JMLR: Workshop and Conference Proceedings vol 49:1–23, 2016

Pure Exploration of Multi-armed Bandit Under Matroid Constraints
[Extended Abstract]*

Lijie Chen[†] and **Anupam Gupta**[‡] and **Jian Li**[†]

Abstract

We study the pure exploration problem subject to a matroid constraint (BEST-BASIS) in a stochastic multi-armed bandit game. In a BEST-BASIS instance, we are given n stochastic arms with unknown reward distributions, as well as a matroid $\mathcal{M}$ over the arms. Let the weight of an arm be the mean of its reward distribution. Our goal is to identify a basis of $\mathcal{M}$ with the maximum total weight, using as few samples as possible. The problem is a significant generalization of the best arm identification problem and the top-k arm identification problem, which have attracted significant attentions in recent years. We study both the exact and PAC versions of BEST-BASIS, and provide algorithms with nearly-optimal sample complexities for these versions. Our results generalize and/or improve on several previous results for the top-k arm identification problem and the combinatorial pure exploration problem when the combinatorial constraint is a matroid.

Keywords: matroid, multi-armed bandit, pure exploration

1. Introduction

The stochastic multi-armed bandit is a classical model for characterizing the exploration-exploitation tradeoff in many decision-making problems in stochastic environments. The popular objectives include maximizing the cumulative sum of rewards, or minimizing the cumulative regret (see e.g., Cesa-Bianchi and Lugosi (2006); Bubeck and Cesa-Bianchi (2012)). However, in many application domains, the exploration phase and the evaluation phase are separated. The decision-maker can perform a *pure-exploration phase* to identify an optimal (or nearly optimal) solution, and then keep exploiting this solution. Such problems arise in application domains such as medical trials Robbins (1985); Audibert and Bubeck (2010), communication network Audibert and Bubeck (2010), crowdsourcing Zhou et al. (2014); Cao et al. (2015). In particular, the problem of identifying the single best arm in a stochastic bandit game has been has received considerable attention in recent years Audibert and Bubeck (2010); Even-Dar et al. (2006); Mannor and Tsitsiklis (2004); Jamieson et al. (2014); Karnin et al. (2013); Chen and Li (2015). The generalization to identifying the top-k arms has also been studied extensively Gabillon et al. (2012); Kalyanakrishnan et al. (2012); Kaufmann and Kalyanakrishnan (2013); Kaufmann et al. (2014); Zhou et al. (2014); Cao et al. (2015). Since these problems are closely related to the problem we study in the paper, we formally define it as follows.

* Extended abstract. Full version appears as [arXiv:1605.07162, v3].

† Institute for Interdisciplinary Information Sciences (IIIS), Tsinghua University, Beijing, China. Research supported in part by the National Basic Research Program of China grants 2015CB358700, 2011CBA00300, 2011CBA00301, and the National NSFC grants 61033001, 61361136003.

‡ Department of Computer Science, Carnegie Mellon University, Pittsburgh, USA. Research partly supported by NSF awards CCF-1016799 and CCF-1319811.

Problem 1 *(*Best-k-Arm*) There are n unknown distributions $\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_n$, all supported on $[0, 1]$. Let the mean of $\mathcal{D}_i$ be μ_i. At each step we choose a distribution and get an i.i.d. sample from the distribution. Our goal is to find the k distributions with the largest means (exactly or approximately), with probability at least $1 - \delta$, using as few samples as possible.*

The distributions above are also called *arms* in the multi-armed bandit literature. We denote the k^{th} largest mean by $\mu_{[k]}$. In addition, we assume $\mu_{[k]}$ and $\mu_{[k+1]}$ are different (so the optimal top-k answer is unique).

In certain applications such as online ad allocations, there is a natural combinatorial constraint over the set of arms, and we can only choose a subset of arms subject to the given constraint (Best-k-Arm simply involves a cardinality constraint). Motivated by such applications, Chen et al. (2014) introduced the *combinatrial pure exploration* problem. They considered the general setting with arbitrary combinatorial constraint, and propose several algorithms. In this paper, we consider the same problem under a *matroid* constraint, one of the most popular combinatorial constraint. The matroid constraint was also discuss in length in Chen et al. (2014).

The notion of matroid (see Section 2 for the definition) is an abstraction of many combinatorial structures, including the sets of linearly independent vectors in a given set of vectors, the sets of spanning forests in an undirected graph and many others. We note that Best-k-Arm is a special case of a matroid constraint, since all subsets of size of at most k form a *uniform matroid*. Now, we formally define the *matroid pure exploration bandit problem* as follows.

Definition 1.1 (Best-Basis) *In a* Best-Basis *instance $\mathcal{S} = (S, \mathcal{M})$, we are given a set S of n arms. Each arm $a \in S$ is associated with an unknown reward distribution $\mathcal{D}_a$, supported on $[0, 1]$, with mean μ_a (which is unknown as well). Without loss of generality, we assume all arms have distinct means.*

We are also given a matroid $\mathcal{M} = (S, \mathcal{I})$ with ground set identified with the set S of arms. The weight function $\mu : S \to \mathbb{R}^+$ simply sets the weight of a to be the mean of $\mathcal{D}_a$; i.e., $\mu(a) = \mu_a$ for all $a \in S$. The weights are initially unknown, and are only learned by sampling arms. Our goal is to find a basis (a.k.a. a maximal independent set) of the matroid with the maximum total weight/cost (exactly or approximately), with probability at least $1 - \delta$, using as few samples as possible.

Besides including Best-k-Arm as a special case, the Best-Basis problem also captures the following natural problems, motivated by various applications.

1. Suppose we have m disjoint groups $G_1, \ldots, G_m$ of arms, and we would like to pick the best k_i arms from group G_i (k_is are given integers). This is exact the best-basis problem for a Partition Matroid. Note that PAC version of the problem is not just a disjoint collection of best-k-arm problems.

   We also note that the special case where $k_i = 1$ has been studied in Gabillon et al. (2011); Bubeck et al. (2012) (under the fixed budget setting). They are motivated by a clinical problem with m subpopulations, where one would like to decide the best k_i treatments from the options available for subjects from each subpopulation.

2. Beside the above constraints for the groups, we may have an additional global constraint on the total number of arms we can choose. This is a special case of Laminar Matroids.

3. An application mentioned in Chen et al. (2014): Consider a network where the delay of the links are stochastic. A network routing system wants to build a minimum spanning tree to connect all nodes, where the weight of each edges are expected delay of that link. A spanning tree is a basis in a Graphical Matroid.

4. Consider a set of workers and a set of tasks. Each worker is able to do only a subset of tasks (which defines a worker-task bipartite graph). Each worker must be assigned to one task (so we need to build a matching between the workers and the tasks) and the reward of a task is stochastic. We would like to identify the set of tasks that can be completed by the set of workers and have maximum total reward. The combinatorial structure (over the subsets of tasks) is a Transversal Matroid. The problem (or variants) may find applications in crowdsourcing or online advertisement.

There are two natural formulations of the BEST-BASIS problem: in one, we need to identify the unique optimal basis with a certain confidence, and in some others, we can settle for an approximate optimal basis (the PAC setting). We now formally define these problems, and present our results.

1.1. Identifying the Exact Optimal basis

Definition 1.2 *(*EXACT-BASIS*) Given a* BEST-BASIS *instance $\mathcal{S} = (S, \mathcal{M})$ and a confidence level $\delta > 0$, the goal is to output the optimal basis of $\mathcal{M}$ (one that maximizes $\sum_{a \in I} \mu_a$) with probability at least $1 - \delta$, using as few samples as possible.*

Without loss of generality, assume that matroid $\mathcal{M}$ has no isolated elements (i.e., elements that are included in every basis) and no loops (i.e., elements that belong to no basis), since we can always include or ignore them without affecting the solution. We use $\mathsf{OPT}(\mathcal{M})$ to denote the optimal basis (as well as the optimal total weight) for matroid $\mathcal{M}$. For a subset of elements $F \subseteq S$, let $\mathcal{M}_F$ denote the restriction of $\mathcal{M}$ to F, and $\mathcal{M}_{/F}$ denote the contraction of $\mathcal{M}$ by F (see Definition 2.3). Note that $\mathsf{OPT}(\mathcal{M}_{/\{e\}}) + \mu(e)$ is the optimal cost among all bases including e.

Naturally, the sample complexity of an algorithm for EXACT-BASIS depends on the parameters of the problem instance. In particular, we need to define the following gap parameter.

Definition 1.3 (Gap) *Given a matroid $\mathcal{M} = (S, \mathcal{I})$ with cost function $\mu : S \to \mathbb{R}^+$, such that all costs are distinct, define the* gap *of an element $e \in S$ to be*

$$\Delta_e^{\mathcal{M},\mu} := \begin{cases} \mathsf{OPT}(\mathcal{M}) - \mathsf{OPT}(\mathcal{M}_{S \setminus \{e\}}) & e \in \mathsf{OPT}(\mathcal{M}) \\ \mathsf{OPT}(\mathcal{M}) - (\mathsf{OPT}(\mathcal{M}_{/\{e\}}) + \mu(e)) & e \notin \mathsf{OPT}(\mathcal{M}) \end{cases}$$

Intuitively, for an element $e \in \mathsf{OPT}(\mathcal{M})$, its gap is the loss if we do not select e, whereas for an element $e \notin \mathsf{OPT}(\mathcal{M})$, its gap is the loss if we are forced to select e. Since we assume that elements have distinct weights, $\Delta_e > 0$ for all arms e. We note that Definition 1.3 is the same as the gap definition in Chen et al. (2014) and generalizes the gaps defined for the BEST-k-ARM problem used in Kalyanakrishnan et al. (2012) (in BEST-k-ARM, the gap of an arm e to be $\Delta_e = \mu_e - \mu_{[k+1]}$ if e is a top-k arm, and $\Delta_e = \mu_{[k]} - \mu_e$ otherwise).

Chen et al. (2014) obtained an algorithm with sample complexity

$$\left(\sum_{e \in S} \Delta_e^{-2} (\ln \delta^{-1} + \ln n + \ln \sum\nolimits_{e \in S} \Delta_e^{-2}) \right),$$

when specialized to EXACT-BASIS. [1] We improve upon their result by proving the following theorem.

Theorem 1.4 (Main Result for Exact Identification) *There is an algorithm for* EXACT-BASIS, *that returns the optimal basis for $\mathcal{S}$, with probability at least $1-\delta$, and uses at most*

$$O\left(\sum_{e\in S}\Delta_e^{-2}(\ln\delta^{-1}+\ln k+\ln\ln\Delta_e^{-1})\right)$$

samples. Here, $k=\operatorname{rank}(\mathcal{M})$ is the size of a basis of $\mathcal{M}$.

Observe that the dependence is now on the rank of the matroid k, rather than the number of elements n which may be much larger than k. Moreover, the dependence on Δ_e is doubly logarithmic.

For the special case of the k-uniform matroid, the problem becomes the BEST-k-ARM problem, for which the current state-of-the-art is $O(\sum_{i=1}^n \Delta_{[i]}^{-2}(\ln\delta^{-1}+\ln\sum_{i=1}^n\Delta_{[i]}^{-2}))$, obtained by Kalyanakrishnan et al. (2012). Theorem 1.4 improves upon this result for the typical case when $\ln\sum_{i=1}^n\Delta_{[i]}^{-2}$ is larger than $\ln k$. Theorem 1.4 also matches the recent upper bound of $O(\sum_{i=2}^n\Delta_{[i]}^{-2}(\ln\ln\Delta_{[i]}^{-1}+\ln\delta^{-1}))$ for BEST-1-ARM, due to Karnin et al. (2013) and Jamieson et al. (2014).

