

On the Parallel Complexity of Finding a Matroid Basis

Sanjeev Khanna*

Aaron Putterman†

Junkai Song‡

July 14, 2025

Abstract

A fundamental question in parallel computation, posed by Karp, Upfal, and Wigderson (FOCS 1985, JCSS 1988), asks: *given only independence-oracle access to a matroid on n elements, how many rounds are required to find a basis using only polynomially many queries?* This question generalizes, among others, the complexity of finding bases of linear spaces, partition matroids, and spanning forests in graphs. In their work, they established an upper bound of $O(\sqrt{n})$ rounds and a lower bound of $\Omega(n^{1/3})$ rounds for this problem, and these bounds have remained unimproved since then.

In this work, we make the first progress in narrowing this gap by designing a parallel algorithm that finds a basis of an arbitrary matroid in $\tilde{O}(n^{7/15})$ rounds (using polynomially many independence queries per round) with high probability, surpassing the long-standing $O(\sqrt{n})$ barrier. Our approach introduces a novel matroid decomposition technique and other structural insights that not only yield this general result but also lead to a much improved new algorithm for the class of *partition matroids* (which underlies the $\tilde{\Omega}(n^{1/3})$ lower bound of Karp, Upfal, and Wigderson). Specifically, we develop an $\tilde{O}(n^{1/3})$ -round algorithm, thereby settling the round complexity of finding a basis in partition matroids.

As a further application, we also improve the parallel complexity of the classic *matroid intersection* problem. By plugging our basis-finding algorithm into a known algorithmic framework for matroid intersection, we obtain an $\tilde{O}(n^{37/45})$ round algorithm for matroid intersection, improving upon the prior $O(n^{5/6})$ bound.

Collectively, these results represent the first progress on the parallel complexity of finding matroid bases in 40 years, and we believe that techniques developed here may prove useful for other problems on matroids.

*School of Engineering and Applied Sciences, University of Pennsylvania, Philadelphia, PA. Supported in part by NSF award CCF-2402284 and AFOSR award FA9550-25-1-0107. Email: sanjeev@cis.upenn.edu.

†School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts, USA. Supported in part by the Simons Investigator Awards of Madhu Sudan and Salil Vadhan, NSF Award CCF 2152413 and AFOSR award FA9550-25-1-0112. Email: aputterman@g.harvard.edu.

‡School of Engineering and Applied Sciences, University of Pennsylvania, Philadelphia, PA. Supported in part by NSF award CCF-2402284 and AFOSR award FA9550-25-1-0107. Email: junkais@cis.upenn.edu.

Contents

1	Introduction	1
1.1	Our Contributions	1
1.2	Technical Overview	2
1.2.1	Prior Work	2
1.2.2	Warming Up with Partition Matroids	3
1.2.3	Extending Notions from Partition Matroids to General Matroids	5
1.2.4	Building a Matroid Decomposition	6
1.2.5	Making Progress through Contraction	8
1.2.6	Making Progress Through Explicit Solving	9
1.2.7	Efficiently Finding Redundant Elements	10
1.2.8	Putting the Pieces Together	11
1.3	Organization	12
2	Preliminaries	12
2.1	Notation	12
2.2	Matroid Theory	12
2.3	Probability Theory	13
3	Tight Bounds for Partition Matroids	13
3.1	A Randomized Algorithm	14
3.2	Derandomization	19
4	Decomposition Algorithm	20
4.1	Finding Sets with Many Circuits	20
4.2	Iterative Matroid Decomposition	24
5	Making Progress With Large α's	28
6	Making Progress With Small α's	28
6.1	Rank Deficiency in Matroids	28
6.1.1	Matroid Decomposition for Quotient Counting Bounds	29
6.1.2	Proof of Rank Deficiency	30
6.2	Efficient Redundant Element Recovery	31
7	Guaranteeing Progress through Decomposition	33
7.1	Subroutines for Making Progress Towards a Basis	34
7.2	Piecing the Subroutines Together	36
8	Conclusions	39
A	Proof of Theorem 1.4	41

1 Introduction

A central pursuit in algorithm design is to understand which problems admit efficient parallel solutions. A common and well-studied measure is the *round complexity*: how many rounds of adaptive queries (or steps) are required to solve a problem, where each round can perform polynomially many operations/queries in parallel. This model has been extensively studied in both theoretical and applied settings, and has yielded parallel algorithms for a range of fundamental problems—including maximal independent sets [Lub86], matchings in graphs [Lov79, KUW86, FGT16, ST17], submodular function minimization [BS20, CCK21, CGJS22] and matroid intersection [GGR22, GT17, Bli22, BT25].

In this work, we focus on parallel computation in matroids. Matroids are a powerful abstraction capturing the structure of independence in many combinatorial settings, including forests in graphs, linearly independent vectors, and feasibility in optimization problems. We revisit a foundational open problem posed by Karp, Upfal, and Wigderson [KUW85, KUW88]:

Given oracle access to a matroid \mathcal{M} on n elements, how many adaptive rounds are required to find a basis of \mathcal{M} ?

We assume the algorithm has access only to an *independence oracle*, which takes a subset $S \subseteq E$ and returns whether $S \in \mathcal{I}$, where \mathcal{I} is the collection of independent sets. This is the most general model of matroid access and captures the full generality of matroid theory. In this setting, no further structure is assumed; even linear or graphic representations of the matroid are unavailable.

The importance of this question stems in part from the broad applicability of matroids. For example, when \mathcal{M} is a graphic matroid, the problem reduces to finding a spanning forest of a graph using only queries to a cycle oracle. For linear matroids, it becomes finding a basis of a subspace without direct access to coordinates. Given the super-exponential number of matroids [BPVdP15], understanding the round complexity in this general oracle model is a natural and foundational challenge.

A *matroid* $\mathcal{M} = (E, \mathcal{I})$ consists of a ground set E and a family of subsets $\mathcal{I} \subseteq 2^E$ satisfying three axioms: (1) $\emptyset \in \mathcal{I}$; (2) hereditary property ($S \subseteq T \in \mathcal{I} \Rightarrow S \in \mathcal{I}$); and (3) exchange property (if $S, T \in \mathcal{I}$ with $|S| < |T|$, there exists $e \in T \setminus S$ such that $S \cup \{e\} \in \mathcal{I}$). A *basis* is a maximal independent set; all bases of a matroid have the same size, called the *rank* of \mathcal{M} .

In their foundational work, Karp, Upfal, and Wigderson [KUW85] gave a parallel algorithm that finds a basis in $O(\sqrt{n})$ adaptive rounds using polynomially many independence queries per round. They also gave an $\tilde{\Omega}(n^{1/3})$ ¹ round lower bound for a family of *partition matroids*, showing that any algorithm using polynomially many queries per round must use at least $\tilde{\Omega}(n^{1/3})$ rounds. These bounds have remained the best known for nearly forty years.

1.1 Our Contributions

We make the first progress on narrowing the gap between upper and lower bounds established in the work of Karp, Upfal, and Wigderson [KUW85]. Our main result is a faster randomized algorithm that breaks the $\Theta(\sqrt{n})$ round barrier:

Theorem 1.1 (Main Result). *There is a randomized algorithm that, with high probability, finds a basis of any n -element matroid in $\tilde{O}(n^{7/15})$ adaptive rounds, using only polynomially many independence queries.*

¹Throughout the paper, we use $\tilde{O}(\cdot)$ and $\tilde{\Omega}(\cdot)$ to hide factors of $\text{polylog}(\cdot)$.

Remark 1.2. As an immediate consequence, given a *weighted* matroid (where every element has an associated weight), there is also an $\tilde{O}(n^{7/15})$ round algorithm for finding a maximum (or minimum) weight basis, as shown in [BT25].

Our algorithm relies on several new techniques and structural results, including, a *matroid decomposition framework* that partitions the ground set into a small number of “components” that are amenable to localized processing; a new characterization of *rank deficiency* in terms of random sampling; and a parallel routine for *recovering redundant elements* that improves over previous deletion-based strategies.

Our techniques (in fact, a simplified version of them) also lead to significantly improved round complexity for the important special case of *partition matroids*.

Partition matroids. Partition matroids form a natural and widely studied subclass of matroids and were used in the lower bound construction of [KUW85]. We show that the known $\tilde{\Omega}(n^{1/3})$ lower bound is essentially tight for this class by providing a matching algorithm:

Theorem 1.3. *There is a deterministic algorithm that finds a basis of any n -element partition matroid in $\tilde{O}(n^{1/3})$ adaptive rounds, using only polynomially many independence queries.*

Matroid intersection. The matroid intersection problem is a common generalization of several fundamental combinatorial optimization problems, including bipartite matching, arborescence in directed graphs, and the colorful spanning tree problem. Recent work by [BT25] showed that matroid intersection can be solved via calls to a matroid basis oracle. Plugging in our improved basis-finding algorithm immediately gives a better round complexity (improving upon [BT25]’s $O(n^{5/6})$ rounds):

Theorem 1.4. *There is a randomized algorithm that, with high probability, finds a maximum (weight) common independent set of two n -element matroids in $\tilde{O}(n^{37/45})$ adaptive rounds, using only polynomially many independence queries.*

We now give an overview of the main technical ideas underlying our results.

1.2 Technical Overview

1.2.1 Prior Work

First, we recap the algorithm and analysis (for the upper bound) of [KUW85]. We start by reviewing two natural operations on matroids:

1. **Contraction:** Contraction in a matroid $\mathcal{M} = (E, \mathcal{I})$ refers to finding a set of independent elements S , and committing to include these in our basis. The contracted matroid (denoted \mathcal{M}/S) has elements $E \setminus S$, and a set $T \subseteq E \setminus S$ is independent *if and only if* $T \cup S$ is independent in \mathcal{M} . For this reason, when we contract on a set S , we can still simulate independence oracle queries to \mathcal{M}/S by simply appending the set S to the query (and then querying \mathcal{M}). Importantly, when we contract on a set S , the remaining number of elements in \mathcal{M}/S decreases by $|S|$, and so in this sense, the problem of finding a basis in \mathcal{M} is reduced to that of finding a basis in a smaller instance.
2. **Deleting Redundant Elements:** The second method of making “progress” towards recovering a basis in a matroid is by deleting redundant elements. Specifically, we say that a set S of elements in a matroid \mathcal{M} is *redundant*, if $\text{rank}(\mathcal{M} \setminus S) = \text{rank}(\mathcal{M})$. Clearly, if we can find such a redundant set, then finding a basis in \mathcal{M} reduces to finding a basis in $\mathcal{M} \setminus S$.

Now, we describe the $\text{poly}(n)$ -query $O(\sqrt{n})$ -round algorithm for finding bases in arbitrary matroids as first presented in [KUW85]. With the previous notions already established, the algorithm itself is quite simple: given a matroid \mathcal{M} on n elements e_1, \dots, e_n , the algorithm first splits the elements up into \sqrt{n} groups: $S_1 = \{e_1, \dots, e_{\sqrt{n}}\}$, $S_2 = \{e_{\sqrt{n}+1}, \dots, e_{2\sqrt{n}}\}$, \dots , $S_{\sqrt{n}} = \{e_{n-\sqrt{n}+1}, \dots, e_n\}$. Now, the queries that the algorithm makes are simply all prefixes of S_i , for every $i \in [\sqrt{n}]$ (i.e., for S_1 , the queries would be $\{e_1\}, \{e_1, e_2\}, \{e_1, e_2, e_3\} \dots \{e_1, e_2, \dots, e_{\sqrt{n}}\}$). There are only two cases for us to consider:

1. If, for any set S_i , we see that $\text{Ind}(S_i) = 1$ (i.e., the entire set is independent), then the algorithm simply contracts on S_i . In particular, this means that we will have recovered an independent set of size \sqrt{n} , and so when we contract on the set S_i , we see that $\text{Rank}(\mathcal{M}/S_i) = \text{Rank}(\mathcal{M}) - \sqrt{n}$. Thus, in future rounds, we simply operate on the new matroid defined by \mathcal{M}/S_i .
2. If there are no sets S_i for which $\text{Ind}(S_i) = 1$, this means that every set S_i was dependent. In particular, because we queried every prefix of S_i , we can also identify the first query which became dependent. I.e., there were two queries where $\{e_{(i-1)*\sqrt{n}+1}, \dots, e_{(i-1)*\sqrt{n}+j}\}$ was independent, but $\{e_{(i-1)*\sqrt{n}+1}, \dots, e_{(i-1)*\sqrt{n}+j+1}\}$ was dependent. A simple fact in matroid theory states that when this happens, $e_{(i-1)*\sqrt{n}+j+1}$ is a *redundant* element for any matroid which contains elements $\{e_{(i-1)*\sqrt{n}+1}, \dots, e_{(i-1)*\sqrt{n}+j}\}$, and can therefore be deleted. Thus, summing across all \sqrt{n} groups, we can delete \sqrt{n} redundant elements.

Thus, we see that in either case we are making progress towards recovering a basis of the matroid. Either we recover \sqrt{n} independent elements towards a basis, or we delete \sqrt{n} redundant elements. In both cases, we reduce the problem to finding a basis on a matroid with $n - \sqrt{n}$ elements, and so a simple recursive calculation reveals that this terminates in $O(\sqrt{n})$ rounds.

1.2.2 Warming Up with Partition Matroids

In the work of [KUW85], their lower bound of $\tilde{\Omega}(n^{1/3})$ rounds for finding a basis relied on the class of *partition matroids*. In this setting, the matroid $\mathcal{M} = (E, \mathcal{I})$ has its elements E partitioned across k parts, denoted P_1, \dots, P_k . Each set P_i is also given a *budget* $b_i \in \mathbb{Z}^{\geq 0}$. Given a set $S \subseteq E$, we define the rank as

$$\text{rank}(S) = \sum_{i=1}^k \min(b_i, |S \cap P_i|).$$

Essentially, as we add elements to a given part, its rank continues to increase until it reaches its budget, at which point the rank remains fixed. Note for a given part P_i , once we find a set of $b_i + 1$ elements that belongs to P_i , we can in fact *completely* recover P_i in a single additional round. If we denote such a set by A_i , this is because we can remove a single element x from A_i , and then query the independence oracle with $(A_i \setminus \{x\}) \cup \{y\}$ for every other element $y \in \mathcal{M}$. The only elements which will make such a query dependent are those that are also in the part P_i .

Given this definition, it is actually quite straightforward to define the lower bound instances that [KUW85] relied on. For an n element ground set, they create $n^{1/3}$ parts, denoted $P_1, \dots, P_{n^{1/3}}$, where each part contains $n^{2/3}$ elements, and the budget of the i -th part is exactly $i \cdot n^{1/3}$. This defines the *structure* of the matroid, but the actual assignment of labels to the elements (i.e., which elements are in which part) is decided uniformly at random. In the first round now, because the labels are decided uniformly at random, any query of elements is effectively a random sample of the n elements of the matroid (and, for notation, let us say the query is denoted by B , and the random sampling is performed at some rate $\beta = |B|/n$).

The key point is that the response to the query B is *with extremely high probability* completely governed by the elements in $B \cap P_1$. Indeed, if the sampling rate $\beta \geq \frac{1.5}{n^{1/3}}$, then with extremely high probability, the query B is dependent, as $> (1 - \epsilon) \cdot \beta \cdot n^{2/3} > n^{1/3}$ elements will be sampled from P_1 which exceeds the budget b_1 . Likewise, if $\beta < \frac{1.5}{n^{1/3}}$, then the budgets among the sets $P_2, \dots, P_{n^{1/3}}$ are *not* exceeded, and so the independence or dependence of the query B is dictated entirely by $B \cap P_1$. Intuitively then, this means that in the first round of queries, the only information that is being leaked relates to the set of elements P_1 . The argument can then be repeated, where in the second round, the only leaked information is from P_2 , and so on. This argument is formalized in [KUW85].

However, the above construction is very suggestive: could we hope for a better lower bound using partition matroids? In fact, a natural starting point for this question would be to consider a partition matroid with \sqrt{n} parts $P_1, \dots, P_{\sqrt{n}}$, where each part has \sqrt{n} elements, and the budget b_i of part P_i is i . If an induction argument as used in the previous setting held true, then we could even hope for an $\Omega(\sqrt{n})$ lower bound, matching the algorithm provided by [KUW85].

It turns out however, that this is not possible. The problem that arises is the following: suppose we have eliminated the first $i - 1$ parts, so now the matroid consists of $P_i, \dots, P_{\sqrt{n}}$. Next, consider a query B which consists of sampling each element independently with probability approximately $\frac{i}{\sqrt{n}}$. In expectation, the number of elements that are sampled from each part is exactly $\frac{i}{\sqrt{n}} \cdot \sqrt{n} = i$. However, the problem that arises is that the number of surviving elements follows a binomial distribution and therefore *anti-concentrates* with non-negligible probability. Indeed, there is a non-negligible probability that P_i receives fewer than i elements, while P_{i+1} receives more than $i + 1$ elements, and thus it is actually P_{i+1} which dictates the response to query B . In fact, because the standard deviation scales as \sqrt{i} , it is even the case that if we make the sampling rate slightly less than $\frac{i}{\sqrt{n}}$, *any* of the parts $P_i, \dots, P_{i+\Omega(\sqrt{i})}$ has a non-negligible chance of dictating the response to the query B . This allows for an algorithm to gain too much information about the underlying parts $P_i, \dots, P_{i+\Omega(\sqrt{i})}$, and can in fact reveal all of their exact identities in a single round. Repeating this argument leads to an algorithm that solves this instance in $O(n^{1/4})$ rounds!

Thus, while this thought experiment does not yield new lower bounds, it does provide an algorithmic insight, namely, instances of partition matroids where each adaptive round can only recover a single part, must have the property that the budgets are “well-separated” (by a standard deviation at least). This ensures that in such an instance with k parts, the budgets must grow to $\Omega(k^2)$. This means that after $n^{1/3}$ rounds of recovery we are dealing with a partition matroid with $\Omega(n^{2/3})$ -size budget and hence $O(n^{1/3})$ remaining parts. It follows that the remaining parts can be discovered in $O(n^{1/3})$ additional rounds. This is essentially the basis for our $\tilde{O}(n^{1/3})$ round algorithm to find basis of any partition matroid. In fact, because the class of partition matroids is only exponentially large, we can even derandomize this algorithm. We do this by repeating our decomposition procedure $\text{poly}(n)$ times in each round, thereby boosting the success probability to be $1 - 2^{-\text{poly}(n)}$, at which point we can take a union bound over *every possible partition matroid*. This implies that there is single, fixed polynomial size set of queries which simultaneously works for recovering a basis *across all* partition matroids. This set of queries constitutes the deterministic algorithm.

However, all of the analysis above is heavily tailored to the setting of partition matroids. The next question is if these insights can be extended to *general* matroids, by developing analogues of the notions of parts and budgets. It turns out that the answer is a “yes” for many of the above notions if we develop a sampling-based view of these concepts. In the remainder of the technical overview, we introduce these new quantities and explain how they can be stitched together to improve on [KUW85]’s long-standing upper bound.

