Topics in Modern Machine Learning: Sequential Decision Making, High-Dimensional Statistics and Differential Privacy

by

Gang Qiao

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Statistics)
in The University of Michigan
2026

Doctoral Committee:

Professor Ambuj Tewari, Chair
Professor Ya'acov Ritov
Professor Clayton Scott
Associate Professor Yuekai Sun

Gang Qiao
qiaogang@umich.edu
ORCID iD: 0009-0006-2852-7869

Dedicated to my father Zongyang Qiao and my mother Wei Yuan.

ACKNOWLEDGEMENTS

This dissertation marks the close of a long, fulfilling chapter of my life. I find myself thinking less about any single theorem or project, and more about the people who made it possible for me to keep going. A dissertation may carry one person's name on the front page, but it is never truly the work of one person alone. It would never have been written without the generosity, patience, and strength of many people, and I have been fortunate to receive guidance, patience, kindness, and love from many people along the way.

First and foremost, I would like to express my deepest gratitude to my advisor, Ambuj Tewari. Ambuj has been an extraordinary mentor throughout my PhD. His intellectual breadth, clarity of thought, and genuine enthusiasm for research have shaped the way I think about statistics, machine learning, and what it means to do meaningful academic work. He gave me the freedom to explore, the guidance to find direction when things felt unclear, and the patience to help me grow as a researcher. I am also profoundly grateful for his support beyond research. During some of the most difficult moments of my PhD, when life outside academia became overwhelming, Ambuj's understanding and encouragement meant more to me than I can adequately express. I will always be thankful not only for what he taught me academically, but also for the steadiness and humanity with which he advised me.

I would also like to thank my committee members, my professors, and the faculty in the Department of Statistics for their guidance, feedback, and support through this journey. Their questions and suggestions helped me see my work more clearly and strengthened this dissertation in many ways. I am grateful as well to the department staff, whose help and kindness made the many practical parts of graduate school much easier.

I am thankful to my fellow PhD students, friends, and collaborators, who made these years warmer, and much less solitary. I have learned a great deal from conversations in offices, hallways, seminars, reading groups, and over many informal discussions that wandered far beyond the problem at hand. Research can often be uncertain and slow, but having generous people nearby made the process more joyful and meaningful. I am grateful to everyone who shared ideas with me, worked through difficulties with me, or simply made Ann Arbor feel more like a home.

Finally and most importantly, I would like to thank my family. To my parents, I owe more than I can say. Through every difficult period, they stood behind me without hesitation and

gave me their full support, encouragement, and love. Their strength has carried me through moments when I did not feel strong myself. Knowing that they always believed in me, even when the path ahead was unclear, gave me the courage to continue.

To everyone named and unnamed who shaped this beautiful journey, thank you.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE

LIST OF APPENDICES

APPENDIX

LIST OF ACRONYMS

CHM Convex Hull Membership

MAB Multi-armed Bandit

DP Differential Privacy

DS Dantzig Selector

HC Higher Criticism

ABSTRACT

Modern machine learning is increasingly applied to make reliable decisions from limited, noisy, high-dimensional, or privacy-constrained data. This dissertation studies several mathematical problems that arise from this demand, with an emphasis on sequential decision making, sparse high-dimensional inference, differential privacy, and global testing. The chapters address several distinct statistical settings, which share a common theme: how to design procedures whose algorithmic behavior is guided by the geometry of the underlying statistical problem and whose performance can be certified by sharp non-asymptotic or asymptotic guarantees, and how do you make better decisions or selections when information is noisy, incomplete, or costly to gather?

The first part of the dissertation concerns stochastic convex hull membership, a pure-exploration problem in which one sequentially samples from a finite collection of distributions in order to decide whether a target point belongs to the convex hull of their unknown means. We first give a complete solution in one dimension, deriving the information-theoretic characteristic time and developing Thompson-CHM, an asymptotically optimal sampling algorithm whose allocation matches the lower bound. We then further extend the whole theory to the higher-dimensional Gaussian setting, where Euclidean geometry makes the least favorable alternatives explicit. The resulting formulas reveal a distinction between sparse geometric certificates, as guaranteed by Carathéodory's theorem, and statistically optimal sampling allocations, which must guard against all low-cost alternatives.

The second part starts from solving differentially private high-dimensional sparse regression using iterative hard thresholding, and further develops an alternative differentially private procedures for the Dantzig selector in high-dimensional linear regression explicitly using the geometry of Dantzig selector. We start by proposing a oneshot private sparse-regression method based on a noisy iterative hard-thresholding oracle. The algorithm preserves sparsity by construction and satisfies privacy, parameter-error, and population excess-risk guarantees, with the main error rate matching the known differentially private minimax benchmark up to logarithmic factors. We then introduce a complementary active-set method that privatizes the Dantzig score more directly: it privately identifies violated score coordinates, refits on a restricted support, and prunes to exact sparsity. This second approach is closer to the

defining feasibility constraint of the Dantzig selector and provides an alternative route to private sparse estimation under stronger sparse-design conditions.

The final part studies sparse-signal detection in high-dimensional regression by combining two classical ideas: knockoffs and higher criticism. Knockoffs provide dependence-adaptive negative controls, while higher criticism is designed to detect rare and weak alternatives near the sharp sparse-mixture boundary. We introduce a multi-knockoff higher-criticism statistic based on Lasso entry times in an augmented design containing multiple knockoff copies per feature. In the orthogonal-design regime, the proposed statistic attains the classical higher-criticism detection boundary for sparse alternatives against the global null. This result suggests a new way to use knockoff constructions beyond false discovery rate control: as a mechanism for calibrating global tests in high-dimensional models with structured dependence.

CHAPTER 1

Stochastic Convex Hull Membership Problem in the One-dimensional Case

In this chapter, we study the convex hull membership (CHM) problem in the pure exploration setting where one aims to efficiently and accurately determine if a given point lies in the convex hull of means of a finite set of distributions. We give a complete characterization of the sample complexity of the CHM problem in the one-dimensional case. We present the first asymptotically optimal algorithm called Thompson-CHM, whose modular design consists of a stopping rule and a sampling rule. In addition, we extend the algorithm to settings that generalize several important problems in the multi-armed bandit literature. Furthermore, we discuss the extension of Thompson-CHM to higher dimensions. Finally, we provide numerical experiments to demonstrate the empirical behavior of the algorithm matches our theoretical results for realistic time horizons.

1.1 Introduction

The multi-armed bandit (MAB) problem is a fundamental problem in sequential decision making where an agent is required to make a series of decisions to pull an arm of a K slot machine in order to maximize the total reward. Each of the arms is associated with a fixed but unknown probability distribution [Auer et al., 2002, Lai et al., 1985]. An enormous literature has accumulated over the past decades on the MAB problem, such as clinical trials and drug testing [Bastani and Bayati, 2020, Durand et al., 2018], recommendation system and online advertising [Bouneffouf et al., 2012, 2014, Nguyen, 2021, Tang et al., 2013, Zhou et al., 2017], information retrieval [Bouneffouf et al., 2013, Losada et al., 2017], and finance [Huo and Fu, 2017, Misra et al., 2019, Mueller et al., 2019, Shen et al., 2015]. The MAB problem was first studied theoretically in the seminal work [Robbins, 1952] and followed by a vast line of work in two canonical settings: regret minimization [Agrawal and Goyal, 2013, Auer, 2002, Auer et al., 2002, Chapelle and Li, 2011, Chu et al., 2011, Dudik et al., 2011, Langford and Zhang, 2007, Li et al., 2010, Srinivas et al., 2009, Valko et al., 2013] and pure

exploration [Chen et al., 2017, Garivier and Kaufmann, 2016, Locatelli et al., 2016, Russo, 2016].

In this chapter, we study the convex hull membership (CHM) problem in a pure exploration setting: testing whether a fixed target point lies in the convex hull of the unknown mean vectors of K distributions as efficiently and accurately as possible. Following the multi-armed sampling model, where each distribution corresponds to an arm, and pulling an arm reveals an independent sample from the distribution. Unlike classical reward-maximizing bandit setting, our goal is to resolve a geometric decision problem under limited stochastic feedback. Pure exploration problems are usually studies in one of two settings: fixed-confidence of success or fixed-budget of samples. We work in the former. The usual non-stochastic version of the CHM problem is well studied in Filippozzi et al. [2023] and has attracted significant attention in different scientific areas and proven its crucial applications in image processing [Jayaram and Fleyeh, 2016, Yang and Cohen, 1999], robot motion planning [Lengyel et al., 1990, Streinu, 2000] and pattern recognition [Katzin, 2018, Roy et al., 2008].

The stochastic CHM problem arises in important applications including fairness [Martinez et al., 2020] and multi-task learning [Lin et al., 2019], where we consider the instance to determine whether a given point $\boldsymbol{\theta}^* \in \Theta \subset \mathbb{R}^d$ lies on the Pareto frontier of a collection of m objective functions $(F_1, \cdots, F_m)$. A point $\boldsymbol{\theta}^*$ is said to be Pareto optimal if no other $\boldsymbol{\theta} \in \Theta$ exists such that $F_i(\boldsymbol{\theta}) \leq F_i(\boldsymbol{\theta}^*)$ for all $i \in [m]$, with strict inequality in at least one coordinate. In is well known that, under differentiability, any Pareto optimal point satisfies a first-order condition: there exists a convex combination $(\lambda_1, \cdots, \lambda_m)$ with summation equal to 1 such that $\sum_{i=1}^m \lambda_i \nabla F_i(\boldsymbol{\theta}^*) = \mathbf{0}$, or equivalently, $\mathbf{0}_d \in \mathrm{Conv}(\{\nabla F_i(\boldsymbol{\theta}^*)\}_{i=1}^m)$. In a multi-task learning setup, $\nabla F_i(\boldsymbol{\theta}) = \mathbb{E}_{P_i}[\nabla l_i(\boldsymbol{\theta})]$ where P_i's and l_i's are the underlying distributions and loss functions of the i-th task for $i \in [m]$. Since P_i's are unknown, we utilize the empirical version of ∇F_i which follows distributions with different means and fits in our stochastic CHM setting. Nevertheless, there is almost no literature on the *stochastic* convex hull membership problem where we have to sample in order to estimate the positions of the means. Recently, Niss et al. [2022] provided the first theoretical bounds for the CHM problem. Unfortunately, their results have significant gaps between the upper and lower complexity bounds. To the best of our knowledge, this fundamental primitive of developing the complexity bounds and an (asymptotically) optimal algorithm for the CHM problem remains open in the literature before this work.

To tackle the aforementioned problem, we introduce Thompson-CHM, a Thompson-Sampling-based algorithm that has asymptotic sample complexity matching the information-theoretic lower bound proved in Garivier and Kaufmann [2016]. The sample complexity lower

bound is modeled as a function of the characteristic time [Garivier and Kaufmann, 2016], which can be captured by the value of a zero-sum pure exploration game between two players [Chernoff, 1959, Degenne et al., 2019]. As discussed in Section 1.4, any successful pure exploration player needs to solve this pure exploration game, and therefore, the intuition behind the game is essential to our algorithm design. The design of the strategy to match the lower bound is based on the individual confidence interval for each of the K arms so that any algorithm using this stopping time can ensure an output of a correct decision with high probability (at least $1 - \delta$) no matter what sampling rule the algorithm applies.

We remark that Kaufmann et al. [2018] first proposed an active sequential testing procedure to study the lowest mean of a finite set of distributions, and provide a conditional modification of the popular heuristic Thompson Sampling (named as *Murphy Sampling*) to tackle the limitations of the Lower Confidence Bound algorithm (LCB) and standard Thompson Sampling in different settings. However, *a major challenge in extending Thompson Sampling to our CHM problem is to study the extreme means (largest and lowest mean in one-dimensional setting) simultaneously.* To tackle this challenge, we borrow a two-arm sampling construction proposed in the Best-Arm Identification setting. Russo [2016] pointed out that Thompson Sampling can have a poor asymptotic performance and this defect can be improved by a top-two arm sampling modification to prevent the algorithm from sampling the arm of interest too frequently. This modification automatically controls the measurement effort of each arm and ensures that the long-term asymptotic behavior is closely linked to the optimal allocation of the algorithm. *To the best of our knowledge, conditional Thompson sampling and two-arm sampling have never been combined before.*

Our novel sampling rule is independent of the confidence parameter δ and ensures the sampled proportion of each arm asymptotically matches the estimated-best allocation design derived by the pure exploration game in the one-dimensional setting. Therefore, it automatically adapts exploration for both feasible and infeasible cases in the CHM problem. We provide a theoretical analysis of the asymptotic optimality and extend it to two more important settings: interval CHM problem (identifying if an interval (γ^-, γ^+) intersects with the convex hull of the means of K arms) and the d-dimensional CHM problem when $d \geq 2$. The first extension generalizes the CHM problem and reproduces the state-of-art results for several important MAB problems in the literature, including thresholding bandit [Locatelli et al., 2016] and sequential test for lowest mean [Kaufmann et al., 2018]. Moreover, the stochastic CHM problem in d-dimensional setting has several important applications but its complete solution remains open.

To highlight and summarize our results, our contribution in this chapter is threefold:

- We prove the information-theoretical lower bound on the sample complexity of the

one-dimensional convex hull membership (CHM) problem and reveal an oracle allocation of different arms for algorithm design.

- We introduce a novel Thompson-Sampling-based algorithm that automatically adapts the right exploration and oracle allocation for both feasible and infeasible cases and we rigorously prove the algorithm is asymptotically optimal and its complexity exactly matches the theoretical lower bound.
- Our final contribution is two important extensions of the Thompson-CHM algorithm. First, we extend the algorithm to the one-dimensional interval CHM problem by presenting the sample complexity bounds and the analogous asymptotically optimal algorithm, and discuss how this extension generalizes several fundamental BAI problems in the literature. We also investigate the potential extension to the d-dimensional CHM problem ($d \geq 2$) by showing the sample complexity bound which shares the same behavior as the one-dimensional case, and defer more details including the variant of the Thompson-CHM algorithm to the appendix.

1.2 Related Work

In this section, we briefly discuss some works and applications that motivate our work and are closely related to the convex hull membership problem in the literature.

Thresholding Bandits: One closely related previous work is a popular combinatorial pure exploration bandit problem known as the thresholding bandit problem where the learner's objective is to find the set of arms whose means are above a threshold. It was first introduced in Chen et al. [2014] and has been extensively studied in both fixed-confidence and fixed-budget settings [Chen et al., 2014, Garivier et al., 2017, Kano et al., 2019, Locatelli et al., 2016, Tao et al., 2019]. Compared to the thresholding bandit problem, the convex hull membership problem only requires a boolean decision and needs the existence for both arms above and below the threshold to guarantee feasibility. A naive approach using the thresholding bandit problem to solve the CHM problem is to find the set of arms with means above and below the threshold by applying a thresholding bandit algorithm twice, and use the results to build a conclusion on if the threshold lies in the convex hull of the set of the arm means. Compared to the proposed Thompson-CHM algorithm, this two-step procedure is sub-optimal and expends unnecessary samples to determine the true sets of arms with means above and below the threshold.

Fair Sampling and Minimax Pareto Fairness: A recent series of works on fairness sampling and minimax Pareto fairness [Abernethy et al., 2020, Anahideh et al., 2022, Martinez et al., 2020, Nargesian et al., 2021] share similar frameworks with the fair data sampling

procedure that is related to the CHM problem. As discussed in Niss et al. [2022], the main challenge of fair data sampling is to collect data of desired distribution requirements, therefore it reveals an appropriate representation of majority and minority groups in the data. In Anahideh et al. [2022], authors propose a fair active learning framework to balance the trade-off between model accuracy and fairness, in order to avoid discrimination in machine learning models. In Martinez et al. [2020], group fairness is formulated as a multi-objective optimization problem and proposes conditions for the classifier to be Pareto-efficient and achieve minimax risk, which is closely related to the stochastic CHM setting.

1.3 Problem Setup and Formulation

We define the problem of efficiently and accurately identifying if a given point lies in (or if a given interval/set intersects with, respectively) the convex hull of means of K probability distributions $\nu_1, \cdots, \nu_K$ in dimension d based on their stochastic sequential samples as the d-dimensional convex hull membership (d-*dim* CHM) problem. In this chapter, we start with the one-dimensional setting where the probability distributions are in the canonical one-dimensional exponential family. In the canonical one-dimensional exponential family, the marginal distribution of a value x given an unknown parameter $\theta \in \mathbb{R}$ takes the form

$$P(x|\theta) = h(x) \exp\{\eta(x)\theta - A(\theta)\}$$

where $h(x)$, $\eta(x)$ and $A(\theta)$ are known functions.

Throughout the chapter, we denote by $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_K)$ the vector of unknown *true* means of the distributions $\nu_1, \cdots, \nu_K$, and $\boldsymbol{\lambda}$ will be used as possible alternatives of the mean vector. The Kullback-Leibler divergence is a standard measure of how one probability distribution P differs from another Q with the form $\sum_x P(x) \ln(P(x)/Q(x))$. For the canonical one-dimensional exponential family, it induces a bijection between the natural parameter and the mean parameter, and we define the Kullback-Leibler divergence of two distributions with means μ_1 and μ_2 as a function $d : (\mu_1, \mu_2) \to \mathbb{R}^+$. Let $\text{Conv}(\boldsymbol{\mu}) = \text{Conv}(\mu_1, \cdots, \mu_K)$ denote the convex hull of the mean vector, which is the smallest convex set that contains all the means $\mu_1, \cdots, \mu_K$. At each time $t = 1, 2, \cdots$, a decision maker chooses one arm $A_t \in \{1, \cdots, K\}$ and independently draws a reward from distribution $X_{t,A_t} \sim \nu_{A_t}$. Let $\mathcal{F}_t$ denote the sigma algebra generated by $(A_1, X_{1,A_1}, \cdots, A_t, X_{t,A_t})$. We aim to design a sequential hypothesis testing procedure that consists of a $\mathcal{F}_{t-1}$-measurable *sampling policy* π_t, a *stopping rule* τ with respect to $\mathcal{F}_t$, and a $\mathcal{F}_\tau$-measurable *decision rule* $I_{\boldsymbol{\pi}}(\gamma) \in \{\text{feasible}, \text{infeasible}\}$.

We now state the formal definition of feasible and infeasible cases.

Definition 1.3.1. (feasibility and infeasibility) Given $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_K)$, where $\mu_i \in \mathbb{R}^d$ for $i = 1, \cdots, K$. For any set S, the problem defined above is S-feasible if the set $S \cap \text{Conv}(\boldsymbol{\mu}) \neq \emptyset$, otherwise, the problem is S-infeasible. When the set only contains a single element $S = \{\gamma\}$, the problem is simply called γ-feasible and γ-infeasible, respectively.

In the one-dimensional case, given a threshold $\gamma \in \mathbb{R}$, our objective is to identify whether the unknown mean vector $\boldsymbol{\mu}$ is γ-feasible, which is equivalent to determining if γ lies in the closed interval between the smallest mean reward $I_*(\boldsymbol{\mu}) = \text{argmin}_{1 \leq i \leq K} \mu_i$ and the largest mean reward $I^*(\boldsymbol{\mu}) = \text{argmax}_{1 \leq i \leq K} \mu_i$ based on the sequential observations while minimizing the expected stopping time τ and maximizing the probability to correctly identify the result. For simplicity, we assume the threshold γ (and the extreme points of the set S, respectively) does not equal to $\mu_{I_*(\boldsymbol{\mu})}$ and $\mu_{I^*(\boldsymbol{\mu})}$. This aims to avoid infinite samples to distinguish the maximum and minimum means of the distributions that are too close to the threshold. This assumption can be easily relaxed by introducing a "precision" term ε to identify an ε-optimal design instead [Locatelli et al., 2016, Russo, 2016]. Additionally, we further assume the extreme points (or the vertices) of the convex hull $\text{Conv}(\boldsymbol{\mu})$ are unique.

In the literature, two distinct settings have been extensively studied. In the *fixed-confidence setting*, given a fixed confidence parameter $\delta \in (0, 1)$, the forecaster aims for a strategy that achieves the confidence δ about the quality of the decision rule while minimizing the sample needed, and in the *fixed-budget setting*, the number of exploration rounds is fixed, and the forecaster tries to maximize the probability of making the right decision. We will focus on the *fixed-confidence setting* in this chapter and introduce the δ-correct strategy.

Definition 1.3.2. (δ-correctness) Let $\mathcal{D}$ be a set of distributions on $\mathbb{R}^d$. Given $\delta \in (0, 1)$, we call an identification strategy δ-correct on the problem class $\boldsymbol{\nu} \in \mathcal{D}^K$ if with probability at least $1 - \delta$, the strategy returns the correct underlying case in a finite expected stopping time, i.e., $\mathbb{P}(\mathbb{E}[\tau] \leq \infty) = 1$, and when $\boldsymbol{\mu}$ is feasible, $\mathbb{P}(I_{\boldsymbol{\pi}}(\gamma) = \text{feasible}) \geq 1 - \delta$, otherwise $\mathbb{P}(I_{\boldsymbol{\pi}}(\gamma) = \text{infeasible}) \geq 1 - \delta$, here $I_{\boldsymbol{\pi}}(\gamma)$ is the decision rule of the strategy.

Before continuing, we pause to introduce some further notations here. We let $N_a(t) = \sum_{s=1}^{t} \mathbb{1}\{A_s = a\}$ be the number of selections of arm a up to round t, and $S_a(t) = \sum_{s=1}^{t} X_s \mathbb{1}\{A_s = a\}$ be the sum of the gathered observations from that arm and $\hat{\mu}_a(t) = S_a(t)/N_a(t)$ be their empirical mean.

1.4 A General Lower Bound

In this section, we extend the general information-theoretical sample complexity lower bound proved in Garivier and Kaufmann [2016] to work for the one-dimensional convex hull mem-

bership problem.

We define $\mathrm{Alt}(\boldsymbol{\mu})$ to be the set of bandit models where the identification result is different from that in $\boldsymbol{\mu}$, and $\Delta = \{\boldsymbol{w} = (w_1, \cdots, w_K) \in \mathbb{R}_+^K | w_1 + \cdots + w_K = 1\}$ is a probability simplex of dimension K. The following bound was proved by Garivier and Kaufmann [2016] that $\mathbb{E}_{\boldsymbol{\mu}}[\tau] \geq T^*(\boldsymbol{\mu})\mathrm{kl}(\delta, 1-\delta)$, where $\mathrm{kl}(x, y) = x \ln(\frac{x}{y}) + (1-x)\ln(\frac{1-x}{1-y})$ denotes the Kullback-Leibler divergence in the binary reward case, and

$$T^*(\boldsymbol{\mu})^{-1} = \sup_{\boldsymbol{w} \in \Delta} \inf_{\boldsymbol{\lambda} \in \mathrm{Alt}(\boldsymbol{\mu})} \sum_a w_a d(\mu_a, \lambda_a).$$

Note that $\mathrm{kl}(\delta, 1-\delta) \sim \ln(1/\delta)$ as $\delta \to 0$, the lower bound above directly implies $\liminf_{\delta \to 0} \frac{\mathbb{E}_{\boldsymbol{\mu}}[\tau]}{\ln(1/\delta)} \geq T^*(\boldsymbol{\mu})$. This max-min problem was first discussed in the seminal work by Chernoff [1959], and the value of $T^*(\boldsymbol{\mu})^{-1}$ can be viewed as the value of a zero-sum simultaneous-move pure exploration game between two players. The player SUP aims to choose an optimal proportion of allocations $\boldsymbol{w}$ as a mixed strategy, and the adversary player INF tries to choose the worst-case alternative arm means that is hard to distinguish from the underlying truth to mislead SUP to an incorrect answer.

This general information-theoretic bound was established to analyze the sample complexity of the Best-Arm Identification problem [Garivier and Kaufmann, 2016], and was studied in different settings [Degenne et al., 2019, 2020] along with its popular variant that tackles pure exploration bandit problems with multiple correct answers [Degenne and Koolen, 2019]. To match this general lower bound, the sampling proportion $\mathbb{E}[N_{\tau_\delta}]/\mathbb{E}_{\boldsymbol{\mu}}[\tau_\delta]$ must converge to the minimizer $\boldsymbol{w}^* \in \Delta$ of the pure exploration game as $\delta \to 0$. This intuition inspires works on different sampling rules and their corresponding threshold functions $\beta(t, \delta)$ to ensure correct recommendation with high probability (at least $1-\delta$) [Degenne et al., 2019, Kaufmann and Koolen, 2021], and novel sampling rules to match the lower bound $\boldsymbol{N}(t)/t \to \boldsymbol{w}^*$, where $\boldsymbol{N}(t)$ is the vector of selection counts [Kaufmann et al., 2018]. With these considerations in mind, we can establish the sample complexity bound and the asymptotically optimal algorithm for the CHM problem. Specifically, following Degenne et al. [2020], we say that a δ-correct algorithm is *asymptotically optimal* if for all $\boldsymbol{\mu}$, $\limsup_{\delta \to 0} \frac{\mathbb{E}_{\boldsymbol{\mu}}[\tau]}{\ln(1/\delta)} \leq T^*(\boldsymbol{\mu})$.

Without loss of generality, in the one-dimensional CHM problem, we assume that $\mu_1 < \mu_2 \leq \mu_3 \leq \cdots \leq \mu_{K-1} < \mu_K$. The strict inequalities come from the aforementioned assumption of unique extreme points of $\mathrm{Conv}\{\boldsymbol{\mu}\}$. We have the following lower bound of any δ-correct algorithm. The proof is provided in the appendix.

Theorem 1. *Given a threshold $\gamma \in \mathbb{R}$, the expected sample complexity $\mathbb{E}_{\boldsymbol{\mu}}[\tau]$ of any δ-correct*

1-dimensional CHM strategy satisfies $\liminf_{\delta\to 0} \frac{\mathbb{E}_{\boldsymbol{\mu}}[\tau]}{\ln(1/\delta)} \geq T^*(\boldsymbol{\mu})$, *where*

$$T^*(\boldsymbol{\mu}) = \begin{cases} \frac{1}{d(\mu_1,\gamma)} + \frac{1}{d(\mu_K,\gamma)} & \gamma \in \mathit{Conv}\{\boldsymbol{\mu}\} \\ \sum_{1\leq i\leq K} \frac{1}{d(\mu_i,\gamma)} & \gamma \notin \mathit{Conv}\{\boldsymbol{\mu}\} \end{cases},$$

and

$$w_a^*(\boldsymbol{\mu}) = \begin{cases} \frac{\frac{1}{d(\mu_a,\gamma)}}{\sum_{i\in\{1,K\}} \frac{1}{d(\mu_i,\gamma)}} \mathbb{1}_{\{a\in\{1,K\}\}} & \gamma \in \mathit{Conv}\{\boldsymbol{\mu}\} \\ \frac{\frac{1}{d(\mu_a,\gamma)}}{\sum_{1\leq i\leq K} \frac{1}{d(\mu_i,\gamma)}} & \gamma \notin \mathit{Conv}\{\boldsymbol{\mu}\} \end{cases}.$$

Surprisingly, the characteristic time and oracle weights that match the general information-theoretic sample complexity show completely different behaviors in feasible and infeasible cases. In the feasible case where τ lies in the convex hull $\mathrm{Conv}(\boldsymbol{\mu})$, the algorithm should only sample the arms with minimum and maximum means, while in the infeasible case, the strategy should sample every single arm with specific fraction inversely proportional to the Kullback-Leibler divergence between its mean and the threshold γ. We remark that the previous work on sequentially testing and learning the lowest mean [Kaufmann et al., 2018] demonstrates a similar phenomenon. In essence, this commonality arises from the fact that the one-dimensional CHM problem generalizes the problem of learning the smallest mean (see section 1.6.1 for details).

1.5 Algorithm

In this section, we introduce an asymptotically optimal Thompson-Sampling-based algorithm for the one-dimensional CHM problem for a given threshold $\gamma \in \mathbb{R}$.

1.5.1 Stopping rule

From a learning point of view, the question of stopping at time t is essentially a classical statistical problem: does the past collected information allow us to assess that the threshold γ lies in or outside the convex hull set $\mathrm{Conv}(\boldsymbol{\mu})$ with risk at most δ? Inspired by Kaufmann et al. [2018], a natural design of the stopping rule is to compare separately each arm to the threshold γ and stop when either one arm lies significantly below γ and one arm lies significantly above γ, or all arms lie significantly below γ, or all arms lie significantly above γ.

We denote $d^+(u,v) = d(u,v)\mathbb{1}\{u \leq v\}$ and $d^-(u,v) = d(u,v)\mathbb{1}\{u \geq v\}$. We define the

first stopping time τ_1 when all arms lie significantly above γ:

$$\tau_1 = \inf\{t \in \mathbb{N}^+ | \forall a, N_a(t) d^-(\hat{\mu}_a(t), \gamma) \geq \text{Thresh}(\delta, N_a(t))\}.$$

Similarly, we define the second stopping time τ_2 when all arms lie significantly below γ:

$$\tau_2 = \inf\{t \in \mathbb{N}^+ | \forall a, N_a(t) d^+(\hat{\mu}_a(t), \gamma) \geq \text{Thresh}(\delta, N_a(t))\}.$$

The third stopping time is when one arm is significantly below γ and another arm lies significantly above γ:

$$\begin{aligned}\tau_3 = &\inf\{t \in \mathbb{N}^+ | \exists a_1, a_2, \text{ such that } N_{a_1}(t) d^+(\hat{\mu}_{a_1}(t), \gamma) \geq \text{Thresh}(\delta, N_{a_1}(t)) \\ &\text{and } N_{a_2}(t) d^-(\hat{\mu}_{a_2}(t), \gamma) \geq \text{Thresh}(\delta, N_{a_2}(t))\}.\end{aligned}$$

Here $\text{Thresh}(\delta, N_a(t))$ is a threshold function to be specified later. Our algorithm stops if any of the three cases happen, i.e., it stops at $\tau = \min\{\tau_1, \tau_2, \tau_3\}$ and returns feasibility or infeasibility based on the case detected. The stopping rule and decision rule ensures that, when the threshold function $\text{Thresh}(\delta, N_a(t))$ is carefully designed and the sampling rule guarantees the sampling allocation proportion converges to the solution $\boldsymbol{w}$ of the max-min problem, the algorithm Thompson-CHM is δ-correct.

Lemma 1.5.1. *Let τ_δ be a stopping rule satisfying $\tau_\delta \leq \tau$. τ is a stopping rule whose threshold function $\text{Thresh}(\delta, r)$ is non-decreasing in r and satisfies the following: $\forall r \geq r_0$, $\text{Thresh}(\delta, r) \leq \ln(r/\delta) + o(\ln(1/\delta))$, then for any $\boldsymbol{\mu}$ and an anytime sampling strategy such that $\frac{\boldsymbol{N}_t}{t} \to w^*(\boldsymbol{\mu})$, we have $\limsup_{\delta \to 0} \frac{\tau_\delta}{\ln(1/\delta)} \leq T^*(\mu)$ almost surely.*

1.5.2 Sampling rule

Our contribution is a sampling rule that extends and generalizes a variant of Thompson sampling (called *Murphy Sampling*) introduced in Kaufmann et al. [2018] to the one-dimensional CHM problem that ensures the algorithm allocates the optimal proportion to each arm asymptotically, therefore guarantees the asymptotical optimality by Lemma 1.5.1. The sampling rule can automatically adapt the asymptotic optimality for both feasible cases and infeasible cases.

We denote by $\Pi_t = \mathbb{P}(\cdot | \mathcal{F}_t)$ the posterior distribution of the mean parameters after t rounds. Inspired by Kaufmann et al. [2018] that introduces *Murphy Sampling* after Murphy's Law, as it performs some conditioning to the "worst event" to learn the smallest mean, we introduce Thompson-CHM (Algorithm 1) to tackle the one-dimensional CHM problem.

Algorithm 1 Thompson-CHM

Input: stopping rule τ with threshold function $\text{Thresh}(\delta, t)$, risk δ, threshold γ, Bernoulli distribution parameter β_t.
Output: decision rule $I_\pi(\boldsymbol{\mu}) \in \{\text{feasible}, \text{infeasible}\}$
for $t = 1, \cdots$ **do**
  if stopping rule τ holds **then**
    if $\tau = \tau_3$ **then**
      return $I_\pi(\boldsymbol{\mu}) = \{\text{feasible}\}$
    else
      return $I_\pi(\boldsymbol{\mu}) = \{\text{infeasible}\}$
    end if
  end if
  Sample $\boldsymbol{\theta}_t = (\theta_{t,1}, \cdots, \theta_{t,K}) \sim \Pi_{t-1}(\cdot|\boldsymbol{\mu}$ feasible).
  Sample $B \sim \text{Bernoulli}\,(\beta_t)$
  if $B = 1$ **then**
    Play arm $A_t = \text{argmin}(\boldsymbol{\theta}_t)$
  else
    Play arm $A_t = \text{argmax}(\boldsymbol{\theta}_t)$
  end if
end for

Note that the top-two Thompson Sampling conditions the standard Thompson Sampling on the event $\text{argmax}\,\boldsymbol{\mu} \neq \text{argmax}\,\boldsymbol{\theta}_t$ with pre-specified probability β [Russo, 2016], and the Murphy Sampling conditions on $\min(\boldsymbol{\mu})$ below the threshold [Kaufmann et al., 2018]. In contrast, Thompson-CHM conditions on the "feasibility" of the underlying mean vector $\boldsymbol{\mu}$ and in each round t, the algorithm proceeds to pull an arm in the sample $\boldsymbol{\theta}_t = (\theta_{t,1}, \cdots, \theta_{t,K})$ with largest or smallest mean based on the previous information $\mathcal{F}_{t-1}$. The posterior is computed explicitly using Bayes' rule with tractable priors (e.g. Beta or uniform). To implement the conditioning, we adopt reject sampling: we repeatedly sample from the unconditioned posterior until a sample satisfying feasibility is obtained. The next theorem guarantees that following this sampling procedure, the algorithm Thompson-CHM can ensure the sampling proportion of each arm converges to the optimal allocation $\boldsymbol{w}^*$ asymptotically, regardless of the position of γ with respect to the convex hull $\text{Conv}\{\boldsymbol{\mu}\}$. Therefore, we can conclude that the algorithm Thompson-CHM is asymptotically optimal in sample complexity.

Theorem 2. *If* $\beta_t = \frac{d(\min \boldsymbol{\theta}_t, \gamma)^{-1}}{d(\min \boldsymbol{\theta}_t, \gamma)^{-1} + d(\max \boldsymbol{\theta}_t, \gamma)^{-1}}$, *then the algorithm Thompson-CHM ensures that* $\frac{\boldsymbol{N}_t}{t} \to w^*(\boldsymbol{\mu})$ *almost surely for any* $\boldsymbol{\mu}$, *and is* δ*-correct for the CHM problem.*

We let $\psi_a(t)$ be the posterior probability of sampling arm a at time t, i.e. $\psi_a(t) = \mathbb{P}(A_t = a|\mathcal{F}_{t-1})$, and define $\Psi_a(t)$ and $\bar{\psi}_a(t)$ as the summation and mean of $\psi_a(t)$ over time t.