Chen et al. (2014) proved an $\Omega(\sum_{e\in S}\Delta_e^{-2}\ln\delta^{-1})$ lower bound for the problem. Moreover, Kalyanakrishnan et al. (2012) showed an $\Omega(n\varepsilon^{-2}(\ln\delta^{-1}+\ln k))$ lower bound for a PAC version (the EXPLORE-k metric, see Section 1.2) of BEST-k-ARM. Indeed, in their lower bound instances, all arms have gap $\Delta_e=\varepsilon$. If we apply our exact algorithm on those instances, the sample complexity is $O(n\varepsilon^{-2}(\ln\delta^{-1}+\ln k+\ln\ln\varepsilon^{-1}))$. Hence, the first two terms of our upper bound are probably necessary in light of the above lower bounds.

1.2. The PAC setting

Next we discuss our results for the PAC setting. Several notions of approximation were used for the special case of BEST-k-ARM, when we return a set I of k arms. Kalyanakrishnan et al. (2012) required that the mean of every arm in I be at least $\mu_{[k]}-\varepsilon$ (The EXPLORE-k metric). Zhou et al. (2014) required that the average mean $\frac{1}{k}\sum_{e\in I}\mu_e$ of I be at least $\frac{1}{k}\sum_{i=1}^k\mu_{[i]}-\varepsilon$; we call such a solution an *average-ε-optimal* solution. Finally, Cao et al. (2015) proposed a stronger metric that required the mean of the i^{th} arm in I be at least $\mu_{[i]}-\varepsilon$, for all $i\in[k]$. This notion, which we call *elementwise-ε-optimality* extends to general matroids: we need that i^{th} largest arm in our solution is at least the i^{th} largest mean in the optimal solution minus ε.

In this paper we introduce the stronger notion of an *ε-optimal* solution.

Definition 1.5 (PAC-BASIS *and ε-optimality) We are given a matroid $\mathcal{M}=(S,\mathcal{I})$ with cost function $\mu: S\to\mathbb{R}^+$. We say a basis I is ε*-optimal *(with respect to μ), if I is an optimal solution for the modified cost function $\mu_{I,\varepsilon}$, defined as follows:*

$$\mu_{I,\varepsilon}(e)=\begin{cases}\mu(e)+\varepsilon & \text{for } e\in I\\ \mu(e) & \text{for } e\notin I.\end{cases}$$

In other words, if we add ε to each element in I, I would become an optimal solution.

1. Their algorithm works for arbitrary combinatorial constraint. The sample complexity depends on a *width* parameter of the constraint, which is roughly the number of elements needed to be exchanged from one feasible solution to another. The width can be as large as n. For a matroid, the width is 2.

The proof of the following proposition can be found in the appendix.

Proposition 1.6 *For a* BEST-BASIS *instance, an ε-optimal solution is also elementwise-ε-optimal. The converse is not necessarily true.*

Theorem 1.7 (Main Result for PAC Setting) *There is an algorithm for* PAC-BASIS *which returns an ε-optimal solution for $\mathcal{S} = (S, \mathcal{M})$, with probability at least $1 - \delta$, and uses at most*

$$O(n\varepsilon^{-2} \cdot (\ln k + \ln \delta^{-1}))$$

samples, where $k = \mathrm{rank}(\mathcal{M})$.

This theorem generalizes and strengthens the results in Kalyanakrishnan et al. (2012); Cao et al. (2015), in which the same sample complexity was obtained for BEST-k-ARM under EXPLORE-k and elementwise-ε-optimality metrics, respectively. In fact, this sample complexity is optimal, since an $\Omega(n\varepsilon^{-2}(\ln k + \ln \delta^{-1}))$ lower bound is known for EXPLORE-k for the special case of BEST-k-ARM, due to Kalyanakrishnan et al. (2012).

1.2.1. AVERAGE-ε-OPTIMALITY

We also consider the weaker notion of average-ε-optimality, which may suffice for certain applications. For this definition, we give another algorithm with a lower sample complexity.

Definition 1.8 *(*PAC-BASIS-AVG*) Given a matroid $\mathcal{M} = (S, \mathcal{I})$ with cost function $\mu : S \to \mathbb{R}^+$. Suppose $k = \mathrm{rank}(\mathcal{M})$. We say a basis I is an average-ϵ-optimal solution (w.r.t. μ), if: $\frac{1}{k}\sum_{e \in I} \mu(e) \ge \frac{1}{k}\mathsf{OPT}(\mathcal{M}) - \varepsilon$.*

Theorem 1.9 *There is an algorithm for* PAC-BASIS-AVG, *which can return an average-ϵ-optimal solution for $\mathcal{S}$, with probability at least $1 - \delta$, and its sample complexity is at most*

$$O\Big(\big(n \cdot (1 + \ln \delta^{-1}/k) + (\ln \delta^{-1} + k)(\ln k \ln\ln k + \ln \delta^{-1} \ln\ln \delta^{-1})\big)\varepsilon^{-2}\Big).$$

In particular, when $k \ln \delta^{-1} \le O(n^{0.99})$ and $\delta \ge \Omega(\exp(-n^{0.49}))$, the sample complexity is

$$O(n\varepsilon^{-2}(1 + \ln \delta^{-1}/k)).$$

Zhou et al. (2014) obtained matching upper and lower bounds of $\Omega(n\varepsilon^{-2}(1+\ln \delta^{-1}/k))$ for BEST-k-ARM under the average metric. Our result matches their result when δ is not extremely small and k is not very close to n. Obtaining tight upper and lower bounds for all range of parameters is left as an interesting open question. We omit the proof of Theorem 1.9 in this extended abstract. The details can be found in the full version of the paper.

1.2.2. PRIOR AND OUR TECHNIQUES

Several prior algorithms for the PAC versions of BEST-1-ARM and BEST-k-ARM (e.g., Karnin et al. (2013); Zhou et al. (2014); Even-Dar et al. (2002)) were elimination-based, roughly using the following framework: In the r^{th} round, we sample each remaining arm Q_r times, [2] and eliminate all arms whose empirical means fall below a certain threshold. This threshold can be either a

2. Typically, Q_r increases exponentially with r.

percentile, as in Even-Dar et al. (2002); Zhou et al. (2014) or an ε-optimal arm obtained by some PAC algorithm, such as in Karnin et al. (2013). After eliminating some arms, we proceed to the next round. Small variations to this procedure are possible, e.g., if the number of remaining arms is not much larger than k, we can directly use the naïve uniform sampling algorithm. A main difference in prior works is in their analysis, due to the different PAC-optimality metrics. However, we cannot easily extend this framework to either PAC-BASIS or PAC-BASIS-AVG, since it is not clear how to eliminate even a small constant fraction of arms while ensuring that the optimal value for the remaining set does not drop. Indeed, due to the combinatorial structure of the matroid, we cannot perform elimination based solely on fixed thresholds.

We resolve the issue by applying a sampling-and-pruning technique developed by Karger, and used by Karger, Klein, and Tarjan in their expected linear-time randomized algorithm for minimum s-panning tree. Here is the high-level idea, in the context of the PAC-BASIS problem. We pick a random subset F by including each arm independently with some small constant probability p, and recursively find an $\varepsilon/3$-optimal basis I for the subset F. The key idea is that this basis I can be used to eliminate a significant proportion of arms, while ensuring that the remaining set still contains a desirable solution. Hence, after eliminating those arms, we can recurse on the remaining arms. Unlike the previous algorithms which eliminate arms based on a single threshold, we perform the elimination based on the solution I of a random subset. We feel this extension of the sampling and pruning technique to bandit problems will find other applications.

Another popular approach for pure exploration problems is based on upper or lower confidence bounds (UCB or LUCB) (see e.g., Kalyanakrishnan et al. (2012); Chen et al. (2014)). While being very flexible and easy to apply, the analysis of all such bounds inevitably requires a union bound of all rounds (which is at least n), thus incurring a $\log n$ factor, which is worse than the optimal $\log k$ factor that we obtain.

1.3. Other Related Work

The problem of identifying the single best arm, a very special case of our problem, has been studied extensively. For the PAC version of the problem, [3] Even-Dar et al. (2002) provided an algorithm with sample complexity $O(n\varepsilon^{-2} \cdot \ln \delta^{-1})$, which is also optimal. For the exact version, Mannor and Tsitsiklis (2004) proved a lower bound of $\Omega(\sum_{i=2}^{n} \Delta_{[i]}^{-2} \ln \delta^{-1})$. Farrell (1964) showed a lower bound of $\Omega(\Delta_{[2]}^{-2} \ln \ln \Delta_{[2]}^{-1})$ even if there are only two arms. Karnin et al. (2013) obtained an upper bound of $O(\sum_{i=2}^{n} \Delta_{[i]}^{-2}(\ln \ln \Delta_{[i]}^{-1} + \ln \delta^{-1}))$, matching Farrell's lower bound for two arms. Jamieson et al. (2014) obtained the same result using a UCB-like algorithm. Very Recently, Chen and Li (2015) provided a new lower bound of $\Omega(\sum_{i=2}^{n} \Delta_{[i]}^{-2} \ln \ln n)$ and an improved upper bound of $O\Big(\Delta_{[2]}^{-2} \ln \ln \Delta_{[2]}^{-1} + \sum_{i=2}^{n} \Delta_{[i]}^{-2} \ln \delta^{-1} + \sum_{i=2}^{n} \Delta_{[i]}^{-2} \ln \ln \min(n, \Delta_{[i]}^{-1})\Big)$.

In all aforementioned results, we require that the (PAC or exact) algorithm returns a correct answer with probability at least $1 - \delta$. This is called the *fixed confidence* setting in the literature. Another popular setting is the *fixed budget* setting, in which the total number of samples is subject to a given budget constraint, and we would like to minimize the failure probability (see e.g., Bubeck et al. (2013); Gabillon et al. (2012); Karnin et al. (2013); Chen et al. (2014)). Some prior work (Audibert and Bubeck (2010); Bubeck et al. (2013); Audibert et al. (2013)) also considered the objective of

3. Since the solution only contains one arm, all different notions of PAC optimality mentioned in Section 1.2 are equivalent.

making the *expected simple regret* at most ε (i.e., $\frac{1}{k}(\sum_{i=1}^{k}\mu_{[i]} - \mathbb{E}[\sum_{a\in T}\mu_a]) \le \varepsilon$), which is a somewhat weaker objective.

