let $R^1, R^2 \subseteq G$ be disjoint p -random subsets, and let $R^3 \subseteq G$ be a p -random subset, sampled independently with R^1 and R^2 . Then, with high probability, the following holds.

Let X, Y, Z be equal-sized subsets of G_A, G_B , and G_C respectively, satisfying the following properties.

- $|(R_A^1 \cup R_B^2 \cup R_C^3) \Delta (X \cup Y \cup Z)| \leq p^{10^{18}} n / \log(n)^{10^{18}}$
- $\sum X + \sum Y + \sum Z = 0$
- Suppose that $0 \notin X \cup Y \cup Z$

Then, $H_G[X, Y, Z]$ contains a perfect matching.

The following is a special case of [30, Lemma 6.25]

Lemma 3.21. *Let $p \geq n^{-1/10^{100}}$ and $3 \leq k \leq 100$. Let R be a p -random subset of an abelian group G . With high probability the following holds.*

Let $X \subseteq G$ with $|X \Delta R| \leq p^{10^{10}} n / \log(n)^{10^{24}}$, $0 \notin X$, $\sum X = 0$, and $|X| \equiv 0 \pmod k$. Then, X can be partitioned into zero-sum sets of size k .

The below is [30, Lemma 6.23].

Lemma 3.22. *Let $p \geq n^{-1/700}$ and let R be a p -random subset of an abelian group G . With high probability the following holds.*

Let $\epsilon \in [2 \log n / \sqrt{n}, p^{800} / 10^{4010}]$. For any m with $|m - pn| \leq \epsilon n$, $g \in G$ and Z with $|Z| \geq m + 3$, $|R \setminus Z| \leq \epsilon n$, there is a set $R' \subseteq Z$ with $|R'| = m$, $|R' \Delta R| \leq 6\epsilon n$, and $\sum R' = g$.

We also deduce the following corollary.

Corollary 3.23. *Let $p \geq n^{-1/10^{100}}$. Let R be a p -random subset of an abelian group G . With high probability the following holds.*

Let $X \subseteq G$ with $|X \Delta R| \leq p^{10^{10}} n / \log(n)^{10^{22}}$. Let $\alpha \in G$. $0 \notin X$, $|X| \equiv 0 \pmod 4$, $\sum X = (|X|/4) \cdot \alpha$. Then, X can be partitioned into sets of size 4 with sum α .

Proof. Let R_1, R_2, R_3, R_4 be disjoint $(p/4)$ -random subsets of G which partition R . Let S be a $(p/4)$ -random subset of G , sampled independently with the previous sets. Note that the set $-S - \alpha$ is also a $(p/4)$ -random subset of G , which is independent with the previous random sets (not including S). With high probability, Theorem 3.20 holds with the sets (R_1, R_2, S) and $(R_3, R_4, -S - \alpha)$, Lemma 3.22 holds for each random set, and by Chernoff's bound, each random set is within a $n^{0.6}$ term of its expectation.

Let $X \subseteq G$ be given. By Lemma 3.22, we can partition X into equal sized sets X_1, X_2, X_3, X_4 such that $|X_i \Delta R_i| \leq p^{10^{10}} n / \log(n)^{10^{21}}$ and $\sum X_1 + \sum X_2 = 0$. This readily implies that $\sum X_3 + \sum X_4 = (|X|/4) \cdot \alpha$ by the sum condition on X . Similarly, via Lemma 3.22, we can fix a set S' with $\sum S' = 0$, $|S'| = |X|/4$, and such that S' has small symmetric difference with S . This implies that $\sum(-S' - \alpha) = -(|X|/4) \cdot \alpha$, and also we have that $-S' - \alpha$ has small symmetric difference with $-S - \alpha$. Thus we have that $\sum X_1 + \sum X_2 + \sum S' = 0$ and $\sum X_3 + \sum X_4 + \sum(-S' - \alpha) = 0$. Also, we remark that S' can be chosen so that both S' and $-S' - \alpha$ do not contain 0. So we can apply Theorem 3.20 twice to deduce that both $H_G[X_1, X_2, S']$ and $H_G[X_3, X_4, -S' - \alpha]$ has a perfect matching. For each $s' \in S'$, consider the edges (x_1, x_2, s') and $(x_4, x_4, -s' - \alpha)$ guaranteed by the two perfect matchings, and observe that $x_1 + x_2 + x_3 + x_4 = \alpha$. Combining 4-tuples of this form, we obtain the desired partition of X . \square

For technical reasons, our absorption strategy for large k requires the assumption that $k \geq 10$. This leaves the case of $3 \leq k \leq 9$ open. The previous lemmas already give us a way to partition sets into k -sets which are zero sum in this regime. Once we have access to such a partition, a natural strategy is to look for an ordering of the k -set yielding a cycle-candidate, in order to be able to perform the cover-down step (see proof overview). We rely on the following result of Alspach and Liversidge to find suitable orderings. Similar results for cyclic groups were obtained in [11, 24].

Lemma 3.24 (Alspach-Liversidge, [3], Corollary 5.2). *Let G be any abelian group (not necessarily finite). Let $S \subseteq G$ be of size at most 9. with $\sum S = 0$. Then, S admits an ordering yielding a rainbow cycle-candidate if $\sum S = 0$, and otherwise S admits an ordering yielding a rainbow path-candidate.*

We can now prove the main lemma of this section.

Lemma 3.25. *There exists an absolute constant $\epsilon_{3,25}$ such that the following holds. Let $3 \leq k \leq 9$. Let G be an abelian group of order n , let $p \geq n^{-\epsilon_{3,25}}$. Let R be a p -random subset of G . With high probability, the following holds. Let $R' \subseteq G$ such that $|R' \Delta R| \leq p^{10^{10}} n / \log(n)^{10^{23}}$. Suppose k divides $|R'|$ and that $0 \notin R'$.*

1. *Suppose that $\sum R' = 0$. Then, R' can be partitioned into k -tuples which are rainbow cycle-candidates.*
2. *Suppose that $k = 4$ and for some $\alpha \in G \setminus \{0\}$, $\sum R' = (|R'|/k) \cdot \alpha$. Then, R' can be partitioned into k -tuples which are rainbow path-candidates with sum α .*

Proof. Choose some $\varepsilon_{3.25} \leq 10^{-1000}$. With high probability, Lemma 3.21 and Corollary 3.23 both hold for R . Let R' be given. For part (1), we apply Lemma 3.21 to partition R' into k -sets which are zero-sum. Then, Lemma 3.24 implies that each of these k sets can be ordered to obtain a rainbow cycle candidate, as $k \leq 9$. For part (2), we apply Corollary 3.23 to partition into 4-tuples each with sum α . As $\alpha \neq 0$, we can order each tuple to be path-candidates by Lemma 3.24. This concludes the proof. \square

3.2.3. Good families of colours

In this section we have some lemmas designed to deal with the $k \geq 10$ case of the cover-down step.

Lemma 3.26. *Let G be an abelian group of order n , let $s \in G \setminus \{0\}$ let \bar{T} be the collection of k -tuples (g_1, \dots, g_k) with $\sum g_i = s$ (note $|\bar{T}| = n^{k-1}$). Suppose $n^{0.01} \geq k \geq 2$ and $n \geq 10^{10}$. Let S be a subset of G of size at most $n/(20k)$. Then, all but at most $n^{k-1}/4$ tuples in \bar{T} are all rainbow path-candidates disjoint with S .*

Proof. As $s \neq 0$, we can count that there are at least $(n-1)(n-2) \cdots (n-k+1) \geq (n-k)^{k-1} \geq n^{k-1}/1.01$ (using that k is small for the final inequality) path-candidates in \bar{T} .

If $k \geq 3$, by a direct counting we can see that there are at most $k^2 nn^{k-3} \leq k^2 n^{k-2}$ tuples in \bar{T} with two coordinates being equal. If $k = 2$, using that G is an abelian group and $s \neq 0$, we see that there are at most $n/2$ tuples (generously) in \bar{T} with two coordinates being equal. In either case, all but $n^{k-1}/2$ tuples in \bar{T} are rainbow (using that $n \gg k$).

For each $g \in G$, there are at most kn^{k-2} elements of \bar{T} having g in some coordinate, here we used that $k \geq 2$. So, there are at most $|S|kn^{k-2} \leq n^{k-1}/20$ many tuples \bar{t} not disjoint with \mathcal{S}_G .

We derive that there are at least $n^{k-1}/1.01 - n^{k-1}/2 - n^{k-1}/20 \geq n^{k-1}/4$ tuples in \bar{T} satisfying all the desired properties. \square

Lemma 3.27. *There exists some absolute constant $C_{3.27}$ such that the following holds. Let G be an abelian group of order n , let $n \geq 10^{100}$, and let k be a integer such that $10 \leq k \leq n^{0.001}$. Then, G contains two families $\mathcal{F}_G = \mathcal{F}_G(k)$ and $\mathcal{S}_G = \mathcal{S}_G(k)$ of disjoint tuples $\mathcal{F}_1, \dots, \mathcal{F}_{\lfloor n/kC_{3.27} \rfloor}$ and $\mathcal{S}_1, \dots, \mathcal{S}_{\lfloor n/kC_{3.27} \rfloor}$ with the following properties.*

- (1) Each \mathcal{F}_i is of size 4 and has the same sum $f = f_{G,k}$.
- (2) Each \mathcal{S}_i has the same size and sum $s = s_{G,k}$. In fact, $|\mathcal{S}_i| =: z_{\mathcal{S}} \in \{2, 3, 4, 5\}$.
- (3) Each \mathcal{S}_i is a rainbow path candidate, and \mathcal{S}_G is near-dissociable.
- (4) $k - 4 - z_{\mathcal{S}}$ is divisible by 4. Furthermore, set $q := q_{G,k} = -((k-4-z_{\mathcal{S}})/4)f - s$. We have that $q \neq 0$.
- (5) Each \mathcal{F}_i can be partitioned into two tuples, $\mathcal{F}_i^+ = (f_i^{+,1}, f_i^{+,2})$ and $\mathcal{F}_i^- = (f_i^{-,1}, f_i^{-,2})$, both of which are rainbow path candidates. The resulting collection of \mathcal{F}_i^+ and \mathcal{F}_i^- are both dissociable. Also, each \mathcal{F}_i is a rainbow path candidate, and \mathcal{F}_G is near-dissociable.
- (6) For each $m \in \{0, 1, 2, \dots, k\}$ and $i \in \lfloor n/kC_{3.27} \rfloor$, we have that \mathcal{F}_i^+ and \mathcal{F}_i^- are separable at a distance $q + mf$.

Proof. Pick some $z_{\mathcal{S}}$ between 2 and 5 so that $k - 4 - z_{\mathcal{S}}$ is positive and divisible by 4, note that this is possible as $k \geq 10$. Pick any $f \neq 0$. Pick some s so that $q + m \cdot f \neq 0$ for any $m \in \{0, 1, 2, \dots, k, k+1\}$ for $q := -((k-4-z_{\mathcal{S}})/4)f - s$. Indeed, there are $0.9n$ such choices of s , as $k \leq n^{0.001}$.

Claim 3.27.1. *We can find \mathcal{S}_G satisfying (2) and (3).*

Proof. Set $k' = z_{\mathcal{S}}$, recalling that $k' \geq 2$. Suppose that we have found a maximal family \mathcal{S}_G satisfying (2) and (3) and suppose that $|\mathcal{S}_G| < n/(kC)$ for some C . We will derive a contradiction for C sufficiently large.

Let \bar{G} be the collection of rainbow path candidate k' -tuples $(g_1, \dots, g_{k'})$ with $g_1 + \dots + g_{k'} = s$ which are disjoint with $\bigcup \mathcal{S}_G$ (observing this set has size at most $5n/(kC)$ by assumption). $|\bar{G}| \geq n^{k'-1}/4$ by Lemma 3.26 (supposing $C \geq 100$). If for some $\bar{t} \in \bar{G}$ we have that \mathcal{S}_G^* becomes nondissociable upon the addition of \bar{t} , there must exist some $\bar{t}' \in \mathcal{S}_G, j, j' \in [k'-1]$ such that $\sum_{i \in [j]} \bar{t}'_i = \sum_{i \in [j']} \bar{t}_i$. There are at most $(n/kC)k'n^{k'-2} \leq n^{k'-1}/C$ such \bar{t} , meaning that there is a $\bar{t} \in \bar{G}$ that we can add to \mathcal{S}_G without breaking (2) and (3), a contradiction. \square

It remains to construct \mathcal{F}_G . Suppose \mathcal{F}_G is a family of 4-tuples satisfying the properties with size at most $n/(kC) - 1$ (where C is a sufficiently large constant). We will show that \mathcal{F}_G can be extended.

Fix some $f^+, f^- \in G \setminus \{0\}$ such that $f^+ + f^- = f$, and the following two properties hold.

1. For any $\mathcal{F}_i \in \mathcal{F}_G$ we have that $f^+, f^- \neq \sum \mathcal{T}_i$ for any $\mathcal{T}_i \subseteq \mathcal{F}_i$.
2. $f^+ + f^- \neq \pm q + mf$ for any $m \in \{0, 1, \dots, k, k + 1\}$.

Such f^+, f^- with the first property exist as long as $C > 50$, as $\sum \mathcal{T}_i^\pm$ can take at most $20n/C$ distinct values due to the assumption on the size of \mathcal{F}_G . Such f^+, f^- automatically satisfy the second property as $q + m \cdot f \neq 0$ for any $m \in \{0, 1, \dots, k\}$.

Let F^+ denote the set of ordered triples with sum f^+ and F^- denote the set of ordered triples with sum f^- , noting $|F^+| = |F^-| = n$.

Claim 3.27.2. *Suppose we delete all pairs from $F \in F^+$ such that the collection $\mathcal{F}_+ \cup \{F\}$ fails to be dissociable. This deletes at most $10n/C$ triples.*

Proof. If for some $F \in F^+$ we have that $\{\mathcal{F}_i^+\} \cup F$ is not dissociable, there must exist some $F' \in \mathcal{F}^+$ and $j \in \{1, 2\}$ and $j' \in \{1, 2\}$ such that $\sum_{i \in [j']} F'(i) = \sum_{i \in [j]} F(i)$. It cannot be that $j = 2$ by the first property coming from our choice of f^+ . By the bound on $|\mathcal{F}_G|$, there are at most $10n/C$ distinct values the quantity $\sum_{i \in [j']} F'(i) =: w$ can take. For each such w , there is at most one $F \in F^+$ with $F(1) = w$. This implies that in total there are at most the claimed number of pairs which make the corresponding collection not dissociable. □

Claim 3.27.3. *Suppose we delete all tuples from $F \in F^+$ such that F and the 1-tuple (f^-) are not separable at a distance $q + m \cdot f$ for some $m \in \{0, \dots, k\}$. This deletes at most $2k$ tuples.*

Proof. If for some $F \in F^+$ we have that F and (f^-) are not separable at a distance $q + m \cdot f$ for some $j \in \{1, 2\}$, then we must have $\sum_{i \in [j]} F(i) + f^- = \pm(q + m \cdot f)$ for some $m \in \{0, \dots, k\}$. Here, $j = 2$ is precluded by the second property of f^+ and f^- . For each of the $2k$ possible values of $\pm(q + m \cdot f) - f^-$, there exists at most one $F \in F^+$ such that $F(1) = \pm(q + m \cdot f) - f^-$, which implies the claim. □

Claim 3.27.4. *There are at least $n/4$ tuples in F^+ which are rainbow path-candidates and which contain no coordinate $F(i)$ also present in an element of \mathcal{S}_G or \mathcal{F}_G .*

Proof. This is immediate by Lemma 3.26 and bounding $|\bigcup \mathcal{S}_G \cup \bigcup \mathcal{F}_G|$. □

Claim 3.27.5. *Deleting all tuples $F \in F^+$ with $F(1) = \pm \sum \mathcal{T}_i$ or $F(2) = \pm \sum \mathcal{T}_i$ for some $\mathcal{T}_i \subseteq \mathcal{F}_i \in \mathcal{F}_G$, we delete at most $80n/C$ elements.*

Proof. There are at most $20n/C$ possible values for the quantity $\pm \sum \mathcal{T}_i$ by the upper bound on the size of \mathcal{F}_G . This implies the claim, as $F(1)$ (and $F(2)$) is a distinct value for each $F \in F^+$. □

By the bounds coming from the claims, we can fix \mathcal{F}_{new}^+ to be a 2-tuple from F^+ which is a rainbow path candidate disjoint with the earlier sets, keeps \mathcal{F}^+ dissociable, and is separable with (f^-) at a distance $q + m \cdot f$ for each $m \leq k$.

