Problems in Pure Exploration Multi-Armed Bandits with Multi-Dimensional Feedback and Criteria

by

Julian Katz-Samuels

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Electrical and Computer Engineering)
in The University of Michigan
2019

Doctoral Committee:

Professor Clayton Scott, Chair
Associate Professor Laura Balzano
Professor Satinder Singh
Associate Professor Ambuj Tewari

Julian Katz-Samuels

jkatzsam@umich.edu

ORCID iD: 0000-0002-9905-4222

Acknowledgements

I have had the great fortune of having Professor Clayton Scott as an adviser. When as a second year in a PhD program in Sociology at the University of Michigan I discovered my interest in machine learning, Clay took a chance on me, enabling me to follow this interest. Throughout my time as his student, he has always been patient with me and pushed me to search for more interesting problems and to come up with better solutions. Finally, I'm also thankful for all of the life advice that he has given me over the years.

I'd also like to thank my parents for their constant support throughout my doctoral studies.

Table of Contents

List of Figures

List of Tables

Abstract

Many applications can be modeled as follows: an agent is given access to several distributions and she wishes to determine those that meet some pre-specified criteria by sampling from the distributions in a sequential experiment. For example internet companies often perform A/B/n testing, which consists of determining which of several website design options is the best (e.g., maximizes the probability of a purchase) by diverting traffic to each of the options. As another example, in crowdsourcing, it is important to identify high-quality workers from a large pool of workers, e.g., those who give the correct answer with the highest probability on a random question. It is common practice to use "gold standard" questions, i.e., questions whose answers are known, to assess the quality of a worker. In these applications, it is of interest to determine the answer as quickly as possible so as to save money from experimentation. Pure exploration multi-armed bandits provide a framework for designing statistically efficient adaptive algorithms for solving these problems. Most of the literature has focused on settings where at each round of the sequential experiment the feedback is scalar-valued (i.e., one random number is observed). In this thesis, I argue that many applications exhibit multi-dimensional feedback and criteria. For example, in the crowdsourcing application, it is natural to require that a worker give the correct answer with high probability and *at a suitable pace* (e.g., within 15 seconds) and when a prospective crowdsourcing worker answers a gold standard question, there is feedback both on whether a worker is correct and her response time. To study and solve applications of this multi-dimensional nature, in Chapters 2 and 3 I introduce

novel pure exploration multi-armed bandit problems and design new algorithms that enjoy both strong theoretical guarantees and excellent empirical performance.

My dissertation also makes contributions to fundamental machine learning problems such as ϵ-good arm identification, mutual contamination models, and preference completion. In Chapter 4, I consider two multi-armed bandit problems: *(i)* given an $\epsilon > 0$, identify an arm with mean that is within ϵ of the largest mean and *(ii)* given a threshold μ_0 and k, minimize the time to identify k arms with means larger than μ_0. Existing lower bounds and algorithms for the PAC framework suggest that both of these problems require $\Omega(n)$ samples. However, I argue that these definitions not only conflict with how these algorithms are used in practice, but also that these results disagree with strong intuition that says *(i)* requires only $\Theta(\frac{n}{m})$ samples where $m = |\{i : \mu_i > \max_{i \in [n]} \mu_i - \epsilon\}|$ and *(ii)* requires $\Theta(\frac{n}{m}k)$ samples where $m = |\{i : \mu_i > \mu_0\}|$. I provide definitions that formalize these intuitions, obtain lower bounds that match the above sample complexities, and develop explicit, practical algorithms that achieve nearly matching upper bounds.

In Chapter 5, I consider the general framework of mutual contamination models. In this framework, the goal is to recover base distributions using only samples from mixtures of these base distributions. This framework models popular machine learning problems such as multiclass classification with label noise, learning with partial labels, and topic modeling. In this work, I provide algorithms for these problems that have theoretical guarantees under very general conditions.

In Chapter 6, I consider the problem of preference completion. In this problem, there is a pool of items and a pool of users. Each user rates a subset of the items and the goal is to recover the personalized ranking of each user over all of the

items. This problem is fundamental to recommender systems, arising in tasks such as movie recommendation and news personalization. In this chapter, I consider a statistical framework for nonparametric preference completion. I propose a simple k-nearest neighbors-like algorithm, and I show that it is consistent in this general nonparametric setting.

Chapter 1

Introduction

Pure exploration multi-armed bandit (MAB) problems provide a framework for designing a statistically efficient sequential experiment. This framework has applications in areas such as crowdsourcing, A/B testing, and clinical trials, and has recently received a surge of interest (Mannor and Tistisklis, 2004; Gabillon et al., 2012; Bubeck et al., 2013; Chen et al., 2014a; Jamieson et al., 2014a; Jamieson and Jain, 2018). Most of the literature has focused on settings where at each round of the sequential experiment the feedback is scalar-valued (i.e., one random number is observed). In this thesis, I argue that many applications exhibit multi-dimensional feedback and criteria, and I introduce novel pure exploration multi-armed bandit problems for modeling these applications. In the subsequent sections of this Introduction, I will describe and motivate these problems in more detail.

1.1 Feasible Arm Identification

Pure exploration multi-armed bandit (MAB) problems provide a framework for determining via a sequential experiment which of a set of distributions meet some criteria. In this setting, there are K distributions $\nu_1, \ldots, \nu_K$ and the agent sequentially chooses from which distribution to sample an observation. At the end of the sampling stage, the agent outputs the distributions believed to meet the desired criteria and the performance of the agent is measured based on the quality of this decision. In the MAB literature, distributions are also referred to as arms, and sampling a realization from a distribution ν_i is referred to as pulling

arm i.

Although the vast majority of work in MAB formulates arms as scalar-valued, in many applications, arms are multi-dimensional and the criteria for a good arm are multi-dimensional. To address this gap in the pure-exploration MAB literature, I introduce the feasible arm identification problem. In this problem, the agent is given a set of D-dimensional arms and a polyhedron $P = \{\boldsymbol{x} : A\boldsymbol{x} \leqslant b\} \subset \mathbb{R}^D$. Pulling an arm gives a random vector and the goal is to determine, using a fixed budget of T pulls, which of the arms have means belonging to P.

I provide a lower bound for the number of pulls that characterizes the problem's difficulty in terms of the distance of the mean of each arm to the boundary of the polyhedron. This lower bound does not depend on the number of constraints of the polyhedron, which suggests that the statistical difficulty of the problem is independent of the number of constraints. Because of this feature of the lower bound and the fact that polyhedra can approximate convex sets arbitrarily well, I am able to derive a lower bound for a generalization of the feasible arm identification problem where P is a convex set.

I also propose three algorithms MD-UCBE, MD-SAR, and MD-APT and provide a unified analysis establishing upper bounds for each of them. Each of the algorithms uses the distance of the empirical mean of an arm to decide whether to pull it. MD-UCBE is a modification of the algorithm UCBE (Upper Confidence Bound Exploration) from Audibert and Bubeck (2010); it uses the technique of confidence bounds to quantify its uncertainty about an estimate. MD-SAR extends the Successive Accepts Rejects algorithm from Bubeck et al. (2013). It divides the budget into phases, samples arms remaining in a set Q uniformly in each phase, and at the end of a phase, it removes from Q the arm about which it has the least uncertainty, i.e., that maximizes the distance of the empirical mean to the boundary. MD-APT is a modification of the APT ((Anytime Parameter-free Thresholdin) algorithm in Locatelli et al. (2016a). The algorithm allocates proportionally to the inverse square of the estimated distance of the mean to the boundary of the polyhedron. In this work, I show that MD-UCBE and MD-APT have upper bounds nearly matching our lower bound, while the upper bound of MD-SAR has a gap.

I also conduct experiments that demonstrate the effectiveness of our algorithms. I use synthetic datasets and real-world datasets based on a clinical trials application and a crowd-

sourcing application (Genovese et al., 2013; Snow et al., 2008). Our experiments show that MD-UCBE has a very difficult-to-tune hyperparameter and that MD-SAR and MD-APT dramatically outperform a uniform allocation approach, often obtaining a probability of error that is better by a factor of 10.

1.2 Top Feasible Arm Identification

In the top arm identification problem in multi-armed bandits, there are K scalar-valued distributions (also referred to as arms) and an agent plays a sequential game where, at each round, the agent chooses (or "pulls") one of the arms and observes an i.i.d. realization from it. At the end of the game, the agent outputs the set of m arms believed to have the largest means. This problem has applications in areas such as crowdsourcing, A/B testing, and clinical trials.

While top arm identification considers settings where the feedback is scalar-valued and the goal is maximization, in many applications, the feedback is multi-dimensional and it is of interest to perform *constrained maximization*. For example, in crowdsourcing, an important challenge is to identify high-quality workers that complete work at a suitable pace (e.g., below 15 seconds on average) and, in clinical trials, it is of interest to efficiently find drugs that are most likely to be effective and have an acceptably low probability of causing an adverse effect.

In this chapter, I propose a new variant of the top arm identification problem, *top feasible arm identification*, where there are K arms associated with D-dimensional distributions and the goal is to find m arms that maximize some known linear function of their means subject to the constraint that their means belong to a given set $P \subset \mathbb{R}^D$. This problem has many applications since in many settings, feedback is multi-dimensional and it is of interest to perform *constrained maximization*. I present problem-dependent lower bounds for top feasible arm identification and upper bounds for several algorithms. Our most broadly applicable algorithm, TF-LUCB-B (Top Feasible Lower Upper Confidence Bound Ball), has an upper bound that is loose by a factor of $O(D \log(K))$. Many problems of practical interest are two-dimensional and, for these, it is loose by a factor of $O(\log(K))$. Finally, I conduct

experiments on synthetic and real-world datasets that demonstrate the effectiveness of our algorithms. Our algorithms are superior both in theory and in practice to a naive two-stage algorithm that first identifies the feasible arms and then applies a best arm identification algorithm to the feasible arms.

1.3 Other Contributions

My dissertation also makes contributions to fundamental machine learning problems such as ϵ-good arm identification, mutual contamination models, and preference completion. In the following sections, I describe this work briefly and elaborate on each of the projects in its own chapter.

1.3.1 The True Sample Complexity of Identifying Good Arms

We consider two multi-armed bandit problems: *(i)* given an $\epsilon > 0$, identify an arm with mean that is within ϵ of the largest mean and *(ii)* given a threshold μ_0 and k, minimize the time to identify k arms with means larger than μ_0. Existing lower bounds and algorithms for the PAC framework suggest that both of these problems require $\Omega(n)$ samples. However, we argue that these definitions not only conflict with how these algorithms are used in practice, but also that these results disagree with strong intuition that says *(i)* requires only $\Theta(\frac{n}{m})$ samples where $m = |\{i : \mu_i > \max_{i \in [n]} \mu_i - \epsilon\}|$ and *(ii)* requires $\Theta(\frac{n}{m}k)$ samples where $m = |\{i : \mu_i > \mu_0\}|$. We provide definitions that formalize these intuitions, obtain lower bounds that match the above sample complexities, and develop explicit, practical algorithms that achieve nearly matching upper bounds.

1.3.2 Decontamination of Mutual Contamination Models

In many machine learning problems, the learner observes several random samples from different mixtures of unknown base distributions, with unknown mixing weights, and the goal is to infer these base distributions. Examples include binary classification with label noise, multiclass classification with label noise, classification with partial labels, and topic mod-

eling. The goal of this chapter is to develop a unified framework and set of tools to study statistical properties of these problems in a very general setting.

To this end, I use the general framework of mutual contamination models (Blanchard and Scott, 2014). In a mutual contamination model, there are L distributions $P_1, \ldots, P_L$ called *base distributions*. The learner observes M random samples

$$X_1^i, \ldots, X_{n_i}^i \overset{i.i.d.}{\sim} \tilde{P}_i = \sum_{j=1}^{L} \pi_{i,j} P_j \tag{1.1}$$

where $i = 1, \ldots, M$, $\pi_{i,j} \geqslant 0$, and $\sum_j \pi_{i,j} = 1$. Here $\pi_{i,j}$ is the probability that an instance of the *contaminated distribution* $\tilde{P}_i$ is a realization of P_j. The $\pi_{i,j}$s and P_js are unknown and the $\tilde{P}_i$s are observed through data. In this work, I avoid parametric models and assume that the sample space is arbitrary.

Our work focuses on three problems: multiclass classification with label noise, demixing of mixed membership models, and classification with partial labels. In multiclass classification with label noise, the learner observes samples whose labels have been randomly flipped. This problem arises in nuclear particle classification (Scott et al., 2013). When one draws samples of a specific particle, it is impossible to remove other types of particles from the background. Thus, each example is drawn from a mixture of the different types of particles.

In demixing of mixed membership models, the goal is to recover the base distributions up to a permutation. This problem arises in the task of automatically uncovering the thematic topics of a corpus of documents. Under the mixed membership model approach, the words of each document are thought of as being drawn from a document-specific mixture of topics. Specifically, documents correspond to the $\tilde{P}_i$s and the topics to the P_is. This approach is also referred to as topic modeling.

In classification with partial labels, each data point is labeled with a *partial label* $S \subset \{1, \ldots, L\}$; the true label is in S, but it is not known which label is the true one. There are many applications of classification with partial labels because often abundant sources of data are naturally associated with information that can be interpreted as partial labels. For example, consider the task of face recognition. On the internet, there are many images with captions that indicate who is in the picture but do not indicate which face belongs to which person. A partial label could be formed by associating each face with the names of

the individuals appearing in the same image (Cour et al., 2011).

In this work, I make several contributions. (i) I give sufficient conditions for identifiability of the three problems. (ii) I establish necessary conditions that in some cases match or are similar to the sufficient conditions. (iii) I introduce novel algorithms for the infinite and finite sample settings. These algorithms are nonparametric in the sense that they do not model the base distributions as a probability vector or other parametric model. (iv) I develop novel estimators for distributions obtained by iteratively applying the κ^* operator (defined below).

1.3.3 Nonparametric Preference Completion

In the preference completion problem, there is a pool of items and a pool of users. Each user rates a subset of the items and the goal is to recover the personalized ranking of each user over all of the items. This problem is fundamental to recommender systems, arising in tasks such as movie recommendation and news personalization. A common approach is to first estimate the ratings through either a matrix factorization method or a neighborhood-based method and to output personalized rankings from the estimated ratings (Koren et al., 2009; Zhou et al., 2008; Ning et al., 2011; Breese et al., 1998). Recent research has observed a number of shortcomings of this approach (Weimer et al., 2007; Liu and Yang, 2008); for example, many ratings-oriented algorithms minimize the RMSE, which does not necessarily produce a good ranking (Cremonesi et al., 2010). This observation has sparked a number of proposals of algorithms that aim to directly recover the rankings (Weimer et al., 2007; Liu and Yang, 2008; Lu and Negahban, 2014; Park et al., 2015; Oh et al., 2015; Gunasekar et al., 2016). Although these ranking-oriented algorithms have strong empirical performance, there are few theoretical guarantees to date and they all make specific distributional assumptions (discussed in more detail below). In addition, these results have focused on low-rank methods, while ranking-oriented neighborhood-based methods have received little theoretical attention.

In this chapter, I consider a statistical framework for nonparametric preference completion. I assume that each item i and each user u have unobserved features x_i and y_u, respectively, and that the associated rating is given by $g_u(f(x_i, y_u))$ where f is Lipschitz and g_u is a monotonic transformation that depends on the user. I make the following contribu-

tions. (i) I propose a simple k-nearest neighbors-like algorithm, (ii) I provide, to the best of our knowledge, the first consistency result for ranking-oriented algorithms in a nonparametric setting, and (iii) I provide a necessary and sufficient condition for the optimality of a solution (defined below) to the preference completion problem.

Chapter 2

Feasible Arm Identification

In this Chapter, I introduce the feasible arm identification problem, a pure exploration multi-armed bandit problem where the agent is given a set of D-dimensional arms and a polyhedron $P = \{\boldsymbol{x} : A\boldsymbol{x} \leqslant b\} \subset \mathbb{R}^D$. Pulling an arm gives a random vector and the goal is to determine, using a fixed budget of T pulls, which of the arms have means belonging to P. I propose three algorithms MD-UCBE, MD-SAR, and MD-APT and provide a unified analysis establishing upper bounds for each of them. I also establish a lower bound that matches up to constants the upper bounds of MD-UCBE and MD-APT. Finally, I demonstrate the effectiveness of our algorithms on synthetic and real-world datasets. This Chapter is joint work with Clayton Scott and was presented at the International Conference on Machine Learning in 2018.

2.1 Introduction

Pure exploration multi-armed bandit (MAB) problems provide a framework for determining via a sequential experiment which of a set of distributions meet some criteria. In this setting, there are K distributions $\nu_1, \ldots, \nu_K$ and the agent sequentially chooses from which distribution to sample an observation. At the end of the sampling stage, the agent outputs the distributions which he believes meet the desired criteria and the performance of the agent is measured based on the quality of this decision. In the MAB literature, distributions are also referred to as arms, and sampling a realization from a distribution ν_i is referred to as

pulling arm i. The most well-studied of these problems is top-k arm identification. In this problem, the goal is to find the k best arms, that is, k arms with the largest means. This problem and other pure exploration problems have applications in a wide range of areas, including crowdsourcing, A/B testing, and online advertising.

In many application domains, the arms and the criteria for a good arm are multi-dimensional in nature. For example, in crowdsourcing it is important to distinguish good workers from bad workers. For a multilabel classification task (where examples are associated with multiple labels), a worker can be modeled as a multi-dimensional arm where each dimension corresponds to her accuracy at identifying a particular label, and a natural definition for a "good worker" is that her accuracy is above some threshold for each label (e.g., 90%). A common approach for finding such workers is to use a collection of examples labeled by domain experts as a set of tests. Since workers are paid for each example that they label, often an organization is only willing to spend a limited number of queries to find good workers and an effective method under this budget constraint is needed. Other examples where this multi-dimensional structure arises include A/B testing and clinical trials.

The pure exploration MAB literature lacks *(i)* a simple framework for describing problems where the arms and criteria are multi-dimensional and *(ii)* practical algorithms for addressing these problems. In this chapter, we aim to address this gap. We introduce the *feasible arm identification problem* in which arms are associated with multi-dimensional distributions and the goal is to find arms whose means belong to a given polyhedron[1] $P = \{\boldsymbol{x} : A\boldsymbol{x} \leqslant b\}$. Polyhedra encompass a large class of regions that can model common user-defined constraints, including thresholds or ranges on individual dimensions and linear constraints involving multiple dimensions. We propose several algorithms for the fixed budget setting and provide upper and lower bounds. Finally, we demonstrate through experiments on synthetic and real-world datasets that by leveraging the geometry of the problem, our methods significantly outperform a uniform allocation strategy. Indeed, in several of our experiments, our methods find the feasible arms with a probability that is a factor of 10 better than that of a uniform allocation strategy. All proofs are contained in the supplementary material.

[1]There are several conflicting definitions of polyhedra. We define a polyhedron as the intersection of a finite number of halfspaces (Boyd and Vandenberghe, 2004).

2.2 Related Work

MABs have received a significant amount of attention. Most work considers minimizing the cumulative regret instead of a pure exploration objective. There have been relatively few works on multi-dimensional arms and criteria in this regime (Drugan and NowÃĺ, 2013; Busa-Fekete et al., 2017; Tekin and Turgay, 2017). Drugan and NowÃĺ (2013) modify a UCB algorithm to find all arms on the Pareto front. Busa-Fekete et al. (2017) use the Generalized Gini Index to optimize all objectives in a fair way. Tekin and Turgay (2017) consider a contextual MAB setting where the goal is to maximize the total reward in a non-dominant objective, subject to the constraint that the total reward in a dominant objective is maximized. These works differ from our work in that (i) they consider the cumulative regret setting, which is fundamentally different from the pure exploration setting (Bubeck et al., 2009), and (ii) they aim to either balance multiple objective functions or find arms on the Pareto front, whereas we aim to find feasible arms, where feasibility is defined by membership in a given polyhedron.

In recent years, there have been many advances in pure exploration MABs in the fixed confidence and fixed budget settings (Mannor and Tistisklis, 2004; Gabillon et al., 2012; Bubeck et al., 2013; Chen et al., 2014a; Jamieson et al., 2014a). A limited number of works have considered multi-dimensional feedback. Auer et al. (2016) considered a variant of the top arm identification problem where arms are multi-dimensional with each dimension corresponding to a distinct objective that an agent wishes to optimize, and the goal is to identify the Pareto front of the arms. In contrast to our work, they consider the fixed confidence setting. More importantly, Pareto front identification and feasible arm identification are mathematically very different problems and apply to distinct situations. Whereas Pareto front identification is relevant to multi-objective optimization, the feasible arm identification problem is useful for situations where there are user-defined criteria for what qualifies as a good arm.

Chen et al. (2017a) recently proposed the *general sampling problem*, which can model a setting where arms are multi-dimensional and the goal is to find arms with means belonging to a given polyhedron. There are several major differences with our work. First, Chen

et al. (2017a) do not consider multi-dimensional feedback as the agent can sample from one dimension of one arm at a time. Second, whereas they study the fixed confidence setting, we study the fixed budget setting. Third, they assume isotropic Gaussian arms, whereas we assume each arm is associated with a multi-dimensional sub-Gaussian distribution. Finally, their proposed algorithm (see their Algorithm 7) is sample-inefficient and impractical since in its first stage, it employs a uniform allocation strategy until the confidence bounds (defined with $\delta = 0.01$) of all of the means either intersect with the given polyhedron or do not intersect with the given polyhedron.

Locatelli et al. (2016a) introduced the thresholding bandit problem (TBP), which is essentially the scalar version of the feasible arm identification problem, and the algorithm APT. In TBP, there are K scalar-valued distributions, a threshold τ, and a budget T. The goal is to identify all of the distributions with means above τ. Our work significantly generalizes TBP by considering multi-dimensional arms and the problem of identifying those arms with means belonging to a given polyhedron. Unlike Locatelli et al. (2016a) who only analyze APT, we provide an unified analysis of three algorithms for the feasible arm identification problem. One of our algorithm, MD-APT, reduces to APT in the one-dimensional thresholding case and our upper bound also reduces to the upper bound of APT (up to constant factors). To deal with this general setting, we introduce a novel complexity measure that characterizes the hardness of determining whether an arm is in P. This measure is essentially the distance of the mean of an arm to the boundary of the polyhedron. In addition, our general setting introduces technical challenges for establishing upper and lower bounds. We overcome these by using tools from convex analysis, properties of multi-dimensional sub-Gaussian distributions, and change-of-measure arguments involving multi-dimensional distributions.

Recently, Zheng et al. (2017) considered a problem with a polyhedral constraint, but their setup is very different from our own. In their setting, the goal is to solve a linear program where either the constraints are not fully known or the cost function is not fully known but can be estimated by adaptive sampling. In our work, the constraints are known and we wish to learn which out of a collection of distributions have feasible means.

2.3 Setup

In this section, we formalize the feasible arm identification problem. To begin, we define some notation. For all $n \in \mathbb{N}$, let $[n] = \{1, \ldots, n\}$. For any $\boldsymbol{x} \in \mathbb{R}^D$ and $A \subset \mathbb{R}^D$, let $\operatorname{dist}(\boldsymbol{x}, A) = \inf_{\boldsymbol{y} \in A} \|\boldsymbol{x} - \boldsymbol{y}\|_2$. Let $\mathbf{1} = (1, \ldots, 1)^T \in \mathbb{R}^D$ and $\mathbf{1}\{\cdot\}$ denote the indicator function. Define $\mathcal{S}^{D-1} = \{\boldsymbol{x} \in \mathbb{R}^D : \|\boldsymbol{x}\|_2 = 1\}$.

Suppose we are given K stochastic arms. When the ith arm is pulled, a reward is drawn i.i.d. from a D-dimensional distribution ν_i. Denote $\boldsymbol{\mu}_i = \mathbb{E}_{\boldsymbol{X} \sim \nu_i} \boldsymbol{X}$. We assume that the agent is given a polyhedron $P = \{\boldsymbol{x} : A\boldsymbol{x} \leqslant \boldsymbol{b}\}$ where $A \in \mathbb{R}^{M \times D}$ such that

$$A = \begin{pmatrix} \boldsymbol{a}_1^t \\ \vdots \\ \boldsymbol{a}_M^t \end{pmatrix}$$

and $\boldsymbol{b} \in \mathbb{R}^M$. By dividing each constraint by $\|\boldsymbol{a}_j\|_2$, we can assume without loss of generality that $\|\boldsymbol{a}_j\|_2 = 1$ for all $j \in [M]$. Let ∂P denote the boundary of P, i.e., $\partial P = \bar{P} \backslash P^\circ$ where $\bar{P}$ denotes the closure of P and P° denotes the interior of P. For simplicity, we assume that P has positive volume. Thus, ∂P is non-empty.

We consider the fixed budget setting. The game is as follows: there are T rounds and at each round t, the agent chooses an arm $I_t \in [K]$ and observes a realization $\boldsymbol{X}_t \sim \nu_{I_t}$. The goal is to identify all of the arms whose means belong to the polyhedron. To define a performance measure, let $\epsilon > 0$ denote the tolerance, and define $S_{P,\epsilon}^{\text{int}} \coloneqq \{i \in [K] : \boldsymbol{\mu}_i \in P \text{ and } \operatorname{dis}(\boldsymbol{\mu}_i, \partial P) \geqslant \epsilon\}$ and $S_{P,\epsilon}^{\text{out}} \coloneqq \{i \in [K] : \operatorname{dis}(\boldsymbol{\mu}_i, P) > \epsilon\}$. $S_{P,\epsilon}^{\text{int}}$ is the set of arms that lie in the interior of P by at least ϵ and $S_{P,\epsilon}^{\text{out}}$ is the set of arms that lie outside of P by at least ϵ. Let $\widehat{S} \subset [K]$ denote the set of arms outputted by an algorithm. We define the following error measure:

$$\mathcal{L}_{T,P,\epsilon}(\widehat{S}) \coloneqq \mathbf{1}\{\widehat{S} \cap S_{P,\epsilon}^{\text{out}} \neq \varnothing \vee \widehat{S}^c \cap S_{P,\epsilon}^{\text{int}} \neq \varnothing\}$$

In words, the goal is to identify all of the arms with means belonging to the polyhedron up to tolerance ϵ in the sense that an algorithm is successful if its output includes every arm i such that $\boldsymbol{\mu}_i \in P$ and $\operatorname{dis}(\boldsymbol{\mu}_i, \partial P) \geqslant \epsilon$ and excludes every arm l such that $\operatorname{dis}(\boldsymbol{\mu}_l, P) > \epsilon$.

We define the *margin* of arm i as

$$\begin{aligned}
\Delta_i^{P,\epsilon} &\coloneqq \operatorname{dist}(\boldsymbol{\mu}_i, \partial P) + \epsilon \\
&= \begin{cases} \min_{j\in[M]} \operatorname{dist}(\boldsymbol{\mu}_i, \{\boldsymbol{x} : \boldsymbol{a}_j^t \boldsymbol{x} = b_j\}) + \epsilon & : \boldsymbol{\mu}_i \in P \\ \operatorname{dist}(\boldsymbol{\mu}_i, P) + \epsilon & : \boldsymbol{\mu}_i \notin P \end{cases} \qquad (2.1) \\
&= \begin{cases} \min_{j\in[M]} b_j - \boldsymbol{a}_j^t \boldsymbol{\mu}_i + \epsilon & : \boldsymbol{\mu}_i \in P \\ \operatorname{dist}(\boldsymbol{\mu}_i, P) + \epsilon & : \boldsymbol{\mu}_i \notin P \end{cases} \qquad (2.2)
\end{aligned}$$

where line (2.1) follows by Lemma 13 and line (2.2) follows by the closed form solution of the distance from a point to a hyperplane and $\|\boldsymbol{a}_j\|_2 = 1$ (Boyd and Vandenberghe, 2004).

The *complexity* of an instance of the feasible arm identification problem is defined to be:

$$H_{P,\epsilon} \coloneqq \sum_{i\in[K]} [\Delta_i^{P,\epsilon}]^{-2}.$$

In words, an instance has low complexity if all of the arms are far from the boundary of the polyhedron and high complexity if some of the arms are very close to the boundary. The intuition behind this complexity measure is that for an algorithm to output the correct answer about arm i, it is sufficient to guarantee that an estimate $\widehat{\boldsymbol{\mu}}_i$ is within a ball centered at $\boldsymbol{\mu}_i$ with radius $\frac{\Delta_i^{P,\epsilon}}{2}$ (see Lemma 3). For the sake of brevity, we usually write $\mathcal{L}_{T,\epsilon}(\widehat{S})$, Δ_i, and H instead of $\mathcal{L}_{T,P,\epsilon}(\widehat{S})$, $\Delta_i^{P,\epsilon}$, and $H_{P,\epsilon}$, respectively.

Our analysis assumes that each ν_i is a multi-dimensional sub-Gaussian distribution, which we now define (see Vershynin et al. (2017) for more details). Let X be a scalar random variable. We say that X is *R-sub-Gaussian* if $\mathbb{E}\exp(\frac{X^2}{R^2}) \leqslant 2$. We define the sub-Gaussian norm of X as the smallest R that satisfies the above requirement:

$$\|X\|_{\psi_2} = \inf\{R > 0 : \mathbb{E}\exp(\frac{X^2}{R^2}) \leqslant 2\}.$$

A random vector $\boldsymbol{X} \in \mathbb{R}^D$ is sub-Gaussian if $\boldsymbol{X}^t\boldsymbol{a}$ is sub-Gaussian for all $\boldsymbol{a} \in \mathbb{R}^D$. The sub-Gaussian norm of $\boldsymbol{X}$ is defined as

$$\|\boldsymbol{X}\|_{\psi_2} = \sup_{\boldsymbol{a}\in\mathcal{S}^{D-1}} \left\|\boldsymbol{X}^t\boldsymbol{a}\right\|_{\psi_2}.$$

We say that a random vector $\boldsymbol{X}$ is *R-sub-Gaussian* if $\|\boldsymbol{X}\|_{\psi_2} \leqslant R$. Henceforth, we assume that $\nu_1, \ldots, \nu_K$ are R-sub-Gaussian. See Vershynin (2012) for a discussion of sub-Gaussian distributions.

2.4 Lower Bound

In this section, we establish a lower bound for the feasible arm identification problem. Our construction takes any polyhedron P and means $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K \in P^\circ$ and produces a collection of problems such that any algorithm makes a mistake on one of the problems with probability at least on the order of $\exp(-c\frac{T}{H})$ (where c is a constant). In fact, this lower bound holds even when the algorithm is given the distance of each arm to the boundary of the polyhedron. If $A \subset \mathbb{R}^D$ is closed and $\boldsymbol{x} \in \mathbb{R}^D$, let $\mathrm{Proj}_A(\boldsymbol{x})$ denote the projection of $\boldsymbol{x}$ onto A.

Theorem 1. *Let $P = \{\boldsymbol{x} \in \mathbb{R}^D : A\boldsymbol{x} \leqslant \boldsymbol{b}\}$ have positive volume and $\epsilon \geqslant 0$ such that $P_\epsilon^\circ \coloneqq \{\boldsymbol{x} \in P : \ \mathrm{dist}(\boldsymbol{x}, \partial P) > \epsilon)\}$ is nonempty. Let $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K \in P_\epsilon^\circ$, $\boldsymbol{\tau}_i \in \mathrm{Proj}_{\partial P}(\boldsymbol{\mu}_i)$ for all $i \in [K]$, and $\boldsymbol{\mu}_i' = \boldsymbol{\mu}_i + 2(\boldsymbol{\tau}_i - \boldsymbol{\mu}_i)$ for all $i \in [K]$. Let ν_i denote the distribution $N(\boldsymbol{\mu}_i, I)$ and ν_i' the distribution $N(\boldsymbol{\mu}_i', I)$. Let $\mathcal{B}^0$ denote the product distribution $\nu_1 \otimes \ldots \otimes \nu_K$ and $\mathcal{B}^i$ denote the product distribution*

$$\nu_1 \otimes \ldots \otimes \nu_{i-1} \otimes \nu_i' \otimes \nu_{i+1} \otimes \ldots \otimes \nu_K.$$

Then, $\mathcal{B}^0, \ldots, \mathcal{B}^K$ have the same problem complexity

$$H = \sum_{i=1}^{K} [\ \mathrm{dist}(\boldsymbol{\mu}_i, \partial P) + \epsilon]^{-2}$$

and for any algorithm,

$$\max_{i \in \{0,\ldots,K\}} \mathbb{E}_{\mathcal{B}^i}(\mathcal{L}_{T,\epsilon}(\widehat{S})) \geqslant \exp(-13\frac{T}{H} - 25D\log(48(\log(T)+1)KD))).$$

This lower bound is equal to the lower bound of Locatelli et al. (2016a) (see their Theorem 1) up to the factor of D and constants. Since $D\log((\log(T)+1)KD))$ grows very slowly as a function of T in comparison with $\frac{T}{H}$, the dependence on D is quite mild. We also note that the lower bound does not depend on the number of constraints M in the polyhedron P, which suggests that the number of constraints of P does not directly affect the statistical difficulty of the feasible arm identification problem. Since polyhedra approximate convex sets arbitrarily well, the independence of our lower bound from M enables us to derive a nearly identical lower bound for the more general version of the feasible arm identification problem where P is convex (see the supplementary material for details).

The proof of Theorem 1 is based on a novel lower bound construction with multidimensional distributions for MABs. Often, lower bounds in the bandit literature modify scalar distributions and the main idea is to perturb the mean of a scalar distribution by making it either larger or smaller. In the feasible arm identification problem, picking a direction to perturb the mean of a distribution is not so simple. Indeed, the direction depends on the polyhedron since for some polyhedra, changing the first coordinate does not produce points lying outside of the polyhedron. In our construction, we interchange a distribution ν_i with mean $\mu_i \in P^\circ$ with a distribution ν_i' with mean μ_i' that is shifted away from μ_i in the direction of its projection onto the boundary of P.

Theorem 1 also implies the following non-asymptotic minimax bound.

Corollary 1. *Let $P = \{\boldsymbol{x} \in \mathbb{R}^D : A\boldsymbol{x} \leqslant \boldsymbol{b}\}$ have positive volume, $\epsilon \geqslant 0$ such that P_ϵ° is nonempty, and $R > 0$. Let $\tilde{H} > 0$ such that there exists $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K \in P_\epsilon^\circ$ with*

$$\tilde{H} = \sum_{i=1}^{K} [\ \mathrm{dist}(\boldsymbol{\mu}_i, \partial P) + \epsilon]^{-2}.$$

Let $\mathcal{B}_{P,\epsilon,\tilde{H},R}$ denote the set of feasible arm identification problems on polyhedron P, with tolerance ϵ, and with K arms such that the distributions are R-sub-Gaussian and the problem complexity is less than $\tilde{H}$. Then, $T \geqslant 25D\log(48(\log(T)+1)KD))$ implies that, for any algorithm,

$$\sup_{\mathcal{B} \in \mathcal{B}_{P,\epsilon,\tilde{H},R}} \mathbb{E}_{\mathcal{B}}(\mathcal{L}_{T,\epsilon}(\widehat{S})) \geqslant \exp(-14\frac{T}{\tilde{H}R^2}).$$

In words, this result says essentially that for any polyhedron P and tolerance $\epsilon \geqslant 0$, the induced class of feasible arm identification problems with P and ϵ has a minimax lower bound on the order of $\exp(-c\frac{T}{HR^2})$ where c is a constant. Henceforth, we say that an algorithm is *nearly optimal* if for large enough T its expected loss decays as $\mathcal{O}(\exp(-c\frac{T}{HR^2}))$ where c is a constant.

2.5 Algorithms

In this section, we extend three algorithms to the feasible arm identification problem, namely, an upper confidence bound based algorithm (UCBE) (Audibert and Bubeck, 2010), a successive accepts and rejects algorithm (SAR) (Bubeck et al., 2013; Chen et al., 2014a), and

the Anytime Parameter-free Thresholding algorithm (APT) (Locatelli et al., 2016a). The main novelty of our approach is that our algorithms estimate the distance of the mean of each arm to the boundary of the polyhedron to decide which arm to pull. To begin, we introduce some notation. Let I_t denote the index of the arm chosen at time t. Let $X_{i,j,t}$ denote the tth realization of the jth coordinate of ν_i, $T_i(t) = \sum_{s=1}^{t-1} \mathbf{1}\{I_s = i\}$ denote the number of pulls of arm i at round t, and $\widehat{\boldsymbol{\mu}}_{i,t}$ denote the estimate of $\boldsymbol{\mu}_i$ after t samples, i.e., $\widehat{\boldsymbol{\mu}}_{i,t} = (\widehat{\mu}_{i,1,t}, \ldots, \widehat{\mu}_{i,D,t})^t$ where $\widehat{\mu}_{i,j,t} = \frac{1}{t}\sum_{s=1}^{t} X_{i,j,s}$.

The key quantity in each of these algorithms is the following empirical estimator of the margin of each arm:

$$\widehat{\Delta}_{i,t} = \begin{cases} \min_{j\in[M]} b_j - \boldsymbol{a}_j^t \widehat{\boldsymbol{\mu}}_{i,t} + \epsilon & : \widehat{\boldsymbol{\mu}}_{i,t} \in P \\ \operatorname{dist}(\widehat{\boldsymbol{\mu}}_{i,t}, P) + \epsilon & : \widehat{\boldsymbol{\mu}}_{i,t} \notin P \end{cases}$$

Given $\widehat{\boldsymbol{\mu}}_{i,t}$, $\operatorname{dist}(\widehat{\boldsymbol{\mu}}_{i,t}, P)$ can be computed by solving a quadratic program and, thus, the interior point method can compute $\widehat{\Delta}_{i,t}$ in runtime polynomial in M and D. Each of our algorithms updates one $\widehat{\Delta}_{i,t}$ in each round, thus solving at most T quadratic programs. Therefore, each algorithm can be implemented efficiently.

Algorithm 1 MD-UCBE: Multi-dimensional Upper Confidence Bound Exploration algorithm

Input: K arms, polyhedron P, tolerance ϵ, budget T, hyperparameter a
for $t = 1, \ldots, T$ **do**
  if $t \leqslant K$ **then**
    Sample $\boldsymbol{X}_t \sim \nu_t$.
  else
    Choose $I_t = \arg\min_i \widehat{\Delta}_{i,T_i(t)} - \sqrt{\frac{a}{T_i(t)}}$ and sample $\boldsymbol{X}_t \sim \nu_{I_t}$.
  end if
end for
Return: $\widehat{S} = \{i \in [K] : \widehat{\boldsymbol{\mu}}_{i,T_i(T+1)} \in P\}$

Next, we describe each of the algorithms and our results. Each algorithm outputs $\widehat{S} = \{i \in [K] : \widehat{\boldsymbol{\mu}}_{i,T_i(T+1)} \in P\}$. The algorithms differ in how they decide which arm to pull. MD-UCBE (Algorithm 1) is a modification of the algorithm UCBE from Audibert and Bubeck

(2010). At each time step t, it pulls an arm i that minimizes $\widehat{\Delta}_{i,T_i(t)} - \sqrt{\frac{a}{T_i(t)}}$ breaking ties arbitrarily where a is a hyperparameter. Theorem 2 gives an upper bound on its expected loss.

Theorem 2. *Let $K \geqslant 0, T \geqslant K$ and $\epsilon \geqslant 0$. Suppose $0 \leqslant a \leqslant \frac{25}{36}\frac{T-K}{H}$. Then, the expected loss of MD-UCBE is bounded as follows:*

$$\mathbb{E}[\mathcal{L}_{T,\epsilon}(\widehat{S})] \leqslant 2(\log(T)+1)K5^D \exp(-\frac{a}{1600R^2}).$$

Paralleling our upper bounds for MD-SAR and MD-APT, this result says that the degree of difficulty of a problem for MD-UCBE depends on H, i.e., the distance of the arms to the boundary of the polyhedron and the tolerance parameter ϵ. Theorem 2 suggests setting $a = \frac{25}{36}\frac{T-K}{H}$, in which case MD-UCBE is nearly optimal. One important shortcoming of this algorithm is that H is not known in practice, so it is unclear how to set the hyperparameter a. Indeed, in our experiments, we show that the performance of MD-UCBE is highly sensitive to the selection of a.

Algorithm 2 MD-SAR: Multi-dimensional Successive Accepts and Rejects algorithm

Input: K arms, polyhedron P, tolerance ϵ, budget T
$\overline{\log}(x) = \frac{1}{2} + \sum_{i=2}^{x} \frac{1}{i}$, $n_0 = 0$, $n_k = \left\lceil \frac{T-K}{\overline{\log}(K)(K+1-k)} \right\rceil$ $(k > 1)$
$Q = [K]$
for $k = 1, \ldots, K-1$ **do**
    Query $n_k - n_{k-1}$ samples from all arms $i \in Q$
    $Q \longleftarrow Q \backslash \arg\max_{i \in Q} \widehat{\Delta}_{i,n_k}$
end for
Return: $\widehat{S} = \{i \in [K] : \widehat{\boldsymbol{\mu}}_{i,T_i(T+1)} \in P\}$

MD-SAR (Algorithm 2) extends the SAR algorithm from Bubeck et al. (2013). It divides the budget T into $K-1$ rounds. In each round, it samples all of the arms belonging to $Q \subset [K]$ the same number of times. At the end of each round, it removes from Q an arm i that maximizes $\widehat{\Delta}_{i,T_i(t)}$. Intuitively, MD-SAR stops sampling from an arm i for which there is the least amount of uncertainty about whether $\boldsymbol{\mu}_i \in P$. Theorem 3 provides an upper bound on the expected loss of MD-SAR. It depends on a different complexity term that is

nevertheless related to H. Let (i) denote the index of the arm with the ith smallest margin so that $\Delta_{(1)} \leqslant \Delta_{(2)} \leqslant \ldots \leqslant \Delta_{(K)}$ and define the complexity parameter

$$H_2 = \max_{i \in [K]} i \Delta_{(i)}^{-2}.$$

The analysis of Audibert and Bubeck (2010) of the analogous quantities immediately implies that $H_2 \leqslant H \leqslant \log(2K) H_2$.

Theorem 3. *Let $K \geqslant 0, T \geqslant K$ and $\epsilon \geqslant 0$. Then, the expected loss of MD-SAR is bounded as follows:*

$$\begin{aligned}\mathbb{E}[\mathcal{L}_{T,\epsilon}(\widehat{S})] \leqslant & 2(\log(T)+1)K5^D \\ & \times \exp(-\frac{T-K}{1296 \log(2K) H_2} \frac{1}{R^2}) \\ & + 4K^3 5^D \exp(-\frac{T-K}{512 R^2 H_2}).\end{aligned}$$

Similar to previous results on SAR-type algorithms in the fixed budget setting (Audibert and Bubeck, 2010; Chen et al., 2014a), our upper bound on MD-SAR is loose by a factor of $\log(K)$ in the exponential. While the guarantee is not tight, it has the significant practical advantage over MD-UCBE that it does not involve a difficult-to-tune hyperparameter. On the other hand, MD-SAR has the limitation that it needs to know T in advance.

Algorithm 3 MD-APT: Multi-dimensional Anytime Parameter-Free Thresholding algorithm

Input: K arms, polyhedron P, tolerance ϵ, budget T
for $t = 1, \ldots, T$ **do**
  if $t \leqslant K$ **then**
    Sample $\boldsymbol{X}_t \sim \nu_t$.
  else
    Choose $I_t = \arg\min_i \widehat{\Delta}_{i,T_i(t)} \sqrt{T_i(t)}$ and sample $\boldsymbol{X}_t \sim \nu_{I_t}$.
  end if
end for
Return: $\widehat{S} = \{i \in [K] : \widehat{\boldsymbol{\mu}}_{i,T_i(t+1)} \in P\}$

MD-APT (Algorithm 3) is a modification of the APT algorithm in Locatelli et al. (2016a). After an initialization phase in which it pulls each arm once, at each round t, it pulls an arm i that minimizes $\widehat{\Delta}_{i,T_i(t)}\sqrt{T_i(t)}$. The intuition behind the algorithm is that if the margins Δ_i were known in advance, then a nearly optimal strategy would allocate samples to the arms proportionally to the Δ_i^2s. For simplicity, let $\epsilon = 0$; the case $\epsilon > 0$ is not as clear since arms whose distance to the boundary is less than ϵ do not need to be sampled at all.

Proposition 1. *Let $\epsilon = 0$. A static allocation strategy with a total of $T_i = \frac{T}{\Delta_i^2 H}$ pulls of the ith arm for all $i \in [K]$ achieves*

$$\mathbb{E}[\mathcal{L}_{T,\epsilon}(\widehat{S})] \leqslant 2K5^D \exp(-\frac{1}{8}\frac{T}{HR^2}).$$

Thus, such a static allocation is nearly optimal. Since the Δ_is are unknown, MD-APT samples the arms proportionally to the estimates $\widehat{\Delta}^2_{i,T_i(t)}$. Theorem 4 gives an upper bound on the expected loss of MD-APT.

Theorem 4. *Let $K \geqslant 0, T \geqslant 2K$, and $\epsilon \geqslant 0$. Then, the expected loss of MD-APT is bounded as follows:*

$$\mathbb{E}[\mathcal{L}_{T,\epsilon}(\widehat{S})] \leqslant 2(\log(T)+1)K5^D \exp(-\frac{T}{1296R^2H}).$$

This Theorem implies that MD-APT is nearly optimal. Further, unlike MD-UCBE, it is parameter-free and, unlike MD-SAR, it is an anytime algorithm in the sense that MD-APT does not require knowledge of the budget T. These properties make MD-ADT practical for many applications (Jun and Nowak, 2016).

We note that although the runtime of our algorithms certainly depends on M, our results suggest that their statistical performance is independent of M. We leverage this result and the fact that polyhedra approximate convex sets arbitrarily well to show that given minimal knowledge about $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K$, there exists a computationally inefficient algorithm with nearly the same guarantee as Theorem 4 for the more general version of the feasible arm identification problem where P is convex (see the supplementary material for details).

2.6 Analysis

Our analyses of the three algorithms are unified through a series of lemmas. The first key idea is a sufficient condition for $\widehat{\Delta}_{i,t}$ to concentrate around Δ_i. Lemma 1 shows that concentration of $\widehat{\boldsymbol{\mu}}_i(t)$ around its mean in the norm sense is sufficient.

Lemma 1. *Let $\gamma > 0$, $i \in [K]$, and $t \in [T]$. If $\|\widehat{\boldsymbol{\mu}}_{i,t} - \boldsymbol{\mu}_i\|_2 \leqslant \gamma$, then*

$$|\widehat{\Delta}_{i,t} - \Delta_i| \leqslant 2\gamma.$$

In the scalar case, concentration of the empirical margin around the true margin often follows by the triangle inequality. In our setting, because of the more complicated relationship between $\widehat{\boldsymbol{\mu}}_{i,t}$ and $\widehat{\Delta}_{i,t}$ such an argument is not sufficient.

The second key idea is that with an appropriately high probability, $\widehat{\boldsymbol{\mu}}_{i,t}$ concentrates around its mean in the norm sense. The main tools are Hoeffding's maximal inequality (see Lemma 5) and an ϵ-net, which we now define (Vershynin et al., 2017).

Definition 1. *Let $A \subset \mathbb{R}^D$ and $\epsilon > 0$. $\mathcal{N} \subset A$ is an ϵ-net of A if $\forall \boldsymbol{x} \in A$, there exists $\boldsymbol{y} \in \mathcal{N}$ such that $\|\boldsymbol{x} - \boldsymbol{y}\|_2 \leqslant \epsilon$. Let $\mathcal{N} \subset A$ be an ϵ-net of A. We say that $\mathcal{N}$ is* minimal *if, for any other ϵ-net $\mathcal{M}$ of A, it holds that $|\mathcal{M}| \geqslant |\mathcal{N}|$.*

Lemma 2. *Let $\mathcal{N}$ be a minimal $\frac{1}{2}$-net on $\mathcal{S}^{D-1}$. Let $\omega > 0$. Define the event*

$$\Xi = \{\forall i \in [K], \forall \boldsymbol{y} \in \mathcal{N}, \forall r \in \{1, \ldots, T\} :$$
$$|\boldsymbol{y}^t(\widehat{\boldsymbol{\mu}}_{i.r} - \boldsymbol{\mu}_i)| \leqslant \sqrt{\frac{\omega^2}{4r}}\}.$$

Then, on Ξ, for all $i \in [K]$ and for all $r \in [T]$,

$$\|\widehat{\boldsymbol{\mu}}_{i.r} - \boldsymbol{\mu}_i\|_2 \leqslant \sqrt{\frac{\omega^2}{r}}$$

and

$$\Pr(\Xi) \geqslant 1 - 2(\log(T) + 1)K5^D \exp(-\frac{\omega^2}{16R^2}).$$

In effect, Lemma 1 and Lemma 2 together imply that with high probability, *(i)* $\widehat{\boldsymbol{\mu}}_{i,t}$ concentrates around $\boldsymbol{\mu}_i$ in the norm sense and *(ii)* $\widehat{\Delta}_{i,t}$ concentrates around Δ_i.

Finally, the third idea is the simple observation that if for all $i \in [K]$, $\widehat{\boldsymbol{\mu}}_{i,t}$ lies in a ball centered at $\boldsymbol{\mu}_i$ with radius $\frac{\Delta_i}{2}$, then an algorithm does not make a mistake.

Lemma 3. *Fix $t \in [T]$ and $i \in [K]$ and suppose that $\|\widehat{\boldsymbol{\mu}}_{i,t} - \boldsymbol{\mu}_i\|_2 < \frac{1}{2}\Delta_i$. Then, $A\boldsymbol{\mu}_i \leqslant \boldsymbol{b} - \epsilon\mathbf{1}$ implies that $A\widehat{\boldsymbol{\mu}}_{i,t} < \boldsymbol{b}$ and $\mathrm{dis}(\boldsymbol{\mu}_i, P) \geqslant \epsilon$ implies that $\widehat{\boldsymbol{\mu}}_{i,t} \notin P$.*

The analysis of each algorithm then proceeds as follows. First, suppose some appropriately defined variant of the event Ξ in Lemma 2. Second, by Lemmas 1 and 2, *(i)* $\widehat{\boldsymbol{\mu}}_{i,t}$ concentrates around $\boldsymbol{\mu}_i$ in the norm sense and *(ii)* $\widehat{\Delta}_{i,t}$ concentrates around Δ_i. Given these concentration results, it is shown that each algorithm pulls each arm a sufficient number of times so that Lemma 3 can be applied.

2.7 Experiments

In this section, we conduct experiments on synthetic and real-world datasets. In addition to the algorithms MD-UCBE, MD-SAR, and MD-APT, we consider a uniform allocation algorithm (UA), which samples the arms in a cyclic fashion. We consider the performance of MD-UCBE under four hyperparameter settings $a_i = i\frac{25}{36}\frac{T-K}{H}$ for $i \in \{.1, 1, 10, 100\}$. Let MD-UCBE[i] denote MD-UCBE with hyperparameter a_i. Note that the larger i is, the more MD-UCBE[i] explores and that our theoretical guarantee in Theorem 2 only covers $i \leqslant 1$. To calculate $\widehat{\Delta}_{i,t}$, we use the quadratic programming solver in the CVXOPT package for python. We average all experiments over 2000 trials.

2.7.1 Synthetic Experiments

Each experiment has 20 5-dimensional arms and is run for 2000 time steps. We use Gaussian distributions with variance $\frac{1}{4}$. For experiments 1, 2, and 3 we use a cube $P = \{\boldsymbol{x} \in \mathbb{R}^5 : 0 \leqslant x_i \leqslant 1\}$. In experiments 4 and 5, we use more complicated feasibility regions. In the following, we say an arm i is *irrelevant* if the error measure $\mathcal{L}_{T,\epsilon}(\cdot)$ does not depend on how i is categorized.

Experiment 1 (Four Groups with Irrelevant Arms): We set $\epsilon = 0.075$ and use $\boldsymbol{\mu}_{0:1} = (.8)^{\otimes 5}$, $\boldsymbol{\mu}_{2:3} = (.9)^{\otimes 5}$, $\boldsymbol{\mu}_{4:5} = (1.1)^{\otimes 5}$, $\boldsymbol{\mu}_{6:7} = (1.2)^{\otimes 5}$, $\boldsymbol{\mu}_8 = (.975)^{\otimes 5}$, $\boldsymbol{\mu}_9 = (1.025)^{\otimes 5}$, $\boldsymbol{\mu}_{10:19} = (.3)^{\otimes 5}$. Note that this problem has two irrelevant arms, $\boldsymbol{\mu}_8$ and $\boldsymbol{\mu}_9$.

Experiment 2 (Four Groups with no Irrelevant Arms): We set $\epsilon = 0$ and use

$\boldsymbol{\mu}_{0:1} = (.8)^{\otimes 5}$, $\boldsymbol{\mu}_{2:3} = (.9)^{\otimes 5}$, $\boldsymbol{\mu}_{4:5} = (1.1)^{\otimes 5}$, $\boldsymbol{\mu}_{6:7} = (1.2)^{\otimes 5}$, $\boldsymbol{\mu}_8 = (.95)^{\otimes 5}$, $\boldsymbol{\mu}_9 = (1.05)^{\otimes 5}$, $\boldsymbol{\mu}_{10:19} = (.3)^{\otimes 5}$. In comparison to experiment 1, we make it slightly easier to determine whether the arms $\boldsymbol{\mu}_8$ and $\boldsymbol{\mu}_9$ belong to the polyhedron because otherwise the difficulty of the problem prevents any algorithm from achieving substantial progress after 2000 time steps.

Experiment 3 (Linear Progression with Irrelevant Arms): We set $\epsilon = 0.075$ and use $\boldsymbol{\mu}_{0:3} = (.75)^{\otimes 5} + (0:3) \times .05$, $\boldsymbol{\mu}_4 = (.975)^{\otimes 5}$, $\boldsymbol{\mu}_5 = (1.025)^{\otimes 5}$, $\boldsymbol{\mu}_{6:9} = (1.25)^{\otimes 5} - (0:3) \times .05$, $\boldsymbol{\mu}_{10:19} = (1.15)^{\otimes 5}$. Note that this problem has two irrelevant arms, $\boldsymbol{\mu}_4$ and $\boldsymbol{\mu}_5$.

Experiment 4 (Four Groups on the Simplex): For this experiment, we use $P = \{\boldsymbol{x} \in \mathbb{R}^5 : x_i \geqslant 0,\ \sum_i x_i \leqslant 2\}$. We set $\epsilon = .1$. Let $\boldsymbol{c} = (.2)^{\otimes 5}$. We use $\boldsymbol{\mu}_{0:4} = \boldsymbol{c}$, $\boldsymbol{\mu}_{5:9} = 1.85 \cdot \boldsymbol{c}$, $\boldsymbol{\mu}_{10:14} = 2.25 \cdot \boldsymbol{c}$, and $\boldsymbol{\mu}_{15:19} = 1.95 \cdot \boldsymbol{c}$. $\boldsymbol{\mu}_{0:9}$ are good arms, $\boldsymbol{\mu}_{10:14}$ are bad arms, and $\boldsymbol{\mu}_{15:19}$ are irrelevant.

Experiment 5 (Ordered Polyhedron): For this experiment, we use $P = \{\boldsymbol{x} \in \mathbb{R}^5 : x_i \leqslant x_{i+1} \forall i \in [4]\}$ and $\epsilon = .1$. We use $\boldsymbol{\mu}_{0:3} = (0, .2, .4, .6, .8)^t$, $\boldsymbol{\mu}_{4:7} = (.0, .15, .3, .45, .6)^t$, $\boldsymbol{\mu}_{8:11} = (0, .2, .15, .6, .8])^t$, $\boldsymbol{\mu}_{12:15} = (0, .2, .05, .6, .8)^t$, and $\boldsymbol{\mu}_{16:19} = (0, .2, .4, .2, 0)^t$. The arms $\boldsymbol{\mu}_{8:11}$ are irrelevant.

The performance of MD-UCBE is very sensitive to the selection of its hyperparameter. MD-UCBE[1] and MD-UCBE[10] tend to do well, but MD-UCBE[100] explores too much so that it tends to perform only slightly better than UA and MD-UCBE[.1] does not explore enough. Although MD-UCBE[.1] has a theoretical guarantee, the constants are too large so that it never makes progress in solving the problems. MD-APT performs better than MD-SAR in experiments 1, 4, and 5 and worse than MD-SAR in experiments 2 and 3. In experiment 2, MD-APT pulls arm 8, which minimizes Δ_i, too frequently. It pulls arm 8 on average 904.8125 times, whereas MD-SAR more evenly spreads out its pulls, pulling arm 8 on average 317.751 times. We observe a similar phenomenon in a variant of experiment 3 where we set $\epsilon = 0$ and which we defer to the supplementary material due to lack of space. This suggests that in certain problems MD-APT focuses too much on specific arms with means near the boundary and does not allocate enough samples to other arms. On the other hand, MD-SAR utilizes knowledge of the time horizon T to effectively spread out samples. MD-APT's agnosticism about T may put it a disadvantage in the regime where some of the Δ_i are very small and T is small relative to H. As suggested by experiment 1, the ϵ

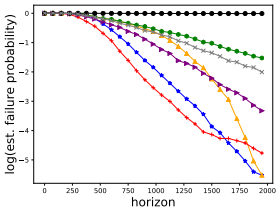


Figure 2.1: Four Groups on Cube with Irrelevant Arms

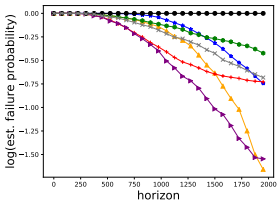


Figure 2.2: Four Groups on Cube, no Irrelevant Arms

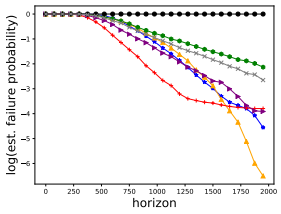


Figure 2.3: Linear Progression on Cube with Irrelevant Arms

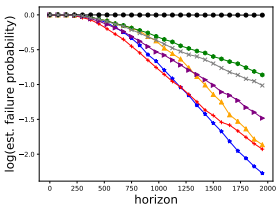


Figure 2.4: Four Groups on a Simplex

parameter can be used to counteract the sensitivity of MD-APT to arms with means near the boundary.

2.7.2 Application 1: Dose-Finding

In clinical trials, an important challenge is determining the appropriate dosage of a drug. The main difficulty is the trade-off that as the dosage increases, the effectiveness of the drug tends to increase, but the likelihood of adverse effects also increases. Thus, one must find a dosage that is sufficiently effective, but does not have too many side effects. We assume a situation where the side effects are mild enough not to be a concern for clinical trials, but could nevertheless be unacceptable for a final commercial product.

We investigate this problem by considering the data in Genovese et al. (2013) (see ARCR20 in week 16 in Table 2 and Table 3). In this study, the authors examine the drug secukinumab for treating rheumatoid arthritis. They consider four dosage levels (25mg, 75mg, 150mg, 300mg) and a placebo. We design a simulation based on their data where each arm

corresponds to a drug and has two attributes, the likelihood of being effective and the likelihood of causing an adverse effect. Let $\mu_{i,1}$ denote the probability of being effective and $\mu_{i,2}$ the probability of causing an adverse effect. Then, dosage levels 25mg, 75mg, 150mg, and 300mg have means $\boldsymbol{\mu}_1 = (.34, .519)^t, \boldsymbol{\mu}_2 = (.469, .612)^t, \boldsymbol{\mu}_3 = (.465, .465)^t, \boldsymbol{\mu}_4 = (.537, .61)^t$, respectively, and the placebo has mean $\boldsymbol{\mu}_5 = (.36, .58)^t$. We suppose that a drug is considered good if the probability of success is above .4 and the probability of adverse effects is below .5 and we set $\epsilon = 0$. Thus, only arm 3 is good and all other arms are bad. We chose these thresholds so that one drug is good; we did not try other threshold settings. We run the experiment for 1000 time steps.

Figure 2.6 gives the results of the experiment. MD-APT and MD-UCBE[10] perform better than the rest of the algorithms. MD-UCBE[1] performs slightly worse than UA, which may be because there are only 5 arms so that UA is not that bad of a strategy and MD-UCBE[1] does not explore sufficiently. MD-SAR only performs slightly better than UA. This may be because the time horizon is only 1000 time steps and there are only 5 arms.

2.7.3 Application 2: Crowdsourcing

We use a real-world dataset for the natural language processing task of affective text analysis (Snow et al., 2008). In this task, workers are asked to rate a short headline on valence and six emotions: disgust, fear, joy, anger, sadness and surprise. A group of experts also provide such ratings for the headlines.

We consider the problem of finding workers that tend to agree with the expert views on each of the tasks. We examine the deviation of a worker's ratings with the experts ratings. We normalize this deviation onto a scale of $[0, 1]$. Let $\mu_{i,j}$ denote the mean of worker i on task j and let $\bar{\mu}_j$ denote the mean of all of the workers on task j. We deem a worker i good if $\mu_{i,j} \leqslant \bar{\mu}_j$ for all $j \in [7]$. In words, a worker is good if for every task, he performs better than the average worker. To make this realistic, we assume that we are in a setting where the average worker performance on each task is known based on another pool of workers. We use a tolerance of $\epsilon = 0.02$. There is a total of 38 workers, where 30 workers are bad arms, 3 workers are good arms, and 5 workers are irrelevant. Because each worker only provides a small number (at least 20) of ratings, whenever an arm is pulled, the algorithm observes an

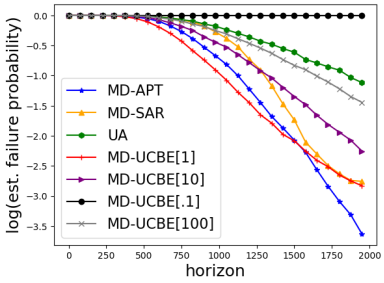


Figure 2.5: Ordered polyhedron

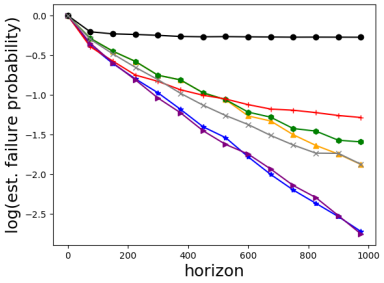


Figure 2.6: Dose-Finding Experiment

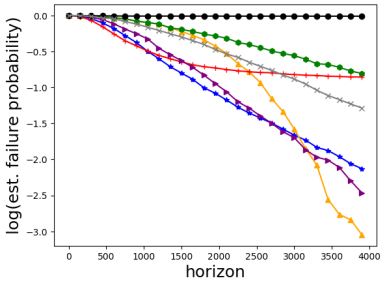


Figure 2.7: Crowdsourcing Experiment

observation chosen uniformly at random with replacement from the data associated with the arm. We run each algorithm for 4000 time steps and in each trial, we randomly permute the samples of each worker. In the supplementary material, we repeat this experiment, but we simulate each arm as a Gaussian distribution (see Section 2.18); the results are very similar.

Figure 2.7 gives the results of the experiment. Until roughly time step 3000, MD-APT and MD-UCBE[10] perform the best. Afterwards, MD-SAR does substantially better than MD-APT and MD-UCBE[10]. MD-UCBE[1] and MD-UCBE[100] perform only marginally better than UA.

2.7.4 Summary of Results

The experiments suggest that although MD-UCBE is a competitive algorithm, it is highly sensitive to hyperparameter selection, which limits its applicability in practice. MD-SAR and MD-APT tend to perform dramatically better than UA. For example, in the crowdsourcing

experiment, UA has a final error rate of roughly 52%, whereas MD-SAR has a final error rate of roughly 5%. Further, our algorithms can handle complicated polyhedra such as the polyhdron that requires that coordinates are sorted in ascending order (see experiment 5). These results suggest that MD-APT tends to perform better than MD-SAR, but in some settings (e.g., some arms with small Δ_i and H large relative to T) MD-APT focuses too much on some of the arms with means near the boundary. Because MD-SAR more evenly spreads out its pulls among the arms, it performs better in this regime.

2.8 Conclusion

In this chapter, we introduced the feasible arm identification problem. This problem provides a flexible framework for settings where arms are multi-dimensional and it is of interest to determine whether each arm satisfies user-defined multi-dimensional criteria. We provided a characterization of the difficulty of these problems that yielded a lower bound and we provided a unified analysis of three algorithms MD-UCBE, MD-SAR, and MD-APT. Our experiments suggest that by leveraging the geometry of the feasible arm identification problem, MD-SAR and MD-APT are able to dramatically outperform a uniform allocation approach.

2.9 Chapter Appendix Outline

In Section 3.10, we prove our lower bound for the feasible identification problem (Theorem 1). In Section 2.12, we prove the upper bound for MD-UCBE (Theorem 2). In Section 2.13, we prove the upper bound for MD-SAR (Theorem 3). In Section 2.14, we prove Proposition 1 and the upper bound for MD-APT (Theorem 4). In Section 2.15, we prove the key lemmas that unify our analyses of the three algorithms, namely, Lemmas 1, 2, and 3. In Section 6.11, we prove some useful technical lemmas. In Section 2.17, we extend our results to the feasible arm identification problem where P is convex. Finally, in Section 2.18, we present additional experimental results.

Regarding the lower and upper bound proofs (Theorems 1, 2, 3, and 4), we note that

we may assume that the realizations for each arm are drawn before the game has begun. Therefore, the empirical mean of an arm after t pulls is well-defined even if that arm has not been pulled t times.

2.10 Notation

Let T_i denote the number of pulls of arm i after T rounds. Let $X_{i,j,t}$ denote the tth realization of the jth coordinate of ν_i.

2.11 Lower Bound Proof

We note that the proof of Theorem 1 has some similarities to the proof of Theorem 1 of Locatelli et al. (2016a). The most important technical differences are *(i)* our novel lower bound construction with multidimensional distributions and *(ii)* our simple "chaining" argument that iteratively applies the well-known change-of-measure equation (6) in Audibert and Bubeck (2010) to relate $\mathcal{B}^0$ and $\mathcal{B}^i$.

Proof of Theorem 1. **Step 1: All of the problems have the same complexity.** The difference between problem $\mathcal{B}^0$ and $\mathcal{B}^i$ is the ith arm, i.e., the distributions ν_i and ν_i'. Since $\boldsymbol{\mu}_i \in P$ and $\boldsymbol{\mu}_i' \notin P$, by definition of H, it suffices to show that $\mathrm{dis}(\boldsymbol{\mu}_i, \partial P) = \mathrm{dis}(\boldsymbol{\mu}_i', P)$. By Lemma 13, there is $m \in [M]$ such that $\boldsymbol{\tau}_m \in \{\boldsymbol{x} : \boldsymbol{a}_m^t \boldsymbol{x} = b_m\}$ and $\boldsymbol{\tau}_m$ is the projection of $\boldsymbol{\mu}_i$ onto $\{\boldsymbol{x} : \boldsymbol{a}_m^t \boldsymbol{x} = b_m\}$. Let $\boldsymbol{\tau}_i'$ denote the projection of $\boldsymbol{\mu}_i'$ onto $\{\boldsymbol{x} : \boldsymbol{a}_m^t \boldsymbol{x} = b_m\}$. We claim that $\boldsymbol{\tau}_i = \boldsymbol{\tau}_i'$. Using the closed form solution of the projection of a vector onto a hyperplane (Boyd and Vandenberghe, 2004),

$$
\begin{aligned}
\boldsymbol{\tau}_i &= \boldsymbol{\mu}_i + (\boldsymbol{a}_m^t \boldsymbol{\mu}_i - b_m)\boldsymbol{a}_m, \\
\boldsymbol{\tau}_i' &= \boldsymbol{\mu}_i' + (b_m - \boldsymbol{a}_m^t \boldsymbol{\mu}_i')\boldsymbol{a}_m \\
&= \boldsymbol{\mu}_i + 2(\boldsymbol{\tau}_i - \boldsymbol{\mu}_i) + (b_m - \boldsymbol{a}_m^t(\boldsymbol{\mu}_i + 2(\boldsymbol{\tau}_i - \boldsymbol{\mu}_i))\boldsymbol{a}_m \\
&= 2\boldsymbol{\tau}_i - \boldsymbol{\mu}_i - (b_m - \boldsymbol{a}_m^t \boldsymbol{\mu}_i)\boldsymbol{a}_m \\
&= \boldsymbol{\tau}_i,
\end{aligned}
$$

establishing the claim.

Then,

$$\mathrm{dis}(\boldsymbol{\mu}_i', P) \geqslant \mathrm{dis}(\boldsymbol{\mu}_i', \{\boldsymbol{x} : \boldsymbol{a}_m^t \boldsymbol{x} \leqslant b_m\}) = \|\boldsymbol{\mu}_i' - \boldsymbol{\tau}_i\|_2 \geqslant \mathrm{dis}(\boldsymbol{\mu}_i', P)$$

where the last inequality follows since $\boldsymbol{\tau}_i \in P$. Thus,

$$\mathrm{dis}(\boldsymbol{\mu}_i', P) = \|\boldsymbol{\mu}_i' - \boldsymbol{\tau}_i\|_2 = \|\boldsymbol{\mu}_i - \boldsymbol{\tau}_i\|_2 = \mathrm{dis}(\boldsymbol{\mu}_i, \partial P).$$

Thus, $\mathcal{B}^0, \ldots, \mathcal{B}^K$ have the same problem complexity.

Step 2: Change of Measure For all $i \in [K]$, since $\mathrm{dis}(\boldsymbol{\mu}_i, \partial P) > \epsilon$, there exists $d_i > 0$ such that $\mathrm{dist}(\boldsymbol{\mu}_i, \partial P) = d_i + \epsilon$. We note that

$$\|\boldsymbol{\mu}_i' - \boldsymbol{\mu}_i\|_2 = 2\|\boldsymbol{\tau}_i - \boldsymbol{\mu}_i\|_2 = 2(d_i + \epsilon) = \Delta_i + d_i \leqslant 2\Delta_i. \tag{2.3}$$

Note that we can write ν_i as a product distribution $\nu_{i,1} \otimes \nu_{i,2} \otimes \ldots \otimes \nu_{i,D}$ where $\nu_{i,j} \sim N(\mu_{i,j}, 1)$ and $\nu_i' \coloneqq \nu_{i,1}' \otimes \nu_{i,2}' \otimes \ldots \otimes \nu_{i,D}'$ where $\nu_{i,j}' \sim N(\mu_{i,j}', 1)$. Let $l \leqslant D$ and define

$$\nu_i'^{,(l)} = \nu_{i,1}' \otimes \nu_{i,2}' \otimes \ldots \otimes \nu_{i,l-1}' \otimes \nu_{i,l}' \otimes \nu_{i,l+1} \otimes \ldots \otimes \nu_{i,D}.$$

Let $\mathcal{B}^{i,(l)}$ denote the product distribution

$$\nu_1 \otimes \ldots \otimes \nu_{i-1} \otimes \nu_i'^{,(l)} \otimes \nu_{i+1} \otimes \ldots \otimes \nu_K.$$

Define

$$\mathrm{KL}_{k,l} \coloneqq \mathrm{KL}(\nu_{k,l}', \nu_{k,l}) = \frac{1}{2}(\mu_{k,l}' - \mu_{k,l})^2$$

where we used the KL-divergence between two multivariate Gaussian random variables. Next, define for $1 \leqslant k \leqslant K$, $1 \leqslant l \leqslant D$, and $1 \leqslant t \leqslant T$,

$$\widehat{\mathrm{KL}}_{k,l,t} \coloneqq \frac{1}{t}\sum_{s=1}^{t} \log\left(\frac{d\nu_{k,l}'(X_{k,l,s})}{d\nu_{k,l}(X_{k,l,s})}\right) = \frac{1}{t}\sum_{s=1}^{t}\left[\frac{1}{2}(\mu_{k,l}^2 - (\mu_{k,l}')^2) + (\mu_{k,l}' - \mu_{k,l})X_{k,l,s}\right]$$

where we used the definition of the pdf of Gaussian random variables. Note that $\mathbb{E}_{\nu_{k,l}}\widehat{\mathrm{KL}}_{k,l,t} = \mathrm{KL}_{k,l}$ and that

$$\mathrm{Var}_{\nu_{k,l}}\left[\frac{1}{2}(\mu_{k,l}^2 - (\mu_{k,l}')^2) + (\mu_{k,l}' - \mu_{k,l})X_{k,l,s}\right] = (\mu_{k,l}' - \mu_{k,l})^2 \mathrm{Var}_{\nu_{k,l}}(X_{k,l,s}) = (\mu_{k,l}' - \mu_{k,l})^2.$$

Define the event

$$\Theta = \{\forall k \leqslant K, \forall t \leqslant T, \forall l \leqslant D : \widehat{\mathrm{KL}}_{k,l,t} - \mathrm{KL}_{k,l} \leqslant 2|\mu_{k,l} - \mu'_{k,l}|\sqrt{\frac{\log(4(\log(T)+1)KD)}{t}}\}.$$

Claim: $\mathbb{P}_{\mathcal{B}^0}(\Theta) \geqslant \frac{3}{4}$. Fix $k \leqslant K$ and $l \leqslant D$. $\widehat{\mathrm{KL}}_{k,l,t} - \mathrm{KL}_{k,l}$ is a sum of centered Gaussian random variables with variance $(\mu'_{k,l} - \mu_{k,l})^2$. Therefore, the sub-Gaussian norm of each term in the sum is $|\mu'_{k,l} - \mu_{k,l}|$. Let $u \in \{0, \ldots, \lceil \log(T) \rceil\}$. By Lemma 5,

$$\mathbb{P}_{\mathcal{B}^0}(\exists t \in [2^u, 2^{u+1}] : \widehat{\mathrm{KL}}_{k,l,t} - \mathrm{KL}_{k,l} \geqslant 2|\mu'_{k,l} - \mu_{k,l}|\sqrt{\frac{\log(4(\log(T)+1)KD)}{t}}) \leqslant \frac{1}{4(\log(T)+1)KD)}.$$

Then a union bound over $k \leqslant K$, $u \in \{0, \ldots, \lceil \log(T) \rceil\}, l \leqslant D$ yields that

$$\mathbb{P}_{\mathcal{B}^0}(\Theta^c) \leqslant \frac{1}{4} \tag{2.4}$$

establishing the claim.

Next, let $i \in [K]$ and define the event $\mathcal{A}_i = \{i \in \widehat{S}\}$. We lower bound $\mathbb{P}_{\mathcal{B}^i}(\mathcal{A}_i)$. Recall that T_i denotes the number of pulls of arm i after T rounds and let

$$t_i = \mathbb{E}_{\mathcal{B}^0} T_i.$$

and define the event

$$\Theta_i = \Theta \cap \mathcal{A}_i \cap \{T_i \leqslant 6t_i\}.$$

We use equation (6) from Audibert and Bubeck (2010), whose argument we briefly restate in the interest of making our chapter more self-contained. Let E denote an event. Then,

$$\begin{aligned}
&\mathbb{E}_{\mathcal{B}^{i,(D-1)}}[\mathbf{1}\{E\}\exp(-T_i\widehat{\mathrm{KL}}_{i,D,T_i})] = \mathbb{E}_{\mathcal{B}^{i,(D-1)}}[\mathbf{1}\{E\}\prod_{s=1}^{T_i}\frac{d\nu_{i,D}}{d\nu'_{i,D}}(X_{i,D,s})] \\
&= \int \ldots \int \mathbf{1}\{E\}\prod_{s=1}^{T_i}\frac{d\nu_{i,D}}{d\nu'_{i,D}}(X_{i,D,s})\Big[\prod_{k\neq i}\prod_{s=1}^{T_k} d\nu_{k,l}(X_{k,l,s})\Big]\Big[\prod_{s=1}^{T_i}\prod_{l\neq D} d\nu_{i,l}(X_{i,l,s})\Big]\prod_{s=1}^{T_i} d\nu'_{i,D}(X_{i,D,s}) \\
&= \mathbb{E}_{\mathcal{B}^i}[\mathbf{1}\{E\}].
\end{aligned} \tag{2.5}$$

We have the following series of inequalities:

$$\begin{aligned}
&\mathbb{P}_{\mathcal{B}^i}(\mathcal{A}_i) \geqslant \mathbb{P}_{\mathcal{B}^i}(\Theta_i) \\
&= \mathbb{E}_{\mathcal{B}^{i,(D-1)}}[\mathbf{1}\{\Theta_i\}\exp(-T_i\widehat{\mathrm{KL}}_{i,D,T_i})] && (2.6)\\
&\geqslant \mathbb{E}_{\mathcal{B}^{i,(D-1)}}[\mathbf{1}\{\Theta_i\}\exp(-T_i\frac{1}{2}(\mu_{i,D}-\mu'_{i,D})^2 - 2|\mu_{i,D}-\mu'_{i,D}|\sqrt{T_i\log(4(\log(T)+1)KD)})] && (2.7)\\
&\geqslant \Pr_{\mathcal{B}^{i,(D-1)}}(\Theta_i)\exp(-3t_i(\mu_{i,D}-\mu'_{i,D})^2 - 2|\mu_{i,D}-\mu'_{i,D}|\sqrt{6t_i\log(4(\log(T)+1)KD)})) && (2.8)
\end{aligned}$$

where equality (2.6) follows by equation (2.5), inequality (2.7) follows by Θ, and inequality (2.8) follows by $\{T_i \leqslant 6t_i\}$. Observe that we can repeat lines (2.6), (2.7), and (2.8) for $\Pr_{\mathcal{B}^{i,(D-1)}}(\Theta_i)$. Continuing in this manner for $l = 1, \ldots, D-1$ yields:

$$\begin{aligned}
&\mathbb{P}_{\mathcal{B}^i}(\mathcal{A}_i) && (2.9)\\
&\geqslant \Pr_{\mathcal{B}^{i,(D-1)}}(\Theta_i)\exp(-3t_i(\mu_{i,D}-\mu'_{i,D})^2 - 2|\mu_{i,D}-\mu'_{i,D}|\sqrt{6t_i\log(4(\log(T)+1)KD)})) \\
&\geqslant \Pr_{\mathcal{B}^0}(\Theta_i)\exp(-3t_i\sum_{l=1}^{D}(\mu_{i,l}-\mu'_{i,l})^2 - |\mu_{i,l}-\mu'_{i,l}|\sqrt{24t_i\log(4(\log(T)+1)KD)})) && (2.10)\\
&\geqslant \Pr_{\mathcal{B}^0}(\Theta_i)\exp(-12t_i\Delta_i^2 - \|\boldsymbol{\mu}_i-\boldsymbol{\mu}'_i\|_1\sqrt{24t_i\log(4(\log(T)+1)KD)})) && (2.11)\\
&\geqslant \Pr_{\mathcal{B}^0}(\Theta_i)\exp(-12t_i\Delta_i^2 - \|\boldsymbol{\mu}_i-\boldsymbol{\mu}'_i\|_2\sqrt{D}\sqrt{24t_i\log(4(\log(T)+1)KD)})) && (2.12)\\
&\geqslant \Pr_{\mathcal{B}^0}(\Theta_i)\exp(-12t_i\Delta_i^2 - 2\Delta_i\sqrt{24t_iD\log(4(\log(T)+1)KD)})) && (2.13)\\
&\geqslant \Pr_{\mathcal{B}^0}(\Theta_i)\exp(-13t_i\Delta_i^2 - 24D\log(4(\log(T)+1)KD))) && (2.14)
\end{aligned}$$

Line (2.11) follows by (2.3), line (2.12) follows by applying the inequality $\|\boldsymbol{x}\|_1 \leqslant \|\boldsymbol{x}\|_2\sqrt{D}$, line (2.13) follows by (2.3), and line (2.14) follows by the inequality $2ab \leqslant a^2 + b^2$ with $a = \Delta_i\sqrt{t_i}$.

Step 3: Lower bounding $\Pr_{\mathcal{B}^0}(\Theta_i)$**.** Suppose that for some i it holds that $\Pr_{\mathcal{B}^0}(\mathcal{A}_i) < \frac{1}{2}$. Then,

$$\Pr_{\mathcal{B}^0}(\cap_{k\in[K]}\mathcal{A}_k) \leqslant \Pr_{\mathcal{B}^0}(\mathcal{A}_i) < \frac{1}{2}.$$

Observe that under $\mathcal{B}^0$, the event $(\cap_{k\in[K]}\mathcal{A}_k)^c$ implies that $\mathcal{L}_{T,\epsilon}(\widehat{S}) = 1$ since for all $k \in [K]$, $\boldsymbol{\mu}_k \in P$ and $\mathrm{dis}(\boldsymbol{\mu}_k, \partial P) \geqslant \epsilon$. Thus, the theorem follows since

$$\max_{i\in\{0,\ldots,K\}}\mathbb{E}_{\mathcal{B}^i}(\mathcal{L}_{T,\epsilon}(\widehat{S})) \geqslant \mathbb{E}_{\mathcal{B}^0}(\mathcal{L}_{T,\epsilon}(\widehat{S})) \geqslant \Pr_{\mathcal{B}^0}((\cap_{k\in[K]}\mathcal{A}_k)^c) > \frac{1}{2}.$$

Therefore, we may suppose for the remainder of the proof that $\Pr_{\mathcal{B}^0}(\mathcal{A}_j) \geqslant \frac{1}{2}$ for all $j \in [K]$. Fix $i \in [K]$. By Markov's inequality,

$$\Pr_{\mathcal{B}^0}(T_i > 6t_i) \leqslant \frac{\mathbb{E}_{\mathcal{B}^0}[T_i]}{6t_i} = \frac{1}{6}.$$

Then, using the above two inequalities and inequality (2.4), by a union bound,

$$\Pr_{\mathcal{B}^0}(\Theta_i^c) \leqslant \frac{1}{4} + \frac{1}{2} + \frac{1}{6} = \frac{11}{12},$$

concluding this step of the proof.

Step 4: Putting it together.

$$\begin{aligned}
\max_{i \in \{1,\dots,K\}} \Pr_{\mathcal{B}^i}(\mathcal{A}_i) &\geqslant \frac{1}{K}\sum_{i=1}^{K} \Pr_{\mathcal{B}^i}(\mathcal{A}_i) \\
&\geqslant \frac{1}{K}\sum_{i=1}^{K} \Pr_{\mathcal{B}^0}(\Theta_i)\exp(-13t_i\Delta_i^2 - 24D\log(4(\log(T)+1)KD))) \qquad (2.15) \\
&\geqslant \frac{1}{12}\frac{1}{K}\sum_{i=1}^{K} \exp(-13t_i\Delta_i^2 - 24D\log(4(\log(T)+1)KD))) \qquad (2.16)
\end{aligned}$$

where in inequality (2.15) we used (2.14), in inequality (2.16) we used $\Pr_{\mathcal{B}^0}(\Theta_i) \geqslant \frac{1}{12}$. We claim that since $\sum_{i=1}^{K} t_i = T$, there exists some j such that $t_j \leqslant \frac{T}{H\Delta_j^2}$. Towards a contradiction, suppose that for all $i \in [K]$ $t_i > \frac{T}{H\Delta_i^2}$. Then,

$$T = \sum_{i \in [K]} t_i > \sum_{i \in [K]} \frac{T}{H\Delta_i^2} = T,$$

which is a contradiction. Then,

$$\begin{aligned}
&\frac{1}{K}\sum_{i=1}^{K} \exp(-13t_i\Delta_i^2 - 24D\log(4(\log(T)+1)KD))) \\
&\geqslant \exp(-13t_j\Delta_j^2 - 24D\log(4(\log(T)+1)KD)) - \log(K)) \\
&\geqslant \exp(-13\frac{T}{H} - 24D\log(4(\log(T)+1)KD)) - \log(K)).
\end{aligned}$$

Observe that under $\mathcal{B}^i$, the event $\mathcal{A}_i$ implies that $\mathcal{L}_{T,\epsilon}(\widehat{S}) = 1$ since $\text{dis}(\boldsymbol{\mu}'_i, P) > \epsilon$. Thus,

$$\begin{aligned}
\max_{i\in\{0,\dots,K\}} \mathbb{E}_{\mathcal{B}^i}(\mathcal{L}_{T,\epsilon}(\widehat{S})) &\geqslant \max_{i\in\{1,\dots,K\}} \mathbb{E}_{\mathcal{B}^i}(\mathcal{L}_{T,\epsilon}(\widehat{S})) \\
&\geqslant \max_{i\in\{1,\dots,K\}} \Pr_{\mathcal{B}^i}(\mathcal{A}_i) \\
&\geqslant \frac{1}{12}\exp(-13\frac{T}{H} - 24D\log(4(\log(T)+1)KD)) - \log(K)) \\
&\geqslant \exp(-13\frac{T}{H} - 25D\log(48(\log(T)+1)KD))).
\end{aligned}$$

□

2.12 MD-UCBE Upper Bound Proof

Proof of Theorem 2. **Step 1: Defining an appropriate event.**

Let $\mathcal{N}$ be a minimal $\frac{1}{2}$-net on $\mathcal{S}^{D-1}$. Let $\delta > 0$ (we choose it later). Define the event

$$\Xi = \{\forall i \in [K], \forall \boldsymbol{y} \in \mathcal{N}, \forall r \in \{1,\dots,T\} : |\boldsymbol{y}^t(\widehat{\boldsymbol{\mu}}_{i,r} - \boldsymbol{\mu}_i)| \leqslant \sqrt{\frac{a\delta^2}{4r}}\}.$$

By Lemma 2, on Ξ, for all $i \in [K]$ and for all $r \in [T]$,

$$\|\widehat{\boldsymbol{\mu}}_{i,r} - \boldsymbol{\mu}_i\|_2 \leqslant \sqrt{\frac{a\delta^2}{r}} \tag{2.17}$$

and

$$\Pr(\Xi) \geqslant 1 - 2(\log(T)+1)K5^D\exp(-a\frac{\delta^2}{16R^2}).$$

For the remainder of the proof, we suppose that Ξ holds.

Step 2: Lower bound the number of pulls for some arm.

Fix T. Recall that T_i denotes the number of pulls of arm i after T rounds. We claim that there exists an arm k such that it has been pulled after initialization and such that $T_k - 1 \geqslant \frac{T-K}{H\Delta_k^2}$ (for the remainder of the proof, let k denote one of these arms). If not, then we obtain the following contradiction.

$$T - K = \sum_{i=1}^{K}(T_i - 1) < \sum_{i=1}^{K}\frac{T-K}{H\Delta_i^2} = T - K.$$

For the remainder of the proof, let t denote the last time at which arm k was pulled. Then,

$$T_k(t) = T_k - 1 \geqslant \frac{T-K}{H\Delta_k^2}. \tag{2.18}$$

Step 3: Lower bound the number of pulls for each arm.

Lemma 1 and event Ξ imply that

$$|\widehat{\Delta}_{i,T_i(t)} - \Delta_i| \leqslant 2\sqrt{\frac{a\delta^2}{T_i(t)}} \tag{2.19}$$

for all $i \in [K]$. We choose $\delta = \frac{1}{10}$.

Arm k was pulled at time t, so that we have for all $i \in [K]$,

$$\widehat{\Delta}_{k,T_k(t)} - \sqrt{\frac{a}{T_k(t)}} \leqslant \widehat{\Delta}_{i,T_i(t)} - \sqrt{\frac{a}{T_i(t)}}. \tag{2.20}$$

Now,

$$\begin{aligned}
\Delta_k + \sqrt{\frac{a}{T_i(t)}} &\leqslant \widehat{\Delta}_{k,T_k(t)} + \sqrt{\frac{a}{T_i(t)}} + \frac{1}{5}\sqrt{\frac{a}{T_k(t)}} && (2.21)\\
&\leqslant \frac{6}{5}\sqrt{\frac{a}{T_k(t)}} + \widehat{\Delta}_{i,T_i(t)} && (2.22)
\end{aligned}$$

where in inequality (2.21) we apply (2.19) and in inequality (2.22) we apply (2.20).

Rearranging (2.22), we obtain

$$\begin{aligned}
\sqrt{\frac{a}{T_i(t)}} &\leqslant \frac{6}{5}\sqrt{\frac{a}{T_k(t)}} - \Delta_k + \widehat{\Delta}_{i,T_i(t)}\\
&\leqslant \Delta_k - \Delta_k + \widehat{\Delta}_{i,T_i(t)} && (2.23)\\
&\leqslant \Delta_i + \frac{1}{5}\sqrt{\frac{a}{T_i(t)}}. && (2.24)
\end{aligned}$$

where inequality (2.23) follows by (2.18) and $0 \leqslant a \leqslant \frac{25}{36}\frac{T-K}{H}$ and inequality (2.24) follows by (2.19).

Rearranging (2.24) implies that

$$\frac{4}{5}\sqrt{\frac{a}{T_i(t)}} \leqslant \Delta_i.$$

Thus,

$$\frac{1}{5}\sqrt{\frac{a}{T_i}} \leqslant \frac{1}{5}\sqrt{\frac{a}{T_i(t)}} \leqslant \frac{\Delta_i}{4} < \frac{\Delta_i}{2} \tag{2.25}$$

Step 4: Putting it together

Combining (2.25), event Ξ, and (2.17) yields for all $i \in [K]$,

$$\left\|\widehat{\boldsymbol{\mu}}_{i,T_i(T+1)} - \boldsymbol{\mu}_i\right\|_2 < \frac{\Delta_i}{2}.$$

Then, by Lemma 3, it follows that $\mathcal{L}_{T,\epsilon}(\widehat{S}) = 0$. □

2.13 MD-SAR Upper Bound Proof

As in Algorithm 2, define $\overline{\log}(x) = \frac{1}{2} + \sum_{i=2}^{x} \frac{1}{i}$.

Proof of Theorem 3. **Step 1: Defining an event and bounding probability** Let $\mathcal{N}$ be a minimal $\frac{1}{2}$-net on $\mathcal{S}^{D-1}$. Let $\delta > 0$ (we choose it later). Define the events

$$\Xi_1 = \{\forall i \in [K], \forall \boldsymbol{y} \in \mathcal{N}, \forall r \in \{1, \ldots, T\} : |\boldsymbol{y}^t(\widehat{\boldsymbol{\mu}}_{i,r} - \boldsymbol{\mu}_i)| \leqslant \sqrt{\frac{(T-K)\delta^2}{\overline{\log}(K) H_2 r}}\},$$

$$\Xi_2 = \{\forall k \in [K-1], \forall l \in \{(K), \ldots, (K+1-k)\} \text{ and } j \in [K] \text{ s.t. } 2\Delta_j < \Delta_l : \widehat{\Delta}_{l,n_k} - \widehat{\Delta}_{j,n_k} > 0\}.$$

Then, by Lemma 2,

$$\begin{aligned}
\Pr(\Xi_1) &\geqslant 1 - 2(\log(T)+1)K5^D \exp(-\frac{T-K}{16\overline{\log}(K) H_2}\frac{\delta^2}{R^2}) \\
&\geqslant 1 - 2(\log(T)+1)K5^D \exp(-\frac{T-K}{16\log(2K) H_2}\frac{\delta^2}{R^2}) \qquad (2.26)
\end{aligned}$$

where line (2.26) follows by $\log(K+1) - \frac{1}{2} \leqslant \overline{\log}(K) \leqslant \log(K) + \frac{1}{2} \leqslant \log(2K)$ (Audibert and Bubeck, 2010).

Next, we bound $\Pr(\Xi_2)$. By a union bound,

$$\Pr(\Xi_2^c) \leqslant \sum_{k \in [K-1]} \sum_{l \in \{(K), \ldots, (K+1-k)\}, j : 2\Delta_j < \Delta_l} \Pr(\widehat{\Delta}_{j,n_k} - \widehat{\Delta}_{l,n_k} \geqslant 0).$$

Fix a round $k \in \{1, \ldots, K-1\}$, let $l \in \{(K), \ldots, (K+1-k)\}$, and let $j \in [K]$ such that

$2\Delta_j < \Delta_l$. Then,

$$\begin{aligned}
\Pr(\widehat{\Delta}_{j,n_k} - \widehat{\Delta}_{l,n_k} \geqslant 0) &= \Pr((\widehat{\Delta}_{j,n_k} - \Delta_j) - (\widehat{\Delta}_{l,n_k} - \Delta_l) \geqslant \Delta_l - \Delta_j) \\
&\leqslant \Pr((\widehat{\Delta}_{j,n_k} - \Delta_j) - (\widehat{\Delta}_{l,n_k} - \Delta_l) > \frac{1}{2}\Delta_l) \\
&\leqslant \Pr(|(\widehat{\Delta}_{j,n_k} - \Delta_j) - (\widehat{\Delta}_{l,n_k} - \Delta_l)| > \frac{1}{2}\Delta_l) \\
&\leqslant \Pr(|\widehat{\Delta}_{j,n_k} - \Delta_j| + |\widehat{\Delta}_{l,n_k} - \Delta_l| > \frac{1}{2}\Delta_l) \\
&\leqslant \Pr(|\widehat{\Delta}_{j,n_k} - \Delta_j| > \frac{1}{4}\Delta_l) + \Pr(|\widehat{\Delta}_{l,n_k} - \Delta_l| > \frac{1}{4}\Delta_l)
\end{aligned}$$

Define the event

$$\Sigma_i = \{\forall \boldsymbol{y} \in \mathcal{N}, : |\boldsymbol{y}^t(\widehat{\boldsymbol{\mu}}_{i,n_k} - \boldsymbol{\mu}_i)| \leqslant \frac{1}{16}\Delta_l\}.$$

Under Σ_i, Lemma 14 implies that

$$\|\widehat{\boldsymbol{\mu}}_{i,n_k} - \boldsymbol{\mu}_i\|_2 \leqslant 2 \sup_{\boldsymbol{y} \in \mathcal{N}} \boldsymbol{y}^t(\widehat{\boldsymbol{\mu}}_{i,n_k} - \boldsymbol{\mu}_i) \leqslant \frac{1}{8}\Delta_l.$$

Thus, by Lemma 1, Σ_i implies that $|\widehat{\Delta}_i(n_k) - \Delta_i| \leqslant \frac{1}{4}\Delta_l$. Using the contrapositive of this implication,

$$\begin{aligned}
\Pr(|\widehat{\Delta}_{j,n_k} - \Delta_j| > \frac{1}{4}\Delta_l) + \Pr(|\widehat{\Delta}_{l,n_k} - \Delta_l| > \frac{1}{4}\Delta_l) &\leqslant \Pr(\Sigma_j^c) + \Pr(\Sigma_l^c) \\
&\leqslant \sum_{\boldsymbol{y} \in \mathcal{N}} [\Pr(|\boldsymbol{y}^t(\widehat{\boldsymbol{\mu}}_{l,n_k} - \boldsymbol{\mu}_l)| > \frac{1}{16}\Delta_l) \\
&\quad + \Pr(|\boldsymbol{y}^t(\widehat{\boldsymbol{\mu}}_{j,n_k} - \boldsymbol{\mu}_j)| > \frac{1}{16}\Delta_l)] \\
&\leqslant 4 \cdot 5^D \exp(-\frac{n_k \Delta_l^2}{512R^2}) \qquad (2.27) \\
&\leqslant 4 \cdot 5^D \exp(-\frac{n_k \Delta_{(K+1-k)}^2}{512R^2}).
\end{aligned}$$

where line (2.27) follows by Lemma 8 and since $\mathcal{N}$ is a $\frac{1}{2}$-net by construction, we have $|\mathcal{N}| \leqslant 5^D$ by Lemma 15. Then,

$$\begin{aligned}
\Pr(\Xi_2^c) &\leqslant \sum_{k \in [K-1]} \sum_{l \in \{(K), \ldots, (K+1-k)\}, j : 2\Delta_j < \Delta_l} 4 \cdot 5^D \exp(-\frac{n_k \Delta_{(K+1-k)}^2}{512R^2}) \\
&\leqslant \sum_{k \in [K-1]} kK^2 4 \cdot 5^D \exp(-\frac{n_k \Delta_{(K+1-k)}^2}{512R^2}) \\
&\leqslant 4K^3 5^D \exp(-\frac{T-K}{512R^2 H_2})
\end{aligned}$$

where we used the fact that

$$n_k \Delta_{(K+1-k)}^2 \geqslant \frac{T-K}{\overline{\log}(K)(K+1-k)\Delta_{(K+1-k)}^{-2}} \geqslant \frac{T-K}{\overline{\log}(K)H_2}.$$

For the remainder of the proof, we suppose $\Xi_1 \cap \Xi_2$ holds.

Step 2: Lower bounding the number of pulls This step of the proof is similar to the proof of Theorem 3 in (Audibert and Bubeck, 2010); we repeat it for the sake of completeness. Consider phase k. At least one of the arms $l \in \{(K), \ldots, (K+1-k)\}$ has not been eliminated. Then, by Ξ_2, we have that $\widehat{\Delta}_{l,n_k} > \widehat{\Delta}_{j,n_k}$ for any arm j satisfying $2\Delta_j < \Delta_{(K+1-k)}$. Thus, at the end of phase k, MD-SAR does not eliminate any arm j such that $2\Delta_j < \Delta_{(K+1-k)}$.

Now, fix an arm j. Recall that T_j denotes the number of pulls of arm j after T rounds. We consider two distinct cases: *(i)* there exists $m \in [K]$ such that $\Delta_{(m-1)} \leqslant 2\Delta_j < \Delta_{(m)}$ and *(ii)* there exists no such $m \in [K]$. Suppose *(i)* holds. Since $2\Delta_j < \Delta_{(m)}$, the arm j is eliminated some time after the $K+2-m$ phase so that

$$\Delta_j^2 T_j \geqslant \Delta_j^2 n_{K+2-m} = \frac{\Delta_j^2}{\Delta_{(m-1)}^2} \frac{T-K}{\overline{\log}(K)(m-1)\Delta_{(m-1)}^{-2}} \geqslant \frac{\Delta_j^2}{\Delta_{(m-1)}^2} \frac{T-K}{\overline{\log}(K)H_2} \geqslant \frac{T-K}{4\overline{\log}(K)H_2}.$$

Next, suppose *(ii)* holds. Then, $2\Delta_j \geqslant \Delta_{(K)}$, so that

$$\Delta_j^2 T_j \geqslant \frac{1}{4}\Delta_{(K)}^2 n_1 = \frac{T-K}{4\overline{\log}(K)K\Delta_{(K)}^{-2}} \geqslant \frac{T-K}{4\overline{\log}(K)H_2}.$$

Thus, we have that for all $j \in [K]$,

$$T_j \geqslant \frac{T-K}{4\overline{\log}(K)H_2\Delta_j^2}. \tag{2.28}$$

Step 3: Putting it together. Using Lemma 2, Ξ_1, and (2.28), we have that for all $i \in [K]$,

$$\left\|\widehat{\boldsymbol{\mu}}_{i,T_i(T+1)} - \boldsymbol{\mu}_i\right\|_2 \leqslant 2\sqrt{\frac{(T-K)\delta^2}{\overline{\log}(K)H_2T_i}} \leqslant 4\delta\Delta_i.$$

We choose $\delta = \frac{1}{9}$. Then, by Lemma 3, it follows that $\mathcal{L}_{T,\epsilon}(\widehat{S}) = 0$.

□

2.14 MD-APT Upper Bound Proof

Proof of Proposition 1. Let $\mathcal{N}$ be a minimal $\frac{1}{2}$-net of $\mathcal{S}^{D-1}$. By Lemma 15, $|\mathcal{N}| \leqslant 5^D$. Then,

$$\begin{aligned}
\mathbb{E}[\mathcal{L}_{T,\epsilon}(\widehat{S})] &\leqslant \Pr(\exists i : \widehat{\boldsymbol{\mu}}_{i,T_i(T+1)} \in P \text{ and } \boldsymbol{\mu}_i \notin P \\
&\qquad \text{or } \widehat{\boldsymbol{\mu}}_{i,T_i(T+1)} \notin P \text{ and } \boldsymbol{\mu}_i \in P) \\
&\leqslant \Pr(\exists i : \left\|\widehat{\boldsymbol{\mu}}_{i,T_i(T+1)} - \boldsymbol{\mu}_i\right\|_2 > \Delta_i) \\
&\leqslant \sum_{i\in[K]} \Pr(\left\|\widehat{\boldsymbol{\mu}}_{i,T_i(T+1)} - \boldsymbol{\mu}_i\right\|_2 > \Delta_i) \\
&\leqslant \sum_{i\in[K]} \sum_{\boldsymbol{y}\in\mathcal{N}} \Pr(|\boldsymbol{y}^t(\widehat{\boldsymbol{\mu}}_{i,T_i(T+1)} - \boldsymbol{\mu}_i)| > \frac{\Delta_i}{2}) \qquad (2.29) \\
&\leqslant \sum_{i\in[K]} \sum_{y\in\mathcal{N}} 2\exp(-\frac{1}{8}\frac{\Delta_i^2 T_i}{R^2}) \qquad (2.30) \\
&= 2K5^D \exp(-\frac{1}{8}\frac{T}{HR^2}).
\end{aligned}$$

Line (2.29) follows since by Lemma 14, if $\left\|\widehat{\boldsymbol{\mu}}_{i,T_i(T+1)} - \boldsymbol{\mu}_i\right\|_2 > \Delta_i$, then there exists $\boldsymbol{y} \in \mathcal{N}$ such that $|\boldsymbol{y}^t(\widehat{\boldsymbol{\mu}}_{i,T_i(T+1)} - \boldsymbol{\mu}_i)| > \frac{\Delta_i}{2}$. Line (2.30) follows by Lemma 8. □

Proof of Theorem 4. **Step 1: Defining an appropriate event.**

Let $\mathcal{N}$ be a minimal $\frac{1}{2}$-net on $\mathcal{S}^{D-1}$. Let $\delta > 0$ (we choose it later). Define the event

$$\Xi = \{\forall i \in [K], \forall \boldsymbol{y} \in \mathcal{N}, \forall r \in \{1,\dots,T\} : |\boldsymbol{y}^t(\widehat{\boldsymbol{\mu}}_{i.r} - \boldsymbol{\mu}_i)| \leqslant \sqrt{\frac{T\delta^2}{4Hr}}\}.$$

By Lemma 2, on Ξ, for all $i \in [K]$ and for all $r \in [T]$

$$\|\widehat{\boldsymbol{\mu}}_{i.r} - \boldsymbol{\mu}_i\|_2 \leqslant \sqrt{\frac{T\delta^2}{Hr}}. \qquad (2.31)$$

and

$$\mathbb{P}(\Xi) \geqslant 1 - 2(\log(T)+1)K5^D \exp(-T\frac{\delta^2}{16R^2H})$$

For the remainder of the proof, we suppose that Ξ holds.

Step 2: Lower bound the number of pulls for some arm.

Fix T. Recall that T_i denotes the number of pulls of arm i after T rounds. We claim that there exists an arm k that has been pulled after initialization and such that $T_k - 1 \geqslant \frac{T-K}{H\Delta_k^2}$

(for the remainder of the proof, let k denote one of these arms). If not, then we obtain the following contradiction:

$$T - K = \sum_{i=1}^{K}(T_i - 1) < \sum_{i=1}^{K} \frac{T-K}{H\Delta_i^2} = T - K.$$

Since $T \geqslant 2K$, $T_k - 1 \geqslant \frac{T}{2H\Delta_k^2}$.

For the remainder of the proof, let $t \leqslant T$ denote the last round that arm k was pulled. Then,

$$T_k(t) = T_k - 1 \geqslant \frac{T}{2H\Delta_k^2}. \tag{2.32}$$

Step 3: Lower bound the number of pulls for each arm.

Lemma 1 and event Ξ imply that

$$|\widehat{\Delta}_{i,T_i(t)} - \Delta_i| \leqslant 2\sqrt{\frac{T\delta^2}{HT_i(t)}} \tag{2.33}$$

for all $i \in [K]$.

At time t, we pulled arm k, so that for all $i \in [K]$, $\widehat{\Delta}_{k,T_k(t)}\sqrt{T_k(t)} \leqslant \widehat{\Delta}_{i,T_i(t)}\sqrt{T_i(t)}$. Then, using (2.32) and (2.33),

$$\begin{aligned}\widehat{\Delta}_{k,T_k(t)}\sqrt{T_k(t)} &\geqslant (\Delta_k - 2\sqrt{\frac{T\delta^2}{HT_k(t)}})\sqrt{T_k(t)} \\ &\geqslant (\Delta_k - 2\sqrt{2\delta^2\Delta_k^2})\sqrt{T_k(t)}. \end{aligned} \tag{2.34}$$

We require that $\delta < \frac{1}{2\sqrt{2}}$ so that (2.34) is positive. Thus, we can apply (2.34) and (2.32) to obtain that

$$\widehat{\Delta}_{k,T_k(t)}\sqrt{T_k(t)} \geqslant (\frac{1}{\sqrt{2}} - 2\delta)\sqrt{\frac{T}{H}}. \tag{2.35}$$

Next, applying (2.33), we obtain

$$\begin{aligned}\widehat{\Delta}_{i,T_i(t)}\sqrt{T_i(t)} &= \widehat{\Delta}_{i,T_i(t)}\sqrt{T_i(t)} \\ &\leqslant (\Delta_i + 2\sqrt{\frac{T\delta^2}{HT_i(t)}})\sqrt{T_i(t)} \\ &\leqslant \Delta_i\sqrt{T_i(t)} + 2\delta\sqrt{\frac{T}{H}}. \end{aligned} \tag{2.36}$$

Combining inequalities $T_i(t) \leqslant T_i$, (2.36), $\widehat{\Delta}_{k,T_k(t)}\sqrt{T_k(t)} \leqslant \widehat{\Delta}_{i,T_i(t)}\sqrt{T_i(t)}$, and (2.35) yields that

$$\begin{aligned}\Delta_i\sqrt{T_i} + 2\delta\sqrt{\frac{T}{H}} &\geqslant \Delta_i\sqrt{T_i(t)} + 2\delta\sqrt{\frac{T}{H}} \\ &\geqslant \widehat{\Delta}_{i,T_i(t)}\sqrt{T_i(t)} \\ &\geqslant \widehat{\Delta}_{k,T_k(t)}\sqrt{T_k(t)} \\ &\geqslant (\frac{1}{\sqrt{2}} - 2\delta)\sqrt{\frac{T}{H}}. \end{aligned} \tag{2.37}$$

Rearranging (2.37) yields for all $i \in [K]$

$$(1 - 4\sqrt{2}\delta)^2\frac{T}{2H\Delta_i^2} \leqslant T_i. \tag{2.38}$$

Step 4: Putting it together.

Combining (2.38) with (2.31) and Ξ respectively, we obtain

$$\left\|\widehat{\boldsymbol{\mu}}_{i,T_i(T+1)} - \boldsymbol{\mu}_i\right\|_2 \leqslant \Delta_i\delta(1 - 4\sqrt{2}\delta).$$

We choose $\delta = \frac{\sqrt{2}}{9}$. Thus, by Lemma 3, $\mathcal{L}_{T,\epsilon}(\widehat{S}) = 0$. □

2.15 Key Lemmas

In this section, we prove the Lemmas of Section 6.5, namely, Lemmas 1, 2, and 3.

Proof of Lemma 1. For the sake of brevity, we write $\widehat{\boldsymbol{\mu}}_i$ instead of $\widehat{\boldsymbol{\mu}}_{i,t}$ and $\widehat{\Delta}_i$ instead of $\widehat{\Delta}_{i,t}$. We separate the analysis into 4 cases.

Case 1: $A\boldsymbol{\mu}_i \leqslant \boldsymbol{b}$ and $A\widehat{\boldsymbol{\mu}}_i \leqslant \boldsymbol{b}$.

Let j be such that $\Delta_i = |b_j - \boldsymbol{a}_j^t\boldsymbol{\mu}_i| + \epsilon$ and let $\widehat{j}$ such that $\widehat{\Delta}_i = |b_{\widehat{j}} - \boldsymbol{a}_{\widehat{j}}^t\widehat{\boldsymbol{\mu}}_i| + \epsilon$. Then, by definition of j and $\widehat{j}$,

$$0 \leqslant b_j - \boldsymbol{a}_j^t\boldsymbol{\mu}_i \leqslant b_{\widehat{j}} - \boldsymbol{a}_{\widehat{j}}^t\boldsymbol{\mu}_i \tag{2.39}$$

$$0 \leqslant b_{\widehat{j}} - \boldsymbol{a}_{\widehat{j}}^t\widehat{\boldsymbol{\mu}}_i \leqslant b_j - \boldsymbol{a}_j^t\widehat{\boldsymbol{\mu}}_i. \tag{2.40}$$

Note that it suffices to bound

$$|\Delta_i - \widehat{\Delta}_i| = ||b_j - \boldsymbol{a}_j^t\boldsymbol{\mu}_i| - |b_{\widehat{j}} - \boldsymbol{a}_{\widehat{j}}^t\widehat{\boldsymbol{\mu}}_i||.$$

Then,

$$(b_j - \boldsymbol{a}_j^t \boldsymbol{\mu}_i) - (b_{\widehat{j}} - \boldsymbol{a}_{\widehat{j}}^t \widehat{\boldsymbol{\mu}}_i) \leqslant (b_{\widehat{j}} - \boldsymbol{a}_{\widehat{j}}^t \boldsymbol{\mu}_i) - (b_{\widehat{j}} - \boldsymbol{a}_{\widehat{j}}^t \widehat{\boldsymbol{\mu}}_i) \tag{2.41}$$

$$= \boldsymbol{a}_{\widehat{j}}^t (\widehat{\boldsymbol{\mu}}_i - \boldsymbol{\mu}_i) \tag{2.42}$$

$$\leqslant \gamma. \tag{2.43}$$

where line (2.41) used line (2.39) and line (2.43) follows by Cauchy-Schwarz inequality, $\|\boldsymbol{a}_j\|_2 = 1$, and the hypothesis. Next,

$$(b_{\widehat{j}} - \boldsymbol{a}_{\widehat{j}}^t \widehat{\boldsymbol{\mu}}_i) - (b_j - \boldsymbol{a}_j^t \boldsymbol{\mu}_i) \leqslant (b_j - \boldsymbol{a}_j^t \widehat{\boldsymbol{\mu}}_i) - (b_j - \boldsymbol{a}_j^t \boldsymbol{\mu}_i) \tag{2.44}$$

$$= \boldsymbol{a}_j^t (\boldsymbol{\mu}_i - \widehat{\boldsymbol{\mu}}_i)$$

$$\leqslant \gamma. \tag{2.45}$$

where line (2.44) used line (2.40) and line (2.45) follows by Cauchy-Schwarz inequality, $\|\boldsymbol{a}_j\|_2 = 1$, and the hypothesis.

Case 2: $A\boldsymbol{\mu}_i \not\leqslant \boldsymbol{b}$ and $A\widehat{\boldsymbol{\mu}}_i \not\leqslant \boldsymbol{b}$. Then,

$$\begin{aligned}
|\Delta_i - \widehat{\Delta}_i| &= |\, \|\boldsymbol{\mu}_i - \mathrm{Proj}_P(\boldsymbol{\mu}_i)\|_2 - \|\widehat{\boldsymbol{\mu}}_i - \mathrm{Proj}_P(\widehat{\boldsymbol{\mu}}_i)\|_2 \,| \\
&\leqslant \|(\boldsymbol{\mu}_i - \widehat{\boldsymbol{\mu}}_i) - (\mathrm{Proj}_P(\boldsymbol{\mu}_i) - \mathrm{Proj}_P(\widehat{\boldsymbol{\mu}}_i))\|_2 && (2.46) \\
&\leqslant \|\boldsymbol{\mu}_i - \widehat{\boldsymbol{\mu}}_i\|_2 + \|\mathrm{Proj}_P(\boldsymbol{\mu}_i) - \mathrm{Proj}_P(\widehat{\boldsymbol{\mu}}_i)\|_2 && (2.47) \\
&\leqslant 2\,\|\boldsymbol{\mu}_i - \widehat{\boldsymbol{\mu}}_i\|_2 && (2.48) \\
&\leqslant 2\gamma
\end{aligned}$$

where line (2.46) used the reverse triangle inequality, (2.47) used the triangle inequality, and (2.48) used the fact that projection onto a convex set is contractive (Proposition 2.2.1 (Bertsekas, 2009)).

Case 3: $A\boldsymbol{\mu}_i \not\leqslant \boldsymbol{b}$ and $A\widehat{\boldsymbol{\mu}}_i \leqslant \boldsymbol{b}$.

