Asymptotically-tight packing and covering with transversal bases in Rota's basis conjecture

Richard Montgomery[*]   Lisa Sauermann[†]

August 7, 2025

Abstract

In 1989, Rota conjectured that, given any n bases $B_1, \ldots, B_n$ of a vector space of dimension n, or more generally a matroid of rank n, it is possible to rearrange these into n disjoint transversal bases. Here, a transversal basis is a basis consisting of exactly one element from each of the original bases $B_1, \ldots, B_n$. Two natural approaches to this conjecture are, to ask in this setting a) how many disjoint transversal bases can we find and b) how few transversal bases do we need to cover all the elements of $B_1, \ldots, B_n$? In this paper, we give asymptotically-tight answers to both of these questions.

For a), we show that there are always $(1 - o(1))n$ disjoint transversal bases, improving a result of Bucić, Kwan, Pokrovskiy, and Sudakov that $(1/2 - o(1))n$ disjoint transversal bases always exist. For b), we show that $B_1 \cup \cdots \cup B_n$ can be covered by $(1 + o(1))n$ transversal bases, improving a result of Aharoni and Berger using instead $2n$ transversal bases, and a subsequent result of the Polymath project on Rota's basis conjecture using $2n - 2$ transversal bases.

1 Introduction

Rota's basis conjecture is a famous conjecture concerning the rearrangeability of bases in a vector space or matroid. Rota posed it in 1989 (see [13, Conjecture 4]), both in the special case of vector spaces, and for general matroids (see Section 2 for the definition and the basic properties of a matroid). Given any n bases $B_1, \ldots, B_n$ of an n-dimensional vector space (or, more generally, in a matroid of rank n), the conjecture states that the elements in each of the bases $B_1, \ldots, B_n$ can be ordered in such a way that the first elements of $B_1, \ldots, B_n$ together form a basis, the second elements of $B_1, \ldots, B_n$ together form a basis, and so on. In other words, the (multi-)set, $B_1 \cup \cdots \cup B_n$ can be decomposed into n disjoint bases each consisting of exactly one element from each of the original bases $B_1, \ldots, B_n$. Saying that such a basis containing exactly one element from each of $B_1, \ldots, B_n$ is a *transversal basis*, Rota's basis conjecture can be restated as follows.

Conjecture 1.1 (Rota's Basis Conjecture). *Any collection of n disjoint bases in a rank-n matroid can be decomposed into n transversal bases.*

Aside from its intriguingly simple statement, an appealing aspect of this conjecture is that it has many unexpected connections. In particular, Huang and Rota [13] discovered a surprising link to the Alon–Tarsi conjecture [3] on the number of odd and even Latin squares of order n (as discussed further below). Additionally, Conjecture 1.1 for graphic matroids can be phrased as a question about rainbow spanning trees in a multi-graph (namely, given a multi-graph on $n + 1$ vertices with n spanning trees in n different colours, does there always exist a decomposition of the coloured edge set into rainbow spanning trees?). Here, a *rainbow* subgraph is one with at most one edge of each colour. Studying the existence of rainbow subgraphs in graphs with various properties, and applications thereof, is more generally a

very active research direction, as can be seen in the recent surveys (on related but different topics) by Pokrovskiy [17], Sudakov [21], and the current first author [14].

Despite a great deal of attention, including as the subject of the 12th Polymath project [19] in 2017, Conjecture 1.1 remains widely open, even for the case of vector spaces (i.e., representable matroids). The conjecture has only been confirmed in some very special cases, including for 'paving matroids' by Geelen and Humphries [10], for 'strongly base orderable matroids' by Wild [22], and for matroids of rank at most 4 by Cheung [6]. The link to the Alon–Tarsi conjecture, discovered by Huang and Rota [13] (and later simplified by Onn [15]), reduces Rota's basis conjecture for real-representable matroids of rank n, for even n, to the Alon–Tarsi conjecture for Latin squares of order n. The Alon–Tarsi conjecture was proved when $n-1$ is prime by Drisko [8] and when $n+1$ is prime by Glynn [12], thus together giving a proof of Conjecture 1.1 for real-representable matroids of rank n when $n-1$ or $n+1$ is an odd prime.

In addition to these special cases, there is some further evidence towards Conjecture 1.1. In 2006, Aharoni and Berger [1] proved that the natural fractional relaxation of Conjecture 1.1 holds. More recently, the current second author [20] showed that the conjecture holds with high probability for representable matroids over finite fields if the n bases are chosen independently and uniformly at random.

There are two natural directions for working towards Conjecture 1.1 in full, as was highlighted by Pokrovskiy [18], and which are the focus of this paper. The first of these is the *packing problem* in this setting: given n bases in a rank-n matroid, how many disjoint transversal bases can we find? The second is the *covering problem* in this setting: given n bases in a rank-n matroid, how many transversal bases do we need to cover all of the elements of these bases? As discussed in Section 1.1 below, in this paper we give asymptotically-tight answers to both of these questions, thus solving Problems 4.2 and 4.3 in [18].

That one transversal basis can be found in the setting of Conjecture 1.1 is an easy consequence of the matroid augmentation property, which can be used to greedily select one element from each basis step by step, while maintaining that the selected elements form an independent set. However, already, the existence of two disjoint transversal bases is not immediate. In 2007, Geelen and Webb [11] gave the first packing result towards Conjecture 1.1, by showing that there are always $\Omega(\sqrt{n})$ disjoint transversal bases in this setting. Geelen and Dong [7] later introduced a beautiful probabilistic argument to show that there are always $\Omega(n/\log n)$ disjoint transversal bases. Finally, in 2020, Bucić, Kwan, Pokrovskiy, and Sudakov [5] made an important breakthrough by showing there is always a linear number of disjoint transversal bases, as follows.

Theorem 1.2 ([5]). *For any $\varepsilon > 0$, the following holds whenever n is sufficiently large with respect to ε. Any collection of n bases of a rank-n matroid has at least $(1/2 - \varepsilon)n$ disjoint transversal bases.*

Theorem 1.2 shows that, given bases $B_1, \ldots, B_n$ in a rank-n matroid, there is a family of disjoint transversal bases using almost one half of the elements in $B_1 \cup \cdots \cup B_n$. Pokrovskiy [18] subsequently showed that if instead of disjoint transversal bases one forms a family of n disjoint rainbow independent sets, then almost all of the elements can be covered, as follows. Here, a *rainbow independent set* is an independent set containing at most one element from each of the given bases.

Theorem 1.3 ([18]). *For any $\varepsilon > 0$, the following holds whenever n is sufficiently large with respect to ε. Any collection of n bases of a rank-n matroid has at least $(1 - \varepsilon)n$ disjoint rainbow independent sets of size at least $(1 - \varepsilon)n$ each.*

We remark that, with a simple greedy argument, one can deduce from Theorem 1.3 that in the same setting one can also find n disjoint rainbow independent sets, each of size at least $(1 - \varepsilon)n$ (this was observed by Kwan, see [18, Section 4]). However, there does not seem to be a simple argument in order to increase the size of the rainbow independent sets in Theorem 1.3.

In the second direction, concerning covering results towards Conjecture 1.1, neither Theorem 1.2 nor Theorem 1.3 imply a strong covering result. Indeed, even if we can find $(1 - o(1))n$ disjoint transversal bases in some matroid, it may still be the case that an additional n transversal bases are needed in order to cover the remaining $o(n^2)$ elements (it is easy to construct examples for this).

Using topological tools, Aharoni and Berger [1, Assertion 8.11] showed in 2006 that $2n$ transversal bases are sufficient to cover all of the elements in the bases in the setting of Conjecture 1.1, as follows (in fact, they proved a more general result [1, Theorem 8.9], see also the discussion just before Theorem 6.6).

Theorem 1.4 ([1]). *Any collection of n bases of a rank-n matroid can be covered by $2n$ transversal bases.*

2

The Polymath 12 project [19] later improved this bound slightly, showing that $2n-2$ transversal bases are already sufficient.

1.1   Our results

In this paper, we give asymptotically-tight bounds for both the packing and the covering problem in the setting of Rota's basis conjecture, as follows, thus improving Theorems 1.2 to 1.4 above.

Theorem 1.5 (Asymptotic Packing Theorem)**.** *For any $\varepsilon > 0$, the following holds whenever n is sufficiently large with respect to ε. Any collection of n bases of a rank-n matroid has at least $(1 - \varepsilon)n$ disjoint transversal bases.*

Theorem 1.6 (Asymptotic Covering Theorem)**.** *For any $\varepsilon > 0$, the following holds whenever n is sufficiently large with respect to ε. Any collection of n bases of a rank-n matroid can be covered by at most $(1 + \varepsilon)n$ transversal bases.*

To prove Theorem 1.5 and 1.6, we modify and extend the methods of [5] and [18], while introducing new randomised approaches to extend families of rainbow independent sets to transversal bases, careful notions of 'density' within subsets of the matroid, and intricate techniques to modify and improve families of disjoint rainbow independent sets. Outlines of our proofs can be found in Section 3.

As discussed by Bucić, Kwan, Pokrovskiy, and Sudakov [5], and by Pokrovskiy [18], a plausible path to a proof of Conjecture 1.1 might be to combine an approximate form of the conjecture with the *absorption method*. Conjecture 1.1 seems to be a particularly complex setting for the use of absorption, but, if the absorption method can be applied, then the two asymptotic forms of Conjecture 1.1 shown here will very likely be helpful.

2   Preliminaries

We begin by recalling the definition and some basic properties of matroids (for more details see, for example, the textbook [16]). A matroid $\mathcal{M}$ is given by a finite set M and a collection $\mathcal{I}$ of subsets of M satisfying the following three conditions:

- $\emptyset \in \mathcal{I}$.

- For any set $S \in \mathcal{I}$, and any subset $S' \subseteq S$, we also have $S' \in \mathcal{I}$.

- For any sets $S, T \in \mathcal{I}$ with $|S| < |T|$, there exists an element $t \in T \setminus S$ with $S \cup \{t\} \in \mathcal{I}$.

The last condition above is often referred to as the *augmentation property*. The set M is called the *ground set* of the matroid and the members of $\mathcal{I}$ are called the *independent sets* in the matroid $\mathcal{M}$. Common examples of matroids are given by considering linearly independent sets in vector spaces, or acyclic subsets of the edges of a graph. Given a matroid $\mathcal{M}$ with ground set M, the *rank* of any subset $S \subseteq M$ is defined by

$$\mathrm{rk}(S) := \max\{|S'| : S' \text{ independent}, S' \subseteq S\}.$$

In other words, the rank of S is the size of the largest independent subset of S. Clearly, we have $\mathrm{rk}(S) \leq |S|$ (with equality if and only if S itself is independent), and furthermore $\mathrm{rk}(S) \leq \mathrm{rk}(T)$ whenever $S \subseteq T$.

The rank of the matroid is the rank of its ground set, i.e., the maximum size of an independent set in the matroid. A maximum size independent set is called a *basis* of the matroid.

Given a matroid $\mathcal{M}$ with ground set M, the *span* of a subset $S \subseteq M$ is defined as

$$\mathrm{span}(S) = \{x \in M : \mathrm{rk}(S \cup \{x\}) = \mathrm{rk}(S)\}.$$

In other words, $\mathrm{span}(S)$ is the set consisting of all elements whose addition to S does not increase the rank. Note that for any basis B of $\mathcal{M}$ we have $\mathrm{span}(B) = M$. Some authors use the name *closure* instead of span. Also note that this concept of span agrees with the linear-algebraic span in matroids obtained from vector spaces (by taking the vector space as the ground set and linearly independent sets as the independent sets of the matroid).

It is not hard to see that $S \subseteq \mathrm{span}(S)$ and $\mathrm{rk}(\mathrm{span}(S)) = \mathrm{rk}(S)$ for every subset $S \subseteq M$ (see [16, Lemma 1.4.2]). One can also show that $\mathrm{span}(\mathrm{span}(S)) = \mathrm{span}(S)$ and $\mathrm{span}(S) \subseteq \mathrm{span}(T)$ for any subsets $S \subseteq T \subseteq M$ (see [16, Lemma 1.4.3]). Finally, for any subsets $S \subseteq T \subseteq M$ with $\mathrm{rk}(S) = \mathrm{rk}(T)$, we must have $\mathrm{span}(S) = \mathrm{span}(T)$. Indeed, if there was an element $x \in \mathrm{span}(T) \setminus \mathrm{span}(S)$, then we would have $\mathrm{rk}(T \cup \{x\}) \geq \mathrm{rk}(S \cup \{x\}) > \mathrm{rk}(S) = \mathrm{rk}(T)$, a contradiction.

In this paper, we are concerned with rainbow independent sets in coloured matroids. A *coloured matroid* is simply a matroid in which every element is assigned a colour. A set is called a *rainbow set* if all of its elements have distinct colours. Thus, a rainbow independent set is a set which is independent and consists of elements of distinct colours (this may be interpreted as a set in the intersection of the underlying matroid with the partition matroid given by the colouring, but this perspective is not important in our paper). Furthermore, a rainbow basis is a basis of the matroid whose elements have distinct colours.

Using the augmentation property above repeatedly, one can show that, for any independent sets S and T in a matroid, one can find an independent set S^* with $S \subseteq S^* \subseteq S \cup T$ and $|S^*| \geq |T|$. This last condition can equivalently be stated as $|(S \cup T) \setminus S^*| \leq |(S \cup T) \setminus T|$, or again equivalently as $|T \setminus S^*| \leq |S \setminus T|$. The following lemma proves a similar statement for rainbow independent sets in coloured matroids.

Lemma 2.1. *Let S and T be rainbow independent sets in a coloured matroid. Then there exists a rainbow independent set S^* with $S \subseteq S^* \subseteq S \cup T$ and $|T \setminus S^*| \leq 2 \cdot |S \setminus T|$.*

We give the simple proof of this lemma in Section 8, in which we collect the proofs of various statements about matroids used in this paper.

In the settings of Theorems 1.5 and 1.6, we are given a matroid $\mathcal{M}$ of rank n and bases $B_1, \ldots, B_n$. We may assume that the bases $B_1, \ldots, B_n$ are disjoint, by "duplicating" elements appearing in more than one of these bases. More formally, if an element x appears in the bases $B_{i(1)}, \ldots, B_{i(k)}$ (with $k \geq 2$), we define a matroid $\mathcal{M}'$ where we replace x by k different elements $x_1, \ldots, x_k$ in the ground set, and we replace x by x_j in each of the bases $B_{i(j)}$ (for $j = 1, \ldots, k$). In this new matroid $\mathcal{M}'$ a set is independent if it contains at most one of the elements $x_1, \ldots, x_k$ and it corresponds to an independent set in $\mathcal{M}$ when identifying $x_1, \ldots, x_k$ with x. Repeating this operation for all elements appearing in more than one of the bases $B_1, \ldots, B_n$, we can ensure that the bases $B_1, \ldots, B_n$ are disjoint.

Thus, in the settings of Theorems 1.5 and 1.6, we may consider a matroid of rank n with disjoint bases $B_1, \ldots, B_n$. We may furthermore restrict the ground set of the matroid to be $B_1 \cup \cdots \cup B_n$ (note that all elements outside $B_1 \cup \cdots \cup B_n$ are irrelevant to the conclusions of Theorems 1.5 and 1.6). Now, having a matroid with ground set $B_1 \cup \cdots \cup B_n$ and disjoint bases $B_1, \ldots, B_n$, we can naturally define a colouring: We simply say that an element $x \in B_1 \cup \cdots \cup B_n$ has colour i if $x \in B_i$. We denote the colour of any element $x \in B_1 \cup \cdots \cup B_n$ by $c(x)$ (note that then we have $x \in B_{c(x)}$). A transversal basis is now simply a rainbow basis.

We close this section by introducing some notational conventions, and stating two more matroid theory lemmas which we will use throughout this paper. For a family $\mathcal{T}$ of disjoint rainbow independent sets in a coloured matroid, we write $E(\mathcal{T}) = \bigcup_{T \in \mathcal{T}} T$ for the set of matroid elements covered by the sets in $\mathcal{T}$. For a subset S of a matroid and an element $x \in S$, we write $S - x$ for the set $S \setminus \{x\}$ obtained when removing x from S (we will not use the notation $S - x$ for $S \setminus \{x\} = S$ if $x \notin S$). Similarly, for an element y of the matroid, we write $S + y$ for the set obtained from S when adding y. In the case, where y is already contained in S, we interpret $S + y$ as a multi-set. Note that, in this case, the set $S + y$ cannot be independent (since it contains the element y twice).

The following statement is a slight strengthening of Lemma 2.7 in [5] and, as noted there, also follows from a result of Brualdi [4]. We include a proof in Section 8.

Lemma 2.2. *Let S be an independent set in some matroid and let B be a basis of the matroid. Then, there is an injection $\phi : S \to B$ such that, for each $x \in S$, the set $S - x + \phi(x)$ is independent, and, for each $b \in B \setminus \phi(S)$, the set $S + b$ is independent.*

Recall that here $S - x + \phi(x)$ and $S + b$ are to be interpreted as multi-sets (namely, obtained from $S - x = S \setminus \{x\}$ when adding $\phi(x)$, and from S when adding b, respectively). That these sets are independent automatically means that they cannot contain an element twice, meaning that $\phi(x) \notin S - x$ and $b \notin S$.

Our final lemma in this section is the following, and is also proved in Section 8.

Lemma 2.3. *For any independent sets S and T in a matroid, at least one of the following statements holds.*

(a) *There is some $x \in S$ such that the set $T + x$ is independent.*

(b) *There is an injection $\phi : S \to T$ such that, for each $x \in S$, the set $T - \phi(x) + x$ is independent and $\mathrm{span}(T - \phi(x) + x) = \mathrm{span}(T)$.*

Again, in (a), the set $T + x$ is to be interpreted as a multi-set. So if (a) holds, this automatically means that $T + x$ cannot contain an element twice, and so we have $x \notin T$. Similarly, if (b) holds, this automatically means that $x \notin T - \phi(x)$ for every $x \in S$.

Our notation follows general convention. In particular, for a positive integer n, we write $[n] = \{1, \ldots, n\}$ as usual. For a bipartite graph with vertex classes X and Y, by slight abuse of notation, we interpret its edges to be pairs of the form (x, y) with $x \in X$ and $y \in Y$, meaning that its edge set can be interpreted as a subset of $X \times Y$.

3 Proof outlines

In this section, we give outlines of the proofs of our main results in Theorems 1.5 and 1.6, and state several key lemmas used in these proofs. The proof of Theorem 1.6 can be found in Section 6, using the results in Section 4. The proof of Theorem 1.5 can be found in Section 7, using the results in Section 5 (which, in turn, build on some of the results in Section 4).

3.1 Covering theorem

As discussed in the previous section, in the setting of our covering theorem (Theorem 1.6), we can consider a coloured matroid of rank n, where the colour classes are bases $B_1, \ldots, B_n$. We then need to show that their union $B_1 \cup \cdots \cup B_n$ can be covered by $(1 + \varepsilon)n$ rainbow bases. This is equivalent to showing that $B_1 \cup \cdots \cup B_n$ can be decomposed into $(1 + \varepsilon)n$ rainbow independent sets (indeed, given such a decomposition into rainbow independent sets, it is not hard to extend these rainbow independent sets to rainbow bases greedily using the augmentation property of the matroid).

Therefore, our goal is to find a family of $(1 + \varepsilon)n$ disjoint rainbow independent sets whose union is the entire ground set $B_1 \cup \cdots \cup B_n$. Instead of finding such a family directly, for some $\lambda < \varepsilon$ (which we will later choose as $\lambda = \varepsilon/3$), we aim to find a family $\mathcal{T}$ of $(1 + \lambda)n$ disjoint rainbow independent sets such that the set of leftover elements not covered by $\mathcal{T}$ can also be decomposed into few (namely, at most $(\varepsilon - \lambda)n$) rainbow independent sets. As discussed in the previous section, we denote the set of elements covered by $\mathcal{T}$ as $E(\mathcal{T}) = \bigcup_{T \in \mathcal{T}} T$. Now, letting $U = (B_1 \cup \cdots \cup B_n) \setminus E(\mathcal{T})$ be the set of uncovered elements, we want to decompose U into few rainbow independent sets. In order to have any hope for such a decomposition, we need U to be sufficiently small (namely, we need $|U| \leq (\varepsilon - \lambda)n^2$ if U is to be decomposed into at most $(\varepsilon - \lambda)n$ rainbow independent sets).

It follows from Pokrovskiy's result [18] stated as Theorem 1.3 that it is possible to find a family $\mathcal{T}$ of $\lfloor (1 + \lambda)n \rfloor$ disjoint rainbow independent sets such that the set $U = (B_1 \cup \cdots \cup B_n) \setminus E(\mathcal{T})$ of uncovered elements is rather small. However, instead of Theorem 1.3, we need the following stronger statement, in which the set U of uncovered elements is small for *every* family $\mathcal{T}$ of $\lfloor (1+\lambda)n \rfloor$ disjoint rainbow independent sets such that $E(\mathcal{T})$ is *inclusion-wise maximal*. Here, $E(\mathcal{T})$ being inclusion-wise maximal means that there does not exist a family $\mathcal{S}$ of $\lfloor (1 + \lambda)n \rfloor$ disjoint rainbow independent sets with $E(\mathcal{T}) \subsetneq E(\mathcal{S})$.

Lemma 3.1. *For any $\lambda, \nu > 0$ the following holds for any sufficiently large n (sufficiently large with respect to λ and ν) and any coloured rank-n matroid with colour classes $B_1, \ldots, B_n$, such that each of the sets B_i for $i = 1, \ldots, n$ is a basis. Let $\mathcal{T}$ be a family of $\lfloor (1 + \lambda)n \rfloor$ disjoint rainbow independent sets such that $E(\mathcal{T})$ is inclusion-wise maximal. Then we have $|E(\mathcal{T})| \geq (1 - \nu)n^2$.*

This lemma is proved in Section 4, and is one of the key ingredients for the proof of Theorem 1.6. It also directly implies Theorem 1.3, as we will show at the end of Section 4 (so Section 4 gives a self-contained proof of Theorem 1.3 which is somewhat simpler than the original proof in [18]). The proof of Lemma 3.1 uses a similar approach as in [18], but with some crucial differences, enabling us to obtain the stronger

statement in Lemma 3.1, which we need for the proof of Theorem 1.6. Some arguments from the proof of Lemma 3.1 will also later be used in the proof of Theorem 1.5.

The conclusion $|E(\mathcal{T})| \geq (1 - \nu)n^2$ of Lemma 3.1 means that $|U| = n^2 - |E(\mathcal{T})| \leq \nu n^2$. Of course, just knowing that U has size at most νn^2 is not enough to decompose U into few rainbow independent sets. To ensure the existence of such a decomposition, we will need two clearly necessary conditions: that U can be decomposed into few rainbow sets and that U can be decomposed into few independent sets. It was shown by Aharoni and Berger [1] in 2006 that these two conditions are actually sufficient for a set U to be decomposable into (almost as) few rainbow independent sets (see Theorem 6.6 for a restatement of their result). In particular, their result implies that, if U can be decomposed both into λn rainbow sets and into λn independent sets, then U can be decomposed in at most $2\lambda n$ rainbow independent sets. If we can achieve this, then, together with the at most $(1 + \lambda)n$ rainbow independent sets in the family $\mathcal{T}$, this would give a decomposition of $B_1 \cup \cdots \cup B_n$ into at most $(1 + 3\lambda)n$ rainbow independent sets (so, taking $\lambda = \varepsilon/3$, this gives the desired conclusion).

The first condition, demanding that U is decomposable into λn rainbow sets, is equivalent to requiring that every colour appears in U at most λn times. On its own, ensuring this would not require significantly more effort than proving Lemma 3.1. Much more difficult will be the second condition about decomposing U into independent sets. A classical theorem of Edmonds [9] implies that U can be decomposed into at most λn independent sets if and only if $|U'| \leq \lfloor \lambda n \rfloor \cdot \mathrm{rk}(U')$ for all subsets $U' \subseteq U$. This is clearly necessary, as every independent subset of U can contain at most $\mathrm{rk}(U')$ elements from U'.