Now, we perform the analogous steps for F^- . Claim 3.27.2 and 3.27.4 (thinking of \mathcal{F}_{new}^+ as an element of \mathcal{F}_G to ensure disjointness) also hold when $+$ is replaced by $-$, giving us at least $n/5$ potential elements of F^- we can select while maintaining dissociability of \mathcal{F}^- , disjointness with previous tuples, and rainbow path candidacy.

In addition, we delete the elements of F^- which are not separable with \mathcal{F}_{new}^+ at a distance $q + mf$ for some $m \in \{0, 1, \dots, k\}$. For any $F \in F^-$, it is already impossible for $\sum_{i \in [j]} \mathcal{F}_{new}^+(i) + \sum_{i \in [j']} F(i) = \pm(q + m \cdot f)$ when $j' = 2$ (by the property from Claim 3.27.3). When $j' = 1$, note that $\pm(q + m \cdot f) - \sum_{i \in [j]} \mathcal{F}_{new}^+(i)$ can take at most $4k$ distinct values v , and we only need to delete at the at most $4k$ many $F \in F^-$ with $F(i) = v$.

For each $F \in F^-$, consider the 4-tuple $\mathcal{F}_{new} = (\mathcal{F}_{new}^+(1), \mathcal{F}_{new}^+(2), F(1), F(2))$, and note that this is always a rainbow sequence. If $F \in F^-$ makes $\mathcal{F}_G \cup \{\mathcal{F}_{new}\}$ not near-dissociable, we delete

F from F^- . To count how many such F there are, suppose that for some $\mathcal{F} \in \mathcal{F}_G$, we have that $\sum_{i \in [j]} \mathcal{F}(i) = \sum_{i \in [j']} \mathcal{F}_{new}(i)$ where $j, j' \in [3]$. It is impossible that $j' = 1$ due to Claim 3.27.5 and it is impossible that $j' = 2$ because of the first property of f^+ . Note there are at most $20n/C$ potential values of $\sum_{i \in [j]} \mathcal{F}(i)$ due to the bound on the size of \mathcal{F}_G . This implies that for the relevant equality to hold, $F(1)$ needs to belong to a set of size $20n/C$, so in this step we delete at most $20n/C$ elements from F^- . Similarly, if $(\mathcal{F}_{new}^+(1), \mathcal{F}_{new}^+(2), F(1), F(2))$ is not a path candidate, it must be that a partial sum of the sequence is 0. This partial sum cannot contain both of $F(1)$ and $F(2)$, as the whole sum is $f \neq 0$, and $\mathcal{F}_{new}^+(2) \neq -F(1) - F(2)$ by Claim 3.27.5, and $(F(1), F(2))$ is a path candidate. But the partial sum has to contain $F(1)$, as $(\mathcal{F}_{new}^+(1), \mathcal{F}_{new}^+(2))$ alone gives a path candidate. This means that at most 2 extra values of $F(1)$ are forbidden if $(\mathcal{F}_{new}^+(1), \mathcal{F}_{new}^+(2), F(1), F(2))$ is to be a rainbow path candidate.

Selecting C large, we can fix a value of $F \in F^-$ so that setting $\mathcal{F}_{new}^- = F$, we successfully extend \mathcal{F}_G , as desired. \square

4. Nibble with some determinism

In this section we give a proof of Lemma 2.5.

Observation 4.1. Let \mathcal{E} be an equation of the form $\pm a \pm b \pm c = 0$. Let H be a tripartite hypergraph obtained by taking three copies of some group G of order n , and letting $(a, b, c) \in G^3$ be an edge whenever it is a solution to \mathcal{E} . Then, H is $(0, 1, n)$ -typical.

Proof. For a proof for when \mathcal{E} is $a + b + c = 0$, see Observation 3.3 in [30]. For other equations of this form, the proof is essentially identical. \square

Typical graphs have the following useful pseudorandomness property (see [30, Lemma 3.7]).

Lemma 4.2. Let $H = (A, B, C)$ be a tripartite linear hypergraph that is $(0, 1, n)$ -typical. Let $p \geq n^{-1/600}$ and let $A' \subseteq A$ be p -random. Then, with probability at least $1 - 1/n^3$, the following holds. For any $B' \subseteq B$, there are at most $n^{9/10}$ vertices $c \in C$ with $e_H(A', B', c) \neq p|B'| \pm n^{9/10}$.

This allows us to deduce the following.

Lemma 4.3. Let $p \geq n^{-1/600}$ and let $X \subseteq G$ be p -random. Then, with probability at least $1 - 8/n^3$, the following holds. For any $Y \subseteq G$, for all but at most $8n^{9/10}$ vertices $g \in G$, and for each equation of the form $\mathcal{E} := \pm g \pm x \pm y = 0$ (where g is a constant and x and y are free variables), we have that there are $p|Y| \pm n^{9/10}$ many $(x, y) \in X \times Y$ such that x, y and g satisfy \mathcal{E} .

Proof. Thanks to Observation 4.1, we can apply Lemma 4.2 to the corresponding hypergraph defined by each of the $2^3 = 8$ possible equations \mathcal{E} , and with probability at least $1 - 8/n^3$, we ensure that the conclusion of Lemma 4.2 holds for each of these hypergraphs. The desired statement follows immediately. \square

We say that $g \in G$ is **generic** if $g \neq e$ and there are at most $n^{1/2}$ solutions to $x^2 = g$ in G . Let $N(G)$ denote the set of nongeneric elements and note that $|N(G)| \leq n^{1/2}$.

Observation 4.4. Let G be an abelian group of order n and let $A \subseteq G$ be a multiset of order k . Consider the sets $A+g$ for each $g \in G$. Then, at most $kn^{-1/10}$ many such sets have more than $n^{3/5}$ many nongeneric elements. Also, there are at most $kn^{-1/10}$ sets $A-g$ with more than $n^{3/5}$ many nongeneric elements.

Proof. There are $\leq kn^{1/2}$ tuples $(a, g) \in A \times G$ where $a+g$ is nongeneric. Let $\#$ be the number of $g \in G$ such that there are $\geq n^{3/5}$ many $a \in A$ such that $a+g$ is nongeneric. Then, $\# \cdot n^{3/5} \leq kn^{1/2}$, so $\# \leq kn^{-1/10}$. The same argument applies when $+$ is replaced by $-$. \square

Recall that given a fixed graph F , a **packing** of F in some other graph G is just a collection of vertex-disjoint copies of F in G . When we talk about rainbow packings, we always mean that there is no colour repetition in edges across all copies of F in the packing.

Lemma 4.5. *There exists an absolute constant $\varepsilon_{4.5} > 0$ such that the following holds. Let $p \geq n^{-\varepsilon_{4.5}}$. Let G be a group of order n . Let $V_2, V_3 \subseteq G$ be disjoint p -random, let $C_1, C_2, C_3 \subseteq G$ be disjoint p -random, sampled independently with V_2, V_3 . The following holds with probability at least $1 - 1/n^{2.9}$.*

Let $V_1, V_4 \subseteq G \setminus (V_2 \cup V_3)$ with $|V_1| = |V_4| = (p \pm n^{-0.1})n$. Let $f: V_1 \rightarrow V_4$ be a bijection. Then, $\tilde{K}_G[V_1, V_2, V_3, V_4; C]$ contains a rainbow packing of at least $n^{1-1/10^5}$ paths of length 3, directed $V_1 \rightarrow V_2 \rightarrow V_3 \rightarrow V_4$, such that for all paths \vec{P} in the packing and $v_1 \in V_1 \cap V(\vec{P})$, we have that $f(v_1) \in V_4 \cap V(\vec{P})$.

Proof. Each of the following holds with probability at least $1 - O(1/n^3)$, thus they all simultaneously hold with probability at least $1 - 1/n^{2.9}$.

- (1) Lemma 4.3 holds for X set to be each of V_2, V_3, C_1, C_2 and C_3 .
- (2) For each $i, |V_i|, |C_i| = (p \pm n^{-0.1})n$, by Chernoff's bound.
- (3) For every colour $c \in G \setminus \{0\}$ and vertex pair $v, w \in G$ such that $v - w - c$ is generic, we have that there exists $(p^4 \pm n^{-0.1})n$ many rainbow paths of length 3 directed $v \rightarrow V_2 \rightarrow V_3 \rightarrow w$ with edge colours (c_1, c, c_3) for some $(c_1, c_3) \in C_1 \times C_3$. If $v - w - c$ is not generic, we have that there exists at most $(p^4 + n^{-0.1})n$ such paths.

Proof. Consider all tuples $(v, v_2, v_3, w, c_1, c, c_3)$ where $v_2, v_3, c_1, c_3 \in G$, and $v - v_2 = c_1, v_2 - v_3 = c$ and $v_3 - w = c_3$. There are n such tuples. Note $c_1 + c_3 = v - v_2 + v_3 - w = v - w - c$ which is generic. This means for all but at most $n^{1/2}$ tuples, $c_1 \neq c_3$. As $c \neq 0$, for all tuples $v_2 \neq v_3$. For each of the $n - n^{1/2}$ tuples where $c_1 \neq c_3$, the probability of $(v_2, v_3, c_1, c_3) \in V_2 \times V_3 \times C_1 \times C_3$ is p^4 . Letting X denote the expected number of paths of the desired form, we obtain that $\mathbb{E}[X] = (p^4 \pm n^{-1/2})n$. Further, X is 2-Lipschitz, so the desired concentration follows from Azuma's inequality. When $v - w - c$ is not generic, the same argument applies except we only have an upper bound on $\mathbb{E}[X]$. \square

- (4) For every pair of vertices $v, w \in G$ such that $v - w$ is generic, we have that there exists $(p^3 \pm n^{-0.05})n$ many rainbow paths directed $v \rightarrow V_3 \rightarrow w$ with edge colours from $C_2 \times C_3$. If $v - w$ is not generic, we have that there exists at most $(p^3 + n^{-0.1})n$ such paths.
- (5) For every pair of vertices $v, w \in G$ such that $v - w$ is generic, we have that there exists $(p^3 \pm n^{-0.05})n$ many rainbow paths directed $v \rightarrow V_2 \rightarrow w$ with edge colours from $C_1 \times C_2$. If $v - w$ is not generic, we have that there exists at most $(p^3 + n^{-0.1})n$ such paths.

The proofs for (4) and (5) are essentially identical to the proof for (3), hence we omit them.

Now, suppose V_i and C_i all of the properties, and fix a bijection $f: V_1 \rightarrow V_4$. Let \mathcal{H} be the hypergraph consisting of edges $(v_1, v_2, v_3, v_4, c_1, c_2, c_3) \in V_1 \times V_2 \times V_3 \times V_4 \times C_1 \times C_2 \times C_3$ where $v_1 - v_2 = c_1, v_2 - v_3 = c_2, v_3 - v_4 = c_3$ and $f(v_1) = v_4$. Our goal is to find a matching covering all but $n^{1-1/1000}$ vertices in this hypergraph. We sometimes refer to the edges of this hypergraph as paths. We will show that there is a set S of $\leq n^{1-1/100}$ vertices we can delete from \mathcal{H} so that the resulting hypergraph \mathcal{H}' is almost regular, that is, for all $v \in V(\mathcal{H}), d(v) = (p^5 \pm n^{-0.05})n^2$.

Towards that goal, set S to include

- o the $\leq 90n^{9/10}$ vertices of G coming from Lemma 4.3 applied with each of V_2, V_3, C_1, C_2, C_3
- o the $\leq 2|V_1|n^{-1/10} \leq 2n^{9/10}$ elements of G coming from Observation 4.4 applied with the multiset $\{v - f(v) : v \in V_1\}$ and both + and -

so we have $|S| \leq 100n^{9/10}$. We will show all vertices of \mathcal{H} not in S have degree $(p^5 \pm n^{-0.05})n^2$. Since S is small, this shows that \mathcal{H} has the desired property. We consider several cases.

Let $v_1 \in V_1 \setminus S$, and set $v_4 = f(v_1) \in V_4 \setminus S$. For all but $n^{1/2}$ many $c_2 \in C_2$ we have that $v - w - c_2$ is generic. For such c_2 , we have by 3. that there are $(p^4 + n^{-0.1})n$ paths passing through both v_1 and c_2 . Combined with the bound on the size of C_2 coming from (2), this shows the desired upper and lower bound on $d(v_1)$ because through the few c_2 such that $v - w - c_2$ is nongeneric, there exists at most $10n$ paths passing through both v_1 and c_2 , giving in total $O(n^{3/2})$ such paths.

If $v_4 \in V_4 \setminus S$, set $v_1 = f^{-1}(v_4)$ and apply the result from the previous paragraph.

Let $c_1 \in C_1 \setminus S$. From Lemma 4.3, we have that there exists $p^2 n \pm n^{9/10}$ directed c_1 coloured edges from V_1 to V_2 . As $c_1 \notin S$, for all but $n^{3/5}$ many $v \in V_1$, $v - w - c_1$ is generic, and so for such v , we have that $(v - c_1) - w$ is generic, so we can apply (4) to obtain that there exists $(p^4 + n^{-0.1})n$ paths to $f(v)$ passing through c_1 . Combined with the bound on $|V_1|$, this gives the desired bound on $d(c_1)$, as the number of paths going through the c_1 such that $v - w - c_1$ is nongeneric is too small to influence the count, as before.

Let $c_3 \in C_3 \setminus S$. This case follows by a symmetric argument with the $c_1 \in C_1 \setminus S$ case, using (5) in place of (4).

Let $c_2 \in C_2 \setminus S$. Let $(v, w = f(v)) \in V_1 \times V_2$ and suppose that $v - w - c_2$ is generic. Then by (3) there are $(p^4 + n^{-0.1})n$ paths passing through v, w , and c_2 . As $c_2 \notin S$, we have that all but $n^{3/5}$ values of $v \in V_1$, $v - w - c_2$ is generic. This, with the bound on $|V_1|$ implies the desired bound on $d(c_2)$, again because there are few paths passing through c_2 with $v - w - c_2$ nongeneric.