We claim that $\mathrm{dis}(\widehat{\boldsymbol{\mu}}_i, \partial P) \leqslant \|\widehat{\boldsymbol{\mu}}_i - \boldsymbol{\mu}_i\|_2$. Suppose not. Then, since $\widehat{\boldsymbol{\mu}}_i \in P$ and $\boldsymbol{\mu}_i \notin P$, there exists $\theta \in [0, 1]$ such that $\boldsymbol{z} = \theta\boldsymbol{\mu}_i + (1-\theta)\widehat{\boldsymbol{\mu}}_i \in \partial P$. Then,

$$\|\widehat{\boldsymbol{\mu}}_i - \boldsymbol{z}\|_2 \leqslant \|\widehat{\boldsymbol{\mu}}_i - \boldsymbol{\mu}_i\|_2 < \mathrm{dis}(\widehat{\boldsymbol{\mu}}_i, \partial P),$$

which is a contradiction. Thus, the claim follows. Then,

$$\mathrm{dis}(\widehat{\boldsymbol{\mu}}_i, \partial P) \leqslant \|\widehat{\boldsymbol{\mu}}_i - \boldsymbol{\mu}_i\|_2 \leqslant \gamma.$$

Next, since $\widehat{\boldsymbol{\mu}}_i \in P$,

$$\operatorname{dis}(\boldsymbol{\mu}_i, P) \leqslant \|\boldsymbol{\mu}_i - \widehat{\boldsymbol{\mu}}_i\|_2 \leqslant \gamma.$$

Thus,

$$|\Delta_i - \widehat{\Delta}_i| = |\operatorname{dis}(\widehat{\boldsymbol{\mu}}_i, \partial P) - \operatorname{dis}(\boldsymbol{\mu}_i, P)| \leqslant \max(\operatorname{dis}(\widehat{\boldsymbol{\mu}}_i, \partial P), \operatorname{dis}(\boldsymbol{\mu}_i, P)) \leqslant \gamma.$$

Case 4: $A\boldsymbol{\mu}_i \leqslant \boldsymbol{b}$ and $A\widehat{\boldsymbol{\mu}}_i \not\leqslant \boldsymbol{b}$. This case is similar to case 3. Since $\boldsymbol{\mu}_i \in P$ and $\widehat{\boldsymbol{\mu}}_i \notin P$,

$$\operatorname{dis}(\widehat{\boldsymbol{\mu}}_i, P) \leqslant \|\widehat{\boldsymbol{\mu}}_i - \boldsymbol{\mu}_i\|_2 \leqslant \gamma.$$

Next, since $\boldsymbol{\mu}_i \in P$ and $\widehat{\boldsymbol{\mu}}_i \notin P$, by a similar argument used in case 3,

$$\begin{aligned}\operatorname{dis}(\boldsymbol{\mu}_i, \partial P) &\leqslant \|\boldsymbol{\mu}_i - \widehat{\boldsymbol{\mu}}_i\|_2 \\ &\leqslant \gamma.\end{aligned}$$

Thus,

$$|\Delta_i - \widehat{\Delta}_i| = |\operatorname{dis}(\widehat{\boldsymbol{\mu}}_i, \partial P) - \operatorname{dis}(\boldsymbol{\mu}_i, P)| \leqslant \max(\operatorname{dis}(\widehat{\boldsymbol{\mu}}_i, \partial P), \operatorname{dis}(\boldsymbol{\mu}_i, P)) \leqslant \gamma.$$

□

Proof of Lemma 2. First, we bound the norm of $\widehat{\boldsymbol{\mu}}_{i,r} - \boldsymbol{\mu}_i$ on Ξ for all $i \in [K]$ and for all $r \in [T]$. Fix $i \in [K]$ and $r \in [T]$. Recall that $\mathcal{N}$ is a minimal $\frac{1}{2}$-net. Using the event Ξ and Lemma 14,

$$\|\widehat{\boldsymbol{\mu}}_{i,r} - \boldsymbol{\mu}_i\|_2 \leqslant 2 \sup_{\boldsymbol{y} \in \mathcal{N}} \boldsymbol{y}^t(\widehat{\boldsymbol{\mu}}_{i,r} - \boldsymbol{\mu}_i) \leqslant \sqrt{\frac{\omega^2}{r}}.$$

Next, we give the probability bound. Since ν_i is R-sub-Gaussian, by definition, we have that if $\boldsymbol{X} \sim \nu_i$, then

$$\sup_{\boldsymbol{y} \in \mathcal{N}} \left\|\boldsymbol{X}^t \boldsymbol{y}\right\|_{\psi_2} \leqslant \sup_{\boldsymbol{y} \in \mathcal{S}^{D-1}} \left\|\boldsymbol{X}^t \boldsymbol{y}\right\|_{\psi_2} = \|\nu_i\|_{\psi_2} \leqslant R.$$

Thus, by Lemma 5 and a union bound, for each $i \in [K]$, $\boldsymbol{y} \in \mathcal{N}$, and $u \in \{0, \ldots, \lfloor \log(T) \rfloor\}$:

$$\Pr(\exists v \in [2^u, 2^{u+1}], |\boldsymbol{y}^t(\widehat{\boldsymbol{\mu}}_{i,v} - \boldsymbol{\mu}_i)| \geqslant \sqrt{\frac{\omega^2}{4v}}) \leqslant 2\exp(-\frac{\omega^2}{16R^2}).$$

Taking a union bound over all $i \in [K]$, $\boldsymbol{y} \in \mathcal{N}$, and $u \in \{0, \ldots, \lfloor \log(T) \rfloor\}$ yields

$$\begin{aligned}\Pr(\Xi) &\geqslant 1 - 2(\log(T)+1)K|\mathcal{N}| \exp(-\frac{\omega^2}{16R^2}) \\ &\geqslant 1 - 2(\log(T)+1)K5^D \exp(-\frac{\omega^2}{16R^2})\end{aligned}$$

where in the last line we used $|\mathcal{N}| \leqslant 5^D$ by Lemma 15. □

Proof of Lemma 3. Fix $i \in [K]$. For the sake of brevity, we write $\widehat{\boldsymbol{\mu}}_i$ instead of $\widehat{\boldsymbol{\mu}}_{i,t}$. First, suppose $A\boldsymbol{\mu}_i \leqslant \boldsymbol{b} - \epsilon\mathbf{1}$. Fix any $j \in [M]$. Then,

$$\begin{aligned}\boldsymbol{a}_j^t \widehat{\boldsymbol{\mu}}_i - b_j &= \boldsymbol{a}_j^t(\widehat{\boldsymbol{\mu}}_i - \boldsymbol{\mu}_i) + \boldsymbol{a}_j^t\boldsymbol{\mu}_i - b_j \\ &< \frac{\Delta_i}{2} + \boldsymbol{a}_j^t\boldsymbol{\mu}_i - b_j \qquad (2.49) \\ &\leqslant \frac{1}{2}(b_j - \boldsymbol{a}_j^t\boldsymbol{\mu}_i + \epsilon) + \boldsymbol{a}_j^t\boldsymbol{\mu}_i - b_j \\ &\leqslant \frac{1}{2}(\boldsymbol{a}_j^t\boldsymbol{\mu}_i - b_j + \epsilon) \\ &\leqslant \frac{1}{2}(-\epsilon + \epsilon) \\ &= 0\end{aligned}$$

where line (2.49) follows by the Cauchy-Schwarz inequality, $\|\boldsymbol{a}_j\|_2 = 1$, and the hypothesis.

Next, suppose $\operatorname{dis}(\boldsymbol{\mu}_i, P) \geqslant \epsilon$. Then,

$$\|\widehat{\boldsymbol{\mu}}_i - \boldsymbol{\mu}_i\|_2 < \frac{\Delta_i}{2} = \frac{1}{2}(\operatorname{dis}(\boldsymbol{\mu}_i, P) + \epsilon) \leqslant \operatorname{dis}(\boldsymbol{\mu}_i, P).$$

Thus, $\widehat{\boldsymbol{\mu}}_i \notin P$ since otherwise we have a contradiction.

□

2.16 Technical Lemmas

Lemma 4. *Let $P = \{\boldsymbol{x} \in \mathbb{R}^D : A\boldsymbol{x} \leqslant \boldsymbol{b}\}$ with $A \in \mathbb{R}^{M \times D}$. Let $\boldsymbol{\mu} \in P$. Then,*

$$\operatorname{dis}(\boldsymbol{\mu}, \partial P) = \min_{i=1,\ldots,M} \operatorname{dis}(\boldsymbol{\mu}, \{\boldsymbol{x} : \boldsymbol{a}_i^t\boldsymbol{x} = b_i\}).$$

Proof. It is not hard to establish that $\partial P = P \cap (\cup_{i=1}^M \{\boldsymbol{x} : \boldsymbol{a}_i^t\boldsymbol{x} = b_i\})$. We claim that

$$\operatorname{dis}(\boldsymbol{\mu}, \cup_{i=1}^M \{\boldsymbol{x} : \boldsymbol{a}_i^t\boldsymbol{x} = b_i\}) = \operatorname{dis}(\boldsymbol{\mu}, P \cap (\cup_{i=1}^M \{\boldsymbol{x} : \boldsymbol{a}_i^t\boldsymbol{x} = b_i\})).$$

Since $\cup_{i=1}^{M}\{\boldsymbol{x} : \boldsymbol{a}_i^t b x = b_i\}$ is closed, there exists $\boldsymbol{y} \in \cup_{i=1}^{M}\{\boldsymbol{x} : \boldsymbol{a}_i^t b x = b_i\}$ such that

$$\|\boldsymbol{\mu} - \boldsymbol{y}\|_2 = \mathrm{dis}(\boldsymbol{\mu}, \cup_{i=1}^{M}\{\boldsymbol{x} : \boldsymbol{a}_i^t b x = b_i\}).$$

We claim that $\boldsymbol{y} \in P$. Suppose not (towards a contradiction). Then, there exists $\theta \in (0,1)$ such that $\boldsymbol{z} = (1-\theta)\boldsymbol{\mu} + \theta\boldsymbol{y} \in \partial P$. Then,

$$\mathrm{dis}(\boldsymbol{\mu}, (\cup_{i=1}^{M}\{\boldsymbol{x} : \boldsymbol{a}_i^t \boldsymbol{x} = b_i\})) \leqslant \|\boldsymbol{z} - \boldsymbol{\mu}\|_2 < \|\boldsymbol{y} - \boldsymbol{\mu}\|_2 = \mathrm{dis}(\boldsymbol{\mu}, \cup_{i=1}^{M}\{\boldsymbol{x} : \boldsymbol{a}_i^t b x = b_i\}),$$

which is a contradiction, establishing the claim. Then,

$$\begin{aligned}\min_{i=1,\ldots,M} \mathrm{dis}(\boldsymbol{\mu}, \{\boldsymbol{x} : \boldsymbol{a}_i^t \boldsymbol{x} = b_i\}) &= \mathrm{dis}(\boldsymbol{\mu}, \cup_{i=1}^{M}\{\boldsymbol{x} : \boldsymbol{a}_i^t \boldsymbol{x} = b_i\}) \\ &= \mathrm{dis}(\boldsymbol{\mu}, P \cap (\cup_{i=1}^{M}\{\boldsymbol{x} : \boldsymbol{a}_i^t \boldsymbol{x} = b_i\})) \\ &= \mathrm{dis}(\boldsymbol{\mu}, \partial P).\end{aligned}$$

□

Lemma 5. *Suppose that $X_1, \ldots, X_T$ are centered scalar R-sub-Gaussian random variables. Then, $\forall u \in \{0, \ldots, \lceil \log(T) \rceil\}$,*

$$\Pr(\exists v \in [2^u, 2^{u+1}] : \frac{1}{v}\sum_{i=1}^{v} X_i \geqslant \sqrt{\frac{x}{v}}) \leqslant \exp(-\frac{x}{4R^2}).$$

Proof. Define $S_v = \sum_{i=1}^{v} X_i$. Fix $u \in \{0, \ldots, \lceil \log(T) \rceil\}$. Let $m = 2^{u+1}$. Hoeffding's maximal inequality yields (see Step 2 of Lemma 1 of (Jamieson et al., 2014a))

$$\begin{aligned}\Pr(\exists v \in [m] : \frac{1}{v}S_v \geqslant \frac{\sqrt{x}}{v}) &= \Pr(\exists v \in [m] : S_v \geqslant \sqrt{x}) \\ &\leqslant \exp(-\frac{x}{2R^2 m}).\end{aligned}$$

Then,

$$\begin{aligned}\Pr(\exists v \in [2^u, 2^{u+1}] : \frac{1}{v}S_v \geqslant \frac{\sqrt{x}}{v}) &\leqslant \Pr(\exists v \in [m] : \frac{1}{v}S_v \geqslant \frac{\sqrt{x}}{v}) \\ &\leqslant \exp(-\frac{x}{2R^2 m}).\end{aligned}$$

Finally,

$$\begin{aligned}\Pr(\exists v \in [2^u, 2^{u+1}] : \frac{1}{v}S_v \geqslant \sqrt{\frac{x}{v}}) &\leqslant \Pr(\exists v \in [2^u, 2^{u+1}] : \frac{1}{v}S_v \geqslant \frac{\sqrt{x2^u}}{v}) \\ &\leqslant \exp(-\frac{x2^u}{2R^2 m}) \\ &= \exp(-\frac{x}{4R^2}).\end{aligned}$$

□

Lemma 6. *Let $\epsilon > 0$ and $\mathcal{N}_\epsilon$ be an ϵ-net of $\mathcal{S}^{D-1}$. For any $\boldsymbol{y} \in \mathbb{R}^D$,*

$$\|\boldsymbol{y}\|_2 \leqslant \frac{1}{1-\epsilon} \sup_{\boldsymbol{z} \in \mathcal{N}_\epsilon} \boldsymbol{y}^t \boldsymbol{z}.$$

Proof. Let $\boldsymbol{z}_0 \in \mathcal{N}_\epsilon$ such that $\left\| \frac{\boldsymbol{y}}{\|\boldsymbol{y}\|_2} - \boldsymbol{z}_0 \right\|_2 \leqslant \epsilon$. Then, by Cauchy-Schwarz,

$$\|\boldsymbol{y}\|_2 = \frac{\boldsymbol{y}^t \boldsymbol{y}}{\|\boldsymbol{y}\|_2} = \boldsymbol{y}^t (\frac{\boldsymbol{y}}{\|\boldsymbol{y}\|_2} - \boldsymbol{z}_0) + \boldsymbol{y}^t \boldsymbol{z}_0 \leqslant \|\boldsymbol{y}\|_2 \left\| \frac{\boldsymbol{y}}{\|\boldsymbol{y}\|_2} - \boldsymbol{z}_0 \right\|_2 + \boldsymbol{y}^t \boldsymbol{z}_0 \leqslant \epsilon \|\boldsymbol{y}\|_2 + \boldsymbol{y}^t \boldsymbol{z}_0.$$

Rearranging the inequality, we obtain

$$\|\boldsymbol{y}\|_2 \leqslant \frac{1}{1-\epsilon} \boldsymbol{y}^t \boldsymbol{z}_0 \leqslant \frac{1}{1-\epsilon} \sup_{\boldsymbol{z} \in \mathcal{N}_\epsilon} \boldsymbol{y}^t \boldsymbol{z}.$$

□

The following Lemma appears in (Vershynin et al., 2017) (see Corollary 4.2.13).

Lemma 7. *Let $\epsilon > 0$ and $\mathcal{N}_\epsilon$ be a minimal ϵ-net of $\mathcal{S}^{D-1}$. Then, $|\mathcal{N}_\epsilon| \leqslant (\frac{2}{\epsilon} + 1)^D$.*

We state without proof general Hoeffding's inequality for sub-Gaussian random variables (see Theorem 2.6.2 in Vershynin et al. (2017)).

Lemma 8. *Suppose that $X_1, \ldots, X_n$ are i.i.d. scalar R-sub-Gaussian random variables with mean $\boldsymbol{\mu} \in \mathbb{R}$. Then, for all $t > 0$,*

$$\Pr(|\frac{1}{n} \sum_{i=1}^{n} X_i - \mu| > t) \leqslant 2 \exp(-\frac{t^2 n}{2R^2}).$$

2.17 Feasible Arm Identification with a Convex Region: Statistical Results

To begin, we introduce some notation. Let $\delta > 0$ and $\boldsymbol{x} \in \mathbb{R}^D$. Define $B_\delta(\boldsymbol{x}) = \{\boldsymbol{x} \in \mathbb{R}^D : \|\boldsymbol{x}\|_2 \leqslant \delta\}$.

Proposition 2. *Let P be a compact convex set with positive volume. There exists a sequence of polyhedra $\{P_n\}$ such that*

$$\lim_{n \longrightarrow \infty} \sup_{\boldsymbol{y} \in \mathbb{R}^D} |\operatorname{dis}(\boldsymbol{y}, \partial P_n) - \operatorname{dis}(\boldsymbol{y}, \partial P)| = 0.$$

Further, if there exists some $\tau > 0$ *such that* $\text{dist}(\boldsymbol{\mu}_i, \partial P) \geqslant \tau \ \forall i \in [K]$*, then for all* $\delta > 0$*, there exists a large enough* n *such that*

$$(1-\delta)H_P \leqslant H_{P_n} \leqslant (1+\delta)H_P.$$

Proof. **Step 1: Defining a sequence of approximations.** Define a *dyadic cube of side length* 2^{-n} as a set of the form

$$[\frac{i_1}{2^{-n}}, \frac{i_1+1}{2^{-n}}] \times \ldots \times [\frac{i_D}{2^{-n}}, \frac{i_D+1}{2^{-n}}]$$

where $i_1, \ldots i_D$ are integers and $n \in \mathbb{N}$. Let E_n denote the set of dyadic cubes with sidelength 2^{-n}. Define

$$P_n = \text{conv}(\underset{E \in E_n, E \subset P}{\cup} E).$$

Note that for any n, P_n is a polyhedron with a finite number of constraints.

Step 2: For large n**,** ∂P_n **is a good approximation of** ∂P**.** Next, we claim that $\forall \delta > 0$, there exists N such that $n \geqslant N$ implies that $\sup_{\boldsymbol{x} \in \partial P_n} \text{dis}(\boldsymbol{x}, \partial P) \leqslant \delta$. Suppose not. Then, there exists $\delta > 0$ such that $\forall n \in \mathbb{N}$ there exists $\boldsymbol{x}_n$ such that $\text{dis}(\boldsymbol{x}_n, \partial P) > \delta$. Since P is compact and $\{\boldsymbol{x}_n\}_{n \in \mathbb{N}} \subset P$, there exists a convergent subsequence $\{\boldsymbol{x}_{n_j}\}$ with limit $\boldsymbol{x} \in P$. Then, $\text{dis}(\boldsymbol{x}, \partial P) \geqslant \delta$, which implies that $\boldsymbol{x} \in P^\circ$ and $B_\delta(\boldsymbol{x}) \subset P$. By definition of P_n, there exists N such that $n \geqslant N$ implies that $B_{\frac{\delta}{2}}(\boldsymbol{x}) \subset P_N$. Thus, $n \geqslant N$ implies that $\text{dis}(\boldsymbol{x}, \partial P_n) \geqslant \frac{\delta}{2}$. Thus, $\boldsymbol{x}_{n_j}$ cannot converge to $\boldsymbol{x}$, which is a contradiction. So, the claim is true.

Next, we claim that $\forall \delta > 0$, there exists N such that $n \geqslant N$ implies that $\sup_{\boldsymbol{x} \in \partial P} \text{dis}(\boldsymbol{x}, \partial P_n) \leqslant \delta$. Suppose not. Then, $\exists \delta > 0$ such that $\forall n \in \mathbb{N}$ there exists $\boldsymbol{x}_n \in \partial P$ such that $\text{dis}(\boldsymbol{x}_n, \partial P_n) > \delta$. ∂P is bounded and closed so that ∂P is compact. Thus, $\{\boldsymbol{x}_n\}$ has a convergent subsequence $\{\boldsymbol{x}_{n_j}\}$ with limit point $\boldsymbol{x} \in \partial P$. $\boldsymbol{x}$ has the property that $\text{dis}(\boldsymbol{x}, \partial P_n) \geqslant \delta$ for all $n \in \mathbb{N}$. Let $\boldsymbol{y} \in P^\circ$ (such a point exists since P has positive volume). Then, since P is convex, by the line segment principle (Proposition 1.4.1 (Bertsekas, 2009)), every point of the form $\boldsymbol{z}_\theta = (1-\theta)\boldsymbol{x} + \theta \boldsymbol{y}$ for $\theta \in (0, 1]$ is such that $\boldsymbol{z}_\theta \in P^\circ$. So there exists $\boldsymbol{w} \in B_\delta(\boldsymbol{x}) \cap P^\circ$. For large enough n, $\boldsymbol{w} \in P_n$. Since $\|\boldsymbol{w} - \boldsymbol{x}\|_2 < \delta$, we have a contradiction and thus the claim follows.

Step 3: Distance to ∂P_n approaches uniformly distance to ∂P. Formally, we show that

$$\lim_{n \longrightarrow \infty} \sup_{\boldsymbol{y} \in \mathbb{R}^D} |\operatorname{dis}(\boldsymbol{y}, \partial P_n) - \operatorname{dis}(\boldsymbol{y}, \partial P)| = 0. \tag{2.50}$$

Let $\delta > 0$. Let n large enough so that $\sup_{\boldsymbol{x} \in \partial P_n} \operatorname{dis}(\boldsymbol{x}, \partial P) \leqslant \delta$ and $\sup_{\boldsymbol{x} \in \partial P} \operatorname{dis}(\boldsymbol{x}, \partial P_n) \leqslant \delta$. Fix $\boldsymbol{y} \in \mathbb{R}^D$. Let $\boldsymbol{x}_p \in \partial P$ such that $\|\boldsymbol{y} - \boldsymbol{x}_p\|_2 = \operatorname{dis}(\boldsymbol{y}, \partial P)$ and $\boldsymbol{x}_{p_n} \in \partial P_n$ such that $\|\boldsymbol{y} - \boldsymbol{x}_{p_n}\|_2 = \operatorname{dis}(\boldsymbol{y}, \partial P_n)$.

Let $\boldsymbol{z} \in \partial P$ such that $\|\boldsymbol{x}_{p_n} - \boldsymbol{z}\|_2 \leqslant \delta$. Then, by the reverse triangle inequality,

$$|\, \|\boldsymbol{z} - \boldsymbol{y}\|_2 - \|\boldsymbol{x}_{p_n} - \boldsymbol{y}\|_2 \,| \leqslant \delta.$$

Then,

$$\operatorname{dist}(\boldsymbol{y}, \partial P) - \|\boldsymbol{y} - \boldsymbol{x}_{p_n}\|_2 \leqslant \operatorname{dist}(\boldsymbol{y}, \partial P) - \|\boldsymbol{z} - \boldsymbol{y}\|_2 + \delta \leqslant \delta.$$

Let $\boldsymbol{w} \in \partial P_n$ such that $\|\boldsymbol{x}_p - \boldsymbol{w}\|_2 \leqslant \delta$. By the reverse triangle inequality,

$$|\, \|\boldsymbol{y} - \boldsymbol{x}_p\|_2 - \|\boldsymbol{y} - \boldsymbol{w}\|_2 \,| \leqslant \delta.$$

Then,

$$\operatorname{dist}(\boldsymbol{y}, \partial P_n) - \|\boldsymbol{y} - \boldsymbol{x}_p\|_2 \leqslant \operatorname{dist}(\boldsymbol{y}, \partial P_n) - \|\boldsymbol{y} - \boldsymbol{w}\|_2 + \delta \leqslant \delta.$$

This establishes (2.50).

Step 4: Approximation of Problem Complexity. Suppose there exists some $\tau > 0$ such that $\operatorname{dist}(\boldsymbol{\mu}_i, \partial P) \geqslant \tau \ \forall i \in [K]$. Let $\gamma_1 = 1 - \frac{1}{\sqrt{1+\delta}}$, $\gamma_2 = \frac{1}{\sqrt{1-\delta}} - 1$ and $\gamma = \min(\gamma_1, \gamma_2)$. Let n large enough such that $\forall \boldsymbol{y} \in \mathbb{R}^D$,

$$|\operatorname{dis}(\boldsymbol{y}, \partial P_n) - \operatorname{dis}(\boldsymbol{y}, \partial P)| \leqslant \gamma\tau \leqslant \gamma \min_i \operatorname{dis}(\boldsymbol{\mu}_i, \partial P).$$

Then,

$$\begin{aligned}
H_P &= \sum_{i\in[K]} [\mathrm{dis}(\boldsymbol{\mu}_i, \partial P) + \epsilon]^{-2} \\
&\leqslant \sum_{i\in[K]} [\mathrm{dis}(\boldsymbol{\mu}_i, \partial P_N)(1-\gamma) + \epsilon]^{-2} \\
&\leqslant \sum_{i\in[K]} [\mathrm{dis}(\boldsymbol{\mu}_i, \partial P_N) + \epsilon]^{-2}(1-\gamma)^{-2} \\
&\leqslant \sum_{i\in[K]} [\mathrm{dis}(\boldsymbol{\mu}_i, \partial P_N) + \epsilon]^{-2}(1+\delta) \\
&= (1+\delta) H_{P_n}
\end{aligned}$$

Similarly, $H_P \geqslant (1-\delta) H_{P_n}$.

□

Theorem 5. *Let P be a convex set with positive volume and $\epsilon \geqslant 0$ such that $P_\epsilon^\circ \coloneqq \{\boldsymbol{x} \in P : \mathrm{dist}(\boldsymbol{x}, \partial P) > \epsilon)\}$ is nonempty. Let $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K \in P_\epsilon^\circ$. Then, for any $\delta > 0$, there exists a collection of $K+1$ problems $\mathcal{B}^0, \ldots, \mathcal{B}^K$ such that for any algorithm,*

$$\max_{i\in\{0,\ldots,K\}} \mathbb{E}_{\mathcal{B}^i}(\mathcal{L}_{T,P,\epsilon}(\widehat{S})) \geqslant \exp(-13\frac{T}{(1-\delta)H_P} - 25D\log(48(\log(T)+1)KD))).$$

where

$$H_P = \sum_{i=1}^{K} [\ \mathrm{dist}(\boldsymbol{\mu}_i, \partial P) + \epsilon]^{-2}.$$

Proof. **Step 1: Reduce convex set to compact convex set.** Let $r_i \coloneqq 2\ \mathrm{dist}(\boldsymbol{\mu}_i, \partial P)$ for all $i \in [K]$. Clearly, there exists large enough $B > 0$ such that $P_0 = P \cap \{\boldsymbol{x} \in \mathbb{R}^D : \|\boldsymbol{x}\|_2 \leqslant B\}$ has the property that for all $i \in [K]$, if we replace $\boldsymbol{\mu}_i$ with any $\tilde{\boldsymbol{\mu}}_i \in B_r(\boldsymbol{\mu}_i)$, then $\mathrm{dist}(\tilde{\boldsymbol{\mu}}_i, \partial P) = \mathrm{dist}(\tilde{\boldsymbol{\mu}}_i, \partial P_0)$ and

$$\mathcal{L}_{T,P,\epsilon}(S) = \mathcal{L}_{T,P_0,\epsilon}(S) \tag{2.51}$$

for all $S \subset [K]$. Further, for the feasible identification problem with $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K$ as the means, $H_P = H_{P_0}$.

Step 2: Define approximation. Since $\mathrm{dist}(\boldsymbol{\mu}_i, \partial P_0) > \epsilon$ for all $i \in [K]$, there exists $\gamma \in (0, \epsilon)$ such that for all $i \in [K]$,

$$\mathrm{dist}(\boldsymbol{\mu}_i, \partial P_0) > \epsilon + \gamma. \tag{2.52}$$

Fix $\delta > 0$. By Proposition 2, there exists a polyhedron P_{approx} such that

$$(1-\delta)H_{P_0} \leqslant H_{P_{approx}} \leqslant (1+\delta)H_{P_0}$$

and

$$\sup_{\boldsymbol{y}\in\mathbb{R}^D} |\operatorname{dis}(\boldsymbol{y}, \partial P_{approx}) - \operatorname{dis}(\boldsymbol{y}, \partial P_0)| < \frac{\gamma}{2}. \tag{2.53}$$

By (2.52) and (2.53), $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K \in P_{approx}$ and for all $i \in [K]$, $\operatorname{dis}(\boldsymbol{\mu}_i, \partial P_{approx}) > \epsilon + \frac{\gamma}{2}$ so that $\mathcal{L}_{T,P_0,\epsilon}(S) = \mathcal{L}_{T,P_{approx},\epsilon}(S)$ for all $S \subset [K]$.

Step 3: Apply lower bound for polyhedra. Apply the lower bound construction from Theorem 1 to P_{approx} to define $K+1$ collections of distributions $\mathcal{B}^i$ for $i \in \{0, 1, \ldots, K\}$ (see Theorem 1 for their definitions). We claim that for every problem $\mathcal{B}^i$, $\mathcal{L}_{T,P_0,\epsilon}(S) = \mathcal{L}_{T,P_{approx},\epsilon}(S)$ for all $S \subset [K]$. We briefly sketch the proof. First,

$$\operatorname{dist}(\boldsymbol{\mu}_i', \partial P_0) + \frac{\gamma}{2} > \operatorname{dist}(\boldsymbol{\mu}_i', \partial P_{approx}) = \operatorname{dist}(\boldsymbol{\mu}_i, \partial P_{approx}) > \operatorname{dist}(\boldsymbol{\mu}_i, \partial P_0) - \frac{\gamma}{2} \geqslant \epsilon + \frac{\gamma}{2}.$$

Thus, $\operatorname{dist}(\boldsymbol{\mu}_i', \partial P_0) > \epsilon$. Further, $\boldsymbol{\mu}_i' \notin P_{approx}$, $\operatorname{dist}(\boldsymbol{\mu}_i', P_{approx}) \geqslant \epsilon + \frac{\gamma}{2}$, and (2.53) imply that $\boldsymbol{\mu}_i' \notin P$. The claim follows from this observation.

Thus, by Theorem 1, for all $\widehat{S} \subset [K]$,

$$\begin{aligned}
\max_{i\in\{0,\ldots,K\}} \mathbb{E}_{\mathcal{B}^i} \mathcal{L}_{T,P,\epsilon}(\widehat{S}) &= \max_{i\in\{0,\ldots,K\}} \mathbb{E}_{\mathcal{B}^i} \mathcal{L}_{T,P_0,\epsilon}(\widehat{S}) \qquad (2.54)\\
&= \max_{i\in\{0,\ldots,K\}} \mathbb{E}_{\mathcal{B}^i} \mathcal{L}_{T,P_{approx},\epsilon}(\widehat{S})\\
&\geqslant \exp(-13\frac{T}{H_{P_{approx}}} - 25D\log(48(\log(T)+1)KD)))\\
&\geqslant \exp(-13\frac{T}{H_P(1-\delta)} - 25D\log(48(\log(T)+1)KD)))
\end{aligned}$$

where line (2.54) follows by (2.51).

□

Theorem 6. *Let P be a convex set with positive volume and $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K \in \mathbb{R}^D$. Suppose that there are some $B > 0$ and $\gamma > 0$ such that it is known that*

1. $\max(\operatorname{dist}(\boldsymbol{\mu}_i, \partial P), \|\boldsymbol{\mu}_i\|_2) \leqslant B$ *for all* $i \in [K]$,

2. $|\operatorname{dis}(\boldsymbol{\mu}_i, \partial P) - \epsilon| \geqslant \gamma$ *for all* $i \in [K]$, *and*

3. $\operatorname{dist}(\boldsymbol{\mu}_i, \partial P) \geqslant \gamma$.

Then, there exists an algorithm that given any $\delta > 0$*, achieves*

$$\mathbb{E}[\mathcal{L}_{T,P,\epsilon}(\widehat{S})] \leqslant 2(\log(T)+1)K5^D \exp(-\frac{T}{1296R^2H_P(1+\delta)}).$$

Proof. **Step 1: Define the algorithm.** The algorithm is as follows:

1. Set $P_0 = P \cap \overline{B_{2B}(\mathbf{0})}$.

2. Use the construction from Proposition 2 to approximate P_0 with P_{approx} such that

$$\sup_{\boldsymbol{y}\in\mathbb{R}^D} |\operatorname{dis}(\boldsymbol{y}, \partial P_0) - \operatorname{dis}(\boldsymbol{y}, \partial P_{approx})| \leqslant \frac{\gamma}{2} \tag{2.55}$$
$$(1-\delta)H_{P_0} \leqslant H_{approx} \leqslant H_{P_0}(1+\delta).$$

3. Run MD-APT with the K given arms, P_{approx}, and ϵ and return its answer $\widehat{S}$.

We note that because it is known that $\operatorname{dist}(\boldsymbol{\mu}_i, \partial P) \geqslant \gamma$, step 2 of the algorithm is valid.

Step 2: Distance of $\boldsymbol{\mu}_i$ to ∂P is equal to the distance of $\boldsymbol{\mu}_i$ to ∂P_0. First, we claim that

$$\operatorname{dist}(\boldsymbol{\mu}_i, \partial P) = \operatorname{dist}(\boldsymbol{\mu}_i, \partial P_0). \tag{2.56}$$

Let $\boldsymbol{x} \in \partial P$ such that $\|\boldsymbol{x} - \boldsymbol{\mu}_i\|_2 = \operatorname{dist}(\boldsymbol{\mu}_i, \partial P)$. Then,

$$\|\boldsymbol{x}\|_2 \leqslant \|\boldsymbol{\mu}_i\|_2 + \|\boldsymbol{x} - \boldsymbol{\mu}_i\|_2 \leqslant 2B.$$

Thus, $\boldsymbol{x} \in P_0$. Since $P_0 \subset P$, $\boldsymbol{x} \in \partial P_0$. Therefore, $\operatorname{dist}(\boldsymbol{\mu}_i, \partial P_0) \leqslant \operatorname{dist}(\boldsymbol{\mu}_i, \partial P)$. Towards a contradiction, suppose that $\operatorname{dist}(\boldsymbol{\mu}_i, \partial P_0) < \operatorname{dist}(\boldsymbol{\mu}_i, \partial P)$. Let $\boldsymbol{y} \in \partial P_0$ such that $\|\boldsymbol{\mu}_i - \boldsymbol{y}\| = \operatorname{dist}(\boldsymbol{\mu}_i, \partial P_0)$. Then,

$$\|\boldsymbol{y}\|_2 \leqslant \|\boldsymbol{\mu}_i\|_2 + \|\boldsymbol{y} - \boldsymbol{\mu}_i\|_2 < 2B. \tag{2.57}$$

Recall the fact

$$\partial(A \cap B) \subset \partial A \cup \partial B.$$

Therefore, by this fact, $\boldsymbol{y} \in \partial P_0$ and (2.57) imply that $\boldsymbol{y} \in \partial P$. Thus, we have a contradiction. This establishes (2.56).

Step 3: $\mathcal{L}_{T,P,\epsilon}(S) = \mathcal{L}_{T,P_0\epsilon}(S)$ for every $S \subset [K]$. Next, we show that $\mathcal{L}_{T,P,\epsilon}(S) = \mathcal{L}_{T,P_0,\epsilon}(S)$ for all $S \subset [K]$. Suppose $\boldsymbol{\mu}_i \in S^{\text{int}}_{P,\epsilon}$. Then, by (2.56), $\text{dist}(\boldsymbol{\mu}_i, \partial P_0) = \text{dist}(\boldsymbol{\mu}_i, \partial P) \geqslant \epsilon$. Further, by hypothesis, $\|\boldsymbol{\mu}_i\|_2 \leqslant B$ so that $\boldsymbol{\mu}_i \in P_0$. Thus, $\boldsymbol{\mu}_i \in S^{\text{int}}_{P_0,\epsilon}$.

Next, suppose that $\boldsymbol{\mu}_i \in S^{\text{out}}_{P,\epsilon}$. Then, $P_0 \subset P$ implies that $\boldsymbol{\mu}_i \notin P_0$ and $\text{dist}(\boldsymbol{\mu}_i, P_0) = \text{dist}(\boldsymbol{\mu}_i, P) > \epsilon$ by (2.56). Thus, $\boldsymbol{\mu}_i \in S^{\text{out}}_{P_0,\epsilon}$.

Next, suppose that $\boldsymbol{\mu}_i \notin S^{\text{out}}_{P,\epsilon}$ and $\boldsymbol{\mu}_i \notin S^{\text{int}}_{P,\epsilon}$. Then, either *(i)* $\boldsymbol{\mu}_i \in P$ and $\text{dis}(\boldsymbol{\mu}_i, \partial P) < \epsilon$ or *(ii)* $\boldsymbol{\mu}_i \notin P$ and $\text{dist}(\boldsymbol{\mu}_i, \partial P) \leqslant \epsilon$. Suppose *(i)*. Then, by (2.56), it follows that $\boldsymbol{\mu}_i \notin S^{\text{out}}_{P_0,\epsilon}$ and $\boldsymbol{\mu}_i \notin S^{\text{int}}_{P_0,\epsilon}$. Suppose *(ii)*. Then, $P_0 \subset P$ and (2.56) imply that $\boldsymbol{\mu}_i \notin S^{\text{out}}_{P_0,\epsilon}$ and $\boldsymbol{\mu}_i \notin S^{\text{int}}_{P_0,\epsilon}$. This establishes the claim.

Step 4: Putting it together. (2.55) and the hypotheses imply that $\mathcal{L}_{T,P_0,\epsilon}(S) = \mathcal{L}_{T,\epsilon,P_{approx}}(S)$ for all $S \subset [K]$.

Thus, let $\widehat{S}$ denote the output of MD-APT with the K given arms, P_{approx}, and ϵ. By Theorem 4,

$$\begin{aligned}
\mathbb{E}[\mathcal{L}_{T,P,\epsilon}(S)] &= \mathbb{E}[\mathcal{L}_{T,P_0,\epsilon}(S)] \\
&= \mathbb{E}[\mathcal{L}_{T,P_{approx},\epsilon}(S)] \\
&\leqslant 2(\log(T)+1)K5^D \exp(-\frac{T}{1296R^2 H_{P_approx}}) \\
&\leqslant 2(\log(T)+1)K5^D \exp(-\frac{T}{1296R^2 H_P(1+\delta)}).
\end{aligned}$$

□

2.18 Additional Experiments

In this section, we present a couple more experiments. First, we present another variant of experiment 3, linear progression of arms on a cube, where there are no irrelevant arms. We set $\epsilon = 0$. We use $\boldsymbol{\mu}_{0:3} = (.75)^{\otimes 5} + (0:3) \times .05$, $\boldsymbol{\mu}_4 = (.95)^{\otimes 5}$, $\boldsymbol{\mu}_5 = (1.05)^{\otimes 5}$, $\boldsymbol{\mu}_{6:9} = (1.25)^{\otimes 5} - (0:3) \times .05$, $\boldsymbol{\mu}_{10:19} = (1.15)^{\otimes 5}$. In comparison to experiment 3, we make it slightly easier to determine whether the arms $\boldsymbol{\mu}_4$ and $\boldsymbol{\mu}_5$ belong to the polyhedron because otherwise

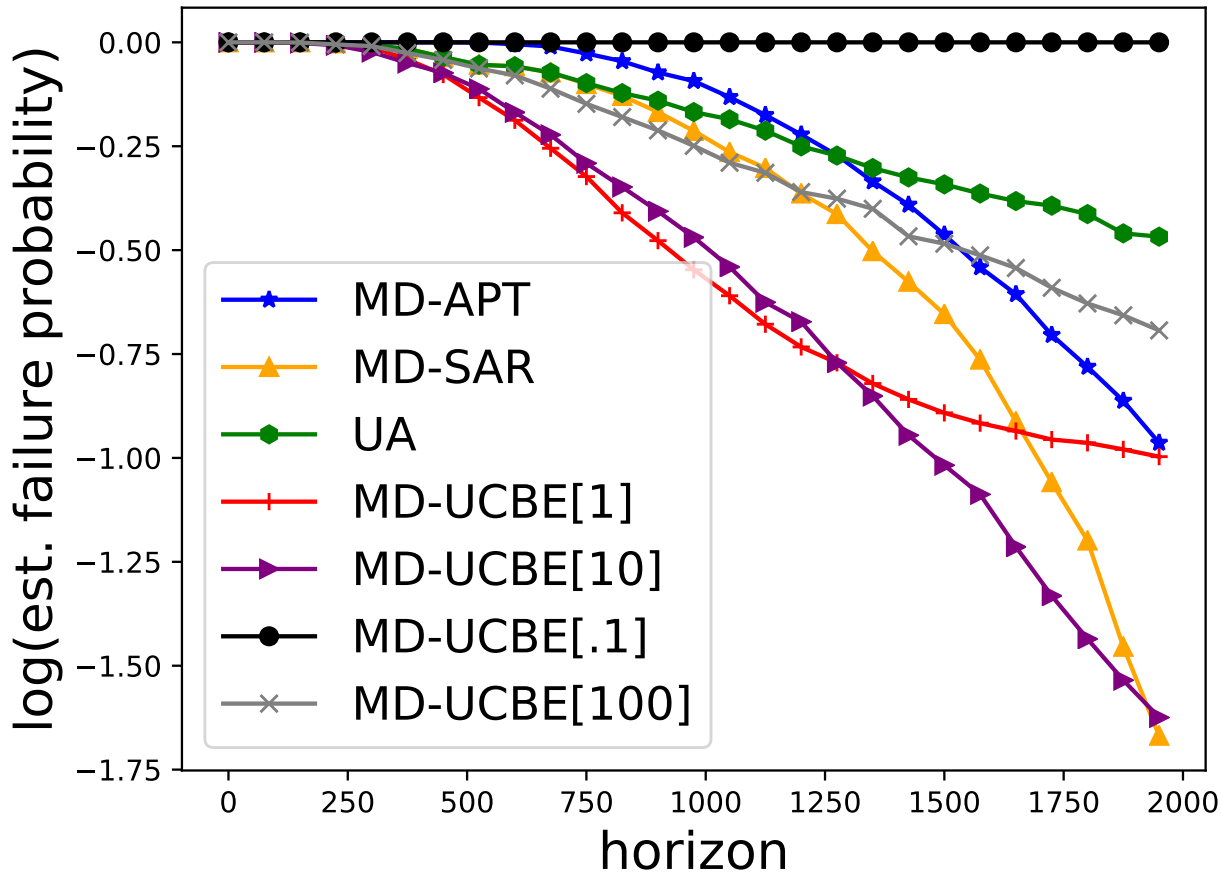


Figure 2.8: Linear Progression on Cube, no Irrelevant Arms

the difficulty of the problem prevents any algorithm from achieving substantial progress after 2000 time steps. Figure 2.8 presents the results. MD-SAR performs substantially better than MD-APT. MD-APT pulls arm 4, which minimizes Δ_i, too much. MD-APT pulls arm 4 time on average 1006.27 times, whereas MD-SAR pulls arm 4 on average 319.59 times.

We also repeat the crowdsourcing experiment with a slightly different setup. Now, we draw samples from a Gaussian distribution for each worker with mean calculated from the dataset in Snow et al. (2008) and variance over all the ratings over all the workers. The results are very similar to the results in the main text.

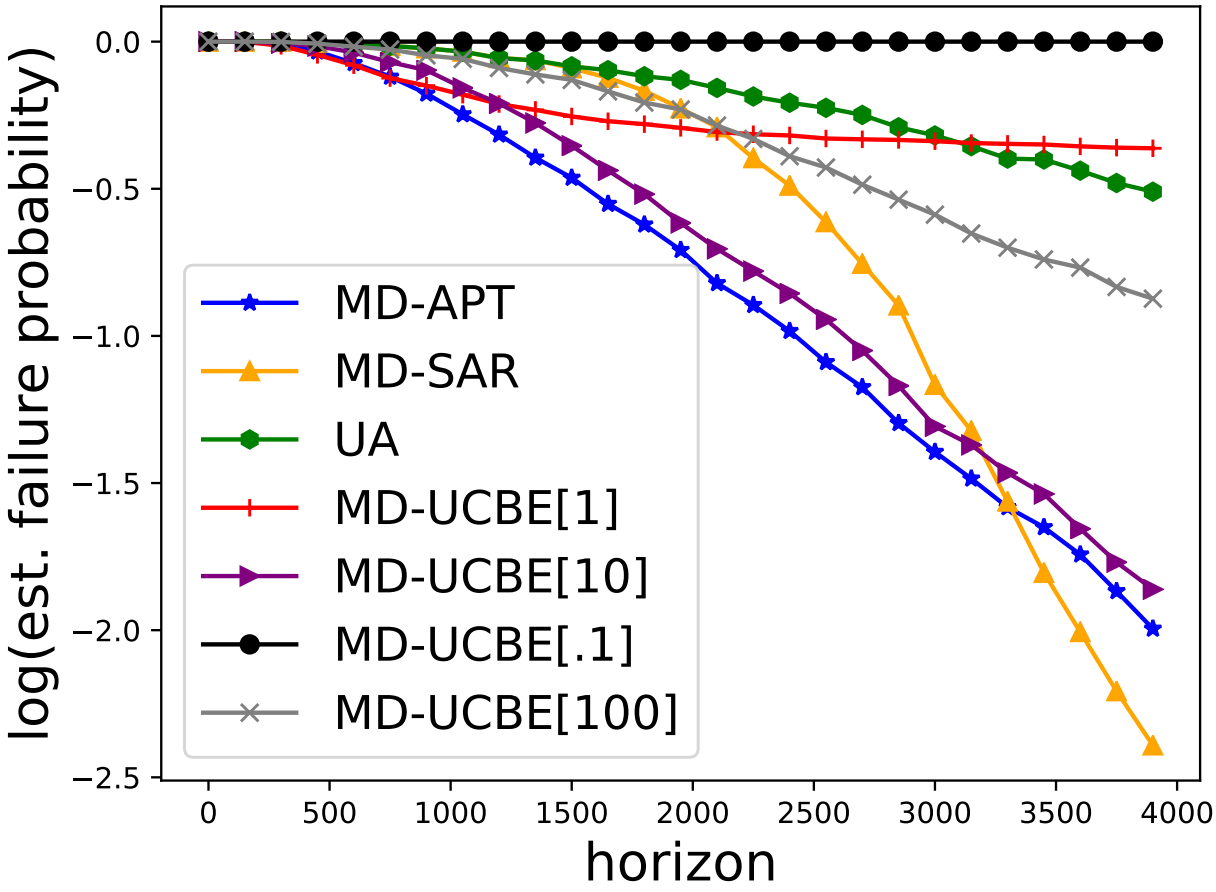


Figure 2.9: Crowdsourcing Experiment with Simulated Data

Chapter 3

Top Feasible Arm Identification

In this chapter, I propose a new variant of the top arm identification problem, *top feasible arm identification*, where there are K arms associated with D-dimensional distributions and the goal is to find m arms that maximize some known linear function of their means subject to the constraint that their means belong to a given set $P \subset \mathbb{R}^D$. This problem has many applications since in many settings, feedback is multi-dimensional and it is of interest to perform *constrained maximization*. I present problem-dependent lower bounds for top feasible arm identification and upper bounds for several algorithms. Our most broadly applicable algorithm, TF-LUCB-B, has an upper bound that is loose by a factor of $O(D \log(K))$. Many problems of practical interest are two-dimensional and, for these, it is loose by a factor of $O(\log(K))$. Finally, I conduct experiments on synthetic and real-world datasets that demonstrate the effectiveness of our algorithms. My algorithms are superior both in theory and in practice to a naive two-stage algorithm that first identifies the feasible arms and then applies a best arm identification algorithm to the feasible arms. This Chapter is joint work with Clayton Scott and was presented at the International Conference on Artificial Intelligence and Statistics in 2019.

3.1 Introduction

In the top arm identification problem in multi-armed bandits, there are K scalar-valued distributions (also referred to as arms) and an agent plays a sequential game where, at each

round, the agent chooses (or "pulls") one of the arms and observes an i.i.d. realization from it. At the end of the game, the agent outputs the set of m arms believed to have the largest means. This problem has applications in areas such as crowdsourcing, A/B testing, and clinical trials.

While top arm identification considers settings where the feedback is scalar-valued and the goal is maximization, in many applications, the feedback is multi-dimensional and it is of interest to perform *constrained maximization*. For example, in crowdsourcing, an important challenge is to identify high-quality workers that complete work at a suitable pace (e.g., below 15 seconds on average) and, in clinical trials, it is of interest to efficiently find drugs that are most likely to be effective and have an acceptably low probability of causing an adverse effect.

In this paper, we introduce *top feasible arm identification* for situations where the feedback is multi-dimensional and the goal is constrained maximization. In this problem, there are K arms and each arm i is associated with a D-dimensional distribution ν_i that has mean $\boldsymbol{\mu}_i$. At each round $t = 1, 2, \ldots$, the agent chooses an arm I_t and observes an independent random vector drawn from ν_{I_t}. For given $P \subset \mathbb{R}^D$, $\boldsymbol{c} \in \mathbb{R}^D$, $m \leqslant K$, and $\delta \in (0, 1)$, the goal of the agent is to identify m arms that maximize $\boldsymbol{c}^\top \boldsymbol{\mu}_i$ subject to the constraint $\boldsymbol{\mu}_i \in P$, with probability at least $1 - \delta$, in the fewest number of samples possible.

We make several contributions to this problem. We prove problem-dependent lower bounds for top feasible arm identification. We also propose a family of algorithms TF-LUCB, where each instance is specified by a test for feasibility TestF, and we prove a master theorem that characterizes an upper bound for TF-LUCB in terms of the subroutine TestF. Finally, we use this master theorem to prove upper bounds for several algorithms. Our most broadly applicable algorithm, TF-LUCB-B, has an upper bound that is loose by a factor of $O(D \log(K))$. Many problems of practical interest are two-dimensional and for these, it is loose by a factor of $O(\log(K))$. Notably, our algorithms are superior both in theory and in practice to a naive two-stage algorithm that first identifies the feasible arms and then applies a best arm identification algorithm to the feasible arms. The sample complexity of such a two-stage algorithm can be arbitrarily larger than the sample complexity of our algorithms and, indeed, in our experiments we improve on such a baseline by as much as a factor of 4.5.

3.2 Related Work

Top arm identification has received a lot of attention in recent years (Mannor and Tistisklis, 2004; Audibert and Bubeck, 2010; Gabillon et al., 2012; Kalyanakrishnan et al., 2012b; Bubeck et al., 2013; Chen et al., 2014a; Jamieson et al., 2014a). Most work considers the case where arms are scalar-valued and, thus, their results cannot be applied to our problem setting. Recently, Chen et al. (2017a) proposed the general sampling problem, which does encompass a variant of top feasible arm identification. Their work differs from ours in several significant ways. First, in the work of Chen et al. (2017a), the agent samples from one dimension of one arm at a time, whereas in our setting pulling an arm yields a random D-dimensional vector. Second, Chen et al. (2017a) assume that the arms are isotropic Gaussian, whereas we assume each arm is a multi-dimensional sub-Gaussian distribution. Finally, their algorithm (see their Algorithm 7) is impractical for moderate values of δ in the fixed confidence setting since its first stage consists of a uniform allocation strategy that terminates when the confidence bounds of all of the means are small enough to determine which of the arms are in the top feasible m with probability at least 0.99.

Auer et al. (2016) also consider a setting where arms are multi-dimensional. Their goal is to determine the Pareto front of the arms, which is quite different from the task of constrained maximization in top feasible arm identification. We also remark that they use an elimination algorithm, whereas we adapt the LUCB algorithm from Kalyanakrishnan et al. (2012b) to our setting.

Recently, Katz-Samuels and Scott (2018) proposed the feasible arm identification problem, in which there are K multi-dimensional distributions and a given polyhedron, and the goal is to determine which of the distributions have means belonging to the polyhedron. By contrast, in top feasible arm identification, the goal is to find a collection of arms whose means are feasible and maximize some linear function. In short, Katz-Samuels and Scott (2018) deal with feasibility while the current paper deals with constrained maximization. Furthermore, whereas Katz-Samuels and Scott (2018) consider the fixed budget setting (in which there is a fixed number of rounds), we consider the fixed confidence setting. These differences require the development of new ideas and algorithms.

We also note that top feasible arm identification differs from best-arm identification in linear bandits (Soare et al., 2014). In best-arm identification in linear bandits, each arm i is associated with a *known* feature vector $\boldsymbol{x}_i$ and the reward of arm i has mean $\boldsymbol{x}_i^\top \boldsymbol{\theta}$ where $\boldsymbol{\theta}$ is unknown. In our setting, each arm is associated with a D-dimensional distribution and the goal is to maximize some known linear function $f : \mathbb{R}^D \longrightarrow \mathbb{R}$ subject to the constraint that $\boldsymbol{\mu}_i \in P$.

3.3 Problem Statement

Notation. For $n \in \mathbb{N}$, let $[n] = \{1, \ldots, n\}$. Let U be a finite set and f be a scalar-valued function with domain containing U, and define $\max_{x \in U}^{(l)} f(x) \coloneqq$

$$\begin{cases} \max\limits_{\{x \in U : |\{y \in U : f(y) \geqslant f(x)\}| \geqslant l-1\}} f(x) & : |U| \geqslant l \\ -\infty & : \text{ otherwise} \end{cases}.$$

In words, $\max_{x \in U}^{(l)} f(x)$ is the value of the lth largest $x \in U$ under $f(\cdot)$ and if $|U| < l$, then it is $-\infty$. For a set $A \subset \mathbb{R}^D$, let ∂A denote the boundary of A, i.e., $\partial A = \bar{A} \backslash A^\circ$ (where $\bar{A}$ denotes the closure of A and A° denotes the interior of A). Let $\boldsymbol{x} \in \mathbb{R}^D$, and define $\text{dist}(\boldsymbol{x}, A) = \inf_{\boldsymbol{y} \in A} \|\boldsymbol{x} - \boldsymbol{y}\|_2$. Let $\gamma > 0$, and define $B_\gamma(\boldsymbol{x}) = \{\boldsymbol{y} : \|\boldsymbol{x} - \boldsymbol{y}\|_2 < \gamma\}$. Let x_i denote the ith entry of $\boldsymbol{x}$ and for $\boldsymbol{y}_i \in \mathbb{R}^D$, let $y_{i,j}$ denote the jth entry of $\boldsymbol{y}_i$. Let $\boldsymbol{e}_i$ denote the ith standard basis vector. We use "whp" for "with high probability" and "wrt" for "with respect to."

Problem Parameters. Suppose that there are K arms associated with distributions $\nu_1, \ldots, \nu_K$ over $\mathbb{R}^D$ that have means $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K \in \mathbb{R}^D$, and let $\nu = (\nu_1, \ldots, \nu_K)$. At each round $t = 1, 2, \ldots$, an agent chooses an arm I_t and observes an independent draw $\boldsymbol{X}_t \sim \nu_{I_t}$.

Let $P \subset \mathbb{R}^D$ denote a nonempty set such that $P \neq \mathbb{R}^D$. Let $\boldsymbol{c}$ denote a reward vector, which is fixed, known, and the same across all arms. We assume $\|\boldsymbol{c}\|_2 = 1$. We say that $\boldsymbol{c}^\top \boldsymbol{\mu}_i$ is the expected *reward* of arm i. Let m denote the number of top feasible arms desired. We denote an instance of the top feasible arm identification problem by $(\nu, P, \boldsymbol{c}, m)$. Let $\Pr_\nu$ ($\mathbb{E}_\nu$) denote the probability measure (expected value) associated with the problem instance $(\nu, P, \boldsymbol{c}, m)$.

Define FEAS $= \{i \in [K] : \boldsymbol{\mu}_i \in P\}$, INFEAS $=$ FEAS^c, and

$$\text{OPT} = \{i \in \text{FEAS} : \boldsymbol{c}^\top \boldsymbol{\mu}_i \geqslant \max^{(m)}_{j \in \text{FEAS}} \boldsymbol{c}^\top \boldsymbol{\mu}_j\},$$
$$\text{SUBOPT} = \{i \in [K] : \boldsymbol{c}^\top \boldsymbol{\mu}_i < \max^{(m)}_{j \in \text{FEAS}} \boldsymbol{c}^\top \boldsymbol{\mu}_j\}.$$

We say that an arm j is *suboptimal* if $\boldsymbol{c}^\top \boldsymbol{\mu}_j < \max^{(m)}_{i \in \text{FEAS}} \boldsymbol{c}^\top \boldsymbol{\mu}_i$; we say that an arm j is *feasible* (*infeasible*) if $\boldsymbol{\mu}_j \in P$ ($\boldsymbol{\mu}_j \notin P$). We note that, in general, SUBOPT and INFEAS are not disjoint, and that when there are fewer than m arms that are feasible ($\boldsymbol{\mu}_i \in P$), SUBOPT $= \varnothing$.

We consider the following class of problems: $\mathcal{M} \coloneqq$

$$\{(\nu, P, \boldsymbol{c}, m) : (\forall i : \boldsymbol{\mu}_i \notin \partial P) \text{ and}$$
$$(\max^{(m)}_{i \in \text{FEAS}} \boldsymbol{c}^\top \boldsymbol{\mu}_i > \max^{(m+1)}_{j \in \text{FEAS}} \boldsymbol{c}^\top \boldsymbol{\mu}_j \vee |\,\text{FEAS}\,| \leqslant m)\}.$$

In words, $\mathcal{M}$ consists of problems where the means of the arms do not belong to the boundary of P and either there are m or fewer feasible arms or the mth largest reward of a feasible arm and the $(m+1)$th largest reward of a feasible arm are distinct. It is possible to drop the assumption $(\nu, P, \boldsymbol{c}, m) \in \mathcal{M}$ by allowing for a tolerance for suboptimality or infeasibility, and we describe this extension in the supplemental material.

Goal. We consider the fixed confidence setting with a novel criterion for correctness. An algorithm $\mathcal{A}$ is associated with a policy that determines which arm $I_t \in [K]$ is chosen at time t, a finite stopping time τ wrt $I_1, \boldsymbol{X}_1, I_2, \boldsymbol{X}_2, \ldots$ (i.e., $\Pr_\nu(\tau < \infty) = 1$) that determines when the algorithm stops, and an outputted partition of the arms $(\widehat{\text{O}}, \widehat{\text{S}}, \widehat{\text{I}})$ with $\widehat{\text{O}} \cup \widehat{\text{S}} \cup \widehat{\text{I}} = [K]$.

A standard criterion of correctness for an algorithm is δ- PAC, which we now define.

Definition 2. *Let $\delta \in (0, 1)$. We say an algorithm $\mathcal{A}$ is δ- PAC wrt $\mathcal{M}$ if for any problem $(\nu, P, \boldsymbol{c}, m)$ belonging to $\mathcal{M}$, $\mathcal{A}$ outputs $\widehat{O} \subset [K]$ such that $\Pr_\nu(\widehat{O} = OPT) \geqslant 1 - \delta$.*

A standard goal is to design algorithms that are δ- PAC wrt $\mathcal{M}$ and that minimize τ. We propose a novel criterion δ- PAC-EXPLANATORY and aim to design algorithms that are δ- PAC-EXPLANATORY wrt $\mathcal{M}$ and that minimize τ.

Definition 3. *Let $\delta \in (0, 1)$. We say an algorithm $\mathcal{A}$ is δ- PAC-EXPLANATORY wrt $\mathcal{M}$ if for any problem $(\nu, P, \boldsymbol{c}, m)$ belonging to $\mathcal{M}$, $\mathcal{A}$ outputs a triple $(\widehat{O}, \widehat{S}, \widehat{I})$ of disjoint sets such*

that $\widehat{O} \cup \widehat{S} \cup \widehat{I} = [K]$ *and*

$$\Pr_\nu(\widehat{O} = \mathit{OPT} \text{ and } (\widehat{S}, \widehat{I}) \in \mathit{Valid\text{-}Partitions}) \geqslant 1 - \delta$$

where $\mathit{Valid\text{-}Partitions} \coloneqq$

$$\{(S, I) : S \subset \mathit{SUBOPT}, I \subset \mathit{INFEAS}, \\ S \cap I = \varnothing, S \cup I = \mathit{OPT}^c\}.$$

To identify arms in OPT, an agent must rule out every $i \in \text{OPT}^c$ as suboptimal or infeasible. When $\text{SUBOPT} \cap \text{INFEAS} \neq \varnothing$, there are arms that can be ruled out in multiple ways. Valid-Partitions captures the various *correct* ways to partition the arms in OPT^c to distinguish them from OPT. Thus, our notion, δ- PAC-EXPLANATORY, is slightly stronger than δ- PAC since it essentially requires that whp *(i)* an algorithm output the correct top m feasible arms and *(ii)* that it provide a correct reason for rejecting each arm (either that it is suboptimal or infeasible). We remark that it is natural to require only one reason for rejecting an arm because once an algorithm identifies an arm as infeasible (suboptimal), there is no reason to keep pulling it to determine whether it is suboptimal (infeasible). Furthermore, in most problems and for most algorithms, if an arm is infeasible and suboptimal, showing one of these is easier than showing the other.

This notion is practically relevant since in many applications it is of interest to provide a reason to reject an arm. For example, in crowdsourcing, it might be necessary to provide a worker with a reason for why she was not hired. In clinical trials, it might be useful for the clinician to know why a drug is rejected. Furthermore, as we discuss in the supplemental material, we conjecture that there is a small gap between δ- PAC and δ- PAC-EXPLANATORY algorithms.

Sub-Gaussian Assumption. We assume that each ν_i is a multi-dimensional sub-Gaussian distribution, which we now define. Let X be a scalar random variable and $\boldsymbol{X} \in \mathbb{R}^D$ a random vector. We say that X is *sub-Gaussian* if $\mathbb{E} \exp(\frac{X^2}{\sigma^2}) \leqslant 2$ for some $\sigma > 0$ and $\boldsymbol{X} \in \mathbb{R}^D$ is sub-Gaussian if for all $\boldsymbol{a} \in \mathbb{R}^D$, $\boldsymbol{X}^\top \boldsymbol{a}$ is sub-Gaussian. The sub-Gaussian norms

of X and $\boldsymbol{X}$ are defined respectively as:

$$\|X\|_{\psi_2} = \inf\{\sigma > 0 : \mathbb{E}\exp(\frac{X^2}{\sigma^2}) \leqslant 2\},$$
$$\|\boldsymbol{X}\|_{\psi_2} = \sup_{\boldsymbol{a}\in\mathbb{R}^D:\|\boldsymbol{a}\|_2=1} \left\|\boldsymbol{X}^\top \boldsymbol{a}\right\|_{\psi_2}.$$

X is said to be σ-*sub-Gaussian* if $\|X\|_{\psi_2} \leqslant \sigma$ and $\boldsymbol{X}$ is said to be σ-*sub-Gaussian* if $\|\boldsymbol{X}\|_{\psi_2} \leqslant \sigma$. For the remainder of this paper, we assume that $\nu_1, \ldots, \nu_K$ are σ-sub-Gaussian. See Vershynin (2012); Vershynin et al. (2017) for more details.

3.4 Lower Bounds

Theorem 7 gives our lower bound for δ- PAC-EXPLANATORY algorithms.

Theorem 7. *Let $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K \in \mathbb{R}^D$ such that $\forall i \neq j \in [K] : \mu_{i,1} \neq \mu_{j,1}$. Define $\nu_i = N(\boldsymbol{\mu}_i, I_D)$ for all $i \in [K]$. Suppose $P = \mathbb{R} \times P'$ for some $P' \subset \mathbb{R}^{D-1}$ and $\forall \boldsymbol{x} \in \partial P, \forall \epsilon > 0 : B_\epsilon(\boldsymbol{x}) \cap P^\circ \neq \varnothing$ and $B_\epsilon(\boldsymbol{x}) \cap (P^c)^\circ \neq \varnothing$. Let $\boldsymbol{c} = \boldsymbol{e}_1$. Assume $(\nu, P, \boldsymbol{e}_1, m) \in \mathcal{M}$ and let $\delta \in (0, 0.1)$. For any $(S, I) \in$ Valid-Partitions, define $L(S, I) \coloneqq$*

$$\sum_{i\in OPT} \max([\min_{j\in S} \boldsymbol{c}^\top(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)]^{-2},\ \mathrm{dist}(\boldsymbol{\mu}_i, \partial P)^{-2})$$
$$+ \sum_{i\in S}[\min_{j\in OPT} \boldsymbol{c}^\top(\boldsymbol{\mu}_j - \boldsymbol{\mu}_i)]^{-2} + \sum_{i\in I} \mathrm{dist}(\boldsymbol{\mu}_i, P)^{-2}.$$

Then, any algorithm $\mathcal{A}$ that is δ- PAC-EXPLANATORY wrt $\mathcal{M}$ has a stopping time τ on the problem $(\nu, P, \boldsymbol{e}_1, m)$ that satisfies

$$\mathbb{E}_\nu[\tau] \geqslant \min_{(S,I)\in \textit{Valid-Partitions}} \frac{2}{15} \ln(\frac{1}{2\delta}) L(S, I).$$

The conditions $P = \mathbb{R} \times P'$ and $\boldsymbol{c} = \boldsymbol{e}_1$ decouple the reward and feasibility of the arms and hold in many applications. The other conditions on P remove pathological cases such as isolated points and nowhere dense sets with positive measure.

The lower bound is the solution of a constrained minimization problem over all the ways to distinguish the arms in OPT^c from OPT, i.e., $(S, I) \in$ Valid-Partitions. If we fix some $(S, I) \in$ Valid-Partitions, there are three main terms in the lower bound reflecting the difficulty of identifying arms as belonging to either OPT, S, or I, respectively. Essentially,

optimal arms must be shown to be feasible and to have reward greater than all arms in S, arms in S must be shown to have reward less than arms in OPT, and arms in I must be shown to be infeasible.

The key observation in the proof is that for a given problem $(\nu, P, \boldsymbol{c}, m)$, we can associate with an algorithm $\mathcal{A}$ a particular $(S, I) \in$ Valid-Partitions such that for every $i \in S$ ($i \in I$), it is likely that $\mathcal{A}$ puts $i \in \widehat{\mathrm{S}}$ ($i \in \widehat{\mathrm{I}}$). Then, using the notion of δ- PAC-EXPLANATORY, it suffices to analyze the difficulty of identifying each arm as belonging either to OPT, S, or I. The result follows by minimizing over $(S, I) \in$ Valid-Partitions.

We also state a similar lower bound for algorithms that are δ- PAC wrt $\mathcal{M}$.

Theorem 8. *Assume the conditions of Theorem 7. Define* $r_o = \min_{j \in OPT} \boldsymbol{c}^\top \boldsymbol{\mu}_j$, $r_s \coloneqq \max_{j \in OPT^c \cap FEAS} \boldsymbol{c}^\top \boldsymbol{\mu}_j$, *and* $L' \coloneqq$

$$\begin{aligned}
&\sum_{i \in INFEAS \cap SUBOPT} \min([r_o - \boldsymbol{c}^\top \boldsymbol{\mu}_i]^{-2},\ \mathrm{dist}(\boldsymbol{\mu}_i, P)^{-2}) \\
&+ \sum_{i \in OPT} \max([\boldsymbol{c}^\top \boldsymbol{\mu}_i - r_s]^{-2},\ \mathrm{dist}(\boldsymbol{\mu}_i, \partial P)^{-2}) \\
&+ \sum_{i \in OPT^c \cap FEAS} [\min_{j \in OPT} \boldsymbol{c}^\top (\boldsymbol{\mu}_j - \boldsymbol{\mu}_i)]^{-2} \\
&+ \sum_{i \in INFEAS \cap SUBOPT^c} \mathrm{dist}(\boldsymbol{\mu}_i, P)^{-2}.
\end{aligned}$$

Then, any algorithm $\mathcal{A}$ *that is* δ- *PAC wrt* $\mathcal{M}$ *has a stopping time* τ *on the problem* $(\nu, P, \boldsymbol{e}_1, m)$ *that satisfies*

$$\mathbb{E}_\nu[\tau] \geqslant \ln(\frac{1}{2.4\delta}) L'.$$

The bound in Theorem 8 suggests that δ- PAC algorithms must show that arms in OPT are feasible and have reward greater than every arm in $\mathrm{OPT}^c \cap \mathrm{FEAS}$, arms in $\mathrm{OPT}^c \cap \mathrm{FEAS}$ have reward less than arms in OPT, arms in $\mathrm{INFEAS} \cap \mathrm{SUBOPT}^c$ are infeasible, and, finally, arms in $\mathrm{INFEAS} \cap \mathrm{SUBOPT}$ are either infeasible or suboptimal.

Since any δ- PAC-EXPLANATORY algorithm wrt $\mathcal{M}$ is δ- PAC wrt $\mathcal{M}$, we expect the lower bound in Theorem 7 to be at least as large as the lower bound in Theorem 8, and this is indeed the case. The main difference between the bounds occurs in the terms corresponding to $i \in$ OPT. Essentially, in Theorem 7, it is required to show that every arm in

OPT has reward greater than all arms that are ruled out as suboptimal (i.e., belong to S), whereas in Theorem 8, these arms must only be shown to have reward greater than arms in $\text{FEAS} \cap \text{OPT}^c$. See the supplemental material for a more detailed discussion.

3.5 TF-LUCB: A Family of Algorithms for Top Feasible Arm Identification

In this section, we introduce an algorithm for the top feasible arm identification problem. To begin, we define some notation. Let $\widehat{\boldsymbol{\mu}}_{i,s}$ denote the empirical mean of arm i after s samples. Let $N_i(t) = \sum_{s=1}^{t-1} \mathbf{1}\{I_s = i\}$ denote the number of times that arm i has been selected up to round t. Let

$$U(t, \delta) = \sigma \sqrt{\frac{2\log(1/\delta) + 6\log\log(1/\delta) + 3\log\log(et)}{t}}$$

denote a confidence bound, which holds uniformly over time (see Lemma 20 in the supplemental material) (Kaufmann et al., 2016b). For the sake of simplicity, we assume henceforth that $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K \in B_{\frac{1}{2}}(\mathbf{0})$ and $P \subset B_{\frac{1}{2}}(\mathbf{0})$.

Challenge. As suggested by Theorem 8, a major challenge in designing a nearly optimal algorithm is how to rule out with nearly *optimal* sample complexity an arm i that is infeasible and whose reward $\boldsymbol{r}^\top \boldsymbol{\mu}_i$ is too small to be among the top m feasible arms (i.e., belongs to $\text{INFEAS} \cap \text{SUBOPT}$). In short, a nearly optimal algorithm must determine which is easier to show: that arm i is infeasible or that it has too small reward. Either of these can be arbitrarily more difficult to show than the other; for example, consider an infeasible arm with mean very close to the set P and a very small reward relative to the other arms. In this case, it is quite easy to show suboptimality, but very difficult to show infeasibility.

Algorithm. TF-LUCB is a family of algorithms, where each instance is specified by a subroutine TestF. $\text{TestF}(i, s)$ considers the first s pulls of arm i and returns True if arm i is feasible whp, returns False if i is infeasible whp, and otherwise returns ?, indicating "don't know." When the context makes it clear which distribution is involved, we simply write $\text{TestF}(s)$. TestF essentially solves what we will call the *set membership problem*, which we now define. In this problem, there is a distribution ξ over $\mathbb{R}^D$ with mean $\boldsymbol{\mu} \in \mathbb{R}^D$ and

a set $P \subset \mathbb{R}^D$. At round $t = 1, 2, \ldots$ an algorithm $\mathcal{B}$ observes $\boldsymbol{X}_t \sim \xi$. An algorithm $\mathcal{B}$ is associated with a stopping time τ wrt $(\boldsymbol{X}_t)_{t\in\mathbb{N}}$, and after τ rounds outputs True if it concludes that $\boldsymbol{\mu} \in P$ and False if it concludes that $\boldsymbol{\mu} \notin P$. We define the following class of set membership problems:

$$\mathcal{N} = \{(\xi, P) : \xi \text{ is a distribution over } \mathbb{R}^D \text{ with mean } \boldsymbol{\mu} \in B_{\frac{1}{2}}(\mathbf{0}), P \subset B_{\frac{1}{2}}(\mathbf{0}), P \neq \varnothing, \boldsymbol{\mu} \notin \partial P\}.$$

We defer our discussion of specific algorithms for the set membership problem until the next section.

Given a subroutine TestF, TF-LUCB is an adaptation of LUCB (Lower Upper Confidence Bound) from Kalyanakrishnan et al. (2012b) to the top feasible arm identification problem. TF-LUCB maintains three sets: arms F_t that are feasible whp, arms G_t that have not been determined whp to be feasible or infeasible, and arms $E_t \coloneqq F_t \cup G_t$ that have not been ruled out as infeasible whp. At round t, TF-LUCB considers TOP_t, the set of m arms that have not been ruled out as infeasible whp (i.e., belong to E_t) and have the top m estimated rewards. TF-LUCB uses $U_{\boldsymbol{c}}(t, \delta) \coloneqq U(t, \frac{\delta}{2K})$ for a confidence bound on the reward associated with an arm. If all of the arms in TOP_t are feasible whp, then it pulls an arm h_t in TOP_t with the smallest lower confidence bound. If only some of the arms in TOP_t are determined to be feasible whp, then to avoid oversampling optimal arms, it chooses the arm h_t instead by picking the arm in $\text{TOP}_t \cap G_t$ with the smallest lower confidence bound, i.e., an arm in the top empirical m for which it is still not determined whp whether it is feasible. We note that because $\text{TOP}_t \cap E_t^c = \varnothing$ by definition of TOP_t, when $\text{TOP}_t \nsubseteq F_t$, $\text{TOP}_t \cap G_t$ is nonempty so that the argmax operator in line 14 is well-defined. If there are arms outside of TOP_t that have not been ruled out as infeasible, then the algorithm pulls an additional arm l_t among these (in $\text{TOP}_t^c \cap E_t$) that maximizes an upper confidence bound on its reward. The algorithm terminates when it determines whp that each arm in TOP_t is feasible and has mean larger than arms in $\text{TOP}_t^c \cap E_t$, or that the arms in TOP_t are feasible and all other arms are infeasible.

For the sake of brevity, define the function $F(x, y) = x^{-2} \log(\log(x^{-2})y)$. Theorem 9 shows that TF-LUCB is δ- PAC-EXPLANATORY with a bound on τ that nearly matches

Algorithm 4 TF-LUCB: Top-m Feasible Lower Upper Confidence Bound algorithm

Input: TestF, sub-Gaussian norm bound σ, confidence δ
for $t = 1, 2, \ldots$ **do**
  $F_t \longleftarrow \{i \in [K] : \text{TestF}(i, N_i(t)) = \text{ True}\}$ // arms that are feasible whp
  $G_t \longleftarrow \{i \in [K] : \text{TestF}(i, N_i(t)) = ?\}$ // arms whose feasibility is unclear whp
  $E_t \longleftarrow F_t \cup G_t$ // arms that are not ruled out as infeasible whp
  $\text{TOP}_t \longleftarrow \arg\max_{Z \subset E_t, |Z| = \min(m, |E_t|)} \sum_{i \in Z} \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{i, N_i(t)}$
  if $\text{TOP}_t = F_t$ and $F_t = E_t$ **then**
    return $(\text{TOP}_t, \text{TOP}_t^c \cap E_t, E_t^c)$
  else if $\text{TOP}_t \subset F_t$ and $\min_{i \in \text{TOP}_t} \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{i, N_i(t)} - U_{\boldsymbol{c}}(N_i(t), \delta) \geqslant \max_{j \in \text{TOP}_t^c \cap E_t} \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{j, N_j(t)} + U_{\boldsymbol{c}}(N_j(t), \delta)$ **then**
    return $(\text{TOP}_t, \text{TOP}_t^c \cap E_t, E_t^c)$
  else if $\text{TOP}_t \subset F_t$ **then**
    $h_t = \arg\min_{i \in \text{TOP}_t} \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{i, N_i(t)} - U_{\boldsymbol{c}}(N_i(t), \delta)$
  else if $\text{TOP}_t \not\subset F_t$ **then**
    $h_t = \arg\min_{i \in \text{TOP}_t \cap G_t} \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{i, N_i(t)} - U_{\boldsymbol{c}}(N_i(t), \delta)$
  end if
  if $\text{TOP}_t^c \cap E_t \neq \varnothing$ **then**
    $l_t = \arg\max_{j \in \text{TOP}_t^c \cap E_t} \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{j, N_j(t)} + U_{\boldsymbol{c}}(N_j(t), \delta)$
    Pull arm l_t
  end if
  Pull arm h_t
end for

the lower bound.

Theorem 9. *Let $\delta \in (0, 1)$ and $(\nu, P, \boldsymbol{c}, m) \in \mathcal{M}$. Suppose that for any set membership problem $(\xi, R) \in \mathcal{N}$ where ξ is σ-sub-Gaussian and has mean $\boldsymbol{\mu}$, with probability at least $1 - \frac{\delta}{2K}$, TestF returns True only if $\boldsymbol{\mu} \in R$ and False only if $\boldsymbol{\mu} \in R^c$, and TestF uses at most $\eta(\xi, R)$ samples, where $\eta(\xi, R)$ is a deterministic function of ξ and R. For any $(S, I) \in$*

Valid-Partitions, define $\mathcal{U}(S, I) \coloneqq$

$$\sum_{i \in S} F(\min_{j \in OPT} \boldsymbol{c}^\top (\boldsymbol{\mu}_j - \boldsymbol{\mu}_i), \frac{K}{\delta}) + \sum_{i \in I} \eta(\nu_i, P)$$
$$+ \sum_{i \in OPT} \max(F(\min_{j \in S} \boldsymbol{c}^\top (\boldsymbol{\mu}_i - \boldsymbol{\mu}_j), \frac{K}{\delta}), \eta(\nu_i, P)).$$

Then, with probability at least $1 - \delta$*, TF-LUCB returns* $(\widehat{O}, \widehat{S}, \widehat{I})$ *such that* $\widehat{O} = OPT$*,* $(\widehat{S}, \widehat{I}) \in$ *Valid-Partitions, and*

$$\tau \leqslant \min_{(S,I) \in \textit{Valid-Partitions}} c\sigma^2 \mathcal{U}(S, I). \tag{3.1}$$

where c is a universal positive constant.

This upper bound has a very similar structure to the lower bound in Theorem 7. It is the solution of a constrained minimization problem over $(S, I) \in$ Valid-Partitions. One can interpret this form as saying that TF-LUCB finds the easiest way to solve a given instance of the top feasible arm identification problem. Ignoring doubly logarithmic factors, the upper bound on the reward-associated terms is loose by a factor of $\log(K)$.[1] Theorem 9 can be interpreted as a reduction of the top feasible arm identification problem to the set membership problem and in the next section we will discuss how various algorithms for the set membership problem affect the sample complexity of TF-LUCB.

In light of Theorem 9, it is instructive to consider a two-stage algorithm that first identifies the collection of feasible arms and then applies a best arm identification algorithm to the feasible arms. The drawback of this two-stage approach is that there may be suboptimal infeasible arms that are much easier to rule out as suboptimal rather than infeasible. Essentially, such a two-stage algorithm solves a problem instance by picking the $(S', I') \in$ Valid-Partitions such that $I' =$ INFEAS, whereas TF-LUCB adapts to the problem instance to choose the best $(S, I) \in$ Valid-Partitions. Thus, the sample complexity of such a two-stage algorithm is at least the sample complexity of TF-LUCB and can be arbitrarily larger than the sample complexity of TF-LUCB. To see this, consider a problem with an arm whose mean is very close to the boundary of P, but has very small reward relative to the other arms.

[1] We note that this logarithmic factor could be improved by adapting LUCB++ (Simchowitz et al., 2017a) instead of LUCB.

The proof of Theorem 9 considers the $(S, I) \in$ Valid-Partitions that minimizes (3.1) and analyzes the sample complexity of TF-LUCB to identify each arm as belonging to either OPT, S, or I. It is shown that at each round t, either h_t or l_t is a needy arm wrt to the sets OPT, S, and I (defined precisely in the supplemental material) in the sense that either it is necessary to determine whether it is feasible or it is necessary to improve our estimate of its reward.

Allowing for a tolerance: It is possible to extend TF-LUCB to allow for a tolerance on suboptimality or infeasibility (see supplemental material). For example, if a *suboptimality gap* of $\epsilon > 0$ is permitted, then $\widehat{\mathrm{O}}$ is correct if it satisfies $\forall i \in \widehat{\mathrm{O}}$, $i \in$ FEAS and $\boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_i + \epsilon \geqslant \min_{j \in \mathrm{OPT}} \boldsymbol{c}^\top \boldsymbol{\mu}_j$.

3.6 Three Instances of TF-LUCB

In this section, we consider three distinct general classes of sets and apply Theorem 9 to derive upper bounds for algorithms for each of these. To begin, we consider a general set P. Since there are in general no known computationally efficient algorithms for such a general setting, we then consider the computationally tractable and very rich class of polyhedra. For this setting, let $P = \{\boldsymbol{x} \in \mathbb{R}^D : A\boldsymbol{x} \leqslant \boldsymbol{b}\}$ denote a polyhedron, where $A \in \mathbb{R}^{M \times D}$ and $\boldsymbol{b} \in \mathbb{R}^M$. Let $\boldsymbol{a}_j^\top$ denote the jth row of A. By dividing each constraint j by $\|\boldsymbol{a}_j\|_2$, we can assume without loss of generality that $\|\boldsymbol{a}_j\|_2 = 1$ for all $j \in [M]$. Finally, we consider the common case where the polyhedron has orthogonal constraints, i.e., $\boldsymbol{a}_i^\top \boldsymbol{a}_j = 0$ for all $i \neq j \in [M]$, which arises for example when there is one constraint per coordinate. Note that in this case, it follows that $M \leqslant D$.

For a general set, we propose the TestF subroutine: TestF-B (see Algorithm 5). It controls $\|\widehat{\boldsymbol{\mu}}_{i,t} - \boldsymbol{\mu}_i\|_2$ with a confidence bound $U_{\mathrm{ball}}(t, \delta) \coloneqq 2U(t, \frac{\delta}{5^D 2K})$ that is constructed based on an ϵ-net argument. TestF-B returns True if the ball centered at $\widehat{\boldsymbol{\mu}}_{i,t}$ with radius $U_{\mathrm{ball}}(t, \delta)$ does not intersect P^c, False if this ball does not intersect P, and otherwise returns ?. The variant of TF-LUCB that uses TestF-B is called *TF-LUCB-B*.

For a polyhedron, we propose the subroutine TestF-CB, which also uses the confidence bound $U_{\mathrm{con}}(t, \frac{\delta}{2}) \coloneqq U(t, \frac{\delta}{4KM})$ (see Algorithm 6). If it determines that $\boldsymbol{\mu}_i$ satisfies all of the

Algorithm 5 TestF-B:

Input: arm index i, number of pulls t

if $\mathrm{dist}(\widehat{\boldsymbol{\mu}}_{i,t}, P^c) > U_{\mathrm{ball}}(t,\delta)$ **then**

  return True

else if $\mathrm{dist}(\widehat{\boldsymbol{\mu}}_{i,t}, P) > U_{\mathrm{ball}}(t,\delta)$ **then**

  return False

else

  return ?

end if

Algorithm 6 TestF-CB:

Input: arm index i, number of pulls t

if $A\widehat{\boldsymbol{\mu}}_{i,t} + U_{\mathrm{con}}(t,\frac{\delta}{2})\mathbf{1} \leqslant \boldsymbol{b}$ **then**

  return True

else if $\mathrm{dist}(\widehat{\boldsymbol{\mu}}_{i,t}, P) > U_{\mathrm{ball}}(t,\frac{\delta}{2})$ **then**

  return False

else

  return ?

end if

Algorithm 7 TestF-C:

Input: arm index i, number of pulls t

if $A\widehat{\boldsymbol{\mu}}_{i,t} + U_{\mathrm{con}}(t,\delta)\mathbf{1} \leqslant \boldsymbol{b}$ **then**

  return True

else if $A\widehat{\boldsymbol{\mu}}_{i,t} - U_{\mathrm{con}}(t,\delta)\mathbf{1} \not\leqslant \boldsymbol{b}$ **then**

  return False

else

  return ?

end if

	TF-LUCB-B	TF-LUCB-CB	TF-LUCB-C
$i \in$ OPT	$DF(\ \mathrm{dist}(\boldsymbol{\mu}_i, \partial P), \frac{K}{\delta})$	$F(\ \mathrm{dist}(\boldsymbol{\mu}_i, \partial P), \frac{KM}{\delta}))$	$F(\ \mathrm{dist}(\boldsymbol{\mu}_i, \partial P), \frac{KM}{\delta})$
$i \in$ INFEAS	$DF(\ \mathrm{dist}(\boldsymbol{\mu}_i, \partial P), \frac{K}{\delta})$	$DF(\ \mathrm{dist}(\boldsymbol{\mu}_i, \partial P), \frac{K}{\delta})$	$v_i F(\ \mathrm{dist}(\boldsymbol{\mu}_i, P), \frac{KM}{\delta})$

Table 3.1: Upper bounds on $\eta(\nu_i, P)$. For the case where P is a polyhedron, let $v_i = |\{j : \boldsymbol{a}_j^\top \boldsymbol{\mu}_i > b_j\}|$.

constraints whp, it returns True, if it determines that the ball centered at $\widehat{\boldsymbol{\mu}}_{i,t}$ with radius $U_{\mathrm{ball}}(t, \frac{\delta}{2})$ does not intersect P whp, it returns False, and otherwise it returns ?. The variant of TF-LUCB that uses TestF-CB is called *TF-LUCB-CB*.

Finally, a polyhedron with orthogonal constraints, we propose the subroutine TestF-C, which uses the confidence bound $U_{\mathrm{con}}(t,\delta)$ (see Algorithm 7). If it determines that $\boldsymbol{\mu}_i$ satisfies all of the constraints whp, it returns True, if it determines that $\boldsymbol{\mu}_i$ violates one of the constraints whp, it returns False, and otherwise it returns ?.

The following theorem establishes upper bounds for TF-LUCB-B, TF-LUCB-CB, and TF-LUCB-C.

Theorem 10. *Let $\delta \in (0,1)$ and $(\nu, P, \boldsymbol{c}, m) \in \mathcal{M}$. Then, with probability at least $1-\delta$,*

- *TF-LUCB-B returns* $(\widehat{O}, \widehat{S}, \widehat{I})$ *such that* $\widehat{O} = OPT$, $(\widehat{S}, \widehat{I}) \in$ *Valid-Partitions,* τ *is bounded as in* (3.1)*, and* $\eta(\nu_i, P)$ *is bounded as in Table 3.1.*

- *If* P *is a polyhedron, TF-LUCB-CB returns* $(\widehat{O}, \widehat{S}, \widehat{I})$ *such that* $\widehat{O} = OPT$, $(\widehat{S}, \widehat{I}) \in$ *Valid-Partitions,* τ *is bounded as in* (3.1)*, and* $\eta(\nu_i, P)$ *is bounded as in Table 3.1.*

- *If* P *is a polyhedron with orthogonal constraints, TF-LUCB-C returns* $(\widehat{O}, \widehat{S}, \widehat{I})$ *such that* $\widehat{O} = OPT$, $(\widehat{S}, \widehat{I}) \in$ *Valid-Partitions,* τ *is bounded as in* (3.1)*, and* $\eta(\nu_i, P)$ *is bounded as in Table 3.1.*

Ignoring doubly logarithmic factors, the terms related to determining feasibility for TF-LUCB-B are loose by a factor of $D\log(K)$ relative to our lower bound. When D is $O(\log K)$, then the bound is loose by a polylogarithmic factor. Since in many applications the dimension of the feedback is not very large, this bound is practically relevant. TF-LUCB-CB only requires $F(\ \mathrm{dist}(\boldsymbol{\mu}_i, \partial P), \frac{KM}{\delta})$ samples to show that an arm $i \in \mathrm{OPT}$ is feasible, which is a significant improvement over the corresponding term for TF-LUCB-B if M is polynomial in D. TF-LUCB-C differs from TF-LUCB-CB in the term for showing infeasibility. The term for determining that arms in I are infeasible is loose by a factor $v_i \log(KM)$, which can be much smaller than $D\log(K)$. In the common setting where the arms are two-dimensional with one coordinate encoding reward and the other a constraint, the upper bound is only loose by a logarithmic factor. See the supplemental material for an upper bound for TF-LUCB-C for the case of a general polyhedron.

3.7 Experiments

In this section, we demonstrate experimentally the effectiveness of our algorithms. We consider the task of identifying $\mathrm{OPT} \subset [K]$.

Synthetic Datasets: In each of the experiments, we use $\delta = 0.1$, the last coordinate determines the reward ($\boldsymbol{c} = (0, \ldots, 0, 1)^\top$), and the rest of the coordinates determine whether $\boldsymbol{x} \in P$. We consider two kinds of reward structures: linearly varying rewards $\boldsymbol{c}^\top \boldsymbol{\mu}_i = .95(1 - \frac{i}{100})$ and polynomially varying rewards $\boldsymbol{c}^\top \boldsymbol{\mu}_i = .95(1 - (\frac{i}{100})^{.3})$. In each trial, we randomly

Experiment	TF-LUCB-C	TF-LUCB-CB	TF-AE-C	FFAF-C	FFAF-CB
Simplex Arithmetic Linear	1.00	1.45	2.84	1.56	3.05
Simplex Arithmetic Polynomial	1.00	1.48	3.12	1.59	3.23
Simplex Groups Linear	1.00	1.25	2.78	1.29	2.14
Simplex Groups Polynomial	1.00	1.32	2.97	1.32	2.14
Ordered Groups Linear	1.00	1.04	1.93	1.15	1.16
Ordered Groups Polynomial	1.00	1.05	2.02	1.43	1.33
Crowdsourcing	1.00	N/A	2.15	2.88	N/A
Medical	1.00	N/A	1.12	4.52	N/A

Table 3.2: Number of samples required, relative to TF-LUCB-C, averaged over 50 trials.

permute the rewards among the arms in the sense that we take a random permutation $\sigma : [K] \longrightarrow [K]$ and set $\mu_{i,D}$ to $\mu_{\sigma(i),D}$.

In one set of experiments, we use 6-dimensional multivariate Gaussian distributions as arms with covariance matrix $\frac{1}{4}I$. We use a simplex $P = \{\boldsymbol{x} \in \mathbb{R}^6 : \sum_{i=1}^{5} x_i \leqslant 2, x_i \geqslant 0 \forall i \in [5]\}$. We consider one setting where there are four groups of arms $\boldsymbol{\mu}_{1:15,1:5} = (.1)^{\otimes 5}$, $\boldsymbol{\mu}_{16:30,1:5} = (.35)^{\otimes 5}$, $\boldsymbol{\mu}_{31:45,1:5} = (.45)^{\otimes 5}$, $\boldsymbol{\mu}_{46:60,1:5} = (-.1)^{\otimes 5}$. Only the arms in $[30]$ are feasible. In another setting, we consider arms with arithmetically changing values. In this setting, for $i \in [30]$, $\boldsymbol{\mu}_{i,1:5} = [(.1 + (\frac{2-0.05}{5} - .1)\frac{i}{30}]^{\otimes 5}$, for $i \in [45]\backslash[30]$, $\boldsymbol{\mu}_{i,1:5} = [2.05/5 + (3/5 - 2.05/5)\frac{i-30}{15}]^{\otimes 5}$, and for $i \in [60]\backslash[45]$, $\boldsymbol{\mu}_{i,1:5} = [-0.05 + (-.3 + 0.05)\frac{i-45}{15}]^{\otimes 5}$. Only the arms in $[30]$ are feasible. We use $\sqrt{1/4}$ as the sub-Gaussian norm for the arms.

In another set of experiments, we use 5-dimensional Bernoulli distributions. We use an ordered polyhedron $P = \{\boldsymbol{x} \in \mathbb{R}^5 : \boldsymbol{x}_i \leqslant \boldsymbol{x}_{i+1} \forall i \in [3]\}$. We consider a setting with three groups: $\boldsymbol{\mu}_{1:30,1:4} = (0.05, 0.35, 0.65, 0.95)^\top$, $\boldsymbol{\mu}_{31:40,1:4} = (0.95, 0.65, 0.35, 0.05)^\top$, and $\boldsymbol{\mu}_{41:50,1:4} = (.7, .6, .5, .4)^\top$. Only the arms in $[30]$ are feasible. We use 1 as the sub-Gaussian norm of the arms.

Crowdsourcing Application: We consider the task of finding the most accurate crowdsourcing workers subject to the constraint that they complete tasks at a suitable average speed. We use a crowdsourcing dataset collected by Venanzi et al. (2016) in which Amazon Mechanical Turk workers determine what kind of a statement a tweet makes regarding the weather: *(i)* positive, *(ii)* neutral, *(iii)* negative, *(iv)* unrelated, or *(v)* can't tell. We only

consider workers that have answered at least 100 questions, leaving a total of 21 workers. Here, $\mu_{i,1}$ is the probability of being correct and $\mu_{i,2}$ is the average amount of time required. We seek the top 3 most accurate workers who on average answer questions within 15 seconds. Whenever an algorithm pulls an arm corresponding to a worker, it samples a datapoint associated with that worker uniformly at random with replacement. We use the standard deviation of the speed measurements (135.86 sec) as the sub-Gaussian norm for the coordinate corresponding to the speed and 1 as the sub-Gaussian norm for the other coordinate. We use $\delta = 0.1$ and allow for a suboptimality gap of 0.05.

Clinical Trials Application: We examine the problem in clinical trials of finding the most effective drugs that also meet some safety threshold. We use data from Genovese et al. (2013) (see ARCR20 in week 16 in Table 2 and Table 3), which studies the drug secukinumab for treating rheumatoid arthritis. Each arm corresponds to a dosage level (25mg, 75mg, 150mg, 300mg, placebo) and has two attributes: the probability of being effective, $\mu_{i,1}$, and the probability of causing an infection or infestation, $\mu_{i,2}$. The dosage levels 25mg, 75mg, 150mg, and 300mg have averages $\boldsymbol{\mu}_1 = (.34, .259)^\top$, $\boldsymbol{\mu}_2 = (.469, .184)^\top$, $\boldsymbol{\mu}_3 = (.465, .209)^\top$, $\boldsymbol{\mu}_4 = (.537, .293)^\top$, respectively, and the placebo has average $\boldsymbol{\mu}_5 = (.36, .36)^\top$. In our experiment, whenever arm i is chosen two Bernoulli random variables with means given by $\boldsymbol{\mu}_i$ are drawn. We assume that a drug is acceptable if the probability of an infection is below .25, we set $m = 1$, and we allow for a suboptimality gap of 0.05. Thus, the correct answer is either arm 2 or arm 3. We use $\delta = 0.05$. We use 1 as the sub-Gaussian norm.

Algorithms: We consider our algorithms TF-LUCB-C and TF-LUCB-CB. We also consider Find-Feasible-Arms-First (FFAF), which is a two-stage algorithm that first determines which of the arms are feasible and then applies LUCB to the feasible arms to find the top arms. FFAF-CB uses TestF-CB to test feasibility, whereas FFAF-C uses TestF-C. We also implement an action elimination algorithm (TF-AE-C) that samples remaining arms in a round-robin fashion, eliminating an arm if it is determined using confidence bounds to be either suboptimal or infeasible. We only consider a variant that uses TestF-C since TF-AE-C has poor performance. For the experiments where $D = 2$, we only run the constraint based algorithms since the ϵ-net approach uses strictly worse (by a constant factor) confidence bounds.

Discussion of Results: Table 3.2 displays our results as the number of samples required, relative to TF-LUCB-C. All algorithms find a correct set of arms on every trial. TF-LUCB-C has the best sample complexity in all of the experiments, beating the FFAF algorithms by a substantial margin in many of them. In particular, FFAF-C requires nearly five times as many samples as TF-LUCB-C on the medical dataset and nearly three times as many samples on the crowdsourcing dataset. The performance gap between TF-LUCB and FFAF depends on the relative difficulty of showing arms to be suboptimal vs. infeasible. In particular, FFAF-C has poor performance on the real-world datasets because on the crowdsourcing dataset the sub-Gaussian norm for showing feasibility is large and on the medical dataset one of the suboptimal infeasible arms is very close to the boundary. TF-AE-C performs so poorly because each suboptimal feasible arm must be pulled until at least m arms are shown to be feasible and have larger reward than it.

3.8 Conclusion

We introduced a novel problem, top feasible arm identification: the first general pure exploration multi-armed bandit problem on constrained optimization. We argued that it has many real-world applications since in many settings there is multi-dimensional feedback and a natural goal is constrained optimization based on this feedback (e.g., safety and effectiveness in clinical trials); thus, we argue that our algorithms are of significant practical interest.

3.9 Chapter Appendix Outline and Notation

In this section, we provide an outline of the supplemental material and define some notation. In Section 3.10, we prove Theorem 7. In Section 3.11, we prove Theorem 9 and the main lemmas used in its proof. In Section 3.12, we prove Theorem 10, splitting it up into three distinct theorems. In Section 3.13, we discuss our conjecture that there is a small gap between δ- PAC and δ- PAC-EXPLANATORY algorithms; we also prove and discuss our lower bound for δ- PAC algorithms. In section 3.14, we prove a number of technical lemmas. In Section 3.15, we present and discuss a version of TF-LUCB that allows for tolerance of infeasibility

and suboptimality. In section 3.16, we provide pseudocode for TF-AE and FFAF.

Define $\mathcal{S}^{D-1} = \{\boldsymbol{x} \in \mathbb{R}^D : \|\boldsymbol{x}\|_2 = 1\}$. Define a function $d : (0,1) \times (0,1) \mapsto \mathbb{R}$ such that $d(x,y) \coloneqq x\log(\frac{x}{y}) + (1-x)\log(\frac{1-x}{1-y})$. Recall that if $U = \varnothing$, then we use the convention $\min_{x\in U} x = \infty$ and $\max_{x\in U} x = -\infty$.

3.10 Lower Bound

For the proof of Theorem 7, we introduce the following notation. For a given problem $(\nu, P, \boldsymbol{c}, m)$, define

$$\begin{aligned}
&\mathrm{FEAS}(\nu,P,\boldsymbol{c},m) = \{i \in [K] : \boldsymbol{\mu}_i \in P\},\ \mathrm{INFEAS}(\nu,P,\boldsymbol{c},m) = \mathrm{FEAS}(\nu,P,\boldsymbol{c},m)^c,\\
&\mathrm{OPT}(\nu,P,\boldsymbol{c},m) = \{i \in \mathrm{FEAS}(\nu,P,\boldsymbol{c},m) : \boldsymbol{c}^\top\boldsymbol{\mu}_i \geqslant \max^{(m)}_{j\in\mathrm{FEAS}(\nu,P,\boldsymbol{c},m)} \boldsymbol{c}^\top\boldsymbol{\mu}_j\},\\
&\mathrm{SUBOPT}(\nu,P,\boldsymbol{c},m) = \{i \in [K] : \boldsymbol{c}^\top\boldsymbol{\mu}_i < \max^{(m)}_{j\in\mathrm{FEAS}(\nu,P,\boldsymbol{c},m)} \boldsymbol{c}^\top\boldsymbol{\mu}_j\}.
\end{aligned}$$

Proof of Theorem 7. **Step 1: Pick a good partition of the arms.** Fix $\delta > 0$. Let $(\nu, P, \boldsymbol{c}, m)$ satisfy the hypotheses of the theorem statement. In the interest of brevity, abbreviate

$$\begin{aligned}
&\mathrm{FEAS} \coloneqq \mathrm{FEAS}(\nu,P,\boldsymbol{c},m),\ \mathrm{INFEAS} \coloneqq \mathrm{INFEAS}(\nu,P,\boldsymbol{c},m),\\
&\mathrm{OPT} \coloneqq \mathrm{OPT}(\nu,P,\boldsymbol{c},m),\ \mathrm{SUBOPT} \coloneqq \mathrm{SUBOPT}(\nu,P,\boldsymbol{c},m).
\end{aligned}$$

Let $\mathcal{A}$ denote a δ- PAC-EXPLANATORY algorithm wrt $\mathcal{M}$ with stopping time τ.

We claim that there exists $(S, I) \in$ Valid-Partitions that satisfies the following property:

$$i \in S \implies \Pr_\nu(i \in \widehat{\mathrm{S}}) \geqslant \frac{1-\delta}{2}; \qquad i \in I \implies \Pr_\nu(i \in \widehat{\mathrm{I}}) \geqslant \frac{1-\delta}{2}. \tag{3.2}$$

As an intermediate step, we claim that for every $i \in \mathrm{OPT}^c$,

$$\max(\Pr_\nu(i \in \widehat{\mathrm{I}}), \Pr_\nu(i \in \widehat{\mathrm{S}})) \geqslant \frac{1-\delta}{2}. \tag{3.3}$$

To see this, fix $i \in \mathrm{OPT}^c$. Define the events

$$\begin{aligned}
B &= \{\widehat{\mathrm{O}} = \mathrm{OPT}, (\widehat{\mathrm{S}}, \widehat{\mathrm{I}}) \in \text{Valid-Partitions}\},\\
B_1 &= B \cap \{i \in \widehat{\mathrm{S}}\},\\
B_2 &= B \cap \{i \in \widehat{\mathrm{I}}\}.
\end{aligned}$$

Note that $B = B_1 \cup B_2$ and $B_1 \cap B_2 = \varnothing$. Since $\mathcal{A}$ is δ- PAC-EXPLANATORY wrt $\mathcal{M}$,

$$\begin{aligned} 1 - \delta &\leqslant \Pr_\nu(B) \\ &= \Pr_\nu(B_1) + \Pr_\nu(B_2) \\ &\leqslant \Pr_\nu(i \in \widehat{\mathrm{S}}) + \Pr_\nu(i \in \widehat{\mathrm{I}}) \\ &\leqslant 2 \max(\Pr_\nu(i \in \widehat{\mathrm{S}}), \Pr_\nu(i \in \widehat{\mathrm{I}})). \end{aligned}$$

This establishes the claim in (3.3). Furthermore, note that if $i \in \mathrm{OPT}^c \setminus \mathrm{INFEAS} = \mathrm{SUBOPT} \cap \mathrm{FEAS}$, then $B_2 = \varnothing$, so that

$$\Pr_\nu(i \in \widehat{\mathrm{S}}) \geqslant \frac{1-\delta}{2}. \tag{3.4}$$

Similarly, if $i \in \mathrm{OPT}^c \setminus \mathrm{SUBOPT}$, then $B_1 = \varnothing$, so that

$$\Pr_\nu(i \in \widehat{\mathrm{I}}) \geqslant \frac{1-\delta}{2}. \tag{3.5}$$

Define

$$\begin{aligned} S &= \{i \in \mathrm{SUBOPT} : \Pr_\nu(i \in \widehat{\mathrm{S}}) \geqslant \frac{1-\delta}{2}\} \\ I &= \mathrm{INFEAS} \setminus \{i \in \mathrm{SUBOPT} : \Pr_\nu(i \in \widehat{\mathrm{S}}) \geqslant \frac{1-\delta}{2}\}. \end{aligned}$$

We claim that $(S, I) \in$ Valid-Partitions. Clearly, $S \subset \mathrm{SUBOPT}$, $I \subset \mathrm{INFEAS}$, $S \cap I = \varnothing$, and $S \cup I \subset \mathrm{OPT}^c$. Therefore, it suffices to show that $\mathrm{OPT}^c \subset S \cup I$. Let $i \in \mathrm{OPT}^c$. If $i \in \mathrm{INFEAS}$, then either $i \in I$ or $i \in S$, so suppose that $i \notin \mathrm{INFEAS}$. Then, $i \in \mathrm{OPT}^c \setminus \mathrm{INFEAS} = \mathrm{SUBOPT} \cap \mathrm{FEAS} \subset S$ by (3.4). Thus, the claim that $(S, I) \in$ Valid-Partitions follows.

We claim that (S, I) has the property (3.2). Let $i \in S$. By definition of S, $\Pr_\nu(i \in \widehat{\mathrm{S}}) \geqslant \frac{1-\delta}{2}$. Next, let $i \in I$. If $i \in \mathrm{SUBOPT}$, then $i \in I \subset \mathrm{INFEAS}$ and $i \notin S$ imply that $\Pr_\nu(i \in \widehat{\mathrm{S}}) < \frac{1-\delta}{2}$. Then, by (3.3) $\Pr_\nu(i \in \widehat{\mathrm{I}}) \geqslant \frac{1-\delta}{2}$. If $i \notin \mathrm{SUBOPT}$, then (3.5) implies that $\Pr_\nu(i \in \widehat{\mathrm{I}}) \geqslant \frac{1-\delta}{2}$. Thus, the claim follows.

Next, we outline the rest of our proof. For the rest of the proof, the S and I that we constructed are fixed. Using the fact that $\tau = \sum_{i=1}^{K} N_i(\tau)$, we will show that for this choice

of S and I,

$$\mathbb{E}_\nu[\tau] = \sum_{i=1}^{K} \mathbb{E}_\nu[N_i(\tau)] \geqslant \frac{1}{15}\ln(\frac{1}{2\delta})[\sum_{i\in \mathrm{OPT}} \max([\min_{j\in S} \boldsymbol{c}^\top(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)]^{-2},\ \mathrm{dist}(\boldsymbol{\mu}_i, \partial P)^{-2}) \quad (3.6)$$

$$+ \sum_{i\in S}[\min_{j\in \mathrm{OPT}} \boldsymbol{c}^\top(\boldsymbol{\mu}_j - \boldsymbol{\mu}_i)]^{-2} + \sum_{i\in I} \mathrm{dist}(\boldsymbol{\mu}_i, P)^{-2}]. \quad (3.7)$$

To this end, we lower bound $\mathbb{E}_\nu[N_i(\tau)]$ for each of the distinct cases ($i \in \mathrm{OPT}$, $i \in S$, $i \in I$). To do this, we construct a related problem by modifying one of the distributions and applying Lemma 11. The result will follow by taking the minimum of the right-hand side of (3.7) over all $(S', I') \in$ Valid-Partitions.

In each of the next steps, we will define a new problem to obtain a lower bound. To avoid notational clutter, we will redefine the symbols $\boldsymbol{\mu}'_i$, ν'_i, and $\nu^{(i)}$ in each step. The context should make their meaning clear.

Step 2.a: reward bound for $i \in$ OPT. Fix $i \in \mathrm{OPT}$. First, we show that

$$\mathbb{E}_\nu[N_i(\tau)] \geqslant \frac{2}{15}\ln(\frac{1}{2\delta})[\min_{j\in S} \boldsymbol{c}^\top(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j) + \epsilon]^{-2}$$

for a sufficiently small $\epsilon > 0$. If $S = \varnothing$, $\min_{j\in S} \boldsymbol{c}^\top(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j) = -\infty$ by definition and there is nothing to show. So, suppose that $S \neq \varnothing$. Define

$$j_0 = \arg\max_{j\in S} \boldsymbol{c}^\top \boldsymbol{\mu}_j. \quad (3.8)$$

Define for all $j \in [K]$

$$\boldsymbol{\mu}'_j = \begin{cases} \begin{pmatrix} \boldsymbol{\mu}_{j_0,1} - \epsilon \\ \boldsymbol{\mu}_{i,2:D} \end{pmatrix} & \text{if } j = i \\ \boldsymbol{\mu}_j & \text{if } j \neq i \end{cases}$$

$$\nu'_j = N(\boldsymbol{\mu}'_j, I_D).$$

where $\epsilon > 0$ is chosen sufficiently small such that for all $\delta \in [0, \epsilon)$, $\boldsymbol{c}^\top\boldsymbol{\mu}'_i + \delta \neq \boldsymbol{c}^\top\boldsymbol{\mu}'_j$ for all $j \neq i$ (which is possible since $\boldsymbol{c}^\top\boldsymbol{\mu}_l \neq \boldsymbol{c}^\top\boldsymbol{\mu}_k$ for all $l \neq k \in [K]$). Define $\nu^{(i)} = (\nu'_1, \ldots, \nu'_K)$ and consider the problem $(\nu^{(i)}, P, \boldsymbol{c}, m)$. We claim that $(\nu^{(i)}, P, \boldsymbol{c}, m) \in \mathcal{M}$. Since $\boldsymbol{\mu}_i \notin \partial P$ and $\partial P = \partial(\mathbb{R} \times P') = \mathbb{R} \times \partial P'$ for some $P' \subset \mathbb{R}^{D-1}$, $\boldsymbol{\mu}'_i \notin \partial P$. Further, by construction,

$\boldsymbol{c}^\top \boldsymbol{\mu}'_i \neq \boldsymbol{c}^\top \boldsymbol{\mu}'_j$ for all $j \neq i$. Thus, none of the arms have means on the boundary of P and all of the rewards of the arms are distinct, so $(\nu^{(i)}, P, \boldsymbol{c}, m) \in \mathcal{M}$.

In the interest of brevity, abbreviate

$$\mathrm{FEAS}_i \coloneqq \mathrm{FEAS}(\nu^{(i)}, P, \boldsymbol{c}, m),\ \mathrm{SUBOPT}_i \coloneqq \mathrm{SUBOPT}(\nu^{(i)}, P, \boldsymbol{c}, m).$$

We claim that $j_0 \notin \mathrm{SUBOPT}_i$. Suppose $j_0 \in \mathrm{FEAS}$. Then,

$$\begin{aligned}
\boldsymbol{c}^\top \boldsymbol{\mu}'_{j_0} &= \boldsymbol{c}^\top \boldsymbol{\mu}_{j_0} \\
&= \max_{l \in S} \boldsymbol{c}^\top \boldsymbol{\mu}_l && (3.9)\\
&\geqslant \max_{l \in \mathrm{SUBOPT} \cap \mathrm{FEAS}} \boldsymbol{c}^\top \boldsymbol{\mu}_l && (3.10)\\
&= \max^{(m+1)}_{l \in \mathrm{FEAS}} \boldsymbol{c}^\top \boldsymbol{\mu}_l && (3.11)\\
&= \max^{(m)}_{l \in \mathrm{FEAS}_i} \boldsymbol{c}^\top \boldsymbol{\mu}'_l && (3.12)
\end{aligned}$$

where line (3.9) follows from (3.8), line (3.10) follows from $S \supset \mathrm{SUBOPT} \cap \mathrm{FEAS}$, (3.11) follows from $j_0 \in \mathrm{FEAS}$ by assumption, and (3.12) follows from the fact that $j_0 \in \mathrm{FEAS}$ and the only difference between ν and $\nu^{(i)}$ is in the ith arm, which now has reward less than the j_0th arm. Thus, if $j_0 \in \mathrm{FEAS}$, then $j_0 \notin \mathrm{SUBOPT}_i$.

On the other hand, if $j_0 \notin \mathrm{FEAS}$, then

$$\begin{aligned}
\boldsymbol{c}^\top \boldsymbol{\mu}'_{j_0} &= \max_{l \in S} \boldsymbol{c}^\top \boldsymbol{\mu}'_l && (3.13)\\
&> \boldsymbol{c}^\top \boldsymbol{\mu}'_i && (3.14)\\
&> \max_{l \in \mathrm{SUBOPT} \cap \mathrm{FEAS}} \boldsymbol{c}^\top \boldsymbol{\mu}'_l && (3.15)
\end{aligned}$$

where line (3.13) follows from (3.8) and $\boldsymbol{\mu}'_l = \boldsymbol{\mu}_l$ for all $l \in S$, line (3.14) follows from $\boldsymbol{\mu}'_i$ is defined to satisfy $\boldsymbol{c}^\top \boldsymbol{\mu}'_i > \max_{l : \boldsymbol{c}^\top \boldsymbol{\mu}'_l < \boldsymbol{c}^\top \boldsymbol{\mu}'_{j_0}} \boldsymbol{c}^\top \boldsymbol{\mu}'_l$, and line (3.15) follows from $S \supset$ $\mathrm{SUBOPT} \cap \mathrm{FEAS}$, $j_0 \notin \mathrm{FEAS}$, and $\{\boldsymbol{c}^\top \boldsymbol{\mu}_l | prime\}_{l \in [K]}$ distinct. (3.15) implies that

$$\max^{(m)}_{l \in \mathrm{FEAS}_i} \boldsymbol{c}^\top \boldsymbol{\mu}'_l = \boldsymbol{c}^\top \boldsymbol{\mu}'_i < \boldsymbol{c}^\top \boldsymbol{\mu}'_{j_0}$$

so that $j_0 \notin \mathrm{SUBOPT}_i$. This establishes the claim that $j_0 \notin \mathrm{SUBOPT}_i$.

Consider the event $B = \{j_0 \in \widehat{\mathrm{S}}\}$. Then, since $\mathcal{A}$ is δ- PAC-EXPLANATORY wrt to $\mathcal{M}$, $(\nu^{(i)}, P, \boldsymbol{c}, m) \in \mathcal{M}$, and arm $j_0 \notin \mathrm{SUBOPT}_i$, we have that

$$\Pr_{\nu^{(i)}}(B) \leqslant \Pr_{\nu^{(i)}}(\widehat{\mathrm{S}} \not\subset \mathrm{SUBOPT}_i) \leqslant \delta. \tag{3.16}$$

Further, by construction of S,

$$\Pr_\nu(B) \geqslant \frac{1-\delta}{2}. \tag{3.17}$$

Then,

$$\begin{aligned}
\frac{1}{2}[\boldsymbol{c}^\top(\boldsymbol{\mu}_i - \boldsymbol{\mu}_{j_0}) + \epsilon]^2 \mathbb{E}_\nu[N_i(\tau)] &= \mathrm{KL}(\nu_i, \nu_i')\mathbb{E}_\nu[N_i(\tau)] && (3.18)\\
&\geqslant d(\Pr_\nu(B), \Pr_{\nu^{(i)}}(B)) && (3.19)\\
&\geqslant d(\Pr_\nu(B), \delta) && (3.20)\\
&\geqslant d(\frac{1-\delta}{2}, \delta) && (3.21)\\
&\geqslant \frac{1}{15}\ln(\frac{1}{2\delta}). && (3.22)
\end{aligned}$$

Line (3.18) follows by the formula for the KL-divergence of two multivariate normal distributions, (3.19) follows by Lemma 11, (3.20) follows since $y \mapsto d(x, y)$ is decreasing when $x > y$, (3.16), (3.17), and $\delta < .1$, (3.21) follows since $x \mapsto d(x, y)$ is increasing when $x > y$, (3.16), (3.17), and $\delta < .1$, and (3.22) follows by Lemma (17). The claim follows by rearranging the inequality.

Step 2.b: feasibility bound for $i \in$ OPT. Next, we show that for sufficiently small $\epsilon > 0$,

$$\mathbb{E}_\nu[N_i(\tau)] \geqslant \frac{2}{15}\ln(\frac{1}{2\delta})[\ \mathrm{dist}(\boldsymbol{\mu}_i, \partial P) + \epsilon]^{-2}.$$

Since P is nonempty and $P \neq \mathbb{R}^D$, by Lemma 16 ∂P is nonempty. Since in addition ∂P is closed, by Lemma 12, there exists $\boldsymbol{\tau}_i \in \mathrm{Proj}_{\partial P}(\boldsymbol{\mu}_i)$. Since $\boldsymbol{\tau}_i \in \partial P$, by the assumptions of the Theorem on P, for all $\epsilon > 0$, $B_\epsilon(\tau) \cap (P^c)^\circ \neq \varnothing$. Thus, for any $\epsilon > 0$, there exists a direction $\boldsymbol{v} \in \mathbb{R}^D$ with $\|\boldsymbol{v}\|_2 = 1$ such that $\boldsymbol{\tau}_i + \epsilon\boldsymbol{v} \in (P^c)^\circ$. Further, since by the assumptions of the Theorem on P, $P = \mathbb{R} \times P'$ for some $P' \subset \mathbb{R}^{D-1}$, we can choose $\boldsymbol{v}$ such that $v_1 = 0$.

Define for $j \in [K]$

$$\boldsymbol{\mu}_j' = \begin{cases} \boldsymbol{\tau}_i + \epsilon\boldsymbol{v} & \text{if } j = i \\ \boldsymbol{\mu}_j & \text{if } j \neq i \end{cases}$$

$$\nu_j' = N(\boldsymbol{\mu}_j', I_D).$$

Define $\nu^{(i)} = (\nu'_1, \ldots, \nu'_K)$ and consider the problem $(\nu^{(i)}, P, \boldsymbol{c}, m)$. We claim that $(\nu^{(i)}, P, \boldsymbol{c}, m) \in \mathcal{M}$. Since $\boldsymbol{\mu}'_i \in (P^c)^\circ$, $\boldsymbol{\mu}'_i \notin \partial P$. Therefore, it suffices to show that $\boldsymbol{c}^\top \boldsymbol{\mu}'_i \neq \boldsymbol{c}^\top \boldsymbol{\mu}'_j$ for all $j \neq i$. To show this, it suffices to show that $\tau_{i,1} = \mu_{i,1}$ since then it follows by our choice of $\boldsymbol{v}$, $\boldsymbol{c} = \boldsymbol{e}_1$, and the fact that for all $j \neq i$, $\boldsymbol{c}^\top \boldsymbol{\mu}_j \neq \boldsymbol{c}^\top \boldsymbol{\mu}_i$. Towards a contradiction, suppose that $\mu_{i,1} \neq \tau_{i,1}$. Define

$$\tau'_{i,j} = \begin{cases} \tau_{i,j} & : j \neq 1 \\ \mu_{i,1} & : \text{ otherwise} \end{cases}.$$

Recall that $P = \mathbb{R} \times P'$ for some $P' \subset \mathbb{R}^{D-1}$ and observe that $\partial P = \partial \mathbb{R} \times P' \cup \mathbb{R} \times \partial P' = \mathbb{R} \times \partial P'$. Thus, $\boldsymbol{\tau}_i \in \partial P$ implies that $\boldsymbol{\tau}'_i \in \partial P$. Further, $\|\boldsymbol{\tau}'_i - \boldsymbol{\mu}_i\|_2 < \|\boldsymbol{\tau}_i - \boldsymbol{\mu}_i\|_2$, which is a contradiction to $\boldsymbol{\tau}_i \in \partial P$. Thus, the claim follows and hence $(\nu^{(i)}, P, \boldsymbol{c}, m) \in \mathcal{M}$.