Effectively, then, we need to make sure that U does not have any "dense spots" U' with $|U'| > \lfloor \lambda n \rfloor \cdot \mathrm{rk}(U')$. However, it turns out that keeping track of all subsets $U' \subseteq U$ with $|U'| > \lfloor \lambda n \rfloor \cdot \mathrm{rk}(U')$ is very difficult. In order to overcome this, we work with a different notion of "dense spots". We call these dense spots *deadlocks*, and introduce this notion in the first part of Section 6. Crucially, the $\lfloor \lambda n \rfloor$-deadlock of U has a non-empty intersection with every subset $U' \subseteq U$ such that $|U'| > \lfloor \lambda n \rfloor \cdot \mathrm{rk}(U')$ (this intersection can be viewed as the "dense part" of U' that causes U' to be dense overall). Thus, if the $\lfloor \lambda n \rfloor$-deadlock of U is empty, then there cannot be any subsets $U' \subseteq U$ with $|U'| > \lfloor \lambda n \rfloor \cdot \mathrm{rk}(U')$, and we will therefore be able to decompose U into $\lfloor \lambda n \rfloor$ independent sets. Our goal is thus to find a family $\mathcal{T}$ of $\lfloor (1 + \lambda)n \rfloor$ disjoint rainbow independent sets such that $U = (B_1 \cup \cdots \cup B_n) \setminus E(\mathcal{T})$ contains every colour at most λn times and such that the $\lfloor \lambda n \rfloor$-deadlock of U is empty.

Using Lemma 3.1, we can find a family $\mathcal{T}$ of $\lfloor (1 + \lambda)n \rfloor$ disjoint rainbow independent sets such that $U = (B_1 \cup \cdots \cup B_n) \setminus E(\mathcal{T})$ satisfies $|U| \leq \nu n^2$. Later, we will denote the $\lfloor \lambda n \rfloor$-deadlock of U by $D_{\lfloor \lambda n \rfloor}(U)$, but in this sketch we will simply call it $D(U)$. This deadlock will have the property that removing any element e in $D(U)$ from U will reduce the size of $D(U)$, a good thing as we wish to reduce $D(U)$ to the empty set. But in order to remove e from U, we need to add it to some $T \in \mathcal{T}$, and there will likely[1] be no such T for which $T + e$ is a rainbow independent set. So, we will need to pick some $T \in \mathcal{T}$ and, when adding e to T, in order to still have a rainbow independent set, we need to remove one or two other elements from T (perhaps one element to maintain rainbowness and perhaps one element to maintain independence). In other words, we replace $T \in \mathcal{T}$ by $(T + e) \setminus F$, where F is a set of at most two elements of T (see part a) of Figure 1). But then the elements in F get added to the uncovered set U, and, while removing e from U decreases the size of the deadlock $D(U)$, adding the elements in F to U may increase the size of $D(U)$ again! Therefore, we will have to choose the element $e \in U$, the set $T \in \mathcal{T}$, and the elements in F very carefully.

Utilising the meticulous definition for our notion of deadlock, it turns out that the right condition for this is to ensure that F is disjoint from the span of the deadlock $D'(U)$ defined by a slightly different parameter (namely, to be precise, the $(\lfloor \lambda n \rfloor - 2)$-deadlock of U). This parameter choice will mean that we have $D(U) \subseteq D'(U)$. However, in order to be able to choose e, T and F such that F is disjoint from the span of the deadlock $D'(U)$, we also want $D'(U)$ not to be too much larger than $D(U)$ (more, precisely, we want the rank of $D'(U)$ not to be too much larger than the rank of $D(U)$).

To ensure this, instead of two deadlocks, we will consider a sequence of deadlocks with gradually decreasing parameters. These deadlocks will be nested, and increasing in size, and thus we will be able to choose a consecutive pair of deadlocks in this sequence where the rank does not grow too much. We can then decrease the size of the smaller of these two deadlocks via a switching operation as described above, replacing some $T \in \mathcal{T}$ by $(T + e) \setminus F$ for some element $e \in U$ (and some set $F \subseteq T$ of size $|F| \leq 2$).

---

[1] If we choose $\mathcal{T}$ as in Lemma 3.1 such that $E(\mathcal{T})$ is inclusion-wise maximal, then definitely there will be no such T.
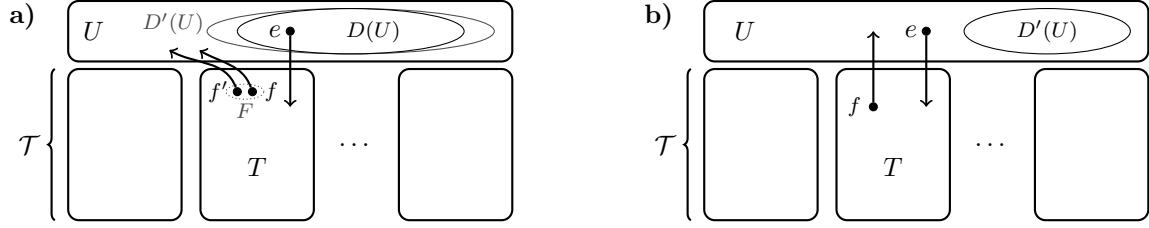
Figure 1: **a)** Given a family $\mathcal{T}$, and its set of uncovered elements U, in order to make the deadlock $D(U)$ smaller, we look for some $e \in D(U)$, a set $T \in \mathcal{T}$, and a subset $F \subseteq T$ of size $|F| \leq 2$ so that $(T + e) \setminus F$ is a rainbow independent set. We then replace $T \in \mathcal{T}$ by $(T + e) \setminus F$. To ensure the size of the deadlock $D(U)$ decreases, we make sure that F is disjoint from the span of the slightly larger deadlock $D'(U)$.

b) Where a colour c appears too often in U, we look for an element $e \in U$ with colour c, a set $T \in \mathcal{T}$ and $f \in T$ so that c does not appear on T and $T + e - f$ is an independent set. We then replace $T \in \mathcal{T}$ by $T + e - f$. To maintain the property that $D(U) = \emptyset$, we ensure that $f \notin \text{span}(D'(U))$. We also make sure that $c(f)$ does not appear that often on U.

Repeating this, we will ultimately be able to decrease the size of the $\lfloor \lambda n \rfloor$-deadlock $D(U)$, and with even more repetitions we obtain $D(U) = \emptyset$.

For clarity, and for a better organisation of the proof, we do not write the proof via a sequence of many such switching steps. Instead, we choose the initial family $\mathcal{T}$ of rainbow independent sets via a certain minimisation, lexicographically minimising the sizes of the relevant sequence of deadlocks. This minimisation will also be chosen in such a way that $E(\mathcal{T})$ is automatically inclusion-wise maximal. Hence, Lemma 3.1 implies that the set U of uncovered elements has size $|U| \leq \nu n^2$. This is crucial for our arguments in order to find suitable elements $e \in U$, $T \in \mathcal{T}$ and $F \subseteq T$ for our desired switching operation replacing T by $(T + e) \setminus F$. Note that this operation can actually increase the size of U (namely, it replaces U by $(U - e) \cup F$ for some set F of size $|F| \leq 2$), so during a long sequence of such switching steps the size of U could a priori grow by a lot. This is why it is very important for us that the conclusion of Lemma 3.1 holds for *every* family $\mathcal{T}$ such that $E(\mathcal{T})$ is inclusion-wise maximal. It would not suffice to just find *some* initial family $\mathcal{T}$ such that $|U| \leq \nu n^2$.

Having found a family $\mathcal{T}$ of $\lfloor (1 + \lambda) n \rfloor$ disjoint rainbow independent sets with $D(U) = \emptyset$ for the set U of uncovered elements (and having completed a lot of background work to achieve this), the remaining task is to modify the family $\mathcal{T}$ such that U does not contain too many elements of any one colour. Suppose we have a colour c that appears at least λn times on U. Picking a set $T \in \mathcal{T}$ without an element of colour c, we would like to find an element $e \in U$ of colour c and an element $f \in T$ such that $T + e - f$ is an independent set (it will automatically be rainbow by the choice of T). Our goal is then to replace T by $T + e - f$ (see part b) of Figure 1), because then the set U of uncovered elements gets replaced by $U - e + f$, which reduces the number of times that colour c appears in U. However, in order to maintain the condition $D(U) = \emptyset$, for some deadlock $D'(U)$ with a slightly different parameter, we need to ensure that $f \notin \text{span}(D'(U))$ (similarly to the discussion above). We also want to make sure that the colour of f does not appear too often in U (because the number of appearances of that colour increases when replacing U by $U - e + f$). Since the rank of $D'(U)$ will be fairly small, and only few colours appear very often on U (both of these properties are due to the fact that again $|U|$ will be small by Lemma 3.1), and the elements in U of colour c have rank at least λn, it will be possible to choose elements e and f with the desired properties. Then, upon replacing T by $T + e - f$ we can reduce the number of times that colour c appears in U (while maintaining $D(U) = \emptyset$ and not increasing the number of appearances of any other colour beyond λn).

Again, for clarity and a better proof organisation, we actually choose the family $\mathcal{T}$ via a certain minimisation (this time minimising $\sum_{c=1}^{n} |U \cap B_c|^2$), subject to the constraint that $D(U) = \emptyset$. We then show that this family $\mathcal{T}$ must have the property that every colour appears at most λn times in U, because otherwise we could improve $\mathcal{T}$ when replacing T by $T + e - f$ as above.

All in all, this yields the desired family $\mathcal{T}$ of $\lfloor (1 + \lambda) n \rfloor$ disjoint rainbow independent sets, such for the set U of uncovered elements the $\lfloor \lambda n \rfloor$-deadlock is empty, and such that U contains every colour at most

λn times. As discussed above, this is sufficient to prove Theorem 1.6.

The detailed proof of Theorem 1.6, as well as our definition of deadlocks and some lemmas about them, can be found in Section 6.

3.2   Packing theorem

As discussed in the previous section, in the setting of our packing theorem (Theorem 1.5), we can again consider a coloured matroid of rank n, where the colour classes are bases $B_1, \dots, B_n$. This time, we need to show that we can we can find at least $(1 - \varepsilon)n$ disjoint rainbow bases in $B_1 \cup \dots \cup B_n$. Our basic strategy to prove Theorem 1.5 is first to reserve a random subset $R \subseteq B_1 \cup \dots \cup B_n$ (the 'reservoir'), then to form a family $\mathcal{T}$ of $\lceil (1 - \varepsilon)n \rceil$ disjoint rainbow independent sets of large total size in the complement of R, and finally to turn $\mathcal{T}$ into a family of disjoint rainbow bases with a procedure using the elements in R.

This strategy relies on two key lemmas, stated below. The first of these lemmas is somewhat similar to Lemma 3.1 and Pokrovskiy's result [18] stated as Theorem 1.3. It gives a family $\mathcal{T}$ of $\lceil (1 - \varepsilon)n \rceil$ disjoint rainbow independent sets of large total size, but with the very important constraint that $\mathcal{T}$ avoids our reservoir $R \subseteq B_1 \cup \dots \cup B_n$ of randomly chosen elements.

Lemma 3.2. *For any $0 < \varepsilon < \eta < 1$ and $\nu > 0$ the following holds for any sufficiently large n (sufficiently large with respect to ε, η and ν) and any coloured rank-n matroid with colour classes $B_1, \dots, B_n$, such that each of the sets B_i for $i = 1, \dots, n$ is a basis. Let $R \subseteq B_1 \cup \dots \cup B_n$ be a set of elements drawn independently at random with probability η. Then, with high probability (more precisely, with probability tending to 1 as $n \to \infty$ with ε, η and ν fixed), there is a family $\mathcal{T}$ of $\lfloor (1-\varepsilon)n \rfloor$ disjoint rainbow independent sets in $(B_1 \cup \dots \cup B_n) \setminus R$ such that $|E(\mathcal{T})| \geq (1 - \eta - \nu)n^2$.*

The proof of Lemma 3.2 can be found in Section 5. It is similar to the proof of Lemma 3.1 in Section 4 and uses some of the results from that section. However, in the setting of Lemma 3.2, we need to be more careful as we must avoid using elements from the set R. Moreover, it is crucial that in Lemma 3.2 we find a family consisting of only $\lfloor (1 - \varepsilon)n \rfloor$ rainbow independent sets, as opposed to the $\lfloor (1 + \lambda)n \rfloor$ rainbow independent sets in Lemma 3.1. Indeed, the fact that the family consist of fewer sets means that the average size of these sets is larger, and so fewer alterations are needed in order to turn these sets into rainbow bases.

Our second key lemma for the proof of Theorem 1.5, stated next, is used in order to turn the family $\mathcal{T}$ obtained from the previous lemma into a family of disjoint rainbow bases. Roughly speaking, this lemma states that for any family $\mathcal{S}$ of $\lceil (1-\varepsilon)n \rceil$ disjoint rainbow independent sets of total size $|E(\mathcal{S})| < \lceil (1-\varepsilon)n \rceil \cdot n$ which does not contain too many elements from R, we can find a family $\mathcal{S}'$ of $\lceil (1 - \varepsilon)n \rceil$ disjoint rainbow independent sets of strictly larger total size which contains not too many more elements from R. Applying this lemma repeatedly, we can keep increasing the total size of the family until we arrive at such a family of total size $\lceil (1 - \varepsilon)n \rceil \cdot n$, i.e., a family of $\lceil (1 - \varepsilon)n \rceil$ disjoint rainbow bases. This proves Theorem 1.5 (the details of the deduction of Theorem 1.5 from the two key lemmas stated here can be found at the start of Section 7).

Lemma 3.3. *For any $0 < \varepsilon < 1/10$, there exist $\sigma = \sigma(\varepsilon) > 0$ and $L = L(\varepsilon) > 0$, such that for any η with $\varepsilon \leq \eta < 1$, any sufficiently large n (sufficiently large with respect to ε and η), and any coloured rank-n matroid with colour classes $B_1, \dots, B_n$ with the property that each of the sets B_i for $i = 1, \dots, n$ is a basis, the following holds with high probability (more precisely, with probability tending to 1 as $n \to \infty$ with ε and η fixed) for a random subset $R \subseteq B_1 \cup \dots \cup B_n$ obtained by drawing elements independently at random with probability η. For any family $\mathcal{S}$ of $\lceil (1-\varepsilon)n \rceil$ rainbow independent sets with $\frac{3}{4}n^2 \leq |E(\mathcal{S})| < \lceil (1-\varepsilon)n \rceil \cdot n$ and $|E(\mathcal{S}) \cap R| \leq \sigma n^2$, there is a family $\mathcal{S}'$ of $\lceil (1-\varepsilon)n \rceil$ rainbow independent sets such that $|E(\mathcal{S}')| > |E(\mathcal{S})|$ and*

$$|E(\mathcal{S}') \cap R| \leq |E(\mathcal{S}) \cap R| + L \cdot \ln\left(\frac{n^2}{\lceil (1 - \varepsilon)n \rceil \cdot n - |E(\mathcal{S})|}\right). \tag{1}$$

The term $\lceil (1 - \varepsilon)n \rceil \cdot n - |E(\mathcal{S})|$ in the denominator here counts the number of elements missing in $\mathcal{S}$ compared to a family of $\lceil (1 - \varepsilon)n \rceil$ rainbow bases. The proof of Lemma 3.3 can be found in Section 7, but we give an outline of the proof here. To prove the lemma we will use some 'cascade-like' switching operations. In order for this to work, we may first have to alter the family $\mathcal{S}$ to get a family $\mathcal{T}$ of $\lceil (1-\varepsilon)n \rceil$

rainbow independent sets of the same total size $|E(\mathcal{T})| = |E(\mathcal{S})|$, before starting these operations. For the purposes of this proof outline, we can assume that $\mathcal{T} = \mathcal{S}$, and we comment on the more technical alteration to $\mathcal{S}$ later, just before the proof of Lemma 3.3 in Section 7.

Now, we have a family $\mathcal{T}$ of $\lceil (1-\varepsilon)n \rceil$ disjoint rainbow independent sets, not all of which are bases, such that $E(\mathcal{T})$ does not contain too many elements from the random reservoir R that we reserved before. As before, we write $U = (B_1 \cup \cdots \cup B_n) \setminus E(\mathcal{T})$ for the set of elements not covered by $\mathcal{T}$. Let us choose a set $T_1 \in \mathcal{T}$ which is not a basis, and a colour c not appearing on T. From the matroid augmentation property applied to T_1 and the basis B_c in colour c, we know that there is an element e of colour c for which $T_1 + e$ is an independent set (and also automatically rainbow). If $e \in U$ then in the family $\mathcal{T}$ we can simply replace T_1 by $T_1 + e$ in order to obtain the desired family $\mathcal{S}'$ in Lemma 3.3. However, we should anticipate that there is some $T_2 \in \mathcal{T} \setminus \{T\}$ with $e \in T_2$. Then we could move e from S to T to increase the size of T, but overall we would not change the total size of $\mathcal{T}$. However, moving e from T_2 to T_1 will mean that T_2 gets replaced by $T_2 - e$, which is no longer a basis. So we can perhaps find some $e' \neq e$ which we can add to $T_2 - e$. We wish to do this iteratively, making some cascade of changes, until the number of elements we can move is so large that one of them must be in U, at which point the corresponding cascade of changes will make an improvement to $\mathcal{T}$. However, it turns out that one cannot hope to succeed simply by moving elements between the different sets in $\mathcal{T}$ until one finds an element in U that can be added to a set in $\mathcal{T}$ without violating rainbowness and independence. Instead, some more complicated switching operations are needed at every step, where in addition to moving elements between different sets $T \in \mathcal{T}$, we also modify the sets $T \in \mathcal{T}$ by exchanging certain elements for other elements in U. The cascade of switches we use is motivated by the proof of Theorem 1.2 by Bucić, Kwan, Pokrovskiy, and Sudakov [5], but crucially different in order to obtain our asymptotically-tight improvement.

To sketch our approach, let us focus on just one step of our switching cascade (as shown in Figure 2). That is, suppose we have found some sets $T_1, \ldots, T_r \in \mathcal{T}$ (for some small r) and some $T \in \mathcal{T} \setminus \{T_1, \ldots, T_r\}$ and $e \in T$ such that e can be moved from T to some modified version of $T_1, \ldots, T_r$, while maintaining $|E(\mathcal{T})|$. Now, $T - e$ is certainly missing some colour, c say. There are at least εn elements $x \in U$ with colour c (recalling that $|\mathcal{T}| = \lceil (1-\varepsilon)n \rceil$, and every set in $\mathcal{T}$ contains at most one element of colour c), and only few of them were used in one of the previous modifications to $T_1, \ldots, T_r$ to allow e to be incorporated. For each of the not-yet-used elements $x \in U$ of colour c, we consider the set $T - e + x$. This set is always rainbow, but might not be independent anymore, so we now choose some $x' \in T - e$ such that $T - e + x - x'$ is independent. Essentially, this means that in the set $T - e$ we exchange the element $x' \in T - e$ against the element $x \in U$. The different choices for x form an independent set, and therefore this will yield many different pairs (x, x') with distinct elements $x' \in T$ (which have many different colours).

For each of these possible pairs (x, x'), we now repeat this procedure for $T - e + x - x'$ instead of $T - e$. Namely, the colour of x' is missing on $T - e + x - x'$, so we can find some elements $q \in U$ with $c(q) = c(x')$ and $q' \in T - e$ such that $T - e + x - x' + q - q'$ is a rainbow independent set (i.e., such that we can exchange q' against q in the rainbow independent set $T - e + x - x'$). However, this time, as there are many different colours appearing among the possible elements x' obtained in the previous step, we also have a lot more possibilities for the element q. In fact, it will suffice to restrain ourselves to choose q from the set $R \cap U$. Using the randomness of R, it is not hard to show that we have a large span of the elements in R whose colours appear on the possible elements x' in the previous step (and in fact, this span is even robustly large when deleting a small number of elements from R). By our assumption that $E(\mathcal{T})$ does not contain many elements of R, most of the elements of R are in U, and so we can show that there are many potential choices for $q \in R \cap U$ and these choices have a large span. This is the crucial part of our argument, and allows us to show that we will have a huge number of possibilities for $q' \in T - e$ (namely, almost all elements of $T - e$ will appear as a possible choice for q' for some appropriately chosen x, x', and q). Then we also have many different options for finding an element e' such that $T - e + x - x' + q - q' + e'$ is a rainbow independent set. Most of these elements e' will be contained in some set in $\mathcal{T} \setminus \{T_1, \ldots, T_r, T\}$. By the pigeonhole principle, we will find some new set $T' \in \mathcal{T} \setminus \{T_1, \ldots, T_r, T\}$ containing a relatively large number of such elements e'. This means that, starting with just one element $e \in T$ that we moved out of T (incorporating e into $T_1, \ldots, T_r$ via some small modifications), we now found a set T' containing several different possible elements e' that can be incorporated into $T_1, \ldots, T_r, T$ via some small modifications. Iterating this, like a cascade, at every step we obtain an increase in the number of possible elements that can be incorporated into the previous sets.
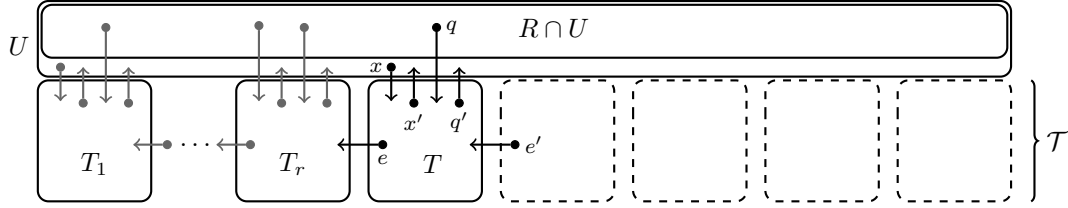
Figure 2: Given an element $e \in T$ which is $(T_1, \ldots, T_r)$-absorbable, we find an element e' which is $(T_1, \ldots, T_r, T)$-absorbable (in fact, we find many such elements). The implicit cascade structure found by our proof is depicted in grey but the critical point is that a small number of alterations to $T_1, \ldots, T_r$ can be made using elements from U to allow e to be incorporated. To incorporate e' into $T - e$, we first swap certain elements $x, q \in U$ against elements $x', q' \in T - e$ so that in the end $T - e + x - x' + q - q' + e'$ is a rainbow independent set.

This number will keep increasing (in fact, it will increase exponentially), until at some point we must find an element in U that can be incorporated into the family $\mathcal{T}$ (after some small modifications), meaning that we found a family $\mathcal{S}'$ improving $\mathcal{T}$, as desired in Lemma 3.3.

To avoid cumbersome notation, we do not wish to remember all of the individual switching steps that are used in this (potentially rather long) cascade. Instead, we will say that an element e is $(T_1, \ldots, T_r)$-*absorbable* if we can make some small alterations to $T_1, \ldots, T_r$ in order to incorporate e, and then we just count the number of $(T_1, \ldots, T_r)$-absorbable elements at every step. Slightly more formally (see Definition 7.4 for the precise definition), e is $(T_1, \ldots, T_r)$-absorbable if there are rainbow independent sets $T_1', \ldots, T_r'$ that can be obtained from $T_1, \ldots, T_r$ via small alterations (where $T_1', \ldots, T_r'$ need to be disjoint from all sets in $\mathcal{T} \setminus \{T_1, \ldots, T_r\}$ apart from potentially containing the element e) such that in total the size $|T_1'| + \cdots + |T_r'|$ increases compared to $|T_1| + \cdots + |T_r|$. While, formally, we do not demand that e is contained in one of the sets $T_1', \ldots, T_r'$, this will automatically follow unless we found an element in U to incorporate into $T_1, \ldots, T_r$ (after some small modifications) to increase their size (in which case we have already found the desired family $\mathcal{S}'$ improving $\mathcal{T}$). At every step, it is essential that we only make 'small' alterations, because we need to iterate this argument, and because we need to bound the number of additional elements in R used by the family $\mathcal{S}'$ compared to the initial families $\mathcal{T}$ and $\mathcal{S}$ in order to satisfy (1). For this reasons, it is also important to bound the total number of steps in the above cascade. In fact, the further away the initial families $\mathcal{S}$ and $\mathcal{T}$ are from a family of $\lceil(1-\varepsilon)n\rceil$ disjoint rainbow bases, the fewer steps we need in the cascade (and therefore fewer additional elements from R are needed). In the proof sketch above, this is because we will be able to start the cascade with some set $T_1 \in \mathcal{T}$ which is further away from a basis, so that already at the next step there will be more possibilities for elements to incorporate into T_1 (after the usual small modifications) and thus will need to iterate our argument less.