There is a large body of work on minimizing the cumulative regret in online multi-armed bandit games with various combinatorial constraints in different feedback settings (see e.g., Cesa-Bianchi and Lugosi (2006); Bubeck and Cesa-Bianchi (2012); Cesa-Bianchi and Lugosi (2012); Audibert et al. (2013); Chen et al. (2013) and the references therein). In an online bandit game, there are T rounds. In the t^{th} round, we can play a combinatorial subset S_i of arms. The goal is to minimize $T\sum_{a\in\mathsf{OPT}}\mu_a - \sum_{t=1}^{T}\sum_{a\in S_t}\mu_a$. We note that it is possible to obtain an expected simple regret of ε for BEST-BASIS, with at most $O(n\varepsilon^{-2})$ samples, using the semi-bandit regret bound in Audibert et al. (2013). In particular, they provided an online mirror descent algorithm and showed a cumulative regret of $\sqrt{knT}$ in the semi-bandit feedback setting (i.e., we can only observe the rewards from the arms we played), where k is the maximum cardinality of a feasible set. By setting $T = nk^{-1}\varepsilon^{-2}$, we get a cumulative regret of n/ε. If we uniformly randomly pick a solution from $\{S_t\}_{t\in[T]}$, we can see that $\mathbb{E}_{t\in[T]}\frac{1}{k}(\sum_{a\in\mathsf{OPT}}\mu_a - \sum_{a\in S_t}\mu_a) \le \varepsilon$. One drawback of their algorithm is that it needs to solve a convex program over the matroid polytope, which can be computationally expensive, while our algorithm is purely combinatorial and very easy to implement.

In recent and concurrent work, Gabillon et al. Gabillon et al. (2016) proposed a new complexity notion for the general combinatorial pure exploration problem, and developed new algorithms in both fixed budget and the fixed confidence setting. They showed that in some cases, the sample complexity of their algorithms is better than that of Chen et al. (2014). While the current implementations of their algorithm have an exponential running time, even for general matroid constraints, it is an interesting problem to get more efficient algorithms, and to combine their notion of complexity with our techniques.

2. Preliminaries

2.1. Useful Facts about Matroids

While there are many equivalent definitions for matroids, we find this one most convenient.

Definition 2.1 (Matroid) *A matroid $\mathcal{M}(S,\mathcal{I})$ consists of a finite set S (called the* ground set*), and a non-empty family $\mathcal{I}$ of subsets of S (with sets in $\mathcal{I}$ being called* independent sets*), satisfying the following:*

i. Any subset of an independent set is an independent set.
ii. Given two sets $I, J \in \mathcal{I}$, if $|I| > |J|$, there exists element $e \in I \setminus J$ such that $J \cup \{e\} \in \mathcal{I}$.

For convenience, we often write $I \in \mathcal{M}$ instead of $I \in \mathcal{I}$ to denote that I is an independent set of $\mathcal{M}$. An independent set is *maximal* if it is not a proper subset of another independent set; a maximal independent set is called a *basis*.

Definition 2.2 (Rank) *Given matroid $\mathcal{M}(S,\mathcal{I})$ and set $A \subseteq S$, the* rank *of A, denoted by* $\mathrm{rank}_{\mathcal{M}}(A)$*, is the cardinality of a maximal independent subset contained in A.*

When $\mathcal{M}$ is clear from context, we merely write $\mathrm{rank}(A)$. All bases of a matroid have the same cardinality. We use $\mathrm{rank}(\mathcal{M})$, instead of $\mathrm{rank}(S)$, to denote the cardinality of every basis of $\mathcal{M}$.

We often need to work with the set of independent sets restricted to a subset of elements. Sometimes we can determine to include some elements as a partial solution, we need to work the the rest of the matroid, conditioning on the partial solution. We need the definitions of matroid restrictions and matroid contractions to formalize the above situations.

Definition 2.3 (Matroid restrictions and contractions) *Let $\mathcal{M}(S, \mathcal{I})$ be a matroid. For $A \subseteq S$, we define the restriction of $\mathcal{M}$ to A as follows: $\mathcal{M}_A$ is also a matroid with ground set A; an independent set of $\mathcal{M}$ which is a subset of A is an independent of $\mathcal{M}_A$.*

The contraction of A is defined as follows: $\mathcal{M}_{/A}$ is the matroid with ground set $S' = \{e \in S \mid \mathrm{rank}(\{e\} \cup A) > \mathrm{rank}(A)\}$, and the independent set family $\mathcal{I}' = \{I \subseteq S' \mid \mathrm{rank}(I \cup A) = |I| + \mathrm{rank}(A)\}$.

Both $\mathcal{M}_A$ and $\mathcal{M}_{/A}$ are indeed matroids. Sometimes, we may also write $\mathcal{M}|A$ and $\mathcal{M}/A$ to avoid successive subscripts. In our paper, we only need to contract an independent set $A \in \mathcal{I}$. In this case, $\mathrm{rank}(A) = |A|$, and the definition simplifies to the following: a set I (disjoint from A) is independent in $\mathcal{M}_{/A}$, if $I \cup A$ is independent in $\mathcal{M}$.

Definition 2.4 (Isolated Elements and Loops) *For a matroid $\mathcal{M} = (S, \mathcal{I})$ and element $e \in S$, we say e is an* isolated element*, if it is contained in all bases of $\mathcal{M}$ (or equivalently, $\mathrm{rank}(S) > \mathrm{rank}(S \setminus \{e\})$). We say e is a* loop *if it belongs to no basis of $\mathcal{M}$.*

Clearly, since the mean of each arm is nonnegative, we can directly select all isolated elements and contract out these elements. Also, we can simply ignore those loops. From now on, we can assume without loss of generality that there is no isolated element or loop in $\mathcal{M}$.

Definition 2.5 (Block) *Let $\mathcal{M}(S, \mathcal{I})$ be a matroid. Given a subset $A \subset S$ and an element e such that $e \notin A$, we say A blocks e, if $\mathrm{rank}_{\mathcal{M}}(A \cup \{e\}) = \mathrm{rank}_{\mathcal{M}}(A)$.*

Intuitively, e is blocked by A if adding e is not useful in increasing the cardinality of the maximal independent set in A. Note that, if $A \subseteq B$, $e \notin B$ and A blocks e, then clearly B also blocks e, due to the submodularity of rank: $\mathrm{rank}(B \cap \{e\}) - \mathrm{rank}(B) \leq \mathrm{rank}(A \cap \{e\}) - \mathrm{rank}(A)$.

We have the following lemma characterizing when a subset A blocks an element e.

Lemma 2.6 *If A blocks e, every basis I of $\mathcal{M}_A$ blocks e.*

Proof Since A blocks e, $\mathrm{rank}_{\mathcal{M}}(A \cup \{e\}) = \mathrm{rank}_{\mathcal{M}}(A)$. Consider a basis I of $\mathcal{M}_A$. $\mathrm{rank}_{\mathcal{M}}(I \cup \{e\}) \leq \mathrm{rank}_{\mathcal{M}}(A \cup \{e\}) = \mathrm{rank}_{\mathcal{M}}(A) = \mathrm{rank}_{\mathcal{M}}(I)$. Hence, I blocks e as well. ■

Then we define what is an optimal solution for a matroid with respect to a cost function μ.

Definition 2.7 *Given a matroid $\mathcal{M}(S, \mathcal{I})$, and an injective cost/weight function $\mu : S \to \mathbb{R}^+$, let $\mu(I) := \sum_{e \in I} \mu(e)$ denote the total weight of elements in the independent set $I \in M$. We say I is an optimal basis (with respect to μ) if $\mu(I)$ has the maximum value among all independent sets in $\mathcal{I}$. We define $\mathsf{OPT}_{\mu}(\mathcal{M}) = \max_{I \in \mathcal{I}} \mu(I)$. With slight abuse of notation, we may also use $\mathsf{OPT}_{\mu}(\mathcal{M})$ to denote the optimal basis. When μ is clear from the context, we simply write $\mathsf{OPT}(\mathcal{M})$.*

From now on, we assume the cost of each element is distinct. It is well known that the optimal basis $\mathsf{OPT}(\mathcal{M})$ is unique (under the distinctness assumption) and can be obtained by a simple greedy algorithm: We first sort the elements in the decreasing order of their cost. Then, we attempt to add the elements greedily one by one in this order, to the current solution, which is initially empty.

We are given matroid $\mathcal{M} = (S, \mathcal{I})$ with cost function $\mu : S \to \mathbb{R}^+$. For a subset $A \subseteq S$, we define

$$A_\mu^{\geq a} := \{e \in A \mid \mu(e) \geq a\}.$$

We define $A_\mu^{>a}, A_\mu^{\leq a}, A_\mu^{<a}$ similarly. Sometimes we omit the subscript μ if it is clear from the context. Finally, the following characterizations of optimal solutions for $\mathcal{M}$ all follow from the greedy procedure.

Lemma 2.8 *For a matroid $\mathcal{M}(S, \mathcal{I})$, cost function $\mu : S \to \mathbb{R}^+$ and basis $I \in \mathcal{I}$, the following statements are equivalent:*

i. I is an optimal basis for $\mathcal{M}$ with respect to μ.
ii. For any $e \in I$, $S^{>\mu(e)}$ does not block e.
iii. For any $e \in S \setminus I$, $I^{\geq \mu(e)}$ blocks e.
iv. For any $r \in \mathbb{R}$, $I^{\geq r}$ is a basis in $\mathcal{M}_{S^{\geq r}}$.

2.2. Uniform Sampling

The following naïve uniform sampling procedure will be used frequently.

Algorithm 1: UniformSample (S, ε, δ)

Data: Arm set S, error bound ε, confidence level δ.
Result: For each arm a, output the empirical mean $\hat{\mu}_a$.
1 For each arm $a \in S$, sample it $\varepsilon^{-2} \ln(2 \cdot \delta^{-1})/2$ times. Let $\hat{\mu}_a$ be the empirical mean.

The following lemma for Algorithm 1, is an immediate consequence of Proposition A.1.1.

Lemma 2.9 *For each arm $a \in S$, we have that* $\Pr[|\mu_a - \hat{\mu}_a| \geq \varepsilon] \leq \delta$.

3. An Optimal PAC Algorithm for the PAC-BASIS Problem

In this section, we prove Theorem 1.7 by presenting an algorithm for PAC-BASIS with optimal sample complexity. The algorithm is also a useful subprocedure for both EXACT-BASIS and PAC-BASIS-AVG.

3.1. Notation

We first introduce an analogue of Lemma 2.8 for ε-optimal solutions.

Lemma 3.1 *For a matroid $\mathcal{M} = (S, \mathcal{I})$ with cost function $\mu : S \to \mathbb{R}^+$, and a basis I, the following statements are equivalent:*

1. *I is ε-optimal for $\mathcal{M}$ with respect to μ.*
2. *For any $e \in S \setminus I$, $I^{\geq \mu(e) - \varepsilon}$ blocks e.*
3. *For any $r \in \mathbb{R}$, let $D_r = (S \setminus I)^{\geq r + \varepsilon} \cup I^{\geq r}$. $I^{\geq r}$ is a basis in $\mathcal{M}_{D_r}$.*

Proof Apply Lemma 2.8 with the cost function $\mu_{I,\varepsilon}$, as defined in Definition 1.5. ■

Definition 3.2 (**ε-Approximation Subset**) *Given a matroid $\mathcal{M} = (S, \mathcal{I})$ and cost function $\mu : S \to \mathbb{R}^+$, let $A \subseteq B$ be two subsets of S. We say A is an ε-approximate subset of B if there exists an independent set $I \in \mathcal{M}_A$ such that I is ε-optimal for $\mathcal{M}_B$ with respect to the cost function μ.*

Lemma 3.3 *Suppose A is an ε-approximate subset of B, and $I \in \mathcal{M}_A$ is ε-optimal for $\mathcal{M}_B$. For any $e \in B \setminus A$, $I^{\geq \mu(e)-\varepsilon}$ blocks e and $A^{\geq \mu(e)-\varepsilon}$ blocks e.*

Proof $I^{\geq \mu(e)-\varepsilon}$ blocks e because Lemma 3.1(2). $I^{\geq \mu(e)-\varepsilon}$ is an independent set of $A^{\geq \mu(e)-\varepsilon}$, so $A^{\geq \mu(e)-\varepsilon}$ blocks e as well. ■

Then we show that "is an ε-approximate subset of" is a transitive relation.

