Efficient Learning in Graphs and in Combinatorial Multi-Armed Bandits

RUO-CHUN TZENG

Doctoral Thesis
Stockholm, Sweden, 2024

KTH Royal Institute of Technology
School of EECS
Division of Theoretical Computer Science

TRITA-EECS-AVL-2024:77
ISBN 978-91-8106-073-7

Academic Dissertation which, with due permission of the KTH Royal Institute of Technology, is submitted for public defence for the Degree of Doctor of Philosophy in D2, Lindstedtsvägen 9, Stockholm.

Tryck: Universitetsservice US AB

Abstract

Graphs have rich structures at both local and global scales. By exploiting structural properties in certain graph problems, it is possible to design computationally efficient algorithms or refine performance analysis. This thesis is divided into two parts: (i) designing new methods for discovering structure from graphs and (ii) studying the interplay between graphs and combinatorial multi-arm bandits. They differ in how structures are defined, discovered, and utilized.

In part (i), we start with a graph-mining problem on signed networks [TOG20], where the graph patterns we aim to detect are groups with mostly positive intra-group edges and mostly negative inter-group edges. We design our objective such that, given a solution, it reflects how well that solution matches the desired pattern. Our proposed algorithm makes no assumptions about the graph. It demonstrates competitive empirical performance in real-world graphs and synthetic graphs. The performance evaluation is conducted through a worst-case analysis, approximating the optimal solution. Moreover, we extend our conflicting-group detection [BGG+19, TOG20] as well as other graph-mining tasks (such as fair densest-subgraph detection [ABF+20], two-community detection [New06]) to a memory-limited and pass-limited setting. Under such a setting, Randomized SVD, which has been proposed by [HMT11], is the most preferable method in the memory-limited and pass-limited setting. However, for an input matrix of size $n \times n$, it has no guarantee in the $o(\ln n)$-pass regime, which is of most interest to practitioners. We hence derive a tighter analysis [TWA+22] for Randomized SVD for positive semi-definite matrices in any number of passes and for indefinite matrices under certain conditions. Furthermore, we initiate the study of a mixture of Johnson-Lindenstrauss distribution and the 0/1 Bernoulli distribution. We show that this mixture helps make the detection of 2-conflicting groups [BGG+19] more efficient.

In part (ii), we study combinatorial multi-arm bandits problems, where the graph properties play important roles but are somewhat hidden in the optimization problem. Instead, one derives the fundamental limit bounds on either the expected sample complexity or the expected cumulative regret, typically in the information-theoretical sense, and then explores the abstract properties associated with those bounds to solve the problem satisfactorily, either statistically or computationally. We focus on combiantorial semi-bandits where the learner observes individual feedbacks for each arm part of the selection. This formulation models many real-world problems, including online ranking [DKC21] in recommendation systems, network routing [CLK+14] in internet service providers, loan assignment [KWA+14], path planning [JMKK21], and influence marketing [Per22]. For best arm identification with fixed confidence, we propose the first polynomial-time algorithm whose sample complexity is instance-specifically optimal in high confidence regime and has polynomial dependency on the number of arms in moderate confidence regime. For regret minimization, we propose the first algorithm whose per-round time complexity is sublinear in the number of arms while matching the instance-specifically gap-dependent lower bound asymptotically.

Sammanfattning

Grafer har rika strukturer både på lokal och global skala. Genom att utnyttja strukturella egenskaper i vissa grafproblem är det möjligt att designa beräkningsmässigt effektiva algoritmer eller förfina prestandaanalysen. Denna avhandling är uppdelad i två delar: (i) att designa nya metoder för att upptäcka struktur från grafer och (ii) att studera samspelet mellan grafer och kombinatoriska multi-arm banditer. De skiljer sig åt i hur strukturer definieras, upptäcks och utnyttjas.

I del (i) börjar vi med ett grafgruvningsproblem på signerade nätverk [TOG20], där de grafmönster vi siktar på att upptäcka är grupper med mestadels positiva intra-gruppkanter och mestadels negativa inter-gruppkanter. Vi utformar vårt mål så att, givet en lösning, det återspeglar hur väl den lösningen matchar det önskade mönstret. Vår föreslagna algoritm gör inga antaganden om grafen. Den visar konkurrenskraftig empirisk prestanda i verkliga grafer och syntetiska grafer. Prestandautvärderingen genomförs genom en värsta fall-analys, som approximativt finner den optimala lösningen. Dessutom utvidgar vi vår konfliktgruppsdetektion [BGG+19, TOG20] samt andra grafgruvningsuppgifter (såsom rättvis tätaste subgrafdetektion [ABF+20], två-samhällesdetektion [New06]) till en minnesbegränsad och passbegränsad inställning. Under en sådan inställning är Randomized SVD, som föreslagits av [HMT11], den mest föredragna metoden i en minnesbegränsad och passbegränsad inställning. Men för en ingångsmatris av storlek $n \times n$ har den ingen garanti i $o(\ln n)$-passregimen, vilket är av största intresse för praktiker. Vi härleder därför en stramare analys [TWA+22] för Randomized SVD för positivt semi-definita matriser i vilket antal pass som helst och för indefinita matriser under vissa villkor. Vidare initierar vi studien av en blandning av Johnson-Lindenstrauss distribution och 0/1 Bernoulli-distributionen. Vi visar att denna blandning hjälper till att göra detektionen av 2-konfliktgrupper [BGG+19] mer effektiv.

I del (ii) studerar vi kombinatoriska multi-arm banditproblem, där grafernas egenskaper spelar viktiga roller men är något dolda i optimeringsproblemet. Istället härleder man de fundamentala gränserna för antingen den förväntade provtagningskomplexiteten eller den förväntade kumulativa ånger, vanligtvis i informations-teoretisk mening, och utforskar sedan de abstrakta egenskaperna associerade med dessa gränser för att lösa problemet på ett tillfredsställande sätt, antingen statistiskt eller beräkningsmässigt. Vi fokuserar på kombinatoriska semi-banditer där läraren observerar individuella återkopplingar för varje arm som är en del av urvalet. Denna formulering modellerar många verkliga problem, inklusive online ranking [DKC21] i rekommendationssystem, nätverksrouting [CLK+14] hos internetleverantörer, lånefördelning [KWA+14], vägplanering [JMKK21], och influencer-marknadsföring [Per22]. För bästa armidentifiering med fast förtroende föreslår vi den första polynomtidalgoritmen vars provtagningskomplexitet är instansspecifikt optimal i hög förtroenderegim och har polynomberoende på antalet armar i måttlig förtroenderegim. För ångerminimering föreslår vi den första algoritmen vars per-rundans tidskomplexitet är sublinjär i antalet armar samtidigt som den matchar den instansspecifikt gapberoende nedre gränsen asymptotiskt.

List of Papers

Papers included

Paper A [TOG20] **Discovering Conflicting Groups in Signed Graphs.** Ruo-Chun Tzeng, Bruno Ordozgoiti, and Aristides Gionis. Advances in Neural Information Processing Systems (2020).

Paper B [TWA+22] **Improved Analysis of Randomized SVD for Top-Eigenvector Approximation.** Ruo-Chun Tzeng, Po-An Wang, Florian Adriaens, Aristides Gionis, and Chi-Jen Lu. International Conference on Artificial Intelligence and Statistics (2022).

Paper C [TWPL23] **Closing the Computational-Statistical Gap in Best Arm Identification for Combinatorial Semi-Bandits.** Ruo-Chun Tzeng, Po-An Wang, Alexandre Proutière, and Chi-Jen Lu. Advances in Neural Information Processing Systems (2023).

Paper D [TOA24] **Matroid Semi-Bandits in Sublinear Time.** Ruo-Chun Tzeng, and Naoto Ohsaka, Kaito Ariu. International Conference on Machine Learning (2024).

Papers not included

Paper E [WTP21] **Fast Pure Exploration via Frank-Wolfe.** Po-An Wang, Ruo-Chun Tzeng, Alexandre Proutière Advances in Neural Information Processing Systems (2021).

Paper F [WTP23] **Best Arm Identification with Fixed Budget: A Large Deviation Perspective.** Po-An Wang, Ruo-Chun Tzeng, Alexandre Proutière. Advances in Neural Information Processing Systems (2023).

Acknowledgement

I have been very fortunate to have three academic mentors, Prof. Aristides Gionis (KTH), and Prof. Chi-Jen Lu (Academia Sinica), and Prof. Alexandre Proutière (KTH), as well as a trustworthy collaborator, Po-An Wang, who is also my husband. They are all incredibly supportive and are always there to help whenever I need it.

Firstly, I would like to express my gratitude towards my main supervisor, Prof. Gionis. In terms of research, he has given me complete trust and the freedom to explore any topic that interests me. In terms of life, he shows care and empathy for his students. For instance, when I was hospitalized due to mental illness, he visited me and patiently waited during my 2-month hospitalization until I recovered. Additionally, he maintains long-term relationships with both academics (for example, universities in the US and Europe) and industry professionals (such as those at the Center for Artificial Intelligence led by Francesco Bonchi). While working in his lab, I never have to worry about running out of funding. He is interested in solving real-world problems that are currently occurring, such as cognitive warfare, information silos, and misinformation. From my perspective, the elegance of a solution comes after the reasonable formulation of a real-world problem. This is a crucial distinction that separates data mining from machine learning or theoretical computer science. I would also like to express my sincere thanks to my major collaborators, Po-An Wang and Prof. Chi-Jen Lu. Their insights into algorithm design and theoretical analysis have deeply ignited my enthusiasm for the field of theoretical machine learning.

My collaboration with Prof. Chi-Jen Lu, Prof. Alexandre Proutière, and my husband Po-An Wang, is more concerned with how well we can fundamentally solve a problem. It is unavoidable to make certain assumptions, starting from the simplest setting and then extending to more realistic ones. Working on problems of this type requires not only the ability to learn new tools quickly but also a rare creativity in algorithm design and theoretical analysis to make a major breakthrough on a long-standing problem. From my experience of collaborating with them, I have clearly seen my own limitations and my immaturity in organizing large-scale proofs (i.e., more than 50 pages of proofs) as well as in writing formal statements. I was impressed by how Prof. Proutière handled the naughty reviewers during the paper rebuttal period, and by his ability to write a concise (though quite dense) proof sketch of our main results.

Kaito Ariu and Yassir Jedra are friends I met through my husband; in other words, they are two hops away in my ego-network. I am grateful to Kaito for inviting me to a research internship at the CyberAgent AI Lab. There, I had a wonderful time discussing various topics with my mentor, Naoto Ohsaka, and meeting several new friends, including Mitsuka Kiyohara, Qiqi Gao, and Hao Wang. I appreciate Yassir for introducing me to the book [Ver18] written by Roman Vershynin, which has been extremely useful in some of my projects. I am also grateful to Kaito for leaving behind many useful books, such as [AB09], and for recommending Mikael Skoglund's excellent lectures on Information Theory and Probability and Random Processes to us.

All of my colleagues in Prof. Gionis's lab are friendly and supportive. I was fortunate to work with Bruno Ordozgoiti, who is enthusiastic about spectral methods, on my onboarding project. There, I applied the special structure of the eigenspace of the graph Laplacian of a complete graph, mentioned in one of the lectures [Dan19] by Daniel A. Spielman. I am grateful to Cigdem Aslay for swapping presentation slots with me when I was busy. I thank Guangyi Zhang for making the effort to present submodular maximization techniques in our group meeting, and I appreciate Stefan Neumann for the offline discussions about dynamic graph techniques. I am grateful to Sijing Tu, Hang Wang, Honglian Wang, and Tianyi Zhou for sharing tips on maintaining personal relationships. I appreciate Florian Adriaens, Ece Calikus, Corinna Coupette, Martino Ciaperoni, Federico Cinus, Erica Coppolillo, Mohit Daga, Yifei Jin, Lutz Oettershagen, Thibault Marette, Antonis Matakos, Ilie Sarpe, Suhas Thejaswi, and Han Xiao for their group presentations, as well as the data science seminar jointly held by Prof. Henrik Boström, Prof. Aristides Gionins, and Prof. Šarūnas Girdzijauskas. The presentations by participants, such as Prof. Karl Meinke, Avavind Nair, and Susanna Pozzoli, broadened my knowledge on different topics. I would like to thank Martino Ciaperoni and Antonis Matakos for giving me a tour around the Aalto University campus. I would also like to express my deepest gratitude towards the HR representative, Anna Olanås Jansson, who has helped me with a lot of miscellaneous things, including arranging hospitalization for urgent mental health care.

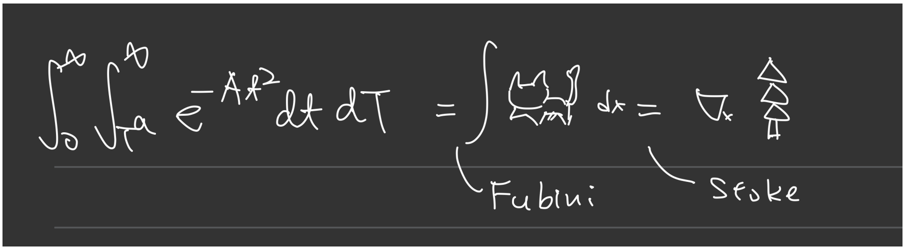
Before starting my PhD at KTH, I had a two-month visit to Prof. Chi-Jen Lu's lab at the Institute of Information Science in Academia Sinica (Taiwan), where I met my lovely husband, Po-An Wang who resembles a mixture of the Olympic gold medallist Tom Daley and the one of the highest-paid actors Tom Cruise, my husband's good friend Dr. Sung-Hsien Hsieh and Dr. Vacuity Hu, Yan-Lin Chen, Jin-Hua Lin and Yi-Shan Wu, Yi-Te Hong, and Ting-Yun Chang, Hong-You Chen, and Hannah Kuo. I also got to know Prof. Po-An Chen, Prof. Yen-Huan Li, Prof. Chuang-Chieh Lin, and Prof. Chen-Yu Wei. However, during that time, I was struggling with self-doubt as my master's thesis had been rejected by all of the top machine learning conferences. I lost faith in myself and decided to go into the industry to work as a software engineer. Luckily, during an interview with Prof. Yarin Gal, he informed me that Microsoft was planning to build an AI Center in Taiwan. I was recruited to a team working on search engine advertisement. There, I met many wonderful colleagues, including Guan-Wen

Chen, Tsung-Hsien Chiang, Edward Chou, Yen-Chn Fu, Ping Huang, Po-Chih Huang, Hung-Ju King, Mike Lee, Guancheng Li, Irene Lin, Hsiang-Hsiang Liu, Juan Liu, Yu-Wen Liu, Kuan-Hao Liao, Yu-Wen Liu, Chi Qi, Po-Yuan Teng, Cheng-Hsiao Tsou, Yu-Lin Tsou, Hsiao-Ling Wang, Xin-Yu Wang, Ashley Wu, Peng Xu, and Mirror Xu. I would like to express my special gratitude to my direct co-workers: Yu-Lin Tsou, who is one of the strongest members on the team but always remains humble and cautious in every project he is involved in, and Edward Chou, who advised me to choose good supervisors over competitive but unsupportive ones. I would also like to thank my direct managers, Juan Liu and Mirror Xu, as well as Mike Lee, who assigned me challenging large-scale machine learning tasks. For example, I was tasked with designing algorithms to solve the cold-start problem for new products and dealing with dynamically changing keywords due to seasonal patterns in textual ads. These challenging tasks reignited my passion to pursue a research career. Coincidentally, around that time, my master's thesis was accepted to ICML 2019. This boosted my confidence and provided the necessary evidence to show others that I am qualified to pursue a research career. I enjoyed attending ICML 2019, both the main conferences and the offline meetups with other researchers interested in Graph Neural Networks, as well as with Prof. Chen Liu.

During my time at National Tsing Hua University (NTHU), I attended the competitive programming club and got to know many highly-motivated individuals including: Yen-Ling Chang, Yi-Ling Cheng, Negolas Cheng, Hanting Chiang, Ming Chia Chung, Cindy Hou, Yen-Chen Lin, Henry Yang, I-Tse Yang, Shih-Yang Su, and Wei-Che Wei. However, I was not particularly interested in winning the competitions. Instead, I befriended a wonderful person, An-An Yu, who introduced me to the burgeoning field of machine learning and told me about the opportunity to do an internship at Academia Sinica, our national research institute. Inspired by her path, I chose Prof. Shan-Hung Wu's lab, which at that time, was one of the few labs focusing on machine learning research. There, I met several extremely competitive Senpai's: Ting-Yu Cheng and Shao-Kan Pi, who were not only excellent in their respective research fields (with independent publications in NeurIPS and in SIGMOD) but also had started their own business (an app called "Forest") during their 2-year master's study! Being in this lab, I learned about various topics in machine learning from our regular paper presentations by Meng-Ren Chen, Wei-Dan Chen, Chi-Chun Chuang, Tsai-Yu Feng, Tsu-Jui Fu, Xin-Yang Gong, Cheng-Yu Hsu, Katy Lee, Chia-Hung Lin, Gui-Guan Lin, Kan-Jun Liu, Tz-Yu Lin, Yen-Chen Lin, Yu-Shan Lin, Hsuan-Yu Guo, Zhufeng Pan, Ching Tsai, Chen-Wei Tseng, Chia-Hsin Yeh, An-An Yu, Ted Yu, Chia-Hung Yuan, Cheng-Hsin Weng, You-Jhih Wong, Ching-Chan Wu, and Yi-Shuan Wu. I completed my master's degree in one-and-a-half years with a thesis oral defense committee consisting of Prof. Hwann-Tzong Chen, Prof. Yuh-Jye Lee, Prof. Hong-Han Shuai, and Prof. Min Sun. I am thankful for Prof. Sun's valuable feedback, which helped to generalize my thesis to broader application scenarios, and for Prof. Chen's feedback, which highlighted the possibility of extending my work to hypergraphs.

Last but not least, regarding doctoral thesis writing, I am drawing inspiration

from the doctoral theses of many people, including Kaito Ariu, Aditya Bhaskara, Jan van den Brand, Edo Liberty, Lijie Chen, Corinna Coupette, Yarin Gal, Yassir Jedra, Yuko Kuroki, M. Sadegh Talebi, Daniel A. Spielman, and Johan Ugander, to organize the thesis structure. This thesis would not have been possible without the emotional support from my friends Chia-Man Hung, Chi-Yun Hsu, Sijing Tu, Honglian Wang, Kaito Ariu, Xin-Yu Wang, Yi-Shan Wu, Holly Wang, Hsin-Wei Gao, Hsuan-Yu Guo, Micki Liao, Linda Tenhu, and Guo-Jhen Wu. Finally, the deepest thanks go to my wonderful husband, Po-An Wang, who has the ambition, courage, and rare creativity to tackle long-standing open problems. He is extremely responsible in every paper he is involved in. Moreover, he not only conducts great research but also manages to maintain a good work-life balance. Working with him is also fun. For example, this is how he playfully responded when I was asking for an integral to bound the error probability in one of my papers [TWPL23]:



Ruo-Chun Tzeng
September, 2024

Contents

Contents

Chapter 1

Introduction

This thesis is centered around two themes: graph mining and combinatorial multi-armed bandits. For both themes, we are interested in designing computationally efficient algorithms with provable guarantees. For graph mining problems, we design an objective function that reflects how well the solution matches the pattern we aim to detect, and design computationally efficient algorithms to find such patterns with provable guarantees. For combinatorial multi-armed bandits, we study the fundamental limit satisfied by any reasonably good algorithm, and design algorithms exploring different trade-offs between statistical efficiency and computational efficiency.

1.1 Graph mining

Graph mining provides a set of tools and techniques that are useful to (i) extract patterns from graph-structured data, such as identifying frequent subgraphs, detecting communities, and predicting links, and (ii) analyze the properties of real-world graphs, such as their degree distribution, clustering coefficient, diameter, and other structural characteristics. These tools and techniques enable us to gain insights into the underlying structures and dynamics of complex systems and enhance our understanding of their behavior. In this thesis, we will be discussing the following graph mining tasks:

- The problem of *densest subgraph detection* is that: given an undirected graph $G = (V, E)$, the goal is to find the node subset $S \subseteq V$ whose edge density $\frac{|E(S)|}{|S|}$ is maximized, where $E(S)$ denotes the set of edges induced by S. Densest subgraph detection can be used to detect (a) tightly connected communities in social networks, (b) fraudulent reviews in a user-product bipartite graph, (c) proteins that are regulating the same process in a protein-protein interaction network, (d) money laundering in a multipartite directed transactions network, etc. See a recent survey [LMFB23] for more applications on densest subgraph detection.

- *Graph clustering* aims to identify groups of nodes that are more similar or more connected to each other than the rest of the network. These densely connected groups, or clusters, often represent communities in social networks, modules in biological networks, or topical clusters in web graphs. Depending on the quality measure of the clustering, graph clustering can be further divided into (a) modularity-based, (b) cut-based, (c) density-based, and (d) other approaches. In this thesis, we will discuss Newsman's modularity-based algorithm [New06] for detecting 2 communities. It works by finding the top-eigenvector $\boldsymbol{u}$ of the modularity matrix $\mathbf{B}$ where $B_{i,j} = A_{i,j} - \frac{\deg(i)\deg(j)}{2|E|}$, and then rounding $\boldsymbol{u}$ to a vector in $\{\pm 1\}^{|V|}$ with 1 (resp. -1) representing the node is in the first (resp. second) community.

Signed graph mining

Signed graph mining deals with the analysis of signed graphs which are graphs where each edge is labeled as either positive or negative. A positive edge represents relationships such as trust, like, or friendship, and a negative edge represents relationships such as distrust, dislike, or enmity. Common signed graph mining tasks include: signed link prediction [WSLX17], signed community detection [CDGT19], node ranking [SJ14], sentiment analysis [CLTL17], and balance theory analysis [LHK10]. These tasks and related methodologies provide valuable insights into the structure and dynamics of signed networks and can be applied in various domains such as social network analysis, e-commerce, bioinformatics, and cybersecurity. Signed graph mining can help understand and predict human behavior, detect communities and conflicts, and analyze sentiment and status.

In this thesis, we study a specific signed graph mining task called *conflicting group detection* in Paper A. In the conflicting group detection, we are given a signed graph, and the goal is to identify k disjoint node subsets $S_1, \cdots, S_k$ such that edges between any two groups S_i and S_j for any $i \neq j$ are mostly negative and edges within each group S_i are mostly positive. A conflicting group can represent a set of individuals with viewpoints conflicting with other groups of users in social networks or political networks. Identifying conflicting groups can provide valuable insights into the structure and dynamics of social, political, or organizational networks. It can help predict future conflicts, understand the root causes of existing conflicts, and design interventions to resolve or mitigate conflicts.

Randomized linear algebra

Randomized Linear Algebra is a field that utilizes randomization techniques to develop efficient algorithms for various linear algebra problems, especially those involving large-scale matrices. The primary goal of randomized linear algebra is to develop algorithms that are faster, more scalable, and require less memory than traditional deterministic methods. Randomized linear algebra is particularly important in the context of big data, where the size of the data often makes traditional methods infeasible. Key techniques include random sampling, random projection, randomized matrix decompositions, sketching, and randomized iterative methods. Applications of randomized linear algebra span multiple domains including machine learning, data mining, statistics, signal processing, and scientific computing. For instance, in machine learning, randomized methods are used for tasks like principal component analysis (PCA), k-means clustering, regression, and community detection.

In this thesis, in Paper B, we will be discussing random projection and randomized matrix decomposition. Randomized SVD [HMT11] provides a memory-efficient way to perform matrix factorization. Given a symmetric matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, Randomized SVD works by (a) projecting $\mathbf{A}^q$ onto a random matrix $\mathbf{S} \sim \mathcal{N}(0,1)^{n \times d}$ to $\mathbf{Y} = \mathbf{A}^q\mathbf{S}$ for some $q \in \mathbb{N}$, (b) performing QR decomposition on $\mathbf{Y} = \mathbf{QR}$, and then (c) computing the approximation of the i-th largest eigenvectors of $\mathbf{A}$ by using $\mathbf{Q}\boldsymbol{u}_i(\mathbf{Q}^T\mathbf{A}\mathbf{Q})$, where $\boldsymbol{u}_i(\cdot)$ denotes the i-th eigenvector of the input matrix. The theoretical guarantee of Randomized SVD using a random matrix $\mathbf{S}$ of size $n \times d$ is:

$$\mathbb{E}\left[\left\|\mathbf{A} - \mathbf{Q}\mathbf{Q}^T\mathbf{A}\right\|_2\right] \le c\sqrt{\lambda_{k+1}(\mathbf{A})}, \tag{1.1}$$

where $\lambda_i(\cdot)$ is the i-th eigenvalue of the input matrix, and $c = (1 + \sqrt{\frac{k}{d-k-1}} + \frac{e\sqrt{d}}{d-k}\sqrt{n-k})^{\frac{1}{2q+1}}$. In Paper B, we focus on the top eigenvector, and aim to analyze the approximated top eigenvector $\hat{\mathbf{u}} \in \mathbb{S}^{n-1}$ returned by Randomized SVD with respect to the multiplicative ratio objective:

$$R(\hat{\mathbf{u}}) = \frac{\hat{\mathbf{u}}^T\mathbf{A}\hat{\mathbf{u}}}{\lambda_1(\mathbf{A})},$$

where the multiplicative ratio $R(\hat{\mathbf{u}})$ is the ratio between the Rayleigh quotient $\hat{\mathbf{u}}^T\mathbf{A}\hat{\mathbf{u}}$ and the top eigenvalue $\lambda_1(\mathbf{A})$. Notice that for a positive semi-definite matrix $\mathbf{A}$, it is possible to convert (1.1) into a guarantee of $R(\hat{\mathbf{u}})$ by (a) using Davis-Kahan theorem [YWS15], yielding $\langle \hat{\mathbf{u}}, \boldsymbol{u}_1(\mathbf{A}) \rangle\rangle \ge \sqrt{1 - \frac{c^2\lambda_2(\mathbf{A})}{(\lambda_1(\mathbf{A}) - \lambda_2(\mathbf{A}))^2}}$, and then (b) using $\hat{\mathbf{u}}^T\mathbf{A}\hat{\mathbf{u}} \ge \lambda_1(\mathbf{A}) \langle \hat{\mathbf{u}}, \boldsymbol{u}_1(\mathbf{A}) \rangle\rangle^2$ to derive a lower bound of $R(\hat{\mathbf{u}})$. However, the converted guarantee of $R(\hat{\mathbf{u}})$ unfortunately depends on the *eigengap* $\lambda_1(\mathbf{A}) - \lambda_2(\mathbf{A})$ of the first and the second largest eigenvalue of $\mathbf{A}$. In Paper B, we develop a tight analysis of Randomized SVD with respect to $R(\hat{\mathbf{u}})$ that does not depend on the eigengap.

1.2 Combinatorial multi-armed bandits

Combinatorial multi-armed bandit is an extension of the classic multi-armed bandit problem. The multi-armed bandit problem models an agent that simultaneously attempts to acquire new knowledge (exploration) and optimize decisions based on existing knowledge (exploitation) while interacting with an environment. The difference between the combinatorial multi-armed bandit and the multi-armed bandit problem is that in each round, the learner selects a subset of arms that satisfy certain combinatorial constraints, instead of choosing a single arm as in the multi-armed bandit. Combinatorial multi-armed bandits can be used to model many real-world tasks such as online ranking [DKC21] in recommendation systems, network routing [CLK$^+$14] in internet service providers, loan assignment [KWA$^+$14], path planning [JMKK21], and influence marketing [Per22].

An instance of combinatorial multi-armed bandit is parameterized by $([K], \mathcal{X}, \boldsymbol{\mu})$, where $[K]$ is the set of arms, $\mathcal{X} \subseteq \{0,1\}^K$ is the set of actions satisfying certain combinatorial constraints, and $\boldsymbol{\mu} \in \mathbb{R}^K$ is a mean vector where each component μ_k is the expectation of arm k's reward distribution ν_k. At each round t, the learner selects an action $\boldsymbol{x}(t) \subseteq [K]$ from the action set $\mathcal{X}$. There are two types of feedback models. In the *semi-bandit* feedback model, the learner observes $y_k(t) \sim \nu_k$ for each of the arm $k \in \mathsf{supp}\,(\boldsymbol{x}(t))$. In the *full-bandit* feedback model, the learner observes $r(\boldsymbol{y}(t), \boldsymbol{x}(t))$ only, where $y_k(t) \sim \nu_k$ for $k \in [K]$ and r is some reward function. The objectives are related to the best action $\boldsymbol{i}^\star(\boldsymbol{\mu}) \in \operatorname{argmax}_{\boldsymbol{x}\in\mathcal{X}} r(\boldsymbol{\mu}, \boldsymbol{x})$ that maximizes the expected reward. There are two popular objectives:

- *Regret minimization* aims to minimize the expected cumulative regret (or equivalently maximize the expected cumulative reward), as compared to an algorithm that knows $\boldsymbol{\mu}$ and always chooses the best action $\boldsymbol{i}^\star(\boldsymbol{\mu})$ that maximizes the expected reward. For this setting, the learner has to balance the trade-off between exploration and exploitation.
- *Pure exploration* aims to identify the correct answer (a) with fixed confidence or (b) with a fixed budget. The correct answer can be the best action $\boldsymbol{i}^\star(\boldsymbol{\mu})$ that maximizes the expected rewards, or any ϵ-good action [BDS22]. For (a), the goal is to identify the correct answer with minimal sample complexity. For (b), the goal is to minimize the error probability after a fixed number of samples.

Regret minimization

In regret minimization, instantaneous regret is defined as the difference between the expected reward of the best action and that of the action chosen. The goal of regret minimization is to minimize the total regret cumulative over all rounds. Regret minimization is a very general framework that can be applied to many different types of decision problems, including online learning [RST10], reinforcement learning [AJO08], and game theory [GGM08].

In this thesis, we focus on combinatorial semi-bandits. At each round, the learner selects a subset of arms that satisfy certain combinatorial constraints, observes individual feedback for each of the arms part of the selection, and incurs an instantaneous regret. The goal is to design algorithms such that the total regret is sublinear in the number of rounds. For matroid structures, Combinatorial Upper Confidence Bound (`CUCB`) [GKJ12, CWY13, KWA$^+$14, KWAS15] is a competitive method whose regret bound matches the gap-dependent lower bound proposed by Kveton et al. [KWA$^+$14], and KL-based Efficient Sampling for Matroids (`KL-OSM`) [TP16] is an optimal algorithm whose regret bound matches the instance-specific lower bound. However, both algorithms require the complexity of the time per round to be at least linear to K, where K is the number of arms. In Paper D, we develop the first matroid semi-bandit algorithm whose per-round time complexity is sublinear in K while matching the gap-dependent lower bound asymptotically. Our method is particularly suitable for large values of K.

Pure exploration

The goal of pure exploration is to identify the set of correct answers (a) with fixed confidence using as few samples as possible, or (b) with a fixed budget achieving as low error probability as possible. Pure exploration is particularly useful in situations where the cost of exploration is low compared to the long-term benefits of finding the correct answers. For example, it might be used in clinical trials to find the most effective treatment, in A/B testing to find the best website design, or in machine learning hyperparameter tuning to find the best set of parameters.

In this thesis, we focus on pure exploration (a) with fixed confidence and semi-bandit feedback under combinatorial structures. In each round, the learner selects a subset of arms that satisfy certain combinatorial constraints and observes individual feedback for each arm that is part of the selection. In the problem of best arm identification, the sample complexity lower bound has been derived by the change-of-measure technique [KCG16, GK16], and the statistically optimal algorithms such as `FWS` [WTP21] and `CombGame` [JMKK21] have been proposed. However, these algorithms are not computationally efficient. In Paper C, we design the first polynomial-time algorithm whose sample complexity is instance-specifically optimal in the high-confidence regime and has a polynomial dependency on the number of arms in the moderate confidence regime.

1.3 Overview of thesis

This thesis contributes to various aspects of graph mining and combinatorial multi-armed bandits. Paper A focuses on one particular signed graph mining task. Paper B extends the first Paper A to the memory-limited setting where techniques from randomized numerical linear algebra are particularly useful. Paper C considers the pure exploration problem for combinatorial semi-bandits. Paper D focuses on the regret minimization for combinatorial semi-bandits. Both Paper C and Paper D contribute to designing faster algorithms for combinatorial semi-bandits. The contributions of each paper are as follows.

In Paper A, a signed graph mining task called k-conflicting group detection is considered. The goal is to identify k disjoint node subsets $S_1, \cdots, S_k$ such that edges between any S_i, S_j for $i \neq j$ are mostly negative and edges within S_i for any $i \in [k]$ are mostly positive. This problem is APX-hard [BCMV12, BGG$^+$19], and we develop a computationally efficient algorithm based on (a) the top-eigenvector of the (modified) signed adjacency matrix, and (b) the rounding technique to assign the group membership of the nodes. Our developed technique is empirically competitive in both real-world signed networks and synthetic stochastic block models.

In Paper B, we analyzed and improved the theoretical guarantee of Randomized SVD (`RSVD`) [HMT11] for graph mining tasks. These tasks include principal component analysis [Jol86], fair dense subgraph detection [ABF$^+$20], 2-community detection [New06], and k-conflicting group detection (Paper A). For the top eigenvector $\hat{\mathbf{u}}$ of a symmetric positive semi-definite matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ approximated by `RSVD`, prior analysis required $\Omega(\ln n)$ passes to ensure a non-trivial result for the multiplicative ratio $R(\hat{\mathbf{u}}) = \hat{\mathbf{u}}^T \mathbf{A} \hat{\mathbf{u}} / \lambda_1(\mathbf{A})$, where $\lambda_1(\cdot)$ denotes the largest eigenvalue of the given matrix. We show that $R(\hat{\mathbf{u}}) = (\Omega(\frac{d}{n}))^{\frac{1}{2q+1}}$ with a probability of at least $1 - e^{-\Omega(d)}$ for any number q of passes and $\mathcal{O}(nd)$ space. This result provides a theoretical guarantee for practical settings with $o(\ln n)$ or even a constant number of passes. In addition, we consider extensions to indefinite matrices and modifications of `RSVD` where the standard Gaussian distribution is replaced with a mixture of standard Gaussian and 0/1 Bernoulli. We show that `RSVD` with such a mixture distribution as a random projection matrix expedites the identification of k-conflicting groups (Paper A) when compared to the standard `RSVD`.

In Paper C, we study the problem of best arm identification with fixed confidence for combinatorial semi-bandits. The sample complexity lower bound satisfied by any reasonably good algorithm is derived from previous work [JMKK21], and there exist statistically optimal algorithms such as `FWS` [WTP21] and `CombGame` [JMKK21]. However, none of these optimal algorithms is computationally efficient. We inspect the optimization problem of the sample complexity lower bound and obtain a useful structural property that allows us to design a computationally efficient oracle (called `MCP`) based on a two-player no-regret learning algorithm. This `MCP` oracle is necessary for the design of optimal best arm identification algorithms. For example, `MCP` is used in the Chernoff stopping rule [GK16] and also in the sampling algorithm of [WTP21, JMKK21]. Based on the `MCP` oracle, we design the first statistically optimal and computationally efficient algorithm for this problem.

In Paper D, matroid semi-bandit is studied. `CUCB` [CWY13] and `KL-OSM` [TP16] are two competitive algorithms for this setting. The regret upper bound of the former achieves the gap-dependent lower bound while that of the latter achieves the instance-specific lower bound. Both algorithms require a time complexity of at least $\Omega(K)$ and require a $O(K(\log K + \mathcal{T}_{\text{member}}))$-time greedy algorithm to compute their sampling strategy, where K is the number of arms and $\mathcal{T}_{\text{member}}$ is the time to query the membership oracle of the given matroid. We develop a sublinear-time algorithm for this problem based on a dynamic algorithm that maintains a maximum-weight base over inner product weights. As all arms change at each round, the existing dynamic algorithms are not directly applicable. Our insight for addressing this issue is that an UCB index can be decomposed into an inner product of (i) a feature, which depends on the arm k and is a pair of empirical reward estimates and radius of the confidence interval (ii) a query, which depends only on the round t. Our proposed dynamic algorithm consists of two speedup techniques. One is feature rounding, which rounds each feature into a few bins to reduce the number of distinct features to consider. The other is the minimum hitting set, which allows us to compute a small number of queries in advance and correctly identify a maximum weight base for any query.

The thesis is organized as follows:

- In Chapter 2, the basic concepts required in Paper A, Paper B, Paper C, and Paper D are introduced, and the important related works in these papers are discussed
- In Chapter 3, an overview of the results of Paper A and Paper B is presented
- In Chapter 4, an overview of the results of Paper C and Paper D is presented
- In Chapter 5, the challenges and open problems left in these papers are discussed

Chapter 2

Background

This chapter serves to provide basic concepts and tools for understanding Paper A, Paper B, Paper C, and Paper D. Section 2.1 provides the definition of approximation algorithms as required in Paper A. Section 2.2 provides an overview of the technique of random projection, which is the main topic of Paper B. Section 2.3, Section 2.4, Section 2.6, Section 2.7, and Section 2.8 are tools used by Paper C. Specifically:

- Section 2.3 provides an introduction to the Lagrangian multiplier methods
- Section 2.4 provides an introduction to the minimax theorem
- Section 2.6 provides an introduction to the Frank-Wolfe algorithm
- Section 2.7 provides an introduction to the envelope theorem
- Section 2.8 provides an introduction to the terminologies of best arm identification

Section 2.5 provides an introduction to matroids as required by Paper D. Finally, the chapter concludes with Section 2.9 and Section 2.10, where state-of-the-art methods are discussed in Section 2.9, and the most important related works are reviewed in Section 2.10.

2.1 Approximation algorithms

Approximation algorithms are polynomial-time algorithms that find solutions to optimization problems with provable guarantees with respect to the optimal solution. There are many established techniques for designing an approximation algorithm, such as greedy strategies, local search, dynamic programming, linear- and convex-programming relaxations, primal-dual methods, dual fitting, metric embedding, and random sampling. For a maximization problem with objective function f, we say that an algorithm A is an *α-approximation algorithm* if A runs in polynomial time and returns a solution x that satisfies

$$\alpha OPT \leq f(x) \leq OPT \text{ and } \alpha < 1,$$

where OPT is an optimal value of the problem. For minimization problems, we say that an algorithm A is a α-*approximation* algorithm if A runs in polynomial time and returns a solution x that satisfies

$$OPT \leq f(x) \leq \alpha OPT \text{ and } \alpha > 1.$$

We say that a problem is *APX-hard* if there does not exist an approximation algorithm with a constant approximation ratio unless $P = NP$.

2.2 Random projection

Random projection refers to the technique of projecting a set of points from a high-dimensional space to a randomly chosen low-dimensional subspace [Vem05]. Let $\boldsymbol{u} = (u_1, \cdots, u_n)^T$ and $\mathbf{R} = [\boldsymbol{r}_1, \cdots, \boldsymbol{r}_d]$ be a $n \times d$ uniform random orthonormal matrix. Then, the projected vector is defined as

$$\boldsymbol{v} = \sqrt{\frac{n}{d}} \mathbf{R}^T \boldsymbol{u} \quad \text{and} \quad \mathbb{E}\left[\|\boldsymbol{v}\|^2\right] = \|\boldsymbol{u}\|^2 .$$

To understand the normalization factor n/d, here we present the intuition due to Sanjoy Dasgupta and Anupam Gupta [DG03]: since each of the n dimensions is equally likely to be chosen, i.e., $\mathbb{E}\left[\langle \boldsymbol{v}, \boldsymbol{z}_i \rangle^2\right] = 1/n$, we have

$$\mathbb{E}\left[\cos^2 \theta(\boldsymbol{u}, \mathbf{R})\right] = \sum_{i=1}^{d} \mathbb{E}\left[\langle \boldsymbol{u}, \boldsymbol{r}_i \rangle^2\right] = \frac{d}{n}.$$

One popular way [HP14] to approximate $\mathbf{R}$ is to sample the entries of $\mathbf{R}$ by i.i.d. from the standard Gaussian $\mathcal{N}(0,1)$ distribution, and the following property holds: Let $\boldsymbol{u} \in \mathbb{R}^n$. Then,

$$\mathbb{P}\left[\cos^2 \theta(\boldsymbol{u}, \mathbf{R}) = \Theta\left(\frac{d}{n}\right)\right] \geq 1 - e^{-\Omega(d)}, \tag{2.1}$$

where $\theta(\boldsymbol{u}, \mathbf{R}) = \cos^{-1}\left(\max_{\boldsymbol{x} \in \text{range}(\mathbf{R})} \frac{\langle \boldsymbol{u}, \boldsymbol{x} \rangle}{\|\boldsymbol{u}\|_2 \|\boldsymbol{x}\|_2}\right)$ is the projection angle. An application of Equation (2.1) is the famous Johnson-Lindenstrauss Lemma (see Theorem 5.3.1 in [Ver18]): Let $\mathcal{X}$ be a set of N points in $\mathbb{R}^n$ and $\epsilon > 0$. If $d = \Omega(\frac{1}{\epsilon^2} \log N)$, then with probability at least $1 - 2\exp(-\Omega(\epsilon^2 d))$,

$$(1 - \epsilon) \|\boldsymbol{x} - \boldsymbol{y}\|_2 \leq \sqrt{\frac{n}{d}} \|\mathbf{R}(\boldsymbol{x} - \boldsymbol{y})\|_2 \leq (1 + \epsilon) \|\boldsymbol{x} - \boldsymbol{y}\|_2$$

holds for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$.

2.3 Lagrangian multiplier method

Here we present the material from Chapter 5 in the book by Vishnoi [Vis21]. Consider a problem:

$$\inf_{\boldsymbol{x} \in \mathbb{R}^n} f(\boldsymbol{x}) \text{ subject to } f_i(\boldsymbol{x}) \leq 0 \text{ for } i = 1, \cdots, m \text{ and } h_i(\boldsymbol{x}) = 0 \text{ for } i = 1, \cdots, p.$$

The idea of the Lagrangian multiplier method is to "move the constraints to the objective". Define the *Lagrangian function*

$$L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\boldsymbol{x}) + \sum_{j=1}^{m} \lambda_i f_i(\boldsymbol{x}) + \sum_{i=1}^{p} \mu_i h_i(\boldsymbol{x}).$$

The relationship between the optimum value $y^* = \inf_{\boldsymbol{x}\in\mathbb{R}^n} \sup_{\boldsymbol{\lambda}\geq 0, \boldsymbol{\mu}} L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})$, and the *Lagrangian dual function*

$$g(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \inf_{\boldsymbol{x}\in\mathbb{R}^n} L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})$$

is that $g(\boldsymbol{\lambda}, \boldsymbol{\mu}) \leq y^*$ holds for any $\boldsymbol{\lambda} \geq 0$ and $\boldsymbol{\mu}$. In other words, the Lagrangian dual function provides a lower bound for the optimum. Thus, we have *weak duality* $\sup_{\boldsymbol{\lambda}\geq 0, \boldsymbol{\mu}} g(\boldsymbol{\lambda}, \boldsymbol{\mu}) \leq y^*$, which is also an application of the max-min inequality in Section 2.4. For the *strong duality*, it is equivalent to the Karush–Kuhn–Tucker (KKT) conditions. Suppose $\boldsymbol{x}^\star \in \mathbb{R}^n$, $\boldsymbol{\lambda}^\star \in \mathbb{R}^m$ and $\boldsymbol{\mu}^\star \in \mathbb{R}^p$ satisfy KKT optimality conditions:

(i) Primal feasibility: $f_j(\boldsymbol{x}^\star) \geq 0$ for $j = 1, \cdots, m$ and $h_i(\boldsymbol{x}^\star) = 0$ for $i = 1, \cdots, p$

(ii) Dual feasibility: $\boldsymbol{\lambda}^\star \geq 0$

(iii) Stationary $\partial_{\boldsymbol{x}} L(\boldsymbol{x}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star) = 0$

(iv) Complementary slackness: $\lambda_j^\star f_j(\boldsymbol{x}^\star) = 0$ for all $j = 1, \cdots, m$

Then, we have the strong duality: $f(\boldsymbol{x}^\star) = g(\boldsymbol{\lambda}^\star, \boldsymbol{\mu}^\star)$.

2.4 Max-min inequality and minimax theorem

For any function $f : X \times Y \to \mathbb{R}$, we have the *max-min inequality*

$$\sup_{\boldsymbol{x}\in X} \inf_{\boldsymbol{y}\in Y} f(\boldsymbol{x}, \boldsymbol{y}) \leq \inf_{\boldsymbol{y}\in Y} \sup_{\boldsymbol{x}\in X} f(\boldsymbol{x}, \boldsymbol{y}).$$

When the equality holds, one says that f, X, and Y satisfies the *strong max-min inequality* (or the saddle-point property). For special case when X and Y are both *compact convex sets*, then one has the *minimax theorem*:

$$\max_{\boldsymbol{x}\in X} \min_{\boldsymbol{y}\in Y} f(\boldsymbol{x}, \boldsymbol{y}) = \min_{\boldsymbol{y}\in Y} \max_{\boldsymbol{x}\in X} f(\boldsymbol{x}, \boldsymbol{y}).$$

2.5 Matroids

Here we present a short introduction taken from [KWA$^+$14]. A *matroid* is a pair $\mathcal{M} = (E, \mathcal{I})$, where $E = \{1, \cdots, K\}$ is called the *ground set*, and $\mathcal{I}$ called the *independent sets* which is the set of the subsets of E. For $\mathcal{I}$, in addition to $\emptyset \in \mathcal{I}$, it has to satisfies two properties:

- (downward closure) For any $Y \in \mathcal{I}$, if $X \subset Y$, then $X \in \mathcal{I}$
- (augmentation property) For any $X, Y \in \mathcal{I}$, if $|X| < |Y|$, then there eixsts $e \in Y$ such that $X \cup \{e\} \in \mathcal{I}$

A set X is called a *basis* of the matroid $\mathcal{M}$ if $|X|$ is maximum. All bases have the same cardinality which is known as the *rank* of the matroid. A *weighted matroid* is a matroid associated with a vector $\boldsymbol{w} \in \mathbb{R}_+^K$. The problem of finding a *maximum-weight basis* of a matroid is

$$\operatorname*{argmax}_{X \in \mathcal{I}} \sum_{e \in X} w_e. \tag{2.2}$$

It is well-known that Equation (2.2) can be solved by the following greedy algorithm:

- $X^\star \leftarrow \emptyset$. Sort $\boldsymbol{w}$ in the non-increasing order such that $w_{i_1} \geq \cdots \geq w_{i_K}$.
- Repeatedly add i_j to $X^\star$ if $X^\star \cup \{i_j\} \in \mathcal{I}$ and then $j = j + 1$ until $|X^\star|$ equals the rank of $\mathcal{M}$.

2.6 Frank-Wolfe algorithm

We present materials from [Jag13] for the Frank-Wolfe algorithm. Suppose we would like to solve the constrained convex optimization problem of the form:

$$\min_{\boldsymbol{x} \in \mathcal{D}} f(\boldsymbol{x}), \tag{2.3}$$

where we assume that the function f is convex and continuously differentiable and the domain $\mathcal{D}$ is a compact convex set of any vector space. For the optimization problem (2.3), one of the earliest and simplest iterative optimizers is given by the Frank-Wolfe method [FW56] (also known as *conditional gradient method*).

The idea of the Frank-Wolfe method (`FW`) is as follows: At the current position $\boldsymbol{x}^{(k)}$, the algorithm considers the linearization of the objective function

$$f(\boldsymbol{x}^{(k)}) + \left\langle \boldsymbol{y} - \boldsymbol{x}^{(k)}, \nabla f(\boldsymbol{x}^{(k)}) \right\rangle \leq f(\boldsymbol{y}),$$

where the left-hand side is the linearization of the objective function, and the inequality is due to the convexity of f. Then, the `FW` method moves towards a minimizer of this linear function since

$$\min_{\boldsymbol{y} \in \mathcal{D}} \left\{ f(\boldsymbol{x}^{(k)}) + \left\langle \boldsymbol{y} - \boldsymbol{x}^{(k)}, \nabla f(\boldsymbol{x}^{(k)}) \right\rangle \right\} \leq \min_{\boldsymbol{y} \in \mathcal{D}} f(\boldsymbol{y}),$$

where the right-hand side is the minimization problem (2.3). Hence, the value

$$g(\boldsymbol{x}) = \max_{\boldsymbol{y} \in \mathcal{D}} \langle \boldsymbol{x} - \boldsymbol{y}, \nabla f(\boldsymbol{x}) \rangle$$

serves as the upper bound of $f(\boldsymbol{x}^{(k)}) - \min_{\boldsymbol{y} \in \mathcal{D}} f(\boldsymbol{y})$. The value $g(\boldsymbol{x}^{(k)})$ is known as the *duality gap* of the current point $\boldsymbol{x}^{(k)}$. To summarize, the `FW` algorithm works as follows:

- Let $\boldsymbol{x}^{(0)} \in \mathcal{D}$.
- For $(k = 0, \cdots, K)$ perform $\begin{cases} \boldsymbol{y} \leftarrow \operatorname{argmin}_{\boldsymbol{y}\in\mathcal{D}} \langle \boldsymbol{y}, \nabla f(\boldsymbol{x}^{(k)}) \rangle \\ \boldsymbol{x}^{(k+1)} = \frac{k}{k+2}\boldsymbol{x}^{(k)} + \frac{2}{k+2}\boldsymbol{y} \end{cases}$.

The convergence of the `FW` algorithm is based on a measure of *non-linearity* of the objective function f in the domain $\mathcal{D}$. The *curvature* C_f of f with respect to $\mathcal{D}$ is defined as

$$\sup_{\substack{\boldsymbol{x},\boldsymbol{y}\in\mathcal{D} \\ \gamma\in[0,1] \\ \boldsymbol{z}=\boldsymbol{x}+\gamma(\boldsymbol{y}-\boldsymbol{x})}} \frac{2}{\gamma^2}\left(f(\boldsymbol{z}) - f(\boldsymbol{x}) + \langle \boldsymbol{z}-\boldsymbol{x}, \nabla f(\boldsymbol{x})\rangle\right). \tag{2.4}$$

A bounded C_f means that $\nabla f(\boldsymbol{z})$ from the linearization of f given by $\nabla f(\boldsymbol{x})$ is bounded. In (2.4), $f(\boldsymbol{z}) - f(\boldsymbol{x}) + \langle \boldsymbol{z}-\boldsymbol{x}, \nabla f(\boldsymbol{x})\rangle$ is also known as the *Bregman divergence* defined by f. The guarantee of `FW` method is that:

Theorem 1. *For any $k \in \mathbb{N}$, the iterate $\boldsymbol{x}^{(k)}$ of the `FW` algorithm satisfies*

$$f(\boldsymbol{x}^{(k)}) - \min_{\boldsymbol{y}\in\mathcal{D}} f(\boldsymbol{y}) \leq \frac{2C_f}{k+2}.$$

The proof of the convergence relies on expressing the improvement per step in terms of the current duality gap.

2.7 Envelope theorem

The *envelope theorem* was used for the first time by economists to solve concave optimization problems appearing in demand theory [MS02]. It was later extended to study incentive constraints in contract theory and game theory to examine nonconvex production problems and to develop the theory of monotone or robust comparative statistics. Here we present a variant of the envelope theorem from Lemma 7 in the paper by Wang et al. [WTP21]:

Theorem 2. *Let $\mathbb{X}$ be a metric space and Y be a non-empty open subset in $\mathbb{R}^K$. Let $u : \mathbb{X} \times Y \to \mathbb{R}$ and assume $\frac{\partial u}{\partial y}$ exists and is continuous in $\mathbb{X} \times Y$. For each $y \in Y$, let $x^\star(y)$ be the minimizer of $u(x, y)$ over $x \in \mathbb{X}$. Set*

$$v(y) = u(x^\star(y), y).$$

Assume that $x^\star : Y \times \mathbb{X}$ is a continuous function. Then, v is continuously differentiable and

$$\frac{d}{dy}v(y) = \frac{\partial u}{\partial y}(x^\star(y), y).$$

In Paper C, the algorithm at each round computes the gradient of the function

$$f(\boldsymbol{\omega}, \boldsymbol{\mu}) = \inf_{\boldsymbol{\lambda}\in C} \sum_{k=1}^{K} \frac{\omega_k(\mu_k - \lambda_k)^2}{2}$$

with respect to the first parameter, where $\mathcal{C}$ is a convex set. The gradient $\nabla_{\boldsymbol{\omega}} f(\boldsymbol{\omega}, \boldsymbol{\mu})$ can be evaluated by envelope theorem, which yields that:

$$\nabla_{\boldsymbol{\omega}} f(\boldsymbol{\omega}, \boldsymbol{\mu}) = \sum_{k=1}^{K} \frac{(\mu_k - \lambda_k^\star)^2}{2} \boldsymbol{e}_k,$$

where $\boldsymbol{e}_k$ is the vector with 1 only on the kth element and 0's elsewhere, $\boldsymbol{\lambda}^\star \in \operatorname{argmin}_{\boldsymbol{\lambda} \in \mathsf{cl}(\mathcal{C})} \sum_{k=1}^{K} \frac{\omega_k(\mu_k, \lambda_k)^2}{2}$, and $\mathsf{cl}\left(\cdot\right)$ is the closure of the given set.

2.8 Best arm identification

This section is served to introduce the terminology of paper C. Refer to the book by Lattimore and Szepesvári [LS20] for a more detailed introduction on the topic.

Best arm identification is one of the pure exploration tasks. There are two settings: *fixed-confidence* and *fixed-budget* settings. For the fixed-confidence setting, the goal is to identify the best answer using as few samples as possible. For the fixed-budget setting, the goal is to minimize the error probability after a fixed number of samples. Paper C focuses on the fixed-confidence setting. There are two notions of efficiency: the *statistical* efficiency and the *computational* efficiency. Statistical efficiency refers to the theoretical guarantee of the expected sample complexity. We focus on the *instance-dependent* sample complexity bounds which means the sample complexity bounds depend on each specific instance, rather than uniformly hold for all the instances. Moreover, we say that an algorithm is *statistically optimal* if its expected sample complexity matches the expected sample complexity lower bound by [GK16] asymptotically as δ goes to 0. Computational efficiency refers to having the algorithm run in time polynomial in the problem parameters (such as the number of arms).

For the best arm identification problem, the arms may have *structures*. We say that a setting has an underlying structure if knowing the expected reward of an arm helps to know the expected rewards of the other arms. For example, in the linear-structured setting, there are K arms with known features $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_K \in \mathbb{R}^d$ and a hidden unknown parameter $\boldsymbol{\theta} \in \mathbb{R}^d$. The expected reward of arm i is $\langle \boldsymbol{a}_i, \boldsymbol{\theta} \rangle$, so pulling one arm reveals part of the information of the other arms. Similarly, we say that a setting is *unstructured* if knowing the expected reward of an arm does not help to know the expected rewards of other arms. This setting is also known as the classical setting which has K arms with mean $\mu_1, \cdots, \mu_K \in \mathbb{R}$ and the goal is to identify the arm with the highest expected reward. In Paper C, we study the combinatorial structure with *semi-bandit* feedbacks. In the semi-bandit feedback model, the individual rewards of all the arms in the pulled action are observed. This contrasts to the *full-bandit* feedback, where only the sum of the rewards of the pulled action is observed.

2.9 Overview of the state-of-the-art results

In this section, we discuss the competitive methods before Paper A, Paper B, Paper C, and Paper D.

Discovering k-conflicting Groups in Signed Networks (NeurIPS'20)

In Paper A, we aim to detect k disjoint groups such that there are mostly positive intra-group edges and mostly negative inter-group edges. Different from the signed clustering [CDGT19], we allow nodes to be *neutral* with respect to the conflicting structure.

Random-Eigensign. For the special case of $k = 2$, the problem was studied by Bonchi et al. [BGG+19]. They proposed the following problem formulation:

$$\max_{\substack{S_1, S_2 \subseteq V \\ S_1 \cap S_2 = \emptyset}} \frac{\sum_{\substack{(i,j) \in E(S_h) \\ h=1,2}} A_{i,j} + \sum_{(i,j) \in E(S_1, S_2)} (-A_{i,j})}{|S_1 \cup S_2|} = \max_{\boldsymbol{x} \in \{0, \pm 1\}^n} \frac{\boldsymbol{x}^T \mathbf{A} \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}}, \quad (2.5)$$

where $\mathbf{A}$ is the signed adjacency matrix. In other words, (2.5) can be compactly written as a Rayleigh quotient optimization problem. They showed that 2-conflicting group detection problem (2.5) is $\mathbf{APX}$-hard by reduction from Correlation Clustering [BBC04], and proposed a $\mathcal{O}(\sqrt{n})$-approximation randomized rounding algorithm as follows:

$$\tilde{u}_k = \begin{cases} \text{Bernoulli}(|u_k|) & \text{if } u_k \geq 0 \\ -\text{Bernoulli}(|u_k|) & \text{if } u_k < 0 \end{cases}, \forall k = 1, \cdots, n,$$

where $\boldsymbol{u}$ is the top eigenvector of $\mathbf{A}$.

QP-Ratio. The problem formulation (2.5) proposed by Bonchi et al. [BGG+19] is also studied by Bhaskara et al. [BCMV12] as a *QP-Ratio* problem. They also showed that solving 2-conflicting group detection problem (2.5) is $\mathbf{APX}$-hard by reducing from *Max k-AND* [Fei02] and ratio version of Unique Games Conjecture. They considered a SDP relaxation as follows:

$$\max_{\boldsymbol{w}_1, \cdots, \boldsymbol{w}_n} \sum_{i,j} A_{ij} \langle \boldsymbol{w}_i, \boldsymbol{w}_j \rangle \text{ subject to } \sum_i \boldsymbol{w}_i^2 = 1 \text{ and } |\langle \boldsymbol{w}_i, \boldsymbol{w}_j \rangle| \leq \|\boldsymbol{w}_i\|_2^2, \forall i, j. \quad (2.6)$$

They proposed a rounding algorithm based on the SDP solution. It achieves $\mathcal{O}(n^{\frac{1}{3}} \ln n)$-approximation for general graphs and $\mathcal{O}(n^{\frac{1}{4}} \ln^2 n)$-approximation for bipartite graphs.

Finding k-Oppositive Cohesive Groups (`FOCG`). Chu et. al. [CWP+16] studies k-conflicting group detection with a different problem formulation. They formulate the problem as trace-maximization, where each group is represented as a simplex with nonzero entries indicating the participation of the nodes in the groups. However, their proposed method `FOCG` finds conflicting groups only within local regions and is sensitive to initialization, often converging to local maxima.

Signed Positive Over Negative Generalized Eigenproblem (`SPONGE`). The state of the art in signed clustering is `SPONGE` [CDGT19]. It is based on a generalized eigenvalue problem for constrained clustering and works especially well on sparse graphs and large k on the stochastic block model. As we will see in Paper A, `SPONGE` does not perform well on real-world graphs.

Improved Analysis of Randomized SVD for Top-Eigenvector Approximation (AISTATS'22)

Prior to our Paper B, the theoretical guarantee of `RSVD` [HMT11] for the metric

$$R(\hat{\mathbf{u}}) = \frac{\hat{\mathbf{u}}^T \mathbf{A} \hat{\mathbf{u}}}{\lambda_1(\mathbf{A})}$$

that we are interested in is shown only for positive semi-definite matrices. Musco and Musco [MM15] showed that `RSVD` using $\mathcal{O}(n \log n)$ space and q passes results in

$$\mathbb{P}\left[R(\hat{\mathbf{u}}) \geq 1 - \mathcal{O}\left(\frac{\ln n}{q}\right)\right] \geq 1 - n^{-\Omega(1)}.$$

This analysis shown by Simchowitz et al. [SEAR18] is tight as there exists a matrix such that `RSVD` fails to find a vector $\hat{\mathbf{u}}$ such that

$$R(\hat{\mathbf{u}}) \geq \frac{23}{24}$$

with high probability within $\mathcal{O}(\ln n)$ passes. We can summarize the status as follows: Let q be the number of passes. Before our work, there was no theoretical analysis for the $o(\ln n)$-pass regime for `RSVD`.

$$\begin{array}{ccl} \text{no guarantee} & R(\hat{\mathbf{u}}) = \Omega(1) & \\ \hline & & \longrightarrow q \\ o(\ln n)\text{-pass} & \Omega(\ln n)\text{-pass} & \end{array}$$

Our contribution is that: for positive semidefinite matrices, we provide the analysis of `RSVD` with $\mathcal{O}(nd)$-space for any $d \in \mathbb{N}$ to have the guarantee of $R(\hat{\mathbf{u}})$ to be

$$\begin{array}{ccl} R(\hat{\mathbf{u}}) = \Omega\left(\left(\frac{d}{n}\right)^{\frac{1}{2q+1}}\right) & R(\hat{\mathbf{u}}) = \Omega(1) & \\ \hline & & \longrightarrow q \\ o(\ln n)\text{-pass} & \Omega(\ln n)\text{-pass} & \end{array}$$

We also provide a non-trivial guarantee for $\hat{\mathbf{u}}$ found by `RSVD` for matric $R(\hat{\mathbf{u}})$ for some indefinite matrices. Our technique is based on establishing a novel connection between $R(\hat{\mathbf{u}})$ for $\hat{\mathbf{u}}$ found by `RSVD` and the length of projecting any vector onto a low-dimensional random subspace. Please refer to Section 2.2 for a brief introduction to the random projection.

Closing the Computational-Statistical Gap in Best Arm Identification for Combinatorial Semi-Bandits (NeurIPS'23)

Here we focus on the combinatorial best arm identification with fixed confidence and semi-bandit feedbacks. Before Paper C, there exist statistically optimal algorithms such as `CombGame` [JMKK21], but there are no computationally efficient algorithms.

CombGame. `CombGame` interprets the optimization problem related to the expected sample complexity lower bound as a two-player game with a $\boldsymbol{\omega}$-player and a $\boldsymbol{\lambda}$-player:

$$T^\star(\boldsymbol{\mu})^{-1} = \sup_{\boldsymbol{\omega}\in\Sigma} F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) \quad \text{with} \quad F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) = \inf_{\boldsymbol{\lambda}\in\mathsf{Alt}(\boldsymbol{\mu})} \sum_{k=1}^{K} \omega_k d(\mu_k, \lambda_k), \tag{2.7}$$

where $\Sigma = \{\sum_{\boldsymbol{x}\in\mathcal{X}} w_{\boldsymbol{x}}\boldsymbol{x} : \boldsymbol{w} \in \Sigma_{|\mathcal{X}|}\}$, $\mathsf{Alt}(\boldsymbol{\mu}) = \{\boldsymbol{\lambda} \in \Lambda : \boldsymbol{i}^\star(\boldsymbol{\lambda}) \neq \boldsymbol{i}^\star(\boldsymbol{\mu})\}$, Σ_N is $(N-1)$-dimensional simplex, and $\boldsymbol{i}^\star(\cdot) \in \operatorname{argmax}_{\boldsymbol{x}\in\mathcal{X}} \langle \cdot, \boldsymbol{x}\rangle$.

An algorithm that computes the optimal solution,

$$\boldsymbol{\lambda}^\star(\boldsymbol{\omega}, \boldsymbol{\mu}) \in \operatorname*{argmin}_{\boldsymbol{\lambda}\in\mathsf{cl}(\mathsf{Alt}(\boldsymbol{\mu}))} \sum_{k=1}^{K} \omega_k d(\mu_k, \lambda_k),$$

which is known as the *most confusing parameter* (`MCP`), to the inner optimization $F_{\boldsymbol{\mu}}(\boldsymbol{\omega})$ in Equation (2.7) is called a `MCP` oracle, where $\mathsf{cl}(\cdot)$ is the closure of the given set. `CombGame` uses Frank-Wolfe algorithms, `OFW` [HK12] and `LLOO` [GH16], for the $\boldsymbol{\omega}$-player and uses a `MCP` oracle for the $\boldsymbol{\lambda}$-player. However, they leave the existence of the `MCP` oracle running in polynomial time as an open problem.

All the existing statistically optimal algorithms such as `CombGame` and `FWS` [WTP21] require the `MCP` oracle. More precisely, the `MCP` oracle is used in their stopping rules (known as the *Chernoff stopping rule* [GK16]) and their respective sampling rules.

One of the key contributions in Paper C is that we design an efficient approximate `MCP` oracle based on a property we discovered while applying the Lagrangian multiplier method (see Section 2.3). Then, based on the proposed approximate `MCP` oracle, we design an algorithm that closes the computational-statistical gap of the problem of combinatorial best arm identification with fixed confidence and semi-bandit feedbacks.

Matroid Semi-Bandits in Sublinear Time (ICML'24)

Before our Paper D, there is no sublinear-time algorithm for matroid semi-bandits. The competitive algorithms for this setting include `CUCB` [CWY13], `KL-OSM` [TP16], and `OSSB` [CMP17]. Let $\mathcal{X} \subseteq \{0,1\}^K$ be the action set on K arms, and $\boldsymbol{i}^\star(\boldsymbol{\mu}) \in \operatorname{argmax}_{\boldsymbol{x}\in\mathcal{X}} \langle \boldsymbol{x}, \boldsymbol{\mu}\rangle$ be an optimal action.

Combinatorial Upper Confidence Bound (CUCB). The sampling rule of CUCB is

$$\boldsymbol{x}(t) \in \operatorname*{argmax}_{\boldsymbol{x}\in\mathcal{X}} \sum_{k=1}^{K} \left(\hat{\mu}_k(t-1) + \frac{1.5\log t}{\sqrt{N_k(t-1)}}\right) x_k, \tag{2.8}$$

where $\hat{\mu}_k(t)$ is the empirical mean estimate of arm k at round t, and $N_k(t)$ is the number of pulls of arm k at round t. CUCB achieves a regret bound $R(T) = \mathcal{O}\left(\frac{(K-D)\log T}{\triangle_{\min}}\right)$ that matches the gap-dependent lower bound, where $\triangle_{\min} = \min_{\boldsymbol{x}\in\mathcal{X}:\triangle_{\boldsymbol{x}}>0} \triangle_{\boldsymbol{x}}$ and $\triangle_{\boldsymbol{x}} = \langle \boldsymbol{i}^\star(\boldsymbol{\mu}) - \boldsymbol{x}, \boldsymbol{\mu}\rangle$.

KL-based Efficient Sampling for Matroids (KL-OSM). Under the assumption of Bernoulli reward, KL-OSM at each round selects

$$\boldsymbol{x}(t) \in \operatorname*{argmax}_{\boldsymbol{x}\in\mathcal{X}} \sum_{k=1}^{K} \omega_k(t) x_k, \tag{2.9}$$

where $\omega_k(t) = \max\{q \in [\hat{\mu}_k(t), 1] : N_k(t)\mathsf{kl}(\hat{\mu}_k(t), q) \le \log t + 3\log\log t\}$. KL-OSM achieves a regret bound that matches the instance-specific lower bound

$$\liminf_{T\to\infty} \frac{R(T)}{\log T} \ge \sum_{k\notin \mathsf{supp}(\boldsymbol{i}^\star(\boldsymbol{\mu}))} \frac{\mu_{\sigma(k)} - \mu_k}{\mathrm{kl}(\mu_k, \mu_{\sigma(k)})}, \tag{2.10}$$

where $\sigma(k) = \operatorname{argmin}_{i\in\mathcal{K}_k} \mu_i$ and $\mathcal{K}_i = \{\ell \in \mathsf{supp}\,(\boldsymbol{i}^\star(\boldsymbol{\mu})) : \mathsf{supp}\,(\boldsymbol{i}^\star(\boldsymbol{\mu}))\setminus\{\ell\} \cup \{i\} \in \mathcal{I}\}$.

However, both algorithms, CUCB and KL-OSM, require per-round time complexity of at least $\Omega(K)$. Their sampling rules (4.1) and (2.9) rely on a greedy algorithm (see Section 2.5) to find a maximum-weight basis (2.2), so the time complexity is upper bounded by $\mathcal{O}(K\ln K)$ plus K membership oracle calls.

Optimal Sampling for Structured Bandits (OSSB). OSSB is an asymptotically instance-specific optimal algorithm. In each round, it solves the Graves-Lai optimization problem [CCG21a], which has a simplified expression for the matroid case, as shown on the right-hand side of Equation (2.10). The computational cost is DK membership oracle calls.

Our contribution of Paper D is that we develop the first algorithm whose per-round time complexity is sublinear in K.

2.10 Discussion of the most important related works

Closing the Computational-Statistical Gap in Best Arm Identification for Combinatorial Semi-Bandits (NeurIPS'23)

In this section, we present the most important works in best arm identification (BAI) with fixed confidence. We start with the unstructured setting and then discuss the state-of-the-art methods in the structured settings.

Track-and-Stop. In the unstructured setting, `Track-and-Stop` [GK16] is the first work to achieve statistical optimality and computational efficiency. Its design has inspired many subsequent works to design a statistically optimal algorithm for other settings such as linear BAI [JP20], ϵ-BAI [GK21], and BAI with multiple correct answers [DK19]. Here we summarize the ideas of `Track-and-Stop`.

Let Λ be the mean parameter space. Garivier and Kaufmann [GK16] derived the expected sample complexity lower bound for the unstructured setting:

$$T^\star(\boldsymbol{\mu})^{-1} = \sup_{\boldsymbol{\omega}\in\Sigma_K} F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) \quad \text{with} \quad F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) = \inf_{\boldsymbol{\lambda}\in\mathsf{Alt}(\boldsymbol{\mu})} \sum_{k=1}^{K} \omega_k d(\mu_k, \lambda_k), \tag{2.11}$$

where Σ_K is the $(K-1)$-dimensional simplex, $\mathsf{Alt}(\boldsymbol{\mu}) = \{\boldsymbol{\lambda} \in \Lambda : \boldsymbol{i}^\star(\boldsymbol{\lambda}) \neq \boldsymbol{i}^\star(\boldsymbol{\mu})\}$ is called the *confusing parameters*, and $\boldsymbol{i}^\star(\boldsymbol{\mu}) \in \operatorname{argmax}_{k\in[K]} \mu_k$ is the optimal arm. Equation (2.11) is derived based on the so-called *transportation lemma* [KCG16].

The optimal solution to the inner optimization $F_{\boldsymbol{\mu}}(\boldsymbol{\omega})$ in Equation (2.11) is called the *most confusing parameter* (`MCP`):

$$\boldsymbol{\lambda}^\star(\boldsymbol{\omega}, \boldsymbol{\mu}) \in \operatorname*{argmin}_{\boldsymbol{\lambda}\in\mathsf{cl}(\mathsf{Alt}(\boldsymbol{\mu}))} \sum_{k=1}^{K} \omega_k d(\mu_k, \lambda_k).$$

The optimal solution to the optimization problem of $T^\star(\boldsymbol{\mu})^{-1}$ is called the *optimal allocation*:

$$\boldsymbol{\omega}^\star(\boldsymbol{\mu}) \in \operatorname*{argmax}_{\boldsymbol{\omega}\in\Sigma_K} F_{\boldsymbol{\mu}}(\boldsymbol{\omega}).$$

They showed that the optimal allocation $\boldsymbol{\omega}^\star(\boldsymbol{\mu})$ is *unique* and computable through the binary search. Moreover, they presented a stopping rule called *Chernoff stopping rule*: Let $\hat{\boldsymbol{\mu}}(t)$ be the empirical mean estimate and $\hat{\boldsymbol{\omega}}(t)$ be the empirical allocation. Then,

$$tF_{\hat{\boldsymbol{\mu}}(t)}(\hat{\boldsymbol{\omega}}(t)) \geq \beta(\delta, t) \implies \mathbb{P}_{\boldsymbol{\mu}}[\boldsymbol{i}^\star(\hat{\boldsymbol{\mu}}(t)) \neq \boldsymbol{i}^\star(\boldsymbol{\mu})] \leq \delta, \tag{2.12}$$

where $\beta(\delta, t)$ is a exploration threshold.

The idea of `Track-and-Stop` is that: since $\boldsymbol{\omega}^\star(\boldsymbol{\mu})$ is unique, by *forced exploration* procedure which uniformly explores all arms and by the law of large numbers, we have $\boldsymbol{\omega}^\star(\hat{\boldsymbol{\mu}}(t)) \to \boldsymbol{\omega}^\star(\boldsymbol{\mu})$ almost surely as $t \to \infty$. It remains to ensure $\hat{\boldsymbol{\omega}}(t) \to \boldsymbol{\omega}^\star(\boldsymbol{\mu})$ almost surely as $t \to \infty$. The way that they ensure $\hat{\boldsymbol{\omega}}(t) \to \boldsymbol{\omega}^\star(\boldsymbol{\mu})$ almost surely as $t \to \infty$ is to use one of the following tracking rules:

- (C-Tracking) Pull

$$A_{t+1} \leftarrow \operatorname*{argmax}_{k=1}^{K} \left(\sum_{s=1}^{t} \omega_k^{\epsilon_s}(\hat{\boldsymbol{\mu}}(s)) - N_k(t) \right),$$

  where $\boldsymbol{\omega}^\epsilon(\boldsymbol{\mu})$ is a L^∞ projection of $\boldsymbol{\omega}^\star(\boldsymbol{\mu})$ onto $\{\boldsymbol{\omega} \in [\epsilon, 1]^K : \sum_{k=1}^{K} \omega_k = 1\}$ and $\epsilon_t = \frac{1}{2\sqrt{K^2+t}}$.

- (D-Tracking) Pull

$$A_{t+1} \leftarrow \operatorname*{argmax}_{k=1}^{K} \left(t\omega_k^\star(\hat{\boldsymbol{\mu}}(t)) - N_k(t)\right).$$

Apart from `Track-and-Stop`, there are two other main frameworks in designing statistically optimal algorithms for structured settings. One is the *gamification* approach [DKM19] and the other is Frank-Wolfe Sampling (`FWS`) [WTP21].

Gamification. For the unstructured setting, Degenn et al. [DKM19] proposed the gamification approach which interprets $T^\star(\boldsymbol{\mu})^{-1}$ as a two-player zero-sum game:

$$T^\star(\boldsymbol{\mu})^{-1} = \sup_{\boldsymbol{\omega}\in\Sigma_K} \inf_{\boldsymbol{\lambda}\in\mathsf{Alt}(\boldsymbol{\mu})} \sum_{k=1}^{K} \omega_k d(\mu_k, \lambda_k) = \inf_{\boldsymbol{q}} \sup_{k\in[K]} \mathbb{E}_{\boldsymbol{\lambda}\sim\boldsymbol{q}}[d(\mu_k, \lambda_k)],$$

where $\boldsymbol{q}$ is a distribution over $\mathsf{Alt}(\boldsymbol{\mu})$. Note that the minimax theorem (see Section 2.4) requires both domains to be convex, but $\mathsf{Alt}(\boldsymbol{\mu})$ is not a convex set. As far as I know, there is no explanation for why the second equality holds. In addition, this approach requires an assumption on the set of mean parameters $\Lambda \subseteq [\mu_{\min}, \mu_{\max}]^K$ that must be bounded. For the stopping rule, they use the Chernoff stopping rule (2.12). For the sampling rule, they propose three schemes, but we limit our focus to the first two, which require the `MCP` oracle. Here, a `MCP` oracle is an algorithm that computes $F_{\boldsymbol{\mu}}(\boldsymbol{\omega})$ in (2.11). The two sampling rule schemes are as follows:

- $\boldsymbol{\omega}$-player plays the first and uses a regret minimization algorithm for linear losses on the simplex to produce $\boldsymbol{\omega}(t) \in \Sigma_K$. The $\boldsymbol{\lambda}$-player uses the *best-response* $\boldsymbol{\lambda}(t) \in \inf_{\boldsymbol{\lambda}\in\mathsf{Alt}(\hat{\boldsymbol{\mu}}(t-1))} \sum_{k=1}^{K} \omega_k(t) d(\hat{\mu}_k(t-1), \lambda_k)$. The computational cost is dominated by one `MCP` oracle call.
- $\boldsymbol{\lambda}$-player plays the first and uses Follow-the-Perturbed-Leader which chooses a distribution of

$$\operatorname*{argmin}_{\boldsymbol{\lambda}\in\mathsf{Alt}(\boldsymbol{\mu})} \left\{ \sum_{s=1}^{t-1} d(\hat{\mu}_{k(s)}(s-1), \lambda_{k(s)}) + \sum_{k=1}^{K} \sigma_k(t) d(\hat{\mu}_k(t-1), \lambda_k) \right\}, \tag{2.13}$$

where $\boldsymbol{\sigma}(t)$ is a random vector with i.i.d. exponentially distributed coordinates. The $\boldsymbol{\omega}$-player plays the best-response $k(t) \in \operatorname{argmax}_{k\in[K]} U_k(t)$, where

$$U_k(t) = \max\left\{ \frac{\ln(t-1)}{N_k(t-1)}, \max_{\xi\in\{\alpha_k(t),\beta_k(t)\}} \mathbb{E}_{\boldsymbol{\lambda}\sim\boldsymbol{q}(t)}[d(\xi, \lambda_k)] \right\}$$

and $[\alpha_k(t), \beta_k(t)] = \{\xi : N_k(t-1)d(\hat{\mu}_k(t-1), \xi) \le \ln(t-1)\}$. $\boldsymbol{\lambda}$-player suffers loss $\mathbb{E}_{\boldsymbol{\lambda}\sim\boldsymbol{q}(t)}\left[d(\hat{\mu}_k(t-1), \lambda_{k(t)})\right]$. They showed in Appendix E.2 that (2.13) can be computed by one `MCP` oracle call, but to estimate the distribution, they use t empirical samples, so the computation cost is dominated by t `MCP` oracle calls.

The expected sample complexity of the gamification approach is asymptotically optimal. However, their sample complexity upper bound is not explicit in problem parameters such as K. This approach has been extended to structured settings, such as linear [Sha21] and combinatorial [JMKK21] settings.

Frank-Wolfe Sampling (`FWS`). `FWS` [WTP21] achieves instance-specifically statistical optimality in pure exploration tasks (e.g., best arm identification [GK16], threshold bandits [LGC16]) under various structures (e.g., unimodal, Lipschitz, convex, linear). Wang et al. [WTP21] provide the first expected sample complexity analysis in the moderate-confidence regime and in the explicit form. Also, in empirical evaluation, the average sample complexity of `FWS` is amongst the smallest of all the statistical optimal algorithms in various structures.

The challenges in designing a generic algorithm lie in:

(i) $\mathsf{Alt}(\boldsymbol{\mu})$ is not a convex set in general

(ii) $F_{\boldsymbol{\mu}}$ is *non-smooth*

(iii) $F_{\boldsymbol{\mu}}$ has unbounded curvature close to the boundary of Σ_K

To deal with the challenge (i), they assume that $\mathsf{Alt}(\boldsymbol{\mu})$ can be partitioned into many convex sets. Then, $F_{\boldsymbol{\mu}}(\boldsymbol{\omega})$ can be written as the minimum of many convex programs

$$F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) = \min_{j\in\mathcal{J}_{\boldsymbol{i}^\star(\boldsymbol{\mu})}} f_j(\boldsymbol{\omega},\boldsymbol{\mu}) \quad \text{and} \quad f_j(\boldsymbol{\omega},\boldsymbol{\mu}) = \inf_{\boldsymbol{\lambda}\in\mathcal{C}_j^{\boldsymbol{i}^\star(\boldsymbol{\mu})}} \sum_{k=1}^{K} \omega_k d(\mu_k,\lambda_k)$$

for $j \in \mathcal{J}_{\boldsymbol{i}^\star(\boldsymbol{\mu})}$ and $\mathcal{C}_j^{\boldsymbol{i}^\star(\boldsymbol{\mu})} = \{\boldsymbol{\mu}\in\Lambda : \mu_j > \mu_{\boldsymbol{i}^\star(\boldsymbol{\mu})}\}$.

For challenge (ii)(iii), observe that each $f_j(\boldsymbol{\omega},\boldsymbol{\mu})$ for each $j\in\mathcal{J}_{\boldsymbol{i}^\star(\boldsymbol{\mu})}$ is smooth, and $F_{\boldsymbol{\mu}}$ is nonsmooth only at the point where $f_j(\boldsymbol{\omega},\boldsymbol{\mu}) = f_j'(\boldsymbol{\omega},\boldsymbol{\mu})$ for $j \neq j'$. They replace the subdifferential at $\boldsymbol{\omega}$ by taking the convex hull of gradients $\nabla f_j(\boldsymbol{\omega},\boldsymbol{\mu})$ satisfying $f_j(\boldsymbol{\omega},\boldsymbol{\mu}) < F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) + r$ for some small $r>0$. They propose a Frank-Wolfe-based sampling rule, which we elaborate on the design here.

To overcome the challenge (ii), in addition to the standard Frank-Wolfe algorithm (see Section 2.6), one can take the minimum of $\langle \boldsymbol{z} - \boldsymbol{x}(t-1), \boldsymbol{h}\rangle$ taken over all $\boldsymbol{h}$ in the subdifferential subspace of $F_{\hat{\boldsymbol{\mu}}(t-1)}(\boldsymbol{x}(t-1))$:

$$\begin{cases} \boldsymbol{z}(t) \leftarrow \operatorname{argmax}_{\boldsymbol{z}\in\Sigma} \min_{\boldsymbol{h}\in\partial F_{\hat{\boldsymbol{\mu}}(t-1)}(\boldsymbol{x}(t-1))} \langle \boldsymbol{z} - \boldsymbol{x}(t-1), \boldsymbol{h}\rangle, \\ \boldsymbol{x}(t) \leftarrow \frac{t-1}{t}\boldsymbol{x}(t-1) + \frac{1}{t}\boldsymbol{z}(t), \end{cases}$$

where $\min_{\boldsymbol{h}\in\partial F_{\hat{\boldsymbol{\mu}}(t-1)}(\boldsymbol{x}(t-1))} \langle \boldsymbol{z} - \boldsymbol{x}(t-1), \boldsymbol{h}\rangle$ is an approximate between the maximum and the current iterate.

To overcome the challenge (iii), they construct the *r-subdifferential subspace*:

$$H_{F_{\boldsymbol{\mu}}}(\boldsymbol{\omega}, r) = \mathsf{cov}\left\{\nabla f_j(\boldsymbol{\omega},\boldsymbol{\mu}) : j\in\mathcal{J}_{\boldsymbol{i}^\star(\boldsymbol{\mu})}, f_j(\boldsymbol{\omega},\boldsymbol{\mu}) < F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) + r\right\},$$

and present the Frank-Wolfe sampling rule as follows:

$$\begin{cases} \boldsymbol{z}(t) \leftarrow \operatorname{argmax}_{\boldsymbol{z}\in\Sigma} \min_{\boldsymbol{h}\in H_{F_{\hat{\boldsymbol{\mu}}(t-1)}}(\boldsymbol{x}(t-1),r_t)} \langle \boldsymbol{z} - \boldsymbol{x}(t-1), \boldsymbol{h} \rangle, \\ \boldsymbol{x}(t) \leftarrow \frac{t-1}{t}\boldsymbol{x}(t-1) + \frac{1}{t}\boldsymbol{z}(t), \end{cases}$$

where $\nabla f_j(\boldsymbol{\omega}, \boldsymbol{\mu})$ can be evaluated by the envelope theorem (see Section 2.7), and the action to be pulled is computed by the following tracking rule:

$$A_t \leftarrow \operatorname*{argmax}_{k\in[K]} \frac{x_k(t)}{\hat{\omega}_k(t-1)}.$$

`FWS` also uses the Chernoff stopping rule (2.12) as the stopping rule.

However, `FWS` is not computationally efficient in combinatorial structure. Specifically, it requires the `MCP` oracle not only in the Chernoff stopping rule (2.12) but also in the construction of the r-subdifferential space $H_{\hat{\boldsymbol{\mu}}(t-1)}(\boldsymbol{x}(t-1), r_t)$. This motivates our work on Paper C [TWPL23], where we develop an efficient `MCP` oracle and use a different smoothing technique.

Chapter 3

Graph mining

In graph mining, we explicitly specify the kind of graph patterns we aim to find. This is achieved by formulating the problem such that the objective function, given a solution, reflects how well that solution matches the desired pattern. The algorithm typically makes no assumptions about the graph, and performance evaluation is performed using a worst-case analysis to approximate the optimal solution. The graphs used for experiments are typically large and originate from real-world data.

In Paper A, we consider a specific signed graph mining task called *conflicting group detection*. We proposed an approximation algorithm to efficiently solve the problem. In Paper B, we consider top-eigenvector computation (which can be used to detect many graph patterns) and sharpen the multiplicative analysis of *Randomized SVD* [HMT11] in the pass-efficient and memory-efficient setting.

3.1 Discovering k-conflicting Groups in Signed Networks (NeurIPS'20)

Summary

In Paper A, we consider the detection of a specific pattern in a signed network. The pattern we aim to detect is called *k-conflicting groups*, which are k mutually-disjoint node sets $S_1, \ldots, S_k \subseteq V$ that have the following informally-stated properties:

Assumption 1. *For all $i, j \in [k]$, with $i \neq j$, the edges in $E(S_i)$ are mostly positive, whereas the edges in $E(S_i, S_j)$ are mostly negative.*

Assumption 2. *There should be a large number of interactions among the nodes of $S_1, \ldots, S_k$ relative to the total number of nodes in these groups. In other words, the subgraph induced by $S_1, \ldots, S_k$ should be as dense as possible.*

To evaluate the quality of the k conflicting groups, we design the objective:

$$\max_{\substack{S_1,\cdots,S_k \subseteq V: \\ S_i \cap S_j = \emptyset, \forall i,j}} \frac{\sum_{\substack{(i,j)\in E(S_h,S_h) \\ h=1,\cdots,k}} A_{i,j} - \frac{1}{k-1}\sum_{h\neq \ell}\sum_{(i,j)\in E(S_h,S_\ell)} A_{i,j}}{\sum_{h\in[k]}|S_h|}. \tag{3.1}$$

The numerator of (3.1) accounts for Assumption 1 and the division by the total sizes of the groups accounts for Assumption 2. This objective is a generalization of a previous work that deals with the special case of $k = 2$ [BGG+19]. We observe that the relationship among the k conflicting groups can be represented by the Laplacian of complete graphs on k nodes. By leveraging the special eigenspace structure of the Laplacian matrix of a complete graph (see Daniel A. Spielman's lecture note [Dan19]), we can rewrite the objective function (3.1) as

$$\max_{\mathbf{Y}\in\mathbb{R}^{n\times(k-1)}\setminus\{\mathbf{0}_{n\times(k-1)}\}} \frac{\mathrm{Tr}(\mathbf{Y}^T\mathbf{A}\mathbf{Y})}{\mathrm{Tr}(\mathbf{Y}^T\mathbf{Y})}, \tag{3.2}$$

subject to $Y_{i,j} = \begin{cases} c_j(k-j) & \text{if } i \in S_j \\ 0 & \text{if } i \in \cup_{h=1}^{j-1} S_h \text{ or } i \notin \cup_{h\in[k]} S_h \\ -c_j & \text{if } i \in \cup_{h=j+1}^{k} S_h \end{cases}$, where $\{c_j\}_{j\in[k-1]}$ are constants. Under the reformulation of (3.2), we extend the approach of [BGG+19] to detect any k conflicting groups [TOG20] in a sequential manner: For any $j \in [k]$, suppose $S_1, \cdots, S_{j-1}$ are detected, and the j-th group is formed as $S_j = \{i \in [n] : \mathbf{u}_i^\star = k - j\}$, where $\mathbf{u}^\star$ is the optimal solution to (3.3):

$$\boldsymbol{u}^\star \in \operatorname{argmax}\left\{\frac{\boldsymbol{x}^T\mathbf{A}^{(j-1)}\boldsymbol{x}}{\boldsymbol{x}^T\boldsymbol{x}} : \boldsymbol{x} \in \{-1, 0, k-j\}^n\setminus\{\mathbf{0}\}\right\}, \tag{3.3}$$

$\mathbf{A}^{(j-1)}$ is the adjacency matrix after removing $\cup_{h\in[j-1]} S_h$, and $\mathbf{A}^{(0)} = \mathbf{A}$. We refer to the problem of (3.3) as `Max-DRQ` problem.

Challenge

`Max-DRQ` (3.3) is an APX-hard problem since a special case of $k = 2$ has been shown to be APX-hard [BGG+19, BCMV12]. For $k = 2$, the eigenvector-based rounding algorithm by [BGG+19] results in $\mathcal{O}(\sqrt{n})$-approximation factor and there is an instance such that any eigenvector-based rounding algorithm suffers from a $\Omega(\sqrt{n})$-approximation factor. The state-of-the-art is a $\mathcal{O}(n^{\frac{1}{3}} \ln n)$-approximation SDP-based rounding algorithm [BCMV12], based on the following SDP relaxation:

$$\max_{\boldsymbol{w}_1,\cdots,\boldsymbol{w}_n} \sum_{i,j} A_{ij}\,\langle \boldsymbol{w}_i, \boldsymbol{w}_j\rangle \text{ subject to } \sum_i \boldsymbol{w}_i^2 = 1 \text{ and } |\langle \boldsymbol{w}_i, \boldsymbol{w}_j\rangle| \leq \|\boldsymbol{w}_i\|_2^2, \forall i, j. \tag{2.6}$$

However, it is hard to generalize (2.6) to the case when $k > 2$.

Our approach

We proposed a $\mathcal{O}((k-j)\sqrt{n})$-approximation eigenvector-based rounding algorithm called `RandomRound` to solve (3.3). Let $q = k - j$, where j goes from 1 to $k - 1$. `RandomRound` rounds an eigenvector $\boldsymbol{u} \in \mathbb{R}^k$ of $\mathbf{A}^{(k-q-1)}$ onto $\{0, -1, q\}^n$ by drawing Bernoulli trials. For each coordinate u_k, we set

$$\tilde{u}_k = \begin{cases} q \cdot \text{Bernoulli}(|u_k|) & \text{if } u_k \geq 0 \\ -1 \cdot \text{Bernoulli}(|u_k|) & \text{if } u_k < 0 \end{cases}, \forall k = 1, \cdots, n. \tag{3.4}$$

In this way, we have $\mathbb{E}[\tilde{\boldsymbol{u}}] = \boldsymbol{u}$. `RandomRound` generalizes [BGG+19], and has the following theoretical guarantee:

Theorem 1. *Let $\boldsymbol{u}$ be the leading eigenvector of the adjacency matrix $\mathbf{A}^{(k-q-1)}$ of a signed graph, and let $q \geq 1$ be a positive integer. Then, the `RandomRound` algorithm with $(\boldsymbol{u}, q)$ as input is a $(q\sqrt{n})$-approximation to the optimum of the corresponding `Max-DRQ` problem.*

Lemma 1. *Let OPT be the optimum solution to the `Max-DRQ` problem. There exists a problem instance such that $\lambda_1(\mathbf{A}) \geq OPT \cdot \Omega(\sqrt{n})$.*

Corollary 1. *The integrality gap of algorithm `RandomRound` is $\Omega(\sqrt{n})$, and thus, the approximation result of Theorem 1 is asymptotically tight up to a factor of q.*

`RandomRound` performs competitively in both the real-world graphs and synthetic networks as compared to the other baselines [CWP+16, CDGT19].

Contributions. A. Gionis and the author of the thesis contributed to the problem formulation. The author of the thesis reformulated the problem and proposed the algorithm. B. Ordozgoiti and the author established the analysis. The initial manuscript was primarily written by the author of the thesis and B. Ordozgoiti, with A. Gionis contributed to subsequent revisions.

3.2 Improved Analysis of Randomized SVD for Top-Eigenvector Approximation (AISTATS'22)

Summary

In Paper B, we study graph mining tasks that can be described as a Rayleigh quotient maximization problem, shown in (3.5), in a pass-efficient and memory-efficient manner.

$$\boldsymbol{x}^\star \in \operatorname{argmax} \left\{ \frac{\boldsymbol{x}^T \mathbf{A} \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} : \boldsymbol{x} \in \mathcal{T} \backslash \{\mathbf{0}\} \right\} \tag{3.5}$$

for some given $\mathcal{T} \subseteq \mathbb{R}^n$ and some given symmetric matrix $\mathbf{A} \subseteq \mathbb{R}^{n \times n}$. Many graph mining tasks, such as fair densest subgraph detection [ABF+20], 2-community detection [New06], conflicting group detection [BGG+19, TOG20] are captured by the formulation of (3.5). A computationally efficient way to solve (3.5) is to (i) compute the top-eigenvector $\hat{\boldsymbol{u}}$ of $\mathbf{A}$ by eigenvector solver and then (ii) round $\hat{\boldsymbol{u}}$ to another vector in $\mathcal{T}$. For step (i), randomized SVD (`RSVD`) [HMT11] is a memory-efficient and pass-efficient method. However, in the $o(\log n)$-pass regime which is of great interest to practitioners, we find that RSVD has no guarantee for the top-eigenvector approximation with respect to the multiplicative ratio objective:

$$R(\hat{\mathbf{u}}) = \lambda_1^{-1} \frac{\hat{\mathbf{u}}^T \mathbf{A} \hat{\mathbf{u}}}{\hat{\mathbf{u}}^T \hat{\mathbf{u}}},$$

where λ_1 is the eigenvalue of the top eigenvector of $\mathbf{A}$. First, the ratio objective $R(\hat{\mathbf{u}})$ provides a finer analysis of the eigenvector-based algorithm. For example, we show in Paper B that: given an approximated top-eigenvector $\hat{\mathbf{u}}$, Random-Eigensign for 2-conflicting group detection (see Section 2.9) is a $\mathcal{O}(\frac{\sqrt{n}}{R(\hat{\mathbf{u}})})$-approximation algorithm. Second, it raises a natural question: whether $\Omega(\log n)$-pass is necessary for RSVD to output an approximate top-eigenvector $\hat{\mathbf{u}}$ with a non-trivial guarantee of $R(\hat{\mathbf{u}})$? We answer this question by sharpening the analysis of RSVD such that it outputs $\hat{\mathbf{u}}$ with a non-trivial guarantee of $R(\hat{\mathbf{u}})$ on positive semi-definite matrices for any number of passes. This result is extended to indefinite matrices under certain conditions. Moreover, we considered a variant of RSVD, which is referred to as RandSum, that uses Johnson-Linderstrauss (JL) distribution and a 0/1-Bernoulli in the random projection step. We demonstrate by experiment that RandSum is helpful to the detection of conflicting groups ([BGG$^+$19] and Paper A) in the sense that the number of passes required to achieve certain performance by RandSum is less than that is required by RSVD.

Challenge

Prior work analyzes the top eigenvector approximation in the additive form. Musco and Musco [MM15] showed that RSVD using $\mathcal{O}(n \log n)$ space and q passes results in $R(\hat{\mathbf{u}}) \geq 1 - \mathcal{O}(\frac{\ln n}{q})$ with probability at least $1 - n^{-\Omega(1)}$. This analysis shown by Simchowitz et al. [SEAR18] is tight as there exists a matrix such that RSVD fails to find a vector $\hat{\mathbf{u}}$ such that $R(\hat{\mathbf{u}}) \geq \frac{23}{24}$ within $\mathcal{O}(\ln n)$ passes. It is unclear whether $q = \Omega(\ln n)$ passes are required for RSVD to output a top-eigenvector approximation $\hat{\mathbf{u}}$ such that the multiplicative gap $R(\hat{\mathbf{u}})$ has a non-trivial guarantee.

Our approach

We avoid using any derivation that will result in an additive form of $R(\hat{\mathbf{u}})$. These include matrix addition and subtraction. Our approach uses only Cauchy-Schwarz inequalities to relate $R(\hat{\mathbf{u}})$ with random projection lemma [HP14]. For positive semi-definite matrices, we have the following results:

Theorem 1. *Let $\mathbf{A}$ be a PSD matrix with $\lambda_1 > 0$ and $\hat{\mathbf{u}} =$ RSVD$(\mathbf{A}, \mathcal{N}(0,1)^{n \times d}, q, d)$ for any $q \in \mathbb{N}$. Then*

$$R(\hat{\mathbf{u}}) = \left(\Omega\left(\frac{d}{n}\right)\right)^{\frac{1}{2q+1}}$$

holds with probability at least $1 - e^{-\Omega(d)}$.

One may wonder whether this analysis is tight, and we show that it is tight up to a constant factor by considering the eigenvalue distribution $\{\alpha_i\}$:

$$1 = \alpha_1 > \alpha_2 = \cdots = \alpha_n = \left(\frac{d}{n}\right)^{\frac{1}{2q+1}},$$

where $\alpha_i = \lambda_i/\lambda_1$.

Theorem 2. *For any $q \in \mathbb{N}$, there exists a positive semidefinite matrix $\mathbf{A}$ with $\lambda_1 > 0$, so that for $\hat{\mathbf{u}} = \texttt{RSVD}(\mathbf{A}, \mathcal{N}(0,1)^{n\times d}, q, d)$, it holds*

$$R(\hat{\mathbf{u}}) = \mathcal{O}\left(\left(\frac{d}{n}\right)^{\frac{1}{2q+1}}\right),$$

with probability at least $1 - e^{-\Omega(d)}$.

While our worst-case analysis is tight, the bad eigenvalue distributions rarely happen in practice. Instead, real-world matrices are often observed to have rapidly decaying singular values [CF06, EG17]. To take this consideration into account, we introduce the following definition to capture whether $\mathbf{A}$ has at least a power-law decay of its singular values $\{\sigma_i\}_{i\geq i_0}^n$.

Definition 1. *Let*

$$i_0 = \begin{cases} \min_{j\in\mathcal{J}} j & \text{if } \mathcal{J} \neq \emptyset, \\ n & \text{otherwise}, \end{cases}$$

where $\mathcal{J} \subseteq [n]$ consists of all the integers $j \in [n]$ such that there exists $\gamma > 1/q$ and $C > 0$ satisfying $\sigma_i/\sigma_1 \leq C \cdot i^{-\gamma}$, for all $i \geq j$.