So S has the desired properties, making \mathcal{H}' almost-regular. Let \mathcal{H}'' be the hypergraph obtained by contracting v and $f(v)$ to a single vertex for each $v_1 \in V_1$. Note that as f is a bijection, and any edge through v has to pass through $f(v)$ as well, this does not change the regularity parameters of any of the other vertices in \mathcal{H}' . To see that this satisfies the hypotheses of Corollary 3.4(1), the only thing left to check is the co-degree condition. This is equivalent to obtaining an upper bound on the number of tuples $(v_1, v_2, v_3, v_4, c_1, c_2, c_3) \in \mathcal{H}$ where the values of 2 coordinates are fixed, and it is not the case that these two coordinates are the first and the fourth (since the corresponding vertices have been contracted). This means that we are counting solutions to a system of equations with 7 free variables and 6 independent constraints, hence there are at most n such solutions. This gives the desired co-degree bound. Corollary 3.4(1) then gives the desired result. \square

Proof of Lemma 2.5. Let $n^{-1/1000} \leq q \leq \varepsilon_{2.5} p^3 / k^{100}$ be a rational number with denominator at most n . We will show that with probability at least $1/n^2$, the statement holds for any $V_D, C_D \subseteq G$ with $|V_D| = |C_D| = qn$, the desired statement then follows by a union bound over the potential values of q .

Let $p \geq n^{-\varepsilon_{2.5}}$ where $\varepsilon_{2.5}$ is some constant such that $\varepsilon_{2.5} \ll \varepsilon_{4.5}$.

Suppose first that $k = 3$. Let $r = (p - 2q)/3$ and let $R_1^{(1)}, R_1^{(2)}, R_1^{(3)}$ be disjoint r, r and $r + 2q$ random (respectively) sets partitioning R_1 . Let $R_2^{(1)}, R_2^{(2)}, R_2^{(3)}, R_2^{(4)}, R_2^{(5)}$ be disjoint q, q, r, r and r -random (respectively) sets partitioning R_2 . Note that $r \geq p/10 \gg n^{-\varepsilon_{4.5}}$ which checks that r is large enough for the upcoming applications of the lemmas to be valid. With probability at least $1 - O(1/n^{2.9})$, Lemma 4.3 holds for $R_2^{(1)}$ and $R_2^{(2)}$ and Lemma 4.5 holds with $(V_2, V_3) = (R_1^{(1)}, R_1^{(2)})$ and $(C_1, C_2, C_3) = (R_2^{(3)}, R_2^{(4)}, R_2^{(5)})$ (noting these are all r -random sets, so r plays the role of p in the statement). Also, by Chernoff's bound the following holds for all cycle-candidate triples (a, b, c) simultaneously with probability at least $1 - 1/n^{10}$: there exists at least $p^3 n / 1000$ vertex-disjoint 3-cycles in $\vec{K}_G[R_1^{(3)}]$ with colour sequence (a, b, c) (this holds with high probability by Lemma 3.19 as well, indeed see Lemma 5.2, but here we cite Chernoff's bound directly to obtain an explicit bound on the probability). Finally, with probability at least $1 - 1/n^{10}$, all random sets are at most $n^{0.6}$ elements away from their expectations. With probability at least $1 - 1/n^2$ all of these properties hold simultaneously.

Now let V_D, C_D be given. Let $H_G[R_2^{(1)}, R_2^{(2)}, C_D]$ denote the 3-partite 3-uniform hypergraph on the indicated parts where triples are edges if and only if they are zero-sum. From Lemma 4.3 applied with the equation $g + x + y = 0$, we have that all but $n^{99/100}$ vertices of $H_G[R_2^{(1)}, R_2^{(2)}, C_D]$ do not satisfy the regularity hypothesis from Corollary 3.4(2). Deleting such vertices, we obtain a $(n^{-0.01}, q^2, qn)$ -regular linear tripartite hypergraph, so Corollary 3.4(2) implies that all but $n^{1-1/700}$ elements of $R_2^{(1)} \cup R_2^{(2)} \cup C_D$ can be covered by disjoint zero-sum triples, denote these triples by \mathcal{T} , noting that $|\mathcal{T}| = qn \pm n^{0.7}$. If necessary, delete at most one edge from \mathcal{T} so that 0 is not used on any triple, meaning that the remaining triples can be ordered to be cycle-candidates (see, for example, Lemma 3.24). Using the property of $R_1^{(3)}$ repeatedly for each triple in \mathcal{T} , we can find a matching M_1 saturating all triples in \mathcal{T} (and nothing else) in $\mathcal{H}_k[R_1^{(3)} \cup \mathcal{T}]$. Now, invoke Lemma 4.5 with $V_1 = V_4 = (R_1^{(3)} \setminus V(M_1)) \cup V_D$ (noting V_1 then has size $(r + 2q - q - q)n \pm n^{0.8} = rn \pm n^{0.8}$, so the application is valid) and f set to be the identity function.

This gives that $\mathcal{H}_k[(R_1^{(1)} \cup R_1^{(2)} \cup R_1^{(3)} \setminus V(M_1)) \cup V_D; R_2^{(3)}, R_2^{(4)}, R_2^{(5)}]$ has a matching covering all but $10n^{1-1/10^5}$ vertices, say M_2 . Then, $M_1 \cup M_2$ is the desired matching.

Union bounding over all potential values of q , we obtain that with high probability, for each q the statement holds. This is sufficient to deduce the assertion, as qn is always an integer.

Now, suppose that $k \geq 4$. For each $i \in [k]$, and $j \in \{1, 2\}$ let $R_j^{(i)}$ be a $((p+q)/k)$ -random set for $i \geq 2$ and $((p+q)/k - q)$ -random set for $i = 1$, partitioning R_j . Similarly to the $k = 3$ case, we will fix some rational $q \leq pn/k^{100}$ with denominator at most n , and prove that the desired statement holds with probability at least $1 - 1/n^{1.1}$.

With probability at least $1 - 1/n^{1.3}$, Lemma 3.5 holds for $(A, B, C) = (R_1^{(i)}, R_1^{(i+1)}, R_2^{(i)})$ with $\ell = ((p+q)/k)n$ for each $i \in [k-3]$ (using that k is polylogarithmic in n for the union bound), with $\gamma := n^{-10^{-5}}$ and $\zeta := n^{-1/2C_{3.5}}$ and a, b, c being the randomness parameters of $(R_1^{(i)}, R_1^{(i+1)}, R_2^{(i)})$. With probability at least $1 - 1/n^{10}$, Lemma 4.5 holds with $(V_1, V_2) = (R_1^{(k-1)}, R_1^{(k)})$, $(C_1, C_2, C_3) = (R_2^{(k-2)}, R_2^{(k-1)}, R_2^{(k)})$. With probability at least $1 - 1/n^2$, each random set is within $n^{0.6}$ elements of its expected size. With probability at least $1 - 1/n^{1.1}$, all these properties hold simultaneously.

Let V_D, C_D be given. Apply Lemma 3.5 with random sets $(R_1^{(1)}, R_1^{(2)}, R_2^{(1)})$ and $(A', B', C') = (R_1^{(1)} \cup V_D, R_1^{(2)}, R_2^{(1)} \cup C_D)$ to find a matching that saturates all but $n^{1-1/10^7}$ vertices. Continue invoking Lemma 3.5 with corresponding random sets and $(A', B', C') = (R_1^{(i)}, R_1^{(i+1)}, R_2^{(i)})$ for each $i \in [k-3] \setminus \{1\}$. In both of these applications, we may delete/add $O(n^{0.78})$ elements from the corresponding sets so that they have size precisely ℓ or $\ell + \lfloor n^{1-10^{-5}} \rfloor$ (depending on whether they are vertex or colours sets), so that the hypotheses of Lemma 3.5 are satisfied, and then if necessary we can delete all edges passing through dummy vertices/colours. Deleting all vertices that fail to be covered by one of the $k-3$ matchings found via the previous applications of Lemma 3.5, we delete at most $kn^{1-1/10^7} \leq n^{1-1/(2 \cdot 10^7)}$ vertices. The remaining vertices form directed paths following sets $R_1^{(1)} \rightarrow R_1^{(2)} \rightarrow \dots \rightarrow R_1^{(k-2)}$. Let $V_1 \subseteq R_1^{(1)}$ and $V_4 \subseteq R_1^{(k-2)}$ be the vertices used by these directed paths, noting $|V_1| = |V_4|$, and let $f: V_1 \rightarrow V_4$ be the bijection induced by the two endpoints of each directed path. Now, Lemma 4.5 allows us to complete all but $n^{1-1/10^5}$ of these paths of length $k-3$ into a k -cycle using the remaining random sets, which gives the desired matching in \mathcal{H}_k . Union bounding over the potential values of q , we obtain the desired result, as in the $k = 3$ case. \square

5. Zero-sum absorption

In this section we prove Lemma 2.4. Throughout this section, whenever a constant C appears inside the statement of a lemma, this should be read as “there is a sufficiently large absolute constant C so that the statement holds with this value of C ”.

5.1. Cover-down step: saturating vertices and colours

5.1.1. Covering vertices

The next lemma gives us a way to find edges of \mathcal{H}_k that pass through a specific set of vertices.

Lemma 5.1. *Let $p \geq n^{-1/700}$. Let $3 \leq k \leq \log^{10} n$. Let R_1, R_2 be p -random subsets of G sampled independently. With high probability, the following holds.*

1. *Let $U \subseteq G$ be a set with $|U| \leq p^{300}n/C_{5.1}$ (recall the convention set in the beginning of Section 5). Let $u, v \in G$ be two vertices, not necessarily distinct. Let k' be such that $2 \leq k' \leq k$. Then, if $u \neq v$, there exists a rainbow path of length k' directed from u to v in $\vec{K}_G[(R_1 \setminus U) \cup \{u, v\}; R_2 \setminus U]$. If $u = v$ and $k' \geq 3$, there exists a directed rainbow cycle of length k' using the vertex v .*
2. *Let $U \subseteq G$ be a set with $|U| \leq p^{300}n/4C_{5.1}$. Let $V \subseteq G$ be a set of vertices with $|V| \leq p^{300}n/(4kC_{5.1})$. Then, $\mathcal{H}_k[(R_1 \setminus U) \cup V; R_2 \setminus U]$ has a matching saturating V where each matched edge uses exactly one vertex from V .*

Proof. Fix a large constant $C \geq 10^7$ and fix $C_{5.1} \gg C$. Fix some distinct vertices of \vec{K}_G , u and v . Fix a set of $n/10$ triples (c_1, x, c_2) where $u \rightarrow x \rightarrow v$ is a rainbow path of length 2 with edge sequence (c_1, c_2) , and the resulting collection of x and $\{c_1, c_2\}$ are both pairwise disjoint. Such a collection exists because there are at least $n/4$ disjoint (c_1, c_2) with $c_1 \neq c_2$ and $c_1 + c_2 = u - v \neq 0$ and G is an abelian group. By Chernoff's bound, with exponentially high probability, $\vec{K}_G[R_1 \cup \{u, v\}; R_2]$ contains at least $p^3 n/10$ such paths. By a union bound over all distinct u and v , we have that $\vec{K}_G[R_1 \cup \{u, v\}; R_2]$ contains at least $p^3 n/10$ such paths for any choice of u and v with high probability. Call this property (*). Also, Lemma 3.19 holds with high probability.

We claim that a stronger version of part (1) holds when $k' \in \{2, 3\}$, with $C_{5.1}$ replaced with C . For $k' = 2$, this already follows from the property (*) as each element of U (other than u or v) can eliminate at most 1 path. We claim the $k' = 3$ case follows from an application of Lemma 3.19. In the case that $u \neq v$, we can see this by defining a pattern that is a directed path of length 3, first vertex labelled u , last vertex labelled v , and colour sequence $(c_1, c_2, u - v - c_1 - c_2)$ (labelled in order of proximity to u) where c_1, c_2 are free variables (and u and v are constants). Note this implies that the vertex sequence is $(u, u - c_1, u - c_1 - c_2, v)$ (so that the third property in the definition of a pattern holds). This is a well-defined pattern as each vertex gets a distinct label. Furthermore, the pattern is well-distributed. The colours are separable by (a), u and v are separable by (b) (as $u \neq v$) and the rest of the vertex pairs are separable by (a). Also, each label is either a constant, or linear in c_1 or c_2 . Then, Lemma 3.19 gives us a copy of this pattern, which corresponds to the desired rainbow path thanks to Observation 3.18. In the case that $u = v$, we can proceed similarly, this time using a pattern that is a directed cycle of length 3, with colour sequence $(c_1, c_2, -c_1 - c_2)$ and vertex sequence $(u, u - c_1, u - c_1 - c_2, u)$, where c_1, c_2 are free variables.

For larger k' , (1) follows from repeated applications of the cases of $k' \in \{2, 3\}$. Indeed, any directed path/cycle of length ≥ 4 can be broken up into directed paths of length 2 and 3. While iteratively invoking (1) with $k' \in \{2, 3\}$, we extend U at each step, in total adding at most k new elements to U . As we know (1) holds with $k' \in \{2, 3\}$ with the smaller constant C , this reduction is valid, as $p^{900}n/C_{5.1} + k < 2p^{900}n/C_{5.1} < p^{900}n/C$ (using that k is small and that $C \ll C_{5.1}$, respectively).

Now, let a U be given as in the part (2) of the statement, and include in U all vertices of V (without relabelling). Given a vertex $z \in V$, we can invoke part (1) with $z = u = v$ to find a cycle of length k using z . For the next iterations, we apply (1) adding to U the vertices we've used so far, which adds to U at most $k|V| \leq p^{300}n/(4C_{5.1})$ elements. This means that U never exceeds a size of $p^{300}n/C_{5.1}$ in any of the iterations, making the applications of (1) valid. \square

5.1.2. Covering colours: Small k

Lemma 5.2. *Let $p \geq n^{-1/700}$. Let $k \leq 50$. Let R_1 be a p -random subset of G . With high probability, the following holds.*

Let S be a k -tuple which is a rainbow path candidate. Let $U \subseteq G$ with $|U| \leq p^{350}n/C_{5.2}$. Then, $\vec{K}_G[R_1 \setminus U]$ contains a path with colour sequence S . Similarly, if S is a cycle candidate, $\vec{K}_G[R_1 \setminus U]$ contains a cycle with colour sequence S .

Proof. With high probability, Lemma 3.19 holds with R_1 (and R_2 set to be a p -random set independent with $R_1 - R_2$ will not be relevant in the proof). Consider a pattern (as in Definition 3.15) consisting of a directed path on $k + 1$ vertices, and label i th edge with c_i and label the first vertex of the path with v where v is a free variable. This is enough information to determine the label of the remaining k vertices: the label of the i th vertex on the path for $2 \leq i \leq k + 1$ has to be $v - \sum_{1 \leq j < i} c_j$ for the pattern to be well-defined (note also that each vertex gets a distinct label).

This pattern is well-distributed. For pairs of colours, separability follows by (b) as $c_i \neq c_j$ when $i \neq j$. For pairs of vertices, separability follows from (b) once again, this time using that S is a path-candidate (recall being a path-candidate implies that all partial sums of S are nonzero). Also, all vertices of this pattern are linear in at least one variable, namely, v . Also, all colours are constants. Therefore, by Lemma 3.19, there exists a copy of this pattern in $\vec{K}_G[R_1 \setminus U, (R_2 \setminus U) \cup S]$, which corresponds to a directed path in \vec{K}_G with S as its colour sequence. To justify this correspondence, recall Observation 3.18.

