Inventiones mathematicae (2025) 240:779–867
https://doi.org/10.1007/s00222-025-01328-x

A random Hall-Paige conjecture

Alp Müyesser[1] · Alexey Pokrovskiy[2]

Received: 14 March 2023 / Accepted: 20 February 2025 / Published online: 5 March 2025

Abstract

A complete mapping of a group G is a bijection $\phi\colon G \to G$ such that $x \mapsto x\phi(x)$ is also bijective. Hall and Paige conjectured in 1955 that a finite group G has a complete mapping whenever $\prod_{x\in G} x$ is the identity in the abelianization of G. This was confirmed in 2009 by Wilcox, Evans, and Bray with a proof using the classification of finite simple groups. In this paper, we give a combinatorial proof of a far-reaching generalisation of the Hall-Paige conjecture for large groups. We show that for random-like and equal-sized subsets A, B, C of a group G, there exists a bijection $\phi\colon A \to B$ such that $x \mapsto x\phi(x)$ is a bijection from A to C whenever $\prod_{a\in A} a \prod_{b\in B} b = \prod_{c\in C} c$ in the abelianization of G. We use this statement as a black-box to settle the following old problems in combinatorial group theory for large groups. (1) We characterise sequenceable groups, that is, groups which admit a permutation π of their elements such that the partial products $\pi_1, \pi_1\pi_2, \pi_1\pi_2\cdots\pi_n$ are all distinct. This resolves a problem of Gordon from 1961 and confirms conjectures made by several authors, including Keedwell's 1981 conjecture that all large non-abelian groups are sequenceable. We also characterise the related R-sequenceable groups, addressing a problem of Ringel from 1974. (2) We confirm in a strong form a conjecture of Snevily from 1999 by characterising large subsquares of multiplication tables of finite groups that admit transversals. Previously, this characterisation was known only for abelian groups of odd order (by a combination of papers by Alon and Dasgupta-Károlyi-Serra-Szegedy and Arsovski). (3) We characterise the abelian groups that can be partitioned into zero-sum sets of specific sizes, solving a problem of Tannenbaum from 1981. This also confirms a recent conjecture of Cichacz. (4) We characterise harmonious groups, that is, groups with an ordering in which the product of each consecutive pair of elements is distinct, solving a problem of Evans from 2015.

✉ A. Pokrovskiy
dralexeypokrovskiy@gmail.com

A. Müyesser
alp.muyesser@new.ox.ac.uk

[1] New College, University of Oxford Oxford, UK

[2] University College London, London, UK

1 Introduction

This paper is about finding large-scale structures in finite groups using techniques from probabilistic combinatorics. The prototypical example of a "large-scale structure" in a group is a complete mapping, or equivalently, a transversal in the Latin square corresponding to the multiplication table of the group. We now define each of these terms. A **complete mapping** of a group G is a bijection $\phi\colon G \to G$ such that the function $x \mapsto x\phi(x)$ is also a bijection. A **Latin array** is an $n \times n$ array filled in with arbitrary symbols such that each symbol appears at most once in each row and column. A **Latin square** is an $n \times n$ Latin array with exactly n symbols. A **transversal** of a Latin array is a collection of n cells which do not share a row, a column, or a symbol. Denote by $M(G)$ the **multiplication table** of the group G whose rows and columns are labelled by the elements of the group, and the entry in row g_i and column g_j is the group element $g_i \cdot g_j$. Observe that $M(G)$ is a Latin square, and that $M(G)$ has a transversal if and only if G has a complete mapping.

The study of transversals in Latin squares began more than two hundred years ago, when Euler posed a problem equivalent to determining for which n there exists a $n \times n$ Latin square whose entries can be partitioned into transversals. This motivates the study of transversals in multiplication tables because if the multiplication table of a group has a transversal, then it can be partitioned into transversals. To see this, we can translate the columns of a transversal by a non-identity group element to produce another transversal, entirely disjoint from the first. However, some multiplication tables do not contain any transversals at all, let alone partitions into transversals. Indeed, if we suppose that the multiplication table of a group has a transversal, equivalently, we know that the group has a complete mapping ϕ. Denote by π the permutation $x \mapsto x\phi(x)$, so we have that $\pi(x) = x\phi(x)$ for every $x \in G$. Multiplying these equations[1] together for each $x \in G$, we obtain the following.

$$\prod_{x\in G} \pi(x) = \prod_{x\in G} x\phi(x) \tag{1}$$

For abelian groups, (1) rearranges into $\prod_{x\in G} x = e$ (where e is the identity element of G), giving a necessary condition for having a complete mapping in abelian groups. An immediate consequence is that even-order cyclic groups do not admit complete mappings (for example in $G = \mathbb{Z}_{2n}$ we have $\prod_{x\in G} x = n \neq e$). For non-abelian groups, by taking the image of (1) in the abelianization of G, we obtain that $\prod_{x\in G} x \in G'$, where G' denotes the commutator subgroup of G. Thus "$\prod_{x\in G} x \in G'$" is a necessary condition for the existence of a complete mapping in a general group. Letting G^{ab} denote the abelianization of G, that is $G^{\text{ab}} = G/G'$, we see that this condition is equivalent to $\prod_{x\in G} \pi_{\text{ab}}(x)$ being equal to the identity in G^{ab} (where $\pi_{\text{ab}} : G \to G^{\text{ab}}$ is the quotient homomorphism). Since G^{ab} is abelian, we write this as "$\sum_{x\in G} x = 0$ in G^{ab}".

The condition that $\sum_{x\in G} x = 0$ in G^{ab} (or equivalently that $\prod_{x\in G} x \in G'$) is known as the Hall-Paige condition [38]. We remark that the Hall-Paige condition

[1] One can fix an arbitrary ordering of G for the product to be well-defined for non-abelian groups. All the statements about products like this that we write in the introduction are independent of the ordering picked.

is sometimes written as "all 2-Sylow subgroups of G are trivial or non-cyclic". This is equivalent to $\prod_{x\in G} x$ being trivial in G^{ab}, as shown by Hall and Paige themselves [38]. Perhaps astonishingly, the Hall-Paige condition is not only necessary, but also sufficient for the existence of a complete mapping. This was first conjectured by Hall and Paige [38]. The Hall-Paige Conjecture has a rich history, we refer the reader to the book of Evans [26] (Chaps. 3-7) for a description of various approaches taken for this problem. The conjecture was finally shown to be true in a combination of papers by Wilcox [62], Evans [23], and Bray [14] in 2009. The original proof of Wilcox, Evans, and Bray uses an inductive argument which relies on the classification of finite simple groups. However, recently, a completely different proof for large groups was found by Eberhard, Manners, and Mrazović using tools from analytic number theory [20].

In this paper, with the goal of giving a unified approach to many related conjectures in the area, we study complete mappings between subsets of groups. For example, given equal sized subsets A, B, C of a group G, is there a bijection $\phi\colon A \to B$ such that the map $\psi(a) := a\phi(a)$ defines a bijection $A \to C$? This corresponds to starting with the multiplication table of a group, and then deleting the rows corresponding to $G \setminus A$, columns corresponding to $G \setminus B$, and symbols corresponding to $G \setminus C$, and then searching for a transversal in the resulting structure, which is simply a Latin array with some missing entries. Generalising the Hall-Paige condition to this set-up, we see that

$$\sum A + \sum B = \sum C \text{ (in } G^{\mathrm{ab}}) \tag{2}$$

is a necessary condition for the existence of such a map ϕ (we use $\sum S$ to denote $\sum_{s\in S} s$, $\prod S$ is defined analogously). Of course, we cannot expect (2) to be a sufficient condition for any triple of equal sized subsets.[2] However, our main theorem essentially states that for most subsets $A, B, C \subseteq G$, (2) is the only obstruction for finding the desired map ϕ. Recall that a p-**random subset** X of a finite set G is a subset sampled by including each element of G in X independently with probability p, and $\triangle$ denotes symmetric difference. When we say that an event holds with high probability, we mean that the probability of that event approaches 1 as n tends to infinity. The letter n throughout the paper always denotes the order of the ambient group G.

Theorem 1.1 (Main result) *Let G be a group of order n. Let $p \geq n^{-1/10^{105}}$. Let $R^1, R^2, R^3 \subseteq G$ be p-random subsets, sampled independently. Then, with high probability, the following holds.*

Let $X, Y, Z \subseteq G$ be equal sized subsets satisfying the following properties.

- $|X \triangle R^1| + |Y \triangle R^2| + |Z \triangle R^3| \leq p^{10^{30}} n/\log(n)^{10^{30}}$
- $\sum X + \sum Y = \sum Z$ *in* G^{ab} *(or equivalently* $\prod X \prod Y (\prod Z)^{-1} \in G'$*)*

Then, there exists a bijection $\phi\colon X \to Y$ such that $x \mapsto x\phi(x)$ is a bijection from X to Z.

[2] For example, consider $G = \mathbb{Z}_{99}$, $A = \{98, 1\}$, $B = \{98, 1\}$, $C = \{49, 50\}$.

We remark that setting $p = 1$ and $X = Y = Z = G$, we see that the Hall-Paige conjecture holds for sufficiently large groups. However, Theorem 1.1 extends far beyond the setting of the Hall-Paige conjecture. Indeed, we can settle several longstanding conjectures at the interface of group theory and combinatorics using the full strength of Theorem 1.1. In the following section, we discuss several such problems. All of these problems are similar in spirit to the Hall-Paige conjecture in the sense that they concern finding large-scale structures in groups with certain desirable properties. However, all of these problems lack the level of symmetry present in the Hall-Paige problem. For example, some of these problems concern finding complete mappings in subsets of groups with limited structure, or they concern finding complete mappings permuting the group elements via a particular cycle type. Both of these constraints seem difficult to reason about using only algebraic techniques.

This perhaps explains why many of the commonly used tools, such as Alon's combinatorial Nullstellensatz, were only able to go so far in addressing these questions. On the other hand, one clearly needs use some of the group theoretic structure. Indeed, the key obstruction for problems of this type ends up being some derivative of the Hall-Paige condition, which is inherently group theoretic. As surveyed by Gowers in [33], there are many problems in combinatorics which are difficult for a similar reason. These are problems where "there is too much choice for constructions to be easy to discover, and too little choice for simple probabilistic arguments to work" [33]. Our proof, similar in spirit to Keevash's celebrated construction of designs [41], uses probabilistic tools but also exploits the algebraic structure of the problem. We give a detailed overview of our strategy in Sect. 2.

From now on, additive notation will always imply that the corresponding operation is taking place in the abelianization of the ambient group (e.g. when for $A \subseteq G$ we write "$\sum A = e$" we mean that $\prod_{a\in A} \pi_{\text{ab}}(a) = e$ where $\pi_{\text{ab}} : G \to G^{\text{ab}}$ is the quotient map to the abelianization of G). Otherwise, operations within non-abelian groups will be denoted multiplicatively. The quantity n always denotes the size of the ambient group. We use e to denote the identity element of a group, and we sometimes use 0 to denote e in the special case of abelian groups.

1.1 Applications

As we already noted, the first application of our main result, Theorem 1.1, is an alternative proof that the Hall-Paige conjecture holds for all sufficiently large groups. This uses only the $p = 1$ case of Theorem 1.1, where we set the subsets X, Y, Z to be the entirety of the group G. We now mention another result we can recover easily, this time setting X, Y, Z to be sets of size $|G| - 1$. Goddyn and Halasz recently proved that multiplication tables contain near transversals, that is, a collection of $n - 1$ cells which do not share a row, column, or a symbol [31]. Perhaps surprisingly, there does not seem to be an easy way of deriving this from the Hall-Paige conjecture directly, and the proof in [31] is somewhat involved. However, we can derive this result from Theorem 1.1 as follows.

Proposition 1.2 *Let M be the multiplication table of a sufficiently large finite group G. Then, M contains a near transversal.*

Proof Applying Theorem 1.1 with $p = 1$, we derive that for $|G|$ large enough, $R^1 = R^2 = R^3$ satisfies the conclusion of Theorem 1.1 with positive probability, and therefore, with probability exactly 1. Let $z \in G$ be some element equal to $\sum G$ in G^{ab}. Setting $X = Y = G \setminus \{e\}$ and $Z = G \setminus \{z\}$, we have that $\sum X + \sum Y = \sum Z$ in G^{ab}. This implies that there is a bijection $\phi\colon X \to Y$ such that $x \to x\phi(x)$ is injective. ϕ then corresponds to a near transversal in M, as desired. □

The $p = 1$ case of Theorem 1.1 when applied with other subsets X, Y, Z has novel applications as well. We discuss such an application in the next section. The other three applications we discuss use the full strength of Theorem 1.1. In fact, for these applications, we rely on an appropriate generalisation of Theorem 1.1 with a more complicated distribution on the sets R^1, R^2, R^3. In Sect. 4, we state this generalised version of Theorem 1.1.

1.1.1 Snevily's conjecture

Deleting k rows and k columns of a multiplication table, we obtain a natural Latin array which we call a **subsquare** of the multiplication table. In analogy with complete mappings, it is natural to ask which subsquares contain transversals. One may suspect that deleting rows and columns should only make it easier to find transversals, and Snevily's conjecture states that this indeed should be the case for abelian groups of odd order [56]. However, for even order abelian groups G, one may delete rows and columns so that the remaining Latin array is in fact the multiplication table of an even order subgroup $H \subseteq G$ (or a translate of such a multiplication table). Such subsquares cannot contain transversals as multiplication tables of even order cyclic groups do no admit transversals. In 1999, Snevily conjectured that this should be the only subtlety for cyclic even order groups. Below, we formally state both cases of Snevily's conjecture. Note that $A \times B$ denotes the subsquare of a group G obtained by keeping only the rows corresponding to A and columns corresponding to B in the multiplication table of G.

Conjecture 1.3 (Snevily, [56]) *Let $S = A \times B$ be a subsquare of the multiplication table of an abelian group G defined by two n-element sets $A, B \subseteq G$.*

1. *If $|G|$ is odd, then S has a transversal.*
2. *If $G \cong \mathbb{Z}_{2k}$ for some $k \in \mathbb{N}$, then S has a transversal unless there exists $g_1, g_2 \in G$ such that $g_1 + A = g_2 + B = H$ for some even order cyclic subgroup $H \subseteq G$.*

The first case of the conjecture was verified by Alon in 2000 for prime order cyclic groups [4]. In 2001, Dasgupta, Károlyi, Serra, and Szegedy generalised Alon's result to arbitrary odd order cyclic groups [19]. Both of these results use the celebrated Combinatorial Nullstellensatz [3]. A decade later, Arsovski fully resolved the first case of Snevily's conjecture, using character theory [7]. On the other hand, as far as the authors are aware, no partial progress has been reported on the second case of the conjecture.

As observed by Wanless [61], the second part of Snevily's conjecture does not generalise straightforwardly to all even abelian groups due to the following construction of Akbari and Alireza [2]. Let $G = (\mathbb{Z}_2)^k$ for some $k \geq 1$, and let $a_1, a_2 \in G$ be distinct and let $b_1, b_2 \in G$ be distinct such that $a_1 + a_2 + b_1 + b_2 = 0$. Then, setting $A = G \setminus \{a_1, a_2\}$, $B = G \setminus \{b_1, b_2\}$, it is a simple exercise to check that the subsquare $A \times B$ does not contain a transversal. We show, perhaps surprisingly, that this is the only other barrier for an abelian subsquare to contain a transversal.

Theorem 1.4 *There exists an $n_0 \in \mathbb{N}$ such that the following holds for all $n \geq n_0$. Let G be an abelian group, and let $A, B \subseteq G$ with $|A| = |B| = n$. Then, $A \times B$ has a transversal, unless there exists some $k \geq 1$, $g_1, g_2 \in G$ and a subgroup $H \subseteq G$ such that one of the following holds.*

1. *$H \cong \mathbb{Z}_{2k} \times H_{odd}$ for some odd-order group H_{odd}, and $H = g_1 A = g_2 B$, i.e. A and B are cosets of H.*
2. *$H \cong (\mathbb{Z}_2)^k$, $g_1 A = H \setminus \{a_1, a_2\}$, $g_2 B = H \setminus \{b_1, b_2\}$ for some distinct $a_1, a_2 \in H$ and distinct $b_1, b_2 \in H$ such that $a_1 + a_2 + b_1 + b_2 = 0$.*

Note that this confirms both cases of Snevily's conjecture for sufficiently large subsquares. We discuss proving a stronger characterisation valid for all n in Sect. 7.

We take Snevily's conjecture further by proving a far more general theorem characterising subsquares without transversals of all groups.

Theorem 1.5 *There exists an $n_0 \in \mathbb{N}$ such that the following holds for all $n \geq n_0$. Let G be a group, and let $A, B \subseteq G$ with $|A| = |B| = n$. Then, $A \times B$ has a transversal, unless there exists some $k \geq 1$, $g_1, g_2 \in G$ and a subgroup $H \subseteq G$ such that one of the following holds.*

1. *H is a group that does not satisfy the Hall-Paige condition, and $A = g_1 H$ and $B = H g_2$.*
2. *$H \cong (\mathbb{Z}_2)^k$, $g_1 A = H \setminus \{a_1, a_2\}$, $g_2 B = H \setminus \{b_1, b_2\}$ for some distinct $a_1, a_2 \in H$ and distinct $b_1, b_2 \in H$ such that $a_1 + a_2 + b_1 + b_2 = 0$.*

Roughly speaking, Theorem 1.5 states that deleting rows and columns from a multiplication table only makes it easier to find transversals (supposing we do not end up with a translate of a multiplication table of a subgroup), except for a very specific scenario where we delete 2 rows and 2 columns summing to zero from the multiplication table of an elementary abelian 2-group. Theorem 1.5 is proved in Sect. 6.1.

1.1.2 Sequenceable and R-sequenceable groups

Given a finite group G, a **sequencing** is an ordering of the elements of G as $b_1, b_2, \ldots, b_n$ where the partial products b_1, $b_1 b_2, \ldots, b_1 b_2 \cdots b_n$ are all distinct. Observe that in a sequencing, $b_1 = e$. A group that admits a sequencing is called **sequenceable**. A similar notion is that of an R-sequencing. An **R-sequencing** is an ordering of the elements of G as $b_1, b_2, \ldots, b_n$ where $b_1 = e$, the partial products $b_1, b_1 b_2, \ldots, b_1 b_2 \cdots b_{n-1}$ are all distinct and $b_1 b_2 \cdots b_n = e$. A group that admits an R-sequencing is called **R-sequenceable**. We will briefly discuss the rich history of the problems relating to these concepts, and we refer the reader to [50] and [26] for a more comprehensive survey.

Gordon introduced the problem of determining which groups are sequenceable in 1961 [32]. His motivation was to construct complete Latin squares, a concept which we now define. A Latin square is called **row-complete** if every pair of distinct symbols appears exactly once in each order in adjacent horizontal cells. A Latin square is called **column-complete** if it has the same property with respect to adjacent vertical cells. A Latin square is **complete** if it is both row-complete and column-complete. Complete Latin squares possess an additional layer of symmetry, making them useful in various contexts. For example, complete Latin squares are useful in graph theory to give decompositions of complete directed graphs into Hamilton paths (see [50] and the references therein). Also, some applications to experimental design are given in [9]. Gordon was motivated by the observation that given a sequenceable group G, we can construct a complete Latin square by considering the multiplication table of the group G.

Ringel had a completely different motivation for studying R-sequenceable groups. Such groups come up naturally in Ringel's celebrated proof of the Heawood map colouring conjecture [52]. Hence, Ringel asked for a classification of all such groups [51].

In Sect. 6.2, we solve both of these problems for large groups. This addresses problems reiterated by several authors [26, 50], and confirms a conjecture of Keedwell [40] (see also Conjecture 7 in [31]). In particular, we show that any large group with the Hall-Paige condition is R-sequenceable and we show that any large non-abelian group is sequenceable which may be surprising in view of the fact that, for example, the nonabelian groups of order 6 and 8 are not sequenceable [32]. Therefore, at least for this problem, some mild assumption on the size of the group is necessary for a clean characterisation. We also remark that several partial results towards this characterisation were obtained by other researchers, see [50] for a survey.

1.1.3 Partitioning Abelian groups into zero-sum sets

We call a subset S of a group **zero-sum** if $\sum S = 0$. Given a sequence $a_1, a_2, \dots, a_k$ ($a_i \geq 2$) with $\sum a_i = n - 1$, where $n = |G|$, when can we partition the non-identity elements of an abelian group into zero-sum sets of size $a_1, a_2, \dots, a_k$? A variant of this natural problem seems to have been first considered in 1957 by Skolem [55], and in 1960 by Hanani [39]. A complete solution for cyclic groups for sequences $a_1, a_2, \dots, a_k$ with $a_i \geq 3$ was given by Friedlander, Gordon, and Tannenbaum in 1981 [28].

Obviously, we need that $\sum G = 0$. This already rules out even-order cyclic groups, for example. There is another very natural necessary condition. Let ℓ denote the number of a_i such that $a_i = 2$. We need that G should contain at least ℓ-many pairs $\{x, -x\}$ where $x \neq -x$ (i.e. at least 2ℓ-elements of order greater than 2). It turns out that for odd order abelian groups, these two conditions are known to be sufficient as well as necessary [57, 63]. However, for even order abelian groups, Tannenbaum observed that additional necessary conditions are required [58]. To see this, we invite the reader to consider the case when $G = \mathbb{Z}_4 \times \mathbb{Z}_2 \times \mathbb{Z}_2$, and the sequence $a_1, a_2, \dots, a_k$ is $(2, 2, 2, 3, 3, 3)$. See [58] for a solution to why G does not have the desired partition in this case.

Motivated by the previous example, Tannenbaum asked to find a set of necessary and sufficient conditions guaranteeing a partition of an even order abelian group into zero-sum sets of prescribed sizes [58]. Despite the apparent lack of structure in the problem as evidenced by a rich family of counterexamples, in Sect. 6.3, we give a complete characterisation of the (large) abelian groups and integer sequences for which the desired partition exists. Various special cases of this problem were investigated by several authors, for example see [17] and all the references therein. In particular, our characterisation confirms the following conjecture of Cichacz [16].

Conjecture 1.6 (Cichacz, [16]) *Let G be an abelian group with at least 3 involutions and suppose we have numbers $r_1, \dots, r_t \geq 3$ with $r_1 + \cdots + r_t = |G| - 1$. Then there is a partition of $G \setminus e = Z_1 \cup \cdots \cup Z_t$ where each Z_i is a zero-sum set of size r_i.*

1.1.4 Harmonious groups

Given a group G, not necessarily abelian, a **harmonious ordering** is an ordering of the elements of G as $a_1, a_2, \dots, a_n$ such that $a_1a_2, a_2a_3, \dots, a_{n-1}a_n, a_na_1$ is also an ordering of the elements of G (i.e. the latter sequence contains no repetitions). Groups which have harmonious orderings are called **harmonious**. Harmonious groups were first introduced in 1991 by Beals, Gallian, Headley, and Jungreis [10]. They gave a characterisation of all abelian harmonious groups. Evans asked for a complete characterisation.

Problem 1.7 (Evans, [25]) Which finite groups are harmonious?

In Sect. 6.2, we give a characterisation of all sufficiently large harmonious groups. It turns out that a (large) group is harmonious if and only if it satisfies the Hall-Paige condition, and is not an elementary abelian 2-group.

1.2 Organisation of the paper

In Sect. 2, we give a extensive proof sketch. Section 3 collects some standard concentration/nibble type results, makes precise some key definitions and notation, and records some group theoretic results we rely on. Section 3.6 in particular is quite central, and is devoted to giving a sufficient set of conditions allowing us to find various gadgets throughout the paper. In Sect. 4, we state a more general version of Theorem 1.1, and we derive several variants including Theorem 1.1 itself. In Sect. 5 the generalised version of the main theorem is proved. Section 6 is devoted to deriving the applications we have listed in Sect. 1.1. In Sect. 7, we discuss further avenues of research.

2 Proof strategy for the main result

In this section, we attempt to give an accessible outline of a special case of our main result, concerning cyclic groups. We conclude in Sect. 2.1.4 by outlining some key difficulties we omit in the simplified discussion.

Firstly, instead of using the language of complete mappings, we will re-frame Theorem 1.1 as a hypergraph matching problem. Given a group G, we define the

3-uniform 3-partite multiplication hypergraph H_G as follows. Set $V(H_G) := G_A \sqcup G_B \sqcup G_C$ where G_* is a copy of G and $\sqcup$ indicates a disjoint union. Set $E(G) := \{(g_A, h_B, k_C) \in G_A \times G_B \times G_C : g_A h_B k_C = e\}$. Given $X, Y, Z \subseteq G$, we denote by $H_G[X, Y, Z]$ the induced subgraph of H_G given by the vertex subset $(G_A \cap X) \sqcup (G_B \cap Y) \sqcup (G_C \cap Z)$. Given equal sized subsets X, Y, Z, observe that finding a bijection $\phi : X \to Y$ such that $x \mapsto x\phi(x)$ is a bijection from X to Z is equivalent to finding a perfect matching in $H_G[X, Y, Z^{-1}]$ where $Z^{-1} = \{z^{-1} : z \in Z\}$. We will outline a proof for the following result. For the rest of the outline, fix $\varepsilon = 1/100$.

Proposition 2.1 *Let $X, Y, Z \subseteq \mathbb{Z}_n$ such that $|X| = |Y| = |Z| = n - O(n^{1-\varepsilon})$, and $\sum X + \sum Y + \sum Z = 0$. Then, $H_{\mathbb{Z}_n}[X, Y, Z]$ has a perfect matching.*

Proposition 2.1 is already novel, and would be sufficient, for example, to deduce Snevily's conjecture for large subsquares. We remark that Proposition 2.1 is a simple corollary of Theorem 1.1 which can be obtained by setting $p = 1$. We remark that throughout the rest of the paper, when we say that a subset $S \subseteq G$ is **zero-sum**, we mean that the product of all elements of S (in any order) is in G', the commutator subgroup. For abelian groups, this corresponds to $\sum S = 0$.

2.1 Absorption

Our main tool is the *absorption method*, a technique codified by Rödl, Ruciński, Szemerédi [53] (see also the earlier work of Erdős, Gyárfás, and Pyber [22]), adapted to the setting of hypergraphs defined by groups. Absorption is a general method that reduces the task of finding spanning structures to finding *almost* spanning structures. In most cases, the latter task is considerably simpler, as evidenced by the celebrated nibble method which roughly states that pseudorandom hypergraphs contain large matchings [5]. The main technical innovation in our paper is developing an absorption strategy for multiplication hypergraphs, which in the setting of Proposition 2.1, culminates in the following lemma.

Lemma 2.2 (Simpler version of Lemma 5.32) *$H_{\mathbb{Z}_n}[X, Y, Z]$ contains a vertex subset $\mathcal{A}$ of size $o(n)$ such that for any $S \subseteq V(H_{\mathbb{Z}_n}[X, Y, Z]) \setminus \mathcal{A}$ of size $O(n^{1-\varepsilon})$ intersecting X, Y and Z in the same number of vertices, and satisfying $\sum S = 0$, $\mathcal{A} \cup S$ has a perfect matching.*

In Lemma 2.2, $\mathcal{A}$ functions as our *absorber*, in the sense that it can *absorb* small enough subsets by forming perfect matchings when combined with them. The key premise of the absorption method is that once a suitable absorber is found and set aside, the only remaining task is to find an *almost perfect matching* in the leftover set. Indeed, $\mathcal{A}$ can absorb whatever small subset S we fail to cover with the almost perfect matching.

We now explain in a bit more detail how Lemma 2.2 reduces the task of proving Proposition 2.1 to finding a matching of size $n - O(n^{1-\varepsilon})$ in $V(H_{\mathbb{Z}_n}[X, Y, Z]) \setminus \mathcal{A}$. First, note that setting $S = \emptyset$ in Lemma 2.2 implies that $\mathcal{A}$ has a perfect matching, and thus $\sum \mathcal{A} = 0$ (if a subset contains a perfect matching, by definition the subset can be partitioned into zero-sum sets, and hence is zero-sum itself). Now, suppose that having fixed the set $\mathcal{A}$, we were able to find a matching M_1 in $V(H_{\mathbb{Z}_n}[X,Y,Z])\setminus\mathcal{A}$ covering all but $O(n^{1-\varepsilon})$ vertices. Let S denote the set of these leftover vertices. As $\mathcal{A}$ and $V(M_1)$ are disjoint zero-sum sets contained in $V(H_{\mathbb{Z}_n}[X,Y,Z])$, and $\sum X+\sum Y+\sum Z=0$ by assumption, we have that $\sum S=0$ as well. So by the property in Lemma 2.2, $\mathcal{A}\cup S$ spans another perfect matching, M_2 say. Then, $M_1\cup M_2$ is the desired perfect matching of $H_{\mathbb{Z}_n}[X,Y,Z]$.

We remark that, in reality, deleting $\mathcal{A}$ from $H_{\mathbb{Z}_n}[X,Y,Z]$ would damage the pseudorandomness properties of the hypergraph too greatly to be able to find the desired M_1 using the Rödl nibble [5]. Therefore, here we actually need a slightly stronger version of Lemma 2.2 which can find $\mathcal{A}$ inside small random sets. This way, deleting $\mathcal{A}$ only spoils the pseudorandomness of a set R much smaller than the multiplication hypergraph itself. $R\setminus\mathcal{A}$ can then be dealt with using standard pseudorandomness arguments, see for example Lemma 3.8.

For Lemma 2.2 to hold, we remark that some condition on the value of $\sum S$ is necessary. Indeed, recall that if a subset admits a perfect matching, it has to be zero-sum. So, if S_1 and S_2 are two sets disjoint with $\mathcal{A}$ such that $\mathcal{A}\cup S_1$ and $\mathcal{A}\cup S_2$ both admit perfect matchings, it follows that

$$0=\sum\mathcal{A}\cup S_1=\sum\mathcal{A}\cup S_2$$

hence $\sum S_1=\sum S_2$. Therefore, the set $\mathcal{A}$ can have the flexibility of combining with any member of a large family of sets $\mathcal{F}$ to produce perfect matchings only if $\sum S$ is fixed for all $S\in\mathcal{F}$. For convenience, we fix this sum to be 0, but Lemma 2.2 would remain true if we replaced 0 with any other fixed element.

2.1.1 Building the absorber from small subgraphs

Our starting point for building the absorber set $\mathcal{A}$, similar in spirit to most applications of the absorption method, is the existence of small subgraphs (gadgets) which give *local variability*. More precisely, we will rely on the existence of $O(1)$-sized gadgets, Q say, that can combine with 2 distinct sets, F_1 and F_2 say, each of size $O(1)$, such that $Q\cup F_1$ and $Q\cup F_2$ both induce perfect matchings in $H_{\mathbb{Z}_n}[X,Y,Z]$. We say that Q can *switch* between F_1 and F_2.

The power of the absorption method rests in the fact that small gadgets such as Q displaying rather limited variability can be combined in a way to build an absorber displaying *global variability*. By global variability, we mean the type of property that $\mathcal{A}$ has in the statement of Lemma 2.2. In particular, we are referring to how $\mathcal{A}$ can combine with essentially any subset of size $O(n^{1-\varepsilon})$, as opposed to just a few of size $O(1)$. To achieve this in our case, we will use a variant of the absorption technique called *distributive absorption*, initially developed by Montgomery [46]. The method has since been applied in numerous settings, notably in the proof of Ringel's conjecture by Montgomery, Pokrovskiy, and Sudakov [48]. For a detailed discussion of how gadgets such as Q can be combined to build an absorber, we refer the reader to the discussion in [48].

For readers who are familiar with the absorption method, we add a quick remark that commonplace methods of building absorbers, such as those used in [53], do

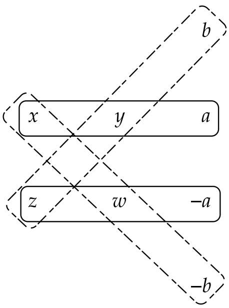

Fig. 1 The gadget Q_x. Matchings M_1 and M_2 depicted in solid and dashed lines, respectively

not work in our context for the following simple reason. Say we have a subset S and we are interested in subsets A of size k such that both A and $A \cup S$ span a perfect matching. In the usual absorption strategy, we would require that the number of subsets A with this property is $\Omega(n^k)$. However, if A spans a perfect matching, then $\sum A = 0$. The number of zero-sum subsets A of order k is $O(n^{k-1})$ – too little to appear in abundance when a positive fraction of k-subsets are randomly sampled. On the other hand, with the distributive absorption strategy, one may build absorbers with fewer gadgets, provided that one can show that the gadgets are well-distributed within the host structure.

2.1.2 Absorption for pairs

Typically, in applications of the distributive absorption method, one works with gadgets Q switching between F_1 and F_2 where F_1 and F_2 are both singletons. This would be impossible to implement in our context as if $Q \cup F_1$ and $Q \cup F_2$ both contain perfect matchings, then $\sum F_1 = \sum F_2$. Thus, if F_1 and F_2 were singletons, F_1 and F_2 would consist of the exact same vertex. Hence, if we want Q to be a gadget that actually gives us some flexibility, we can only hope to switch between sets of size at least 2. This motivates us to search for disjoint sets Q, F_1 and F_2 such that $Q \cup F_1$ and $Q \cup F_2$ both span perfect matchings, $|F_1| = |F_2| = 2$, and $\sum F_1 = \sum F_2 = 0$, where the final equality is chosen for convenience as in the statement of Lemma 2.2.

Finding Q with this property turns out to be rather easy. For example, fix disjoint sets $F_1 := \{a, -a\}$ and $F_2 := \{b, -b\}$ where neither a nor b is equal to its own inverse. Let us view F_1 and F_2 as subsets of G_C. Consider some $x \in G_A$. Set $y := -x - a \in G_B$, set $z = x + a - b \in G_A$ and $w = -x + b \in G_B$. Suppose that $x \neq z$ and $y \neq w$. Observe that $M_1 = \{(x, y, a), (z, t, -a)\}$ is a matching of $\{x, y, z, w\} \cup F_1$ and $M_2 = \{(x, w, -b), (z, y, b)\}$ is a matching of $\{x, y, z, w\} \cup F_2$. Hence, $Q_x := \{x, y, z, w\}$ is a gadget with the desirable property of switching between F_1 and F_2. See Fig. 1 for an illustration. Moreover, it is not hard to see that there are many choices of x for which the corresponding sets Q_x are all disjoint, which is critical for the distributive absorption strategy.

Building on the idea detailed in the previous paragraph, we can show that there exists gadgets like Q_x which can combine with any one of 100 (as opposed to just 2) pairs of inverses to produce a matching. This, combined with the usual distributive absorption strategy, is already sufficient to prove a version of Lemma 2.2 with additional hypotheses on the set S. Namely, one can show the following.

Lemma 2.3 (Simpler version of Lemma 5.29) *Let S_1 be a $o(n)$-sized vertex subset of $H_{\mathbb{Z}_n}[X, Y, Z]$. Then, $H_{\mathbb{Z}_n}[X, Y, Z]$ contains a vertex subset $\mathcal{A}'$ of size $o(n)$ such that for any $S_2 \subseteq S_1$ of size $O(n^{1-\varepsilon})$ intersecting X, Y and Z in the same number of vertices, closed under the function $x \to -x$, and containing no elements of order ≤ 2, we have that $\mathcal{A}' \cup (S_1 \setminus S_2)$ has a perfect matching.*

In the above lemma, it would arguably be more natural to insist that $\mathcal{A}' \cup S_2$ has a perfect matching, as opposed to $\mathcal{A}' \cup (S_1 \setminus S_2)$. Such a version of the lemma would also be correct, and be even easier to prove. We formulate the lemma in the form above for a reason that will become clear shortly.

2.1.3 From absorption for pairs to absorption for arbitrary zero-sum sets

Now, we discuss how we can derive Lemma 2.2 from Lemma 2.3. The key idea is encapsulated in the following lemma.

Lemma 2.4 (Simpler version of Lemma 5.31) *$H_{\mathbb{Z}_n}[X, Y, Z]$ contains a vertex subset $\mathcal{T}$ of size $o(n)$ such that for any $S \subseteq V(H_{\mathbb{Z}_n}[X, Y, Z]) \setminus \mathcal{T}$ of size $O(n^{1-\varepsilon})$ intersecting X, Y and Z in the same number of vertices, and satisfying $\sum S = 0$, there exists a matching M of order $O(n^{1-\varepsilon})$ with $\mathcal{T} \cup S \supseteq V(M) \supseteq S$ satisfying also that $V(M) \setminus S$ is closed under $x \to -x$.*

Deriving Lemma 2.2 from Lemma 2.3 and Lemma 2.4 is a simple exercise. Indeed, let $\mathcal{T}$ be a vertex subset of $H_{\mathbb{Z}_n}[X, Y, Z]$) with the property in Lemma 2.4. Apply Lemma 2.3 with $S_1 = \mathcal{T}$ to obtain a vertex subset $\mathcal{A}'$. Set $\mathcal{A} := \mathcal{A}' \cup \mathcal{T}$. We invite the reader to check that $\mathcal{A}$ then satisfies the property that Lemma 2.2 requires.

To see how we prove a version of Lemma 2.4, we invite the reader to see Sect. 5.3.1. Similar in spirit to the proof of Lemma 2.3, our main trick here is to reduce (a version of) Lemma 2.4 to the existence of many small matchings with certain desirable properties (see Lemma 5.30).

2.1.4 Additional difficulties

We now point out several complications we omitted in the previous discussion, along with some technicalities that arise in the level of generality of our main theorem.

Ensuring distinctness A fair portion of our arguments rely on the existence of constant sized matchings with specific properties, such as being closed under the map $x \to -x$. Often, the properties we require can be written as solutions to a particular system of linear equations. For example, consider Q_x defined in Sect. 2.1.2. The requirements from $\{x, y, z, w\}$ could be written as:

$$x + y + a = 0 \qquad z + w - a = 0 \qquad x + w + b = 0 \qquad z + y - b = 0$$

This is a system of equations with 4 variables and 4 constraints; however, any 3 of these equations imply the fourth, allowing us to easily deduce that there are many x, y, z, w with the desired properties. However, a solution to the system of equations is useful to us only when $x \neq z$ and $y \neq w$, as otherwise $\{(x, -x+b, -b), (x+a-b, -x-a, b)\}$ would not be a matching in $H_{\mathbb{Z}_n}$.

In the specific scenario outlined above, getting the relevant coordinates to be distinct is not particularly challenging. However, throughout the paper, we will require gadgets with properties significantly more complicated than those of Q_x. Furthermore, our main theorem works with subsets $X, Y, Z \subseteq G$ which are disjoint. Thus, we would have no chance of locating Q_x within $H_G[X, Y, Z]$ unless all coordinates x, y, z, w are distinct. Due to these technicalities, finding gadgets with the desired properties can become quite delicate.

Section 3.6 is entirely devoted to obtaining a sufficient set of conditions for a system of relations to yield solutions where each coordinate is distinct. The key result of that section, Lemma 3.32, is used in abundance throughout the paper.

The elementary Abelian 2-group The strategy of working with pairs of inverses $\{x, -x\}$ with $x \neq -x$ fails for obvious reasons in the elementary abelian 2-group. This turns out to be not a serious complication, as for general groups G, we will work with pairs $\{x, q_\phi x^{-1}\}$ for a carefully chosen q_ϕ so that $x \to q_\phi x^{-1}$ does not create too many fixed points.

Nonabelian groups Although it turns out that the case of general abelian groups is not significantly more complicated than cyclic groups, there are serious issues to overcome with nonabelian groups. As just one example, suppose that we wish to find a gadget similar to Q_x from Sect. 2.1.2 in H_G, where G is a non-abelian group. Suppose that $F_1 = \{a, q_\phi a^{-1}\}$, $F_2 = \{b, q_\phi b^{-1}\}$. Our goal is then to find (many) Q such that $Q \cup F_1$ and $Q \cup F_2$ both can be perfectly matched. It is quite instructive to try to construct such Q, and we invite the reader to try to do so.

Some reflection shows that while it is difficult to construct Q switching between $\{a, q_\phi a^{-1}\}$ and $\{b, q_\phi b^{-1}\}$, it is considerably simpler to find *some* b' such that b' is in the same G'-coset as $q_\phi b^{-1}$ such that we can find a Q switching between $\{a, q_\phi a^{-1}\}$ and $\{b, b'\}$. This motivates us to search for gadgets not only switching between pairs, but also switching between elements of the same G'-coset. Achieving this latter task is considerably more technical. We accomplish this by first devising a strategy for switching between commutator elements (as opposed to arbitrary elements in the commutator subgroup). To switch between arbitrary elements of G', we have to rely on a non-trivial fact from representation theory (see Theorem 5.11). The statement we require is that arbitrary elements in the commutator subgroup can be written as products of just $O(\log n)$ commutators. Arguably, this is the only point in the proof of the main theorem where we use a group theoretic result beyond the undergraduate level. Due to bounds coming from Theorem 5.11 (which are tight), we are obliged to use gadgets of logarithmic size to switch between elements of the same G'-coset. This inflates the error rate in our main theorem by a polylogarithmic factor. We refer the reader to Sect. 5.1 for more details.

3 Preliminaries

3.1 Probabilistic tools

3.1.1 Concentration inequalities

The below is a standard bound that can be found in many probability textbooks, for example see [5]. We will refer to it 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 $|pn - |R|| \leq \log n\sqrt{n}$.

In almost all instances, the Chernoff bound will be all we need. Otherwise, we will make use of Azuma's inequality, which we now state. Given a product probability space $\Omega = \prod_{i\in[n]} \Omega_i$, a random variable $X : \Omega \to \mathbb{R}$ is called C-Lipschitz if $|X(\omega) - X(\omega')| \leq C$ whenever ω and ω' differ in at most 1-coordinate. We will refer to the following standard bound as Azuma's inequality.

Lemma 3.2 (Azuma's inequality) *Let X be 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 Pseudorandom graphs and the Rödl nibble

Here, we give some tools to find matchings in pseudorandom hypergraphs covering all but a few vertices. Our approach here is complicated by the fact that we need to find large matchings in subsets of hypergraphs, and some subsets could be arbitrary and we may as well suppose they were chosen adversarially. This comes from the first step of the proof where we set aside an absorber, whose complement could potentially have poor pseudorandomness properties. The most important result from this section is Lemma 3.8, which tells us that subsets of multiplication tables contain large matchings whenever the subset is obtained by taking unions of two random sets, and one (potentially adversarially chosen) deterministic set. We now give the details.

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 in U which are in the neighbourhood of both u and v, i.e. the number of vertices z in U such that there exists $v, w \in V(H)$ such that $\{u, z, v\}$ and $\{v, z, w\}$ are both edges of H. For a 2-uniform, bipartite graph 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 mutual neighbours of u and v in U.

We say that a r-partite r-uniform hypergraph H (where r will be either 2 or 3) is (γ, p, n)-**regular** if every part has $(1 \pm \gamma)n$ vertices and every vertex has degree

$(1 \pm \gamma)pn$. We say that H is (γ, p, n)-**typical** if, additionally, every pair of vertices x, y in the same part of H have pair degree $(1 \pm \gamma)p^2n$ into every other part of H. We say that a hypergraph is **linear** if through every pair of vertices, there is at most one edge. Multiplication hypergraphs have all these properties.

Observation 3.3 *For a group G of order n, the multiplication hypergraph H_G is $(0, 1, n)$-typical and linear.*

Proof It is immediate that all parts have size n. For any two vertices u, v in different parts, there is a unique edge through u and v. If $u \in A$, $v \in B$, then this edge is $(u, v, v^{-1}u^{-1})$. If $u \in B$, $v \in C$, then this edge is $(v^{-1}u^{-1}, u, v)$. If $u \in A$, $v \in C$, then this edge is $(u, u^{-1}v^{-1}, v)$, hence H_G is linear. This shows that all vertices have degree exactly n and pair degree exactly n to each part i.e. that the hypergraph is $(0, 1, n)$-typical. □

Frankl and Rödl [27] (also Pippenger, unpublished) showed that for all ε, $p \gg \gamma \gg n^{-1}$ every (γ, p, n)-regular hypergraph has a matching of size $(1 - \epsilon)n$. We need a well known variant of this where ε, p, γ have polynomial dependencies on n.