Theorem 3. *Let $\mathbf{A}$ be a positive semidefinite matrix, $\hat{\mathbf{u}} = \texttt{RSVD}(\mathbf{A}, \mathcal{N}(0,1)^{n\times d}, q, d)$ for any $q \in \mathbb{N}$, and i_0 be defined as in Definition 1. Then*

$$R(\hat{\mathbf{u}}) = \Omega\left(\left(\frac{d}{d+i_0}\right)^{\frac{1}{2q+1}}\right)$$

holds with probability at least $1 - e^{-\Omega(d)}$.

If $\mathbf{A}$ has negative eigenvalues, we expect to have a guarantee of $R(\hat{\mathbf{u}})$ similar to that of Theorem 3 if the negative eigenvalues are not too large. We introduce the following technical assumption and generalize Theorem 3 to indefinite matrices. Let λ_i and $\mathbf{u}_i$ be the i-th largest eigenvalue and the corresponding eigenvector.

Theorem 4. *Assume there exists a constant $\kappa \in (0,1]$ such that $\sum_{i=2}^n \lambda_i^{2q+1} \geq \kappa \sum_{i=2}^n |\lambda_i|^{2q+1}$. Let $\hat{\mathbf{u}} = \texttt{RSVD}(\mathbf{A}, \mathcal{N}(0,1)^{n\times d}, q, d)$ for any $q \in \mathbb{N}$. Then, there is a constant $c_\kappa \in (0,1]$ that depends on κ such that*

$$R(\hat{\mathbf{u}}) = \Omega\left(c_\kappa\left(\frac{d}{d+i_0}\right)^{\frac{1}{2q+1}}\right),$$

with probability at least $1 - e^{-\Omega(\sqrt{d}\kappa^2)}$.

Moreover, we consider an extension of RSVD by replacing the random projection matrix $\mathbf{S} \sim \mathcal{N}(0,1)^{n \times d}$ with $\mathbf{S} = [\mathbf{S}_1, \mathbf{S}_2]$, where $\mathbf{S}_1 \sim \mathcal{N}(0,1)^{n \times \lceil \frac{d}{2} \rceil}$ and $\mathbf{S}_2 \sim \text{Bernoulli}(p)^{n \times \lfloor \frac{d}{2} \rfloor}$. This algorithm is called RandSum, and we show that it has the following guarantees:

Theorem 5. *Let* $\mathbf{A}$ *be a positive semindefinite matrix with* $\lambda_1 > 0$*, and* $\hat{\mathbf{u}} = \texttt{RandSum}(\mathbf{A}, q, d, p)$ *for any constant* $p \in (0,1)$ *and integer* $d \geq 2$*. Then,*

$$R(\hat{\mathbf{u}}) = \left(\Omega\left(\frac{\max\{d, \langle \mathbf{u}_1, \mathbf{1}_n \rangle^2\}}{n}\right)\right)^{\frac{1}{2q+1}}$$

holds with probability at least $1 - e^{-\Omega(d)}$.

Theorem 6. *Assume (i) there exists a constant* $\kappa \in (0,1]$ *such that* $\sum_{i=2}^{n} \lambda_i^{2q+1} \geq \kappa \sum_{i=2}^{n} |\lambda_i|^{2q+1}$*; (ii)* $\langle \mathbf{u}_1, \mathbf{1}_n \rangle^2 = \Omega(1)$*; (iii) there exists a constant* $\kappa' \in (0,1]$ *such that* $\sum_{i=2}^{n} \lambda_i^{2q+1} \xi_i \geq \kappa' \sum_{i=2}^{n} |\lambda_i|^{2q+1} \xi_i$*, where* $\xi_i = \mathbb{E}\left[\left\langle \mathbf{S}^T \mathbf{u}_i, \frac{\mathbf{1}_d}{\sqrt{d}} \right\rangle^2\right]$*, for all* $i \in [n]$*. Then,*

$$R(\hat{\mathbf{u}}) = \Omega\left(\left(\max\left\{\frac{d}{d+i_0}, \frac{\langle \mathbf{u}_1, \mathbf{1}_n \rangle^2}{n}\right\}\right)^{\frac{1}{2q+1}}\right)$$

holds with probability at least $1 - e^{-\Omega(\sqrt{d})}$.

While such a random matrix is rarely used in the literature of random projections, we show that there exist applications such as conflicting group detection [BGG+19, TOG20] that are especially suitable for this technique. Several properties of such a random matrix that we derived in the paper may be of independent interest.

Contributions. C.-J. Lu and the author of the thesis contributed to the problem formulation and the theoretical analysis. The initial manuscript was written by the author of the thesis, with P.-A. Wang, A. Gionis, and F. Adriaens contributing to subsequent revisions.

Chapter 4

Combinatorial multi-armed bandits

In this thesis, Paper C focuses on (ii) pure exploration (a) with fixed confidence, and Paper D focuses on (i) regret minimization. In all of these papers, we consider the stochastic environment and linear reward function and aim for an algorithm that is both statistically efficient and computationally efficient. In Paper C, we consider an open problem in combinatorial best arm identification with semi-bandit feedback: "Is it possible to design an algorithm that is both statistically optimal and runs in time polynomial in problem-specific parameters?". We answer the question affirmatively by providing an algorithm that achieves statistical optimality and computational efficiency. In Paper D, we consider matroid semi-bandits in the regret minimization setting and propose the first algorithm that runs in time sublinear in the number of arms for common classes of matroids.

4.1 Closing the Computational-Statistical Gap in Best Arm Identification for Combinatorial Semi-Bandits (NeurIPS'23)

Summary

In Paper C, we confirm the conjecture left by Jourda et al. [JMKK21] that there is no computational-statistical gap in best arm identification in stochastic combinatorial semi-bandits with fixed-confidence for uncorrelated Gaussian rewards. Stochastic combinatorial semi-bandits is an online learning problem where, at each round, the learner pulls an action that is a subset of arms satisfying certain combinatorial constraints (e.g., m-sets, spanning trees, and matchings) and observes noisy rewards for all the arms contained in the selected action. The objective of best arm identification with fixed confidence is to identify the best action with a given confidence level while using as few samples as possible. Our proposed method, `Perturbed-FWS`, addresses the computational inefficiency of the `FWS` method [WTP21] (see Section 2.10). Our main contribution is the proposal of a computationally efficient `MCP` algorithm for solving the inner optimization associated with the sample complexity lower bound problem [GK16]. For the outer optimization of the lower bound problem, we replace the smoothing technique of `FWS`, which constructs a space spanned by potentially $\mathcal{O}(2^K)$ gradient vectors (where K is the number of arms), with the standard stochastic smoothing technique [DBW12]. As a result, all `Perturbed-FWS` needs is a linear maximization oracle, making it the first optimal algorithm that runs in time polynomial in K.

Challenge

For combinatorial best arm identification with uncorrelated Gaussian rewards, we have the sample complexity lower bound $\mathbb{E}_{\boldsymbol{\mu}}[\tau] \geq T^\star(\boldsymbol{\mu})\mathsf{kl}(\delta, 1-\delta)$. We call $T^\star(\boldsymbol{\mu})^{-1}$ the *lower-bound problem* which is defined as

$$T^\star(\boldsymbol{\mu})^{-1} = \sup_{\boldsymbol{\omega}\in\Sigma} F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) \quad \text{with} \quad F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) = \inf_{\boldsymbol{\lambda}\in\mathsf{Alt}(\boldsymbol{\mu})} \left\langle \boldsymbol{\omega}, \frac{(\boldsymbol{\mu}-\boldsymbol{\lambda})^2}{2} \right\rangle,$$

where $\Sigma = \{\sum_{\boldsymbol{x}\in\mathcal{X}} w_{\boldsymbol{x}}\boldsymbol{x} : \boldsymbol{w} \in \Sigma_{|\mathcal{X}|}\}$, $\mathsf{kl}(a, b)$ is the KL-divergence between two Bernoulli distributions with respective means a and b, and $\mathsf{Alt}(\boldsymbol{\mu}) = \{\boldsymbol{\lambda} \in \Lambda : \boldsymbol{i}^\star(\boldsymbol{\lambda}) \neq \boldsymbol{i}^\star(\boldsymbol{\mu})\}$ is the set of *confusing* parameters, and $\boldsymbol{i}^\star(\cdot) \in \operatorname{argmax}_{\boldsymbol{x}\in\mathcal{X}} \langle \boldsymbol{x}, \cdot\rangle$ is a linear maximization (`LM`) oracle. For pure exploration in structured bandits, the existing statistical optimal algorithms [WTP21, JMKK21] which are designed based on solving $T^\star(\boldsymbol{\mu})^{-1}$ with $\boldsymbol{\mu}$ plugged-in with the empirical estimated mean $\hat{\boldsymbol{\mu}}(t-1)$. For the setting we consider, the computational challenge comes from $F_{\boldsymbol{\mu}}(\boldsymbol{\omega})$ because a naïve way of solving $F_{\boldsymbol{\mu}} = \min_{\boldsymbol{x}\neq\boldsymbol{i}^\star(\boldsymbol{\mu})} f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{\mu})$ has to solve $|\mathcal{X}| - 1$ many convex programs $f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{\mu}) = \inf_{\boldsymbol{\lambda}:\langle \boldsymbol{x}-\boldsymbol{i}^\star(\boldsymbol{\mu}), \boldsymbol{\lambda}\rangle>0} \left\langle \boldsymbol{\omega}, \frac{(\boldsymbol{\mu}-\boldsymbol{\lambda})^2}{2} \right\rangle$, and $|\mathcal{X}|$ might be exponential in the number K of arms.

Our approach

We address this computational challenge by proposing an efficient no-regret algorithm, referred to as `MCP` algorithm, for computing $F_{\boldsymbol{\mu}}(\boldsymbol{\omega})$. The design of the `MCP` algorithm is based on the observation that the Lagrangian dual function $g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha)$ of $f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{\mu})$ is linear in the action $\boldsymbol{x}$ and concave in the Lagrangian multiplier α (shown in Proposition 1), and hence we rewrite

$$F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) = \min_{\boldsymbol{x}\neq\boldsymbol{i}^\star(\boldsymbol{\mu})} \max_{\alpha\geq 0} g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha).$$

Please refer to Section 2.3 for an introduction to the Lagrangian multiplier method.

Assumption 1. *(i) There exists a polynomial-time algorithm identifying $\boldsymbol{i}^\star(\boldsymbol{v})$ for any $\boldsymbol{v} \in \mathbb{R}^K$; (ii) $\mathcal{X}$ is inclusion-wise maximal, i.e., there is no $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{X}$ s.t. $\boldsymbol{x} < \boldsymbol{x}'$; (iii) for each $k \in [K]$, there exists $\boldsymbol{x} \in \mathcal{X}$ such that $x_k = 1$; (iv) $|\mathcal{X}| \geq 2$.*

Proposition 1. *Let $(\boldsymbol{\omega}, \boldsymbol{\mu}) \in \Sigma_+ \times \Lambda$ and $\boldsymbol{x} \in \mathcal{X} \setminus \{\boldsymbol{i}^\star(\boldsymbol{\mu})\}$.*
(a) The Lagrange dual function is linear in $\boldsymbol{x}$. More precisely,

$$g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha) = c_{\boldsymbol{\omega},\boldsymbol{\mu}}(\alpha) + \langle \boldsymbol{\ell}_{\boldsymbol{\omega},\boldsymbol{\mu}}(\alpha), \boldsymbol{x}\rangle,$$

where $c_{\boldsymbol{\omega},\boldsymbol{\mu}}(\alpha) = \alpha \left\langle \boldsymbol{\mu} - \frac{\alpha}{2}\boldsymbol{\omega}^{-1}, \boldsymbol{i}^\star(\boldsymbol{\mu}) \right\rangle$ *and* $\boldsymbol{\ell}_{\boldsymbol{\omega},\boldsymbol{\mu}}(\alpha) = -\alpha \left(\boldsymbol{\mu} + \frac{\alpha}{2}\boldsymbol{\omega}^{-1} \odot (\mathbf{1}_K - 2\boldsymbol{i}^\star(\boldsymbol{\mu}))\right)$.
(b) $g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \cdot)$ *is strictly concave (for any fixed* $\boldsymbol{x}$*).*
(c) $f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{\mu}) = \max_{\alpha \geq 0} g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha)$ *is attained by* $\alpha^\star_{\boldsymbol{x}} = \frac{\triangle_{\boldsymbol{x}}(\boldsymbol{\mu})}{\langle \boldsymbol{x} \oplus \boldsymbol{i}^\star(\boldsymbol{\mu}), \boldsymbol{\omega}^{-1}\rangle}$.
(d) $\|\boldsymbol{\ell}_{\boldsymbol{\omega},\boldsymbol{\mu}}(\alpha^\star_{\boldsymbol{x}})\|_1 \leq L_{\boldsymbol{\omega},\boldsymbol{\mu}} = 4D^2K \|\boldsymbol{\mu}\|_\infty^2 \left\|\boldsymbol{\omega}^{-1}\right\|_\infty$.

But, different from the literature of the two-player game, we not only want to estimate $F_{\boldsymbol{\mu}}$ but also want the *equilibrium action* $\boldsymbol{x}_e$ such that $F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) = f_{\boldsymbol{x}_e}(\boldsymbol{\omega}, \boldsymbol{\mu})$. We need $\boldsymbol{x}_e$ because we would like to solve the outer optimization $\sup_{\boldsymbol{\omega} \in \Sigma} F_{\boldsymbol{\mu}}(\boldsymbol{\omega})$ by first-order methods and $\boldsymbol{x}_e$ is required to compute the subgradients. Hence, we design `MCP` algorithm from scratch shown in Algorithm 1:

Algorithm 1: (ϵ, θ)-`MCP`$(\boldsymbol{\omega}, \boldsymbol{\mu})$

```
initialization: n = 1, F̂ = ∞,
  c_θ = L_{ω,μ} ( 4 sqrt(K(ln K + 1)) + sqrt(ln(θ^{-1})/2) );
while (n = 1) or (n > 1 and sqrt(n) ≤ c_θ(1 + ε)/(ε F̂)) do
    Sample Z_n ~ exp(1)^K and set η_n = sqrt( K(ln K + 1)/(4n L²_{ω,μ}) );
    x^(n) ← argmin_{x ≠ i*(μ)} ( Σ_{m=1}^{n-1} g_{ω,μ}(x, α^(m)) + ⟨Z_n, x⟩ / η_n );
    α^(n) ← argmax_{α ≥ 0} g_{ω,μ}(x^(n), α);
    if g_{ω,μ}(x^(n), α^(n)) < F̂ then (F̂, x̂) ← (g_{ω,μ}(x^(n), α^(n)), x^(n)) ;
    n ← n + 1;
end
return (F̂, x̂);
```

To be able to estimate $F_{\boldsymbol{\mu}}$ and the equilibrium action $\boldsymbol{x}_e$ simultaneously, in our `MCP` algorithm, we let the $\boldsymbol{x}$-player use follow-the-perturbed-leader (`FTPL`) rule, and let the α-player use the best-response rule. The stopping criterion of `MCP` algorithm is designed such that when the criterion is met, the estimated $(\hat{F}, \hat{\boldsymbol{x}})$ is a good approximation to $(F_{\boldsymbol{\mu}}(\boldsymbol{\omega}), \boldsymbol{x}_e)$ with high probability. Our result of `MCP` is summarized as follows:

Theorem 1. *Let* $\epsilon, \theta \in (0, 1)$*. Under Assumption 1, for any* $(\boldsymbol{\omega}, \boldsymbol{\mu}) \in \Sigma_+ \times \Lambda$*, the* (ϵ, θ)*-`MCP`*$(\boldsymbol{\omega}, \boldsymbol{\mu})$ *algorithm outputs* $(\hat{F}, \hat{\boldsymbol{x}})$ *satisfying*

$$\mathbb{P}\Big[F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) \leq \hat{F} \leq (1+\epsilon) F_{\boldsymbol{\mu}}(\boldsymbol{\omega})\Big] \geq 1 - \theta \quad \textit{and} \quad \hat{F} = \max_{\alpha \geq 0} g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\hat{\boldsymbol{x}}, \alpha).$$

Moreover, the number of `LM` Oracle calls the algorithm does is almost surely at most

$$\left\lceil \frac{c_\theta^2 (1+\epsilon)^2}{\epsilon^2 F_{\boldsymbol{\mu}}(\boldsymbol{\omega})^2} \right\rceil = \mathcal{O}\left(\frac{\|\boldsymbol{\mu}\|_\infty^4 \left\|\boldsymbol{\omega}^{-1}\right\|_\infty^2 K^3 D^5 \ln K \ln \theta^{-1}}{\epsilon^2 F_{\boldsymbol{\mu}}(\boldsymbol{\omega})^2} \right).$$

Based on `MCP` algorithm, we design `Perturbed-FWS` which is the first statistically optimal and computationally efficient algorithm for combinatorial best arm identification with fixed-confidence and semi-bandit feedback for uncorrelated Gaussian rewards. The sample complexity of `Perturbed-FWS` is instance-specifically optimal in the high-confidence regime and has a polynomial dependency in K in the moderate-confidence regime.

Theorem 2. *For any $\delta \in (0,1)$, `P-FWS` is δ-PAC, and for any $(\epsilon, \tilde{\epsilon}) \in (0,1)$ small enough, its sample complexity satisfies:*

$$\mathbb{E}_{\boldsymbol{\mu}}[\tau] \leq \frac{(1+\tilde{\epsilon})^2}{T^\star(\boldsymbol{\mu})^{-1} - \epsilon} \times H\left(\frac{1}{\delta} \cdot \frac{c(1+\tilde{\epsilon})^2}{T^\star(\boldsymbol{\mu})^{-1} - \epsilon}\right) + \Psi(\epsilon, \tilde{\epsilon}),$$

where $H(x) = \ln(x) + \ln\ln(x)$, $c > 0$ is a universal constant, and $\Psi(\epsilon, \tilde{\epsilon})$ is polynomial in ϵ^{-1}, $\tilde{\epsilon}^{-1}$, K, $\|\boldsymbol{\mu}\|_\infty$, and $\triangle_{\min}^{-1}$, where $\triangle_{\min} = \min_{\boldsymbol{x} \neq \boldsymbol{i}^\star(\boldsymbol{\mu})} \langle \boldsymbol{i}^\star(\boldsymbol{\mu}) - \boldsymbol{x}, \boldsymbol{\mu} \rangle$. Under `P-FWS`, the number of `LM` Oracle calls per round is at most polynomial in $\ln \delta^{-1}$ and K. The total expected number of these calls is also polynomial.

Contributions. P.-A. Wang contributed to the problem formulation. C.-J. Lu, P. Wang, and the author of the thesis contributed to the algorithm design and theoretical analysis. The initial manuscript was written by the author of the thesis, with A. Proutiere, C.-J. Lu, and P.-A. Wang contributed to subsequent revisions.

4.2 Matroid Semi-Bandits in Sublinear Time (ICML'24)

Summary

In Paper D, we study matroid semi-bandits: given $[K] = \{1, \cdots, K\}$ arms with unknown mean $\boldsymbol{\mu} \in (0,1)^K$ and the set $\mathcal{X}$ of the bases of a given matroid, the goal is to identify the best action $\boldsymbol{i}^\star(\boldsymbol{\mu}) = \operatorname{argmax}_{\boldsymbol{x} \in \mathcal{X}} \langle \boldsymbol{x}, \boldsymbol{\mu} \rangle$ while minimizing the expected cumulative regret. Existing works such as `CUCB` [CWY13] and `KL-OSM` [TP16] have a per-round time complexity of at least $\Omega(K)$ which is inefficient when K is large. To address this issue, we propose `FasterCUCB` whose per-round time complexity is sublinear in K for common classes of matroids: $\mathcal{O}(D \,\mathsf{polylog}\,(K,T))$ for uniform matroid, partition matroid, and graphical matroid, and $\mathcal{O}(D\sqrt{K}\,\mathsf{polylog}\,(T))$ for transversal matroid, where D is the rank of the matroid and T is the horizon. Our technique is based on dynamic maintenance of an approximate maximum-weight basis over inner-product weights. Although the introduction of an approximate maximum-weight basis presents a challenge in regret analysis, we can still guarantee an upper bound on regret as tight as `CUCB` in the sense that it matches the gap-dependent lower bound by Kveton et al. [KWA+14] asymptotically.

Challenge

There is no known prior work whose per-round time complexity is sublinear in K. Also, the challenge to speed up `CUCB`, whose sampling rule is shown in (4.1), is that every arm's weight changes at each round.

$$\boldsymbol{x}(t) \in \operatorname*{argmax}_{\boldsymbol{x} \in \mathcal{X}} \langle \boldsymbol{f}_k, \boldsymbol{q} \rangle \, x_k, \tag{4.1}$$

where $\boldsymbol{f} = (\hat{\mu}_k(t-1), \frac{1}{\sqrt{N_k(t-1)}})$ is a pair of the empirical mean estimate and the radius of the confidence interval, and $\boldsymbol{q} = (1, \lambda_t)$ is a pair of 1 and the parameter that controls the confidence interval.

Our approach

Observe that when $\boldsymbol{q}$ is fixed, then there are only D arms in (4.1) that will change after each round. In this way, we can use a dynamic algorithm that supports fast arm weight change and fast computation of the maximum-weight base. So, the key question is how can we implement as if $\boldsymbol{q}$ is fixed? Our idea is to construct a *minimum hitting set* $\mathcal{H}$ which has the property that: for any $\boldsymbol{q} \in \mathbb{R}_+^2$, there exists $\boldsymbol{h} \in \mathcal{H}$ such that

$$\langle \boldsymbol{f}_k, \boldsymbol{h} \rangle > \langle \boldsymbol{f}_\ell, \boldsymbol{h} \rangle \implies \langle \boldsymbol{f}_k, \boldsymbol{q} \rangle > \langle \boldsymbol{f}_\ell, \boldsymbol{q} \rangle$$

for any $k \neq \ell$. What this means is that when replacing $\langle \boldsymbol{f}_k, \boldsymbol{q} \rangle$ with $\langle \boldsymbol{f}_k, \boldsymbol{h} \rangle$, the greedy algorithm will output the same solution. However, this naive solution will require $|\mathcal{H}| = \mathcal{O}(K^2)$. We overcome this issue by *rounding*. Specifically, we put the K arms into $\mathsf{polylog}\,(T)$ bins and use one point called *dominating point* to present each bin. The minimum hitting set is constructed using those dominating points, and for each $\boldsymbol{h} \in \mathcal{H}$, we create a dynamic algorithm with arm's weight $\langle \mathrm{dom}(\boldsymbol{f}_k), \boldsymbol{h} \rangle$, where $\mathrm{dom}(\boldsymbol{f}_k)$ denotes the dominating point of $\boldsymbol{f}_k$. When an arm's feature has to change, we need to change the weights stored in all the dynamic algorithms of each $h \in \mathcal{H}$. So, the procedure of our `FasterCUCB` is that we compute (4.1) by first finding the $\boldsymbol{h} \in \mathcal{H}$ that preserves the pairwise ordering with respect to $\boldsymbol{q}$, and then calling its dynamic algorithm to output a maximum-weight base. Then, after receiving the semi-bandit feedbacks, we update the corresponding arms' weight in all the dynamic algorithms of each $\boldsymbol{h} \in \mathcal{H}$. The procedures are summarized as follows:

INITIALIZE: Given lower and upper bounds $[\alpha_{\text{lb}}, \alpha_{\text{ub}}]$ and $[\beta_{\text{lb}}, \beta_{\text{ub}}]$, K features $(\alpha_k, \beta_k)_{k \in [K]}$, a matroid $\mathcal{M} = ([K], \mathcal{I})$, a dynamic algorithm $\mathcal{A}$ for maximum-weight base maintenance, and a precision parameter ϵ, this procedure initializes the data structure used in the remaining two procedures.

FIND-BASE: Given a query $\boldsymbol{q}$, this procedure is supposed to return a $(1+\epsilon)$-approximate maximum-weight base of $\mathcal{M}$, where arm k's weight is defined as $\langle \boldsymbol{f}_k, \boldsymbol{q} \rangle$ for the up-to-date k's feature $\boldsymbol{f}_k$.

UPDATE-FEATURE: Given an arm k and a new feature $\boldsymbol{f}'_k$, this procedure reflects the change of arm k's feature on the data structure.

Theorem 1. *There exist implementations of* **INITIALIZE**, **FIND-BASE**, *and* **UPDATE-FEATURE** *such that the following are satisfied:* **FIND-BASE** *always returns a* $(1+\epsilon)$*-approximate maximum-weight base of a matroid* $\mathcal{M}$ *with arm k's weight defined as* $\langle \boldsymbol{f}_k, \boldsymbol{q} \rangle$ *for an up-to-date k's feature* $\boldsymbol{f}_k$ *and a query* $\boldsymbol{q}$*. Moreover,* **INITIALIZE** *runs* *in* $\mathcal{O}(K + \mathsf{poly}(W) \cdot \mathcal{T}_{\mathrm{init}}(\mathcal{A}; \frac{\epsilon}{3}))$ *time,* **FIND-BASE** *runs in* $\mathcal{O}(\mathsf{poly}(W) + D)$ *time, and* **UPDATE-FEATURE** *runs in* $\mathcal{O}(\mathsf{poly}(W) \cdot \mathcal{T}_{\mathrm{update}}(\mathcal{A}; \frac{\epsilon}{3}))$ *time, where*

$$W = \mathcal{O}\left(\epsilon^{-1} \cdot \log\left(\frac{\alpha_{ub}}{\alpha_{lb}} \cdot \frac{\beta_{ub}}{\beta_{lb}}\right)\right). \tag{4.2}$$

Since $W = \mathcal{O}\left(\log^m T \log\left(\frac{b}{a}\sqrt{T}\right)\right) = \mathcal{O}\left(\log^{m+1} T\right)$, the per-round time complexity of `FasterCUCB` is $\mathcal{O}\left(D\, \mathsf{polylog}(T)\ \mathcal{T}_{\mathrm{update}}\left(\mathcal{A}; \frac{\epsilon}{3}\right)\right)$. Here, we will set $\epsilon = \frac{1}{\log^m T}$ for the regret analysis.

Theorem 2. *Let* $\lambda_t = \sqrt{1.5(b-a)^2 \log t}$ *and* $m \in \mathbb{N}$. *Define* $T_0 \triangleq \max\{K, \exp((\frac{b}{\triangle_{\min}})^{\frac{1}{m}})\}$. *For* $T \in \mathbb{N}$, *the expected regret of* `FasterCUCB` *is upper bounded by*

$$R(T) \leq \sum_{k \notin \mathsf{supp}(\boldsymbol{i}^\star)} \left(\sum_{j=1}^{d_k} \triangle_{\overline{j},k} T_0 + \frac{12 \triangle_{\overline{d_k},k} (b-a)^2 \log T}{\left(\frac{\mu_{\overline{d_k}}}{1+\log^{-m} T} - \mu_k\right)^2} \right) + \sum_{k \notin \mathsf{supp}(\boldsymbol{i}^\star)} \sum_{j=1}^{d_k} \triangle_{\overline{j},k} \left(\frac{1}{T} + \frac{\pi^2}{6}\right) + DbT_0,$$

where $\{\overline{j}\}_{j=1}^{D}$ *be the permutation of* $\mathsf{supp}(\boldsymbol{i}^\star)$ *such that* $\mu_{\overline{1}} \geq \cdots \geq \mu_{\overline{D}}$, $\triangle_{j,k} \triangleq \mu_j - \mu_k$ *and* $d_k \triangleq \max\{j \in [D] : \triangle_{\overline{j},k} > 0\}$ *for* $j \in \mathsf{supp}(\boldsymbol{i}^\star)$ *and* $k \notin \mathsf{supp}(\boldsymbol{i}^\star)$, *and* $\triangle_{\min} \triangleq \min_{k \notin \mathsf{supp}(\boldsymbol{i}^\star)} \triangle_{\overline{d_k},k}$.

As a consequence of Theorem 2, setting $T \to \infty$ yields:

$$\lim_{T\to\infty} \frac{R(T)}{\log T} \leq \sum_{k \notin \mathsf{supp}(\boldsymbol{i}^\star)} \frac{12(b-a)^2}{\triangle_{\overline{d_k},k}} \leq \mathcal{O}\left(\frac{K-D}{\triangle_{\min}}\right),$$

which matches Theorem 4 in [KWA+14], $\liminf_{T\to\infty} \frac{R(T)}{\log T} = \Omega(\frac{K-D}{\triangle_{\min}})$, asymptotically up to a constant factor. Note that `FasterCUCB` is faster than `CUCB` when $\triangle_{\min} = \Omega(\frac{1}{\mathsf{polylog}(K)})$ and when $T = \mathsf{poly}(K)$. Also, similar to [CCG21b], our per-round time complexity also goes to infinity as $T \to \infty$, one way to address this issue is to use `CUCB` when the per-round time complexity of ours is larger than that of `CUCB`.

Contributions. The author of the thesis contributed to the problem formulation. N. Ohsaka contributed to the algorithm design. N. Ohsaka and the author of the thesis contributed to the theoretical analysis. The initial manuscript was written by N. Ohsaka and the author of the thesis, with K. Ariu contributing to subsequent revisions.

Chapter 5

Challenges and open problems

5.1 Some technical challenges

Challenges in performing tight analysis

Challenges in pass-efficient analysis and 0/1-Bernoulli random projection

In Paper B, we focus on eigensolvers with $\mathcal{O}(n \log n)$ space and q passes over a input matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$. Suppose the algorithm outputs $\hat{\mathbf{u}} \in \mathbb{S}^{n-1}$ in q passes. The performance is measured against the multiplicative ratio:

$$R(\hat{\mathbf{u}}) = \frac{\hat{\mathbf{u}}^T \mathbf{A} \hat{\mathbf{u}}}{\lambda_1},$$

where λ_1 is the largest eigenvalue of $\mathbf{A}$. Deriving a tight upper bound of $R(\hat{\mathbf{u}})$ satisfied by any algorithm using $\mathcal{O}(n \log n)$ space and q passes over $\mathbf{A}$ is an open problem. See Table 5.1.

Table 5.1: Multiplicative ratio $R(\hat{\mathbf{u}})$ for any algorithm.

$R(\hat{\mathbf{u}})$	lower bound	upper bound
Positive semidefnite matrices $\mathcal{O}(nd)$ space and q passes	$\Omega\left(\left(\frac{d}{n}\right)^{\frac{1}{2q+1}}\right)$ by Randomized SVD Paper B [TWA+22]	open

Furthermore, we proposed a variant of `RSVD` called `RandSum` which uses 0/1-Bernoulli with Gaussian distribution in the random projection step. It is challenging to derive a tight analysis of the 0/1-Bernoulli random projection due to the lack of appropriate tools in probability theory. This is because the 0/1-Bernoulli distribution is not uniformly distributed with respect to the Haar measure on the Grassmannian of all $\log n$-dimensional subspaces of $\mathbb{R}^n$, which is a crucial property satisfied by all Johnson-Lindenstrauss (JL) distributions, so the existing results on the JL distributions do not apply to the 0/1-Bernoulli random projection.

Challenges in designing statistically optimal and computationally efficient algorithms for combinatorial multi-armed bandit problem.

There generally exists computational-statistical gaps in combinatorial multi-armed bandits for the (a) regret minimization problem and for the (b) best arm identification problem with fixed confidence. See Table 5.2 for a summary of the current progress. We say a problem has a computational-statistical gap if there exist statistically optimal algorithms for the problem, but no computationally efficient implementations. In problems (a) and (b), a statistically optimal algorithms [CMP17, GK16, WTP21, JMKK21] is designed from the insights gained from the regret lower bound [CTMSP+15] or sample complexity lower bound [GK16]. A regret or sample complexity lower bound is characterized by an optimization problem, and a statistically optimal algorithm consists of a subprocedure for solving the lower bound optimization problem. However, in combinatorial multi-armed bandits, solving the lower bound problem is often computationally challenging. In the following, we list the lower bound problem for the (a) regret minimization and for the (b) best arm identification with fixed confidence.

Table 5.2: Computational-statistical gap in combinatorial semi-bandits.

	Reward Distr.	Comput. Efficient	Stat. Optimal
combinatorial BAI semi-bandit	Gaussian	Paper C [TWPL23]	[JMKK21, WTP21]
	Bernoulli	open	
combinatorial BAI full-bandit	Gaussian Bernoulli	open	[WTP21]
combinatorial regret minimization	Gaussian Bernoulli	open	[CMP17]

Challenges in solving the regret lower bound problem

The regret lower bound [CTMSP+15] is characterized by the Graves-Lai optimization problem:

$$\begin{aligned} &\min_{\boldsymbol{a}\in\mathbb{R}_+^{|\mathcal{X}|}} \sum_{\boldsymbol{x}\in\mathcal{X}} a_{\boldsymbol{x}} \langle \boldsymbol{i}^\star(\boldsymbol{\mu}) - \boldsymbol{x}, \boldsymbol{\mu}\rangle \\ &\text{subject to } \sum_{\boldsymbol{x}\in\mathcal{X}} a_{\boldsymbol{x}} \sum_{k\in\mathsf{supp}(\boldsymbol{x})} d(\mu_k,\lambda_k) \geq 1, \forall \boldsymbol{\lambda}\in\Lambda, \end{aligned} \tag{5.1}$$

where $\Lambda = \{\boldsymbol{\lambda}\in\mathbb{R}^K : |\boldsymbol{i}^\star(\boldsymbol{\lambda})| = 1, \boldsymbol{i}^\star(\boldsymbol{\lambda}) \neq \boldsymbol{i}^\star(\boldsymbol{\mu}), \lambda_k = \mu_k, \forall k \in \mathsf{supp}(\boldsymbol{i}^\star(\boldsymbol{\mu}))\}$, $\boldsymbol{i}^\star(\cdot) \in \operatorname{argmax}_{\boldsymbol{x}\in\mathcal{X}} \langle \cdot, \boldsymbol{x}\rangle$, and $d(a,b)$ is the KL divergence of the two distributions parameterized under a and b. `OSSB` [CMP17] is a statistically optimal algorithm that requires solving (5.1) each round with $\boldsymbol{\mu}$ replaced with the empirical mean estimates, but the authors did not propose an efficient procedure to solve (5.1). [CCG21a] aims to address this computational issue for `OSSB` by proposing a subprocedure called `GLPG` that performs proximal gradient descent for solving (5.1). `GLPG` requires a subprocedure for solving a budgeted linear maximization (`BLM`) problem:

$$\max_{\boldsymbol{x}\in\mathcal{X}} \langle \boldsymbol{a}, \boldsymbol{x}\rangle \text{ subject to } \langle \boldsymbol{b}, \boldsymbol{x}\rangle \geq s. \tag{5.2}$$

There exists an exact `BLM` algorithm for m-sets and source-destination paths, but for spanning trees and matchings, only approximate `BLM` algorithms exist. As a consequence, the approach by Cuvelier et al. [CCG21a] managed to maintain statistical optimality only for m-sets and source-destination paths, and under uncorrelated Gaussian reward distributions. How to design an algorithm that closes the computational-statistical gap for combinatorial regret minimization for common combinatorial structures (including m-sets, spanning trees, and matchings) is an open problem.

Challenges in solving the sample complexity lower bound problem

The sample complexity lower bound [JMKK21] for best arm identification with fixed confidence for combinatorial semi-bandits which can be derived by change-of-measure technique [KCG16] is characterized by $T^\star(\boldsymbol{\mu})$ defined as:

$$T^\star(\boldsymbol{\mu})^{-1} = \sup_{\boldsymbol{\omega}\in\Sigma} F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) \quad \text{with} \quad F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) = \inf_{\boldsymbol{\lambda}\in\mathsf{Alt}(\boldsymbol{\mu})} \sum_{k=1}^{K} \omega_k d(\mu_k, \lambda_k), \tag{2.7}$$

where $\Sigma = \{\sum_{\boldsymbol{x}\in\mathcal{X}} w_{\boldsymbol{x}} \boldsymbol{x} : \boldsymbol{w} \in \Sigma_{|\mathcal{X}|}\}$, and $\mathsf{Alt}(\boldsymbol{\mu}) = \{\boldsymbol{\lambda} \in \Lambda : \boldsymbol{i}^\star(\boldsymbol{\lambda}) \neq \boldsymbol{i}^\star(\boldsymbol{\mu})\}$.

There exist statistically optimal algorithms such as `FWS` [WTP21] and `CombGame` [JMKK21] that are designed based on solving (2.7) at each round with $\boldsymbol{\mu}$ and $\boldsymbol{\omega}$ replaced with the empirical estimates of the mean and the arm allocations. However, the authors of `FWS` and `CombGame` did not propose a computationally efficient procedure for solving (2.7). The challenge of solving (2.7) comes from the evaluation of $F_{\boldsymbol{\mu}}(\boldsymbol{\omega})$ because $\mathsf{Alt}(\boldsymbol{\mu})$ is a convex union of a potentially $\mathcal{O}\left(2^K\right)$ many convex set. In other words, one may need to solve $\mathcal{O}\left(2^K\right)$ many convex programs to evaluate $F_{\boldsymbol{\mu}}(\boldsymbol{\omega})$, which is not computationally efficient.

In Paper C, we proposed an efficient procedure for evaluating $F_{\boldsymbol{\mu}}(\boldsymbol{\omega})$ on common combinatorial structures for uncorrelated Gaussian rewards, and use such procedure to design a statistically optimal and computationally efficient algorithm called `P-FWS`. The efficiency of such a procedure for evaluating $F_{\boldsymbol{\mu}}$ relies on the property of KL divergence of the uncorrelated Gaussian rewards, and unfortunately does not generalize to other distributions. So, how to design an algorithm that closes the computational-statistical gap for combinatorial best arm identification with semi-bandit feedback and reward distributions other than uncorrelated Gaussian is an open problem.

Moreover, beyond semi-bandits, statistically optimal algorithms exist [WTP21] for best arm identification with full-bandit feedback and also for linear best arm identification, but none are computationally efficient. It remains an open problem to design a statistically optimal and computationally efficient algorithm for these problems.

5.2 Open problems left in the included papers

Discovering k-conflicting Groups in Signed Networks (NeurIPS'20)

Several questions are left by this work. (i) It remains open whether we can improve the $\mathcal{O}((k-j)\sqrt{n})$-approximation for (3.3) using an approach that does not rely on rounding the leading eigenvector, such as by extending the SDP-based algorithm in [BCMV12]; (ii) The modified Stochastic Block Model (`m-SSBM`) is a special case of Label Stochastic Block Model (`LSBM`) [HLM12]. It would be relevant to analyze the recovery guarantee of our proposed method in `m-SSBM` concerning the fundamental limit results [YP16] and the interplay with the Bethe-Hessian operator [SKZ14] in the sparse regime; (iii) The difference in the empirical performance of our two rounding techniques and the spectral clustering baseline `SPONGE` [CDGT19] in the real-world networks and the synthetic network is somewhat striking. Some properties or structures may exist in the real-world networks but not in the synthetic networks. An interesting question is to explain this behavior analytically, in particular concerning properties of real-world networks.

Improved Analysis of Randomized SVD for Top-Eigenvector Approximation (AISTATS'22)

Several questions are left by this work. (i) It is an open problem to characterize the fundamental limit of maximizing $R(\hat{\mathbf{u}})$ for any algorithm with fixed number of pass and $\mathcal{O}(n \log n)$ space; (ii) Our results may be extended in different ways. For example, we may relax the requirement on the input matrix from symmetric to stochastic, to analyze approximations of PageRank [PBMW99]. Or, we may extend `RandSum` to use any non-centered subgaussian distribution combined with JL distribution, and we conjecture this yields similar results. (iii) Another direction is to extend our analysis to top-k eigenvectors; since there are already several methods for computing top-k eigenvectors [HMT11, Mac08, AZL16], the most challenging part is to define the proper metric to maximize, as a generalization of $R(\hat{\mathbf{u}})$.

Closing the Computational-Statistical Gap in Best Arm Identification for Combinatorial Semi-Bandits (NeurIPS'23)

(i) In this work, we design an efficient algorithm for computing $F_{\boldsymbol{\mu}}$ by exploiting the property of the KL divergence of the Gaussian distribution. For Bernoulli distribution, whether one can design an efficient algorithm for computing $F_{\boldsymbol{\mu}}$ is an open problem; (ii) We have studied the computational-statistical trade-off through the analysis of the optimization problem leading to instance-specific sample complexity lower bounds. This approach can be extended to study the computational-statistical gap in other learning tasks, such as combinatorial bandits with bandit feedback [KHS20], linear bandits [DMSV20, JP20], and RL in linear or low-rank MDPs [AKKS20]). Most results on these problems are concerned with statistical efficiency and ignore computational issues.

Matroid Semi-Bandits in Sublinear Time (ICML'24)

There are many semi-bandit algorithms. In this work, we have accelerated the `CUCB` algorithm to match the gap-dependent regret lower bound. However, for the `KL-OSM` algorithm, which matches the instance-specific lower bound, it is currently unknown how to convert it into a sublinear-time algorithm. Similarly, for UCB-based algorithms such as `LinUCB` [LCLS10] in the full-bandit feedback, determining how to convert them into sublinear-time algorithms remains an open problem.

References

[AB09] Sanjeev Arora and Boaz Barak. *Computational complexity: a modern approach*. Cambridge University Press, 2009.

[ABF$^+$20] Aris Anagnostopoulos, Luca Becchetti, Adriano Fazzone, Cristina Menghini, and Chris Schwiegelshohn. Spectral relaxations and fair densest subgraphs. In *Proc. of CIKM*, 2020.

[AJO08] Peter Auer, Thomas Jaksch, and Ronald Ortner. Near-optimal regret bounds for reinforcement learning. *Proc. of NeurIPS*, 2008.

[AKKS20] Alekh Agarwal, Sham Kakade, Akshay Krishnamurthy, and Wen Sun. Flambe: Structural complexity and representation learning of low rank mdps. In *Proc. of NeurIPS*, 2020.

[AZL16] Zeyuan Allen-Zhu and Yuanzhi Li. Lazysvd: Even faster svd decomposition yet without agonizing pain. In *Proc. of NeurIPS*, 2016.

[BBC04] Nikhil Bansal, Avrim Blum, and Shuchi Chawla. Correlation clustering. *Machine learning*, 2004.

[BCMV12] Aditya Bhaskara, Moses Charikar, Rajsekar Manokaran, and Aravindan Vijayaraghavan. On quadratic programming with a ratio objective. In *Proc. of ICALP*, 2012.

[BDS22] Noa Ben-David and Sivan Sabato. A fast algorithm for pac combinatorial pure exploration. In *Proc. of AAAI*, 2022.

[BGG$^+$19] Francesco Bonchi, Edoardo Galimberti, Aristides Gionis, Bruno Ordozgoiti, and Giancarlo Ruffo. Discovering polarized communities in signed networks. In *Proc. of CIKM*, 2019.

[CCG21a] Thibaut Cuvelier, Richard Combes, and Eric Gourdin. Asymptotically optimal strategies for combinatorial semi-bandits in polynomial time. In *Proc. of ALT*, 2021.

[CCG21b] Thibaut Cuvelier, Richard Combes, and Eric Gourdin. Statistically efficient, polynomial-time algorithms for combinatorial semi-bandits. *Proc. of SIGMETRICS*, 2021.

[CDGT19] Mihai Cucuringu, Peter Davies, Aldo Glielmo, and Hemant Tyagi. Sponge: A generalized eigenproblem for clustering signed networks. In *Proc. of AISTATS*, 2019.

[CF06] Deepayan Chakrabarti and Christos Faloutsos. Graph mining: Laws, generators, and algorithms. *ACM computing surveys (CSUR)*, 2006.

[CLK$^+$14] Shouyuan Chen, Tian Lin, Irwin King, Michael R Lyu, and Wei Chen. Combinatorial pure exploration of multi-armed bandits. In *Proc. of NeurIPS*, 2014.

[CLTL17] Kewei Cheng, Jundong Li, Jiliang Tang, and Huan Liu. Unsupervised sentiment analysis with signed social networks. In *Proc. of AAAI*, 2017.

[CMP17] Richard Combes, Stefan Magureanu, and Alexandre Proutiere. Minimal exploration in structured stochastic bandits. In *Proc. of NeurIPS*, 2017.

[CTMSP$^+$15] Richard Combes, Mohammad Sadegh Talebi Mazraeh Shahi, Alexandre Proutiere, et al. Combinatorial bandits revisited. In *Proc. of NeurIPS*, 2015.

[CWP$^+$16] Lingyang Chu, Zhefeng Wang, Jian Pei, Jiannan Wang, Zijin Zhao, and Enhong Chen. Finding gangs in war from signed networks. In *Proc. of SIGKDD*, 2016.

[CWY13] Wei Chen, Yajun Wang, and Yang Yuan. Combinatorial multi-armed bandit: General framework and applications. In *Proc. of ICML*, 2013.

[Dan19] Daniel A. Spielman. Spectral Graph Theory. `https://www.cs.yale.edu/homes/spielman/462/462schedule.html`, 2019. Online; accessed 31 October 2023.

[DBW12] John C Duchi, Peter L Bartlett, and Martin J Wainwright. Randomized smoothing for stochastic optimization. *SIAM Journal on Optimization*, 2012.

[DG03] Sanjoy Dasgupta and Anupam Gupta. An elementary proof of a theorem of johnson and lindenstrauss. *Random Structures & Algorithms*, 2003.

[DK19] Rémy Degenne and Wouter M Koolen. Pure exploration with multiple correct answers. In *Proc. of NeurIPS*, 2019.

[DKC21] Yihan Du, Yuko Kuroki, and Wei Chen. Combinatorial pure exploration with full-bandit or partial linear feedback. In *Proc. of AAAI*, 2021.

[DKM19] Rémy Degenne, Wouter M Koolen, and Pierre Ménard. Non-asymptotic pure exploration by solving games. In *Proc. of NeurIPS*, 2019.

[DMSV20] Rémy Degenne, Pierre Ménard, Xuedong Shang, and Michal Valko. Gamification of pure exploration for linear bandits. In *Proc. of ICML*, 2020.

[EG17] Nicole Eikmeier and David F Gleich. Revisiting power-law distributions in spectra of real world networks. In *Proc. of SIGKDD*, 2017.

[Fei02] Uriel Feige. Relations between average case complexity and approximation complexity. In *Proc. of STOC*, 2002.

[FW56] Marguerite Frank and Philip Wolfe. An algorithm for quadratic programming. *Naval Research Logistics Quarterly*, 1956.

[GGM08] Geoffrey J Gordon, Amy Greenwald, and Casey Marks. No-regret learning in convex games. In *Proc. of ICML*, 2008.

[GH16] Dan Garber and Elad Hazan. A linearly convergent variant of the conditional gradient algorithm under strong convexity, with applications to online and stochastic optimization. *SIAM Journal on Optimization*, 2016.

[GK16] Aurélien Garivier and Emilie Kaufmann. Optimal best arm identification with fixed confidence. In *Proc. of COLT*, 2016.

[GK21] Aurélien Garivier and Emilie Kaufmann. Nonasymptotic sequential tests for overlapping hypotheses applied to near-optimal arm identification in bandit models. *Sequential Analysis*, 2021.

[GKJ12] Yi Gai, Bhaskar Krishnamachari, and Rahul Jain. Combinatorial network optimization with unknown variables: Multi-armed bandits with linear rewards and individual observations. *IEEE/ACM Transactions on Networking*, 2012.

[HK12] Elad Hazan and Satyen Kale. Projection-free online learning. In *Proc. of ICML*, 2012.

[HLM12] Simon Heimlicher, Marc Lelarge, and Laurent Massoulié. Community detection in the labelled stochastic block model. *Proc. of NIPS Workshop*, 2012.

[HMT11] Nathan Halko, Per-Gunnar Martinsson, and Joel A Tropp. Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions. *SIAM review*, 2011.

[HP14] Moritz Hardt and Eric Price. The noisy power method: a meta algorithm with applications. In *Proc. of NeurIPS*, 2014.

[Jag13] Martin Jaggi. Revisiting frank-wolfe: Projection-free sparse convex optimization. In *Proc. of ICML*, 2013.

[JMKK21] Marc Jourdan, Mojmír Mutnỳ, Johannes Kirschner, and Andreas Krause. Efficient pure exploration for combinatorial bandits with semi-bandit feedback. In *Proc. of ALT*, 2021.

[Jol86] Ian T Jolliffe. Principal components in regression analysis. In *Principal component analysis*. Springer, 1986.

[JP20] Yassir Jedra and Alexandre Proutiere. Optimal best-arm identification in linear bandits. In *Proc. of NeurIPS*, 2020.

[KCG16] Emilie Kaufmann, Olivier Cappé, and Aurélien Garivier. On the complexity of best-arm identification in multi-armed bandit models. *JMLR*, 2016.

[KHS20] Yuko Kuroki, Junya Honda, and Masashi Sugiyama. Combinatorial pure exploration with full-bandit feedback and beyond: Solving combinatorial optimization under uncertainty with limited observation. *arXiv preprint arXiv:2012.15584*, 2020.

[KWA+14] Branislav Kveton, Zheng Wen, Azin Ashkan, Hoda Eydgahi, and Brian Eriksson. Matroid bandits: fast combinatorial optimization with learning. In *Proc. of UAI*, 2014.

[KWAS15] Branislav Kveton, Zheng Wen, Azin Ashkan, and Csaba Szepesvari. Tight regret bounds for stochastic combinatorial semi-bandits. In *Proc. of AISTATS*, 2015.

[LCLS10] Lihong Li, Wei Chu, John Langford, and Robert E Schapire. A contextual-bandit approach to personalized news article recommendation. In *Proc. of The Web Conference*, 2010.

[LGC16] Andrea Locatelli, Maurilio Gutzeit, and Alexandra Carpentier. An optimal algorithm for the thresholding bandit problem. In *Proc. of ICML*, 2016.

[LHK10] Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. Predicting positive and negative links in online social networks. In *Proc. of The Web Conference*, 2010.

[LMFB23] Tommaso Lanciano, Atsushi Miyauchi, Adriano Fazzone, and Francesco Bonchi. A survey on the densest subgraph problem and its variants. *arXiv preprint arXiv:2303.14467*, 2023.

[LS20] Tor Lattimore and Csaba Szepesvári. *Bandit algorithms*. Cambridge University Press, 2020.

[Mac08] Lester Mackey. Deflation methods for sparse pca. In *Proc. of NeurIPS*, 2008.

[MM15] Cameron Musco and Christopher Musco. Randomized block krylov methods for stronger and faster approximate singular value decomposition. In *Proc. of NeurIPS*, 2015.

[MS02] Paul Milgrom and Ilya Segal. Envelope theorems for arbitrary choice sets. *Econometrica*, 2002.

[New06] Mark EJ Newman. Modularity and community structure in networks. *Proc. of NAS*, 2006.

[PBMW99] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: Bringing order to the web. Technical report, Stanford InfoLab, 1999.

[Per22] Pierre Perrault. When combinatorial thompson sampling meets approximation regret. In *Proc. of NeurIPS*, 2022.

[RST10] Alexander Rakhlin, Karthik Sridharan, and Ambuj Tewari. Online learning: Random averages, combinatorial parameters, and learnability. *Proc. of NeurIPS*, 2010.

[SEAR18] Max Simchowitz, Ahmed El Alaoui, and Benjamin Recht. Tight query complexity lower bounds for pca via finite sample deformed wigner law. In *Proc. of STOC*, 2018.

[Sha21] Xuedong Shang. Linbai: Gamification of pure exploration for linear bandits. `https://github.com/xuedong/LinBAI.jl`, 2021. [Online; accessed 09-May-2021].

[SJ14] Moshen Shahriari and Mahdi Jalili. Ranking nodes in signed social networks. *Social network analysis and mining*, 2014.

[SKZ14] Alaa Saade, Florent Krzakala, and Lenka Zdeborová. Spectral clustering of graphs with the bethe hessian. In *Proc. of NeurIPS*, 2014.

[TOA24] Ruo-Chun Tzeng, Naoto Ohsaka, and Kaito Ariu. Matroid semi-bandits in sublinear time. In *Proc. of ICML*, 2024.

[TOG20] Ruo-Chun Tzeng, Bruno Ordozgoiti, and Aristides Gionis. Discovering conflicting groups in signed networks. In *Proc. of NeurIPS*, 2020.

[TP16] Mohammad Sadegh Talebi and Alexandre Proutiere. An optimal algorithm for stochastic matroid bandit optimization. In *Proc. of AAMAS*, 2016.

[TWA$^+$22] Ruo-Chun Tzeng, Po-An Wang, Florian Adriaens, Aristides Gionis, and Chi-Jen Lu. Improved analysis of randomized svd for top-eigenvector approximation. In *Proc. of AISTATS*, 2022.

[TWPL23] Ruo-Chun Tzeng, Po-An Wang, Alexandre Proutiere, and Chi-Jen Lu. Closing the computational-statistical gap in best arm identification for combinatorial semi-bandits. In *Proc. of NeurIPS*, 2023.

[Vem05] Santosh S Vempala. *The random projection method*. American Mathematical Soc., 2005.

[Ver18] Roman Vershynin. *High-dimensional probability: An introduction with applications in data science*. Cambridge university press, 2018.

[Vis21] Nisheeth K. Vishnoi. *Algorithms for Convex Optimization*. Cambridge University Press, 2021.

[WSLX17] Jing Wang, Jie Shen, Ping Li, and Huan Xu. Online matrix completion for signed link prediction. In *Proc. of WSDM*, 2017.

[WTP21] Po-An Wang, Ruo-Chun Tzeng, and Alexandre Proutiere. Fast pure exploration via frank-wolfe. In *Proc. of NeurIPS*, 2021.

[WTP23] Po-An Wang, Ruo-Chun Tzeng, and Alexandre Proutiere. Best arm identification with fixed budget: A large deviation perspective. In *Proc. of NeurIPS*, 2023.

[YP16] Se-Young Yun and Alexandre Proutiere. Optimal cluster recovery in the labeled stochastic block model. In *Proc. of NeurIPS*, 2016.

[YWS15] Yi Yu, Tengyao Wang, and Richard J Samworth. A useful variant of the davis–kahan theorem for statisticians. *Biometrika*, 2015.

Appendix A

Discovering k-conflicting groups in signed networks

Figure: Two groups of cats fighting each other.

We study the problem of k-conflicting group detection in signed networks, where conflicting groups are node subsets such that inter-group edges are mostly negative while intra-group edges are mostly positive. We derive a formulation of the problem such that each conflicting group is naturally characterized by the maximum discrete Rayleigh's quotient (`MAX-DRQ`) problem, and present an eigenvector-based algorithm with provable guarantee to the `MAX-DRQ` problem.

Discovering conflicting groups in signed networks

Ruo-Chun Tzeng
KTH Royal Institute of Technology
rctzeng@kth.se

Bruno Ordozgoiti
Aalto University
bruno.ordozgoiti@aalto.fi

Aristides Gionis
KTH Royal Institute of Technology
argioni@kth.se

Abstract

Signed networks are graphs where edges are annotated with a positive or negative sign, indicating whether an edge interaction is friendly or antagonistic. Signed networks can be used to study a variety of social phenomena, such as mining polarized discussions in social media, or modeling relations of trust and distrust in online review platforms. In this paper we study the problem of detecting k conflicting groups in a signed network. Our premise is that each group is positively connected internally and negatively connected with the other $k-1$ groups. A distinguishing aspect of our formulation is that we are not searching for a complete partition of the signed network; instead, we allow a subset of nodes to be neutral with respect to the conflict structure we are searching. As a result, the problem we tackle differs from previously-studied problems, such as correlation clustering and k-way partitioning. To solve the conflicting-group discovery problem, we derive a novel formulation in which each conflicting group is naturally characterized by the solution to the maximum discrete Rayleigh's quotient (MAX-DRQ) problem. We present two spectral methods for finding approximate solutions to the MAX-DRQ problem, which we analyze theoretically. Our experimental evaluation shows that, compared to state-of-the-art baselines, our methods find solutions of higher quality, are faster, and recover ground-truth conflicting groups with higher accuracy.

1 Introduction

Signed networks are graphs where each edge is labeled either positive or negative. The introduction of edge signs, which goes back to the 50's, was motivated by the study of friendly and antagonistic social relationships [22]. The representation power of signed networks comes at the cost of significant differences in fundamental graph properties, and thus, algorithmic techniques employed to analyze unsigned networks are usually not directly applicable to their signed counterparts. These differences have spurred significant interest in a variety of analysis tasks in signed networks [20, 36] such as signed network embeddings [7, 24, 25, 39], signed clustering [8, 14, 27, 32], and signed link prediction [9, 28, 38, 41] in recent years.

In this paper we study the problem of detecting k conflicting groups in signed networks. In more detail, we are interested in finding a collection of k vertex subsets, each of which is positively connected internally, and negatively connected to the other $k-1$ subsets. In social networks where edge signs indicate positive or negative interactions, identifying conflicting groups may help in the study of polarization [1, 31, 40, 43], echo chambers [17, 19] and the spread of fake news [12, 35, 42].

Detecting k conflicting groups is challenging due to various reasons. First, conflicting groups are not simply dense subgraphs, so community-detection techniques for unsigned graphs are not effective. Second, in real applications we can expect a majority of the network nodes to be neutral with respect

34th Conference on Neural Information Processing Systems (NeurIPS 2020), Vancouver, Canada.

to the conflicting structure. As an example, consider a social network where a heated discussion is taking place between different political factions. Most users might not get involved in the quarrel, and thus their interactions are not necessarily consistent with this division. For this reason, methods for signed networks like correlation clustering and k-way partitioning may not be effective.

Our approach for detecting k-conflicting groups in signed networks extends the formulation of Bonchi et al. [4], to arbitrary values of k, addressing an open problem left in that work, which studied only the case $k = 2$. We argue that simply rounding the principal eigenvectors of the adjacency matrix might yield unsatisfactory results. Instead, we show that the proposed objective can be interpreted in terms of the Laplacian of a complete graph, and rely on the spectral properties of this matrix to derive a novel optimization framework, *spectral conflicting group detection* (SCG). By carefully examining the invariant subspaces of the aforementioned Laplacian, we reformulate the problem as a maximum discrete Rayleigh quotient (MAX-DRQ) objective, which is an $\mathbf{APX}$-hard problem. We propose two algorithms, one deterministic, and one randomized with approximation guarantees. We show that the obtained approximation is essentially the best possible, when using the largest eigenvalue as an upper bound.

We perform an extensive set of experiments to compare the performance of our approach to that of multiple alternatives from the literature, on a variety of synthetic and real datasets. Our algorithms generally run faster, yield solutions of higher quality, and exhibit a better ability to find ground-truth groups than competing methods. In addition, we discuss how to select the number of groups k in practical scenarios.

2 Related work

Signed graph partition. Typical formulations partition the nodes of the signed graph into k sets so that intra-edges are mostly positive and inter-edges are mostly negative. This is a special case of k conflicting group detection with no neutral nodes. Spectral methods are competitive and we review several representatives here. The signed Laplacian has been used for clustering [27], but resulting clusters tend to behave like in unsigned spectral clustering [37]. k-way balanced normalized cut (BNC) was proposed to address the issue [8]. Signed Laplacians [8, 27] were recently generalized through matrix power means [32]. The state of the art method SPONGE [14] is based on a generalized eigenvalue problem for constrained clustering [13] and works well on sparse graphs and large k. All these methods partition the network and are ineffective in the presence of many neutral nodes.

Correlation clustering methods partition the entire network, but allow k to be unspecified. The standard objective [2, 5, 16, 6] counts the number of edges that agree (disagree) with the partition, i.e., positive (negative) intra-group edges and negative (positive) inter-group edges, and aims to maximize (minimize) agreement (disagreement). The problem is $\mathbf{APX}$-hard for general graphs and has many variants. Giotis et al. [21] consider the case of fixed k and Puleo et al. [33] measure per-node error. Our work is inspired by a recent variant [4], which formulates the discrete eigenvector problem by maximizing the gap between agreement and disagreement with respect to the total size of two conflicting groups. They propose a randomized $\mathcal{O}(\sqrt{n})$-approximation algorithm. However, their approach does not extend to $k > 2$, as the two groups are identified by the sign of the optimal vector. In fact, the discrete eigenvector problem is $\mathbf{APX}$-hard and the best known result achieves an approximation guarantee of $\tilde{\mathcal{O}}(n^{1/3})$ using an SDP-based approach [3]. The latter SDP formulation cannot be extended to $k > 2$ as well. In this paper, we generalize the problem as MAX-DRQ and present two algorithms for $k \geq 2$.

Antagonistic group mining focuses on the setting with two groups. These works can be divided into direct or indirect. Direct methods [18, 29, 30] search for structures such as bi-cliques or balanced triads. Indirect methods [45, 46] find frequent conflicting patterns in database transactions. These approaches cannot be easily extended to finding $k > 2$ conflicting groups.

To our knowledge, our only direct competitor is the KOCG method [11]. They formulate the problem as trace-maximization, where each group is represented as a simplex with nonzero entries indicating the participation of the nodes in the groups. However, their method finds conflicting groups only within local regions and is sensitive to initialization, often converging to local maxima. Our approach is fundamentally different and experimentally is shown to consistently outperform this baseline.

3 Preliminaries

We focus on simple undirected signed graphs. We denote $G = (V, E)$ to be a signed graph, with $E = E_+ \cup E_-$ consisting of the sets of positive edges E_+ and negative edges E_-. The signed adjacency matrix of G is denoted by $A \in \{-1, 0, 1\}^{n\times n}$ with $A_{i,j}$ being $+1$ if $(i, j) \in E_+$, -1 if $(i, j) \in E_-$ and 0 otherwise. We use $n = |V|$ and $m = |E|$ to indicate the number of nodes and edges of the signed graph G. We use $E(V_1, V_2)$ to denote the set of edges between two subsets $V_1, V_2 \subseteq V$, where V_1, V_2 are not required to be disjoint. We define $E(V_1)$ to be $E(V_1, V_1)$, for any $V_1 \subseteq V$.

We consider the eigenvalues $\lambda_1(M) \geq \ldots \geq \lambda_n(M)$ of a symmetric matrix $M \in \mathbb{R}^{n\times n}$, arranged in non-increasing order and listed with multiplicities. We denote the corresponding eigenvectors $\mathbf{v}_1(M), \ldots, \mathbf{v}_n(M)$, with $\mathbf{v}_i(M)$ associated with eigenvalue $\lambda_i(M)$. By convention, $\mathbf{v}_1(M)$ is the leading eigenvector and $\{\mathbf{v}_1(M), \ldots, \mathbf{v}_i(M)\}$ are the i principal eigenvectors.

We denote by I_n the identity matrix of size $n \times n$, and by J_n the $n \times n$ matrix with all elements being 1. For a matrix $M \in \mathbb{R}^{n\times n}$, we use $M_{i,:}$ to indicate its i-th row, and $M_{:,j}$ to refer to its j-th column. We also use $M_{i:,j:}$ to indicate the submatrix of M that consists of rows i to n, and columns j to n. We use $\mathrm{tr}(\cdot)$ to denote the trace of a matrix, $\langle\cdot,\cdot\rangle_F$ to denote the Frobenius product between two matrices, and $\langle\cdot,\cdot\rangle$ to denote the dot product between two vectors. We use $\theta(\mathbf{u}, \mathbf{v}) = \arccos(\langle\mathbf{u}, \mathbf{v}\rangle/(\|\mathbf{u}\|_2\|\mathbf{v}\|_2)) \in [0, \pi]$ to indicate the angle between two nonzero vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. Finally, we write $[n]$ to denote the set $\{1, \ldots, n\}$.

Note. All omitted proofs can be found in the supplementary material.

4 Problem formulation

Given a signed graph $G = (V, E)$ and an integer k, our goal is to find k mutually-disjoint node sets $S_1, \ldots, S_k \subseteq V$ that have the following informally-stated properties:

Property 1 *For all $i, j \in [k]$, with $i \neq j$, the edges in $E(S_i)$ are mostly positive, whereas the edges in $E(S_i, S_j)$ are mostly negative.*

Property 2 *There should be a large number of interactions among the nodes of $S_1, \ldots, S_k$ relative to the total number of nodes in these groups. In other words, the subgraph induced by $S_1, \ldots, S_k$ should be as dense as possible.*

Inspired by the formulation of Bonchi et al. [4], our objective function is also a variant of the correlation-clustering problem [2], but with certain differences that we discuss below. For a set of groups $S_1, \ldots, S_k$ as a candidate solution, we quantify Property 1 by using the objective

$$f(S_1, \ldots, S_k) = \sum_{h\in[k]} \sum_{(i,j)\in E(S_h)} A_{i,j} + \frac{1}{k-1} \sum_{\substack{h,\ell\in[k]\\ h\neq\ell}} \sum_{(i,j)\in E(S_h,S_\ell)} (-A_{i,j}). \tag{1}$$

Compared to the standard objective of correlation clustering [2], which treats all edges equally, our objective in Equation (1) weighs an intra-group edge $k - 1$ times more heavily than an inter-group edge. The rationale is as follows: suppose the group sizes and edge densities stay fixed as k increases. Since the number of inter-group edges grows quadratically with k and the number of intra-group edges grows linearly with k, the weighting in Equation (1) prevents the inter-group edges from dominating the objective. The value of $(k - 1)$ in the denominator is chosen so that our objective reduces to the standard case, i.e., the formulation of Bonchi et al. [4], when $k = 2$.

By introducing an indicator matrix $X \in \{0, 1\}^{n\times k}$ with $X_{i,j} = 1$ if node $i \in S_j$ and 0 otherwise, our objective in Equation (1) can be rewritten as

$$f(S_1, \ldots, S_k) = \langle A, XX^T\rangle_F - \frac{\langle A, XJ_kX^T\rangle_F - \langle A, XX^T\rangle_F}{k-1} = \frac{\langle A, XL_kX^T\rangle_F}{k-1}, \tag{2}$$

where $L_k = kI_k - J_k$. The term XL_kX^T in Equation (2) captures explicitly the relationship between the k groups as $(XL_kX^T)_{i,j}$ is positive (negative) whenever nodes i and j are in the same (different) groups. Also, $(XL_kX^T)_{i,j} = 0$ if either node i or node j does not belong to any of the groups.

Hence, the value of the Frobenius product $\langle A, XL_kX^T\rangle_F$ quantifies Property (1) for the groups $S_1, \ldots, S_k$ (which are encoded in matrix X).

Next, we analyze the matrix L_k, which is a fixed matrix not depending on the input signed graph. Let $L_k = UDU^T$ be the eigendecomposition of L_k, where $D = \text{diag}([0, k, \ldots, k]) \in \mathbb{R}^{k\times k}$, and $U \in \mathbb{R}^{k\times k}$ is a real-valued orthogonal matrix. As the geometric multiplicity of eigenvalue k is $k-1$, the matrix U is not unique. For the rest of the paper, we restrict our choice of U to be the following

$$\begin{array}{lll} (U_{:,1})^T = 1/\sqrt{k}\,[1, \ldots, 1], & & (U_{:,2})^T = c_1\,[k-1, -1, \ldots, -1], \\ (U_{:,3})^T = c_2\,[0, k-2, -1, \ldots, -1], & \ldots & (U_{:,k})^T = c_{k-1}\,[0, \ldots, 0, 1, -1], \end{array} \tag{3}$$

where $c_i = 1/\sqrt{(k-i+1)(k-i)}$, for $i = 1, \ldots, k-1$.

By the change of variables $Y = XU$, we can rewrite our objective in Equation (2) as

$$\langle A, XL_kX^T\rangle_F = \langle A, Y\text{diag}([0, k, \ldots, k])Y^T\rangle_F = k\,\text{tr}((Y_{:,2:})^T A(Y_{:,2:})). \tag{4}$$

To account for Property (2) we normalize our objective with the total number of nodes in the groups $S_1, \ldots, S_k$, which can be written as

$$\sum_{i\in[k]} |S_i| = \text{tr}(Y^TY) = k\,(Y_{:,1})^T(Y_{:,1}) = \frac{k}{k-1}\,\text{tr}((Y_{:,2:})^T(Y_{:,2:})). \tag{5}$$

Finally we replace the constraints on the indicator matrix X with the constraint that the rows of Y should take values in the set $\{\mathbf{0}, U_{1,:}, \ldots, U_{k,:}\}$. The equivalence holds since $X_{i,:}$ picks the j-th row of U if $i \in S_j$. Putting all this together, we can now give the final formulation of our problem:

$$\begin{aligned} &\max_{Y\in\mathbb{R}^{n\times k}\setminus\{\mathbf{0}\}} \frac{\text{tr}((Y_{:,2:})^T A(Y_{:,2:}))}{\text{tr}((Y_{:,2:})^T(Y_{:,2:}))}, \\ &\text{subject to} \quad Y_{i,:} \in \{\mathbf{0}, U_{1,:}, \ldots, U_{k,:}\}, \text{ for all } i = 1, \ldots, n. \end{aligned} \tag{6}$$

Intuitively, our objective aims to find small-size conflicting groups with many edges satisfying Property (1). Note that if we ignore the weighting between the inter-group and intra-group edges, Equation (6) can be expressed as ($\#\{$edges satisfying Property (1)$\} - \#\{$edges violating Property (1)$\}$) divided by $|\cup_{h\in[k]} S_h|$.

Also, note that our optimization problem, as formulated above, is different from the trace-maximization problem [26], which given two $n\times n$ matrices M and A, seeks to find an $n \times d$ matrix Z to maximize the form $\text{tr}(Z^TAZ)$, subject to the constraint $Z^TMZ = I_d$. The reason is that since we have no constraint on the group sizes, there is no predefined matrix M to require $X^TMX = I_k$.

5 Proposed spectral approach

The problem we study has been shown to be **APX**-hard for the special case of $k = 2$ [3]. Here we consider a generalization for any $k \geq 2$. In this section we present an efficient spectral algorithm by leveraging the problem formulation (6).

Our starting point is that matrix U, as seen in Equations (3), is almost lower-triangular. We can use this observation to partition $Y_{:,2:}$ column-wise, and reformulate the constraints in problem formulation (6) as follows:

$$Y_{:,2} \in \{0, -c_1, c_1(k-1)\}^n \text{ implies } Y_{i,2} = \begin{cases} c_1(k-1) & \text{if } i \in S_1, \\ -c_1 & \text{if } i \in \cup_{h=2}^k S_h, \end{cases}$$

$$Y_{:,3} \in \{0, -c_2, c_2(k-2)\}^n \text{ implies } Y_{i,3} = \begin{cases} 0 & \text{if } i \in S_1, \\ c_2(k-2) & \text{if } i \in S_2, \\ -c_2 & \text{if } i \in \cup_{h=3}^k S_h, \end{cases}$$

$$\vdots$$

$$Y_{:,k} \in \{0, -c_{k-1}, c_{k-1}\}^n \text{ implies } Y_{i,k} = \begin{cases} 0 & \text{if } i \in \cup_{h=1}^{k-2} S_h, \\ c_{k-1} & \text{if } i \in S_{k-1}, \\ -c_{k-1} & \text{if } i \in S_k. \end{cases}$$

Algorithm 1: SCG (A, k) `Spectral Conflicting Group detection`

Input : A is the adjacency matrix of the signed network; k is the number of groups.
Output : Groups $S_1, \ldots, S_k$.

$A^{(0)} \leftarrow A$;
for $t = 1, \ldots, k-1$ **do**
  $\mathbf{r}^{(t)} \leftarrow$ Solve-Max-DRQ $(A^{(t-1)}, k-t)$; `// See Algorithm 2`
  if $t < k-1$ **then**
    $S_t \leftarrow \{i \notin \cup_{j=1}^{t-1} S_j : |\mathbf{r}_i^{(t)}| = (k-t)\}$;
    $A^{(t)} \leftarrow A^{(t-1)}$;
    $A_{i,:}^{(t)} \leftarrow \mathbf{0}_{1\times n}$ and $A_{:,i}^{(t)} \leftarrow \mathbf{0}_{n\times 1}$ for all $i \in S_t$; `// Remove edges` $E(S_t, V)$
  else $S_{k-1} \leftarrow \{i \notin \cup_{j=1}^{t-1} S_j : \mathbf{r}_i^{(t)} = 1\}$ and $S_k \leftarrow \{i \notin \cup_{j=1}^{t-1} S_j : \mathbf{r}_i^{(t)} = -1\}$;
end
return $S_1, \ldots, S_k$;

Notice that $Y_{i,j} = 0$ for all $i \in \cup_{h=1}^{j-2} S_h$ and $Y_{i,j} = -c_{j-1}$ for all $i \in \cup_{h=j}^{k} S_h$. We let $A^{(0)} = A$, and we define $A^{(t)}$ to be the adjacency matrix that results after removing from $A^{(t-1)}$ all entries that correspond to edges incident to nodes in S_t. Then, the objective function (6) is equivalent to

$$\frac{\mathrm{tr}((Y_{:,2:})^T A(Y_{:,2:}))}{\mathrm{tr}((Y_{:,2:})^T (Y_{:,2:}))} = \sum_{t=1}^{k-1} w_t \frac{(Y_{:,t+1})^T A(Y_{:,t+1})}{(Y_{:,t+1})^T (Y_{:,t+1})} = \sum_{t=1}^{k-1} w_t \frac{(Y_{:,t+1})^T A^{(t-1)}(Y_{:,t+1})}{(Y_{:,t+1})^T (Y_{:,t+1})}, \tag{7}$$

where $w_t = (Y_{:,t+1})^T (Y_{:,t+1}) / \mathrm{tr}((Y_{:,2:})^T (Y_{:,2:})) \in [0,1]$ and $\sum_{t=1}^{k-1} w_t = 1$. In other words, Equation (7) shows that the objective function (6) is a convex combination of $k-1$ discrete Rayleigh quotients. Moreover, Equation (7) also suggests that the solution $Y_{:,t+1}$ characterizes the group S_t that conflicts the most with the (not yet decided) rest of groups S_h for $h > t$. Based on this observation, we propose a scheme SCG (spectral conflicting groups), shown as Algorithm 1.

SCG executes $k-1$ iterations. At the t-th iteration, for each $t \in [k-1]$, we find the vector $Y_{:,t+1}$ that maximizes the discrete Rayleigh quotient of $A^{(t-1)}$, while satisfying the constraints set on matrix Y. We refer to this problem as MAX-DRQ:

$$\mathbf{r}^{(t)} = \underset{\mathbf{x} \in \{0,-1,k-t\}^n \setminus \{\mathbf{0}\}}{\mathrm{argmax}} \frac{\mathbf{x}^T A^{(t-1)} \mathbf{x}}{\mathbf{x}^T \mathbf{x}}. \tag{8}$$

The vector $Y_{:,t+1}$ is then given by $Y_{:,t+1} = c_t\, \mathbf{r}^{(t)}$. We note that our scheme works with any method that solves the MAX-DRQ problem. In Algorithm 1 (SCG) we refer to such a general method as Solve-Max-DRQ. Strategies to solve MAX-DRQ are presented in Section 6. Once the MAX-DRQ problem is solved in the t-th iteration, the vector $\mathbf{r}^{(t)}$ is obtained. If $t < k-1$, the t-th group is recovered by $S_t = \{i \notin \cup_{j=1}^{t-1} S_j : |\mathbf{r}_i^{(t)}| = (k-t)\}$, and if $t = k-1$ (last iteration), the last two groups are recovered by $S_{k-1} = \{i \notin \cup_{j=1}^{t-1} S_j : \mathbf{r}_i^{(t)} = 1\}$ and $S_k = \{i \notin \cup_{j=1}^{t-1} S_j : \mathbf{r}_i^{(t)} = -1\}$.

Note that Equation (7) justifies why it is not a good idea to use the $k-1$ principal vectors of A to identify the conflicting groups: the reason is that the coefficients $[w_t]$ are not fixed values.

6 Solving the maximum discrete Rayleigh quotient problem

In this section we present two solutions for MAX-DRQ. Our first solution is a deterministic algorithm presented in Section 6.1. The second solution is a randomized algorithm presented in Section 6.2. Both solutions first compute the leading eigenvector $\mathbf{v}_1$ of the input matrix $A^{(t-1)}$, and then round $\mathbf{v}_1$ to the appropriate discrete form. The difference is the rounding method. We refer to this generic algorithm as Solve-Max-DRQ, and it is the procedure used in the iterative step of SCG. Pseudocode for Solve-Max-DRQ is given as Algorithm 2.

Algorithm 2: Solve-Max-DRQ (A, q) `Find maximum discrete Rayleigh quotient`

Input : Square and symmetric matrix A, and positive integer q.
Output : The rounded vector $\mathbf{r} \in \{0, -1, q\}^n$.
$\mathbf{v} \leftarrow$ the leading eigenvector of A;
$(d_1, \mathbf{r}_1) \leftarrow$ Round$(\mathbf{v}, q)$; // $d_1 = \sin\theta(\mathbf{v}, \mathbf{r}_1)$
$(d_2, \mathbf{r}_2) \leftarrow$ Round$(-\mathbf{v}, q)$; // $d_2 = \sin\theta(-\mathbf{v}, \mathbf{r}_2)$
if $d_1 \leq d_2$ **then** $\mathbf{r} \leftarrow \mathbf{r}_1$;
else $\mathbf{r} \leftarrow \mathbf{r}_2$;
return $\mathbf{r}$;

Algorithm 3: MinAngleRound $(\mathbf{v}, q)$ `Deterministic rounding by minimum-angle heuristic`

Input : Vector $\mathbf{v} \in \mathbb{R}^n$ and positive integer q.
Output : Vector $\mathbf{u}^* \in \{0, -1, q\}^n$ with min angle to $\mathbf{v}$.
$\{i_k\}_{k=1}^n \leftarrow$ Sort $\mathbf{v}$ and return the indexes such that $\mathbf{v}_{i_1} \geq \ldots \geq \mathbf{v}_{i_n}$;
$(d, \mathbf{u}^*) \leftarrow (\infty, \mathbf{0})$;
$(k_1, k_2) \leftarrow (0, n+1)$;
while $k_1 < k_2$ **do**
  $\mathbf{u}_1 \leftarrow$ set the i_{k_1+1}-th element of $\mathbf{u}^*$ to q;
  $\mathbf{u}_2 \leftarrow$ set the i_{k_2-1}-th element of $\mathbf{u}^*$ to -1;
  if $\min\{\sin\theta(\mathbf{v}, \mathbf{u}_1), \sin\theta(\mathbf{v}, \mathbf{u}_2)\} \geq d$ **then** break;
  if $\sin\theta(\mathbf{v}, \mathbf{u}_1) < \sin\theta(\mathbf{v}, \mathbf{u}_2)$ **then** $(k_1, d, \mathbf{u}^*) \leftarrow (k_1 + 1, \sin\theta(\mathbf{v}, \mathbf{u}_1), \mathbf{u}_1)$;
  else $(k_2, d, \mathbf{u}^*) \leftarrow (k_2 - 1, \sin\theta(\mathbf{v}, \mathbf{u}_2), \mathbf{u}_2)$;
end
return $(d, \mathbf{u}^*)$;

6.1 Deterministic rounding

Our goal is to find a discrete vector $\mathbf{v}^* \in \{0, -1, q\}^n$ that maximizes the quotient $\mathbf{x}^T A^{(t-1)} \mathbf{x} / (\mathbf{x}^T \mathbf{x})$. Let $\mathbf{v}$ be the leading eigenvector of $A^{(t-1)}$, i.e., the real-valued maximizer of $\mathbf{x}^T A^{(t-1)} \mathbf{x} / (\mathbf{x}^T \mathbf{x})$. The idea is to round $\mathbf{v}$ to a discrete vector $\mathbf{u}^* \in \{0, -1, q\}^n$ that minimizes $\sin\theta(\mathbf{v}, \mathbf{u})$, among all vectors $\mathbf{u} \in \{0, -1, q\}^n$. It can be shown that such $\mathbf{u}^*$ can be found by restricting the search over $\mathcal{O}(n^2)$ thresholded candidate vectors obtained by $\mathbf{v}$. We formalize this below.

Definition 1 *Let $\mathbf{v} \in \mathbb{R}^n$, $q \in [k-1]$ and $a, b \in \mathbb{R}$ be given. Define a threshold function $\sigma_{a,b} : \mathbb{R}^n \to \mathbb{R}^n$ that maps $\mathbf{v}$ to a new vector $\sigma_{a,b}(\mathbf{v})$, whose i-th coordinate is*

$$\sigma_{a,b}(\mathbf{v})_i = \begin{cases} q & \text{if } \mathbf{v}_i \geq a > 0, \\ -1 & \text{if } \mathbf{v}_i \leq b < 0, \\ 0 & \text{otherwise}, \end{cases}$$

and denote $\mathcal{T} = \{t_i\}_{i=0}^{n+1}$ the sequence of all possible thresholds over the coordinates of $\mathbf{v}$, that is, $t_0 = \infty$, $t_{n+1} = -\infty$ and t_i is the i-th largest coordinate of $\mathbf{v}$, for $i \in [n]$. Then, the set of all possible thresholded vectors of $\mathbf{v}$ is denoted by $\Gamma(\mathbf{v}) = \{\sigma_{a,b}(\mathbf{v}) \ : \text{for all } a, b \in \mathcal{T}\}$.

Given a vector $\mathbf{v}$, the discrete vector $\mathbf{u}^* \in \{0, -1, q\}^n$ that minimizes $\sin\theta(\mathbf{v}, \mathbf{u})$ can be computed by using the following result.

Lemma 1 *Let $\mathbf{v} \in \mathbb{R}^n$ and $q \in [k-1]$ be given. The minimizer of $\sin\theta(\mathbf{v}, \mathbf{u})$ over all $\mathbf{u} \in \{0, -1, q\}^n$ is equal to the minimizer of $\sin\theta(\mathbf{v}, \mathbf{u})$ over all $\mathbf{u} \in \Gamma(\mathbf{v}) \cup \Gamma(-\mathbf{v})$.*

Since the size of the set $\Gamma(\mathbf{v}) \cup \Gamma(-\mathbf{v})$ is $\mathcal{O}(n^2)$, enumerating all vectors to find the optimal $\mathbf{u}$ is not efficient for large datasets. To make our method scalable, we propose a linear-time rounding heuristic in Algorithm 3, which finds a local optimum.

This heuristic works by initializing two indexes k_1, k_2, the indexes of the two thresholds, which are initially set to 0 and $n+1$, respectively. At each iteration, we move only 1 threshold, we either increase k_1 by 1 or decrease k_2 by 1. This is determined by comparing $\sin\theta(\mathbf{v}, \sigma_{t_{k_1+1}, t_{k_2}}(\mathbf{v}))$ and $\sin\theta(\mathbf{v}, \sigma_{t_{k_1}, t_{k_2-1}}(\mathbf{v}))$ and choosing the smaller option.

Algorithm 4: RandomRound $(\mathbf{v}, q)$ `Randomized rounding`

Input : Vector $\mathbf{v} \in \mathbb{R}^n$ and positive integer q.
Output : Vector $\mathbf{u} \in \{0, -1, q\}^n$, a randomized rounded vector of $\mathbf{v}$.
$\mathbf{u} \leftarrow \mathbf{0}$;
for $i = 1, \dots, n$ **do**
  if $\mathbf{v}_i > 0$ **then** $\mathbf{u}_i \leftarrow q\,\text{Bernoulli}(|\mathbf{v}_i|/q)$;
  else if $\mathbf{v}_i < 0$ **then** $\mathbf{u}_i \leftarrow (-1)\,\text{Bernoulli}(|\mathbf{v}_i|)$;
end
$d \leftarrow \sin\theta(\mathbf{v}, \mathbf{u})$;
return $(d, \mathbf{u})$;

6.2 Randomized rounding

Our second algorithm for maximizing $\mathbf{x}^T A^{(t-1)}\mathbf{x}/(\mathbf{x}^T\mathbf{x})$ in $\{0, -1, q\}^n$ is a randomized-rounding scheme starting with the eigenvector $\mathbf{v}$ of $A^{(t-1)}$. Pseudocode is shown in Algorithm 4.

In more detail, we round $\mathbf{v}$ onto $\{0, -1, q\}^n$ by drawing Bernoulli trials. For each positive coordinate $\mathbf{v}_i$ we set $\mathbf{u}_i \sim q\,\text{Bernoulli}(|\mathbf{v}_i|/q)$, for each negative coordinate $\mathbf{v}_i$ we set $\mathbf{u}_i \sim (-1)\,\text{Bernoulli}(|\mathbf{v}_i|)$, and if $\mathbf{v}_i = 0$ we set $\mathbf{u}_i = 0$. In this way, we have $\mathbb{E}[\mathbf{u}] = \mathbf{v}$. By applying similar arguments to the ones presented by Bonchi et al. [4], we can show that the randomized-rounding algorithm provides a $\mathcal{O}(q\sqrt{n})$-approximation guarantee to the MAX-DRQ problem. We present this result as Theorem 1. Furthermore, Corollary 1 states that this result is tight for $k = 2$.

Theorem 1 *Let $\mathbf{v}$ be the leading eigenvector of the adjacency matrix A of a signed graph, and let $q \geq 1$ be a positive integer. Then, the* RandomRound *algorithm with $(\mathbf{v}, q)$ as input is a $(q\sqrt{n})$-approximation to the optimum of the corresponding* MAX-DRQ *problem.*

Lemma 2 *Let OPT be the optimum solution to the* MAX-DRQ *problem. There exists a problem instance such that $\lambda_1(A) \geq OPT \cdot \Omega(\sqrt{n})$.*

Corollary 1 *The integrality gap of algorithm* RandomRound *is $\Omega(\sqrt{n})$, and thus, the approximation result of Theorem 1 is asymptotically tight up to a factor of q.*

7 Experimental evaluation

In this section, we evaluate our framework with both synthetic and real-world graphs. All the experiments are executed on a machine with Intel Core i5 at 1.8 GHz with 8 GB RAM. All methods have been implemented in Python 3.[1] The datasets we have used are all publicly available and the detailed information can be found in Supplementary § D.1. Beyond the experiments discussed here, we present more results in Supplementary § D, including execution times, and a discussion on deciding the number of groups k.

Proposed methods. Our approach (SCG) is a framework that admits different methods to solve MAX-DRQ. We have instantiated our framework with the following routines. *Minimum angle*: the deterministic rounding algorithm presented in Section 6.1; *Randomized rounding*: the randomized rounding algorithm presented in Section 6.2; *Maximum objective*: a generalization of EigenSign [4], that rounds $\mathbf{v}_1(A)$ by finding an optimal threshold to maximize the objective; *Bansal*: motivated by the pivot for correlation clustering [2], which finds two conflicting groups by considering the neighborhood of a single node, and using the node that results in the maximum value of the objective. These instantiations are denoted by SCG-MA, SCG-R, SCG-MO, and SCG-B, respectively.

Baselines. We use the following baselines: KOCG [11] is a method designed for a similar formulation to ours. We use the authors' implementation [10] with default hyperparameters $\alpha = 1/(k-1)$, $\beta = 50$, and $\ell = 5000$. As KOCG returns a ranked list of disjoint subgraphs, each containing k conflicting groups, we pick the k groups contained in their *top-1* and *top-r* subgraphs. We choose r so that the total group size equals the one returned by SCG-MA. We use two spectral algorithms: BNC [8], which optimizes *balanced normalized cut*; and SPONGE [14], a method particularly suitable

[1] https://github.com/rutzeng/SCG-NeurIPS2020.

Table 1: Polarity objective (Equation (6)) achieved by the proposed methods and the baselines on real-world signed graphs, for two different values of k: the number of conflicting groups to be detected. Dashes indicate that a method exceeded the memory limit.

		`WoW-EP8`	`Bitcoin`	`WikiVot`	`Referendum`	`Slashdot`	`WikiCon`	`Epinions`	`WikiPol`
	$\lvert V\rvert$	790	5 881	7 115	10 884	82 140	116 717	131 580	138 587
	$\lvert E\rvert$	116 009	21 492	100 693	251 406	500 481	2 026 646	711 210	715 883
	$\lvert E_-\rvert/\lvert E\rvert$	0.2	0.2	0.2	0.1	0.2	0.6	0.2	0.1
$k=2$	SCG-MA	**236.6**	**28.8**	**71.5**	**172.2**	**77.5**	**155.2**	128.3	**82.8**
	SCG-MO	**236.6**	**29.5**	**71.7**	**174.1**	**79.7**	**175.7**	**128.7**	**88.4**
	SCG-B	200.6	21.7	37.6	116.3	61.0	129.3	**156.4**	46.5
	SCG-R	218.3	14.9	55.7	119.6	29.9	100.2	70.9	36.0
	KOCG-top-1	9.0	3.6	4.0	4.3	1.0	6.2	4.2	1.0
	KOCG-top-r	18.2	3.8	2.5	14.0	3.7	2.4	6.2	0.9
	BNC-k	184.6	5.3	15.8	41.5	—	—	—	—
	BNC-$(k+1)$	-0.7	-10.8	-1.0	-1.0	—	—	—	—
	SPONGE-k	191.4	5.1	15.8	41.5	—	—	—	—
	SPONGE-$(k+1)$	88.0	1.0	1.0	1.0	—	—	—	—
$k=6$	SCG-MA	207.3	**14.6**	**45.5**	**84.9**	**37.8**	**102.6**	88.8	**57.5**
	SCG-MO	**226.9**	**15.2**	**47.0**	55.6	34.6	**111.6**	**129.2**	41.8
	SCG-B	**211.6**	9.3	23.3	**116.2**	**47.7**	46.1	**94.5**	**46.0**
	SCG-R	198.1	5.0	9.7	39.8	7.3	16.2	39.4	5.5
	KOCG-top-1	7.0	4.4	5.5	8.8	2.6	4.5	8.7	4.8
	KOCG-top-r	8.5	2.9	2.9	5.0	3.6	4.0	6.5	3.0
	BNC-k	185.2	5.2	15.8	41.5	—	—	—	—
	BNC-$(k+1)$	-0.2	-4.2	-1.1	-0.8	—	—	—	—
	SPONGE-k	58.5	5.0	15.8	41.5	—	—	—	—
	SPONGE-$(k+1)$	48.1	0.8	1.0	1.0	—	—	—	—

for sparse graphs and large k. To detect k conflicting groups using the spectral clustering algorithms, we compare with two approaches. The first approach is to directly apply BNC and SPONGE to detect k clusters and return all the detected clusters as conflicting groups. The second approach is to detect $(k+1)$ clusters, then heuristically treat the largest cluster as the non-conflicting cluster, and return the k smallest clusters as the detected conflicting groups. Let BNC-k and SPONGE-k denote SPONGE and BNC with the first approach and let BNC-$(k+1)$ and SPONGE-$(k+1)$ denote the two with the second approach. We use a publicly-available implementation [15] for BNC and SPONGE.

Results on real-world networks. We first measure the quality of the proposed methods and baselines with respect to the polarity objective, i.e., Equation (6), on real-world signed graphs. The results are shown in Table 1. The running times of all methods are listed in Supplementary § D.2. We observe that mostly, SCG-MA and SCG-MO achieve the best polarity scores. They are also the fastest, and usually find larger groups. An example of the sizes of the groups found by all methods is given in Supplementary § D.3. The SCG-B algorithm identifies conflicting groups by exploring local neighborhoods, and its detected groups tend to be located around high-degree nodes. Although SCG-B achieves the largest polarity on `Referendum` for $k=6$, it only detects 2 groups, already covered by SCG-MA and SCG-MO. As the groups are not necessarily the high-degree nodes, SCG-B performs less competitive on `WikiVot` and `WikiCon` for $k=6$. Finally, SCG-R is not as competitive as SCG-MA or SCG-MO and is slower due to random sampling.

With respect to our direct competitor KOCG, the KOCG-top-1 variant performs slightly better than KOCG-top-r when $k=6$. As KOCG finds groups in local regions, KOCG-top-1 returns much smaller groups than the other methods. On the contrary, KOCG-top-r intersects several local groups in different graph regions but remains ineffective compared to SCG-MA under the same total group size. All KOCG settings perform worse than BNC and SPONGE on the first 4 datasets.

Finally, the spectral-clustering methods BNC and SPONGE exceed the memory limit (caused by k-means) on large datasets. The k groups returned by BNC-k and SPONGE-k usually consist of one large group with many non-conflicting nodes and $k-1$ very small groups. Since BNC-$(k+1)$ and SPONGE-$(k+1)$ can use the spare cluster to put the non-conflicting nodes, we expect they perform better than BNC-k and SPONGE-k but it turns out to be worse on all 4 real-world networks. Despite of the unexpected results, both versions of BNC and SPONGE are less effective than SCG-MA and SCG-MO at finding conflicting groups in real-world graphs.

Results on synthetic graphs. In our second experiment, we use synthetic graphs to measure how well the methods recover ground-truth conflicting groups. We use the *modified signed stochastic*

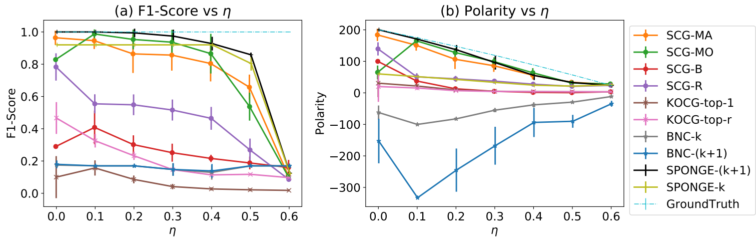


Figure 1: F_1-score (left) and polarity score (right) as a function of the parameter η. The input signed graphs are generated by the m-SSBM model, for a graph of size $n = 2\,000$, with $k = 6$ ground-truth groups, each having size $\ell = 100$.

block model (m-SSBM) [4], which has 4 parameters; n: the graph size; k: the number of conflicting groups; ℓ: the size of each of the conflicting groups (all have the same size); and $\eta \in [0, 1]$: a parameter that controls the edge probabilities. Edges in the same group are positive with probability $1 - \eta$ and negative or absent with probability $\eta/2$. Edges between distinct groups are negative with probability $1 - \eta$ and positive or absent with probability $\eta/2$. All other edges have equal probability of $\min(\eta, 1/2)$ of being positive or negative. Hence, the smaller the value of η, the denser the conflicting groups and the lower the noise level. Note that the conflicting groups only emerge when $\eta \leq 2/3$, since m-SSBM is expected to have more negative edges in the groups and more positive edge between groups if $\eta > 2/3$.

In this experiment we measure the recovery rate of the ground-truth groups using the F_1 score, with precision and recall averaged over all groups. In Figure 1 we report the results of the m-SSBM model with parameters $n = 2\,000$, $k = 6$, $\ell = 100$, and $\eta = 0 : 0.1 : 0.6$. Each setting is repeated 20 times, and we report the average F_1 score and polarity scores.

As seen in Figure 1, the recovery rate (F_1 score) for all methods declines with η, since the graph becomes sparser and more noisy. It is clear that SCG-MA and SCG-MO are robust methods, handling very well the increasing noise level. It is worth noting that SPONGE-$(k + 1)$ performs the best in this experiment with respect to both F_1 and polarity. We also see that SCG-B is less competitive here, as in this data the conflicting groups are not concentrated around high-degree nodes. In summary, under the m-SSBM model, our polarity score is consistent with the F_1 score, and our proposed methods SCG-MA and SCG-MO are effective in detecting the ground-truth conflicting groups.

8 Conclusions and future work

We propose an efficient method for detecting k conflicting groups in a signed network. Our approach relies on interpreting the problem objective in terms of the Laplacian of a complete graph, characterizing the spectral properties of this matrix, and deriving a novel formulation in which each conflicting group is characterized by the solution to the maximum discrete Rayleigh quotient problem.

Our work opens several exciting directions for future work. First, it remains open whether we can improve the $\mathcal{O}(\sqrt{n})$-approximation for the maximum discrete Rayleigh quotient problem, using an approach that does not rely on rounding the leading eigenvector, such as by extending the SDP-based algorithm in [3]. Second, it would be interesting to explore the applicability of our approach to unsigned graphs for the task of detecting dense subgraphs. Third, the modified Stochastic Block Model (m-SSBM) is actually a special case of Label Stochastic Block Model (LSBM) [23]. It would be relevant to analyze the recovery guarantee of our proposed method in m-SSBM with respect to the fundamental limit results [44] and the interplay with the Bethe-Hessian operator [34] in the sparse regime. Finally, the difference in the empirical performance of our two rounding techniques and the spectral clustering baseline SPONGE [14] in the real-world networks and the synthetic network is somewhat striking. It is possible that some properties or structures exist in the real-world networks but not in the synthetic networks. An interesting question is to explain this behavior analytically, in particular with respect to properties of real-world networks.

Acknowledgments

We thank the anonymous reviewers for their insightful feedback. We thank Stefan Neumann for pointing to us the work of Bhaskara et al. [3]. This research is supported by the Academy of Finland projects AIDA (317085) and MLDB (325117), the ERC Advanced Grant REBOUND (834862), the EC H2020 RIA project SoBigData (871042), and the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation.

Broader Impact

As the task we tackle in this paper belongs to the broad category of data mining, and as our study is mainly of theoretical nature, the impact of our work to the society is indirect. With respect to positive consequences, we name two possible applications that could impact the modern society. First, the rise of polarization and fake news is related to the existence of conflicting groups. Thus, having an efficient characterization tool is the first step to mitigate the situation. Second, both collaboration and competition exist in a diverse environment and detecting conflicting groups helps to understand the interplay of the two. With respect to negative consequences, we do not foresee specific issues when applying our method.

References

[1] Lada A Adamic and Natalie Glance. The political blogosphere and the 2004 us election: divided they blog. In *Proc. of Link discovery workshop*, 2005.

[2] Nikhil Bansal, Avrim Blum, and Shuchi Chawla. Correlation clustering. *Machine learning*, 2004.