Lemma 3.4 *Let $A \subseteq B \subseteq C$. Suppose A is an ε_1-approximate subset of B, and B is an ε_2-approximate subset of C. Then A is an $(\varepsilon_1 + \varepsilon_2)$-approximate subset of C.*

Proof Let $I \in \mathcal{M}_A$ be ε_1-optimal for $\mathcal{M}_B$. We prove it is $(\varepsilon_1 + \varepsilon_2)$-optimal for $\mathcal{M}_C$. For any element $e \in B \setminus A$, $I^{\geq \mu(e)-\varepsilon_1}$ blocks e. So $I^{\geq \mu(e)-(\varepsilon_1+\varepsilon_2)}$ blocks e as well. For $e \in C \setminus B$, we have $B^{\geq \mu(e)-\varepsilon_2}$ blocks e, by Lemma 3.3. Set $r = \mu(e) - (\varepsilon_1 + \varepsilon_2)$. Using Lemma 3.1(3) with $D_r = (B \setminus I)^{\geq \mu(e)-\varepsilon_2} \cup I^{\geq \mu(e)-(\varepsilon_1+\varepsilon_2)}$, we can see that $I^{\geq \mu(e)-(\varepsilon_1+\varepsilon_2)}$ is a basis in $\mathcal{M}_{D_r}$. Clearly $B^{\geq \mu(e)-\varepsilon_2} \subseteq D_r$. So D_r blocks e, which implies $I^{\geq \mu(e)-(\varepsilon_1+\varepsilon_2)}$ blocks e. Hence, by Lemma 3.1, I is $(\varepsilon_1 + \varepsilon_2)$-optimal for $\mathcal{M}_C$. ■

3.2. Naïve Uniform Sampling Algorithm

We start with a naïve uniform sampling algorithm, which samples each arm enough times to ensure that with high probability the empirical means are all within $\varepsilon/2$ from the true means, and then outputs the optimal solution with respect to the empirical means. The algorithm is a useful procedure in our final algorithm.

Algorithm 2: Naïve-I $(\mathcal{S}, \varepsilon, \delta)$

Data: A PAC-BASIS instance $\mathcal{S} = (S, \mathcal{M})$, with $\mathrm{rank}(\mathcal{M}) = k$, approximation error ε, confidence level δ.

Result: A basis I in $\mathcal{M}$.

```
1 μ̂ ← UniformSample(S, ε/2, δ/|S|)
2 Return The optimal solution I with respect to the empirical means.
```

Lemma 3.5 *The Naïve-I $(\mathcal{S}, \varepsilon, \delta)$ algorithm outputs an ε-optimal solution for $\mathcal{S}$ with probability at least $1 - \delta$. The number of samples is $O(|S|\varepsilon^{-2} \cdot (\ln \delta^{-1} + \ln |S|))$.*

Proof By Lemma 2.9 and a simple union bound, we have $|\mu_e - \hat{\mu}_e| \leq \varepsilon/2$ simultaneously for all arms $e \in S$ with probability $1 - \delta$. Conditioning on that event, let I be the returned basis. For an arm $e \notin I$, we have $I_{\hat{\mu}}^{\geq \hat{\mu}_e}$ blocks e. Note that for all arm $a \in I$, if $\hat{\mu}_a \geq \hat{\mu}_e$, we must have $\mu_a \geq \mu_e - \varepsilon$. Hence, $I_{\hat{\mu}}^{\geq \hat{\mu}_e} \subseteq I_{\mu}^{\geq \mu_e - \varepsilon}$. So $I_{\mu}^{\geq \mu_e - \varepsilon}$ blocks e. Then we have I is ε-optimal by Lemma 3.1. The sample complexity follows from the algorithm statement. ■

3.3. Sampling and Pruning

Our optimal PAC algorithm applies the sampling and pruning technique, initially developed in the celebrated work of Karger, Klein and Tarjan. They used the technique to obtain an expected linear-time algorithm for computing the minimum spanning tree.

We first describe the high level idea from Karger et al. (1995), which will be instructive for our later development. Suppose we want to find the maximum spanning tree (MST). We first construct a subgraph F by sampling each edge with probability p; this subgraph may not be connected, so we solve the maximum-weight spanning forest I of F. The key idea is this: we can use I to prune a lot "useless" edges in the original graph. Formally, an edge $e = (u, v)$ is useless if edges with larger cost in I can connect u and v: this is because the cheapest edge in a cycle does not belong to the MST). (In other words, e is useless if it is blocked by $I^{>\mu(e)}$.) Having removed these useless edges, we again recurse on the remaining graph, which now has much fewer edges, to find the MST. A crucial ingredient of the analysis in Karger et al. (1995) is to show that I can indeed prune a lot of edges.

A proof from Karger et al. (1995); Karger (1998) or (Motwani and Raghavan, 2010, pp. 299-300) shows that an optimal solution from a random subset can help us prune a substantial amount of elements.

Lemma 3.6 *((Karger et al., 1995, Lemma 2.1 and Remark 2.3)) Given a matroid $\mathcal{M} = (S, \mathcal{I})$ with an injective cost function $\mu : S \to \mathbb{R}^+$, sample a subset F of S by selecting each element independently with probability p. An element $e \in S$ is called F-*good *if $F^{>\mu(e)}$ does not block e, else it is F-*bad*. If the r.v. X denotes the number of F-good elements in S, then X is stochastically dominated by* $\mathsf{NegBin}(\mathrm{rank}(\mathcal{M}); p)$.

We also introduce a lemma which shows an ε-optimal solution I in F can be used to eliminate some sub-optimal arms.

Lemma 3.7 *For a matroid $\mathcal{M} = (S, \mathcal{I})$ with cost function $\mu : S \to \mathbb{R}^+$, Let $F \subseteq S$ be a subset, and I be an α-optimal basis for $\mathcal{M}_F$ for some $\alpha > 0$. If an element $e \in S \setminus I$ is F-bad, $I^{\geq\mu(e)-\alpha}$ blocks e.*

Proof As e is F-bad, $F^{\geq\mu(e)}$ blocks e. Let $r = \mu(e) - \alpha$, and

$$D = (F \setminus I)^{\geq r+\alpha} \cup I^{\geq r} = (F \setminus I)^{\geq\mu(e)} \cup I^{\geq\mu(e)-\alpha}.$$

(In other words, we first add α to the cost of every element in I, then consider all element with cost at least $\mu(e)$ in F). Then by Lemma 2.8(4) and the fact I is α-optimal for $\mathcal{M}_F$, $I^{\geq\mu(e)-\alpha}$ is maximal for $\mathcal{M}_D$ (in fact, it is optimal for $\mathcal{M}_D$ w.r.t. the modified cost function). Clearly $F^{\geq\mu(e)} \subseteq D$, so D blocks e as well. Hence $I^{\geq\mu(e)-\alpha}$ also blocks e, by Lemma 2.6. ■

3.4. Our Optimal PAC Algorithm

Now, we present our algorithm for the PAC case, which is based on the sampling-and-pruning technique discussed above. Let $p = 0.01$. The algorithm runs as follows: If the number of arms $|S|$ is sufficiently small, we simply run the naïve uniform sampling algorithm. Otherwise, we sample

Algorithm 3: PAC-SamplePrune $(\mathcal{S}, \varepsilon, \delta)$

Data: A PAC-BASIS instance $\mathcal{S} = (S, \mathcal{M})$, with $\operatorname{rank}(\mathcal{M}) = k$, approximation error ε, confidence level δ.

Result: A basis I in $\mathcal{M}$.

```
1 if |S| ≤ 2p^{-2} · max(4 · ln 8δ^{-1}, k) then
2     Return Naïve-I (S, ε, δ)
3
4 Sample a subset F ⊆ S by choosing each element with probability p independently.
5 α ← ε/3, λ ← ε/12
6 I ← PAC-SamplePrune(S_F = (F, M_F), α, δ/8)
7 μ̂ ← UniformSample (S, λ, δ · p/8k)
8 S' ← I ∪ {e ∈ S \ I | I_{μ̂}^{≥ μ̂_e − α − 2λ} does not block e}
9 Return PAC-SamplePrune (S_{S'} = (S', M_{S'}), α, δ/4)
```

a subset F of S by selecting each arm with probability p independently, and recurse on the sub-instance $\mathcal{S}_F = (F, \mathcal{M}_F)$ to find an α-optimal solution I, where $\alpha = \varepsilon/3$. Next, we uniformly sample each arm in S by calling UniformSample $(S, \lambda, \delta \cdot p/8k)$, where $\lambda = \varepsilon/12$. Then, we use I to eliminate those sub-optimal arms in $S \setminus I$. More precisely, a sub-optimal arm e is blocked by the arms of I with empirical values larger than $\hat{\mu}_e - \alpha - 2\lambda$. Finally, we invoke the algorithm recursively on the remaining arms to find an α-optimal solution, which we output as the final result. The pseudo-code can be found in Algorithm 3.

Note that UniformSample (step 6) is the only step in which we take samples from the arms. Also note that in both recursive calls we set the approximation error to be $\alpha = \varepsilon/3$. Effectively, this makes sure that an arms surviving in deeper recursive call are sampled more times. This feature is shared by other elimination-based method, such as Even-Dar et al. (2002); Zhou et al. (2014). However, the way we choose which arms should be eliminated is quite different.

3.5. Analysis of the sample complexity

In this subsection, we analyze PAC-SamplePrune and prove Theorem 1.7.

Theorem 1.7 (rephrased) *Given a* PAC-BASIS *instance $\mathcal{S} = (S, \mathcal{M})$, Algorithm PAC-SamplePrune $(\mathcal{S}, \varepsilon, \delta)$ returns an ε-optimal solution, with probability at least $1 - \delta$, and uses at most*

$$O(n\varepsilon^{-2} \cdot (\ln k + \ln \delta^{-1}))$$

samples. Here $k = \operatorname{rank}(\mathcal{M})$, and $n = |S|$.

Let c_1, c_2 be two constants to be specified later. We will prove by induction on $|S|$ that with probability at least $1 - \delta$, PAC-SamplePrune $(\mathcal{S} = (S, \mathcal{M}), \varepsilon, \delta)$ returns an ε-optimal solution, using at most $c_1 \cdot (|S|\varepsilon^{-2} \cdot (\ln k + \ln \delta^{-1} + c_2))$ samples. Remember that $p = 0.01$.