1.2.3 Extending Notions from Partition Matroids to General Matroids

To extend the algorithmic approach used for partition matroids to arbitrary matroids, we must address a fundamental question:

What is the appropriate analogue of a “part” in a general matroid, and how can we discover such structure via random sampling?

Our main tool is a new decomposition framework that simulates the process of “peeling off” components from the matroid, in analogy to how we recovered parts in the partition matroid case. Instead of relying on explicit part definitions, we use random permutations to expose dependence structure. The core idea is that, under random sampling, elements or sets that frequently cause early dependence can be interpreted as playing the role of “tight” components.

Circuits from random prefixes. Let $\mathcal{M} = (E, \mathcal{I})$ be a matroid of rank r . Consider a random permutation π of the ground set E . Define a prefix process as before: let $S_j := \{\pi(1), \dots, \pi(j)\}$, and find the smallest index j^* such that S_{j^*} is dependent. Because j^* is the smallest such index, S_{j^*-1} is independent, and the dependent set S_{j^*} contains a unique minimal dependent subset, i.e., a *circuit* $C_\pi \subseteq S_{j^*}$. In particular, we can actually *exactly* recover the elements in this circuit that has formed: If we query with $S_{j^*} - \{x\}$ for $x \in C_\pi$, then this query will in fact be independent. Likewise, if we query with $S_{j^*} - \{x\}$ for $x \notin C_\pi$, the resulting query is dependent, as the circuit C_π is still contained in S_{j^*} .

We will view the circuit C_π as a random variable: it is the unique first circuit encountered over the randomness of π . We now define the following notions:

- For a subset $S \subseteq E$, let $q_{\mathcal{M}}(S) := \Pr_\pi[C_\pi \subseteq S]$ be the probability that the first circuit lies entirely in S .
- For an element $x \in E$, let $p_{\mathcal{M}}(x) := \Pr_\pi[x \in C_\pi]$ be the probability that x appears in the first circuit.

We can interpret $p_{\mathcal{M}}(x)$ as the “circuit participation mass” of element x under random prefix sampling.

Greedily-optimal sets. We now define the building blocks of our decomposition, namely, *greedily-optimal sets*. Informally speaking, a subset $S \subseteq E$ is a greedily-optimal set if it maximizes $q_{\mathcal{M}}(S)$ (the probability that the first circuit lies inside S), subject to a stability condition: removing any element from S significantly reduces $q_{\mathcal{M}}(S)$.

Definition 1.5. We say that a set $S^* \subseteq E$ is *greedily-optimal* if

$$q_{\mathcal{M}}(S^*) \geq 1 - 2^{-20} + \frac{\sum_{i=1}^{|S^*|} \frac{1}{i}}{2^{20} \log n},$$

and there is no element $x \in S^*$ such that

$$q_{\mathcal{M}}(S^* \setminus \{x\}) \geq 1 - 2^{-20} + \frac{\sum_{i=1}^{|S^*|-1} \frac{1}{i}}{2^{20} \log n}.$$

Note that here 2^{20} is just a sufficiently large constant to ensure our probabilistic arguments go through.

There are two key useful properties of greedily-optimal sets. First, they can be found efficiently. This is because we can start with the complete set $S^* = E$, and as long as there exists an element x which satisfies the second condition, we update $S^* \leftarrow S^* - \{x\}$. All that we need to know for this are the probabilities $q_{\mathcal{M}}(S)$ for every S , and we can estimate these to very high accuracy by empirically evaluating them on a large sample of random permutations (and this takes only a single round). Second, once we find a greedily optimal set S^* , this means $q_{\mathcal{M}}(S^*) \geq 1 - 2^{-20} + \frac{\sum_{i=1}^{|S^*|} \frac{1}{i}}{2^{20} \log n}$, but for any element we delete, $q_{\mathcal{M}}(S^* - \{x\}) < 1 - 2^{-20} + \frac{\sum_{i=1}^{|S^*|-1} \frac{1}{i}}{2^{20} \log n}$. Importantly, this means for every element $x \in S^*$, it intuitively participates in a large fraction of the circuits that appear. Formally, we let $p_{\mathcal{M}|_{S^*}}(x) = \Pr_{\pi}[x \in C_{\pi}]$, where (now C_{π} refers to the circuits that appear when sampling a permutation restricted to S^* , *not* the original matroid \mathcal{M}). We show that, in a very strong sense, this second piece of intuition is true:

Claim 1.6. *Let $\mathcal{M} = (E, \mathcal{I})$ be a matroid on n elements, and let S^* be a greedily-optimal set. Then, for every $x \in S^*$,*

$$p_{\mathcal{M}|_{S^*}}(x) \geq \frac{1}{2 \cdot 2^{20} |S^*| \log n}.$$

These greedily-optimal sets will play the role of “components” in our general decomposition, mimicking the parts P_i in partition matroids. While a greedily-optimal set S does not come with an explicit budget, it can be characterized by an analogous notion using a parameter $\alpha(S)$ that measures how early dependences appear in S under random sampling.

The $\alpha(S)$ parameter. For a set $S \subseteq E$, we define the parameter $\alpha(S)$ as the median prefix length (within S) required to trigger dependence:

$$\alpha(S) := \text{median}_{\pi} \{j : \{\pi(1), \dots, \pi(j)\} \subseteq S \text{ is dependent, } \{\pi(1), \dots, \pi(j-1)\} \subseteq S \text{ is independent}\},$$

where π is a random permutation of S .

Informally, $\alpha(S)$ captures how “tight” the set S is: smaller $\alpha(S)$ implies dependence tends to occur early when sampling from S . This plays the role of the “budget” in partition matroids: if $|S| = \ell$ and $\alpha(S) \approx b + 1$, then the behavior of S resembles a part of size ℓ and budget b .

We use this parameter to stratify components of the matroid and guide their ordering in the decomposition.

1.2.4 Building a Matroid Decomposition

Our decomposition does the following: given a matroid \mathcal{M} , it finds a greedily optimal set S_1 of \mathcal{M} in a single round. It peels this set off, and then repeats the decomposition starting with the matroid $\mathcal{M} \setminus S_1$. This procedure continues repeating until the matroid is exhausted (has no elements remaining). Note that a greedily-optimal set cannot be empty, so we are also guaranteed that this procedure terminates in a finite number of steps. We denote the resulting greedily-optimal sets that are peeled off by S_1, \dots, S_k .

The informal claim below helps bound the number of sets that we peel off and allows us to relate these sets to one another.

Claim 1.7. *[Informal] Let \mathcal{M} be a matroid, and let S_i, S_j be any two sets that are peeled off in the course of the above decomposition, with $j > i$. Then,*

$$\alpha(S_j) = \frac{\alpha(S_i)|S_j|}{|S_i|} + \Omega\left(\sqrt{\frac{\alpha(S_i)|S_j|}{|S_i|}}\right).$$

To start, we give a brief sketch which describes why the above claim is true. Let S_i, S_j be given, and let \mathcal{M}_{i-1} denote the state of the matroid \mathcal{M} after $i-1$ iterations. Then, S_i is a greedily-optimal set with respect to \mathcal{M}_{i-1} . In particular, because S_i is greedily-optimal, $q_{\mathcal{M}}(S_i) \geq 1 - 1/2^{20}$: this means that when we sample a random permutation π and add elements from \mathcal{M}_{i-1} until there is a circuit, there is a $\geq 1 - 1/2^{20}$ probability that the resulting circuit is completely contained in S_i . That is to say, it is exceedingly rare for there *ever* to be a circuit which includes elements from outside S_i .

Now, instead of sampling a random permutation π and adding elements in this order, we consider a slightly different procedure: we let $p = \frac{\alpha(S_i)}{|S_i|}$, and we consider what happens when we sample the matroid \mathcal{M}_{i-1} at rate p .

1. Within the set S_i , we expect exactly $\alpha(S_i)$ elements to survive. In fact, with probability $\Omega(1)$, there are even $\leq \alpha(S_i) - 1$ elements selected from S_i . Because $\alpha(S_i)$ is defined as the median number of samples needed before a circuit appears in S_i , if $\leq \alpha(S_i) - 1$ elements are sampled from S_i , then with probability $\geq 1/2$, no circuit forms completely inside S_i .
2. At the same time, we consider what happens when we sample the set S_j . We expect $\frac{\alpha(S_i)|S_j|}{|S_i|}$ elements to be sampled from S_i . Indeed, because the number of elements sampled follows a binomial distribution, we even know that the standard deviation of the number of sampled elements is $\sqrt{\frac{\alpha(S_i)|S_j|}{|S_i|}}$. Thus, with constant probability, there are

$$\geq \frac{\alpha(S_i)|S_j|}{|S_i|} + \Omega\left(\sqrt{\frac{\alpha(S_i)|S_j|}{|S_i|}}\right)$$

elements that survive sampling in S_j .

Now, we are ready to prove the above claim: when we sample \mathcal{M}_{i-1} at rate p , there is a constant probability of both (1) no circuit in S_i , and (2) at least $\frac{\alpha(S_i)|S_j|}{|S_i|} + \Omega\left(\sqrt{\frac{\alpha(S_i)|S_j|}{|S_i|}}\right)$ elements sampled from S_j . If we assume for the sake of contradiction that $\alpha(S_j) = \frac{\alpha(S_i)|S_j|}{|S_i|} + o\left(\sqrt{\frac{\alpha(S_i)|S_j|}{|S_i|}}\right)$, then point (2) would imply that there is also a constant probability of there being a circuit completely contained in S_j . However, this would yield a contradiction, as it implies that there is a constant probability of recovering a circuit in S_j *before* we recover a circuit in S_i . But, by our definition of S_i being greedily optimal, we know that when we add random elements, a $\geq 1 - 1/2^{20}$ fraction of circuits appear only in S_i . The above argument would otherwise suggest that there is a larger than $1/2^{20}$ probability of circuits appearing *outside* S_i .

Having established [Claim 1.7](#), we can already observe some interesting behavior about our decomposition. For instance, if we have two sets $|S_i| = |S_j|$, then $\alpha(S_j) = \alpha(S_i) + \Omega(\sqrt{\alpha(S_i)})$. If we generalize this to more than just two sets, and instead suppose we have $|S_1| = |S_2| = \dots = |S_\ell|$, then we get a chain of growth in the alpha values:

$$\alpha(S_2) = \alpha(S_1) + \Omega\left(\sqrt{\alpha(S_1)}\right), \quad \alpha(S_3) = \alpha(S_2) + \Omega\left(\sqrt{\alpha(S_2)}\right),$$

and so on. Ultimately, this means that if we have ℓ sets of the same size, the α value of the final set is $\Omega(\ell^2)$.

In fact, we can generalize this argument to sets whose sizes are within constant factors of one another:

Lemma 1.8. *Let \mathcal{M} be a matroid, and let S_1, \dots, S_k be a sequence of sets that are peeled off. Let $\ell \in [\log n]$ be an integer, let $T = \{i \in [k] : |S_i| \in [2^\ell, 2^{\ell+1} - 1]\}$, let $\gamma = |T|$, and let a_γ be the largest index in T . Then it must be the case that*

$$\alpha(S_{a_\gamma}) = \Omega(\gamma^2).$$

Likewise, because $\alpha(S) \leq |S|$, this also means that there can only ever be $O(\sqrt{2^\ell})$ sets whose sizes are in the range $[2^\ell, 2^{\ell+1} - 1]$. This immediately gives us a bound on the number of sets that can be returned by our decomposition: first, there must be $\leq n^{1/3}$ sets of size $\geq n^{2/3}$ (otherwise all the elements of the matroid are removed). Now, for any $\ell \in [\frac{2}{3} \log n]$, there can only be $O(2^{\ell/2})$ sets of size $[2^\ell, 2^{\ell+1}]$. Summing over the values of ℓ , the number of sets returned by the decomposition is bounded by $O(n^{1/3})$, exactly as we saw in the partition matroid case.

With these basic facts established about the decomposition, we now show how the decomposition gives us the power to make progress towards finding a basis in many different ways.

1.2.5 Making Progress through Contraction

To simplify the above discussion, we consider an abridged version of the above decomposition, where instead of continuing until the matroid \mathcal{M} is empty, we stop the decomposition as soon as there are $< n/2$ elements remaining. As before, we still denote the sets that are recovered by S_1, \dots, S_k . Our key observation now is the following:

Claim 1.9. *In one additional round, we can recover an independent set of size*

$$\max_{i \in [k]} \Omega\left(\frac{\alpha(S_i)}{|S_i|} \cdot n\right).$$

Thus, if this ratio $\frac{\alpha(S_i)}{|S_i|}$ every becomes too large (i.e., imagine it becomes $\Omega(1)$), then in just a single additional round, we can recover many independent elements, thereby making progress towards recovering an independent set.

To see why the claim is true, consider any matroid \mathcal{M} along with a greedily optimal set S of \mathcal{M} . As S is greedily optimal, it should be the case that S contains the vast majority of circuits that appear when random sampling. So if we sample \mathcal{M} at rate $\frac{\alpha(S)}{10|S|}$, then we know that there is a constant probability of *no* circuits in S , and therefore there must also not be any circuits in \mathcal{M} (with high probability), which in turn implies the sampled set is independent. In order to amplify this success probability, we can just repeat this sampling procedure for $\text{poly}(n)$ times (in parallel) and obtain an independent set of size $\Omega\left(\frac{\alpha(S)}{|S|} |\mathcal{M}|\right)$ with high probability. To get the above claim in its exact form, we can simply repeat this procedure for *every* S_i that is peeled off; i.e., with the matroid \mathcal{M} and S_1 , with $\mathcal{M}_1 = (\mathcal{M} \setminus S_1)$ and S_2 , and so on. Because we stopped the decomposition before \mathcal{M} reaches $< n/2$ elements, in each case, the remaining elements have cardinality $|\mathcal{M}| \geq \Omega(n)$. This yields the claim.

On an intuitive level, the benefit from being able to contract is that we can now assume that $\frac{\alpha(S_i)}{|S_i|} = o(1)$, for if not, there is a simple mechanism for making progress. Even better, since our only objective is to beat $n^{0.5}$ rounds, we can even assume the gap is non-trivially large, some n^ϵ for $\epsilon > 0$. In the next section, we will show how we can make progress towards recovering redundant elements when this “ α -gap” is large.

1.2.6 Making Progress Through Explicit Solving

As discussed in the previous section, we will assume that $\frac{\alpha(S_i)}{|S_i|} = o(1)$ is significantly smaller than 1. We are motivated by the following observation: if, in the course of doing our decomposition, there are many of the sets S_1, \dots, S_k that are of size $[\tau/2, \tau]$, then we can in fact *explicitly* find bases of these sets by investing $O(\sqrt{\tau})$ extra rounds, using the algorithm of [KUW85]. Unfortunately, as already remarked by [KUW85], there is no guarantee that *combining* bases together is making meaningful progress towards finding a basis of the entire matroid. Instead, we have to argue that each one of these sets has many *redundant* elements that we can delete *in parallel* to make progress.

If we were only guaranteed that $\frac{\alpha(S_i)}{|S_i|} = o(1)$, it is hard to argue that there are many redundant elements to be deleted (consider for instance matroid that has n^ϵ elements with rank 0). It is here where we first use the second key property of a greedily-optimal set: i.e., that *every element* in the set S_i has a large probability (approximately $\tilde{\Omega}(1/|S_i|)$) of being in the first recovered circuit when we randomly sample elements from S_i . It turns out that these two conditions suffice for non-trivially bounding the number of redundant elements in each S_i :

Theorem 1.10. *Let S be a set of elements in a matroid \mathcal{M} such that $\text{rank}(S) = r$ and*

1. $\alpha(S) \leq \frac{|S|}{100 \log |S|}$.
2. *For every element $x \in E$, $p_{\mathcal{M}|_S}(x) \geq \frac{1}{|S|^{10}}$.*

Then, $|S| - r = \Omega(|S|/\log |S|)$.

To see why this theorem is true, we need several additional tools and pieces of notation: we let $\mathcal{M}' = \mathcal{M}|_S$ be the matroid with only elements in S , and we let $(\mathcal{M}')^*$ denote the dual matroid of \mathcal{M}' . This is the matroid on the same set of elements, whose bases are the *complements* of bases in \mathcal{M}' , and whose rank is therefore $|S| - \text{rank}(S)$. Next, we require the notion of a *quotient* of a matroid: for a set $R \subseteq \mathcal{M}'$, we say that the quotient of R (denoted $Q(R)$) is the set of elements $\mathcal{M}' - \text{span}(R)$, where $\text{span}(R) = \{x \in \mathcal{M}' : \text{rank}(\{x\} \cup R) = \text{rank}(R)\}$. A key fact in matroid theory [Oxl06] is that if a set of elements C forms a circuit in \mathcal{M}' , then the *same* set of elements forms a quotient in $(\mathcal{M}')^*$. Finally, we require a useful theorem from the work of Quanrud [Qua24], which we invoke on the *dual matroid* $(\mathcal{M}')^*$: for any parameter $d \in \mathbb{Z}^+$, there is a set P of $\leq d \cdot \text{rank}((\mathcal{M}')^*)$ elements, such that in the matroid $(\mathcal{M}')^* \setminus P$, for any $\alpha \in \mathbb{Z}^+$, there are at most $|S|^{2\alpha}$ quotients of size $\leq \alpha \cdot d$. For intuition, this result is a strengthening of the *cut-counting bound* first derived in the work of Karger [Kar93]. The quotients correspond to the cuts of a graph, and the elements we remove correspond to the *small cuts* that we remove in order to make the minimum cut larger. With these results established, we are ready to give an outline of the proof of our theorem.

To start, recall that our goal is to show that $|S| - \text{rank}(S)$ is large. We can immediately see that this value is exactly the rank of the dual matroid $(\mathcal{M}')^*$, and so it suffices to lower bound $\text{rank}((\mathcal{M}')^*)$. We then proceed by contradiction: if it happens to be the case that $\text{rank}((\mathcal{M}')^*) = o(|S|/\log |S|)$, then by the result of [Qua24], for $d = 100 \log |S|$, there exists a set $P \subseteq (\mathcal{M}')^*$ of size $|P| \leq d \cdot \text{rank}((\mathcal{M}')^*) = o(|S|)$, such that removing this set of elements leads to a “quotient counting bound” with parameter d in the matroid $(\mathcal{M}')^* \setminus P$. Importantly, this counting bound implies that if we sample the elements of $(\mathcal{M}')^* \setminus P$ at rate $1/2$, then the probability that there is *any quotient* in $(\mathcal{M}')^* \setminus P$ for which *every element* is sampled is bounded by $2/|S|^{97}$. This is because there is an implicit tradeoff: smaller quotients are more likely to survive sampling, but there are fewer of them. As we increase the quotient size, there are more quotients, but their “survival” probability

decreases proportionately. Note that by circuit-quotient duality this means that if we sample the matroid $((\mathcal{M}')^* \setminus P)^*$ at rate $1/2$ we expect no *circuits* to survive sampling.