If S was a cycle-candidate instead, an analogous argument works, this time starting with a pattern that is a directed cycle of length k , with colour labels coming from S , and one of the vertices labelled v which is a free variable. Separability of colours and vertices are by (b), using the definition of a cycle-candidate. \square

Our cover-down statement for small groups of colours is the following. The proof is omitted, it follows easily by iteratively invoking Lemma 5.2.

Lemma 5.3. *Let $p \geq n^{-1/700}$. Let $3 \leq k \leq 9$. Let R_1 be a p -random subset of G . With high probability, the following holds.*

Let $C \subseteq G$ be a set of $k\ell$ colours, admitting a partition into tuples C_1, \dots, C_ℓ where each C_i is a rainbow cycle-candidate. Suppose $\ell \leq p^{400}n/kC_{5.3}$. Then, $\mathcal{H}_k[R_1 \setminus U; C]$ contains a matching of size ℓ .

5.1.3. Saturating colours: Large k

When k is large, the strategy in the previous section fails for two reasons. Firstly, assuming that the C_i are cycle-candidates as in Lemma 5.2 would be too much of an ask, as in general we do not have a good way of finding orderings of zero-sum sets in this way (for $k \leq 9$, we get to assume this without loss of generality, relying on Lemma 3.24). Even if we were able to find such orderings, there is a second issue which is probabilistic which comes up only when $k \geq \log n$. The issue is that the expected number of k -cycles using a particular colour sequence contained in a p -random set is $\leq p^k n$ – this is too small to have any useful analogue of Lemma 3.19. Therefore, we pursue a more complicated strategy as follows.

Recall that the families \mathcal{F}_G and \mathcal{S}_G as well as the parameters $q_{G,k}, f_{G,k}, s_{G,k}$ were defined in Lemma 3.27.

Lemma 5.4. *Let $10 \leq k \leq \log^{10} n$. Let G be a group of order n , where n is sufficiently large. The following statements both hold.*

1. *Let $v, w \in G$ be distinct vertices with $v - w = q_{G,k} + m \cdot f_{G,k}$ for some natural $m \leq k$. Then, there exists a family \mathcal{P} of size $\geq n/(kC_{5.4})$ of pairs of rainbow paths (P, P') in \vec{K}_G with the following properties.*
 - (a) *For each $(P, P') \in \mathcal{P}$, P starts on v , and P' ends on w .*
 - (b) *Each path in \mathcal{P} is rainbow, pairwise colour disjoint, and pairwise vertex disjoint except on $\{v, w\}$.*
 - (c) *For each $(P, P') \in \mathcal{P}$, $C(P) \cup C(P') \in \mathcal{F}_G$*
2. *Let $v, w \in G$ be distinct vertices with $v - w = s_{G,k}$. Then, there exists a family of rainbow paths \mathcal{P} of size $\geq n/(kC_{5.4})$ in \vec{K}_G with the following properties.*
 - (a) *For each $P \in \mathcal{P}$, P starts on v , and P ends on w .*
 - (b) *Each path in \mathcal{P} is rainbow, pairwise colour disjoint, and pairwise vertex disjoint except on $\{v, w\}$.*
 - (c) *For each $P \in \mathcal{P}$, $C(P) \in \mathcal{S}_G$*

Proof. We suppose $C_{5.4}$ is sufficiently large for the following calculations to go through.

Let \mathcal{P} be a maximal family with properties 1(a), 1(b), and 1(c). Suppose $|\mathcal{P}| < n/(kC_{5.4})$. Supposing that $C_{5.4} \ll C_{3.27}$, there exists at least $n/(2kC_{3.27})$ many $\mathcal{F}_i \in \mathcal{F}_G$ which are unused by \mathcal{P} . For each such unused $\mathcal{F}_i \in \mathcal{F}_G$, consider the path P_i with vertex sequence $P_{out}(v, \mathcal{F}_i^+)$ and path P'_i with vertex sequence $P_{in}(w, \mathcal{F}_i^-)$ (these are defined in Section 3.2). Note that P_i and P'_i are in fact paths as the corresponding colour sequences are path-candidates by Lemma 3.27.

Also by Lemma 3.27, the paths P_i and P'_i are vertex disjoint except possibly on v and w , as \mathcal{F}_i^+ and \mathcal{F}_i^- are separable at a distance $q + mf$. By the dissociability property coming from Lemma 3.27, for each P_i, P_j with $i \neq j$, P_i and P_j are vertex disjoint except on v , and similarly P'_i, P'_j are vertex-disjoint except on w . Using these two properties and assuming that $C_{5.4} \ll C_{3.27}$ we can find some i for which \mathcal{F}_i is unused in \mathcal{P} , (P_i, P'_i) is vertex-disjoint with all vertices included in \mathcal{P} (note there are at most $10n/kC_{5.4}$ such vertices), and $V(P_i) \cap V(P'_i) = \emptyset$, contradicting maximality of \mathcal{P} .

Now, let \mathcal{P} be a maximal family with properties 2(a), 2(b), and 2(c). Suppose $|\mathcal{P}| < n/(kC_{5.4})$. Supposing that $C_{5.4} \ll C_{3.27}$, there exists at least $n/(2kC_{3.27})$ many $\mathcal{S}_i \in \mathcal{S}_G$ which are unused by \mathcal{P} .

For each such unused $\mathcal{S}_i \in \mathcal{S}_G$, consider the path P_i with vertex sequence $P_{out}(v, \mathcal{S}_i)$ which terminates on vertex w as $v - w = s_{G,k} = \sum \mathcal{S}_i$. Note that P_i is a rainbow path because \mathcal{S}_i is a rainbow path candidate. Each P_i uses a disjoint colour set as the \mathcal{S}_i are disjoint, and all vertices of the P_i , save the endpoints v and w are disjoint, by definition of near-dissociable. Now one such path P_i has to internally vertex-disjoint with all the paths in \mathcal{P} , as the latter paths span at most $5n/(kC_{5.4}) \ll n/(2kC_{3.27})$ vertices. This means that \mathcal{P} can be extended, as desired. \square

Lemma 5.4 combined with Chernoff's bound implies the following easily.

Lemma 5.5. *Let $p \geq n^{-1/700}$, let $10 \leq k \leq \log^{10} n$. Let R_1, R_2 be p -random subsets of G sampled independently. With high probability, the following holds.*

Let $U \subseteq G$ be a set with $|U| \leq p^{50}n/(kC_{5.5})$. Let $v, w \in G$ be distinct vertices.

1. *Suppose $v - w = q_{G,k} + m \cdot f_{G,k}$ for some $0 \leq m \leq k$. There exists some $\mathcal{F} \in \mathcal{F}_G$ such that $\vec{K}_G[(R_1 \setminus U) \cup \{v, w\}; \mathcal{F}]$ has two vertex disjoint rainbow directed paths of length 2, one directed away from v , one directed into w . Furthermore, \mathcal{F} is disjoint with U .*
2. *Suppose $v - w = s_{G,k}$. There exists some $\mathcal{S} \in \mathcal{S}_G$ such that $\vec{K}_G[(R_1 \setminus U) \cup \{v, w\}; \mathcal{S}]$ contains a rainbow directed path (of length z_S) from v to w . Furthermore, \mathcal{S} is disjoint with U .*

The next lemma summarises our cover-down strategy for large k . The proof simply iterates parts (1) and (2) of Lemma 5.5, and this works due to properties acquired in Lemma 3.27.

Lemma 5.6. *Let $p \geq n^{-1/700}$, let $10 \leq k \leq \log^{10} n$. Let R_1, R_2 be p -random subsets of G sampled independently. With high probability, the following holds.*

Let $U \subseteq G$ be a set with $|U| \leq p^{600}n/(kC_{5.6})$. Suppose ℓ is some positive integer with $\ell \leq p^{600}n/(k^2C_{5.6})$. Let $C \subseteq G$ be a set of 4ℓ colours admitting a partition into 4-tuples C_1, \dots, C_ℓ where each C_i is a rainbow path candidate with sum $q_{G,k}$. Then, $\mathcal{H}_k[R_1 \setminus U; (R_2 \setminus U) \cup C]$ has a matching saturating C , and the set of colours C^ used on the matching other than C is closed under \mathcal{F}_G and \mathcal{S}_G . Moreover, C^* uses exactly $\ell(k - 4 - z_S)/4$ tuples from \mathcal{F}_G and exactly ℓ tuples from \mathcal{S}_G . As a consequence, the matching consists of exactly ℓ edges of \mathcal{H}_k .*

Proof. Fix some large K and some $C_{5.6} \gg K$. With high probability, Lemma 5.5 and Lemma 5.2 both hold.

We will first prove the statement when $\ell = 1$ with $C_{5.6}$ replaced with K (note this strengthens the statement). Initialise $U_1 = U$, and for each $i \in \{1, \dots, \ell\}$, do the following.

In $\vec{K}_G[R_1 \setminus U_i; (R_2 \setminus U_i) \cup C]$, find a rainbow path P_0 with colours C_1 via Lemma 5.2. Note that $\text{start}(P_0) - \text{end}(P_0) = q_{G,k} + m \cdot f_{G,k}$ for some natural $m \leq k$ (in fact $m = 0$). For a natural i , while $k - |P_i| > z_S$, apply Lemma 5.5(1) to extend P_i into another rainbow path P_{i+1} using a set of 4 extra colours coming from an element of \mathcal{F}_G (disjoint with U_i). Note that this preserves that $\text{start}(P_{i+1}) - \text{end}(P_{i+1}) = q_{G,k} + m \cdot f_{G,k}$ for some $0 \leq m \leq k$. Update U_{i+1} to include the new colours and vertices used in P_{i+1} .

In at most k steps, this procedure yields a path P with $k - |E(P)| = z_S$ due to the divisibility conditions coming from Lemma 3.27, note also $|U_i| \leq |U| + 10i$, so we add at most $10k$ new elements to U throughout the process. At this point, we can apply Lemma 5.5(2) to complete P into a rainbow cycle, say \mathcal{C} , and therefore an edge of \mathcal{H}_k . Also, observe that while building \mathcal{C} , we used $(k - 4 - z_S)/4$ tuples from \mathcal{F}_G and one tuple from \mathcal{S}_G .

When $2 \leq \ell \leq p^{600}n/(k^2C_{5.6})$, we can repeat the above procedure for each of C_2, \dots, C_ℓ , at each iteration including in U the set of at most $10k$ vertices and colours used in the previous step. This would add at most $10k\ell \leq 10p^{600}n/kC_{5.6}$ new elements to U , which means U will never exceed a size of $11p^{600}n/kC_{5.6} \ll p^{600}n/kK$ throughout the process. Hence at each step, we can invoke (the stronger version of) the $\ell = 1$ case. \square

5.2. Distributive absorption in \mathcal{H}_k

Definition 5.7. Let H be a hypergraph and let $\mathcal{F} = \{S_1, \dots, S_t\}$ be a family of subsets of $V(H)$. We say that a set of vertices R **absorbs** \mathcal{F} if for every subfamily $\mathcal{F}' \subseteq \mathcal{F}$ of size m , there is a hypergraph matching whose vertex set is exactly $R \cup \bigcup_{S_i \in \mathcal{F}'} S_i$.

We will build the desired absorbing structures by finding collections of small subgraphs with certain properties. Each structure found in this section is formed by combining patterns coming from Figure 1. Therefore, the following result is crucial, as it will allow us to use Lemma 3.19 numerous times throughout the section (we remark that $P_{\mathcal{F}}(V)$ is excluded in the statement of the following lemma as we treat this pattern separately).

Lemma 5.8. *In Figure 1, all patterns depicted except for $P_{\mathcal{F}}(V)$ are well-defined patterns which are well-distributed.*

Proof. To remind the reader the definitions, let us check everything explicitly for $P_x(c_1, c_2)$. All vertices and colours get labels as words in $(G * F_1)^{\text{ab}}$, vertices get distinct labels, and for each of the three directed edges \vec{e} from v to w , we have that the label of \vec{e} is the difference of the labels of v and w . This verifies that $P_x(c_1, c_2)$ is a well-defined pattern. All colour labels are linear in either c_1 or c_2 (or both), and each vertex label is either the constant x , or linear in c_1 , and of course there are at most 20 distinct free variables and each label word has length at most 20. For any two distinct colour labels, there exists a free variable (c_1 or c_2) that exists in one label with coefficient ± 1 , and doesn't exist in the other. This means that the difference of the labels is linear in at least one free variable, checking (a). Same property holds for the vertex labels as well. This verifies that $P_x(c_1, c_2)$ is a well-distributed pattern.

Generally, the definition of a pattern and the second and third part of the definition of well-distributed is easy to verify by inspecting the figure, so we focus on checking pairwise separability for vertices and colours (viewed as a set, not a multiset).

- $P_{x,z}(y, \ell_1, \ell_2)$. For this pattern and the next, we implicitly assume that $x \neq z$, making x and z separable by (b). All other pairs of vertices are separable by (a). Except for the two pairs of colours $(2\ell_1 - y - x, y - \ell_1)$ and $(2\ell_2 - y - z, z - \ell_2)$ which are separable by (c), all pairs of colours are separable by (a).
- $P'_{x,z}(y, \ell_1, \ell_2)$. This pattern is well-distributed as it is a strict subset of $P_{x,z}(y, \ell_1, \ell_2)$.
- $P_x(c_1, c_2)$. All pairs of vertices and colours are separable by (a).
- $P_x^{(4)}(c_1, c_2, c_3)$. All pairs of vertices and colours are separable by (a).
- $P_a(v, d_1, d_2)$. All pairs of vertices and colours are separable by (a).

Hence, each pattern is well-distributed as desired. □

The following remark is crucial to keep in mind as we will make many references to Figure 1 in the rest of the section.

Remark 5.9. On numerous occasions, we will consider multiple *instances* of the same *pattern type* in Figure 1. For example, $P_{x,z'}(y', \ell_1, \ell_2)$ refers to the pattern depicted in Figure 1 denoted $P_{x,z}(y, \ell_1, \ell_2)$ with y replaced with y' , z replaced with z' , but ℓ_1 and ℓ_2 unchanged. So $P_{x,z}(y, \ell_1, \ell_2)$ and $P_{x,z}(y', \ell_1, \ell_2)$ are different patterns with some overlap in the vertex and edge labels they receive.

5.2.1. Vertex-switchers

We first show how to 1-absorb a set of 2 vertices. We emphasise that when we say a vertex of \mathcal{H}_k in this section, we specifically mean a vertex which corresponds to a vertex of \vec{K}_G , as opposed to a colour of \vec{K}_G . This is crucial, as in fact it is impossible to build a 1-absorber for a set of 2 vertices of \mathcal{H}_k which correspond to colours of \vec{K}_G . This follows from Observation 2.1.