For the feasible case, the first step of the sampling rule performs the same as Thompson

Sampling, and the probability of drawing the first arm at time t can be written as a weighted sum (with weights β_t and $1-\beta_t$) of the posterior probabilities that the first sample in $\boldsymbol{\theta}_t$ is the maximum and minimum. The asymptotic convergence of sample proportions $N_1(t)/t \to w_1^*(\boldsymbol{\mu})$ can be derived by the combination of facts that the former probability converges to 1 and β_t converges to $w_1^*(\boldsymbol{\mu})$. The proof of $N_K(t)/t \to w_K^*(\boldsymbol{\mu})$ is symmetric.

For the infeasible case when the lowest mean is larger than threshold γ or the largest mean is smaller than γ, the core idea of the proof is based on the following proposition, and the complete proofs of Theorem 1, Lemma 1.5.1, and Theorem 2 are deferred to the appendix.

Proposition 1.5.2. *(Simplified version of Lemma 12 of Russo [2016]) Consider any sampling rule, if for any arm $a \in [K]$ and all $c > 0$, $\sum_t \psi_a(t)\mathbb{1}\{\bar{\psi}_a(t) \geq w_a^* + c\} < \infty$, then $\bar{\psi}(t) \to \boldsymbol{w}^*$.*

The above result gives a sufficient condition in which $\bar{\psi}(t)$ converges to the optimal allocation $\boldsymbol{w}^*$, and implies that for any arm a that meets $\bar{\psi}_a(t) \geq w_a^* + c$, the arm has been over-allocated compared to the optimal proportion w_a^*. Hence the total measurements the arm gets must be bounded in order to reduce towards w_a^* for optimality. The rest of the proof is to establish the condition holds for Thompson-CHM algorithm. We develop the conclusion by showing that, if arm a has been over-allocated compared to w_a^*, then $\Pi_t(\theta_{t,a} < \gamma < \theta_{t,b})$ is exponentially small compared to $\max_{a,b} \Pi_t(\theta_{t,a} < \gamma < \theta_{t,b})$. Based on the known result, for any open set $\tilde{\Theta} \subset \Theta$, the posterior concentrates at rate $\Pi_t(\tilde{\Theta}) \doteq \exp\left(-t \min_{\boldsymbol{\lambda} \in \tilde{\Theta}} \sum_a \bar{\psi}_a(t) d(\mu_a, \lambda_a)\right)$, where $x_t \doteq y_t$ means $\frac{1}{t} \ln \frac{x_t}{y_t} \to 0$. Combined with the properties of $T^*(\boldsymbol{\mu})$ in the pure exploration game and the concentration rate of the posterior, we can show that there exists $\delta' > 0$ such that,

$$\psi_a(t) \sim \frac{\Pi_t(\theta_{t,a} < \gamma < \theta_{t,b})}{\max_{a,b} \Pi_t(\theta_{t,a} < \gamma < \theta_{t,b})} \leq \exp(-t(\delta' + \varepsilon_t)),$$

where ε_t is a sequence converging to 0. This implies for any arm a such that $\bar{\psi}_a(t) \geq w_a^* + c$, $\psi_a(t)$ has an exponential decay rate, and Proposition 1.5.2 immediately yields $\bar{\psi}(t) \to \boldsymbol{w}^*$.

It is worth mentioning that one can tackle the one-dimensional CHM problem by first checking if γ is smaller than the minimum mean and then checking if γ is larger than the maximum mean. Using the results in Kaufmann et al. [2018], this strategy's sample complexity is at most two times the sample complexity stated in Theorem 1. However, this procedure has obvious drawbacks compared to our solution. First, this procedure does not generalize to higher dimensions since minimum and maximum means have no analogs in higher dimensions. Moreover, even in the one-dimensional case, this procedure incurs

sub-optimality in its sample complexity in the infeasible case. By sequentially checking the one-sided setting twice, the arm that is farthest away from γ will be sampled more than the optimal $\boldsymbol{w}^*(\boldsymbol{\mu})$ (and all other arms will be sampled less than $\boldsymbol{w}^*(\boldsymbol{\mu})$, respectively), especially when the arms are not spread out significantly. This demonstrates the sub-optimality of this easy solution as our main results indicate an algorithm matching the theoretical lower bound should follow the optimal allocation $\boldsymbol{w}^*(\boldsymbol{\mu})$. More details are discussed in the appendix.

1.6 Extensions of the Thompson-CHM Algorithm

1.6.1 Interval CHM problem

In this section, we show that our results in Section 1.4 and Section 1.5 are fully generalizable to the interval feasibility setting, where our goal is to determine if the open set (γ^-, γ^+) intersects with the convex hull set of $\boldsymbol{\mu}$. Here, we allow γ^- to be $-\infty$ and γ^+ to be $+\infty$ for better generalization results.

1.6.1.1 Asymptotic Optimality and the Algorithm

We build the first result on the general sample complexity lower bound.

Theorem 3. *Given thresholds* $-\infty \leq \gamma^- \leq \gamma^+ \leq +\infty$, *let* $(\gamma^*, \mu^*) = \operatorname{argmin}_{\gamma\in\{\gamma^-,\gamma^+\},\mu\in\boldsymbol{\mu}} |\gamma - \mu|$. *The expected sample complexity* $\mathbb{E}_{\boldsymbol{\mu}}[\tau]$ *of any* δ*-correct 1-dimensional CHM strategy satisfies* $\liminf_{\delta\to 0} \frac{\mathbb{E}_{\boldsymbol{\mu}}[\tau]}{\ln(1/\delta)} \geq T^*(\boldsymbol{\mu})$, *where*

$$T^*(\boldsymbol{\mu}) = \begin{cases} \frac{1}{d(\mu_1,\gamma^+)} + \frac{1}{d(\mu_K,\gamma^-)} & (\gamma^-,\gamma^+) \cap \mathit{Conv}\{\boldsymbol{\mu}\} \neq \emptyset \\ \sum_{1\leq i\leq K} \frac{1}{d(\mu_i,\gamma^*)} & (\gamma^-,\gamma^+) \cap \mathit{Conv}\{\boldsymbol{\mu}\} = \emptyset \end{cases},$$

and

$$w_a^*(\boldsymbol{\mu}) = \begin{cases} \frac{\frac{1}{d(\mu_1,\gamma^+)}\mathbb{1}\{a=1\} + \frac{1}{d(\mu_K,\gamma^-)}\mathbb{1}\{a=K\}}{\frac{1}{d(\mu_1,\gamma^+)} + \frac{1}{d(\mu_K,\gamma^-)}} & (\gamma^-,\gamma^+) \cap \mathit{Conv}\{\boldsymbol{\mu}\} \neq \emptyset \\ \frac{\frac{1}{d(\mu_a,\gamma^*)}}{\sum_{1\leq i\leq K} \frac{1}{d(\mu_i,\gamma^*)}} & (\gamma^-,\gamma^+) \cap \mathit{Conv}\{\boldsymbol{\mu}\} = \emptyset \end{cases}$$

The stopping rule is similar with minor adjustments. To be more specific, we again define the first stopping time τ_1 when all arms lie significantly above γ^+:

$$\tau_1 = \inf\{t \in \mathbb{N}^+ | \forall a, N_a(t) d^-(\hat{\mu}_a(t), \gamma^+) \geq \mathrm{Thresh}(\delta, N_a(t))\}.$$

Similarly, we define the second stopping time τ_2 when all arms lie significantly below γ^-:

$$\tau_2 = \inf\{t \in \mathbb{N}^+ | \forall a, N_a(t) d^+(\hat{\mu}_a(t), \gamma^-) \geq \text{Thresh}(\delta, N_a(t))\}.$$

To identify the feasible case, the third stopping time is when one arm is significantly below γ^+ and another arm lies significantly above γ^-:

$$\begin{aligned}\tau_3 = &\inf\{t \in \mathbb{N}^+ | \exists a_1, a_2, \text{ such that } N_{a_1}(t) d^+(\hat{\mu}_{a_1}(t), \gamma^+) \geq \text{Thresh}(\delta, N_{a_1}(t)) \\ &\text{and } N_{a_2}(t) d^-(\hat{\mu}_{a_2}(t), \gamma^-) \geq \text{Thresh}(\delta, N_{a_2}(t))\}.\end{aligned}$$

Again, the algorithm stops if any of the three cases happens and $\tau = \min\{\tau_1, \tau_2, \tau_3\}$, and the following lemma ensures that the algorithm Thompson-CHM is δ-correct in the interval feasibility framework.

Lemma 1.6.1. *Let τ_δ be a stopping rule satisfying $\tau_\delta \leq \tau$. τ is a stopping rule whose threshold function $\text{Thresh}(\delta, r)$ is non-decreasing in r and satisfies the following: $\forall r \geq r_0$, $\text{Thresh}(\delta, r) \leq \ln(r/\delta) + o(\ln(1/\delta))$, then for any $\boldsymbol{\mu}$ and an anytime sampling strategy such that $\frac{\boldsymbol{N}_t}{t} \to w^*(\boldsymbol{\mu})$, we have $\limsup_{\delta \to 0} \frac{\tau_\delta}{\ln(1/\delta)} \leq T^*(\mu)$ almost surely.*

The sampling rule in the interval CHM problem remains the same and we have the next theorem.

Theorem 4. *If $\beta_t = \frac{d\left(\min \boldsymbol{\theta}_t, \gamma^+\right)^{-1}}{d(\min \boldsymbol{\theta}_t, \gamma^+)^{-1} + d(\max \boldsymbol{\theta}_t, \gamma^-)^{-1}}$, then the algorithm Thompson-CHM ensures that $\frac{\boldsymbol{N}_t}{t} \to w^*(\boldsymbol{\mu})$ almost surely for any $\boldsymbol{\mu}$, and is δ-correct for the CHM problem.*

1.6.1.2 Connections with Other State-of-art Results

We comment on some important connections of Section 1.6.1 with the previous state-of-art results in the MAB literature. Trivially, when $\gamma^- = \gamma^+$, we immediately derive the same results as the CHM problem, implying a direct generalization to the regular CHM problem. If we set $\gamma^- = -\infty$, the Bernoulli parameter β_t becomes 0, and this reproduces the same Murphy Sampling results from the state-of-art sequential test paper for learning the minima mean [Kaufmann et al., 2018].

On the other hand, by setting $\gamma^+ = +\infty$, the interval CHM problem shares the same setup with the thresholding bandit problem with the threshold γ^-. In the infeasible case when γ^- is larger than the largest mean in $\boldsymbol{\mu}$, testing if there exists an arm with a mean above the threshold is essentially equivalent to finding all arms with means above the threshold since to identify both questions, one needs to traverse all arms to conclude that the means of all arms are actually below the threshold, and our complexity bound exactly matches the state-of-art

optimal bound of thresholding bandit [Locatelli et al., 2016]. When μ is feasible, the CHM problem is strictly easier than the thresholding bandit, and our complexity is strictly smaller than the state-of-art result. Notably, the Thompson-CHM algorithm adapts both feasible and infeasible cases for the thresholding bandit problem without knowing any information on the threshold as a priori.

1.6.2 Convex hull membership problem in higher dimensions

We now investigate and discuss the extensions of the Thompson-CHM to d-dimensional setting where $d \geq 2$. Before proceeding, we define the vertices set (or extreme point set) $\text{Vert}(S)$ of a convex set S to be the union of points that do not fall on any line segment connecting any two unique points in set S. The following theorem states that the lower bound for the CHM problem exhibits a shared behavior in all dimensions: *in the feasible case, the optimal strategy should only sample arms whose means are extreme points, and in the infeasible case, it should sample all arms.*

Theorem 5. *Let* $\text{Vert}(\text{Conv}\{\boldsymbol{\mu}\}) = (\mu_{s_1}, \cdots, \mu_{s_m})$ *be the vertices set of* $\text{Conv}\{\boldsymbol{\mu}\}$*. Given* $\gamma \in \mathbb{R}^d$ *where* $2 \leq d < \infty$*, the expected sample complexity* $\mathbb{E}_{\boldsymbol{\mu}}[\tau]$ *of any* δ*-correct* d*-dimensional CHM strategy satisfies* $\liminf_{\delta \to 0} \frac{\mathbb{E}_{\boldsymbol{\mu}}[\tau]}{\ln(1/\delta)} \geq T^*(\boldsymbol{\mu})$*, where*

$$T^*(\boldsymbol{\mu}) = \begin{cases} f_0(\mu_{s_1}, \cdots, \mu_{s_m}, \gamma) & \gamma \in \text{Conv}\{\boldsymbol{\mu}\} \\ \sum_{1 \leq i \leq K} \frac{1}{d(\mu_i, \gamma)} & \gamma \notin \text{Conv}\{\boldsymbol{\mu}\} \end{cases},$$

and

$$w_a^*(\boldsymbol{\mu}) = \begin{cases} \sum_{i=1}^m f_i(\mu_{s_1}, \cdots, \mu_{s_m}, \gamma) \mathbb{1}_{\{a = s_i\}} & \gamma \in \text{Conv}\{\boldsymbol{\mu}\} \\ \frac{\frac{1}{d(\mu_a, \gamma)}}{\sum_{1 \leq i \leq K} \frac{1}{d(\mu_i, \gamma)}} & \gamma \notin \text{Conv}\{\boldsymbol{\mu}\} \end{cases}.$$

Here $f_0, f_1, \cdots, f_m$ *are non-negative real-value functions, and* $\sum_{i=1}^m f_i(\mu_{s_1}, \cdots, \mu_{s_m}, \gamma) = 1$*.*

Using Theorem 5, we can generalize the Thompson-CHM algorithm to higher dimensions by simply replacing the Bernoulli distribution with a categorical distribution with parameters $\beta_i = f_i(\mu_{s_1}, \cdots, \mu_{s_m}, \gamma)$ and assuming oracle access to the functions f_i for $1 \leq i \leq k$. We discuss more details of the Thompson-CHM algorithm in d-dimensional setting ($d \geq 2$) in the appendix.

1.7 Numerical Results

The chapter's main results are reflected in some numerical experiments in this section. We consider 7 Bernoulli bandits with means $\boldsymbol{\mu} = (0.1, 0.2, \cdots, 0.7)$ and $\delta = e^{-3}$, and we consider Beta(1,1) prior and different γ's to compare the sample complexity and sample weights to the theoretical results in both feasible and infeasible cases. We use the threshold function developed in Kaufmann et al. [2018]: $\text{Thresh}(\delta, r) = \ln(1 + \ln(r)) + T(\ln(1/\delta))$ and $T : \mathbb{R}^+ \to \mathbb{R}^+$ is a function defined by $T(x) = 2h^{-1}\left(1 + \frac{h^{-1}(1+x)+\ln\zeta(2)}{2}\right)$, where $h(u) = u - \ln(u)$ for $u \geq 1$ and $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$. We also adopt the empirical implementation of $T(x)$ from Kaufmann et al. [2018] in our experiments. The property of the threshold function is verified in Kaufmann et al. [2018].

We first pick different values of γ to compare the theoretical sample complexity and the sample complexity of Thompson-CHM in both feasible and infeasible cases. We choose γ to be $(0.15, 0.25, 0.35, 0.45, 0.55, 0.65)$ for the feasible case and $(0.75, 0.8, 0.85, 0.9, 0.95)$ for the infeasible case. Figure 1.1 demonstrates the efficiency of the algorithm Thompson-CHM. In both feasible and infeasible cases, the sample complexity of Thompson-CHM matches the theoretical results proved in Theorem 1 well for realistic time horizons.

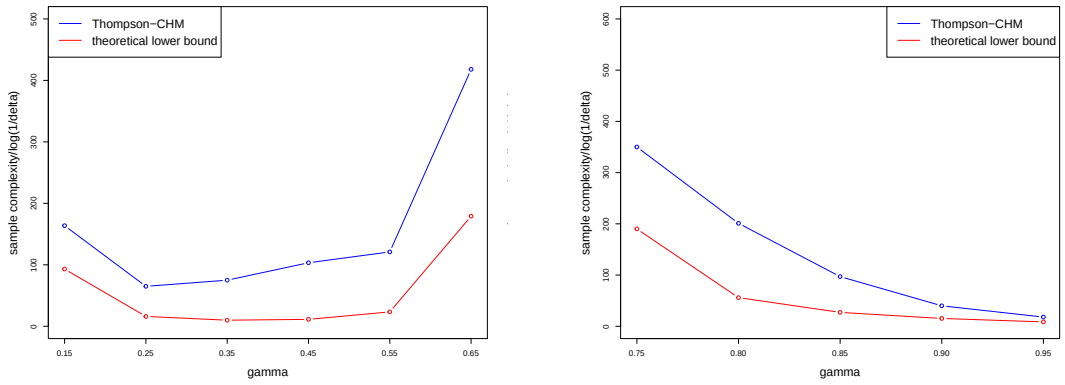

Figure 1.1: Sample complexity for different γ's in feasible cases (left) and infeasible cases (right).

Figure 1.2 provides insights into the asymptotic convergence performance of the sampling proportions $N_a(\tau)/\tau$ in Thompson-CHM in both feasible and infeasible cases. In the feasible case when $\gamma = 0.25$, we note that Thompson-CHM spent the most fraction of time sampling the side arms (especially the minimum arm since γ is much closer to the arm with minimum mean compared to the arm with maximum mean). In the infeasible case when $\gamma = 0.9$, we can observe that the sampling proportion of the algorithm almost matches the theoretical optimal $\boldsymbol{w}^*(\boldsymbol{\mu})$ in Theorem 1.

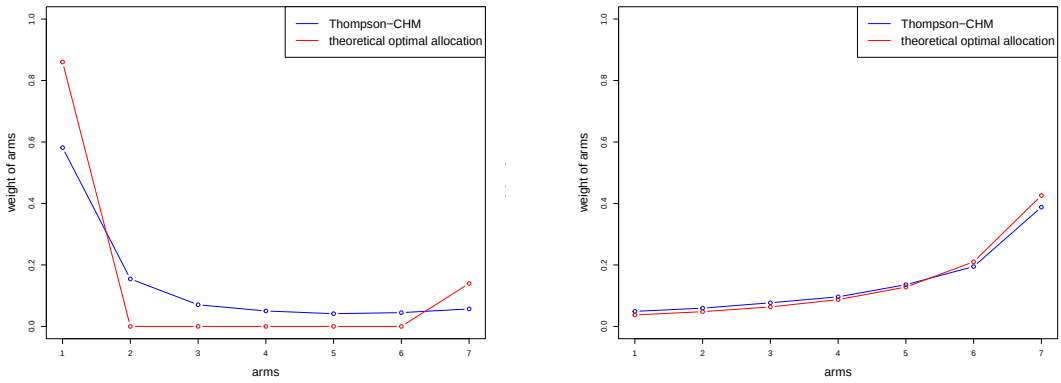

Figure 1.2: Empirical proportion of samples compared to optimal allocation $\boldsymbol{w}^*(\boldsymbol{\mu})$ in feasible cases (left) and infeasible cases (right) estimated using 100 repetitions.

We further investigate how the sample complexity of Thompson-CHM scales with the confidence parameter δ, as motivated by the logarithmic dependence predicted by our theoretical results. We choose $\gamma = 0.4$ for feasible case and $\gamma = 0.9$ for infeasible case, and a varying range of $-\log(\delta)$ to evenly range between 1 to 5. For each δ, we run 100 independent trials and report the average sample complexity. As shown in Figure 1.3, the empirical sample complexity grows linearly with $\log(1/\delta)$, closely tracking the theoretical lower bound up to a constant factor. This supports the sharpness of our bound and confirms that the Thompson-CHM algorithm achieves asymptotic optimal sample complexity in terms of its δ dependence.

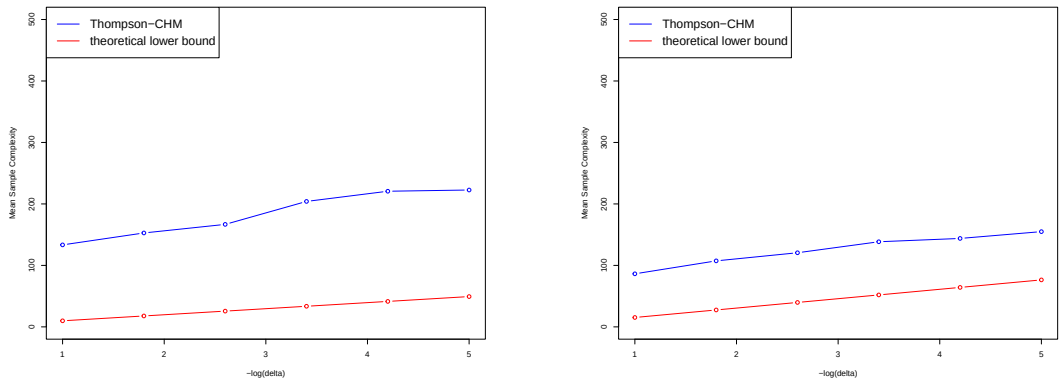

Figure 1.3: Sample complexity for different δ's in feasible cases (left) and infeasible cases (right).

1.8 Discussion and Conclusion

This chapter thoroughly investigates the convex hull membership (CHM) problem in the pure exploration setting and mostly focused on 1-dimensional setting. We propose a novel asymptotically optimal algorithm to tackle this problem, which we refer to as Thompson-CHM algorithm. The sampling rule combines the ideas of top-two Thompson sampling [Russo, 2016] and Murphy sampling [Kaufmann et al., 2018], and it can automatically guarantee the sampled proportion of each arm converges to the optimal allocation derived by the information-theoretical lower bound in the one-dimensional setting, regardless of relative position between the threshold and the arm mean set. Moreover, we extend our results to the interval CHM setting that generalizes some important MAB problems in the literature and investigate the extensions of the Thompson-CHM algorithm in higher dimensions.

In the next chapter, we will attempt to derive a complete solution to d-dimensional CHM problems with broader settings when $d \geq 2$. We conjecture that the sample complexity bounds and the asymptotically optimal algorithm are identical to the one-dimensional case. The current theoretical results reveal challenges in the feasible d-dimensional setting due to the complex geometric structure in the "alternative" set. We will fully understand the CHM problem in d-dimensional case when $d \geq 2$ for the Gaussian case.

CHAPTER 2

Generalization of Stochastic Convex Hull Membership Problem to the Higher-Dimensional Case

This chapter extends the stochastic convex hull membership problem discussed in Chapter 1 to dimensions $d \geq 2$. We focus on the Gaussian case with known identity covariance, where the information-theoretic game admits closed geometric forms. In the feasible case, the characteristic time is governed by an optimal design problem over separating hyperplanes, and non-vertex arms are dominated. In the infeasible case, the least confusing alternative is obtained by moving a small convex combination of arms to the target, yielding a genuinely higher-dimensional barycentric game. These formulas clarify the role of Carathéodory's theorem: a feasibility certificate may use at most $d + 1$ arms, but an asymptotically optimal sampling allocation need not be supported on a single certificate. We then give a high-dimensional generalized likelihood-ratio stopping rule, an oracle Track-and-Stop algorithm that is asymptotically optimal under standard regularity assumptions, and a Thompson-CHM extension whose feasible-case optimality follows directly and whose infeasible-case optimality reduces to a geometric large-deviation gap condition.

The motivation of the generalization comes from the fixed-confidence pure exploration program initiated by Chernoff's sequential design viewpoint and developed in modern bandit identification through information-theoretic games, Track-and-Stop procedures, and game-based pure exploration algorithms [Chernoff, 1959, Garivier and Kaufmann, 2016, Degenne et al., 2019, 2020]. The geometric decision problem considered here is also closely connected to deterministic convex hull membership and to recent stochastic formulations of convex feasibility sampling [Filippozzi et al., 2023, Niss et al., 2022].

2.1 Introduction

The one-dimensional CHM problem conveys a particularly transparent solution because convex hull membership reduces to two scalar comparisons: whether the target is below the largest mean and above the smallest mean. This total ordering is precisely what makes

the optimal allocation explicit. In the feasible case, the two extreme arms are the only statistically relevant arms; in the infeasible case, every arm must be protected against the possibility that it alone crosses the threshold. Chapter 1 develops a complete lower bound, stopping rule, and Thompson-sampling algorithm based on this observation.

In dimension $d \geq 2$, the same feasible/infeasible dichotomy survives, but the mechanism behind it changes. Feasibility is no longer witnessed by a pair of ordered extremes, it is witnessed by a convex combination. Infeasibility is no longer witnessed by all means lying on one side of a scalar threshold, it is witnessed by a separating hyperplane. Thus the natural replacement for the two one-dimensional sides of γ is the dual pair of certificates furnished by convex geometry: barycentric certificates for feasibility and separating certificates for infeasibility. This is why the higher-dimensional characteristic time is best written as a game over either directions $u \in \mathbb{S}^{d-1}$ or convex weights $q \in \Delta_K$, rather than as a formula involving only arm-wise distances to the target.

This chapter develops the Gaussian distribution case in detail. In this case, KL projections are Euclidean projections and the geometry is visible without additional notation. The resulting formulas should be viewed as the higher-dimensional analogue of the one-dimensional inverse-KL expressions. They also clarify the precise role of Carathéodory's theorem and the separating hyperplane theorem, two classical tools in finite-dimensional convexity [Rockafellar, 1997, Barvinok, 2025]. Carathéodory's theorem guarantees that a feasible point has a sparse certificate involving at most $d+1$ arms, but the statistical design problem is more demanding: an optimal allocation must gather evidence against all low-cost alternatives. Consequently, a single sparse feasibility certificate need not determine the asymptotic sample complexity.

The contribution of the chapter is organized as follows. Section 2.2 fixes notation and regularity. Section 2.3 derives the exact Gaussian characteristic-time game. Sections 2.4 and 2.5 interpret the feasible and infeasible sides of the game. Section 2.6 gives the generalized likelihood-ratio stopping rule. Section 2.7 gives an asymptotically optimal Track-CHM benchmark. Finally, Section 2.8 returns to the Thompson-CHM philosophy of Chapter 1 and identifies the remaining large-deviation gap needed for a complete higher-dimensional Thompson proof.

2.2 Setup and Regularity

We begin with a formulation of the problem in translated coordinates. The translation by γ is not merely a notational convenience: it makes the two alternatives geometrically symmetric. The feasible case becomes the statement that the origin belongs to a convex hull, while

the infeasible case becomes the statement that the origin can be strictly separated from that convex hull. This convention will be used throughout the chapter and in the appendix proofs.

We consider the d-dimensional stochastic convex hull membership problem with $d \geq 2$. There are K arms. Pulling arm i produces an independent observation

$$Y_{i,s} \sim \mathcal{N}(\mu_i, I_d), \qquad \mu_i \in \mathbb{R}^d,$$

where I_d is the $d \times d$ identity matrix. The target point is denoted by $\gamma \in \mathbb{R}^d$. As in the one-dimensional problem, the goal is to determine, with fixed confidence, whether

$$\gamma \in \mathrm{Conv}(\mu_1, \ldots, \mu_K).$$

Throughout this chapter we translate the target to the origin. Define

$$x_i = \mu_i - \gamma, \qquad i = 1, \ldots, K.$$

Then the problem is equivalent to deciding whether

$$0 \in \mathrm{Conv}(x_1, \ldots, x_K).$$

We write

$$X = (x_1, \ldots, x_K) \in (\mathbb{R}^d)^K.$$

The feasible and infeasible sets are

$$\mathcal{F} = \left\{ X \in (\mathbb{R}^d)^K : 0 \in \mathrm{Conv}(x_1, \ldots, x_K) \right\},$$

and

$$\mathcal{I} = \left\{ X \in (\mathbb{R}^d)^K : 0 \notin \mathrm{Conv}(x_1, \ldots, x_K) \right\}.$$

The Gaussian KL divergence between two arms with means $u, v \in \mathbb{R}^d$ is

$$d_G(u, v) = \frac{1}{2} \|u - v\|^2 .$$

Therefore the information-theoretic characteristic time is

$$T_G^*(\mu)^{-1} = \sup_{w \in \Delta_K} \inf_{\lambda \in \mathrm{Alt}(\mu)} \frac{1}{2} \sum_{i=1}^{K} w_i \|\mu_i - \lambda_i\|^2 , \tag{2.1}$$

where

$$\Delta_K = \left\{ w \in \mathbb{R}_+^K : \sum_{i=1}^K w_i = 1 \right\},$$

and $\mathrm{Alt}(\mu)$ is the set of mean vectors whose convex hull membership answer is different from that of μ.

The expression above is the direct analogue of the lower-bound game used throughout fixed-confidence pure exploration literature [Garivier and Kaufmann, 2016, Degenne et al., 2019, Degenne and Koolen, 2019]. The allocation vector w represents the asymptotic proportions, whereas the alternative vector λ represents the least distinguishable instance with the opposite answer. Thus $T_G^*(\mu)$ is not only a lower-bound constant, it is the object from which the target sampling proportions are read.

We call an instance regular if either

$$0 \in \operatorname{int} \operatorname{Conv}(x_1, \ldots, x_K), \tag{2.2}$$

or

$$0 \notin \operatorname{Conv}(x_1, \ldots, x_K). \tag{2.3}$$

The boundary case

$$0 \in \partial \operatorname{Conv}(x_1, \ldots, x_K)$$

is excluded. This is the higher-dimensional analogue of excluding the one-dimensional case where the threshold equals an arm mean. At such boundary instances the decision is not locally stable and the characteristic time may degenerate.

The regularity assumption is the higher-dimensional version of a positive margin condition. In the feasible case it asks that the target be surrounded by the convex hull in all directions; in the infeasible case compactness gives a positive Euclidean distance from the target to the convex hull. This distinction is important because the generalized likelihood-ratio statistic below has a positive linear growth rate exactly away from the boundary.

Remark 2.2.1. If the observations are Gaussian with known positive definite covariance Σ, one may apply the whitening map $z \mapsto \Sigma^{-1/2} z$. All statements remain valid after replacing the Euclidean norm by the Mahalanobis norm

$$\|z\|_{\Sigma^{-1}}^2 = z^\top \Sigma^{-1} z.$$

We keep the identity-covariance assumption in the main statements to avoid obscuring the main ideas. For exponential-family arms, the same variational formulas hold with the corresponding KL divergence, but the projections that appear below become KL projections

rather than Euclidean projections. The Gaussian model is therefore the cleanest setting in which to see the new geometry.

2.3 Exact Gaussian Characteristic Time

The central object in higher dimensions is not an ordering of arms but a pair of dual geometric certificates. Feasibility is certified by a convex combination, while infeasibility is certified by a separating hyperplane. This leads to two different explicit games.

The next theorem starts from the standard information game and eliminates the infinite-dimensional alternative set by two elementary geometric projections. When the true instance is feasible, the adversary tries to make the empirical convex hull miss the origin, by separation, this means pushing all arms into a common closed halfspace. When the true instance is infeasible, the adversary tries to make the origin enter the convex hull, by barycentricity, this means finding a convex combination that can be moved to zero. The two resulting formulas are the exact higher-dimensional analogues of the one-dimensional lower-bound expressions.

Theorem 6. *Assume $d \geq 2$, let $Y_{i,s} \sim \mathcal{N}(\mu_i, I_d)$, and write $x_i = \mu_i - \gamma$.*

(i) Feasible case. *Assume*

$$0 \in \operatorname{int} \operatorname{Conv}(x_1, \dots, x_K).$$

Then

$$T_G^*(\mu)^{-1} = \sup_{w \in \Delta_K} \inf_{u \in \mathbb{S}^{d-1}} \frac{1}{2} \sum_{i=1}^{K} w_i \left(-u^\top x_i\right)_+^2. \tag{2.4}$$

The vector u is the normal vector of a separating hyperplane through the target.

(ii) Infeasible case. *Assume*

$$0 \notin \operatorname{Conv}(x_1, \dots, x_K).$$

Then

$$T_G^*(\mu)^{-1} = \sup_{w \in \Delta_K} \inf_{q \in \Delta_K} \frac{\left\| \sum_{i=1}^K q_i x_i \right\|^2}{2 \sum_{i=1}^K q_i^2 / w_i}, \tag{2.5}$$

with the convention that $q_i^2/w_i = +\infty$ when $q_i > 0$ and $w_i = 0$. Under this convention, a denominator equal to $+\infty$ gives value zero.

A useful way to read (2.4) is as a continuous optimal-design problem. Each direction u represents a possible separating certificate against feasibility, and one chooses w to maximize the weakest accumulated evidence over all such directions. Formula (2.5) has the dual flavor: each q represents a possible convex-combination certificate that would make the

target feasible, and the denominator records how expensive it is to distribute the required movement across arms that are sampled with proportions w_i. The appearance of q_i^2/w_i is the usual quadratic price of moving arm i when arm i receives information at rate w_i.

It is important to emphasize what is new in dimension $d \geq 2$ compared to the one-dimensional case. In one dimension, an infeasible instance has all translated means x_i on the same side of the origin. Then the adversary's cheapest feasible alternative is obtained by moving one arm to the threshold. In higher dimensions, the adversary may move several arms jointly so that their convex combination reaches the target. Thus the variable q in (2.5) is not merely a technical artifact, it is the barycentric certificate of the adversarial feasible alternative.

2.3.1 Reduction to the one-dimensional formula

Before developing the new geometry, we verify that the barycentric game is consistent with the formula already obtained in Chapter 1. This step shows that the higher-dimensional expression is not an unrelated relaxation, but an exact extension of the original inverse-KL allocation.

The exact game above reduces to the one-dimensional result in Chapter 1. Suppose $d = 1$ and the instance is infeasible with $x_i > 0$ for all i. For fixed w, (2.5) becomes

$$\inf_{q \in \Delta_K} \frac{\left(\sum_i q_i x_i\right)^2}{2 \sum_i q_i^2 / w_i}.$$

Let

$$m_w = \min_{1 \leq i \leq K} w_i x_i^2.$$

Since $q_i \geq 0$ and $x_i > 0$,

$$\left(\sum_i q_i x_i\right)^2 \geq \sum_i q_i^2 x_i^2 \geq m_w \sum_i q_i^2 / w_i.$$

Hence the infimum is at least $m_w/2$, and equality is attained by choosing $q = e_j$, where $j \in \operatorname{argmin}_i w_i x_i^2$. Therefore

$$\inf_{q \in \Delta_K} \frac{\left(\sum_i q_i x_i\right)^2}{2 \sum_i q_i^2 / w_i} = \min_i w_i \frac{x_i^2}{2} = \min_i w_i d_G(\mu_i, \gamma).$$

Maximizing over $w \in \Delta_K$ gives

$$T_G^*(\mu) = \sum_{i=1}^{K} \frac{1}{d_G(\mu_i, \gamma)},$$

with optimal allocation proportional to $d_G(\mu_i, \gamma)^{-1}$. The case $x_i < 0$ for all i is identical.

Thus the inverse-KL allocation in the one-dimensional infeasible case is a special case of the barycentric game (2.5), attained when the least-cost convex-combination certificate is a single arm.