4   Inclusion-wise maximal families are large: proof of Lemma 3.1

Given a family $\mathcal{T}$ of disjoint rainbow independent sets in a matroid coloured with colours $1, \ldots, n$, the *colour-availability graph* $\mathcal{A}(\mathcal{T})$ is the bipartite graph with vertex classes $\mathcal{T}$ and $[n]$ and edges (T, c) with $T \in \mathcal{T}$ and $c \in [n]$ exactly when c is a colour that does not appear on T. In other words, the colour-availability graph is the auxiliary graph recording which colours are missing, i.e., still available, on which of the rainbow independent sets in $\mathcal{T}$. Note that for any $T \in \mathcal{T}$ we have $\deg_{\mathcal{A}(\mathcal{T})}(T) = n - |T|$ and for any colour $c \in [n]$, if B_c is the set of elements of colour c, then we have $\deg_{\mathcal{A}(\mathcal{T})}(c) = |\mathcal{T}| - |E(\mathcal{T}) \cap B_c| \geq |\mathcal{T}| - |B_c|$.

For our arguments, we need to consider edges (T, c) in the colour-availability graph $\mathcal{A}(\mathcal{T})$ of a family $\mathcal{T}$ of disjoint rainbow independent sets such that $\deg_{\mathcal{A}(\mathcal{T})}(c)$ can be bounded in terms of $\deg_{\mathcal{A}(\mathcal{T})}(T)$. The following graph-theoretic lemma enables us to find many such edges.

Lemma 4.1. *Let $0 < \alpha < \beta$ and $\delta > 0$. Let G be a bipartite graph with vertex classes X and Y and edge set $E(G) \subseteq X \times Y$. Assume that $|X| \leq \alpha \cdot |Y|$ and that every vertex $y \in Y$ has degree $\deg_G(y) \geq \delta \cdot |X|$. Then, there are at least $(\delta(\beta - \alpha)/\beta) \cdot |X| \cdot |Y|$ edges $(x, y) \in E(G)$ with $\deg_G(y) \leq \beta \cdot \deg_G(x)$.*

Proof. Assign each edge $(x, y) \in E(G)$ with $x \in X$ and $y \in Y$ the weight $w(x, y) = 1/\deg_G(y)$, and call the edge (x, y) *good* if $\deg_G(y) \leq \beta \cdot \deg_G(x)$, and *bad* otherwise. We count the weight on the edges from each side of the partition. Firstly,

$$\sum_{(x,y)\in E(G)} w(x,y) = \sum_{y\in Y} \sum_{x\in N_G(y)} \frac{1}{\deg_G(y)} = \sum_{y\in Y} 1 = |Y|. \tag{2}$$

Secondly, letting $e_{\mathrm{gd}}(G)$ be the number of good edges in G, we have

$$\sum_{(x,y)\in E(G)} w(x,y) = \sum_{\substack{(x,y)\in E(G) \\ xy \text{ good}}} \frac{1}{\deg_G(y)} + \sum_{\substack{(x,y)\in E(G) \\ xy \text{ bad}}} \frac{1}{\deg_G(y)}$$

$$\leq \sum_{\substack{(x,y)\in E(G) \\ xy \text{ good}}} \frac{1}{\delta|X|} + \sum_{\substack{(x,y)\in E(G) \\ xy \text{ bad}}} \frac{1}{\beta \cdot \deg_G(x)}$$

$$\leq \frac{e_{\mathrm{gd}}(G)}{\delta|X|} + \sum_{x\in X} \sum_{y\in N_G(x)} \frac{1}{\beta \cdot \deg_G(x)} = \frac{e_{\mathrm{gd}}(G)}{\delta|X|} + \frac{|X|}{\beta} \leq \frac{e_{\mathrm{gd}}(G)}{\delta|X|} + \frac{\alpha|Y|}{\beta}.$$

Combining this with (2), and rearranging, we indeed obtain

$$e_{\mathrm{gd}}(G) \geq \delta|X| \cdot \left(1 - \frac{\alpha}{\beta}\right) \cdot |Y| = \frac{\delta(\beta - \alpha)}{\beta} \cdot |X| \cdot |Y|. \qquad \square$$

Instead of using Lemma 4.1 directly, we will use the following corollary. The set E in this corollary will simply be obtained by taking the set of edges satisfying the property in Lemma 4.1.

Corollary 4.2. *For any $\alpha, \mu, \delta, \sigma, \lambda > 0$ with $\sigma + \lambda \geq \alpha > \sigma$, there is some $\rho > 0$ such that the following holds. Let G be a bipartite graph with vertex classes X and Y and edge set $E(G) \subseteq X \times Y$, such that $|X| \leq \alpha \cdot |Y|$. Suppose that $\deg_G(y) \geq \delta \cdot |X|$ for every vertex $y \in Y$ and $\deg_G(x) \leq (1 - \mu) \cdot |Y|$ for every vertex $x \in X$. Then, there is a set $E \subseteq E(G)$ such that*

$$\sigma \cdot \sum_{(x,y)\in E} \deg_G(x) \geq \rho \cdot |X| \cdot |Y|^2 + \sum_{(x,y)\in E} \Big(\deg_G(y) - \lambda \cdot |Y|\Big).$$

Proof. Note that $(\alpha - \sigma)(1 - \mu) < \alpha - \sigma \leq \lambda$, so we can choose $\beta > \alpha$ such that $(\beta - \sigma)(1 - \mu) < \lambda$. Now, let us define $\rho = (\delta(\beta - \alpha)/\beta) \cdot (\lambda - (\beta - \sigma)(1 - \mu)) > 0$. By Lemma 4.1, we can find a set $E \subseteq E(G)$ of size $|E| \geq (\delta(\beta - \alpha)/\beta) \cdot |X| \cdot |Y|$ such that $\deg_G(y) \leq \beta \cdot \deg_G(x)$ for each $(x, y) \in E$. Now, noting that $\beta > \alpha > \sigma$, and thus $(\beta - \sigma)\deg_G(x) \leq (\beta - \sigma) \cdot (1 - \mu)|Y|$ for each $x \in X$, we have

$$\sigma \cdot \sum_{(x,y)\in E} \deg_G(x) - \sum_{(x,y)\in E} \Big(\deg_G(y) - \lambda \cdot |Y|\Big) = \sum_{(x,y)\in E} \Big(\sigma \cdot \deg_G(x) - \deg_G(y) + \lambda \cdot |Y|\Big)$$

$$\geq \sum_{(x,y)\in E} \Big(\sigma \cdot \deg_G(x) - \beta \cdot \deg_G(x) + \lambda \cdot |Y|\Big)$$

$$\geq \sum_{(x,y)\in E} \Big(\lambda \cdot |Y| - (\beta - \sigma) \cdot (1 - \mu) \cdot |Y|\Big)$$

$$= |E| \cdot (\lambda - (\beta - \sigma)(1 - \mu)) \cdot |Y|$$

$$\geq \frac{\delta(\beta - \alpha)}{\beta}(\lambda - (\beta - \sigma)(1 - \mu)) \cdot |X| \cdot |Y|^2 = \rho \cdot |X| \cdot |Y|^2.$$

Rearranging yields the desired inequality. $\qquad \square$

The following definition, due to Pokrovskiy [18] will also play an important role in our proofs.

Definition 4.3. Given a family $\mathcal{T}$ of disjoint rainbow independent sets in a coloured matroid, and a positive integer ℓ, the ℓ-*reduction* $\mathcal{T}'$ of $\mathcal{T}$ is defined as follows. We consider all elements $x \in E(\mathcal{T})$ with the property that there exist ℓ different sets $T \in \mathcal{T}$ such that $T + x$ is a rainbow independent set. Now, let $\mathcal{T}'$ be the family of disjoint rainbow independent sets obtained from $\mathcal{T}$ by deleting all elements $x \in E(\mathcal{T})$ with this property from their respective rainbow independent sets in $\mathcal{T}$.

Recall that $T + x$ is to be interpreted as a multi-set here. Hence, $T + x$ being a rainbow independent set in particular means $x \notin T$. Intuitively, for every element $e \in E(\mathcal{T})$ with the property in the definition above, there is plenty of flexibility to move e to different member of $\mathcal{T}$ without violating rainbowness or independence. This intuition is formalised by the following lemma, also due to Pokrovskiy [18] (this lemma is essentially a special case of [18, Lemma 11]). For the reader's convenience we include a simple proof of this lemma here.

Lemma 4.4. *Fix integers $r \geq 0$ and $\ell > 3^r$. Consider a family $\mathcal{T} = \{T_1, \ldots, T_m\}$ of disjoint rainbow independent sets in a coloured matroid, and define families $\mathcal{T}^{(0)}, \ldots, \mathcal{T}^{(r)}$ of disjoint rainbow independent sets by setting $\mathcal{T}^{(0)} = \mathcal{T}$ and taking $\mathcal{T}^{(i)}$ to be the ℓ-reduction of the family $\mathcal{T}^{(i-1)}$ for $i = 1, \ldots, r$. Suppose that there is an element $e \notin E(\mathcal{T})$ of the matroid such that $T + e$ is a rainbow independent set for some $T \in \mathcal{T}^{(i)}$ for some index $i \in \{0, \ldots, r\}$. Then there exists a family $\mathcal{S} = \{S_1, \ldots, S_m\}$ of disjoint rainbow independent sets with $E(\mathcal{S}) = E(\mathcal{T}) + e$, such that $\sum_{j=1}^{m} |S_j \setminus T_j| \leq 3^r$.*

Proof. We prove the lemma by induction on r. In the case $r = 0$, by assumption $T + e$ is a rainbow independent set for some $T \in \mathcal{T}^{(0)} = \mathcal{T}$. Then we can take $\mathcal{S}$ to be the family obtained from $\mathcal{T}$ by replacing T with $T + e$ (i.e., by adding e to the set T).

So let us now assume that $r \geq 1$ and that we already proved the statement in the lemma for $r - 1$. First note that $\mathcal{T}^{(1)}$ is the ℓ-reduction of the family $\mathcal{T}^{(0)} = \mathcal{T}$. This means that, letting $\mathcal{T}^{(1)} = \{T_1', \ldots, T_m'\}$, for $j = 1, \ldots, m$ we have $T_j' \subseteq T_j$, and for every element $x \in E(\mathcal{T}) \setminus E(\mathcal{T}^{(1)})$ there are ℓ different indices $j \in [m]$ such that $T_j + x$ is a rainbow independent set.

By assumption, there exist $i \in \{0, \ldots, r\}$ and $T \in \mathcal{T}^{(i)}$ such that $T + e$ is a rainbow independent set. If $i \leq r - 1$, the desired statement already follows from the induction hypothesis for $r - 1$, so let us assume that $i = r$. By applying the induction hypothesis to the family $\mathcal{T}^{(1)} = \{T_1', \ldots, T_m'\}$ we find a family $\mathcal{S}' = \{S_1', \ldots, S_m'\}$ of disjoint rainbow independent sets with $E(\mathcal{S}') = E(\mathcal{T}^{(1)}) + e$, such that $\sum_{j=1}^{m} |S_j' \setminus T_j'| \leq 3^{r-1}$.

Now, we can extend $\mathcal{S}'$ to a family $\mathcal{S}^* = \{S_1^*, \ldots, S_m^*\}$ of disjoint rainbow independent sets such that for $j = 1, \ldots, m$ we have $S_j' \subseteq S_j^* \subseteq S_j' \cup (T_j \setminus T_j')$ and $|T_j \setminus (T_j' \cup S_j^*)| \leq 2 \cdot |S_j' \setminus T_j'|$. Indeed, for each $j = 1, \ldots, m$ let us add the elements of $S_j' \setminus T_j'$ to the rainbow independent set $(S_j' \cap T_j') \cup (T_j \setminus T_j')$, reinstating independence and resolving colouring conflicts by deleting at most $2 \cdot |S_j' \setminus T_j'|$ elements chosen from $T_j \setminus T_j'$ (formally this follows from Lemma 2.1 applied to the rainbow independent sets S_j' and $(S_j' \cap T_j') \cup (T_j \setminus T_j')$). Then we have

$$\sum_{j=1}^{m} |S_j^* \setminus T_j| = \sum_{j=1}^{m} |S_j' \setminus T_j| \leq \sum_{j=1}^{m} |S_j' \setminus T_j'| \leq 3^{r-1}. \tag{3}$$

Furthermore, note that $E(\mathcal{T}^{(1)}) + e = E(\mathcal{S}') \subseteq E(\mathcal{S}^*) \subseteq E(\mathcal{T}) + e$ and therefore

$$|E(\mathcal{T}) \setminus E(\mathcal{S}^*)| = |E(\mathcal{T}) \setminus (E(\mathcal{T}^{(1)}) \cup E(\mathcal{S}^*))| \leq \sum_{j=1}^{m} |T_j \setminus (T_j' \cup S_j^*)| \leq 2 \sum_{j=1}^{m} |S_j' \setminus T_j'| \leq 2 \cdot 3^{r-1}.$$

Defining $X = E(\mathcal{T}) \setminus E(\mathcal{S}^*)$, let $X = \{x_1, \ldots, x_z\}$, and note that then $z = |X| \leq 2 \cdot 3^{r-1}$. We will now show that we can distribute the elements in X among the sets in $\mathcal{S}^*$ while maintaining rainbowness and independence.

Claim 4.5. *For each $h \in [z]$, $S_j^* + x_h$ is a rainbow independent set for at least z indices $j \in [m]$.*

Using this claim, we can complete the proof easily. Indeed, by Claim 4.5, we can greedily choose distinct indices $j(1), \ldots, j(z) \in [m]$ such that $S_{j(h)}^* + x_h$ is a rainbow independent set for each $h \in [z]$. Then, letting $S_{j(h)} = S_{j(h)}^* + x_{j(h)}$ for each $h \in [z]$, and $S_j = S_j^*$ for each $j \in [m] \setminus \{j(1), \ldots, j(z)\}$, the

12

family $\mathcal{S} = \{S_1, \ldots, S_m\}$ of disjoint rainbow independent sets satisfies $E(\mathcal{S}) = E(\mathcal{S}^*) \cup X = E(\mathcal{T}) + e$ (as $e \in E(\mathcal{S}') \subseteq E(\mathcal{S}^*)$). By (3), we furthermore have $\sum_{j=1}^{m} |S_j \setminus T_j| \leq |X| + \sum_{j=1}^{m} |S_j^* \setminus T_j| \leq 2 \cdot 3^{r-1} + 3^{r-1} = 3^r$, as desired. Thus, it only remains to prove Claim 4.5.

Proof of Claim 4.5. Let $h \in [z]$. As $E(\mathcal{T}^{(1)}) + e \subseteq E(\mathcal{S}^*)$, we have that $X \subseteq E(\mathcal{T}) \setminus E(\mathcal{T}^{(1)})$ and, hence, that $T_j + x_h$ is a rainbow independent set for at least ℓ values of $j \in [m]$ by the definition of $\mathcal{T}^{(1)}$. From (3), we can observe that $S_j^* \subseteq T_j$ for all but at most 3^{r-1} indices $j \in [m]$. Thus, for at least $\ell - 3^{r-1} \geq 2 \cdot 3^{r-1} \geq z$ indices $j \in [m]$ we have that $S_j^* + x_h \subseteq T_j + x_h$ and $T_j + x_h$ (and hence $S_j^* + x_h$) is a rainbow independent set. This completes the proof of the claim and hence the lemma. $\qquad \boxdot \quad \square$

We will deduce Lemma 3.1 from Lemma 4.4 and the following lemma. The statement of this lemma has some similarities to the statement of [18, Lemma 10] due to Pokrovskiy, but there are several important differences between the two statements. These differences are crucial in order to obtain the stronger statement in Lemma 3.1 compared to Theorem 1.3 from [18].

Lemma 4.6. *For any $\mu > 0$ and $0 < \lambda < 1$ there exists $\gamma = \gamma(\lambda, \mu) > 0$ such that for any positive integer ℓ and any sufficiently large n (sufficiently large with respect to λ, μ and ℓ) the following holds. Consider a coloured rank-n matroid with colour classes $B_1, \ldots, B_n$, such that each of the sets B_i for $i = 1, \ldots, n$ is a basis, and let $\mathcal{T}$ be a family of $\lfloor (1 + \lambda)n \rfloor$ disjoint rainbow independent sets. Then, at least one of the following three statements holds.*

(i) *There is some $T \in \mathcal{T}$ of size $|T| \leq \mu n$.*

(ii) *There is some $T \in \mathcal{T}$ and some colour $c \in [n]$ such that there are strictly more than $\deg_{\mathcal{A}(\mathcal{T})}(c) - \lfloor \lambda n \rfloor$ elements $e \in B_c \setminus E(\mathcal{T})$ such that $T + e$ is a rainbow independent set.*

(iii) *The ℓ-reduction $\mathcal{T}'$ of the family $\mathcal{T}$ satisfies $|E(\mathcal{T}')| \leq |E(\mathcal{T})| - \gamma n^2$.*

Proof. Let $\rho > 0$ be such that the statement in Corollary 4.2 holds with $\alpha = 1 + \lambda$ and $\delta = \lambda/4$ and $\sigma = 1$ (as well as the given values of λ and μ), noting that $1 + \lambda \geq \alpha > 1$. Now, let $\gamma = \rho/3 > 0$, let ℓ be a positive integer, and assume that n is sufficiently large with respect to all the other parameters. Consider a coloured matroid and a family $\mathcal{T}$ of disjoint rainbow independent sets as in the assumptions of the lemma, and let us assume for a contradiction that none of (i), (ii), or (iii) holds.

Recall the definition of the colour-availability graph $\mathcal{A}(\mathcal{T})$ from the beginning of this section, and note that for every colour $c \in [n]$ we have

$$\deg_{\mathcal{A}(\mathcal{T})}(c) = |\mathcal{T}| - |E(\mathcal{T}) \cap B_c| \geq |\mathcal{T}| - n = \lfloor (1 + \lambda)n \rfloor - n = \lfloor \lambda n \rfloor \geq \frac{\lambda}{2} \cdot n \geq \frac{\lambda}{4} \cdot |\mathcal{T}|$$

(using that n is sufficiently large). Furthermore, for each $T \in \mathcal{T}$, as (i) does not hold and thus $|T| > \mu n$, we have $\deg_{\mathcal{A}(\mathcal{T})}(T) = n - |T| < (1 - \mu)n$. Thus, by the property of ρ from its choice via Corollary 4.2, applied with $G = \mathcal{A}(\mathcal{T})$, $X = \mathcal{T}$ and $Y = [n]$, there is some set $E \subseteq E(\mathcal{A}(\mathcal{T}))$ such that

$$\sum_{(T,c) \in E} \deg_{\mathcal{A}(\mathcal{T})}(T) \geq \rho \cdot \lfloor (1 + \lambda)n \rfloor \cdot n^2 + \sum_{(T,c) \in E} (\deg_{\mathcal{A}(\mathcal{T})}(c) - \lambda n)$$

$$\geq \rho n^3 - |E| + \sum_{(T,c) \in E} (\deg_{\mathcal{A}(\mathcal{T})}(c) - \lfloor \lambda n \rfloor)$$

$$\geq 2\gamma n^3 + \sum_{(T,c) \in E} |\{e \in B_c \setminus E(\mathcal{T}) : T + e \text{ rainbow independent}\}|, \qquad (4)$$

where in the last step we have used that (ii) does not hold (and that $\rho n^3 - |E| \geq 3\gamma n^3 - (1+\lambda)n^2 \geq 2\gamma n^3$).

Now, for each $e \in E(\mathcal{T}')$, by Definition 4.3 there are at most ℓ different sets $T \in \mathcal{T}$ such that $T + e$ is

a rainbow independent set. Thus,

$$\sum_{(T,c)\in E} |\{e \in B_c \cap E(\mathcal{T}) : T + e \text{ rainbow independent}\}| \leq \sum_{T \in \mathcal{T}} |\{e \in E(\mathcal{T}) : T + e \text{ rainbow independent}\}|$$

$$= \sum_{e \in E(\mathcal{T})} |\{T \in \mathcal{T} : T + e \text{ rainbow independent}\}|$$

$$\leq |E(\mathcal{T}')| \cdot \ell + |E(\mathcal{T}) \setminus E(\mathcal{T}')| \cdot |\mathcal{T}|$$

$$\leq n^2 \cdot \ell + \gamma n^2 \cdot (1 + \lambda)n < 2\gamma n^3,$$

where we have used that (iii) does not hold and therefore $|E(\mathcal{T}) \setminus E(\mathcal{T}')| = |E(\mathcal{T})| - |E(\mathcal{T}')| \leq \gamma n^2$. Combining this with (4) yields

$$\sum_{(T,c)\in E} \deg_{\mathcal{A}(\mathcal{T})}(T) > \sum_{(T,c)\in E} |\{e \in B_c : T + e \text{ rainbow independent}\}|. \tag{5}$$

On the other hand, for each pair $(T, c) \in E$, there is an edge in $\mathcal{A}(\mathcal{T})$ between T and c, so the colour c does not appear on T. Furthermore, as B_c is an independent set, there are at least $|B_c| - |T| = n - |T| = \deg_{\mathcal{A}(\mathcal{T})}(T)$ different elements $e \in B_c$ such that $T + e$ is independent. For all of these elements e, the set $T + e$ is then a rainbow independent set. Therefore we have

$$\sum_{(T,c)\in E} |\{e \in B_c : T + e \text{ rainbow independent}\}| \geq \sum_{(T,c)\in E} \deg_{\mathcal{A}(\mathcal{T})}(T),$$

a contradiction to (5). □

Finally here, let us combine Lemmas 4.4 and 4.6 to prove Lemma 3.1.

Proof of Lemma 3.1. Let $\mu = \nu/2$ and let $\gamma > 0$ be chosen as in Lemma 4.6 for λ and μ. Fix an integer $r > 1/\gamma$ and let $\ell = 3^r + 1$. We may assume that n is sufficiently large with respect to λ, γ, r and ℓ. Suppose, then, that $\mathcal{T}$ is a family of $\lfloor (1 + \lambda)n \rfloor$ disjoint rainbow independent sets such that $E(\mathcal{T})$ is inclusion-wise maximal and assume for contradiction that $|E(\mathcal{T})| < (1 - \nu)n^2$.

Let $U = (B_1 \cup \cdots \cup B_n) \setminus E(\mathcal{T})$ be the set of elements of $B_1 \cup \cdots \cup B_n$ not covered by $\mathcal{T}$, so that $|U| = n^2 - |E(\mathcal{T})| > \nu n^2$. Now, let us define families $\mathcal{T}^{(0)}, \ldots, \mathcal{T}^{(r)}$ of disjoint rainbow independent sets as in Lemma 4.4 by setting $\mathcal{T}^{(0)} = \mathcal{T}$ and taking $\mathcal{T}^{(i)}$ to be the ℓ-reduction of the family $\mathcal{T}^{(i-1)}$ for $i = 1, \ldots, r$.

If there is some element $e \in U = (B_1 \cup \cdots \cup B_n) \setminus E(\mathcal{T})$, index $i \in \{0, \ldots, r\}$, and $T \in \mathcal{T}^{(i)}$ such that $T + e$ is a rainbow independent set, then, by Lemma 4.4 there is a family $\mathcal{S}$ of $\lfloor (1 + \lambda)n \rfloor$ disjoint rainbow independent sets with $E(\mathcal{S}) = E(\mathcal{T}) \cup \{e\}$, which contradicts the inclusion-wise maximality of $E(\mathcal{T})$. Thus, we have the observation that there is no element $e \in U = (B_1 \cup \cdots \cup B_n) \setminus E(\mathcal{T})$ such that $T + e$ is a rainbow independent set for some $T \in \mathcal{T}^{(i)}$ for some index $i \in \{0, \ldots, r\}$.

For each $i = 0, \ldots, r - 1$, let us now apply Lemma 4.6 to the family $\mathcal{T}^{(i)}$. Then, for each $i = 0, \ldots, r-1$, one of the three statements (i) to (iii) in Lemma 4.6 holds. We will now show that (i), and afterwards (ii), cannot hold for any index $i \in \{0, \ldots, r - 1\}$, before using that (iii) holds for every $i = 0, \ldots, r - 1$ to reach our final contradiction.