Lemma 3.4 *Let n be sufficiently large. Every (γ, δ, n)-regular linear tripartite hypergraph has a matching covering all but at most $n^{1-1/500} + 3\gamma n$ vertices.*

Proof There are various ways of proving this. We will deduce it from a result of Molloy-Reed — Theorem 1 from [45]. Applying that theorem with $k = 3$, $\Delta = (1 + \gamma)\delta n$ gives us a decomposition of our (γ, δ, n)-regular hypergraph into $\Delta + c_k \Delta^{1-\frac{1}{k}} \log^4 \Delta \leq \delta n + \gamma \delta n + n^{6/7}$ matchings (for some constant c_k). By the pigeonhole principle one of these has at least $\frac{e(H)}{\delta n + \gamma \delta n + n^{6/7}} \geq \frac{(1-\gamma)^2 \delta n^2/3}{\delta n + \gamma \delta n + n^{6/7}} \geq n - (n^{1-1/500} + 3\gamma n)$ edges as required. □

Typical graphs have the following well-known pseudorandomness condition, which dates back to work of Thomason [59]. Note that for the rest of the section, we assume that bipartite graphs come with a partition of their vertex set as (A, B) and similarly tripartite hypergraphs come with a partition (A, B, C).

Lemma 3.5 *Let G be a (γ, δ, n)-typical bipartite graph. Then for every $A' \subseteq A$, $B' \subseteq B$, we have $e(A', B') = \delta|A'||B'| \pm 5\sqrt{(\delta + \gamma)n^3} + \gamma n^2$.*

Proof This will be a consequence of Theorem 2 of [59]. First, delete at most γn vertices from one side of the graph to obtain a balanced bipartite graph. Now with parameters $p := \delta - \gamma$ and $\mu := 5\gamma n$ it is easy to see that the hypothesis of Theorem 2 are satisfied. From the conclusion of Theorem 2, for every $A' \subseteq A$, $B' \subseteq B$, we have $e(A', B') = p|A'||B'| \pm 5\sqrt{(\delta + \gamma)n^3}$. With another γn^2 term, we can account for the deleted vertices in the beginning, implying the desired bound. □

Typicality is preserved by taking random subsets, in the following sense.

Lemma 3.6 *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 bipartite graph between B and C consisting of edges passing through A' is $(n^{-1/5}, p, n)$-typical.*

Proof For some $c, c' \in C$, let $d_{A'}(c) = e(A', B, c)$ be the degree of c into A' and let $d_{A'}(c, c')$ denote the pair degree of (c, c') into A'. We have $d_A(c) = d_A(c, c') = n$ (by $(0, 1, n)$-typicality of H).

Note that $\mathbb{E}(d_{A'}(c)) = pd_A(c) = pn$ and $\mathbb{E}(d_{A'}(c, c')) = p^2 d_A(c, c') = p^2 n$ (using linearity of H) for all c, c'. Set $\gamma := n^{-1/5}$. By Chernoff's bound and a union bound, with probability at least $1 - 1/n^4$, for all c we have $d_{A'}(c) = \mathbb{E}(d_{A'}(c)) \pm \gamma n = pn \pm \gamma n$. Note that $d_{A'}(c, c')$ is 2-Lipschitz. This is because for each a, there is exactly one b with abc an edge, and one b with abc' an edge (using linearity of H). Hence by Azuma's inequality and a union bound, we have that with probability at least $1 - 1/n^4$, for each pair $c, c' \in C$, $d_{A'}(c, c') = \mathbb{E}(d_{A'}(c, c')) \pm \gamma n = p^2 n \pm \gamma n$. Corresponding bounds hold for $b, b' \in B$. With probability at least $1 - 3/n^4$ all these properties hold simultaneously. Whenever these properties all hold, we have that the bipartite graph (B, C) consisting of edges through A' is (γ, p, n)-typical as desired. □

Using the previous pseudorandomness property, we can derive the following lemma which states most vertices send approximately the expected number of edges through a random set and a deterministic set.

Lemma 3.7 *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 $e_H(A', B', c) \neq p|B'| \pm n^{9/10}$.*

Proof Set $\gamma = n^{-1/5}$. By Lemma 3.6, with probability at least $1 - 1/n^3$, we have that the bipartite graph between B and C consisting of edges passing through A' is $(n^{-1/5}, p, n)$-typical. Supposing this property holds, by Lemma 3.5, for any $B' \subseteq B$, $C' \subseteq C$, we have $e_H(A', B', C') = p|B'||C'| \pm 5\gamma^{1/2} n^2$. Let C^- be the set of vertices with $e_H(A', B', c) < p|B'| - \gamma^{1/4} n$. We have that $e_H(A', B', C^-) < p'|B'||C^-| - \gamma^{1/4} n|C^-|$ and $e_H(A', B', C^-) = p'|B'||C^-| \pm 5\gamma^{1/2} n^2$ implying $|C^-| \leq 10\gamma^{1/4} n$. Similarly letting C^+ be the set of vertices with $e_H(A', B', c) > p|B'| + \gamma^{1/2} n$, we get $|C^+| \leq 10\gamma^{1/4} n$. Plugging in the value of γ, this implies the lemma. □

The following lemma will allow us to find a large matching whenever we are given two random subsets and a deterministic subset.

Lemma 3.8 *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, and let B' a p-random subset of B, where A' and B' are not necessarily independent. Then, with probability at least $1 - 10n^{-3}$, the following holds.*

For any $C' \subseteq C$ of size $(1 \pm n^{-0.2})pn$, there is a matching covering all but $2n^{1-1/500}$ vertices in $A' \cup B' \cup C'$.

Proof With probability at least $1 - 10n^{-3}$, A' and B' satisfy the conclusion of Lemma 3.7 and have size $(1 \pm n^{-0.2})pn$ (by Chernoff's bound). This means that $A' \cup B' \cup C'$ has $\leq n^{9/10}$ vertices with degree $\neq p^2 n \pm n^{19/20}$ in $H[A', B', C']$.

Deleting all such vertices gives a hypergraph satisfying the hypothesis of Lemma 3.4 with $\gamma := n^{-0.01}$, hence the desired matching exists. □

In some applications the following formulation which allows for three deterministic subsets as opposed to just one will be more convenient. The result follows simply by applying the previous result four times.

Lemma 3.9 *Let $H = H_G$ be a multiplication hypergraph. Let $p \geq n^{-1/650}$. Let A', B', C' be p-random subsets of A, B, C respectively, not necessarily independent. Set $R := A' \cup B' \cup C'$. With high probability the following holds. Let $q \leq 5p$. For any $X \subseteq V(H) \setminus R$ with $|X \cap A|, |X \cap B|, |X \cap C| = (1 \pm n^{-0.25})qn$, there is a matching in $R \cup X$ covering all but at most $n^{1-10^{-4}}$ vertices of $R \cup X$.*

Proof Let q be a fixed rational number between 0 and 1 and denominator at most n. Suppose first that $q \leq n^{-1/600}$. We have that Lemma 3.8 holds for A' and B' with probability $\geq 1 - n^{1.5}$ and by Chernoff's bound $C' = (1 \pm n^{-0.2})pn$, with probability $\geq 1 - n^{-2}$. Both properties hold simultaneously with probability $\geq 1 - n^{-1.49}$. Then, using Lemma 3.8, $A' \cup B' \cup C'$ has a matching covering all but $n^{1-10^{-3}}$ vertices. Together with X, this gives $n^{1-10^{-3}} + 2n^{599/600} \leq n^{1-10^{-4}}$ vertices.

Suppose now that $q \geq n^{-1/600}$. For each $\diamond \in \{A, B, C\}$, partition $\diamond'$ into a q-random set $\diamond_1$, and a q-random set $\diamond_2$, and a $(p - 2q)$-random set $\diamond_3$. As $q, p - 2q \geq n^{-1/600}$ by assumption, with probability $\geq 1 - n^{-1.49}$ the pairs (A_1, B_1), (A_2, C_1), (B_2, C_2) and (A_3, B_3) satisfy the property of Lemma 3.8, and $| \diamond_i | = (1 + n^{-0.22})\mathbb{E}[| \diamond_i |]$ for each $\diamond_i$.

Let $X_A = X \cap A$, $X_B = X \cap B$, $X_C = X \cap C$ to get sets of size $(1 \pm n^{-0.2})qn$. Using Lemma 3.8, we have matchings M_1, M_2, M_3, M_4 covering all, but at most $n^{1-10^{-3}}$ vertices of $A_1 \cup B_1 \cup X_C$, $A_2 \cup X_B \cup C_1$, $X_A \cup B_2 \cup C_2$, and $A_3 \cup B_3 \cup C_3$ respectively. In total, the number of uncovered vertices does not exceed $4n^{1-10^{-3}} \leq n^{1-10^{-4}}/10$.

Taking a union bound over all rational q with denominator at most n, we have shown that with high probability, there exists a matching covering all but $n^{1-10^{-4}}/10$ vertices. The statement for real values of q follows simply by using the property for the closest rational value q' to q with denominator at most n. Indeed, leaving out or deleting few elements, we can ensure that the set X has $|X \cap A|$, $|X \cap B|$, $|X \cap C| = (1 \pm n^{-0.25})q'n$, thereby obtaining a matching that covers all but $n^{1-10^{-4}}/10 + n^{1-10^{-4}}/10 \leq n^{1-10^{-4}}$ elements of the original sets $X \cup R$. □

3.2 The multiplication hypergraph

We make some clarifications regarding our notation with the multiplication hypergraph H_G of a group (recall that this was defined in Sect. 2). While referring to the parts of the multiplication hypergraph, we often omit the $A/B/C$ subscripts and think of A_G, B_G, C_G simply as copies of G. For example, whenever we have some vertices $v_1, v_2, \ldots, v_k \in V(H_G)$, we write $v_1 v_2 \ldots v_k$ to mean the product of the corresponding group elements in G (as opposed to in any of the copies A_G, B_G, C_G). Similarly, if $v \in V(H_G)$ and U is a subset of G, then we use $v \in U$ to mean that v is an element of G after dropping the $A/B/C$ subscript. For any subset $S \subseteq G$, we use $S_A/S_B/S_C$ to denote the corresponding subsets of $G_A/G_B/G_C$ respectively.

A subset S of $V(H_G)$ is called **balanced**, if $|S \cap G_\diamond|$ does not depend on the value of $\diamond \in \{A, B, C\}$.

3.3 Basic group theory definitions and results

Throughout the paper we use e to denote the identity element of a group G. An involution is an element of order exactly 2. Recall that we use G' to denote the commutator subgroup of a group G. That is, G' is the subgroup generated by elements of the form $[g,h] := ghg^{-1}h^{-1}$ where $g, h \in G$. We denote the abelianization of G (quotient of G by G') as G^{ab}. We sometimes call the elements of G^{ab} G'-**cosets**. For $g \in G$, we use $[g]$ to denote the unique G'-coset that g is a member of. When $g \in V(H_G)$, we think of $[g]$ as the G'-coset that resides in the same part $G_A/G_B/G_C$ that g resides in. As mentioned in the introduction, whenever we use additive notation together with elements of G, all operations take place in G^{ab}. We do this so that we don't have to use the $[g]$ notation excessively.

Lemma 3.10 *$g \in G'$ if, and only if, g can be written as $g = g_1 \ldots g_t$ such that there is a permutation σ of $[t]$ with $g_{\sigma(1)} \ldots g_{\sigma(t)} = e$.*

Proof To see the "only if" direction, write g as a product of commutators as

$$g = [a_1, b_1] \ldots [a_t, b_t] = a_1 b_1 a_1^{-1} b_1^{-1} \ldots a_t b_t a_t^{-1} b_t^{-1}.$$

Clearly, the latter product can be permuted as $a_1 a_1^{-1} b_1 b_1^{-1} \ldots a_t a_t^{-1} b_t b_t^{-1} = e$, as required. For the "if" direction, consider some $g_1 \ldots g_t$ which rearranges into $g_{\sigma(1)} \ldots g_{\sigma(t)} = e$. Consider the quotient homomorphism $\phi : G \to G/G'$. Then since G/G' is abelian, we have $\phi(g_1 \ldots g_t) = \phi(g_1) \ldots \phi(g_t) = \phi(g_{\sigma(1)}) \ldots \phi(g_{\sigma(t)}) = \phi(g_{\sigma(1)} \ldots g_{\sigma(t)}) = \phi(e) = e$ i.e. $g_1 \ldots g_t \in ker(\phi) = G'$ as required. □

The following is a well-known property of finite abelian groups.

Theorem 3.11 (Fundamental theorem of finite abelian groups) *Let G be an abelian group. Then, G is isomorphic to a product of cyclic prime-power order groups.*

Given $g \in G$, $s(g)$ denotes the size of the set $\{x \in G : x^2 = g\}$.

Proposition 3.12 ([30]) *Let G be group such that there exists some $g \in G$ with $s(g) > (3/4)|G|$. Then, G is an elementary abelian 2-group and $g = 0$.*

Proof We express our gratitude for all participants of the active discussion that took place on Math Overflow including Emil Jeřábek, Derek Holt, Saúl Rodríguez Martín, and Terry Tao. Here, we reproduce the argument of GH from MO [30]. Let $g \in G$, and suppose that $|s(g)| > (3/4)|G|$. Fix some $y \in G$, and define $S = \{x \in G : x^2 = g\}$ and $T = \{x \in S : xy \in S\}$. Observe that $|G \setminus T| \leq 2|G \setminus S|$, since $x \notin T$ only if $x \notin S$ or $xy \notin S$, and there are at most $|G \setminus S|$ many x of either type. By assumption $|G \setminus S| < |G|/4$, so $|G \setminus T| < |G|/2$. Consequently, $|T| > |G|/2$.

Observe that for any $x \in T$, we have that $(xy)^2 = g = x^2$, and so we have

$$xyx^{-1} = (xy)(xy)^{-2}(xy)^2x^{-1} = (xy)^{-1}x^2x^{-1} = y^{-1}.$$

It is an easy exercise to check that $C = \{x \in G : xyx^{-1} = y^{-1}\}$ is a coset of $C(y)$, the centralizer subgroup of y.

As $|C| > |G|/2$, and C is a coset, $C = G$ by Lagrange's theorem. This implies that $y = y^{-1}$ for each y, hence G is an elementary abelian 2-group. It follows that $g = 0$, as $s(g) = 0$ for all $g \neq 0$ in an elementary abelian 2-group. □

3.4 Generic elements, and choice of a_ϕ, b_ϕ, c_ϕ

The following definition is critical.

Definition 3.13 A group element $g \in G$ is **generic** if $g \neq e$ and there are at most $n/10^{9000}$ solutions to $x^2 = g$ in G.

Let $N(G)$ denote the set of non-generic elements and note that $|N(G)| \leq 10^{9000}$, simply as in a group with n elements, the number of g with more than k square-roots is at most n/k. Similarly, we call vertices of H_G generic if the corresponding group element is generic.

As described in Sect. 2, it is critical to our method to pair up group elements with a fixed sum (which can be viewed as defining a suitable involution ϕ), where the fixed sum has some desirable properties. The following lemma serves to show that the pairing we desire exists.

Lemma 3.14 *For every group G, there exist $a_\phi, b_\phi, c_\phi \in G$ such that the following all hold.*

(*a*) $a_\phi b_\phi c_\phi = e$.
(*b*) *There are at most* 30 *values of $x \in G$ such that $x^2 \in \{a_\phi, b_\phi, c_\phi\}$. In particular, a_ϕ, b_ϕ, c_ϕ are generic.*
(*c*) *There are at most* $30|G'|$ *values of $x \in G$ such that $x^2 \in [a_\phi] \cup [b_\phi] \cup [c_\phi]$.*
(*d*) *If $|G'| \leq 10^{-9}n$, then $a_\phi, b_\phi, c_\phi \notin G'$.*

Proof Choose a_ϕ, and b_ϕ uniformly at random and set $c_\phi := (a_\phi b_\phi)^{-1}$, noting c_ϕ is then also sampled uniformly at random. For a random $g \in G$, the probability that g has more than 30 square-roots is at most $1/30$. By the same argument applied to G^{ab}, the probability that $[g]$ has more than 30 square-roots in G^{ab} is also at most $1/30$. The probability that $g \in G'$ is at most 10^{-9} if $|G'| \leq 10^{-9}n$. Then, with positive probability, all the conditions are satisfied for a_ϕ, b_ϕ, and c_ϕ by a union bound. □

Definition 3.15 We say that an element x is **ϕ-generic** if x, $a_\phi^{\pm1}x$, $b_\phi^{\pm1}x$, $c_\phi^{\pm1}x$, $a_\phi^{\pm1}b_\phi^{\pm1}x$, $b_\phi^{\pm1}c_\phi^{\pm1}x$, $c_\phi^{\pm1}a_\phi^{\pm1}x$ are all generic.

Notice that the number of elements which aren't ϕ-generic is $\leq 30|N(G)| \leq 10^{9010}$. We call a subset ϕ-generic if all elements of that subset are ϕ-generic.

For each group G, we fix a triple a_ϕ, b_ϕ, c_ϕ with the properties as in Lemma 3.14. We call two vertices v and w of H_G coming from the same part a **pair** if $v \cdot w \in [x_\phi]$ where $x = a, b, c$ depending on whether $v, w \in A, B, C$.

We call a subset S of $V(H_G)$ **coset-paired** if S can be partitioned into a disjoint union of pairs. We call a coset $[x]$ of G, viewed as a subset of G_A, G_B or G_C, **self-paired** if $[x^2]$ is equal to $[a_\phi]$, $[b_\phi]$, or $[c_\phi]$, respectively. Note that Lemma 3.14(c) implies in particular that there are at most 30 self-paired cosets. A subset $S \subseteq V(H_G)$ being coset-paired is equivalent to the following statement: $|S \cap [g]| = |S \cap [x_\phi g^{-1}]|$ for every non-self-paired coset $[g]$ and $|S \cap [g]|$ is even for every self-paired coset $[g]$.

3.5 Symmetric sets

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 $R_1, \ldots, R_k \subseteq S$ 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. Considering such disjoint distributions complicates our approach as it makes various gadgets significantly more difficult to find. The reason we are interested in such distributions is the applications we give later on in the paper. Indeed, all applications we give other than the alternative proof of Hall-Paige conjecture and Snevily's conjecture require that we work with such disjoint distributions.

In fact, we need to generalise the concept of a disjoint distribution even further so that we can work with random sets X, Y, Z where X, Y and $Z^{-1} = \{z^{-1} : z \in Z\}$ are sampled disjointly. This need comes from the applications to sequenceability and R-sequenceability. Thankfully, this generalisation does not create many additional combinatorial difficulties. However, we still need the following definitions to state a single theorem that covers all of the applications we want to give.

For $g \in G$, define $\hat{g} := \{g, g^{-1}\}$, noting that $\hat{g}$ has size 1 or 2 (depending on whether g has order ≤ 2 or not). For a subset $T \subseteq G$, let $\hat{T} = \{\hat{t} : t \in T\}$ and $\bigcup \hat{T} = \bigcup_{t \in T} \hat{t} = T \cup T^{-1}$. We say that a subset $T \subseteq G$ is **symmetric** if $T^{-1} = T$ (or equivalently if $T = \bigcup \hat{T}$). We call a subset $S \subseteq V(H_G)$ symmetric if $S \cap A$, $S \cap B$, $S \cap C$ are all symmetric. We say that R is a **symmetric** p-**random subset** of G if R is always symmetric and $\hat{R}$ is a p-random subset of $\hat{G}$ (or equivalently if R is formed by flipping an independent coin for each $\hat{g} \in \hat{G}$ and taking the union of all group elements for which heads comes up). We say that R^1, R^2, R^3 are **disjoint symmetric** p-**random subsets** of G if additionally the joint distribution of $\hat{R}^1$, $\hat{R}^2$, $\hat{R}^3$ is that of disjoint p-random subsets of $\hat{G}$. The following two lemmas are useful as they allow us to jump between these definitions.

Lemma 3.16 *Let R_1, R_2, R_3 be disjoint p-random sets. Then there are $S_1 \subseteq R_1$, $S_2 \subseteq R_2$, $S_3 \subseteq R_3$ so that the joint distribution on S_1, S_2, S_3 is that of disjoint symmetric p^2-random sets.*

Proof Let T be a p-random subset of G, independent of R_1, R_2, R_3. For non-involutions, non-identity elements g, place g, g^{-1} into S_i whenever $\{g, g^{-1}\} \subseteq R_i$. For g an involution or the identity, place g into S_i whenever $g \in R_i \cap T$. Notice that for each $g \in G$, this gives $\mathbb{P}(\hat{g} \subseteq S) = p^2$. Also, these events are mutually independent for different $\hat{g}$, $\hat{h}$ since they depend on different coordinates. □

Lemma 3.17 *Let Q_1, Q_2, Q_3, R be random sets with Q_1, Q_2, Q_3 being disjoint symmetric q-random and R being p-random and independent of Q_1, Q_2, Q_3. Then, there are $S_1 \subseteq Q_1 \cap R$, $S_2 \subseteq Q_2 \cap R$, $S_3 \subseteq Q_3 \cap R$ so that the joint distribution on S_1, S_2, S_3 is that of disjoint symmetric p^2q-random sets.*

Proof Let T be a p-random subset of G, independent of R, Q_1, Q_2, Q_3. For g which is not an involution or the identity, place g, g^{-1} into S_i whenever $\{g, g^{-1}\} \subseteq Q_i \cap R$. For g an involution or the identity, place g into S_i whenever $g \in Q_i \cap R \cap T$. Notice that for each $\hat{g} \in \hat{G}$, this gives $\mathbb{P}(\hat{g} \subseteq S_i) = p^2q$. Also, these events are mutually independent for different $\hat{g}$, $\hat{h}$ since they depend on different coordinates. □

3.6 Free products

The goal of this section is to prove Lemma 3.32, which is our main tool to find vertex-disjoint copies of constant sized substructures in multiplication hypergraphs. We first remind the reader of some standard group theoretic terminology. We use F_k to denote the free group on k generators $v_1, \dots, v_k$.

For a group G, we use $G * F_k$ to denote the free product of G and F_k. We call $v_1, \dots, v_k$ the **free variables** of the free product $G * F_k$. A **word** is simply an element of $G * F_k$. For $w \in G * F_k$, we define the length of w to be the minimum number ℓ needed to write $w = x_1x_2 \dots x_\ell$ for $x_i \in \{v_i, v_i^{-1} : i = 1, \dots, k\} \cup G$. We call a presentation of a word $g_0x_1g_1 \dots x_tg_t = w \in G * F_k$ **reduced** if it cannot be made shorter using the group operations, i.e. if it doesn't contain consecutive elements of G, doesn't contain e, and doesn't contain consecutive v_i, v_i^{-1} or v_i^{-1}, v_i for any i. It is a standard property of free products that every $g \in G * F_k$ is uniquely expressible as a reduced word.

Lemma 3.18 *For $w \in G * F_k$, there's a unique way of writing $w = g_0x_1g_1 \dots x_tg_t$ with $x_i \in \{v_1, \dots, v_k, v_1^{-1}, \dots, v_k^{-1}\}$ and $g_i \in G$ such that we don't have "$x_i = x_{i+1}^{-1}$ and $g_i = e$" for any $i \in \{0, 1, \dots, t-1\}$.*

Proof Let W be the set of words which can be written of the form $g_0x_1g_1 \dots x_tg_t$ with $x_i \in \{v_1, \dots, v_k, v_1^{-1}, \dots, v_k^{-1}\}$ and $g_i \in G$ such that we don't have "$x_i = x_{i+1}^{-1}$ and $g_i = e$" for any $i \in \{0, 1, \dots, t-1\}$. Let R be the set of reduced words of $G * F_k$. Let $f : W \to R$ be defined by mapping $w = g_0x_1g_1 \dots x_tg_t$ to the word formed by removing all copies of e from w. Let $g : R \to W$ be defined by mapping $w' = y_1y_2 \dots y_k$ to the word formed by inserting e between any y_i, y_{i+1} which are both in $\{v_1, \dots, v_k, v_1^{-1}, \dots, v_k^{-1}\}$. It is easy to see that $f(g(w')) = w'$ and $g(f(w)) = w$ for any $w \in W$ and $w' \in R$ i.e. both functions are bijections. Thus, since every $g \in G * F_k$ is uniquely expressible as a reduced word, it is also uniquely expressible as a word in W. □

From the above, we get that every $w \in G * F_k$ can be written as $w = g_0 x_1 g_1 \dots x_t g_t$ with $x_i \in \{v_1, \dots, v_k, v_1^{-1}, \dots, v_k^{-1}\}$ and $g_i \in G$. We say that $w \in G * F_k$ is **linear in** v_i there's a way of writing w like this with precisely one occurrence of v_i (meaning one occurrence v_i or v_i^{-1}, but not both). We say that w is **linear** if all variables occur at most once in w, and some variable occurs exactly once in w. A useful fact is that for a linear w, there's a unique way of writing it as $w = g_0 x_1 g_1 \dots x_t g_t$ (with $x_i \in \{v_1, \dots, v_k, v_1^{-1}, \dots, v_k^{-1}\}$ and $g_i \in G$) such that there's at most one occurrence of each variable. This comes from Lemma 3.18, because in a linear $w = g_0 x_1 g_1 \dots x_t g_t$ it is impossible to have $x_i = x_{i+1}^{-1}$, $g_i = e$ (since this would create two occurrences of the free variable x_i).

A homomorphism $\pi : G * F_k \to G$ is a **projection** if $\pi(g) = g$ for all $g \in G$.

Lemma 3.19 *For each function $f : \{v_1, \dots, v_k\} \to G$, there is precisely one projection $\pi_f : G * F_k \to G$ which agrees with f on $\{v_1, \dots, v_k\}$. In particular, there are precisely n^k projections $G * F_k \to G$.*

Proof Let $f : \{v_1, \dots, v_k\} \to G$ be a function. By the universal property of free groups, there is a unique homomorphism $\phi_f : F_k \to G$ which agrees with f on $\{v_1, \dots, v_k\}$. Let $e_G : G \to G$ be the identity homomorphism. By the universal property of free products there is a unique homomorphism π_f which agrees with ϕ_f on F_k and agrees with e_G on G. Such a homomorphism is exactly a projection from $G * F_k$ to G.

For the "in particular" part, note that the number of functions $\{v_1, \dots, v_k\} \to G$ is exactly n^k, and so there are this many projections $G * F_k \to G$ by the first part. □

We use $\pi_0 : G * F_k \to G$ to denote the projection which maps all $w \in F_k$ to e (i.e. the map coming from Lemma 3.19 via the function f mapping all v_i to e).

Observation 3.20 *For all $g, h \in G$ there are the same number of solutions to $x^2 \in [g]$ and $x^2 \in [h^2 g]$.*

In particular, for any $w, w' \in G * F_k$, there are the same number of solutions to $x^2 \in [\pi_0(w^{-1}w')]$ an $x^2 \in [\pi_0(ww')]$ in G).

Proof We have that $x^2 \in [g]$ if, and only if $(hx)^2 \in [h^2 g]$. Therefore, if S is the set of solutions to $x^2 \in [g]$, then hS is the set of solutions to $(hx)^2 \in [h^2 g]$. Since S and hS always have the same size in a group, this gives what we want.

The "in particular" part follows from the above by taking $g = \pi_0(w^{-1}w')$, $h = \pi_0(w)$, and noting that, since π_0 is a homomorphism, we have $\pi_0(ww') = h^2 g$ □

Definition 3.21 Let $w, w' \in G * F_k$. We say that w and w' are **strongly separable** if any of the following hold.

(a) A free variable v_i appears once in one of w/w', and never in the other.
(b) w, w' are linear and there is a $g \in G$ with g generic so that $w' \in \{gw, g^{-1}w, gw^{-1}, g^{-1}w^{-1}, wg, wg^{-1}, w^{-1}g, w^{-1}g^{-1}\}$.
(c) All of the following hold.
  - $|G'| \leq 10^{-9} n$

- w and w' are linear and have the same free variables (potentially with different signs).
- We either have $\pi_0(ww') \notin G'$ or some free variable occurs with the same sign in w, w'.
- We either have $\pi_0(w^{-1}w') \notin G'$ or some free variable occurs with the opposite sign in w, w'.
- There are $\leq 90|G'|$ solutions to $x^2 \in [\pi_0(ww')]$ (or equivalently by Observation 3.20, there are $\leq 90|G'|$ solutions to $x^2 \in [\pi_0(w^{-1}w')]$ in G).

We remark that strong separability is a symmetric relation.

Definition 3.22 We say that two sets $S, T \subseteq G$ are **strongly separable** if every pair of elements $s \in S$, $t \in T$ are strongly separable.

The following observation is quite critical, and justifies why considering symmetric disjoint random sets doesn't create additional complications when compared to disjoint random sets.

Observation 3.23 *w, w' are strongly separable $\iff$ w^{-1}, w' are strongly separable $\iff$ $\hat{w}, \hat{w}'$ are strongly separable.*

Proof Within this proof, we say separable to mean strongly separable. For "w, w' are separable $\iff$ w^{-1}, w' are separable": Consider the possible cases of Definition 3.21. If w, w' are separable by (a), then w^{-1}, w' are also separable by (a) (this is immediate when one considers the "a free variable v_i appears once in one of w/w', but not both" version of (a)). If w, w' are separable by (b), then w^{-1}, w' are also separable by (b) — this is true because the set $\{gw, g^{-1}w, gw^{-1}, g^{-1}w^{-1}, wg, wg^{-1}, w^{-1}g, w^{-1}g^{-1}\}$ doesn't change if you replace each "w" with "w^{-1}". If w, w' are separable by (c), then w^{-1}, w' are also separable by (c). (The first bullet point doesn't involve w, w'. The second bullet point doesn't change by replacing w with w^{-1} because w, w^{-1} are always linear in the same free variables. The 3rd and 4th bullet points get exchanged when replacing w by w^{-1}. The 5th bullet point doesn't change when replacing w by w^{-1} since there is the same number of solutions to $x^2 \in [\pi_0(ww')]$ and $x^2 \in [\pi_0(w^{-1}w')]$ by Observation 3.20).

The direction "$\hat{w}$, $\hat{w}'$ are separable $\implies$ w, w' are separable" is immediate from the definition of "S, T are separable". For "w, w' are separable $\implies$ $\hat{w}$, $\hat{w}'$ are separable", note that once we know that w, w' are separable, we also know that the pairs (w^{-1}, w'), $(w, (w')^{-1})$, $(w^{-1}, (w')^{-1})$ are separable (all coming from "w, w' are separable $\iff$ w^{-1}, w' are separable"). This gives that $\hat{w}$, $\hat{w}'$ are separable. □

Observation 3.24 *Let $S \subseteq G * F_k$ be a set of linear elements. For each $T \subseteq \{v_1, \dots, v_k\}$, let S_T be the set of $w \in W$ such that the free variables in w are exactly the set T. Then the sets S_T, $S_{T'}$ are strongly separable for distinct T, T'.*

Proof If T, T' are distinct, then there's some $v_i \in T \Delta T'$ say $v_i \in T \setminus T'$. For any $w \in S_T$, $w' \in S_{T'}$, we have that v_i appears in w (just once by linearity) but not w', and so part (a) of the definition of "strongly separable" applies. □

Definition 3.25 Let $w, w' \in G * F_k$. We say that w and w' are **weakly separable** if either they are strongly separable, or they satisfy the following property.

(b') For some non-identity element g, the equation $w = w'$ rearranges into $e = g$, meaning that $w^{-1}w'$ is conjugate to an element of G.

We remark that the key difference between strong and weak separability comes from the property (b') not necessarily holding when w is replaced with w^{-1}, i.e. Observation 3.23 fails for weak separability. This will not be an issue while we search for gadgets, as we rely on weak separation only to separate elements coming from the same set A_G, B_G, C_G.

Definition 3.26 Let $S \subseteq G * F_k$. We say that a homomorphism $\phi : G * F_k \to G$, **separates** S if for every weakly-separable $w, w' \in S$ we have $\phi(w) \neq \phi(w')$.

Recall that $(G * F_k)'$ denotes the commutator subgroup of $(G * F_k)$.

Lemma 3.27 *For a group G, let $w, w' \in G * F_k$ satisfy part (c) of the definition of "strongly separable". Then $w = w'$ rearranges to $u^2 = \pi_0(w^{-1}w')y$ for some $y \in (G * F_k)'$ and some $u \in F_k$ which is either linear or equals e. Additionally, $u = e$ only if the free variables occur with the same signs in w, w'.*

Proof Let $w = g_0x_1g_1 \ldots x_tg_t$ and $w' = g_0'x_1'g_1' \ldots x_{t'}'g_{t'}'$ (with $x_i, x_i' \in \{v_1, \ldots, v_k, v_1^{-1}, \ldots, v_k^{-1}\}$ and $g_i, g_i' \in G$). Note that the assumption "w and w' are linear and have the same free variables" implies that $t = t'$, and that there is a permutation σ of $[t]$ so that $x_i' \in \{x_{\sigma(i)}, x_{\sigma(i)}^{-1}\}$ for $i = 1, \ldots, t$. Partition $[t] = I^+ \cup I^-$ with $I^+ = \{i : x_i' = x_{\sigma(i)}\}$ and $I^- = \{i : x_i' = x_{\sigma(i)}^{-1}\}$. Set $u = \prod_{i \in I^-}(x_i')^{-1}$, noting that u is in F_k, and that $u = e \iff I^- = \emptyset \iff$ the free variables occur with the same signs in w, w'. This also implies that if $u \neq e$, then u is linear. Note that $w = w'$ rearranges to $u^2 = \pi_0(w^{-1}w')y$ where $y = \pi_0(w^{-1}w')^{-1}u^2w'w^{-1}$. Notice that

$$\begin{aligned}\pi_0(w^{-1}w')^{-1}u^2w'w^{-1} &= \pi_0(w')^{-1}\pi_0(w)u^2w'w^{-1}\\ &= (g_t'^{-1} \ldots g_1'^{-1}g_0'^{-1})(g_0 \ldots g_{t-1}g_t)(\prod_{i \in I^-}(x_i')^{-1})\\ &\quad \times (\prod_{i \in I^-}(x_i')^{-1})(g_0'x_1'g_1' \ldots x_t'g_t')(g_t^{-1}x_t^{-1}g_{t-1}^{-1} \ldots x_1^{-1}g_0^{-1})\end{aligned}$$

Notice that the above product can be permuted into the identity. Indeed each $g_i'^{-1}$ in the 1st bracket cancels with g_i' in the 5th bracket, each g_i in the 2nd bracket cancels with g_i^{-1} 6th bracket, for $i \in I^-$ each $(x_i')^{-1}$ in the 3rd bracket cancels with x_i' in the 5th bracket, each $(x_i')^{-1}$ in the 4th bracket cancels with $(x_{\sigma(i)})^{-1} = x_i'$ in the 6th bracket, and for $i \in I^+$ each x_i' in the 5th bracket cancels with $x_{\sigma(i)}^{-1} = (x_i')^{-1}$ in the 6th bracket. Thus, by Lemma 3.10, $t \in (G * F_k)'$. □

Lemma 3.28 *Let $w \in G * F_k$ be linear in some free variable v_i and let $g \in G$. Then there are exactly n^{k-1} projections $\pi : G * F_k \to G$ having $\pi(w) = g$.*

Proof Without loss of generality, suppose $i = k$. By Lemma 3.19 there are exactly n^{k-1} projections $\pi : G * F_{k-1} \to G$. We will show that for every such projection, there is a unique projection $\pi' : G * F_k \to G$ that agrees with π on $G * F_{k-1}$ and additionally has $\pi'(w) = g$. To see this, note that the equation $w = g$ in the group $G * F_k$ rearranges into $v_k = h$ where $h \in G * F_k$ such that the free variable v_k doesn't occur in h (this is possible because w is linear in v_k). This shows that for any projection π', the equation $\pi'(w) = g$ is equivalent to $\pi'(v_k) = \pi'(h)$ (using that for any $g \in G$ we have $\pi'(g) = g$ for any projection π'). Since $h \in G * F_{k-1}$, the image $\pi(h) \in G$ is defined. Thus a projection π' agrees with π on $F_{k-1} * G$ and also has $\pi'(w) = g$ if and only if $\pi'(v_1) = \pi(v_i), \dots, \pi'(v_{k-1}) = \pi(v_{k-1})$, and also $\pi'(v_k) = \pi(h)$. By Lemma 3.19, there is a unique projection satisfying this, as required. □

Lemma 3.29 *Let $S \subseteq G * F_k$ be a set of elements which are each linear in at least one variable, and let $U \subseteq G$. Then the number of projections $\pi : G * F_k \to G$ for which $\pi(S)$ intersects U is $\leq |S||U|n^{k-1}$.*

Proof For each $w \in S$ and $u \in U$, by Lemma 3.28, there are n^{k-1} projections $\pi : G * F_k \to G$ with $\pi(w) = u$. Thus, summing over all w, u, there are at most $\leq |S||U|n^{k-1}$ projections with $\pi(S)$ intersecting U. □

Lemma 3.30 *Let n be sufficiently large. Let $S \subseteq G * F_k$ be a set of size ≤ 1000. Then there are at most $0.1n^k$ projections $\pi : G * F_k \to G$ which do not separate S. If $|G'| > 10^{-9}n$, then this can be improved to "at most $10^{-8900}n^k$ projections".*

Proof The total number of projections $\pi : G * F_k \to G$ is n^k by Lemma 3.19. Consider two weakly separable words $w, w' \in S$. We will count the number of projections for which $\pi(w) = \pi(w')$. There are four cases, depending on which part of the definition of weakly separable applies to w, w'.

(a) Note that the equation $\pi(w) = \pi(w')$ can be rearranged into $\pi(w^{-1}w') = e$. Since in (a), $w^{-1}w'$ is linear in some variable, Lemma 3.28 tells us that there are $n^{k-1} \leq n^k/10^{9000}$ projections $\pi : G * F_k \to G$ for which $\pi(w^{-1}w') = e$.

(b') Since $w = w'$ rearranges into $e = g$, the equation $\pi(w) = \pi(w')$ rearranges into $\pi(e) = \pi(g)$. But this is impossible for a projection π (since the definition of "projection" gives that $\pi(e) = e$ and $\pi(g) = g$). Thus there are zero projections with $\pi(w) = \pi(w')$ in this case.

(b) If we are not in some case covered by the previous bullet point, then the equation $w = w'$ rearranges into $(w)^2 = g$ for a generic group element $g \in G$. Let T_g be the set of solutions to $x^2 = g$ to get a set of size $\leq n/10^{9000}$, using the definition of a generic group element. For a projection π to have $\pi(w) = \pi(w')$, it must have (using that π is a homomorphism) $\pi(w)^2 = \pi(g) = g$ and so $\pi(w) \in T_g$. By Lemma 3.29, there are $|T_g|n^{k-1} \leq n^k/10^{9000}$ projections with $\pi(w) \in T_g$.

(c) Using Lemma 3.27, $w = w'$ rearranges into $u^2 = \pi_0(w^{-1}w')k$ where $u \in F_k$ is either linear or $u = e$, and $k \in (G * F_k)'$. If $u = e$, then we know that free variables occur with the same signs in w, w', which tells us that $\pi_0(w^{-1}w') \notin$

G'. This gives that $[\pi_0(w^{-1}w')k] \neq G'$, and so $e \neq \pi_0(w^{-1}w')k = u^2 = e$, a contradiction. Thus in this case, there are no projections with $\pi(w) = \pi(w')$. So we can assume that $u \neq e$ which implies by Lemma 3.27 that u is linear.

Let T be set of solutions to $x^2 \in [\pi_0(w^{-1}w')]$ in G. Since we are in (c), we have $|T| \leq 90|G'| \leq 90 \cdot 10^{-9}n \leq 10^{-7}n$. For $\pi(w) = \pi(w')$ to hold, we must have (using that π is a homomorphism) $\pi(u^2) = \pi(\pi_0(w^{-1}w')k)$, which implies (using that π is a projection) $\pi(u)^2 = \pi_0(w^{-1}w')\pi(k)$. Since $\pi_0(w^{-1}w')\pi(k) \in [\pi_0(w^{-1}w')]$ for any projection π (we have $\phi(H') \subseteq G'$ for any group homomorphism $\phi : H \to G$, since if x is a commutator in H, then $\phi(x)$ is a commutator in G), this would imply that $\pi(w) \in T$. By Lemma 3.29, there are $|T|n^{k-1} \leq 10^{-7}n^k$ projections with $\pi(w) \in T$ as required.

There are at most $\binom{|S|}{2} \leq \binom{1000}{2} \leq 10^6$ pairs of weakly separable $w, w' \in S$, and for each of them there are $\leq n^k/10^7$ projections with $w = w'$. Thus in total there are $\leq 10^6 n^k/10^7 = 0.1n^k$ projections which don't separate S. This implies the lemma when $|G'| \leq 10^{-9}n$. When $|G'| > 10^{-9}n$, note that case (c) can't occur, so we actually have $\leq n^k/10^{9000}$ projections with $w = w'$ for each separable w, w'. This gives a total of $\leq 2 \cdot 10^6 n^k/10^{9000} \leq n^k/10^{8900}$ projections which don't separate S. □

Lemma 3.31 *Let n be sufficiently large, $k \leq 200$, and $S \subseteq G * F_k$ a set of ≤ 1000 elements which are linear in at least one variable. There are projections $\pi_1, \dots, \pi_{10^{-3000}n}$ which separate S and have $\pi_1(S), \dots, \pi_{10^{-3000}n}(S)$ disjoint. If $|G'| > 10^{-9}n$, then we can additionally ensure that $\pi_j(v_i) \in G'$ for all free variables v_i and projections π_j.*

Proof For $|G'| \leq 10^{-9}n$ say that a projection π is good if it separates S. For $|G'| > 10^{-9}n$ say that a projection π is good if it separates S and has $\pi(v_i) \in G'$ for all $v_i \in F_k$. Our task is to find $10^{-3000}n$ good projections $\pi_i(S)$, which have $\pi_i(S)$ disjoint for different i. Consider a maximal family $\pi_1, \dots, \pi_t$ of good projections. which have $\pi_1(S), \dots, \pi_t(S)$ disjoint. Let $T = \pi_1(S) \cup \dots \cup \pi_t(S)$. By maximality, we have that all good projections π have $\pi(S) \cap T \neq \emptyset$. Lemma 3.29 tells us that the number of projections with $\pi(S) \cap T \neq \emptyset$ is $\leq |T|n^{k-1} \leq t|S|n^{k-1} \leq 1000tn^{k-1}$. Thus we established that there are $\leq 1000tn^{k-1}$ good projections.

Suppose $|G'| \leq 10^{-9}n$. Lemma 3.30 tells us that there are $\leq 0.1n^k$ bad projections, and hence $\geq n^k - 0.1n^k = 0.9n^k$ good projections. Thus $t \geq 0.9n/1000 \geq 10^{-3000}n$.

Now suppose $|G'| > 10^{-9}n$: There are $|G'|^k$ projections with $\pi(v_i) \in G'$ for all $v_i \in F_k$ (using Lemma 3.19). By Lemma 3.30, there are $\leq n^k/10^{8900} \leq 0.1(10^{-9}n)^k \leq 0.1|G'|^k$ projections that don't separate S. Hence there remain $\geq 0.9|G'|^k \geq 0.9 \cdot 10^{-9k}n^k$ projections that both separate S and have all $\pi(v_i) \in G'$ i.e. there are $\geq 0.9 \cdot 10^{-9k}n^k$ good projections. Combining with "there are $\leq 1000tn^{k-1}$ good projections", this gives $t \geq 0.9 \cdot 10^{-9k}n/1000 \geq 10^{-3000}n$. □

Lemma 3.32 *Let $p \geq n^{-1/700}$. Let R_A, R_B, R_C be disjoint p-random symmetric subsets of G and set $R = R_A \cup R_B \cup R_C$. With high probability, the following holds:*

Let $k \leq 200$, $S \subseteq G * F_k$ a set of ≤ 600 elements of length ≤ 200 which are each linear in at least one variable, and $U \subseteq G$ with $|U| \leq p^{800}n/10^{4000}$. Then there is a projection $\pi : G * F_k \to G$ which separates S and has $\pi(S) \subseteq R \setminus U$. Moreover:

- *For any $S_A, S_B, S_C \subseteq S$, with S_A, S_B, S_C being pairwise strongly separable sets, we can ensure $\pi(S_A) \subseteq R_A, \pi(S_B) \subseteq R_B, \pi(S_C) \subseteq R_C$.*
- *If $|G'| > 10^{-9}n$, then we can additionally ensure that $\pi(v_i) \in G'$ for all free variables v_i.*

Proof We can assume that n is sufficiently large (otherwise the lemma is vacuous since "with high probability" wouldn't mean anything). First fix $k \le 200$, and some set $S \subseteq G$, and $S_A, S_B, S_C \subseteq S$ as in the lemma. Note $|\bigcup \hat{S}| \le 2|S| \le 1200$. Fixing $m := n/10^{3000}$, apply Lemma 3.31 to get projections $\pi_1, \dots, \pi_m$ which separate $\bigcup \hat{S}$ and have $\pi_1(\bigcup \hat{S}), \dots, \pi_m(\bigcup \hat{S})$ disjoint (and additionally, when $|G'| > 10^{-9}n$, having that $\pi_j(v_i) \in G'$ for all free variables v_i and projections π_j).