2.4 Geometry of the Feasible Case

Theorem 6 shows that the feasible case is an optimal design problem over separating directions. This section records two consequences: non-vertex arms are dominated, and the optimizer is characterized by a dual distribution over separating hyperplanes.

The feasible case is the side on which the higher-dimensional problem most visibly differs from the one-dimensional problem. In the one-dimensional case, there are only two separating directions, corresponding to the two sides of the threshold. In dimension $d \geq 2$, the adversary may choose any direction on the unit sphere. Therefore, one must build a design that is robust to a continuum of possible separating hyperplanes. The results in this section explain which arms can matter and how the optimal design balances the worst separating directions.

2.4.1 Non-vertices are dominated

Let

$$V = \operatorname{Vert} \operatorname{Conv}(x_1, \ldots, x_K)$$

be the set of vertices of the convex hull.

The next statement formalizes the extreme-point principle. If an arm lies in the convex hull of other arms, then along every separating direction its squared negative projection is no larger than the corresponding convex combination of the squared negative projections of the vertices. Sampling such an interior arm therefore never improves the worst-direction evidence. This is the exact higher-dimensional analogue of the fact that, in the one-dimensional feasible case, only the minimum and maximum means need be sampled.

Proposition 2.4.1. *Assume*

$$0 \in \operatorname{int} \operatorname{Conv}(x_1, \ldots, x_K).$$

In the Gaussian feasible game (2.4)*, there exists an optimal allocation* w^* *supported on the vertices of* $\mathrm{Conv}(x_1, \ldots, x_K)$*. Equivalently, every non-vertex arm can be assigned zero sampling mass without decreasing the value of the game.*

Remark 2.4.2. This proposition gives the precise mathematical intuition of the qualitative statement that, in the feasible case, the optimal strategy samples only extreme points. It does not say that every vertex must be sampled. Which vertices are sampled is determined by the saddle point of the separation game.

We note that the statement is an existence theorem for optimal allocations, not a prescription to sample every vertex. A vertex may still be irrelevant if it is never active in a least favorable separating direction. Conversely, in symmetric configurations many vertices may be simultaneously active, even though only a small number of them are needed to certify feasibility in the deterministic sense.

2.4.2 Dual form of the feasible game

The feasible game admits a useful dual interpretation. One player chooses an allocation over arms, the adversary chooses a separating direction. Equivalently, the adversary may randomize over separating directions.

The following dual form is a useful diagnostic for optimality. It is a finite-arm, continuous-direction minimax identity in the spirit of classical zero-sum design problems, the exchange of sup and inf is justified by Sion's minimax theorem [Sion, 1958]. The dual distribution ρ should be interpreted as a least favorable distribution over separating hyperplanes. At an optimum, every sampled arm has the same average exposure to this least favorable distribution, while ρ concentrates on directions that are tight for the chosen allocation.

Proposition 2.4.3. *Let*

$$\ell_i(u) = \frac{1}{2}\left(-u^\top x_i\right)_+^2.$$

Then

$$\sup_{w\in\Delta_K} \inf_{u\in\mathbb{S}^{d-1}} \sum_{i=1}^K w_i \ell_i(u) = \inf_{\rho\in\mathcal{P}(\mathbb{S}^{d-1})} \max_{1\le i\le K} \int_{\mathbb{S}^{d-1}} \ell_i(u)\,\rho(\mathrm{d}u), \tag{2.6}$$

where $\mathcal{P}(\mathbb{S}^{d-1})$ *denotes the set of probability measures on the unit sphere. Moreover, if* w^* *and* ρ^* *form a saddle point and the value is* V*, then*

$$\int_{\mathbb{S}^{d-1}} \ell_i(u)\,\rho^*(\mathrm{d}u) = V \qquad \text{for every } i \text{ with } w_i^* > 0, \tag{2.7}$$

and ρ^* is supported on directions u satisfying

$$\sum_{i=1}^{K} w_i^* \ell_i(u) = V. \tag{2.8}$$

The equalization conditions (2.7)–(2.8) play the same conceptual role as the equalization of inverse-KL terms in the one-dimensional lower bound. In the one-dimensional case, the saddle point makes the two sides of the threshold equally difficult to refute. In higher dimensions, the saddle point equalizes a family of directional risks. This viewpoint is often the most convenient way to reason about any implementations, because it separates the arm weights from the least favorable geometry.

Remark 2.4.4. Let $s_1, \dots, s_m$ be the indices of the vertices of $\mathrm{Conv}(x_1, \dots, x_K)$. In the Gaussian feasible case, the oracle functions f_i appearing in the higher-dimensional lower-bound statement can be identified as any measurable optimizer of the game

$$w^F(x) \in \mathrm{argmax}_{w \in \Delta_K} \inf_{u \in \mathbb{S}^{d-1}} \frac{1}{2} \sum_{i=1}^{K} w_i \left(-u^\top x_i\right)_+^2. \tag{2.9}$$

By Proposition 2.4.1, there is an optimizer with

$$w_i^F(x) = 0 \qquad \text{whenever } x_i \notin \mathrm{Vert}\,\mathrm{Conv}(x_1, \dots, x_K).$$

Thus one may take

$$f_j(x_{s_1}, \dots, x_{s_m}, 0) = w_{s_j}^F(x), \qquad j = 1, \dots, m.$$

When the optimizer is not unique, f_j should be interpreted as a choice of measurable selector.

2.4.3 Carathéodory certificates and optimal designs

We next separate two notions that are easily conflated. A deterministic certificate of feasibility answers the yes/no geometric question: can the target be written as a convex combination of observed means? An optimal sampling design answers a different statistical question: how should evidence be collected so that every nearby infeasible alternative becomes distinguishable? Interestingly, Carathéodory's theorem solves the first question by giving a sparse certificate, but it does not solve the second question by itself.

Carathéodory's theorem says that, if

$$0 \in \mathrm{Conv}(x_1, \dots, x_K),$$

then there is a subset $S \subseteq [K]$ with

$$|S| \leq d + 1$$

such that

$$0 \in \mathrm{Conv}\{x_i : i \in S\}.$$

This gives a small deterministic certificate of feasibility. However, the optimal sequential sampling problem is not merely to find one feasible certificate, it is to collect evidence against all separating hyperplanes. Consequently, a single Carathéodory certificate need not be statistically optimal.

This phenomenon is a real departure from the one-dimensional theory. In the one-dimensional case, the feasible certificate necessarily consists of the two extremes, and these are also the statistically optimal arms. In higher dimensions, there may be many different simplices containing the target. The optimal allocation may be better understood as a mixture over such local certificates, chosen so that no separating direction remains weakly protected.

For a subset $S \subseteq [K]$, define the restricted feasible value

$$V(S) = \sup_{w \in \Delta(S)} \inf_{u \in \mathbb{S}^{d-1}} \frac{1}{2} \sum_{i \in S} w_i \left(-u^\top x_i\right)_+^2, \tag{2.10}$$

where $\Delta(S) = \{w \in \Delta_K : \mathrm{supp}(w) \subseteq S\}$.

The following proposition is a concrete example to show that a Carathéodory certificate need not be optimal. The regular hexagon example below is deliberately elementary. Its purpose is to rule out the tempting conjecture that asymptotic optimality can always be reduced to identifying one favorable simplex of size $d + 1$. The example shows that statistical optimality may require averaging information over more vertices than any single Carathéodory representation uses.

Proposition 2.4.5. *There exists a feasible instance in dimension $d = 2$ for which the optimal allocation uses more than $d + 1 = 3$ vertices. In particular, a single Carathéodory certificate need not achieve the optimal characteristic time.*

Remark 2.4.6. Carathéodory's theorem is a statement about the existence of a feasibility certificate. The characteristic time is instead a robust certification quantity. It asks how much evidence must be gathered so that every separating hyperplane is ruled out. In asymmetric instances, the optimal design may indeed concentrate on one small simplex. In symmetric or nearly symmetric instances, the optimal design may mix over many local certificates and sample many vertices.

For algorithm design, this suggests that a higher-dimensional CHM method should not first commit to a single simplex unless additional structure is known. A simplex-based rule can be locally valid but globally sub-optimal. The separation game avoids this premature commitment by optimizing directly over all separating directions.

2.5 Geometry of the Infeasible Case

The infeasible game has a different structure. The adversary tries to create feasibility by moving the means so that the origin becomes a convex combination. The variable q in (2.5) is precisely this attempted convex combination.

The infeasible case reverses the roles of sparsity. The least confusing alternative may be witnessed by a sparse convex combination, but one cannot know in advance which arm or small group of arms will form the most dangerous certificate. Therefore, the adversary's witness may be sparse while one's optimal allocation must have full support. This is the same qualitative phenomenon as in the one-dimensional infeasible case, now expressed through barycentric geometry.

The next proposition proves a fundamental fact that every arm is needed in the infeasible case, which coincides with the one-dimensional case. This fact is intuitive, since we only need the additional of one arm to make an infeasible case to become feasible. Therefore, in order to identify an infeasible case, we need to traverse all arms.

More precisely, if an arm receives zero asymptotic mass, then the information cost of moving that arm is zero in the lower-bound game. The adversary can exploit this by using that unsampled arm as a cheap route toward a feasible alternative. Full support is therefore not a numerical artifact of the formula, it is forced by the logical structure of proving infeasibility.

Proposition 2.5.1. *Assume*

$$0 \notin \mathrm{Conv}(x_1, \ldots, x_K).$$

Then every optimizer w^* of the Gaussian infeasible game (2.5) *has full support:*

$$w_i^* > 0, \qquad i = 1, \ldots, K.$$

The next result gives the complementary side of the story. Although one must sample every arm, the adversary does not need to move every arm to create a feasible alternative. Carathéodory's theorem implies that, once feasibility is created, it can be certified by at most $d + 1$ moved means. Thus higher dimension produces an asymmetry: full support on the learning side, sparse certificates on the adversarial side.

Proposition 2.5.2. *In the infeasible Gaussian game, the infimum over $q \in \Delta_K$ in* (2.5) *may be restricted to convex combinations supported on at most $d+1$ arms. That is,*

$$\inf_{q \in \Delta_K} \frac{\|\sum_i q_i x_i\|^2}{2\sum_i q_i^2/w_i} = \inf_{\substack{q \in \Delta_K \\ |\operatorname{supp}(q)| \le d+1}} \frac{\|\sum_i q_i x_i\|^2}{2\sum_i q_i^2/w_i}. \tag{2.11}$$

Remark 2.5.3. Proposition 2.5.2 is the correct higher-dimensional role of Carathéodory's theorem in the infeasible case: the least confusing alternative can be witnessed by at most $d+1$ moved arms. This does not contradict Proposition 2.5.1. The adversary may use a sparse certificate at any fixed alternative, but one must sample every arm because any unsampled arm could be moved at zero information cost.

This sparse-adversary/full-support contrast is a useful organizing principle for the rest of the chapter. It explains why the stopping statistic involves an infimum over q, while an optimal infeasible allocation still assigns positive mass to every coordinate. It also anticipates the difficulty of the Thompson proof: rare posterior feasibility is indexed by small barycentric certificates, not by individual crossing events.

2.6 A High-dimensional Stopping Rule

The one-dimensional stopping rule has three explicit cases: all arms above the threshold, all arms below the threshold, or one arm on each side. In higher dimensions these cases are replaced by convex and separating certificates. The natural stopping statistic is a generalized likelihood ratio.

The stopping rule is where the geometric reduction becomes operational. Rather than maintaining separate confidence statements for the minimum and maximum arms, we ask whether the empirical instance can be separated from the opposite decision region by likelihood evidence. This is the same principle underlying Chernoff-style stopping and modern Track-and-Stop algorithms [Chernoff, 1959, Garivier and Kaufmann, 2016], but the alternative set is now a convex-geometric object.

Let $N_i(t)$ be the number of samples from arm i up to time t, and let

$$\widehat{\mu}_i(t) = \frac{1}{N_i(t)} \sum_{s:A_s=i} Y_{i,s}$$

be the empirical mean. Define

$$\widehat{x}_i(t) = \widehat{\mu}_i(t) - \gamma.$$

The empirical decision is

$$\widehat{I}_t = \begin{cases} \text{feasible}, & 0 \in \mathrm{Conv}(\widehat{x}_1(t), \dots, \widehat{x}_K(t)), \\ \text{infeasible}, & 0 \notin \mathrm{Conv}(\widehat{x}_1(t), \dots, \widehat{x}_K(t)). \end{cases}$$

Define the GLR statistic

$$Z(t) = \inf_{\lambda:\, I(\lambda) \neq \widehat{I}_t} \frac{1}{2} \sum_{i=1}^{K} N_i(t) \left\| \widehat{\mu}_i(t) - \lambda_i \right\|^2 . \tag{2.12}$$

By Theorem 6, this statistic has explicit forms.

These two forms are the empirical counterparts of the two population games. Equation (2.13) asks for the least costly separating hyperplane against empirical feasibility. Equation (2.14) asks for the least costly empirical convex combination that would reverse empirical infeasibility. In this sense, the stopping rule does not introduce a new object, it simply plugs empirical means and sample counts into the same geometry that defines the lower bound.

If the empirical instance is feasible, then

$$Z(t) = \inf_{u \in \mathbb{S}^{d-1}} \frac{1}{2} \sum_{i=1}^{K} N_i(t) \left(-u^\top \widehat{x}_i(t) \right)_+^2 . \tag{2.13}$$

If the empirical instance is infeasible, then

$$Z(t) = \inf_{q \in \Delta_K} \frac{\left\| \sum_{i=1}^{K} q_i \widehat{x}_i(t) \right\|^2}{2 \sum_{i=1}^{K} q_i^2 / N_i(t)}. \tag{2.14}$$

The threshold assumption below abstracts the concentration inequality needed for correctness. One may obtain such thresholds from Gaussian mixture martingales or related anytime confidence sequences, the use of mixture martingales in sequential tests and confidence intervals is developed, for example, by Kaufmann and Koolen [2021]. For the asymptotic statements in this chapter, the essential requirement is that the leading term be $\log(1/\delta)$, matching the information-theoretic lower bound.

Assumption 7. The threshold $\beta(t, \delta)$ satisfies, for every Gaussian instance μ,

$$\mathbb{P}_\mu \left(\exists t \geq 1 : \frac{1}{2} \sum_{i=1}^{K} N_i(t) \left\| \widehat{\mu}_i(t) - \mu_i \right\|^2 \geq \beta(t, \delta) \right) \leq \delta, \tag{2.15}$$

and, as $\delta \downarrow 0$,

$$\beta(t, \delta) = \log(1/\delta) + o(\log(1/\delta)) \tag{2.16}$$

along stopping times of order $\log(1/\delta)$.

The stopping and decision rules are

$$\tau_\delta = \inf\{t \geq 1 : Z(t) \geq \beta(t, \delta)\}, \tag{2.17}$$

and

$$I_{\tau_\delta} = \widehat{I}_{\tau_\delta}.$$

The next theorem is the correctness part of the stopping argument. It says that the GLR statistic is compatible with any anytime threshold controlling the Gaussian likelihood confidence set. The proof is purely set-theoretic: if the algorithm stops and returns the wrong answer, then the true parameter itself is an alternative to the empirical decision and must therefore have been excluded by the confidence set.

Theorem 8. *Under Assumption 7, the stopping rule* (2.17) *with decision* $I_{\tau_\delta} = \widehat{I}_{\tau_\delta}$ *is* δ*-correct.*

Correctness alone does not determine the sample complexity. The next lemma identifies the linear growth rate of the GLR statistic under any sampling rule whose empirical allocation converges. This is the bridge between allocation design and stopping time: once $N(t)/t$ approaches a target vector w, the stopping statistic grows at rate exactly equal to the inner value $C(w, \mu)$ of the characteristic-time game.

Lemma 2.6.1. *Let* μ *be a regular instance. Suppose an anytime sampling rule satisfies*

$$\frac{N_i(t)}{t} \to w_i \qquad \textit{almost surely}$$

for every i*, and* $N_i(t) \to \infty$ *for every* i*. Then*

$$\frac{Z(t)}{t} \to C(w, \mu) \qquad \textit{almost surely,} \tag{2.18}$$

where $C(w, \mu)$ *is the inner value of the corresponding game:*

$$C(w, \mu) = \begin{cases} \inf\limits_{u \in \mathbb{S}^{d-1}} \dfrac{1}{2} \sum\limits_i w_i \left(-u^\top x_i\right)_+^2, & 0 \in \operatorname{int} \operatorname{Conv}(x_1, \ldots, x_K), \\ \inf\limits_{q \in \Delta_K} \dfrac{\left\|\sum_i q_i x_i\right\|^2}{2 \sum_i q_i^2 / w_i}, & 0 \notin \operatorname{Conv}(x_1, \ldots, x_K). \end{cases}$$

The following corollary is the high-dimensional analogue of the stopping-time upper bound in Chapter 1. It shows that the stopping rule is asymptotically sharp provided the sampling

rule tracks the optimizer of the lower-bound game. Thus the remaining algorithmic task is not to redesign the stopping rule, but to construct a sampling rule whose allocation converges to $w^*(\mu)$.

Corollary 2.6.2. *Assume the conditions of Lemma 2.6.1, and suppose*

$$C(w, \mu) > 0.$$

Then the GLR stopping rule satisfies

$$\limsup_{\delta \downarrow 0} \frac{\tau_\delta}{\log(1/\delta)} \leq \frac{1}{C(w, \mu)} \qquad \textit{almost surely.} \tag{2.19}$$

In particular, if $w = w^*(\mu)$ is an optimizer of the characteristic-time game, then

$$\limsup_{\delta \downarrow 0} \frac{\tau_\delta}{\log(1/\delta)} \leq T_G^*(\mu) \qquad \textit{almost surely.}$$

2.7 An asymptotically Optimal High-dimensional Track-CHM Algorithm

The exact characteristic-time formulas suggest a direct oracle algorithm: estimate the current instance, solve the corresponding game, track the estimated optimal allocation, and stop using the GLR statistic.

This section gives a benchmark algorithm rather than the final Thompson-style construction. Track-and-Stop algorithms show that, once the characteristic-time game is known and regularity holds, asymptotic optimality can be obtained by tracking the plug-in estimate of the optimal allocation [Garivier and Kaufmann, 2016, Degenne et al., 2019, 2020]. In the present problem, the novelty lies in the form of the game, after that game is identified, the tracking principle follows the established pure-exploration technique.

Let

$$\mathcal{W}(\mu) = \operatorname{argmax}_{w \in \Delta_K} C(w, \mu)$$

be the set of optimal allocations. In this section we assume that the optimizer is unique at the true instance and denote it by $w^*(\mu)$.

The local uniqueness assumption is standard in asymptotic optimality statements for tracking rules. It rules out discontinuities caused by several equally good saddle-point allocations. When uniqueness fails, one can still work with set-valued tracking or approachability-type arguments, but the statement becomes heavier and obscures the main geometric mes-

Algorithm 2 HD-Track-CHM

Input: target γ, confidence δ, threshold $\beta(t, \delta)$

Pull each arm once.
for $t = K, K+1, \ldots$ **do**
  Compute empirical means $\widehat{\mu}_i(t)$ and translated means $\widehat{x}_i(t) = \widehat{\mu}_i(t) - \gamma$.
  Compute $Z(t)$ from (2.13) or (2.14), according to the empirical answer.
  if $Z(t) \geq \beta(t, \delta)$ **then**
    Stop and return feasible iff $0 \in \mathrm{Conv}(\widehat{x}_1(t), \ldots, \widehat{x}_K(t))$.
  end if
  if there exists i with $N_i(t) < \sqrt{t}$ **then**
    Pull such an arm i.
  else
    Compute a measurable selector

$$\widehat{w}(t) \in \mathcal{W}(\widehat{\mu}(t)).$$

    Pull

$$A_{t+1} \in \mathrm{argmax}_{i \in [K]} \left\{ \sum_{s=K}^{t} \widehat{w}_i(s) - N_i(t) \right\}.$$

  end if
end for

sage.

Assumption 9 (Local uniqueness and continuity). The true instance μ is regular. The optimizer $w^*(\mu)$ is unique, and there exists a neighborhood $\mathcal{U}$ of μ such that a measurable choice $w^*(\nu) \in \mathcal{W}(\nu)$ is continuous at μ.

The tracking lemma records the deterministic part of the allocation argument. Forced exploration guarantees consistency of all empirical means, while cumulative tracking prevents the realized counts from drifting linearly away from the estimated target proportions. This separation between statistical estimation and deterministic tracking is exactly the reason Track-and-Stop methods are robust across many pure-exploration problems.

Lemma 2.7.1 (Tracking)**.** *Under Assumption 9, HD-Track-CHM satisfies*

$$\frac{N(t)}{t} \to w^*(\mu) \qquad \textit{almost surely.}$$

The following theorem combines three ingredients established above: the GLR rule is correct, the tracking rule converges to the optimal allocation, and the GLR statistic grows at the value of the game. In this sense, HD-Track-CHM gives a complete higher-dimensional asymptotic solution for regular Gaussian instances, conditional only on the usual continuity

of the game optimizer.

Theorem 10. *Assume Assumptions 7 and 9. Then HD-Track-CHM is δ-correct. Moreover,*

$$\limsup_{\delta \downarrow 0} \frac{\tau_\delta}{\log(1/\delta)} \leq T_G^*(\mu) \qquad \text{almost surely.} \tag{2.20}$$

If, in addition, the family

$$\{\tau_\delta / \log(1/\delta) : \delta \in (0, 1/2)\}$$

is uniformly integrable under $\mathbb{P}_\mu$, then

$$\lim_{\delta \downarrow 0} \frac{\mathbb{E}_\mu[\tau_\delta]}{\log(1/\delta)} = T_G^*(\mu). \tag{2.21}$$

Remark 2.7.2. The Track-CHM algorithm is not intended to replace Thompson-CHM as the conceptual algorithm. Rather, it gives a clean benchmark theorem: once the higher-dimensional game is solved, tracking the estimated saddle-point allocation and using the GLR stopping rule yields an asymptotically optimal fixed-confidence procedure. The Thompson construction below is the natural analogue of the algoriothm in Chapter 1.

This distinction is useful for exposition. HD-Track-CHM proves that the characteristic-time formulas are operationally meaningful. HD-Thompson-CHM, discussed next, is closer to the algorithmic philosophy of the previous chapter: rather than solving only the empirical game, it samples from a posterior conditioned on the event whose rare fluctuations determine the optimal allocation.

2.8 A Higher-dimensional Thompson-CHM Rule

The Appendix A proposes a higher-dimensional Thompson-CHM algorithm with categorical weights f_i. The Gaussian game identifies the feasible-case version of these weights.

The sampling rule in Chapter 1 is motivated by two related Bayesian ideas: top-two Thompson sampling for best-arm identification [Russo, 2016] and Murphy sampling for testing the lowest mean [Kaufmann et al., 2018]. Both methods use posterior conditioning to force samples from the statistically relevant alternative event. The same principle is natural here: draw a posterior instance conditioned on feasibility, then sample according to the optimal feasible allocation of that posterior draw.

The following definition is the promised identification of the functions f_i. In one dimension, a feasible posterior draw has a sampled minimum and maximum, and the Bernoulli parameter in Thompson-CHM balances these two extremes. In higher dimensions, the sam-

pled feasible instance has a set of vertices and a continuum of separating directions, the correct categorical distribution is therefore the optimizer of the feasible separation game.

For a feasible parameter vector $\theta = (\theta_1, \dots, \theta_K)$, write

$$z_i = \theta_i - \gamma.$$

Define

$$w^F(\theta) \in \operatorname{argmax}_{w \in \Delta_K} \inf_{u \in \mathbb{S}^{d-1}} \frac{1}{2} \sum_{i=1}^{K} w_i \left(-u^\top z_i\right)_+^2. \tag{2.22}$$

By Proposition 2.4.1, one may choose $w^F(\theta)$ supported on the vertices of

$$\operatorname{Conv}(\theta_1 - \gamma, \dots, \theta_K - \gamma).$$

This is the higher-dimensional analogue of choosing the sampled minimum or sampled maximum in the one-dimensional Thompson-CHM rule.

Thus the algorithm does not choose an arbitrary vertex of the posterior convex hull. It chooses according to the posterior instance's own optimal design. This is important because, as the regular hexagon example shows, the statistically relevant vertices need not coincide with one preselected Carathéodory simplex.

The procedure below is the literal high-dimensional analogue of Thompson-CHM. The posterior conditioning event remains feasibility, while the binary top-two choice is replaced by a categorical draw from $w^F(\theta_t)$. The stopping rule is unchanged from Section 2.6, all geometric novelty is contained in the sampling distribution.

Algorithm 3 HD-Thompson-CHM

Input: target γ, confidence δ, threshold $\beta(t, \delta)$

Pull each arm once.
for $t = K, K+1, \dots$ **do**
  Compute $Z(t)$ from (2.13) or (2.14).
  if $Z(t) \geq \beta(t, \delta)$ **then**
    Stop and return feasible iff $0 \in \operatorname{Conv}(\widehat{\mu}_1(t) - \gamma, \dots, \widehat{\mu}_K(t) - \gamma)$.
  end if
  Draw

$$\theta_t \sim \Pi_{t-1}\left(\cdot \,|\, \gamma \in \operatorname{Conv}(\theta_1, \dots, \theta_K)\right).$$

  Compute $w^F(\theta_t)$ from (2.22).
  Draw $A_t \sim \text{Categorical}(w^F(\theta_t))$, and pull arm A_t.
end for

The feasible case is the clean part of the Thompson analysis. Since the true instance lies in the interior of the feasible region, conditioning the posterior on feasibility is asymptotically

harmless: the conditioning event has posterior probability tending to one. The algorithm therefore behaves like a plug-in rule using the feasible-game optimizer at the true parameter.

Proposition 2.8.1 (Feasible-case consistency of HD-Thompson-CHM)**.** *Assume*

$$0 \in \operatorname{int} \operatorname{Conv}(x_1, \ldots, x_K).$$

Assume further that the optimizer $w^F(\theta)$ in (2.22) is unique and continuous at $\theta = \mu$. Then HD-Thompson-CHM satisfies

$$\frac{N(t)}{t} \to w^F(\mu) \qquad \textit{almost surely.}$$

Consequently, with the GLR stopping rule, HD-Thompson-CHM is asymptotically optimal in the feasible case.

2.8.1 The infeasible Thompson problem

In the one-dimensional infeasible case, a conditioned feasible posterior sample is produced by an arm crossing the threshold. The rare event is therefore indexed by arms. In higher dimensions, the rare event is produced by a convex combination reaching the target, and the indexing object is the barycentric vector q. This shift from arm-indexed crossings to certificate-indexed crossings is the main technical obstacle.

The infeasible case is the genuinely new higher-dimensional difficulty. In one dimension, conditioning on feasibility forces one arm to cross the threshold. In higher dimensions, a feasible posterior draw may be produced by a joint movement of several arms:

$$\sum_i q_i(\theta_i - \gamma) = 0.$$

Thus rare feasible posterior samples are indexed by barycentric certificates q, not only by arms.

The following conditional theorem isolates the missing geometric ingredient. We state the result conditionally because it identifies exactly what remains to be proved for a complete Thompson theory. The required condition is a large-deviation gap: if an arm has already been sampled more often than its target allocation, then feasible posterior samples that would select that arm must be exponentially less likely than the dominant feasible posterior samples. This is the higher-dimensional counterpart of the over-allocation argument used in the one-dimensional proof and in related Thompson-sampling analyses [Russo, 2016, Kaufmann et al., 2018].

Let

$$g_i(\theta) = w_i^F(\theta)$$

be the categorical selection function in HD-Thompson-CHM. For an allocation vector $\alpha \in \Delta_K$, define the posterior large-deviation rate

$$R_\alpha(\theta) = \frac{1}{2} \sum_{i=1}^{K} \alpha_i \left\| \mu_i - \theta_i \right\|^2 . \tag{2.23}$$

Let

$$\mathcal{F} = \{\theta : \gamma \in \mathrm{Conv}(\theta_1, \ldots, \theta_K)\} .$$

For each arm i, define

$$\mathcal{F}_i = \{\theta \in \mathcal{F} : g_i(\theta) > 0\} .$$

The first assumption is a posterior large-deviation principle with the empirical allocation as the rate multiplier. For Gaussian arms, this is the natural quadratic analogue of the KL large-deviation rates appearing in exponential-family bandit identification. It says that, on the exponential scale, the posterior mass of a set is governed by the cheapest way to move the true means into that set under the current sampling proportions.

Assumption 11. For every compact set $B \subseteq (\mathbb{R}^d)^K$, uniformly over allocation vectors α in compact subsets of Δ_K,

$$\Pi_t(B) \doteq \exp\left(-t \inf_{\theta \in B} R_{\bar{\psi}(t)}(\theta) \right), \tag{2.24}$$

where

$$\bar{\psi}(t) = \frac{1}{t} \sum_{s=1}^{t} \psi(s),$$

and $a_t \doteq b_t$ means

$$\frac{1}{t} \log(a_t / b_t) \to 0.$$

The second assumption is the geometric heart of the idea. The set $\mathcal{F}_i$ consists of feasible posterior instances for which the categorical rule would assign positive probability to arm i. The gap condition says that, once arm i is over-allocated, the most likely feasible posterior samples no longer belong to this active set. In other words, over-sampling an arm makes all feasible certificates that continue to use that arm exponentially unattractive.

Assumption 12. Let $w^*(\mu)$ be the unique optimizer of the infeasible game (2.5). For every arm i and every $\varepsilon > 0$, there exists $c = c(i, \varepsilon) > 0$ such that, for every $\alpha \in \Delta_K$ satisfying

$$\alpha_i \geq w_i^*(\mu) + \varepsilon,$$

one has

$$\inf_{\theta \in \mathrm{cl}(\mathcal{F}_i)} R_\alpha(\theta) \geq \inf_{\theta \in \mathcal{F}} R_\alpha(\theta) + c. \tag{2.25}$$

The next lemma is the allocation mechanism behind the Thompson proof. It is a Russo-type over-allocation criterion: to prove convergence of the empirical allocation, it is enough to show that every arm receives only summably many conditional pulls after it has exceeded its target proportion by a fixed margin [Russo, 2016]. The lemma separates the probabilistic large-deviation estimate from the deterministic consequence for sampling proportions.

Lemma 2.8.2. *Let* $\psi_i(t) = \mathbb{P}(A_t = i \mid \mathcal{F}_{t-1})$ *and*

$$\bar{\psi}_i(t) = \frac{1}{t} \sum_{s=1}^{t} \psi_i(s).$$

If for every arm i *and every* $\varepsilon > 0$*,*

$$\sum_{t=1}^{\infty} \psi_i(t) \mathbf{1}\{\bar{\psi}_i(t) \geq w_i^*(\mu) + \varepsilon\} < \infty, \tag{2.26}$$

then

$$\bar{\psi}(t) \to w^*(\mu),$$

and hence

$$\frac{N(t)}{t} \to w^*(\mu) \qquad \textit{almost surely.}$$

The conditional theorem now follows the same logical structure as the one-dimensional Thompson-CHM analysis. Posterior large deviations convert the active-certificate gap into exponential decay of over-allocated pulls. The Russo-type criterion converts this decay into allocation convergence. The GLR stopping rule then converts allocation convergence into asymptotic optimality.

Theorem 13. *Assume the true instance is infeasible:*

$$0 \notin \mathrm{Conv}(x_1, \ldots, x_K).$$

Assume that the infeasible optimizer $w^*(\mu)$ *is unique, and that Assumptions 11 and 12 hold. Then HD-Thompson-CHM satisfies*

$$\frac{N(t)}{t} \to w^*(\mu) \qquad \textit{almost surely.}$$

Consequently, with the GLR stopping rule, HD-Thompson-CHM is asymptotically optimal in the infeasible case.

We record the remaining geometric statement as a conjecture. It is deliberately phrased at the level of certificates rather than algorithms: the issue is not the stopping rule, nor the form of the feasible-game optimizer, but the comparison of posterior rates among different barycentric ways of making an infeasible instance appear feasible.

Conjecture 14 (Geometric gap for the infeasible Thompson rule). For regular Gaussian infeasible CHM instances with unique saddle point, the active-certificate gap in Assumption 12 holds for the selection function $g_i = w_i^F$ induced by the feasible separation game.

Remark 2.8.3. The conjecture is the precise higher-dimensional analogue of the key large-deviation comparison in the one-dimensional proof. In one dimension, a rare feasible posterior draw is indexed by the identity of the arm crossing the threshold, and the gap follows from the explicit inverse-KL allocation. In higher dimensions, rare feasible posterior draws are indexed by barycentric certificates q. The conjecture asserts that, when an arm has been over-allocated relative to the saddle point of the infeasible game, all feasible posterior certificates that would cause the Thompson rule to select that arm become exponentially less likely than the dominant feasible certificate.

2.9 Discussion

We close this chapter by summarizing the conceptual message of the higher-dimensional extension. The one-dimensional theory is not wrong in higher dimension, rather, it must be reinterpreted through the two dual languages of convex geometry. Extremes become vertices and supporting directions. Threshold crossings become barycentric certificates. The lower-bound game remains the organizing principle, but its explicit solution is no longer a scalar inverse-KL calculation. The higher-dimensional problem preserves the qualitative dichotomy of the one-dimensional CHM problem but changes its geometry.

In the feasible case, the relevant alternatives are separating hyperplanes. This leads to the separation game

$$\sup_{w\in\Delta_K} \inf_{u\in\mathbb{S}^{d-1}} \frac{1}{2}\sum_i w_i \left(-u^\top x_i\right)_+^2 .$$

Non-vertex arms are dominated, so an optimal allocation exists on the vertices of the convex hull. However, Carathéodory's theorem does not imply that the optimal allocation is supported on only $d+1$ arms. A small simplex may certify feasibility, but one must gather evidence against all separating hyperplanes. The regular hexagon example shows that the optimal design may use strictly more than $d+1$ vertices.

In the infeasible case, the relevant alternatives are feasible convex combinations. The

exact game is

$$\sup_{w\in\Delta_K}\inf_{q\in\Delta_K}\frac{\left\|\sum_i q_i x_i\right\|^2}{2\sum_i q_i^2/w_i}.$$

The adversary's certificate can be chosen with support at most $d+1$, but one must sample every arm. This is the higher-dimensional analogue of the one-dimensional inverse-KL allocation, with the important difference that the least confusing alternative may move several arms jointly.

The GLR stopping rule based on convex and separating certificates gives a rigorous high-dimensional analogue of the three-case one-dimensional stopping rule. Combined with tracking of the exact saddle-point allocation, it yields an asymptotically optimal high-dimensional algorithm under standard uniqueness and continuity assumptions. The Thompson-CHM extension is equally natural: condition the posterior on feasibility, then sample according to the feasible-game optimizer of the posterior draw. Its feasible-case optimality follows directly. The infeasible case reduces to a geometric large-deviation gap over active barycentric certificates; proving this gap appears to be the central remaining step toward a full Thompson-sampling theory in dimensions $d \geq 2$.