We first consider the simple case where $|S|$ is not much larger than k. When $|S| \le 2p^{-2} \cdot \max(4 \cdot \ln 8\delta^{-1}, k)$, we have that $\ln |S| = O(\ln \delta^{-1} + \ln k)$. So the number of samples of Naïve-I is $O(|S|\varepsilon^{-2} \cdot (\ln |S| + \ln \delta^{-1})) = O(|S|\varepsilon^{-2} \cdot (\ln k + \ln \delta^{-1}))$; by Lemma 3.5, the returned basis is ε-optimal with probability at least $1 - \delta$. Hence the theorem holds in this case.

Now consider the case where $|S| > 2p^{-2} \cdot \max(4 \cdot \ln 8\delta^{-1}, k)$, and inductively assume that the theorem is true for all instances of size smaller than $|S|$. We first need the following lemma, which describes the good events that happen with high probability.

Lemma 3.8 *Let O be the unique optimal solution for $\mathcal{S} = (S, \mathcal{M})$. With probability at least $1-\delta/2$, the following statements hold simultaneously.*

1. $|F| \le 2p \cdot |S|$ *(F is obtained in Line 3).*
2. *There are at most $p \cdot |S|$ F-good elements in S.*
3. $|\mu_e - \hat{\mu}_e| \le \lambda$*, for all elements $e \in O \cup I$ (I is obtained in Line 5).*
4. *I is an α-optimal solution for F.*

Proof Let $n = |S|$. By Corollary A.2, we have that

$$\Pr[|F| > 2pn] = \Pr[\mathsf{Bin}(n,p) > 2pn] \le e^{-pn/3} \le \delta/8,$$

for $n \ge 8p^{-2}(\ln 8\delta^{-1})$. Moreover, let X be the r.v. denoting the number of F-good elements in S. By Lemma 3.6, X is dominated by $\mathsf{NegBin}(k;p)$, and hence

$$\Pr[X > pn] \le \Pr[\mathsf{NegBin}(k;p) > pn] = \Pr[\mathsf{Bin}(pn,p) < k] \le \Pr[\mathsf{Bin}(pn,p) < \frac{1}{2}p^2 n] \le e^{-\frac{1}{8}p^2 n} \le \delta/8.$$

The second inequality holds since $p^2 n \ge 2k$, while the last inequality is due to $\frac{1}{8}p^2 n \ge \ln \delta^{-1} + \ln 8$. In addition, by Lemma 2.9 and a trivial union bound over all arms in $O \cup I$, the third statement holds with probability at least $1 - (p \cdot \delta/8k) \cdot (2k) \ge 1 - \delta/8$.

Finally, conditioning on the first statement, we have $|F| < |S|$, and hence by the induction hypothesis, with probability at least $1-\delta/8$, I is an α-optimal solution for F. Putting them together, all four statements hold with probability at least $1 - \delta/8 \cdot 4 = 1 - \delta/2$. ∎

Now let $\mathcal{E}$ denote the event that all statements in Lemma 3.8 hold. We show each F-bad element in $S \setminus I$ has a constant probability to be eliminated.

Lemma 3.9 *Conditioning on $\mathcal{E}$, for an F-bad element $e \in S \setminus I$, $\Pr[e \in S'] \le \delta \cdot p/8k$.*

Proof Conditioning on $\mathcal{E}$, I is α-optimal for F. Hence, by Lemma 3.7, for an F-bad element $e \in S \setminus I$, $I_{\mu}^{\ge \mu_e - \alpha}$ blocks e. By Lemma 3.5, $|\hat{\mu}_e - \mu_e| \le \lambda$ with probability $1 - p \cdot \delta/8k$. Moreover, conditioning on $\mathcal{E}$, we have $|\hat{\mu}_a - \mu_a| \le \lambda$ for every element $a \in I$ (by Lemma 3.8.3). Consequently, $I_{\mu}^{\ge \mu_e - \alpha} \subseteq I_{\hat{\mu}}^{\ge \hat{\mu}_e - \alpha - 2\lambda}$, which implies that $I_{\hat{\mu}}^{\ge \hat{\mu}_e - \alpha - 2\lambda}$ blocks e. By the definition of S', this means $e \notin S'$. Hence, we have $\Pr[e \in S' \mid \mathcal{E}] \le \delta \cdot p/8k$. ∎

Now, we show that with high probability, S' is a $2\varepsilon/3$-approximate subset of S, and the size of S' is much smaller than $|S|$.

Lemma 3.10 *Conditioning on $\mathcal{E}$, $|S'| \le 2p|S|$, and S' is an $(\alpha + 4\lambda)$-approximate subset of S, with probability $1-\delta/4$.*

Proof Conditioned on event $\mathcal{E}$, there are at most $p \cdot |S|$ F-good elements in S (by Lemma 3.8.2). If X denotes the number of F-bad elements in $S \setminus I$ which remain in S', Lemma 3.9 implies $\mathbb{E}[X] \leq \delta \cdot (p/8k) \cdot |S \setminus I| \leq \delta \cdot p/8 \cdot |S|$. By Markov's inequality, we have $\Pr[X \geq 0.5p|S|] \leq \Pr[X \geq 4 \cdot \delta^{-1}\mathbb{E}[X]] \leq \delta/4$. So there are at most $|I| + 1.5p \cdot |S| \leq k + 1.5p|S| \leq 2p|S|$ elements in S' with probability at least $1 - \delta/4$.

For the second part, observe that $O \cup S' \subseteq S$ is a 0-approximate subset of S, so by Lemma 3.4, it suffices to show S' is an $(\alpha + 4\lambda)$-approximate subset for $O \cup S'$. Still conditioned on event $\mathcal{E}$, for all arms $e \in I \cup O$, we have $|\mu_e - \hat{\mu}_e| \leq \lambda$. So for an arm $e \in O \setminus S'$, we have $I_{\hat{\mu}}^{\geq \hat{\mu}_e - \alpha - 2\lambda}$ blocks e (otherwise, e should be included in S'), which implies $I_{\mu}^{\geq \mu_e - \alpha - 4\lambda}$ blocks e. Since $I \subseteq S'$, we can see S' is an $(\alpha + 4\lambda)$-approximate subset of $S' \cup O$ by Definition 3.2. ■

Finally, we are ready to prove Theorem 1.7.

Proof [Proof of Theorem 1.7] Let $\mathcal{E}_G$ be the intersection of the event $\mathcal{E}$, the event that Lemma 3.10 holds and the event that PAC-SamplePrune (line 8) outputs correctly. Conditioning on event $\mathcal{E}$, $|S'| < |S|$, so by the induction hypothesis, the last event happens with probability at least $1 - \delta/4$. Hence, $\Pr[\mathcal{E}_G] \geq 1 - \delta/2 - \delta/4 - \delta/4 = 1 - \delta$. We condition our following argument on $\mathcal{E}_G$.

First we show the algorithm is correct. By Lemma 3.10, S' is an $(\alpha + 4\lambda)$-approximate subset of S, and the returned basis J is an α-optimal solution of S', hence also an α-approximate subset of S'. By the "transitivity" property of Lemma 3.4, J is an $(\alpha + \alpha + 4\lambda)$-approximate subset of S. This is an ε-optimal solution of S since $\alpha + \alpha + 4\lambda = \varepsilon$.

By Lemma 3.8 and Lemma 4.1, we have $|F| \leq 2p \cdot |S|$ and $|S'| \leq 2p \cdot |S|$. By the induction hypothesis, the total number of samples in both recursive calls (line 6 and line 8) can be bounded by

$$c_1 \cdot 4p|S|\varepsilon^{-2}(\ln \delta^{-1} + \ln k + c_2) \cdot 9 \leq 36c_1 p \cdot |S|\varepsilon^{-2}(\ln \delta^{-1} + \ln k + c_2).$$

The number of samples incurred by UniformSample (line 7) can be bounded by

$$|S|\lambda^{-2} \cdot (\ln \delta^{-1} + \ln 16 + \ln p^{-1} + \ln k)/2 \leq 72|S|\varepsilon^{-2} \cdot (\ln \delta^{-1} + \ln 16 + \ln p^{-1} + \ln k).$$

Now, let $c_2 = \ln 16 + \ln p^{-1}$, which is a constant. Then the total number of samples is bounded by

$$(36p \cdot c_1 + 72)|S|\varepsilon^{-2} \cdot (\ln \delta^{-1} + \ln k + c_2).$$

Setting $c_1 = \max(120, c_0)$, and plugging in $p = 0.01$, we can see the above quantity is bounded by $c_1 \cdot |S|\varepsilon^{-2} \cdot (\ln \delta^{-1} + \ln k + c_2)$, which completes the proof. ■

4. An Algorithm for the Exact-Basis Problem

We now turn to the Exact-Basis problem, and prove Theorem 1.4. If we denote the unique optimal basis by OPT, and let BAD be the set of all other arms in $S \setminus$ OPT, our goal for the Exact-Basis problem is to find this set OPT with confidence $1 - \delta$ using as few samples as possible.

Our algorithm Exact-ExpGap is based on our previous PAC result for PAC-Basis, and also borrow some idea from the Exponential-Gap-Eliminating algorithm by Karnin et al. (2013). It will run in rounds. In each round, it either tries to eliminate some arms in BAD (we call such a round an *elimination-round*), or adds some arms from OPT into our solution and removes them from further consideration (we call such a round a *selection-round*). Let us give some details about these two kinds of rounds. Let $\mathcal{M}_{\mathrm{cur}}$ be the current matroid defined over the remaining arms, n_{opt} be the number of remaining arms in OPT, and n_{bad} be the number of remaining arms in BAD.

1. (elimination-round) When $n_{\mathrm{opt}} \le n_{\mathrm{bad}}$, we are in an elimination-round. In the r^{th} elimination-round, first we find an ε_r-optimal solution I for the current matroid $\mathcal{M}_{\mathrm{cur}}$ by calling PAC-SamplePrune $(\mathcal{M}_{\mathrm{cur}}, \varepsilon_r, \delta_r)$ (i.e., the PAC algorithm from Section 3) with $\varepsilon_r = 2^{-r}/4$ and $\delta_r = \delta/100r^3$. We sample each arm in I by calling UniformSample$(I, \varepsilon_r/2, \delta_r/n_{\mathrm{opt}})$ to estimate their means. We do the same for arms $S_{\mathrm{cur}} \setminus I$ by calling UniformSample$(S_{\mathrm{cur}} \setminus I, \varepsilon_r, \delta_r/n_{\mathrm{opt}})$. Note the confidence parameter is not low enough to give accurate estimations for all arms in $S_{\mathrm{cur}} \setminus I$ with high probability: that would require us reducing the parameter to $\delta_r/|S_{\mathrm{cur}} \setminus I|$. However, we will be satisfied with being accurate only for arms in OPT $\setminus I$ with probability δ_r.

   Finally, we use I to eliminate some sub-optimal arms. In particular, an arm e should be eliminated if $I_{\hat{\mu}}^{\ge \hat{\mu}_e + 1.5\varepsilon_r}$ blocks e, where $\hat{\mu}$ is the cost function defined by the empirical means obtained from the above UniformSample procedures.

2. (selection-round) When $n_{\mathrm{bad}} < n_{\mathrm{opt}}$, we are in a selection-round. In the r^{th} selection-round, we sample all the arms in $\mathcal{M}_{\mathrm{cur}}$ by calling UniformSample$(S_{\mathrm{cur}}, \varepsilon_r, \delta_r/|S_{\mathrm{cur}}|)$. We then select into our solution Ans those elements e which are not blocked by all other elements in $\mathcal{M}_{\mathrm{cur}}$ with larger empirical means, even if we slightly decrease e's empirical mean by $2\varepsilon_r$. Having contracted these selected arms, we proceed to the next round.