[3] Aditya Bhaskara, Moses Charikar, Rajsekar Manokaran, and Aravindan Vijayaraghavan. On quadratic programming with a ratio objective. In *Proc. of ICALP*. Springer, 2012.

[4] Francesco Bonchi, Edoardo Galimberti, Aristides Gionis, Bruno Ordozgoiti, and Giancarlo Ruffo. Discovering polarized communities in signed networks. In *Proc. of CIKM*, 2019.

[5] Moses Charikar and Anthony Wirth. Maximizing quadratic programs: Extending grothendieck's inequality. In *Proc. of FOCS*, 2004.

[6] Moses Charikar, Venkatesan Guruswami, and Anthony Wirth. Clustering with qualitative information. *Journal of Computer and System Sciences*, 2005.

[7] Yiqi Chen, Tieyun Qian, Huan Liu, and Ke Sun. " bridge" enhanced signed directed network embedding. In *Proc. of CIKM*, 2018.

[8] Kai-Yang Chiang, Joyce Jiyoung Whang, and Inderjit S Dhillon. Scalable clustering of signed networks using balance normalized cut. In *Proc. of CIKM*, 2012.

[9] Kai-Yang Chiang, Cho-Jui Hsieh, Nagarajan Natarajan, Inderjit S Dhillon, and Ambuj Tewari. Prediction and clustering in signed networks: a local to global perspective. *JMLR*, 2014.

[10] Lingyang Chu. Source code: Finding gangs in war from signed networks. `https://github.com/lingyangchu/KOCG.SIGKDD2016`, 2016.

[11] Lingyang Chu, Zhefeng Wang, Jian Pei, Jiannan Wang, Zijin Zhao, and Enhong Chen. Finding gangs in war from signed networks. In *Proc. of KDD*, 2016.

[12] Nicole A Cooke. *Fake news and alternative facts: Information literacy in a post-truth era*. American Library Association, 2018.

[13] Mihai Cucuringu, Ioannis Koutis, Sanjay Chawla, Gary Miller, and Richard Peng. Simple and scalable constrained clustering: a generalized spectral method. In *Proc. of AISTATS*, 2016.

[14] Mihai Cucuringu, Peter Davies, Aldo Glielmo, and Hemant Tyagi. Sponge: A generalized eigenproblem for clustering signed networks. In *Proc. of AISTATS*, 2019.

[15] Peter Davies and Aldo Glielmo. Signet: A package for clustering of signed networks. `https://github.com/alan-turing-institute/signet`, 2019.

[16] Erik D Demaine, Dotan Emanuel, Amos Fiat, and Nicole Immorlica. Correlation clustering in general weighted graphs. *Proc. of STOC*, 2006.

[17] Seth Flaxman, Sharad Goel, and Justin M Rao. Filter bubbles, echo chambers, and online news consumption. *Public opinion quarterly*, 2016.

[18] Ming Gao, Ee-Peng Lim, David Lo, and Philips Kokoh Prasetyo. On detecting maximal quasi antagonistic communities in signed graphs. *Data mining and knowledge discovery*, 2016.

[19] R Kelly Garrett. Echo chambers online?: Politically motivated selective exposure among internet news users. *Journal of Computer-Mediated Communication*, 2009.

[20] Aristides Gionis, Antonis Matakos, Bruno Ordozgoiti, and Han Xiao. Mining signed networks: Theory and applications: Tutorial proposal for the web conference 2020. In *Companion Proceedings of the Web Conference 2020*, WWW '20, 2020.

[21] Ioannis Giotis and Venkatesan Guruswami. Correlation clustering with a fixed number of clusters. *Proc. of STOC*, 2006.

[22] Frank Harary et al. On the notion of balance of a signed graph. *The Michigan Mathematical Journal*, 1953.

[23] Simon Heimlicher, Marc Lelarge, and Laurent Massoulié. Community detection in the labelled stochastic block model. *Proc. of NIPS Workshop*, 2012.

[24] Amin Javari, Tyler Derr, Pouya Esmailian, Jiliang Tang, and Kevin Chen-Chuan Chang. Rose: Role-based signed network embedding. In *Proc. of The Web Conference*, 2020.

[25] Junghwan Kim, Haekyu Park, Ji-Eun Lee, and U Kang. Side: representation learning in signed directed networks. In *Proc. of WWW*, 2018.

[26] Effrosini Kokiopoulou, Jie Chen, and Yousef Saad. Trace optimization and eigenproblems in dimension reduction methods. *Numerical Linear Algebra with Applications*, 2011.

[27] Jérôme Kunegis, Stephan Schmidt, Andreas Lommatzsch, Jürgen Lerner, Ernesto W De Luca, and Sahin Albayrak. Spectral analysis of signed graphs for clustering, prediction and visualization. In *Proc. of ICDM*, 2010.

[28] Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. Signed networks in social media. In *Proc. of SIGCHI*, 2010.

[29] David Lo, Didi Surian, Kuan Zhang, and Ee-Peng Lim. Mining direct antagonistic communities in explicit trust networks. In *Proc. of CIKM*, 2011.

[30] David Lo, Didi Surian, Philips Kokoh Prasetyo, Kuan Zhang, and Ee-Peng Lim. Mining direct antagonistic communities in signed social networks. *Information processing & management*, 2013.

[31] Michael McCluskey and Young Mie Kim. Moderatism or polarization? representation of advocacy groups' ideology in newspapers. *Journalism & Mass Communication Quarterly*, 2012.

[32] Pedro Mercado, Francesco Tudisco, and Matthias Hein. Spectral clustering of signed graphs via matrix power means. In *Proc. of ICML*, 2019.

[33] Gregory J Puleo and Olgica Milenkovic. Correlation clustering and biclustering with locally bounded errors. In *Proc. of ICML*, 2016.

[34] Alaa Saade, Florent Krzakala, and Lenka Zdeborová. Spectral clustering of graphs with the bethe hessian. In *Proc. of NIPS*, 2014.

[35] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. Fake news detection on social media: A data mining perspective. *Proc. of SIGKDD*, 2017.

[36] Jiliang Tang, Yi Chang, Charu Aggarwal, and Huan Liu. A survey of signed network mining in social media. *ACM Computing Surveys (CSUR)*, 49(3):1–37, 2016.

[37] Ulrike Von Luxburg. A tutorial on spectral clustering. *Statistics and computing*, 2007.

[38] Jing Wang, Jie Shen, Ping Li, and Huan Xu. Online matrix completion for signed link prediction. In *Proc. of WSDM*, 2017.

[39] Suhang Wang, Charu Aggarwal, Jiliang Tang, and Huan Liu. Attributed signed network embedding. In *Proc. of CIKM*, 2017.

[40] Han Xiao, Bruno Ordozgoiti, and Aristides Gionis. Searching for polarization in signed graphs: a local spectral approach. In *Proc. of The Web Conference*, 2020.

[41] Pinghua Xu, Wenbin Hu, Jia Wu, and Bo Du. Link prediction with signed latent factors in signed social networks. In *Proc. of KDD*, 2019.

[42] Shuo Yang, Kai Shu, Suhang Wang, Renjie Gu, Fan Wu, and Huan Liu. Unsupervised fake news detection on social media: A generative approach. In *Proc. of AAAI*, 2019.

[43] Sarita Yardi and Danah Boyd. Dynamic debates: An analysis of group polarization over time on twitter. *Bulletin of science, technology & society*, 2010.

[44] Se-Young Yun and Alexandre Proutiere. Optimal cluster recovery in the labeled stochastic block model. In *Proc. of NIPS*, 2016.

[45] Kuan Zhang, David Lo, and Ee-Peng Lim. Mining antagonistic communities from social networks. In *Proc. of PAKDD*, pages 68–80, 2010.

[46] Kuan Zhang, David Lo, Ee-Peng Lim, and Philips Kokoh Prasetyo. Mining indirect antagonistic communities from social interactions. *Knowledge and information systems*, 2013.

Discovering conflicting groups in signed networks Supplementary material

Ruo-Chun Tzeng
KTH Royal Institute of Technology
rctzeng@kth.se

Bruno Ordozgoiti
Aalto University
bruno.ordozgoiti@aalto.fi

Aristides Gionis
KTH Royal Institute of Technology
argioni@kth.se

A Proof of Lemma 1

Proof: We have $\|\mathbf{v}\|_2 = 1$ and, without loss of generality, we can assume that the coordinates of $\mathbf{v}$ are sorted in non-increasing order. Let $\mathcal{T} = \{t_i\}_{i=0}^{n+1}$ be all possible thresholds for $\mathbf{v}$ and $\mathcal{T}' = \{t'_i\}_{i=0}^{n+1}$ be all possible thresholds for $-\mathbf{v}$. Recall the definition of $\theta(\cdot,\cdot)$ from Section 3 that $\theta(\mathbf{a},\mathbf{b}) = \arccos(\langle \mathbf{a},\mathbf{b}\rangle / \|\mathbf{a}\|_2\|\mathbf{b}\|_2) \in [0,\pi]$ for any two nonnegative vectors $\mathbf{a},\mathbf{b} \in \mathbb{R}^n$, so $\sin\theta(\mathbf{a},\mathbf{b})$ is always non-negative. Let $\mathbf{u}^*$ be the minimizer of $\sin\theta(\mathbf{v},\mathbf{u})$ over all $\mathbf{u} \in \Gamma(\mathbf{v}) \cup \Gamma(-\mathbf{v})$.

For simplicity, we assume $\mathbf{u}^* \in \Gamma(\mathbf{v})$ and $\langle \mathbf{v},\mathbf{u}^*\rangle \geq 0$. This is because if the dot product is negative, we can make it positive by reversing the sign of $\mathbf{v}$. Let k_1^*, k_2^* be the two thresholds such that $\mathbf{u}^* = \sigma_{t_{k_1^*}, t_{k_2^*}}(\mathbf{v})$. We will show that $\sin\theta(\mathbf{v},\mathbf{u}) \geq \sin\theta(\mathbf{v},\mathbf{u}^*)$ for any $\mathbf{u} \in \{0,-1,q\}^n$.

Fix any $\mathbf{u} \in \{0,-1,q\}^n$. Our first step is to identify the coordinates that $\mathbf{u}_i \neq \mathbf{u}_i^*$, denoted by $\mathcal{I} = \{j : \mathbf{u}_j \neq \mathbf{u}_j^*\}$. Moreover, since $\mathbf{u}_j^* = q$ for all $j \leq k_1^*$, $\mathbf{u}_j^* = -1$ for all $j \geq k_2^*$, and $\mathbf{u}_j^* = 0$ for all $j \in (k_1^*, k_2^*)$, we further divide $\mathcal{I}$ into 6 disjoint subsets:

$$
\begin{aligned}
\mathcal{I}_{11} &= \{j \in \mathcal{I} : \mathbf{u}_j = 0, j \leq k_1^*\}, & \mathcal{I}_{12} &= \{j \in \mathcal{I} : \mathbf{u}_j = -1, j \leq k_1^*\}, \\
\mathcal{I}_{21} &= \{j \in \mathcal{I} : \mathbf{u}_j = q, j \in (k_1^*, k_2^*)\}, & \mathcal{I}_{22} &= \{j \in \mathcal{I} : \mathbf{u}_j = -1, j \in (k_1^*, k_2^*)\}, \\
\mathcal{I}_{31} &= \{j \in \mathcal{I} : \mathbf{u}_j = 0, j \geq k_2^*\}, & \mathcal{I}_{32} &= \{j \in \mathcal{I} : \mathbf{u}_j = q, j \geq k_2^*\}.
\end{aligned}
$$

Denote the overall division by k_1^* and k_2^* by $\mathcal{I}_1 = \mathcal{I}_{11} \cup \mathcal{I}_{12}$, $\mathcal{I}_2 = \mathcal{I}_{21} \cup \mathcal{I}_{22}$, and $\mathcal{I}_3 = \mathcal{I}_{31} \cup \mathcal{I}_{32}$.

We claim that for any such $\mathbf{u}$, there exists a vector $\tilde{\mathbf{u}} \in \Gamma(\mathbf{v}) \cup \Gamma(-\mathbf{v})$ such that $\sin\theta(\mathbf{v},\mathbf{u}) \geq \sin\theta(\mathbf{v},\tilde{\mathbf{u}})$, which is sufficient to complete the proof since $\mathbf{u}^*$ is the minimizer of $\sin\theta(\mathbf{v},\mathbf{u})$ for all $\mathbf{u} \in \Gamma(\mathbf{v}) \cup \Gamma(-\mathbf{v})$. We will show how to find such vector $\tilde{\mathbf{u}}$ by examining the following two cases:

(Case 1) $\langle \mathbf{v},\mathbf{u}\rangle \geq 0$:

Let $c_1 = |\mathcal{I}_{21}| - |\mathcal{I}_1|$ and $c_2 = |\mathcal{I}_{22}| - |\mathcal{I}_3|$. The claim is proved by setting $\tilde{\mathbf{u}} = \sigma_{t_{k_1^*+c_1}, t_{k_2^*-c_2}}(\mathbf{v})$, which is justified by the following two observations.

First, observe that $\|\tilde{\mathbf{u}}\|_2 \leq \|\mathbf{u}\|_2$ because $\|\tilde{\mathbf{u}}\|_2^2 + |\mathcal{I}_{12}| + q^2|\mathcal{I}_{32}| = \|\mathbf{u}\|_2^2$.

Second, write $\langle \mathbf{v},\mathbf{u}\rangle$ as

$$
\langle \mathbf{v},\mathbf{u}\rangle = \langle \mathbf{v},\mathbf{u}^*\rangle + q\left(-\sum_{j\in\mathcal{I}_1} \mathbf{v}_j + \sum_{j\in\mathcal{I}_{21}\cup\mathcal{I}_{32}} \mathbf{v}_j\right) + \left(\sum_{j\in\mathcal{I}_3} \mathbf{v}_j + \sum_{j\in\mathcal{I}_{12}\cup\mathcal{I}_{22}} (-\mathbf{v}_j)\right). \tag{1}
$$

34th Conference on Neural Information Processing Systems (NeurIPS 2020), Vancouver, Canada.

Notice that some terms of the summation in Equation (1) are negative, in particular,

$$\sum_{j\in\mathcal{I}_{12}}(-\mathbf{v}_j)+\sum_{j\in\mathcal{I}_{32}}q\mathbf{v}_j<0,$$

since $\mathbf{v}_j>0$ for all $j\in\mathcal{I}_{12}$, and $\mathbf{v}_j<0$ for all $j\in\mathcal{I}_{32}$.

Therefore, we have

$$\langle\mathbf{v},\mathbf{u}\rangle\le\langle\mathbf{v},\mathbf{u}^*\rangle+q\left(-\sum_{j\in\mathcal{I}_1}\mathbf{v}_j+\sum_{j\in\mathcal{I}_{21}}\mathbf{v}_j\right)+\left(\sum_{j\in\mathcal{I}_3}\mathbf{v}_j+\sum_{j\in\mathcal{I}_{22}}(-\mathbf{v}_j)\right). \tag{2}$$

Since $\mathbf{v}$ is sorted non-increasingly, the latter two terms in (2) are smaller than

$$q\left(-\sum_{j=1}^{|\mathcal{I}_1|}\mathbf{v}_{k_1^*-j}+\sum_{j=1}^{|\mathcal{I}_{21}|}\mathbf{v}_{k_1^*-|\mathcal{I}_1|+j}\right)+\left(\sum_{j=1}^{|\mathcal{I}_3|}\mathbf{v}_{k_2^*+j}+\sum_{j=1}^{|\mathcal{I}_{22}|}(-\mathbf{v}_{k_2^*+|\mathcal{I}_3|-j})\right).$$

That is,

$$\begin{aligned}\langle\mathbf{v},\mathbf{u}\rangle&\le\langle\mathbf{v},\mathbf{u}^*\rangle+q\left(-\sum_{j=1}^{|\mathcal{I}_1|}\mathbf{v}_{k_1^*-j}+\sum_{j=1}^{|\mathcal{I}_{21}|}\mathbf{v}_{k_1^*-|\mathcal{I}_1|+j}\right)+\left(\sum_{j=1}^{|\mathcal{I}_3|}\mathbf{v}_{k_2^*+j}+\sum_{j=1}^{|\mathcal{I}_{22}|}(-\mathbf{v}_{k_2^*+|\mathcal{I}_3|-j})\right)\\&=\langle\mathbf{v},\tilde{\mathbf{u}}\rangle\end{aligned}$$

Hence, we have $0\le\cos\theta(\mathbf{v},\mathbf{u})\le\cos\theta(\mathbf{v},\tilde{\mathbf{u}})$, which is equivalent to $\sin\theta(\mathbf{v},\mathbf{u})\ge\sin\theta(\mathbf{v},\tilde{\mathbf{u}})$ due to the non-negativity of $\sin\theta(\cdot,\cdot)$.

(Case 2) $\langle\mathbf{v},\mathbf{u}\rangle<0$:

Let $c_1=|\{j\in\mathcal{I}_{21}:\mathbf{v}_j<0\}|+|\mathcal{I}_{32}|$ and $c_2=|\{j\in\mathcal{I}_{22}:\mathbf{v}_j>0\}|+|\mathcal{I}_{12}|$. The claim is proved by setting $\tilde{\mathbf{u}}=\sigma_{t'_{c_1},t'_{c_2}}(-\mathbf{v})$, which is justified in the below two observations.

First, observe that $\|\tilde{\mathbf{u}}\|_2\le\|\mathbf{u}\|_2$ because

$$\|\tilde{\mathbf{u}}\|_2^2+q^2|\{j\in\mathcal{I}_{21}:\mathbf{v}_j\ge0\}|+|\{j\in\mathcal{I}_{22}:\mathbf{v}_j\le0\}|=\|\mathbf{u}\|_2^2.$$

Second, write $\langle\mathbf{v},\mathbf{u}\rangle$ by Equation (1) as

$$\langle\mathbf{v},\mathbf{u}\rangle=\langle\mathbf{v},\mathbf{u}^*\rangle+q\left(-\sum_{j\in\mathcal{I}_1}\mathbf{v}_j+\sum_{j\in\mathcal{I}_{21}\cup\mathcal{I}_{32}}\mathbf{v}_j\right)+\left(\sum_{j\in\mathcal{I}_3}\mathbf{v}_j+\sum_{j\in\mathcal{I}_{12}\cup\mathcal{I}_{22}}(-\mathbf{v}_j)\right). \tag{3}$$

Notice that some terms of the summation in Equation (3) are non-negative, in particular

$$\sum_{j\in\mathcal{I}_{21},\mathbf{v}_j\ge0}q\mathbf{v}_j+\sum_{j\in\mathcal{I}_{22},\mathbf{v}_j\le0}(-\mathbf{v}_j)\ge0.$$

Therefore, by letting $\mathcal{I}_{21}^-=\{i\in\mathcal{I}_{21},\mathbf{v}_i<0\}$ and $\mathcal{I}_{22}^+=\{i\in\mathcal{I}_{22},\mathbf{v}_i>0\}$, we have

$$\begin{aligned}\langle\mathbf{v},\mathbf{u}\rangle&\ge\langle\mathbf{v},\mathbf{u}^*\rangle+q\left(-\sum_{j\in\mathcal{I}_1}\mathbf{v}_j+\sum_{j\in\mathcal{I}_{32}\cup\mathcal{I}_{21}^-}\mathbf{v}_j\right)+\left(\sum_{j\in\mathcal{I}_3}\mathbf{v}_j-\sum_{j\in\mathcal{I}_{12}\cup\mathcal{I}_{22}^+}\mathbf{v}_j\right) &(4)\\&\ge q\sum_{j\in\mathcal{I}_{32}\cup\mathcal{I}_{21}^-}\mathbf{v}_j-\sum_{j\in\mathcal{I}_{12}\cup\mathcal{I}_{22}^+}\mathbf{v}_j &(5)\\&\ge-\left(q\sum_{j=1}^{|\mathcal{I}_{32}\cup\mathcal{I}_{21}^-|}t'_j-\sum_{j=|\mathcal{I}_{12}\cup\mathcal{I}_{22}^+|+1}^{n}t'_j\right)=-\langle\mathbf{v},\tilde{\mathbf{u}}\rangle,\end{aligned}$$

where Inequalities (4) and (5) hold because $\mathcal{I}_1\subseteq[k_1^*]$ and $\mathcal{I}_3\subseteq[k_2^*,\cdots,n]$.

Hence, we have $0\ge\cos\theta(\mathbf{v},\mathbf{u})\ge\cos\theta(\mathbf{v},\tilde{\mathbf{u}})$, which is equivalent to $\sin\theta(\mathbf{v},\mathbf{u})\ge\sin\theta(\mathbf{v},\tilde{\mathbf{u}})$ due to the non-negativity of $\sin\theta(\cdot,\cdot)$. □

B Proof of Theorem 1

Proof: Let $\mathbf{u}$ be the random variable defined in Section 6.2, such that $\mathbf{u}_i \sim q\,\text{Bernoulli}(|\mathbf{v}_i|/q)$ for positive coordinates $\mathbf{v}_i > 0$, $\mathbf{u}_i \sim (-1)\text{Bernoulli}(|\mathbf{v}_i|)$ for negative coordinates $\mathbf{v}_i < 0$, and $\mathbf{u}_i = 0$ if $\mathbf{v}_i = 0$. For convenience, we define

$$g(x) = \begin{cases} q, & x > 0 \\ -1, & x < 0 \\ 0, & x = 0. \end{cases}$$

We are interested in analyzing the expectation of $\mathbf{u}^T A\mathbf{u}/\mathbf{u}^T\mathbf{u}$, which is given by

$$\begin{aligned} \mathbb{E}\left[\frac{\mathbf{u}^T A\mathbf{u}}{\mathbf{u}^T\mathbf{u}}\right] &= \sum_{(k_1,k_2):1\le k_1+k_2\le n} \mathbb{E}\left[\frac{\mathbf{u}^T A\mathbf{u}}{\mathbf{u}^T\mathbf{u}} \mid \mathbf{u}^T\mathbf{u} = qk_1 + k_2\right]\mathbb{P}(\mathbf{u}^T\mathbf{u} = qk_1 + k_2) \\ &= \sum_{(k_1,k_2):1\le k_1+k_2\le n} \frac{\mathbb{E}\left[\mathbf{u}^T A\mathbf{u} \mid \mathbf{u}^T\mathbf{u} = qk_1 + k_2\right]\mathbb{P}(\mathbf{u}^T\mathbf{u} = qk_1 + k_2)}{qk_1 + k_2}. \end{aligned} \tag{6}$$

The term $\mathbb{E}\left[\mathbf{u}^T A\mathbf{u} \mid \mathbf{u}^T\mathbf{u} = qk_1 + k_2\right]\mathbb{P}(\mathbf{u}^T\mathbf{u} = qk_1 + k_2)$ in Equation (6), can be written as

$$\sum_{i\neq j} A_{i,j} g(\mathbf{v}_i)g(\mathbf{v}_j)\mathbb{P}(\mathbf{u}_i = g(\mathbf{v}_i), \mathbf{u}_j = g(\mathbf{v}_j) \mid \mathbf{u}^T\mathbf{u} = qk_1 + k_2)\mathbb{P}(\mathbf{u}^T\mathbf{u} = qk_1 + k_2), \tag{7}$$

and using Bayes' theorem we can re-write Equation (7) as

$$\sum_{i\neq j} A_{i,j} g(\mathbf{v}_i)g(\mathbf{v}_j)\mathbb{P}(\mathbf{u}^T\mathbf{u} = qk_1{+}k_2 \mid \mathbf{u}_i = g(\mathbf{v}_i), \mathbf{u}_j = g(\mathbf{v}_j))\mathbb{P}(\mathbf{u}_i = g(\mathbf{v}_i), \mathbf{u}_j = g(\mathbf{v}_j)). \tag{8}$$

By Equations (6) and (8) and since $g(\mathbf{v}_i)g(\mathbf{v}_j)\mathbb{P}(\mathbf{u}_i = g(\mathbf{v}_i), \mathbf{u}_j = g(\mathbf{v}_j)) = \mathbf{v}_i\mathbf{v}_j$, we have

$$\begin{aligned} &\sum_{(k_1,k_2):1\le k_1+k_2\le n} \frac{\sum_{i\neq j} A_{i,j}\mathbf{v}_i\mathbf{v}_j\mathbb{P}(\mathbf{u}^T\mathbf{u} = qk_1 + k_2 \mid \mathbf{u}_i = g(\mathbf{v}_i), \mathbf{u}_j = g(\mathbf{v}_j))}{qk_1 + k_2} \\ &= \sum_{i\neq j} A_{i,j}\mathbf{v}_i\mathbf{v}_j \sum_{(k_1,k_2):1\le k_1+k_2\le n} \frac{\mathbb{P}(\mathbf{u}^T\mathbf{u} = qk_1 + k_2 \mid \mathbf{u}_i = g(\mathbf{v}_i), \mathbf{u}_j = g(\mathbf{v}_j))}{qk_1 + k_2} \\ &= \sum_{i\neq j} A_{i,j}\mathbf{v}_i\mathbf{v}_j\mathbb{E}\left[\frac{1}{\mathbf{u}^T\mathbf{u}} \mid \mathbf{u}_i = g(\mathbf{v}_i), \mathbf{u}_j = g(\mathbf{v}_j)\right]. \end{aligned} \tag{9}$$

As the reciprocal function is convex, we apply Jensen's inequality to Equation (9) to obtain

$$\mathbb{E}\left[\frac{\mathbf{u}^T A\mathbf{u}}{\mathbf{u}^T\mathbf{u}}\right] \ge \frac{\sum_{i\neq j} A_{i,j}\mathbf{v}_i\mathbf{v}_j}{\mathbb{E}\left[\mathbf{u}^T\mathbf{u} \mid \mathbf{u}_i = g(\mathbf{v}_i), \mathbf{u}_j = g(\mathbf{v}_j))\right]}. \tag{10}$$

To estimate the denominator in Equation (10), we compute

$$\begin{aligned} \mathbb{E}\left[\mathbf{u}^T\mathbf{u} \mid \mathbf{u}_i = g(\mathbf{v}_i), \mathbf{u}_j = g(\mathbf{v}_j))\right] &= g(\mathbf{v}_i)^2 + g(\mathbf{v}_j)^2 + \sum_{k\neq i, k\neq j} g(\mathbf{v}_k)^2 \cdot \frac{|\mathbf{v}_k|}{|g(\mathbf{v}_k)|} \\ &\le \max\left(q\sqrt{n-2}, 2q^2 + q\frac{n-2}{\sqrt{n}}\right). \end{aligned} \tag{11}$$

Combining (10) and (12) we get

$$\mathbb{E}\left[\frac{\mathbf{u}^T A\mathbf{u}}{\mathbf{u}^T\mathbf{u}}\right] \ge \frac{\sum_{i\neq j} A_{i,j}\mathbf{v}_i\mathbf{v}_j}{\max\left(q\sqrt{n-2}, 2q^2 + q\frac{n-2}{\sqrt{n}}\right)} = \frac{\lambda_1(A)}{\max\left(q\sqrt{n-2}, 2q^2 + q\frac{n-2}{\sqrt{n}}\right)}. \tag{12}$$

Hence, the expected approximation ratio is

$$\mathcal{O}(q\sqrt{n})\mathbb{E}\left[\frac{\mathbf{u}^T A\mathbf{u}}{\mathbf{u}^T\mathbf{u}}\right] \ge \lambda_1(A) \ge OPT,$$

where OPT is the optimum of MAX-DRQ. □

C Proof of Lemma 2

Proof: Consider a graph $G = (V, E)$ consisting of $|V| = n = 2c+1$ nodes, for some $c \geq 1$, where $2c$ nodes form a negative clique and the extra node v is negatively connected to c of the nodes in the clique. Let A be the signed adjacency matrix of G. We will show the problem instance defined on G results in an optimal value of MAX-DRQ equal to $OPT = \mathcal{O}(1)$, while $\lambda_1(A)$ is $\Omega(\sqrt{n})$.

Any solution $\mathbf{u} \in \{0, -1, q\}^n$ to MAX-DRQ defines the two sets $S_p = \{i : \mathbf{u}_i = q\}$ and $S_n = \{i : \mathbf{u}_i = -1\}$. We claim that $\max_{\mathbf{u}\in\{0,-1,q\}^n} \mathbf{u}^T A\mathbf{u}/\mathbf{u}^T\mathbf{u} \leq 2$, and will show it by considering 3 cases:

<u>(Case 1) $v \notin S_p \cup S_n$:</u>

$$\begin{aligned}\frac{\mathbf{u}^T A\mathbf{u}}{\mathbf{u}^T\mathbf{u}} &= \frac{-q^2 \overbrace{|S_p|(|S_p|-1)}^{2|E(S_p)|} + q\overbrace{2|S_p||S_n|}^{2|E(S_p,S_n)|} - \overbrace{|S_n|(|S_n|-1)}^{2|E(S_n)|}}{q^2|S_p| + |S_n|} \\ &= \frac{-(q|S_p| - |S_n|)^2 + q^2|S_p| + |S_n|}{q^2|S_p| + |S_n|}. \end{aligned} \tag{13}$$

Let $r = q|S_p| - |S_n|$ and let $\epsilon = r/q|S_p| \leq 1$. Then, Equation (13) can be written as

$$\begin{aligned}\frac{\mathbf{u}^T A\mathbf{u}}{\mathbf{u}^T\mathbf{u}} &= \frac{q(q+1)|S_p| - r(r+1)}{|S_p|q(q+1) - r} \\ &= \frac{(q+1) + \frac{1}{4(q|S_p|)}}{(q+1)-\epsilon} - \frac{(r+\frac{1}{2})^2}{q|S_p|(q+1-\epsilon)} \\ &\leq \frac{(q+1) + \frac{1}{4(q|S_p|)}}{(q+1)-\epsilon} \\ &\leq \frac{q+2}{q} \leq 2 = \mathcal{O}(1).\end{aligned}$$

<u>(Case 2) $v \in S_p$:</u>

$$\begin{aligned}\mathbf{u}^T A\mathbf{u} =& -q^2\left(\underbrace{(|S_p|-1)(|S_p|-2)}_{2|E(S_p\setminus\{v\})|} + 2|E(\{v\}, S_p)|\right) \\ &+ q\left(\underbrace{2(|S_p|-1)|S_n|}_{2|E(S_p\setminus\{v\},S_n)|} + 2|E(\{v\}, S_n)|\right) - \underbrace{|S_n|(|S_n|-1)}_{2|E(S_n)|} \\ =& -\left(q(|S_p|-1) - |S_n|\right)^2 + |S_n| \\ &+ q^2(|S_p|-1) + 2q|E(\{v\}, S_n)| - 2q^2|E(\{v\}, S_p)| \\ \leq& -\left(q(|S_p|-1) - |S_n|\right)^2 + |S_n| + q^2(|S_p|-1) + 2q|S_n|. \end{aligned} \tag{14}$$

Let $r = q(|S_p|-1) - |S_n|$ and write Equation (14) as

$$\begin{aligned}\mathbf{u}^T A\mathbf{u} &= -(r-q)^2 + q(q+3)(|S_p|-1) + q^2 - r \\ &\leq q(q+3)(|S_p|-1) + q^2 - r. \end{aligned} \tag{15}$$

By (15) and letting $\epsilon = r/q(|S_p|-1) \leq 1$, we have

$$\begin{aligned}\frac{\mathbf{u}^T A\mathbf{u}}{\mathbf{u}^T\mathbf{u}} &\leq \frac{q(q+3)(|S_p|-1) + q^2 - r}{q(q+1)(|S_p|-1) + (q^2 - r)} \\ &= 1 + \frac{2}{(q+1) + (q/(|S_p|-1) - \epsilon)} \leq 2 = \mathcal{O}(1).\end{aligned}$$

(Case 3) $v \in S_n$:

$$\begin{aligned}
\mathbf{u}^T A\mathbf{u} = q^2|S_p|(|S_p|-1) + q\left(\underbrace{|S_p|(|S_n|-1)}_{2|E(S_n\setminus\{v\},S_p)|} + 2|E(\{v\},S_p)|\right) \\
- \left(\underbrace{(|S_n|-1)(|S_n|-2)}_{2|E(S_n\setminus\{v\})|} + 2|E(\{v\},S_n)|\right) \\
= -\left(q|S_p| - (|S_n|-1)\right)^2 + q^2|S_p| + |S_n| - 1 \\
+ 2q|E(\{v\},S_p)| - 2|E(\{v\},S_n)| \\
\leq -\left(q|S_p| - (|S_n|-1)\right)^2 + q^2|S_p| + |S_n| - 1 + 2q|S_p|. \qquad (16)
\end{aligned}$$

Let $r = q|S_p| - (|S_n|-1)$ and write Inequality (16) as

$$\begin{aligned}
\mathbf{u}^T A\mathbf{u} &= -(r+\frac{1}{2})^2 + q(q+3)|S_p| + \frac{1}{4} \\
&\leq q(q+3)|S_p| + \frac{1}{4}. \qquad (17)
\end{aligned}$$

By (17) and letting $\epsilon = (r+1)/q|S_p| \leq 1$, we have

$$\begin{aligned}
\frac{\mathbf{u}^T A\mathbf{u}}{\mathbf{u}^T\mathbf{u}} &\leq \frac{q(q+3)|S_p| + \frac{1}{4}}{q(q+1)|S_p| - (r+1)} \\
&= 1 + \frac{2 + 1/(4q|S_p|) + \epsilon}{(q+1) - \epsilon} \\
&\leq 2 = \mathcal{O}(1).
\end{aligned}$$

Therefore, we know that the optimal solution OPT of MAX-DRQ is $\mathcal{O}(1)$. However, consider a vector $\mathbf{x} \in \mathbb{R}^n$ such that

$$\mathbf{x} = \left[\sqrt{\frac{n+1}{2n}}, \underbrace{\frac{1}{\sqrt{2n}}, \cdots, \frac{1}{\sqrt{2n}}}_{c \text{ entries}}, \underbrace{\frac{-1}{\sqrt{2n}}, \cdots, \frac{-1}{\sqrt{2n}}}_{c \text{ entries}}\right], \qquad (18)$$

where the first entry of $\mathbf{x}$ corresponds to v. Then, the vector $\mathbf{x}$ defined in Equation (18) gives

$$\begin{aligned}
\frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}} &= \frac{\sqrt{n+1}(n-1)}{2n} + \frac{n-1}{2n} \\
&= \frac{\sqrt{n+1}+1}{2} - \frac{\sqrt{n+1}+1}{2n} \\
&= \Omega(\sqrt{n}).
\end{aligned}$$

As $\lambda_1(A) = \max_{\mathbf{x}\in\mathbb{R}^n\setminus\{0\}} \mathbf{x}^T A\mathbf{x}/\mathbf{x}^T\mathbf{x}$, we have shown $\lambda_1(A) \geq OPT \cdot \Omega(\sqrt{n})$. $\square$

D Experiment Results

D.1 Dataset

`WoW-EP8` [1] is the interaction network of authors in the 8th legislature of the EU Parliament, where edge signs indicate if two authors are collaborative or competitive to each other. `Bitcoin` [4] is the trust-distrust network of users trading on the Bitcoin OTC platform. `WikiVot` [4] collects the positive and negative votes for electing Wikipedia admins. `Referendum` [3] collects the tweets about the Italian constitutional referendum in 2016, and edge signs indicate if two users are classified to have the same stance or not. `Slashdot` [4] is a friend-foe network collected from the Slashdot Zoo feature. `WikiCon` [2] collects the positive and negative iterations of users editing the English Wikipedia. `Epinions` [4] is the trust-distrust network of users on the online social network Epinions. `WikiPol` [5] is the interaction network of users who have edited the English Wikipedia pages about politics.

D.2 Execution Time

Table 2: Running times for the results shown in Table 1. All times are shown in seconds. Dashes indicate that a method cannot finish execution due to memory limit exceeded.

		`WoW-EP8`	`Bitcoin`	`WikiVot`	`Referendum`	`Slashdot`	`WikiCon`	`Epinions`	`WikiPol`
	$\lvert V\rvert$	790	5 881	7 115	10 884	82 140	116 717	131 580	138 587
	$\lvert E\rvert$	116 009	21 492	100 693	251 406	500 481	2 026 646	711 210	715 883
	$\lvert E_-\rvert/\lvert E\rvert$	0.2	0.2	0.2	0.1	0.2	0.6	0.2	0.1
$k=2$	SCG-MA	2	1	2	4	10	217	109	25
	SCG-MO	2	1	2	4	11	70	94	15
	SCG-B	13	9	21	44	693	3 584	1 906	1 624
	SCG-R	4	3	6	17	70	485	37	217
	KOCG	3	11	16	25	1 243	3 269	3 208	3 506
	BNC-k	2	1	2	4	—	—	—	—
	BNC-$(k+1)$	2	1	2	4	—	—	—	—
	SPONGE-k	2	5	3	4	—	—	—	—
	SPONGE-$(k+1)$	2	11	4	9	—	—	—	—
$k=6$	SCG-MA	3	1	6	16	75	394	132	136
	SCG-MO	3	1	6	18	74	229	107	139
	SCG-B	17	29	78	201	3 280	10 637	5 455	5 714
	SCG-R	3	5	9	21	118	415	219	892
	KOCG	1	5	8	14	690	1 837	1 845	1 724
	BNC-k	2	1	2	4	—	—	—	—
	BNC-$(k+1)$	2	1	2	4	—	—	—	—
	SPONGE-k	2	7	6	20	—	—	—	—
	SPONGE-$(k+1)$	2	5	4	26	—	—	—	—

D.3 Detected Group Sizes

Figure 2, extracted from the `Referendum` dataset, shows the typical distribution of the group sizes for all the comparison methods. This pattern is similar to all other datasets except `WoW-EP8`. That is, SCG-MA, SCG-MO, and SCG-R return the largest groups while KOCG-top-1, BNC-$(k+1)$, and SPONGE-$(k+1)$ return the smallest groups.

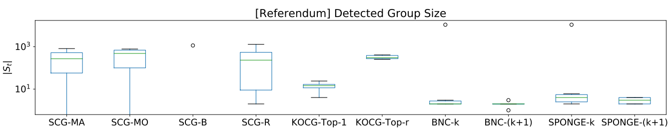


Figure 2: The typical group size distribution on all the datasets except `WoW-EP8` when $k=6$.

On the other hand, `WoW-EP8` shows a different group-size distribution, which is shown in Figure 3. All SCG methods and BNC-k find one giant group. By checking the polarity (Table 1 in main paper), their scores are high, so this probably suggests there exists a giant conflicting group in the network.

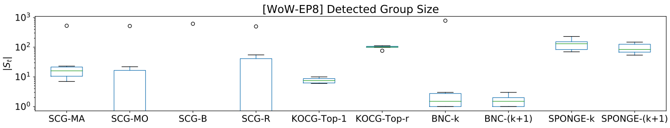


Figure 3: The group size distribution on WoW-EP8 when $k = 6$.

D.4 Deciding k

We present a heuristic similar to Elbow Method [6] to decide k, which consists of the following steps:

1. Run SCG multiple times with different k.
2. Draw a DRQ-Plot, where the Discrete Rayleigh Quotient (DRQ) values in each run are sorted, and then plot the i-th largest DRQ value at the i-th location.
3. Decide k to be one of the "knees" of the curve.

The reason why the heuristic works is that, if there exist conflicting groups and the noise-level is not too high, then the leading eigenvector should be indicative of the true conflicting groups and have large DRQ values in the first $k - 1$ iterations, while the leading eigenvector only captures noise structures and has low DRQ value after the k-th iteration. Therefore, it is expected to see knees of the curve at the $(k - 1)$-th iteration.

First, we evaluate the heuristic using m-SSBM under the same setting ($k = 6$, $\ell = 100$, and $n = 2\,000$) by varying $\eta = 0 : 0.1 : 0.6$. The result of detecting the conflicting groups by SCG-MA is depicted in Figure 4. As expected, the most prominent knee is at the 5-th iteration when the noise-level is not too high ($\eta \leq 0.3$). As the noise-level increases ($\eta \geq 0.4$), the knee at the 5-th iteration becomes less obvious and some artificial knees that fit the random noise emerge.

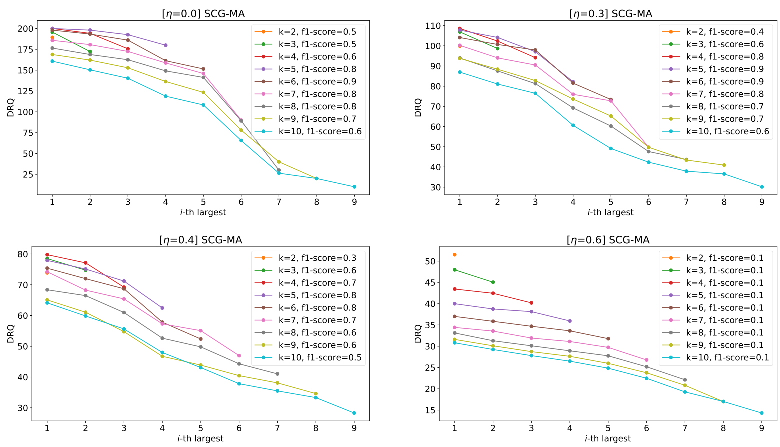


Figure 4: Run SCG-MA with different k on networks generated by m-SSBM ($k = 6$, $\ell = 100$, and $n = 2\,000$). Each setting is repeated 20 times and reported the average.

Finally, we use the heuristic on the real-world datasets to decide k and show the result in Figure 5. Our analysis suggests that Referendum has 4 conflicting groups, because the most prominent knee appears at the 3-th iteration, while on `Epinions`, there are two prominent knees at the 3rd and the 4-th iterations, so there are probably 4 or 5 conflicting groups in the network.

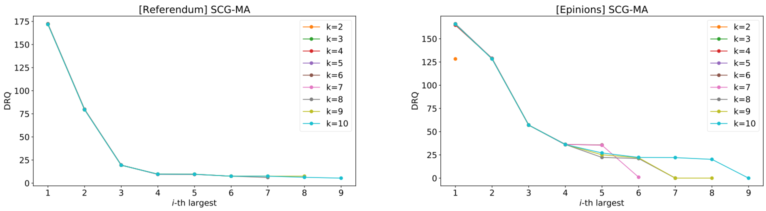


Figure 5: Run SCG-MA Real-world networks with different k.

References

[1] Victor Kristof, Matthias Grossglauser, and Patrick Thiran. War of words: The competitive dynamics of legislative processes. *Proc. of The Web Conference*, 2020.

[2] Jérôme Kunegis. Konect: the koblenz network collection. In *Proc of WWW*, 2013.

[3] Mirko Lai, Viviana Patti, Giancarlo Ruffo, and Paolo Rosso. Stance evolution and twitter interactions in an italian political debate. In *Proc. of NLDB*. Springer, 2018.

[4] Jure Leskovec and Andrej Krevl. SNAP Datasets: Stanford large network dataset collection. `http://snap.stanford.edu/data`, 2014.

[5] Silviu Maniu, Talel Abdessalem, and Bogdan Cautis. Casting a web of trust over wikipedia: an interaction-based approach. In *Proc. of WWW Companion*, 2011.

[6] Robert L Thorndike. Who belongs in the family? *Psychometrika*, 1953.

Appendix B

Improved analysis of RSVD for top-eigenvector approximation

Figure: A cat looking into a mirror that randomly projects the cat's image.

We analyze the approximated leading eigenvector $\hat{\mathbf{u}}$ returned by Randomized SVD (`RSVD`) for a given symmetric matrix $\mathbf{A}$, and study its theoretical guarantee with respect to the ratio $R(\hat{\mathbf{u}}) = \frac{\hat{\mathbf{u}}^T \mathbf{A} \hat{\mathbf{u}}}{\lambda_1}$, where λ_1 is the leading eigenvalue. By relating to the random projection lemma, we sharpen the theoretical guarantee of $R(\hat{\mathbf{u}})$ using `RSVD` with any number of iterations.

Improved analysis of randomized SVD for top-eigenvector approximation

Ruo-Chun Tzeng[1] **Po-An Wang**[2] **Florian Adriaens**[1]

Aristides Gionis[1] **Chi-Jen Lu**[3]

[1]Division of Theoretical Computer Science
[2]Division of Decision and Control Systems
KTH Royal Institute of Technology, Sweden

[3]Institute of Information Science
Academia Sinica, Taiwan

Abstract

Computing the top eigenvectors of a matrix is a problem of fundamental interest to various fields. While the majority of the literature has focused on analyzing the reconstruction error of low-rank matrices associated with the retrieved eigenvectors, in many applications one is interested in finding one vector with high Rayleigh quotient. In this paper we study the problem of approximating the top-eigenvector. Given a symmetric matrix $\mathbf{A}$ with largest eigenvalue λ_1, our goal is to find a vector $\hat{\mathbf{u}}$ that approximates the leading eigenvector $\mathbf{u}_1$ with high accuracy, as measured by the ratio $R(\hat{\mathbf{u}}) = \lambda_1^{-1}\hat{\mathbf{u}}^T\mathbf{A}\hat{\mathbf{u}}/\hat{\mathbf{u}}^T\hat{\mathbf{u}}$. We present a novel analysis of the randomized SVD algorithm of Halko et al. (2011b) and derive tight bounds in many cases of interest. Notably, this is the first work that provides non-trivial bounds for approximating the ratio $R(\hat{\mathbf{u}})$ using randomized SVD with any number of iterations. Our theoretical analysis is complemented with a thorough experimental study that confirms the efficiency and accuracy of the method.

1 INTRODUCTION

Spectral methods, which typically rely on computing the leading eigenvectors of an appropriately-designed matrix, have been shown to provide high-quality solutions to a variety of problems in the fields of data analysis, optimization, clustering and learning (Kannan and Vempala, 2009). From a computational perspective, randomized approaches for spectral methods often give good estimates of leading eigenvectors and low-rank structures, opening up the possibility of dealing with truly large datasets (Halko et al., 2011a).

Proceedings of the 25[th] International Conference on Artificial Intelligence and Statistics (AISTATS) 2022, Valencia, Spain. PMLR: Volume 151.

In this paper, we study the problem of approximating the leading eigenvector of a matrix while using a small amount of memory and making a limited number of passes over the input matrix. More concretely, given a symmetric matrix $\mathbf{A}$ with largest eigenvalue λ_1, our goal is to find a vector $\hat{\mathbf{u}}$ that maximizes the ratio

$$R(\hat{\mathbf{u}}) = \lambda_1^{-1}\frac{\hat{\mathbf{u}}^T\mathbf{A}\hat{\mathbf{u}}}{\hat{\mathbf{u}}^T\hat{\mathbf{u}}}. \tag{1}$$

Note that since λ_1 is fixed given $\mathbf{A}$, it can be omitted from the definition of R; it is used only for convenience, to ensure that $R \leq 1$. Often, in different applications, in addition to having to select which matrix $\mathbf{A}$ to use, it is also required that $\hat{\mathbf{u}} \in \mathcal{T} \subseteq \mathbb{R}^n$, where $\mathcal{T}$ is typically a discrete subspace of $\mathbb{R}^n$. A common strategy in this case, is to first compute an approximation of the leading eigenvector in $\mathbb{R}^n$ and then "round" the solution in $\mathcal{T}$. Below we outline some prominent examples of this scheme.

(1) The most direct example is PCA, where $\mathbf{A}$ is the covariance matrix (Jolliffe, 1986); in this case $\mathcal{T} = \mathbb{R}^n$, and no rounding is required; **(2)** In the community-detection problem we can partition a network into two communities (and then recursively find more communities) by maximizing modularity (Newman, 2006), which can be mapped to our setting by taking $\mathbf{A}$ to be the modularity matrix and $\mathcal{T} = \{\pm 1\}^n$; **(3)** The problem of finding k conflicting groups in signed networks can be formulated by taking $\mathbf{A}$ to be the adjacency matrix of the signed network and $\mathcal{T} = \{0, -1, \ell\}^n$, for $\ell \in [k-1]$ (Bonchi et al., 2019; Tzeng et al., 2020); **(4)** For the fair densest subgraph, Anagnostopoulos et al. (2020) consider $\mathcal{T} = \{0, 1\}^n$ and obtain $\mathbf{A}$ after projecting the adjacency matrix onto the subspace orthogonal to a given fairness labeling $z \in \{\pm 1\}^n$; **(5)** In several other applications, a solution to our problem is used as an intermediate step in the proposed method (Abdullah et al., 2014; Hopkins et al., 2016; Allen-Zhu and Li, 2016; Silva et al., 2018).

Despite numerous pass-efficient algorithms for computing top eigenvectors proposed in the literature, prior analysis of $R(\hat{\mathbf{u}})$ have strong limitations when applied in practice. The main shortcoming is that most works provide *additive bounds* and require $\Omega(\ln n)$ passes to be meaningful (Simchowitz et al., 2018), whereas a smaller number of passes (constant or even a single pass) is critical in practical settings. It is unclear in the state-of-the-art whether $\Omega(\ln n)$ passes is necessary for previous methods, or whether such a bound is an artifact of the analysis.

In this paper we demonstrate that the requirement of $\Omega(\ln n)$ passes in the analysis of prior works is artificial. We show this by giving a *multiplicative bound* for $R(\hat{\mathbf{u}})$ achieved by the randomized SVD method (RSVD) of Halko et al. (2011b), which is one of the most prominent and widely-implemented pass-efficient algorithms (Pedregosa et al., 2011; Řehůřek and Sojka, 2010; Corporation, 2021; Erichson et al., 2019; Terray and Pinsard, 2021; Liutkus, 2021)

Our analysis shows that for any positive semidefinite matrix, RSVD returns with high probability a vector $\hat{\mathbf{u}}$ satisfying $R(\hat{\mathbf{u}}) = \Omega\left((d/n)^{1/(2q+1)}\right)$ after $q \in \mathbb{N}$ iterations (Theorem 1), using $\mathcal{O}(dn)$ space for $d \in \mathbb{N}$, where typically $d \ll n$ (e.g., $d = \mathcal{O}(\ln n)$). Theorem 2 shows that our analysis is tight. Notably, our analysis subsumes the guarantee by prior works in the regime of $\Omega(\ln n)$ passes (Remark 1), and to the best of our knowledge, provides the first non-trivial guarantee of $R(\hat{\mathbf{u}})$ in the literature of pass-efficient algorithms for $o(\ln n)$ passes. Moreover, we show that under some natural conditions satisfied by real-world datasets, it is even possible to achieve $R(\hat{\mathbf{u}}) = \Omega(1)$ with a single pass (Remark 2).

Our core technical argument is a reduction from the optimization problem of maximizing R over a random subspace to the problem of estimating the projection length of a vector onto a random subspace. By using our technique, we derive the first non-trivial guarantee of $R(\hat{\mathbf{u}})$ for any number of passes for indefinite matrices (Theorem 4), under mild conditions (Assumption 1).

In addition, we propose an extension of the RSVD method, called RandSum, by using a random matrix sampled from Bernoulli(p) with mean $p \in (0, 1)$. While such a random matrix is rarely used in the literature of random projections, we show that there exist applications (Bonchi et al., 2019; Tzeng et al., 2020) especially suitable for this technique, and we show several properties of such a random matrix, which may be of independent interest.

2 RELATED WORK

For lack of space, we provide a brief overview of the existing literature, focusing on the most relevant works to our paper. For a general introduction on pass-efficient algorithms for matrix approximations, we refer the reader to Mahoney et al. (2011); Woodruff et al. (2014); Martinsson and Tropp (2020).

The study of $R(\hat{\mathbf{u}})$ for pass-efficient algorithms can be dated back to Kuczyński and Woźniakowski (1992) who analyzed two classical methods: the power method and the Lanczos method with random start. For any positive semidefinite matrix, they showed that the power method (respectively, Lanczos method) with random start, after $q \geq 2$ iterations returns an approximated top-eigenvector $\hat{\mathbf{u}}$ with $\mathbb{E}\left[R(\hat{\mathbf{u}})\right] \geq 1 - 0.871\frac{\ln n}{q-1}$ (respectively, $\mathbb{E}\left[R(\hat{\mathbf{u}})\right] \geq 1 - 2.575(\frac{\ln n}{q-1})^2$).

The aforementioned methods are generalized to randomized SVD (Halko et al., 2011b) and block-Krylov methods (Musco and Musco, 2015), and a similar additive analysis of $R(\hat{\mathbf{u}})$ by Musco and Musco (2015) showed that for any positive semidefinite matrix, RSVD (respectively, randomized block-Krylov method) using $\mathcal{O}(nd)$ space (respectively, $\mathcal{O}(ndq)$ space) and after q iterations, returns an approximate top-eigenvector $\hat{\mathbf{u}}$ with $R(\hat{\mathbf{u}}) \geq 1 - \mathcal{O}(\frac{\ln n}{q})$ (respectively, $R(\hat{\mathbf{u}}) \geq 1 - \mathcal{O}((\frac{\ln n}{q})^2)$), with probability at least $1 - e^{-\Omega(d)}$.[1]

The analysis of the previous works (Kuczyński and Woźniakowski, 1992; Musco and Musco, 2015) is tight, as shown by Simchowitz et al. (2018) for a class of methods (which include RSVD and block Krylov), which with high probability fail to find a vector $\hat{\mathbf{u}}$ with $R(\hat{\mathbf{u}}) \geq 23/24$ within $q = \mathcal{O}(\ln n)$ passes.

In the aforementioned works there are two limitations. First, the bounds of Kuczyński and Woźniakowski (1992) and Musco and Musco (2015) are additive, and unfortunately require $\Omega(\ln n)$ passes to be meaningful. In contrast, our analysis provides a multiplicative bound for $R(\hat{\mathbf{u}})$ and offers non-trivial guarantees for any number of passes. Second, the applicability of the methods of Kuczyński and Woźniakowski (1992) and Musco and Musco (2015) is limited to positive semidefinite matrices. Instead, we provide a sharp analysis of randomized SVD for positive semidefinite matrices and

[1] Musco and Musco (2015) showed that the aforementioned results hold with constant probability, which could be improved to hold with probability $1 - e^{-\Omega(d)}$ by using stronger concentration results (Rudelson and Vershynin, 2010) in their proofs of Lemma 4 and Lemma 9.

show that our proof techniques generalize to indefinite matrices under mild conditions.

To complement our study, we briefly compare the measure $R(\hat{\mathbf{u}})$ with other classical metrics. Note here that, even though it is possible to covert an error guarantee for classical metrics (Xu et al., 2018; Drineas et al., 2018; Ghashami et al., 2016; Chen et al., 2017; Musco and Woodruff, 2017; Huang, 2018) into a lower bound for $R(\hat{\mathbf{u}})$ by matrix perturbation theory (Stewart and Guang Sun, 1990; Yu et al., 2015), the resulting bound is additive and depends on the eigengap. We also note that classical metrics typically compare the approximation $\hat{\mathbf{u}}$ to the top-eigenvector $\mathbf{u}_1$ of $\mathbf{A}$, however, such a comparison is not meaningful in our setting as small distance between $\hat{\mathbf{u}}$ and $\mathbf{u}_1$[2] is a sufficient but not necessary condition for having large $R(\hat{\mathbf{u}})$.

3 PRELIMINARIES

Let $\mathbb{N}$ be the set of natural numbers excluding 0. Let $\mathbb{R}$ be the set of real numbers, $\mathbb{S}^{m-1} = \{\mathbf{x} \in \mathbb{R}^m : \mathbf{x}^T\mathbf{x} = 1\}$, and $[m] = \{1, \dots, m\}$. Let range($\mathbf{M}$) denote the column space of matrix $\mathbf{M}$, and $\|\cdot\|_F$ and $\|\cdot\|_2$ denote the Frobenius norm and the spectral norm, respectively. For a square matrix $\mathbf{M}$, let $\lambda_i(\mathbf{M})$ be its i-th largest eigenvalue and $\mathbf{u}_i(\mathbf{M})$ the corresponding eigenvector, and let $\sigma_i(\mathbf{M})$ be the i-th largest singular value. In all subsequent sections, we use boldface $\mathbf{A}$ to denote the input matrix, and abbreviate $\lambda_i = \lambda_i(\mathbf{A})$, $\mathbf{u}_i = \mathbf{u}_i(\mathbf{A})$, and $\sigma_i = \sigma_i(\mathbf{A})$. We use $\langle \cdot, \cdot \rangle$ to denote the vector inner product. Finally, we use $\mathbf{1}_n = [1, \dots, 1]^T$ to denote the n-dimensional vector of all 1's and $\mathbf{0}_n = [0, \dots, 0]^T$ to denote the n-dimensional vector of all 0's.

For simplicity, we assume that the input matrix $\mathbf{A}$ is real-valued and symmetric, with $\lambda_1 > 0$.

Definition 1 (Vector projection onto subspace)**.** *Let $\mathbf{v} \in \mathbb{R}^n$ be a nonzero vector and $\mathcal{X} \subseteq \mathbb{R}^n$ be a non-empty subspace. The projection length of $\mathbf{v}$ onto $\mathcal{X}$ is given by $\cos\theta(\mathbf{v}, \mathcal{X})$, where*

$$\theta(\mathbf{v}, \mathcal{X}) = \cos^{-1}\left(\max_{\mathbf{x}\in\mathcal{X}} \frac{\langle \mathbf{v}, \mathbf{x}\rangle}{\|\mathbf{v}\|_2 \|\mathbf{x}\|_2}\right)$$

is the projection angle. For a matrix $\mathbf{X}$, we use $\theta(\mathbf{v}, \mathbf{X})$ to denote the projection angle of $\mathbf{v}$ onto the range of $\mathbf{X}$.

It is well-known that projecting any vector $\mathbf{v} \in \mathbb{R}^n$ onto the range($\mathbf{S}$) of a random matrix $\mathbf{S} \sim \mathcal{N}(0,1)^{n\times d}$ results in $\cos^2\theta(\mathbf{v}, \mathbf{S}) \approx d/n$ with high probability.

Lemma 1. *(Hardt and Price, 2014) Let $\mathbf{v} \in \mathbb{R}^n$ be a nonzero vector and $\mathbf{S} \sim \mathcal{N}(0,1)^{n\times d}$, where $n, d \in \mathbb{N}$ and $n \geq d$. Then,*

$$\cos^2\theta(\mathbf{v}, \mathbf{S}) = \Theta\left(\frac{d}{n}\right),$$

with probability at least $1 - e^{-\Omega(d)}$.

[2]More precisely, the distance between $\hat{\mathbf{u}}$ and the eigenspace associated with the largest eigenvalue λ_1 of $\mathbf{A}$.

Algorithm 1: $\mathsf{RSVD}(\mathbf{A}, \mathcal{D}, q, d)$

$\mathbf{Y} \leftarrow \mathbf{A}^q\mathbf{S}$ where $\mathbf{S} \sim \mathcal{D}$;
$\mathbf{Y} = \mathbf{QR}$;
$\mathbf{B} \leftarrow \mathbf{Q}^T\mathbf{AQ}$;
$\hat{\mathbf{u}} = \mathbf{Q}\,\mathbf{u}_1(\mathbf{B})$;
return $\hat{\mathbf{u}}$;

For completeness, we provide the proof of Lemma 1 in Appendix A.1. The proof idea is to observe that $\frac{\|\mathbf{S}^T\mathbf{v}\|_2}{\sigma_1(\mathbf{S})} \leq \cos\theta(\mathbf{v}, \mathbf{S}) \leq \frac{\|\mathbf{S}^T\mathbf{v}\|_2}{\sigma_d(\mathbf{S})}$ and use the concentration of the extreme singular values of a Gaussian random matrix.

More generally, Lemma 1 holds for any random matrix $\mathbf{S}$ whose range is uniformly distributed with respect to the Haar measure on Grassmannian $\mathcal{G}_{n,d}$ of all the d-dimensional subspaces of $\mathbb{R}^n$, written as range($\mathbf{S}$) $\sim$ Uniform($\mathcal{G}_{n,d}$). The reader may refer to Achlioptas (2001) and Halko et al. (2011b) for other choices of $\mathbf{S}$ and Vershynin (2018, Section 5) for a general introduction to this phenomenon.

4 RANDOMIZED SVD

We briefly review the following variant of the randomized SVD (RSVD) algorithm, as proposed by Halko et al. (2011b), and shown in Algorithm 1. The algorithm returns an estimate $\hat{\mathbf{u}}$ of the leading eigenvector $\mathbf{u}_1$ of the input matrix $\mathbf{A}$. It uses $\mathcal{O}(dn)$ space and requires $q+1$ passes over the matrix $\mathbf{A}$, where $q \in \mathbb{N}$.[3] The distribution $\mathcal{D}$ is over $\mathbb{R}^{n\times d}$, and one particular instance of the algorithm sets $\mathcal{D} = \mathcal{N}(0,1)^{n\times d}$. The algorithm begins with a random projection $\mathbf{Y} = \mathbf{A}^q\mathbf{S}$. The eigenvectors of $\mathbf{A}^q$ are the same as $\mathbf{A}$, but the eigenvalues of $\mathbf{A}^q$ have much stronger decay. Thus intuitively, by taking powers of the input matrix, the relative weight of the eigenvectors associated with the small eigenvalues is reduced, which is helpful in the basis identification for input matrices whose eigenvalues decay slowly. After projecting, the algorithm efficiently approximates the top-eigenvector of $\mathbf{A}$ by

$$\hat{\mathbf{u}} \in \operatorname{argmax}\{\mathbf{v}^T\mathbf{A}\mathbf{v} : \mathbf{v} \in \operatorname{range}(\mathbf{Y}) \cap \mathbb{S}^{n-1}\}. \tag{2}$$

[3]More precisely, RSVD requires q passes when $d = 1$ and $q + 1$ passes when $d > 1$ as there is no need to compute $\mathbf{u}_1(B)$ when $d = 1$.

Indeed, any $\mathbf{v} \in \text{range}(\mathbf{Y})$ of unit length can be written as $\mathbf{v} = \mathbf{Q}\mathbf{a}$ for some $\mathbf{a} \in \mathbb{S}^{d-1}$, where $\mathbf{Q}$ is an $n \times d$ orthonormal basis given by a QR decomposition of $\mathbf{Y}$. So it follows that

$$\max_{\mathbf{v}\in\text{range}(\mathbf{Y})\cap\mathbb{S}^{n-1}} \mathbf{v}^T\mathbf{A}\mathbf{v} = \max_{\mathbf{a}\in\mathbb{S}^{d-1}} \mathbf{a}^T\mathbf{B}\mathbf{a} = \lambda_1(\mathbf{B}).$$

Thus, the vector $\hat{\mathbf{u}} = \mathbf{Q}\mathbf{u}_1(\mathbf{B})$ maximizes expression (2), and $\mathbf{u}_1(\mathbf{B})$ can be efficiently computed as the matrix $\mathbf{B}$ is of dimension $d \times d$.

4.1 Analysis of RSVD

We now derive lower and upper bounds for $R(\hat{\mathbf{u}})$, where $\hat{\mathbf{u}}$ is the output of Algorithm 1, and $R(\mathbf{v}) = \lambda_1^{-1}\frac{\mathbf{v}^T\mathbf{A}\mathbf{v}}{\mathbf{v}^T\mathbf{v}}$ is defined for any nonzero vector $\mathbf{v} \in \mathbb{R}^n$. Note that due to expression (2), $\hat{\mathbf{u}}$ maximizes R over the column space range$(\mathbf{Y})$ of $\mathbf{Y}$. Since range$(\mathbf{Y}) = \{\mathbf{Y}\mathbf{a} : \mathbf{a} \in \mathbb{R}^d\}$, we can rewrite $R(\hat{\mathbf{u}})$ as

$$R(\hat{\mathbf{u}}) = \max_{\mathbf{v}\in\text{range}(\mathbf{Y})\setminus\{\mathbf{0}_n\}} R(\mathbf{v}) = \max_{\mathbf{a}\in\mathbb{S}^{d-1}} R(\mathbf{Y}\mathbf{a}),$$

where the latter equality follows from the scale invariance of R. For notational convenience, we denote $R_{\mathbf{a}} = R(\mathbf{Y}\mathbf{a})$. After substituting $\mathbf{Y} = \mathbf{A}^q\mathbf{S}$ in the definition of R, we can evaluate $R_{\mathbf{a}}$ as

$$R_{\mathbf{a}} = \frac{1}{\lambda_1}\frac{(\mathbf{S}\mathbf{a})^T\mathbf{A}^{2q+1}(\mathbf{S}\mathbf{a})}{(\mathbf{S}\mathbf{a})^T\mathbf{A}^{2q}(\mathbf{S}\mathbf{a})}. \tag{3}$$

Since $\mathbf{A}$ is real and symmetric, it has a real-valued eigen-decomposition $\mathbf{A} = \sum_{i=1}^n \lambda_i\mathbf{u}_i\mathbf{u}_i^T$, with $\{\mathbf{u}_i\}_{i=1}^n$ being orthonormal. Hence $\mathbf{A}^k = \sum_{i=1}^n \lambda_i^k\mathbf{u}_i\mathbf{u}_i^T$, for any $k \in \mathbb{N}$, and we further expand Equation (3) as

$$R_{\mathbf{a}} = \frac{1}{\lambda_1}\frac{\sum_i \lambda_i^{2q+1}\langle\mathbf{S}^T\mathbf{u}_i,\mathbf{a}\rangle^2}{\sum_i \lambda_i^{2q}\langle\mathbf{S}^T\mathbf{u}_i,\mathbf{a}\rangle^2} = \frac{\sum_i \alpha_i^{2q+1}\langle\mathbf{S}^T\mathbf{u}_i,\mathbf{a}\rangle^2}{\sum_i \alpha_i^{2q}\langle\mathbf{S}^T\mathbf{u}_i,\mathbf{a}\rangle^2}, \tag{4}$$

where $\alpha_i = \lambda_i/\lambda_1$, for all $i \in [n]$. This is well-defined since $\lambda_1 > 0$. For our analysis of $R(\hat{\mathbf{u}}) = \max_{\mathbf{a}\in\mathbb{S}^{d-1}} R_{\mathbf{a}}$, we first consider the case when $\mathbf{A}$ is positive semi-definite (p.s.d.). The proof strategy and arguments serve as a building block for the indefinite case, discussed in Section 4.3.

4.2 Positive semidefinite matrices

Our first result, is a guarantee on the performance of RSVD, asserted by the following.

Theorem 1. *Let $\mathbf{A}$ be a positive semidefinite matrix with $\lambda_1 > 0$ and $\hat{\mathbf{u}} = \mathsf{RSVD}(\mathbf{A}, \mathcal{N}(0,1)^{n\times d}, q, d)$ for any $q \in \mathbb{N}$. Then*

$$R(\hat{\mathbf{u}}) = \left(\Omega\left(\frac{d}{n}\right)\right)^{\frac{1}{2q+1}}$$

holds with probability at least $1 - e^{-\Omega(d)}$.

Proof. If $\mathbf{A}$ is p.s.d. we have $\alpha_i \geq 0$, and thus (assuming $q \in \mathbb{N}$) we can repeatedly apply the Cauchy-Schwarz inequality to Equation (4) and get

$$R_{\mathbf{a}} \geq \frac{\sum_i \alpha_i^{2q}\langle\mathbf{S}^T\mathbf{u}_i,\mathbf{a}\rangle^2}{\sum_i \alpha_i^{2q-1}\langle\mathbf{S}^T\mathbf{u}_i,\mathbf{a}\rangle^2} \geq \cdots \geq \frac{\sum_i \alpha_i\langle\mathbf{S}^T\mathbf{u}_i,\mathbf{a}\rangle^2}{\sum_i\langle\mathbf{S}^T\mathbf{u}_i,\mathbf{a}\rangle^2}. \tag{5}$$

The key observation is that by repeatedly using Equation (5) results in

$$\sum_{i=1}^n\langle\mathbf{S}^T\mathbf{u}_i,\mathbf{a}\rangle^2 \geq R_{\mathbf{a}}^{-1}\sum_{i=1}^n \alpha_i\langle\mathbf{S}^T\mathbf{u}_i,\mathbf{a}\rangle^2 \geq \cdots \geq R_{\mathbf{a}}^{-(2q+1)}\sum_{i=1}^n \alpha_i^{2q+1}\langle\mathbf{S}^T\mathbf{u}_i,\mathbf{a}\rangle^2$$

which implies

$$R_{\mathbf{a}}^{2q+1} \geq \frac{\sum_{i=1}^n \alpha_i^{2q+1}\langle\mathbf{S}^T\mathbf{u}_i,\mathbf{a}\rangle^2}{\sum_{i=1}^n\langle\mathbf{S}^T\mathbf{u}_i,\mathbf{a}\rangle^2} \geq \frac{\langle\mathbf{S}^T\mathbf{u}_1,\mathbf{a}\rangle^2}{\sum_{i=1}^n\langle\mathbf{S}^T\mathbf{u}_i,\mathbf{a}\rangle^2}. \tag{6}$$

Finally, by $R(\hat{\mathbf{u}}) = \max_{\mathbf{a}\in\mathbb{S}^{d-1}} R_{\mathbf{a}}$ and Definition 1 we have

$$R(\hat{\mathbf{u}})^{2q+1} \geq \max_{\mathbf{a}\in\mathbb{S}^{d-1}}\frac{\langle\mathbf{S}^T\mathbf{u}_1,\mathbf{a}\rangle^2}{\sum_{i=1}^n\langle\mathbf{S}^T\mathbf{u}_i,\mathbf{a}\rangle^2} = \cos^2\theta(\mathbf{u}_1,\mathbf{S}), \tag{7}$$

and invoking Lemma 1 proves the claim. □

We offer a few remarks. First note that the fact that Equation (7) implies Theorem 1 can be proven by estimating $R_{\mathbf{a}}$ only on $\mathbf{a} = \frac{\mathbf{S}^T\mathbf{u}_1}{\|\mathbf{S}^T\mathbf{u}_1\|_2}$, since we essentially prove Lemma 1 on such a vector $\mathbf{a}$ — see our discussion in Section 3 or Appendix A.1. Second, Equation (6) can also be shown by Hölder's inequality — see a simplified proof of Theorem 1 in Appendix A.2. Third, from Theorem 1, we see that increasing the number of passes q makes $R(\hat{\mathbf{u}})$ approaching to 1 exponentially fast, while increasing the dimension d leads to stronger concentration of $R(\hat{\mathbf{u}})$ around the slowly increased mean $\Omega((d/n)^{1/(2q+1)})$. Finally, we have:

Remark 1. *The guarantee by Theorem 1 can be written as $R(\hat{\mathbf{u}}) = e^{-\mathcal{O}(\ln n/(2q+1))} \geq 1 - \mathcal{O}(\ln n/q)$, and hence, subsumes the result of Musco and Musco (2015).*

One may wonder if our analysis is tight. The next theorem confirms the tightness of Theorem 1 up to a constant factor.

Theorem 2. *For any $q \in \mathbb{N}$, there exists a positive semidefinite matrix $\mathbf{A}$ with $\lambda_1 > 0$, so that for $\hat{\mathbf{u}} = \mathsf{RSVD}(\mathbf{A}, \mathcal{N}(0,1)^{n\times d}, q, d)$, it holds*

$$R(\hat{\mathbf{u}}) = \mathcal{O}\left(\left(\frac{d}{n}\right)^{\frac{1}{2q+1}}\right),$$

with probability at least $1 - e^{-\Omega(d)}$.

We prove Theorem 2 in Appendix A.2 by considering the following eigenvalue distribution $\{\alpha_i\}$:

$$1 = \alpha_1 > \alpha_2 = \cdots = \alpha_n = \left(\frac{d}{n}\right)^{\frac{1}{2q+1}}. \tag{8}$$

While our worst-case analysis is tight, Equation (8) rarely happens in practice. Instead, real-world matrices are often observed to have rapidly decaying singular values (Chakrabarti and Faloutsos, 2006; Eikmeier and Gleich, 2017). To take this consideration into account, we introduce the following definition to capture whether$\mathbf{A}$ has at least power-law decay of its singular values $\{\sigma_i\}_{i\geq i_0}^n$.

Definition 2. *Let*

$$i_0 = \begin{cases} \min_{j\in\mathcal{J}} j & \text{if } \mathcal{J} \neq \emptyset, \\ n & \text{otherwise}, \end{cases}$$

where $\mathcal{J} \subseteq [n]$ consists of all the integers $j \in [n]$ such that there exists $\gamma > 1/q$ and $C > 0$ satisfying $\sigma_i/\sigma_1 \leq C \cdot i^{-\gamma}$, for all $i \geq j$.

Theorem 3. *Let $\mathbf{A}$ be a positive semidefinite matrix, $\hat{\mathbf{u}} = \mathsf{RSVD}(\mathbf{A}, \mathcal{N}(0,1)^{n\times d}, q, d)$ for any $q \in \mathbb{N}$, and i_0 be defined as in Definition 2. Then*

$$R(\hat{\mathbf{u}}) = \Omega\left(\left(\frac{d}{d+i_0}\right)^{\frac{1}{2q+1}}\right)$$

holds with probability at least $1 - e^{-\Omega(d)}$.

The proof of Theorem 3 can be found in Appendix A.2. The idea is to estimate $R_{\mathbf{a}}$ on $\mathbf{a} = \frac{\mathbf{S}^T\mathbf{u}_1}{\|\mathbf{S}^T\mathbf{u}_1\|_2}$ and check two possible cases. If i_0 is large, the analysis reduces to Theorem 1, while if i_0 is small, we invoke Bernstein-type inequalities and show that $R_{\mathbf{a}} = \Omega(1)$ with high probability. So, the overall guarantee of $R_{\mathbf{a}}$ is determined by the former case, and recalling $R(\hat{\mathbf{u}}) \geq \max_{\mathbf{a}\in\mathbb{S}^{d-1}} R_{\mathbf{a}}$ yields Theorem 3.

Remark 2. *Theorem 3 subsumes Theorem 1 up to a constant factor as $d + i_0 = \mathcal{O}(n)$, and provides a much better guarantee if $\mathbf{A}$ has singular values having at least power-law decay. In particular, if $i_0 = \mathcal{O}(d)$ then $R(\hat{\mathbf{u}}) = \Omega(1)$ with high probability, even with a single pass when $q = 1$ and $d = 1$.*

4.3 Indefinite matrices

If $\mathbf{A}$ has negative eigenvalues, the Inequality (5) in the proof of Theorem 1 is not valid anymore. Nevertheless, we expect to have a guarantee of $R(\hat{\mathbf{u}})$ similar to that of Theorem 1 if the negative eigenvalues are not too large. We introduce the following technical assumption.

Assumption 1. *Assume there exists a constant $\kappa \in (0,1]$ such that $\sum_{i=2}^n \lambda_i^{2q+1} \geq \kappa \sum_{i=2}^n |\lambda_i|^{2q+1}$.*

An important observation is that Theorems 1 and 3 can be proved by estimating $R_{\mathbf{a}}$ only on one specific vector $\mathbf{a} = \frac{\mathbf{S}^T\mathbf{u}_1}{\|\mathbf{S}^T\mathbf{u}_1\|_2}$; see Section 4.2. Hence, it suffices to use the following lemma (proved in Appendix A.3) to generalize our results in Section 4.2 to indefinite matrices satisfying Assumption 1.

Lemma 2. *Assume that matrix $\mathbf{A}$ satisfies Assumption 1 and $\mathbf{S} \sim \mathcal{N}(0,1)^{n\times d}$. There exists a constant $c_\kappa \in (0,1]$ such that with probability at least $1 - e^{-\Omega(\sqrt{d}\kappa^2)}$, it holds*

$$\sum_{i=1}^n \lambda_i^{2q+1} \langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2 \geq c_\kappa \sum_{i=1}^n |\lambda_i|^{2q+1} \langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2.$$

Lemma 2 essentially states that any indefinite matrix $\mathbf{A}$ satisfying Assumption 1 has $R_{\mathbf{a}} = \Theta(\bar{R}_{\mathbf{a}})$ on such a vector $\mathbf{a} = \frac{\mathbf{S}^T\mathbf{u}_1}{\|\mathbf{S}^T\mathbf{u}_1\|_2}$, where

$$\bar{R}_{\mathbf{a}} = \frac{\sum_{i=1}^n |\alpha_i|^{2q+1} \langle \mathbf{S}^T\mathbf{u}_i, \mathbf{a}\rangle^2}{\sum_{i=1}^n \alpha_i^{2q} \langle \mathbf{S}^T\mathbf{u}_i, \mathbf{a}\rangle^2}. \tag{9}$$

The next theorem, proven in Appendix A.3, follows from Lemma 2 and the proof of Theorem 3.

Theorem 4. *Assume that matrix $\mathbf{A}$ satisfies Assumption 1. Let $\hat{\mathbf{u}} = \mathsf{RSVD}(\mathbf{A}, \mathcal{N}(0,1)^{n\times d}, q, d)$ for any $q \in \mathbb{N}$. Then,*

$$R(\hat{\mathbf{u}}) = \Omega\left(c_\kappa \left(\frac{d}{d+i_0}\right)^{\frac{1}{2q+1}}\right),$$

with probability at least $1 - e^{-\Omega(\sqrt{d}\kappa^2)}$.

Remark 3. *As discussed in Section 3, all the theorems shown in this section, i.e., Theorems 1, 2, 3, and 4, can be easily extended to any random matrix $\mathbf{S}$ satisfying $\mathbf{S} \sim$ Uniform$(\mathcal{G}_{n,d})$.*

5 EXTENSION: COMBINING WITH PROJECTION FROM BERNOULLI

In this section, we propose an extension of Randomized SVD, which we name $\mathsf{RandSum}$, and show as Algorithm 2. In $\mathsf{RandSum}$, half of the columns of $\mathbf{S}$ are replaced with i.i.d. samples from a Bernoulli distribution with mean $p \in (0,1)$.[4] We can show that the guarantee achieved by the $\mathsf{RandSum}$ algorithm for $R(\hat{\mathbf{u}})$ is no worse than that by the RSVD algorithm, since

[4] Bernoulli$(p)^{n\times d}$ does not belong to the class of distributions mentioned in Section 3 to which Lemma 1 applies.

Algorithm 2: RandSum ($\mathbf{A}$, q, d, p)

$\mathbf{S}_1 \sim \mathcal{N}(0,1)^{n\times\lceil \frac{d}{2}\rceil}$, $\mathbf{S}_2 \sim \text{Bernoulli}(p)^{n\times\lfloor \frac{d}{2}\rfloor}$;
$\mathbf{S} \leftarrow \begin{bmatrix}\mathbf{S}_1 & \mathbf{S}_2\end{bmatrix}$;
return $\mathsf{RSVD}(\mathbf{A}, \mathbf{S}, q, d)$;

half of the coulmns of $\mathbf{S}$ come from a normal distribution. To study the additional benefits due to the submatrix drawn from the Bernoulli, we derive the following lemma as an analog of Lemma 1 for a Bernoulli random matrix. The proof is in Appendix B.1.

Lemma 3. *Let $\mathbf{v} \in \mathbb{S}^{n-1}$, $d \leq n/3$, and $\mathbf{S} \sim$ Bernoulli$(p)^{n\times d}$ for a constant $p \in (0,1)$ Then,*

$$\cos^2\theta(\mathbf{v},\mathbf{S}) = \Omega\left(\frac{\max\{1, \langle \mathbf{v}, \mathbf{1}_n\rangle^2\}}{n}\right)$$

holds with probability at least $1 - e^{-\Omega(d)}$.

The next theorem, which holds for any p.s.d. matrix $\mathbf{A}$, is a direct consequence of Lemmas 1 and 3 and applying the techniques introduced in Theorem 1. The proof is in Appendix B.2.

Theorem 5. *Let $\mathbf{A}$ be a positive semindefinite matrix with $\lambda_1 > 0$, and $\hat{\mathbf{u}} = \mathsf{RandSum}(\mathbf{A}, q, d, p)$ for any constant $p \in (0,1)$ and integer $d \geq 2$. Then,*

$$R(\hat{\mathbf{u}}) = \left(\Omega\left(\frac{\max\{d, \langle \mathbf{u}_1, \mathbf{1}_n\rangle^2\}}{n}\right)\right)^{\frac{1}{2q+1}}$$

holds with probability at least $1 - e^{-\Omega(d)}$.

Theorem 5 shows that $R(\hat{\mathbf{u}}) = \Theta(1)$ with high probability when $\langle \mathbf{u}_1, \mathbf{1}_n\rangle^2 = \Theta(n)$, which is acheviable as the maximum possible value of $\langle \mathbf{u}_1, \mathbf{1}_n\rangle^2$ is n.

Remark 4. *For certain tasks such as conflicting-group detection (Bonchi et al., 2019; Tzeng et al., 2020), one could expect to have large $\langle \mathbf{u}_1, \mathbf{1}_n\rangle^2$, since $\langle \mathbf{u}_1, \mathbf{1}_n\rangle^2$ naturally corresponds to the size of the subgraph, which is located by $\mathbf{u}_1$.[5] However, for tasks such as community detection, $\langle \mathbf{u}_1, \mathbf{1}_n\rangle^2 \approx 0$ is often the case.*

Finally, we consider the generalization of Theorem 5 to indefinite matrices. To derive Lemma 4, the analog of Lemma 2 for Bernoulli random matrices, we introduce Assumption 2, where (i) is merely for the ease of presentation and (ii) generalizes Assumption 1 as $\xi_i = 1$ for $\mathbf{S} \sim \mathcal{N}(0,1)^{n\times d}$. The proof of Lemma 4 can be found in Appendix B.3.

[5]We say that $\mathbf{u}_1$ is located around some indices $\mathcal{I} \subseteq [n]$ if the magnitude of $(\mathbf{u}_1)_i$ for any $i \in \mathcal{I}$ is much larger than those not in $\mathcal{I}$.

Assumption 2. *Assume that (i) $\langle \mathbf{u}_1, \mathbf{1}_n\rangle^2 = \Omega(1)$ and (ii) there exists a constant $\kappa' \in (0,1]$ such that*

$$\sum_{i=2}^{n} \lambda_i^{2q+1}\xi_i \geq \kappa' \sum_{i=2}^{n} |\lambda_i|^{2q+1}\xi_i,$$

where $\xi_i = \mathbb{E}\left[\langle \mathbf{S}^T\mathbf{u}_i, \frac{\mathbf{1}_d}{\sqrt{d}}\rangle^2\right]$, for all $i \in [n]$.

Lemma 4. *Assume that $\mathbf{A}$ satisfies Assumption 2. Let $\mathbf{S} \sim$ Bernoulli$(p)^{n\times d}$ for a constant $p \in (0,1)$. There exists a constant $c_{\kappa'} \in (0,1]$, such that*

$$\sum_{i=1}^{n} \lambda_i^{2q+1}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2 \geq c_{\kappa'} \sum_{i=1}^{n} |\lambda_i|^{2q+1}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2,$$

with probability at least $1 - e^{-\Omega(\sqrt{d}\kappa'^2)}$.

Our last result, Theorem 6, immediately follows from Theorems 4 and 5 and Lemma 4. The proof and the full version are in Appendix B.3.

Theorem 6. *Assume that $\mathbf{A}$ satisfies Assumptions 1 and 2. Let $\hat{\mathbf{u}} = \mathsf{RandSum}(\mathbf{A}, q, d, p)$ for any constant $p \in (0,1)$ and any $q \in \mathbb{N}$, and i_0 be defined as in Definition 2. Then,*

$$R(\hat{\mathbf{u}}) = \Omega\left(\left(\max\left\{\frac{d}{d+i_0}, \frac{\langle \mathbf{u}_1, \mathbf{1}_n\rangle^2}{n}\right\}\right)^{\frac{1}{2q+1}}\right)$$

holds with probability at least $1 - e^{-\Omega(\sqrt{d})}$. (For the full dependency on κ, κ', c_κ, and $c_{\kappa'}$, see Appendix B.3.)

6 EXPERIMENTS

In this section we evaluate the randomized algorithms we analyze in this paper using synthetic and real-world datasets. In Section 6.1, we use synthetic datasets to benchmark the RSVD algorithm with respect to the R measure, and study the effect of its parameters. In Section 6.2, we employ RSVD and RandSum as subroutines of spectral approaches for specific knowledge-discovery tasks on real-world datasets.

Settings. We use LanczosMethod, provided by the ARPACK library (Lehoucq et al., 1998), for computing λ_1, which is required for measuring R. We fix $q = 1$ while varying $d \in \{1, 5, 10, 25, 50\}$ to study the effect of d, and fix $d = 10$ while varying $q \in \{1, 2, 4, 8, 16\}$ to study the effect of q. Each setting is repeated 100 times and the average is reported. All experiments are performed on an Intel Core i5 machine at 1.8 GHz with 8 GB RAM. All methods are implemented in Python 3.7.4.[6]

[6]The code is available at the github repo `https://bit.ly/34dI4Nl`.

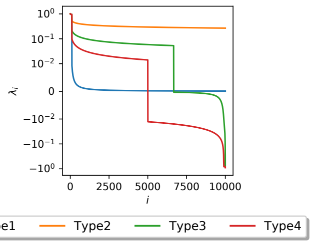


Figure 1: Different types of eigenvalue distributions.

6.1 Evaluation with synthetic data

We consider different types of eigenvalue distributions, also illustrated in Figure 1. The size of the input matrix is set to $n = 10\,000$ and $i_0 = 100$ (see Definition 2). For all types of synthetic matrices we set $\lambda_i = i^{-0.01}$, for $i < i_0$, and the rest of the eigenvalues $\{\lambda_i\}_{i \geq i_0}^n$ are specified as follows:

- Type 1: $\lambda_i = i^{-1}$ for $i \geq i_0$.
- Type 2: $\lambda_i = i^{-\frac{1}{7}}$ for $i \geq i_0$.
- Type 3: $\lambda_i = \begin{cases} i^{-\frac{1}{3}} & \text{if } i \in [i_0, \frac{2n}{3}], \\ -(i - \frac{2n}{3})^{-1} & \text{if } i > \frac{2n}{3}. \end{cases}$
- Type 4: $\lambda_i = \begin{cases} i^{-\frac{1}{2}} & \text{if } i \in [i_0, \frac{n}{2}], \\ -\frac{9}{10}(i - \frac{n}{2})^{-\frac{1}{2}} & \text{if } i \in (\frac{n}{2}, n - i_0), \\ -\frac{9}{10} i^{-0.01} & \text{if } i \geq n - i_0. \end{cases}$

For the value of κ in Assumption 1, we compute κ with $q = 1$ and get: $\kappa = 1$ for Type 1 and Type 2, $\kappa = 0.99$ for Type 3, and $\kappa = 0.22$ for Type 4. For each type of eigenvalue distribution, we generate a random $n \times n$ input matrix by sampling the eigenvectors uniformly from the space of orthogonal matrices.

Figure 2 shows the value of R for the vector $\hat{\mathbf{u}}$ computed by $\mathsf{RSVD}(\mathbf{A}, \mathcal{N}(0,1)^{n \times d}, q, d)$, and the speedup in running time against $\mathsf{LanczosMethod}$.

For matrices of Type 1, it is expected that RSVD performs the best as the eigenvalues of such matrices have the fastest decay and $\kappa = 1$.

For matrices of Type 2, we notice that $R(\hat{\mathbf{u}})$ is very close to 1 when $q \geq 4$. This result is better than what our analysis predicts, since by Theorem 3 it holds that $R(\hat{\mathbf{u}}) = \Omega(1)$ with high probability after $q = 7$ (since the decay rate of Type 2 is $1/7$).

For matrices of Type 3, despite being indefinite, the magnitude of the negative eigenvalues is almost negligible ($\kappa = 0.99$). By Theorem 4 and Lemma 2, $R(\hat{\mathbf{u}})$ is nearly identical to its counterpart $\bar{R}$ (see (9)), so it is expected that RSVD performs better on data of Type 3 than on data of Type 2, as the eigenvalue-distribution decay rate is faster.

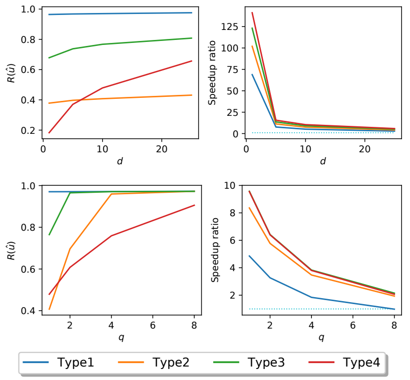


Figure 2: The value of $R(\hat{\mathbf{u}})$ for $\hat{\mathbf{u}}$ computed by $\mathsf{RSVD}(\mathbf{A}, \mathcal{N}(0,1)^{n \times d}, q, d)$. Top row shows dependence with d. Bottom row shows dependence with q. The speedup is measured against $\mathsf{LanczosMethod}$.

For matrices of Type 4, although the eigenvalues decay faster than those of Type 3 matrices, the magnitudes of the negative eigenvalues are much larger ($\kappa = 0.22$). By Theorem 4 and Lemma 2, $R(\hat{\mathbf{u}})$ is upper-bounded by a factor of κ when increasing q, and the results indeed show that the performance of RSVD is worse for Type 4 matrices, compared to Type 3 ($\kappa = 0.99$).

6.2 Applications on real-world data

We use publicly-available networks from the SNAP collection (Leskovec and Krevl, 2014). Statistics of the datasets are listed in Tables 1 and 2.

6.2.1 Detection of 2 conflicting groups

The problem of 2-conflicting group detection aims to find two optimal groups that maximize the polarity objective $P(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} / \mathbf{x}^T \mathbf{x}$, where $\mathbf{A}$ is the signed adjacency matrix and $x \in \mathcal{T} = \{0, \pm 1\}^n \backslash \{\mathbf{0}_n\}$. Bonchi et al. (2019) propose a tight $\mathcal{O}(n^{1/2})$-approximation algorithm based on the leading eigenvector $\mathbf{u}_1$. In Appendix D we show that applying their approach on the approximated top-eigenvector $\hat{\mathbf{u}}$ yields an $\mathcal{O}(n^{1/2}/R(\hat{\mathbf{u}}))$-approx algorithm.

Datasets. The statistics of datasets we use for this

Table 1: Datasets for conflicting group detection.

	WikiVot	Referendum	Slashdot	WikiCon
$\lvert V\rvert$	7 115	10 884	82 140	116 717
$\lvert E\rvert$	100 693	251 406	500 481	2 026 646
(γ, i_0)	(4.6, 15)	(4.5, 16)	(5.3, 17)	(2.8, 22)
κ	0.397	0.620	0.204	0.034
$\cos\theta(\mathbf{u}_1, \mathbf{1}_n)$	0.378	0.399	0.194	0.193

Table 2: Datasets for community detection.

	FBArtist	Gnutella31	YouTube	RoadCA
$\lvert V\rvert$	50 515	62 586	1 134 890	1 965 206
$\lvert E\rvert$	819 306	147 892	2 987 624	2 766 607

experiment are listed in Table 1. We observe that all datasets have rapidly-decaying singular values. To measure the parameters γ and i_0 (see Definition 2), due to memory limitations, we compute the top 1 000 eigenvalues (in magnitude) of its signed adjacency matrix by LanczosMethod, and fit the parameters (γ, i_0) by an MLE-based method (Clauset et al., 2009). Moreover, we test the validity of Assumption 1 by computing κ with $q = 1$, and also computing $\langle \mathbf{u}_1, \mathbf{1}_n\rangle$.

Results. Figure 3 illustrates the results obtained by applying the spectral algorithm of Bonchi et al. (2019) on the top-eigenvector $\hat{\mathbf{u}}$ returned by RSVD and RandSum. Due to the value of κ, the result is that, as expected, both algorithms perform the best on Referendum. Due to the value of $\cos\theta(\mathbf{u}_1, \mathbf{1}_n)$, the superiority of RandSum over RSVD is, as expected, more pronounced on WikiVot and Referendum than on Slashdot and WikiCon.

6.2.2 Detection of 2 communities

For the task of detecting two communities in a graph, Newman (2006) proposed an efficient algorithm by maximizing the modularity score $Q(\mathbf{x}) = \mathbf{x}^T\mathbf{M}\mathbf{x}/4|E|$, where $\mathbf{M}_{i,j} = \mathbf{A}_{i,j} - \deg(i)\deg(j)/2|E|$, $\mathbf{A}$ is the adjacency matrix of the input graph, and the two communities are determined by the sign of the top eigenvector of $\mathbf{M}$.

Datasets. The datasets used for evaluating this task are listed in Table 2. As the modularity matrix $\mathbf{M}$ is dense and the networks are large, LanczosMethod runs out-of-memory on our machine when trying to compute the top eigenvalues, and hence, unlike Table 1, the number κ and the parameters (γ, i_0) are not displayed in Table 2.

Results. Figure 4 shows the results by applying the spectral algorithm of Newman (2006) on the top-eigenvector $\hat{\mathbf{u}}$ returned by RSVD and RandSum. Notice that on this task, RandSum has no advantage over

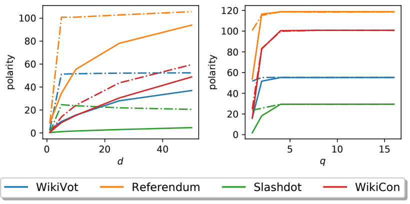


Figure 3: Results on the task of detecting 2 conflicting groups. Results for RSVD (resp. RandSum) are plotted with a solid (resp. dashed) line.

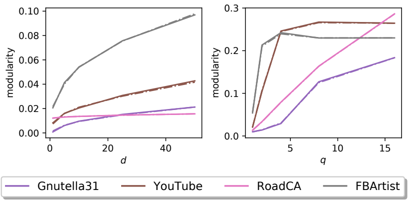


Figure 4: Results on the task of detecting 2 communities. Results for RSVD (in solid line) and RandSum (in dotted line) are nearly the same.

RSVD since $\mathbf{M}\mathbf{1}_n = 0$, and thus $\langle \mathbf{u}_1, \mathbf{1}_n\rangle = \mathbf{0}_n$ if $\lambda_1 \geq 0$. When fixing $d = 10$ and increasing q, the modularity scores converge much faster on FBArtist and YouTube than on Gnutella31 and RoadCA, suggesting that it could be hard to discover community structures in Gnutella31 and RoadCA. This is an expected result. For Gnutella (Gnutella31) the design of the network prevents the formation of large communities so as to enable reliable communication For the road network of California (RoadCA) the reason is the grid-like structure of the network (Leskovec et al., 2009).

7 CONCLUSION

In this paper, we study the problem of approximating the leading eigenvector of a matrix with limited number of passes. The problem is of interest in many applications. We provide a tight theoretical analysis of the popular randomized SVD method, with respect to the metric $R(\hat{\mathbf{u}}) = \lambda_1^{-1}\hat{\mathbf{u}}^T\mathbf{A}\hat{\mathbf{u}}/\hat{\mathbf{u}}^T\hat{\mathbf{u}}$. Our results substantially improve the analysis of randomized SVD in the regime of $o(\ln n)$ passes and recover the analysis of prior works in the regime of $\Omega(\ln n)$ passes. A new technique is introduced to transform the problem of maximizing $R(\hat{\mathbf{u}})$ into a well-studied problem in the literature of high-dimensional probability.

Our work opens several interesting directions. First, it is an open problem to characterize the fundamental limit of maximizing $R(\hat{\mathbf{u}})$ for any algorithm with fixed number of pass and $\mathcal{O}(n)$ space. Second, our results may be extended in different ways. For example, we may relax the requirement on the input matrix from symmetric to stochastic, so as to analyze approximations of PageRank (Page et al., 1999). Or, we may extend RandSum to use any non-centered subgaussian distribution for drawing $\mathbf{S}_2$, and we conjecture this yields similar results. Another direction is to extend our analysis to top-k eigenvectors; since there are already several methods for computing top-k eigenvectors (Halko et al., 2011b; Mackey, 2008; Allen-Zhu and Li, 2016), the most challenging part is to define the proper metric to maximize, as a generalization of $R(\hat{\mathbf{u}})$.

Acknowledgements

We thank the anonymous reviewers for their insightful feedback. This research is supported by the ERC Advanced Grant REBOUND (834862), the EC H2020 RIA project SoBigData++ (871042), and the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation.

References

Amirali Abdullah, Alexandr Andoni, Ravindran Kannan, and Robert Krauthgamer. Spectral approaches to nearest neighbor search. In *Proc. of FOCS*. IEEE, 2014.

Dimitris Achlioptas. Database-friendly random projections. In *Proc. of PODS*, 2001.

Zeyuan Allen-Zhu and Yuanzhi Li. Lazysvd: Even faster svd decomposition yet without agonizing pain. In *Proc. of NeurIPS*, 2016.

Aris Anagnostopoulos, Luca Becchetti, Adriano Fazzone, Cristina Menghini, and Chris Schwiegelshohn. Spectral relaxations and fair densest subgraphs. In *Proc. of CIKM*, 2020.

Francesco Bonchi, Edoardo Galimberti, Aristides Gionis, Bruno Ordozgoiti, and Giancarlo Ruffo. Discovering polarized communities in signed networks. In *Proc. of CIKM*, 2019.

Deepayan Chakrabarti and Christos Faloutsos. Graph mining: Laws, generators, and algorithms. *ACM computing surveys (CSUR)*, 2006.

Xixian Chen, Irwin King, and Michael R Lyu. Frosh: Faster online sketching hashing. In *Proc. of UAI*, 2017.

Aaron Clauset, Cosma Rohilla Shalizi, and Mark EJ Newman. Power-law distributions in empirical data. *SIAM review*, 2009.

IBM Reseach Corporation. libskylark: Sketching-based distributed matrix computations for machine learning. `https://github.com/xdata-skylark/libskylark`, 2021. [Online; accessed 27-April-2021].

Petros Drineas, Ilse CF Ipsen, Eugenia-Maria Kontopoulou, and Malik Magdon-Ismail. Structural convergence results for approximation of dominant subspaces from block krylov spaces. *SIAM Journal on Matrix Analysis and Applications*, 2018.