Note that for each i we have $\pi_i(\hat{S}_A)$, $\pi_i(\hat{S}_B)$, $\pi_i(\hat{S}_C)$ pairwise disjoint. Indeed if say $\pi_i(\hat{S}_A) \cap \pi_i(\hat{S}_B) \neq \emptyset$, then there would be some $a \in S_A$, $b \in S_B$ with $\pi_i(\hat{a}) = \pi_i(\hat{b})$. We know that a, b separable, which implies that $\hat{a}, \hat{b}$ are separable (by Observation 3.23). But π_i separates $\bigcup \hat{S}$, which shows that $\pi(\hat{a}), \pi(\hat{b})$ are disjoint.

For each i, $\hat{g} \in \pi_i(\hat{S})$, let

$$ABC(\hat{g}) = \begin{cases} A \text{ if } \hat{g} \in \pi_i(\hat{S}_A) \\ B \text{ if } \hat{g} \in \pi_i(\hat{S}_B) \\ C \text{ if } \hat{g} \in \pi_i(\hat{S}_C) \\ C \text{ otherwise} \end{cases}.$$

Note that this is well defined by the previous paragraph. For each $i, \hat{g} \in \pi_i(S)$, let $E_i^{\hat{g}}$ be the event"$\hat{g} \subseteq R_{ABC(\hat{g})}$". Note that we have $P(E_i^{\hat{g}}) = p$. Note that $E_i^{\hat{g}}$, $E_j^{\hat{h}}$ are independent for $\hat{g} \neq \hat{h}$ ($E_i^{\hat{g}}$ depends only on the coordinate $\hat{g}$, $E_j^{\hat{h}}$ depends only on $\hat{h}$). Let $E_i = \bigcap_{\hat{g} \in \pi_i(\hat{S})} E_i^{\hat{g}}$. We have $\mathbb{P}(E_i) = \prod_{\hat{g} \in \pi_i(\hat{S})} \mathbb{P}(E_i^{\hat{g}}) \ge p^{|S|}$. For all $i = 1, \dots, m$, the events E_i are independent (since E_i, E_j depend on the coordinates in $\pi_i(\bigcup \hat{S})$, $\pi_j(\bigcup \hat{S})$ respectively. These are disjoint for $i \neq j$). By linearity of expectation, the expected number of indices i for which E_i occurs is at least $p^{|S|}m$. By Chernoff's Bound, there are at least $p^{|S|}m/2$ indices for which E_i occurs with probability

$$\ge 1 - 2e^{\frac{1}{6}p^{|S|}m} \ge 1 - 2e^{\frac{1}{6}(n^{-1/700})^{600}n/10^{3000}} \ge 1 - 2e^{-n^{1/9}/10^{6000}}.$$

Since $|U| \le p^{800}n/10^{4000} \le p^{600}n/(2 \cdot 10^{3000}) = p^{|S|}m/2$, there is at least one such index with $\pi_i(S) \cap U = \emptyset$. This projection π_i satisfies the lemma.

To get the lemma for all possible families $\{k, S, S_A, S_B, S_C\}$, notice that there are $o(e^{n^{1/7}/36{,}000})$ such families. Indeed, there are 200 choices for k and for each k, there are $\le 201(400n)^{201}$ length ≤ 200 elements $w \in G * F_k$. There are $\le (201(400n)^{201})^{600}$ sets of ≤ 600 such words. Hence, there are $\le ((201(400 \times n)^{201})^{600})^4 = o(e^{n^{1/37}/36{,}000})$ families of 4-tuples of such subsets. So we can take a union bound over all such families. □

Corollary 3.33 *Let $p \geq 3n^{-1/1400}$. Let R be random set which is either p-random or symmetric p-random. With high probability, the following holds*:

Let $k \leq 200$, $S \subseteq G * F_k$ a set of ≤ 600 elements of length ≤ 200 which are each linear in at least one variable, and $U \subseteq G$ with $|U| \leq p^{1600}n/10^{4002}$. Then there is a projection $\pi : G * F_k \to G$ which separates S and has $\pi(S) \subseteq R \setminus U$.

Proof Using Lemma 3.16, we can choose random sets $R_A, R_B, R_C \subseteq R$ such that the joint distribution on R_A, R_B, R_C is that of disjoint symmetric $p^2/9$-random sets. Now, the lemma follows from Lemma 3.32. □

We end with a simple application of the above lemma for later use.

Lemma 3.34 *Let $p \geq n^{-1/700}$. Let G be a group R a symmetric p-random subset of G. With high probability, the following holds.*

For any generic $x_\phi \in G$ and $U \subseteq V(H_G)$ with $|U| \leq p^{800}n/10^{4001}$, there are distinct and ϕ-generic $x, x' \in R \setminus U$ with $xx' = x_\phi$.

Proof With high probability, Lemma 3.32 applies. Let x_ϕ, U be as in the lemma. Add all non-ϕ-generic elements to U in order to get a set U' with $|U'| \leq p^{800}n/10^{4000}$. Let $x_\phi \in G$ and consider the set $\{v_1, x_\phi v_1^{-1}\} \subseteq G * F_1$. Note that $v_1, x_\phi v_1^{-1}$ are separable (by part (b) of the definition), and so Lemma 3.32 gives a projection $\pi : G * F_1 \to G$ with $\pi(v_1)$, $\pi(x_\phi v_1^{-1})$ distinct and contained in $R \setminus U'$ setting $x = \pi(x_\phi v_1^{-1})$, $x' = \pi(v_1)$ gives the lemma. □

4 The main theorem and its variants

This section is devoted to stating the main technical result of the paper, and collecting various consequences thereof (including Theorem 1.1) which will be more convenient to use for various applications we give. For the applications, we need variants of Theorem 1.1, where the probability distributions on R^1, R^2, R^3 are different from the one given (e.g. it is sometimes useful to sample R^1, R^2, R^3 disjointly rather than independently). We begin by stating a more technical version of Theorem 1.1 which covers all the different distributions of these sets that we might need. The following definition is the most general case.

Definition 4.1 Let G be a group and let R^1, R^2, R^3 be random subsets of G. We say that R^1, R^2, R^3 are q-**slightly-independent**, if there are random subsets $Q^1 \subseteq R^1$, $Q^2 \subseteq R^2$, $Q^3 \subseteq R^3$ such that the joint distribution on Q^1, Q^2, Q^3 is that of disjoint symmetric, q-random subsets of G.

Note that q-slightly-independent sets R^1 and R^2 do not necessarily have the same size (even in expectation). The following observation about this definition is useful.

Observation 4.2 *If R^1, R^2, R^3 are q-slightly-independent, then so are R^1, $(R^2)^{-1}$, R^3.*

Proof We have $Q^1 \subseteq R^1$, $Q^2 \subseteq R^2$, $Q^3 \subseteq R^3$ such that the joint distribution on Q^1, Q^2, Q^3 is that of disjoint, symmetric, q-random subsets of G. Note that since Q^2 is symmetric, $Q^2 = (Q^2)^{-1}$. Also, $(Q^2)^{-1} \subseteq (R^2)^{-1}$. Thus, Q^1, Q^2, Q^3 witness R^1, $(R^2)^{-1}$, R^3 being q-slightly-independent. □

The following is the strongest version of the main theorem that we prove in this paper. It is proved in Sect. 5.

Theorem 4.3 *Let $q \geq n^{-1/10^{101}}$. Let G be a group of order n. Let R^1, R^2, $R^3 \subseteq G$ be p-random, q-slightly-independent subsets. 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.

- $|(R^1_A \cup R^2_B \cup R^3_C) \triangle (X \cup Y \cup Z)| \leq q^{10^{17}} n/\log(n)^{10^{17}}$
- $\sum X + \sum Y + \sum Z = 0$ *(in* G^{ab}*)*
- $e_G \notin X \cup Y \cup Z$

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

We now state and prove a number of consequences of this theorem, where the distributions on R^1, R^2, R^3 are more natural. Firstly, the following theorem easily implies Theorem 1.1 (we will prove this formally later on in the section).

Theorem 4.4 *Let $p \geq n^{-1/10^{102}}$. Let G be a group of order n. Let R^1, R^2, $R^3 \subseteq G$ be p-random subsets, sampled independently. 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.

- $|(R^1_A \cup R^2_B \cup R^3_C) \triangle (X \cup Y \cup Z)| \leq p^{10^{18}} n/\log(n)^{10^{18}}$
- $\sum X + \sum Y + \sum Z = 0$ *(in* G^{ab}*)*

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

When R^1, R^2, R^3 are sampled disjointly, then the statement of the theorem needs to change slightly. In this case, when the group is $\mathbb{Z}_2^k$, then it is impossible to cover e with a hyperedge contained in $R^1 \cup R^2 \cup R^3$. Thus to get a matching, we need to additionally have the condition "$e_G \notin X \cup Y \cup Z$ if $G \cong \mathbb{Z}_2^k$" (Proposition 3.12 shows that this is the only local obstruction of this type).

Theorem 4.5 *Let $p \geq n^{-1/10^{102}}$. Let G be a group of order n. Let R^1, R^2, $R^3 \subseteq G$ be p-random disjoint subsets. 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.

- $|(R^1_A \cup R^2_B \cup R^3_C) \triangle (X \cup Y \cup Z)| \leq p^{10^{18}} n/\log(n)^{10^{18}}$
- $\sum X + \sum Y + \sum Z = 0$ *(in* G^{ab}*)*
- $e_G \notin X \cup Y \cup Z$ *if* $G \cong \mathbb{Z}_2^k$.

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

Finally, we have two versions of the theorem which are intermediate between the previous two. First one which asks R^1, R^2 to be sampled disjointly, and R^3 to be sampled independently.

Theorem 4.6 *Let $p \geq n^{-1/10^{102}}$. Let G be a 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 subsets of G_A, G_B, and G_C be equal sized subsets satisfying the following properties.

- $|(R^1_A \cup R^2_B \cup R^3_C)\triangle(X \cup Y \cup Z)| \leq p^{10^{18}} n/\log(n)^{10^{18}}$
- $\sum X + \sum Y + \sum Z = 0$ *(in the abelianization of G)*
- *If $G \cong (\mathbb{Z}_2)^k$ for some k, suppose that $e \notin Z$.*

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

The next theorem is almost the same as the previous one, with the difference that we sample $(R^1)^{-1}$, R^2 disjointly (as opposed to R^1, R^2).

Theorem 4.7 *Let $p \geq n^{-1/10^{102}}$. Let G be a group of order n. Let $(R^1)^{-1}, R^2 \subseteq G$ 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 subsets of G_A, G_B, and G_C be equal sized subsets satisfying the following properties.

- $|(R^1_A \cup R^2_B \cup R^3_C)\triangle(X \cup Y \cup Z)| \leq p^{10^{18}} n/\log(n)^{10^{18}}$
- $\sum X + \sum Y + \sum Z = 0$ *(in the abelianization of G)*
- *If $G \cong (\mathbb{Z}_2)^k$ for some k, suppose that $e \notin Z$.*

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

These theorems all have almost the same proof. The idea is to first show that the distribution of R_1, R_2, R_3 is q-slightly-independent (and so Theorem 4.3) applies. When the sets X, Y, Z don't contain the identity, then this already gives what we want. When X, Y, Z do contain the identity, then we first find a small matching covering all copies of the identity in X, Y, Z, and then find another matching covering all remaining vertices using Theorem 4.3.

Proof of Theorem 4.5 Note that R^1, R^2, R^3 are q-slightly-independent for $q = p^2/9$. To see this, consider disjoint $1/3$-random sets $T^1, T^2, T^3 \subseteq G$, chosen independently of R^1, R^2, R^3. We have that $R^1 \cap T^1$, $R^2 \cap T^2$, $R^3 \cap T^3$ are disjoint $p/3$-random subsets of G. Indeed for every $g \in G$, we have $\mathbb{P}(g \in R^i \cap T^i) = \mathbb{P}(g \in R^i)\mathbb{P}(g \in T^i) = p/3$, and $\mathbb{P}(g \in R^i \cap T^i \cap R^j \cap T^j) \leq \mathbb{P}(g \in T^i \cap T^j) = 0$. These imply $\mathbb{P}(g \notin \bigcup_{i=1}^3 R^i \cap T^i) = 1 - 3\alpha p$. Also for different g, h, their locations are independent of each other (since there was no correlations between distinct g, h, in any of R^1, R^2, R^3, T^1, T^2, T^3), giving that $R^1 \cap T^1$, $R^2 \cap T^2$, $R^3 \cap T^3$ are disjoint $p/3$-random subsets of G. Use Lemma 3.16 to pick disjoint, symmetric, q-random subsets $Q^1 \subseteq R^1 \cap T^1$, $Q^2 \subseteq R^2 \cap T^2$, $Q^3 \subseteq R^3 \cap T^3$. Now Q^1, Q^2, Q^3 demonstrate R^1, R^2, R^3 being q-slightly-independent.

So Theorem 4.3 applies to R^1, R^2, R^3. We also have the following property if $G \neq \mathbb{Z}_2^k$:

P: For each $i, j \in \{1, 2, 3\}$, $i \neq j$, $R^i \times R^j$ contains $\geq p^2 n/1000$ pairs of the form (x, x^{-1}) where for each such pair of pairs $\{x, x^{-1}\} \cap \{y, y^{-1}\} = \emptyset$ (this follows by Chernoff's bound).

Now consider sets X, Y, Z as in the theorems. Use (P) to pick distinct elements g_A, g_B, g_C with $g_A \in R^2$, $g_A^{-1} \in R^3$, $g_B \in R^1$, $g_B^{-1} \in R^3$, $g_C \in R^1$, $g_C^{-1} \in R^2$. Noting that we have $p^2 n/1000$ choices for each g_A, g_B, g_C, we can choose them to have $\{g_A, g_A^{-1}\}$, $\{g_B, g_B^{-1}\}$, $\{g_C, g_C^{-1}\}$ disjoint from each other and from $R^1 \setminus X$, $R^2 \setminus Y$, $R^3 \setminus Z$. The result is that the three edges $f_A := (e_A, g_A, g_A^{-1})$, $f_B := (g_B, e_B, g_B^{-1})$, $f_C := (g_C, g_C^{-1}, e_C)$ form a matching with all vertices, other than possibly e_B, e_B, e_C contained in $X \cup Y \cup Z$. Let $N = \{f_i : e_i \in X \cup Y \cup Z\}$ to get a matching contained in $X \cup Y \cup Z$ covering all copies of the identity in $X \cup Y \cup Z$.

Let $X' = X \setminus N$, $Y' = Y \setminus N$, $Z' = Z \setminus N$, noting that these have the same size and have $\sum X' + \sum Y' + \sum Z' = 0$ in G^{ab} (due to N being a matching). Thus the property of Theorem 4.3 applies to give a perfect matching M in $H_G[X', Y', Z']$. Now $M \cup N$ satisfies the theorem. □

Next we show how to modify the above proof to obtain Theorems 4.4, 4.6, 4.7.

Proof of Theorems 4.4, 4.6, 4.7 First notice that in all three theorems, we have that R^1, R^2, R^3 are q-slightly-independent. In Theorems 4.4, 4.6 this is exactly the first paragraph of the proof of Theorem 4.5. For Theorem 4.7, that paragraph shows that $(R^1)^{-1}$, R^2, R^3 are q-slightly-independent. But, then by Observation 4.2, we have that R^1, R^2, R^3 are q-slightly-independent.

Next note that property (P) holds in the following cases:

- Theorem 4.4: here property (P) always holds.
- Theorem 4.6: here property (P) always holds for $(i, j) = (2, 3)$ and $(i, j) = (1, 3)$. For $(i, j) = (1, 2)$ property (P) holds when $G \neq \mathbb{Z}_2^k$.
- Theorem 4.7: here property (P) always holds for $(i, j) = (2, 3)$ and $(i, j) = (1, 3)$. For $(i, j) = (1, 2)$ property (P) holds when $G \neq \mathbb{Z}_2^k$.

The rest of the proofs are the same as in Theorem 4.5 — property (P) produces a matching of size ≤ 3 covering all copies of the identity in $X \cup Y \cup Z$, and then the property of Theorem 4.3 gives a matching covering the rest of $X \cup Y \cup Z$. □

Finally we show how to derive Theorem 1.1 as stated in the introduction.

Proof of Theorem 1.1 via Theorem 4.4 Let $R^1, R^2, R^3 \subseteq G$ be p-random subsets, independently sampled. Observe that $R^1, R^2, (R^3)^{-1} \subseteq G$ are also p-random subsets, independently sampled, so with high probability Theorem 4.4 applies. Let $X, Y, Z \subseteq G$ be subsets with the properties as in the statement of Theorem 1.1. Then, the sets $X, Y, Z^{-1} \subseteq G$ clearly satisfy the two properties required by Theorem 4.4 with respect to $R^1, R^2, (R^3)^{-1} \subseteq G$. Thus, $H_G[X, Y, Z^{-1}]$ contains a perfect matching, say M. Define the bijection $\phi \colon X \to Y$ so that x maps to the unique element y of Y such that x and y are contained in an edge together in M. As for each edge $(x, \phi(x), z)$ of M, $x\phi(x)z = e$, $x \mapsto x\phi(x)$ is a bijection $X \to (Z^{-1})^{-1} = Z$, as desired. □

4.1 Complete mappings and orthomorphisms

We conclude this section with a version of the main theorem that fits better with the results proved in Sect. 6.2. Given a triple of subsets of a group G as (X, Y, Z), a **complete mapping** is a bijection $\phi\colon X \to Y$ such that the induced map from X to Z via $x \to x\phi(x)$ is also a bijection, whereas an **orthomorphism** is a bijection $\phi\colon X \to Y$ such that the induced map from X to Z via $x \to x^{-1}\phi(x)$ is also a bijection.

Observation 4.8 *Let X, Y, Z be subsets of a group G. Then,*

- *(X, Y, Z) admits a complete mapping if and only if $H_G[X, Y, Z^{-1}]$ has a perfect matching.*
- *(X, Y, Z) admits an orthomorphism if and only if $H_G[X^{-1}, Y, Z^{-1}]$ has a perfect matching.*

Proof The proof is routine and we refer the reader to an earlier version of the paper available at arXiv:2204.09666v2 for a proof. □

Theorem 4.9 *Let $p \ge n^{-1/10^{102}}$. Let G be a 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 satisfying the following properties.

- $|(R^1_A \cup R^2_B \cup R^3_C)\triangle(X \cup Y \cup Z)| \le p^{10^{18}} n/\log(n)^{10^{18}}$
- *One of the following identities holds in the abelianization of G.*
  C. $\sum X + \sum Y = \sum Z$
  O. $\sum Y - \sum X = \sum Z$
- *If $G = (\mathbb{Z}_2)^k$ for some k, then $e \notin Z$.*

*Then, if **C** holds, (X, Y, Z) admits a complete mapping, and if **O** holds, (X, Y, Z) admits an orthomorphism.*

Proof We refer the reader to an earlier version of the paper available at arXiv:2204.09666v2 for a short and routine proof. □

5 Proof of the main theorem

In this section, we prove the main result of the paper, Theorem 4.3. We begin by clarifying that for every (sufficiently large) group G, we fix $a_\phi, b_\phi, c_\phi \in G$ with properties as in Sect. 3.4. Definitions such as ϕ-generic, pair, and coset-paired are with respect to these fixed choices of $a_\phi, b_\phi, c_\phi \in G$ given by Lemma 3.14.

5.1 Absorbers

In this section we give constructions of absorbers i.e. subsets $R \subseteq V(H_G)$ which can be extended into a matching in several different ways. Lemma 5.24 is the main result of this section, and the only result we need for the rest of the paper. The following definition precisely describes what we will be looking for.

Definition 5.1 Let $\mathcal{F} = \{S_1, \dots, S_t\}$ be a family of subsets of $V(H)$ for a hypergraph H. We say that a set of vertices R m**-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 $R \cup \bigcup_{S_i \in \mathcal{F}'} S_i$.

It will be convenient to note that when $t = 2$ and $m = 1$, then the above definition is equivalent to "Let X, Y be sets of vertices in a hypergraph H. We say that a set of vertices R 1**-absorbs** $\{X, Y\}$ if there are hypergraph matchings R^-, R^+ whose vertex sets are exactly $V(R^-) = R \cup X$ and $V(R^+) = R \cup Y$."

The following lemma shows how the parameter h changes when we pass to a subfamily of $\mathcal{F}$.

Lemma 5.2 *Let $\mathcal{F}$ be a family of disjoint subsets of $V(H)$ and $\mathcal{F}' \subseteq F$ a subfamily with $|\mathcal{F} \setminus \mathcal{F}'| = t$. If R h-absorbs $\mathcal{F}$ then*

1. *R h-absorbs $\mathcal{F}'$.*
2. *$R \cup \bigcup_{S \in \mathcal{F} \setminus \mathcal{F}'} S$ $(h - t)$-absorbs $\mathcal{F}'$.*

Proof Part (1) is immediate from the definition of absorbing. For part (2), notice that for any subfamily $\mathcal{F}'' \subseteq \mathcal{F}'$ of size $h - t$, the subfamily $\mathcal{F}'' \cup (\mathcal{F} \setminus \mathcal{F}')$ has size h. Therefore there is a matching with vertex set $R \cup \bigcup_{S \in \mathcal{F}'' \cup (\mathcal{F} \setminus \mathcal{F}')} S = (R \cup \bigcup_{S \in \mathcal{F} \setminus \mathcal{F}'} S) \cup \bigcup_{S \in \mathcal{F}''} S$. □

We build larger absorbers from smaller ones. The following lemma allows us to take unions of 1-absorbers to get another 1-absorber.

Lemma 5.3 *Suppose that $\{R_1, \dots, R_t\}$, $\{X_1, \dots, X_t\}$, $\{Y_1, \dots, Y_t\}$ are three families of disjoint sets with $\bigcup R_i$ disjoint from $\bigcup (X_i \cup Y_i)$. Suppose that R_i 1-absorbs $\{X_i, Y_i\}$ for $i = 1, \dots, t$. Set $Z = (\bigcup X_i) \cap (\bigcup Y_i)$. Then, $\bigcup R_i \cup Z$ 1-absorbs $\{\bigcup X_i \setminus Z, \bigcup Y_i \setminus Z\}$.*

Proof By the definition of 1-absorbs we have matchings R_i^- and R_i^+ with vertices $V(R_i^-) = R_i \cup X_i$ and $V(R_i^+) = R_i \cup Y_i$. Note that $\bigcup R_i^-$ and $\bigcup R_i^+$ are matchings (since $\{R_1, \dots, R_t\}$, $\{X_1, \dots, X_t\}$, $\{Y_1, \dots, Y_t\}$ are families of disjoint sets with $\bigcup R_i$ disjoint from $\bigcup (X_i \cup Y_i)$). Also $V(\bigcup R_i^-) = \bigcup R_i \cup \bigcup X_i = (\bigcup R_i \cup Z) \cup (\bigcup X_i \setminus Z)$ and $V(\bigcup R_i^+) = \bigcup R_i \cup \bigcup Y_i = (\bigcup R_i \cup Z) \cup (\bigcup Y_i \setminus Z)$. Thus $\bigcup R_i^-$ and $\bigcup R_i^+$ are matchings satisfying the definition of "$\bigcup R_i \cup Z$ 1-absorbs $\{\bigcup X_i \setminus Z, \bigcup Y_i \setminus Z\}$". □

We state a technical consequence of the above lemma for later use.

Lemma 5.4 *Suppose that we have distinct vertices $a, b, c, d \in V(H_G)$, vertices $w_1, \dots, w_k \in V(H_G) \setminus \{a, b, c, d\}$ and disjoint sets $R_0, R_1, \dots, R_{k-1} \subseteq V(H_G) \setminus \{w_1, \dots, w_k, a, b, c, d\}$ with R_0 1-absorbing $\{\{a, w_1, b\}, \{a, w_k, b\}\}$ and R_i 1-absorbing $\{w_i, w_{i+1}\}$ for $i = 1, \dots, k - 1$. Then, there is a subset $R' \subseteq \bigcup_{i=0}^{k-1} R_i \cup \{w_1, \dots, w_k\}$ which 1-absorbs $\{\{a, b\}, \{c, d\}\}$.*

Proof Without loss of generality, we can assume that $w_1, \dots, w_k$ are distinct (by passing to a subset of $\{w_1, \dots, w_k\}$ of distinct vertices, and a corresponding subfamily of $\{R_1, \dots, R_{k-1}\}$). Now, the lemma follows from Lemma 5.3 with $\{R_0, \dots, R_k\}$,

$$\{X_0, \dots, X_k\} := \{\{a, w_1, b\}, \{w_2\}, \{w_3\}, \dots, \{w_k\}\},$$

and

$$\{Y_1, \dots, Y_t\} := \{\{c, w_k, d\}, \{w_1\}, \{w_2\}, \dots, \{w_{k-1}\}\}. \qquad \square$$

5.1.1 Constructing 1-absorbing sets

The following lemma shows that there are many edges through generic vertices in H_G.

Lemma 5.5 *Let $p \geq n^{-1/700}$. Let H_G be a multiplication hypergraph, R^1, R^2, R^3 disjoint, symmetric p-random subsets of G and set $R = R^1_A \cup R^2_B \cup R^3_C$. With high probability, the following holds:*

For any generic $v \in V(H_G)$ and $U \subseteq V(H_G)$ with $|U| \leq p^{800}n/10^{4000}$, there is an edge e of H_G passing through v and having the other two vertices in $R \setminus U$.

Proof With high probability, Lemma 3.32 applies. We'll just look at the case $v \in G_A$, the other two cases are symmetric. Fix some generic $v \in G_A$. Thinking of b as a free variable, consider the set $S = \{b, b^{-1}v^{-1}\} \subseteq G * F_1$. Note that this is a set of linear elements with $w = b$, $w' = b^{-1}v^{-1}$ separable (by (b) since v^{-1} is generic). Lemma 3.32 now implies what we want. $\square$

The following is the basic building block for the absorbers that we construct. It shows that we can 1-absorb sets of the form $\{[a, b]c, c\}$.

Lemma 5.6 *Let $p \geq n^{-1/700}$. Let H_G be a multiplication hypergraph, R^1, R^2, R^3 disjoint, symmetric p-random subsets of G and set $R = R^1_A \cup R^2_B \cup R^3_C$. With high probability, the following holds:*

For any $a, b, c \in G_C$ with c, $c^{-1}bab^{-1}a^{-1}$, $c^{-1}bab^{-1}$, $c^{-1}ba$, $c^{-1}b$ generic and $U \subseteq V(H_G)$ with $|U| \leq p^{800}n/10^{4010}$, there is a set $R' \subseteq R \setminus U$ of size ≤ 14 which 1-absorbs $\{[a, b]c, c\}$.

Proof With high probability, the properties of Lemmas 3.32 and 5.5 hold. Fix some $a, b, c \in G_C$ with

$$c, c^{-1}bab^{-1}a^{-1}, c^{-1}bab^{-1}, c^{-1}ba, c^{-1}b$$

all being generic and $U \subseteq V(H_G)$ with $|U| \leq p^{800}n/10^{4010}$. First suppose that $[a, b] = e$ i.e. that $[a, b]c = c$. From Lemma 5.5 we know that there is an edge $f = \{s, t, c\}$ of H_G with $s, t \in R \setminus U$. Then $R' = \{s, t\}$ satisfies the lemma (with both M^- and M^+ for the definition of "1-absorbs" equal to the single edge f).

Now suppose $[a, b] \neq e$. Notice that this implies that $a, b, bab^{-1}a^{-1}, bab^{-1}, ba$, $b^{-1}a^{-1}, ab^{-1}a^{-1}, ab^{-1} \neq e$ also. Thinking of x, y, z as free variables in $G * F_3$,

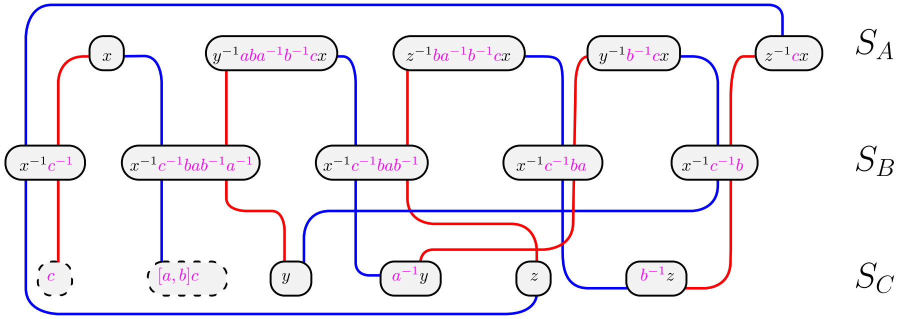

Fig. 2 The set $S \subseteq G * F_3$ in Lemma 5.6. Black letters x, y, z are free variables, while pink letters are elements of G. The two elements $[a, b]c$, c are not part of S (and are just pictured to show how S 1-absorbs $\{[a, b]c, c\}$) (Color figure online)

S_x :

	x	$x^{-1}c^{-1}$	$x^{-1}c^{-1}bab^{-1}a^{-1}$	$x^{-1}c^{-1}bab^{-1}$	$x^{-1}c^{-1}ba$	$x^{-1}c^{-1}b$
x		c^{-1}	$c^{-1}bab^{-1}a^{-1}$	$c^{-1}bab^{-1}$	$c^{-1}ba$	$c^{-1}b$
$x^{-1}c^{-1}$	c^{-1}		$bab^{-1}a^{-1}$	bab^{-1}	ba	b
$x^{-1}c^{-1}bab^{-1}a^{-1}$	$c^{-1}bab^{-1}a^{-1}$	$bab^{-1}a^{-1}$		a^{-1}	$b^{-1}a^{-1}$	$ab^{-1}a^{-1}$
$x^{-1}c^{-1}bab^{-1}$	$c^{-1}bab^{-1}$	bab^{-1}	a^{-1}		b^{-1}	ab^{-1}
$x^{-1}c^{-1}ba$	$c^{-1}ba$	ba	$b^{-1}a^{-1}$	b^{-1}		a
$x^{-1}c^{-1}b$	$c^{-1}b$	b	$ab^{-1}a^{-1}$	ab^{-1}	a	

$S_{x,y}$:

	$y^{-1}aba^{-1}b^{-1}cx$	$y^{-1}b^{-1}cx$
$y^{-1}aba^{-1}b^{-1}cx$		aba^{-1}
$y^{-1}b^{-1}cx$	aba^{-1}	

S_y :

	y	$a^{-1}y$
y		a^{-1}
$a^{-1}y$	a^{-1}	

$S_{x,z}$:

	$z^{-1}ba^{-1}b^{-1}cx$	$z^{-1}cx$
$z^{-1}ba^{-1}b^{-1}cx$		$ba^{-1}b^{-1}$
$z^{-1}cx$	$ba^{-1}b^{-1}$	

S_z :

	z	$b^{-1}z$
z		b^{-1}
$b^{-1}z$	b^{-1}	

Fig. 3 Proofs for weak/strong separability and linearity of all pairs $w, w' \in S$ in Lemma 5.6 to justify the application of Lemma 3.32. For strong separability, first we have partitioned S into five subsets $S = S_x \cup S_{x,y} \cup S_{x,z} \cup S_z \cup S_y$ based on which free variables appear in each $w \in S$ (as in Observation 3.24). By Observation 3.24, any w, w' in different subsets are strongly separable by part (a) of the definition. For each of the sets S_x, $S_{x,y}$, $S_{x,z}$, S_z, S_y we give a table explaining why the w, w' in that set are strongly/weakly-separable. Note that for words coming from different $S_A/S_B/S_C$ we need to show strong separability, but for words coming from the same part, weak separability suffices. Blue cells represent w, w' being strongly separable via part (b) of the definition, green cells represent w, w' being weakly separable via part (b'), and grey cells represent w, w' not being separable/weakly-separable. The group element inside each blue cell is a generic element g so that $w' \in \{gw, g^{-1}w, gw^{-1}, g^{-1}w^{-1}, wg, wg^{-1}, w^{-1}g, w^{-1}g^{-1}\}$ (thus checking (b) for w, w'). The group element inside the green cells is a non-identity element g so that $w = w'$ rearranges into $e = g$ (thus checking (b') for w, w'). Observe that green cells are used only between pairs of words coming from the same part of $S_A/S_B/S_C$, meaning that we have strong separation for pairs of words coming from different parts $S_A/S_B/S_C$, as needed. To see that every $s \in S$ is linear notice that every word pictured has no repetitions of black letters (Color figure online)

consider the set S given in Fig. 2, with partition $S = S_A \cup S_B \cup S_C$. Notice that all words in S are linear, all pairs of words weakly-separable, and S_A, S_B, S_C are pairwise strongly separable (see Fig. 3 for justification). Using Lemma 3.32, there is some projection $\pi : G * F_3 \to G$ which separates S and has $\pi(S) \subseteq R \setminus U$. Any such $\pi(S)$ 1-absorbs $\{[a, b]c, c\}$ (using the red/blue matchings in Fig. 2). □

The following lemma is identical to the previous one, except that it weakens the assumption on what elements are generic.

Lemma 5.7 *Let $p \geq n^{-1/700}$. Let H_G be a multiplication hypergraph, R^1, R^2, R^3 disjoint, symmetric p-random subsets of G and set $R = R^1_A \cup R^2_B \cup R^3_C$. With high probability, the following holds:*

For any $a, b, x \in G_B$ with x, $x[a, b]$ generic and $U \subseteq V(H_G)$ with $|U| \leq p^{800}n/10^{4020}$, there is a set $R' \subseteq R \setminus U$ of size ≤ 16 which 1-absorbs $\{x[a, b], x\}$.

Proof With high probability, the properties of Lemmas 3.32, 5.5 and 5.6 hold. Fix some $a, b, x \in G_B$ with x, $x[a, b]$ generic and $U \subseteq V(H_G)$ with $|U| \leq p^{800}n/10^{4020}$. As in Lemma 5.6, the conclusion trivially follows from Lemma 5.5 if $[a, b] = e$, so assume that this doesn't happen. With v the free variable in $G * F_1$, consider the sets of words $T := \{v, [b, a]x^{-1}v^{-1}, x^{-1}v^{-1}\}$ and $S = \{x^{-1}v^{-1}, vxaba^{-1}b^{-1}, vxaba^{-1}, vxab, vxa\}$. Define $T_A = \{v\}$, $T_B = \emptyset$, $T_C = \{[b, a]x^{-1}v^{-1}, x^{-1}v^{-1}\}$ to get a partition of T. It is easy to check that all $w \in T \cup S$ are linear (since v appears precisely once in each $w \in T \cup S$), that T_A, T_C are strongly separable (by part (b) of the definition, using that x, $x[a, b]$ are generic), and that $[b, a]x^{-1}v^{-1}$, $x^{-1}v^{-1}$ are weakly separable (since $[b, a]x^{-1}v^{-1} = x^{-1}v^{-1}$ rearranges into $[b, a] = e$). By Lemma 3.32, there is a projection π which separates $T \cup S$ and has $\pi(T \cup S) \subseteq R \setminus (U \cup N(G))$. Since T_A, T_B, T_C are pairwise strongly separable, we additionally get $\pi(T_A) \subseteq R^1_A$, $\pi(T_B) \subseteq R^2_B$, $\pi(T_C) \subseteq R^3_C$. Combining this with the fact that all vertices in $\pi(T)$ are distinct (which comes from all pairs of words in T being weakly separable, and so $\pi(t) \neq \pi(t')$ for distingt $t, t' \in T$), we have that $e^- = (\pi(v), x[a, b], [b, a]x^{-1}\pi(v)^{-1})$ and $e^+ = (\pi(v), x, x^{-1}\pi(v)^{-1})$ are edges of H_G contained in $R \cup \{x, x[a, b]\}$.

Using Lemma 5.6 with $a' = b$, $b' = a$, $c = x^{-1}\pi(v^{-1})$ we find a set Q disjoint from $U \cup \pi(T)$ which 1-absorbs $\{[b, a]x^{-1}\pi(v)^{-1}, x^{-1}\pi(v)^{-1}\}$ (all the required elements are generic for that lemma as a consequence of $\pi(S) \cap N(G) = \emptyset$). Now $Q \cup \pi(T)$ 1-absorbs $\{x[a, b], x\}$. This can be seen directly or by first noticing that $\{\pi(v)\}$ 1-absorbs $\{\{x[a, b], [b, a]x^{-1}\pi(v)^{-1}\}, \{x, x^{-1}\pi(v)^{-1}\}\}$ (the single-edge matchings e^-, e^+ witness this). Then Lemma 5.3 shows that $Q \cup \pi(T)$ 1-absorbs $\{x[a, b], x\}$. □

For two sets $X, Y \subseteq G$ and $s \in G$, use XsY to denote $\{xsy : x \in X, y \in Y\}$. The following lemma lets us 1-absorb a pair of size 3 sets.

Lemma 5.8 *Let $p \geq n^{-1/700}$. Let H_G be a multiplication hypergraph, R^1, R^2, R^3 disjoint, symmetric p-random subsets of G and set $R = R^1_A \cup R^2_B \cup R^3_C$. With high probability, the following holds:*

For any distinct, generic $a, b, c, d \in G_B$ and $X, Y, U \subseteq V(H_G)$ with $|U| \leq p^{800}n/10^{4000}$, $|X|$, $|Y| \leq 5$, there is a $R' \subseteq R \setminus U$ of size 6 and s with $XsY \subseteq R \setminus (U \cup R')$ such that R' 1-absorbs $\{\{a, dsc, b\}, \{c, bsa, d\}\}$.

Proof With high probability, the property of Lemma 3.32 holds. Suppose we have generic $a, b, c, d \in G_B$ and $X, Y, U \subseteq V(H_G)$ with $|U| \leq p^{800}n/10^{4000}$, $|X|$, $|Y| \leq 5$. Let $X' = X \cup \{b, d\}$, $Y' = Y \cup \{a, c\}$. For free variables u, w, consider the sets of words $S = \{u, a^{-1}u^{-1}, c^{-1}u^{-1}, w, b^{-1}w^{-1}, d^{-1}w^{-1}\}$ with $S_A = \{u, b^{-1}w^{-1}, d^{-1}w^{-1}\}$, $S_B = \emptyset$, $S_C = \{w, a^{-1}u^{-1}, c^{-1}u^{-1}\}$. Notice that all words

S_u :

	u	$a^{-1}u^{-1}$	$c^{-1}u^{-1}$
u		a^{-1}	c^{-1}
$a^{-1}u^{-1}$	a^{-1}		ac^{-1}
$c^{-1}u^{-1}$	c^{-1}	ac^{-1}	

S_w :

	w	$w^{-1}b^{-1}$	$w^{-1}d^{-1}$
w		b^{-1}	d^{-1}
$w^{-1}b^{-1}$	b^{-1}		bd^{-1}
$w^{-1}d^{-1}$	d^{-1}	bd^{-1}	

Fig. 4 Justification for linearity and separability of pairs (w, w') in S in Lemma 5.8. For separability, first split S into $S_u = \{u, a^{-1}u^{-1}, c^{-1}u^{-1}\}$, $S_w = \{w, b^{-1}w^{-1}, d^{-1}w^{-1}\}$ based on which free variables appear in the elements (as in Observation 3.24). By Observation 3.24, pairs (w, w') with w, w' in different sets S_u/S_w fall under part (a) of the definition of strongly separable, so it remains to check pairs inside S_u and S_w. Justification for this is given in the two tables above (with the same conventions as in Fig. 3, in particular, green cells are used only between pairs of vertices coming from the same part $S_A/S_B/S_C$). Relevant elements are generic/non-identity as a consequence of a, b, c, d being distinct and generic). To see that each $w \in S$ is linear, note that there are no repetitions of black letters in each w (and each $w \in S$ contains at least one black letter) (Color figure online)

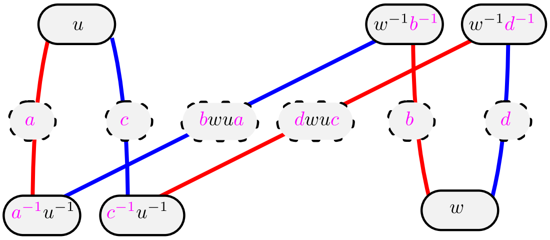

Fig. 5 The set of words $S \subseteq G * F_2$ for Lemma 5.8. First, second, and third row of words represent S_A, S_B, and S_C, respectively. Black letters u, w are free variables of F_2, while pink letters are elements of G. The elements a, b, c, d, $bwua$, $dwuc$ are not part of S (and are pictured just to demonstrate how the absorption works) (Color figure online)

in $S \cup (X'wuY')$ are linear, all pairs of words in S are weakly separable (see Fig. 4), and S_A, S_B, S_C are pairwise strongly separable. Also the pair of sets $(S, X'wuY')$ is strongly separable (using part (a) of the definition of separable). By Lemma 3.32, there is some projection π separating $S \cup (X'wuY')$ and having $S \cup (X'wuY') \subseteq R \setminus U$. For any such projection, $R' = \pi(S)$ and $s := \pi(wu)$ satisfy the lemma (with the red/blue matchings in Fig. 5 satisfying the definition of R' 1-absorbing $\{\{a, dsc, b\}, \{c, bsa, d\}\}$. Note that this works regardless of whether $dwuc = bwua$ or not). □

The following technical lemma allows us to 1-absorb certain pairs of sets of size 2.

Lemma 5.9 *Let $p \geq n^{-1/700}$. Let H_G be a multiplication hypergraph, R^1, R^2, R^3 disjoint, symmetric p-random subsets of G and set $R = R^1_A \cup R^2_B \cup R^3_C$. With high probability, the following holds:*

For any $g, x, y, a, b \in G_B$ with yg, x, y, $x[a, b]g$ distinct, generic and $U \subseteq V(H_G)$ with $|U| \leq p^{800}n/10^{4030}$, there is a set $R' \subseteq R \setminus U$ of size ≤ 60 which 1-absorbs $\{\{yg, x\}, \{y, x[a, b]g\}\}$.

Proof With high probability, the properties of Lemmas 5.7 and 5.8 hold. Let g, x, y, a, b and U be as in the lemma. Fix $a' = yg$, $b' = x$, $c' = y$, $d' =$

$x[a,b]g$ and note these are distinct and generic by assumption. Define $X = \{x, x[a,b]g, x[a,b], x[a,b]yg, e\}$, $Y = \{yg, y, g[a,b], yg[a,b], e\}$. By Lemma 5.8, we get R' and s with $R_0 \cup (XsY) \subseteq V(H_G) \setminus (U \cup N(G) \cup \{a', b', c', d'\})$ with R_0 1-absorbing $\{\{a', d'sc', b'\}, \{c', b'sa', d'\}\}$.

Define $w_1 = d'sc' = x[a,b]gsy = x[a,b]ygs[y, gs]$, $w_2 = x[a,b]ygs = x[a, b]syg[s, yg]$, $w_3 = x[a,b]syg = xsyg[a,b][syg,[a,b]]$, $w_4 = xsyg[a,b]$, and $w_4 = b'sa' = xsyg$, noting that all of these are in XsW and so are generic and disjoint from $R_0 \cup \{a', b', c', d'\}$. So we can use Lemma 5.7 to find disjoint sets $R_1, R_2, R_3, R_4 \subseteq R \setminus (R' \cup U)$ with R_i 1-absorbing $\{w_i, w_{i+1}\}$. We also have R_0 1-absorbing $\{\{a', w_1, b'\}, \{c', w_5, d'\}\}$. By Lemma 5.4, there is a subset $R' \subseteq \bigcup_{i=0}^{4} R_i \cup \{w_1, w_2, w_3, w_4, w_5\}$ which 1-absorbs $\{\{a', b'\}, \{c', d'\}\}$. □

The following lemma gives us a collection of distinct, generic elements.