This stochastic convex hull membership problem leads to important and interesting real world applications. For example, in quant finance, one way to view the problem is as testing whether an observed price, return, or strategy profile lies inside an attainable set generated by a collection of uncertain expected-return distributions. In that sense, the framework can be thought of as a statistically principled way to reason about whether an observed pattern is explainable by an existing opportunity set, or whether it may reflect a potential inefficiency that may point to a possible arbitrage opportunity. More broadly, the topic is also about adaptive sampling and efficient allocation of research effort when one is trying to resolve an uncertain but economically meaningful decision as quickly as possible.

From a broader perspective, the chapter suggests a useful plan for future work. First, one may seek sharper structural descriptions of the saddle points in (2.4) and (2.5), especially conditions under which the feasible optimizer is supported on a small number of vertices or, conversely, must spread over many vertices. Second, one may develop efficient numerical procedures for the two games, using either explicit best responses or no-regret dynamics. Third, and most directly connected to Thompson-CHM, one may prove the active-certificate gap by analyzing the geometry of least-cost barycentric perturbations. A proof of this final statement would turn the conditional Thompson theorem into a complete higher-dimensional analogue of the one-dimensional result.

CHAPTER 3

A Oneshot Differentially Private NIHT Algorithm for High–Dimensional Sparse Linear Regression

Iterative Hard Thresholding (IHT) is a simple yet powerful routine for obtaining sparse coefficients in linear regression. In this chapter we consider the *differentially private* version of IHT. We introduce a one-shot mechanism that injects calibrated Laplace noise *once per iterate* into the magnitude vector, selects the top s coordinates, and then hard–thresholds after a gradient step. Compared to the existing private IHT algorithms that peel off the maximum element at each iteration, our approach selects the top coordinates in a oneshot style, which is much more computational efficient. We refer to the resulting procedure as **Oneshot DP-NIHT with Sparsifier**. The algorithm keeps its output *exactly* s–sparse without requiring oracle knowledge of s. Theoretically we establish (ε, δ)–differential privacy and show that the estimator achieves the *minimax parameter-error rate up to logarithmic factors*. We also provide a matching excess-risk bound and illustrate the empirical performance of Oneshot DP-NIHT with Sparsifier on synthetic data.

3.1 Introduction

Feature selection is a fundamental topic in pattern recognition, statistics and machine learning, which aims to find a subset of the most useful features from the given feature set. A common challenge in this domain is when the sample size is much smaller than the number of features, i.e. $p \gg n$. A large body of work has developed tools for feature selection with large p in the literature, most prominently the Lasso Tibshirani [1996], Yuan and Lin [2006], Zou [2006], Meinshausen [2007], Wang et al. [2007], Chen et al. [2013], the elastic net Zou and Hastie [2005], least angle regressionEfron et al. [2004], and the Dantzig Selector Candes and Tao [2007].

Although feature-selection methods are widely used in real-world data analysis, modern deployments increasingly involve sensitive covariates such as genomics, mobility traces, clinical logs, and financial records; in such settings, releasing even a support set can leak

personal information. In this context, the successful privacy-preserving framework differential privacy Dwork et al. [2006a,c] provides a principled safeguard: it ensures that any single sample has limited influence on the released output, is preserved under post–processing, and composes across multiple releases.

Differentially private feature selection has been widely studied Talwar et al. [2015], Cai et al. [2021], Dandekar et al. [2018], Zhang et al. [2012], Thakurta and Smith [2013], Li and Lu [2024], nevertheless, a rigorous exploration of a differentially private procedure for the high-dimensional sparse regression has received little attention in the literature. While several differentially private ERM algorithms exist Chaudhuri et al. [2011], Bassily et al. [2014], most produce *dense* outputs and rely on generic gradient perturbation, incurring privacy noise that scales with the ambient dimension d. In high dimensions this leads to sub-optimal error or calls for post-hoc sparsification that is not privacy-accounted. Our goal is to design an efficient *explicitly sparse*, high-dimensional DP estimator whose noise depends only on the target sparsity s.

In this chapter, we introduce Oneshot DP-NIHT with Sparsifier, a differentially private, hard–thresholding based procedure. At a high level, Oneshot DP-NIHT with Sparsifier runs a noisy iterative hard–thresholding oracle that applies an efficient *one–shot* top-s mechanism to the *magnitudes* of the gradient-updated iterate, and projects onto an ℓ_1 ball and retains the top coordinates. This design keeps the output sparse by construction and avoids the well-known tendency of generic DP solvers for linear programs to yield dense solutions. On the theory side, we rigorously establish the privacy results of the procedure and provide finite-sample guarantees. Our main statistical result is a finite-sample *parameter-error* bound that is *minimax-optimal up to logarithmic factors* under standard RE/RSC conditions and sub-Gaussian noise, matching the (ε, δ)-DP lower bound of Cai et al. [2021]. In parallel, we establish a nonasymptotic *population excess-risk* bound with the expected structure: the non-private fast rate plus an explicit privacy penalty that decays as n^{-2} and scales linearly with the target sparsity.

We remark that a practical challenge in sparse feature selection is that the target sparsity is rarely known in advance. A common approach is to estimate the sparsity using cross validation, which can be computationally expensive in large data sets. To address this challenge, motivated by Khanna et al. [2023], we privatize the model sparsity via a clipped count with double–geometric noise (a "Sparsifier") and apply the privatized sparsity to the hard-thresholding oracle. The Sparsifier is simple to implement, in the utility theory below its role is privacy accounting, while the estimation bound is stated conditional on the released sparsity lying in the overspecified regime required for IHT contraction. Empirically, we show that adding the "Sparsifier" generally outperforms the hard-thresholding oracle without this

step.

We summarize and highlight our main contributions below.

- We introduce a novel one-shot hard–thresholding based algorithm. The algorithm privately estimates the sparsity level to avoid relying on oracle sparsity information and returns a sparse output. We also provide the privacy results for the algorithm.
- Under standard conditions, we prove a non-asymptotic parameter error bound that matches the known (ε, δ)–DP minimax rate for high-dimensional sparse regression up to logarithmic factors.
- We establish a population risk bound that matches a non-private fast rate up to the provable privacy penalty and a vanishing optimization tail.

3.2 Related Work

Differential privacy for ERM and regression. Differentially private empirical risk minimization (ERM) traces back to objective/noise-perturbation schemes Chaudhuri et al. [2011], with subsequent advances yielding tighter rates and efficient algorithms for smooth losses Bassily et al. [2014], Kifer et al. [2012], Abadi et al. [2016]. In high-dimensional linear models, a central theme is quantifying the *cost of privacy* relative to the non-private minimax rates. For sparse regression, the benchmark lower bounds showing that privacy induces an additive penalty on the order of $(s \log d)^2/(n^2\varepsilon^2)$ (up to logarithmic factors) were established by Cai et al. [2021]. Most DP-ERM procedures, however, produce *dense* outputs or require post hoc thresholding that is not aligned with their privacy accounting. By contrast, Oneshot DP-NIHT with Sparsifier maintains sparsity at each iterate and calibrates privacy noise exactly at the support-selection step.

Private feature selection and private top-k. Closer to support recovery, there is a line of work on differentially private selection and filtering, including the sparse-vector technique and its refinements, as well as mechanisms for privately selecting (and releasing) the top coordinates of a vector. Of particular relevance is the *one-shot* differentially private top-k mechanism Qiao et al. [2021], which adds a single Laplace vector for support selection and *fresh* value noise on the chosen coordinates, achieving (ε, δ)-DP with a sharp scale. Peeling-based private hard-thresholding procedures form a complementary approach by iteratively selecting small batches. Oneshot DP-NIHT with Sparsifier intentionally avoids peeling in favor of a single-shot calibration that composes cleanly across iterations and preserves sparsity throughout.

3.3 Preliminaries

Differential privacy In modern data analysis, differential privacy Dwork et al. [2006a,b] provides quantifiable guarantees that the outcome of a computation reveals minimal information about any single individual's data. It ensures that the output of a randomized algorithm keeps almost unchanged when a single entry in the dataset is modified. We start with some basic concepts in differential privacy.

Definition 3.3.1. Data sets $\mathcal{D}, \mathcal{D}'$ are said to be *neighbors*, or *adjacent*, if they differ in exactly one record

Consider neighboring datasets $\mathcal{D}, \mathcal{D}' \in \mathcal{Z}^n$, each containing n data points from the sample space $\mathcal{Z}$. A randomized algorithm $\mathcal{M}$ takes a dataset as input, and differential privacy requires the output distributions under $\mathcal{D}$ and $\mathcal{D}'$ to be nearly indistinguishable.

Definition 3.3.2 (Differential privacy Dwork et al. [2006a,b])**.** A randomized mechanism $\mathcal{M}$ is (ε, δ)-differentially private if for all adjacent $\mathcal{D}, \mathcal{D}'$, and for any subset of possible outputs S: $\mathbb{P}(\mathcal{M}(\mathcal{D}) \in S) \leq \mathrm{e}^{\varepsilon}\mathbb{P}(\mathcal{M}(\mathcal{D}') \in S) + \delta$. Pure differential privacy is the special case of approximate differential privacy in which $\delta = 0$.

Definition 3.3.3 (Sensitivity)**.** Let $f = (f_1, \cdots, f_m)$ be m real valued functions that takes a database as input. The sensitivity of f, denoted as s, is defined as

$$s_f = \max_{\mathcal{D},\mathcal{D}'} \max_{1 \leq i \leq m} |f_i(\mathcal{D}) - f_i(\mathcal{D}')|,$$

where the maximum is taken over any adjacent data sets $\mathcal{D}$ and $\mathcal{D}'$.

We use the following notation throughout. We observe $(x_i, y_i) \in \mathbb{R}^d \times \mathbb{R}$ for $i = 1, \ldots, n$, and the design matrix $X \in \mathbb{R}^{n \times d}$ is the collection of rows $x_i^\top$, and $y \in \mathbb{R}^n$ denotes the response. For $v \in \mathbb{R}^d$, we use $\|v\|_p$ to denote its ℓ_p norms. We use $[d] := \{1, \ldots, d\}$, denote the support of v by $\mathrm{supp}(v)$, and for $T \subseteq [d]$ let v_T be the vector such that v_T matches v on indices in T and 0 elsewhere. For $C > 0$, $\Pi_C(\cdot)$ denotes the Euclidean projection onto the ℓ_1 ball $\{\theta : \|\theta\|_1 \leq C\}$. The hard-thresholding operator at a set T is $H_T(v) := (v_T, 0_{T^c})$, its size-s version $H_s(v)$ keeps the s largest entries of $|v|$ and zeros the rest. Finally, the clipping map $\mathrm{clip}_R : \mathbb{R} \to \mathbb{R}$ is given by $\mathrm{clip}_R(t) := \mathrm{sign}(t) \min\{|t|, R\}$, and clip_R on a vector means applying the clipping map coordinate-wise to the vector.

3.4 Oneshot DP-NIHT Algorithm

In this section, we introduce Oneshot DP-NIHT, an (ε, δ)-DP hard-thresholding procedure motivated by high-dimensional sparse regression. We start from a one-shot differentially private algorithm for top-k selection Qiao et al. [2021] as an oracle. Next, we show how we can derive a one-shot version of differentially private iterative hard thresholding (IHT) algorithm, which we refer to as Oneshot DP-NIHT. Taking Oneshot DP-NIHT as a key step, we build the Oneshot DP-NIHT with Sparsifier. Finally, we provide privacy results for both Oneshot DP-NIHT and Oneshot DP-NIHT with Sparsifier.

3.4.1 Oneshot DP top-k algorithm

Qiao et al. [2021] proposed an efficient one-shot algorithm for privately selecting the top-k elements. By adding a single i.i.d. Laplace vector $g \sim \mathrm{Lap}(\lambda)^d$ to z, select the indices of the k largest coordinates of $z + g$, and then add freshly sampled, independent Laplace noise to the values released on the selected support. In Qiao et al. [2021] it is proved that, by choosing

$$\lambda \gtrsim s_f \, \frac{\sqrt{k \, \log(d/\delta)}}{\varepsilon}$$

the algorithm ensures (ε, δ)-DP for the selection-and-release in a single shot. This mechanism serves as a building block of the algorithms we will propose: we use this mechanism inside each iteration, applying it to the magnitudes of the gradient-updated vector. We denote the output of the oneshot top-k algorithm applied to a vector $v \in \mathbb{R}^d$ as $\mathrm{TOP}_k(v; \varepsilon, \delta, \lambda)$, where ε, δ are privacy parameters and λ is the parameter of the Laplace noises in the algorithm. All details and theoretical results for the oneshot top-k algorithm are deferred to the appendix (see also the discussions and technical details of Algorithm 1 and Theorem 2.2 in Qiao et al. [2021]).

3.4.2 Oneshot DP-NIHT algorithm and the Sparsifier

In this section, we describe the privatized hard–thresholding oracle Oneshot DP-NIHT and how it builds into Oneshot DP-NIHT with Sparsifier that releases a sparsity level before calling the oracle. We assume coordinate-wise bounded design $\|x_i\|_\infty \le x_{\max}$ and project all iterates onto the ℓ_1 ball of radius C. We first introduce the Oneshot DP-NIHT algorithm. Given $s \in \{1, \ldots, d\}$, define the top-s index set of v by

$$S_s(v) \;\in\; \arg \max_{T \subseteq [d],\ |T| = s} \|v_T\|_2 \quad \text{(ties broken arbitrarily)}.$$

The hard-thresholding map is $H_s(v) := H_{S_s(|v|)}(v)$, i.e., selection by magnitude with signed values retained. We will apply the oneshot DP top-s mechanism to $|v|$ with scale λ and then add freshly sampled noise to the reported values on the selected support, this is the only privacy-consuming step per iteration. Finally, we project onto the ℓ_1 ball: $\theta \mapsto \Pi_C(\theta)$. As a post-processing step, the projection incurs no privacy cost.

Algorithm 4 Oneshot DP-NIHT (Oneshot Noisy Iterative Hard Thresholding)

Input: Data X, y, sparsity s, privacy level ε, δ, number of iterations M, truncation level R, noise scale rule B, step size $\eta > 0$, projection level C.

Initialize $\theta^{(0)} = \mathbf{0} \in \mathbb{R}^d$.

for $m = 0, \ldots, M-1$ **do**

$\theta^{(m+0.5)} = \theta^{(m)} - \eta \left(\nabla L(\theta^{(m)} \mid X, \mathrm{clip}_R(y)) \right)$.

$\tilde{\theta}^{(m+1)} = \mathrm{TOP}_s \left(\theta^{(m+0.5)};\ \varepsilon/M,\ \delta/M,\ \eta B/n \right)$, where TOP_s selects the support using $|\theta^{(m+0.5)}|$ and returns the signed selected coordinates with fresh value noise.

Euclidean project onto $\|\cdot\|_1 \le C$:

$$\theta^{(m+1)} \ \leftarrow\ \Pi_C\left(\tilde{\theta}^{(m+1)} \right).$$

end for

return $\hat{\theta} = \theta^{(M)}$.

In Oneshot DP-NIHT, each iteration takes a gradient step using $y_i^{\mathrm{clip}} = \mathrm{clip}_R(y_i)$ to control sensitivity, then performs a oneshot private top-s selection on the *magnitudes* of the pre-thresholded iterate, and finally projects back to $\{\|\theta\|_1 \le C\}$. The oneshot operator adds a single i.i.d. Laplace vector to select a support and then adds fresh independent Laplace noise to the reported values on that support.

To formulate a differentially private algorithm for the high-dimensional sparse regression, a practical difficulty is that most algorithms take the sparsity s as prior knowledge and use it directly. This is not available in many applications. To address this issue, motivated by the *Sparsifier* in Khanna et al. [2023], we add a private sparsity-estimation step. The *Sparsifier* releases a privatized sparsity level via clipped counting plus double-geometric noise. Oneshot DP-NIHT with Sparsifier first releases a privatized sparsity $\tilde{c}$ (via a clipped count with double-geometric noise), runs Oneshot DP-NIHT with $s = \tilde{c}$, and post-processes by keeping the $\tilde{c}$ largest entries in magnitude.

We choose a clipping level $R > 0$ and a projection radius $C \ge \|\theta^\star\|_1$, and we set a step size $\eta \in (0, 2/L_k]$ at the appropriate sparsity level k. These choices enforce a uniform ℓ_∞ sensitivity bound for the vector fed to the oneshot oracle and yield the restricted contraction inequalities used later in the error and risk analyses.

Algorithm 5 Oneshot DP-NIHT with Sparsifier

Input: Data X, y, privacy level $\varepsilon_1, \varepsilon_2, \delta$, number of iterations M, truncation level R, noise scale rule B, step size $\eta > 0$, projection level C, clipping parameters α, β, precision parameter $\rho > 0$.

$\widehat{w}_{NP} \leftarrow$ Nonprivate pilot estimator
$c \leftarrow$ Count Nonzero Components($\widehat{w}_{NP}$)
$c \leftarrow \mathrm{Clip}(c, \alpha, \beta)$
$c \leftarrow c + Z$, where $\mathbb{P}(Z = z) \propto \exp\{-\varepsilon_1|z|/(\beta - \alpha)\}$ for $z \in \mathbb{Z}$
$c \leftarrow \mathrm{Clip}(c, \alpha, \beta)$
$c \leftarrow \mathrm{round}(c)$
$c \leftarrow \mathrm{Clip}(c \cdot \rho, 0, d)$
$c \leftarrow \mathrm{round}(c)$
Choose $B = B(c)$ satisfying (3.2) with $s = c$, or use a fixed B calibrated with $s_{\max} = \min\{d, \lceil \rho\beta \rceil\}$.
$\widehat{\theta}_P \leftarrow$ Oneshot DP-NIHT$(X, y, c, \varepsilon_2, \delta, M, R, B, \eta, C)$
$\widehat{\theta} \leftarrow$ Keep Nonzero$(\widehat{\theta}_P, c)$, here Keep Nonzero operation keeps the c largest entries in magnitude and zeros the rest
Output $\widehat{\theta}$

3.4.3 Privacy guarantees

The oneshot top-k mechanism operates under a per-coordinate ℓ_∞ sensitivity bound. For the pre-thresholded vector $v^{(m)} = \theta^{(m)} - \eta \nabla L\Big(\theta^{(m)} \;\Big|\; X, \mathrm{clip}_R(y)\Big)$, replace-one adjacency yields

$$\|v^{(m)}(D) - v^{(m)}(D')\|_\infty \;\le\; \frac{2\eta}{n}\, x_{\max}\Big(R + x_{\max}C\Big) \;=:\; s_\infty. \tag{3.1}$$

We allocate per-iteration budgets $(\varepsilon_2/M, \delta/M)$ and use $\lambda = \eta B/n$ for both the selection and the fresh value noise. The calibration of B below is precisely what is required by the one-shot DP theorem when applied to $|v^{(m)}|$. First, we present the privacy result for Algorithm 4 and provide a proof sketch. We leave the full proof to the appendix.

Theorem 15 (Privacy of Oneshot DP-NIHT)**.** *Let $M \ge 1$, $\varepsilon_2 > 0$, $\delta \in (0, 1)$, and we assume $\varepsilon_2/M \le 0.2$, $\delta/M \le 0.05$ and $\|x_i\|_\infty \le x_{\max}$. If the Laplace scale is $\lambda = \eta B/n$ with*

$$B \;\ge\; \frac{16\, M\, x_{\max}\Big(R + x_{\max}C\Big)\sqrt{s\,\log(dM/\delta)}}{\varepsilon_2}, \tag{3.2}$$

then Algorithm 4 is (ε_2, δ)-differentially private.

Proof sketch. By (3.1), the input to the selection step has per-coordinate sensitivity s_∞. Adding i.i.d. Lap(λ) to $|v^{(m)}|$, selecting the top s, and then adding freshly sampled Laplace noise Lap(λ) to the reported values on the selected coordinates yields $(\varepsilon_2/M, \delta/M)$-DP for

iteration m when λ meets the oneshot condition, and (3.2) is equivalent to that condition with $\lambda = \eta B/n$. Notice that the projection onto the ℓ_1 ball is a post-processing step, therefore basic composition over the M iterations ensures (ε_2, δ)-DP.

We continue to state the following theorem for the privacy guarantee of Algorithm 5.

Theorem 16 (Privacy of Oneshot DP-NIHT with Sparsifier)**.** *Algorithm 5 is $(\varepsilon_1 + \varepsilon_2, \delta)$-differentially private, provided that the Oneshot DP-NIHT noise scale is calibrated either to the realized private sparsity c after the private sparsity release, or to the deterministic upper bound $s_{\max} := \min\{d, \lceil \rho\beta \rceil\}$.*

Note that by (3.1), the clipping of y and the ℓ_1 projection together bound the per-iteration sensitivity by s_∞, enabling a single Laplace draw to determine the support. The requirement of fresh value noise is essential for the privacy guarantee and the one-shot proof and is the only place where the value-level noise appears, all subsequent operations are post-processing. Algorithm 5 exposes only $\tilde{c}$ and the private estimate. Choosing $\rho \geq 1$ in Algorithm 5 avoids underspecifying the sparsity relative to the contraction regime used later in the utility analysis. The proofs and technical details of Theorem 15 and Theorem 16 are left to the Appendix.

3.5 Error bound for Oneshot DP-NIHT with Sparsifier

In this section, we study the theoretical properties of Algorithm 5. The first fundamental result of this chapter is that Algorithm 5 achieves a minimax-optimal error bound up to logarithmic factors. Some key quantities and conditions are needed to establish the result. We use the conventional restricted eigenvalue constants of $\Sigma := \frac{1}{n} X^\top X$:

$$\underline{\kappa}(k) := \min_{\substack{h \neq 0 \\ \|h\|_0 \leq k}} \frac{h^\top \Sigma h}{\|h\|_2^2}, \qquad \overline{\kappa}(k) := \max_{\substack{h \neq 0 \\ \|h\|_0 \leq k}} \frac{h^\top \Sigma h}{\|h\|_2^2}, \qquad \kappa(k) := \frac{\overline{\kappa}(k)}{\underline{\kappa}(k)}.$$

and assume the following standard conditions:

(A1) **Restricted strong convexity/smoothness.** $\underline{\kappa}(k_0) > 0$ and $\overline{\kappa}(k_0) < \infty$ for $k_0 = 2s + s^\star$.

(A2) **Gradient at the truth is sub-Gaussian coordinatewise.** There is $\bar{g} > 0$ such that $\|\nabla L(\theta^\star)\|_\infty \leq \bar{g}$ with probability at least $1 - \alpha_1$, where L is the clipped empirical loss used by the algorithm. For the un-clipped linear model with sub-Gaussian noise and $\|x_i\|_\infty \leq x_{\max}$, the usual choice $\bar{g} \lesssim \sigma x_{\max} \sqrt{(2\log(2d/\alpha_1))/n}$ is valid on the event

that no response is clipped; equivalently, the probability of clipping is included in the failure probability.

(A3) **Adequate sparsity fed to Oneshot DP-NIHT.** The s used by the algorithm satisfies $s \geq s^\star$ and $s \geq C_0\,\kappa(k_0)^2\,s^\star$ for a universal C_0 (the usual overspecification condition ensuring contraction).

(A4) **Privacy scales.** $\lambda = \eta B/n$ with B satisfying the oneshot calibration (as stated in the privacy section), and the per-iteration budgets lie in the regime required by the oneshot top-k theorem. If the sparsity is random as in Algorithm 5, this calibration is imposed conditionally on the realized private sparsity or uniformly over $s \leq s_{\max}$.

We write $b_{\max}$ for any deterministic upper bound on $\|\theta^\star\|_1$; one may take $b_{\max} = C$.

We work on the intersection of the following events and union-bound at the end.

$$\begin{aligned}
\mathcal{E}_\nabla &:= \{\|\nabla L(\theta^\star)\|_\infty \leq \bar{g}\}\,,\\
\mathcal{E}_{\text{sel}} &:= \left\{\max_{0\leq m<M} \left\|g^{(m)}\right\|_\infty \leq \lambda\,\tau\right\} \quad \text{with} \quad \tau := \log\frac{2dM}{\alpha_2},\\
\mathcal{E}_{\text{val}} &:= \left\{\max_{0\leq m<M} \left\|\xi^{(m)}_{T_{m+1}}\right\|_2^2 \leq c_\xi\, s\lambda^2\,\tau_\xi\right\} \quad \text{with} \quad \tau_\xi := \log^2\frac{2sM}{\alpha_3},\\
\mathcal{E}_{\text{RSC}} &:= \{\text{(A1) holds at level } k_0 = 2s + s^\star\},\\
\mathcal{E}_c &:= \{\text{(A3) holds}\}.
\end{aligned}$$

Here $c_\xi > 0$ is a universal constant. The Laplace tail $\mathbb{P}(|\text{Lap}(\lambda)| > t) = e^{-t/\lambda}$ gives $\mathbb{P}(\mathcal{E}_{\text{sel}}^c) \leq \alpha_2$ by a union bound over the dM selection-noise coordinates. Conditional on the selected supports, the released value noises are independent Laplace variables; hence another union bound gives

$$\mathbb{P}\Big(\max_{0\leq m<M} \left\|\xi^{(m)}_{T_{m+1}}\right\|_2^2 > c_\xi s\lambda^2 \log^2\frac{2sM}{\alpha_3}\Big) \leq \alpha_3.$$

This is the only concentration statement needed for the squared Laplace value noise. In particular, we do not use a sub-exponential claim for ξ^2. Finally, let $\alpha := \alpha_1 + \alpha_2 + \alpha_3$. The stochastic bounds above give failure probability at most α for $\mathcal{E}_\nabla$, $\mathcal{E}_{\text{sel}}$, and $\mathcal{E}_{\text{val}}$. The restricted-eigenvalue event $\mathcal{E}_{\text{RSC}}$ and the sparsity-adequacy event $\mathcal{E}_c$ are treated as conditioning events unless verified separately. Thus, all upcoming statements hold on $\mathcal{E} := \mathcal{E}_\nabla \cap \mathcal{E}_{\text{sel}} \cap \mathcal{E}_{\text{val}} \cap \mathcal{E}_{\text{RSC}} \cap \mathcal{E}_c$; if $\mathbb{P}(\mathcal{E}_{\text{RSC}}^c) \leq \alpha_{\text{RSC}}$ and $\mathbb{P}(\mathcal{E}_c^c) \leq \alpha_c$, then the final failure probability is enlarged by $\alpha_{\text{RSC}} + \alpha_c$. Throughout this subsection we abbreviate $\kappa = \kappa(k_0)$ and recall that $\lambda = \eta B/n$. We also keep the auxiliary quantities $\tau > 0$ and $\tau_\xi > 0$ generic inside the theorem (they will be instantiated immediately after the proof). We now state the parameter-error bound for the output of Algorithm 5.

Theorem 17. *Assume* (A1)–(A4) *and condition on* $\mathcal{E}_{\mathrm{RSC}} \cap \mathcal{E}_c$. *Let* $\eta = \frac{1}{2\,\overline{\kappa}(k_0)}$ *and* $\lambda = \eta B/n$. *If* $M \geq c_0\,\kappa(k_0)\,\log\!\big(b_{\max}^2 n\big)$ *for a universal constant* $c_0 > 0$, *then, with probability at least* $1-\alpha$ *over the gradient and privacy-noise events, the* ONESHOT DP-NIHT WITH SPARSIFIER *output* $\widehat{\theta} = H_s(\theta^{(M)})$ *satisfies*

$$\left\|\widehat{\theta} - \theta^\star\right\|_2^2 \;\leq\; \frac{1}{n} \;+\; C_4\,\frac{(s+s^\star)\,\bar{g}^2}{\underline{\kappa}(k_0)^2} \;+\; C_5\,\kappa(k_0)^2\,s\,\lambda^2\,\tau^2 \;+\; C_6\,\kappa(k_0)^2\,s\,\lambda^2\,\tau_\xi\,, \tag{3.3}$$

for suitably large universal constants $C_4, C_5, C_6 > 0$.

Corollary 3.5.1. *On the event* $\mathcal{E}$ *we may take*

$$\tau \;=\; \log\!\Big(\tfrac{2dM}{\alpha}\Big), \qquad \tau_\xi \;=\; c_\tau\,\log^2\!\Big(\tfrac{2sM}{\alpha}\Big)$$

for a universal $c_\tau > 0$, *and set* $\lambda = \eta B/n = \frac{B}{2\,\overline{\kappa}(k_0)\,n}$. *With these substitutions and* $M \asymp \kappa(k_0)\log(b_{\max}^2 n)$, *the bound* (3.3) *becomes*

$$\left\|\widehat{\theta} - \theta^\star\right\|_2^2 \;\leq\; \frac{1}{n} \;+\; C_1\,\frac{(s+s^\star)\,\bar{g}^2}{\underline{\kappa}(k_0)^2} \;+\; \widetilde{C}\;\frac{\kappa(k_0)^2\,s\,B^2}{\overline{\kappa}(k_0)^2\,n^2}\left[\log^2\!\Big(\tfrac{2dM}{\alpha}\Big) + \log^2\!\Big(\tfrac{2sM}{\alpha}\Big)\right],$$

where $C_1, \widetilde{C} > 0$ *are universal constants. This is the same form as the usual noisy-IHT bound, with the one-shot selection and value-release errors collected in the final bracket.*

Remark 3.5.2. The rate in Theorem 17 is minimax optimal up to logarithmic factors after the privacy-calibrated scale is substituted. Under the standing RE/RSC assumptions, take $\eta \asymp 1/\overline{\kappa}(k_0)$, choose $R \gtrsim x_{\max}\|\theta^\star\|_1 + \sigma\sqrt{\log(n/\alpha)}$ so that response clipping is negligible with high probability, take $C \geq \|\theta^\star\|_1$, and choose the actual one-shot scale according to Theorem 15, namely

$$B \;\asymp\; \frac{M\,x_{\max}(R + x_{\max}C)\sqrt{s\log(dM/\delta)}}{\varepsilon_2}.$$

If $s \asymp s^\star$, the restricted condition number is bounded, and $x_{\max}(R + x_{\max}C) \lesssim \sigma\,\mathrm{polylog}^{1/2}(n, d, 1/\alpha)$, then (3.3) gives

$$\|\widehat{\theta} - \theta^\star\|_2^2 \;\lesssim\; \underbrace{\frac{\sigma^2 s^\star \log d}{n}}_{\text{statistical}} + \underbrace{\frac{\sigma^2 (s^\star \log d)^2}{n^2\varepsilon_2^2}\cdot \mathrm{polylog}(n, d, 1/\delta, 1/\alpha)}_{\text{privacy}},$$

with high probability. Passing to the Σ_x-norm gives the same rate up to the restricted eigenvalue constants. This matches the (ε, δ)-DP minimax lower bound presented as Theorem 4.3

in Cai et al. [2021] for high-dimensional regression, $\sigma^2\{s^\star \log d/n + (s^\star \log d)^2/(n^2\varepsilon_2^2)\}$, up to logarithmic factors. See also Theorem 4.1 and Appendix E.2 in Chakraborty et al. [2024] for an upper bound of this form under a related private IHT model.

3.6 Excess Risk Bounds

In this section, we analyze the excess risk bound of Algorithm 5. The argument is a direct conversion of the parameter bound in Theorem 17 through restricted smoothness of the population risk. This keeps the empirical optimization event, which drives the IHT recursion, separate from the population quantity appearing in the risk.

Let

$$F(\theta) = \mathbb{E}\left[(y - x^\top\theta)^2\right]/2, \qquad \Sigma_{\text{pop}} := \mathbb{E}[xx^\top],$$

so that $F(\theta) - F(\theta^\star) = \frac{1}{2}(\theta - \theta^\star)^\top \Sigma_{\text{pop}}(\theta - \theta^\star)$ under the correctly specified linear model. Define the population restricted smoothness constant

$$L_k^{\text{pop}} := \max_{\|u\|_2=1,\ \|u\|_0\le k} u^\top \Sigma_{\text{pop}} u.$$

The empirical restricted eigenvalues are those in Section 5; to avoid ambiguity in this section we write

$$\mu_k^{(n)} := \underline{\kappa}(k), \qquad L_k^{(n)} := \overline{\kappa}(k), \qquad \kappa_k^{(n)} := L_k^{(n)}/\mu_k^{(n)}.$$

Throughout this section we continue to work on the event $\mathcal{E}$ from Section 5 and assume $s \ge s^\star$, so that $k_0 = 2s + s^\star \le 4s$.

Lemma 3.6.1 (Population risk from parameter error)**.** *If θ and $\theta^\star$ satisfy $\|\theta - \theta^\star\|_0 \le k$, then*

$$F(\theta) - F(\theta^\star) \le \frac{L_k^{\text{pop}}}{2}\,\|\theta - \theta^\star\|_2^2.$$

3.6.1 Main Result

The following result is our main result on the excess risk bound of Algorithm 5.

Theorem 18 (Excess risk of ONESHOT DP-NIHT WITH SPARSIFIER)**.** *Let $\lambda = \eta B/n$ be the per-iteration Laplace scale used by the oneshot selection oracle, let $\eta = 1/(2L_{4s}^{(n)})$, and take $M \ge c_0\,\kappa_{4s}^{(n)} \log(b_{\max}^2 n)$. Assume the hypotheses of Theorem 17 hold with the empirical restricted eigenvalue constants evaluated at level $4s$; in particular, $s \ge C_0(\kappa_{4s}^{(n)})^2 s^\star$. Then, on the same event as in Theorem 17, the ONESHOT DP-NIHT WITH SPARSIFIER output $\widehat{\theta}$*

satisfies

$$F(\widehat{\theta})-F(\theta^\star) \;\le\; C_{\mathrm{stat}}\,\frac{L_{4s}^{\mathrm{pop}}}{(\mu_{4s}^{(n)})^2}\cdot\frac{s\,\sigma^2 x_{\max}^2\log(d/\alpha)}{n} \;+C_{\mathrm{priv}}\,L_{4s}^{\mathrm{pop}}\,(\kappa_{4s}^{(n)})^2\,s\lambda^2\Big(\tau^2+\tau_\xi\Big)\;+\frac{C_{\mathrm{opt}}L_{4s}^{\mathrm{pop}}}{n}, \tag{3.4}$$

for universal constants $C_{\mathrm{stat}}, C_{\mathrm{priv}}, C_{\mathrm{opt}} > 0$, where τ and τ_ξ are the oneshot selection and value-noise envelopes from Section 5.

A full proof of Theorem 18 is provided in the Appendix C. Additionally we have the following corollary.