Finally, the algorithm terminates when either $n_{\mathrm{opt}} = 0$ or $n_{\mathrm{bad}} = 0$. The pseudo-code is given as Algorithm 4.

4.1. Analysis of the algorithm

Now, we prove the main theorem of this section by analyzing the correctness and sample complexity of Exact-ExpGap.

Theorem 1.4 (rephrased) Given an EXACT-BASIS instance $\mathcal{S}(S, \mathcal{M})$, Exact-ExpGap$(\mathcal{S}, \varepsilon, \delta)$ returns the optimal basis of $\mathcal{M}$, with probability at least $1 - \delta$, and uses at most

$$O\left(\sum_{e \in S} \Delta_e^{-2}(\ln \delta^{-1} + \ln k + \ln \ln \Delta_e^{-1})\right)$$

samples. Here, $k = \mathrm{rank}(\mathcal{M})$ is the size of a basis of $\mathcal{M}$.

We first recall that the gap of an element e (throughout this section, we only consider the cost function μ for gap) is defined to be

$$\Delta_e^{\mathcal{M}} := \begin{cases} \mathsf{OPT}(\mathcal{M}) - \mathsf{OPT}(\mathcal{M}_{S \setminus \{e\}}), & e \in \mathsf{OPT} \text{ and } e \text{ is not isolated;} \\ +\infty, & e \text{ is isolated;} \\ \mathsf{OPT}(\mathcal{M}) - \mathsf{OPT}(\mathcal{M}_{/\{e\}}) - \mu_e, & e \notin \mathsf{OPT}. \end{cases}$$

Algorithm 4: Exact-ExpGap $(\mathcal{S}, \varepsilon, \delta)$

Data: An EXACT-BASIS instance $\mathcal{S} = (S, \mathcal{M})$, with $\mathrm{rank}(\mathcal{M}) = k$, approx. error ε, confidence level δ.

Result: A basis I in $\mathcal{M}$.

1 $r_{\text{elim}} \leftarrow 1$, $r_{\text{sele}} \leftarrow 1$
2 **while** *True* **do**
3     $S_{\text{cur}} \leftarrow$ the arm set of $\mathcal{M}_{\text{cur}}$
4     $n_{\text{opt}} \leftarrow \mathrm{rank}(\mathcal{M}_{\text{cur}})$, $n_{\text{bad}} \leftarrow |S_{\text{cur}}| - n_{\text{opt}}$
5   **if** $n_{\text{opt}} \leq n_{\text{bad}}$ **then**
6   **if** $n_{\text{opt}} = 0$ **then break**
7         $r \leftarrow r_{\text{elim}}$
8         $\varepsilon_r \leftarrow 2^{-r}/4$, $\delta_r \leftarrow \delta/100r^3$, $r_{\text{elim}} \leftarrow r_{\text{elim}} + 1$
9         $I \leftarrow$ PAC-SamplePrune$(\mathcal{S}_{\text{cur}} = (S_{\text{cur}}, \mathcal{M}_{\text{cur}}), \varepsilon_r, \delta_r)$
10         $\hat{\mu} \leftarrow$ UniformSample$(I, \varepsilon_r/2, \delta_r/n_{\text{opt}})$
11         $\hat{\mu} \leftarrow$ UniformSample$(S_{\text{cur}} \setminus I, \varepsilon_r, \delta_r/n_{\text{opt}})$
12         $S_{\text{new}} \leftarrow I \cup \{e \in S_{\text{cur}} \setminus I \mid I_{\hat{\mu}}^{\geq \hat{\mu}_e + 1.5\varepsilon_r}$ does not block e in $\mathcal{M}_{\text{cur}}\}$
13         $\mathcal{M}_{\text{cur}} \leftarrow \mathcal{M}_{\text{cur}} | S_{\text{new}}$
14   **else**
15   **if** $n_{\text{bad}} = 0$ **then**
16   Ans $\leftarrow$ Ans $\cup\, S_{\text{cur}}$
17   **break**
18         $r \leftarrow r_{\text{sele}}$
19         $\varepsilon_r \leftarrow 2^{-r}/4$, $\delta_r \leftarrow \delta/100r^3$, $r_{\text{sele}} \leftarrow r_{\text{sele}} + 1$
20         $\hat{\mu} \leftarrow$ UniformSample$(S_{\text{cur}}, \varepsilon_r, \delta_r/|S_{\text{cur}}|)$
21         $U \leftarrow \{e \in S_{\text{cur}} \mid (S_{\text{cur}} \setminus \{e\})_{\hat{\mu}}^{\geq \hat{\mu}_e - 2\varepsilon_r}$ does not block e in $\mathcal{M}_{\text{cur}}\}$
22   Ans $\leftarrow$ Ans $\cup\, U$
23         $\mathcal{M}_{\text{cur}} \leftarrow \mathcal{M}_{\text{cur}}/U$
24
25 **Return** Ans

Note that we extend the definition to the isolated elements, since the restrictions and contractions may result in such element (note that no loop is introduced during the process). We also need the following equivalent definition (the equivalence follows from Lemma 2.8), which may be convenient in some case:

$$\Delta_e^{\mathcal{M}} := \begin{cases} \max\{w \in R \mid (S \setminus \{e\})^{>\mu_e - w} \text{ does not block } e\} & \text{for } e \in \mathsf{OPT}; \\ \max\{w \in R \mid S^{\geq \mu_e + w} \text{ blocks } e\} & \text{for } e \notin \mathsf{OPT} \end{cases} \tag{1}$$

First, we prove that our algorithm returns the optimal basis with high probability. In the following lemma, We specify a few events on which we condition our later discussion.

Lemma 4.1 *With probability at least $1 - \delta/5$, all of the following statements hold:*

1. *In all elimination-rounds,* PAC-SamplePrune *(line 9) returns correctly.*
2. *In all elimination-rounds, for all element $u \in I$, $|\mu_u - \hat{\mu}_u| < \varepsilon_r/2$.*

3. *In all elimination-rounds, for all element* $u \in \mathsf{OPT}(\mathcal{M}_{\text{cur}})$, $|\mu_u - \hat{\mu}_u| < \varepsilon_r$.
4. *In all selection-rounds, for all element* $u \in S_{\text{cur}}$, $|\mu_u - \hat{\mu}_u| < \varepsilon_r$.

We use $\mathcal{E}$ *to denote the event that all above statements are true.*

Proof In the r^{th} elimination-round, the specification of the failure probabilities of PAC-SamplePrune and UniformSample imply the first three statements hold with probability $1 - 3\delta_r$. In the r^{th} selection-round, the last statement holds with probability $1 - \delta_r$. A trivial union bound over all rounds gives

$$\Pr[\neg\mathcal{E}] \le \sum_{r=1}^{+\infty}(3\delta_r + \delta_r) = \sum_{r=1}^{+\infty} 4\delta/100r^3 \le \delta/5,$$

and the lemma follows immediately. ■

Lemma 4.2 *Conditioning on* $\mathcal{E}$, *the subset* Ans *returned by the algorithm is the optimal basis* OPT.

Proof We condition on $\mathcal{E}$ in the following discussion. We show that the algorithm only deletes arms from BAD in every elimination-round, and it only selects arms from OPT in every selection-round. We say a round is correct if it satisfies the above requirements. We prove all rounds are correct by induction. Consider a round, and suppose all previous rounds are correct. Hence, at the beginning of the current round, Ans clearly is a subset of OPT, and $\mathsf{OPT}(\mathcal{M}_{\text{cur}}) = \mathsf{OPT} \setminus \mathsf{Ans}$. There are two cases:

If the current round is an elimination-round, consider an arm $u \in \mathsf{OPT}(\mathcal{M}_{\text{cur}}) = \mathsf{OPT} \setminus \mathsf{Ans}$. We can see that $I_{\mu}^{>\mu_u}$ does not block u in $\mathcal{M}_{\text{cur}}$, by the characterization of the matroid optimal solutions in Lemma 2.8.2. Since $|\hat{\mu}_u - \mu_u| < \varepsilon_r$ for all $u \in \mathsf{OPT}(\mathcal{M}_{\text{cur}})$, and $|\hat{\mu}_e - \mu_e| < \varepsilon_r/2$ for all $e \in I$, we also have $I_{\hat{\mu}}^{\ge \hat{\mu}_u + 1.5\varepsilon_r}$ does not block u in $\mathcal{M}_{\text{cur}}$. Hence $u \in S_{new}$, and it is not eliminated.

Next, consider a selection-round and an arm $u \in \mathsf{BAD} \cap S_{\text{cur}}$. By the induction hypothesis, in the beginning of the round, $u \notin \mathsf{OPT}(\mathcal{M}_{\text{cur}})$. Then, we can see u is blocked by $(S_{\text{cur}} \setminus \{u\})_{\mu}^{\ge \mu_u}$ in $\mathcal{M}_{\text{cur}}$ by Lemma 2.8.3. Again, for all arms e in S_{cur}, $|\mu_e - \hat{\mu}_e| < \varepsilon_r$, so u is also blocked by $(S_{\text{cur}} \setminus \{u\})_{\hat{\mu}}^{\ge \hat{\mu}_u - 2\varepsilon_r}$ in $\mathcal{M}_{\text{cur}}$. Hence $u \notin U$, and is not selected into Ans.

Finally, if the algorithms returns, we have $|\mathsf{Ans}| = |\mathsf{OPT}| = \mathrm{rank}(\mathcal{M})$. Since $\mathsf{Ans} \subseteq \mathsf{OPT}$, it must be the case that $\mathsf{Ans} = \mathsf{OPT}$. ■

4.1.1. ANALYSIS OF SAMPLE COMPLEXITY

To analyze the sample complexity, we need some additional notation. Let n_{opt}^r (resp. n_{bad}^r) denote n_{opt} (resp. n_{bad}) at the beginning of r^{th} elimination-round (resp. selection-round). Also, let S_{elim}^r denote the arm set of $\mathcal{M}_{\text{cur}}$ at the end of the r^{th} elimination-round, and S_{sele}^r denote the arm set of $\mathcal{M}_{\text{cur}}$ at the end of the r^{th} selection-round. We partition the arms in OPT and BAD based on their gaps, as follows:

$$\mathsf{OPT}_s = \{u \in \mathsf{OPT} \mid 2^{-s} \le \Delta_u < 2^{-s+1}\},$$

$$\mathsf{BAD}_s = \{u \in \mathsf{BAD} \mid 2^{-s} \le \Delta_u < 2^{-s+1}\}.$$

Moreover, we define $\mathsf{OPT}_{r,s} := S_{\text{sele}}^r \cap \mathsf{OPT}_s$, i.e., the set of arms in OPT_s not selected in the r^{th} selection-round—recall that in a selection-round we aim to select those arms into OPT. Similarly, define $\mathsf{BAD}_{r,s} := S^r_{\text{elim}} \cap \mathsf{BAD}_s$ as the set of arms in BAD_s not eliminated the r^{th} elimination-round—again, in an elimination-round we aim to delete those arms in BAD.