In the interest of brevity, abbreviate

$$\mathrm{FEAS}_i \coloneqq \mathrm{FEAS}(\nu^{(i)}, P, \boldsymbol{c}, m), \ \mathrm{OPT}_i \coloneqq \mathrm{OPT}(\nu_P^{(i)}, P, \boldsymbol{c}, m).$$

Define the event $B = \{i \in \widehat{\mathrm{O}}\}$. Then, $i \notin \mathrm{FEAS}_i$, so that the event B implies that the algorithm $\mathcal{A}$ makes a mistake. Since $\mathcal{A}$ is δ- PAC-EXPLANATORY wrt $\mathcal{M}$, $\Pr_{\nu^{(i)}}(B) \leqslant \Pr_{\nu^{(i)}}(\widehat{\mathrm{O}} \not\subseteq \mathrm{OPT}_i) \leqslant \delta$. Further, since $i \in \mathrm{OPT}$ and $\mathcal{A}$ is δ- PAC-EXPLANATORY wrt $\mathcal{M}$,

$$\Pr_\nu(B) \geqslant \Pr_\nu(\widehat{\mathrm{O}} = \mathrm{OPT}) \geqslant 1 - \delta \geqslant \frac{1-\delta}{2}.$$

Thus,

$$\begin{aligned} \frac{1}{2}(\ \mathrm{dist}(\boldsymbol{\mu}_i, \partial P) + \epsilon)^2 \mathbb{E}_\nu[N_i(\tau)] &= \frac{1}{2}(\|\boldsymbol{\tau}_i - \boldsymbol{\mu}_i\|_2 + \epsilon)^2 \mathbb{E}_\nu[N_i(\tau)] && (3.23) \\ &\geqslant \frac{1}{2}\|\boldsymbol{\tau}_i + \epsilon \boldsymbol{v} - \boldsymbol{\mu}_i\|_2^2 \, \mathbb{E}_\nu[N_i(\tau)] && (3.24) \\ &= \mathrm{KL}(\nu_i, \nu'_i) \mathbb{E}_\nu[N_i(\tau)] && (3.25) \\ &\geqslant \frac{1}{15} \ln(\frac{1}{2\delta}). && (3.26) \end{aligned}$$

Line (3.23) follows by the definition of $\boldsymbol{\tau}_i$, line (3.24) follows by the triangle inequality and $\|v\|_2 = 1$, line (3.25) follows by the definition of the KL divergence for multivariate normal distributions, and line (3.26) follows by a similar series of inequalities as (3.18)-(3.22).

Step 3: $i \in S$**.** If $S = \varnothing$, then there is nothing to show in this step. So, suppose that $S \neq \varnothing$. Then, $S \neq \varnothing$ implies that there are at least m feasible arms. Let $j_0 \in [K]$ such

that $\boldsymbol{\mu}_{j_0} \in P$ and

$$\boldsymbol{c}^\top \boldsymbol{\mu}_{j_0} = \min_{l \in \mathrm{OPT}} \boldsymbol{c}^\top \boldsymbol{\mu}_l.$$

Define for $j \in [K]$

$$\boldsymbol{\mu}'_j = \begin{cases} \begin{pmatrix} \mu_{i,1} + \mu_{j_0,1} + \epsilon \\ \boldsymbol{\mu}_{i,2:D} \end{pmatrix} & \text{if } j = i \\ \boldsymbol{\mu}_j & \text{if } j \neq i \end{cases}$$

$$\nu'_j = N(\boldsymbol{\mu}'_j, I_D).$$

where $\epsilon > 0$ is chosen sufficiently small so that for any $\delta \in [0, \epsilon)$, $\boldsymbol{c}^\top \boldsymbol{\mu}'_i - \delta \neq \boldsymbol{c}^\top \boldsymbol{\mu}'_j$ for all $j \neq i$ (which is possible since $\boldsymbol{c}^\top \boldsymbol{\mu}_l \neq \boldsymbol{c}^\top \boldsymbol{\mu}_k$ for all $l \neq k \in [K]$). Define $\nu^{(i)} = (\nu'_1, \ldots, \nu'_K)$ and consider the problem $(\nu^{(i)}, P, \boldsymbol{c}, m)$. It follows that $(\nu^{(i)}, P, \boldsymbol{c}, m) \in \mathcal{M}$ by a similar argument that showed in Step 2.a that when $i \in \mathrm{OPT}$, $(\nu^{(i)}, P, \boldsymbol{c}, m) \in \mathcal{M}$.

In the interest of brevity, abbreviate

$$\mathrm{SUBOPT}_i \coloneqq \mathrm{SUBOPT}(\nu^{(i)}, P, \boldsymbol{c}, m).$$

Define $B = \{i \in \widehat{\mathrm{S}}\}$. Note that arm $i \notin \mathrm{SUBOPT}_i$ by construction. Thus, since $\mathcal{A}$ is δ- PAC-EXPLANATORY wrt $\mathcal{M}$, we have that $\Pr_{\nu^{(i)}}(B) \leqslant \delta$. Further, by construction of S, $\Pr_\nu(B) \geqslant \frac{1-\delta}{2}$. Therefore, by a similar series of inequalities as (3.18)-(3.22), it follows that

$$\frac{1}{15} \ln(\frac{1}{2\delta}) \leqslant \frac{1}{2}[\boldsymbol{c}^\top(\boldsymbol{\mu}_{j_0} - \boldsymbol{\mu}_i) + \epsilon]^2 \mathbb{E}_\nu[N_i(\tau)]. \tag{3.27}$$

Step 4: $i \in I$**.** Since $P \neq \mathbb{R}^D$ and P is nonempty, by Lemma 16 ∂P is nonempty. Since in addition ∂P is closed, by Lemma 12, there exists $\boldsymbol{\tau}_i \in \mathrm{Proj}_{\partial P}(\boldsymbol{\mu}_i)$. By the assumptions of the Theorem on P, since $\boldsymbol{\tau}_i \in \partial P$, for every $\epsilon > 0$, $B_\epsilon(\boldsymbol{\tau}_i) \cap P^\circ \neq \varnothing$. Thus, for sufficiently small $\epsilon > 0$, there exists a direction $\boldsymbol{v} \in \mathbb{R}^D$ with $\|\boldsymbol{v}\|_2 = 1$ such that $\boldsymbol{\tau}_i + \epsilon \boldsymbol{v} \in P^\circ$. Since by the assumptions of the Theorem on P, $P = \mathbb{R} \times P'$ for some $P' \subset \mathbb{R}^{D-1}$, we can choose $\boldsymbol{v}$

such that $v_1 = 0$. Define for $j \in [K]$

$$\boldsymbol{\mu}'_j = \begin{cases} \boldsymbol{\tau}_i + \epsilon \boldsymbol{v} & \text{if } j = i \\ \boldsymbol{\mu}_j & \text{if } j \neq i \end{cases}$$
$$\nu'_j = N(\boldsymbol{\mu}'_j, I_D).$$

Define $\nu^{(i)} = (\nu'_1, \ldots, \nu'_K)$ and consider the problem $(\nu^{(i)}, P, \boldsymbol{c}, m)$. It follows that $(\nu^{(i)}, P, \boldsymbol{c}, m) \in \mathcal{M}$ by a similar argument that showed in step 2.b that when $i \in \text{OPT}$, $(\nu^{(i)}, P, \boldsymbol{c}, m) \in \mathcal{M}$.

In the interest of brevity, abbreviate

$$\text{INFEAS}_i \coloneqq \text{INFEAS}(\nu^{(i)}, P, \boldsymbol{c}, m).$$

Define the event $B = \{i \in \widehat{\text{I}}\}$. Observe that $i \notin \text{INFEAS}_i$. Thus, since $\mathcal{A}$ is δ- PAC-EXPLANATORY wrt $\mathcal{M}$,

$$\text{Pr}_{\nu^{(i)}}(B) \leqslant \text{Pr}_{\nu^{(i)}}(\widehat{\text{I}} \not\subseteq \text{INFEAS}_i) \leqslant \delta.$$

Further, by construction of I, $\text{Pr}_\nu(i \in \widehat{\text{I}}) \geqslant \frac{1-\delta}{2}$. Therefore, by a similar series of inequalities as (3.23)-(3.26), it follows that

$$\frac{1}{15} \ln(\frac{1}{2\delta}) \leqslant \frac{1}{2}(\ \text{dist}(\boldsymbol{\mu}_i, P) + \epsilon)^2 \mathbb{E}_\nu[N_i(\tau)]. \tag{3.28}$$

Step 5: Putting it together. Using $\mathbb{E}_\nu[\tau] = \sum_{i=1}^K \mathbb{E}_\nu[N_i(\tau)]$ and inequalities (3.22), (3.26), (3.27), and (3.28), we establish for all sufficiently small $\epsilon > 0$,

$$\begin{aligned} \mathbb{E}_\nu[\tau] \geqslant & \frac{2}{15} \ln(\frac{1}{2\delta})[\sum_{i \in \text{OPT}} \max([\min_{j \in S} \boldsymbol{c}^\top(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j) + \epsilon]^{-2}, [\ \text{dist}(\boldsymbol{\mu}_i, \partial P) + \epsilon]^{-2}) \\ & + \sum_{i \in S}[\min_{j \in \text{OPT}} \boldsymbol{c}^\top(\boldsymbol{\mu}_j - \boldsymbol{\mu}_i) + \epsilon]^{-2} + \sum_{i \in I}[\ \text{dist}(\boldsymbol{\mu}_i, P) + \epsilon]^{-2}]. \end{aligned}$$

Since this bound holds for all $\epsilon > 0$ sufficiently small, letting $\epsilon \longrightarrow 0$ on the RHS of the above inequality establishes (3.7).

□

3.11 Proof of Theorem 9

To begin, we introduce some notation. Fix $(S, I) \in$ Valid-Partitions. We will bound the number of samples required to identify each arm as belonging to either OPT, S, or I. Define

$$d(S) = \frac{\min_{i \in \mathrm{OPT}} \boldsymbol{c}^\top \boldsymbol{\mu}_i + \max_{j \in S} \boldsymbol{c}^\top \boldsymbol{\mu}_j}{2}.$$

If either $|\,\mathrm{FEAS}\,| < m$ or $S = \varnothing$, then define $d(S) \coloneqq -\infty$. Next, we introduce a notion, which captures when arm i needs to be pulled more. Define for all $i \in [K]$,

$$\begin{aligned}\mathrm{NEEDY}_i^t(S, I) = [\{i \in \mathrm{OPT}\} \wedge (\{i \in G_t\} \vee \{\boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{i,N_i(t)} - U_{\boldsymbol{c}}(N_i(t), \delta) \leqslant d(S)\})] \\ \vee [\{i \in S\} \wedge \{\boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{i,N_i(t)} + U_{\boldsymbol{c}}(N_i(t), \delta) \geqslant d(S)\}] \\ \vee [\{i \in I\} \wedge \{i \in G_t\}]\end{aligned}$$

In words, if arm i is optimal, then it needs to be pulled more if either it has not been determined whp that $\boldsymbol{\mu}_i \in P$ or the lower bound on its reward is below $d(S)$. If i is in S, then it needs to be pulled more if the upper bound on its reward is above $d(S)$, and if i is in I, then it needs to be pulled more if it has not been determined that $\boldsymbol{\mu}_i \notin P$.

Next, we state the two main lemmas that we use in the proof of Theorem 9.

Lemma 9. *Fix $\delta > 0$ and a problem $(\nu, P, \boldsymbol{c}, m) \in \mathcal{M}$. Fix $(S, I) \in$ Valid-Partitions. Suppose that for all $i \in [K]$ and for all $t \geqslant 1$,* (i) *it holds that*

$$|\boldsymbol{c}^\top (\boldsymbol{\mu}_i - \widehat{\boldsymbol{\mu}}_{i,t})| \leqslant U_{\boldsymbol{c}}(t, \delta), \tag{3.29}$$

and (ii) *TestF(i, t) = True implies that $\boldsymbol{\mu}_i \in P$ and TestF(i, t) = False implies that $\boldsymbol{\mu}_i \notin P$. Then, for all t prior to termination (i.e., $t < \tau$), NEEDY$^t_{l_t}(S, I) \vee$ NEEDY$^t_{h_t}(S, I)$ is true.*

Lemma 9 essentially says that provided *(i)* $U_{\boldsymbol{c}}(t, \delta)$ bounds the deviation $|\boldsymbol{c}^\top (\boldsymbol{\mu}_i - \widehat{\boldsymbol{\mu}}_{i,t})|$ and *(ii)* TestF does not make a mistake, then every round prior to termination, at least one of the pulled arms is "needy."

The second main lemma states that provided that *(i)* $U_{\boldsymbol{c}}(t, \delta)$ bounds the deviation $|\boldsymbol{c}^\top (\boldsymbol{\mu}_i - \widehat{\boldsymbol{\mu}}_{i,t})|$ and *(ii)* TestF does not make a mistake, the algorithm returns a correct answer, i.e., returns $(\widehat{\mathrm{O}}, \widehat{\mathrm{S}}, \widehat{\mathrm{I}})$ such that $\widehat{\mathrm{O}} = \mathrm{OPT}$, $\widehat{\mathrm{S}} \subset \mathrm{SUBOPT}$, and $\widehat{\mathrm{I}} \subset \mathrm{INFEAS}$.

Lemma 10. *Fix $\delta > 0$ and a problem $(\nu, P, \boldsymbol{c}, m) \in \mathcal{M}$. Suppose that for all $i \in [K]$ and $t \in \mathbb{N}$,* (i) *it holds that*

$$|\boldsymbol{c}^\top(\boldsymbol{\mu}_i - \widehat{\boldsymbol{\mu}}_{i,t})| \leqslant U_{\boldsymbol{c}}(t, \delta), \tag{3.30}$$

and (ii) *TestF(i,t) = True implies that $\boldsymbol{\mu}_i \in P$ and TestF(i,t) = False implies that $\boldsymbol{\mu}_i \notin P$. Then, TF-LUCB($\delta$) returns $(\widehat{O}, \widehat{S}, \widehat{I})$ such that $\widehat{O} = OPT$, and $(\widehat{S}, \widehat{I}) \in$ Valid-Partitions.*

The proofs of the two lemmas are given in Section 3.11.1.

Next, we prove Theorem 9. The proof has three main steps. First, we show that whp for every arm i *(i)* TestF does not make a mistake about the feasibility of arm i, *(ii)* after arm i has been pulled $\eta(\nu_i, P)$ times, TestF determines whether arm i is feasible, and *(iii)* $U_{\boldsymbol{c}}(t, \delta)$ controls the deviation of the empirical mean reward to the expected reward for arm i. Second, we apply Lemma 10 to conclude that the algorithm returns the correct answer. Finally, we upper bound the sample complexity, τ, of the algorithm by essentially upper bounding how many times an arm must be pulled before no longer being "needy."

Proof of Theorem 9. **Step 1: Defining the event.** Let $(S, I) \in$ Valid-Partitions that achieves the minimum in the upper bound (3.1) stated in Theorem 9. For the sake of brevity, we write NEEDY_i^t and d instead of $\text{NEEDY}_i^t(S, I)$ and $d(S)$, respectively.

If $\boldsymbol{\mu}_i \in P$, let

$$\begin{aligned} B_i = &\{\forall t \in \mathbb{N} : \text{TestF}(i,t) \neq \text{ False}\} \cap \{\forall t \geqslant \eta(\nu_i, P) : \text{TestF}(i,t) = \text{ True}\} \\ &\cap \{\forall t \in \mathbb{N} : |\boldsymbol{c}^\top(\widehat{\boldsymbol{\mu}}_{i,t} - \boldsymbol{\mu}_i)| \leqslant U_{\boldsymbol{c}}(t, \delta)\}. \end{aligned}$$

If $\boldsymbol{\mu}_i \notin P$, let

$$\begin{aligned} B_i = &\{\forall t \in \mathbb{N} : \text{TestF}(i,t) \neq \text{ True}\} \cap \{\forall t \geqslant \eta(\nu_i, P) : \text{TestF}(i,t) = \text{ False}\} \\ &\cap \{\forall t \in \mathbb{N} : |\boldsymbol{c}^\top(\widehat{\boldsymbol{\mu}}_{i,t} - \boldsymbol{\mu}_i)| \leqslant U_{\boldsymbol{c}}(t, \delta)\}. \end{aligned}$$

In words, when $\boldsymbol{\mu}_i \in P$, B_i says that *(i)* TestF does not make the mistake of concluding that arm i is infeasible, *(ii)* after arm i has been pulled $\eta(\nu_i, P)$ times, TestF determines that arm i is feasible, and *(iii)* $U_{\boldsymbol{c}}(t, \delta)$ controls the deviation of the empirical mean reward to the expected reward of arm i. For $\boldsymbol{\mu}_i \notin P$, B_i is the analogous event.

Observe that since $\|\boldsymbol{c}\|_2 = 1$ and ν_i is σ-sub-Gaussian, if $\boldsymbol{X} \sim \nu_i$, then

$$\left\|\boldsymbol{c}^\top \boldsymbol{X}\right\|_{\psi_2} \leqslant \|\boldsymbol{X}\|_{\psi_2} \leqslant \sigma$$

so that $\boldsymbol{c}^\top \boldsymbol{X}$ is σ-sub-Gaussian.

Then, by the union bound,

$$\begin{aligned}
&\Pr(\cup_{i=1}^K B_i^c) && (3.31)\\
&\leqslant \sum_{i\in[K]} \Pr(B_i^c) && (3.32)\\
&\leqslant \sum_{i\in\text{FEAS}} \Pr([\{\forall t \in \mathbb{N} : \text{TestF}(i,t) \neq \text{ False}\} \cap \{\forall t \geqslant \eta(\nu_i, P) : \text{TestF}(i,t) = \text{ True}\}]^c) && (3.33)\\
&+ \sum_{i\in\text{INFEAS}} \Pr([\{\forall t \in \mathbb{N} : \text{TestF}(i,t) \neq \text{ True}\} \cap \{\forall t \geqslant \eta(\nu_i, P) : \text{TestF}(i,t) = \text{ False}\}]^c) && (3.34)\\
&+ \sum_{i\in[K]} \Pr(\exists t \in \mathbb{N} : |\boldsymbol{c}^\top(\widehat{\boldsymbol{\mu}}_{i,t} - \boldsymbol{\mu}_i)| > U_{\boldsymbol{c}}(t,\delta)) && (3.35)\\
&\leqslant \sum_{i\in[K]} 2\frac{\delta}{2K} && (3.36)\\
&= \delta, && (3.37)
\end{aligned}$$

where line (3.36) follows by Lemma 20 and the assumption on TestF that for any set membership problem $(\xi, R) \in \mathcal{N}$ where ξ is σ-sub-Gaussian and has mean $\boldsymbol{\mu}$, with probability at least $1 - \frac{\delta}{2K}$, TestF returns True only if $\boldsymbol{\mu} \in R$ and False only if $\boldsymbol{\mu} \in R^c$ and uses at most $\eta(\xi, R)$ samples. For the rest of the proof, we assume $\cap_{i\in[K]} E_i$.

Step 2: Correctness. On event $\cap_{i\in[K]} B_i$, the conditions of Lemma 10 are satisfied, so that TF-LUCB returns $(\widehat{\text{O}}, \widehat{\text{S}}, \widehat{\text{I}})$ such that $\widehat{\text{O}} = \text{OPT}$, $\widehat{\text{S}} \subset \text{SUBOPT}$ and $\widehat{\text{I}} \subset \widehat{\text{I}}$.

Step 3: Sample Complexity. Next, we bound the sample complexity of TF-LUCB, i.e., prove (3.1) in the statement of Theorem 9. If $i \in \text{OPT}$, let ρ_i denote the smallest integer such that $\forall t \geqslant \rho_i$

$$U_{\boldsymbol{c}}(t,\delta) < \frac{\min_{j\in S} \boldsymbol{c}^\top(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)}{4}. \tag{3.38}$$

We claim that for all $i \in \text{OPT}$ and $s \in \mathbb{N}$, if $N_i(s) \geqslant \max(\rho_i, \eta(\nu_i, P))$, then $\text{NEEDY}_i^s = 0$. Let $i \in \text{OPT}$. Let $N_i(s) \geqslant \max(\rho_i, \eta(\nu_i, P))$. Then, on event B_i, $\text{TestF}(i, N_i(s)) = \text{ True}$,

which implies that $i \notin G_s$. Further,

$$\boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{i,N_i(s)} - U_{\boldsymbol{c}}(N_i(s),\delta) \geqslant \boldsymbol{c}^\top \boldsymbol{\mu}_i - 2\, U_{\boldsymbol{c}}(N_i(s),\delta) \tag{3.39}$$

$$\geqslant \boldsymbol{c}^\top \boldsymbol{\mu}_i - \frac{\min_{j\in S} \boldsymbol{c}^\top(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)}{2} \tag{3.40}$$

$$\geqslant d \tag{3.41}$$

where line (3.39) follows by event B_i and line (3.40) follows by (3.38). Thus, $\text{NEEDY}_i^s = 0$.

If $i \in S$, let ρ_i denote the smallest integer such that $\forall t \geqslant \rho_i$

$$U_{\boldsymbol{c}}(t,\delta) < \frac{\min_{j\in \text{OPT}} \boldsymbol{c}^\top \boldsymbol{\mu}_j - \boldsymbol{c}^\top \boldsymbol{\mu}_i}{4}. \tag{3.42}$$

We claim that for all $i \in S$ and $s \in \mathbb{N}$, if $N_i(s) \geqslant \rho_i$, then $\text{NEEDY}_i^s = 0$. Observe that

$$\boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{i,N_i(s)} + U_{\boldsymbol{c}}(N_i(s),\delta) \leqslant \boldsymbol{c}^\top \boldsymbol{\mu}_i + 2\, U_{\boldsymbol{c}}(N_i(s),\delta) \tag{3.43}$$

$$\leqslant \boldsymbol{c}^\top \boldsymbol{\mu}_i + \frac{\min_{j\in \text{OPT}} \boldsymbol{c}^\top(\boldsymbol{\mu}_j - \boldsymbol{\mu}_i)}{2} \tag{3.44}$$

$$\leqslant d \tag{3.45}$$

where line (3.43) follows by event B_i, and (3.44) follows by (3.42). Thus, $\text{NEEDY}_i^s = 0$.

Finally, let $i \in I$. Then, $N_i(s) \geqslant \eta(\nu_i, P)$ implies by event B_i that $\text{TestF}(i,t) =$ False, so that $i \notin G_s$. Thus, $\text{NEEDY}_i^s = 0$.

Then,

$$\tau - 1 \leqslant \sum_{t=1}^{\infty} \mathbf{1}\{\text{NEEDY}_{h_t}^t = 1 \text{ or } \text{NEEDY}_{l_t}^t = 1\} \tag{3.46}$$

$$\leqslant \sum_{t=1}^{\infty} \sum_{i=1}^{K} \mathbf{1}\{h_t = i \text{ or } l_t = i\} \mathbf{1}\{\text{NEEDY}_i^t = 1\} \tag{3.47}$$

$$\leqslant \sum_{t=1}^{\infty} \sum_{i\in \text{OPT}} \big[\mathbf{1}\{h_t = i \text{ or } l_t = i\} \mathbf{1}\{N_i(t) \leqslant \max(\rho_i, \eta(\nu_i,P))\} \tag{3.48}$$

$$+ \sum_{i\in S} \mathbf{1}\{h_t = i \text{ or } l_t = i\} \mathbf{1}\{N_i(t) \leqslant \rho_i\} \tag{3.49}$$

$$+ \sum_{i\in I} \mathbf{1}\{h_t = i \text{ or } l_t = i\} \mathbf{1}\{N_i(t) \leqslant \eta(\nu_i,P)\}\big] \tag{3.50}$$

$$\leqslant \sum_{i\in \text{OPT}} \max(\rho_i, \eta(\nu_i,P)) + \sum_{i\in S} \rho_i + \sum_{i\in I} \eta(\nu_i,P). \tag{3.51}$$

Line (3.46) follows by Lemma 9; line (3.48) follows by the contrapositive of the claim that for $i \in \text{OPT}$ and $s \in \mathbb{N}$, if $N_i(s) \geqslant \max(\rho_i, \eta(\nu_i,P))$, then $\text{NEEDY}_i^s = 0$; lines (3.49) and

(3.50) follow by the contrapositives of the analogous claims for $i \in S$ and $i \in I$; line (3.51) follows by exchanging the summations via Tonelli's theorem for series and if $h_t = i$ or $l_t = i$, then $N_i(t+1) = N_t(t) + 1$.

By Lemma 21, for $i \in \text{OPT}$,

$$\rho_i \leqslant c\sigma^2[\min_{j \in S} \boldsymbol{c}^\top(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)]^{-2}) \log(\log([\min_{j \in S} \boldsymbol{c}^\top(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)]^{-2}) \frac{K}{\delta}).$$

where c is a universal positive constant. By Lemma 21, for $i \in S$,

$$\rho_i \leqslant c\sigma^2[\min_{j \in \text{OPT}} \boldsymbol{c}^\top(\boldsymbol{\mu}_j - \boldsymbol{\mu}_i)]^{-2} \log(\log([\min_{j \in \text{OPT}} \boldsymbol{c}^\top(\boldsymbol{\mu}_j - \boldsymbol{\mu}_i)]^{-2}) \frac{K}{\delta})$$

where c is a universal positive constant. The result follows. □

3.11.1 Main Lemmas

Define the sets

$$\begin{aligned}
\text{ABOVE}_t(S) &= \{i \in [K] : \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{i,N_i(t)} - U_{\boldsymbol{c}}(N_i(t), \delta) > d(S)\} \\
\text{BELOW}_t(S) &= \{i \in [K] : \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{i,N_i(t)} + U_{\boldsymbol{c}}(N_i(t), \delta) < d(S)\} \\
\text{MIDDLE}_t(S) &= [K] \backslash (\text{ABOVE}_t(S) \cup \text{BELOW}_t(S))
\end{aligned}$$

Recall that $d(S)$ is the average of the smallest reward among the arms in OPT and the largest reward among the arms in S. Note that $d(S)$ is not known to the agent. Hence, $\text{ABOVE}_t(S)$ are the arms that at time t it is clear that whp their rewards are greater than the rewards of the arms in S and, similarly, $\text{BELOW}_t(S)$ are the arms that at time t it is clear that whp their rewards are less than the rewards of the arms in OPT. $\text{MIDDLE}_t(S)$ are the arms for which more evidence must be collected about their rewards to determine whether their reward is greater than or less than $d(S)$.

Proof of Lemma 9. Fix $(S, I) \in$ Valid-Partitions. Let t be some round prior to termination, i.e., $t < \tau$. For the sake of brevity, we write NEEDY_i^t, d, ABOVE_t, BELOW_t, and MIDDLE_t instead of $\text{NEEDY}_i^t(S, I)$, $d(S)$, $\text{ABOVE}_t(S)$, $\text{BELOW}_t(S)$, and $\text{MIDDLE}_t(S)$ respectively.

Case 1: $|\,\textbf{FEAS}\,| < m$**.** Then, SUBOPT $= \varnothing$ so that $S = \varnothing$. We claim that $h_t \in G_t$. Towards a contradiction, suppose that $h_t \notin G_t$. Since $h_t \in \text{TOP}_t \subset E_t$, if $h_t \notin G_t$, then

$h_t \in F_t$. Then, by lines 11 and 13 of the algorithm, $\text{TOP}_t \subset F_t$. Either *(i)* $|\,\text{TOP}_t\,| < m$ or *(ii)* $|\,\text{TOP}_t\,| = m$. Suppose $|\,\text{TOP}_t\,| < m$. Then, the definition of TOP_t implies that

$$E_t = \underset{t}{\text{TOP}} \subset F_t \subset E_t,$$

so that $\text{TOP}_t = F_t = E_t$. Thus, that t is the last round, i.e., $t = \tau$, which is a contradiction. Next, assume that $|\,\text{TOP}_t\,| = m$. Since by assumption $|\,\text{FEAS}\,| < m$ there exists $i \in \text{INFEAS}$ such that $\text{TestF}(i,t) = \text{True}$, which is a contradiction. Thus, $h_t \in G_t$.

Since $S = \varnothing$, $h_t \in \text{OPT} \cup I$, which implies $\text{NEEDY}_{h_t}^t = 1$.

Case 2: $|\,\textbf{FEAS}\,| \geqslant m$**.** We split the rest of the proof up into cases, where in each case we show either that $\text{NEEDY}_{h_t}^t = 1$, $\text{NEEDY}_{l_t}^t = 1$, or there is a contradiction. We briefly make two useful observations that follow from the assumption that $\text{TestF}(i,t) = \text{True}$ implies that $\boldsymbol{\mu}_i \in P$ and $\text{TestF}(i,t) = \text{False}$ implies that $\boldsymbol{\mu}_i \notin P$. First, the assumption implies that $\text{FEAS} \subset E_t$ for all t, so that $m \leqslant |\,\text{FEAS}\,| \leqslant |E_t|$ and furthermore by the definition of TOP_t, $|\,\text{TOP}_t\,| \geqslant m$. Second, if $i \in I \subset \text{INFEAS}$, the assumption implies that $\text{TestF}(i,t) \neq \text{True}$, so that $i \in G_t$ for all $t \in \mathbb{N}$.

- Suppose $\text{TOP}_t^c \cap E_t = \varnothing$. Then, $|E_t^c| \geqslant n - m$, which implies that

  $$m \leqslant |\,\text{FEAS}\,| \leqslant |E_t| \leqslant m.$$

  Since $E_t = \text{TOP}_t$ by definition of TOP_t, $\text{FEAS} \subset E_t = \text{TOP}_t$, so that $\text{TOP}_t = \text{FEAS} = \text{OPT}$. Either $\text{TOP}_t \subset F_t$ or $\text{TOP}_t \not\subset F_t$. If $\text{TOP}_t \subset F_t$, then

  $$E_t = \underset{t}{\text{TOP}} \subset F_t \subset E_t,$$

  so that $\text{TOP}_t = F_t = E_t$. Thus, that t is the last round, i.e., $t = \tau$, which is a contradiction. If $\text{TOP}_t \not\subset F_t$, then $h_t \in G_t$ by line 13 of the algorithm, so that $\text{NEEDY}_{h_t}^t = 1$. For the remainder of the proof, we will assume $\text{TOP}_t^c \cap E_t \neq \varnothing$.

- Suppose $h_t \in \text{BELOW}_t$ and $l_t \in \text{ABOVE}_t$. Then,

  $$\boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{h_t, N_{h_t}(t)} \leqslant \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{h_t, N_{h_t}(t)} + U_{\boldsymbol{c}}(N_{h_t}(t), \delta) \tag{3.52}$$

  $$< d \tag{3.53}$$

  $$< \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{l_t, N_{l_t}(t)} - U_{\boldsymbol{c}}(N_{l_t}(t), \delta) \tag{3.54}$$

  $$\leqslant \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{l_t, N_{l_t}(t)} \tag{3.55}$$

where line (3.53) follows since $h_t \in \text{BELOW}_t$ and line (3.54) follows since $l_t \in \text{ABOVE}_t$. Thus, $\boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{h_t, N_{h_t}(t)} < \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{l_t, N_{l_t}(t)}$. However, $h_t \in \text{TOP}_t$ and $l_t \in \text{TOP}_t \cap E_t$ imply $\boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{h_t, N_{h_t}(t)} \geqslant \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{l_t, N_{l_t}(t)}$ and thus we have a contradiction.

- Suppose that $h_t \in \text{BELOW}_t$ and $l_t \in \text{BELOW}_t$; we will derive a contradiction. We claim that $\text{OPT} \cap \text{TOP}_t^c = \varnothing$. Suppose that there exists $i \in \text{OPT} \cap \text{TOP}_t^c$. Since $\text{TestF}(i,t) \neq \text{False}$ for all $i \in \text{OPT}$, $i \in E_t$. Then,

$$\begin{aligned} \boldsymbol{c}^\top \boldsymbol{\mu}_i &\leqslant \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{i, N_i(t)} + U_{\boldsymbol{c}}(N_i(t), \delta) && (3.56) \\ &\leqslant \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{l_t, N_{l_t}(t)} + U_{\boldsymbol{c}}(N_{l_t}(t), \delta) && (3.57) \\ &< d, && (3.58) \end{aligned}$$

  where (3.56) follows by (3.29), (3.57) follows by $i \in E_t$, and (3.58) follows by $l_t \in \text{BELOW}_t$. $\boldsymbol{c}^\top \boldsymbol{\mu}_i < d$ is a contradiction, so that $\text{OPT} \cap \text{TOP}_t^c = \varnothing$. Thus, for all $i \in \text{TOP}_t^c$, either $\boldsymbol{c}^\top \boldsymbol{\mu}_i < d$ or $\boldsymbol{\mu}_i \notin P$. Furthermore, observe that

$$\begin{aligned} \boldsymbol{c}^\top \boldsymbol{\mu}_{h_t} &\leqslant \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{h_t, N_{h_t}(t)} + U_{\boldsymbol{c}}(N_{h_t}(t), \delta) && (3.59) \\ &\leqslant d && (3.60) \end{aligned}$$

  where line (3.59) follows by (3.29) and line (3.60) follows by $h_t \in \text{BELOW}_t$. Thus, there are at least $K - m + 1$ arms that are either suboptimal or infeasible. But, this is a contradiction since by assumption $|\,\text{FEAS}\,| \geqslant m$, there are exactly $K - m$ arms that are suboptimal or infeasible.

- Suppose $h_t \in \text{ABOVE}_t$ and $\text{TOP}_t \not\subset F_t$. Since $\text{TOP}_t \not\subset F_t$, $h_t \in G_t$ so that if $h_t \in \text{OPT} \cup I$, then $\text{NEEDY}_{h_t}^t = 1$. So, suppose that $h_t \in S$. If $h_t \in S$, then

$$\boldsymbol{c}^\top \boldsymbol{\mu}_{h_t} \geqslant \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{h_t, N_{h_t}(t)} - U_{\boldsymbol{c}}(N_{h_t}(t), \delta) > d$$

  where the first inequality follows by (3.29) and the second inequality follows by $h_t \in \text{ABOVE}_t$. But, $\boldsymbol{c}^\top \boldsymbol{\mu}_{h_t} > d$ is a contradiction since $h_t \in S$.

- Suppose $h_t \in \text{ABOVE}_t$, $\text{TOP}_t \subset F_t$, and $l_t \in \text{BELOW}_t$. Then, $\text{TOP}_t \subset F_t$, $h_t \in \text{ABOVE}_t$, and $l_t \in \text{BELOW}_t$ imply that the termination condition is satisfied so that $t = \tau$, which is a contradiction.

- Suppose $h_t \in \text{ABOVE}_t$, $\text{TOP}_t \subset F_t$, and $l_t \in \text{ABOVE}_t$. First, we claim that $\text{TOP}_t \subset \text{OPT}$. Let $i \in \text{TOP}_t$. Then, $i \in F_t$, which implies that $\text{TestF}(i,t) =$ True, so that $i \notin I$. Further, $h_t \in \text{ABOVE}_t$ implies that

$$\boldsymbol{c}^\top \boldsymbol{\mu}_i \geqslant \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{i,N_i(t)} - U_{\boldsymbol{c}}(N_i(t),\delta) > d,$$

  where the first inequality follows by (3.29) and the second inequality follows by $h_t \in \text{ABOVE}_t$. Therefore, $i \notin S$. Thus, $i \in \text{OPT}$, proving that $\text{TOP}_t \subset \text{OPT}$.

  There are three cases: either $l_t \in \text{OPT}$, $l_t \in S$, or $l_t \in I$. $l_t \in \text{OPT}$ implies that there are $m+1$ optimal feasible arms since $|\,\text{TOP}_t\,| \geqslant m$ as established earlier and $\text{OPT} \supset \text{TOP}_t$, which is a contradiction. Since $l_t \in \text{ABOVE}_t$, we have by (3.29),

$$\boldsymbol{c}^\top \boldsymbol{\mu}_{l_t} \geqslant \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{l_t,N_{l_t}(t)} - U_{\boldsymbol{c}}(N_{l_t}(t),\delta) > d,$$

  which implies that $l_t \notin S$. Thus, $l_t \in I$. Since $l_t \in G_t$ as established earlier, we have that $\text{NEEDY}^t_{l_t} = 1$.

- If $l_t \in \text{MIDDLE}_t$, then $l_t \notin \text{ABOVE}_t \cup \text{BELOW}_t$ so if $l_t \in \text{OPT} \cup S$, then $\text{NEEDY}^t_{l_t} = 1$. Further, if $l_t \in I$, then as argued previously $l_t \in G_t$, so that $\text{NEEDY}^t_{l_t} = 1$.

- If $h_t \in \text{MIDDLE}_t$, the argument is identical to the previous case.

□

Proof of Lemma 10. First, we observe that $\widehat{\text{O}} = \text{TOP}_\tau$, $\widehat{\text{S}} = (\text{TOP}_\tau \cup E_\tau^c)^c$, and $\widehat{\text{I}} = E_\tau^c$. Note that $\widehat{\text{S}} \cap \widehat{\text{I}} = \varnothing$ by definition of the algorithm.

Step 1: $\text{TOP}_\tau = \text{OPT}$.

To begin, we make two useful observations. *(i)* We claim that $\text{TOP}_\tau \subset \text{FEAS}$. Let $i \in \text{TOP}_\tau$. Then, since at termination, $\text{TOP}_\tau \subset F_\tau$, we have that $\text{TestF}(i,\tau) = \text{True}$. Then, by the hypothesis, $\boldsymbol{\mu}_i \in P$, so that $i \in \text{FEAS}$. *(ii)* We claim that $\text{OPT} \subset E_\tau$. Let $i \in \text{OPT}$. Since by assumption $\text{TestF}(i,\tau) \neq \text{False}$ for all i such that $\boldsymbol{\mu}_i \in P$, it follows that $i \in E_\tau$, establishing $\text{OPT} \subset E_\tau$.

Case 1: $|\,\text{FEAS}\,| < m$. Notice that since there are fewer than m feasible arms, $\text{SUBOPT} = \varnothing$ and we have that $\text{OPT} = \text{FEAS}$.

By our observation *(i)*, $\mathrm{TOP}_\tau \subset \mathrm{FEAS} = \mathrm{OPT}$.

Next, we show that $\mathrm{OPT} \subset \mathrm{TOP}_\tau$. Let $i \in \mathrm{OPT}$. Then, by observation *(ii)* $i \in E_\tau$. Since $\mathrm{TOP}_\tau \subset \mathrm{OPT}$ and $|\,\mathrm{OPT}\,| < m$, $|\,\mathrm{TOP}_\tau\,| < m$. Since $|\,\mathrm{TOP}_\tau\,| < m$, the definition of TOP_τ in line 6 of the algorithm implies that $|\,\mathrm{TOP}_\tau\,| = |E_\tau|$ and $\mathrm{TOP}_\tau \subset E_\tau$. Therefore, $\mathrm{TOP}_\tau = E_\tau$ so $i \in \mathrm{TOP}_\tau$, which establishes the claim.

Case 2: $|\,\mathbf{FEAS}\,| \geqslant m$. By observation *(i)*, $\mathrm{TOP}_\tau \subset \mathrm{FEAS}$, which implies that $\mathrm{TOP}_\tau \cap \mathrm{INFEAS} = \varnothing$. Next, we show that $\mathrm{TOP}_\tau \cap \mathrm{SUBOPT} = \varnothing$. Towards a contradiction, suppose that there exists $i \in \mathrm{TOP}_\tau \cap \mathrm{SUBOPT}$. Then, since $|\,\mathrm{OPT}\,| = m$ and $|\,\mathrm{TOP}_\tau\,| = m$ by *(ii)*, there exists $j \in \mathrm{OPT} \cap \mathrm{TOP}_\tau^c$. Since $\mathrm{OPT} \subset E_\tau$ by observation *(ii)*, $j \in E_\tau$. Then, by line 6 defining TOP_τ, $|E_\tau| > m$, so the algorithm must terminate with the stopping condition: $\mathrm{TOP}_t \subset F_t$ and $\min_{i \in \mathrm{TOP}_t} \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{i,N_i(t)} - U_{\boldsymbol{c}}(N_i(t), \delta) \geqslant \max_{j \in \mathrm{TOP}_t^c \cap E_t} \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{j,N_j(t)} + U_{\boldsymbol{c}}(N_j(t), \delta)$. By the stopping condition, we have that

$$\begin{aligned}
\boldsymbol{c}^\top \boldsymbol{\mu}_i &\geqslant \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{i,N_i(\tau)} - U_{\boldsymbol{c}}(N_i(\tau), \delta) && (3.61)\\
&\geqslant \min_{l \in \mathrm{TOP}_\tau} \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{l,N_k(\tau)} - U_{\boldsymbol{c}}(N_l(\tau), \delta) && (3.62)\\
&\geqslant \max_{k \in \mathrm{TOP}_\tau^c \cap E_\tau} \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{k,N_k(\tau)} + U_{\boldsymbol{c}}(N_k(\tau), \delta) && (3.63)\\
&\geqslant \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{j,N_j(\tau)} + U_{\boldsymbol{c}}(N_j(\tau), \delta) && (3.64)\\
&\geqslant \boldsymbol{c}^\top \boldsymbol{\mu}_j && (3.65)
\end{aligned}$$

where lines (3.61) and (3.65) follow by (3.30) and (3.63) follows by the stopping condition. Thus, $\boldsymbol{c}^\top \boldsymbol{\mu}_i \geqslant \boldsymbol{c}^\top \boldsymbol{\mu}_j$, which is contradicts the assumption $(\nu, P, \boldsymbol{c}, m) \in \mathcal{M}$. Therefore, the claim $\mathrm{TOP}_\tau \cap \mathrm{SUBOPT} = \varnothing$ follows.

Note that $\mathrm{TOP}_\tau \cap \mathrm{INFEAS} = \varnothing$ and $\mathrm{TOP}_\tau \cap \mathrm{SUBOPT} = \varnothing$ imply that $\mathrm{TOP}_\tau \subset \mathrm{OPT}$. Since $\mathrm{OPT} \subset E_\tau$ and $|\,\mathrm{FEAS}\,| \geqslant m$, $|\,\mathrm{TOP}_\tau\,| = m$. Thus, it follows that $\mathrm{TOP}_\tau = \mathrm{OPT}$ and correctness follows.

Step 2: $\widehat{\mathbf{S}} \subset \mathbf{SUBOPT}$ and $\widehat{\mathbf{I}} \subset \mathbf{INFEAS}$. First, we show that $\widehat{\mathrm{S}} \subset \mathrm{SUBOPT}$. If $\widehat{\mathrm{S}} = \varnothing$, there is nothing to show so suppose that $\widehat{\mathrm{S}} \neq \varnothing$. Let $i \in \widehat{\mathrm{S}}$. Since $i \in \widehat{\mathrm{S}} = \mathrm{TOP}_\tau^c \cap E_\tau$, we cannot have that $\mathrm{TOP}_\tau = F_\tau$ and $F_\tau = E_\tau$. So, the algorithm terminates with the stopping condition: $\mathrm{TOP}_t \subset F_t$ and $\min_{i \in \mathrm{TOP}_t} \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{i,N_i(t)} - U_{\boldsymbol{c}}(N_i(t), \delta) \geqslant \max_{j \in \mathrm{TOP}_t^c \cap E_t} \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{j,N_j(t)} +$

$U_{\boldsymbol{c}}(N_j(t), \delta)$. Then, using the stopping condition,

$$\boldsymbol{c}^\top \boldsymbol{\mu}_i \leqslant \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{i,N_i(\tau)} + U_{\boldsymbol{c}}(N_j(\tau), \delta) \tag{3.66}$$

$$\leqslant \min_{k \in \mathrm{TOP}_t} \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{k,N_k(t)} - U_{\boldsymbol{c}}(N_k(t), \delta) \tag{3.67}$$

$$\leqslant \min_{k \in \mathrm{OPT}} \boldsymbol{c}^\top \boldsymbol{\mu}_k \tag{3.68}$$

where lines (3.66) and (3.68) follow by (3.30) and line (3.67) follows by the stopping condition. Thus, $i \in$ SUBOPT by the assumption $(\nu, P, \boldsymbol{c}, m) \in \mathcal{M}$.

Next, we show that $\widehat{\mathrm{I}} \subset$ INFEAS. Let $i \in \widehat{\mathrm{I}}$. Then, TestF$(i, N_\tau(i))$ = False. By hypothesis, this implies that $\boldsymbol{\mu}_i \notin P$, so $i \in$ INFEAS.

□

3.12 Upper Bounds for Three Instances of TF-LUCB

In the following three sections, we prove Theorem 10. We prove a separate theorem for each statement in Theorem 10: namely, Theorem 11, Theorem 12, and Theorem 13. Each proof has a similar structure: *(i)* define a good event that holds whp, *(ii)* show that on this event, the TestF subroutine in question does not return the wrong answer, and *(iii)* show that after enough samples have been taken from the distribution, the TestF subroutine in question determines whether the mean of the distribution belongs to the set.

We introduce the following definition.

Definition 4. *Let $Z \subset \mathbb{R}^D$ and $\epsilon > 0$. $\mathcal{N} \subset Z$ is an ϵ-net of Z if for all $\boldsymbol{x} \in Z$, there exists $\boldsymbol{y} \in \mathcal{N}$ such that $\|\boldsymbol{x} - \boldsymbol{y}\|_2 \leqslant \epsilon$. Let $\mathcal{N} \subset Z$ be an ϵ-net of Z. We say that $\mathcal{N}$ is* minimal *if, for any other ϵ-net $\mathcal{O}$ of Z, it holds that $|\mathcal{O}| \geqslant |\mathcal{N}|$.*

3.12.1 Proof of Upper Bound for TF-LUCB-B

Theorem 11. *Let $\delta > 0$ and $(\nu, P, \boldsymbol{c}, m) \in \mathcal{M}$. With probability at least $1-\delta$, TF-LUCB-B returns $(\widehat{O}, \widehat{S}, \widehat{I})$ such that $\widehat{O} = OPT$, $\widehat{S} \subset SUBOPT$, $\widehat{I} \subset INFEAS$, and*

$$\tau \leqslant \min_{(S,I) \in Valid\text{-}Partitions} c\sigma^2 \Big[\sum_{i \in S} F(\min_{j \in OPT} \boldsymbol{c}^\top(\boldsymbol{\mu}_j - \boldsymbol{\mu}_i), \frac{K}{\delta}) + \sum_{i \in I} DF(\ \mathrm{dist}(\boldsymbol{\mu}_i, \partial P), \frac{K}{\delta}) \tag{3.69}$$

$$+ \sum_{i \in OPT} \max(F(\min_{j \in S} \boldsymbol{c}^\top(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j), \frac{K}{\delta}), DF(\ \mathrm{dist}(\boldsymbol{\mu}_i, \partial P), \frac{K}{\delta})) \Big]. \tag{3.70}$$

where c is a universal positive constant.

Proof. By Theorem 9, it suffices to show that for any $(\xi, R) \in \mathcal{N}$ where ξ is σ-sub-Gaussian and has mean $\boldsymbol{\mu} \in \mathbb{R}^D$, with probability at least $1 - \frac{\delta}{2K}$, TestF-B returns True only if $\boldsymbol{\mu} \in R$ and returns False only if $\boldsymbol{\mu} \notin R$ and after at most

$$c\sigma^2 D\ \mathrm{dist}(\boldsymbol{\mu}, \partial R)^{-2} \log(\log(\ \mathrm{dist}(\boldsymbol{\mu}, \partial R)^{-2}) \frac{K}{\delta})$$

pulls for some universal positive constant c, it returns either True or False.

Step 1: Define the event. Let $\widehat{\boldsymbol{\mu}}_t$ denote the empirical mean of ξ after t samples. Define the event $B = \{\forall t \in \mathbb{N} : \|\widehat{\boldsymbol{\mu}}_t - \boldsymbol{\mu}\|_2 \leqslant U_{\mathrm{ball}}(t, \delta)\}$. Let $\mathcal{N}$ be a minimal $\frac{1}{2}$-net of $\mathcal{S}^{D-1}$.

Observe that since for any $\boldsymbol{y} \in \mathcal{N}$, $\|\boldsymbol{y}\|_2 = 1$ and ν_i is σ-sub-Gaussian, if $\boldsymbol{X} \sim \nu_i$, then

$$\left\|\boldsymbol{y}^\top \boldsymbol{X}\right\|_{\psi_2} \leqslant \|\boldsymbol{X}\|_{\psi_2} \leqslant \sigma$$

so that $\boldsymbol{y}^\top \boldsymbol{X}$ is σ-sub-Gaussian.

Then,

$$\Pr(B^c) = \Pr(\exists t \in \mathbb{N} : \|\widehat{\boldsymbol{\mu}}_t - \boldsymbol{\mu}\|_2 > U_{\mathrm{ball}}(t, \delta)) \tag{3.71}$$

$$= \Pr(\exists t \in \mathbb{N}, \exists \boldsymbol{y} \in \mathcal{N} : |\boldsymbol{y}^\top(\widehat{\boldsymbol{\mu}}_t - \boldsymbol{\mu})| > \frac{1}{2} U_{\mathrm{ball}}(t, d)) \tag{3.72}$$

$$\leqslant \sum_{\boldsymbol{y} \in \mathcal{N}} \Pr(\exists t \in \mathbb{N} : |\boldsymbol{y}^\top(\widehat{\boldsymbol{\mu}}_t - \boldsymbol{\mu})| > \frac{1}{2} U_{\mathrm{ball}}(t, d)) \tag{3.73}$$

$$\leqslant 5^D \frac{\delta}{5^D 2K} \tag{3.74}$$

$$\leqslant \frac{\delta}{2K}, \tag{3.75}$$

where line (3.72) follows by Lemma 14 and line (3.74) follows by Lemma 20 and since Lemma 15 implies that $|\mathcal{N}| \leqslant 5^D$. So, $\Pr(B) \geqslant 1 - \frac{\delta}{2K}$. For the remainder of the proof, we suppose that B occurs.

Step 2: An incorrect answer is never returned. First, we consider the case $\boldsymbol{\mu} \in R$. First, we show that TestF-B returns only either True or ?. Towards a contradiction, suppose that TestF-B(t) = False. Then, since $\boldsymbol{\mu} \in R$ and event B,

$$U_{\text{ball}}(t, \delta) < \text{ dist}(\widehat{\boldsymbol{\mu}}_t, R) \leqslant \|\widehat{\boldsymbol{\mu}}_t - \boldsymbol{\mu}\|_2 \leqslant U_{\text{ball}}(t, \delta),$$

which is a contradiction. Thus, TestF-B returns either True or ?.

Next, consider the case $\boldsymbol{\mu} \in R^c$; the proof is very similar to the case $\boldsymbol{\mu} \in R$. Towards a contradiction, suppose that TestF-B(t) = True. Then, since $\boldsymbol{\mu} \in R^c$ and event B,

$$U_{\text{ball}}(t, \delta) < \text{ dist}(\widehat{\boldsymbol{\mu}}_t, R^c) \leqslant \|\widehat{\boldsymbol{\mu}}_t - \boldsymbol{\mu}\|_2 \leqslant U_{\text{ball}}(t, \delta),$$

which is a contradiction. Thus, TestF-B returns either False or ?.

Step 3: Bound the sample complexity. Next, we show that TestF-B(t) = returns either True or False for all

$$t \geqslant c\sigma^2 D \text{ dist}(\boldsymbol{\mu}, \partial R)^{-2} \log(\log(\frac{\text{dist}(\boldsymbol{\mu}, \partial R)^{-2})K}{\delta})$$

where c is a universal positive constant. Let ρ denote the smallest integer such that

$$U_{\text{ball}}(\rho, \delta) < \frac{\text{dist}(\boldsymbol{\mu}, \partial R)}{2}.$$

By Lemma 21, $\rho \leqslant c\sigma^2 D \text{ dist}(\boldsymbol{\mu}, \partial R)^{-2} \log(\frac{\log(\text{ dist}(\boldsymbol{\mu}, \partial R)^{-2})2K}{\delta})$ for some universal positive constant c. Let $t \geqslant \rho$. Towards a contradiction, suppose that TestF-B(i, t) = ?. Then, $\text{dist}(\widehat{\boldsymbol{\mu}}_t, R) \leqslant U_{\text{ball}}(t, \delta)$ and $\text{dist}(\widehat{\boldsymbol{\mu}}_t, R^c) \leqslant U_{\text{ball}}(t, \delta)$ so that by Lemma 19, there exists $\boldsymbol{x} \in \partial R$ such that $\|\widehat{\boldsymbol{\mu}}_t - \boldsymbol{x}\|_2 \leqslant U(t, \delta)$. Then, by the triangle inequality and event B,

$$\begin{aligned}
\|\boldsymbol{\mu} - \boldsymbol{x}\|_2 &\leqslant \|\boldsymbol{\mu} - \widehat{\boldsymbol{\mu}}_t\|_2 + \|\widehat{\boldsymbol{\mu}}_t - \boldsymbol{x}\|_2 \\
&\leqslant U_{\text{ball}}(t, \delta) + U_{\text{ball}}(t, \delta) \\
&< \text{ dist}(\boldsymbol{\mu}, \partial R) \\
&\leqslant \|\boldsymbol{\mu} - \boldsymbol{x}\|_2 ,
\end{aligned}$$

which is a contradiction. Thus, for all $t \geqslant \rho$, TestF-B(t) returns True or False. The result follows.

□

3.12.2 Proof of Upper Bound for TF-LUCB-CB

Define

$$\mathcal{N}_{poly} = \{(\xi, R) \in \mathcal{N} : \ R \text{ is a polyhedron}\}.$$

Theorem 12. *Let $\delta > 0$, $P = \{\boldsymbol{x} \in \mathbb{R}^D : A\boldsymbol{x} \leqslant \boldsymbol{b}\}$, and $(\nu, P, \boldsymbol{c}, m) \in \mathcal{M}$. With probability at least $1-\delta$, TF-LUCB-CB returns $(\widehat{O}, \widehat{S}, \widehat{I})$ such that $\widehat{O} = OPT$, $\widehat{S} \subset SUBOPT$, $\widehat{I} \subset INFEAS$, and*

$$\tau \leqslant \min_{(S,I) \in Valid\text{-}Partitions} c\sigma^2 \Big[\sum_{i \in S} F(\min_{j \in OPT} \boldsymbol{c}^\top(\boldsymbol{\mu}_j - \boldsymbol{\mu}_i), \frac{K}{\delta}) + \sum_{i \in I} DF(\ \mathrm{dist}(\boldsymbol{\mu}_i, \partial P), \frac{K}{\delta}) \tag{3.76}$$

$$+ \sum_{i \in OPT} \max(F(\min_{j \in S} \boldsymbol{c}^\top(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j), \frac{K}{\delta}), F(\ \mathrm{dist}(\boldsymbol{\mu}_i, \partial P), \frac{KM}{\delta})) \Big]. \tag{3.77}$$

where c is a universal positive constant.

Proof. By Theorem 9, it suffices to show that for any $(\xi, R) \in \mathcal{N}_{poly}$ where ξ is σ-sub-Gaussian and has mean $\boldsymbol{\mu} \in \mathbb{R}^D$, with probability at least $1 - \frac{\delta}{2K}$, if $\boldsymbol{\mu} \in R$, then TestF-CB only returns either ? or True and for all

$$t \geqslant c\sigma^2 \ \mathrm{dist}(\boldsymbol{\mu}, \partial R)^{-2} \log(\log(\ \mathrm{dist}(\boldsymbol{\mu}, \partial R)^{-2}) \frac{K}{\delta})$$

where c is a universal positive constant, TestF-CB(t) returns True, and if $\boldsymbol{\mu} \notin R$, then TestF-CB only returns either ? or False and for all

$$t \geqslant c\sigma^2 D \ \mathrm{dist}(\boldsymbol{\mu}, \partial R)^{-2} \log(\log(\ \mathrm{dist}(\boldsymbol{\mu}, \partial R)^{-2}) \frac{KM}{\delta})$$

where c is a universal positive constant, TestF-CB(t) returns False.

Step 1: Define the event. For the sake of brevity, let $U_{\mathrm{ball}}(t) \coloneqq U_{\mathrm{ball}}(t, \frac{\delta}{2})$ and $U_{\mathrm{con}}(t) \coloneqq U_{\mathrm{con}}(t, \frac{\delta}{2})$. Let $\widehat{\boldsymbol{\mu}}_t$ denote the empirical mean of ξ after t samples. Define the event

$$\begin{aligned} B =& \{\forall t \in \mathbb{N} : \|\widehat{\boldsymbol{\mu}}_t - \boldsymbol{\mu}\|_2 \leqslant U_{\mathrm{ball}}(t)\} \\ & \cap \{\forall t \in \mathbb{N}, \forall s \in [M] : |\boldsymbol{a}_s^\top \widehat{\boldsymbol{\mu}}_t - \boldsymbol{\mu}| \leqslant U_{\mathrm{con}}(t)\}. \end{aligned}$$

Let $\mathcal{N}$ be a minimal $\frac{1}{2}$-net of $\mathcal{S}^{D-1}$. Observe that since for any $\boldsymbol{y} \in \mathcal{N}$, $\|\boldsymbol{y}\|_2 = 1$, for any $j \in [M]$, $\|\boldsymbol{a}_j\|_2 = 1$ and ν_i is σ-sub-Gaussian, if $\boldsymbol{X} \sim \nu_i$, then for $\boldsymbol{z} \in \mathcal{N} \cup \{\boldsymbol{a}_j : j \in [M]\}$

$$\left\|\boldsymbol{z}^\top \boldsymbol{X}\right\|_{\psi_2} \leqslant \|\boldsymbol{X}\|_{\psi_2} \leqslant \sigma$$

so that $\boldsymbol{z}^\top \boldsymbol{X}$ is σ-sub-Gaussian.

By the union bound, Lemma 20, and a similar argument as in (3.75),

$$\Pr(B^c) \leqslant \frac{\delta}{4K} + \frac{\delta}{4K} = \frac{\delta}{2K}.$$

For the remainder of the proof, suppose that B occurs.

Step 2: $\boldsymbol{\mu} \in R$. Suppose $\boldsymbol{\mu} \in R$. First, we show that TestF-CB returns only either True or ?. Towards a contradiction, suppose that TestF-CB(t) = False. Then, since $\boldsymbol{\mu} \in R$ and event B,

$$U_{\text{ball}}(t) < \ \text{dist}(\widehat{\boldsymbol{\mu}}_t, R) \leqslant \|\widehat{\boldsymbol{\mu}}_t - \boldsymbol{\mu}\|_2 \leqslant U_{\text{ball}}(t),$$

which is a contradiction. Thus, TestF-CB returns either True or ?.

Next, we show that TestF-CB(t) = True for all

$$t \geqslant c\sigma^2 \ \text{dist}(\boldsymbol{\mu}, \partial R)^{-2} \log(\log(\ \text{dist}(\boldsymbol{\mu}, \partial R)^{-2}) \frac{K}{\delta})$$

where c is a universal positive constant. Let ρ denote the smallest integer such that

$$U_{\text{con}}(\rho) < \frac{\text{dist}(\boldsymbol{\mu}, \partial R)}{2} = \frac{\min_{s \in [M]} b_s - \boldsymbol{a}_s^\top \boldsymbol{\mu}}{2}.$$

where the equality follows by Lemma 13. By Lemma 21,

$$\rho \leqslant c\sigma^2 \ \text{dist}(\boldsymbol{\mu}, \partial R)^{-2} \log(\frac{\log(\ \text{dist}(\boldsymbol{\mu}, \partial R)^{-2}) KM}{\delta})$$

for some universal positive constant c. Let $t \geqslant \rho$. Fix $r \in [M]$. Then, by event B,

$$\begin{aligned} \boldsymbol{a}_r^\top \widehat{\boldsymbol{\mu}}_t + U_{\text{con}}(t, \delta) &\leqslant \boldsymbol{a}_r^\top \boldsymbol{\mu} + 2\, U_{\text{con}}(t) \\ &\leqslant \boldsymbol{a}_r^\top \boldsymbol{\mu} + b_r - \boldsymbol{a}_r^\top \boldsymbol{\mu} \\ &= b_r. \end{aligned}$$

Thus, TestF-CB(t) = True.

Step 3: $\boldsymbol{\mu} \in R^c$. Suppose $\boldsymbol{\mu} \in R^c$. Towards a contradiction, suppose that TestF-CB(t) returns True. Then, for all $s \in [M]$, $\boldsymbol{a}_s^\top \widehat{\boldsymbol{\mu}}_t + U_{\text{con}}(t) \leqslant b_s$. Then, by the event B,

$$b_s \geqslant \boldsymbol{a}_s^\top \widehat{\boldsymbol{\mu}}_t + U_{\text{con}}(t) \geqslant \boldsymbol{a}_s^\top \boldsymbol{\mu}$$

which contradicts the assumption that $\boldsymbol{\mu} \notin R$. Thus, TestF-CB($t$) only returns ? or False.

Next, we show that TestF-CB(t) returns False for all

$$t \geqslant cD \ \operatorname{dist}(\boldsymbol{\mu}, \partial R)^{-2} \log(\log(\frac{\operatorname{dist}(\boldsymbol{\mu}, \partial R)^{-2})K}{\delta})$$

where c is a universal positive constant. Let ρ denote the smallest integer such that

$$U_{\text{ball}}(\rho) < \frac{\operatorname{dist}(\boldsymbol{\mu}, \partial R)}{2}.$$

By Lemma 21, $\rho \leqslant c\sigma^2 D \ \operatorname{dist}(\boldsymbol{\mu}, \partial R)^{-2} \log(\frac{\log(\ \operatorname{dist}(\boldsymbol{\mu}, \partial R)^{-2})2K}{\delta})$ for some universal positive constant c. Let $t \geqslant \rho$. Towards a contradiction, suppose that TestF-B(t) = ?. Then, $\operatorname{dist}(\widehat{\boldsymbol{\mu}}_t, R) \leqslant U_{\text{ball}}(t)$ and $\operatorname{dist}(\widehat{\boldsymbol{\mu}}_t, R^c) \leqslant U_{\text{ball}}(t)$, so there exists $\boldsymbol{x} \in \partial R$ such that $\|\widehat{\boldsymbol{\mu}}_t - \boldsymbol{x}\|_2 \leqslant U_{\text{ball}}(t)$. Then, by the triangle inequality and event B,

$$\begin{aligned}
\|\boldsymbol{\mu} - \boldsymbol{x}\|_2 &\leqslant \|\boldsymbol{\mu} - \widehat{\boldsymbol{\mu}}_t\|_2 + \|\widehat{\boldsymbol{\mu}}_t - \boldsymbol{x}\|_2 \\
&\leqslant U_{\text{ball}}(t) + U_{\text{ball}}(t) \\
&< \operatorname{dist}(\boldsymbol{\mu}, \partial R) \\
&\leqslant \|\boldsymbol{\mu} - \boldsymbol{x}\|_2 ,
\end{aligned}$$

which is a contradiction. Thus, for all $t \geqslant \rho$, TestF-CB(t) returns False. The result follows.

□

3.12.3 Proof of Upper Bound for TF-LUCB-C

First, we prove a more general version of Theorem 14 that allows for any polyhedron.

Theorem 13. *Let $\delta > 0$, $P = \{\boldsymbol{x} \in \mathbb{R}^D : A\boldsymbol{x} \leqslant \boldsymbol{b}\}$, and $(\nu, P, \boldsymbol{c}, m) \in \mathcal{M}$. For all $i \in [K]$ such that $\boldsymbol{\mu}_i \notin P$, let $\tilde{\Delta}_i = \max_{s \in [M]} \boldsymbol{a}_s^\top \boldsymbol{\mu}_i - b_s$. With probability at least $1 - \delta$, TF-LUCB-C returns TOP_τ such that $TOP_\tau = OPT$ and*

$$\tau \leqslant \min_{(S,I) \in \textit{Valid-Partitions}} c\sigma^2 \Big[\sum_{i \in S} F(\min_{j \in OPT} \boldsymbol{c}^\top(\boldsymbol{\mu}_j - \boldsymbol{\mu}_i), \frac{K}{\delta}) + \sum_{i \in I} F(\tilde{\Delta}_i, \frac{KM}{\delta}) \tag{3.78}$$

$$+ \sum_{i \in OPT} \max(F(\min_{j \in S} \boldsymbol{c}^\top(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j), \frac{KM}{\delta})), F(\ \operatorname{dist}(\boldsymbol{\mu}_i, \partial P), \frac{KM}{\delta})) \Big]. \tag{3.79}$$

where c is a universal positive constant.

Proof. By Theorem 9, it suffices to show that for any $(\xi, R) \in \mathcal{N}_{poly}$ where ξ is σ-sub-Gaussian and has mean $\boldsymbol{\mu} \in \mathbb{R}^D$, with probability at least $1 - \frac{\delta}{2K}$, if $\boldsymbol{\mu} \in R$, then TestF-C only returns either ? or True and for all

$$t \geqslant c\sigma^2 \ \operatorname{dist}(\boldsymbol{\mu}, \partial R)^{-2} \log(\log(\ \operatorname{dist}(\boldsymbol{\mu}, \partial R)^{-2}) \frac{K}{\delta}$$

where c is a universal positive constant, TestF-C(t) returns True, and if $\boldsymbol{\mu} \notin R$, then TestF-C only returns either ? or False and for all

$$t \geqslant c\sigma^2 \tilde{\Delta}^{-2} \log(\log(\tilde{\Delta}_i \frac{KM}{\delta})$$

where $\tilde{\Delta} = \max_{s \in [M]} \boldsymbol{a}_s^\top \boldsymbol{\mu} - b_s$ and c is a universal positive constant, TestF-C(t) returns False.

Step 1: Define the event. Let $\widehat{\boldsymbol{\mu}}_t$ denote the empirical mean of ξ after t samples. Define the event $B = \{\forall t \in \mathbb{N}, \forall s \in [M] : |\boldsymbol{a}_s^\top (\widehat{\boldsymbol{\mu}}_t - \boldsymbol{\mu})| \leqslant U_{\text{con}}(t, \delta)\}$.

Observe that since for any $s \in [M]$, $\|\boldsymbol{a}_s\|_2 = 1$ and ν_i is σ-sub-Gaussian, if $\boldsymbol{X} \sim \nu_i$, then

$$\left\|\boldsymbol{a}_s^\top \boldsymbol{X}\right\|_{\psi_2} \leqslant \|\boldsymbol{X}\|_{\psi_2} \leqslant \sigma$$

so that $\boldsymbol{a}_s^\top \boldsymbol{X}$ is σ-sub-Gaussian.

Then, by Lemma 20,

$$\begin{aligned}
\Pr(B^c) &= \Pr(\exists t \in \mathbb{N}, \exists s \in [M] : |\boldsymbol{a}_s^\top (\widehat{\boldsymbol{\mu}}_t - \boldsymbol{\mu})| > U_{\text{con}}(t, \delta)) \\
&= M \Pr(\exists t \in \mathbb{N} : |\boldsymbol{a}_s^\top (\widehat{\boldsymbol{\mu}}_t - \boldsymbol{\mu})| > U_{\text{con}}(t, \delta)) \\
&\leqslant M \frac{\delta}{2KM} \\
&= \frac{\delta}{2K}.
\end{aligned}$$

So, $\Pr(B) \geqslant 1 - \frac{\delta}{2K}$. For the remainder of the proof, we suppose that B occurs.

Step 2: $\boldsymbol{\mu} \in R$. Suppose $\boldsymbol{\mu} \in R$. Towards a contradiction, suppose that TestF-C(t) = False. Then, there exists $s \in [M]$ such that $\boldsymbol{a}_s^\top \widehat{\boldsymbol{\mu}}_t - U_{\text{con}}(t, \delta) > b_s$. Then, by the event B,

$$b_s < \boldsymbol{a}_s^\top \widehat{\boldsymbol{\mu}}_t - U_{\text{con}}(t, \delta) \leqslant \boldsymbol{a}_s^\top \boldsymbol{\mu}$$

which contradicts the assumption that $\boldsymbol{\mu} \in R$. Thus, TestF-C only returns either ? or True.

Next, we show that TestF-C(t) = True for all

$$t \geqslant c\sigma^2 \ \mathrm{dist}(\boldsymbol{\mu}, \partial R)^{-2} \log(\log(\ \mathrm{dist}(\boldsymbol{\mu}, \partial R)^{-2}) \frac{K}{\delta}$$

where c is a universal positive constant. Let ρ denote the smallest integer such that

$$U_{\mathrm{con}}(\rho, \delta) < \frac{\mathrm{dist}(\boldsymbol{\mu}, \partial R)}{2} = \frac{\min_{s \in [M]} b_s - \boldsymbol{a}_s^\top \boldsymbol{\mu}}{2}.$$

where the equality follows by Lemma 13. By Lemma 21,

$$\rho \leqslant c\sigma^2 \ \mathrm{dist}(\boldsymbol{\mu}, \partial R)^{-2} \log(\frac{\log(\ \mathrm{dist}(\boldsymbol{\mu}, \partial R)^{-2}) KM}{\delta})$$

for some universal positive constant c. Let $t \geqslant \rho$. Fix $r \in [M]$. Then, by the event B,

$$\begin{aligned} \boldsymbol{a}_r^\top \widehat{\boldsymbol{\mu}}_t + U_{\mathrm{con}}(t, \delta) &\leqslant \boldsymbol{a}_r^\top \boldsymbol{\mu} + 2\, U_{\mathrm{con}}(t, \delta) \\ &\leqslant \boldsymbol{a}_r^\top \boldsymbol{\mu} + b_r - \boldsymbol{a}_r^\top \boldsymbol{\mu} \\ &= b_r. \end{aligned}$$

Thus, TestF-C(t) = True.

Step 3: $\boldsymbol{\mu} \in R^c$. Next, suppose $\boldsymbol{\mu} \in R^c$. Let $s \in [M]$ such that $\tilde{\Delta} = \boldsymbol{a}_s^\top \boldsymbol{\mu} - b_s$. Towards a contradiction, suppose that TestF-C(t) returns True. Then, $\boldsymbol{a}_s^\top \widehat{\boldsymbol{\mu}}_t + U_{\mathrm{con}}(t, \delta) \leqslant b_s$. Then, by the event B,

$$b_s \geqslant \boldsymbol{a}_s^\top \widehat{\boldsymbol{\mu}}_t + U_{\mathrm{con}}(t, \delta) \geqslant \boldsymbol{a}_s^\top \boldsymbol{\mu}$$

which contradicts the assumption that $\boldsymbol{\mu} \notin R$ and our choice of $s \in [M]$. Thus, TestF-C(t) only returns ? or False.

Next, we show that TestF-C(t) = False for all $t \geqslant c\sigma^2 \tilde{\Delta}^{-2} \log(\frac{\log(\tilde{\Delta}^{-2}) KM}{\delta})$ where c is a universal positive constant. Let ρ denote the smallest integer such that

$$U_{\mathrm{con}}(\tau, \delta) < \frac{\tilde{\Delta}}{2}.$$

By Lemma 21, $\rho \leqslant c\sigma^2 \tilde{\Delta}^{-2} \log(\frac{\log(\tilde{\Delta}^{-2}) KM}{\delta})$ for some universal positive constant c. Let $t \geqslant \rho$. Then, by the event B

$$\begin{aligned} \boldsymbol{a}_s^\top \widehat{\boldsymbol{\mu}}_t - U_{\mathrm{con}}(t, \delta) &\geqslant \boldsymbol{a}_s^\top \boldsymbol{\mu} - 2\, U_{\mathrm{con}}(t, \delta) \\ &\geqslant \boldsymbol{a}_s^\top \boldsymbol{\mu} - (\boldsymbol{a}_s^\top \boldsymbol{\mu} - b_s) \\ &= b_s. \end{aligned}$$

Thus, TestF-C(t) = False.

□

In general, $\tilde{\Delta}_i$ can be arbitrarily smaller than $\text{dist}(\boldsymbol{\mu}_i, P)$, as indicated by the following Proposition.

Proposition 3. *For all $M > 0$ and for all $\epsilon > 0$, there exists a polyhedron $P = \{\boldsymbol{x} \in \mathbb{R}^D : A\boldsymbol{x} \leqslant b\}$ and $\boldsymbol{x}_0 \in \mathbb{R}^D$ such that $\text{dist}(\boldsymbol{x}_0, P) \geqslant M$ and $\max_{i=1,\dots,M} \text{dist}(\boldsymbol{x}_0, \{\boldsymbol{x} \in \mathbb{R}^D : \boldsymbol{a}_i^\top \boldsymbol{x} \leqslant b_i\}) \leqslant \epsilon$.*

Proof of Proposition 12. Consider the case $D = 2$. Fix $M > 0$ and $\epsilon > 0$. Consider

$$P_\alpha = \{\boldsymbol{x} \in \mathbb{R}^2 : \boldsymbol{e}_2^\top \boldsymbol{x} \geqslant 0, (\alpha \boldsymbol{e}_1 + (1-\alpha)\boldsymbol{e}_2)^\top \boldsymbol{x} \geqslant 0\}$$

where $\alpha \in (0,1)$. Let $\boldsymbol{x}_0 = -M\boldsymbol{e}_1$. Then, for sufficiently small $\alpha \in (0,1)$, we have that $\text{dist}(\boldsymbol{x}_0, P) \geqslant M$ and $\text{dist}(\boldsymbol{x}_0, \{\boldsymbol{x} \in \mathbb{R}^D : (\alpha \boldsymbol{e}_1 + (1-\alpha)\boldsymbol{e}_2)^\top \boldsymbol{x} \geqslant 0\}) \leqslant \epsilon$ □

However, the Theorem 14 shows that it has good performance in the setting where $\boldsymbol{a}_i^\top \boldsymbol{a}_j = 0$ for all $i \neq j \in [K]$.

Theorem 14. *Let $\delta > 0$, $P = \{\boldsymbol{x} \in \mathbb{R}^D : A\boldsymbol{x} \leqslant \boldsymbol{b}\}$ such that for any $l \neq k \in [K]$, $\boldsymbol{a}_l^\top \boldsymbol{a}_k = 0$, and $(\nu, P, \boldsymbol{c}, m) \in \mathcal{M}$. For each $i \in [K]$ such that $\boldsymbol{\mu}_i \notin P$, define $v_i = |\{j : \boldsymbol{a}_j^\top \boldsymbol{\mu}_i > b_j\}|$. With probability at least $1-\delta$, TF-LUCB-C returns $(\widehat{O}, \widehat{S}, \widehat{I})$ such that $\widehat{O} = OPT$, $\widehat{S} \subset SUBOPT$, $\widehat{I} \subset INFEAS$, and $\tau \leqslant$*

$$\min_{(S,I) \in \textit{Valid-Partitions}} c\sigma^2 \Big[\sum_{i \in S} F(\min_{j \in OPT} \boldsymbol{c}^\top (\boldsymbol{\mu}_j - \boldsymbol{\mu}_i), \frac{K}{\delta}) + \sum_{i \in I} v_i F(\ \text{dist}(\boldsymbol{\mu}_i, P), \frac{KM}{\delta}) \quad (3.80)$$

$$+ \sum_{i \in OPT} \max(F(\min_{j \in S} \boldsymbol{c}^\top (\boldsymbol{\mu}_i - \boldsymbol{\mu}_j), \frac{KM}{\delta}), F(\ \text{dist}(\boldsymbol{\mu}_i, \partial P), \frac{KM}{\delta}) \Big]. \quad (3.81)$$

Proof of Theorem 14. Let $l \in [K]$ such that $\boldsymbol{\mu}_l \notin P$. Without loss of generality, by relabeling $\boldsymbol{a}_1, \dots, \boldsymbol{a}_M$, let

$$[r] = \{j \in [K] : \boldsymbol{a}_j^\top \boldsymbol{\mu}_i > b_j\}.$$

Define

$$S_i = \{\boldsymbol{x} \in \mathbb{R}^D : \boldsymbol{a}_i^\top \boldsymbol{x} = b_i\}$$

$$S = \cap_{i=1,\dots,r} S_i$$

We will show that

$$\operatorname{dist}(\boldsymbol{\mu}_l, P)^2 \leqslant \operatorname{dist}(\boldsymbol{\mu}_l, S)^2 \leqslant r \max_{i=1,\ldots,r} \operatorname{dist}(\boldsymbol{\mu}_l, S_i)^2;$$

Then, the result will following by plugging the above inequality into the upper bound (3.79) in the statement of Theorem 13. By relabeling the subspaces, we may assume without loss of generality that

$$\max_{i=1,\ldots,r} \operatorname{dist}(\boldsymbol{\mu}_l, S_i) = \operatorname{dist}(\boldsymbol{\mu}_l, S_1).$$

Define

$$\begin{aligned}
\boldsymbol{x}_0 &= \boldsymbol{\mu}_l, \\
\boldsymbol{x}_1 &= \operatorname{Proj}_{S_1}(\boldsymbol{x}_0), \\
\boldsymbol{x}_{i+1} &= \operatorname{Proj}_{S_{i+1}}(\boldsymbol{x}_i).
\end{aligned}$$

We claim that for all $i \in [r]$, $\boldsymbol{x}_i = \boldsymbol{x}_0 + \sum_{j=1}^{i}(b_j - \boldsymbol{a}_j^\top \boldsymbol{x}_0)\boldsymbol{a}_j$. We prove this inductively. By the closed form solution of the distance from a point to a hyperplane and $\|\boldsymbol{a}_j\|_2 = 1$ for all $j \in [M]$ (Boyd and Vandenberghe, 2004),

$$\boldsymbol{x}_1 = \boldsymbol{x}_0 + (b_1 - \boldsymbol{a}_1^\top \boldsymbol{x}_0)\boldsymbol{a}_1,$$

which shows that base case. Next, we show the inductive step; suppose $\boldsymbol{x}_i = \boldsymbol{x}_0 + \sum_{j=1}^{i}(b_j - \boldsymbol{a}_j^\top \boldsymbol{x}_0)\boldsymbol{a}_j$. Then,

$$\begin{aligned}
\boldsymbol{x}_{i+1} &= \boldsymbol{x}_i + (b_{i+1} - \boldsymbol{a}_{i+1}^\top \boldsymbol{x}_i)\boldsymbol{a}_{i+1} \\
&= \boldsymbol{x}_0 + \sum_{j=1}^{i}(b_j - \boldsymbol{a}_j^\top \boldsymbol{x}_0)\boldsymbol{a}_j + (b_{i+1} - \boldsymbol{a}_{i+1}^\top [\boldsymbol{x}_0 + \sum_{j=1}^{i}(b_j - \boldsymbol{a}_j^\top \boldsymbol{x}_0)\boldsymbol{a}_j])\boldsymbol{a}_{i+1} \\
&= \boldsymbol{x}_0 + \sum_{j=1}^{i+1}(b_j - \boldsymbol{a}_j^\top \boldsymbol{x}_0)\boldsymbol{a}_j
\end{aligned}$$

where we used the assumption that $\boldsymbol{a}_{i+1}^\top \boldsymbol{a}_j = 0$ for all $j \neq i+1$. Thus, the claim follows. Note that this implies that $\boldsymbol{x}_r \in S$.

Next, we note that for $i \neq j$,

$$(\boldsymbol{x}_i - \boldsymbol{x}_{i+1})^\top (\boldsymbol{x}_j - \boldsymbol{x}_{j+1}) = [-(b_{i+1} - \boldsymbol{a}_{i+1}^\top \boldsymbol{x}_0)\boldsymbol{a}_{i+1}]^\top [-(b_{j+1} - \boldsymbol{a}_{j+1}^\top \boldsymbol{x}_0)\boldsymbol{a}_{j+1}] = 0.$$

Then, by the pythagorean theorem,

$$\begin{aligned}
\operatorname{dist}(\boldsymbol{x}_0, S)^2 &\leqslant \|\boldsymbol{x}_0 - \boldsymbol{x}_r\|_2^2 \\
&= \|(\boldsymbol{x}_0 - \boldsymbol{x}_1) + (\boldsymbol{x}_1 - \boldsymbol{x}_2) + \ldots + (\boldsymbol{x}_{r-1} - \boldsymbol{x}_r)\|_2^2 \\
&= \sum_{i=1}^{r} \|\boldsymbol{x}_{i-1} - \boldsymbol{x}_r\|_2^2 \\
&\leqslant r \ \operatorname{dist}(\boldsymbol{x}_0, S_1) \\
&= r \max_{i=1,\ldots,r} \operatorname{dist}(\boldsymbol{x}_0, S_i).
\end{aligned}$$

Next, we show that $\operatorname{dist}(\boldsymbol{\mu}_l, P) \leqslant \operatorname{dist}(\boldsymbol{\mu}_l, S)$. It suffices to show that $\boldsymbol{x}_r \in P$. For $s \in [r]$, $\boldsymbol{a}_s^\top \boldsymbol{x}_r = b_s$ by construction, so let $s \in [M] \backslash [r]$. Then, since $\boldsymbol{a}_s^\top \boldsymbol{a}_k = 0$ for all $k \in [r]$, it follows that

$$\boldsymbol{a}_s^\top \boldsymbol{x}_r = \boldsymbol{a}_s^\top \boldsymbol{x}_0 + \sum_{j=1}^{r} (b_j - \boldsymbol{a}_j^\top \boldsymbol{x}_0) \boldsymbol{a}_s^\top \boldsymbol{a}_i \leqslant b_s + 0.$$

Thus, it follows that $\operatorname{dist}(\boldsymbol{\mu}_l, P) \leqslant \operatorname{dist}(\boldsymbol{\mu}_l, S)$. □

3.13 Alternative Lower Bound

To begin, we discuss our conjecture that there is a small gap between δ- PAC and δ- PAC-EXPLANATORY algorithms. Essentially a δ- PAC algorithm that is not δ- PAC-EXPLANATORY is allowed to rule out suboptimal feasible arms by incorrectly concluding that they are infeasible and to make the analogous mistake for infeasible arms with reward greater than $\max_{j \in \mathrm{FEAS}}^{(m)} \boldsymbol{c}^\top \boldsymbol{\mu}_j$. We do not believe that this affords significant savings in sample complexity since δ- PAC algorithms typically use confidence bounds and to satisfy the δ- PAC criterion, these confidence bounds must be strong enough to determine that arms in OPT are feasible and have optimal rewards and to rule out every arm in OPT^c as either suboptimal or infeasible–all without prior knowledge of the number of infeasible or suboptimal arms. Nevertheless, we leave this as an open question.

Next, we discuss the differences between Theorems 7 and 8. Since any δ- PAC-EXPLANATORY algorithm wrt $\mathcal{M}$ is δ- PAC wrt $\mathcal{M}$, we expect the lower bound in Theorem 7 to be at least as large as the lower bound in Theorem 8, and this is in fact the case. The main difference between the bounds occurs in the terms corresponding to $i \in \mathrm{OPT}$. The term

$\min_{j \in \text{OPT}^c \cap \text{FEAS}} \boldsymbol{c}^\top(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)$ in Theorem 8 is replaced with $\min_{j \in S} \boldsymbol{c}^\top(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)$ where $S \supseteq \text{OPT}^c \cap \text{FEAS}$. Essentially, in Theorem 7, it is required to show that every arm in OPT has reward greater than all arms that are ruled out as suboptimal (i.e., belong to S), whereas in Theorem 8, these arms must only be shown to have reward greater than arms in $\text{FEAS} \cap \text{OPT}^c$. We conjecture that Theorem 8 is loose in this respect since intuitively if an algorithm rules out an arm by concluding that it is suboptimal, then regardless of whether the arm is feasible, the algorithm must determine that the arms in OPT have reward greater than it. To see the difference between theorems 7 and 8, consider the case where $K = 3$, $m = 1$, $\boldsymbol{c}^\top\boldsymbol{\mu}_1 > \boldsymbol{c}^\top\boldsymbol{\mu}_2 > \boldsymbol{c}^\top\boldsymbol{\mu}_3$, arms 1 and 3 are feasible and arm 2 is feasible. If arm 2 is very close to the boundary, then it may be much easier to show that arm 2 is suboptimal than to show that it is infeasible. In this case, the term reflecting the difficulty of showing that arm 1 is optimal will differ in the two theorems. Specifically, in this case, $\text{OPT}^c \cap \text{FEAS} = \{3\}$ and $S = \{2, 3\}$, so

$$\min_{j \in \text{OPT}^c \cap \text{FEAS}} \boldsymbol{c}^\top(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_j) = \boldsymbol{c}^\top(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_3) > \boldsymbol{c}^\top(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) = \min_{j \in S} \boldsymbol{c}^\top(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j).$$

Next, we prove Theorem 8. The proof has many similarities with the proof of Theorem 7. Recall the notation that for a given problem $(\nu, P, \boldsymbol{c}, m)$, we define

$$\begin{aligned}
&\text{FEAS}(\nu, P, \boldsymbol{c}, m) = \{i \in [K] : \boldsymbol{\mu}_i \in P\},\ \text{INFEAS}(\nu, P, \boldsymbol{c}, m) = \text{FEAS}(\nu, P, \boldsymbol{c}, m)^c,\\
&\text{OPT}(\nu, P, \boldsymbol{c}, m) = \{i \in \text{FEAS}(\nu, P, \boldsymbol{c}, m) : \boldsymbol{c}^\top\boldsymbol{\mu}_i \geqslant \max^{(m)}_{j \in \text{FEAS}(\nu, P, \boldsymbol{c}, m)} \boldsymbol{c}^\top\boldsymbol{\mu}_j\},\\
&\text{SUBOPT}(\nu, P, \boldsymbol{c}, m) = \{i \in [K] : \boldsymbol{c}^\top\boldsymbol{\mu}_i < \max^{(m)}_{j \in \text{FEAS}(\nu, P, \boldsymbol{c}, m)} \boldsymbol{c}^\top\boldsymbol{\mu}_j\}.
\end{aligned}$$

Proof of Theorem 8. Fix $\delta > 0$. Let $(\nu, P, \boldsymbol{c}, m)$ satisfy the hypotheses of the Theorem statement; note that these properties imply that $(\nu, P, \boldsymbol{c}, m) \in \mathcal{M}$. Let $\mathcal{A}$ denote a δ- PAC algorithm with stopping time τ.

In each of the next steps, we will define a new problem to obtain a lower bound. To avoid notational clutter, we will redefine the symbols $\boldsymbol{\mu}'_i$, ν'_i, and $\nu^{(i)}$ in each step. The context should make their meaning clear.

Step 1.a: reward bound for $i \in$ OPT. Fix $i \in \text{OPT}$. First, we show that

$$\mathbb{E}_\nu[N_i(\tau)] \geqslant 2\ln(\frac{1}{2.4\delta})\big[\min_{j \in \text{FEAS} \cap \text{OPT}^c} \boldsymbol{c}^\top(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j) + \epsilon\big]^{-2}$$

for a sufficiently small $\epsilon > 0$. If $\text{FEAS} \cap \text{OPT}^c = \varnothing$, $\min_{j \in \text{FEAS} \cap \text{OPT}^c} \boldsymbol{c}^\top(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j) = -\infty$ by definition and there is nothing to show. So, suppose that $\text{FEAS} \cap \text{OPT}^c \neq \varnothing$. Define

$$j_0 = \underset{j \in \text{FEAS} \cap \text{OPT}^c}{\arg\max} \boldsymbol{c}^\top \boldsymbol{\mu}_j.$$

Define for all $j \in [K]$

$$\boldsymbol{\mu}'_j = \begin{cases} \begin{pmatrix} \mu_{i,1} - \mu_{j_0,1} - \epsilon \\ \boldsymbol{\mu}_{i,2:D} \end{pmatrix} & \text{if } j = i \\ \boldsymbol{\mu}_j & \text{if } j \neq i \end{cases}$$

$$\nu'_j = N(\boldsymbol{\mu}'_j, I_D).$$

where $\epsilon > 0$ is chosen sufficiently small such that for any $\delta \in [0, \epsilon)$ $\boldsymbol{c}^\top \boldsymbol{\mu}'_i + \delta \neq \boldsymbol{c}^\top \boldsymbol{\mu}'_j$ for all $j \neq i$ (which is possible since $\boldsymbol{c}^\top \boldsymbol{\mu}_l \neq \boldsymbol{c}^\top \boldsymbol{\mu}_k$ for all $l \neq k \in [K]$). Define $\nu^{(i)} = (\nu'_1, \ldots, \nu'_K)$ and consider the problem $(\nu^{(i)}, P, \boldsymbol{c}, m)$. We claim that $(\nu^{(i)}, P, \boldsymbol{c}, m) \in \mathcal{M}$. Since $\boldsymbol{\mu}_i \notin \partial P$ and $\partial P = \partial(\mathbb{R} \times P') = \mathbb{R} \times \partial P'$ for some $P' \subset \mathbb{R}^{D-1}$, $\boldsymbol{\mu}'_i \notin \partial P$. Further, by construction, $\boldsymbol{c}^\top \boldsymbol{\mu}'_i \neq \boldsymbol{c}^\top \boldsymbol{\mu}'_j$ for all $j \neq i$. Thus, none of the arms have means on the boundary of P and all of the rewards of the arms are distinct, so $(\nu^{(i)}, P, \boldsymbol{c}, m) \in \mathcal{M}$.

Consider the event $B = \{i \in \widehat{\text{O}}\}$. Define $\text{OPT}_i = \text{OPT}(\nu^{(i)}, P, \boldsymbol{c}, m)$ and $\text{FEAS}_i = \text{FEAS}(\nu^{(i)}, P, \boldsymbol{c}, m)$. Observe that $i \notin \text{OPT}_i$ since $j_0 \in \text{FEAS}_i$ and $\boldsymbol{c}^\top \boldsymbol{\mu}'_i < \boldsymbol{c}^\top \boldsymbol{\mu}'_{j_0} = \max_{j \in \text{FEAS} \cap \text{OPT}^c} \boldsymbol{c}^\top \boldsymbol{\mu}_j$., so that there are m feasible arms with reward greater than $\boldsymbol{c}^\top \boldsymbol{\mu}'_i$.

Then, since $\mathcal{A}$ is δ- PAC wrt to $\mathcal{M}$, $(\nu^{(i)}, P, \boldsymbol{c}, m) \in \mathcal{M}$, and arm $i \notin \text{OPT}_i$, we have that

$$\Pr_{\nu^{(i)}}(B) \leqslant \Pr_{\nu^{(i)}}(\text{OPT} \neq \widehat{\text{O}}) \leqslant \delta. \tag{3.82}$$

Further, since $\mathcal{A}$ is δ- PAC wrt $\mathcal{M}$,

$$\Pr_{\nu}(i \in \widehat{\text{O}}) \geqslant \Pr_{\nu}(\text{OPT} = \widehat{\text{O}}) \geqslant 1 - \delta. \tag{3.83}$$

Then,

$$\frac{1}{2}[\boldsymbol{c}^\top(\boldsymbol{\mu}_i - \boldsymbol{\mu}_{j_0}) + \epsilon]^2 \mathbb{E}_\nu[N_i(\tau)] = \mathrm{KL}(\nu_i, \nu_i')\mathbb{E}_\nu[N_i(\tau)] \tag{3.84}$$

$$\geqslant d(\Pr_\nu(B), \Pr_{\nu^{(i)}}(B)) \tag{3.85}$$

$$\geqslant d(\Pr_\nu(B), \delta) \tag{3.86}$$

$$\geqslant d(\frac{1-\delta}{2}, \delta) \tag{3.87}$$

$$\geqslant \ln(\frac{1}{2.4\delta}). \tag{3.88}$$

Line (3.84) follows by the formula for the KL-divergence of two multivariate normal distributions, (3.85) follows by Lemma 11, (3.86) follows since $x \mapsto d(x, y)$ is increasing when $x > y$, (3.82), (3.83), and $\delta < .1$, (3.87) follows since $y \mapsto d(x, y)$ is decreasing when $x > y$, (3.82), (3.83), and $\delta < .1$, and (3.88) follows by Lemma 18. The claim follows by rearranging the inequality.

Step 1.b: feasibility bound for $i \in$ OPT. A similar argument to step 2.b from the proof of Theorem 7 yields

$$\frac{1}{2}(\ \mathrm{dist}(\boldsymbol{\mu}_i, \partial P) + \epsilon)^2 \mathbb{E}_\nu[N_i(\tau)] \geqslant \ln(\frac{1}{2.4\delta}). \tag{3.89}$$

Step 2: $i \in$ FEAS $\cap$ OPT^c.

This step is very similar to step 3 of the proof of Theorem 7 and yields

$$\ln(\frac{1}{2.4\delta}) \leqslant \frac{1}{2}[\boldsymbol{c}^\top(\boldsymbol{\mu}_{j_0} - \boldsymbol{\mu}_i) + \epsilon]^2 \mathbb{E}_\nu[N_i(\tau)]. \tag{3.90}$$

Step 3: $i \in$ INFEAS $\cap$ SUBOPT^c. Since $P \neq \mathbb{R}^D$ and P is nonempty, by Lemma 16 ∂P is nonempty. Since in addition ∂P is closed, by Lemma 12, there exists $\boldsymbol{\tau}_i \in \mathrm{Proj}_{\partial P}(\boldsymbol{\mu}_i)$. By definition of $\mathcal{M}$, since $\boldsymbol{\tau}_i \in \partial P$, for every $\epsilon > 0$, $B_\epsilon(\boldsymbol{\tau}_i) \cap P^\circ \neq \varnothing$. Thus, for sufficiently small $\epsilon > 0$, there exists a direction $\boldsymbol{v} \in \mathbb{R}^D$ with $\|\boldsymbol{v}\|_2 = 1$ such that $\boldsymbol{\tau}_i + \epsilon\boldsymbol{v} \in P^\circ$. Since by definition of $\mathcal{M}$, $P = \mathbb{R} \times P'$ for some $P' \subset \mathbb{R}^{D-1}$, we can choose $\boldsymbol{v}$ such that $v_1 = 0$. Define for all $j \in [K]$

$$\boldsymbol{\mu}_j' = \begin{cases} \boldsymbol{\tau}_i + \epsilon\boldsymbol{v} & \text{if } j = i \\ \boldsymbol{\mu}_j & \text{if } j \neq i \end{cases}$$

$$\nu_j' = N(\boldsymbol{\mu}_j', I_D).$$

Define $\nu^{(i)} = (\nu'_1, \ldots, \nu'_K)$ and consider the problem $(\nu^{(i)}, P, \boldsymbol{c}, m)$. It follows that $(\nu^{(i)}, P, \boldsymbol{c}, m) \in \mathcal{M}$ by a similar argument that showed in step 2.b of the proof of Theorem 7 that when $i \in \text{OPT}$, $(\nu^{(i)}, P, \boldsymbol{c}, m) \in \mathcal{M}$.

Define the event $B = \{i \notin \widehat{\text{O}}\}$. Define $\text{OPT}_i = \text{OPT}(\nu^{(i)}, P, \boldsymbol{c}, m)$ and $\text{SUBOPT}_i = \text{SUBOPT}(\nu^{(i)}, P, \boldsymbol{c}, m)$. Then, $i \in \text{OPT}_i$ since $\boldsymbol{\mu}'_i \in P$ and $i \in \text{SUBOPT}_i^c$ implies that $\boldsymbol{c}^\top \boldsymbol{\mu}'_i \geqslant \max^{(m)}_{l \in \text{FEAS}} \boldsymbol{c}^\top \boldsymbol{\mu}'_l$. Thus, since $\mathcal{A}$ is δ- PAC wrt $\mathcal{M}$,

$$\text{Pr}_{\nu^{(i)}}(B) \leqslant \text{Pr}_{\nu^{(i)}}(\widehat{\text{O}} \neq \text{OPT}) \leqslant \delta, \quad \text{and} \quad \text{Pr}_{\nu}(B) \geqslant 1 - \delta.$$

Therefore, by a series of inequalities similar to those in (3.23)-(3.26) in ste 2.b of the proof of Theorem 7,

$$\ln(\frac{1}{2.4\delta}) \leqslant \frac{1}{2}(\ \text{dist}(\boldsymbol{\mu}_i, P) + \epsilon)^2 \mathbb{E}_\nu[N_i(\tau)]. \tag{3.91}$$

Step 4: $i \in$ **INFEAS** $\cap$ **SUBOPT.** If $\text{INFEAS} \cap \text{SUBOPT} = \varnothing$, there is nothing to show. Thus, we may suppose without loss of generality that $\text{INFEAS} \cap \text{SUBOPT} \neq \varnothing$. Then, since in particular $\text{SUBOPT} \neq \varnothing$, there are m feasible arms and we may define

$$j_0 = \arg\ \max^{(m)}_{l \in \text{FEAS}} \boldsymbol{c}^\top \boldsymbol{\mu}_l.$$

By the same argument at the beginning of Step 3, there exists $\boldsymbol{\tau}_i \in \text{Proj}_{\partial P}(\boldsymbol{\mu}_i)$ and for sufficiently small $\epsilon > 0$, there exists a direction $\boldsymbol{v} \in \mathbb{R}^D$ with $\|\boldsymbol{v}\|_2 = 1$ and $v_1 = 0$ such that $\boldsymbol{\tau}_i + \epsilon \boldsymbol{v} \in P^\circ$. Define for all $j \in [K]$

$$\boldsymbol{\mu}'_j = \begin{cases} \begin{pmatrix} \mu_{i,1} + \mu_{j_0,1} + \epsilon \\ \boldsymbol{\tau}_{i,2:D} + \epsilon \boldsymbol{v}_{2,D} \end{pmatrix} & \text{if } j = i \\ \boldsymbol{\mu}_j & \text{if } j \neq i \end{cases}$$

$$\nu'_j = N(\boldsymbol{\mu}'_j, I_D).$$

where we choose $\epsilon > 0$ sufficiently small so that for any $\delta \in [0, \epsilon)$, $\boldsymbol{c}^\top \boldsymbol{\mu}'_i - \delta \neq \boldsymbol{c}^\top \boldsymbol{\mu}'_j$ for all $j \neq i$ (which is possible since $\boldsymbol{c}^\top \boldsymbol{\mu}_l \neq \boldsymbol{c}^\top \boldsymbol{\mu}_k$ for all $l \neq k \in [K]$). Then, define $\nu^{(i)} = (\nu'_1, \ldots, \nu'_K)$ and consider the problem $(\nu^{(i)}, P, \boldsymbol{c}, m)$. Using arguments similar to those in step 1, it follows that $(\nu^{(i)}, P, \boldsymbol{c}, m) \in \mathcal{M}$.

Consider the event $B = \{i \notin \widehat{\mathrm{O}}\}$. Then, $\Pr_\nu(B) \geqslant 1 - \delta$. Define for the sake of brevity $\mathrm{OPT}_i = \mathrm{OPT}(\nu^{(i)}, P, \boldsymbol{c}, m)$. Observe that $\boldsymbol{\mu}_i' \in P$ and $\boldsymbol{c}^\top \boldsymbol{\mu}_i' > \boldsymbol{c}^\top \boldsymbol{\mu}_{j_0}'$, so that $i \in \mathrm{OPT}_i$. Then, since $\mathcal{A}$ is δ- PAC wrt $\mathcal{M}$, $\Pr_{\nu^{(i)}}(B) \leqslant \delta$. Then,

$$\begin{aligned}
\ln(\frac{1}{2.4\delta}) &\leqslant \mathrm{KL}(\nu_i, \nu_{i,P,\boldsymbol{c}}^{(\epsilon)}) \mathbb{E}_\nu[N_i(\tau)] && (3.92)\\
&= \frac{1}{2}(\ \mathrm{dist}(\boldsymbol{\mu}_i, P) + \epsilon)^2 + [\boldsymbol{c}^t(\boldsymbol{\mu}_i - \boldsymbol{\mu}_{j_0}) + \epsilon]^2] \mathbb{E}_\nu[N_i(\tau)] && (3.93)\\
&\leqslant \max((\ \mathrm{dist}(\boldsymbol{\mu}_i, P) + \epsilon)^2, [\boldsymbol{c}^t(\boldsymbol{\mu}_i - \boldsymbol{\mu}_{j_0}) + \epsilon]^2) \mathbb{E}_\nu[N_i(\tau)] && (3.94)
\end{aligned}$$

where line (3.92) follows by a series of inequalities similar to (3.84)-(3.88), and line (3.93) follows by the definition of KL divergence of multivariate normal distributions.

Step 5: Putting it together. Using $\mathbb{E}_\nu[\tau] = \sum_{i=1}^K \mathbb{E}_\nu[N_i(\tau)]$ and inequalities (3.88). (3.89), (3.90), and (3.91), we establish for all sufficiently small $\epsilon > 0$,

$$\begin{aligned}
\mathbb{E}_\nu[\tau] \geqslant & 2 \ln(\frac{1}{2.4\delta}) \Big[\sum_{i \in \mathrm{OPT}} \max([\min_{j \in \mathrm{OPT}^c \cap \mathrm{FEAS}} \boldsymbol{c}^\top(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j) + \epsilon]^{-2}, [\ \mathrm{dist}(\boldsymbol{\mu}_i, \partial P) + \epsilon]^{-2})\\
& + \sum_{i \in \mathrm{OPT}^c \cap \mathrm{FEAS}} [\min_{j \in \mathrm{OPT}} \boldsymbol{c}^\top(\boldsymbol{\mu}_j - \boldsymbol{\mu}_i) + \epsilon]^{-2} + \sum_{i \in \mathrm{INFEAS} \cap \mathrm{SUBOPT}^c} [\ \mathrm{dist}(\boldsymbol{\mu}_i, P) + \epsilon]^{-2}\\
& + \sum_{i \in \mathrm{INFEAS} \cap \mathrm{SUBOPT}} \frac{1}{2} \min([\min_{j \in \mathrm{OPT}} \boldsymbol{c}^\top(\boldsymbol{\mu}_j - \boldsymbol{\mu}_i) + \epsilon]^{-2}, [\ \mathrm{dist}(\boldsymbol{\mu}_i, P) + \epsilon]^{-2}) \Big].
\end{aligned}$$

Since this bound holds for all $\epsilon > 0$ sufficiently small, letting $\epsilon \longrightarrow 0$ on the RHS of the above inequality establishes the result.

□

3.14 Technical Lemmas

We use the following lemma from Kaufmann et al. (2016b). Although they prove it for the case where arms are associated with scalar distributions, the proof generalizes to multi-dimensional distributions by simply replacing the scalar-valued distributions in the proof with vector-valued distributions. Let $I_t \in [K]$ denote the arm chosen by an agent at time t and $\boldsymbol{X}_t \sim \nu_{I_t}$. Let $\mathcal{F}_t = \sigma(I_1, \boldsymbol{X}_1, \ldots, I_t, \boldsymbol{X}_t)$, i.e., the sigma-algebra generated by $I_1, \boldsymbol{X}_1, \ldots, I_t, \boldsymbol{X}_t$.

Lemma 11. *Let ν and ν' be two bandit models with K arms such that for all a, the distributions ν_a and ν'_a are mutually absolutely continuous. Let τ denote a stopping time wrt $(\mathcal{F}_t)$. Then,*

$$\sum_{i=1}^{K} \mathbb{E}_\nu[N_i(\tau)]\, KL(\nu_a, \nu'_a) \geqslant \sup_{E \in \mathcal{F}_\tau} d(\Pr_\nu(E), \Pr_{\nu'}(E))$$

Lemma 12. *Let $\boldsymbol{x} \in \mathbb{R}^D$ and $A \subset \mathbb{R}^D$ be a closed nonempty set. Then, $\text{Proj}_A(\boldsymbol{x})$ is nonempty.*

Proof. Let $r > 0$ large enough such that $\bar{B}_r(\boldsymbol{x}) \cap A \neq \varnothing$. Then, observe that there exists $\boldsymbol{y} \in \text{Proj}_{A \cap \bar{B}_r(\boldsymbol{x})}(\boldsymbol{x})$ since $A \cap \bar{B}_r(\boldsymbol{x})$ is a compact set and $\|\cdot\|_2$ is continuous. Towards a contradiction, suppose there exists $\boldsymbol{z} \in A$ such that

$$\|\boldsymbol{z} - \boldsymbol{x}\|_2 < \|\boldsymbol{y} - \boldsymbol{x}\|_2 .$$

Then, $\boldsymbol{z} \in A \cap \bar{B}_r(\boldsymbol{x})$, which implies that $\boldsymbol{y} \notin \text{Proj}_{A \cap \bar{B}_r(\boldsymbol{x})}(\boldsymbol{x})$, a contradiction. Thus, for all $\boldsymbol{z} \in A$,

$$\|\boldsymbol{y} - \boldsymbol{x}\|_2 \leqslant \|\boldsymbol{x} - \boldsymbol{y}\|_2 .$$

Thus, $\boldsymbol{y} \in \text{Proj}_A(\boldsymbol{x})$. □

Lemmas 13 and 14 appear in Katz-Samuels and Scott (2018). For the sake of completeness, we restate the proof.

Lemma 13. *Let $P = \{\boldsymbol{x} \in \mathbb{R}^D : A\boldsymbol{x} \leqslant \boldsymbol{b}\}$ with $A \in \mathbb{R}^{M \times D}$. Let $\boldsymbol{\mu} \in P$. Then,*

$$\text{dis}(\boldsymbol{\mu}, \partial P) = \min_{i=1,\dots,M} \text{dis}(\boldsymbol{\mu}, \{\boldsymbol{x} : \boldsymbol{a}_i^\top \boldsymbol{x} = b_i\}).$$

Proof. It is not hard to establish that $\partial P = P \cap (\cup_{i=1}^M \{\boldsymbol{x} : \boldsymbol{a}_i^\top \boldsymbol{x} = b_i\})$. We claim that

$$\text{dis}(\boldsymbol{\mu}, \cup_{i=1}^M \{\boldsymbol{x} : \boldsymbol{a}_i^\top \boldsymbol{x} = b_i\}) = \text{dis}(\boldsymbol{\mu}, P \cap (\cup_{i=1}^M \{\boldsymbol{x} : \boldsymbol{a}_i^\top \boldsymbol{x} = b_i\})).$$

Since $\cup_{i=1}^M \{\boldsymbol{x} : \boldsymbol{a}_i^\top bx = b_i\}$ is closed, there exists $\boldsymbol{y} \in \cup_{i=1}^M \{\boldsymbol{x} : \boldsymbol{a}_i^\top bx = b_i\}$ such that

$$\|\boldsymbol{\mu} - \boldsymbol{y}\|_2 = \text{dis}(\boldsymbol{\mu}, \cup_{i=1}^M \{\boldsymbol{x} : \boldsymbol{a}_i^\top bx = b_i\}).$$

We claim that $\boldsymbol{y} \in P$. Suppose not (towards a contradiction). Then, there exists $\theta \in (0,1)$ such that $\boldsymbol{z} = (1-\theta)\boldsymbol{\mu} + \theta\boldsymbol{y} \in \partial P$. Then,

$$\text{dis}(\boldsymbol{\mu}, (\cup_{i=1}^{M}\{\boldsymbol{x} : \boldsymbol{a}_i^\top \boldsymbol{x} = b_i\})) \leqslant \|\boldsymbol{z}-\boldsymbol{\mu}\|_2 < \|\boldsymbol{y}-\boldsymbol{\mu}\|_2 = \text{dis}(\boldsymbol{\mu}, \cup_{i=1}^{M}\{\boldsymbol{x} : \boldsymbol{a}_i^\top b x = b_i\}),$$

which is a contradiction, establishing the claim. Then,

$$\begin{aligned}\min_{i=1,\ldots,M} \text{dis}(\boldsymbol{\mu}, \{\boldsymbol{x} : \boldsymbol{a}_i^\top \boldsymbol{x} = b_i\}) &= \text{dis}(\boldsymbol{\mu}, \cup_{i=1}^{M}\{\boldsymbol{x} : \boldsymbol{a}_i^\top \boldsymbol{x} = b_i\}) \\ &= \text{dis}(\boldsymbol{\mu}, P \cap (\cup_{i=1}^{M}\{\boldsymbol{x} : \boldsymbol{a}_i^\top \boldsymbol{x} = b_i\})) \\ &= \text{dis}(\boldsymbol{\mu}, \partial P).\end{aligned}$$

□

Lemma 14. *Let $\epsilon > 0$ and $\mathcal{N}_\epsilon$ be an ϵ-net of $\mathcal{S}^{D-1}$. For any $\boldsymbol{y} \in \mathbb{R}^D$,*

$$\|\boldsymbol{y}\|_2 \leqslant \frac{1}{1-\epsilon} \sup_{\boldsymbol{z} \in \mathcal{N}_\epsilon} \boldsymbol{y}^\top \boldsymbol{z}.$$

Proof. Let $\boldsymbol{z}_0 \in \mathcal{N}_\epsilon$ such that $\left\| \frac{\boldsymbol{y}}{\|\boldsymbol{y}\|_2} - \boldsymbol{z}_0 \right\|_2 \leqslant \epsilon$. Then, by Cauchy-Schwarz,

$$\|\boldsymbol{y}\|_2 = \frac{\boldsymbol{y}^\top \boldsymbol{y}}{\|\boldsymbol{y}\|_2} = \boldsymbol{y}^\top (\frac{\boldsymbol{y}}{\|\boldsymbol{y}\|_2} - \boldsymbol{z}_0) + \boldsymbol{y}^\top \boldsymbol{z}_0 \leqslant \|\boldsymbol{y}\|_2 \left\| \frac{\boldsymbol{y}}{\|\boldsymbol{y}\|_2} - \boldsymbol{z}_0 \right\|_2 + \boldsymbol{y}^\top \boldsymbol{z}_0 \leqslant \epsilon \|\boldsymbol{y}\|_2 + \boldsymbol{y}^\top \boldsymbol{z}_0.$$

Rearranging the inequality, we obtain

$$\|\boldsymbol{y}\|_2 \leqslant \frac{1}{1-\epsilon} \boldsymbol{y}^\top \boldsymbol{z}_0 \leqslant \frac{1}{1-\epsilon} \sup_{\boldsymbol{z} \in \mathcal{N}_\epsilon} \boldsymbol{y}^\top \boldsymbol{z}.$$

□

The following Lemma appears in Vershynin et al. (2017) (see Corollary 4.2.13).

Lemma 15. *Let $\epsilon > 0$ and $\mathcal{N}_\epsilon$ be a minimal ϵ-net of $\mathcal{S}^{D-1}$. Then, $|\mathcal{N}_\epsilon| \leqslant (\frac{2}{\epsilon} + 1)^D$.*

Lemma 16. *Suppose $A \subset \mathbb{R}^D$ is nonempty and $A \neq \mathbb{R}^D$. Then, A has nonempty boundary.*

Proof. Suppose that A has empty boundary. Then, for every $\boldsymbol{x} \in A$, there exists a sufficiently small ball B containing $\boldsymbol{x}$ such that $B \subset A$ and for every $\boldsymbol{y} \in A^c$, there exists a sufficiently small ball B' containing $\boldsymbol{y}$ such that $B' \subset A^c$. Then, A and A^c are both open sets, which contradicts the assumption that A is nonempty and $A \neq \mathbb{R}^D$. □

Recall that $d(x, y) \coloneqq x \log(\frac{x}{y}) + (1-x)\log(\frac{1-x}{1-y})$.

Lemma 17. *For $x \leqslant .1$,*

$$d(\frac{1-x}{2}, x) = \frac{1-x}{2}\ln(\frac{1-x}{2x}) + \frac{1+x}{2}\ln(\frac{1+x}{2(1-x)}) \geqslant \frac{1}{15}\ln(\frac{1}{2x}).$$

Proof. We note that the term

$$\frac{1+x}{2}\ln(\frac{1+x}{2(1-x)}) = \frac{1+x}{2}(\ln(1+x) - \ln(2(1-x)))$$

is increasing in $x \in (0,1)$. Thus, for all $x \in (0,1)$,

$$\frac{1+x}{2}(\ln(1+x) - \ln(2(1-x))) \geqslant \frac{1}{2}\ln(1/2) \geqslant -.35.$$

Next, for $x \leqslant .1$,

$$\begin{aligned}
\frac{1-x}{2}\ln(\frac{1-x}{2x}) &= \frac{1-x}{2}[\ln(1-x) + \ln(\frac{1}{2x})] \\
&\geqslant \frac{1-x}{2}[\ln(0.9) + \ln(\frac{1}{2x})] \\
&\geqslant \frac{1}{2}\cdot(-0.106) + \frac{1-x}{2}\ln(\frac{1}{2x}) \\
&\geqslant \frac{1}{2}\cdot(-0.106) + \frac{1}{3}\ln(\frac{1}{2x}).
\end{aligned}$$

Then, putting it together, for $x \leqslant .1$,

$$\begin{aligned}
\frac{1+x}{2}\ln(\frac{1+x}{2(1-x)}) + \frac{1-x}{2}\ln(\frac{1-x}{2x}) &\geqslant \frac{1}{3}\ln(\frac{1}{2x}) - 0.35 - \frac{1}{2}\cdot(0.106) \\
&\geqslant \frac{1}{3}\ln(\frac{1}{2x}) - \frac{4}{15}\ln(\frac{1}{2x}) \\
&= \frac{1}{15}\ln(\frac{1}{2x}).
\end{aligned}$$

where we used the fact that $\frac{4}{15}\ln(\frac{1}{2x}) \geqslant 0.403$ for all $x \leqslant 0.1$. □

The following Lemma is from Kaufmann et al. (2016b).

Lemma 18. *For any $x \in [0,1]$, $d(x, 1-x) \geqslant \ln(\frac{1}{2.4x})$.*

Lemma 19. *Let $P \subset \mathbb{R}^D$ and let $\boldsymbol{x} \in P$ and $\boldsymbol{y} \in P^c$. Then, there exists $\theta \in [0,1]$ such that $\theta\boldsymbol{x} + (1-\theta)\boldsymbol{y} \in \partial P$.*

Proof. Since $\boldsymbol{x} \in P$ and $\boldsymbol{y} \in P^c$, by Lemma 16, $\partial P \neq \varnothing$. Consider the following sequence, which resembles binary search.

$$\begin{aligned}
\boldsymbol{x}_0 &= \boldsymbol{x} \\
\boldsymbol{x}_1 &= \boldsymbol{y} \\
\boldsymbol{x}_2 &= \frac{1}{2}(\boldsymbol{x} + \boldsymbol{y}) \\
\boldsymbol{x}_n &= \begin{cases} \frac{1}{2}\boldsymbol{x}_{n-1} + \frac{1}{2}\boldsymbol{x}_{\min(k:\forall l \in \{k+1,\ldots,n-1\}, \boldsymbol{x}_l \in P)} & : \boldsymbol{x}_{n-1} \in P \\ \frac{1}{2}\boldsymbol{x}_{n-1} + \frac{1}{2}\boldsymbol{x}_{\min(k:\forall l \in \{k+1,\ldots,n-1\}, \boldsymbol{x}_l \in P^c)} & : \boldsymbol{x}_{n-1} \in P^c \end{cases}.
\end{aligned}$$

$\{\boldsymbol{x}_n\}$ is clearly a Cauchy sequence so that it has a a limit $\bar{\boldsymbol{x}} = \theta \boldsymbol{x} + (1-\theta)\boldsymbol{y} \in \partial P$ for some $\theta \in [0,1]$. If for every $N \in \mathbb{N}$, there exist $n, m \geqslant N$ such that $\boldsymbol{x}_n \in P$ and $\boldsymbol{x}_m \in P^c$, then it is clear that $\bar{\boldsymbol{x}} \in \partial P$. Suppose that there exists N such $\boldsymbol{x}_N \in P$ and for every $n > N$, $\boldsymbol{x}_n \notin P$ (the other case is similar). Then, it is clear that $\bar{\boldsymbol{x}} = \boldsymbol{x}_N$ and that every open ball containing $\bar{\boldsymbol{x}}$ contains some point not in P, so that $\bar{\boldsymbol{x}} \in \partial P$.