To reach a contradiction, we have to understand this matroid $((\mathcal{M}')^* \setminus P)^*$. A well-known property of matroids (see, for instance [Oxl06]), is that deleting elements is actually *dual* to contracting on elements. Thus, the matroid $((\mathcal{M}')^* \setminus P)^* = ((\mathcal{M}')^*)^* / P = \mathcal{M}' / P$ (i.e., our starting matroid \mathcal{M}' contracted on P). By the previous paragraph, this implies that when we sample \mathcal{M}' / P at rate $1/2$, with extremely high probability, there are no circuits that survive the sampling. However, using assumption (2) of our theorem statement, for any element $x \notin P$, we can show that sampling \mathcal{M}' at rate $1/2$ yields a circuit involving x with probability $\geq 1/|S|^{10}$. After this contraction on P , it is in fact the case that these elements in $\mathcal{M}' \setminus P$ are *only more likely* to form circuits when sampling at rate $1/2$ compared to in \mathcal{M}' , and thus this probability is still $\geq 1/|S|^{10}$. But, this yields our contradiction: our first bound says that with probability $\geq 1 - 2/|S|^{97}$, no circuits appear at this sampling rate, while our second bound says that circuits do appear with probability $\geq 1/|S|^{10}$. Therefore, it must be the case that $\text{rank}((\mathcal{M}')^*) = \Omega(|S|/\log |S|)$, as we desire.

Note that with this theorem in hand, we are guaranteed that for each set S_i of size $\lceil \tau/2, \tau \rceil$ that we explicitly find a basis for, we can essentially delete $\tilde{\Omega}(\tau)$ redundant elements, thereby guaranteeing some form of progress. Unfortunately, this method has a key drawback, which is that we *explicitly* solve for a basis, which requires investing even more rounds of adaptivity. In the next section, we present our final approach for deleting redundant elements which deletes fewer redundant elements, but also requires only a single round of adaptivity.

1.2.7 Efficiently Finding Redundant Elements

As mentioned above, our final method for recovering redundant elements recovers fewer redundant elements, but does so in only a single round, in a sense forming the analog of the procedure for recovering elements in a single part as outlined in Section 1.2.2. We encapsulate the behavior of this routine below:

Lemma 1.11. *Let \mathcal{M} be a matroid, and S be a greedily-optimal set. Then, there is a 1-round algorithm which recovers $\tilde{\Omega}\left(\min\left\{|S|, \frac{|S|^2}{\alpha(S)^2}\right\}\right)$ redundant elements.*

We omit a complete proof, but include the intuition here. Recall that because S is greedily-optimal, every element $x \in S$ satisfies $p_{\mathcal{M}|_S}(x) = \tilde{\Omega}(1/|S|)$ (i.e., they frequently appear in the first circuit that arises under random sampling). We create an even more fine-grained understanding: we define $p_{x,\ell}$ to be the probability that x appears in the first circuit that arises under random sampling *conditioned* on x being the ℓ -th element included. In particular, we can observe that if $\ell \gg \alpha(S)$, this probability is essentially 0, as the first circuit will already have formed by the time we add x . We can also see that these probabilities are monotonely decreasing, as x is only more likely to participate in the first circuit when it is added earlier. These two facts are enough to derive that $p_{x,1} = \tilde{\Omega}\left(\frac{|S|}{\alpha(S)} \cdot p_{\mathcal{M}|_S}(x)\right) = \tilde{\Omega}\left(\frac{|S|}{\alpha(S) \cdot |S|}\right)$.

Now, we consider the following simple process: we sample a random set A_1 of size approximately $100\alpha(S) \log |S| \gg \alpha(S)$. Now, for each element $x \in S$, we query the independence oracle with sets $\{x\} \cup W$, for W being every prefix of the set A_1 . Because A_1 is a completely random set, we see that the probability x participates in the first circuit as we add elements to W is essentially $p_{x,1}$. If x does appear in a circuit with W , then x is in fact a *redundant* element conditioned on A_1 . Thus, for our deletion procedure, we simply keep all the elements in A_1 , and remove all the elements *outside* A_1 that formed circuits with A_1 .

However, instead of stopping here, we repeat this procedure multiple times: we sample sets A_2, A_3, \dots and so on. Ultimately, we can only sample $\approx \frac{|S|}{\alpha(S)}$ many sets before *every* element is in

one of the A_i 's (this would be an issue, as we cannot remove the elements inside the A_i 's). But, this is enough to boost the deletion probability of each element x to be approximately

$$\Omega\left(\min\left(1, p_{x,1} \cdot \frac{|S|}{\alpha(S)}\right)\right) = \Omega\left(\min\left(1, \frac{|S|^2}{\alpha^2(S)|S|}\right)\right).$$

Summing across all elements in S , this yields the deletion of $\Omega\left(\min\left(|S|, \frac{|S|^2}{\alpha^2(S)}\right)\right)$ elements, as we desired.

With this routine now established, we are ready to complete the proof of our overall algorithm.

1.2.8 Putting the Pieces Together

As before, we consider running our decomposition procedure until there are $< n/2$ elements remaining. We let the recovered sets be denoted by S_1, \dots, S_k . By a simple pigeonhole argument, we are also guaranteed that there is some choice of $\ell \in [\log n]$ for which there are $\geq \frac{k}{\log n}$ of the sets S_1, \dots, S_k are of size $[2^\ell, 2^{\ell+1} - 1]$. We let τ denote this value 2^ℓ , and let γ denote this number of sets of size $[\tau, 2\tau]$. We let T refer to the indices of the sets whose sizes are in this range, and we let $\beta = \tau \cdot (\max_{i \in [k]} \alpha(S_i) / |S_i|)$.

To summarize the above subsections, we have the following methods of making progress:

1. From [Claim 1.9](#), we can invest a single additional round beyond those $k = \tilde{O}(\gamma)$ invested for the decomposition, and find an independent set of size $\Omega(\frac{n\beta}{\tau})$.
2. From [Theorem 1.10](#), we can invest $O(\sqrt{\tau})$ extra rounds and find $\gamma \cdot \tilde{\Omega}(\tau)$ redundant elements.
3. From [Lemma 1.11](#), we can invest a single additional round and find $\gamma \cdot \tilde{\Omega}\left(\min\left(\tau, \frac{\tau^2}{\beta^2}\right)\right)$ redundant elements.
4. Lastly, also from [Lemma 1.11](#), we can recover $\sum_{i=1}^k \tilde{\Omega}\left(|S_i|, \frac{\tau^2}{\beta^2}\right) = \tilde{\Omega}\left(\min\left(n, \frac{\tau^2}{\beta^2}\right)\right)$ redundant elements in a single additional rounds.

We summarize this below:

	Progress	Round Complexity
Claim 1.9	$\tilde{\Omega}(n\beta/\tau)$	$\tilde{O}(\gamma)$
Theorem 1.10	$\tilde{\Omega}(\gamma\tau)$	$\tilde{O}(\gamma + \tau^{1/2}) = \tilde{O}(\tau^{1/2})$
Lemma 1.11	$\tilde{\Omega}\left(\gamma \cdot \min\left(\tau, \frac{\tau^2}{\beta^2}\right)\right)$	$\tilde{O}(\gamma)$
Lemma 1.11	$\tilde{\Omega}\left(\min\left(n, \frac{\tau^2}{\beta^2}\right)\right)$	$\tilde{O}(\gamma)$

To obtain [Theorem 1.1](#), we simply do a case analysis based on β, τ and γ . We show that *for any* setting of these parameters, there is a choice of one of the above sub-routines which guarantees a Progress to Round ratio of at least $\tilde{\Omega}(n^{8/15})$. If we let κ denote the number of rounds invested in the sub-routine, this means the round complexity is governed by the recurrence $T(n) = \kappa + T(n - \kappa \cdot \tilde{\Omega}(n^{8/15}))$, and a simple calculation shows then that the algorithm terminates in $\tilde{O}(n^{7/15})$ rounds.

Remark 1.12. In fact, we can observe that [Theorem 1.10](#) relies on using the $O(\sqrt{n})$ round algorithm of [\[KUW85\]](#) to find bases of an arbitrary matroid. Now that we have an algorithm with complexity $\tilde{O}(n^{7/15})$ rounds, we can actually *improve* the round complexity of this sub-routine, and thereby achieve a better complexity than $\tilde{O}(n^{7/15})$ rounds in the global algorithm.

1.3 Organization

In [Section 2](#) we present some preliminary facts that we will make use of throughout our work. For the interested reader, in [Section 3](#) we present a complete analysis deriving [Theorem 1.3](#) in a stand-alone manner. Subsequent sections of the paper focus on the general case, and can be read independently of [Section 3](#). In [Section 4](#), we present a formal analysis of our decomposition algorithm. In [Section 5](#), we show how to make progress by recovering large independent sets. In [Section 6](#), we present a formal analysis of our techniques discussed above for recovering redundant elements. In [Section 7](#), we show how to trade-off between all of our subroutines for making progress to achieve our improved round complexity, thereby proving [Theorem 1.1](#).

2 Preliminaries

2.1 Notation

For a set S and an element x , we let $S + x = S \cup \{x\}$ and $S - x = S \setminus \{x\}$ for short. We use $\binom{S}{i}$ to denote the set of all subsets of S of cardinality i . For a matroid $\mathcal{M} = (E, \mathcal{I})$ and $S \subseteq E$, we denote $\mathcal{M}|_S$ as the matroid restricted to set S , and \mathcal{M}/S as the matroid after contracting S .

Throughout this paper, we denote E_0 as the ground set of the original matroid, E as the ground set of current matroid, and let $n_0 = |E_0|$, $n = |E|$. We always assume n_0 is sufficiently large.

2.2 Matroid Theory

Definition 2.1 (Matroids). A *matroid* $\mathcal{M} = (E, \mathcal{I})$ is a pair where E is a finite ground set and $\mathcal{I} \subseteq 2^E$ is a collection of independent sets with the following properties: (i) $\emptyset \in \mathcal{I}$ (non-triviality), (ii) for every $S \in \mathcal{I}$ and $S' \subset S$, $S' \in \mathcal{I}$ (downward-closedness), and (iii) for every $S, S' \in \mathcal{I}$ and $|S'| < |S|$, there exists some $x \in S \setminus S'$ such that $S + x \in \mathcal{I}$ (exchange property).

Definition 2.2 (Independent Sets, Circuits, Bases). For a matroid $\mathcal{M} = (E, \mathcal{I})$, we say a set $S \subseteq E$ is *independent* if $S \in \mathcal{I}$ and *dependent* otherwise. We call a set B a *basis* if it is a maximal independent set, i.e. for any $x \notin B$, $B + x \notin \mathcal{I}$. We call a set C a *circuit* if it is a minimal dependent set, i.e. for any $x \in C$, $C - x \in \mathcal{I}$.

We have the following fact:

Fact 2.3. For a matroid $\mathcal{M} = (E, \mathcal{I})$ and $S \subseteq E$, $x \in E \setminus S$, if $S \in \mathcal{I}$ and $S + x \notin \mathcal{I}$, then there is a unique circuit C in $S + x$ where $x \in C$. Moreover, for every $y \in C \setminus x$, $S - y + x \in \mathcal{I}$.

Definition 2.4 (Rank). For a matroid $\mathcal{M} = (E, \mathcal{I})$, we define the *rank* of \mathcal{M} as $\text{rank}(\mathcal{M}) = \max_{S \in \mathcal{I}} |S|$. Further, for any $S \subseteq E$, we define $\text{rank}_{\mathcal{M}}(S) = \max_{T \subseteq S, T \in \mathcal{I}} |T|$. The rank function of a matroid is submodular.

Definition 2.5 (Span, Flats). In a matroid $\mathcal{M} = (E, \mathcal{I})$, we define $\text{span}(S)$ as

$$\text{span}(S) = \{x \in E \mid \text{rank}(S \cup \{x\}) = \text{rank}(S)\}.$$

We say that a set $S \subseteq E$ is a *flat* if $S = \text{span}(S)$. Furthermore, when a set S is a flat and satisfies $\text{rank}(S) = \text{rank}(\mathcal{M}) - 1$, we call S a *hyperplane*.

Definition 2.6 (Quotients). Let $\mathcal{M} = (E, \mathcal{I})$ be a matroid. Then, the set $Q = \{E \setminus S : S \text{ is a flat in } \mathcal{M}\}$ is called the *set of quotients* in a matroid.

Definition 2.7 (Dual Matroid). Let $\mathcal{M} = (E, \mathcal{I})$ be a matroid, the dual matroid $\mathcal{M}^* = (E, \mathcal{I}^*)$ is defined as

$$\mathcal{I}^* = \{S \subseteq E \mid \exists B \subseteq (E \setminus S) \text{ s.t. } B \text{ is a basis of } \mathcal{M}\}.$$

In particular, $\text{rank}(\mathcal{M}^*) = |E| - \text{rank}(\mathcal{M})$, and $(\mathcal{M}^*)^* = \mathcal{M}$.

There are also the following useful facts about dual matroids:

Fact 2.8. [Oxl06] Let $\mathcal{M} = (E, \mathcal{I})$ be a matroid, and $S \subseteq E$. Then,

$$(\mathcal{M}/S)^* = \mathcal{M}^* \setminus S.$$

Furthermore, for all $T \subseteq E \setminus S$

$$\text{rank}_{\mathcal{M}/S}(T) = \text{rank}_{\mathcal{M}}(T \cup S) - \text{rank}_{\mathcal{M}}(S).$$

Fact 2.9. [Oxl06] Let $\mathcal{M} = (E, \mathcal{I})$ be a matroid, and let \mathcal{M}^* be its dual matroid. A set of elements $C \subseteq E$ is a circuit in \mathcal{M} if and only if $E \setminus C$ is a hyperplane in \mathcal{M}^* .

2.3 Probability Theory

In the analysis of our algorithm, we will need an anti-concentration bound for the hypergeometric distribution. We will approximate the hypergeometric distribution using a binomial distribution, for which the following anti-concentration bound is established in the literature:

Lemma 2.10 ([Doe20]). Let $X \sim \text{Bin}(n, p)$ with $p \leq 1/2$, $\text{Var}[X] = np(1-p) \geq 1$, then

$$\Pr \left[X \geq \mathbb{E}[X] + \frac{1}{5} \sqrt{\text{Var}[X]} \right] \geq \frac{1}{108}, \quad \Pr \left[X \leq \mathbb{E}[X] - \frac{1}{5} \sqrt{\text{Var}[X]} \right] \geq \frac{1}{108}$$

We formally quantify the difference between the hypergeometric distribution and binomial distribution by bounding the total variation distance between them.

Definition 2.11 (Total Variation Distance). Let P, Q be two probability distribution on \mathbb{Z} , the total variation distance between P and Q is defined as

$$\delta(P, Q) = \sup_{A \subseteq \mathbb{Z}} |P(A) - Q(A)|.$$

We use the following result from [Ehm91].

Lemma 2.12 ([Ehm91]). For $\text{Hyp}(n, m, k)$ and $\text{Bin}(k, p)$ where $p = \frac{m}{n}$ and $kp(1-p) \geq 1$,

$$\delta(\text{Hyp}(n, m, k), \text{Bin}(k, p)) \leq \frac{k-1}{n-1}.$$

3 Tight Bounds for Partition Matroids

Recall the definition of partition matroids.

Definition 3.1 (Partition Matroid). A partition matroid $\mathcal{M} = (E, \mathcal{I})$ is defined by a ground set E being partitioned into disjoint sets A_1, \dots, A_m and m integers b_1, \dots, b_m where $0 \leq b_i \leq |A_i|$. A set $S \subseteq E$ is independent iff $|S \cap A_i| \leq b_i$ for every $0 \leq i \leq m$. We refer to A_i as a *part*, and b_i as the *budget* of the part.

In this section, we present an $O(n^{1/3} \log n)$ round algorithms for finding bases in partition matroids, which matches the lower bound $\tilde{\Omega}(n^{1/3})$ of [KUW85] up to logarithmic factors.

3.1 A Randomized Algorithm

In a partition matroid $\mathcal{M} = (E, \mathcal{I})$ with $|E| = n$, recovering a single part in each round is straightforward, but very inefficient as there can be $\Omega(n)$ parts. The main idea of our algorithm is to recover multiple parts simultaneously in a single round. Specifically, if we consider adding elements in a random order until the budget of some part is exceeded, we will show that we can efficiently identify and recover this part. If we repeat this $\text{poly}(n)$ times in parallel, we can collect and identify all the parts that ever caused a dependency and remove them from the matroid. Intuitively, in the next iteration, we expect that the dependence will occur later when adding elements in random order, as the parts that are more likely to cause circuits have already been removed (i.e., the small budget parts). Thus, the expected number of elements needed to observe a dependency increases. We formally quantify this growth and thereby show that in $\tilde{O}(n^{1/3})$ rounds, we can recover all parts in the matroid.

We begin by presenting a simple algorithm to recover a single part. The algorithm adds elements according to the order of a permutation π , and stops once a dependence occurs. This implies that for exactly one part, the budget has been exceeded by 1, and allows us to uniquely identify this part.

Algorithm 1: RecoverSinglePart($\mathcal{M} = (E, \mathcal{I}), \pi : \text{bijection } [n] \rightarrow E$)

```

1 for  $i \in [n]$  in parallel do
2   | Query Ind( $\{\pi(1), \dots, \pi(i)\}$ )
3 end
4 Let  $t$  be the smallest index such that  $S = \{\pi(1), \dots, \pi(t-1)\} \in \mathcal{I}$  and  $S + \pi(t) \notin \mathcal{I}$ 
5  $I \leftarrow \emptyset$ 
6 for  $i = 1, \dots, t-1$  do
7   | Query Ind( $S - \pi(i) + \pi(t)$ )
8   | if Ind( $S - \pi(i) + \pi(t)$ ) = 1 then
9     |    $I \leftarrow I \cup \pi(i)$ 
10  | end
11 end
12  $T \leftarrow I \cup \pi(t)$ 
13 for  $i = t+1, \dots, n$  do
14   | Query Ind( $I + \pi(i)$ )
15 end
16 if Ind( $I + \pi(i)$ ) = 1 then
17   |  $T \leftarrow T \cup \pi(i)$ 
18 end
19 return  $T$  and  $|I|$ 

```

Claim 3.2. In a partition matroid \mathcal{M} where every part has budget ≥ 1 , [Algorithm 1](#) finds the part that contains $\pi(t)$ (t as defined in [Line 4](#)), and can be implemented in 2 rounds.

Proof. Let A_ℓ be the part that contains $\pi(t)$. Given that $\text{Ind}(S) = 1$ and $\text{Ind}(S + \pi(t)) = 0$, it follows that $|S \cap A_\ell| = b_\ell$ and $|S \cap A_j| \leq b_j$ for every $j \neq \ell$. Therefore, for every $i < t$, $S - \pi(i) + \pi(t)$ is independent if and only if $\pi(i) \in A_\ell$. As $|S \cap A_\ell| = b_\ell$, we have $I \subseteq A_\ell$ and $|I| = b_\ell$. Moreover, for every $i > t$, we see that $I + \pi(i)$ is dependent if and only if $\pi(i) \in A_\ell$. Thus, we conclude that $T = A_\ell$ and $|I| = b_\ell$.