If for some index $i \in \{0, \ldots, r - 1\}$ statement (i) holds, then there is a set $T \in \mathcal{T}^{(i)}$ with $|T| \leq \mu n = \nu n/2$. If for some colour c not appearing on T we had $|U \cap B_c| > |T|$, then we could find an element $e \in U \cap B_c$ such that $T + e$ is independent. But then $T + e$ would be a rainbow independent set, which is a contradiction to our observation above. Therefore, for every colour c not appearing on T we have $|U \cap B_c| \leq |T|$. For the $|T|$ colours c appearing on T, we clearly have $|U \cap B_c| \leq n$, and so in total we obtain

$$|U| = \sum_{c=1}^{n} |U \cap B_c| \leq n \cdot |T| + |T| \cdot n = |T| \cdot 2n \leq \nu n^2,$$

contradicting $|U| > \nu n^2$. Thus, (i) cannot hold for any index $i \in \{0, \ldots, r - 1\}$.

If for some index $i \in \{0, \ldots, r-1\}$ statement (ii) holds, then there is a set $T \in \mathcal{T}^{(i)}$ and some colour $c \in [n]$ such that there are strictly more than $\deg_{\mathcal{A}(\mathcal{T}^{(i)})}(c) - \lfloor \lambda n \rfloor$ elements $e \in B_c \setminus E(\mathcal{T}^{(i)})$ such that $T + e$ is a rainbow independent set. By our above observation we must have $e \in E(\mathcal{T})$ and hence $e \in B_c \cap (E(\mathcal{T}) \setminus E(\mathcal{T}^{(i)}))$ for all of these elements e, and consequently

$$|(B_c \cap E(\mathcal{T})) \setminus (B_c \cap E(\mathcal{T}^{(i)}))| = |B_c \cap (E(\mathcal{T}) \setminus E(\mathcal{T}^{(i)}))| > \deg_{\mathcal{A}(\mathcal{T}^{(i)})}(c) - \lfloor \lambda n \rfloor. \tag{6}$$

On the other hand, $\deg_{\mathcal{A}(\mathcal{T}^{(i)})}(c) = |\mathcal{T}| - |B_c \cap E(\mathcal{T}^{(i)})| = \lfloor (1 + \lambda)n \rfloor - |B_c \cap E(\mathcal{T}^{(i)})|$, and therefore

$$\deg_{\mathcal{A}(\mathcal{T}^{(i)})}(c) - \lfloor \lambda n \rfloor = \lfloor (1 + \lambda)n \rfloor - |B_c \cap E(\mathcal{T}^{(i)})| - \lfloor \lambda n \rfloor = n - |B_c \cap E(\mathcal{T}^{(i)})|$$
$$\geq |B_c \cap E(\mathcal{T})| - |B_c \cap E(\mathcal{T}^{(i)})| = |(B_c \cap E(\mathcal{T})) \setminus (B_c \cap E(\mathcal{T}^{(i)}))|.$$

This contradicts (6), and thus (ii) does not hold for any index $i \in \{0, \ldots, r-1\}$.

Thus, (iii) must hold for all $i = 0, \ldots, r-1$. But this means that $|E(\mathcal{T}^{(i+1)})| \leq |E(\mathcal{T}^{(i)})| - \gamma n^2$ for all $i = 0, \ldots, r-1$. Hence $|E(\mathcal{T}^{(r)})| \leq |E(\mathcal{T}^{(0)})| - r \cdot \gamma n^2 = |E(\mathcal{T})| - r \cdot \gamma n^2 \leq n^2 - r \cdot \gamma n^2 < 0$ (recalling that $r > 1/\gamma$), which is again a contradiction. This finishes the proof of Lemma 3.1. $\qquad \square$

We end this section by showing how Theorem 1.3 can be deduced from Lemma 3.1.

Proof of Theorem 1.3. Let $\lambda = \varepsilon^2/2$ and $\nu = \varepsilon^2/2$, and let n be sufficiently large such that the statement in Lemma 3.1 holds. Let us choose a family $\mathcal{T}$ of $\lfloor (1 + \lambda)n \rfloor$ disjoint rainbow independent sets such that $E(\mathcal{T})$ is inclusion-maximal among all such families. By Lemma 3.1 we have $|E(\mathcal{T})| \geq (1 - \nu)n^2$. Let m be the number of sets $T \in \mathcal{T}$ with size at least $(1 - \varepsilon)n$. Then, the m largest sets $T \in \mathcal{T}$ each have size at most $|T| \leq n$, and the remaining $\lfloor (1 + \lambda)n \rfloor - m \leq (1 + \lambda)n - m$ sets $T \in \mathcal{T}$ each have size $|T| < (1 - \varepsilon)n$, so the total size $|E(\mathcal{T})|$ of the family $\mathcal{T}$ satisfies

$$(1 - \nu)n^2 \leq |E(\mathcal{T})| \leq m \cdot n + ((1 + \lambda)n - m) \cdot (1 - \varepsilon)n = (1 + \lambda)(1 - \varepsilon)n^2 + m \cdot \varepsilon n \leq (1 - \varepsilon + \lambda)n^2 + m \cdot \varepsilon n.$$

This implies $m \geq (\varepsilon - \lambda - \nu)n^2/(\varepsilon n) = (\varepsilon - \varepsilon^2)n/\varepsilon = (1 - \varepsilon)n$, meaning that $\mathcal{T}$ contains at least $(1 - \varepsilon)n$ sets $T \in \mathcal{T}$ of size $|T| \geq (1 - \varepsilon)n$, each of which is rainbow and independent, so the conclusion of the theorem holds. $\qquad \square$

5 Large families avoiding a random set: proof of Lemma 3.2

The proof of Lemma 3.2 is similar to the proof of Lemma 3.1 in the previous section, and will in particular also rely on Corollary 4.2 and Lemma 4.4. First, we show that with high probability the random set $R \subseteq B_1 \cup \cdots \cup B_n$ in Lemma 3.2 has the properties in the following two lemmas. The property ($\blacklozenge$) in the first lemma estimates how often each colour appears in R, and the property ($\bigstar$) in the second lemma states, roughly speaking, that, for every independent set T and any set $C \subseteq [n]$, there are many ways to extend T to a larger independent set by adding an element in $\bigcup_{c \in C} B_c \setminus R$.

Lemma 5.1. *For any $0 < \eta < 1$ and $\gamma > 0$, the following holds for any coloured rank-n matroid with colour classes $B_1, \ldots, B_n$ of size n. Let $R \subseteq B_1 \cup \cdots \cup B_n$ be a set of elements drawn independently at random with probability η. Then, with high probability (more precisely, with probability tending to 1 as $n \to \infty$ with η and γ fixed), the following holds.*

($\blacklozenge$) *For every colour $c \in [n]$, we have $(\eta - \gamma)n \leq |B_c \cap R| \leq (\eta + \gamma)n$.*

Proof. For each colour $c \in [n]$, note that $|B_c \cap R|$ is a sum of n independent Bernoulli random variables and $\mathbb{E}[|B_c \cap R|] = \eta n$. Thus, by a Chernoff bound (see e.g. [2, Theorem A.1.4]) we have

$$\P[||B_c \cap R| - \eta n| > \gamma n] \leq 2 \cdot e^{-2(\gamma n)^2/n} = 2 \cdot e^{-2\gamma^2 n}.$$

Taking a union bound over all $c \in [n]$ shows that with probability at least $1 - n \cdot 2 \cdot e^{-2\gamma^2 n} = 1 - o(1)$ we indeed have $(\eta - \gamma)n \leq |B_c \cap R| \leq (\eta + \gamma)n$ for all $c \in [n]$. $\qquad \square$

Lemma 5.2. *For any $0 < \eta < 1$ and $\gamma > 0$, the following holds for any coloured rank-n matroid with colour classes $B_1, \ldots, B_n$, such that each of the sets B_i for $i = 1, \ldots, n$ is a basis. Let $R \subseteq B_1 \cup \cdots \cup B_n$ be a set of elements drawn independently at random with probability η. Then, with high probability (more precisely, with probability tending to 1 as $n \to \infty$ with η and γ fixed), the following holds.*

($\bigstar$) *For every independent set $T \subseteq B_1 \cup \cdots \cup B_n$, and every colour subset $C \subseteq [n]$, there are at least $(1 - \eta) \cdot (n - |T|) \cdot |C| - \gamma n^2$ elements $e \in \bigcup_{c \in C} B_c \setminus R$ such that $T + e$ is independent.*

Proof. For each independent set $T \subseteq B_1 \cup \cdots \cup B_n$ and each colour subset $C \subseteq [n]$, there are at least $(n - |T|) \cdot |C|$ elements $e \in \bigcup_{c \in C} B_c$ such that $T + e$ is independent (since there are at least $n - |T|$ such elements in every colour $c \in C$); say the set of those elements e is $E_{T,C}$. Now, property ($\bigstar$) holds for T and C precisely when $|E_{T,C} \setminus R| \geq (1 - \eta)(n - |T|)|C| - \gamma n^2$. Note that we have $\mathbb{E}[|E_{T,C} \setminus R|] = (1 - \eta)|E_{T,C}| \geq (1 - \eta)(n - |T|)|C|$. Thus, observing that $|E_{T,C} \setminus R|$ is a sum of $|E_{T,C}| \leq n^2$ independent Bernoulli random variables, by a Chernoff bound (see e.g. [2, Theorem A.1.4]) we have

$$\P\left[|E_{T,C} \setminus R| \leq (1-\eta)(n-|T|)|C| - \gamma n^2\right] \leq \P\left[|E_{T,C} \setminus R| \leq \mathbb{E}[|E_{T,C} \setminus R|] - \gamma n^2\right] \leq e^{-2(\gamma n^2)^2/n^2} = e^{-2\gamma^2 n^2}.$$

for each independent set $T \subseteq B_1 \cup \cdots \cup B_n$ and each colour subset $C \subseteq [n]$. Note there are at most $\sum_{t=0}^{n} \binom{n^2}{t} \leq (n^2)^n = n^{2n}$ independent sets $T \subseteq B_1 \cup \cdots \cup B_n$ and only 2^n colour subsets $C \subseteq [n]$. Thus, the probability that property ($\bigstar$) fails for some choice of T and C is at most $n^{2n} \cdot 2^n \cdot e^{-2\gamma^2 n^2} = o(1)$, as required. $\qquad\square$

The following lemma is the analogue of Lemma 4.6 needed in the proof of Lemma 3.2.

Lemma 5.3. *For any $0 < \varepsilon < \eta < 1$ and any $\mu > 0$ there exists $\gamma = \gamma(\varepsilon, \eta, \mu) > 0$ such that for any positive integer ℓ and any sufficiently large n (sufficiently large with respect to ε, η, μ and ℓ) the following holds. Consider a coloured rank-n matroid with colour classes $B_1, \ldots, B_n$, such that each of the sets B_i for $i = 1, \ldots, n$ is a basis, and let $R \subseteq B_1 \cup \cdots \cup B_n$ satisfy properties ($\blacklozenge$) and ($\bigstar$) above. Finally, let $\mathcal{T}$ be a family of $\lfloor (1 - \varepsilon)n \rfloor$ disjoint rainbow independent sets in $(B_1 \cup \cdots \cup B_n) \setminus R$. Then at least one of the following three statements holds.*

(i) *There is some $T \in \mathcal{T}$ of size $|T| \leq \mu n$.*

(ii) *For some $T \in \mathcal{T}$ and some colour $c \in [n]$ there are strictly more than $\deg_{\mathcal{A}(\mathcal{T})}(c) - |B_c \cap R| + \lceil \varepsilon n \rceil$ elements $e \in B_c \setminus (E(\mathcal{T}) \cup R)$ such that $T + e$ is a rainbow independent set.*

(iii) *The ℓ-reduction $\mathcal{T}'$ of the family $\mathcal{T}$ satisfies $|E(\mathcal{T}')| \leq |E(\mathcal{T})| - \gamma n^2$.*

Note that if this statement holds for some value of $\gamma = \gamma(\varepsilon, \eta, \mu) > 0$, it automatically also holds for all smaller positive values of γ (indeed, the assumptions ($\blacklozenge$) and ($\bigstar$) get harder to satisfy for smaller γ and the conclusion in (iii) gets easier to satisfy for smaller γ).

The proof of the lemma is very similar to the proof of Lemma 4.6 in the previous section (but it requires some adaptations to incorporate the set R).

Proof of Lemma 5.3. Let $\rho > 0$ be such that the statement in Corollary 4.2 holds with $\alpha = 1 - \varepsilon$, $\delta = (\eta - \varepsilon)/2$, $\sigma = 1 - \eta$ and $\lambda = \eta - \varepsilon$ (as well as the given value of μ), noting that $(1 - \eta) + (\eta - \varepsilon) \geq \alpha > 1 - \eta$. Now, let $\gamma = \min(\rho(1 - \varepsilon)/4, (\eta - \varepsilon)/4) > 0$, let ℓ be a positive integer, and assume that n is sufficiently large with respect to all the other parameters. Let us consider a coloured matroid and a family $\mathcal{T}$ of disjoint rainbow independent sets as in the assumptions of the lemma, and assume for a contradiction that none of (i), (ii), or (iii) holds.

Recall the definition of the colour-availability graph $\mathcal{A}(\mathcal{T})$ from the beginning of Section 4, and note that for every colour $c \in [n]$ we have

$$\deg_{\mathcal{A}(\mathcal{T})}(c) = |\mathcal{T}| - |E(\mathcal{T}) \cap B_c| \geq \lfloor (1-\varepsilon)n \rfloor - |B_c \setminus R| \geq (1-\varepsilon-\gamma)n - (1-\eta+\gamma)n = (\eta-\varepsilon-2\gamma)\cdot n \geq \frac{\eta - \varepsilon}{2} \cdot |\mathcal{T}|,$$

using that $|B_c \setminus R| = n - |B_c \cap R| \leq (1 - \eta + \gamma)n$ by ($\blacklozenge$) and that n is sufficiently large. Furthermore, for each $T \in \mathcal{T}$, as (i) does not hold and thus $|T| > \mu n$, we have $\deg_{\mathcal{A}(\mathcal{T})}(T) = n - |T| < (1 - \mu)n$. Thus, by the statement in Corollary 4.2, there is some set $E \subseteq E(\mathcal{A}(\mathcal{T}))$ such that

$$
\begin{aligned}
(1 - \eta) \cdot \sum_{(T,c) \in E} \deg_{\mathcal{A}(\mathcal{T})}(T) &\geq \rho \cdot \lfloor (1 - \varepsilon)n \rfloor \cdot n^2 + \sum_{(T,c) \in E} (\deg_{\mathcal{A}(\mathcal{T})}(c) - (\eta - \varepsilon)n) \\
&\geq \rho(1 - \varepsilon)n^3 - n^2 - |E| \cdot (1 + \gamma n) + \sum_{(T,c) \in E} (\deg_{\mathcal{A}(\mathcal{T})}(c) - (\eta - \gamma)n + \lceil \varepsilon n \rceil) \\
&\geq \rho(1 - \varepsilon)n^3 - 2\gamma n^3 + \sum_{(T,c) \in E} (\deg_{\mathcal{A}(\mathcal{T})}(c) - |B_c \cap R| + \lceil \varepsilon n \rceil) \\
&\geq 2\gamma n^3 + \sum_{(T,c) \in E} |\{e \in B_c \setminus (E(\mathcal{T}) \cup R) : T + e \text{ rainbow independent}\}|, \quad (7)
\end{aligned}
$$

where in the third step we used that that $|B_c \cap R| \geq (\eta - \gamma)n$ by ($\blacklozenge$) and in the last step that (ii) does not hold (and that $\rho(1 - \varepsilon) \geq 4\gamma$).

Now, for each $e \in E(\mathcal{T}')$, there are at most ℓ different sets $T \in \mathcal{T}$ such that $T + e$ is a rainbow independent set. Thus,

$$
\begin{aligned}
\sum_{(T,c) \in E} |\{e \in B_c \cap E(\mathcal{T}) : T + e \text{ rainbow independent}\}| &\leq \sum_{T \in \mathcal{T}} |\{e \in E(\mathcal{T}) : T + e \text{ rainbow independent}\}| \\
&= \sum_{e \in E(\mathcal{T})} |\{T \in \mathcal{T} : T + e \text{ rainbow independent}\}| \\
&\leq |E(\mathcal{T}')| \cdot \ell + |E(\mathcal{T}) \setminus E(\mathcal{T}')| \cdot |\mathcal{T}| \\
&\leq n^2 \cdot \ell + \gamma n^2 \cdot (1 - \varepsilon)n < \gamma n^3,
\end{aligned}
$$

where we have used that (iii) does not hold and therefore $|E(\mathcal{T}) \setminus E(\mathcal{T}')| = |E(\mathcal{T})| - |E(\mathcal{T}')| \leq \gamma n^2$. Combining this with (7) yields

$$
(1 - \eta) \cdot \sum_{(T,c) \in E} (n - |T|) = (1 - \eta) \cdot \sum_{(T,c) \in E} \deg_{\mathcal{A}(\mathcal{T})}(T) > \gamma n^3 + \sum_{(T,c) \in E} |\{e \in B_c \setminus R : T + e \text{ rainbow independent}\}|. \quad (8)
$$

On the other hand, for each $T \in \mathcal{T}$, let $C(T) \subseteq [n]$ be the set of colours $c \in [n]$ with $(T, c) \in E$. Then, for every $c \in C(T)$, there is an edge in $\mathcal{A}(\mathcal{T})$ between T and c, so the colour c does not appear on T. Furthermore, for each $T \in \mathcal{T}$, by ($\bigstar$) there are at least $(1 - \eta) \cdot (n - |T|) \cdot |C(T)| - \gamma n^2$ elements $e \in \bigcup_{c \in C(T)} B_c \setminus R$ such that $T + e$ is independent. For all of these elements e, the set $T + e$ is then a rainbow independent set. Therefore, for each $T \in \mathcal{T}$, we have

$$
\sum_{c \in C(T)} |\{e \in B_c \setminus R : T + e \text{ rainbow independent}\}| \geq (1 - \eta) \cdot (n - |T|) \cdot |C(T)| - \gamma n^2.
$$

Summing this up for all $T \in \mathcal{T}$, we obtain

$$
\begin{aligned}
\sum_{(T,c) \in E} |\{e \in B_c \setminus R : T + e \text{ rainbow independent}\}| &\geq \sum_{T \in \mathcal{T}} \left((1 - \eta) \cdot (n - |T|) \cdot |C(T)| - \gamma n^2 \right), \\
&> (1 - \eta) \cdot \sum_{(T,c) \in E} (n - |T|) - \gamma n^3,
\end{aligned}
$$

which contradicts (8). $\qquad \square$

Again, we can combine Lemmas 4.4 and 5.3 to prove Lemma 3.2.

Proof of Lemma 3.2. Let $\mu = \nu/3$ and let $\gamma > 0$ be chosen as in Lemma 5.3. We may assume without loss of generality that $\gamma < \nu/3$ (since making γ smaller will not violate the statement in Lemma 5.3). Fix an integer $r > 1/\gamma$ and let $\ell = 3^r + 1$. We may assume that n is sufficiently large with respect to ε, η, γ,

r and ℓ. Take, then, any coloured rank-n matroid with colour classes $B_1, \ldots, B_n$, such that each of the sets B_i for $i = 1, \ldots, n$ is a basis, and let $R \subseteq B_1 \cup \cdots \cup B_n$ be a set of elements drawn independently at random with probability η.

By Lemmas 5.1 and 5.2, with high probability the random set $R \subseteq B_1 \cup \cdots \cup B_n$ satisfies ($\blacklozenge$) and ($\bigstar$). Thus, it suffices to prove that whenever these properties are satisfied, there exists a family $\mathcal{T}$ of $\lfloor (1 - \varepsilon)n \rfloor$ disjoint rainbow independent sets in $(B_1 \cup \cdots \cup B_n) \setminus R$ of total size $|E(\mathcal{T})| \geq (1 - \eta - \nu)n^2$.

Let us assume then that ($\blacklozenge$) and ($\bigstar$) hold, and let $\mathcal{T}$ be a family of $\lfloor (1 - \varepsilon)n \rfloor$ disjoint rainbow independent sets in $(B_1 \cup \cdots \cup B_n) \setminus R$ maximising $|E(\mathcal{T})|$. Let us suppose for contradiction that $|E(\mathcal{T})| < (1 - \eta - \nu)n^2$. Then the set $U = (B_1 \cup \cdots \cup B_n) \setminus E(\mathcal{T})$ of elements of $B_1 \cup \cdots \cup B_n$ not covered by $\mathcal{T}$ has size $|U| = n^2 - |E(\mathcal{T})| > (\eta + \nu)n^2$ and therefore

$$|U \setminus R| = |U| - |R| > (\eta + \nu)n^2 - \sum_{c=1}^{n} |B_c \cap R| \overset{(\blacklozenge)}{\geq} (\eta + \nu)n^2 - n \cdot (\eta + \gamma)n > (2\nu/3)n^2.$$

Now, we again define families $\mathcal{T}^{(0)}, \ldots, \mathcal{T}^{(r)}$ of disjoint rainbow independent sets as in Lemma 4.4 by setting $\mathcal{T}^{(0)} = \mathcal{T}$ and taking $\mathcal{T}^{(i)}$ to be the ℓ-reduction of the family $\mathcal{T}^{(i-1)}$ for $i = 1, \ldots, r$. We observe that there cannot be an element $e \in U \setminus R = (B_1 \cup \cdots \cup B_n) \setminus (E(\mathcal{T}) \cup R)$ such that $T + e$ is a rainbow independent set for some $T \in \mathcal{T}^{(i)}$ for some index $i \in \{0, \ldots, r\}$ (since otherwise by Lemma 4.4 we would contradict the maximality of $|E(\mathcal{T})|$). Now, for each $i = 0, \ldots, r - 1$, we apply Lemma 5.3 to the family $\mathcal{T}^{(i)}$, and conclude that one of the three statements (i) to (iii) holds.

If for some index $i \in \{0, \ldots, r - 1\}$ statement (i) holds, then there is a set $T \in \mathcal{T}^{(i)}$ with $|T| \leq \mu n = \nu n/3$. For every colour c not appearing on T we must have $|(U \setminus R) \cap B_c| \leq |T|$ (otherwise we could find an element $e \in (U \setminus R) \cap B_c$ such that $T + e$ is a rainbow independent set). For the $|T|$ colours c appearing on T, we have $|(U \setminus R) \cap B_c| \leq n$, and so we can conclude

$$|U \setminus R| = \sum_{c=1}^{n} |(U \setminus R) \cap B_c| \leq n \cdot |T| + |T| \cdot n = |T| \cdot 2n \leq (2\nu/3)n^2.$$

Since this contradicts $|U \setminus R| > (2\nu/3)n^2$, statement (i) cannot hold for any index $i \in \{0, \ldots, r - 1\}$.

If for some index $i \in \{0, \ldots, r - 1\}$ statement (ii) holds, then there is a set $T \in \mathcal{T}^{(i)}$ and some colour $c \in [n]$ such that there are strictly more than $\deg_{\mathcal{A}(\mathcal{T}^{(i)})}(c) - |B_c \cap R| + \lceil \varepsilon n \rceil$ elements $e \in B_c \setminus (E(\mathcal{T}^{(i)}) \cup R)$ such that $T + e$ is a rainbow independent set. By our above observation we have $e \in E(\mathcal{T})$ and hence $e \in B_c \cap (E(\mathcal{T}) \setminus E(\mathcal{T}^{(i)}))$ for all of these elements e, so

$$|B_c \cap (E(\mathcal{T}) \setminus E(\mathcal{T}^{(i)}))| = |(B_c \cap E(\mathcal{T})) \setminus (B_c \cap E(\mathcal{T}^{(i)}))| > \deg_{\mathcal{A}(\mathcal{T}^{(i)})}(c) - |B_c \cap R| + \lceil \varepsilon n \rceil. \quad (9)$$

On the other hand, $\deg_{\mathcal{A}(\mathcal{T}^{(i)})}(c) = |\mathcal{T}| - |B_c \cap E(\mathcal{T}^{(i)})| = \lfloor (1 - \varepsilon)n \rfloor - |B_c \cap E(\mathcal{T}^{(i)})|$, and therefore

$$\deg_{\mathcal{A}(\mathcal{T}^{(i)})}(c) - |B_c \cap R| + \lceil \varepsilon n \rceil = \lfloor (1 - \varepsilon)n \rfloor - |B_c \cap E(\mathcal{T}^{(i)})| - |B_c \cap R| + \lceil \varepsilon n \rceil$$

$$= n - |B_c \cap R| - |B_c \cap E(\mathcal{T}^{(i)})| = |B_c \setminus R| - |B_c \cap E(\mathcal{T}^{(i)})|$$

$$\geq |B_c \cap E(\mathcal{T})| - |B_c \cap E(\mathcal{T}^{(i)})| = |(B_c \cap E(\mathcal{T})) \setminus (B_c \cap E(\mathcal{T}^{(i)}))|.$$

This contradiction to (9) shows that (ii) does not hold for any index $i \in \{0, \ldots, r - 1\}$.