Lemma 5.10. *Let $p \geq n^{-1/700}$. Let $3 \leq k \leq \log^{10} n$. Let R_1, R_2 be p -random subsets of G sampled independently. With high probability, the following holds.*

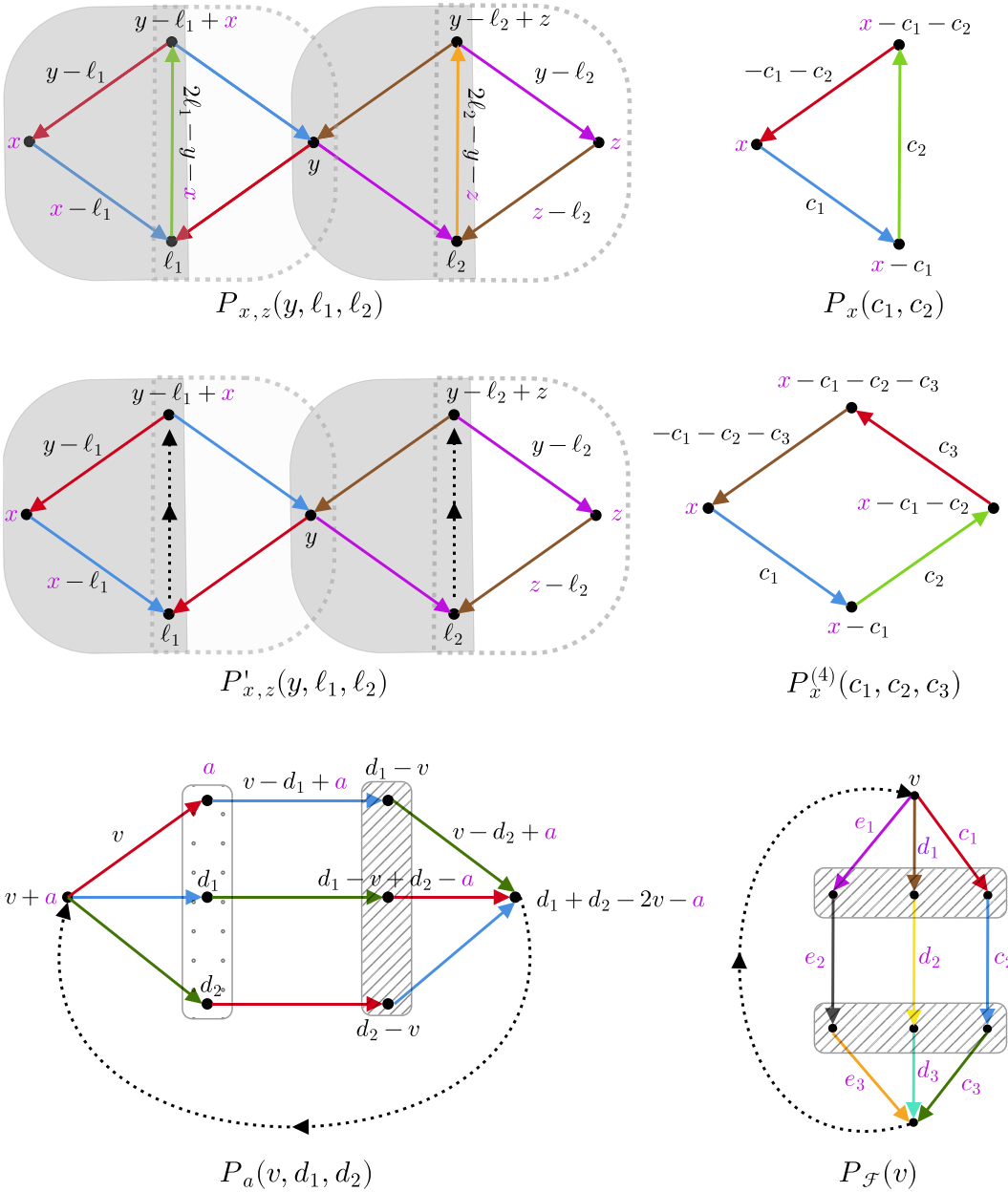


Figure 1. Several patterns (see Definition 3.15) used in Section 5.2. Edge colours correspond to edge labels, so for a collection of edges with the same colour, the label is written for only one of the edges. Free variables are denoted in black letters, and elements of G are denoted in pink letters. The dashed arrows indicate that after a copy of the pattern is found, a rainbow directed path (of the appropriate length) between the indicated vertices will be found. For triples tiled with the diagonal lines, in the proof we find a 2-absorber for the indicated vertices.

For any distinct vertices $x, z \in \mathcal{H}_k$ and $U \subseteq G$ with $|U| \leq p^{300}n/C_{5.10}$, $\mathcal{H}_k[(R_1 \setminus U) \cup \{x, z\}, (R_2 \setminus U)]$ contains a subgraph of size at most $10k$ that 1-absorbs $\{x, z\}$.

Proof. With high probability, Lemma 3.19 and Lemma 5.1 both hold. Suppose first that $k = 3$. Consider the pattern $P = P_{x,z}(y, \ell_1, \ell_2)$ from Figure 1. A copy of P can be found in $K_G[(R_1 \setminus U) \cup \{x, z\}, (R_2 \setminus U)]$

by Lemma 3.19. Inspecting the two sets of matchings (the solid matching and the dashed matching) in Figure 1 (top-left), we see that such a copy of P corresponds to a desired absorbing subgraph of $\mathcal{H}_k[(R_1 \setminus U) \cup \{x, z\}, (R_2 \setminus U)]$ of size ≤ 30 . Recall Observation 3.18 to justify this correspondence.

Suppose now that $4 \leq k \leq \log^{10} n$. In this case consider a copy of the pattern $P = P'_{x,z}(y, \ell_1, \ell_2)$ from Figure 1 given by Lemma 3.19. To complete the absorber, we need to find rainbow paths of size $k - 2 \geq 2$ from ℓ_1 to $y - \ell_1 + x$ and from ℓ_2 to $y - \ell_2 + x$ (when we say ℓ_1 , we mean the copy of the vertex with the label ℓ_1 in $P'_{x,z}(y, \ell_1, \ell_2)$, similarly for the other variables. We use this convention in the rest of this section). Two applications of Lemma 5.1(1) allows us to find these paths. As in the previous case, the solid matching and the dashed matching demonstrates that the desired absorption property holds. \square

Chaining together gadgets which 1-absorb pairs as in the previous lemma, we can construct gadgets which $(s - 1)$ -absorb sets of size s .

Lemma 5.11. *Let $p \geq n^{-1/700}$. Let $3 \leq k \leq \log^{10} n$. Let R_1, R_2 be p -random subsets of G sampled independently. With high probability, the following holds.*

Let S be a vertex-subset of size at most 100 and let $U \subseteq G$ with $|U| \leq p^{310}n/C_{5.11}$. Then, there are sets $V' \subseteq R_1 \setminus U$ and $C' \subseteq R_2 \setminus U$ of size $\leq 10^4k$ such that $V' \cup C'(|S| - 1)$ -absorbs S in \mathcal{H}_k .

Proof. With high probability, Lemma 5.10 holds.

Let $S = \{a_1, \dots, a_\ell\}$ be given for some ℓ with $2 \leq \ell \leq 100$. For every $i \in \{1, \dots, \ell - 1\}$, apply Lemma 5.10 with $(a, b) = (a_i, a_{i+1})$, each time finding a subgraph F_i which 1-absorbs $\{a_i, a_{i+1}\}$ disjointly with U . By extending U in each application to include F_1, \dots, F_{i-1} , we can also ensure the collection of F_i are vertex and colour disjoint (except for elements of S). The union of the subgraphs F_i ($i \in [\ell]$) now has the desired absorption property. For an illustration of the case when $k = \ell = 3$, see Figure 2 (left). \square

Now we show how to 1-absorb sets of size $s \leq 100$. For small k , essentially all the necessary ideas are included in the previous two lemmas, however for large k we introduce some new ideas.

Lemma 5.12. *Let $p \geq n^{-1/700}$ and let $3 \leq k \leq \log^{10} n$. Let R_1, R_2 be p -random subsets of G sampled independently. With high probability, the following holds.*

Let S be a vertex-subset of size at most 100 and let $U \subseteq G$ with $|U| \leq p^{340}n/C_{5.11}$, there are sets $V' \subseteq R_1 \setminus U$ and $C' \subseteq R_2 \setminus U$ of size $\leq 10^8k^2$ such that $V' \sqcup C'$ 1-absorbs S in \mathcal{H}_k .

Proof. With high probability, Lemma 3.19, Lemma 5.1 and Lemma 5.11 all hold. Denote $S = \{a_1, a_2, \dots, a_\ell\}$ where $2 \leq \ell \leq 100$.

Case 1: $k \in \{3, 4\}$. We write the details of the argument for $k = 3$, for $k = 4$ a proof can be obtained by replacing $P_{a_i}(c_1, c_2)$ with $P_{a_i}^{(4)}(c_1, c_2, c_3)$ in the below argument³.

Consider a pattern P_S formed by the union of patterns $P_{a_i}(c_1, c_2)$ for each $i \in [\ell]$ (recall Remark 5.9). This pattern is well-defined as the vertices of $P_{a_i}(c_1, c_2)$ and $P_{a_j}(c_1, c_2)$ get different labels when $i \neq j$ (as $a_i \neq a_j$). To see that this pattern is well-distributed we only need to check vertices and edges belonging to different copies, as each $P_{a_i}(c_1, c_2)$ is well-distributed by Lemma 5.8 already. There is nothing to check with edge labels as they are all identical. Vertices in the same position in the triangle are separable by (b) (this is as $a_i \neq a_j$ when $i \neq j$), and vertices in different positions are separable by (a). By Lemma 3.19, P_S admits a copy in $\vec{K}_G[(R_1 \setminus U) \cup S; R_2 \setminus U]$, say P'_S . For $j \in \{2, 3\}$, denote by S_2 the vertices of F'_S coming from copies of top vertices of $P_{a_i}(c_1, c_2)$ and denote by S_3 the vertices coming from bottom vertices of $P_{a_i}(c_1, c_2)$ for each $i \in [\ell]$. Apply Lemma 5.11 first with $S = S_2$, and then $S = S_3$, to find sets A_2 and $A_3(|S| - 1)$ -absorbing S_2 and S_3 , respectively. We can also ensure that A_2, A_3 and F'_S are vertex and colour disjoint by extending U in each successive application of Lemma 5.11. Now $F'_S \cup A_2 \cup A_3$ is the desired absorber. See Figure 2 for a demonstration of this when $\ell = 3$, the boxes shaded with diagonal lines represent the sets S_2 and S_3 which we can 2-absorb.

³The same argument works for each $k = O(1)$, the reason why the second case exists is the range when $k = \Omega(\log n)$.

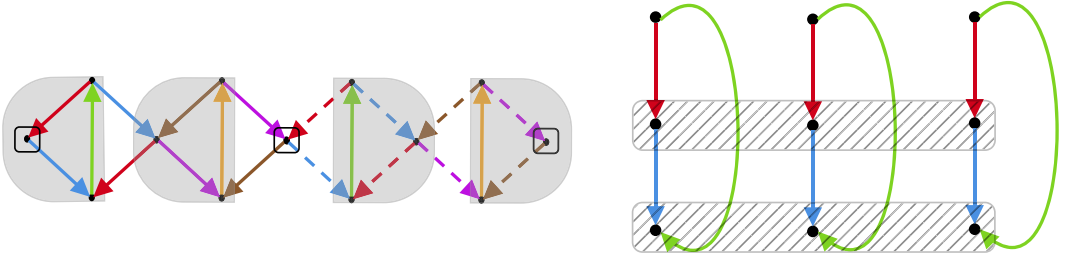


Figure 2. On the left: A 2-absorber for the 3 vertices contained in boxes (from the proof of Lemma 5.10). Dashed versions of a coloured edge are to be interpreted as having a distinct colour. The matching which absorbs the outer two vertices is indicated. On the right: The pattern P_S (from the proof of Lemma 5.12 Case 1) consisting of 3 copies of the pattern $P_x(c_1, c_2)$ ($\ell = 3$), and the sets S_2 and S_3 shaded with the diagonal lines. For the vertices covered with the diagonal lines we have a 2-absorber for the indicated vertices. The union of these two 2-absorbers and the illustrated directed graph 1-absorbs the 3-vertices on the top row of the diagram.

Case 2: $k \geq 5$. We begin by the following observation which will help to motivate the choice of parameters for the rest of the argument. We remind the reader that when we say $d_1 + d_2 - 2v - a$ below, what we mean is the copy of the vertex of $P_a(v, d_1, d_2)$ with the label $d_1 + d_2 - 2v - a$, and similarly for the other expressions.

Observation 5.13. Consider a copy P of $P_a(v, d_1, d_2)$ from Figure 1. Let Q be a rainbow path of length $k - 3$ (recall $k - 3 \geq 2$) from $d_1 + d_2 - 2v - a$ to $v + a$, colour-disjoint with P , and vertex-disjoint with P except on the endpoints. Let A be a set that 2-absorbs $Y := \{d_1 - v, d_1 - v + d_2 - a, d_2 - v\}$ (the column shaded by the diagonal lines), vertex and colour disjoint with $P \cup Q$, except on Y . Then, $P \cup A \cup Q \setminus \{a, d_1, d_2\}$ 1-absorbs the vertex-triple $\{a, d_1, d_2\}$ in \mathcal{H}_k .

Proof. Let $w \in \{a, d_1, d_2\}$. To find a matching in \mathcal{H}_k that uses w and each of $P \cup A \cup Q$ but not any of $\{a, d_1, d_2\} \setminus \{w\}$, we start by matching the edge corresponding to the rainbow k -cycle using Q and the (unique) 3-edge path in P using the vertex w , note this path also uses a unique vertex of Y . Using the definition of A2-absorbing Y , there is a matching saturating A and exactly the two unused vertices of Y . This matching combined with the matched edge corresponding to the k -cycle passing through w is the desired matching in \mathcal{H}_k verifying the definition of 1-absorbing $\{a, d_1, d_2\}$. \square

The previous observation essentially shows that we can find sets 1-absorbing triples, provided that 2 elements of the triples are *free variables* (it is important that d_1 and d_2 are free variables while ensuring that $P_a(v, d_1, d_2)$ is well distributed, and that $d_1 + d_2 - 2v - a$ is linear in at least one variable). The remainder of the proof is focused on using this property many times with carefully chosen sets of triples to find subgraphs that 1-absorb sets of size at most 100 constants (not free variables). The following observation motivates the choice of triples.

Observation 5.14. Consider a collection of sets $\{a_1, d_\ell, d_1\}$, $\{a_2, d_1, d_2\}$, $\{a_3, d_2, d_3\}$, \dots , $\{a_{\ell-1}, d_{\ell-2}, d_{\ell-1}\}$, $\{a_\ell, d_{\ell-1}\}$. Suppose for some a_i , the set containing a_i is deleted from the collection. Then, there exists a choice of an element from each of the remaining sets in the collection so that overall the chosen elements are precisely $\{d_1, \dots, d_{\ell-1}\}$.

Proof. When $i = 3$, the correct choices are displayed in the below table.

a_1	a_2	a_3	a_4	\dots	$a_{\ell-1}$	a_ℓ
d_ℓ	d_1	d_2	d_3	\dots	$d_{\ell-2}$	$d_{\ell-1}$
d_1	d_2	d_3	d_4	\dots	$d_{\ell-1}$	

For other values of i , one can similarly choose the value on the bottom row for the j th column where $j < i$, and choose the value on the middle row for the j th column where $j > i$. \square

Consider a pattern P formed by the union of patterns $P_{a_1} := P_{a_1}(v_1, d_\ell, d_1), P_{a_2} := P_{a_2}(v_2, d_1, d_2), P_{a_3} := P_{a_3}(v_3, d_2, d_3), \dots, P_{a_{\ell-1}} := P_{a_{\ell-1}}(v_{\ell-1}, d_{\ell-2}, d_{\ell-1})$ (recall Remark 5.9). Formally, P is obtained by taking the (disjoint) union of each of the graphs P_{a_i} and identifying together vertices which share the same label (this identification step ensures that the second property in the definition of pattern is satisfied, so P is indeed a well-defined pattern). We remark that the vertex/edge labels live in the set $(G * F_{2\ell-1})^{ab}$.