Note that the only adaptivity we need is to find I . The for loops before [Line 12](#) can be implemented in 1 round in parallel (by querying all prefixes and prefixes with one element excluded), and the for loop after [Line 12](#) can also be implemented in 1 round. \square

We also show a similar subroutine to recover every part with budget at most 50.

Claim 3.3. *Algorithm 2 eliminates all parts with budget at most 50, and can be implemented in 1 round.*

Algorithm 2: RemoveSmallParts($\mathcal{M} = (E, \mathcal{I})$)

```

1 for  $i \in [50], S \in \binom{E}{i}$  in parallel do
2   | Query Ind( $S$ )
3 end
4  $B \leftarrow \emptyset$ 
5 for  $i \in [49]$  do
6   | for  $S \in \binom{E}{i}, x \in E \setminus S$  do
7     | if Ind( $S$ ) = 1  $\wedge$  Ind( $S + x$ ) = 0 then
8       |    $T \leftarrow S$  for  $y \in E \setminus S$  do
9         |     if Ind( $S + y$ ) = 1 then
10          |        $T \leftarrow T + y$ 
11          |     end
12          | end
13          |  $B \leftarrow B \cup S$ 
14          |  $\mathcal{M} \leftarrow \mathcal{M} \setminus T$ 
15        | end
16      | end
17 end
18 return  $\mathcal{M}, B$ 

```

Now we are ready to present our main procedure, detailed in [Algorithm 3](#). It begins by invoking [Algorithm 2](#) to remove parts with small budgets. This is for a technical reason in the analysis of our algorithm. It also checks the trivial case that the whole ground set is independent. Then, the algorithm draws $\text{poly}(n_0)$ random permutations π in parallel and checks if the first $n/1000$ elements of π form an independent set. If the set is independent, the algorithm adds it to the solution, contracts it from the matroid, and then terminates. Otherwise, we invoke [Algorithm 1](#) to recover the part which contains the first circuit that appears in the matroid when elements are added in the order of π . The recovered parts are recorded and removed from the matroid at the end of each iteration.

Let S_1, \dots, S_k be the sets recovered by [Algorithm 3](#) in k iterations. Note that each S_i is the union of one or more parts. For every $i \in [k]$, we define $\alpha(S_i)$ as the smallest integer ℓ such that

$$\Pr_{T \in \binom{S_i}{\ell}} [\text{Ind}(T) = 1] \leq \frac{1}{2}.$$

I.e., the probability that a random subset of S_i of cardinality ℓ is independent is less than $1/2$. It follows that within S_i , a random subset of cardinality less than $\alpha(S_i)$ is independent with probability at least $1/2$, and a random subset of cardinality at least $\alpha(S_i)$ is dependent with probability at least $1/2$.

Algorithm 3: RecoverMultipleParts($\mathcal{M} = (E, \mathcal{I})$)

```
1  $\mathcal{M}, B \leftarrow \text{RemoveSmallParts}(\mathcal{M})$ 
2 if  $\text{Ind}(E) = 1$  then
3   | return  $\emptyset, E$ 
4 end
5  $k \leftarrow 0$ 
6 while  $\mathcal{M} \neq \emptyset$  do
7   |  $\mathcal{A} \leftarrow \emptyset, I \leftarrow \emptyset$ 
8   | for  $i \in [n_0^{10}]$  in parallel do
9     | Draw a random permutation (bijection)  $\pi : [n] \rightarrow E$ 
10    | Query  $\text{Ind}(\{\pi(1), \dots, \pi(n/1000)\})$ 
11    | if  $\text{Ind}(\{\pi(1), \dots, \pi(n/1000)\}) = 1$  then
12      | |  $I \leftarrow \{\pi(1), \dots, \pi(n/1000)\}$ 
13    | end
14    |  $T, \ell \leftarrow \text{RecoverSinglePart}(\mathcal{M}, \pi)$ 
15    |  $\mathcal{A} \leftarrow \mathcal{A} \cup \{(T, \ell)\}$ 
16  | end
17  | if  $I \neq \emptyset$  then
18    | | return  $\mathcal{M}/I, B \cup I$ 
19  | end
20  |  $k \leftarrow k + 1, S_k \leftarrow \emptyset$ 
21  | for  $(T, \ell) \in \mathcal{A}$  do
22    | |  $S_k \leftarrow S_k \cup T$ 
23    | | Pick an arbitrary  $I \in \binom{T}{\ell}, B \leftarrow B \cup I$ 
24  | end
25  |  $\mathcal{M} \leftarrow \mathcal{M} \setminus S_k$ 
26 end
27 return  $B$ 
```

We analyze the performance of [Algorithm 3](#) in the following claims. Note that the algorithm requires $O(k)$ rounds of adaptivity, as all queries are made within the for loop on [Line 8](#) and can be executed in parallel, where each parallel instance requires $O(1)$ rounds by [Claim 3.2](#). Thus, our ultimate goal is to bound $k = \tilde{O}(n^{1/3})$.

Claim 3.4. For every $i \in [k]$, $\alpha(S_i) \geq 50$.

Proof. Recall that we invoke [Algorithm 2](#) to remove all parts with budget less than or equal to 50. Since in a partition matroid, the size of a dependent set must be at least the budget of some part, we see that $\alpha(S_i) \geq 50$ according to the definition of α . \square

Claim 3.5. In the i -th iteration of [Algorithm 3](#), let A_1, \dots, A_ℓ be the remaining parts. For every $j \in [\ell]$, let p_j denote the probability that [Algorithm 1](#), when given a random permutation, returns A_j . Then with probability at least $1 - 2^{-n_0^7}$, we have for any $j \in [\ell]$ with $p_j \geq 1/n^2$, $A_j \subseteq S_i$.

Proof. Consider any $j \in [\ell]$ with $p_j \geq 1/n^2$, the probability that the algorithm fails to find it in the n_0^{10} parallel instances of [Algorithm 1](#) is at most

$$\left(1 - \frac{1}{n^2}\right)^{n_0^{10}} \leq e^{-n_0^8}.$$

Taking a union bound over at most $\ell \leq n \leq n_0$ parts, we see that the algorithm fails to find any such part with probability at most $e^{-n_0^8} \cdot n_0 \leq 2^{-n_0^7}$. \square

It immediately follows from the above claim that with probability at least $1 - 1/n$, the first circuit will appear inside S_i when adding elements according to the order of a random permutation π .

Claim 3.6. *In the i -th iteration of [Algorithm 3](#), suppose with probability at least $1 - 1/n$, the first circuit appears inside S_i when adding elements according to the order of a random permutation π . If*

$$\frac{\alpha(S_i)}{|S_i|} \geq \frac{1}{250},$$

then the algorithm will terminate on [Line 18](#) in this iteration with probability at least $1 - 2^{-n_0^9}$.

Proof. Let $\ell = \frac{\alpha(S_i)}{4|S_i|}n \geq n/1000$. Suppose U is a random subset of E of cardinality ℓ , we define random variable $X_i = |U \cap S_i|$. Note that $X_i \sim \text{Hyp}(n, |S_i|, \ell)$, and $\mathbb{E}[X_i] = \alpha(S_i)/4$. By Markov's inequality,

$$\Pr[X_i \geq \alpha(S_i)] \leq \frac{1}{4}.$$

Since $\ell \geq n/1000$, we see that

$$\begin{aligned} \Pr_{\pi}[\text{Ind}(\{\pi(1), \dots, \pi(n/1000)\}) = 0] &\leq \Pr_{\pi}[\text{Ind}(\{\pi(1), \dots, \pi(\ell)\}) = 0] \\ &\leq \frac{1}{n} + \Pr_U[\text{Ind}(U \cap S_i) = 0] \\ &\leq \frac{1}{n} + \left(\Pr_{T \in \binom{S_i}{X_i}}[\text{Ind}(T) = 0 \mid X_i < \alpha(S_i)] + \Pr[X_i \geq \alpha(S_i)] \right) \\ &\leq \frac{1}{n} + \frac{1}{2} + \frac{1}{4} \leq \frac{7}{8}. \end{aligned}$$

Once the first $n/1000$ elements of π form an independent set, the algorithm will contract the set and terminate on [Line 18](#). Therefore, the probability that the algorithm does not terminate is at most $(7/8)^{n_0^{10}} \leq 2^{-n_0^9}$. \square

By taking a simple union bound, the algorithm achieves all the above guarantees with probability at least $1 - 2^{n_0^7} - 2_{n_0}^9 \geq 1 - 2^{n_0^6}$. We make subsequent claims conditioned on this event.

Claim 3.7. *Let S_1, \dots, S_k be the sets obtained by [Algorithm 3](#). For any $i < j$ where $|S_i| \leq 2|S_j|$ and $|S_i|, |S_j| \leq n^{2/3}$, we have*

$$\alpha(S_j) \geq \frac{\alpha(S_i)|S_j|}{|S_i|} + \frac{1}{5\sqrt{2}} \sqrt{\frac{\alpha(S_i)|S_j|}{|S_i|}}.$$

Proof. Let $\ell = \frac{\alpha(S_i)}{|S_i|}n$. Suppose U is a subset of the ground set E of cardinality ℓ , we define random variables $X_i = |U \cap S_i|$, $X_j = |U \cap S_j|$. Note that $X_i \sim \text{Hyp}(n, |S_i|, \ell)$, $X_j \sim \text{Hyp}(n, |S_j|, \ell)$ and they are negatively correlated.

For the sake of contradiction, suppose

$$\alpha(S_j) \leq \frac{\alpha(S_i)|S_j|}{|S_i|} + \frac{1}{5\sqrt{2}} \sqrt{\frac{\alpha(S_i)|S_j|}{|S_i|}}.$$

We aim to show that

$$\Pr[X_i < \alpha(S_i) \wedge X_j \geq \alpha(S_j)] = \Omega(1).$$

Consider the i -th iteration of [Algorithm 3](#), the above implies that for a random permutation π of $[n]$, there is a constant probability that there are less than $\alpha(S_i)$ elements from S_i and more than $\alpha(S_j)$ elements from S_j in the first ℓ elements of π . Conditioned on this, with at least $1/4$ probability, there is no circuit within S_i but there is a circuit in S_j . This implies that with $\Omega(1)$ probability, the first circuit appears outside of S_i when adding elements according to the order of a random permutation π , which contradicts [Claim 3.5](#).

To estimate $X_i \sim \text{Hyp}(n, |S_i|, \ell)$, we define $Y_i \sim \text{Bin}(\ell, |S_i|/n)$. It gives a good estimation since $\ell/n = \alpha(S_i)/|S_i| \leq 1/250$ by [Claim 3.6](#). Since $|S_i| \leq n^{2/3}$, we see that

$$\mathbb{E}[Y_i] = \alpha(S_i), \quad \mathbf{Var}[Y_i] = \ell \frac{|S_i|}{n} \left(1 - \frac{|S_i|}{n}\right) \geq \frac{\alpha(S_i)}{2} \geq 25$$

It follows from [Lemma 2.10](#) that

$$\Pr[Y_i < \alpha(S_i)] = \Pr[Y_i \leq \mathbb{E}[Y_i] - 1] \geq \Pr\left[Y_i \leq \mathbb{E}[Y_i] - \frac{1}{5}\sqrt{\mathbf{Var}[Y_i]}\right] \geq \frac{1}{108}.$$

Combined with [Lemma 2.12](#) and $\ell/n \leq 1/250$, we conclude that

$$\Pr[X_i < \alpha(S_i)] \geq \frac{1}{108} - \frac{\ell - 1}{n - 1} \geq \frac{1}{200}.$$

To estimate $X_j \sim \text{Hyp}(n, |S_j|, \ell)$, we also define $Y_j \sim \text{Bin}(\ell, |S_j|/n)$. Since $|S_j| \leq n^{2/3}$, $|S_i| \leq 2|S_j|$, we see that

$$\mathbb{E}[Y_j] = \frac{\alpha(S_i)|S_j|}{|S_i|}, \quad \mathbf{Var}[Y_j] = \ell \frac{|S_j|}{n} \left(1 - \frac{|S_j|}{n}\right) \geq \frac{\alpha(S_i)|S_j|}{2|S_i|} \geq \frac{\alpha(S_i)}{4} \geq 1.$$

It follows from [Lemma 2.10](#) that

$$\Pr\left[Y_j \geq \mathbb{E}[Y_j] + \frac{1}{5}\sqrt{\mathbf{Var}[Y_j]}\right] \geq \frac{1}{108}.$$

Combined with [Lemma 2.12](#) and $\ell/n \leq 1/250$, we have

$$\Pr\left[X_j \geq \mathbb{E}[Y_j] + \frac{1}{5}\sqrt{\mathbf{Var}[Y_j]}\right] \geq \frac{1}{108} - \frac{\ell - 1}{n - 1} \geq \frac{1}{200}.$$

We conclude that

$$\Pr[X_j \geq \alpha(S_j)] \geq \Pr\left[X_j \geq \frac{\alpha(S_i)|S_j|}{|S_i|} + \frac{1}{5\sqrt{2}}\sqrt{\frac{\alpha(S_i)|S_j|}{|S_i|}}\right] \geq \Pr\left[X_j \geq \mathbb{E}[Y_j] + \frac{1}{2}\sqrt{\mathbf{Var}[Y_j]}\right] \geq \frac{1}{200}.$$

The penultimate inequality follows from

$$\mathbb{E}[Y_j] = \frac{\alpha(S_i)|S_j|}{|S_i|}, \quad \mathbf{Var}[Y_j] \geq \frac{\alpha(S_i)|S_j|}{2|S_i|}.$$

Since X_i, X_j are negatively correlated, we have

$$\Pr[X_i < \alpha(S_i) \wedge X_j \geq \alpha(S_j)] \geq \Pr[X_i < \alpha(S_i)] \cdot \Pr[X_j \geq \alpha(S_j)] \geq \frac{1}{200^2}.$$

as desired. \square

Lemma 3.8. *Throughout the course of [Algorithm 3](#), $k = O(n^{1/3})$.*

Proof. Consider S_1, \dots, S_k obtained from [Algorithm 3](#). We first note that there are at most $n^{1/3}$ S_i 's with cardinality at least $n^{2/3}$. Therefore, we will drop every set of cardinality larger than $n^{2/3}$ and assume that $|S_i| \leq n^{2/3}$ for every $i \in [k]$.

For any $\ell \in [2/3 \cdot \log n]$, let $S_{i_1}, \dots, S_{i_{q(\ell)}}$ be all the S_i 's with $|S_i| \in [2^{\ell-1}, 2^\ell - 1]$ and assume $i_1 < \dots < i_{q(\ell)}$. As $|S_{i_j}| \leq 2|S_{i_{j+1}}|$ and $|S_{i_j}|, |S_{i_{j+1}}| \leq n^{2/3}$, it follows from [Claim 3.7](#) that

$$\frac{\alpha(S_{i_{j+1}})}{|S_{i_{j+1}}|} \geq \frac{\alpha(S_{i_j})}{|S_{i_j}|} + \frac{1}{5\sqrt{2}} \sqrt{\frac{\alpha(S_{i_j})}{|S_{i_j}| |S_{i_{j+1}}|}} \geq \frac{\alpha(S_{i_j})}{|S_{i_j}|} + \frac{1}{5\sqrt{2} \cdot 2^{\ell/2}} \sqrt{\frac{\alpha(S_{i_j})}{|S_{i_j}|}}.$$

Since $\alpha(S_{i_1})/|S_{i_1}| \geq 1/n$, $\alpha(S_{i_{q(\ell)}})/|S_{i_{q(\ell)}}| \leq 1$, the recursion gives $q(\ell) = O(2^{\ell/2})$. Therefore,

$$k \leq \sum_{\ell=0}^{\frac{2}{3} \log n} q(\ell) = \sum_{\ell=0}^{\frac{2}{3} \log n} O(2^{\ell/2}) = O(n^{1/3}).$$

This concludes the proof. □

We obtain our algorithm for finding a basis by iteratively executing [Algorithm 3](#).

Algorithm 4: FindBasis($\mathcal{M} = (E, \mathcal{I})$)

```

1  $B \leftarrow \emptyset$ 
2 while  $\mathcal{M} \neq \emptyset$  do
3    $\mathcal{M}, I \leftarrow \text{RecoverMultipleParts}(\mathcal{M})$ 
4    $B \leftarrow B \cup I$ 
5 end
6 return  $B$ 
```

Theorem 3.9. *For a partition matroid $\mathcal{M} = (E, \mathcal{I})$ with $|E| = n$, [Algorithm 4](#) requires $O(n^{1/3} \log n)$ rounds to recover a basis in \mathcal{M} with high probability.*

Proof. Throughout the algorithm, we either contract an independent set of size $\Omega(n)$ or recover parts. When a part is recovered, an independent set of size equal to its budget is always included before removing the part from the matroid. Therefore, the final set obtained is indeed a basis of the matroid.

In each execution of [Algorithm 3](#), it either finds an independent set and terminates at [Line 18](#), or it successfully recovers the entire matroid. If it terminates at [Line 18](#), it implies that an independent set of size $n/1000$ has been found and the matroid has been contracted. It is straightforward to check that the matroid remains a partition matroid after contraction. Since this can happen $O(\log n)$ times, [Algorithm 3](#) is executed at most $O(\log n)$ times. Combined with [Lemma 3.8](#), the total rounds of adaptivity required is $O(n^{1/3} \log n)$. By a simple union bound, we can also bound the total failure probability by $2^{-\Omega(n)}$. □

3.2 Derandomization

We also show that the above algorithm can be derandomized.

Claim 3.10. *There exists a universal family of permutations $\pi_1, \dots, \pi_{n_0^{10}}$ such that for any partition matroid $\mathcal{M} = (E, \mathcal{I})$ with $|E| = n \leq n_0$, the guarantees of [Claim 3.5](#) and [Claim 3.6](#) are satisfied after running [Algorithm 3](#) using these permutations instead of random permutations.*

Proof. By [Claim 3.5](#) and [Claim 3.6](#), we see that a random collection of $\pi_1, \dots, \pi_{n_0^{10}}$ achieves the guarantees with probability at least $1 - 2^{-n_0^6}$. Since there are at most $n^n \cdot n! \leq 2^{n^2}$ possible partition matroids, we obtain that a random collection of $\pi_1, \dots, \pi_{n_0^{10}}$ achieves the guarantees for any partition matroid with probability at least $1 - 2^{-n_0^6} \cdot 2^{n^2} \geq 1 - 2^{-n_0^5}$. This implies that there exists a deterministic choice of $\pi_1, \dots, \pi_{n_0^{10}}$ that achieves the guarantees. \square

Theorem 3.11. *There is a deterministic algorithm that finds a basis of any partition matroid \mathcal{M} in $\tilde{O}(n^{1/3})$ adaptive rounds, using only polynomially many independence queries per round.*

Proof. For every $i \in [n]$, by [Claim 3.10](#), there exists a feasible family of permutations $\mathcal{P}_i = \{\pi_1, \dots, \pi_{n_0^{10}}\}$ for matroids on i elements, which we encode non-uniformly. The remainder of the algorithm is identical to [Algorithm 4](#), except that we replace the random permutations in [Algorithm 3](#) with the deterministic family of permutations obtained earlier. \square

Note that this derandomization is non-uniform. While such derandomizations are typical in some settings, such as $\mathbf{BPP} \subseteq \mathbf{P} / \mathbf{poly}$, non-uniform derandomizations are not always possible with query complexity bounds, where randomization can sometimes be essential.