Nicole Eikmeier and David F Gleich. Revisiting power-law distributions in spectra of real world networks. In *Proc of SIGKDD*, 2017.

N Benjamin Erichson, Sergey Voronin, Steven L Brunton, and J Nathan Kutz. Randomized matrix decompositions using r. *Journal of Statistical Software*, 2019.

Mina Ghashami, Edo Liberty, Jeff M Phillips, and David P Woodruff. Frequent directions: Simple and deterministic matrix sketching. *SIAM Journal on Scientific Computing*, 2016.

Allan Gut. Multivariate random variables. In *An Intermediate Course in Probability*. Springer, 2009.

Nathan Halko, Per-Gunnar Martinsson, Yoel Shkolnisky, and Mark Tygert. An algorithm for the principal component analysis of large data sets. *SIAM Journal on Scientific Computing*, 2011a.

Nathan Halko, Per-Gunnar Martinsson, and Joel A Tropp. Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions. *SIAM review*, 2011b.

Moritz Hardt and Eric Price. The noisy power method: a meta algorithm with applications. In *Proc. of NeurIPS*, 2014.

Samuel B Hopkins, Tselil Schramm, Jonathan Shi, and David Steurer. Fast spectral algorithms from sum-of-squares proofs: tensor decomposition and planted sparse vectors. In *Proc. of STOC*, 2016.

Roger A Horn, Roger A Horn, and Charles R Johnson. *Topics in matrix analysis.* Cambridge university press, 1994.

Zengfeng Huang. Near optimal frequent directions for sketching dense and sparse matrices. In *Proc. of ICML*. PMLR, 2018.

Ian T Jolliffe. Principal components in regression analysis. In *Principal component analysis*. Springer, 1986.

Ravindran Kannan and Santosh Vempala. *Spectral algorithms*. Now Publishers Inc., 2009.

Jacek Kuczyński and Henryk Woźniakowski. Estimating the largest eigenvalue by the power and lanczos algorithms with a random start. *SIAM Journal on Matrix Analysis and Applications*, 1992.

Richard B Lehoucq, Danny C Sorensen, and Chao Yang. *ARPACK users' guide: solution of large-scale eigenvalue problems with implicitly restarted Arnoldi methods*. SIAM, 1998.

Jure Leskovec and Andrej Krevl. SNAP Datasets: Stanford large network dataset collection. `http://snap.stanford.edu/data`, 2014.

Jure Leskovec, Kevin J Lang, Anirban Dasgupta, and Michael W Mahoney. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Mathematics*, 2009.

Antoine Liutkus. randomized singular value decomposition in matlab central file exchange. 47835-randomized-singular-value-decomposition, 2021. [Online; accessed 27-April-2021].

Lester Mackey. Deflation methods for sparse pca. In *Proc. of NIPS*, 2008.

Michael W Mahoney et al. Randomized algorithms for matrices and data. *Foundations and Trends® in Machine Learning*, 2011.

Per-Gunnar Martinsson and Joel A Tropp. Randomized numerical linear algebra: Foundations and algorithms. *Acta Numerica*, 2020.

Cameron Musco and Christopher Musco. Randomized block krylov methods for stronger and faster approximate singular value decomposition. In *Proc. of NeurIPS*, 2015.

Cameron Musco and David P Woodruff. Sublinear time low-rank approximation of positive semidefinite matrices. In *Proc. of FOCS*. IEEE, 2017.

Mark EJ Newman. Modularity and community structure in networks. *Proc. of NAS*, 2006.

Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: Bringing order to the web. Technical report, Stanford InfoLab, 1999.

Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. Scikit-learn: Machine learning in python. *JMLR*, 2011.

Radim Řehůřek and Petr Sojka. Software framework for topic modelling with large corpora. In *Proc. of LREC 2010 Workshop*, 2010. `https://radimrehurek.com/gensim/models/lsimodel.html#gensim.models.lsimodel.stochastic_svd`.

Mark Rudelson and Roman Vershynin. Smallest singular value of a random rectangular matrix. *Communications on Pure and Applied Mathematics*, 2009.

Mark Rudelson and Roman Vershynin. Non-asymptotic theory of random matrices: extreme singular values. In *Proc. of International Congress of Mathematicians*, 2010.

Arlei Silva, Ambuj Singh, and Ananthram Swami. Spectral algorithms for temporal graph cuts. In *Proc. of WWW*, 2018.

Max Simchowitz, Ahmed El Alaoui, and Benjamin Recht. Tight query complexity lower bounds for pca via finite sample deformed wigner law. In *Proc. of STOC*, 2018.

G. W. Stewart and Ji Guang Sun. *Matrix Perturbation Theory*. Academic Press, 1990.

Pascal Terray and Françoise Pinsard. statpack: Eig procedures. `https://terray.locean-ipsl.upmc.fr/statpack2.1/manuals/Module_Eig_Procedures.html`, 2021. [Online; accessed 17-September-2021].

Ruo-Chun Tzeng, Bruno Ordozgoiti, and Aristides Gionis. Discovering conflicting groups in signed networks. In *Proc. of NeurIPS*, 2020.

Roman Vershynin. *High-dimensional probability: An introduction with applications in data science*. Cambridge university press, 2018.

David P Woodruff et al. Sketching as a tool for numerical linear algebra. *Foundations and Trends® in Theoretical Computer Science*, 2014.

Peng Xu, Bryan He, Christopher De Sa, Ioannis Mitliagkas, and Chris Re. Accelerated stochastic power iteration. In *Proc. of AISTATS*. PMLR, 2018.

Yi Yu, Tengyao Wang, and Richard J Samworth. A useful variant of the davis–kahan theorem for statisticians. *Biometrika*, 2015.

Supplementary Material: Improved analysis of randomized SVD for top-eigenvector approximation

A Proofs of RSVD

A.1 Large deviation of projection length for Gaussian random matrix

This subsection is devoted to proving Lemma 1 restated below.

Lemma 1. *Let $\mathbf{v} \in \mathbb{R}^n$ be a nonzero vector and $\mathbf{S} \sim \mathcal{N}(0,1)^{n\times d}$ where $n, d \in \mathbb{N}$ and $n \geq d$. Then,*

$$\cos^2\theta(\mathbf{v},\mathbf{S}) = \Theta\left(\frac{d}{n}\right)$$

with probability at least $1 - e^{-\Omega(d)}$.

This lemma stems from the observations that $\frac{\sigma_1(\mathbf{S}^T\mathbf{v})}{\sigma_1(\mathbf{S})} \leq \cos\theta(\mathbf{v},\mathbf{S}) \leq \frac{\sigma_1(\mathbf{S}^T\mathbf{v})}{\sigma_d(\mathbf{S})}$ and the distribution of $\frac{\mathbf{S}^T\mathbf{v}}{\|\mathbf{v}\|_2}$ is exactly $\mathcal{N}(0,1)^{d\times 1}$. The proof relies on the union bound of concentration inequalities on the extreme singular values of Gaussian random matrix, Lemma 5, and Lemma 6. Similar inequalities shown in the previous works, e.g. Hardt and Price (2014), also rely on this observation.

Lemma 5 (Theorem 4.4.5 (Vershynin, 2018))**.** *Let $\mathbf{S}$ be a $n \times d$ random matrix whose entries are i.i.d. zero-mean subgaussian r.v.'s.*

$$\text{For all } t > 0, \quad \mathbb{P}\left[\sigma_1(\mathbf{S}) \geq c\,(\sqrt{n} + \sqrt{d} + t)\right] \leq 2e^{-t^2},$$

where $c > 0$ depends linearly only on $\|\mathbf{S}_{1,1}\|_{\psi_2}$ (see Definition 4 of ψ_2-norm in Appendix C.1).

Lemma 6 (Theorem 1.1 (Rudelson and Vershynin, 2009))**.** *Let $\mathbf{S}$ be a $n \times d$ random matrix whose entries are i.i.d. zero-mean subgaussian r.v.'s and $n \geq d$.*

$$\text{For all } \delta > 0, \quad \mathbb{P}\left[\sigma_d(\mathbf{S}) \leq \delta\,(\sqrt{n} - \sqrt{d-1})\right] \leq (c_1\delta)^{n-d+1} + e^{-c_2 n},$$

where $c_1, c_2 > 0$ have polynomial dependence on $\|\mathbf{S}_{1,1}\|_{\psi_2}$ (see Definition 4 of ψ_2-norm in Appendix C.1).

Proof of Lemma 1: For the simplicity of presentation, we assume $\|\mathbf{v}\|_2 = 1$ as $\cos\theta(\cdot,\cdot)$ is scale-invariant.

(i) $\cos\theta(\mathbf{v},\mathbf{S}) = \Omega(\sqrt{d/n})$:

Recall that $\cos\theta(\mathbf{v},\mathbf{S}) = \max_{\mathbf{a}\in\mathbb{S}^{d-1}} \frac{\langle \mathbf{v},\mathbf{S}\mathbf{a}\rangle}{\|\mathbf{S}\mathbf{a}\|_2}$. Let $\mathbf{a} = \mathbf{S}^T\mathbf{v}/\left\|\mathbf{S}^T\mathbf{v}\right\|$. We get

$$\cos\theta(\mathbf{v},\mathbf{S}) \geq \frac{\langle \mathbf{v}, \mathbf{S}\mathbf{S}^T\mathbf{v}\rangle}{\|\mathbf{S}\mathbf{S}^T\mathbf{v}\|_2} = \frac{\left\|\mathbf{S}^T\mathbf{v}\right\|_2^2}{\|\mathbf{S}\mathbf{S}^T\mathbf{v}\|_2} \geq \frac{\left\|\mathbf{S}^T\mathbf{v}\right\|_2}{\sigma_1(\mathbf{S})} = \frac{\sigma_1(\mathbf{S}^T\mathbf{v})}{\sigma_1(\mathbf{S})},$$

where the second inequality directly follows from the definitions of the largest singular value. Because $\mathbf{S}^T\mathbf{v} \sim \mathcal{N}(0,1)^{d\times 1}$, invoking Lemma 6 with $\delta = e^{-1}$ yields that $\mathbb{P}\left[\sigma_1(\mathbf{S}^T\mathbf{v}) \geq \sqrt{d}/e\right] \geq 1 - e^{-\Omega(d)}$. Meanwhile, Lemma 5 with $t = \sqrt{n} - \sqrt{d}$ implies that $\mathbb{P}[\sigma_1(\mathbf{S}) \le 2c\sqrt{n}] \ge 1 - e^{-\Omega(n)}$. We hence conclude (i) by applying the union bound.

(ii) $\cos\theta(\mathbf{v}, \mathbf{S}) = \mathcal{O}(\sqrt{d/n})$:

Due to $\sigma_d(\mathbf{S}) \le \|\mathbf{S}\|_2$ and $\langle \mathbf{v}, \mathbf{Sa}\rangle \le \left\|\mathbf{S}^T\mathbf{v}\right\|_2 \|\mathbf{a}\|_2 = \sigma_1(\mathbf{S}^T\mathbf{v})$, for all $\mathbf{a} \in \mathbb{S}^{d-1}$,

$$\cos\theta(\mathbf{v}, \mathbf{S}) = \max_{\mathbf{a}\in\mathbb{S}^{d-1}} \frac{\langle \mathbf{v}, \mathbf{Sa}\rangle}{\|\mathbf{Sa}\|_2} \le \frac{\sigma_1(\mathbf{S}^T\mathbf{v})}{\sigma_d(\mathbf{S})}.$$

For the denominator, Lemma 6 with $\delta = e^{-1}$ is applied to permit that $\mathbb{P}\left[\sigma_d(\mathbf{S}) \ge \frac{\sqrt{n}-\sqrt{d-1}}{e}\right] \ge 1 - e^{-\Omega(n-d+1)} - e^{-\Omega(n)}$. For the numerator, as $\mathbf{S}^T\mathbf{v} \sim \mathcal{N}(0,1)^{d\times 1}$, Lemma 5 with $t = \sqrt{d}$ shows that $\mathbb{P}\left[\sigma_1(\mathbf{S}^T\mathbf{v}) \le 2\sqrt{d}\right] \ge 1 - e^{-\Omega(d)}$. Thus, (ii) holds by applying the union bound. □

A.2 RSVD with positive semidefinite matrices

Lemma 7. *Let $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ and $\mathbf{y} = (\mathbf{y}_1, \ldots, \mathbf{y}_n)$ be two vectors in $\mathbb{R}^n$ satisfying (i) there exists $i \in [n]$ s.t. $\mathbf{x}_i\mathbf{y}_i \ne 0$, and (ii) there exists $j \in [n]$ s.t. $\mathbf{y}_j \ne 0$. Then for all $q \in \mathbb{N}$,*

$$\frac{\sum_{i=1}^n |\mathbf{x}_i|^{2q+1} \mathbf{y}_i^2}{\sum_{i=1}^n |\mathbf{x}_i|^{2q} \mathbf{y}_i^2} \ge \left(\frac{\sum_{i=1}^n |\mathbf{x}_i|^{2q} \mathbf{y}_i^2}{\sum_{i=1}^n \mathbf{y}_i^2}\right)^{\frac{1}{2q}}.$$

Proof For any n-dimensional vectors $\mathbf{a} = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$, $\mathbf{v} = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ satisfying that (i)$'$ there exists $i \in [n]$ s.t. $\mathbf{a}_i\mathbf{b}_i \ne 0$, and (ii)$'$ there exists $j \in [n]$ s.t. $\mathbf{b}_j \ne 0$, Hölder's inequality implies that

$$\left(\frac{\sum_{i=1}^n |\mathbf{a}_i|^r}{\sum_{i=1}^n |\mathbf{a}_i\mathbf{b}_i|}\right)^{\frac{1}{r}} \ge \left(\frac{\sum_{i=1}^n |\mathbf{a}_i\mathbf{b}_i|}{\sum_{i=1}^n |\mathbf{b}_i|^s}\right)^{\frac{1}{s}}, \tag{10}$$

where $r, s \in [1, \infty]$ with $1/r + 1/s = 1$. Let $\mathbf{a}_i = |\mathbf{x}_i|^{2q}\mathbf{y}_i^{2/r}$ and $\mathbf{b}_i = |\mathbf{y}_i|^{2/s}$, for all $i \in [n]$, then (i) and (ii) imply (i)$'$ and (ii)$'$ respectively. Hence, (10) with $r = (2q+1)/2q$, $s = 2q+1$ gives us that

$$\left(\frac{\sum_{i=1}^n \left(|\mathbf{x}_i|^{2q} \mathbf{y}_i^{\frac{4q}{2q+1}}\right)^{\frac{2q+1}{2q}}}{\sum_{i=1}^n |\mathbf{x}_i|^{2q} \mathbf{y}_i^2}\right)^{\frac{2q}{2q+1}} \ge \left(\frac{\sum_{i=1}^n |\mathbf{x}_i|^{2q} \mathbf{y}_i^2}{\sum_{i=1}^n \left(\mathbf{y}_i^{\frac{2}{2q+1}}\right)^{2q+1}}\right)^{\frac{1}{2q+1}}.$$

We conclude this lemma by rearranging the above inequality. □

Theorem 1. *Let $\mathbf{A}$ be a positive semidefinite matrix with $\lambda_1 > 0$ and $\hat{\mathbf{u}} = \mathsf{RSVD}(\mathbf{A}, \mathcal{N}(0,1)^{n\times d}, q, d)$ for any $q \in \mathbb{N}$. Then,*

$$R(\hat{\mathbf{u}}) = \left(\Omega\left(\frac{d}{n}\right)\right)^{\frac{1}{2q+1}}$$

holds with probability at least $1 - e^{-\Omega(d)}$.

Proof Thanks to Lemma 1, the proof follows if the following inequality holds almost surely

$$R(\hat{\mathbf{u}})^{2q+1} \ge \max_{\mathbf{a}\in\mathbb{S}^{d-1}} \frac{\langle \mathbf{S}^T\mathbf{u}_1, \mathbf{a}\rangle^2}{\sum_{i=1}^n \langle \mathbf{S}^T\mathbf{u}_i, \mathbf{a}\rangle^2} = \cos^2\theta(\mathbf{u}_1, \mathbf{S}), \tag{11}$$

where the equation is due to Definition 1. We show (11) by Lemma 7 and the alternating form of $R(\hat{\mathbf{u}})$ follows by (4) in Section 4.2,

$$R(\hat{\mathbf{u}}) = \max_{\mathbf{a}\in\mathbb{S}^{d-1}} R_{\mathbf{a}} = \max_{\mathbf{a}\in\mathbb{S}^{d-1}} \frac{\sum_{i=1}^n \alpha_i^{2q+1}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{a}\rangle^2}{\sum_{i=1}^n \alpha_i^{2q}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{a}\rangle^2}. \tag{12}$$

Let $\mathbf{x}_i = \alpha_i$ and $\mathbf{y}_i = \langle \mathbf{S}^T\mathbf{u}_i, \mathbf{a}\rangle$, for all $i \in [n]$, because $\langle \mathbf{S}^T\mathbf{u}_1, a\rangle \neq 0$ a.e., the conditions of Lemma 7, (i) and (ii)., hold a.e.. Therefore, it holds almost surely that

$$\begin{aligned} R_{\mathbf{a}} &= \frac{\sum_{i=1}^n \alpha_i^{2q+1}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{a}\rangle^2}{\sum_{i=1}^n \alpha_i^{2q}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{a}\rangle^2} \geq \left(\frac{\sum_{i=1}^n \alpha_i^{2q}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{a}\rangle^2}{\sum_{i=1}^n \langle \mathbf{S}^T\mathbf{u}_i, \mathbf{a}\rangle^2}\right)^{\frac{1}{2q}} \\ &= \left(\frac{\sum_{i=1}^n \alpha_i^{2q}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{a}\rangle^2}{\sum_{i=1}^n \alpha_i^{2q+1}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{a}\rangle^2}\frac{\sum_{i=1}^n \alpha_i^{2q+1}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{a}\rangle^2}{\sum_{i=1}^n \langle \mathbf{S}^T\mathbf{u}_i, \mathbf{a}\rangle^2}\right)^{\frac{1}{2q}} = \left(R_{\mathbf{a}}^{-1}\frac{\sum_{i=1}^n \alpha_i^{2q+1}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{a}\rangle^2}{\sum_{i=1}^n \langle \mathbf{S}^T\mathbf{u}_i, \mathbf{a}\rangle^2}\right)^{\frac{1}{2q}}, \end{aligned}$$

where the last equation follows from (4) in Section 4.2 again. Rearranging the above inequality, we get that

$$R_{\mathbf{a}}^{2q+1} \geq \frac{\sum_{i=1}^n \alpha_i^{2q+1}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{a}\rangle^2}{\sum_{i=1}^n \langle \mathbf{S}^T\mathbf{u}_i, \mathbf{a}\rangle^2} \geq \frac{\langle \mathbf{S}^T\mathbf{u}_1, \mathbf{a}\rangle^2}{\sum_{i=1}^n \langle \mathbf{S}^T\mathbf{u}_i, \mathbf{a}\rangle^2}, \text{ a.e.,} \tag{13}$$

where the second inequality is leveraged the fact that $\sum_{i\neq 1} \alpha_i^{2q+1}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{a}\rangle^2 \geq 0$. (13) and the definition $R(\hat{\mathbf{u}}) = \max_{\mathbf{a}\in\mathbb{S}^{d-1}} R_{\mathbf{a}}$ imply (11) as desired and hence the proof completes. □

Theorem 2. *For any $q \in \mathbb{N}$, there exists a positive semi-definite matrix $\mathbf{A}$ with $\lambda_1 > 0$, so that for $\hat{\mathbf{u}} = \mathsf{RSVD}(\mathbf{A}, \mathcal{N}(0,1)^{n\times d}, q, d)$, it holds*

$$R(\hat{\mathbf{u}}) = \mathcal{O}\left(\left(\frac{d}{n}\right)^{\frac{1}{2q+1}}\right),$$

with probability at least $1 - e^{-\Omega(d)}$.

Proof Let $\mathbf{A}$ be a diagonal matrix with $\mathbf{A}_{1,1} = 1$ and $\mathbf{A}_{i,i} = (d/n)^{\frac{1}{2q+1}}$, for all $i \neq 1$. Apparently, $\mathbf{A} = \mathbf{e}_1^T\mathbf{e}_1 + \sum_{i=2}^n \alpha\mathbf{e}_i^T\mathbf{e}_i$, where $\alpha = (d/n)^{\frac{1}{2q+1}}$ and $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is the canonical basis in $\mathbb{R}^n$. As discussed in Section 4, $R(\hat{\mathbf{u}}) = \max_{\mathbf{a}\in\mathbb{S}^{d-1}} R_{\mathbf{a}}$ and the alternating expression of $R_{\mathbf{a}}$, (4) in Section 4.2, can be rewritten as

$$\text{for all } \mathbf{a} \in \mathbb{S}^{d-1}, \quad R_{\mathbf{a}} = \frac{\langle \mathbf{S}^T\mathbf{e}_1, \mathbf{a}\rangle^2}{\langle \mathbf{S}^T\mathbf{e}_1, \mathbf{a}\rangle^2 + \sum_{i=2}^n \alpha^{2q}\langle \mathbf{S}^T\mathbf{e}_i, \mathbf{a}\rangle^2} + \frac{\sum_{i=2}^n \alpha^{2q+1}\langle \mathbf{S}^T\mathbf{e}_i, \mathbf{a}\rangle^2}{\langle \mathbf{S}^T\mathbf{e}_1, \mathbf{a}\rangle^2 + \sum_{i=2}^n \alpha^{2q}\langle \mathbf{S}^T\mathbf{e}_i, \mathbf{a}\rangle^2}. \tag{14}$$

On the one hand, as $1 > (d/n)^{\frac{2q}{2q+1}} = \alpha^{2q}$, the first term in (14) is upper bounded as:

$$\frac{\langle \mathbf{S}^T\mathbf{e}_1, \mathbf{a}\rangle^2}{\langle \mathbf{S}^T\mathbf{e}_1, a\rangle^2 + \sum_{i=2}^n \alpha^{2q}\langle \mathbf{S}^T\mathbf{e}_i, \mathbf{a}\rangle^2} \leq \frac{\langle \mathbf{S}^T\mathbf{e}_1, \mathbf{a}\rangle^2}{\sum_{i=1}^n \alpha^{2q}\langle \mathbf{S}^T\mathbf{e}_i, \mathbf{a}\rangle^2} \leq \alpha^{-2q}\cos^2\theta(\mathbf{e}_1, \mathbf{S}), \tag{15}$$

where the second inequality follows directly from the definition of $\cos^2\theta(\mathbf{e}_1, \mathbf{S})$. On the other hand, the second term in (14) is upper bounded as:

$$\frac{\sum_{i=2}^n \alpha^{2q+1}\langle \mathbf{S}^T\mathbf{e}_i, \mathbf{a}\rangle^2}{\langle \mathbf{S}^T\mathbf{e}_1, \mathbf{a}\rangle^2 + \sum_{i=2}^n \alpha^{2q}\langle \mathbf{S}^T\mathbf{e}_i, \mathbf{a}\rangle^2} \leq \frac{\sum_{i=2}^n \alpha^{2q+1}\langle \mathbf{S}^T\mathbf{e}_i, \mathbf{a}\rangle^2}{\sum_{i=2}^n \alpha^{2q}\langle \mathbf{S}^T\mathbf{e}_i, \mathbf{a}\rangle^2} = \alpha. \tag{16}$$

By substituting (15) and(16) into (14), we derive that $R_{\mathbf{a}} \leq \alpha^{-2q}\cos^2\theta(\mathbf{e}_1, \mathbf{S}) + \alpha$, for all $\mathbf{a} \in \mathbb{S}^{d-1}$, which provides an upper bound of $R(\hat{\mathbf{u}})$. Finally, invoking Lemma 1, which states that $\cos^2\theta(\mathbf{e}_1, \mathbf{S}) = \Theta(d/n)$ with high probability, and recalling that $\alpha = (d/n)^{\frac{1}{2q+1}}$ yields the conclusion. □

Theorem 3. *Let $\mathbf{A}$ be a positive semi-definite matrix, $\hat{\mathbf{u}} = \mathsf{RSVD}(\mathbf{A}, \mathcal{N}(0,1)^{n\times d}, q, d)$ for any $q \in \mathbb{N}$, and i_0 be defined as in Definition 2 in Section 4.2. Then*

$$R(\hat{\mathbf{u}}) = \Omega\left(\left(\frac{d}{d+i_0}\right)^{\frac{1}{2q+1}}\right)$$

holds with probability at least $1 - e^{-\Omega(d)}$.

Proof If $i_0 = n$, then we subsume the result by Theorem 1 directly. Hence we assume that $i_0 < n$ below. By applying Corollary 1 in Appendix C.2 with $\delta = \frac{1}{3}, \mathbf{x} = \mathbf{u}_1$, we have probability $1 - e^{-\Omega(d)}$ that

$$\frac{2d}{3} \le \left\|\mathbf{S}^T\mathbf{u}_1\right\|_2^2 \le \frac{4d}{3}, \tag{17}$$

which directly implies that $\mathbf{S}^T\mathbf{u}_1 \neq 0$. In the following, we consider (i). $\sum_{i=1}^{i_0} \alpha_i^{2q}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2 > \sum_{i=i_0+1}^{n} \alpha_i^{2q}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2$; (ii). otherwise. Then, we show the claimed lower bound in (i). and (ii). separately by invoking Lemma 12 in Appendix C.3, which gives the bounds for the weighted sum with high probability.

(i). $\sum_{i=1}^{i_0} \alpha_i^{2q}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2 > \sum_{i=i_0+1}^{n} \alpha_i^{2q}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2$.

Roughly speaking in (i), the top i_0 terms dominate, hence one can expect the similar proof for Theorem 1 without the last $n - i_0$ terms will help us reason. The alternating form of $R(\hat{\mathbf{u}})$ follows by (4) in Section 4.1,

$$R(\hat{\mathbf{u}}) = \max_{\mathbf{a}\in\mathbb{S}^{d-1}} \frac{\sum_{i=1}^{n} \alpha_i^{2q+1}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{a}\rangle^2}{\sum_{i=1}^{n} \alpha_i^{2q}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{a}\rangle^2} \ge \frac{\sum_{i=1}^{n} \alpha_i^{2q+1}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2}{\sum_{i=1}^{n} \alpha_i^{2q}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2} > \frac{\sum_{i=1}^{i_0} \alpha_i^{2q+1}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2}{2\sum_{i=1}^{i_0} \alpha_i^{2q}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2}, \tag{18}$$

where the first inequality comes from the fact that $\mathbf{S}^T\mathbf{u}_1/\left\|\mathbf{S}^T\mathbf{u}_1\right\|_2 \in \mathbb{S}^{d-1}$ and the last one uses that $\sum_{i=i_0+1}^{n} \alpha_i^{2q+1}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2 \ge 0$ and (i). From (18), we repeat the deduction of (6) in Section 4.1, by viewing $R_{\mathbf{a}}$ as $2R(\hat{\mathbf{u}})$, $\alpha_i = \alpha_i$ for $i = 1, \ldots, i_0$, $\alpha_i = 0$ for $i > i_0$, and $\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{a}\rangle^2 = \langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2$ to conclude that (an alternative way is to use Lemma 7 as shown Appendix A.2)

$$\frac{\sum_{i=1}^{i_0} \alpha_i^{2q+1}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2}{\sum_{i=1}^{i_0} \alpha_i^{2q}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2} > \left((2R(\hat{\mathbf{u}}))^{-1}\frac{\sum_{i=1}^{i_0} \alpha_i^{2q+1}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2}{\sum_{i=1}^{i_0}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2}\right)^{\frac{1}{2q}}. \tag{19}$$

Rearranging the inequalities (18) and (19), we get

$$(2R(\hat{\mathbf{u}}))^{2q+1} > \frac{\langle \mathbf{S}^T\mathbf{u}_1, \mathbf{S}^T\mathbf{u}_1\rangle^2}{\sum_{i=1}^{i_0}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2} \ge \frac{4d^2}{16d^2 + 9\sum_{1\le i\le i_0}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2}, \tag{20}$$

where the last inequality is a consequence of (17). By applying Lemma 12 in Appendix C.3 with $\epsilon = \frac{1}{3}, \delta = \frac{1}{3}$, $\beta_1 = \ldots = \beta_{i_0} = 1$ and $\beta_{i_0+1} = \ldots = \beta_n = 0$, then we have probability $1 - e^{-\Omega(d)}$ that $\sum_{1\le i\le i_0}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2 \le \frac{16di_0}{9}$. Together with (20), the proof is derived by the union bound.

(ii). $\sum_{i=1}^{i_0} \alpha_i^{2q}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2 \le \sum_{i=i_0+1}^{n} \alpha_i^{2q}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2$.

As $\mathbf{S}^T\mathbf{u}_1/\left\|\mathbf{S}^T\mathbf{u}_1\right\|_2 \in \mathbb{S}^{d-1}$, (4) in Section 4.2 yields that

$$R(\hat{\mathbf{u}}) \ge \frac{\sum_{i=1}^{n} \alpha_i^{2q+1}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2}{\sum_{i=1}^{n} \alpha_i^{2q}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2} \ge \frac{\langle \mathbf{S}^T\mathbf{u}_1, \mathbf{S}^T\mathbf{u}_1\rangle^2}{2\sum_{i=i_0+1}^{n} \alpha_i^{2q}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2} \ge \frac{2d^2}{9\sum_{i=i_0+1}^{n} \alpha_i^{2q}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2},$$

where the second inequality is due to (ii); the last is a result of (17) . By Lemma 12 with $\delta = d, \epsilon = \frac{1}{2}$, $\beta_2 = \ldots = \beta_{i_0} = 0$, and $\beta_i = \alpha_i^{2q}$ for all $i = i_0 + 1, \ldots, n$, we have

$$\mathbb{P}\left[\sum_{i=i_0+1}^{n} \alpha_i^{2q}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2 \le \frac{3d(d+1)}{2}\sum_{i=i_0+1}^{n}\alpha_i^{2q}\right] \le 1 - e^{-\Omega(d)}.$$

By Definition 2, since $\gamma > 1/q$,

$$\sum_{i=i_0+1}^{n} \alpha_i^{2q} \le C\int_1^\infty x^{-2q\gamma}dx < C\int_1^\infty x^{-2}dx = C.$$

Hence, the union bound yields $R(\hat{\mathbf{u}}) = \Omega(1)$ with probability at least $1 - e^{-\Omega(d)}$. □

A.3 RSVD with indefinite matrices

Assumption 1 is restated here for convenience.

Assumption 1. *Assume there exists a constant $\kappa \in (0,1]$ such that $\sum_{i=2}^n \lambda_i^{2q+1} \geq \kappa \sum_{i=2}^n |\lambda_i|^{2q+1}$.*

Lemma 2. *Assume that matrix $\mathbf{A}$ satisfies Assumption 1 and $\mathbf{S} \sim \mathcal{N}(0,1)^{n\times d}$. There exists a constant $c_\kappa \in (0,1]$ such that*

$$\mathbb{P}\left[\sum_{i=1}^n \lambda_i^{2q+1}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2 \geq c_\kappa \sum_{i=1}^n |\lambda_i|^{2q+1}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2\right] \geq 1 - e^{-\Omega(\sqrt{d}\kappa^2)}.$$

Proof Recall $\alpha_i = \lambda_i/\lambda_1$ for all $i \in [n]$ and introduce $\mathcal{I}_+ = \{i \in [n]\backslash\{1\} : \alpha_i > 0\}$ and $\mathcal{I}_- = \{i \in [n] : \alpha_i < 0\}$. It is natural to assume $\mathcal{I}_- \neq \emptyset$, as otherwise this lemma trivially holds. Also, for simplicity $\kappa \in (0,1]$ is assumed to be the number such that $\sum_{i=2}^n \alpha_i^{2q+1} = \kappa \sum_{i=2}^n |\alpha_i|^{2q+1}$ (this number can be found always).

Apparently in both sums of interest, $\sum_{i=1}^n \alpha_i^{2q+1}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2$ and $\sum_{i=1}^n |\alpha_i|^{2q+1}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2$, the largest single term is the first one, $\langle \mathbf{S}^T\mathbf{u}_1, \mathbf{S}^T\mathbf{u}_1\rangle^2$, so our initial step is to derive its high probability bound. Applying Corollary 1 in Appendix C.2 with $\mathbf{x} = \mathbf{u}_1$ and $\delta = 1 - \sqrt{1-\kappa/4}$ (resp. $\delta = \sqrt{1+\kappa/4} - 1$) for the lower-tail (resp. upper-tail) yields

$$\mathbb{P}\left[d^2\left(1 - \frac{\kappa}{4}\right) \leq \langle \mathbf{S}^T\mathbf{u}_1, \mathbf{S}^T\mathbf{u}_1\rangle^2 \leq d^2\left(1 + \frac{\kappa}{4}\right)\right] \geq 1 - e^{-\Omega(d\kappa^2)}. \tag{21}$$

However, as the other terms highly depend on the decay rate of eigenvalues, to derive a high probability bound, we need to carefully choose the parameters when applying concentration inequalities. In what follows, we define $s_+ = \sum_{i\in\mathcal{I}_+} \alpha_i^{2q+1}$ and $s_- = \sum_{i\in\mathcal{I}_-} |\alpha_i|^{2q+1}$, and then prove in two cases: either (i). $s_- = \Omega(\sqrt{d})$ or (ii). $s_- = o(\sqrt{d})$.

(i). $s_- = \Omega(\sqrt{d})$. Applying the lower-tail (resp. upper tail) of Lemma 12 in Appendix C.2 with $\delta = \epsilon = 1 - \sqrt{1-\kappa/2}$ (resp. $\delta = \epsilon = \sqrt{1+\kappa/2} - 1$), $\beta_i = \alpha_i$ for $i \in \mathcal{I}_+$, and $\beta_i = 0$ otherwise, we get

$$\mathbb{P}\left[d\left(1 - \frac{\kappa}{2}\right)s_+ \leq \sum_{i\in\mathcal{I}_+} \alpha_i^{2q+1}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2 \leq d\left(1 + \frac{\kappa}{2}\right)s_+\right] \geq 1 - e^{-\Omega(\sqrt{d}\kappa^2)}. \tag{22}$$

In addition, using Lemma 12 with $\delta = \epsilon = \sqrt{1+\kappa/2} - 1$ in Appendix C.2 , $\beta_i = |\alpha_i|$ for $i \in \mathcal{I}_-$, and $\beta_i = 0$ otherwise, we derive

$$\mathbb{P}\left[\sum_{i\in\mathcal{I}_-} |\alpha_i|^{2q+1}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2 \leq d\left(1 + \frac{\kappa}{2}\right)s_-\right] \geq 1 - e^{-\Omega(\sqrt{d}\kappa^2)}. \tag{23}$$

Now, we prove our assertion. The lower-tails in (21)(22) and the upper-tail in (23) implies

$$\begin{aligned}\sum_{i=1}^n \alpha_i^{2q+1}\langle \mathbf{S}\mathbf{u}_i, \mathbf{S}\mathbf{u}_1\rangle^2 &\geq d\left(d(1 - \frac{\kappa}{4}) + (1 - \frac{\kappa}{2})s_+ - (1 + \frac{\kappa}{2})s_-\right)\\ &\stackrel{(a)}{=} d\left(\frac{d(4-\kappa)}{4} + \frac{\kappa(s_+ + s_-)}{2}\right)\\ &\stackrel{(b)}{\geq} \frac{\kappa}{3}\left(d\left(d(1 + \frac{\kappa}{4}) + (1 + \frac{\kappa}{2})(s_+ + s_-)\right)\right) \stackrel{(c)}{\geq} \frac{\kappa}{3}\sum_{i=1}^n |\alpha_i|^{2q+1}\langle \mathbf{S}\mathbf{u}_i, \mathbf{S}\mathbf{u}_1\rangle^2,\end{aligned}$$

where (a) is due to $(1-\kappa)s_+ = (1+\kappa)s_-$ (rearranged from $\sum_{i=2}^n \alpha_i^{2q+1} = \kappa\sum_{i=2}^n |\alpha_i|^{2q+1}$), (b) is easily checked by comparing the coefficients, and (c) follows from the upper-tails in (21)(22)(23). Therefore, a union bound completes the proof with $c_\kappa = \frac{\kappa}{3}$ in this case.

(ii). $s_- = o(\sqrt{d})$. This is equivalent to say that there exists a constant $c > 0$ such that $s_- \le c\sqrt{d}$. Notice that from $\sum_{i=1}^n \alpha_i^{2q+1}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2$ to $\sum_{i=1}^n |\alpha_i|^{2q+1}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2$, only the term with index $i \in \mathcal{I}_-$ changes its sign and we show our assertion in the sense that the terms with indices in $\mathcal{I}_-$ do not affect too much with high probability.

Invoking Lemma 12 in Appendix C.3 with $\delta = \frac{\kappa\sqrt{d}}{8c}$, $\epsilon = \frac{\delta}{1+\delta}$, $\beta_i = |\alpha_i|$ for $i \in \mathcal{I}_-$, and $\beta_i = 0$ otherwise , we get

$$\mathbb{P}\left[\sum_{i\in\mathcal{I}_-} |\alpha_i|^{2q+1}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2 \le d\left(1 + \frac{\kappa\sqrt{d}}{4c}\right)s_-\right] \ge 1 - e^{-\Omega(\sqrt{d}\kappa^2)}. \tag{24}$$

On the one hand, the lower-tail in (21) and the upper-tail in (24) yield that with high probability

$$\sum_{i=1}^n \alpha_i^{2q+1}\langle \mathbf{S}\mathbf{u}_i, \mathbf{S}\mathbf{u}_1\rangle^2 \ge d^2\left(1 - \frac{\kappa}{4}\right) - d\left(1 + \frac{\kappa\sqrt{d}}{4c}\right)s_- + \sum_{i\in\mathcal{I}_+} \alpha_i^{2q+1}\langle \mathbf{S}\mathbf{u}_i, \mathbf{S}\mathbf{u}_1\rangle^2. \tag{25}$$

On the other hand, the upper-tails in (21) and (24) imply that with high probability

$$\sum_{i=1}^n |\alpha_i|^{2q+1}\langle \mathbf{S}\mathbf{u}_i, \mathbf{S}\mathbf{u}_1\rangle^2 \le d^2(1 + \frac{\kappa}{4}) + d(1 + \frac{\kappa\sqrt{d}}{4c})s_- + \sum_{i\in\mathcal{I}_+} \alpha_i^{2q+1}\langle \mathbf{S}\mathbf{u}_i, \mathbf{S}\mathbf{u}_1\rangle^2. \tag{26}$$

Finally with a union bound on (25) and (26) , we have probability at least $1 - e^{-\Omega(\sqrt{d}\kappa^2)}$ that for any $d \ge \left(\frac{14c}{10-7\kappa}\right)^2 = \Theta(1)$,

$$\begin{aligned}\sum_{i=1}^n \alpha_i^{2q+1}\langle \mathbf{S}\mathbf{u}_i, \mathbf{S}\mathbf{u}_1\rangle^2 - \frac{\sum_{i=1}^n |\alpha_i|^{2q+1}\langle \mathbf{S}\mathbf{u}_i, \mathbf{S}\mathbf{u}_1\rangle^2}{6} &\ge d^2\left(1 - \frac{\kappa}{4}\right) - d\left(1 + \frac{\kappa\sqrt{d}}{4c}\right)s_- - \frac{d^2(1+\frac{\kappa}{4}) + d(1+\frac{\kappa\sqrt{d}}{4c})s_-}{6} \\ &\ge \frac{(10-7\kappa)d^2 - 14cd\sqrt{d}}{12} \ge 0,\end{aligned}$$

where the second inequality stems from $s_- \le c\sqrt{d}$. Hence, the proof is completed with $c_\kappa = \frac{1}{6}$ in this case. □

Theorem 4. *Assume* $\mathbf{A}$ *satisfies Assumption 1. Let* $\hat{\mathbf{u}} = \mathsf{RSVD}(\mathbf{A}, \mathcal{N}(0,1)^{n\times d}, q, d)$ *for any* $q \in \mathbb{N}$*. Then,*

$$R(\hat{\mathbf{u}}) = \Omega\left(c_\kappa\left(\frac{d}{d+i_0}\right)^{\frac{1}{2q+1}}\right)$$

with probability at least $1 - e^{-\Omega(\sqrt{d}\kappa^2)}$.

Proof Evaluating $R_{\mathbf{a}}$ defined in (4) in Section 4.1 on $\mathbf{a} = \mathbf{S}^T\mathbf{u}_1 / \left\|\mathbf{S}^T\mathbf{u}_1\right\|_2$ and by Lemma 2 there exists a constant $c_\kappa \in (0, 1]$ such that

$$R_{\mathbf{a}} = \frac{\sum_{i=1}^n \alpha_i^{2q+1}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2}{\sum_{i=1}^n \alpha_i^{2q}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2} \ge c_\kappa \frac{\sum_{i=1}^n |\alpha_i|^{2q+1}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2}{\sum_{i=1}^n \alpha_i^{2q}\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2} = c_\kappa \bar{R}_{\mathbf{a}},$$

where $\bar{R}_{\mathbf{a}}$ is introduced in (9) in Section 4.3, with probability at least $1 - e^{-\Omega(\sqrt{d}\kappa^2)}$. Repeating the arguments in the proof of Theorem 3 in Appendix A.2 with replacing $R(\hat{\mathbf{u}})$ by $\bar{R}_{\mathbf{a}}$ yields:

$$\bar{R}_{\mathbf{a}} = \Omega\left(c_\kappa\left(\frac{d}{d+i_0}\right)^{\frac{1}{2q+1}}\right)$$

with probability at least $1 - e^{-\Omega(\sqrt{d}\kappa^2)}$, and hence the desired result follows by the union bound. □

Ruo-Chun Tzeng, Po-An Wang, Florian Adriaens, Aristides Gionis, Chi-Jen Lu

B Proofs of RandSum

B.1 Large deviation of projection length for Bernoulli random matrix

This subsection is used to prove Lemma 3, which serves as an intermediate step for Theorem 5, restated below. The proof relies on a simple but powerful concept, ε-net. As its usefulness, the definition and related theorems can be found in literature of random matrix. Here we shortly define it and state its important property below Lemma 3. Interested reader are referred to the reference therein.

Definition 3 (ε-net, Definition 4.2.1 in (Vershynin, 2018))**.** *Let $(\mathbb{S}^{d-1}, \|\cdot\|_2)$ be a metric space and $\varepsilon > 0$. A subset $\mathcal{N}_\varepsilon \subseteq \mathbb{S}^{d-1}$ is called ε-net if*

$$\forall x, y \in \mathcal{N}_\varepsilon, \|x - y\|_2 \leq \varepsilon.$$

Lemma 8 (Corrollary 4.2.13 in (Vershynin, 2018))**.** *For any $\varepsilon \in (0,1)$, the size of $\mathcal{N}_\varepsilon$ is bounded by*

$$|\mathcal{N}_\varepsilon| \leq 3^d \varepsilon^{-d}.$$

We are ready to prove Lemma 3 restated below.

Lemma 3. *Let $\mathbf{v} \in \mathbb{S}^{n-1}$, $d \leq n/3$, and $\mathbf{S} \sim$ Bernoulli$(p)^{n\times d}$ for a constant $p \in (0,1)$ Then,*

$$\cos^2 \theta(\mathbf{v}, \mathbf{S}) = \Omega\left(\frac{\max\{1, \langle \mathbf{v}, \mathbf{1}_n\rangle^2\}}{n}\right)$$

holds with probability at least $1 - e^{-\Omega(d)}$.

Proof As it is easy to see that $\mathbf{S}$ is a nonzero matrix with probability $1 - e^{-nd}$, the following deduction will be made under $\|\mathbf{S}\|_2 > 0$.
By the second inequality in Corollary 2 in Appendix C.2 with $\mathbf{x} = \mathbf{v}$ and $\delta = 1/2$, we deduce that

$$\mathbb{P}\left[\left\|\mathbf{S}^T\mathbf{v}\right\|_2 \geq \sqrt{\frac{dp(1-p+p\langle \mathbf{v}, \mathbf{1}_n\rangle^2)}{2}}\right] \geq 1 - e^{-\Omega(d)}. \tag{27}$$

Recall that $\cos\theta(\mathbf{v},\mathbf{S}) = \max_{\mathbf{a}\in\mathbb{S}^{d-1}} \frac{\langle \mathbf{v}, \mathbf{S}\mathbf{a}\rangle}{\|\mathbf{S}\mathbf{a}\|_2}$, (27) allows us to substitute $\mathbf{a} = \mathbf{S}^T\mathbf{v}/\left\|\mathbf{S}^T\mathbf{v}\right\|$ and have

$$\cos\theta(\mathbf{v},\mathbf{S}) \geq \frac{\left\|\mathbf{S}^T\mathbf{v}\right\|_2^2}{\left\|\mathbf{S}\mathbf{S}^T\mathbf{v}\right\|_2} \geq \frac{\left\|\mathbf{S}^T\mathbf{v}\right\|_2}{\|\mathbf{S}\|_2} \geq \frac{\sqrt{dp(1-p+p\langle \mathbf{v}, \mathbf{1}_n\rangle^2)}}{\sqrt{2}\,\|\mathbf{S}\|_2},$$

where the second inequality is due to submultiplicativity of $\|\cdot\|_2$, namely $\left\|\mathbf{S}\mathbf{S}^T\mathbf{v}\right\|_2 \leq \|\mathbf{S}\|_2 \left\|\mathbf{S}^T\mathbf{v}\right\|_2$, and the last one is a consequence of (27). It remains to show that $\|\mathbf{S}\|_2 \leq \mathcal{O}\left(\sqrt{nd}\right)$ w.h.p., then the proof is done. For this goal, we use the ε-net technique, introduced in the beginning of this subsection, and give a bound in two steps:

(i). Let $\mathcal{N}_\varepsilon$ be an ε-net defined on $(\mathbb{S}^{d-1}, \|\cdot\|_2)$ for some $\varepsilon \in (0,1)$ to be determined later. We claim that

$$\|\mathbf{S}\|_2 \leq \frac{1}{1-\varepsilon} \sup_{\mathbf{x}\in\mathcal{N}_\varepsilon} \|\mathbf{S}\mathbf{x}\|_2. \tag{28}$$

Let $\mathbf{w}^* \in \operatorname{argmax}_{\mathbf{x}\in\mathbb{S}^{d-1}} \|\mathbf{S}\mathbf{x}\|_2$, and since there exists $\mathbf{x}^* \in \mathcal{N}_\varepsilon$ satisfying $\|\mathbf{w}^* - \mathbf{x}^*\|_2 \leq \varepsilon$, by submultiplicativity and triangle inequality, we get

$$\varepsilon\|\mathbf{S}\|_2 \geq \|\mathbf{S}(\mathbf{w}^* - \mathbf{x}^*)\|_2 \geq \|\mathbf{S}\|_2 - \|\mathbf{S}\mathbf{x}^*\|_2 \geq \|\mathbf{S}\|_2 - \sup_{\mathbf{x}\in\mathcal{N}_\varepsilon} \|\mathbf{S}\mathbf{x}\|_2,$$

and rearranging the terms yields (28).

(ii). Show that

$$\mathbb{P}\left[\sup_{\mathbf{x}\in\mathcal{N}_\varepsilon}\|\mathbf{S}\mathbf{x}\|_2 \le \left(\frac{3}{2}np(1-p+pd)\right)^{\frac{1}{2}}\right] \ge 1-3^d\varepsilon^{-d}e^{-\Omega(n)} \ge 1-e^{-\Omega(n+d\ln\frac{\varepsilon}{3})}. \tag{29}$$

For each $\mathbf{x}\in\mathcal{N}_\varepsilon$, the first inequality in Corollary 2 in Appendix C.2 with $\mathbf{x}=\mathbf{x}$ and $\delta=\frac{1}{2}$ (here n and d are reversed) implies that we have probability at least $1-e^{-\Omega(n)}$

$$\|\mathbf{S}\mathbf{x}\|_2 \le \left(\frac{3}{2}np(1-p+p\langle\mathbf{x},\mathbf{1}_d\rangle^2)\right)^{\frac{1}{2}} \le \left(\frac{3}{2}np(1-p+pd)\right)^{\frac{1}{2}},$$

where the last inequality is due to $\langle\mathbf{x},\mathbf{1}_d\rangle^2\le d$. As the size of $\mathcal{N}_\varepsilon$ is upper bounded by $3^d\varepsilon^{-d}$ (see Lemma 8 on the top of this subsection), the union bound over all $\mathbf{x}\in\mathcal{N}_\varepsilon$ yields (29).

Finally, setting $\varepsilon=1/e$ in (28)-(29) and assumption $n-2d>d$ lead to $\|\mathbf{S}\mathbf{x}\|_2\le\mathcal{O}\left(\sqrt{nd}\right)$ holds with probability at least $1-e^{-\Omega(n-2d)}>1-e^{-\Omega(d)}$. The union bound completes our proof as desired. □

B.2 RandSum with positive semidefinite matrices

Theorem 5. *Let* $\mathbf{A}$ *be a positive semi-definite matrix with* $\lambda_1>0$ *and* $\hat{\mathbf{u}}=\mathsf{RandSum}(\mathbf{A},q,d,p)$ *for any constant* $p\in(0,1)$*, any* $q\in\mathbb{N}$*, and* $d\ge 2$*. Then,*

$$R(\hat{\mathbf{u}}) = \left(\Omega\left(\frac{\max\{d,\langle\mathbf{u}_1,\mathbf{1}_n\rangle^2\}}{n}\right)\right)^{\frac{1}{2q+1}}$$

with probability at least $1-e^{-\Omega(d)}$.

Proof Define $\mathcal{A}_1=\left\{\begin{bmatrix}\mathbf{a_1}\\ \mathbf{0}_{\lfloor\frac{d}{2}\rfloor}\end{bmatrix}:\mathbf{a_1}\in\mathbb{S}^{\lceil\frac{d}{2}\rceil-1}\right\}$ and $\mathcal{A}_2=\left\{\begin{bmatrix}\mathbf{0}_{\lceil\frac{d}{2}\rceil}\\ \mathbf{a_2}\end{bmatrix}:\mathbf{a_2}\in\mathbb{S}^{\lfloor\frac{d}{2}\rfloor-1}\right\}$. Since $R(\hat{\mathbf{u}})\ge\max_{\mathbf{a}\in\mathbb{S}^{d-1}}R_{\mathbf{a}}$, where $R_{\mathbf{a}}$ is introduce on Section 4.1 and has an expression (4), we can conclude that

$$R(\hat{\mathbf{u}}) \ge \max\left\{\max_{\mathbf{a}\in\mathcal{A}_1}R_{\mathbf{a}},\max_{\mathbf{a}\in\mathcal{A}_2}R_{\mathbf{a}}\right\} \ge \max\left\{\cos^2\theta(\mathbf{u}_1,\mathbf{S_1}),\cos^2\theta(\mathbf{u}_1,\mathbf{S_2})\right\}^{\frac{1}{2q+1}},$$

where the last inequality is an application of (7) in Section 4.1. The proof is completed by Lemma 1 and Lemma 3. □

B.3 RandSum with indefinite matrices

Assumption 2 is restated here for convenience.

Assumption 2. *Assume that (i)* $\langle\mathbf{u}_1,\mathbf{1}_n\rangle^2=\Omega(1)$ *and (ii) there exists a constant* $\kappa'\in(0,1]$ *such that*

$$\sum_{i=2}^n\lambda_i^{2q+1}\xi_i \ge \kappa'\sum_{i=2}^n|\lambda_i|^{2q+1}\xi_i,$$

where $\xi_i=\mathbb{E}\left[\langle\mathbf{S}^T\mathbf{u}_i,\frac{\mathbf{1}_d}{\sqrt{d}}\rangle^2\right]=p(1-p+pd\langle\mathbf{u}_i,\mathbf{1}_n\rangle^2)$, $\forall i\in[n]$.

Lemma 4. *Assume that* $\mathbf{A}$ *satisfies Assumption 2. Let* $\mathbf{S}\sim$ *Bernoulli*$(p)^{n\times d}$ *for a constant* $p\in(0,1)$*. There exists a constant* $c_{\kappa'}\in(0,1]$ *such that*

$$\mathbb{P}\left[\sum_{i=1}^n\lambda_i^{2q+1}\langle\mathbf{S}^T\mathbf{u}_i,\mathbf{S}^T\mathbf{u}_1\rangle^2 \ge c_{\kappa'}\sum_{i=1}^n|\lambda_i|^{2q+1}\langle\mathbf{S}^T\mathbf{u}_i,\mathbf{S}^T\mathbf{u}_1\rangle^2\right] \ge 1-e^{-\Omega(\sqrt{d}\kappa'^2)}.$$

Proof Here we introduce

$$\mu_i = \mathbb{E}\left[\left\|\mathbf{S}_{:,1}^T\mathbf{u}_i\right\|_2^2\right] = p(1-p+p\langle\mathbf{u}_i,\mathbf{1}_n\rangle^2),\ \forall i\in[n].$$

Recall that $\alpha_i = \lambda_i/\lambda_1$ for all $i\in[n]$. From Assumption 2, we have:

- By (i), there exists a constant $\nu\in(0,1]$ such that $\langle\mathbf{u}_1,\mathbf{1}_n\rangle^2 \geq \nu$. It follows that

$$\xi_1 \geq p^2 d\langle\mathbf{u}_1,\mathbf{1}_n\rangle^2 = pd\cdot p((1-p)\langle\mathbf{u}_1,\mathbf{1}_n\rangle^2 + p\langle\mathbf{u}_1,\mathbf{1}_n\rangle^2) \geq p\nu d\mu_1. \tag{30}$$

- By (ii), there exists $\kappa'\in(0,1]$ such that $\sum_{i=2}^n \alpha_i^{2q+1}\xi_i = \kappa'\sum_{i=2}^n|\alpha_i|^{2q+1}\xi_i$.

We then partition $[n]$ into three subsets, $[n] = \{1\}\cup\mathcal{I}_+\cup\mathcal{I}_-$, where $\mathcal{I}_+ = \{i\in[n]\backslash\{1\} : \alpha_i > 0\}$ and $\mathcal{I}_- = \{i\in[n] : \alpha_i < 0\}$. It is natural to assume $\mathcal{I}_-\neq\emptyset$, as otherwise this lemma trivially holds. As similar to what we proceed in the proof of Lemma 2 in Appendix A.3, two important quantities follows from this partition: $s_+ = \sum_{i\in\mathcal{I}_+}\alpha_i^{2q+1}\xi_i$ and $s_- = \sum_{i\in\mathcal{I}_-}|\alpha_i|^{2q+1}\xi_i$.

Firstly, for the term $\langle\mathbf{S}^T\mathbf{u}_1,\mathbf{S}^T\mathbf{u}_1\rangle^2$, applying Corollary 2 in Appendix C.2 with $\mathbf{x}=\mathbf{u}_1$ and $\delta = 1-\sqrt{1-\kappa'/4}$ (resp. $\delta=\sqrt{1+\kappa'/4}-1$) for the lower-tail (resp. upper-tail) yields that

$$\mathbb{P}\left[d^2\mu_1^2\left(1-\frac{\kappa'}{4}\right) \leq \langle\mathbf{S}^T\mathbf{u}_1,\mathbf{S}^T\mathbf{u}_1\rangle^2 \leq d^2\mu_1^2\left(1+\frac{\kappa'}{4}\right)\right] \geq 1-e^{-\Omega(d\kappa'^2)}. \tag{31}$$

As for the remaining terms, we carefully apply concentration inequalities under two scenarios: either (i). $s_- = \Omega(\sqrt{d})$ or (ii). $s_- = o(\sqrt{d})$.

<u>(i). $s_- = \Omega(\sqrt{d})$</u>. Invoking the lower-tail (resp. upper-tail) of Lemma 13 in C.3 with $\delta=\epsilon=\sqrt{1+\kappa'/2}-1$ (resp. $\delta=\epsilon=\sqrt{1-\kappa'/2}-1$), $\beta_i=\alpha_i$ for $i\in\mathcal{I}_+$, and $\beta_i = 0$ otherwise, we get

$$\mathbb{P}\left[\xi_1\left(1-\frac{\kappa'}{2}\right)s_+ \leq \sum_{i\in\mathcal{I}_+}\alpha_i^{2q+1}\langle\mathbf{S}^T\mathbf{u}_i,\mathbf{S}^T\mathbf{u}_1\rangle^2 \leq d\mu_1\left(1+\frac{\kappa'}{2}\right)s_+\right] \geq 1-e^{-\Omega(\sqrt{d}\kappa'^2)}. \tag{32}$$

Again, using Lemma 13 with $\delta=\epsilon=\sqrt{1+\kappa'/2}-1$, $\beta_i=|\alpha_i|$ for $i\in\mathcal{I}_-$, and $\beta_i=0$ otherwise leads to

$$\mathbb{P}\left[\sum_{i\in\mathcal{I}_-}|\alpha_i|^{2q+1}\langle\mathbf{S}^T\mathbf{u}_i,\mathbf{S}^T\mathbf{u}_1\rangle^2 \leq d\mu_1\left(1+\frac{\kappa'}{2}\right)s_-\right] \geq 1-e^{-\Omega(\sqrt{d}\kappa'^2)}. \tag{33}$$

Now, we prove our assertion. The lower-tails in (31)(32) and upper-tail in (33) imply that

$$\begin{aligned}
\sum_{i=1}^n \alpha_i^{2q+1}\langle\mathbf{S}\mathbf{u}_i,\mathbf{S}\mathbf{u}_1\rangle^2 &\geq d\mu_1\left(d\mu_1(1-\frac{\kappa'}{4}) + \frac{\xi_1}{d\mu_1}(1-\frac{\kappa'}{2})s_+ - (1+\frac{\kappa'}{2})s_-\right)\\
&\overset{(a)}{\geq} p\nu\cdot d\mu_1\left(d\mu_1(1-\frac{\kappa'}{4}) + (1-\frac{\kappa'}{2})s_+ - (1+\frac{\kappa'}{2})s_-\right)\\
&\overset{(b)}{=} p\nu\cdot d\mu_1\left(\frac{d(4-\kappa')\mu_1}{4} + \frac{\kappa'(s_++s_-)}{2}\right)\\
&\overset{(c)}{\geq} \frac{p\nu\kappa'}{3}\left(d\mu_1\left(\frac{d(4+\kappa')\mu_1}{4} + (1+\frac{\kappa'}{2})(s_++s_-)\right)\right) \overset{(d)}{\geq} \frac{p\nu\kappa'}{3}\sum_{i=1}^n|\alpha_i|^{2q+1}\langle\mathbf{S}\mathbf{u}_i,\mathbf{S}\mathbf{u}_1\rangle^2,
\end{aligned}$$

where (a) uses (30), (b) is due to $(1-\kappa')s_+ = (1+\kappa')s_-$ (rearranged from $\sum_{i\neq1}\alpha_i^{2q+1}\xi_i = \kappa'\sum_{i\neq1}|\alpha_i|^{2q+1}\xi_i$), (c) is easily checked by comparing the coefficients, and (d) follows from the upper-tails in (31)(32)(33). Therefore, a union bound completes the proof with $c_{\kappa'} = \frac{p\nu\kappa'}{3}$ in this case.

(ii). $s_- = o(\sqrt{d})$. There exists a constant $c > 0$ such that $s_- \leq c\sqrt{d}$. Observe that for two summations of interest, $\sum_{i=1}^{n} \alpha_i^{2q+1} \langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1 \rangle^2$ and $\sum_{i=1}^{n} |\alpha_i|^{2q+1} \langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1 \rangle^2$, only the terms in $\mathcal{I}_-$ change their signs. Our assertion follows in the sense that the terms with indices in $\mathcal{I}_-$ do not affect too much with high probability.

Invoking Lemma 13 in C.3 with $\delta = \frac{\kappa'(\sqrt{d}-1)\mu_1}{4c}$, $\epsilon = \frac{\kappa'\mu_1}{4c(1+\delta)}$, $\beta_i = |\alpha_i|$ for $i \in \mathcal{I}_-$, and $\beta_i = 0$ otherwise , we get

$$\mathbb{P}\left[\sum_{i \in \mathcal{I}_-} |\alpha_i|^{2q+1} \langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1 \rangle^2 \leq d\mu_1 \left(1 + \frac{\kappa'\sqrt{d}\mu_1}{4c}\right) s_- \right] \geq 1 - e^{-\Omega(\sqrt{d}\kappa'^2)}. \tag{34}$$

On the one hand, the lower-tail in (31) and the upper-tail in (34) yield that

$$\sum_{i=1}^{n} \alpha_i^{2q+1} \langle \mathbf{S}\mathbf{u}_i, \mathbf{S}\mathbf{u}_1 \rangle^2 \geq d^2\mu_1^2 \left(1 - \frac{\kappa'}{4}\right) - d\mu_1 \left(1 + \frac{\kappa'\sqrt{d}\mu_1}{4c}\right) s_- + \sum_{i \in \mathcal{I}_+}^{n} \alpha_i^{2q+1} \langle \mathbf{S}\mathbf{u}_i, \mathbf{S}\mathbf{u}_1 \rangle^2 \tag{35}$$

On the other hand, the upper-tails in (31)(34) imply that

$$\sum_{i=1}^{n} |\alpha_i|^{2q+1} \langle \mathbf{S}\mathbf{u}_i, \mathbf{S}\mathbf{u}_1 \rangle^2 \leq d^2\mu_1^2 (1 + \frac{\kappa'}{4}) + d\mu_1 (1 + \frac{\kappa'\sqrt{d}\mu_1}{4c}) s_- + \sum_{i \in \mathcal{I}_+}^{n} \alpha_i^{2q+1} \langle \mathbf{S}\mathbf{u}_i, \mathbf{S}\mathbf{u}_1 \rangle^2 \tag{36}$$

As a consequence of a union bound on (35)(36), we have with probability at least $1 - e^{-\Omega(\sqrt{d}\kappa'^2)}$,

$$\begin{aligned}
&\sum_{i=1}^{n} \alpha_i^{2q+1} \langle \mathbf{S}\mathbf{u}_i, \mathbf{S}\mathbf{u}_1 \rangle^2 - \frac{1}{6} \cdot \sum_{i=1}^{n} |\alpha_i|^{2q+1} \langle \mathbf{S}\mathbf{u}_i, \mathbf{S}\mathbf{u}_1 \rangle^2 \\
&\geq d^2\mu_1^2 \left(1 - \frac{\kappa'}{4}\right) - d\mu_1 \left(1 + \frac{\kappa'\sqrt{d}\mu_1}{4c}\right) s_- - \frac{d^2\mu_1^2(1 + \frac{\kappa'}{4}) + d\mu_1(1 + \frac{\kappa'\sqrt{d}\mu_1}{4c}) s_-}{6} \\
&\geq \frac{1}{12} \left((10 - 7\kappa') d^2\mu_1^2 - 14cd\sqrt{d}\mu_1 \right) \geq 0,
\end{aligned}$$

where the second inequality is due to $s_- \leq c\sqrt{d}$, for any $d \geq \left(\frac{14c}{(10-7\kappa')\mu_1}\right)^2 = \Theta(1)$. Hence, the proof is completed with $c_{\kappa'} = \frac{1}{6}$ in this case. □

Theorem 6. *Assume that* $\mathbf{A}$ *satisfies Assumptions 1 and 2. Let* $\hat{\mathbf{u}} = \mathsf{RandSum}(\mathbf{A}, q, d, p)$ *for any constant* $p \in (0, 1)$ *and any* $q \in \mathbb{N}$, *and* i_0 *be defined as in Definition 2 in Section 4.2. Then,*

$$R(\hat{\mathbf{u}}) = \Omega \left(\max \left\{ c_\kappa \left(\frac{d}{d + i_0}\right)^{\frac{1}{2q+1}}, c_{\kappa'} \left(\frac{\max\{d, \langle \mathbf{u}_1, \mathbf{1}_n \rangle^2\}}{n}\right)^{\frac{1}{2q+1}} \right\} \right)$$

with probability at least $1 - e^{-\Omega(\sqrt{d}\min(\kappa,\kappa')^2)}$.

Proof Let

$$\mathbf{a}_1 = \begin{bmatrix} \frac{\mathbf{S}_1^T\mathbf{u}_1}{\|\mathbf{S}_1^T\mathbf{u}_1\|_2} \\ \mathbf{0}_{\lfloor \frac{d}{2} \rfloor} \end{bmatrix} \quad \text{and} \quad \mathbf{a}_2 = \begin{bmatrix} \mathbf{0}_{\lceil \frac{d}{2} \rceil} \\ \frac{\mathbf{S}_2^T\mathbf{u}_1}{\|\mathbf{S}_2^T\mathbf{u}_1\|_2} \end{bmatrix}.$$

A union bound of Lemma 2 and Lemma 4 implies that there exist constants c_κ and $c_{\kappa'}$ such that

$$R_{\mathbf{a}_1} \geq c_\kappa \bar{R}_{\mathbf{a}_1} \quad \text{and} \quad R_{\mathbf{a}_2} \geq c_{\kappa'} \bar{R}_{\mathbf{a}_2},$$

with probability at least $1 - e^{-\Omega(\sqrt{d}\kappa^2)} - e^{-\Omega(\sqrt{d}\kappa'^2)}$, where $R_{\mathbf{a}}$ and $\bar{R}_{\mathbf{a}}$ are defined in (4) in Section 4.1 and (9) in Section 4.3, respectively. Hence,

$$\mathbb{P}\left[R(\hat{\mathbf{u}}) \geq \max\left\{c_\kappa \bar{R}_{\mathbf{a}_1}, c_{\kappa'} \bar{R}_{\mathbf{a}_2}\right\}\right] \geq 1 - e^{-\Omega(\sqrt{d}\min(\kappa,\kappa')^2)}.$$

Finally, applying similar argument in the proof of Theorem 3 (see Appendix A.2) to lower bound $\bar{R}_{\mathbf{a}_1}$ and Theorem 5 to lower bound $\bar{R}_{\mathbf{a}_2}$ completes the proof. □

C Concentration inequalities

Before showing our lemmas on both Gaussian and Bernoulli random variables, there are some necessary definition and standard concentration inequalities to be introduced. For the random variables considered in this work, sub-gaussian and sub-exponential norms are useful to quantify the probabilities of rare events. In C.1, we introduce them for completeness and list the concentration inequalities (Hoeffding, Bernstein, and Hanson-Wright inequalities) used in the following proofs. In C.2, we provide two corollaries yielded by Bernstein inequality for Gaussian and Bernoulli distributions respectively. Finally, our technical lemmas for these two random variables will be shown in C.3.

C.1 Sub-gaussian norm and sub-exponential norm

Definition 4 (Definition 2.5.6 (Vershynin, 2018))**.** *The sub-gaussian norm $\|\cdot\|_{\psi_2}$ is a norm on the space of sub-gaussian random variables. For any sub-gaussian random variable X,*

$$\|X\|_{\psi_2} = \inf\{t > 0 : \mathbb{E}\left[\exp\left(X^2/t^2\right)\right] \leq 2\}.$$

The sum of sub-gaussian random variables is still a sub-gaussian random variable, and its norm can be characterized by the following Proposition.

Proposition 1 (Proposition 2.6.1 (Vershynin, 2018))**.** *Let $X_1, \cdots, X_m$ be a zero-mean sub-gaussian random variables. Then,*

$$\left\| \sum_{i\in[m]} X_i \right\|_{\psi_2}^2 = \mathcal{O}\left(\sum_{i\in[m]} \|X_i\|_{\psi_2}^2 \right),$$

where $\mathcal{O}$ hides an absolute constant.

Definition 5 (Definition 2.7.5 (Vershynin, 2018))**.** *The sub-exponential norm $\|\cdot\|_{\psi_1}$ is a norm on the space of sub-exponential random variables. For any sub-exponential random variable X,*

$$\|X\|_{\psi_1} = \inf\{t > 0 : \mathbb{E}\left[\exp\left(|X|/t\right)\right] \leq 2\}.$$

If X is sub-gaussian random variable, then X is also a sub-exponential random variable. Besides, there is one well-known property for these two norms.

Proposition 2 (Lemma 2.7.6 (Vershynin, 2018))**.** *Let X be a zero-mean sub-gaussian random variable. Then,*

$$\|X\|_{\psi_2}^2 = \left\|X^2\right\|_{\psi_1}.$$

For concreteness, we compute sub-gaussian norms for two basic variables.

Example 1. *Here we evaluate the values of $\|\cdot\|_{\psi_2}$ and $\|\cdot\|_{\psi_1}$ for the sub-gaussian random variables which will be used later*

- *If $X \sim \mathcal{N}(0, \sigma^2)$, for some $\sigma \in \mathbb{R}_+$, then $\|X\|_{\psi_2} = 2\sigma$.*
- *If $Y \sim$ Bernoulli(p), for some $p \in (0, 1)$, then $\|Y\|_{\psi_2} = \frac{1}{\sqrt{\ln(1+p^{-1})}}$ and $\|Y\|_{\psi_1} = \frac{1}{\ln(1+p^{-1})}$.*

Proof For any $t > \sqrt{2}\sigma$, we observe that

$$\mathbb{E}\left[\exp\left(X^2/t^2\right)\right] = \frac{1}{\sigma\sqrt{2\pi}}\int_{x\in\mathbb{R}} \exp\left(-\frac{x^2}{2\sigma^2}+\frac{x^2}{t^2}\right)dx = \frac{1}{\sigma\sqrt{\frac{1}{2\sigma^2}-\frac{1}{t^2}}},$$

which is 2 when $t = 2\sigma$, hence $\|X\|_{\psi_2} = 2\sigma$. As for Y, elementary calculus shows that

$$\|Y\|_{\psi_2} = \inf\left\{t > 0 : p\exp(t^{-2}) + (1-p) \leq 2\right\} = \inf\left\{t > 0 : \exp(t^{-2}) \leq \frac{1+p}{p}\right\} = \frac{1}{\sqrt{\ln(1+p^{-1})}},$$

and that

$$\|Y\|_{\psi_1} = \inf\left\{t > 0 : p\exp(t^{-1}) + (1-p) \leq 2\right\} = \inf\left\{t > 0 : \exp(t^{-1}) \leq \frac{1+p}{p}\right\} = \frac{1}{\ln(1+p^{-1})}.$$

□

Here is the list of concentration inequalities we will use later. The first proposition is an immediate result from Definition 4 and 5, the others are standard concentration inequalities characterized by these two norms.

Proposition 3 (Proposition 2.5.2 and Proposition 2.7.1 in (Vershynin, 2018))**.** *Let X and Y be a sub-gaussian and a sub-exponential random variables, respectively. Then for any $t \geq 0$, we have*

$$\mathbb{P}\left[|X - \mathbb{E}[X]| \geq t\right] \leq \exp\left(-\Omega\left(\frac{t^2}{\|X\|_{\psi_2}^2}\right)\right) \quad \textit{and} \quad \mathbb{P}\left[|Y - \mathbb{E}[Y]| \geq t\right] \leq \exp\left(-\Omega\left(\frac{t}{\|Y\|_{\psi_1}}\right)\right).$$

Lemma 9 (Hoeffding's inequality (Theorem 2.6.3 in (Vershynin, 2018)))**.** *Let $m \in \mathbb{N}$, $X_1, \cdots, X_m$ be i.i.d. zero-mean sub-gaussian random variables, and $\mathbf{a} \in \mathbb{R}^m$ be a nonzero vector. Then,*

$$\forall t \geq 0, \quad \mathbb{P}\left[\left|\sum_{i=1}^{m} \mathbf{a}_i X_i\right| > t\right] \leq \exp\left(-\Omega\left(\frac{t^2}{K\|\mathbf{a}\|_2^2}\right)\right),$$

where $K = \|\mathbf{X}_1\|_{\psi_2}^2$.

Lemma 10 (Bernstein's inequality (Theorem 2.8.2 in (Vershynin, 2018)))**.** *Let $m \in \mathbb{N}$ and $\mathbf{a} = (\mathbf{a}_1, \cdots, \mathbf{a}_m) \in \mathbb{R}^m \setminus \{\mathbf{0}_n\}$. Let $X_1, \cdots, X_m$ be independent sub-gaussian r.v.'s. Then there exists a universal constant $c > 0$ such that for any $t > 0$,*

$$\mathbb{P}\left[\left|\sum_{i=1}^{m} \mathbf{a}_i\left(X_i^2 - \mathbb{E}\left[X_i^2\right]\right)\right| \geq t\right] \leq 2\exp\left(-c\min\left\{\frac{t^2}{K^2\|\mathbf{a}\|_2^2}, \frac{t}{K\|\mathbf{a}\|_\infty}\right\}\right),$$

where $K = \max_{i\in[m]} \left\|X_i^2 - \mathbb{E}\left[X_i^2\right]\right\|_{\psi_1}$.

Lemma 11 (Hanson-Wright inequality (Theorem 6.2.1 in (Vershynin, 2018)))**.** *Let $m \in \mathbb{N}$ and $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_m)$ be a random vector with i.i.d zero-mean sub-gaussian entries and $\mathbf{M} \in \mathbb{R}^{m\times m} \setminus \{\mathbf{0}_{m\times m}\}$. Then,*

$$\forall t > 0, \quad \mathbb{P}\left[\left|\sum_{i,j\in[m]} \mathbf{M}_{i,j}\mathbf{X}_i\mathbf{X}_j - \mathbb{E}\left[\sum_{i,j\in[m]} \mathbf{M}_{i,j}\mathbf{X}_i\mathbf{X}_j\right]\right| > t\right] \leq \exp\left(-\Omega\left(\min\left\{\frac{t^2}{K^2\|\mathbf{M}\|_F^2}, \frac{t}{K\|\mathbf{M}\|_2}\right\}\right)\right),$$

where $K = \|\mathbf{X}_1\|_{\psi_2}^2$.

C.2 Useful lemmas derived from Bernstein's inequality

In this subsection, we will use Lemma 10 in C.1 to derive two Bernstein-type concentration inequalities. Corollary 1 (resp. Corollary 2) provides tail bounds on the length $\|\mathbf{Sx}\|_2$ of Gaussian (resp. Bernoulli) random matrix S with linear combination weights $\mathbf{x}$ of its columns.

Corollary 1. *Let* $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}_n\}$ *and* $\mathbf{S} \sim \mathcal{N}(0,1)^{n\times d}$*. Then,* $\forall \delta > 0$*,*

$$\mathbb{P}\left[\left\|\mathbf{S}^T\mathbf{x}\right\|_2^2 \geq d(1+\delta)\left\|\mathbf{x}\right\|_2^2\right] \leq e^{-\Omega\left(d\min\left\{\delta,\delta^2\right\}\right)}, \quad \text{and} \quad \mathbb{P}\left[\left\|\mathbf{S}^T\mathbf{x}\right\|_2^2 \leq d(1-\delta)\left\|\mathbf{x}\right\|_2^2\right] \leq e^{-\Omega\left(d\min\left\{\delta,\delta^2\right\}\right)}.$$

Proof For each $i = 1, \ldots, d$, the i-th column of $\mathbf{S}$ is denoted as $\mathbf{S}_{:,i}$. Because $\langle \mathbf{S}_{:,1}, \frac{\mathbf{x}}{\|\mathbf{x}\|_2}\rangle, \ldots, \langle \mathbf{S}_{:,d}, \frac{\mathbf{x}}{\|\mathbf{x}\|_2}\rangle$ are i.i.d. random variable drawn from $\mathcal{N}(0,1)$, the application of Lemma 10 with $m = d$, $\mathbf{a} = \mathbf{1}_d$, $t = \delta d$, and $X_i = \langle \mathbf{S}_{:,i}, \mathbf{x}/\left\|\mathbf{x}\right\|_2\rangle$ for $i = 1, \ldots, d$, implies that there is a universal constant $c > 0$ such that

$$\mathbb{P}\left[\left|\sum_{i=1}^{d}\langle \mathbf{S}_{:,i}, \frac{\mathbf{x}}{\|\mathbf{x}\|_2}\rangle^2 - d\right| \geq \delta \cdot d\right] \leq 2\exp\left(-c\min\left\{\frac{\delta^2 d}{K^2}, \frac{\delta d}{K}\right\}\right) = \exp\left(-\Omega\left(d\min\left\{\delta,\delta^2\right\}\right)\right),$$

where $K = \left\|\mathbf{X}_1^2 - \mathbb{E}\left[\mathbf{X}_1^2\right]\right\|_{\psi_1}$. A triangle inequality on ψ_1 norm gives the of Kas:

$$K \leq \left\|\mathbf{X}_1^2\right\|_{\psi_1} + \left\|\mathbb{E}\left[\mathbf{X}_1^2\right]\right\|_{\psi_1} \leq \left\|\mathbf{X}_1\right\|_{\psi_2}^2 + \frac{1}{\ln 2} \leq 2 + \frac{1}{\ln 2},$$

where the second inequality is a consequence of Proposition 2 and the last one is shown in Example 1. As $\left\|\mathbf{S}^T\mathbf{x}\right\|_2^2 = \sum_{i=1}^d \langle \mathbf{S}_{:,i}, \mathbf{x}\rangle^2$, the two claimed inequalities hold by rearranging the above inequality. □

Corollary 2. *Let* $\mathbf{x} \in \mathbb{R}^n$ *and* $\mathbf{S} \sim$ *Bernoulli*$(p)^{n\times d}$ *for a constant* $p \in (0,1)$*. Then,* $\forall \delta > 0$*,*

$$\mathbb{P}\left[\left\|\mathbf{S}^T\mathbf{x}\right\|_2^2 \geq d(1+\delta)\mu\right] \leq e^{-\Omega\left(d\min\left\{\delta,\delta^2\right\}\right)} \quad \text{and} \quad \mathbb{P}\left[\left\|\mathbf{S}^T\mathbf{x}\right\|_2^2 \leq d(1-\delta)\mu\right] \leq e^{-\Omega\left(d\min\left\{\delta,\delta^2\right\}\right)},$$

where $\mu = p(1-p)\left\|\mathbf{x}\right\|_2^2 + p^2\langle \mathbf{x}, \mathbf{1}_n\rangle^2$.

Proof For each $i = 1, \ldots, d$, we denote the i-th column of $\mathbf{S}$ as $\mathbf{S}_{:,i}$. Since $\langle \mathbf{S}_{:,1}, \mathbf{x}\rangle, \ldots, \langle \mathbf{S}_{:,d}, \mathbf{x}\rangle$ are i.i.d., Lemma 10 with $m = d$, $\mathbf{a} = \mathbf{1}_d$, $t = \delta d\mu$, and $X_i = \langle \mathbf{S}_{:,i}, \mathbf{x}\rangle$ for $i = 1, \ldots, d$, implies that there exists a universal constant $c > 0$ such that

$$\mathbb{P}\left[\left|\sum_{i=1}^{d}\langle \mathbf{S}_{:,i}, \mathbf{x}\rangle^2 - d\mathbb{E}\left[\langle \mathbf{S}_{:,1}, \mathbf{x}\rangle^2\right]\right| \geq \delta \cdot d\mu\right] \leq 2\exp\left(-c\min\left\{\frac{d\mu^2\delta^2}{K^2}, \frac{d\mu\delta}{K}\right\}\right),$$

where $K = \left\|\langle \mathbf{S}_{:,1}, \mathbf{x}\rangle^2 - \mathbb{E}\left[\langle \mathbf{S}_{:,1}, \mathbf{x}\rangle^2\right]\right\|_{\psi_1}$. The proof is done by showing (i). $\mathbb{E}\left[\langle \mathbf{S}_{:,1}, \mathbf{x}\rangle^2\right] = \mu$, and (ii). $K = \Theta(\mu)$.

(i). Show $\mathbb{E}\left[\langle \mathbf{S}_{:,1}, \mathbf{x}\rangle^2\right] = \mu$: By using linearity of expectation repeatedly, we obtain that

$$\begin{aligned}\mathbb{E}\left[\langle \mathbf{S}_{:,1}, \mathbf{x}\rangle^2\right] = \mathbb{E}\left[\left(\sum_{i=1}^{n}\mathbf{S}_{i,1}\mathbf{x}_i\right)^2\right] &= \sum_{i=1}^{n}\mathbb{E}\left[(\mathbf{S}_{i,1}\mathbf{x}_i)^2\right] + \sum_{i\neq j}\mathbb{E}\left[(\mathbf{S}_{i,1}\mathbf{x}_i)(\mathbf{S}_{j,1}\mathbf{x}_j)\right] \\ &= p\left\|\mathbf{x}\right\|_2^2 + p^2(\langle \mathbf{x}, \mathbf{1}_n\rangle^2 - \left\|\mathbf{x}\right\|_2^2) = p(1-p)\left\|\mathbf{x}\right\|_2^2 + p^2\langle \mathbf{x}, \mathbf{1}_n\rangle^2 = \mu.\end{aligned}$$

(ii). Show $K = \Theta(\mu)$: Let $\mathbf{Z} = \mathbf{S}_{:,1} - p\mathbf{1}_n$. As verified in (i), $\mathbb{E}\left[\langle \mathbf{S}_{:,1}, \mathbf{x}\rangle^2\right] = \mu = p(1-p)\left\|\mathbf{x}\right\|_2^2 + p^2\langle \mathbf{x}, \mathbf{1}_n\rangle^2$, we get

$$\begin{aligned}K = \left\|\langle \mathbf{S}_{:,1}, \mathbf{x}\rangle^2 - \mathbb{E}\left[\langle \mathbf{S}_{:,1}, \mathbf{x}\rangle^2\right]\right\|_{\psi_1} &= \left\|\langle \mathbf{Z}, \mathbf{x}\rangle^2 + 2p\langle \mathbf{Z}, \mathbf{x}\rangle\langle \mathbf{x}, \mathbf{1}_n\rangle - p(1-p)\left\|\mathbf{x}\right\|_2^2\right\|_{\psi_1} \\ &\leq \left\|\langle \mathbf{Z}, \mathbf{x}\rangle^2\right\|_{\psi_1} + 2p|\langle \mathbf{x}, \mathbf{1}_n\rangle|\left\|\langle \mathbf{Z}, \mathbf{x}\rangle\right\|_{\psi_1} + \frac{p(1-p)\left\|\mathbf{x}\right\|_2^2}{\ln 2}.\end{aligned}$$

Since $\mathbf{Z}$ has i.i.d. entries and $p = \Theta(1)$, we evaluate

$$\left\|\langle \mathbf{Z}, \mathbf{x}\rangle^2\right\|_{\psi_1} = \|\langle \mathbf{Z}, \mathbf{x}\rangle\|_{\psi_2}^2 = \sum_{i=1}^n \mathbf{x}_i^2 \|\mathbf{Z}_i\|_{\psi_2}^2 = \|x\|_2^2 \|\mathbf{Z}_1\|_{\psi_2}^2 \le \|x\|_2^2 \left(\|\mathbf{S}_{1,1}\|_{\psi_2} + \|p\|_{\psi_2}\right)^2,$$

$$\text{and } \|\langle \mathbf{Z}, \mathbf{x}\rangle\|_{\psi_1} = |\langle \mathbf{x}, \mathbf{1}_n\rangle| \|\mathbf{Z}_1\|_{\psi_1} \le |\langle \mathbf{x}, \mathbf{1}_n\rangle| \left(\|\mathbf{S}_{1,1}\|_{\psi_1} + \|p\|_{\psi_1}\right).$$

Because $\|\mathbf{S}_{1,1}\|_{\psi_2} = \frac{1}{\sqrt{\ln(1+p^{-1})}} = \Theta(1)$, $\|\mathbf{S}_{1,1}\|_{\psi_1} = \frac{1}{\ln(1+p^{-1})} = \Theta(1)$ (see Example 1 for ψ_1 and ψ_2 norm), $\|p\|_{\psi_2} = \frac{p}{\sqrt{\ln 2}} = \Theta(1)$, and $\|p\|_{\psi_1} = \frac{p}{\ln 2} = \Theta(1)$, combining all yields $K = \Theta(\mu)$. □

C.3 Techinical Lemmas

Lemma 12. *Let $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_n) \in [0,1]^n$ s.t $(\beta_2, \dots, \beta_n) \neq \mathbf{0}_{n-1}$, $\mathbf{U} = [\mathbf{u}_1, \cdots, \mathbf{u}_n] \in \mathbb{R}^{n\times n}$ be an orthonormal matrix, and $\mathbf{S} \sim \mathcal{N}(0,1)^{n\times d}$. Then, for any $\delta > 0$ and $\epsilon \in (0,1)$,*

$$\mathbb{P}\left[\sum_{i=2}^n \beta_i \langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2 \ge d(1+\epsilon)(1+\delta)\sum_{i=2}^n \beta_i\right] \le \exp\left(-\Omega\left(\max\left\{1, \sum_{i=2}^n \beta_i\right\}\min\left\{\delta, \delta^2\right\}\right)\right) + e^{-\Omega(d\epsilon^2)},$$

$$\textit{and} \quad \mathbb{P}\left[\sum_{i=2}^n \beta_i \langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2 \le d(1-\epsilon)(1-\delta)\sum_{i=2}^n \beta_i\right] \le \exp\left(-\Omega\left(\max\left\{1, \sum_{i=2}^n \beta_i\right\}\min\left\{\delta, \delta^2\right\}\right)\right) + e^{-\Omega(d\epsilon^2)}.$$

Proof In the following, we only focus on the upper-tail bound as the others will hold by symmetry.

For the simplicity of presentation, we introduce a set $\mathcal{V}_\epsilon = \{\mathbf{v} \in \mathbb{R}^d : 0 < \|\mathbf{v}\|_2^2 \le d(1+\epsilon)\}$ and the events

$$E = \mathbb{I}\left\{\sum_{i=2}^n \beta_i \langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2 \ge d(1+\epsilon)(1+\delta)\sum_{i=2}^n \beta_i\right\} \text{ and } G(\mathbf{v}) = \mathbb{I}\left\{\mathbf{S}^T\mathbf{u}_1 = \mathbf{v}\right\}, \forall \mathbf{v} \in \mathcal{V}_\epsilon.$$

Using Corollary 1 in C.2 with $\mathbf{x} = \mathbf{u}_1, \delta = \epsilon < 1$ and the fact that $\left\|\mathbf{S}^T\mathbf{u}_1\right\|_2^2 > 0$ a.e. yield that $\mathbb{P}\left[\neg\left(\cup_{\mathbf{v}\in\mathcal{V}_\epsilon} G(\mathbf{v})\right)\right] = \mathbb{P}\left[\left\|\mathbf{S}^T\mathbf{u}_1\right\|_2^2 > d(1+\epsilon)\right] \le e^{-\Omega(d\epsilon^2)}$, which explicitly says that $\cup_{\mathbf{v}\in\mathcal{V}_\epsilon} G(\mathbf{v})$ happens with high probability. As a consequence, we have

$$\mathbb{P}[E] \le \mathbb{P}\left[E \cap \neg\left(\cup_{\mathbf{v}\in\mathcal{V}_\epsilon} G(\mathbf{v})\right)\right] + \int_{\mathbf{v}\in\mathcal{V}_\epsilon} \mathbb{P}[E \cap G(\mathbf{v})]\, d\mathbb{P}[G(\mathbf{v})] \le e^{-\Omega(d\epsilon^2)} + \sup_{\mathbf{v}\in\mathcal{V}_\epsilon} \mathbb{P}[E \cap G(\mathbf{v})], \tag{37}$$

where the last inequality follows from $\mathbb{P}\left[\cup_{\mathbf{v}\in\mathcal{V}_\epsilon} G(\mathbf{v})\right] \le 1$ and the upper bound of $\mathbb{P}\left[\neg\left(\cup_{\mathbf{v}\in\mathcal{V}_\epsilon} G(\mathbf{v})\right)\right]$ proved above. By (37), it is sufficient to show that for any $\mathbf{v} \in \mathcal{V}_\epsilon$,

$$\mathbb{P}[E \cap G(\mathbf{v})] \le \exp\left(-\Omega\left(\max\left\{1, \sum_{i=2}^n \beta_i\right\}\min\left\{\delta, \delta^2\right\}\right)\right). \tag{38}$$

Show (38) **for any** $\mathbf{v} \in \mathcal{V}_\epsilon$

Since $\|\mathbf{v}\|_2^2 \le d(1+\epsilon)$ for each $\mathbf{v} \in \mathcal{V}_\epsilon$,

$$\begin{aligned}\mathbb{P}[E \cap G(\mathbf{v})] &= \mathbb{P}\left[\sum_{i=2}^n \beta_i \langle \mathbf{S}^T\mathbf{u}_i, \mathbf{v}\rangle^2 \ge \left(\sum_{i=2}^n \beta_i\right) d(1+\delta)(1+\epsilon)\right] \\ &\le \mathbb{P}\left[\sum_{i=2}^n \beta_i \langle \mathbf{S}^T\mathbf{u}_i, \mathbf{v}\rangle^2 \ge \left(\sum_{i=2}^n \beta_i\right)(1+\delta)\|\mathbf{v}\|_2^2\right].\end{aligned} \tag{39}$$

Because for each $i = 2, \dots, n$, $\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{v}\rangle = \sum_{r=1}^n \sum_{s=1}^d \mathbf{S}_{r,s}(\mathbf{u}_i)_r \mathbf{v}_s$ is a random variable from $\mathcal{N}(0, \|\mathbf{v}\|_2^2)$ (a linear combination of normal distributions is a normal distribution again), Example 1 in C.1 shows that

$\left\|\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{v}\rangle\right\|_{\psi_2} = 2\left\|\mathbf{v}\right\|_2$. Moreover, the assumption $\mathbf{U}$ is an orthonormal matrix implies that $\{\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{v}\rangle\}_{i=2}^n$ are independent (see Theorem 8.1, Chap 5(Gut, 2009)). By applying Lemma 10 in C.1 with $m = n-1$, $X_i = \langle \mathbf{S}^T\mathbf{u}_{i+1}, \mathbf{v}\rangle$, $\forall i = 1, \cdots, n-1$, $\mathbf{a} = (\beta_2, \ldots, \beta_n)$, and $t = \delta \cdot (\sum_{i=2}^n \beta_i)\left\|\mathbf{v}\right\|_2^2$, we give an upper bound of right-hand side of (39) as below (it is already shown that $\mathbb{E}[\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{v}\rangle^2] = \left\|\mathbf{v}\right\|_2^2$ before):

$$\begin{aligned}
\mathbb{P}\left[\sum_{i=2}^n \beta_i\left(\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{v}\rangle^2 - \|\mathbf{v}\|_2^2\right) \geq \delta\left(\sum_{i=2}^n \beta_i\right)\|\mathbf{v}\|_2^2\right] &\leq \mathbb{P}\left[\left|\sum_{i=2}^n \beta_i\left(\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{v}\rangle^2 - \|\mathbf{v}\|_2^2\right)\right| \geq \delta\left(\sum_{i=2}^n \beta_i\right)\|\mathbf{v}\|_2^2\right] \\
&\leq 2\exp\left(-c\min\left\{\frac{\delta^2\left(\sum_{i=2}^n \beta_i\right)^2\|\mathbf{v}\|_2^4}{\|\mathbf{v}\|_2^4\sum_{i=2}^n \beta_i^2}, \frac{\delta\cdot\left(\sum_{i=2}^n \beta_i\right)\|\mathbf{v}\|_2^2}{\|\mathbf{v}\|_2^2\max_{i\neq 1}\beta_i}\right\}\right) \\
&= 2\exp\left(-c\min\left\{\frac{(\sum_{i=2}^n \beta_i)^2\delta^2}{\sum_{i=2}^n \beta_i^2}, \frac{\sum_{i=2}^n \beta_i\delta}{\max_{i\neq 1}\beta_i}\right\}\right). \qquad (40)
\end{aligned}$$

Combining (39)(40), it remains to show that

$$\text{(i)}\frac{(\sum_{i=2}^n \beta_i)^2}{\sum_{i=2}^n \beta_i^2} \geq \max\left\{1, \sum_{i=2}^n \beta_i\right\}, \text{ and (ii).}\frac{\sum_{i=2}^n \beta_i}{\max_{i\neq 1}\beta_i} \geq \max\left\{1, \sum_{i=2}^n \beta_i\right\}.$$

For (i). As $(\sum_{i=2}^n \beta_i)^2 = \sum_{i=2}^n \beta_i^2 + \sum_{i\neq j}\beta_i\beta_j \geq \sum_{i=2}^n \beta_i^2$, and $\sum_{i=2}^n \beta_i^2 \leq \sum_{i=2}^n \beta_i$, (i) holds by using these two inequalities in numerator and denominator respectively.
For (ii). As $\sum_{i=2}^n \beta_i \geq \max_{i\neq 1}\beta_i$, and $\max_{i\neq 1}\beta_i \leq 1$, (ii) follows by using these two inequalities in numerator and denominator respectively. □

Lemma 13. *Let $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_n) \in [0,1]^n$ s.t. $(\beta_2, \ldots, \beta_n) \neq \mathbf{0}_{n-1}$, $\mathbf{U} = [\mathbf{u}_1, \cdots, \mathbf{u}_n] \in \mathbb{R}^{n\times n}$ be an orthonormal matrix with $\langle \mathbf{u}_1, \mathbf{1}_n\rangle^2 = \Omega(1)$, and $\mathbf{S} \sim$ Bernoulli$(p)^{n\times d}$ with some constant $p \in (0,1)$. Then, for any $\delta > 0$ and $\epsilon \in (0,1)$, we have probability at least $1 - e^{-\Omega\left(\max\left\{1, \sum_{i=2}^n \beta_i\xi_i\right\}\frac{(1-\epsilon)^2}{(1+\epsilon)^2}\min\left\{\delta, \delta^2\right\}\right)} - e^{-\Omega(\min(d,\xi_1)\epsilon^2)}$ that*

$$(1-\delta)(1-\epsilon)\sum_{i=2}^n \beta_i\xi_i\xi_1 \leq \sum_{i=2}^n \beta_i\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2 \leq d(1+\delta)(1+\epsilon)\sum_{i=2}^n \beta_i\xi_i\mu_1,$$

where

$$\mu_i = p(1-p+p\langle \mathbf{u}_i, \mathbf{1}_n\rangle^2), \quad \textit{and} \quad \xi_i = p(1-p+pd\langle \mathbf{u}_i, \mathbf{1}_n\rangle^2), \quad \forall i \in [n].$$

Proof Similar to the proof C.3 of Lemma 12, we introduce the set

$$\mathcal{V}_\epsilon = \{\mathbf{v} \in \mathbb{R}^d : d\mu_1(1-\epsilon) \leq \|\mathbf{v}\|_2^2 \leq d\mu_1(1+\epsilon) \text{ and } d\xi_1(1-\epsilon) \leq \langle \mathbf{v}, \mathbf{1}_d\rangle_2^2 \leq d\xi_1(1+\epsilon)\}$$

and the events

$$\begin{aligned}
E = \mathbb{I}\Bigg(&\left\{\sum_{i=2}^n \beta_i\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2 \leq (1-\delta)(1-\epsilon)\sum_{i=2}^n \beta_i\xi_i\xi_1\right\} \cup \\
&\left\{\sum_{i=2}^n \beta_i\langle \mathbf{S}^T\mathbf{u}_i, \mathbf{S}^T\mathbf{u}_1\rangle^2 \geq d(1+\delta)(1+\epsilon)\sum_{i=2}^n \beta_i\xi_i\mu_1\right\}\Bigg)
\end{aligned}$$

and $G(\mathbf{v}) = \mathbb{I}\left\{\mathbf{S}^T\mathbf{u}_1 = \mathbf{v}\right\}, \forall \mathbf{v} \in \mathcal{V}_\epsilon$. The sum rule of probability implies that

$$\mathbb{P}[E] = \mathbb{P}[E \cap \neg(\cup_{\mathbf{v}\in\mathcal{V}_\epsilon} G(\mathbf{v}))] + \int_{\mathbf{v}\in\mathcal{V}_\epsilon}\mathbb{P}[E\cap G(\mathbf{v})]\, d\mathbb{P}[G(\mathbf{v})] \leq \mathbb{P}[\neg(\cup_{\mathbf{v}\in\mathcal{V}_\epsilon} G(\mathbf{v}))] + \sup_{\mathbf{v}\in\mathcal{V}_\epsilon}\mathbb{P}[E\cap G(\mathbf{v})]. \qquad (41)$$

To bound the first term in (41), we claim that

$$\begin{aligned}
\text{(i). } &\mathbb{P}\left[\left|\left\|\mathbf{S}^T\mathbf{u}_1\right\|_2^2 - d\mu_1\right| \geq \epsilon\cdot d\mu_1\right] \leq e^{-\Omega(d\epsilon^2)}, \\
\text{(ii). } &\mathbb{P}\left[\left|\langle \mathbf{S}^T\mathbf{u}_1, \mathbf{1}_d\rangle^2 - d\xi_1\right| \geq \epsilon\cdot d\xi_1\right] \leq e^{-\Omega(\xi_1\epsilon^2)},
\end{aligned}$$

and then an application of a union bound of (i)(ii) yields $\mathbb{P}\left[\neg\left(\cup_{\mathbf{v}\in\mathcal{V}_\epsilon} G(\mathbf{v})\right)\right] \le e^{-\Omega(d\epsilon^2)} + e^{-\Omega(\xi_1\epsilon^2)}$. As for the second term in (41), one consequence of Lemma 14 at the end of this section is that

$$\sup_{\mathbf{v}\in\mathcal{V}_\epsilon} \mathbb{P}\left[E\cap G(\mathbf{v})\right] \le e^{-\Omega\left(\max\left\{1,\sum_{i=2}^n \beta_i\xi_i\right\}\min\left\{1,\frac{\langle\mathbf{v},\mathbf{1}_d\rangle^4}{d^2\|\mathbf{v}\|_2^4},\frac{\langle\mathbf{v},\mathbf{1}_d\rangle^2}{d\|\mathbf{v}\|_2^2}\right\}\min\left\{\delta,\delta^2\right\}\right)}$$

in which $\min\left\{1,\frac{\langle\mathbf{v},\mathbf{1}_d\rangle^4}{d^2\|\mathbf{v}\|_2^4},\frac{\langle\mathbf{v},\mathbf{1}_d\rangle^2}{d\|\mathbf{v}\|_2^2}\right\} = \min\left\{1,\frac{(1-\epsilon)^2}{(1+\epsilon)^2}\frac{\xi_1^2}{d^2\mu_1^2}\right\} = \Omega\left(\frac{(1-\epsilon)^2}{(1+\epsilon)^2}\right)$. Hence, combining all by union bound gives the desired.