Lemma 5.10 *Let $p \geq n^{-1/700}$ and $t \leq p^{800}n/10^{4020}$. Let G be a group and R a symmetric p-random subset. With high probability, the following holds*:

Let $g_1, \ldots, g_t \neq e$ be distinct and $U \subseteq G$ with $|U| \leq p^{800}n/10^{4020}$. There are $y_1, \ldots, y_{t-1} \in G$ such that the elements $y_1, y_2, \ldots, y_{t-1}, y_1g_2, y_2g_3, \ldots, y_{t-1}g_t$ are distinct, *generic elements in $R \setminus U$.*

Proof Lemma 3.32 applies with $R_A = R$, and R_B, R_C arbitrary disjoint symmetric p-random subsets (which won't be used in the proof). Let $g_1, \ldots, g_t \neq e$ be distinct and $U \subseteq G$ with $|U| \leq p^{800}n/10^{4020}$. Thinking of y as the free variable in $G * F_1$, let $S_i = \{y, yg_i\}$. Note all $w, w' \in S_i$ are linear and strongly separable (by part (b) since g_i is generic). For $i = 1, \ldots, t$, use Lemma 3.32 to pick projections $\pi_1, \ldots, \pi_i$ such that $\pi_i(S_i)$ is separated and disjoint from $U' := U \cup N(G) \cup \bigcup_{j<i} \pi_j(S_j)$ (noting that $|U'| \leq |U| + |N(G)| + 2t \leq 3p^{800}n/10^{4020} + 10^{9000} \leq n/10^{4010}$). Now $y_1 = \pi_1(y), \ldots, y_{t-1} = \pi_{t-1}(y)$ satisfy the lemma. □

The following theorem of Gallagher is key to our approach and shows that elements of the commutator subgroup can be written as a product of a small number of commutators. Its proof uses character theory.

Theorem 5.11 (Gallagher, [29]) *Let G be a group. Any $g \in G'$ can be written as $g = \prod_{i=1}^{t}[a_i, b_i]$ for some $a_i, b_i \in G$ and $t \leq \log_4 |G'| \leq 10 \log n$.*

We now prove one of the main lemmas in this section. It strengthens several earlier lemmas and shows that we can 1-absorb any pair of elements as long as they are in the same coset of G'.

Lemma 5.12 *Let $p \geq n^{-1/700}$. Let H_G be a multiplication hypergraph, R^1, R^2, R^3 disjoint*, *symmetric p-random subsets of G and set $R = R^1_A \cup R^2_B \cup R^3_C$. With high probability*, *the following holds*:

For any generic $h, k \in G_A, G_B$, or G_C with $[h] = [k]$ and $U \subseteq V(H_G)$ with $|U| \leq p^{800}n/10^{4040}$, there is a set $R' \subseteq R \setminus U$ of size $\leq 400 \log n$ which 1-absorbs $\{h, k\}$.

Proof With high probability, the properties of Lemmas 5.5, 5.7, 5.9, and 5.10 hold. We just prove the lemma for $h, k \in G_B$, the other two cases follow by symmetry. Let $h, k \in G_B$ be generic with $[h] = [k]$ and $U \subseteq V(H_G)$ with $|U| \le p^{800}n/10^{4040}$. As in Lemma 5.6, the conclusion trivially follows from Lemma 5.5 if $h = k$, so assume $h \neq k$. Write $h = kg$ for some $g \in G'$. By Theorem 5.11, we have $g = \prod_{i=1}^{t}[a_i, b_i]$ for some $a_i, b_i \in G$ and $t \le 10\log n$. Assume this product is as short as possible i.e. that t is minimal. For each $s = 0, \dots, t-1$, define $g_s = \prod_{i=s}^{t}[a_i, b_i]$. By minimality of t, we have $[a_i, b_i] \neq e$ and $g_i \neq e$ for all i. Set $y_0 = k$ and note that y_0, y_0g_1 are distinct, generic (since k, h are distinct, generic).

Note that $t \le 10\log n \le n^{1/8}/10^{4010} \le p^{800}n/10^{4010}$ (using that n is large which follows from "with high probability"). Use Lemma 5.10 with $U' = U \cup \{y_0, y_0g_1\}$ to get elements $y_1, \dots, y_{t-1}$, noting that now

$$y_0, y_1, \dots, y_{t-1}, y_0g_1, y_1g_2, \dots, y_{t-1}g_t$$

are distinct, generic. $U' = U \cup \{y_i, y_ig_{i+1} : i = 1, \dots, t-1\}$. Using Lemma 5.9 with $y = y_i$, $x = y_{i-1}$, $g = g_{i+1}$, $a = a_i$, $b = b_i$ for all i (the conditions "y_ig_{i+1}, y_{i-1}, y_i, $y_{i-1}[a_i, b_i]g_{i+1}$ distinct, generic" coming from Lemma 5.10), gives a set R_i which 1-absorbs $\{\{y_ig_{i+1}, y_{i-1}\}, \{y_i, y_{i-1}[a_i, b_i]g_{i+1}\}\} = \{\{y_ig_{i+1}, y_{i-1}\}, \{y_i, y_{i-1}g_i\}\}$. Using Lemma 5.7 with $x = y_{t-1}$, $a = a_t$, $b = b_t$ (condition "y_{t-1}, $y_{t-1}[a_t, b_t]$ generic" coming from Lemma 5.10), gives a set R_t which 1-absorbs $\{y_{t-1}[a_t, b_t], y_{t-1}\} = \{y_{t-1}g_t, y_{t-1}\}$. By enlarging the set U during the application of these lemmas, we can assume that $R_1, \dots, R_{t-1}, R_t, \{y_i, y_ig_{i+1} : i = 1, \dots, t-1\}$ are all disjoint (there's space to do this because $|R_1 \cup \dots \cup R_{i-1}| \le 400\log n < 10^{800}n^{1/8}/10^{4020} \le p^{800}n/10^{4030}$).

Set $R = R_1 \cup \dots \cup R_{t-1} \cup R_t \cup \{y_i, y_ig_{i+1} : i = 1, \dots, t-1\}$. Set $X_i = \{y_ig_{i+1}, y_{i-1}\}$, $Y_i = \{y_i, y_{i-1}g_i\}$ for $i = 1, \dots, t-1$, and $X_t = \{y_{t-1}\}$, $Y_t = \{y_{t-1}g_t\}$. Notice that the sets in the families $\{R_1, \dots, R_{t-1}\}$, $\{X_1, \dots, X_{t-1}\}$, $\{Y_1, \dots, Y_{t-1}\}$ are disjoint, that $\bigcup R_i$ is disjoint from $\bigcup X_i$, $\bigcup Y_i$ and that $(\bigcup X_i) \cap (\bigcup Y_i) = \{y_i, y_ig_{i+1} : i = 1, \dots, t-1\}$. By Lemma 5.3 R 1-absorbs $\{y_0g_1, y_0\} = \{kg, k\} = \{h, k\}$. □

The following lemma is similar to the previous one, except it allows us to 1-absorb pairs of sets of size 2.

Lemma 5.13 *Let $p \ge n^{-1/700}$. Let H_G be a multiplication hypergraph, R^1, R^2, R^3 disjoint, symmetric p-random subsets of G and set $R = R^1_A \cup R^2_B \cup R^3_C$. With high probability, the following holds*:

For any distinct, generic $a, b, c, d \in G_A, G_B$, or G_C with $[ab] = [cd]$ and $U \subseteq V(H_G)$ with $|U| \le p^{800}n/10^{4050}$, there is a set $R' \subseteq R \setminus U$ of size $\le 500\log n$ which 1-absorbs $\{\{a, b\}, \{c, d\}\}$.

Proof With high probability, the properties of Lemmas 5.8 and 5.12 hold. We just prove the lemma for $a, b, c, d \in G_B$. The other two cases follow by symmetry. Suppose that we have distinct, generic $a, b, c, d \in B(H_G)$ with $[ab] = [cd]$, and $U \subseteq G$ with $|U| \le p^{800}n/10^{21}$. Let $X = \{d, b\}$, $Y = \{c, a\}$. Using Lemma 5.8, pick some R', s with $R' \cup XsY \subseteq R \setminus (U \cup N(G))$ and R' 1-absorbing $\{\{a, dsc, b\}, \{c, bsa, d\}\}$.

When $dsc = bsa$, $R' \cup dsc$ 1-absorbs $\{\{a,b\},\{c,d\}\}$ and so satisfies the lemma. So suppose $dsc \neq bsa$. Notice that $[ab] = [cd]$ implies $[bsa] = [dsc]$ and so using Lemma 5.12 we can choose a $Q \subseteq V(H_G) \setminus (U \cup R' \cup \{dsc, bsa\})$ which 1-absorbs $\{bsa, dsc\}$ (bsa, dsc are generic because they are contained in XsC). Set $R'' = R' \cup Q \cup \{bsa, dsc\}$ to get a set which 1-absorbs $\{\{a,b\},\{c,d\}\}$ by Lemma 5.3. □

5.1.2 Distributive absorption for pairs

In this section we prove a variety of lemmas about h-absorbing sets of the form $\{\{a_1,b_1\},\dots,\{a_k,b_k\}\}$ where $[a_1b_1] = \dots = [a_kb_k]$. The following does this for $h = k-1$.

Lemma 5.14 *Let $p \geq n^{-1/700}$. Let H_G be a multiplication hypergraph, R^1, R^2, R^3 disjoint, symmetric p-random subsets of G and set $R = R^1_A \cup R^2_B \cup R^3_C$. With high probability, the following holds:*

Let $k \leq 200$ and $x_1, y_1, \dots, x_k, y_k \in G_A, G_B$ or G_C be distinct, generic with $[x_1y_1] = \dots = [x_ky_k]$ and let $U \subseteq G$ with $|U| \leq p^{800}n/10^{4060}$. Then there is a set $R' \subseteq R \setminus U$ of size $\leq 600k\log n$ which $(k-1)$-absorbs $\{\{x_1,y_1\},\dots,\{x_k,y_k\}\}$.

Proof With high probability, the property of Lemma 5.13 holds. Let $x_1, y_1, \dots, x_k, y_k$ be distinct, generic with $[x_1y_1] = \dots = [x_ky_k]$ and let $U \subseteq G$ with $|U| \leq p^{800}n/10^{4060}$. Apply the property of Lemma 5.12 to get disjoint sets $R_1, \dots, R_{k-1} \subseteq R \setminus U$ which 1-absorb $\{\{x_i,y_i\},\{x_{i+1},y_{i+1}\}\}$ for each $i \in [k-1]$ (for disjointness use $U' = U \cup \bigcup_{j=1}^{i-1} R_j$ at each application. This set has size $\leq |U| + k500\log n \leq p^{800}n/10^{4060}$). Set $R' = R_1 \cup \dots \cup R_{k-1}$. For each i, we have matchings R_i^- and R_i^+ with vertex sets $R_i \cup \{x_i, y_i\}$ and $R_i \cup \{x_{i+1}, y_{i+1}\}$. Now the matchings $M_i = (\bigcup_{j=1}^{i-1} R_i^-) \cup (\bigcup_{j=i}^{|Y|-1} R_i^+)$ have vertex sets exactly $V(M_i) = R' \cup Y \setminus \{x_i, y_i\}$, and so they satisfy the definition of R' $(k-1)$-absorbing Y. □

We'll need the following lemma which finds a common neighbour of a set of vertices in H_k.

Lemma 5.15 *Let $p \geq n^{-1/700}$ and let H_G be a multiplication hypergraph, R^1, R^2, R^3 disjoint, symmetric p-random subsets of G and set $R = R^1_A \cup R^2_B \cup R^3_C$. With high probability, the following holds:*

Let $k \leq 200$ and $a_1, \dots, a_k \in G_A$ be distinct and generic and let $U \subseteq V(H_G)$ with $|U| \leq p^{800}n/10^{4001}$. Then there are distinct, generic $b, c_1, \dots, c_k \in G \setminus U$ such that for each i, $\{a_i, b, c_i\}$ is an edge of H_G.

Proof With b the free variable in $G * F_1$, consider the set $S = \{b, b^{-1}a_1^{-1}, \dots, b^{-1}a_k^{-1}\}$, $S_A = \emptyset$, $S_B = \{b\}$, $S_C = \{b^{-1}a_1^{-1}, \dots, b^{-1}a_k^{-1}\}$. Notice that all $w \in S$ are linear, that S_A, S_B, S_C are pairwise separable (by part (b) of the definition since $a_1, \dots, a_k$ are generic), and all $w, w' \in S_C$ are weakly separable (by (b') since equalities between w, w' rearrange to $e = a_ia_j^{-1}$ and we know $a_ia_j^{-1} \neq e$ by distinctness). Thus the lemma follows from Lemma 3.32 applied with $U' = U \cup N(G)$. □

The following lemma 1-absorbs a set of k pairs.

Lemma 5.16 *Let $p \geq n^{-1/700}$. Let H_G be a multiplication hypergraph, R^1, R^2, R^3 disjoint, symmetric p-random subsets of G and set $R = R^1_A \cup R^2_B \cup R^3_C$. With high probability, the following holds:*

Let $k \leq 200$ and $x_1, y_1, \dots, x_k, y_k \in G_A$ be distinct, generic with $[x_1 y_1] = \dots = [x_k y_k]$ and let $U \subseteq G$ with $|U| \leq p^{800} n/10^{4070}$. Then there is a set $R' \subseteq R$ of size $\leq 600k \log n$ which 1-absorbs $\{\{x_1, y_1\}, \dots, \{x_k, y_k\}\}$.

Proof With high probability, the conclusions of Lemmas 5.15 and 5.14 apply. Let $x_1, y_1, \dots, x_k, y_k \in G_A$ be distinct, generic with $[x_1 y_1] = \dots = [x_k y_k]$ and let $U \subseteq G$ with $|U| \leq p^{800} n/10^{4070}$. Use the conclusion of Lemma 5.15 twice to get elements b^x, b^y and $c^x_1, c^y_1, \dots, c^x_k, c^y_k$ outside U, such that $x_1, y_1, \dots, x_k, y_k$, b^x, b^y, $c^x_1, c^y_1, \dots, c^x_k, c^y_k$ are all distinct, generic and also $x_i b^x c^x_i$, $y_i b^y c^y_i$ are edges for all i (to get distinctness of all the vertices, enlarge U to include previously found vertices between the two applications). Notice that this implies that $[c^x_i c^y_i] = [x_i^{-1} b_x^{-1} y_i^{-1} b_y^{-1}] = [x_1^{-1} b_x^{-1} y_1^{-1} b_y^{-1}]$ for all i (using $[x_1 y_1] = \dots = [x_k y_k]$). Thus using the conclusion of Lemma 5.14, we get a set R which $(k-1)$-absorbs $\{\{c^x_1, c^y_1\}, \dots, \{c^x_k, c^y_k\}\}$. In other words we have matchings $M_1, \dots, M_k$ with $V(M_i) = R \cup \{\{c^x_j, c^y_j\} : j \neq i\}$. Set $R'' = R' \cup \{b^x, b^y, c^x_1, c^y_1 \dots, c^x_k, c^y_k\}$. Now for each i, $M_i \cup \{x_i b^x c^x_i, y_i b^y c^y_i\}$ is a matching with vertex set exactly $R'' \cup \{x_i, y_i\}$, verifying the definition of "1-absorbs". □

The below was shown by Montgomery in [46], and is the essence of the distributive absorption approach.

Lemma 5.17 (Montgomery, [46]) *There is a constant 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 $|X| = 3h$, $|Y| = |Y'| = 2h$, so that the following holds. For any $Y_0 \subseteq Y'$ with $|Y_0| = h$, there is a perfect matching between X and $Y \cup Y_0$.*

Graphs produced by this lemma are called **robustly matchable bipartite graphs**. Combining this lemma with the previous one, we can h-absorb sets of pairs.

Lemma 5.18 *Let $p \geq n^{-1/700}$. Let H_G be a multiplication hypergraph, R^1, R^2, R^3 disjoint, symmetric p-random subsets of G and set $R = R^1_A \cup R^2_B \cup R^3_C$. With high probability, the following holds:*

Consider sets Y, Y' of disjoint pairs of generic elements $(a_1, a_2) \in (A \setminus e) \times (A \setminus e)$ having $[a_1 a_2] = [a'_1 a'_2]$ for $(a_1, a_2), (a'_1, a'_2) \in Y, Y'$ and also $2h = |Y| = |Y|' \leq \frac{p^{800} n}{10^{4080} \log n}$. Let $U \subseteq G$ with $|U| \leq p^{800} n/10^{4080}$. Then there is a subset $R' \subseteq R \setminus U$ of size $\leq 800|Y| \log n$ such that $R' \cup Y'$ h-absorbs Y.

Proof With high probability the property of Lemma 5.16 holds.

Let Y, Y' be disjoint sets of pairs of elements $(a_1, a_2) \in A \times A$ having $[a_1 a_2] = [a'_1 a'_2]$ for $(a_1, a_2), (a'_1, a'_2) \in Y$ and also $2h := |Y| = |Y|' \leq \frac{p^{800} n}{10^{4080} \log n}$. Let $U \subseteq G$

with $|U| \le p^{800}n/10^{4080}$. Consider a robustly matchable bipartite graph D with $\Delta(D) \le 100$ whose sets are Y, Y' and X as in Lemma 5.17 (here X is just an abstract set of size $3h$ unrelated to the hypergraph we have). For all $x \in X$ use Lemma 5.16 to pick a set $R_x \subseteq R \setminus (U \cup Y \cup Y')$ which 1-absorbs $N(x)$. We can choose all these sets to be disjoint by enlarging U to contain the union of previously picked sets at each application (whose total size is at most $100|X| \times 700 \log n \le p^{800}n/10^{4070}$).

Letting $R' \bigcup_{x\in X} R_x$ we claim that $R' \cup Y'$ will h-absorb Y, and so satisfy the lemma. Let $Y_0 \subseteq Y$ be a set of $|Y|/2 = h$ pairs. Since D is robustly matchable, there is a matching M with vertex set $Y_0 \cup Y' \cup X$. For each $x \in X$, let xy_x be the matching edge of M through x, and let N_x be a matching with vertex set $R_x \cup y_x$ (which exists because R_x 1-absorbs $N(x)$). Now $\bigcup_{x\in X} N_x$ is a matching with vertex set $Y_0 \cup R' \cup Y'$. □

The following is a version of the previous lemma with more versatility with choosing h.

Lemma 5.19 *Let $p \ge n^{-1/700}$. Let H_G be a multiplication hypergraph, R^1, R^2, R^3 disjoint, symmetric p-random subsets of G and set $R = R^1_A \cup R^2_B \cup R^3_C$. With high probability, the following holds:*

Consider sets Y, Z of disjoint pairs of generic elements $(a_1, a_2) \in A \times A$ having $[a_1a_2] = [a'_1a'_2]$ for $(a_1, a_2), (a'_1, a'_2) \in Y, Z$ and $4|Y| \le |Z| = \frac{p^{800}n}{10^{4090}\log n}$. Let $h \in \mathbb{N}$. Let $U \subseteq G$ with $|U| \le p^{800}n/10^{4090}$. Then there is a subset $R' \subseteq R \setminus U$, $Y' \subseteq Z$ of size $\le 800|Y|\log n$ such that $R' \cup Y'$ h-absorbs Y.

Proof With high probability Lemma 5.18 applies. If $h > |Y|$, then the definition of "h-absorbs Y" is vacuous, so we can suppose that $h \le |Y|$. If $|Y|/2 \le h \le |Y|$, pick subsets $\hat{Y}, Y' \subseteq Z$ with $|\hat{Y}| = 2h - |Y|$ and $|Y'| = 2h$. Apply Lemma 5.18 to g, $Y \cup \hat{Y}$, Y', and h to get a set R' such that $R' \cup Y'$ h-absorbs $Y \cup \hat{Y}$. By Lemma 5.2 (1), $R' \cup Y'$ h-absorbs Y also.

If $h \le |Y|/2$, pick subsets $\hat{Y}, Y' \subseteq Z$ with $|\hat{Y}| = |Y| - 2h$ and $|Y'| = 2|Y| - 2h$. Apply Lemma 5.18 to g, $Y \cup \hat{Y}$, Y', and $h' = |Y| - h$ to get a set R' such that $R' \cup Y'$ $(|Y| - h)$-absorbs $Y \cup \hat{Y}$. By Lemma 5.2 (2) with $t = |\hat{Y}| = |Y| - 2h$, $R' \cup Y' \cup \hat{Y}$ $(|Y| - h - |\hat{Y}|)$-absorbs Y. This implies the lemma since $|Y| - h - |\hat{Y}| = h$. □

5.1.3 Distributive absorption for singletons

Everything in this section is almost identical to the previous one (though often a bit easier). We give constructions of h-absorbers of sets of vertices Y contained in a coset of G'. The following lemma does this with $h = |Y| - 1$.

Lemma 5.20 *Let $p \ge n^{-1/700}$. Let H_G be a multiplication hypergraph, R^1, R^2, R^3 disjoint, symmetric p-random subsets of G and set $R = R^1_A \cup R^2_B \cup R^3_C$. With high probability, the following holds:*

For $g \in G$, let $Y \subseteq G_A, G_B$, or G_C with $Y \subseteq [g]$ be a set of generic elements with $|Y| \le \frac{p^{800}n}{10^{4100}\log n}$ and let $U \subseteq G$ with $|U| \le p^{800}n/10^{4100}$. Then there is a set $R' \subseteq R \setminus U$ of size $\le 700|Y|\log n$ which $(|Y| - 1)$-absorbs Y.

Proof With high probability, the property of Lemma 5.12 holds.

Let $g \in G$ and $Y = \{y_1, \dots, y_{|Y|}\} \subseteq [g]$ with $|Y| \leq \frac{p^{800}n}{10^{4100}\log n}$ and U with $|U| \leq p^{800}n/10^{17}$. Apply the property of Lemma 5.12 to get disjoint sets $R_1, \dots, R_{|Y|-1} \subseteq R \setminus U$ which 1-absorb $\{y_i, y_{i+1}\}$ (for disjointness use $U' = U \cup \bigcup_{j=1}^{i-1} R_j$ at each application. This set has size $\leq |U| + |Y|500\log n \leq p^{800}n/10^{4090}$). Set $R' = R_1 \cup \dots \cup R_{|Y|-1}$. For each i, we have matchings R_i^- and R_i^+ with vertex sets $R_i \cup \{y_i\}$ and $R_i \cup \{y_{i+1}\}$. Now the matchings $M_i = (\bigcup_{j=1}^{i-1} R_i^-) \cup (\bigcup_{j=i}^{|Y|-1} R_i^+)$ have vertex sets exactly $V(M_i) = R' \cup Y \setminus \{y_i\}$, and so they satisfy the definition of R' $(|Y|-1)$-absorbing Y. □

The next lemma builds 1-absorbers of small sets Y.

Lemma 5.21 *Let $p \geq n^{-1/700}$. Let H_G be a multiplication hypergraph, R^1, R^2, R^3 disjoint, symmetric p-random subsets of G and set $R = R_A^1 \cup R_B^2 \cup R_C^3$. With high probability, the following holds:*

Let $k \leq 200$ and $s_1, \dots, s_k \in G_A$ be distinct, generic with $[s_1] = \dots = [s_k]$ and let $U \subseteq G$ with $|U| \leq p^{800}n/10^{4110}$. Then there is a set $R' \subseteq R \setminus U$ of size $\leq 700k\log n$ which 1-absorbs $\{s_1, \dots, s_k\}$.

Proof With high probability, the conclusions of Lemmas 5.15 and 5.20 apply. Let $s_1, \dots, s_k \in G$ be distinct, generic with $[s_1] = \dots = [s_k]$ and let $U \subseteq G$ with $|U| \leq p^{800}n/10^{4110}$. Use the conclusion of Lemma 5.15 to get distinct, generic elements b and $c_1, \dots, c_k \in G \setminus U$, such that also $s_i bc_i$ are edges for all i. Thus using the conclusion of Lemma 5.20, we get a set R which $(k-1)$-absorbs $\{c_1, \dots, c_k\}$. From the definition of $(k-1)$-absorbs, we have matchings $M_1, \dots, M_k$ with $V(M_i) = R \cup \{c_1, \dots, c_{i-1}, c_{i+1}, \dots, c_k\}$. Set $R'' = R' \cup \{b, c_1, \dots, c_k\}$. Now for each i, $M_i \cup s_i bc_i$ is a matching with vertex set exactly $R'' \cup \{s_i\}$, verifying the definition of “1-absorbs”. □

The next lemma uses distributive absorption to build h-absorbers.

Lemma 5.22 *Let $p \geq n^{-1/700}$. Let H_G be a multiplication hypergraph, R^1, R^2, R^3 disjoint, symmetric p-random subsets of G and set $R = R_A^1 \cup R_B^2 \cup R_C^3$. With high probability, the following holds:*

Let $g \in G$, and consider disjoint sets $Y, Y' \subseteq [g]$ of generic elements of G_A of size $2h := |Y| = |Y'| \leq \frac{p^{800}n}{10^{4120}\log n}$. Let $U \subseteq G$ with $|U| \leq p^{800}n/10^{4120}$. Then there is are subset $R' \subseteq R \setminus U$ of size $\leq 10^5|Y|\log n$ such that $R' \cup Y'$ h-absorbs Y.

Proof With high probability the property of Lemma 5.21 holds. Let $g \in G$, every disjoint sets $Y, Y' \subseteq [g]$ of size $2h = |Y| = |Y'| \leq \frac{p^{800}n}{10^{4120}\log n}$. Let $U \subseteq G$ with $|U| \leq p^{800}n/10^{4120}$. Consider a robustly matchable bipartite graph D with $\Delta(D) \leq 100$ whose sets are Y, Y' and X as in Lemma 5.17 (here X is just an abstract set of size $3h$ unrelated to the hypergraph we have). For all $x \in X$ fix a set $R_x \subseteq R$ which 1-absorbs $N(x)$. We can choose all these sets to be disjoint by letting U be the union of previously picked sets at each application (whose total size is at most $100|X| \times 700 \log n \le p^{800}n/10^{4110}$).

We claim that for $R' := \bigcup_{x \in X} R_x$, we have $R' \cup Y'$ h-absorbing Y, and so satisfying the lemma. Let $Y_0 \subseteq Y$ be a set of $|Y|/2 = h$ pairs. Since D is robustly matchable, there is a matching M with vertex set $Y_0 \cup Y' \cup X$. For each $x \in X$, let xy_x be the matching edge of M through x, and let N_x be a matching with vertex set $R_x \cup y_x$ (which exists because R_x 1-absorbs $N(x)$). Now $\bigcup_{x \in X} N_x$ is a matching with vertex set $Y_0 \cup R' \cup Y'$. □

The next lemma is a version of the previous one which allows for more flexibility in the value of h.

Lemma 5.23 *Let $p \ge n^{-1/700}$. Let H_G be a multiplication hypergraph, R^1, R^2, R^3 disjoint, symmetric p-random subsets of G and set $R = R^1_A \cup R^2_B \cup R^3_C$. With high probability, the following holds:*

Let $g \in G$, and consider disjoint sets $Y, Z \subseteq [g]$ of generic elements of G_A with $4|Y| \le |Z| = \frac{p^{800}n}{10^{4130} \log n}$. Let $h \in \mathbb{N}$. Let $U \subseteq G$ with $|U| \le p^{800}n/10^{4130}$. Then there is are subset $R' \subseteq R \setminus U$, $Y' \subseteq Z$ of size $\le 10^5|Y| \log n$ such that $R' \cup Y'$ h-absorbs Y.

Proof With high probability Lemma 5.22 applies. If $h > |Y|$, then the definition of "h-absorbs Y" is vacuous, so we can suppose that $h \le |Y|$. If $|Y|/2 \le h \le |Y|$, pick subsets $\hat{Y}, Y' \subseteq Z$ with $|\hat{Y}| = 2h - |Y|$ and $|Y'| = 2h$. Apply Lemma 5.22 to g, $Y \cup \hat{Y}$, Y', and h to get a set R' such that $R' \cup Y'$ h-absorbs $Y \cup \hat{Y}$. By Lemma 5.2 (1), $R' \cup Y'$ h-absorbs Y also.

If $h \le |Y|/2$, pick subsets $\hat{Y}, Y' \subseteq Z$ with $|\hat{Y}| = |Y| - 2h$ and $|Y'| = 2|Y| - 2h$. Apply Lemma 5.22 to g, $Y \cup \hat{Y}$, Y', and $h' = |Y| - h$ to get a set R' such that $R' \cup Y'$ $(|Y| - h)$-absorbs $Y \cup \hat{Y}$. By Lemma 5.2 (2) with $t = |\hat{Y}| = |Y| - 2h$, $R' \cup Y' \cup \hat{Y}$ $(|Y| - h - |\hat{Y}|)$-absorbs Y. This implies the lemma since $|Y| - h - |\hat{Y}| = h$. □

5.1.4 Main absorption lemma

The following lemma summarises everything from Sect. 5.1 that we use in other sections.

Lemma 5.24 *Let $p \ge n^{-1/700}$. Let H_G be a multiplication hypergraph, R^1, R^2, R^3 disjoint, symmetric p-random subsets of G and set $R = R^1_A \cup R^2_B \cup R^3_C$. With high probability, the following holds:*

Let $U \subseteq G$ with $|U| \le p^{800}n/10^{4140}$. Let $* \in \{A, B, C\}$. Suppose that Y and h are one of the following:

(1) *For some $g \in G$, $Y \subseteq [g] \cap * \setminus N(G)$ with $|Y| \le \frac{p^{800}|G'|}{10^{4140} \log n}$, and $h \le |Y|$.*
(2) *For some $g \in G$, $Y \subseteq [g] \cap * \setminus N(G)$ with $|Y| \le \frac{p^{800}n}{10^{4140} \log n}$, and $h = |Y| - 1$.*
(3) *For some generic a_ϕ, we have $Y \subseteq G \setminus G$ a set of $|Y| \le \frac{p^{800}n}{10^{4140} \log n}$ disjoint pairs of generic elements, with $[a_1a_2] = [a_\phi]$ for $(a_1, a_2) \in Y$, and $h \le |Y|$.*

Then there is a subset $R' \subseteq R \setminus U$ of size $\leq 10^6|Y|\log n$ which h-absorbs Y.

Proof With high probability, Lemmas 3.34, 5.20, 5.19, and 5.23 apply. Additionally when $|G'| \geq 10^{4140}p^{-800}\log n$, we can assume that $|R_i \cap [g]| \geq p|G'|/2 \geq 4 \times p^{800}|G'|/10^{4140}\log n$ for all cosets $[g]$ and $i = 1, 2, 3$ (using Chernoff's bound).

(1) We can suppose that $|G'| \geq 10^{4140}p^{-800}\log n$, since otherwise Y would have to be empty and the conclusion is vacuous. In this case, we have that for every g, there is a subset $Z_g \subseteq [g]$ of size $4 \times p^{800}|G'|/10^{4140}\log n$. Let $g \in G$, $Y \subseteq [g] \cap * $ of size $|Y| \leq \frac{p^{800}|G'|}{10^{4140}\log n}$, and $h \leq |Y|$. The result follows from Lemma 5.23 applied to Y, Z_g, U.
(2) This is strictly weaker than Lemma 5.20.
(3) We'll just prove the lemma when $Y \subseteq G_A \times G_A$, the other cases are symmetric. Use Lemma 3.34 in order to find a set $Z \subseteq (R_A \setminus U) \times (R_A \setminus U)$ of $4 \times \frac{p^{800}n}{10^{4140}\log n}$ pairs (a, a') of generic elements having $aa' = a_\phi$. Using Lemma 5.19, there is a $R' \subseteq R \setminus (U \cup Z)$ and $Z' \subseteq Z$ with $R' \cup Z'$ h-absorbing Y. □

5.2 Absorbing coset-paired sets

In this section we prove results about absorbing coset-paired sets. The main lemmas in this section are all along the lines of "given a set Q, there exists a set R' with the property that for every coset-paired $S \subseteq Q$, there is a matching with vertex set $R' \cup (Q \setminus S)$".

For all $g \in G$, define $i_{[g]} = 2$ if $[g]$ is self-paired and 1 otherwise. The following lemma allows us to absorb the complement of a coset-paired set.

Lemma 5.25 *Let $p \geq n^{-1/700}$. Let H_G be a multiplication hypergraph, R^1, R^2, R^3 disjoint, symmetric p-random subsets of G and set $R = R^1_A \cup R^2_B \cup R^3_C$. With high probability, the following holds:*

Let $s \in (0, 1]$. Let $m \in \mathbb{N}$. Let $U \subseteq G$ and $Q \subseteq G_\diamond$ (for some $\diamond \in \{A, B, C\}$) be disjoint, and suppose $|U| \leq p^{800}n/10^{4150}$. Suppose also that Q is generic, $|Q| \leq \frac{p^{800}n}{10^{4150}\log n}$, and that Q satisfies:

(i) For all pairs g, h either $Q \cap [g] = Q \cap [h] = \emptyset$ or $|Q \cap [g]|, |Q \cap [h]| \geq i_{[g]}\lceil 12s|G'|\rceil$.
(ii) If $|G'| > s^{-1}/12$ then for all pairs g, h $|Q \cap [g]|, |Q \cap [h]| \leq \frac{p^{800}|G'|}{10^{4150}\log n}$
(iii) If $|G'| \leq s^{-1}/12$, then Q doesn't intersect any self-paired cosets.
Then, there exists an $R' \subseteq R \setminus U$ such that for all coset-paired, $\lceil s|G'|\rceil$-coset-bounded $S \subseteq Q$ of size m, there is a matching with vertex set $R' \cup (Q \setminus S)$.

Proof With high probability Lemma 5.24 applies to R. Let s, m, U, Q be as in the lemma. Let K be the set of cosets $[g]$ with $Q \cap [g]$ nonempty. Partition $Q = Q_1 \cup Q_2$ where, for each coset $[g] \in K$, we have $|Q_1 \cap [g]| = i_{[g]}\lceil 12s|G'|\rceil$ (and so $|Q_2 \cap [g]| = i_{[g]}|Q \cap [g]| - \lceil 12s|G'|\rceil$). Note that Q_1 is coset-paired (since it has even intersection with self-paired cosets and intersection $\lceil 12s|G'|\rceil$ with other pairs of cosets) and so we can partition it into a set of $|Q_1|/2 \leq |Q| \leq \frac{p^{800}n}{10^{4150}\log n}$ pairs P_1. For each g, let $h_g = |Q \cap [g]| - i_{[g]}\lceil s|G'|\rceil$, noting that this is nonnegative by (i). Use Lemma 5.24 (1) or (2) to pick a subset $R_g \subseteq R$ which h_g-absorbs $Q \cap [g]$ (if $|G'| \geq s^{-1}/12$, then, using condition (ii), part (1) of that lemma gives this. Otherwise, we have that Q intersects no self-paired cosets and $\lceil s|G'|\rceil = 1$ and so part (2) gives it). Set $h_P = \sum_{[g]\in K} i_{[g]}\lceil s|G'|\rceil$, noting that this is even. Also note that we can assume that m is even and $\leq h_g$ (otherwise there could be no coset-paired, $\lceil s|G'|\rceil$-coset-bounded S of size m contained in Q). So, we can use Lemma 5.24 (3) to pick a subset R_P which $\frac{1}{2}(h_P - m)$-absorbs P_1. We can ensure that these are all disjoint by enlarging U as we choose them (the maximum total size of sets we need to avoid is $|U| + \sum_{[g]\in K} |R_g| \leq |U| + \sum_{[g]\in K} 10^6 \log n|Q \cap [g]| \leq |U| + 10^6 \log n|Q| \leq p^{800}n/10^{4150} + 10^6 \log n(\frac{p^{800}n}{10^{4150}\log n}) \leq p^{800}n/10^{4140}$). Let $R' = R_P \cup \bigcup P_1 \cup \bigcup_{g\in K} R_g$.

Consider some coset-paired, $\lceil s|G'|\rceil$-coset-bounded $S \subseteq Q$ of size m. Since S is coset-paired, $|S \cap [g]|/i_{[g]}$ is an integer for all $[g]$.

Claim 5.25.1 *There is a subset $P_1' \subseteq P_1$ of pairs disjoint from S which have $\lceil s|G'|\rceil - |S \cap [g]|/i_{[g]}$ pairs intersecting each $[g] \in K$.*

Proof First suppose that $\lceil s|G'|\rceil = 1$. Note that for each $[g] \in K$ this means that we need to select at most one pair intersecting $[g]$, and we only need to do so when $S \cap [g], S \cap [\phi(g)] = \emptyset$. In this case we have $|Q_1 \cap [g]| = |Q_1 \cap [\phi(g)]| = i_{[g]}|\lceil 12s|G'|\rceil$ and so there is at least one pair in P_1 contained $[g] \cup [\phi(g)]$ (and choosing that pair works).

Next suppose that $\lceil s|G'|\rceil > 1$, which implies that $s|G'| \geq \lceil s|G'|\rceil - 1 \geq \lceil s|G'|\rceil/2$. Then $|Q_1 \cap [g]| = i_{[g]}\lceil 12s|G'|\rceil \geq 12s|G'| \geq 6\lceil s|G'|\rceil$. Thus there are at least $3\lceil s|G'|\rceil$ pairs contained in $[g] \cup [\phi(g)]$. At most $2\lceil s|G'|\rceil$ of them can intersect $S \cap ([g] \cup [\phi(g)])$ and so we can choose $\lceil s|G'|\rceil - |S \cap [g]|/i_{[g]}$ of them disjointly from S. □

Note that $|P_1'| = \frac{1}{2}|\bigcup P_1'| = \frac{1}{2}\sum_{[g]\in K} i_{[g]}(\lceil s|G'|\rceil - |S \cap [g]|/i_{[g]}) = \frac{1}{2}(h_P - m)$. By the property of R_P, there is a matching M_P with vertex set $R_P \cup \bigcup P_1'$. For each $g \in K$, note that $|[g] \cap (S \cup \bigcup P_1')| = i_{[g]}\lceil s|G'|\rceil$ and so $|[g] \cap (Q \setminus (S \cup \bigcup P_1'))| = h_g$. By the property of R_g, there is a matching M_g with vertex set $R_g \cup ([g] \cap (Q \setminus (S \cup \bigcup P_1')))$. The union of these matchings has vertex set $(R_P \cup \bigcup P_1') \cup \bigcup_{[g]\in K} R_g \cup ([g] \cap (Q \setminus (S \cup \bigcup P_1'))) = R' \cup \bigcup P_1' \cup (Q \setminus (S \cup \bigcup P_1')) = R' \cup (Q \setminus S)$ as required. □

The following simple lemma covers generic sets by matchings.

Lemma 5.26 *Let $p \geq n^{-1/700}$. Let H_G be a multiplication hypergraph, R^1, R^2, R^3 disjoint, symmetric p-random subsets of G and set $R = R^1_A \cup R^2_B \cup R^3_C$. With high probability, the following holds:*

For any $U \subseteq V(H_G)$ and any generic $X \subseteq V(H_G)$ with $|X|, |U| \leq p^{800}n/10^{4010}$, there is a matching M of size $|X|$ in H_G covering X and having all other vertices in $R \setminus U$.

Proof With high probability, the property of Lemma 5.5 applies i.e. for any generic $v \in V(H_G)$ and $U \subseteq V(H_G)$ with $|U| \leq p^{800}n/10^{4000}$, there is an edge e of H_G passing through v and having the other two vertices in $R \setminus U$. Now let $U \subseteq V(H_G)$ and any generic $X \subseteq V(H_G)$ with $|X|, |U| \leq p^{800}n/10^{4010}$. Applying the property of Lemma 5.5 to each $v \in X$ we get edges e_v passing through v and having other vertices in $R \setminus U$. By enlarging U as we choose these to include previously found vertices, we can ensure that all e_v are disjoint i.e. they give us a matching like the lemma asks for. □

The following is a variant of Lemma 5.25 where Q is a random set.

Lemma 5.27 *Let $p \geq n^{-1/701}$ and $q \leq \frac{p^{800}}{10^{4160}\log n}$. Let H_G be a multiplication hypergraph, R^1, R^2, R^3 disjoint, symmetric p-random subsets of G and Q a disjoint q-random subset. Set $R = R^1_A \cup R^2_B \cup R^3_C$. With high probability, the following holds:*

Let $s \leq \frac{q^2 p^{800}}{10^{4160}\log^3 n}$. For every $m \in \mathbb{N}$ and $U \subseteq G$ with $|U| \leq q^2 p^{800} n/10^{4160}$, there is a U' with $U \subseteq U' \subseteq U \cup Q$, $|U'| \leq 5q^{-1}|U| + 50s^{-1}$, and $R' \subseteq R \setminus U$ such that for all coset-paired, $\lceil s|G'|\rceil$-coset-bounded $S \subseteq Q \setminus U'$ of size m, there is a matching with vertex set $R' \cup (Q \setminus (S \cup U'))$.

Proof Randomly split each R^i into two disjoint, symmetric $p/2$-random sets R^i_-, R^i_+. Set $R^- = R^1_{-A} \cup R^2_{-B} \cup R^3_{-C}$ and $R^+ = R^1_{+A} \cup R^2_{+B} \cup R^3_{+C}$. Note that the following hold with high probability.

(A1) R^- satisfies Lemmas 5.25 while R^+ satisfies Lemma 5.26.
(A2) When $|G'| \geq q^{-2}\log^2 n$, then $|Q \cap [g]| \in [q|G'|/2, 2q|G'|]$ for all $g \in G$ (from Chernoff's bound).
(A3) $|Q| \leq qn + \sqrt{n}\log n \leq \frac{p^{800}n}{10^{4150}\log n}$ (from Chernoff's bound).

Now let $s \leq \frac{q^2 p^{800}}{10^{4160}\log^3 n} \leq q^2 \log^{-2} n/12$, $m \in \mathbb{N}$, and $U \subseteq G$ with $|U| \leq q^2 p^{800} n/10^{4160}$. When $|G'| \leq s^{-1}/12$, let U_1 be the union of self-paired cosets, noting that $|U_1| \leq 30|G'| \leq 30s^{-1}$. When $|G'| > s^{-1}/12$, let U_1 be the union of cosets $[g]$ and their pairs for which $|(U \cup N(G)) \cap [g]| \geq q|G'|/10$, noting that $|U_1| \leq 2|G'|\frac{|U|+|N(G)|}{q|G'|/10}$. Set $U' = U \cup N(G) \cup U_1$, noting that $|U'| \leq 5q^{-1}|U| + 6q^{-1}|N(G)| + 30s^{-1} \leq 5q^{-1}|U| + 6q^{-1}10^{9000} + 30s^{-1} \leq 5q^{-1}|U| + 50s^{-1}$. Set $Q_1 = Q \setminus U'$. Let $Q_2 \subseteq Q_1$ be the subset formed by deleting all pairs of cosets for which $Q_1 \cap [g]$ or $Q_1 \cap [\phi(g)]$ is empty. Note that Q_2 satisfies all of the following properties.

1. Q_2 is generic (since it's disjoint from $N(G)$).
2. $|Q_2| \leq \frac{p^{800}n}{10^{4150}\log n}$ (by (A3) and $Q_2 \subseteq Q$).
3. When $|G'| \leq s^{-1}/12$, Q_2 doesn't intersect self-paired cosets (since it's disjoint from U_1).
4. When $|G'| \leq s^{-1}/12$, for all $g \in G$ we have that either $Q_2 \cap [g] = Q_2 \cap [h] = \emptyset$ or $i_{[g]}\lceil 12s|G'|\rceil = 1 \leq |Q_2 \cap [g]|, |Q_2 \cap [h]|$ (by construction of Q_2 from Q_1).

5. When $|G'| > s^{-1}/12$, for all $[g]$ either $Q_2 \cap [g] = Q_2 \cap [h] = \emptyset$ or we have $i_{[g]}12s|G'| \leq 24s|G'| \leq q|G'|/4 \leq q|G'|/2 - q|G'|/10 \leq |Q_2 \cap [g]|$, $|Q_2 \cap [h]| \leq 2q|G'| \leq \frac{p^{800}|G'|}{10^{4150}\log n}$ (by (A2), since Q_2 is disjoint from U_1, and since $s \leq \frac{q^2 p^{800}}{10^{4160}\log^3 n}$, $q \leq \frac{p^{800}}{10^{4160}\log n}$).

Thus Lemma 5.25 applies to Q_2. This gives us a set $R'^{-} \subseteq R^{-} \setminus U$. By Lemma 5.26, there is a matching M_1 covering $Q_1 \setminus Q_2$ with $V(M_1) \setminus (Q_1 \setminus Q_2) \subseteq R^{+} \setminus U$. Set $R' = R'^{-} \cup (M_1 \setminus Q) \subseteq R \setminus U$. This is possible as $Q_1 \setminus Q_2 \subseteq Q$ and (A3).