Claim 5.14.1. P is well-distributed.

Proof. Each $P_{a_i}(\cdot, \cdot, \cdot)$ is well-distributed by Lemma 5.8, so we only need to check separability for pairs of vertices and pairs of colours coming from different $P_{a_i}(\cdot, \cdot, \cdot)$. For pairs of such colours, the word corresponding to a label includes v_i as a free variable for some i , and the other label does not include v_i . This makes the pair separable by (a). Similarly, for pairs of vertices, if one of the words has a v_i as a free variable for some i , we have separability by (a). If this is not the case the pair of words could be of the form (d_i, d_j) for some i and j . If $i \neq j$, we have separability by **a**, and otherwise, the words are identical and in this case we do not need to check separability, as in the definition of the union we identify such vertices together. The pair of words could also be of the form $(d_i, d_j + d_{j+1} - 2v_j - a_j)$, in which case either d_j or d_{j+1} is different with d_i , giving separability by **a**. Finally, pair of words could also be of the form $(d_i + d_{i+1} - 2v_i - a_i, d_j + d_{j+1} - 2v_j - a_j)$ where i and j are distinct, so we again have separability by **a**. \square

By Lemma 3.19, we can find a copy of P , say P' , in $\vec{K}_G[R_1 \setminus U; R_2 \setminus U]$. In addition, by Lemma 5.10 applied with $\{a, b\} = \{a_\ell, d_{\ell-1}\}$, we obtain a subgraph A which 1-absorbs $\{a_\ell, d_{\ell-1}\}$. By extending U in this application, we can ensure A is disjoint with P' .

Recall that $k - 3 \geq 2$ by assumption. For each $i \in [\ell - 1]$, apply Lemma 5.1 with u set to be the copy of the rightmost vertex in P_{a_i} and v set to be the copy of the leftmost vertex in P_{a_i} to find a rainbow path of length $k - 3$ directed from u to v that does not clash with any of the forbidden colours or vertices (see the dashed line in Figure 1). We can achieve this by iteratively invoking Lemma 5.1, extending U at each step. Note that $k \leq \log^{10} n$, so we never add more than $200 \log^{10} n$ elements to U during this process. Again, for each $i \in [\ell - 1]$, apply Lemma 5.11 with S set to be the subset corresponding to the copies of the 3 vertices of P_{a_i} which correspond to the column indicated with the dots to obtain a subgraph A_{a_i} which 2-absorbs this subset. We can ensure that the collection of A_{a_i} are pairwise disjoint (except for the vertices corresponding to highlighted vertices plugged into S), and also disjoint with U , again by extending U in each application of Lemma 5.11. By the bound coming from Lemma 5.11, we never have to extend U by more than $10^6 k^2 \ell$ elements during this process.

For each $i \in [\ell - 1]$, let P'_{a_i} be the copy of P_{a_i} combined with the path found by applying Lemma 5.1 and A_{a_i} , and let Z_{a_i} be the set of vertices corresponding to the copies of the vertices of P_{a_i} (the column indicated by dots). The following is a rephrasing of Observation 5.13.

Observation 5.15. We have that $P'_{a_i} \setminus Z_{a_i}$ 1-absorbs Z_{a_i} .

Now, we claim that $(\bigcup P'_{a_i} \setminus S) \cup A$ is the desired absorber. To see this, let $a_i \in S$. We wish to show that the vertices and colours of $(\bigcup P'_{a_i} \setminus S) \cup A \cup \{a_i\}$ induce a perfect matching in \mathcal{H} . To see this, first take a perfect matching in $(P'_{a_i} \setminus Z_{a_i}) \cup \{a_i\}$ (which exists by definition of 1-absorbing and Observation 5.15).

Now, from each set Z_{a_j} (for $i \neq j$) and $\{a_\ell, d_{\ell-1}\}$ (coming from the 1-absorbing A), we can select exactly one $d_{j'}$ so that each $d_{j'}$ (for each $j' \in [\ell - 1]$) is selected precisely once, using Observation 5.14. Using either Observation 5.15 or the 1-absorbing property of A , we can find a perfect matching of $(\bigcup P'_{a_i} \setminus S) \cup A \cup \{a_i\}$ as required. \square

5.2.2. Colour-switchers

Lemma 5.16. *Let $p \geq n^{-1/700}$. Let $4 \leq k \leq \log^{10} n$. Let R_1, R_2 be p -random subsets of G sampled independently. With high probability, the following holds.*

Let $\alpha \in G$, and $s \in \mathbb{N}$ with $2 \leq s \leq \min\{k - 2, 100\}$. Let S be a disjoint and near-dissociable family of rainbow s -tuples of colours, each tuple sums to α , and each tuple is a path-candidate, and suppose $|S| \leq 100$. Let $U \subseteq G$ with $|U| \leq p^{400}n/C_{5.16}$. Then, there are sets $V' \subseteq R_1 \setminus U$ and $C' \subseteq R_2 \setminus U$ of size $\leq 10^{11}k^2$ such that $V' \cup C'$ 1-absorbs S in \mathcal{H}_k .

Proof. With high probability, Lemma 3.19, Lemma 5.11 and Lemma 5.1 hold. Let S and U be given as in the statement. Consider a pattern P constructed as follows. Take $|S|$ directed paths, each of length s , with the same start and end vertices, but internally vertex-disjoint. Label the start vertex with the free variable v . Label the edges of the i th directed path with the i th s -tuple of S (counting in order of proximity to v). Note this induces a labelling on each of the remaining vertices of P . This labelling is well-defined on the end-vertex, because each tuple in S has the same sum, namely, α . In particular, the end-vertex receives the label $v - \alpha$. For an illustration of the pattern when $s = 3$ and $S = \{(e_1, e_2, e_3), (d_1, d_2, d_3), (c_1, c_2, c_3)\}$, inspect the bottom-right pattern in Figure 1.

Observation 5.17. P is well-distributed.

Proof. Each pair of colours is separable by \mathbf{b} as distinct coordinates of elements of S are distinct, and the elements of S are pairwise disjoint. We claim each pair of vertices is separable by \mathbf{b} . For vertices belonging to the same directed path, this follows as elements of S are path-candidates. For vertices belonging to different directed paths, this follows as S is near-dissociable. \square

Thus, we may apply Lemma 3.19 to find a copy of P , say P' , in $\vec{K}_G[R_1 \setminus U; (R_2 \setminus U) \cup \cup S]$. For each $i \in [s - 1]$, let P_i denote the vertices of distance i from v , noting $|P_i| = |S| \leq 100$. Apply Lemma 5.11 for each P_i to find (disjointly) sets P'_i which $|S| - 1$ absorb P_i . Finally, noting that $k - s \geq 2$, apply Lemma 5.1 with u as the end-vertex of the path of length s from v , and $v = v$, and $k' = k - s$, to find a rainbow path of length $k - s$. It is easy to see that the resulting structure has the desired absorption property. \square

5.2.3. Putting the gadgets together

So far we have lemmas allowing us to find sets 1-absorbing arbitrary sets of size ≤ 100 . How can we go from to sets which can h absorb sets of size $(1 + \beta)h$, where h and βh are potentially linear in n ? The bipartite graph given in the next lemma gives us a nice collection of subsets of size at most 100 to 1-absorb, which together have the desired property. In this sense the utility of this bipartite graph is similar in spirit to that of the sequence constructed in Observation 5.14 (which could be viewed as a bipartite graph with maximum degree 3 with much weaker properties).

Lemma 5.18 (Montgomery, [28]). *Let $0 < \beta \leq 1$. There is a positive integer h_0 such that for every $h \geq h_0$ there exists a bipartite graph K with maximum degree at most 100 and vertex classes X and $Y \cup Y'$ with $|X| = 3h$, $|Y| = 2h$, $|Y'| = h + \beta h$ so that the following holds. For any $Y_0 \subseteq Y'$ with $|Y_0| = h$, there is a perfect matching between X and $Y \cup Y'$.*

Graphs produced by this lemma are called *robustly matchable bipartite graphs*.

Lemma 5.19. *Let $p \geq n^{-1/700}$, $3 \leq k \leq \log^{10} n$. Let R_1, R_2 be p -random subsets of G sampled independently. With high probability, the following holds.*

Let $0 < \beta \leq 1$. Let $h \in \mathbb{N}$ with $\sqrt{n} \leq h \leq p^{400}n/C_{5.19}k^2$. Let $U \subseteq G$ with $|U| \leq n^{999/1000}$. Let Y' be a subset of size $(1 + \beta)h$ with one of the following forms.

1. $Y' \subseteq \mathcal{H}_k$ is a vertex-subset of \mathcal{H}_k .
2. Let $\alpha \in G$, and $s \in \mathbb{N}$ with $2 \leq s \leq \min\{k - 2, 100\}$. Y' is a disjoint and near-dissociable family of s -tuples of colours of \mathcal{H}_k where each tuple sums to α , and each tuple is a path-candidate (hence $\alpha \neq 0$).

Then, there exists a set $A \subseteq (R_1 \cup R_2) \setminus U$ of size $\leq 10^{15}k^2 3h$ where Ah -absorbs Y' .

Proof. With high probability, Lemma 5.12 holds and R_2 has size at least $10h$. Also, Lemma 3.19 holds.

We show how to prove part (1) of the statement using Lemma 5.12 first.

Let β, h, U , and Y' be given. As h is sufficiently large, we can apply Lemma 5.18 to construct a bipartite graph G with parameters β and h with vertex classes X and $Y \cup Y'$. Here, we associate the given set of vertices Y' with the Y' that denotes a set of vertices of G . We arbitrarily associate, disjointly with Y' and U , a subset of vertices of \mathcal{H}_k from R_2 with Y ($|R_2| \geq 10h$, so there is space to do this). Now, each element $x \in X$ is linked, via the graph G , to a subset of vertices of \mathcal{H}_k of size at most 100, that is, the neighbourhood which we denote $N_G(x)$. For each element of $x \in X$, we will find a $A_x \subseteq \mathcal{H}_k$ that 1-absorbs $N_G(x)$, and the collection of A_x we find will be disjoint except on elements of $Y \cup Y'$. The property from Lemma 5.12 allows us to do this greedily, extending U with $10^8 k^2$ elements at each step, adding to U at most

$$10^8 k^2 |X| \leq 10^8 k^2 3h \leq 3 \cdot 10^8 k^2 p^{400} n / (C_{5.19} k^2) \leq 3 \cdot 10^8 p^{400} n / C_{5.19}$$

elements. Combined with the initial elements of U , this means that U never exceeds a size of $p^{340} n / C_{5.12}$ if $C_{5.19}$ is sufficiently large. This means that the applications of Lemma 5.12 are valid.

We claim that the union of the A_x with Y have the desired absorption property. To see this, take some subset $Y_0 \subseteq Y'$ of size h . In G , we have perfect matching f from X to $Y \cup Y_0$. For each A_x , use the matching of $A_x \cup \{f(x)\}$ which exists by the absorption property of A_x . These matchings together give a matching of $\bigcup A_x \cup Y \cup Y_0$, as required.

The proof for part (2) is essentially the same, using Lemma 5.16 instead. Note in this case we may assume $k \geq 4$, otherwise there is no s for which we have to prove anything. Let β, h, U , and Y' be given as a near-dissociable family of s -tuples (with an associated α and $s \geq 2$). As h is sufficiently large, we can apply Lemma 5.18 to construct a bipartite graph G with parameters β and h with vertex classes X and $Y \cup Y'$ (associate arbitrarily the elements of the two sets Y' as before). Arbitrarily associate to Y disjoint s -tuples with the same structural properties as Y' , that is, near-dissociable and with sum $\alpha \neq 0$ (to find such tuples, as $s \leq 100$, we may represent such tuples as well-distributed patterns and use Lemma 3.19). As before, each vertex $x \in X$ is associated to a subset of at most 100 many s -tuples, and we may find a $A_x \subseteq \mathcal{H}_k$ so that A_x 1-absorbs the s -tuples that appear in its neighbourhood in G . Via Lemma 5.16, we may find the A_x disjointly (except on $Y \cup Y'$). By a similar calculation, the union of the A_x with the s -tuples in Y has the desired property. \square

5.3. Proof of Lemma 2.4

Now we combine the distributive absorption strategy with the cover-down strategy to give a proof of Lemma 2.4. Recall the convention about random subsets of random sets given before the proof of Theorem 2.3.

Proof. Let $K = K_{2.4} \geq 1$ be sufficiently large, and fix $\varepsilon = \varepsilon_{2.4} \ll 1/K$, so that in particular, $\varepsilon K \leq 10^{-10}$ holds, and $n^{-\varepsilon_{2.4}}$ is sufficiently large so that the upcoming randomness parameters are large enough that the applications of the various lemmas are valid. Fix some $p \geq n^{-\varepsilon_{2.4}}$.

Case 1: $3 \leq k \leq 9$. Write $p = p_1 + p_2$ where $p_2 = p^{500} / (10C_{5.19} k^3)$. Partition R_i into disjoint p_1 and p_2 -random sets $R_i^{(1)}$ and $R_i^{(2)}$ for each $i \in [2]$. Let $R_i^* \subseteq R_i^{(2)}$ be a r -random subset of G where $r = p^{500} / (1000C_{5.19} k^3)$. With high probability, $R_1^{(2)}$ satisfies Lemma 5.3, R_2^* satisfies Lemma 3.25 as well as Lemma 3.22, (R_1^*, R_2^*) satisfies Lemma 5.1, and $(R_1^{(1)}, R_2^{(1)})$ satisfies Lemma 5.19 (the necessary lower bounds for the corresponding randomness parameters in each of these applications is satisfied for a small enough value of $\varepsilon_{2.4}$). With high probability, the size of each random set is at most $n^{0.6} \log n$ away from its expectation. All these properties hold simultaneously with high probability.

Now, let $U \subseteq G$ with $|U| \leq n^{4/5}$, without relabelling, include 0 in U . By Lemma 3.22, we can find a subset $R_2^{**} \subseteq R_2^* \setminus U$ using all but at most k elements of $R_2^* \setminus U$ such that $\sum R_2^{**} = 0$ and k divides

$|R_2^{**}|^4$. Set β and h so that they satisfy the two identities $(1 + \beta)h = |R_1^{(2)} \setminus U|$ and $\beta h = |R_2^{**}|$ (so $h = |R_1^{(2)} \setminus U| - |R_2^{**}| \geq \sqrt{n}$, and $0 < \beta \leq 1$ by choice of r). Apply Lemma 5.19(1) with these values of β and h and $Y' := R_1^{(2)} \setminus U$ to obtain an absorbing set A contained in $R_1^{(1)} \cup R_2^{(1)} \setminus U$ (the necessary upper bound of $h \leq p^{400}n/C_{5.19}k^2$ holds by definition of p_1, p_2 (we simply need $p_1 \geq p/2 \geq n^{-\varepsilon_{2.4}}/2$) using also that each random set has size close to its expectation).