It remains to show (i) and (ii). For convenience, let $\mathbf{Z} = \mathbf{S} - p\mathbf{1}_n\mathbf{1}_d^T$, a zero-mean matrix.

(i). This is a direct result of the first inequality in Corollary 2 in C.2 with $\mathbf{x} = \mathbf{u}_1$ and $\delta = \epsilon$.

(ii). To show $\mathbb{P}\left[\left|\langle\mathbf{S}^T\mathbf{u}_1,\mathbf{1}_d\rangle^2 - d\xi_1\right| \ge \epsilon\cdot d\xi_1\right] \le e^{-\Omega(\xi_1\epsilon^2)}$, where $\xi = p(1-p+pd\langle\mathbf{u}_1,\mathbf{1}_n\rangle^2)$. As the lower tail is proved in a similar to the upper tail, in what follows, we will pay attention on the upper tail only. Firstly, it is easy to verified that

$$\langle\mathbf{S}^T\mathbf{u}_1,\mathbf{1}_d\rangle^2 = \langle\mathbf{Z}^T\mathbf{u}_1,\mathbf{1}_d\rangle^2 + 2pd\langle\mathbf{Z}^T\mathbf{u}_1,\mathbf{1}_d\rangle\langle\mathbf{u}_1,\mathbf{1}_n\rangle + p^2d^2\langle\mathbf{u}_1,\mathbf{1}_n\rangle^2. \tag{42}$$

To bound the first (resp. the second) term in (42), we will use Proposition 3 in C.1 for sub-exponential (resp. sub-gaussian) r.v., which is quantified the sub-gaussian norm, denoted by $K = \left\|\langle\mathbf{Z}^T\mathbf{u}_1,\mathbf{1}_d\rangle\right\|_{\psi_2}$ (recall that the sub-exponential norm can be obtained by sub-gaussian norm, and vice versa, see Propsition 2 in C.1). By Proposition 1 and Example 1 in C.1, we have

$$K^2 = \mathcal{O}\left(d\sum_{i\in[n]}(\mathbf{u}_1)_i^2\left\|\mathbf{Z}_{1,1}\right\|_{\psi_2}^2\right) = \mathcal{O}\left(d\left\|\mathbf{Z}_{1,1}\right\|_{\psi_2}^2\right) = \mathcal{O}\left(d\left(\left\|\mathbf{S}_{i,j}\right\|_{\psi_2} + \|p\|_{\psi_2}\right)^2\right) = \mathcal{O}(d).$$

Additionally, one can evaluate $\langle\mathbf{Z}^T\mathbf{u}_1,\mathbf{1}_d\rangle^2 = dp(1-p)$ by repeatedly use the linearity of expectation and the fact that the entries of $\mathbf{S}$ are i.i.d. drawn from Bernoulli(p). Hence, invoking the concentration inequality for sub-exponential (resp. sub-gaussian) in Proposition 3 in C.1 with $t = \frac{d\xi_1\epsilon}{3}$ (resp. $t = \frac{\epsilon\sqrt{d\xi_1}}{3}$) on $\langle\mathbf{Z}^T\mathbf{u}_1,\mathbf{1}_d\rangle^2$ (resp. $\langle\mathbf{Z}^T\mathbf{u}_1,\mathbf{1}_d\rangle$) yields that

$$\mathbb{P}\left[\left|\langle\mathbf{Z}^T\mathbf{u}_1,\mathbf{1}_d\rangle^2 - dp(1-p)\right| \ge \frac{d\xi_1\epsilon}{3}\right] \le e^{-\Omega(\frac{d\xi_1\epsilon}{K^2})} \le e^{-\Omega(\xi_1\epsilon^2)}, \tag{43}$$

$$\mathbb{P}\left[\left|\langle\mathbf{Z}^T\mathbf{u}_1,\mathbf{1}_d\rangle\right| \ge \frac{\epsilon\sqrt{d\xi_1}}{3}\right] \le e^{-\Omega(\frac{d\xi_1\epsilon^2}{K^2})} = e^{-\Omega(\xi_1\epsilon^2)}. \tag{44}$$

Plugging these (43) and (44) into (42), a union bound gives us that

$$\begin{aligned}\langle\mathbf{Z}^T\mathbf{u}_1,\mathbf{1}_d\rangle^2 + 2pd\langle\mathbf{Z}^T\mathbf{u}_1,\mathbf{1}_d\rangle\langle\mathbf{u}_1,\mathbf{1}_n\rangle + p^2d^2\langle\mathbf{u}_1,\mathbf{1}_n\rangle^2 &\le dp(1-p) + p^2d^2\langle\mathbf{u}_1,\mathbf{1}_n\rangle^2 + \frac{d\xi_1\epsilon}{3} + 2pd\langle\mathbf{u}_1,\mathbf{1}_n\rangle\frac{\epsilon\sqrt{d\xi_1}}{3}\\ &\le d\xi_1 + d\xi_1\epsilon,\end{aligned}$$

where the second inequality is yielded by $\xi_1 = p(1-p+pd\langle\mathbf{u}_1,\mathbf{1}_n\rangle^2)$ and $pd\langle\mathbf{u}_1,\mathbf{1}_n\rangle \le \sqrt{d\xi_1}$. Then we conclude this lemma with (ii) as desired. □

Lemma 14. *Let* $\mathbf{v}\in\mathbb{R}^d\setminus\{\mathbf{0}_d\}$, $(\beta_2,\ldots,\beta_n)\in[0,1]^{n-1}\setminus\{\mathbf{0}_{n-1}\}$, $[\mathbf{u}_1,\ldots,\mathbf{u}_n]\in\mathbb{R}^{n\times n}$ *be an orthonormal matrix,* $\mathbf{S}\sim$ *Bernoulli*$(p)^{n\times d}$ *with some constant* $p\in(0,1)$*, and* $\xi_i = p(1-p+pd\langle\mathbf{u}_i,\mathbf{1}_n\rangle^2), \forall i\in[n]$*. Then,*

$$\mathbb{P}\left[(1-\delta)\eta_1 \le \sum_{i=2}^n \beta_i\langle\mathbf{S}^T\mathbf{u}_i,\mathbf{v}\rangle^2 \le (1+\delta)\eta_2\right] \ge 1 - e^{-\Omega\left(\max\left\{1,\sum_{i=2}^n\beta_i\xi_i\right\}\min\left\{1,\frac{\langle\mathbf{v},\mathbf{1}_n\rangle^4}{d^2\|\mathbf{v}\|_2^4},\frac{\langle\mathbf{v},\mathbf{1}_n\rangle^2}{d\|\mathbf{v}\|_2^2}\right\}\min\left\{\delta,\delta^2\right\}\right)},$$

where $\eta_1 = \sum_{i=2}^n \beta_i\xi_i\frac{\langle\mathbf{v},\mathbf{1}_d\rangle^2}{d}$ *and* $\eta_2 = \sum_{i=2}^n\beta_i\xi_i\|\mathbf{v}\|_2^2$.

Proof An elementary calculation of evaluating the expectation of $\sum_{i=2}^n \beta_i \langle \mathbf{S}^T\mathbf{u}_i, \mathbf{v}\rangle^2$ leads to

$$\mathbb{E}\left[\sum_{i=2}^n \beta_i \langle \mathbf{S}^T\mathbf{u}_i, \mathbf{v}\rangle^2\right] = p(1-p)\sum_{i=2}^n \beta_i \|\mathbf{v}\|_2^2 + p^2 \sum_{i=2}^n \beta_i \langle \mathbf{u}_i, \mathbf{1}_n\rangle^2 \langle \mathbf{v}, \mathbf{1}_d\rangle^2.$$

After applying Cauchy inequality, $\langle \mathbf{v}, \mathbf{1}_d\rangle^2 \le d\|\mathbf{v}\|_2^2$, twice, we get that $\eta_1 \le \mathbb{E}\left[\sum_{i=2}^n \beta_i \langle \mathbf{S}^T\mathbf{u}_i, \mathbf{v}\rangle^2\right] \le \eta_2$. This observation inspires us to give high probability lower bound in term of η_1 and upper bound in term of η_2 respectively.

Define $\mathbf{Z} = \mathbf{S} - p\mathbf{1}_n\mathbf{1}_d^T$, $\mathbf{M} = \sum_{i=2}^n \beta_i \mathbf{u}_i\mathbf{u}_i^T$ and $\mathbf{B} = \mathbf{M} \otimes \mathbf{v}\mathbf{v}^T$ where $\otimes$ is the Kronecker product. With these definition, we can express the weighted sum as:

$$\begin{aligned} \sum_{i=2}^n \beta_i \langle \mathbf{S}^T\mathbf{u}_i, \mathbf{v}\rangle^2 &= \sum_{(i_1,j_1),(i_2,j_2)\in[n]\times[d]} \mathbf{B}_{(i_1,j_1),(i_2,j_2)} \mathbf{S}_{i_1,j_1}\mathbf{S}_{i_2,j_2} = \text{(I)} + \text{(II)}, \\ \text{where} \quad \text{(I)} &= \sum_{(i_1,j_1),(i_2,j_2)\in[n]\times[d]} \mathbf{B}_{(i_1,j_1),(i_2,j_2)} \mathbf{Z}_{i_1,j_1}\mathbf{Z}_{i_2,j_2} \\ \text{and} \quad \text{(II)} &= \sum_{(i_1,j_1),(i_2,j_2)\in[n]\times[d]} \mathbf{B}_{(i_1,j_1),(i_2,j_2)} (\mathbf{Z}_{i_1,j_1} + \mathbf{Z}_{i_2,j_2} + p)p. \end{aligned} \tag{45}$$

Such decomposition allows us to bound (I) by Lemma 11 and bound (II) by Lemma 9 in C.1, which require us to evaluate the necessary quantities.

- $\|\mathbf{B}\|_F^2 = \sum_{i_1,i_2\in[n],j_1,j_2\in[d]} \left(\mathbf{M}_{i_1,i_2}\mathbf{v}_{j_1}\mathbf{v}_{j_2}\right)^2 = \|\mathbf{v}\|_2^4 \|\mathbf{M}\|_F = \|\mathbf{v}\|_2^4 \sum_{i=2}^n \beta_i^2$,
  where the last equation is due to $\mathbf{M} = \sum_{i=2}^n \beta_i \mathbf{u}_i\mathbf{u}_i^T$ is an eigenvalue decomposition of $\mathbf{M}$.
- $\|\mathbf{B}\|_2 = \|\mathbf{M}\|_2 \left\|\mathbf{v}\mathbf{v}^T\right\|_2 = \max_{i\neq 1} \beta_i \|\mathbf{v}\|_2^2$,
  where the first equation is a property of Kronecker product (see e.g. Theorem 4.2.15 in (Horn et al., 1994)).

To bound (I), invoking Lemma 11 with $m = nd$, $\mathbf{M} = \mathbf{B}$, $\mathbf{X}_{(i-1)d+j} = \mathbf{Z}_{i,j}$, $\forall i \in [n], j \in [d]$ and $t = \delta\eta_1/2$ (resp. $t = \delta\eta_2/2$) for the lower- (resp. upper-) tail bounds yields that

$$\mathbb{P}\left[\neg\left\{-\frac{\delta\eta_1}{2} < \text{(I)} - \mathbb{E}\left[\text{(II)}\right] < \frac{\delta\eta_2}{2}\right\}\right] \le \exp\left(-\Omega\left(\min\left\{\frac{\eta_2^2\delta^2}{\|\mathbf{B}\|_F^2}, \frac{\eta_2\delta}{\|\mathbf{B}\|_2}\right\}\right)\right) + \exp\left(-\Omega\left(\min\left\{\frac{\eta_1^2\delta^2}{\|\mathbf{B}\|_F^2}, \frac{\eta_1\delta}{\|\mathbf{B}\|_2}\right\}\right)\right). \tag{46}$$

To bound (II), applying Lemma 9 with $t = \delta\eta_1/4$ (resp. $t = \delta\eta_2/4$) for the lower- (resp. upper-) tail bounds yields

$$\mathbb{P}\left[\neg\left\{-\frac{\delta\eta_1}{2} < \text{(II)} - \mathbb{E}\left[\text{(II)}\right] < \frac{\delta\eta_2}{2}\right\}\right] \le \exp\left(-\Omega\left(\frac{\eta_2^2\delta^2}{\|\mathbf{B}\|_F^2}\right)\right) + \exp\left(-\Omega\left(\frac{\eta_1^2\delta^2}{\|\mathbf{B}\|_F^2}\right)\right). \tag{47}$$

In what follows, we will show

(i). $\min\left\{\frac{\eta_2^2}{\|\mathbf{B}\|_F^2}, \frac{\eta_2}{\|\mathbf{B}\|_2}\right\} = \Omega\left(\max\left\{1, \sum_{i=2}^n \beta_i\xi_i\right\}\right)$, and

(ii). $\min\left\{\frac{\eta_1^2}{\|\mathbf{B}\|_F^2}, \frac{\eta_1}{\|\mathbf{B}\|_2}\right\} = \Omega\left(\max\left\{1, \sum_{i=2}^n \beta_i\xi_i\right\} \min\left\{\frac{\langle \mathbf{v}, \mathbf{1}_d\rangle^4}{d^2\|\mathbf{v}\|_2^4}, \frac{\langle \mathbf{v}, \mathbf{1}_d\rangle^2}{d\|\mathbf{v}\|_2^2}\right\}\right)$.

Then this proof is done by using a union bound of (46) and (47) into (45).

<u>(i).</u> From the definition of η_2 and our above computations, we get $\frac{\eta_2^2}{\|\mathbf{B}\|_F^2} = \Omega\left(\frac{(\sum_{i=2}^n \beta_i\xi_i)^2}{\sum_{i=2}^n \beta_i^2}\right)$ and $\frac{\eta_2}{\|\mathbf{B}\|_2} = \Omega\left(\frac{\sum_{i=2}^n \beta_i\xi_i}{\max_{i\neq 1}\beta_i}\right)$. It is done by the following claims:

$$\text{(a).}\frac{(\sum_{i=2}^n \beta_i\xi_i)^2}{\sum_{i=2}^n \beta_i^2} = \Omega\left(\max\left\{1, \sum_{i=2}^n \beta_i\xi_i\right\}\right), \text{ and (b).}\frac{\sum_{i=2}^n \beta_i\xi_i}{\max_{i\neq 1}\beta_i} = \Omega\left(\max\left\{1, \sum_{i=2}^n \beta_i\xi_i\right\}\right).$$

<u>(a).</u> stems from $(\sum_{i=2}^n \beta_i\xi_i)^2 \ge p^2(1-p)^2\sum_{i=2}^n \beta_i^2$, and $p(1-p)\sum_{i=2}^n \beta_i^2 \le p(1-p)\sum_{i=2}^n \beta_i \le \sum_{i=2}^n \beta_i\xi_i$.
<u>(b).</u> holds since $\sum_{i=2}^n \beta_i\xi \ge p(1-p)\max_{i\neq 1}\beta_i$, and $\max_{i\neq 1}\beta_i \le 1$.

(ii). From the definition of η_2 and our above computations, we get

$$\min\left\{\frac{\eta_1^2}{\|\mathbf{B}\|_F^2}, \frac{\eta_1}{\|\mathbf{B}\|_2}\right\} = \Omega\left(\min\left\{\frac{(\sum_{i=2}^n \beta_i\xi_i)^2}{\sum_{i=2}^n \beta_i^2}, \frac{\sum_{i=2}^n \beta_i\xi_i}{\max_{i\neq 1}\beta_i}\right\}\min\left\{\frac{\langle \mathbf{v}, \mathbf{1}_d\rangle^4}{d^2\|\mathbf{v}\|_2^4}, \frac{\langle \mathbf{v}, \mathbf{1}_d\rangle^2}{d\|\mathbf{v}\|_2^2}\right\}\right).$$

We then deduce (ii). by (a). and (b). and conclude this proof. □

D Conflicting group detection: approximation ratio

Algorithm 3: RandomEigenSign $(\mathbf{v})$ by Bonchi et al. (2019)

for $i = 1 \to n$ **do**
  $\mathbf{r}_i = \text{sign}(\mathbf{v}_i) \cdot \text{Bernoulli}(|\mathbf{v}_i|)$;
end
return $\mathbf{r}$;

Theorem 7. *For any* $\hat{\mathbf{u}} \in \mathbb{S}^{n-1}$, RandomEigenSign$(\hat{\mathbf{u}})$ *is an* $\mathcal{O}(n^{1/2}/R(\hat{\mathbf{u}}))$*-approx algorithm to* 2*-conflicting group detection.*

Proof The proof strategy is similar to the analysis in (Bonchi et al., 2019).

Let $\mathbf{r} = \text{RandomEigenSign}(\hat{\mathbf{u}})$ and $\mathbf{s} = \text{sign}(\hat{\mathbf{u}})$ where $\mathbf{s}_i = 1$ if $\hat{\mathbf{u}}_i > 0$ otherwise $\mathbf{s}_i = 0$, $\forall i \in [n]$. We have

$$\begin{aligned}
\mathbb{E}\left[\frac{\mathbf{r}^T\mathbf{A}\mathbf{r}}{\mathbf{r}^T\mathbf{r}}\right] &= \sum_k \mathbb{E}\left[\frac{\mathbf{r}^T\mathbf{A}\mathbf{r}}{\mathbf{r}^T\mathbf{r}}\middle|\mathbf{r}^T\mathbf{r} = k\right]\mathbb{P}\left[\mathbf{r}^T\mathbf{r} = k\right] = \sum_k \frac{1}{k}\sum_{i,j\in[n]} \mathbf{A}_{i,j}\mathbf{s}_i\mathbf{s}_j\mathbb{P}\left[\mathbf{r}_i\mathbf{r}_j = \mathbf{s}_i\mathbf{s}_j\middle|\mathbf{r}^T\mathbf{r} = k\right]\mathbb{P}\left[\mathbf{r}^T\mathbf{r} = k\right] \\
&\overset{(a)}{=} \sum_k \frac{1}{k}\sum_{i,j\in[n]} \mathbf{A}_{i,j}\mathbf{s}_i\mathbf{s}_j\mathbb{P}\left[\mathbf{r}^T\mathbf{r} = k\middle|\mathbf{r}_i\mathbf{r}_j = \mathbf{s}_i\mathbf{s}_j\right]\mathbb{P}\left[\mathbf{r}_i\mathbf{r}_j = \mathbf{s}_i\mathbf{s}_j\right] = \sum_{i,j\in[n]} \mathbf{A}_{i,j}\hat{\mathbf{u}}_i\hat{\mathbf{u}}_j\mathbb{E}\left[\frac{1}{r^Tr}\middle|\mathbf{r}_i\mathbf{r}_j = \mathbf{s}_i\mathbf{s}_j\right] \\
&\overset{(b)}{\geq} \sum_{i,j\in[n]} \mathbf{A}_{i,j}\hat{\mathbf{u}}_i\hat{\mathbf{u}}_j\frac{1}{\mathbb{E}\left[r^Tr\middle|\mathbf{r}_i\mathbf{r}_j = \mathbf{s}_i\mathbf{s}_j\right]},
\end{aligned}$$

where (a) results from applying Bayes' rule, and (b) uses conditional Jensen's inequality. By

$$\mathbb{E}\left[\mathbf{r}^T\mathbf{r}\middle|\mathbf{r}_i\mathbf{r}_j = \mathbf{s}_i\mathbf{s}_j\right] = 2 + \sum_{\ell\in[n]\setminus\{i,j\}} \mathbb{P}\left[\mathbf{r}_\ell = \mathbf{s}_\ell\right] \leq 2 + \sqrt{n-2},$$

(b) and $\hat{\mathbf{u}}^T\mathbf{A}\hat{\mathbf{u}} = R(\hat{\mathbf{u}})$, we get that $\mathbb{E}\left[\frac{\mathbf{r}^T\mathbf{A}\mathbf{r}}{\mathbf{r}^T\mathbf{r}}\right] \geq \frac{\hat{\mathbf{u}}^T\mathbf{A}\hat{\mathbf{u}}}{2+\sqrt{n-2}} = \frac{R(\hat{\mathbf{u}})\lambda_1}{2+\sqrt{n-2}}$. □

Appendix C

Closing the computational-statistical gap in combinatorial BAI

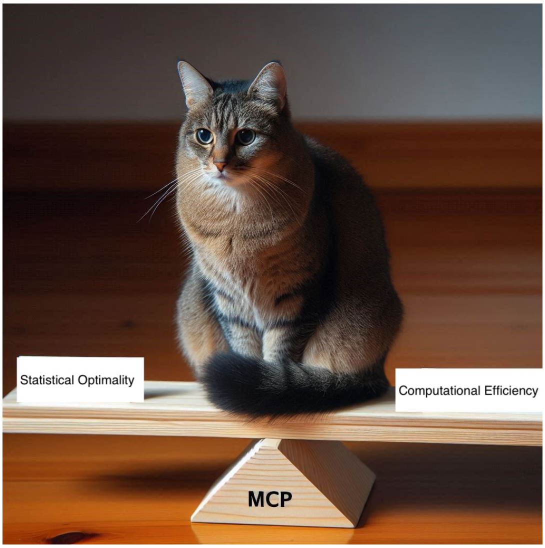


Figure: A cat closing the gap between statistical optimality and computational efficiency using the `MCP` algorithm.

In the combinatorial best arm identification with fixed confidence and semi-bandit feedbacks, there exists a statistically optimal but not computationally efficient algorithm. The computational bottleneck is the computation for the most confusing parameter (`MCP`). We design an approximate `MCP` algorithm and use it to design an algorithm that is both statistically optimal and computationally efficient.

Closing the Computational-Statistical Gap in Best Arm Identification for Combinatorial Semi-bandits

Ruo-Chun Tzeng
EECS
KTH, Stockholm, Sweden
rctzeng@kth.se

Po-An Wang
EECS
KTH, Stockholm, Sweden
wang9@kth.se

Alexandre Proutiere
EECS and Digital Futures
KTH, Stockholm, Sweden
alepro@kth.se

Chi-Jen Lu
Institute of Information Science
Academia Sinica, Taiwan
cjlu@iis.sinica.edu.tw

Abstract

We study the best arm identification problem in combinatorial semi-bandits in the fixed confidence setting. We present Perturbed Frank-Wolfe Sampling (P-FWS), an algorithm that (i) runs in polynomial time, (ii) achieves the instance-specific minimal sample complexity in the high confidence regime, and (iii) enjoys polynomial sample complexity guarantees in the moderate confidence regime. To the best of our knowledge, even for the vanilla bandit problems, no algorithm was able to achieve (ii) and (iii) simultaneously. With P-FWS, we close the computational-statistical gap in best arm identification in combinatorial semi-bandits. The design of P-FWS starts from the optimization problem that defines the information-theoretical and instance-specific sample complexity lower bound. P-FWS solves this problem in an online manner using, in each round, a single iteration of the Frank-Wolfe algorithm. Structural properties of the problem are leveraged to make the P-FWS successive updates computationally efficient. In turn, P-FWS only relies on a simple linear maximization oracle.

1 Introduction

An efficient method to design statistically optimal algorithms solving active learning tasks (e.g., regret minimization or pure exploration in bandits and reinforcement learning) consists in the following two-step procedure. The first step amounts to deriving, through change-of-measure arguments, tight information-theoretical fundamental limits satisfied by a wide class of learning algorithms. These limits are often expressed as the solution of an optimization problem, referred in this paper to as the *lower-bound problem*. Interestingly, this solution specifies the instance-specific optimal exploration process: it characterizes the limiting behavior of the adaptive sampling rule any statistically optimal algorithm should implement. In the second step, the learning algorithm is designed so that its exploration process approaches the solution of the lower-bound problem. This design yields statistically optimal algorithms, but typically requires to repeatedly solve the lower-bound problem. This method has worked remarkably well for simple learning tasks such as regret minimization or best-arm identification with fixed confidence in classical stochastic bandits [Lai87, GC11, GK16], but also in bandits whose arm-to-average reward function satisfies simple structural properties (e.g., Lipschitz, unimodal) [MCP14, WTP21].

The method also provides a natural way of studying the computational-statistical gap [KLLM22] for active learning tasks. Indeed, if solving the lower-bound problem in polynomial time is possible, one

37th Conference on Neural Information Processing Systems (NeurIPS 2023).

may hope to devise learning algorithms that are both statistically optimal and computationally efficient. As of now, however, the computational complexity of the lower-bound problem remains largely unexplored, except for simple learning tasks. For example, in the case of best policy identification in tabular Markov Decision Processes, the lower-bound problem is non-convex [AMP21] and its complexity and approximability are unclear.

In this paper, we leverage the aforementioned two-step procedure to assess the computational-statistical gap for the best arm identification in combinatorial semi-bandits in the fixed confidence setting. We establish that, essentially, this gap does not exist (a result that was conjectured in [JMKK21]). Specifically, we present an algorithm that enjoys the following three properties: (i) it runs in polynomial time, (ii) its sample complexity matches the fundamental limits asymptotically in the high confidence regime, and (iii) its sample complexity is at most polynomial in the moderate confidence regime. Next, after formally introducing combinatorial semi-bandits, we describe our contributions and techniques in detail.

Best arm identification in combinatorial semi-bandits. In combinatorial semi-bandits [CBL12, CTMSP+15], the learner sequentially selects an action from a combinatorial set $\mathcal{X} \subset \{0,1\}^K$. When in round t, the action $\boldsymbol{x}(t) = (x_1(t), \dots, x_K(t)) \in \mathcal{X}$ is chosen, the environment samples a K-dimensional vector $\boldsymbol{y}(t)$ whose distribution is assumed to be Gaussian $\mathcal{N}(\boldsymbol{\mu}, \mathbf{I})$. The learner then receives the detailed reward vector $\boldsymbol{x}(t) \odot \boldsymbol{y}(t)$ where $\odot$ denotes the element-wise product (in other words, the learner observes the individual reward $y_k(t)$ of the arm k if and only if this arm is selected in round t, i.e., $x_k(t) = 1$). The parameter $\boldsymbol{\mu}$ characterizing the average rewards of the various arms is initially unknown. The goal of a learner is to identify the best action $\boldsymbol{i}^\star(\boldsymbol{\mu}) = \operatorname{argmax}_{\boldsymbol{x}\in\mathcal{X}} \langle \boldsymbol{x}, \boldsymbol{\mu} \rangle$ with a given level of confidence $1-\delta$, for some $\delta > 0$ while minimizing the expected number of rounds needed. We assume that the best action is unique and denote by $\Lambda = \{\boldsymbol{\mu} \in \mathbb{R}^K : |\boldsymbol{i}^\star(\boldsymbol{\mu})| = 1\}$ the set of parameters satisfying this assumption. The learner strategy is defined by three components: (i) a sampling rule dictating the sequence of the selected actions; (ii) a stopping time τ defining the last round where the learner interacts with the environment; (iii) a decision rule specifying the action $\hat{\boldsymbol{\imath}} \in \mathcal{X}$ believed to be optimal based on the data gathered until τ.

The sample complexity lower-bound problem. Consider the set of δ-PAC algorithms such that for any $\boldsymbol{\mu} \in \Lambda$, the best action is identified correctly with probability at least $1-\delta$. We wish to find a δ-PAC algorithm with minimal expected sample complexity $\mathbb{E}_{\boldsymbol{\mu}}[\tau]$. To this aim, using classical change-of-measure arguments [GK16], we may derive a lower bound of the expected sample complexity satisfied by any δ-PAC algorithm. This lower bound is given by[1] $\mathbb{E}_{\boldsymbol{\mu}}[\tau] \geq T^\star(\boldsymbol{\mu})\mathsf{kl}(\delta, 1-\delta)$. The characteristic time $T^\star(\boldsymbol{\mu})$ is defined as the value of the following problem

$$T^\star(\boldsymbol{\mu})^{-1} = \sup_{\boldsymbol{\omega}\in\Sigma} \inf_{\boldsymbol{\lambda}\in\mathsf{Alt}(\boldsymbol{\mu})} \left\langle \boldsymbol{\omega}, \frac{(\boldsymbol{\mu}-\boldsymbol{\lambda})^2}{2} \right\rangle, \tag{1}$$

where[2] $\Sigma = \{\sum_{\boldsymbol{x}\in\mathcal{X}} w_{\boldsymbol{x}} \boldsymbol{x} : \boldsymbol{w} \in \Sigma_{|\mathcal{X}|}\}$, $\mathsf{kl}(a,b)$ is the KL-divergence between two Bernoulli distributions with respective means a and b, and $\mathsf{Alt}(\boldsymbol{\mu}) = \{\boldsymbol{\lambda} \in \Lambda : \boldsymbol{i}^\star(\boldsymbol{\lambda}) \neq \boldsymbol{i}^\star(\boldsymbol{\mu})\}$ is the set of *confusing* parameters. As it turns out (see Lemma 1), $T^\star(\boldsymbol{\mu})$ is at most quadratic in K, and hence the sample complexity lower bound is polynomial. (1) is a concave program over Σ [WTP21], and a point $\boldsymbol{\omega}^\star$ in its solution set corresponds to an optimal allocation of action draws: an algorithm sampling actions according to $\boldsymbol{\omega}^\star$ and equipped with an appropriate stopping rule would yield a sample complexity matching the lower bound. In this paper, we provide computationally efficient algorithms to solve (1) and show how these can be used to devise a δ-PAC best action identification algorithm with minimal sample complexity and running in polynomial time. We only assume that we have access to a computationally efficient Oracle, referred to as the `LM` (Linear Maximization) Oracle, identifying the best action should $\boldsymbol{\mu}$ be known (but for any possible $\boldsymbol{\mu}$). This assumption, made in all previous work on combinatorial semi-bandits (see e.g. [JMKK21, PBVP20]), is crucial as indeed, if there is no computationally efficient algorithm solving the offline problem $\operatorname{argmax}_{\boldsymbol{x}\in\mathcal{X}} \langle \boldsymbol{x}, \boldsymbol{\mu} \rangle$ with known $\boldsymbol{\mu}$, there is no hope to solve its online version with unknown $\boldsymbol{\mu}$ in a computationally efficient manner. The assumption holds for a large array of combinatorial sets of actions [S+03], including m-sets, matchings, (source–destination)-paths, spanning trees, matroids (refer to [CCG21b] for a thorough discussion).

[1] We present proof in Appendix K for completeness – see also [JMKK21].

[2] Σ_N denotes the $(N-1)$ dimensional simplex.

The Most-Confusing-Parameter (`MCP`) algorithm. The difficulty of solving (1) lies in the inner optimization problem, i.e., in evaluating the objective function:

$$F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) = \inf_{\boldsymbol{\lambda} \in \mathsf{Alt}(\boldsymbol{\mu})} \left\langle \boldsymbol{\omega}, \frac{(\boldsymbol{\mu} - \boldsymbol{\lambda})^2}{2} \right\rangle = \min_{\boldsymbol{x} \neq \boldsymbol{i}^\star(\boldsymbol{\mu})} f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{\mu}) \tag{2}$$

where $f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{\mu}) = \inf_{\boldsymbol{\lambda} \in \mathcal{C}_{\boldsymbol{x}}} \langle \boldsymbol{\omega}, \frac{(\boldsymbol{\mu}-\boldsymbol{\lambda})^2}{2} \rangle$ and $\mathcal{C}_{\boldsymbol{x}} = \{\boldsymbol{\lambda} \in \mathbb{R}^K : \langle \boldsymbol{\lambda}, \boldsymbol{i}^\star(\boldsymbol{\mu}) - \boldsymbol{x} \rangle < 0\}$. Evaluating $F_{\boldsymbol{\mu}}(\boldsymbol{\omega})$ is required to solve (1), but also in the design of an efficient stopping rule. Our first contribution is `MCP` (Most-Confusing-Parameter), a polynomial time algorithm able to approximate $F_{\boldsymbol{\mu}}(\boldsymbol{\omega})$ for any given $\boldsymbol{\mu}$ and $\boldsymbol{\omega}$. The algorithm's name refers to the fact that by computing $F_{\boldsymbol{\mu}}(\boldsymbol{\omega})$, we implicitly identify the *most confusing parameter* $\boldsymbol{\lambda}^\star \in \arg\inf_{\boldsymbol{\lambda} \in \mathsf{Alt}(\boldsymbol{\mu})} \langle \boldsymbol{\omega}, \frac{(\boldsymbol{\mu}-\boldsymbol{\lambda})^2}{2} \rangle$. The design of `MCP` leverages a Lagrangian relaxation of the optimization problem defining $f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{\mu})$ and exploits the fact that the Lagrange dual function linearly depends on $\boldsymbol{x}$. In turn, this linearity allows us to make use of the `LM` Oracle. From these observations, we show that computing $F_{\boldsymbol{\mu}}(\boldsymbol{\omega})$ boils down to solving a two-player game, for which one of the players can simply update her strategy using the `LM` Oracle. We prove the following informally stated theorem quantifying the performance of the `MCP` algorithm (see Theorem 3 for a more precise statement).

Theorem 1. *For any $(\boldsymbol{\omega}, \boldsymbol{\mu})$, the `MCP` algorithm with precision ϵ and certainty parameter θ returns $\hat{F}$ and $\hat{\boldsymbol{x}}$ satisfying $\mathbb{P}_{\boldsymbol{\mu}}[F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) \leq \hat{F} \leq (1+\epsilon) F_{\boldsymbol{\mu}}(\boldsymbol{\omega})] \geq 1 - \theta$ and $\hat{F} = f_{\hat{\boldsymbol{x}}}(\boldsymbol{\omega}, \boldsymbol{\mu})$. The number of calls to the `LM` Oracle is, almost surely, at most polynomial in K, ϵ^{-1}, and $\ln \theta^{-1}$.*

The Perturbed Frank-Wolfe Sampling (`P-FWS`) algorithm. The `MCP` algorithm allows us to solve the lower-bound problem (1) for any given $\boldsymbol{\mu}$. The latter is initially unknown, but could be estimated. A Track-and-Stop algorithm [GK16] solving (1) with this plug-in estimator in each round would yield asymptotically minimal sample complexity. It would however be computationally expensive. To circumvent this difficulty, as in [WTP21], our algorithm, `P-FWS`, performs a single iteration of the Frank-Wolfe algorithm for the program (1) instantiated with an estimator of $\boldsymbol{\mu}$. To apply the Frank-Wolfe algorithm, `P-FWS` uses stochastic smoothing techniques to approximate the non-differentiable objective function $F_{\boldsymbol{\mu}}$ by a smooth function. To estimate the gradient of the latter, `P-FWS` leverages both the `LM` Oracle and the `MCP` algorithm (more specifically its second output $\hat{\boldsymbol{x}}$). Finally, `P-FWS` stopping rule takes the form of a classical Generalized Likelihood Ratio Test (GLRT) where the estimated objective function is compared to a time-dependent threshold. Hence the stopping rule also requires the `MCP` algorithm. We analyze the sample and computational complexities of `P-FWS`. Our main results are summarized in the following theorem (refer to Theorem 4 for details).

Theorem 2. *For any $\delta \in (0,1)$, `P-FWS` is δ-PAC, and for any $(\epsilon, \tilde{\epsilon}) \in (0,1)$ small enough, its sample complexity satisfies:*

$$\mathbb{E}_{\boldsymbol{\mu}}[\tau] \leq \frac{(1+\tilde{\epsilon})^2}{T^\star(\boldsymbol{\mu})^{-1} - \epsilon} \times H\left(\frac{1}{\delta} \cdot \frac{c(1+\tilde{\epsilon})^2}{T^\star(\boldsymbol{\mu})^{-1} - \epsilon}\right) + \Psi(\epsilon, \tilde{\epsilon}),$$

where $H(x) = \ln(x) + \ln\ln(x)$, $c > 0$ is a universal constant, and $\Psi(\epsilon, \tilde{\epsilon})$ is polynomial in ϵ^{-1}, $\tilde{\epsilon}^{-1}$, K, $\|\boldsymbol{\mu}\|_\infty$, and $\triangle_{\min}^{-1}$, where $\triangle_{\min} = \min_{\boldsymbol{x} \neq \boldsymbol{i}^\star(\boldsymbol{\mu})} \langle \boldsymbol{i}^\star(\boldsymbol{\mu}) - \boldsymbol{x}, \boldsymbol{\mu} \rangle$. Under `P-FWS`, the number of `LM` Oracle calls per round is at most polynomial in $\ln \delta^{-1}$ and K. The total expected number of these calls is also polynomial.

To the best of our knowledge, `P-FWS` is the first polynomial time best action identification algorithm with minimal sample complexity in the high confidence regime (when δ tends to 0). Its sample complexity is also polynomial in K in the moderate confidence regime.

2 Preliminaries

We start by introducing some notation. We use bold lowercase letters (e.g., $\boldsymbol{x}$) for vectors, and bold uppercase letter (e.g., $\mathbf{A}$) for matrices. $\odot$ (resp. $\oplus$) denotes the element-wise product (resp. sum over $\mathbb{Z}_2$). For $\boldsymbol{x} \in \mathbb{R}^K, i \in \mathbb{N}$, $\boldsymbol{x}^i = (x_k^i)_{k \in [K]}$ is the i-th element-wise power of $\boldsymbol{x}$. $D = \max_{\boldsymbol{x} \in \mathcal{X}} \|\boldsymbol{x}\|_1$ denotes the maximum number of arms part of an action. For any $\boldsymbol{\mu} \in \Lambda$, we define the sub-optimality gap of $\boldsymbol{x} \in \mathcal{X}$ as $\triangle_{\boldsymbol{x}}(\boldsymbol{\mu}) = \langle \boldsymbol{i}^\star(\boldsymbol{\mu}) - \boldsymbol{x}, \boldsymbol{\mu} \rangle$, and the minimal gap as $\triangle_{\min}(\boldsymbol{\mu}) = \min_{\boldsymbol{x} \neq \boldsymbol{i}^\star(\boldsymbol{\mu})} \triangle_{\boldsymbol{x}}(\boldsymbol{\mu})$. $\mathbb{P}_{\boldsymbol{\mu}}$ (resp. $\mathbb{E}_{\boldsymbol{\mu}}$) denotes the probability measure (resp. expectation) when the arm rewards are parametrized by $\boldsymbol{\mu}$. Whenever it is clear from the context, we will drop $\boldsymbol{\mu}$ for simplicity, e.g. $\boldsymbol{i}^\star = \boldsymbol{i}^\star(\boldsymbol{\mu})$, $\triangle_{\boldsymbol{x}} = \triangle_{\boldsymbol{x}}(\boldsymbol{\mu})$, and $\triangle_{\min} = \min_{\boldsymbol{x} \neq \boldsymbol{i}^\star} \triangle_{\boldsymbol{x}}$. Refer to Appendix A for an exhaustive table of notation.

2.1 The lower-bound problem

Classical change-of-measure arguments lead to the asymptotic sample complexity lower bound $\mathbb{E}_{\boldsymbol{\mu}}[\tau] \geq T^\star(\boldsymbol{\mu})\mathsf{kl}(\delta, 1-\delta)$ where the characteristic time is defined in (1). To have a chance to develop a computationally efficient best action identification algorithm, we need that the sample complexity lower bound grows at most polynomially in K. This is indeed the case as stated in the following lemma, whose proof is provided in Appendix K.

Lemma 1. *For any $\boldsymbol{\mu} \in \Lambda$, $T^\star(\boldsymbol{\mu}) \leq 4KD\triangle_{\min}(\boldsymbol{\mu})^{-2}$.*

We will use first-order methods to solve the lower-bound problem, and to this aim, we will need to evaluate the gradient w.r.t. $\boldsymbol{\omega}$ of $f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{\mu})$. We can apply the envelop theorem [WTP21] to show that for $(\boldsymbol{\omega}, \boldsymbol{\mu}) \in \Sigma_+ \times \Lambda$,

$$\nabla_{\boldsymbol{\omega}} f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{\mu}) = \frac{(\boldsymbol{\mu} - \boldsymbol{\lambda}^\star_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}))^2}{2},$$

where $\Sigma_+ = \Sigma \cap \mathbb{R}^K_{>0}$, $\boldsymbol{\lambda}^\star_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}) = \mathrm{argmin}_{\boldsymbol{\lambda} \in \mathsf{cl}(\mathcal{C}_{\boldsymbol{x}})} \langle \boldsymbol{\omega}, \frac{(\boldsymbol{\mu}-\boldsymbol{\lambda})^2}{2} \rangle$ and $\mathsf{cl}\,(\mathcal{C}_{\boldsymbol{x}})$ is the closure of $\mathcal{C}_{\boldsymbol{x}}$ (refer to Lemma 19 in Appendix G.2).

2.2 The Linear Maximization Oracle

As mentioned earlier, we assume that we have access to a computationally efficient Oracle, referred to as the `LM` (Linear Maximization) Oracle, identifying the best action if $\boldsymbol{\mu}$ is known. More precisely, as in existing works in combinatorial semi-bandits [KWA+14, PPV19, PBVP20], we make the following assumption.

Assumption 1. *(i) There exists a polynomial-time algorithm identifying $\boldsymbol{i}^\star(\boldsymbol{v})$ for any $\boldsymbol{v} \in \mathbb{R}^K$; (ii) $\mathcal{X}$ is inclusion-wise maximal, i.e., there is no $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{X}$ s.t. $\boldsymbol{x} < \boldsymbol{x}'$; (iii) for each $k \in [K]$, there exists $\boldsymbol{x} \in \mathcal{X}$ such that $x_k = 1$; (iv) $|\mathcal{X}| \geq 2$.*

Assumption 1 holds for combinatorial sets including m-sets, spanning forests, bipartite matching, s-t paths. For completeness, we verify the assumption for these action sets in Appendix J. In the design of our `MCP` algorithm, we will actually need to solve for some $\boldsymbol{v} \in \mathbb{R}^K$ the linear maximization problem $\max\langle \boldsymbol{x}, \boldsymbol{v}\rangle$ over $\mathcal{X} \setminus \{\boldsymbol{i}^\star(\boldsymbol{\mu})\}$; in other words, we will probably need to identify the second best action. Fortunately, this can be done in a computationally efficient manner under Assumption 1. The following lemma formalizes this observation. Its proof, presented in Appendix J, is inspired by Lawler's m-best algorithm [Law72].

Lemma 2. *Let $\boldsymbol{v} \in \mathbb{R}^K$ and $\boldsymbol{x} \in \mathcal{X}$. Under Assumption 1, there exists an algorithm that solves $\max_{\boldsymbol{x}' \in \mathcal{X}: \boldsymbol{x}' \neq \boldsymbol{x}} \langle \boldsymbol{v}, \boldsymbol{x}' \rangle$ by only making at most D queries to the `LM` Oracle.*

3 Solving the lower bound problem: the `MCP` algorithm

Solving the lower bound problem first requires to evaluate its objective function $F_{\boldsymbol{\mu}}(\boldsymbol{\omega})$. A naive approach, enumerating $f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{\mu})$ for all $\boldsymbol{x} \in \mathcal{X} \setminus \{\boldsymbol{i}^\star\}$, would be computationally infeasible. In this section, we present and analyze `MCP`, an algorithm that approximates $F_{\boldsymbol{\mu}}(\boldsymbol{\omega})$ by calling the `LM` Oracle a number of times growing at most polynomially in K.

3.1 Lagrangian relaxation

The first step towards the design of `MCP` consists in considering the Lagrangian relaxation of the optimization problem defining $f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{\mu}) = \inf_{\boldsymbol{\lambda} \in \mathcal{C}_{\boldsymbol{x}}} \langle \boldsymbol{\omega}, \frac{(\boldsymbol{\mu}-\boldsymbol{\lambda})^2}{2} \rangle$ (see e.g., [BV04, Vis21]). For any $(\boldsymbol{\omega}, \boldsymbol{\mu}) \in \Sigma_+ \times \Lambda$ and $\boldsymbol{x} \neq \boldsymbol{i}^\star$, the Lagrangian $\mathcal{L}_{\boldsymbol{\omega},\boldsymbol{\mu}}$ and Lagrange dual function $g_{\boldsymbol{\omega},\boldsymbol{\mu}}$ of this problem are defined as, $\forall \alpha \geq 0$,

$$\mathcal{L}_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{\lambda}, \boldsymbol{x}, \alpha) = \left\langle \boldsymbol{\omega}, \frac{(\boldsymbol{\mu}-\boldsymbol{\lambda})^2}{2} \right\rangle + \alpha \left\langle \boldsymbol{i}^\star - \boldsymbol{x}, \boldsymbol{\lambda} \right\rangle \quad \text{and} \quad g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha) = \inf_{\boldsymbol{\lambda} \in \mathbb{R}^K} \mathcal{L}_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{\lambda}, \boldsymbol{x}, \alpha),$$

respectively. The following proposition, proved in Appendix C.1, provides useful properties of $g_{\boldsymbol{\omega},\boldsymbol{\mu}}$:

Proposition 1. *Let $(\boldsymbol{\omega}, \boldsymbol{\mu}) \in \Sigma_+ \times \Lambda$ and $\boldsymbol{x} \in \mathcal{X} \setminus \{\boldsymbol{i}^\star(\boldsymbol{\mu})\}$.*
(a) The Lagrange dual function is linear in $\boldsymbol{x}$. More precisely, $g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha) = c_{\boldsymbol{\omega},\boldsymbol{\mu}}(\alpha) + \langle \boldsymbol{\ell}_{\boldsymbol{\omega},\boldsymbol{\mu}}(\alpha), \boldsymbol{x} \rangle$

where $c_{\boldsymbol{\omega},\boldsymbol{\mu}}(\alpha) = \alpha \left\langle \boldsymbol{\mu} - \frac{\alpha}{2}\boldsymbol{\omega}^{-1}, \boldsymbol{i}^\star(\boldsymbol{\mu}) \right\rangle$ *and* $\boldsymbol{\ell}_{\boldsymbol{\omega},\boldsymbol{\mu}}(\alpha) = -\alpha\left(\boldsymbol{\mu} + \frac{\alpha}{2}\boldsymbol{\omega}^{-1} \odot (\mathbf{1}_K - 2\boldsymbol{i}^\star(\boldsymbol{\mu}))\right)$.
(b) $g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \cdot)$ *is strictly concave (for any fixed* $\boldsymbol{x}$*).*
(c) $f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{\mu}) = \max_{\alpha \geq 0} g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha)$ *is attained by* $\alpha^\star_{\boldsymbol{x}} = \frac{\triangle_{\boldsymbol{x}}(\boldsymbol{\mu})}{\langle \boldsymbol{x} \oplus \boldsymbol{i}^\star(\boldsymbol{\mu}), \boldsymbol{\omega}^{-1}\rangle}$.
(d) $\left\|\boldsymbol{\ell}_{\boldsymbol{\omega},\boldsymbol{\mu}}(\alpha^\star_{\boldsymbol{x}})\right\|_1 \leq L_{\boldsymbol{\omega},\boldsymbol{\mu}} = 4D^2K \left\|\boldsymbol{\mu}\right\|^2_\infty \left\|\boldsymbol{\omega}^{-1}\right\|_\infty$.

From Proposition 1 (c), strong duality holds for the program defining $f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{\mu})$, and we conclude:

$$F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) = \min_{\boldsymbol{x} \neq \boldsymbol{i}^\star} \max_{\alpha \geq 0} g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha). \tag{3}$$

$F_{\boldsymbol{\mu}}(\boldsymbol{\omega})$ can hence be seen as the value in a two-player game. The aforementioned properties of the Lagrange dual function will help to compute this value.

3.2 Solving the two-player game with no regret

There is a rich and growing literature on solving zero-sum games using no-regret algorithms, see for example [RS13, ALLW18, DFG21, ZODS21]. Our game has the particularity that the $\boldsymbol{x}$-player has a discrete combinatorial action set whereas the α-player has a convex action set. Importantly, for this game, we wish not only to estimate its value $F_{\boldsymbol{\mu}}(\boldsymbol{\omega})$ but also an *equilibrium* action $\boldsymbol{x}_e$ such that $F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) = \max_{\alpha \geq 0} g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}_e, \alpha)$. Indeed, an estimate of $\boldsymbol{x}_e$ will be needed when implementing the Frank-Wolfe algorithm and more specifically when estimating the gradient of $F_{\boldsymbol{\mu}}(\boldsymbol{\omega})$. To return such an estimate, one could think of leveraging results from the recent literature on last-iterate convergence, see e.g. [DP19, GPDO20, LNP+21, WLZL21, APFS22, AAS+23]. However, most of these results concern saddle-point problems only, and are not applicable in our setting. Here, we adopt a much simpler solution, and take advantage of the properties of the Lagrange dual function $g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha)$ to design an iterative procedure directly leading to estimates of $(F_{\boldsymbol{\mu}}(\boldsymbol{\omega}), \boldsymbol{x}_e)$. In this procedure, the two players successively update their actions until a stopping criterion is met, say up to the N-th iterations. The procedure generates a sequence $\{(\boldsymbol{x}^{(n)}, \alpha^{(n)})\}_{1 \leq n \leq N}$, and from this sequence, estimates $(\hat{F}, \hat{\boldsymbol{x}})$ of $(F_{\boldsymbol{\mu}}(\boldsymbol{\omega}), \boldsymbol{x}_e)$. The details of the resulting `MCP` algorithm are presented in Algorithm 1.

$\boldsymbol{x}$-player. We use a variant of the Follow-the-Perturbed-Leader (FTPL) algorithm [Han57, KV05]. The $\boldsymbol{x}$-player updates her action as follows:

$$\boldsymbol{x}^{(n)} \in \operatorname*{argmin}_{\boldsymbol{x} \neq \boldsymbol{i}^\star} \left(\sum_{m=1}^{n-1} g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha^{(m)}) + \left\langle \frac{\boldsymbol{\mathcal{Z}}_n}{\eta_n}, \boldsymbol{x} \right\rangle \right) = \operatorname*{argmin}_{\boldsymbol{x} \neq \boldsymbol{i}^\star} \left(\left\langle \sum_{m=1}^{n-1} \boldsymbol{\ell}_{\boldsymbol{\omega},\boldsymbol{\mu}}(\alpha^{(m)}) + \frac{\boldsymbol{\mathcal{Z}}_n}{\eta_n}, \boldsymbol{x} \right\rangle \right),$$

where $\boldsymbol{\mathcal{Z}}_n$ is a random vector, exponentially distributed and with unit mean ($\{\boldsymbol{\mathcal{Z}}_n\}_{n \geq 1}$ are i.i.d.). Compared to the standard FTPL algorithm, we vary learning rate η_n over time to get *anytime* guarantees (as we do not know a priori when the iterative procedure will stop). This kind of time-varying learning rate was also used in [Neu15] with a similar motivation. Note that thanks to the linearity of $g_{\boldsymbol{\omega},\boldsymbol{\mu}}$ and Lemma 2, the $\boldsymbol{x}$-player update can be computed using at most D calls to the `LM` Oracle.

α-player and `MCP` outputs. From Proposition 1, $f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{\mu}) = \max_{\alpha \geq 0} g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha)$. This suggests that the α-player can just implement a best-response strategy: after the $\boldsymbol{x}$-player action $\boldsymbol{x}^{(n)}$ is selected, the α-player chooses $\alpha^{(n)} = \alpha^\star_{\boldsymbol{x}^{(n)}} = \frac{\triangle_{\boldsymbol{x}^{(n)}}(\boldsymbol{\mu})}{\langle \boldsymbol{x}^{(n)} \oplus \boldsymbol{i}^\star(\boldsymbol{\mu}), \boldsymbol{\omega}^{-1}\rangle}$. This choice ensures that $f_{\boldsymbol{x}^{(n)}}(\boldsymbol{\omega}, \boldsymbol{\mu}) = g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}^{(n)}, \alpha^{(n)})$, and suggests natural outputs for `MCP`: should it stops after N iterations, it can return $\hat{F} = \min_{n \in [N]} g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}^{(n)}, \alpha^{(n)})$ and $\hat{\boldsymbol{x}} \in \operatorname{argmin}_{n \in [N]} g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}^{(n)}, \alpha^{(n)})$.

Stopping criterion. The design of the `MCP` stopping criterion relies on the convergence analysis and regret from the $\boldsymbol{x}$-player perspective of the above iterative procedure, which we present in the next subsection. This convergence will be controlled by $\boldsymbol{\ell}_{\boldsymbol{\omega},\boldsymbol{\mu}}(\alpha^\star_{\boldsymbol{x}})$ and its upper bound $L_{\boldsymbol{\omega},\boldsymbol{\mu}}$ derived in Proposition 1. Introducing $c_\theta = L_{\boldsymbol{\omega},\boldsymbol{\mu}}(4\sqrt{K(\ln K + 1)} + \sqrt{\ln(\theta^{-1})/2})$, the `MCP` stopping criterion is: $\sqrt{n} > c_\theta(1+\epsilon)/(\epsilon \hat{F})$. Since $\sqrt{n}$ strictly increases with n and since $\hat{F} \geq F_{\boldsymbol{\mu}}(\boldsymbol{\omega})$, this criterion ensures that the algorithm terminates in a finite number of iterates. Moreover, as shown in the next subsection, it also ensures that $\hat{F}$ returned by `MCP` is an $(1+\epsilon)$-approximation of $F_{\boldsymbol{\mu}}(\boldsymbol{\omega})$ with probability at least $1 - \theta$.

Algorithm 1: (ϵ, θ)-`MCP`$(\boldsymbol{\omega}, \boldsymbol{\mu})$

initialization: $n = 1, \hat{F} = \infty, c_\theta = L_{\boldsymbol{\omega},\boldsymbol{\mu}} \left(4\sqrt{K(\ln K + 1)} + \sqrt{\ln(\theta^{-1})/2}\right)$**;**

while $(n = 1)$ ***or*** $(n > 1$ ***and*** $\sqrt{n} \leq c_\theta(1+\epsilon)/(\epsilon\hat{F}))$ **do**

  Sample $\boldsymbol{\mathcal{Z}}_n \sim \exp(1)^K$ and set $\eta_n = \sqrt{K(\ln K + 1)/(4nL^2_{\boldsymbol{\omega},\boldsymbol{\mu}})}$;

  $\boldsymbol{x}^{(n)} \leftarrow \mathrm{argmin}_{\boldsymbol{x} \neq \boldsymbol{i}^\star(\boldsymbol{\mu})} \left(\sum_{m=1}^{n-1} g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha^{(m)}) + \langle \boldsymbol{\mathcal{Z}}_n, \boldsymbol{x} \rangle / \eta_n\right)$ *(ties broken arbitrarily)*;

  $\alpha^{(n)} \leftarrow \mathrm{argmax}_{\alpha \geq 0}\, g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}^{(n)}, \alpha)$ *(uniqueness ensured by Proposition 1 (c))*;

  if $g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}^{(n)}, \alpha^{(n)}) < \hat{F}$ **then** $(\hat{F}, \hat{\boldsymbol{x}}) \leftarrow (g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}^{(n)}, \alpha^{(n)}), \boldsymbol{x}^{(n)})$;

  $n \leftarrow n + 1$;

end

return $(\hat{F}, \hat{\boldsymbol{x}})$;

3.3 Performance analysis of the `MCP` algorithm

We start the analysis by quantifying the regret from the $\boldsymbol{x}$-player perspective of `MCP` before its stops. The following lemma is proved in Appendix C.3.

Lemma 3. *Let $N \in \mathbb{N}$. Under (ϵ, θ)-`MCP`$(\boldsymbol{\omega}, \boldsymbol{\mu})$,*

$$\mathbb{P}\left[\frac{1}{N}\sum_{n=1}^{N} g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}^{(n)}, \alpha^{(n)}) - \frac{1}{N}\min_{\boldsymbol{x} \neq \boldsymbol{i}^\star} \sum_{n=1}^{N} g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha^{(n)}) \leq \frac{c_\theta}{\sqrt{N}}\right] \geq 1 - \theta.$$

Observe that on the one hand,

$$\frac{1}{N}\sum_{n=1}^{N} g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}^{(n)}, \alpha^{(n)}) \geq \min_{n \in [N]} g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}^{(n)}, \alpha^{(n)}) = \hat{F} \tag{4}$$

always holds. On the other hand, if $\boldsymbol{x}_e \in \mathrm{argmin}_{\boldsymbol{x} \neq \boldsymbol{i}^\star} \max_{\alpha \geq 0} g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha)$, then we have:

$$\frac{1}{N}\min_{\boldsymbol{x} \neq \boldsymbol{i}^\star} \sum_{n=1}^{N} g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha^{(n)}) \leq \frac{1}{N}\sum_{n=1}^{N} g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}_e, \alpha^{(n)}) \leq \max_{\alpha \geq 0} g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}_e, \alpha) = F_{\boldsymbol{\mu}}(\boldsymbol{\omega}). \tag{5}$$

We conclude that for N such that $\sqrt{N} \geq \frac{c_\theta(1+\epsilon)}{\epsilon\hat{F}}$, Lemma 3 together with the inequalities (4) and (5) imply that $\hat{F} - F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) \leq \frac{c_\theta}{\sqrt{N}} \leq \frac{\epsilon\hat{F}}{1+\epsilon}$ holds with probability at least $1 - \theta$. Hence $\mathbb{P}\left[\hat{F} \leq (1+\epsilon)F_{\boldsymbol{\mu}}(\boldsymbol{\omega})\right] \geq 1 - \theta$. From this observation, we essentially deduce the following theorem, whose complete proof is given in Appendix C.2.

Theorem 3. *Let $\epsilon, \theta \in (0, 1)$. Under Assumption 1, for any $(\boldsymbol{\omega}, \boldsymbol{\mu}) \in \Sigma_+ \times \Lambda$, the (ϵ, θ)-`MCP`$(\boldsymbol{\omega}, \boldsymbol{\mu})$ algorithm outputs $(\hat{F}, \hat{\boldsymbol{x}})$ satisfying $\mathbb{P}\left[F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) \leq \hat{F} \leq (1+\epsilon)F_{\boldsymbol{\mu}}(\boldsymbol{\omega})\right] \geq 1 - \theta$ and $\hat{F} = \max_{\alpha \geq 0} g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\hat{\boldsymbol{x}}, \alpha)$. Moreover, the number of* `LM` *Oracle calls the algorithm does is almost surely at most $\left\lceil \frac{c_\theta^2(1+\epsilon)^2}{\epsilon^2 F_{\boldsymbol{\mu}}(\boldsymbol{\omega})^2} \right\rceil = \mathcal{O}\left(\frac{\|\boldsymbol{\mu}\|_\infty^4 \left\|\boldsymbol{\omega}^{-1}\right\|_\infty^2 K^3 D^5 \ln K \ln \theta^{-1}}{\epsilon^2 F_{\boldsymbol{\mu}}(\boldsymbol{\omega})^2}\right)$.*

4 The Perturbed Frank-Wolfe Sampling (`P-FWS`) algorithm

To identify an optimal sampling strategy, rather than solving the lower-bound problem in each round as a Track-and-Stop algorithm would [GK16], we devise `P-FWS`, an algorithm that performs a single iteration of the Frank-Wolfe algorithm for the lower-bound problem instantiated with an estimator of $\boldsymbol{\mu}$. This requires us to first smooth the objective function $F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) = \min_{\boldsymbol{x} \neq \boldsymbol{i}^\star} f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{\mu})$ (the latter is not differentiable at points $\boldsymbol{\omega}$ where the $\min$ is achieved for several sub-optimal actions $\boldsymbol{x}$). To this aim, we cannot leverage the same technique as in [WTP21], where r-subdifferential subspaces are built from gradients of $f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{\mu})$. These subspaces could indeed be generated by a number of vectors (here gradients) exponentially growing with K. Instead, to cope with the combinatorial decision sets, `P-FWS` applies more standard stochastic smoothing techniques as described in the next subsection. All the ingredients of `P-FWS` are gathered in §4.2. By design, the algorithm just leverages the `MCP` algorithm as a subroutine, and hence only requires the `LM` Oracle. In §4.3, we analyze the performance of `P-FWS`.

4.1 Smoothing the objective function $F_{\boldsymbol{\mu}}$

Here, we present and analyze a standard stochastic technique to smooth a function Φ. In `P-FWS`, this technique will be applied to the objective function $\Phi = F_{\boldsymbol{\mu}}$. Let $\Phi : \mathbb{R}^K_{>0} \mapsto \mathbb{R}$ be a concave and ℓ-Lipschitz function. Assume that the set of points where Φ is not differentiable is of Lebesgue-measure zero. To smooth Φ, we can take its average value in a neighborhood of the point considered, see e.g. [FKM05]. Formally, we define the *stochastic smoothed* approximate of Φ as:

$$\bar{\Phi}_\eta(\boldsymbol{\omega}) = \mathbb{E}_{\boldsymbol{\mathcal{Z}} \sim \text{Uniform}(B_2)}[\Phi(\boldsymbol{\omega} + \eta \boldsymbol{\mathcal{Z}})] \,, \tag{6}$$

where $B_2 = \{\boldsymbol{v} \in \mathbb{R}^K : \|\boldsymbol{v}\|_2 \leq 1\}$ and $\eta \in (0, \min_{k \in [K]} \omega_k)$. The following proposition lists several properties of this smoothed function, and gathers together some of the results from [DBW12], see Appendix H for more details.

Proposition 2. *For any $\boldsymbol{\omega} \in \Sigma_+$ and $\eta \in (0, \min_{k\in[K]} \omega_k)$, $\bar{\Phi}_\eta$ satisfies: (i) $\Phi(\boldsymbol{\omega}) - \eta\ell \leq \bar{\Phi}_\eta(\boldsymbol{\omega}) \leq \Phi(\boldsymbol{\omega})$; (ii) $\nabla \bar{\Phi}_{\boldsymbol{\mu},\eta}(\boldsymbol{\omega}) = \mathbb{E}_{\boldsymbol{\mathcal{Z}} \sim \text{Uniform}(B_2)}[\nabla \Phi_{\boldsymbol{\mu}}(\boldsymbol{\omega} + \eta \boldsymbol{\mathcal{Z}})]$; (iii) $\bar{\Phi}_\eta$ is $\frac{\ell K}{\eta}$-smooth; (iv) if $\eta > \eta' > 0$, then $\bar{\Phi}_{\eta'}(\boldsymbol{\omega}) \geq \bar{\Phi}_\eta(\boldsymbol{\omega})$.*

Note that with (i), we may control the approximation error between $\bar{\Phi}_\eta$ and Φ by η. (ii) and (iii) ensure the differentiability and smoothness of $\bar{\Phi}_\eta$ respectively. (iii) is equivalent to $\bar{\Phi}_\eta(\boldsymbol{\omega}') \leq \bar{\Phi}_\eta(\boldsymbol{\omega}) + \langle \nabla \bar{\Phi}_\eta(\boldsymbol{\omega}), \boldsymbol{\omega}' - \boldsymbol{\omega} \rangle + \frac{\ell K}{2\eta} \|\boldsymbol{\omega} - \boldsymbol{\omega}'\|_2^2 , \ \forall \boldsymbol{\omega}, \boldsymbol{\omega}' \in \Sigma_+$. Finally, (iv) stems from the concavity of Φ, and implies that the value $\bar{\Phi}_\eta(\boldsymbol{\omega})$ monotonously increases while η decreases, and it is upper bounded by $\Phi(\boldsymbol{\omega})$ thanks to (i). The above results hold for $\Phi = F_{\boldsymbol{\mu}}$. Indeed, first it is clear that the definition (2) of $F_{\boldsymbol{\mu}}$ can be extended to $\mathbb{R}^K$; then, it can be shown that $F_{\boldsymbol{\mu}}$ is Lipschitz-continuous and almost-everywhere differentiable – refer to Appendices I and H for formal proofs.

4.2 The algorithm

Before presenting `P-FWS`, we need to introduce the following notation. For $t \geq 1$, $k \in [K]$, we define $N_k(t) = \sum_{s=1}^t \mathbb{1}\{x_k(s) = 1\}$, $\hat{\omega}_k(t) = N_k(t)/t$, and $\hat{\mu}_k(t) = \sum_{s=1}^t y_k(s) \mathbb{1}\{x_k(s) = 1\} / N_k(t)$ when $N_k(t) > 0$.

Sampling rule. The design of the sampling rule is driven by the following objectives: (i) the empirical allocation should converge to the solution of the lower-bound problem (1), and (ii) the number of calls to the `LM` Oracle should be controlled. To meet the first objective, we need in the Frank-Wolfe updates to plug an accurate estimator of $\boldsymbol{\mu}$ in. The accuracy of our estimator will be guaranteed by alternating between *forced exploration* and *FW update* sampling phases. Now for the second objective, we also use forced exploration phases when in a Frank-Wolfe update, the required number of calls to the `LM` Oracle predicted by the upper bound presented in Theorem 3 is too large. In view of Lemma 1 and Theorem 3, this happens in round t if $\|\hat{\boldsymbol{\mu}}(t-1)\|_\infty$ or $\triangle_{\min}(\hat{\boldsymbol{\mu}}(t-1))^{-1}$ is too large. Next, we describe the forced exploration and Frank-Wolfe update phases in detail.

Forced exploration. Initially, `P-FWS` applies the `LM` Oracle to compute the *forced exploration set* $\mathcal{X}_0 = \{\boldsymbol{i}^\star(\boldsymbol{e}_k) : k \in [K]\}$, where $\boldsymbol{e}_k$ is the K-dimensional vector whose k-th component is equal to one and zero elsewhere. `P-FWS` then selects each action in $\mathcal{X}_0$ once. Note that Assumption 1 (iii) ensures that the k-th component of $\boldsymbol{i}^\star(\boldsymbol{e}_k)$ is equal to one. In turn, this ensures that $\mathcal{X}_0$ is a $[K]$-covering set, and that playing actions from $\mathcal{X}_0$ is enough to estimate $\boldsymbol{\mu}$. `P-FWS` starts an exploration phase at rounds t such that $\sqrt{t/|\mathcal{X}_0|}$ is an integer or such that the maximum of $\triangle_{\min}(\hat{\boldsymbol{\mu}}(t-1))^{-1}$ and $\|\hat{\boldsymbol{\mu}}(t-1)\|_\infty$ is larger than $\sqrt{t-1}$. Whenever this happens, `P-FWS` pulls each $\boldsymbol{x} \in \mathcal{X}_0$ once.

Frank-Wolfe updates. When in round t, the algorithm is not in a forced exploration phase, it implements an iteration of the Frank-Wolfe algorithm applied to maximize the smoothed function $\bar{F}_{\hat{\boldsymbol{\mu}}(t-1),\eta_t}(\hat{\boldsymbol{\omega}}(t-1)) = \mathbb{E}_{\boldsymbol{\mathcal{Z}} \sim \text{Uniform}(B_2)}\left[F_{\hat{\boldsymbol{\mu}}(t-1)}(\hat{\boldsymbol{\omega}}(t-1) + \eta_t \boldsymbol{\mathcal{Z}})\right]$. The sequence of parameters $\{\eta_t\}_{t \geq 1}$ is chosen to ensure that η_t chosen in $(0, \min_k \hat{\omega}_k(t))$, and hence $\hat{\boldsymbol{\omega}}(t-1) + \eta_t \boldsymbol{\mathcal{Z}} \in \mathbb{R}^K_{>0}$. Also note that in a round t where the algorithm is not in a forced exploration phase,

Algorithm 2: P-FWS $(\{(\epsilon_t, \eta_t, n_t, \rho_t, \theta_t)\}_t)$

initialization:
  for $k = 1, \ldots, K$ **do**
    $\mathcal{X}_0 \leftarrow \mathrm{argmax}_{\boldsymbol{x}\in\mathcal{X}} \langle \boldsymbol{e}_k, \boldsymbol{x}\rangle$ (tie broken arbitrarily)
  end
  Sample $\boldsymbol{x} \in \mathcal{X}_0$ in a round-robin manner for $4|\mathcal{X}_0|$ rounds; update $\hat{\boldsymbol{\mu}}(4|\mathcal{X}_0|)$ and $\hat{\boldsymbol{\omega}}(4|\mathcal{X}_0|)$;
for $t = 4|\mathcal{X}_0| + 1, \cdots$ **do**
  if $\sqrt{t/|\mathcal{X}_0|} \in \mathbb{N}$ **or** $\max\{\triangle_{\min}(\hat{\boldsymbol{\mu}}(t-1))^{-1}, \|\hat{\boldsymbol{\mu}}(t-1)\|_\infty\} > \sqrt{t-1}$ **then**
    Sample each $\boldsymbol{x} \in \mathcal{X}_0$ once, update $\hat{\boldsymbol{\mu}}(t)$ and $\hat{\boldsymbol{\omega}}(t)$, and $t \leftarrow t + |\mathcal{X}_0| - 1$;
  else
    Compute $\nabla \tilde{F}_{\hat{\boldsymbol{\mu}}(t-1),\eta_t,n_t}(\hat{\boldsymbol{\omega}}(t-1))$ by (ρ_t, θ_t)-MCP algorithm;
    $\boldsymbol{x}(t) \leftarrow \boldsymbol{i}^\star\left(\nabla \tilde{F}_{\hat{\boldsymbol{\mu}}(t-1),\eta_t,n_t}(\hat{\boldsymbol{\omega}}(t-1))\right)$;
    Sample $\boldsymbol{x}(t)$ and update $\hat{\boldsymbol{\mu}}(t)$ and $\hat{\boldsymbol{\omega}}(t)$;
  end
  if $\max\{\triangle_{\min}(\hat{\boldsymbol{\mu}}(t))^{-1}, \|\hat{\boldsymbol{\mu}}(t)\|_\infty\} \le \sqrt{t}$ **then**
    $\hat{F}_t \leftarrow \left(\epsilon_t, \delta/t^2\right)$-MCP$(\hat{\boldsymbol{\omega}}(t), \hat{\boldsymbol{\mu}}(t))$;
    if $t\hat{F}_t > (1+\epsilon_t)\,\beta\left(t, \left(1 - \frac{1}{4|\mathcal{X}_0|}\right)\delta\right)$ **then** break;
end
return $\hat{\boldsymbol{\imath}} = \boldsymbol{i}^\star(\hat{\boldsymbol{\mu}}(t))$;

by definition $\triangle_{\min}(\hat{\boldsymbol{\mu}}(t-1)) > 0$. This implies that $\hat{\boldsymbol{\mu}}(t-1) \in \Lambda$ and that $F_{\hat{\boldsymbol{\mu}}(t-1)}$ and $\bar{F}_{\hat{\boldsymbol{\mu}}(t-1),\eta_t}(\hat{\boldsymbol{\omega}}(t-1))$ are well-defined. Now an ideal FW update would consist in playing an action $\boldsymbol{i}^\star(\nabla \bar{F}_{\hat{\boldsymbol{\mu}}(t-1),\eta_t}(\hat{\boldsymbol{\omega}}(t-1))) = \mathrm{argmax}_{\boldsymbol{x}\in\mathcal{X}} \left\langle \nabla \bar{F}_{\hat{\boldsymbol{\mu}}(t-1),\eta_t}(\hat{\boldsymbol{\omega}}(t-1)), \boldsymbol{x}\right\rangle$, see e.g. [Jag13]. Unfortunately, we do not have access to $\nabla \bar{F}_{\hat{\boldsymbol{\mu}}(t-1),\eta_t}(\hat{\boldsymbol{\omega}}(t-1))$. But the latter can be approximated, as suggested in Proposition 2 (ii), by $\nabla \tilde{F}_{\hat{\boldsymbol{\mu}}(t-1),\eta_t,n_t}(\hat{\boldsymbol{\omega}}(t-1)) = \frac{1}{n_t}\sum_{m=1}^{n_t} \nabla f_{\hat{\boldsymbol{x}}_m}(\hat{\boldsymbol{\omega}}(t-1) + \eta_t \boldsymbol{\mathcal{Z}}_m, \hat{\boldsymbol{\mu}}(t-1))$, where $\boldsymbol{\mathcal{Z}}_1, \cdots, \boldsymbol{\mathcal{Z}}_{n_t} \overset{i.i.d.}{\sim} \mathrm{Uniform}(B_2)$, $\hat{\boldsymbol{x}}_m$ is the action return by (ρ_t, θ_t)-MCP$(\hat{\boldsymbol{\omega}}(t-1) + \eta_t \boldsymbol{\mathcal{Z}}_m, \hat{\boldsymbol{\mu}}(t-1))$. P-FWS uses this approximation and the LM Oracle to select the action: $\boldsymbol{x}(t) \in \boldsymbol{i}^\star\left(\nabla \tilde{F}_{\hat{\boldsymbol{\mu}}(t-1),\eta_t,n_t}(\hat{\boldsymbol{\omega}}(t-1))\right)$. The choices of the parameters η_t, n_t, ρ_t and θ_t do matter. η_t impacts the sample complexity and should converge to 0 as $t \to \infty$ so that $\bar{F}_{\boldsymbol{\mu},\eta_t}(\boldsymbol{\omega}) \to F_{\boldsymbol{\mu}}(\boldsymbol{\omega})$ at any point $\boldsymbol{\omega} \in \Sigma_+$ (this is a consequence of Proposition 2 (i)(iv)). η_t should not decay too fast however as it would alter the smoothness of $\bar{F}_{\boldsymbol{\mu},\eta_t}$. We will show that η_t should actually decay as $1/\sqrt{t}$. (n_t, ρ_t, θ_t) impact the trade-off between the sample complexity and the computational complexity of the algorithm. We let $n_t \to \infty$ and $(\rho_t, \theta_t) \to 0$ as $t \to \infty$ so that $\left\langle \nabla \tilde{F}_{\boldsymbol{\mu},\eta_t,n_t}(\boldsymbol{\omega}) - \nabla \bar{F}_{\boldsymbol{\mu},\eta_t}(\boldsymbol{\omega}), \boldsymbol{x}\right\rangle \to 0$ for any $(\boldsymbol{\omega}, \boldsymbol{x}) \in \Sigma_+ \times \mathcal{X}$.

Stopping and decision rule. As often in best arm identification algorithms, the P-FWS stopping rule takes the form of a GLRT:

$$\tau = \inf\left\{ t > 4|\mathcal{X}_0| : \frac{t\hat{F}_t}{1+\epsilon_t} > \beta\left(t, \left(1 - \frac{1}{4|\mathcal{X}_0|}\right)\delta\right), \max\left\{\triangle_{\min}(\hat{\boldsymbol{\mu}}(t))^{-1}, \|\hat{\boldsymbol{\mu}}(t)\|_\infty\right\} \le \sqrt{t}\right\}, \tag{7}$$

where $\epsilon_t \in \mathbb{R}_{>0}$, $\hat{F}_t$ is returned by the $(\epsilon_t, \delta/t^2)$-MCP$(\hat{\boldsymbol{\omega}}(t), \hat{\boldsymbol{\mu}}(t))$ algorithm. The function β satisfies

$$\forall t \ge 1, \quad \left(tF_{\hat{\boldsymbol{\mu}}(t)}(\boldsymbol{\omega}(t)) \ge \beta(t,\delta)\right) \implies \left(\mathbb{P}_{\boldsymbol{\mu}}[\boldsymbol{i}^\star(\hat{\boldsymbol{\mu}}(t)) \ne \boldsymbol{i}^\star(\boldsymbol{\mu})] \le \delta\right), \tag{8}$$

$$\exists c_1, c_2 > 0 : \quad \forall t \ge c_1, \beta(t,\delta) \le \ln\left(\frac{c_2 t}{\delta}\right). \tag{9}$$

Examples of function β satisfying the above conditions can be found in [GK16, JP20, KK21]. The condition (8) will ensure the δ-correctness of P-FWS, whereas (9) will control its sample complexity. Finally, the action returned by P-FWS is simply defined as $\hat{\boldsymbol{\imath}} = \boldsymbol{i}^\star(\hat{\boldsymbol{\mu}}(\tau))$. The complete pseudo-code of P-FWS is presented in Algorithm 2.[3]

[3] Our Julia implementation could be found at https://github.com/rctzeng/NeurIPS2023-PerturbedFWS.

4.3 Non-asymptotic performance analysis of `P-FWS`

The following theorem provides an upper bound of the sample complexity of `P-FWS` valid for any confidence level δ, as well as the computational complexity of the algorithm.

Theorem 4. *Let $\boldsymbol{\mu} \in \Lambda$ and $\delta \in (0,1)$. If* `P-FWS` *is parametrized using*

$$(\epsilon_t, \eta_t, n_t, \rho_t, \theta_t) = \left(t^{-\frac{1}{9}}, \frac{1}{4\sqrt{t|\mathcal{X}_0|}}, \left\lceil t^{\frac{1}{4}} \right\rceil, \frac{1}{16tD^2|\mathcal{X}_0|}, \frac{1}{t^{\frac{1}{4}} e^{\sqrt{t}}} \right), \tag{10}$$

then (i) the algorithm finishes in finite time almost surely and $\mathbb{P}_{\boldsymbol{\mu}}[\hat{\boldsymbol{\imath}} \neq \boldsymbol{i}^\star(\boldsymbol{\mu})] \leq \delta$; (ii) its sample complexity satisfies $\mathbb{P}_{\boldsymbol{\mu}}\left[\limsup_{\delta \to 0} \frac{\tau}{\ln \delta^{-1}} \leq T^\star(\boldsymbol{\mu})\right] = 1$ and for any $\epsilon, \tilde{\epsilon} \in (0,1)$ with $\epsilon < \min\{1, \frac{2D^2 \triangle_{\min}^2}{K}, \frac{D^2\|\boldsymbol{\mu}\|_\infty^2}{3}\}$,

$$\mathbb{E}_{\boldsymbol{\mu}}[\tau] \leq \frac{(1+\tilde{\epsilon})^2}{T^\star(\boldsymbol{\mu})^{-1} - 6\epsilon} \times H\left(\frac{1}{\delta} \cdot \frac{4c_2}{3} \cdot \frac{(1+\tilde{\epsilon})^2}{T^\star(\boldsymbol{\mu})^{-1} - 6\epsilon} \right) + \Psi(\epsilon, \tilde{\epsilon}),$$

where $H(x) = \ln x + \ln\ln x + 1$ and $\Psi(\epsilon, \tilde{\epsilon})$ (refer to (34) for a detailed expression) is polynomial in ϵ^{-1}, $\tilde{\epsilon}^{-1}$, K, $\|\boldsymbol{\mu}\|_\infty$, and $\triangle_{\min}(\boldsymbol{\mu})^{-1}$; (iii) the expected number of `LM` *Oracle calls is upper bounded by a polynomial in $\ln \delta^{-1}$, K, $\|\boldsymbol{\mu}\|_\infty$, and $\triangle_{\min}(\boldsymbol{\mu})^{-1}$.*

The above theorem establishes the statistical asymptotic optimality of `P-FWS` since it implies that $\limsup_{\delta \to 0} \mathbb{E}_{\boldsymbol{\mu}}[\tau] / \ln(1/\delta) \leq (1+\tilde{\epsilon})^2/(T^\star(\boldsymbol{\mu})^{-1} - 6\epsilon)$. This upper bound matches the sample complexity lower bound (1) when $\epsilon \to 0$ and $\tilde{\epsilon} \to 0$.

Proof sketch. The complete proof of Theorem 4 is presented in Appendix D.

(i) Correctness. To establish the δ-correctness of the algorithm, we introduce the event $\mathcal{G}$ under which $\hat{F}_t$, computed by $(\epsilon_t, \delta/t^2)$-`MCP`$(\hat{\boldsymbol{\omega}}(t), \hat{\boldsymbol{\mu}}(t))$, is an $(1+\epsilon_t)$-approximation of $F_{\hat{\boldsymbol{\mu}}(t)}(\hat{\boldsymbol{\omega}}(t))$ in each round $t \geq 4|\mathcal{X}_0| + 1$. From Theorem 3, we deduce that $\mathbb{P}_{\boldsymbol{\mu}}[\mathcal{G}^c] \leq \sum_{t=4|\mathcal{X}_0|+1}^{\infty} \delta/t^2 \leq \delta/4|\mathcal{X}_0|$. In view of (8), this implies that $\mathbb{P}_{\boldsymbol{\mu}}[\hat{\boldsymbol{\imath}} \neq \boldsymbol{i}^\star(\boldsymbol{\mu})] \leq \delta$.

(ii) Non-asymptotic sample complexity upper bound.

Step 1. (Concentration and certainty equivalence) We first define two *good* events, $\mathcal{E}_t^{(1)}$ and $\mathcal{E}_t^{(2)}$. $\mathcal{E}_t^{(2)}$ corresponds to the event where $\hat{\boldsymbol{\mu}}(t)$ is close to $\boldsymbol{\mu}$, and its occurrence probability can be controlled using the forced exploration rounds and concentration inequalities. Under $\mathcal{E}_t^{(1)}$, the selected action $\boldsymbol{x}(t)$ is close to the ideal FW update. Again using concentration results and the performance guarantees of `MCP` given in Theorem 3, we can control the occurrence probability of $\mathcal{E}_t^{(2)}$. Overall, we show that $\sum_{t=1}^{\infty} \mathbb{P}_{\boldsymbol{\mu}}[(\mathcal{E}_t^{(1)} \cap \mathcal{E}_t^{(2)})^c] < \infty$. To this aim, we derive several important continuity results presented in Appendix G. These results essentially allow us to study the convergence of the smoothed FW updates as if the certainty equivalence principle held, i.e., as if $\hat{\boldsymbol{\mu}}(t) = \boldsymbol{\mu}$.

Step 2. (Convergence of the smoothed FW updates) We study the convergence assuming that $(\mathcal{E}_t^{(1)} \cap \mathcal{E}_t^{(2)})$ holds. We first show that $\bar{F}_{\boldsymbol{\mu}, \eta_t}$ is ℓ-Lipschitz and smooth for $\ell = 2D^2 \|\boldsymbol{\mu}\|_\infty^2$, see Appendices H and I. Then, in Appendix E, we establish that the dynamics of $\phi_t = \max_{\boldsymbol{\omega} \in \Sigma} F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) - F_{\boldsymbol{\mu}}(\hat{\boldsymbol{\omega}}(t))$ satisfy $t\phi_t \leq (t-1)\phi_{t-1} + \ell\left(\eta_{t-1} + \frac{K^2}{2t\eta_t}\right)$. Observe that, as mentioned earlier, $1/\sqrt{t}$ is indeed the optimal scaling choice for η_t. We deduce that after a certain finite number T_1 of rounds, ϕ_t is sufficiently small and $\max\{\triangle_{\min}(\hat{\boldsymbol{\mu}}(t))^{-1}, \|\hat{\boldsymbol{\mu}}(t)\|_\infty\} \leq \sqrt{t}$.

Step 3. Finally, we observe that $\mathbb{E}_{\boldsymbol{\mu}}[\tau] \leq T_1 + \sum_{t=T_1}^{\infty} \mathbb{P}_{\boldsymbol{\mu}}\left[t\hat{F}_t \leq (1+\epsilon_t)\beta\left(t, (1 - \frac{1}{4|\mathcal{X}_0|})\delta\right)\right] + \sum_{t=T_1+1}^{\infty} \mathbb{P}_{\boldsymbol{\mu}}\left[(\mathcal{E}_t^{(1)} \cap \mathcal{E}_t^{(2)})^c\right]$, and show that the second term in the r.h.s. in this inequality is equivalent to $T^\star(\boldsymbol{\mu}) \ln(1/\delta)$ as $\delta \to 0$ using the property of the function β defining the stopping threshold and similar arguments as those used in [GK16, WTP21].

(iii) Expected number of `LM` *Oracle calls.* The `MCP` algorithm is called to compute $\hat{F}_t$ and to perform the FW update only in rounds t such that $\max\{\triangle_{\min}(\hat{\boldsymbol{\mu}}(t))^{-1}, \|\hat{\boldsymbol{\mu}}(t)\|_\infty\} \leq \sqrt{t}$. Thus, from Theorem 3 and Lemma 1, the number of `LM` Oracle calls per-round is a polynomial in t and K. As the $\mathbb{E}_{\boldsymbol{\mu}}[\tau]$ is polynomial (in $\ln \delta^{-1}$, K, $\|\boldsymbol{\mu}\|_\infty$ and $\triangle_{\min}^{-1}$), the expected number of `LM` Oracle calls is also polynomial in the same variables.

5 Related Work

We provide an exhaustive survey of the related literature in Appendix B. To summarize, to the best of our knowledge, CombGame [JMKK21] is the state-of-the-art algorithm for BAI in combinatorial semi-bandits in the high confidence regimes. A complete comparison to `P-FWS` is presented in Appendix B. CombGame was initially introduced in [DKM19] for classical bandit problems. There, the lower-bound problem is casted as a two-player game and the authors propose to use no-regret algorithms for each player to solve it. [JMKK21] adapts the algorithm for combinatorial semi-bandits, and provides a non-asymptotic sample complexity upper bound matching (1) asymptotically. However, the resulting algorithm requires to call an oracle solving the Most-Confusing-Parameter problem as our `MCP` algorithm. The authors of [JMKK21] conjectured the existence of such a computationally efficient oracle, and we establish this result here.

6 Conclusion

In this paper, we have presented `P-FWS`, the first computationally efficient and statistically optimal algorithm for the best arm identification problem in combinatorial semi-bandits. For this problem, we have studied the computational-statistical trade-off through the analysis of the optimization problem leading to instance-specific sample complexity lower bounds. This approach can be extended to study the computational-statistical gap in other learning tasks. Of particular interest are problems with an underlying structure (e.g. linear bandits [DMSV20, JP20], or RL in linear / low rank MDPs [AKKS20]). Most results on these problems are concerned with statistical efficiency, and ignore computational issues.

Acknowledgments and Disclosure of Funding

We thank Aristides Gionis and the anonymous reviewers for their valuable feedback. The research is funded by ERC Advanced Grant REBOUND (834862), the Wallenberg AI, Autonomous Systems and Software Program (WASP), and Digital Futures.

References

[AAS+23] Kenshi Abe, Kaito Ariu, Mitsuki Sakamoto, Kentaro Toyoshima, and Atsushi Iwasaki. Last-iterate convergence with full-and noisy-information feedback in two-player zero-sum games. In *Proc. of AISTATS*, 2023.

[AKKS20] Alekh Agarwal, Sham Kakade, Akshay Krishnamurthy, and Wen Sun. Flambe: Structural complexity and representation learning of low rank mdps. In *Proc. of NeurIPS*, 2020.

[ALLW18] Jacob Abernethy, Kevin A Lai, Kfir Y Levy, and Jun-Kun Wang. Faster rates for convex-concave games. In *Proc. of COLT*, 2018.

[AMP21] Aymen Al Marjani and Alexandre Proutiere. Adaptive sampling for best policy identification in markov decision processes. In *Proc. of ICML*, 2021.

[APFS22] Ioannis Anagnostides, Ioannis Panageas, Gabriele Farina, and Tuomas Sandholm. On last-iterate convergence beyond zero-sum games. In *Proc. of ICML*, 2022.

[BBGS11] André Berger, Vincenzo Bonifaci, Fabrizio Grandoni, and Guido Schäfer. Budgeted matching and budgeted matroid intersection via the gasoline puzzle. *Mathematical Programming*, 2011.

[BGK22] Antoine Barrier, Aurélien Garivier, and Tomáš Kocák. A non-asymptotic approach to best-arm identification for gaussian bandits. In *Proc. of AISTATS*, 2022.

[BLM13] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press, 2013.

[BV04] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[CBL06] Nicolo Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge university press, 2006.

[CBL12] Nicolo Cesa-Bianchi and Gábor Lugosi. Combinatorial bandits. *Journal of Computer and System Sciences*, 2012.

[CCG21a] Thibaut Cuvelier, Richard Combes, and Eric Gourdin. Asymptotically optimal strategies for combinatorial semi-bandits in polynomial time. In *Proc. of ALT*, 2021.

[CCG21b] Thibaut Cuvelier, Richard Combes, and Eric Gourdin. Statistically efficient, polynomial-time algorithms for combinatorial semi-bandits. *Proc. of SIGMETRICS*, 2021.

[CGL16] Lijie Chen, Anupam Gupta, and Jian Li. Pure exploration of multi-armed bandit under matroid constraints. In *Proc. of COLT*, 2016.

[CGL+17] Lijie Chen, Anupam Gupta, Jian Li, Mingda Qiao, and Ruosong Wang. Nearly optimal sampling algorithms for combinatorial pure exploration. In *Proc. of COLT*, 2017.

[CLK+14] Shouyuan Chen, Tian Lin, Irwin King, Michael R Lyu, and Wei Chen. Combinatorial pure exploration of multi-armed bandits. In *Proc. of NeurIPS*, 2014.

[CMP17] Richard Combes, Stefan Magureanu, and Alexandre Proutiere. Minimal exploration in structured stochastic bandits. In *Proc. of NeurIPS*, 2017.

[CTMSP+15] Richard Combes, Mohammad Sadegh Talebi Mazraeh Shahi, Alexandre Proutiere, et al. Combinatorial bandits revisited. In *Proc. of NeurIPS*, 2015.

[DBW12] John C Duchi, Peter L Bartlett, and Martin J Wainwright. Randomized smoothing for stochastic optimization. *SIAM Journal on Optimization*, 2012.

[DFG21] Constantinos Daskalakis, Maxwell Fishelson, and Noah Golowich. Near-optimal no-regret learning in general games. In *Proc. of NeurIPS*, 2021.

[DKC21] Yihan Du, Yuko Kuroki, and Wei Chen. Combinatorial pure exploration with full-bandit or partial linear feedback. In *Proc. of AAAI*, 2021.

[DKM19] Rémy Degenne, Wouter M Koolen, and Pierre Ménard. Non-asymptotic pure exploration by solving games. In *Proc. of NeurIPS*, 2019.

[DMSV20] Rémy Degenne, Pierre Ménard, Xuedong Shang, and Michal Valko. Gamification of pure exploration for linear bandits. In *Proc. of ICML*, 2020.

[DP19] Constantinos Daskalakis and Ioannis Panageas. Last-iterate convergence: Zero-sum games and constrained min-max optimization. *Proc. of ITCS*, 2019.

[FKM05] Abraham D Flaxman, Adam Tauman Kalai, and H Brendan McMahan. Online convex optimization in the bandit setting: gradient descent without a gradient. In *Proc. of SODA*, 2005.

[FKV14] Eugene A Feinberg, Pavlo O Kasyanov, and Mark Voorneveld. Berge's maximum theorem for noncompact image sets. *Journal of Mathematical Analysis and Applications*, 2014.

[GC11] A. Garivier and O. Cappé. The KL-UCB algorithm for bounded stochastic bandits and beyond. In *Proc. of COLT*, 2011.

[GH16] Dan Garber and Elad Hazan. A linearly convergent variant of the conditional gradient algorithm under strong convexity, with applications to online and stochastic optimization. *SIAM Journal on Optimization*, 2016.

[GK16] Aurélien Garivier and Emilie Kaufmann. Optimal best arm identification with fixed confidence. In *Proc. of COLT*, 2016.

[GPDO20] Noah Golowich, Sarath Pattathil, Constantinos Daskalakis, and Asuman Ozdaglar. Last iterate is slower than averaged iterate in smooth convex-concave saddle point problems. In *Proc. of COLT*, 2020.

[H+16] Elad Hazan et al. Introduction to online convex optimization. *Foundations and Trends® in Optimization*, 2016.

[Han57] James Hannan. Approximation to bayes risk in repeated play. *Contributions to the Theory of Games*, 1957.

[HK12] Elad Hazan and Satyen Kale. Projection-free online learning. In *Proc. of ICML*, 2012.

[Jag13] Martin Jaggi. Revisiting frank-wolfe: Projection-free sparse convex optimization. In *Proc. of ICML*, 2013.

[JMKK21] Marc Jourdan, Mojmír Mutnỳ, Johannes Kirschner, and Andreas Krause. Efficient pure exploration for combinatorial bandits with semi-bandit feedback. In *Proc. of ALT*, 2021.

[JP20] Yassir Jedra and Alexandre Proutiere. Optimal best-arm identification in linear bandits. In *Proc. of NeurIPS*, 2020.

[KCG16] Emilie Kaufmann, Olivier Cappé, and Aurélien Garivier. On the complexity of best-arm identification in multi-armed bandit models. *JMLR*, 2016.

[KK21] Emilie Kaufmann and Wouter M Koolen. Mixture martingales revisited with applications to sequential tests and confidence intervals. *JMLR*, 2021.

[KLLM22] Daniel Kane, Sihan Liu, Shachar Lovett, and Gaurav Mahajan. Computational-statistical gap in reinforcement learning. In *Proc. of COLT*, 2022.

[KSJJ+20] Julian Katz-Samuels, Lalit Jain, Kevin G Jamieson, et al. An empirical process approach to the union bound: Practical algorithms for combinatorial and linear bandits. In *Proc. of NeurIPS*, 2020.

[KTAS12] Shivaram Kalyanakrishnan, Ambuj Tewari, Peter Auer, and Peter Stone. Pac subset selection in stochastic multi-armed bandits. In *Proc. of ICML*, 2012.

[KV05] Adam Kalai and Santosh Vempala. Efficient algorithms for online decision problems. *Journal of Computer and System Sciences*, 2005.

[KWA+14] Branislav Kveton, Zheng Wen, Azin Ashkan, Hoda Eydgahi, and Brian Eriksson. Matroid bandits: fast combinatorial optimization with learning. In *Proc. of UAI*, 2014.

[Lai87] Tze Leung Lai. Adaptive treatment allocation and the multi-armed bandit problem. *The annals of statistics*, 1987.

[Law72] Eugene L Lawler. A procedure for computing the k best solutions to discrete optimization problems and its application to the shortest path problem. *Management science*, 1972.

[LNP+21] Qi Lei, Sai Ganesh Nagarajan, Ioannis Panageas, et al. Last iterate convergence in no-regret learning: constrained min-max optimization for convex-concave landscapes. In *Proc. of AISTATS*, 2021.

[MCP14] Stefan Magureanu, Richard Combes, and Alexandre Proutiere. Lipschitz bandits: Regret lower bounds and optimal algorithms. In *Proc. of COLT*, 2014.

[Neu15] Gergely Neu. First-order regret bounds for combinatorial semi-bandits. In *Proc. of COLT*, 2015.

[Oka73] Masashi Okamoto. Distinctness of the eigenvalues of a quadratic form in a multivariate sample. *The Annals of Statistics*, 1973.

[PBVP20] Pierre Perrault, Etienne Boursier, Michal Valko, and Vianney Perchet. Statistical efficiency of thompson sampling for combinatorial semi-bandits. In *Proc. of NeurIPS*, 2020.

[Per22] Pierre Perrault. When combinatorial thompson sampling meets approximation regret. In *Proc. of NeurIPS*, 2022.

[PPV19] Pierre Perrault, Vianney Perchet, and Michal Valko. Exploiting structure of uncertainty for efficient matroid semi-bandits. In *Proc. of ICML*, 2019.

[RG96] Ram Ravi and Michel X Goemans. The constrained minimum spanning tree problem. In *Scandinavian Workshop on Algorithm Theory*. Springer, 1996.

[RS13] Sasha Rakhlin and Karthik Sridharan. Optimization, learning, and games with predictable sequences. In *Proc. of NeurIPS*, 2013.

[S$^+$03] Alexander Schrijver et al. *Combinatorial optimization: polyhedra and efficiency*, volume 24. Springer, 2003.

[SN20] Arun Suggala and Praneeth Netrapalli. Follow the perturbed leader: Optimism and fast parallel algorithms for smooth minimax games. In *Proc. of NeurIPS*, 2020.

[Vis21] Nisheeth K. Vishnoi. *Algorithms for Convex Optimization*. Cambridge University Press, 2021.

[WLZL21] Chen-Yu Wei, Chung-Wei Lee, Mengxiao Zhang, and Haipeng Luo. Linear last-iterate convergence in constrained saddle-point optimization. In *Prof. of ICLR*, 2021.

[WTP21] Po-An Wang, Ruo-Chun Tzeng, and Alexandre Proutiere. Fast pure exploration via frank-wolfe. In *Proc. of NeurIPS*, 2021.

[ZODS21] Tom Zahavy, Brendan O'Donoghue, Guillaume Desjardins, and Satinder Singh. Reward is enough for convex mdps. In *Proc. of NeurIPS*, 2021.