Very roughly speaking, the s^{th} round is dedicated to deal with those arms with gap roughly $O(2^{-s})$ (namely, an arm in OPT_s is likely to be selected in the s^{th} selection-round and an arm in BAD_s is likely to be eliminated in the s^{th} elimination-round). Now, we prove a crucial lemma, which states that all elements in OPT_s should be selected in or before the s^{th} selection-round, and the number of remaining elements in BAD_s should drop exponentially after the s^{th} elimination-round.

Lemma 4.3 *Conditioning on event $\mathcal{E}$, with probability at least $1 - 4\delta/5$, we have*

$$|\mathsf{OPT}_{r,s}| = 0 \quad \text{and} \quad |\mathsf{BAD}_{r,s}| \le \frac{1}{8}|\mathsf{BAD}_{r-1,s}| \quad \text{for all } 1 \le s \le r.$$

Proving Lemma 4.3 requires some preparations. All the following arguments are conditioned on event $\mathcal{E}$. We first prove a useful lemma which roughly states that if we select some elements in OPT and remove some elements in BAD, the gap of the remaining instance does not decrease.

Lemma 4.4 *For two subsets A, B of S such that $A \subseteq \mathsf{OPT} \subseteq B$, consider the matroid $\mathcal{M}' = (\mathcal{M}|B)/A$. Let S' be its ground set. For all element $u \in S'$, we have that $\Delta_u^{\mathcal{M}'} \ge \Delta_u^{\mathcal{M}}$.*

Proof By the definition of matroid contraction and the fact that A is independent, we can see for any subset $U \subseteq S'$, U is independent in $\mathcal{M}'$ iff $U \cup A$ is independent in $\mathcal{M}$. We also have $\text{rank}_{\mathcal{M}'}(S') = \text{rank}_{\mathcal{M}}(S) - |A|$ and $\mathsf{OPT} \setminus A$ is the unique optimal solution for $\mathcal{M}'$.

Now, let $u \in S'$. Suppose $u \in \mathsf{OPT}(\mathcal{M}') = \mathsf{OPT} \setminus A$. Suppose for contradiction that $\Delta_u^{\mathcal{M}'} < \Delta_u^{\mathcal{M}}$. Then we have a basis I contained in $S' \setminus \{u\}$ in $\mathcal{M}'$ such that

$$\mu(I) = \mu(\mathsf{OPT} \setminus A) - \Delta_u^{\mathcal{M}'} > \mu(\mathsf{OPT} \setminus A) - \Delta_u^{\mathcal{M}}.$$

But this means $I \cup A$ is a basis contained in $S \setminus \{u\}$ in $\mathcal{M}$ such that $\mu(I \cup A) > \mathsf{OPT}(\mathcal{M}) - \Delta_u^{\mathcal{M}}$, contradicting to the definition of $\Delta_u^{\mathcal{M}}$. Note that a non-isolated element in $\mathcal{M}$ may become isolated in $\mathcal{M}'$ (for which $\Delta_u^{\mathcal{M}'} = +\infty$).

Then, we consider the case $u \notin \mathsf{OPT}(\mathcal{M}') = \mathsf{OPT} \setminus A$. The argument is quite similar. Suppose for contradiction that $\Delta_u^{\mathcal{M}'} < \Delta_u^{\mathcal{M}}$. This means that there exists a basis I in $\mathcal{M}'$ such that $u \in I$ and $\mu(I) > \mathsf{OPT}(\mathcal{M}') - \Delta_u^{\mathcal{M}} = \mu(\mathsf{OPT} \setminus A) - \Delta_u^{\mathcal{M}}$. But this means $A \cup I$ is a basis in $\mathcal{M}$ such that $\mu(A \cup I) > \mathsf{OPT} - \Delta_u^{\mathcal{M}}$. Since $u \in (A \cup I)$, this contradicts the definition of $\Delta_u^{\mathcal{M}}$. ■

Proof [Proof of Lemma 4.3] We first prove $|\mathsf{OPT}_{r,s}| = 0$ for $r \ge s$. Suppose we are at the beginning of the r^{th} selection-round. Let $A = \mathsf{Ans}$ and $B = S_{\text{cur}} \cup \mathsf{Ans}$. We can see $A \subseteq OPT \subseteq B$ and $\mathcal{M}_{\text{cur}} = (\mathcal{M}|B)_{/A}$.

For any arm $u \in \mathsf{OPT}_{r-1,s}$ such that $s \le r$, we have $\Delta_u \ge 2^{-s} \ge 2^{-r} \ge 4\varepsilon_r$. By Lemma 4.4, we have $\Delta_u^{\mathcal{M}_{\text{cur}}} \ge \Delta_u^{\mathcal{M}} \ge 4\varepsilon_r$, which means $(S_{\text{cur}} \setminus \{u\})_{\mu}^{>\mu_u - 4\varepsilon_r}$ does not block u. Note that conditioning on $\mathcal{E}$, $|\mu_e - \hat{\mu}_e| < \varepsilon_r$ for all $e \in S_{\text{cur}}$. This implies that $(S_{\text{cur}} \setminus \{u\})_{\hat{\mu}}^{\ge \hat{\mu}_u - 2\varepsilon_r}$ does not block u as well. So $u \in U$ (U is defined in line 21) and consequently $|\mathsf{OPT}_{r,s}| = 0$.

Now, we prove the second part of the lemma. We claim that for $1 \le s \le r$, we have that

$$\Pr[|\mathsf{BAD}_{r,s}| \le \frac{1}{8}|\mathsf{BAD}_{r-1,s}|] \ge 1 - 8\delta_r \tag{2}$$

The lemma follows directly from the claim by taking a union bound over all s, r such that $1 \le s \le r$:

$$\sum_{r=1}^{+\infty}\sum_{s=1}^{r} 8\delta_r = \sum_{r=1}^{+\infty} 8\delta/100r^2 \le 4\delta/5.$$

What remains to prove is the claim (Inequality (2)). Suppose we are at the beginning of the r^{th} elimination-round. Let $A = \mathsf{Ans}$ and $B = S_{\text{cur}} \cup \mathsf{Ans}$. Conditioning on event $\mathcal{E}$, we can see that $A \subseteq \mathsf{OPT} \subseteq B$ and $\mathcal{M}_{\text{cur}} = (\mathcal{M}|B)/A$.

For any arm $u \in \mathsf{BAD}_{r-1,s}$ such that $s \le r$, we have $\Delta_u \ge 2^{-s} \ge 2^{-r} \ge 4\varepsilon_r$. By Lemma 4.4, we have $\Delta_u^{\mathcal{M}_{\text{cur}}} \ge \Delta_u^{\mathcal{M}} \ge 4\varepsilon_r$. So $(S_{\text{cur}})_{\mu}^{\ge \mu_u + 4\varepsilon_r}$ blocks u in $\mathcal{M}_{\text{cur}}$ by the definition of $\Delta_u^{\mathcal{M}_{\text{cur}}}$. As I is ε_r-optimal for $\mathcal{M}_{\text{cur}}$, we also have u is blocked by $I_{\mu}^{\ge \mu_u + 3\varepsilon_r}$ in $\mathcal{M}_{\text{cur}}$. This implies that $u \notin I$.

Since we have $|\mu_u - \hat{\mu}_u| < \varepsilon_r$ with probability $1 - \delta_r$, combining with the fact that $|\mu_e - \hat{\mu}_e| < \varepsilon_r/2$ for all $e \in I$ (guaranteed by $\mathcal{E}$), u is blocked by $I_{\hat{\mu}}^{>\hat{\mu}_u + 1.5\varepsilon_r}$ with probability $1 - \delta_r$. This implies that $u \notin \mathsf{BAD}_{r,s}$ (u should be eliminated in line 12). From the above, we can see that

$$\mathbb{E}[|\mathsf{BAD}_{r,s}|] \le \delta_r |\mathsf{BAD}_{r-1,s}|.$$

By Markov inequality, we have $\Pr[|\mathsf{BAD}_{r,s}| \ge \frac{1}{8}|\mathsf{BAD}_{r-1,s}|] \le 8\delta_r$, which concludes the proof. ■

Finally, everything is in place to prove Theorem 1.4.

Proof [Proof of Theorem 1.4] Let $\mathcal{E}_G$ be the intersection of event $\mathcal{E}$ and the event that Lemma 4.3 holds. By Lemma 4.1 and Lemma 4.3, $\Pr[\mathcal{E}_G] \ge 1 - \delta$. Now we condition on this event.

The correctness has been proved in Lemma 4.2. We only need to bound the sample complexity of Exact-ExpGap.

We first consider the number samples taken by the UniformSample procedure. We handle the samples taken by PAC-SamplePrune later. Now, we bound the total number of samples taken from arms in OPT_s in all selection-round s. Notice that we can safely ignore all samples on arms in BAD since $n_{\text{opt}} \ge n_{\text{bad}}$. By Lemma 4.3, $|\mathsf{OPT}_{r,s}| = 0$ for $r \ge s$. So it can be bounded as:

$$\begin{aligned} O\left(\sum_{r=1}^{s} |\mathsf{OPT}_{r-1,s}| \cdot (\ln n_{\text{opt}} + \ln \delta_r^{-1})\varepsilon_r^{-2}\right) &\le O\left(\sum_{r=1}^{s} |\mathsf{OPT}_s| \cdot (\ln k + \ln \delta^{-1} + \ln r)\varepsilon_r^{-2}\right) \\ &\le O\left(|\mathsf{OPT}_s| \cdot (\ln k + \ln \delta^{-1} + \ln s) \cdot 4^s\right). \end{aligned}$$

Next, we consider the number of samples from elimination-round s. In an elimination-round, since $n_{\text{opt}} \le n_{\text{bad}}$, we only need to bound the number of samples from BAD. The total number of samples taken from arms in BAD_s in the first s elimination-rounds can be bounded as:

$$O\left(\sum_{r=1}^{s} |\mathsf{BAD}_{r-1,s}| \cdot (\ln |S_{\text{cur}}| + \ln \delta_r^{-1})\varepsilon_r^{-2}\right) \le O\left(|\mathsf{BAD}_s| \cdot (\ln k + \ln \delta^{-1} + \ln s) \cdot 4^s\right)$$

The inequality holds since $|S_{\text{cur}}| \le n_{\text{opt}} + n_{\text{bad}} \le 2n_{\text{opt}} \le 2k$ in a selection-round. Now, we bound the number of samples from the remaining rounds. Since $|\mathsf{BAD}_{r,s}| \le \frac{1}{8}|\mathsf{BAD}_{r-1,s}|$ when

$r \geq s$, we have:

$$
\begin{aligned}
O\left(\sum_{r=s+1}^{+\infty} |\mathsf{BAD}_{r-1,s}| \cdot (\ln |S_{\mathrm{cur}}| + \ln \delta_r^{-1}) \varepsilon_r^{-2}\right) &= O\left(\sum_{r=s+1}^{+\infty} \frac{1}{8^{r-s}} \cdot |\mathsf{BAD}_s| \cdot (\ln k + \ln \delta^{-1} + \ln r) \varepsilon_r^{-2}\right) \\
&= O\left(|\mathsf{BAD}_s| \cdot (\ln k + \ln \delta^{-1} + \ln s) \cdot 4^s\right)
\end{aligned}
$$

Putting them together, we can see the number of samples incurred by UniformSample is bounded by:

$$
O\left(\sum_{s=1}^{+\infty} (|\mathsf{BAD}_s| + |\mathsf{OPT}_s|) \cdot (\ln k + \ln \delta^{-1} + \ln s) \cdot 4^s\right),
$$

which simplifies to $O\left(\sum_{e \in S} \Delta_e^{-2} (\ln k + \ln \delta^{-1} + \ln \ln \Delta_e^{-1})\right)$. Finally, we consider the number of samples taken by PAC-SamplePrune. Noticing $n_{\mathrm{opt}} = \mathrm{rank}(\mathcal{M}_{\mathrm{cur}})$, the number of samples is $O(|S_{\mathrm{cur}}|(\ln n_{\mathrm{opt}} + \ln \delta_r^{-1}) \varepsilon_r^{-2})$. So PAC-SamplePrune does not affect the sample complexity, and we finish our proof. ■

5. Future Work

In this paper, we present nearly-optimal algorithms for both the exact and PAC versions of the pure-exploration problem subject to a matroid constraint in a stochastic multi-armed bandit game: given a set of arms with a matroid constraint on them, pick a basis of the matroid whose weight (the sum of expectations over arms in this basis) is as large as possible, with high probability.