Corollary 3.6.2. *Under the hypotheses of Theorem 18, take*

$$\eta=\frac{1}{2L_{4s}^{(n)}},\qquad \lambda=\frac{\eta B}{n}=\frac{B}{2L_{4s}^{(n)}n},\qquad \tau=\log\Big(\frac{2dM}{\alpha}\Big),\qquad \tau_\xi=\Theta\Big(\log^2\frac{2sM}{\alpha}\Big).$$

Then

$$\begin{aligned}F(\widehat{\theta})-F(\theta^\star)\;&\le C_{\mathrm{stat}}\,\frac{L_{4s}^{\mathrm{pop}}}{(\mu_{4s}^{(n)})^2}\cdot\frac{s\,\sigma^2 x_{\max}^2\log(d/\alpha)}{n}\\&\quad+\widetilde{C}_{\mathrm{priv}}\;L_{4s}^{\mathrm{pop}}\,(\kappa_{4s}^{(n)})^2\frac{sB^2}{(L_{4s}^{(n)})^2n^2}\left[\log^2\Big(\frac{2dM}{\alpha}\Big)+\log^2\Big(\frac{2sM}{\alpha}\Big)\right]\;+\frac{C_{\mathrm{opt}}L_{4s}^{\mathrm{pop}}}{n}.\end{aligned}$$

for universal constants $C_{\mathrm{stat}}, \widetilde{C}_{\mathrm{priv}}, C_{\mathrm{opt}} > 0$.

We note that the first term matches the non-private fast rate up to restricted-eigenvalue constants under a sub-Gaussian gradient at $\theta^\star$, and the second term is the privacy penalty stemming from the one-shot selection and value noises. Once the scale B is chosen according to the privacy theorem and $M \asymp \log n$, this privacy term has the same $n^{-2}\varepsilon_2^{-2}$ behavior as the private sparse-regression benchmark, up to logarithmic factors.

3.7 Experiments

We implement a controlled synthetic study on logistic classification using Algorithm 5.

Data generation. We draw features $X \in \mathbb{R}^{n\times d}$ with entries $X_{ij} \sim \mathcal{N}(0,1)$, standardize each column, and then clip the resulting matrix to $[-x_{\max}, x_{\max}]$ before applying the private mechanism. We fix a $s^\star$-sparse ground truth $\theta^\star \in \mathbb{R}^d$ by sampling a support of size $s^\star$ uniformly at random and setting nonzeros to magnitude 2.0 with random signs. Labels are

sampled from the logistic model

$$y \mid X \sim \text{Bernoulli}\Big(\sigma(X\theta^\star)\Big), \qquad \sigma(z) = \frac{1}{1+e^{-z}}.$$

where $n_{\text{train}} = 4000$, $n_{\text{test}} = 1000$, $d = 300$, $s^\star = 20$, and $x_{\max} = 1$. Both methods iterate a single gradient step (logistic loss) followed by hard thresholding on magnitudes and Euclidean projection onto the ℓ_1 ball $\{\|\theta\|_1 \le C\}$ with $C = 1.2\,\|\theta^\star\|_1$. For the gradient step we use step size $\eta = 1/(2L)$ with $L = 0.25\,\lambda_{\max}\Big((1/n)X^\top X\Big)$, the standard Lipschitz constant for the logistic gradient.

Privacy calibration. At each release we apply the oneshot top-s mechanism to the magnitudes, adding (i) a single i.i.d. Laplace vector for *selection* and (ii) *fresh* independent Laplace noise to the values on the selected support, coordinates not selected are set to zero. We allocate per-iteration budgets $(\varepsilon_m, \delta_m) = (\varepsilon_2/M, \delta/M)$ and set the Laplace scale by Theorem 2.2 of Qiao et al. [2021]:

$$\lambda \;\ge\; \frac{8\,s_\infty\sqrt{s\log\Big(d/\delta_m\Big)}}{\varepsilon_m}, \qquad s_\infty \;=\; \frac{2\eta}{n}\,x_{\max},$$

where s_∞ is the ℓ_∞ sensitivity of the map $D \mapsto \theta - \eta\nabla\ell_{\text{logistic}}(\theta; D)$ under replace-one adjacency (because $|(p-y)| \le 1$ and $\|x_i\|_\infty \le x_{\max}$). We employ basic composition across iterations of M releases. For the reported runs we fix $M = 1$ so that the per-release ε_m is not unduly small, this keeps the oneshot mechanism in a practically informative regime while preserving a clean link to theorems.

Oneshot DP-NIHT with Sparsifier uses the same oneshot mechanism at each release and projects onto the ℓ_1 ball. Oneshot DP-NIHT with Sparsifier privatizes the sparsity via the Sparsifier: we clip an initial count to $[\alpha, \beta]$, add two-sided geometric noise with parameter $1 - \exp\{-\varepsilon_1/(\beta-\alpha)\}$, re-clip/round, then run Oneshot DP-NIHT with $s = c$ and finally keep the top-c magnitudes. To isolate the effect of count privatization, we initialize the count at $s^\star$ (oracle) before adding noise, and take $(\alpha, \beta) = (1, 3s^\star)$ and a precision factor $\rho = 1$ (applied after noising and rounding, then clipped to $[1, d]$). We sweep total privacy budgets $\varepsilon_{\text{total}} \in \{0.5, 1, 2, 4\}$ with $\delta = 10^{-6}$. For Oneshot DP-NIHT with Sparsifier we set $\varepsilon_1 = \min(0.2,\, 0.25\,\varepsilon_{\text{total}})$ and $\varepsilon_2 = \varepsilon_{\text{total}} - \varepsilon_1$, while Oneshot DP-NIHT uses $\varepsilon_2 = \varepsilon_{\text{total}}$ and $\varepsilon_1 = 0$. (When $\varepsilon_m > 0.2$ we use the same scale formula as a practical extrapolation beyond the small-ε regime of the theorem.) We report test Accuracy and ROC/AUC, averaged over 5 independent seeds per $\varepsilon_{\text{total}}$; the table shows mean $\pm$ standard deviation. The column "Avg s used" records the average sparsity level actually used by each

$\varepsilon_{\text{total}}$	Accuracy	AUC	Avg s used
0.5	0.514 ± 0.019	0.531 ± 0.025	6.6
1.0	0.542 ± 0.033	0.562 ± 0.044	4.8
2.0	0.611 ± 0.034	0.659 ± 0.049	5.8
4.0	0.664 ± 0.060	0.714 ± 0.078	4.8

Table 3.1: Synthetic logistic classification. Test Accuracy and ROC–AUC (mean ± sd over five seeds). $n_{\text{train}} = 4000$, $n_{\text{test}} = 1000$, $d = 300$, $s^\star = 20$, $M = 1$, $\delta = 10^{-6}$.

method (for ONESHOT DP-NIHT WITH SPARSIFIER, this is $s^\star$).

Results Table 3.1 shows that performance improves overall as $\varepsilon_{\text{total}}$ increases for ONESHOT DP-NIHT WITH SPARSIFIER, with ROC–AUC rising from ≈ 0.53 at $\varepsilon_{\text{total}} = 0.5$ to ≈ 0.71 at $\varepsilon_{\text{total}} = 4$. The oneshot selection noise dominates the single update when $\varepsilon_{\text{total}}$ (hence ε_m) is small, and in this regime ONESHOT DP-NIHT and ONESHOT DP-NIHT WITH SPARSIFIER are both near chance. For moderate budgets ($\varepsilon_{\text{total}} \geq 1$), ONESHOT DP-NIHT WITH SPARSIFIER typically outperforms ONESHOT DP-NIHT. The column "Avg s used" reflects the privatized count c; at small budgets the counting noise (and the post-noise precision factor) can push c below $s^\star$ or inflate it, illustrating the same bias–variance tradeoff that underpins the analysis of the Sparsifier. Using $M = 1$ keeps the per-release calibration at a meaningful scale under simple composition; increasing M would require tighter accounting (e.g., moments accountant or zCDP) or larger n to avoid swamping the signal with selection noise, which we leave for follow-up experiments.

3.8 Discussion

In this chapter, we study the feature selection problem in high-dimensional sparse linear regression under the differential privacy framework. A natural approach is to view the problem as a LP, and apply a developed private LP algorithm, which typically does not preserve sparsity, which is the core spirit of sparsity. Rather than privatizing a generic linear-programming solver, in this chapter we propose Oneshot DP-NIHT with Sparsifier, an efficient hard-thresholding-based procedure that returns sparse estimates and attains a minimax-optimal parameter-error rate up to logarithmic factors. The method uses a one-shot private top-s mechanism instead of a peeling step, together with a Sparsifier that privately releases a data-dependent sparsity level. We proved the privacy result of Oneshot DP-NIHT with Sparsifier, and additionally, we also provide a population excess-risk upper bound obtained from the parameter-error guarantee through restricted population smoothness.

CHAPTER 4

A Differentially Private Dantzig Selector Algorithm

4.1 Introduction

In this chapter, we aim to explore explicitly on a differentially private Dantzig Selector algorithm. In chapter 3, we develop a private sparse-regression procedure through a noisy hard-thresholding oracle, which can be directly applied to solve Dantzig selector differentially privately. The present chapter takes a more literal route. It returns to the score constraint that defines the Dantzig selector, namely the requirement that the empirical correlations between the residual vector and all covariates be uniformly small, and asks whether this score geometry can be privatized directly while retaining exact sparsity. This point of view is close in spirit to the original Dantzig selector of Candes and Tao [2007] and to the sparse-recovery intuition behind compressed sensing and high-dimensional model selection Donoho [2006], Candes et al. [2006], Bickel et al. [2009].

The difficulty is that the Dantzig selector is naturally written as a high-dimensional linear program. Generic private linear-programming methods protect the optimization output, but they do not exploit the sparse statistical structure of the regression model and need not return a sparse vector Hsu et al. [2014], Mangasarian [2011]. Sparsity is not merely a computational preference in this problem, it is part of the statistical meaning of the estimator. The selected support is the interpretable object in feature selection, and the sharp high-dimensional rates depend on controlling errors on sets whose cardinality is comparable to the true sparsity.

The construction below therefore separates the Dantzig selector into two operations. First, it privately identifies the coordinates on which the empirical Dantzig score is most strongly violated. This is done with the one-shot top-k mechanism Qiao et al. [2021], which is well suited to sparse support selection because its noise depends logarithmically on the ambient dimension. Second, after forming a small private active set, the algorithm releases only the restricted score map on this active set and performs a low-dimensional Dantzig-score refit. A final hard-pruning step guarantees exact sparsity. Thus the procedure is neither a private

linear-programming solver nor a private hard-thresholding oracle for the squared loss, it is an active-set method driven directly by the Dantzig score.

The theoretical guarantees follow the same structure. The privacy proof uses adaptive composition: one part of the budget is spent on private score-violation identification and the other on a Gaussian release of the restricted Gram-score pair. The statistical proof shows that, under a sparse-RIP condition, the privately selected violated score coordinates capture the part of the current error outside the active set. The restricted refit then contracts the estimation error up to the ordinary stochastic score error and the two privacy errors. Under the standard sparse-row normalization, the resulting parameter rate matches the known high-dimensional differentially private sparse-regression benchmark of Cai et al. [2021] up to logarithmic factors.

4.2 A Direct Private Active-Set Dantzig Selector

The procedure analyzed in Chapter 3 is built from a private hard-thresholding oracle. In this chapter, we develop a complementary construction which starts more directly from the Dantzig selector itself. Recall that, for the clipped response vector $y^R = \mathrm{clip}_R(y)$, the empirical Dantzig score is

$$q_D(\theta) := \frac{1}{n} X^\top (y^R - X\theta) = z - G\theta, \qquad z := \frac{1}{n} X^\top y^R, \quad G := \frac{1}{n} X^\top X. \tag{4.1}$$

The classical Dantzig selector controls $\|q_D(\theta)\|_\infty$. Instead of solving the full linear program privately, the method in this chapter privately searches for the largest violated Dantzig score coordinates, forms a small active set, and solves a private restricted Dantzig-score refit on that active set. The active-set restriction is essential: it keeps the output sparse and avoids releasing either a dense coefficient vector or the full $d \times d$ Gram matrix.

The algorithm will be called DP-ASDS, for differentially private active-set Dantzig selector. It is not a generic private linear-programming routine. The only high-dimensional private operation is a one-shot top-k search over the Dantzig score coordinates, and the only matrix quantity released is a low-dimensional score map on a privately selected active set. This is the main structural distinction between DP-ASDS and a black-box differentially private solver for the Dantzig linear program.

This distinction is useful to keep in mind throughout the chapter. The active set is chosen because the Dantzig score itself reveals where the residual remains correlated with the design. The refit is then performed only on the coordinates that the private score comparison has judged statistically relevant. In this way, the method combines a Dantzig-type feasibility

principle with the active-set philosophy familiar from sparse approximation: the algorithm searches for a small set on which the score equation should be enforced, rather than solving all d score inequalities at once.

4.2.1 Notation and standing normalization

The notation in this section is chosen to make explicit which objects are global and which are released only after restriction to a private active set. The full score vector $q_D(\theta)$ is a d-dimensional object, but the algorithm never releases it. Instead, it uses the score only to select a support and then releases a low-dimensional pair (G_{UU}, z_U) on that support. The restricted-isometry notation below is used to formalize the principle that, on sparse subsets, the empirical Gram matrix behaves like a well-conditioned identity perturbation.

For $T \subset [d]$, we write X_T for the submatrix of X with columns indexed by T, $G_{TT} = X_T^\top X_T/n$, and $z_T = X_T^\top y^R/n$. We use $H_s(v)$ for the vector obtained by keeping the s largest coordinates of $|v|$ and setting the rest to zero. All ties are broken deterministically. Throughout this section, s is the target sparsity level used by the algorithm and $s^\star = \|\theta^\star\|_0$ is the true sparsity.

For the utility analysis we use the sparse restricted-isometry constants of the empirical Gram matrix. For an integer k, let δ_k be the smallest number such that

$$(1-\delta_k)\|u\|_2^2 \le u^\top G_{TT} u \le (1+\delta_k)\|u\|_2^2 \tag{4.2}$$

for all $T \subset [d]$ with $|T| \le k$ and all $u \in \mathbb{R}^T$. We also define the sparse row bound

$$K_{3s} := \max_{1\le i\le n} \max_{T\subset[d]:|T|\le 3s} \|x_{i,T}\|_2. \tag{4.3}$$

The quantity K_{3s} is the natural sensitivity parameter for the restricted Gram-score release. In the usual normalized sparse-regression regime, $K_{3s} = O(1)$, if one assumes only a coordinatewise bound $\|x_i\|_\infty \le x_{\max}$, then $K_{3s} \le \sqrt{3s}\, x_{\max}$, and the private release of a restricted Gram matrix becomes correspondingly more expensive.

This distinction between $x_{\max}$ and K_{3s} is one of the few new normalization issues introduced by the direct Dantzig construction. The score-identification step is coordinatewise and is governed by $x_{\max}$. The restricted refit, however, releases a small Gram matrix, and its sensitivity is governed by the Euclidean size of sparse rows. The minimax discussion later should therefore be read under the sparse-row normalization $K_{3s} = O(1)$, or under an equivalent normalization that keeps the restricted Gram release from dominating the privacy cost.

4.2.2 Algorithm

The algorithm has two private primitives. The first is a support-search primitive: it asks which Dantzig score coordinates are currently too large. The second is a restricted refitting primitive: after a small active set has been selected, it privately releases the score map needed to refit on that active set. This division is important because it prevents the algorithm from paying privacy cost for the full Gram matrix and, at the same time, prevents the output from becoming dense.

We now specify the two privacy budgets used by the algorithm. The first budget, $(\varepsilon_I, \delta_I)$, is used to identify large Dantzig-score coordinates. The second budget, $(\varepsilon_F, \delta_F)$, is used to release the restricted score map on the active set. Let

$$\Delta_q := \frac{2x_{\max}(R + x_{\max}C)}{n}, \qquad \Delta_F := \frac{2K_{3s}\sqrt{K_{3s}^2 + R^2}}{n}. \tag{4.4}$$

For M iterations, set

$$b_I := \frac{8\Delta_q\sqrt{2s\log(dM/\delta_I)}}{\varepsilon_I/M}, \qquad \sigma_F := \frac{\Delta_F\sqrt{2\log(1.25M/\delta_F)}}{\varepsilon_F/M}. \tag{4.5}$$

The scale b_I is the one-shot top-$2s$ scale from Qiao et al. [2021], the scale σ_F is the Gaussian mechanism scale for the restricted Gram-score query.

Algorithm 6 DP-ASDS: private active-set Dantzig selector

Input: Data (X, y); sparsity s; privacy parameters $(\varepsilon_I, \delta_I)$ and $(\varepsilon_F, \delta_F)$; iterations M; clipping level R; radius C; bounds $x_{\max}, K_{3s}$.

1: Set $y^R = \mathrm{clip}_R(y)$, $G = X^\top X/n$, $z = X^\top y^R/n$, and initialize $\theta^{(0)} = 0$.

2: **for** $m = 0, 1, \ldots, M-1$ **do**

3: Compute the Dantzig score

$$q^{(m)} = z - G\theta^{(m)}.$$

4: Run the one-shot top-$2s$ mechanism on $|q^{(m)}|$ with privacy budget $(\varepsilon_I/M, \delta_I/M)$ and Laplace scale b_I. Retain only the selected support and denote it by J_m.

5: Form the active set

$$U_m = \mathrm{supp}(\theta^{(m)}) \cup J_m.$$

6: Release the restricted score map on U_m:

$$\widetilde{G}_m = G_{U_m U_m} + W_m, \qquad \widetilde{z}_m = z_{U_m} + v_m,$$

where the entries of W_m and v_m are independent $N(0, \sigma_F^2)$ random variables.

7: Compute the private restricted Dantzig-score refit

$$\widetilde{\beta}_m \in \mathrm{argmin}_{\beta \in \mathbb{R}^{U_m}:\ \|\beta\|_1 \le C} \left\| \widetilde{z}_m - \widetilde{G}_m \beta \right\|_2.$$

8: Embed $\widetilde{\beta}_m$ into $\mathbb{R}^d$ by setting coordinates outside U_m to zero; call the embedded vector $\bar{\theta}^{(m+1)}$.

9: Prune to sparsity s:

$$\theta^{(m+1)} = H_s(\bar{\theta}^{(m+1)}).$$

10: **end for**

Output: $\widehat{\theta} = \theta^{(M)}$.

Before continuing to the analysis, we discuss several features of Algorithm 6. First, the support J_m is selected from the empirical Dantzig score, not from a gradient-updated coefficient vector. Second, the active set U_m has size at most $3s$, so the subsequent private matrix release is low-dimensional. Third, the final pruning step is part of the algorithm rather than an afterthought, it ensures that the released estimator is exactly sparse, which is the central structural feature inherited from the Dantzig selector and the Lasso Tibshirani [1996], Candes and Tao [2007].

The refit in line 6 is a least-violation version of the Dantzig score equation on the privately

selected active set. One could replace the displayed ℓ_2 score norm by an ℓ_∞ score norm to obtain a still closer finite-dimensional Dantzig program. The ℓ_2 refit is used here because it yields a stable perturbation analysis under a constant sparse-RIP condition; an ℓ_∞ refit introduces an additional $\sqrt{s}$ factor in the cross-correlation term unless one imposes a stronger incoherence condition.

The use of the ℓ_2 score residual should not be interpreted as abandoning the Dantzig viewpoint. The Dantzig selector is fundamentally a score-based method: it asks the residual to be nearly orthogonal to the design. The present refit enforces this score equation on a private active set in a norm that is stable under Gaussian perturbation of the restricted Gram-score pair. The selected-score feasibility theorem below makes this interpretation precise.

4.2.3 Exact sparsity and active-set interpretation

Algorithm 6 returns an exactly sparse estimate. Indeed, if the input sparsity is s, then every iterate satisfies $\|\theta^{(m)}\|_0 \le s$ by induction, because the only step that produces the next public iterate is $\theta^{(m+1)} = H_s(\bar{\theta}^{(m+1)})$. The restricted refit may temporarily live on a set of size at most $3s$, but this intermediate vector is not the final iterate. The hard-pruning step restores the sparsity budget before the next score computation and before the final release.

This exact sparsity guarantee is more than a presentational feature. In high-dimensional regression, the statistical analysis is local: restricted eigenvalue, restricted smoothness, and sparse-RIP assumptions are meaningful only when all relevant error vectors have small support. Exact pruning keeps the entire trajectory within this sparse regime. It also preserves the feature-selection interpretation of the output, which is one of the main reasons for studying the Dantzig selector rather than a generic private convex program.

4.2.4 Privacy guarantee

The privacy result and its proof follows the usual two principles of differential privacy: sensitivity calibration and adaptive composition Dwork et al. [2006a,b,c]. The only point requiring care is that the active set is itself private and data-dependent. The argument conditions on the transcript already released, proves privacy of the next step uniformly over all possible active sets of the allowed size, and then composes the resulting guarantees over the M iterations.

Theorem 19 (Privacy of DP-ASDS)**.** *Assume $\|x_i\|_\infty \le x_{\max}$ and $|y_i^R| \le R$. Since the refit is constrained to the ℓ_1 ball of radius C and pruning preserves the ℓ_1 norm, every iterate of*

Algorithm 6 satisfies $\|\theta^{(m)}\|_1 \le C$. *Suppose* $\varepsilon_I/M \le 0.2$, $\delta_I/M \le 0.05$, *and* $\varepsilon_F/M \le 1$. *If* b_I *and* σ_F *are chosen as in* (4.5), *then Algorithm 6 is* $(\varepsilon_I + \varepsilon_F, \delta_I + \delta_F)$*-differentially private.*

The theorem separates the two sources of privacy loss. The budget $(\varepsilon_I, \delta_I)$ protects the identity of the violated score coordinates, while $(\varepsilon_F, \delta_F)$ protects the restricted empirical score map used in the refit. Once these noisy objects have been released, the optimization step and the pruning step are pure post-processing. This is why the privacy accounting does not depend on the numerical method used to solve the small refitting problem.

4.2.5 Error bound

The error analysis is organized around a simple contraction principle. If the current error is $h^{(m)} = \theta^{(m)} - \theta^\star$, then the Dantzig score satisfies $q^{(m)} = e - Gh^{(m)}$ on the no-clipping event. Under sparse RIP, the coordinates of $q^{(m)}$ with large magnitude reveal the coordinates on which $h^{(m)}$ has not yet been captured by the active set. The private top-$2s$ step therefore acts as a noisy active-set discovery step, and the restricted refit contracts the remaining error up to stochastic and privacy perturbations.

We now prove a parameter error bound. The theorem is stated conditionally on a fixed design matrix satisfying the sparse-RIP and sparse-row bounds. For random designs, the probability of these design events can be added to the final failure probability.

Assume the linear model

$$y_i = x_i^\top \theta^\star + \varepsilon_i, \qquad \|\theta^\star\|_0 = s^\star, \qquad \|\theta^\star\|_1 \le C, \tag{4.6}$$

where ε_i are independent mean-zero σ-sub-Gaussian variables, conditionally on X. Let $\mathcal{E}_R$ be the no-clipping event $\{y^R = y\}$. On $\mathcal{E}_R$, write

$$e := \frac{1}{n} X^\top \varepsilon. \tag{4.7}$$

For a failure probability $\alpha_n \in (0,1)$, define

$$\lambda_n := \sigma x_{\max} \sqrt{\frac{2\log(4d/\alpha_n)}{n}}. \tag{4.8}$$

Then $\|e\|_\infty \le \lambda_n$ with probability at least $1 - \alpha_n$, conditionally on X.

The privacy-noise envelopes are

$$\omega_I := b_I \log\left(\frac{4dM}{\alpha_I}\right), \qquad \omega_F := c_F(1+C)\sigma_F\left(\sqrt{s} + \sqrt{\log\left(\frac{4M}{\alpha_F}\right)}\right), \tag{4.9}$$

where $c_F > 0$ is a sufficiently large universal constant. We will also use

$$\nu_F := c_F \sigma_F \left(\sqrt{s} + \sqrt{\log\left(\frac{4M}{\alpha_F}\right)} \right). \tag{4.10}$$

The condition $\nu_F \leq (1 - \delta_{3s})/2$ ensures that each noisy restricted Gram matrix remains well-conditioned.

Theorem 20. *Assume the linear model in* (4.6)*, $s \geq s^\star$, $\delta_{4s} \leq \delta_0$ for a sufficiently small universal constant δ_0, and $\nu_F \leq (1-\delta_{3s})/2$. Suppose $M \geq c_0 \log(C^2 n)$ for a sufficiently large universal constant c_0. Then, on the no-clipping event $\mathcal{E}_R$, Algorithm 6 satisfies*

$$\|\widehat{\theta} - \theta^\star\|_2^2 \leq \frac{1}{n} + C_1 s\lambda_n^2 + C_2 s\omega_I^2 + C_3 \omega_F^2 \tag{4.11}$$

with probability at least $1 - \alpha_n - \alpha_I - \alpha_F$, where $C_1, C_2, C_3 > 0$ are universal constants. If $\mathbb{P}(\mathcal{E}_R^c) \leq \alpha_R$, the same bound holds with probability at least $1 - \alpha_R - \alpha_n - \alpha_I - \alpha_F$.

The three terms in (4.11) have distinct interpretations. The term $s\lambda_n^2$ is the ordinary non-private high-dimensional score error. The term $s\omega_I^2$ is the price of privately identifying the active coordinates. The term ω_F^2 is the price of privately releasing the low-dimensional restricted refit problem. This decomposition is useful because it isolates the new cost of the direct Dantzig construction: unlike the hard-thresholding oracle of Chapter 3, the present method releases a restricted Gram-score map, and the sensitivity of that release is measured by K_{3s}.

The proofs of the following lemmas are placed in the appendix. The statements are kept in the main chapter because they describe the mechanism of the argument: first a noisy score-comparison lemma, then an active-set identification lemma, then a restricted refit lemma, and finally the elementary pruning inequality.

The proof is deterministic on three high-probability events. First,

$$\mathcal{E}_n := \{\|e\|_\infty \leq \lambda_n\}.$$

Second, if $g^{(m)}$ denotes the Laplace vector used in the one-shot score selection at iteration m, then

$$\mathcal{E}_I := \left\{ \max_{0 \leq m < M} \|g^{(m)}\|_\infty \leq \omega_I \right\}.$$

Since $\mathbb{P}(|\operatorname{Lap}(b_I)| > t) = \exp(-t/b_I)$, a union bound gives $\mathbb{P}(\mathcal{E}_I^c) \leq \alpha_I$. Third, standard Gaussian concentration and a union bound over $m = 0, \ldots, M-1$ give, with probability at

least $1 - \alpha_F$, the event

$$\mathcal{E}_F := \{\|v_m\|_2 + C\|W_m\|_{\mathrm{op}} \le \omega_F, \quad \|W_m\|_{\mathrm{op}} \le \nu_F, \quad 0 \le m < M\}. \tag{4.12}$$

Indeed, for $p \le 3s$, if $v \sim N(0, \sigma_F^2 I_p)$ then

$$\mathbb{P}\{\|v\|_2 > \sigma_F(\sqrt{p} + t)\} \le e^{-t^2/2},$$

while if W is a $p \times p$ matrix with independent $N(0, \sigma_F^2)$ entries, then

$$\mathbb{P}\{\|W\|_{\mathrm{op}} > \sigma_F(2\sqrt{p} + t)\} \le 2e^{-t^2/2}.$$

Taking t of order $\sqrt{\log(4M/\alpha_F)}$ proves (4.12) after increasing c_F.

We now give the deterministic part of the proof.

The first deterministic ingredient records two standard consequences of sparse RIP. On sparse sets, the empirical Gram matrix is close to the identity; between disjoint sparse sets, the cross-Gram block is small. These two facts are repeatedly used to convert score information into coefficient-error information, a familiar step in sparse recovery analyses Candes et al. [2006], Bickel et al. [2009].

Lemma 4.2.1 (Restricted consequences of sparse RIP)**.** *If* (4.2) *holds at level k, then for every $T \subset [d]$ with $|T| \le k$,*

$$\|(G_{TT} - I)u\|_2 \le \delta_k \|u\|_2, \qquad u \in \mathbb{R}^T.$$

Moreover, if $A, B \subset [d]$ are disjoint and $|A \cup B| \le k$, then

$$\|G_{AB} v\|_2 \le \delta_k \|v\|_2, \qquad v \in \mathbb{R}^B.$$

The next lemma is a deterministic abstraction of the private support-selection step. If the top-k comparison is made after adding bounded coordinatewise noise, then the unselected part of any competing k-sparse set cannot have much larger score mass than the selected part. This is the precise place where the one-shot top-k mechanism enters the estimation proof.

Lemma 4.2.2 (Noisy score comparison)**.** *Let $a \in \mathbb{R}^d$, let $g \in \mathbb{R}^d$ satisfy $\|g\|_\infty \le \omega$, and let J be the indices of the k largest coordinates of $|a| + g$. Then, for every $T \subset [d]$ with $|T| \le k$,*

$$\|a_{T \setminus J}\|_2 \le \|a_{J \setminus T}\|_2 + 2\sqrt{k}\,\omega. \tag{4.13}$$

The private identification lemma translates the preceding score comparison into a statement about the coefficient error. Since $q^{(m)} = e - Gh^{(m)}$, large missed components of $h^{(m)}$ would create large score coordinates unless they are hidden by stochastic noise, privacy noise, or the sparse-RIP distortion. Thus the private active set captures all but a controlled tail of the current error.

Lemma 4.2.3 (Private identification of violated Dantzig coordinates)**.** *We work on $\mathcal{E}_R \cap \mathcal{E}_n \cap \mathcal{E}_I$ and assume $\delta_{4s} < 1$. Let*

$$h^{(m)} := \theta^{(m)} - \theta^\star, \qquad U_m = \operatorname{supp}(\theta^{(m)}) \cup J_m.$$

Then, for every m,

$$\|h^{(m)}_{U_m^c}\|_2 \le c_1 \delta_{4s} \|h^{(m)}\|_2 + c_2 \sqrt{s}(\lambda_n + \omega_I), \tag{4.14}$$

where $c_1, c_2 > 0$ are universal constants.

After the active set has captured the relevant coordinates, the restricted refit solves the selected score equation with a privately perturbed Gram-score pair. The following lemma is the stability statement for that refit. Its conditioning assumption on $\widetilde{G}_m$ is indispensable: without it, a small perturbation of the restricted score map could be amplified by an ill-conditioned active-set Gram matrix.

Lemma 4.2.4 (Restricted private Dantzig-score refit)**.** *we work on $\mathcal{E}_R \cap \mathcal{E}_n \cap \mathcal{E}_F$. Let $U = U_m$, $|U| \le 3s$, and let $\bar{\theta}^{(m+1)}$ be the embedded unpruned refit produced in line 7 of Algorithm 6. If $\|W_m\|_{\mathrm{op}} \le (1 - \delta_{3s})/2$, then*

$$\|\bar{\theta}^{(m+1)} - \theta^\star\|_2 \le c_3 \|\theta^\star_{U^c}\|_2 + c_4 \sqrt{s} \lambda_n + c_5 \omega_F, \tag{4.15}$$

where $c_3, c_4, c_5 > 0$ are universal constants.

The final deterministic ingredient is the standard best-s approximation inequality. It is included to make explicit why the algorithm may refit on a set of size at most $3s$ and still return to the sparse parameter space without losing the order of the error.

Lemma 4.2.5 (Pruning)**.** *If $s \ge s^\star$, then for every $b \in \mathbb{R}^d$,*

$$\|H_s(b) - \theta^\star\|_2 \le 2\|b - \theta^\star\|_2. \tag{4.16}$$

Combining these lemmas gives the one-step recursion

$$\|\theta^{(m+1)} - \theta^\star\|_2 \le \rho \|\theta^{(m)} - \theta^\star\|_2 + C\sqrt{s}(\lambda_n + \omega_I) + C\omega_F,$$

with $\rho < 1$ when δ_{4s} is sufficiently small. Iterating the recursion produces Theorem 20. The proof is deliberately parallel to classical active-set sparse-recovery arguments, except that the score comparison and the restricted refit are both privatized.

4.2.6 Expanded rate and minimax comparison

The abstract error bound is most informative after substituting the privacy scales. This substitution also clarifies the relationship between the direct Dantzig construction and the minimax theory for private sparse regression. The target benchmark is the usual non-private sparse rate plus the privacy penalty identified in the cost-of-privacy literature Cai et al. [2021].

Substituting (4.5) and (4.9) into Theorem 20 gives the following explicit bound:

$$\begin{aligned}\|\widehat{\theta} - \theta^\star\|_2^2 \le \frac{1}{n} &+ C\frac{\sigma^2 x_{\max}^2 s \log(4d/\alpha_n)}{n} \\ &+ C\frac{s^2 M^2 x_{\max}^2 (R + x_{\max}C)^2}{n^2\varepsilon_I^2} \log\left(\frac{dM}{\delta_I}\right) \log^2\left(\frac{4dM}{\alpha_I}\right) \\ &+ C\frac{(1+C)^2 K_{3s}^2 (K_{3s}^2 + R^2) M^2}{n^2\varepsilon_F^2} \log\left(\frac{1.25M}{\delta_F}\right)\left(s + \log\left(\frac{4M}{\alpha_F}\right)\right). \end{aligned} \tag{4.17}$$

The first line is the usual non-private sparse-regression term. The second line is the privacy cost of identifying the large Dantzig-score coordinates. The third line is the privacy cost of releasing the restricted score map used for refitting.

The second line has the same qualitative form as the privacy term in sparse private selection: the ambient dimension appears only through logarithms, while the leading sparsity dependence is quadratic after squaring the support-identification noise. The third line is specific to the direct active-set Dantzig approach. It is lower order in the sparse-row normalized regime, but it records the fact that a Dantzig-score refit must learn a local Gram map, not only a coefficient vector.

In the sparse-row normalized regime $K_{3s} = O(1)$, with $x_{\max}(R + x_{\max}C)$ of constant or logarithmic order, $s \asymp s^\star$, $M \asymp \log n$, and $\varepsilon_I \asymp \varepsilon_F \asymp \varepsilon$, the dominant privacy term in (4.17) is

$$\frac{(s^\star \log d)^2}{n^2\varepsilon^2} \cdot \mathrm{polylog}(n, d, 1/\delta, 1/\alpha).$$

Thus DP-ASDS attains

$$\|\widehat{\theta} - \theta^\star\|_2^2 \lesssim \frac{\sigma^2 s^\star \log d}{n} + \frac{(s^\star \log d)^2}{n^2\varepsilon^2} \cdot \mathrm{polylog}(n, d, 1/\delta, 1/\alpha), \tag{4.18}$$

up to restricted-isometry constants and logarithmic factors. This is the same high-

dimensional private sparse-regression benchmark established by Cai et al. [2021]. If one assumes only $\|x_i\|_\infty \le x_{\max}$ with $x_{\max} = O(1)$, then K_{3s} may grow as $\sqrt{s}$ and the restricted-Gram release in the third line of (4.17) may be larger. The minimax comparison above is therefore a statement for the standard sparse-row normalization under which private high-dimensional regression upper bounds are usually

Consequently, the active-set Dantzig method attains the same benchmark as Chapter 3 when the restricted matrix release is statistically inexpensive. If the design rows are only coordinatewise bounded and no sparse-row normalization is imposed, the restricted Gram release may introduce an additional sparsity factor. The result should therefore be read as a sharp theorem for the normalized sparse-design regime, not as a claim that arbitrary design scalings can be privatized at no cost..