□

We use the anytime confidence interval from Kaufmann et al. (2016b).

Lemma 20. *Let $X_1, X_2, \ldots$ be i.i.d. zero-mean sub-Gaussian random variables with scale $\sigma > 0$ and $\delta \in (0,1)$. Then,*

$$\Pr(\exists t : |\frac{1}{t}\sum_{s=1}^{\top} X_s| \geqslant \sigma\sqrt{\frac{2\log(1/\delta) + 6\log\log(1/\delta) + 3\log\log(et)}{t}}) \leqslant \delta.$$

Recall that $U(t,\delta) = \sigma\sqrt{\frac{2\log(1/\delta)+6\log\log(1/\delta)+3\log\log(et)}{t}}$. We use the following fact from Jamieson and Jain (2018).

Lemma 21. *Let $\Delta \in (0,1)$ and $\delta \in (0,1)$. There is a universal constant $c > 0$ such that if*

$$N \geqslant c\Delta^{-2}\log(\frac{\log(\Delta^{-2})}{s})$$

then $U(N,s) \leqslant \Delta$.

3.15 TF-LUCB with Tolerance

In this section, we present a variant of TF-LUCB that tolerates some violation of the constraints and some suboptimality: TF-LUCB-Tol. TF-LUCB-Tol also takes as input two

scalars ϵ_P and $\epsilon_{\boldsymbol{c}}$, which quantify how much the algorithm tolerates a violation of the constraints and suboptimality, respectively. The main difference is that TestF-Tol also takes as input ϵ_P and the stopping condition associated with the rewards is now

$$\min_{i \in \mathrm{TOP}_t} \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{i,N_i(t)} - U_{\boldsymbol{c}}(N_i(t),\delta) + \epsilon_{\boldsymbol{c}} \geqslant \max_{j \in \mathrm{TOP}_t^c \cap E_t} \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{j,N_j(t)} + U_{\boldsymbol{c}}(N_j(t),\delta).$$

Next, we introduce variants of TestF-B and TestF-C that allow for a tolerance. TestF-B-Tol now returns True if $B_{U_{\mathrm{ball}}(t,\delta)}(\widehat{\boldsymbol{\mu}}_{i,t})$ intersects P and P^c and $U_{\mathrm{ball}}(t,\delta) \leqslant \frac{\epsilon_P}{2}$. Since the diameter of $B_{U_{\mathrm{ball}}(t,\delta)}(\widehat{\boldsymbol{\mu}}_{i,t})$ is $2\,U_{\mathrm{ball}}(t,\delta)$, this guarantees that on an event where the confidence bounds work appropriately, we only accept $\boldsymbol{\mu}_i$ such that $\mathrm{dist}(\boldsymbol{\mu}_i, P) \leqslant \epsilon_P$. TestF-C-Tol tolerates violations on a constraint-basis instead. Now, it accepts arms if $U_{\mathrm{con}}(t,\delta) \leqslant \frac{\epsilon_P}{2}$. Thus, assuming an event on which the confidence bounds work appropriately, it only tolerates mistakes on arms such that for every constraint $j \in [M]$, $\boldsymbol{a}_j^\top \boldsymbol{\mu}_i \leqslant b_j + \epsilon_P$.

Algorithm 9 TestF-B-Tol:

Input: arm index i, number of pulls t, ϵ_P
if $\mathrm{dist}(\widehat{\boldsymbol{\mu}}_{i,t}, P^c) > U_{\mathrm{ball}}(t,\delta)$ **then**
  return True
else if $\mathrm{dist}(\widehat{\boldsymbol{\mu}}_{i,t}, P) > U_{\mathrm{ball}}(t,\delta)$ **then**
  return False
else if $U_{\mathrm{ball}}(t,\delta) \leqslant \frac{\epsilon_P}{2}$ **then**
  return True
else
  return ?
end if

Algorithm 10 TestF-C-Tol:

Input: arm index i, number of pulls t, ϵ_P
if $A\widehat{\boldsymbol{\mu}}_{i,t} + U_{\mathrm{con}}(t,\delta)\mathbf{1} \leqslant \boldsymbol{b}$ **then**
  return True
else if $A\widehat{\boldsymbol{\mu}}_{i,t} - U_{\mathrm{con}}(t,\delta)\mathbf{1} \not\leqslant \boldsymbol{b}$ **then**
  return False
else if $U_{\mathrm{con}}(t,\delta) \leqslant \frac{\epsilon_P}{2}$ **then**
  return True
else
  return ?
end if

Proving the upper bound for this algorithm would have a similar structure to what we have done in this paper. One subtlety is that finding the top m feasible arms depends on which arms we consider to be feasible so that accepting as feasible an arm that is in fact infeasible might make the problem more difficult. We conjecture that the upper bound would reflect this subtlety. We leave the proof of an upper bound of this to future work.

However, as a practical consideration, we also note that accepting as feasible an arm that is in fact infeasible might make the problem much easier. We conjecture that in most

Algorithm 8 TF-LUCB-Tol: Top-m Feasible Upper Confidence Bound algorithm

Input: TestF, sub-Gaussian norm bound σ, confidence δ, ϵ_P, $\epsilon_{\boldsymbol{c}}$
for $t = 1, 2, \ldots$ **do**
  $F_t \longleftarrow \{i \in [K] : \text{TestF}(i, N_i(t), \epsilon_P) = \text{True}\}$ # arms that are determined to be feasible whp
  $G_t \longleftarrow \{i \in [K] : \text{TestF}(i, N_i(t), \epsilon_P) = ?\}$# arms that have not be determined to be feasible or infeasible whp
  $E_t \longleftarrow F_t \cup G_t$ # arms that are not ruled out as infeasible whp
  $\text{TOP}_t \longleftarrow \arg\max_{Z \subset E_t, |Z| = \min(m, |E_t|)} \sum_{i \in Z} \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{i, N_i(t)}$
  if $\text{TOP}_t = F_t$ and $F_t = E_t$ **then**
    return $(\text{TOP}_t, \text{TOP}_t^c \cap E_t, E_t^c)$
  else if $\text{TOP}_t \subset F_t$ and $\min_{i \in \text{TOP}_t} \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{i, N_i(t)} - U_{\boldsymbol{c}}(N_i(t), \delta) + \epsilon_{\boldsymbol{c}} \geqslant \max_{j \in \text{TOP}_t^c \cap E_t} \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{j, N_j(t)} + U_{\boldsymbol{c}}(N_j(t), \delta)$ **then**
    return $(\text{TOP}_t, \text{TOP}_t^c \cap E_t, E_t^c)$
  else if $\text{TOP}_t \subset F_t$ **then**
    $h_t = \arg\min_{i \in \text{TOP}_t} \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{i, N_i(t)} - U_{\boldsymbol{c}}(N_i(t), \delta)$
  else if $\text{TOP}_t \not\subset F_t$ **then**
    $h_t = \arg\min_{i \in \text{TOP}_t \cap G_t} \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{i, N_i(t)} - U_{\boldsymbol{c}}(N_i(t), \delta)$
  end if
  if $\text{TOP}_t^c \cap E_t \neq \varnothing$ **then**
    $l_t = \arg\max_{j \in \text{TOP}_t^c \cap E_t} \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{j, N_j(t)} + U_{\boldsymbol{c}}(N_j(t), \delta)$
    Pull arm l_t
  end if
  Pull arm h_t
end for

applications, there is no a priori reason to believe that doing this would make the problem easier or more difficult. Furthermore, this issue could be somewhat alleviated by allowing a tolerance for suboptimality.

3.16 Pseudocode for algorithms TF-AE and FFAF

Algorithm 11 TF-AE: Top-m Feasible Action Elimination

Input: TestF, sub-Gaussian norm bound σ, confidence δ
$t \longleftarrow 1$
while True **do**

  $F_t \longleftarrow \{i \in [K] : \text{TestF}(i, N_i(t)) = \text{ True}\}$ # arms that are determined to be feasible whp
  $G_t \longleftarrow \{i \in [K] : \text{TestF}(i, N_i(t)) = ?\}$# arms that have not be determined to be feasible or infeasible whp
  $E_t \longleftarrow F_t \cup G_t$ # arms that are not ruled out as infeasible whp
  $H_t \longleftarrow \{i \in [K] : |\{j \in F_t : \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{j,N_j(t)} - U_{\boldsymbol{c}}(N_j(t), \delta) \geqslant \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{i,N_i(t)} + U_{\boldsymbol{c}}(N_i(t), \delta)\}| < m\}$
  $Q_t \longleftarrow E_t \cap H_t$
  for $i \in Q_t$ **do**
    pull arm i
    $t \longleftarrow t + 1$
  end for
  if $E_t = F_t$ and $|F_t| < m$ **then**
    return F_t
  else if $Q_t \subset F_t$ and $|Q_t| = m$ **then**
    return Q_t
  end if
end while

For FFAF, we require that it find the the feasible arms with probability at least $1 - \frac{\delta}{2}$ and, then, to find the best arms among those with probability at least $1 - \frac{\delta}{2}$. Thus, we require that TestF output the correct answer with probability at least $1 - \frac{\delta}{2K}$. We modify the confidence bound for the rewards in the second stage since in that stage there are only $|F_t|$ arms among which the m arms with the largest rewards must be identified.

Algorithm 12 FFAF: Find Feasible Arms First

Input: TestF, sub-Gaussian norm bound σ, confidence δ
$t \longleftarrow 1$
for $i \in [K]$ **do**
  while TestF$(i, N_i(t)) =?$ **do**
    Pull arm i
    $t \longleftarrow t + 1$
  end while
end for
$F_t \longleftarrow \{i \in [K] : \text{TestF}(i, N_i(t)) = \text{ True}\}$ # arms that are determined to be feasible whp
if $|F_t| \leqslant m$ **then**
  return F_t
end if
$U_{\boldsymbol{c}}(s, \delta) \coloneqq U(s, \frac{\delta}{2|F_t|})$
while True **do**
  $\text{TOP}_t = \arg\max_{Z \subset F_t, |Z|=m} \sum_{i \in Z} \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{i,N_i(t)}$
  if $\min_{i \in \text{TOP}_t} \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{i,N_i(t)} - U_{\boldsymbol{c}}(N_i(t), \delta) \geqslant \max_{j \in \text{TOP}_t^c \cap F_t} \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{j,N_j(t)} + U_{\boldsymbol{c}}(N_j(t), \delta)$ **then**
    return TOP_t
  end if
  $h_t = \arg\min_{i \in \text{TOP}_t} \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{i,N_i(t)} - U_{\boldsymbol{c}}(N_i(t), \delta)$
  $l_t = \arg\max_{j \in \text{TOP}_t^c \cap F_t} \boldsymbol{c}^\top \widehat{\boldsymbol{\mu}}_{j,N_j(t)} + U_{\boldsymbol{c}}(N_j(t), \delta)$
  Pull arms h_t and l_t
  $t \longleftarrow t + 1$
end while

Chapter 4

The True Sample Complexity of Identifying Good Arms

We consider two multi-armed bandit problems: *(i)* given an $\epsilon > 0$, identify an arm with mean that is within ϵ of the largest mean and *(ii)* given a threshold μ_0 and k, minimize the time to identify k arms with means larger than μ_0. Existing lower bounds and algorithms for the PAC framework suggest that both of these problems require $\Omega(n)$ samples. However, we argue that these definitions not only conflict with how these algorithms are used in practice, but also that these results disagree with strong intuition that says *(i)* requires only $\Theta(\frac{n}{m})$ samples where $m = |\{i : \mu_i > \max_{i \in [n]} \mu_i - \epsilon\}|$ and *(ii)* requires $\Theta(\frac{n}{m}k)$ samples where $m = |\{i : \mu_i > \mu_0\}|$. We provide definitions that formalize these intuitions, obtain lower bounds that match the above sample complexities, and develop explicit, practical algorithms that achieve nearly matching upper bounds. This Chapter is joint work with Kevin Jamieson.

4.1 Introduction

Define a multi-armed *bandit instance* ρ as a collection of n distributions over $\mathbb{R}$ where the jth sample from the ith distribution ρ_i is an iid random variable $X_{i,j} \sim \rho_i$ with $\mathbb{E}[X_{i,j}] = \mu_i$. Proceeding in rounds, at round t a player selects an index $I_t \in [n] := \{1, \dots, n\}$, immediately observes $X_{I_t,t}$, and then outputs a set $\widehat{S}_t \subseteq [n]$. With respect to the filtration $(\mathcal{F}_t)_{t \in \mathbb{N}}$

where $\mathcal{F}_t = \{(I_s, X_{I_s,s}, \widehat{S}_s) : 1 \leqslant s \leqslant t\}$, I_t is $\mathcal{F}_{t-1}$ measurable while $X_{I_s,s}$ and $\widehat{S}_t$ are $\mathcal{F}_t$ measurable, each with possibly additional external sources of randomness. Informed by past observations $\mathcal{F}_{t-1}$ up to time $t-1$, at round t the player chooses which distribution ρ_{I_t} to observe a sample from in order to accomplish a predetermined objective for $\widehat{S}_t$ as quickly as possible (i.e., minimizing the total number of observed samples).

Two important problems that arise in this setting are *(i)* identifying an arm with the largest mean (commonly referred to as best arm identification) and *(ii)* identifying all of the arms with means above a given threshold $\mu_0 \in \mathbb{R}$. Applications of *(i)* include drug or material design in the presence of noisy experiments. Applications of *(ii)* include genetic screens where individual genes are inhibited to infer a causal dependence to some experimentally measured phenotype relative to a control wild-type phenotype; typically multiple genes are identified as influencing the phenotype.

Unfortunately solving these problems requires an impractical number of samples, especially if the number of arms is large. In practice, it is often sufficient to solve relaxations of these problems. For example, if it enables significant savings, identifying a nearly optimal material design is satisfactory and, similarly, a scientist is typically content to discover a few of the important genes for a phenotype. These relaxations can be formalized as follows:

1. **Identifying an ϵ-good mean**: for a given $\epsilon > 0$, minimize τ such that the index $\widehat{S}_t \in [n]$ satisfies $\mu_{\widehat{S}_t} > \max_{i \in [n]} \mu_i - \epsilon$ for all $t \geqslant \tau$ with high probability.

2. **Identifying means above a threshold μ_0**: for a given threshold $\mu_0 \in \mathbb{R}$ and $k \in [n]$, minimize τ_k such that the set $\widehat{S}_t \subseteq [n]$ satisfies $|\widehat{S}_t \cap \{i : \mu_i > \mu_0\}| \geqslant \min(k, |\{i : \mu_i > \mu_0\}|)$ for every $t \geqslant \tau_k$ subject to $\widehat{S}_s \cap \{i : \mu_i \leqslant \mu_0\} = \varnothing$ for all times s with high probability[1].

Note that in the first problem we allow $\widehat{S}_t \notin \{i : \mu_i > \mu_1 - \epsilon\}$ for some times t but for the second we require $\widehat{S}_t \subseteq \{i : \mu_i > \mu_0\}$ for all times t with high probability. That is, the first problem allows mistakes while the second does not.

[1]The constraint $\widehat{S}_s \cap \{i : \mu_i \leqslant \mu_0\} = \varnothing$ is known as a family-wise error rate (FWER) condition. We will also consider a more relaxed condition known as false discovery rate (FDR) which controls the size of $\mathbb{E}[|\widehat{S}_s \cap \{i : \mu_i \leqslant \mu_0\}|/|\widehat{S}_s|]$

The goal of this work is to characterize the number of samples that are necessary and sufficient to achieve these objectives. The difficulty of the objectives are closely related, which we now describe. For a fixed set of means $\mu_1 \geqslant \cdots \geqslant \mu_n$ and any threshold μ_0 there exists an $\epsilon = \mu_1 - \mu_0$ so that $\{\mu_i : \mu_i > \mu_1 - \epsilon\} = \{\mu_i : \mu_i > \mu_0\}$. Thus, identifying k arms above the threshold μ_0 is equivalent to identifying k ϵ-good means for $\epsilon = \mu_1 - \mu_0$. Consequently, if $m = |\{i \in [n] : \mu_i > \mu_1 - \epsilon\}|$ then we can study *lower bounds* on the sample complexity of both problems simultaneously by considering the necessary number of samples required to identify k of the m largest means (i.e., to have $\widehat{S}_t \subset [m]$ with $|\widehat{S}_t| = k$) for any value of $1 \leqslant k \leqslant m$. However, while considering m is a useful tool for analysis, it should be stressed that ***m* is never known to the algorithm and must be adapted to** and, in fact, such knowledge significantly simplifies this problem. For the following discussion, one can take $k = 1$ for simplicity.

If there is just a single ϵ-good distribution out of the total n so that $m = 1$, then any strategy will have to sample from about n distributions before coming across the unique ϵ-good distribution. However, if there are $1 < m \leqslant n$ means which are ϵ-good then choosing to sample from distributions uniformly at random will eventually take a sample from an ϵ-good distribution about m times faster than if there were just a single good distribution. Thought of a different way, if at least n/m indices are drawn uniformly at random from $\{1, \ldots, n\}$ then with constant probability at least one of them will be ϵ-good. Thus, when there are m ϵ-good distributions, one should expect the number of samples to identify an ϵ-good distribution to scale as n/m, not n. **In an extreme case, if m is $\Theta(n)$ so that a linear number of the distributions are ϵ-good, then one should expect the number of samples to identify an ϵ-good distribution to be *constant* with respect to *n*.** Similarly, if there are m means above the threshold μ_0, then one would expect the number of samples required to identify at least $1 \leqslant k \leqslant m$ of them should scale like $k\frac{n}{m}$, not n.

Unfortunately, confirmation of this intuition in sample complexities that depend on m and are potentially independent of n is almost entirely absent from the existing literature for both of the above problems. In the case of identifying arms above a threshold, only the case of identifying all $k = m$ arms above the threshold has been addressed, ignoring the reality that only rarely is it possible to identify all of m as they require often impractical $\Omega(n)$

sample complexities. In the case of identifying an ϵ-good arm, using the definitions accepted hitherto by the community, one can prove information theoretic lower bounds that conflict with the above intuition, as we discuss below. **This chapter aims to develop definitions, algorithms, and lower bounds that confirm the necessary and sufficient conditions for obtaining the intuitive sample complexities expected.** Identifying means above a threshold will be addressed in the coming sections. But for now we turn our attention to identifying an ϵ-good arm.

4.1.1 Revisiting identifying an ϵ-good arm

Identifying an ϵ-good arm with high probability in the multi-armed bandit literature has been studied by dozens of works over the past several decades. We begin by considering the standard definition under which the sample complexity of identifying an ϵ-good arm is characterized.

Definition 5. *We say an algorithm is* $(\boldsymbol{\epsilon, \delta})$***-PAC (Probably Approximately Correct)*** *on $\mathcal{P}$ if, when executed on* any *bandit instance $\rho \in \mathcal{P}$, with probability at least $1 - \delta$ the algorithm terminates at a stopping time τ_{PAC} with respect to $(\mathcal{F})_{t \in \mathbb{N}}$ and outputs an ϵ-good arm.*

One typically takes $\mathcal{P}$ to be quite large, like the space of all distributions with sub-Gaussian tails, or for lower bounds $\mathcal{P} = \{\mathcal{N}(\mu', \sigma^2 I) : \mu' \in \mathbb{R}^n\}$. For a given ϵ, δ, instance ρ, and an instance space $\mathcal{P}$, one lower bounds $\mathbb{E}_\rho[\tau_{PAC}]$ for any (ϵ, δ)-PAC algorithm. Specifically, known lower bounds Kaufmann et al. (2016a); Mannor et al. (2004) imply that for a sufficiently rich $\mathcal{P}$ we have that **any $(\boldsymbol{\epsilon, \delta})$-PAC algorithm has an expected sample complexity $\mathbb{E}[\tau_{PAC}]$ of at least $\boldsymbol{n}$, regardless of $\boldsymbol{m}$** the number of ϵ-good distribuions there are among the n. This is necessary because an (ϵ, δ)-PAC algorithm must prove that any output arm $\widehat{i}$ satisfies $\mu_{\widehat{i}} \geqslant \mu_1 - \epsilon$, but the value of μ_1 is not known a priori, so the algorithm must pull *every* arm about at least once to *verify* that $\widehat{i}$ is indeed ϵ-good[2]. Contrast this with the above discussion where we were merely concerned with how quickly an

[2]We note that if $\mathcal{P}$ is structured enough like in linear bandits, it is possible to identify and verify an ϵ-good arm with a sample complexity that does not scale linearly with n.

algorithm could start outputting an ϵ-good arm, with no condition of *verifying* that it is ϵ-good.

In this work, we argue that *simple regret* is the metric of merit Bubeck et al. (2011).

Definition 6. *We say an algorithm is* $(\boldsymbol{\epsilon, \delta})$***-SimplePAC*** *on* $\mathcal{P}$ *if, when executed on* any *bandit instance* $\rho \in \mathcal{P}$*, there exists a stopping time* τ_{simple} *wrt* $(\mathcal{F})_{t\in\mathbb{N}}$ *such that* $\mathbb{P}(\tau_{simple} < \infty) = 1$ *and with probability at least* $1 - \delta$ *the algorithm outputs an* ϵ*-good arm at every* $t \geqslant \tau_{simple}$*.*

Analogous to above, for a given $\mathcal{P}$ and $\rho \in \mathcal{P}$ we would like to lower bound $\mathbb{E}_\rho[\tau_{simple}]$ for any (ϵ, δ)-SimplePAC algorithm. We emphasize τ_{simple} is for *analysis purposes only* and *is unknown to the algorithm*, and the algorithm never stops taking samples and recommending sets $\widehat{S}_t$. Clearly, if an algorithm is (ϵ, δ)-PAC for an instance ρ then it is also (ϵ, δ)-SimplePAC for ρ since we may take $\tau_{simple} = \tau_{PAC}$ and output the arm identified at τ_{PAC} at all $t \geqslant \tau_{PAC}$. However, the above discussion suggests that τ_{simple} may be significantly smaller than τ_{PAC} since the former is just a time after which an algorithm only outputs ϵ-good arms, and the latter is when the algorithm provides an additional guarantee that that recommended arm is actually ϵ-good. Indeed, when $m = n/2$ then we will see examples where $\tau_{simple} = O(1)$ while $\tau_{PAC} = \Omega(n)$. To define such a stopping time τ_{simple}, we will use knowledge of the behavior of the algorithm (e.g., how $\widehat{S}_t$ is constructed) and of the true means of the problem instance. In particular, we will essentially define τ_{simple} to be the first time that some ϵ-good arm i is pulled on the order of $(\mu_1 - \epsilon - \mu_i)^{-2}$ times and we will show that whp for all $t \geqslant \tau_{simple}$ the algorithm only outputs ϵ-good arms. The next section provides real-world case studies that illustrate how algorithms designed for PAC are inferior in practice relative to our proposed algorithm that is designed for SimplePAC.

4.1.2 Related work

Lower bounds: $(\epsilon, \delta) - PAC$ **for identifying** k ϵ**-good arms**

Kaufmann et al. (2016a) proved the following theorem which characterizes the sample complexity for ϵ-good arm identification $k = 1, m \geqslant 1$ and multiple identifications above a

threshold μ_0 in the special case of $k = m$ (in general, we are interested in any $1 \leqslant k \leqslant m$) in the (ϵ, δ)-PAC setting.

Theorem 15 (Kaufmann et al. (2016a)). *Fix $\epsilon, \delta > 0$, and a vector $\mu \in \mathbb{R}^n$. Fix a bandit instance ρ of n arms where the ith distribution equals $\rho_i(\mu) = \mathcal{N}(\mu_i, 1)$, a Gaussian distribution with mean μ_i and variance 1. Assume without loss of generality that $\mu_1 \geqslant \mu_2 \geqslant \cdots \geqslant \mu_n$ and let $m = |\{i \in [n] : \mu_i \geqslant \mu_1 - \epsilon\}|$ so that $\mu_i \geqslant \mu_1 - \epsilon$ for all $i \in [m]$. If algorithm $\mathcal{A}$ returns $k = 1$ arms of the top m arms and is (ϵ, δ)-PAC on $\mathcal{P} = \{\mathcal{N}(\mu', I) : \mu' \in \mathbb{R}^n\}$ then*

$$\mathbb{E}_\rho\Big[\sum_{i=1}^n T_i(\tau_{PAC})\Big] \geqslant \tfrac{1}{2}\log(1/2.4\delta)\Big((m-1)\epsilon^{-2} + \sum_{i=m+1}^n (\mu_1 - \mu_i)^{-2}\Big) \qquad (k=1)$$

Under the same conditions, if $\mathcal{A}$ returns $k = m$ arms then

$$\mathbb{E}_\rho\Big[\sum_{i=1}^n T_i(\tau_{PAC})\Big] \geqslant 2\log(1/2.4\delta)\Big(\sum_{i=1}^m (\mu_i - \mu_{m+1})^{-2} + \sum_{i=m+1}^n (\mu_m - \mu_i)^{-2}\Big) \quad (k=m)$$

Note that by the definition of m we have that $\mu_m - \mu_{m+1} > 0$. We emphasize that the sample complexity of Theorem 15 for both $k = 1$ or $k = m$ is necessarily $\Omega(n)$ regardless of the number of ϵ-good arms m. As discussed below, the $k = 1$ lower bound is achievable up to $\log\log$ factors Karnin et al. (2013). The special case of $k = m$ is notably the TOP-k identification problem where lower bounds were recently sharpened with additional log factors independently by Simchowitz et al. (2017b); Chen et al. (2017b). In particular, if for some μ_0 we have $\mu_i = \mu_0 + \epsilon$ for $i \leqslant m$ and $\mu_i = \mu_0$ for $i < m$ then their lower bounds on the expected sample complexity scale like $k\epsilon^{-2}\log(n-k) + (n-k)\epsilon^{-2}\log(k)$, which is always larger than $n\epsilon^{-2}$ that is predicted by the above theorem.

Lower bounds: Towards (ϵ, δ)-SimplePAC

The limitations of a definition like PAC that requires *verifying* that the correct hypothesis is returned has been observed before in the active binary classification literature Balcan et al. (2010). To illustrate, consider a "noiseless" bandit game where sampling the ith arm returns a deterministic value $\mu_i \in \{0, 1\}$. If the player knew that the μ_i were non-increasing (i.e., $\mu_i = 1$ for $i \leqslant m$ and $\mu_i = 0$ for $i > m$) but not the number of 1's equal to m, then binary search could identify the top-m arms (and the index of the transition) with just $\log_2(n)$ total

samples. However, if the player *didn't* know they were non-decreasing, then to even *verify* that the first m were indeed 1 and the last $n-m$ means were 0 would require n samples. The idea is that the player could run binary search and in the favorable case in which the first m arms were equal to 1 and the remaining 0, the algorithm could start outputting the top-m arms with just $\log_2(n)$ pulls (SimplePAC) but not be able to *verify* it with a certificate until n pulls are taken (PAC).

In the Simple-PAC definition, for a given instance ρ we wish to lower bound $\mathbb{E}_\rho[\tau_{simple}]$. However, we would like to avoid trivial algorithms that output an index deterministically that happens to be an ϵ-good arm by sheer luck. To remove these trivial cases, we adopt the random permutation model of Simchowitz et al. (2017b); Chen et al. (2017b). Let $\mathbb{S}^n$ denote the set of permutations over $[n]$ so that for any $\pi \in \mathbb{S}^n$, $\pi(i)$ denotes the index that i is mapped to under π. For a bandit instance $\rho = (\rho_1, \ldots, \rho_n)$ let $\pi(\rho) = (\pi(\rho_1), \pi(\rho_2), \ldots, \pi(\rho_n))$ so that $E_{\pi(\rho)}[T_{\pi(i)}(t)]$ denotes the expected number of samples taken by the algorithm up to time t from the arm with mean μ_i when run on instance $\pi(\rho)$. At the start of the game the algorithm is given a permutation π of the arms chosen uniformly at random from $\mathbb{S}^n$ which we denote as $\pi \sim \mathbb{S}^n$. Thus, the sample complexity of interest is the expected number of samples taken by the algorithm under $\pi(\rho)$ averaged over all possible $\pi \in \mathbb{S}^n$, which we denote as $\mathbb{E}_{\pi\sim\mathbb{S}^n}[\mathbb{E}_{\pi(\rho)}[\tau_{Simple}]] = \mathbb{E}_{\pi\sim\mathbb{S}^n}[\mathbb{E}_{\pi(\rho)}[\sum_{i=1}^n T_{\pi(i)}(\tau_{Simple})]]$.

The works of Simchowitz et al. (2017b); Chen et al. (2017b) considered this model independently to improve the lower bounds of the Top-k problem discussed above (i.e., $k = m$). Remarkably, they proved lower bounds for algorithms that knew the set of means $\{\mu_1, \ldots, \mu_n\}$ but just not which mean is assigned to which arm. The conclusions of these works was that if merely constant confidence is desired (e.g., $\delta = 0.01$) then the sample complexity for these two settings (the means known a priori, or not) are equal up to constants. To obtain this result, at the heart of the analysis of Chen et al. (2017b) is a clever application of Lemma 1 of Kaufmann et al. (2016a). On the other hand, Simchowitz et al. (2017b) obtained the result for best-arm identification using a more general framework known as the *Simulator* argument that we appeal to in this work.

This previous paragraph is still in the PAC setting and by the above theorem, cases $k = 1$ and $k = m$ have a sample complexity of $\Omega(n)$. As suggested in the introduction, when

$k < m$ we should expect that sample complexities of $\Theta(kn/m)$ are possible for the Simple-PAC setting. In the context of proving lower bounds, going from identifying *all* ϵ-good arms ($k = m$) to merely $1 \leqslant k < m$ ϵ-good arms is non-trivial because whereas in the former there was only one way for the algorithm to be correct, in the latter there are precisely $\binom{m}{k}$ subsets of ϵ-good arms that are equally correct.

Related problem settings have also aimed to identify multiple arms Chaudhuri and Kalyanakrishnan (2017, 2019). Specifically, given a tolerance $\eta \geqslant 0$, they say an arm i is (η, m)-optimal if $\mu_i \geqslant \mu_m - \eta$. The objective, given m and η as inputs to the algorithm, is to identify k (η, m)-optimal arms with probability at least $1 - \delta$. The case when $\eta = 0$ and $m = |\{i : \mu_i > \mu_1 - \epsilon\}|$ coincides with our setting, with the critical difference that in our setting the algorithm never has knowledge of m. Note that our setting is significantly more difficult because we have no guide to how many arms we need to sample in order to obtain an ϵ-good arm. Nevertheless, this setting is still relevant from a lower bound perspective for which they prove *worst-case* instance results for $\eta > 0$. In contrast, in this work we prove instance-specific lower-bounds that directly apply to their setting.

Upper bounds: Algorithms for ϵ-good identification

While for lower bounds we can characterize the difficulty of identifying $k = 1$ ϵ-good arm or $1 \leqslant k \leqslant m$ arms above a threshold μ_0 with a single unifying problem statement as the difficulty of identifying k of the m largest arms, for upper bounds we require explicit algorithms that differ slightly depending on the problem in question. We first address the problem of identifying an ϵ-good arm (i.e., $k = 1$ arms within top m).

With a union bound, one can easily design an (ϵ, δ)-PAC algorithm with sample complexity of $n\epsilon^{-2}\log(n/\delta)$ that samples each arm the same number of times and can identify any number $1 \leqslant k \leqslant m$ of ϵ-good arms. Remarkably, Even-Dar et al. (2006) proposed Median Elimination, a (ϵ, δ)-PAC algorithm with sample complexity $n\epsilon^{-2}\log(1/\delta)$. In addition, Even-Dar et al. (2006) proposed a very simple elimination or "Chernoff racing" strategy that has a sample complexity of $\sum_{i=1}^{n} \Delta_{i,\epsilon}^{-2} \log(n \log(\Delta_{i,\epsilon}^{-2}))$ where $\Delta_{i,\epsilon} = \max\{\max_j \mu_j - \mu_i, \epsilon\}$. A decade later Karnin et al. (2013) showed a gap-dependent sample complexity without the extraneous $\log(n)$, i.e., $\sum_{i=1}^{n} \Delta_{i,\epsilon}^{-2} \log(\log(\Delta_{i,\epsilon}^{-2}))$ which notably used Median Elimination as

a sub-routine but was impractical for real-world use (the leading constants are in the tens of thousands and are reflected in the empirical performance of the algorithm). In the years that followed, for the best-arm identification scenario with $\epsilon = 0$, numerous works proposed alternative schemes to Karnin et al. (2013) that achieved equally favorable sample complexity upper bounds (up to constants) but were practically useful Jamieson et al. (2014b), shaved off $\log\log$ factors Chen et al. (2017b), and optimized asymptotic constants Garivier and Kaufmann (2016).

A closely related problem is known as the **infinite armed-bandit problem** where the player has access to an infinite pool of arms such that when a new arm is requested, its mean is drawn iid from a distribution ν. Nearly all of this work makes parametric assumptions about ν in some way Berry et al. (1997); Wang et al. (2009); Carpentier and Valko (2015); Chandrasekaran and Karp (2014); Jamieson et al. (2016). For example, for a drawn arm with random mean μ it is assumed $\mathbb{P}(\mu \leqslant x) \geqslant c(x - \mu_*)^\beta$ for some fixed parameters c, μ_*, β that are known (or not). Given an infinite armed bandit algorithm that can identify an ϵ-good arm for arbitrary arm distribution ν could be used, in principle, to solve the problem of interest of this paper by taking $\nu(x) = \frac{1}{n}\sum_{i=1}^n \mathbf{1}\{\mu_i \leqslant x\}$. The only algorithm that we are aware of that provides guarantees for general arm distribution ν is Hyperband of Li et al. (2017). However, this algorithm was designed for a much more general setting and cannot leverage the particular stochastic structure of our problem (e.g., empirical Bernstein confidence intervals) and was not designed to take a fixed confidence δ as input, and it thus could waste samples trying to drive its error probability to zero when just δ sufficed. Nevertheless, our proposed algorithms are inspired by Hyperband's hedging strategy of considering different-sized sets of randomly sampled arms.

Upper bounds: Algorithms for identifying means above μ_0.

As alluded to before, in the special case of $k = m$ this is known as the Top-k problem. However, algorithms for Top-k require k as input, but of course, m is unknown so it is impossible to set $k = m$ a priori, making these algorithm inapplicable in practice. Thus, the relevant problem statement here is identifying $1 \leqslant k \leqslant m$ arms above the known threshold μ_0.

Identifying *all* arms with means above a threshold μ_0 with high probability is known as the **threshold bandit problem** Locatelli et al. (2016b); Mukherjee et al. (2017). Specifically, these works take a time-horizon T and threshold μ_0 as input, and sample arms in such a way as to maximize the probability of classifying all arms correctly above or below the threshold μ_0 (i.e., the $k = m$ setting). These works explicitly assume no arms are *equal* to μ_0 and penalize incorrectly predicting a mean above or below the threshold equally. If an arm's mean is arbitrarily close to, or equal to, μ_0 the required time horizon T of these works to obtain a non-trivial probability of success is unbounded.

The most related work to this paper is Jamieson and Jain (2018) which proposes an algorithm that takes a confidence δ and threshold μ_0 as input and characterizes the total number of samples taken before all $k = m$ arms with means above the threshold are output with probability at least $1 - \delta$ for all future times, that is, the $k = m$ SimplePAC setting. Witnessed by the lower bounds of Simchowitz et al. (2017b); Chen et al. (2014b), the proven sample complexity for this $k = m$ case is essentially tight up to $\log \log$ factors. However, a weakness of this work is that it is silent on the issue of identifying just a subset of size $k \leqslant m$ means above the threshold. This is a significant gap because if some means above the threshold are much larger than others, they will clearly be detected earlier, and it is not clear that the proposed algorithm identifies any $1 \leqslant k \leqslant m$ in a near-optimal way or just all m at a particular time. The work of Jamieson and Jain (2018) also notably relaxed the family-wise error rate (FWER) $\mathbb{P}(\cap_{t=1}^{\infty}\{\widehat{S}_t \subseteq \{i : \mu_i \leqslant \mu_0\}\}) \geqslant 1 - \delta$ to merely a bounded false-discovery rate (FDR): $\max_t \mathbb{E}\left[\frac{|\widehat{S}_t \cap \{i:\mu_i \leqslant \mu_0\}|}{|\widehat{S}_t|}\right] \leqslant \delta$. Using the more relaxed error criterion of FDR versus FWER, it was shown that nearly all m arms above the threshold could be identified substantially faster.

4.2 Lower bounds

As pointed out in the introduction, for any threshold μ_0 there exists an $\epsilon = \mu_1 - \mu_0$. Thus, if $m = |\{i : \mu_i > \mu_1 - \epsilon\}|$ then a lower bound for identifying an ϵ-good arm or k arms above a threshold μ_0 is implied by a lower bound to identify k arms among the m largest means for any $1 \leqslant k \leqslant m$. The next theorem handles all $1 \leqslant k \leqslant m$ cases simultaneously for a specific

instance (i.e., not worst-case).

Theorem 16. *Fix $\epsilon, \delta > 0$, and a vector $\mu \in \mathbb{R}^n$. Consider n arms where rewards from the ith arm are distributed according to $\mathcal{N}(\mu_i, 1)$, a Gaussian distribution with mean μ_i and variance 1. Assume without loss of generality that $\mu_1 \geqslant \mu_2 \geqslant \cdots \geqslant \mu_n$ and let $m = |\{i \in [n] : \mu_i > \mu_1 - \epsilon\}|$ so that $\mu_i > \mu_1 - \epsilon$ for all $i \in [m]$. For every permutation $\pi \in \mathbb{S}^n$ let τ_π be a stopping time with respect to the filtration generated by the algorithm playing on instance $\pi(\rho)$ at which time the algorithm outputs a set $\widehat{S}_{\tau_\pi} \subseteq [n]$ with $|\widehat{S}_{\tau_\pi}| = k$. If $\mathbb{P}_{\pi(\rho)}(\widehat{S}_{\tau_\pi} \subset \pi([m])) \geqslant 1 - \delta$, then*

$$\mathbb{E}_{\pi \sim \mathbb{S}^n} \mathbb{E}_{\pi(\rho)} \Big[\sum_{i=1}^{n} T_{\pi(i)}(\tau_\pi) \Big] \geqslant (1/8 - \delta)^2 \Big(-(\mu_1 - \mu_{m+1})^{-2} + \tfrac{k}{m} \sum_{i=m+1}^{n} (\mu_1 - \mu_i)^{-2} \Big).$$

Remark 1. *By definition, $(\mu_1 - \mu_{m+1})^{-2} \leqslant \epsilon^{-2}$ so aside from pathological cases such as $\mu_1 - \mu_i \gg \epsilon$ for all $i > m + 1$ the lower bound will be positive and non-trivial. For example, suppose $m \leqslant n/2$ arms have means equal to $\mu_0 + \epsilon$ while the remaining have means equal to μ_0. Then Theorem 16 implies that to identify any k of the top m arms requires about $k\frac{n}{m}\epsilon^{-2}$ samples.*

4.3 Algorithm and Upper Bounds

Algorithm 13 simultaneously handles both the identification of an ϵ-good arm (Line 14) and the identification of multiple arms above a threshold μ_0 (Line 17). This emphasizes the fact that until the first arm is determined to be above a threshold, the algorithm is essentially acting identically to a near-optimal best-arm identification algorithm Jamieson et al. (2014b) since the same arm I_t is pulled for all objectives. Furthermore, if the multiple identifications variant of the algorithm has identified a subset of the arms as belonging to $\mathcal{H}_1$, then it continues to act as a near optimal best arm identification algorithm on the remaining arms. The similarities between the algorithms for these distinct problems reflect the deep connection between these two problems.

We introduce some notation. Let $N_{i,r}(t)$ be the number of times arm i has been pulled in bracket r up to time r. Let $\widehat{\mu}_{i,r,t}$ be the empirical mean of arm i in bracket r after t pulls (for the purposes of simplifying the analysis of the algorithm, observations from arms are not

shared across brackets). Let $U(t,\delta) = c\sqrt{\frac{1}{t}\log(\log(t)/\delta)}$ be an anytime confidence bound (thus, satisfiying for any $r \in \mathbb{N}$ and $i \in [n]$ $\mathbb{P}(\cap_{t=1}^{\infty}|\widehat{\mu}_{i,r,t} - \mu_i| \leqslant U(t,\delta)) \geqslant 1-\delta$) based on the law of the iterated logarithm (LIL), a fundamental component of pure-exploration stochastic bandits for which the constant c is ever-improving Jamieson et al. (2014b); Kaufmann et al. (2016a). We use the term *bracket* to denote a subset of the arms.

The algorithm opens progressively larger brackets over time. It cycles through the open brackets, at each round pulling an arm in the chosen bracket with the largest upper confidence bound. Each bracket l consist of 2^l arms chosen uniformly at random from $[n]$ and is opened after $(l-1)2^{l-1}$ rounds, that is, after the algorithm has pulled each arm in the $l-1$ bracket once. We note that while samples are not shared between brackets to facilitate the theoretical analysis, in practice samples should be shared.

Depending on the objective, the algorithm uses different criteria for suggesting sets of arms. If the goal is to output an ϵ-good arm, then the algorithm chooses an arm O_t that maximizes a lower confidence bound (Line 14). For the problem of multiple identifications above a threshold, various guarantees are possible. In the main body of the paper, we focus on a kind of guarantee that we call *FDR-TPR* (false discovery rate-true positive rate) (Jamieson and Jain, 2018), which guarantees approximate identification in the sense that the expected number of correctly identified arms is large and the expected number of mistakes is small (see Theorem 20 for a precise statement). For this goal, the algorithm builds a set $\mathcal{S}_t$ (Line 17) based on the Benjamini-Hochberg procedure developed for multi-armed bandits in Jamieson and Jain (2018). In the Appendix, we provide algorithms that have stronger guarantees but are less practical.

4.3.1 Upper Bound for Identifying an ϵ-good mean

To make our upper bound more digestible, we use "$\lesssim$" to hide constants and doubly logarithmic factors. In Appendix, we will restate the the theorems with doubly logarithmic factors. For the sake of simplicity, we will assume that the distributions are sub-gaussian with sub-gaussian norm at most 1.

As argued in the introduction, in the SimplePAC setting if there are m ϵ-good arms, one

Algorithm 13 Subsampling Algorithm: ϵ-good arm identification and FDR-TPR

$\delta_r = \frac{\delta}{r^2}$, $\delta'_r = \frac{\delta_r}{6.4\log(36/\delta_r)}$ $R_0 = 0$, $l = 0$, $\mathcal{S}_0 = \varnothing$
for $t = 1, 2, \ldots$ **do**
  if $t \geqslant 2^l l$ **then**
    Let A_{l+1} be a random set of size $\min(2^{l+1}, n)$ in $[n]$
    $l = l + 1$
  end if
  $R_t = [(R_{t-1} + 1) \mathrm{mod} l] + 1$
  if there exists $i \in A_{R_t} \backslash \mathcal{S}_t$ such that $N_{i,R_t}(t) == 0$ **then**
    Pull arm $I_t = \mathrm{argmin}_{i \in A_{R_t} \backslash \mathcal{S}_t} N_{i,R_t}(t)$
  else
    Pull arm $I_t = \mathrm{argmax}_{i \in A_{R_t} \backslash \mathcal{S}_t} \widehat{\mu}_{i,R_t,N_{i,R_t}(t)} + U(N_{i,R_t}(t), \delta)$
  end if
  if Best Arm Identification **then**
    $O_t = \mathrm{argmax}_{i \in A_r \text{ for some } r \leqslant l} \widehat{\mu}_{i,r,N_{i,r}(t)} - U(N_{i,r}(t), \frac{\delta}{|A_r| r^2})$ % Best-arm Thm.17
  else if FDR-TPR **then**
    $s(p) = \{i \in A_{R_t} : \widehat{\mu}_{i,R_t,N_{i,R_t}(t)} - U(N_{i,R_t}(t), \frac{p}{|A_{R_t}|}\delta'_{R_t} \geqslant \mu_0\}$
    $\mathcal{S}_{t+1} = \mathcal{S}_t \cup s(\widehat{p})$ where $\widehat{p} = \max\{p \in [|A_{R_t}|] : |s(p)| \geqslant p\}$ % FDR Thm.20
  end if
end for

would expect to identify an ϵ-good arm about m times faster than if there were just one. It turns out that its a bit more subtle than that because many of those ϵ-good arms may be nearly indistinguishable from μ_{m+1}, the largest mean that is not ϵ-good. To capture this observation in the sample complexity of our algorithm we introduce a novel definition for the gaps of the arms.

Let $\epsilon > 0$ and $\gamma \in (0, \epsilon)$. Define $G_\gamma = \{i \in [n] : \mu_i \geqslant \mu_1 - \gamma\}$ and define the gaps wrt to ϵ and γ as follows:

$$\Delta_{i,\epsilon,\gamma} = \begin{cases} \mu_i - \mu_{m+1} & i \leqslant |G_\gamma| \\ \max(\mu_{|G_\gamma|} - \mu_i, \mu_i - \mu_{m+1}) & |G_\gamma| < i \leqslant m \\ \mu_{|G_\gamma|} - \mu_i & i \geqslant m+1 \end{cases}.$$

We note that as γ decreases on $(\epsilon, 0)$ we have that $|G_\gamma|$ decreases from m to 1. On the other hand, the defined gaps only increase as γ decreases. The gaps of the arms in $\{|G_\gamma|+1, \ldots, m\}$

behave essentially as $\mu_{|G_\gamma|} - \mu_{m+1}$ as demonstrated by the inequality

$$\frac{\mu_{|G_\gamma|} - \mu_{m+1}}{2} \leqslant \max(\mu_{|G_\gamma|} - \mu_i, \mu_i - \mu_{m+1}) \leqslant \mu_{|G_\gamma|} - \mu_{m+1}.$$

Figure 4.1 illustrates the $|G_\gamma|$ sets making it easy to visualize the gaps.

Next, we introduce a function to characterize the number of samples required by an individual bracket to identify an ϵ-good arm. For all $\epsilon > 0$ define

$$\widetilde{U}_\epsilon(\gamma) := \frac{1}{|G_\gamma|}\Big(\sum_{j=1}^{m} \Delta_{j,\epsilon,\gamma}^{-2} \log(\tfrac{n}{|G_\gamma|}/\delta) + \sum_{j=m+1}^{n} \Delta_{j,\epsilon,\gamma}^{-2} \log(1/\delta)\Big),$$
$$\widetilde{U}_\epsilon := \widetilde{U}_\epsilon(\epsilon) \lesssim \frac{1}{m}\Big(\sum_{j=1}^{m} (\mu_i - \mu_{m+1})^{-2} \log(\tfrac{n}{m}/\delta) + \sum_{j=m+1}^{n} (\mu_m - \mu_j)^{-2} \log(1/\delta)\Big)$$

$\widetilde{U}_\epsilon(\gamma)$ bounds the expected number of samples required by a bracket of size $\Theta(\frac{n}{|G_\gamma|})$ to identify an ϵ-good arm when *(i)* one of its arms is γ-good and *(ii)* the empirical means of the arms in the bracket concentrate well enough.

Now, we present our upper bound for $\boldsymbol{\epsilon}$-good identification.

Theorem 17 ($\boldsymbol{\epsilon}$-good identification)**.** *Let $\delta \leqslant 0.025$ and $\epsilon > 0$. Then, Algorithm 13 has the property that there exists a stopping time τ wrt $(\mathcal{F})_{t\in\mathbb{N}}$ such that*

$$\mathbb{E}[\tau] \lesssim \min_{\gamma\in(0,\epsilon)} \widetilde{U}_\epsilon(\gamma) \log(\widetilde{U}_\epsilon(\gamma) + \Delta_{m,\epsilon,\gamma}^{-2}) \tag{4.1}$$
$$\lesssim \widetilde{U}_\epsilon \log(\widetilde{U}_\epsilon + (\mu_m - \mu_{m+1})^{-2}) \tag{4.2}$$

and $\mathbb{P}(\exists s \geqslant \tau : \mu_{O_s} \leqslant \mu_1 - \epsilon) \leqslant 2\delta$.

Remark 2. *Assume the setting of Theorem 17. If m arms have means equal to $\mu_0 + \epsilon$ while the remaining have means equal to μ_0 then defining $\bar{U} := \frac{n}{m}\epsilon^{-2}\log(1/\delta)$, we have that $\mathbb{E}[\tau] \lesssim \bar{U}\log(\bar{U})$. This intuitively agrees with the sample complexity arguments made in the introduction of this paper and matches the lower bound of above up to* log *factors.*

To interpret Theorem 17, we begin by considering the more digestible bound of equation (4.2). The term $\widetilde{U}_\epsilon$ bounds the expected number of rounds it takes for a bracket of size $\Theta(\frac{n}{m})$ to identify an ϵ-good arm when one of these arms is ϵ-good. The extra $\log(\widetilde{U}_\epsilon)$ factor reflects the cost of not knowing m in advance and having to adapt to it.

In many situations the bound of Equation 4.2 is woefully loose because while it suffices to sample just $\Theta(\frac{n}{m})$ arms to get an ϵ-good arm with constant probability, it may be advantageous to sample many more than this in hopes of getting an arm with a mean much closer to μ_1 than $\mu_1 - \epsilon$. Figure 4.1 illustrates a bandit instance that demonstrates this tradeoff for a particular $\gamma \in (0, \epsilon)$. Informally, if one randomly chooses $\frac{n}{m}$ arms then one expects the highest mean amongst these to be have an index in $\{|G_\gamma| + 1, \ldots, m\}$ whose mean is very close to the bottom $n-m$ arms requiring an enormous number of samples to distinguish this top arm from the bottom $n-m$. On the other hand, if one randomly chooses $\frac{n}{|G_\gamma|}$ arms then one expects the highest mean amongst these to have an index in $\{1, \ldots, |G_\gamma|\}$ whose mean is substantially larger than the means of the bottom $n-m$ arms requiring far few samples to distinguish it from the bottom $n-m$. Thus, there is a problem-dependent tradeoff between how many arms to sample and an effective strategy must naturally adapt to it.

The minimization problem in the bound in Equation 4.1 says that Algorithm 13 uses a bracket with the optimal number of arms to identify an ϵ-good arm. The case $\gamma = \epsilon$ captures the case where the algorithm considers $\Theta(\frac{n}{m})$ arms and identifies one of these as ϵ-good. As γ decreases to 0, the algorithm considers more arms but the gaps of the arms increase, reflecting that it is likely to find a better arm to identify as ϵ-good.

One potential drawback of our framework is that a practitioner might feel uncomfortable not knowing when to stop the algorithm. We have several responses to this concern. First, the lower confidence bound of the suggested arm O_t provides a lower bound for μ_{O_t} and thus provides an indication of its quality, so the practitioner may decide to stop once this lower bound is suitably large. Second, it is possible to design an algorithm that nearly optimal according to both the PAC and the SimplePAC criteria. Consider the following Theorem on Algorithm 15 described in the Appendix, which combines Algorithm 13 and LUCB from Kalyanakrishnan et al. (2012a).

Theorem 18. *Let ρ be a problem instance and let $\delta \leqslant 0.025$ and $\epsilon_1, \epsilon_2 > 0$. Let $(\mathcal{F}_t)_{t\in\mathbb{N}}$ be the filtration generated by running Algorithm 15 with input ϵ_1 on ρ. There is a stopping time τ_{simple} wrt $(\mathcal{F}_t)_{t\in\mathbb{N}}$ such that*

$$\mathbb{E}[\tau_{simple}] \lesssim \min_{\gamma\in(0,\epsilon_2)} U_{\epsilon_2}(\gamma) \log(U_{\epsilon_2}(\gamma) + \Delta^{-2}_{m,\epsilon_2,\gamma}) \tag{4.3}$$

and $\mathbb{P}(\exists s \geqslant \tau_{simple} : \mu_{\widehat{i}_s} \leqslant \mu_1 - \epsilon_2) \leqslant 2\delta$. *Furthermore, there exists a stopping time* τ_{PAC} *wrt* $(\mathcal{F}_t)_{t\in\mathbb{N}}$ *such that*

$$\mathbb{E}[\tau_{PAC}] \lesssim H^{\epsilon/2} \log(\frac{H^{\epsilon/2}}{\delta}) \tag{4.4}$$

where $H^{\gamma} = \sum_{i\in[n]} \max(\mu_1 - \mu_i, \gamma)^{-2}$ *and at time* τ_{PAC} *the Algorithm 15 terminates and returns an arm* $\widehat{i}_{\tau_{PAC}}$ *such that* $\mathbb{P}(\mu_{\widehat{i}_{\tau_{PAC}}} \leqslant \mu_1 - \min(\epsilon_1, \epsilon_2)) \leqslant 3\delta$.

To interpret the Theorem 18, suppose that $\epsilon_1 > \epsilon_2 > 0$ are such that $\mathbb{E}[\tau_{simple}] \leqslant \mathbb{E}[\tau_{simple}]$. Then, Theorem 18 says that Algorithm 15 with input ϵ_1 starts outputting an ϵ_2-good arm in nearly optimal time *and* certifies that it is an ϵ_1-good arm in nearly optimal optimal. Thus, Algorithm 15 achieves the best of both worlds.

Finally, there is another interpretation of Algorithm 13 when it is applied in a fixed horizon setting. Consider the following problem that lies between the fixed confidence setting and fixed budget setting: given a failure probability $\delta > 0$, time horizon $T \in \mathbb{N}$, and permissible approximation $\epsilon > 0$, find an arm $\widehat{i}$ that satisfies $\mu_{\widehat{i}} \geqslant \mu_1 - \epsilon$ with probability at least $1-\delta$ using at most T samples. Algorithm 13 essentially solves this problem, as demonstrated by the following Theorem.

Theorem 19. *Let* $\delta \leqslant 0.025$ *and* $T \in \mathbb{N}$. *Define*

$$UB(\epsilon') = \min_{\gamma\in(0,\epsilon')} \frac{1}{\delta}\widetilde{U}_{\epsilon'}(\gamma) \log(\widetilde{U}_{\epsilon'}(\gamma) + \Delta^{-2}_{m,\epsilon',\gamma}),$$

$$\epsilon = \min(\epsilon' : \epsilon' > 0 : UB(\epsilon') \leqslant T).$$

Then, $\mathbb{P}(\mu_{O_T} \leqslant \mu_1 - \epsilon) \leqslant 3\delta$.

Thus, the practitioner can interpret Algorithm 13 as an anytime algorithm that minimizes simple regret over a time horizon T subject to a probability of error constraint δ. It is worth noting that an algorithm that optimizes the (ϵ, δ)-PAC does not admit such an interpretation since it minimizes *certified* simple regret subject to a probability of error constraint δ.

4.3.2 Upper Bound for Identifying means above a threshold μ_0

Next we present our upper bounds on multiple identifications above a threshold with a FDR-TPR guarantee, deferring our stronger guarantees (for our less practical algorithms) until

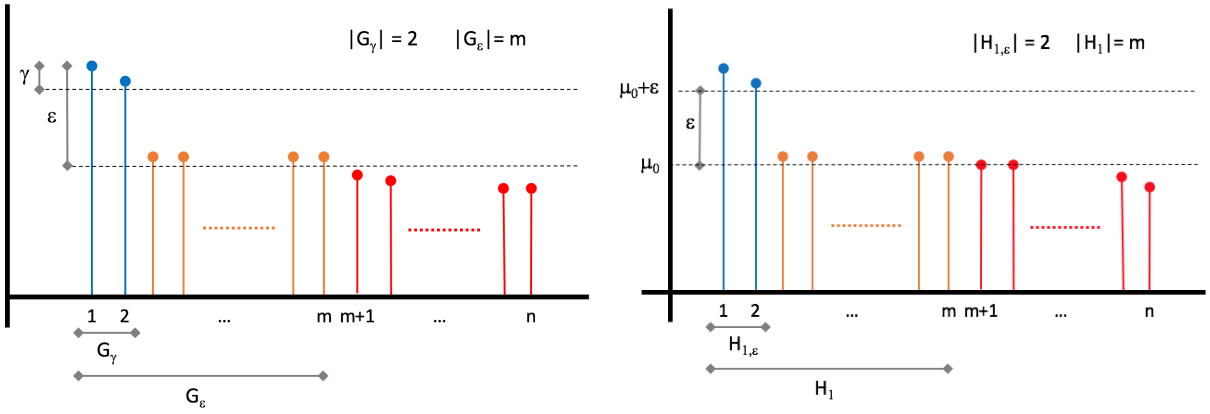


Figure 4.1: Our sample complexity result relies on a trade-off between the number of arms above a certain critical value and the defined gaps as a function of this critical value. (Left) Illustration of G_γ sets that define gaps for identifying an ϵ-good arm. (Right) Illustration of $\mathcal{H}_{1,\epsilon}$ sets that define gaps for identifying means above a threshold μ_0.

the Appendix. To begin, we introduce some notation. Define

$$\mathcal{H}_1 = \{i \in [n] : \mu_i > \mu_0\} \quad \text{and} \quad \mathcal{H}_0 = \{i \in [n] : \mu_i \leqslant \mu_0\}.$$

$\mathcal{H}_1$ consists of the arms that we wish to identify and $\mathcal{H}_0$ of all the other arms. Finally, we define

$$\mathcal{H}_{1,\epsilon} = \{i \in [n] : \mu_i \geqslant \mu_0 + \epsilon\}, \quad \Delta_i = \Delta_{i,0}, \quad \text{and} \quad \Delta_{i,\epsilon} = \begin{cases} \max(\mu_i - \mu_0, \epsilon) & i \in \mathcal{H}_1 \\ \mu_0 - \mu_i + \max(\epsilon, \Delta) & i \in \mathcal{H}_0 \end{cases}$$

where $\Delta = \min_{i \in \mathcal{H}_1} \mu_i - \mu_0$ is the gap of the arm in $\mathcal{H}_1$ that is closest to μ_0. The set $\mathcal{H}_{1,\epsilon}$ (illustrated in Figure 4.1) and the gaps $\Delta_{i,\epsilon}$ will respectively play an analogous role to the set G_γ and gaps $\Delta_{i,\epsilon,\gamma}$ in our results for finding ϵ-good arms.

Next, we introduce two functions to characterize the number of samples required for an individual bracket to identify k arms in $\mathcal{H}_1$. For all $\epsilon > 0$ define

$$\widetilde{S}_k(\epsilon) := \frac{k}{|\mathcal{H}_{1,\epsilon}|}\Big(\sum_{i \in \mathcal{H}_1} \Delta_{i,\epsilon}^{-2} \log(\tfrac{n}{|\mathcal{H}_{1,\epsilon}|}k/\delta) + \sum_{i \in \mathcal{H}_0} \Delta_{i,\epsilon}^{-2} \log(1/\delta)\Big)$$

$$\widetilde{S}_k := \widetilde{S}_k(\Delta) = \frac{k}{|\mathcal{H}_1|}\Big(\sum_{i \in \mathcal{H}_1} \Delta_i^{-2} \log(\tfrac{n}{|\mathcal{H}_1|}k/\delta) + \sum_{i \in \mathcal{H}_0} \Delta_i^{-2} \log(1/\delta)\Big)$$

$$\widetilde{T}_k(\epsilon) := \frac{n}{|\mathcal{H}_{1,\epsilon}|} k \max(\epsilon, \Delta)^{-2} \log(1/\delta), \quad \widetilde{T}_k := T_k(\Delta) = \frac{n}{|\mathcal{H}_1|} k \Delta^{-2} \log(1/\delta).$$

$\widetilde{S}_k(\epsilon)$ bounds the expected number of samples a bracket of size $\Theta(\frac{nk}{|\mathcal{H}_{1,\epsilon}|})$ requires to identify k arms satisfying $\mu_i > \mu_0$ when *(i)* at least k of its arms have means satisfying $\mu_i \geqslant \mu_0 + \epsilon$ and *(ii)* the empirical means of the arms in the bracket concentrate well. $\widetilde{T}_k(\epsilon)$ plays a similar role but removes a logarithmic factor on the arms in $\mathcal{H}_1$ at the cost of losing the dependence on the individual gaps.

Now, we have the tools to present our Theorem.

Theorem 20 (FDR-TPR)**.** *Let $\delta \leqslant .025$. Let $k \leqslant |\mathcal{H}_1|$. Then, Algorithm 13 has the property that for all $t \in \mathbb{N}$, $\mathbb{E}[\frac{|\mathcal{S}_t \cap \mathcal{H}_0|}{|\mathcal{S}_t| \wedge 1}] \leqslant 2\delta$ and there exists a stopping time τ_k wrt $(\mathcal{F})_{t\in\mathbb{N}}$ such that*

$$\mathbb{E}[\tau_k] \lesssim \min_{\epsilon \geqslant \Delta : |\mathcal{H}_{1,\epsilon}| \geqslant k} \widetilde{S}_k(\epsilon) \log[\widetilde{S}_k(\epsilon) + \epsilon^{-2})] \tag{4.5}$$

$$\lesssim \widetilde{S}_k \log[\widetilde{S}_k + \Delta^{-2})], \ \textit{and} \tag{4.6}$$

$$\mathbb{E}[\tau_k] \lesssim \min_{\epsilon \geqslant \Delta : |\mathcal{H}_{1,\epsilon}| \geqslant k} \widetilde{T}_k(\epsilon) \log(\widetilde{T}_k(\epsilon)) \tag{4.7}$$

$$\lesssim \widetilde{T}_k \log(\widetilde{T}_k) \tag{4.8}$$

and for all $t \geqslant \tau_k$, $\mathbb{E}[|\mathcal{S}_t \cap \mathcal{H}_1|] \geqslant (1-\delta)k$.

Consider the gap-dependent bounds in the inequalities (4.5) and (4.6). Analogous to the ϵ-good arm identification upper bound, the bound in inequality (4.6) is more accessible (but potentially much looser) than the bound in line (4.5). The quantity $\widetilde{S}_k$ upper bounds the expected number of samples required for a bracket of size $\Theta(\frac{n}{m}k)$ to identify k arms with means $\mu_i > \mu_0$. The extra $\log(\widetilde{S}_k)$ reflects the cost of not knowing $|\mathcal{H}_1|$ and adapting to it. Paralleling ϵ-good arm identification, in the problem of identifying k arms with means above a threshold it may be useful to consider more than $\Theta(\frac{n}{m}k)$ arms in order to find k that are easier to identify as being larger than the threshold. Indeed, the example in and discussion concerning Figure 4.1 apply directly to this problem as well. The minimization problems in the bounds in line (4.5) shows that the algorithm adapts to a problem instance, finding the optimal number of arms to consider. As ϵ increases, more arms are considered but the gaps of the arms grow. For example, for arms with $i \in \mathcal{H}_1$, the gaps are $\max(\mu_i - \mu_0, \epsilon)$ while for arms with $i \in \mathcal{H}_0$, the gaps are $\mu_0 + \max(\epsilon, \Delta) - \mu_i$.

The inequalities (4.7) and (4.8) sacrifice the dependence on the individual gaps in order to eliminate an additional logarithmic factor on the arms in $\mathcal{H}_1$. Similarly to the discussion

concerning the inequalities (4.5) and (4.6), the inequality (4.7) is far tighter than (4.8) but less digestible.

4.3.3 A note about our proof techniques

While our proofs are relegated to the appendices due to space restrictions, we briefly comment on their novel aspects. First, our upper bounds scale as $\log(1/\delta)$ which arises due to requiring concentration of measure on subsets of the observations of the arms. We do *not* rely on a high probability event that a particular bracket includes some number of good arms which would result in a $\log^2(1/\delta)$ that is common in other related results Chaudhuri and Kalyanakrishnan (2017); Aziz et al. (2018); Chaudhuri and Kalyanakrishnan (2019); this unfortunately results in a remarkably more technical and challenging proof. Second, our lower bounds for general $1 \leqslant k \leqslant m$ employed the use of a slightly extended version of the so-called *Simulator* argument that is a general framework for developing lower bounds for adaptive sampling problems Simchowitz et al. (2017b). This is the first instance to our knowledge that the argument has been used to lower bound the number of samples for such combinatorial settings where many outcomes are potentially correct (i.e., any k of m) and shows great promise for proving tight lower bounds for other pure-exploration combinatorial bandit settings Cao et al. (2015); Chen et al. (2014b).

4.4 Proof of lower bounds

We now briefly provide some intuition behind the proof. Suppose $m > 1$ and $k = 1$ and consider the easier problem where the permutation set averaged over is just the identity permutation $\pi_1 = (1, 2, \ldots, n)$ and the permutation π_2 that swaps $\{1, \ldots, m\}$ and some fixed $\sigma \subset [n] \backslash [m]$ with $|\sigma| = m$. That is, the algorithm knows the instance it is playing is either $\pi_1(\rho) = \rho$ or $\pi_2(\rho)$ where ρ is known but the permutation π_1 or π_2 is not. Information theoretic arguments say that at least $\tau \approx \min_{i \in \sigma}(\mu_1 - \mu_i)^{-2}$ observations from $[m] \cup \sigma$ are necessary in order to determine whether the underlying instance is $\pi_1(\rho)$ versus $\pi_2(\rho)$. But if the algorithm cannot distinguish between π_1 and π_2 with fewer than τ samples, then we can also argue that if π_1 and π_2 are chosen with equal probability, then taking nearly τ samples

from the arms in σ with sub-optimal means is unavoidable in expectation. The choice of σ was arbitrary and there are $\frac{n}{m}-1$ disjoint choices (e.g., $\{m+1,\dots,2m\},\{2m+1,\dots,3m\},\dots$) resulting in a lower bound of about $\frac{1}{m}\sum_{i=m+1}^{n}(\mu_1-\mu_i)^{-2}$.

The $k>1$ case is trickier because if we used just π_1 and π_2 as above, as soon as we found just one ϵ-good arm (and thus being able to accurately discern whether the instance is $\pi_1(\rho)$ or $\pi_2(\rho)$) the algorithm would immediately know of $m-1$ other ϵ-good arms. To overcome this, we choose a large enough set $\sigma\subset[m]$ such that $\sigma\cap\widehat{S}$ is non-empty with constant probability on the identity permutation. This way, if we swap this set $\sigma\subset[m]$ with some other set in $[n]\backslash[m]$ of size $|\sigma|$, then the algorithm would error with constant probability on this alternative permutation. The next lemma guarantees the existence of such a set of size $\lceil m/k\rceil$ and the final result follows from the fact that there are about $\frac{n}{\lceil m/k\rceil}$ such disjoint choices in $[n]\backslash[m]$.

We introduce the following notation: for any $j\leqslant m$ let $\binom{[m]}{j}$ denote all subsets of $\{1,\dots,m\}$ of size j.

Lemma 22. *Fix $m\in\mathbb{N}$ and let S be a random subset of size $k\leqslant m$ drawn from an arbitrary distribution over $\binom{[m]}{k}$. For any $\ell\leqslant m-k$ there exists a subset $\sigma\subset[m]$ with $|\sigma|=\ell$ such that*

$$\mathbb{P}(\sigma\cap S\neq\varnothing)\geqslant 1-\binom{m-k}{\ell}\Big/\binom{m}{\ell}\geqslant 1-e^{-\ell k/m}$$

If $\ell>m-k$ then $\mathbb{P}(\sigma\cap S\neq\varnothing)=1$.

Proof. Because the max of a set of positive numbers is always at least the average, we have

$$\begin{aligned}
\max_{\sigma\in\binom{[m]}{\ell}}\mathbb{P}(\sigma\cap S\neq\varnothing)&\geqslant\frac{1}{\binom{m}{\ell}}\sum_{\sigma\in\binom{[m]}{\ell}}\mathbb{P}(\sigma\cap S\neq\varnothing)\\
&=\frac{1}{\binom{m}{\ell}}\sum_{\sigma\in\binom{[m]}{\ell}}\sum_{s\in\binom{[m]}{k}}\mathbb{P}(S=s)\mathbf{1}\{\sigma\cap s\neq\varnothing)\\
&=\frac{1}{\binom{m}{\ell}}\sum_{s\in\binom{[m]}{k}}\mathbb{P}(S=s)\sum_{\sigma\in\binom{[m]}{\ell}}\mathbf{1}\{\sigma\cap s\neq\varnothing)\\
&=\frac{1}{\binom{m}{\ell}}\sum_{s\in\binom{[m]}{k}}\mathbb{P}(S=s)\left(\binom{m}{\ell}-\binom{m-k}{\ell}\right)\\
&=1-\binom{m-k}{\ell}\Big/\binom{m}{\ell}
\end{aligned}$$

where the last line follows from the fact that $\sum_{s\in\binom{[m]}{k}} \mathbb{P}(S=s)=1$ because it is a probability distribution. Now

$$\begin{aligned}\binom{m-k}{\ell}\Big/\binom{m}{\ell} &= \frac{(m-k)!\,(m-\ell)!}{(m-k-\ell)!\,m!} \\ &= \prod_{i=0}^{k-1}\frac{m-i-\ell}{m-i} = \prod_{i=0}^{k-1}\left(1-\frac{\ell}{m-i}\right) \leqslant \prod_{i=0}^{k-1}\left(1-\frac{\ell}{m}\right) \leqslant e^{-\ell k/m}.\end{aligned}$$

□

Fix any $\sigma \subset [m]$ with $|\sigma| = \lceil m/k \rceil$ that satisfies $\mathbb{P}_\rho\left(\widehat{S}\cap\sigma\neq\varnothing\right) \geqslant 1-e^{-1}$ (which must exist by the above lemma). Now fix any $\sigma' \subset [n]\backslash[m]$ with $|\sigma'| = |\sigma|$ and define ρ' as swapping the arms of σ and σ', maintaining their relative ordering of the indices within the sets. Note that by the correctness assumption at the relative stopping times of ρ and ρ' we have

$$\mathbb{P}_\rho(\widehat{S}\subset[m]) \geqslant 1-\delta, \qquad \mathbb{P}_{\rho'}(\widehat{S}\cap\sigma\neq\varnothing) \leqslant \delta, \qquad \mathbb{P}_\rho(\widehat{S}\cap\sigma\neq\varnothing) \geqslant 1-e^{-1}$$

which implies

$$\mathrm{TV}(\mathbb{P}_\rho,\mathbb{P}_{\rho'}) = \sup_{\mathcal{E}}|\mathbb{P}_\rho(\mathcal{E})-\mathbb{P}_{\rho'}(\mathcal{E})| \geqslant |\mathbb{P}_\rho(\widehat{S}\cap\sigma\neq\varnothing)-\mathbb{P}_{\rho'}(\widehat{S}\cap\sigma\neq\varnothing)| \geqslant 1-\delta-e^{-1}. \tag{4.9}$$

Remark 3. *Given* (4.9)*, one is tempted to apply Pinsker's inequality to obtain the right-hand-side of Lemma 1 from Kaufmann et al. (2016a) and then provide a lower bound on* $\mathbb{E}_\rho[\sum_{i\in\sigma\cup\sigma'} T_i]$*. The difficulty here is that once we cover* $[n]\backslash[m]$ *with alternative* σ' *sets, they would all share the same* σ *in this lower bound, which will end up being* $o(n)$*. Alternatively, one could consider using the technique of Chen et al. (2017b) which compares a given instance to a degenerate instance where the means of* σ' *would be copied to* σ *and argue that the probability of error is at least* $1/2$ *since there truly is no difference. This strategy is successful if* $k=1$ *so that* $|\sigma|=m$ *but breaks down when* $k>1$ *because one cannot reason about what the algorithm would have to do if the means of* σ *were changed like one could if* $k=1$*. Consequently, we employ the use of the Simulator argument from Simchowitz et al. (2017b) that is much more powerful at the cost of the introduction of some machinery.*

The Simulator (background)

The simulator argument is a kind of thought experiment where the player is playing against a non-stationary distribution. In the real game when the player pulls arm $I_t = i$ arm at time t she observes a sample from the ith distribution of instance ρ: $X_{i,t} \sim \rho_i$. However, when playing against the simulator she observes a sample form the ith distribution of an instance denoted $\text{Sim}(\rho, \{I_1, \ldots, I_t\})$ that depends on all past requests: $X_{i,t} \sim \text{Sim}(\rho, \{I_1, \ldots, I_t\})_i$ with probability law Q given ρ, $\{I_s = i_s\}_{s=1}^t$. That is, instead of receiving rewards from a stationary distribution ρ at each time t, the simulator is an instance that depends on all the indices of past pulls (but not their values). For any set $A \subset \mathbb{R}$ define

$$\mathbb{P}_{\text{Sim}(\rho,(i_1,\ldots,i_t))}\left(X_{i_t} \in A\right) \coloneqq Q\left(X_{i_t} \in A | \rho, \{I_s = i_s\}_{s=1}^t\right).$$

We allow the algorithm to have internal randomness with probability law P so that for $B \subset [n]$ define

$$\mathbb{P}_{\text{Alg}((i_1,x_1,\ldots,i_{t-1},x_{t-1}))}\left(I_t \in B\right) \coloneqq P\left(I_t \in B | \{I_s = i_s, X_{I_s} = x_s\}_{s=1}^{t-1}\right)$$

so that for any event $E \in \mathcal{F}_T$ we define

$$\begin{aligned}
&\mathbb{P}_{\text{Alg},\text{Sim}(\rho)}(E) \\
&\coloneqq \sum_{i_1,\ldots,i_T} \int_{x_1,\ldots,x_T} \mathbf{1}_E \prod_{t=1}^T Q\left(X_{I_t} = x_t | \rho, \{I_s = i_s\}_{s=1}^t\right) P\left(I_t = i_t | \{I_s = i_s, X_{I_s} = x_s\}_{s=1}^{t-1}\right) dx_1 \ldots dx_T \\
&= \sum_{i_1,\ldots,i_T} \int_{x_1,\ldots,x_T} \mathbf{1}_E \prod_{t=1}^T \mathbb{P}_{\text{Sim}(\rho,(i_1,\ldots,i_t))}\left(X_{I_t} = x_t\right) \mathbb{P}_{\text{Alg}((i_1,x_1,\ldots,i_{t-1},x_{t-1}))}\left(I_t = i_t\right) dx_1 \ldots dx_T
\end{aligned}$$

so that for any T we have $KL\left(\mathbb{P}_{\text{Alg,Sim}(\rho)}, \mathbb{P}_{\text{Alg,Sim}(\rho')}\right) =$

$$\sum_{i_1,\ldots,i_T} \int_{x_1,\ldots,x_T} \mathbb{P}_{\text{Alg,Sim}(\rho)}(\{I_s = i_s, X_{I_s} = x_s\}_{s=1}^T) \log\left(\frac{\mathbb{P}_{\text{Alg,Sim}(\rho)}(\{I_s = i_s, X_{I_s} = x_s\}_{s=1}^T)}{\mathbb{P}_{\text{Alg,Sim}(\rho')}(\{I_s = i_s, X_{I_s} = x_s\}_{s=1}^T)}\right) dx_1 \ldots dx_T$$

$$= \sum_{i_1,\ldots,i_T} \int_{x_1,\ldots,x_T} \mathbb{P}_{\text{Alg,Sim}(\rho)}(\{I_s = i_s, X_{I_s} = x_s\}_{s=1}^T) \log\left(\frac{\prod_{t=1}^T \mathbb{P}_{\text{Sim}(\rho,(i_1,\ldots,i_t))}(X_{I_t} = x_t)}{\prod_{t=1}^T \mathbb{P}_{\text{Sim}(\rho',(i_1,\ldots,i_t))}(X_{I_t} = x_t)}\right) dx_1 \ldots dx_T$$

$$= \sum_{t=1}^T \sum_{i_1,\ldots,i_T} \int_{x_1,\ldots,x_T} \mathbb{P}_{\text{Alg,Sim}(\rho)}(\{I_s = i_s, X_{I_s} = x_s\}_{s=1}^T) \log\left(\frac{\mathbb{P}_{\text{Sim}(\rho,(i_1,\ldots,i_t))}(X_{I_t} = x_t)}{\mathbb{P}_{\text{Sim}(\rho',(i_1,\ldots,i_t))}(X_{I_t} = x_t)}\right) dx_1 \ldots dx_T$$

$$= \sum_{t=1}^T \sum_{i_1,\ldots,i_T} \mathbb{P}_{\text{Alg,Sim}(\rho)}(\{I_s = i_s\}_{s=1}^T) \int_{x_t} \mathbb{P}_{\text{Sim}(\rho,(i_1,\ldots,i_t))}(X_{I_t} = x_t) \log\left(\frac{\mathbb{P}_{\text{Sim}(\rho,(i_1,\ldots,i_t))}(X_{I_t} = x_t)}{\mathbb{P}_{\text{Sim}(\rho',(i_1,\ldots,i_t))}(X_{I_t} = x_t)}\right) dx_t$$

$$= \sum_{t=1}^T \sum_{i_1,\ldots,i_T} \mathbb{P}_{\text{Alg,Sim}(\rho)}(\{I_s = i_s\}_{s=1}^T) KL\left(\mathbb{P}_{\text{Sim}(\rho,(i_1,\ldots,i_t))}, \mathbb{P}_{\text{Sim}(\rho',(i_1,\ldots,i_t))}\right)$$

$$= \sum_{i_1,\ldots,i_T} \mathbb{P}_{\text{Alg,Sim}(\rho)}(\{I_s = i_s\}_{s=1}^T) \sum_{t=1}^T KL\left(\mathbb{P}_{\text{Sim}(\rho,(i_1,\ldots,i_t))}, \mathbb{P}_{\text{Sim}(\rho',(i_1,\ldots,i_t))}\right)$$

$$\leqslant \max_{i_1,\ldots,i_T} \sum_{t=1}^T KL\left(\mathbb{P}_{\text{Sim}(\rho,(i_1,\ldots,i_t))}, \mathbb{P}_{\text{Sim}(\rho',(i_1,\ldots,i_t))}\right)$$

The simulator will be defined so that the right hand side is always finite for any T. When it is clear from context we will simply write $\mathbb{P}_\rho(E)$ or $\mathbb{P}_{\text{Sim}(\rho)}(E)$ to represent $\mathbb{P}_{\text{Alg},\rho}(E)$ or $\mathbb{P}_{\text{Alg,Sim}(\rho)}(E)$, respectively. Let $\Omega_t = \{I_1, \ldots, I_t\}$ denote the history of all arm pulls requested by the player up to time t. Note that Ω_t is a multi-set so that $|\Omega_t| = t$.

Definition 7. *We say an event W is* truthful *under a simulator* Sim *with respect to instance ρ if for all events $E \in \mathcal{F}_T$*

$$\mathbb{P}_\rho(E \cap W) = \mathbb{P}_{\text{Sim}(\rho,\Omega_T)}(E \cap W).$$

Lemma 23 (Simchowitz et al. (2017b))**.** *Let $\rho^{(1)}$ and $\rho^{(2)}$ be two instances,* $\text{Sim}(\cdot,\cdot)$ *be a simulator, and let W_i be two truthful $\mathcal{F}_T$-measureable events under* $\text{Sim}(\rho^{(i)}, \Omega_T)$ *for $i = 1, 2$ where Ω_T is the history of pulls up to a stopping time T. Then*

$$\mathbb{P}_{\rho^{(1)}}(W_1^c) + \mathbb{P}_{\rho^{(2)}}(W_2^c) \geqslant \text{TV}(\rho^{(1)}, \rho^{(2)}) - Q\left(KL\left(\mathbb{P}_{\text{Alg,Sim}(\rho^{(1)})}, \mathbb{P}_{\text{Alg,Sim}(\rho^{(2)})}\right)\right)$$

where $Q(\beta) = \min\{1 - \frac{1}{2}e^{-\beta}, \sqrt{\beta/2}\}$.

Constructing the Simulator

Recall the definitions of ρ, ρ' and σ, σ' from above. For some $\tau \in \mathbb{N}$ and multiset Ω of requested arm pulls, define $W_\sigma(\Omega) = \{\sum_{i\in\Omega} \mathbf{1}\{i \in \sigma\} \leqslant \tau\}$ and $W_{\sigma'}(\Omega) = \{\sum_{i\in\Omega} \mathbf{1}\{i \in \sigma'\} \leqslant \tau\}$. For these events, an instance $\nu \in \{\rho, \rho'\}$, and any multiset Ω_t denoting the indices the player has played up to the current time t, define a simulator

$$\mathrm{Sim}(\nu, \Omega_t)_i = \begin{cases} \nu_i & \text{if } i \notin \sigma \cup \sigma' \\ \nu_i & \text{if } i \in \sigma \cup \sigma',\ W_\sigma(\Omega_t) \cap W_{\sigma'}(\Omega_t) \\ \rho_i & \text{if } i \in \sigma,\ W^c_\sigma(\Omega_t) \cup W^c_{\sigma'}(\Omega_t) \\ \rho_{\sigma(\sigma'^{-1}(i))} & \text{if } i \in \sigma',\ W^c_\sigma(\Omega_t) \cup W^c_{\sigma'}(\Omega_t) \end{cases}$$

where $\sigma(i)$ denotes the ith element of σ and $\sigma^{-1}(i) \in \{1, \ldots, |\sigma|\}$ so that $\sigma(\sigma^{-1}(i)) = i$ for any $i \in \sigma$, and $\sigma(\sigma'^{-1}(i)) \in \sigma$ for any $i \in \sigma'$. Note that $\mathrm{Sim}(\nu, \Omega_t)_i$ and ν_i potentially differ only on those arms $i \in \sigma'$, and only if $W^c_\sigma(\Omega_t) \cup W^c_{\sigma'}(\Omega_t) := \max\{\sum_{j\in\Omega_t} \mathbf{1}\{j \in \sigma\}, \sum_{j\in\Omega_t} \mathbf{1}\{j \in \sigma'\}\} > \tau$. That is, if $\max\{\sum_{j\in\Omega_t} \mathbf{1}\{j \in \sigma\}, \sum_{j\in\Omega_t} \mathbf{1}\{j \in \sigma'\}\} > \tau$ then $\mathrm{Sim}(\rho, \Omega_t)_i = \mathrm{Sim}(\rho', \Omega_t)_i$ for all $i \in [n]$. On the other hand, if $W_\sigma(\Omega_t) \cap W_{\sigma'}(\Omega_t) := \max\{\sum_{j\in\Omega_t} \mathbf{1}\{j \in \sigma\}, \sum_{j\in\Omega_t} \mathbf{1}\{j \in \sigma'\}\} \leqslant \tau$ then $\mathrm{Sim}(\nu, \Omega_t)_i = \nu$ for $\nu \in \{\rho, \rho'\}$. Thus, $W_\sigma(\Omega_t)$ and $W_{\sigma'}(\Omega_t)$ are both truthful under $\mathrm{Sim}(\nu, \Omega_t)_i$ for $\nu \in \{\rho, \rho'\}$. Using these observations, we can easily upper bound the KL divergence:

$$\begin{aligned} \max_{i_1,\ldots,i_T\in[n]} \sum_{t=1}^{T} KL\left(\mathrm{Sim}(\rho, \{i_s\}_{s=1}^t), \mathrm{Sim}(\rho', \{i_s\}_{s=1}^t)\right) &\leqslant \max_{i\in\sigma} \tau KL(\rho_i, \rho'_i) + \max_{j\in\sigma'} \tau KL(\rho_j, \rho'_j) \\ &= \max_{i=1,\ldots,\ell} \tau(\mu_{\sigma(i)} - \mu_{\sigma'(i)})^2. \end{aligned}$$

As shown in (Simchowitz et al., 2017b, Lemma 1) averaging over all permutations is equivalent to constructing a symmeterized version of the algorithm such that given any bandit instance, the algorithm randomly permutes the arms internally and then after making its set selection, returns the set inverted by the randomly chosen permutation. This modified algorithm is symmetric in the sense that

$$\mathbb{P}_\rho((i_1, \ldots, i_T, s) = (I_1, \ldots, I_T, \widehat{S})) = \mathbb{P}_{\pi(\rho)}((i_1, \ldots, i_T, , s) = (\pi(I_1), \ldots, \pi(I_T), \pi(\widehat{S}))).$$

In what follows, we assume the algorithm is symmetric which, in particular, implies

$$\mathbb{P}_\rho(W^c_{\sigma'}) + \mathbb{P}_{\rho'}(W^c_\sigma) = 2\mathbb{P}_\rho(W^c_{\sigma'}).$$

Putting all the pieces together we have

$$\begin{aligned}\mathbb{P}_\rho\left(\sum_{i\in\sigma'} T_i > \tau\right) = \mathbb{P}_\rho(W^c_{\sigma'}) &= \frac{1}{2}\left(\mathbb{P}_\rho(W^c_{\sigma'}) + \mathbb{P}_{\rho'}(W^c_\sigma)\right) \\ &\geqslant \frac{1}{2}\left(1 - \delta - e^{-1} - \sqrt{\tau \max_{i=1,\dots,\ell}(\mu_{\sigma(i)} - \mu_{\sigma'(i)})^2/2}\right) \\ &> \frac{1}{8} - \delta\end{aligned}$$

if $\tau = \frac{1}{2\max_{i=1,\dots,\ell}(\mu_{\sigma(i)} - \mu_{\sigma'(i)})^2}$. By Markov's inequality, $\mathbb{E}_\rho[\sum_{i\in\sigma'} T_i] \geqslant \tau\mathbb{P}_\rho\left(\sum_{i\in\sigma'} T_i > \tau\right)$. Noting that $\sigma' \subset [n]\backslash[m]$ was arbitrary, we apply the above calculation for all connected subsets of size $\lceil m/k \rceil$.

$$\begin{aligned}\mathbb{E}_\rho\left[\sum_{i=m+1}^{n} T_i\right] &\geqslant (1/8 - \delta) \sum_{r=1}^{(n-m)k/m} (\mu_1 - \mu_{m+rm/k})^{-2} \\ &\geqslant (1/8 - \delta)\frac{k}{m} \sum_{i=m+m/k+1}^{n} (\mu_1 - \mu_i)^{-2} \\ &\geqslant (1/8 - \delta)\left[-(\mu_1 - \mu_{m+1})^{-2} + \frac{k}{m}\sum_{i=m+1}^{n} (\mu_1 - \mu_i)^{-2}\right]\end{aligned}$$

4.5 Additional Algorithms

In this section, we briefly introduce two additional algorithms that are very similar to the Algorithm 13 presented earlier but have stronger guarantees for the task of identifying means above a threshold. A FWER-TPR (family-wise error rate-true positive rate) guarantee outputs a set $\mathcal{Q}_t$ such that $\mathbb{P}(\exists t : \mathcal{Q}_t \cap \mathcal{H}_0 \neq \varnothing) \leqslant c\delta$ and $\mathbb{E}[|\mathcal{Q}_t \cap \mathcal{H}_1|] \geqslant (1-\delta)k$ for large enough t. In words, it does not allow mistakes, but it allows for only identification of the arms above the threshold only in expectation. A FWER-FWPD (family-wise error rate-family-wise probability of detection) guarantee is stronger since it requires that the outputted set $\mathbb{R}_t$ satisfies $\mathbb{P}(\exists t : \mathbb{R}_t \cap \mathcal{H}_0 \neq \varnothing) \leqslant c\delta$ and $|\mathbb{R}_t \cap \mathcal{H}_1| \geqslant k$ for large enough t. For more formal examples of these guarantees, see Theorems 22 and 24.

The algorithm suggests different sets depending on the objective. If FWER-TPR is desired, the algorithm maintains a set $\mathcal{Q}_t$ and adds arms whose lower confidence bounds are above the threshold μ_0 (Line 12). If FWER-FWPD is the goal, then an additional arm J_t is

Algorithm 14 Subsampling Algorithm: FWER-TPR and FWER-FWPD

1: $\delta_r = \frac{\delta}{r^2}$, $\delta'_r = \frac{\delta_r}{6.4\log(36/\delta_r)}$ $R_0 = 0$, $l = 0$, $\mathcal{S}_0 = \varnothing$, $\mathcal{Q}_0 = \varnothing$
2: **for** $t = 1, 2, \ldots$ **do**
3: **if** $t \geqslant 2^l l$ **then**
4: Let A_{l+1} be a random set of size $\min(2^{l+1}, n)$ in $[n]$
5: $l = l + 1$
6: **end if**
7: $R_t = [(R_{t-1} + 1) \mathrm{mod} l] + 1$
8: **if** there exists $i \in A_{R_t} \backslash \mathcal{S}_t$ such that $N_{i,R_t}(t) == 0$ **then**
9: Pull arm $I_t = \mathrm{argmin}_{i \in A_{R_t} \backslash \mathcal{S}_t} N_{i,R_t}(t)$
10: **else if** FWER-TPR **then**
11: Pull arm $I_t = \mathrm{argmax}_{i \in A_{R_t} \backslash \mathcal{Q}_t} \widehat{\mu}_{i,R_t,N_{i,R_t}(t)} + U(N_{i,R_t}(t), \delta)$
12: $\mathcal{Q}_{t+1} = \mathcal{Q}_t \cup \{i \in A_{R_t} : \widehat{\mu}_{i,R_t,N_{i,R_t}(t)} - U(N_{i,R_t}(t), \frac{\delta}{|A_{R_t}|R_t^2}) \geqslant \mu_0\}$ % FWER Thm.22
13: **else if** FWER-FWPD **then**
14: $\xi_{t,R_t} = \max\{2|\mathcal{S}_t \cap A_{R_t}|, \frac{5}{3(1-4\delta_{R_t})}\log(1/\delta_{R_t})R_t^2\}$
15: Pull arm $I_t = \mathrm{argmax}_{i \in A_{R_t} \backslash \mathcal{S}_t} \widehat{\mu}_{i,R_t,N_{i,R_t}(t)} + U(N_{i,R_t}(t), \frac{\delta}{\xi_{t,R_t}})$
16: $s(p) = \{i \in A_{R_t} : \widehat{\mu}_{i,R_t,N_{i,R_t}(t)} - U(N_{i,R_t}(t), \frac{p}{|A_{R_t}|}\delta'_{R_t} \geqslant \mu_0\}$
17: $\mathcal{S}_{t+1} = \mathcal{S}_t \cup s(\widehat{p})$ where $\widehat{p} = \max\{p \in [|A_{R_t}|] : |s(p)| \geqslant p\}$
18: **if** $\mathcal{S}_t \cap A_{R_t} \neq \varnothing$ **then**
19: $\nu_{t,R_t} = \max(|\mathcal{S}_t \cap A_{R_t}|, 1)$
20: Pull arm $J_t = \mathrm{argmax}_{i \in \mathcal{S}_t \cap A_{R_t} \backslash \mathcal{R}_t} \widehat{\mu}_{i,R_t,N_{i,R_t}(t)} + U(N_{i,R_t}(t), \frac{\delta_{R_t}}{\nu_{t,R_t}})$
21: $\chi_{t,R_t} = |A_{R_t}| - (1 - 2\delta'_{R_t}(1 + 4\delta'_{R_t}))|\mathcal{S}_t \cap A_{R_t}| + \frac{4(1+4\delta'_{R_t})}{3}\log(5\log_2(|A_{R_t}|/\delta'_{R_t})/\delta'_{R_t})$
22: $\mathcal{R}_{t+1} = \mathcal{R}_t \cup \{i \in \mathcal{S}_t \cap A_{R_t} : \widehat{\mu}_{i,R_t,N_{i,R_t}(t)} - U(N_{i,R_t}(t), \frac{\delta}{\chi_{t,R_t}}) \geqslant \mu_0\}$ % FWER Thm.24
23: **end if**
24: **end if**
25: **end for**

pulled each time based on an upper confidence bound criterion and arms are accepted into the set $\mathcal{R}_{t+1}$ (Line 22) if their lower confidence bound is above the threshold μ_0.

4.6 Proofs of Upper Bounds

The proofs for the FDR-TPR result (The proof of Theorem 21 in Section 4.6.1 should be read first. Then, one can read the proofs for any of the other results. We introduce some notation that we used throughout the proofs. We use c to denote a positive constant whose

value may change from line to line. Define

$$\rho_{i,r} = \sup\{\rho \in (0,1] : \cap_{t=1}^{\infty}\{|\widehat{\mu}_{i,r,t} - \mu_i| \leqslant U(t,\rho)\}\}.$$

We note that $\{\rho_{i,r}\}_{i\in[n],r\in\mathbb{N}}$ are independent and $\mathbb{P}(\rho_{i,r} \leqslant \delta) \leqslant \delta$ since by definition of $U(\cdot,\cdot)$ for any bracket $r \in \mathbb{N}$ and $\alpha \in (0,1)$, $\mathbb{P}(\cap_{t=1}^{\infty}\{|\widehat{\mu}_{i,r,t} - \mu_i| \leqslant U(t,\alpha)) \geqslant 1-\alpha$. We define

$$\mathcal{I}_r = \{i \in \mathcal{H}_1 \cap A_r : \rho_{i,r} \leqslant \delta\}.$$

to be those arms whose empirical means concetrate well in the sense that $\rho_{i,r} \leqslant \delta$. We also define $U^{-1}(\gamma,\delta) = \min(t : U(t,\delta) \leqslant \gamma)$. It can be shown for a sufficiently large constant c that $U^{-1}(\gamma,\delta) \leqslant c\gamma^{-2}\log(\log(\gamma^{-2})/\delta)$.

4.6.1 Proof of FDR-TPR

Recall the following definitions:

$$\mathcal{H}_{1,\epsilon} = \{i \in [n] : \mu_i \geqslant \mu_0 + \epsilon\}, \quad \text{and} \quad \Delta_{i,\epsilon} = \begin{cases} \max(\mu_i - \mu_0, \epsilon) & i \in \mathcal{H}_1 \\ \mu_0 - \mu_i + \max(\epsilon, \Delta) & i \in \mathcal{H}_0 \end{cases}$$

and $\Delta = \min_{i\in[n]} \mu_i - \mu_0$. We restate Theorem 20 with the doubly logarithmic terms. We only consider the gap-independent upper bound here (inequality (4.7)); in the following section, we will prove a stronger result, which will imply inequality (4.5).

Theorem 21. *Let $\delta \leqslant .025$. Let $k \leqslant |\mathcal{H}_1|$. For $\epsilon \geqslant \Delta$ define*

$$T_k(\epsilon) := \frac{n}{|\mathcal{H}_{1,\epsilon}|} k\epsilon^{-2} \log\left(\log(\tfrac{n}{|\mathcal{H}_{1,\epsilon}|}k)\log(\epsilon^{-2})/\delta\right)$$

Then, Algorithm 13 has the property that for all $t \in \mathbb{N}$, $\mathbb{E}[\frac{|\mathcal{S}_t \cap \mathcal{H}_0|}{|\mathcal{S}_t| \wedge 1}] \leqslant 2\delta$ and there exists a stopping time τ_k wrt $(\mathcal{F})_{t\in\mathbb{N}}$ such that

$$\mathbb{E}[\tau_k] \leqslant c \min_{\epsilon \geqslant \Delta : |\mathcal{H}_{1,\epsilon}| \geqslant k} T_k(\epsilon)\log(T_k(\epsilon)) \tag{4.10}$$

where c is a universal constant and for all $t \geqslant \tau_k$, $\mathbb{E}[|\mathcal{S}_t \cap \mathcal{H}_1|] \geqslant (1-\delta)k$.

We briefly sketch the proof. Let $\epsilon_0 \geqslant \Delta$ such that $|\mathcal{H}_{1,\epsilon_0}| \geqslant k$ minimize the upper bound (4.10). Then, there exists a bracket r_0 with size $\Theta(\frac{n}{|\mathcal{H}_{1,\epsilon_0}|}k)$ such that with constant

probability A_{r_0} has at least k arms in $\mathcal{H}_{1,\epsilon_0}$ and the empirical means concentrate well enough (defined formally in Lemma 25 as the event $E_{r_0} \coloneqq E_{\epsilon_0,r_0} \cap E_{0,r_0} \cap E_{1,r_0}$). The argument controls $\mathbb{E}[\tau_k]$ by partitioning the sample space according to which bracket $r_0 + s$ is the first such that the good event E_{r_0+s} occurs, i.e., according to $\{E_{r_0}, E_{r_0}^c \cap E_{r_0+1}, E_{r_0}^c \cap E_{r_0+1}^c \cap E_{r_0+2}, ...\}$. Lemma 25 shows that $\mathbb{E}[\mathbf{1}\{E_{r_0}\}\tau_{r_0,k}]$ has the same upper bound as (4.10) and that $\mathbb{E}[\mathbf{1}\{E_{r_0+s}\}\tau_{r_0+s,k}]$ has an upper bound that is larger than line (4.10) by a factor exponential in s. On the other hand, because the brackets are independent and growing in size, the probability of $E_{r_0+s} \cap (\cap_{r=0}^{s-1} E_{r_0+r}^c)$ decreases exponentially in s, enabling control of the exponential increase in $\mathbb{E}[\mathbf{1}\{E_{r_0+s}\}\tau_{r_0+s,k}]$ and, by extension, $\mathbb{E}[\tau_k]$.

Lemma 24 bounds the false discovery rate of Algorithm 13.

Lemma 24. *For all $t \in \mathbb{N}$, $\mathbb{E}[\frac{|\mathcal{S}_t \cap \mathcal{H}_0|}{|\mathcal{S}_t| \wedge 1}] \leqslant 2\delta$.*

Proof.

$$\begin{aligned}
\mathbb{E}[\frac{|\mathcal{S}_t \cap \mathcal{H}_0|}{|\mathcal{S}_t| \wedge 1}] &\leqslant \mathbb{E}[\frac{\sum_{l=1}^{\infty} |\mathcal{S}_t \cap A_l \cap \mathcal{H}_0|}{|\mathcal{S}_t| \wedge 1}] \\
&\leqslant \sum_{l=1}^{\infty} \mathbb{E}[\frac{|\mathcal{S}_t \cap A_l \cap \mathcal{H}_0|}{|\mathcal{S}_t \cap A_l| \wedge 1}] \\
&\leqslant \delta \sum_{l=1}^{\infty} \frac{1}{l^2} \\
&= \delta \frac{\pi^2}{6}
\end{aligned}$$

where we used Lemma 1 of Jamieson and Jain (2018).

□

Lemma 25, below, is the key result for establishing Theorem 21. For some $\epsilon_0 \geqslant \Delta$ such that $|\mathcal{H}_{1,\epsilon_0}| \geqslant k$, it bounds the expected number of iterations that it takes bracket r to add k arms to the set $\mathcal{S}_t$ when the events $E_{\epsilon_0,r} \cap E_{0,r} \cap E_{1,r}$ occur where

$$\begin{aligned}
E_{\epsilon_0,r} &= \{|\mathcal{H}_{1,\epsilon_0} \cap A_r| \geqslant k\}, \\
E_{0,r} &= \{\sum_{i \in \mathcal{H}_0 \cap A_r} \Delta_{i,\epsilon_0}^{-2} \log(\frac{1}{\rho_{i,r}}) \leqslant 5 \sum_{i \in \mathcal{H}_0 \cap A_r} \Delta_{i,\epsilon_0}^{-2} \log(\frac{1}{\delta})\}, \\
E_{1,r} &= \{\sum_{i \in \mathcal{H}_{1,\epsilon_0} \cap A_r} \Delta_{i,\epsilon_0}^{-2} \log(\frac{1}{\rho_{i,r}}) \leqslant 5 \sum_{i \in \mathcal{H}_{1,\epsilon_0} \cap A_r} \Delta_{i,\epsilon_0}^{-2} \log(\frac{1}{\delta})\}.
\end{aligned}$$

Event $E_{\epsilon_0,r}$ says that there are at least k arms in A_r with $\mu_i \geqslant \mu_0 + \epsilon_0$. The event $E_{0,r}$ says that the empirical means of the arms in $\mathcal{H}_0 \cap A_r$ concentrate well on the whole; event $E_{1,r}$ makes the analogous claim about $\mathcal{H}_{1,\epsilon_0} \cap A_r$. We remark that the the events $E_{0,r}$ and $E_{1,r}$ allow us to avoid using a union bound.

Lemma 25. *Fix $\delta < 0.025$. Fix $k \leqslant |\mathcal{H}_1|$ and any $\epsilon_0 > 0$ such that $|\mathcal{H}_{1,\epsilon_0}| \geqslant k$. Then, there exists a random variable τ_k such that for all $t \geqslant \tau_k$, $\mathbb{E}[|\mathcal{S}_t \cap \mathcal{H}_1|] \geqslant (1-\delta)k$, and*

$$\mathbb{E}[\mathbf{1}\{E_{\epsilon_0,r} \cap E_{0,r} \cap E_{1,r}\}\tau_k] \leqslant c[2^{r-1}(r-1) + |A_r|\epsilon_0^{-2}\log(r\frac{\log(\epsilon_0^{-2})}{\delta})\log(|A_r|\epsilon_0^{-2}\log(r\frac{\log(\epsilon_0^{-2})}{\delta}))] \tag{4.11}$$

where c is a universal constant.

Proof. **Step 1: Define stopping time.** Define

$$\tau_k = \min(t \in \mathbb{N} \cup \{\infty\} : \exists s \text{ such that } |A_s \cap \mathcal{H}_{1,\epsilon_0}| \geqslant k \text{ and } \mathcal{I}_s \cap A_s \cap \mathcal{H}_1 \subset \mathcal{S}_t).$$

Observe that for all $t \geqslant \tau_k$, $\mathbb{E}[|\mathcal{S}_t \cap \mathcal{H}_1|] \geqslant (1-\delta)k$ since for $t \geqslant \tau_k$

$$\mathbb{E}[|\mathcal{S}_t \cap \mathcal{H}_1|] \geqslant \mathbb{E}[|\mathcal{I}_s \cap A_s \cap \mathcal{H}_1|] \geqslant (1-\delta)|A_s \cap \mathcal{H}_1| \geqslant (1-\delta)k.$$

Step 2: Relate to bracket r.

Fix $r \in \mathbb{N}$. In the interest of brevity, define $E \coloneqq E_{\epsilon_0,r} \cap E_{0,r} \cap E_{1,r}$ and since we will only focus on bracket r, write $\widehat{\mu}_{i,t}$, $N_i(t)$, $\mathcal{I}$, and ρ_i instead of $\widehat{\mu}_{i,r,t}$, $N_{i,r}(t)$, $\mathcal{I}_r$, and $\rho_{i,r}$. We will bound the number of rounds until $\mathcal{I} \cap A_r \cap \mathcal{H}_1 \subset \mathcal{S}_t$. Define

$$T = |\{t \in \mathbb{N} : \mathcal{I} \cap A_r \cap \mathcal{H}_{1,\epsilon_0} \nsubseteq \mathcal{S}_t \text{ and } R_t = r\}|,$$

i.e., the number of rounds that the algorithm works on the rth bracket and $\mathcal{I} \cap A_r \cap \mathcal{H}_1 \nsubseteq \mathcal{S}_t$.

Next, we bound the number of brackets $r+s$ that are opened before $\mathcal{I} \cap A_r \cap \mathcal{H}_{1,\epsilon_0} \subset \mathcal{S}_t$. The $r+1$ bracket is opened after bracket r is sampled 2^r times and similarly the $r+s$th bracket is opened after bracket r is sampled $\sum_{i=0}^{s-1} 2^{r+i} \geqslant 2^{r+s-1}$ times. Thus,

$$2^{r+s-1} \leqslant T \implies r+s-1 \leqslant \log(T).$$

So while $\mathcal{I} \cap A_r \cap \mathcal{H}_1 \nsubseteq \mathcal{S}_t$, every time bracket r is sampled, at most $\log(T)$ total brackets are sampled. Thus, we have that once the algorithm starts working on bracket r, after

$$\log(T)T \tag{4.12}$$

additional rounds, we have that $\mathcal{I} \cap A_r \cap \mathcal{H}_{1,\epsilon_0} \subset \mathcal{S}_t$.

We note that after $2^{r-1}(r-1)$ rounds, the algorithm starts working on bracket r. Thus,

$$\begin{aligned}\mathbf{1}\{E\}\tau_k &\leqslant [2^{r-1}(r-1) + \mathbf{1}\{E\}\log(T)T] \\ &= [2^{r-1}(r-1) + \log(\mathbf{1}\{E\}T)\mathbf{1}\{E\}T] \end{aligned} \tag{4.13}$$

Step 3: Bounding $\mathbf{1}\{E\}T$. Note that we can write

$$\begin{aligned}\mathbf{1}\{E\}T &= \mathbf{1}\{E\}\sum_{t=1}^{\infty}\mathbf{1}\{\mathcal{H}_{1,\epsilon_0} \cap \mathcal{I} \cap A_r \not\subset \mathcal{S}_t, R_t = r\} \\ &= \mathbf{1}\{E\}\sum_{t:R_t=r}^{\infty}\mathbf{1}\{\mathcal{H}_{1,\epsilon_0} \cap \mathcal{I} \cap A_r \not\subset \mathcal{S}_t\} \\ &\leqslant \mathbf{1}\{E\}\sum_{t:R_t=r}^{\infty}\mathbf{1}\{\mathcal{H}_{1,\epsilon_0} \cap \mathcal{I} \cap A_r \not\subset \mathcal{S}_t, I_t \in \mathcal{H}_0\} \\ &\qquad + \mathbf{1}\{\mathcal{H}_{1,\epsilon_0} \cap \mathcal{I} \cap A_r \not\subset \mathcal{S}_t, I_t \in \mathcal{H}_1 \cap \mathcal{H}_{1,\epsilon_0}^c\} + \mathbf{1}\{I_t \in \mathcal{H}_{1,\epsilon_0}\} \\ &\leqslant \mathbf{1}\{E\}\sum_{t:R_t=r}^{\infty}\mathbf{1}\{\mathcal{H}_{1,\epsilon_0} \cap \mathcal{I} \cap A_r \not\subset \mathcal{S}_t, I_t \in \mathcal{H}_0\} \\ &\qquad + \mathbf{1}\{\mathcal{H}_{1,\epsilon_0} \cap \mathcal{I} \cap A_r \not\subset \mathcal{S}_t, I_t \in \mathcal{H}_1 \cap \mathcal{H}_{1,\epsilon_0}^c, \widehat{\mu}_{I_t,N_{I_t}(t)} < \mu_0 + \frac{\epsilon_0}{2}\} \\ &\qquad + \mathbf{1}\{I_t \in \mathcal{H}_1 \cap \mathcal{H}_{1,\epsilon_0}^c, \widehat{\mu}_{I_t,N_{I_t}(t)} \geqslant \mu_0 + \frac{\epsilon_0}{2}\} + \mathbf{1}\{I_t \in \mathcal{H}_{1,\epsilon_0}\}\end{aligned}$$

To begin, we bound the first sum.

For any $j \in \mathcal{I} \cap \mathcal{H}_{1,\epsilon_0} \cap A_r$ we have $\rho_j \geqslant \delta$ by definition, so

$$\widehat{\mu}_{j,N_i(t)} + U(N_j(t),\delta) \geqslant \mu_j - U(N_j(t),\rho_j) + U(N_j(t),\delta) \geqslant \mu_j \geqslant \mu_0 + \epsilon_0.$$

For any $i \in \mathcal{H}_0 \cap A_r$,

$$\widehat{\mu}_{i,N_i(t)} + U(N_i(t),\delta) \leqslant \mu_i + U(N_i(t),\rho_i) + U(N_i(t),\delta) \leqslant \mu_i + 2U(N_i(t),\rho_i\delta).$$

Thus, recalling the definition $\Delta = \min_{j\in\mathcal{H}_1}\mu_j - \mu_0$ and $\Delta_{i,\epsilon_0} \geqslant \mu_0 - \mu_i + \max\{\Delta,\epsilon_0\}$, we have that $\widehat{\mu}_{i,N_i(t)} + U(N_i(t),\delta) \leqslant \mu_0 + \epsilon_0$ if $N_i(t) \geqslant U^{-1}(\frac{\Delta_{i,\epsilon_0}}{2},\rho_i\delta)$, so that arm i would not be

pulled this many times as long as $\mathcal{H}_{1,\epsilon_0} \cap \mathcal{I} \cap A_r \not\subseteq \mathcal{S}_t$. Thus,

$$\begin{aligned}
&\mathbf{1}\{E\} \sum_{t:R_t=r}^{\infty} \mathbf{1}\{\mathcal{H}_{1,\epsilon_0} \cap \mathcal{I} \cap A_r \not\subseteq \mathcal{S}_t, I_t \in \mathcal{H}_0\} \qquad (4.14)\\
&\leqslant \mathbf{1}\{E\} \sum_{i \in \mathcal{H}_0 \cap A_r} U^{-1}(\frac{\Delta_{i,\epsilon_0}}{2}, \rho_i \delta)\\
&\leqslant \mathbf{1}\{E\} \sum_{i \in \mathcal{H}_0 \cap A_r} c\Delta_{i,\epsilon_0}^{-2} \log(\frac{\log(\Delta_{i,\epsilon_0}^{-2})}{\delta \rho_i})\\
&= \mathbf{1}\{E\} \sum_{i \in \mathcal{H}_0 \cap A_r} c\Delta_{i,\epsilon_0}^{-2} \log(\frac{\log(\Delta_{i,\epsilon_0}^{-2})}{\delta}) + c\Delta_{i,\epsilon_0}^{-2} \log(\frac{1}{\rho_i})\\
&\leqslant \mathbf{1}\{E\} \sum_{i \in \mathcal{H}_0 \cap A_r} c'\Delta_{i,\epsilon_0}^{-2} \log(\frac{\log(\Delta_{i,\epsilon_0}^{-2})}{\delta})\\
&\leqslant \sum_{i \in \mathcal{H}_0 \cap A_r} c'\Delta_{i,\epsilon_0}^{-2} \log(\frac{\log(\Delta_{i,\epsilon_0}^{-2})}{\delta}) \qquad (4.15)
\end{aligned}$$

where the second to last inequality follows from $E_{r,0} \subseteq E$.

Next, we consider the second sum. If $\mathcal{H}_{1,\epsilon_0} \cap \mathcal{I} \cap A_r \not\subseteq \mathcal{S}_t$, for any arm i satisfying $\widehat{\mu}_{i,N_i(t)} < \mu_0 + \frac{\epsilon_0}{2}$, we have that

$$\widehat{\mu}_{i,N_i(t)} + U(N_i(t), \delta) < \mu_0 + \frac{\epsilon_0}{2} + U(N_i(t), \delta)$$

so that if $N_i(t) \geqslant U^{-1}(\frac{\epsilon_0}{2}, \delta)$, then arm i is not pulled again until $\mathcal{H}_{1,\epsilon_0} \cap \mathcal{I} \cap A_r \subset \mathcal{S}_t$. Thus,

$$\begin{aligned}
&\sum_{t:R_t=r}^{\infty} \mathbf{1}\{\mathcal{H}_{1,\epsilon_0} \cap \mathcal{I} \cap A_r \not\subseteq \mathcal{S}_t, I_t \in \mathcal{H}_1 \cap \mathcal{H}_{1,\epsilon_0}^c \cap A_r, \widehat{\mu}_{I_t,N_{I_t}(t)} < \mu_0 + \frac{\epsilon_0}{2}\}\\
&\qquad \leqslant \sum_{i \in \mathcal{H}_1 \cap \mathcal{H}_{1,\epsilon_0}^c \cap A_r} U^{-1}(\frac{\epsilon_0}{2}, \delta)\\
&\qquad \leqslant c|\mathcal{H}_1 \cap \mathcal{H}_{1,\epsilon_0}^c \cap A_r| \epsilon_0^{-2} \log(\frac{\log(\epsilon_0^{-2})}{\delta}).
\end{aligned}$$

Next, we bound the final summands

$$\mathbf{1}\{E\} \sum_{t:R_t=r}^{\infty} \mathbf{1}\{I_t \in \mathcal{H}_1 \cap \mathcal{H}_{1,\epsilon_0}^c, \widehat{\mu}_{I_t,N_{I_t}(t)} \geqslant \mu_0 + \frac{\epsilon_0}{2}\} + \mathbf{1}\{I_t \in \mathcal{H}_{1,\epsilon_0}\}.$$

Let $p \leqslant |A_r|$. If $j \in \mathcal{H}_1 \cap \mathcal{H}_{1,\epsilon_0}^c \cap A_r$ and $\widehat{\mu}_{j,N_j(t)} \geqslant \mu_0 + \frac{\epsilon_0}{2}$, then

$$\widehat{\mu}_{j,N_j(t)} - U(N_j(t), \delta_r' \frac{p}{|A_r|}) \geqslant \mu_0 + \frac{\epsilon_0}{2} - U(N_j(t), \delta_r' \frac{p}{|A_r|})$$

so that $\widehat{\mu}_{j,N_j(t)} - U(N_j(t), \delta_r' \frac{p}{|A_r|}) \geqslant \mu_0$ if $N_i(t) \geqslant U^{-1}(\frac{\epsilon}{2}, \delta_r' \frac{p}{|A_r|})$, which implies that $j \in s(p)$.

Next, if $j \in \mathcal{H}_{1,\epsilon_0} \cap A_r$, then

$$\begin{aligned}\widehat{\mu}_{j,N_j(t)} - U(N_j(t), \delta_r' \frac{p}{|A_r|}) &\geqslant \mu_j - U(N_j(t), \rho_j) - U(N_j(t), \delta_r' \frac{p}{|A_r|}) \\ &\geqslant \mu_j - 2U(N_j(t), \rho_j \delta_r' \frac{p}{|A_r|})\end{aligned}$$

so that $\widehat{\mu}_{j,N_j(t)} - U(N_j(t), \delta_r' \frac{p}{|A_r|}) \geqslant \mu_0$ if $N_i(t) \geqslant U^{-1}(\frac{\mu_j - \mu_0}{2}, \rho_j \delta_r' \frac{p}{|A_r|})$, which implies that $j \in s(p)$.