We claim now that $A \cup R_2^{**} \cup (R_1^{(2)} \setminus U)$ has the desired property (of $V \cup C$ in the statement). To see this, take $V', C' \subseteq G$ with $|V'| = |C'| = m$ as in the statement of the lemma. As $m \ll r^{300}n/kC_{5.1}$ (supposing $K_{2.4}$ is sufficiently large), by Lemma 5.1(2), there exists a matching M_1 of size exactly m in \mathcal{H}_k saturating V' and using exactly $(k - 1)m$ vertices from $R_1^* \setminus U$ and km vertices from R_2^{**} . $C'' := C' \cup (R_2^{**} \setminus V(M_1))$ is a zero-sum set whose order is divisible by k with small symmetric difference with R_2^* (note that $|C''| = m + |R_2^{**}| - km$ (which is a quantity divisible by k), so $|C'' \Delta R_2^*| \leq 10km + n^{4/5} \leq 10k(p/k \log n)^K n + n^{4/5} \leq r^{1010} n / \log(n)^{10^{23}}$ supposing K is sufficiently large). Hence C'' can be partitioned into k -sets which are cycle-candidates by Lemma 3.25(1). This partition allows us to apply Lemma 5.35 to deduce that there exists a matching M_2 saturating the colours C'' using exactly $|C''|$ vertices from the set $R_1^{(2)} \setminus U \cup V(M_1)$. Observe that in total we used exactly $|R_2^{**}| = \beta h$ vertices from $R_1^{(2)} \setminus U$, and therefore the remaining vertices in $R_1^{(2)} \setminus U$ combined with A admits a perfect matching M_3 by the absorption property of A . Then, $M_1 \cup M_2 \cup M_3$ is the desired perfect matching of $A \cup R_2^{**} \cup V' \cup C' \cup (R_1^{(2)} \setminus U)$.

Case 2: $k \geq 10$. Set $q_1 = p^{500}/(10^{10}C_{5.19}k^2)$, $q_2 = (p - q_1)/3$, $r_* = q_1/1000k^{10}$. Let $R_1^{(1)}, R_1^{(2)}, R_1^{(3)}, R_1^{(4)}$ be disjoint subsets of R_1 , and q_1, q_2, q_2, q_2 -random, respectively. Let $R_1^{(1,1)}, R_1^{(1,2)} \subseteq R_1^{(1)}$ be r_* -random and $(q_1 - r_*)$ -random and disjoint. Let $R_2^*, R_2^{(1)}, R_2^{(2)}, R_2^{(3)}, R_2^{(4)}$ be disjoint subsets of R_2 and $r_*, (q_1 - r_*)$, q_2, q_2 and q_2 -random, respectively.

With high probability, Lemma 5.1 holds for $(R_1^{(1,1)}, R_2^*)$, Lemma 3.25 holds for R_2^* , Lemma 5.6 holds for $(R_1^{(1,2)}, R_2^{(1)})$, Lemma 5.19 holds for each of $(R_1^{(2)}, R_2^{(2)})$, $(R_1^{(3)}, R_2^{(3)})$, $(R_1^{(4)}, R_2^{(4)})$, Lemma 3.22 holds for R_2^* , and the size of each random set is at most $n^{0.6} \log n$ away from its expectation. These applications are valid supposing $\varepsilon_{2.4}$ is small enough, that is, p is large enough, just like in the previous case.

Let U be given, as before, include 0 in U . Fix f to be the largest integer bounded above by $|R_2^* \setminus U|$ with the property that $f - (k - 1)m$ is divisible by 4. By Lemma 3.22, we can fix a f -subset $R_2^{**} \subseteq R_2^* \setminus U$ using all but at most 4 vertices from the latter set such that $\sum R_2^{**} = ((f - (k - 1)m)/4) \cdot q_{G,k}$ (recall Lemma 3.27 for definition of $q_{G,k}$, and here we can select $\varepsilon := n^{-0.001}$ and $m := |R_2^* \setminus U| - 4$ in the application of Lemma 3.22).

Set β_1 and h_1 be so that $(1 + \beta_1)h_1 = |R_1^{(1)} \setminus U|$ and $\beta_1 h_1 = (k - 1)m + k(f - (k - 1)m)/4$. Denote by \mathcal{F}'_G the family of sets from \mathcal{F}_G which are entirely contained in $R_2^{(1)} \setminus U$. Similarly, denote by \mathcal{S}'_G the family of sets from \mathcal{S}_G which are entirely contained in $R_2^{(1)} \setminus U$. Set β_2 and h_2 so that $(1 + \beta_2)h_2 = |\mathcal{F}'_G|$ and $\beta_2 h_2 = ((f - (k - 1)m)/4)(k - 4 - z_S)/4$ (recall this is an integer by Lemma 3.27). Set β_3 and h_3 so that $(1 + \beta_3)h_3 = |\mathcal{S}'_G|$ and $\beta_3 h_3 = (f - (k - 1)m)/4$.

Apply Lemma 5.19(1) with $(R_1^{(2)}, R_2^{(2)})$ and $Y' = R_1^{(1)} \setminus U$ with parameters (β_1, h_1) to obtain a set A_1 (disjoint with U) with a vertex-absorption property. Apply Lemma 5.19(2) with $(R_1^{(3)}, R_2^{(3)})$ and $Y' := \mathcal{F}'_G$ with parameters (β_2, h_2) to obtain a set A_2 (disjoint with U and A_1) with a colour-absorption property. Similarly, apply Lemma 5.19(2) with $(R_1^{(4)}, R_2^{(4)})$ and $Y' := \mathcal{S}'_G$ with parameters (β_3, h_3) to obtain a set A_3 (disjoint with U, A_1 , and A_2) with a colour-absorption property. For the last two applications, we use that \mathcal{F}_G and \mathcal{S}_G are near-dissociable, contain only path-candidates, and

⁴Lemma 3.22 can be applied for example with $\varepsilon = n^{-0.001}$ (note $k, |U| \leq 9 \leq n^{9/10}$), $g = 0$, $m := |R_2^* \setminus U| - k$, $Z = R_2^* \setminus U$

⁵Note that the hypothesis $\ell \leq p^{400}n/kC_{5.3}$ because $r = p^{500}/(1000C_{5.19}k^3)$.

that $k - 2 \geq z_S, 4$ as $k \geq 10$. These properties come from Lemma 3.27. For all three applications, the necessary upper bound on h holds by definition of q_1, q_2 using that each random set has size close to its expectation. The lower bounds on h and that $0 < \beta \leq 1$ for the latter two applications follow from lower bounds on the sizes of \mathcal{F}'_G and \mathcal{S}'_G which can be derived from Lemma 5.6 (this is done implicitly in the rest of the argument).

We claim that $A_1 \cup A_2 \cup A_3 \cup R_2^{**} \cup (R_1^{(1)} \setminus U) \cup \mathcal{F}'_G \cup \mathcal{S}'_G$ has the desired absorption property. To see this, let V' and C' be given as in the lemma. By Lemma 5.1(2), there exists a matching M_1 in \mathcal{H}_k saturating V' and using exactly $(k - 1)m$ vertices from $R_1^{(1,1)} \setminus U$ and km vertices from R_2^{**} . $C' \cup (R_2^{**} \setminus V(M_1)) := C''$ then has size $m + f - km = f - (k - 1)m$ which is divisible by 4 by choice of the integer f . Furthermore, $\sum C'' = (|C''|/4) \cdot q_{G,k}$ by the sum property on the set R_2^{**} . Hence, C'' can be partitioned into 4-tuples with sum $q_{G,k}$ (recall this is not 0) which are path-candidates by Lemma 3.25(2) (as in the previous case, to check that C'' has small symmetric difference with R_2^{**} , recall that K is sufficiently large). This partition of C'' allows us to apply Lemma 5.6 (with $\ell = (f - (k - 1)m)/4$) to deduce that there exists a matching M_2 saturating C'' using (exactly $k\ell = k(f - (k - 1)m)/4$ many) vertices from $R_1^{(1,2)} \setminus U$ and colours from $R_2^{(1)} \setminus U$ which are closed under the families \mathcal{F}_G and \mathcal{S}_G , and hence also closed under the families \mathcal{F}'_G and \mathcal{S}'_G (as the colours come from the set $R_2^{(1)}$). Lemma 5.6 also guarantees that M_2 uses $\ell(k - 4 - z_S)/4$ elements of \mathcal{F}'_G and ℓ elements of \mathcal{S}'_G . Thus, there are exactly h_1 elements of $R_1^{(1)} \setminus U$, h_2 elements of \mathcal{F}'_G , and h_3 elements of \mathcal{S}'_G that are unused by $M_1 \cup M_2$, so the leftovers of these sets combine with A_1, A_2 and A_3 (respectively) to produce perfect matchings, say M_3, M_4 and M_5 . Then, $\bigcup_{i \in [5]} M_i$ is the desired matching. \square

6. The high-girth case

In this section, we show how the high-girth case of the FGT conjecture follows by results from [30].

Lemma 6.1 [30]. *Let $1/n \ll p \leq 1$, let t be a positive integer between $\log^7(n)$ and $\log^8(n)$, and let q satisfy $p = (t - 1)q$. Let G be an abelian group of order n . Let $V_{str}, V_{mid}, V_{end}$ be disjoint random subsets with V_{str}, V_{end} q -random and V_{mid} p -random. Let C be a $(q + p)$ -random subset, sampled independently with the previous sets. Then, with high probability, the following holds.*

Let $V'_{str}, V'_{end}, V'_{mid}$ be disjoint subsets of G , let C' be a subset of G , and let $\ell = |V'_{mid}|/(t - 1)$. Suppose all of the following hold.

1. *For each random set $R \in \{V_{str}, V_{mid}, V_{end}, C\}$, we have that $|R\Delta R'| \leq n^{0.6}$.*
2. $\sum V'_{str} - \sum V'_{end} = \sum C'$
3. *$e \notin C'$ if G is an elementary abelian 2-group.*
4. $\ell := |V'_{str}| = |V'_{end}| = |V'_{mid}|/(t - 1) = |C'|/t$

Then, given any bijection $f: V'_{str} \rightarrow V'_{end}$, we have that $\vec{K}_G[V'_{str} \cup V'_{end} \cup V'_{mid}; C']$ has a rainbow \vec{P}_t -factor where each path starts on some $v \in V'_{str}$ and ends on $f(v) \in V'_{end}$.

Theorem 6.2. *Let G be an abelian group of order n , where n is sufficiently large. Suppose k is some integer such that $k \geq \log^9 n$, and k divides $n - 1$. Suppose $\sum G = 0$. Then, $\mathcal{H}_k[G \setminus \{0\}; C \setminus \{0\}]$ has a perfect matching.*

Proof. If $k \leq n^{1/10^{10}}$, set $s = k$, otherwise set $s = \lceil \log^{10} n \rceil$.

Partition the group G into disjoint sets twice, independently, as V_1, \dots, V_s and C_0, \dots, C_{s-1} where each set is $(1/s)$ -random, noting $1/s \geq n^{-1/10^{10}}$ in either case for n large. Set $t := \lceil \log^7 n \rceil$.

Lemma 6.1 holds with high probability with $t, V_{mid} = \bigcup_{1 \leq i \leq t-2} V_i$ and $C = \bigcup_{0 \leq i \leq t-2} V_i$. Lemma 3.5 holds with random sets (V_i, V_{i+1}, C_i) for each i (where indices are viewed in a cyclic order) and each integer value of $\ell = n/s \pm n^{1-1/10^{10}}$ (we achieve this via a union bound over many applications of Lemma 3.5). Also with high probability, all random sets are within $n^{0.6}$ elements of their expectations via

Chernoff's bound. By the probabilistic method, fix the random sets so they have all the aforementioned properties.

Suppose first that $k \leq n^{1/10^{10}}$, so $s = k$. By the divisibility assumption and the property coming from Chernoff's bound, we can move $O(n^{0.7})$ elements between the sets V_i without relabelling so that each set V_i has size exactly $(n-1)/k$. Similarly, moving around at most $O(n^{1-10^6})$ elements, we can make sure each C_i where $i \geq t-1$ has $(n-1)/k + \lfloor n^{1-10^5} \rfloor$ elements. Now, apply Lemma 3.5 (with $\ell = (n-1)/k$) with the triples

$$(V_{t-1}, V_t, C_{t-1}), (V_t, V_{t+1}, C_t), \dots, (V_{k-1}, V_k, C_{k-1})$$

to find rainbow matchings saturating the corresponding vertex sets (and missing $\lfloor n^{1-10^5} \rfloor$ colours from each C_i , $i \geq t-1$). Note that the union of the matchings found give a rainbow \vec{P}_{k-t} factor where each path is directed from V_{t-1} to V_k . Now, we apply Lemma 6.1 with $V_{str} = V_k$ and $V_{end} = V_{t-1}$, and C' set to be the union of C and the $(k-t+1)\lfloor n^{1-10^5} \rfloor$ unused colours in each C_i , $i \geq t-1$. V_{mid} remains unchanged. All but the second hypothesis of Lemma 6.1 follow easily from our choice of sets. To see that $\sum V_k - \sum V_{t-1} = \sum C'$, first note that $\sum V_{t-1} - \sum V_k = \sum C''$ where C'' is all of the colours used via applications of Lemma 3.5 (this comes from the fact that we have a rainbow directed path factor in \vec{K}_G where each path starts in V_{t-1} and ends in V_k). As $\sum G \setminus \{0\} = 0$ by assumption, and $C' = G \setminus \{0\} \setminus C''$, the desired equality follows. Thus we can indeed apply Lemma 6.1. In our application, we set f to be the bijection that maps the last endpoint of each \vec{P}_{k-t} to the first endpoint of the same directed path. This allows us to complete each \vec{P}_{k-t} into a cycle of length k , giving us a cycle-factor that corresponds to the desired matching in \mathcal{H}_k .

Suppose now that $k > n^{1/10^{10}}$, so $s = \lceil \log^{10} n \rceil$. If it was the case that s divides k , then we can proceed exactly like the previous case, with the only difference being in the choice of f in the previous paragraph (we would choose f so that when the connecting paths are found we end up with a C_k -factor as opposed to a C_s -factor). So suppose that r , the remainder when k is divided by s , is positive, noting that $r < s$. We start by finding $(n-1)/k$ vertex/colour disjoint rainbow \vec{P}_r in \vec{K}_G , calling this collection of paths \mathcal{P} . Note this can be done greedily, and the resulting collection of paths occupies $2n^{1-1/10^{10}}$ vertices, due to our assumption on k and s . Let P_1 and P_2 denote collection of first endpoints of each of the paths in \mathcal{P} , respectively. We remove the vertices in $\mathcal{P} \setminus P_2$ from the graph, and proceed exactly as in the previous case to redistribute the sets so that they are of the right size, with the additional condition that $P_2 \subseteq V_{t-1}$. In the end, while applying Lemma 6.1, we set V_{end} to be $(V_{t-1} \setminus P_2) \cup P_1$. We can then select an appropriate bijection f so that after an application of Lemma 6.1, the resulting structure is a \vec{C}_k -factor. \square

7. Concluding remarks

In this section we outline some directions for further research.