Contents

A Table of Notation

Problem setting	
K	Number of arms
$[m]$ for any $m \in \mathbb{N}$	The set $\{1, 2 \ldots, m\}$
δ	Required uncertainty
$\boldsymbol{\mu} \in \mathbb{R}^K$	Vector of the expected rewards of the various arms
$\mathbb{E}_{\boldsymbol{\mu}}$ and $\mathbb{P}_{\boldsymbol{\mu}}$	The expectation and probability measure corresponding to $\boldsymbol{\mu}$
Λ	$\{\boldsymbol{\mu} \in \mathbb{R}^K : \lvert \boldsymbol{i}^\star(\boldsymbol{\mu}) \rvert = 1\}$
$\boldsymbol{i}^\star(\boldsymbol{\mu})$	Best arm under parameter $\boldsymbol{\mu}$
$\mathcal{X}$	Set of actions in $\{0, 1\}^K$
D	$\max_{\boldsymbol{x} \in \mathcal{X}} \lVert \boldsymbol{x} \rVert_1$
$\triangle_{\boldsymbol{x}}(\boldsymbol{\mu})$	$\langle \boldsymbol{i}^\star(\boldsymbol{\mu}) - \boldsymbol{x}, \boldsymbol{\mu} \rangle$
$\triangle_{\min}(\boldsymbol{\mu})$	$\min_{\boldsymbol{x} \neq \boldsymbol{i}^\star(\boldsymbol{\mu})} \triangle_{\boldsymbol{x}}(\boldsymbol{\mu})$
Notation related to a given algorithm	
$N_k(t)$	Number of pulls of arm k up to time t
$\hat{\omega}_k(t)$	$N_k(t)/t$
$\boldsymbol{x}(t)$	The action taken in time t
$y_k(t)$	Random reward received if $x_k(t) = 1$
$\hat{\mu}_k(t)$	$\sum_{s=1}^t y_k(s) \mathbb{1}\{x_k(s) = 1\}/N_k(t)$
τ	Stopping time
$\hat{\boldsymbol{\imath}}$	Recommended action
Notation used for sets and vectors	
$\odot$	Elementwise product
$\oplus$	Elementwise sum over $\mathbb{Z}_2$
$\boldsymbol{x}^i$	The i-th elementwise power of $\boldsymbol{x} \in \mathbb{R}^K$, i.e., $(x_k^i)_{k \in [K]}$
$\mathsf{cl}\,(\mathcal{S})$	The closure of set $\mathcal{S}$
$\boldsymbol{e}_k$	the K-dimensional vector whose k-th component is equal to one and zero elsewhere
Properties for lower bound	
$d(\mu, \mu')$	KL divergence between the distributions parametrized by μ and μ'
$\mathsf{kl}(a, b)$	KL divergence between two Bernoulli distributions of means a and b
$\mathsf{Alt}(\boldsymbol{\mu})$	$\{\boldsymbol{\lambda} \in \Lambda : \boldsymbol{i}^\star(\boldsymbol{\lambda}) \neq \boldsymbol{i}^\star(\boldsymbol{\mu})\}$
Σ	$\{\sum_{\boldsymbol{x} \in \mathcal{X}} w_{\boldsymbol{x}} \boldsymbol{x} : \boldsymbol{w} \in \Sigma_{\lvert\mathcal{X}\rvert}\}$ where Σ_N is a $(N-1)$-dimensional simplex
Σ_+	$\Sigma \cap \mathbb{R}^K_{>0}$
Notation for `MCP`	
$F_{\boldsymbol{\mu}}(\boldsymbol{\omega})$	$\min_{\boldsymbol{x} \neq \boldsymbol{i}^\star(\boldsymbol{\mu})} f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{\mu})$
$f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{\mu})$	$\inf_{\boldsymbol{\lambda} \in \mathcal{C}_{\boldsymbol{x}}} \langle \boldsymbol{\omega}, \frac{(\boldsymbol{\mu} - \boldsymbol{\lambda})^2}{2} \rangle$, where $\mathcal{C}_{\boldsymbol{x}} = \{\boldsymbol{\lambda} \in \mathbb{R}^K : \langle \boldsymbol{\lambda}, \boldsymbol{i}^\star(\boldsymbol{\mu}) - \boldsymbol{x} \rangle < 0\}$
$\mathcal{L}_{\boldsymbol{\omega}, \boldsymbol{\mu}}(\boldsymbol{\lambda}, \boldsymbol{x}, \alpha)$	$\langle \boldsymbol{\omega}, \frac{(\boldsymbol{\mu} - \boldsymbol{\lambda})^2}{2} \rangle + \alpha \langle \boldsymbol{i}^\star(\boldsymbol{\mu}) - \boldsymbol{x}, \boldsymbol{\lambda} \rangle$
$g_{\boldsymbol{\omega}, \boldsymbol{\mu}}(\boldsymbol{x}, \alpha)$	$\inf_{\boldsymbol{\lambda} \in \mathbb{R}^K} \mathcal{L}_{\boldsymbol{\omega}, \boldsymbol{\mu}}(\boldsymbol{\lambda}, \boldsymbol{x}, \alpha)$
Notation for `P-FWS`	
$\mathcal{X}_0$	A $[K]$-covering set
$\hat{F}_t$	`MCP`-approximated value of $F_{\hat{\boldsymbol{\mu}}(t)}(\hat{\boldsymbol{\omega}}(t))$ for stopping rule
$\bar{F}_{\boldsymbol{\mu}, \eta}(\cdot)$	$\mathbb{E}_{\boldsymbol{\mathcal{Z}} \sim \text{Uniform}(B_2)}[\nabla F_{\boldsymbol{\mu}}(\cdot + \eta \boldsymbol{\mathcal{Z}})]$ where $B_2 = \{\boldsymbol{v} \in \mathbb{R}^K : \lVert \boldsymbol{v} \rVert_2 \leq 1\}$
$\bar{F}_{\boldsymbol{\mu}, \eta, n}$	The empirical n-sample estimate of $\bar{F}_{\boldsymbol{\mu}, \eta}$
ℓ	Lipschitz constant of $F_{\boldsymbol{\mu}}$

B Further related work

Combinatorial semi-bandits [CBL12] have found numerous applications including online ranking [DKC21], network routing [CLK+14, KWA+14], loan assignment [KWA+14], path planning problem [JMKK21], and influence marketing [Per22]). We do not discuss these applications here, but rather focus the literature that is the most relevant to our analysis and results.

Solving the lower-bound problem in combinatorial semi-bandits. We are not aware of any computationally efficient algorithm to solve the lower-bound problem, or to compute its objective function. To the best of our knowledge, `MCP` is the first algorithm to do so. A work closed to ours is [CCG21a] for combinatorial semi-bandits but in the regret minimization. Regret minimization yields a different lower-bound problem. There exits a statistically optimal algorithm [CMP17], called `OSSB`, that matches the regret lower bound by [CTMSP+15]. `OSSB` requires to solve the lower-bound problem in each round, and the authors [CCG21a] are the first to investigate whether this is at all possible in a computationally efficient way. They establish that if budgeted-linear maximization (BLM) [RG96, BBGS11] can be solved within an ε-approximation factor for the combinatorial set $\mathcal{X}$, then the lower-bound problem can be approximately solved with a precision depending on ε. As a consequence, the approach leads to an algorithm with asymptotically minimal regret only if one has access to an exact BLM solver. This is the case for m-sets and s-t paths but this is not the case for spanning trees and perfect matchings. For the latter case, as mentioned [CCG21a], an algorithm using an approximately correct BLM solver would not be statistically optimal.

Best arm identification in combinatorial semi-bandits. Many tasks related to combinatorial best arm identification are formulated in the *transductive* setting [JMKK21], where the set $\mathcal{A} \subseteq \{0,1\}^K$ available for exploration is not necessarily the same as the set $\mathcal{X} \subseteq \{0,1\}^K$ for decision. The minimal sample complexity in the transductive setting is exactly (1) with Σ replaced with $\{\sum_{\boldsymbol{x}\in\mathcal{A}} w_{\boldsymbol{x}}\boldsymbol{x} : \boldsymbol{\omega} \in \Sigma_{|\mathcal{A}|}\}$ - see (58) in Appendix L for details. Two most studied tasks are combinatorial multi-arm bandit (C-MB) where $\mathcal{A} = \{\boldsymbol{e}_k\}_{k\in[K]}$ and the best action identification (C-BAI) where $\mathcal{A} = \mathcal{X}$. The former is arguably simpler than the latter if we compare the corresponding minimal sample complexities (note that $\Sigma_K \supseteq \Sigma$). We note that our results for C-BAI can be easily generalized to the transductive setting (see Appendix L).

Prior works mainly focus on the C-MB task. UCB-based [KTAS12, CLK+14] and elimination-based [CGL16, CGL+17, KSJJ+20] approaches are popular. Among these, `EfficientGapElim` [CGL+17] achieves the lowest sample complexity $\mathcal{O}(T^\star(\boldsymbol{\mu})(\ln\delta^{-1} + \ln^2 \triangle_{\min}^{-1}(\ln\ln\triangle_{\min}^{-1} + \ln|\mathcal{X}|))$ with high probability[4], but its computational complexity is hard to analyze. `Peace` [KSJJ+20], another elimination-based approach by experimental design, requires with high probability a polynomial number of the `LM` Oracle calls in total. The sample complexity of `Peace` has a δ-dependent term (scaling as $KT^\star(\boldsymbol{\mu})\ln\delta^{-1}$) worse than `EfficientGapElim`. Overall, none of these are statistically optimal when $\delta \to 0$. Note that algorithms for linear best-arm identification [DMSV20, WTP21] are applicable to C-MB but not to C-BAI and the general transductive setting.

For the task of C-BAI, we are only aware of two works: `GCB-PE` [DKC21] and `CombGame` [JMKK21]. `GCB-PE` is a UCB-based algorithm with guarantees on the sample complexity and computational complexity valid with high probability only. `CombGame` [JMKK21] is proposed for the transductive setting, and its design inherits from [DKM19] that interprets the lower-bound problem and more precisely $T^\star(\boldsymbol{\mu})^{-1}$ as the value of a two-player game (a $\boldsymbol{\omega}$-player and a $\boldsymbol{\lambda}$-player)[5]. Assuming that an `MCP` oracle is available, `CombGame` leverages Frank-Wolfe algorithms, namely OFW [HK12] and LLOO [GH16], for the $\boldsymbol{\omega}$-player and the `MCP` algorithm for the $\boldsymbol{\lambda}$-player. [JMKK21] leaves the existence of such an oracle running in polynomial time as an open problem. Our `MCP` algorithm resolves this issue. `CombGame` is statistically optimal in the high confidence regime but has no clear guarantees in the moderate regime [BGK22].

We wish to finally mention an algorithm that has inspired the design of `P-FWS`. This algorithm is referred to as Frank-Wolfe Sampling (`FWS`) [WTP21]. `FWS` is optimal in high confidence regime

[4] In Section 4.5 in [CGL+17], the authors provide a lemma stating that: if *parallel simulation* is additionally allowed, then any high-probability sample complexity upper bound can be converted to an upper bound in expectation.

[5] Note that this two-player game is different than the two-player game involved in our algorithm `MCP`.

but is not computationally efficient for combinatorial semi-bandits. For example, to deal with the non-smoothness issue of the objective function $F_{\boldsymbol{\mu}}$, FWS needs to construct the so-called r-subdifferentiable spaces and to optimize a linear function on these spaces. Unfortunately, these spaces can be generated by a number of vectors exponentially increasing with K in combinatorial semi-bandits. Moreover, in moderate confidence regime, the sample complexity upper bound derived in [WTP21] has an exponential dependence in K.

All the relevant algorithms, their sample complexity guarantees and computational complexity are summarized in Table 1.

Table 1: Algorithms for best-arm identification in combinatorial semi-bandits with fixed confidence and their performance.

Algorithm	Task	Instance-specific Sample Complexity		Computational Complexity	
		Non-asympt.	Asympt. Opt.	Needed (Provided)	Total LM oracle calls
Peace	C-MB	poly $\left(K, \triangle_{\min}^{-1}, \ln \delta^{-1}\right)$ w.h.p.	✗	LP solver (✓)	poly $\left(K, \triangle_{\min}^{-1}, \delta^{-1}\right)$ w.h.p.
GCB-PE	C-BAI	poly $\left(K, \triangle_{\min}^{-1}, \ln \delta^{-1}\right)$ w.h.p.	✗	-	poly $\left(K, \triangle_{\min}^{-1}, \ln \delta^{-1}\right)$ w.h.p.
CombGame	Trans.	✗ (incomparable)	✓	MCP (✗)	✗
P-FWS	Trans.	poly $\left(K, \triangle_{\min}^{-1}, \ln \delta^{-1}\right)$	✓	MCP (✓)	poly $\left(K, \triangle_{\min}^{-1}, \ln \delta^{-1}\right)$

C Results related to our (ϵ, θ)-MCP algorithm

C.1 Properties of Lagrangian dual of $f_{\boldsymbol{x}}$

Proposition 1. *Let $(\boldsymbol{\omega}, \boldsymbol{\mu}) \in \Sigma_+ \times \Lambda$ and $\boldsymbol{x} \in \mathcal{X} \setminus \{\boldsymbol{i}^\star(\boldsymbol{\mu})\}$.*
(a) The Lagrange dual function is linear in $\boldsymbol{x}$. More precisely, $g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha) = c_{\boldsymbol{\omega},\boldsymbol{\mu}}(\alpha) + \langle \boldsymbol{\ell}_{\boldsymbol{\omega},\boldsymbol{\mu}}(\alpha), \boldsymbol{x}\rangle$ where $c_{\boldsymbol{\omega},\boldsymbol{\mu}}(\alpha) = \alpha \left\langle \boldsymbol{\mu} - \frac{\alpha}{2}\boldsymbol{\omega}^{-1}, \boldsymbol{i}^\star(\boldsymbol{\mu})\right\rangle$ and $\boldsymbol{\ell}_{\boldsymbol{\omega},\boldsymbol{\mu}}(\alpha) = -\alpha\left(\boldsymbol{\mu} + \frac{\alpha}{2}\boldsymbol{\omega}^{-1} \odot (\mathbf{1}_K - 2\boldsymbol{i}^\star(\boldsymbol{\mu}))\right)$.
(b) $g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \cdot)$ is strictly concave (for any fixed $\boldsymbol{x}$).
(c) $f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{\mu}) = \max_{\alpha \geq 0} g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha)$ is attained by $\alpha^\star_{\boldsymbol{x}} = \frac{\triangle_{\boldsymbol{x}}(\boldsymbol{\mu})}{\langle \boldsymbol{x} \oplus \boldsymbol{i}^\star(\boldsymbol{\mu}), \boldsymbol{\omega}^{-1}\rangle}$.
(d) $\|\boldsymbol{\ell}_{\boldsymbol{\omega},\boldsymbol{\mu}}(\alpha^\star_{\boldsymbol{x}})\|_1 \leq L_{\boldsymbol{\omega},\boldsymbol{\mu}} = 4D^2K\|\boldsymbol{\mu}\|_\infty^2 \left\|\boldsymbol{\omega}^{-1}\right\|_\infty$.

Proof Fix any $(\boldsymbol{\omega}, \boldsymbol{\mu}) \in \Sigma_+ \times \Lambda$ and let $\boldsymbol{i}^\star = \boldsymbol{i}^\star(\boldsymbol{\mu})$ for short. For convenience, the definition of $\mathcal{L}_{\boldsymbol{\omega},\boldsymbol{\mu}}$ and $g_{\boldsymbol{\omega},\boldsymbol{\mu}}$ are restated:

$$\mathcal{L}_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{\lambda}, \boldsymbol{x}, \alpha) = \left\langle \boldsymbol{\omega}, \frac{(\boldsymbol{\mu} - \boldsymbol{\lambda})^2}{2}\right\rangle + \alpha\langle \boldsymbol{i}^\star - \boldsymbol{x}, \boldsymbol{\lambda}\rangle \quad \text{and} \quad g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha) = \inf_{\boldsymbol{\lambda} \in \mathbb{R}^K} \mathcal{L}_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{\lambda}, \boldsymbol{x}, \alpha).$$

<u>Proof of (a): linearity of $g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\cdot, \alpha)$</u>: Let $\boldsymbol{\lambda}^\star_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha) \in \arg\inf_{\boldsymbol{\lambda} \in \mathbb{R}^K} \mathcal{L}_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{\lambda}, \boldsymbol{x}, \alpha)$. The first-order condition implies that $\mathbf{0}_K = \nabla_{\boldsymbol{\lambda}} \mathcal{L}_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{\lambda}^\star_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha), \boldsymbol{x}, \alpha) = \boldsymbol{\omega} \odot (\boldsymbol{\lambda}^\star_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha) - \boldsymbol{\mu}) + \alpha(\boldsymbol{i}^\star - \boldsymbol{x})$, which directly yields (as $\boldsymbol{\omega} > \mathbf{0}_K$)

$$\boldsymbol{\lambda}^\star_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha) = \boldsymbol{\mu} + \alpha\boldsymbol{\omega}^{-1} \odot (\boldsymbol{x} - \boldsymbol{i}^\star). \tag{11}$$

We plug (11) into $\mathcal{L}_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{\lambda}^\star_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha), \boldsymbol{x}, \alpha)$ and directly obtain that

$$\begin{aligned} g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha) &= \left\langle \boldsymbol{\omega}, \frac{\alpha^2}{2}\boldsymbol{\omega}^{-2} \odot (\boldsymbol{x} - \boldsymbol{i}^\star)^2 \right\rangle + \alpha\langle \boldsymbol{\mu}, \boldsymbol{i}^\star - \boldsymbol{x}\rangle - \alpha^2\left\langle \boldsymbol{\omega}^{-1}, (\boldsymbol{x} - \boldsymbol{i}^\star)^2\right\rangle \\ &= \alpha\langle \boldsymbol{\mu}, \boldsymbol{i}^\star - \boldsymbol{x}\rangle - \frac{\alpha^2}{2}\left\langle \boldsymbol{\omega}^{-1}, (\boldsymbol{x} - \boldsymbol{i}^\star)^2\right\rangle \qquad (12) \\ &= c_{\boldsymbol{\omega},\boldsymbol{\mu}}(\alpha) + \langle \boldsymbol{\ell}_{\boldsymbol{\omega},\boldsymbol{\mu}}(\alpha), \boldsymbol{x}\rangle, \qquad (13) \end{aligned}$$

where (13) follows from a fact that $(\boldsymbol{x} - \boldsymbol{i}^\star)^2 = \boldsymbol{i}^\star - 2\boldsymbol{x} \odot \boldsymbol{i}^\star + \boldsymbol{x} = \boldsymbol{i}^\star + \boldsymbol{x} \odot (\mathbf{1}_K - 2\boldsymbol{i}^\star)$.

<u>Proof of (b): strict concavity of $g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \cdot)$</u>: This is trivial from (12).

<u>Proof of (c): $f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{\mu}) = \max_{\alpha \geq 0} g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha)$ is attained by $\alpha^\star_{\boldsymbol{x}} = \frac{\triangle_{\boldsymbol{x}}(\boldsymbol{\mu})}{\langle \boldsymbol{x} \oplus \boldsymbol{i}^\star, \boldsymbol{\omega}^{-1}\rangle}$</u> : For a fixed $\boldsymbol{x} \neq \boldsymbol{i}^\star$, by the first-order condition of (12), we find that the maximum of $g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \cdot)$ is reached at

$$\alpha^\star_{\boldsymbol{x}} = \frac{\triangle_{\boldsymbol{x}}(\boldsymbol{\mu})}{\langle \boldsymbol{\omega}^{-1}, (\boldsymbol{x} - \boldsymbol{i}^\star)^2\rangle} = \frac{\triangle_{\boldsymbol{x}}(\boldsymbol{\mu})}{\langle \boldsymbol{x} \oplus \boldsymbol{i}^\star, \boldsymbol{\omega}^{-1}\rangle}, \tag{14}$$

where for the second equality, we use the assumption that $\boldsymbol{i}^\star$ and $\boldsymbol{x}$ are binary vectors and hence $(\boldsymbol{x} - \boldsymbol{i}^\star)^2 = \boldsymbol{x} \oplus \boldsymbol{i}^\star$. We now verify that $(\alpha^\star_{\boldsymbol{x}}, \boldsymbol{\lambda}^\star_{\boldsymbol{x}})$, where $\boldsymbol{\lambda}^\star_{\boldsymbol{x}} = \boldsymbol{\lambda}^\star_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha^\star_{\boldsymbol{x}})$ (see (11)), satisfies KKT conditions, which is equivalent to strong duality (refer to [Vis21, BV04]) under Slater's condition (there exists a $\boldsymbol{\lambda} \in \mathbb{R}^K$ such that the constraint is strict). Since $\boldsymbol{x}$ is a suboptimal action, $\triangle_{\boldsymbol{x}}(\boldsymbol{\mu})$ is positive, so is $\alpha^\star_{\boldsymbol{x}}$ (dual feasibility). To verify $\langle \boldsymbol{\lambda}^\star_{\boldsymbol{x}}, \boldsymbol{i}^\star - \boldsymbol{x}\rangle \leq 0$ (primal feasibility), the definition of $\boldsymbol{\lambda}^\star_{\boldsymbol{\omega},\boldsymbol{\mu}}(\cdot, \cdot)$, (11), yields

$$\begin{aligned} \langle \boldsymbol{\lambda}^\star_{\boldsymbol{x}}, \boldsymbol{i}^\star - \boldsymbol{x}\rangle &= \triangle_{\boldsymbol{x}}(\boldsymbol{\mu}) + \alpha^\star_{\boldsymbol{x}}\left\langle \boldsymbol{\omega}^{-1} \odot (\boldsymbol{x} - \boldsymbol{i}^\star), (\boldsymbol{i}^\star - \boldsymbol{x})\right\rangle \\ &= \triangle_{\boldsymbol{x}}(\boldsymbol{\mu}) - \alpha^\star_{\boldsymbol{x}}\langle \boldsymbol{\omega}^{-1}, \boldsymbol{x} \oplus \boldsymbol{i}^\star\rangle = 0, \end{aligned}$$

which implies that $\alpha^\star_{\boldsymbol{x}}\langle \boldsymbol{i}^\star - \boldsymbol{x}, \boldsymbol{\lambda}^\star_{\boldsymbol{x}}\rangle = 0$ (complementary slackness). Finally, stationarity holds automatically as $\nabla_{\boldsymbol{\lambda}} \mathcal{L}_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{\lambda}^\star_{\boldsymbol{x}}, \boldsymbol{x}, \alpha) = 0$ for all α.

<u>Proof of (d): $\|\boldsymbol{\ell}_{\boldsymbol{\omega},\boldsymbol{\mu}}(\alpha^\star_{\boldsymbol{x}})\|_1 \leq L_{\boldsymbol{\omega},\boldsymbol{\mu}} = 4D^2K\|\boldsymbol{\mu}\|_\infty^2 \left\|\boldsymbol{\omega}^{-1}\right\|_\infty$</u>: Following from the expression of $\boldsymbol{\ell}_{\boldsymbol{\omega},\boldsymbol{\mu}}(\alpha)$, we have $\boldsymbol{\ell}_{\boldsymbol{\omega},\boldsymbol{\mu}}(\alpha^\star_{\boldsymbol{x}}) = -\alpha^\star_{\boldsymbol{x}}\boldsymbol{\mu} + \frac{\alpha^{\star 2}_{\boldsymbol{x}}}{2}\boldsymbol{\omega}^{-1} \odot (\mathbf{1}_K - 2\boldsymbol{i}^\star)$. Observe that $\|\boldsymbol{\mu}\|_1 \leq K\|\boldsymbol{\mu}\|_\infty \leq$

$K \left\|\boldsymbol{\omega}^{-1}\right\|_\infty \|\boldsymbol{\mu}\|_\infty$ (as $\boldsymbol{\omega} \in \Sigma_+$) and the coordinate of $\mathbf{1}_K - 2\boldsymbol{i}^\star$ is either 1 or -1, a simple application of triangle inequality leads to

$$\|\boldsymbol{\ell}_{\boldsymbol{\omega},\boldsymbol{\mu}}(\alpha^\star_{\boldsymbol{x}})\|_1 \leq K \left\|\boldsymbol{\omega}^{-1}\right\|_\infty \left(\|\boldsymbol{\mu}\|_\infty + \frac{\alpha^\star_{\boldsymbol{x}}}{2}\right) \alpha^\star_{\boldsymbol{x}}.$$

As for $\alpha^\star_{\boldsymbol{x}}$ (see (14)), $\triangle_{\boldsymbol{x}}(\boldsymbol{\mu}) \leq 2D \|\boldsymbol{\mu}\|_\infty$ and $\left\langle \boldsymbol{\omega}^{-1}, \boldsymbol{x} \oplus \boldsymbol{i}^\star \right\rangle \geq \min_k \omega_k^{-1} \geq 1$, hence we conclude that $\|\boldsymbol{\ell}_{\boldsymbol{\omega},\boldsymbol{\mu}}(\alpha^\star_{\boldsymbol{x}})\|_1 \leq 2D(D+1)K \|\boldsymbol{\mu}\|^2_\infty \left\|\boldsymbol{\omega}^{-1}\right\|_\infty \leq L_{\boldsymbol{\omega},\boldsymbol{\mu}}$. □

C.2 Analysis of MCP

Theorem 3. *Let $\epsilon, \theta \in (0,1)$. Under Assumption 1, for any $(\boldsymbol{\omega}, \boldsymbol{\mu}) \in \Sigma_+ \times \Lambda$, the (ϵ,θ)-MCP$(\boldsymbol{\omega}, \boldsymbol{\mu})$ algorithm outputs $(\hat{F}, \hat{\boldsymbol{x}})$ satisfying*

$$\mathbb{P}\left[F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) \leq \hat{F} \leq (1+\epsilon) F_{\boldsymbol{\mu}}(\boldsymbol{\omega})\right] \geq 1 - \theta \quad \textit{and} \quad \hat{F} = \max_{\alpha \geq 0} g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\hat{\boldsymbol{x}}, \alpha).$$

Moreover, the number of LM Oracle calls the algorithm does is almost surely at most

$$\left\lceil \frac{c_\theta^2 (1+\epsilon)^2}{\epsilon^2 F_{\boldsymbol{\mu}}(\boldsymbol{\omega})^2} \right\rceil = \mathcal{O}\left(\frac{\|\boldsymbol{\mu}\|^4_\infty \left\|\boldsymbol{\omega}^{-1}\right\|^2_\infty K^3 D^5 \ln K \ln \theta^{-1}}{\epsilon^2 F_{\boldsymbol{\mu}}(\boldsymbol{\omega})^2} \right).$$

Proof Fix any $(\boldsymbol{\omega}, \boldsymbol{\mu}) \in \Sigma_+ \times \Lambda$ and denote by $\boldsymbol{i}^\star = \boldsymbol{i}^\star(\boldsymbol{\mu})$. Suppose Algorithm 1 reaches the stopping criterion at the N-th iteration.

Guarantees on the outputs of MCP: By Proposition 1 (a),

$$\sum_{n=1}^N g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}^{(n)}, \alpha^{(n)}) - \min_{\boldsymbol{x} \neq \boldsymbol{i}^\star} \sum_{n=1}^N g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha^{(n)}) = \sum_{n=1}^N \left\langle \boldsymbol{\ell}_{\boldsymbol{\omega},\boldsymbol{\mu}}(\alpha^{(n)}), \boldsymbol{x}^{(n)} \right\rangle - \min_{\boldsymbol{x} \neq \boldsymbol{i}^\star} \sum_{n=1}^N \left\langle \boldsymbol{\ell}_{\boldsymbol{\omega},\boldsymbol{\mu}}(\alpha^{(n)}), \boldsymbol{x} \right\rangle.$$

The regret of $\boldsymbol{x}$-player can be bounded by applying Lemma 3, resulting in:

$$\mathbb{P}\left[\sum_{n=1}^N g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}^{(n)}, \alpha^{(n)}) - \min_{\boldsymbol{x} \neq \boldsymbol{i}^\star} \sum_{n=1}^N g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha^{(n)}) \leq c_\theta \sqrt{N} \right] \geq 1 - \theta. \tag{15}$$

To relate $F_{\boldsymbol{\mu}}(\boldsymbol{\omega})$ with (15), let $\boldsymbol{x}_e$ be the minimizer attaining $F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) = f_{\boldsymbol{x}_e}(\boldsymbol{\omega}, \boldsymbol{\mu})$. Then,

$$\min_{\boldsymbol{x} \neq \boldsymbol{i}^\star} \sum_{n=1}^N g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha^{(n)}) \leq \sum_{n=1}^N g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}_e, \alpha^{(n)}) \leq N \max_{\alpha \geq 0} g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}_e, \alpha) = N F_{\boldsymbol{\mu}}(\boldsymbol{\omega}). \tag{16}$$

Recall that $\alpha^{(n)}$ is chosen as the best response $\max_{\alpha \geq 0} g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}^{(n)}, \alpha) = g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}^{(n)}, \alpha^{(n)})$ and that $\hat{F} = \min_{n \in [N]} g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}^{(n)}, \alpha^{(n)})$. These together with (16) imply that

$$N(\hat{F} - F_{\boldsymbol{\mu}}(\boldsymbol{\omega})) \leq \sum_{n=1}^N g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}^{(n)}, \alpha^{(n)}) - \min_{\boldsymbol{x} \neq \boldsymbol{i}^\star} \sum_{n=1}^N g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}, \alpha^{(n)}). \tag{17}$$

A simple rearrangement on (15) and (17) implies that: with probability at least $1 - \theta$,

$$\hat{F} - F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) \leq \frac{c_\theta}{\sqrt{N}} \leq \epsilon(\hat{F} - \frac{c_\theta}{\sqrt{N}}) \leq \epsilon F_{\boldsymbol{\mu}}(\boldsymbol{\omega}),$$

where the second inequality follows from the stopping criterion that $\sqrt{N} > c_\theta(1+\epsilon)/\epsilon \hat{F}$, and the last inequality simply comes from the rearrangement of the first inequality.

Computational cost: From the stopping criterion of MCP, we know that

$$N = \left\lceil \frac{c_\theta^2 (1+\epsilon)^2}{\epsilon^2 \hat{F}^2} \right\rceil \leq \left\lceil \frac{c_\theta^2 (1+\epsilon^{-1})^2}{F_{\boldsymbol{\mu}}(\boldsymbol{\omega})^2} \right\rceil = \mathcal{O}\left(\frac{L^2_{\boldsymbol{\omega},\boldsymbol{\mu}} \left(\sqrt{K \ln K} + \sqrt{\ln \theta^{-1}}\right)^2}{\epsilon^2 F_{\boldsymbol{\mu}}(\boldsymbol{\omega})^2} \right)$$

since $\hat{F} \geq F_{\boldsymbol{\mu}}(\boldsymbol{\omega})$ and $c_\theta = L_{\boldsymbol{\omega},\boldsymbol{\mu}}\left(4\sqrt{K(\ln K+1)} + \sqrt{\ln(\theta^{-1})/2}\right)$. Finally, as computing each $\boldsymbol{x}^{(n)}$ takes at most D calls to `LM` Oracle, the total number of `LM` Oracle calls is

$$\mathcal{O}\left(\frac{L_{\boldsymbol{\omega},\boldsymbol{\mu}}^2 D\left(\sqrt{K\ln K} + \sqrt{\ln\theta^{-1}}\right)^2}{\epsilon^2 F_{\boldsymbol{\mu}}(\boldsymbol{\omega})^2}\right) = \mathcal{O}\left(\frac{\|\boldsymbol{\mu}\|_\infty^4 \left\|\boldsymbol{\omega}^{-1}\right\|_\infty^2 K^3 D^5 \ln K \ln\theta^{-1}}{\epsilon^2 F_{\boldsymbol{\mu}}(\boldsymbol{\omega})^2}\right)$$

by recalling $L_{\boldsymbol{\omega},\boldsymbol{\mu}} = 4D^2K\|\boldsymbol{\mu}\|_\infty^2 \left\|\boldsymbol{\omega}^{-1}\right\|_\infty$ from Proposition 1 (d) and $(\sqrt{\ln K} + \sqrt{\ln\theta^{-1}})^2 = \mathcal{O}(\ln K \ln\theta^{-1})$. □

C.3 Regret analysis of Follow-the-Perturbed-Leader

In this subsection, we aim at proving Lemma 3, which is a direct consequence of Lemma 4. One can find similar proofs in e.g. [KV05, Neu15, SN20]. However, the parameter η_n in our `MCP` algorithm is varying and carefully chosen (without the knowledge of the last round), which makes the proof slightly more complicated.

Lemma 3. *Let $N \in \mathbb{N}$. Under (ϵ,θ)-`MCP`$(\boldsymbol{\omega},\boldsymbol{\mu})$, then*

$$\mathbb{P}\left[\frac{1}{N}\sum_{n=1}^N g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}^{(n)},\alpha^{(n)}) - \frac{1}{N}\min_{\boldsymbol{x}\neq\boldsymbol{i}^\star}\sum_{n=1}^N g_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x},\alpha^{(n)}) \leq \frac{c_\theta}{\sqrt{N}}\right] \geq 1-\theta.$$

Lemma 4. *Let $\theta \in (0,1)$ and $\mathcal{M} \subseteq \{0,1\}^K$. Given an arbitrary sequence $\{\boldsymbol{\ell}_n\}_{n\geq 1}$ of vectors in $\mathbb{R}^K$ whose length $\|\boldsymbol{\ell}_n\|_1$ is bounded by $L>0$ for all $n \in \mathbb{N}$. Suppose $\{\boldsymbol{x}^{(n)}\}_{n\geq 1}$ is generated by*

$$\boldsymbol{x}^{(n)} \in \operatorname*{argmin}_{\boldsymbol{x}\in\mathcal{M}}\left(\sum_{m=1}^{n-1}\langle\boldsymbol{\ell}_m,\boldsymbol{x}\rangle + \left\langle\frac{\boldsymbol{\mathcal{Z}}_n}{\eta_n},\boldsymbol{x}\right\rangle\right),$$

where $\boldsymbol{\mathcal{Z}}_n = (\mathcal{Z}_{1,n},\cdots,\mathcal{Z}_{K,n})$ is a random vector with uncorrelated exponentially distributed (with unit mean) components, and $\eta_n = \sqrt{K(\ln K+1)/(4nL^2)}$. Then, for any $N \in \mathbb{N}$,

$$\mathbb{P}\left[\sum_{n=1}^N\left\langle\boldsymbol{\ell}_n,\boldsymbol{x}^{(n)}\right\rangle - \min_{\boldsymbol{x}\in\mathcal{M}}\sum_{n=1}^N\langle\boldsymbol{\ell}_n,\boldsymbol{x}\rangle \leq L\sqrt{N}\left(4\sqrt{K(\ln K+1)} + \sqrt{\frac{\ln\theta^{-1}}{2}}\right)\right] \geq 1-\theta.$$

Proof of Lemma 4: We will prove this lemma as if $\{\boldsymbol{\ell}_n\}_n$ is chosen in advance since there exists a standard technique for extending regret against oblivious player to the one against nonoblivious one (see Lemma 4.1 in [CBL06]). For convenience, we introduce the following notation. Let $\boldsymbol{m}^\star(\cdot) = \operatorname{argmin}_{\boldsymbol{x}\in\mathcal{M}}\langle\cdot,\boldsymbol{x}\rangle$. Finally, further define global minimizer $\boldsymbol{x}_\star = \boldsymbol{m}^\star\left(\sum_{n=1}^N\boldsymbol{\ell}_n\right)$ and an auxiliary vector $\boldsymbol{b}^{(n)} = \boldsymbol{m}^\star\left(\sum_{m=1}^n\boldsymbol{\ell}_m + \boldsymbol{\mathcal{Z}}_1/\eta_n\right)$.

It suffices to show the expected regret bound (18).

$$\mathbb{E}\left[\sum_{n=1}^N\left\langle\boldsymbol{\ell}_n,\boldsymbol{x}^{(n)}\right\rangle\right] - \min_{\boldsymbol{x}\in\mathcal{M}}\sum_{n=1}^N\langle\boldsymbol{\ell}_n,\boldsymbol{x}\rangle \leq 4L\sqrt{NK(\ln K+1)}. \tag{18}$$

This is because $\{\langle\boldsymbol{\ell}_n,\boldsymbol{x}^{(n)}\rangle - \mathbb{E}\left[\langle\boldsymbol{\ell}_n,\boldsymbol{x}^{(n)}\rangle\right]\}_n$ forms a sequence of bounded martingale difference, so an application of a concentration inequality (Lemma 6) with $V_n = \langle\boldsymbol{\ell}_n,\boldsymbol{x}^{(n)}\rangle - \mathbb{E}\left[\langle\boldsymbol{\ell}_n,\boldsymbol{x}^{(n)}\rangle\right]$, $r_n = L$ for $n \in [N]$, and $s = L\sqrt{N\ln\theta^{-1}/2}$ gives that

$$\mathbb{P}\left[\sum_{n=1}^N\left\langle\boldsymbol{\ell}_n,\boldsymbol{x}^{(n)}\right\rangle - \mathbb{E}\left[\sum_{n=1}^N\left\langle\boldsymbol{\ell}_n,\boldsymbol{x}^{(n)}\right\rangle\right] > L\sqrt{\frac{N\ln\theta^{-1}}{2}}\right] \leq \theta$$

and combining this with (18) completes the proof.

Proof of (18): We decompose the regret into two terms:

$$\sum_{n=1}^{N}\left\langle \boldsymbol{\ell}_n, \boldsymbol{x}^{(n)}-\boldsymbol{x}_\star\right\rangle = \sum_{n=1}^{N}\left\langle \boldsymbol{b}^{(n)}-\boldsymbol{x}_\star, \boldsymbol{\ell}_n\right\rangle + \sum_{n=1}^{N}\left\langle \boldsymbol{x}^{(n)}-\boldsymbol{b}^{(n)}, \boldsymbol{\ell}_n\right\rangle .$$

(i). We show that $\mathbb{E}\left[\sum_{n=1}^{N}\left\langle \boldsymbol{b}^{(n)}-\boldsymbol{x}_\star,\boldsymbol{\ell}_n\right\rangle\right] \le \frac{K(\ln K+1)}{\eta_N}$**.** Invoking Lemma 5 with $\boldsymbol{x}=\boldsymbol{x}_\star$ results in

$$\begin{aligned}
\mathbb{E}\left[\sum_{n=1}^{N}\left\langle \boldsymbol{b}^{(n)}-\boldsymbol{x}_\star,\boldsymbol{\ell}_n\right\rangle\right] &\le \mathbb{E}\left[\left\langle \frac{\boldsymbol{x}_\star}{\eta_N}-\left(\frac{\boldsymbol{b}^{(1)}}{\eta_1}+\sum_{n=2}^{N}\left(\frac{1}{\eta_n}-\frac{1}{\eta_{n-1}}\right)\boldsymbol{b}^{(n)}\right),\boldsymbol{\mathcal{Z}}_1\right\rangle\right]\\
&\le \mathbb{E}\left[\|\boldsymbol{\mathcal{Z}}_1\|_\infty\left\|\frac{\boldsymbol{x}_\star}{\eta_N}-\left(\frac{\boldsymbol{b}^{(1)}}{\eta_1}+\sum_{n=2}^{N}\left(\frac{1}{\eta_n}-\frac{1}{\eta_{n-1}}\right)\boldsymbol{b}^{(n)}\right)\right\|_1\right],
\end{aligned}$$

where the last inequality uses Hölder's inequality. As all the components of $\boldsymbol{x}_\star$ and $\frac{\eta_N\boldsymbol{b}^{(1)}}{\eta_1}+\sum_{n=2}^{N}\eta_N\left(\frac{1}{\eta_n}-\frac{1}{\eta_{n-1}}\right)\boldsymbol{b}^{(n)}$ are nonnegative and bounded by 1, the 1-norm of their difference is bounded by K. It remains to show

$$\mathbb{E}[\|\boldsymbol{\mathcal{Z}}_1\|_\infty] = \int_0^\infty \mathbb{P}\left[\max_i \mathcal{Z}_{1,i}\ge x\right]dx \le \int_0^{\ln K}\mathbb{P}\left[\max_i \mathcal{Z}_{1,i}\ge x\right]dx+\int_{\ln K}^\infty Ke^{-x}dx \le \ln K+1.$$

(ii). We show that $\mathbb{E}\left[\sum_{n=1}^{N}\left\langle \boldsymbol{x}^{(n)}-\boldsymbol{b}^{(n)},\boldsymbol{\ell}_n\right\rangle\right] \le 2L^2\sum_{n=1}^{N}\eta_n$**.** Let the pdf of $\exp(1)$ be $\pi(\cdot)=e^{-\|\cdot\|_1}$.

$$\begin{aligned}
\mathbb{E}\left[\left\langle \boldsymbol{b}^{(n)},\boldsymbol{\ell}_n\right\rangle\right] &= \int_{\boldsymbol{z}\in\mathbb{R}^K}\left\langle \boldsymbol{m}^\star\left(\eta_n\sum_{m=1}^{n}\boldsymbol{\ell}_m+\boldsymbol{z}\right),\boldsymbol{\ell}_n\right\rangle d\pi(\boldsymbol{z})\\
&= \int_{\boldsymbol{y}\in\mathbb{R}^K}\left\langle \boldsymbol{m}^\star\left(\eta_n\sum_{m=1}^{n-1}\boldsymbol{\ell}_m+\boldsymbol{y}\right),\boldsymbol{\ell}_n\right\rangle d\pi(\boldsymbol{y}-\eta_n\boldsymbol{\ell}_n)\\
&= \int_{\boldsymbol{y}\in\mathbb{R}^K}\left\langle \boldsymbol{m}^\star\left(\eta_n\sum_{m=1}^{n-1}\boldsymbol{\ell}_m+\boldsymbol{y}\right),\boldsymbol{\ell}_n\right\rangle e^{-\|\boldsymbol{y}-\eta_n\boldsymbol{\ell}_n\|_1+\|\boldsymbol{y}\|_1}d\pi(\boldsymbol{y}).
\end{aligned}$$

Notice that the triangular inequality implies $-\|\boldsymbol{y}-\eta_n\boldsymbol{\ell}_n\|_1+\|\boldsymbol{y}\|_1\le\|\eta_n\boldsymbol{\ell}_n\|_1\le\eta_n L$ and $e^x\le 1+2x$ for all $x\in(0,1)$ (Taylor expansion), so recalling $\boldsymbol{x}^{(n)}=\boldsymbol{m}^\star\left(\eta_n\sum_{m=1}^{n-1}\boldsymbol{\ell}_m+\boldsymbol{\mathcal{Z}}_1\right)$, we deduce that

$$\begin{aligned}
\sum_{n=1}^{N}\mathbb{E}\left[\left\langle \boldsymbol{x}^{(n)}-\boldsymbol{b}^{(n)},\boldsymbol{\ell}_n\right\rangle\right] &\le \sum_{n=1}^{N}2\eta_n L\int_{\boldsymbol{z}\in\mathbb{R}^K}\left\langle \boldsymbol{m}^\star\left(\eta_n\sum_{m=1}^{n}\boldsymbol{\ell}_m+\boldsymbol{z}\right),\boldsymbol{\ell}_n\right\rangle d\pi(\boldsymbol{z})\\
&\le \sum_{n=1}^{N}2\eta_n L\int_{\boldsymbol{z}\in\mathbb{R}^K}\left\|\boldsymbol{m}^\star\left(\eta_n\sum_{m=1}^{n}\boldsymbol{\ell}_m+\boldsymbol{z}\right)\right\|_\infty\|\boldsymbol{\ell}_n\|_1\,d\pi(\boldsymbol{z})\le 2L^2\sum_{n=1}^{N}\eta_n.
\end{aligned}$$

Finally, plugging $\eta_n=\sqrt{\frac{K(\ln K+1)}{4nL^2}}$ into (i). and (ii). directly concludes the proof. □

The following lemma is a result that can be found in [CBL06, H+16], we rewrite it here for completeness.

Lemma 5. *According to* $\boldsymbol{b}^{(n)}=\boldsymbol{m}^\star\left(\eta_n\sum_{m=1}^{n}\boldsymbol{\ell}_m+\boldsymbol{\mathcal{Z}}_1\right)$, *we can have*

$$\forall\boldsymbol{x}\in\mathcal{M},\quad \sum_{n=1}^{N}\left\langle \boldsymbol{b}^{(n)}-\boldsymbol{x},\boldsymbol{\ell}_n\right\rangle \le \left\langle \frac{\boldsymbol{x}}{\eta_N},\boldsymbol{\mathcal{Z}}_1\right\rangle-\left\langle \frac{\boldsymbol{b}^{(1)}}{\eta_1}+\sum_{n=2}^{N}\left(\frac{1}{\eta_n}-\frac{1}{\eta_{n-1}}\right)\boldsymbol{b}^{(n)},\boldsymbol{\mathcal{Z}}_1\right\rangle. \quad (19)$$

Proof This is done by induction. For the base case, $N = 1$, as $\boldsymbol{b}^{(1)} = \boldsymbol{m}^\star\left(\boldsymbol{\ell}_1 + \boldsymbol{\mathcal{Z}}_1/\eta_1\right)$

$$\left\langle \boldsymbol{b}^{(1)}, \boldsymbol{\ell}_1 + \boldsymbol{\mathcal{Z}}_1/\eta_1 \right\rangle \le \left\langle \boldsymbol{x}, \boldsymbol{\ell}_1 + \boldsymbol{\mathcal{Z}}_1/\eta_1 \right\rangle$$

for any $\boldsymbol{x} \in \mathcal{M}$. A simple rearrangement yields (19). While considering $N+1$, we suppose (19) holds for all integers smaller than $N+1$. For an arbitrary $\boldsymbol{x} \in \mathcal{M}$, the fact $\boldsymbol{b}^{(N+1)} = \boldsymbol{m}^\star\left(\sum_{n=1}^{N+1} \boldsymbol{\ell}_n + \boldsymbol{\mathcal{Z}}_1/\eta_{N+1}\right)$ directly implies that

$$\begin{aligned}
\left\langle \boldsymbol{x}, \sum_{n=1}^{N+1} \boldsymbol{\ell}_n + \frac{\boldsymbol{\mathcal{Z}}_1}{\eta_{N+1}} \right\rangle \ge \left\langle \boldsymbol{b}^{(N+1)}, \sum_{n=1}^{N+1} \boldsymbol{\ell}_n + \frac{\boldsymbol{\mathcal{Z}}_1}{\eta_{N+1}} \right\rangle& \\
= \left\langle \boldsymbol{b}^{(N+1)}, \boldsymbol{\ell}_{N+1} + \left(\frac{1}{\eta_{N+1}} - \frac{1}{\eta_N}\right)\boldsymbol{\mathcal{Z}}_1 \right\rangle + \left\langle \boldsymbol{b}^{(N+1)}, \sum_{n=1}^{N} \boldsymbol{\ell}_n + \frac{\boldsymbol{\mathcal{Z}}_1}{\eta_N} \right\rangle& \\
\ge \sum_{n=1}^{N+1} \left\langle \boldsymbol{b}^{(n)}, \boldsymbol{\ell}_n \right\rangle + \left\langle \frac{\boldsymbol{b}^{(1)}}{\eta_1} + \sum_{n=2}^{N+1}\left(\frac{1}{\eta_n} - \frac{1}{\eta_{n-1}}\right)\boldsymbol{b}^{(n)}, \boldsymbol{\mathcal{Z}}_1 \right\rangle&,
\end{aligned}$$

where the last inequality comes from applying the hypothesis (19) with $\boldsymbol{x} = \boldsymbol{b}^{(N+1)}$ on the second inner product. Rearrange the above inequality, our induction is completed. □

Lemma 6 (Hoeffding-Azuma). *Let $N \in \mathbb{N}$, $V_1, V_2, \cdots, V_N$ be a bounded martingale difference sequence w.r.t. $X_1, X_2, \cdots, X_N$ such that for any $n \in [N]$ $V_n \in [A_n, A_n + r_n]$ for some random variable A_n, measurable w.r.t. $X_1, \cdots, X_{n-1}$ and a positive constant r_n. Then, for any $s > 0$,*

$$\mathbb{P}\left[\sum_{n\in[N]} V_n > s\right] \le \exp\left(-\frac{2s^2}{\sum_{n\in[N]} r_n^2}\right) \quad \textit{and} \quad \mathbb{P}\left[\sum_{n\in[N]} V_n < -s\right] \le \exp\left(-\frac{2s^2}{\sum_{n\in[N]} c_n^2}\right).$$

D Analysis of P-FWS

In this appendix, we prove our main theorem.

Theorem 4. *Let $\boldsymbol{\mu} \in \Lambda$ and $\delta \in (0,1)$. If* P-FWS *is parametrized using*

$$(\epsilon_t, \eta_t, n_t, \rho_t, \theta_t) = \left(t^{-\frac{1}{9}},\, \frac{1}{4\sqrt{t|\mathcal{X}_0|}},\, \left\lceil t^{\frac{1}{4}} \right\rceil,\, \frac{1}{16tD^2|\mathcal{X}_0|},\, \frac{1}{t^{\frac{1}{4}} e^{\sqrt{t}}}\right), \tag{10}$$

then (i) the algorithm finishes in finite time almost surely and $\mathbb{P}_{\boldsymbol{\mu}}[\hat{\boldsymbol{\imath}} \neq \boldsymbol{i}^\star(\boldsymbol{\mu})] \leq \delta$; (ii) its sample complexity satisfies $\mathbb{P}_{\boldsymbol{\mu}}\big[\limsup_{\delta\to 0} \frac{\tau}{\ln \delta^{-1}} \leq T^\star(\boldsymbol{\mu})\big] = 1$ and for any $\epsilon, \tilde{\epsilon} \in (0,1)$ with $\epsilon < \min\{1, \frac{2D^2\triangle_{\min}^2}{K}, \frac{D^2\|\boldsymbol{\mu}\|_\infty^2}{3}\}$,

$$\mathbb{E}_{\boldsymbol{\mu}}[\tau] \leq \frac{(1+\tilde{\epsilon})^2}{T^\star(\boldsymbol{\mu})^{-1} - 6\epsilon} \times H\left(\frac{1}{\delta} \cdot \frac{4c_2}{3} \cdot \frac{(1+\tilde{\epsilon})^2}{T^\star(\boldsymbol{\mu})^{-1} - 6\epsilon}\right) + \Psi(\epsilon, \tilde{\epsilon}),$$

where $H(x) = \ln x + \ln\ln x + 1$ and $\Psi(\epsilon, \tilde{\epsilon})$ (refer to (34) *for a detailed expression) is polynomial in ϵ^{-1}, $\tilde{\epsilon}^{-1}$, K, $\|\boldsymbol{\mu}\|_\infty$ and $\triangle_{\min}(\boldsymbol{\mu})^{-1}$; (iii) the expected number of* LM *Oracle calls is upper bounded by a polynomial in $\ln \delta^{-1}$, K, $\|\boldsymbol{\mu}\|_\infty$ and $\triangle_{\min}(\boldsymbol{\mu})^{-1}$.*

D.1 δ-correctness (Theorem 4 (i))

Recall that P-FWS stopping rule is:

$$\tau = \inf\left\{ t > 4|\mathcal{X}_0| : \frac{t\hat{F}_t}{1+\epsilon_t} > \beta\left(t, \left(1 - \frac{1}{4|\mathcal{X}_0|}\right)\delta\right), \max\left\{\frac{1}{\triangle_{\min}(\hat{\boldsymbol{\mu}}(t))}, \|\hat{\boldsymbol{\mu}}(t)\|_\infty\right\} \leq \sqrt{t} \right\}, \tag{7}$$

where $\hat{F}_t$ is computed by $(\epsilon_t, \delta/t^2)$-MCP$(\hat{\boldsymbol{\omega}}(t), \hat{\boldsymbol{\mu}}(t))$. Let $\hat{\boldsymbol{\imath}} = \boldsymbol{i}^\star(\hat{\boldsymbol{\mu}}(\tau))$ be the output of P-FWS. Define the good event $\mathcal{G} = \bigcap_{t=4|\mathcal{X}_0|+1}^{\infty} \{\hat{F}_t \leq (1+\epsilon_t) F_{\hat{\boldsymbol{\mu}}(t)}(\hat{\boldsymbol{\omega}}(t))\}$. Hence, it follows from the guarantee of $(\epsilon_t, \delta/t^2)$-MCP algorithm that

$$\mathbb{P}_{\boldsymbol{\mu}}[\mathcal{G}^c] \leq \delta \sum_{t=4|\mathcal{X}_0|+1}^{\infty} t^{-2} \leq \delta \int_{4|\mathcal{X}_0|}^{\infty} x^{-2} dx \leq \frac{\delta}{4|\mathcal{X}_0|}.$$

Besides, under the event $\mathcal{G}$,

$$(1+\epsilon_\tau)\tau F_{\hat{\boldsymbol{\mu}}(\tau)}(\hat{\boldsymbol{\omega}}(\tau)) \geq \tau \hat{F}_\tau \geq (1+\epsilon_\tau)\beta\left(\tau, \left(1 - \frac{1}{4|\mathcal{X}_0|}\right)\delta\right)$$

holds, implying that $\tau F_{\hat{\boldsymbol{\mu}}(\tau)}(\hat{\boldsymbol{\omega}}(\tau)) \geq \beta\left(\tau, \left(1 - \frac{1}{4|\mathcal{X}_0|}\right)\delta\right)$. So, by (8)-(9), $\hat{\boldsymbol{\imath}} = \boldsymbol{i}^\star(\hat{\boldsymbol{\mu}}(t))$ satisfies:

$$\mathbb{P}_{\boldsymbol{\mu}}[\hat{\boldsymbol{\imath}} \neq \boldsymbol{i}^\star(\boldsymbol{\mu}), \mathcal{G}] \leq \left(1 - \frac{1}{4|\mathcal{X}_0|}\right)\delta,$$

and thus $\mathbb{P}_{\boldsymbol{\mu}}[\hat{\boldsymbol{\imath}} \neq \boldsymbol{i}^\star(\boldsymbol{\mu})] \leq \mathbb{P}_{\boldsymbol{\mu}}[\hat{\boldsymbol{\imath}} \neq \boldsymbol{i}^\star(\boldsymbol{\mu}), \mathcal{G}] + \mathbb{P}_{\boldsymbol{\mu}}[\mathcal{G}^c] \leq \delta$.

D.2 Almost-sure upper bound (Theorem 4 (ii))

In this section, we show Theorem 4 (ii) an almost-sure upper bound on the sample complexity for P-FWS. Our proof is based on the continuity of $F_{\boldsymbol{\mu}}$ in $\boldsymbol{\mu}$ (as in [GK16, WTP21]) and also on the following observations:

(a) $\{\hat{\boldsymbol{\mu}}(t) \stackrel{t\to\infty}{\longrightarrow} \boldsymbol{\mu}\}$ and $\{\nabla \tilde{F}_{\boldsymbol{\mu},\eta_t,n_t}(\boldsymbol{\omega}) \stackrel{t\to\infty}{\longrightarrow} \nabla \bar{F}_{\boldsymbol{\mu},\eta_t}(\boldsymbol{\omega}), \forall \boldsymbol{\omega} \in \Sigma_+\}$ happen almost surely,

(b) $\hat{F}_t \geq F_{\hat{\boldsymbol{\mu}}(t)}(\hat{\boldsymbol{\omega}}(t))$.

For (a), by the law of large numbers, $\hat{\boldsymbol{\mu}}(t) \stackrel{t\to\infty}{\longrightarrow} \boldsymbol{\mu}$ as $N_k(t) \stackrel{t\to\infty}{\longrightarrow} \infty$ for all $k \in [K]$ yielded by forced exploration rounds involved in P-FWS (Lemma 14 in Appendix F), $\nabla \tilde{F}_{\boldsymbol{\mu},\eta_t,n_t}(\boldsymbol{\omega}) \stackrel{t\to\infty}{\longrightarrow}$

$\nabla \bar{F}_{\boldsymbol{\mu},\eta_t}(\boldsymbol{\omega})$, $\forall \boldsymbol{\omega} \in \Sigma_+$ is a direct consequence that $n_t \overset{t\to\infty}{\longrightarrow} \infty$. (b) is immediately derived from the definition of $\hat{F}_t$ as $\hat{F}_t = f_{\hat{\boldsymbol{x}}}(\hat{\boldsymbol{\omega}}(t), \hat{\boldsymbol{\mu}}(t))$ for some action $\hat{\boldsymbol{x}} \neq \boldsymbol{i}^\star(\hat{\boldsymbol{\mu}}(t))$ and $F_{\hat{\boldsymbol{\mu}}(t)}(\hat{\boldsymbol{\omega}}(t)) = \min_{\boldsymbol{x}\in\mathcal{X}\setminus \boldsymbol{i}^\star(\hat{\boldsymbol{\mu}}(t))} f_{\boldsymbol{x}}(\hat{\boldsymbol{\omega}}(t), \hat{\boldsymbol{\mu}}(t))$.

Introduce the event

$$\mathcal{E} = \left\{ F_{\boldsymbol{\mu}}(\hat{\boldsymbol{\omega}}(t)) \overset{t\to\infty}{\longrightarrow} \max_{\boldsymbol{\omega}\in\Sigma} F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) \text{ and } \hat{\boldsymbol{\mu}}(t) \overset{t\to\infty}{\longrightarrow} \boldsymbol{\mu} \right\}.$$

Because of (a), Theorem 5 in Appendix E ensures that $\mathbb{P}_{\boldsymbol{\mu}}[\mathcal{E}] = 1$. Also, by the uniform continuity of $F_{\boldsymbol{\mu}}(\boldsymbol{\omega})$ in $\boldsymbol{\mu}$ for an arbitrary $\boldsymbol{\omega} \in \Sigma_+$ (Lemma 7 in D.3.3),

$$\max_{\boldsymbol{\omega}\in\Sigma_+} \left| F_{\hat{\boldsymbol{\mu}}(t)}(\boldsymbol{\omega}) - F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) \right| \overset{t\to\infty}{\longrightarrow} 0$$

almost surely, and hence by the triangle inequality, this implies that

$$\mathbb{P}_{\boldsymbol{\mu}}\left[F_{\hat{\boldsymbol{\mu}}(t)}(\hat{\boldsymbol{\omega}}(t)) \overset{t\to\infty}{\longrightarrow} \max_{\boldsymbol{\omega}\in\Sigma} F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) \right] = 1.$$

For any $\epsilon \in (0,1)$, under $\mathcal{E}$, there exists a positive integer $T_\epsilon > \max\{c_1, 4|\mathcal{X}_0|\}$ such that for any $t \geq T_\epsilon$, we have

$$F_{\hat{\boldsymbol{\mu}}(t)}(\hat{\boldsymbol{\omega}}(t)) \geq (1-\epsilon) \max_{\boldsymbol{\omega}\in\Sigma} F_{\boldsymbol{\mu}}(\boldsymbol{\omega}),\ \max\left\{ \frac{1}{\triangle_{\min}(\hat{\boldsymbol{\mu}}(t))}, \|\hat{\boldsymbol{\mu}}(t)\|_\infty \right\} \leq \sqrt{t},\ \text{and } \epsilon_t \leq \epsilon, \tag{20}$$

where the second inequality is due to (a) and the third is because $\epsilon_t \to 0$. So, the stopping time (7) can be upper bounded by

$$\begin{aligned}
\tau &\leq T_\epsilon + \inf\left\{ t > T_\epsilon : t\hat{F}_t > (1+\epsilon)\,\beta\left(t, \frac{(4|\mathcal{X}_0|-1)\delta}{4|\mathcal{X}_0|}\right) \right\} \\
&\leq T_\epsilon + \inf\left\{ t > T_\epsilon : tF_{\hat{\boldsymbol{\mu}}(t)}(\hat{\boldsymbol{\omega}}(t)) > (1+\epsilon)\,\beta\left(t, \frac{(4|\mathcal{X}_0|-1)\delta}{4|\mathcal{X}_0|}\right) \right\} \\
&\leq T_\epsilon + \inf\left\{ t > T_\epsilon : t(1-\epsilon)\max_{\boldsymbol{\omega}\in\Sigma} F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) > (1+\epsilon)\,\beta\left(t, \frac{(4|\mathcal{X}_0|-1)\delta}{4|\mathcal{X}_0|}\right) \right\} \\
&\leq T_\epsilon + \inf\left\{ t > T_\epsilon : \frac{(1-\epsilon)t}{(1+\epsilon)T^\star(\boldsymbol{\mu})} > \ln\left(\frac{c_2 t}{\delta} \cdot \frac{4|\mathcal{X}_0|}{4|\mathcal{X}_0|-1} \right) \right\} \\
&\leq 2T_\epsilon + \left(\frac{1+\epsilon}{1-\epsilon}\right) T^\star(\boldsymbol{\mu}) H\left(\frac{1}{\delta} \cdot \frac{8c_2}{7} \left(\frac{1+\epsilon}{1-\epsilon}\right) T^\star(\boldsymbol{\mu}) \right).
\end{aligned} \tag{21}$$

where the first inequality uses the last two inequalities of (20), the second inequality uses (b), the third inequality is based on the first inequality of (20), the fourth uses $T^\star(\boldsymbol{\mu})^{-1} = \max_{\boldsymbol{\omega}\in\Sigma} F_{\boldsymbol{\mu}}(\boldsymbol{\omega})$ and (9), and the last inequality results from $(4|\mathcal{X}_0|)/(4|\mathcal{X}_0|-1) \leq 8/7$ (as $|\mathcal{X}_0| \geq 2$), and an application of Lemma 9 with

$$\alpha = 1, \quad b_1 = \frac{1-\epsilon}{1+\epsilon} \cdot \frac{1}{T^\star(\boldsymbol{\mu})} \quad \text{and} \quad b_2 = \frac{8c_2}{7} \cdot \frac{1}{\delta}.$$

Finally, as $\epsilon \in (0,1)$ can be arbitrarily small, (21) implies that

$$\mathbb{P}_{\boldsymbol{\mu}}\left[\limsup_{\delta\to 0} \frac{\tau}{\ln \delta^{-1}} \leq T^\star(\boldsymbol{\mu}) \right] = 1.$$

D.3 Non-asymptotic sample complexity (Theorem 4 (ii))

We establish the following non-asymptotic upper bound on $\mathbb{E}_{\boldsymbol{\mu}}[\tau]$: for any $\tilde{\epsilon}, \epsilon \in (0,1)$ small enough,

$$\mathbb{E}_{\boldsymbol{\mu}}[\tau] \leq \frac{(1+\tilde{\epsilon})^2}{T^\star(\boldsymbol{\mu})^{-1} - 6\epsilon} H\left(\frac{1}{\delta} \cdot \frac{8c_2}{7} \cdot \frac{(1+\tilde{\epsilon})^2}{T^\star(\boldsymbol{\mu})^{-1} - 6\epsilon} \right) + \Psi(\epsilon, \tilde{\epsilon}),$$

where $H(x) = \ln x + \ln\ln x + 1$ and $\Psi(\epsilon, \tilde{\epsilon})$ is defined in (34).

Note that this directly implies the asymptotic optimality. Indeed, when $\delta \to 0$, we get:

$$\limsup_{\delta \to 0} \frac{\mathbb{E}_{\boldsymbol{\mu}}[\tau]}{\ln \delta^{-1}} \le \frac{(1+\tilde{\epsilon})^2}{T^\star(\boldsymbol{\mu})^{-1} - 6\epsilon}.$$

As $\epsilon, \tilde{\epsilon}$ can be set arbitrarily small and $\mathsf{kl}(\delta, 1-\delta) \approx \ln \delta^{-1}$ as $\delta \to 0$, it matches the sample complexity lower bound (1) (Theorem 7 in Appendix K) asymptotically.

Throughout this section, we assume $\boldsymbol{\mu} \in \Lambda$ is given and take any $\epsilon \in (0,1)$ satisfying the following:

$$\epsilon < \min\left\{1, \frac{2D^2\triangle_{\min}^2}{K}, \frac{1}{6T^\star(\boldsymbol{\mu})}\right\} \le \min\left\{1, \frac{2D^2\triangle_{\min}^2}{K}, \frac{D^2\|\boldsymbol{\mu}\|_\infty^2}{3}\right\}, \tag{22}$$

where the second inequality is because $T^\star(\boldsymbol{\mu})^{-1} \le \ell = 2D^2\|\boldsymbol{\mu}\|_\infty^2$ by Lemma 22 in Appendix I. The assumption of $\epsilon < \min\{1, \frac{2D^2\triangle_{\min}^2}{K}, \frac{D^2\|\boldsymbol{\mu}\|_\infty^2}{3}\}$ is used to define the good events introduced in D.3.1 as well as to derive several necessary technical lemmas summarized in D.3.3.

D.3.1 Good events

Since in early rounds, the estimation of $\hat{\boldsymbol{\mu}}(t)$ is noisy, we introduce two threshold functions, $\underline{h}$ and $\overline{h}$, on the round index T:

$$\begin{cases} \underline{h}(T) & = \min\{t \in \mathbb{N} : t \ge T^a, \sqrt{t/|\mathcal{X}_0|} \in \mathbb{N}\} \\ \overline{h}(T) & = \min\{t \in \mathbb{N} : t \ge T^b\underline{h}(T), \sqrt{t/|\mathcal{X}_0|} \in \mathbb{N}\} \end{cases}, \tag{23}$$

where $a, b \in (0,1)$ and $a+b<1$ will be explained later in (27). Now, we define our good events:

$$\mathcal{E}_{1,\epsilon}(T) = \bigcap_{t=\underline{h}(T)}^{T} \mathcal{E}_{1,\epsilon}^{(t)} \quad \text{and} \quad \mathcal{E}_{2,\epsilon}(T) = \bigcap_{t=\underline{h}(T)}^{T} \mathcal{E}_{2,\epsilon}^{(t)}, \tag{24}$$

where $\mathcal{E}_{1,\epsilon}^{(t)} = \left\{\langle \nabla \bar{F}_{\hat{\boldsymbol{\mu}}(t-1),\eta_t}(\hat{\boldsymbol{\omega}}(t-1)), \boldsymbol{x}(t)\rangle \ge \max_{\boldsymbol{x}\in\mathcal{X}} \langle \nabla \bar{F}_{\hat{\boldsymbol{\mu}}(t-1),\eta_t}(\hat{\boldsymbol{\omega}}(t-1)), \boldsymbol{x}\rangle - \epsilon\right\}$ and $\mathcal{E}_{2,\epsilon}^{(t)} = \left\{\|\hat{\boldsymbol{\mu}}(t-1) - \boldsymbol{\mu}\|_\infty < \frac{\epsilon}{24D^3\|\boldsymbol{\mu}\|_\infty}\right\}$.

$\mathcal{E}_{1,\epsilon}^{(t)}$ is the event when the solution of FW update is bounded by at most ϵ, and $\mathcal{E}_{2,\epsilon}^{(t)}$ is the event when the empirical estimate of $\boldsymbol{\mu}$ is sufficiently accurate. Under $\mathcal{E}_{2,\epsilon}^{(t)}$, the uniform continuity shown in Lemma 7 in D.3.3 ensures that:

$$|F_{\hat{\boldsymbol{\mu}}(t-1)}(\boldsymbol{\omega}) - F_{\boldsymbol{\mu}}(\boldsymbol{\omega})| < \epsilon, \quad \forall \boldsymbol{\omega} \in \Sigma_+,$$
$$|\langle \nabla \bar{F}_{\hat{\boldsymbol{\mu}}(t-1),\eta}(\boldsymbol{\omega}) - \nabla \bar{F}_{\boldsymbol{\mu},\eta}(\boldsymbol{\omega}), \boldsymbol{x} - \boldsymbol{\omega}\rangle| < \epsilon, \quad \forall (\boldsymbol{\omega}, \boldsymbol{x}) \in \Sigma_+ \times \mathcal{X}, \forall \eta \in (0, \min_{k\in[K]} \omega_k).$$

The second inequality enables the duality gap of FW algorithm to be controlled, leading to the convergence of `P-FWS`. Let

$$M = \max\left\{(4|\mathcal{X}_0|)^{\frac{1}{a}}, \left(\frac{4K^2}{\epsilon^2 D^2|\mathcal{X}_0|}\right)^{\frac{1}{a}}, \left(\frac{2}{\triangle_{\min}(\boldsymbol{\mu})}\right)^{\frac{2}{a}}, \left(\frac{3\|\boldsymbol{\mu}\|_\infty}{2}\right)^{\frac{2}{a}}\right\} + \max\left\{\left(\frac{\ell}{\epsilon}\right)^{\frac{1}{b}}, \left(\frac{5\ell K^2}{\epsilon\sqrt{|\mathcal{X}_0|}}\right)^{\frac{2}{a+b}}\right\}, \tag{25}$$

then overall, we have (Theorem 5 in D.3.3): for any $t \ge \overline{h}(M)$,

$$\max_{\boldsymbol{\omega}\in\Sigma} F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) - F_{\boldsymbol{\mu}}(\hat{\boldsymbol{\omega}}(t)) \le 5\epsilon,\ \triangle_{\min}(\hat{\boldsymbol{\mu}}(t)) \ge \frac{\triangle_{\min}(\boldsymbol{\mu})}{2},\ \text{and}\ \|\hat{\boldsymbol{\mu}}(t)\|_\infty \le \frac{3\|\boldsymbol{\mu}\|_\infty}{2}. \tag{26}$$

Finally, the values of a, b are set to the following:

$$a = \frac{7}{9} \quad \text{and} \quad b = \frac{1}{9}. \tag{27}$$

This choices will balance the leading order between ϵ^{-1} and $\tilde{\epsilon}^{-1}$ in the δ-independent terms (34) of the non-asymptotic upper bound (which will be shown later).

D.3.2 Proof of non-asymptotic sample complexity

Let $\delta \in (0,1)$. We claim that:

$$\mathbb{E}_{\boldsymbol{\mu}}[\tau] \leq \sum_{T=1}^{\infty} \mathbb{P}_{\boldsymbol{\mu}}[\tau \geq T] \leq T_0(\delta) + \sum_{T=M+1}^{\infty} \mathbb{P}_{\boldsymbol{\mu}}[(\mathcal{E}_{1,\epsilon}(T) \cap \mathcal{E}_{2,\epsilon}(T))^c], \tag{28}$$

where $T_0(\delta) = \inf\left\{T \geq M : \overline{h}(T) + \frac{(1+\epsilon_T)}{T^\star(\boldsymbol{\mu})^{-1}-6\epsilon}\beta\left(T, \frac{(4|\mathcal{X}_0|-1)\delta}{4|\mathcal{X}_0|}\right) \leq T\right\}$. The proof is completed by bounding each term in the right-hand side of (28).

Proof of (28): Suppose $T \geq M$ and $\mathcal{E}_{1,\epsilon}(T) \cap \mathcal{E}_{2,\epsilon}(T)$ holds. Observe that

$$\min\{\tau, T\} \leq \overline{h}(T) + \sum_{t=\lceil \overline{h}(T) \rceil}^{T} \mathbb{1}\{\tau > t\}.$$

To derive an upper bound of $\sum_{t=\lceil \overline{h}(T)\rceil}^{T} \mathbb{1}\{\tau > t\}$, recall the stopping rule (7) that

$$\begin{aligned}
\tau &= \inf\left\{t > 4|\mathcal{X}_0| : \frac{t\hat{F}_t}{1+\epsilon_t} > \beta\left(t, \frac{(4|\mathcal{X}_0|-1)\delta}{4|\mathcal{X}_0|}\right), \max\left\{\frac{1}{\triangle_{\min}(\hat{\boldsymbol{\mu}}(t))}, \|\hat{\boldsymbol{\mu}}(t)\|_\infty\right\} \leq \sqrt{t}\right\} \\
&\leq \inf\left\{t \geq \overline{h}(M) : t\hat{F}_t > (1+\epsilon_t)\,\beta\left(t, \frac{(4|\mathcal{X}_0|-1)\delta}{4|\mathcal{X}_0|}\right)\right\} \\
&\leq \inf\left\{t \geq \overline{h}(M) : t(T^\star(\boldsymbol{\mu})^{-1} - 6\epsilon) > (1+\epsilon_t)\,\beta\left(t, \frac{(4|\mathcal{X}_0|-1)\delta}{4|\mathcal{X}_0|}\right)\right\},
\end{aligned}$$

where the first inequality uses (26), and the second follows from Lemma 7 and Theorem 5 in D.3.3:

$$|F_{\hat{\boldsymbol{\mu}}(t-1)}(\hat{\boldsymbol{\omega}}(t-1)) - F_{\boldsymbol{\mu}}(\hat{\boldsymbol{\omega}}(t-1))| < \epsilon \quad \text{and} \quad T^\star(\boldsymbol{\mu})^{-1} - F_{\boldsymbol{\mu}}(\hat{\boldsymbol{\omega}}(t)) \leq 5\epsilon, \tag{29}$$

and the fact that $\hat{F}_t \geq F_{\hat{\boldsymbol{\mu}}(t)}(\hat{\boldsymbol{\omega}}(t))$. Hence, $\sum_{t=\overline{h}(T)}^{T} \mathbb{1}\{\tau > t\}$ is upper bounded by

$$\sum_{t=\overline{h}(T)}^{T} \mathbb{1}\left\{t(T^\star(\boldsymbol{\mu})^{-1} - 6\epsilon) \leq (1+\epsilon_t)\,\beta\left(t, \frac{(4|\mathcal{X}_0|-1)\delta}{4|\mathcal{X}_0|}\right)\right\} \leq \frac{(1+\epsilon_T)}{T^\star(\boldsymbol{\mu})^{-1} - 6\epsilon}\beta\left(T, \frac{(4|\mathcal{X}_0|-1)\delta}{4|\mathcal{X}_0|}\right).$$

By defining $T_0(\delta)$ as done in (28), we get (28), i.e.,

$$\mathbb{E}_{\boldsymbol{\mu}}[\tau] \leq \sum_{T=1}^{\infty} \mathbb{P}_{\boldsymbol{\mu}}[\tau \geq T] \leq T_0(\delta) + \sum_{T=M+1}^{\infty} \mathbb{P}_{\boldsymbol{\mu}}[(\mathcal{E}_{1,\epsilon}(T) \cap \mathcal{E}_{2,\epsilon}(T))^c]$$

because $\mathcal{E}_{1,\epsilon}(T) \cap \mathcal{E}_{2,\epsilon}(T) \subseteq \{\tau \leq T\}$ for any $T \geq T_0(\delta)$.

Now, we proceed with the proof by upper-bounding each term in the right-hand side of (28).

Bounding $T_0(\delta)$: Introduce $\tilde{\epsilon} \in (0,1)$ that can be chosen arbitrarily small. Notice that

$$T - \overline{h}(T) = T - T^{a+b} \geq \frac{T}{1+\tilde{\epsilon}} \quad \text{when} \quad T \geq \left(1 + \frac{1}{\tilde{\epsilon}}\right)^{\frac{1}{1-(a+b)}}, \tag{30}$$

$$\epsilon_T = T^{-\frac{1}{9}} \leq \tilde{\epsilon} \quad \text{when} \quad T \geq \left(\frac{1}{\tilde{\epsilon}}\right)^9, \tag{31}$$

where the first inequality results from a simple rearrangement, and the second substitutes $\epsilon_t = t^{-1/9}$. Then, it follows from (9) that:

$$\begin{aligned}
T_0(\delta) &\leq \inf\left\{T \geq \max\left\{M, (1+\tilde{\epsilon}^{-1})^{\frac{1}{1-(a+b)}}, \tilde{\epsilon}^{-9}\right\} : \frac{(1+\tilde{\epsilon})\,\beta(T, \frac{3\delta}{4})}{T^\star(\boldsymbol{\mu})^{-1} - 6\epsilon} \leq \frac{T}{1+\tilde{\epsilon}}\right\} \\
&\leq \inf\left\{T \geq \max\left\{M, (1+\tilde{\epsilon}^{-1})^{\frac{1}{1-(a+b)}}, \tilde{\epsilon}^{-9}, c_1\right\} : \ln\left(\frac{4c_2 T}{3\delta}\right) \leq \frac{T^\star(\boldsymbol{\mu})^{-1} - 6\epsilon}{(1+\tilde{\epsilon})^2} \cdot T\right\} \\
&\leq \max\left\{M, (1+\tilde{\epsilon}^{-1})^{\frac{1}{1-(a+b)}}, \tilde{\epsilon}^{-9}, c_1\right\} + \frac{(1+\tilde{\epsilon})^2}{T^\star(\boldsymbol{\mu})^{-1} - 6\epsilon} \times H\left(\frac{4c_2}{3\delta} \cdot \frac{(1+\tilde{\epsilon})^2}{T^\star(\boldsymbol{\mu})^{-1} - 6\epsilon}\right),
\end{aligned} \tag{32}$$

where the first inequality uses (30)-(31) and $\frac{4|\mathcal{X}_0|-1}{4|\mathcal{X}_0|} \geq \frac{3}{4}$ (as $|\mathcal{X}_0| \geq 2$ is shown in Lemma 23 in Appendix J), the second inequality is due to (9), and the last results from an application of Lemma 9 in Appendix D.3.3 with

$$\alpha = 1, \quad b_1 = \frac{T^\star(\boldsymbol{\mu})^{-1} - 6\epsilon}{(1+\tilde{\epsilon})^2}, \quad \text{and} \quad b_2 = \frac{4c_2}{3\delta}.$$

Bounding $\sum_{T=M+1}^{\infty} \mathbb{P}_{\boldsymbol{\mu}}[(\mathcal{E}_{1,\epsilon}(T) \cap \mathcal{E}_{2,\epsilon}(T))^c]$: By Lemma 8 in Appendix D.3.3, it is upper bounded by

$$2K\left(\frac{3}{\min\{1, \frac{\epsilon^2}{8\ell^2 K^3 D^2}\}^{2+\frac{2}{a}}} + 2\left(\frac{2304 D^6 \|\boldsymbol{\mu}\|_\infty^2 \sqrt{|\mathcal{X}_0|}}{\epsilon^2}\right)^{2+\frac{2}{a}}\right)\Gamma\left(2+\frac{2}{a}\right). \tag{33}$$

Putting things together: Finally, substituting $(a, b) = (\frac{7}{9}, \frac{1}{9})$ into (25)-(32)-(33) yields that:

- $T_0(\delta) \leq M + \left(1+\frac{1}{\tilde{\epsilon}}\right)^9 + \left(\frac{1}{\tilde{\epsilon}}\right)^9 + c_1 + \frac{(1+\tilde{\epsilon})^2}{T^\star(\boldsymbol{\mu})^{-1}-6\epsilon} \times H\left(\frac{4c_2}{3\delta} \cdot \frac{(1+\tilde{\epsilon})^2}{T^\star(\boldsymbol{\mu})^{-1}-6\epsilon}\right)$
- $M \leq \max\{(4|\mathcal{X}_0|)^{\frac{9}{7}}, (\frac{4K^2}{\epsilon^2 D^2 |\mathcal{X}_0|})^{\frac{9}{7}}, (\frac{4}{\triangle_{\min}^2})^{\frac{9}{7}}, (\frac{9\|\boldsymbol{\mu}\|_\infty^2}{4})^{\frac{9}{7}}\} + \max\{(\frac{\ell}{\epsilon})^9, (\frac{5\ell K^2}{\epsilon\sqrt{|\mathcal{X}_0|}})^{2.25}\}$
- $\sum_{T=M}^{\infty} \mathbb{P}_{\boldsymbol{\mu}}[(\mathcal{E}_{1,\epsilon}(T) \cap \mathcal{E}_{2,\epsilon}(T))^c] < \frac{78K}{\epsilon^{10}}\left(2^{15}K^{15}D^{10}\ell^{10} + 2^{41}3^9 D^{30}\|\boldsymbol{\mu}\|_\infty^{10}|\mathcal{X}_0|^{2.5}\right)$

where simplifications are obtained remarking that $\Gamma(2+\frac{2}{a}) \leq 13$ and $2+\frac{2}{a} < 5$. Therefore, substituting $\ell = 2D^2\|\boldsymbol{\mu}\|_\infty^2$ (defined in Appendix I) and $78 < 2^7$, $3^9 \leq 2^{15}$, and $4^{9/7} < 6$ lead to:

$$\mathbb{E}_{\boldsymbol{\mu}}[\tau] \leq \frac{(1+\tilde{\epsilon})^2}{T^\star(\boldsymbol{\mu})^{-1}-6\epsilon} \times H\left(\frac{4c_2}{3\delta} \cdot \frac{(1+\tilde{\epsilon})^2}{T^\star(\boldsymbol{\mu})^{-1}-6\epsilon}\right) + \Psi(\epsilon, \tilde{\epsilon}),$$

where

$$\Psi(\epsilon,\tilde{\epsilon}) = 6\max\left\{|\mathcal{X}_0|, \frac{K^2}{\epsilon^2 D^2|\mathcal{X}_0|}, \frac{1}{\triangle_{\min}^2}, \|\boldsymbol{\mu}\|_\infty^2\right\}^{\frac{9}{7}} + \max\left\{\frac{2^9 D^{18}\|\boldsymbol{\mu}\|_\infty^{18}}{\epsilon^9}, \frac{10^{2.25}D^{4.5}\|\boldsymbol{\mu}\|_\infty^{4.5}K^2}{\epsilon^{2.25}|\mathcal{X}_0|^{1.125}}\right\}$$
$$+\left(1+\frac{1}{\tilde{\epsilon}}\right)^9 + \left(\frac{1}{\tilde{\epsilon}}\right)^9 + c_1 + \frac{2^{32}KD^{30}\|\boldsymbol{\mu}\|_\infty^{10}\left(K^{15}\|\boldsymbol{\mu}\|_\infty^{10} + 2^{31}|\mathcal{X}_0|^{2.5}\right)}{\epsilon^{10}}. \tag{34}$$

D.3.3 Technical lemmas

The most important step in Theorem 4 is to bound the term $\sum_{T=M+1}^{\infty} \mathbb{P}_{\boldsymbol{\mu}}[(\mathcal{E}_{1,\epsilon}(T) \cap \mathcal{E}_{2,\epsilon}(T))^c]$ in (28) explicitly in terms of K, $\|\boldsymbol{\mu}\|_\infty$ and ϵ. For this purpose, inspired by Assumption 3 in [WTP21], we developed Proposition 4 (see Appendix G.2 for the proof) and combine it with the mean-valued theorem to derive our main continuity results in Lemma 7 (see Appendix G for the proof). Throughout this section, we fix $\boldsymbol{\mu} \in \Lambda$ and denote $\triangle_{\min}(\boldsymbol{\mu})$ by $\triangle_{\min}$.

Lemma 7. *Let $\epsilon \in (0, \frac{2D^2\triangle_{\min}^2}{K})$. Then, any $\boldsymbol{\pi} \in \mathbb{R}^K$ with $\|\boldsymbol{\pi} - \boldsymbol{\mu}\|_\infty < \frac{\epsilon}{24D^3\|\boldsymbol{\mu}\|_\infty}$ satisfies the following:*

$$|F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) - F_{\boldsymbol{\pi}}(\boldsymbol{\omega})| < \epsilon, \quad \forall \boldsymbol{\omega} \in \Sigma_+ \tag{35}$$
$$|\langle \nabla \bar{F}_{\boldsymbol{\pi},\eta}(\boldsymbol{\omega}) - \nabla \bar{F}_{\boldsymbol{\mu},\eta}(\boldsymbol{\omega}), \boldsymbol{x} - \boldsymbol{\omega}\rangle| < \epsilon, \quad \forall(\boldsymbol{\omega}, \boldsymbol{x}) \in \Sigma_+ \times \mathcal{X}, \forall \eta \in (0, \min_{k\in[K]} \omega_k). \tag{36}$$

Our main concentration result with error specified explicitly in terms of ϵ is (see Appendix F for the proof):

Lemma 8. *Let $\epsilon \in (0, \frac{2D^2\triangle_{\min}^2}{K})$ and M be defined as in (25). Then,*

$$\sum_{T=M+1}^{\infty} \mathbb{P}_{\boldsymbol{\mu}}[(\mathcal{E}_{1,\epsilon}(T) \cap \mathcal{E}_{2,\epsilon}(T))^c] < 2K\left(\frac{3}{A_1(\epsilon)^{2+\frac{2}{a}}} + \frac{2}{A_2(\epsilon)^{2+\frac{2}{a}}}\right)\Gamma\left(2+\frac{2}{a}\right),$$

where $A_1(\epsilon) = \min\{1, \frac{\epsilon^2}{8\ell^2K^3D^2}\}$, $A_2(\epsilon) = \frac{\epsilon^2}{2304D^6\|\boldsymbol{\mu}\|_\infty^2\sqrt{|\mathcal{X}_0|}}$, and Γ denotes the gamma function $\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt$ for any $z > 0$.

We remark that Lemma 8 sharpens a similar result, Lemma 2 in [WTP21], by a factor of e^K after performing a more careful analysis.

Under good events $\mathcal{E}_{1,\epsilon}(T) \cap \mathcal{E}_{2,\epsilon}(T)$, we show Theorem 5, the convergence of `P-FWS` when $\hat{\boldsymbol{\mu}}(t)$ is replaced with $\boldsymbol{\mu}$. As shown in Appendix D.3.2, the extra error due to this replacement is controlled, thanks to Lemma 7.

Theorem 5. *Let* $\epsilon \in (0, \min\{1, \frac{2D^2\triangle_{\min}^2}{K}\})$ *and* T *be an integer at least larger than*

$$\max\left\{(4|\mathcal{X}_0|)^{\frac{1}{a}}, \left(\frac{4K^2}{\epsilon^2 D^2|\mathcal{X}_0|}\right)^{\frac{1}{a}}, \left(\frac{2}{\triangle_{\min}}\right)^{\frac{2}{a}}, \left(\frac{3\|\boldsymbol{\mu}\|_\infty}{2}\right)^{\frac{2}{a}}\right\} + \max\left\{\left(\frac{\ell}{\epsilon}\right)^{\frac{1}{b}}, \left(\frac{5\ell K^2}{\epsilon\sqrt{|\mathcal{X}_0|}}\right)^{\frac{2}{a+b}}\right\}.$$

Under $\mathcal{E}_{1,\epsilon}(T) \cap \mathcal{E}_{2,\epsilon}(T)$*, Algorithm 2 with* (10) *satisfies that: for any* $t = \overline{h}(T), \overline{h}(T)+1\cdots, T$,

$$\text{(i) } \max_{\boldsymbol{\omega}\in\Sigma} F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) - F_{\boldsymbol{\mu}}(\hat{\boldsymbol{\omega}}(t)) \le 5\epsilon, \text{ (ii) } \triangle_{\min}(\hat{\boldsymbol{\mu}}(t)) \ge \frac{\triangle_{\min}}{2}, \text{ and (iii) } \|\hat{\boldsymbol{\mu}}(t)\|_\infty \le \frac{3\|\boldsymbol{\mu}\|_\infty}{2}.$$

The proof of Theorem 5 is given in Appendix E.

Finally, the last ingredient is Lemma 9.

Lemma 9 (Lemma 18 in [GK16])**.** *Let* $\alpha \in [1, \frac{e}{2}]$ *and* $b_1, b_2 > 0$*. Then,*

$$x = \frac{1}{b_1}\left(\ln\left(\frac{b_2 e}{b_1^\alpha}\right) + \ln\ln\left(\frac{b_2}{b_1^\alpha}\right)\right)$$

satisfies $b_1 x \ge \ln(b_2 x^\alpha)$.

D.4 Computational complexity (Theorem 4 (iii))

In this section, we analyze the computational complexity of `P-FWS` running with (10) in terms of the number of calls to `LM` Oracle. We will show that the expected number of `LM` Oracle calls is upper bounded by a polynomial in $\ln\delta^{-1}$, K, $\|\boldsymbol{\mu}\|_\infty$ and $\triangle_{\min}(\boldsymbol{\mu})^{-1}$.

Proof The construction of $\mathcal{X}_0$ and computation of $\hat{\boldsymbol{\imath}}$ merely takes $\mathcal{O}(KD)$ calls to `LM` Oracle. The overall complexity is dominated by the `LM` Oracle calls performed from $4|\mathcal{X}_0|+1$ to round τ, analyzed as follows.

Per-round complexity: Fix $t \in \{4|\mathcal{X}_0|+1, \cdots, \tau\}$. Recall from `P-FWS` that the FW update in round t and the stopping rule in round $t-1$ are computed only if:

$$\max\{\triangle_{\min}(\hat{\boldsymbol{\mu}}(t-1))^{-1}, \|\hat{\boldsymbol{\mu}}(t-1)\|_\infty\} \le \sqrt{t-1}. \tag{37}$$

Otherwise, forced-exploration procedure is invoked. Verifying (37) takes at most $D+1$ calls.[6] The computation of $\hat{F}_{t-1}$ and $\boldsymbol{i}^\star(\nabla\tilde{F}_{\hat{\boldsymbol{\mu}}(t-1),\eta_t,n_t}(\hat{\boldsymbol{\omega}}(t-1)))$ by Theorem 3 in Appendix C.2 takes at most

$$\mathcal{O}\left(D + \frac{\|\hat{\boldsymbol{\mu}}(t-1)\|_\infty^4 \left\|\hat{\boldsymbol{\omega}}(t-1)^{-1}\right\|_\infty^2 K^3 D^5 \ln K}{F_{\hat{\boldsymbol{\mu}}(t-1)}(\hat{\boldsymbol{\omega}}(t-1))^2}\left(\frac{\ln(t^2\delta^{-1})}{\epsilon_t^2} + \frac{n_t \ln\theta_t^{-1}}{\rho_t^2}\right)\right) \tag{38}$$

calls to `LM` Oracle. To evaluate (38), we need a lower bound on $F_{\hat{\boldsymbol{\mu}}(t-1)}(\hat{\boldsymbol{\omega}}(t-1))$. By Proposition 1 (c) in Appendix C.1, one evaluates $F_{\hat{\boldsymbol{\mu}}(t-1)}(\hat{\boldsymbol{\omega}}(t-1))$ in closed-form:

$$F_{\hat{\boldsymbol{\mu}}(t-1)}(\hat{\boldsymbol{\omega}}(t-1)) = \min_{\boldsymbol{x}\neq\boldsymbol{i}^\star(\hat{\boldsymbol{\mu}}(t-1))} \frac{\triangle_{\boldsymbol{x}}(\hat{\boldsymbol{\mu}}(t-1))^2}{2\left\langle \boldsymbol{x}\oplus\boldsymbol{i}^\star(\hat{\boldsymbol{\mu}}(t-1)), \hat{\boldsymbol{\omega}}(t-1)^{-1}\right\rangle} \ge \frac{\min_{k\in[K]}\hat{\omega}_k(t-1)}{4D(t-1)},$$

where the inequality results from (37) that $\triangle_{\min}(\hat{\boldsymbol{\mu}}(t-1)) \ge \frac{1}{\sqrt{t-1}}$, $\|\boldsymbol{x}\oplus\boldsymbol{x}'\|_1 \le 2D$ for any $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{X}$, and $\langle \boldsymbol{y}, \boldsymbol{z}\rangle \le \|\boldsymbol{y}\|_1 \|\boldsymbol{z}\|_\infty$ for any $\boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^K$. Further, combining with Lemma 14 (which states $\min_{k\in[K]}\hat{\omega}_k(t-1) \ge \frac{1}{2\sqrt{(t-1)|\mathcal{X}_0|}}$) in Appendix F yields

$$F_{\hat{\boldsymbol{\mu}}(t-1)}(\hat{\boldsymbol{\omega}}(t-1)) \ge \frac{1}{8D\sqrt{|\mathcal{X}_0|}(t-1)^{1.5}}. \tag{39}$$

[6] For any $\boldsymbol{\pi} \in \Lambda$, $\triangle_{\min}(\boldsymbol{\pi})$ requires to compute $\boldsymbol{i}^\star(\boldsymbol{\pi})$ and and solve $\max_{\boldsymbol{x}\neq\boldsymbol{i}^\star(\boldsymbol{\pi})}\langle\boldsymbol{\pi}, \boldsymbol{x}\rangle$, where the latter requires at most D calls to the `LM` Oracle by Lemma 2 in § 2.2.

From (39), $\|\hat{\boldsymbol{\mu}}(t-1)\|_\infty \le \sqrt{t-1}$, Lemma 14, and substituting the parameters (10) into (38), we know that the number of `LM` Oracle calls performed at any round $t \ge 4|\mathcal{X}_0|+1$ is at most

$$\begin{aligned}
&\mathcal{O}\bigg(t^6|\mathcal{X}_0|^2K^3D^7\ln K\left(\frac{\ln(t^2\delta^{-1})}{\epsilon_t^2}+\frac{n_t\ln\theta_t^{-1}}{\rho_t^2}\right)\bigg)\\
&\quad=\mathcal{O}\bigg(t^6|\mathcal{X}_0|^2K^3D^7\ln K\left(t^{\frac{2}{9}}\ln\left(\frac{t}{\delta}\right)+t^{2.75}D^4|\mathcal{X}_0|^2\right)\bigg)\\
&\quad=\mathcal{O}\bigg(t^{8.75}\ln\left(\frac{t}{\delta}\right)|\mathcal{X}_0|^4D^{11}K^3\ln K\bigg). \qquad (40)
\end{aligned}$$

<u>Overall complexity</u>: Invoking Theorem 4 in D.3 with $\tilde{\epsilon}=0.1$ and $\epsilon=\frac{1}{12T^\star(\boldsymbol{\mu})}$ results in

$$\mathbb{E}_{\boldsymbol{\mu}}[\tau]=\mathcal{O}\Bigg(T^\star(\boldsymbol{\mu})\ln\bigg(\frac{T^\star(\boldsymbol{\mu})}{\delta}\bigg)+\frac{1}{\triangle_{\min}^{\frac{18}{7}}}+K^{16}D^{30}\|\boldsymbol{\mu}\|_\infty^{20}T^\star(\boldsymbol{\mu})^{10}\Bigg)$$

which after using $T^\star(\boldsymbol{\mu})\le 4KD/\triangle_{\min}^2$ (Lemma 1 in §2.1) becomes

$$\mathbb{E}_{\boldsymbol{\mu}}[\tau]=\mathcal{O}\Bigg(\frac{KD}{\triangle_{\min}^2}\ln\bigg(\frac{KD}{\delta\triangle_{\min}^2}\bigg)+\frac{K^{26}D^{40}\|\boldsymbol{\mu}\|_\infty^{20}}{\triangle_{\min}^{20}}\Bigg). \qquad (41)$$

Hence, by a summation of (40) over $t=4|\mathcal{X}_0|+1$ to $\mathbb{E}_{\boldsymbol{\mu}}[\tau]$, the expected total number of the `LM` Oracle calls is upper bounded by

$$\mathcal{O}\bigg(\mathbb{E}_{\boldsymbol{\mu}}[\tau]^{9.75}\ln\left(\frac{\mathbb{E}_{\boldsymbol{\mu}}[\tau]}{\delta}\right)|\mathcal{X}_0|^4D^{11}K^3\ln K\bigg), \qquad (42)$$

where the inequality uses integral by parts $\int t^{8.75}\ln t dt = \mathcal{O}(t^{9.75}\ln t)$. Remind that $\max\{D,|\mathcal{X}_0|\}\le K$. Thus, we conclude that (42) is bounded by a polynomial function in $\ln\delta^{-1}$, $\|\boldsymbol{\mu}\|_\infty$, $\triangle_{\min}^{-1}$, and K (due to (41), $\mathbb{E}_{\boldsymbol{\mu}}[\tau]$ is bounded by a polynomial function in the same variables). $\square$

E Convergence of P-FWS under the good events

Throughout this section, we assume that $\boldsymbol{\mu}$ is fixed and drop $\boldsymbol{\mu}$ from the notation, e.g., $F = F_{\boldsymbol{\mu}}$, $\bar{F}_\eta = \bar{F}_{\boldsymbol{\mu},\eta}$, $\tilde{F}_{\eta,t} = \tilde{F}_{\boldsymbol{\mu},\eta,t}$, and $\triangle_{\min} = \triangle_{\min}(\boldsymbol{\mu})$. Also, we will use $\boldsymbol{\omega}^\star \in \operatorname{argmax}_{\boldsymbol{\omega}\in\Sigma} F(\boldsymbol{\omega})$ to denote any optimal allocation and let $\boldsymbol{i}^\star = \boldsymbol{i}^\star(\boldsymbol{\mu})$. Recall that $\underline{h}(T) \geq T^a$ and $\overline{h}(T) \geq T^{a+b}$ is defined in (23) in Appendix D.3.1 for some $a, b \in (0,1)$.

Theorem 5. *Let $\epsilon \in (0, \min\{1, \frac{2D^2\triangle_{\min}^2}{K}\})$ and T be an integer at least larger than*

$$\max\left\{(4|\mathcal{X}_0|)^{\frac{1}{a}}, \left(\frac{4K^2}{\epsilon^2 D^2|\mathcal{X}_0|}\right)^{\frac{1}{a}}, \left(\frac{2}{\triangle_{\min}}\right)^{\frac{2}{a}}, \left(\frac{3\,\|\boldsymbol{\mu}\|_\infty}{2}\right)^{\frac{2}{a}}\right\} + \max\left\{\left(\frac{\ell}{\epsilon}\right)^{\frac{1}{b}}, \left(\frac{5\ell K^2}{\epsilon\sqrt{|\mathcal{X}_0|}}\right)^{\frac{2}{a+b}}\right\}.$$

Under $\mathcal{E}_{1,\epsilon}(T)\cap\mathcal{E}_{2,\epsilon}(T)$, Algorithm 2 with (10) *satisfies that: for any $t = \overline{h}(T), \overline{h}(T)+1, \cdots, T$,*

$$\textit{(i) } F(\boldsymbol{\omega}^\star) - F(\hat{\boldsymbol{\omega}}(t)) \leq 5\epsilon, \textit{ (ii) } \triangle_{\min}(\hat{\boldsymbol{\mu}}(t)) \geq \frac{\triangle_{\min}}{2}, \textit{ and (iii) } \|\hat{\boldsymbol{\mu}}(t)\|_\infty \leq \frac{3\,\|\boldsymbol{\mu}\|_\infty}{2}.$$

Proof Fix arbitrary ϵ and T that satisfy the conditions in the statement, and suppose $\mathcal{E}_{1,\epsilon}(T)\cap\mathcal{E}_{2,\epsilon}(T)$ holds. (ii)(iii) directly follows from Lemma 10 (where one can verified that its assumption of Lemma 10 on T is satisfied). With (ii)(iii), the analysis of FW convergence will be greatly simplified as (ii)(iii) ensure that

$$\max\left\{\frac{1}{\triangle_{\min}(\hat{\boldsymbol{\mu}}(t-1))}, \|\hat{\boldsymbol{\mu}}(t-1)\|\right\} \leq \sqrt{t-1}.$$

This means that the forced-exploration procedure will only be invoked by the condition of $\sqrt{t/|\mathcal{X}_0|}$ when $t \geq \underline{h}(T) = T^a$.