Thus, (iii) holds for all $i = 0, \ldots, r - 1$, so $|E(\mathcal{T}^{(i+1)})| \leq |E(\mathcal{T}^{(i)})| - \gamma n^2$ for all $i = 0, \ldots, r - 1$. This implies $|E(\mathcal{T}^{(r)})| \leq |E(\mathcal{T}^{(0)})| - r \cdot \gamma n^2 \leq n^2 - r \cdot \gamma n^2 < 0$ (as $r > 1/\gamma$), again a contradiction. $\qquad \square$

6 Covering: proof of Theorem 1.6

In addition to Lemma 3.1, the second key for proving Theorem 1.6 is to find a suitable notion of "dense spots" in the set of uncovered elements to make the approach outlined in Section 3 work. Our notion will be introduced in Definition 6.3 below. In order to make this definition, we first need some preparations.

Definition 6.1. For a positive integer k, a subset S of a matroid is called *k-overcrowded* if for every subset $S' \subseteq S$ we have

$$|S \setminus S'| \geq k \cdot (\mathrm{rk}(S) - \mathrm{rk}(S')).$$

18

Note that for any k-overcrowded S, the set S is automatically also k'-overcrowded for any positive integer $k' < k$. It turns out that the union of two k-overcrowded subsets S_1 and S_2 of some matroid is automatically also k-overcrowded.

Lemma 6.2. *Let k be a positive integer, and consider k-overcrowded subsets S_1 and S_2 of some matroid. Then their union $S_1 \cup S_2$ is also k-overcrowded.*

The proof of this lemma is fairly simple and can be found in Section 8, in which we collect the proofs of various simple matroid lemmas needed in this paper. Lemma 6.2 now enables us to make the following definition.

Definition 6.3. For a positive integer k, and a subset U of a matroid, let the *k-deadlock* $D_k(U)$ of U be the unique inclusion-wise maximal k-overcrowded subset of U.

Note that $D_k(U) \subseteq U$, that the k-deadlock $D_k(U)$ is k-overcrowded, and that for every k-overcrowded set $S \subseteq U$ we have $S \subseteq D_k(U)$. Furthermore note that for any subset $U' \subseteq U$ we have $D_k(U') \subseteq D_k(U)$ and for any positive integer $k' < k$ we have $D_k(U) \subseteq D_{k'}(U)$

By our next lemma, the k-deadlock of a set U in a matroid can detect whether there exists a subset $S \subseteq U$ with $|S| > k \cdot \text{rk}(S)$. This is important, because the existence of such a subset S is equivalent to U not being decomposable into k independent sets (see also the discussion below).

Lemma 6.4. *Let k be a positive integer, and let U be a subset of some matroid. If there exists a set $S \subseteq U$ with $|S| > k \cdot \text{rk}(S)$, then S has a non-empty k-overcrowded subset, and so in particular we have $D_k(U) \neq \emptyset$ for the k-deadlock of U.*

Proof. Let $S^* \subseteq S$ be a minimal subset of S with the property that $|S^*| > k \cdot \text{rk}(S^*)$ (such a subset exists, since S itself has this property). Clearly, S^* must be non-empty. Now, for any subset $S' \subsetneq S^*$ we have $|S'| \leq k \cdot \text{rk}(S')$ and hence

$$|S^* \setminus S'| = |S^*| - |S'| \geq k \cdot \text{rk}(S^*) - k \cdot \text{rk}(S') = k \cdot (\text{rk}(S^*) - \text{rk}(S')).$$

Noting that this inequality also holds for $S' = S^*$, we can conclude that S^* is k-overcrowded. The second part of the conclusion follows by noting that $S^* \subseteq D_k(U)$ (since $D_k(U)$ contains all k-overcrowded subsets of $S \subseteq U$). $\qquad\square$

Due to the classical theorem of Edmonds [9] that a matroid can be decomposed into k independent sets if and only if $|S| \leq k \cdot \text{rk}(S)$ for all subsets of the matroid, we obtain the following corollary of Lemma 6.4.

Corollary 6.5. *Let k be a positive integer, and let U be a subset of some matroid. If $D_k(U) = \emptyset$, then U can be decomposed into k independent sets.*

Proof. If $D_k(U) = \emptyset$, then by Lemma 6.4 we have $|S| \leq k \cdot \text{rk}(S)$ for all subsets $S \subseteq U$. Thus, by Edmonds' matroid decomposition theorem [9], U can be decomposed into k independent sets. $\qquad\square$

Aharoni and Berger [1, Theorem 8.9] proved that in order to decompose a coloured matroid into few rainbow independent sets it is sufficient to assume that the underlying (uncoloured) matroid is decomposable into few independent sets and that every colour appears only few times. In fact, they actually proved a more general result for the intersection of two arbitrary matroids, but we state their result only in the special case relevant to us (where one of the two matroids is a partition matroid described by the colouring).

Theorem 6.6 ([1])**.** *Let k be a positive integer, and U a subset of a coloured matroid, such that U can be decomposed into k independent sets and such that every colour appears at most k times in U. Then U can be decomposed into $2k$ rainbow independent sets.*

We remark that this directly implies the result of Aharoni and Berger [1] stated as Theorem 1.4 in the introduction.

By combining Corollary 6.5 and Theorem 6.6, we see that a set U in a coloured matroid can always be decomposed into $2k$ rainbow independent sets if $D_k(U) = \emptyset$ and every colour appears at most k times in

U. Our approach for proving Theorem 1.6 is to find a family $\mathcal{T}$ of $\lfloor(1+\lambda)n\rfloor$ disjoint rainbow independent sets (for some suitable parameter λ), such that the set of uncovered elements $U = (B_1 \cup \cdots \cup B_n) \setminus E(\mathcal{T})$ satisfies these two conditions, and then to decompose U into few additional rainbow independent sets. In order to ensure these conditions for U, i.e. to ensure that $D_k(U) = \emptyset$ and that every colour appears only few times in U, we use switching arguments which exchange certain elements in a set $T \in \mathcal{T}$ by other elements. It is crucial to keep track of how $D_k(U)$ changes during these switching processes. The following lemma allows us to control this.

Lemma 6.7. *Consider positive integers $k' < k$ and disjoint subsets U and U' of some matroid such that $|U'| \leq k - k'$. If $U' \cap \operatorname{span}(D_{k'}(U)) = \emptyset$, then we have $D_k(U) = D_k(U \cup U')$.*

We postpone the proof of Lemma 6.7 to Section 8. Our next result shows that, in the setting relevant for Theorem 1.6, we can find a family of $\lfloor(1+\lambda)n\rfloor$ rainbow independent sets so that the set of uncovered elements has an empty $\lfloor\lambda n\rfloor$-deadlock, as long as n is sufficiently large with respect to $\lambda > 0$.

Lemma 6.8. *For any $0 < \lambda < 1$ the following holds for any sufficiently large n (sufficiently large with respect to λ). In any coloured rank-n matroid with colour classes $B_1, \ldots, B_n$, such that each of the sets B_i for $i = 1, \ldots, n$ is a basis, there is a family $\mathcal{T}$ of $\lfloor(1+\lambda)n\rfloor$ disjoint rainbow independent sets such that, setting $U = (B_1 \cup \cdots \cup B_n) \setminus E(\mathcal{T})$, we have $D_{\lfloor\lambda n\rfloor}(U) = \emptyset$.*

Proof. Let $\nu = \lambda^2/10$, and recall that we are assuming that n is sufficiently large with respect to λ (in particular, large enough that the statement in Lemma 3.1 holds for λ and $\nu = \lambda^2/10$). Let z be the unique positive integer with $\lambda n - 4 < z \leq \lambda n$ which is divisible by 4. Since $z \leq \lfloor\lambda n\rfloor$, we then have $D_{\lfloor\lambda n\rfloor}(U) \subseteq D_z(U)$ for any subset U of the matroid. Thus, it suffices to find a family $\mathcal{T}$ of $\lfloor(1+\lambda)n\rfloor$ disjoint rainbow independent sets such that we have $D_z(U) = \emptyset$ for $U = (B_1 \cup \cdots \cup B_n) \setminus E(\mathcal{T})$.

To this end, let us choose a family $\mathcal{T}$ of $\lfloor(1+\lambda)n\rfloor$ disjoint rainbow independent sets such that, when denoting the set of uncovered elements by $U = (B_1 \cup \cdots \cup B_n) \setminus E(\mathcal{T})$, the $(z/4 + 2)$-tuple

$$\Big(|D_z(U)|, |D_{z-2}(U)|, |D_{z-4}(U)|, \ldots, |D_{z/2}(U)|, |U|\Big) \in \mathbb{Z}^{(z/4)+2} \tag{10}$$

is lexicographically minimised (this means that among all families of $\lfloor(1+\lambda)n\rfloor$ disjoint rainbow independent sets, we choose $\mathcal{T}$ such that size of the z-deadlock $|D_z(U)|$ of U is minimised, and among all such choices such that $|D_{z-2}(U)|$ is minimised, and so on).

Then, in particular, the set $E(\mathcal{T})$ is inclusion-wise maximal among all families $\mathcal{T}$ of $\lfloor(1+\lambda)n\rfloor$ disjoint rainbow independent sets. Indeed, if there was a family $\mathcal{T}^*$ of $\lfloor(1+\lambda)n\rfloor$ disjoint rainbow independent sets with $E(\mathcal{T}) \subsetneq E(\mathcal{T}^*)$, then for its set of uncovered elements $U^* = (B_1 \cup \cdots \cup B_n) \setminus E(\mathcal{T}^*)$ we would have $U^* \subsetneq U$. This means that $|D_k(U^*)| \leq |D_k(U)|$ for all positive integers k, and furthermore $|U^*| < |U|$, contradicting our choice of $\mathcal{T}$ to lexicographically minimise (10). Thus, $E(\mathcal{T})$ is indeed inclusion-wise maximal among all families $\mathcal{T}$ of $\lfloor(1+\lambda)n\rfloor$ disjoint rainbow independent sets.

Consequently, by Lemma 3.1, we have $|E(\mathcal{T})| \geq (1-\nu)n^2$ and hence

$$|U| = n^2 - |E(\mathcal{T})| \leq \nu n^2 = \lambda^2 n^2/10.$$

Let us assume for contradiction that $D_z(U) \neq \emptyset$, then we have $\operatorname{rk}(D_z(U)) \geq 1$ (as each element of $B_1 \cup \cdots \cup B_n$ forms an independent set by itself). Defining $r_k = \operatorname{rk}(D_k(U))$ for $k \in \{z, z-2, z-4, \ldots, z/2\}$, we then have $1 \leq r_z \leq r_{z-2} \leq r_{z-4} \leq \cdots \leq r_{z/2} \leq n$. Since $(1+\lambda/10)^{z/4} \geq (1+\lambda/10)^{\lambda n/4 - 1} > n$ (by our assumption that n is sufficiently large with respect to λ), there is an index $j \in \{z/2 + 2, z/2 + 4, \ldots, z\}$ such that $r_{j-2} \leq (1+\lambda/10)r_j$. Note that, as $D_j(U)$ is j-overcrowded, by the condition in Definition 6.1 with $S = D_j(U)$ and $S' = \emptyset$ we have $|D_j(U)| \geq j \cdot \operatorname{rk}(D_j(U))$, and, hence

$$r_j = \operatorname{rk}(D_j(U)) \leq \frac{|D_j(U)|}{j} \leq \frac{|U|}{j} \leq \frac{\lambda^2 n^2/10}{z/2 + 2} \leq \frac{\lambda^2 n^2/10}{\lambda n/2} = \frac{\lambda n}{5}. \tag{11}$$

Now, let $S = \operatorname{span} D_{j-2}(U)$, and note that then $\operatorname{rk}(S) = \operatorname{rk}(D_{j-2}(U)) = r_{j-2}$ and

$$|S| = \sum_{c=1}^{n} |B_c \cap S| \leq \sum_{c=1}^{n} \operatorname{rk}(S) = r_{j-2} \cdot n, \tag{12}$$

20

using that each of the sets $B_1, \dots, B_n$ is independent. Furthermore, let $W \subseteq D_j(U) \subseteq U$ be a maximal independent set in $D_j(U)$, so that $|W| = \mathrm{rk}(D_j(U)) = r_j$.

As discussed in the proof outline in Section 3.1, our goal is to move some element $e \in W \subseteq D_j(U) \subseteq U$ into one of the sets $T \in \mathcal{T}$, while removing up to two other elements from T in order to maintain rainbowness and independence. However, we need to ensure that these removed elements do not lie in $S = \mathrm{span}\, D_{j-2}(U)$. The following claim will allow us to find an element $e \in W$ and a set $T \in \mathcal{T}$ such that this is possible.

Claim 6.9. There is some $e \in W$ and $T \in \mathcal{T}$ such that $(T \cap S) + e$ is a rainbow independent set.

Proof. We need to show that there is a pair $(e, T) \in W \times \mathcal{T}$ such that $e \notin \mathrm{span}(T \cap S)$ and the colour $c(e)$ does not appear on $T \cap S$. Note that in total there are $|W| \cdot |\mathcal{T}| = r_j \cdot \lfloor (1 + \lambda)n \rfloor$ pairs $(e, T) \in W \times \mathcal{T}$. Our proof proceeds via a counting argument, bounding above the number of pairs $(e, T) \in W \times \mathcal{T}$ with $e \in \mathrm{span}(T \cap S)$, as well as the number of pairs $(e, T) \in W \times \mathcal{T}$ such that $c(e)$ appears on $T \cap S$.

First, the number of pairs $(e, T) \in W \times \mathcal{T}$ with $e \in \mathrm{span}(T \cap S)$ can be bounded above by

$$\sum_{T \in \mathcal{T}} |W \cap \mathrm{span}(T \cap S)| \leq \sum_{T \in \mathcal{T}} \mathrm{rk}(\mathrm{span}(T \cap S)) = \sum_{T \in \mathcal{T}} \mathrm{rk}(T \cap S) \leq \sum_{T \in \mathcal{T}} |T \cap S| \leq |S| \overset{(12)}{\leq} r_{j-2} \cdot n,$$

where in the first step we used that W is an independent set.

Furthermore, there are at most $|W| \cdot r_{j-2} = r_j \cdot r_{j-2}$ pairs $(e, T) \in W \times \mathcal{T}$ such that the colour $c(e)$ of e appears on $T \cap S$. Indeed, for each element $e \in W$, we have $|S \cap B_{c(e)}| \leq \mathrm{rk}(S) = r_{j-2}$ (recalling that $B_{c(e)}$ is an independent set) and therefore there can be at most r_{j-2} different sets $T \in \mathcal{T}$ with $T \cap S \cap B_{c(e)} \neq \emptyset$ (meaning that $T \cap S$ contains an element of $B_{c(e)}$, i.e., an element of colour $c(e)$).

Thus, the number of pairs $(e, T) \in W \times \mathcal{T}$ with $e \notin \mathrm{span}(T \cap S)$ such that $c(e)$ does not appear on $T \cap S$ is at least

$$r_j \cdot \lfloor (1 + \lambda)n \rfloor - r_{j-2} \cdot n - r_j \cdot r_{j-2} = r_j \cdot \lfloor (1 + \lambda)n \rfloor - r_{j-2} \cdot (n + r_j)$$
$$\overset{(11)}{\geq} r_j \cdot (1 + \lambda/2)n - r_{j-2} \cdot (n + \lambda n/5)$$
$$\geq r_j \cdot (1 + \lambda/2)n - (1 + \lambda/10)r_j \cdot (1 + \lambda/5)n > 0. \qquad \boxdot$$

As in the claim, let $e \in W$ and $T \in \mathcal{T}$ be chosen such that $(T \cap S) + e$ is a rainbow independent set. Now, we can form a rainbow independent set T^* of the form $T^* = (T + e) \setminus F$ for some set $F \subseteq T \setminus S$ of size $|F| \leq 2$, by adding the element e to the rainbow independent set T and reinstating independence and resolving colouring conflicts by deleting a set $F \subseteq T \setminus S$ consisting of at most two elements (it suffices to delete one element to ensure independence and one element to resolve a potential colouring conflict with e, more formally this follows from Lemma 2.1 applied to the rainbow independent sets $(T \cap S) + e$ and T).

Now, let $\mathcal{T}^*$ be the family of $\lfloor (1 + \lambda)n \rfloor$ disjoint rainbow independent sets obtained from $\mathcal{T}$ when replacing $T \in \mathcal{T}$ by $T^* = (T + e) \setminus F$. Then we have $E(\mathcal{T}^*) = (E(\mathcal{T}) \setminus F) + e$ and the set $U^* = (B_1 \cup \dots \cup B_n) \setminus E(\mathcal{T}^*)$ of elements not covered by $\mathcal{T}^*$ can be described as $U^* = (U \cup F) - e$.

Now, for any integer $k \geq j$ we have

$$D_k(U^*) = D_k((U \cup F) - e) \subseteq D_k(U \cup F) = D_k(U),$$

where in the last step we used Lemma 6.7, recalling that $F \subseteq T \setminus S = T \setminus \mathrm{span}(D_{j-2}(U))$ and $|F| \leq 2 \leq k - (j - 2)$. Furthermore, for $k = j$ we have

$$D_j(U^*) = D_j((U \cup F) - e) \subsetneq D_j(U \cup F) = D_j(U),$$

where the strict inclusion in the second step follows from the fact that $e \in D_j(U) \subseteq D_j(U \cup F)$, but $e \notin D_j((U \cup F) - e)$ as $e \notin (U \cup F) - e$. This shows that for all positive integers $k \geq j$ we have $|D_k(U^*)| \leq |D_k(U)|$, and for $k = j$ we even have $|D_k(U^*)| < |D_k(U)|$. This contradicts our choice of $\mathcal{T}$ to lexicographically minimise (10), recalling that $j \in \{z, z-2, z-4, \dots, z/2 + 2\}$. Thus, we have $D_z(U) = \emptyset$. $\qquad \square$

Using Lemmas 3.1 and 6.8, we can now prove Theorem 1.6.

Proof of Theorem 1.6. We may assume that $0 < \varepsilon < 1$. Let $\lambda = \varepsilon/3$ and $\nu = \lambda^2/4 = \varepsilon^2/36$, and recall that we are assuming that n is sufficiently large with respect to ε (in particular, large enough such that the statement in Lemma 3.1 holds for $\lambda = \varepsilon/3$ and $\nu = \varepsilon^2/36$, and the statement in Lemma 6.8 holds for $\lambda = \varepsilon/3$).

Consider a rank-n matroid with n given bases $B_1, \ldots, B_n$ (we may assume that $B_1, \ldots, B_n$ are disjoint by introducing additional copies of any element that appears in more than one of these bases). By restricting the matroid to the set $B_1 \cup \cdots \cup B_n$, and colouring each of the bases $B_1, \ldots, B_n$ with a different colour, we obtain a coloured matroid. Using Lemma 6.8, let $\mathcal{T}$ be a family of $\lfloor (1+\lambda)n \rfloor$ disjoint rainbow independent sets such that, when denoting the set of uncovered elements by $U = (B_1 \cup \cdots \cup B_n) \setminus E(\mathcal{T})$, we have $D_{\lfloor \lambda n \rfloor}(U) = \emptyset$ and, subject to this, $\sum_{c=1}^n |B_c \cap U|^2$ is minimised.

Now, similarly as in the proof of Lemma 6.8, the set $E(\mathcal{T})$ is inclusion-wise maximal among all families $\mathcal{T}$ of $\lfloor (1+\lambda)n \rfloor$ disjoint rainbow independent sets (indeed, if there was such a family $\mathcal{T}^*$ with $E(\mathcal{T}) \subsetneq E(\mathcal{T}^*)$, then for its set of uncovered elements $U^* = (B_1 \cup \cdots \cup B_n) \setminus E(\mathcal{T}^*)$ we would have $U^* \subsetneq U$ and therefore $D_{\lfloor \lambda n \rfloor}(U^*) \subseteq D_{\lfloor \lambda n \rfloor}(U) = \emptyset$ and $\sum_{c=1}^n |B_c \cap U^*|^2 < \sum_{c=1}^n |B_c \cap U|^2$). Consequently, by Lemma 3.1, we have $|E(\mathcal{T})| \geq (1-\nu)n^2$ and hence

$$|U| = n^2 - |E(\mathcal{T})| \leq \nu n^2 = \lambda^2 n^2/4.$$

In particular, there can be at most $\lambda n/2$ different colours $c \in \{1, \ldots, n\}$ with $|B_c \cap U| \geq \lambda n/2$.

Furthermore, as the set $D_{\lfloor \lambda n \rfloor - 1}(U) \subseteq U$ is $(\lfloor \lambda n \rfloor - 1)$-overcrowded, it follows from Definition 6.1 that (similarly to (11) in the proof of Lemma 6.8)

$$\mathrm{rk}(D_{\lfloor \lambda n \rfloor - 1}(U)) \leq \frac{\nu n^2}{\lfloor \lambda n \rfloor - 1} \leq \frac{\lambda^2 n^2/4}{\lambda n/2} = \frac{\lambda n}{2}.$$

Claim 6.10. *For every colour $c = 1, \ldots, n$, we have $|B_c \cap U| \leq \lambda n$.*

Proof. Assume, for contradiction, that there is some colour $c' \in [n]$ with $|B_{c'} \cap U| > \lambda n$. Let $T \in \mathcal{T}$ be a set on which colour c' does not appear, and apply Lemma 2.3 to the independent sets $S = B_{c'} \cap U = B_{c'} \setminus E(\mathcal{T})$ and T. If there was an element $x \in B_{c'} \setminus E(\mathcal{T})$ as in (a) in Lemma 2.3, then $T + x$ would be a rainbow independent set, and so one could obtain a family $\mathcal{S}$ of $\lfloor (1+\lambda)n \rfloor$ rainbow independent sets with $E(\mathcal{S}) = E(\mathcal{T}) + x$ from the family $\mathcal{T}$ simply by adding x to the set $T \in \mathcal{T}$. This would contradict $E(\mathcal{T})$ being inclusion-wise maximal, and so (b) in Lemma 2.3 must hold. This means that there is an injection $\phi : B_{c'} \cap U \to T$ such that for every element $x \in B_{c'} \cap U$, the set $T - \phi(x) + x$ is a rainbow independent set.

We claim that there must be an element $x \in B_{c'} \cap U$ with $\phi(x) \notin \mathrm{span}(D_{\lfloor \lambda n \rfloor - 1}(U))$ and $|B_{c(\phi(x))} \cap U| < \lambda n/2$. Indeed, since T is an independent set, we have $|T \cap \mathrm{span}(D_{\lfloor \lambda n \rfloor - 1}(U))| \leq \mathrm{rk}(\mathrm{span}(D_{\lfloor \lambda n \rfloor - 1}(U))) = \mathrm{rk}(D_{\lfloor \lambda n \rfloor - 1}(U)) \leq \lambda n/2$, so there are at most $\lambda n/2$ elements $x \in B_{c'} \cap U$ with $\phi(x) \in \mathrm{span}(D_{\lfloor \lambda n \rfloor - 1}(U))$. Furthermore, recalling that there are at most $\lambda n/2$ colours $c \in [n]$ with $|B_c \cap U| \geq \lambda n/2$, there are at most $\lambda n/2$ elements $x \in B_{c'} \cap U$ with $|B_{c(\phi(x))} \cap U| \geq \lambda n/2$. Thus, since $|B_{c'} \cap U| > \lambda n$, there must indeed be an element $x \in B_{c'} \cap U$ with $\phi(x) \notin \mathrm{span}(D_{\lfloor \lambda n \rfloor - 1}(U))$ and $|B_{c(\phi(x))} \cap U| < \lambda n/2$.