While there is *some* p associated with each arm when it is added to $s(p)$ and then consequently to $\mathcal{S}_t$, we don't know the order in or time at which particular arms are added. However, in the worst case, the arms of $\mathcal{H}_1$ are added one at a time to $\mathcal{S}_t$ instead of in a big group so that the first reqires $p = 1$, the second $p = 2$, etc. Letting $\Gamma = \{f : f : \mathcal{H}_1 \longrightarrow$

$[|H_1|]$ is a bijection$\}$,

$$
\begin{aligned}
&\mathbf{1}\{E\} \sum_{t:R_t=r}^{\infty} \mathbf{1}\{I_t \in \mathcal{H}_1 \cap \mathcal{H}_{1,\epsilon_0}^c, \widehat{\mu}_{I_t,N_{I_t}(t)} \geqslant \mu_0 + \frac{\epsilon_0}{2}\} + \mathbf{1}\{I_t \in \mathcal{H}_{1,\epsilon_0}\} \\
&\leqslant \mathbf{1}\{E\} c \max_{\sigma\in\Gamma} \left(\sum_{j\in\mathcal{H}_1\cap\mathcal{H}_{1,\epsilon_0}^c\cap A_r} U^{-1}(\frac{\epsilon}{2}, \delta_r' \frac{\sigma(j)}{|A_r|}) + \sum_{j\in\mathcal{H}_{1,\epsilon_0}\cap A_r} U^{-1}(\frac{\mu_j-\mu_0}{2}, \rho_j\delta_r' \frac{\sigma(j)}{|A_r|}) \right) \\
&\leqslant \mathbf{1}\{E\} c \max_{\sigma\in\Gamma} \left(\sum_{j\in\mathcal{H}_1\cap\mathcal{H}_{1,\epsilon_0}^c\cap A_r} \Delta_{j,\epsilon_0}^{-2} \log(\frac{|A_r|}{\sigma(j)} \frac{\log(\Delta_{j,\epsilon_0}^{-2})}{\delta_r'}) \right. \qquad (4.16) \\
&\qquad \left. + \sum_{j\in\mathcal{H}_{1,\epsilon_0}\cap A_r} \Delta_{j,\epsilon_0}^{-2} \log(\frac{|A_r|}{\sigma(j)} \frac{\log(\Delta_{j,\epsilon_0}^{-2})}{\rho_j\delta_r'}) \right) \\
&= \mathbf{1}\{E\} c \max_{\sigma\in\Gamma} \left(\sum_{j\in\mathcal{H}_1\cap\mathcal{H}_{1,\epsilon_0}^c\cap A_r} \Delta_{j,\epsilon_0}^{-2} \log(\frac{|A_r|}{\sigma(j)} \frac{\log(\Delta_{j,\epsilon_0}^{-2})}{\delta_r'}) \right. \\
&\qquad \left. + \sum_{j\in\mathcal{H}_{1,\epsilon_0}\cap A_r} \Delta_{j,\epsilon_0}^{-2} \log(\frac{|A_r|}{\sigma(j)} \frac{\log(\Delta_{j,\epsilon_0}^{-2})}{\delta_r'}) + \sum_{j\in\mathcal{H}_{1,\epsilon_0}\cap A_r} \Delta_{j,\epsilon_0}^{-2} \log(\frac{1}{\rho_j}) \right) \\
&= \mathbf{1}\{E\} c \max_{\sigma\in\Gamma} \left(\sum_{j\in\mathcal{H}_1\cap\mathcal{H}_{1,\epsilon_0}^c\cap A_r} \Delta_{j,\epsilon_0}^{-2} \log(\frac{|A_r|}{\sigma(j)} \frac{\log(\Delta_{j,\epsilon_0}^{-2})}{\delta_r'}) \right. \\
&\qquad \left. + \sum_{j\in\mathcal{H}_{1,\epsilon_0}\cap A_r} \Delta_{j,\epsilon_0}^{-2} \log(\frac{|A_r|}{\sigma(j)} \frac{\log(\Delta_{j,\epsilon_0}^{-2})}{\delta_r'}) + 5 \sum_{j\in\mathcal{H}_{1,\epsilon_0}\cap A_r} \Delta_{j,\epsilon_0}^{-2} \log(\frac{1}{\delta}) \right) \\
&\leqslant c' \max_{\sigma\in\Gamma} \sum_{i\in\mathcal{H}_1\cap A_r} \Delta_{i,\epsilon_0}^{-2} \log(\frac{|A_r|}{\sigma(i)} r^2 \frac{\log(\Delta_{i,\epsilon_0}^{-2})}{\delta}) \qquad (4.17) \\
&\leqslant c' \sum_{i=1}^{|\mathcal{H}_1\cap A_r|} \epsilon_0^{-2} \log(\frac{|A_r|}{i} r^2 \frac{\log(\epsilon_0^{-2})}{\delta}) \\
&\leqslant c'' |A_r| \epsilon_0^{-2} \log(r \frac{\log(\epsilon_0^{-2})}{\delta}) \qquad (4.18)
\end{aligned}
$$

where the last line follows from the fact that for any $p \leqslant |A_r|$, $\sum_{i=1}^{p} \log(\frac{|A_r|}{i}) \leqslant |A_r|$.

Step 4: finishing bound (4.11)**.** Using lines (4.18) and (4.13),

$$\mathbf{1}\{E\}\tau_{r,k} \leqslant c'[2^{r-1}(r-1) + \log(|A_r|\epsilon_0^{-2}\log(r\frac{\log(\epsilon_0^{-2})}{\delta}))|A_r|\epsilon_0^{-2}\log(r\frac{\log(\epsilon_0^{-2})}{\delta})]$$

deterministically, which yields line (4.11).

□

Proof of Theorem 21. As in the proof of Lemma 25, define

$$\tau_k = \min(t \in \mathbb{N} \cup \{\infty\} : \exists s \text{ such that } |A_s \cap \mathcal{H}_{1,\epsilon_0}| \geqslant k \text{ and } \mathcal{I}_s \cap A_s \cap \mathcal{H}_1 \subset \mathcal{S}_t).$$

As was argued in the proof of Lemma 25, for all $t \geqslant \tau_k$, $\mathbb{E}[|\mathcal{S}_t \cap \mathcal{H}_1|] \geqslant (1-\delta)k$ since for $t \geqslant \tau_k$.

Step 1: A lower bound on the probability of a good event. Define $E_r = E_{\epsilon_0,r} \cap E_{0,r} \cap E_{1,r}$. We note that since $\{\rho_{i,r}\}_{i\in[n],r\in\mathbb{N}}$ are independent, $\{E_r\}_{r\in\mathbb{N}}$ are independent events. Let $\epsilon_0 \in [0, 1-\mu_0)$ such that $n_{\epsilon_0} \geqslant k$ minimize (4.10). Let r_0 be the smallest integer such that

$$\min(40\frac{n}{n_{\epsilon_0}}k, n) \leqslant 2^{r_0} \leqslant 80\frac{n}{n_{\epsilon_0}}k,$$

Note that if $2^{r_0} \geqslant n$, then the bracket r_0 has n arms.

Next, we bound $\mathbb{P}(E^c_{\epsilon_0,r_0})$. If $2^{r_0} \geqslant n$, then $\mathbb{P}(E^c_{\epsilon_0,r_0}) = 0$, so assume that $2^{r_0} < n$. Note that since the elements of A_{r_0} are chosen uniformly from $[n]$ and $|A_{r_0}| = 2^{r_0} \geqslant 40\frac{n}{|\mathcal{H}_{1,\epsilon_0}|}k$ we have that

$$\begin{aligned}\mathbb{E}[|\mathcal{H}_{1,\epsilon_0} \cap A_{r_0}|] &= \frac{|\mathcal{H}_{1,\epsilon_0}|}{n}|A_{r_0}| \\ &\geqslant 40k.\end{aligned}$$

Then, by a Chernoff bound for hypergeometric random variables,

$$\mathbb{P}(|\mathcal{H}_{1,\epsilon_0} \cap A_{r_0}| \leqslant 20k) \leqslant \exp(-\frac{1}{8}40k) \leqslant \exp(-5).$$

Thus, E_{ϵ_0,r_0} occurs with probability at least $1-\exp(-5)$. Furthermore, we note that for any $r \geqslant r_0$, $\mathbb{P}(E^c_{\epsilon_0,r}) \leqslant \exp(-5)$.

Furthermore, by Lemma 8 of Jamieson and Jain (2018), for any $r \in \mathbb{N}$ and $i = 0, 1$,

$$\mathbb{P}(E^c_{i,r}) = \mathbb{E}[\mathbb{P}(E^c_{i,r}|A_r)] \leqslant \delta.$$

Finally, note that for every $r \geqslant r_0$ and any $\delta \leqslant 0.025$ we have

$$\mathbb{P}(E^c_r) \leqslant \exp(-5) + 2\delta \leqslant \frac{1}{16}.$$

Furthermore, we claim that $\mathbb{P}(\cap_{l=r_0}^{\infty} E^c_l) = 0$. Let $s \geqslant r_0$; then, using the independence between brackets,

$$\mathbb{P}(\cap_{l=r_0}^{\infty} E^c_l) \leqslant \mathbb{P}(\cap_{l=r_0}^{s} E^c_l) = \frac{1}{16^s} \longrightarrow 0$$

as $s \longrightarrow \infty$, proving the claim.

Step 2: Gap-Independent bound on the number of samples. For the sake of brevity, write τ instead of τ_k. Then, by the independence between brackets, $\cup_{r=r_0}^{\infty} E_r \cap (\cap_{r_0 \leqslant l < r} E_l^c)$ occurs with probability 1, and line 4.11 of Lemma 25,

$$
\begin{aligned}
&\mathbb{E}[\tau] \\
&= \mathbb{E}[\tau \mathbf{1}\{\cup_{r=r_0}^{\infty} E_r \cap (\cap_{r_0 \leqslant l < r} E_l^c)\}] \\
&\leqslant \sum_{r=r_0}^{\infty} \mathbb{E}[\tau \mathbf{1}\{E_r \cap (\cap_{r_0 \leqslant l < r} E_l^c)\}] \\
&= \sum_{r=r_0}^{\infty} \mathbb{E}[\tau \mathbf{1}\{E_r\}] \mathbb{P}(\cap_{r_0 \leqslant l < r} E_l^c) \\
&\leqslant \sum_{r=r_0}^{\infty} [2^{r-1}(r-1) + \log(|A_r| \epsilon_0^{-2} \log(r \frac{\log(\epsilon_0^{-2})}{\delta})) |A_r| \epsilon_0^{-2} \log(r \frac{\log(\epsilon_0^{-2})}{\delta}))] \frac{1}{16^{r-r_0}} \\
&\leqslant \sum_{s=0}^{\infty} [2^{r_0} \cdot 2^s (r_0 + s) \\
&\quad + \log(2^s |A_{r_0}| \epsilon_0^{-2} \log((r_0+s) \frac{\log(\epsilon_0^{-2})}{\delta})) 2^s |A_{r_0}| \epsilon_0^{-2} \log((r_0+s) \frac{\log(\epsilon_0^{-2})}{\delta}))] \frac{1}{16^s}
\end{aligned}
$$

We note that

$$
\begin{aligned}
\log((r_0+s) \frac{\log(\epsilon_0^{-2})}{\delta})) &\leqslant c[\log(r_0 \frac{\log(\epsilon_0^{-2})}{\delta})) + \log(s \frac{\log(\epsilon_0^{-2})}{\delta}))] \\
&\leqslant c' \log(r_0 \frac{\log(\epsilon_0^{-2})}{\delta})) + c \log(s)
\end{aligned}
$$

and

$$
\begin{aligned}
&\log(2^s |A_{r_0}| \epsilon_0^{-2} \log((r_0+s) \frac{\log(\epsilon_0^{-2})}{\delta})) \\
&= \log(|A_{r_0}| \epsilon_0^{-2} \log((r_0+s) \frac{\log(\epsilon_0^{-2})}{\delta})) + s \\
&\leqslant \log(|A_{r_0}| \epsilon_0^{-2} c' \log(r_0 \frac{\log(\epsilon_0^{-2})}{\delta}) + c \log(s)) + s \\
&\leqslant c'' \log(|A_{r_0}| \epsilon_0^{-2} \log(r_0 \frac{\log(\epsilon_0^{-2})}{\delta}))) + c''' \log(\log(s)) + s \\
&\leqslant c'' \log(|A_{r_0}| \epsilon_0^{-2} \log(r_0 \frac{\log(\epsilon_0^{-2})}{\delta}))) + c'''' s
\end{aligned}
$$

Then,

$$\sum_{s=0}^{\infty} \log(2^s |A_{r_0}| \epsilon_0^{-2} \log((r_0+s)\frac{\log(\epsilon_0^{-2})}{\delta}))2^s |A_{r_0}| \epsilon_0^{-2} \log((r_0+s)\frac{\log(\epsilon_0^{-2})}{\delta}))\frac{1}{16^s}$$
$$\leqslant \sum_{s=0}^{\infty} [c'' \log(|A_{r_0}| \epsilon_0^{-2} \log(r_0 \frac{\log(\epsilon_0^{-2})}{\delta}))) + c'''' s] |A_{r_0}| \epsilon_0^{-2} [c' \log(r_0 \frac{\log(\epsilon_0^{-2})}{\delta})) + c \log(s)] \frac{1}{8^s}$$
$$\leqslant c'' c' \log(|A_{r_0}| \epsilon_0^{-2} \log(r_0 \frac{\log(\epsilon_0^{-2})}{\delta})))|A_{r_0}| \epsilon_0^{-2} \log(r_0 \frac{\log(\epsilon_0^{-2})}{\delta}))$$
$$+ c'''' \log(|A_{r_0}| \epsilon_0^{-2} \log(r_0 \frac{\log(\epsilon_0^{-2})}{\delta})))|A_{r_0}| \epsilon_0^{-2}$$
$$+ c''''' |A_{r_0}| \epsilon_0^{-2} \log(r_0 \frac{\log(\epsilon_0^{-2})}{\delta})) + + c''''' |A_{r_0}| \epsilon_0^{-2}$$
$$\leqslant c''''' \log(|A_{r_0}| \epsilon_0^{-2} \log(r_0 \frac{\log(\epsilon_0^{-2})}{\delta})))|A_{r_0}| \epsilon_0^{-2} \log(r_0 \frac{\log(\epsilon_0^{-2})}{\delta}))$$

Plugging in $|A_{r_0}|$ and r_0 yields the gap independent bound.

□

4.6.2 Proof of FWER-TPR

In this section, we prove an upper bound for the FWER-TPR version of our Algorithm (see Algorithm 14). We note that the upper bound (4.5) in Theorem 20 follows as a corollary since whenever the FWER-TPR version of our Algorithm 14 accepts an arm, the FDR-TPR version of our Algorithm 13 accepts the same arm.

Theorem 22. *Let $\delta \leqslant .025$. Let $k \leqslant |\mathcal{H}_1|$. For all $\epsilon > 0$ define*

$$S_k(\epsilon) := \frac{k}{|\mathcal{H}_{1,\epsilon}|}\Big(\sum_{i \in \mathcal{H}_1} \Delta_{i,\epsilon}^{-2} \log(\tfrac{n}{|\mathcal{H}_{1,\epsilon}|} k \log(\Delta_{i,\epsilon}^{-2})/\delta) + \sum_{i \in \mathcal{H}_0} \Delta_{i,\epsilon}^{-2} \log(\log(\Delta_{i,\epsilon}^{-2})/\delta)\Big).$$

Then, Algorithm 14 has the property that $\mathbb{P}(\exists t : \mathcal{Q}_t \cap \mathcal{H}_0 \neq \varnothing) \leqslant 2\delta$ and there exists a stopping time τ_k wrt $(\mathcal{F})_{t \in \mathbb{N}}$ such that

$$\mathbb{E}[\tau_k] \leqslant c \min_{\epsilon \geqslant \Delta : |\mathcal{H}_{1,\epsilon}| \geqslant k} S(\epsilon) \log[S(\epsilon) + \epsilon^{-2} \log(\tfrac{n}{|\mathcal{H}_{1,\epsilon}|} k \log(\epsilon^{-2})/\delta)] \tag{4.19}$$

and for all $t \geqslant \tau_k$, $\mathbb{E}[|\mathcal{Q}_t \cap \mathcal{H}_1|] \geqslant (1-\delta)k$.

Lemma 26. *Fix $\delta < 0.025$. Fix $k \leqslant |\mathcal{H}_1|$ and any $\epsilon_0 > 0$ such that $|\mathcal{H}_{1,\epsilon_0}| \geqslant k$. Define for any $r \in \mathbb{N}$,*

$$U_r := \frac{\min(2^r, n)}{n}[\sum_{i \in \mathcal{H}_1} \Delta_{i,\epsilon_0}^{-2} \log(\min(2^r, n) r \frac{\log(\Delta_{i,\epsilon_0}^{-2})}{\delta})) + \sum_{i \in \mathcal{H}_0} \Delta_{i,\epsilon_0}^{-2} \log(\frac{\log(\Delta_{i,\epsilon_0}^{-2})}{\delta})].$$

Then, there exists a random variable τ_k such that for all $t \geqslant \tau_k$, $\mathbb{E}[|\mathcal{Q}_t \cap \mathcal{H}_1|] \geqslant (1-\delta)k$, and

$$\mathbb{E}[\mathbf{1}\{E_{\epsilon_0,r} \cap E_{0,r} \cap E_{1,r}\}\tau_k] \leqslant c[2^{r-1}(r-1) + U_r \log(U_r + \epsilon_0^{-2} \log(\min(2^r, n) r \frac{\log(\epsilon_0^{-2})}{\delta})))] \tag{4.20}$$

where c is a universal constant.

Proof. **Step 1: Define stopping time.** Define

$$\tau_k = \min(t \in \mathbb{N} \cup \{\infty\} : \exists s \text{ such that } |A_s \cap \mathcal{H}_{1,\epsilon_0}| \geqslant k \text{ and } \mathcal{I}_s \cap A_s \cap \mathcal{H}_1 \subset \mathcal{Q}_t).$$

Observe that for all $t \geqslant \tau_k$, $\mathbb{E}[|\mathcal{Q}_t \cap \mathcal{H}_1|] \geqslant (1-\delta)k$ since for $t \geqslant \tau_k$

$$\mathbb{E}[|\mathcal{Q}_t \cap \mathcal{H}_1|] \geqslant \mathbb{E}[|\mathcal{I}_s \cap A_s \cap \mathcal{H}_1|] \geqslant (1-\delta)|A_s \cap \mathcal{H}_1| \geqslant (1-\delta)k.$$

Let $r \in \mathbb{N}$. Define

$$T = |\{t \in \mathbb{N} : \mathcal{I} \cap A_r \cap \mathcal{H}_{1,\epsilon_0} \not\subset \mathcal{Q}_t \text{ and } R_t = r\}|,$$

By the same argument used in Lemma (25) to obtain line (4.13),

$$\mathbf{1}\{E\}\tau_k \leqslant [2^{r-1}(r-1) + \log(\mathbf{1}\{E\}T)\mathbf{1}\{E\}T]. \tag{4.21}$$

We can use the same argument that was used to obtain line (4.17) in Lemma 25 by the lower bounds $1 \leqslant \sigma(i)$ and $p \geqslant 1$ to obtain

$$\begin{aligned}
\mathbf{1}\{E\}T &\leqslant c\Bigg(\sum_{i \in \mathcal{H}_0 \cap A_r} \Delta_{i,\epsilon_0}^{-2} \log(\frac{\log(\Delta_{i,\epsilon_0}^{-2})}{\delta}) && (4.22)\\
&\qquad + |\mathcal{H}_1 \cap \mathcal{H}_{1,\epsilon_0}^c \cap A_r| \epsilon_0^{-2} \log(\frac{\log(\epsilon_0^{-2})}{\delta}) + \sum_{i \in \mathcal{H}_1 \cap A_r} \Delta_{i,\epsilon_0}^{-2} \log(|A_r| r^2 \frac{\log(\Delta_{i,\epsilon_0}^{-2})}{\delta}) \Bigg) \\
&\leqslant c'[\sum_{i \in \mathcal{H}_0 \cap A_r} \Delta_{i,\epsilon_0}^{-2} \log(\frac{\log(\Delta_{i,\epsilon_0}^{-2})}{\delta}) + \sum_{i \in \mathcal{H}_1 \cap A_r} \Delta_{i,\epsilon_0}^{-2} \log(|A_r| r \frac{\log(\Delta_{i,\epsilon_0}^{-2})}{\delta})] && (4.23)\\
&:= c' S_r && (4.24)
\end{aligned}$$

where the second inequality follows from the fact that $\Delta_{i,\epsilon_0} = \epsilon_0$ for all $i \in \mathcal{H}_1 \cap \mathcal{H}_{1,\epsilon_0}^c \cap A_r$ so the third term absorbs the second.

Using lines (4.23) and (4.21),

$$\mathbf{1}\{E\}\tau_k \leqslant c[2^{r-1}(r-1) + \log(S_r)S_r]$$

but note that now the bound depends on the particular random elements of $A_r \cap \mathcal{H}_0$ and $A_r \cap \mathcal{H}_1$.

Step 2: Bounding $\mathbb{E}[\log(S_r)S_r]$. Next, taking the expectation of both sides and focusing on the expectation of the second term,

$$\mathbb{E}[\log(S_r)S_r] = \sum_{i\in\mathcal{H}_0} \Delta_{i,\epsilon_0}^{-2} \log(\frac{\log(\Delta_{i,\epsilon_0}^{-2})}{\delta})\mathbb{E}[\mathbf{1}\{i \in A_r\}\log(S_r)] + \sum_{i\in\mathcal{H}_1} \Delta_{i,\epsilon_0}^{-2} \log(|A_r|r^2\frac{\log(\Delta_{i,\epsilon_0}^{-2})}{\delta})\mathbb{E}[\mathbf{1}\{i \in A_r\}\log(S_r)].$$

It suffices to bound the first sum since the argument for the second is the same.

$$\begin{aligned}
\mathbb{E}[\mathbf{1}\{j \in A_r\}\log(S_r)] &= \mathbb{E}[\log(S_r)|j \in A_r]\frac{\min(2^r, n)}{n} && (4.25)\\
&\leqslant \log(\mathbb{E}[S_r|j \in A_r])\frac{\min(2^r, n)}{n} && (4.26)\\
&= \log(\frac{\min(2^r-1, n-1)}{n-1}[\sum_{i\in\mathcal{H}_1} \Delta_{i,\epsilon_0}^{-2}\log(\min(2^r, n)r\frac{\log(\Delta_{i,\epsilon_0}^{-2})}{\delta})) \\
&\quad + \sum_{i\in\mathcal{H}_0\setminus j} \Delta_{i,\epsilon_0}^{-2}\log(\frac{\log(\Delta_{i,\epsilon_0}^{-2})}{\delta})] + \Delta_{j,\epsilon_0}^{-2}\log(\frac{\log(\Delta_{j,\epsilon_0}^{-2})}{\delta}))\frac{\min(2^r, n)}{n} && (4.27)\\
&\leqslant \log(S_r + \Delta_{j,\epsilon_0}^{-2}\log(\frac{\log(\Delta_{j,\epsilon_0}^{-2})}{\delta}))\frac{\min(2^r, n)}{n}, && (4.28)
\end{aligned}$$

where line (4.25) follows by the law of total expectation, line (4.26) follows by Jensen's inequality, and line (4.28) follows since $\frac{a}{b} \leqslant \frac{a+1}{b+1}$ if $a \leqslant b$. Thus, collecting terms,

$$\mathbb{E}[\mathbf{1}\{E_{\epsilon_0,r} \cap E_{0,r} \cap E_{1,r}\}\tau_{r,k}] \leqslant U_r \log(U_r + \epsilon_0^{-2}\log(\min(2^r, n)r\frac{\log(\epsilon_0^{-2})}{\delta}))$$

yielding line (4.20). □

Proof of Theorem 22. **Step 1: Showing $\mathbb{P}(\exists t : \mathcal{Q}_t \cap \mathcal{H}_0 \neq \varnothing) \leqslant 2\delta$.** First, we show that

$\mathbb{P}(\exists t : \mathcal{Q}_t \cap \mathcal{H}_0 \neq \varnothing) \leqslant 2\delta$.

$$\begin{aligned}
\mathbb{P}(\exists t : \mathcal{Q}_t \cap \mathcal{H}_0 \neq \varnothing) &\leqslant \sum_{r=1}^{\infty} \mathbb{P}(\exists t : \mathcal{Q}_t \cap A_r \cap \mathcal{H}_0 \neq \varnothing) \\
&\leqslant \sum_{r=1}^{\infty} \mathbb{P}(\exists t \in \mathbb{N} \text{ and } i \in \mathcal{H}_0 \cap A_r : \widehat{\mu}_{i,r,N_{i,r}(t)} - U(N_{i,r}(t), \frac{\delta}{|A_r| r^2}) \geqslant \mu_0) \\
&\leqslant \sum_{r=1}^{\infty} \mathbb{P}(\exists t \in \mathbb{N} \text{ and } i \in \mathcal{H}_0 \cap A_r : \widehat{\mu}_{i,r,N_{i,r}(t)} - U(N_{i,r}(t), \frac{\delta}{|A_r| r^2}) \geqslant \mu_i) \\
&\leqslant \sum_{r=1}^{\infty} |A_r \cap \mathcal{H}_0| \frac{\delta}{|A_r| r^2} \\
&\leqslant \sum_{r=1}^{\infty} \frac{\delta}{r^2} \\
&\leqslant \delta \frac{\pi^2}{6}
\end{aligned}$$

Step 2: Defining the stopping time. As in the proof of Lemma 26, define

$$\tau_k = \min(t \in \mathbb{N} \cup \{\infty\} : \exists s \text{ such that } |A_s \cap \mathcal{H}_{1,\epsilon_0}| \geqslant k \text{ and } \mathcal{I}_s \cap A_s \cap \mathcal{H}_1 \subset \mathcal{Q}_t).$$

As was argued in the proof of Lemma 26, for all $t \geqslant \tau_k$, $\mathbb{E}[|\mathcal{Q}_t \cap \mathcal{H}_1|] \geqslant (1-\delta)k$ since for $t \geqslant \tau_k$.

Step 3: A lower bound on the probability of a good event. Define $E_r = E_{\epsilon_0,r} \cap E_{0,r} \cap E_{1,r}$. We note that since $\{\rho_{i,r}\}_{i\in[n],r\in\mathbb{N}}$ are independent, $\{E_r\}_{r\in\mathbb{N}}$ are independent events. Let $\epsilon_0 > 0$ such that $|\mathcal{H}_{1,\epsilon_0}| \geqslant k$ minimize (4.19). Let r_0 be the smallest integer such that

$$\min(40\frac{n}{n_{\epsilon_0}}k, n) \leqslant 2^{r_0} \leqslant 80\frac{n}{n_{\epsilon_0}}k,$$

Note that if $2^{r_0} \geqslant n$, then the bracket r_0 has n arms.

As was argued in the proof of Theorem 21 we have that

$$\mathbb{P}(E_r^c) \leqslant \exp(-5) + 2\delta \leqslant \frac{1}{16}.$$

and that $\mathbb{P}(\cap_{l=r_0}^{\infty} E_l^c) = 0$.

Step 4: Gap-Dependent bound on the number of samples. For the sake of brevity, write τ instead of τ_k. Then, by the independence between brackets and $\cup_{r=r_0}^{\infty} E_r \cap (\cap_{r_0 \leqslant l < r} E_l^c)$

occurs with probability 1,

$$\begin{aligned}
\mathbb{E}[\tau] &= \mathbb{E}[\tau \mathbf{1}\{\cup_{r=r_0}^{\infty} E_r \cap (\cap_{r_0 \leqslant l < r} E_l^c)\}] \\
&\leqslant \sum_{r=r_0}^{\infty} \mathbb{E}[\tau \mathbf{1}\{E_r \cap (\cap_{r_0 \leqslant l < r} E_l^c)\}] \\
&= \sum_{r=r_0}^{\infty} \mathbb{E}[\tau \mathbf{1}\{E_r\}] \mathbb{P}(\cap_{r_0 \leqslant l < r} E_l^c) \\
&\leqslant \sum_{r=r_0}^{\infty} c[2^{r-1}(r-1) + U_r \log(U_r + \epsilon_0^{-2} \log(\min(2^r, n) r \frac{\log(\epsilon_0^{-2})}{\delta}))] \frac{1}{16^{r-r_0}} \\
&\leqslant \sum_{r=r_0}^{\infty} c[2^{r-r_0} \cdot 2^{r_0-1}(r-1) + 4^{r-r_0} U_{r_0} \log(4^{r-r_0} U_{r_0} + \epsilon_0^{-2} \log(2^{r_0} \cdot 2^{r-r_0} r \frac{\log(\epsilon_0^{-2})}{\delta}))] \frac{1}{16^{r-r_0}}
\end{aligned}$$

where we used Lemma 25 and the fact that $4^s U_r \geqslant U_{r+s}$ for any $s \geqslant 1$, which holds by the following argument

$$\begin{aligned}
4^s U_r &= 4^s \frac{\min(2^r, n)}{n} [\sum_{i \in \mathcal{H}_1} \Delta_{i,\epsilon_0}^{-2} \log(\min(2^r, n) r \frac{\log(\Delta_{i,\epsilon_0}^{-2})}{\delta})) + \sum_{i \in \mathcal{H}_0} \Delta_{i,\epsilon_0}^{-2} \log(\frac{\log(\Delta_{i,\epsilon_0}^{-2})}{\delta})] \\
&\geqslant \frac{\min(2^{r+s}, n)}{n} [\sum_{i \in \mathcal{H}_1} \Delta_{i,\epsilon_0}^{-2} \log(\min(2^{r2^s}, n) r \frac{\log(\Delta_{i,\epsilon_0}^{-2})}{\delta})) + \sum_{i \in \mathcal{H}_0} \Delta_{i,\epsilon_0}^{-2} \log(\frac{\log(\Delta_{i,\epsilon_0}^{-2})}{\delta})] \\
&\geqslant \frac{\min(2^{r+s}, n)}{n} [\sum_{i \in \mathcal{H}_1} \Delta_{i,\epsilon_0}^{-2} \log(\min(2^{r+s}, n) r \frac{\log(\Delta_{i,\epsilon_0}^{-2})}{\delta})) + \sum_{i \in \mathcal{H}_0} \Delta_{i,\epsilon_0}^{-2} \log(\frac{\log(\Delta_{i,\epsilon_0}^{-2})}{\delta})] \\
&= U_{r+s}.
\end{aligned}$$

Next, we bound the first term.

$$\begin{aligned}
\sum_{r=r_0}^{\infty} 2^{r-r_0} \cdot 2^{r_0-1}(r-1) \frac{1}{16^{r-r_0}} &\leqslant \sum_{r=r_0}^{\infty} 2^{r_0-1}(r_0 + r - r_0) \frac{1}{8^{r-r_0}} \\
&\leqslant c 2^{r_0-1} r_0 + \sum_{r=r_0}^{\infty} \frac{2^{r_0-1}(r-r_0)}{8^{r-r_0}} \\
&\leqslant c' 2^{r_0-1} r_0
\end{aligned}$$

Next, we bound the second term.

$$\begin{aligned}
&\sum_{r=r_0}^{\infty} \frac{1}{4^{r-r_0}} U_{r_0} \log(4^{r-r_0} U_{r_0} + \epsilon_0^{-2} \log(2^{r_0} \cdot 2^{r-r_0}(s+r_0) \frac{\log(\epsilon_0^{-2})}{\delta}))) \\
&= \sum_{s=0}^{\infty} \frac{1}{4^s} U_{r_0} \log(4^s U_{r_0} + \epsilon_0^{-2} \log(2^{r_0} \cdot 2^s (s+r_0) \frac{\log(\epsilon_0^{-2})}{\delta}))) \\
&\leqslant \sum_{s=0}^{\infty} \frac{1}{4^s} U_{r_0} \log(4^s U_{r_0} + c' \epsilon_0^{-2} \log(2^{r_0} r_0 \frac{\log(\epsilon_0^{-2})}{\delta})) + c' s) \\
&\leqslant U_{r_0} \sum_{s=0}^{\infty} \frac{1}{4^s} [c'' \log(4^s U_{r_0} + c' \epsilon_0^{-2} \log(2^{r_0} r_0 \frac{\log(\epsilon_0^{-2})}{\delta}))) + c''' \log(s)] \\
&\leqslant U_{r_0} \sum_{s=0}^{\infty} \frac{1}{4^s} [c'' \log(U_{r_0} + c' \epsilon_0^{-2} \log(2^{r_0} r_0 \frac{\log(\epsilon_0^{-2})}{\delta}))) + c'' \log(4^s) + c''' \log(s)] \\
&\leqslant c'''' U_{r_0} \log(U_{r_0} + \epsilon_0^{-2} \log(2^{r_0} r_0 \frac{\log(\epsilon_0^{-2})}{\delta})))
\end{aligned}$$

where we used

$$\begin{aligned}
\log(2^{r_0} \cdot 2^s (s+r_0) \frac{\log(\epsilon_0^{-2})}{\delta})) &= \log(2^{r_0} (s+r_0) \frac{\log(\epsilon_0^{-2})}{\delta})) + cs \\
&\leqslant c' \log(2^{r_0} r_0 \frac{\log(\epsilon_0^{-2})}{\delta})) + c' \log(2^{r_0} s \frac{\log(\epsilon_0^{-2})}{\delta})) + cs \\
&\leqslant c'' \log(2^{r_0} r_0 \frac{\log(\epsilon_0^{-2})}{\delta})) + c'''' s
\end{aligned}$$

Finally, the result follows from noting that $U_{r_0} \leqslant cS(\epsilon_0)$ for some universal constant c and

$$2^{r_0} r_0 \leqslant c \frac{n}{n_{\epsilon_0}} k \log(\frac{n}{n_{\epsilon_0}} k) \leqslant cS(\epsilon_0) \log(S(\epsilon_0))$$

where the last inequality follows since $\Delta_{i,\epsilon_0}^{-2} \geqslant 1$ for all $i \in [n]$.

□

4.6.3 Proof of ϵ-Good Arm Identification

Recall the relevant notation. For a fixed $\epsilon > 0$ define $m = \{i \in [n] : \mu_i > \mu_i - \epsilon\}$. For any $\gamma \in (0, \epsilon)$ define $G_\gamma = \{i \in [n] : \mu_i \geqslant \mu_1 - \gamma\}$ and define the gaps wrt to ϵ and γ as follows:

$$\Delta_{i,\epsilon,\gamma} = \begin{cases} \mu_i - \mu_{m+1} & i \leqslant |G_\gamma| \\ \max(\mu_{|G_\gamma|} - \mu_i, \mu_i - \mu_{m+1}) & |G_\gamma| < i \leqslant m \cdot \\ \mu_{|G_\gamma|} - \mu_i & i \geqslant m+1 \end{cases}$$

Note that as γ decreases on $(\epsilon, 0)$ the gaps only *increase* but $|G_\gamma|$ *decreases*. We restate Theorem 17 with the doubly logarithmic terms.

Theorem 23. *Let $\delta \in (0,1)$. For all $\epsilon > 0$ define*

$$U_\epsilon(\gamma) := \frac{1}{|G_\gamma|}\Big(\sum_{j=1}^{m} \Delta_{j,\epsilon,\gamma}^{-2} \log(\tfrac{n}{|G_\gamma|}\log(\Delta_{j,\epsilon,\gamma}^{-2})/\delta) + \sum_{j=m+1}^{n} \Delta_{j,\epsilon,\gamma}^{-2} \log(\log(\Delta_{j,\epsilon,\gamma}^{-2})/\delta)\Big).$$

Then, Algorithm 13 has the property that there exists a stopping time τ wrt $(\mathcal{F})_{t\in\mathbb{N}}$ such that

$$\mathbb{E}[\tau] \leqslant c \min_{\gamma\in(0,\epsilon)} U_\epsilon(\gamma)\log(U_\epsilon(\gamma) + \Delta_{m,\epsilon,\gamma}^{-2}\log(\tfrac{n}{|G_\gamma|}\log(\Delta_{m,\epsilon,\gamma}^{-2})/\delta)) \tag{4.29}$$

$$\leqslant cU_\epsilon \log(U_\epsilon + (\mu_m - \mu_{m+1})^{-2}\log(\tfrac{n}{m}\log((\mu_m-\mu_{m+1})^{-2})/\delta)) \tag{4.30}$$

and $\mathbb{P}(\exists s \geqslant \tau : \mu_{O_s} \leqslant \mu_1 - \epsilon) \leqslant 2\delta$.

Lemma 27 is the key intermediate result in the proof of Theorem 23; its role is similar to that of Lemma 25 in the proof of Theorem 21 and the proof is technically similar to the proof of Lemma 25. For any $r \in \mathbb{N}$ define the events

$$F_{r,1} = \{A_r \cap G_\gamma \neq \varnothing\}$$

$$F_{r,2} = \{ \sum_{i\in A_r : \mu_i < \frac{\mu_{|G_\gamma|}+\mu_{m+1}}{2}} \Delta_{i,\gamma}^{-2}\log(\frac{1}{\rho_{i,r}}) \leqslant 5 \sum_{i\in A_r:\mu_i<\frac{\mu_{|G_\gamma|}+\mu_{m+1}}{2}} \Delta_{i,\gamma}^{-2}\log(\frac{1}{\delta})\},$$

$$F_{r,3} = \{\exists i_0 \in A_r \cap G_\gamma \text{ s.t.} \forall t \in \mathbb{N} : |\widehat{\mu}_{i_0,r,t} - \mu_{i_0}| \leqslant U(t,\delta)\}.$$

$F_{r,1}$ says that there is at least one γ-good arm in bracket r. $F_{r,2}$ allows us to avoid a union bound and says that most of the arms in bracket r have large $\rho_{i,r}$. Finally, $F_{r,3}$ says that at least one of the arms that is γ-good and and in rth bracket concentrates well in the sense that $\rho_{i,r} \geqslant \delta$.

Lemma 27. *Let $\epsilon > 0$, $\gamma \in (0,\epsilon)$, and $r \in \mathbb{N}$. Define*

$$Y_r = \frac{\min(2^r, n)}{n}[\sum_{j:\mu_j\leqslant\mu_1-\epsilon} \Delta_{j,\epsilon,\gamma}^{-2}\log(\frac{\log(\Delta_{j,\epsilon,\gamma}^{-2})}{\delta}) + \sum_{j:\mu_j>\mu_1-\epsilon}\Delta_{j,\epsilon,\gamma}^{-2}\log(|A_r|r\frac{\log(\Delta_{j,\epsilon,\gamma}^{-2})}{\delta})]$$

Then, there exists a random variable τ such that $\mathbb{P}(\exists s \geqslant \tau : \mu_{O_s} \leqslant \mu_1 - \epsilon) \leqslant 2\delta$, and

$$\mathbb{E}[\mathbf{1}\{F_{r,1}\cap F_{r,2}\cap F_{r,3}\}\tau] \leqslant c[2^{r-1}(r-1) + Y_r\log(Y_r + \Delta_{m,\epsilon,\gamma}^{-2}\log(|A_r|r\frac{\log(\Delta_{m,\epsilon,\gamma}^{-2})}{\delta})]. \tag{4.31}$$

Proof. **Step 1: Define stopping time.** Our strategy is to define a stopping time τ that says that some arm i that is ϵ-good has been sampled enough times so that its confidence bound is sufficiently small and then to show that with high probability for all $t \geqslant \tau$, *(i)* the lower confidence bound of arm i is above μ_{m+1} and *(ii)* the algorithm always outputs an ϵ-good arm. To this end, define

$$\tau = \min\{t \in \mathbb{N} \cup \{\infty\} : \exists s \in \mathbb{N} \text{ and } \exists i \in A_s \text{ s.t. } \mu_i \geqslant \frac{\mu_{|G_\gamma|} + \mu_{m+1}}{2} \\ \text{and } N_{i,s}(t) \geqslant U^{-1}(\frac{\Delta_{i,\epsilon,\gamma}}{4}, \frac{\delta}{|A_s|s^2})\}.$$

We claim that $\mathbb{P}(\exists t \geqslant \tau : \mu_{O_t} < \mu_1 - \epsilon) \leqslant 2\delta$. Define the event

$$F = \{\forall t \in \mathbb{N}, s \in \mathbb{N}, \text{ and } i \in A_s : |\widehat{\mu}_{i,s,t} - \mu_i| \leqslant U(t, \frac{\delta}{|A_s|s^2})\}.$$

By a union bound, F occurs with probability at least $1-2\delta$. Suppose F occurs and let $t \geqslant \tau$. Then, since $t \geqslant \tau$, there exists a bracket s and an arm $i \in A_s$ such that $\mu_i \geqslant \frac{\mu_{|G_\gamma|}+\mu_{m+1}}{2}$ and $N_{i,s}(t) \geqslant U^{-1}(\frac{\Delta_{i,\epsilon,\gamma}}{4}, \frac{\delta}{|A_s|s^2})$. Then by event F,

$$\begin{aligned}\widehat{\mu}_{i,s,N_{i,s}(t)} - U(N_{i,s}(t), \frac{\delta}{|A_s|s^2}) &\geqslant \mu_i - 2U(N_{i,s}(t), \frac{\delta}{|A_s|s^2}) \\ &> \mu_i - \frac{\Delta_{i,\epsilon,\gamma}}{2} \\ &\geqslant \mu_{m+1}.\end{aligned}$$

Towards a contradiction, suppose that there exists a bracket $s_0 \in \mathbb{N}$ and another arm $j \in A_{s_0}$ $(j \neq i)$ such that $\mu_j \leqslant \mu_1 - \epsilon$ and the algorithm outputs j at time t. Then, by event F,

$$\mu_j \geqslant \widehat{\mu}_{j,s_0,N_{j,s_0}(t)} - U(N_{j,s_0}(t), \frac{\delta}{|A_{s_0}|s_0^2}) \geqslant \widehat{\mu}_{i,s,N_{i,s}(t)} - U(N_{i,s}(t), \frac{\delta}{|A_s|s^2}) > \mu_{m+1} \geqslant \mu_j,$$

which is a contradiction. Thus, $\mathbb{P}(\exists t \geqslant \tau : \mu_{O_t} < \mu_1 - \epsilon) \leqslant 2\delta$.

Step 2: Relating τ to bracket r. Next, we bound $\mathbb{E}[\mathbf{1}\{F_{r,1} \cap F_{r,2} \cap F_{r,3}\}\tau]$. For the sake of brevity, define $F_r \coloneqq F_{r,1} \cap F_{r,2} \cap F_{r,3}$ and since we will only focus on bracket r, write $\widehat{\mu}_{i,t}$, $N_i(t)$, and ρ_i instead of $\widehat{\mu}_{i,r,t}$, $N_{i,r}(t)$, and $\rho_{i,r}$. Define

$$T = |\{t \in \mathbb{N} : R_t = r \text{ and } \nexists i \in A_r \text{ s.t. } \mu_i \geqslant \frac{\mu_{|G_\gamma|} + \mu_{m+1}}{2} \text{ and } N_i(t) \geqslant U^{-1}(\frac{\Delta_{i,\epsilon,\gamma}}{4}, \frac{\delta}{|A_r|r^2})\}|,$$

i.e., the number of rounds that the algorithm works on the rth bracket and there does not exist $i \in A_r$ s.t. $\mu_i \geqslant \frac{\mu_{|G_\gamma|}+\mu_{m+1}}{2}$ and $N_i(t) \geqslant U^{-1}(\frac{\Delta_{i,\epsilon,\gamma}}{2}, \frac{\delta}{|A_r|r^2})$. By the same argument given in line (4.13) in Lemma 25, we have that

$$\mathbf{1}\{F_r\}\tau \leqslant c[2^{r-1}(r-1) + \log(T\mathbf{1}\{F_r\})T\mathbf{1}\{F_r\}].$$

Step 3: Bounding $T\mathbf{1}\{F_r\}$. In the interest of brevity, define $F(t) = \{\nexists i \in A_r \text{ s.t. } \mu_i \geqslant \frac{\mu_{|G_\gamma|}+\mu_{m+1}}{2} \text{ and } N_i(t) \geqslant U^{-1}(\frac{\Delta_{i,\epsilon,\gamma}}{4}, \frac{\delta}{|A_r|r^2})\}$. Then,

$$\begin{aligned}\mathbf{1}\{F_r\}T &\leqslant \mathbf{1}\{F_r\}\sum_{t=1}^{\infty}\mathbf{1}\{R_t = r, F(t)\} \\ &\leqslant \mathbf{1}\{F_r\}\sum_{t:R_t=r}^{\infty}\mathbf{1}\{\mu_{I_t} < \frac{\mu_{|G_\gamma|}+\mu_{m+1}}{2}\} + \mathbf{1}\{\mu_{I_t} \geqslant \frac{\mu_{|G_\gamma|}+\mu_{m+1}}{2}, F(t)\}\end{aligned}$$

We bound each sum separately. Note that by $F_{r,3}$ there exists an $i_0 \in A_r \cap G_\gamma$ such that

$$\widehat{\mu}_{i_0,N_{i_0}(t)} + U(N_{i_0}(t),\delta) \geqslant \mu_{i_0} \geqslant \mu_1 - \gamma. \tag{4.32}$$

Let j such that $\mu_j < \frac{\mu_{|G_\gamma|}+\mu_{m+1}}{2}$. Then,

$$\widehat{\mu}_{j,N_j(t)} + U(N_j(t),\delta) \leqslant \mu_j + U(N_j(t),\rho_j) + U(N_j(t),\delta) \leqslant \mu_j + 2U(N_j(t),\rho_j\delta).$$

Thus, line (4.32) implies that if $N_j(t) \geqslant U^{-1}(\frac{\Delta_{j,\epsilon,\gamma}}{4}, \rho_j\delta)$, arm j is not pulled since in that case

$$\widehat{\mu}_{j,N_j(t)} + U(N_j(t),\delta) \leqslant \mu_j + 2U(N_j(t),\rho_j\delta) \leqslant \mu_j + \frac{\Delta_{j,\epsilon,\gamma}}{2} \leqslant \mu_{|G_\gamma|}.$$

Thus, by arguments made throughout this paper (e.g., line (4.15) of the proof of Lemma 25) and the event $F_{r,2}$,

$$\sum_{t:R_t=r}^{\infty}\mathbf{1}\{\mu_{I_t} \leqslant \mu_1 - \epsilon\} \leqslant c\sum_{j\in A_r:\mu_j<\frac{\mu_{|G_\gamma|}+\mu_{m+1}}{2}}\Delta_{j,\epsilon,\gamma}^{-2}\log(\frac{\log(\Delta_{j,\epsilon,\gamma}^{-2})}{\delta})$$

Finally, by event F we clearly have

$$\sum_{t:R_t=r}^{\infty}\mathbf{1}\{\mu_{I_t} \geqslant \frac{\mu_{|G_\gamma|}+\mu_{m+1}}{2}, F(t)\} \leqslant c\sum_{j\in A_r:\mu_j\geqslant\frac{\mu_{|G_\gamma|}+\mu_{m+1}}{2}}\Delta_{j,\epsilon,\gamma}^{-2}\log(|A_r|r\frac{\log(\Delta_{j,\epsilon,\gamma}^{-2})}{\delta})$$

Thus,

$$\begin{aligned}\mathbf{1}\{F_r\}T &\leqslant c[\sum_{j\in A_r:\mu_j<\frac{\mu_{|G_\gamma|}+\mu_{m+1}}{2}}\Delta_{j,\epsilon,\gamma}^{-2}\log(\frac{\log(\Delta_{j,\epsilon,\gamma}^{-2})}{\delta})\\&+\sum_{j\in A_r:\mu_j\geqslant\frac{\mu_{|G_\gamma|}+\mu_{m+1}}{2}}\Delta_{j,\epsilon,\gamma}^{-2}\log(|A_r|r\frac{\log(\Delta_{j,\epsilon,\gamma}^{-2})}{\delta})]\\&\leqslant c[\sum_{j\in A_r:\mu_j\leqslant\mu_1-\epsilon}\Delta_{j,\epsilon,\gamma}^{-2}\log(\frac{\log(\Delta_{j,\epsilon,\gamma}^{-2})}{\delta})+\sum_{j\in A_r:\mu_j>\mu_1-\epsilon}\Delta_{j,\epsilon,\gamma}^{-2}\log(|A_r|r\frac{\log(\Delta_{j,\epsilon,\gamma}^{-2})}{\delta})]\\&:=cX_r\end{aligned}$$

Then, using the same argument from lines (4.25)-(4.28), we have that

$$\mathbb{E}X_r\log(X_r)\leqslant cY_r\log(Y_r+\Delta_{m,\epsilon,\gamma}^{-2}\log(|A_r|r\frac{\log(\Delta_{m,\epsilon,\gamma}^{-2})}{\delta})]$$

Thus, putting it together,

$$\mathbb{E}[\mathbf{1}\{\mathbf{1}\{F_r\}\tau]\leqslant c[2^{r-1}(r-1)+Y_r\log(Y_r+\Delta_{m,\epsilon,\gamma}^{-2}\log(|A_r|r\frac{\log(\Delta_{m,\epsilon,\gamma}^{-2})}{\delta})]$$

□

Proof of Theorem 23. Let $\gamma_0\in(0,\epsilon)$ minimize the optimization problem in line (4.30). Let r_0 such that be the smallest integer such that

$$\min(40\frac{n}{|G_{\gamma_0)|}},n)\leqslant 2^{r_0}\leqslant 80\frac{n}{|G_{\gamma_0|}}.$$

Next, we bound $\mathbb{P}((F_{r_0,1}\cap F_{r_0,2}\cap F_{r_0,3})^c)$. By a union bound and the law of total probability,

$$\begin{aligned}\mathbb{P}((F_{r_0,1}\cap F_{r_0,2}\cap F_{r_0,3})^c)&\leqslant\mathbb{P}(F_{r_0,1}^c\cap F_{r_0,3}^c)+\mathbb{P}(F_{r_0,2}^c)\\&\leqslant\mathbb{P}(F_{r_0,1}^c)+\mathbb{P}(F_{r_0,3}^c|F_{r,1})+\mathbb{P}(F_{r_0,2}^c)\\&\leqslant 2\delta+\mathbb{P}(F_{r_0,1}^c)\\&\leqslant 2\delta+\exp(-5)\\&\leqslant\frac{1}{16}\end{aligned}$$

The rest of the proof proceeds as the proof of Theorem 21 starting at step 2. □

4.6.4 Proof of FWER-FWPD

Finally, we present a Theorem for the FWER-FWPD version of Algorithm 14. Although it is possible to use the ideas from the other upper bound proofs to establish a result that depends on the distribution of the arms in $\mathcal{H}_1$, for simplicity our upper bound is in terms of $\Delta = \min_{i \in \mathcal{H}_1} \mu_i - \mu_0$ and m.

Theorem 24. *Let $\delta \in (0, \frac{1}{600})$. Let $k \leqslant |\mathcal{H}_1|$. Define*

$$\begin{aligned}\widetilde{V}_k &:= (\frac{n}{m}k - k)\Delta^{-2}\log(\max(k, \log\log(\frac{n}{m}k\frac{1}{\delta}))\log(\Delta^{-2})\log(\frac{n}{m}k)/\delta) \\ &\quad + k\log(\max(\frac{n}{m}k - (1 - 2\delta(1+4\delta))k, \log\log(\frac{n}{m}k\frac{1}{\delta}))\log(\Delta^{-2})\log(\frac{n}{m}k)/\delta)] \\ &\lesssim (\frac{n}{m}k - k)\Delta^{-2}\log(k/\delta) + k\log(\frac{\frac{n}{m}k - (1-2\delta(1+4\delta))k}{\delta})\end{aligned}$$

Furthermore, define

$$\lambda_k = \min(t \in \mathbb{N} : |\mathcal{R}_t \cap \mathcal{H}_1| \geqslant k).$$

Then, Algorithm 14 has the property that $\mathbb{P}(\exists t \in \mathbb{N} : \mathcal{R}_t \cap \mathcal{H}_0 \neq \varnothing) \leqslant 10\delta$ and

$$\mathbb{E}[\lambda_k] \leqslant c\log(\widetilde{V}_k)\widetilde{V}_k.$$

Lemma 28. *Let $\delta \in (0, .01)$. Let $k \leqslant |\mathcal{H}_1|$. Let $r \in \mathbb{N}$ such that $2^r \geqslant k$. Define*

$$\lambda_r = \min(t \in \mathbb{N} : |\mathcal{R}_t \cap A_r \cap \mathcal{H}_1| \geqslant k).$$

Define

$$\begin{aligned}V_r &:= (2^r - \min(|\mathcal{H}_1|, \frac{|\mathcal{H}_1|}{n}2^r))\Delta^{-2}\log(\max(\min(|\mathcal{H}_1|, \frac{|\mathcal{H}_1|}{n}2^r), \log\log(r2^r/\delta))\log(\Delta^{-2})r/\delta) \\ &\quad + \min(|\mathcal{H}_1|, \frac{|\mathcal{H}_1|}{n}2^r)\log(\max(2^r - (1-2\delta(1+4\delta))\min(|\mathcal{H}_1|, \frac{|\mathcal{H}_1|}{n}2^r), \log\log(\frac{r2^r}{\delta}))\log(\Delta^{-2})r/\delta)]\end{aligned}$$

Then with probability at least $1 - 6\delta - 2\exp(-2^{r-3}) - \mathbb{P}(|A_r \cap \mathcal{H}_1| < k)$,

$$\lambda_r \leqslant c(2^{r-1}(r-1) + \log(V_r)V_r).$$

Proof. **Step 1: Definitions and events.** Recall R_t is the bracket chosen at time t and define

$$T = |\{t \in \mathbb{N} : A_r \cap \mathcal{H}_1 \nsubseteq \mathcal{R}_t \text{ and } R_t = r\}|,$$

i.e., the number of rounds that the algorithm works on the rth bracket and $A_r \cap \mathcal{H}_1 \nsubseteq \mathcal{R}_t$. Define the events

$$\Sigma_{r,1} = \{|A_r \cap \mathcal{H}_1| \geqslant k\}$$
$$\Sigma_{r,2} = \{|A_r \cap \mathcal{H}_1| \leqslant \min(|\mathcal{H}_1|, \tfrac{|\mathcal{H}_1|}{n} 2^{r+1}))\}$$
$$\Sigma_{r,3} = \{|A_r \cap \mathcal{H}_1| \geqslant \min(|\mathcal{H}_1|, \tfrac{|\mathcal{H}_1|}{n} 2^{r-1}))\}$$

If $2^{r+1} \geqslant n$, then $|A_r \cap \mathcal{H}_1| \leqslant |\mathcal{H}_1|$ implies $\mathbb{P}(\Sigma_{r,2}^c) = 0$. Therefore, suppose $2^{r+1} < n$. Then, by multiplicative Chernoff for hypergeometric random variables,

$$\mathbb{P}(\Sigma_{r,2}^c) = \mathbb{P}(|A_r \cap \mathcal{H}_1| > \frac{|\mathcal{H}_1|}{n} 2^{r+1}) \leqslant \exp(-\frac{|\mathcal{H}_1|}{n} 2^{r-2}) \leqslant \exp(-2^{r-2})$$

Similarly, if $2^r \geqslant n$, then $|A_r| = n$ and $\mathbb{P}(\Sigma_{r,2}^c) = 0$. Therefore, suppose $2^r < n$.

$$\mathbb{P}(\Sigma_{r,3}^c) = \mathbb{P}(|A_r \cap \mathcal{H}_1| < \frac{|\mathcal{H}_1|}{n} 2^{r-1})) \leqslant \exp(-\frac{|\mathcal{H}_1|}{n} 2^{r-3}) \leqslant \exp(-2^{r-3})$$

Since the algorithm essentially runs the FWER-FWDP version of the algorithm from Jamieson and Jain (2018) on each bracket r with confidence δ/r^2, we can apply Theorem 4 of Jamieson and Jain (2018) directly to obtain that there exists an event $\Sigma_{r,4}$, which only depends on the samples of the arms in bracket r, such that $\mathbb{P}(\Sigma_{r,4}^c) \leqslant 6\delta$ and on $\Sigma_{r,4}$

$$\begin{aligned} T \leqslant c[&(|A_r| - |A_r \cap \mathcal{H}_1|)\Delta^{-2} \log(\max(|A_r \cap \mathcal{H}_1|, \log\log(|A_r|/\delta_r)) \log(\Delta^{-2})/\delta_r) \\ &+ |A_r \cap \mathcal{H}_1|\Delta^{-2} \log(\max(|A_r| - (1 - 2\delta_r(1 + 4\delta_r)|A_r \cap \mathcal{H}_1|, \log\log(\frac{|A_r|}{\delta_r})) \log(\Delta^{-2})/\delta_r)]. \end{aligned}$$

This roughly says $T \lesssim (|A_r| - |A_r \cap \mathcal{H}_1|)\Delta^{-2} \log(|A_r \cap \mathcal{H}_1|/\delta) + |A_r \cap \mathcal{H}_1|\Delta^{-2} \log((|A_r| - |A_r \cap \mathcal{H}_1|)/\delta)$.

Step 2: Bounding λ_r. In what follows, assume $\Sigma_{r,1} \cap \Sigma_{r,2} \cap \Sigma_{r,3} \cap \Sigma_{r,4}$ occur which happens with probability at least

$$1 - 6\delta - 2\exp(-2^{r-3}) - \mathbb{P}(\Sigma_{r,1}^c).$$

By the same argument given in lines (4.12) and (4.13), event $\Sigma_{r,1}$ implies that

$$\lambda_r \leqslant c(2^{r-1}(r-1) + \log(T)T).$$

Furthermore, using $\Sigma_{r,2} \cap \Sigma_{r,3} \cap \Sigma_{r,4}$,

$$\begin{aligned}
T \leqslant & c[(|A_r| - |A_r \cap \mathcal{H}_1|)\Delta^{-2} \log(\max(|A_r \cap \mathcal{H}_1|, \log\log(|A_r|/\delta_r)) \log(\Delta^{-2})/\delta_r) \\
& + |A_r \cap \mathcal{H}_1|\Delta^{-2} \log(\max(|A_r| - (1 - 2\delta_r(1 + 4\delta_r))|A_r \cap \mathcal{H}_1|, \log\log(\frac{|A_r|}{\delta_r})) \log(\Delta^{-2})/\delta_r)] \\
\leqslant & c'[(|A_r| - |A_r \cap \mathcal{H}_1|)\Delta^{-2} \log(\max(|A_r \cap \mathcal{H}_1|, \log\log(r|A_r|/\delta)) \log(\Delta^{-2})r/\delta) \\
& + |A_r \cap \mathcal{H}_1|\Delta^{-2} \log(\max(|A_r| - (1 - 2\delta(1 + 4\delta))|A_r \cap \mathcal{H}_1|, \log\log(\frac{r|A_r|}{\delta})) \log(\Delta^{-2})r/\delta)] \\
\leqslant & c''[(2^r - \min(|\mathcal{H}_1|, \frac{|\mathcal{H}_1|}{n}2^r))\Delta^{-2} \log(\max(\min(|\mathcal{H}_1|, \frac{|\mathcal{H}_1|}{n}2^r), \log\log(r2^r/\delta)) \log(\Delta^{-2})r/\delta) \\
& + \min(|\mathcal{H}_1|, \frac{|\mathcal{H}_1|}{n}2^r)\Delta^{-2} \log(\max(2^r - (1 - 2\delta(1 + 4\delta)) \\
& \cdot \min(|\mathcal{H}_1|, \frac{|\mathcal{H}_1|}{n}2^r), \log\log(\frac{r2^r}{\delta})) \log(\Delta^{-2})r/\delta)]
\end{aligned}$$

□

Proof of Theorem 24. We note that the algorithm essentially runs the FWER-FWDP version of the algorithm from Jamieson and Jain (2018) on each bracket r with confidence δ/r^2. Therefore, by Theorem 4 from Jamieson and Jain (2018),

$$\mathbb{P}(\exists t \in \mathbb{N} : A_r \cap \mathcal{R}_t \cap \mathcal{H}_0 \neq \varnothing) \leqslant 6\frac{\delta}{r^2}$$

Thus,

$$\begin{aligned}
\mathbb{P}(\exists t \in \mathbb{N} : \mathcal{R}_t \cap \mathcal{H}_0 \neq \varnothing) &\leqslant \mathbb{P}(\exists t \in \mathbb{N}, r \in \mathbb{N} : A_r \cap \mathcal{R}_t \cap \mathcal{H}_0 \neq \varnothing) \\
&\leqslant \sum_{r \in \mathbb{N}} \mathbb{P}(\exists t \in \mathbb{N} : A_r \cap \mathcal{R}_t \cap \mathcal{H}_0 \neq \varnothing) \\
&\leqslant \sum_{r \in \mathbb{N}} 6\frac{\delta}{r^2} \\
&\leqslant 10\delta.
\end{aligned}$$

Let $r_0 \in \mathbb{N}$ be the smallest integer such that $r_0 \geqslant 6$ and

$$\min(40\frac{n}{m}k, n) \leqslant 2^{r_0} \leqslant 80\frac{n}{m}k.$$

If $2^{r_0} \geqslant n$, then $\mathbb{P}(|A_r \cap \mathcal{H}_1| < k) = 0$. Otherwise,by multiplicative Chernoff for hypergeometric random variables,

$$\mathbb{P}(|A_r \cap \mathcal{H}_1| < k) \leqslant \exp(-5).$$

In the interest of brevity, define $\Sigma_r = \Sigma_{r,1} \cap \Sigma_{r,2} \cap \Sigma_{r,3} \cap \Sigma_{r,4}$. Observe that $\{\Sigma_r\}_{r\in\mathbb{N}}$ are mutually independent. Further, using $\delta \in (0, \frac{1}{600})$, for all brackets $r \geqslant r_0$, the events occur which happens with probability at least

$$\mathbb{P}(\Sigma_r^c) \leqslant 6\delta + 2\exp(-2^{r-3}) + \mathbb{P}(\Sigma_{r,1}^c) \leqslant \frac{1}{16}$$

The rest of the proof proceeds as in Step 2 of the proof of Theorem 21. □

4.7 Best of both Worlds Algorithm for ϵ-Good Arm Identification

We consider the version of LUCB from Kalyanakrishnan et al. (2012a). Let LUCB(ϵ) denote the LUCB algorithm that terminates once it finds an ϵ-good arm. Let $\beta(t, \delta)$ denote the confidence bound used in Kalyanakrishnan et al. (2012a); although, it is possible to tighten these confidence bounds, for the sake of simplicity and brevity we use theirs so that we can appeal to their sample complexity results. Algorithm 15 takes a desired tolerance $\epsilon > 0$ as input, runs LUCB(ϵ) and the ϵ-good arm identification version of Algorithm 13 in parallel without sample sharing between the algorithms,[3], and outputs an arm $\widehat{i}_t$ at every iteration. This arm $\widehat{i}_t$ is the arm O_t suggested by Algorithm 13 for every iteration until the termination condition of LUCB(ϵ) obtains at which point algorithm 15 decides whether to output O_t or the arm suggested by LUCB(ϵ). Let $\widehat{\mu}_{i,t}$ denote the empirical mean at time t of arm i based on the samples collected by LUCB(ϵ) and $N_{i,t}$ denote the number of pulls of arm i at time t by LUCB(ϵ).

Proof of Theorem 18. Theorem 6 of Kalyanakrishnan et al. (2012a) implies that there exists a stopping time τ_{PAC} wrt $(\mathcal{F}_t)_{t\in\mathbb{N}}$ such that at time τ_{PAC} the Algorithm 15 terminates and (4.4) holds. Theorem 17 implies the existence of stopping time τ_{simple} wrt $(\mathcal{F}_t)_{t\in\mathbb{N}}$ such that (4.3) holds and $\mathbb{P}(\exists s \geqslant \tau_{simple} : \mu_{O_s} \leqslant \mu_1 - \epsilon_2) \leqslant 2\delta$.

It remains to show that when the Algorithm 15 terminates at $t = \tau_{PAC}$, $\mathbb{P}(\mu_{\widehat{i}_{\tau_{PAC}}} \leqslant$

[3]Samples should be shared in practice.

$\mu_1 - \min(\epsilon_1, \epsilon_2)) \leqslant 3\delta$. Define the event

$$F = \{\forall t \in \mathbb{N}, s \in \mathbb{N}, \text{ and } i \in A_s : |\widehat{\mu}_{i,s,t} - \mu_i| \leqslant U(t, \frac{\delta}{|A_s|s^2})\}.$$

By a union bound, F occurs with probability at least $1 - 2\delta$. By the argument in Step 1 of the proof of Lemma 27, on F, for all $t \geqslant \tau_{simple}$

$$\max_{r \in \mathbb{N}} \widehat{\mu}_{O_t,r,N_{O_t,r}(t)} - U(N_{O_t,r}(t), \frac{\delta}{|A_r|r^2}) > \max_{i:\mu_i \leqslant \mu_1 - \epsilon_2} \mu_i.$$

Next, define the event

$$E = \{\forall t \in \mathbb{N} \text{ and } \forall i \in [n] : |\widehat{\mu}_{i,t} - \mu_i| \leqslant \beta(t, \delta)\}$$

By Theorem 1 of Kalyanakrishnan et al. (2012a), $\mathbb{P}(E) \geqslant 1 - \delta$ and on E,

$$\widehat{\mu}_{\widehat{j},N_{\widehat{j}}(\tau_{PAC})} - \beta(N_{\widehat{j}}(\tau_{PAC}), \delta) > \mu_1 - \epsilon_1$$

Suppose F and E occur, which by a union bound occur with probability at least $1 - 3\delta$. Either $\widehat{i}_{\tau_{PAC}} = \widehat{j}$ or $\widehat{i}_{\tau_{PAC}} = O_{\tau_{PAC}}$. Suppose $\widehat{i}_{\tau_{PAC}} = \widehat{j}$. Then,

$$\begin{aligned}
\mu_{\widehat{i}_{\tau_{PAC}}} &= \mu_{\widehat{j}} \\
&\geqslant \widehat{\mu}_{\widehat{j},N_{\widehat{j}}(\tau_{PAC})} - \beta(N_{\widehat{j}}(\tau_{PAC}), \delta) \\
&> \max_{r \in \mathbb{N}} \widehat{\mu}_{O_t,r,N_{O_t,r}(t)} - U(N_{O_t,r}(t), \frac{\delta}{|A_r|r^2}) \\
&\geqslant \max_{i:\mu_i \leqslant \mu_1 - \epsilon_2} \mu_i,
\end{aligned}$$

which implies that $\mu_{\widehat{i}_{\tau_{PAC}}} \geqslant \mu_1 - \min(\epsilon_1, \epsilon_2)$. A similar argument proves the case $\widehat{i}_{\tau_{PAC}} = O_{\tau_{PAC}}$.

□

Algorithm 15 Best of both Worlds Algorithm: ϵ-Good Arm Identification

Input: $\epsilon > 0$
for $t = 1, 2, \ldots$ **do**
  Pull arm according to sampling rule given by the ϵ-good arm identification version of Algorithm 13
  Pull arm according to sampling rule given by LUCB(ϵ)
  Let O_t be the arm returned by the ϵ-good arm identification version of Algorithm 13
  if LUCB(ϵ) terminates **then**
    Let $\widehat{j}$ denote the arm returned by LUCB(ϵ)
    $r_0 = \text{argmax}_{r \in \mathbb{N}} \widehat{\mu}_{O_t, r, N_{O_t, r}(t)} - U(N_{O_t, r}(t), \frac{\delta}{|A_r| r^2})$
    if $\widehat{\mu}_{O_t, r_0, N_{i,r}(t)} - U(N_{O_t, r_0}(t), \frac{\delta}{|A_{r_0}| r^2}) \geqslant \widehat{\mu}_{\widehat{j}, N_{\widehat{j}}(t)} - \beta(N_{\widehat{j}}(t), \delta)$ **then**
      Set $\widehat{i}_t = O_t$
    else
      Set $\widehat{i}_t = \widehat{j}$
    end if
    Output $\widehat{i}_t$ and terminate.
  else
    Set $\widehat{i}_t = O_t$
    Output $\widehat{i}_t$
  end if
end for

Chapter 5

Decontamination of Mutual Contamination Models

Many machine learning problems can be characterized by *mutual contamination models*. In these problems, one observes several random samples from different convex combinations of a set of unknown base distributions and the goal is to infer these base distributions. This Chapter considers the general setting where the base distributions are defined on arbitrary probability spaces. I examine three popular machine learning problems that arise in this general setting: multiclass classification with label noise, demixing of mixed membership models, and classification with partial labels. In each case, I give sufficient conditions for identifiability and present algorithms for the infinite and finite sample settings, with associated performance guarantees. This Chapter is joint work with Clayton Scott and Gilles Blanchard, and it was published in the Journal of Machine Learning Research in 2019.

5.1 Introduction

In many machine learning problems, the learner observes several random samples from different mixtures of unknown base distributions, with unknown mixing weights, and the goal is to infer these base distributions. Examples include binary classification with label noise, multiclass classification with label noise, classification with partial labels, and topic modeling. The goal of this paper is to develop a unified framework and set of tools to study

statistical properties of these problems in a very general setting.

To this end, we use the general framework of mutual contamination models (Blanchard and Scott, 2014). In a mutual contamination model, there are L distributions $P_1, \ldots, P_L$ called *base distributions*. The learner observes M random samples

$$X_1^i, \ldots, X_{n_i}^i \overset{i.i.d.}{\sim} \tilde{P}_i = \sum_{j=1}^{L} \pi_{i,j} P_j \tag{5.1}$$

where $i = 1, \ldots, M$, $\pi_{i,j} \geqslant 0$, and $\sum_j \pi_{i,j} = 1$. Here $\pi_{i,j}$ is the probability that an instance of the *contaminated distribution* $\tilde{P}_i$ is a realization of P_j. The $\pi_{i,j}$s and P_js are unknown and the $\tilde{P}_i$s are observed through data. In this work, we avoid parametric models and assume that the sample space is arbitrary. The model can be stated concisely as

$$\tilde{\boldsymbol{P}} = \boldsymbol{\Pi} \boldsymbol{P} \tag{5.2}$$

where $\boldsymbol{P} = (P_1, \ldots, P_L)^T$, $\tilde{\boldsymbol{P}} = (\tilde{P}_1, \ldots, \tilde{P}_M)^T$, and $\boldsymbol{\Pi} = \left(\pi_{i,j}\right)$ is an $M \times L$ matrix that we call the *mixing matrix*.

In this paper we study *decontamination* of mutual contamination models, which is the problem of recovering, or estimating, the base distributions $\boldsymbol{P}$ from the contaminated distributions $\tilde{\boldsymbol{P}}$ from which data are observed, without knowledge of the mixing matrix $\boldsymbol{\Pi}$. We focus our attention on three specific types of mutual contamination models, all of which describe modern problems in machine learning: multiclass classification with label noise, demixing of mixed membership models and classification with partial labels. We will demonstrate that these three decontamination problems can be addressed using a common set of concepts and techniques. Before elaborating our contributions in detail, we first offer an overview of the three specific mutual contamination models, and associated decontamination problems, that we study.

Multiclass Classification with Label Noise: In multiclass classification with label noise, $M = L$ and the goal is to recover $\boldsymbol{P}$. Each P_i represents the distribution of a class of examples. The learner observes training examples with noisy labels, that is, realizations from the $\tilde{P}_j$s. This problem arises in nuclear particle classification (Scott et al., 2013). When one draws samples of a specific particle, it is impossible to remove other types of particles from the background. Thus, each example is drawn from a mixture of the different types of particles.

Demixing of Mixed Membership Models: We consider the following decontamination problem in mixed membership models: given a sample from each $\tilde{P}_i$, recover $\boldsymbol{P}$ up to a permutation. We refer to this decontamination problem as *demixing of mixed membership models*. This problem arises in the task of automatically uncovering the thematic topics of a corpus of documents. Under the mixed membership model approach, the words of each document are thought of as being drawn from a document-specific mixture of topics. Specifically, documents correspond to the $\tilde{P}_i$s and the topics to the P_is. This approach is also referred to as topic modeling. As we discuss in the next section, our theory significantly generalizes existing topic modeling theoretical guarantees.

Classification with Partial Labels:[1] In classification with partial labels, each data point is labeled with a *partial label* $Y \subset \{1, \ldots, L\}$; the true label is in Y, but it is not known which label is the true one. In our setup, we view the ith random sample as having partial label $Y_i \coloneqq \{j : \boldsymbol{\pi}_{i,j} > 0\}$ and being distributed according to $\tilde{P}_i = \sum_{j \in Y_i} \pi_{i,j} P_j$. Thus, the learner observes training examples from the contaminated distributions $\tilde{\boldsymbol{P}}$ and the *partial label matrix* $\boldsymbol{\Pi}^+ = (\mathbf{1}\{\boldsymbol{\Pi}_{i,j} > 0\})$, and the goal is to recover $\boldsymbol{P}$.

There are many applications of classification with partial labels because often abundant sources of data are naturally associated with information that can be interpreted as partial labels. For example, consider the task of face recognition. On the internet, there are many images with captions that indicate who is in the picture but do not indicate which face belongs to which person. A partial label could be formed by associating each face with the names of the individuals appearing in the same image (Cour et al., 2011).

Although our work emphasizes recovery of $\boldsymbol{P}$, it is also possible to think of decontamination of mutual contamination models as concerned with estimation of the mixing matrix $\boldsymbol{\Pi}$. This estimate of $\boldsymbol{\Pi}$ could be used as a plug-in for recently developed debiased losses for multiclass classification with label noise and classification with partial labels, which require knowledge of $\boldsymbol{\Pi}$ (Cid-Sueiro, 2012; Menon et al., 2015b; van Rooyen and Williamson, 2015; Patrini et al., 2017).

In this paper, we make the following contributions: (i) We give sufficient conditions on

[1] Classification with partial labels has also been referred to as the "superset learning problem" or the "multiple label problem" (Liu and Dietterich, 2014).

$\boldsymbol{P}$, $\boldsymbol{\Pi}$, and $\boldsymbol{\Pi}^+$ for identifiability of the three problems. (ii) We establish necessary conditions that in some cases match or are similar to the sufficient conditions. (iii) We introduce novel algorithms for the infinite and finite sample settings. These algorithms are nonparametric in the sense that they do not model P_i as a probability vector or other parametric model. Our algorithmic contributions show that while all three problems can be described in a unified way, the special structure of multiclass classification with label noise allows for a substantially simpler algorithm. (iv) We develop novel estimators for distributions obtained by iteratively applying the κ^* operator (defined below). (v) Finally, our framework gives rise to several novel geometric insights about each of these three problems and leverages concepts from affine geometry, multilinear algebra, and probability.

5.1.1 Notation

Let $\mathbb{Z}^+$ denote the positive integers. For $n \in \mathbb{Z}^+$, let $[n] = \{1, \ldots, n\}$. If $\boldsymbol{x} \in \mathbb{R}^K$, let x_i denote the ith entry of $\boldsymbol{x}$. If $\boldsymbol{x}_j \in \mathbb{R}^K$, then $x_{j,i}$ denotes the ith entry of $\boldsymbol{x}_j$. Let $\boldsymbol{e}_i$ denote the length L vector with 1 in the ith position and zeros elsewhere. Let $\boldsymbol{\pi}_i \in \Delta_L \subset \mathbb{R}^L$ be the transpose of the ith row of $\boldsymbol{\Pi}$ where Δ_L denotes the $(L-1)$-dimensional simplex, i.e., $\Delta_L = \{\boldsymbol{\mu} = (\mu_1, \ldots, \mu_L)^T \in \mathbb{R}^L \mid \sum_{i=1}^L \mu_i = 1 \text{ and } \forall i : \mu_i \geqslant 0\}$. Let Δ_L^M denote the product of M $(L-1)$-dimensional simplices, viewed as the space of $M \times L$ row-stochastic matrices. Let $\mathcal{P}$ denote the space of probability distributions on a measurable space $(\mathcal{X}, \mathcal{C})$. Let $\text{supp}(F)$ denote the support of a distribution F on a Borel space.

5.2 Related Work

Our work makes various contributions to the statistical understanding of multiclass classification with label noise, demixing of mixed membership models, and classification with partial labels. In the following subsections, we discuss how our results improve upon and relate to previous results in the literature.

5.2.1 Multiclass Classification with Label Noise

There has not been much work on classification with multiclass label noise. By contrast, label noise in the binary setting has received a fair amount of attention. For a review of work prior to 2013, see Scott et al. (2013). More recently, Natarajan et al. (2013) considered the binary label noise case where the label noise rates are known (in our case, the label noise rates are unknown). van Rooyen and Williamson (2015) generalized the work of Natarajan et al. (2013) to the multiclass case, but again assumed that the mixing proportions are known. Recent work has proposed various algorithms for the binary setting where the label noise rates are unknown (Scott, 2015; van Rooyen et al., 2015; Menon et al., 2015a), but these algorithms have not been generalized to the multiclass case. Menon et al. (2016) consider the binary setting with instance-dependent corruption, but they assume that the class probability functions take the form of a single-index model, whereas we make no parametric assumptions on the P_is. Ghosh et al. (2017) consider multiclass label noise, but they make two restrictive assumptions: *(i)* in the infinite sample setting, they assume that there exists some function belonging to the chosen hypothesis class that attains 0 risk and *(ii)* in the finite sample setting, they assume that the label noise is symmetric, i.e., there exists a constant $c \in (0, 1)$ such that $\pi_{i,j} = \frac{c}{L-1}$ for all $i \neq j$. Patrini et al. (2017) also study the multiclass setting, but they assume that if their neural network has access to sufficiently many samples, it can perfectly model $\Pr(\tilde{Y} = k \,|\, \boldsymbol{x})$ where $\boldsymbol{x}$ is a given feature vector and $\tilde{Y}$ is a corrupted label. Unlike most previous work that aims to learn a classifier, our focus is on estimating the base distributions. Given these estimates, one could then design a classifier to optimize some performance measure. See, for example, Section 4.3 of our initial work on this subject (Blanchard and Scott, 2014).

Another approach for modeling random label noise, in addition to the mutual contamination model, is the label flipping model. Indeed, several of the above-cited papers adopt this setting. In this model, the label Y of a data point is flipped independently of its features X and

$$\mu_{l,k} \coloneqq \Pr(\tilde{Y} = k \,|\, Y = l)$$

gives the probability that a data point with true label $Y = l$ is corrupted to have an observed

label $\tilde{Y} = k$. Under the assumption that Y and X are jointly distributed, the $\mu_{l,k}$s can be related to the $\pi_{i,j}s$ via Bayes' rule. We choose to study the mutual contamination model because we find it more convenient to study the question of identifiability.

In this paper, we extend Scott et al. (2013), which examined binary classification with label noise (the case where $M = L = 2$). The multiclass setting is significantly more challenging and, as such, requires novel sufficient conditions and mathematical notions. In particular, Scott et al. (2013) use the notion of *irreducibility* of distributions as one of their sufficient conditions.

Definition 8. *For distributions G and H, we say that G is* irreducible *with respect to H if it is not possible to write $G = \gamma H + (1-\gamma)F$ where F is a distribution and $0 < \gamma \leqslant 1$.*

Definition 9. *For distributions G and H, we say that G and H are* mutually irreducible *if G is irreducible with respect to H and H is irreducible with respect to G. We denote*

$$IR = \{(G, H) : G \text{ and } H \text{ are mutually irreducible distributions}\}.$$

Scott et al. (2013) require that P_1 and P_2 are mutually irreducible. To treat the multiclass setting, we introduce a generalization of mutual irreducibility, namely *joint irreducibility*.

The work presented below on multiclass label noise originally appeared in a conference paper (Blanchard and Scott, 2014). The purpose of the present paper is to demonstrate that the framework developed in that paper can be extended to the other two decontamination problems, and to provide a unified presentation of the three settings. In particular, the joint irreducibility assumption plays a pivotal role in all three settings, as does the task of mixture proportion estimation. However, the decontamination procedures for the latter two problems are substantially more complicated than for multiclass classification with label noise.

5.2.2 Demixing Mixed Membership Models

Mixed membership models have become a powerful modeling tool for data where data points are associated with multiple distributions. Applications have appeared in a wide range of fields including image processing (Li and Perona, 2005), population genetics (Pritchard et al., 2000), document analysis (Blei et al., 2003), and surveys (Berkman et al., 1989).

One particularly popular application is topic modeling on a corpus of documents, such as the articles published in the journal Science. Topic modeling is closely related to demixing of mixed membership models and our work may be viewed as studying topic modeling on general domains.

In topic modeling, the base distributions P_i correspond to topics and the contaminated distributions $\tilde{P}_i$ to documents, which are regarded as mixtures of topics. In most cases, the P_is are assumed to have a finite sample space. A variety of approaches have been proposed for topic modeling. The most common approach assumes a generative model for a corpus of documents and determines the maximum likelihood fit of the model given data. However, because maximum likelihood is NP-hard, these approaches must rely on heuristics that can get stuck in local minima (Arora et al., 2012).

Recently, a trend towards algorithms for topic modeling with provable guarantees has emerged. Most of these methods rely on the separability assumption **(SEP)** and its variants (Donoho and Stodden, 2003; Arora et al., 2012, 2013; Ding et al., 2013, 2014; Recht et al., 2012; Huang et al., 2016). According to **(SEP)**, $P_1, \ldots, P_L$ are distributions on a finite sample space and for every $i \in \{1, \ldots, L\}$, there exists a word $x \in \text{supp}(P_i)$ such that $x \notin \cup_{j \neq i} \text{supp}(P_j)$. Our requirement that $P_1, \ldots, P_L$ are jointly irreducible is a natural generalization of separability of $P_1, \ldots, P_L$, as we will argue below. Specifically, if $P_1, \ldots, P_L$ have discrete sample spaces, separability and joint irreducibility coincide; however, if $P_1, \ldots, P_L$ are continuous, under joint irreducibility, $P_1, \ldots, P_L$ can have the same support.

A key ingredient in these algorithms is to use the assumption of a finite sample space to view the distributions as probability vectors in Euclidean space; this leads to approaches based on non-negative matrix factorization (NMF), linear programs, and random projections (Donoho and Stodden, 2003; Arora et al., 2012, 2013; Ding et al., 2013, 2014; Recht et al., 2012; Huang et al., 2016). However, more general distributions cannot be viewed as finite-dimensional vectors. Therefore, topic modeling on general domains requires new techniques. Our work seeks to provide such techniques.

Topic modeling on general domains has several applications, including in high-energy physics (Metodiev and Thaler, 2018a,b). In collider data, quantum chromodynamics causes data samples to be a mixture of different types of particles, where the underlying fraction of

the particle type is unknown. In this setting, it is of interest to recover information about each of the particles. Recently, Metodiev and Thaler (2018b) applied the Demix algorithm, Algorithm 19 in the current paper, to this problem in the case $M = L = 2$.

Topic modeling on general domains is also relevant to recent empirical research on topic modeling with word embeddings, e.g., (Das et al., 2015; Li et al., 2016b,a; Xun et al., 2017; Zhao et al., 2018). Word embeddings map words to vectors in $\mathbb{R}^d$ in a semantically and syntactically meaningful way. Their use has been pivotal to the state-of-art performance of many algorithms in NLP (Luong et al., 2013). Several algorithms for topic modeling with word embeddings model the topics as multivariate Gaussian distributions in order to handle words that do not belong to the vocabulary of the training dataset (Das et al., 2015; Xun et al., 2017). Whereas current topic modeling algorithms with theoretical guarantees do not cover such a modeling approach, the generality of our algorithms does.

5.2.3 Classification with Partial Labels

Classification with partial labels has had two main formulations in previous work (Liu and Dietterich, 2014). In one formulation **(PL-1)**, instances from each class are drawn independently and the partial label for each instance is drawn independently from a set-valued distribution. In another formulation **(PL-2)**, training data are in the form of bags where each bag is a set of instances and the bag has a set of labels. Each instance belongs to a single class, and the set of labels associated with the bag is given by the union of the labels of the instances in the bag. Our framework is similar to **(PL-2)**, although it does not assume a joint distribution on the features of instances and the partial labels.

Most work takes an empirical risk minimization approach to classification with partial labels (Jin and Ghahramani, 2002; Nyugen and Caruana, 2008; Cour et al., 2011; Liu and Dietterich, 2012). Typically, these algorithms aim to pick a classifier that minimizes the *partial label error*: the probability that a given classifier assigns a label to a training instance that is not contained in the partial label associated with the training instance. By contrast, our approach is to estimate the base distributions. One could then use these estimates to train a classifier under some performance measure.

There has not been much theoretical work on developing a statistical understanding of

classification with partial labels. Cid-Sueiro (2012) and van Rooyen and Williamson (2015) develop methods for classification with partial labels that require knowledge of the mixing proportions, e.g., the probability that a label is in a partial label, given the true label. In this work, we make the more realistic assumption that the mixing proportions are unknown.

Liu and Dietterich (2014) consider the question of learnability where the mixing proportions are unknown. They consider two main sufficient conditions for learnability of a partial label problem. First, they require that for every label $l \in [L]$, the probability that l occurs with any particular distinct label l' is less than 1. Our condition on the partial label (described in the next Section) is considerably weaker. For example, it permits the case where there are two labels $l \neq l'$ such that whenever l occurs in a partial label, l' also occurs.

The second sufficient condition of Liu and Dietterich (2014) is based on the class distributions, partial label distributions *and* the hypothesis class of choice. It requires that every hypothesis that attains zero partial label error also attains zero true error. While this condition may be useful for the selection of a suitable hypothesis class for an ERM approach, it is important to develop interpretable sufficient conditions that only depend on the characteristics of a partial label problem. Our work provides such conditions.

We also note that Liu and Dietterich (2014) consider the realizable case, that is, the case where the supports of $P_1, \ldots, P_L$ do not overlap. By contrast, we make the significantly weaker assumption that $P_1, \ldots, P_L$ are jointly irreducible, which allows $P_1, \ldots, P_L$ to have the same support. Thus, our work addresses the agnostic case in classification with partial labels.

5.3 Sufficient Conditions for Identifiability

We can think of each problem as requiring a specific factorization of $\tilde{\boldsymbol{P}}$ in terms of $\boldsymbol{P}$ and $\boldsymbol{\Pi}$. We say $\tilde{\boldsymbol{P}}$ is *factorizable* if there exists $(\boldsymbol{\Pi}, \boldsymbol{P}) \in \Delta_L^M \times \mathcal{P}^L$ such that $\tilde{\boldsymbol{P}} = \boldsymbol{\Pi}\boldsymbol{P}$; we call $(\boldsymbol{\Pi}, \boldsymbol{P})$ a *factorization* of $\tilde{\boldsymbol{P}}$. Multiclass classification with label noise requires a specific ordering of the elements of $\boldsymbol{P}$; classification with partial labels requires that $\boldsymbol{\Pi}$ is consistent with $\boldsymbol{\Pi}^+$ and a specific ordering of the elements of $\boldsymbol{P}$.

A factorization is not guaranteed to exist. For example, there is no factorization in the

case where $M = 3$, $L = 2$, and $\tilde{P}_1, \tilde{P}_2, \tilde{P}_3$ are linearly independent. When a factorization exists, in general it is not unique. For instance, consider the case where $L = M$, $(\mathbf{\Pi}, \boldsymbol{P})$ solves (5.2), and $\mathbf{\Pi}$ is not a permutation matrix. Then, another solution is $\tilde{\boldsymbol{P}} = \boldsymbol{I}\tilde{\boldsymbol{P}}$. Furthermore, there are infinitely many solutions in the following general case.

Proposition 4. *Suppose that $\tilde{\boldsymbol{P}}$ has at least two distinct $\tilde{P}_j$s and has a factorization $(\mathbf{\Pi}, \boldsymbol{P})$. If there is some $\tilde{P}_i$ in the interior of* $\operatorname{conv}(P_1, \ldots, P_L)$*, then there are infinitely many distinct non-trivial factorizations of $\tilde{\boldsymbol{P}}$.*

Proof. Without loss of generality, suppose that $i = 1$ and $\tilde{P}_1 \neq \tilde{P}_2$. Then, since $\tilde{P}_1$ is in the interior of $\operatorname{conv}(P_1, \ldots, P_L)$, there is some $\delta > 0$ such that for any $\alpha \in (1, 1 + \delta)$, $Q_\alpha = \alpha\tilde{P}_1 + (1-\alpha)\tilde{P}_2$ is a distribution. Then, $\operatorname{conv}(\tilde{P}_1, \ldots, \tilde{P}_L) \subseteq \operatorname{conv}(Q_\alpha, \tilde{P}_2, \ldots, \tilde{P}_L)$ and, consequently, there is some $\mathbf{\Pi}' \in \Delta_L^L$ such that $(\mathbf{\Pi}', (Q_\alpha, \tilde{P}_2, \ldots, \tilde{P}_L)^T)$ solves (5.2). Clearly, by varying α, there are infinitely many solutions to (5.2). □

Identifiability of each problem is equivalent to the existence of a unique factorization for that problem. Therefore, to establish identifiability for the three problems, we must impose conditions on $(\mathbf{\Pi}, \boldsymbol{P})$ and $\mathbf{\Pi}^+$. To this end, we use the notion of joint irreducibility of distributions.

Definition 10. *The distributions $\{P_i\}_{1 \leqslant i \leqslant L}$ are* jointly irreducible *iff the following equivalent conditions hold*

(a) For all $I \subset [L]$ such that $1 \leqslant |I| < L$, and ϵ_i such that $\epsilon_i \geqslant 0$ and $\sum_{i \in I} \epsilon_i = \sum_{i \notin I} \epsilon_i = 1$,

$$(\sum_{i \in I} \epsilon_i P_i, \sum_{i \notin I} \epsilon_i P_i) \in IR.$$

(b) $\sum_{i=1}^{L} \gamma_i P_i$ is a distribution implies that $\gamma_i \geqslant 0 \, \forall i$.

Conditions *(a)* and *(b)*, whose equivalence was established by Blanchard and Scott (2014), give two ways to think about joint irreducibility. Condition *(a)* says that every convex combination of a subset of the P_is is irreducible (see Section 5.2.1) with respect to every convex combination of the other P_is. Condition *(b)* says that if a distribution is in the span of $P_1, \ldots, P_L$, it is in their convex hull. Joint irreducibility holds when each P_i has a region

of positive probability that does not belong to the support of any of the other P_is; thus, separability (see Section 5.2.2) of the P_is entails joint irreducibility of $P_1, \ldots, P_L$. However, the converse is not true: the P_is can have the same support and still be jointly irreducible (e.g., P_is Gaussian with a common variance and distinct means (Scott et al., 2013)).

For all three problems, we assume that

(A) $P_1, \ldots, P_L$ are jointly irreducible.

Henceforth, unless we say otherwise, $P_1, \ldots, P_L$ are assumed to be jointly irreducible. In Appendix 5.13, we provide experiments on real-world datasets that suggest that this assumption is reasonable.

We make different assumptions on $\mathbf{\Pi}$ for each of the three problems. For multiclass classification with label noise, we assume that

(B1) $\mathbf{\Pi}$ is invertible and $\mathbf{\Pi}^{-1}$ is a matrix with strictly positive diagonal entries and non-positive off-diagonal entries.

According to Lemma 29 below, this assumption essentially says that the problem has low noise in the sense that for each i, $\tilde{P}_i$ mostly comes from P_i. In particular, each P_i can be recovered by subtracting small multiples of $\tilde{P}_j$, $j \neq i$ from $\tilde{P}_i$. For example, consider the following case where $\mathbf{\Pi}$ satisfies **(B1)**. Suppose that there is a "common background noise" $\boldsymbol{c} \in \Delta_L$ that appears in different proportions in each of the distributions; formally, we have $\boldsymbol{\pi}_i = \gamma_i \boldsymbol{c} + (1 - \gamma_i)\boldsymbol{e}_i$ with $\gamma_i \in [0, 1)$. In other words, we shift each of the vertices $\boldsymbol{e}_i$ towards a common point $\boldsymbol{c}$ (see panel (iii) of Figure 5.1). See Blanchard and Scott (2014) for a proof that this setup satisfies **(B1)**. In the binary case where $M = L = 2$, **(B1)** is equivalent to the simple condition that $\pi_{1,1} + \pi_{2,2} < 1$. This assumption roughly says that in expectation the majority of labels are correct. In Section 5.4.3, we present Lemma 29, which gives a geometric interpretation of **(B1)**.

For the demixing problem, we assume that

(B2) $\mathbf{\Pi}$ has full column rank.

We note that **(B2)** is considerably weaker than **(B1)**, e.g., it allows $M > L$. Of course, it is natural to demand a weaker sufficient condition for demixing the mixed membership

problem than muliticlass classification with label noise because the goal of the former problem is to recover any permutation of $\boldsymbol{P}$ while the goal of the latter is to recover $\boldsymbol{P}$ exactly. Nevertheless, the identifiability analysis to establish **(B2)** as a sufficient condition is also significantly more involved than the analysis of **(B1)**.

For classification with partial labels, we assume that

(B3) $\boldsymbol{\Pi}$ has full column rank and the columns of $\boldsymbol{\Pi}^+$ are unique.

The assumption that the columns of $\boldsymbol{\Pi}^+$ are unique says that there are no two classes that always appear together in the partial labels. In Appendix 5.9, we argue that several of the above conditions are also necessary, or are not much stronger than what is necessary.

5.4 Algorithms for the Population Case

In this section, to establish that the above conditions are indeed sufficient for identifiability, we give a population case analysis of the three problems. The results on multiclass classification with label noise appeared in a conference paper (Blanchard and Scott, 2014); we refer the reader to that paper for the proofs.

5.4.1 Background

This paper relies on the following quantity from Blanchard et al. (2010).

Definition 11. *Given probability distributions F_0, F_1, define*

$$\kappa^*(F_0 \,|\, F_1) = \max\{\kappa \in [0,1] | \, \exists \text{ a distribution } G \text{ s.t. } F_0 = (1-\kappa)G + \kappa F_1\}.$$

The following Proposition from Blanchard et al. (2010) establishes some useful properties of κ^*.

Proposition 5. *Given probability distributions F_0, F_1 on a measurable space $(\mathcal{X}, \mathcal{C})$, if $F_0 \neq F_1$, then $\kappa^*(F_0 \,|\, F_1) < 1$ and the above maximum is attained for a unique distribution G (which we refer to as the* residue of F_0 wrt. F_1*). Furthermore, the following equivalent characterization holds:*

$$\kappa^*(F_0 \,|\, F_1) = \inf_{C \in \mathcal{C}, F_1(C) > 0} \frac{F_0(C)}{F_1(C)}.$$

Note that $\kappa^*(F_0 \,|\, F_1) = 0$ iff F_0 is irreducible wrt F_1. $\kappa^*(F_0 \,|\, F_1)$ can be thought of as the maximum possible proportion of F_1 in F_0. We can think of $1 - \kappa^*(F_0 \,|\, F_1)$ as a statistical distance since it is non-negative and equal to zero if and only if $F_0 = F_1$. We refer to κ^* as the two-sample κ^* operator. To obtain the residue of F_0 wrt F_1, one computes Residue$(F_0 \,|\, F_1)$ (see Algorithm 16); this is well-defined under Proposition 5 when $F_0 \neq F_1$.

In order to gain intuition about κ^*, we briefly discuss how it can be used to recover $\mathbf{\Pi}^{-1}$ in the case $L = 2$. Under conditions discussed above (Scott et al., 2013), it holds that

$$\tilde{P}_1 = (1 - \kappa_1)P_1 + \kappa_1 \tilde{P}_2, \text{ and}$$
$$\tilde{P}_2 = (1 - \kappa_2)P_2 + \kappa_2 \tilde{P}_1.$$

and $\kappa_1 = \kappa^*(\tilde{P}_1 \,|\, \tilde{P}_2)$ and $\kappa_2 = \kappa^*(\tilde{P}_2 \,|\, \tilde{P}_1)$. By rearranging this system of equations, we can write

$$\boldsymbol{P} = \mathbf{\Pi}^{-1}\tilde{\boldsymbol{P}} = \begin{pmatrix} \frac{1}{1-\kappa_1} & -\frac{\kappa_1}{1-\kappa_1} \\ -\frac{\kappa_2}{1-\kappa_2} & \frac{1}{1-\kappa_2} \end{pmatrix} \tilde{\boldsymbol{P}}.$$

Next, we turn to the multi-sample generalization of κ^*, which we call the multi-sample κ^* operator.

Definition 12. *Given distributions $F_0, \ldots, F_K$, define*

$$\kappa^*(F_0 \,|\, F_1, \ldots, F_K) = \max_{\boldsymbol{\mu} \in \Delta_K} \kappa^*(F_0 \,|\, \sum_{i=1}^{K} \mu_i F_i)$$
$$= \max \Big(\sum_{i=1}^{K} \nu_i : \nu_i \geqslant 0, \sum_{i=1}^{K} \nu_i \leqslant 1, \exists \text{ distribution } G \text{ s.t. } F_0 = (1 - \sum_{i=1}^{K} \nu_i)G + \sum_{i=1}^{K} \nu_i F_i \Big). \quad (5.3)$$

Blanchard and Scott (2014) establish the equivalence in line (5.3), as well as Lemma 43, which shows that the outer maximum is always attained at some $\boldsymbol{\mu} \in \Delta_K$, i.e., κ^* is well-defined. Although there is always a G achieving the max, it is not necessarily unique. Any G attaining the maximum is called a *maximizer* of $\kappa^*(F_0 \,|\, F_1, \ldots, F_K)$. The algorithm MultiResidue$(F_0 \,|\, \{F_1, \ldots, F_K\})$ returns one of these G (see Algorithm 17). If G is unique, we call G the *multi-sample residue of F_0 with respect to $\{F_1, \ldots, F_K\}$*. Under our proposed sufficient conditions, certain residues are shown to exist, and our decontamination methods compute such residues via Algorithm 17. In Section 5.4.3, we discuss Lemma 29, which

Algorithm 16 Residue($F_0 \,|\, F_1$)

1: $\kappa \longleftarrow \kappa^*(F_0 \,|\, F_1)$

2: **return** $\frac{F_0 - \kappa F_1}{1-\kappa}$

Algorithm 17 MultiResidue($F_0 \,|\, \{F_1, \ldots, F_K\}$)

1: $(\nu_1, \ldots, \nu_K)^T \longleftarrow (\nu_1', \ldots, \nu_K')^T$ achieving the maximum in $\kappa^*(F_0 \,|\, F_1, \ldots, F_K)$

2: **return** $\frac{F_0 - \sum_{i=1}^K \nu_i F_i}{1 - \sum_{i=1}^K \nu_i}$

establishes useful conditions under which a multi-sample residue exists and is equal to one of the vertices of Δ_L.

In general, one cannot express the multi-sample version of κ^* in terms of the two-sample version. However, it is possible in some special cases. For example, if one had access to feasible $\nu_1, \ldots, \nu_K$ that attain the optimum in (5.3), then it holds that $\kappa^*(F_0 \,|\, F_1, \ldots, F_K) = \kappa^*(F_0 \,|\, \frac{\sum_{i=1}^K \nu_i F_i}{\sum_{i=1}^K \nu_i})$. Further, it is possible to replace the multi-sample κ^* with several calls of the two-sample κ^* when $K = L-1$, $F_i = P_i$ for all $i \neq 0$ and $F_0 = \sum_{i=1}^L \alpha_i P_i$ where $\sum_i \alpha_i = 1$ and $\forall i \ \alpha_i > 0$ (see Lemmas 34 and 41).

We remark that in previous work that assumes P_i are probability vectors, distributions are compared using l_p distances. By contrast, in our setting of general probability spaces, we use κ^* to compare different distributions.

5.4.2 Mixture Proportions

Recall that we assume that $P_1, \ldots, P_L$ are jointly irreducible. If $\boldsymbol{\eta} \in \mathbb{R}^L$ and $Q = \boldsymbol{\eta}^T \boldsymbol{P}$, we say that $\boldsymbol{\eta}$ is the *mixture proportion* of Q. Since by Lemma 44, joint irreducibility of $P_1, \ldots, P_L$ implies linear independence of $P_1, \ldots, P_L$, mixture proportions are well-defined, i.e., the mixture proportions are unique.

An important feature of our decontamination strategy is recovering various mixture proportions in the simplex Δ_L. To make this precise, we introduce the following definitions. If $i \in [L]$, we say that $\mathrm{conv}(\{\boldsymbol{e}_j : j \neq i\})$ is a *face* of the simplex Δ_L; if $A \subset [L]$ and $|A| = k$, we also say that $\mathrm{conv}(\{\boldsymbol{e}_j : j \in A\})$ is a *k-face* of Δ_L. If $\boldsymbol{\eta} \in \mathbb{R}^L$, Q is a distribution, and $Q = \boldsymbol{\eta}^T \boldsymbol{P}$, we say that $\mathcal{N}(Q) = \mathcal{N}(\boldsymbol{\eta}) = \{j : \eta_j > 0\}$ is the *support set* of $\boldsymbol{\eta}$ or the support

set of Q. Note that by joint irreducibity, $\mathcal{N}(\boldsymbol{\eta})$ consists of the indices of all the nonzero entries in the mixture proportion $\boldsymbol{\eta}$. Finally, for $\boldsymbol{\eta}_i \in \Delta_L$, and $Q_i = \boldsymbol{\eta}_i^T \boldsymbol{P}$ for $i = 1, 2$, we say that the distributions Q_1 and Q_2 (or the mixture proportions $\boldsymbol{\eta}_1$ and $\boldsymbol{\eta}_2$) are *on the same face* of the simplex Δ_L if there exists $j \in [L]$ such that $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2 \in \text{conv}(\{\boldsymbol{e}_k : k \neq j\})$.

The heart of our approach is that under joint irreducibility, one can interchange distributions $Q_1, \ldots, Q_K$ and their mixture proportions $\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_K$, as indicated by the following Proposition. We note that it is valid to to apply the κ^* operator to $\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_K$ since they can be viewed as discrete probability distributions over $[L]$.

Proposition 6. *Let $Q_i = \boldsymbol{\eta}_i^T \boldsymbol{P}$ for $i \in [L]$ and $\boldsymbol{\eta}_i \in \Delta_L$. Suppose $\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_L$ are linearly independent and $P_1, \ldots, P_L$ are jointly irreducible. Then,*

1. *for any $i \in [L]$ and $A \subseteq [L] \backslash \{i\}$, $\kappa^*(Q_i \,|\, \{Q_j : j \in A\}) = \kappa^*(\boldsymbol{\eta}_i \,|\, \{\boldsymbol{\eta}_j : j \in A\}) < 1$,*
2. *for any $i \in [L]$ and $A \subseteq [L] \backslash \{i\}$, a maximizer of $\kappa^*(Q_i \,|\, \{Q_j : j \in A\})$ exists, and*
3. *$\boldsymbol{\gamma} \in \Delta_L$ is a maximizer to $\kappa^*(\boldsymbol{\eta}_i \,|\, \{\boldsymbol{\eta}_j : j \in A\})$ if and only if $G = \boldsymbol{\gamma}^T \boldsymbol{P}$ is a maximizer to $\kappa^*(Q_i \,|\, \{Q_j : j \in A\})$.*

In words, this proposition says that the optimization problem given by $\kappa^*(Q_i \,|\, \{Q_j : j \in A\})$ is equivalent to the optimization problem given by $\kappa^*(\boldsymbol{\eta}_i \,|\, \{\boldsymbol{\eta}_j : j \in A\})$. *Thus, joint irreducibility of $P_1, \ldots, P_L$ and linear independence of the mixture proportions enable a reduction of each of the three problems to a geometric problem where the goal is to recover the vertices of a simplex by applying κ^* to points (i.e., the mixture proportions) in the simplex.* This makes the figures below valid for general distributions (see Figures 5.1, 5.2, 5.3, and 5.4).

5.4.3 Multiclass Classification with Label Noise

Our algorithm for multiclass classification with label noise is by far the simplest of the three. It simply computes a maximizer of $\kappa^*(\tilde{P}_i \,|\, \{\tilde{P}_j : j \neq i\})$ for every $i \in [L]$.

Theorem 25. *Let $P_1, \ldots, P_L$ be jointly irreducible and $\boldsymbol{\Pi}$ satisfy **(B1)**. Then, Multiclass($\tilde{P}_1, \ldots, \tilde{P}_L$) returns $\boldsymbol{Q} \in \mathcal{P}^L$ such that $\boldsymbol{Q} = \boldsymbol{P}$.*

Algorithm 18 Multiclass($\tilde{P}_1, \ldots, \tilde{P}_L$)

1: **for** $i = 1, \ldots, L$ **do**

2: $\quad Q_i \longleftarrow \text{MultiResidue}(\tilde{P}_i \,|\, \{\tilde{P}_j : j \neq i\})$

3: **end for**

4: **return** $(Q_1, \ldots, Q_L)^T$

The proof of this result has two main ideas. First, it applies the one-to-one correspondence established in Proposition 6 between the maximizers of $\kappa^*(\tilde{P}_i \,|\, \{\tilde{P}_j : j \neq i\})$ and the maximizers of $\kappa^*(\boldsymbol{\pi}_i \,|\, \{\boldsymbol{\pi}_j : j \neq i\})$.

Second, the proof shows that $\kappa^*(\boldsymbol{\pi}_i \,|\, \{\boldsymbol{\pi}_j : j \neq i\})$ is well-behaved in the sense that the residue of $\boldsymbol{\pi}_i$ wrt $\{\boldsymbol{\pi}_j : j \neq i\}$ is $\boldsymbol{e}_i$. The key idea is encapsulated in the following Lemma from Blanchard and Scott (2014).

Lemma 29. (Blanchard and Scott, 2014) *The following conditions on $\boldsymbol{\pi}_1, \ldots, \boldsymbol{\pi}_L$ are equivalent:*

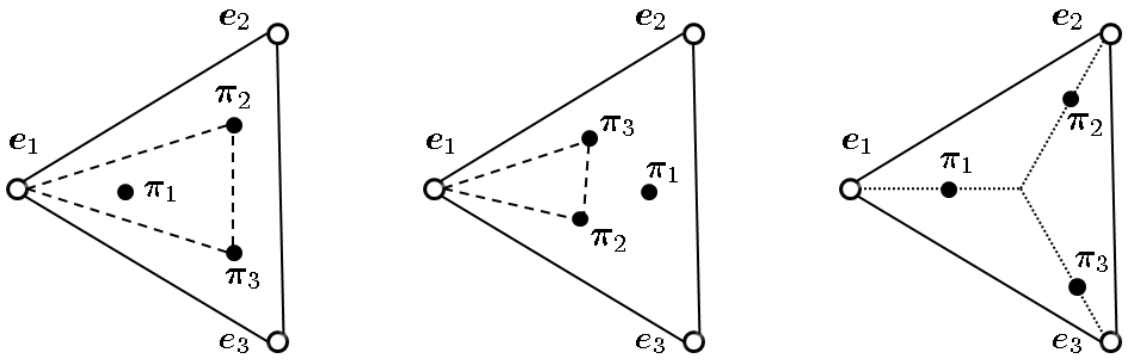
1. *For each i, the residue of $\boldsymbol{\pi}_i$ with respect to $\{\boldsymbol{\pi}_j, j \neq i\}$ is $\boldsymbol{e}_i$.*

2. *For every i there exists a decomposition $\boldsymbol{\pi}_i = \kappa_i \boldsymbol{e}_i + (1 - \kappa_i)\boldsymbol{\pi}'_i$ where $\kappa_i > 0$ and $\boldsymbol{\pi}'_i$ is a convex combination of $\boldsymbol{\pi}_j$ for $j \neq i$.*

3. *$\boldsymbol{\Pi}$ is invertible and $\boldsymbol{\Pi}^{-1}$ is a matrix with strictly positive diagonal entries and nonpositive off-diagonal entries.*

This lemma establishes that under **(B1)**, for each i, the residue of $\boldsymbol{\pi}_i$ with respect to $\{\boldsymbol{\pi}_j, j \neq i\}$ is $\boldsymbol{e}_i$. The main step in the proof of this Lemma is establishing that *3* implies *1*. The argument identifies the residue of $\boldsymbol{\pi}_i$ with respect to $\{\boldsymbol{\pi}_j\}_{j \neq i}$ by reformulating the linear program in $\kappa^*(\boldsymbol{\pi}_i \,|\, \{\boldsymbol{\pi}_j\}_{j \neq i})$ such that the objective is to maximize $\boldsymbol{e}_i^t \boldsymbol{\Pi}^{-1} \gamma$ subject to some appropriately defined constraint. By the structure of $\boldsymbol{\Pi}^{-1}$ assumed in **(B1)**, it follows that the $\gamma \in \Delta_L$ that maximizes this objective is $\boldsymbol{e}_i$, and it can further be shown that this maximizer satisfies the other constraints.

Thus, combining the above two ideas yields the result. In addition, Lemma 29 provides geometric intuition as to when **(B1)** is satisfied through condition *2*. Figure 5.1 illustrates



(i) (ii) (iii)

Figure 5.1: Illustration of the **(B1)** when there are $L = 3$ classes where $\boldsymbol{e}_i$ denotes the ith unit vector. Panel (i): Low noise, $\boldsymbol{\Pi}$ recoverable. Each $\boldsymbol{\pi}_l$ can be written as a convex combination of $\boldsymbol{e}_l$ and the other two $\boldsymbol{\pi}_j$ (with a positive weight on $\boldsymbol{e}_l$), depicted here for $l = 1$. Panel (ii): High noise, $\boldsymbol{\Pi}$ not recoverable. Panel (iii): The setting of "common background noise" described in the text.

the case $L = 3$. See Panel (i) for an example where condition (b) is satisfied and Panel (ii) for an example where (b) is not satisfied.

5.4.4 Demixing Mixed Membership Models

In this section, we assume that $M = L$; we consider the nonsquare case in the appendix. For certain simple cases of mixture proportions, a straightforward resampling strategy can be used to reduce the problem of demixing mixed membership models to multiclass classification with label noise. For example, suppose that there are $L = 3$ classes and

$$\boldsymbol{\Pi} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}. \tag{5.4}$$

Inspection shows that the inverse of $\boldsymbol{\Pi}$ does not satisfy the condition in **(B1)** and, therefore, one cannot simply apply Algorithm 18. A simple procedure to circumvent this issue is to resample from the contaminated distributions to obtain the following distributions:

$$\tilde{Q}_1 = \frac{1}{2}\tilde{P}_1 + \frac{1}{2}\tilde{P}_2,\ \tilde{Q}_2 = \frac{1}{2}\tilde{P}_1 + \frac{1}{2}\tilde{P}_3,\ \text{and}\ \tilde{Q}_3 = \frac{1}{2}\tilde{P}_2 + \frac{1}{2}\tilde{P}_3.$$

Then, it can be shown that the resulting mixing matrix

$$\tilde{\mathbf{\Pi}} = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}$$

associated with the $\tilde{Q}_i$s satisfies the conditions of Lemma 29 so that Multiclass$(\tilde{Q}_1, \tilde{Q}_2, \tilde{Q}_3)$ gives the desired solution. However, this approach breaks down for most possible mixing matrices. Thus, the challenge is to develop an algorithm that works for a large class of mixture proportions and does not rely on knowledge of the mixture proportions. To meet this challenge, we propose the Demix algorithm.

The Demix algorithm is recursive. Let $S_1, \ldots, S_K$ denote K contaminated distributions. In the base case, the algorithm takes as its input two contaminated distributions S_1 and S_2. It returns Residue$(S_1 \,|\, S_2)$ and Residue$(S_2 \,|\, S_1)$, which are a permutation of the two base distributions (see Figure 5.2). When $K > 2$, Demix uses a subroutine FindFace (see Algorithm 20) to find $K-1$ distributions $R_2, \ldots, R_K$ on the same $(K-1)$-face. FindFace iteratively generates candidates for distributions on the same $(K-1)$-face, which it tests using FaceTest (see Algorithm 21). FaceTest$(S_1, \ldots, S_{K-1})$ determines whether a set of distributions $S_1, \ldots, S_{K-1}$ are on the interior of the same face by using the two-sample κ^* operator; equivalently, it tests whether there exists a pair of distributions S_i and S_j such that S_i is irreducible with respect to S_j. Once Demix finds $K-1$ distributions $R_2, \ldots, R_K$ on the same $(K-1)$-face, it recursively applies Demix to $R_2, \ldots, R_K$ to obtain distributions $Q_1, \ldots, Q_{K-1}$ that are a permutation of $K-1$ of the base distributions. Subsequently, the algorithm computes a maximizer Q_K of $\kappa^*(\frac{1}{K}\sum_{i=1}^K S_i \,|\, Q_1, \ldots, Q_{K-1})$. Since $Q_1, \ldots, Q_{K-1}$ are a permutation of $K-1$ of the base distributions, the maximizer Q_K is guaranteed to be unique and to be the remaining base distribution (see Figure 5.3 for an execution of the algorithm).

A number of remarks are in order regarding the Demix algorithm. First, although we compute the residue of $\frac{1}{n}S_i + \frac{n-1}{n}Q$ wrt S_1 for each $i \neq 1$, there is nothing special about the distribution S_1. We could replace S_1 with any S_j where $j \in [K]$, provided that we adjust the rest of the algorithm accordingly. Second, we can replace the sequence $\{\frac{n-1}{n}\}_{n=1}^{\infty}$ with any sequence $\alpha_n \nearrow 1$. Finally, we could replace line 7 with the following sequence of steps: for

Algorithm 19 Demix$(S_1, \ldots, S_K)$

Input: $S_1, \ldots, S_K$ are distributions

if $K = 2$ **then**
  return $(\text{Residue}(S_1 \,|\, S_2), \text{Residue}(S_2 \,|\, S_1))^T$
else
  $(R_2, \ldots, R_K)^T \longleftarrow \text{FindFace}(S_1, \ldots, S_K)$
  $(Q_1, \ldots, Q_{K-1})^T \longleftarrow \text{Demix}(R_2, \ldots, R_K)$
  $Q_K \longleftarrow \frac{1}{K} \sum_{i=1}^{K} S_i$
  $Q_K \longleftarrow \text{MultiResidue}(Q_K \,|\, Q_1, \ldots, Q_{K-1})$
  return $(Q_1, \ldots, Q_K)^T$
end if

Algorithm 20 FindFace$(S_1, \ldots, S_K)$

Input: $S_1, \ldots, S_K$ are distributions

$Q \longleftarrow$ uniformly distributed element in $\text{conv}(S_2, \ldots, S_K)$
for $n = 1, 2, \ldots$ **do**
  Set $R_i \longleftarrow \text{Residue}(\frac{1}{n} S_i + \frac{n-1}{n} Q \,|\, S_1)$ for all $i \in \{2, \ldots, K\}$
  if $\text{FaceTest}(R_2, \ldots, R_K)$ **then**
    return $(R_2, \ldots, R_K)^T$
  end if
end for

$i = 1, \ldots, K-1$, compute $Q_K \longleftarrow \text{Residue}(Q_K \,|\, Q_i)$ (see Lemma 41). Then, the algorithm would only use the two-sample κ^* operator. We use such an algorithm in the finite-sample setting.

We also remark that a simplified version of Demix solves the demixing mixed membership models problem if we assume **(B1)** from the multiclass label noise setting. In that case, finding $L - 1$ distributions on the same face can be accomplished by simply computing $Q_i \longleftarrow \text{MultiResidue}(\tilde{P}_i \,|\, \{\tilde{P}_j\}_{j \neq i})$ for $i = 2, \ldots, L$. Indeed, then, each Q_i is equal to P_i and P_1 can be obtained by computing $\text{MultiResidue}(\tilde{P}_1 \,|\, \{Q_j\}_{j=2,\ldots,L})$.

We establish the following theorem.

Algorithm 21 FaceTest$(S_1, \ldots, S_K)$

Set $\boldsymbol{Z}_{i,j} \coloneqq \mathbf{1}\{\kappa^*(S_i \,|\, S_j) > 0\}$ for all i and j
if $\boldsymbol{Z}$ has a zero off-diagonal entry **then**
  return 0
else
  return 1
end if

Theorem 26. *Let $P_1, \ldots, P_L$ be jointly irreducible and $\boldsymbol{\Pi}$ have full column rank. Then, with probability* 1*, Demix($\tilde{\boldsymbol{P}}$) returns a permutation of $\boldsymbol{P}$.*

We briefly sketch three key aspects of the proof. First, in the FindFace subroutine, sampling Q uniformly at random from $\mathrm{conv}(S_2, \ldots, S_K)$ ensures that w.p. 1 $\mathrm{Residue}(Q \,|\, S_1)$ is on the interior of a face of the simplex. Then, conditional on this event, we show that by a continuity property of $\mathrm{Residue}(\cdot \,|\, S_1)$ there is a large enough n such that $R_2, \ldots, R_K$ are on the same face of the simplex Δ_{K-1} (see panels (c) and (d) of Figure 5.3). Second, Proposition 11 in Appendix 5.10 establishes that the subroutine FaceTest$(R_2, \ldots, R_K)$ returns 1 if and only if $R_2, \ldots, R_K$ are on the same face of the simplex. Combining the above two observations implies that eventually FindFace$(S_1, \ldots, S_K)$ terminates at which point $\{R_k\}_{k \in [K] \setminus \{1\}} \subset \{P_k\}_{k \in [K] \setminus \{l\}}$ for some $l \in [K]$. The final key observation is that $\{R_k\}_{k \in [K] \setminus \{1\}}$ and $\{P_k\}_{k \in [K] \setminus \{l\}}$ form an instance of the demixing problem that satisfies the sufficient conditions **(A)** and **(B2)** (see Figure 5.2). Therefore, this instance can be solved recursively.