Proof of (i) $F(\boldsymbol{\omega}^\star) - F(\hat{\boldsymbol{\omega}}(t)) \leq 5\epsilon$: Fix $t \geq \underline{h}(T)$. As mentioned above, for such t, the forced-exploration procedure will be invoked only when $\sqrt{t/|\mathcal{X}_0|} \in \mathbb{N}$. To specify the rounds performing FW udpates, introduce $s(t) = \lfloor\sqrt{t/|\mathcal{X}_0|}\rfloor - 1$ and define

$$p(t) = (s(t)^2+1)|\mathcal{X}_0| \quad \text{and} \quad q(t) = (s(t)+1)^2|\mathcal{X}_0| - 1.$$

Notice that $p(t)$ and $q(t)$ are respectively the starting (including) and the ending (including) round of a successive FW update rounds with no forced exploration in between. Let $\phi_t = F(\boldsymbol{\omega}^\star) - \bar{F}_{\eta_t}(\hat{\boldsymbol{\omega}}(t))$ be the error. By a careful analysis, we derive a recursive relationship satisfied by ϕ_t (Lemma 11):

$$\begin{cases} t\phi_t \leq (t-|\mathcal{X}_0|)\phi_{t-|\mathcal{X}_0|} + 2\ell\sqrt{D}|\mathcal{X}_0|^2 & \text{if } t = p(t)-1, \\ t\phi_t \leq (t-1)\phi_{t-1} + 3\epsilon + \ell\left(\eta_{t-1} + \frac{K^2}{2t\eta_t}\right) & \text{if } t \in [p(t), q(t)]. \end{cases} \tag{43}$$

The first case (in round $t = p(t)-1$) is exactly the ending round of a forced-exploration procedure (from $t-|\mathcal{X}_0|, \cdots, t$), and the second case (in round $t \in [p(t), q(t)]$) is a FW-update round. By repeatedly applying (43), we have

$$\begin{aligned} \overline{h}(T)\phi_{\overline{h}(T)} &\leq \underline{h}(T)\phi_{\underline{h}(T)} + 2\ell\sqrt{D}|\mathcal{X}_0|^2\left(s(\overline{h}(T)) - s(\underline{h}(T))\right) + 3\sum_{t=\underline{h}(T)}^{\overline{h}(T)}\left(\frac{\ell K^2}{\sqrt{t|\mathcal{X}_0|}} + \epsilon\right) \\ &\leq \underline{h}(T)\ell + \ell\left(2\sqrt{D}|\mathcal{X}_0|^{1.5} + \frac{3K^2}{\sqrt{|\mathcal{X}_0|}}\right)\left(\sqrt{\overline{h}(T)} - \sqrt{\underline{h}(T)}\right) + 3\epsilon(\overline{h}(T) - \underline{h}(T)), \end{aligned}$$

where the second inequality follows from $\phi_{\overline{h}(T)} \leq \max_{\boldsymbol{\omega}\in\Sigma} F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) \leq \ell$ (Lemma 22 in Appendix I), $s(\overline{h}(T)) - s(\underline{h}(T)) \leq \frac{\sqrt{\overline{h}(T)} - \sqrt{\underline{h}(T)}}{\sqrt{\mathcal{X}_0}}$ and $\sum_{t=\underline{h}(T)}^{\overline{h}(T)} \frac{1}{\sqrt{t}} \leq \sqrt{\overline{h}(T)} - \sqrt{\underline{h}(T)}$. Substituting $\underline{h}(T)$ and $\overline{h}(T)$ from (23) and simplifying the terms, we get:

$$F(\boldsymbol{\omega}^\star) - F(\hat{\boldsymbol{\omega}}(t)) \leq \phi_{\overline{h}(T)} \leq \ell T^{-b} + \frac{5\ell K^2}{\sqrt{|\mathcal{X}_0|}}T^{-\frac{a+b}{2}} + 3\epsilon(1 - T^{-b}) \leq 5\epsilon$$

when

$$T \geq \max\left\{(4|\mathcal{X}_0|)^{\frac{1}{a}}, \left(\frac{4K^2}{\epsilon^2 D^2 |\mathcal{X}_0|}\right)^{\frac{1}{a}}, \left(\frac{2}{\triangle_{\min}}\right)^{\frac{2}{a}}, \left(\frac{3\|\boldsymbol{\mu}\|_\infty}{2}\right)^{\frac{2}{a}}\right\} + \max\left\{\left(\frac{\ell}{\epsilon}\right)^{\frac{1}{b}}, \left(\frac{5\ell K^2}{\epsilon\sqrt{|\mathcal{X}_0|}}\right)^{\frac{2}{a+b}}\right\},$$

where the first inequality is due to $F(\hat{\boldsymbol{\omega}}(t)) \geq \bar{F}_{\eta_t}(\hat{\boldsymbol{\omega}}(t))$ by Proposition 2 (i) in §4.1. $\square$

Lemma 10. *Let $\epsilon \in (0, \min\{1, 2D^2\triangle_{\min}^2/K\})$ and T be a positive integer s.t.*

$$T \geq \max\left\{(4|\mathcal{X}_0|)^{\frac{1}{a}}, \left(\frac{4K^2}{\epsilon^2 D^2 |\mathcal{X}_0|}\right)^{\frac{1}{a}}, \left(\frac{2}{\triangle_{\min}}\right)^{\frac{2}{a}}, \left(\frac{3\|\boldsymbol{\mu}\|_\infty}{2}\right)^{\frac{2}{a}}\right\}. \tag{44}$$

Suppose $\mathcal{E}_{1,\epsilon}(T) \cap \mathcal{E}_{2,\epsilon}(T)$ holds. Then, for any $t \geq \underline{h}(T)$,

$$\triangle_{\min}(\hat{\boldsymbol{\mu}}(t)) \geq \frac{\triangle_{\min}}{2} \quad \textit{and} \quad \|\hat{\boldsymbol{\mu}}(t)\|_\infty \leq \frac{3\|\boldsymbol{\mu}\|_\infty}{2}. \tag{45}$$

Proof Fix any T satisfying (45) and suppose $\mathcal{E}_{1,\epsilon}(T) \cap \mathcal{E}_{2,\epsilon}(T)$ holds. Consider any $t \geq T^a$ and hence $t \geq 4|\mathcal{X}_0|$. To show the first inequality of (45), from $\mathcal{E}_{2,\epsilon}(T)$ and $\epsilon < 2D^2\triangle_{\min}^2/K$, we have

$$\triangle_{\min}(\hat{\boldsymbol{\mu}}(t-1)) \geq \triangle_{\min} - \frac{2D\epsilon}{24D^3\|\boldsymbol{\mu}\|_\infty} > \triangle_{\min} - \frac{\triangle_{\min}^2}{6K\|\boldsymbol{\mu}\|_\infty} > \frac{\triangle_{\min}}{2},$$

where the last inequality is because $\triangle_{\min} \leq 2D\|\boldsymbol{\mu}\|_\infty$ and $D \leq K$. To show the second inequality of (45), observe that

$$\|\hat{\boldsymbol{\mu}}(t)\|_\infty \leq \|\boldsymbol{\mu}\|_\infty + \frac{\epsilon}{24D^3\|\boldsymbol{\mu}\|_\infty} < \|\boldsymbol{\mu}\|_\infty + \frac{\triangle_{\min}^2}{12KD\|\boldsymbol{\mu}\|_\infty} < \frac{3\|\boldsymbol{\mu}\|_\infty}{2},$$

where the first inequality is because of $\mathcal{E}_{2,\epsilon}(T)$, the second is due to $\epsilon < 2D^2\triangle_{\min}^2/K$, and the last uses $\triangle_{\min} \leq 2D\|\boldsymbol{\mu}\|_\infty$. $\square$

Lemma 11. *Let $\epsilon > 0$ and $t \in \mathbb{N}$ be such that* (45) *holds. Then, under the event $\mathcal{E}_{1,\epsilon}^{(t)} \cap \mathcal{E}_{2,\epsilon}^{(t)}$,*

$$\begin{cases} t\phi_t \leq (t - |\mathcal{X}_0|)\phi_{t-|\mathcal{X}_0|} + 2\ell\sqrt{D}|\mathcal{X}_0|^2 & \textit{if } t = p(t) - 1 \\ t\phi_t \leq (t-1)\phi_{t-1} + 3\epsilon + \ell\left(\eta_{t-1} + \frac{K^2}{2t\eta_t}\right) & \textit{if } t \in [p(t), q(t)] \end{cases}, \tag{43}$$

where $p(t) = (s(t)^2 + 1)|\mathcal{X}_0|$, $q(t) = (s(t)+1)^2|\mathcal{X}_0| - 1$, and $s(t) = \lfloor\sqrt{t/|\mathcal{X}_0|}\rfloor - 1$.

Proof The first case basically follows from the Lipschitzness of F (Appendix I), whereas the second relies on results on stochastic smoothing (Appendix H).

Case $t = p(t) - 1$: In this case, round t is exactly the end (including) round of a forced-exploration procedure. By ℓ-Lipschitzness of F (Lemma 21 in Appendix I),

$$F(\hat{\boldsymbol{\omega}}(t)) - F(\hat{\boldsymbol{\omega}}(t - |\mathcal{X}_0|)) \geq -\ell\|\hat{\boldsymbol{\omega}}(t) - \hat{\boldsymbol{\omega}}(t - |\mathcal{X}_0|)\|_2 \geq -\frac{\ell\sqrt{D}|\mathcal{X}_0|^2}{t},$$

where the second inequality stems from $\hat{\boldsymbol{\omega}}(t) = \frac{t-|\mathcal{X}_0|}{t}\hat{\boldsymbol{\omega}}(t - |\mathcal{X}_0|) + \frac{|\mathcal{X}_0|}{t}\sum_{\boldsymbol{x}\in\mathcal{X}_0}\boldsymbol{x}$ after performing the forced exploration. It then follows that $\|\hat{\boldsymbol{\omega}}(t) - \hat{\boldsymbol{\omega}}(t - |\mathcal{X}_0|)\|_2 \leq \frac{\sqrt{D}|\mathcal{X}_0|^2}{t}$. By $\max_{\boldsymbol{\omega}\in\Sigma} F(\boldsymbol{\omega}) \leq \ell$ (Lemma 22 in Appendix I) and a rearrangement of the above yields

$$t\phi_t \leq t\phi_{t-|\mathcal{X}_0|} + \ell\sqrt{D}|\mathcal{X}_0|^2 \leq (t - |\mathcal{X}_0|)\phi_{t-|\mathcal{X}_0|} + \ell|\mathcal{X}_0|(\sqrt{D}|\mathcal{X}_0| + 1).$$

The proof is completed after simplifying the terms.

Case: $t \in [p(t), q(t)]$: In this case, round t performs a FW update. For brevity, let $\boldsymbol{z} = \hat{\boldsymbol{\omega}}(t)$ and $\boldsymbol{y} = \hat{\boldsymbol{\omega}}(t-1)$. By $\frac{\ell K}{\eta_t}$-smoothness of $\bar{F}_{\eta_t}$ (Proposition 2 (iii) in §4.1) and $\boldsymbol{z} - \boldsymbol{y} = \frac{1}{t}(\boldsymbol{x}(t) - \boldsymbol{y})$,

$$\bar{F}_{\eta_t}(\boldsymbol{z}) \geq (*) - \frac{\ell K}{2\eta_t}\|\boldsymbol{z} - \boldsymbol{y}\|_2^2 \geq (*) - \frac{\ell K}{2t^2\eta_t}\|\boldsymbol{x}(t) - \boldsymbol{y}\|_2^2 \geq (*) - \frac{\ell K^2}{2t^2\eta_t},$$

where $(*) = \bar{F}_{\eta_t}(\boldsymbol{y}) + \langle \nabla \bar{F}_{\eta_t}(\boldsymbol{y}), \boldsymbol{z} - \boldsymbol{y} \rangle = \bar{F}_{\eta_t}(\boldsymbol{y}) + \frac{1}{t} \langle \nabla \bar{F}_{\eta_t}(\boldsymbol{y}), \boldsymbol{x}(t) - \boldsymbol{y} \rangle$. It follows from $\mathcal{E}_{1,\epsilon}^{(t)} \cap \mathcal{E}_{2,\epsilon}^{(t)}$ and the continuity argument (Lemma 7 in Appendix G.1) that

$$\begin{aligned}\langle \nabla \bar{F}_{\eta_t}(\boldsymbol{y}), \boldsymbol{x}(t) - \boldsymbol{y} \rangle &\geq \langle \nabla \bar{F}_{\hat{\boldsymbol{\mu}}(t-1),\eta_t}(\boldsymbol{y}), \boldsymbol{x}(t) - \boldsymbol{y} \rangle - \epsilon \\ &\geq \max_{\boldsymbol{x} \in \mathcal{X}} \langle \nabla \bar{F}_{\hat{\boldsymbol{\mu}}(t-1),\eta_t}(\boldsymbol{y}), \boldsymbol{x} - \boldsymbol{y} \rangle - 2\epsilon \geq \max_{\boldsymbol{x} \in \mathcal{X}} \langle \nabla \bar{F}_{\eta_t}(\boldsymbol{y}), \boldsymbol{x} - \boldsymbol{y} \rangle - 3\epsilon.\end{aligned}$$

Then, the duality gap [Jag13] and the ℓ-Lipschitzness of F (Lemma 21 in Appendix I) yield

$$\begin{aligned}\max_{\boldsymbol{x} \in \mathcal{X}} \langle \nabla \bar{F}_{\eta_t}(\boldsymbol{y}), \boldsymbol{x} - \boldsymbol{y} \rangle &\geq \max_{\boldsymbol{\omega} \in \Sigma} \bar{F}_{\eta_t}(\boldsymbol{\omega}) - \bar{F}_{\eta_t}(\boldsymbol{y}) \\ &\geq (F(\boldsymbol{\omega}^\star) - \eta_t \ell) - (\bar{F}_{\eta_{t-1}}(\boldsymbol{y}) + \ell(\eta_{t-1} - \eta_t)) = \phi_{t-1} - \ell \eta_{t-1}.\end{aligned}$$

Therefore, $\bar{F}_{\eta_t}(\boldsymbol{z}) \geq \bar{F}_{\eta_t}(\boldsymbol{y}) + \frac{\phi_{t-1} - \ell\eta_{t-1} - 3\epsilon}{t} - \frac{\ell K^2}{2t^2 \eta_t}$ and subtracting $F(\boldsymbol{\omega}^\star)$ on both sides,

$$\begin{aligned}\phi_t &= F(\boldsymbol{\omega}^\star) - \bar{F}_{\eta_t}(\boldsymbol{z}) \\ &\leq (F(\boldsymbol{\omega}^\star) - \bar{F}_{\eta_t}(\boldsymbol{y})) + \frac{-\phi_{t-1} + \ell\eta_{t-1} + 3\epsilon}{t} + \frac{\ell K^2}{2t^2\eta_t} \\ &= \frac{t-1}{t}\phi_{t-1} + \frac{1}{t}\left(3\epsilon + \ell\left(\eta_{t-1} + \frac{K^2}{2t\eta_t}\right)\right),\end{aligned}$$

which completes the proof. $\square$

F Upper bound of $\sum_{T=M+1}^{\infty} \mathbb{P}_{\boldsymbol{\mu}}[(\mathcal{E}_{1,\epsilon}(T) \cap \mathcal{E}_{2,\epsilon}(T))^c]$ under P-FWS

Recall from (24) that $\mathcal{E}_{1,\epsilon}(T) = \cap_{t=\underline{h}(T)}^{T} \mathcal{E}_{1,\epsilon}^{(t)}$ and $\mathcal{E}_{2,\epsilon}(T) = \cap_{t=\underline{h}(T)}^{T} \mathcal{E}_{2,\epsilon}^{(t)}$, where

$$\mathcal{E}_{1,\epsilon}^{(t)} = \left\{ \left\langle \nabla \bar{F}_{\hat{\boldsymbol{\mu}}(t-1),\eta_t}(\hat{\boldsymbol{\omega}}(t-1)), \boldsymbol{x}(t) \right\rangle \geq \max_{\boldsymbol{x}\in\mathcal{X}} \left\langle \nabla \bar{F}_{\hat{\boldsymbol{\mu}}(t-1),\eta_t}(\hat{\boldsymbol{\omega}}(t-1)), \boldsymbol{x} \right\rangle - \epsilon \right\},$$

$$\mathcal{E}_{2,\epsilon}^{(t)} = \left\{ \|\hat{\boldsymbol{\mu}}(t-1) - \boldsymbol{\mu}\|_\infty < \frac{\epsilon}{24D^3 \|\boldsymbol{\mu}\|_\infty} \right\},$$

$T \geq M$ and M is defined in (25). Also, recall $\boldsymbol{x}(t) \in \operatorname{argmax}_{\boldsymbol{x}\in\mathcal{X}} \left\langle \nabla \tilde{F}_{\hat{\boldsymbol{\mu}}(t-1),\eta_t,n_t}(\hat{\boldsymbol{\omega}}(t-1)), \boldsymbol{x} \right\rangle$, where $\nabla \tilde{F}_{\hat{\boldsymbol{\mu}}(t-1),\eta_t,n_t}(\hat{\boldsymbol{\omega}}(t-1))$ is computed by (ρ_t, θ_t)-MCP algorithm with

$$(\eta_t, n_t, \rho_t, \theta_t) = \left(\frac{1}{4\sqrt{t|\mathcal{X}_0|}}, \left\lceil t^{\frac{1}{4}} \right\rceil, \frac{1}{16tD^2\mathcal{X}_0}, \frac{1}{t^{\frac{1}{4}} e^{\sqrt{t}}} \right). \tag{10}$$

Our main result Lemma 8 is built by bounding $\mathbb{P}_{\boldsymbol{\mu}}[\mathcal{E}_{1,\epsilon}(T)]$ and $\mathbb{P}_{\boldsymbol{\mu}}[\mathcal{E}_{2,\epsilon}(T)]$ separately with Lemma 12 in F.1 and Lemma 14 in F.2.

Lemma 8. *Let $\epsilon \in (0, 2D^2\triangle_{\min}^2/K)$ and M be defined as in* (25) *Then,*

$$\sum_{T=M+1}^{\infty} \mathbb{P}_{\boldsymbol{\mu}}[(\mathcal{E}_{1,\epsilon}(T) \cap \mathcal{E}_{2,\epsilon}(T))^c] < 2K \left(\frac{3}{A_1(\epsilon)^{2+\frac{2}{a}}} + \frac{2}{A_2(\epsilon)^{2+\frac{2}{a}}} \right) \Gamma\left(2 + \frac{2}{a}\right),$$

where $A_1(\epsilon) = \min\{1, \frac{\epsilon^2}{8\ell^2K^3D^2}\}$, $A_2(\epsilon) = \frac{\epsilon^2}{2304D^6\|\boldsymbol{\mu}\|_\infty^2\sqrt{|\mathcal{X}_0|}}$, and Γ denotes the gamma function $\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt$ for any $z > 0$.

Proof Fix $\epsilon > 0$. For all $T \geq M$, we have

$$\mathbb{P}_{\boldsymbol{\mu}}[(\mathcal{E}_{1,\epsilon}(T) \cap \mathcal{E}_{2,\epsilon}(T))^c] \leq \mathbb{P}_{\boldsymbol{\mu}}[\mathcal{E}_{1,\epsilon}(T)^c] + \mathbb{P}_{\boldsymbol{\mu}}[\mathcal{E}_{2,\epsilon}(T)^c].$$

Bounding $\mathbb{P}_{\boldsymbol{\mu}}[\mathcal{E}_{1,\epsilon}(T)^c]$: This is done by using Lemma 12 with $\boldsymbol{v} = \hat{\boldsymbol{\omega}}(t)$ and $(\eta, n) = (\eta_t, n_t)$. Before applying Lemma 12, we verify that our chosen ρ_t in (10) satisfies the assumption of Lemma 12:

$$\frac{(\min_{k\in[K]} \hat{\omega}_k(t) - \eta_t)^2}{D^2} \geq \frac{1}{D^2} \left(\frac{1}{2\sqrt{t|\mathcal{X}_0|}} - \frac{1}{4\sqrt{t|\mathcal{X}_0|}} \right)^2 = \frac{1}{16tD^2|\mathcal{X}_0|} = \rho_t,$$

where the inequality is because of $\min_{k\in[K]} \hat{\omega}_k(t) \geq \frac{1}{2\sqrt{t|\mathcal{X}_0|}}$ (Lemma 14) and that $\eta_t = \frac{1}{4\sqrt{t|\mathcal{X}_0|}}$ in (10). Then, applying Lemma 12 with $\boldsymbol{v} = \hat{\boldsymbol{\omega}}(t)$, $(v, \eta, n) = (\frac{1}{2\sqrt{t|\mathcal{X}_0|}}, \frac{1}{4\sqrt{t|\mathcal{X}_0|}}, \lceil t^{\frac{1}{4}} \rceil)$ yields that:

$$\mathbb{P}_{\boldsymbol{\mu}}[\mathcal{E}_{1,\epsilon}(T)^c] \leq \sum_{t=\underline{h}(T)}^{T} K \left(2\exp\left(-\frac{\epsilon^2\sqrt{t}}{8\ell^2K^3D^2} \right) + \exp\left(-\sqrt{t}\right) \right) \leq \sum_{t=\underline{h}(T)}^{T} 3K \exp\left(-\sqrt{t}A_1(\epsilon)\right),$$

where $A_1(\epsilon) = \min\{1, \frac{\epsilon^2}{8\ell^2K^3D^2}\}$.

Bounding $\mathbb{P}_{\boldsymbol{\mu}}[\mathcal{E}_{2,\epsilon}(T)^c]$: As Lemma 14 provides a lower bound on the number of pulls, $\min_{k\in[K]} N_k(t) \geq \frac{1}{2}\sqrt{\frac{t}{|\mathcal{X}_0|}}$, for all arms, using this lower bound of $N_k(t)$ as the number of i.i.d. samples in the application of Chernoff bound leads to:

$$\mathbb{P}_{\boldsymbol{\mu}}\left[|\hat{\mu}_k(t) - \mu_k| \geq \frac{\epsilon}{24D^3\|\boldsymbol{\mu}\|_\infty} \right] \leq 2\exp\left(-\sqrt{t}A_2(\epsilon)\right).$$

Hence, $\mathbb{P}_{\boldsymbol{\mu}}[\mathcal{E}_{2,\epsilon}(T)^c] \leq 2K\sum_{t=\underline{h}(T)}^{T} \exp\left(-\sqrt{t}A_2(\epsilon)\right)$. Then, we have

$$\begin{aligned}\sum_{T=M+1}^{\infty} \mathbb{P}_{\boldsymbol{\mu}}[(\mathcal{E}_{1,\epsilon}(T) \cap \mathcal{E}_{2,\epsilon}(T))^c] &\leq \int_{M+1}^{\infty}\int_{T^a}^{\infty} \left(3Ke^{-\sqrt{t}A_1(\epsilon)} + 2Ke^{-\sqrt{t}A_2(\epsilon)}\right) dt\, dT \\ &\leq 2K \left(\frac{3}{A_1(\epsilon)^{2+\frac{2}{a}}} + \frac{2}{A_2(\epsilon)^{2+\frac{2}{a}}} \right) \Gamma\left(2 + \frac{2}{a}\right),\end{aligned}$$

where the second inequality uses Lemma 15. □

F.1 Lemmas for bounding $\mathbb{P}_{\boldsymbol{\mu}}[\mathcal{E}_{1,\epsilon}(T)^c]$

The following lemma is a result of two concentration inequalities, one bounds how much the empirical average deviates from the expectation (Proposition 3), and the other bounds the error incurred by MCP (Lemma 13).

Lemma 12. *Let $(\boldsymbol{\pi}, \boldsymbol{\omega}, \theta) \in \Lambda \times \Sigma_+ \times (0,1)$, $v \in (0, \min_{k\in[K]} \omega_k)$, and $\eta \in (0, v)$. Then, $\forall \epsilon \in (0, 4K(v-\eta)/D)$,*

$$\mathbb{P}\left[\left\langle \nabla \bar{F}_{\boldsymbol{\pi},\eta}(\boldsymbol{\omega}), \tilde{\boldsymbol{x}}^\star - \boldsymbol{\omega} \right\rangle \geq \max_{\boldsymbol{x}\in\mathcal{X}} \left\langle \nabla \bar{F}_{\boldsymbol{\pi},\eta}(\boldsymbol{\omega}), \boldsymbol{x} - \boldsymbol{\omega} \right\rangle - \epsilon\right] \geq 1 - K\left(2\exp\left(-\frac{\epsilon^2 n^2}{8\ell^2 K^3 D^2}\right) + n\theta\right),$$

where $\nabla \tilde{F}_{\boldsymbol{\pi},\eta,n}(\boldsymbol{\omega})$ is computed by $\left(\frac{(v-\eta)^2}{D^2}, \theta\right)$-MCP, and $\tilde{\boldsymbol{x}}^\star \in \operatorname{argmax}_{\boldsymbol{x}\in\mathcal{X}} \left\langle \nabla \tilde{F}_{\boldsymbol{\pi},\eta,n}(\boldsymbol{\omega}), \boldsymbol{x}\right\rangle$.

Proof Let $\boldsymbol{x}^\star \in \operatorname{argmax}_{\boldsymbol{x}\in\mathcal{X}} \left\langle \nabla \bar{F}_{\boldsymbol{\pi},\eta}(\boldsymbol{\omega}), \boldsymbol{x}\right\rangle$. From $\tilde{\boldsymbol{x}}^\star \in \operatorname{argmax}_{\boldsymbol{x}\in\mathcal{X}} \left\langle \nabla \tilde{F}_{\boldsymbol{\pi},\eta,n}(\boldsymbol{\omega}), \boldsymbol{x}\right\rangle$,

$$\begin{aligned}\left\langle \nabla \bar{F}_{\boldsymbol{\pi},\eta}(\boldsymbol{\omega}), \boldsymbol{x}^\star - \tilde{\boldsymbol{x}}^\star \right\rangle &\leq \left\langle \nabla \bar{F}_{\boldsymbol{\pi},\eta}(\boldsymbol{\omega}), \boldsymbol{x}^\star - \tilde{\boldsymbol{x}}^\star \right\rangle + \left\langle \nabla \tilde{F}_{\boldsymbol{\pi},\eta}(\boldsymbol{\omega}), \tilde{\boldsymbol{x}}^\star - \boldsymbol{x}^\star \right\rangle \\ &= \left\langle \nabla \bar{F}_{\boldsymbol{\pi},\eta}(\boldsymbol{\omega}) - \nabla \tilde{F}_{\boldsymbol{\pi},\eta,n}(\boldsymbol{\omega}), \boldsymbol{x}^\star - \tilde{\boldsymbol{x}}^\star \right\rangle.\end{aligned}$$

Fix $\epsilon > 0$. Recall that $\nabla \tilde{F}_{\boldsymbol{\pi},\eta,n}(\boldsymbol{\omega}) = \frac{1}{n}\sum_{m=1}^n \nabla_{\boldsymbol{\omega}} f_{\hat{\boldsymbol{x}}_m}(\boldsymbol{\omega} + \eta \boldsymbol{\mathcal{Z}}_m, \boldsymbol{\pi})$ where each $\hat{\boldsymbol{x}}_m$ is computed by $(\frac{(v-\eta)^2}{D^2}, \theta)$-MCP$(\boldsymbol{\omega} + \eta\boldsymbol{\mathcal{Z}}_m, \boldsymbol{\pi})$, and each $\boldsymbol{\mathcal{Z}}_m$ is independently sampled from Uniform(B_2). Now, consider any fixed $\boldsymbol{x} = \boldsymbol{e}_k$ for any $k \in [K]$. Invoking Proposition 3 with $\epsilon = \frac{\epsilon}{4K}$ and $\boldsymbol{x} = \boldsymbol{e}_k$, we get:

$$\mathbb{P}\left[\left|\left\langle \nabla \bar{F}_{\boldsymbol{\pi},\eta}(\boldsymbol{\omega}) - \frac{1}{n}\sum_{m=1}^n \nabla F_{\boldsymbol{\pi}}(\boldsymbol{\omega} + \eta\boldsymbol{\mathcal{Z}}_m), \boldsymbol{e}_k \right\rangle\right| \geq \frac{\epsilon}{4K}\right] \leq 2\exp\left(-\frac{\epsilon^2 n^2}{8\ell^2 K^3 D^2}\right).$$

Also, for $\nabla \tilde{F}_{\boldsymbol{\pi},\eta,n}(\boldsymbol{\omega})$ computed by the $((v-\eta)^2/D^2, \theta)$-MCP algorithm, Lemma 13 with $\boldsymbol{x} = \boldsymbol{e}_k$, and $\theta = \theta$ and the assumption that $\epsilon \in (0, 4K(v-\eta)/D)$ implies that:

$$\mathbb{P}\left[\left|\left\langle \frac{1}{n}\sum_{m=1}^n \nabla F_{\boldsymbol{\pi}}(\boldsymbol{\omega} + \eta\boldsymbol{\mathcal{Z}}_m) - \nabla \tilde{F}_{\boldsymbol{\pi},\eta,n}(\boldsymbol{\omega}), \boldsymbol{e}_k \right\rangle\right| \geq \frac{\epsilon}{4K}\right] \leq n\theta.$$

Combining the two inequalities leads to:

$$\mathbb{P}\left[\left|\left\langle \nabla \bar{F}_{\boldsymbol{\pi},\eta}(\boldsymbol{\omega}) - \nabla \tilde{F}_{\boldsymbol{\pi},\eta,n}(\boldsymbol{\omega}), \boldsymbol{e}_k \right\rangle\right| \leq \frac{\epsilon}{2K}\right] \geq 1 - \left(2\exp\left(-\frac{\epsilon^2 n^2}{8\ell^2 K^3 D^2}\right) + n\theta\right).$$

Then, an application of a union bound over all $\{\boldsymbol{e}_k\}_{k\in[K]}$ gives

$$\mathbb{P}\left[\left\langle \nabla \bar{F}_{\boldsymbol{\pi},\eta}(\boldsymbol{\omega}) - \nabla \tilde{F}_{\boldsymbol{\pi},\eta,n}(\boldsymbol{\omega}), \boldsymbol{x}^\star - \tilde{\boldsymbol{x}}^\star \right\rangle \leq \epsilon\right] \geq 1 - K\left(2\exp\left(-\frac{\epsilon^2 n^2}{8\ell^2 K^3 D^2}\right) + n\theta\right). \quad (46)$$

Observe $\left\langle -\nabla \tilde{F}_{\boldsymbol{\pi},\eta,n}(\boldsymbol{\omega}), \boldsymbol{x}^\star - \tilde{\boldsymbol{x}}^\star \right\rangle \geq 0$ implies

$$\left\{\left\langle \nabla \bar{F}_{\boldsymbol{\pi},\eta}(\boldsymbol{\omega}) - \nabla \tilde{F}_{\boldsymbol{\pi},\eta,n}(\boldsymbol{\omega}), \boldsymbol{x}^\star - \tilde{\boldsymbol{x}}^\star \right\rangle \leq \epsilon\right\} \subseteq \left\{\left\langle \nabla \bar{F}_{\boldsymbol{\pi},\eta}(\boldsymbol{\omega}), \boldsymbol{x}^\star - \tilde{\boldsymbol{x}}^\star \right\rangle \leq \epsilon\right\}. \quad (47)$$

From (46)-(47), we conclude that the r.h.s. of (47) happens with probability at least $1 - K\left(2\exp\left(-\frac{\epsilon^2 n^2}{8\ell^2 K^3 D^2}\right) + n\theta\right)$. The proof is completed by simply rearranging the r.h.s. of (47). $\square$

Lemma 13. *Let $(\boldsymbol{\pi}, \boldsymbol{\omega}, \boldsymbol{x}, \theta) \in \Lambda \times \Sigma_+ \times \{0,1\}^K \times (0,1)$ with $\|\boldsymbol{x}\|_1 \leq D$ and $v \in (0, \min_{k\in[K]} \omega_k)$.*

$$\forall (\eta, \boldsymbol{z}) \in (0, v) \times B_2, \quad \mathbb{P}\left[\left|\left\langle \nabla_{\boldsymbol{\omega}} f_{\boldsymbol{x}_\star}(\boldsymbol{\omega} + \eta\boldsymbol{z}) - \nabla_{\boldsymbol{\omega}} f_{\hat{\boldsymbol{x}}}(\boldsymbol{\omega} + \eta\boldsymbol{z}), \boldsymbol{x}\right\rangle\right| \leq \frac{v-\eta}{D}\right] \geq 1 - \theta,$$

where $\boldsymbol{x}_\star$ is some action satisfying $f_{\boldsymbol{x}_\star}(\boldsymbol{\omega} + \eta\boldsymbol{z}) = F_{\boldsymbol{\pi}}(\boldsymbol{\omega} + \eta\boldsymbol{z})$, and $\hat{\boldsymbol{x}}$ is the returned action of $((v-\eta)^2/D^2, \theta)$-MCP$(\boldsymbol{\omega} + \eta\boldsymbol{z}, \boldsymbol{\pi})$.

Proof This basically follows from a direct calculation. Let $\epsilon > 0$ and fix any $(\boldsymbol{\pi}, \boldsymbol{\omega}, \boldsymbol{x}) \in \Lambda \times \Sigma_+ \times \{0,1\}^K$, $\|\boldsymbol{x}\|_1 \le D$, and any $(\eta, \boldsymbol{z}) \in (0, v) \times B_2$. Then, for $\hat{\boldsymbol{x}}$ computed by (ρ, θ)-$\texttt{MCP}(\boldsymbol{\omega} + \eta \boldsymbol{z}, \boldsymbol{\pi})$ with $\rho = (v - \eta)^2/D^2$, we have with probability at least $1 - \theta$

$$\begin{aligned}\rho &\ge |\langle \nabla_{\boldsymbol{\omega}} f_{\boldsymbol{x}_\star}(\boldsymbol{\omega} + \eta \boldsymbol{z}) - \nabla_{\boldsymbol{\omega}} f_{\hat{\boldsymbol{x}}}(\boldsymbol{\omega} + \eta \boldsymbol{z}), \boldsymbol{\omega} + \eta \boldsymbol{z} \rangle| \\ &\ge \min_{k \in [K]} (\boldsymbol{\omega} + \eta \boldsymbol{z})_k \left\| \nabla_{\boldsymbol{\omega}} f_{\boldsymbol{x}_\star}(\boldsymbol{\omega} + \eta \boldsymbol{z}) - \nabla_{\boldsymbol{\omega}} f_{\hat{\boldsymbol{x}}}(\boldsymbol{\omega} + \eta \boldsymbol{z}) \right\|_\infty .\end{aligned}$$

Hence, remarking that $\min_{k \in [K]} (\boldsymbol{\omega} + \eta \boldsymbol{z})_k \ge v - \eta > 0$, we get: with probability at least $1 - \theta$,

$$|\langle \nabla_{\boldsymbol{\omega}} f_{\boldsymbol{x}_\star}(\boldsymbol{\omega} + \eta \boldsymbol{z}) - \nabla_{\boldsymbol{\omega}} f_{\hat{\boldsymbol{x}}}(\boldsymbol{\omega} + \eta \boldsymbol{z}), \boldsymbol{x} \rangle| \le \frac{\rho D}{v - \eta} = \frac{v - \eta}{D},$$

where we used the fact that $\|\boldsymbol{x}\|_1 \le D$ and Hölder's inequality. □

Proposition 3. *Let $(\boldsymbol{\pi}, \boldsymbol{\omega}, \boldsymbol{x}) \in \Lambda \times \Sigma_+ \times \{0,1\}^K$, $\eta \in (0, \min_{k \in [K]} \omega_k)$, and $\|\boldsymbol{x}\|_1 \le D$. Then,*

$$\forall \epsilon > 0, \quad \mathbb{P}\left[\left| \left\langle \nabla \bar{F}_{\boldsymbol{\pi}, \eta}(\boldsymbol{\omega}) - \frac{1}{n} \sum_{m=1}^n \nabla F_{\boldsymbol{\pi}}(\boldsymbol{\omega} + \eta \boldsymbol{\mathcal{Z}}_m), \boldsymbol{x} \right\rangle \right| \ge \epsilon \right] \le 2 \exp\left(- \frac{2\epsilon^2 n^2}{\ell^2 K D^2} \right),$$

where $\boldsymbol{\mathcal{Z}}_1, \cdots, \boldsymbol{\mathcal{Z}}_n$ are independently sampled from Uniform(B_2).

Proof Fix $(\boldsymbol{\pi}, \boldsymbol{\omega}, \boldsymbol{x}) \in \Lambda \times \Sigma_+ \times \{0,1\}^K$ where $\|\boldsymbol{x}\|_1 \le D$, and fix $\eta \in (0, \min_{k \in [K]} \omega_k)$. Define

$$\phi(\boldsymbol{z}_1, \cdots, \boldsymbol{z}_t) = \left\langle \nabla \bar{F}_{\boldsymbol{\pi}, \eta}(\boldsymbol{\omega}) - \frac{1}{n} \sum_{m=1}^n \nabla F_{\boldsymbol{\pi}}(\boldsymbol{\omega} + \eta \boldsymbol{z}_m), \boldsymbol{x} \right\rangle.$$

Note that $\mathbb{E}_{\boldsymbol{\mathcal{Z}}_1, \cdots, \boldsymbol{\mathcal{Z}}_n}[\phi(\boldsymbol{\mathcal{Z}}_1, \cdots, \boldsymbol{\mathcal{Z}}_n)] = 0$ by definition. Now we also observe that:

$$\max_{\boldsymbol{z}_1, \cdots, \boldsymbol{z}_n, \boldsymbol{z}' \in B_2, m \in [n]} |\phi(\boldsymbol{z}_1, \cdots, \boldsymbol{z}_n) - \phi(\boldsymbol{z}_1, \cdots, \boldsymbol{z}_{m-1}, \boldsymbol{z}', \boldsymbol{z}_{m+1}, \cdots, \boldsymbol{z}_n)| \le \frac{\ell D}{n}$$

due to the ℓ-Lipschitzness of $F_{\hat{\boldsymbol{\mu}}}$ (Lemma 21 in Appendix I) and $\max_{\boldsymbol{x} \in \mathcal{X}} \|\boldsymbol{x}\|_1 \le D$. Hence it follows from McDiarmid's inequality (Lemma 16 in F.3) that

$$\forall \epsilon > 0, \quad \mathbb{P}[|\phi(\boldsymbol{\mathcal{Z}}_1, \cdots, \boldsymbol{\mathcal{Z}}_n)| \ge \epsilon] \le 2 \exp\left(- \frac{2\epsilon^2}{K(\frac{\ell D}{n})^2} \right) = 2 \exp\left(- \frac{2\epsilon^2 n^2}{\ell^2 K D^2} \right).$$

□

F.2 Lemmas for bounding $\mathbb{P}_{\boldsymbol{\mu}}[\mathcal{E}_{2,\epsilon}(T)^c]$

Lemma 14 (forced exploration)**.** *Let $\mathcal{X}_0 \subseteq \mathcal{X}$ be any set covering all arms $[K]$ and $t \ge 4|\mathcal{X}_0|$. Any algorithm with forced-exploration procedure satisfies*

$$\hat{\boldsymbol{\omega}}(t) \in \Sigma_{\sqrt{\frac{1}{t|\mathcal{X}_0|}} - \frac{1}{t}} \subset \Sigma_{\frac{1}{2}\sqrt{\frac{1}{t|\mathcal{X}_0|}}}, \quad \forall t \ge 4|\mathcal{X}_0|.$$

Proof Fix any $k \in [K]$. By merely counting the rounds before t performing forced exploration,

$$N_k(t) \ge \sum_{s \in [t] : \lfloor \sqrt{\frac{s}{|\mathcal{X}_0|}} \rfloor \in \mathbb{N}} \sum_{\boldsymbol{x} \in \mathcal{X}_0} \boldsymbol{x}_k \ge \sqrt{\frac{t}{|\mathcal{X}_0|}} - 1 \ge \frac{1}{2} \sqrt{\frac{t}{|\mathcal{X}_0|}},$$

where the last inequality holds for any $t \ge 4|\mathcal{X}_0|$. □

F.3 Technical lemmas

Lemma 15 ([WTP21])**.** *Let $\alpha \in (0,1)$ and $A, \beta > 0$. Then,*

$$\int_0^\infty \left(\int_{T^\alpha}^\infty e^{-At^\beta} dt \right) dT = \frac{\Gamma(\frac{1}{\alpha\beta} + \frac{1}{\beta})}{\beta A^{\frac{1}{\alpha\beta} + \frac{1}{\beta}}}.$$

Proof The result of Lemma 5 [WTP21] is stated for $\alpha, \beta \in (0,1)$ but it actually applies for the case of $\beta > 0$ as well. Here we provide a proof for completeness.

$$\int_0^\infty \left(\int_{T^\alpha}^\infty e^{-At^\beta} dt \right) dT = \int_0^\infty \alpha T^\alpha e^{-AT^{\alpha\beta}} dT = \frac{1}{\beta} \int_0^\infty x^{\frac{1}{\alpha\beta} + \frac{1}{\beta} - 1} e^{-Ax} dx = \frac{\Gamma(\frac{1}{\alpha\beta} + \frac{1}{\beta})}{\beta A^{\frac{1}{\alpha\beta} + \frac{1}{\beta}}}.$$

□

The below Lemma 16, also known as bounded different inequality, can be found in many textbooks, e.g., Theorem 6.2 in [BLM13].

Lemma 16 (McDiarmid's inequality)**.** *Let $\boldsymbol{\mathcal{Z}} = (\mathcal{Z}_1, \cdots, \mathcal{Z}_n)$ be independent random variables, and $\phi : \mathbb{R}^n \mapsto \mathbb{R}$ be a measurable function. Suppose $\phi(\mathbf{z})$ changes by at most $c_i > 0$ under an arbitrary change of the i-th coordinate. Then,*

$$\forall \epsilon > 0, \quad \mathbb{P}[\phi(\boldsymbol{\mathcal{Z}}) - \mathbb{E}[\phi(\boldsymbol{\mathcal{Z}})] \geq \epsilon] \leq \exp\left(-\frac{2\epsilon^2}{\sum_{i=1}^n c_i^2} \right).$$

G Continuity arguments

In this section, we establish the continuity of $F_{\boldsymbol{\pi}}(\boldsymbol{\omega})$ and $\nabla \bar{F}_{\boldsymbol{\pi},\eta}(\boldsymbol{\omega})$ in $\boldsymbol{\pi}$ for any fixed $\boldsymbol{\omega} \in \Sigma_+$, where $\nabla \bar{F}_{\boldsymbol{\pi},\eta}(\boldsymbol{\omega})$ denotes the gradient $\nabla_{\boldsymbol{\omega}} \bar{F}_{\boldsymbol{\pi},\eta}(\boldsymbol{\omega})$ taken w.r.t. the input $\boldsymbol{\omega}$. As the consequence of the continuity of $F_{\boldsymbol{\pi}}$ and $\nabla \bar{F}_{\boldsymbol{\pi},\eta}$ in $\boldsymbol{\pi}$, we can show the point-wise convergence of $F_{\hat{\boldsymbol{\mu}}(t)} \to F_{\boldsymbol{\mu}}$ and $\nabla \bar{F}_{\hat{\boldsymbol{\mu}}(t),\eta} \to \nabla \bar{F}_{\boldsymbol{\mu},\eta}$ given that $\hat{\boldsymbol{\mu}}(t) \to \boldsymbol{\mu}$ almost surely.

Notation. Throughout this section, we define $\nabla \bar{F}_{\boldsymbol{\pi},\eta}(\boldsymbol{\omega}) = \mathbf{0}_K$ if $\eta \geq \min_{k\in[K]} \omega_k$ for any $(\boldsymbol{\pi}, \boldsymbol{\omega}) \in \Lambda \times \Sigma_+$. Moreover, for any $(\boldsymbol{v}, \boldsymbol{\omega}) \in \mathbb{R}^K \times \Sigma_+$, we will use $\nabla_{\boldsymbol{\pi}} F_{\boldsymbol{v}}(\boldsymbol{\omega})$ (resp. $\nabla_{\boldsymbol{\pi}}(\frac{\partial \bar{F}_{\boldsymbol{v},\eta}(\boldsymbol{\omega})}{\partial \omega_k})$) to denote the gradient of the function $\boldsymbol{\pi} \mapsto F_{\boldsymbol{\pi}}(\boldsymbol{\omega})$ (resp. $\boldsymbol{\pi} \mapsto \frac{\partial \bar{F}_{\boldsymbol{\pi},\eta}(\boldsymbol{\omega})}{\partial \omega_k}$) evaluated at the point $\boldsymbol{v}$.

The main result in this section, Lemma 7, is derived based on Lemma 17 in Appendix G.1 (which asserts the continuity of the function $\psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}(\boldsymbol{\pi}) = \langle \nabla \bar{F}_{\boldsymbol{\pi},\eta}(\boldsymbol{\omega}), \boldsymbol{x} - \boldsymbol{\omega} \rangle$ on $\mathbb{R}^K$) and Proposition 4 in Appendix G.2 (which upper bounds the length of $\nabla f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{\mu})$).

Lemma 7. *Let $\boldsymbol{\mu} \in \Lambda$ and $\epsilon \in (0, \frac{2D^2 \triangle_{\min}(\boldsymbol{\mu})^2}{K})$. Then, any $\boldsymbol{\pi} \in \mathbb{R}^K$ with $\|\boldsymbol{\pi} - \boldsymbol{\mu}\|_\infty < \frac{\epsilon}{24D^3 \|\boldsymbol{\mu}\|_\infty}$ satisfies the following:*

$$|F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) - F_{\boldsymbol{\pi}}(\boldsymbol{\omega})| < \epsilon, \quad \forall \boldsymbol{\omega} \in \Sigma_+ \tag{35}$$

$$|\langle \nabla \bar{F}_{\boldsymbol{\pi},\eta}(\boldsymbol{\omega}) - \nabla \bar{F}_{\boldsymbol{\mu},\eta}(\boldsymbol{\omega}), \boldsymbol{x} - \boldsymbol{\omega} \rangle| < \epsilon, \quad \forall (\boldsymbol{\omega}, \boldsymbol{x}) \in \Sigma_+ \times \mathcal{X}, \forall \eta \in (0, \min_{k\in[K]} \omega_k). \tag{36}$$

Proof Inspired by Lemma 14 in [WTP21], we prove this lemma using Proposition 4 and applying the mean-value theorem to $\psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}$.

Fix $(\boldsymbol{\omega}, \boldsymbol{\mu}) \in \Sigma_+ \times \Lambda$, and let $\boldsymbol{i}^\star = \boldsymbol{i}^\star(\boldsymbol{\mu})$ and $\triangle_{\boldsymbol{x}} = \triangle_{\boldsymbol{x}}(\boldsymbol{\mu})$ for any $\boldsymbol{x} \in \mathcal{X} \backslash \{\boldsymbol{i}^\star\}$. Fix $\epsilon \in (0, \frac{2D^2 \triangle_{\min}^2}{K})$ and $\boldsymbol{\pi} \in \mathbb{R}^K$ such that $\|\boldsymbol{\pi} - \boldsymbol{\mu}\|_\infty < \frac{\epsilon}{24D^3 \|\boldsymbol{\mu}\|_\infty}$. One may check that this $\boldsymbol{\pi}$ satisfies the assumption of Proposition 4 as

$$\|\boldsymbol{\pi} - \boldsymbol{\mu}\|_\infty < \frac{\epsilon}{24D^3 \|\boldsymbol{\mu}\|_\infty} < \frac{2D^2 \triangle_{\min}^2}{24KD^3 \|\boldsymbol{\mu}\|_\infty} = \frac{\triangle_{\min}^2}{12KD \|\boldsymbol{\mu}\|_\infty} \leq \frac{\triangle_{\min}}{6K} < \frac{\triangle_{\min}}{\sqrt{2KD}},$$

where the second inequality stems from the choice of ϵ and the second last is because $\triangle_{\min} \leq 2D \|\boldsymbol{\mu}\|_\infty$. In what follows, we will be applying the mean-value theorem to $\psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}$ (whose continuity is stated in Lemma 17). For convenience, introduce the function $\boldsymbol{r}(\beta) = (1 - \beta)\boldsymbol{\mu} + \beta \boldsymbol{\pi}$ for any $\beta \in (0, 1)$.

Proof of (35): For any $\boldsymbol{x} \in \mathcal{X} \backslash \{\boldsymbol{i}^\star\}$, by the mean-value theorem, there exists a $\beta \in (0, 1)$ such that

$$\begin{aligned}
|f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{\pi}) - f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{\mu})| &= |\langle \nabla_{\boldsymbol{\pi}} f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{r}(\beta)), \boldsymbol{\pi} - \boldsymbol{\mu} \rangle| \\
&= \left| \sum_{k\in[K]} \omega_k \left\langle \nabla_{\boldsymbol{\pi}} \left(\frac{\partial f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{r}(\beta))}{\partial \omega_k} \right), \boldsymbol{\pi} - \boldsymbol{\mu} \right\rangle \right| \\
&\leq \sum_{k\in[K]} \omega_k \left\| \nabla_{\boldsymbol{\pi}} \left(\frac{\partial f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{r}(\beta))}{\partial \omega_k} \right) \right\|_1 \|\boldsymbol{\pi} - \boldsymbol{\mu}\|_\infty < \epsilon,
\end{aligned} \tag{48}$$

where the last inequality uses $\boldsymbol{\omega} \in \Sigma_+$, $\|\boldsymbol{\pi} - \boldsymbol{\mu}\|_\infty < \frac{\epsilon}{24D^3 \|\boldsymbol{\mu}\|_\infty}$, $\left\| \nabla_{\boldsymbol{\pi}} \left(\frac{\partial f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{r}(\beta))}{\partial \omega_k} \right) \right\|_1 \leq 12D^2 \|\boldsymbol{\mu}\|_\infty$ (Proposition 4). Hence, from a substitution of $\boldsymbol{x}$ in (48) with $\boldsymbol{x}_e \in \operatorname{argmin}_{\boldsymbol{x} \neq \boldsymbol{i}^\star} f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{\mu})$ and the fact that $F_{\boldsymbol{\pi}}(\boldsymbol{\omega}) \leq f_{\boldsymbol{x}_e}(\boldsymbol{\omega}, \boldsymbol{\pi})$, we derive

$$F_{\boldsymbol{\pi}}(\boldsymbol{\omega}) - F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) \leq f_{\boldsymbol{x}_e}(\boldsymbol{\omega}, \boldsymbol{\pi}) - f_{\boldsymbol{x}_e}(\boldsymbol{\omega}, \boldsymbol{\mu}) < \epsilon.$$

The other inequality of $F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) - F_{\boldsymbol{\pi}}(\boldsymbol{\omega}) < \epsilon$ can be derived similarly. This proves (35).

Proof of (36): Recall that $\psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}(\boldsymbol{\pi}) = \langle \nabla \bar{F}_{\boldsymbol{\pi},\eta}(\boldsymbol{\omega}), \boldsymbol{x} - \boldsymbol{\omega} \rangle$ is continuous on $\mathbb{R}^K$ (Lemma 17). By the mean-value theorem, there exists $\beta \in (0, 1)$ such that

$$\begin{aligned}
|\psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}(\boldsymbol{\pi}) - \psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}(\boldsymbol{\mu})| &= |\langle \nabla_{\boldsymbol{\pi}} \psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}(\boldsymbol{r}(\beta)), \boldsymbol{\pi} - \boldsymbol{\mu} \rangle| \\
&\leq \|\nabla_{\boldsymbol{\pi}} \psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}(\boldsymbol{r}(\beta))\|_1 \|\boldsymbol{\pi} - \boldsymbol{\mu}\|_\infty .
\end{aligned} \tag{49}$$

To bound $\|\nabla_{\boldsymbol{\pi}}\psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}(\boldsymbol{r}(\beta))\|_1$, we write

$$\nabla_{\boldsymbol{\pi}}\psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}(\boldsymbol{r}(\beta)) = \sum_{k\in[K]} \nabla_{\boldsymbol{\pi}}\left(\frac{\partial \bar{F}_{\boldsymbol{r}(\beta),\eta}(\boldsymbol{\omega})}{\partial \omega_k}\right)(x_k - \omega_k).$$

Then it follows from the fundamental theorem of calculus that: the gradient $\nabla_{\boldsymbol{\pi}}$ and the expectation operators are exchangeable, i.e.,

$$\forall k \in [K], \quad \nabla_{\boldsymbol{\pi}}\left(\frac{\partial \bar{F}_{\boldsymbol{r}(\beta),\eta}(\boldsymbol{\omega})}{\partial \omega_k}\right) = \mathbb{E}_{\boldsymbol{\mathcal{Z}}\sim\text{Uniform}(B_2)}\left[\nabla_{\boldsymbol{\pi}}\left(\frac{\partial F_{\boldsymbol{r}(\beta)}(\boldsymbol{\omega}+\eta\boldsymbol{\mathcal{Z}})}{\partial \omega_k}\right)\right].$$

As shown in Appendix H, $\frac{\partial F_{\boldsymbol{r}(\beta)}(\boldsymbol{\omega}+\eta\boldsymbol{\mathcal{Z}})}{\partial \omega_k}$ exists almost surely. When such gradient exists, Proposition 4 bounds its 1-norm length by

$$\left\|\nabla_{\boldsymbol{\pi}}\left(\frac{\partial F_{\boldsymbol{r}(\beta)}(\boldsymbol{\omega}+\eta\boldsymbol{\mathcal{Z}})}{\partial \omega_k}\right)\right\|_1 \leq 12D^2\|\boldsymbol{\mu}\|_\infty,$$

so it follows that $\left\|\nabla_{\boldsymbol{\pi}}\left(\frac{\partial \bar{F}_{\boldsymbol{r}(\beta),\eta}(\boldsymbol{\omega})}{\partial \omega_k}\right)\right\|_1 \leq 12D^2\|\boldsymbol{\mu}\|_\infty$ as well. Hence, substituting the above back to $\nabla_{\boldsymbol{\pi}}\psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}(\boldsymbol{r}(\beta))$ yields:

$$\|\nabla_{\boldsymbol{\pi}}\psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}(\boldsymbol{r}(\beta))\|_1 \leq \max_{k\in[K]}\left\|\nabla_{\boldsymbol{\pi}}\left(\frac{\partial \bar{F}_{\boldsymbol{r}(\beta),\eta}(\boldsymbol{\omega})}{\partial \omega_k}\right)\right\|_1 \|\boldsymbol{x}-\boldsymbol{\omega}\|_1 \leq 24D^3\|\boldsymbol{\mu}\|_\infty,$$

where the first inequality use Hölder's inequality. Finally, plugging the above into (49) and recalling that $\|\boldsymbol{\pi}-\boldsymbol{\mu}\|_\infty < \frac{\epsilon}{24D^3\|\boldsymbol{\mu}\|_\infty}$, we have

$$|\psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}(\boldsymbol{\pi}) - \psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}(\boldsymbol{\mu})| < \epsilon.$$

This concludes the proof. □

G.1 An application of the maximum theorem

Recall that $\psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}(\boldsymbol{\pi}) = \langle \nabla \bar{F}_{\boldsymbol{\pi},\eta}(\boldsymbol{\omega}), \boldsymbol{x}-\boldsymbol{\omega}\rangle$.

Lemma 17. *For any $\epsilon > 0$, there exists a constant $\xi_\epsilon > 0$ such that if $\|\boldsymbol{\pi}-\boldsymbol{\mu}\|_\infty < \xi_\epsilon$, then*

$$|\psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}(\boldsymbol{\pi}) - \psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}(\boldsymbol{\mu})| < \epsilon, \quad \forall(\boldsymbol{\omega},\boldsymbol{x}) \in \Sigma_+ \times \mathcal{X},\ \forall \eta \in (0, \min_{k\in[K]} \omega_k). \tag{50}$$

The proof of Lemma 17 replies on the celebrated maximum theorem [FKV14], which is introduced below. After that, we then show its proof.

Maximum Theorem: Here we briefly introduce the maximum theorem and Lemma 17 will be proved at the end of this section. The definitions and results are taken from [FKV14] (see also Appendix K.1 of [WTP21]).

Definition 1. *Let $U \neq \emptyset$ be a subset of a topological space and $h : U \mapsto \mathbb{R}$ be a function. Define the level sets of h for $y \in \mathbb{R}$ as*

$$L_h(y,U) = \{x \in U : h(x) \leq y\} \quad \textit{and} \quad L_h^<(y,U) = \{x \in U : h(x) < y\}.$$

The function h is said to be lower semi-continuous (resp. upper semi-continuous) on U if $L_h(y,U)$ are closed (resp. $L_h^<(y,U)$ are open) for all $y \in \mathbb{R}$; h is said to be inf-compact on U if $L_h(y,U)$ and $L_h^<(y,U)$ are compact for all $y \in \mathbb{R}$.

Definition 2. *Let $\mathbb{X}$ and $\mathbb{Y}$ be Hausdorff topological spaces and $\Phi : \mathbb{X} \rightrightarrows \mathbb{S}(\mathbb{Y})$ be a set-valued function, where $\mathbb{S}(\mathbb{Y})$ is the set of non-empty subsets of $\mathbb{Y}$. Define*

$$Gr_U(\Phi) = \{(x,y) \in U \times \mathbb{Y} : y \in \Phi(x)\}$$

as the graph of Φ restricted to U. The function $u : \mathbb{X}\times\mathbb{Y} \mapsto \mathbb{R}$ is said to be $\mathbb{K}$-inf-compact on $Gr_{\mathbb{X}}(\Phi)$ if for all non-empty compact subset C of $\mathbb{X}$, u is inf-compact on $Gr_C(\Phi)$.

Theorem 6 (Maximum theorem)**.** *Suppose $\mathbb{X}$ is compactly generated, $\Phi : \mathbb{X} \rightrightarrows \mathbb{S}(\mathbb{Y})$ is lower hemicontinuous, and $u : \mathbb{X}\times\mathbb{Y} \mapsto \mathbb{R}$ is $\mathbb{K}$-inf-compact and upper semi-continuous on $Gr_{\mathbb{X}}(\Phi)$. Then, the function $v(x) = \inf_{y\in\Phi(x)} u(x,y)$ is continuous and the set of its optimal solutions $\Phi^\star(x) = \{y \in \Phi(x) : u(x,y) = v(x)\}$ is upper hemicontinuous and compact-valued.*

Proof of Lemma 17: Fix any $\boldsymbol{\mu} \in \Lambda$ and let $\boldsymbol{i}^\star = \boldsymbol{i}^\star(\boldsymbol{\mu})$. The goal is to show that for any $\epsilon > 0$, there exists a constant $\xi_\epsilon > 0$ such that if $\|\boldsymbol{\pi} - \boldsymbol{\mu}\|_\infty < \xi_\epsilon$, then

$$|\psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}(\boldsymbol{\pi}) - \psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}(\boldsymbol{\mu})| < \epsilon, \quad \forall(\boldsymbol{\omega},\boldsymbol{x}) \in \Sigma_+ \times \mathcal{X}, \, \forall \eta \in (0, \min_{k\in[K]} \omega_k), \tag{50}$$

where $\psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}(\boldsymbol{\pi}) = \langle \nabla \bar{F}_{\boldsymbol{\pi},\eta}(\boldsymbol{\omega}), \boldsymbol{x} - \boldsymbol{\omega} \rangle$. In what follows, we will use p to denote the probability distribution of Uniform(B_2). We will first show that $\psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}$ is continuous for each fixed $(\boldsymbol{\omega},\boldsymbol{x},\eta) \in \Sigma_+ \times \mathcal{X} \times (0,1)$, and then use Theorem 6 to show (50).

Continuity of $\psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}$: Fix $(\boldsymbol{\omega},\boldsymbol{x},\eta) \in \Sigma_+ \times \mathcal{X} \times (0,1)$. Let $U_\eta = \{\boldsymbol{z} \in B_2 : |\partial F_{\boldsymbol{\mu}}(\boldsymbol{\omega} + \eta \boldsymbol{z})| > 1\}$ which is a measure-zero set under p (Lemma 20 in Appendix H). For its complement set $B_2 \backslash U_\eta$, we split $B_2 \backslash U_\eta = \cup_{\boldsymbol{y} \neq \boldsymbol{i}^\star} B_\eta(\boldsymbol{y})$ into possibly overlapping sets $B_\eta(\boldsymbol{y}) = \{\boldsymbol{z} \in B_2 \backslash U_\eta : \nabla F_{\boldsymbol{\pi},\eta}(\boldsymbol{\omega} + \eta\boldsymbol{z}) = \nabla_{\boldsymbol{\omega}} f_{\boldsymbol{y}}(\boldsymbol{\omega} + \eta \boldsymbol{z}, \boldsymbol{\pi})\}$, and define $\psi_{\boldsymbol{\omega},\boldsymbol{y},\eta}(\boldsymbol{y},\cdot) = \int_{\boldsymbol{z}\in B_\eta(\boldsymbol{y})} \langle \nabla_{\boldsymbol{\omega}} f_{\boldsymbol{y}}(\boldsymbol{\omega} + \eta \boldsymbol{z}, \cdot), \boldsymbol{x} - \boldsymbol{\omega} \rangle \, dp(\boldsymbol{z})$ on each of these sets $B_\eta(\boldsymbol{y})$. Observe that for any $\boldsymbol{\pi} \in \mathbb{R}^K$, we have

$$\psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}(\boldsymbol{\pi}) = \int_{\boldsymbol{z}\in B_2 \backslash U_\eta} \langle \nabla F_{\boldsymbol{\pi},\eta}(\boldsymbol{\omega} + \eta \boldsymbol{z}), \boldsymbol{x} - \boldsymbol{\omega} \rangle \, dp(\boldsymbol{z}) = \sum_{\boldsymbol{y} \neq \boldsymbol{i}^\star} \psi_{\boldsymbol{\omega},\boldsymbol{y},\eta}(\boldsymbol{y},\boldsymbol{\pi}).$$

To show the continuity of $\psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}(\boldsymbol{\pi})$, it suffices to show that each $\psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}(\boldsymbol{y},\cdot)$ is continuous. Fix $\boldsymbol{y} \in \mathcal{X} \backslash \{\boldsymbol{i}^\star\}$ and any sequence $\{\boldsymbol{\pi}_n\}_{n=1}^\infty$ converging to $\boldsymbol{\mu}$. Then, for any $\forall \boldsymbol{z} \in B_2$, we have

(i) $|\langle \nabla_{\boldsymbol{\omega}} f_{\boldsymbol{y}}(\boldsymbol{\omega} + \eta \boldsymbol{z}, \boldsymbol{\pi}_n), \boldsymbol{x} - \boldsymbol{\omega} \rangle| \leq \|\nabla_{\boldsymbol{\omega}} f_{\boldsymbol{y}}(\boldsymbol{\omega} + \eta \boldsymbol{z}, \boldsymbol{\pi}_n)\|_\infty \|\boldsymbol{x} - \boldsymbol{\omega}\|_1 \leq 2D\ell$

(ii) $\lim_{n\to\infty} \langle \nabla_{\boldsymbol{\omega}} f_{\boldsymbol{y}}(\boldsymbol{\omega} + \eta \boldsymbol{z}, \boldsymbol{\pi}_n), \boldsymbol{x} - \boldsymbol{\omega} \rangle = \langle \nabla_{\boldsymbol{\omega}} f_{\boldsymbol{y}}(\boldsymbol{\omega} + \eta \boldsymbol{z}, \boldsymbol{\mu}), \boldsymbol{x} - \boldsymbol{\omega} \rangle$. This is because $\nabla_{\boldsymbol{\omega}} f_{\boldsymbol{y}}(\boldsymbol{\omega} + \eta \boldsymbol{z}, \cdot) = \frac{\langle \boldsymbol{i}^\star - \boldsymbol{y}, \cdot \rangle^2 (\boldsymbol{i}^\star \oplus \boldsymbol{y}) \odot (\boldsymbol{\omega} + \eta \boldsymbol{z})^{-2}}{2\langle \boldsymbol{i}^\star \oplus \boldsymbol{y}, (\boldsymbol{\omega} + \eta \boldsymbol{z})^{-1} \rangle^2}$ by Lemma 19 and Proposition 1 (Appendix C.1) is obviously continuous and that function composition preserves continuity.

From (i) and (ii), the dominated convergence theorem implies that

$$\psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}(\boldsymbol{y},\boldsymbol{\mu}) = \lim_{n\to\infty} \int_{\boldsymbol{z}\in B_\eta(\boldsymbol{y})} \langle \nabla_{\boldsymbol{\omega}} f_{\boldsymbol{y}}(\boldsymbol{\omega} + \eta \boldsymbol{z}, \boldsymbol{\pi}_n), \boldsymbol{x} - \boldsymbol{\omega} \rangle \, dp(\boldsymbol{z}).$$

This shows the continuity of $\psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}(\boldsymbol{y},\cdot)$ for each $\boldsymbol{y} \neq \boldsymbol{i}^\star$, and thus $\psi_{\boldsymbol{\omega},\boldsymbol{y},\eta}$ is continuous.

Application of the maximum theorem (Theorem 6): For this part, we take the approach similar to Lemma 6 in [WTP21]. Define

$$\phi(\boldsymbol{\pi}) = \min \left\{ -|\psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}(\boldsymbol{\pi}) - \psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}(\boldsymbol{\mu})| : (\boldsymbol{\omega},\boldsymbol{x},\eta) \in \Sigma_+ \times \mathcal{X} \times (0,1) \right\}.$$

We prove the continuity of ϕ on $\mathcal{S} = \mathbb{R}^K \backslash \mathsf{cl}\,(\mathsf{Alt}(\boldsymbol{\mu}))$ by invoking Theorem 6 with the following substitutions:

- $\mathbb{X} = \mathcal{S}$,
- $\mathbb{Y} = \Sigma_+ \times \mathcal{X} \times (0,1)$,
- $\Phi = \Sigma_+ \times \mathcal{X} \times (0,1)$,
- $u(\boldsymbol{\pi},\boldsymbol{\omega},\boldsymbol{x},\eta) = -|\psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}(\boldsymbol{\pi}) - \psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}(\boldsymbol{\mu})|$.

Here we verify that the assumptions of Theorem 6 are satisfied. $\mathbb{X}$ is compactly generated as $\mathcal{S}$ is a metric space; Φ is continuous as it is a constant map; u is continuous due to the continuity of $\psi_{\boldsymbol{\omega},\boldsymbol{x},\eta}$. To show that u is $\mathbb{K}$-inf compact, consider any compact set $C \subset \mathcal{S}$ and any $y \in \mathbb{R}$. We see that $L_u(y, C \times \Sigma_+ \times \mathcal{X} \times (0,1))$ is compact because it is bounded (as $\Sigma_+ \times \mathcal{X} \times (0,1)$ is bounded and C is compact) and closed (as u is continuous and the preimage of $[0,y]$ is closed). Hence, ϕ is continuous on $\mathcal{S}$ by Theorem 6. Finally, by $\phi(\boldsymbol{\mu}) = 0$ and the continuity of ϕ, there exists $\xi_\epsilon > 0$ such that $\phi(\boldsymbol{\pi}) > -\epsilon$ for any $\|\boldsymbol{\pi} - \boldsymbol{\mu}\|_\infty < \xi_\epsilon$. This completes the proof of (50). $\square$

G.2 The length of gradients

Throughout this subsection, we fix $\boldsymbol{\mu} \in \Lambda$ and denote $\boldsymbol{i}^\star = \boldsymbol{i}^\star(\boldsymbol{\mu})$, $\triangle_{\boldsymbol{x}} = \triangle_{\boldsymbol{x}}(\boldsymbol{\mu})$, and $\triangle_{\min}(\boldsymbol{\mu}) = \triangle_{\min}$ for short. Here we aim to present Proposition 4, in which (i) quantifies how close an estimate $\boldsymbol{\pi}$ of $\boldsymbol{\mu}$ should be such that $\boldsymbol{i}^\star(\boldsymbol{\pi}) = \boldsymbol{i}^\star$, and (ii) asserts the continuity of any component of $\nabla_{\boldsymbol{\omega}} f_{\boldsymbol{x}}(\boldsymbol{\omega},\boldsymbol{\pi})$ in $\boldsymbol{\pi}$, and that its gradient with respective to $\boldsymbol{\pi}$ is bounded.

Proposition 4. *Any $\boldsymbol{\pi} \in \mathbb{R}^K$ such that $\|\boldsymbol{\pi} - \boldsymbol{\mu}\|_\infty < \frac{\triangle_{\min}}{\sqrt{2KD}}$ satisfies* *(i)* $\boldsymbol{i}^\star(\boldsymbol{\pi}) = \boldsymbol{i}^\star$,

(ii) $\forall \boldsymbol{x} \in \mathcal{X}\backslash\{\boldsymbol{i}^\star\}$ *and all* $k \in [K]$, $\frac{\partial f_{\boldsymbol{x}}(\boldsymbol{\omega},\boldsymbol{\pi})}{\partial \omega_k}$ *is continuous in* $\boldsymbol{\pi}$ *and*

$$\left\| \nabla_{\boldsymbol{\pi}} \left(\frac{\partial f_{\boldsymbol{x}}(\boldsymbol{\omega},\boldsymbol{\pi})}{\partial \omega_k} \right) \right\|_1 \leq 12D^2 \left\|\boldsymbol{\mu}\right\|_\infty .$$

Proof Proof of (i): Lemma 18 is equivalent to that: any $\boldsymbol{\pi} \in \mathbb{R}^K$ satisfying $\|\boldsymbol{\mu} - \boldsymbol{\pi}\|_\infty < \frac{\triangle_{\min}}{\sqrt{2KD}}$ implies that $\boldsymbol{\pi} \notin \mathsf{cl}\left(\mathsf{Alt}(\boldsymbol{\mu})\right)$. As closure of finite union equals union of closures,

$$\begin{aligned}
\mathbb{R}^K\backslash \mathsf{cl}\left(\mathsf{Alt}(\boldsymbol{\mu})\right) &= \mathbb{R}^K\backslash \left(\cup_{\boldsymbol{x}\neq \boldsymbol{i}^\star} \mathsf{cl}\left(\{\boldsymbol{\lambda} \in \mathbb{R}^K : \langle \boldsymbol{i}^\star - \boldsymbol{x}, \boldsymbol{\lambda}\rangle < 0\}\right)\right) \\
&= \mathbb{R}^K\backslash \left(\cup_{\boldsymbol{x}\neq \boldsymbol{i}^\star} \{\boldsymbol{\lambda} \in \mathbb{R}^K : \langle \boldsymbol{i}^\star - \boldsymbol{x}, \boldsymbol{\lambda}\rangle \leq 0\}\right) \\
&= \{\boldsymbol{\lambda} \in \mathbb{R}^K : \boldsymbol{i}^\star(\boldsymbol{\lambda}) = \boldsymbol{i}^\star\}.
\end{aligned}$$

Thus, $\boldsymbol{\pi} \notin \mathsf{cl}\left(\mathsf{Alt}(\boldsymbol{\mu})\right)$ is equivalent to $\boldsymbol{i}^\star(\boldsymbol{\pi}) = \boldsymbol{i}^\star$. This concludes the proof of (i).

Proof of (ii): Fix any $\boldsymbol{\pi} \in \mathbb{R}^K$ satisfying $\|\boldsymbol{\mu} - \boldsymbol{\pi}\|_\infty < \frac{\triangle_{\min}}{\sqrt{2KD}}$. By Lemma 19 and $\boldsymbol{i}^\star(\boldsymbol{\pi}) = \boldsymbol{i}^\star$,

$$\forall k \in [K], \quad \frac{\partial f_{\boldsymbol{x}}(\boldsymbol{\omega},\boldsymbol{\pi})}{\partial \omega_k} = \frac{\langle \boldsymbol{i}^\star - \boldsymbol{x}, \boldsymbol{\pi}\rangle^2 (x_k \oplus i^\star_k)}{2\left\langle \boldsymbol{x} \oplus \boldsymbol{i}^\star, \boldsymbol{\omega}^{-1}\right\rangle^2 \omega_k^2}.$$

Fix $k \in [K]$. Note that the function $\boldsymbol{\pi} \mapsto \frac{\partial f_{\boldsymbol{x}}(\boldsymbol{\omega},\boldsymbol{\pi})}{\partial \omega_k}$ is continuous and differentiable since it consists of inner products, element-wise products, and since its denumerator is always positive. For its derivative,

$$\begin{aligned}
\left\| \nabla_{\boldsymbol{\pi}} \left(\frac{\partial f_{\boldsymbol{x}}(\boldsymbol{\omega},\boldsymbol{\pi})}{\partial \omega_k} \right) \right\|_1 &= \left\| \frac{(\boldsymbol{i}^\star - \boldsymbol{x}) \langle \boldsymbol{i}^\star - \boldsymbol{x}, \boldsymbol{\pi}\rangle (x_k \oplus i^\star_k)}{\left\langle \boldsymbol{x} \oplus \boldsymbol{i}^\star, \boldsymbol{\omega}^{-1}\right\rangle^2 \omega_k^2} \right\|_1 \\
&\leq \left\|(\boldsymbol{i}^\star - \boldsymbol{x}) \langle \boldsymbol{i}^\star - \boldsymbol{x}, \boldsymbol{\pi}\rangle (x_k \oplus i^\star_k)\right\|_1 \\
&\leq \|\boldsymbol{i}^\star - \boldsymbol{x}\|_1 \left|\langle \boldsymbol{i}^\star - \boldsymbol{x}, \boldsymbol{\pi}\rangle\right| \leq 4D^2 \|\boldsymbol{\pi}\|_\infty \leq 12D^2 \|\boldsymbol{\mu}\|_\infty ,
\end{aligned}$$

where the first inequality is because $\left\langle \boldsymbol{x} \oplus \boldsymbol{i}^\star, \boldsymbol{\omega}^{-1}\right\rangle \omega_k \geq 1$ if $(x_k \oplus i^\star_k) = 1$; the second is because $x_k \oplus i^\star_k \leq 1$; the third uses $\|\boldsymbol{i}^\star - \boldsymbol{x}\|_1 \leq 2D$ and $|\langle \boldsymbol{i}^\star - \boldsymbol{x}, \boldsymbol{\pi}\rangle| \leq \|\boldsymbol{i}^\star - \boldsymbol{x}\|_1 \|\boldsymbol{\pi}\|_\infty$; the last uses the triangle inequality:

$$\|\boldsymbol{\pi}\|_\infty \leq \|\boldsymbol{\mu}\|_\infty + \|\boldsymbol{\mu} - \boldsymbol{\pi}\|_\infty \leq \|\boldsymbol{\mu}\|_\infty + \frac{\triangle_{\min}}{\sqrt{2KD}} \leq 3\|\boldsymbol{\mu}\|_\infty ,$$

where the last inequality is due to an application of Hölder's inequality to

$$\triangle_{\min} \leq \min_{\boldsymbol{x}\neq \boldsymbol{i}^\star} \|\boldsymbol{i}^\star - \boldsymbol{x}\|_1 \|\boldsymbol{\mu}\|_\infty \leq 2D \|\boldsymbol{\mu}\|_\infty .$$

□

Lemma 18. $\inf_{\boldsymbol{\lambda}\in \mathsf{Alt}(\boldsymbol{\mu})} \|\boldsymbol{\mu} - \boldsymbol{\lambda}\|_\infty \geq \frac{\triangle_{\min}}{\sqrt{2KD}}$.

Proof We claim that

$$\inf_{\boldsymbol{\lambda}\in\Lambda:\langle \boldsymbol{\lambda}, \boldsymbol{i}^\star - \boldsymbol{x}\rangle<0} \|\boldsymbol{\mu} - \boldsymbol{\lambda}\|_2^2 = \frac{\triangle_{\boldsymbol{x}}}{\|\boldsymbol{i}^\star \oplus \boldsymbol{x}\|_2}, \forall \boldsymbol{x} \neq \boldsymbol{i}^\star. \tag{51}$$

Observe that the proof immediately follows from (51) because the facts that $\|\boldsymbol{y}\|_2 \leq \sqrt{K}\,\|\boldsymbol{y}\|_\infty$ for any $\boldsymbol{y} \in \mathbb{R}^K$, $\mathsf{Alt}(\boldsymbol{\mu}) = \cup_{\boldsymbol{x}\neq \boldsymbol{i}^\star} \{\boldsymbol{\lambda} \in \Lambda : \langle \boldsymbol{\lambda}, \boldsymbol{i}^\star - \boldsymbol{x}\rangle < 0\}$, and $\|\boldsymbol{i}^\star \oplus \boldsymbol{x}\|_2 \leq \sqrt{2D}$.

Proof of (51): By solving the stationary conditions, i.e., $\nabla_{\boldsymbol{\lambda}} \mathcal{L}_{\boldsymbol{x}}(\boldsymbol{\lambda}^\star_{\boldsymbol{x}}, \alpha^\star) = 2(\boldsymbol{\mu} - \boldsymbol{\lambda}^\star_{\boldsymbol{x}}) + \alpha^\star(\boldsymbol{i}^\star - \boldsymbol{x}) = 0$ and $\nabla_\alpha \mathcal{L}_{\boldsymbol{x}}(\boldsymbol{\lambda}^\star_{\boldsymbol{x}}, \alpha^\star) = \langle \boldsymbol{\lambda}^\star_{\boldsymbol{x}}, \boldsymbol{i}^\star - \boldsymbol{x}\rangle = 0$, we find

$$\boldsymbol{\lambda}^\star_{\boldsymbol{x}} = \boldsymbol{\mu} - \frac{\triangle_{\boldsymbol{x}}(\boldsymbol{\mu}) \odot (\boldsymbol{i}^\star - \boldsymbol{x})}{\|\boldsymbol{i}^\star \oplus \boldsymbol{x}\|_2^2}$$

is a minimizer for $\inf_{\boldsymbol{\lambda}\in\Lambda:\langle \boldsymbol{\lambda}, \boldsymbol{i}^\star - \boldsymbol{x}\rangle<0} \|\boldsymbol{\mu} - \boldsymbol{\lambda}\|_2$. (51) follows by plugging $\boldsymbol{\lambda}^\star_{\boldsymbol{x}}$ into $\|\boldsymbol{\mu} - \boldsymbol{\lambda}\|_2^2$. □

Remind that $\nabla_{\boldsymbol{\omega}} f_{\boldsymbol{x}}$ can be evaluated by the following Lemma 19.

Lemma 19 (Envelope theorem)**.** *Let* $(\boldsymbol{\omega}, \boldsymbol{\mu}) \in \Sigma_+ \times \Lambda$ *and* $\boldsymbol{x} \in \mathcal{X} \setminus \{\boldsymbol{i}^\star\}$*. Define* $\boldsymbol{\lambda}^\star_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}) \in \operatorname{argmin}_{\boldsymbol{\lambda} \in \mathrm{cl}(\mathcal{C}_{\boldsymbol{x}})} \left\langle \boldsymbol{\omega}, \frac{(\boldsymbol{\mu}-\boldsymbol{\lambda})^2}{2} \right\rangle$*. Then,*

$$\nabla_{\boldsymbol{\omega}} f_{\boldsymbol{x}}(\boldsymbol{\omega}, \boldsymbol{\mu}) = \frac{(\boldsymbol{\mu} - \boldsymbol{\lambda}^\star_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}))^2}{2} = \frac{\triangle_{\boldsymbol{x}}(\boldsymbol{\mu})^2 (\boldsymbol{x} \oplus \boldsymbol{i}^\star) \odot \boldsymbol{\omega}^{-2}}{2 \left\langle \boldsymbol{x} \oplus \boldsymbol{i}^\star, \boldsymbol{\omega}^{-1} \right\rangle^2}.$$

Proof The first equality is an application of Lemma 6 and Proposition 1 of [WTP21] with $\mathcal{I} = \mathcal{X}$, $\mathcal{J}_{\boldsymbol{x}} = \{\boldsymbol{x}\}$, $\Sigma = \Sigma_K$, $\mathcal{S}_{\boldsymbol{x}} = \{\boldsymbol{\lambda} \in \Lambda : \boldsymbol{i}^\star(\boldsymbol{\lambda}) = \boldsymbol{x}\}$ (see Appendix K.2 and Appendix K.4 in [WTP21] for more details). The second equality substitutes $\boldsymbol{\lambda}^\star_{\boldsymbol{\omega},\boldsymbol{\mu}}(\boldsymbol{x}) = \boldsymbol{\mu} + \frac{\triangle_{\boldsymbol{x}}(\boldsymbol{\mu})(\boldsymbol{x} - \boldsymbol{i}^\star) \odot \boldsymbol{\omega}^{-1}}{\langle \boldsymbol{x} \oplus \boldsymbol{i}^\star, \boldsymbol{\omega}^{-1} \rangle}$ by using (11)-(14). $\square$

H Stochastic smoothing

This section is devoted to present Proposition 2 and verify the assumptions required for applying Proposition 2 to our objective $F_{\boldsymbol{\mu}}$.

Stochastic smoothing [FKM05, DBW12] is a well-studied technique and has been widely applied to online convex nonsmooth optimization [HK12, H$^+$16]. Proposition 2 is a restatement of existing results. In particular, Proposition 2 (i), (ii) and (iii) directly follow from Lemma E.2 in [DBW12] with $(L_0, u) = (\ell, \eta)$, $f = -\Phi$ and $f_u = -\bar{\Phi}_\eta(\cdot)$, and Proposition 2 (iv) can be established by Jensen's inequality as done in the proof of Theorem 2.1 [DBW12].

Proposition 2. *Assume that $\Phi : \mathbb{R}^K_{>0} \mapsto \mathbb{R}$ is concave, ℓ-Lipschitz, and differentiable almost everywhere. Let $B_2 = \{\boldsymbol{v} \in \mathbb{R}^K : \|\boldsymbol{v}\|_2 \le 1\}$. For any $\boldsymbol{\omega} \in \Sigma_+$ and $\eta \in (0, \min_{k\in[K]} \omega_k)$, define*

$$\bar{\Phi}_\eta(\boldsymbol{\omega}) = \mathbb{E}_{\boldsymbol{\mathcal{Z}}\sim \textit{Uniform}(B_2)}[\Phi(\boldsymbol{\omega} + \eta\boldsymbol{\mathcal{Z}})] \,. \tag{6}$$

Then, $\bar{\Phi}_\eta(\boldsymbol{\omega})$ satisfies that:

(i) $\Phi(\boldsymbol{\omega}) - \eta\ell \le \bar{\Phi}_\eta(\boldsymbol{\omega}) \le \Phi(\boldsymbol{\omega})$

(ii) $\nabla\bar{\Phi}_{\boldsymbol{\mu},\eta}(\boldsymbol{\omega}) = \mathbb{E}_{\boldsymbol{\mathcal{Z}}\sim \textit{Uniform}(B_2)}[\nabla\Phi_{\boldsymbol{\mu}}(\boldsymbol{\omega} + \eta\boldsymbol{\mathcal{Z}})]$

(iii) $\bar{\Phi}_\eta$ *is* $\frac{\ell K}{\eta}$*-smooth*

(iv) if $\eta > \eta' > 0$, then $\bar{\Phi}_{\eta'}(\boldsymbol{\omega}) \ge \bar{\Phi}_\eta(\boldsymbol{\omega})$

Now, we validate assumptions of Proposition 2 on $F_{\boldsymbol{\mu}}$. The concavity of $F_{\boldsymbol{\mu}}$, which is shown by [WTP21], follows from the facts that each $f_{\boldsymbol{x}}(\cdot, \boldsymbol{\mu})$ is concave and that $F_{\boldsymbol{\mu}}$ is a minimum of these functions $f_{\boldsymbol{x}}(\cdot, \boldsymbol{\mu})$ over all possible $\boldsymbol{x}$. The Lipschitzness of $F_{\boldsymbol{\mu}}$ is shown in Lemma 21 in Appendix I). Hence, it remains to show the almost-everywhere differentiability of $F_{\boldsymbol{\mu}}$. To show that the set of non-differentiable points of $F_{\boldsymbol{\mu}}$, i.e.,

$$\bigcup_{\boldsymbol{x},\boldsymbol{x}'\in\mathcal{X}\setminus\{\boldsymbol{i}^\star(\boldsymbol{\mu})\},\, \boldsymbol{x}\neq\boldsymbol{x}'} \{\boldsymbol{z} \in B_2 : f_{\boldsymbol{x}}((\boldsymbol{\omega} + \eta\boldsymbol{z}, \boldsymbol{\mu}) = f_{\boldsymbol{x}'}(\boldsymbol{\omega} + \eta\boldsymbol{z}, \boldsymbol{\mu})\} \,,$$

is measure-zero under Uniform(B_2), it suffices to show the following lemma.

Lemma 20. *Let $\boldsymbol{\mu} \in \Lambda$ and $\boldsymbol{x}_1, \boldsymbol{x}_2$ be distinct actions in $\mathcal{X}\setminus\{\boldsymbol{i}^\star(\boldsymbol{\mu})\}$. Then under the probability measure of Uniform(B_2),*

$$\{\boldsymbol{z} \in B_2 : f_{\boldsymbol{x}_1}(\boldsymbol{\omega} + \eta\boldsymbol{z}, \boldsymbol{\mu}) = f_{\boldsymbol{x}_2}(\boldsymbol{\omega} + \eta\boldsymbol{z}, \boldsymbol{\mu})\}$$

is a measure-zero set.

Proof To simplify the notation, let $\boldsymbol{i}^\star = \boldsymbol{i}^\star(\boldsymbol{\mu})$ and $\triangle_{\boldsymbol{x}} = \triangle_{\boldsymbol{x}}(\boldsymbol{\mu})$. Thanks to the close-form expressions of $f_{\boldsymbol{x}_1}$ and $f_{\boldsymbol{x}_2}$, $\boldsymbol{z} \in B_2$ such that $f_{\boldsymbol{x}_1}(\boldsymbol{\omega} + \eta\boldsymbol{z}, \boldsymbol{\mu}) = f_{\boldsymbol{x}_2}(\boldsymbol{\omega} + \eta\boldsymbol{z}, \boldsymbol{\mu})$ are the points satisfying that:

$$\frac{\triangle^2_{\boldsymbol{x}_1}}{2\left\langle \boldsymbol{x}_1 \oplus \boldsymbol{i}^\star, (\boldsymbol{\omega} + \eta\boldsymbol{z})^{-1}\right\rangle} = \frac{\triangle^2_{\boldsymbol{x}_2}}{2\left\langle \boldsymbol{x}_2 \oplus \boldsymbol{i}^\star, (\boldsymbol{\omega} + \eta\boldsymbol{z})^{-1}\right\rangle}.$$

In other words, the set of interests is

$$\left\{ \boldsymbol{z} \in B_2 : \sum_{k=1}^{K} a_k \prod_{k'\neq k} (\omega_{k'} + \eta z_{k'}) = 0 \right\}, \tag{52}$$

where $a_k = (\boldsymbol{x}_2 \oplus \boldsymbol{i}^\star)_k \triangle^2_{\boldsymbol{x}_1} - (\boldsymbol{x}_1 \oplus \boldsymbol{i}^\star)_k \triangle^2_{\boldsymbol{x}_2}$ for all $k \in [K]$. We claim that $\boldsymbol{a}$ is a non-zero vector. Otherwise, $a_k = 0, \forall k \in [K]$, which together with the fact that $\triangle^2_{\boldsymbol{x}_1}, \triangle^2_{\boldsymbol{x}_2} > 0$ directly imply $(\boldsymbol{x}_2 \oplus \boldsymbol{i}^\star)_k = 0$ if and only if $(\boldsymbol{x}_1 \oplus \boldsymbol{i}^\star)_k = 0$. That means $\boldsymbol{x}_1 = \boldsymbol{x}_2$, but this becomes a contradiction. Therefore, the set in (52) are the roots of a non-zero polynomial inside B_2, and hence it is a measure-zero set (see e.g. Lemma in [Oka73]). □

I Lischitzness of $F_{\boldsymbol{\mu}}$ and boundness of $F_{\boldsymbol{\mu}}$ on $\Sigma_K \cap \mathbb{R}^K_{>0}$

In this section, we show the Lipschitzness of $F_{\boldsymbol{\mu}}(\boldsymbol{v}) = \min_{\boldsymbol{x}\neq \boldsymbol{i}^\star} f_{\boldsymbol{x}}(\boldsymbol{v},\boldsymbol{\mu})$ for $\boldsymbol{v} \in \mathbb{R}^K_{>0}$. Let $\boldsymbol{x}_e$ be an equilibrium action such that $F_{\boldsymbol{\mu}}(\boldsymbol{v}) = f_{\boldsymbol{x}_e}(\boldsymbol{v},\boldsymbol{\mu})$. We will use the envelope theorem (Lemma 19 in Appendix G) to evaluate $\nabla_{\boldsymbol{\omega}} f_{\boldsymbol{x}_e}(\boldsymbol{v},\boldsymbol{\mu})$ in closed-form, and then bound its length. We will also derive an upper bound of $F_{\boldsymbol{\mu}}(\boldsymbol{v})$ valid for any positive vector $\boldsymbol{v}$ in the $(K-1)$-dimensional simplex Σ_K. In what below, we denote $\boldsymbol{i}^\star = \boldsymbol{i}^\star(\boldsymbol{\mu})$ and $\triangle_{\boldsymbol{x}} = \triangle_{\boldsymbol{x}}(\boldsymbol{\mu})$ for any $\boldsymbol{x} \neq \boldsymbol{i}^\star$ for short.

Lemma 21. *Let $\boldsymbol{\mu} \in \Lambda$ and $\ell = 2D^2 \|\boldsymbol{\mu}\|_\infty^2$. Then, $F_{\boldsymbol{\mu}}$ is ℓ-Lipschitz with respect to $\|\cdot\|_\infty$ on $\mathbb{R}^K_{>0}$,*

Proof Let $\boldsymbol{v} \in \mathbb{R}^K_{>0}$. Recall that $F_{\boldsymbol{\mu}}(\boldsymbol{v}) = \min_{\boldsymbol{x}\neq \boldsymbol{i}^\star} f_{\boldsymbol{x}}(\boldsymbol{v},\boldsymbol{\mu})$, and each $f_{\boldsymbol{x}}(\boldsymbol{v},\boldsymbol{\mu})$ is differentiable (proven in Lemma 19 in Appendix G.2). Hence if $\boldsymbol{x}$ is the action such that $F_{\boldsymbol{\mu}}(\boldsymbol{v}) = f_{\boldsymbol{x}}(\boldsymbol{v},\boldsymbol{\mu})$, the concavity of $F_{\boldsymbol{\mu}}(\boldsymbol{v})$ and the fact that $\nabla_{\boldsymbol{\omega}} f_{\boldsymbol{x}}(\boldsymbol{v},\boldsymbol{\mu})$ is the subdifferential of $F_{\boldsymbol{\mu}}$ on $\boldsymbol{v}$ yield that

$$\forall \boldsymbol{v}' \in \mathbb{R}^K_{>0}, |F_{\boldsymbol{\mu}}(\boldsymbol{v}) - F_{\boldsymbol{\mu}}(\boldsymbol{v}')| \leq |\langle \nabla_{\boldsymbol{\omega}} f_{\boldsymbol{x}}(\boldsymbol{v},\boldsymbol{\mu}), \boldsymbol{v}-\boldsymbol{v}'\rangle| \leq \|\nabla_{\boldsymbol{\omega}} f_{\boldsymbol{x}}(\boldsymbol{v},\boldsymbol{\mu})\|_1 \|\boldsymbol{v}-\boldsymbol{v}'\|_\infty ,$$

where the last inequality stems from Hölder's inequality. From the above, the ℓ-Lipschitz can be derived by upper bounding $\|\nabla_{\boldsymbol{\omega}} f_{\boldsymbol{x}}(\boldsymbol{v},\boldsymbol{\mu})\|_1$ by ℓ. Now applying Lemma 19 in Appendix G.2 yields

$$\|\nabla_{\boldsymbol{\omega}} f_{\boldsymbol{x}}(\boldsymbol{v},\boldsymbol{\mu})\|_1 = \left\| \frac{(\boldsymbol{\mu} - \boldsymbol{\lambda}^\star_{\boldsymbol{v},\boldsymbol{\mu}}(\boldsymbol{x},\alpha^\star_{\boldsymbol{x}}))^2}{2} \right\|_1 = \frac{\left\|\boldsymbol{v}^{-2} \odot (\boldsymbol{x} \oplus \boldsymbol{i}^\star)\right\|_1 \triangle^2_{\boldsymbol{x}}}{2\left\langle \boldsymbol{x} \oplus \boldsymbol{i}^\star, \boldsymbol{v}^{-1}\right\rangle^2}. \tag{53}$$

To simplify the above, we observe that

$$\left\langle \boldsymbol{x} \oplus \boldsymbol{i}^\star, \boldsymbol{v}^{-1}\right\rangle^2 = \left(\sum_{k=1}^K v_k^{-1} \mathbb{1}\{x_k \neq \boldsymbol{i}^\star_k\}\right)^2 \geq \sum_{k=1}^K v_k^{-2} \mathbb{1}\{x_k \neq \boldsymbol{i}^\star_k\} = \left\|\boldsymbol{v}^{-2} \odot (\boldsymbol{x} \oplus \boldsymbol{i}^\star)\right\|_1 , \tag{54}$$

where the inequality uses the fact that $v_k > 0$ for all $k \in [K]$. Also,

$$\triangle_{\boldsymbol{x}} = \langle \boldsymbol{i}^\star - \boldsymbol{x}, \boldsymbol{\mu}\rangle \leq \|\boldsymbol{i}^\star - \boldsymbol{x}\|_1 \|\boldsymbol{\mu}\|_\infty \leq 2D \|\boldsymbol{\mu}\|_\infty . \tag{55}$$

Thus, (53)-(54)-(55) yields that $\|\nabla_{\boldsymbol{\omega}} f_{\boldsymbol{x}}(\boldsymbol{v},\boldsymbol{\mu})\|_1 \leq 2D^2 \|\boldsymbol{\mu}\|^2_\infty$. □

Lemma 22. *Let $\boldsymbol{\mu} \in \Lambda$ and $\ell = 2D^2 \|\boldsymbol{\mu}\|_\infty^2$. Then, $\max_{\boldsymbol{\omega}\in\Sigma_K \cap \mathbb{R}^K_{>0}} F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) \leq \ell$.*

Proof Observe that $f_{\boldsymbol{x}}(\boldsymbol{v},\boldsymbol{\mu}) = \langle \boldsymbol{\omega}, \nabla_{\boldsymbol{\omega}} f_{\boldsymbol{x}}(\boldsymbol{v},\boldsymbol{\mu})\rangle$ for any $\boldsymbol{x} \neq \boldsymbol{i}^\star$. Combining this observation with the fact that $\triangle_{\boldsymbol{x}} \leq 2D \|\boldsymbol{\mu}\|_\infty$ (as argued in (54) in proof of Lemma 21) implies:

$$F_{\boldsymbol{\mu}}(\boldsymbol{v}) = \min_{\boldsymbol{x}\neq \boldsymbol{i}^\star} \frac{\triangle^2_{\boldsymbol{x}}}{2\left\langle \boldsymbol{x} \oplus \boldsymbol{i}^\star, \boldsymbol{v}^{-1}\right\rangle} \leq \frac{(2D\|\boldsymbol{\mu}\|_\infty)^2}{2} = \ell,$$

where the first inequality is because $\left\langle \boldsymbol{x} \oplus \boldsymbol{i}^\star, \boldsymbol{v}^{-1}\right\rangle \geq \min_{k\in[K]} v_k^{-1} \geq 1$ (as $\boldsymbol{v} \in \Sigma_K$ and $v_k > 0$ for all $k \in [K]$).The proof is completed since $\boldsymbol{v}$ is taken arbitrarily. □

J Proofs related to combinatorial sets

Assumption 1. *(i) There exists a polynomial-time algorithm identifying $\boldsymbol{i}^\star(\boldsymbol{v})$ for any $\boldsymbol{v} \in \mathbb{R}^K$; (ii) $\mathcal{X}$ is inclusion-wise maximal, i.e., there is no $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{X}$ s.t. $\boldsymbol{x} < \boldsymbol{x}'$; (iii) for each $k \in [K]$, there exists $\boldsymbol{x} \in \mathcal{X}$ such that $x_k = 1$; (iv) $|\mathcal{X}| \geq 2$.*

As claimed in §2.2, Assumption 1 holds for the following combinatorial sets:

- *m-sets*: $\mathcal{X} = \{\boldsymbol{x} \in \{0,1\}^K : \|\boldsymbol{x}\|_1 = m\}$
- *spanning forests*: $\mathcal{X}$ is a set of all spanning forests in a given graph
- *bipartite matchings*: $\mathcal{X}$ is a set of all maximal matchings in a given bipartite graph
- *s-t paths*: $\mathcal{X}$ is the set of all source-destination paths in a directed acyclic graph

In what below, we present a simple proof for the above examples.

Proof Suppose (iii) (iv) hold (as we can always achieve (iii) by removing arms not covered by $\mathcal{X}$ and (iv) holds for non-trivial sets). For (i), it is well-known that a polynomial-time LM Oracle, i.e., $\boldsymbol{i}^\star(\cdot)$, exists for each of the discussed combinatorial structures. For example, see Chapter 39 in [S$^+$03] for the greedy algorithm for matroids (applicable to m-set and spanning forests), Chapter 41 in [S$^+$03] for the augmentation-based algorithm for 2-matroid intersection (applicable to bipartite matchings), and algorithms such as Dijkstra's algorithm for s-t paths.