4.2.7 Selected Dantzig-score feasibility

The parameter bound is the primary statistical guarantee, but it is also useful to record what the algorithm achieves at the level of the Dantzig score itself. Because the final pruning step is imposed to guarantee exact sparsity, the strongest score statement is naturally stated for the unpruned restricted refit. This is analogous to active-set methods in numerical optimization: the refit enforces the optimality equations on the current working set, while the pruning step enforces the structural constraint.

The next statement records the sense in which the method directly enforces the Dantzig score equations. It concerns the unpruned refit $\bar{\theta}^{(m+1)}$, because pruning is imposed for exact sparsity and may slightly perturb the selected score residual.

Theorem 21 (Selected Dantzig-score feasibility)**.** *Under the assumptions and on the event of Theorem 20, for every $m = 0, \dots, M-1$,*

$$\left\| q_D(\bar{\theta}^{(m+1)})_{U_m} \right\|_2 \le C\delta_{4s}\|h^{(m)}\|_2 + C\sqrt{s}(\lambda_n + \omega_I) + C\omega_F. \tag{4.19}$$

Consequently, for the final refit, if $M \ge c_0 \log(C^2 n)$ and $\delta_{4s} \le \delta_0$ is sufficiently small,

$$\left\| q_D(\bar{\theta}^{(M)})_{U_{M-1}} \right\|_2 \le C\sqrt{s}(\lambda_n + \omega_I) + C\omega_F + Cn^{-1/2}. \tag{4.20}$$

This result is not the same as global exact Dantzig feasibility, which would require a uniform bound on $\|q_D(\widehat{\theta})\|_\infty$ over all coordinates after pruning. Instead, it captures the active-set nature of the method: the score equations are enforced on the coordinates that have been privately identified as relevant. Together with the error theorem, this gives a

statistically meaningful Dantzig-type guarantee without solving or privatizing the full linear program.

4.2.8 Population excess risk

The population risk bound is a direct consequence of the parameter error bound. This is the natural route in the well-specified squared-loss model: once the estimator is close to $\theta^\star$ in Euclidean norm and the error vector is sparse, restricted population smoothness converts estimation error into excess prediction risk. No separate stability argument is needed for this quadratic risk.

Let

$$F(\theta) = \frac{1}{2}\mathbb{E}\{(y - x^\top\theta)^2\}, \qquad \Sigma_{\text{pop}} = \mathbb{E}[xx^\top].$$

Under the correctly specified linear model,

$$F(\theta) - F(\theta^\star) = \frac{1}{2}(\theta - \theta^\star)^\top \Sigma_{\text{pop}}(\theta - \theta^\star).$$

Define the population restricted smoothness constant

$$L_{2s}^{\text{pop}} := \max_{\|u\|_2=1,\ \|u\|_0\le 2s} u^\top \Sigma_{\text{pop}} u.$$

Since both $\widehat{\theta}$ and $\theta^\star$ are supported on at most s coordinates, $\|\widehat{\theta} - \theta^\star\|_0 \le 2s$. Therefore Theorem 20 immediately implies the following bound.

Theorem 22 (Population risk of DP-ASDS)**.** *Under the assumptions of Theorem 20, on the same event,*

$$F(\widehat{\theta}) - F(\theta^\star) \le \frac{L_{2s}^{\text{pop}}}{2}\left(\frac{1}{n} + C_1 s\lambda_n^2 + C_2 s\omega_I^2 + C_3\omega_F^2\right). \tag{4.21}$$

The factor L_{2s}^{pop} appears because the final estimator and the truth are both sparse. Hence the risk conversion only needs population smoothness on $2s$-sparse directions. This is consistent with the local nature of the entire chapter: privacy, optimization, and statistical curvature are all measured on sparse subsets rather than on the full ambient space.

4.2.9 Discussion

This chapter aims to demonstrate that the Dantzig selector can be privatized in a way that respects its score geometry and its sparsity structure. The preceding results show that this is

possible without using a generic private linear-programming algorithm and without treating the hard-thresholding oracle of Chapter 3 as a black box.

The construction above gives a differentially private procedure that is closer to the Dantzig selector than the hard-thresholded gradient oracle analyzed in Chapter 3. The iteration is driven by the empirical score $q_D(\theta) = X^\top(y^R - X\theta)/n$, and the selected coordinates are precisely the coordinates with large private Dantzig-score violations. The restricted refit then approximately solves the score equation on the private active set. Thus the algorithm follows the geometry of the Dantzig constraint rather than the geometry of a generic gradient projection step.

The proof also explains the statistical price of this more direct Dantzig construction. The private score-identification term has the same order as the top-s privacy term in private sparse regression. The additional restricted-refit term is harmless under sparse-row normalization, but it should not be ignored: if the row normalization is weakened so that K_{3s} grows with s, then privately releasing the restricted Gram matrix may become the dominant cost. This distinction is intrinsic to direct Dantzig-score refitting, since the Dantzig score depends on the Gram matrix $X^\top X/n$.

Finally, the algorithm can be combined with the private Sparsifier step from Section 4 when s is unknown. The privacy proof is unchanged by composition and post-processing. The utility theorem should then be read on the event that the privatized sparsity lies in the adequate regime $s \geq s^\star$ and $\delta_{4s} \leq \delta_0$, exactly as in the conditional sparsity treatment of the main Oneshot DP-NIHT with Sparsifier result.

From a methodological perspective, DP-ASDS gives a second route to private sparse estimation. Chapter 3 follows an iterative hard-thresholding path and obtains broad RE/RSC-type guarantees. The present chapter follows a more direct active-set Dantzig path and obtains a clean selected-score interpretation under a stronger sparse-RIP condition. These two procedures therefore complement each other: the former is closer to the modern private sparse-regression algorithmic toolkit, while the latter is closer to the defining constraint of the Dantzig selector. This distinction also suggests several possible refinements. One may replace the ℓ_2 restricted refit by an ℓ_∞ refit under stronger incoherence assumptions, use a private Sparsifier as in Chapter 3 when s is unknown Khanna et al. [2023], or investigate adaptive stopping rules based on privately monitored selected-score residuals. The analysis in this chapter isolates the essential ingredients needed for such extensions: private score identification, stable restricted refitting, exact pruning, and sparse-design curvature.

The novel active set approach introduced bridges the last missing piece among *privacy cost independent of* d, *exact sparsity*, and *minimax risk*: existing differentially private sparse regression algorithms either add d-dimensional noise at every step, return a dense estimate

that still requires post-processing, or lose a factor in sample size through data partitioning. By contrast, DP-ASDS keeps the linear program geometry of the classical Dantzig selector and marries it with a dimension-free privacy mechanism: each iteration releases only a $3s \times 3s$ sub-Gram matrix and a length $3s$ score vector, so the privacy noise scales with the target sparsity s rather than the ambient dimension d. This design achieves the minimax rate $\sigma^2 s^* \log d/n + \sigma^2 (s^* \log d)^2/(n^2\varepsilon^2)$ while preserving exact s-sparsity, something the literature did not offer. In addition, the oneshot top-$2s$ selection plus noisy restricted refit constitutes a reusable template for other LP-based estimators, showing that privacy can coexist with fast, geometry-aware algorithms instead of forcing generic quadratic objectives or heavy sample splitting. For these reasons, DP-ASDS is not simply "another DP algorithm", it demonstrates a qualitatively new way to pay privacy only where information is concentrated and thereby advances both the theoretical frontier and the practical toolbox for DP high-dimensional inference.

APPENDIX A

Appendix for Chapter 1

A.1 Proofs and Supplementary Materials

A.1.1 Proofs for one-dimensional γ-CHM problem

We provide proofs of Theorem 1, Lemma 1.5.1 and Theorem 2 in this Section. As discussed in Section 1.6.1.2, the interval CHM problem fully generalizes the regular CHM problem, and the analogous results in the interval CHM problem follow the exactly same proofs and can be derived by properly adjusting γ to γ^- and γ^+ in this section. Therefore, we only prove Theorem 1, Lemma 1.5.1 and Theorem 2 for the ease of extensions to the Gaussian bandit setting with unknown variances.

A.1.1.1 Proof of Theorem 1

We recall $\mu_1 = \min\{\mu_1, \cdots, \mu_K\}$ and $\mu_K = \max\{\mu_1, \cdots, \mu_K\}$. In the one-dimensional case, the γ-CHM problem is to test if $\gamma \in \text{Conv}(\boldsymbol{\mu})$. For the feasible case,

$$\text{Alt}(\boldsymbol{\mu}) = \{\boldsymbol{\lambda} | I_\pi(\boldsymbol{\lambda}) = \text{infeasible}\} = \{\boldsymbol{\lambda} | \max_{1\le i\le K} \lambda_i < \gamma \text{ or } \min_{1\le i\le K} \lambda_i > \gamma\}.$$

Therefore,

$$\begin{aligned}
T^*(\boldsymbol{\mu})^{-1} =& \max_{\boldsymbol{w}\in\Delta} \min_{\boldsymbol{\lambda}\in\text{Alt}(\boldsymbol{\mu})} \sum_a w_a d(\mu_a, \lambda_a) \\
=& \max_{\boldsymbol{w}\in\Delta} \min \left(\sum_{a:\mu_a<\gamma} w_a d(\mu_a, \gamma), \sum_{a:\mu_a>\gamma} w_a d(\mu_a, \gamma) \right) \\
=& \max_{w_1+w_K=1} \min \left(w_1 d(\mu_1, \gamma), w_K d(\mu_K, \gamma) \right) \\
=& \frac{d(\mu_1, \gamma) d(\mu_K, \gamma)}{d(\mu_1, \gamma) + d(\mu_K, \gamma)} \\
=& \frac{1}{d(\mu_1, \gamma)^{-1} + d(\mu_K, \gamma)^{-1}} .
\end{aligned}$$

From the derivation, we can see the optimization problem derives its optimal solution when the strategy only samples arms with maximum and minimum mean with proportions $w_1 = \frac{d(\mu_1,\gamma)^{-1}}{d(\mu_1,\gamma)^{-1}+d(\mu_K,\gamma)^{-1}}$ and $w_K = \frac{d(\mu_K,\gamma)^{-1}}{d(\mu_1,\gamma)^{-1}+d(\mu_K,\gamma)^{-1}}$. Now we consider the infeasible case. Without loss of generality, we assume $\mu_1 > \tau$ (the case when $\mu_K < \tau$ can be proved in the same way due to symmetry). In this case,

$$\text{Alt}(\boldsymbol{\mu}) = \{\boldsymbol{\lambda} | I_\pi(\boldsymbol{\lambda}) = \text{feasible}\} = \{\boldsymbol{\lambda} | \lambda_1 < \gamma < \lambda_K\}.$$

and

$$\begin{aligned} T^*(\boldsymbol{\mu})^{-1} &= \max_{\boldsymbol{w}\in\Delta} \min_{\boldsymbol{\lambda}\in\text{Alt}(\boldsymbol{\mu})} \sum_a w_a d(\mu_a, \lambda_a) \\ &= \max_{\boldsymbol{w}\in\Delta} \min_{1\le a\le K} \; w_a d(\mu_a, \gamma) \\ &= \frac{1}{\sum_{1\le a\le K} d(\mu_a,\gamma)^{-1}}\,. \end{aligned}$$

We can see from that in the infeasible case, the decision-maker should sample all arms, and for each arm $a \in \{1, \cdots, K\}$, it should be sampled with proportion $w_a = \frac{\frac{1}{d(\mu_a,\gamma)}}{\sum_{1\le i\le K}\frac{1}{d(\mu_i,\gamma)}}$.

A.1.1.2 Proof of Lemma 1.5.1

Now we proceed to prove Lemma 1.5.1. We will make use of the following proposition.

Proposition A.1.1. *(Lemma 22 of Kaufmann et al. [2016]) For every $\beta, \eta > 0$, if*

$$x \ge \frac{1}{\beta}\ln\left(\frac{e\ln(1/\beta\eta)}{\beta\eta}\right),$$

then we have

$$\beta x \ge \ln\left(\frac{x}{\eta}\right) + o\left(\ln\left(\frac{1}{\eta}\right)\right).$$

Now for the feasible case, on the event that

$$\frac{N_1(t)}{t} \to w_1^*(\boldsymbol{\mu}), \quad \hat{\mu}_1(t) \to \mu_1,$$

and

$$\frac{N_K(t)}{t} \to w_K^*(\boldsymbol{\mu}), \quad \hat{\mu}_K(t) \to \mu_K,$$

for any $\varepsilon > 0$, there exists $t_0 > 0$, such that for any $t > t_0$, we have

$$N_1(t) d^+(\hat{\mu}_1(t), \gamma) \ge (1-\varepsilon) t w_1^*(\boldsymbol{\mu}) d(\mu_1, \gamma),$$

and

$$N_K(t)d^-(\hat{\mu}_K(t),\gamma) \geq (1-\varepsilon)tw_K^*(\boldsymbol{\mu})d(\mu_K,\gamma).$$

Following this,

$$\begin{aligned}
\tau &\leq \tau_3 \\
&\leq \inf\{t|N_1(t)d^+(\hat{\mu}_1(t),\gamma) \geq \text{Thresh}(\delta,N_1(t)) \text{ and } N_K(t)d^-(\hat{\mu}_K(t),\gamma) \geq \text{Thresh}(\delta,N_K(t))\} \\
&\leq \inf\left\{t|N_1(t)d^+(\hat{\mu}_1(t),\gamma) \geq \text{Thresh}(\delta,t) \text{ and } N_K(t)d^+(\hat{\mu}_K(t),\gamma) \geq \text{Thresh}(\delta,t)\right\} \\
&\leq \inf\{t|(1-\varepsilon)tw_1^*(\boldsymbol{\mu})d(\mu_1,\gamma) \geq \text{Thresh}(\delta,t) \text{ and } (1-\varepsilon)tw_K^*(\boldsymbol{\mu})d(\mu_K,\gamma) \geq \text{Thresh}(\delta,t)\} \\
&\leq \inf\left\{t|(1-\varepsilon)t\min\left(w_1^*(\boldsymbol{\mu})d(\mu_1,\gamma), w_K^*(\boldsymbol{\mu})d(\mu_K,\gamma)\right) \geq \ln\left(\frac{t}{\delta}\right) + o\left(\ln\left(\frac{1}{\delta}\right)\right)\right\} \\
&\leq \inf\left\{t|(1-\varepsilon)tT^*(\boldsymbol{\mu})^{-1} \geq \ln\left(\frac{t}{\delta}\right) + o\left(\ln\left(\frac{1}{\delta}\right)\right)\right\}.
\end{aligned}$$

Hence we have $\tau(1-\varepsilon)T^*(\boldsymbol{\mu})^{-1} \leq \ln\left(\frac{t}{\delta}\right) + o\left(\ln\left(\frac{1}{\delta}\right)\right)$. By setting $\beta = (1-\varepsilon)T^*(\boldsymbol{\mu})^{-1}$, $x = \tau$ and $\eta = \delta$, Proposition A.1.1 directly yields

$$\begin{aligned}
\tau &\leq \frac{1}{(1-\varepsilon)T^*(\boldsymbol{\mu})^{-1}} \ln\left(\frac{e\ln\left(\frac{1}{(1-\varepsilon)T^*(\boldsymbol{\mu})^{-1}\delta}\right)}{(1-\varepsilon)T^*(\boldsymbol{\mu})^{-1}\delta}\right) \\
&\leq \frac{1}{(1-\varepsilon)T^*(\boldsymbol{\mu})^{-1}} \left(\ln\left(\frac{e}{(1-\varepsilon)T^*(\boldsymbol{\mu})^{-1}}\right) + \ln\left(\frac{1}{\delta}\right) + \ln\ln\left(\frac{1}{(1-\varepsilon)T^*(\boldsymbol{\mu})^{-1}\delta}\right)\right) \\
&\leq \frac{1}{(1-\varepsilon)T^*(\boldsymbol{\mu})^{-1}} \ln\left(\frac{1}{\delta}\right) + o\left(\ln\left(\frac{1}{\delta}\right)\right)
\end{aligned}$$

Notice that ε is arbitrary, we have

$$\limsup_{\delta\to 0} \frac{\tau}{\ln(1/\delta)} \leq T^*(\boldsymbol{\mu}).$$

For the infeasible case, WLOG we assume $\mu_1 > \gamma$ (proof of the symmetric case $\mu_K < \gamma$ is identical). On the event that for any arm $a \in \{1,\cdots,K\}$,

$$\frac{N_a(t)}{t} \to w_a^*(\boldsymbol{\mu}), \quad \hat{\mu}_a(t) \to \mu_a,$$

and for arbitrary $\varepsilon > 0$, similarly, there exists t_0, and for any $t > t_0$, we have the following inequality to hold:

$$N_a(t)d^-(\hat{\mu}_a(t),\gamma) \geq (1-\varepsilon)tw_a^*(\boldsymbol{\mu})d(\mu_a,\gamma).$$

Using a parallel statement of the feasible case,

$$\begin{aligned}
\tau \leq \tau_3 &\leq \inf\{t|\forall a, N_a(t)d(\hat{\mu}_a(t),\gamma) \geq \text{Thresh}(\delta, N_1(t))\} \\
&\leq \inf\left\{t|\forall a, N_a(t)d^-(\hat{\mu}_a(t),\gamma) \geq \text{Thresh}(\delta, t)\right\} \\
&\leq \inf\left\{t|\forall a, (1-\varepsilon)tw_a^*(\boldsymbol{\mu})d(\mu_a,\gamma) \geq \ln\left(\frac{t}{\delta}\right) + o\left(\ln\left(\frac{1}{\delta}\right)\right)\right\} \\
&\leq \inf\left\{t|(1-\varepsilon)t\min_{1\leq a\leq K}\{w_a^*(\boldsymbol{\mu})d(\mu_a,\gamma)\} \geq \ln\left(\frac{t}{\delta}\right) + o\left(\ln\left(\frac{1}{\delta}\right)\right)\right\} \\
&\leq \inf\left\{t|(1-\varepsilon)tT^*(\boldsymbol{\mu})^{-1} \geq \ln\left(\frac{t}{\delta}\right) + o\left(\ln\left(\frac{1}{\delta}\right)\right)\right\}.
\end{aligned}$$

Following the same statements and by applying Proposition A.1.1, we have

$$\tau \leq \frac{1}{(1-\varepsilon)T^*(\boldsymbol{\mu})^{-1}}\ln\left(\frac{1}{\delta}\right) + o\left(\ln\left(\frac{1}{\delta}\right)\right).$$

Thus, in the infeasible case,

$$\limsup_{\delta\to 0}\frac{\tau}{\ln(1/\delta)} \leq T^*(\boldsymbol{\mu})$$

also holds.

A.1.1.3 Proof of Theorem 2

We again consider feasible and infeasible cases separately. We recall the following notations

$$\psi_a(t) = \mathbb{P}(A_t = a|\mathcal{F}_{t-1}), \quad \Psi_a(t) = \sum_{i=1}^{t}\psi_a(i), \quad \text{and} \quad \bar{\psi}_a(t) = \frac{1}{t}\Psi_a(t).$$

Our proof is based on a classic result (see Corollary 1 of Russo [2016]) that for any arm $a \in [K]$, if $\Psi_a(t) \to \infty$, then

$$\frac{w_a(t)}{\bar{\psi}_a(t)} = \frac{N_a(t)}{\Psi_a(t)} \to 1 \quad \text{a.s}$$

and the following result from Kaufmann et al. [2018].

Proposition A.1.2. *(Theorem 12 of Kaufmann et al. [2018]) Given a threshold γ, for any $\boldsymbol{\mu} = \{(\mu_1, \cdots, \mu_K)|\mu_1 \leq \cdots \leq \mu_K,$ and $\mu_1 < \gamma\}$. If we sequentially sample as follows: for any $t \in \mathbb{N}^+$, sample $\boldsymbol{\theta}_t \sim \Pi_{t-1}(\cdot|\min_{1\leq i\leq K}\mu_i < \gamma)$, then play the arm A_t with lowest mean in $\boldsymbol{\theta}_t$. Then the sampling procedure ensures that the sampling frequencies satisfy*

$$\frac{N_1(t)}{t} \to 1,$$

and for any $2 \leq a \leq K$,

$$\frac{N_a(t)}{t} \to 0$$

almost surely.

Back to our proof, now we consider the feasible case first. In this case, we have $\boldsymbol{\mu}$ with property $\mu_1 < \gamma < \mu_K$. For any $n \in \mathbb{N}^+$,

$$\psi_1(t) = \beta_t \Pi_t \left(\theta_{t,1} < \min_{j \neq 1} \theta_{t,j} | \boldsymbol{\mu} \text{ feasible} \right) + (1 - \beta_t) \Pi_t \left(\theta_{t,1} > \max_{j \neq 1} \theta_{t,j} | \boldsymbol{\mu} \text{ feasible} \right).$$

Notice that $\{\text{argmin}_a \, \theta_a = 1\}$ and $\{\mu_K > \gamma\}$ are independent events,

$$\Pi_t \left(\theta_{t,1} < \min_{j \neq 1} \theta_{t,j} | \boldsymbol{\mu} \text{ feasible} \right) \to 1 \quad a.s..$$

And $\Pi_t \left(\theta_{t,1} < \min_{j \neq 1} \theta_{t,j} | \boldsymbol{\mu} \text{ feasible} \right) + \Pi_t \left(\theta_{t,1} > \min_{j \neq 1} \theta_{t,j} | \boldsymbol{\mu} \text{ feasible} \right) \leq 1$ directly yields

$$\Pi_t \left(\theta_{t,1} > \max_{j \neq 1} \theta_{t,j} | \boldsymbol{\mu} \text{ feasible} \right) \to 0 \quad a.s..$$

Combine with the fact that $\beta_t \to \frac{d^{-1}(\mu_1, \gamma)}{d^{-1}(\mu_1, \gamma) + d^{-1}(\mu_K, \gamma)}$, we get $\frac{N_1(t)}{t} \to w_1^*(\boldsymbol{\mu})$. Similarly,

$$\psi_K(t) = \beta_t \Pi_t \left(\theta_{t,K} < \min_{j \neq 1} \theta_{t,j} | \boldsymbol{\mu} \text{ feasible} \right) + (1 - \beta_t) \Pi_t \left(\theta_{t,K} > \max_{j \neq 1} \theta_{t,j} | \boldsymbol{\mu} \text{ feasible} \right).$$

With the facts that $\Pi_t \left(\theta_{t,K} < \min_{j \neq 1} \theta_{t,j} | \boldsymbol{\mu} \text{ feasible} \right) \to 0$ and $\Pi_t \left(\theta_{t,K} > \max_{j \neq 1} \theta_{t,j} | \boldsymbol{\mu} \text{ feasible} \right) \to 1$ almost surely, this leads to $\frac{N_K(t)}{t} \to w_K^*(\boldsymbol{\mu})$. Notice that $w_1^*(\boldsymbol{\mu}) + w_K^*(\boldsymbol{\mu}) = 1$, we have shown that $\frac{\boldsymbol{N}(t)}{t} \to w^*(\boldsymbol{\mu})$ almost sure in the feasible case.

For the infeasible case, we use the following proposition.

Proposition A.1.3. *(Simplified version of Lemma of Russo [2016]) Consider any sampling rule, if for any arm* $a \in [K]$ *and all* $c > 0$,

$$\sum_t \psi_a(t) \mathbb{1}\{\bar{\psi}_a(t) \geq w_a^* + c\} < \infty,$$

then $\bar{\psi}(t) \to \boldsymbol{w}^*$.

By applying a similar proof strategy in Russo [2016] and Kaufmann et al. [2018], we aim to prove the precondition in Proposition 1.5.2. For any $a \in [K]$ and $c > 0$, consider any

round n where $\bar{\psi}_a(t) \geq w_a^* + c$, we have

$$\begin{aligned}
\psi_a(t) &= \beta_t \frac{\Pi_{t-1}(a = \operatorname{argmin}_i \theta_{t,i}, b = \operatorname{argmax}_i \theta_{t,i}, \min_i \theta_{t,i} < \gamma < \max_i \theta_{t,i})}{\Pi_{t-1}(\min_i \theta_{t,i} < \gamma < \max_i \theta_{t,i})} \\
&\quad + (1-\beta_t) \frac{\Pi_{t-1}(a = \operatorname{argmax}_i \theta_{t,i}, b = \operatorname{argmin}_i \theta_{t,i}, \min_i \theta_{t,i} < \gamma < \max_i \theta_{t,i})}{\Pi_{t-1}(\min_i \theta_{t,i} < \gamma < \max_i \theta_{t,i})} \\
&\leq \beta_t \frac{\Pi_{t-1}(\theta_{t,a} < \gamma < \theta_{t,b})}{\max_{a,b} \Pi_{t-1}(\theta_{t,a} < \gamma < \theta_{t,b})} + (1-\beta_t) \frac{\Pi_{t-1}(\theta_{t,b} < \gamma < \theta_{t,a})}{\max_{a,b} \Pi_{t-1}(\theta_{t,b} < \gamma < \theta_{t,a})}.
\end{aligned}$$

Following Russo [2016], recall we use $x_t \doteq y_t$ to denote that $t^{-1} \ln(x_t/y_t) \to 0$. Based on any known posterior concentration rate result (for example, Proposition 5 in Russo [2016]) that for any open set $\tilde{\Theta} \subset \Theta$, the posterior concentrates at the rate $\Pi_t(\tilde{\Theta}) \doteq \exp\left(-t \min_{\boldsymbol{\lambda} \in \tilde{\Theta}} \sum_a \bar{\psi}_a(t) d(\mu_a, \lambda_a)\right)$. Moreover, for any $a, b \in [K]$,

$$\begin{aligned}
\Pi_t(\theta_{t,a} < \gamma < \theta_{t,b}) &\doteq \exp\left(-t \min_{\boldsymbol{\theta}_t \text{ feasible}} \sum_a \bar{\psi}_a(t) d(\mu_a, \theta_{t,a})\right) \\
&= \exp\left(-t \min\left(\sum_{a:\mu_a<\gamma} \bar{\psi}_a(t) d(\mu_a, \gamma), \sum_{a:\mu_a>\gamma} \bar{\psi}_a(t) d(\mu_a, \gamma)\right)\right).
\end{aligned}$$

This means, there is a sequence $\varepsilon_t \to 0$ such that for any t,

$$\Pi_t(\theta_a < \gamma < \theta_b) \in \exp\left(-t\left(\min\left(\sum_{a:\mu_a<\gamma} \bar{\psi}_a(t) d(\mu_a, \gamma), \sum_{a:\mu_a>\gamma} \bar{\psi}_a(t) d(\mu_a, \gamma)\right)\right) \pm \varepsilon_t\right),$$

which implies

$$
\begin{aligned}
\psi_a(t) \le \beta_t \frac{\Pi_{t-1}(\theta_{t,a} < \gamma < \theta_{t,b})}{\max_{a,b} \Pi_{t-1}(\theta_{t,a} < \gamma < \theta_{t,b})} + (1-\beta_t) \frac{\Pi_{t-1}(\theta_{t,b} < \gamma < \theta_{t,a})}{\max_{a,b} \Pi_{t-1}(\theta_{t,b} < \gamma < \theta_{t,a})}.& \\
= \frac{\exp\left(-t\left(\min\left(\sum_{a:\mu_a<\gamma} \bar{\psi}_a(t) d(\mu_a,\gamma), \sum_{a:\mu_a>\gamma} \bar{\psi}_a(t) d(\mu_a,\gamma)\right)\right) - \varepsilon_t\right)}{\max_a \exp\left(-t\left(\min\left(\sum_{a:\mu_a<\gamma} \bar{\psi}_a(t) d(\mu_a,\gamma), \sum_{a:\mu_a>\gamma} \bar{\psi}_a(t) d(\mu_a,\gamma)\right)\right) + \varepsilon_t\right)}& \\
= \exp\Bigg\{-t\Bigg[\min\Big(\sum_{a:\mu_a<\gamma} \bar{\psi}_a(t) d(\mu_a,\gamma), \sum_{a:\mu_a>\gamma} \bar{\psi}_a(t) d(\mu_a,\gamma)\Big)& \\
- \min_a \min\Big(\sum_{a:\mu_a<\gamma} \bar{\psi}_a(t) d(\mu_a,\gamma), \sum_{a:\mu_a>\gamma} \bar{\psi}_a(t) d(\mu_a,\gamma)\Big)\Bigg] - 2\varepsilon_t\Bigg\}& \\
\le \exp\Bigg\{-t\Bigg[\min\Big(\sum_{a:\mu_a<\gamma} (w_a^* + c) d(\mu_a,\gamma), \sum_{a:\mu_a>\gamma} (w_a^* + c) d(\mu_a,\gamma)\Big)& \\
- \min\Big(\sum_{a:\mu_a<\gamma} w_a^* d(\mu_a,\gamma), \sum_{a:\mu_a>\gamma} w_a^* d(\mu_a,\gamma)\Big)\Bigg] - 2\varepsilon_t\Bigg\}.& \\
\le \exp\left\{-t\left[c \min\Big(\sum_{a:\mu_a<\gamma} d(\mu_a,\gamma), \sum_{a:\mu_a>\gamma} d(\mu_a,\gamma)\Big) - 2\varepsilon_t\right]\right\}&
\end{aligned}
$$

When $\varepsilon_t \to 0$ the $\psi_a(t)$ is bounded by an exponential decay term, therefore

$$\sum_t \psi_a(t) \mathbb{1}\{\bar{\psi}_a(t) \ge w_a^* + c\} \le \infty.$$

Therefore, we have $\bar{\psi}(t) \to \boldsymbol{w}^*$, and by the conclusions above, $\boldsymbol{N}(t)/t \to \boldsymbol{w}^*$.

A.1.2 d-dimensional CHM problem when $d \ge 2$

In this section, we provide further details and discussions about potential extensions of Thompson-CHM algorithm to the higher dimensional case. Before moving forward, we first prove Theorem 5, which gives insight into how to generalize our algorithm.

A.1.2.1 Proofs of Theorem 5

For the infeasible case when $\gamma \notin \mathrm{Conv}(\boldsymbol{\mu})$, the proof is identical to the one-dimensional case and we omit it here.

For the feasible case when $\gamma \in \mathrm{Conv}(\boldsymbol{\mu})$. We assume $\boldsymbol{\lambda}^* \in \mathrm{Alt}(\boldsymbol{\mu})$ is the optimal solution in the game $T^*(\boldsymbol{\mu})^{-1} = \sup_{\boldsymbol{w}\in\Delta} \inf_{\boldsymbol{\lambda}\in\mathrm{Alt}(\boldsymbol{\mu})} \sum_a w_a d(\mu_a, \lambda_a)$. For any $\mu_i \in \{\mu_1, \cdots, \mu_K\} \backslash \mathrm{Vert}(\mathrm{Conv}\{\boldsymbol{\mu}\})$, we consider two different cases.

- Case 1: if $\mu_i \in \mathrm{Conv}(\boldsymbol{\lambda})$, then we have $\lambda_i = \mu_i$, and therefore $w_i = 0$.
- Case 2: if $\mu_i \in \mathrm{Conv}(\boldsymbol{\mu}) \backslash \mathrm{Conv}(\boldsymbol{\lambda})$, in this case, $\lambda_i \ne \mu_i$, WLOG we assume

$\mu_1, \mu_2, \cdots, \mu_s$ are the means that differs from those in the optimal solution $\boldsymbol{\lambda}^*$, i.e. $\{1, 2, \cdots, s\} = \{j | \mu_j \neq \lambda_j^*\}$, so $1 \leq i \leq s$. If $w_i \neq 0$, then $d(\mu_i, \lambda_i^*) > d(\mu_j, \lambda_j^*)$ for all $j \in \{1, 2, \cdots, s\} \backslash \{i\}$, this contradicts with the fact that $\mu_i \in \text{Conv}\{\mu_1, \cdots, \mu_s, \lambda_1, \cdots, \lambda_s\}$.

Combining the statements above, we can see that for all μ_i's that is not one of the vertices of Conv$\{\boldsymbol{\mu}\}$, in order to win the optimization game $T^*(\boldsymbol{\mu})^{-1}$, no proportion of the corresponding arm should be sampled.

A.1.2.2 Potential extension of Thompson-CHM algorithm

As discussed in Section 1.6.2, the Thompson-CHM algorithm outperforms the trivial solution (first checking if γ is smaller than the minimum mean and then checking if γ is larger than the maximum mean) in both generalizability and optimality. For the generalizability part, the Thompson-CHM can possibly generalize to higher dimensional cases and we will discuss more details in this section. For the optimality part, by the results in Kaufmann et al. [2018], the allocations of different arms in the trivial solution are not in align with the optimal $\boldsymbol{w}^*(\boldsymbol{\mu})$ and therefore, lead to a sub-optimal sample complexity compared to the Thompson-CHM algorithm. For example, in the infeasible case when $\gamma < \mu_1 < \cdots < \mu_K$, by utilizing the result in Kaufmann et al. [2018] twice, sampling proportion of arm K is $\frac{2d(\mu_K, \gamma)^{-1}}{\sum_{i=1}^K d(\mu_i, \gamma)^{-1} + d(\mu_K, \gamma)^{-1}}$, and for arm j satisfying $1 \leq j \leq K-1$, its sampling proportion is $\frac{d(\mu_j, \gamma)^{-1}}{\sum_{i=1}^K d(\mu_i, \gamma)^{-1} + d(\mu_K, \gamma)^{-1}}$. This is a direct example of the sub-optimality of the trivial solution to the CHM problem in the one-dimensional case.

Theorem 5 demonstrates an important phenomenon that shares in all dimensions: *in the feasible case, the optimal strategy should only sample arms whose means are extreme points, and in the infeasible case, it should sample all arms.* And if we can prove analogs of the results for the stopping rule, it is possible to fully extend Thompson-CHM algorithm to higher dimensions. We call λ^* the point that is on the boundary of Conv$(\boldsymbol{\mu})$ that minimizes the l_2 distance between γ and γ^*, and we vertically project all the means $\mu_1, \cdots, \mu_K$ to the line that connects γ and γ^*, and denote the projected points to be $\mu_1^*, \cdots, \mu_K^*$, then the d-dimensional distributions with means $\mu_1, \cdots, \mu_K$ are feasible (infeasible) with respect to γ if and only if the 1-dimensional distributions with means $\mu_1^*, \cdots, \mu_K^*$ are feasible (infeasible) with respect to γ on the line that connects γ and γ^*. With this important fact, it is possible to prove our conjecture that the analog of Thompson-CHM (described below) is also asymptotically optimal in higher dimensions using similar techniques in the one-dimensional case.

We now generalize the Thompson-CHM algorithm to higher dimensions by replacing the Bernoulli distribution with a categorical distribution with parameters $\beta_i = f_i(\mu_{s_1}, \cdots, \mu_{s_m}, \gamma)$. Assuming oracle access to the functions f_i for $1 \leq i \leq k$, the analog

of Thompson-CHM algorithm in d-dimensional case is stated below. The future work is to find the exact form of functions f_i and prove the asymptotic optimality of d-dimensional Thompson-CHM algorithm.