Now consider a coset-paired, $\lceil s|G'|\rceil$-coset-bounded $S \subseteq Q \setminus U' = Q_1$ of size m. Note that we must have $S \subseteq Q_2$ because S is coset paired (and for every $g \in Q_1 \setminus Q_2$ we have $[\phi(g)] \cap Q_1$ empty). Therefore, Lemma 5.25 gives us a matching M_2 with vertex set $R'^{-} \cup (Q_2 \setminus S)$. Combining this with M_1 gives a matching with vertex set $R'^{-} \cup (Q_2 \setminus S) \cup (M_1 \setminus Q) \cup (Q_1 \setminus Q_2) = R' \cup (Q_1 \setminus S) = R' \cup (Q \setminus (S \cup U'))$ as required. □

The following shows that large coset-paired matchings exist inside random sets.

Lemma 5.28 *Let $p \geq n^{-1/700}$. Let H_G be a multiplication hypergraph, R^1, R^2, R^3 disjoint, symmetric p-random subsets of G and set $R = R^1_A \cup R^2_B \cup R^3_C$. With high probability, the following holds:*

Let $U \subseteq G$ with $|U| \leq p^{800}n/10^{4140}$. For any $k \in [1, |G'|]$, there is a coset-paired, k-coset-bounded matching of size $\frac{p^{800}nk}{10^{4140}|G'|}$ in $R \setminus U$.

Proof With high probability, Lemma 3.32 applies. When $|G'| \leq n/\log^{10^{20}} n$, consider the set

$$S = \{x, x^{-1}a_\phi, y, b_\phi y^{-1}, y^{-1}x^{-1}, yc_\phi x\} \subseteq G * F_2.$$

When $|G'| > n/\log^{10^{20}} n$, consider the set

$$S = \{x_1, x_2 a_\phi, y_1, y_2 b_\phi, y_1^{-1}x_1^{-1}, y_2 c_\phi x_2\} \subseteq G * F_4.$$

Note that in either case all words in S are linear and all pairs of words are separable (using part (a) or (c) of the definition of "separable". For checking (c), we use that for any $(w, w') = (x, x^{-1}a_\phi), (y, b_\phi y^{-1}), (y^{-1}x^{-1}, yc_\phi x)$ we have $\pi_0(ww'), \pi_0(w^{-1}w') \in \{a_\phi, b_\phi, c_\phi, a_\phi^{-1}, b_\phi^{-1}, c_\phi^{-1}\}$, which all satisfy (c) from Lemma 3.14).

Let $U \subseteq G$ with $|U| \leq p^{800}n/10^{4120}$. If $|G'| \leq p^{800}n/10^{4120}$ add all self-paired cosets to U in order to get a set U' with $|U'| \leq p^{800}n/10^{4110}$ (otherwise set $U' = U$). Use Lemma 3.32 to get a projection π which separates S and has $\pi(S) \subseteq R \setminus U'$. This gives us a coset-paired matching of size 2 as in the lemma (whose edges are $(\pi(x), \pi(y), \pi(y^{-1}x^{-1}))$ and $(\pi(x^{-1}a_\phi), \pi(b_\phi y^{-1}), \pi(yc_\phi x))$). To get one of size $\frac{p^{800}nk}{10^{4140}|G'|}$, keep selecting multiple matchings like this, enlarging U at every step in order to keep them disjoint (we can do this as long as $3|G'||M|k^{-1} \leq p^{800}n/10^{4130}$.

Indeed for a matching M, letting U_M be the union of cosets $[g]$ having $|V(M) \cap [g]| \geq k$, note that $|U_M|/|G'| \leq |V(M)|/k$ which gives $|U_M| \leq 3|G'||M|/k$. For any pair of edges outside $U \cup U_M$, adding them to M gives a bigger matching like we want). □

Now we arrive at the main lemma of this section. The following again shows that we can absorb the complement of a coset-paired set. Unlike Lemmas 5.25 and 5.27, this one allows the coset-paired set to have variable size.

Lemma 5.29 *Let $q, p \geq n^{-1/701}$ with $q \leq \frac{p^{801}}{10^{4170} \log n}$. Let H_G be a multiplication hypergraph, let R^1, R^2, R^3, Q_1, Q_2, Q_3 be disjoint, symmetric subsets of G with R^1, R^2, R^3 p-random and Q_1, Q_2, Q_3 q-random subsets. Set $R = R^1_A \cup R^2_B \cup R^3_C$ and $Q = Q^1_A \cup Q^2_B \cup Q^3_C$. With high probability, the following holds:*

Let $s \leq \frac{q^{1600} p^{10^7}}{10^{10^7} \log^{2400} n}$, $U \subseteq G$ with $|U| \leq q^9 p^{800} n/10^{4170}$. There is a $U' \subseteq U$ with $|U'| \leq 500q^{-3}|U| + 10^{10}s^{-3}$, and $R' \subseteq R \setminus U$ such that for all coset-paired, $\lceil s|G'|\rceil$-coset-bounded, balanced $S \subseteq Q \setminus U'$, with $|S| \leq s^2 n$ there is a matching with vertex set $R' \cup (Q \setminus (S \cup U'))$.

Proof Set $m := 8\lceil sn \rceil$ and $s' := s^{1/800} 10^{4140}$. Split R into three disjoint, symmetric $p/3$-random subsets R_1, R_2, R_3 of $V(H_G)$ (by placing each $\hat{g}$ in $R_1/R_2/R_3$ with probability $1/3$, and making these choices independently of those for the random sets R^1, R^2, R^3 in the lemma. So now for all i $R_i \cap R^1_A$, $R_i \cap R^2_B$, $R_i \cap R^3_C$ are three disjoint, symmetric $p/3$-random subsets of G partitioning R_i). Lemma 5.27 applies to each pair $R' = R_i$, $Q' = Q_i$ for $i = 1, 2, 3$ (with $R'^1 = R_i \cap R^1_A$, $R'^2 = R_i \cap R^2_B$, $R'^3 = R_i \cap R^3_C$). Lemma 5.28 applies to Q. Let $U_0 := U$ be as in the lemma. For $i = 1, 2, 3$ build U_1, U_2, U_3, R_1, R_2, R_3 by applying Lemma 5.27 to R_i, Q_i with $U = U_{i-1}$, $m = m$, $s = 4s'$. The end result is sets $R'_1 \subseteq R_1$, $R'_2 \subseteq R_2$, $R'_3 \subseteq R_3$ and a set U' with $|U'| \leq 500q^{-3}|U| + 10^{10}s^{-3}$ such that for all coset-paired, $\lceil 4s'|G'|\rceil$-coset-bounded $S_i \subseteq Q_i \setminus U'$ of size m, there are matchings with vertex sets $R'_i \cup (Q_i \setminus (S_i \cup U'))$. Set $R' = R'_1 \cup R'_2 \cup R'_3$. Use Lemma 5.28 with $k = \lceil s'|G'|\rceil$ to find a coset-paired, $\lceil s'|G'|\rceil$-coset-bounded matching M_0 of size $\frac{p^{800} nk}{10^{4140}|G'|} \geq q^{800} s' n/10^{4140} \geq 8m$ in $Q_1 \cup Q_2 \cup Q_3 \setminus U'$.

Now, consider a coset-paired, $\lceil s|G'|\rceil$-coset-bounded, balanced $S \subseteq Q \setminus U'$ with $|S| \leq s^2 n$. Let $m_0 = |S \cap G_A| = |S \cap G_B| = |S \cap G_C| \leq s^2 n$, noting that these are even by coset-pairedness.

Claim 5.29.1 *There is a coset-paired submatching $M'_0 \subseteq M_0$ of size $m - m_0$ disjoint from S. Additionally when $|G'| \leq s^{-1}$, then M'_0 and S never intersect the same cosets.*

Proof When $|G'| \leq s^{-1}$ let S' be the union of cosets intersecting S, otherwise let $S' = S$. Note that in either case $|S'| \leq |S|s^{-1} \leq sn \leq m$. Since $|M_0| \geq 8m$, there is a paired submatching $M'_0 \subseteq M_0$ of size $m - m_0$ disjoint from S'. □

Now set $S'_1 := (S \cup M'_0) \cap G_A$, $S'_2 := (S \cup M'_0) \cap G_B$, $S'_3 := (S \cup M'_0) \cap G_C$ and note that these are coset-paired, $\lceil 4s'|G'|\rceil$-coset-bounded sets of size m. By the properties of R'_1, R'_2, R'_3, we get matchings M_1, M_2, M_3 with vertex sets $V(M_i) = R'_i \cup (Q_i \setminus (S'_i \cup U'))$. Now $M = M'_0 \cup M_1 \cup M_2 \cup M_3$ is a matching with vertex set $V(M'_0) \cup \bigcup_{i=1}^{3} R'_i \cup (Q_i \setminus (S'_i \cup U')) = R' \cup V(M'_0) \cup (Q \setminus (S \cup M'_0 \cup U')) = R' \cup (Q \setminus (S \cup U'))$ as required. □

5.3 Absorbing zero-sum sets

The goal of this section is to prove a lemma which can absorb arbitrary balanced zero-sum sets (Lemma 5.32). To do this we first prove results about covering zero-sum sets using coset-paired matchings.

5.3.1 Covering zero-sum sets

Here we prove Lemma 5.31, which roughly states that given a random vertex subset R of H_G and a small zero-sum set S, there exists a coset-paired set $R' \subseteq R$ such that $S \cup R'$ contains a perfect matching. As explained in Sect. 2, this will allow us to reduce the task of absorbing arbitrary zero-sum sets to absorbing coset-paired sets.

The following lemma does the majority of the work for this section. It allows us to find the desired set R' iteratively by reducing the size of the set of vertices we need to cover by 3 at every step.

Lemma 5.30 *Let $p \geq n^{-1/700}$. Let H_G be a multiplication hypergraph, R^1, R^2, R^3 disjoint, symmetric p-random subsets of G and set $R = R^1_A \cup R^2_B \cup R^3_C$. With high probability, the following holds.*

Let $F \subseteq V(H_G)$ be a balanced ϕ-generic set of size 6, $h \in [\prod F]$, and $U \subseteq V(H_G)$ with $|U| \leq p^{800}n/10^{4001}$. Then there is a matching M in H_G of size 15 and a disjoint set $\{a, b, c\}$ such that $F \subseteq V(M)$, $\{a, b, c\} \cup V(M) \setminus F \subseteq R \setminus U$, $\{a, b, c\} \cup V(M) \setminus F$ is coset-paired and $abc = h$. Furthermore, if $|G'| \leq \log^5 n/p^{2000}$, then for each $\diamond \in \{A, B, C\}$, and $g \in G$ we have that $|(\{a, b, c\} \cup V(M) \setminus F) \cap G_\diamond \cap [g]| \leq 1$.

Proof With high probability, the property of Lemma 3.32 holds. Let $F \subseteq V(H_G)$ be a balanced ϕ-generic set of size 6, $h \in [\prod F]$, and $U \subseteq V(H_G)$ with $|U| \leq p^{800}n/10^{4001}$. Let $F = \{a_1, a_2, b_1, b_2, c_1, c_2\}$ with $a_i \in G_A$, $b_i \in G_B$, $c_i \in G_C$. Let $k = ha_2^{-1}c_1^{-1}c_2^{-1}b_2^{-1}b_1^{-1}a_1^{-1}$ and note that by the assumption that $h \in [\prod F]$, we know that $k \in G'$. Depending on whether $|G'| \geq \log^5 n/p^{2000}$ or not, consider the set of words S given in Fig. 6 or 7, noting that blue/green/red vertices represent a partition as $S_A/S_B/S_C$. Note that in either case, all the words in S are linear and all pairs of words in S are weakly separable and words in S coming from different $S_A/S_B/S_C$ are strongly separable (see the figure captions for justification). In fact, it will be the case that every pair of words are strongly separable, meaning we never use condition (b'), so this distinction will not be essential in the justification.

For $* = A, B, C$, let $T_* = \{w^{-1}w' : w, w' \in S_*$ and w, w' do not have the same free variables$\}$, noting that $|T_*| \leq \binom{|S_*|}{2} = \binom{14}{2} = 91$ and that elements of T_* are all linear in at least one variable. If $|G'| > \log^5 n/p^{2000}$, let $U' = U \cup F$. If $|G'| \leq \log^5 n/p^{2000}$, then let U' be U together with all the self-paired cosets and all the

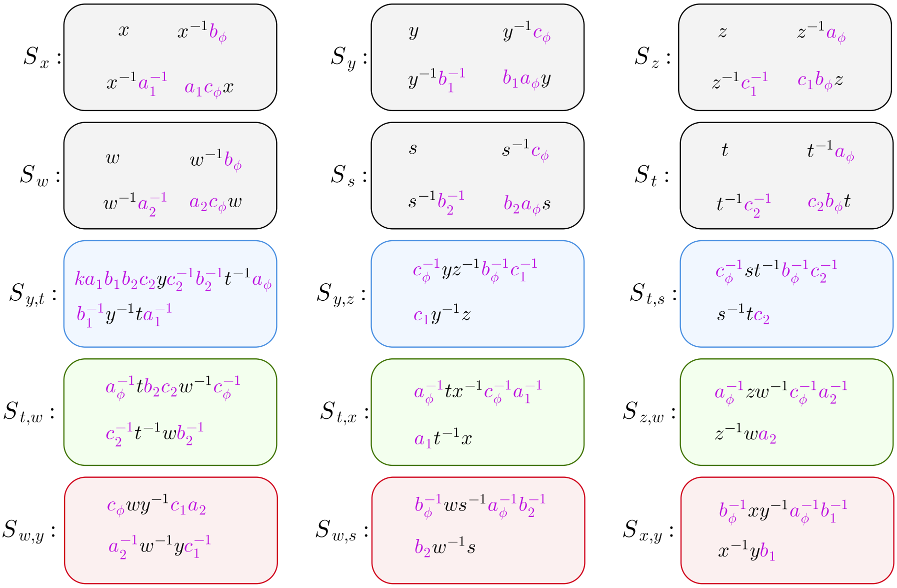
S_x : x $x^{-1}b_\phi$ $x^{-1}a_1^{-1}$ $a_1c_\phi x$
S_y : y $y^{-1}c_\phi$ $y^{-1}b_1^{-1}$ $b_1a_\phi y$
S_z : z $z^{-1}a_\phi$ $z^{-1}c_1^{-1}$ $c_1b_\phi z$
S_w : w $w^{-1}b_\phi$ $w^{-1}a_2^{-1}$ $a_2c_\phi w$
S_s : s $s^{-1}c_\phi$ $s^{-1}b_2^{-1}$ $b_2a_\phi s$
S_t : t $t^{-1}a_\phi$ $t^{-1}c_2^{-1}$ $c_2b_\phi t$
$S_{y,t}$: $ka_1b_1b_2c_2yc_2^{-1}b_2^{-1}t^{-1}a_\phi$ $b_1^{-1}y^{-1}ta_1^{-1}$
$S_{y,z}$: $c_\phi^{-1}yz^{-1}b_\phi^{-1}c_1^{-1}$ $c_1y^{-1}z$
$S_{t,s}$: $c_\phi^{-1}st^{-1}b_\phi^{-1}c_2^{-1}$ $s^{-1}tc_2$
$S_{t,w}$: $a_\phi^{-1}tb_2c_2w^{-1}c_\phi^{-1}$ $c_2^{-1}t^{-1}wb_2^{-1}$
$S_{t,x}$: $a_\phi^{-1}tx^{-1}c_\phi^{-1}a_1^{-1}$ $a_1t^{-1}x$
$S_{z,w}$: $a_\phi^{-1}zw^{-1}c_\phi^{-1}a_2^{-1}$ $z^{-1}wa_2$
$S_{w,y}$: $c_\phi wy^{-1}c_1a_2$ $a_2^{-1}w^{-1}yc_1^{-1}$
$S_{w,s}$: $b_\phi^{-1}ws^{-1}a_\phi^{-1}b_2^{-1}$ $b_2w^{-1}s$
$S_{x,y}$: $b_\phi^{-1}xy^{-1}a_\phi^{-1}b_1^{-1}$ $x^{-1}yb_1$

Fig. 6 The set S when $|G'| \leq 10^{-9}n$. Black letters represent free variables, while pink ones represent elements of G. The words are grouped into rectangles S_T based on which free variables occur where, as in Observation 3.24. To see that all words in S are linear, check that there are no repetitions of black letters in any word (and every word has at least one black letter). To see that any pair w, w' is strongly separable, note that by Observation 3.24, any w, w' coming from different rectangles fall under part (a) of the definition of separable. In the coloured rectangles, there are always two words w, w' which fall under part (c) of the definition of strongly separable (to check this first notice that w, w' have the same free variables, verifying the 2nd bullet point. The free variables occur with opposite signs in w, w' verifying the 4th bullet point. We have $\pi_0(ww') \in [a_\phi] \cup [b_\phi] \cup [c_\phi]$ so in particular $\pi_0(ww') \notin G'$, verifying the 3rd bullet point. Finally, since $\pi_0(ww') \in [a_\phi] \cup [b_\phi] \cup [c_\phi]$, there are $\leq 90|G'|$ solutions to $x^2 \in [a_\phi] \cup [b_\phi] \cup [c_\phi]$ verifying the last bullet point). This leaves the grey rectangles. In each such rectangle the four elements are v, $v^{-1}d_i^{-1}$, $v^{-1}e_\phi$, $d_i f_\phi v$ for a free variable v, $i \in \{1, 2\}$, and (d, e, f) some permutation of (a, b, c).
- The pair $w = v^{-1}d_i^{-1}$, $w' = d_i f_\phi v$ falls under (c) because $\pi_0(ww') = \pi_0(v^{-1}d_i^{-1}d_i f_\phi v) = f_\phi$.
- All other pairs fall under (b), as witnessed by the following equations $v = v^{-1}d_i^{-1}$, $v = d_i f_\phi v$, $v = v^{-1}e_\phi$, $v^{-1}d_i^{-1} = (v^{-1}e_\phi)(e_\phi^{-1}d_i^{-1})$, $v^{-1}e_\phi = (d_i f_\phi v)^{-1}(d_i f_\phi e_\phi)$. The elements g in these equations are e_ϕ, d_i, $d_i f_\phi$, $e_\phi^{-1}d_i^{-1}$, $d_i f_\phi e_\phi$, which are all generic (since d_i is ϕ-generic) (Color figure online)

cosets intersecting F, and G'. Note that in both cases, $|U'| \leq |U| + 36\log^5 n/p^{2000} \leq p^{800}n/10^{4000}$.

Use Lemma 3.32 to get a projection π with $\pi(S \cup T_A \cup T_B \cup T_C) \subseteq R \setminus (U \cup F)$ that separates $S \cup T_A \cup T_B \cup T_C$ and ensures that $S_A \subseteq R_1$, $S_B \subseteq R_2$, $S_C \subseteq R_3$ (so the matching is contained in R).

First, we prove the lemma without the “furthermore” part. Since all $w, w' \in S$ are separable, this means that all the vertices of $\pi(S)$ are distinct. Recall that $S_A/S_B/S_C$ are the blue/green/red vertices in the figures. Note that $F \cup \pi(S)$ has a partition into a matching M (given by the black triangles in Figs. 8 and 9) and a set $\{a, b, c\}$ having $abc = h$ (given by the pink triangle in the same figures where a is the top left vertex, b is the bottom vertex, and c is the top right vertex). We claim that this matching M

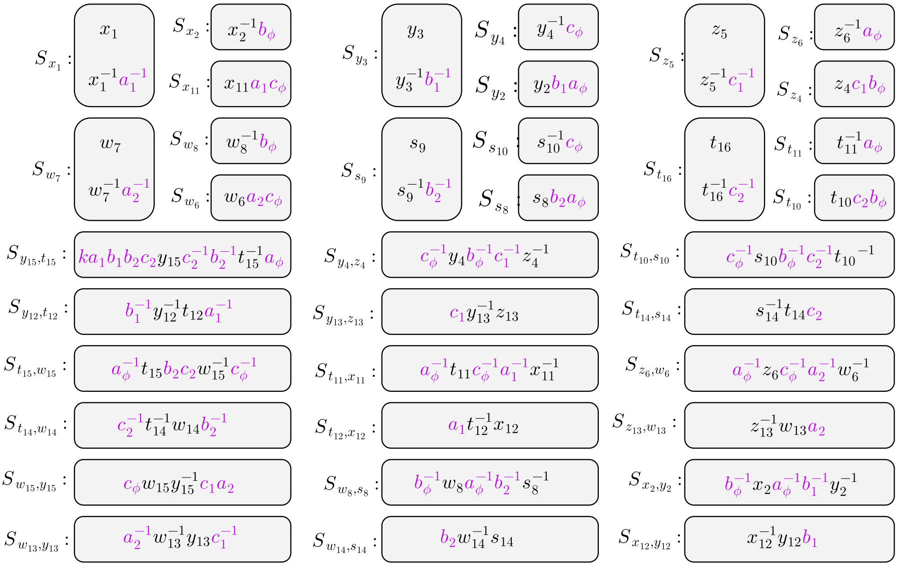
S_{x_1} : x_1 $x_1^{-1}a_1^{-1}$
S_{x_2} : $x_2^{-1}b_\phi$
$S_{x_{11}}$: $x_{11}a_1c_\phi$
S_{y_3} : y_3 $y_3^{-1}b_1^{-1}$
S_{y_4} : $y_4^{-1}c_\phi$
S_{y_2} : $y_2b_1a_\phi$
S_{z_5} : z_5 $z_5^{-1}c_1^{-1}$
S_{z_6} : $z_6^{-1}a_\phi$
S_{z_4} : $z_4c_1b_\phi$
S_{w_7} : w_7 $w_7^{-1}a_2^{-1}$
S_{w_8} : $w_8^{-1}b_\phi$
S_{w_6} : $w_6a_2c_\phi$
S_{s_9} : s_9 $s_9^{-1}b_2^{-1}$
$S_{s_{10}}$: $s_{10}^{-1}c_\phi$
S_{s_8} : $s_8b_2a_\phi$
$S_{t_{16}}$: t_{16} $t_{16}^{-1}c_2^{-1}$
$S_{t_{11}}$: $t_{11}^{-1}a_\phi$
$S_{t_{10}}$: $t_{10}c_2b_\phi$
$S_{y_{15},t_{15}}$: $ka_1b_1b_2c_2y_{15}c_2^{-1}b_2^{-1}t_{15}^{-1}a_\phi$
S_{y_4,z_4} : $c_\phi^{-1}y_4b_\phi^{-1}c_1^{-1}z_4^{-1}$
$S_{t_{10},s_{10}}$: $c_\phi^{-1}s_{10}b_\phi^{-1}c_2^{-1}t_{10}{}^{-1}$
$S_{y_{12},t_{12}}$: $b_1^{-1}y_{12}^{-1}t_{12}a_1^{-1}$
$S_{y_{13},z_{13}}$: $c_1y_{13}^{-1}z_{13}$
$S_{t_{14},s_{14}}$: $s_{14}^{-1}t_{14}c_2$
$S_{t_{15},w_{15}}$: $a_\phi^{-1}t_{15}b_2c_2w_{15}^{-1}c_\phi^{-1}$
$S_{t_{11},x_{11}}$: $a_\phi^{-1}t_{11}c_\phi^{-1}a_1^{-1}x_{11}^{-1}$
S_{z_6,w_6} : $a_\phi^{-1}z_6c_\phi^{-1}a_2^{-1}w_6^{-1}$
$S_{t_{14},w_{14}}$: $c_2^{-1}t_{14}^{-1}w_{14}b_2^{-1}$
$S_{t_{12},x_{12}}$: $a_1t_{12}^{-1}x_{12}$
$S_{z_{13},w_{13}}$: $z_{13}^{-1}w_{13}a_2$
$S_{w_{15},y_{15}}$: $c_\phi w_{15}y_{15}^{-1}c_1a_2$
S_{w_8,s_8} : $b_\phi^{-1}w_8a_\phi^{-1}b_2^{-1}s_8^{-1}$
S_{x_2,y_2} : $b_\phi^{-1}x_2a_\phi^{-1}b_1^{-1}y_2^{-1}$
$S_{w_{13},y_{13}}$: $a_2^{-1}w_{13}^{-1}y_{13}c_1^{-1}$
$S_{w_{14},s_{14}}$: $b_2w_{14}^{-1}s_{14}$
$S_{x_{12},y_{12}}$: $x_{12}^{-1}y_{12}b_1$

Fig. 7 The set S when $|G'| \geq n/10^9$. Black letters represent free variables, while pink ones represent elements of G. The words here are exactly the same as in Fig. 6, except that there are more free variables e.g. instead of the free variable x, we have free variables x_1, x_2, x_{11}, x_{12} with x replaced by one of x_1, x_2, x_{11}, x_{12} wherever it occurred. The words are grouped into rectangles S_T based on which free variables occur where, as in Observation 3.24. To see that all words in S are linear, check that there are no repetitions of black letters in every word (and every word has at least one black letter). To see that all $w, w' \in S$ are strongly separable, note that from Observation 3.24 this holds when w, w' come from different rectangles. This leaves only the pairs inside $S_{x_1}, S_{y_3}, S_{z_5}, S_{w_7}, S_{s_9}, S_{t_{16}}$, which are separable by (b) since $a_1, a_2, b_1, b_2, c_1, c_2$ are generic (Color figure online)

together with the set $\{a, b, c\}$ satisfy the lemma. The things that need to be verified by inspecting the figure are as follows.

- Each black triangle has that if its vertices are multiplied in the order blue/green/red ($G_A/G_B/G_C$), we obtain e. This shows that the black triangles are in fact edges of H_G.
- The product of the top left, bottom, and top right vertices of the yellow (central) triangle is h (here, use the definition of k). This gives that $abc = h$.
- The product of the blue/green/red edges belongs to $[a_\phi]/[b_\phi]/[c_\phi]$. In the case where $|G'| \geq 10^{-9}n$, note that we select the black (free) variables from G', hence they can be ignored while performing this check. This shows that $\{a, b, c\} \cup V(M) \setminus F$ is coset-paired.

For the "furthermore" part, we have that $|G'| \leq \log^5 n/p^{2000}$. For $* = A, B, C$, notice that for all $w, w' \in S_*$ we either have that $(\pi(w), \pi(w'))$ is a pair, or there is a free variable which appears in one of w/w', but not both (to check this, note that for any vertices w, w' of the same colour in Fig. 8, either w, w' are joined by an edge or w and w' have different combinations of black letters). In the first case we have that $[\pi(w)] \neq [\pi(w')]$ since $\pi(S)$ is disjoint from all self-paired cosets. In the second case

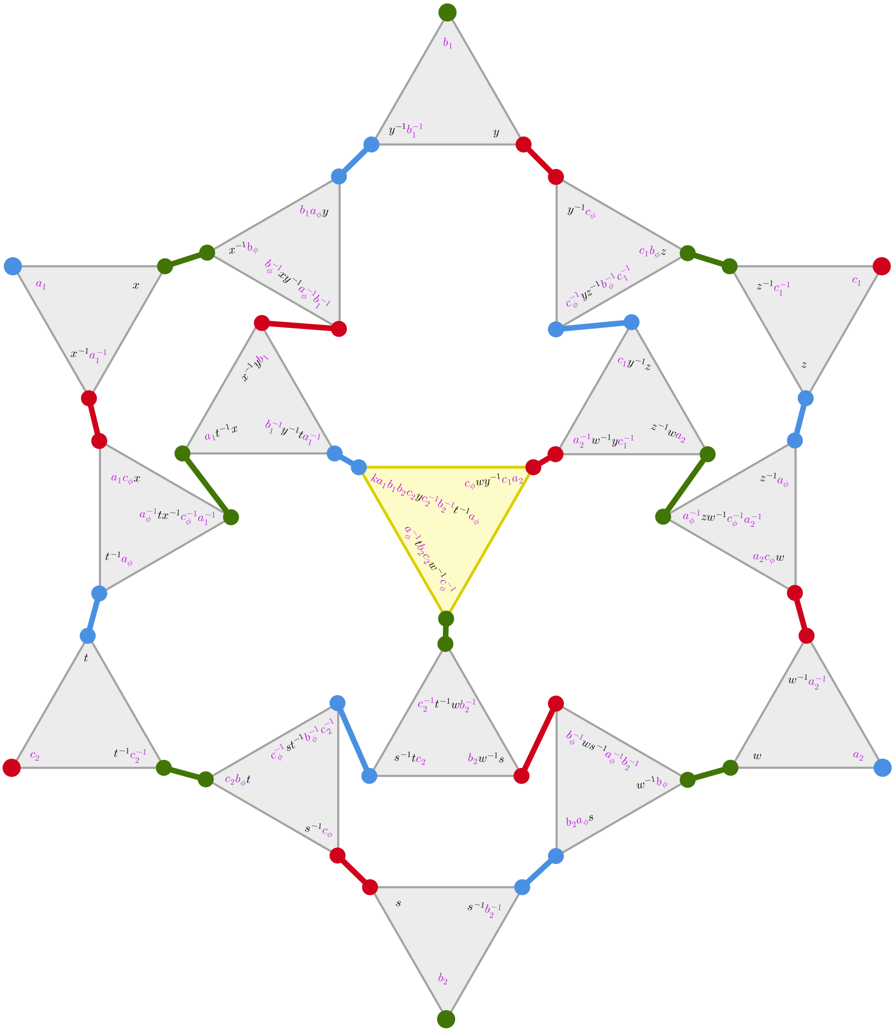

Fig. 8 The set S when $|G'| \leq 10^{-9}n$. Blue, green, and red vertices give a partition of S into sets $S_A/S_B/S_C$ which are pairwise strongly separable. Black letters represent free variables, while pink ones represent elements of G. The elements a_1, a_2, b_1, b_2, c_1, c_2 are not part of S, they are depicted only to illustrate the matching. The grey triangles are matching edges, while the yellow triangle is the set $\{a, b, c\}$. Note that the yellow triangle is an edge if and only if $k = e$ (Color figure online)

we have $w^{-1}w' \in T_*$, which implies $\pi(w^{-1}w') \notin G'$ (since $G' \subseteq U'$), or equivalently $[\pi(w)] \neq [\pi(w')]$. □

Now, we may prove the main result of this section.

Lemma 5.31 *Let $p \geq n^{-1/10^{20}}$. Let H_G be a multiplication hypergraph, R^1, R^2, R^3 disjoint, symmetric p-random subsets of G and set $R = R^1_A \cup R^2_B \cup R^3_C$. With high probability, the following holds:*

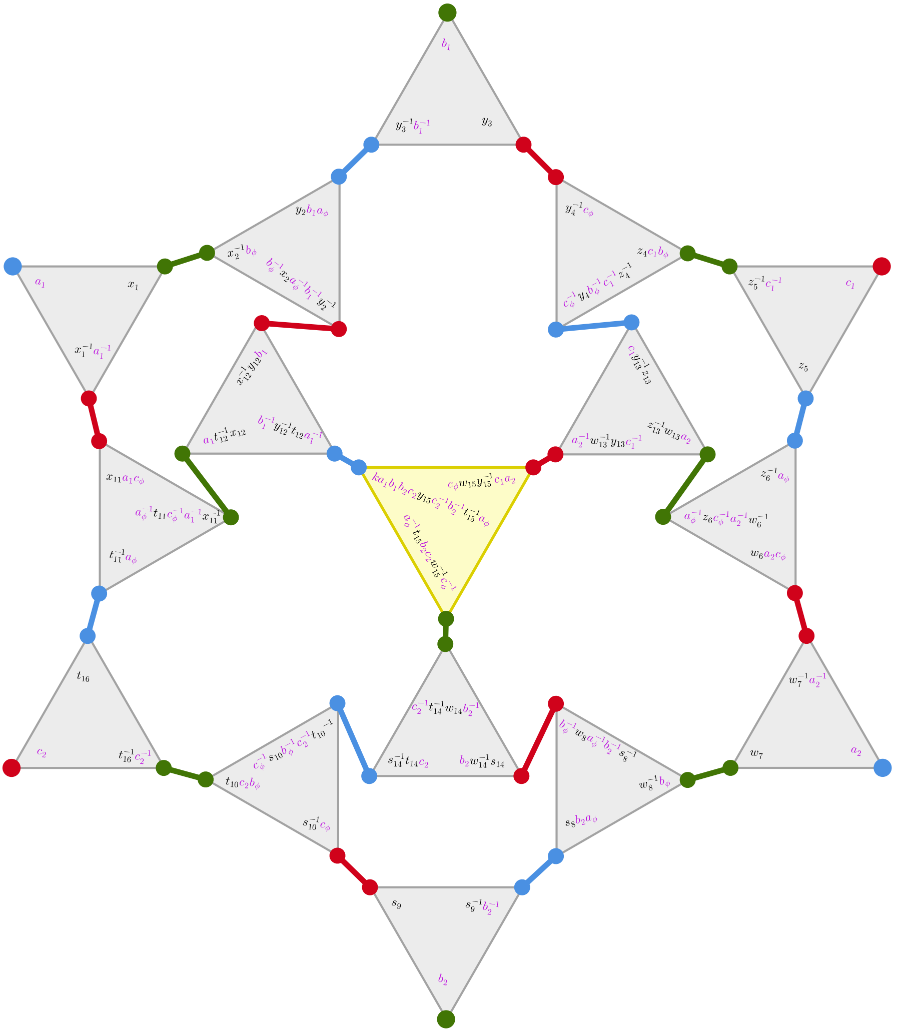
b_1
$y_3^{-1}b_1^{-1}$
y_3
$y_2b_1a_\phi$
$x_2^{-1}b_\phi$
$b_\phi^{-1}x_2a_\phi^{-1}b_1^{-1}y_2^{-1}$
$y_4^{-1}c_\phi$
$z_4c_1b_\phi$
$c_\phi^{-1}y_4b_\phi^{-1}c_1^{-1}z_4^{-1}$
a_1
x_1
$x_1^{-1}a_1^{-1}$
$z_5^{-1}c_1^{-1}$
c_1
z_5
$x_{12}^{-1}y_{12}b_1$
$a_1t_{12}^{-1}x_{12}$
$b_1^{-1}y_{12}^{-1}t_{12}a_1^{-1}$
$c_1y_{13}^{-1}z_{13}$
$a_2^{-1}w_{13}^{-1}y_{13}c_1^{-1}$
$z_{13}^{-1}w_{13}a_2$
$x_{11}a_1c_\phi$
$a_\phi^{-1}t_{11}c_\phi^{-1}a_1^{-1}x_{11}^{-1}$
$t_{11}^{-1}a_\phi$
$ka_1b_1b_2c_2y_{15}c_2^{-1}b_2^{-1}t_{15}^{-1}a_\phi$
$c_\phi w_{15}y_{15}^{-1}c_1a_2$
$a_\phi^{-1}t_{15}b_2c_2w_{15}^{-1}c_\phi^{-1}$
$z_6^{-1}a_\phi$
$a_\phi^{-1}z_6c_\phi^{-1}a_2^{-1}w_6^{-1}$
$w_6a_2c_\phi$
t_{16}
c_2
$t_{16}^{-1}c_2^{-1}$
$c_\phi^{-1}s_{10}b_\phi^{-1}c_2^{-1}t_{10}^{-1}$
$t_{10}c_2b_\phi$
$s_{10}^{-1}c_\phi$
$c_2^{-1}t_{14}^{-1}w_{14}b_2^{-1}$
$s_{14}^{-1}t_{14}c_2$
$b_2w_{14}^{-1}s_{14}$
$b_\phi^{-1}w_8a_\phi^{-1}b_2^{-1}s_8^{-1}$
$w_8^{-1}b_\phi$
$s_8b_2a_\phi$
$w_7^{-1}a_2^{-1}$
w_7
a_2
s_9
$s_9^{-1}b_2^{-1}$
b_2

Fig. 9 The set S when $|G'| > 10^{-9}n$. Blue, green, and red vertices give a partition of S into sets $S_A/S_B/S_C$ which are pairwise strongly separable. Black letters represent free variables, while pink ones represent elements of G. The elements a_1, a_2, b_1, b_2, c_1, c_2 are not part of S. This set is exactly the same as Fig. 8, except that there are more free variables. More precisely, for every grey/yellow triangle we introduce new variables and use them only in words occurring inside that triangle. The grey triangles are matching edges, while the yellow triangle is the set $\{a, b, c\}$ where a is the top left vertex, b is the bottom vertex, and c is the top right vertex (Color figure online)

Let $S \subseteq V(H_G)$ be a balanced and ϕ-generic subset with $\prod S \in G'$ and $|S| \leq \frac{p^{10^{13}}n}{10^{10^6}\log(n)^{10^8}}$. Let $U \subseteq V(H_G)$ with $|U| \leq p^{800}n/10^{4100}$. Then, there exists a matching M in H_G with the following properties.

Q1 $S \subseteq V(M)$

Q2 $V(M) \setminus S \subseteq R \setminus U$

Q3 *$V(M)\setminus S$ is coset-paired*

Q4 *If $|G'|\le \log(n)^{8000}/p^{10^{10}}$, then for each $\diamond\in\{A,B,C\}$, and $g\in G$ we have that $|(V(M)\setminus S)\cap G_\diamond\cap[g]|\le 1$.*

Proof With high probability, the property of Lemma 5.30 holds and Lemma 3.34 applies to each R^1, R^2, R^3. Let $S\subseteq V(H_G)$ be a balanced and ϕ-generic subset with $\prod S\in G'$ and $|S|\le \frac{p^{10^{13}}n}{10^{10^6}\log(n)^{10^8}}$ and $U\subseteq V(H_G)$ with $|U|\le p^{800}n/10^{4010}$. Without loss of generality, we can assume that $|S|\ge 6$ (if this doesn't hold, then pick ϕ-generic vertices $a,a',b,b',c,c'\in R\setminus(S\cup U)$ with $aa'=a_\phi$, $bb'=b_\phi$, $cc'=c_\phi$ using Lemma 3.34, and define $S'=S\cup\{a,a',b,b',c,c'\}$. Now S' is a set with $|S'|\ge 6$, so we can continue the proof with S' rather than S. (In the case when $|G'|\le \log(n)^{8000}/p^{10^{10}}$ we can include in U the $30|G'|\ll\sqrt{n}$ many elements of G in self-paired cosets before applying Lemma 3.34 to guarantee that the pairs we are adding come from distinct cosets. This way, **Q4** is not violated.)

Set $t=|S|/3$. Since S is balanced we can write $S=\{a_1,\dots,a_t,b_1,\dots,b_t,c_1,\dots,c_t\}$ with $a_i\in G_A$, $b_i\in G_B$, $c_i\in G_C$.

We build matchings $M_2,\dots,M_t$ and ϕ-generic vertices $\{a_2',\dots,a_t',b_2',\dots,b_t',c_2',\dots,c_t'\}$ as follows. Define $a_2'=a_1$, $b_2'=b_1$, $c_2'=c_1$. Given a subset $S\subseteq G$, denote by $\psi(S)$ the set of all g in G such that there exists some $s\in S$ with $[s]=[g]$. Observe that $|\psi(S)|\le|S|\cdot|G'|$. For $i=2,\dots,t-1$ apply Lemma 5.30 to $F_i=\{a_i,b_i,c_i,a_i',b_i',c_i'\}$ and

$$U=U\cup N(G)\cup\bigcup_{j<i}\psi(V(M_i))$$

in the case when $|G'|\le\log(n)^{8000}/p^{10^{10}}$ and

$$U=U\cup N(G)\cup\bigcup_{j<i}V(M_i)$$

otherwise. This way, we obtain a matching M_i and a set $\{a_{i+1}',b_{i+1}',c_{i+1}'\}$ (use an arbitrary choice of h from the corresponding G'-coset for these applications). Further, **Q4** is maintained for $\bigcup_{j\le i}M_i$ in the case when $|G'|\le\log(n)^{8000}/p^{10^{10}}$ by our inclusion in U of all previously used G'-cosets. To see that the necessary upper bound on U holds for this application, note that when $|G'|\le\log(n)^{8000}/p^{10^{10}}$, we have that $|\bigcup_{j<i}\psi(V(M_i))|\le 100t|G'|\le 100\left(\frac{p^{10^{13}}n}{10^{10^6}\log(n)^{10^8}}\right)(\log(n)^{8000}/p^{10^{10}})\ll p^{800}n/10^{4001}$. Otherwise, when $|G'|$ is large, we have that $|\bigcup_{j<i}V(M_i)|\le 100t\ll p^{800}n/10^{4001}$ as well.

Notice that $a_t'b_t'c_t'a_tb_tc_t\in G'$ (to see this, show that for all i we have $a_i'b_i'c_i'a_ib_i\times c_ia_{i+1}b_{i+1}c_{i+1}\dots a_tb_tc_t\in G'$ by induction. The initial case is just the assumption $\prod S\in G'$. For the induction step use that Lemma 5.30 gives $[a_i'b_i'c_i'a_ib_ic_i]=[a_{i+1}'b_{i+1}'c_{i+1}']$). Now apply Lemma 5.30 to $F_t=\{a_t,b_t,c_t,a_t',b_t',c_t'\}$ with $h=e\in G'$ and $U=U\cup N(G)\cup\bigcup_{j<t}V(M_i)$ in order to obtain a matching M_t and a set $\{a,b,c\}$ with $abc=e$. Let $M=\bigcup_{i=2}^t M_i\cup\{abc\}$ in order to get a matching satisfying

the lemma. Checking **Q1-Q2** is routine. To verify **Q3**, note that Lemma 5.30 tells us that $N_i := \{a'_{i+1}, b'_{i+1}, c'_{i+1}\} \cup V(M_i) \setminus F_i$ is coset-paired for $i = 2, \dots, t-1$ and also that $N_t := \{a, b, c\} \cup M_t \setminus F_t$ is coset-paired. Note that $\bigcup_{j<i} N_i \subseteq F_i \cup \bigcup_{j<i} V(M_i)$ for each i, which shows that $N_1, \dots, N_t$ are disjoint (N_i is trivially disjoint from F_i, and is disjoint from $\bigcup_{j<i} V(M_i)$ by choice of U when applying Lemma 5.30). Also, $\bigcup_{i=1}^{t} N_i = V(M) \setminus S$, which shows that this set is coset-paired as well. □

5.3.2 The zero-sum absorption lemma

Now we arrive at the main lemma of this section.

Lemma 5.32 *Let $n^{-1/10^{100}} \leq p$. Let $R^1, R^2, R^3 \subseteq G$ be disjoint, symmetric p-random subsets and set $R = R^1_A \cup R^2_B \cup R^3_C$. With high probability, the following holds.*

Let $U \subseteq G$ with $|U| \leq p^{10^{14}} n/\log(n)^{10^{14}}$. Then, there exists a subset $R' \subseteq R \setminus U$ such that for all balanced and ϕ-generic subsets $S \subseteq V(H_G) \setminus R'$ with $|S| \leq \frac{p^{10^{13}}}{10^{10^8} \log(n)^{10^8}}$, $\sum S = 0$, there exists a matching with vertex set $R' \cup S$.

We remark that the set R' in this lemma is always balanced and zero-sum. To see this, notice that the conclusion of the lemma applies with $S = \emptyset$ (since the empty set is balanced and zero-sum). This gives a matching with vertex set $R' \cup \emptyset = R'$. Since matchings are balanced and zero-sum, we get that R' is balanced and zero-sum.

Proof Let $q = \frac{p^{801}}{10^{4180} \log n}$ and $s = \frac{q^{1600} p^{10^7}}{10^{10^6} \log^{2400} n}$. For each $i \in \{1, 2, 3\}$, let Q^i, W^i be q-random and s^4-random, symmetric subsets of G respectively, satisfying $W^i \subseteq Q^i \subseteq R^i$. Note that Q^1, Q^2, Q^3 and W^1, W^2, W^3 are disjoint, symmetric q-random and disjoint, symmetric s-random subsets of G, respectively. The following all hold simultaneously with high probability.

1. Lemma 5.29 holds for R^1, R^2, R^3 and Q^1, Q^2, Q^3.
2. Lemma 5.31 holds for W^1, W^2, W^3 (observe that $s^4 \geq n^{-1/10^{20}}$).
3. If $|G'| \geq s^{-5}$, each W^i is $\lceil 2s^4|G'|\rceil$-coset-bounded. (This follows by Chernoff's bound, as $1/s \gg \log(n)^{10}$.)
4. $|\bigcup W_i| \leq 100s^4 n$, by Chernoff's bound.

Fix some $U \subseteq G$ such that $|U| \leq p^{10^{14}} n/\log(n)^{10^{14}}$ as in the statement of the lemma. Note that it also follows that $|U| \leq q^9 p^{800} n/10^{4170}$. Then, Lemma 5.29 gives us a set $U' \supseteq U$ with $|U'| \leq 500q^{-3}|U| + 10^{10}s^{-3}$ and $R' \subseteq R \setminus U$ with the property that for all coset-paired, $\lceil s|G'|\rceil$-coset-bounded, balanced, $S \subseteq Q \setminus U'$, with $|S| \leq s^2 n$ there is a matching with vertex set $R' \cup (Q \setminus (S \cup U'))$.