Now, let us consider the family $\mathcal{T}^*$ of $\lfloor (1+\lambda)n \rfloor$ rainbow independent sets obtained from $\mathcal{T}$ when replacing $T \in \mathcal{T}$ by $T - \phi(x) + x$. Note that for the set $U^* = (B_1 \cup \cdots \cup B_n) \setminus E(\mathcal{T}^*)$ we then have $U^* = U - x + \phi(x)$. In particular, we obtain $B_{c'} \cap U^* = (B_{c'} \cap U) - x$ and $B_{c(\phi(x))} \cap U^* = (B_{c(\phi(x))} \cap U) + \phi(x)$ (and $B_c \cap U^* = B_c \cap U$ for all other colours c).

Using $\phi(x) \notin \mathrm{span}(D_{\lfloor \lambda n \rfloor - 1}(U))$, Lemma 6.7 applied to $U' = \{\phi(x)\}$, U, $k = \lfloor \lambda n \rfloor$ and $k' = \lfloor \lambda n \rfloor - 1$ yields

$$D_{\lfloor \lambda n \rfloor}(U^*) = D_{\lfloor \lambda n \rfloor}(U - x + \phi(x)) \subseteq D_{\lfloor \lambda n \rfloor}(U \cup \{\phi(x)\}) = D_{\lfloor \lambda n \rfloor}(U) = \emptyset.$$

Furthermore, we have

$$\sum_{c=1}^n |B_c \cap U|^2 - \sum_{c=1}^n |B_c \cap U^*|^2 = |B_{c'} \cap U|^2 - |B_{c'} \cap U^*|^2 - |B_{c(\phi(x))} \cap U^*|^2 + |B_{c(\phi(x))} \cap U|^2$$

$$= |B_{c'} \cap U|^2 - (|B_{c'} \cap U| - 1)^2 - (|B_{c(\phi(x))} \cap U| + 1)^2 + |B_{c(\phi(x))} \cap U|^2 > 0,$$

using $|B_{c(\phi(x))} \cap U| < \lambda n/2 < \lambda n - 1 < |B_{c'} \cap U| - 1$ and the general inequality $a^2 + b^2 > (a-1)^2 + (b+1)^2$ for all real numbers a, b with $a - b > 1$. In other words, we obtain $\sum_{c=1}^n |B_c \cap U^*|^2 < \sum_{c=1}^n |B_c \cap U|^2$, contradicting our choice of $\mathcal{T}$. □

22

Now, as $D_{\lfloor \lambda n \rfloor}(U) = \emptyset$, by Corollary 6.5, the set U can be decomposed into $\lfloor \lambda n \rfloor$ independent sets. Furthermore, by Claim 6.10, every colour appears at most $\lfloor \lambda n \rfloor$ times in U. Thus, by Theorem 6.6, the set $U = (B_1 \cup \cdots \cup B_n) \setminus E(\mathcal{T})$ can be decomposed into $2 \cdot \lfloor \lambda n \rfloor = 2 \cdot \lfloor \varepsilon n/3 \rfloor \leq (2\varepsilon/3)n$ rainbow independent sets. Together with the $\lfloor (1 + \lambda)n \rfloor \leq (1 + \varepsilon/3)n$ rainbow independent sets in the family $\mathcal{T}$, this forms a decomposition of $B_1 \cup \cdots \cup B_n$ into at most $(1 + \varepsilon)n$ rainbow independent sets. Each of these rainbow independent sets can greedily be extended to a rainbow basis. Thus, we obtain a covering of $B_1 \cup \cdots \cup B_n$ by at most $(1 + \varepsilon)n$ rainbow bases (i.e., transversal bases). $\qquad\square$

7 Packing: proof of Lemma 3.3 and Theorem 1.5

Let us start this section by showing how Theorem 1.5 can be deduced from Lemmas 3.2 and 3.3. The rest of this section will then be devoted to proving Lemma 3.3 (recall that we already proved Lemma 3.2 in Section 5).

Proof of Theorem 1.5. We may assume without loss of generality that $\varepsilon < 1/10$. Given $0 < \varepsilon < 1/10$, let $\sigma = \sigma(\varepsilon) > 0$ and $L = L(\varepsilon) > 0$ be as in Lemma 3.3. Now, choose $0 < \nu < 1/20$ such that $3\nu \cdot (\ln(1/\nu) + 1) < \sigma/L$, define $\eta = \varepsilon + \nu$, and assume that n is sufficiently large with respect to the other parameters.

Consider a rank-n matroid and n bases $B_1, \ldots, B_n$ of it (we may assume that $B_1, \ldots, B_n$ are disjoint by introducing additional copies of any element that appears in more than one of these bases). By restricting the matroid to the set $B_1 \cup \cdots \cup B_n$, and colouring each of the bases $B_1, \ldots, B_n$ with a different colour, we obtain a coloured matroid. Our goal is to show that there is a family of $\lceil (1 - \varepsilon)n \rceil$ disjoint rainbow bases in $B_1 \cup \cdots \cup B_n$.

Now, let $R \subseteq B_1 \cup \cdots \cup B_n$ be a set of elements drawn independently at random with probability η, then R satisfies the statements in Lemmas 3.2 and 3.3 with high probability. In particular, we can fix an outcome of R satisfying both of these statements.

By the statement in Lemma 3.2, we can find a family $\mathcal{T}$ of $\lceil (1 - \varepsilon)n \rceil$ disjoint rainbow independent sets in $(B_1 \cup \cdots \cup B_n) \setminus R$ such that $|E(\mathcal{T})| \geq (1 - \eta - \nu)n^2 > \frac{3}{4}n^2$ (note that a priori the statement in Lemma 3.2 gives such a family $\mathcal{T}$ consisting of $\lfloor (1 - \varepsilon)n \rfloor$ sets, but in case $\lfloor (1 - \varepsilon)n \rfloor \neq \lceil (1 - \varepsilon)n \rceil$, we can add a copy of the empty set to the family $\mathcal{T}$, or partition one of the sets $T \in \mathcal{T}$ into two subsets).

Let $m = \lceil (1 - \varepsilon)n \rceil \cdot n$ be the number of elements that a family of $\lceil (1 - \varepsilon)n \rceil$ disjoint rainbow bases would have. By our assumption that n is sufficiently large, we have $m \leq (1 - \varepsilon + \nu)n^2$ and therefore $m - |E(\mathcal{T})| \leq (1 - \varepsilon + \nu)n^2 - (1 - \eta - \nu)n^2 = (\eta - \varepsilon + 2\nu)n^2 = 3\nu n^2$. Furthermore, note that for every family $\mathcal{S}$ of $\lceil (1 - \varepsilon)n \rceil$ disjoint rainbow independent sets we have $|E(\mathcal{S})| \leq \lceil (1 - \varepsilon)n \rceil \cdot n = m$.

Now, let $\mathcal{S}$ be a family of $\lceil (1 - \varepsilon)n \rceil$ disjoint rainbow independent sets in $B_1 \cup \cdots \cup B_n$ maximising $|E(\mathcal{S})|$ subject to the constraint

$$|E(\mathcal{S}) \cap R| \leq \sum_{i=|E(\mathcal{T})|}^{|E(\mathcal{S})|-1} L \cdot \ln \left(\frac{n^2}{m - i} \right).$$

The family $\mathcal{T}$ satisfies this constraint (as $|E(\mathcal{T}) \cap R| = 0$), so $\mathcal{S}$ is well-defined and we have $|E(\mathcal{S})| \geq |E(\mathcal{T})| > \frac{3}{4}n^2$.

If $|E(\mathcal{S})| = m$, then $\mathcal{S}$ must be a family of $\lceil (1 - \varepsilon)n \rceil$ disjoint rainbow bases, finishing the proof of Theorem 1.5. Assume, then, for contradiction that $|E(\mathcal{S})| < m$. Note that then we have

$$|E(\mathcal{S}) \cap R| \leq \sum_{i=|E(\mathcal{T})|}^{|E(\mathcal{S})|-1} L \cdot \ln \left(\frac{n^2}{m - i} \right) = L \cdot \sum_{j=m-|E(\mathcal{S})|+1}^{m-|E(\mathcal{T})|} \ln \left(\frac{n^2}{j} \right) \leq L \cdot \sum_{j=2}^{\lfloor 3\nu n^2 \rfloor} \ln \left(\frac{n^2}{j} \right) \leq L \cdot \frac{\sigma n^2}{L} = \sigma n^2,$$

using that

$$\sum_{j=2}^{\lfloor 3\nu n^2 \rfloor} \ln \left(\frac{n^2}{j} \right) \leq \int_1^{3\nu n^2} \ln \left(\frac{n^2}{x} \right) \mathrm{d}\, x = \left[x \ln \left(\frac{n^2}{x} \right) + x \right]_{x=1}^{3\nu n^2} \leq 3\nu \ln(1/\nu) \cdot n^2 + 3\nu n^2 \leq \frac{\sigma n^2}{L}.$$

Thus, we can apply Lemma 3.3 to the family $\mathcal{S}$ (also recalling that $\frac{3}{4}n^2 < |E(\mathcal{S})| < m = \lceil(1-\varepsilon)n\rceil \cdot n$), and conclude that there is a family $\mathcal{S}'$ of $\lceil(1-\varepsilon)n\rceil$ disjoint rainbow independent sets in $B_1 \cup \cdots \cup B_n$ with $|E(\mathcal{S}')| > |E(\mathcal{S})|$ and

$$|E(\mathcal{S}') \cap R| \leq |E(\mathcal{S}) \cap R| + L \cdot \ln\left(\frac{n^2}{\lceil(1-\varepsilon)n\rceil \cdot n - |E(\mathcal{S})|}\right) = |E(\mathcal{S}) \cap R| + L \cdot \ln\left(\frac{n^2}{m - |E(\mathcal{S})|}\right)$$

$$\leq \sum_{i=|E(\mathcal{T})|}^{|E(\mathcal{S})|-1} L \cdot \ln\left(\frac{n^2}{m-i}\right) + L \cdot \ln\left(\frac{n^2}{m - |E(\mathcal{S})|}\right) \leq \sum_{i=|E(\mathcal{T})|}^{|E(\mathcal{S}')|-1} L \cdot \ln\left(\frac{n^2}{m-i}\right).$$

This contradicts our choice of the family $\mathcal{S}$. $\qquad\square$

It remains to prove Lemma 3.3. The following lemma characterises a property of the random set R needed in the proof (as we will show later, the statement in Lemma 3.3 holds for any subset $R \subseteq B_1 \cup \cdots \cup B_n$ satisfying property ($\spadesuit$) in the lemma below with γ chosen sufficiently small with respect to ε).

Lemma 7.1. *For any $0 < \eta < 1$ and $\gamma > 0$, the following holds for any coloured rank-n matroid with colour classes $B_1, \ldots, B_n$, such that each of the sets B_i for $i = 1, \ldots, n$ is a basis. Let $R \subseteq B_1 \cup \cdots \cup B_n$ be a set of elements drawn independently at random with probability η. Then, with high probability (more precisely, with probability tending to 1 as $n \to \infty$ with η and γ fixed), the following holds.*

($\spadesuit$) *For every independent set $T \subseteq B_1 \cup \cdots \cup B_n$, and every colour subset $C \subseteq [n]$, there are at least $\eta \cdot (n - |T|) \cdot |C| - \gamma n^2$ elements $e \in \bigcup_{c \in C}(B_c \cap R)$ such that $T + e$ is independent.*

Proof. The proof is completely analogous to the proof of Lemma 5.2. Alternatively, the lemma here can be deduced directly from the statement of Lemma 5.2 applied to the random set $(B_1 \cup \cdots \cup B_n) \setminus R$ (in which elements are drawn with probability $1 - \eta$). $\qquad\square$

The relevance of property ($\spadesuit$) is actually that it implies property ($\blacktriangle$) in the following lemma. Roughly speaking, this property states that, for any reasonably large colour subset $C \subseteq [n]$, the union $\bigcup_{c \in C}(B_c \cap R)$ has large rank, even after deleting up to γn^2 elements from R.

Lemma 7.2. *For $0 < \eta < 1$ and $\gamma, \varepsilon' > 0$ with $(\varepsilon')^2\eta > 2\gamma$, consider a coloured rank-n matroid with colour classes $B_1, \ldots, B_n$ and a subset $R \subseteq B_1 \cup \cdots \cup B_n$ satisfying ($\spadesuit$) above. Then the following holds.*

($\blacktriangle$) *For any subset $Q \subseteq R$ of size $|Q| \geq |R| - \gamma n^2$ and any colour subset $C \subseteq [n]$ of size $|C| \geq \varepsilon'n$ we have $\mathrm{rk}(\bigcup_{c \in C}(B_c \cap Q)) \geq (1 - \varepsilon')n$.*

Proof. Suppose for contradiction that $|\mathrm{rk}(\bigcup_{c \in C}(B_c \cap Q))| < (1-\varepsilon')n$, and let T be a maximal independent subset of $\bigcup_{c \in C}(B_c \cap Q)$. Then $|T| = |\mathrm{rk}(\bigcup_{c \in C}(B_c \cap Q))| < (1 - \varepsilon')n$, and so by ($\spadesuit$) there are at least $\eta \cdot (n - |T|) \cdot |C| - \gamma n^2 \geq \eta \cdot \varepsilon'n \cdot \varepsilon'n - \gamma n^2 > \gamma n^2$ elements $e \in \bigcup_{c \in C}(B_c \cap R)$ such that $T + e$ is independent. By the maximality of T, for each of these elements e, we have $e \notin \bigcup_{c \in C}(B_c \cap Q)$. Thus, there are more than γn^2 elements in $\bigcup_{c \in C}(B_c \cap R) \setminus \bigcup_{c \in C}(B_c \cap Q) \subseteq R \setminus Q$, and we must have $|R \setminus Q| > \gamma n^2$, contradicting our assumptions on Q. $\qquad\square$

We also need the following lemma, stating that if we can make two different switches to an independent set, then we can also combine these switches in certain ways.

Lemma 7.3. *Let T be an independent set in a matroid and let x, x', q, q' be elements of the matroid with $x, q \notin T$ and $x', q' \in T$, such that $T - x' + x$ and $T - q' + q$ are independent sets with $\mathrm{span}(T - x' + x) = \mathrm{span}(T - q' + q) = \mathrm{span}(T)$. Then at least one of the following statements holds.*

(a) *The set $T - q' + x$ is independent with $\mathrm{span}(T - q' + x) = \mathrm{span}(T)$.*

(b) *We have $x' \neq q'$, and the set $T - x' - q' + x + q$ is independent with $\mathrm{span}(T - x' - q' + x + q) = \mathrm{span}(T)$.*

Proof. If $x' = q'$, then $T - q' + x = T - x' + x$ and so (a) holds by assumption. So let us assume that $x' \neq q'$. Now, note that $x \in \mathrm{span}(T - x' + x) = \mathrm{span}(T)$, and $q \in \mathrm{span}(T - q' + q) = \mathrm{span}(T)$, so we can conclude that $\mathrm{span}(T - q' + x) \subseteq \mathrm{span}(T)$ and $\mathrm{span}(T - x' - q' + x + q) \subseteq \mathrm{span}(T)$. Since we also have $|T - q' + x| = |T - x' - q' + x + q| = |T|$, it suffices to show that the set $T - q' + x$ is independent or the set $T - x' - q' + x + q$ is independent.

Note that the set $T - x' - q' + x \subseteq T - x' + x$ is independent and has size $|T - x' - q' + x| = |T| - 1$, which is strictly smaller than the size $|T - q' + q| = |T|$ of the independent set $T - q' + q$. Thus, there exists an element $e \in (T - q' + q) \setminus (T - x' - q' + x) = \{q, x'\}$ such that $T - x' - q' + x + e$ is an independent set. If $e = q$, this means $T - x' - q' + x + q$ is independent, and if $e = x'$, this means that $T - q' + x$ is independent. $\square$

The following definition is crucial for our proof of Lemma 3.3. Roughly speaking, for distinct members $T_1, \ldots, T_r$ of some family $\mathcal{T}$ of rainbow independent sets, it describes the situation that an element e outside $T_1 \cup \cdots \cup T_r$ can be absorbed into $T_1, \ldots, T_r$, by slightly modifying the sets $T_1, \ldots, T_r$ in such a way that the total size $|T_1| + \cdots + |T_r|$ increases (but without touching any elements in $E(\mathcal{T}) \setminus (T_1 \cup \cdots \cup T_r)$ except e). See also the outline in Section 3.2 for the motivation behind this definition. Our definition here is inspired by the notion of cascade-addability in [5, Definition 2.8]. However, our notion of absorbability here is more flexible, and this additional flexibility is crucial in our proof of Lemma 3.3 (the switching operations considered for cascade-addability in [5, Definition 2.8] are not strong enough for our purposes).

Definition 7.4. Consider a family of disjoint rainbow independent sets $\mathcal{T}$ in a coloured matroid, and let U be the set of matroid elements outside of $E(\mathcal{T})$ (i.e., the uncovered set). For distinct $T_1, \ldots, T_r \in \mathcal{T}$, we say that an element e of the matroid with $e \notin T_1 \cup \cdots \cup T_r$ is $(T_1, \ldots, T_r)$-*absorbable with respect to* $\mathcal{T}$ if there are disjoint rainbow independent sets $T_1', \ldots, T_r' \subseteq U \cup T_1 \cup \cdots \cup T_r \cup \{e\}$ such that $|T_i' \setminus T_i| \leq 3$ for $i = 1, \ldots r$, and $|T_1'| + \cdots + |T_r'| \geq |T_1| + \cdots + |T_r| + 1$.

Furthermore, for a set $T_{r+1} \in \mathcal{T} \setminus \{T_1, \ldots, T_r\}$, let us denote by $m_{\mathcal{T}}(T_1, \ldots, T_r, T_{r+1})$ the number of elements $e \in T_{r+1}$ such that e is $(T_1, \ldots, T_r)$-absorbable with respect to $\mathcal{T}$.

Note that in Definition 7.4 we do not demand that $e \in T_1' \cup \cdots \cup T_r'$. However, in the setting of the proof of Lemma 3.3 we will automatically have that $e \in T_1' \cup \cdots \cup T_r'$. Indeed, if we have $e \notin T_1' \cup \cdots \cup T_r'$, then we can obtain a family of disjoint rainbow independent sets from $\mathcal{T}$ with strictly larger total size, simply replacing $T_1, \ldots, T_r$ by $T_1', \ldots, T_r'$. In our proof of Lemma 3.3, this would trivially give the desired family $\mathcal{S}'$. So we may assume that $e \in T_1' \cup \cdots \cup T_r'$ (then $T_1', \ldots, T_r'$ are not disjoint from the sets in $\mathcal{T} \setminus \{T_1, \ldots, T_r\}$), which motivates the word "absorbable".

The following lemma states that under certain assumptions for a given $T \in \mathcal{T}$ we can find many (T)-absorbable elements (i.e. many elements satisfying the condition in Definition 7.4 for $r = 1$ and $T_1 = T$).

Lemma 7.5. *Let $\gamma > 0$ and $\varepsilon' > 0$. Consider a family $\mathcal{T}$ of at most $(1 - \varepsilon')n$ rainbow independent sets in a coloured rank-n matroid with colour classes $B_1, \ldots, B_n$, such that each of the sets B_i for $i = 1, \ldots, n$ is a basis, and let $U = (B_1 \cup \cdots \cup B_n) \setminus E(\mathcal{T})$. Let $R \subseteq B_1 \cup \cdots \cup B_n$ be a subset satisfying ($\blacktriangle$), and assume that $|E(\mathcal{T}) \cap R| \leq \gamma n^2$. Let $T \in \mathcal{T}$ have size $|T| < n$, and assume that there is no rainbow independent set $S \subseteq U \cup T$ with $|S| > |T|$ and $|S \setminus T| \leq 2$.*

Then there is a set of colours $C \subseteq [n]$ of size $|C| \geq (1 - \varepsilon')n$, such that the following condition holds: For every colour $c \in C$, every element in $B_c \setminus \mathrm{span}(T)$ is (T)-absorbable with respect to $\mathcal{T}$.

In other words, the condition on the set C here states that every element e such that $T + e$ is independent and e has a colour in C must be (T)-absorbable with respect to $\mathcal{T}$. The proof of this lemma relies on switching arguments like in the proof outline in Section 3.2 (see also Figure 2).

Proof of Lemma 7.5. Since $|T| < n$, we can fix a colour $c^* \in [n]$ not appearing on T. Note that we have $B_{c^*} \cap E(\mathcal{T}) \leq |\mathcal{T}| \leq (1 - \varepsilon')n$ and therefore $|B_{c^*} \cap U| \geq \varepsilon' n$.

Let us now apply Lemma 2.3 to the independent sets $B_{c^*} \cap U$ and T. If option (a) there holds, there would be an element $x \in B_{c^*} \cap U$ such that $T + x$ is independent. But then $T + x \subseteq U \cup T$ is a rainbow independent set with $|T + x| > |T|$ and $|(T + x) \setminus T| < 2$. This would be a contradiction to our assumption.

Thus, we must have option (b) in Lemma 2.3, so there is an injection $\phi^* : B_{c^*} \cap U \to T$ such that, for each $x \in B_{c^*} \cap U$, the set $T - \phi^*(x) + x$ is independent and $\mathrm{span}(T - \phi^*(x) + x) = \mathrm{span}(T)$. Note

that the set $T - \phi^*(x) + x$ is also rainbow, since the colour c^* of x does not appear on T, and we have $T - \phi^*(x) + x \subseteq U \cup T$.

Now, let $C^* \subseteq [n]$ be the set of colours appearing on the elements in the image of ϕ^*, i.e., the elements $\phi^*(x)$ for $x \in B_{c^*} \cap U$. Then $|C^*| = |B_{c^*} \cap U| \geq \varepsilon'n$. Furthermore, taking $Q = R \cap U$, note that $|Q| = |R| - |E(\mathcal{T}) \cap R| \geq |R| - \gamma n^2$. Thus, by (▲) we have $\mathrm{rk}(\bigcup_{c \in C^*}(B_c \cap Q)) \geq (1 - \varepsilon')n$. So we can find an independent set $Q^* \subseteq \bigcup_{c \in C^*}(B_c \cap Q)$ of size $|Q^*| \geq (1 - \varepsilon')n$. By definition of C^*, for every element $q \in Q^*$ there exists an element $x \in B_{c^*} \cap U$ such that $\phi^*(x)$ has the same colour as q. Note that then the set $T - \phi^*(x) + x + q$ is rainbow.

Now, let us again apply Lemma 2.3, this time to the independent sets Q^* and T. If option (a) holds, there is some $q \in Q^* \subseteq Q \subseteq U$ such that the set $T + q$ is independent. This means that $q \notin \mathrm{span}(T) = \mathrm{span}(T - \phi^*(x) + x)$, where $x \in B_{c^*} \cap U$ is such that $\phi^*(x)$ has the same colour as q. But then $T - \phi^*(x) + x + q \subseteq U \cup T$ is a rainbow independent set of size $|T| + 1$ with $|(T - \phi^*(x) + x + q) \setminus T| \leq 2$. This again contradicts our assumption.

Thus, again option (b) in Lemma 2.3 must hold. So we obtain an injection $\phi' : Q^* \to T$ such that, for each $q \in Q^*$, the set $T - \phi'(q) + q$ is independent and $\mathrm{span}(T - \phi'(q) + q) = \mathrm{span}(T)$. Finally, let $C \subseteq [n]$ be the set of colours appearing on the image $\phi'(Q^*) \subseteq T$. Then we indeed have $|C| = |\phi'(Q^*)| = |Q^*| \geq (1 - \varepsilon')n$.