Algorithm 7 d-dimensional Thompson-CHM ($d \geq 2$)

Input: stopping rule τ with threshold function $\text{Thresh}(\delta, t)$, risk δ, threshold γ, Categorical distribution parameter $\beta_1, \cdots, \beta_K$.
Output: decision rule $I_\pi(\boldsymbol{\mu}) \in \{\text{feasible}, \text{infeasible}\}$
for $t = 1, \cdots$ **do**
  if stopping rule τ holds **then**
    if $\tau = \tau_3$ **then**
      return $I_\pi(\boldsymbol{\mu}) = \{\text{feasible}\}$
    else
      return $I_\pi(\boldsymbol{\mu}) = \{\text{infeasible}\}$
    end if
  end if
  Sample $\boldsymbol{\theta}_t = (\theta_{t,1}, \cdots, \theta_{t,K}) \sim \Pi_{t-1}(\cdot|\boldsymbol{\mu}$ feasible).
  Sample $B \sim \text{Categorical}\,(\beta_1, \cdots, \beta_K)$
  if $B = i$ **then**
    Play arm $A_t = i$
  end if
end for

APPENDIX B

Appendix for Chapter 2

B.1 Proofs and Supplementary Materials

B.1.1 Proof of Theorem 6

We prove the two cases separately.

Feasible case. Let $y_i = \lambda_i - \gamma$. The closure of the infeasible set in the translated coordinates is

$$\mathrm{cl}(\mathcal{I}) = \bigcup_{u \in \mathbb{S}^{d-1}} \left\{ (y_1, \ldots, y_K) : u^\top y_i \geq 0,\ i = 1, \ldots, K \right\}. \tag{B.1}$$

Indeed, if $0 \notin \mathrm{Conv}(y_1, \ldots, y_K)$, the strong separating hyperplane theorem gives a vector $u \in \mathbb{S}^{d-1}$ such that $u^\top y_i > 0$ for all i. If 0 lies on the boundary of the convex hull, the supporting hyperplane theorem gives $u^\top y_i \geq 0$ for all i. Conversely, if such a u exists, then replacing y_i by $y_i + \varepsilon u$ gives $u^\top (y_i + \varepsilon u) > 0$ for all i, hence the perturbed instance is infeasible. Therefore the original point is in the closure of the infeasible set.

Since the objective in (2.1) is continuous in λ, the infimum over the alternative set equals the infimum over its closure. For fixed u, the least-cost way to move x_i into the halfspace

$$H_u^+ = \{ z \in \mathbb{R}^d : u^\top z \geq 0 \}$$

is the Euclidean projection of x_i onto H_u^+. Thus

$$\inf_{y_i : u^\top y_i \geq 0} \frac{1}{2} \left\| x_i - y_i \right\|^2 = \frac{1}{2} \left(-u^\top x_i \right)_+^2.$$

Summing with weights w_i, and then taking the infimum over $u \in \mathbb{S}^{d-1}$, gives (2.4).

Infeasible case. Now the alternative set is the feasible set. A translated mean vector

$Y = (y_1, \dots, y_K)$ is feasible if and only if there exists $q \in \Delta_K$ such that

$$\sum_{i=1}^{K} q_i y_i = 0.$$

For fixed q, write $y_i = x_i + h_i$. We must solve

$$\inf_{h_1,\dots,h_K} \left\{ \frac{1}{2} \sum_{i=1}^{K} w_i \|h_i\|^2 : \sum_{i=1}^{K} q_i h_i = -\sum_{i=1}^{K} q_i x_i \right\}. \tag{B.2}$$

Assume first that $w_i > 0$ for all i. Let

$$m(q) = \sum_{i=1}^{K} q_i x_i.$$

The Lagrangian is

$$L(h, \eta) = \frac{1}{2} \sum_{i=1}^{K} w_i \|h_i\|^2 + \eta^\top \left(\sum_{i=1}^{K} q_i h_i + m(q) \right).$$

The first-order conditions give

$$w_i h_i + q_i \eta = 0, \qquad h_i = -\frac{q_i}{w_i} \eta.$$

Substituting into the constraint yields

$$-\left(\sum_{i=1}^{K} \frac{q_i^2}{w_i} \right) \eta + m(q) = 0,$$

and therefore

$$\eta = \frac{m(q)}{\sum_{i=1}^{K} q_i^2 / w_i}.$$

The optimal value of (B.2) is

$$\frac{1}{2} \frac{\|m(q)\|^2}{\sum_{i=1}^{K} q_i^2 / w_i}.$$

Taking the infimum over $q \in \Delta_K$ gives (2.5) when w has full support. If some $w_i = 0$, the same formula follows by continuity from allocations $w^\varepsilon = (1 - \varepsilon)w + \varepsilon(1/K, \dots, 1/K)$, with the convention stated in the theorem. This completes the proof.

B.1.2 Proof of Proposition 2.4.1

For a fixed direction $u \in \mathbb{S}^{d-1}$, define

$$\ell_i(u) = \frac{1}{2}\left(-u^\top x_i\right)_+^2.$$

The feasible game is

$$\sup_{w \in \Delta_K} \inf_{u \in \mathbb{S}^{d-1}} \sum_{i=1}^K w_i \ell_i(u).$$

Let x_j be a non-vertex. Then there exist vertices $x_v \in V$ and coefficients $\alpha_v \geq 0$, $\sum_{v \in V} \alpha_v = 1$, such that

$$x_j = \sum_{v \in V} \alpha_v x_v.$$

The map

$$s \mapsto \frac{1}{2}(-s)_+^2$$

is convex on $\mathbb{R}$. Hence, for every $u \in \mathbb{S}^{d-1}$,

$$\ell_j(u) = \frac{1}{2}\left(-u^\top \sum_{v \in V} \alpha_v x_v\right)_+^2 \leq \sum_{v \in V} \alpha_v \frac{1}{2}\left(-u^\top x_v\right)_+^2 = \sum_{v \in V} \alpha_v \ell_v(u).$$

Therefore, if an allocation w assigns mass w_j to arm j, we may remove this mass and redistribute it to the vertices by adding $w_j \alpha_v$ to each vertex v. The quantity

$$\sum_i w_i \ell_i(u)$$

does not decrease for any u. Thus the infimum over u does not decrease. Repeating this argument for all non-vertices gives an optimal allocation supported on vertices.

B.1.3 Proof of Proposition 2.4.3

For any probability measure $\rho \in \mathcal{P}(\mathbb{S}^{d-1})$, define

$$\Phi(w, \rho) = \sum_{i=1}^K w_i \int_{\mathbb{S}^{d-1}} \ell_i(u)\, \rho(\mathrm{d}u).$$

Since the infimum over $u \in \mathbb{S}^{d-1}$ equals the infimum over point masses $\rho = \delta_u$, we have

$$\inf_{u \in \mathbb{S}^{d-1}} \sum_i w_i \ell_i(u) = \inf_{\rho \in \mathcal{P}(\mathbb{S}^{d-1})} \Phi(w, \rho).$$

The function Φ is bilinear in (w, ρ), and both Δ_K and $\mathcal{P}(\mathbb{S}^{d-1})$ are compact convex sets under the usual weak topology. By Sion's minimax theorem,

$$\sup_{w \in \Delta_K} \inf_{\rho \in \mathcal{P}(\mathbb{S}^{d-1})} \Phi(w, \rho) = \inf_{\rho \in \mathcal{P}(\mathbb{S}^{d-1})} \sup_{w \in \Delta_K} \Phi(w, \rho).$$

For fixed ρ,

$$\sup_{w \in \Delta_K} \Phi(w, \rho) = \max_{1 \le i \le K} \int_{\mathbb{S}^{d-1}} \ell_i(u)\, \rho(\mathrm{d}u),$$

which proves (2.6).

Let w^*, ρ^* be a saddle point with value V. Then

$$\Phi(w, \rho^*) \le V \le \Phi(w^*, \rho) \qquad \forall w \in \Delta_K,\ \rho \in \mathcal{P}(\mathbb{S}^{d-1}).$$

The left inequality implies

$$\max_i \int \ell_i(u) \rho^*(\mathrm{d}u) = V.$$

Since

$$\sum_i w_i^* \int \ell_i(u) \rho^*(\mathrm{d}u) = V,$$

any arm with $w_i^* > 0$ must satisfy (2.7). The right inequality, applied to point masses $\rho = \delta_u$, implies

$$\sum_i w_i^* \ell_i(u) \ge V \qquad \forall u \in \mathbb{S}^{d-1}.$$

Since

$$\int \sum_i w_i^* \ell_i(u) \rho^*(\mathrm{d}u) = V,$$

ρ^* can put mass only on directions for which equality holds. This proves (2.8).

B.1.4 Proof of Proposition 2.4.5

Let $\gamma = 0$, $d = 2$, and let the six arms be the vertices of the regular hexagon on the unit circle:

$$x_k = \left(\cos \frac{k\pi}{3}, \sin \frac{k\pi}{3} \right), \qquad k = 0, \dots, 5.$$

Consider the uniform allocation $w_k = 1/6$. For any unit vector u, the six vertices form three antipodal pairs. From each antipodal pair, exactly one squared projection contributes to

$$\sum_{k=0}^{5} \left(-u^\top x_k \right)_+^2,$$

and the three representative directions $0, \pi/3, 2\pi/3$ satisfy

$$\sum_{j=0}^{2} v_j v_j^\top = \frac{3}{2} I_2.$$

Therefore

$$\sum_{k=0}^{5} \left(-u^\top x_k\right)_+^2 = \frac{3}{2} \qquad \forall u \in \mathbb{S}^1.$$

Thus

$$\inf_{u \in \mathbb{S}^1} \frac{1}{2} \sum_{k=0}^{5} \frac{1}{6} \left(-u^\top x_k\right)_+^2 = \frac{1}{8}. \tag{B.3}$$

No allocation can achieve value larger than $1/8$. Indeed, if U is uniform on $\mathbb{S}^1$, then for every fixed unit vector x,

$$\mathbb{E}\left[\left(-U^\top x\right)_+^2\right] = \frac{1}{4}.$$

Hence, for any allocation $w \in \Delta_6$,

$$\inf_{u \in \mathbb{S}^1} \frac{1}{2} \sum_{k=0}^{5} w_k \left(-u^\top x_k\right)_+^2 \le \mathbb{E}_U \left[\frac{1}{2} \sum_{k=0}^{5} w_k \left(-U^\top x_k\right)_+^2\right] = \frac{1}{8}.$$

Combining this upper bound with (B.3), the uniform allocation over all six vertices is optimal.

Now consider any Carathéodory certificate S with $|S| \le 3$. If $|S| \le 2$, or if 0 lies on the boundary of $\mathrm{Conv}\{x_i : i \in S\}$, then there exists a direction u for which $u^\top x_i \ge 0$ for all $i \in S$, and hence $V(S) = 0$. The only three-point subsets of the regular hexagon containing the origin in their interior are the two equilateral triangles. For such a triangle and any allocation $w \in \Delta(S)$, choose $j \in S$ with $w_j = \max_{i \in S} w_i \ge 1/3$, and take $u = x_j$. The other two vertices have inner product $-1/2$ with u, while $u^\top x_j = 1$. Therefore

$$\frac{1}{2} \sum_{i \in S} w_i \left(-u^\top x_i\right)_+^2 = \frac{1}{2} \cdot \frac{1}{4}(1 - w_j) = \frac{1 - w_j}{8} \le \frac{1}{12}.$$

Thus $V(S) \le 1/12$ for every full-dimensional Carathéodory certificate S. Since $1/12 < 1/8$, no single Carathéodory certificate is statistically optimal.

B.1.5 Proof of Proposition 2.5.1

Let

$$\rho = \mathrm{dist}(0, \mathrm{Conv}(x_1, \ldots, x_K)).$$

Since the instance is infeasible and the convex hull is compact, $\rho > 0$.

The uniform allocation $w_i = 1/K$ has strictly positive value. Indeed, for every $q \in \Delta_K$,

$$\left\| \sum_i q_i x_i \right\| \geq \rho,$$

and

$$\sum_i \frac{q_i^2}{1/K} = K \sum_i q_i^2 \leq K.$$

Therefore

$$\inf_{q \in \Delta_K} \frac{\|\sum_i q_i x_i\|^2}{2 \sum_i q_i^2/(1/K)} \geq \frac{\rho^2}{2K} > 0.$$

Thus the optimal value is positive.

Now suppose $w_j = 0$ for some arm j. In (2.5), choose $q = e_j$. Then the denominator contains $q_j^2/w_j = +\infty$, so the value is zero. Hence $C(w, \mu) = 0$, and such a w cannot be optimal because a strictly positive value is achievable. Therefore every optimal allocation has full support.

B.1.6 Proof of Proposition 2.5.2

It is enough to argue at the level of the original variational problem. For fixed w, the infeasible-to-feasible cost is

$$\inf_{\lambda:\, 0 \in \mathrm{Conv}(\lambda_1 - \gamma, \ldots, \lambda_K - \gamma)} \frac{1}{2} \sum_i w_i \|\lambda_i - \mu_i\|^2 . \tag{B.4}$$

For any feasible λ, Carathéodory's theorem gives a vector $q \in \Delta_K$ with

$$|\operatorname{supp}(q)| \leq d + 1, \qquad \sum_i q_i(\lambda_i - \gamma) = 0.$$

Thus every feasible alternative is feasible through a certificate supported on at most $d + 1$ arms. Therefore the infimum in (B.4) is unchanged if one restricts to such certificates. Applying the fixed-q calculation from the proof of Theorem 6 gives (2.11).

B.1.7 Proof of Theorem 8

Let

$$\mathcal{E}_\delta = \left\{ \forall t \geq 1 : \frac{1}{2} \sum_{i=1}^K N_i(t) \|\widehat{\mu}_i(t) - \mu_i\|^2 < \beta(t, \delta) \right\}.$$

By Assumption 7,

$$\mathbb{P}_\mu(\mathcal{E}_\delta) \geq 1 - \delta.$$

On $\mathcal{E}_\delta$, suppose the algorithm stops at time t and makes an incorrect decision. Then the true mean vector μ belongs to the alternative set relative to the empirical answer $\widehat{I}_t$. Therefore, by definition of $Z(t)$,

$$Z(t) \le \frac{1}{2}\sum_{i=1}^{K} N_i(t)\,\|\widehat{\mu}_i(t) - \mu_i\|^2 < \beta(t,\delta).$$

This contradicts the stopping condition $Z(t) \ge \beta(t,\delta)$. Hence no error occurs on $\mathcal{E}_\delta$, and the error probability is at most δ.

B.1.8 Proof of Lemma 2.6.1

ince $N_i(t) \to \infty$, the strong law of large numbers gives

$$\widehat{\mu}_i(t) \to \mu_i \qquad \text{almost surely}$$

for every arm. Hence $\widehat{x}_i(t) \to x_i$.

If the true instance is feasible and regular, then $0 \in \operatorname{int}\operatorname{Conv}(x_1,\dots,x_K)$. Since the empirical points converge to the true points, the empirical instance is feasible for all sufficiently large t. Therefore (2.13) applies eventually, and

$$\frac{Z(t)}{t} = \inf_{u\in\mathbb{S}^{d-1}} \frac{1}{2}\sum_i \frac{N_i(t)}{t}\left(-u^\top \widehat{x}_i(t)\right)_+^2.$$

The functions inside the infimum converge uniformly over the compact sphere $\mathbb{S}^{d-1}$, because they are continuous in u and the coefficients and empirical means converge. This proves (2.18) in the feasible case.

If the true instance is infeasible, then

$$\operatorname{dist}(0, \operatorname{Conv}(x_1,\dots,x_K)) > 0.$$

Again by convergence of the empirical means, the empirical instance is infeasible for all sufficiently large t. Formula (2.14) applies eventually. If w has full support, then

$$\frac{Z(t)}{t} = \inf_{q\in\Delta_K} \frac{\|\sum_i q_i \widehat{x}_i(t)\|^2}{2\sum_i q_i^2/(N_i(t)/t)}.$$

The objective converges uniformly over the compact simplex Δ_K, giving (2.18). If some $w_i = 0$, the same conclusion holds in the extended sense induced by the convention of Theorem 6; for the optimal infeasible allocation this issue does not arise by Proposition 2.5.1.

B.1.9 Proof of Corollary 2.6.2

By Lemma 2.6.1,

$$Z(t) = tC(w, \mu) + o(t) \qquad \text{almost surely.}$$

For every $\varepsilon > 0$, eventually

$$Z(t) \geq t(C(w, \mu) - \varepsilon).$$

Combining this with the asymptotic property (2.16) of $\beta(t, \delta)$, the stopping condition holds for all

$$t \geq \frac{1 + \varepsilon}{C(w, \mu) - \varepsilon} \log(1/\delta)$$

when δ is sufficiently small. Letting $\varepsilon \downarrow 0$ gives (2.19).

B.1.10 Proof of Lemma 2.7.1

The forced-exploration rule guarantees $N_i(t) \to \infty$ for every arm. Hence

$$\widehat{\mu}_i(t) \to \mu_i \qquad \text{almost surely}$$

for all i. By Assumption 9,

$$\widehat{w}(t) \to w^*(\mu) \qquad \text{almost surely.}$$

Therefore

$$\frac{1}{t} \sum_{s=K}^{t} \widehat{w}(s) \to w^*(\mu). \tag{B.5}$$

The number of forced-exploration rounds up to time t is $o(t)$, since each forced pull is used only to keep $N_i(t) \geq \sqrt{t}$. On the non-forced rounds, the algorithm uses cumulative tracking. The standard tracking argument shows that the discrepancy between the actual number of pulls and the cumulative target allocation is $o(t)$:

$$\max_i \left| N_i(t) - \sum_{s=K}^{t} \widehat{w}_i(s) \right| = o(t). \tag{B.6}$$

Indeed, at every non-forced round the arm with largest positive deficit is selected, so the maximal deficit cannot grow linearly; the forced rounds contribute only $o(t)$. Combining (B.5) and (B.6) gives

$$\frac{N(t)}{t} \to w^*(\mu).$$

B.1.11 Proof of Theorem 10

Correctness follows from Theorem 8. By Lemma 2.7.1,

$$\frac{N(t)}{t} \to w^*(\mu) \qquad \text{almost surely.}$$

Applying Corollary 2.6.2 gives (2.20).

The information-theoretic lower bound gives, for every δ-correct algorithm,

$$\liminf_{\delta \downarrow 0} \frac{\mathbb{E}_\mu[\tau_\delta]}{\log(1/\delta)} \geq T_G^*(\mu).$$

The almost-sure upper bound (2.20), together with uniform integrability, gives

$$\limsup_{\delta \downarrow 0} \frac{\mathbb{E}_\mu[\tau_\delta]}{\log(1/\delta)} \leq T_G^*(\mu).$$

Thus (2.21) follows.

B.1.12 Proof of Proposition 2.8.1

Since the true instance is feasible and belongs to the interior of the feasible set, the posterior probability of the conditioning event tends to one:

$$\Pi_{t-1}\left(\gamma \in \mathrm{Conv}(\theta_1, \ldots, \theta_K)\right) \to 1 \qquad \text{almost surely.}$$

Thus the conditional posterior still concentrates at μ. By continuity,

$$w^F(\theta_t) \to w^F(\mu)$$

in posterior probability, conditionally on the past. Therefore the conditional sampling probabilities

$$\psi_i(t) = \mathbb{P}(A_t = i \mid \mathcal{F}_{t-1}) = \mathbb{E}\left[w_i^F(\theta_t) \mid \mathcal{F}_{t-1}, \theta_t \text{ feasible}\right]$$

satisfy

$$\psi_i(t) \to w_i^F(\mu) \qquad \text{almost surely.}$$

The martingale strong law for bounded martingale differences gives

$$\frac{1}{t}\sum_{s=1}^{t} \left(\mathbf{1}\{A_s = i\} - \psi_i(s)\right) \to 0 \qquad \text{almost surely.}$$

Hence

$$\frac{N_i(t)}{t} = \frac{1}{t}\sum_{s=1}^{t} \psi_i(s) + o(1) \to w_i^F(\mu) \qquad \text{almost surely.}$$

The asymptotic optimality statement follows from Corollary 2.6.2.

B.1.13 Proof of Lemma 2.8.2

The first statement is the standard over-allocation criterion used in top-two Thompson sampling and Murphy sampling analyses. Intuitively, (2.26) says that once an arm has been sampled more often than its target proportion by a fixed margin, it receives only summably many additional conditional pulls. This prevents persistent over-allocation of any coordinate and forces $\bar{\psi}(t)$ to converge to $w^*(\mu)$. The convergence of $N(t)/t$ then follows from the martingale strong law for bounded martingale differences:

$$\frac{1}{t}\sum_{s=1}^{t} (\mathbf{1}\{A_s = i\} - \psi_i(s)) \to 0 \qquad \text{almost surely.}$$

B.1.14 Proof of Theorem 13

Fix an arm i and $\varepsilon > 0$. Consider times t such that

$$\bar{\psi}_i(t) \geq w_i^*(\mu) + \varepsilon.$$

By Assumption 12, there exists $c > 0$ such that

$$\inf_{\theta \in \mathrm{cl}(\mathcal{F}_i)} R_{\bar{\psi}(t)}(\theta) \geq \inf_{\theta \in \mathcal{F}} R_{\bar{\psi}(t)}(\theta) + c.$$

By the posterior large-deviation principle,

$$\Pi_t(\mathcal{F}_i) \leq \exp\left(-t\left[\inf_{\theta \in \mathcal{F}} R_{\bar{\psi}(t)}(\theta) + c + o(1)\right]\right),$$

whereas

$$\Pi_t(\mathcal{F}) = \exp\left(-t\left[\inf_{\theta \in \mathcal{F}} R_{\bar{\psi}(t)}(\theta) + o(1)\right]\right).$$

Therefore

$$\mathbb{P}(\theta_t \in \mathcal{F}_i \mid \theta_t \in \mathcal{F}, \mathcal{F}_{t-1}) \leq \exp(-ct + o(t)).$$

Since $g_i(\theta) = 0$ outside $\mathcal{F}_i$ and $0 \leq g_i \leq 1$,

$$\psi_i(t) = \mathbb{E}[g_i(\theta_t) \mid \theta_t \in \mathcal{F}, \mathcal{F}_{t-1}] \leq \exp(-ct + o(t)).$$

Thus

$$\sum_{t=1}^{\infty} \psi_i(t) \mathbf{1}\{\bar{\psi}_i(t) \geq w_i^*(\mu) + \varepsilon\} < \infty.$$

Lemma 2.8.2 gives

$$\bar{\psi}(t) \to w^*(\mu), \qquad \frac{N(t)}{t} \to w^*(\mu) \quad \text{almost surely.}$$

The asymptotic optimality follows from Corollary 2.6.2.

APPENDIX C

Appendix for Chapter 3

C.1 Proofs and Supplementary Materials

In this appendix, we include a more detailed introduction to the oneshot top-k selection mechanism from Qiao et al. [2021].

C.1.1 the Oneshot Differentially Private Mechanism

We recall the one-shot differentially private mechanism Qiao et al. [2021] and the theoretical results needed for our analysis.

Algorithm 8 The oneshot Mechanism $\mathcal{M}^{\text{os}}$ for Privately Reporting Minimum k Elements

Input: database D, functions $f = (f_1, \cdots, f_m)$ with sensitivity s_f, parameter k, and the noise scale λ

Output: indices $i_1, \ldots, i_k$ and approximations to $f_{i_1}(D), \cdots, f_{i_k}(D)$

1: **for** $i = 1$ to m **do**
2: set $y_i = f_i(D) + g_i$ where g_i is sampled i.i.d. from $\text{Lap}(\lambda)$
3: **end for**
4: sort $y_1, \ldots, y_m$ from low to high, $y_{i_1} \leq y_{i_2} \leq \cdots \leq y_{i_m}$
5: return the set $\{i_1, \ldots, i_k\}$ and $f_{i_j}(D) + g'_{i_j}$, where $1 \leq j \leq k$ and g'_{i_j} are fresh independent random noise sampled from $\text{Lap}(\lambda)$

In Qiao et al. [2021], the authors provided the privacy guarantees for both pure differential privacy and (ε, δ)-differential privacy for the mechanism as below.

Theorem 23. *The oneshot mechanism is $(\varepsilon, 0)$-differentially private if we set $\lambda = 2ks_f/\varepsilon$ or larger.*

Theorem 24. *Given $\varepsilon \le 0.2$, $\delta \le 0.05$ and $m \ge 2$, the oneshot mechanism is (ε, δ)-differentially private if we set*

$$\lambda_{\text{oneshot}} = \frac{8 s_f \sqrt{k \log(m/\delta)}}{\varepsilon}$$

or larger.

The TOP_k operator is the *oneshot differentially private top-k* mechanism: it (i) adds a single i.i.d. Laplace vector to $|v|$ to choose the top-k support and then (ii) adds *fresh* independent Laplace noise to the *reported* values on that support (remaining coordinates set to zero). We denote the output of the oneshot top-k algorithm applied to a vector $v \in \mathbb{R}^d$ as $\text{TOP}_k(v; \varepsilon, \delta, \lambda)$, where ε, δ are privacy parameters and λ is the parameter of the Laplace noises in the algorithm.

C.1.2 Proof of Theorem 15 and Theorem 16

We begin by revisiting some of the setting and notations used in the algorithm statements. The dataset is $D = \{(x_i, y_i)\}_{i=1}^n$ with $x_i \in \mathbb{R}^d$, $y_i \in \mathbb{R}$. We call two datasets adjacent $D \sim D'$ if they differ in exactly one record. We assume $\|x_i\|_\infty \le x_{\max}$ for all i, and we use response clipping $y_i^{\text{clip}} := \text{clip}_R(y_i) \in [-R, R]$. The empirical squared loss and its gradient are

$$L(\theta) = \frac{1}{2n} \sum_{i=1}^n \Big(y_i^{\text{clip}} - x_i^\top \theta\Big)^2, \qquad \nabla L(\theta) = -\frac{1}{n} \sum_{i=1}^n \Big(y_i^{\text{clip}} - x_i^\top \theta\Big)\, x_i.$$

Within iteration m, DP–NIHT computes $v = \theta^{(m)} - \eta \nabla L(\theta^{(m)} \mid D)$, then applies the oneshot DP TOP_s operator to the magnitudes $|v|$ using i.i.d. Laplace noise of scale $\lambda_m = \eta B/n$, and finally projects onto the ℓ_1-ball of radius C: $\theta^{(m+1)} = \Pi_C(\widetilde{\theta}^{(m+1)})$. Throughout the M iterations, per-iteration privacy budgets are $(\varepsilon_m, \delta_m) = (\varepsilon_2/M, \delta/M)$ with $\varepsilon_2/M \le 0.2$ and $\delta/M \le 0.05$.

We present a few basic lemmas we will use in the proofs.

Lemma C.1.1 (Per-iteration ℓ_∞ sensitivity of the thresholding input)**.** *Fix an iteration m and assume $\|\theta^{(m)}\|_1 \le C$. Then the map $D \mapsto v(D) = \theta^{(m)} - \eta \nabla L(\theta^{(m)} \mid D)$ obeys*

$$\|v(D) - v(D')\|_\infty \;\le\; s_\infty \quad \textit{for all } D \sim D', \qquad s_\infty := \frac{2\eta}{n}\, x_{\max}\Big(R + x_{\max} C\Big).$$

Lemma C.1.2 (Post-processing invariance)**.** *If a (possibly randomized) map M is (ε, δ)-DP and Φ is any (possibly randomized) map, then $\Phi \circ M$ is also (ε, δ)-DP.*

Lemma C.1.3 (Basic composition)**.** *Suppose we run M mechanisms sequentially on the same dataset D: for each m, conditional on all previous outputs, the m-th mechanism M_m is $(\varepsilon_m, \delta_m)$-DP. Then the joint mechanism (that outputs the entire transcript, or any function thereof) is $\left(\sum_{m=1}^{M} \varepsilon_m,\ \sum_{m=1}^{M} \delta_m\right)$-DP.*

Proof of Theorem 15 (privacy of DP–NIHT). Fix any iteration $m \in \{0, \dots, M-1\}$, by Lemma C.1.1, the vector fed to the oneshot mechanism has ℓ_∞ sensitivity at most $s_\infty = \frac{2\eta}{n} x_{\max}(R + x_{\max} C)$. Because absolute value is 1–Lipschitz, the same s_∞ applies to $|v|$. By construction, the oneshot mechanism in line TOP$_s$ uses $\lambda_m = \eta B/n$. Under the calibration

$$\frac{\eta B}{n} \ \geq\ \frac{8\, s_\infty \sqrt{s\, \log(d/\delta_m)}}{\varepsilon_m} \quad \text{with} \quad (\varepsilon_m, \delta_m) = \left(\frac{\varepsilon_2}{M}, \frac{\delta}{M}\right),$$

By Theorem 24 we have that $D \mapsto \widetilde{\theta}^{(m+1)}$ is $(\varepsilon_2/M, \delta/M)$-DP. The subsequent projection $\theta^{(m+1)} = \Pi_C(\widetilde{\theta}^{(m+1)})$ is a post-processing step, so by Lemma C.1.2 the same per-iteration privacy holds for $\theta^{(m+1)}$. Applying Lemma C.1.3 over $m = 0, \dots, M-1$ yields that the joint transcript $(\theta^{(1)}, \dots, \theta^{(M)})$ is (ε_2, δ)-DP. Finally, releasing only the last iterate $\widehat{\theta}_P = \theta^{(M)}$ is a measurable function of the transcript, hence also (ε_2, δ)-DP by post-processing invariance. Hence we have proved Theorem 15. □

Proof of Theorem 16 (privacy of DP–DS). The DP–DS algorithm releases two objects: a privatized sparsity level $\tilde{c} \in \mathbb{Z}$ and then a DP–NIHT estimate followed by a final keep–c truncation. We prove Theorem 16 by three steps.

Step 1: privacy of $\tilde{c}$. Let $c_{\text{np}}(D) \in \mathbb{Z}$ denote the nonprivate count (number of nonzeros) computed internally from a nonprivate estimator. Define the *clipped* count $c(D) = \min\{\beta, \max\{\alpha, c_{\text{np}}(D)\}\} \in [\alpha, \beta] \cap \mathbb{Z}$. For any $D \sim D'$, $|c(D) - c(D')| \leq \beta - \alpha$ trivially because both values lie in $[\alpha, \beta]$. Thus the global ℓ_1 sensitivity of $c(\cdot)$ is at most $\Delta := \beta - \alpha$. The mechanism releases $\tilde{c} = c(D) + Z$, where Z is *double–geometric* (discrete Laplace) noise with

$$\Pr(Z = z) = \frac{1 - e^{-\varepsilon_1/\Delta}}{1 + e^{-\varepsilon_1/\Delta}}\, e^{-\varepsilon_1 |z|/\Delta}, \qquad z \in \mathbb{Z},$$

optionally followed by re-clipping and rounding. For any two integers t,

$$\frac{\Pr[\tilde{c} = t \mid D]}{\Pr[\tilde{c} = t \mid D']} = \exp\left\{-\frac{\varepsilon_1}{\Delta}\Big(|t - c(D)| - |t - c(D')|\Big)\right\} \ \leq\ \exp\left\{\frac{\varepsilon_1}{\Delta}|c(D) - c(D')|\right\} \ \leq\ e^{\varepsilon_1},$$

where the last inequality uses $|c(D) - c(D')| \leq \Delta$. Therefore $\tilde{c}$ is $(\varepsilon_1, 0)$-DP. Any subsequent clipping/rounding of $\tilde{c}$ is post-processing and does not affect the privacy.

Step 2: privacy of the DP–NIHT stage conditional on $\tilde{c}$. Condition on an arbitrary realized value of $\tilde{c}$ (hence, of $s := c$ used inside DP–NIHT). By Theorem 15, provided the oneshot scale satisfies $\lambda_m = \eta B/n \geq \dfrac{8\, s_\infty \sqrt{s\, \log(dM/\delta)}}{\varepsilon_2/M}$ with $s_\infty = \frac{2\eta}{n} x_{\max}(R + x_{\max} C)$, the map $D \mapsto \widehat{\theta}_P$ is (ε_2, δ)-DP for that fixed s. If B is fixed before the sparsity release, the same conclusion holds by calibrating with $s_{\max} = \min\{d, \lceil \rho\beta \rceil\}$. If B is chosen after the private sparsity has been released, it is a post-processing function of $\tilde{c}$ and the conditional guarantee applies for every realized value.

Step 3: composition and truncation. We have now two sequential releases: first $\tilde{c}$ with $(\varepsilon_1, 0)$-DP, then $\widehat{\theta}_P$ with (ε_2, δ)-DP (conditionally). By Lemma C.1.3, the pair $(\tilde{c}, \widehat{\theta}_P)$ is $(\varepsilon_1 + \varepsilon_2, \delta)$-DP. The final estimator $\widehat{\theta} = \text{Keep-Nonzero}(\widehat{\theta}_P, \tilde{c})$ is a measurable function of $(\tilde{c}, \widehat{\theta}_P)$ and therefore enjoys the same privacy by Lemma C.1.2. □

C.1.3 Proof of Theorem 17

C.1.3.1 A one-shot hard-thresholding inequality

For a vector z, let $S_s(z)$ denote the indices of the s largest entries of $|z|$. If T is the support selected by the one-shot mechanism from the noisy scores $|z_j| + g_j$, write $H_T(z) = z_T$.

Lemma C.1.4 (Noisy hard-thresholding stability)**.** *Let $z, t \in \mathbb{R}^d$ with $\|t\|_0 \leq s^\star \leq s$. Let T be the size-s set selected from $|z| + g$, where $\|g\|_\infty \leq u$, and let ξ_T be any vector supported on T. For any $a, b, c > 0$,*

$$\|H_T(z) + \xi_T - t\|_2^2 \leq A(a,b,c)\, \|z_{T \cup \text{supp}(t)} - t\|_2^2 + C(a,b,c)\, s u^2 + (1+a)\|\xi_T\|_2^2,$$

where

$$A(a,b,c) = \left(1 + \frac{1}{a}\right)\left[(1+b) + \left(1 + \frac{1}{b}\right)(1+c)\frac{s^\star}{s}\right]$$

and $C(a,b,c) > 0$ depends only on a, b, c.

Proof. Let $I = T \cup \text{supp}(t)$ and let S be the exact top-s support of z_I. Since $|I| \leq s + s^\star$, the standard deterministic top-s counting lemma for IHT (see, e.g., Chakraborty et al. [2024]) gives, for every $c > 0$,

$$\|z_{I \setminus S}\|_2^2 \leq (1+c)\frac{s^\star}{s}\, \|z_I - t\|_2^2.$$

The noisy support T can only lose an element of S if it is replaced by an element whose noisy magnitude is at least as large. Pairing indices in $S \setminus T$ with indices in $T \setminus S$ and using

$\|g\|_\infty \le u$ gives

$$\|z_{I\setminus T}\|_2^2 \le (1+c)\|z_{I\setminus S}\|_2^2 + C_c s u^2 \le C_c' \frac{s^\star}{s}\|z_I - t\|_2^2 + C_c s u^2,$$

where the constants depend only on c. Now

$$\|H_T(z) - t\|_2^2 = \|(H_T(z) - z_I) + (z_I - t)\|_2^2 \le (1+b)\|z_I - t\|_2^2 + \Big(1 + \frac{1}{b}\Big)\|z_{I\setminus T}\|_2^2.$$

Finally apply Young's inequality to the fresh value noise,

$$\|H_T(z) + \xi_T - t\|_2^2 \le \Big(1 + \frac{1}{a}\Big)\|H_T(z) - t\|_2^2 + (1+a)\|\xi_T\|_2^2,$$

and combine the preceding displays. □

C.1.3.2 Contractive recursion

Write $S_m = \mathrm{supp}(\theta^{(m)})$, $S^\star = \mathrm{supp}(\theta^\star)$, and T_{m+1} for the support selected at iteration m. Put $I_m = T_{m+1} \cup S^\star$ and $J_m = I_m \cup S_m$, so that $|I_m| \le s + s^\star$ and $|J_m| \le 2s + s^\star = k_0$.