We claim that $R' \cup \bigcup_{i\in[3]} Q^i \setminus U'$ satisfies the property required of R' in the statement of the lemma, and the rest of the proof will justify this. Fix some set S' with the properties of S as in the statement of the lemma, that is, S' is ϕ-generic, balanced, disjoint with $R' \cup \bigcup_{i\in[3]} W^i$, $|S'| \leq \frac{p^{10^{13}}}{10^{10^8} \log(n)^{10^8}}$, and S' is zero-sum. Note it easily follows that $|S| \leq p^{4000}n/\log(n)^{10}$, and we also have that $|U'| \leq p^{800}/10^{4100}$, so we may invoke the property of W^i coming from Lemma 5.31 with $S = S'$ and $U = U'$ to deduce the existence of a matching M_1 with properties **Q1-Q4**. Then, the set $S'' := \bigcup W^i \cap (V(M_1) \setminus S')$ is coset-paired by **Q3**, $\lceil s|G'|\rceil$-coset-bounded (if $|G'| \geq s^{-5}$, this follows by (3), and if $|G'| < s^{-5} \leq \log(n)^{8000}/p^{10^{10}}$ this follows by **Q4**, as we have 1-coset-boundedness in this case), balanced, and have size at most s^2n (as $S'' \subseteq \bigcup W^i$ and (4)). By property of the set R', we have that $R' \cup (Q \setminus (S \cup U'))$ has a perfect matching, M_2 say. Then, $M_1 \cup M_2$ is a perfect matching of $(R' \cup \bigcup_{i\in[3]} Q^i \setminus U') \cup S'$, as desired. □

5.4 Proof of Theorem 4.3

Our goal in this section is to prove the main result of the paper, Theorem 4.3, as stated in Sect. 4. We first prove the following slight variant which changes the 3rd bullet point from "$e \notin X, Y, Z$" to the more restrictive condition "X, Y, Z are ϕ-generic".

Theorem 5.33 *Let $\alpha \geq n^{-1/10^{101}}$. Let G be a group of order n. Let $R^1, R^2, R^3 \subseteq G$ be p-random, α-slightly-independent subsets. Then, with high probability, the following holds.*

Let X, Y, Z be equal-sized subsets of R^1_A, R^2_B, and R^3_C respectively, satisfying the following properties.

- $|(R^1_A \cup R^2_B \cup R^3_C) \setminus (X \cup Y \cup Z)| \leq \alpha^{10^{15}}n/\log(n)^{10^{15}}$
- $\sum X + \sum Y + \sum Z = 0$ *(in G^{ab})*
- *All elements of X, Y and Z are ϕ-generic.*

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

Proof By the definition of R^1, R^2, R^3 being α-slightly-independent, we have $Q^1 \subseteq R^1$, $Q^2 \subseteq R^2$, $Q^3 \subseteq R^3$ such that the joint distribution of Q^1, Q^2, Q^3 is that of disjoint α-random subsets of G.

Partition G into a $(1/10^5)$-random subset G_- and disjoint $(1-1/10^5)$-random subset G_+, making the choices independently of $R^1, R^2, R^3, Q^1, Q^2, Q^3$. For each $i = 1, 2, 3$, let $R^i_+ = R^i \cap G_+$ to get three $(1/10^5)p$-random subsets. Use Lemma 3.17 to pick $Q^1_- \subseteq Q^1 \cap G_-$, $Q^2_- \subseteq Q^2 \cap G_-$, $Q^3_- \subseteq Q^3 \cap G_-$, so that their joint distribution is that of disjoint, symmetric $\alpha/10^{10}$-random subsets of G. Let $Q_- := (Q^1_-)_A \cup (Q^2_-)_B \cup (Q^3_-)_C$, $R := R^1_A \cup R^2_B \cup R^3_C$, and $R_+ := (R^1_+)_A \cup (R^2_+)_B \cup (R^3_+)_C$. With high probability, the property in Lemma 5.32 holds for Q_-. As the multiplication hypergraph H_G is $(0, 1, n)$-typical, with high probability, the property in Lemma 3.9 holds with the random sets $(A', B', C') = (R^1_+, R^2_+, R^3_+)$. Finally, as a consequence of Chernoff's bound, with high probability, $|Q_-| \leq |R_+|/100$, and for i we have $|R^i_+| = (p(1-/10^5) \pm n^{-0.27})n$ and $|R^i|(p \pm n^{-0.27})n$

Now, given subsets X, Y, Z, define $U := (R^1_A \cup R^2_B \cup R^3_C) \setminus (X \cup Y \cup Z)$ and note $|U| \leq \alpha^{10^{15}}n/\log(n)^{10^{15}}$. From the property in Lemma 5.32, we find a set $Q'_- \subseteq Q_- \setminus U$ that can combine with balanced zero-sum sets to produce perfect matchings. By the remark after Lemma 5.32, Q'_- is balanced and zero-sum.

Let $L := R \setminus (Q'_- \cup R_+)$. Note that as R^1_+, R^2_+, R^3_+ each have size $(p(1-/10^5) \pm n^{-0.27})n$, and Q'_- is a balanced set contained in R, there exists a $q \leq p/10^5$ such that

$L \cap X$, $L \cap Y$, $L \cap Z$ each have size $(q \pm n^{-0.26})n$. Note that $q \leq (1 - 10^{-5})p/100$, so we may apply Lemma 3.9, to conclude that there exists a matching M_1' with $V(M_1') \subseteq L \cup R_+$ covering all but $n^{1-10^{-4}}$ vertices of $L \cup R_+$. Of the edges forming this matching, at most $\alpha^{10^{15}} n/\log(n)^{10^{15}}$ many of them can meet U, hence we find a matching M_1 of $H_G[X, Y, Z]$ covering all but at most $\alpha^{10^{15}} n/\log(n)^{10^{15}} + n^{1-10^{-4}} \leq (\alpha/10^{10})^{10^{14}}/\log(n)^{10^{14}}$ vertices of $(X \cup Y \cup Z) \setminus Q'_{-}$. Call this uncovered set of vertices S.

We claim that S is a balanced zero-sum set. This is as the sets Q'_{-}, $V(M_1)$, and $X \cup Y \cup Z$ are each balanced and zero-sum, and S is obtained by removing all of the former two sets from the latter set, and the former two sets are disjoint. Further, S is ϕ-generic, as by assumption, $X \cup Y \cup Z$ is ϕ-generic. Then, by property of the set Q'_{-}, $S \cup Q'_{-}$ spans a perfect matching, M_2 say. $M_1 \cup M_2$ then is the desired perfect matching of $X \cup Y \cup Z$. □

It remains to prove Theorem 4.3. We will need the following intermediary result.

Lemma 5.34 *Let $p \geq n^{-1/10^{100}}$. Let G be a group of order n. Let $Q^1, Q^2, Q^3 \subseteq G$ be disjoint, symmetric p-random subsets and set $Q = Q^1_A \cup Q^2_B \cup Q^3_C$. With high probability, the following holds.*

Let $S, U \subseteq V(H_G)$ with $|S|, |U| \leq p^{10^{10}} n/\log(n)^{10^{10}}$ and $e \notin S$. Then, there exists a matching M in $H_G[Q] \setminus U$ of size at most $2p^{10^{10}} n/\log(n)^{10^{10}}$ saturating S, meaning $S \subseteq V(M)$.

Proof Let $g \in G$ be an arbitrary element of the group such that $g \neq e$. Then, by Proposition 3.12, it follows that g is contained in at least $n/5$ edges $\{g, x, y\}$ such that $x \notin \{y, y^{-1}\}$, regardless of whether $g \in G_A$, G_B, or G_C. For any i, j, and any such $\{x, y\}$, $x \in Q^i$ and $y \in Q^j$ with probability p^2. For each of the $n/5$ such edges, the corresponding pairs $\{x, y\}$ are disjoint. Hence, we may apply Chernoff's bound to deduce that with probability at least $1 - 1/n^2$, g has degree at least $p^2 n/1000$ in $H_G[Q^1_A, Q^2_B, Q^3_C]$, as long as $g \in V(H_G[Q^1_A, Q^2_B, Q^3_C])$. By a union bound, this property holds for all $g \in G$ such that $g \neq e$ in the case that $G = (\mathbb{Z}_2)^k$ for some k.

Now, let S, U be given. By the minimum degree property from the previous paragraph, all elements of S have degree at least $p^2 n/1000 - 2p^{10^{10}} n/\log(n)^{10^{10}} \geq p^2 n/2000$ in $H_G[Q] \setminus U$. As $p^2 n \gg |S|$, we may greedily pick a matching saturating S and disjoint from U of size $|S| \leq 2p^{10^{10}} n/\log(n)^{10^{10}}$ as claimed. □

We can now give the final proof of the section.

Proof of Theorem 4.3 from Theorem 5.33 and Lemma 5.34 By the definition of q-slightly-independent, we have disjoint, symmetric, q-random sets Q^1, Q^2, Q^3 with $Q^1 \subseteq R^1$, $Q^2 \subseteq R^2$, $Q^3 \subseteq R^3$. With high probability Theorem 5.33 holds R^1, R^2 and R^3 and Lemma 5.34 holds for Q^1, Q^2, Q^3. Given X, Y, Z, let $S = (X \cup Y \cup Z) \setminus (R^1_A \cup R^2_B \cup R^3_C)$ together with the non-ϕ-generic elements of $X \cup Y \cup Z$ (of which there are at most 10^{9010} many). Let $U = (R^1_A \cup R^2_B \cup R^3_C) \setminus (X \cup Y \cup Z)$. Note that $|S|, |U| \leq p^{10^{16}} n/\log(n)^{10^{16}}$ and $e \notin S$ in the case that $G = \mathbb{Z}_2^k$. Apply

Lemma 5.34 to find a matching M, and set X', Y', Z' to be $X \setminus V(M)$, $Y \setminus V(M)$, $Z \setminus V(M)$ respectively. Observe that X', Y', Z' satisfy the hypothesis of X, Y, Z in Theorem 4.3. Indeed, $X'/Y'/Z'$ is contained in $R^1/R^2/R^3$ and does not contain non-ϕ-generic elements since all such elements are in $S \subseteq V(M)$. We have $\sum X' + \sum Y' + \sum Z' = 0$, as $\sum X + \sum Y + \sum Z = 0$, and $\sum V(M) = 0$ (as $V(M)$ spans a perfect matching). We have $|X'| = |Y'| = |Z'|$ for the same reason. Hence, $H_G[X', Y', Z']$ spans a perfect matching by Theorem 4.3, and this together with M gives a perfect matching of $H_G[X, Y, Z]$ as desired. □

6 Applications

6.1 Characterising subsquares with transversals

The goal of this section is to prove a far-reaching generalisation of Snevily's conjecture as stated in the introduction (Theorem 1.5).

6.1.1 Preliminaries

The below lemma allowing us to find zero-sum sets of prescribed size will also be useful in Sect. 6.3.

Lemma 6.1 *Let $k \in \{3, 4, 5\}$. Let $p \geq 3n^{-1/1400}$. Let G be an abelian group, and let R be a p-random subset of G. Then, the following holds with high probability.*

Let $X \subseteq R$ with $|R \setminus X| \leq p^{8000}n/10^{6000}$. Then, there exists at least $p^{800}n/10^{4010}$ many disjoint zero-sum sets of size k in X.

Proof With high probability, Corollary 3.33 applies to R. If it is the case that $k = 3$, consider the set $S = \{xy, x^{-1}, y^{-1}\} \subseteq G * F_2$ and note that all pairs of words in S are linear and separable (by part (a) of the definition of separable). Thus, setting $U := R \setminus X$ there is a projection $\pi : G * F_2 \to G$ which has $\pi(S) \subseteq X$ and which separates S. Resetting U to include the vertices in the projection, and invoking Corollary 3.32 iteratively, we can conclude that there exists at least $(p^{800}n/10^{4000} - p^{8000}n/10^{6000})/10 \gg p^{800}n/10^{4010}$ disjoint zero-sum sets of size 3 contained in X. If $k = 4$ set $S = \{xyz, x^{-1}, y^{-1}, z^{-1}\}$, and if $k = 5$ set $S = \{xyzw, x^{-1}, y^{-1}, z^{-1}, w^{-1}\}$ to obtain sets of words which are linear and pairwise strongly separable, and proceed in the same way as the case $k = 3$. □

For this section, we will only need the following easy corollary of Lemma 6.1

Lemma 6.2 *Let G be a group, and let $3 \leq t \leq n/3 - 1$, and let $g \in G$. Then, there exists distinct $c_1, \ldots, c_t \in G$ such that $\sum c_i = g$ in G^{ab}.*

Proof Let $c_4, \ldots, c_t$ be arbitrary distinct elements of G, and set $g' = g(c_4 \cdots c_t)^{-1}$. Observe there exists at least $n^2 - 3n$ choices of c_1, c_2 and c_3 such that $c_1c_2c_3 = g'$ and c_1, c_2, c_3 are distinct. Among those choices, at most $3tn$ of them have $\{c_1, c_2, c_3\} \cap \{c_4, \ldots, c_t\} \neq \emptyset$. This implies that if $n^2 - 3n - 3tn > 0$, i.e. if $t \leq n/3 - 1$, there exists a choice of $c_1, \ldots, c_t$ with the desired properties. □

The below result has essentially appeared in [47]. For our application here, we provide a more precise formulation as well as a full proof for the sake of completeness. We remark that the following proof was suggested to us by an anonymous referee.

Lemma 6.3 *Let n be sufficiently large and set $\gamma := 1/\log(n)^{10^{100}}$. Let $S = A \times B$ be a subsquare of a multiplication table of a group G defined by two n-element sets $A, B \subseteq G$. Then, either S has at most $(1-\gamma)n$ symbols occurring more than $(1-\gamma)n$ times or there is a subgroup $H \subseteq G$ and elements $g, g' \in G$ such that $|A \Delta gH|, |B \Delta Hg'| \leq \gamma^{1/10}n$.*

Proof We will use a theorem of Fournier as stated as Theorem 1.3.3 in the lecture notes of Green [36]. Towards this goal, we recall the definition of **multiplicative energy** of a subset A of a group. It is defined to be the quantity $E(A) := |\{(a_1, a_2, b_1, b_2) \in A^4 \colon a_1a_2^{-1} = b_1b_2^{-1}\}|$. We set $E(A, B) := |\{(a_1, a_2, b_1, b_2) \in A^2 \times B^2 \colon a_1a_2^{-1} = b_1b_2^{-1}\}|$, note that $E(A) = E(A, A)$. Denote by $r_A(x) := \{(a_1, a_2) \in A^2 \colon a_1a_2^{-1} = x\}$, and $r_B(x)$ is defined analogously. Then, by the Cauchy-Schwarz inequality, we have that

$$E(A^{-1}, B) = \sum_{x \in G} r_{A^{-1}}(x) r_B(x) \leq E(A^{-1})^{1/2} E(B)^{1/2}.$$

Note also that $E(A^{-1}, B) := \{(a_1, a_2, b_1, b_2) \in A^2 \times B^2 \colon a_1b_1 = a_2b_2\}$. Suppose now that S has more than $(1-\gamma)n$ symbols occurring more that $(1-\gamma)n$ times. Thus, $E(A^{-1}, B) \geq (1-2\gamma)^3n^3$. As $E(A^{-1}), E(B) \leq n^3$, we can deduce using the above that $E(A^{-1}), E(B) \geq (1-2\gamma)^6n^3$. From Theorem 1.3.3 in [36], it follows that there exists a subgroup H and $g \in G$ such that $|A \Delta gH| \leq \gamma^{1/9}n$. Consider an element t of S that repeats at least $(1-\gamma)n$ times. So we have $a_1b_1 = \cdots = a_{(1-\gamma)n}b_{(1-\gamma)n} = t$. Let $i \in [(1-\gamma)n]$ so that $a_i \in gH$, and note that there are at least $(1-2\gamma^{1/9})n$ indices i with this property. So we have that $t = (gh)b_i$ for some $h \in H$. Then, $b_i = h^{-1}g^{-1}t$, so $b_i \in H(g^{-1}t)$ for each such i. Then, the statement holds for $g' = g^{-1}t$. □

Observation 6.4 *Let $A \times B$ be a subsquare of a multiplication table of a group G defined by two n-element sets $A, B \subseteq G$. Let $g, g' \in G$. Suppose that $A \times B$ contains a transversal. Then, $gA \times Bg'$ also contains a transversal.*

Proof Let $(a_i, \phi(a_i))$, $i \in [n]$ be a transversal of $A \times B$, where ϕ is a bijection $A \to B$. Then, $(ga_i, \phi(a_i)g')$ is a transversal of $gA \times Bg'$, since $a_i\phi(a_i) \neq a_j\phi(a_j)$ implies that $ga_i\phi(a_i)g' \neq ga_j\phi(a_j)g'$. □

Lemma 6.5 *Let $\varepsilon > 0$ be sufficiently small and let n be sufficiently large. Let $\gamma \geq n^{-\varepsilon}$. Let S be a subsquare of a group G of order n such that S has at most $(1-\gamma)n$ symbols occuring more than $(1-\gamma)n$ times. Then, S contains a transversal.*

Proof As subsquares of G correspond to edge-coloured balanced complete graphs, this lemma is a direct corollary of Lemma 8.1.3 from [47]. □

6.1.2 The characterisation

We first prove the following weakening of the main theorem which characterises when subsquares which are close to the whole multiplication table have transversals.

Lemma 6.6 *Let $t \le n/\log(n)^{10^{30}}$. Let G be a sufficiently large group and $A, B \subseteq G$ with $|A| = |B| = n - t$. Then $A \times B$ has a transversal unless one of the following holds.*

1. *G is a group that does not satisfy the Hall-Paige condition, and $A = B = G$.*
2. *$G \cong (\mathbb{Z}_2)^k$, $A = G \setminus \{a_1, a_2\}$, $B = G \setminus \{b_1, b_2\}$ for some distinct $a_1, a_2 \in G$ and distinct $b_1, b_2 \in G$ such that $a_1 + a_2 + b_1 + b_2 = 0$.*

Proof If $t = 0$ then there is a transversal (by any proofs of the Hall-Paige Conjecture e.g. by Theorem 4.4 with $p = 1$). Thus, we can suppose that $1 \le t \le n/\log(n)^{10^{30}}$. Let $\alpha := \prod G(\prod A \prod B)^{-1}$ (where the terms in the products are multiplied in any fixed order).

Claim 6.6.1 *There are distinct elements $c_1, \dots, c_t \in G$ such that $c_1 + \cdots + c_t = \alpha$ in G^{ab}.*

Proof If $t = 1$, one can trivially choose $c_1 = \alpha$. If $t = 2$, and $\alpha \neq e$, then we could select $c_1 = e$ and $c_2 = \alpha$. The case $t = 2$, $\alpha = e$, $G = (\mathbb{Z}_2)^k$ is excluded by the hypothesis. Indeed, in this case we'd have $A = G \setminus \{a_1, a_2\}$, $B = G \setminus \{b_1, b_2\}$ for some distinct $a_1, a_2 \in H$ and distinct $b_1, b_2 \in G$. But then in G^{ab} we'd have $0 = \alpha = \sum G - \sum A - \sum B = a_1 + a_2 + b_1 + b_2 - \sum(\mathbb{Z}_2)^k = a_1 + a_2 + b_1 + b_2$. When $t = 2$, $\alpha = e$, $G \neq (\mathbb{Z}_2)^k$, then, there must exist distinct $c_1, c_2 \in G$ such that $c_1 c_2 = e$. Finally, when $t \ge 3$, then the claim follows from Lemma 6.2. □

Let $C = G \setminus \{c_1, \dots, c_t\}$ to get a set with $|C| = |A| = |B|$. Now Theorem 4.4 applies with $p = 1$ to give a perfect matching $\{(a_1, b_1, c_1), \dots, (a_{n-1}, b_{n-1}, c_{n-1})\}$ in $H_G[A, B, C^{-1}]$. The entries $(a_1, b_1), \dots, (a_{n-t}, b_{n-t})$ give a transversal in $A \times B$. □

We now prove the main result of this section.

Theorem 6.7 *There exists a $n_0 \in \mathbb{N}$ such that the following holds for all $n \ge n_0$. Let G be a group, and let $A, B \subseteq G$ with $|A| = |B| = n$. Then, $A \times B$ has a transversal, unless there exists some $k \ge 1$, $g_1, g_2 \in G$ and a subgroup $H \subseteq G$ such that one of the following holds.*

1. *H is a group that does not satisfy the Hall-Paige condition, and $A = g_1 H$ and $B = H g_2$.*
2. *$H \cong (\mathbb{Z}_2)^k$, $g_1 A = H \setminus \{a_1, a_2\}$, $g_2 B = H \setminus \{b_1, b_2\}$ for some distinct $a_1, a_2 \in H$ and distinct $b_1, b_2 \in H$ such that $a_1 + a_2 + b_1 + b_2 = 0$.*

Proof of Theorem 6.7 Set $\gamma = 1/\log(n)^{10^{100}}$. Suppose that $A \times B$ does not have a transversal. Then, by Lemma 6.5, we may assume that $A \times B$ contains more than

$(1-\gamma)n$ symbols occuring more than $(1-\gamma)n$ times. By Lemma 6.3, it follows that there is a subgroup $H \subseteq G$ and elements $g, g' \in G$ such that $|A \Delta gH|, |B \Delta Hg'| \leq \gamma^{1/10}n$. As $A \times B$ does not contain a transversal, by Observation 6.4, $g^{-1}A \times Bg'^{-1}$ does not contain a transversal either. Set $A' \times B' = g^{-1}A \times Bg'^{-1}$ and observe that $|A'\Delta H|, |B'\Delta H| \leq \gamma^{1/10}n$.

Set $A_1 = A' \cap H$, $A_2 = A' \setminus H$, $B_1 = B' \cap H$, $B_2 = B' \setminus H$, noting that $|A_2|, |B_2| \leq \gamma^{1/10}n$. If $A_2 = B_2 = \emptyset$, then the theorem follows from Lemma 6.6 applied with $G = H$. Thus we can suppose that A_2 and/or B_2 are nonempty. Note that all the elements in the multiplication table in $A_1 \times B_2$ and $A_2 \times B_1$ are outside H. Let $A_2 = \{a_1, \dots, a_{|A_2|}\}$, $B_2 = \{b_1, \dots, b_{|B_2|}\}$. We can greedily select a partial transversal $T_1 = \{(a_1, b'_1), \dots, (a_{|A_2|}, b'_{|A_2|}), (a'_1, b_1), \dots, (a'_{|B_2|}, b_{|B_2|})\}$ by selecting elements $b'_1, \dots, b'_{|A_2|} \in B_1, a'_1, \dots, a'_{|B_2|} \in A_1$ in order (to see this note that there are at least $\min(|A_1|, |B_1|) \geq n/2$ choices for each element and so there's room to avoid the $|A_2| + |B_2| \leq \gamma^{1/10}n$ rows/columns/symbols previously used). Note that since there are at least $n/4$ choices for the last element $a'_{|B_2|}$, we can additionally ensure that $\sum A_1 \setminus \{a'_1, \dots, a'_{|B_2|}\} + \sum B_1 \setminus \{b'_1, \dots, b'_{|A_2|}\} \neq 0$ in H^{ab} in the case where H^{ab} has at least 100 elements. Thus Lemma 6.6 applies to give a transversal T_2 in $(A_1 \setminus \{a'_1, \dots, a'_{|B_2|}) \times (B_1 \setminus \{b'_1, \dots, b'_{|A_2|}\})$ (to apply Lemma 6.6, we need to know that we are not in cases 1 and 2. We're not in case 1 because we're assuming $A_2 \cup B_2 \neq \emptyset$. We're not in case 2 because in this case we have $|H^{ab}| \geq 100$, and also that $\sum A + \sum B = 0$, and we selected a'_i and b'_i to avoid this scenario). Now $T_1 \cup T_2$ is a transversal in $A' \times B'$ as required. □

6.2 Path-like structures in groups

The goal of this section is to give a characterisation of sequenceable, R-sequenceable, and harmonious groups which are sufficiently large. It will be convenient to rephrase all three of these problems as finding rainbow structures in edge-coloured digraphs. Towards this aim, we give some definitions.

Given a group G, by K_G^+ we denote the complete directed edge-coloured graph with vertex set G, edge set $\{\vec{ab}\colon a \neq b \text{ and } a, b \in G\}$, where the edge $\vec{ab}$ gets assigned the colour $ab \in G$. We call this the **multiplication digraph of** G. Similarly, by K_G^- we denote the **division digraph of** G. In the division digraph, the edge $\vec{ab}$ gets assigned the colour $a^{-1}b \in G$, and all other properties of K_G^- are same with those of K_G^+. We sometimes use the notation $K_G^\pm$ to make statements and definitions about K_G^+ and K_G^- simultaneously. For subsets $R, R' \subseteq G$, we will use $K_G^\pm[R; R']$ to denote the subgraph of $K_G^\pm$ induced on vertex set R consisting of all edges of colours in R'. For disjoint subsets $V_1, V_2 \subseteq G$, by $K_G^\pm[V_1, V_2; R']$ we denote the bipartite subgraph of $K_G^\pm$ obtained by keeping only the edges between V_1 and V_2 with colour in R'.

Recall that a subgraph of an edge-coloured graph is called **rainbow** if all edges have distinct colours. The definitions of sequenceable, R-sequenceable, and harmonious were given in Sect. 1.1. The following is straightforward to derive from the definitions.

Observation 6.8 *A group G is harmonious if and only if K_G^+ has a directed rainbow Hamilton cycle, sequenceable if and only if K_G^- has a directed Hamilton path with colour set $G \setminus e$, and R-sequenceable if and only if K_G^- has a directed rainbow cycle with colour set $G \setminus e$.*

The main trick in this section is to use our main theorem iteratively to build rainbow paths out of rainbow matchings. One key issue with this idea is that this does not give us to freedom to construct paths connecting specified end-points, which is critical for building Hamilton paths, instead of an arbitrary path/cycle-factor. To remedy this, in Sect. 6.2.1, based on ideas of Kühn, Lapinskas, Osthus, and Patel [44], we introduce a way of building path systems allowing us to construct path-factors with specified end-points. The key result of that section, combined with a variant of our main theorem, allows us to deduce the following theorem.

Theorem 6.9 *Let G be a sufficiently large group on n vertices. Let $V, C \subseteq G$ and $x, y \in G$ be such that $|V| + 1 = |C| \geq n - n^{1/2}$, $x \neq y$, $x, y \notin V$, and further suppose that $e \notin C$ if G is an elementary abelian 2-group. Then, $K_G^-[\{x, y\} \cup V; C]$ has a directed rainbow Hamilton path from x to y if $\sum C = y - x$ in G^{ab}, and $K_G^+[\{x, y\} \cup V; C]$ has a directed rainbow Hamilton path from x to y if $\sum C = x + y + 2\sum V$ in G^{ab}.*

This implies the following characterisation of which rainbow Hamilton paths can be found in K_G^+.

Corollary 6.10 *Let G be a sufficiently large group not isomorphic to an elementary abelian 2-group. Let $c, x_1, x_2 \in G$ such that $0 = \sum G + c - x_1 - x_2$ in G^{ab}, and $x_1 \neq x_2$. Then, K_G^+ contains a directed rainbow Hamilton path using the colour set $G - c$ and with endpoints x_1 and x_2.*

Proof This is immediate by applying Theorem 6.9 with $V := G \setminus \{x_1, x_2\}$ and $(x, y) = (x_1, x_2)$ and $C = G \setminus \{c\}$. □

This gives a characterisation of groups where K_G^+ has a directed rainbow Hamilton cycle which is equivalent to G being harmonious.

Corollary 6.11 *Let G be a sufficiently large group satisfying the Hall-Paige condition, and suppose G is not isomorphic to an elementary abelian 2-group. Then, K_G^+ has a directed rainbow Hamilton cycle, i.e. G is harmonious.*

Proof Let x_1 and x_2 be such that $x_1 x_2 = e$ and $x_1 \neq x_2$. Such x_1 and x_2 exist as G is not isomorphic to an elementary abelian 2-group. By assumption, $\sum G = e$ in G^{ab}. Then, $0 = e - x_1 - x_2 = \sum G + e - x_1 - x_2$ in G^{ab}, hence by Corollary 6.10 we have a directed rainbow Hamilton path from x_1 to x_2 using all colours but e. Combined with the edge $x_2 \to x_1$, this gives the desired directed rainbow Hamilton cycle. □

We now characterise large sequenceable groups. Note that for abelian groups, the below result was proved by Gordon [32] (with no assumption on the size of the group).

Theorem 6.12 *Let G be a sufficiently large group. If G is abelian, suppose that $\sum G \neq 0$, or equivalently, G has a unique element of order 2. Then, G is sequenceable.*

Proof First suppose G is abelian. Let $k \neq 0$ be the unique element of order 2 in G. Apply Theorem 6.9 with $* = -$, $C = G \setminus \{0\}$, $(x, y) = (0, k)$, $V = G \setminus \{0, k\}$, noting this is possible as $y - x - k - 0 = k = \sum G = \sum C$. This gives us the desired directed rainbow Hamilton path with colour set $G \setminus \backslash\{0\}$.

If G is nonabelian, let y be such that $y = \sum G$ in G^{ab} and $y \neq e$. Such a y exists since each coset of the commutator subgroup has at least 2 elements. We can now invoke Theorem 6.9 with $C = G - e$, $(x, y) = (e, y)$ and $V = G - e - y$, this is possible as $y - e = \sum C = \sum G$ in G^{ab} by choice of y. This again gives us the desired directed rainbow Hamilton path. □

Using similar ideas, we characterise R-sequenceable groups.

Theorem 6.13 *Let G be a sufficiently large group satisfying the Hall-Paige condition, that is, $\sum G = 0$ in G^{ab}. Then, G is R-sequenceable.*

Proof Let x, y be two distinct elements of the group G, and apply Theorem 6.9 (the K_G^- case) with $C = G \setminus \{e, y^{-1}x\}$, $(x, y) = (x, y)$ and $V = G \setminus \{x, y, e\}$, this application is valid as $\sum G = 0$. This gives a path from x to y, and combined with the edge from $y \to x$, we obtain a directed rainbow cycle in K_G^- using all colours but e, meaning that G is R-sequenceable. □

In the rest of this section, we are focused on proving Theorem 6.9.

6.2.1 Sorting networks

In [44] (see in particular Lemma 4.3), an ingenious method was introduced in order to construct path systems which can connect specified endpoints. The key idea is to use an appropriate *sorting network* as a template while building the path system. In this section, we adapt the arguments from [44] to our context. First, we introduce some terminology. For a more detailed treatment, we refer the reader to [18].

Definition 6.14 A **comparison network** is a union of four types of objects: input nodes $x_1, \ldots, x_m$, output nodes $y_1, \ldots, y_m$, comparators $C_1, \ldots, C_t$ and wires $w_1, \ldots, w_s$.

- Comparators are sets of 4 nodes $C_i = \{y_i^-, y_i^+, x_i^-, x_i^+\}$ (which are disjoint with the input and output nodes).
- Each wire joins an x-node to a y-node. Additionally, each node is in precisely one wire, and the directed graph formed by contracting comparators into single nodes is acyclic.

Definition 6.15 A **comparison sorting network** is a comparison network with the following additional property. Let σ be any permutation of $[m]$. Assign each x_i the value $v(x_i) = \sigma(i)$. Assign the values of the other nodes via the following rules.

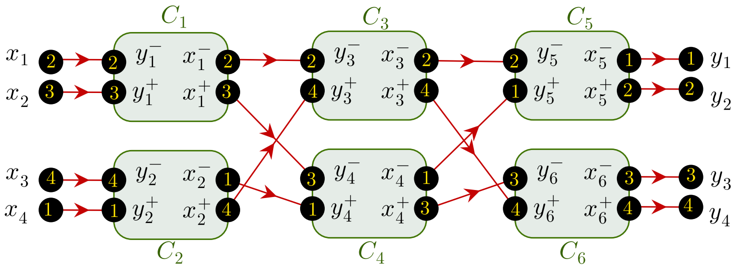
C_1
C_3
C_5
x_1 2 2 y_1^- x_1^- 2 2 y_3^- x_3^- 2 2 y_5^- x_5^- 1 1 y_1
x_2 3 3 y_1^+ x_1^+ 3 4 y_3^+ x_3^+ 4 1 y_5^+ x_5^+ 2 2 y_2
x_3 4 4 y_2^- x_2^- 1 3 y_4^- x_4^- 1 3 y_6^- x_6^- 3 3 y_3
x_4 1 1 y_2^+ x_2^+ 4 1 y_4^+ x_4^+ 3 4 y_6^+ x_6^+ 4 4 y_4
C_2
C_4
C_6

Fig. 10 A comparison sorting network for sorting four numbers. The black circles represent nodes, the green rectangles are comparators, and the red arrows are wires. Here the network was given input values $v(x_1) = 1$, $v(x_2) = 3$, $v(x_3) = 4$, $v(x_4) = 1$ (represented by the yellow numbers inside those nodes). Then all other nodes get a value based on the rules in Definition 6.15 (represented by the yellow numbers in the other nodes). The network correctly sorted the numbers, which can be seen by the fact that each y_i contains yellow number i (Color figure online)

1. If xy is a wire then $v(y) := v(x)$.
2. If $C_i = \{y_i^-, y_i^+, x_i^-, x_i^+\}$ is a comparator, then $v(x_i^-) := \min(v(y_i^-), v(y_i^+))$ and $v(x_i^+) := \max(v(y^-), v(y^+))$.

Then, all nodes get assigned a value and moreover, $v(y_i) = i$ for $i = 1, \dots, m$.

See Fig. 10 for an example of these definitions. A classical result due to Batcher [8] states that for all $m \in \mathbb{N}$, there is a sorting network with m input/outputs and $100m \log^2 m$ comparators. In fact, there are sorting networks with $O(m \log m)$ comparators thanks to a celebrated result of Ajtai, Komlós, and Szemerédi [1] but we will not need this sharper bound here. However, it will be convenient for us to have sorting networks with symmetry in the following sense.

Lemma 6.16 ([1]) *For all $m \in \mathbb{N}$, there is a sorting network such that the length of every path from x_i to y_j is exactly $\lceil 100 \log^2 m \rceil$ (in the directed graph formed by contracting every comparator into a single node).*

The above can be proved by inspecting any common method of constructing a sorting network, for example the method of Batcher [8]. Indeed, the bound of Batcher is in terms of the *depth* of the network as opposed to the total number of comparators, so we can simply add redundant comparators to ensure the conclusion of Lemma 6.16.

We now show how to simulate the task of a comparator in a sorting network via a collection of paths.

Lemma 6.17 *Let $p \geq n^{-1/700}$. Let R^1, R^2 be p-random subsets of G, sampled independently.*

With high probability, for any $U \subseteq G$ with $|U| \leq p^{800}n/10^{4010}$, there is a subgraph $C \subseteq K_G^{\pm}[R^1 \setminus U; R^2 \setminus U]$ consisting of 12 vertices and 10 colours containing vertices x^-, x^+, y^-, y^+ and directed paths Q_{x^-,y^-}, Q_{x^+,y^+}, Q_{x^-,y^+}, Q_{x^+,y^-} with each $Q_{x,y}$ having length 5 and going from x to y. Additionally the vertices and colours of the path pairs $(Q_{x^-,y^-}, Q_{x^+,y^+})$ and $(Q_{x^-,y^+}, Q_{x^+,y^-})$ both partition the 12 vertices and 10 colours of C.

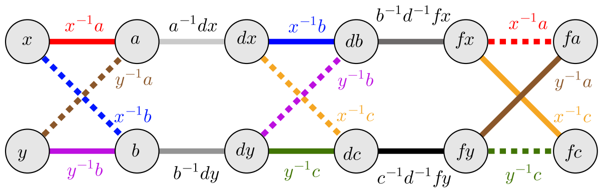

Fig. 11 The analogue of Fig. 12 when proving Lemma 6.17 for K_G^-. Aside from replacing S with the set of elements given in this figure, the proof for K^- is identical to the one given for K^+ (Color figure online)

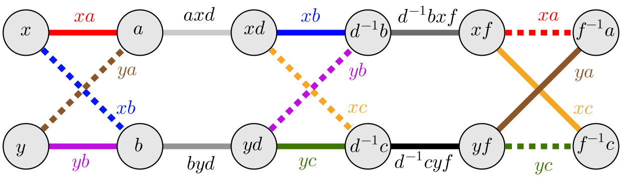

Fig. 12 The coloured graph produced by Lemma 6.17 for K_G^+. Each edge is directed towards the right. Notice that all vertices/edges are labelled by elements of S. Since all $w, w' \in S$ are separable, this means that the π-image of this graph has all vertices/colours distinct (and so in particular has 12 vertices and 10 colours as required). To see that it satisfies the lemma, we need to exhibit paths Q_{x^-,y^-}, Q_{x^+,y^+}, Q_{x^-,y^+}, Q_{x^+,y^-} between $x^- = x$, $x^+ = y$, $y^- = f^{-1}a$, $y^+ = f^{-1}c$. The solid lines in the picture give the two paths Q_{x^-,y^+}, Q_{x^+,y^-}. Replacing the coloured solid lines for the coloured dashed lines (and keeping all grey lines) gives the two paths Q_{x^-,y^-}, Q_{x^+,y^+} (Color figure online)

Proof We prove the lemma when $K_G^{\pm} = K_G^+$. A slight change in variables proves the lemma also when $K_G^{\pm} = K_G^-$ (see Fig. 11). With high probability, Corollary 3.33 applies with $R = R^1 \cap R^2$. Thinking of x, y, a, b, c, d, f, as free variables consider the set

$$S = \{x, y, a, b, xd, yd, d^{-1}b, d^{-1}c, xf, yf, f^{-1}a, f^{-1}c, xa, yb, axd, byd, xb,$$
$$ya, xc, yc, d^{-1}bxf, d^{-1}cyf\}$$

(see Fig. 2). Note that all pairs w, w' are linear and separable (by (a), since they're all linear in different combinations of free variables). Lemma 3.32 gives a projection π which separates S and has $\pi(S) \subseteq R$. This means that $\pi(w)$ are distinct for all $w \in S$. Now the graph given in Fig. 12 satisfies the lemma with $x^- = \pi(x)$, $x^+ = \pi(y)$, $y^- = \pi(f^{-1}a)$, $y^+ = \pi(f^{-1}c)$, with paths as shown. □

We now show how to simulate the task of a wire in a sorting network via short paths.

Lemma 6.18 *Let $p \geq n^{-1/700}$. Let H_G be a multiplication hypergraph, R^1, R^2 p-random subsets of G, sampled independently. With high probability, for any $x, y \in V(K_G^{\pm})$, $U \subseteq G$ with $|U| \leq p^{800}n/10^{4010}$, there is a length 3 x to y path $xuvy$ in $K_G^{\pm}$ with $u, v \in R_1 \setminus U$, $c(xv), c(vy) \in R_2 \setminus U$.*

Proof We prove the lemma when $K_G^{\pm} = K_G^+$. To prove it for $K_G^{\pm} = K_G^-$ replace the definition of S below by $S = \{u, v, u^{-1}v, x^{-1}u, v^{-1}y\}$.

With high probability, Corollary 3.33 applies with $R = R^1 \cap R^2$. Consider the set $S = \{u, v, uv, xu, vy\} \subseteq G * F_3$. Note that all pairs w, w' are linear and the pairs (u, v), (xu, uy), (xu, uv), (uv, vy) are separable (since they're all linear in different combinations of free variables). Lemma 3.32 gives a projection π which separates S. Now the path $P = x\pi(u)\pi(v)y$ satisfies the lemma — the vertices are distinct since u, v are separable, while the colours are distinct since uv, xu, uy are pairwise separable. □

We now prove the analogue of Lemma 4.3 from [44] adapted to a setting where the host structure is $K_G^{\pm}$.

Lemma 6.19 *Let $p \geq n^{-1/800}$. Let $t = 8\lceil 100 \log^2 n \rceil$. Let R_V and R_C be p-random subsets of G, sampled independently. Then, the following holds with high probability.*

Let $A, B \subseteq G$ be disjoint subsets with $|A| = |B| \leq \frac{p^{1000}n}{10^{5000}\log(n)^3}$, and let $U \subseteq G$ with $|U| \leq \frac{p^{1000}n}{10^{5000}}$. Then, there exists $V \subseteq R_V \setminus U$ and $C \subseteq R_C \setminus U$ such that for any bijection $\phi\colon A \to B$, there exists a system of paths using exactly the vertices/colours of $K_G^{\pm}[A \cup B \cup V; C]$ from A to B, each of length t, and connecting a to $\phi(a)$ for each $a \in A$.

Proof With high probability, Lemmas 6.17 and 6.18 apply. Let N be a sorting network given by Lemma 6.16, with $m := |A| = |B|$, noting this sorting network has $\leq 200m \log^2 m$ comparators. For each comparator $C_i = \{x_i^-, x_i^+, y_i^-, y_i^+\}$, use Lemma 6.17 to find a subgraph C_i' in $R \setminus (A \cup B)$. Identify the nodes x_i^-, x_i^+, y_i^-, y_i^+ of the comparator with the vertices x_i^-, x_i^+, y_i^-, y_i^+ of C_i'. Let $A = \{x_1, \dots, x_m\}$, $B = \{y_1, \dots, y_m\}$. For each wire xy of the sorting network use Lemma 6.18 to find a rainbow length 3 path P_{xy} joining corresponding vertices of $A \cup B \cup \bigcup C_i$. By enlarging the set U at all these applications, we can ensure that the subgraphs C_i' are all disjoint, and that the paths P_{xy} are all internally disjoint and colour-disjoint from the subgraphs C_i' and from each other. We claim that $V = A \cup B \cup \bigcup V(C_i) \cup \bigcup V(P_w)$ and $C = \bigcup C(C_i) \cup \bigcup C(P_w)$ satisfy the lemma.

Consider a bijection $\phi\colon A \to B$. This gives a permutation σ of $[m]$ so that $\phi(x_i) = y_{\sigma(i)}$. Assign value $v(a_i) := \sigma(i)$. This gives a value to each node and wire of the sorting network as in Definition 6.15. We now translate this into values for the corresponding paths/vertices in $K_G^{\pm}$. The values we assign come from the set $\{1, \dots, m\} \cup \{0\}$. For each wire xy of the sorting network, define $v(P_{xy}) = v(x) = v(y)$ and give all vertices/edges of P_{xy} this value. For each comparator $C = \{y_i^-, y_i^+, x_i^-, x_i^+\}$ we have either $v(y_i^-) = v(x_i^-)$, $v(y_i^+) = v(x_i^+)$ or $v(y_i^-) = v(x_i^+)$, $v(y_i^+) = v(x_i^-)$. In the former case, define $v(Q_{y_i^-, x_i^-}) = v(y_i^-) = v(x_i^-)$, $v(Q_{y_i^+, x_i^+}) = v(y_i^+) = v(x_i^+)$, and give the vertices and edges of the corresponding paths the same value. Give the paths $Q_{y_i^-, x_i^+}$ and $v(Q_{y_i^+, x_i^-})$ as well as the unused edges of C_i' (those edges of C_i' not on $Q_{y_i^-, x_i^-}$ or $Q_{y_i^+, x_i^+}$) value 0. In the latter case, define $v(Q_{y_i^-, x_i^+}) = v(y_i^-) = v(x_i^+)$, $v(Q_{y_i^+, x_i^-}) = v(y_i^+) = v(x_i^-)$, and give the vertices and edges of the corresponding paths the same value. As before, give the other two paths and the unused edges value 0. Note that this way every vertex/edge of U gets a value, and these values match those that corresponding nodes/wires have in N. For every $i = 1, \dots, m$, let

$$P_i = \bigcup_{v(P_{xy})=i} P_{xy} \cup \bigcup_{v(Q_{y_i^{\diamond_1}, x_i^{\diamond_2}})=i} Q_{y_i^{\diamond_1}, x_i^{\diamond_2}}.$$

We clarify that above xy is quantified over the set of wires, and $\diamond_j$ is quantified over $\{+, -\}$. We claim that $P_1, \dots, P_m$ are each paths and have all the required properties.

First note that every vertex in $V \setminus (A \cup B)$ has exactly one in-going edge of non-zero value and exactly one outgoing edge of non-zero value. The vertices a_i have no in-going edges and one out-going edge (whose value is $\sigma(i)$). The vertices b_i have no out-going edges and one in-going edge (whose value is i). Combined with the whole graph being acyclic (which holds due to the sorting network being acyclic), this shows that $\bigcup P_i$ is a union of paths.