4 Decomposition Algorithm

In this section, we present our decomposition algorithm, and derive all of the relevant results that we will need when designing our improved algorithm for finding bases.

4.1 Finding Sets with Many Circuits

To start, we present an algorithm which adds elements in the order of a permutation π until a circuit forms:

Algorithm 5: FindCircuit($\mathcal{M} = (E, \mathcal{I}), \pi : \text{bijection } [n] \rightarrow E$)

```

1 for  $i \in [n]$  in parallel do
2   | Query Ind( $\{\pi(1), \dots, \pi(i)\}$ )
3 end
4 Let  $t$  be the smallest index such that  $S = \{\pi(1), \dots, \pi(t-1)\} \in \mathcal{I}$  and  $S + \pi(t) \notin \mathcal{I}$ .
5  $C_\pi \leftarrow \{\pi(t)\}$ 
6 for  $i = 1, \dots, t-1$  in parallel do
7   | Query Ind( $S - \pi(i) + \pi(t)$ ).
8   | if Ind( $S - \pi(i) + \pi(t)$ ) = 1 then
9     |    $C_\pi \leftarrow C_\pi + \pi(i)$ 
10  | end
11 end
12 return  $C_\pi$ 

```

The following claim follows immediately from [Fact 2.3](#).

Claim 4.1. *Algorithm 5 can be implemented in 1 round and returns the first circuit appears when adding elements in the order of a permutation π .*

Using the above procedure, we define the following quantities:

Definition 4.2. For a matroid $\mathcal{M} = (E, \mathcal{I})$ and an element $x \in E$, we let

$$p_{\mathcal{M}}(x) = \Pr_{\pi}[x \in \text{FindCircuit}(\mathcal{M}, \pi)]$$

Similarly, for a subset $S \subseteq E$, we let

$$q_{\mathcal{M}}(S) = \Pr_{\pi}[\text{FindCircuit}(\mathcal{M}, \pi) \subseteq S]$$

We let $\widehat{q}_{\mathcal{M}}(S)$ denote the estimate of this probability that results from running Algorithm 5 on n_0^{10} random permutations.

Remark 4.3. We sometimes omit the subscript \mathcal{M} when the underlying matroid is clear from the context.

A simple application of a Chernoff bound yields the following statement:

Claim 4.4. *For a matroid $\mathcal{M} = (E, \mathcal{I})$ and every subset $S \subseteq E$,*

$$|\widehat{q}(S) - q(S)| \leq \frac{1}{n_0^2},$$

with probability $1 - 2^{-n_0}$.

Remark 4.5. Note that because the error probability is $1 - 2^{-n_0}$, we will often present intermediate claim / theorem statements without quantifying their success probability. Ultimately, our algorithm will only ever perform $\text{poly}(n_0)$ invocations of the decomposition, and thus this error probability is negligible.

Definition 4.6. For a matroid $\mathcal{M} = (E, \mathcal{I})$ and $S \subseteq E$, we define $\alpha(S)$ as the smallest integer ℓ such that

$$\Pr_{T \in \binom{S}{\ell}}[\text{Ind}(T) = 1] \leq \frac{1}{2}.$$

I.e., the probability that a random subset of S of cardinality ℓ is independent is less than $1/2$. Equivalently, $\alpha(S)$ is the median number of elements required until a circuit appears when running $\text{FindCircuit}(\mathcal{M}|_S, \pi)$ on a random permutation π . We also use $\widehat{\alpha}(S)$ to denote the estimate of $\alpha(S)$ results from running Algorithm 5 on n_0^{10} random permutations and taking median.

We have the following useful properties.

Claim 4.7. *For any set S and any integer $d > 1$, a random subset of S of cardinality $d \cdot \alpha(S)$ is dependent with probability at least $1 - 2^{-d}$.*

Proof. Let T be a random subset of S of cardinality $d \cdot \alpha(S)$, and R_1, \dots, R_d be independently drawn random subsets of S of cardinality $\alpha(S)$. We see that

$$\Pr[\text{Ind}(T) = 1] \leq \Pr\left[\text{Ind}\left(\bigcup_{i \in [d]} R_i\right) = 1\right] \leq \prod_{i=1}^d \Pr[\text{Ind}(R_i) = 1] \leq (1/2)^d.$$

The last inequality follows from the definition of $\alpha(S)$. □

Claim 4.8. $(\alpha(S) - 1)/2 \leq \hat{\alpha}(S) \leq 2 \cdot \alpha(S)$ with probability at least $1 - 2^{-n_0}$.

Proof. By Claim 4.7, the probability that a random subset of S of cardinality $2 \cdot \alpha(S)$ is independent is at most $1/4$. Thus, we see that $\hat{\alpha}(S) \leq 2 \cdot \alpha(S)$ with high probability by a straightforward Chernoff bound. On the other hand, we show that the probability a random subset of S of cardinality $(\alpha(S) - 1)/2$ is independent is at least $1/\sqrt{2}$. This is due to an argument similar to Claim 4.7: let T be a random subset of S of cardinality $\alpha(S) - 1$ and R_1, R_2 be 2 independently drawn random subsets of S of cardinality $(\alpha(S) - 1)/2$. We have

$$\frac{1}{2} \leq \Pr[\text{Ind}(T) = 1] \leq \Pr[\text{Ind}(R_1 \cup R_2) = 1] \leq \Pr[\text{Ind}(R_1) = 1]^2,$$

and thus $\Pr[\text{Ind}(R_1) = 1] \geq 1/\sqrt{2}$. Again, we have $\hat{\alpha}(S) \geq (\alpha(S) - 1)/2$ with high probability by a Chernoff bound. \square

Now, we will let $S^* \subseteq E$ be a *greedily-optimal* set in the following sense:

Definition 4.9. For a matroid $\mathcal{M} = (E, \mathcal{I})$, we say that a set $S^* \subseteq E$ is *greedily-optimal* if

$$\hat{q}(S^*) \geq 1 - 2^{-20} + \frac{\sum_{i=1}^{|S^*|} \frac{1}{i}}{2^{20} \log n},$$

and there is no element $x \in S^*$ such that

$$\hat{q}(S^* - x) \geq 1 - 2^{-20} + \frac{\sum_{i=1}^{|S^*|-1} \frac{1}{i}}{2^{20} \log n}.$$

Observe that there is a simple algorithm for creating greedily-optimal sets: we start with the complete set n , and continue to delete elements until the condition no longer holds:

Algorithm 6: FindGreedlyOptimal($\mathcal{M} = (E, \mathcal{I})$)

```

1 Multiset  $\mathcal{C} \leftarrow \emptyset$ 
2 for  $i \in [n_0^{10}]$  in parallel do
3   | Draw a random permutation (bijection)  $\pi : [n] \rightarrow E$ 
4   |  $C_\pi \leftarrow \text{FindCircuit}(\mathcal{M}, \pi)$ 
5   |  $\mathcal{C} \leftarrow \mathcal{C} \cup \{C_\pi\}$ 
6 end
7  $S^* \leftarrow E$ .
8 while True do
9   | for  $x \in S^*$  do
10    |  $\hat{q}(S^* - x) \leftarrow \frac{|\{C \in \mathcal{C} : C \subseteq S^* - \{x\}\}|}{|\mathcal{C}|}$ 
11    | end
12    | if  $\exists x \in S^*$  s.t.  $\hat{q}(S^* - x) \geq 1 - 2^{-20} + \frac{\sum_{i=1}^{|S^*|-1} \frac{1}{i}}{2^{20} \log n}$  then
13      | Let  $x^*$  be the first such element
14      |  $S^* \leftarrow S^* - x^*$ 
15    | end
16    | else
17      | return  $S^*$ 
18    | end
19 end

```

Claim 4.10. *Algorithm 6 finds a greedily-optimal set S^* in \mathcal{M} .*

Proof. Note that at the initialization of Algorithm 6, $S^* = E$, and so it must be the case that $1 = \widehat{q}(S^*) = \widehat{q}(E) \geq 1 - 2^{-20} + \frac{\sum_{i=1}^n \frac{1}{i}}{2^{20} \log n}$, as

$$1 - 2^{-20} + \frac{\sum_{i=1}^n \frac{1}{i}}{2^{20} \log n} \leq 1 - 2^{-20} + \frac{1 + \ln n}{2^{20} \log n} \leq 1 - 2^{-20} + \frac{1}{2^{20} \log n} + \frac{\log n}{2^{20} \log(e)} < 1.$$

Now, in each iteration of Algorithm 6, we only remove elements x from S^* that ensure that S^* continues to satisfy

$$\widehat{q}(S^*) \geq 1 - 2^{-20} + \frac{\sum_{i=1}^{|S^*|} \frac{1}{i}}{2^{20} \log n}.$$

At the termination of the above algorithm, we are also guaranteed that there is no $x \in S^*$ for which

$$\widehat{q}(S^* - x) \geq 1 - 2^{-20} + \frac{\sum_{i=1}^{|S^*|-1} \frac{1}{i}}{2^{20} \log n},$$

thereby yielding our claim. \square

Now, we establish the following claim, which seeks to understand the *marginal* probabilities that an element $x \in S^*$ participates in a circuit.

Claim 4.11. *Let $\mathcal{M} = (E, \mathcal{I})$ be a matroid on n elements, and let S^* be a greedily-optimal set. Let $\mathcal{M}' = \mathcal{M}|_{S^*}$ be the matroid restricted to S^* . Then, for every $x \in S^*$,*

$$p_{\mathcal{M}'}(x) \geq \frac{1}{2 \cdot 2^{20} |S^*| \log n}.$$

Proof. First, observe that by definition of being greedily-optimal, for every element $x \in S^*$, it must be that

$$\widehat{q}_{\mathcal{M}}(S^* - x) < 1 - 2^{-20} + \frac{\sum_{i=1}^{|S^*|-1} \frac{1}{i}}{2^{20} \log n},$$

and that

$$\widehat{q}_{\mathcal{M}}(S^*) \geq 1 - 2^{-20} + \frac{\sum_{i=1}^{|S^*|} \frac{1}{i}}{2^{20} \log n}.$$

In particular, this means

$$\widehat{q}_{\mathcal{M}}(S^*) - \widehat{q}_{\mathcal{M}}(S^* - x) \geq \frac{1}{2^{20} |S^*| \log n}.$$

By our bound relating p and \widehat{p} (Claim 4.4), we also know that with overwhelmingly high probability,

$$q_{\mathcal{M}}(S^*) - q_{\mathcal{M}}(S^* - x) \geq \frac{1}{2 \cdot 2^{20} |S^*| \log n}.$$

Now, let us define some auxiliary values: for a matroid $\mathcal{M} = (E, \mathcal{I})$ and $x \in T, T \subseteq E$, $p_{\mathcal{M}}(x, T)$ is defined as

$$p_{\mathcal{M}}(x, T) = \Pr_{\pi}[\{x\} \subseteq \text{FindCircuit}(\mathcal{M}, \pi) \subseteq T]$$

where the permutation π is drawn uniformly at random. We can observe that $p_{\mathcal{M}'}(x) = p_{\mathcal{M}'}(x, S^*) \geq p_{\mathcal{M}}(x, S^*)$. The inequality is because whenever we sample in accordance to a permutation π of $[n]$, and recover a circuit C_{π} such that $x \in C_{\pi}$ and $C_{\pi} \subseteq S^*$, the same permutation, if restricted to S^* and used to sample elements of S^* , would have given a circuit such that $x \in S^*$.

Finally, we can observe that $p_{\mathcal{M}}(x, S^*) = q_{\mathcal{M}}(S^*) - q_{\mathcal{M}}(S^* - x)$, as $q_{\mathcal{M}}(S^*) - q_{\mathcal{M}}(S^* - x)$ is exactly the probability that a circuit, when sampled from \mathcal{M} , is contained in S^* and uses the element x . Thus, we conclude that $p_{\mathcal{M}'}(x) \geq p_{\mathcal{M}}(x, S^*) \geq \frac{1}{2 \cdot 2^{20} |S^*| \log n}$. \square

4.2 Iterative Matroid Decomposition

Now that we have established how to recover greedily-optimal sets, we show how we can repeat this procedure to iteratively decompose our starting matroid. To start, we remove all circuits of length ≤ 50 to ensure that our algorithm has a non-trivial starting point.

Algorithm 7: RemoveSmallCircuits($\mathcal{M} = (E, \mathcal{I})$)

```

1 for  $i \in [50], S \in \binom{E}{i}$  in parallel do
2   | Query Ind( $S$ ).
3 end
4 Fix an arbitrary bijection  $\pi : E \rightarrow [n]$ 
5 for  $i \in [50], S \in \binom{E}{i}$  do
6   |  $x \leftarrow \arg \min_{y \in S} \pi(y)$ 
7   | if Ind( $S - x$ ) = 1  $\wedge$  Ind( $S$ ) = 0 then
8   |   |  $\mathcal{M} \leftarrow \mathcal{M} \setminus \{x\}$ .
9   | end
10 end
11 return  $\mathcal{M}$ 

```

Claim 4.12. *Algorithm 7 can be implemented in 1 round and removes all circuits of size ≤ 50 in \mathcal{M} while ensuring the rank of \mathcal{M} remains unchanged.*

Proof. It is clear that the algorithm finds all circuits of size ≤ 50 . Since it removes at least 1 element from each such circuit, these circuits are indeed eliminated. We show the rank of the matroid remains unchanged in the following.

Let E be the ground set of the input matroid and S be the set of elements we deleted from \mathcal{M} during Algorithm 7. We denote the elements in S as $e_1, \dots, e_{|S|}$ and order them such that $\pi(e_1) < \dots < \pi(e_{|S|})$. We prove $\text{rank}(E) = \text{rank}(E \setminus S)$ by induction on the size of S : Suppose $\text{rank}(E) = \text{rank}\left(E \setminus \bigcup_{j < i} \{e_j\}\right)$, we see that e_i must be in a circuit C which is disjoint from $\bigcup_{j < i} \{e_j\}$ since we always delete the smallest element w.r.t. π . Thus, we have

$$e_i \in \text{span}(C \setminus \{e_i\}) \subseteq \text{span}\left(\left(E \setminus \bigcup_{j < i} \{e_j\}\right) \setminus \{e_i\}\right) = \text{span}\left(E \setminus \bigcup_{j \leq i} \{e_j\}\right)$$

and

$$\text{rank}\left(E \setminus \bigcup_{j \leq i} \{e_j\}\right) = \text{rank}\left(E \setminus \bigcup_{j < i} \{e_j\}\right) = \text{rank}(E).$$

□

Having removed all short circuits, we next consider repeatedly running Algorithm 6, peeling off sets S_1, S_2, \dots :

Algorithm 8: Peel(\mathcal{M})

```
1  $S \leftarrow \text{FindGreedilyOptimal}(\mathcal{M})$ .  
2  $\mathcal{M} \leftarrow \mathcal{M} \setminus S$   
3 return  $S, \mathcal{M}$ 
```

Algorithm 9: IterativePeel(\mathcal{M})

```
1  $\mathcal{M} \leftarrow \text{RemoveSmallCircuits}(\mathcal{M})$   
2  $k = 0$ .  
3 while  $\mathcal{M} \neq \emptyset$  do  
4    $k \leftarrow k + 1$ .  
5    $S_k, \mathcal{M} \leftarrow \text{Peel}(\mathcal{M})$ .  
6   if  $\alpha(S_k) \geq 1/\log n$  or  $|S_k| > n/2$  then  
7     return  $S_1, \dots, S_{k-1}$   
8   end  
9 end  
10 return  $S_1, \dots, S_k$ 
```

We first observe that since we invoked [Algorithm 7](#) to eliminate every circuit of size ≤ 50 at the beginning, we always have $\alpha(S_i) \geq 50$. We provide the following characterization of how the α -value of the sets changes:

Claim 4.13. *Let \mathcal{M} be a matroid, and let S_1, \dots, S_k be a sequence of sets that are peeled off in accordance with [Algorithm 9](#). For any $i < j$ where $|S_i| \leq 2|S_j|$, we have*

$$\alpha(S_j) \geq \frac{\alpha(S_i)|S_j|}{|S_i|} + \frac{1}{5\sqrt{2}} \sqrt{\frac{\alpha(S_i)|S_j|}{|S_i|}}.$$

Proof. Let $\ell = \frac{\alpha(S_i)}{|S_i|}n$. Suppose U is a subset of the ground set E of cardinality ℓ , we define random variables $X_i = |U \cap S_i|$, $X_j = |U \cap S_j|$. Note that $X_i \sim \text{Hyp}(n, |S_i|, \ell)$, $X_j \sim \text{Hyp}(n, |S_j|, \ell)$ and they are negatively correlated.

For the sake of contradiction, suppose

$$\alpha(S_j) \leq \frac{\alpha(S_i)|S_j|}{|S_i|} + \frac{1}{5\sqrt{2}} \sqrt{\frac{\alpha(S_i)|S_j|}{|S_i|}}.$$

We aim to show that

$$\Pr[X_i < \alpha(S_i) \wedge X_j \geq \alpha(S_j)] > 4 \cdot 2^{-19}.$$

Consider the i -th iteration of [Algorithm 9](#), the above implies that for a random permutation π of $[n]$, there is a $> 4 \cdot 2^{-19}$ probability that there are less than $\alpha(S_i)$ elements from S_i and more than $\alpha(S_j)$ elements from S_j in the first ℓ elements of π . Conditioned on this, with at least $1/4$ probability, there is no circuit within S_i but there is a circuit in S_j . This implies that with $> 2^{-19}$ probability, the first circuit appears outside of S_i when adding elements according to the order of a random permutation π . But on the other hand, since S_i is a greedily-optimal set in the i -th iteration, by [Claim 4.4](#), we have $q(S) \geq 1 - 2^{-20} - 1/n_0^2 \geq 1 - 2^{-19}$. This is a contradiction.