Lemma C.1.5 (One-step bound)**.** *Assume* (A1)–(A4) *and work on* $\mathcal{E}$*. If* $s \ge C_0\kappa(k_0)^2 s^\star$ *for a sufficiently large universal* C_0*, then there exist universal constants* $c, C > 0$ *such that for all* m*,*

$$\begin{aligned}\left\|\theta^{(m+1)} - \theta^\star\right\|_2^2 \le \left(1 - \frac{c}{\kappa(k_0)}\right)\left\|\theta^{(m)} - \theta^\star\right\|_2^2 + C\kappa(k_0)\eta^2(2s + s^\star)\bar{g}^2 \\ + C\kappa(k_0)s\lambda^2\tau^2 + C\kappa(k_0)\left\|\xi^{(m)}_{T_{m+1}}\right\|_2^2.\end{aligned} \tag{C.1}$$

Proof. Because $C \ge \|\theta^\star\|_1$, Euclidean projection onto the ℓ_1 ball is non-expansive around $\theta^\star$:

$$\|\theta^{(m+1)} - \theta^\star\|_2 \le \|H_{T_{m+1}}(v^{(m+1/2)}) + \xi^{(m)}_{T_{m+1}} - \theta^\star\|_2.$$

Apply Lemma C.1.4 with $z = v^{(m+1/2)}$, $t = \theta^\star$, and $u = \lambda\tau$ on $\mathcal{E}_{\mathrm{sel}}$. Choose $a \asymp \kappa(k_0)$ and $b \asymp 1/\kappa(k_0)$. Since $s \ge C_0\kappa(k_0)^2 s^\star$, the coefficient $A(a,b,c)$ in Lemma C.1.4 is at most $1 + c_1/\kappa(k_0)$ for a sufficiently small numerical constant $c_1 > 0$ by taking C_0 large enough. Thus

$$\|\theta^{(m+1)} - \theta^\star\|_2^2 \le \left(1 + \frac{c_1}{\kappa(k_0)}\right)\|v^{(m+1/2)}_{I_m} - \theta^\star\|_2^2 + C\kappa(k_0)s\lambda^2\tau^2 + C\kappa(k_0)\|\xi^{(m)}_{T_{m+1}}\|_2^2. \tag{C.2}$$

For the gradient step, the vector $\theta^{(m)} - \theta^\star$ is supported on $S_m \cup S^\star \subseteq J_m$. Since $|J_m| \le k_0$

and $\eta = 1/(2\overline{\kappa}(k_0))$,

$$\|(I - \eta\Sigma)(\theta^{(m)} - \theta^\star)\|_{2,I_m} \le \left(1 - \frac{1}{2\kappa(k_0)}\right) \|\theta^{(m)} - \theta^\star\|_2.$$

Moreover, on $\mathcal{E}_\nabla$, $\|\nabla L(\theta^\star)_{I_m}\|_2^2 \le |I_m|\bar{g}^2 \le (s+s^\star)\bar{g}^2$. Hence, for a numerical constant $C > 0$,

$$\|v_{I_m}^{(m+1/2)} - \theta^\star\|_2^2 \le \left(1 - \frac{3c_2}{\kappa(k_0)}\right) \|\theta^{(m)} - \theta^\star\|_2^2 + C\kappa(k_0)\eta^2(s+s^\star)\bar{g}^2,$$

with $c_2 > 0$ universal. Combining this display with (C.2), and taking c_1 small relative to c_2, yields (C.1). □

Lemma C.1.6 (Unrolling)**.** *Under the conditions of Lemma C.1.5, for $M \ge 1$,*

$$\left\|\theta^{(M)} - \theta^\star\right\|_2^2 \le \left(1 - \frac{c}{\kappa(k_0)}\right)^M \left\|\theta^{(0)} - \theta^\star\right\|_2^2 + C\frac{(s+s^\star)\bar{g}^2}{\underline{\kappa}(k_0)^2} + C\kappa(k_0)^2 s\lambda^2\tau^2 + C\kappa(k_0)^2 \max_{0\le m<M} \left\|\xi_{T_{m+1}}^{(m)}\right\|_2^2. \tag{C.3}$$

Proof. Iterate (C.1) and use $\sum_{j\ge 0}(1 - c/\kappa)^j \le C\kappa$. The gradient term is bounded as $\kappa(k_0)^2\eta^2(s+s^\star)\bar{g}^2 \le C(s+s^\star)\bar{g}^2/\underline{\kappa}(k_0)^2$. For the value-noise term, the additional factor κ in (C.1) and the geometric sum together give the multiplier κ^2 in (C.3). □

C.1.3.3 Proof of the main theorem

From (C.3) and $\theta^{(0)} = 0$,

$$\left\|\theta^{(M)} - \theta^\star\right\|_2^2 \le \left(1 - \frac{c}{\kappa}\right)^M \|\theta^\star\|_2^2 + C\frac{(s+s^\star)\bar{g}^2}{\underline{\kappa}(k_0)^2} + C\kappa^2 s\lambda^2\tau^2 + C\kappa^2 \max_{0\le m<M} \left\|\xi_{T_{m+1}}^{(m)}\right\|_2^2,$$

where $\kappa = \kappa(k_0)$. Since $\|\theta^\star\|_2^2 \le \|\theta^\star\|_1^2 \le b_{\max}^2$, choosing $M \ge c_0\kappa\log(b_{\max}^2 n)$ makes the geometric tail at most $1/n$. On $\mathcal{E}_{\mathrm{val}}$,

$$\max_{0\le m<M} \left\|\xi_{T_{m+1}}^{(m)}\right\|_2^2 \le c_\xi s\lambda^2\tau_\xi.$$

The projection of an s-sparse vector onto an ℓ_1 ball remains supported on the same coordinates, so the iterates are s-sparse and $H_s(\theta^{(M)}) = \theta^{(M)}$. Absorbing constants gives (3.3).

C.1.4 Proof of Lemma 3.6.1

The lemma follows immediately from the quadratic form of the population risk. If $h = \theta - \theta^\star$ and $\|h\|_0 \le k$, then

$$F(\theta) - F(\theta^\star) = \frac{1}{2} h^\top \Sigma_{\text{pop}} h \le \frac{L_k^{\text{pop}}}{2} \|h\|_2^2.$$

C.1.5 Proof of Theorem 18

Since $\widehat{\theta}$ is s-sparse and $\theta^\star$ is $s^\star$-sparse, the difference $\widehat{\theta} - \theta^\star$ is supported on at most $s + s^\star \le 4s$ coordinates. Lemma 3.6.1 therefore gives

$$F(\widehat{\theta}) - F(\theta^\star) \le \frac{L_{4s}^{\text{pop}}}{2} \|\widehat{\theta} - \theta^\star\|_2^2.$$

Apply Theorem 17 with the empirical restricted eigenvalue constants taken at level $4s$ (which is admissible since $2s + s^\star \le 4s$):

$$\|\widehat{\theta} - \theta^\star\|_2^2 \le \frac{1}{n} + C_4 \frac{(s + s^\star)\bar{g}^2}{(\mu_{4s}^{(n)})^2} + C_5 (\kappa_{4s}^{(n)})^2 s \lambda^2 \tau^2 + C_6 (\kappa_{4s}^{(n)})^2 s \lambda^2 \tau_\xi.$$

Because $s \ge s^\star$, $(s + s^\star) \le 2s$. Under the sub-Gaussian gradient event,

$$\bar{g}^2 \lesssim \sigma^2 x_{\max}^2 \frac{\log(d/\alpha)}{n}.$$

Substituting this bound into the preceding display and multiplying by $L_{4s}^{\text{pop}}/2$ yields (3.4). The corollary follows by substituting $\lambda = B/(2L_{4s}^{(n)} n)$ and the displayed choices of τ and τ_ξ. □

APPENDIX D

Appendix for Chapter 4

D.1 Proofs and Supplementary Materials

D.1.1 Proof of Theorem 19

Fix an iteration m and condition on the transcript before the Dantzig-score selection step. Then $\theta^{(m)}$ is fixed. For adjacent data sets D, D' differing in one record,

$$\begin{aligned}\left\|q_D(\theta^{(m)}) - q_{D'}(\theta^{(m)})\right\|_\infty &\le \frac{1}{n}\left(|y_i^R - x_i^\top\theta^{(m)}|\|x_i\|_\infty + |y_i'^R - x_i'^\top\theta^{(m)}|\|x_i'\|_\infty\right)\\ &\le \frac{2x_{\max}(R + x_{\max}C)}{n} = \Delta_q.\end{aligned}$$

The absolute-value map is 1-Lipschitz in ℓ_∞, so the same sensitivity bound applies to $|q_D(\theta^{(m)})|$. The one-shot top-$2s$ mechanism with scale b_I is therefore $(\varepsilon_I/M, \delta_I/M)$-DP by Theorem 6 of Appendix A.1.

Next condition on the active set U_m produced by the previous private step. For this fixed set, $|U_m| \le 3s$. The query

$$D \mapsto (G_{U_mU_m}, z_{U_m})$$

enjoys the ℓ_2 sensitivity bound

$$\begin{aligned}&\frac{1}{n}\left(\left\|x_{i,U_m}x_{i,U_m}^\top - x'_{i,U_m}{x'_{i,U_m}}^\top\right\|_F^2 + \left\|x_{i,U_m}y_i^R - x'_{i,U_m}y_i'^R\right\|_2^2\right)^{1/2}\\ &\quad\le \frac{1}{n}\left(\left(\|x_{i,U_m}\|_2^2 + \|x'_{i,U_m}\|_2^2\right)^2 + \left(R\|x_{i,U_m}\|_2 + R\|x'_{i,U_m}\|_2\right)^2\right)^{1/2}\\ &\quad\le \frac{2K_{3s}\sqrt{K_{3s}^2 + R^2}}{n} = \Delta_F.\end{aligned}$$

Thus the Gaussian mechanism with standard deviation σ_F in (4.5) is $(\varepsilon_F/M, \delta_F/M)$-DP for the restricted score-map release. The restricted refit and the pruning step are post-

processing.

By adaptive composition over M iterations, the full transcript is $(\varepsilon_I + \varepsilon_F, \delta_I + \delta_F)$-DP. Releasing only $\widehat{\theta}$ is another post-processing step, completing the proof.

D.1.2 Proof of Lemma 4.2.1

The first display is exactly the operator-norm form of (4.2) on the principal submatrix G_{TT}. For the second display, apply the first display on $T = A \cup B$ to a vector supported on B and then restrict the result to the coordinates in A. Since A and B are disjoint, the identity part contributes nothing on A.

D.1.3 Proof of Lemma 4.2.2

Let $A = T \setminus J$ and $B = J \setminus T$. Since $|J| = k$ and $|T| \le k$, we have $|B| \ge |A|$. Choose any subset $B_0 \subset B$ with $|B_0| = |A|$ and pair the elements of A and B_0 so that each paired (j, i) satisfies the top-k ordering relation

$$|a_j| + g_j \le |a_i| + g_i.$$

Hence $|a_j| \le |a_i| + 2\omega$ for every pair. Squaring and summing is not necessary; directly,

$$\|a_A\|_2 \le \|a_{B_0}\|_2 + 2\sqrt{|A|}\,\omega \le \|a_B\|_2 + 2\sqrt{k}\,\omega,$$

which proves the claim.

D.1.4 Proof of Lemma 4.2.3

Let $A_m = \mathrm{supp}(h^{(m)})$. Since $\theta^{(m)}$ is s-sparse and $s \ge s^\star$, $|A_m| \le 2s$. On $\mathcal{E}_R$,

$$q^{(m)} = z - G\theta^{(m)} = e - Gh^{(m)}.$$

Apply Lemma 4.2.2 to $a = q^{(m)}$, $T = A_m$, $k = 2s$, and $J = J_m$. With $A = A_m \setminus J_m$ and $B = J_m \setminus A_m$,

$$\|q_A^{(m)}\|_2 \le \|q_B^{(m)}\|_2 + 2\sqrt{2s}\,\omega_I. \tag{D.1}$$

For the left-hand side, since $A \subset A_m$,

$$\begin{aligned}\|q_A^{(m)}\|_2 &= \|e_A - (Gh^{(m)})_A\|_2 \\ &\geq \|h_A^{(m)}\|_2 - \|e_A\|_2 - \|((G-I)h^{(m)})_A\|_2 \\ &\geq \|h_A^{(m)}\|_2 - \sqrt{2s}\lambda_n - \delta_{4s}\|h^{(m)}\|_2.\end{aligned}$$

For the right-hand side, $B \cap A_m = \varnothing$, so $h_B^{(m)} = 0$ and

$$\begin{aligned}\|q_B^{(m)}\|_2 &\leq \|e_B\|_2 + \|(Gh^{(m)})_B\|_2 \\ &= \|e_B\|_2 + \|((G-I)h^{(m)})_B\|_2 \\ &\leq \sqrt{2s}\lambda_n + \delta_{4s}\|h^{(m)}\|_2.\end{aligned}$$

Combining these bounds with (D.1) yields

$$\|h_A^{(m)}\|_2 \leq 2\delta_{4s}\|h^{(m)}\|_2 + 2\sqrt{2s}\lambda_n + 2\sqrt{2s}\omega_I.$$

Finally, $\operatorname{supp}(\theta^{(m)}) \subset U_m$, and therefore $h_{U_m^c}^{(m)} = -\theta_{U_m^c}^\star$. Moreover, $U_m^c \cap \operatorname{supp}(h^{(m)}) \subset A_m \setminus J_m = A$. Hence $\|h_{U_m^c}^{(m)}\|_2 \leq \|h_A^{(m)}\|_2$, proving the lemma.

D.1.5 Proof of Lemma 4.2.4

Let $\beta_U^\star = \theta_U^\star$. Since $\|\theta^\star\|_1 \leq C$, $\beta_U^\star$ is feasible for the refit problem. On $\mathcal{E}_R$,

$$z_U - G_{UU}\beta_U^\star = G_{UU^c}\theta_{U^c}^\star + e_U.$$

By Lemma 4.2.1, valid because $|U| + |\operatorname{supp}(\theta_{U^c}^\star)| \leq 4s$,

$$\|G_{UU^c}\theta_{U^c}^\star\|_2 \leq \delta_{4s}\|\theta_{U^c}^\star\|_2.$$

Also $\|e_U\|_2 \leq \sqrt{3s}\lambda_n$. Hence, on $\mathcal{E}_F$,

$$\begin{aligned}\|\widetilde{z}_m - \widetilde{G}_m\beta_U^\star\|_2 &\leq \delta_{4s}\|\theta_{U^c}^\star\|_2 + \sqrt{3s}\lambda_n + \|v_m\|_2 + \|W_m\beta_U^\star\|_2 \\ &\leq \delta_{4s}\|\theta_{U^c}^\star\|_2 + \sqrt{3s}\lambda_n + \omega_F.\end{aligned}$$

By optimality of $\widetilde{\beta}_m$, its noisy residual is no larger than the preceding display. Therefore

$$\begin{aligned}\|\widetilde{G}_m(\widetilde{\beta}_m - \beta_U^\star)\|_2 &\le \|\widetilde{z}_m - \widetilde{G}_m\widetilde{\beta}_m\|_2 + \|\widetilde{z}_m - \widetilde{G}_m\beta_U^\star\|_2 \\ &\le 2\left(\delta_{4s}\|\theta^\star_{U^c}\|_2 + \sqrt{3s}\lambda_n + \omega_F\right).\end{aligned}$$

The singular values of G_{UU} are between $1-\delta_{3s}$ and $1+\delta_{3s}$. Since $\|W_m\|_{\mathrm{op}} \le (1-\delta_{3s})/2$,

$$s_{\min}(\widetilde{G}_m) \ge 1 - \delta_{3s} - \|W_m\|_{\mathrm{op}} \ge \frac{1-\delta_{3s}}{2}.$$

It follows that

$$\|\widetilde{\beta}_m - \beta_U^\star\|_2 \le c\left(\delta_{4s}\|\theta^\star_{U^c}\|_2 + \sqrt{s}\lambda_n + \omega_F\right).$$

Adding the tail $\|\theta^\star_{U^c}\|_2$ gives (4.15); the constant c_3 absorbs the factor $1 + c\delta_{4s}$.

D.1.6 Proof of Lemma 4.2.5

Let S be the support of $H_s(b)$. Since $H_s(b)$ is the best s-term approximation to b in Euclidean norm and $\theta^\star$ is feasible for the same approximation problem,

$$\|b - H_s(b)\|_2 \le \|b - \theta^\star\|_2.$$

The triangle inequality gives

$$\|H_s(b) - \theta^\star\|_2 \le \|H_s(b) - b\|_2 + \|b - \theta^\star\|_2 \le 2\|b - \theta^\star\|_2.$$

D.1.7 Proof of Theorem 20

We work on $\mathcal{E}_R \cap \mathcal{E}_n \cap \mathcal{E}_I \cap \mathcal{E}_F$. Combining Lemmas 4.2.3, 4.2.4, and 4.2.5, we obtain

$$\begin{aligned}\|h^{(m+1)}\|_2 &= \|\theta^{(m+1)} - \theta^\star\|_2 \\ &\le 2\|\bar{\theta}^{(m+1)} - \theta^\star\|_2 \\ &\le C\|\theta^\star_{U_m^c}\|_2 + C\sqrt{s}\lambda_n + C\omega_F \\ &= C\|h^{(m)}_{U_m^c}\|_2 + C\sqrt{s}\lambda_n + C\omega_F \\ &\le C\delta_{4s}\|h^{(m)}\|_2 + C\sqrt{s}(\lambda_n + \omega_I) + C\omega_F.\end{aligned}$$

Choose the universal constant δ_0 in the theorem small enough that $C\delta_{4s} \le 1/2$. Then

$$\|h^{(m+1)}\|_2 \le \frac{1}{2}\|h^{(m)}\|_2 + C\sqrt{s}(\lambda_n + \omega_I) + C\omega_F. \tag{D.2}$$

Unrolling (D.2) and using $\theta^{(0)} = 0$ gives

$$\|h^{(M)}\|_2 \le 2^{-M}\|\theta^\star\|_2 + C\sqrt{s}(\lambda_n + \omega_I) + C\omega_F.$$

Since $\|\theta^\star\|_2 \le \|\theta^\star\|_1 \le C$, the choice $M \ge c_0 \log(C^2 n)$ ensures $2^{-2M}\|\theta^\star\|_2^2 \le 1/n$. Squaring the preceding display and increasing constants yields (4.11). The probability statement follows from the bounds for $\mathcal{E}_n$, $\mathcal{E}_I$, and $\mathcal{E}_F$, and from adding $\mathbb{P}(\mathcal{E}_R^c)$ if clipping may occur.

D.1.8 Proof of Theorem 21

By the definition of $\bar{\theta}^{(m+1)}$ and the event $\mathcal{E}_F$,

$$\begin{aligned}\|q_D(\bar{\theta}^{(m+1)})_{U_m}\|_2 &= \|z_{U_m} - G_{U_m U_m}\widetilde{\beta}_m\|_2 \\ &\le \|\widetilde{z}_m - \widetilde{G}_m\widetilde{\beta}_m\|_2 + \|v_m\|_2 + C\|W_m\|_{\mathrm{op}} \\ &\le \|\widetilde{z}_m - \widetilde{G}_m\theta^\star_{U_m}\|_2 + \omega_F.\end{aligned}$$

The same calculation as in the proof of Lemma 4.2.4 gives

$$\|\widetilde{z}_m - \widetilde{G}_m\theta^\star_{U_m}\|_2 \le \delta_{4s}\|\theta^\star_{U_m^c}\|_2 + \sqrt{3s}\lambda_n + \omega_F.$$

Using $\theta^\star_{U_m^c} = -h^{(m)}_{U_m^c}$ and Lemma 4.2.3 proves (4.19). The final display follows by applying the recursion (D.2) up to iteration $M - 1$.

D.1.9 Proof of Theorem 22

The claim follows from

$$F(\widehat{\theta}) - F(\theta^\star) = \frac{1}{2}(\widehat{\theta} - \theta^\star)^\top \Sigma_{\mathrm{pop}}(\widehat{\theta} - \theta^\star) \le \frac{L^{\mathrm{pop}}_{2s}}{2}\|\widehat{\theta} - \theta^\star\|_2^2,$$

and Theorem 20.

BIBLIOGRAPHY

M. Abadi, A. Chu, I. Goodfellow, H. B. McMahan, I. Mironov, K. Talwar, and L. Zhang. Deep learning with differential privacy. In *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*, pages 308–318, 2016.

J. Abernethy, P. Awasthi, M. Kleindessner, J. Morgenstern, and J. Zhang. Adaptive sampling to reduce disparate performance. *arXiv e-prints*, pages arXiv–2006, 2020.

S. Agrawal and N. Goyal. Thompson sampling for contextual bandits with linear payoffs. In *International conference on machine learning*, pages 127–135. PMLR, 2013.

H. Anahideh, A. Asudeh, and S. Thirumuruganathan. Fair active learning. *Expert Systems with Applications*, 199:116981, 2022.

P. Auer. Using confidence bounds for exploitation-exploration trade-offs. *Journal of Machine Learning Research*, 3(Nov):397–422, 2002.

P. Auer, N. Cesa-Bianchi, and P. Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine learning*, 47(2):235–256, 2002.

A. Barvinok. *A course in convexity*, volume 54. American Mathematical Society, 2025.

R. Bassily, A. Smith, and A. Thakurta. Private empirical risk minimization: Efficient algorithms and tight error bounds. In *2014 IEEE 55th annual symposium on foundations of computer science*, pages 464–473. IEEE, 2014.

H. Bastani and M. Bayati. Online decision making with high-dimensional covariates. *Operations Research*, 68(1):276–294, 2020.

P. J. Bickel, Y. Ritov, and A. B. Tsybakov. Simultaneous analysis of lasso and dantzig selector. 2009.

D. Bouneffouf, A. Bouzeghoub, and A. L. Gançarski. A contextual-bandit algorithm for mobile context-aware recommender system. In *International conference on neural information processing*, pages 324–331. Springer, 2012.

D. Bouneffouf, A. Bouzeghoub, and A. L. Gançarski. Contextual bandits for context-based information retrieval. In *International Conference on Neural Information Processing*, pages 35–42. Springer, 2013.

D. Bouneffouf, R. Laroche, T. Urvoy, R. Féraud, and R. Allesiardo. Contextual bandit for active learning: Active thompson sampling. In *International Conference on Neural Information Processing*, pages 405–412. Springer, 2014.

T. T. Cai, Y. Wang, and L. Zhang. The cost of privacy: Optimal rates of convergence for parameter estimation with differential privacy. *The Annals of Statistics*, 49(5):2825–2850, 2021.

E. Candes and T. Tao. The dantzig selector: Statistical estimation when p is much larger than n. 2007.

E. J. Candes, J. K. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 59(8):1207–1223, 2006.

S. Chakraborty, S. Roy, and D. Basu. Fliphat: Joint differential privacy for high dimensional sparse linear bandits. *arXiv preprint arXiv:2405.14038*, 2024.

O. Chapelle and L. Li. An empirical evaluation of thompson sampling. *Advances in neural information processing systems*, 24, 2011.

K. Chaudhuri, C. Monteleoni, and A. D. Sarwate. Differentially private empirical risk minimization. *Journal of Machine Learning Research*, 12(3), 2011.

L. Chen, J. Li, and M. Qiao. Towards instance optimal bounds for best arm identification. In *Conference on Learning Theory*, pages 535–592. PMLR, 2017.

S. Chen, T. Lin, I. King, M. R. Lyu, and W. Chen. Combinatorial pure exploration of multi-armed bandits. *Advances in neural information processing systems*, 27, 2014.

S.-B. Chen, C. Ding, B. Luo, and Y. Xie. Uncorrelated lasso. In *Proceedings of the AAAI conference on artificial intelligence*, volume 27, pages 166–172, 2013.

H. Chernoff. Sequential design of experiments. *The Annals of Mathematical Statistics*, 30 (3):755–770, 1959.

W. Chu, L. Li, L. Reyzin, and R. Schapire. Contextual bandits with linear payoff functions. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, pages 208–214. JMLR Workshop and Conference Proceedings, 2011.

A. Dandekar, D. Basu, and S. Bressan. Differential privacy for regularised linear regression. In *International Conference on Database and Expert Systems Applications*, pages 483–491. Springer, 2018.

R. Degenne and W. M. Koolen. Pure exploration with multiple correct answers. *Advances in Neural Information Processing Systems*, 32, 2019.

R. Degenne, W. M. Koolen, and P. Ménard. Non-asymptotic pure exploration by solving games. *Advances in Neural Information Processing Systems*, 32, 2019.

R. Degenne, P. Ménard, X. Shang, and M. Valko. Gamification of pure exploration for linear bandits. In *International Conference on Machine Learning*, pages 2432–2442. PMLR, 2020.

D. L. Donoho. Compressed sensing. *IEEE Transactions on information theory*, 52(4):1289–1306, 2006.

M. Dudik, D. Hsu, S. Kale, N. Karampatziakis, J. Langford, L. Reyzin, and T. Zhang. Efficient optimal learning for contextual bandits. *arXiv preprint arXiv:1106.2369*, 2011.

A. Durand, C. Achilleos, D. Iacovides, K. Strati, G. D. Mitsis, and J. Pineau. Contextual bandits for adapting treatment in a mouse model of de novo carcinogenesis. In *Machine learning for healthcare conference*, pages 67–82. PMLR, 2018.

C. Dwork, K. Kenthapadi, F. McSherry, I. Mironov, and M. Naor. Our data, ourselves: Privacy via distributed noise generation. In *Annual international conference on the theory and applications of cryptographic techniques*, pages 486–503. Springer, 2006a.

C. Dwork, F. McSherry, K. Nissim, and A. Smith. Calibrating noise to sensitivity in private data analysis. In *Theory of cryptography conference*, pages 265–284. Springer, 2006b.

C. Dwork, F. McSherry, K. Nissim, and A. Smith. Calibrating noise to sensitivity in private data analysis. In *Theory of cryptography conference*, pages 265–284. Springer, 2006c.

B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani. Least angle regression. 2004.

R. Filippozzi, D. S. Gonçalves, and L.-R. Santos. First-order methods for the convex hull membership problem. *European Journal of Operational Research*, 306(1):17–33, 2023.

A. Garivier and E. Kaufmann. Optimal best arm identification with fixed confidence. In *Conference on Learning Theory*, pages 998–1027. PMLR, 2016.

A. Garivier, P. Ménard, L. Rossi, and P. Menard. Thresholding bandit for dose-ranging: The impact of monotonicity. *arXiv preprint arXiv:1711.04454*, 2017.

J. Hsu, A. Roth, T. Roughgarden, and J. Ullman. Privately solving linear programs. In *International Colloquium on Automata, Languages, and Programming*, pages 612–624. Springer, 2014.

X. Huo and F. Fu. Risk-aware multi-armed bandit problem with application to portfolio selection. *Royal Society open science*, 4(11):171377, 2017.

M. Jayaram and H. Fleyeh. Convex hulls in image processing: a scoping review. *American Journal of Intelligent Systems*, 6(2):48–58, 2016.

H. Kano, J. Honda, K. Sakamaki, K. Matsuura, A. Nakamura, and M. Sugiyama. Good arm identification via bandit feedback. *Machine Learning*, 108(5):721–745, 2019.

N. Katzin. Convex hull as a heuristic. 2018.

E. Kaufmann and W. M. Koolen. Mixture martingales revisited with applications to sequential tests and confidence intervals. *J. Mach. Learn. Res.*, 22:246–1, 2021.

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, 2016.

E. Kaufmann, W. M. Koolen, and A. Garivier. Sequential test for the lowest mean: From thompson to murphy sampling. *Advances in Neural Information Processing Systems*, 31, 2018.

A. Khanna, F. Lu, E. Raff, and B. Testa. Sparse private lasso logistic regression. *arXiv preprint arXiv:2304.12429*, 2023.

D. Kifer, A. Smith, and A. Thakurta. Private convex empirical risk minimization and high-dimensional regression. In *Conference on Learning Theory*, pages 25–1. JMLR Workshop and Conference Proceedings, 2012.

T. L. Lai, H. Robbins, et al. Asymptotically efficient adaptive allocation rules. *Advances in applied mathematics*, 6(1):4–22, 1985.

J. Langford and T. Zhang. The epoch-greedy algorithm for multi-armed bandits with side information. *Advances in neural information processing systems*, 20, 2007.

J. Lengyel, M. Reichert, B. R. Donald, and D. P. Greenberg. Real-time robot motion planning using rasterizing computer graphics hardware. *ACM Siggraph Computer Graphics*, 24(4): 327–335, 1990.

J. Li and L. Lu. A novel differentially private online learning algorithm for group lasso in big data. *IET Information Security*, 2024(1):5553292, 2024.

L. Li, W. Chu, J. Langford, and R. E. Schapire. A contextual-bandit approach to personalized news article recommendation. In *Proceedings of the 19th international conference on World wide web*, pages 661–670, 2010.

X. Lin, H.-L. Zhen, Z. Li, Q.-F. Zhang, and S. Kwong. Pareto multi-task learning. *Advances in neural information processing systems*, 32, 2019.

A. Locatelli, M. Gutzeit, and A. Carpentier. An optimal algorithm for the thresholding bandit problem. In *International Conference on Machine Learning*, pages 1690–1698. PMLR, 2016.

D. E. Losada, J. Parapar, and A. Barreiro. Multi-armed bandits for adjudicating documents in pooling-based evaluation of information retrieval systems. *Information Processing & Management*, 53(5):1005–1025, 2017.

O. L. Mangasarian. Privacy-preserving linear programming. *Optimization Letters*, 5(1): 165–172, 2011.

N. Martinez, M. Bertran, and G. Sapiro. Minimax pareto fairness: A multi objective perspective. In *International Conference on Machine Learning*, pages 6755–6764. PMLR, 2020.

N. Meinshausen. Relaxed lasso. *Computational Statistics & Data Analysis*, 52(1):374–393, 2007.

K. Misra, E. M. Schwartz, and J. Abernethy. Dynamic online pricing with incomplete information using multiarmed bandit experiments. *Marketing Science*, 38(2):226–252, 2019.

J. W. Mueller, V. Syrgkanis, and M. Taddy. Low-rank bandit methods for high-dimensional dynamic pricing. *Advances in Neural Information Processing Systems*, 32, 2019.

F. Nargesian, A. Asudeh, and H. Jagadish. Tailoring data source distributions for fairness-aware data integration. *Proceedings of the VLDB Endowment*, 14(11):2519–2532, 2021.

T. T. Nguyen. On the edge and cloud: Recommendation systems with distributed machine learning. In *2021 International Conference on Information Technology (ICIT)*, pages 929–934. IEEE, 2021.

L. Niss, Y. Sun, and A. Tewari. Achieving representative data via convex hull feasibility sampling algorithms. *arXiv preprint arXiv:2204.06664*, 2022.

G. Qiao, W. Su, and L. Zhang. Oneshot differentially private top-k selection. In *International Conference on Machine Learning*, pages 8672–8681. PMLR, 2021.

H. Robbins. Some aspects of the sequential design of experiments. *Bulletin of the American Mathematical Society*, 58(5):527–535, 1952.

R. T. Rockafellar. *Convex analysis*, volume 28. Princeton university press, 1997.

P. P. Roy, U. Pal, J. Lladós, and F. Kimura. Convex hull based approach for multi-oriented character recognition from graphical documents. In *2008 19th international conference on pattern recognition*, pages 1–4. IEEE, 2008.

D. Russo. Simple bayesian algorithms for best arm identification. In *Conference on Learning Theory*, pages 1417–1418. PMLR, 2016.

W. Shen, J. Wang, Y.-G. Jiang, and H. Zha. Portfolio choices with orthogonal bandit learning. In *Twenty-fourth international joint conference on artificial intelligence*, 2015.

M. Sion. On general minimax theorems. 1958.

N. Srinivas, A. Krause, S. M. Kakade, and M. Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. *arXiv preprint arXiv:0912.3995*, 2009.

I. Streinu. A combinatorial approach to planar non-colliding robot arm motion planning. In *Proceedings 41st Annual Symposium on Foundations of Computer Science*, pages 443–453. IEEE, 2000.

K. Talwar, A. Guha Thakurta, and L. Zhang. Nearly optimal private lasso. *Advances in Neural Information Processing Systems*, 28, 2015.

L. Tang, R. Rosales, A. Singh, and D. Agarwal. Automatic ad format selection via contextual bandits. In *Proceedings of the 22nd ACM international conference on Information & Knowledge Management*, pages 1587–1594, 2013.

C. Tao, S. Blanco, J. Peng, and Y. Zhou. Thresholding bandit with optimal aggregate regret. *Advances in Neural Information Processing Systems*, 32, 2019.

A. G. Thakurta and A. Smith. Differentially private feature selection via stability arguments, and the robustness of the lasso. In *Conference on Learning Theory*, pages 819–850. PMLR, 2013.

R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(1):267–288, 1996.

M. Valko, N. Korda, R. Munos, I. Flaounas, and N. Cristianini. Finite-time analysis of kernelised contextual bandits. *arXiv preprint arXiv:1309.6869*, 2013.

H. Wang, G. Li, and G. Jiang. Robust regression shrinkage and consistent variable selection through the lad-lasso. *Journal of Business & Economic Statistics*, 25(3):347–355, 2007.

Z. Yang and F. S. Cohen. Image registration and object recognition using affine invariants and convex hulls. *IEEE Transactions on Image Processing*, 8(7):934–946, 1999.

M. Yuan and Y. Lin. Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 68(1):49–67, 2006.

J. Zhang, Z. Zhang, X. Xiao, Y. Yang, and M. Winslett. Functional mechanism: Regression analysis under differential privacy. *arXiv preprint arXiv:1208.0219*, 2012.

Q. Zhou, X. Zhang, J. Xu, and B. Liang. Large-scale bandit approaches for recommender systems. In *International Conference on Neural Information Processing*, pages 811–821. Springer, 2017.

H. Zou. The adaptive lasso and its oracle properties. *Journal of the American statistical association*, 101(476):1418–1429, 2006.

H. Zou and T. Hastie. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 67(2):301–320, 2005.