Since $v(a_i) = \sigma(i)$ and $v(b_i) = i$, path $P_{\sigma(i)}$ goes from a_i to $b_{\sigma(i)}$. In particular, each P_i is a path. Also, this shows that the paths partition the vertices V and have the correct endpoints. We now show that their union is rainbow using exactly the colours C. The fact that every colour of $\bigcup C(P_{xy})$ is used exactly once comes from the fact that every wire has a value, and so every edge of each P_{xy} as a value. So such colours are used at least once (and hence exactly once because these colours occur once in the whole graph). The colours on the comparators C_i' are used once as a consequence of Lemma 6.17. To see this, note each such colour comes up exactly twice — once in the paths $Q_{x_i^-, y_i^-} \cup Q_{x_i^+, y_i^+}$ and once in the paths $Q_{x_i^-, y_i^+} \cup Q_{x_i^+, y_i^-}$. By the assignment of the values to the comparator one of these always has value 0.

Finally, to see that each P_i has length exactly t, observe that by Lemma 6.16 we know each P_i has length $\lceil 100 \log^2 n \rceil$ when viewed as a path in the sorting network. As each wire gadget corresponds to a path of length 3 and each comparator corresponds to a path of length 5, it follows that each P_i has length $8\lceil 100 \log^2 n \rceil$, as required. □

6.2.2 Deducing Lemma 6.22

We will need the following technical lemma, which allows us to use the nibble method to saturate large sets of non-random vertices. This is necessary, as after applying Lemma 6.19 to a random subset, we will be left with a large set of non-random vertices which do not immediately fit into the setting of our main theorem.

Lemma 6.20 *There exists $C = C_{6.20} \geq 10$ 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 $K_G^{\pm}[A', B'; C']$.

To prove the above lemma, we will make use of the following result of Montgomery, Pokrovskiy, and Sudakov. We use the notation $x \overset{\text{POLY}}{\ll} y$ to mean that $x, y \in (0, 1]$ and there is some absolute constant $C \geq 1$ such that the proof works with $x \leq y^C/C$. Recall that an edge-coloured graph is **globally** K**-bounded** if each colour occurs at most K times in the colouring. When we say an edge-coloured bipartite graph is typical, we refer to the typicality of the underlying uncoloured bipartite graph. We remark that in [47], the following lemma is stated with the additional hypothesis that $n^{-1} \overset{\text{POLY}}{\ll} \gamma$, but this is easily seen to be redundant as γ being smaller only makes the statement easier to prove (as a (γ, δ, n)-typical graph is also $(2\gamma, \delta, n)$-typical).

Lemma 6.21 ([47], Corollary 8.12) *Let n, δ, p, γ be such that $n^{-1}, \gamma \overset{\text{POLY}}{\ll} p, \delta \leq 1$. Every properly coloured, (γ, δ, n)-typical, globally $(1-p)\delta n$-bounded, balanced bipartite graph G of order $2n$ has $(1-p)\delta n$ edge-disjoint rainbow perfect matchings.*

Proof of Lemma 6.20 Recall $m := \max\{an, bn, cn\}$ and note that $m \leq \ell + m^{1-\gamma}$. Set $p := m/n$. Note that combining the given inequalities we obtain that $a + b + c \geq 12(\ell/n)/5 \geq 11p/5$. We define the following random sets.

- Let $A_1, A_2 \subseteq A$ be disjoint $p - b$ and $(p - c + \zeta)$-random subsets of A. Let $A_3 := A \setminus A_1 \setminus A_2$ noting that this set is $(a - 2p + b + c - \zeta) =: \alpha$-random.
- Let $B_1, B_2 \subseteq B$ be $(p - a)$ and $(p - c + \zeta)$-random subsets of B. Let $B_3 := B \setminus B_1 \setminus B_2$ noting that this set is $(b - 2p + a + c - \zeta) = \alpha$-random.
- Let $C_1, C_2 \subseteq C$ be $(p - a)$ and $(p - b)$-random subsets. Let $C_3 := C \setminus C_1 \setminus C_2$ noting that this set is $(c - 2p + a + b) =: \beta$-random.

Note there is space to sample these sets disjointly due to the assumptions on the size of ℓ and a, b, c. In particular, using that $a + b + c \geq 11p/5$ and $\zeta \leq p/100$, we have that $\alpha, \beta \geq p/5 - \zeta \geq p/10 = 1/10 \log^C n$ and $p - a, p - c + \zeta, p - b \geq 0$ so the random sets with these parameters are well-defined. For each pair of random sets (A_2, B_2), (A_1, C_2), and (B_1, C_1) such that the corresponding randomness parameters are both at least $n^{-1/600}$ (the randomness parameter for each pair is in fact equal), we have that with probability at least $1 - 100/n^3$, Lemma 3.8 holds (the linearity and typicality of the corresponding 3-uniform hypergraph of $K_G^{\pm}$ follows by a straightforward modification of Observation 3.3, which we omit). By Chernoff's bound, the sizes of all these random sets are at most $n^{0.6}$ away from their expectations with probability at least $1 - 1/n^5$.

Set $\nu := n^{-0.25}$. For any $v \in G$, the expected number of (out-)neighbours of v in $K_G^{\pm}[\{v\} \cup B_3; C_3]$ is $\alpha\beta(n-1)$, as C is sampled independently with B and the analogous statement holds for expected number of neighbours of v in $K_G^{\pm}[A_3 \cup \{v\}; C_3]$. By Chernoff's bound and a union bound, both of these random variables are at most νn away from their expectation with probability at least $1 - 1/n^3$. Let $d_A(a, a')$ denote the pair degree of a and a' in $K_G^{\pm}[\{a, a'\} \cup B_3; C_3]$ (viewed as an uncoloured bipartite graph), and let $d_B(b, b')$ denote the pair degree of b and b' in $K_G^{\pm}[A_3 \cup \{b, b'\}; C_3]$. For each a, a' and b, b', we have that $\mathbb{E}(d_A(a, a')) = \mathbb{E}(d_B(b, b')) =$

$\alpha\beta^2 n$. Further, observe that $d_A(a,a')$ and $d_B(b,b')$ are 2-Lipschitz random variables. Hence, by Azuma's inequality and a union bound, with probability at least $1-1/n^3$, for each a, a' and b, b', $d_B(b,b')=d_A(a,a')=(\alpha\beta^2 \pm \nu)n$. This establishes in particular that $K_G^{\pm}[A_3,B_3;C_3]$ is $(\nu,\beta,\alpha n)$-typical as an uncoloured bipartite graph. Let $d(c)$ denote the number of times c occurs in $K_G^{\pm}[A_3,B_3,\{c\}]$. Then $\mathbb{E}(d(c)) \le \alpha^2 n$ for each c. By Chernoff's bound and a union bound, $d(c) \le (1\pm\nu)\alpha^2 n$ for each $c \in G$. In particular, this implies that $K_G^{\pm}[A_3,B_3,C_3]$ is globally $(\alpha^2+\nu)n$-bounded, which implies that it is globally $(1-\zeta/10)\beta(\alpha n)$-bounded (using bounds on α, β obtained earlier).

With probability $\ge 1-1/n^{2.5}$, all of the previous properties hold. Let A', B', C' be given with the indicated properties. By properties coming from Lemma 3.8, $(A_2,B_2,C'\setminus C)$, $(A_1,B'\setminus B,C_2)$ and $(A'\setminus A,B_1,C_1)$ each contain matchings covering all but $n^{1-\gamma/2}$ vertices. Here, if the corresponding random sets have parameters smaller than $n^{-1/600}$ (so we cannot apply Lemma 3.8), we simply take an empty matching. This is sufficient as in this case by Chernoff's bound, the random sets themselves cannot contain more than $n^{1-1/603}$ elements. Accounting for differences $A\setminus A'$, $B\setminus B'$, $C\setminus C'$ (each of size $\le n^{1-\gamma}$), we have a matching covering all but $10n^{1-\gamma/2}$ vertices of $(A_2\cap A', B_2\cap B', C'\setminus C)$, $(A_1\cap A', B'\setminus B, C_2\cap C')$ and $(A'\setminus A, B_1\cap B', C_1\cap C')$.

The set of leftover vertices of A', B' and C' have a symmetric difference with A_3, B_3 and C_3 (respectively) of size at most $100n^{1-\gamma/2}$. Set $\gamma' := n^{-\gamma/100}$. By the pair-degree and vertex-degree bounds we obtained earlier, this implies that the associated properly coloured bipartite graph with the leftover vertices is $(\gamma',\beta,\alpha n)$-typical and globally $(1-\zeta/100)\beta(\alpha n)$-bounded. Then, by Lemma 6.21 we can find a matching saturating the remaining vertices of A' and B' using the leftover colours from C', as desired. To see that the hypothesis of Lemma 6.21 are satisfied, note first that $(\alpha n)^{-1} \overset{\text{POLY}}{\ll} \beta$ holds as $\beta \ge p/10 \ge 1/(10\log^C n)$ and n is sufficiently large. Also, $(\alpha n)^{-1} \overset{\text{POLY}}{\ll} \zeta/100$ follows from $1/m \le \zeta^C/C$ (supposing C is sufficiently large) and $\alpha n \ge 10m$ (as $\alpha \ge p/10$). □

We can now give the proof of the key lemma of this section.

Lemma 6.22 *Let $1/n \ll \gamma, p \le 1$, let t be a positive integer between $\log^7(n)$ and $\log^8(n)$, and set $q := p/(t-1)$. Let G be a 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 $|R\Delta R'| \le n^{1-\gamma}$.*
2. *Either **O.** $\sum V'_{end} - \sum V'_{str} = \sum C'$ or **C.** $\sum V'_{end} + \sum V'_{str} + 2\sum V'_{mid} = \sum C'$ holds in the abelianization of G.*
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 $f\colon V'_{str} \to V'_{end}$, we have that $K^*_G[V'_{str} \cup V'_{end} \cup V'_{mid}; C']$ has a rainbow $\vec{P}_t$-factor where each path starts on some $v \in V'_{str}$ and ends on $f(v) \in V'_{end}$ where $* = +$ if $\boldsymbol{C}$ holds and $* = -$ if $\boldsymbol{O}$ holds.

Proof Set $m = 8\lceil 100\log^2 n\rceil$, $r = 10^{-100}$. Partition V_{mid} into random subsets $R_V, V_1, \ldots, V_{t-m}$ where R_V is rp-random, V_1, V_2, V_3 and V_{t-m} are q-random and the rest are $((1-r)p-4q)/(t-m-4)$-random. Note that $9q/10 \le ((1-r)p-4q)/(t-m-4) \le (1-r^{100})q$. Independently with the previous sets, partition C into $R_C, C_1, \ldots, C_{t-m}$ where R_C is rp-random, C_1, C_2 and C_{t-m} are q-random, and the rest are $((1-r)p-3q)/(t-m-3)$-random. Note similarly that $9q/10 \le ((1-r)p-3q)/(t-m-3) \le (1-r^{100})q$.

With high probability, Theorem 4.9 applies with the q-random sets (V_2, V_3, C_2), Lemma 6.19 applies with R_V, R_C and Lemma 6.20 applies with all potential values of ℓ and all rational values of ζ with denominator at most n for all triples of random sets

$$\begin{aligned}&(V_{str}, V_1, C_1), (V_3, V_4, C_3), (V_4, V_5, C_4), \ldots,\\ &(V_{t-m-1}, V_{t-m}, C_{t-m-1}), (V_{t-m}, V_{end}, C_{t-m}).\end{aligned} \tag{3}$$

We remark that here we use a union bound over all of the $\le n$ potential values of ℓ and p as well as all of the $\le \log^5 n$ listed triples (note $a, b, c \ge 1/\log^9(n)$ in each of these applications). This is possible as the failure probability of Lemma 6.20 is at most $1/n^{2.5}$. Finally, with high probability, all the random sets have sizes within $n^{0.6}$ of their expectations.

Fix all the random sets and the integer ℓ so they have all of the collected properties. Fix also the sets V'_{str}, V'_{mid}, V'_{end}, C' so that they satisfy properties (1-4). Note that as the random sets have size close to their expectations, and property (1) holds for $V'_{str}/V'_{mid}/V'_{end}$, this implies in particular that $\ell = qn \pm n^{1-\gamma/2}$. Recall that the maximum and minimum values of the randomness parameters of the random sets listed in (3) are q and $9q/10$ respectively. These imply that ℓ satisfies the requirements for Lemma 6.20 (with $\gamma/2$ playing the role of γ) to be applicable for each of the triples in (3)). We define ζ to be a rational number with denominator at most n as close as possible to $1/\log^{10^{10}} n$ as possible while ensuring that $\ell/(1-\zeta)$ is an integer, noting ζ also satisfies the constraints of Lemma 6.20 (for any triple in (3)) if n is sufficiently large.

We can find some $R'_V \subseteq R_V$ and $R'_C \subseteq R_C$ such that these sets have the property from Lemma 6.19 with respect to V'_1 and V'_2. That is, for any bijection $\phi\colon V'_1 \to V'_2$, there exists a system of rainbow paths using exactly the vertices/colours of $K^{\pm}_G[V'_1 \cup V'_2 \cup R'_V; C']$ from V'_1 to V'_2, each of length $m = 8\lceil 100\log^2 n\rceil$, and connecting v to $\phi(v)$ for each $v \in V'_1$. Note that the relevant inequalities hold while applying Lemma 6.19 with $(A, B) = (V'_1, V'_2)$ and R_V and R_C, since $|V'_1| = |V'_2| = \ell \le n/\log^7 n \ll (rp)^{1000} n/\log^4 n$ since $1/n \ll rp$.

Distributing the leftover vertices in $R_V \setminus R'_V$ into V_i (for $i \notin \{1, 2, 3, t-m\}$), we can find disjoint subsets $V'_1, V'_2, \ldots, V'_{t-m} \subseteq G$ with the following properties.

1. $R'_V, V'_1, V'_2, \ldots, V'_{t-m}$ partition V'_{mid}

2. For each $i \in [t-m]$, $|V_i'| = \ell$ and $|V_i \setminus V_i'| \leq n^{1-\gamma/2}$
3. For each $i \in \{1, 2, 3, t-m\}$, $|V_i \Delta V_i'| \leq n^{1-\gamma/4}$

We can perform a similar distribution, adding to each C_i with $i \neq 2$ ζ-fraction more colours than necessary (to be able to apply Lemma 6.20) borrowed from C_2, we can find disjoint subsets $C_1', C_2', \dots, C_{t-m}' \subseteq G$ partitioning $C' \setminus R_C'$ with the following property: $|C_i'| = \ell/(1-\zeta)$ for each $i \neq 2$, $|C_i \setminus C_i'| \leq n^{1-\gamma/4}$ for each $i \neq 2$ and $|C_2 \Delta C_2'| \leq \log^9 n \zeta n$.

Invoke Lemma 6.20 with sets

$$(V_{str}', V_1', C_1'), (V_3', V_4', C_3'), (V_4', V_5', C_4'), \dots,$$
$$(V_{t-m-1}', V_{t-m}', C_{t-m-1}'), (V_{t-m}', V_{end}', C_{t-m}')$$

to find perfect matchings (saturating the vertex sets, and using all but ζ-fraction of the colours from each of the colour sets). For each $i \neq 2$, denote the colour subset of C_i' used in the corresponding perfect matching by C_i''. Let C_2'' be the union of C_2' and all of the unused colours from each C_i', that is, $C_i' \setminus C_i''$ $(i \neq 2)$, noting that there are $pn(t-m-1) \ll \log^{10} n(\zeta n)$ such unused colours, meaning that C_2'' is defined to retake the borrowed colours from earlier.

Claim 6.22.1 *$|C_0''| = \ell$ and if* $\mathbf{O}$ *holds, we have $\sum V_3' - \sum V_2' = \sum C_2''$, and if* $\mathbf{C}$ *holds, we have $\sum V_3' + \sum V_2' = \sum C_2''$.*

Proof To see this, it is convenient to apply the property of R_V' and R_C' coming from Lemma 6.19 with an arbitrary choice of ϕ. This gives a packing of rainbow paths in $K_G^{\pm}$ with the set of endpoints of the paths being V_1', V_2' and R_C' being the set of colours used on the paths. This, together with assumption (4), and that the colours in $C' \setminus R_C' \setminus C_2''$ have been used to find perfect matchings in the specified sets implies the first part of the claim. In the case that we work with the division digraph, considering all the directed paths we have found so far, this implies that $\sum V_2' - \sum V_{str} = \sum C_1'' + \sum R_C'$ and $\sum V_{end} - \sum V_3' = \sum C_3'' \cup C_4'' \cup \dots \cup C_{t-m}''$. Combining these equalities with $\mathbf{O}$, we obtain the second part of the claim. In the case that we work with the multiplication digraph instead, a similar argument shows the second part of the claim. □

The previous claim allows us to apply Theorem 4.9 with (V_2', V_3', C_2'') to find a perfect matching directed from V_2' towards V_3' both in the case of division digraphs and multiplication digraphs (note C_2'' cannot contain e if G is Boolean by assumption). Now, we invoke the property of the sets R_V' and R_C' with a choice of ϕ so that we produce a $\vec{P}_t$-factor connecting each $v \in V_{str}$ to $f(v) \in V_{end}$. To define such a ϕ, for each $v \in V_1$, let v' be the matched neighbour of v in V_{str}. Consider the vertex v'' which is obtained by starting with $f(v') \in V_{end}$, and following each matched edge until we reach a vertex of V_2. Set $\phi(v) = v''$. It is easy to see that this function ϕ has the desired behaviour, concluding the proof. □

6.2.3 Deducing Theorem 6.9

Proof Set $t = \lfloor 2\log^7(n) \rfloor$. Let w be the remainder when n is divided by $t+1$, and set ℓ so that $n = (t+1)\ell + w$. Note that $w \le t+1 \le 2\log^7 n$. Set $q := (\ell - 1)/n$ and set $p := (t-1)q$, noting $2q + p = 1 \pm \frac{3t}{n}$. Take random disjoint sets V_{str}, V_{end}, and V_{mid} where the former two are q-random and the latter is p-random (noting that this is essentially a partition of G). Independently, take disjoint random sets C_1 and C_0 where C_1 is $(q+p)$-random and C_0 is q-random (hence C_0 and C_1 almost partition G). With high probability, Lemma 6.22 applies with V_{str}, V_{end}, V_{mid} and C_1 (set to be C in the statement of Lemma 6.22) and t. With high probability, Lemma 6.20 applies with (V_{str}, V_{end}, C_0) and $\ell - 1$ (in this application, $a = b = c = q$ and $\ell := \ell - 1$ so the relevant inequalities hold for Lemma 6.20 to apply). With high probability, all the random sets have size close to their expectations. Fix sets V, C and vertices x, y as in the statement of the theorem.

As $w = o(n)$, we can greedily find a rainbow directed path using vertices from $V \setminus \{y\}$ and colours from C, say P_0, on $w+1$ vertices, starting on x and terminating on some x' where x'. Now, we may partition $(V \setminus V(P_0)) \cup \{x'\}$ as V'_{str}, V'_{end} and V'_{mid} so that $x' \in V'_{str}$ and $y \in V'_{end}$, $\ell = |V'_{str}| = |V'_{end}| = |V'_{mid}|/(t-1)$ (possible due to the divisibility condition coming from the size of P_0) and for each random set R, R and R' have small symmetric difference. For this, it is important that $2q + p = 1 \pm \frac{3t}{n}$. Set $k := \lfloor n^{1-10^{-5}} \rfloor$. We may partition $C \setminus C(P_0)$ as C' and C'_0 so that $C'_0 = \ell + k - 1$, and $C' := C \setminus C'_0$ where the pairs of sets (C'_0, C_0) and (C_1, C') have small symmetric difference.

Now, invoke Lemma 6.20 with the sets $(V'_{end} - y, V'_{str} - x', C'_0)$ to find a perfect directed matching from $V'_{end} - y$ to $V'_{str} - x'$ using $\ell - 1$ colours coming from C'_0, calling this set C''_0. Re-define (without relabelling) C' to include the k unused colours from C'_0, noting that C' still has small symmetric difference with C_1. Note that $C' := C \setminus C(P_0) \setminus C''_0$.

Claim 6.22.2 *Hypothesis* 2. *of Lemma* 6.22 *holds with the sets* V'_{str}, V'_{end}, V'_{mid}, C'.

Proof Suppose first that $K^{\pm}_G = K^-_G$, that is, we work with the division digraph. Then, it must be that $\sum(V'_{str} - x') - \sum(V'_{end} - y) = \sum C''_0$ from the perfect matching we found. We also have $\sum C''_0 = \sum C - \sum C' - \sum C(P_0)$, and that $\sum C(P_0) = x' - x$ (again using that we work with the division digraph). By assumption of the theorem, we also have $\sum C = y - x$. Adding up all of these equalities, we obtain $\sum V'_{end} - \sum V'_{str} = \sum C'$ as desired. An analogous argument works for multiplication digraphs and we omit the details. □

The other hypotheses are easy to verify and thus we can invoke Lemma 6.22 with an appropriate choice of a bijection $f\colon V'_{str} \to V'_{end}$ so that the union of the path system we obtain combined with the initial perfect matching creates a path directed from x' to y. Combined with P_0, we obtain the desired path from x to y, concluding the proof. □

6.3 Zero sum partitions

In this section we apply our main theorem to solve a conjecture of Cichacz and a problem of Tannenbaum about partitions of abelian groups into zero-sum sets.

6.3.1 The Cichacz conjecture

The following lemma makes our main theorem easier to apply by slightly modifying a set into zero-sum sets of specified size.

Lemma 6.23 *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 $|R'| = m$, $|R' \triangle R| \leq 6\epsilon n$, and $\sum R' = g$.

Proof By Chernoff's bound, with high probability $||R| - pn| \leq \sqrt{n}\log n \leq \epsilon n$. Also with high probability, Corollary 3.33 applies to R. Let $\epsilon \in [2\log n/\sqrt{n}, p^{800}/10^{4010}]$, m with $|m-pn| \leq \epsilon n$, $g \in G$ and Z with $|Z| \geq m+3$, $|R \setminus Z| \leq \epsilon n$. Let X be a subset of Z of order $m+3$ which either contains or is contained in $R \cap Z$ (depending on whether $|R \cap Z| \leq m+3$ or not). Note that in both cases $|X \setminus R| \leq m+3-|R \cap Z|$ and $|R \setminus X| \leq |R \setminus Z|$. These give $|X \triangle R| = |R \setminus X| + |X \setminus R| \leq |R \setminus Z| + (m+3-|R \cap Z|) = 2|R \setminus Z| + 3 + m - |R| \leq 2|R \setminus Z| + 3 + |m - |R|| \leq 2|R \setminus Z| + 3 + ||R| - pn| + |m - pn| \leq 5\epsilon n$. Set $U = R \setminus X$ to get a set of size $\leq 5\epsilon n \leq p^{800}n/10^{4005}$. Let h be an element such that $h = \sum X - g$ in G^{ab}. Thinking of x and y as free variables in $G * F_2$, consider the set $S = \{hxy, x^{-1}, y^{-1}\} \subseteq G * F_2$ and note that all pairs of words in S are linear and separable (by part (a) of the definition of separable). Thus there is a projection $\pi : G * F_2 \to G$ which has $\pi(S) \subseteq R \setminus U = R \cap X$ and which separates S (i.e. has $\pi(hxy)$, $\pi(x^{-1})$, $\pi(y^{-1})$ distinct). Thus setting $R' = X \setminus \{\pi(hxy), \pi(x^{-1}), \pi(y^{-1})\}$ gives a set of size $|X| - 3 = m$ with $\sum R' = \sum X - h = g$. □

The first result about zero-sum partitions that we prove gives a zero-sum partition of a subset of a group into sets of some small fixed size.

Lemma 6.24 *Let $p \geq n^{-1/10^{100}}$ and $k = 3, 4,$ or 5. Let R be a p-random subset of an abelian group G. With high probability the following holds.*

Let $X \subseteq G$ with $|X \triangle R| \leq p^{10^{10}}n/\log(n)^{10^{18}}$, $0 \notin X$, $\sum X = 0$, and $|X| \equiv 0 \pmod{k}$. Then, X can be partitioned into zero-sum sets of size k.

Proof Partition R into (p/k)-random sets $R_1, \dots, R_k$ by placing each vertex of R in each set independently with probability $1/k$. Let S_1, S_2 be (p/k)-random, and chosen independently of these and each other. Note that for $i = 1, 2$, the sets $S_i^{-1} = \{-s : s \in S_i\}$ are also (p/k)-random and independent of $R_1, \dots, R_k$. By Chernoff's bound, with high probability, $|R_i| = pn/k \pm \sqrt{n}\log n$. With high probability Lemma 6.23 applies to all these sets. Also with high probability, the conclusion of Theorems 4.5, 4.4, or 4.6 applies to any choice of 3 sets out of $R_1, \dots, R_k, S_1, S_2, S_1^{-1}, S_2^{-1}$ aside from choices which use both S_i and S_i^{-1} for $i = 1, 2$.

Let $X \subseteq G$ with $|X \triangle R| \leq p^{10^{10}} n/\log(n)^{10^{18}}$, $\sum X = 0$, and $|X| \equiv 0 \pmod{k}$. Partition X into sets $X_1, \dots, X_k$ satisfying $|X_i| = |X|/k$, $|X_i \triangle R_i| \leq p^{10^{10}} n/\log(n)^{10^{16}}$, $\sum X_i = 0$ for $i = 1, \dots, k$ (to do this, use Lemma 6.23 $k - 1$ times, applying it to the pairs $(R, Z) = (R_1, X), (R_2, X \setminus X_1), \dots, (R_{k-1}, X \setminus \bigcup_{i=1}^{k-2} X_i)$ with $p' = p/k$, $m = |X|/k$, $\epsilon = p^{10^{10}} n/\log(n)^{10^{17}}$, $g = 0$. Afterwards setting $X_k = X \setminus \bigcup_{i=1}^{k-1} X_i$ gives a set with $\sum X_k = 0$, $|X_k| = |X|/k$, and $|X_i \triangle R_i| \leq p^{10^{16}} n/\log(n)^{10^{10}}$). Use Lemma 6.23 to choose Y_1, Y_2 with $|Y_1 \triangle S_1|, |Y_2 \triangle S_1| \leq p^{10^{10}} n/\log(n)^{10^{16}}$, $|Y_1| = |Y_2| = |X|/k$ and $\sum Y_1, \sum Y_2 = 0$ (for this application, use $Z = G$). We proceed differently based on the value of k:

- If $k = 3$, then we have $\sum X_1 + \sum X_2 + \sum X_3 = 0$ and so Theorem 4.6 applies to these sets giving a perfect matching (whose edges give a partition of X into zero-sum sets of size 3). Note that here we used that $0 \notin X$ in the case that G is an elementary 2-group.
- If $k = 4$, then $\sum X_1 + \sum X_2 + \sum Y_1 = 0$ and $\sum X_3 + \sum X_3 + \sum Y_1^{-1} = 0$. Thus Theorem 4.6 gives perfect matchings M_1, M_2 in $H_G[X_1, X_1, Y_1]$, $H_G[X_3, X_4, Y_1^{-1}]$ respectively. For each $y \in Y_1$, let (a_y, b_y, y) and $(c_y, d_y, -y)$ be the edges of M_1/M_2 through $y/-y$ respectively. Notice that for each y, we have $a_y + b_y + c_y + d_y = (a_y + b_y + y) + (c_y + d_y - y) = 0$ and so $\{(a_y, b_y, c_y, d_y) : y \in Y_1\}$ gives a partition of X into zero-sum sets of size 4.
- If $k = 5$, then $\sum X_1 + \sum X_2 + \sum Y_1 = 0$, $\sum X_4 + \sum X_5 + \sum Y_2 = 0$, and $\sum X_3 + \sum Y_1^{-1} + \sum Y_2^{-1} = 0$. Thus Theorems 4.4, 4.6 give perfect matchings M_1, M_2, M_3 in $H_G[X_1, X_1, Y_1]$, $H_G[X_4, X_5, Y_2]$, $H_G[X_3, Y_1^{-1}, Y_2^{-1}]$ respectively. For each $x \in X_3$, let $(x, y_x, z_x) \in M_3$ be the edge of M_3 through x and let $(a_x, b_x, -y_x)$, $(c_x, d_x, -z_x)$ be the edges of M_1/M_2 through $-y_x/-z_x$ respectively. Notice that for each x, we have $a_x + b_x + x + c_x + d_x = (a_x + b_x - y_x) + (x + y_x + z_x) + (c_x + d_x - z_x) = 0$ and so $\{(a_y, b_y, x, c_y, d_y) : x \in X_3\}$ gives a partition of X into zero-sum sets of size 5. □

We now prove a version of the previous result where the sizes of the zero-sum sets are not fixed. We remark that the $p = 1$, $X = G \setminus e$ case of the following result proves Conjecture 1.6 for sufficiently large groups. We'll work with multisets in this section and will use $m_i(M)$ to denote the number of occurrences of element i in a multiset M.

Lemma 6.25 *Let $p \geq 1/\log(n)^{10^{99}}$ and $M \subseteq \{3, 4, 5, \dots\}$ be a multiset. Let R be a p-random subset of an abelian group G. With high probability the following holds.*

Let $X \subseteq G$ with $|X \triangle R| \leq p^{10^{10}} n/\log(n)^{10^{24}}$, $0 \notin X$, $\sum X = 0$, and $|X| = \sum M$. Then, X has a zero-sum M-partition.

Proof Without loss of generality, we can suppose that $M \subseteq \{3, 4, 5\}$. Otherwise consider a set M' formed by replacing each $y \in M$ with $y > 5$ by $y_1, \dots, y_t$ with

$y_1 + \cdots + y_t = y$ and $y_i \in \{3, 4, 5\}$. It is easy to see that M' still satisfies the hypotheses of the lemma and a M'-partition of G gives a M-partition by combining $y_1, \ldots, y_t$-sets into a single y-set for each y.

For $i = 3, 4, 5$, fix $m_i = m_i(M)$. Partition $R = R_3 \cup R_4 \cup R_5$ into disjoint sets with R_i being q_i-random for $q_i := \frac{im_i}{3m_3+4m_4+5m_5}$. For each $i \in \{3, 4, 5\}$, either $q_i \leq n^{-1/10^{100}}$, or with high probability, R_i satisfies the conclusions of Lemmas 6.23 and 6.24. We'll suppose that $q_3 \geq q_4 \geq q_5$ — the proof is identical if these are ordered differently (switching the roles of $R_3/R_4/R_5$ correspondingly). In particular, we will suppose that R_3 satisfies the conclusions of Lemmas 6.23, 6.24, as well as Lemma 6.1 with high probability. Also, with high probability, by Chernoff's bound, we have $|R_i| = q_i pn \pm \sqrt{n} \log n$.

Let $X \subseteq G$ with $|X \triangle R| \leq p^{10^{10}} n/\log(n)^{10^{24}}$, $\sum X = 0$, and $|X| = 3m_3 + 4m_4 + 5m_5$. Notice that since $|R| = |R_3| + |R_4| + |R_5| = (q_3 + q_4 + q_5)pn \pm 3\sqrt{n} \log n = pn \pm 3\sqrt{n} \log n$ and $|R| = |X| \pm p^{10^{10}} n/\log(n)^{10^{23}}$, we have $pn = 3m_3 + 4m_4 + 5m_5 \pm p^{10^{10}} n/\log(n)^{10^{22}}$. Using the definitions of q_i, this implies that $|R_i| = q_i pn \pm \sqrt{n} \log n = im_i \pm p^{10^{10}} n/\log(n)^{10^{21}}$. In particular, these imply $|(R_4 \cup R_3) \cap X| \geq 4m_4 + 3m_3 - 2p^{10^{10}} n/\log(n)^{10^{20}} \geq 4m_4 + 0.1pn$, and $|(R_5 \cup R_3) \cap X| \geq 5m_5 + 3m_3 - 2p^{10^{10}} n/\log(n)^{10^{20}} \geq 5m_5 + 0.1pn$ (using $q_3 \geq q_4, q_5$).

If q_4 (or q_5) is less than $n^{-1/10^{100}}$, we will argue that we may assume that $m_4 = 0$ (or $m_5 = 0$), only at the expense of working with a set X with a slightly larger symmetric difference with X, i.e. a set X for which $|X \triangle R| \leq p^{10^{10}} n/\log(n)^{10^{23}}$. To see this, note if $q_4 \leq n^{-1/10^{100}}$, then we can find $4m_4 \leq n^{1-\varepsilon}$ many zero-sum sets of size 4. We can achieve this by invoking the property from Lemma 6.1. As $n^{1-\varepsilon} \ll p^{10^{10}} n/\log(n)^{10^{100}}$, and the set of vertices we delete are zero-sum, we can continue the argument with $m_4 = 0$, and the remaining vertices of X (which are zero-sum). If q_5 is also small, the same operation can be performed disjointly to find a small number of zero-sum sets of size 5.

If $m_i = 0$ for some $i \in \{4, 5\}$, set $X_i = \emptyset$. Otherwise, by the previous paragraph, the corresponding lemmas apply to R_i, and X_4 and X_5 are defined via the following operations. By Lemma 6.23 applied with $R = R_4$, $Z = (R_4 \cup R_3) \cap X$, pick a zero-sum set $X_4 \subseteq (R_4 \cup R_3) \cap X$ with $|X_4| = 4m_4$ and having $|X_4 \triangle R_4| \leq p^{10^{10}} n/\log(n)^{10^{19}}$. By Lemma 6.23 applied with $R = R_5$, $Z = (R_5 \cup R_3) \cap X \setminus X_4$, pick a zero-sum set $X_5 \subseteq (R_5 \cup R_3) \cap X \setminus X_4$ with $|X_5| = 5m_5$ and having $|X_5 \triangle R_5| \leq p^{10^{10}} n/\log(n)^{10^{19}}$. Let $X_3 = X \setminus (X_4 \cup X_5)$ and notice that $|X_3| = 3m_3$, $|X_3 \triangle R_3| \leq p^{10^{10}} n/\log(n)^{10^{18}}$, and $\sum X_3 = \sum X - \sum X_4 - \sum X_5 = 0$. From Lemma 6.24 we have that each X_i has a zero-sum partition into m_i sets of size i. Together, all of these sets give a zero-sum M-partition of X. □

6.3.2 Tannenbaum's problem

In the remainder of this section we characterise the multisets M for which it is possible to find a zero-sum M-partition of $G \setminus 0$ for an abelian group G, thereby resolving an old problem of Tannenbaum [58]. It is well-known that abelian groups with just one involution (an order 2 element of G) have that $\sum G \neq 0$, so a zero-sum partition in this case is impossible. Abelian groups without involutions have odd order, and for such groups a characterisation of the multisets M for which G has a zero-sum M-partition was given by Tannenbaum [57] (the necessary and sufficient condition is that $M \subseteq \{2, 3, \dots\}$ and $\sum M = n - 1$). Hence, we will be concerned with abelian groups with at least 3 involutions for the rest of the section. For groups with 3 involutions, a characterization was found by Zeng [63] (the necessary and sufficient condition is again that $M \subseteq \{2, 3, \dots\}$ and $\sum M = n - 1$). The characterization for groups with > 3 involutions turns out to be substantially more involved, see Theorem 6.36.

We begin with the following technical lemma that allows us to work with a random partition which will be critical for the proof of Lemma 6.27.

Lemma 6.26 *Let G be a size n set and suppose G is partitioned as $G = \{g_1, h_1\} \cup \dots \{g_m, h_m\} \cup I$ (with $|I| + 2m = n$). Let Y be a p-random subset of $[m]$ and set $X = I \cup \bigcup_{i \in Y} \{g_i, h_i\}$. Then, we can partition $X = Q \cup R \cup S$ where Q, R are disjoint and $p/2$-random subsets of G, and S is a $(1 - p)$-random subset of I.*

Proof For any i, notice that $P(\{g_i, h_i\} \cap X = \emptyset) = 1 - p$ and $P(\{g_i, h_i\} \subseteq X) = p$, and $P(|\{g_i, h_i\} \cap X| = 1) = 0$, with these events being independent for different i. In order to make sure that elements end up in $Q/R/S$ independently, we need to define them carefully as follows: Define Q, R, S conditional on the outcome of Y as follows: For each $\{g_i, h_i\}$ such that $i \in Y$, place both g_i, h_i in Q with probability $a = p/4$, place both g_i, h_i in R with probability $a = p/4$, place g_i in Q and h_i in R with probability $b = 1/2 - p/4$, and place g_i in R and h_i in Q with probability $b = 1/2 - p/4$ (noting that $2a + 2b = 1$). Additionally place each $k \in I$ into Q with probability $p/2$, into R with probability $p/2$, and into S with probability $(1 - p)$. Do these latter set of choices independently of each other, and independently of the former choices made while choosing Y.

It remains to show that Q and R are disjoint $p/2$-random subsets of G, as that Q, R, S partitions X and that S is a $(1 - p)$-random subset of I follows directly. First, note that each element g of G is included in Q with probability $p(a + b) = p/2$, included in R with probability $p(a + b) = p/2$, and included in both Q and R with probability 0, regardless of whether $g \in I$. Now, we show that the collection of such events for each $g \in G$ are independent. Since choices for different i are already done independently, it is sufficient to show that for all i, $\{g_i, h_i\} \cap Q$ and $\{g_i, h_i\} \cap R$ are $p/2$-random subsets of $\{g_i, h_i\}$. Fix some i, and note that $P(g_i \in Q) = P(h_i \in Q) = p(a + b) = p/2$, $P(g_i, h_i \in Q) = pa = (p/2)^2$, $P(g_i \in Q, h_i \notin Q) = P(g_i \notin Q, h_i \in Q) = pb = (p/2)(1 - p/2)$, and $P(g_i, h_i \notin Q) = 1 - p + pa = (1 - p/2)^2$. The same holds for R (since Q and R are symmetric), and so we get that $\{g_i, h_i\} \cap Q$ and $\{g_i, h_i\} \cap R$ are $p/2$-random subsets of $\{g_i, h_i\}$ as required. □

Let $I(G)$ be the set of order 2 elements of G, noting that $I(G) \cup \{0\}$ is isomorphic to $\mathbb{Z}_2^k$ for some k. Define $f(G) = (|G| - |I(G)| - 1)/2$ and note that $f(G)$ is the number of inverse pairs in G (and so in particular an integer). A trivial necessary condition to have a zero-sum M-partition of $G \setminus 0$ is that $f(G) \geq m_2(M)$ (otherwise some involution would have to be contained in a zero-sum 2-set of the partition, which is impossible). The following lemma shows that for any multiset M it is possible to find a zero-sum M-partition of a group G as long as $m_2(M)$ is significantly smaller than $f(G)$.

Lemma 6.27 *Let G be a sufficiently large abelian group with $|I(G)| \geq 3$, and $M \subseteq \{2, 3, 4, 5, \dots\}$ a multiset with $\sum M = n - 1$ and $m_2(M) \leq f(G) - 0.0001n$. Then $G \setminus \{0\}$ has a zero-sum M-partition.*

Proof As in Lemma 6.25, without loss of generality, we can assume that $M \subseteq \{2, 3, 4, 5\}$. Define $m_i := m_i(M)$ for $i = 2, 3, 4, 5$. The basic idea will be to partition use Lemma 6.25 to get $m_3/m_4/m_5$ sets of orders $3/4/5$ in such a way that the remaining elements form inverse pairs. We will apply Lemma 6.25 three times in order to achieve this, twice to random subsets R, Q in the group G, and once to a random subset S in the group $I(G) \cup e$. In the paragraphs that follow, we construct these random sets, and check the properties needed for Lemma 6.25.

Note that $\sum_{i \geq 3} im_i = \sum M - 2m_2 = n - 1 - 2m_2 = 2f(G) + |I(G)| - 2m_2 \geq |I(G)| + 0.0002n$. Let $p = \frac{3m_3+4m_4+5m_5-|I(G)|}{2f(G)} = \frac{\sum M - 2m_2 - |I(G)|}{2f(G)} = 1 - \frac{m_2}{f(G)} \geq \frac{f(G)-m_2(G)}{n} \geq 0.0001$. If $|I(G)| \geq p^{10^{10}} n/\log(n)^{10^{26}}$, then set $m_i^Q = \lceil \frac{1}{2} \frac{pm_i n}{n-1-2m_2} \rceil$, $m_i^R = \lceil \frac{1}{2} \frac{pm_i n}{n-1-2m_2} \rceil$, $m_i^S = m_i - 2\lceil \frac{1}{2} \frac{pm_i n}{n-1-2m_2} \rceil$ for $i = 3, 4, 5$. If $|I(G)| < p^{10^{10}} n/\log(n)^{10^{26}}$, then set $m_i^Q = \lfloor m_i/2 \rfloor$, $m_i^R = m_i - \lfloor m_i/2 \rfloor$ and $m_i^S = 0$ for $i = 3, 4, 5$. Note that in both cases, we have $m_i = m_i^Q + m_i^R + m_i^S$ for each $i = 3, 4, 5$. Also note that when $|I(G)| \geq p^{10^{10}} n/\log(n)^{10^{26}}$, we have $\sum_{i=3}^5 m_i^S \pm 12 = \sum_{i=3}^5 i(m_i^S \pm 1) = \sum_{i=3}^5 (im_i - \frac{pim_i n}{n-1-2m_2}) = |I(G)| + 2f(G) - 2m_2 - pn = |I(G)|(1-p) - p \leq |I(G)| - 3$, and that in all cases we have $\sum_{i=3}^5 im_i^Q \leq \sum_{i=3}^5 im_i - \sum_{i=3}^5 im_i^S - 3$. For $* = Q, R, S$, define $M^* = \{m_3^* \times 3, m_4^* \times 4, m_5^* \times 5\}$ and note that $M^Q \cup M^R \cup M^S \cup \{m_2 \times 2\} = M$.

Recall $2f(G) = n - |I(G)| - 1$ and enumerate $G \setminus (I(G) \cup \{e\})$ as $g_1, g_1^{-1}, \dots, g_{f(G)}, g_{f(G)}^{-1}$. Pick a p-random set Y of $\{1, \dots, f(G)\}$, and set $X = I(G) \cup \bigcup_{i \in Y} \{g_i, g_i^{-1}\}$, noting that $\mathbb{E}(|X|) = 2pf(G) + |I(G)| = 3m_3 + 4m_4 + 5m_5$. Use Lemma 6.26 to partition $X = Q \cup R \cup S$ with Q, R $p/2$-random and S a $(1-p)$-random subset of $I(G)$.