5.4.5 Classification with Partial Labels

As in the case of demixing mixed membership models, a simple resampling strategy works in certain nice settings of classification with partial labels. For example, consider an instance of classification with partial labels with the mixing matrix from equation (5.4). The resampling procedure that yields $\tilde{Q}_1, \tilde{Q}_2, \tilde{Q}_3$ (described in Section 5.4.4) also works here. Nevertheless, as in demixing mixed membership models, this approach does not meet our goal of an algorithm that solves a broad class of mixing matrices and partial labels.

Indeed, we observe that the partial labels do not provide enough information for choosing

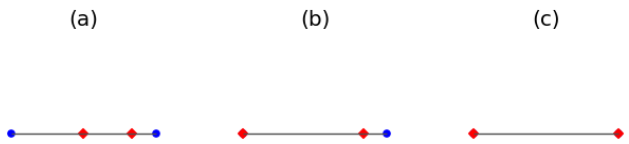


Figure 5.2: In (a), we consider a demixing problem where there are two classes and $M = L$ (the base case of Algorithm 19). The diamonds represent the mixture proportions of $\tilde{P}_1$ and $\tilde{P}_2$. The circles represent the base distributions. In (b), the residue of a contaminated distribution wrt the other contaminated distribution is computed (line 3), yielding a base distribution. In (c), the residue is computed again switching the roles of the contaminated distributions (line 4); this yields the remaining base distribution.

the resampling weights. Consider another instance of the problem with the same partial label matrix:

$$\mathbf{\Pi} = \begin{pmatrix} \frac{1}{10} & \frac{9}{10} & 0 \\ \frac{9}{10} & 0 & \frac{1}{10} \\ 0 & \frac{1}{10} & \frac{9}{10} \end{pmatrix}. \tag{5.5}$$

Applying the resampling approach to (5.5) can be shown to fail by observing that the inverse of the resampled mixing matrix does not satisfy condition *3* of Lemma 29. Thus, although the problem instances (5.4) and (5.5) have the same partial label matrix, the resampling procedure only works for one of these.

Next, we turn to presenting an algorithm that solves classification with partial labels for a wide class of mixing matrices and partial labels. We propose the PartialLabel algorithm (see Algorithm 22). PartialLabel proceeds by iteratively creating sets of candidate distributions $\boldsymbol{W} \coloneqq (W_1, \ldots, W_L)^T$ via the subroutine CreateCandidates (see Algorithm 23). Given each $\boldsymbol{W}$, it runs an algorithm VertexTest (see Algorithm 24) that uses $\tilde{\boldsymbol{P}}$ and the partial label matrix $\mathbf{\Pi}^+$ to determine whether $\boldsymbol{W}$ is a permutation of the base distributions $\boldsymbol{P}$. If $\boldsymbol{W}$ is a permutation of $\boldsymbol{P}$, VertexTest constructs the corresponding permutation matrix for relating these distributions. If not, it reports failure and the PartialLabel algorithm increments k

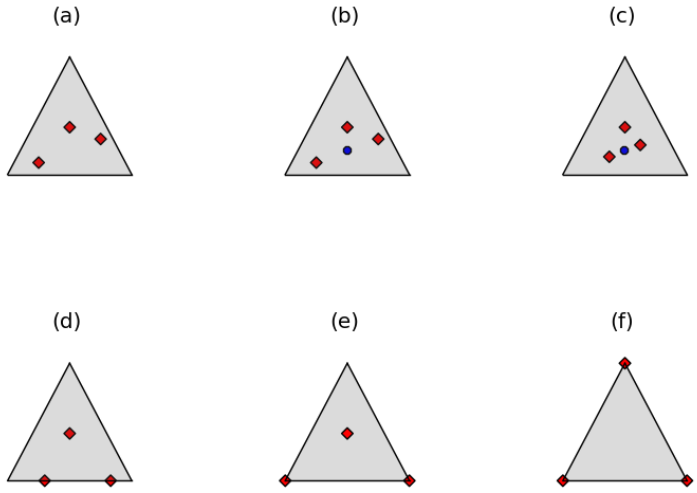


Figure 5.3: In (a), we consider a demixing problem where there are three classes and $M = L$. The diamonds represent the mixture proportions of $\tilde{P}_1, \tilde{P}_2$ and $\tilde{P}_3$. In (b), the blue circle is a random distribution chosen in the convex hull of two of the distributions (line 7). In (c), two of the distributions are resampled so that their residues wrt the other distribution are on the same face of the simplex (lines 12-15). In (d), these particular residues are computed (lines 12-15). In (e), two of the distributions are demixed (lines 3-5). In (f), the residue of the final distribution wrt the final two demixed distributions is computed to obtain the final demixing (line 18-21).

and finds another set of candidate distributions.

The VertexTest algorithm proceeds as follows on a vector of candidate distributions $\boldsymbol{Q} \coloneqq (Q_1, \ldots, Q_L)^T$. First, it determines whether there are two distinct distributions Q_i, Q_j such that Q_i is not irreducible wrt Q_j, in which case $\boldsymbol{Q}$ cannot be a permutation of $\boldsymbol{P}$. If there is such a pair, it reports failure. Otherwise, it forms the matrix $Z_{i,j} \coloneqq \mathbf{1}\{\kappa^*(\tilde{P}_i \,|\, Q_j) > 0\}$ and uses any algorithm that finds a permutation $\boldsymbol{C}$ (if it exists) of the columns of $\boldsymbol{Z}$ to match the columns of $\boldsymbol{\Pi}^+$. If such a permutation $\boldsymbol{C}$ exists, it returns $\boldsymbol{C}^T$ and, as we show in Lemma 35, $\boldsymbol{C}^T\boldsymbol{Q} = \boldsymbol{P}$; otherwise, VertexTest reports failure.

We remark that finding the permutation in line 6 of Algorithm 24 is not NP-hard. One algorithm (but most likely not the most efficient) proceeds as follows: define a total ordering on the columns of binary matrices. Sort the columns of $\boldsymbol{Z}$ and $\boldsymbol{\Pi}^+$ according to this total

Algorithm 22 PartialLabel($\mathbf{\Pi}^+, (\tilde{P}_1, \ldots, \tilde{P}_M)^T$)

1: **for** $i = 1, \ldots, L$ **do**
2: $\quad Q_i \longleftarrow$ uniformly random distribution in $\text{conv}(\tilde{P}_1, \ldots, \tilde{P}_M)$
3: **end for**
4: **for** $k = 2, 3, \ldots$ **do**
5: $\quad (W_1, \ldots, W_L)^T \longleftarrow$ GenerateCandidates$(k, (Q_1, \ldots, Q_L)^T)$
6: $\quad$ (FoundVertices, $\boldsymbol{C}$) $\longleftarrow$ VertexTest$(\mathbf{\Pi}^+, \tilde{P}_1, \ldots, \tilde{P}_M, W_1, \ldots, W_L)$
7: $\quad$ **if** FoundVertices **then**
8: $\quad\quad$ **return** $\boldsymbol{C}(W_1, \ldots, W_L)^T$
9: $\quad$ **end if**
10: **end for**

Algorithm 23 GenerateCandidates$(k, (Q_1, \ldots, Q_L)^T)$

1: Set $W_i \longleftarrow Q_i$ for all $i \in [K]$
2: **for** $i = 1, \ldots, L$ **do**
3: $\quad \bar{Q}_i \longleftarrow \frac{1}{L-1}[\sum_{j>i} Q_j + \sum_{j<i} W_j]$
4: $\quad W_i \longleftarrow$ MultiResidue$(\frac{1}{k}Q_i + (1 - \frac{1}{k})\bar{Q}_i \,|\, \{Q_j\}_{j>i} \cup \{W_j\}_{j<i})$
5: **end for**
6: **return** $(W_1, \ldots, W_L)^T$

ordering. Check whether the resulting matrices are equal.

The following theorem gives our identification result for classification with partial labels.

Theorem 27. *Suppose that $P_1, \ldots, P_L$ are jointly irreducible, $\mathbf{\Pi}$ has full column rank, and the columns of $\mathbf{\Pi}^+$ are unique. Then, PartialLabel$(\mathbf{\Pi}^+, (\tilde{\boldsymbol{P}})^T)$ returns $\boldsymbol{R} \in \mathcal{P}^L$ such that $\boldsymbol{R} = \boldsymbol{P}$.*

There are two key ideas to the proof of Theorem 27. First, the randomization in line 2 of Algorithm 22 ensures through a linear independence argument that with probability 1, the operation MultiResidue$(\frac{1}{k}Q_i + (1 - \frac{1}{k})\bar{Q}_i \,|\, \{Q_j\}_{j>i} \cup \{W_j\}_{j<i})$ in line 4 of Algorithm 23 is well-defined. Second, in the GenerateCandidates algorithm, let $Q_j = \boldsymbol{\tau}_j^T \boldsymbol{P}$ and $W_j = \boldsymbol{\gamma}_j^T P$. We make the simple observation that the affine hyperplane given by $\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_{i-1}, \boldsymbol{\tau}_{i+1}, \ldots, \boldsymbol{\tau}_L$ bisects Δ_L such that $\boldsymbol{\tau}_i$ and a nonempty subset of $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_L\} \backslash \{\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_{i-1}\}$ are in the same

Algorithm 24 VertexTest($\mathbf{\Pi}^+, (\tilde{P}_1, \ldots, \tilde{P}_M)^T, (Q_1, \ldots, Q_L)^T$)

Form the matrix $Y_{i,j} \coloneqq \mathbf{1}\{\kappa^*(Q_i \,|\, Q_j) > 0\}$
if $\boldsymbol{Y}$ has a non-zero off-diagonal entry **then**
  return $(0, \mathbf{0})$
end if
Form the matrix $Z_{i,j} \coloneqq \mathbf{1}\{\kappa^*(\tilde{P}_i \,|\, Q_j) > 0\}$
Use any algorithm that finds a permutation matrix $\boldsymbol{C}$ such that $\boldsymbol{Z}\boldsymbol{C} = \mathbf{\Pi}^+$ (if it exists)
if such a permutation matrix $\boldsymbol{C}$ exists **then**
  return $(1, \boldsymbol{C}^T)$
else
  return $(0, \mathbf{0})$
end if

halfspace. Using this observation, we show that for large enough k, W_i is one of the base distributions and is distinct from all W_j with $j < i$.

The VertexTest algorithm connects the demixing problem and classification with partial labels by showing that any algorithm that solves the demixing problem can be used as a subroutine to solve classification with partial labels. For example, consider the following algorithm for classification with partial labels. First, use the Demix algorithm to obtain a permutation $\boldsymbol{Q}$ of the base distributions $\boldsymbol{P}$. Second, use VertexTest to find the permutation matrix relating $\boldsymbol{Q}$ and $\boldsymbol{P}$. This alternate algorithm is the basis of our finite sample algorithm for classification with partial labels (see Section 5.5.3 for a more thorough discussion).

5.5 Estimators for the Finite Sample Setting

In this section, we develop the estimation theory to treat the three problems in the finite sample setting. Let $\mathcal{X} = \mathbb{R}^d$ be equipped with the standard Borel σ-algebra $\mathcal{C}$ and $\tilde{P}_1, \ldots, \tilde{P}_L$ be probability distributions on this space. Suppose that we observe for $i = 1, \ldots, L$,

$$X_1^i, \ldots, X_{n_i}^i \overset{i.i.d.}{\sim} \tilde{P}_i.$$

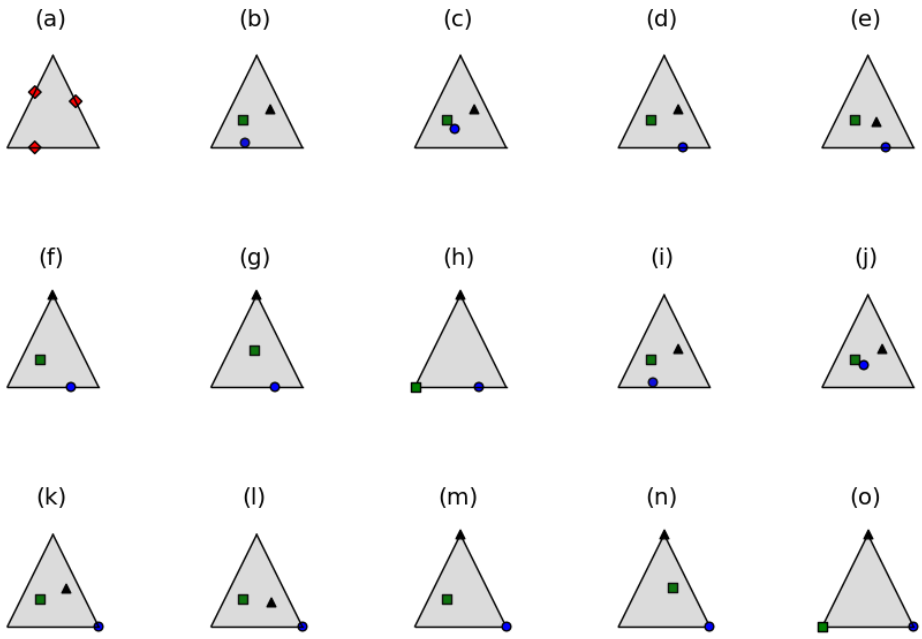


Figure 5.4: (a) depicts an instance of the partial label problem where there are $L = 3$ classes, $M = 3$ partial labels, and each partial label only contains two of the classes. In (a), the red diamonds represent the mixture proportions of the distributions $\tilde{P}_1, \tilde{P}_2, \tilde{P}_3$. In (b), three distributions Q_1, Q_2, Q_3 are sampled uniformly randomly from the convex hull of $\tilde{P}_1, \tilde{P}_2, \tilde{P}_3$; the blue circle, black triangle, and green square represent their mixture proportions. Figures (c)-(h) show how the algorithm generates a set of candidate distributions $(W_1^{(2)}, W_2^{(2)}, W_3^{(2)})^T$ with $k = 2$. In (h), PartialLabel runs VertexTest on $(W_1^{(2)}, W_2^{(2)}, W_3^{(2)})^T$ and determines that $(W_1^{(2)}, W_2^{(2)}, W_3^{(2)})^T$ is not a permutation of $(P_1, P_2, P_3)^T$. In (i)-(o), PartialLabel begins again with Q_1, Q_2, Q_3 and executes the same series of steps with $k = 3$, generating $(W_1^{(3)}, W_2^{(3)}, W_3^{(3)})^T$. In (o), it runs VertexTest on $(W_1^{(3)}, W_2^{(3)}, W_3^{(3)})^T$ and determines that $(W_1^{(3)}, W_2^{(3)}, W_3^{(3)})^T$ is a permutation of $(P_1, P_2, P_3)^T$.

Let $\mathcal{E}$ be any Vapnik-Chervonenkis (VC) class with VC-dimension $V < \infty$, containing the set of all open balls, all open rectangles, or some other collection of sets that generates the Borel σ-algebra $\mathcal{C}$. For example, $\mathcal{E}$ could be the set of all open balls wrt the Euclidean distance, in which case $V = d + 1$. Define $\epsilon_i(\delta_i) \equiv 3\sqrt{\frac{V \log(n_i+1) - \log(\delta_i/2)}{n_i}}$ for $i = 1, \ldots, L$. Our estimators

are based on the VC inequality (Devroye et al., 1996). This inequality says that for each $i \in [L]$, and $\delta > 0$, the following holds with probability at least $1 - \delta$:

$$\sup_{E \in \mathcal{E}} |\tilde{P}_i(E) - \tilde{P}_i^\dagger(E)| \leqslant \epsilon_i(\delta)$$

where the *empirical distribution* is given by $\tilde{P}_i^\dagger(E) = \frac{1}{n_i} \sum_{j=1}^{n_i} \mathbf{1}\{X_j^i \in E\}$.

5.5.1 Multiclass Classification with Label Noise

Let $F_0^\dagger, \ldots, F_M^\dagger$ denote the empirical distributions based on i.i.d. random samples from respective distributions $F_0, \ldots, F_M$. We introduce the following estimator of the multi-sample κ^*:

$$\widehat{\kappa}(F_0^\dagger \,|\, F_1^\dagger, \ldots, F_M^\dagger) = \max_{\boldsymbol{\mu} \in \Delta_M} \inf_{E \in \mathcal{E}} \frac{F_0^\dagger(E) + \epsilon_0(\frac{1}{n_0})}{(\sum_{i=1}^M \mu_i F_i^\dagger(E) - \sum_i \mu_i \epsilon_i(\frac{1}{n_i}))_+} \tag{5.6}$$

where the ratio is defined to be ∞ when the denominator is zero. This estimator arises from applying the VC inequality to the following expression:

$$\kappa^*(F_0 \,|\, F_1, \ldots, F_M) = \max_{\boldsymbol{\mu} \in \Delta_M} \kappa^*(F_0 \,|\, \sum_{i=1}^M \mu_i F_i) = \max_{\boldsymbol{\mu} \in \Delta_M} \inf_{E \in \mathcal{E}, \sum_{i=1}^M \mu_i F_i(E) > 0} \frac{F_0(E)}{\sum_{i=1}^M \mu_i F_i(E)},$$

where the last equality uses Proposition 5. Let $\widehat{\boldsymbol{\mu}}$ denote a point where the maximum is achieved in (5.6). Then, $\widehat{\boldsymbol{\nu}} \coloneqq \widehat{\kappa}\widehat{\boldsymbol{\mu}}$ estimates the vector $(\nu_1, \ldots, \nu_M)$ attaining the maximum in (5.3). See Proposition 2 of Blanchard and Scott (2014) to find a proof that the proposed estimator is consistent.

Based on this estimator, we introduce estimators that under the assumptions of Theorem 25 converge to the base distributions uniformly in probability. Let $\boldsymbol{n} \equiv (n_1, \ldots, n_L)$; we write $\boldsymbol{n} \longrightarrow \infty$ to indicate that $\min_i n_i \longrightarrow \infty$.

Theorem 28. *Let $(\widehat{\nu}_{i,j})_{j \neq i}$ be a vector attaining the maximum in the definition of $\widehat{\kappa}_i \coloneqq \widehat{\kappa}(\tilde{P}_i^\dagger \,|\, \{\tilde{P}_j^\dagger : j \neq i\})$ and*

$$\widehat{Q}_i = \frac{\tilde{P}_i^\dagger - \sum_{j \neq i} \widehat{\nu}_{i,j} \tilde{P}_j^\dagger}{1 - \widehat{\kappa}_i}.$$

Then, under the assumptions of Theorem 25, $\forall i = 1, \ldots, L$, $\sup_{E \in \mathcal{E}} |\widehat{Q}_i(E) - P_i(E)| \xrightarrow{i.p.} 0$ as $\boldsymbol{n} \longrightarrow \infty$.

Algorithm 25 ResidueHat($\widehat{F} \,|\, \widehat{H}$)

Input: $\widehat{F}, \widehat{H}$ are estimates of F, H

1: $\widehat{\kappa} \longleftarrow \widehat{\kappa}(\widehat{F} \,|\, \widehat{H})$

2: **return** $\frac{\widehat{F} - \widehat{\kappa}(1-\widehat{H})}{1-\widehat{\kappa}}$

5.5.2 Demixing Mixed Membership Models

In this section, we develop a novel estimator that can be used to extend the Demix algorithm to the finite sample case. Uniform convergence results typically assume access to i.i.d. samples. The challenge of developing an estimator for Demix is that because of the recursive nature of the Demix algorithm, we cannot assume access to i.i.d. samples to estimate every distribution that arises. Nonetheless, we show that uniform convergence of distributions propagates through the algorithm if we employ an estimator of κ^* with a known rate of convergence.

Let $\widehat{F}$ and $\widehat{H}$ be estimates of distributions F and H, respectively. We introduce the following estimator:

$$\widehat{\kappa}(\widehat{F} \,|\, \widehat{H}) = \inf_{E \in \mathcal{E}} \frac{\widehat{F}(E) + \gamma_{\boldsymbol{n}}}{(\widehat{H}(E) - \gamma_{\boldsymbol{n}})_+}$$

where $\gamma_{\boldsymbol{n}} = \sum_{i=1}^{L} \epsilon_i(\frac{1}{n_i})$. Our estimator is closely related to the estimator from Blanchard et al. (2010): if $\widehat{F}$ and $\widehat{H}$ are empirical distributions, e.g., $\widehat{F} = \tilde{P}_i^{\dagger}$ and $\widehat{H} = \tilde{P}_j^{\dagger}$, then their estimator for $\kappa^*(F \,|\, H)$ is $\inf_{E \in \mathcal{E}} \frac{\widehat{F}(E) + \epsilon_i(\frac{1}{n_i})}{(\widehat{H}(E) - \epsilon_j(\frac{1}{n_j}))_+}$. Note that our proofs only require that $\gamma_{\boldsymbol{n}}$ include the terms $\epsilon_i(\frac{1}{n_i})$ corresponding to $\tilde{P}_i$ that the estimators $\widehat{F}$ and $\widehat{H}$ use samples from; to simplify presentation, however, we include all the terms, which leads to bounds that are looser by only a constant factor.

Based on the estimator $\widehat{\kappa}$, we introduce the following estimator of the residue of F wrt H.

Definition 13. *Let $\widehat{F}$ and $\widehat{H}$ be estimators of F and H, respectively, where $F \neq H$ and let $G \longleftarrow$ Residue($F \,|\, H$) and $\widehat{G} \longleftarrow$ ResidueHat($\widehat{F} \,|\, \widehat{H}$). We call $\widehat{G}$ a* ResidueHat estimator *of order $k \geqslant 1$ if (i) $F, H \in \mathrm{conv}(P_1, \ldots, P_L)$, and (ii) at least one of $\widehat{F}$ and $\widehat{H}$ is a ResidueHat estimator of order $k-1$ and the other is either an empirical distribution or a ResidueHat*

Algorithm 26 DemixHat$(\widehat{S}_1, \ldots, \widehat{S}_K \,|\, \epsilon)$

Input: $\widehat{S}_1, \ldots, \widehat{S}_K$ are ResidueHat estimates

if $K = 2$ **then**
  return $(\text{ResidueHat}(\widehat{S}_1 \,|\, \widehat{S}_2), \text{ResidueHat}(\widehat{S}_2 \,|\, \widehat{S}_1))^T$
else
  $(R_2, \ldots, R_K)^T \longleftarrow \text{FindFaceHat}(\widehat{S}_1, \ldots, \widehat{S}_K \,|\, \epsilon)$
  $(\widehat{Q}_1, \cdots, \widehat{Q}_{K-1})^T \longleftarrow \text{DemixHat}(\widehat{R}_2, \cdots, \widehat{R}_K)$
  $\widehat{Q}_K \longleftarrow \frac{1}{K} \sum_{i=1}^{K} \widehat{S}_i$
  for $i = 1, \ldots, K-1$ **do**
    $\widehat{Q}_K \longleftarrow \text{ResidueHat}(\widehat{Q}_K \,|\, \widehat{Q}_i)$
  end for
  return $(\widehat{Q}_1, \cdots, \widehat{Q}_K)^T$
end if

estimator of order less than or equal to $k-1$. *We call* $\widehat{G}$ *a* ResidueHat estimator *of order* 0 *if (i) holds, and* $\widehat{F}$ *and* $\widehat{H}$ *are empirical distributions.*

Note that the above definition is recursive and matches the recursive structure of the Demix algorithm. We suppress the qualifier "of order k" when it is not relevant.

To use ResidueHat estimators to estimate the P_is, we build on the rate of convergence result from Scott (2015). In Scott (2015), a rate of convergence was established for an estimator of κ^* using empirical distributions; we extend these results to our setting of recursive estimators and achieve the same rate of convergence. To ensure that this rate of convergence holds for every estimate in our algorithm, we introduce the following condition.

(A″) $P_1, \ldots, P_L$ are such that $\forall i$ $\text{supp}(P_i) \nsubseteq \cup_{j \neq i} \text{supp}(P_j)$.

Note that this assumption implies joint irreducibility and is a natural generalization of the separability assumption.

The following result establishes sufficient conditions under which ResidueHat estimates converge uniformly.

Algorithm 27 FindFaceHat($\widehat{S}_1, \ldots, \widehat{S}_K \,|\, \epsilon$)

Input: $\widehat{S}_1, \ldots, \widehat{S}_K$ are ResidueHat estimates

1: $\widehat{Q} \longleftarrow$ uniformly distributed random element from $\text{conv}(\widehat{S}_2, \ldots, \widehat{S}_K)$
2: **for** $n = 2, 3, \ldots$ **do**
3: Set $\widehat{R}_i \longleftarrow \text{ResidueHat}(\frac{1}{n}\widehat{S}_i + \frac{n-1}{n}\widehat{Q}) \,|\, \widehat{S}_1)$ for all $i \in \{2, \ldots, K\}$
4: **if** FaceTestHat($\widehat{R}_2, \cdots, \widehat{R}_K \,|\, \epsilon$) **then**
5: **return** $(\widehat{R}_2, \cdots, \widehat{Q}_K)^T$
6: **end if**
7: **end for**

Algorithm 28 FaceTestHat($\widehat{Q}_1, \cdots, \widehat{Q}_K \,|\, \epsilon$)

1: Set $\boldsymbol{Z}_{i,j} \coloneqq \mathbf{1}\{\widehat{\kappa}(\widehat{Q}_i \,|\, \widehat{Q}_j) > \epsilon\}$ for $i \neq j$
2: **if** $\boldsymbol{Z}$ has a zero off-diagonal entry **then**
3: **return** 0
4: **else**
5: **return** 1
6: **end if**

Proposition 7. *If $P_1, \ldots, P_L$ satisfy **(A″)** and $\widehat{G}$ is a ResidueHat estimator of a distribution $G \in \text{conv}(P_1, \ldots, P_L)$, then $\sup_{E \in \mathcal{E}} |\widehat{G}(E) - G(E)| \xrightarrow{i.p.} 0$ as $\boldsymbol{n} \longrightarrow \infty$.*

Based on the ResidueHat estimators, we introduce an empirical version of the Demix algorithm—DemixHat (see Algorithm 26). The main differences are that *(i)* we replace the Residue function with the ResidueHat function, *(ii)* we replace line 7 in the Demix algorithm with a sequence of applications of the two-sample κ^* operator, as mentioned just before Theorem 26, and *(iii)* DemixHat requires specification of a hyperparameter $\epsilon \in (0, 1)$. We replace the multi-sample κ^* with the two-sample κ^* because there is no known estimator with a rate of convergence for the multi-sample κ^*, and the rate of convergence is essential to our consistency proof. The hyperparameter ϵ gives a tradeoff between runtime and accuracy. The runtime increases with increasing ϵ, but the amount of uncertainty about whether DemixHat executes successfully decreases with increasing ϵ.

Algorithm 29 PartialLabelHat($\mathbf{\Pi}^+, (\tilde{P}_1^\dagger, \ldots, \tilde{P}_M^\dagger)^T \,|\, \epsilon$)

1: $(\widehat{Q}_1, \ldots, \widehat{Q}_L)^T \longleftarrow$ DemixHat($\tilde{P}_1^\dagger, \ldots, \tilde{P}_M^\dagger \,|\, \epsilon$)

2: (FoundVertices, $\boldsymbol{C}$) $\longleftarrow$ VertexTestHat($\mathbf{\Pi}^+, (\tilde{P}_1^\dagger, \ldots, \tilde{P}_M^\dagger)^T, (\widehat{Q}_1, \ldots, \widehat{Q}_L)^T$)

3: **return** $\boldsymbol{C}(\widehat{Q}_1, \ldots, \widehat{Q}_L)^T$

We now state our main estimation result.

Theorem 29. *Let $\delta > 0$ and $\epsilon \in (0,1)$. Suppose that $P_1, \ldots, P_L$ satisfy **(A″)** and $\mathbf{\Pi}$ has full rank. Then, with probability tending to 1 as $\boldsymbol{n} \longrightarrow \infty$, DemixHat($\tilde{P}_1^\dagger, \ldots, \tilde{P}_L^\dagger \,|\, \epsilon$) returns $(\widehat{Q}_1, \ldots, \widehat{Q}_L)$ for which there exists a permutation $\sigma : [L] \longrightarrow [L]$ such that for every $i \in [L]$,*

$$\sup_{E \in \mathcal{E}} |\widehat{Q}_i(E) - P_{\sigma(i)}(E)| < \delta.$$

5.5.3 Classification with Partial Labels

In this section, we present a finite sample algorithm for the decontamination of a partial label model (see Algorithm 29). This algorithm is based on a different approach from PartialLabel (Algorithm 22): it combines DemixHat with an empirical version of the VertexTest algorithm (see Algorithm 32). The reason for this is that we have an estimator with a rate of convergence for the two-sample κ^*, whereas there is no known estimator with a rate of convergence for the multi-sample κ^*. We leverage this rate of convergence to prove the consistency of our algorithm.

We make an assumption that simplifies our algorithm: $\mathbf{\Pi}^+$ satisfies

(D) there does not exist $i, j \in [L]$ such that $\mathbf{\Pi}^+_{i,:} = \boldsymbol{e}_j^T$.

In words, this says that there is no contaminated distribution $\tilde{P}_i$ and base distribution P_j such that $\tilde{P}_i = P_j$. We emphasize that we make this assumption only to simplify the presentation and development of the algorithm; one can reduce any instance of a partial label model satisfying **(B3)** and **(A)** to an instance of a partial label model that also satisfies **(D)**. We defer the sketch of this reduction to Section 5.11.3.

We now state our main estimation result for classification with partial labels.

Theorem 30. *Let $\delta > 0$ and $\epsilon \in (0,1)$. Suppose that $P_1, \ldots, P_L$ satisfy **(A″)**, $\mathbf{\Pi}$ has full rank, the columns of $\mathbf{\Pi}^+$ are unique and $\mathbf{\Pi}^+$ satisfies **(D)**. Then, with probability tending to 1 as $\boldsymbol{n} \longrightarrow \infty$, PartialLabelHat($\mathbf{\Pi}^+, (\tilde{P}_1^\dagger, \ldots, \tilde{P}_M^\dagger)^T \,|\, \epsilon$) returns $(\widehat{Q}_1, \ldots, \widehat{Q}_L)^T$ such that for every $i \in [L]$,*

$$\sup_{E \in \mathcal{E}} |\widehat{Q}_i(E) - P_i(E)| < \delta.$$

5.5.4 Sieve Estimators

In the preceding, we have assumed a fixed VC class to simplify the presentation. However, these results easily extend to the setting where $\mathcal{E} = \mathcal{E}_k$ and $k \longrightarrow \infty$ at a suitable rate depending on the growth of the VC dimensions V_k. This allows for the P_is to be estimated uniformly on arbitrarily complex events, e.g., $\mathcal{E}_k$ is the set of unions of k open balls.

5.6 Discussion

In this paper, we have studied the problem of how to recover the base distributions $\boldsymbol{P}$ from the contaminated distributions $\tilde{\boldsymbol{P}}$ without knowledge of the mixing matrix $\mathbf{\Pi}$. We used a common set of concepts and techniques to solve three popular machine learning problems that arise in this setting: multiclass classification with label noise, demixing of mixed membership models, and classification with partial labels. Our technical contributions include: (i) We provide sufficient and sometimes necessary conditions for identifiability for all three problems. (ii) We give nonparametric algorithms for the infinite and finite sample settings. (iii) We provide a new estimator for iterative applications of κ^* that is of independent interest. (iv) Finally, our work provides a novel geometric perspective on each of the three problems.

Our results improve on what was previously known for all three problems. For multiclass classification with label noise and unknown $\mathbf{\Pi}$, previous work had only considered the case $M = L = 2$. Our work achieves a generalization to arbitrarily many distributions. For demixing of mixed membership models, previous algorithms with theoretical guarantees required a finite sample space. Our work allows for a much more general set of distributions. Finally, for classification with partial labels, previous work on learnability assumed

the realizable case (non-overlapping $P_1, \ldots, P_L$) and assumed strong conditions on label co-occurence in partial labels. Our analysis covers the agnostic case and a much wider set of partial labels.

Our work has also highlighted the advantages and disadvantages associated with the two-sample κ^* operator and multi-sample κ^* operator, respectively. Algorithms that only use the two-sample κ^* operator have the following two advantages: (i) the geometry of the two-sample κ^* operator is simpler than the geometry of the multi-sample κ^* operator and, as such, can be more tractable. Indeed, in recent years, several practical algorithms for estimating the two-sample κ^* have been developed (see Jain et al. (2016) and references therein). (ii) We have estimators with established rates of convergence for the two-sample κ^* operator, but not for the multi-sample κ^* operator. On the other hand, algorithms that use the multi-sample κ^* operator have fewer steps.

The aims of this work are mainly theoretical, but we believe that our work can inform practical algorithms. First, we note that while we have emphasized recovery of $\boldsymbol{P}$, another interpretation is that our work deals with estimating $\mathbf{\Pi}$. One can then plug our estimate of $\mathbf{\Pi}$ into corrected losses for multiclass classification with label noise and classification with partial labels that require knowledge of $\mathbf{\Pi}$ (Cid-Sueiro, 2012; Menon et al., 2015b; van Rooyen and Williamson, 2015; Patrini et al., 2017). Thus, in general, our work can be applied in this two-stage approach. Second, when L or M are small, the Demix algorithm is practical. For example, Metodiev and Thaler (2018b) apply the Demix algorithm to a high-energy physics application where $L = M = 2$. It is of interest to examine more generally whether variants of Demix work when M or L are small. Third, we conjecture that our analysis suggests novel principles for designing algorithms. For example, an alternative approach to the three problems in question is to embed the contaminated distributions in a reproducing kernel Hilbert Space and to estimate the $\mathbf{\Pi}$ matrix by setting up an optimization problem (e.g., see Ramaswamy et al. (2016) for the setting where there are two distributions). One of our necessary conditions, maximality (see Appendix 5.9), suggests formulating the optimization problem to search for base distributions that *(i)* explain the observed contaminated distributions and *(ii)* whose convex hull is as large as possible. In this way, we believe that our general treatment of these three problems that arise in mutual contamination models

provides intuitions that could be useful for designing new algorithms.

Although our experiments in Appendix 5.13 suggest that joint irreducibility of the base distributions is a reasonable assumption, it is nevertheless worthwhile to consider what questions arise if the base distributions are not exactly jointly irreducible. We see two possible research directions. First, one could perform a stability analysis: when the base distributions are not jointly irreducible, but are nearly jointly irreducible (in some sense that would need to be defined precisely), does the estimate of $\mathbf{\Pi}$ remain close to the true $\mathbf{\Pi}$? A second research question is to reinterpret the problem of demixing of mixed membership models as a dimensionality reduction problem. That is, given a large set of distributions, one could seek to represent them as convex combinations of a small set of irreducible base distributions. Then, the challenge would be to define an appropriate measure of approximation quality and to determine whether our approach could be useful for designing a consistent algorithm for the best approximation.

5.7 Outline of Chapter Appendix

To begin, we introduce additional notation for the appendices. In Section 5.9, we discuss how strong our sufficient conditions are and present factorization results that suggest that they are reasonable. In Section 5.10, we give our identifiability analysis of demixing mixed membership models and classification with partial labels. In Section 5.11, we prove our results on the ResidueHat estimator, as well as the finite sample algorithms for demixing mixed membership models and classification with partial labels. In Section 5.12, we state some lemmas from related papers that we use in our arguments.

5.8 Notation for Appendices

Let $A \subset \mathbb{R}^d$ be a set. Let $\operatorname{aff} A$ denote the affine hull of A, i.e., $\operatorname{aff} A = \{\sum_{i=1}^K \theta_i \boldsymbol{x}_i | \boldsymbol{x}_1, \ldots, \boldsymbol{x}_K \in A, \sum_{i=1}^K \theta_i = 1\}$. A° denotes the relative interior of A, i.e., $A^\circ = \{x \in A | B(x, r) \cap \operatorname{aff} A \subseteq A \text{ for some } r > 0\}$. Then, ∂A denotes the relative boundary of A, i.e., $\partial A = A \backslash A^\circ$. In addition, let $\|\cdot\|$ denote an arbitrary finite-dimensional norm on $\mathbb{R}^L$. For two vectors, $\boldsymbol{x}, \boldsymbol{y} \in$

$\mathbb{R}^K$, define

$$\min(\boldsymbol{x}^T, \boldsymbol{y}^T) = (\min(x_1, y_1), \ldots, \min(x_K, y_K)).$$

$\boldsymbol{x} \geqslant \boldsymbol{y}$ means $x_i \geqslant y_i \ \forall i \in [K]$.

For distributions $Q_1, \ldots, Q_k$, we use $\text{conv}(Q_1, \ldots, Q_K)^\circ$ to denote the relative interior of their convex hull and have that

$$\text{conv}(Q_1, \ldots, Q_K)^\circ = \{\sum_{i=1}^K \alpha_i Q_i : \alpha_i > 0, \sum_{i=1}^K \alpha_i = 1\}.$$

Note that when $Q_1, \ldots, Q_K$ are discrete distributions, this definition coincides with the definition of the relative interior of a set of Euclidean vectors.

We use the following affine mapping throughout the paper: $m_\nu(\boldsymbol{x}, \boldsymbol{y}) = (1-\nu)\boldsymbol{x} + \nu\boldsymbol{y}$ where $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^L$ and $\nu \in [0, 1]$. Overloading notation, when Q_1 and Q_2 are distributions, we define $m_\nu(Q_1, Q_2) = (1-\nu)Q_1 + \nu Q_2$. Note that if $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2 \in \Delta_L$ and $Q_1 = \boldsymbol{\eta}_1^T \boldsymbol{P}$ and $Q_2 = \boldsymbol{\eta}_2^T \boldsymbol{P}$, then $m_\nu(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2)$ is the mixture proportion for the distribution $m_\nu(Q_1, Q_2)$.

5.9 Factorization Results

In this section, we discuss whether our sufficient conditions are necessary. For the problems of demixing mixed membership models and classification with partial labels, we provide factorization results that suggest that our sufficient conditions are not much stronger than what is necessary.

5.9.1 Multiclass Classification with Label Noise

Our sufficient condition **(B1)** for multiclass classification with label noise is not necessary. Rather, **(B1)** is one of several possible sufficient conditions, and one that reflects a low noise assumption as illustrated in Figure 5.1. Consider the case $L = M = 2$ where **(B1)** is equivalent to $\pi_{1,2} + \pi_{2,1} < 1$. Recovery is still possible if $\pi_{1,2} + \pi_{2,1} > 1$ since one can simply swap $\tilde{P}_1$ and $\tilde{P}_2$ in a decontamination procedure. $\pi_{1,2} + \pi_{2,1} < 1$ is only necessary if one assumes that most of the training labels are correct, which is what $\pi_{1,2} + \pi_{2,1} < 1$ essentially says. For larger $L = M$, **(B1)** says in a sense that most of the data from $\tilde{P}_i$ come from P_i

for every i. Other sufficient conditions are possible (as in the binary case), but these would require at least one $\tilde{P}_i$ to contain a significant portion of some P_j, $j \neq i$. Regarding **(A)**, Blanchard et al. (2016) study the question of necessity for joint irreducibility in the case $L = M = 2$ and show that under mild assumptions on the decontamination procedure, joint irreducibility is necesssary.

5.9.2 Demixing Mixed Membership Models

Recall the definition of a factorization: $\tilde{\boldsymbol{P}}$ is factorizable if there exists $(\boldsymbol{\Pi}, \boldsymbol{P}) \in \Delta_L^M \times \mathcal{P}^L$ such that $\tilde{\boldsymbol{P}} = \boldsymbol{\Pi}\boldsymbol{P}$; we call $(\boldsymbol{\Pi}, \boldsymbol{P})$ a factorization of $\tilde{\boldsymbol{P}}$.

Our sufficient conditions are not much stronger than what is required by factorizations that satisfy the two forthcoming desirable properties.

Definition 14. *We say a factorization* $(\boldsymbol{\Pi}, \boldsymbol{P})$ *of* $\tilde{\boldsymbol{P}}$ *is* maximal ***(M)*** *iff for all factorizations* $(\boldsymbol{\Pi}', \boldsymbol{P}')$ *of* $\tilde{\boldsymbol{P}}$ *with* $\boldsymbol{P}' = (P'_1, \ldots, P'_L)^T \in \mathcal{P}^L$*, it holds that* $\{P'_1, \ldots, P'_L\} \subseteq \text{conv}(P_1, \ldots, P_L)$.

In words, $\boldsymbol{P} = (P_1, \ldots, P_L)^T$ is a maximal collection of base distributions if it is not possible to move any of the P_is outside of $\text{conv}(P_1, \ldots, P_L)$ and represent $\tilde{\boldsymbol{P}}$.

Definition 15. *We say a factorization* $(\boldsymbol{\Pi}, \boldsymbol{P})$ *of* $\tilde{\boldsymbol{P}}$ *is* linear ***(L)*** *iff* $\{P_1, \ldots, P_L\} \subseteq \text{span}(\tilde{P}_1, \ldots, \tilde{P}_M)$.

We believe that **(L)** is a reasonable requirement because it holds in the common situation in which there exist $\boldsymbol{\pi}_{i_1}, \ldots, \boldsymbol{\pi}_{i_L}$ that are linearly independent. Then for $I = \{i_1, \ldots, i_L\}$, we can write $\tilde{\boldsymbol{P}}_I = \boldsymbol{\Pi}_I \boldsymbol{P}$ where $\boldsymbol{\Pi}_I$ is the submatrix of $\boldsymbol{\Pi}$ containing only the rows indexed by I and $\tilde{\boldsymbol{P}}_I$ is similarly defined. Then, $\boldsymbol{\Pi}_I$ is invertible and $\boldsymbol{P} = \boldsymbol{\Pi}_I^{-1} \tilde{\boldsymbol{P}}_I$.

Factorizations that satisfy **(A)** and **(B2)** are maximal and linear.

Proposition 8. *Let* $(\boldsymbol{\Pi}, \boldsymbol{P})$ *be a factorization of* $\tilde{\boldsymbol{P}}$*. If* $(\boldsymbol{\Pi}, \boldsymbol{P})$ *satisfies* ***(A)*** *and* ***(B2)****, then* $(\boldsymbol{\Pi}, \boldsymbol{P})$ *satisfies* ***(M)*** *and* ***(L)****.*

Proof. We first show that $(\boldsymbol{\Pi}, \boldsymbol{P})$ satisfies **(L)**. By hypothesis, $P_1, \ldots, P_L$ are jointly irreducible. By Lemma 44, $P_1, \ldots, P_L$ are linearly independent. Since by hypothesis $\boldsymbol{\Pi}$ has full rank, there exist L rows in $\boldsymbol{\Pi}$, $\boldsymbol{\pi}_{i_1}, \ldots, \boldsymbol{\pi}_{i_L}$, that are linearly independent. By Lemma 44,

$\tilde{P}_{i_1}, \ldots, \tilde{P}_{i_L}$ are linearly independent. Since $\boldsymbol{\Pi P} = \tilde{\boldsymbol{P}}$, $\text{span}(\tilde{P}_1, \ldots, \tilde{P}_M) \subseteq \text{span}(P_1, \ldots, P_L)$. Since $\dim \text{span}(\tilde{P}_1, \ldots, \tilde{P}_M) \geqslant L$, we have $\text{span}(\tilde{P}_1, \ldots, \tilde{P}_M) = \text{span}(P_1, \ldots, P_L)$. Therefore, $(\boldsymbol{\Pi}, \boldsymbol{P})$ satisfies **(L)**.

Now, we show that $(\boldsymbol{\Pi}, \boldsymbol{P})$ satisfies **(M)**. Suppose that there is another solution $(\boldsymbol{\Pi'}, \boldsymbol{P'})$ with $\boldsymbol{P'} = (P'_1, \ldots, P'_L)^T$ such that $\boldsymbol{\Pi' P'} = \tilde{\boldsymbol{P}}$ and with some P'_i such that $P'_i \notin \text{conv}(P_1, \ldots, P_L)$. We claim that $P'_i \notin \text{span}(P_1, \ldots, P_L)$. Towards a contradiction, suppose that $P'_i \in \text{span}(P_1, \ldots, P_L)$ so that we can write $P'_i = \sum_{i=1}^{L} a_i P_i$. Then, at least one of the a_i is negative since, by assumption, $P'_i \notin \text{conv}(P_1, \ldots, P_L)$. But, by joint irreducibility of $P_1, \ldots, P_L$, P'_i is not a distribution, which is a contradiction. So, the claim follows. But, then, since $\text{span}(P_1, \ldots, P_L) = \text{span}(\tilde{P}_1, \ldots, \tilde{P}_M)$, we must have that $\text{span}(\tilde{P}_1, \ldots, \tilde{P}_M) \subseteq \text{span}(P'_1, \ldots, P'_{i-1}, P'_{i+1}, \ldots, P_L)$, which is impossible since $\dim \text{span}(\tilde{P}_1, \ldots, \tilde{P}_M) = L$. □

Maximal and linear factorizations imply conditions that are not much weaker than our sufficient conditions.

Theorem 31. *Let* $(\boldsymbol{\Pi}, \boldsymbol{P})$ *be a factorization of* $\tilde{\boldsymbol{P}}$. *If* $(\boldsymbol{\Pi}, \boldsymbol{P})$ *satisfies* ***(M)****, then*

(A′) $\forall i$, P_i *is irreducible with respect to every distribution in* $\text{conv}(\{P_j : j \neq i\})$.

If $(\boldsymbol{\Pi}, \boldsymbol{P})$ *satisfies* ***(L)****, then*

(B′) $\text{rank}(\boldsymbol{\Pi}) \geqslant \dim \text{span}(P_1, \ldots, P_L)$.

Proof. **(A′)** We prove the contrapositive. Suppose that there is some P_i and $Q \in \text{conv}(\{P_j : j \neq i\})$ with $Q = \sum_{j \neq i} \beta_j P_j$ such that P_i is not irreducible wrt Q. Then, there is some distribution G and $\gamma \in (0, 1]$ such that $P_i = \gamma Q + (1 - \gamma) G$.

Suppose $\gamma = 1$. Then, $P_i = Q \in \text{conv}(\{P_k : k \neq i\})$. But, then $\tilde{P}_1, \ldots, \tilde{P}_M \in \text{conv}(\{P_j : j \neq i\} \cup \{R\})$ for any distribution $R \notin \text{conv}(P_1, \ldots, P_L)$. This shows that $(\boldsymbol{\Pi}, \boldsymbol{P})$ does not satisfy **(M)**.

Therefore, assume that $\gamma \in (0, 1)$. Either $G \in \text{conv}(P_1, \ldots, P_L)$ or $G \notin \text{conv}(P_1, \ldots, P_L)$. Suppose that $G \in \text{conv}(P_1, \ldots, P_L)$. Then, there exist $\alpha_1, \ldots, \alpha_L$ all nonnegative and summing to 1 such that

$$P_i = \gamma Q + (1 - \gamma)(\alpha_1 P_1 + \ldots + \alpha_L P_L).$$

Therefore, $P_i \in \text{conv}(\{P_k : k \neq i\})$. Then, by the argument in the previous paragraph, $(\boldsymbol{\Pi}, \boldsymbol{P})$ does not satisfy **(M)**.

Now, suppose that $G \notin \text{conv}(P_1, \ldots, P_L)$. Since $P_i \in \text{conv}(G, Q)$ and $Q \in \text{conv}(\{P_j : j \neq i\})$, we have that $\text{conv}(\{P_j : j \neq i\} \cup \{G\}) \supset \text{conv}(P_1, \ldots, P_L)$. Then, $\tilde{P}_1, \ldots, \tilde{P}_M \in \text{conv}(\{P_j : j \neq i\} \cup \{G\})$. This shows that $(\boldsymbol{\Pi}, \boldsymbol{P})$ does not satisfy **(M)**. The result follows.

(B′) Clearly, $\dim \text{span}(\tilde{P}_1, \ldots, \tilde{P}_M) \leqslant \dim \text{span}(P_1, \ldots, P_L)$ since $\tilde{P}_i = \boldsymbol{\pi}_i^T \boldsymbol{P}$ for all $i \in [M]$. Since $(\boldsymbol{\Pi}, \boldsymbol{P})$ satisfies **(L)**, $\text{span}(P_1, \ldots, P_L) \subset \text{span}(\tilde{P}_1, \ldots, \tilde{P}_L)$, which implies that

$$\dim \text{span}(P_1, \ldots, P_L) \leqslant \dim \text{span}(\tilde{P}_1, \ldots, \tilde{P}_M).$$

Therefore, $\dim \text{span}(\tilde{P}_1, \ldots, \tilde{P}_M) = \dim \text{span}(P_1, \ldots, P_L)$. Then, since $\tilde{P}_1, \ldots, \tilde{P}_M \in \text{range}(\boldsymbol{\Pi})$, $\dim \text{span}(P_1, \ldots, P_L) \leqslant \dim \text{range}\, \boldsymbol{\Pi}$. By Result 3.117 of Axler (2015),

$$\text{rank}(\boldsymbol{\Pi}) = \dim \text{range}\, \boldsymbol{\Pi} \geqslant \dim \text{span}(P_1, \ldots, P_L).$$

□

As a corollary, Theorem 31 implies that if there is a linear factorization $(\boldsymbol{\Pi}, \boldsymbol{P})$ of $\tilde{P}$ and $P_1, \ldots, P_L$ are linearly independent, then there must be at least as many contaminated distributions as base distributions, i.e., $M \geqslant L$. Also, note that **(A′)** appears as a sufficient condition in Sanderson and Scott (2014).

By comparing **(A)** with **(A′)** and **(B2)** with **(B′)**, we see that the proposed sufficient conditions are not much stronger than **(M)** and **(L)** require. Since joint irreducibility of $P_1, \ldots, P_L$ entails their linear independence by Lemma 44, under **(A)**, **(B2)** and **(B′)** are the same. **(A)** differs from **(A′)** in that it requires that a slightly larger set of distributions are irreducible with respect to convex combinations of the remaining distributions. Specifically, under **(A)**, *every convex combination* of a subset of the P_is is irreducible with respect to every convex combination of the other P_is whereas **(A′)** only requires that every P_i be irreducible with respect to every convex combination of the other P_is.

5.9.3 Classification with Partial Labels

Most of our definitions and results for classification with partial labels parallel those of demixing mixed membership models. We say that $\tilde{\boldsymbol{P}}$ is $\boldsymbol{\Pi}^+$-factorizable if there exists a pair $(\boldsymbol{\Pi}, \boldsymbol{P}) \in \Delta_L^M \times \mathcal{P}^L$ that solves (5.2) such that $\boldsymbol{\Pi}$ is consistent with $\boldsymbol{\Pi}^+$; we call $(\boldsymbol{\Pi}, \boldsymbol{P})$ an $\boldsymbol{\Pi}^+$-factorization of $\tilde{\boldsymbol{P}}$. We say a partial label model is identifiable if given $(\tilde{\boldsymbol{P}}, \boldsymbol{\Pi}^+)$, $\tilde{\boldsymbol{P}}$ has a unique $\boldsymbol{\Pi}^+$-factorization $(\boldsymbol{\Pi}, \boldsymbol{P})$.

Our definitions of maximal and linear $\boldsymbol{\Pi}^+$-factorizations resemble definitions 14 and 15.

Definition 16. *We say a $\boldsymbol{\Pi}^+$-factorization $(\boldsymbol{\Pi}, \boldsymbol{P})$ of $\tilde{\boldsymbol{P}}$ is* maximal ***(M)*** *iff for all $\boldsymbol{\Pi}^+$-factorizations $(\boldsymbol{\Pi}', \boldsymbol{P}')$ of $\tilde{\boldsymbol{P}}$ with $\boldsymbol{P}' = (P_1', \ldots, P_L')^T \in \mathcal{P}^L$, it holds that $\{P_1', \ldots, P_L'\} \subseteq \operatorname{conv}(P_1, \ldots, P_L)$.*

Definition 17. *We say a $\boldsymbol{\Pi}^+$-factorization $(\boldsymbol{\Pi}, \boldsymbol{P})$ of $\tilde{\boldsymbol{P}}$ is* linear ***(L)*** *iff $\{P_1, \ldots, P_L\} \subseteq \operatorname{span}(\tilde{P}_1, \ldots, \tilde{P}_M)$.*

Similarly, $\boldsymbol{\Pi}^+$-factorizations that satisfy **(A)** and **(B3)** are maximal and linear. The proof is identical and, accordingly, omitted.

Proposition 9. *Let $(\boldsymbol{\Pi}, \boldsymbol{P})$ be a $\boldsymbol{\Pi}^+$-factorization of $\tilde{\boldsymbol{P}}$. If $(\boldsymbol{\Pi}, \boldsymbol{P})$ satisfies **(A)** and **(B3)**, then $(\boldsymbol{\Pi}, \boldsymbol{P})$ satisfies **(M)** and **(L)**.*

Linear $\boldsymbol{\Pi}^+$-factorizations must satisfy **(B′)**; indeed, the proof is identical to the proof for linear factorizations. However, maximal $\boldsymbol{\Pi}^+$-factorizations need not satisfy **(A′)**. Consider the following counterexample. Let $Q_1 \sim \text{unif}(0, 2)$ and $Q_2 \sim \text{unif}(1, 3)$. Let $P_1 = \frac{2}{3}Q_1 + \frac{1}{3}Q_2$, $P_2 = \frac{1}{3}Q_1 + \frac{2}{3}Q_2$, $\tilde{P}_1 = P_1$, and $\tilde{P}_2 = P_2$. Then, $\boldsymbol{\Pi}^+ = \boldsymbol{I}_2$—the identity matrix. Then, any $(\boldsymbol{\Pi}', \boldsymbol{P}')$ that satisfies (5.2) and is consistent with $\boldsymbol{\Pi}^+$ must be such that $(P_1, P_2)^T = \boldsymbol{P}'$. Therefore, **(M′)** is satisfied. But, clearly, **(A′)** is not satisfied.

In summary, we are unable to offer a necessary condition that is close to **(A)**. On the other hand, **(B3)** is necessary.

Proposition 10. *Let $(\boldsymbol{\Pi}, \boldsymbol{P})$ be an $\boldsymbol{\Pi}^+$-factorization of $\tilde{\boldsymbol{P}}$. If $(\tilde{\boldsymbol{P}}, \boldsymbol{\Pi}^+)$ is identifiable, then the columns of $\boldsymbol{\Pi}^+$ are distinct.*

Proof. First, suppose $(\tilde{\boldsymbol{P}}, \boldsymbol{\Pi}^+)$ is identifiable. Then, we can write $\tilde{\boldsymbol{P}} = \boldsymbol{\Pi}\boldsymbol{P}$ where $\boldsymbol{\Pi}$ is consistent with $\boldsymbol{\Pi}^+$. We claim that for all $i \neq j$, $P_i \neq P_j$. To the contrary, suppose that there exists $i \neq j$ such that $P_i = P_j$. Without loss of generality, suppose $i = 1, j = 2$. Then, we can write

$$\tilde{\boldsymbol{P}} = \begin{pmatrix} 2\boldsymbol{\Pi}_{:,1} & \boldsymbol{\Pi}_{:,3} & \dots & \boldsymbol{\Pi}_{:,L} \end{pmatrix} \begin{pmatrix} P_1 \\ P_3 \\ \vdots \\ P_L \end{pmatrix},$$

which contradicts the uniqueness of $\boldsymbol{P}$ and $\boldsymbol{\Pi}$.

Next, we give a proof by contraposition. Suppose that there exists $i \neq j$ such that $\boldsymbol{\Pi}^+_{:,i} = \boldsymbol{\Pi}^+_{:,j}$. Without loss of generality, let $i = 1$ and $j = 2$. Suppose that $(\boldsymbol{\Pi}, \boldsymbol{P})$ is consistent with $\boldsymbol{\Pi}^+$ and solves $\tilde{\boldsymbol{P}} = \boldsymbol{\Pi}\boldsymbol{P}$. Then, the pair $(\boldsymbol{\Pi}', \boldsymbol{P}')$ given by

$$\boldsymbol{\Pi}' = \begin{pmatrix} \boldsymbol{\Pi}_{:,2} & \boldsymbol{\Pi}_{:,1} & \boldsymbol{\Pi}_{:,3} & \dots & \boldsymbol{\Pi}_{:,L} \end{pmatrix}$$
$$\boldsymbol{P}' = \begin{pmatrix} P_2 \\ P_1 \\ P_3 \\ \vdots \\ P_L \end{pmatrix}$$

solves $\tilde{\boldsymbol{P}} = \boldsymbol{\Pi}'\boldsymbol{P}'$ and is consistent with $\boldsymbol{\Pi}^+$. If $P_1 = P_2$, then $(\tilde{\boldsymbol{P}}, \boldsymbol{\Pi}^+)$ is not identifiable, so we may rule out this case. Therefore, $\boldsymbol{P}' \neq \boldsymbol{P}$, yielding the result. □

5.10 Identification

In this section, we establish our identification results, i.e., Theorems 26 and 27. We begin by proving Proposition 6. Second, we prove a set of useful lemmas in Section 5.10.2. Third, we present our results on demixing mixed membership models in Section 5.10.3. Finally, we present our results on classification with partial labels in Section 5.10.4.

5.10.1 Proof of Proposition 6

Proof. We prove the claims in order.

1. Without loss of generality, suppose $i = 1$ and let $A = [L]\backslash\{1\}$. There is at least one point attaining the maximum in the optimization problem $\kappa^*(\boldsymbol{\eta}_1 \,|\, \{\boldsymbol{\eta}_j : j \neq 1\})$ by Lemma 43. Take a G that achieves the maximum in $\kappa^*(Q_1 \,|\, \{Q_j : j \neq 1\})$, which exists also by Lemma 43. Then, we can write:

$$Q_1 = (1 - \sum_{j \geqslant 2} \mu_j)G + \sum_{j \geqslant 2} \mu_j Q_j. \tag{5.7}$$

   Note that since $\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_L$ are linearly independent and $P_1, \ldots, P_L$ are jointly irreducible, $Q_1, \ldots, Q_L$ are linearly independent by Lemma 44. Therefore, $\kappa^*(Q_1 \,|\, \{Q_j : j \neq 1\}) = \sum_{j \geqslant 2} \mu_j < 1$ because, if not, $Q_1 = \sum_{j \geqslant 2} \mu_j Q_j$.

   Further, any G that satisfies (5.7) has the form $\sum_{i=1}^{L} \gamma_i P_i$ because (5.7) implies that $G \in \text{span}(Q_1, \ldots, Q_L)$ and each $Q_i \in \text{conv}(P_1, \ldots, P_L)$ by hypothesis. The γ_i must sum to one, and we have that they are nonnegative by joint irreducibility. That is, $\boldsymbol{\gamma} \equiv \Big(\gamma_1, \ldots, \gamma_L\Big)^T$ is a discrete distribution. Then, the above equation is equivalent to

$$\boldsymbol{\eta}_1^T \boldsymbol{P} = (1 - \sum_{j \geqslant 2} \mu_j)\boldsymbol{\gamma}^T \boldsymbol{P} + \sum_{j \geqslant 2} \mu_j \boldsymbol{\eta}_j^T \boldsymbol{P}. \tag{5.8}$$

   Since $P_1, \ldots, P_L$ are jointly irreducible, $P_1, \ldots, P_L$ are linearly independent by Lemma 44. By linear independence of $P_1, \ldots, P_L$, we obtain

$$\boldsymbol{\eta}_1 = (1 - \sum_{j \geqslant 2} \mu_j)\boldsymbol{\gamma} + \sum_{j \geqslant 2} \mu_j \boldsymbol{\eta}_j. \tag{5.9}$$

   Consequently, $\kappa^*(Q_1 \,|\, \{Q_j : j \neq 1\}) = \kappa^*(\boldsymbol{\eta}_1 \,|\, \{\boldsymbol{\eta}_j : j \neq 1\}) < 1$. This completes the proof of statement 1.

2. This result follows immediately from Lemma 43.

3. By equations (5.8) and (5.9), there is a one-to-one correspondence between the maximizer G to $\kappa^*(Q_1 \,|\, \{Q_j : j \neq 1\})$ and the maximizer $\boldsymbol{\gamma}$ to $\kappa^*(\boldsymbol{\eta}_1 \,|\, \{\boldsymbol{\eta}_j : j \neq 1\})$. The one-to-one correspondence is given by $G = \boldsymbol{\gamma}^T \boldsymbol{P}$.

□

5.10.2 Lemmas for Identification

We present some technical results that are used repeatedly for our identification results. Lemma 30 gives us some useful properties of the two-sample κ^* that we exploit in the PartialLabel and Demix algorithms. Statement *1* gives an alternative form of κ^*. Statement *2* gives the intuitive result that the residues lie on the boundary of the simplex. Statement *3* gives a useful relation for determining whether two mixture proportions are on the same face; we use this relation extensively in our algorithms.

Lemma 30. *Let $F_1, \ldots, F_K$ be jointly irreducible distributions with $\boldsymbol{F} = (F_1, \ldots, F_K)^T$, Q_1, Q_2 be two distributions such that $Q_i = \boldsymbol{\eta}_i^T \boldsymbol{F}$ where $\boldsymbol{\eta}_i \in \Delta_K$ for $i = 1, 2$ and $\boldsymbol{\eta}_1 \neq \boldsymbol{\eta}_2$. Let R be the residue of Q_1 wrt Q_2 and $R = \boldsymbol{\mu}^T \boldsymbol{F}$.*

1. *There is a one-to-one correspondence between the optimization problem in $\kappa^*(Q_1 \,|\, Q_2)$ and the optimization problem*

$$\max(\alpha \geqslant 1 | \exists \text{ a distribution } G \text{ s.t } G = Q_2 + \alpha(Q_1 - Q_2))$$

   via $\alpha = (1 - \kappa)^{-1}$.

2. *$\boldsymbol{\mu} \in \partial\Delta_K$.*

3. *$\mathcal{N}(\boldsymbol{\eta}_2) \nsubseteq \mathcal{N}(\boldsymbol{\eta}_1)$ if and only if $R = Q_1$ if and only if $\kappa^*(Q_1 \,|\, Q_2) = 0$.*

Proof. We note that we may assume that $R = \boldsymbol{\mu}^T \boldsymbol{F}$ since by definition of the residue, $R \in \text{span}(Q_1, Q_2)$ and $Q_1, Q_2 \in \text{conv}(F_1, \ldots, F_K)$.

1. Consider the linear relation: $Q_1 = (1 - \kappa)G + \kappa Q_2$ where $\kappa \in [0, 1]$. Since $F_1, \ldots, F_K$ are jointly irreducible and $\boldsymbol{\eta}_1$ and $\boldsymbol{\eta}_2$ are linearly independent, Q_1 and Q_2 are linearly independent by Lemma 44. Therefore, $\kappa < 1$. We can rewrite the relation as

$$G = \frac{1}{1 - \kappa} Q_1 - \frac{\kappa}{1 - \kappa} Q_2 = \alpha Q_1 + (1 - \alpha) Q_2$$

   where $\alpha = \frac{1}{1-\kappa}$. The equivalence follows.

2. Since R is the residue of Q_1 wrt Q_2, by Proposition 6, $\boldsymbol{\mu}$ is the residue of $\boldsymbol{\eta}_1$ wrt $\boldsymbol{\eta}_2$ and $\boldsymbol{\mu} \in \Delta_K$. Therefore, by statement *1* in Lemma 30, $\boldsymbol{\mu}$ is such that α^* is maximized subject to the following constraints:

$$\begin{aligned} \boldsymbol{\mu} &= (1-\alpha^*)\boldsymbol{\eta}_2 + \alpha^*\boldsymbol{\eta}_1 \\ \alpha^* &\geqslant 1 \\ \boldsymbol{\mu} &\in \Delta_K. \end{aligned}$$

Suppose that $\min_i \mu_i > 0$. Then,

$$\boldsymbol{\mu} = (1-\alpha^*)\boldsymbol{\eta}_2 + \alpha^*\boldsymbol{\eta}_1 = \boldsymbol{\eta}_2 + \alpha(\boldsymbol{\eta}_1 - \boldsymbol{\eta}_2) > 0$$

so that there is some $\epsilon > 0$ such that

$$\begin{aligned} \boldsymbol{\mu}' &= (1-\alpha^*-\epsilon)\boldsymbol{\eta}_2 + (\alpha^*+\epsilon)\boldsymbol{\eta}_1 \\ \alpha^* + \epsilon &\geqslant 1 \\ \boldsymbol{\mu}' &\in \Delta_K. \end{aligned}$$

But, this contradicts the definition of α^* and $\boldsymbol{\mu}$. Therefore, $\min_i \mu_i = 0$. Consequently, $\boldsymbol{\mu} \in \partial\Delta_K$.

3. By definition of κ^*, it is clear that $R = Q_1$ if and only if $\kappa^*(Q_1 \,|\, Q_2) = 0$. Therefore, it suffices to show that $\mathcal{N}(\boldsymbol{\eta}_2) \nsubseteq \mathcal{N}(\boldsymbol{\eta}_1)$ if and only if $\kappa^*(Q_1 \,|\, Q_2) = 0$. Suppose $\mathcal{N}(\boldsymbol{\eta}_2) \nsubseteq \mathcal{N}(\boldsymbol{\eta}_1)$. Then, there must be $i \in [K]$ such that $\eta_{2,i} > 0$ and $\eta_{1,i} = 0$. For any $\alpha > 1$,

$$\min_{i\in[K]} (1-\alpha)\eta_{2,i} + \alpha\eta_{1,i} < 0;$$

but, this violates the constraint of the optimization problem. Therefore, $\alpha = 1$. By statement *1* in Lemma 30, $\kappa^*(\boldsymbol{\eta}_1 \,|\, \boldsymbol{\eta}_2) = 0$. By statement *1* of Proposition 6, $\kappa^*(Q_1 \,|\, Q_2) = \kappa^*(\boldsymbol{\eta}_1 \,|\, \boldsymbol{\eta}_2) = 0$.

Now, suppose $\mathcal{N}(\boldsymbol{\eta}_2) \subseteq \mathcal{N}(\boldsymbol{\eta}_1)$. Then, for any $i \in [K]$, if $\eta_{2,i} > 0$, then $\eta_{1,i} > 0$. Then, there is $\alpha > 1$ sufficiently close to 1 such that

$$\min_{i\in[K]} \eta_{2,i} + \alpha(\eta_{1,i} - \eta_{2,i}) \geqslant 0.$$

By statement *1* in this Lemma, $\kappa^*(\boldsymbol{\eta}_1 \,|\, \boldsymbol{\eta}_2) > 0$. By statement *1* of proposition 6, $\kappa^*(Q_1 \,|\, Q_2) = \kappa^*(\boldsymbol{\eta}_1 \,|\, \boldsymbol{\eta}_2) > 0$.

□

Lemma 31. *Let $0 \leqslant k < L$. If $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \in \Delta_L$ are linearly independent and $\boldsymbol{w}_{k+1}, \ldots, \boldsymbol{w}_L \in \Delta_L$ are random vectors drawn independently from the uniform distribution on a set $A \subset \Delta_L$ with positive $(L-1)$-dimensional Lebesgue measure, then $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k, \boldsymbol{w}_{k+1}, \ldots, \boldsymbol{w}_L$ are linearly independent with probability* 1.

Proof. We prove the result inductively. To begin, we prove the base case, i.e. $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k, \boldsymbol{w}_{k+1}$ are linearly independent w.p. 1. It suffices to show that $\boldsymbol{w}_{k+1} \notin \operatorname{span}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k)$ w.p. 1. Thus, it is enough to show that $\operatorname{span}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k) \cap A$ has $(L-1)$-dimensional Lebesgue measure 0. Since

$$\operatorname{span}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k) \cap A \subset \operatorname{span}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k) \cap \Delta_L,$$

it suffices to show that $\operatorname{span}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k) \cap \Delta_L$ has $(L-1)$-dimensional Lebesgue measure 0.

Next, we claim that $\operatorname{span}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k) \cap \Delta_L \subseteq \operatorname{aff}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k)$. Let $\sum_{i=1}^k \alpha_i \boldsymbol{v}_i \in \Delta_L$. Since $\boldsymbol{v}_i \in \Delta_L$ for all $i \in [k]$, we can write $\boldsymbol{v}_i = \sum_{j=1}^L \beta_{i,j} \boldsymbol{e}_i$ where $\beta_{i,j} \geqslant 0$ and $\sum_{j=1}^L \beta_{i,j} = 1$. Then, since $\sum_{i=1}^k \alpha_i \boldsymbol{v}_i = \sum_{i=1}^k \alpha_i \sum_{j=1}^L \beta_{i,j} \boldsymbol{e}_j \in \Delta_L$, it holds that

$$1 = \sum_{i=1}^k \alpha_i \sum_{j=1}^L \beta_{i,j} = \sum_{i=1}^k \alpha_i.$$

Thus, $\sum_{i=1}^k \alpha_i \boldsymbol{v}_i \in \operatorname{aff}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k)$, establishing the claim.

Thus, it suffices to show that $\operatorname{aff}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k)$ has $(L-1)$-dimensional Lebesgue measure 0. $\operatorname{aff}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k)$ has affine dimension at most $k-1$. Since it is not possible to fit a $(L-1)$ dimensional ball in an affine subspace of affine dimension $k-1 < L-1$, $\operatorname{aff}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k)$ has $(L-1)$-dimensional Lebesgue measure 0. Thus, with probability 1, $\boldsymbol{w}_{k+1} \notin \operatorname{span}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k)$. This establishes the base case.

The inductive step follows by a union bound and a similar argument to the base case. Thus, the result follows. □

5.10.3 Demixing Mixed Membership Models

In this section, we prove our identification result for demixing mixed membership models, i.e., Theorem 26. First, we present technical lemmas in Section 5.10.3. Second, in Section 5.10.3, we present the key subroutine FaceTest and prove that it behaves as desired. Third, we prove Theorem 26 in Section 5.10.3. Finally, in Section 5.10.3, we extend our results to the nonsquare case (where $M > L$).

Lemmas

Lemma 32 establishes an intuitive continuity property of the two-sample version of κ^* and the residue. Recall that $\|\cdot\|$ denotes an arbitrary finite-dimensional norm on $\mathbb{R}^L$.

Lemma 32. *Let $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2 \in \Delta_L$ be distinct vectors and let $\boldsymbol{\mu}$ be the residue of $\boldsymbol{\eta}_2$ wrt $\boldsymbol{\eta}_1$. Let $\boldsymbol{\gamma}_n \in \Delta_L$ be a sequence such that $\|\boldsymbol{\gamma}_n - \boldsymbol{\eta}_2\| \longrightarrow 0$ as $n \longrightarrow \infty$, and let $\boldsymbol{\tau}_n$ be the residue of $\boldsymbol{\gamma}_n$ wrt $\boldsymbol{\eta}_1$. Then,*

1. $\lim_{n\longrightarrow\infty} \kappa^*(\boldsymbol{\gamma}_n \,|\, \boldsymbol{\eta}_1) = \kappa^*(\boldsymbol{\eta}_2 \,|\, \boldsymbol{\eta}_1)$, *and*
2. $\lim_{n\longrightarrow\infty} \|\boldsymbol{\tau}_n - \boldsymbol{\mu}\| = 0$.
3. *If, in addition, $\boldsymbol{\rho}_n \in \Delta_L$ is a sequence such that $\|\boldsymbol{\rho}_n - \boldsymbol{\eta}_2\| \longrightarrow 0$ as $n \longrightarrow \infty$ and $\mathcal{N}(\boldsymbol{\eta}_2) = \mathcal{N}(\boldsymbol{\gamma}_n) = \mathcal{N}(\boldsymbol{\rho}_n)$ for all n. Then, $\lim_{n\longrightarrow\infty} \kappa^*(\boldsymbol{\gamma}_n \,|\, \boldsymbol{\rho}_n) = 1$.*

Proof. 1. In order to apply the residue operator κ^* to $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\gamma}_n$ we think of $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\gamma}_n$ as discrete probability distributions. By Proposition 5,

$$\kappa^*(\boldsymbol{\eta}_2 \,|\, \boldsymbol{\eta}_1) = \min_{i,\, \eta_{1,i}>0} \frac{\eta_{2,i}}{\eta_{1,i}}.$$

Clearly, there is a constant $\delta > 0$ such that $\min_{i,\, \boldsymbol{\eta}_{1,i}>0} \boldsymbol{\eta}_{1,i} > \delta$. Let $\epsilon > 0$. By the equivalence of norms on finite-dimensional vector spaces, there exists a constant $C > 0$ such that $\|\cdot\|_\infty \leqslant C\,\|\cdot\|$ where $\|\cdot\|_\infty$ denotes the supremum norm. Thus, since $\|\boldsymbol{\gamma}_n - \boldsymbol{\eta}_2\| \longrightarrow 0$ as $n \longrightarrow \infty$, we can let n large enough such that $|\boldsymbol{\gamma}_{n,i} - \boldsymbol{\eta}_{2,i}| \leqslant \epsilon$ for all $i \in [L]$. Then,

$$\begin{aligned}\kappa^*(\boldsymbol{\gamma}_n \,|\, \boldsymbol{\eta}_1) = \min_{i,\eta_{1,i}>0} \frac{\gamma_{n,i}}{\eta_{1,i}} &\leqslant \min_{i,\eta_{1,i}>0} \frac{\eta_{2,i}+\epsilon}{\eta_{1,i}} \\ &\leqslant \kappa^*(\boldsymbol{\eta}_2 \,|\, \boldsymbol{\eta}_1) + \frac{\epsilon}{\delta}.\end{aligned}$$

Similarly,

$$\begin{aligned}\kappa^*(\boldsymbol{\gamma}_n \,|\, \boldsymbol{\eta}_1) &= \min_{i,\eta_{1,i}>0} \frac{\gamma_{n,i}}{\eta_{1,i}} \\ &\geqslant \frac{\eta_{2,i} - \epsilon}{\eta_{1,i}} \\ &\geqslant \kappa^*(\boldsymbol{\eta}_2 \,|\, \boldsymbol{\eta}_1) - \frac{\epsilon}{\delta}.\end{aligned}$$

Since $\epsilon > 0$ was arbitrary, statement *1* follows.

2. Write $\boldsymbol{\mu} = \kappa\boldsymbol{\eta}_2 + (1-\kappa)\boldsymbol{\eta}_1$ and $\boldsymbol{\tau}_n = \kappa_n\boldsymbol{\gamma}_n + (1-\kappa_n)\boldsymbol{\eta}_1$ where $\kappa = \kappa^*(\boldsymbol{\eta}_2 \,|\, \boldsymbol{\eta}_1)$ and $\kappa_n = \kappa^*(\boldsymbol{\gamma}_n \,|\, \boldsymbol{\eta}_1)$. Then, by the triangle inequality,

$$\begin{aligned}\|\boldsymbol{\mu} - \boldsymbol{\tau}_n\| &\leqslant \|\kappa\boldsymbol{\eta}_2 - \kappa_n\boldsymbol{\gamma}_n\| + \|(1-\kappa)\boldsymbol{\eta}_1 - (1-\kappa_n)\boldsymbol{\eta}_1\| \\ &\leqslant |\kappa - \kappa_n|\,\|\boldsymbol{\eta}_2\| + |\kappa_n|\,\|\boldsymbol{\eta}_2 - \boldsymbol{\gamma}_n\| + |\kappa - \kappa_n|\,\|\boldsymbol{\eta}_1\| \longrightarrow 0\end{aligned}$$

as $n \longrightarrow \infty$.

3. W.l.o.g., suppose that $[K] = \mathcal{N}(\boldsymbol{\eta}_2) = \mathcal{N}(\boldsymbol{\gamma}_n) = \mathcal{N}(\boldsymbol{\rho}_n)$ where $K \leqslant L$. Observe that

$$\kappa^*(\boldsymbol{\gamma}_n \,|\, \boldsymbol{\rho}_n) = \min_{i,\,\rho_{n,i}>0} \frac{\gamma_{n,i}}{\rho_{n,i}} = \min_{i\in[K]} \frac{\gamma_{n,i}}{\rho_{n,i}}$$

There exists a constant $\delta > 0$ such that $\min_{i\in[K]} \eta_{2,i} \geqslant \delta$. Let $\delta > \epsilon > 0$. By the equivalence of norms on finite-dimensional vector spaces, we can let n large enough such that $|\gamma_{n,i} - \eta_{2,i}| \leqslant \epsilon$ and $|\rho_{n,i} - \eta_{2,i}| \leqslant \epsilon$ for all $i \in [L]$. Then,

$$\begin{aligned}\kappa^*(\boldsymbol{\gamma}_n \,|\, \boldsymbol{\rho}_n) &= \min_{i\in[K]} \frac{\gamma_{n,i}}{\rho_{n,i}} \\ &\leqslant \min_{i\in[K]} \frac{\eta_{2,i} + \epsilon}{\eta_{2,i} - \epsilon} \\ &\leqslant \frac{\eta_{2,i} + \epsilon}{\eta_{2,i} - \epsilon}\end{aligned}$$

for any $i \in [K]$, which goes to 1 as $\epsilon \longrightarrow 0$. Similarly,

$$\begin{aligned}\kappa^*(\boldsymbol{\gamma}_n \,|\, \boldsymbol{\rho}_n) &= \min_{i\in[K]} \frac{\gamma_{n,i}}{\rho_{n,i}} \\ &\geqslant \min_{i\in[K]} \frac{\eta_{2,i} - \epsilon}{\eta_{2,i} + \epsilon}\end{aligned}$$

Since for any $i \in [K]$,

$$\frac{\eta_{2,i} - \epsilon}{\eta_{2,i} + \epsilon} \longrightarrow 1$$

as $\epsilon \longrightarrow 0$, the above lower bound goes to 1 as $\epsilon \longrightarrow 0$. Thus, statement *3* follows.

□

Lemma 33 guarantees that certain operations in the Demix algorithm preserve linear independence of the mixture proportions. The proof uses tools from multilinear algebra.

Lemma 33. *Let $\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_K \in \Delta_K$ be linearly independent and $P_1, \ldots, P_K$ be jointly irreducible. Let $Q_i = \boldsymbol{\tau}_i^T \boldsymbol{P}$ for $i \in [K]$. Then for any $i, j \in [K]$ such that $i \neq j$,*

1. *If $\boldsymbol{\eta} = \sum_{k=1}^K a_k \boldsymbol{\tau}_k$ with $a_j \neq 0$, then $\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_{j-1}, \boldsymbol{\eta}, \boldsymbol{\tau}_{j+1}, \ldots, \boldsymbol{\tau}_K$ are linearly independent.*

2. *Let R_k be the residue of Q_k with respect to Q_j for all $k \in [K] \backslash \{j\}$. Then, $R_k = \boldsymbol{\eta}_k^T \boldsymbol{P}$ where $\boldsymbol{\eta}_k \in \Delta_L$ and $\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_{j-1}, \boldsymbol{\tau}_j, \boldsymbol{\eta}_{j+1}, \ldots, \boldsymbol{\eta}_K$ are linearly independent.*

3. *Let $\boldsymbol{\tau}^* \in \text{conv}(\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_k)^\circ$ and $\boldsymbol{\eta}_i \in \text{conv}(\boldsymbol{\tau}_i, \boldsymbol{\tau}^*)^\circ$ for $i \in [k]$ where $k \leqslant K$. Then,*

$$\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \ldots, \boldsymbol{\eta}_k, \boldsymbol{\tau}_{k+1}, \ldots, \boldsymbol{\tau}_K$$

are linearly independent.

Proof. We use the multilinear expansion and usual properties of determinants.

1. Viewing each $\boldsymbol{\tau}_i$ as a column vector,

$$\det(\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_{j-1}, \sum_k a_k \boldsymbol{\tau}_k, \boldsymbol{\tau}_{j+1}, \ldots, \boldsymbol{\tau}_K) = a_j \det(\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_K) \neq 0.$$

2. Linear independence of $\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_K$ implies that the $Q_1, \ldots, Q_K$ are distinct. Hence, by Proposition 5, we can write $\boldsymbol{\eta}_k = (1 - \alpha_k)\boldsymbol{\tau}_j + \alpha_k \boldsymbol{\tau}_k$ where $\alpha_k \neq 0 \, \forall k \neq j$. Then, it holds that

$$\det(\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_{j-1}, \boldsymbol{\tau}_j, \boldsymbol{\eta}_{j+1}, \ldots, \boldsymbol{\eta}_K) = \left(\prod_{i \neq j} \alpha_i \right) \det(\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_K) \neq 0.$$

3. Since $\boldsymbol{\eta}_j = (1-\alpha_j)\boldsymbol{\tau}^* + \alpha_j \boldsymbol{\tau}_j$ where $\alpha_j \in (0,1)$ for all $j \leqslant k$, and $\boldsymbol{\tau}^* = \sum_i \beta_i \boldsymbol{\tau}_i$, it holds

$$\det(\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_{k-1}, \boldsymbol{\tau}_k, \ldots, \boldsymbol{\tau}_K) = \left(1 + \sum_{j=1}^{k} \frac{(1-\alpha_j)}{\alpha_j}\beta_j\right)\left(\prod_{i=1}^{k}\alpha_i\right)\det(\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_K) \neq 0.$$

□

Lemma 34 gives a condition on the mixture proportions under which the multi-sample residue is unique. Lemma 2 in Blanchard and Scott (2014) is very similar and is proved in a very similar way. We give a useful generalization here that reproduces many of the same details.

Lemma 34. *Let $l, k \in [L]$. Let $\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_L \in \Delta_L$ be linearly independent. We have that condition 1 implies condition 2 and condition 2 implies condition 3.*

1. *There exists a decomposition*

$$\boldsymbol{\tau}_l = \kappa \boldsymbol{e}_k + (1-\kappa)\boldsymbol{\tau}_l'$$

where $\kappa > 0$ and $\boldsymbol{\tau}_l' \in \operatorname{conv}(\{\boldsymbol{\tau}_j : j \neq l\})$. Further, for every $\boldsymbol{e}_i$ such that $i \neq k$, there exists a decomposition

$$\boldsymbol{e}_i = \sum_{j=1}^{L} a_j \boldsymbol{\tau}_j$$

such that $a_l < \frac{1}{\kappa}$.

2. *Let*

$$\boldsymbol{T} = \begin{pmatrix} \boldsymbol{\tau}_1^T \\ \vdots \\ \boldsymbol{\tau}_L^T \end{pmatrix};$$

the matrix $\boldsymbol{T}$ is invertible and $\boldsymbol{T}^{-1}$ is such that $(\boldsymbol{T}^{-1})_{l,k} > 0$ and $(\boldsymbol{T}^{-1})_{l,i} \leqslant 0$ for $i \neq k$ and $(\boldsymbol{T}^{-1})_{l,k} > (\boldsymbol{T}^{-1})_{j,k}$ for $j \neq l$. In words, the (l,k)th entry in $\boldsymbol{T}^{-1}$ is positive, every other entry in the lth row of $\boldsymbol{T}^{-1}$ is nonpositive and every other entry in the kth column of $\boldsymbol{T}^{-1}$ is strictly less than the (l,k)th entry. [2]

[2] $(\boldsymbol{T}^{-1})_{i,j}$ is the $i \times j$ entry in the matrix $\boldsymbol{T}^{-1}$.

3. The residue of $\boldsymbol{\tau}_l$ with respect to $\{\boldsymbol{\tau}_j, j \neq l\}$ is $\boldsymbol{e}_k$.

Proof. Without loss of generality, let $l = 1$ and $k = 2$. By relabeling the vectors $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_L$, we can assume without loss of generality that $k = 1$. First, we show that condition *1* implies condition *2*. Suppose that condition *1* holds. Then, there exists $\kappa > 0$ such that

$$\boldsymbol{\tau}_1 = \kappa \boldsymbol{e}_1 + (1 - \kappa) \sum_{i=2}^{L} \mu_i \boldsymbol{\tau}_i$$

with $\mu_i \geqslant 0$ for $i \in [L] \backslash \{1\}$. Then,

$$\boldsymbol{e}_1 = \frac{1}{\kappa}(\boldsymbol{\tau}_1 - \sum_{i \geqslant 2}(1 - \kappa)\mu_i \boldsymbol{\tau}_i).$$

Hence, the first row of $\boldsymbol{T}^{-1}$ is given by $\frac{1}{\kappa}(1, -(1-\kappa)\mu_2, \cdots, -(1-\kappa)\mu_L)$. This shows that the first row is such that $(\boldsymbol{T}^{-1})_{1,1} > 0$ and $(\boldsymbol{T}^{-1})_{1,i} \leqslant 0$ for $i \neq 1$.

Consider $\boldsymbol{e}_i$ such that $i \neq 1$. Then, we have the relation: $\boldsymbol{e}_i = \sum_{j=1}^{L} a_j \boldsymbol{\tau}_j$, which gives the ith row of $\boldsymbol{T}^{-1}$. By assumption, $a_1 < \frac{1}{\kappa}$, so the $(i,1)$th entry is strictly less than the $(1,1)$th entry. Hence, *2* follows.

Now, we prove that condition *2* implies condition *3*. Suppose condition *2* is true. Consider the optimization problem

$$\max_{\boldsymbol{\nu}, \boldsymbol{\gamma}} \sum_{i=2}^{L} \nu_i \ s.t. \ \boldsymbol{\tau}_1 = (1 - \sum_{i \geqslant 2} \nu_i)\boldsymbol{\gamma} + \sum_{i=2}^{L} \nu_i \boldsymbol{\tau}_i$$

over $\boldsymbol{\gamma} \in \Delta_L$ and $\boldsymbol{\nu} = (\nu_2, \cdots, \nu_L) \in C_{L-1} = \{(\nu_2, \cdots, \nu_L) : \nu_i \geqslant 0; \sum_{i=2}^{L} \nu \leqslant 1\}$.

By the same argument given in the proof of Lemma 2 of Blanchard and Scott (2014), this optimization problem is equivalent to the program

$$\max_{\boldsymbol{\gamma} \in \Delta_L} \boldsymbol{e}_1^T (\boldsymbol{T}^T)^{-1} \boldsymbol{\gamma} \ s.t. \ \boldsymbol{\nu}((\boldsymbol{T}^T)^{-1}\boldsymbol{\gamma}) \in C_{L-1}$$

where $\boldsymbol{\nu}(\boldsymbol{\eta}) \coloneqq \eta_1^{-1}(-\eta_2, \cdots, -\eta_L)$. The above objective is of the form $\boldsymbol{a}^T \boldsymbol{\gamma}$ where $\boldsymbol{a}$ is the first column of $\boldsymbol{T}^{-1}$. Since $l = 1$, by assumption, for every $i \neq 1$, $\boldsymbol{T}_{1,1}^{-1} > \boldsymbol{T}_{i,1}^{-1}$. Therefore, the unconstrained maximum over $\boldsymbol{\gamma} \in \Delta_L$ is attained uniquely by $\boldsymbol{\gamma} = \boldsymbol{e}_1$. Notice that $(\boldsymbol{T}^T)^{-1}\boldsymbol{e}_1$ is the first row of $\boldsymbol{T}^{-1}$. Denote this vector $\boldsymbol{b} = (b_1, \cdots, b_L)$. We show that $\boldsymbol{\nu}(\boldsymbol{b}) = b_1^{-1}(-b_2, \cdots, -b_L) \in C_{L-1}$. By assumption, $\boldsymbol{b}$ has its first coordinate positive and the

other coordinates are nonpositive. Therefore, all of the components of $\boldsymbol{\nu}(\boldsymbol{b})$ are nonnegative. Furthermore, the sum of the components of $\boldsymbol{\nu}(\boldsymbol{b})$ is

$$\sum_{i=2}^{L} \frac{-b_i}{b_1} = 1 - \frac{\sum_{i=1}^{L} b_i}{b_1} = 1 - \frac{1}{b_1} \leqslant 1.$$

The last equality follows because the rows of $\boldsymbol{T}^{-1}$ sum to 1 since $\boldsymbol{T}$ is a stochastic matrix. Then, we have $\boldsymbol{\nu}((\boldsymbol{T}^T)^{-1}\boldsymbol{e}_1) \in C_{L-1}$. Consequently, the unique maximum of the optimization problem is attained for $\boldsymbol{\gamma} = \boldsymbol{e}_1$. This establishes *3*. □

The FaceTest Algorithm

Next, we consider the main subroutine in the Demix algorithm: the FaceTest algorithm (see Algorithm 21). Proposition 11 establishes that FaceTest$(Q_1, \ldots, Q_K)$ returns 1 if and only if $Q_1, \ldots, Q_K$ are in the relative interior of the same face of the simplex.

Proposition 11. *Let $Q_j = \boldsymbol{\eta}_j^T \boldsymbol{P}$ for $\boldsymbol{\eta}_j \in \Delta_K$ and all $j \in [K]$. Let $P_1, \ldots, P_K$ be jointly irreducible, $Q_1, \ldots, Q_K \in \operatorname{conv}(P_1, \ldots, P_K)$ be distinct, and for each $i \in [K]$, let $\boldsymbol{\eta}_i$ lie in the relative interior of one of the faces of Δ_K. FaceTest$(Q_1, \ldots, Q_K)$ returns 1 if and only if $\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_K$ lie in the relative interior of the same face of Δ_K.*

Proof. Suppose that $\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_K$ lie on the relative interior of the same face of Δ_K. Then, $\mathcal{N}(Q_1) = \ldots = \mathcal{N}(Q_K)$. By statement 3 of Lemma 30, $\kappa^*(Q_i \,|Q_j) > 0$ for all $i \neq j$. Hence, FaceTest$(Q_1, \ldots, Q_K)$ returns 1.

Suppose that $Q_1, \ldots, Q_K$ do not all lie on the relative interior of the same face. Then, there exists Q_i, Q_j $(i \neq j)$ that do not lie on the relative interior of the same face. Without loss of generality, suppose that $\mathcal{N}(Q_j) \nsubseteq \mathcal{N}(Q_i)$. Then, by statement 3 of Lemma 30, $\kappa^*(Q_i \,|\, Q_j) = 0$. Hence, FaceTest$(Q_1, \ldots, Q_K)$ returns 0. □

The Demix Algorithm

Proof of Theorem 26. Let $K \leqslant L$, $\boldsymbol{\gamma}_i \in \Delta_L$ for all $i \in [K]$, $S_i = \boldsymbol{\gamma}_i^T \boldsymbol{P}$ for all $i \in [K]$, and

$$\boldsymbol{\Gamma} = \begin{pmatrix} \boldsymbol{\gamma}_1^T \\ \vdots \\ \boldsymbol{\gamma}_K^T \end{pmatrix}.$$

We claim that for any $\{i_1, \ldots, i_K\} \subset [L]$ and $\{S_1, \ldots, S_K\} \subset \text{conv}(P_{i_1}, \ldots, P_{i_K})$, if $P_1, \ldots, P_L$ are jointly irreducible, and $\mathbf{\Gamma}$ has full row rank, then w.p. 1 Demix$(S_1, \ldots, S_K)$ returns a permutation of $(P_{i_1}, \ldots, P_{i_K})$. If the claim holds, then setting $K = L$ and putting $\tilde{P}_i = S_i$ yields the result. We prove the claim by induction on K.

Consider the base case: $K = 2$. Suppose that $\{S_1, S_2\} \subset \text{conv}(P_1, P_2)$ (the other cases are similar). Note that $\boldsymbol{\gamma}_1 \neq \boldsymbol{\gamma}_2$ by linear independence of $\boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_2$. Either $\boldsymbol{\gamma}_1 \in \text{conv}(\boldsymbol{e}_1, \boldsymbol{\gamma}_2)$ or $\boldsymbol{\gamma}_1 \in \text{conv}(\boldsymbol{e}_2, \boldsymbol{\gamma}_2)$. Suppose $\boldsymbol{\gamma}_1 \in \text{conv}(\boldsymbol{e}_1, \boldsymbol{\gamma}_2)$. Condition *2* of Lemma 29 is satisfied so that $\boldsymbol{e}_1$ is the residue of $\boldsymbol{\gamma}_1$ with respect to $\boldsymbol{\gamma}_2$ and $\boldsymbol{e}_2$ is the residue of $\boldsymbol{\gamma}_2$ with respect to $\boldsymbol{\gamma}_1$. Thus, by statement *3* of Proposition 6, P_1 is the residue of S_1 with respect to S_2 and P_2 is the residue of S_2 with respect to S_1. If $\boldsymbol{\gamma}_1 \in \text{conv}(\boldsymbol{e}_2, \boldsymbol{\gamma}_2)$, then similar reasoning establishes that P_2 is the residue of S_1 with respect to S_2 and P_1 is the residue of S_2 with respect to S_1. Thus, the base case follows.

Suppose $L \geqslant K > 2$. The inductive hypothesis is:

Inductive Hypothesis: for any $\{i_1, \ldots, i_{K-1}\} \subset [L]$ and $\{S_1, \ldots, S_{K-1}\} \subset \text{conv}(P_{i_1}, \ldots, P_{i_{K-1}})$, if $P_1, \ldots, P_L$ are jointly irreducible and $\mathbf{\Gamma}$ has full row rank, then w.p. 1 Demix$(S_1, \ldots, S_{K-1})$ returns a permutation of $(P_{i_1}, \ldots, P_{i_{K-1}})$.

Suppose that $\{S_1, \ldots, S_K\} \subset \text{conv}(P_1, \ldots, P_K)$ (the other cases are similar). Set $\Xi = \text{conv}(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_K)$. With probability 1, $Q \in \text{conv}(S_2, \ldots, S_K)^\circ$. We can write $Q = \boldsymbol{\eta}^T \boldsymbol{P}$ where $\boldsymbol{\eta}$ is a uniformly distributed random vector in $\text{conv}(\boldsymbol{\gamma}_2, \ldots, \boldsymbol{\gamma}_K)$. Let R be the residue of Q with respect to S_1. By statement *3* of Proposition 6, we can write $R = \boldsymbol{\lambda}^T \boldsymbol{P}$ where $\boldsymbol{\lambda}$ is the residue of $\boldsymbol{\eta}$ with respect to $\boldsymbol{\gamma}_1$. By statement *2* of Lemma 30, $\boldsymbol{\lambda} \in \partial\Xi$.

Step 1: We claim that with probability 1, there is $l \in [K]$ such that $\boldsymbol{\lambda} \in \text{conv}(\{\boldsymbol{e}_j : j \in [K] \backslash \{l\}\})^\circ$. Let $B_{i,j} = \text{conv}(\{\boldsymbol{\gamma}_1\} \cup \{\boldsymbol{e}_k : k \in [K] \backslash \{i, j\}\})$ where $i, j \in [K]$ and $i \neq j$ and let $C = \text{conv}(\boldsymbol{\gamma}_2, \ldots, \boldsymbol{\gamma}_K)$. First, we argue that $C \cap B_{i,j}$ has affine dimension at most $K-3$.[3] Since $\boldsymbol{\gamma}_2, \ldots, \boldsymbol{\gamma}_K$ are linearly independent, C has affine dimension $K-2$. Since $\{\boldsymbol{e}_k : k \in [K] \backslash \{i, j\}\}$ are linearly independent, $B_{i,j}$ has affine dimension $K-2$ or $K-3$. If $B_{i,j}$ has affine dimension $K-3$, then $C \cap B_{i,j}$ has affine dimension at most $K-3$. So,

[3]Note that if $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in \mathbb{R}^L$ are linearly independent and $n \leqslant L$, then aff$(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ has affine dimension $n-1$.

suppose that $B_{i,j}$ has affine dimension $K-2$. If $C \cap B_{i,j}$ has affine dimension $K-2$, then $\operatorname{aff} C = \operatorname{aff} B_{i,j}$. Then, in particular, $\boldsymbol{\gamma}_1 \in \operatorname{aff} C$. But, this contradicts the linear independence of $\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_K$. Therefore, $C \cap B_{i,j}$ has affine dimension at most $K-3$.

Because C has affine dimension $K-2$ and $\boldsymbol{\eta}$ is a uniformly distributed random vector in C, with probability 1, $\boldsymbol{\eta} \notin \cup_{i,j\in[K],i\neq j} B_{i,j}$. Since $\boldsymbol{\gamma}_1 \in B_{i,j}$ for all $i,j \in [K]$ and $\boldsymbol{\eta} \in \operatorname{conv}(\boldsymbol{\lambda}, \boldsymbol{\gamma}_1)$ by definition, the convexity of $B_{i,j}$ implies that $\boldsymbol{\lambda} \notin \cup_{i,j\in[K],i\neq j} B_{i,j}$. Since $\boldsymbol{\lambda} \in \partial\Xi$, the claim follows.

Step 2: Let $R_i^{(n)}$ be the residue of $m_{\frac{n-1}{n}}(S_i, Q)$ with respect to S_1. We claim that there is some finite integer $N \geqslant 2$ such that for all $n \geqslant N$,

$$\text{FaceTest}(R_2^{(n)}, \ldots, R_K^{(n)})$$

returns 1. By Proposition 11, this is equivalent to the statement that there exists $N \geqslant 2$ such that for all $n \geqslant N$, the mixture proportions of $R_2^{(n)}, \ldots, R_K^{(n)}$ are on the relative interior of the same face. Let $m_{\frac{n-1}{n}}(S_i, Q) = (\boldsymbol{\tau}_i^{(n)})^T \boldsymbol{P}$ for $i \in [K]\backslash\{1\}$; note that $\boldsymbol{\tau}_i^{(n)} = \frac{1}{n}\boldsymbol{\gamma}_i + \frac{n-1}{n}\boldsymbol{\eta}$ and, consequently, $\boldsymbol{\tau}_i^{(n)} \in \Xi$. Since $\boldsymbol{\eta} \in \operatorname{conv}(\boldsymbol{\gamma}_2, \ldots, \boldsymbol{\gamma}_K)^\circ$ with probability 1, $\boldsymbol{\tau}_i^{(n)} \in \operatorname{conv}(\boldsymbol{\gamma}_i, \boldsymbol{\eta})^\circ$ for all $i \in [K]\backslash\{1\}$ and $n \in \mathbb{N}$, and $\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_K$ are linearly independent, it follows that for all $n \in \mathbb{N}$ with probability 1, $\boldsymbol{\gamma}_1, \boldsymbol{\tau}_2^{(n)}, \ldots, \boldsymbol{\tau}_K^{(n)}$ are linearly independent by statement *3* in Lemma 33. Fix $i \in [K]\backslash\{1\}$. It suffices to show that there is large enough N such that for $n \geqslant N$, $\text{Residue}(m_{\frac{n-1}{n}}(S_i, Q) \,|\, S_1) = R_i^{(n)}$ is on the same face as R. Let $R_i^{(n)} = (\boldsymbol{\mu}_i^{(n)})^T \boldsymbol{P}$; by statement *3* of Proposition 6, $\boldsymbol{\mu}_i^{(n)}$ is the residue of $\boldsymbol{\tau}_i^{(n)}$ with respect to $\boldsymbol{\gamma}_1$ and by statement 2 of Lemma 30 $\boldsymbol{\mu}_i^{(n)} \in \Xi$. It suffices to show that $\mathcal{N}(\boldsymbol{\mu}_i^{(n)}) = \mathcal{N}(\boldsymbol{\lambda})$, i.e., every $\boldsymbol{\mu}_i^{(n)}$ is on the same face as $\boldsymbol{\lambda}$. As $n \longrightarrow \infty$, $\boldsymbol{\tau}_i^{(n)} = (1 - \frac{n-1}{n})\boldsymbol{\gamma}_i + \frac{n-1}{n}\boldsymbol{\eta} \longrightarrow \boldsymbol{\eta}$, hence by statement *2* in Lemma 32, $\left\|\boldsymbol{\mu}_i^{(n)} - \boldsymbol{\lambda}\right\| \longrightarrow 0$. Since with probability 1, $\boldsymbol{\lambda} \in \operatorname{conv}(\{\boldsymbol{e}_j : j \in [K]\backslash\{l\}\})^\circ$ for some l (step 1), it follows that for large enough n, $\boldsymbol{\mu}_i^{(n)} \in \operatorname{conv}(\{\boldsymbol{e}_j : j \in [K]\backslash\{l\}\})^\circ$.

Step 3: Assume that n is sufficiently large such that $R_2^{(n)}, \ldots, R_K^{(n)}$ are on the same face. The algorithm recurses on $R_2^{(n)}, \ldots, R_K^{(n)}$. Since $\boldsymbol{\gamma}_1, \boldsymbol{\tau}_2^{(n)}, \ldots, \boldsymbol{\tau}_K^{(n)}$ are linearly independent, it follows by statement *2* in Lemma 33 that $\boldsymbol{\mu}_2^{(n)}, \ldots, \boldsymbol{\mu}_K^{(n)}$ are linearly independent. Suppose wlog that $\{R_2^{(n)}, \ldots, R_K^{(n)}\} \subset \{P_1, \ldots, P_{K-1}\}$. Then, by the inductive hypothe-

Algorithm 30 NonSquareDemix($\tilde{P}_1, \ldots, \tilde{P}_M$)

1: $R_1, \ldots, R_L \longleftarrow$ independently uniformly distributed elements in $\operatorname{conv}(\tilde{P}_1, \ldots, \tilde{P}_M)$

2: $(Q_1, \ldots, Q_L)^T \longleftarrow \operatorname{Demix}(R_1, \ldots, R_L)$

3: **return** $(Q_1, \ldots, Q_L)^T$

sis, if $(Q_1, \ldots, Q_{K-1}) \longleftarrow \operatorname{Demix}(R_2^{(n)}, \ldots, R_K^{(n)})$, then $(Q_1, \ldots, Q_{K-1})$ is a permutation of $(P_1, \ldots, P_{K-1})$. Note that $\frac{1}{K}\sum_{i=1}^{K} S_i \in \operatorname{conv}(P_1, \ldots, P_K)^\circ$ since $\mathbf{\Gamma}$ has full rank by assumption.

Write $Q_i = \boldsymbol{\rho}_i^T \boldsymbol{P}$ for $i \in [K]$. Then, there exists of a permutation $\sigma : [K-1] \longrightarrow [K-1]$ such that $\boldsymbol{\rho}_i = \boldsymbol{e}_{\sigma(i)}$. Since $\boldsymbol{\rho}_K \in \Xi^\circ$ and $\boldsymbol{\rho}_i = \boldsymbol{e}_{\sigma(i)}$ for $i \leqslant K-1$, the conditions in statement *1* of Lemma 34 are satisfied. Therefore, by Lemma 34, the residue of $\boldsymbol{\rho}_K$ with respect to $\{\boldsymbol{\rho}_1, \ldots, \boldsymbol{\rho}_{K-1}\}$ is $\boldsymbol{e}_K$. Then, by statement *3* of Proposition 6, the residue of Q_K with respect to $\{Q_1, \ldots, Q_{K-1}\}$ is P_K. This completes the inductive step.

□

The Non-Square Demix Algorithm

Now, we examine the non-square case of the demixing problem ($M > L$). Note that knowledge of L is needed since one must resample exactly L distributions in order to run the square Demix algorithm.

Corollary 2. *Suppose $M > L$. Let $P_1, \ldots, P_L$ be jointly irreducible and $\mathbf{\Pi}$ have full rank. Then, with probability 1, NonSquareDemix($\tilde{P}_1, \ldots, \tilde{P}_M$) returns $(Q_1, \ldots, Q_L)$ such that $(Q_1, \ldots, Q_L)$ is a permutation of $(P_1, \ldots, P_L)$.*

Proof. We can write $R_i = \boldsymbol{\tau}_i^T \boldsymbol{P}$ where $\boldsymbol{\tau}_i \in \Delta_L$ and $i = 1, \ldots, L$. $\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_L$ are drawn uniformly independently from a set with positive $(L-1)$-dimensional Lebesgue measure since $\mathbf{\Pi}$ has full rank by hypothesis. By Lemma 31, $\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_L$ are linearly independent with probability 1. Then, by Theorem 26, with probability 1, Demix($R_1, \ldots, R_L$) returns a permutation of $(P_1, \ldots, P_L)$. □

5.10.4 Classification with Partial Labels

In this section, we present our identification result for classification with partial labels, i.e., Theorem 27. To begin, in Section 5.10.4, we prove an important lemma for the main subroutine of the algorithm PartialLabel: VertexTest (algorithm 24). Second, in Section 5.10.4, we present the proof of Theorem 27.

VertexTest Algorithm

Lemma 35 establishes that the VertexTest algorithm determines whether one vector of distributions is a permutation of another vector of distributions.

Lemma 35. *Let* $\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_L \in \Delta_L$ *and* $Q_i = \boldsymbol{\eta}_i^T \boldsymbol{P}$ *for* $i \in [L]$ *and* $\boldsymbol{Q} = (Q_1, \ldots, Q_L)^T$. *Suppose that* $P_1, \ldots, P_L$ *are jointly irreducible,* $\boldsymbol{\Pi}$ *has full column rank, and the columns of* $\boldsymbol{\Pi}^+$ *are unique. Then, VertexTest*$(\boldsymbol{\Pi}^+, \tilde{\boldsymbol{P}}, \boldsymbol{Q})$ *returns* $(1, \boldsymbol{C}^T)$ *with* $\boldsymbol{C}$ *a permutation matrix if and only if* $\boldsymbol{Q}$ *is a permutation of* $\boldsymbol{P}$. *Further, if VertexTest*$(\boldsymbol{\Pi}^+, \tilde{\boldsymbol{P}}, \boldsymbol{Q})$ *returns* $(1, \boldsymbol{C}^T)$, *then* $\boldsymbol{C}^T \boldsymbol{Q} = \boldsymbol{P}$.

Proof. If $\boldsymbol{Q} = (Q_1, \ldots, Q_L)^T$ is such that $\boldsymbol{D}^T \boldsymbol{Q} = \boldsymbol{P}$ where $\boldsymbol{D}$ is a permutation matrix, then it is clear that VertexTest$(\boldsymbol{\Pi}^+, (\tilde{P}_1, \ldots, \tilde{P}_M)^T, (Q_1, \ldots, Q_L)^T)$ returns $(1, \boldsymbol{C}^T)$ for some permutation matrix $\boldsymbol{C}$ since the entries of $\boldsymbol{\Pi}^+$ are $\boldsymbol{\Pi}^+_{i,j} = \mathbf{1}\{\kappa^*(\tilde{P}_i \,|\, P_j) > 0\}$. Since $\boldsymbol{D}^T \boldsymbol{Q} = \boldsymbol{P}$, clearly, $\boldsymbol{Z}\boldsymbol{D} = \boldsymbol{\Pi}^+$. But since the columns of $\boldsymbol{\Pi}^+$ are unique, there is a unique permutation of the columns of $\boldsymbol{Z}$ to obtain the columns of $\boldsymbol{\Pi}^+$. Therefore, $\boldsymbol{D} = \boldsymbol{C}$.

Consider the "only if" direction. We use the notation from Algorithm 24. Suppose Algorithm 24 has returned $(1, \boldsymbol{C}^T)$ where $\boldsymbol{C}$ is a permutation matrix. W.l.o.g. (reordering the Q_i) we can assume that $\boldsymbol{C}$ is the identity and thus $\boldsymbol{Z} = \boldsymbol{\Pi}^+$.

In the sequel denote $\phi(x) := \mathbf{1}\{x > 0\}$ and $\phi(\boldsymbol{M})$ the entry-wise application of ϕ to the matrix or vector $\boldsymbol{M}$. We denote $\boldsymbol{v} \preceq \boldsymbol{w}$ when all entries of $\boldsymbol{v}$ are less than or equal to the corresponding entries of $\boldsymbol{w}$ (where $\boldsymbol{v}$ and $\boldsymbol{w}$ are vectors). This is a partial order, which will be used only for $0-1$ vectors below (essentially to denote support inclusion). W.l.o.g. (reordering the P_i) we can assume that the columns of $\boldsymbol{\Pi}^+$ are reordered in some sequence compatible with $\preceq$ in decreasing order, i.e. such that if $\boldsymbol{\Pi}^+_{:,j} \preceq \boldsymbol{\Pi}^+_{:,i}$, then $i \leqslant j$.

Introduce the following additional notation: let $\mathbf{\Lambda}$ be the matrix with rows $\mathbf{\Lambda}_{i,:} = \phi(\boldsymbol{\eta}_i^T)$. Observe that by statement 3 of Lemma 30, for any i, j, k, $\kappa^*(\tilde{P}_i|Q_j) > 0$ and $\kappa^*(Q_j|P_k) > 0$ implies $\kappa^*(\tilde{P}_i|P_k) > 0$. Note that we can write $\mathbf{\Lambda}_{j,k} = \mathbf{1}\{\kappa^*(Q_j \,|\, P_k) > 0\}$ and $\mathbf{\Pi}^+_{j,k} = \mathbf{1}\{\kappa^*(\tilde{P}_j \,|\, P_k) > 0\}$. Thus, we must have $\phi(\boldsymbol{Z}\mathbf{\Lambda}) \preceq \mathbf{\Pi}^+$.

We now argue that this implies that $\mathbf{\Lambda}$ is sub-diagonal, i.e., $\mathbf{\Lambda}_{ij} = 0$ for $i < j$. Let $i < j$. If $\mathbf{\Lambda}_{ij} > 0$, then $\boldsymbol{Z}_{:,i} \preceq \mathbf{\Pi}^+_{:,j}$ by the above relation. Since $\boldsymbol{Z} = \mathbf{\Pi}^+$, this implies $\mathbf{\Pi}^+_{:,i} \preceq \mathbf{\Pi}^+_{:,j}$, which implies $j \leqslant i$ by the assumed ordering of the columns of $\mathbf{\Pi}^+$, a contradiction. Hence $\mathbf{\Lambda}_{ij} = 0$ for $i < j$.

Now, since the matrix $\boldsymbol{Y}$ (line 1 of Algorithm 24) is diagonal, Statement 3 of Lemma 1 gives that for any $i \neq j$ we have $\mathbf{\Lambda}_{i,:} \not\preceq \mathbf{\Lambda}_{j,:}$. One can conclude by a straightforward recursion that since $\mathbf{\Lambda}$ is sub-diagonal, this implies that $\mathbf{\Lambda}$ is in fact diagonal. Start with the first row $\mathbf{\Lambda}_{1,:}$ which must be $(1, 0, \ldots, 0)^T$ (by sub-diagonality). Since $\mathbf{\Lambda}_{1,:} \not\preceq \mathbf{\Lambda}_{j,:}$ for $j > 1$, this implies the first column $\mathbf{\Lambda}_{:,1}$ is also $(1, 0, \ldots, 0)$. The subsequent columns/rows are handled in the same way.

Hence $\mathbf{\Lambda}$ is the identity, which implies that $\boldsymbol{Q} = \boldsymbol{P}$. □

Proof of Theorem 27

Proof. We adopt the notation from the description of Algorithm 22 with the exception that we make explicit the dependence on k by writing $W_i^{(k)}$ instead of W_i and $\bar{Q}_i^{(k)}$ instead of $\bar{Q}_i$. We show that there is a K such that for all $k \geqslant K$, $(W_1^{(k)}, \ldots, W_L^{(k)})^T$ is a permutation of $(P_1, \ldots, P_L)^T$. Then, the result will follow from Lemma 35.

Let $Q_i = \boldsymbol{\tau}_i^T \boldsymbol{P}$, $\bar{Q}_i^{(k)} = \bar{\boldsymbol{\tau}}_i^{(k)^T} \boldsymbol{P}$, and $W_i^{(k)} = \boldsymbol{\gamma}_i^{(k)^T} \boldsymbol{P}$. Further, let $0 \leqslant n < L$, $\{i_1, \ldots, i_n\} \subset [L]$, $l \neq j \in [L]$, and define the following events wrt the randomness of $\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_L$:

$$
\begin{aligned}
E_{i_1,\ldots,i_n} &= \{\boldsymbol{e}_{i_1}, \ldots, \boldsymbol{e}_{i_n}, \boldsymbol{\tau}_{n+1}, \ldots, \boldsymbol{\tau}_L \text{ are linearly independent}\} \\
E &= \cap_{\{i_1,\ldots,i_n\} \subset [L], 0 \leqslant n < L} E_{i_1,\ldots,i_n} \\
F_{l,j} &= \{\boldsymbol{e}_l - \boldsymbol{e}_j, \boldsymbol{\tau}_2, \ldots, \boldsymbol{\tau}_L \text{ are linearly independent}\} \\
F &= \cap_{l \neq j \in [L]} F_{l,j} \\
G^{l,j}_{i_1,\ldots,i_{n-1}} &= \{\boldsymbol{e}_l - \boldsymbol{e}_j, \boldsymbol{e}_{i_1}, \ldots, \boldsymbol{e}_{i_{n-1}}, \boldsymbol{\tau}_{n+1}, \ldots, \boldsymbol{\tau}_L \text{ are linearly independent}\} \\
G &= \cap_{\{i_1,\ldots,i_{n-1}\} \subset [L], 0 \leqslant n < L, l \neq j \in [L] \setminus \{i_1,\ldots,i_{n-1}\}} G^{l,j}_{i_1,\ldots,i_{n-1}}.
\end{aligned}
$$

By Lemma 31, for any $0 \leqslant n < L$, $\{i_1, \ldots, i_n\} \subset [L]$, the event $E_{i_1,\ldots,i_n}$ occurs with probability 1. Similarly, by Lemma 31, for any $l \neq j \in [L]$, the event $F_{l,j}$ occurs with probability 1. Finally, by Lemma 31, for any $0 \leqslant n < L$, $\{i_1, \ldots, i_{n-1}\} \subset [L]$ and any $l \neq j \in [L] \backslash \{i_1, \ldots, i_{n-1}\}$, the event $G^{l,j}_{i_1,\ldots,i_{n-1}}$ occurs with probability 1. Hence, the event $E \cap F \cap G$ occurs with probability 1. For the remainder of the proof, assume event $E \cap F \cap G$ occurs.

We prove the claim inductively. We show that for all $n \leqslant L$ there exists K_n such that if $k \geqslant K_n$, then $W_1^{(k)}, \ldots, W_n^{(k)}$ are distinct base distributions.