7.1. Further applications in graph labelling

As alluded to previously, our methods are quite flexible and can potentially be used to make advances on other embedding problems with an algebraic flavour. The language of *graph labellings* makes this connection explicit. Suppose we have a graph T with m edges, and a set L of $\geq m$ labels with some algebraic structure. For example, when L is the cyclic group on m elements, a labelling of the vertices of T with L is called a **harmonious labelling** if for each edge $e = \{x, y\}$ of T , we have that $x + y \pmod{m}$ is distinct. Similarly, if L is the first m positive integers, a vertex-labelling where for each edge we have that $|x - y|$ is distinct is called a **graceful labelling**. The famous harmonious tree and the graceful tree conjectures respectively assert that all trees have harmonious and graceful labellings

[21, 18]. Harmonious and graceful labellings are also heavily investigated for other classes of graphs due to intimate connections with the theory of error correcting codes (here we refer the reader to [18]).

Section 2.1 establishes that the FGT conjecture can be framed as a problem about labelling the vertices of a collection of vertex-disjoint directed cycles, where the edge labelling rule is $a - b$ (if ab is an edge of a directed cycle) within an abelian group G . We develop general methods for labellings of short paths and cycles in the current paper, and it is likely that our methods could give well-behaved labellings of other sparse structures build up of short paths and cycles.

For example, we would not expect that working with the labelling rule $a + b$ (as opposed to $a - b$) creates dramatic complications (see Remark 1.5), therefore our methods seem applicable to the study of harmonious labellings. Harmonious labellings of various sparse graph classes (such as powers of paths, which in particular contain collections of vertex-disjoint cycles) have been separately investigated, using ad-hoc methods, and ‘few general results are known’ as Gallian notes in his survey [18]. We believe the methods we develop here could be used as a unifying method to produce classes of harmonious graphs (those graphs which admit harmonious labellings).

The connection with the study of graceful labellings is less direct, as the function $(x, y) \rightarrow |x - y|$ over the integers behaves rather differently compared to the function $x + y$ over the cyclic group. That said, the function $(x, y) \rightarrow |x - y|$ still exhibits a lot of algebraic symmetry, hence there is potential for our methods to be applicable.

7.2. Nonabelian groups, Latin squares, and Ryser’s conjecture

Now that the Friedlander–Gordon–Tannenbaum conjecture is verified, at least for sufficiently large groups, we propose the following extension for general groups.

Conjecture 7.1. *Let G be a sufficiently large group satisfying the Hall–Paige condition, and suppose $k \geq 3$ and k divides $n - 1$. Then, there exists an orthomorphism of G that fixes the identity element, and permutes the remaining elements of disjoint cycles of length k .*

It is also sensible to replace orthomorphisms with complete mappings in the above conjecture, due to the assumption that $k \geq 3$ (recall Remark 1.5 from the Introduction). One way of attacking the above conjecture would be to try to combine the methods from this paper with the methods developed for nonabelian groups in [30]. Also, we remark that in Section 6, we did not actually use that the group G is abelian. Therefore, the above conjecture is true in the high-girth case.

More generally, we can turn our attention to Latin squares, which are also known as quasi-groups. These objects can be described as n by n arrays filled with n symbols so that no symbol repeats in a row or a column. For us, it will be more natural to view Latin squares in the following way (see the survey by Pokrovskiy from [31] for a more detailed discussion). We first take a complete directed graph \vec{K}_n with edges in both directions between all vertices and a loop at every vertex. We then equip this graph with a proper edge-colouring using n colours. The most famous conjecture in the area is the following.

Conjecture 7.2 (Ryser’s conjecture). *Suppose n is odd. Then, \vec{K}_n contains a rainbow spanning subgraph where every vertex has in-degree and out-degree equal to one. Equivalently, \vec{K}_n can be packed with directed cycles in a rainbow fashion.*

In analogy with the Friedlander–Gordon–Tannenbaum conjecture, it makes sense to strengthen Ryser’s conjecture to ask for cycles of specific lengths. There are numerous conjectures in this area which focus on finding a single cycle which covers the entirety of the vertex set, which is analogous to the $k = n - 1$ case of the Friedlander–Gordon–Tannenbaum conjecture. For more information about these conjectures, we refer the reader to Pokrovskiy’s survey about rainbow subgraphs in [31] and a recent paper by Gould and Kelly which includes a nice unifying conjecture [19]. We pose a conjecture in the other extreme, where the cycle lengths are as small as possible. This is analogous to the $k = 3$ case of the Friedlander–Gordon–Tannenbaum conjecture.

Conjecture 7.3. *Let K_n be a complete graph properly coloured with n colours. Then, K_n contains a rainbow subgraph which is a disjoint union of triangles covering all but at most C vertices, for some absolute constant C .*

We are not aware of any examples that would rule out the possibility that one can take $C = 2$ above. In the other direction, one can prove a relaxed version of the above conjecture with C replaced with $n^{1-\varepsilon}$ for some $\varepsilon > 0$ by using the Rödl nibble (see for example Corollary 3.4(1)). Improving this bound, for example by replacing C with a polylogarithmic term, could be an interesting challenge, see [26] for an analogous result in the setting of Ryser’s conjecture.

7.3. Other cycle types

To prove the FGT conjecture, we only used the $p = 1$ case of Theorem 2.3. Applying Theorem 2.3 with different values of p , we can derive that many other cycle types for orthomorphisms are possible. Suppose for example that G is an abelian group of order n with the Hall–Paige property, $n - 1 = 3k + 4\ell$, and we want to find an orthomorphism fixing the identity and permuting the remaining elements as k many disjoint 3-cycles and ℓ many 4-cycles. Let’s also suppose for simplicity that $k, \ell = \Omega(n)$. Then, we can partition the vertices of \vec{K}_G into a $3k/n$ -random set V_1 and $4\ell/n$ -random set V_2 , and we can partition the colours of \vec{K}_G into $3k/n$ -random set C_1 and $4\ell/n$ -random set C_2 . With positive probability, Theorem 2.3 holds with (V_1, C_1) , $k = 3$ and with (V_2, C_2) , $k = 4$. We can then do a few exchanges between the sets of vertices and colours so that they satisfy the divisibility condition as well as the sum condition $\sum C_1 = \sum C_2 = 0$. Then, by Theorem 2.3 we obtain the desired cycle partition.

We can go further and ask the following question. Suppose that s_1, s_2, \dots, s_j is a sequence of integers where $s_i \geq 2$ and $\sum s_i = n - 1$, and suppose that G is an abelian group with the Hall–Paige property. When is it true that G has an orthomorphism fixing the identity and permuting the remaining elements as cycles of lengths s_1, s_2, \dots, s_j ? Note that a necessary condition for the existence of such an orthomorphism is a partition of $G \setminus \{0\}$ into zero-sum sets of size s_1, s_2, \dots, s_j (recall Observation 2.1). Characterising pairs of sequences s_1, s_2, \dots, s_j and abelian groups that admit such a partition is known as Tannenbaum’s problem. This problem was solved for large groups in [30]. Perhaps the methods from the current paper could be sufficient to solve the more general problem of characterising which cycle types are feasible for orthomorphisms.

7.4. Other equations

As discussed in Section 2, there is a connection between the Hall–Paige conjecture, the FGT conjecture, and toroidal version of the n -queens problem [7]. We can make this connection more formal as follows. Suppose A is a $\ell \times m$ matrix with integer entries, and G is an abelian group of order n . Can we find a collection of n -many vectors \vec{v} in G^m with $A \cdot \vec{v} = \vec{0}_\ell$ (meaning the ℓ -dimensional 0-vector) such that for each $i \in \{1, 2, \dots, m\}$, the collection of i th coordinates of the vectors \vec{v} is equal to G (i.e., contains no repetitions). If this is possible, let us call the pair (A, G) **matchable**. This term is motivated by the fact that we can equivalently phrase this as a hypergraph matching problem in m -partite m -uniform hypergraphs where the edge set is governed by a collection of ℓ linear equations given by the matrix A .

For example, in the Hall–Paige conjecture, the corresponding matrix A is $[[1, -1, -1]]$, in the $k = 3$ case of the FGT conjecture, the matrix is $[[1, -1, 0, -1, 0, 0], [0, 1, -1, 0, -1, 0], [-1, 0, 1, 0, 0, -1]]$, and in the n -queens problem, the matrix is $[[1, 1, -1, 0], [1, -1, 0, -1]]$. Characterising integer matrices A and abelian groups G such that (A, G) is matchable is a natural unifying problem. This would be interesting already when A consists only of $\{-1, 0, 1\}$ -entries.

7.5. Controlling the cycle type of both bijections

It is also natural to investigate the existence of orthomorphisms/complete mappings ϕ where one makes a restriction on the cycle type of ϕ as well as the cycle type of the permutation $g \rightarrow g^{-1}\phi(g)$. Several

partial results as well as open problems in this direction are given in [5, 6] by Bors and Wang. It would be interesting to see if our methods can be adapted to address this more restrictive variant of the problem.

Acknowledgments. The author thanks Alexey Pokrovskiy for providing feedback on an early version of this manuscript. The author also thanks an anonymous referee for several corrections and very useful suggestions.

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

References

- [1] M. Ajtai, J. Komlós and E. Szemerédi, ‘Sorting in $c \log n$ parallel steps’, *Combinatorica* **3**(1) (Jan 1983), 1–19.
- [2] B. Alspach, D. L. Kreher and A. Pastine, ‘The Friedlander–Gordon–Miller conjecture is true’, *Australas. J Comb.* **67** (2017), 11–24.
- [3] B. Alspach and G. Liversidge, ‘On strongly sequenceable abelian groups’, *Art Discrete Appl. Math.*, **3**(1) (2020), 1–19.
- [4] K. E. Batchler, ‘Sorting networks and their applications’, in *Proceedings of the April 30–May 2, 1968, Spring Joint Computer Conference*, AFIPS ‘68 (Spring), page 307–314, New York, NY, USA, 1968. Association for Computing Machinery.
- [5] A. Bors and Q. Wang, ‘Coset-wise affine functions and cycle types of complete mappings’, *Finite Fields and Their Applications*, **83** (2022), 102088.
- [6] A. Bors and Q. Wang, ‘Cycle types of complete mappings of finite fields’, *J Algebra*, **591** (2022), 577–610.
- [7] C. Bowtell and P. Keevash, ‘The n -queens problem’, arXiv preprint, [arXiv:2109.08083](https://arxiv.org/abs/2109.08083), 2021.
- [8] J. N. Bray, Q. Cai, P. J. Cameron, P. Spiga, and H. Zhang, ‘The Hall–Paige conjecture, and synchronization for affine and diagonal groups’, *J. Algebra*, **545** (2020), 27–42.
- [9] S. Costa, S. Della Fiore, and M. Ollis, ‘Sequencings in semidirect products via the polynomial method’, arXiv preprint, [arXiv:2301.09367](https://arxiv.org/abs/2301.09367), 2023.
- [10] S. Costa, S. Della Fiore, M. Ollis, and S. Z. Rovner-Frydman, ‘On sequences in cyclic groups with distinct partial sums’, arXiv preprint, [arXiv:2203.16658](https://arxiv.org/abs/2203.16658), 2022.
- [11] S. Costa and M. A. Pellegrini, ‘Some new results about a conjecture by Brian Alspach’, *Archiv der Mathematik*, **115**(5) (2020), 479–488.
- [12] S. Eberhard, F. Manners, and R. Mrazović, ‘An asymptotic for the Hall–Paige conjecture’, *Adv. Math.* **404** (2022), 108423.
- [13] S. Ehard, S. Glock, and F. Joos, ‘Pseudorandom hypergraph matchings’, *Combinatorics, Probability and Computing*, **29**(6) (2020), 868–885.
- [14] A. Evans, ‘The admissibility of sporadic simple groups’, *J. Algebra*, **321**(1) (2009), 105–116.
- [15] A. B. Evans, *Orthogonal Latin Squares Based on Groups*, volume 57 (Springer, Cham, 2018).
- [16] R. Friedlander, B. Gordon, and P. Tannenbaum, ‘Partitions of groups and complete mappings’, *Pacific J Math.*, **92**(2) (1981), 283–293.
- [17] R. J. Friedlander, B. Gordon, and M. D. Miller, ‘On a group sequencing problem of Ringel’, *Congr. Numer.*, **21** (1978), 307–321.
- [18] J. Gallian, ‘A Dynamic Survey of Graph Labeling’, *The Electronic Journal of Combinatorics*, (2023), 1–712. <https://www.combinatorics.org/files/Surveys/ds6/ds6v27-2024.pdf>.
- [19] S. Gould and T. Kelly, ‘Hamilton transversals in random Latin squares’, arXiv preprint [arXiv:2104.12718](https://arxiv.org/abs/2104.12718), 2021.
- [20] R. Graham, ‘On sums of integers taken from a fixed sequence’, in *Proceedings, Washington State University Conference on Number Theory*, pages **22–40**, 1971.
- [21] R. L. Graham and N. J. A. Sloane, ‘On additive bases and harmonious graphs’, *SIAM Journal on Algebraic Discrete Methods*, **1**(4) (1980), 382–404.
- [22] M. Hall and L. Paige, ‘Complete mappings of finite groups’, *Pacific J. Math.*, **5** (1955), 541–549.
- [23] J. Haviland and A. Thomason, ‘On testing the ‘pseudo-randomness’ of a hypergraph’, *Discrete Math.*, **103**(3) (1992), 321–327.
- [24] J. Hicks, M. Ollis, and J. R. Schmitt, ‘Distinct partial sums in cyclic groups: polynomial method and constructive approaches’, *Journal of Combinatorial Designs*, **27**(6) (2019), 369–385.
- [25] E. C. Johnsen and T. Storer, ‘Combinatorial structures in loops I. elements of the decomposition theory’, *Journal of Combinatorial Theory, Series A*, **14**(2) (1973), 149–166.
- [26] P. Keevash, A. Pokrovskiy, B. Sudakov, and L. Yepremyan, ‘New bounds for Ryser’s conjecture and related problems’, *Transactions of the American Mathematical Society, Series B*, **9**(08) (2022), 288–321.
- [27] M. Kwan, A. Sah, M. Sawhney, and M. Simkin, ‘High-girth steiner triple systems’, arXiv preprint [arXiv:2201.04554](https://arxiv.org/abs/2201.04554), 2022.
- [28] R. Montgomery, ‘Spanning trees in random graphs’, *Adv. Math.*, **356** (2019).
- [29] R. Montgomery, A. Pokrovskiy, and B. Sudakov, ‘A proof of Ringel’s conjecture’, *Geometric and Functional Analysis*, **31** (2021).
- [30] A. Müyesser and A. Pokrovskiy, ‘A random Hall–Paige conjecture’, arXiv preprint [arXiv:2204.09666](https://arxiv.org/abs/2204.09666), 2022.
- [31] A. Nixon and S. Prendiville, *Surveys in Combinatorics 2022*, volume 481 (Cambridge University Press, 2022).
- [32] M. Ollis, ‘Sequenceable groups and related topics’, *The Electronic Journal of Combinatorics*, **1000**:DS10–Aug, 2002.

- [33] G. Ringel, 'Cyclic arrangements of the elements of a group', *Notices of the American Mathematical Society*, **21**(1) (1974), A–95.
- [34] G. Ringel, *Map Color Theorem*, volume 209 (Springer Science & Business Media, 2012).
- [35] C. Wang, 'On harmoniousness and complete mappings decomposable into disjoint cycles of the same length', *Combinatorics, Graph Theory, Algorithms and Applications (Beijing, 1993)*, pages 347–353, 1994.
- [36] S. Wilcox, 'Reduction of the Hall-Paige conjecture to sporadic simple groups', *J. Algebra*, **321**(5) (2009), 1407–1428.