An immediate direction for investiation is to extend our results to other polynomial-time-computable combinatorial constraints: s-t paths, matchings (or more generally, the intersection of two matroids), etc. The model also extends to NP-hard combinatorial constraints, but there we would likely compare our solution against α-approximate solutions, instead of the optimal solution. Considering non-linear functions of the means is another natural next step. Yet another, perhaps more challenging, direction is to consider stochastic optimization problems, where the solution may depend on other details of the distributions than just the means.

References

Jean-Yves Audibert and Sébastien Bubeck. Best arm identification in multi-armed bandits. In *COLT-23th Conference on Learning Theory-2010*, pages 13–p, 2010.

Jean-Yves Audibert, Sébastien Bubeck, and Gábor Lugosi. Regret in online combinatorial optimization. *Mathematics of Operations Research*, 39(1):31–45, 2013.

Sébastien Bubeck and Nicolo Cesa-Bianchi. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *arXiv preprint arXiv:1204.5721*, 2012.

Sébastien Bubeck, Tengyao Wang, and Nitin Viswanathan. Multiple identifications in multi-armed bandits. *arXiv preprint arXiv:1205.3181*, 2012.

Séebastian Bubeck, Tengyao Wang, and Nitin Viswanathan. Multiple identifications in multi-armed bandits. In *Proceedings of The 30th International Conference on Machine Learning*, pages 258–265, 2013.

Wei Cao, Jian Li, Yufei Tao, and Zhize Li. On top-k selection in multi-armed bandits and hidden bipartite graphs. In *Advances in Neural Information Processing Systems*, pages 1036–1044, 2015.

Nicolo Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge university press, 2006.

Nicolo Cesa-Bianchi and Gábor Lugosi. Combinatorial bandits. *Journal of Computer and System Sciences*, 78(5):1404–1422, 2012.

Lijie Chen and Jian Li. On the optimal sample complexity for best arm identification. *arXiv preprint arXiv:1511.03774*, 2015.

Shouyuan Chen, Tian Lin, Irwin King, Michael R Lyu, and Wei Chen. Combinatorial pure exploration of multi-armed bandits. In *Advances in Neural Information Processing Systems*, pages 379–387, 2014.

Wei Chen, Yajun Wang, and Yang Yuan. Combinatorial multi-armed bandit: General framework and applications. In *Proceedings of the 30th International Conference on Machine Learning*, pages 151–159, 2013.

Eyal Even-Dar, Shie Mannor, and Yishay Mansour. Pac bounds for multi-armed bandit and markov decision processes. In *Computational Learning Theory*, pages 255–270. Springer, 2002.

Eyal Even-Dar, Shie Mannor, and Yishay Mansour. Action elimination and stopping conditions for the multi-armed bandit and reinforcement learning problems. *The Journal of Machine Learning Research*, 7:1079–1105, 2006.

RH Farrell. Asymptotic behavior of expected sample size in certain one sided tests. *The Annals of Mathematical Statistics*, pages 36–72, 1964.

Victor Gabillon, Mohammad Ghavamzadeh, Alessandro Lazaric, and Sébastien Bubeck. Multi-bandit best arm identification. In *Advances in Neural Information Processing Systems*, pages 2222–2230, 2011.

Victor Gabillon, Mohammad Ghavamzadeh, and Alessandro Lazaric. Best arm identification: A unified approach to fixed budget and fixed confidence. In *Advances in Neural Information Processing Systems*, pages 3212–3220, 2012.

Victor Gabillon, Alessandro Lazaric, Mohammad Ghavamzadeh, Ronald Ortner, and Peter Bartlett. Improved learning complexity in combinatorial pure exploration bandits. In *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*, pages 1004–1012, 2016.

Kevin Jamieson, Matthew Malloy, Robert Nowak, and Sébastien Bubeck. lil'ucb: An optimal exploration algorithm for multi-armed bandits. *COLT*, 2014.

Shivaram Kalyanakrishnan, Ambuj Tewari, Peter Auer, and Peter Stone. Pac subset selection in stochastic multi-armed bandits. In *Proceedings of the 29th International Conference on Machine Learning (ICML-12)*, pages 655–662, 2012.

David R Karger. Random sampling and greedy sparsification for matroid optimization problems. *Mathematical Programming*, 82(1-2):41–81, 1998.

David R Karger, Philip N Klein, and Robert E Tarjan. A randomized linear-time algorithm to find minimum spanning trees. *Journal of the ACM (JACM)*, 42(2):321–328, 1995.

Zohar Karnin, Tomer Koren, and Oren Somekh. Almost optimal exploration in multi-armed bandits. In *Proceedings of the 30th International Conference on Machine Learning (ICML-13)*, pages 1238–1246, 2013.

Emilie Kaufmann and Shivaram Kalyanakrishnan. Information complexity in bandit subset selection. In *Conference on Learning Theory*, pages 228–251, 2013.

Emilie Kaufmann, Olivier Cappé, and Aurélien Garivier. On the complexity of best arm identification in multi-armed bandit models. *arXiv preprint arXiv:1407.4443*, 2014.

Shie Mannor and John N Tsitsiklis. The sample complexity of exploration in the multi-armed bandit problem. *The Journal of Machine Learning Research*, 5:623–648, 2004.

Rajeev Motwani and Prabhakar Raghavan. *Randomized algorithms*. Chapman & Hall/CRC, 2010.

Herbert Robbins. Some aspects of the sequential design of experiments. In *Herbert Robbins Selected Papers*, pages 169–177. Springer, 1985.

Yuan Zhou, Xi Chen, and Jian Li. Optimal pac multiple arm identification with applications to crowdsourcing. In *Proceedings of the 31st International Conference on Machine Learning (ICML-14)*, pages 217–225, 2014.

Appendix A. Preliminaries in Probability

We first introduce the following versions of the standard Chernoff-Hoeffding bounds.

Proposition A.1 *Let $X_i (1 \le i \le n)$ be n independent random variables with values in $[0, 1]$. Let $X = \frac{1}{n}\sum_{i=1}^{n} X_i$. The following statements hold:*

1. *For every $t > 0$, we have that*

$$\Pr[X - \mathbb{E}[X] \ge t] \le \exp(-2t^2 n), \; and$$

$$\Pr[X - \mathbb{E}[X] \le -t] \le \exp(-2t^2 n).$$

2. *For any $\epsilon > 0$, we have that*

$$\Pr[X < (1-\epsilon)\mathbb{E}[X]] \le \exp(-\epsilon^2 n\mathbb{E}[X]/2), \; and$$

$$\Pr[X > (1+\epsilon)\mathbb{E}[X]] \le \exp(-\epsilon^2 n\mathbb{E}[X]/3).$$

Applying the above Proposition, we can get useful upper bound for the binomial distribution.

Corollary A.2 *Suppose the random variable X follows the binomial distribution $\mathsf{Bin}(n, p)$, i.e., $\Pr[X = k] = \binom{n}{k} p^k (1-p)^{n-k}$ for $k \in \{0, 1, \ldots, n\}$. It holds that for any $\epsilon > 0$,*

$$\Pr[X < (1-\epsilon)pn] \le \exp(-\epsilon^2 pn/2), \; and$$

$$\Pr[X > (1+\epsilon)pn] \le \exp(-\epsilon^2 pn/3).$$

We also need the following Chernoff-type concentration inequality (see Proposition A.4. in Zhou et al. (2014)).

Proposition A.3 *Let $X_i (1 \le i \le k)$ be independent random variables. Each X_i takes value a_i $(a_i \ge 0)$ with probability at most $\exp(-a_i^2 t)$ for some $t \ge 0$, and 0 otherwise. Let $X = \frac{1}{k}\sum_{i=1}^{k} X_i$. For every $\epsilon > 0$, when $t \ge \frac{2}{\epsilon^2}$, we have that*

$$\Pr[X > \epsilon] < \exp(-\epsilon^2 tk/2).$$

We need to introduce the definition of the negative binomial distributions (see e.g., (Motwani and Raghavan, 2010, pp.446)).

Definition A.4 *Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with the common distribution being the geometric distribution with parameter p. The random variable $X = X_1 + X_2 + \ldots + X_n$ denotes the number of coin flips (each one has probability p to be HEAD) needed to obtain n HEADS. The random variable X has the negative binomial distribution with parameters n and p, denote as $X \sim \mathsf{NegBin}(n;p)$.*

Lemma A.5 $\Pr[\mathsf{NegBin}(n;p) > r] = \Pr[\mathsf{Bin}(r;p) < n]$.

Proof Consider the event $\mathsf{NegBin}(n;p) > r$. By the definition of $\mathsf{NegBin}(n;p)$, it is equivalent to the event that during the first r coin flips, there are less than n HEADS. The lemma follows immediately. ■

Definition A.6 *(stochastic dominance) We say a random variable X* stochastically dominates *another random variable Y if for all $r \in \mathbb{R}$, we have $\Pr[X > r] \ge \Pr[Y > r]$.*

Appendix B. Missing Proofs

Proof [Proof of Proposition 1.6] Let I be an ε-optimal solution. We show it is also elementwise-ε-optimal. Let o_i be the arm with the i^{th} largest mean in OPT and a_i be the arm with the i^{th} largest mean in I. Suppose for contradiction that $\mu(a_i) < \mu(o_i) - \varepsilon$ for some $i \in [k]$ where $k = \operatorname{rank}(\mathcal{S})$. Now, consider the sorted list of the arms according to the modified cost function $\mu_{I,\varepsilon}$. The arm a_i is ranked after o_i and all o_j with $j < i$. Let P be the set of all arms with mean no less than o_i with respect to $\mu_{I,\varepsilon}$. Clearly, $\operatorname{rank}(P) \ge i$. So the greedy algorithm should select at least i elements in P, while I only has at most $i - 1$ elements in P, contradicting the optimality of I with respect to $\mu_{I,\varepsilon}$.

For the second part, take a BEST-k-ARM ($k = 2$) instance with four arms: $\mu(a_1) = 0.91, \mu(a_2) = 0.9, \mu(a_3) = 0.89, \mu(a_4) = 0.875$. The set $\{a_3, a_4\}$ is elementwise-0.3-optimal, but not 0.3-optimal. ■