Base Case: $n = 1$**.** We will apply Lemma 34. By event E, $\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_L$ are linearly independent. Therefore, $\text{aff}(\boldsymbol{\tau}_2, \ldots, \boldsymbol{\tau}_L)$ gives a hyperplane with an associated open halfspace $\boldsymbol{H}$ that contains $\boldsymbol{\tau}_1$ and at least one $\boldsymbol{e}_j$. Inspection of Line 3 of Algorithm 23 shows that $\bar{\boldsymbol{\tau}}_1^{(k)}$ is simply the average of $\boldsymbol{\tau}_2, \ldots, \boldsymbol{\tau}_L$ and does not depend on k. Thus, there exists K_1 such that for all $k \geqslant K_1$, if $\boldsymbol{e}_j \in \boldsymbol{H}$, then $\boldsymbol{\lambda}_k \coloneqq \frac{1}{k}\boldsymbol{\tau}_1 + \frac{k-1}{k}\bar{\boldsymbol{\tau}}_1^{(k)} \in \text{conv}(\boldsymbol{e}_j, \boldsymbol{\tau}_2, \ldots, \boldsymbol{\tau}_L)^\circ$. Fix $k \geqslant K_1$. Then, by event E, for all $\boldsymbol{e}_j \in \boldsymbol{H}$, there exists a unique $\kappa_j > 0$ and unique $a_{j,2}, \ldots, a_{j,L}$ such that

$$\begin{aligned}\boldsymbol{\lambda}_k &= \kappa_j \boldsymbol{e}_j + \sum_{i=2}^{L} a_{j,i}\boldsymbol{\tau}_i \\ &= \kappa_j \boldsymbol{e}_j + (1-\kappa_j)\tilde{\boldsymbol{\tau}}_j\end{aligned}$$

where $\tilde{\boldsymbol{\tau}}_j \in \text{conv}(\boldsymbol{\tau}_2, \ldots, \boldsymbol{\tau}_L)$ is unique. We claim that for all $i \neq j$ and $\{\boldsymbol{e}_i, \boldsymbol{e}_j\} \subset \boldsymbol{H}$, $\kappa_i \neq \kappa_j$. Suppose to the contrary that there is $i \neq j$ such that $\{\boldsymbol{e}_i, \boldsymbol{e}_j\} \subset \boldsymbol{H}$ and $\kappa_i = \kappa_j = \kappa$. Then,

$$\begin{aligned}\boldsymbol{\lambda}_k &= \kappa \boldsymbol{e}_i + (1-\kappa)\tilde{\boldsymbol{\tau}}_i \\ \boldsymbol{\lambda}_k &= \kappa \boldsymbol{e}_j + (1-\kappa)\tilde{\boldsymbol{\tau}}_j.\end{aligned}$$

Then, $(1-\kappa)(\tilde{\boldsymbol{\tau}}_j - \tilde{\boldsymbol{\tau}}_i) - \kappa(\boldsymbol{e}_i - \boldsymbol{e}_j) = 0$, from which it follows that $\boldsymbol{e}_i - \boldsymbol{e}_j \in \text{span}(\boldsymbol{\tau}_2, \ldots, \boldsymbol{\tau}_L)$. But, by event F, $\boldsymbol{e}_i - \boldsymbol{e}_j, \boldsymbol{\tau}_2, \ldots, \boldsymbol{\tau}_L$ are linearly independent and, hence, we have a contradiction. Thus, the claim follows.

Consequently, there is a unique j that minimizes κ_j. Note that for all $\boldsymbol{e}_i \notin \boldsymbol{H}$, if we write $\boldsymbol{e}_i = \sum_{l \geqslant 2} a_l \boldsymbol{\tau}_l + a_1 \boldsymbol{\lambda}_k$, then $a_1 \leqslant 0$. Then, by Lemma 34, $\boldsymbol{e}_j$ is the residue of $\boldsymbol{\lambda}_k$ with respect to $\boldsymbol{\tau}_2, \ldots, \boldsymbol{\tau}_L$. Therefore, by Proposition 6, $\text{MultiResidue}(\frac{1}{k}Q_1 + (1-\frac{1}{k})\bar{Q}_1 \,|\, \{Q_j\}_{j>1})$ is well-defined and if $W_1^{(k)} \longleftarrow \text{MultiResidue}(\frac{1}{k}Q_1 + (1-\frac{1}{k})\bar{Q}_1 \,|\, \{Q_j\}_{j>1})$, $W_1^{(k)}$ is one of the base distributions. This establishes the base case.

The Inductive Step: The proof is similar to the base case. Suppose that there exists K_{n-1} such that for all $k \geqslant K_{n-1}$, $W_1^{(k)}, \ldots, W_{n-1}^{(k)}$ are distinct base distributions. Let $\{i_1, \ldots, i_{n-1}\} \subset [L]$ denote the indices of the base distributions that are equal to $W_1^{(k)}, \ldots, W_{n-1}^{(k)}$ under the inductive hypothesis. By the event E, $\boldsymbol{e}_{i_1}, \ldots, \boldsymbol{e}_{i_{n-1}}, \boldsymbol{\tau}_n, \ldots, \boldsymbol{\tau}_L$ are linearly independent. Hence, $\text{aff}(\boldsymbol{e}_{i_1}, \ldots, \boldsymbol{e}_{i_{n-1}}, \boldsymbol{\tau}_{n+1}, \ldots, \boldsymbol{\tau}_L)$ gives a hyperplane with an associated open halfspace $\boldsymbol{H}_{i_1,\ldots,i_{n-1}}$ such that $\boldsymbol{\tau}_n \in \boldsymbol{H}_{i_1,\ldots,i_{n-1}}$. We claim that there is $\boldsymbol{e}_j \notin \{\boldsymbol{e}_{i_1}, \ldots, \boldsymbol{e}_{i_{n-1}}\}$ such that $\boldsymbol{e}_j \in \boldsymbol{H}_{i_1,\ldots,i_{n-1}}$. Suppose not. Then, $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_L \in \boldsymbol{H}^c_{i_1,\ldots,i_{n-1}}$ and $\boldsymbol{\tau}_n \in \boldsymbol{H}_{i_1,\ldots,i_{n-1}}$, which implies that $\boldsymbol{\tau}_n \notin \Delta_{L-1}$. This is a contradiction, so the claim follows.

Define

$$\boldsymbol{\lambda}_k^{(i_1,\ldots,i_{n-1})} \coloneqq \frac{1}{k}\boldsymbol{\tau}_n + [\frac{k-1}{k}]\frac{1}{L-1}(\sum_{s>n}\boldsymbol{\tau}_s + \sum_{s<n}\boldsymbol{e}_{i_s}).$$

There exists an integer $K_n^{(i_1,\ldots,i_{n-1})}$ such that if $k \geqslant K_n^{(i_1,\ldots,i_{n-1})}$, then for all $\boldsymbol{e}_j \in \boldsymbol{H}_{i_1,\ldots,i_{n-1}}$, $\boldsymbol{\lambda}_k^{(i_1,\ldots,i_{n-1})} \in \text{conv}(\boldsymbol{e}_j, \boldsymbol{e}_{i_1}, \ldots, \boldsymbol{e}_{i_{n-1}}, \boldsymbol{\tau}_{n+1}, \ldots, \boldsymbol{\tau}_L)^\circ$. Set

$$K_n \coloneqq \max(\max_{\{i_1,\ldots,i_{n-1}\}\subset[L]}(K_n^{(i_1,\ldots,i_{n-1})}), K_{n-1}).$$

Fix $k \geqslant K_n$. Define

$$\begin{aligned}\boldsymbol{\lambda}_k &\coloneqq \frac{1}{k}\boldsymbol{\tau}_n + [\frac{k-1}{k}]\bar{\boldsymbol{\tau}}_n^{(k)} \\ &= \frac{1}{k}\boldsymbol{\tau}_n + [\frac{k-1}{k}]\frac{1}{L-1}(\sum_{s>n}\boldsymbol{\tau}_s + \sum_{s<n}\boldsymbol{\gamma}_s^{(k)}).\end{aligned}$$

By the inductive hypothesis, $k \geqslant K_{n-1}$, and Proposition 6, there exists $\{i_1, \ldots, i_{n-1}\} \subset [L]$ such that $\boldsymbol{\gamma}_j^{(k)} = \boldsymbol{e}_{i_j}$ for all $j \in [n-1]$. For the sake of abbreviation, let $\boldsymbol{H} = \boldsymbol{H}_{i_1,\ldots,i_{n-1}}$. Thus, $\boldsymbol{\tau}_n \in \boldsymbol{H}$ and there exists $\boldsymbol{e}_j \in \boldsymbol{H}$ such that $\boldsymbol{e}_j \notin \{\boldsymbol{e}_{i_1}, \ldots, \boldsymbol{e}_{i_{n-1}}\}$. Hence, by our choice of K_n, for every $\boldsymbol{e}_j \in \boldsymbol{H}$

$$\boldsymbol{\lambda}_k = \boldsymbol{\lambda}_k^{(i_1,\ldots,i_{n-1})} \in \text{conv}(\boldsymbol{e}_j, \boldsymbol{e}_{i_1}, \ldots, \boldsymbol{e}_{i_{n-1}}, \boldsymbol{\tau}_{n+1}, \ldots, \boldsymbol{\tau}_L)^\circ.$$

By event E for all $\boldsymbol{e}_j \in \boldsymbol{H}$, there is a unique $\kappa_j > 0$ and unique $a_{j,1}, \ldots, a_{j,n-1}, a_{j,n+1}, \ldots, a_{j,L} > 0$ such that

$$\begin{aligned}\boldsymbol{\lambda}_k &= \kappa_j\boldsymbol{e}_j + \sum_{l<n}a_{j,i}\boldsymbol{e}_{i_l} + \sum_{l>n}a_{j,l}\boldsymbol{\tau}_l \\ &= \kappa_j\boldsymbol{e}_j + (1-\kappa_j)\tilde{\boldsymbol{\tau}}_j\end{aligned}$$

where $\tilde{\boldsymbol{\tau}}_j \in \text{conv}(\boldsymbol{e}_{i_1}, \ldots, \boldsymbol{e}_{i_{n-1}}, \boldsymbol{\tau}_{n+1}, \ldots, \boldsymbol{\tau}_L)$ is unique.

We claim that for all $l \neq j$ such that $\{\boldsymbol{e}_l, \boldsymbol{e}_j\} \subset \boldsymbol{H}$, $\kappa_l \neq \kappa_j$. Suppose to the contrary that there exists $l \neq j$ such that $\{\boldsymbol{e}_l, \boldsymbol{e}_j\} \subset \boldsymbol{H}$ and $\kappa_l = \kappa_j = \kappa$. Then,

$$\boldsymbol{\lambda}_k = \kappa \boldsymbol{e}_l + (1-\kappa)\tilde{\boldsymbol{\tau}}_i = \kappa \boldsymbol{e}_j + (1-\kappa)\tilde{\boldsymbol{\tau}}_j.$$

This implies that $\boldsymbol{e}_l - \boldsymbol{e}_j \in \text{span}(\boldsymbol{e}_{i_1}, \ldots, \boldsymbol{e}_{i_{n-1}}, \boldsymbol{\tau}_{n+1}, \ldots, \boldsymbol{\tau}_L)$. Observe that $\{\boldsymbol{e}_l, \boldsymbol{e}_j\} \subset \boldsymbol{H}$ implies that $\boldsymbol{e}_l \notin \{\boldsymbol{e}_{i_1}, \ldots, \boldsymbol{e}_{i_{n-1}}\}$ and $\boldsymbol{e}_j \notin \{\boldsymbol{e}_{i_1}, \ldots, \boldsymbol{e}_{i_{n-1}}\}$. Thus, event G implies that $\boldsymbol{e}_l - \boldsymbol{e}_j, \boldsymbol{e}_{i_1}, \ldots, \boldsymbol{e}_{i_{n-1}}, \boldsymbol{\tau}_{n+1}, \ldots, \boldsymbol{\tau}_L$ are linearly independent. Therefore, we have a contradiction, establishing the claim.

Consequently, there is a unique j that minimizes κ_j. Note that for all $\boldsymbol{e}_l \notin \boldsymbol{H}$, if we write $\boldsymbol{e}_l = \sum_{m<n} a_m \boldsymbol{e}_{i_m} + \sum_{m>n} a_m \boldsymbol{\tau}_m + a_n \boldsymbol{\lambda}_k$, then $\boldsymbol{a}_n \leqslant 0$. Then, by Lemma 34, $\boldsymbol{e}_j$ is the residue of $\boldsymbol{\lambda}_k$ with respect to $\boldsymbol{\gamma}_1^{(k)}, \ldots, \boldsymbol{\gamma}_{n-1}^{(k)}, \boldsymbol{\tau}_{n+1}, \ldots, \boldsymbol{\tau}_L$. Therefore, by Proposition 6, MultiResidue$(\frac{1}{k}Q_n + (1-\frac{1}{k})\bar{Q}_n \,|\, \{Q_j\}_{j>n} \cup \{W_j^{(k)}\}_{j<n})$ is well-defined and if $W_n \longleftarrow$ MultiResidue$(\frac{1}{k}Q_n + (1-\frac{1}{k})\bar{Q}_n \,|\, \{Q_j\}_{j>n} \cup \{W_j^{(k)}\}_{j<n})$, $W_n^{(k)}$ is one of the base distributions. Since $\boldsymbol{e}_j \in \boldsymbol{H}$ implies that $\boldsymbol{e}_j \notin \{\boldsymbol{e}_{i_1}, \ldots, \boldsymbol{e}_{i_{n-1}}\}$, it follows that $W_1^{(k)}, \ldots, W_n^{(k)}$ are distinct base distributions. This establishes the inductive step.

The result follows from applying Lemma 35. □

5.11 Estimation

In this section, we present the estimation results of our paper. To begin, in Section 5.11.1, we present the proof of sufficient conditions under which ResidueHat estimators converge uniformly in probability (Proposition 7). Second, in Section 5.11.2, we prove our main estimation result for demixing mixed membership models (Theorem 29). Finally, in Section 5.11.3, we prove our main estimation result for classification with partial labels (Theorem 30).

5.11.1 ResidueHat Results

Let $A_1, A_2, \ldots$ denote positive constants whose values may change from line to line. We introduce the following definitions.

Definition 18. *Let $\widehat{F}$ and $\widehat{H}$ be ResidueHat estimators of F and H, respectively, where $F \neq H$ and let $G \longleftarrow Residue(F \mid H)$ and $\widehat{G} \longleftarrow ResidueHat(\widehat{F} \mid \widehat{H})$. If $\widehat{G}$ is a ResidueHat estimator of order 0, we say its* distributional ancestors *are $\{F, H\}$ and define $ancestors(\widehat{G}) \coloneqq \{F, H\}$. If $\widehat{G}$ is a ResidueHat estimator of the kth order, we define its distributional ancestors to be $ancestors(\widehat{G}) = ancestors(\widehat{F}) \cup ancestors(\widehat{H})$.*

The constants in our bounds depend on the distributional ancestors.

Definition 19. *We say that the distribution F satisfies the* support condition ***(SC)*** *with respect to H if there exists a distribution G and $\gamma \in [0, 1)$ such that $\operatorname{supp}(H) \nsubseteq \operatorname{supp}(G)$ and $F = (1 - \gamma)G + \gamma H$.*

Definition 20. *If*

$$\sup_{E \in \mathcal{E}} |\widehat{F}(E) - F(E)| \xrightarrow{i.p.} 0$$

as $\boldsymbol{n} \longrightarrow \infty$, we say that $\widehat{F} \longrightarrow F$ uniformly (or $\widehat{F}$ converges uniformly to F) with respect to $\mathcal{E}$.

Definition 21. *Let $\widehat{F}$ be a ResidueHat estimator of a distribution F. We say that $\widehat{F}$ satisfies a Uniform Deviation Inequality* ***(UDI)*** *with respect to $\mathcal{E}$ if for all $\epsilon > 0$, there exist constants $A_{1,\epsilon}, A_{2,\epsilon} > 0$ and $\boldsymbol{N}$ depending on $ancestors(\widehat{F})$ such that if $\boldsymbol{n} \geqslant \boldsymbol{N}$, then for all $E \in \mathcal{E}$*

$$|\widehat{F}(E) - F(E)| < A_{1,\epsilon}\gamma_{\boldsymbol{n}} + \epsilon$$

with probability at least $1 - A_{2,\epsilon} \sum_{i \in [L]} \frac{1}{n_i}$

Henceforth, for the purposes of abbreviation, we will only say that a ResidueHat estimator satisfies a Uniform Deviation Inequality **(UDI)** and omit "with respect to $\mathcal{E}$" because the context makes this clear.

Definition 22. *Let $\widehat{F}$ and $\widehat{H}$ be ResidueHat estimators. We say that $\widehat{\kappa}(\widehat{F} \mid \widehat{H})$ satisfies a Rate of Convergence* ***(RC)*** *with respect to $\mathcal{E}$ if for all $\epsilon > 0$, there exists constants $A_{1,\epsilon}, A_{2,\epsilon} > 0$ and $\boldsymbol{N}$ depending on $ancestors(\widehat{F}) \cup ancestors(\widehat{H})$ such that for $\boldsymbol{n} \geqslant \boldsymbol{N}$,*

$$|\widehat{\kappa}(\widehat{F} \mid \widehat{H}) - \kappa^*(F \mid H)| \leqslant A_{1,\epsilon}\gamma_{\boldsymbol{n}} + \epsilon$$

with probability at least $1 - A_{2,\epsilon} \sum_{i \in [L]} \frac{1}{n_i}$.

Lemma 36 gives sufficient conditions under which F satisfies **(SC)** with respect to H.

Lemma 36. *Let $P_1, \ldots, P_L$ satisfy **(A″)** and let $F, H \in \text{conv}(P_1, \ldots, P_L)$ such that $F \neq H$. Then, F satisfies **(SC)** with respect to H.*

Proof. Let $A = \arg\min(|B| : B \subseteq \{P_1, \ldots, P_L\}, F, H \in \text{conv}(B))$. Without loss of generality, suppose that $A = \{P_1, \ldots, P_K\}$. F either lies on the boundary of $\text{conv}(P_1, \ldots, P_K)$ or doesn't. If F lies on the boundary of $\text{conv}(P_1, \ldots, P_K)$, then $H \in \text{conv}(P_1, \ldots, P_K)^\circ$ by minimality of A. Then, we pick $G = F$ and $\gamma = 0$ to obtain $F = (1-\gamma)F + \gamma H$. Since $P_1, \ldots, P_L$ satisfy **(A″)**, $\text{supp}(H) \nsubseteq \text{supp}(F)$.

Now, suppose that $F \in \text{conv}(P_1, \ldots, P_K)^\circ$. Let $G \longleftarrow \text{Residue}(F \,|\, H)$; we can write $F = (1-\gamma)G + \gamma H$ for $\gamma \in [0, 1)$ since $F \neq H$. Then, by Statement *2* of Lemma 30 and statement *3* of Proposition 6, G is on the boundary of $\text{conv}(P_1, \ldots, P_K)$. Without loss of generality, suppose that $G \in \text{conv}(P_1, \ldots, P_{K-1})$. Since $F = (1-\gamma)G + \gamma H \in \text{conv}(P_1, \ldots, P_K)^\circ$, and $G \in \text{conv}(P_1, \ldots, P_{K-1})$, $H \notin \text{conv}(P_1, \ldots, P_{K-1})$. Since $P_1, \ldots, P_L$ satisfy **(A″)**, $\text{supp}(H) \nsubseteq \text{supp}(G)$. This completes the proof. □

Lemma 37 gives sufficient conditions under which an estimator $\widehat{G}$ satisfies a **(UDI)**.

Lemma 37. *Let*

1. *F and H be distributions such that $F \neq H$,*
2. *$G \longleftarrow$ Residue$(F \,|\, H)$, and*
3. *$\widehat{G} \longleftarrow$ ResidueHat$(\widehat{F} \,|\, \widehat{H})$.*

*If $\widehat{\kappa}(\widehat{F} \,|\, \widehat{H})$ satisfies a **(RC)**, $\widehat{H}$ satisfies a **(UDI)**, and $\widehat{F}$ satisfies a **(UDI)**, then $\widehat{G}$ satisfies a **(UDI)**.*

Proof. For the sake of abbreviation, let $\widehat{\kappa} = \widehat{\kappa}(\widehat{F} \,|\, \widehat{H})$, $\kappa^* = \kappa^*(F \,|\, H)$, $\widehat{\alpha} = \frac{1}{1-\widehat{\kappa}}$ and $\alpha^* = \frac{1}{1-\kappa^*}$. Let $\epsilon > 0$. We claim that there are constants $A_{1,\epsilon}, A_{2,\epsilon} > 0$ such that for sufficiently large $\boldsymbol{n}$,

$$\Pr(|\widehat{\alpha} - \alpha^*| < A_{1,\epsilon}\gamma_{\boldsymbol{n}} + \epsilon) \geqslant 1 - A_{2,\epsilon} \sum_{i \in [L]} \frac{1}{n_i}. \tag{5.10}$$

Let $\delta = \frac{\epsilon(1-\kappa^*)^2}{2}$. Since $\widehat{\kappa}$ satisfies a **(RC)**, there exists constants $A_{1,\delta}, A_{2,\delta} > 0$ such that for large enough $\boldsymbol{n}$,

$$|\widehat{\kappa} - \kappa^*| \leqslant A_{1,\delta}\gamma_{\boldsymbol{n}} + \delta$$

with probability at least $1 - A_{2,\delta}\sum_{i\in[L]}\frac{1}{n_i}$. Since $F \neq H$, $\kappa^* < 1$ by Proposition 5, so we can let $\boldsymbol{n}$ large enough so that

$$\frac{1}{(1-\kappa^*)(1-\widehat{\kappa})} \leqslant 2\frac{1}{(1-\kappa^*)^2}$$

with high probability. Then, on this same event, for large enough $\boldsymbol{n}$,

$$\begin{aligned}
\left|\frac{1}{1-\kappa^*} - \frac{1}{1-\widehat{\kappa}}\right| &\leqslant \frac{A_{1,\delta}\gamma_{\boldsymbol{n}} + \delta}{(1-\kappa^*)(1-\widehat{\kappa})} \\
&\leqslant 2\frac{A_{1,\delta}\gamma_{\boldsymbol{n}} + \delta}{(1-\kappa^*)^2} \\
&\leqslant 2\frac{A_{1,\delta}\gamma_{\boldsymbol{n}}}{(1-\kappa^*)^2} + \epsilon.
\end{aligned}$$

Thus, we obtain the claim.

We can write $G = \alpha F + (1-\alpha)H$ with $\alpha \geqslant 1$. Then, by the triangle inequality,

$$\begin{aligned}
|\widehat{G} - G| &= |\widehat{\alpha}\widehat{F} + (1-\widehat{\alpha})\widehat{H} - \alpha F - (1-\alpha)H| \\
&\leqslant |\widehat{\alpha}\widehat{F} - \alpha F| + |(1-\widehat{\alpha})\widehat{H} - (1-\alpha)H| \\
&= |\widehat{\alpha}\widehat{F} - \widehat{\alpha}F + \widehat{\alpha}F - \alpha F| + |(1-\widehat{\alpha})\widehat{H} - (1-\widehat{\alpha})H + (1-\widehat{\alpha})H - (1-\alpha)H| \\
&\leqslant |\widehat{\alpha}||\widehat{F} - F| + |\widehat{\alpha} - \alpha| + |1-\widehat{\alpha}||\widehat{H} - H| + |\widehat{\alpha} - \alpha|.
\end{aligned}$$

Since $\widehat{F}$ satisfies a **(UDI)**, $\widehat{H}$ satisfies a **(UDI)**, inequality (5.10) holds, and $|\widehat{\alpha}|$ and $|1-\widehat{\alpha}|$ are bounded in probability, the result follows by an application of a union bound and picking the ϵs in the uniform deviation inequalities appropriately for each term. □

Lemma 38 gives sufficient conditions under which $\widehat{\kappa}$ satisfies **(RC)**.

Lemma 38. *Let F and H be distributions such that $F \neq H$. If*

- *F satisfies **(SC)** with respect to H,*
- *$\widehat{F}$ satisfies **(UDI)**, and*

- $\widehat{H}$ *satisfies* ***(UDI)***,

then $\widehat{\kappa}(\widehat{F} \,|\, \widehat{H})$ *satisfies* ***(RC)***.

Proof. For abbreviation, let $\kappa^* = \kappa(F \,|\, H)$ and $\widehat{\kappa} = \widehat{\kappa}(\widehat{F} \,|\, \widehat{H})$.

We first prove the upper bound. F satisfies **(SC)** with respect to H, so there exists a distribution G such that $F = (1-\gamma)G + \gamma H$ for some $\gamma \in [0, 1)$ and $\text{supp}(H) \nsubseteq \text{supp}(G)$. Therefore, we have that G is irreducible with respect to H and, by Proposition 5, $\kappa^* = \gamma$.

Let $\delta > 0$ (to be chosen later). Since by hypothesis $\widehat{F}$ and $\widehat{H}$ satisfy **(UDI)**, there exist constants $A_{1,\delta}, A_{2,\delta} > 0$ such that for large enough $\boldsymbol{n}$, with probability at least $1 - A_{1,\delta}[\sum_{i\in[L]} \frac{1}{n_i}]$, for all $E \in \mathcal{E}$,

$$|\widehat{F}(E) - F(E)| < A_{2,\delta}\gamma_{\boldsymbol{n}} + \delta \tag{5.11}$$

$$|\widehat{H}(E) - H(E)| < A_{2,\delta}\gamma_{\boldsymbol{n}} + \delta. \tag{5.12}$$

Without loss of generality, let $A_{1,\delta}, A_{2,\delta} > 1$.

Pick $R \in \mathcal{E}$ such that $H(R) > 0$. By inequality (5.12), there exists $\boldsymbol{N}_1$ such that $\boldsymbol{n} \geqslant \boldsymbol{N}_1$ implies that $\widehat{H}(R) - \gamma_{\boldsymbol{n}} > 0$ with high probability. This implies that for $\boldsymbol{n} \geqslant \boldsymbol{N}_1$, $\widehat{\kappa}$ is finite. Let $\epsilon > 0$. By definition of $\widehat{\kappa}$, there exists $E \in \mathcal{E}$ such that

$$\frac{\epsilon}{2} + \widehat{\kappa} \geqslant \frac{\widehat{F}(E) + \gamma_{\boldsymbol{n}}}{(\widehat{H}(E) - \gamma_{\boldsymbol{n}})_+}.$$

Since $\widehat{\kappa}$ is finite, we have that $\widehat{H}(E) > \gamma_{\boldsymbol{n}}$ and $H(E) > 0$. Then,

$$\begin{aligned}
\frac{\epsilon}{2} + \widehat{\kappa} &\geqslant \frac{\widehat{F}(E) + \gamma_{\boldsymbol{n}}}{\widehat{H}(E) - \gamma_{\boldsymbol{n}}} \\
&\geqslant \frac{F(E) - (A_{2,\delta} - 1)\gamma_{\boldsymbol{n}} - \delta}{H(E) + (A_{2,\delta} - 1)\gamma_{\boldsymbol{n}} + \delta} \\
&\geqslant \frac{\gamma H(E)}{H(E) + (A_{2,\delta} - 1)\gamma_{\boldsymbol{n}} + \delta} - \frac{(A_{2,\delta} - 1)\gamma_{\boldsymbol{n}}}{H(E) + (A_{2,\delta} - 1)\gamma_{\boldsymbol{n}} + \delta} - \frac{\delta}{H(E) + (A_{2,\delta} - 1)\gamma_{\boldsymbol{n}} + \delta} \\
&\geqslant \frac{\gamma H(E)}{H(E)} - \frac{(A_{2,\delta} - 1)\gamma_{\boldsymbol{n}} + \delta}{H(E)} - \frac{(A_{2,\delta} - 1)\gamma_{\boldsymbol{n}}}{H(E) + (A_{2,\delta} - 1)\gamma_{\boldsymbol{n}} + \delta} - \frac{\delta}{H(E) + (A_{2,\delta} - 1)\gamma_{\boldsymbol{n}} + \delta} \\
&\geqslant \kappa^* - 2\frac{(A_{2,\delta} - 1)\gamma_{\boldsymbol{n}}}{H(E)} - 2\frac{\delta}{H(E)}
\end{aligned}$$

where in the second to last inequality we used the elementary fact that if $a, b, c > 0$ and $a \leqslant b$, then $\frac{a}{b+c} \geqslant \frac{a}{b} - \frac{c}{b}$. Picking $\delta = \frac{H(E)\epsilon}{4}$, we obtain the upper bound.

The proof of the other direction of the inequality is very similar to the proof of Theorem 2 in Scott (2015). By hypothesis, F satisfies **(SC)** with respect to H, so there exists a distribution G such that $F = (1-\gamma)G + \gamma H$ for some $\gamma \in [0, 1)$ and $\text{supp}(H) \nsubseteq \text{supp}(G)$. Therefore, we have that G is irreducible with respect to H and, by Proposition 5, $\kappa^*(F \,|\, H) = \gamma$. For abbreviation, let $\kappa^* = \kappa^*(F \,|\, H)$ and $\widehat{\kappa} = \widehat{\kappa}(\widehat{F} \,|\, \widehat{H})$. Since $\text{supp}(H) \nsubseteq \text{supp}(G)$, there exists an open set O such that

$$\frac{F(O)}{H(O)} = (1-\gamma)\frac{G(O)}{H(O)} + \gamma = \kappa^*.$$

Then, since $\mathcal{E}$ contains a generating set for the standard topology on $\mathbb{R}^d$, there exists $E \in \mathcal{E}$ such that

$$\frac{F(E)}{H(E)} = \kappa^*.$$

Let $\delta > 0$ such that $\delta \leqslant \frac{1}{4}H(E)$. Since by hypothesis $\widehat{F}$ and $\widehat{H}$ satisfy **(UDI)**, there exist constants $A_{3,\delta}, A_{4,\delta} > 0$ such that for large enough $\boldsymbol{n}$, with probability at least $1 - A_{3,\delta}[\sum_{i\in[L]} \frac{1}{n_i}]$,

$$\begin{aligned}\widehat{\kappa} &\leqslant \frac{F(E) + A_{4,\delta}\gamma_{\boldsymbol{n}} + \delta}{(H(E) - A_{4,\delta}\gamma_{\boldsymbol{n}} - \delta)_+} \\ &\leqslant \frac{F(E) + \epsilon}{(H(E) - \epsilon)_+}\end{aligned}$$

where $\epsilon = 2A_{4,\delta}\gamma_{\boldsymbol{n}} + \delta$. The rest of the proof is identical to the proof of Theorem 2 from Scott (2015) and, therefore, we omit it. □

The following theorem gives sufficient conditions under which a ResidueHat estimator satisfies **(UDI)**. It is the basis of Proposition 7.

Lemma 39. *If $P_1, \ldots, P_L$ satisfy **(A″)** and $\widehat{G}$ is a ResidueHat estimator of order k of a distribution $G \in \text{conv}(P_1, \ldots, P_L)$, then $\widehat{G}$ satisfies **(UDI)**.*

Proof. Let $\widehat{G} \longleftarrow \text{ResidueHat}(\widehat{F} \,|\, \widehat{H})$ where $G \longleftarrow \text{ResidueHat}(F \,|\, H)$, $F \neq H$, $F, H \in \text{conv}(P_1, \ldots, P_L)$ and $\widehat{F}, \widehat{H}$ are ResidueHat estimators of F and H respectively. We use

induction on k. Suppose $k = 0$. Then, $\widehat{F}$ and $\widehat{H}$ are empirical distributions. Therefore, the VC inequality applies to $\widehat{F}$ and $\widehat{H}$. Consequently, $\widehat{F}$ and $\widehat{H}$ satisfy **(UDI)**. Since $P_1, \ldots, P_L$ satisfy **(A″)**, $F, H \in \text{conv}(P_1, \ldots, P_L)$ and $F \neq H$, by Lemma 36, F satisfies **(SC)** with respect to H. Then, by Lemma 38, $\widehat{\kappa}(\widehat{F} \,|\, \widehat{H})$ satisfies **(RC)**. Then, all of the assumptions of Lemma 37 are satisfied, so $\widehat{G}$ satisfies **(UDI)**. Note that $G \in \text{conv}(P_1, \ldots, P_L)$ by Proposition 6.

The inductive step ($k > 0$) follows by similar reasoning. The difference is that instead of applying the VC inequality to $\widehat{F}$ and $\widehat{H}$, we use the fact that $\widehat{F}$ and $\widehat{H}$ are ResidueHat estimators of order $k - 1$ and, therefore, satisfy **(UDI)** by the inductive hypothesis. □

Proof of Proposition 7. Let $0 < \delta < \epsilon$. By Lemma 39, $\widehat{G}$ satisfies **(UDI)**. Consequently, there exist constants $A_{1,\delta}, A_{2,\delta} > 0$ such that for large enough $\boldsymbol{n}$ with probability at least $1 - A_{1,\delta} \sum_{i \in [L]} \frac{1}{n_i}$, $\widehat{G}$ satisfies for every $E \in \mathcal{E}$,

$$|\widehat{G}(E) - G(E)| \leqslant A_{2,\delta}\gamma_{\boldsymbol{n}} + \delta = A_{2,\delta} \sum_{i \in [L]} \epsilon_i(\frac{1}{n_i}) + \delta \longrightarrow \delta < \epsilon.$$

□

5.11.2 Demixing Mixed Membership Models

In this section, we prove our main estimation result for demixing mixed membership models, i.e., Theorem 29. First, in Section 5.11.2, we present an important lemma for FaceTestHat. Second, in Section 5.11.2, we present an empirical version of Demix and prove Theorem 29.

The FaceTestHat Algorithm

The following establishes that FaceTestHat behaves as desired.

Lemma 40. *Let $\epsilon \in (0, 1)$. For all $j \in [K]$, let $Q_j = \boldsymbol{\eta}_j^T \boldsymbol{P}$ and $\boldsymbol{\eta}_j \in \Delta_K$ such that every $\boldsymbol{\eta}_j$ lies in the relative interior of the same face of Δ_K. Let $P_1, \ldots, P_K$ satisfy **(A″)**, and $Q_1, \ldots, Q_K \in \text{conv}(P_1, \ldots, P_K)$ be distinct. Let $\widehat{Q}_i$ be a ResidueHat estimate of Q_i $\forall i \in [K]$.*

1. *With probability tending to 1 as $\boldsymbol{n} \longrightarrow \infty$, if FaceTestHat($\widehat{Q}_1, \cdots, \widehat{Q}_K \,|\, \epsilon$) returns 1, then $\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_K$ are in the relative interior of the same face.*

2. *Let* $\kappa^*_{i,j} = \kappa^*(Q_i \,|\, Q_j)$. *If* $\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_K$ *are in the relative interior of the same face and* $\min_{i,j} \kappa^*_{i,j} > \epsilon$, *then with probability tending to* 1 *as* $\boldsymbol{n} \longrightarrow \infty$, *FaceTestHat*$(\widehat{Q}_1, \cdots, \widehat{Q}_K \,|\, \epsilon)$ *returns* 1.

Proof. Let $\epsilon > 0$, $\kappa^*_{i,j} = \kappa^*(Q_i \,|\, Q_j)$ and $\widehat{\kappa}_{i,j} = \widehat{\kappa}(\widehat{Q}_i \,|\, \widehat{Q}_j)$. Since $P_1, \ldots, P_K$ satisfy **(A″)** and $Q_i \neq Q_j$, by Lemma 36, Q_i satisfies **(SC)** wrt Q_j. Since $\widehat{Q}_i$ and $\widehat{Q_j}$ are ResidueHat estimators, $\widehat{Q}_i$ and $\widehat{Q}_j$ satisfy **(UDI)** (Lemma 39). Then, by Lemma 38 $\widehat{\kappa}_{i,j}$ satisfies **(RC)**.

1. We prove the contrapositive. Suppose that $Q_1, \ldots, Q_K$ are not in the relative interior of the same face. Then, by Proposition 11, FaceTest$(Q_1, \ldots, Q_K)$ returns 0, which occurs if and only if there exist $i \neq j$ such that $\kappa^*_{i,j} = 0$. Since $\widehat{\kappa}_{i,j}$ satisfies **(RC)**, as $\boldsymbol{n} \longrightarrow \infty$, with probability tending to 1, $\widehat{\kappa}_{i,j} \longrightarrow 0$. This completes the proof.

2. If $\min_{i,j} \kappa^*_{i,j} > \epsilon$, then as $\boldsymbol{n} \longrightarrow \infty$, with probability tending to 1, $\min_{i,j} \widehat{\kappa}_{i,j} > \epsilon$.

□

The DemixHat Algorithm

The DemixHat algorithm (see Algorithm 26) differs from the Demix algorithm in that *(i)* it requires the specification of a constant $\epsilon \in (0, 1)$ and *(ii)* it only uses the two-sample κ^* operator. In the interest of clarity, we state the population version of the algorithm DemixHat, which we call Demix2. The only difference between Demix and Demix2 is that line 7 in Demix has been replaced with lines 6-8 in Demix2.

Lemma 41 establishes that it is possible to replace line 7 of the Algorithm 19 with the sequence of applications of the two-sample κ^* in lines 6-8 of Algorithm 31, without changing the conclusion of Theorem 26.

Lemma 41. *Let* $\{i_1, \ldots, i_K\} \subset [L]$ *be distinct indices. Let* $P_1, \ldots, P_L$ *be jointly irreducible,* $(Q_1, \ldots, Q_{K-1})$ *be a permutation of* $(P_{i_1}, \ldots, P_{i_{K-1}})$ *and* $Q^1_K \in \text{conv}(P_{i_1}, \ldots, P_{i_K})^\circ$. *Define the sequence*

$$Q^i_K \longleftarrow \textit{Residue}(Q^{i-1}_K \,|\, Q_{i-1});$$

then, $Q^K_K = P_{i_K}$.

Algorithm 31 Demix2$(S_1, \ldots, S_K)$

Input: $S_1, \ldots, S_K$ are distributions

if $K = 2$ **then**
  return $(\text{Residue}(S_1 \,|\, S_2), \text{Residue}(S_2 \,|\, S_1))^T$
else
  $(R_2, \ldots, R_K)^T \longleftarrow \text{FindFace}(S_1, \ldots, S_K)$
  $(Q_1, \ldots, Q_{K-1})^T \longleftarrow \text{Demix2}(R_2, \ldots, R_K)$
  for $i = 1, \ldots, K-1$ **do**
    $Q_K \longleftarrow \text{Residue}(Q_K \,|\, Q_i)$
  end for
  return $(Q_1, \ldots, Q_K)^T$
end if

Proof. Relabel the distributions so that $Q_j = P_j$. Let $\boldsymbol{\mu}_i$ denote the mixture proportion of Q_K^i and $\boldsymbol{e}_j$ the mixture proportion of P_j. Write $\boldsymbol{\mu}_1 = \sum_{i=1}^K \alpha_i \boldsymbol{e}_i$. We claim that $\boldsymbol{\mu}_k = \frac{\sum_{i \geqslant k} \alpha_i \boldsymbol{e}_i}{\sum_{i \geqslant k} \alpha_i}$ for all $k \leqslant K$. We prove this inductively. The base case $k = 1$ follows since $\sum_{i \geqslant 1} \alpha_i = 1$. Next, we prove the inductive step. Suppose that $\boldsymbol{\mu}_{k-1} = \frac{\sum_{i \geqslant k-1} \alpha_i \boldsymbol{e}_i}{\sum_{i \geqslant k-1} \alpha_i}$. By Proposition 6, the mixture proportion of Q_K^k, $\boldsymbol{\mu}_k$, is the residue of $\boldsymbol{\mu}_{k-1}$ with respect to $\boldsymbol{e}_{k-1}$. By statement 1 of Lemma 30, we can write

$$\begin{aligned} \boldsymbol{\mu}_k &= \boldsymbol{e}_{k-1} + \alpha^*(\boldsymbol{\mu}_{k-1} - \boldsymbol{e}_{k-1}) \\ &= \frac{[\sum_{i \geqslant k-1} \alpha_i (1 - \alpha^*) + \alpha^* \alpha_{k-1}] \boldsymbol{e}_{k-1} + \alpha^* \sum_{i \geqslant k} \alpha_i \boldsymbol{e}_i}{\sum_{i \geqslant k-1} \alpha_i} \end{aligned}$$

where

$$\alpha^* = \frac{1}{1 - \kappa^*(\boldsymbol{\mu}_{k-1} \,|\, \boldsymbol{e}_{k-1})}$$

and we have used the inductive hypothesis $\boldsymbol{\mu}_{k-1} = \frac{\sum_{i \geqslant k-1} \alpha_i \boldsymbol{e}_i}{\sum_{i \geqslant k-1} \alpha_i}$. α^* is the value of the following optimization problem (in statement 1 of Lemma 30):

$$\max(\alpha \geqslant 1 \,|\, \exists G, G = \boldsymbol{\mu}_{k-1} + \alpha(\boldsymbol{e}_{k-1} - \boldsymbol{\mu}_{k-1})).$$

Inspection of the above optimization problem reveals that $\alpha^* = \frac{\sum_{i \geqslant k-1} \alpha_i}{\sum_{i \geqslant k} \alpha_i}$. Plugging this into the above equation gives $\boldsymbol{\mu}_k = \frac{\sum_{i \geqslant k} \alpha_i \boldsymbol{e}_i}{\sum_{i \geqslant k} \alpha_i}$. This establishes the claim.

Setting $k = K$, it follows that $\mu_K = \frac{\alpha_K \boldsymbol{e}_K}{\alpha_K} = \boldsymbol{e}_K$. □

Corollary 3. *Let $P_1, \ldots, P_L$ be jointly irreducible and $\boldsymbol{\Pi}$ have full column rank. Then, with probability* 1*, Demix2($\tilde{\boldsymbol{P}}$) returns a permutation of $\boldsymbol{P}$.*

Proof of Theorem 29. Note that every estimator of a distribution in the DemixHat algorithm is a ResidueHat estimator since *(i)* the Demix2 algorithm 31 only considers distributions that are in $\text{conv}(P_1, \ldots, P_L)$ and *(ii)* only computes $\text{Residue}(F \,|\, H)$ if $F \neq H$. To see why *(ii)* is true, consider: the Demix2 algorithm computes $\text{Residue}(\cdot \,|\, \cdot)$ at lines 2 and 7 in Demix2 and line 3 in FindFace. In the proof of Theorem 26, we showed that $S_1, \ldots, S_K$ are always linearly independent and therefore distinct. This implies that in lines 2 and 7 in Demix2 and line 3 in FindFace, the residue function is called on distinct distributions. Thus, every estimator of a distribution of the DemixHat algorithm satisfies the assumptions of Lemma 39.

First, we argue that the order of the ResidueHat estimators is bounded; this implies that the constants in the uniform deviation inequalities associated with the ResidueHat estimators are bounded. We give a very loose bound. DemixHat calls itself at most $L-1$ times and in each call recurses on at most $L-1$ ResidueHat estimators and calculates at most $L-1$ more ResidueHat estimators. Therefore, each ResidueHat estimator has order at most $(L-1)^3$.

Second, let A_i denote the event that DemixHat recurses on i distributions lying in the relative interior of an i-face in the $(L-i)$th recursive call. We show that the event $\cap_{i=2}^{L-1} A_i$ occurs with probability tending to 1 as $\boldsymbol{n} \longrightarrow \infty$. Consider A_{L-1}. Let $\widehat{R}_i^{(n)}$ denote the estimate of the ith distribution in line 3 in the nth iteration of the for loop in Algorithm 27 and let $R_i^{(n)}$ denote the corresponding distribution. Let $\kappa^*_{i,j,n} = \kappa^*(R_i^{(n)} \,|\, R_j^{(n)})$. From the proof of Theorem 26, there exists an integer $N_1 \geqslant 0$ such that for $n \geqslant N_1$, $R_i^{(n)}$ lies in the relative interior of the same face for all $i = 2, \ldots, L$. Further, using the notation from the proof of Theorem 26, we have that the mixture proportions of the $R_i^{(n)}$s, i.e., the $\boldsymbol{\mu}_i^{(n)}$s, converge to a common $\boldsymbol{\lambda}$ on this face, i.e., for all $i = 2, \ldots, L$, $\left\| \boldsymbol{\mu}_i^{(n)} - \boldsymbol{\lambda} \right\| \longrightarrow 0$. Thus, by statement 3 of Lemma 32, for all $i \neq j \in [L] \backslash \{1\}$ $\kappa^*(\boldsymbol{\mu}_i^{(n)} \,|\, \boldsymbol{\mu}_j^{(n)}) \longrightarrow 1$. Hence, there exists $N_2 \geqslant N_1$ such that $\kappa^*(\boldsymbol{\mu}_i^{(N_2)} \,|\, \boldsymbol{\mu}_j^{(N_2)}) > \epsilon$ for all $i \neq j$. By statement 1 of Lemma 40 and a union bound argument, with probability increasing to 1, $\text{FaceTestHat}(\widehat{R}_2^{(n)}, \ldots, \widehat{R}_L^{(n)} \,|\, \epsilon)$

returns 0 for all $n < N_1$ since $R_2^{(n)}, \ldots, R_L^{(n)}$ are not on the relative interior of the same face. Thus, with probability tending to 1, FaceTest does not make the mistake to return 1 before the distributions $R_2^{(n)}, \ldots, R_L^{(n)}$ are on the relative interior of the same face. By statement 2 of Lemma 40, with probability tending to 1 as $\boldsymbol{n} \longrightarrow \infty$, FaceTestHat$(\widehat{R}_2^{(N_2)}, \ldots, \widehat{R}_L^{(N_2)} \,|\, \epsilon)$ returns 1. Hence, with probability increasing to 1, the event A_{L-1} occurs. Applying the same argument to A_i for $i < L-1$ and taking the union bound shows that $\cap_{i=2}^{L-1} A_i$ occurs with probability tending to 1 as $\boldsymbol{n} \longrightarrow \infty$.

Now, we can complete the proof. Under the assumptions of Theorem 26, there is a permutation σ such that for each distribution Q_i estimated by $\widehat{Q}_i$, $P_{\sigma(i)} = Q_i$. By Proposition 7, as $\boldsymbol{n} \longrightarrow \infty$, $\widehat{Q}_i$ converges uniformly to Q_i. The result follows. □

5.11.3 Classification with Partial Labels

In this section, we prove Theorem 30. To begin, we briefly sketch an argument that one can reduce any instance of a partial label model satisfying **(B3)** and **(A)** to an instance of a partial label model that also satisfies **(D)**. Let $J = \{i : \mathbf{\Pi}^+_{i,:} = \boldsymbol{e}_j^T \text{ for some } j \in [L]\} = \{j_1, \ldots, j_k\}$, the set of indices of contaminated distributions that are equal to some base distribution. Compute Residue$(\tilde{P}_i \,|\, \tilde{P}_{j_1})$ for $i \in [L] \backslash J$ if there is l such that $\mathbf{\Pi}^+_{i,l} = \mathbf{\Pi}^+_{j_1,l} = 1$. Replace $\tilde{P}_i$ with Residue$(\tilde{P}_i \,|\, \tilde{P}_{j_1})$ (and call it $\tilde{P}_i$ for simplicity of presentation). Update $\mathbf{\Pi}^+$ and remove j_1 from J. Repeat this procedure until J is empty. Then, there will be $(L - |J|)$ $\tilde{P}_i$ lying in a $(L - |J|)$-face of Δ_L that are not equal to any of the base distributions and the other contaminated distributions will be equal to base distributions. Then, it suffices to solve the instance of the partial label model on the $(L - |J|)$-face, which satisfies **(D)**.

Next, we introduce VertexTestHat (Algorithm 32), an empirical version of VertexTest, and prove that it satisfies a useful consistency property.

Lemma 42. *Suppose that $P_1, \ldots, P_L$ satisfy **(A″)**, $\mathbf{\Pi}$ has full column rank, the columns of $\mathbf{\Pi}^+$ are unique and $\mathbf{\Pi}^+$ satisfies **(D)**. Let $\widehat{Q}_1, \ldots, \widehat{Q}_L$ be ResidueHat estimators of $Q_1, \ldots, Q_L$, respectively. Suppose that $(Q_1, \ldots, Q_L)$ is a permutation of $(P_1, \ldots, P_L)$. Then, with probability tending to 1 as $\boldsymbol{n} \longrightarrow \infty$, VertexTestHat$(\mathbf{\Pi}^+, (\tilde{P}_1^\dagger, \ldots, \tilde{P}_M^\dagger)^T, (\widehat{Q}_1, \ldots, \widehat{Q}_L)^T)$ returns a permutation matrix $\boldsymbol{C}$ such that $\forall i$, $\boldsymbol{C}_{i,:}(\widehat{Q}_1, \ldots, \widehat{Q}_L)^T$ is a ResidueHat estimator*

Algorithm 32 VertexTestHat($\mathbf{\Pi}^+, (\tilde{P}_1^\dagger, \ldots, \tilde{P}_M^\dagger)^T, (\widehat{Q}_1, \ldots, \widehat{Q}_L)^T$)

1: Form the matrix $\widehat{\boldsymbol{M}}_{i,j} \coloneqq \widehat{\kappa}(\tilde{P}_i^\dagger \,|\, \widehat{Q}_j)$
2: Let $|\mathbf{\Pi}^+|$ denote the number of nonzero entries in $\mathbf{\Pi}^+$
3: Form the matrix $\widehat{\boldsymbol{Z}}$ by making the $|\mathbf{\Pi}^+|$ largest entries of $\widehat{\boldsymbol{M}}$ equal to 1 and the rest of its entries equal to 0
4: Use any algorithm that finds a permutation matrix $\boldsymbol{C}$ such that $\widehat{\boldsymbol{Z}}\boldsymbol{C} = \mathbf{\Pi}^+$ (if it exists)
5: **if** such a permutation matrix $\boldsymbol{C}$ exists **then**
6: **return** $(1, \boldsymbol{C}^T)$
7: **else**
8: **return** $(0, \mathbf{0})$
9: **end if**

of P_i.

Proof. Define $\widehat{\kappa}_{i,j} \coloneqq \widehat{\kappa}(\tilde{P}_i^\dagger \,|\, \widehat{Q}_j)$ and $\kappa^*_{i,j} \coloneqq \kappa^*(\tilde{P}_i \,|Q_j)$. We claim that $\widehat{\kappa}_{i,j}$ satisfies a **(RC)**. Since $P_1, \ldots, P_L$ satisfy **(A″)** and by assumption **(D)** $Q_j \neq \tilde{P}_i$, by Lemma 36, $\tilde{P}_i$ satisfies **(SC)** wrt Q_j. Since $\tilde{P}_i^\dagger$ is an empirical distribution, $\tilde{P}_i^\dagger$ satisfies a **(UDI)**. Since $\widehat{Q}_j$ is a ResidueHat estimator, $\widehat{Q}_j$ satisfies a **(UDI)** by Lemma 39. Therefore, the hypotheses of Lemma 38 are satisfied and $\widehat{\kappa}_{i,j}$ satisfies a **(RC)**.

Form the matrix $Z_{i,j} = \mathbf{1}\{\kappa^*(\tilde{P}_i \,|\, Q_j) > 0\}$ as in Algorithm 24. Since $Q_1, \ldots, Q_L$ are a permutation of $P_1, \ldots, P_L$, $\boldsymbol{Z}$ is formed by permuting the columns of $\mathbf{\Pi}^+$ appropriately. Thus, there are $|\mathbf{\Pi}^+|$ (i, j) pairs such that $\kappa_{i,j} > 0$ and the rest are such that $\kappa_{i,j} = 0$. Then, using Lemma 38 and a union bound, with probability tending to 1 as $\boldsymbol{n} \longrightarrow \infty$, $\widehat{\kappa}_{i,j}$ is among the $|S|$ largest values in the matrix $\widehat{M}$ if and only if $\kappa_{i,j} > 0$. On this event, VertexTestHat($\mathbf{\Pi}^+, (\tilde{P}_1^\dagger, \ldots, \tilde{P}_M^\dagger)^T, (\widehat{Q}_1, \ldots, \widehat{Q}_L)^T$) and VertexTest($\mathbf{\Pi}^+, (\tilde{P}_1, \ldots, \tilde{P}_M)^T, (Q_1, \ldots, Q_L)^T$) return the same output. Since $(Q_1, \ldots, Q_L)$ is a permutation of $(P_1, \ldots, P_L)$ by hypothesis, by Lemma 35 VertexTest($\mathbf{\Pi}^+, (\tilde{P}_1, \ldots, \tilde{P}_M)^T, (Q_1, \ldots, Q_L)^T$) returns $(1, \boldsymbol{C}^T)$ such that $\boldsymbol{C}^T(Q_1, \ldots, Q_L)^T = \boldsymbol{P}$. The result follows. □

Proof of Theorem 30. Let $(\widehat{Q}_1, \ldots, \widehat{Q}_L) \longleftarrow$ DemixHat($\tilde{P}_1^\dagger, \ldots, \tilde{P}_M^\dagger \,|\, \epsilon$). By Theorem 29, w.p. tending towards 1 as $\boldsymbol{n} \longrightarrow \infty$, there exists a permutation $\sigma : [L] \longrightarrow [L]$ such that for every

$i \in [L]$,

$$\sup_{E \in \mathcal{E}} |\widehat{Q}_i(E) - P_{\sigma(i)}(E)| < \delta.$$

From the proof of Theorem 29, each $\widehat{Q}_i$ is a ResidueHat estimator. The assumptions of Lemma 42 are satisfied. The result follows immediately from Lemma 42. □

5.12 Previous Results

Lemma 43 (Lemma A.1 (Blanchard and Scott, 2014))**.** *The maximum operation in the definition of κ^* and $\widehat{\kappa}$ (lines (5.3) and (5.6), respectively) is well-defined, that is, the outside supremum is attained at at least one point.*

Lemma 44 (Lemma B.1 (Blanchard and Scott, 2014))**.** *If Π satisfies **(B1)**, then $\boldsymbol{\pi}_1, \ldots, \boldsymbol{\pi}_L$ are linearly independent. If $P_1, \ldots, P_L$ are jointly irreducible, then they are linearly independent. If $\boldsymbol{\pi}_1, \ldots, \boldsymbol{\pi}_L$ are linearly independent and $P_1, \ldots, P_L$ are linearly independent, then $\tilde{P}_1, \ldots, \tilde{P}_L$ are linearly independent.*

5.13 Experiments

In this Section, we perform experiments that suggest that joint irreducibility of $P_1, \ldots, P_L$ is a reasonable assumption. In particular, our experiments suggest that on the datasets in question, **(A″)** holds (which is a strictly stronger condition than joint irreducibility). We consider three datasets: classes 1, 2, and 3 of MNIST (LeCun et al., 1998), the Iris dataset (Fisher, 1936), and the Breast Cancer Wisconsin (Diagnostic) Data Set (Dheeru and K. Taniskidou, 2017). We use the Spectral Support Estimation algorithm (De Vito et al., 2010; Rudi et al., 2014) to estimate the support of each class in each dataset. We split each dataset into training, validation, and test sets, applying the algorithm to the training set, using the validation set to pick the hyperparameters, and evaluating the performance on the test set. We average our results over 60 trials where in each trial we randomly permute the dataset, thus altering the training, validation, and test sets. Let $\widehat{S}_i$ denote an estimate of the support of class i. Tables 5.1, 5.3, and 5.5 display an estimate of the probability that

a point sampled from P_i belongs to the estimate of the support $\widehat{S}_i$. They indicate that the Spectral Support Estimation has reasonably good performance in producing $\widehat{S}_i$s containing the support of the associated class. Tables 5.2, 5.4, and 5.6 use the $\widehat{S}_i$ to estimate the quantity $\Pr_{\boldsymbol{x}\sim P_i}(\boldsymbol{x} \in \cup_{j\neq i} \operatorname{supp}(P_j))$, which must be strictly less than 1 for $\mathbf{(A'')}$ to hold. We find that our estimates are considerably less than 1, which suggests that joint irreducibility holds on these datasets.

	$i=1$	$i=2$
$\widehat{\Pr}_{\boldsymbol{x}\sim P_i}(\boldsymbol{x}\in\widehat{S}_i)$	0.87	0.89

Table 5.1: Cancer Support Results.

	$i=1$	$i=2$
$\widehat{\Pr}_{\boldsymbol{x}\sim P_i}(\boldsymbol{x}\in\cup_{j\neq i}\widehat{S}_j)$	0.18	0.38

Table 5.2: Cancer Separability Results.

	$i=1$	$i=2$	$i=3$
$\widehat{\Pr}_{\boldsymbol{x}\sim P_i}(\boldsymbol{x}\in\widehat{S}_i)$	0.86	0.84	0.84

Table 5.3: Iris Support Results.

	$i=1$	$i=2$	$i=3$
$\widehat{\Pr}_{\boldsymbol{x}\sim P_i}(\boldsymbol{x}\in\cup_{j\neq i}\widehat{S}_j)$	0.0	0.17	0.19

Table 5.4: Iris Separability Results.

	$i=1$	$i=2$	$i=3$
$\widehat{\Pr}_{\boldsymbol{x}\sim P_i}(\boldsymbol{x}\in\widehat{S}_i)$	0.98	0.87	0.83

Table 5.5: MNIST Support Results.

	$i=1$	$i=2$	$i=3$
$\widehat{\Pr}_{\boldsymbol{x}\sim P_i}(\boldsymbol{x}\in\cup_{j\neq i}\widehat{S}_j)$	0.08	0.17	0.14

Table 5.6: MNIST Separability Results.

Chapter 6

Nonparametric Preference Completion

In this Chapter, I consider the task of collaborative preference completion: given a pool of items, a pool of users and a partially observed item-user rating matrix, the goal is to recover the *personalized ranking* of each user over all of the items. Our approach is nonparametric: I assume that each item i and each user u have unobserved features x_i and y_u, and that the associated rating is given by $g_u(f(x_i, y_u))$ where f is Lipschitz and g_u is a monotonic transformation that depends on the user. I propose a k-nearest neighbors-like algorithm and prove that it is consistent. To the best of our knowledge, this is the first consistency result for the collaborative preference completion problem in a nonparametric setting. Finally, I demonstrate the performance of my algorithm with experiments on the Netflix and Movielens datasets. This Chapter is joint work with Clayton Scott and was presented at the International Conference on Artificial Intelligence and Statistics in 2018.

6.1 Introduction

In the preference completion problem, there is a pool of items and a pool of users. Each user rates a subset of the items and the goal is to recover the personalized ranking of each user over all of the items. This problem is fundamental to recommender systems, arising in tasks such as movie recommendation and news personalization. A common approach is to first

estimate the ratings through either a matrix factorization method or a neighborhood-based method and to output personalized rankings from the estimated ratings (Koren et al., 2009; Zhou et al., 2008; Ning et al., 2011; Breese et al., 1998). Recent research has observed a number of shortcomings of this approach (Weimer et al., 2007; Liu and Yang, 2008); for example, many ratings-oriented algorithms minimize the RMSE, which does not necessarily produce a good ranking (Cremonesi et al., 2010). This observation has sparked a number of proposals of algorithms that aim to directly recover the rankings (Weimer et al., 2007; Liu and Yang, 2008; Lu and Negahban, 2014; Park et al., 2015; Oh et al., 2015; Gunasekar et al., 2016). Although these ranking-oriented algorithms have strong empirical performance, there are few theoretical guarantees to date and they all make specific distributional assumptions (discussed in more detail below). In addition, these results have focused on low-rank methods, while ranking-oriented neighborhood-based methods have received little theoretical attention.

In this chapter, we consider a statistical framework for nonparametric preference completion. We assume that each item i and each user u have unobserved features x_i and y_u, respectively, and that the associated rating is given by $g_u(f(x_i, y_u))$ where f is Lipschitz and g_u is a monotonic transformation that depends on the user. We make the following contributions. (i) We propose a simple k-nearest neighbors-like algorithm, (ii) we provide, to the best of our knowledge, the first consistency result for ranking-oriented algorithms in a nonparametric setting, and (iii) we provide a necessary and sufficient condition for the optimality of a solution (defined below) to the preference completion problem.

6.2 Related Work

The two main approaches to preference completion are matrix factorization methods (e.g., low-rank approximation) and neighborhood-based methods. Recently, there has been a surge of research with many theoretical advances in low-rank approximation for collaborative filtering, e.g., (Recht, 2011; Keshavan et al., 2010). These methods tend to focus on minimizing the RMSE even though applications usually use ranking measures to evaluate performance. While recent work has developed ranking-oriented algorithms that outperform

ratings-oriented algorithms (Gunasekar et al., 2016; Liu and Yang, 2008; Rendle et al., 2009; Pessiot et al., 2007; Cremonesi et al., 2010; Weimer et al., 2007), many of these proposals lack basic theoretical guarantees such as consistency. A recent line of work has begun to fill this gap by establishing theoretical results under specific generative models. Lu and Negahban (2014) and Park et al. (2015) provided consistency guarantees using a low-rank approach and the Bradley-Terry-Luce model. Similarly, Oh et al. (2015) established a consistency guarantee using a low rank approach and the MultiNomial Logit model. By contrast, our approach forgoes such strong parametric assumptions.

Neighborhood-based algorithms are popular methods, e.g. (Das et al., 2007), because they are straightforward to implement, do not require expensive model-training, and generate interpretable recommendations (Ning et al., 2011). There is an extensive experimental literature on neighborhood-based collaborative filtering methods. The most common approach is the user-based model; it is based on the intuition that if two users give similar ratings to items in the observed data, then their unobserved ratings are likely to be similar. This approach employs variants of k nearest-neighbors. Popular similarity measures include the Pearson Correlation coefficient and cosine similarity. There are a large number of schemes for predicting the unobserved ratings using the k nearest neighbors, including taking a weighted average of the ratings of the users and majority vote of the users (Ning et al., 2011).

Recently, researchers have sought to develop neighborhood-based collaborative filtering algorithms that aim to learn a personalized ranking for each user instead of each user's ratings (Liu and Yang, 2008; Wang et al., 2014, 2016). Eigenrank, proposed by Liu and Yang (2008), is structurally similar to our algorithm. It measures the similarity between users with the Kendall rank correlation coefficient, a measure of the similarity of two rankings. Then, it computes a utility function $\psi : [n_1] \times [n_1] \longrightarrow \mathbb{R}$ for each user that estimates his pairwise preferences over the items. From the estimated pairwise preferences, it constructs a personalized ranking for each user by either using a greedy algorithm or random walk model. In contrast, our algorithm uses the average number of agreements on pairs of items to measure similarity between users and a majority vote approach to predict pairwise preferences.

Neighborhood-based collaborative filtering has not received much theoretical attention.

Kleinberg and Sandler (2003, 2004) model neighborhood-based collaborative filtering as a latent mixture model and prove consistency results in this specific generative setting. Recently, Lee et al. (2016), who inspired the framework in the current chapter, studied rating-oriented neighborhood-based collaborative filtering in a more general nonparametric setting. Their approach assumes that each item i and each user u have unobserved features x_i and y_u, respectively, and that the associated rating is given by $f(x_i, y_u)$ where f is Lipschitz, whereas we assume that the associated rating is given by $g_u(f(x_i, y_u))$ where g_u is a user-specific monotonic transformation. As we demonstrate in our experiments, their algorithm is not robust to monotonic transformations of the columns, but this robustness is critical for many applications. For example, consider the following implicit feedback problem (Hu et al., 2008). A recommender system for news articles measures how long users read articles as a proxy for item-user ratings. Because reading speeds and attention spans vary dramatically, two users may actually have very similar preferences despite substantial differences in reading times.

Even though our method is robust to user-specific monotonic transformations, we do not require observing many more entries of the item-user matrix than Lee et al. (2016) in the regime where there are many more users than items (e.g., the Netflix dataset). If there are n_1 items and n_2 users, Lee et al. (2016) requires that there exists $\frac{1}{2} > \alpha > 0$ such that the probability of observing an entry is greater than $\max(n_1^{-\frac{1}{2}+\alpha}, n_2^{-1+\alpha})$, whereas we require that this probability is greater than $\max(n_1^{-\frac{1}{2}+\alpha}, n_2^{-\frac{1}{2}+\alpha})$.

Our work is also related to the problem of Monotonic Matrix Completion (MMC) where a single monotonic Lipschitz function is applied to a low rank matrix and the goal is rating estimation (Ganti et al., 2015). In contrast, we allow for distinct monotonic, possibly non-Lipschitz functions for every user and pursue the weaker goal of preference completion.

To the best of our knowledge, there is no theoretically supported, nonparametric method for preference completion. Our work seeks to address this issue.

6.3 Setup

Notation: Define $[n] = \{1, \ldots, n\}$. Let $\Omega \subset [n_1] \times [n_2]$. If $X \in \mathbb{R}^{n_1 \times n_2}$, let $\mathcal{P}_\Omega(X) \in (\mathbb{R} \cup \{?\})^{n_1 \times n_2}$ be defined as $[\mathcal{P}_\Omega(X)]_{i,j} = \begin{cases} X_{i,j} & \text{if } (i,j) \in \Omega \\ ? & \text{if } (i,j) \notin \Omega \end{cases}$. If f is some function and U a finite collection of objects belonging to the domain of f, let $\max_{u \in U}^{(l)} f(u)$ denote the lth largest value of f over U. Let $\text{Bern}(p)$ denote a realization of a Bernoulli random variable with parameter p. For a metric space $\mathcal{M}$ with metric $d_\mathcal{M}$, let $B_\epsilon(z) = \{z' \in \mathcal{M} : d_\mathcal{M}(z, z') < \epsilon\}$. We use bold type to indicate random variables. For example, $\boldsymbol{z}$ denotes a random variable and z a realization of $\boldsymbol{z}$.

Nonparametric Model: Suppose that there are n_1 items and n_2 users. Furthermore,

1. The items are associated with unobserved features $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n_1} \in \mathcal{X}$, and the users are associated with unobserved features $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{n_2} \in \mathcal{Y}$ where $\mathcal{X}$ and $\mathcal{Y}$ are compact metric spaces with metrics $d_\mathcal{X}$ and $d_\mathcal{Y}$, respectively.

2. $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n_1}, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{n_2}$ are independent random variables such that $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n_1} \overset{i.i.d.}{\sim} \mathcal{P}_\mathcal{X}$ and $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{n_2} \overset{i.i.d.}{\sim} \mathcal{P}_\mathcal{Y}$ where $\mathcal{P}_\mathcal{X}$ and $\mathcal{P}_\mathcal{Y}$ denote Borel probability measures over $\mathcal{X}$ and $\mathcal{Y}$, respectively. We assume that for all $\epsilon > 0$ and $y \in \mathcal{Y}$, $\mathcal{P}_\mathcal{Y}(B_\epsilon(y)) > 0$.

3. The complete ratings matrix is $H \coloneqq [h_u(x_i, y_u)]_{i \in [n_1], u \in [n_2]}$ where $h_u = g_u \circ f$, $f : \mathcal{X} \times \mathcal{Y} \longrightarrow \mathbb{R}$ is a Lipschitz function with respect to the induced metric $d_{\mathcal{X} \times \mathcal{Y}}((x_1, y_1), (x_2, y_2)) \coloneqq \max(d_\mathcal{X}(x_1, x_2), d_\mathcal{Y}(y_1, y_2))$ with Lipschitz constant 1,[1] i.e., $\forall y_1, y_2 \in \mathcal{Y}$ and $\forall x_1, x_2 \in \mathcal{X}$, $|f(x_1, y_1) - f(x_2, y_2)| \leqslant \max(d_\mathcal{X}(x_1, x_2), d_\mathcal{Y}(y_1, y_2))$, and g_u is a nondecreasing function. Note that each h_u need not be Lipschitz.

4. Each entry of the matrix H is observed independently with probability p. Let $\boldsymbol{\Omega} \subset [n_1] \times [n_2]$ be a random variable denoting the indices of the observed ratings.

Whereas Lee et al. (2016) considers the task of completing a partially observed matrix $F \coloneqq [f(x_i, y_u)]_{i \in [n_1], u \in [n_2]}$ when $\{x_i\}_{i \in [n_1]}$ and $\{y_u\}_{u \in [n_2]}$ are unobserved, we aim to recover the ordering of the elements in each column of H when $\{x_i\}_{i \in [n_1]}$ and $\{y_u\}_{u \in [n_2]}$ are unobserved.

[1] We could develop our framework with an arbitrary Lipschitz constant L, but for ease of presentation, we fix $L = 1$.

In our setup, we view F as an ideal preference matrix representing how much users like items and H as how those preferences are expressed based on user-specific traits (see the news recommender system example in Section 6.2).

This framework subsumes various parametric models. For example, consider a matrix factorization model that assumes that there is a matrix $H \in \mathbb{R}^{n_1 \times n_2}$ of rank $d \leqslant \min(n_1, n_2)$ such that user u prefers item i to item j if and only if $H_{i,u} > H_{j,u}$. Then, we can factorize H such that $H_{i,u} = x_i^t y_u$ where $x_i, y_u \in \mathbb{R}^d$ for all $i \in [n_1]$ and $u \in [n_2]$. In our setup, we have $f(x_i, y_u) = x_i^t y_u$ and $g_u(z) = z$.

Task: Let $\mathcal{S}^{n_1} = \{\sigma : \sigma : [n_1] \longrightarrow [n_1]$, σ is a permutation$\}$ denote the set of permutations on n_1 objects. We call $\sigma \in \mathcal{S}^{n_1}$ a *ranking*. Let $\mathcal{S}^{n_1 \times n_2} = (\mathcal{S}^{n_1})^{n_2}$. That is, $\sigma \in \mathcal{S}^{n_1 \times n_2}$ if $\sigma : [n_1] \times [n_2] \longrightarrow [n_1]$ and for fixed $u \in [n_2]$, $\sigma(\cdot, u)$ is a permutation on $[n_1]$. We call $\sigma \in \mathcal{S}^{n_1 \times n_2}$ a *collection of rankings*. Let $\epsilon > 0$. Our goal is to learn $\sigma \in \mathcal{S}^{n_1 \times n_2}$ that minimizes the number of pairwise ranking disagreements per user with some slack, i.e.,

$$\mathrm{dis}_\epsilon(\sigma, H) = \sum_{u=1}^{n_2} \sum_{i<j} \mathbf{1}\{|f(x_i, y_u) - f(x_j, y_u)| > \epsilon\} \mathbf{1}\{(h_u(x_i, y_u) - h_u(x_j, y_u))(\sigma(i, u) - \sigma(j, u)) < 0\}.$$

6.4 Algorithm

Our algorithm, Multi-Rank (Algorithm 33), has two stages: first it estimates the pairwise preferences of each user and, second, it constructs a full ranking for each user from its estimated pairwise preferences. In the first stage, Multi-Rank computes $A \in \{0, 1\}^{n_2 \times n_1 \times n_1}$ where $A_{u,i,j} = 1$ denotes that user u prefers item i to item j and $A_{u,i,j} = 0$ denotes that user u prefers item j to item i. If a user has provided distinct ratings for a pair of items, Multi-Rank fills in the corresponding entries of A. Otherwise, Multi-Rank uses a subroutine called Pairwise-Rank that we will describe shortly. Once Multi-Rank has constructed A, it applies the Copeland ranking procedure to the pairwise preferences of each user (discussed at the end of the section).

The Pairwise-Rank algorithm predicts whether a user u prefers item i to item j or vice versa. It is similar to k-nearest neighbors where we use the forthcoming ranking measure as

Algorithm 33 Multi-Rank

Input: $\mathcal{P}_\Omega(H), \beta \geqslant 2, k > 0$
for $u \in [n_2]$, $i, j \in [n_1], i < j$ **do**
  if $(i,u) \in \Omega$, $(j,u) \in \Omega$ and $H_{i,u} \neq H_{j,u}$ **then**
    Set $A_{u,i,j} = \mathbf{1}\{H_{i,u} > H_{j,u}\}$
    Set $A_{u,j,i} = 1 - A_{u,i,j}$
  else
    Set $A_{u,i,j} = \text{Pairwise-Rank}(u, i, j, \beta, k)$
    Set $A_{u,j,i} = 1 - A_{u,i,j}$
  end if
end for
for $u \in [n_2]$ **do**
  $\widehat{\sigma}_u = \text{Copeland}(A_{u,:,:})$
end for
return $\widehat{\sigma} \coloneqq (\widehat{\sigma}_1, \ldots, \widehat{\sigma}_{n_2})$

our distance measure. Let $N(u)$ denote the set of items that user u has rated, i.e.,

$$N(u) = \{l : (l,u) \in \Omega\},$$

and $N(u,v) = N(u) \cap N(v)$ denote the set of items that users u and v have both rated. Viewing $N(u,v)$ as an ordered array where $N(u,v)[\ell]$ denotes the $(\ell+1)$th element, let

$$I(u,v) = \{(s,t) : s = N(u,v)[\ell], t = N(u,v)[\ell+1] \text{ for some } \ell \in \{2k : k \in \mathbb{N} \cup \{0\}\}\}.$$

In words, $I(u,v)$ is formed by sorting the indices of $N(u,v)$ and selecting nonoverlapping pairs in the given order. Note that there is no overlap between the indices in the pairs in $I(u,v)$.[2] Fix $y_u, y_v \in \mathcal{Y}$. If $I(u,v) = \varnothing$, define $R_{u,v} = 0$ and if $I(u,v) \neq \varnothing$, let $R_{u,v} \coloneqq$

$$\frac{1}{|I(u,v)|} \sum_{(s,t) \in I(u,v)} \mathbf{1}\{(h_u(\boldsymbol{x}_s, y_u) - h_u(\boldsymbol{x}_t, y_u))(h_v(\boldsymbol{x}_s, y_v) - h_v(\boldsymbol{x}_t, y_v)) \geqslant 0\}$$

denote the fraction of times that users u and v agree on the relative ordering of item pairs belonging to $I(u,v)$. In practice, one can simply compute this statistic over all pairs of

[2] We select nonoverlapping pairs to preserve independence in the estimates for the forthcoming analysis.

commonly rated items. Observe that $\rho(y_u, y_v) \coloneqq$

$$\mathbb{E}[R_{u,v}|I(u,v) \neq \varnothing, \boldsymbol{y}_u = y_u, \boldsymbol{y}_v = y_v] = \Pr_{\boldsymbol{x}_s, \boldsymbol{x}_t \sim P_{\mathcal{X}}}([h_u(\boldsymbol{x}_s, y_u) - h_u(\boldsymbol{x}_t, y_u)] \times [h_v(\boldsymbol{x}_s, y_v) - h_v(\boldsymbol{x}_t, y_v)] \geqslant 0)$$

i.e., $\rho(y_u, y_v)$ is the probability that users u and v with features y_u and y_v order two random items in the same way.

We apply Pairwise-Rank (Algorithm 34) to a user u and a pair of items (i, j) if the user has not provided distinct ratings for items i and j. Pairwise-Rank(u, i, j, β, k) finds users that have rated items i and j, and have rated at least β items in common with u. If there are no such users, Pairwise-Rank flips a coin to predict the relative preference ordering. If there are such users, then it sorts the users in decreasing order of $R_{u,v}$ and takes a majority vote over the first k users about whether item i or item j is preferred. If the vote results in a tie, Pairwise-Rank flips a coin to predict the relative preference ordering.

Algorithm 34 Pairwise-Rank

Input: $u \in [n_2], i \in [n_1], j \in [n_1], \beta \geqslant 2, k \in \mathbb{N}$
$W_u^{i,j}(\beta) = \{v \in [n_2] : |N(u,v)| \geqslant \beta, (i,v), (j,v) \in \Omega\}$
Sort $W_u^{i,j}(\beta)$ in decreasing order of $R_{u,v}$ and let V be the first k elements.
if $V = \varnothing$ **then**
    return Bern$(\frac{1}{2})$
end if
$\forall v \in V$, set $P_v = \mathbf{1}\{h_v(x_i, y_v) > h_v(x_j, y_v)\} - \mathbf{1}\{h_v(x_i, y_v) < h_v(x_j, y_v)\}$
if $\sum_{v \in V} P_v > 0$ **then**
    return 1
else if $\sum_{v \in V} P_v < 0$ **then**
    return 0
else
    return Bern$(\frac{1}{2})$
end if

Next, Multi-Rank converts the pairwise preference predictions of each user into a full estimated ranking for each user. It applies the Copeland ranking procedure (Algorithm 35)–

an algorithm for the feedback arc set problem in tournaments (Copeland, 1951; Coppersmith et al., 2006) to each user-specific set of pairwise preferences. The Copeland ranking procedure simply orders the items by the number of times an item is preferred to another item. It is possible to use other approximation algorithms for the feedback arc set problem such as Fas-Pivot from Ailon et al. (2008).

Algorithm 35 Copeland

Input: $A \in \{0,1\}^{n_1 \times n_1}$
for $j \in [n_1]$ **do**
    $I_j = \sum_{i=1, i \neq j}^{n_1} A_{j,i}$
end for
return $\sigma \in \mathcal{S}^{n_1}$ that orders items in decreasing order of I_j

6.5 Analysis of Algorithm

The main idea behind our algorithm is to use pairwise agreements about items to infer whether two users are close to each other in the feature space. However, this is not possible in the absence of further distributional assumptions. The Lipschitz condition on f only requires that if users u and v are close to each other, then $\max_z |f(z, y_u) - f(z, y_v)|$ is small. Proposition 12 shows that there exist functions arbitrarily close to each other that disagree about the relative ordering of almost every pair of points.

Proposition 12. *Let $\mathcal{X} = [0,1]$ and $\mathcal{P}_{\mathcal{X}}$ be the Lebesgue measure over $\mathcal{X}$. For every $\epsilon > 0$, there exist functions $f, g : \mathcal{X} \longrightarrow \mathbb{R}$ such that $\max_{x \in [0,1]} |f(x) - g(x)| = \|f - g\|_\infty \leqslant \epsilon$ and for almost every pair of points $(x, x') \in [0,1]^2$, $f(x) > f(x')$ iff $g(x) < g(x')$.*

Thus, we make the following mild distributional assumption.

Definition 23. *Fix $y \in \mathcal{Y}$ and let $f_y(x) \coloneqq f(x, y)$. Let r be a positive nondecreasing function. We say y is r-*discerning *if $\forall \epsilon > 0$, $\Pr_{\boldsymbol{x}_1, \boldsymbol{x}_2 \sim \mathcal{P}_{\mathcal{X}}}(|f_y(\boldsymbol{x}_1) - f_y(\boldsymbol{x}_2)| \leqslant 2\epsilon) < r(\epsilon)$.*

This assumption says that the probability that $f_y(\boldsymbol{x}_1)$ and $f_y(\boldsymbol{x}_2)$ are within ϵ of each other decays at some rate given by r. In a sense, it means that users perceive some difference

between most randomly selected items with different features, although the difference might be masked by the transformation g_u.

We also assume that if two users are not close to each other in the latent space, then they must have some disagreements. Definition 24 requires that the nonparametric model is economical (i.e., not redundant) in the sense that different parts of the feature space correspond to different preferences.

Definition 24. *Fix $y \in \mathcal{Y}$. Let $\epsilon, \delta > 0$. We say that y is* (ϵ, δ)-discriminative *if $z \in B_\epsilon(y)^c$ implies that $\rho(y, z) < 1 - \delta$.*

These assumptions are satisfied under many parametric models. Proposition 13 provides two illustrative examples under a matrix factorization model. We briefly note that, as we show in the supplementary material, $f(x, y) = x^t y$ and $f(x, y) = \|x - y\|_2$ are equivalent models by adding a dimension.

Proposition 13. *Consider $(\mathbb{R}^d, \|\cdot\|_2)$. Let $f(x, y) = \|x - y\|_2$ and $g_u(\cdot)$ be strictly increasing $\forall u \in [n_2]$.*

1. *Let $\mathcal{X} = \mathcal{Y} = \{x \in \mathbb{R}^d : \|x\|_2 \leqslant 1\}$, $\mathcal{P}_\mathcal{X}$ be the uniform distribution and for all $y \in \mathcal{Y}$ define $r_y(\epsilon) = \sup_{z \in [0,2]} \mathcal{P}_\mathcal{X}(B_z(y) \backslash B_{z-4\epsilon}(y))$. Then, for all $y \in \mathcal{Y}$, y is r_y-discerning. Further, define for all $\epsilon > 0$, $\delta_\epsilon = \inf_{v \in \mathcal{Y}} 2\mathcal{P}_\mathcal{X}(B_{\frac{\epsilon}{2}}(v))^2$. Then, for all $y_u \in \mathcal{Y}$ and for all $\epsilon > 0$, y_u is $(\epsilon, \delta_\epsilon)$-discriminative.*

2. *Let $\mathcal{X} \subset \mathbb{R}^d$ be a finite collection of points, $\mathcal{P}_\mathcal{X}$ be uniform over $\mathcal{X}$, and for all $y \in \mathcal{Y}$ define $r_y(\epsilon) = \frac{|\{(x,x') \in \mathcal{X} \times \mathcal{X} : |\|y-x\| - \|y-x'\|| \leqslant 2\epsilon\}|}{|\mathcal{X}|^2}$. Then, for all $y \in \mathcal{Y}$, y is r_y-discerning. Next, suppose $\mathcal{Y}$ is a finite collection of points and every pair of distinct $y, y' \in \mathcal{Y}$ disagree about at least C pairs of items. Let $\delta = \frac{C}{|\mathcal{X}|^2}$. For all $y_u \in \mathcal{Y}$ and for all $\epsilon > 0$, y_u is (ϵ, δ)-discriminative.*

Our analysis uses two functions to express problem-specific constants. First, let $\tau : \mathbb{R}_{++} \longrightarrow (0, 1]$ be defined as $\tau(\epsilon) = \inf_{y_0 \in \mathcal{Y}} \Pr_{\boldsymbol{y} \sim \mathcal{P}_\mathcal{Y}}(d_\mathcal{Y}(y_0, \boldsymbol{y}) \leqslant \epsilon)$. Second, let $\kappa : \mathbb{R}_{++} \longrightarrow (0, 1]$ be such that $\kappa(\epsilon) = \inf_{y_0 \in \mathcal{Y}} \Pr_{\boldsymbol{y} \sim \mathcal{P}_\mathcal{Y}}(d_\mathcal{Y}(y_0, \boldsymbol{y}) > \epsilon)$. Our assumption that for all $\delta > 0$ and $y \in \mathcal{Y}$, $\mathcal{P}_\mathcal{Y}(B_\delta(y)) > 0$ ensures that $\tau(\cdot) > 0$ and $\kappa(\cdot) < 1$ (see Lemma 47). If $\mathcal{P}_\mathcal{Y}$ is

uniform over the unit cube in $(\mathbb{R}^d, \|\cdot\|_\infty)$, then $\tau(\epsilon) = \min(1, \epsilon)^d$ and if $\mathcal{Y}$ is a finite collection of points, then $\tau(\epsilon) = \min_{y \in \mathcal{Y}} \mathcal{P}_{\mathcal{Y}}(y)$ (Lee et al., 2016).

Our model captures the intrinsic difficulty of a problem instance as follows. $r(\cdot)$ and $\tau(\cdot)$ together control the probability of sampling nearby users with similar preferences. (ϵ, δ)-discriminative captures how often users u and v must agree in order to infer that y_u and y_v are close in the latent space and, thus, $\max_z |f(z, y_u) - f(z, y_v)| \leqslant \epsilon$.

6.5.1 Continuous Ratings Setting

Our analysis deals with the case of continuous ratings and the case of discrete ratings separately. In this section, we prove theorems dealing with the continuous case and in the next section we give analogous results with similar proofs for the discrete case. Theorem 32 establishes that with probability tending to 1 as $n_2 \longrightarrow \infty$, Multi-Rank outputs $\widehat{\sigma} \in \mathcal{S}^{n_1 \times n_2}$ such that $\mathrm{dis}_{2\epsilon}(\widehat{\sigma}, H) = 0$.

Theorem 32. *Suppose $\forall u \in [n_2]$, $g_u(z)$ is strictly increasing. Let $\epsilon, \delta > 0$, $\eta \in (0, \frac{\epsilon}{2})$. Suppose that almost every $y \in \mathcal{Y}$ is $(\frac{\epsilon}{2}, \delta)$-discriminative. Let r be a positive nondecreasing function such that $r(\frac{\epsilon}{2}) \geqslant \delta$ and $r(\eta) < \frac{\delta}{2}$. Suppose that almost every $y \in \mathcal{Y}$ is r-discerning. Let $0 < \alpha < \frac{1}{2}$. If $p \geqslant \max(n_1^{-\frac{1}{2}+\alpha}, n_2^{-\frac{1}{2}+\alpha})$, $n_1 p^2 \geqslant 16$, and n_2 is sufficiently large, then Multi-Rank with $k = 1$ and $\beta = \frac{p^2 n_1}{2}$ outputs $\widehat{\sigma} \in \mathcal{S}^{n_1 \times n_2}$ such that*

$$\begin{aligned}\Pr_{\{\boldsymbol{x}_i\},\{\boldsymbol{y}_u\},\boldsymbol{\Omega}}(\mathrm{dis}_{2\epsilon}(\widehat{\sigma}, H) > 0) \leqslant & n_2 \binom{n_1}{2} [2\exp(-\frac{(n_2-1)p^2}{12}) + (n_2 - 1)\exp(-\frac{n_1 p^2}{8}) \\ & + \exp(-(\frac{(n_2-1)p^2}{2})\tau(\eta)) + 3(n_2-1)p^2 \exp(-\frac{\delta^2 n_1 p^2}{20})].\end{aligned}$$

A couple of remarks are in order. First, if ϵ and δ are small, then η must be correspondingly small. η represents how close a user y_v must be to a user y_u in the feature space to guarantee that the ratings of y_v can be used to make inferences about the ranking of user y_u. Second, whereas we require that $p \geqslant n_2^{-\frac{1}{2}+\alpha}$, Lee et al. (2016) require that $p \geqslant n_2^{-1+\alpha}$. We conjecture that this stronger requirement is fundamental to our algorithm since $v \in W_u^{i,j}(\beta)$ only if v has rated both items i and j, which v does with probability p^2. However, there may be another algorithm that circumvents this issue. Theorem 32 implies the following Corollary.

Corollary 4. *Assume the setting of Theorem 32. If* $n_2 \longrightarrow \infty$, $p \geqslant \max(n_1^{-\frac{1}{2}+\alpha}, n_2^{-\frac{1}{2}+\alpha})$, *and* $n_2^{C_1} \geqslant n_1 \geqslant C_2 \log(n_2)^{\frac{1}{2\alpha}}$ *for any constant* $C_1 > 0$ *and some constant* $C_2 > 0$ *depending on* α, *then* $\Pr_{\{\boldsymbol{x}_i\},\{\boldsymbol{y}_u\},\Omega}(\mathrm{dis}_{2\epsilon}(\widehat{\sigma}, H) > 0) \longrightarrow 0$ *as* $n_2 \longrightarrow \infty$.

Note that the growth rates of n_1, n_2 and p imply that the average number of rated items by each user pn_1 must grow as $C \log(n_2)^{\frac{1}{2}+\frac{1}{4\alpha}}$ for some universal constant $C > 0$.

Next, we sketch the proof. The main part of the analysis deals with establishing a probability bound of a mistake by Pairwise-Rank for a specific user u and a pair of items i and j when $|f(\boldsymbol{x}_i, \boldsymbol{y}_u) - f(\boldsymbol{x}_j, \boldsymbol{y}_u)| > \epsilon$. First, we establish that w.h.p. $|W_u^{i,j}(\beta)|$ is large, i.e., there are many users that have rated i and j and many other items in common with u. Second, using standard concentration bounds, it is shown that for every $v \in W_u^{i,j}(\beta)$, $R_{u,v}$ concentrates around $\rho(u, v)$. Since $\beta \longrightarrow \infty$, this estimate converges to $\rho(u, v)$. Third, we show that eventually we sample a point from $B_\eta(\boldsymbol{y}_u)$. Further, if $\boldsymbol{y}_v \in B_\eta(\boldsymbol{y}_u)$ and $\boldsymbol{y}_w \in B_{\frac{\epsilon}{2}}(\boldsymbol{y}_u)^c$ (note $\eta \leqslant \frac{\epsilon}{2}$), then since $\boldsymbol{y}_u$ is $(\frac{\epsilon}{2}, \delta)$-discriminative w.p. 1, by our choice of η, $\rho(\boldsymbol{y}_u, \boldsymbol{y}_v) > \rho(\boldsymbol{y}_u, \boldsymbol{y}_w) + \frac{\delta}{2}$. Thus, by concentration bounds, $R_{u,v} > R_{u,w}$. Therefore, Pairwise-Rank with $k = 1$ uses the preference ordering of a user in $B_{\frac{\epsilon}{2}}(\boldsymbol{y}_u)$ on items i and j to make the prediction. The Lipschitzness of f and our assumption that g_v is strictly increasing imply that this prediction is correct. It is possible to extend this argument to handle the case when $k > 1$.

6.5.2 Discrete Ratings Setting

Let $N > 0$ and suppose that $|f(x, y)| \leqslant N \ \forall x \in \mathcal{X}, \ \forall y \in \mathcal{Y}$. Suppose that there are L distinct ratings and let $\mathcal{G}$ denote the set of all step functions of the form

$$g_u(x) = \begin{cases} 1 & : x \in [-N, a_{u,1}) \\ 2 & : x \in [a_{u,1}, a_{u,2}) \\ \vdots & \\ L & : x \in [a_{u,L-1}, N] \end{cases}.$$

We assume that for all $u \in [n_2]$, $g_u \in \mathcal{G}$ and that the rating thresholds are random, i.e., $(\boldsymbol{a}_{1,1}, \ldots, \boldsymbol{a}_{1,L-1}), \ldots, (\boldsymbol{a}_{n_2,1}, \ldots, \boldsymbol{a}_{n_2,L-1}) \overset{i.i.d.}{\sim} \mathcal{P}_{[-N,N]^{L-1}}$. We write $\boldsymbol{g}_1, \ldots, \boldsymbol{g}_{n_2} \overset{i.i.d.}{\sim} \mathcal{P}_{\mathcal{G}}$ and

we assume that $\{\boldsymbol{g}_u\}_{u\in[n_2]}$ is independent from $\{\boldsymbol{x}_i\}_{i\in[n_1]}$, $\{\boldsymbol{y}_u\}_{u\in[n_2]}$, and $\boldsymbol{\Omega}$. Let $\mathcal{P}_l$ denote the marginal distribution of $\boldsymbol{a}_{u,l}$ for all $u \in [n_2]$. We make the following assumption.

Definition 25. *We say that $\mathcal{P}_{\mathcal{G}}$ is* diverse *if for every open interval $I \subset [-N, N]$ there exists l such that $\mathcal{P}_l(I) > 0$.*

Let $d_{\mathbb{R}}$ denote a metric on $\mathbb{R}$; fix $u \in [n_2]$ and let $\gamma(\epsilon) = \inf_{z\in[-N,N]} \mathcal{P}_{\{\boldsymbol{a}_{u,l}\}_{l\in[L-1]}}(\exists l \in [L-1] : d_{\mathbb{R}}(z, \boldsymbol{a}_{u,l}) \leqslant \epsilon)$. The aforementioned assumption ensures via a measure theoretic argument that $\gamma(\epsilon) > 0$ for all $\epsilon > 0$ (see Lemma 47 in the Appendix).

Theorem 33. *Let $\epsilon, \delta > 0$ and $\eta \in (0, \frac{\epsilon}{4})$. Suppose that $\mathcal{P}_{\mathcal{G}}$ is diverse and that almost every $y \in \mathcal{Y}$ is $(\frac{\epsilon}{4}, \delta)$-discriminative. Let r be a positive nondecreasing function such that $r(\frac{\epsilon}{4}) \geqslant \delta$ and $r(\eta) < \frac{\delta}{2}$. Suppose that almost every $y \in \mathcal{Y}$ is r-discerning. Let $\frac{1}{2} > \alpha > \alpha' > 0$. If $p \geqslant \max(n_1^{-\frac{1}{2}+\alpha}, n_2^{-\frac{1}{2}+\alpha})$, $n_1 p^2 \geqslant 16$, $n_1 \geqslant C_1 \log(n_2)^{\frac{1}{2\alpha}}$ for some constant C_1, and n_2 is sufficiently large, Multi-Rank with $k = n_2^{\alpha'}$ and $\beta = \frac{p^2 n_1}{2}$ outputs $\widehat{\sigma}$ such that*

$$
\begin{aligned}
&\Pr_{\{\boldsymbol{x}_i\},\{\boldsymbol{y}_u\},\{\boldsymbol{a}_{u,l}\},\boldsymbol{\Omega}}(\mathrm{dis}_{2\epsilon}(\widehat{\sigma}, H) > 0) \\
&\qquad \leqslant n_2 \binom{n_1}{2} [2\exp(-\frac{(n_2-1)p^2}{12}) + (n_2-1)\exp(-\frac{n_1 p^2}{8}) + 2\exp(-\gamma(\frac{\epsilon}{4})k) \\
&\qquad + \frac{1}{1-r(\frac{\epsilon}{2})}[3(n_2-1)p^2\exp(-\frac{\delta^2 n_1 p^2}{20}) \\
&\qquad + \exp([1-\kappa(\frac{\epsilon}{4}) + \tau(\eta) + \log(3\frac{(n_2-1)p^2}{2})]k - k\log(k) - \tau(\eta)\frac{(n_2-1)p^2}{2})]].
\end{aligned}
$$

Corollary 5. *Assume the setting of Theorem 33. If $p \geqslant \max(n_1^{-\frac{1}{2}+\alpha}, n_2^{-\frac{1}{2}+\alpha})$, $k = n_2^{\alpha'}$, and $n_2^{C_1} \geqslant n_1 \geqslant C_2 \log(n_2)^{\frac{1}{2\alpha}}$ for any constant $C_1 > 0$ and some constant $C_2 > 0$ depending on α, then $\Pr_{\{\boldsymbol{x}_i\},\{\boldsymbol{y}_u\},\Omega}(\mathrm{dis}_{2\epsilon}(\widehat{\sigma}, H) > 0) \longrightarrow 0$ as $n_2 \longrightarrow \infty$.*

The bulk of the analysis for the discrete ratings setting is similar to the continuous rating setting and, once again, mainly deals with the analysis of Pairwise-Rank for a user u and items i and j. Since the ratings are discrete, although users that are sufficiently close to user u in the feature space agree about the ordering of items i and j, we need to show that at least one of these neighbors does not give the same rating to items i and j. To this end, we show that eventually k nearby points are sampled: $\boldsymbol{y}_{v_1}, \ldots, \boldsymbol{y}_{v_k} \in B_\eta(\boldsymbol{y}_u)$. Conditional on $|f(\boldsymbol{x}_i, \boldsymbol{y}_u) - f(\boldsymbol{x}_j, \boldsymbol{y}_u)| > \epsilon$, using the Lipschitzness of f, $(f(\boldsymbol{x}_i, \boldsymbol{y}_{v_q}), f(\boldsymbol{x}_j, \boldsymbol{y}_{v_q}))$ has length at

least $\frac{\epsilon}{2}$. Finally, since $\mathcal{P}_{\mathcal{G}}$ is diverse, a concentration argument wrt $\boldsymbol{g}_{v_1}, \ldots, \boldsymbol{g}_{v_k}$ implies that w.h.p. there exists $q \in [k]$ and $l \in [L-1]$ such that $\boldsymbol{a}_{v_q,l} \in (f(\boldsymbol{x}_i, \boldsymbol{y}_{v_q}), f(\boldsymbol{x}_j, \boldsymbol{y}_{v_q}))$. Thus, user v_q provides distinct ratings for items i and j.

6.6 A Necessary and Sufficient Condition for $\mathrm{dis}_\epsilon(\sigma, H) = 0$

In this section, we characterize the class of optimal collections of rankings, i.e., $\sigma \in \mathcal{S}^{n_1 \times n_2}$ such that $\mathrm{dis}_\epsilon(\sigma, H) = 0$. We show roughly that a collection of rankings σ is optimal in the sense that $\mathrm{dis}_\epsilon(\sigma, H) = 0$ if and only if σ agrees with the observed data and σ gives the same ranking to users that are close to each other in the latent space $\mathcal{Y}$. To study this question, we consider the regime where the number of items n_1 is fixed, the probability of an entry being revealed p is fixed, and the number of users n_2 goes to infinity.

Consider the following notion, which is the main ingredient in our necessary and sufficient condition:

Definition 26. *Let $\epsilon > 0$ and $T \subset [n_1] \times [n_1] \times [n_2]$. $\sigma \in \mathcal{S}^{n_1 \times n_2}$ is an ϵ-consistent collection of rankings over T if $\forall i \neq j \in [n_1], u \neq v \in [n_2]$ such that $(i,j,u), (i,j,v) \in T$ and $d_{\mathcal{Y}}(y_u, y_v) \leqslant \epsilon$, it holds that $\sigma(i,u) < \sigma(j,u) \iff \sigma(i,v) < \sigma(j,v)$. If σ is an ϵ-consistent collection of rankings over $[n_1] \times [n_1] \times [n_2]$, then we simply say that σ is an ϵ-consistent collection of rankings.*

In words, a collection of rankings is ϵ-consistent if it gives the same ranking to users that are within ϵ of each other in the latent space.

We introduce the following objective function: $\widehat{\mathrm{dis}}(\sigma, H) \coloneqq$

$$\sum_{u=1}^{n_2} \sum_{i<j:(i,u),(j,u)\in\Omega} \mathbf{1}\{(h_u(x_i, y_u) - h_u(x_j, y_u))(\sigma(i,u) - \sigma(j,u)) < 0\}.$$

Once again, we analyze separately the continuous rating and discrete rating settings. With respect to the continuous rating setting, Theorems 34 and 36 *roughly* imply that with probability tending to 1 as $n_2 \longrightarrow \infty$, a collection of rankings $\sigma \in \mathcal{S}^{n_1 \times n_2}$ that minimizes $\widehat{\mathrm{dis}}(\cdot, H)$

is $\frac{\epsilon}{2}$-consistent if and only if $\text{dis}_\epsilon(\sigma, H) = 0$. A similar statement holds for the discrete rating setting.

To begin, we present our sufficient conditions.

Theorem 34. *Assume the continuous rating setting. Let $\epsilon > 0$ and suppose that for all $u \in [n_2]$, $g_u(\cdot)$ is strictly increasing. With probability increasing to 1 as $n_2 \longrightarrow \infty$, if $\sigma \in \mathcal{S}^{n_1 \times n_2}$ is $\frac{\epsilon}{2}$-consistent and minimizes $\widehat{\text{dis}}(\cdot, H)$, then $\text{dis}_\epsilon(\sigma, H) = 0$.*

Theorem 35. *Assume the discrete rating setting and that $\mathcal{P}_\mathcal{G}$ is diverse. Let $\epsilon > 0$. With probability increasing to 1 as $n_2 \longrightarrow \infty$, if $\sigma \in \mathcal{S}^{n_1 \times n_2}$ is $\frac{\epsilon}{8}$-consistent and minimizes $\widehat{\text{dis}}(\cdot, H)$, then $\text{dis}_\epsilon(\sigma, H) = 0$.*

The proofs for the continuous and discrete cases are similar. We briefly sketch the argument for the continuous case. Since $\mathcal{Y}$ is compact, there is a finite subcover of $\mathcal{Y}$ with open balls with diameter at most $\frac{\epsilon}{2}$. As $n_2 \longrightarrow \infty$, with probability increasing to 1, for every open ball $\mathcal{O}$ belonging to the finite subcover and for every pair of distinct items $i, j \in [n_1]$, there is some user $u \in \mathcal{O}$ that has rated i and j. Then, on this event, it can be shown that if $\text{dis}_\epsilon(\sigma, H) > 0$, then $\widehat{\text{dis}}(\sigma, H) > 0$. Thus, using the contrapositive, the result follows.

Theorem 36 gives our necessary condition.

Theorem 36. *Let $\epsilon > 0$ and $\sigma \in \mathcal{S}^{n_1 \times n_2}$ such that $\text{dis}_\epsilon(\sigma, H) = 0$. Let $T = \{(i, j, u) \in [n_1] \times [n_1] \times [n_2] : |f(x_i, y_u) - f(x_j, y_u)| > \epsilon,\ h(x_i, y_u) \neq h(x_j, y_u)\}$. Then, σ is an ϵ-consistent collection of rankings over T.*

Theorem 36 shows that in our general setting, learning the correct collection of rankings requires giving the same ranking to nearby users. In particular, this provides an intuition on the kind of embedding that matrix factorization learns. Theorem 36 only applies to items i, j and user u if there is a large enough difference in the underlying values given by f. The proof follows by the Lipschitzness of f and algebra.

6.7 Experiments

In this section, we examine the empirical performance of Multi-Rank. It is well-known that matrix factorization methods tend to outperform neighborhood-based methods. Neverthe-

Dataset	Method	Kendall Tau	Spearman Rho	NDCG@5	Precision@5
	MRW	0.3156 (0.0021)	0.4012 (0.0029)	0.7104 (0.0010)	0.4492 (0.0018)
	MR	0.3105 (0.0021)	0.3963 (0.0030)	0.7063 (0.0038)	0.4457 (0.0044)
	LA	0.3271 (0.0018)	0.4153 (0.0022)	0.7136 (0.0026)	0.4570 (0.0041)
	AltSVM	0.3271 (0.0007)	0.4173 (0.0008)	0.7022 (0.0015)	0.4365 (0.0036)
Netflix	RMC	0.3288 (0.0017)	0.4178 (0.0020)	0.7204 (0.0006)	0.4581 (0.0048)
	MRW	0.3933 (0.0010)	0.5009 (0.0013)	0.7769 (0.0066)	0.6083 (0.0096)
	MR	0.3924 (0.0011)	0.4999 (0.0013)	0.7735 (0.0061)	0.6021 (0.0063)
	LA	0.3993 (0.0009)	0.5075 (0.0012)	0.7767 (0.0058)	0.6071 (0.0080)
	AltSVM	0.4099 (0.0008)	0.5219 (0.0010)	0.8002 (0.0042)	0.6417 (0.0067)
Movielens	RMC	0.4041 (0.0004)	0.5139 (0.0006)	0.8068 (0.0030)	0.6485 (0.0029)

Table 6.1: Netflix and MovieLens Results. On the Netflix dataset, MR usually used $\beta = 5$ and $k \in [13, 19]$. MRW usually used $\beta = 9$ and $k \in [16, 23]$. On the MovieLens dataset, MR usually used $\beta = 10$ and $k \in [7, 13]$. MRW usually used $\beta = 12$ and $k \in [13, 17]$.

less, neighborhood-based methods remain popular in situations where practitioners want an easy-to-implement method, to avoid expensive model-building, and to be able to interpret predictions easily (Ning et al., 2011). Furthermore, it has been observed that for the task of matrix completion, (i) matrix factorization methods and neighborhood-based methods have complementary strengths and weaknesses and (ii) performance gains can be achieved by merging these methods into a single algorithm (Bell and Koren, 2007; Koren, 2008). Yet, it is non-trivial to generalize ideas for combining matrix factorization and neighborhood-based methods in the matrix completion setting to the preference completion setting. In light of this discussion, the purpose of our experiments is not to demonstrate the superiority of our method over matrix factorization methods, but to compare the performance of our algorithm with the state-of-the-art.

We compared the performance of our algorithm (MR) and a weighted version of our algorithm (MRW) where votes are weighted by $R_{u,v}$ against Alternating SVM (AltSVM) (Park et al., 2015), Retargeted Matrix Completion (RMC) (Gunasekar et al., 2016), and the proposed algorithm in (Lee et al., 2016) (LA). We chose AltSVM and RMC because they are state-of-the-art matrix factorization methods for preference completion and we chose LA because its theoretical guarantees are similar to our guarantees for Multi-Rank and it was

Dataset	Method	Kendall Tau	Spearman Rho	NDCG@5	Precision@5
	MRW	0.2736 (0.0017)	0.3327 (0.0021)	0.8063 (0.0031)	0.7849 (0.0034)
	MR	0.2677 (0.0019)	0.3255 (0.0023)	0.7980 (0.0034)	0.7764 (0.0008)
	LA	0.2786 (0.0022)	0.3387 (0.0027)	0.8024 (0.0024)	0.7843 (0.0018)
	AltSVM	0.2743 (0.0015)	0.3335 (0.0018)	0.7949 (0.0023)	0.7768 (0.0023)
Netflix	RMC	0.2856 (0.0017)	0.3473 (0.0021)	0.8052 (0.0038)	0.7861 (0.0032)
	MRW	0.3347 (0.0015)	0.4090 (0.0018)	0.8903 (0.0059)	0.8810 (0.0059)
	MR	0.3343 (0.0017)	0.4085 (0.0021)	0.8879 (0.0052)	0.8792 (0.0061)
	LA	0.3395 (0.0017)	0.4149 (0.0020)	0.8908 (0.0085)	0.8845 (0.0078)
	AltSVM	0.3451 (0.0016)	0.4217 (0.0020)	0.9070 (0.0056)	0.8982 (0.0056)
Movielens	RMC	0.3504 (0.0014)	0.4281 (0.0017)	0.9140 (0.0026)	0.9051 (0.0032)

Table 6.2: Quantized Netflix and MovieLens Results. On the Netflix dataset, MR usually used $\beta = 5$ and $k = 22$. MRW usually used $\beta \in [9, 10]$ and $k \in [27, 31]$. On the MovieLens dataset, MR usually used $\beta \in [10, 13]$ and $k \in [10, 19]$. MRW usually used $\beta \in [8, 11]$ and $k \in [16, 23]$.

Dataset	Method	Kendall Tau	Spearman Rho	NDCG@5	Precision@5
Netflix	LA	0.1798 (0.0034)	0.2300 (0.0040)	0.5962 (0.0022)	0.3322 (0.0053)
MovieLens	LA	0.2404 (0.0098)	0.3092 (0.0123)	0.6543 (0.0138)	0.4435 (0.0163)

Table 6.3: Monotonically Transformed Netflix and MovieLens Results. We only display the results for LA since the other methods are invariant to monotonic transformations of the columns.

shown to be superior to item-based and user-based neighborhood methods (Lee et al., 2016). We used grid search to optimize the hyperparameters for each of the algorithms using a validation set.

We use the ranking metrics Kendall Tau, Spearman Rho, NDCG@5, and Precision@5. Kendall Tau and Spearman Rho measure how correlated the predicted ranking is with the true ranking. The other metrics measure the quality of the predicted ranking at the top of the list. For Precision@5, we deem an item relevant if it has a score of 5. For all of these metrics, higher scores are better. See Liu (2009) for a more detailed discussion of these metrics. The numbers in parentheses are standard deviations.

We use the Netflix and MovieLens 1M datasets. We pre-process the data in a similar way

to Liu and Yang (2008). For the Netflix dataset, we take the 2000 most popular movies and randomly selected 4000 users that had rated at least 100 of these movies. For both datasets, we randomly subsample the ratings 5 times in the following way: we randomly shuffled the (user-id, movie, rating) triples and split 40% into a training set, 15% into a validation set, and 45% into a test set. For the Netflix dataset, we drop users if they have fewer than 50 ratings in the training set and fewer than 10 ratings in either the validation set or the test set. For the MovieLens dataset, we drop users if they have fewer than 100 ratings in the training set and fewer than 50 ratings in either the validation set or the test set. Table 6.1 shows that although MRW does not have the best performance, it outperforms AltSVM on NDCG@5 and Precision@5 on the Netflix dataset and LA on NDCG@5 and Precision@5 on the MovieLens dataset.

In addition, we quantized the scores of both datasets to 1 if the true rating is less than or equal to 3 and to 5 otherwise (see Table 6.2). Here, MR and MRW have the same amount of information as LA and RMC. On the Netflix dataset, MRW performed the best on the NDCG@5 measure.

Finally, we considered a setting where a company performs A/B testing on various rating scales (e.g., 1-5, 1-10, 1-50, 1-100) and wishes to use all of the collected data to predict preferences. To model this situation, for each user, we randomly sampled a number $a \in \{1, 2, 10, 20\}$ and $b \in [a-1] \cup \{0\}$, and transformed the rating $r \mapsto a \cdot r - b$. Table 6.3 shows that on the monotonically transformed versions of the Netflix and MovieLens datasets, LA performs dramatically worse. This is unsurprising since it is well-known that the performance of rating-oriented neighborhood-based methods like LA suffers when there is rating scale variance (Ning et al., 2011).

6.8 Chapter Appendix Outline

In Section 6.9, we give the counterexample establishing Proposition 12 and give theorem proofs for the continuous rating setting. In Section 6.10, we give theorem proofs for the discrete rating setting. In Section 6.11, we prove the lemmas used in our theorem proofs, beginning with lemmas common to both the continuous rating setting and discrete rating

setting and, then, presenting the lemmas on the continuous rating setting and discrete rating setting, separately. In Section 6.12, we provide the proofs of the necessary and sufficient conditions. In Section 6.13, we prove Proposition 13 and that the models $f(x,y) = x^t y$ and $f(x,y) = \|x-y\|_2$ are equivalent by adding a dimension. Finally, in Section 6.14, we give some bounds that we use in the proofs for reference.

Unless otherwise indicated, all probability statements are with respect to $\{\boldsymbol{x}_i\}_{i\in[n_1]} \cup \{\boldsymbol{y}_u\}_{u\in[n_2]} \cup \boldsymbol{\Omega}$ in the continuous ratings setting and with respect to $\{\boldsymbol{x}_i\}_{i\in[n_1]} \cup \{\boldsymbol{y}_u\}_{u\in[n_2]} \cup \{\boldsymbol{a}_{u,l}\}_{u\in[n_2],l\in[L-1]} \cup \boldsymbol{\Omega}$ in the discrete ratings setting.

6.9 Proofs for Section 6.5.1

To begin, we introduce some additional notation. When $\boldsymbol{y}_u$ and $\boldsymbol{y}_v$ are random, we write $R_{\boldsymbol{u},\boldsymbol{v}}$ instead of $R_{u,v}$ for emphasis.

Proof of Proposition 12. Consider the functions

$$f(z) = \begin{cases} \epsilon z & : z \in [0, \frac{1}{2}] \\ \epsilon(1-z) & : z \in (\frac{1}{2}, 1] \end{cases}$$

and

$$g(z) = \begin{cases} -\epsilon z & : z \in [0, \frac{1}{2}] \\ \epsilon(z-1) & : z \in (\frac{1}{2}, 1] \end{cases}$$

□

Next, we analyze Pairwise-Rank (PR), bounding the probability that Pairwise-Rank cannot distinguish between items i and j when $|f(x_i, y_u) - f(x_j, y_u)| > \epsilon$, i.e., the event

$$\begin{aligned} D^{\epsilon}_{u,i,j} \coloneqq &\{f(\boldsymbol{x}_i, \boldsymbol{y}_u) + \epsilon < f(\boldsymbol{x}_j, \boldsymbol{y}_u)\} \cap \{\mathrm{PR}(u,i,j,\beta,k) = 1)\}) \\ &\cup \{f(\boldsymbol{x}_i, \boldsymbol{y}_u) > f(\boldsymbol{x}_j, \boldsymbol{y}_u) + \epsilon\} \cap \{\mathrm{PR}(u,i,j,\beta,k) = 0\}). \end{aligned}$$

Theorem 37. *Suppose $\forall u \in [n_2]$, $g_u(z)$ is strictly increasing. Let $\epsilon, \delta > 0$ and $\eta \in (0, \frac{\epsilon}{2})$. Suppose that almost every $y \in \mathcal{Y}$ is $(\frac{\epsilon}{2}, \delta)$-discriminative. Let r be a positive nondecreasing function such that $r(\frac{\epsilon}{2}) \geqslant \delta$ and $r(\eta) < \frac{\delta}{2}$. Suppose that almost every $y \in \mathcal{Y}$ is r-discerning.*

Let $0 < \alpha < \frac{1}{2}$. *If* $p \geqslant \max(n_1^{-\frac{1}{2}+\alpha}, n_2^{-\frac{1}{2}+\alpha})$, $n_1 p^2 \geqslant 16$, *and* n_2 *is sufficiently large, for all* $u \in [n_2]$ *and* $i \neq j \in [n_1]$, *the output of Pairwise-Rank with* $k = 1$ *and* $\beta = \frac{p^2 n_1}{2}$ *is such that*

$$\begin{aligned}\Pr_{\{\boldsymbol{x}_i\},\{\boldsymbol{y}_u\},\boldsymbol{\Omega}}(D^{\epsilon}_{u,i,j}) \leqslant\ & 2\exp(-\frac{(n_2-1)p^2}{12}) + (n_2-1)\exp(-\frac{n_1 p^2}{8}) \\ & + \exp(-(\frac{(n_2-1)p^2}{2})\tau(\eta)) + 3(n_2-1)p^2\exp(-\frac{\delta^2 n_1 p^2}{20}).\end{aligned}$$

The structure of the proof of Theorem 37 is similar to the proof of Theorem 1 from Lee et al. (2016). The lemmas are distinct, however.

Proof of Theorem 37. Fix $u \in [n_2]$, $i, j \in [n_1]$ such that $i \neq j$. Define:

$$W_u^{i,j}(\beta) = \{v \in [n_2] : |N(u,v)| \geqslant \beta, (i,v), (j,v) \in \Omega\}.$$

Further, define the events:

$$\begin{aligned}A &= \{|W_u^{i,j}(\beta)| \in [\frac{(n_2-1)p^2}{2}, \frac{3(n_2-1)p^2}{2}]\}, \\ B &= \{\max_{v \in W_u^{i,j}(\beta)} \rho(\boldsymbol{y}_u, \boldsymbol{y}_v) \geqslant 1 - \frac{\delta}{2}\}, \\ C &= \{|R_{\boldsymbol{uv}} - \rho(\boldsymbol{y}_u, \boldsymbol{y}_v)| \leqslant \frac{\delta}{4},\ \forall v \in W_u^{i,j}(\beta)\}.\end{aligned}$$

By several applications of the law of total probability, we have that

$$\begin{aligned}\Pr(D^{\epsilon}_{u,i,j}) &= \Pr(D^{\epsilon}_{u,i,j}|A,B,C)\Pr(A,B,C) + \Pr(D^{\epsilon}_{u,i,j}|(A\cap B\cap C)^c)\Pr((A\cap B\cap C)^c) \\ &\leqslant \Pr(D^{\epsilon}_{u,i,j}|A,B,C) + \Pr(A^c) + \Pr((A\cap B\cap C)^c|A) \\ &\leqslant \Pr(D^{\epsilon}_{u,i,j}|A,B,C) + \Pr(A^c) + \Pr(B^c|A) + \Pr(C^c|A,B).\end{aligned}$$

We will upper bound each term in the above bound. By Lemma 51, $\Pr(D^{\epsilon}_{u,i,j}|A,B,C) = 0$.