To estimate $X_i \sim \text{Hyp}(n, |S_i|, \ell)$, we define $Y_i \sim \text{Bin}(\ell, |S_i|/n)$. It gives a good estimation as $\ell/n = \alpha(S_i)/|S_i| \leq 1/\log n$. Since $|S_i| \leq n/2$, we see that

$$\mathbb{E}[Y_i] = \alpha(S_i), \quad \text{Var}[Y_i] = \ell \frac{|S_i|}{n} \left(1 - \frac{|S_i|}{n}\right) \geq \frac{\alpha(S_i)}{2} \geq 25$$

It follows from [Lemma 2.10](#) that

$$\Pr[Y_i < \alpha(S_i)] = \Pr[Y_i \leq \mathbb{E}[Y_i] - 1] \geq \Pr\left[Y_i \leq \mathbb{E}[Y_i] - \frac{1}{5}\sqrt{\mathbf{Var}[Y_i]}\right] \geq \frac{1}{108}.$$

Combined with [Lemma 2.12](#) and $\ell/n \leq 1/\log n$, we conclude that

$$\Pr[X_i < \alpha(S_i)] \geq \frac{1}{108} - \frac{\ell-1}{n-1} \geq \frac{1}{200}.$$

To estimate $X_j \sim \text{Hyp}(n, |S_j|, \ell)$, we also define $Y_j \sim \text{Bin}(\ell, |S_j|/n)$. Since $|S_j| \leq n/2, |S_i| \leq 2|S_j|$, we see that

$$\mathbb{E}[Y_j] = \frac{\alpha(S_i)|S_j|}{|S_i|}, \quad \mathbf{Var}[Y_j] = \ell \frac{|S_j|}{n} \left(1 - \frac{|S_j|}{n}\right) \geq \frac{\alpha(S_i)|S_j|}{2|S_i|} \geq \frac{\alpha(S_i)}{4} \geq 1.$$

It follows from [Lemma 2.10](#) that

$$\Pr\left[Y_j \geq \mathbb{E}[Y_j] + \frac{1}{5}\sqrt{\mathbf{Var}[Y_j]}\right] \geq \frac{1}{108}$$

Combined with [Lemma 2.12](#) and $\ell/n \leq 1/\log n$, we have

$$\Pr\left[X_j \geq \mathbb{E}[Y_j] + \frac{1}{2}\sqrt{\mathbf{Var}[Y_j]}\right] \geq \frac{1}{108} - \frac{\ell-1}{n-1} \geq \frac{1}{200}.$$

We conclude that

$$\Pr[X_j \geq \alpha(S_j)] \geq \Pr\left[X_j \geq \frac{\alpha(S_i)|S_j|}{|S_i|} + \frac{1}{5\sqrt{2}}\sqrt{\frac{\alpha(S_i)|S_j|}{|S_i|}}\right] \geq \Pr\left[X_j \geq \mathbb{E}[Y_j] + \frac{1}{5}\sqrt{\mathbf{Var}[Y_j]}\right] \geq \frac{1}{200}.$$

The penultimate inequality follows from

$$\mathbb{E}[Y_j] = \frac{\alpha(S_i)|S_j|}{|S_i|}, \quad \mathbf{Var}[Y_j] \geq \frac{\alpha(S_i)|S_j|}{2|S_i|}.$$

Since X_i, X_j are negatively correlated, we have

$$\Pr[X_i < \alpha(S_i) \wedge X_j \geq \alpha(S_j)] \geq \Pr[X_i < \alpha(S_i)] \cdot \Pr[X_j \geq \alpha(S_j)] \geq \frac{1}{200^2} > 4 \cdot 2^{-19}.$$

as desired. \square

Now, we will bound the growth of the α values as a function of the number of sets that are peeled off:

Lemma 4.14. *Let \mathcal{M} be a matroid, and let S_1, \dots, S_k be a sequence of sets that are peeled off in accordance with [Algorithm 9](#). Now, let $\ell \in [\log n]$ be an integer, let $T = \{i \in [k] : |S_i| \in [2^\ell, 2^{\ell+1} - 1]\}$, let $\gamma = |T|$, and let a_1, \dots, a_γ denote the indices in T . Then it must be the case that*

$$\alpha(S_{a_\gamma}) = \Omega(\gamma^2), \quad \gamma = O\left(\sqrt{2^\ell}\right).$$

Proof. As $|S_{a_i}| \leq 2|S_{a_{i+1}}|$, we have

$$\alpha(S_{a_{i+1}}) \geq \frac{\alpha(S_{a_i})|S_{a_{i+1}}|}{|S_{a_i}|} + \frac{1}{5\sqrt{2}} \sqrt{\frac{\alpha(S_{a_i})|S_{a_{i+1}}|}{|S_{a_i}|}}.$$

by [Claim 4.13](#). Now, we multiply both sides with $2^\ell/|S_{a_{i+1}}|$:

$$\frac{\alpha(S_{a_{i+1}})}{|S_{a_{i+1}}|} \cdot 2^\ell \geq \frac{\alpha(S_{a_i})}{|S_{a_i}|} \cdot 2^\ell + \frac{1}{5\sqrt{2}} \sqrt{\frac{\alpha(S_{a_i})}{|S_{a_i}|} \cdot \frac{2^\ell}{|S_{a_{i+1}}|}} \cdot 2^\ell \geq \frac{\alpha(S_{a_i})}{|S_{a_i}|} \cdot 2^\ell + \frac{1}{10} \sqrt{\frac{\alpha(S_{a_i})}{|S_{a_i}|}} \cdot 2^\ell$$

If we set $X_i = \frac{\alpha(S_{a_i})}{|S_{a_i}|} \cdot 2^\ell$. We get the relationship that

$$X_{i+1} = X_i + \frac{\sqrt{X_i}}{10}.$$

As $X_1 \geq 1$, the recurrence implies that $X_\gamma = \Omega(\gamma^2)$. Therefore, we conclude that

$$\alpha(S_{a_\gamma}) = X_\gamma \cdot \frac{|S_{a_\gamma}|}{2^\ell} = \Omega(\gamma^2).$$

As $\alpha(S_{a_\gamma}) \leq |S_{a_\gamma}| \leq 2^{\ell+1}$, we have $\gamma = O(\sqrt{2^\ell})$. □

Claim 4.15. *Let \mathcal{M} be a matroid, and let S_1, \dots, S_k be a sequence of sets that are peeled off in accordance with [Algorithm 9](#). We have $k = O(n^{1/3})$.*

Proof. For every $\ell \in [\log n]$, we let $T_\ell = \{i \in [k] : |S_i| \in [2^\ell, 2^{\ell+1} - 1]\}$. It follows from [Lemma 4.14](#) that $|T_\ell| = O(\sqrt{2^\ell})$. We can also trivially bound $|T_\ell| = O(\frac{n}{2^\ell})$ as the total size of the sets is at most n . Therefore, we see that

$$k = \sum_{\ell=1}^{\log n} |T_\ell| = \sum_{\ell=1}^{(2/3)\log n} |T_\ell| + \sum_{\ell=(2/3)\log n+1}^{\log n} |T_\ell| = \sum_{\ell=1}^{(2/3)\log n} O(\sqrt{2^\ell}) + n^{1/3} = O(n^{1/3}).$$

□

Combining [Claim 4.15](#) and [Lemma 4.14](#) gives the following theorem (which we state explicitly, as it may be of independent interest):

Theorem 4.16 (Decomposition Theorem). *Let \mathcal{M} be a matroid, and let S_1, \dots, S_k be a sequence of sets that are peeled off in accordance with [Algorithm 9](#). Now, let $\ell \in [\log n]$ be an integer, let $T = \{i \in [k] : |S_i| \in [2^\ell, 2^{\ell+1} - 1]\}$, let $\gamma = |T|$, and let a_1, \dots, a_γ denote the indices in T . Then it must be the case that*

1. $\alpha(S_{a_\gamma}) = \Omega(\gamma^2)$.
2. $\gamma = O(\sqrt{2^\ell})$.
3. $k = O(n^{1/3})$.

5 Making Progress With Large α 's

In this section, we will describe how we can make progress by *contracting a large independent set* when we obtain a set S with large $\alpha(S)$ during our decomposition. In the subsequent section, we will show that when the $\alpha(S)$ values are small, there are different ways of making progress. Collectively, these results will allow for a case-based analysis that allows us to beat $O(\sqrt{n})$ rounds no matter what is returned during the decomposition, which we present in [Section 7](#).

Claim 5.1. *Let \mathcal{M} be a matroid on n elements, and let S be a greedily-optimal set in \mathcal{M} . Then, for $\ell = \frac{\alpha(S)}{10|S|}n$, we have*

$$\Pr_{\pi}[\text{Ind}(\{\pi(1), \dots, \pi(\ell)\}) = 1] \geq \frac{1}{4}.$$

Proof. Suppose U is a random subset of E of cardinality ℓ . We define a random variable $X = |U \cap S|$. Note that $X \sim \text{Hyp}(n, |S|, \ell)$, and $\mathbb{E}[X] = \alpha(S)/10$. By Markov's inequality,

$$\Pr[X \geq \alpha(S)] \leq \frac{1}{10}.$$

In the i -th iteration of [Algorithm 9](#), as S is a greedily-optimal set, we have $q_{\mathcal{M}}(S) \geq 1 - 2^{-20} - 1/n_0^2$ by [Claim 4.4](#). In particular, if there is no circuit in $S \cap U$, this bounds the probability of a circuit appearing outside of S by $2^{-20} + 1/n_0^2$. Using this, we obtain:

$$\begin{aligned} \Pr_{\pi}[\text{Ind}(\{\pi(1), \dots, \pi(\ell)\}) = 0] &\leq \frac{1}{2^{20}} + \frac{1}{n_0^2} + \Pr_U[\text{Ind}(U \cap S) = 0] \\ &\leq \frac{1}{2^{20}} + \frac{1}{n_0^2} + \left(\Pr_{T \in \binom{S}{\ell}}[\text{Ind}(T) = 0 \mid X < \alpha(S)] + \Pr[X \geq \alpha(S)] \right) \\ &\leq \frac{1}{2^{20}} + \frac{1}{n_0^2} + \frac{1}{2} + \frac{1}{10} \leq \frac{3}{4}. \end{aligned}$$

The first line of the proof follows by seeing that if there is a dependence in $\{\pi(1), \dots, \pi(\ell)\}$, then either (1) this is a circuit in $U \cap S$, or (2) there is no circuit in $U \cap S$, but there is a circuit outside $U \cap S$. Our bound on $q_{\mathcal{M}}(S)$ states that the probability of this second case is bounded by $\frac{1}{2^{20}} + \frac{1}{n_0^2}$. This concludes the proof. \square

Given the above claim, it follows that we can recover an independent set of size $\Omega\left(\frac{\alpha(S)}{|S|}n\right)$ with probability at least $1 - 2^{-n_0}$ by sampling $\text{poly}(n_0)$ many random permutations.

6 Making Progress With Small α 's

In this section, we will describe how we can make progress on *deleting redundant elements* when we obtain a set S with a small $\alpha(S)$ value. The reader may see [Section 2.2](#) for the definitions of some of the quantities we make use of in this section.

6.1 Rank Deficiency in Matroids

Our first result is a general structural result which governs the so-called *rank deficiency* of matroids.

Definition 6.1. For a matroid \mathcal{M} on n elements of rank r we say that the rank-deficiency is $n - r$.

6.1.1 Matroid Decomposition for Quotient Counting Bounds

Now, recall the work of [Qua24] showed the following key theorem:

Theorem 6.2. [Lemma 3.2 in [Qua24]] Let $\mathcal{M} = (E, \mathcal{I})$ be a matroid of rank r with $|E| = n$, and let

$$\kappa(\mathcal{M}) = \min_{S \subseteq E} \frac{n - |S|}{r - \text{rank}(S)}.$$

Then, for any $\eta \in \mathbb{Z}^+$, there are $\leq n^{\eta+1} r^\eta$ quotients of \mathcal{M} of size $\leq \eta \cdot \kappa(\mathcal{M})$.

Remark 6.3. The expression

$$\kappa(\mathcal{M}) = \min_{S \subseteq E} \frac{n - |S|}{r - \text{rank}(S)}$$

is always minimized for $S = \text{span}(S)$, i.e. S is a flat. Otherwise, WLOG we can set $S = \text{span}(S)$, and the denominator is unchanged while the numerator decreases.

Now, we have the following key claim:

Claim 6.4. Let $\mathcal{M} = (E, \mathcal{I})$ be a matroid of rank r with $|E| = n$ and let d be a given parameter. Then, there exists a set T of size $|T| \leq r \cdot d$ such that the new matroid $\mathcal{M} \setminus T$ is either empty or satisfies $\kappa(\mathcal{M} \setminus T) \geq d$.

Proof. Consider the following iterative process: We start with $T = \emptyset$ and $\mathcal{M}_1 = \mathcal{M}$. In the i -th iteration:

- If $\mathcal{M}_i = \emptyset$ or $\kappa(\mathcal{M}_i) \geq d$, we are done.
- Otherwise, let $\mathcal{M}_i = (E_i, \mathcal{I}_i)$. It implies there exist a flat $S \subseteq E_i$ such that

$$\frac{|E_i| - |S|}{\text{rank}(\mathcal{M}_i) - \text{rank}(S)} < d.$$

We set $\mathcal{M}_{i+1} \leftarrow \mathcal{M}_i|_S$, $T \leftarrow T \cup (E_i \setminus S)$. Note that $\text{rank}(\mathcal{M}_{i+1}) = \text{rank}(S)$ and the size of T has increased by $|E_i| - |S| < d \cdot (\text{rank}(\mathcal{M}_i) - \text{rank}(S)) = d \cdot (\text{rank}(\mathcal{M}_i) - \text{rank}(\mathcal{M}_{i+1}))$.

Suppose the process terminates after t iterations. We finish the proof by observing that

$$|T| < \sum_{i=1}^t d(\text{rank}(\mathcal{M}_i) - \text{rank}(\mathcal{M}_{i+1})) = d(\text{rank}(\mathcal{M}_1) - \text{rank}(\mathcal{M}_t)) \leq d \cdot \text{rank}(\mathcal{M}_1) = d \cdot r.$$

□

Theorem 6.5. Let $\mathcal{M} = (E, \mathcal{I})$ be a matroid on n elements of rank r , and let $d \in \mathbb{Z}^+$ be a parameter of our choosing. Then, there is a set of elements $T \subseteq E$ of size $|T| \leq r \cdot d$ such that the matroid $\mathcal{M}' = \mathcal{M} \setminus T$ is either empty, or for any $\eta \in \mathbb{Z}^+$, \mathcal{M}' has at most $n^{2\eta+1}$ quotients of size $\leq \eta d$.

Proof. This follows by letting T be the set of elements recovered by Claim 6.4, and then invoking Theorem 6.2 on the matroid $\mathcal{M} \setminus T$ (if non-empty), where $\kappa(\mathcal{M} \setminus T) \geq d$. □

6.1.2 Proof of Rank Deficiency

With all of this groundwork established, we are now ready to prove our theorem for general matroids:

Theorem 6.6. *Let $\mathcal{M} = (E, \mathcal{I})$ be a matroid on n elements. For a set $S \subseteq E$ and matroid $\mathcal{M}' = \mathcal{M}|_S$ such that*

1. $\alpha(S) \leq \frac{|S|}{100 \log n}$.
2. For every element $x \in S$, $p_{\mathcal{M}'}(x) \geq \frac{1}{n^{10}}$.

Then, $|S| - \text{rank}(S) = \Omega(|S|/\log n)$, where $\text{rank}(S) = \text{rank}(\mathcal{M}')$.

Proof. We invoke [Theorem 6.2](#) on the dual matroid $(\mathcal{M}')^*$, with parameter $d = 100 \log n$. The result tells us that there is a set $T \subseteq S$ of size $|T| \leq (|S| - \text{rank}(S))100 \log n$ such that the resulting matroid $(\mathcal{M}')^* \setminus T$ is either empty or satisfies a quotient counting bound: specifically, that the number of quotients of size $\leq \eta d$ is at most $|S|^{2\eta+1}$. We show that it must be the former case by showing a contradiction assuming $(\mathcal{M}')^* \setminus T$ is non-empty.

In particular, let us now consider sampling this matroid $(\mathcal{M}')^* \setminus T$ at rate $1/2$. Our goal is to bound the probability that there exists *any quotient* which survives (here, we take surviving to mean that *every* element in the quotient is selected during the sampling) this sampling procedure. We see that this is bounded by

$$\begin{aligned} \Pr[\exists c \in Q((\mathcal{M}')^* \setminus T) : c \text{ survives sampling}] &\leq \sum_{c \in Q((\mathcal{M}')^* \setminus T)} \left(\frac{1}{2}\right)^{|c|} \\ &\leq \sum_{\eta \in \mathbb{Z}^+} n^{2\eta+1} \cdot \left(\frac{1}{2}\right)^{\eta d} \\ &\leq n \cdot \sum_{\eta \in \mathbb{Z}^+} \frac{1}{n^{98\eta}} \leq \frac{2}{n^{97}}. \end{aligned}$$

It follows from [Fact 2.8](#) that $((\mathcal{M}')^* \setminus T)^* = \mathcal{M}'/T$. By [Fact 2.9](#), we know that any circuit C of \mathcal{M}'/T is the complement to a hyperplane of $(\mathcal{M}'/T)^* = (\mathcal{M}')^* \setminus T$. In particular, because hyperplanes are flats, we know that any circuit of \mathcal{M}'/T is a quotient of $(\mathcal{M}')^* \setminus T$. Importantly, this means that when we sample the matroid \mathcal{M}'/T at rate $1/2$, there is a $\leq 2/n^{97}$ chance of any circuit surviving.

Now, consider the elements $S \setminus T$. Recall that in the original matroid, we know that $\forall x \in S$, $p_{\mathcal{M}'}(x) \geq \frac{1}{n^{10}}$. In particular, because $\alpha(S) \leq \frac{|S|}{100 \log n}$, if we sample \mathcal{M}' at rate $1/2$, we expect $\geq 50 \log n \cdot \alpha(S)$ elements to survive the sampling. It follows from a simple Chernoff bound that with probability at least $1 - n^{-5}$, we have $\geq 25 \log n \cdot \alpha(S)$ elements that are selected during sampling. Conditioned on this, by [Claim 4.7](#), we see that a circuit will survive with probability at least $1 - n^{-25}$.