It remains to show that, for each $c \in C$, every element in $B_c \setminus \mathrm{span}(T)$ is (T)-absorbable with respect to $\mathcal{T}$. So let $c \in C$, then there exists an element $q \in Q^* \subseteq U$ such that its image $\phi'(q)$ has colour c. Let $x \in B_{c^*} \cap U$ be such that $\phi^*(x)$ has the same colour as q. Recall that $T - \phi^*(x) + x$ is an independent set with $\mathrm{span}(T - \phi^*(x) + x) = \mathrm{span}(T)$, and that $T - \phi'(q) + q$ is an independent set with $\mathrm{span}(T - \phi'(q) + q) = \mathrm{span}(T)$. Thus, by Lemma 7.3, we have that $T - \phi'(q) + x$ is an independent set with $\mathrm{span}(T - \phi'(q) + x) = \mathrm{span}(T)$, or that $T - \phi^*(x) - \phi'(q) + x + q$ is an independent set with $\mathrm{span}(T - \phi^*(x) - \phi'(q) + x + q) = \mathrm{span}(T)$. Noting that both of these sets are rainbow, in either case we find a rainbow independent set $T' \subseteq U \cup T$ of size $|T'| = |T|$ with $\mathrm{span}(T') = \mathrm{span}(T)$ and $|T' \setminus T| \leq 2$, such that the colour $c = c(\phi'(q))$ does not appear on T'. Now, for each $e \in B_c \setminus \mathrm{span}(T)$, note that $T' + e$ is a rainbow independent set (since $e \notin \mathrm{span}(T) = \mathrm{span}(T')$ and the colour c of e does not appear on T'). So e is (T)-absorbable with respect to $\mathcal{T}$, since we can take $T_1' = T' + e$ in Definition 7.4. □

Finally, we are ready to prove Lemma 3.3. As outlined in Section 3.2, the proof strategy is to find a cascade of sets $T_1, \ldots, T_{r_{\max}} \in \mathcal{T}$ such that the number $m_{\mathcal{T}}(T_1, \ldots, T_r, T_{r+1})$ of $(T_1, \ldots, T_r)$-absorbable elements in T_{r+1} grows exponentially with r. Here, $\mathcal{T}$ is either identical to the original family $\mathcal{S}$ given in Lemma 3.3, or $\mathcal{T}$ is a family of $\lceil(1 - \varepsilon)n\rceil$ rainbow independent sets of the same total size $|E(\mathcal{T})| = |E(\mathcal{S})|$ as $\mathcal{S}$, where $|E(\mathcal{T}) \cap R|$ is not much larger than $|E(\mathcal{S}) \cap R|$, and which contains sets $T_1, T_2 \in \mathcal{T}$ such that $m_{\mathcal{T}}(T_1, T_2)$ is very large (i.e., such that T_2 contains a lot of (T_1)–absorbable elements). Finding such $T_1, T_2 \in \mathcal{T}$ with very large $m_{\mathcal{T}}(T_1, T_2)$ is very helpful for starting our cascade with growing $m_{\mathcal{T}}(T_1, \ldots, T_r, T_{r+1})$. If no family $\mathcal{T}$ with these conditions containing such $T_1, T_2 \in \mathcal{T}$ exists, then we simply take $\mathcal{T} = \mathcal{S}$. Then at every step, we will have a set $T_{r+1} \in \mathcal{T} = \mathcal{S}$ containing many $(T_1, \ldots, T_r)$-absorbable elements. As outlined in Section 3.2, we can then make some small modifications to T_{r+1} to incorporate many different possible new elements, in order to show that there are many $(T_1, \ldots, T_r, T_{r+1})$-absorbable elements. However, if most of these possible new elements to be added to T_{r+1} (after small modifications) lie in $T_1 \cup \cdots \cup T_r$, they are by definition not $(T_1, \ldots, T_r, T_{r+1})$-absorbable, so this would be problematic. In the case where we take $\mathcal{T} = \mathcal{S}$, this cannot happen because otherwise one of the sets $T_1, \ldots, T_r$ would contain many (T'_{r+1})-absorbable elements (where T'_{r+1} denotes the modified version of T_{r+1}), and this would allow us to choose $\mathcal{T}$ differently.

Proof of Lemma 3.3. Let $0 < \varepsilon < 1/10$, and note that then $1 + \varepsilon/4 \geq e^{\varepsilon/8}$ (since the inequality $1 + 2x \geq e^x$ holds for all $x \in [0, 1]$). Define $\sigma = \varepsilon^3/20$ and $L = 10^7/\varepsilon^5$. As in the statement of Lemma 3.3, let $\varepsilon \leq \eta < 1$, assume that n is large with respect to ε and η, and consider a coloured rank-n matroid where the colour classes $B_1, \ldots, B_n$ are bases. Given Lemma 7.1, it suffices to prove that the desired conclusion in Lemma 3.3 holds for every subset $R \subseteq B_1 \cup \cdots \cup B_n$ satisfying (♠) with $\gamma = 2\sigma = \varepsilon^3/10$. Now, defining $\varepsilon' = \varepsilon/2$, note that $(\varepsilon')^2\eta \geq \varepsilon^3/4 > 2\gamma$, and so (▲) holds by Lemma 7.2.

Finally, as in the statement of Lemma 3.3, let $\mathcal{S}$ be a family of $\lceil(1 - \varepsilon)n\rceil$ rainbow independent sets with $\frac{3}{4}n^2 \leq |E(\mathcal{S})| < \lceil(1 - \varepsilon)n\rceil \cdot n$ and $|E(\mathcal{S}) \cap R| \leq \sigma n^2$. Let $s = \lceil(1 - \varepsilon)n\rceil \cdot n - |E(\mathcal{S})| \geq 1$, and assume for

contradiction that there is no family $\mathcal{S}'$ of $\lceil(1-\varepsilon)n\rceil$ rainbow independent sets such that $|E(\mathcal{S}')| > |E(\mathcal{S})|$ and $|E(\mathcal{S}') \cap R| \leq |E(\mathcal{S}) \cap R| + L \cdot \ln(n^2/s)$.

Let $\mathcal{T}$ be a family of $\lceil(1-\varepsilon)n\rceil$ rainbow independent sets with $|E(\mathcal{T})| = |E(\mathcal{S})|$ and $|E(\mathcal{T}) \cap R| \leq |E(\mathcal{S}) \cap R| + (L/2) \cdot \ln(n^2/s)$ such that there are distinct $T_1, T_2 \in \mathcal{T}$ with $m_{\mathcal{T}}(T_1, T_2) \geq \varepsilon^4 n/10^5$, if such a family $\mathcal{T}$ exists. In this case, also define $r_{\min} = 2$. If no such family exists, then define instead $\mathcal{T} = \mathcal{S}$ and $r_{\min} = 1$, and choose $T_1 \in \mathcal{T} = \mathcal{S}$ such that $n - |T_1|$ is maximal. Note that $\sum_{T \in \mathcal{S}}(n - |T|) = \lceil(1-\varepsilon)n\rceil \cdot n - |E(\mathcal{S})| = s$, so in this case we have $n - |T_1| \geq s/n > 0$.

Note that in either case we have $|E(\mathcal{T})| = |E(\mathcal{S})|$ and $|E(\mathcal{T}) \cap R| \leq |E(\mathcal{S}) \cap R| + (L/2) \cdot \ln(n^2/s)$. Let $U = (B_1 \cup \cdots \cup B_n) \setminus E(\mathcal{T})$ and $r_{\max} = \lceil(L/8) \cdot \ln(n^2/s)\rceil + 2$. Note that $s \leq n^2 - |E(\mathcal{S})| \leq n^2/4$, and hence $\ln(n^2/s) \geq \ln(4) > 1$ and $r_{\max} - 2 \geq L/8 > 10^6/\varepsilon^5$.

Now, for any $r_{\min} \leq r < r_{\max}$, let us recursively define $T_{r+1} \in \mathcal{T} \setminus \{T_1, \ldots, T_r\}$ to be the set maximising $m_{\mathcal{T}}(T_1, \ldots, T_r, T_{r+1})$. Our goal is to show that $m_{\mathcal{T}}(T_1, \ldots, T_r, T_{r+1})$ must grow very quickly with r, contradicting the fact that we have $m_{\mathcal{T}}(T_1, \ldots, T_r, T_{r+1}) \leq |T_{r+1}| \leq n$ for all $r_{\min} \leq r < r_{\max}$.

More specifically, we claim that for any $r_{\min} \leq r < r_{\max}$ we have

$$m_{\mathcal{T}}(T_1, \ldots, T_r, T_{r+1}) \geq \left(1 + \frac{\varepsilon}{4}\right) \cdot \left(m_{\mathcal{T}}(T_1, \ldots, T_{r-1}, T_r) + n - |T_r|\right). \tag{13}$$

(where in the case $r_{\min} = 1$ we define $m_{\mathcal{T}}(T_1) = 0$). Noting that $m_{\mathcal{T}}(T_1, \ldots, T_{r_{\max}-1}, T_{r_{\max}}) \leq |T_{r_{\max}}| \leq n$, this is sufficient to get a contradiction. Indeed, in the case $r_{\min} = 2$, (13) implies

$$m_{\mathcal{T}}(T_1, \ldots, T_{r_{\max}}) \geq \left(1 + \frac{\varepsilon}{4}\right)^{r_{\max}-2} \cdot m_{\mathcal{T}}(T_1, T_2) \geq (e^{\varepsilon/8})^{10^6/\varepsilon^5} \cdot \frac{\varepsilon^4 n}{10^5} > e^{10^5/\varepsilon^4} \cdot \frac{\varepsilon^4 n}{10^5} > \frac{10^5}{\varepsilon^4} \cdot \frac{\varepsilon^4 n}{10^5} = n,$$

and in the case $r_{\min} = 1$, (13) implies

$$m_{\mathcal{T}}(T_1, \ldots, T_{r_{\max}}) \geq \left(1 + \frac{\varepsilon}{4}\right)^{r_{\max}-1} \cdot (n - |T_1|) \geq (e^{\varepsilon/8})^{(10^6/\varepsilon^5) \cdot \ln(n^2/s)} \cdot \frac{s}{n} > \frac{n^2}{s} \cdot \frac{s}{n} = n.$$

Thus, it suffices to prove (13).

We show (13) by induction on r. So, for the rest of this proof, take some $r \in \{r_{\min}, \ldots, r_{\max} - 1\}$ and assume (13) holds for all smaller values of r in $\{r_{\min}, \ldots, r_{\max} - 1\}$. Then, if $r \geq 2$, we in particular have

$$m_{\mathcal{T}}(T_1, \ldots, T_r) \geq \left(1 + \frac{\varepsilon}{4}\right) \cdot m_{\mathcal{T}}(T_1, \ldots, T_{r-1}) \geq \cdots \geq \left(1 + \frac{\varepsilon}{4}\right)^{r-2} m_{\mathcal{T}}(T_1, T_2).$$

So in the case $r_{\min} = 2$, we conclude $m_{\mathcal{T}}(T_1, \ldots, T_r) \geq m_{\mathcal{T}}(T_1, T_2) \geq \varepsilon^4 n/10^5$, and in the case $r_{\min} = 1$ we conclude $m_{\mathcal{T}}(T_1, \ldots, T_r) \geq (1 + \varepsilon/4)^{r-2} m_{\mathcal{T}}(T_1, T_2) \geq (1 + \varepsilon/4)^{r-1}(n - |T_1|) \geq (1 + \varepsilon/4)^{r-1}$ if $r \geq 2$.

In order to show (13), we will show that there are many elements $e' \in E(\mathcal{T}) \setminus (T_1 \cup \cdots \cup T_r)$ which are $(T_1, \ldots, T_r)$-absorbable with respect to $\mathcal{T}$. In the special case $r = 1$, this is shown by the following relatively easy claim (in contrast, in the case $r \geq 2$, this is only shown in Claim 7.9 and requires significantly more arguments).

Claim 7.6. If $r = 1$, then there are at least $(1 - \varepsilon/2)n \cdot (n - |T_1|)$ elements $e' \in E(\mathcal{T}) \setminus T_1$ which are (T_1)-absorbable with respect to $\mathcal{T}$.

Proof. Recall that $|E(\mathcal{T}) \cap R| \leq |E(\mathcal{S}) \cap R| + (L/2) \cdot \ln(n^2/s)$ and $|E(\mathcal{S}) \cap R| \leq \sigma n^2$, so $|E(\mathcal{T}) \cap R| \leq 2\sigma n^2 = \gamma n^2$. Furthermore, note that $|T_1| < n$ (since $r = 1$ implies $r_{\min} = 1$) and that there cannot be a rainbow independent set $S \subseteq U \cup T_1$ with $|S| > |T_1|$ and $|S \setminus T_1| \leq 3$ (since otherwise taking $\mathcal{S}'$ to be the family obtained from $\mathcal{T}$ when replacing $T_1 \in \mathcal{T}$ by S would contradict our assumption). Therefore, for any element e' which is (T_1)-absorbable with respect to $\mathcal{T}$, we must have $e \in E(\mathcal{T}) \setminus T_1$. Furthermore, by Lemma 7.5 there is a set of colours $C \subseteq [n]$ of size $|C| \geq (1 - \varepsilon')n = (1 - \varepsilon/2)n$, such that, for each colour $c \in C$, every element in $B_c \setminus \text{span}(T_1)$ is (T_1)-absorbable with respect to $\mathcal{T}$. Note that, for every colour $c \in C$, we have $|B_c \cap \text{span}(T_1)| \leq \text{rk}(\text{span}(T_1)) = |T_1|$ and hence $|B_c \setminus \text{span}(T_1)| = n - |B_c \cap \text{span}(T_1)| \geq n - |T_1|$. Thus in total there are at least $|C| \cdot (n - |T_1|) \geq (1 - \varepsilon/2)n \cdot (n - |T_1|)$ elements $e' \in E(\mathcal{T}) \setminus T_1$ which are (T_1)-absorbable with respect to $\mathcal{T}$. ◻

Our goal is to show that also in the case $r \geq 2$ there are many $(T_1, \ldots, T_r)$-absorbable elements.

Let $E^* \subseteq T_r$ be the set of elements $e \in T_r$ which are $(T_1, \ldots, T_{r-1})$-absorbable with respect to $\mathcal{T}$, and note that $|E^*| = m_\mathcal{T}(T_1, \ldots, T_r)$. For each $e \in E^*$, we can find a collection of disjoint rainbow independent sets $T_1^{(e)}, \ldots, T_{r-1}^{(e)} \subseteq U \cup T_1 \cup \cdots \cup T_{r-1} \cup \{e\}$ such that $|T_i^{(e)} \setminus T_i| \leq 3$ for $i = 1, \ldots r-1$, and $|T_1^{(e)}| + \cdots + |T_{r-1}^{(e)}| \geq |T_1| + \cdots + |T_{r-1}| + 1$. Define $\mathcal{T}^{(e)}$ to be the family of $\lceil (1-\varepsilon)n \rceil$ disjoint rainbow independent sets obtained from $\mathcal{T}$ when replacing $T_1, \ldots, T_r$ by $T_1^{(e)}, \ldots, T_{r-1}^{(e)}, T_r - e$. Note that

$$|E(\mathcal{T}^{(e)})| - |E(\mathcal{T})| = |T_1^{(e)}| + \ldots, +|T_{r-1}^{(e)}| + (|T_r| - 1) - (|T_1| + \cdots + |T_r|) \geq 0$$

and

$$|E(\mathcal{T}^{(e)}) \cap R| \leq |E(\mathcal{T}) \cap R| + \sum_{i=1}^{r-1} |T_i^{(e)} \setminus T_i| \leq |E(\mathcal{T}) \cap R| + 3r \leq |E(\mathcal{T}) \cap R| + (L/2) \cdot \ln(n^2/s) - 3$$

(as $3r \leq 3r_{\max} \leq (3/8)L \cdot \ln(n^2/s) + 9 \leq (L/2) \cdot \ln(n^2/s) - 3$). Thus, we in particular have $|E(\mathcal{T}^{(e)})| \geq |E(\mathcal{T})| = |E(\mathcal{S})|$ and $|E(\mathcal{T}^{(e)}) \cap R| \leq |E(\mathcal{S}) \cap R| + L \cdot \ln(n^2/s) - 3$ for all $e \in E^*$.

For each $e \in E^*$, let $U^{(e)} = (B_1 \cup \cdots \cup B_n) \setminus E(\mathcal{T}^{(e)})$ be the set of elements uncovered by $\mathcal{T}^{(e)}$, and note that $U^{(e)} \subseteq U \cup T_1 \cup \cdots \cup T_r$. Recalling that $|E(\mathcal{S}) \cap R| \leq \sigma n^2$, we can also observe that $|E(\mathcal{T}^{(e)}) \cap R| \leq 2\sigma n^2 = \gamma n^2$. Furthermore $|T_r - e| < n$, and there cannot be a rainbow independent set $S \subseteq U^{(e)} \cup (T_r - e)$ with $|S| > |T_r - e|$ and $|S \setminus (T_r - e)| \leq 2$ (since otherwise taking $\mathcal{S}'$ to be the family obtained from $\mathcal{T}^{(e)}$ when replacing $T_r - e \in \mathcal{T}^{(e)}$ by S would yield a contradiction to our assumption). Thus, by Lemma 7.5, there is a set of colours $C^{(e)} \subseteq [n]$ of size $|C^{(e)}| \geq (1 - \varepsilon')n$, such that, for each colour $c \in C^{(e)}$, every element in $B_c \setminus \mathrm{span}(T_r - e)$ is $(T_r - e)$-absorbable with respect to $\mathcal{T}^{(e)}$.

The following claim shows, roughly speaking, that in order to get many $(T_1, \ldots, T_r)$-absorbable elements, it suffices to find many elements which are $(T_r - e)$-absorbable with respect to $\mathcal{T}^{(e)}$ for some $e \in E^*$.

Claim 7.7. Let $e \in E^*$, and let e' be $(T_r - e)$-absorbable with respect to $\mathcal{T}^{(e)}$. Then, unless $e' \in T_1^{(e)} \cup \cdots \cup T_{r-1}^{(e)}$, the element e' is $(T_1, \ldots, T_r)$-absorbable with respect to $\mathcal{T}$ and we have $e' \in E(\mathcal{T}) \setminus (T_1 \cup \cdots \cup T_r)$.

Proof. By assumption, there exists a rainbow independent set $T_r' \subseteq U^{(e)} \cup (T_r - e) \cup \{e'\}$ with $|T_r' \setminus (T_r - e)| \leq 3$ and $|T_r'| \geq |T_r - e| + 1 = |T_r|$. We claim that $e' \notin U^{(e)}$. Indeed, if $e' \in U^{(e)}$, then the family $\mathcal{S}'$ obtained from $\mathcal{T}^{(e)}$ when replacing $T_r - e \in \mathcal{T}^{(e)}$ by T_r' would be a family of $\lceil (1-\varepsilon)n \rceil$ disjoint rainbow independent sets with $|E(\mathcal{S}')| > |E(\mathcal{T}^{(e)})| \geq |E(\mathcal{S})|$ and

$$|E(\mathcal{S}') \cap R| \leq |E(\mathcal{T}^{(e)}) \cap R| + 3 \leq |E(\mathcal{S}) \cap R| + L \cdot \ln(n^2/s)$$

(noting that $|E(\mathcal{S}') \cap R| - |E(\mathcal{T}^{(e)}) \cap R| = |T_r' \cap R| - |(T_r - e) \cap R| \leq |T_r' \setminus (T_r - e)| \leq 3$). This would be a contradiction to our assumption in the beginning that no such family $\mathcal{S}'$ exists.

So indeed $e' \notin U^{(e)}$, which means that $e' \in E(\mathcal{T}^{(e)}) = T_1^{(e)} \cup \cdots \cup T_{r-1}^{(e)} \cup (T_r - e) \cup E(\mathcal{T} \setminus \{T_1, \ldots, T_r\})$. Since e' is $(T_r - e)$-absorbable, we have $e' \notin T_r - e$ (see Definition 7.4). Thus, unless $e' \in T_1^{(e)} \cup \cdots \cup T_{r-1}^{(e)}$, we therefore have $e' \in E(\mathcal{T} \setminus \{T_1, \ldots, T_r\}) = E(\mathcal{T}) \setminus (T_1 \cup \cdots \cup T_r)$.

It remains to show that e' is $(T_1, \ldots, T_r)$-absorbable with respect to $\mathcal{T}$ if $e' \notin T_1^{(e)} \cup \cdots \cup T_{r-1}^{(e)}$. To this end, note that $T_1^{(e)}, \ldots, T_{r-1}^{(e)}, T_r' \subseteq U \cup U^{(e)} \cup T_1 \cup \cdots \cup T_r \cup \{e'\} = U \cup T_1 \cup \cdots \cup T_r \cup \{e'\}$ are disjoint rainbow independent sets with $|T_1^{(e)}| + \cdots + |T_{r-1}^{(e)}| + |T_r'| \geq |T_1| + \cdots + |T_{r-1}| + 1 + |T_r|$ and $|T_r' \setminus T_r| \leq |T_r' \setminus (T_r - e)| \leq 3$ (as well as $|T_i^{(e)} \setminus T_i| \leq 3$ for $i = 1, \ldots r-1$). Thus, by Definition 7.4, the element e' is $(T_1, \ldots, T_r)$-absorbable with respect to $\mathcal{T}$ if $e' \notin T_1^{(e)} \cup \cdots \cup T_{r-1}^{(e)}$. $\boxdot$

Note that we obtain the conclusion in Claim 7.7 only for $(T_r - e)$-absorbable elements e' with $e' \notin T_1^{(e)} \cup \cdots \cup T_{r-1}^{(e)}$. Our next claim gives an upper bound for the number of $(T_r - e)$-absorbable elements e' with $e' \in T_1^{(e)} \cup \cdots \cup T_{r-1}^{(e)}$ if $r_{\min} = 1$.

Claim 7.8. Let $e \in E^*$, and assume that $r_{\min} = 1$. Then, for every $i = 1, \ldots, r-1$, there are at most $\varepsilon^4 n/10^5$ elements $e' \in T_i^{(e)}$ which are $(T_r - e)$-absorbable with respect to $\mathcal{T}^{(e)}$.

28

Proof. Suppose for contradiction that for some $i \in \{1, \ldots, r-1\}$ there are more than $\varepsilon^4 n/10^5$ elements $e' \in T_i^{(e)}$ which are $(T_r - e)$-absorbable with respect to $\mathcal{T}^{(e)}$. This means that $m_{\mathcal{T}^{(e)}}(T_r - e, T_i^{(e)}) > \varepsilon^4 n/10^5$.

Now, recall that $\mathcal{T}^{(e)}$ is a family of $\lceil(1-\varepsilon)n\rceil$ disjoint rainbow independent sets with $|E(\mathcal{T}^{(e)})| \geq |E(\mathcal{S})|$ and $|E(\mathcal{T}^{(e)}) \cap R| \leq |E(\mathcal{T}) \cap R| + (L/2) \cdot \ln(n^2/s) = |E(\mathcal{S}) \cap R| + (L/2) \cdot \ln(n^2/s)$ (where in the last step we used that $\mathcal{T} = \mathcal{S}$ by our assumption $r_{\min} = 1$).

Thus, if $|E(\mathcal{T}^{(e)})| > |E(\mathcal{S})|$, this contradicts our assumption that there is no family $\mathcal{S}'$ of $\lceil(1-\varepsilon)n\rceil$ disjoint rainbow independent sets with $|E(\mathcal{S}')| > |E(\mathcal{S})|$ and $|E(\mathcal{S}') \cap R| \leq |E(\mathcal{S}) \cap R| + L \cdot \ln(n^2/s)$. On the other hand, if $|E(\mathcal{T}^{(e)})| = |E(\mathcal{S})|$, we obtain a contradiction to $r_{\min} = 1$ and our choice of $\mathcal{T}$ (since we should have taken $\mathcal{T}$ to be $\mathcal{T}^{(e)}$, making $r_{\min} = 2$). $\boxdot$

Combining the two previous claims, we can now give the following lower bound for the number of $(T_1, \ldots, T_r)$-absorbable elements.

Claim 7.9. There are at least $(1 - 3\varepsilon/4)n \cdot (m_{\mathcal{T}}(T_1, \ldots, T_r) + n - |T_r|)$ elements $e' \in E(\mathcal{T}) \setminus (T_1 \cup \cdots \cup T_r)$ which are $(T_1, \ldots, T_r)$-absorbable with respect to $\mathcal{T}$.