Setting $\lambda = \frac{1}{2}$ in Lemma 45 yields that

$$\begin{aligned}\Pr(A^c) &= \Pr(|W_u^{i,j}(\beta)| \notin [\frac{(n_2-1)p^2}{2}, \frac{3(n_2-1)p^2}{2}]) \\ &\leqslant 2\exp(-\frac{(n_2-1)p^2}{12}) + (n_2-1)\exp(-\frac{n_1 p^2}{8}).\end{aligned}$$

Lemma 49 yields that

$$\begin{aligned}\Pr(B^c|A) = \Pr(\max_{v \in W_u^{i,j}(\beta)} \rho(\boldsymbol{y}_u, \boldsymbol{y}_v) < 1 - \frac{\delta}{2}|A) &\leqslant \Pr(\max_{v \in W_u^{i,j}(\beta)} \rho(\boldsymbol{y}_u, \boldsymbol{y}_v) < 1 - r(\eta)|A) \\ &\leqslant [1-\tau(\eta)]^{\frac{(n_2-1)p^2}{2}} \qquad (6.1) \\ &\leqslant \exp(-(\frac{(n_2-1)p^2}{2})\tau(\eta)). \qquad (6.2)\end{aligned}$$

Line (6.1) follows by Lemma 49 since conditional on A, $W_u^{i,j}(\beta) \geqslant \frac{(n_1-1)p^2}{2}$ and line (6.2) follows by the inequality $1 - x \leqslant \exp(-x)$. Since by hypothesis $\alpha \in (0, \frac{1}{2})$ is fixed such that $p \geqslant \max(n_1^{-\frac{1}{2}+\alpha}, n_2^{-\frac{1}{2}+\alpha})$, there exists a sufficiently large n_2 such that line (6.2) is less than $\frac{1}{2}$. Then, by Bayes rule, the union bound, and Lemma 50,

$$\begin{aligned}
\Pr(C^c|A,B) &\leqslant \frac{\Pr(C^c|A)}{\Pr(B|A)} \leqslant 2\Pr(C^c|A) \\
&= 2\Pr(\exists v \in W_u^{i,j}(\beta), |R_{\boldsymbol{uv}} - \rho(\boldsymbol{y}_u, \boldsymbol{y}_v)| > \frac{\delta}{4}|A) \\
&\leqslant 3(n_2-1)p^2 \exp(-\frac{\delta^2}{4}\left\lfloor \frac{\beta}{2} \right\rfloor) \\
&= 3(n_2-1)p^2 \exp(-\frac{\delta^2}{4}\left\lfloor \frac{n_1p^2}{4} \right\rfloor) \\
&\leqslant 3(n_2-1)p^2 \exp(-\frac{\delta^2 n_1 p^2}{20})
\end{aligned}$$

where the last line follows because $n_1p^2 \geqslant 16$ and $\forall x \geqslant 16$, $\lfloor \frac{x}{4} \rfloor \geqslant \frac{x}{5}$. Putting it all together, we have

$$\begin{aligned}
\Pr(D_{u,i,j}^{\epsilon}) \leqslant\ & 2\exp(-\frac{(n_2-1)p^2}{12}) + (n_2-1)\exp(-\frac{n_1p^2}{8}) \\
& + \exp(-(\frac{(n_2-1)p^2}{2})\tau(\eta)) + 3(n_2-1)p^2\exp(-\frac{\delta^2 n_1 p^2}{20})
\end{aligned}$$

□

Proof of Theorem 32. For any $u \in [n_2]$, $i \neq j \in [n_1]$, define the event

$$\begin{aligned}
\text{Error}_{u,i,j}^{\epsilon} = & (\{f(\boldsymbol{x}_i, \boldsymbol{y}_u) + \epsilon < f(\boldsymbol{x}_j, \boldsymbol{y}_u)\} \cap \{A_{u,i,j} = 1\}) \\
& \cup (\{f(\boldsymbol{x}_i, \boldsymbol{y}_u) > f(\boldsymbol{x}_j, \boldsymbol{y}_u) + \epsilon\} \cap \{A_{u,i,j} = 0\}).
\end{aligned}$$

Suppose that there exists $u \in [n_2]$ and distinct $i, j \in [n_1]$ such that $\text{Error}_{u,i,j}^{\epsilon}$ occurs. Without loss of generality suppose that $f(\boldsymbol{x}_i, \boldsymbol{y}_u) + \epsilon < f(\boldsymbol{x}_j, \boldsymbol{y}_u)$, and $A_{u,i,j} = 1$. Then, inspection of the Multi-Rank algorithm reveals that $1 = A_{u,i,j} = \text{Pairwise-Rank}(u, i, j, \beta, k)$. Thus, $D_{u,i,j}^{\epsilon}$

occurs. Therefore, by Theorem 38 and the union bound,

$$\begin{aligned}
&\Pr(\exists u \in [n_2], i \neq j \in [n_1] \text{ s.t. } \text{Error}^{\epsilon}_{u,i,j}) \\
&\quad \leqslant \Pr(\exists u \in [n_2], i \neq j \in [n_1] \text{ s.t. } D^{\epsilon}_{u,i,j}) \\
&\quad \leqslant n_2 \binom{n_1}{2} [2\exp(-\frac{(n_2-1)p^2}{12}) + (n_2-1)\exp(-\frac{n_1 p^2}{8}) \\
&\quad + \exp(-(\frac{(n_2-1)p^2}{2})\tau(\eta)) + 3(n_2-1)p^2\exp(-\frac{\delta^2 n_1 p^2}{20})].
\end{aligned}$$

Now, suppose that $\forall u \in [n_2]$ and $i, j \in [n_1]$ such that $i \neq j$, $(\text{Error}^{\epsilon}_{u,i,j})^c$ occurs. Then, by Lemma 46, $\widehat{\sigma} = (\widehat{\sigma}_1, \ldots, \widehat{\sigma}_{n_2})$ with $\widehat{\sigma}_u = \text{Copeland}(A_{u,:,:})$ satisfies $\text{dis}_{2\epsilon}(\widehat{\sigma}, H) = 0$. □

Proof of Corollary 4. Ignoring constants, the two dominant terms in the bound in Theorem 32 are of the form $n_1^2 n_2 \exp(-n_2 p^2)$ and $n_1^2 n_2^2 \exp(-n_1 p^2)$. Then, under the conditions of Theorem 37, as $n_2 \longleftarrow \infty$

$$\begin{aligned}
n_1^2 n_2 \exp(-n_2 p^2) &\leqslant \exp(2\log(n_1) + \log(n_2) - n_2^{2\alpha}) \\
&\leqslant \exp((1+2C_1)\log(n_2) - n_2^{2\alpha}) \longrightarrow 0.
\end{aligned}$$

Now, observe that

$$\begin{aligned}
n_1^2 n_2^2 \exp(-n_1 p^2) &= \exp(2\log(n_2) + 2\log(n_1) - n_1 p^2) \\
&\leqslant \exp(2\log(n_2) + 2\log(n_1) - n_1^{2\alpha}) \\
&\leqslant \exp(4\max(\log(n_2), \log(n_1)) - n_1^{2\alpha})
\end{aligned}$$

Suppose that $n_1 \geqslant n_2$. Then, clearly, the limit of the RHS as $n_2 \longrightarrow \infty$ is 0. Now, suppose that $n_1 < n_2$. Then, if $C_2^{2\alpha} > 4$, then as $n_2 \longrightarrow \infty$,

$$\begin{aligned}
n_2^2 n_1^2 \exp(-n_1 p^2) &\leqslant \exp(4\log(n_2) - n_1^{2\alpha}) \\
&\leqslant \exp([4 - C_2^{2\alpha}]\log(n_2)) \longrightarrow 0.
\end{aligned}$$

□

6.10 Proofs for Section 6.5.2

To begin, because the model for the discrete ratings section is different, we introduce new notation in the interest of clarity. Fix $y_u, y_v \in \mathcal{Y}$. Define

$$\rho'(y_u, y_v) = \Pr_{\boldsymbol{g}_u, \boldsymbol{g}_v, \boldsymbol{x}_s, \boldsymbol{x}_t}[\boldsymbol{g}_u(f(\boldsymbol{x}_s, y_u)) - \boldsymbol{g}_u(f(\boldsymbol{x}_t, y_u))][\boldsymbol{g}_v(f(\boldsymbol{x}_s, y_v)) - \boldsymbol{g}_v(f(\boldsymbol{x}_t, y_v))] \geqslant 0).$$

Note that in this setting, the meaning of (ϵ, δ)-discriminative is slightly different.

Definition 27. *Fix $y \in \mathcal{Y}$. Let $\epsilon, \delta > 0$. We say that y is* (ϵ, δ)-discriminative *if $z \in B_\epsilon(y)^c$ implies that $\rho'(y, z) < 1 - \delta$.*

In a sense, the notion requires in addition that the distribution of the monotonic functions reveals some differences in the preferences of the users.

Unless otherwise indicated, all probability statements are with respect to $\{\boldsymbol{x}_i\}_{i\in[n_1]} \cup \{\boldsymbol{y}_u\}_{u\in[n_2]} \cup \{\boldsymbol{a}_{u,l}\}_{u\in[n_2],l\in[L-1]} \cup \boldsymbol{\Omega}$. Next, we prove a theorem that is analogous to Theorem 37. Recall the notation:

$$\begin{aligned} D^{\epsilon}_{u,i,j} \coloneqq &(\{f(\boldsymbol{x}_i, \boldsymbol{y}_u) + \epsilon < f(\boldsymbol{x}_j, \boldsymbol{y}_u)\} \cap \{\mathrm{PR}(u, i, j, \beta, k) = 1)\}) \\ &\cup (\{f(\boldsymbol{x}_i, \boldsymbol{y}_u) > f(\boldsymbol{x}_j, \boldsymbol{y}_u) + \epsilon\} \cap \{\mathrm{PR}(u, i, j, \beta, k) = 0\}). \end{aligned}$$

Theorem 38. *Let $\epsilon, \delta > 0$ and $\eta \in (0, \frac{\epsilon}{4})$. Suppose that $\mathcal{P}_{\mathcal{G}}$ is diverse and that almost every $y \in \mathcal{Y}$ is $(\frac{\epsilon}{4}, \delta)$-discriminative. Let r be a positive nondecreasing function such that $r(\frac{\epsilon}{4}) \geqslant \delta$ and $r(\eta) < \frac{\delta}{2}$. Suppose that almost every $y \in \mathcal{Y}$ is r-discerning. Let $\frac{1}{2} > \alpha > \alpha' > 0$. If $p \geqslant \max(n_1^{-\frac{1}{2}+\alpha}, n_2^{-\frac{1}{2}+\alpha})$, $n_1 p^2 \geqslant 16$, $n_1 \geqslant C_1 \log(n_2)^{\frac{1}{2\alpha}}$ for some suitable universal constant C_1, and n_2 is sufficiently large, for all $u \in [n_2]$ and $i \neq j \in [n_1]$, the output of Pairwise-Rank with $k = n_2^{\alpha'}$ and $\beta = \frac{p^2 n_1}{2}$ is such that*

$$\begin{aligned} \Pr_{\{\boldsymbol{x}_i\},\{\boldsymbol{y}_u\},\{\boldsymbol{a}_{u,l}\},\boldsymbol{\Omega}}(D^{\epsilon}_{u,i,j}) \leqslant &2\exp(-\frac{(n_2-1)p^2}{12}) + (n_2-1)\exp(-\frac{n_1 p^2}{8}) + 2\exp(-\gamma(\frac{\epsilon}{4})k) \\ &+ \frac{1}{1 - r(\frac{\epsilon}{2})}[3(n_2-1)p^2\exp(-\frac{\delta^2 n_1 p^2}{20}) \\ &+ \exp([1 - \kappa(\frac{\epsilon}{2}) + \tau(\eta) + \log(\frac{3(n_2-1)p^2}{2})]k \\ &- k\log(k) - \tau(\eta)\frac{(n_2-1)p^2}{2})]. \end{aligned}$$

Proof of Theorem 38. Fix $u \in [n_2]$, $i, j \in [n_1]$ such that $i \neq j$. Define:

$$W_u^{i,j}(\beta) = \{v \in [n_2] : |N(u,v)| \geqslant \beta, (i,v), (j,v) \in \Omega\}.$$

Further, define the events:

$$\begin{aligned}
A &= \{|W_u^{i,j}(\beta)| \in [\frac{(n_2-1)p^2}{2}, \frac{3(n_2-1)p^2}{2}]\},\\
B &= \{\max\nolimits^{(k)}_{v \in W_u^{i,j}(\beta)} \rho'(\boldsymbol{y}_u, \boldsymbol{y}_v) \geqslant 1 - \frac{\delta}{2}\},\\
C &= \{|R_{\boldsymbol{uv}} - \rho'(\boldsymbol{y}_u, \boldsymbol{y}_v)| \leqslant \frac{\delta}{4}, \forall v \in W_u^{i,j}(\beta)\}\\
E &= \{|f(\boldsymbol{x}_i, \boldsymbol{y}_u) - f(\boldsymbol{x}_j, \boldsymbol{y}_u)| > \epsilon\}\\
M &= \{\exists v \in W_u^{i,j}(\beta) \text{ s.t. } \rho'(\boldsymbol{y}_u, \boldsymbol{y}_v) \geqslant 1 - \frac{\delta}{2} \text{ and } \exists l \in [L-1] \text{ s.t. } \boldsymbol{a}_{v,l} \in (f(\boldsymbol{x}_j, \boldsymbol{y}_v), f(\boldsymbol{x}_i, \boldsymbol{y}_v))\}
\end{aligned}$$

By several applications of the law of total probability, we have that

$$\begin{aligned}
\Pr(D^{\epsilon}_{u,i,j}) &\leqslant \Pr(D^{\epsilon}_{u,i,j}|E) + \Pr(D^{\epsilon}_{u,i,j}|E^c)\\
&= \Pr(D^{\epsilon}_{u,i,j}|E)\\
&\leqslant \Pr(D^{\epsilon}_{u,i,j}|A,B,C,M,E) + \Pr(A^c|E) + \Pr(B^c|A,E)\\
&+ \Pr(C^c|A,B,E) + \Pr(M^c|A,B,C,E)\\
&= \Pr(D^{\epsilon}_{u,i,j}|A,B,C,M,E) + \Pr(A^c) + \Pr(B^c|A,E) \qquad (6.3)\\
&+ \Pr(C^c|A,B,E) + \Pr(M^c|A,B,C,E)
\end{aligned}$$

Line (6.3) follows from the independence of $\boldsymbol{\Omega}$ from $\{\boldsymbol{x}_s\}_{s\in[n_1]}$ and $\{\boldsymbol{y}_v\}_{v\in[n_2]}$. We will bound each term in the above upper bound. By Lemma 56,

$$\Pr(D^{\epsilon}_{u,i,j}|A,B,C,M,E) = 0. \qquad (6.4)$$

Setting $\lambda = \frac{1}{2}$ in Lemma 45 yields that

$$\begin{aligned}
\Pr(A^c) &= \Pr(|W_u^{i,j}(\beta)| \notin [\frac{(n_2-1)p^2}{2}, \frac{3(n_2-1)p^2}{2}])\\
&\leqslant 2\exp(-\frac{(n_2-1)p^2}{12}) + (n_2-1)\exp(-\frac{n_1p^2}{8}). \qquad (6.5)
\end{aligned}$$

Next, we bound $\Pr(B^c|A,E)$. By Bayes theorem,

$$\begin{aligned}\Pr(B^c|A,E) &\leqslant \frac{\Pr(B^c|A)}{\Pr(E|A)} \\ &= \frac{\Pr(B^c|A)}{\Pr(E)} \qquad (6.6) \\ &< \frac{\Pr(B^c|A)}{1-r(\frac{\epsilon}{2})}. \qquad (6.7)\end{aligned}$$

Line (6.6) follows from the independence of $\boldsymbol{\Omega}$ from $\{\boldsymbol{x}_s\}_{s\in[n_1]}$ and $\{\boldsymbol{y}_v\}_{v\in[n_2]}$. Line (6.7) follows since by hypothesis almost every $y \in \mathcal{Y}$ is r-discerning.

Since almost every $y \in \mathcal{Y}$ is $(\frac{\epsilon}{4},\delta)$-discriminative and r-discerning, and $\eta > 0$ is such that $r(\eta) < \frac{\delta}{2}$, Lemma 53 yields that

$$\begin{aligned}\Pr(B^c|A) &= \Pr(\max^{(k)}_{v\in W_u^{i,j}(\beta)}\rho'(\boldsymbol{y}_u,\boldsymbol{y}_v) < 1-\frac{\delta}{2}|A) \\ &\leqslant \Pr(\max^{(k)}_{v\in W_u^{i,j}(\beta)}\rho'(\boldsymbol{y}_u,\boldsymbol{y}_v) < 1-r(\eta)|A) \\ &\leqslant \exp((1-\kappa(\frac{\epsilon}{4})+\tau(\eta)+\log(3\frac{(n_2-1)p^2}{2}))k - k\log(k) - \tau(\eta)\frac{(n_2-1)p^2}{2})).\end{aligned} \qquad (6.8)$$

Next, we bound $\Pr(C^c|A,B,E)$. By Bayes theorem,

$$\Pr(C^c|A,B,E) \leqslant \frac{\Pr(C^c|A,B)}{\Pr(E|A,B)}.$$

Fix $\boldsymbol{y}_u = y_u$ r-discerning such that A and B occur. Then, since $\{\boldsymbol{x}_s\}_{s\in[n_1]}$, $\{\boldsymbol{y}_v\}_{v\in[n_2]}$, and $\boldsymbol{\Omega}$ are independent and y_u is r-discerning,

$$\begin{aligned}&\Pr_{\{\boldsymbol{y}_v\}_{v\in[n_2]},\{\boldsymbol{x}_s\}_{s\in[n_1]},\boldsymbol{\Omega}}(|f(\boldsymbol{x}_i,y_u)-f(\boldsymbol{x}_j,y_u)|>\epsilon|\boldsymbol{y}_u=y_u) \\ &\quad= \Pr_{\boldsymbol{x}_i,\boldsymbol{x}_j}(|f(\boldsymbol{x}_i,y_u)-f(\boldsymbol{x}_j,y_u)|>\epsilon|\boldsymbol{y}_u=y_u) \\ &\quad= \Pr_{\boldsymbol{x}_i,\boldsymbol{x}_j}(|f(\boldsymbol{x}_i,y_u)-f(\boldsymbol{x}_j,y_u)|>\epsilon) > 1-r(\frac{\epsilon}{2}).\end{aligned}$$

Since the above bound holds for all y_u such that $A\cap B$ holds, taking the expectation of the above bound with respect to $\boldsymbol{y}_u$ over the set $A\cap B$ gives

$$\Pr(|f(\boldsymbol{x}_i,\boldsymbol{y}_u)-f(\boldsymbol{x}_j,\boldsymbol{y}_u)|>\epsilon|A,B) > 1-r(\frac{\epsilon}{2}).$$

Thus,

$$\Pr(C^c|A,B,E) < \frac{\Pr(C^c|A,B)}{1-r(\frac{\epsilon}{2})}. \qquad (6.9)$$

Since by hypothesis $\frac{1}{2} > \alpha > \alpha' > 0$, $p \geqslant \max(n_1^{-\frac{1}{2}+\alpha}, n_2^{-\frac{1}{2}+\alpha})$ and $k = n_2^{\alpha'}$, if n_2 is sufficiently large, the bound in line (6.8) is less than $\frac{1}{2}$. Then, by Bayes rule, the union bound, and Lemma 54,

$$\begin{aligned}
\Pr(C^c|A,B) &\leqslant \frac{\Pr(C^c|A)}{\Pr(B|A)} \leqslant 2\Pr(C^c|A) \\
&= 2\Pr(\exists v \in W_u^{i,j}(\beta), |R_{\boldsymbol{uv}} - \rho'(\boldsymbol{y}_u, \boldsymbol{y}_v)| > \frac{\delta}{4}|A) \\
&\leqslant 3(n_2-1)p^2 \Pr(|R_{\boldsymbol{uv}} - \rho'(\boldsymbol{y}_u, \boldsymbol{y}_v)| > \frac{\delta}{4}|A) \qquad (6.10) \\
&\leqslant 3(n_2-1)p^2 \exp(-\frac{\delta^2}{4}\left\lfloor\frac{\beta}{2}\right\rfloor) \\
&= 3(n_2-1)p^2 \exp(-\frac{\delta^2}{4}\left\lfloor\frac{n_1p^2}{4}\right\rfloor) \\
&\leqslant 3(n_2-1)p^2 \exp(-\frac{\delta^2 n_1 p^2}{20}) \qquad (6.11)
\end{aligned}$$

where line (6.10) follows by the union bound and line (6.11) follows because $n_1p^2 \geqslant 16$ and $\forall x \geqslant 15$, $\lfloor\frac{x}{4}\rfloor \geqslant \frac{x}{5}$.

Since by hypothesis $\frac{1}{2} > \alpha > 0$, $p \geqslant \max(n_1^{-\frac{1}{2}+\alpha}, n_2^{-\frac{1}{2}+\alpha})$, and $n_1 \geqslant C_1 \log(n_2)^{\frac{1}{2\alpha}}$ for some constant C_1, if n_2 is sufficiently large, the bound in line (6.9) is eventually less than $\frac{1}{2}$. Thus, using Bayes rule and Lemma 55,

$$\begin{aligned}
\Pr(M^c|A,B,C,E) &\leqslant \frac{\Pr(M^c|A,B,E)}{\Pr(C|A,B,E)} \\
&\leqslant 2\Pr(M^c|A,B,E) \\
&\leqslant 2\exp(-\gamma(\frac{\epsilon}{4})k). \qquad (6.12)
\end{aligned}$$

Putting together lines (6.3), (6.4), (6.5), (6.7), (6.8), (6.9), (6.11), and (6.12) we have

$$\begin{aligned}
\Pr(D^{\epsilon}_{u,i,j}) \leqslant & 2\exp(-\frac{(n_2-1)p^2}{12}) + (n_2-1)\exp(-\frac{n_1p^2}{8}) + 2\exp(-\gamma(\frac{\epsilon}{4})k) \\
& + \frac{1}{1-r(\frac{\epsilon}{2})}[3(n_2-1)p^2\exp(-\frac{\delta^2 n_1 p^2}{20}) \\
& + \exp([1-\kappa(\frac{\epsilon}{4}) + \tau(\eta) + \log(3\frac{(n_2-1)p^2}{2})]k - k\log(k) - \tau(\eta)\frac{(n_2-1)p^2}{2})].
\end{aligned}$$

□

Proof of Theorem 33. The proof follows the same steps as the proof of Theorem 32, but applies Theorem 38 instead of Theorem 37. □

Proof of Corollary 5. The only new term that did not appear in Corollary 5 is, ignoring constants, of the form

$$n_2^2 n_1^2 \exp(\log(n_2 p^2)k - k\log(k) - n_2 p^2).$$

Using $\alpha > \alpha'$ and $n_1 \leqslant C_1 n_2$, as $n_2 \longrightarrow \infty$,

$$\begin{aligned} n_2^2 n_1^2 \exp(\log(n_2 p^2)k - k\log(k) - n_2 p^2) & \\ &\leqslant \exp(2\log(n_2) + 2\log(n_1) + \log(n_2^{2\alpha})n_2^{\alpha'} - n_2^{\alpha'}\log(n_2)\alpha' - n_2^{2\alpha}) \\ &\leqslant \exp((2 + 2C_1)\log(n_2) + (2\alpha - \alpha')\log(n_2)n_2^{\alpha'} - n_2^{2\alpha}) \\ &\longrightarrow 0 \end{aligned}$$

□

6.11 Technical Lemmas

We separate the lemmas into three sections: lemmas for both the continuous and discrete rating settings, lemmas for the continuous rating setting, and lemmas for the discrete rating setting.

6.11.1 Lemmas Common to the Continuous Rating Setting and the Discrete Rating Setting

Lemma 45 establishes that for a user $u \in [n_2]$ and distinct items $i, j \in [n_1]$, with high probability there are many other users that have rated items i and j and many items in common with user u. It is similar to Lemma 1 from Lee et al. (2016).

Lemma 45. *Fix $u \in [n_2]$, $i \neq j \in [n_1]$, and let $\lambda > 0$ and $2 \leqslant \beta \leqslant \frac{n_1 p^2}{2}$. Let $W_u^{i,j}(\beta) = \{v \in [n_2] : |N(u,v)| \geqslant \beta, (i,v), (j,v) \in \Omega\}$. Then,*

$$\begin{aligned} &\Pr_{\boldsymbol{\Omega}}(|W_u^{i,j}(\beta)| \notin [(1-\lambda)(n_2-1)p^2, (1-\lambda)(n_2-1)p^2]) \\ &\leqslant 2\exp(-\frac{\lambda^2(n_2-1)p^2}{3}) + (n_2-1)\exp(-\frac{n_1 p^2}{8}). \end{aligned}$$

Proof. Define the following binary variables for all $v \in [n_2]\backslash\{u\}$. $E_v = 1$ if $|N(u,v)| \geqslant \beta$ and 0 otherwise, $F_v = 1$ if $(i,v) \in \mathbf{\Omega}$ and 0 otherwise, and $G_v = 1$ if $(j,v) \in \mathbf{\Omega}$ and 0 otherwise. Observe that $|W_u^{i,j}(\beta)| = \sum_{v\neq u} E_v F_v G_v$. Fix $0 \leqslant a < b \leqslant n_2 - 1$. Observe that if $\sum_{v\neq u} F_v G_v \in [a,b]$ and $\sum_{v\neq u} E_v = n_2 - 1$, then $|W_u^{i,j}(\beta)| \in [a,b]$. Thus, the contrapositive implies that for any $0 \leqslant a < b \leqslant n_2 - 1$,

$$\begin{aligned}\Pr_{\mathbf{\Omega}}(|W_u^{i,j}(\beta)| \notin [a,b]) &\leqslant \Pr_{\mathbf{\Omega}}(\sum_{v\neq u} F_v G_v \notin [a,b] \cup \sum_{v\neq u} E_v < n_2 - 1) \\ &\leqslant \Pr_{\mathbf{\Omega}}(\sum_{v\neq u} F_v G_v \notin [a,b]) + \Pr_{\mathbf{\Omega}}(\sum_{v\neq u} E_v < n_2 - 1).\end{aligned}$$

$\sum_{v\neq u} F_v G_v$ is a binomial random variable with parameters $n_2 - 1$ and p^2. Letting $a = (1-\lambda)(n_2-1)p^2$ and $b = (1+\lambda)(n_2-1)p^2$, Chernoff's multiplicative bound (Proposition 17) yields that

$$\Pr_{\mathbf{\Omega}}(\sum_{v\neq u} F_v G_v \notin [(1-\lambda)(n_2-1)p^2, (1+\lambda)(n_2-1)p^2]) \leqslant 2\exp(-\frac{\lambda^2(n_2-1)p^2}{3}).$$

Since $N(u,v)$ is binomial with parameters n_1 and p^2, by Chernoff's multiplicative bound (Proposition 17),

$$\begin{aligned}\Pr_{\mathbf{\Omega}}(E_v = 0) &= \Pr_{\mathbf{\Omega}}(N(u,v) \leqslant \beta) \\ &\leqslant \Pr_{\mathbf{\Omega}}(N(u,v) \leqslant \frac{n_1 p^2}{2}) \\ &\leqslant \exp(-\frac{n_1 p^2}{8}).\end{aligned}$$

Then, by the union bound,

$$\begin{aligned}\Pr_{\mathbf{\Omega}}(\sum_{v\neq u} E_v < n_2 - 1) &= \Pr_{\mathbf{\Omega}}(\exists v \in [n_2]\backslash\{u\} : E_v = 0) \\ &\leqslant (n_2 - 1)\exp(-\frac{n_1 p^2}{8}).\end{aligned}$$

□

To convert the pairwise comparisons to a ranking, we use the Copeland ranking procedure (Algorithm 35 in the main document). Lemma 46 establishes that if the output of the Pairwise-Rank algorithm is such that for all $i,j \in [n_1]$ and $u \in [n_2]$, $D^{\epsilon}_{u,i,j}$ does not occur, then applying the Copeland ranking procedure to A (as defined in Multi-Rank) yields a $\widehat{\sigma}$ such that $\mathrm{dis}_{2\epsilon}(\widehat{\sigma}, H) = 0$.

Lemma 46. *Let $\epsilon > 0$, $u \in [n_2]$, A as defined in Multi-Rank (Algorithm 33), and $\widehat{\sigma}_u =$ Copeland$(A_{u,:,:})$. If for all $i \neq j \in [n_1]$ $f(x_i, y_u) > f(x_j, y_u) + \epsilon$ implies that $A_{u,i,j} = 1$, then for all $i \neq j \in [n_1]$ $h_u(x_i, y_u) > h_u(x_j, y_u)$ and $f(x_i, y_u) > f(x_j, y_u) + 2\epsilon$ implies that $\widehat{\sigma}_u(i) > \widehat{\sigma}_u(j)$.*

Proof. Let $i \neq j \in [n_1]$ such that $h_u(x_i, y_u) > h_u(x_j, y_u)$ and $f(x_i, y_u) > f(x_j, y_u) + 2\epsilon$. Let $l \in [n_1]$ such that $l \neq i$ and $l \neq j$. We claim that if $A_{u,i,l} = 0$, then $A_{u,j,l} = 0$. If $A_{u,i,l} = 0$, then by the hypothesis $f(x_i, y_u) \leqslant f(x_l, y_u) + \epsilon$. Then,

$$f(x_j, y_u) + 2\epsilon < f(x_i, y_u) \leqslant f(x_l, y_u) + \epsilon$$

so that $f(x_j, y_u) + \epsilon < f(x_l, y_u)$. Then, by the hypothesis, $A_{u,j,l} = 0$, establishing the claim.

The contrapositive of the claim is that if $A_{u,j,l} = 1$, then $A_{u,i,l} = 1$. Then,

$$I_j = \sum_{l=1, l\neq j}^{n_1} A_{u,j,l} = \sum_{l=1, l\notin\{j,i\}}^{n_1} A_{u,j,l} \leqslant \sum_{l=1, l\notin\{j,i\}}^{n_1} A_{u,i,l} = I_i - 1 < I_i$$

so that $\widehat{\sigma}_u(i) > \widehat{\sigma}_u(j)$. □

Recall the definition of our problem-specific constants: $\tau(\epsilon) = \inf_{y_0 \in \mathcal{Y}} \Pr_{\boldsymbol{y}_u}(d_{\mathcal{Y}}(y_0, \boldsymbol{y}_u) \leqslant \epsilon)$, $\kappa(\epsilon) = \inf_{y_0 \in \mathcal{Y}} \Pr_{\boldsymbol{y}_u}(d_{\mathcal{Y}}(y_0, \boldsymbol{y}_u) > \epsilon)$, and $\gamma(\epsilon) = \inf_{z \in [-N,N]} \mathcal{P}_{\{\boldsymbol{a}_{u,l}\}_{l\in[L-1]}}(\exists l \in [L-1] : d_{\mathbb{R}}(z, \boldsymbol{a}_{u,l}) \leqslant \epsilon)$. Lemma 47 establishes that under our assumptions, for all $\epsilon > 0$, $\tau(\epsilon) > 0$, $\kappa(\epsilon) < 1$, and $\gamma(\epsilon) > 0$.

Lemma 47. *If there exists $\epsilon > 0$ such that $\tau(\epsilon) = 0$, or $\kappa(\epsilon) = 1$, then there exists a point $z \in \mathcal{Y}$ such that $\mathcal{P}_{\mathcal{Y}}(B_\epsilon(z)) = 0$. Similarly, if there exists $\epsilon > 0$ such that $\gamma(\epsilon) = 0$, then there exists $z \in [-N, N]$ such that $\mathcal{P}_l(B_\epsilon(z)) = 0$ for all $l \in [L-1]$.*

Proof. Let $\epsilon > 0$ and suppose $\tau(\epsilon) = 0$. Then, there exists a sequence of points $z_1, z_2, \ldots \in \mathcal{Y}$ such that for every n, $\mathcal{P}_{\mathcal{Y}}(B_\epsilon(z_n)) \leqslant \frac{1}{n}$. Since $\mathcal{Y}$ is compact by assumption, there exists a convergent subsequence $z_{i_1}, z_{i_2}, \ldots$ to z.

We claim that for all $z' \in \mathcal{Y}$, there exists a sufficiently large N such that $z' \in B_\epsilon(z_{i_N})$ if and only if $z' \in B_\epsilon(z)$. Fix $z' \in B_\epsilon(z)$. Since $B_\epsilon(z)$ is open, there exists $\delta > 0$ such that $d_{\mathcal{Y}}(z, z') < \delta < \epsilon$. Let N large enough such that $d_{\mathcal{Y}}(z, z_{i_N}) \leqslant \epsilon - \delta$. Then, by the triangle inequality,

$$d(z', z_{i_N}) \leqslant d(z', z) + d(z_{i_N}, z) \leqslant \delta + \epsilon - \delta = \epsilon$$

so that $z' \in B_\epsilon(z_{i_N})$. A similar argument shows the other direction of the claim. Since a probability space has finite measure, by the dominated convergence theorem,

$$\mathcal{P}_{\mathcal{Y}}(B_\epsilon(z)) = \lim_{n \longrightarrow \infty} \mathcal{P}_{\mathcal{Y}}(B_\epsilon(z_{i_n})) \leqslant \lim_{n \longrightarrow \infty} \frac{1}{i_n} = 0.$$

Next, suppose $\kappa(\epsilon) = 1$. Then, there exists a sequence of points $z_1, z_2, \ldots \in \mathcal{Y}$ such that for every n, $\mathcal{P}_{\mathcal{Y}}(B_\epsilon(z_n)^c) \geqslant 1 - \frac{1}{n}$. Then, for every n, $\mathcal{P}_{\mathcal{Y}}(B_\epsilon(z_n)) \leqslant \frac{1}{n}$ A similar argument from the $\tau(\cdot)$ case using the dominated convergence theorem shows that $\mathcal{P}_{\mathcal{Y}}(B_\epsilon(z)) = 0$.

Since $[-N, N]$ is compact and γ has a similar definition to τ, the result for $\gamma(\cdot)$ follows by an argument similar to the one used for the $\tau(\cdot)$ case. □

6.11.2 Lemmas for Continuous Rating Setting

Lemma 48 uses the notion of r-discerning to relate the distance between points in $\mathcal{Y}$ and to a lower bound on $\rho(y_u, y_v)$.

Lemma 48. *Let r be a positive nondecreasing function. If $y_u \in \mathcal{Y}$ is r-discerning, then for any $\epsilon > 0$, if $y_v \in B_\epsilon(y_u)$, then $\rho(y_u, y_v) > 1 - r(\epsilon)$.*

Proof. Suppose that $d(y_u, y_v) \leqslant \epsilon$. Suppose that $\boldsymbol{x}_i = x_i$ and $\boldsymbol{x}_j = x_j$ such that $|f(x_i, y_u) - f(x_j, y_u)| > 2\epsilon$ and without loss of generality suppose that $h_u(x_i, y_u) \geqslant h_u(x_j, y_u)$. Then, since f is Lipschitz,

$$f(x_i, y_v) \geqslant f(x_i, y_u) - \epsilon > f(x_j, y_u) + \epsilon \geqslant f(x_j, y_v).$$

Hence, $h_v(x_i, y_v) \geqslant h_v(x_j, y_v)$. Thus,

$$\rho(y_u, y_v) \geqslant \Pr_{\boldsymbol{x}_i, \boldsymbol{x}_j}(|f(\boldsymbol{x}_i, y_u) - f(\boldsymbol{x}_j, y_u)| > 2\epsilon) > 1 - r(\epsilon),$$

where the last inequality follows from the hypothesis that y_u is r-discerning. Thus, we conclude the result. □

Lemma 49 establishes that if $S \subset [n_2] \backslash \{u\}$ is a large enough set, then with high probability there is at least one element $\boldsymbol{y}_v$ in S that tends to agree with $\boldsymbol{y}_u$.

Lemma 49. *Let r be a positive non-decreasing function and suppose that almost every $y \in \mathcal{Y}$ is r-discerning. Let $S \subset [n_2] \backslash \{u\}$. Then, $\forall \epsilon > 0$,*

$$\Pr_{\boldsymbol{y}_v, \boldsymbol{y}_u}(\max_{v \in [S]} \rho(\boldsymbol{y}_v, \boldsymbol{y}_u) \leqslant 1 - r(\epsilon)) \leqslant [1 - \tau(\epsilon)]^{|S|}.$$

Proof. Fix $\boldsymbol{y}_u = y_u \in \mathcal{Y}$ that is r-discerning. By Lemma 48, if $\boldsymbol{y}_v = y_v$ is such that $d(y_u, y_v) \leqslant \epsilon$, then $\rho(y_u, y_v) > 1 - r(\epsilon)$. Hence,

$$\Pr_{\boldsymbol{y}_v}(d(y_u, \boldsymbol{y}_v)) \leqslant \epsilon) \leqslant \Pr_{\boldsymbol{y}_v}(\rho(y_u, \boldsymbol{y}_v) > 1 - r(\epsilon)).$$

Then,

$$\Pr_{\boldsymbol{y}_v}(\rho(y_u, \boldsymbol{y}_v) \leqslant 1 - r(\epsilon)) \leqslant \Pr_{\boldsymbol{y}_v}(d(y_u, \boldsymbol{y}_v)) > \epsilon) = 1 - \Pr_{\boldsymbol{y}_v}(d(y_u, \boldsymbol{y}_v)) \leqslant \epsilon) \leqslant 1 - \tau(\epsilon).$$

The RHS does not depend on y_u, and $\boldsymbol{y}_v, \boldsymbol{y}_u$ are independent and almost every $y \in \mathcal{Y}$ is r-discerning, so we can take the expectation with respect to $\boldsymbol{y}_u$ to obtain

$$\Pr_{\boldsymbol{y}_v, \boldsymbol{y}_u}(\rho(\boldsymbol{y}_v, \boldsymbol{y}_u) \leqslant 1 - r(\epsilon)) \leqslant 1 - \tau(\epsilon). \tag{6.13}$$

Finally,

$$\begin{aligned}\Pr_{\{\boldsymbol{y}_v\}_{v \in S}, \boldsymbol{y}_u}(\max_{v \in [S]} \rho(\boldsymbol{y}_v, \boldsymbol{y}_u) \leqslant 1 - r(\epsilon)) &= \Pr_{\boldsymbol{y}_v, \boldsymbol{y}_u}(\rho(\boldsymbol{y}_v, \boldsymbol{y}_u) \leqslant 1 - r(\epsilon))^{|S|} \\ &\leqslant [1 - \tau(\epsilon)]^{|S|},\end{aligned}$$

where the first equality follows from the independence of $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{n_2}$ and the inequality follows from line (6.13). □

Lemma 50 establishes that $R_{\boldsymbol{u},\boldsymbol{v}}$ concentrates around $\rho(\boldsymbol{y}_u, \boldsymbol{y}_v)$.

Lemma 50. *Let $u \neq v \in [n_2]$, $i \neq j \in [n_1]$, $\eta > 0$, $\beta \geqslant 2$, and $W_u^{i,j}(\beta)$ be defined as in Lemma 45. Then,*

$$\Pr(|R_{\boldsymbol{u},\boldsymbol{v}} - \rho(\boldsymbol{y}_u, \boldsymbol{y}_v)| > \frac{\eta}{4} | v \in W_u^{i,j}(\beta)) \leqslant 2 \exp(-\frac{\eta^2}{4} \left\lfloor \frac{\beta}{2} \right\rfloor).$$

Proof. Fix $\boldsymbol{y}_u = y_u$ and $\boldsymbol{y}_v = y_v$. Recall that if $I(u, v) \neq \varnothing$, then

$$R_{u,v} = \frac{1}{|I(u,v)|} \sum_{(s,t) \in I(u,v)} \mathbf{1}\{(h_u(\boldsymbol{x}_s, y_u) - h_u(\boldsymbol{x}_t, y_u))(h_v(\boldsymbol{x}_s, y_v) - h_v(\boldsymbol{x}_t, y_v)) \geqslant 0\}.$$

Since $I(u,v)$ consists of pairs of indices that do not overlap, conditioned on $\boldsymbol{y}_u = y_u, \boldsymbol{y}_v = y_v$, and any nonempty $I(u,v)$, $\{\mathbf{1}\{(h_u(\boldsymbol{x}_s, y_u) - h_u(\boldsymbol{x}_t, y_u))(h_v(\boldsymbol{x}_s, y_v) - h_v(\boldsymbol{x}_t, y_v)) \geqslant 0\} : (s,t) \in I(u,v)\}$ is a set of independent random variables. Further, each has mean $\rho(y_u, y_v)$. Thus, by Chernoff's bound (Proposition 16),

$$\Pr(|R_{u,v} - \rho(y_u, y_v)| > \frac{\eta}{4} | \boldsymbol{y}_u = y_u, \boldsymbol{y}_v = y_v, I(u,v)) \leqslant \exp(-\frac{\eta^2}{2}|I(u,v)|)$$

When $v \in v \in W_u^{i,j}(\beta)$, $|I(u,v)| \geqslant \left\lfloor \frac{\beta}{2} \right\rfloor$. Since the above bound holds for all y_u, y_v, it follows that

$$\Pr(|R_{\boldsymbol{u},\boldsymbol{v}} - \rho(\boldsymbol{y}_u, \boldsymbol{y}_v)| > \frac{\eta}{4} | v \in W_u^{i,j}(\beta)) \leqslant 2\exp(-\frac{\eta^2}{4}\left\lfloor \frac{\beta}{2} \right\rfloor).$$

□

Lemma 51 establishes that conditional on A, B, C (defined in the proof of Theorem 37), the event $D^{\epsilon}_{u,i,j}$ does not occur with probability 1.

Lemma 51. *Under the setting described in Theorem 37, let $u \in [n_2]$ and $i \neq j \in [n_1]$. Then, $\Pr(D^{\epsilon}_{u,i,j} | A, B, C) = 0$.*

Proof. Define the events

$$E_1 = \{f(\boldsymbol{x}_i, \boldsymbol{y}_u) + \epsilon < f(\boldsymbol{x}_j, \boldsymbol{y}_u)\}$$
$$E_2 = \{f(\boldsymbol{x}_i, \boldsymbol{y}_u) > f(\boldsymbol{x}_j, \boldsymbol{y}_u) + \epsilon\}$$

By the union bound and law of total probability,

$$\begin{aligned}
\Pr(D^{\epsilon}_{u,i,j} | A, B, C) &\leqslant \Pr(\mathrm{PR}(u,i,j,\beta,k) = 1 \cap E_1 | A, B, C) \\
&\quad + \Pr(\mathrm{PR}(u,i,j,\beta,k) = 0 \cap E_2 | A, B, C) \\
&\leqslant \Pr(\mathrm{PR}(u,i,j,\beta,k) = 1 | A, B, C, E_1) \\
&\quad + \Pr(\mathrm{PR}(u,i,j,\beta,k) = 0 | A, B, C, E_2).
\end{aligned}$$

The argument for bounding each of these is similar and, thus, we bound the term $\Pr(\mathrm{PR}(u,i,j,\beta,k) = 1 | A, B, C, E_1)$.

Fix $\{\boldsymbol{y}_v = y_v\}_{v \in [n_2]}$ r-discerning and $(\frac{\epsilon}{2}, \delta)$-discriminative, $\{\boldsymbol{x}_s = x_s\}_{s \in [n_1]}$, and $\boldsymbol{\Omega} = \Omega$ such that the event $A \cap B \cap C \cap E_1$ occurs. We claim that Pairwise-Rank puts $V = \{v\}$ (see

Algorithm 34 for definition of V) such that $y_v \in B_{\frac{\epsilon}{2}}(y_u)$. On the event B, there is $v \in W_u^{i,j}(\beta)$ with $\rho(y_u, y_v) \geqslant 1 - \frac{\delta}{2}$. Since y_u is $(\frac{\epsilon}{2}, \delta)$-discriminative, it follows that $y_v \in B_{\frac{\epsilon}{2}}(y_u)$. Suppose that $w \in W_u^{i,j}(\beta)$ such that $y_w \in B_{\frac{\epsilon}{2}}(y_u)^c$. Since y_u is $(\frac{\epsilon}{2}, \delta)$-discriminative, $\rho(y_w, y_u) < 1 - \delta$. Then,

$$\begin{aligned} R_{w,u} &\leqslant \rho(y_w, y_u) + \frac{\delta}{4} && (6.14) \\ &< 1 - \frac{3}{4}\delta \\ &\leqslant \rho(y_u, y_v) - \frac{\delta}{4} \\ &\leqslant R_{u,v} && (6.15) \end{aligned}$$

where lines (6.14) and (6.15) follow by event C and $v, w \in W_u^{i,j}(\beta)$. Thus, the claim follows. Conditional on E_1, we have that $f(x_i, y_u) + \epsilon < f(x_j, y_u)$. Then, using the Lipschitzness of f,

$$f(x_i, y_v) \leqslant f(x_i, y_u) + \frac{\epsilon}{2} < f(x_j, y_u) - \frac{\epsilon}{2} \leqslant f(x_j, y_v).$$

Since g_v is strictly increasing by hypothesis, $h_v(x_i, y_v) < h_v(x_j, y_v)$. Thus, Pairwise-Rank with $k = 1$ outputs 0. Consequently,

$$\Pr(\mathrm{PR}(u, i, j, \beta, k) = 1 | A, B, C, E_1, \{\boldsymbol{y}_v = y_v\}_{v \in [n_2]} \{\boldsymbol{x}_s = x_s\}_{s \in [n_1]}, \boldsymbol{\Omega} = \Omega) = 0$$

Since almost every $y \in \mathcal{Y}$ is r-discerning and $(\frac{\epsilon}{2}, \delta)$-discriminative, taking the expectation wrt $\{\boldsymbol{y}_v\}_{v \in [n_2]}$, $\{\boldsymbol{x}_s\}_{s \in [n_1]}$, $\boldsymbol{\Omega}$ on the set $A \cap B \cap C \cap E_1$ of the last equality gives the result. $\square$

6.11.3 Lemmas for Discrete Rating Setting

Lemma 52 is the analogoue of Lemma 48 for the discrete case. The proof is very similar.

Lemma 52. *Let r be a positive non-decreasing function. If $y_u \in \mathcal{Y}$ is r-discerning, then for any $\epsilon > 0$, if $y_v \in B_\epsilon(y_u)$, then $\rho'(y_u, y_v) > 1 - r(\epsilon)$.*

Proof. Suppose y_v is such that $d(y_u, y_v) \leqslant \epsilon$. We claim that under this assumption

$$\rho'(y_u, y_v) \geqslant \Pr_{\boldsymbol{x}_i, \boldsymbol{x}_j}(|f(\boldsymbol{x}_i, y_u) - f(\boldsymbol{x}_j, y_u)| > 2\epsilon). \tag{6.16}$$

Fix $\boldsymbol{g}_u = g_u$ and $\boldsymbol{g}_v = g_v$, and $\boldsymbol{x}_i = x_i$ and $\boldsymbol{x}_j = x_j$ such that $|f(x_i, y_u) - f(x_j, y_u)| > 2\epsilon$. Without loss of generality, suppose that $h_u(x_i, y_u) \geqslant h_u(x_j, y_u)$. Then, since f is Lipschitz,

$$f(x_i, y_v) \geqslant f(x_i, y_u) - \epsilon > f(x_j, y_u) + \epsilon \geqslant f(x_j, y_v).$$

Hence, $h_v(x_i, y_v) \geqslant h_v(x_j, y_v)$, establishing that

$$\begin{aligned}
&\rho'(y_u, y_v | \boldsymbol{g}_u = g_u, \boldsymbol{g}_v = g_v) \\
&\quad = \Pr_{\boldsymbol{x}_i, \boldsymbol{x}_j}([g_u(f(\boldsymbol{x}_i, y_u)) - g_u(f(\boldsymbol{x}_j, y_u))][g_v(f(\boldsymbol{x}_i, y_v)) - g_u(f(\boldsymbol{x}_j, y_v))] \geqslant 0) \\
&\quad \geqslant \Pr_{\boldsymbol{x}_i, \boldsymbol{x}_j}(|f(\boldsymbol{x}_i, y_u) - f(\boldsymbol{x}_j, y_u)| > 2\epsilon). \qquad (6.17)
\end{aligned}$$

Since $\{\boldsymbol{g}_u, \boldsymbol{g}_v, \boldsymbol{x}_i, \boldsymbol{x}_j\}$ are independent, taking the expectation with respect to $\boldsymbol{g}_u$ and $\boldsymbol{g}_v$ in line (6.17) establishes line (6.16). Thus,

$$\rho'(y_u, y_v) \geqslant \Pr_{\boldsymbol{x}_i, \boldsymbol{x}_j}(|f(\boldsymbol{x}_i, y_u) - f(\boldsymbol{x}_j, y_u)| > 2\epsilon) > 1 - r(\epsilon),$$

where the last inequality follows from the hypothesis that y_u is r-discerning. □

Lemma 53 is the analogoue of Lemma 49 for the discrete case.

Lemma 53. *Let $\epsilon, \delta > 0$. Let r be a positive nondecreasing function such that $r(\epsilon) \geqslant \delta$ and $r(\eta) < \delta$ for some $\eta > 0$. Suppose that almost every $y \in \mathcal{Y}$ is (ϵ, δ)-discriminative and r-discerning. Let $R_2 \geqslant R_1 \geqslant 0$ be constants. Then, for any $S \subset [n_2]$ depending on $\boldsymbol{\Omega}$ and $k \leqslant R_1$,*

$$\begin{aligned}
&\Pr_{\boldsymbol{y}_v, \boldsymbol{y}_u}(\max\nolimits^{(k)}_{v \in [S]} \rho'(\boldsymbol{y}_v, \boldsymbol{y}_u) \leqslant 1 - r(\eta) \,|R_1 \leqslant |S| \leqslant R_2) \\
&\quad \leqslant \exp((1 - \kappa(\epsilon) + \tau(\eta) + \log(R_2))k - k\log(k) - \tau(\eta)R_1) | R_1 \leqslant |S| \leqslant R_2).
\end{aligned}$$

Proof. Let $C_\eta = \Pr_{\boldsymbol{y}_v, \boldsymbol{y}_u}(\rho'(\boldsymbol{y}_v, \boldsymbol{y}_u) \leqslant 1 - r(\eta))$.

Claim: $C_\eta \leqslant 1 - \tau(\eta)$.

Fix $\boldsymbol{y}_u = y_u \in \mathcal{Y}$ r-discerning. By Lemma 52, if $\boldsymbol{y}_v = y_v$ is such that $d(y_u, y_v) \leqslant \epsilon$, then $\rho'(y_u, y_v) > 1 - r(\epsilon)$. Hence,

$$\Pr_{\boldsymbol{y}_v}(d(y_u, \boldsymbol{y}_v)) \leqslant \epsilon) \leqslant \Pr_{\boldsymbol{y}_v}(\rho'(y_u, \boldsymbol{y}_v) > 1 - r(\epsilon)).$$

Then,

$$\Pr_{\boldsymbol{y}_v}(\rho'(y_u, \boldsymbol{y}_v) \leqslant 1 - r(\epsilon)) \leqslant \Pr_{\boldsymbol{y}_v}(d(y_u, \boldsymbol{y}_v)) > \epsilon) = 1 - \Pr_{\boldsymbol{y}_v}(d(y_u, \boldsymbol{y}_v)) \leqslant \epsilon) \leqslant 1 - \tau(\epsilon),$$

where the last inequality follows by the definition of $\tau(\cdot)$. The RHS does not depend on y_u, and $\boldsymbol{y}_v, \boldsymbol{y}_u$ are independent, so we can take the expectation with respect to $\boldsymbol{y}_u$ to establish the claim.

Claim: $1 - C_\eta \leqslant 1 - \kappa(\epsilon)$.

Since almost every $y \in \mathcal{Y}$ is (ϵ, δ)-discriminative and $r(\eta) < \delta$, $\mathcal{Y}$ is almost-everywhere $(\epsilon, r(\eta))$-discriminative. Fix $\boldsymbol{y}_u = y_u$ such that y_u is $(\epsilon, r(\eta))$-discriminative. Then, $\forall y_v \in \mathcal{Y}$, $\rho'(y_u, y_v) > 1 - r(\eta)$ implies that $d_{\mathcal{Y}}(y_u, y_v) \leqslant \epsilon$. Thus,

$$\begin{aligned}
\Pr_{\boldsymbol{y}_v}(\rho'(y_u, \boldsymbol{y}_v) > 1 - r(\eta)) &\leqslant \Pr_{\boldsymbol{y}_v}(d_{\mathcal{Y}}(y_u, \boldsymbol{y}_v \leqslant \epsilon) \\
&= 1 - \Pr_{\boldsymbol{y}_v}(d_{\mathcal{Y}}(y_u, \boldsymbol{y}_v) > \epsilon) \\
&\leqslant 1 - \kappa(\epsilon).
\end{aligned}$$

Since the RHS does not depend on y_u, and $\boldsymbol{y}_u$ and $\boldsymbol{y}_v$ are independent, we can take the expectation with respect to $\boldsymbol{y}_u$ to establish the claim.

Main Probability Bound: Fix $\mathbf{\Omega} = \Omega$ such that $R_1 \leqslant |S| \leqslant R_2$.

$$\begin{aligned}
\Pr_{\boldsymbol{y}_v, \boldsymbol{y}_u}(\max\nolimits_{v\in[S]}^{(k)} \rho'(\boldsymbol{y}_v, \boldsymbol{y}_u) \leqslant & 1 - r(\eta) | \mathbf{\Omega} = \Omega) \\
&= \sum_{l=0}^{k-1} \binom{|S|}{l} C_\eta^{|S|-l} (1 - C_\eta)^l \\
&\leqslant k \max_{l\in\{0,\dots,k-1\}} \binom{|S|}{l} C_\eta^{|S|-l} (1 - C_\eta)^l \\
&\leqslant k \max_{l\in[k-1]\cup\{0\}} \binom{|S|}{l} (1 - \tau(\eta))^{|S|-l} (1 - \kappa(\epsilon))^l \\
&\leqslant k \max_{l\in\{0,\dots,k-1\}} (\frac{|S|e}{l})^l (1 - \tau(\eta))^{|S|-l} (1 - \kappa(\epsilon))^l \qquad (6.18) \\
&\leqslant k \max_{l\in\{0,\dots,k-1\}} \exp(l + l\log(\frac{|S|}{l}) - \tau(\eta)[|S| - l] - \kappa(\epsilon) l) \qquad (6.19) \\
&= k \max_{l\in\{0,\dots,k-1\}} \exp([1 - \kappa(\epsilon) + \tau(\eta)] l + l\log(\frac{|S|}{l}) - \tau(\eta)|S|)) \\
&\leqslant k \exp([1 - \kappa(\epsilon) + \tau(\eta)] k + k\log(\frac{|S|}{k}) - \tau(\eta)|S|)) \qquad (6.20) \\
&= \exp([1 - \kappa(\epsilon) + \tau(\eta) + \log(|S|)] k - k\log(k) - \tau(\eta)|S|)) \\
&\leqslant \exp([1 - \kappa(\epsilon) + \tau(\eta) + \log(R_1)] k - k\log(k) - \tau(\eta) R_2))
\end{aligned}$$

where line (6.18) follows from the the inequality $\binom{n}{k} \leqslant (\frac{ne}{k})^k$, line (6.19) follows from the inequality $(1-x) \leqslant \exp(-x)$, and line (6.20) follows since $|S| \geqslant k$ and $1-\kappa(\epsilon) > 0$ by Lemma 47. Finally, we can take the expectation with respect to $\mathbf{\Omega} = \Omega$ over the set $R_1 \leqslant |S| \leqslant R_2$ to conclude the result. □

Lemma 54 is the analogoue of Lemma 50 for the discrete case.

Lemma 54. *Consider the discrete ratings setting. Let $u \neq v \in [n_2]$, $i \neq j \in [n_1]$, $\eta > 0$, $\beta \geqslant 2$, and $W_u^{i,j}(\beta)$ be defined as in Lemma 45. Then,*

$$\Pr(|R_{\boldsymbol{u},\boldsymbol{v}} - \rho'(\boldsymbol{y}_u, \boldsymbol{y}_v)| > \frac{\eta}{4} | v \in W_u^{i,j}(\beta)) \leqslant 2\exp(-\frac{\eta^2}{4} \left\lfloor \frac{\beta}{2} \right\rfloor).$$

Proof. Fix $\boldsymbol{y}_u = y_u$, $\boldsymbol{y}_v = y_v$, and $\boldsymbol{g}_u = g_u, \boldsymbol{g}_v = g_v$. Recall that if $I(u,v) \neq \varnothing$, then

$$R_{u,v} = \frac{1}{|I(u,v)|} \sum_{(s,t)\in I(u,v)} \mathbf{1}\{[h_u(\boldsymbol{x}_s, y_u) - h_u(\boldsymbol{x}_t, y_u)][h_v(\boldsymbol{x}_s, y_v) - h_v(\boldsymbol{x}_t, y_v)] \geqslant 0\}.$$

Since $I(u,v)$ consists of pairs of indices that do not overlap, conditioned on $\boldsymbol{y}_v = y_v$, $\boldsymbol{y}_u = y_u$, $\boldsymbol{g}_u = g_u$, $\boldsymbol{g}_v = g_v$ and any nonempty $I(u,v)$,

$$\{\mathbf{1}\{(g_u(f(\boldsymbol{x}_s, y_u)) - g_u(f(\boldsymbol{x}_t, y_u)))(g_v(f(\boldsymbol{x}_s, y_v)) - g_v(f(\boldsymbol{x}_t, y_v))) \geqslant 0\} : (s,t) \in I(u,v)\}$$

is a set of independent random variables. Further, each has mean $\rho'(y_u, y_v | \boldsymbol{g}_u = g_u, \boldsymbol{g}_v = g_v)$. Thus, by Chernoff's bound (Proposition 16),

$$\begin{aligned}\Pr(|R_{u,v} - \rho'(y_u, y_v|\boldsymbol{g}_u = g_u, \boldsymbol{g}_v = g_v)| > \frac{\eta}{4} | \boldsymbol{y}_u = y_u, \boldsymbol{y}_v = y_v, \boldsymbol{g}_u = g_u, \boldsymbol{g}_v = g_v, I(u,v)) \\ \leqslant \exp(-\frac{\eta^2}{2}|I(u,v)|)\end{aligned}$$

When $v \in v \in W_u^{i,j}(\beta)$, $|I(u,v)| \geqslant \left\lfloor \frac{\beta}{2} \right\rfloor$. Since the above bound holds for all $\boldsymbol{y}_u, \boldsymbol{y}_v, \boldsymbol{g}_u \boldsymbol{g}_v$, it follows that

$$\Pr(|R_{\boldsymbol{u},\boldsymbol{v}} - \rho'(\boldsymbol{y}_u, \boldsymbol{y}_v)| > \frac{\eta}{4} | v \in W_u^{i,j}(\beta)) \leqslant 2\exp(-\frac{\eta^2}{4}\left\lfloor \frac{\beta}{2} \right\rfloor).$$

□

Lemma 55. *Let $\epsilon, \delta > 0$, $\frac{1}{2} > \alpha > \alpha' > 0$, and r be a positive nondecreasing function such that $r(\frac{\epsilon}{4}) \geqslant \delta$ and $r(\eta) < \frac{\delta}{2}$ for some $\eta > 0$. Suppose that almost every $y \in \mathcal{Y}$ is r-discerning and $(\frac{\epsilon}{4}, \delta)$-discriminative. Fix $u \in [n_2]$, $i \neq j \in [n_1]$, and $k \leqslant \frac{(n_2-1)p^2}{2}$. As in the proof of Theorem 38, define*

$$\begin{aligned}A &= \{|W_u^{i,j}(\beta)| \in [\frac{(n_2-1)p^2}{2}, \frac{3(n_2-1)p^2}{2}]\}, \\ B &= \{\max\nolimits_{v \in W_u^{i,j}(\beta)}^{(k)} \rho'(\boldsymbol{y}_u, \boldsymbol{y}_v) \geqslant 1 - \frac{\delta}{2}\}, \\ E &= \{|f(\boldsymbol{x}_i, \boldsymbol{y}_u) - f(\boldsymbol{x}_j, \boldsymbol{y}_u)| > \epsilon\} \\ M &= \{\exists v \in W_u^{i,j}(\beta) \text{ s.t. } \rho'(\boldsymbol{y}_u, \boldsymbol{y}_v) \geqslant 1 - \frac{\delta}{2} \text{ and } \exists l \in [L-1] \text{ s.t. } \boldsymbol{a}_{v,l} \in (f(\boldsymbol{x}_j, \boldsymbol{y}_v), f(\boldsymbol{x}_i, \boldsymbol{y}_v))\}.\end{aligned}$$

Then,

$$\Pr(M^c | A, B, E) \leqslant \exp(-\gamma(\frac{\epsilon}{4})k).$$

Proof. Fix $\{\boldsymbol{y}_v = y_v\}_{v \in [n_2]}$ r-discerning and $(\frac{\epsilon}{4}, \delta)$-discriminative, $\boldsymbol{\Omega} = \Omega$, and $\{\boldsymbol{x}_s = x_s\}_{s \in [n_1]}$ such that $A \cap B \cap E$ holds. Let $R = \{v \in [n_2] \backslash \{u\} : v \in W_u^{i,j}(\beta) \text{ and } \rho'(y_u, y_v) \geqslant 1 - \frac{\delta}{2}\}$.

Events A and B imply that $|R| \geqslant k$. Since y_u is $(\frac{\epsilon}{4}, \delta)$-discriminative and for all $v \in R$, $\rho'(y_u, y_v) \geqslant 1 - \frac{\delta}{2}$, it follows that for all $v \in R$, $y_v \in B_{\frac{\epsilon}{4}}(y_u)$.

By E, $|f(x_i, y_u) - f(x_j, y_u)| > \epsilon$. Suppose that $f(x_i, y_u) > f(x_j, y_u) + \epsilon$ (the other case is similar). Then, by Lipschitzness of f, for all $v \in R$

$$f(x_j, y_v) \leqslant f(x_j, y_u) + \frac{\epsilon}{4} < f(x_i, y_u) - \frac{3}{4}\epsilon \leqslant f(x_i, y_v) - \frac{\epsilon}{2}.$$

Thus, for all $v \in R$, $(f(x_j, y_v), f(x_i, y_v))$ is an open interval of length at least $\frac{\epsilon}{2}$. Fix $v' \in [n_2] \backslash \{u\}$. Since R is a finite set, the following is well-defined:

$$I \coloneqq \operatorname*{arg\,min}_{J \in \{(f(x_j, y_v), f(x_i, y_v)) : v \in R\}} \Pr_{\{\boldsymbol{a}_{v',l}\}_{l \in [L-1]}}(\exists l \in [L-1] \text{ s.t. } a_{v',l} \in J). \tag{6.21}$$

Then,

$$\begin{aligned}
&\Pr_{\{\boldsymbol{a}_{v,l}\}}(\forall v \in R, \forall l \in [L-1],\, \boldsymbol{a}_{v,l} \notin (f(x_j, y_v), f(x_i, y_v)) | \{\boldsymbol{y}_v = y_v\}_{v \in [n_2]}, \boldsymbol{\Omega} = \Omega, \{\boldsymbol{x}_s = x_s\}_{s \in [n_2]}) \\
&\quad = \Pr_{\{\boldsymbol{a}_{v,l}\}}(\forall v \in R, \forall l \in [L-1],\, \boldsymbol{a}_{v,l} \notin (f(x_j, y_v), f(x_i, y_v))) && (6.22) \\
&\quad \leqslant \Pr_{\{\boldsymbol{a}_{v,l}\}}(\forall v \in R, \forall l \in [L-1],\, \boldsymbol{a}_{v,l} \notin I) && (6.23) \\
&\quad = \Pr_{\{\boldsymbol{a}_{v',l}\}_{l \in [L-1]}}(\forall l \in [L-1],\, \boldsymbol{a}_{v',l} \notin I)^k && (6.24) \\
&\quad = [1 - \Pr_{\{\boldsymbol{a}_{v',l}\}_{l \in [L-1]}}(\exists l \in [L-1] \text{ s.t. } \boldsymbol{a}_{v',l} \in I)]^k \\
&\quad \leqslant (1 - \gamma(\frac{\epsilon}{4}))^k && (6.25) \\
&\quad \leqslant \exp(-\gamma(\frac{\epsilon}{4})k). && (6.26)
\end{aligned}$$

Line (6.22) follows from the independence of $\{\boldsymbol{y}_v\}_{v \in [n_2]}$, $\boldsymbol{\Omega}$, and $\{\boldsymbol{x}_s\}_{s \in [n_1]}$ from $\{\boldsymbol{a}_{v,l}\}_{v \in [n_2], l \in [L-1]}$. Line (6.23) follows from the definition of I in line (6.21) and because the monotonic functions $\{\boldsymbol{g}_v\}_{v \in [n_2]}$ are identically distributed. Line (6.24) follows since $\{\boldsymbol{g}_v\}_{v \in R}$ are i.i.d., line (6.25) follows from the definition of γ, and line (6.26) follows from the inequality $1 - x \leqslant \exp(-x)$. Note that since $\mathcal{P}_{\mathcal{G}}$ is diverse by hypothesis, by Lemma 47, $\gamma(\frac{\epsilon}{4}) > 0$.

Since $\{\boldsymbol{y}_v\}_{v \in [n_2]}$, $\boldsymbol{\Omega} \cup \{\boldsymbol{x}_s\}_{s \in [n_1]}$, and $\{\boldsymbol{a}_{v,l}\}_{v \in [n_2], l \in [L-1]}$ are independent and almost every $y \in \mathcal{Y}$ is r-discerning and $(\frac{\epsilon}{4}, \delta)$-discriminative, taking the expectation of line (6.26) with respect to $\{\boldsymbol{y}_v\}_{v \in [n_2]}$, $\boldsymbol{\Omega}$, and $\{\boldsymbol{x}_s\}_{s \in [n_1]}$ over $A \cap B \cap E$ finishes the proof.

□

Lemma 56 gives a bound on the probability of $D^{\epsilon}_{u,i,j}$ conditional on $A \cap B \cap C \cap E \cap M$ (defined in the proof of Theorem 38).

Lemma 56. *Under the setting described in Theorem 38, let $u \in [n_2]$ and $i \neq j \in [n_1]$. Then,*

$$\Pr(D^{\epsilon}_{u,i,j} | A, B, C, E, M) = 0.$$

Proof. Define the sets

$$E_1 = \{f(\boldsymbol{x}_i, \boldsymbol{y}_u) + \epsilon < f(\boldsymbol{x}_j, \boldsymbol{y}_u)\}$$
$$E_2 = \{f(\boldsymbol{x}_i, \boldsymbol{y}_u) > f(\boldsymbol{x}_j, \boldsymbol{y}_u) + \epsilon\}.$$

Then, by the union bound and the law of total probability,

$$\begin{aligned}
\Pr(D^{\epsilon}_{u,i,j} | A, B, C, E, M) &\leqslant \Pr(\mathrm{PR}(u, i, j, \beta, k) = 1 \cap E_1 | A, B, C, E, M) \\
&\quad + \Pr(\mathrm{PR}(u, i, j, \beta, k) = 0 \cap E_2 | A, B, C, E, M) \\
&\leqslant \Pr(\mathrm{PR}(u, i, j, \beta, k) = 1 | A, B, C, E_1, M) \\
&\quad + \Pr(\mathrm{PR}(u, i, j, \beta, k) = 0 | A, B, C, E_2, M).
\end{aligned}$$

The argument for bounding each of these terms is similar, so we only bound $\Pr(\mathrm{PR}(u, i, j, \beta, k) = 1 | A, B, C, E_1, M)$.

Fix $\{\boldsymbol{y}_v = y_v\}_{v \in [n_2]}$ r-discerning and $(\frac{\epsilon}{4}, \delta)$-discriminative, $\{\boldsymbol{x}_s = x_s\}_{s \in [n_1]}$, $\boldsymbol{\Omega} = \Omega$, and $\{\boldsymbol{a}_{v,l} = a_{v,l}\}_{v \in [n_2], l \in [L-1]}$ such that $A \cap B \cap C \cap E_1 \cap M$ occurs . We claim that the set V in Pairwise-Rank consists of $v_1, \ldots, v_k \in W^{i,j}_u(\beta)$ such that for all $l \in [k]$, $y_{v_l} \in B_{\frac{\epsilon}{4}}(y_u)$. The event B implies that there are $v_1, \ldots, v_k$ such that for all $l \in [k]$, $\rho'(y_u, y_{v_l}) \geqslant 1 - \frac{\delta}{2}$. Then, since y_u is $(\frac{\epsilon}{4}, \delta)$-discriminative, it follows that $y_{v_1}, \ldots, y_{v_k} \in B_{\frac{\epsilon}{4}}(y_u)$. Suppose that $w \in W^{i,j}_u(\beta)$ such that $y_w \in B_{\frac{\epsilon}{4}}(y_u)^c$. Then, since y_u is $(\frac{\epsilon}{4}, \delta)$-discriminative, it follows that that $\rho'(y_u, y_w) < 1 - \delta$. Then, for all $l \in [k]$,

$$\begin{aligned}
R_{w,u} &\leqslant \rho'(y_w, y_u) + \frac{\delta}{4} && (6.27) \\
&< 1 - \frac{3}{4}\delta \\
&\leqslant \rho'(y_u, y_{v_l}) - \frac{\delta}{4} \\
&\leqslant R_{u,v_l} && (6.28)
\end{aligned}$$

where lines (6.27) and (6.28) follow by event C and $v_l, w \in W_u^{i,j}(\beta)$. Thus, Pairwise-Rank selects $v_1, \ldots, v_k \in W_u^{i,j}(\beta)$ such that for all $l \in [k]$, $y_{v_l} \in B_{\frac{\epsilon}{4}}(y_u)$. Thus, the claim follows.

Event E_1 implies that $f(x_i, y_u) + \epsilon < f(x_j, y_u)$. Fix $l \in [k]$. Then, by the Lipschitzness of f,

$$f(x_i, y_{v_l}) \leqslant f(x_i, y_u) + \frac{\epsilon}{4} < f(x_j, y_u) - \frac{3\epsilon}{4} \leqslant f(x_j, y_{v_l}) - \frac{\epsilon}{2}.$$

Hence, $\forall l \in [k]$, $f(x_i, y_{v_l}) + \frac{\epsilon}{2} < f(x_j, y_{v_l})$ and $h_{v_l}(x_i, y_{v_l}) \leqslant h_{v_l}(x_j, y_{v_l})$. Then, event M implies that there is some $l \in [k]$ such that $h_{v_l}(x_i, y_{v_l}) < h_{v_l}(x_j, y_{v_l})$. Thus, the majority vote outputs the correct result. Thus,

$$\begin{aligned}\Pr(\mathrm{PR}(u, i, j, \beta, k) = 1 | A, B, C, E_1, M, \\ \{\boldsymbol{y}_v = y_v\}_{v\in[n_2]}, \{\boldsymbol{x}_s = x_s\}_{s\in[n_1]}, \\ \boldsymbol{\Omega} = \Omega, \{\boldsymbol{a}_{v,l} = a_{v,l}\}_{v\in[n_2], l\in[L-1]}) = 0. \end{aligned} \tag{6.29}$$

Since line (6.29) holds for all $\{y_v\}_{v\in[n_2]}$ r-discerning and $(\frac{\epsilon}{4}, \delta)$-discriminative, $\{a_{v,l}\}_{v\in[n_2],l\in[L-1]}$, $\{x_s\}_{s\in[n_1]}$, Ω conditioned on the set the set $A \cap B \cap C \cap E_1 \cap M$ and almost every $y \in \mathcal{Y}$ is r-discerning and $(\frac{\epsilon}{4}, \delta)$-discriminative, the result follows. □

6.12 Proofs for Section 6.6

Proof of Theorem 34. By compactness of $\mathcal{Y}$, there exists a finite subcover $\{C_1, \ldots, C_n\}$ of $\mathcal{Y}$ where each open ball C_i has diameter $\frac{\epsilon}{2}$. Since by assumption, for all $r > 0$ and $y \in \mathcal{Y}$, $\mathcal{P}_{\mathcal{Y}}(B_r(y)) > 0$, we have that $\mathcal{P}_{\mathcal{Y}}(C_i) > 0$ for all $i = 1, \ldots, n$. Let Q_{n_2} denote the event that for every $l \in [n]$ and $i, j \in [n_1]$, there exists $u \in [n_2]$ such that $\boldsymbol{y}_u \in C_l$ and we observe $(i, u) \in \Omega$ and $(j, u) \in \boldsymbol{\Omega}$. Since $p > 0$, as $n_2 \longrightarrow \infty$, $\Pr(Q_{n_2}) \longrightarrow 1$.

Let $\{\boldsymbol{x}_i = x_i\}_{i\in[n_1]}$, $\{\boldsymbol{y}_u = y_u\}_{u\in[n_2]}$, and $\boldsymbol{\Omega} = \Omega$ such that Q_{n_2} occurs. Let $\sigma \in \mathcal{S}^{n_1 \times n_2}$ be an $\frac{\epsilon}{2}$-consistent minimizer of $\widehat{\mathrm{dis}}(\cdot, H)$ over the sample. Towards a contradiction, suppose there exists y_u and $i \neq j \in [n_1]$ such that $\sigma(i, u) < \sigma(j, u)$, $h_u(x_i, y_u) > h_u(x_j, y_u)$, and $f(x_i, y_u) > f(x_j, y_v) + \epsilon$. Without loss of generality, suppose that $y_u \in C_1$.

Since Q_{n_2} occurs by assumption, there exists $v \in [n_2]$ such that $y_v \in C_1$ and $(i, v), (j, v) \in \Omega$. Since σ is an $\frac{\epsilon}{2}$-consistent collection of rankings and the diameter of C_1 is $\frac{\epsilon}{2}$, σ gives the

same ranking to y_u and y_v. Then, since $\sigma(i,u) < \sigma(j,u)$, it follows that $\sigma(i,v) < \sigma(j,v)$. By Lipschitzness of f,

$$f(x_i, y_v) \geqslant f(x_i, y_u) - \frac{\epsilon}{2} > f(x_j, y_u) + \frac{\epsilon}{2} \geqslant f(x_j, y_v). \tag{6.30}$$

Since g_v is strictly increasing, line (6.30) implies that $h_v(x_i, y_v) > h_v(x_j, y_v)$. Thus, σ is not a minimizer of $\widehat{\text{dis}}(\cdot, H)$–a contradiction. Thus, $\forall u \in [n_2]$ and $i \neq j \in [n_1]$ if $\sigma(i,u) < \sigma(j,u)$ and $h_u(x_i, y_u) > h_u(x_j, y_u)$, then $f(x_i, y_u) \leqslant f(x_j, y_u) + \epsilon$, implying that $\text{dis}_\epsilon(\sigma, H) = 0$. $\square$

Proof of Theorem 35. Fix $\{\boldsymbol{x}_i = x_i\}_{i\in[n_1]}$. By compactness of $\mathcal{Y}$, there exists a finite subcover $\{C_1, \ldots, C_n\}$ of $\mathcal{Y}$ where each open ball C_i has diameter $\frac{\epsilon}{8}$. For every $l \in [n]$, fix $z_l \in C_l$ and define $P_l = \{(i,j) : f(x_i, z_l) > f(x_j, z_l) + \frac{\epsilon}{2}\}$.

Fix $l \in [n]$ and $(i,j) \in P_l$. Let $Q_{n_2}^{l,i,j}$ denote the event that there exists $\boldsymbol{y}_u \in C_l$ with $(i,u), (j,u) \in \boldsymbol{\Omega}$ and $\boldsymbol{a}_{u,q} \in (f(x_j, \boldsymbol{y}_u), f(x_i, \boldsymbol{y}_u))$ for some $q \in [L-1]$. Further, define

$$Q_{n_2} = \cap_{l\in[n],(i,j)\in P_l} Q_{n_2}^{l,i,j}.$$

Observe that by the Lipschitzness of f, for every $z \in C_l$, if $(i,j) \in P_l$, then $f(x_i, z) > f(x_j, z) + \frac{\epsilon}{4}$. Since n is fixed and finite, $|P_l|$ is fixed and finite, and the probability of observing a rating, p, is fixed, there exists a positive constant $C > 0$ such that $\Pr_{\boldsymbol{y}_u, \boldsymbol{\Omega}}(Q_{n_2}^{l,i,j} \mid \{\boldsymbol{x}_s = x_s\}_{s\in[n_1]}) \geqslant C$. Thus, $\Pr(Q_{n_2}^{l,i,j} \mid \{\boldsymbol{x}_s = x_s\}_{s\in[n_1]}) \longrightarrow 1$ as $n_2 \longrightarrow \infty$. Then, by the union bound,

$$\begin{aligned}\lim_{n_2\longrightarrow\infty} \Pr_{\boldsymbol{y}_u, \boldsymbol{\Omega}}([Q_{n_2}]^c \mid \{\boldsymbol{x}_s = x_s\}_{s\in[n_1]}) &\leqslant \lim_{n_2\longrightarrow\infty} n\binom{n_1}{2} \Pr_{\boldsymbol{y}_u, \boldsymbol{\Omega}}([Q_{n_2}^{l,i,j}]^c \mid \{\boldsymbol{x}_s = x_s\}_{s\in[n_1]}) \\ &= 0.\end{aligned}$$

Since $\mathbb{E}[\mathbf{1}\{Q_{n_2}\}|\{\boldsymbol{x}_i\}_{i\in[n_1]}] \leqslant 1$, by the dominated convergence theorem,

$$\begin{aligned}\lim_{n_2\longrightarrow\infty} \Pr(Q_{n_2}) &= \lim_{n_2\longrightarrow\infty} \mathbb{E}_{\{\boldsymbol{x}_i\}} \mathbb{E}[\mathbf{1}\{Q_{n_2}\}|\{\boldsymbol{x}_i\}_{i\in[n_1]}] \\ &= \mathbb{E}_{\{\boldsymbol{x}_i\}} \lim_{n_2\longrightarrow\infty} \mathbb{E}[\mathbf{1}\{Q_{n_2}\}|\{\boldsymbol{x}_i\}_{i\in[n_1]}] \\ &= 1\end{aligned}$$

Now, condition on $\{\boldsymbol{x}_i = x_i\}_{i\in[n_1]}, \{\boldsymbol{y}_u = y_u\}_{u\in[n_2]}, \boldsymbol{\Omega} = \Omega, \{\boldsymbol{a}_{u,l} = a_{u,l}\}_{u\in[n_2],l\in[L-1]}$ such that Q_{n_2} happens. Let $\sigma \in \mathcal{S}^{n_1\times n_2}$ be an $\frac{\epsilon}{8}$-consistent minimizer of $\widehat{\text{dis}}(\cdot, H)$. Towards a contradiction, suppose there exists y_u and $i \neq j \in [n_1]$ such that $\sigma(i,u) < \sigma(j,u)$, $h_u(x_i, y_u) >$

$h_u(x_j, y_u)$, and $f(x_i, y_u) > f(x_j, y_v) + \epsilon$. Without loss of generality, suppose that $y_u \in C_1$. We have that $(i, j) \in P_1$ since

$$\begin{aligned} f(x_i, z_1) &\geqslant f(x_i, y_u) - \frac{\epsilon}{8} \\ &\geqslant f(x_j, y_u) + \frac{7}{8}\epsilon \\ &\geqslant f(x_j, z_1) + \frac{3}{4}\epsilon. \end{aligned}$$

Therefore, the event Q_{n_2} implies that there exists $y_v \in C_1$ such that $(i, v), (j, v) \in \Omega$ and there exists $a_{v,q} \in (f(x_j, y_v), f(x_i, y_v))$. By the Lipschitzness of f, $f(x_j, y_v) < f(x_i, y_v)$, so that $h(x_j, y_v) < h(x_i, y_v)$. Since σ is $\frac{\epsilon}{8}$-consistent, $\sigma(i, v) < \sigma(j, v)$. But, then σ is not a minimizer of $\widehat{\text{dis}}(\cdot, H)$ over the sample–a contradiction. Thus, $\forall u \in [n_2]$ and $i \neq j \in [n_1]$ if $\sigma(i, u) < \sigma(j, u)$ and $h_u(x_i, y_u) > h_u(x_j, y_u)$, then $f(x_i, y_u) \leqslant f(x_j, y_u) + \epsilon$, implying that $\text{dis}_\epsilon(\sigma, H) = 0$. □

Proof of Theorem 36. Let $\boldsymbol{x}_1 = x_1, \ldots, \boldsymbol{x}_{n_1} = x_{n_1}, \boldsymbol{y}_1 = y_1, \ldots, \boldsymbol{y}_{n_2} = y_{n_2}$. Towards a contradiction, suppose that σ is not an ϵ-consistent collection of rankings over T. Then, there exists $i, j \in [n_1]$ and $u, v \in [n_2]$ such that $(i, j, u), (i, j, v) \in T$ and

$$d_{\mathcal{Y}}(y_u, y_v) \leqslant \epsilon, \tag{6.31}$$

$$\sigma(j, u) < \sigma(i, u), \tag{6.32}$$

$$\sigma(j, v) > \sigma(i, v). \tag{6.33}$$

Further, by definition of T,

$$|f(x_j, y_u) - f(x_i, y_u)| > \epsilon \tag{6.34}$$

$$|f(x_i, y_v) - f(x_j, y_v)| > \epsilon. \tag{6.35}$$

$$h(x_i, y_u) \neq h(x_j, y_u) \tag{6.36}$$

$$h(x_i, y_v) \neq h(x_j, y_v) \tag{6.37}$$

Since $\text{dis}_\epsilon(\sigma, H) = 0$ by hypothesis, and by inequalities (6.32), (6.33), (6.34), (6.35), (6.36), and (6.37) it follows that $h(x_j, y_u) < h(x_i, y_u)$ and $h(x_i, y_v) < h(x_j, y_v)$. Thus, by mono-

tonicity of g_u, g_v,

$$\epsilon + f(x_j, y_u) < f(x_i, y_u),$$
$$\epsilon + f(x_i, y_v) < f(x_j, y_v).$$

Then,

$$\begin{aligned} f(x_i, y_u) - f(x_i, y_v) &= f(x_i, y_u) - f(x_j, y_u) + f(x_j, y_u) - f(x_j, y_v) + f(x_j, y_v) - f(x_i, y_v) \\ &> 2\epsilon + f(x_j, y_u) - f(x_j, y_v). \end{aligned}$$

Then, rearranging the above equation and applying the Lipschitzness of f, we have that

$$2\epsilon < f(x_j, y_v) - f(x_j, y_u) + f(x_i, y_u) - f(x_i, y_v) \leqslant 2d_{\mathcal{Y}}(y_v, y_u),$$

which contradicts inequality (6.31). □

6.13 Proof of Proposition 13 and other Results

In the following proposition, we give a simple illustrative example of a 1-Lipschitz function that is (ϵ, δ)-discriminative and r-discerning.

Proposition 14. *Let $\mathcal{X} = [0,1]$, $\mathcal{Y} = [0,1]$, $\mathcal{P}_{\mathcal{X}}$ be the Lebesgue measure over $\mathcal{X}$, and $\mathcal{P}_{\mathcal{Y}}$ be the Lebesgue measure over $\mathcal{Y}$. Suppose that for all $u \in [n_2]$, g_u is strictly increasing. Consider the function*

$$f(x,y) = \begin{cases} x & : x \in [0,y] \\ y - x & : x \in (y,1] \end{cases}$$

Then, for all $1 > \epsilon > 0$, every $y \in \mathcal{Y}$ is (ϵ, ϵ^2)-discriminative. Further, there exists a positive nondecreasing r such that $\lim_{r \longrightarrow 0} r(z) = 0$ and every $y \in \mathcal{Y}$ is r-discerning.

Proof. Let $\epsilon \in (0,1)$ and suppose that $|y_1 - y_2| = \epsilon$. Without loss of generality, suppose that $y_1 < y_2$. Then, when $x_1 < x_2 \in (y_1, y_1 + \epsilon)$, $f(x_1, y_1) > f(x_2, y_1)$ and $f(x_1, y_2) < f(x_2, y_2)$. Since g_u is strictly increasing, $h_1(x_1, y_1) > h_1(x_2, y_1)$ and $h_2(x_1, y_2) < h_2(x_2, y_2)$. Since $\mathcal{P}_{\mathcal{X}} \times \mathcal{P}_{\mathcal{X}}((y_1, y_1 + \epsilon) \times (y_1, y_1 + \epsilon)) = \epsilon^2$, it follows that $\rho(y_1, y_2) < 1 - \epsilon^2$.

Clearly, there exists a positive nondecreasing r such that $\lim_{r \longrightarrow 0} r(z) = 0$ and every $y \in \mathcal{Y}$ is r-discerning. □

This example can easily be generalized to $f(x,y) = \|x-y\|_2$. The following proposition shows that by adding a dimension, the model $f(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{x}^t\boldsymbol{y}$ with $\boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^d$ is a special case of the model $f(\tilde{\boldsymbol{x}},\tilde{\boldsymbol{y}}) = \|\tilde{\boldsymbol{x}}-\tilde{\boldsymbol{y}}\|_2$ with $\tilde{\boldsymbol{x}},\tilde{\boldsymbol{y}} \in \mathbb{R}^{d+1}$. A similar construction in the other direction exists.

Proposition 15. *Let $\boldsymbol{x}_1,\dots,\boldsymbol{x}_{n_1} \in \mathbb{R}^d$ and $\boldsymbol{y}_1,\dots,\boldsymbol{y}_{n_2} \in \mathbb{R}^d$. There exist $\tilde{\boldsymbol{x}}_1,\dots,\tilde{\boldsymbol{x}}_{n_1} \in \mathbb{R}^{d+1}$ and $\tilde{\boldsymbol{y}}_1,\dots,\tilde{\boldsymbol{y}}_{n_2} \in \mathbb{R}^{d+1}$ such that $\forall u \in [n_2]$ and $\forall i \neq j \in [n_1]$, $\boldsymbol{x}_i^t\boldsymbol{y}_u > \boldsymbol{x}_j^t\boldsymbol{y}_u$ if and only if $\|\tilde{\boldsymbol{x}}_i-\tilde{\boldsymbol{y}}_u\|_2 > \|\tilde{\boldsymbol{x}}_j-\tilde{\boldsymbol{y}}_u\|_2$.*

Proof. Let $B = \max_{i\in[n_1]}\|\boldsymbol{x}_i\|_2$. For all $i \in [n_1]$, there exists $\gamma_i \geqslant 0$ such that $\tilde{\boldsymbol{x}}_i \coloneqq (\boldsymbol{x}_i^t,\gamma_i)^t$ and $\|\tilde{\boldsymbol{x}}_i\|_2 = B$ (by continuity and monotonicity of $\|\cdot\|_2$). For all $u \in [n_2]$, define $\tilde{\boldsymbol{y}}_u = (-\boldsymbol{y}_u^t, 0)^t$.

Fix $u \in [n_2]$ and $i \neq j \in [n_1]$. Then,

$$\begin{aligned}\|\tilde{\boldsymbol{x}}_i-\tilde{\boldsymbol{y}}_u\|_2^2 - \|\tilde{\boldsymbol{x}}_j-\tilde{\boldsymbol{y}}_u\|_2^2 &= \|\tilde{\boldsymbol{x}}_i\|_2^2 + \|\tilde{\boldsymbol{y}}_u\|_2^2 - 2\tilde{\boldsymbol{x}}_i^t\tilde{\boldsymbol{y}}_u - (\|\tilde{\boldsymbol{x}}_j\|_2^2 + \|\tilde{\boldsymbol{y}}_u\|_2^2 - 2\tilde{\boldsymbol{x}}_j^t\tilde{\boldsymbol{y}}_u)\\ &= -2\tilde{\boldsymbol{x}}_i^t\tilde{\boldsymbol{y}}_u + 2\tilde{\boldsymbol{x}}_j^t\tilde{\boldsymbol{y}}_u\\ &= \boldsymbol{x}_i^t\boldsymbol{y}_u - \boldsymbol{x}_j^t\boldsymbol{y}_u.\end{aligned}$$

The result follows. □

Proof of Proposition 13. 1. Consider a fixed $y \in \mathcal{Y}$. Fix $\boldsymbol{x}_2 = x_2 \in \mathcal{X}$. Then,

$$\begin{aligned}\Pr_{\boldsymbol{x}_1}(|\,\|\boldsymbol{x}_1-y\|_2 - \|\boldsymbol{x}_2-y\|_2\,| \leqslant 2\epsilon) &\leqslant \Pr_{\boldsymbol{x}_1}(\boldsymbol{x}_1 \in B_{\|\boldsymbol{x}_2-y\|+2\epsilon}(y)\backslash B_{\|\boldsymbol{x}_2-y\|-2\epsilon}(y))\\ &\leqslant \sup_{z\in[0,2]} \mathcal{P}_{\mathcal{X}}(B_z(y)\backslash B_{z-4\epsilon}(y))\\ &= r(\epsilon)\end{aligned}$$

Taking the expectation with respect to $\boldsymbol{x}_2$ establishes the first part of this result.

Fix $y_u \in \mathcal{Y}$ and $\epsilon > 0$ and set $\delta = 2\mathcal{P}_{\mathcal{X}}(B_{\frac{\epsilon}{2}}(y_u))^2$. Fix $y_v \in B_\epsilon(y_u)^c \cap \mathcal{Y}$. If $\boldsymbol{x}_1 = x_1 \in B_{\frac{\epsilon}{2}}(y_u)$ and $\boldsymbol{x}_2 = x_2 \in B_{\frac{\epsilon}{2}}(y_v)$, then

$$[f(x_1,y_u)-f(x_2,y_u)][f(x_1,y_v)-f(x_2,y_v)] < 0.$$

A similar argument applies to the case $\boldsymbol{x}_1 = x_1 \in B_{\frac{\epsilon}{2}}(y_v)$ and $\boldsymbol{x}_2 = x_2 \in B_{\frac{\epsilon}{2}}(y_u)$. Thus, since by hypothesis, g_u is strictly increasing for all $u \in [n_2]$,

$$\rho(y_u,y_v) < 1-\delta.$$

2. Both results follow immediately.

□

6.14 Useful Bounds

Proposition 16 (Chernoff-Hoeffding's Bound)**.** *Let $X_1, \ldots, X_n$ be independent random variables with $X_i \in [a_i, b_i]$. Let $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Then,*

$$\Pr(|\bar{X} - \mathbb{E}[\bar{X}]| \geqslant t) \leqslant 2\exp(-\frac{2n^2t^2}{\sum_{i=1}^{n}(b_i - a_i)^2}).$$

Proposition 17 (Chernoff's multiplicative bound)**.** *Let $X_1, \ldots, X_n$ be independent random variables with values in $[0, 1]$. Let $X = \sum_{i=1}^{n} X_i$. Then, for any $\epsilon > 0$,*

$$\Pr(X > (1+\epsilon)\mathbb{E}[X]) < \exp(-\frac{\epsilon^2\mathbb{E}[X]}{3}),$$
$$\Pr(X < (1-\epsilon)\mathbb{E}[X]) < \exp(-\frac{\epsilon^2\mathbb{E}[X]}{2}).$$

Chapter 7

Conclusion: Future Work

7.1 Any m-Feasible Arm Identification

An important limitation of the feasible arm identification problem is that it requires the identification of all satisfactory arms, i.e., with means belonging to a given polyhedron. In many applications, it suffices to find a given number of arms that satisfy the criteria. For example, in crowdsourcing, it is often of interest to hire specific number of workers that have a sufficiently high probability of giving the correct answer and respond on average at a suitable pace. As another example, in A/B testing, a common goal is to identify one of a set of options as performing better than a baseline on several metrics.

For future work, I plan to consider the *any-m feasible arm identification* problem: there are K multidimensional arms and a polyhedron P, and the goal is to find m arms with means belonging to the polyhedron. Presumably, one could design much more efficient algorithms than the algorithms that we designed for any m-feasible arm identification.

7.2 Nonparametric Preference Completion from Pairwise Comparisons

While our work on preference completion considers the setting where ratings of items are available, often judgments are made in the form of pairwise comparisons. Thus, it is natural

to consider the problem of nonparametric preferences with pairwise comparisons as the input.

Next, we formalize the problem. Suppose there are n_1 items and n_2 users. Let

$$\mathcal{M} = \{M \in [0,1]^{n_1 \times n_1} : M + M^T = 1\}$$
$$\mathcal{M}_{st} = \{M \in [0,1]^{n_1 \times n_1} : M_{i,j} > \frac{1}{2} \text{ and } M_{j,k} > \frac{1}{2} \Longrightarrow M_{i,k} > \frac{1}{2}\}.$$

For a given $M \in \mathcal{M}$, we interpret $M_{i,j}$ as the probability that item i is preferred to item j. $\mathcal{M}_{st}$ is the set of such matrices that satisfies a transitivity property, namely, that if item i tends to be preferred to item j and item j tends to be preferred to item k then item i tends to be preferred to item k.

Let $\mathcal{Y}$ denote a compact metric space with metric $d(\cdot, \cdot)$. Suppose we have a mapping $P : \mathcal{Y} \longrightarrow \mathcal{M}_{s,t}$ that is 1-Lipschitz, i.e., satisfies the following property:

$$\|P(y) - P(y')\|_F \leqslant d(y, y').$$

Fix $y_i \in \mathcal{Y}$. We interpret $P(y)_{i,j}$ in the following way:

$$P(y)_{i,j} = \Pr(\text{ user i with feature vector } y \text{ prefers item } i \text{ to item } j).$$

Let $\mathcal{P}$ be a probability measure on $\mathcal{Y}$. We assume that $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{n_2} \sim \mathcal{P}_{\mathcal{Y}}$ (where $\boldsymbol{y}_i$s are unobserved) and for every $u \in [n_2]$ and distinct $i, j \in [n_1]$, we observe $X_{u,i,j} \sim \text{Bern}(P(\boldsymbol{y}_u)_{i,j})$ with probability p.

An algorithm outputs $\sigma : [n_1] \times [n_2] \longrightarrow [n_1]$ where for fixed $u \in [n_2]$, $\sigma(\cdot, u)$ is a permutation on $[n_1]$. The performance measure is:

$$\text{dis}(\sigma, \{P(\boldsymbol{y}_u)\}_{u \in [n_2]}) = \sum_{u \in [n_2]} \sum_{i<j} \mathbf{1}\{P(\boldsymbol{y}_u)_{i,j} > \frac{1}{2} \text{ and } (\sigma(u,i) < \sigma(u,j)\}$$
$$+\mathbf{1}\{P(\boldsymbol{y}_u)_{i,j} < \frac{1}{2} \text{ and } (\sigma(u,i) > \sigma(u,j)\}.$$

We conjecture that it is possible to show that a variant of Multi-Rank from our work on nonparametric preference completion is consistent in this setting.

Bibliography

N. Ailon, M. Charikar, and A. Newman. Aggregating inconsistent information: Ranking and clustering. *Journal of the ACM*, 2008.

S. Arora, R. Ge, and A. Moitra. Learning topic models–going beyond svd. *Foundations of Computer Science*, 2012.

S. Arora, R. Ge, Y. Halpern, D. Mimno, A. Moitra, D. Sontag, Y. Wu, and M. Zhu. A practical algorithm for topic modeling with provable guarantees. *International Conference on Machine Learning*, 2013.

J.-Y. Audibert and S. Bubeck. Best arm identification in multi-armed bandits. *Conference on Learning Theory*, 2010.

P. Auer, C.-K. Chiang, R. Ortner, and M. Drugan. Pareto front identification from stochastic bandit feedback. *Artificial Intelligence and Statistics*, pages 939–947, 2016.

S. Axler. *Linear Algebra Done Right*. Springer, 3rd edition, 2015.

M. Aziz, J. Anderton, E. Kaufmann, and J. Aslam. Pure exploration in infinitely-armed bandit models with fixed-confidence. In *ALT 2018-Algorithmic Learning Theory*, 2018.

M.-F. Balcan, S. Hanneke, and J. W. Vaughan. The true sample complexity of active learning. *Machine learning*, 80(2-3):111–139, 2010.

R. Bell and Y. Koren. Lessons from the netflix prize challenge. *ACM SIGKDD Explorations Newsletter*, pages 75–79, 2007.

L. Berkman, B. H. Singer, and K. Manton. Black/white differences in health status and mortality among the Elderly. *Demography*, 26:661–678, 1989.

D. A. Berry, R. W. Chen, A. Zame, D. C. Heath, and L. A. Shepp. Bandit problems with infinitely many arms. *Ann. Statist.*, 25(5):2103–2116, 10 1997. doi: 10.1214/aos/1069362389.

D. Bertsekas. *Convex optimization theory*. Belmont: Athena Scientific, 2009.

G. Blanchard and C. Scott. Decontamination of mutually contaminated models. *International Conference of Artificial Intelligence and Statistics*, 2014.

G. Blanchard, G. Lee, and C. Scott. Semi-supervised novelty detection. *Journal of Machine Learning Research*, 11:2973–3009, 2010.

G. Blanchard, M. Flaska, G. Handy, S. Pozzi, and C. Scott. Classification with asymmetric label noise: Consistency and maximal denoising. *Electronic Journal of Statistics*, 10: 2780–2824, 2016.

D. Blei, A. Ng, and M. Jordan. Latent dirichlet allocation. *Journal of Machine Learning research*, 3:993–1022, 2003.

S. Boyd and L. Vandenberghe. Convex optimization. *Cambridge University press*, 2004.

J. Breese, D. Heckerman, and C. Kadie. Empirical analysis of predictive algorithms for collaborative filtering. *In Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence*, 1998.

S. Bubeck, R. Munos, and G. Stolz. Pure Âť exploration in multi-armed bandits problems. *Algorithmic Learning Theory*, pages 23–37, 2009.

S. Bubeck, R. Munos, and G. Stoltz. Pure Exploration in Finitely Armed and Continuous Armed Bandits. *Theoretical Computer Science 412, 1832-1852*, 412:1832–1852, 2011.

S. Bubeck, T. Wang, and N. Viswanathan. Multiple identifications in multi-armed bandits. *International Conference on Machine Learning*, pages 258–265, 2013.

R. Busa-Fekete, B. Szorenyi, P. Weng, and S. Mannor. Multi-objective bandits: Optimizing the generalized gini index. *ICML*, pages 625–634, 2017.

W. Cao, J. Li, Y. Tao, and Z. Li. On top-k selection in multi-armed bandits and hidden bipartite graphs. In C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, R. Garnett, and R. Garnett, editors, *Advances in Neural Information Processing Systems 28*, pages 1036–1044. Curran Associates, Inc., 2015.

A. Carpentier and M. Valko. Simple regret for infinitely many armed bandits. *CoRR*, abs/1505.04627, 2015.

K. Chandrasekaran and R. Karp. Finding a most biased coin with fewest flips. In *Conference on Learning Theory*, pages 394–407, 2014.

A. R. Chaudhuri and S. Kalyanakrishnan. Pac identification of a bandit arm relative to a reward quantile. In *AAAI*, pages 1777–1783, 2017.

A. R. Chaudhuri and S. Kalyanakrishnan. Pac identification of many good arms in stochastic multi-armed bandits. *arXiv preprint arXiv:1901.08386*, 2019.

L. Chen, A. Gupta, J. Li, M. Qiao, and R. Wang. Nearly optimal sampling algorithms for combinatorial pure exploration. *Proceedings of Machine Learning Research*, 65:1–53, 2017a.

L. Chen, J. Li, and M. Qiao. Nearly instance optimal sample complexity bounds for top-k arm selection. In *Artificial Intelligence and Statistics*, pages 101–110, 2017b.

S. Chen, T. Lin, I. King, M. Lyu, and W. Chen. Combinatorial pure exploration of multi-armed bandits. *Advances in Neural Information Processing Systems*, pages 379–387, 2014a.

S. Chen, T. Lin, I. King, M. R. Lyu, and W. Chen. Combinatorial pure exploration of multi-armed bandits. In *Advances in Neural Information Processing Systems 27: Annual Conference on Neural Information Processing Systems 2014, December 8-13 2014, Montreal, Quebec, Canada*, pages 379–387, 2014b.

J. Cid-Sueiro. Proper losses for learning from partial labels. *Advances in Neural Information Processing Systems*, 2012.

A. H. Copeland. A reasonable social welfare function. *Seminar on Mathematics in Social Sciences, University of Michigan*, 1951.

D. Coppersmith, L. Fleischer, and A. Rudra. Ordering by weighted number of wins gives a good ranking for weighted tournaments. *Proceedings of the 17th Annual ACM-SIAM Symposium on Discrete Algorithms*, 2006.

T. Cour, B. Sapp, and B. Taskar. Learning from partial labels. *Journal of Machine Learning Research*, 12:1501–1536, 2011.

P. Cremonesi, Y. Koren, and R. Turrin. Proceedings of the fourth acm conference on recommender systems. *Performance of recommender algorithms on top-n recommendation tasks*, 2010.

A. S. Das, M. Datar, A. Garg, and S. Rajaram. Google news personalization: scalable online collaborative filtering. *Proceedings of the 16th international conference on World Wide Web*, pages 271–280, 2007.

R. Das, M. Zaheer, and C. Dyer. Gaussian lda for topic models with word embeddings. 2015.

E. De Vito, L. Rosasco, and A. Toigo. Spectral regularization for support estimation. *Advances in neural information processing systems*, 2010.

L. Devroye, L. Györfi, and G. Lugosi. *A Probabilistic Theory of Pattern Recognition*. Springer, 1996.

D. Dheeru and E. K. Taniskidou. UCI machine learning repository, 2017. URL `http://archive.ics.uci.edu/ml`.

W. Ding, M. Rohban, P. Ishwar, and V. Saligrama. Topic discovery through data dependent and random projections. *International Conference on Machine Learning*, 2013.

W. Ding, M. Rohban, P. Ishwar, and V. Saligrama. Efficient distributed topic modeling with provable guarantees. *International Conference on Artificial Intelligence and Statistics*, 2014.

D. Donoho and V. Stodden. When does non-negative matrix factorization give a correct decomposition into parts? *Advances in neural information processing systems*, 2003.

M. Drugan and A. NowÃĭ. Designing multi-objective multi-armed bandits algorithms: A study. *Neural Networks (IJCNN), The 2013 International Joint Conference on*, pages 1–8, 2013.

E. Even-Dar, S. Mannor, and Y. Mansour. Action Elimination and Stopping Conditions for the Multi-Armed Bandit and Reinforcement Learning Problems. *Journal of Machine Learning Research*, 7:1079–1105, 2006.

R. Fisher. The use of multiple measurements in taxonomic problems. *Annual Eugenics*, pages 179–188, 1936.

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

R. S. Ganti, L. Balzano, and R. Willett. Matrix completion under monotonic single index models. *Advances in Neural Information Processing Systems*, pages 1873–1881, 2015.

A. Garivier and E. Kaufmann. Optimal best arm identification with fixed confidence. In *Proceedings of the 29th Conference On Learning Theory*, 2016.

M. Genovese, P. Durez, H. Richards, J. Supronik, E. Dokoupilova, V. Mazurov, and J. Aelion. Efficacy and safety of secukinumab in patients with rheumatoid arthritis: a phase ii, dose-finding, double-blind, randomised, placebo controlled study. *Annals of the rheumatic diseases*, 72:863–869, 2013.

A. Ghosh, H. Kumar, and P. Sastry. Robust loss functions under label noise for deep neural networks. *AAAI*, 2017.

S. Gunasekar, O. Koyejo, and J. Ghosh. Preference completion from partial rankings. *Advances in Neural Information Processing Systems*, 2016.

Y. Hu, Y. Koren, and C. Volinsky. Collaborative filtering for implicit feedback datasets. *Data Mining*, pages 263–272, 2008.

K. Huang, X. Fu, and N. D. Sidiropoulos. Anchor-free correlated topic modeling. *Advances in Neural Information Processing Systems*, 2016.

S. Jain, M. White, M. W. Trosset, and P. Radivojac. Nonparametric semi-supervised learning of class proportions. *arXiv preprint arXiv:1601.01944*, 2016.

K. Jamieson, M. Malloy, R. Nowak, and S. Bubeck. lil'ucb: An optimal exploration algorithm for multi-armed bandits. *Conference on Learning Theory*, pages 424–439, 2014a.

K. Jamieson, M. Malloy, R. Nowak, and S. Bubeck. lilÃćÂĂÂŹucb: An optimal exploration algorithm for multi-armed bandits. In *Conference on Learning Theory*, pages 423–439, 2014b.

K. G. Jamieson and L. Jain. A bandit approach to sequential experimental design with false discovery control. In *Advances in Neural Information Processing Systems*, pages 3660–3670, 2018.

K. G. Jamieson, D. Haas, and B. Recht. The power of adaptivity in identifying statistical alternatives. In *Advances in Neural Information Processing Systems*, pages 775–783, 2016.

R. Jin and Z. Ghahramani. Learning with multiple labels. *Advances in Neural Information Processing Systems*, 2002.

K.-S. Jun and R. Nowak. Anytime exploration for multi-armed bandits using confidence information. *ICML*, pages 974–982, 2016.

S. Kalyanakrishnan, A. Tewari, P. Auer, and P. Stone. PAC subset selection in stochastic multi-armed bandits. In *Proceedings of the 29th International Conference on Machine Learning, ICML 2012, Edinburgh, Scotland, UK, June 26 - July 1, 2012*, 2012a.

S. Kalyanakrishnan, A. Tewari, P. Auer, and P. Stone. PAC subset selection in stochastic multi-armed bandits. *ICML*, pages 655–662, 2012b.

Z. Karnin, T. Koren, and O. Somekh. Almost optimal exploration in multi-armed bandits. In S. Dasgupta and D. Mcallester, editors, *Proceedings of the 30th International Conference on Machine Learning (ICML-13)*, volume 28, pages 1238–1246. JMLR Workshop and Conference Proceedings, May 2013.

J. Katz-Samuels and C. Scott. Feasible arm identification. In J. Dy and A. Krause, editors, *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 2535–2543, StockholmsmÃdssan, Stockholm Sweden, 10–15 Jul 2018. PMLR. URL `http://proceedings.mlr.press/v80/katz-samuels18a.html`.

E. Kaufmann, O. Cappé, and A. Garivier. On the complexity of best-arm identification in multi-armed bandit models. *The Journal of Machine Learning Research*, 17(1):1–42, 2016a.

E. Kaufmann, O. CappÃĭ, and A. Garivier. On the complexity of best-arm identification in multi-armed bandit models. *The Journal of Machine Learning Research*, 17:1–42, 2016b.

R. H. Keshavan, A. Montanai, and S. Oh. Matrix completion from a few entries. *IEEE Transactions on Information Theory*, 2010.

J. Kleinberg and M. Sandler. Convergent algorithms for collaborative filtering. *Proceedings of the 4th ACM conference on Electronic commerce*, 2003.

J. Kleinberg and M. Sandler. Using mixture models for collaborative filtering. *Proceedings of the thirty-sixth annual ACM symposium on Theory of computing*, 2004.

Y. Koren. Factorization meets the neighborhood: a multifaceted collaborative filtering model. *In Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 426–434, 2008.

Y. Koren, R. Bell, and C. Volinsky. Matrix factorization techniques for recommender systems. *IEEE Computer*, 2009.

Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *IEEE*, pages 2278–2324, 1998.

C. E. Lee, Y. Li, D. Shah, and D. Song. Blind regression: Nonparametric regression for latent variable models via collaborative filtering. *Advances in Neural Information Processing Systems*, 2016.

C. Li, H. Wang, Z. Zhang, A. Sun, and Z. Ma. Topic modeling for short texts with auxiliary word embeddings. *ACM SIGIR conference on Research and Development in Information Retrieval*, 2016a.

F. Li and P. Perona. A bayesian hierarchical model for learning natural scene categories. *Computer Vision and Pattern Recognition*, 2005.

L. Li, K. G. Jamieson, G. DeSalvo, A. Rostamizadeh, and A. Talwalkar. Hyperband: A novel bandit-based approach to hyperparameter optimization. *Journal of Machine Learning Research*, 18:185–1, 2017.

S. Li, T. Chua, and C. Miao. Generative topic embedding: a continuous representation of documents. *ACL*, 2016b.

L.-P. Liu and T. G. Dietterich. A conditional multinomial mixture model for superset label learning. *Advances in Neural Information Processing Systems*, 2012.

L.-P. Liu and T. G. Dietterich. Learnability of the superset label learning problem. *International Conference on Machine Learning*, 32, 2014.

N. N. Liu and Q. Yang. Eigenrank: A ranking-oriented approach to collaborative filtering. *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, 2008.

T.-Y. Liu. Learning to rank for information retrieval. *Foundations and Trends in Information Retrieval*, pages 225–331, 2009.

A. Locatelli, M. Gutzeit, and A. Carpentier. An optimal algorithm for the thresholding bandit problem. *Proceedings of The 33rd International Conference on Machine Learning*, pages 1690–1698, 2016a.

A. Locatelli, M. Gutzeit, and A. Carpentier. An optimal algorithm for the thresholding bandit problem. In *International Conference on Machine Learning*, pages 1690–1698, 2016b.

Y. Lu and S. Negahban. Individualized rank aggregation using nuclear norm regularization. *Technical Report, Department of Statistics Yale University*, 2014.

M. Luong, R. Socher, and C. Manning. Better word representations with recursive neural networks for morphology. *Conference on Computational Natural Language Learning*, 2013.

S. Mannor and J. Tistisklis. The sample complexity of exploration in the multi-armed bandit problem. *Journal of Machine Learning Research*, pages 623–648, 2004.

S. Mannor, J. N. Tsitsiklis, K. Bennett, and N. Cesa-bianchi. The sample complexity of exploration in the multi-armed bandit problem. *Journal of Machine Learning Research*, 5:2004, 2004.

A. Menon, B. van Rooyen, C. Ong, and B. Williamson. Learning from corrupted binary labels via class-probability estimation. *International Conference on Machine Learning*, 2015a.

A. Menon, B. van Rooyen, C. S. Ong, and R. Williamson. Learning from corrupted binary labels via class-probability estimation. 2015b.

A. Menon, B. van Rooyen, and N. Natarajan. Learning from binary labels with instance-dependent corruption. *arXiv preprint*, 2016.

E. Metodiev and J. Thaler. On the topic of jets. *arXiv preprint arXiv:1802.00008*, 2018a.

E. Metodiev and J. Thaler. Jet topics: Disentangling quarks and gluons at colliders. *Physical Review Letters*, 120(24):241602, 2018b.

S. Mukherjee, N. K. Purushothama, N. Sudarsanam, and B. Ravindran. Thresholding bandits with augmented ucb. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, pages 2515–2521. AAAI Press, 2017.

N. Natarajan, I. S. Dhillon, P. Ravikumar, and A. Tewari. Learning with noisy labels. *Advances in Neural Information Processing Systems*, 2013.

X. Ning, C. Desrosiers, and G. Karypis. *A comprehensive survey of neighborhood-based recommendation methods.* 2011.

N. Nyugen and R. Caruana. Classification with partial labels. *International Conference on Knowledge Discovery and Data Mining*, 2008.

S. Oh, K. Thekumparampil, and J. Xu. Collaboratively learning preferences from ordinal data. *In Advances in Neural Information Processing Systems*, pages 1909–1917, 2015.

D. Park, J. Neeman, J. Zhang, S. Sanghavi, and I. Dhillon. Preference completion: Large-scale collaborative ranking from pairwise comparisons. *Proceedings of the 32nd International Conference on Machine Learning*, 2015.

G. Patrini, A. Rozza, A. Menon, R. Nock, and L. Qu. Making deep neural networks robust to label noise: a loss correction approach. *CVPR*, 2017.

J.-F. Pessiot, T.-V. Truong, N. Usunier, M.-R. Amini, and P. Gallinari. Learning to rank for collaborative filtering. *ICEIS*, 2007.

J. K. Pritchard, M. Stephens, N. A. Rosenberg, and P. Donnelly. Association mapping in structured populations. *American Journal of Human Genetics*, 67:170–181, 2000.

H. Ramaswamy, C. Scott, and A. Tewari. Mixture proportion estimation via kernel embeddings of distributions. *International Conference on Machine Learning*, 2016.

B. Recht. A simpler approach to matrix completion. *Journal of Machine Learning Research*, 2011.

B. Recht, C. Re, J. Tropp, and V. Bittorf. Factoring non-negative matrices with linear programs. *Advances in Neural Information Processing Systems*, 2012.

S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme. Bpr:bayesian personalized ranking from implicit feedback. *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*, pages 452–461, 2009.

A. Rudi, F. Odone, and E. De Vito. Geometrical and computational aspects of spectral support estimation for novelty detection. *Pattern Recognition Letters*, 2014.

T. Sanderson and C. Scott. Class proportion estimation with application to multiclass anomaly rejection. *International Conference on Artificial Intelligence and Statistics*, 2014.

C. Scott. A rate of convergence for mixture proportion estimation, with application to learning from noisy labels. *International Conference on Artificial Intelligence and Statistics*, 2015.

C. Scott, G. Blanchard, and G. Handy. Classification with asymmetric label noise: Consistency and maximal denoising. *Conference on Learning Theory*, 2013.

M. Simchowitz, K. Jamieson, and B. Recht. The simulator: Understanding adaptive sampling in the moderate-confidence regime. *Conference on Learning Theory*, pages 1794–1834, 2017a.

M. Simchowitz, K. Jamieson, and B. Recht. The simulator: Understanding adaptive sampling in the moderate-confidence regime. In *Conference on Learning Theory*, pages 1794–1834, 2017b.

R. Snow, B. O'Connor, D. Jurafsky, and A. Ng. Cheap and fast—but is it good?: evaluating non-expert annotations for natural language tasks. *Proceedings of the conference on empirical methods in natural language processing*, pages 254–263, 2008.

M. Soare, A. Lazaric, and R. Munos. Best-arm identification in linear bandits. In *Advances in Neural Information Processing Systems*, pages 828–836, 2014.

C. Tekin and E. Turgay. Multi-objective contextual multi-armed bandit problem with a dominant objective. *arXiv preprint arXiv:1708.05655*, 2017.

B. van Rooyen and R. Williamson. Learning in the presence of corruption. *arXiv preprint*, 2015.

B. van Rooyen, A. Menon, and R. Williamson. Learning with symmetric label noise: The importance of being unhinged. *Advances in Neural Information Processing Systems*, 2015.

M. Venanzi, J. Guiver, P. Kohli, and N. Jennings. Time-sensitive bayesian information aggregation for crowdsourcing systems. *Journal of Artificial Intelligence Research*, pages 517–545, 2016.

R. Vershynin. Introduction to the non-asymptotic analysis of random matrices. In Y. Eldar and G. Kutyniok, editors, *Compressed Sensing: Theory and Applications*, pages 210–268. Cambridge University Press, 2012.

R. Vershynin, P. Hsu, C. Ma, J. Nelson, E. Schnoor, D. Stoger, T. Sullivan, and T. Tao. High-dimensional probability: An introduction with applications in data science. 2017.

S. Wang, J. Sun, B. Gao, and J. Ma. Vsrank: A novel framework for ranking-based collaborative filtering. *ACM Transactions on Intelligent Systems and Technology*, 2014.

S. Wang, S. Huang, T.-Y. Liu, J. Ma, Z. Chen, and J. Veijalainen. Ranking-oriented collaborative filtering: A listwise approach. *ACM Transactions on Information Systems*, 2016.

Y. Wang, J. yves Audibert, and R. Munos. Algorithms for infinitely many-armed bandits. In D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou, editors, *Advances in Neural Information Processing Systems 21*, pages 1729–1736. Curran Associates, Inc., 2009.

M. Weimer, A. Karatzoglou, Q. Viet Le, and A. Smola. Maximum margin matrix factorization for collaborative ranking. *Advances in neural information processing systems*, pages 1–8, 2007.

G. Xun, Y. Li, W. X. Zhao, J. Gao, and A. Zhang. A correlated topic model using word embeddings. 2017.

H. Zhao, L. Du, W. Buntine, and M. Zhou. Inter and intra topic structure learning with word embeddings. 2018.

S. Zheng, B. Waggoner, Y. Liu, and Y. Chen. Active information acquisition for linear optimization. *arXiv preprint arXiv:1709.10061*, 2017.

Y. Zhou, D. Wilkinson, R. Schreiber, and R. Pan. Large-scale parallel collaborative filtering for the netflix prize. *In International Conference on Algorithmic Applications in Management*, 2008.