The final key observation is the following: if $T \neq S$, there must exist elements $x \in S \setminus T$. Originally, these elements satisfied $p_{\mathcal{M}'}(x) \geq \frac{1}{n^{10}}$. Thus, if we sample the original matroid \mathcal{M}' at rate $1/2$, we would have that the probability that there is a circuit in the resulting sample which includes x is at least

$$\left(1 - \frac{1}{n^5}\right) \left(1 - \frac{1}{n^{25}}\right) \frac{1}{n^{10}} \geq \frac{1}{n^{11}}.$$

However, we are now no longer sampling the matroid \mathcal{M}' , but rather sampling the matroid \mathcal{M}'/T . For analysis, we consider correlating the two sampling procedures; i.e., we let P denote the sample

of elements received in \mathcal{M}' , and we let \tilde{P} denote the same set of sampled elements in \mathcal{M}'/T , so $\tilde{P} = P \setminus T$. The point now is that if there is a circuit $C \subseteq P$ such that $C \setminus T \neq \emptyset$, *there must also be a circuit $\tilde{C} \subseteq C \setminus T \subseteq \tilde{P}$* . To see this, we note that $\text{rank}_{\mathcal{M}'}(C) = |C| - 1$ as C is a circuit. Likewise, $\text{rank}_{\mathcal{M}'}(C \cap T) = |C \cap T|$ as $C \cap T \subsetneq T$ and thus it must be independent (otherwise it contracts that C is a minimal dependent set). By [Fact 2.8](#), we have

$$\begin{aligned} \text{rank}_{\mathcal{M}'/T}(C \setminus T) &= \text{rank}_{\mathcal{M}'}((C \setminus T) \cup T) - \text{rank}_{\mathcal{M}'}(T) \\ &= \text{rank}_{\mathcal{M}'}(C \cup T) - \text{rank}_{\mathcal{M}'}(T) \\ &\leq \text{rank}_{\mathcal{M}'}(C) + \text{rank}_{\mathcal{M}'}(T) - \text{rank}_{\mathcal{M}'}(C \cap T) - \text{rank}_{\mathcal{M}'}(T) \\ &\leq |C| - 1 - |C \cap T| \\ &\leq |C \setminus T| - 1. \end{aligned}$$

The first inequality follows from the submodularity of the rank function of a matroid. Therefore, we conclude that $C \setminus T$ is dependent in \mathcal{M}'/T and there is a circuit inside it.

But this leads to a contradiction that

$$\begin{aligned} \frac{2}{n^{97}} &\geq \Pr[\exists \text{ circuit in } \mathcal{M}'/T \text{ when sampling with rate } 1/2] \\ &\geq \Pr[\exists \text{ circuit } C \text{ in } \mathcal{M}' \text{ when sampling with rate } 1/2 \text{ s.t. } C \setminus T \neq \emptyset] \geq \frac{1}{n^{11}} \end{aligned}$$

In particular, this means that the matroid $(\mathcal{M}')^* \setminus T$ must be empty. Therefore $|S| = |(\mathcal{M}')^*| = |T| \leq (|S| - \text{rank}(S))100 \log n$, and so $|S| - \text{rank}(S) = \Omega(|S|/\log n)$. \square

6.2 Efficient Redundant Element Recovery

In this section, we present an algorithm which can efficiently recover redundant elements, provided the α values are significantly smaller than the set size.

Algorithm 10: RecoverRedundantElements(\mathcal{M}, S)

```

1 Let  $\hat{\alpha}(S)$  be the estimation of  $\alpha(S)$  given by Claim 4.8
2  $t \leftarrow 20 \log n \cdot \hat{\alpha}(S), \ell \leftarrow \frac{|S|}{4t}$ 
3 for  $i \in [\ell]$  in parallel do
4   Draw a random permutation (bijection)  $\pi : [|S|] \rightarrow S$ 
5    $A_i \leftarrow \{\pi(1), \dots, \pi(t)\}, B_i \leftarrow \emptyset$ 
6   for  $j \in [t]$  in parallel do
7     for  $x \in E \setminus A_i$  in parallel do
8       Query  $\text{Ind}(\{\pi(1), \dots, \pi(j)\})$  and  $\text{Ind}(\{\pi(1), \dots, \pi(j)\} \cup \{x\})$ 
9       if  $\text{Ind}(\{\pi(1), \dots, \pi(j)\}) = 1 \wedge \text{Ind}(\{\pi(1), \dots, \pi(j)\} \cup \{x\}) = 0$  then
10         $B_i \leftarrow B_i \cup \{x\}$ 
11      end
12    end
13  end
14 end
15 return  $\bigcup_{i \in [\ell]} B_i \setminus (\bigcup_{i \in [\ell]} A_i)$ 

```

We establish the following lemma:

Lemma 6.7. *Let \mathcal{M} be a matroid, and S be a greedily-optimal set and $\alpha(S) \leq \frac{|S|}{100 \log n}$. Then, Algorithm 10 recovers $\tilde{\Omega} \left(\min \left\{ |S|, \frac{|S|^2}{\alpha(S)^2} \right\} \right)$ redundant elements with probability $1 - 1/2^n$.*

Proof. First, we note that

$$\left(\bigcup_{i \in [\ell]} B_i \setminus \left(\bigcup_{i \in [\ell]} A_i \right) \right) \subseteq \text{span} \left(\bigcup_{i \in [\ell]} A_i \right)$$

Thus, the recovered set is indeed redundant. We now focus on bounding the size of this set in the following.

As S is a greedily-optimal set in \mathcal{M} , by Claim 4.11, we have for any $x \in S$, $p_{\mathcal{M}|S}(x) \geq \frac{1}{2^{21}|S| \log n}$. For every $x \in S$, we say that $p_{x,r}$ is the probability over a random order of picking elements such that x is the r -th element added, that x participates in the first circuit that appears.

In particular, we can establish some simple inequalities:

1. $p_{x,r} \geq p_{x,r+1}$.
2. $p_{\mathcal{M}|S}(x) = \frac{1}{|S|} \sum_{r=1}^{|S|} p_{x,r}$.
3. $\sum_{r=t+1}^{|S|} p_{x,r} \leq 1/n^{10}$.

The third inequality follows from Claim 4.7: a random subset of S of cardinality $t = 20\hat{\alpha}(S) \log n \geq 10\alpha(S) \log n$ is dependent with probability at least $1 - 1/n^{10}$.

With this, we can see that for any $x \in S$,

$$\frac{1}{2^{21}|S| \log n} \leq p_x = \frac{1}{|S|} \sum_{r=1}^{|S|} p_{x,r} = \frac{1}{|S|} \left(\sum_{r=1}^t p_{x,r} + \frac{1}{n} \right) \leq \frac{t}{|S|} \cdot p_{x,1} + \frac{1}{n^{10}}.$$

In particular, this implies that

$$p_{x,1} \geq \frac{|S|}{t} \cdot \left(\frac{1}{2^{21}|S| \log n} - \frac{1}{n^{10}} \right) \geq \frac{1}{2^{22}t \log n}$$

Now, let us revisit the above algorithm. Our first step will be to understand the probability that an element x appears in one of the sets A_1, \dots, A_ℓ . For this, observe that each set A_i is of size t . Thus,

$$\Pr \left[x \notin \bigcup_{i \in [\ell]} A_i \right] = \Pr[x \notin A_1]^\ell = \left(1 - \frac{t}{|S|} \right)^\ell = \left(1 - \frac{t}{|S|} \right)^{\frac{|S|}{4t}} \geq e^{-1/2} \geq 1/2.$$

The first inequality is because $\frac{t}{|S|} = \frac{20 \log n \cdot \hat{\alpha}(S)}{|S|} \leq \frac{40 \log n \cdot \alpha(S)}{|S|} \leq \frac{1}{2}$ and for every $0 \leq x \leq \frac{1}{2}$, $1 - x \geq e^{-2x}$.

Now, let us introduce the value q_i such that

$$q_x = \Pr \left[x \notin \bigcup_{i \in [\ell]} A_i \wedge x \in \bigcup_{i \in [\ell]} B_i \right] = \Pr \left[x \notin \bigcup_{i \in [\ell]} A_i \right] \cdot \Pr \left[x \in \bigcup_{i \in [\ell]} B_i \mid x \notin \bigcup_{i \in [\ell]} A_i \right].$$

Note that the samples A_1, \dots, A_ℓ are all done independently of one another. Hence,

$$\Pr \left[x \in \bigcup_{i \in [\ell]} B_i \mid x \notin \bigcup_{i \in [\ell]} A_i \right] = 1 - \Pr \left[x \notin \bigcup_{i \in [\ell]} B_i \mid x \notin \bigcup_{i \in [\ell]} A_i \right] = 1 - \Pr[x \notin B_1 \mid x \notin A_1]^\ell.$$

Now, let us understand $\Pr[x \in B_1 \mid x \notin A_1]$. This is exactly the probability of x appearing in the first circuit when we randomly add the set A_1 of elements to x . Since A_1 is disjoint from x , this is exactly $p_{i,1}$. Hence, we obtain that

$$\begin{aligned} 1 - \Pr[x \notin B_1 \mid x \notin A_1]^\ell &= 1 - (1 - p_{x,1})^\ell \\ &\geq 1 - \left(1 - \frac{1}{2^{22}t \log n} \right)^{\frac{|S|}{4t}} \\ &\geq 1 - \exp \left(-\frac{|S|}{2^{24}t^2 \log n} \right) \\ &\geq \left\{ \frac{1}{2}, \frac{|S|}{2^{25}t^2 \log n} \right\} \end{aligned}$$

The last inequality follows from the fact that $1 - e^{-x} \geq \min\{1/2, x/2\}$ when $x \geq 0$.

To conclude, we obtain that

$$q_x \geq \frac{1}{2} \cdot (1 - \Pr[x \notin B_1 \mid x \notin A_1]^\ell) \geq \min \left\{ \frac{1}{4}, \frac{|S|}{2^{26}t^2 \log n} \right\}$$

Finally then, we see that

$$\begin{aligned} \mathbb{E} \left[\left| \bigcup_{i \in [\ell]} B_i \setminus \left(\bigcup_{i \in [\ell]} A_i \right) \right| \right] &= \sum_{x \in S} q_x \\ &\geq \sum_{x \in S} \min \left\{ \frac{1}{4}, \frac{|S|}{2^{26}t^2 \log n} \right\} \\ &\geq \min \left\{ \frac{|S|}{4}, \frac{|S|^2}{2^{26}t^2 \log n} \right\} \\ &= \tilde{\Omega} \left(\min \left\{ |S|, \frac{|S|^2}{\widehat{\alpha}(S)^2} \right\} \right) \\ &= \tilde{\Omega} \left(\min \left\{ |S|, \frac{|S|^2}{\alpha(S)^2} \right\} \right). \end{aligned}$$

Repeating the above $\text{poly}(n_0)$ times achieves at least this expectation with probability at least $1 - 2^{-n_0}$ by a Hoeffding's inequality. □

7 Guaranteeing Progress through Decomposition

We now establish several warm-up claims about when it is easy to make progress during our decomposition. First, to establish uniform notation, we consider the following process:

Algorithm 11: EarlyStopDecomposition(\mathcal{M})

```
1  $\mathcal{M} \leftarrow \text{RemoveSmallCircuits}(\mathcal{M})$ 
2  $k \leftarrow 0$ 
3 while  $|E| \geq n/2$  do
4    $k \leftarrow k + 1$ 
5    $S_k \leftarrow \text{Peel}(\mathcal{M})$ 
6    $\mathcal{M} \leftarrow \mathcal{M} \setminus S_k$ .
7   Let  $\hat{\alpha}(S_k)$  be the estimation of  $\alpha(S_k)$  given by Claim 4.8
8   if  $\hat{\alpha}(S_k) = \Omega(|S_k|/\log n)$  then
9     return  $k - 1, S_1, \dots, S_{k-1}, \mathcal{M}$ 
10  end
11 end
12 return  $k, S_1, \dots, S_k, \mathcal{M}$ 
```

7.1 Subroutines for Making Progress Towards a Basis

We define

$$i^* = \arg \max_{i \in [k]} \frac{\alpha(S_i)}{|S_i|}.$$

For $\ell \in [\log n]$, let $J_\ell = \{i \in [k] \mid |S_i| \in [2^\ell, 2^{\ell+1} - 1]\}$. We see that at least one J_{ℓ^*} satisfies $|J_{\ell^*}| \geq k/\log n$. We denote the sets S_i for $i \in J_{\ell^*}$ by T_1, \dots, T_γ , and let $\tau = 2^{\ell^*}$, $\beta = \tau \cdot \alpha(S_{i^*})/|S_{i^*}|$.

Claim 7.1. *We have the following properties:*

1. $\gamma = \tilde{\Omega}(k)$.
2. For any $i \in [\gamma]$, $\tau \leq |T_i| \leq 2\tau$.
3. For any $i \in [\gamma]$, $\alpha(T_i) \leq \frac{|T_i|}{100 \log n}$.
4. For any $i \in [k]$, $\frac{\alpha(S_i)}{|S_i|} \leq \frac{\beta}{\tau}$. For any $i \in [\gamma]$, $\alpha(T_i) = O(\beta)$.
5. $\beta = \Omega(\gamma^2)$, $\tau = \Omega(\beta)$.

Proof. Item 1 follows from our notational choices, as J_{ℓ^*} satisfies $\gamma = |J_{\ell^*}| \geq k/\log n$. Item 2 again follows by our definition of τ . Item 3 is because if $\alpha(T_i) > \frac{|T_i|}{100 \log n}$, the algorithm would have returned on Line 9 as $\hat{\alpha}(T_i) = \Theta(\alpha(T_i)) = \Omega(|T_i|/\log n)$. Item 4 follows from the maximality of $\alpha(S_{i^*})/|S_{i^*}|$ and $\alpha(T_i) \leq \frac{\beta}{\tau} |T_i| = O(\beta)$. For item 5, it follows from Lemma 4.14 that $\alpha(T_\gamma) = \Omega(\gamma^2)$. Given $\alpha(T_\gamma) = O(\beta)$ by item 3, we see that $\beta = \Omega(\gamma^2)$. Additionally, we have $\tau = \Omega(\beta)$ as $\frac{\beta}{\tau} = \frac{\alpha(S_k)}{S_k} < 1$. \square

Claim 7.2. *If the algorithm returns on Line 9, then there is an efficient procedure for recovering an independent set of size $\tilde{\Omega}(n)$ with probability $1 - 2^{\Omega(n_0)}$ by investing a single additional round with $\text{poly}(n_0)$ queries.*

Proof. Let \mathcal{M} be the matroid before we peel off S_k . Since S_k is a greedily-optimal set in \mathcal{M} and $\alpha(S_k) = \Theta(\hat{\alpha}(S_k)) = \Omega(|S_k|/\log n)$, by Claim 5.1, for $\ell = \frac{n\alpha(S_k)}{10|S_k|} = \Omega(n)$,

$$\Pr_{\pi}[\text{Ind}(\{\pi(1), \dots, \pi(\ell)\}) = 1] \geq \frac{1}{4}.$$

Thus, by sampling $\text{poly}(n_0)$ many random permutations, we can find an independent set of size ℓ with high probability in 1 additional round. \square

By the above claim, we see that if the algorithm returns on [Line 9](#), we are in a good situation since we have invested $O(k) = O(n^{1/3})$ rounds ([Claim 4.15](#)) and can recover an independent set of size $\tilde{\Omega}(n)$. Therefore, we assume that the algorithm *does not* return on [Line 9](#) in the following. We present 4 different ways of making progress.

Lemma 7.3. *There is an efficient procedure for recovering an independent set of size $\Omega(n\beta/\tau)$ with probability $1 - 2^{-\Omega(n_0)}$ by investing a single additional round with $\text{poly}(n_0)$ queries.*

Proof. Let \mathcal{M} be the matroid right before we peel off S_{i^*} . Since S_{i^*} is a greedily-optimal set in \mathcal{M} and $\alpha(S_{i^*})/|S_{i^*}| = \beta/\tau$, by [Claim 5.1](#), for $\ell = \frac{n\beta}{10\tau}$

$$\Pr_{\pi}[\text{Ind}(\{\pi(1), \dots, \pi(\ell)\}) = 1] \geq \frac{1}{4}.$$

Thus, by sampling $\text{poly}(n_0)$ many random permutations, we can find an independent set of size ℓ with high probability in 1 additional round. \square

Lemma 7.4. *There is an efficient procedure for deleting $\tilde{\Omega}(\gamma\tau)$ redundant elements by investing $O(\tau^{1/2})$ additional rounds and making only $\text{poly}(n_0)$ queries.*

Proof. For any $i \in [\gamma]$, we have $\alpha(T_i) \leq \frac{|T_i|}{100 \log n}$. Thus, by [Theorem 6.6](#), $|T_i| - \text{rank}(T_i) = \tilde{\Omega}(|T_i|)$ for every $i \in [\gamma]$. Therefore, we can use the $O(n^{1/2})$ round algorithm of [\[KUW85\]](#) to find an independent set I_i for each T_i in parallel in $O(|T_i|^{1/2}) = O(\tau^{1/2})$ additional rounds. In total, we can remove

$$\sum_{i=1}^{\gamma} |T_i| - |I_i| = \sum_{i=1}^{\gamma} \tilde{\Omega}(|T_i|) = \tilde{\Omega}(\gamma\tau)$$

redundant elements. \square

Lemma 7.5. *There is an efficient procedure for deleting $\tilde{\Omega}\left(\gamma \cdot \min\left(\tau, \frac{\tau^2}{\beta^2}\right)\right)$ redundant elements with probability $1 - 2^{-\Omega(n_0)}$ by investing a single additional round with $\text{poly}(n_0)$ queries.*

Proof. For every $i \in [\gamma]$, we have $\alpha(T_i) \leq \frac{|T_i|}{100 \log n}$. Therefore, we can invoke [Lemma 6.7](#) (in parallel) across all T_i 's for $i \in [\gamma]$. This lemma guarantees that for each T_i , we recover

$$\tilde{\Omega}\left(\min\left\{|T_i|, \frac{|T_i|^2}{\alpha(T_i)^2}\right\}\right) \geq \tilde{\Omega}\left(\min\left\{\tau, \frac{\tau^2}{\beta^2}\right\}\right).$$

In total, we recover

$$\sum_{i \in [\gamma]} \tilde{\Omega}\left(\min\left\{\tau, \frac{\tau^2}{\beta^2}\right\}\right) = \tilde{\Omega}\left(\gamma \cdot \min\left\{\tau, \frac{\tau^2}{\beta^2}\right\}\right).$$

redundant elements. \square

Lemma 7.6. *There is an efficient procedure for deleting $\tilde{\Omega}\left(\min\left(n, \frac{\tau^2}{\beta^2}\right)\right)$ redundant elements with probability $1 - 2^{-\Omega(n_0)}$ by investing a single additional round with $\text{poly}(n_0)$ queries.*

Proof. We apply [Lemma 6.7](#) (in parallel) across every S_i for $i \in [k]$. As in the previous lemma's proof, [Lemma 6.7](#) guarantees that we can recover

$$\tilde{\Omega} \left(\min \left\{ |S_i|, \frac{|S_i|^2}{\alpha(S_i)^2} \right\} \right) \geq \tilde{\Omega} \left(\min \left\{ |S_i|, \frac{\tau^2}{\beta^2} \right\} \right)$$

redundant elements from each S_i . In total, we recover

$$\sum_{i \in [k]} \tilde{\Omega} \left(\min \left(|S_i|, \frac{\tau^2}{\beta^2} \right) \right) = \tilde{\Omega} \left(\min \left(n, \frac{\tau^2}{\beta^2} \right) \right)$$

redundant elements then. Note that this follows because in the sum, either in every term $|S_i| \leq \frac{\tau^2}{\beta^2}$, in which case the sum behaves like $\sum_i |S_i| = \Omega(n)$, or for at least one term $|S_i| > \frac{\tau^2}{\beta^2}$, and so the sum contains at least one contribution of $\frac{\tau^2}{\beta^2}$. \square

7.2 Piecing the Subroutines Together

To summarize, we have 4 ways of making progress:

	Progress	Round Complexity
Lemma 7.3	$\tilde{\Omega}(n\beta/\tau)$	$\tilde{O}(\gamma)$
Lemma 7.4	$\tilde{\Omega}(\gamma\tau)$	$\tilde{O}(\gamma + \tau^{1/2}) = \tilde{O}(\tau^{1/2})$
Lemma 7.5	$\tilde{\Omega} \left(\gamma \cdot \min \left(\tau, \frac{\tau^2}{\beta^2} \right) \right)$	$\tilde{O}(\gamma)$
Lemma 7.6	$\tilde{\Omega} \left(\min \left(n, \frac{\tau^2}{\beta^2} \right) \right)$	$\tilde{O}(\gamma)$

Remark 7.7. Note that we know $\gamma = O(\beta^{1/2}) = O(\tau^{1/2})$ because of [Claim 7.1](#).