Proof. In the case $r = 1$, the desired statement follows from Claim 7.6 (recalling that $m_{\mathcal{T}}(T_1) = 0$). So let us assume that $r \geq 2$, which in particular implies $|E^*| = m_{\mathcal{T}}(T_1, \ldots, T_r) \geq 1$.

Now let E' be the set of elements $e' \in B_1 \cup \cdots \cup B_n$ for which there is some $e \in E^*$ so that e' is $(T_r - e)$-absorbable with respect to $\mathcal{T}^{(e)}$. Furthermore, let E'' be the set of elements $e' \in B_1 \cup \cdots \cup B_n$ for which there is some $e \in E^*$ so that e' is $(T_r - e)$-absorbable with respect to $\mathcal{T}^{(e)}$ and $e' \in T_1^{(e)} \cup \cdots \cup T_{r-1}^{(e)}$. Note that $E'' \subseteq E'$. By Claim 7.7, every element $e' \in E' \setminus E''$ is $(T_1, \ldots, T_r)$-absorbable with respect to $\mathcal{T}$ and satisfies $e' \in E(\mathcal{T}) \setminus (T_1 \cup \cdots \cup T_r)$. Thus it suffices to prove that $|E' \setminus E''| \geq (1 - 3\varepsilon/4) \cdot (m_{\mathcal{T}}(T_1, \ldots, T_r) + n - |T_r|)$.

Let us now show a lower bound for $|E'|$ (afterwards, we use Claim 7.8 to show an upper bound for $|E''|$). For every colour $c \in [n]$, let us apply Lemma 2.2 to the independent set T_r and the basis B_c, to find an injection $\phi_c : T_r \to B_c$ such that, for each $e \in T_r$, the set $T_r - e + \phi_c(e)$ is independent, and for each $b \in B_c \setminus \phi_c(T_r)$, the set $T_r + b$ is independent (and hence also the set $T_r - e + b$ for each $e \in T_r$). So, for any $c \in [n]$ we have $\phi_c(e) \in B_c \setminus \mathrm{span}(T_r - e)$ and $B_c \setminus \phi_c(T_r) \subseteq B_c \setminus \mathrm{span}(T_r - e)$ for all $e \in T_r$ (and in particular for all $e \in E^*$).

Recall that for every $e \in E^*$ there is a set of colours $C^{(e)} \subseteq [n]$ of size $|C^{(e)}| \geq (1 - \varepsilon')n$, such that for each colour $c \in C^{(e)}$ every element in $B_c \setminus \mathrm{span}(T_r - e)$ is $(T_r - e)$-absorbable with respect to $\mathcal{T}^{(e)}$. This means that $B_c \setminus \mathrm{span}(T_r - e) \subseteq E'$ for every $e \in E^*$ and $c \in C^{(e)}$, and hence $\phi_c(e) \in E'$ and $B_c \setminus \phi_c(T_r) \subseteq E'$. Therefore, since $|C^{(e)}| \geq (1 - \varepsilon')n$ for all $e \in E^*$ and each map $\phi_c : T_r \to B_c$ is an injection, we can find $|E^*| \cdot (1 - \varepsilon')n$ distinct elements of the form $\phi_c(e)$ in E'. In addition, when fixing an arbitrary element $e^* \in E^*$ (recalling that $|E^*| \geq 1$), we have $B_c \setminus \phi_c(T_r) \subseteq E'$ for the at least $(1 - \varepsilon')n$ colours $c \in C^{(e^*)}$ (and $|B_c \setminus \phi_c(T_r)| = n - |T_r|$). This gives at least $(1 - \varepsilon')n \cdot (n - |T_r|)$ elements in E' which do not lie in $\phi_1(T_r) \cup \cdots \cup \phi_n(T_r)$. Thus, we can conclude that

$$|E'| \geq |E^*| \cdot (1 - \varepsilon')n + (1 - \varepsilon')n \cdot (n - |T_r|) = (1 - \varepsilon')n \cdot (|E^*| + n - |T_r|).$$

Next, we claim that $|E''| \leq (\varepsilon n/4) \cdot |E^*|$. If $|E^*| \geq 8r/\varepsilon$, this automatically holds as

$$E'' \subseteq \bigcup_{e \in E^*} (T_1^{(e)} \cup \cdots \cup T_{r-1}^{(e)}) \subseteq (T_1 \cup \cdots \cup T_{r-1}) \cup \bigcup_{e \in E^*} \bigcup_{i=1}^{r-1} (T_1^{(e)} \setminus T_i),$$

and therefore

$$|E''| \leq |T_1 \cup \cdots \cup T_{r-1}| + \sum_{e \in E^*} \sum_{i=1}^{r-1} |T_1^{(e)} \setminus T_i| \leq (r-1) \cdot n + |E^*| \cdot (r-1) \cdot 3 \leq \frac{\varepsilon n}{8} \cdot |E^*| + \frac{\varepsilon n}{8} \cdot |E^*| = \frac{\varepsilon n}{4} \cdot |E^*|$$

by our assumption $|E^*| \geq 8r/\varepsilon$ and by $r \leq r_{\max} \leq L \cdot \ln(n^2) \leq \varepsilon n/24$. So in particular this automatically holds if $r_{\min} = 2$ (since then $|E^*| = m_{\mathcal{T}}(T_1, \ldots, T_r) \geq \varepsilon^4 n/10^5 \geq (8/\varepsilon)L \ln(n^2) \geq 8r_{\max}/\varepsilon \geq 8r/\varepsilon$). It also automatically holds if $r_{\min} = 1$ and $r \geq 10^4/\varepsilon^3$ (since then $|E^*| \geq m_{\mathcal{T}}(T_1, \ldots, T_r) \geq (1 + \varepsilon/4)^{r-1} \geq e^{(\varepsilon/8)(r-1)} \geq (e^{\varepsilon r/20})^2 \geq (1 + \varepsilon r/20)^2 \geq \varepsilon^2 r^2/400 \geq 8r/\varepsilon$). Finally, if $r_{\min} = 1$ and $r \leq 10^4/\varepsilon^3$, observe that, by Claim 7.8 for each $e \in E^*$ and each $i = 1, \ldots, r-1$, there are at most $\varepsilon^4 n/10^5$ elements

29

$e' \in T_i^{(e)}$ that are $(T_r - e)$-absorbable with respect to $\mathcal{T}^{(e)}$. Thus, by the definition of E'', we have $|E''| \leq |E^*| \cdot (r-1) \cdot \varepsilon^4 n/10^5 \leq (\varepsilon n/4) \cdot |E^*|$. Thus, we have shown $|E''| \leq (\varepsilon n/4) \cdot |E^*|$ in any case.

All in all, we can conclude (recalling that $\varepsilon' = \varepsilon/2$ and $|E^*| = m_{\mathcal{T}}(T_1, \ldots, T_r)$)

$$|E' \setminus E''| = |E'| - |E''| \geq (1 - \varepsilon')n \cdot (|E^*| + n - |T_r|) - (\varepsilon n/4) \cdot |E^*|$$
$$\geq (1 - 3\varepsilon/4)n \cdot (|E^*| + n - |T_r|) = (1 - 3\varepsilon/4)n \cdot \Big(m_{\mathcal{T}}(T_1, \ldots, T_r) + n - |T_r|\Big). \boxdot$$

By Claim 7.9, at least $(1 - 3\varepsilon/4)n \cdot (m_{\mathcal{T}}(T_1, \ldots, T_r) + n - |T_r|)$ elements $e' \in E(\mathcal{T}) \setminus (T_1 \cup \cdots \cup T_r)$ $= \bigcup_{T \in \mathcal{T} \setminus \{T_1, \ldots, T_r\}} T$ are $(T_1, \ldots, T_r)$-absorbable with respect to $\mathcal{T}$. In particular, there exists some $T \in \mathcal{T} \setminus \{T_1, \ldots, T_r\}$ containing at least

$$\frac{(1 - 3\varepsilon/4)n}{|\mathcal{T}| - r} \cdot \Big(m_{\mathcal{T}}(T_1, \ldots, T_r) + n - |T_r|\Big) \geq (1 + \varepsilon/4) \cdot \Big(m_{\mathcal{T}}(T_1, \ldots, T_r) + n - |T_r|\Big)$$

$(T_1, \ldots, T_r)$-absorbable elements with respect to $\mathcal{T}$ (here, we used that $|\mathcal{T}| - r = \lceil (1-\varepsilon)n \rceil - r \leq (1-\varepsilon)n$ and $(1 + \varepsilon/4)(1 - \varepsilon) \leq 1 - 3\varepsilon/4$). This means that $m_{\mathcal{T}}(T_1, \ldots, T_r, T) \geq (1 + \varepsilon/4) \cdot (m_{\mathcal{T}}(T_1, \ldots, T_r) + n - |T_r|)$. Recalling that $m_{\mathcal{T}}(T_1, \ldots, T_r, T_{r+1}) \geq m_{\mathcal{T}}(T_1, \ldots, T_r, T)$ by the choice of T_{r+1}, this implies (13), as desired. $\square$

8 Proofs of simple matroid lemmas

This section contains the proofs of several simple lemmas concerning matroids, rainbow independent sets and deadlocks, which we omitted in the previous sections. We start by proving Lemma 2.1.

Proof of Lemma 2.1. First, we can find an independent set S' with $S \subseteq S' \subseteq S \cup T$ of size $|S'| \geq |T|$ by starting with the independent set S and adding elements from the independent set T using the augmentation property (adding elements from T one at a time until $|S'| \geq |T|$). Note that then we have $|T \setminus S'| = |T \cup S'| - |S'| = |T \cup S| - |S'| \leq |T \cup S| - |T| = |S \setminus T|$. Every colour appears at most twice in S', since each of the sets S and T is rainbow. Furthermore, if a colour appears twice in S', then it must appear once in $S \setminus T \subseteq S'$ (and once in $S' \setminus S$). So there are at most $|S \setminus T|$ colours appearing twice in S', and for each of them we can delete an element of $S' \setminus S$ to remove the colouring conflict. This way, we obtain a rainbow independent set S^* with $S \subseteq S^* \subseteq S' \subseteq S \cup T$ and $|S' \setminus S^*| \leq |S \setminus T|$. Note that then we have $|T \setminus S^*| \leq |T \setminus S'| + |S' \setminus S^*| \leq 2 \cdot |S \setminus T|$. $\square$

Next, we will prove Lemmas 2.2 and 2.3. We will deduce both of these lemmas from the following statement due to Brualdi [4] (for the reader's convenience we also include a proof here).

Lemma 8.1 ([4]). *Let B and B' be bases of some matroid. Then, there is a bijection $\psi : B \to B'$ such that, for each $x \in B$, the set $B' - \psi(x) + x$ is independent.*

Proof. For each $x \in B$, let $S_x \subseteq B'$ be the set of all $x' \in B'$ such that $B' - x' + x$ is independent. We claim that $S_x + x$ cannot be an independent set. Indeed, if $S_x + x$ was an independent set for some $x \in B$, then we could find an independent set T with $S_x + x \subseteq T \subseteq (S_x + x) \cup B' = B' + x$ of size $|T| = |B'|$ via the augmentation property by successively adding elements from B'. But then the independent set T is of the form $B' - x' + x$ for some $x' \notin S_x$, which is a contradiction to the definition of S_x.

So indeed, for each $x \in B$, the set $S_x + x$ is not independent. Thus we have $x \in \mathrm{span}(S_x)$ for every $x \in B$. Then, for any subset $X \subseteq B$, we have $X \subseteq \mathrm{span}(\bigcup_{x \in X} S_x)$ and hence $|\bigcup_{x \in X} S_x| \geq \mathrm{rk}(\bigcup_{x \in X} S_x) = \mathrm{rk}(\mathrm{span}(\bigcup_{x \in X} S_x)) \geq \mathrm{rk}(X) = |X|$. Thus, by Hall's Theorem, we can find a bijection $\psi : B \to B'$ with $\psi(x) \in S_x$ for all $x \in B$. By the definition of S_x, this means that $B' - \psi(x) + x$ is an independent set for each $x \in B$. $\square$

Using Lemma 8.1, it is not hard to deduce Lemmas 2.2 and 2.3.

Proof of Lemma 2.2. Let B' be a basis of the matroid with $S \subseteq B'$. Apply Lemma 8.1 to find a bijection $\psi : B \to B'$ such that $B' - \psi(x) + x$ is an independent set for every $x \in B$. Letting $\phi : B' \to B$ be the inverse map of ψ, the set $B' - x + \phi(x)$ is independent for every $x \in B'$. In particular, this means that for

every $x \in S$, the set $S - x + \phi(x)$ is independent. Furthermore, for each $b \in B \setminus \phi(S)$, the set $B' - x + b$ is independent for some $x \in B' \setminus S$ (namely, the element x with $\phi(x) = b$). Thus, the set $S + b \subseteq B' - x + b$ is also independent. Finally, note that restricting ϕ to S gives an injection from S into B. $\qquad\square$

Proof of Lemma 2.3. Let B and B' be bases of the matroid with $S \subseteq B$ and $T \subseteq B'$. By Lemma 8.1, there is a bijection $\psi : B \to B'$ such that $B' - \psi(x) + x$ is an independent set for every $x \in B$. If there is some $x \in S \subseteq B$ with $\psi(x) \notin T$, then this means that $T + x \subseteq B' - \psi(x) + x$ is an independent set, and thus (a) holds. So let us assume that $\psi(x) \in T$ for all $x \in S$. Then, by restricting ψ to S, we get an injection $\phi : S \to T$ such that $B' - \phi(x) + x$, and hence also $T - \phi(x) + x$, is an independent set for every $x \in S$.

Furthermore, we may assume that $x \in \operatorname{span}(T)$ for every $x \in S$. Indeed, if there is some $x \in S$ with $x \notin \operatorname{span}(T)$, then $T + x$ is an independent set and (a) holds again. So, for any $x \in S$ we may assume that $x \in \operatorname{span}(T)$, which implies $T \cup \{x\} \subseteq \operatorname{span}(T)$ and hence $\operatorname{span}(T \cup \{x\}) = \operatorname{span}(T)$ (since $\operatorname{span}(T) \subseteq \operatorname{span}(T \cup \{x\}) \subseteq \operatorname{span}(\operatorname{span}(T)) = \operatorname{span}(T)$ by the basic matroid rules in Section 2). Thus,

$$\operatorname{span}(T - \phi(x) + x) \subseteq \operatorname{span}(T \cup \{x\}) = \operatorname{span}(T),$$

and together with $\operatorname{rk}(T - \phi(x) + x) = |T - \phi(x) + x| = |T| = \operatorname{rk}(T)$, this implies $\operatorname{span}(T - \phi(x) + x) = \operatorname{span}(T)$. Thus, (b) holds. $\qquad\square$

It remains to prove Lemmas 6.2 and 6.7. To do so, we first need some preparation. First, the following well-known lemma states that the rank function in a matroid is submodular (see for example [16, Lemma 1.3.1]). Again, we include a short proof for completeness.

Lemma 8.2. *For any subsets A and B of a matroid, we have $\operatorname{rk}(A \cap B) + \operatorname{rk}(A \cup B) \leq \operatorname{rk}(A) + \operatorname{rk}(B)$.*

Proof. Let $S \subseteq A \cap B$ be a maximal independent subset of $A \cap B$, then $|S| = \operatorname{rk}(A \cap B)$. Let us now extend S to a maximal independent subset S_A of A, then $S \subseteq S_A \subseteq A$ and $|S_A| = \operatorname{rk}(A)$. Similarly, let us extend S to a maximal independent subset S_B of B, then $S \subseteq S_B \subseteq B$ and $|S_B| = \operatorname{rk}(B)$. Now, $S_A \cup S_B = S \cup (S_A \setminus S) \cup (S_B \setminus S)$ is a set of size

$$|S| + |S_A \setminus S| + |S_B \setminus S| = \operatorname{rk}(A \cap B) + (\operatorname{rk}(A) - \operatorname{rk}(A \cap B)) + (\operatorname{rk}(B) - \operatorname{rk}(A \cap B)) = \operatorname{rk}(A) + \operatorname{rk}(B) - \operatorname{rk}(A \cup B).$$

Thus, as $A \cup B \subseteq \operatorname{span}(S_A) \cup \operatorname{span}(S_B) \subseteq \operatorname{span}(S_A \cup S_B)$, we have

$$\operatorname{rk}(A \cup B) \leq \operatorname{rk}(\operatorname{span}(S_A \cup S_B)) = \operatorname{rk}(S_A \cup S_B) \leq |S_A \cup S_B| = \operatorname{rk}(A) + \operatorname{rk}(B) - \operatorname{rk}(A \cup B),$$

and rearranging yields the desired inequality. $\qquad\square$

Next, we prove Lemma 6.2.

Proof of Lemma 6.2. Let $S' \subseteq S_1 \cup S_2$, then we have

$$\begin{aligned}
|(S_1 \cup S_2) \setminus S'| &= |S_1 \setminus (S' \cap S_1)| + |S_2 \setminus ((S' \cup S_1) \cap S_2)| \\
&\geq k \cdot (\operatorname{rk}(S_1) - \operatorname{rk}(S' \cap S_1)) + k \cdot (\operatorname{rk}(S_2) - \operatorname{rk}((S' \cup S_1) \cap S_2)) \\
&\geq k \cdot (\operatorname{rk}(S_1) - \operatorname{rk}(S' \cap S_1)) + k \cdot (\operatorname{rk}((S' \cup S_1) \cup S_2) - \operatorname{rk}(S' \cup S_1)) \\
&= k \cdot (\operatorname{rk}(S_1 \cup S_2) + \operatorname{rk}(S_1) - \operatorname{rk}(S' \cap S_1) - \operatorname{rk}(S' \cup S_1)) \\
&\geq k \cdot (\operatorname{rk}(S_1 \cup S_2) - \operatorname{rk}(S')),
\end{aligned}$$

where in the third step we used Lemma 8.2 (the submodularity of the rank function) for $A = S_2$ and $B = S' \cup S_1$ and in the last step we used it for $A = S_1$ and $B = S'$. Thus, $S_1 \cup S_2$ is k-overcrowded. $\quad\square$

Our last remaining task is to prove Lemma 6.7.

Proof of Lemma 6.7. Let $D = D_k(U \cup U')$, so that $D \subseteq U \cup U'$ is k-overcrowded and we wish to show that $D_k(U) = D$. First, we show that $D \setminus U'$ is k'-overcrowded. For this, let $S' \subseteq D \setminus U'$. If $\mathrm{rk}(S') = \mathrm{rk}(D \setminus U')$, then $|(D \setminus U') \setminus S'| \geq k' \cdot (\mathrm{rk}(D \setminus U') - \mathrm{rk}(S'))$ trivially holds. So assume $\mathrm{rk}(S') < \mathrm{rk}(D \setminus U')$, then $\mathrm{rk}(S') \leq \mathrm{rk}(D \setminus U') - 1 \leq \mathrm{rk}(D) - 1$. As D is k-overcrowded and $S' \subseteq D$, we have (using $|U'| \leq k - k'$)

$$|(D \setminus U') \setminus S'| \geq |D \setminus S'| - |U'| \geq k(\mathrm{rk}(D) - \mathrm{rk}(S')) - (k - k') \geq k'(\mathrm{rk}(D) - \mathrm{rk}(S')) \geq k'(\mathrm{rk}(D \setminus U') - \mathrm{rk}(S')).$$

Therefore, $D \setminus U'$ is indeed k'-overcrowded.

Since $D \setminus U' \subseteq (U \cup U') \setminus U' = U$, this implies $D \setminus U' \subseteq D_{k'}(U)$ and hence $D \setminus (D_{k'}(U) \cap D) \subseteq U'$. Thus, we obtain (again using that D is k-overcrowded)

$$k > k - k' \geq |U'| \geq |D \setminus (D_{k'}(U) \cap D)| \geq k \cdot \Big(\mathrm{rk}(D) - \mathrm{rk}(D_{k'}(U) \cap D) \Big).$$

This implies $\mathrm{rk}(D) - \mathrm{rk}(D_{k'}(U) \cap D) < 1$, meaning that $\mathrm{rk}(D) = \mathrm{rk}(D_{k'}(U) \cap D)$ and therefore $\mathrm{span}(D) = \mathrm{span}(D_{k'}(U) \cap D)$. In particular, we obtain $D \subseteq \mathrm{span}(D_{k'}(U) \cap D) \subseteq \mathrm{span}(D_{k'}(U))$. Recalling that $U' \cap \mathrm{span}(D_{k'}(U)) = \emptyset$, this implies $D \cap U' = \emptyset$ and hence $D \subseteq U$. So we can conclude that $D \subseteq D_k(U)$, as D is k-overcrowded. Since we also have $D_k(U) \subseteq D_k(U \cup U') = D$, we obtain $D_k(U \cup U') = D = D_k(U)$. $\square$

References

[1] R. Aharoni and E. Berger. The intersection of a matroid and a simplicial complex. *Transactions of the American Mathematical Society*, 358(11):4895–4917, 2006.

[2] N. Alon and J. H. Spencer. *The Probabilistic Method*. John Wiley & Sons, 2016.

[3] N. Alon and M. Tarsi. Colorings and orientations of graphs. *Combinatorica*, 12:125–134, 1992.

[4] R. A. Brualdi. Comments on bases in dependence structures. *Bulletin of the Australian Mathematical Society*, 1(2):161–167, 1969.

[5] M. Bucić, M. Kwan, A. Pokrovskiy, and B. Sudakov. Halfway to Rota's basis conjecture. *International Mathematics Research Notices*, 2020(21):8007–8026, 2020.

[6] M. Cheung. Computational proof of Rota's basis conjecture for matroids of rank 4. Unpublished Manuscript, `http://educ.jmu.edu/~duceyje/undergrad/2012/mike.pdf`.

[7] S. Dong and J. Geelen. Improved bounds for Rota's basis conjecture. *Combinatorica*, 39(2):265–272, 2019.

[8] A. A. Drisko. On the number of even and odd Latin squares of order $p + 1$. *Advances in Mathematics*, 128(1):20–35, 1997.

[9] J. Edmonds. Minimum partition of a matroid into independent subsets. *J. Res. Nat. Bur. Standards Sect. B*, 69B:67–72, 1965.

[10] J. Geelen and P. J. Humphries. Rota's basis conjecture for paving matroids. *SIAM Journal on Discrete Mathematics*, 20(4):1042–1045, 2006.

[11] J. Geelen and K. Webb. On Rota's basis conjecture. *SIAM Journal on Discrete Mathematics*, 21(3):802–804, 2007.

[12] D. G. Glynn. The conjectures of Alon-Tarsi and Rota in dimension prime minus one. *SIAM Journal on Discrete Mathematics*, 24(2):394–399, 2010.

[13] R. Huang and G.-C. Rota. On the relations of various conjectures on Latin squares and straightening coefficients. *Discrete Mathematics*, 128(1-3):225–236, 1994.

[14] R. Montgomery. *Transversals in Latin squares*, page 131–158. London Mathematical Society Lecture Note Series. Cambridge University Press, 2024.

[15] S. Onn. A colorful determinantal identity, a conjecture of Rota, and Latin squares. *The American mathematical monthly*, 104(2):156–159, 1997.

[16] J. Oxley. *Matroid theory*, volume 21 of *Oxford Graduate Texts in Mathematics*. Oxford University Press, Oxford, second edition, 2011.

[17] A. Pokrovskiy. *Rainbow Subgraphs and their Applications*, page 191–214. London Mathematical Society Lecture Note Series. Cambridge University Press, 2022.

[18] A. Pokrovskiy. Rota's basis conjecture holds asymptotically. *Proceedings of the American Mathematical Society*, to appear.

[19] D. H. J. Polymath. Rota's basis conjecture: Polymath 12. Polymath Wiki Page, `https://polymathprojects.org/2017/03/06/rotas-basis-conjecture-polymath-12-2/`.

[20] L. Sauermann. Rota's basis conjecture holds for random bases of vector spaces. *European Journal of Combinatorics*, 121:103836, 2024.

[21] B. Sudakov. Restricted subgraphs of edge-colored graphs and applications. *arXiv preprint arXiv:2412.13945*, 2024.

[22] M. Wild. On Rota's problem about n bases in a rank n matroid. *Advances in Mathematics*, 108(2):336–345, 1994.