Claim 6.27.1 *With positive probability, the following all hold.*

1. $f(G) - m_2 - \sqrt{n} \log n \leq |Y| \leq f(G) - m_2$.
2. $3m_3 + 4m_4 + 5m_5 - 2\sqrt{n} \log n \leq |X| \leq 3m_3 + 4m_4 + 5m_5$
3. $|Q|, |R|, |S| = 3m_3^* + 4m_4^* + 5m_5^* \pm p^{10^{10}} n/\log(n)^{10^{25}}$ (*for* $* = Q, R, S$).
4. *Q, R satisfy Lemmas 6.23 and 6.25 with respect to G.*
5. *If $|I(G)| \geq p^{10^{10}} n/\log(n)^{10^{26}}$, then S satisfies Lemmas 6.23 and 6.25 with respect to the subgroup $I(G) \cup e$.*

Proof We'll show that with probability $\geq 1/2$ the upper bounds of (1) and (2) both hold, whilst all other parts of the claim hold with high probability. For the upper bound of (1), note that $|Y|$ is a binomial random variable with expectation $pf(G) = (3m_3 + 4m_4 + 5m_5 - |I(G)|)/2 = f(G) - m_2 \in \mathbb{Z}$. For binomial random variables, if the expectation is an integer, then it is also the median. Thus we have that the median of $|Y|$ is $f(G) - m_2$ which shows that the upper bound (1) holds with probability $\geq 1/2$. The lower bound of (1) comes from Chernoff's bound. Note that $|X| = 2|Y| + |I(G)|$ and $2m_2 + 3m_3 + 4m_4 + 5m_5 = n - 1 = 2f(G) + |I(G)|$ show that (1) implies (2). For (3), note that Chernoff's bound gives $|Q|, |R| = pn/2 \pm \sqrt{n}\log n$ and $|S| = (1-p)|I(G)| \pm \sqrt{n}\log n$. When $|I(G)| \geq p^{10^{10}} n/\log(n)^{10^{26}}$, then $pn/2 = \frac{1}{2}\sum_{i=3}^{5} i\frac{pm_i n}{n-1-2m_2} = \sum_{i=3}^{5} im_i^Q \pm 3 = \sum_{i=3}^{5} im_i^R \pm 3$ and $(1-p)|I(G)| = \sum_{i=3}^{5} im_i^S \pm 13$ imply (3). When $|I(G)| < p^{10^{10}} n/\log(n)^{10^{26}}$, then $|S| \leq |I(G)|$ gives (3) for "S", while the result for Q, R then follows from (2) and $|Q| = |R| \pm 2\sqrt{n}\log n$. Properties (4), (5) are immediate from Lemmas 6.23 and 6.25. □

Partition $G \setminus (X \cup \{0\})$ into $f(G) - |Y|$ zero-sum sets of size 2 as $Z^2_{|Y|+1}, \dots, Z^2_{f(G)}$. Notice that we have $f(G) \geq f(G) - m_2 \geq |Y|$, and so we can set $J = Z^2_{|Y|+1} \cup \dots \cup Z^2_{f(G)-m_2}$ to get a zero-sum set of size exactly $2f(G) - 2m_2 - 2|Y| = n - 1 - 2m_2 - |X| = 3m_3 + 4m_4 + 5m_5 - |Q| - |R| - |S| \leq 3\sqrt{n}\log n$. When $|I(G)| \geq p^{10^{10}} n/\log(n)^{10^{26}}$, use Lemma 6.23 (with $Z = I(G)$) to pick a subset $S' \subseteq I(G)$ with $|S \triangle S'| \leq p^{10^{10}} n/\log(n)^{10^{25}}$, $|S'| = 3m_3^S + 4m_4^S + 5m_5^S$ and $\sum S' = 0$. Otherwise set $S' = \emptyset$ (In this application we use $Z = I(G)$ and $m = \sum_{i=3}^{5} m_i^S$ which satisfy $m \leq |Z| - 3$ from the first paragraph). Use Lemma 6.23 to find a set $Q' \subseteq (X \cup J) \setminus S'$ with $|Q \triangle Q'| \leq p^{10^{10}} n/\log(n)^{10^{24}}$, $|Q'| = 3m_3^Q + 4m_4^Q + 5m_5^Q$ and $\sum Q' = 0$ (In this application we use $Z = (X \cup J) \setminus S'$ and $m = \sum_{i=3}^{5} m_i^Q$ which satisfy $m \leq |Z| - 3$ from the first paragraph). Set $R' = (X \cup J) \setminus (Q' \cup S')$ and note that $|R \triangle R'| \leq p^{10^{10}} n/\log(n)^{10^{23}}$, $|R'| = 3m_3^R + 4m_4^R + 5m_5^R$, and $\sum R' = 0$. Apply Lemma 6.25 to $* = Q', R', S'$ to get M^*-partitions of these sets respectively. Putting the partitions of Q', R', S' together with the sets $Z^2_{m_2}, \dots, Z^2_{f(G)}$, we get a zero-sum M-partition of G. □

In the rest of the section we deal with groups and multisets having $f(G) - 0.0001 \leq m_2(M) \leq f(G)$. We say that a subset of an abelian group S is Σ-generic if for every proper non-empty subset $A \subseteq S$, we have that $\sum A$ is generic.

Lemma 6.28 *Let $3 \leq k \leq 10$ and let G be a sufficiently large abelian group with $|I(G)| \geq n/10^{9000}$. Then $I(G)$ contains a zero-sum set S of size k which is Σ-generic (in G).*

Proof Let $H = I(G) \cup e$ and recall that this is a subgroup. Let $v_1, \dots, v_k$ be free variables in $H * F_{k-1}$. Let $y = v_{k-1}^{-1} \dots v_1^{-1}$ and set $S = \{v_1, \dots, v_{k-1}, y\}$, noting that pairs $w, w' \in S$ are linear and separable (using part (a) of the definition of "separable"). For all nonempty $A \subseteq \{v_1, \dots, v_{k-1}\}$, define $g_A = \prod A$ and set $T = \{g_A : \{v_1, \dots, v_{k-1}\} \supseteq A \neq \emptyset\}$. Notice that all elements of T are linear. By Lemma 3.32, there is a projection $\pi : H * F_{k-1} \to H$ which separates $\pi(S \cup T)$ and has $\pi(S \cup T) \subseteq H \setminus N(G)$ (for this application, we have $p = 1$, $n' = |H| \geq n/10^{9000}$ and $U = N(G) \leq 10^{-9000} \leq p^{800} n/10^{4000}$ since n is sufficiently large). We have that $\pi(S)$ is a zero-sum set of size k (since $\prod S = e$ and all pairs $w, w' \in S$ are separable). The fact that $\pi(S)$ is Σ-generic follows from $\pi(T)$ being generic (for $A \subseteq S \setminus v_k$, the definition of "$\Sigma$-generic" is immediate for $\pi(A)$ due to $\pi(g_A)$ being generic. For A

containing v_k, note that $\sum \pi(A) = -\sum \pi(S \setminus A)$ which is generic because $\pi(g_{S\setminus A})$ is generic). □

The following theorem was poved independently by Caccetta-Jia, Engawa, and Tannenbaum and classifies what sorts of zero-sum partitions the groups $\mathbb{Z}_2^m$ has. We use it as a black box.

Theorem 6.29 (Cacceta-Jia [15]; Engawa [21]; Tannenbaum [58]) *For every multiset $M \subseteq \{3, 4, \dots\}$ with $\sum M = 2^m - 1$, the group $\mathbb{Z}_2^m \setminus \{0\}$ has a zero-sum M-partition.*

The following lemma is basically a version of this theorem, but additionally guarantees that one set in the partition is Σ-generic.

Lemma 6.30 *Let $m \geq 2$ and G be a sufficiently large abelian group of order n with $|I(G)| \geq 3$. Let $M \subseteq \{2, \dots, 10\}$ be a multiset with $\sum M = n - 1$ and $m_2(M) = f(G)$. Fix some $x \in M$. Then, there is a zero-sum M-partition of $G \setminus \{0\}$. Additionally we can assume that a size x set in this partition is Σ-generic (in G).*

Proof Let $M' \subseteq M$ be the sub-multiset consisting of all elements of size at least 3 and let $M'' = \{m_2(G) \times 2\}$. Note that $\sum M' = \sum M - 2m_2(M) = n - 1 - 2f(G) = |I(G)| - 1$ and that $G \setminus (I(G) \cup \{0\})$ has a zero-sum M''-partition $\mathcal{P}$.

If $|I(G)| \leq n/10^{9000}$, then note that for every g, the number of solutions to $x^2 = g$ in G is either $|I(G)| + 1$ or 0 (for any two such solutions x, x', we have $y = x^{-1}x'$ is a solution to $y^2 = e$ and the set of such solutions is exactly $I(G) \cup e$). This means that when $|I(G)| \leq n/10^{9000}$, everything is generic, and so all sets are Σ-generic. To get the lemma, use Theorem 6.29 to get a M'-partition of $I(G)$. Extend this to a M-partition satisfying the lemma by adding $\mathcal{P}$. The same argument works if $|I(G)| \geq n/10^{9000}$ and $x = 2$ (since there are at most $10^{9000} \leq f(G)$ non-generic elements, one of the sets in $\mathcal{P}$ always has only generic elements).

So, suppose that $|I(G)| \geq n/10^{9000}$ and $x \geq 3$. Let $H = I(G) \cup e$, recalling that this is a subgroup. Use Lemma 6.28 to find a Σ-generic set S_1 of size x. Let $X = I(G) \setminus (S_1 \cup e)$ and note that $|X \triangle H| \leq |H|/\log^{10^{24}}|H|$ and $\sum X = 0$. By Lemma 6.25 applied to H with $p = 1$, X has a zero-sum $(M' \setminus \{x\})$-partition. Together with S_1 and $\mathcal{P}$, this gives a zero-sum M-partition of G. □

The following two lemma deals with the case when m_2 and $f(G)$ are within a small additive constant of each other.

Lemma 6.31 *Let G be a sufficiently large abelian group with $|I(G)| \geq 3$. Let $M \subseteq \{2, \dots, 10\}$ be a multiset with $\sum M = n - 1$, $t := f(G) - m_2(M) \in \{0, \dots, 10\}$. Suppose that the t largest elements of M are all ≥ 3 and add up to $\geq 2t + 3$. Then $G \setminus \{0\}$ has a zero-sum M-partition.*

Proof If $f(G) = 0$, then $G \cong \mathbb{Z}_2^k$ and $m_2(M) = 0$, and so the lemma follows from Theorem 6.29. So, suppose $f(G) \geq 1$. Since $f(G) = (n - |I(G) \cup \{0\}|)/2 > 0$ and $I(G) \cup \{e\}$ is a subgroup, Lagrange's Theorem tells us that $|I(G) \cup \{0\}| \leq n/2$.

Let $y_1, \dots, y_t$ be the t largest elements of M, noting $y_1, \dots, y_t \geq 3$ and $y_1 + \dots + y_t \geq 2t + 3$. Let $y_0' = y_1 + \dots + y_t - 2t \geq 3$, $y_1', \dots, y_t' = 2$, and set $M' = (M \setminus \{y_1, \dots, y_t\}) \cup \{y_0', y_1', \dots, y_t'\}$. Note that $m_2(M') = m_2(M) + t = f(G)$ and $\sum M' = \sum M = n - 1$. Use Lemma 6.30 to find a zero-sum M-partition of G having a Σ generic size y_0' set A. Note that $y_0' = y_1 + \dots + y_t - 2t$ means that we can partition $A = A_1 \cup \dots \cup A_t$ into sets of size $|A_i| = y_i - 2$. Label $A_i = \{a_i^1, \dots, a_i^{y_i - 2}\}$ and set $a_i = \prod A_i = a_i^1 \dots a_i^{y_i - 2}$. Thinking of x as the free variable in $G * F_1$, define $S = \{x, a_1 x, a_2 a_1 x, \dots, a_{t-1} \dots a_2 a_1 x\}$. Also define the partition $S \cup S^{-1} \cup A = T_1 \cup T_2 \cup \dots \cup T_t$ with $T_1 = \{x, x^{-1} a_1^{-1}, a_1^1, \dots, a_1^{y_1 - 2}\}$, $T_2 = \{a_1 x, x^{-1} a_1^{-1} a_2^{-1}, a_2^1, \dots, a_2^{y_2 - 2}\}, \dots, T_t = \{a_{t-1} \dots a_2 a_1 x, x^{-1}, a_t^1, \dots, a_t^{y_t - 2}\}$ (more formally, if we denote $x_i := a_i \dots a_1 x$, then we have $S = \{x_0, \dots, x_{t-1}\}$ and $T_i = A_i \cup \{x_{i-1}, x^{-1}_{i \pmod{t-1}}\}$).

Note that for all $w \in S \cup S^{-1}$ are linear and all non-inverse $w, w' \in S \cup S^{-1}$ are separable (for this, first notice that by Σ-genericness, all partial product $a_i a_{i-1} \dots a_j$ are generic unless $i = t$, $j = 1$. Now if $w, w' \in S$ or $w, w' \in S^{-1}$ then the pair is separable by (c), while if $w \in S$, $w' \in S^{-1}$ the pair is separable by (b)). Using Lemmas 3.29 and 3.30, find a projection π which separates S and has $\pi(x) \notin I(G) \cup \{e\}$ (Lemma 3.30 gives us $0.9n$ projections which separate S, while Lemma 3.29 tells us that there are $|I(G) \cup \{0\}| \leq n/2$ projections with $\pi(x) \in I(G) \cup \{0\}$). Note that $\pi(x) \notin I(G) \cup \{0\}$ implies that $\pi(S)$ is disjoint from $I(G) \cup \{0\}$ (using that $a_1, \dots, a_k \in I(G) \cup \{0\}$ and $I(G) \cup \{0\}$ is a subgroup, we have that all $y \in \pi(S)$ are in one of the cosets $\pi(x)(I(G) \cup \{0\})$ or $\pi(x^{-1})(I(G) \cup \{0\})$). Note that this implies that all the elements in $\pi(S)$ are distinct (all non-inverse pairs $w, w' \in S$ are separable giving $\pi(w) \neq \pi(w')$. For $w, w' \in S$ with $w' = w^{-1}$, we cannot have $\pi(w) = \pi(w')$ since $\pi(S)$ is disjoint from $I(G) \cup \{e\}$). Now, in each case replace $\pi(S) \cup \pi(S^{-1}) \cup A$ in the original M'-partition, by $\pi(T_1), \dots, \pi(T_t)$. This gives a M-partition of G. □

The next lemma is a stronger version of the previous one with the "t largest elements of M are all ≥ 3" condition dropped.

Lemma 6.32 *Let G be a sufficiently large abelian group with $|I(G)| \geq 3$. Let $M \subseteq \{2, \dots, 10\}$ be a multiset with $\sum M = n - 1$, $t := f(G) - m_2(M) \in \{0, \dots, 5\}$. Suppose that the t largest elements of M add up to $\geq 2t + 3$. Then $G \setminus \{0\}$ has a zero-sum M-partition.*

Proof The proof is by induction on t. The initial cases are $t = 0$ and $t = 1$ which follow trivially from Lemma 6.31. Suppose that $t \geq 2$ and that the lemma is true for smaller t. Let $m_1 \geq \dots \geq m_t$ be the t largest elements of M and note that if $m_t \geq 3$, then the lemma follows from Lemma 6.31. Thus we can assume that $m_t = 2$. Note that $\sum_{m_i \geq 3} i m_i = |I(G)| + 2t \geq 2t + 3$ which gives $m_1 + \dots + m_t \geq 2t + 5$. Note that we cannot have $m_1 \leq 3$ since then $m_1 + \dots + m_{t-1} \leq 3(t - 1)$, giving $2t + 5 \leq m_1 + \dots + m_t \leq 3t - 1$ which is impossible for $t \leq 5$.

We have established that $m_1 \geq 4$ and $m_1 + \dots + m_t \geq 2t + 5$. Consider the multiset $M' = (M \setminus \{m_1\}) \cup \{m_1 - 2, 2\}$. If $m_1 \geq 5$, then we have $f(G) - m_2(M') = t - 1$ and the $t - 1$ largest elements of M' add up to $\geq (m_1 - 2) + m_2 + \dots + m_{t-1} =$

$(m_1+m_2+\cdots+m_t)-4\geq(2t+5)-4=2(t-1)+3$. If $m_1=4$, then we have $f(G)-m_2(M')=t-2$ and the $t-2$ largest elements of M' add up to $\geq m_2+\cdots+m_{t-1}=(m_1+m_2+\cdots+m_t)-6\geq(2t+5)-6=2(t-2)+3$. In either case, by induction, there is a zero-sum M'-partition of G. Combining the size m_1-2 and 2 sets in this partition into a size m_1 set produces a M-partition. □

We'll need the following result about zero-sum 3-sets.

Lemma 6.33 *Let $G\neq\mathbb{Z}_2^m,\mathbb{Z}_2^m\times\mathbb{Z}_4$ be a sufficiently large abelian group. Then there are $n/100$ disjoint triples of distinct non-involutions x,y,z with $x+y+z=0$.*

Proof In any large abelian group G, Lemma 6.25 applied with $p=1$, and $X\subseteq G\setminus\{0\}$ a zero-sum set of cardinality at least $n-4$ and divisible by 3 gives $0.33332n$ disjoint zero-sum 3-sets (that such a subset X exists follows from Lemma 6.23). If $G\neq\mathbb{Z}_2^m,\mathbb{Z}_2^m\times\mathbb{Z}_4,\mathbb{Z}_2^m\times\mathbb{Z}_3$, then $|I(G)\cup\{0\}|\leq 0.25n$. At least $0.33332n-0.25n\geq 0.01n$ of these zero-sum 3-sets must also be disjoint from $I(G)\cup\{0\}$. All of these sets satisfy the lemma, so it remains to just look at $\mathbb{Z}_2^m\times\mathbb{Z}_3$. In this case, we know from Theorem 6.29 that $\mathbb{Z}_2^m$ has $0.1n$ disjoint zero-sum 3-sets $\{x_1,y_1,z_1\},\ldots,\{x_{0.1n},y_{0.1n},z_{0.1n}\}$. Let $(0,1)$ be the generator of the $\mathbb{Z}_3$ subgroup of $\mathbb{Z}_2^m\times\mathbb{Z}_3$. Now $\{(x_1,1),(y_1,1),(z_1,1)\},\ldots,\{(x_{0.1n},1),(y_{0.1n},1),(z_{0.1n},1)\}$ are zero-sum 3-sets of non-involutions. □

Somewhat annoyingly, our proof doesn't quite work with the group $\mathbb{Z}_2^m\times\mathbb{Z}_4$, so we deal with this group by hand in the following lemma.

Lemma 6.34 *For sufficiently large m, let $G=\mathbb{Z}_2^m\times\mathbb{Z}_4$. Let $t\in[3,0.001f(G)]$ with $|I(G)|+2t\equiv 0\pmod 3$ and consider the multiset $M=\{(f(G)-t)\times 2,\frac{1}{3}(|I(G)|+2t)\times 3\}$. Then $(\mathbb{Z}_2^m\times\mathbb{Z}_4)\setminus\{0\}$ has a zero-sum M-partition.*

Proof Note $f(G)=2^m$ and $|I(G)|=2^{m+1}-1$. Write $t=k+3s$ for $k\in\{3,4,5\}$ and $s\in\mathbb{Z}$. Note that $|I(G)|+2k\equiv|I(G)|+2(k+3s)\equiv|I(G)|+2t\pmod 3$, and so we can define a multiset $M'=\{(f(G)-k)\times 2,\frac{1}{3}(|I(G)|+2k)\times 3\}$. Note that $\sum M'=2(f(G)-k)+\frac{3}{3}(|I(G)|+2k)=n-1$, and that the k largest elements of M' add up to $\geq\min(3k,|I(G)|+2k)\geq 2k+3$. Thus Lemma 6.32 gives us a zero-sum M'-partition of G. Let $\mathcal{S}$ be the family of 3-sets in this partition. Let $K=\bigcup\mathcal{S}\setminus I(G)$, noting that $|K|=2k$.

Let $(0,1)$ be the generator of the $\mathbb{Z}_4$ subgroup of $\mathbb{Z}_2^m\times\mathbb{Z}_4$. Note that $I(G)=\{(x,0),(x,2):x\in\mathbb{Z}_2^m\}\setminus(0,0)$. Let $K'=K\cup(K+(0,1))\cup(K+(0,2))\cup(K+(0,3))$ noting that $|K'|\leq 8k$. Let $\mathcal{S}'\subseteq\mathcal{S}$ be the subfamily of 3-sets which are disjoint from K' and of the form "$\{(x,0),(y,0),(z,0)\}$". We claim that $|\mathcal{S}'|\geq 0.01n$. To see this, first not that there are at most $|K'|\leq 24$ 3-sets intersecting K'. All the remaining $\geq\frac{1}{3}(|I(G)|+2k)-24=\frac{1}{3}2^{m+1}-24$ sets must be of the form $\{(x,0),(y,0),(z,0)\}$ or $\{(x,2),(y,2),(z,0)\}$ for some $x,y,z\in\mathbb{Z}_2^m$. There are at most $2^m/2$ 3-sets of the form $\{(x,2),(y,2),(z,0)\}$ in $\mathcal{S}$ (since $|\{(x,2):x\in\mathbb{Z}_2^m\}|=2^m$ and every triple of this form used two elements of $\{(x,2):x\in\mathbb{Z}_2^m\}$), leaving us with $\frac{1}{3}2^{m+1}-24-2^m/2\geq 0.01n$ 3-sets in $\mathcal{S}'$.

For each $S \in \mathcal{S}'$, let $S = \{a_S, b_S, c_S\}$ and note that $A_S = \{a_S + (0,1), a_S + (0,3)\}$, $B_S = \{b_S + (0,1), b_S + (0,3)\}$, $C_S = \{c_S + (0,1), c_S + (0,3)\}$ must be distinct 2-sets in the partition (using that $S \in \mathcal{S}'$). Define three zero-sum sets $T_S^a = \{a_S, b_S + (0,1), c_S + (0,3)\}$, $T_S^b = \{b_S, c_S + (0,1), a_S + (0,3)\}$, $T_S^a = \{c_S, a_S + (0,1), b_S + (0,3)\}$. Now for each $S \subseteq \mathcal{S}'$, replace S, A_S, B_S, C_S in the original partition by T_S^a, T_S^b, T_S^c to get a M-partition of G. □

The following lemma finds zero-sum M-partitions for multisets M with $f(G) - m_2(G)$ outside of the ranges considered by Lemmas 6.27 and 6.32.

Lemma 6.35 *Let G be a sufficiently large abelian group with $|I(G)| \geq 3$. Let $M \subseteq \{2, \dots, 10\}$ be a multiset with $f(G) - m_2(M) \in [3, 0.001n]$. Then $G \setminus e$ has a zero-sum M-partition.*

Proof As in Lemma 6.31, we can assume that $|I(G) \setminus 0| \leq n/2$. The proof is by induction on t. The initial cases are when $t = 3, 4, 5$, in which case the lemma follows from Lemma 6.32 (to see this, we need to know that the condition "the t largest elements of M add up to at least $2t + 3$" holds for these values of t. Letting $y_1 \geq \dots \geq y_t$ be the t largest elements, note that $\sum_{i=1}^t y_t \geq t y_t$ and $\sum_{i=1}^t y_t \geq \sum_{i > y_t} i m_i(M)$ both hold. Now, using $\sum_{i \geq 3} i m_i(M) = n - 1 - 2m_2(M) = |I(G)| + 2t \geq 2t + 3$, we get $\sum_{i=1}^t y_i \geq \min(3t, \sum_{i \geq 3} i m_i(M)) \geq \min(3t, 2t + 3) = 2t + 3$). Now suppose that $t \geq 6$ and that the lemma holds for multisets M' with $f(G) - m_2(M') < t$ and let $M \subseteq \{2, \dots, 10\}$ be a multiset with $f(G) - m_2(M) = t$, and $\sum M = n - 1$.

Suppose that $\max M \geq 4$. Let $M' = (M \setminus \{\max M\}) \cup \{\max M - 2, 2\}$. Note that $f(G) - m_2(M') = t - 1$ or $t - 2$ (depending on whether $\max M = 4$ or not). By induction, there is a zero-sum M'-partition of G. Combining the size $\max M - 2$ and 2 sets in this partition into a size $\max M$ set produces a M-partition.

Suppose that $\max M \leq 3$, which implies that $M \subseteq \{2, 3\}$. If $G = \mathbb{Z}_2^m \times \mathbb{Z}_4$, then the lemma follows from Lemma 6.34, so suppose this doesn't happen. Let $M' = (M \setminus \{3, 3\}) \cup \{2, 2, 2\}$. Note that $f(G) - m_2(M') = t - 3$, so by induction, we get an M'-partition of G. In this partition there are $2(t - 3) \leq 0.01n$ non-involutions in sets of size ≥ 3. Therefore Lemma 6.33 gives us a zero sum set $\{x, y, z\}$ of non-involutions so that $\{x, -x\}$, $\{y, -y\}$, $\{z, -z\}$ are sets in the M'-partition. Now replace $\{x, -x\}$, $\{y, -y\}$, $\{z, -z\}$ by $\{x, y, z\}$, $\{-x, -y, -z\}$ to get an M-partition of G. □

The following is the main result of this section. Together with earlier results of Tannenbaum [57] and Zeng [63], it answers a problem of Tannenbaum from 1983 [58].

Theorem 6.36 *Let G be a sufficiently large abelian group with $|I(G)| \geq 3$ and let $M \subseteq \{2, 3, 4, \dots\}$ be a multiset. Then there is a zero-sum M-partition of $G \setminus \{0\}$ if, and only if, all of the following are true.*

1. *$\sum M = n - 1$ and $f(G) \geq m_2(M)$.*
2. *If $f(G) = m_2(M) + 1$, then $\max M \geq 5$.*

Proof "If" direction: Let $t = f(G) - m_2(M)$ and note that this is ≥ 0 by (1). As in Lemma 6.25, without loss of generality, we can suppose that $\max M \leq 10$. Note that if $t \geq 3$, then we get a zero-sum M-partition by Lemma 6.27 or Lemma 6.35. Otherwise, we apply Lemma 6.32, where we additionally need the condition that "the t largest elements of M add up to at least $2t + 3$". For $t = 1$, we have $\max M \geq 5 \geq 2 \times 1 + 3$. For $t = 2$, let $x_1, x_2 \in M$ be the two largest elements of M. We have $x_1, x_2 \geq 3$ (since $\sum_{x_i \geq 3} x_i = |I(G)| + 2t \geq 3 + 2 \times 2$ which implies that $x_1, x_2 \neq 2$). We also have $x_1 \geq 4$ (otherwise we'd have $M \subseteq \{2, 3\}$ giving $3m_3(M) = \sum M - 2m_2(M) = n - 1 - 2m_2 = |I(G)| + 2f(G) - 2m_2 = |I(G)| + 4$, which gives a contradiction since on one hand $3m_3(M) \equiv 0 \pmod 3$, but on another hand $|I(G)| + 1 = 2^j$ for some j and so $|I(G)| + 4 = 2^j + 3 \equiv 1$ or $2 \pmod 3$)). These give $x_1 + x_2 \geq 4 + 3 \geq 2 \times 2 + 3$.

"Only if" direction: Consider some zero-sum M-partition $G \setminus e = S_1 \cup \dots S_k$. Then, by the definition of "M-partition", we have the equality between multisets $\{|S_i| : i = 1, \dots, k\} = M$, which gives $\sum M = |G \setminus e| = n - 1$. Suppose that $S_1, \dots, S_{m_2(M)}$ are the size 2 sets of this partition. Then for each i, we must have $S_i = \{g_i, g_i^{-1}\}$, where g_i are distinct non-involutions of G. This means that $S_1 \cup \dots \cup S_{m_2(M)} \subseteq G \setminus (I(G) \cup e)$ giving $2m_2 = |S_1 \cup \dots \cup S_{m_2(M)}| \leq n - 1 - |I(G)| = 2f(G)$. Thus we have established (1).

If $f(G) = m_2(M) + 1$, then we learn that there is some unique non-involution g with $g, -g \notin S_1 \cup \dots \cup S_{m_2(M)}$. Note that it's impossible for a zero-sum set to have precisely one non-involution (since $I(G) \cup e$ is a subgroup of G), and so we get that $g, -g \subseteq S_j$ for some zero-sum set S_j in the partition. Letting $S_j = \{g, -g, a_1, \dots, a_{|S_j|-2}\}$, we get that $a_1 + \dots + a_{|S_j|-2} = 0$. This implies that $|S_j| \geq 5$, since it's impossible for 1 or 2 distinct non-identity involutions to sum to 0 in an abelian group. □

7 Concluding remarks

7.1 Further applications

Our methods here have implications for several other conjectures/problems in the area. Bors and Wang [12] used some of our results to study the group generated by all complete mappings of a finite group G. The first author [49] used our results to prove the Friedlander-Gordon-Tannenbaum Conjecture about possible cycle types of orthomorphisms in groups. Another relevant problem is the conjecture of Graham and Sloane on harmonious labellings of trees [35]. As this application requires novel ideas, we defer exploring it to a future paper. For the interested reader wishing to get involved in the area, we now list several directions of further research that we believe to be of interest.

7.2 Counting aspect

Using more refined arguments in place of the nibble-type arguments we use in the current paper (see for example [11]), the authors anticipate that Theorem 1.1 could be strengthened to give the existence of exponentially many complete mappings, as opposed to just one. Counting spanning structures with this approach has been quite fruitful recently to address the famous n-queens problem [13, 54]. We did not pursue this direction in this paper, as the proof of the Hall-Paige conjecture by Eberhard, Manners, and Mrazović [20] gives a much more precise count on the number of complete mappings than we could hope to accomplish with our methods.

However, we are curious if Fourier-analytic methods used in [20] could be used to count sequencings in sequenceable groups, or transversals in subsquares of multiplication tables. This would potentially give alternative resolutions to conjectures of Ringel [51] and Snevily [56].

7.3 Small n

Our methods only work for large n, and it seems hopeless to extend them to work for all n. However some of our theorems likely do extend to all n — and it would be interesting to reprove them for all n using different methods. One such theorem is our classification of subsquares without transversals.

Conjecture 7.1 *The following holds for all n. Let G be a group, and let $A, B \subseteq G$ with $|A| = |B| = n$. Then, $A \times B$ has a transversal, unless there exists some $k \geq 1$, $g_1, g_2 \in G$ and a subgroup $H \subseteq G$ such that one of the following holds.*

1. *H is a group that does not satisfy the Hall-Paige condition, and $A = g_1 H$ and $B = H g_2$.*
2. *$H \cong (\mathbb{Z}_2)^k$ and $g_1 A = H \setminus \{a_1, a_2\}$, $g_2 B = H \setminus \{b_1, b_2\}$ for some distinct $a_1, a_2 \in H$ and distinct $b_1, b_2 \in H$ such that $a_1 + a_2 + b_1 + b_2 = 0$*

It wouldn't make sense to extend Theorem 4.4 to all n (due to the probabilistic nature of the statement of that theorem). However if we set $p = 1$ and $|(R_A^1 \cup R_B^2 \cup R_C^3) \Delta (X \cup Y \cup Z)|$ to be very small, we expect that the following should hold.

Conjecture 7.2 *Let G be a group of order n. Let X, Y, Z be equal-sized subsets of G_A, G_B, and G_C of size $n-1$ having $\sum X + \sum Y + \sum Z = 0$ (in G^{ab}). Then, $H_G[X, Y, Z]$ contains a perfect matching.*

Note that the above implies the Hall-Paige conjecture for all n, by setting X, Y, Z to be $G \setminus \{e\}$.

7.4 Bounds in the main theorem

It would be interesting to sharpen the bounds in e.g. Theorem 4.4 in various ways. One interesting parameter to optimize is the size of the symmetric difference $|(R_A^1 \cup R_B^2 \cup R_C^3) \Delta (X \cup Y \cup Z)|$. We proved this with the upper bound $|(R_A^1 \cup R_B^2 \cup R_C^3) \Delta (X \cup Y \cup Z)| \leq p^{10^{10}} n / \log(n)^{10^{11}}$, and it would be interesting to figure out what the best possible bound one could impose here is.

One improvement that can be made with essentially no modifications to the proof is to use the upper bound $|(R_A^1 \cup R_B^2 \cup R_C^3) \Delta (X \cup Y \cup Z)| \leq p^{10^{10}} n / \log(|G'|)^{10^{11}}$. To obtain this, simply replace instances of $\log n$ by $\log |G'|$ throughout the proof (and observe that the ultimate source of all logarithms is Theorem 5.11).

From the other side, it is easy to see that Theorem 4.4 isn't true with $|(R^1_A \cup R^2_B \cup R^3_C)\Delta(X \cup Y \cup Z)| \gg p^2n$ — this is because under the assumptions of that theorem, vertices have $\approx p^2n$ edges going into $R^1_A \cup R^2_B \cup R^3_C$. If $|(R^1_A \cup R^2_B \cup R^3_C)\Delta(X \cup Y \cup Z)| \gg p^2n$, then it would be possible to choose X, Y, Z satisfying this where some vertices have no edges going into $X \cup Y \cup Z$ (and so there will be no perfect matching). It would be interesting to sharpen these bounds.

It would also be interesting to improve the bounds on Theorem 4.4 in the special case when $p = 1$. In this case, the theorem reduces to a non-probabilistic statement.

Problem 7.3 For each n, what is the largest number $f(n)$ such that the following is true?

Let G be a group of order n. Let X, Y, Z be subsets of G_A, G_B, and G_C respectively, each of size $n - f(n)$ with $\sum X + \sum Y + \sum Z = 0$ (in G^{ab}). Then, $H_G[X, Y, Z]$ contains a perfect matching.

From Theorem 4.4, we get that for large n, we have $f(n) \geq n/\log(n)^{10^{11}}$. As discussed above, the logarithmic factors are redundant for abelian groups, so in this case, one could hope for a much more precise understanding of the corresponding function $f(n)$. Concretely, we propose the following.

Problem 7.4 For each n, what is the largest number $g(n)$ such that the following is true?

Let G be the cyclic group of order n. Let X, Y, Z be subsets of G_A, G_B, and G_C respectively, each of size $n - g(n)$ with $\sum X + \sum Y + \sum Z = 0$. Then, $H_G[X, Y, Z]$ contains a perfect matching.

From our results, it follows that $g(n) = \Omega(n)$, and it is not hard to see that $g(n) \leq n/2$. It would already be interesting to determine $g(n)$ asymptotically. One can also consider the non-partite version of the problem. That is, what is the smallest subset $S \subseteq \mathbb{Z}_n$ with $|S|$ divisible by 3, $\sum S = 0$, and S cannot be partitioned into triples with zero-sum?

7.5 Strong complete mappings

An **orthomorphism** of a group G is a bijection $\psi : G \to G$ such that $x \to x^{-1}\psi(x)$ is also bijective. A **strong complete mapping** of a group G is a complete mapping which is also an orthomorphism. Evans raised the fascinating problem of characterising the groups G which contain strong complete mappings (see [24]). We remark that a strong complete mapping of cyclic groups corresponds to the placement of non-attacking queens on a toroidal chessboard. Using this correspondence, a recent result of Bowtell and Keevash [13] can be interpreted as an estimation on the number of strong complete mappings of cyclic groups. It would be very interesting to see if methods we develop in this paper for general groups can be combined with the strategies in [13] to make progress on Evans' problem.

7.6 Mappings of groups with other properties

Let S be a multiset with elements coming from a group G. When is there a bijection $\phi\colon G \to G$ such that the multiset $\{x\phi(x)\colon x \in G\}$ is equal to S? The Hall-Paige conjecture corresponds to the case when $S = G$. In [37], Hall answers this question for abelian groups. It would be interesting to generalise these results to non-abelian groups. In [6], a conjecture in this direction is given in the setting of Latin squares (quasi-groups). This seems like an exciting direction to generalise the Ryser-Brualdi-Stein conjecture [42].

7.7 The Kézdy-Snevily conjecture

We end with another problem similar in spirit to the previous one, but this time concerned only with cyclic groups. It was proposed initially by Snevily in 2000 [56], and reiterated by Kézdy and Snevily [43].

Problem 7.5 For any positive k and n with $k < n$, show that any sequence $a_1, a_2, \ldots, a_k$ of not necessarily distinct elements of $\mathbb{Z}_n$ admits a permutation π such that the sequence $a_{\pi(1)} + 1, a_{\pi(2)} + 2, \ldots, a_{\pi(k)} + k$ are all distinct (in $\mathbb{Z}_n$).

The results in the current paper can be used to address the above problem for large k, whenever the sequence $a_1, a_2, \ldots, a_k$ does not contain repetitions. Alon addressed the above problem whenever n is prime [4]. The previously stated result of Hall [37] can be used to address the problem when $k = n - 1$. Kézdy and Snevily [43] solved the problem when $2k \le n + 1$. Otherwise, the problem seems to be wide open.

We refer the reader to a 2013 survey by Ron Graham that includes an amusing interpretation of the above problem [34]. The survey by Ullman and Velleman [60] also contains a nice exposition for results of this flavour, including the aforementioned result of Hall.

Acknowledgements We would like to thank an anonymous referee for a careful reading of this manuscript and helpful suggestions that simplified several of our proofs.

References

1. Ajtai, M., Komlós, J., Szemerédi, E.: Sorting in $c \log n$ parallel steps. Combinatorica **3**(1), 1–19 (1983)
2. Akbari, S., Alipour, A.: Transversals and multicolored matchings. J. Comb. Des. **12**(5), 325–332 (2004)

3. Alon, N.: Combinatorial nullstellensatz. Comb. Probab. Comput. **8**(1–2), 7–29 (1999)
4. Alon, N.: Additive Latin transversals. Isr. J. Math. **117**(1), 125–130 (2000)
5. Alon, N., Spencer, J.: The Probabilistic Method. Wiley, New York (2004)
6. Anastos, M., Fabian, D., Müyesser, A., Szabó, T.: Splitting matchings and the Ryser-Brualdi-Stein conjecture for multisets (2022). arXiv:2212.03100. arXiv preprint
7. Arsovski, B.: A proof of Snevily's conjecture. Isr. J. Math. **182**(1), 505–508 (2011)
8. Batcher, K.E.: Sorting networks and their applications. In: Proceedings of the April 30–May 2, 1968, Spring Joint Computer Conference, AFIPS '68 (Spring), pp. 307–314. Association for Computing Machinery, New York (1968)
9. Bate, S., Jones, B.: A review of uniform cross-over designs. J. Stat. Plan. Inference **138**(2), 336–351 (2008)
10. Beals, R., Gallian, J.A., Headley, P., Jungreis, D.: Harmonious groups. J. Comb. Theory, Ser. A **56**(2), 223–238 (1991)
11. Bennett, P., Bohman, T.: A natural barrier in random greedy hypergraph matching. Comb. Probab. Comput. **28**(6), 816–825 (2019)
12. Bors, A., Wang, Q.: Compositions and parities of complete mappings and of orthomorphisms. J. Comb. Theory, Ser. A **196**, 105723 (2023)
13. Bowtell, C., Keevash, P.: The n-queens problem (2021). arXiv:2109.08083. arXiv preprint
14. Bray, J.N., Cai, Q., Cameron, P.J., Spiga, P., Zhang, H.: The Hall–Paige conjecture, and synchronization for affine and diagonal groups. J. Algebra **545**, 27–42 (2020)
15. Caccetta, L., Jia, R.-Z.: Binary labeling of graphs. Graphs Comb. **13**(2), 119–137 (1997)
16. Cichacz, S.: Zero sum partition of Abelian groups into sets of the same order and its applications. Electron. J. Comb. **25**(1) (2018)
17. Cichacz, S., Suchan, K.: Zero-sum partitions of Abelian groups of order 2^n. Discret. Math. Theor. Comput. Sci. **25**(Combinatorics) (2023)
18. Cormen, T.H.: Introduction to Algorithms. The MIT Press (2009)
19. Dasgupta, S., Károlyi, G., Serra, O., Szegedy, B.: Transversals of additive Latin squares. Isr. J. Math. **126**(1), 17–28 (2001)
20. Eberhard, S., Manners, F., Mrazović, R.: An asymptotic for the Hall–Paige conjecture. Adv. Math. **404**, 108423 (2022)
21. Egawa, Y.: Graph labelings in elementary Abelian 2-groups. Tokyo J. Math. **20**(2), 365–379 (1997)
22. Erdős, P., Gyárfás, A., Pyber, L.: Vertex coverings by monochromatic cycles and trees. J. Comb. Theory, Ser. B **51**(1), 90–95 (1991)
23. Evans, A.: The admissibility of sporadic simple groups. J. Algebra **321**(1), 105–116 (2009)
24. Evans, A.B.: The existence of strong complete mappings of finite groups: a survey. Discrete Math. **313**(11), 1191–1196 (2013)
25. Evans, A.: Applications of complete mappings and orthomorphisms of finite groups. Quasigr. Relat. Syst. **23**(1), 5–30 (2015)
26. Evans, A.B.: Orthogonal Latin Squares Based on Groups, vol. 57. Springer, Berlin (2018)
27. Frankl, P., Rödl, V.: Near perfect coverings in graphs and hypergraphs. Eur. J. Comb. **6**(4), 317–326 (1985)
28. Friedlander, R., Gordon, B., Tannenbaum, P.: Partitions of groups and complete mappings. Pac. J. Math. **92**(2), 283–293 (1981)
29. Gallagher, P.: Group characters and commutators. Math. Z. **79**(1), 122–126 (1962)
30. GH from MO (https://mathoverflow.net/users/11919/gh-from mo). How many square roots can a non-identity element in a group have? MathOverflow. https://mathoverflow.net/q/410789 (version: 2021-12-15)
31. Goddyn, L., Halasz, K.: All group-based Latin squares possess near transversals. J. Comb. Des. **28**(5), 358–365 (2020)
32. Gordon, B.: Sequences in groups with distinct partial products. Pac. J. Math. **11**(4), 1309–1313 (1961)
33. Gowers, W.: Probabilistic combinatorics and the recent work of Peter Keevash. Bull. Am. Math. Soc. **54**(1), 107–116 (2017)
34. Graham, R.: Juggling mathematics and magic. Not. Int. Consort. Chin. Math. **1**(1), 7–9 (2013)
35. Graham, R.L., Sloane, N.J.A.: On additive bases and harmonious graphs. SIAM J. Algebraic Discrete Methods **1**(4), 382–404 (1980)
36. Green, B., Wigderson, A.: Lecture notes for the 22nd McGill invitational workshop on computational complexity (2010)
37. Hall, M.: A combinatorial problem on Abelian groups. Proc. Am. Math. Soc. **3**(4), 584–587 (1952)

38. Hall, M., Paige, L.: Complete mappings of finite groups. Pac. J. Math. **5**, 541–549 (1955)
39. Hanani, H.: A note on Steiner triple systems. Math. Scand. **8**, 154–156 (1960)
40. Keedwell, A.D.: On the sequenceability of non-Abelian groups of order pq. Discrete Math. **37**(2–3), 203–216 (1981)
41. Keevash, P.: The existence of designs (2014). arXiv:1401.3665. arXiv preprint
42. Keevash, P., Pokrovskiy, A., Sudakov, B., Yepremyan, L.: New bounds for ryser's conjecture and related problems. Trans. Amer. Math. Soc. Ser. B **9**(8), 288–321 (2022)
43. Kézdy, A.E., Snevily, H.S.: Distinct sums modulo n and tree embeddings. Comb. Probab. Comput. **11**(1), 35–42 (2002)
44. Kühn, D., Lapinskas, J., Osthus, D., Patel, V.: Proof of a conjecture of Thomassen on Hamilton cycles in highly connected tournaments. Proc. Lond. Math. Soc. **109**(3) (2014)
45. Molloy, M., Reed, B.: Near-optimal list colorings. Random Struct. Algorithms **17**(3–4), 376–402 (2000)
46. Montgomery, R.: Spanning trees in random graphs. Adv. Math. **356** (2019)
47. Montgomery, R., Pokrovskiy, A., Sudakov, B.: Decompositions into spanning rainbow structures. Proc. Lond. Math. Soc. **119**, 04 (2019)
48. Montgomery, R., Pokrovskiy, A., Sudakov, B.: A proof of Ringel's conjecture. Geom. Funct. Anal. **31** (2021)
49. Müyesser, A.: Cycle type in Hall-Paige: a proof of the Friedlander-Gordon-Tannenbaum conjecture (2023). arXiv:2303.16157. arXiv preprint
50. Ollis, M.: Sequenceable groups and related topics. Electron. J. Comb. **1000**, DS10 (2002)
51. Ringel, G.: Cyclic arrangements of the elements of a group. Not. Am. Math. Soc. **21**(1), A-95 (1974)
52. Ringel, G.: Map Color Theorem, vol. 209. Springer, Berlin (2012)
53. Rödl, V., Szemerédi, E., Ruciński, A.: An approximate Dirac-type theorem for k-uniform hypergraphs. Combinatorica **28**(2), 229–260 (2008)
54. Simkin, M., Luria, Z.: A lower bound for the n-queens problem. In: Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), pp. 2185–2197. SIAM, Philadelphia (2022)
55. Skolem, T.: On certain distributions of integers in pairs with given differences. Math. Scand. **5**(1), 57–68 (1957)
56. Snevily, H.S.: The Cayley addition table of $\mathbb{Z}_n$. Am. Math. Mon. **106**(6), 584–585 (1999)
57. Tannenbaum, P.: Partitions of Abelian groups into sets with zero sums. Congr. Numer. **33**, 453–461 (1981)
58. Tannenbaum, P.: Partitions of $\mathbb{Z}_2^n$. SIAM J. Algebraic Discrete Methods **4**(1), 22–29 (1983)
59. Thomason, A.: Dense expanders and pseudo-random bipartite graphs. In: In Annals of Discrete Mathematics, vol. 43, pp. 381–386. Elsevier, Amsterdam (1989)
60. Ullman, D.H., Velleman, D.J.: Differences of bijections. Am. Math. Mon. **126**(3), 199–216 (2019)
61. Wanless, I.: Transversals in Latin Squares: A Survey. London Mathematical Society Lecture Note Series, pp. 403–437. Cambridge University Press, Cambridge (2011)
62. Wilcox, S.: Reduction of the Hall-Paige conjecture to sporadic simple groups. J. Algebra **321**(5), 1407–1428 (2009)
63. Zeng, X.: On zero-sum partitions of Abelian groups. Integers **15**, A44 (2015)