Algorithmically, we will always choose the one that maximizes the average progress per round, which is given by

$$\max \left\{ \frac{n\beta}{\tau\gamma}, \gamma\tau^{1/2}, \min \left\{ \tau, \frac{\tau^2}{\beta^2} \right\}, \min \left\{ \frac{n}{\gamma}, \frac{\tau^2}{\beta^2\gamma} \right\} \right\}.$$

Notationally, we will let $\mathbf{Progress}_1(\beta, \tau, \gamma)$, $\mathbf{Rounds}_1(\beta, \tau, \gamma)$ denote the progress and round complexity guaranteed by [Lemma 7.3](#), $\mathbf{Progress}_2(\beta, \tau, \gamma)$, $\mathbf{Rounds}_2(\beta, \tau, \gamma)$ that of [Lemma 7.4](#), $\mathbf{Progress}_3(\beta, \tau, \gamma)$, $\mathbf{Rounds}_3(\beta, \tau, \gamma)$ that of [Lemma 7.5](#), and $\mathbf{Progress}_4(\beta, \tau, \gamma)$, $\mathbf{Rounds}_4(\beta, \tau, \gamma)$ that of [Lemma 7.6](#).

Now, we have the following lemma:

Lemma 7.8. *Let β, τ, γ be the parameters resulting from running [Algorithm 11](#). Then, for any possible β, τ, γ , there is a choice of sub-routine $q \in [4]$ such that*

$$\frac{\mathbf{Progress}_q(\beta, \tau, \gamma)}{\mathbf{Rounds}_q(\beta, \tau, \gamma)} = \tilde{\Omega}(n^{8/15}).$$

Proof. We use a case analysis based on the values of τ, β .

1. When $\beta \geq n^{2/5}$:

- If $\tau \leq \beta^2/n^{2/15}$, then we immediately have that $\frac{\beta}{\sqrt{\tau}} \geq n^{1/15}$. Thus, we can write that

$$n^{8/15} \leq n^{7/15} \cdot \frac{\beta}{\sqrt{\tau}},$$

which implies that

$$\frac{n^{8/15}}{\tau^{1/2}} \leq \frac{n^{7/15}\beta}{\tau}.$$

Now again, we do a case analysis based on the value of γ : if $\gamma \geq \frac{n^{8/15}}{\tau^{1/2}}$, then

$$\frac{\mathbf{Progress}_2(\beta, \tau, \gamma)}{\mathbf{Rounds}_2(\beta, \tau, \gamma)} = \tilde{\Omega}(\gamma\tau^{1/2}) = \tilde{\Omega}(n^{8/15}).$$

Otherwise, if $\gamma \leq \frac{n^{7/15}\beta}{\tau}$:

$$\frac{\mathbf{Progress}_1(\beta, \tau, \gamma)}{\mathbf{Rounds}_1(\beta, \tau, \gamma)} = \tilde{\Omega}\left(\frac{n\beta}{\tau\gamma}\right) \geq n^{8/15}.$$

- Now, we consider when $\tau \geq \beta^2/n^{2/15} \geq n^{2/3} \geq n^{8/15}$. Immediately, this implies that

$$\frac{\tau^2}{\beta^2} \geq \frac{\beta^2}{n^{4/15}} \geq n^{8/15}.$$

Thus, we have

$$\frac{\mathbf{Progress}_3(\beta, \tau, \gamma)}{\mathbf{Rounds}_3(\beta, \tau, \gamma)} = \tilde{\Omega}\left(\min\left\{\tau, \frac{\tau^2}{\beta^2}\right\}\right) = \tilde{\Omega}(n^{8/15}).$$

2. When $n^{2/15} \leq \beta \leq n^{2/5}$: Because $n^{1/15} \leq \sqrt{\beta} \leq n^{1/5}$, this means that

$$n^{4/15}\beta \leq n^{7/15}\beta^{1/2}$$

and

$$n^{8/15} \leq n^{7/15}\beta^{1/2}.$$

Now, again we do a case analysis. If $\frac{n\beta^{1/2}}{\tau} \geq n^{8/15}$, then

$$\tau \leq n^{7/15}\beta^{1/2},$$

which means that

$$\frac{\mathbf{Progress}_1(\beta, \tau, \gamma)}{\mathbf{Rounds}_1(\beta, \tau, \gamma)} = \tilde{\Omega}\left(\frac{n\beta}{\tau\gamma}\right) = \tilde{\Omega}\left(\frac{n\beta^{1/2}}{\tau}\right) \geq \tilde{\Omega}(n^{8/15}),$$

where in the second equality we have used that $\gamma = O(\sqrt{\beta})$ as per [Remark 7.7](#). Now, in the second case, we consider what happens if $\frac{n\beta^{1/2}}{\tau} \leq n^{8/15} \leq n^{7/15}\beta^{1/2}$. Then, we obtain that $\frac{n}{\tau} \leq n^{7/15}$, which means $\tau \geq n^{8/15}$. Likewise, because $\beta \leq n^{2/5}$, this means that

$$\frac{\tau^2}{\beta^2} \geq n^{8/15}.$$

Together, this implies that

$$\frac{\mathbf{Progress}_3(\beta, \tau, \gamma)}{\mathbf{Rounds}_3(\beta, \tau, \gamma)} = \tilde{\Omega}\left(\min\left\{\tau, \frac{\tau^2}{\beta^2}\right\}\right) = \tilde{\Omega}(n^{8/15}).$$

3. Finally, we consider what happens when $\beta \leq n^{2/15}$. Immediately, this implies that $\beta^{3/4} \leq n^{1/10}$, so we have

$$n^{4/15} \beta^{5/4} \leq n^{7/15} \beta^{1/2}.$$

Likewise, from [Remark 7.7](#), we know that $\gamma = O(\sqrt{\beta})$, so this means

$$\frac{n}{\gamma} = \Omega\left(\frac{n}{\beta^{1/2}}\right) = \Omega(n^{14/15}) \geq n^{8/15}.$$

Again, we have two cases. First, if $\tau \leq n^{7/15} \beta^{1/2}$, then

$$\frac{\tau}{\beta^{1/2}} \leq n^{7/15}.$$

Then, we can see

$$\frac{\mathbf{Progress}_1(\beta, \tau, \gamma)}{\mathbf{Rounds}_1(\beta, \tau, \gamma)} = \tilde{\Omega}\left(\frac{n\beta}{\tau\gamma}\right) = \tilde{\Omega}\left(\frac{n\beta^{1/2}}{\tau}\right) = \tilde{\Omega}(n^{8/15}),$$

where in the second equality we have used that $\gamma = O(\sqrt{\beta})$ as per [Remark 7.7](#). Otherwise, if $\tau \geq n^{4/15} \beta^{5/4}$, this means that

$$\frac{\tau^2}{\beta^{5/2}} \geq n^{8/15}.$$

In this case, we see that

$$\frac{\mathbf{Progress}_4(\beta, \tau, \gamma)}{\mathbf{Rounds}_4(\beta, \tau, \gamma)} = \tilde{\Omega}\left(\min\left\{\frac{n}{\gamma}, \frac{\tau^2}{\beta^2}\right\}\right) = \tilde{\Omega}(n^{8/15}).$$

Thus, in every case, we see that there is some choice of sub-routine which guarantees a progress to round ratio of $\tilde{\Omega}(n^{8/15})$. \square

Finally, we can conclude with our main theorem:

Theorem 7.9. *There is a randomized algorithm that, with high probability, finds a basis of any n -element matroid \mathcal{M} in $\tilde{O}(n^{7/15})$ adaptive rounds, using only polynomially many independence queries per round.*

Proof. We simply run [Algorithm 11](#), recovering the parameters γ, β, τ and sets S_1, \dots, S_k . If the algorithm does not return on [Line 9](#), we then invoke [Lemma 7.8](#): based on the values of β, γ, τ we can choose which subroutine to run. The result is that we have invested some $\kappa \geq 1$ adaptive rounds (and using $\text{poly}(n)$ queries in each round), but by [Lemma 7.8](#), we are guaranteed to have made at least $\tilde{\Omega}(\kappa \cdot n^{8/15})$ progress. In particular, this means that we have either recovered $\tilde{\Omega}(\kappa \cdot n^{8/15})$ redundant elements (which we simply delete), or we have recovered $\tilde{\Omega}(\kappa \cdot n^{8/15})$ independent elements, which we then contract on. In either case, we reduce the problem of computing a basis of \mathcal{M} on n elements, to computing a basis on a matroid with $n - \tilde{\Omega}(\kappa \cdot n^{8/15})$ elements.

On the other hand, if the algorithm returns on [Line 9](#), by [Claim 7.2](#), we can recover an independent set of size $\tilde{\Omega}(n)$ and contract. As we have invested $\kappa = O(k) = O(n^{1/3})$ rounds by [Claim 4.15](#), we also reduce the problem of computing a basis of \mathcal{M} on n elements, to computing a basis on a matroid with $n - \tilde{\Omega}(n) \leq n - \tilde{\Omega}(\kappa \cdot n^{8/15})$ elements.

Thus, the total number of adaptive rounds required (denoted $T(n)$) obeys the recurrence

$$T(n) \leq \kappa + T\left(n - \tilde{\Omega}(\kappa \cdot n^{8/15})\right).$$

In particular, this means that $T(n) = \tilde{O}(n^{7/15})$ adaptive rounds suffice, as we desire. Note that the failure probability in each subroutine is exponentially small $1/2^{\Omega(n)}$ (and can even be amplified by repeating each subroutine $\text{poly}(n)$ times in parallel), and so there is no concern of cascading errors becoming too large through rounds. Likewise, via [Claim 4.4](#), the error probability of our decomposition is also exponentially small. This yields the theorem. \square

8 Conclusions

We have presented the first progress in nearly four decades on the parallel complexity of finding bases in general matroids under independence-oracle access. Our main result improves the longstanding $O(\sqrt{n})$ upper bound of [\[KUW85\]](#), achieving an $\tilde{O}(n^{7/15})$ -round algorithm with polynomial query complexity. As a corollary, we obtain a new upper bound for matroid intersection by integrating our techniques into the reduction of [\[BT25\]](#). We also match the $\Omega(n^{1/3})$ lower bound for partition matroids by giving an $\tilde{O}(n^{1/3})$ -round algorithm for this class.

Our results are achieved through a new matroid decomposition framework, a probabilistic analysis of rank deficiency via random sampling, and efficient parallel routines for contraction and deletion. We believe these techniques will pave the way for further algorithmic developments in the independence-oracle model for matroids.

References

- [Bli22] Joakim Blikstad. Sublinear-round parallel matroid intersection. In Mikolaj Bojanczyk, Emanuela Merelli, and David P. Woodruff, editors, *49th International Colloquium on Automata, Languages, and Programming, ICALP 2022, July 4-8, 2022, Paris, France*, volume 229 of *LIPIcs*, pages 25:1–25:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.
- [BPVdP15] Nikhil Bansal, Rudi A. Pendavingh, and Jorn G. Van der Pol. On the number of matroids. *Combinatorica*, 35:253–277, 2015.
- [BS20] Eric Balkanski and Yaron Singer. A lower bound for parallel submodular minimization. In Konstantin Makarychev, Yury Makarychev, Madhur Tulsiani, Gautam Kamath, and Julia Chuzhoy, editors, *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing, STOC 2020, Chicago, IL, USA, June 22-26, 2020*, pages 130–139. ACM, 2020.
- [BT25] Joakim Blikstad and Ta-Wei Tu. Efficient matroid intersection via a batch-update auction algorithm. In Ioana Oriana Bercea and Rasmus Pagh, editors, *2025 Symposium on Simplicity in Algorithms, SOSA 2025, New Orleans, LA, USA, January 13-15, 2025*, pages 226–237. SIAM, 2025.
- [CCK21] Deeparnab Chakrabarty, Yu Chen, and Sanjeev Khanna. A polynomial lower bound on the number of rounds for parallel submodular function minimization. In *62nd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2021, Denver, CO, USA, February 7-10, 2022*, pages 37–48. IEEE, 2021.
- [CGJS22] Deeparnab Chakrabarty, Andrei Graur, Haotian Jiang, and Aaron Sidford. Improved lower bounds for submodular function minimization. In *63rd IEEE Annual Symposium*

on *Foundations of Computer Science, FOCS 2022, Denver, CO, USA, October 31 - November 3, 2022*, pages 245–254. IEEE, 2022.

- [Doe20] Benjamin Doerr. Probabilistic tools for the analysis of randomized optimization heuristics. *Theory of evolutionary computation: Recent developments in discrete optimization*, pages 1–87, 2020.
- [Ehm91] Werner Ehm. Binomial approximation to the poisson binomial distribution. *Statistics & Probability Letters*, 11(1):7–16, 1991.
- [FGT16] Stephen A. Fenner, Rohit Gurjar, and Thomas Thierauf. Bipartite perfect matching is in quasi-nc. In Daniel Wichs and Yishay Mansour, editors, *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*, pages 754–763. ACM, 2016.
- [GGR22] Sumanta Ghosh, Rohit Gurjar, and Roshan Raj. A deterministic parallel reduction from weighted matroid intersection search to decision. In Joseph (Seffi) Naor and Niv Buchbinder, editors, *Proceedings of the 2022 ACM-SIAM Symposium on Discrete Algorithms, SODA 2022, Virtual Conference / Alexandria, VA, USA, January 9 - 12, 2022*, pages 1013–1035. SIAM, 2022.
- [GT17] Rohit Gurjar and Thomas Thierauf. Linear matroid intersection is in quasi-nc. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 821–830, 2017.
- [Kar93] David R. Karger. Global min-cuts in **RNC**, and other ramifications of a simple min-cut algorithm. In Vijaya Ramachandran, editor, *Proceedings of the Fourth Annual ACM/SIGACT-SIAM Symposium on Discrete Algorithms, 25-27 January 1993, Austin, Texas, USA*, pages 21–30. ACM/SIAM, 1993.
- [KUW85] Richard M. Karp, Eli Upfal, and Avi Wigderson. The complexity of parallel computation on matroids. In *26th Annual Symposium on Foundations of Computer Science, Portland, Oregon, USA, 21-23 October 1985*, pages 541–550. IEEE Computer Society, 1985.
- [KUW86] Richard M. Karp, Eli Upfal, and Avi Wigderson. Constructing a perfect matching is in random NC. *Comb.*, 6(1):35–48, 1986.
- [KUW88] Richard M. Karp, Eli Upfal, and Avi Wigderson. The complexity of parallel search. *J. Comput. Syst. Sci.*, 36(2):225–253, 1988.
- [Lov79] László Lovász. On determinants, matchings, and random algorithms. In Lothar Budach, editor, *Fundamentals of Computation Theory, FCT 1979, Proceedings of the Conference on Algebraic, Arithmetic, and Categorical Methods in Computation Theory, Berlin/Wendisch-Rietz, Germany, September 17-21, 1979*, pages 565–574. Akademie-Verlag, Berlin, 1979.
- [Lub86] Michael Luby. A simple parallel algorithm for the maximal independent set problem. *SIAM J. Comput.*, 15(4):1036–1053, 1986.
- [Oxl06] James G Oxley. *Matroid theory*, volume 3. Oxford University Press, USA, 2006.

- [Qua24] Kent Quanrud. *Quotient sparsification for submodular functions*, pages 5209–5248. SIAM, 2024.
- [ST17] Ola Svensson and Jakub Tarnawski. The matching problem in general graphs is in quasi-nc. In Chris Umans, editor, *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*, pages 696–707. IEEE Computer Society, 2017.

A Proof of Theorem 1.4

To start, we recall the following lemma of [BT25]:

Lemma A.1 ([BT25], Fact 2.4 and Corollary 3.8). *Let $\mathcal{M}_1 = (E, \mathcal{I}_1), \mathcal{M}_2 = (E, \mathcal{I}_2)$ be 2 matroids. Let $n = |E|$ and r be the size of the size of the largest independent set of $\mathcal{M}_1, \mathcal{M}_2$. Then*

- *There is an $\tilde{O}\left(\frac{nT(n)}{\varepsilon\Delta}\right)$ rounds independence-query algorithm that finds a common independent set $S \in \mathcal{I}_1 \cap \mathcal{I}_2$ of size $|S| \geq r - (\varepsilon r + \Delta)$, given that there is a $T(n)$ round independence-query algorithm that finds a maximum weight basis of a matroid on n elements.*
- *Given $S \in \mathcal{I}_1 \cap \mathcal{I}_2$, in a single round of independence query, one can compute an $S' \in \mathcal{I}_1 \cap \mathcal{I}_2$ of size $|S'| = |S| + 1$ or decide that S is of maximum possible size.*

Now, we show Theorem 1.4:

Theorem A.2 (see also [BT25] Theorem 1.4). *[Theorem 1.4 restated] There is a randomized algorithm that, with high probability, finds a maximum common independent set of two n -element matroids in $\tilde{O}(n^{37/45})$ adaptive rounds, using only polynomially many independence queries per round.*

Proof. First, as in Lemma 2.2 of [BT25], observe that an r -round independence query algorithm for computing the basis of an n element matroid immediately yields an r round algorithm for computing a *maximum weight* basis of an n element matroid. Thus, we can use our $\tilde{O}(n^{7/15})$ round algorithm to also find maximum weight bases.

To proceed, we then set $\varepsilon = n^{22/45}r^{-2/3}$ and $\Delta = \varepsilon r = n^{22/45}r^{1/3}$. We can first find an $S \in \mathcal{I}_1 \cap \mathcal{I}_2$ of size $|S| \geq r - (\varepsilon r + \Delta)$ in

$$\tilde{O}\left(\frac{n \cdot n^{7/15}}{\varepsilon\Delta}\right) = \tilde{O}(n^{22/45}r^{1/3})$$

rounds and then augment it to optimal in $O(\varepsilon r + \Delta) = O(n^{22/45}r^{1/3})$ rounds. As $r \leq n$, the total rounds of adaptivity needed is $\tilde{O}(n^{37/45})$. \square