It remains to verify (ii) the inclusion-wise maximal property of $\mathcal{X}$. For $\mathcal{X}$ as m-sets, the inclusion-wise maximal property clearly holds because any binary vector $\boldsymbol{x}' > \boldsymbol{x}$ (resp. $\boldsymbol{x}' < \boldsymbol{x}$) for some $\boldsymbol{x} \in \mathcal{X}$ must have $\sum_{k\in[K]} x'_k > m$ (resp. $< m$) and thus $\boldsymbol{x}' \notin \mathcal{X}$. The case is similar for $\mathcal{X}$ as spanning forests since the number of edges of any spanning forests of a graph is the same. For $\mathcal{X}$ as maximal matchings in which the term 'maximal' exactly refers to being inclusion-wise maximal, (ii) directly follows from the definition. For $\mathcal{X}$ as the set of all source-destination paths in an acyclic graph, if there exists any source-destination path $\boldsymbol{x}' > \boldsymbol{x}$ for some $\boldsymbol{x} \in \mathcal{X}$ then $\boldsymbol{x}'$ must contain a cycle, so inclusion-wise maximal property holds. ☐

Lemma 2. *Let $\boldsymbol{v} \in \mathbb{R}^K$ and $\boldsymbol{x} \in \mathcal{X}$. Under Assumption 1, there exists an algorithm that solves $\max_{\boldsymbol{x}'\in\mathcal{X}:\boldsymbol{x}'\neq\boldsymbol{x}} \langle \boldsymbol{v}, \boldsymbol{x}' \rangle$ by only making at most D queries to the LM Oracle.*

Proof Fix $\boldsymbol{x} \in \mathcal{X}$. Assume $\boldsymbol{v} \neq \boldsymbol{0}_K$ (as otherwise, any $\boldsymbol{x}' \neq \boldsymbol{x}$ is a second-best action). Inspired by Lawler-Murty's m-best algorithm [Law72], we will prove this lemma by considering the algorithm described as follows. It first computes $\boldsymbol{i}^\star(\boldsymbol{v})$ by the LM Oracle, and returns it as the output if $\boldsymbol{i}^\star(\boldsymbol{v}) \neq \boldsymbol{x}$. Otherwise, we identify the second-best action by the program below:

$$\max_{k\in[K]:x_k=1} \left\langle \boldsymbol{v}, \boldsymbol{i}^\star\left(\boldsymbol{v}^{(k)}\right)\right\rangle, \quad \text{where} \quad v_i^{(k)} = \begin{cases} -3\left\|\boldsymbol{v}\right\|_1 & \text{if } i = k \\ v_i & \text{otherwise.} \end{cases} \tag{56}$$

Intuitively, for each arm k of $\boldsymbol{x}$, the action $\boldsymbol{i}^\star\left(\boldsymbol{v}^{(k)}\right)$ represents the best one among all actions without k (we have a strong negative weight on the k-th component of $\boldsymbol{v}^{(k)}$). In the following, we will show that at least one of $\{\boldsymbol{i}^\star\left(\boldsymbol{v}^{(k)}\right) : k \in [K], x_k = 1\}$ is the second-best action.

More precisely, we will show that for any maximizer $a \in [K]$ to (56), $\boldsymbol{i}^\star\left(\boldsymbol{v}^{(a)}\right)$ is a second-best action. Consider if $(\boldsymbol{i}^\star\left(\boldsymbol{v}^{(a)}\right))_a = 0$, then the claim follows from the fact that $\boldsymbol{i}^\star\left(\boldsymbol{v}^{(a)}\right)$ is the best among all actions without a and also the best in $\{\boldsymbol{i}^\star\left(\boldsymbol{v}^{(k)}\right) : k \in [K], x_k = 1\}$. It suffices to show that $(\boldsymbol{i}^\star\left(\boldsymbol{v}^{(a)}\right))_a = 1$ cannot happen. If $(\boldsymbol{i}^\star\left(\boldsymbol{v}^{(a)}\right))_a = 1$, then it follows from Assumption 1 (iv) $|\mathcal{X}| \geq 2$ and (ii) the inclusion-wise maximality of $\mathcal{X}$ that there is another action $\boldsymbol{x}'$ such that $x'_k = 0$ but $x_k = 1$ for some $k \in [K]$. So, by $\boldsymbol{i}^\star\left(\boldsymbol{v}^{(a)}\right)_a = 1$, $\boldsymbol{v} \neq \boldsymbol{0}_K$ and the definition of $\boldsymbol{v}^{(a)}$, we get

$$\left\langle \boldsymbol{v}, \boldsymbol{i}^\star\left(\boldsymbol{v}^{(a)}\right)\right\rangle = \sum_{j\in[K]:\boldsymbol{i}^\star\left(\boldsymbol{v}^{(a)}\right)_j=1, j\neq a} v_j - 3\left\|\boldsymbol{v}\right\|_1 \leq -2\left\|\boldsymbol{v}\right\|_1 < \langle \boldsymbol{v}, \boldsymbol{x}' \rangle \leq \left\langle \boldsymbol{v}, \boldsymbol{i}^\star\left(\boldsymbol{v}^{(k)}\right)\right\rangle,$$

which contradicts the optimality of a (as it would imply that $\boldsymbol{i}^\star\left(\boldsymbol{v}^{(k)}\right)$ is better).

Finally, as $\|\boldsymbol{x}\|_1 \leq D$, the number of `LM` Oracle calls required for solving (56) is at most D. $\square$

Finally, we present the property of $\mathcal{X}_0$ briefly argued in § 4.2.

Lemma 23. *Let $\boldsymbol{e}_k$ is the k-th column of an identity matrix. Under Assumption 1, $\mathcal{X}_0$ is a $[K]$-covering set and $|\mathcal{X}_0| \geq 2$.*

Proof Showing that $\mathcal{X}_0$ covers $[K]$: Assumption 1 (iii) ensures $\{\boldsymbol{x} \in \mathcal{X} : x_k = 1\} \neq \emptyset$, and it follows that $\max_{\boldsymbol{x}\in\mathcal{X}} \langle \boldsymbol{x}, \boldsymbol{e}_k \rangle = 1$, i.e., $(\boldsymbol{i}^\star(\boldsymbol{e}_k))_k = 1$. As $(\boldsymbol{i}^\star(\boldsymbol{e}_k))_k = 1$ holds for all k, the proof is completed.

Showing that $|\mathcal{X}_0| \geq 2$: Suppose on the contrary, $|\mathcal{X}_0| = 1$. Thanks to Assumption 1 (iv) $|\mathcal{X}| \geq 2$, there exists $\boldsymbol{x} \in \mathcal{X}$ such that $x_k = \langle \boldsymbol{e}_k, x \rangle \geq \langle \boldsymbol{e}_k, x' \rangle = x'_k$ for all $k \in [K]$, $x' \neq x$. Together with Assumption 1 (iii), one can easily deduce that $x_k = 1$ for all $k \in [K]$. However, this implies $\boldsymbol{x}' < \boldsymbol{x}$ for any $\boldsymbol{x}' \neq \boldsymbol{x}$ and hence contradicts to Assumption 1 (ii) that $\mathcal{X}$ is inclusion-wise maximal. $\square$

K Sample complexity lower bound

In this section, we assume $\boldsymbol{\mu} \in \Lambda$ and $\delta \in (0,1)$ is fixed, and show Theorem 7 by adapting Lemma 19 in [KCG16].

Lemma 24 ([KCG16])**.** *Any δ-PAC algorithm satisfies*

$$\forall \boldsymbol{\lambda} \in \mathsf{Alt}(\boldsymbol{\mu}), \sum_{k\in[K]} \sum_{\boldsymbol{x}\in\mathcal{X}:x_k=1} \mathbb{E}_{\boldsymbol{\mu}}[N_{\boldsymbol{x}}(\tau)] \frac{(\mu_k-\lambda_k)^2}{2} \geq \mathsf{kl}(\delta, 1-\delta). \tag{57}$$

Theorem 7. *Any δ-PAC strategy satisfies*

$$\mathbb{E}_{\boldsymbol{\mu}}[\tau] \geq T^\star(\boldsymbol{\mu}) \mathsf{kl}(\delta, 1-\delta) \quad \textit{with} \quad T^\star(\boldsymbol{\mu})^{-1} = \sup_{\boldsymbol{\omega}\in\Sigma} \inf_{\boldsymbol{\lambda}\in\mathsf{Alt}(\boldsymbol{\mu})} \left\langle \boldsymbol{\omega}, \frac{(\boldsymbol{\mu}-\boldsymbol{\lambda})^2}{2} \right\rangle, \tag{1}$$

where $\Sigma = \{\sum_{\boldsymbol{x}\in\mathcal{X}} w_{\boldsymbol{x}} : \boldsymbol{w} \in \Sigma_{|\mathcal{X}|}\}$ and $\mathsf{Alt}(\boldsymbol{\mu}) = \{\boldsymbol{\lambda} \in \Lambda : \boldsymbol{i}^\star(\boldsymbol{\lambda}) \neq \boldsymbol{i}^\star(\boldsymbol{\mu})\}$.

Proof We have: under any algorithm,

$$\sup_{\boldsymbol{\omega}\in\Sigma} \inf_{\boldsymbol{\lambda}\in\mathsf{Alt}(\boldsymbol{\mu})} \sum_{k\in[K]} \omega_k \frac{(\mu_k-\lambda_k)^2}{2} \geq \inf_{\boldsymbol{\lambda}\in\mathsf{Alt}(\boldsymbol{\mu})} \sum_{k\in[K]} \sum_{\boldsymbol{x}\in\mathcal{X}:x_k=1} \frac{\mathbb{E}_{\boldsymbol{\mu}}[N_{\boldsymbol{x}}(\tau)]}{\mathbb{E}_{\boldsymbol{\mu}}[\tau]} \frac{(\mu_k-\lambda_k)^2}{2},$$

Hence if the algorithm is δ-PAC, by Lemma 24,

$$\begin{aligned}\mathbb{E}_{\boldsymbol{\mu}}[\tau] \sup_{\boldsymbol{\omega}\in\Sigma} \inf_{\boldsymbol{\lambda}\in\mathsf{Alt}(\boldsymbol{\mu})} \sum_{k\in[K]} \omega_k \frac{(\mu_k-\lambda_k)^2}{2} &\geq \inf_{\boldsymbol{\lambda}\in\mathsf{Alt}(\boldsymbol{\mu})} \sum_{k\in[K]} \sum_{\boldsymbol{x}\in\mathcal{X}:x_k=1} \mathbb{E}_{\boldsymbol{\mu}}[N_{\boldsymbol{x}}(\tau)] \frac{(\mu_k-\lambda_k)^2}{2} \\ &\geq \mathsf{kl}(\delta, 1-\delta).\end{aligned}$$

□

Lemma 1. *For any $\boldsymbol{\mu} \in \Lambda$, $T^\star(\boldsymbol{\mu}) \leq 4KD\triangle_{\min}(\boldsymbol{\mu})^{-2}$.*

Proof Take $\boldsymbol{\omega}_0 = \sum_{\boldsymbol{x}\in\mathcal{X}_0} \boldsymbol{x}/|\mathcal{X}_0| \in \Sigma$, where $\mathcal{X}_0 = \{\boldsymbol{i}^\star(\boldsymbol{e}_k) : k \in [K]\}$. Observe that $\boldsymbol{\omega}_0 \geq \mathbf{1}_K/K$ by Lemma 23 (which leads to $\sum_{\boldsymbol{x}\in\mathcal{X}_0} \boldsymbol{x} \geq \mathbf{1}_K$ and $1/|\mathcal{X}_0| \geq 1/K$). Thus,

$$F_{\boldsymbol{\mu}}(\boldsymbol{\omega}_0) = \min_{\boldsymbol{x}\neq\boldsymbol{i}^\star(\boldsymbol{\mu})} \frac{\triangle_{\boldsymbol{x}}(\boldsymbol{\mu})^2}{2\left\langle \boldsymbol{x}\oplus\boldsymbol{i}^\star(\boldsymbol{\mu}), \boldsymbol{\omega}_0^{-1}\right\rangle} \geq \frac{\triangle_{\min}(\boldsymbol{\mu})^2}{4KD},$$

where we used Proposition 1 in §3.1 to obtain the equality, and the last inequality is because

$$\left\langle \boldsymbol{x}\oplus\boldsymbol{i}^\star(\boldsymbol{\mu}), \boldsymbol{\omega}_0^{-1}\right\rangle \leq \|\boldsymbol{x}\oplus\boldsymbol{i}^\star(\boldsymbol{\mu})\|_1 \left\|\boldsymbol{\omega}_0^{-1}\right\|_\infty \leq \frac{2D}{\min_{k\in[K]}(\boldsymbol{\omega}_0)_k} \leq 2KD.$$

As $T^\star(\boldsymbol{\mu})^{-1} = \max_{\boldsymbol{\omega}\in\Sigma} F_{\boldsymbol{\mu}}(\boldsymbol{\omega}) \geq F_{\boldsymbol{\mu}}(\boldsymbol{\omega}_0)$, we then have $T^\star(\boldsymbol{\mu}) \leq \frac{4KD}{\triangle_{\min}(\boldsymbol{\mu})^2}$. □

L Extension to the transductive setting

In this section, we extend our results to the transductive combinatorial semi-bandits. In transductive best-arm identification with fixed confidence with semi-bandit feedback [JMKK21], the decision maker is given an exploration set $\mathcal{A} \subseteq \{0,1\}^K$ and a decision set $\mathcal{X} \subseteq \{0,1\}^K$ ($\mathcal{A}$ might differ from $\mathcal{X}$), and at each round, she selects an action in $\mathcal{A}$ to receive a semi-bandit feedback. Her goal is to identify the best action in $\mathcal{X}$ using as few samples as possible.

Notation. Let $\mathcal{M} \subseteq \{0,1\}^K$ be any set of actions. We use $\boldsymbol{i}^\star_{\mathcal{M}}(\boldsymbol{\mu})$ to denote any maximizer in $\mathcal{M}$ of the linear maximization $\max_{\boldsymbol{x}\in\mathcal{M}} \langle \boldsymbol{x}, \boldsymbol{\mu}\rangle$. We also use $\Sigma_{\mathcal{M}} = \{\sum_{\boldsymbol{x}\in\mathcal{M}} w_{\boldsymbol{x}} : \boldsymbol{w} \in \Sigma_{|\mathcal{M}|}\}$.

Sample complexity lower bound. The generalization of Theorem 7 to the transductive setting has been made in [JMKK21]: any δ-PAC algorithm satisfies

$$\mathbb{E}_{\boldsymbol{\mu}}[\tau] \geq T^\star(\boldsymbol{\mu})\mathsf{kl}(\delta, 1-\delta) \quad \text{with} \quad T^\star(\boldsymbol{\mu})^{-1} = \sup_{\boldsymbol{\omega}\in\Sigma_{\mathcal{A}}} \inf_{\boldsymbol{\lambda}\in\mathsf{Alt}(\boldsymbol{\mu})} \left\langle \boldsymbol{\omega}, \frac{(\boldsymbol{\mu}-\boldsymbol{\lambda})^2}{2} \right\rangle. \tag{58}$$

The inner optimization is still with respect to $\mathcal{X}$ while the outer optimization is with respect to the exploration set $\mathcal{A}$. Refer to Appendix C in [JMKK21] for the proof.

Transductive `P-FWS` algorithm. Assumption 1 has to be extended. It now needs to ensure that $\boldsymbol{i}^\star_{\mathcal{A}}(\boldsymbol{v})$ for any $\boldsymbol{v} \in \mathbb{R}^K$ can be computed in polynomial-time. The `P-FWS` algorithm also needs to be adapted to the transductive setting. This is done by the following two modifications:

- $[K]$-covering set: $\mathcal{X}_0 \leftarrow \{\boldsymbol{i}^\star_{\mathcal{A}}(\boldsymbol{e}_k) : k \in [K]\}$
- FW update: $\boldsymbol{x}(t) \leftarrow \boldsymbol{i}^\star_{\mathcal{A}}\left(\nabla \tilde{F}_{\hat{\boldsymbol{\mu}}(t-1),\eta_t,n_t}(\hat{\boldsymbol{\omega}}(t-1))\right)$

Analysis of `P-FWS`. Let $D_{\mathcal{A}} = \max_{\boldsymbol{x}\in\mathcal{A}} \|\boldsymbol{x}\|_1$. The analysis is easily extended by replacing $(D, \mathcal{X})$ with $(D_{\mathcal{A}}, \mathcal{A})$ in Appendix D, Appendix E, Appendix F and Appendix G whenever the context is subject to the exploration set rather than the decision set.

Appendix D

Matroid semi-bandits in sublinear time

Figure: A cat being rounded and having the minimum hitting set to play with.

D

We study the matroid semi-bandit problem and propose an algorithm that runs in time sublinear to the number of arms for the common matroids while achieving the gap-dependent regret lower bound by Kveton et al. [KWA+14]. The main technique is rounding and the minimum hitting set.

Matroid Semi-Bandits in Sublinear Time

Ruo-Chun Tzeng [1 §] **Naoto Ohsaka** [2] **Kaito Ariu** [2]

Abstract

We study the matroid semi-bandits problem, where at each round the learner plays a subset of K arms from a feasible set, and the goal is to maximize the expected cumulative linear rewards. Existing algorithms have per-round time complexity at least $\Omega(K)$, which becomes expensive when K is large. To address this computational issue, we propose `FasterCUCB` whose sampling rule takes time sublinear in K for common classes of matroids: $\mathcal{O}(D\,\mathrm{polylog}\,(K)\;\mathrm{polylog}\,(T))$ for uniform matroids, partition matroids, and graphical matroids, and $\mathcal{O}(D\sqrt{K}\mathrm{polylog}\,(T))$ for transversal matroids. Here, D is the maximum number of elements in any feasible subset of arms, and T is the horizon. Our technique is based on dynamic maintenance of an approximate maximum-weight basis over inner-product weights. Although the introduction of an approximate maximum-weight basis presents a challenge in regret analysis, we can still guarantee an upper bound on regret as tight as `CUCB` in the sense that it matches the gap-dependent lower bound by Kveton et al. (2014a) asymptotically.

1. Introduction

Matroid semi-bandits model many real-world tasks. An instance of matroid semi-bandit is described by $([K], \mathcal{X}, \boldsymbol{\mu})$, where $[K] \triangleq \{1, \cdots, K\}$ is the ground set, each $k \in [K]$ is associated with a probability distribution ν_k with mean μ_k, and $\mathcal{X} \subseteq \{0,1\}^K$ is the set of bases of a given matroid $\mathcal{M} = ([K], \mathcal{I})$ of rank D. At each round $t \in [T]$, the learner pulls an action $\boldsymbol{x}(t) \in \mathcal{X}$ and observes a *semi-bandit* feedback, i.e., $y_k(t) \sim \nu_k$ iff $x_k(t) = 1$. This formulation can be used to model online advertisting and news selection (Kale et al., 2010) with $\mathcal{M}$ as a uniform matroid. Ad placement (Bubeck et al., 2013; Streeter et al., 2009) and diversified recommendation (Abbassi et al., 2013) can be modeled with $\mathcal{M}$ as a partition matroid. Network routing (Kveton et al., 2014a) can be modeled with $\mathcal{M}$ as a graphical matroid. Task assignment (Chen et al., 2016) can be modeled with $\mathcal{M}$ as a transversal matroid.

Popular algorithms include Combinatorial Upper Confidence Bound (`CUCB`) (Gai et al., 2012; Chen et al., 2013; Kveton et al., 2014a; 2015), Combinatorial Thompson Sampling (`CTS`) (Wang and Chen, 2018; Kong et al., 2021; Perrault, 2022), and the instance-specifically optimal algorithm KL-based Efficient Sampling for Matroids (`KL-OSM`) (Talebi and Proutiere, 2016). All of these algorithms rely on a greedy algorithm (see Algorithm 1) to determine the action to be pulled. The greedy algorithm takes time at least $\Omega(K)$ and at most $\mathcal{O}(K(\log K + \mathcal{T}_{\mathrm{member}}))$, where $\mathcal{T}_{\mathrm{member}}$ is the time taken to determine whether $\boldsymbol{x} + \boldsymbol{e}_k \in \mathcal{I}$ for some $(\boldsymbol{x}, k) \in \mathcal{I} \times [K]$, and $\boldsymbol{e}_k$ is the k-th canonical unit vector. However, when the number K of arms is large, performing the greedy algorithm at each round can become expensive. There is a need to develop a matroid semi-bandit algorithm with per-round time complexity sublinear in K.

In this work, we present `FasterCUCB` (Algorithm 5), the first sublinear-time algorithm for matroid semi-bandit. The design of `FasterCUCB` is based on `CUCB`, but with a much faster sampling rule which takes time sublinear in K for many classes of matroids. For uniform matroids, partition matroids, and graphical matroids, it has per-round time complexity of $\mathcal{O}(D\,\mathrm{polylog}\,(K)\;\mathrm{polylog}\,(T))$, which is optimal up to a polylogarithmic factor as compared to the trivial lower bound of $\Omega(D)$. For transversal matroids, the per-round time complexity is $\mathcal{O}(D\sqrt{K}\,\mathrm{polylog}\,(T))$, which is still sublinear in K when $D = \mathcal{O}(K^{\frac{1}{2}-\epsilon})$ for any $\epsilon > 0$. `FasterCUCB` trades the accuracy for computational efficiency. In other words, the action computed by the sampling rule of `FasterCUCB` is an *approximation* to the optimal solution computed by the sampling rule of `CUCB`. This introduces difficulty in the regret analysis because prior analysis of `CUCB` (Kveton et al., 2014a) requires the exact solution. What is interesting is that we can still guarantee the same regret upper bound asymptotically as prior analysis of `CUCB`.

To develop a sublinear-time sampling rule, we present a dynamic algorithm for maintaining maximum-weight base

§Work done during an internship at CyberAgent. [1]EECS, KTH Royal Institue of Technology, Sweden [2]AI Lab, CyberAgent, Japan. Correspondence to: Ruo-Chun Tzeng <rubys88684@gmail.com>.

Proceedings of the 41st International Conference on Machine Learning, Vienna, Austria. PMLR 235, 2024.

	CUCB	FasterCUCB
Per-round Time Complexity	$\mathcal{O}(K(\log K + \mathcal{T}_{\text{member}}))$	$\mathcal{O}(D \,\text{polylog}\,(T)\, \mathcal{T}_{\text{update}}(\mathcal{A}))$
Uniform Matroid	$\mathcal{O}(K \log K)$	$\mathcal{O}(D \log K \,\text{polylog}\,(T))$
Partition Matroid	$\mathcal{O}(K \log K)$	$\mathcal{O}(D \log K \,\text{polylog}\,(T))$
Graphical Matroid	$\mathcal{O}(K \log K)$	$\mathcal{O}(D \,\text{polylog}\,(K)\, \text{polylog}\,(T))$
Transversal Matroid	$\mathcal{O}(K(\log K + DK))$	$\mathcal{O}(D\sqrt{K} \,\text{polylog}\,(T))$

Table 1. Per-round time complexity of `CUCB` (Kveton et al., 2014a) and `FasterCUCB` (Algorithm 5) for different classes of matroids. K is the number of arms and D is the maximum number of elements in any action in $\mathcal{X}$. $\mathcal{T}_{\text{member}}$ for different matroids is discussed in Appendix C. $\mathcal{T}_{\text{update}}(\mathcal{A})$ for different matroids is discussed in Section 3.

over inner product weights (Section 4). There have been many sublinear-time algorithms for dynamic maximum-weight base maintenance (see Section 3), which, however, may not be directly used in `FasterCUCB` because *all* arm weights representing the UCB index can change simultaneously at each round. Our insight for addressing this issue is that the UCB index of each arm k at round t can be decomposed into an inner product of the following two-dimensional vectors: (1) a *feature*, which depends on k and is supposedly a pair of the empirical reward estimate and radius of confidence interval, and (2) a *query*, which depends only on round t. Our proposed dynamic algorithm consists of two speeding-up techniques. One is *feature rounding*, which rounds each feature into a few bins so as to reduce the number of distinct features to consider. The other is the *minimum hitting set* technique, which allows us to compute a small number of queries in advance and correctly identify an (approximate) maximum-weight base for *any* query.

Sections are organized as follows. We introduce matroid semi-bandits and basic concepts in Section 2. We review relevant literature in Section 3. We develop a dynamic algorithm for maintaining a maximum-weight base over inner product weights in Section 4. We propose `FasterCUCB` based on the algorithms developed in Section 4 and analyzed its regret and time complexity in Section 5.

2. Preliminaries

We use $[n]$ to denote the set $\{1, \cdots, n\}$. We use $\boldsymbol{i}^\star(\boldsymbol{\mu})$ to denote any element in $\text{argmax}_{\boldsymbol{x}\in\mathcal{X}} \langle \boldsymbol{\mu}, \boldsymbol{x}\rangle$, and when it is clear from the context, we drop $\boldsymbol{\mu}$ from $\boldsymbol{i}^\star(\boldsymbol{\mu})$ and write $\boldsymbol{i}^\star$. We use $\text{supp}\,(\cdot)$ to denote the support set of a given vector. We use $\boldsymbol{e}_k$ to denote the vector with 1 only on the k-th row and 0's elsewhere, and use $\boldsymbol{0}_K$ to denote a K-dimensional vector with 0 on every row. We use $\log$ with base e. See Appendix A for a table of notation.

Matroid. A matroid is described by $\mathcal{M} \triangleq ([K], \mathcal{I})$, where $[K]$ is called the *ground set* and $\mathcal{I} \subseteq \{0,1\}^K$ is the set of *independent sets* satisfying (i) *hereditary property*, i.e., if $\text{supp}\,(\boldsymbol{y}) \subset \text{supp}\,(\boldsymbol{x})$ and $\boldsymbol{x} \in \mathcal{I}$, then $\boldsymbol{y} \in \mathcal{I}$; and (ii) *augmentation property*, i.e., if $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{I}$ and $\text{supp}\,(\boldsymbol{y}) \subset \text{supp}\,(\boldsymbol{x})$, then there exists $j \in \text{supp}\,(\boldsymbol{x}) \setminus \text{supp}\,(\boldsymbol{y})$ such that $\boldsymbol{y} + \boldsymbol{e}_j \in \mathcal{I}$. We said $\boldsymbol{x} \in \mathcal{I}$ is a basis if $\text{supp}\,(\boldsymbol{x})$ is *maximal*, i.e., there does not exist $\boldsymbol{y} \in \mathcal{I}$ such that $\text{supp}\,(\boldsymbol{x}) \subset \text{supp}\,(\boldsymbol{y})$. All bases have the same cardinality, which is called the *rank* of the matroid. For $\boldsymbol{v} \in \mathbb{R}_+^K$, a maximum-weight basis $\boldsymbol{i}^\star(\boldsymbol{v}) \in \text{argmax}_{\boldsymbol{x}\in\mathcal{X}} \sum_{k=1}^K v_k x_k$ can be found by a greedy algorithm (Algorithm 1) in $\mathcal{O}(K(\log K + \mathcal{T}_{\text{member}}))$ time, where $\mathcal{T}_{\text{member}}$ is the time taken for the membership oracle to determine whether $\boldsymbol{x} + \boldsymbol{e}_k \in \mathcal{I}$ for some $\boldsymbol{x} \in \mathcal{I}$ and some $k \in [K] \setminus \text{supp}\,(\boldsymbol{x})$.

Algorithm 1 A greedy maximum-weight basis algorithm

Input: $\boldsymbol{v} \in \mathbb{R}^K$ and the bases $\mathcal{X} \subseteq \{0,1\}^K$.
Sort $\boldsymbol{v}$ in non-increasing order: $v_{\gamma(1)} \geq \cdots \geq v_{\gamma(K)}$;
$\boldsymbol{x} = \boldsymbol{0}_K$; $i = 1$;
while $\|\boldsymbol{x}\|_0 < D$ **do**
  if $\boldsymbol{x} + \boldsymbol{e}_{\gamma(i)} \in \mathcal{I}$ **then**
    $\boldsymbol{x} \leftarrow \boldsymbol{x} + \boldsymbol{e}_{\gamma(i)}$
  $i \leftarrow i + 1$;
end

Matroid semi-bandits. An instance of matroid semi-bandit is described by $([K], \mathcal{X}, \boldsymbol{\mu})$, where $[K]$ is the ground set, $\mathcal{X} \subseteq \{0,1\}^K$ is the set of bases of the given matroid $\mathcal{M} \triangleq ([K], \mathcal{I})$ of rank D, and $\mathcal{I} \subseteq \{0,1\}^K$ is the set of independent sets. Each $k \in [K]$ is associated with a distribution ν_k with mean μ_k. The learner knows the matroid $\mathcal{M}$, and aims to learn the best action $\boldsymbol{i}^\star(\boldsymbol{\mu}) \in \text{argmax}_{\boldsymbol{x}\in\mathcal{X}} \langle \boldsymbol{\mu}, \boldsymbol{x}\rangle$ by playing a game with the environment: At each round $t \in \mathbb{N}$, the learner plays an action $\boldsymbol{x}(t)$, and the environment draws a noisy reward $y_k(t) \sim \nu_k$ for each arm $k \in [K]$ and reveals $y_k(t)$ to the learner iff $k \in \text{supp}\,(\boldsymbol{x}(t))$. We assume arms' rewards are bounded:

Assumption 2.1. Assume the support of each arm ν_k is a subset of $[a, b]$ and $0 < a < b < \infty$.

The performance is measured by expected regret:

$$R(T) \triangleq T \langle \boldsymbol{\mu}, \boldsymbol{i}^\star(\boldsymbol{\mu})\rangle - \mathbb{E}\left[\sum_{t=1}^T \langle \boldsymbol{\mu}, \boldsymbol{x}(t)\rangle\right],$$

which is the difference between the expected cumulative reward of the learner and that of an algorithm who knows $\boldsymbol{\mu}$ and always selects the best action $\boldsymbol{i}^\star(\boldsymbol{\mu})$.

Common classes of matroids. Refer to Chapter 39 (Schrijver, 2003) or Chapter 1 (Oxley, 2011) for more details.

- A *uniform matroid* $([K], \mathcal{I})$ of rank D has independent sets $\mathcal{I} = \{S \subseteq [K] : |S| \leq D\}$ and the bases $\mathcal{X}$ consist of subsets whose cardinalities are exactly D.
- A *partition matroid* $([K], \mathcal{I})$ of rank D is given a partition $S_1, \cdots, S_D$ of the ground set $[K]$, the independent sets $\mathcal{I} = \{S : |S \cap S_i| \leq 1, \forall i \in [D]\}$, and the bases $\mathcal{X}$ are subsets that choose exactly one element from each of the D sets.
- A *graphical matroid* is given a graph $G = (V, E)$ with K edges, the bases $\mathcal{X}$ consist of all spanning forests in G, and the rank D is $|V|$ minus the number of connected components in G.
- A *transversal matroid* is given a bipartite graph $G = ([K] \cup V, E)$ with a bipartition $([K], V)$, $|V| \leq K$, the independent sets $\mathcal{I}$ consist of $S \subseteq [K]$ such that there is a matching of S to $|S|$ vertices in V, and $\mathcal{X}$ is the set of endpoints in $[K]$ of all maximum matchings in G.

We discuss the query time $\mathcal{T}_{\text{member}}$ of membership oracle in Appendix C. For more examples on semi-bandits under different matroid constraints, we refer the readers to (Kveton et al., 2014a) for linear matroids, and (Kveton et al., 2014b) for polymatroid semi-bandits.

CUCB. The action selected by `CUCB` (Gai et al., 2012; Chen et al., 2013; Kveton et al., 2014b) at round $t \in \mathbb{N}$ is:

$$\boldsymbol{x}(t) \in \operatorname*{argmax}_{\boldsymbol{x} \in \mathcal{X}} \sum_{k=1}^{K} \left(\hat{\mu}_k(t-1) + \frac{\lambda_t}{\sqrt{N_k(t-1)}} \right) x_k, \tag{1}$$

where $\hat{\mu}_k(t) \triangleq \frac{1}{N_k(t)} \sum_{s=1}^{t} y_k(s) \mathbb{1}\{x_k(s) = 1\}$ is the empirical reward estimate, $N_k(t) \triangleq \sum_{s=1}^{t} \mathbb{1}\{x_k(s) = 1\}$ is the number of pulls of arm k, and $\lambda_t > 0$ controls the confidence width. The value $\hat{\mu}_k(t-1) + \frac{\lambda_t}{\sqrt{N_k(t-1)}}$ is called the *UCB index* of arm k. In (Kveton et al., 2014a), Eq. (1) is solved by a $\mathcal{O}(K(\log K + \mathcal{T}_{\text{member}}))$-time greedy algorithm shown in Algorithm 1. In Section 4, we will develop a faster algorithm for solving Eq. (1) with the following reformulation:

$$\boldsymbol{x}(t) \in \operatorname*{argmax}_{\boldsymbol{x} \in \mathcal{X}} \sum_{k=1}^{K} \langle \boldsymbol{f}_k, \boldsymbol{q} \rangle \, x_k,$$

where $\boldsymbol{f}_k = (\hat{\mu}_k(t-1), \frac{1}{\sqrt{N_k(t-1)}})$ and $\boldsymbol{q} = (1, \lambda_t)$.

3. Related Works

Semi-bandits and sublinear-time bandits. We provide an extensive survey on related bandit literature in Appendix B. To summarize here, for semi-bandit algorithms, `CUCB` (Gai et al., 2012; Chen et al., 2013; Kveton et al., 2014a), `CTS` (Wang and Chen, 2018; Kong et al., 2021; Perrault, 2022) and `KL-OSM` (Talebi and Proutiere, 2016) all rely on a $\mathcal{O}(K(\log K + \mathcal{T}_{\text{member}}))$-time greedy algorithm to compute the action to be pulled. In contrast, our `FasterCUCB`, as far as we know, is the first semi-bandit algorithm having per-round time complexity of $o(K)$. For linear bandits, there exist several works (Jun et al., 2017; Yang et al., 2022) on reducing the per-round complexity to be sublinear in the number of arms. But, their results transferred to our setting are worse than what we have obtained both in terms of regret bound and the time complexity (see the discussion in Appendix B).

Dynamic maintenance of maximum-weight base of a matroid Here, we review existing dynamic algorithms for maintaining a maximum-weight base of a matroid. In a standard (fully-)dynamic setting, we are given a weighted matroid $\mathcal{M} = ([K], \mathcal{I})$, where each arm's weight dynamically changes over time in an online manner. The objective is to maintain any (exact or approximate) maximum-weight basis of $\mathcal{M}$ over up-to-date arm weights as efficiently as possible. We use $\mathcal{T}_{\text{update}}(\mathcal{A})$ to denote the time complexity of a dynamic algorithm $\mathcal{A}$ required for updating an (approximate) maximum-weight base according to the change of a single arm weight. The best-known bound of $\mathcal{T}_{\text{update}}(\mathcal{A})$ for each matroid class is summarized as follows: For graphic matroids, a maximum-weight basis can be updated in $\mathcal{O}(\sqrt{K})$ worst-case time (Frederickson, 1985; Eppstein et al., 1997) and in $\mathcal{O}(\text{polylog}\,(K))$ amortized time (Holm et al., 2001). For laminar matroids (which include uniform and partition matroids as special cases), the worst-case time complexity for exact dynamic algorithms is bounded by $\mathcal{O}(\log K)$ (Henzinger et al., 2023). For transversal matroids, a $\mathcal{O}(K^{1.407})$-time exact dynamic algorithm is known (van den Brand et al., 2019), while a $(1+\eta)$-approximation dynamic algorithm runs in $\mathcal{O}(\eta^{-2}\sqrt{K})$ time (Gupta and Peng, 2013). We can safely assume that, after updating multiple arm weights, $\mathcal{A}$ returns an (approximate) maximum-weight basis in $\mathcal{O}(D)$ time, where D is the rank of a matroid.

4. Dynamic Maintenance of Maximum-weight Base over Inner Product Weight

In this section, we develop a sublinear-time sampling rule, which is used as a subroutine in `FasterCUCB`. Recall that *any* static algorithm that solves the linear maximization of Eq. (1) from scratch requires at least $\Omega(K)$ time, which is computationally expensive. To circumvent this issue, we present a dynamic algorithm for maintaining an (approximate) maximum-weight base of a matroid where arm weights change over time. The next subsection begins with formalizing the problem setting.

4.1. Problem Setting and Technical Result

Consider the following problem setting: Let $\mathcal{M} = ([K], \mathcal{I})$ be a matroid of rank D over K arms, and $\mathcal{X}$ be the set of its bases. Each arm $k \in [K]$ has a (nonnegative) two-dimensional vector $\boldsymbol{f}_k = (\alpha_k, \beta_k) \in \mathbb{R}_+^2$ referred to as a *feature*, which may change as time goes by. Given a (nonnegative) two-dimensional vector $\boldsymbol{q} \in \mathbb{R}_+^2$ as a *query*, we are required to find any (approximate) maximum-weight base of $\mathcal{M}$, where arm k's weight is given by projecting its feature onto $\boldsymbol{q}$, i.e., $\langle \boldsymbol{f}_k, \boldsymbol{q} \rangle$.

Observe that in the matroid semi-bandit setting, each arm k's feature $\boldsymbol{f}_k = (\alpha_k, \beta_k)$ corresponds to a pair of the empirical reward estimate $\alpha_k = \hat{\mu}_k(t-1)$ and radius $\beta_k = \frac{1}{\sqrt{N_k(t-1)}}$ of confidence interval, and a query is $\boldsymbol{q} = (1, \lambda_t)$ at round t, both of which change over rounds.

Hereafter, we make the following two assumptions.

Assumption 4.1. Lower and upper bounds, denoted by α_{lb} and α_{ub} (resp. β_{lb} and β_{ub}), on the possible positive values of α_k's (resp. β_k's) at anytime are known; namely, it always holds that $\alpha_k \in \{0\} \cup [\alpha_{\text{lb}}, \alpha_{\text{ub}}]$ and $\beta_k \in \{0\} \cup [\beta_{\text{lb}}, \beta_{\text{ub}}]$. The precise values of these bounds will be discussed in Section 5.

Assumption 4.2. There exists a dynamic algorithm $\mathcal{A}$ for maintaining a $(1+\eta)$-approximate maximum-weight base of $\mathcal{M}$ with arm weights changing over time, where parameter $\eta \in (0,1)$ specifies the approximation guarantee. Denote by $\mathcal{T}_{\text{init}}(\mathcal{A}; \eta)$ and $\mathcal{T}_{\text{update}}(\mathcal{A}; \eta)$ the time complexity of $\mathcal{A}$ required for initializing the data structure and updating a single arm's weight, respectively. We can safely assume that after updating multiple arm weights, $\mathcal{A}$ returns a maximum-weight base in $\mathcal{O}(D)$ time. See Section 3 for existing implementations.

Our dynamic algorithm is parameterized by a *precision parameter* $\epsilon \in (0,1)$, and consists of the following three procedures:

INITIALIZE: Given lower and upper bounds $[\alpha_{\text{lb}}, \alpha_{\text{ub}}]$ and $[\beta_{\text{lb}}, \beta_{\text{ub}}]$ as in Assumption 4.1, K features $(\alpha_k, \beta_k)_{k \in [K]}$, a matroid $\mathcal{M} = ([K], \mathcal{I})$, a dynamic algorithm $\mathcal{A}$ for maximum-weight base maintenance as in Assumption 4.2, and a precision parameter ϵ, this procedure initializes the data structure used in the remaining two procedures.

FIND-BASE: Given a query $\boldsymbol{q}$, this procedure is supposed to return a $(1+\epsilon)$-approximate maximum-weight base of $\mathcal{M}$, where arm k's weight is defined as $\langle \boldsymbol{f}_k, \boldsymbol{q} \rangle$ for the up-to-date k's feature $\boldsymbol{f}_k$.

UPDATE-FEATURE: Given an arm k and a new feature $\boldsymbol{f}'_k$, this procedure reflects the change of arm k's feature on the data structure.

Remark 4.3. Our problem setting is different from a canonical setting of dynamic maximum-weight base maintenance in a sense that the arm weights are revealed when a query is issued in **FIND-BASE**. Consequently, existing dynamic algorithms may not be used directly.

The technical result in this section is stated below.

Theorem 4.4 ($*$). *There exist implementations of* **INITIALIZE**, **FIND-BASE**, *and* **UPDATE-FEATURE** *such that the following are satisfied:* **FIND-BASE** *always returns a $(1+\epsilon)$-approximate maximum-weight base of a matroid $\mathcal{M}$ with arm k's weight defined as $\langle \boldsymbol{f}_k, \boldsymbol{q} \rangle$ for an up-to-date k's feature $\boldsymbol{f}_k$ and a query $\boldsymbol{q}$. Moreover,* **INITIALIZE** *runs in $\mathcal{O}(K + \text{poly}(W) \cdot \mathcal{T}_{\text{init}}(\mathcal{A}; \frac{\epsilon}{3}))$ time,* **FIND-BASE** *runs in $\mathcal{O}(\text{poly}(W) + D)$ time, and* **UPDATE-FEATURE** *runs in $\mathcal{O}(\text{poly}(W) \cdot \mathcal{T}_{\text{update}}(\mathcal{A}; \frac{\epsilon}{3}))$ time, where*

$$W = \mathcal{O}\left(\epsilon^{-1} \cdot \log\left(\frac{\alpha_{ub}}{\alpha_{lb}} \cdot \frac{\beta_{ub}}{\beta_{lb}}\right)\right). \tag{2}$$

Remark 4.5. The proof of Theorem 4.4 can be easily adapted to the case when (the update procedure of) dynamic algorithm $\mathcal{A}$ has only amortized complexity. In such case, Theorem 4.4 holds in the *amortized* sense rather than the *worst-case* sense.

The remainder of this section is organized as follows: In Section 4.2, we apply a rounding technique to arm features to reduce the number of distinct features to consider, in Section 4.3, we investigate the representability of permutations induced by inner product weights to deal with multiple queries efficiently, and Section 4.4 finally develops our dynamic algorithm for maximum-weight base maintenance. All proofs of the lemmas appearing in this section are deferred to Appendix D.

4.2. Rounding Arm Features

Here, we apply a rounding technique to arm features so as to reduce the number of distinct features to consider. Hereafter, let $\eta \triangleq \frac{\epsilon}{3}$, so that $(1+\eta)^2 \leq 1+\epsilon$ for $\epsilon \in (0,1)$. Define

$$\begin{aligned} W &\triangleq \max\left\{\left\lceil \log_{1+\eta}\left(\frac{\alpha_{\text{ub}}}{\alpha_{\text{lb}}}\right)\right\rceil, \left\lceil \log_{1+\eta}\left(\frac{\beta_{\text{ub}}}{\beta_{\text{lb}}}\right)\right\rceil\right\}, \\ \mathbb{W} &\triangleq \{-\infty\} \cup [W] = \{-\infty, 1, 2, 3, \dots, W\}. \end{aligned} \tag{3}$$

Since any features are within $(\{0\} \cup [\alpha_{\text{lb}}, \alpha_{\text{ub}}]) \times (\{0\} \cup [\beta_{\text{lb}}, \beta_{\text{ub}}])$ as guaranteed by Assumption 4.2, we shall partition the possible region of the features into $|\mathbb{W}|^2$ bins. For each $q, r \in \mathbb{W}$, define $\mathrm{BIN}_{q,r} \subset \mathbb{R}_+^2$ as

$$\begin{aligned}\mathrm{BIN}_{q,r} \triangleq & (\alpha_{\text{lb}}(1+\eta)^{q-1}, \alpha_{\text{lb}}(1+\eta)^{q}] \\ & \times (\beta_{\text{lb}}(1+\eta)^{r-1}, \beta_{\text{lb}}(1+\eta)^{r}],\end{aligned} \tag{4}$$

$$\text{where } (\alpha_{\text{lb}}(1+\eta)^{-\infty}, \alpha_{\text{lb}}(1+\eta)^{-\infty}] \triangleq \{0\}, \tag{5}$$

$$(\beta_{\text{lb}}(1+\eta)^{-\infty}, \beta_{\text{lb}}(1+\eta)^{-\infty}] \triangleq \{0\}. \tag{6}$$

Note that these bins are pairwise disjoint, and that $\cup_{q,r\in\mathbb{W}} \mathrm{BIN}_{q,r}$ covers $(\{0\} \cup [\alpha_{\text{lb}}, \alpha_{\text{ub}}]) \times (\{0\} \cup [\beta_{\text{lb}}, \beta_{\text{ub}}])$; i.e., any possible feature belongs to a unique $\mathrm{BIN}_{q,r}$. For each $q, r \in \mathbb{W}$, let $\mathrm{dom}_{q,r} \in \mathbb{R}_+^2$ denote the unique *dominating point* of $\mathrm{BIN}_{q,r}$; namely,

$$\mathrm{dom}_{q,r} \triangleq (\alpha_{\text{lb}}(1+\eta)^{q}, \beta_{\text{lb}}(1+\eta)^{r}). \tag{7}$$

For any feature $\boldsymbol{f}_k = (\alpha_k, \beta_k)$, we will use $\mathrm{dom}(\boldsymbol{f}_k) = \mathrm{dom}(\alpha_k, \beta_k)$ to denote the dominating point $\mathrm{dom}_{q,r}$ such that $\boldsymbol{f}_k \in \mathrm{BIN}_{q,r}$. See Figure 1 in Appendix D for illustration of $\mathrm{BIN}_{q,r}$'s, $\mathrm{dom}_{q,r}$'s, and $\mathrm{dom}(\boldsymbol{f}_k)$'s.

Observe easily that for any feature $\boldsymbol{f}_k \in \mathbb{R}_+^2$ and query $\boldsymbol{q} \in \mathbb{R}_+^2$,

$$\frac{1}{1+\eta} \cdot \langle \mathrm{dom}(\boldsymbol{f}_k), \boldsymbol{q} \rangle < \langle \boldsymbol{f}_k, \boldsymbol{q} \rangle \le \langle \mathrm{dom}(\boldsymbol{f}_k), \boldsymbol{q} \rangle. \tag{8}$$

By Eq. (8), we can replace each arm's feature by its dominating point without much deteriorating the quality of the (approximate) maximum-weight base, as shown below.

Lemma 4.6 $(*)$**.** *Let $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_K \in (\{0\} \cup [\alpha_{lb}, \alpha_{ub}]) \times (\{0\} \cup [\beta_{lb}, \beta_{ub}])$ be K features, $\boldsymbol{q} \in \mathbb{R}_+^2$ be a query, and $\boldsymbol{x}^*_{\mathrm{dom}}$ be a $(1+\eta)$-approximate maximum-weight base of matroid $\mathcal{M}$ with arm k's weight defined as $\langle \mathrm{dom}(\boldsymbol{f}_k), \boldsymbol{q} \rangle$. Then, for any base $\boldsymbol{x}$ of $\mathcal{M}$, it holds that*

$$\sum_{k \in \mathit{supp}(\boldsymbol{x}^*_{\mathrm{dom}})} \langle \boldsymbol{f}_k, \boldsymbol{q} \rangle \ge \frac{1}{1+\epsilon} \cdot \sum_{k \in \mathit{supp}(\boldsymbol{x})} \langle \boldsymbol{f}_k, \boldsymbol{q} \rangle. \tag{9}$$

In particular, $\boldsymbol{x}^*_{\mathrm{dom}}$ is a $(1+\epsilon)$-approximate maximum-weight base with arm k's weight defined as $\langle \boldsymbol{f}_k, \boldsymbol{q} \rangle$.

4.3. Handling Multiple Queries

From weighting to permutation. Now we deal with multiple queries. Our idea is that, if two queries $\boldsymbol{q}_1, \boldsymbol{q}_2 \in \mathbb{R}_+^2$ are "very close" to each other, then they should derive the common maximum-weight base (provided that features are fixed). This intuition can be justified with respect to *orderings* of arms. For two total orders $\preceq$ and $\preceq^\circ$ over $[K]$, we say that $\preceq^\circ$ is *consistent with* $\preceq$ if $k \preceq^\circ k'$ implies $k \preceq k'$ for any $k \neq k'$. The following fact is easy to confirm:

Lemma 4.7 $(*)$**.** *Let $\boldsymbol{w} = (w_1, \ldots, w_K) \in \mathbb{R}_+^K$ be K arm weights and $\preceq$ be a total order over $[K]$ such that $k \succeq k'$ if and only if $w_k \ge w_{k'}$. Let $\preceq^\circ$ be a total order over $[K]$ that is consistent with $\preceq$. If $\boldsymbol{x}^\circ$ is a base of matroid $\mathcal{M}$ obtained by running the greedy algorithm over any ordering of $[K]$ consistent with $\preceq^\circ$, it is a maximum-weight base of $\mathcal{M}$ over arm k's weight w_k; namely, for any base $\boldsymbol{x}$ of $\mathcal{M}$,*

$$\langle \boldsymbol{x}^\circ, \boldsymbol{w} \rangle \ge \langle \boldsymbol{x}, \boldsymbol{w} \rangle. \tag{10}$$

Lemma 4.7 implies that any maximum-weight base can be obtained by running the greedy algorithm over some total order $\preceq^\circ$; moreover, we can safely assume that $\preceq^\circ$ is *strict* (i.e., $k \prec^\circ k'$ or $k \succ^\circ k'$ for all $k \neq k'$), or equivalently, a *permutation* over $[K]$. Our strategy for dealing with multiple queries is: (1) we enumerate all possible permutations in advance, and (2) we guess a permutation consistent with the arm weights determined based on a query. To this end, the following question arises: *What kind of and how many permutations are representable given a fixed set of features?*

Characterizing representable permutations. To answer the above question, we characterize representable permutations. Hereafter, let $\mathfrak{S}_K$ denote the set of all permutations over $[K]$, and $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_K \in \mathbb{R}_+^2$ be any fixed, distinct K features. We say that a query $\boldsymbol{q} \in \mathbb{R}^2$ over $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_K$ *represents* a permutation $\pi \in \mathfrak{S}_K$ if $\langle \boldsymbol{f}_{\pi(i)}, \boldsymbol{q} \rangle > \langle \boldsymbol{f}_{\pi(j)}, \boldsymbol{q} \rangle$ for all $1 \le i < j \le K$,[1] and that π is *representable* if such $\boldsymbol{q}$ exists.

For a permutation $\pi \in \mathfrak{S}_K$ to be representable, we wish for some query $\boldsymbol{q} \in \mathbb{R}^2$ to ensure that for any $i < j$, arm $\pi(i)$'s weight is (strictly) higher than arm $\pi(j)$'s weight. This requirement is equivalent to $\langle \boldsymbol{f}_{\pi(i)} - \boldsymbol{f}_{\pi(j)}, \boldsymbol{q} \rangle > 0$; thus, if the following system of linear inequalities is feasible, any of its solutions $\boldsymbol{q}$ represents π:

$$\langle \boldsymbol{f}_{\pi(i)} - \boldsymbol{f}_{\pi(j)}, \boldsymbol{q} \rangle > 0 \ \text{ for all } \ 1 \le i < j \le K. \tag{11}$$

Observe now that the above system is feasible *if and only if* the intersection of $\mathcal{P}_{i,j}$ for all $i < j$ is nonempty, where $\mathcal{P}_{i,j}$ is an open half-plane defined as

$$\mathcal{P}_{i,j} \triangleq \{\boldsymbol{q} \in \mathbb{R}^2 : \langle \boldsymbol{f}_{\pi(i)} - \boldsymbol{f}_{\pi(j)}, \boldsymbol{q} \rangle > 0\}. \tag{12}$$

Because each $\mathcal{P}_{i,j}$ is obtained by dividing $\mathbb{R}^2$ by a unique line that is orthogonal to line $\overleftrightarrow{\boldsymbol{f}_{\pi(i)}\boldsymbol{f}_{\pi(j)}}$ and intersects the origin $\mathbf{0}$, the set of feasible solutions for Eq. (11) is equal to (the interior of) a *polyhedral cone* defined by the boundaries of a particular pair of $\mathcal{P}_{i,j}$'s.

Here, we characterize the representable permutations by the concept of arrangement of lines. Let $\mathcal{L} \triangleq \{\boldsymbol{l}_1, \ldots, \boldsymbol{l}_{\binom{K}{2}}\}$ denote $\binom{K}{2}$ lines, each of which is orthogonal to line $\overleftrightarrow{\boldsymbol{f}_k\boldsymbol{f}_{k'}}$ for some $k \neq k'$ and intersects $\mathbf{0}$. Given such $\mathcal{L}$, a *cell* $\mathcal{C}$

[1] This definition does not allow "ties"; i.e., no pair $k \neq k'$ satisfies $\langle \boldsymbol{f}_k, \boldsymbol{q} \rangle = \langle \boldsymbol{f}_{k'}, \boldsymbol{q} \rangle$.

in arrangement of $\mathcal{L}$ is defined as a maximum connected region of $\mathbb{R}^2$ that does not intersect with $\mathcal{L}$ (which is the interior of a polyhedral cone). Then, for each cell $\mathcal{C}$, every query $\boldsymbol{q}$ in $\mathcal{C}$ represents the same permutation $\pi_{\mathcal{C}} \in \mathfrak{S}_K$ depending only on $\mathcal{C}$; namely, *there is a bijection between the representable permutations and the cells in arrangement of* $\mathcal{L}$. See Figure 2 in Appendix D for illustration.

With this connection in mind, we demonstrate that reserving a single vector for each cell suffices to cover all representable permutations. A *minimum hitting set* of the cells in arrangement of $\mathcal{L}$ is defined as any minimum set $\mathcal{H}$ of vectors in $\mathbb{R}^2$ such that $\mathcal{H}$ and each cell have a non-empty intersection.

Lemma 4.8 ($*$). *Let $\mathcal{H}$ be a minimum hitting set of the cells in arrangement of $\mathcal{L}$. Then, for any query $\boldsymbol{q} \in \mathbb{R}^2$, there exists a vector $\boldsymbol{h} \in \mathcal{H}$ such that for any $k \neq k'$,*

$$\langle \boldsymbol{f}_k, \boldsymbol{h} \rangle > \langle \boldsymbol{f}_{k'}, \boldsymbol{h} \rangle \implies \langle \boldsymbol{f}_k, \boldsymbol{q} \rangle \geq \langle \boldsymbol{f}_{k'}, \boldsymbol{q} \rangle. \quad (13)$$

Lemma 4.8 along with Lemma 4.7 ensure that for any query $\boldsymbol{q} \in \mathbb{R}^2$, there is a vector $\boldsymbol{h}$ in $\mathcal{H}$ such that a maximum-weight base with arm k's weight $\langle \boldsymbol{f}_k, \boldsymbol{h} \rangle$ is a maximum-weight base with arm k's weight $\langle \boldsymbol{f}_k, \boldsymbol{q} \rangle$.

Generating a minimum hitting set. Subsequently, we generate a minimum hitting set. One may think that it requires exponentially long time because the number of permutations in $\mathfrak{S}_K$ is $K!$. However, it turns out that the number of cells in arrangement of $\mathcal{L}$ is $\mathcal{O}(K^2)$ (i.e., so is the number of representable permutations), and a minimum hitting set can be constructed in poly(K) time.

Lemma 4.9 ($*$). *The number of cells in arrangement of $\mathcal{L}$ is at most $\mathcal{O}(K^2)$. Moreover, we can generate a minimum hitting set in poly(K) time (by using Algorithm 6 in Appendix D).*

4.4. Putting It All Together: Algorithm Description and Complexity

We are now ready to implement the three procedures. We here stress that applying either of the feature rounding or minimum hitting set technique *separately* does not make sense: On one hand, if we only apply feature rounding, we would have to recompute each arm's weight *every time* a query is issued, which is expensive. On the other hand, if the minimum hitting set technique is only applied (to raw features $\boldsymbol{f}_k$'s), then a minimum hitting set $\mathcal{H}$ cannot be constructed in advance due to a dynamic nature of features, and its size would be $\mathcal{O}(K^2)$, which is prohibitive.

By applying both techniques, (1) we know *a priori* the set of possible dominating points, whose size $\mathcal{O}(W^2)$ depends *only* on W; moreover, (2) we can create a minimum hitting set $\mathcal{H}$ of size $\mathcal{O}(W^4)$ beforehand at initialization. Pseudocodes of **INITIALIZE**, **FIND-BASE**, and **UPDATE-FEATURE** are described in Algorithms 2 to 4, respectively. The proof of Theorem 4.4 follows from Lemmas 4.6 to 4.9, whose details are deferred to Appendix D.

In **INITIALIZE**, we construct a hitting set $\mathcal{H}$ of $\mathrm{dom}_{q,r}$'s and $\frac{1}{1+\eta} \cdot \mathrm{dom}_{q,r}$'s rather than solely of $\mathrm{dom}_{q,r}$'s, which incurs a constant-factor blowup in the time complexity. Though this change is not needed in the proof of Theorem 4.4, the following immediate corollary of Lemma 4.8 is crucial in the regret analysis of Section 5.

Corollary 4.10 ($*$). *Let $\mathcal{H}$ be a minimum hitting set constructed in Algorithm 2. Then, for any query $\boldsymbol{q} \in \mathbb{R}^2$, there exists a vector $\boldsymbol{h} \in \mathcal{H}$ such that for any* $\mathrm{dom} = \mathrm{dom}_{q,r}$ *and* $\mathrm{dom}' = \mathrm{dom}_{q',r'}$ *with $q, r, q', r' \in \mathbb{W}$,*

$$\langle \mathrm{dom}, \boldsymbol{h} \rangle > \langle \mathrm{dom}', \boldsymbol{h} \rangle \implies \langle \mathrm{dom}, \boldsymbol{q} \rangle \geq \langle \mathrm{dom}', \boldsymbol{q} \rangle,$$

$$\langle \mathrm{dom}, \boldsymbol{h} \rangle > \frac{\langle \mathrm{dom}', \boldsymbol{h} \rangle}{1+\eta} \implies \langle \mathrm{dom}, \boldsymbol{q} \rangle \geq \frac{\langle \mathrm{dom}', \boldsymbol{q} \rangle}{1+\eta}.$$

Algorithm 2 INITIALIZE.

Input: lower and upper bounds $[\alpha_{\mathrm{lb}}, \alpha_{\mathrm{ub}}]$ and $[\beta_{\mathrm{lb}}, \beta_{\mathrm{ub}}]$; K features $(\boldsymbol{f}_k)_{k \in [K]}$; precision parameter $\epsilon \in (0, 1)$.
Define $\mathbb{W}$ by Eq. (3);
for *each* $q, r \in \mathbb{W}$ **do**
  Define $\mathrm{BIN}_{q,r}$ and $\mathrm{dom}_{q,r}$ by Eqs. (4) and (7);
end
Define $\eta \triangleq \frac{\epsilon}{3}$;
Construct a minimum hitting set $\mathcal{H}$ of size $\mathcal{O}(W^4)$ for $\mathrm{dom}_{q,r}$'s and $\frac{1}{1+\eta} \cdot \mathrm{dom}_{q,r}$'s by Lemma 4.9;
for *each* $\boldsymbol{h} \in \mathcal{H}$ **do**
  Create an instance $\mathcal{A}_{\boldsymbol{h}}$ of dynamic maximum-weight base algorithm with precision parameter $\eta \triangleq \frac{\epsilon}{3}$, $\mathcal{M}$, and arm k's weight $\langle \mathrm{dom}(\boldsymbol{f}_k), \boldsymbol{h} \rangle$;
end

Algorithm 3 FIND-BASE.

Input: query $\boldsymbol{q} \in \mathbb{R}^2_+$.
Find $\boldsymbol{h} \in \mathcal{H}$ s.t. $\boldsymbol{q}$ and $\boldsymbol{h}$ belong to (the closure of) the same cell in arrangement of $\mathcal{V}$;
Call $\mathcal{A}_{\boldsymbol{h}}$ and return the maximum-weight base $\boldsymbol{x}^\circ$ of $\mathcal{M}$ with arm k's weight $\langle \mathrm{dom}(\boldsymbol{f}_k), \boldsymbol{h} \rangle$;

Algorithm 4 UPDATE-FEATURE.

Input: arm $k \in [K]$; new feature $\boldsymbol{f}'_k \in \mathbb{R}^2_+$.
for *each* $\boldsymbol{h} \in \mathcal{H}$ **do**
  Change arm k's weight stored in $\mathcal{A}_{\boldsymbol{h}}$ to $\langle \mathrm{dom}(\boldsymbol{f}'_k), \boldsymbol{h} \rangle$;
end

5. Our Proposed Algorithm: `FasterCUCB`

In this section, we present `FasterCUCB` in Algorithm 5, which uses procedures introduced in Section 4.

The purpose of initialization procedure is to ensure each arm is pulled at least once. It takes at most K rounds, and in each round, the computation of $\boldsymbol{i}^\star(\boldsymbol{e}_k)$ takes $\mathcal{O}(K \cdot \mathcal{T}_{\text{member}})$ time because the permutation γ in Algorithm 1 can be specified explicitly as $\gamma(j) = k$ if $j = 1$, $\gamma(j) = j-1$ if $2 \leq j < k$, and $\gamma(j) = j+1$ if $j \geq k$. So, it only require to compute at most K membership tests. After the initialization, the computation of each round t consists of one call to **FIND-BASE** for computing the action $\boldsymbol{x}(t)$, the update of $\hat{\mu}_k(t)$ and $N_k(t)$ for each $k \in \operatorname{supp}(\boldsymbol{x}(t))$, and one call to **UPDATE-FEATURE** for updating the feature of each arm $k \in \operatorname{supp}(\boldsymbol{x}(t))$ stored in the instances of the dynamic maxium-weight base algorithm.

Algorithm 5 `FasterCUCB`

Input: the total number of rounds T, λ_t, and $m \in \mathbb{N}$
Initialization:
  $t = 0$;
  while $\exists k \in [K]$ *such that* $N_k(t) = 0$ **do**
    Pull $\boldsymbol{i}^\star(\boldsymbol{e}_k)$; $t = t+1$;
  end
  INITIALIZE$\left(a, b, \frac{1}{\sqrt{T}}, 1, (\hat{\mu}_k(t), N_k(t))_{k \in [K]}, \frac{1}{\log^m T}\right)$
while $t < T$ **do**
  $\boldsymbol{x}(t) \leftarrow$**FIND-BASE**$((1, \lambda_t))$;
  Pull $\boldsymbol{x}(t)$ and receive $y_k(t) \sim \nu_k$ for each $k \in \operatorname{supp}(\boldsymbol{x}(t))$;
  for $k \in$ *supp* $(\boldsymbol{x}(t))$ **do**
    $N_k(t) \leftarrow N_k(t-1) + 1$;
    $\hat{\mu}_k(t) \leftarrow \frac{t-1}{t}\hat{\mu}_k(t-1) + \frac{1}{t} y_k(t)$;
    UPDATE-FEATURE$\left(k, \left(\mu_k(t), \frac{1}{\sqrt{N_k(t)}}\right)\right)$;
  end
  $t = t+1$;
end

5.1. Per-round Time Complexity

By Theorem 4.4, one call to **FIND-BASE** takes $\mathcal{O}(\operatorname{poly}(W) + D)$ and D calls to **UPDATE-FEATURE** take $\mathcal{O}(D \operatorname{poly}(W)\, \mathcal{T}_{\text{update}}(\mathcal{A}; \frac{\epsilon}{3}))$. Since

$$W = \mathcal{O}\left(\log^m T \log\left(\frac{b}{a}\sqrt{T}\right)\right) = \mathcal{O}\left(\log^{m+1} T\right),$$

the per-round time complexity of Algorithm 5 is

$$\mathcal{O}\left(D \operatorname{polylog}(T)\, \mathcal{T}_{\text{update}}\left(\mathcal{A}; \frac{\epsilon}{3}\right)\right)$$

. Here, we will set $\epsilon = \frac{1}{\log^m T}$ for the regret analysis.

5.2. Regret Upper Bound

Notation. Fix $\boldsymbol{\mu} \in \Lambda$ and $\boldsymbol{i}^\star \in \operatorname{argmax}_{\boldsymbol{x} \in \mathcal{X}} \langle \boldsymbol{\mu}, \boldsymbol{x} \rangle$. We introduce a few notation. Let $\{\bar{j}\}_{j=1}^D$ be the permutation of $\operatorname{supp}(\boldsymbol{i}^\star)$ such that $\mu_{\bar{1}} \geq \cdots \geq \mu_{\bar{D}}$. Define $\triangle_{j,k} \triangleq \mu_j - \mu_k$ and $d_k \triangleq \max\{j \in [D] : \triangle_{\bar{j},k} > 0\}$ for $j \in \operatorname{supp}(\boldsymbol{i}^\star)$ and $k \notin \operatorname{supp}(\boldsymbol{i}^\star)$, and $\triangle_{\min} \triangleq \min_{k \notin \operatorname{supp}(\boldsymbol{i}^\star)} \triangle_{\overline{d_k},k}$.

Theorem 5.1. *Let $\lambda_t = \sqrt{1.5(b-a)^2 \log t}$ and $m \in \mathbb{N}$. Define $T_0 \triangleq \max\{K, \exp((\frac{b}{\triangle_{\min}})^{\frac{1}{m}})\}$. For $T \in \mathbb{N}$, the expected regret of Algorithm 5 is upper bounded by*

$$R(T) \leq \sum_{k \notin \operatorname{supp}(\boldsymbol{i}^\star)} \left(\sum_{j=1}^{d_k} \triangle_{\bar{j},k} T_0 + \frac{12 \triangle_{\overline{d_k},k} (b-a)^2 \log T}{\left(\frac{\mu_{\overline{d_k}}}{1+\log^{-m} T} - \mu_k\right)^2} \right)$$
$$+ \sum_{k \notin \operatorname{supp}(\boldsymbol{i}^\star)} \sum_{j=1}^{d_k} \triangle_{\bar{j},k} \left(\frac{1}{T} + \frac{\pi^2}{6}\right) + DbT_0.$$

As a consequence of Theorem 5.1, setting $T \to \infty$ yields:

$$\lim_{T \to \infty} \frac{R(T)}{\log T} \leq \sum_{k \notin \operatorname{supp}(\boldsymbol{i}^\star)} \frac{12(b-a)^2}{\triangle_{\overline{d_k},k}} \leq \mathcal{O}\left(\frac{K-D}{\triangle_{\min}}\right),$$

which matches Theorem 4 in (Kveton et al., 2014a), $\liminf_{T \to \infty} \frac{R(T)}{\log T} = \Omega(\frac{K-D}{\triangle_{\min}})$, asymptotically up to a constant factor. Note that `FasterCUCB` is faster than `CUCB` when $\triangle_{\min} = \Omega(\frac{1}{\operatorname{polylog}(K)})$ and when $T = \operatorname{poly}(K)$. Also, similar to (Cuvelier et al., 2021b), our per-round time complexity also goes to infinity as $T \to \infty$, one way to address this issue is to use `CUCB` when the per-round time complexity of ours is larger than that of `CUCB`.

Useful lemmas. Here we present two lemmas that will be used to show Theorem 5.1 in Section 5.3. First, inspired by Kveton et al. (2014a), we define a bijection g_t from $\operatorname{supp}(\boldsymbol{i}^\star)$ to $\operatorname{supp}(\boldsymbol{x}(t))$ with the following properties:

Lemma 5.2. *There exists a bijection $g_t : \operatorname{supp}(\boldsymbol{i}^\star) \to \operatorname{supp}(\boldsymbol{x}(t))$ such that (i) $g_t(j) = j$ for $j \in \operatorname{supp}(\boldsymbol{i}^\star) \cap \operatorname{supp}(\boldsymbol{x}(t))$; (ii) for any $j \in \operatorname{supp}(\boldsymbol{i}^\star) \setminus \operatorname{supp}(\boldsymbol{x}(t))$,*

$$x_{g_t(j)}(t) = 1 \implies \left\langle \operatorname{dom}(\boldsymbol{f}_{g_t(j)}), \boldsymbol{h} \right\rangle \geq \frac{\langle \operatorname{dom}(\boldsymbol{f}_j), \boldsymbol{h} \rangle}{1 + \frac{1}{3 \log^m T}}.$$

The proof of Lemma 5.2 is in Appendix E.1, where an explicit construction of g_t is provided. Property (i) allows us to decompose the instantaneous regret $\langle \boldsymbol{i}^\star - \boldsymbol{x}(t), \boldsymbol{\mu} \rangle = \sum_{k \notin \operatorname{supp}(\boldsymbol{i}^\star)} \sum_{j \in \operatorname{supp}(\boldsymbol{i}^\star)} \triangle_{j,k} \mathbb{1}\{g_t(j) = k\}$, and Property (ii), Lemma 5.2, allows us to derive a bound of $\sum_{t=1}^T \mathbb{1}\{g_t(j) = k\}$ that relates with UCB values.

Second, for technical reasons, we need the precision parameter $\epsilon = \log^{-m} T$ to be small enough so that $\triangle_{i,j}$ and $\mu_i - (1+\epsilon)\mu_j$ have the same sign. The below lemma (proved in Appendix E.2) gives the threshold to make it happen:

Lemma 5.3. *Let $\epsilon < \frac{\triangle_{\min}}{b}$. Then, for any $i \in \operatorname{supp}(\boldsymbol{i}^\star)$ and any $j \notin \operatorname{supp}(\boldsymbol{i}^\star)$, $\mu_i - \mu_j > 0 \implies \frac{\mu_i}{1+\epsilon} - \mu_j > 0$.*

5.3. Proof of Theorem 5.1

For $T \leq T_0$, $R(T)$ is trivially bounded by $T \langle \boldsymbol{\mu}, \boldsymbol{i}^\star \rangle \leq DbT_0$. In the following, we assume $T > T_0$.

As g_t is a bijection from $\text{supp}(\boldsymbol{i}^\star)$ to $\text{supp}(\boldsymbol{x}(t))$ and $g_t(j) = j$ for $j \in \text{supp}(\boldsymbol{i}^\star) \cap \text{supp}(\boldsymbol{x}(t))$, we can rewrite

$$\mathbb{E}[\langle \boldsymbol{i}^\star - \boldsymbol{x}(t), \boldsymbol{\mu}\rangle] = \sum_{k \notin \text{supp}(\boldsymbol{i}^\star)} \sum_{j \in \text{supp}(\boldsymbol{i}^\star)} \triangle_{j,k} \mathbb{E}[\mathbb{1}\{g_t(j) = k\}]$$
$$\leq \sum_{k \notin \text{supp}(\boldsymbol{i}^\star)} \sum_{j=1}^{d_k} \triangle_{\bar{j},k} \mathbb{E}\left[\mathbb{1}\{g_t(\bar{j}) = k\}\right]$$

so that the expected regret is bounded from the above by:

$$R(T) \leq \sum_{k \notin \text{supp}(\boldsymbol{i}^\star)} \sum_{j=1}^{d_k} \triangle_{\bar{j},k} \mathbb{E}\left[\sum_{t=1}^{T} \mathbb{1}\{g_t(\bar{j}) = k\}\right]$$
$$= \sum_{k \notin \text{supp}(\boldsymbol{i}^\star)} \sum_{j=1}^{d_k} \triangle_{\bar{j},k} \left((I)_{\bar{j},k} + (II)_{\bar{j},k}\right),$$

where $\begin{cases} (I)_{\bar{j},k} = \sum_{t=1}^{T} \mathbb{E}\left[\mathbb{1}\left\{g_t(\bar{j}) = k, N_k(t) \leq n_{\bar{j},k}\right\}\right] \\ (II)_{\bar{j},k} = \sum_{t=1}^{T} \mathbb{E}\left[\mathbb{1}\left\{g_t(\bar{j}) = k, N_k(t) > n_{\bar{j},k}\right\}\right] \end{cases}$

and $n_{j,k} = \max\left\{\frac{6(b-a)^2 \log T}{(\frac{\mu_j}{1+\log^{-m} T} - \mu_k)^2}, T_0\right\}$. The proof is completed by bounding related terms of $(I)_{\bar{j},k}$ and $(II)_{\bar{j},k}$ by Lemma 5.4 (proved in Appendix E.2) and Lemma 5.5.

Lemma 5.4. *Let $k \notin \text{supp}(\boldsymbol{i}^\star)$ and $j \in [d_k]$. For $T > T_0$,*

$$\sum_{j=1}^{d_k} \triangle_{\bar{j},k} (I)_{\bar{j},k} \leq \sum_{j=1}^{d_k} \triangle_{\bar{j},k} T_0 + \frac{12(b-a)^2 \triangle_{\overline{d_k},k} \log T}{(\frac{\mu_{\overline{d_k}}}{1+\log^{-m} T} - \mu_k)^2}.$$

Lemma 5.5. *Let $k \notin \text{supp}(\boldsymbol{i}^\star)$ and $j \in [d_k]$. For $T > T_0$,*

$$(II)_{\bar{j},k} \leq \frac{1}{T} + \frac{\pi^2}{6}.$$

Proof sketch: See Appendix E.2 for the entire proof. Let $\epsilon \triangleq \frac{1}{\log^m T}$. First, we claim:

$$g_t(\bar{j}) = k \implies u_k(N_k(t-1), T) \geq \frac{\min_{s<t} u_{\bar{j}}(s,t)}{1+\epsilon}, \tag{14}$$

where $u_k(s,t) = \tilde{\mu}_k(s) + \frac{\lambda_t}{\sqrt{s}}$ and $\tilde{\mu}_k(t) = \frac{1}{t}\sum_{s=1}^{t} y_k(s)$.

Show Eq. (14): Observe that $g_t(\bar{j}) = k$ implies:

$$\left(1 + \frac{\epsilon}{3}\right) \langle \boldsymbol{f}_k, \boldsymbol{q}\rangle \geq \langle \text{dom}(\boldsymbol{f}_k), \boldsymbol{q}\rangle$$
$$\geq \frac{\left\langle \text{dom}(\boldsymbol{f}_{\bar{j}}), \boldsymbol{q}\right\rangle}{1+\frac{\epsilon}{3}} \geq \frac{\left\langle \boldsymbol{f}_{\bar{j}}, \boldsymbol{q}\right\rangle}{1+\frac{\epsilon}{3}}, \tag{15}$$

where Eq. (8) is used in the first and the last inequality, and the second uses Lemma 5.2 and Corollary 4.10. By $(1+\frac{\epsilon}{3})^2 \leq 1+\epsilon$ and expanding $\boldsymbol{f}_k = (\hat{\mu}_k(t-1), \frac{1}{\sqrt{N_k(t-1)}})$ and $\boldsymbol{q} = (1, \lambda_t)$, we derive from (15) that:

$$u_k(N_k(t-1), t) \geq \frac{u_{\bar{j}}(N_{\bar{j}}(t-1), t)}{1+\epsilon},$$

and further by $\log T > \log t$ and $N_{\bar{j}}(t-1) \in [t-1]$,

$$u_k(N_k(t-1), T) \geq \frac{u_{\bar{j}}(N_{\bar{j}}(t-1), t)}{1+\epsilon} \geq \frac{\min_{s<t} u_{\bar{j}}(s,t)}{1+\epsilon},$$

which shows Eq. (14). Second, define

$$\mathcal{T}_{\bar{j},k} = \{t \in \{n_{\bar{j},k}+1, \cdots, T\} : g_t(\bar{j}) = k, N_k(t-1) > n_{\bar{j},k}\}.$$

From Eq. (14), $(II)_{\bar{j},k}$ is upper bounded by

$$\mathbb{E}\left[\sum_{t \in \in \mathcal{T}_{\bar{j},k}} \mathbb{1}\left\{u_k(N_k(t-1), T) \geq \frac{\min_{s<t}\{u_{\bar{j}}(s,t)\}}{1+\epsilon}\right\}\right]$$
$$\leq \mathbb{E}\left[\sum_{t \in \mathcal{T}_{\bar{j},k}} \sum_{s<t} (\mathbb{1}\{\mathcal{A}_{1,t,s}\} + \mathbb{1}\{\mathcal{A}_{2,t,s}\} + \mathbb{1}\{\mathcal{A}_{3,t,s}\})\right], \tag{16}$$

where $\begin{cases} \mathcal{A}_{1,t,s} = \left\{\tilde{\mu}_k(N_k(t-1)) \geq \mu_k + \frac{\lambda_T}{\sqrt{N_k(t-1)}}\right\} \\ \mathcal{A}_{2,t,s} = \left\{\mu_{\bar{j}} \geq \tilde{\mu}_{\bar{j}}(s) + \frac{\lambda_t}{\sqrt{s}}\right\} \\ \mathcal{A}_{3,t,s} = \left\{\mu_k + \frac{2\lambda_T}{\sqrt{N_k(t-1)}} > \frac{\mu_{\bar{j}}}{1+\epsilon}\right\} \end{cases}.$

See Appendix E.2 for the derivation of Eq. (16). Observe when $t \in \mathcal{T}_{\bar{j},k}$,

$$\mathbb{1}\{\mathcal{A}_{3,t,s}\} \leq \mathbb{1}\left\{\mu_k + \frac{2\lambda_T}{\sqrt{n_{\bar{j},k}+1}} > \frac{\mu_{\bar{j}}}{1+\epsilon}\right\} = 0,$$

where the inequality is because $N_k(t-1) > n_{\bar{j},k}$, and the equality is because

$$n_{\bar{j},k} \geq \frac{4\lambda_T^2}{(\frac{\mu_{\bar{j}}}{1+\epsilon} - \mu_k)^2} \implies \frac{4\lambda_T^2}{n_{\bar{j},k}+1} < \left(\frac{\mu_{\bar{j}}}{1+\epsilon} - \mu_k\right)^2,$$

and also $\frac{\mu_{\bar{j}}}{1+\epsilon} - \mu_k > 0$ which is ensured by Lemma 5.3 as $T > T_0$. Finally, from Eq. (16) and using Hoeffding's inequality, we get

$$(II)_{\bar{j},k} \leq \mathbb{E}\left[\sum_{t \in \mathcal{T}_{\bar{j},k}} \sum_{s<t} (\mathbb{1}\{\mathcal{A}_{1,t,s}\} + \mathbb{1}\{\mathcal{A}_{2,t,s}\})\right]$$
$$\leq \sum_{t=n_{\bar{j},k}+1}^{T} \sum_{s<t} \left(e^{-3\log T} + e^{-3\log t}\right).$$

See Appendix E.2 for the derivation of the second inequality. The proof is completed by evaluating

$$\begin{cases} \sum_{t=1}^{T} \sum_{s<t} e^{-3\log T} \leq \sum_{t=1}^{T} \frac{t}{T^3} \leq \frac{T(T+1)}{2T^3} \leq \frac{1}{T} \\ \sum_{t=1}^{T} \sum_{s<t} e^{-3\log t} \leq \sum_{t=1}^{\infty} \frac{t}{t^3} \leq \sum_{t=1}^{\infty} \frac{1}{t^2} \leq \frac{\pi^2}{6} \end{cases}.$$

□

6. Conclusion

In this paper, we have presented `FasterCUCB`, the first sublinear-time algorithm for matroid semi-bandits. Several possible future directions. First, one might extend our approach to speed up UCB-style algorithms for different problems such as combinatorial best-arm identification (Chen et al., 2014; Du et al., 2021) and nonstationary semi-bandits (Zhou et al., 2020; Chen et al., 2021). Second, another direction is to study the possibility of speeding up other forms of weights, such as those derived from gradients (Tzeng et al., 2023) and those in the follow-the-perturbed-leader algorithm (Neu and Bartók, 2016).

Acknowledgements

Ruo-Chun Tzeng's research is supported by the ERC Advanced Grant REBOUND (834862). Kaito Ariu's research is supported by JSPS KAKENHI Grant No. 23K19986.

Impact Statements

This paper develops an algorithm for matroid semi-bandits. The societal consequences of this work indirectly come from the applications of matroid semi-bandits, and none of which we think should be discussed here.

References

Zeinab Abbassi, Vahab S Mirrokni, and Mayur Thakur. Diversity maximization under matroid constraints. In *Proc. of KDD*, 2013.

Alper Atamtürk and Andrés Gómez. Maximizing a class of utility functions over the vertices of a polytope. *Operations Research*, 2017.

Sébastian Bubeck, Tengyao Wang, and Nitin Viswanathan. Multiple identifications in multi-armed bandits. In *Proc. of ICML*, 2013.

Lijie Chen, Anupam Gupta, and Jian Li. Pure exploration of multi-armed bandit under matroid constraints. In *Proc. of COLT*, 2016.

Shouyuan Chen, Tian Lin, Irwin King, Michael R Lyu, and Wei Chen. Combinatorial pure exploration of multi-armed bandits. In *Proc. of NeurIPS*, 2014.

Wei Chen, Yajun Wang, and Yang Yuan. Combinatorial multi-armed bandit: General framework and applications. In *Proc. of ICML*, 2013.

Wei Chen, Liwei Wang, Haoyu Zhao, and Kai Zheng. Combinatorial semi-bandit in the non-stationary environment. In *Proc. of UAI*, 2021.

Sayak Ray Chowdhury, Gaurav Sinha, Nagarajan Natarajan, and Amit Sharma. Combinatorial categorized bandits with expert rankings. In *Proc. of UAI*, 2023.

Richard Combes, Mohammad Sadegh Talebi Mazraeh Shahi, Alexandre Proutiere, and Marc Lelarge. Combinatorial bandits revisited. In *Proc. of NeurIPS*, 2015.

Richard Combes, Stefan Magureanu, and Alexandre Proutiere. Minimal exploration in structured stochastic bandits. In *Proc. of NeurIPS*, 2017.

Thibaut Cuvelier, Richard Combes, and Eric Gourdin. Asymptotically optimal strategies for combinatorial semi-bandits in polynomial time. In *Proc. of ALT*, 2021a.

Thibaut Cuvelier, Richard Combes, and Eric Gourdin. Statistically efficient, polynomial-time algorithms for combinatorial semi-bandits. In *Proc. of SIGMETRICS*, 2021b.

Rémy Degenne and Vianney Perchet. Combinatorial semi-bandit with known covariance. In *Proc. of NeurIPS*, 2016.

Yihan Du, Yuko Kuroki, and Wei Chen. Combinatorial pure exploration with full-bandit or partial linear feedback. In *Proc. of AAAI*, 2021.

Jack Edmonds. Matroids and the greedy algorithm. *Mathematical Programming*, 1971.

David Eppstein, Zvi Galil, Giuseppe F. Italiano, and Amnon Nissenzweig. Sparsification—a technique for speeding up dynamic graph algorithms. *Journal of the ACM*, 1997.

Greg N. Frederickson. Data structures for on-line updating of minimum spanning trees, with applications. *SIAM Journal on Computing*, 1985.

Yi Gai, Bhaskar Krishnamachari, and Rahul Jain. Combinatorial network optimization with unknown variables: Multi-armed bandits with linear rewards and individual observations. *IEEE/ACM Transactions on Networking*, 2012.

Todd L Graves and Tze Leung Lai. Asymptotically efficient adaptive choice of control laws incontrolled markov chains. *SIAM Journal on Control and Optimization*, 1997.

Manoj Gupta and Richard Peng. Fully dynamic $(1+\epsilon)$-approximate matchings. In *Proc. of FOCS*, 2013.

Monika Henzinger, Paul Liu, Jan Vondrák, and Da Wei Zheng. Faster submodular maximization for several classes of matroids. In *Proc. of ICALP*, 2023.

Jacob Holm, Kristian De Lichtenberg, and Mikkel Thorup. Poly-logarithmic deterministic fully-dynamic algorithms for connectivity, minimum spanning tree, 2-edge, and biconnectivity. *Journal of the ACM*, 2001.

Shinji Ito. Hybrid regret bounds for combinatorial semi-bandits and adversarial linear bandits. In *Proc. of NeurIPS*, 2021.

Kwang-Sung Jun, Aniruddha Bhargava, Robert Nowak, and Rebecca Willett. Scalable generalized linear bandits: Online computation and hashing. In *Proc. of NeurIPS*, 2017.

Satyen Kale, Lev Reyzin, and Robert E Schapire. Non-stochastic bandit slate problems. In *Proc. of NeurIPS*, 2010.

Jon Kleinberg and Éva Tardos. *Algorithm design*. Pearson Education India, 2006.

Fang Kong, Yueran Yang, Wei Chen, and Shuai Li. The hardness analysis of Thompson sampling for combinatorial semi-bandits with greedy oracle. In *Proc. of NeurIPS*, 2021.

Branislav Kveton, Zheng Wen, Azin Ashkan, Hoda Eydgahi, and Brian Eriksson. Matroid bandits: Fast combinatorial optimization with learning. In *Proc. of UAI*, 2014a.

Branislav Kveton, Zheng Wen, Azin Ashkan, and Michal Valko. Learning to act greedily: Polymatroid semi-bandits. *arXiv preprint arXiv:1405.7752*, 2014b.

Branislav Kveton, Zheng Wen, Azin Ashkan, and Csaba Szepesvari. Tight regret bounds for stochastic combinatorial semi-bandits. In *Proc. of AISTATS*, 2015.

Gergely Neu and Gábor Bartók. Importance weighting without importance weights: An efficient algorithm for combinatorial semi-bandits. *JMLR*, 2016.

James G Oxley. Matroid theory, 2011.

Orestis Papadigenopoulos and Constantine Caramanis. Recurrent submodular welfare and matroid blocking semi-bandits. In *Proc. of NeurIPS*, 2021.

Pierre Perrault. When combinatorial Thompson sampling meets approximation regret. In *Proc. of NeurIPS*, 2022.

Pierre Perrault, Vianney Perchet, and Michal Valko. Exploiting structure of uncertainty for efficient matroid semi-bandits. In *Proc. of ICML*, 2019.

Pierre Perrault, Etienne Boursier, Michal Valko, and Vianney Perchet. Statistical efficiency of thompson sampling for combinatorial semi-bandits. In *Proc. of NeurIPS*, 2020a.

Pierre Perrault, Michal Valko, and Vianney Perchet. Covariance-adapting algorithm for semi-bandits with application to sparse outcomes. In *Proc. of COLT*, 2020b.

Alexander Schrijver. *Combinatorial Optimization: Polyhedra and Efficiency*. Springer, 2003.

Matthew Streeter, Daniel Golovin, and Andreas Krause. Online learning of assignments. In *Proc. of NeurIPS*, 2009.

Mohammad Sadegh Talebi and Alexandre Proutiere. An optimal algorithm for stochastic matroid bandit optimization. In *Proc. of AAMAS*, 2016.

Taira Tsuchiya, Shinji Ito, and Junya Honda. Further adaptive best-of-both-worlds algorithm for combinatorial semi-bandits. In *Proc. of AISTATS*, 2023.

Ruo-Chun Tzeng, Po-An Wang, Alexandre Proutiere, and Chi-Jen Lu. Closing the computational-statistical gap in best arm identification for combinatorial semi-bandits. In *Proc. of NeurIPS*, 2023.

Jan van den Brand, Danupon Nanongkai, and Thatchaphol Saranurak. Dynamic matrix inverse: Improved algorithms and matching conditional lower bounds. In *Proc. of FOCS*, 2019.

Siwei Wang and Wei Chen. Thompson sampling for combinatorial semi-bandits. In *Proc. of ICML*, 2018.

Zheng Wen, Branislav Kveton, and Azin Ashkan. Efficient learning in large-scale combinatorial semi-bandits. In *Proc. of ICML*, 2015.

Shuo Yang, Tongzheng Ren, Sanjay Shakkottai, Eric Price, Inderjit S Dhillon, and Sujay Sanghavi. Linear bandit algorithms with sublinear time complexity. In *Proc. of ICML*, 2022.

Huozhi Zhou, Lingda Wang, Lav Varshney, and Ee-Peng Lim. A near-optimal change-detection based algorithm for piecewise-stationary combinatorial semi-bandits. In *Proc. of AAAI*, 2020.

A. Notation

Problem setting	
K	the number of arms
$\mathcal{X}$	the bases of the given matroid $([K], \mathcal{I})$
D	$\max_{\boldsymbol{x}\in\mathcal{X}} \\|\boldsymbol{x}\\|_0$
$\boldsymbol{\mu}$	the mean vector of the K arms $\nu_1, \cdots, \nu_K$
$\boldsymbol{i}^\star(\boldsymbol{\mu})$	an action attaining $\max_{\boldsymbol{x}\in\mathcal{X}} \langle \boldsymbol{\mu}, \boldsymbol{x} \rangle$
Notation related to `FasterCUCB`	
$N_k(t)$	the number of arm pulls of arm k
$\boldsymbol{x}(t)$	the action selected by the algorithm at round t
$\boldsymbol{y}(t)$	the reward vector at round t
$\hat{\mu}_k(t)$	the empirical reward $\frac{1}{N_k(t)} \sum_{s=1}^{t} y_k(t)\mathbb{1}\{x_k(t) = 1\}$ of arm k at round t
λ_t	the parameter that controls the confidence interval
Notation related to dynamic algorithm	
$\boldsymbol{f}_k = (\alpha_k, \beta_k)$	a nonnegative two-dimensional feature of arm k
$\boldsymbol{q}$	a nonnegative two-dimensional query
$(\alpha_{\text{lb}}, \alpha_{\text{ub}})$	lower and upper bounds of α_k's
$(\beta_{\text{lb}}, \beta_{\text{ub}})$	lower and upper bounds of β_k's
W	(the square root of) the number of bins
$\text{BIN}_{q,r}$	bins that partition the possible region of the features
$\text{dom}_{q,r}$	dominating point of $\text{BIN}_{q,r}$
$\text{dom}(\boldsymbol{f}_k)$	dominating point of $\text{BIN}_{q,r}$ to which $\boldsymbol{f}_k$ belongs
$\mathcal{L} = \{\boldsymbol{l}_1, \ldots, \boldsymbol{l}_{\binom{K}{2}}\}$	the set of $\binom{K}{2}$ lines, each of which is orthogonal to line $\overleftrightarrow{\boldsymbol{f}_k \boldsymbol{f}_{k'}}$ for some $k \neq k'$ and intersects $\mathbf{0}$
$\mathcal{H}$	a minimum hitting set of the cells in arrangement of $\mathcal{L}$
Notation related to regret analysis	
$\{\bar{j}\}_{j=1}^{D}$	the permutation of $\text{supp}(\boldsymbol{i}^\star)$ such that $\mu_{\bar{1}} \geq \cdots \geq \mu_{\bar{D}}$
ϵ	the precision parameter which is set to $\frac{1}{\log^m T}$ in `FasterCUCB` (Algorithm 5)
$g_t(j)$	the mapping from $\text{supp}(\boldsymbol{i}^\star)$ to $\text{supp}(\boldsymbol{x}(t))$ such that (i) $g_t(j) = j$ if $j \in \text{supp}(\boldsymbol{i}^\star) \cap \text{supp}(\boldsymbol{x}(t))$ (ii) $x_{g_t(j)}(t) = 1$ implies $\left\langle \text{dom}(\boldsymbol{f}_{g_t(j)}), \boldsymbol{h} \right\rangle \geq \frac{1}{1+\frac{\epsilon}{3}} \left\langle \text{dom}(\boldsymbol{f}_j), \boldsymbol{h} \right\rangle$
$\triangle_{j,k}$	the difference $\mu_j - \mu_k$ between arm j's and arm k's expected reward
$\triangle_{\min}$	the smallest positive gap $\triangle_{i,j}$ between any pair of $i \in \text{supp}(\boldsymbol{i}^\star)$ and $j \notin \text{supp}(\boldsymbol{i}^\star)$
d_k	the largest $j \in [D]$ such that $\triangle_{\delta(j),k} > 0$
$\tilde{\mu}_k(t)$	the average $\frac{1}{t} \sum_{s=1}^{t} y_k(s)$ of rewards of arm k in the first t rounds
$u_k(s,t)$	the UCB value of $\tilde{\mu}_k(s) + \frac{\lambda_t}{\sqrt{s}}$ under s samples of arm k and with confidence parameter λ_t

Table 2. Table of notation.

B. Further Related Works

In this section, we review relevant literatures on combinatorial semi-bandits and sublinear-time bandits. We focus on the stochastic setting. For ease of comparison, we assume the best action $\boldsymbol{i}^\star \in \operatorname{argmax}_{\boldsymbol{x}\in\mathcal{X}} \langle \boldsymbol{x}, \boldsymbol{\mu} \rangle$ is unique, and define $\triangle_{\min} \triangleq \min_{j,k\in[K]:i^\star_j=1,i^\star_k=0,\mu_j-\mu_k>0}(\mu_j - \mu_k)$, and $\triangle \triangleq \min_{\boldsymbol{x}\neq\boldsymbol{i}^\star:\langle \boldsymbol{i}^\star-\boldsymbol{x},\boldsymbol{\mu}\rangle>0} \langle \boldsymbol{i}^\star - \boldsymbol{x}, \boldsymbol{\mu} \rangle$.

Matroid semi-bandits. Kveton et al. (2014a) showed an instance such that any uniformly good algorithm[2] suffer $R(T) = \Omega\left(\frac{(K-D)\log T}{\triangle_{\min}}\right)$. They also showed that `CUCB` (Gai et al., 2012; Chen et al., 2013) have a regret bound scaling as $\mathcal{O}\left(\frac{(K-D)\log T}{\triangle_{\min}}\right)$. Talebi and Proutiere (2016) showed an instance-specific lower bound $\liminf_{T\to\infty} \frac{R(T)}{\log T} \geq c(\boldsymbol{\mu})$ for uniformly good algorithms, where $c(\boldsymbol{\mu})$ is the optimum of a semi-infinite linear program (Graves and Lai, 1997; Combes et al., 2015), and proposed `KL-OSM` whose regret upper bound matches this lower bound. The per-round compleixty of `KL-OSM` is K line search for generating the indices plus the time for solving a linear maximization. Both `CUCB` and `KL-OSM` rely on the greedy algorithm (Algorithm 1) to solve the linear maximization for determing the action to be pulled. The time complexity of the greedy algorithm is upper bounded by $\mathcal{O}(K(\log K + \mathcal{T}_{\text{member}}))$ time and lower bounded by $\Omega(K)$. (Perrault et al., 2019) showed that the sampling rule of many combinatorial semi-bandit algorithms is a maximization problem over a summation of a linear function and a submodular function, and proposed two efficient algorithms for matroid semi-bandits: One is based on local search and the other is a greedy algorithm. Both have per-round time complexity at least $\Omega(KD)$. In contrast, our `FasterCUCB` is the first matroid semi-bandit algorithm with per-round time complexity sublinear in K for many classes of matroids.

Combinatorial semi-bandits. Here, we review works that focus on the standard setting of stochastic combinatorial semi-bandits. These consider a linear reward function and any action sets $\mathcal{X}$, where linear maximization $\max_{\boldsymbol{x}\in\mathcal{X}} \langle \boldsymbol{x}, \boldsymbol{v} \rangle$ for any $\boldsymbol{v} \in \mathbb{R}^K$ can be solved in time polynomial in K. We omit the discussion on works that focus on a specific action set (Chowdhury et al., 2023), with additional structural assumptions on the rewards (Wen et al., 2015; Perrault et al., 2020b), or with a different reward function (Papadigenopoulos and Caramanis, 2021). Perrault et al. (2020a) showed that `CTS` has a regret bound of $\mathcal{O}(\frac{K\log^2 D\log T}{\triangle})$ for mutually independent gaussian rewards and a regret bound of $\mathcal{O}(\frac{KD\log^2 D\log T}{\triangle})$ for correlated gaussian rewards. Perrault (2022) sharpen the regret bound of `CTS` for the case of mutually independent gaussian rewards to be $\mathcal{O}(\frac{K\log D\log T}{\triangle})$. The per-round time complexity of `CTS` is at least $\Omega(K)$ due to sampling from the posterior distributions. Degenne and Perchet (2016) showed that `ESCB2` has regret bound of $\mathcal{O}(\frac{K\log^2 D\log T}{\triangle})$ for independent subgaussian rewards, but its sampling rule is NP-hard (Atamtürk and Gómez, 2017) to optimize. Cuvelier et al. (2021b) proposed `AESCB` that approximates `ESCB2` with per-round time complexity of $\mathcal{O}(KD\log^3 K \text{ poly}(\log T))$ while maintaining the same regret bound. Their technique is based on rounding and budgeted-linear maximization. `OSSB` (Combes et al., 2017) is an asymptotically instance-specifically optimal algorithm for general structured bandits, including combinatorial semi-bandits, but at each round, it requires to solve a semi-infinite linear program (Graves and Lai, 1997). Cuvelier et al. (2021a) developed a method that runs in time polynomial in K to solve the semi-infinite linear program for Gaussian rewards. They managed to maintain `OSSB`'s asymptotic optimality for m-sets, but not for spanning trees and bipartite matchings. Ito (2021) and Tsuchiya et al. (2023) proposed algorithms based on the optimistic FTRL framework that achieve $\mathcal{O}(\frac{KD\log T}{\triangle})$ regret in the stochastic setting and $\mathcal{O}(\sqrt{KDT\log T})$ in the adversarial setting. At each round, the proposed algorithms first use FTRL rule to obtain a vector $\boldsymbol{a}(t)$ in the convex hull of $\mathcal{X}$ and then sample an action $\boldsymbol{x}(t)$ based on $\boldsymbol{a}(t)$. Tsuchiya et al. (2023) mentioned that the computational efficiency of the sampling step has long been a problem in semi-bandits using the optimistic FTRL framework.

Sublinear-time linear bandits. Several works (Jun et al., 2017; Yang et al., 2022) focusing on making per-round complexity of linear bandits sublinear in the number of arms. Maximum Inner Product Search (MIPS) is the primary tool used to design such algorithms. For N arms in $\mathbb{R}^d$, `Q-GLOC` (Jun et al., 2017) achieves a high-probability regret bound of $\tilde{\mathcal{O}}(d^{\frac{5}{4}}\sqrt{T})$ and per-round time complexity of $\tilde{\mathcal{O}}(d^2 N^\rho \log N)$ for some $\rho \in (0, 1)$, where $\tilde{\mathcal{O}}$ hides polylogarithmic factors in T and d. Yang et al. (2022) considered the setting with arms addition (resp. addition and deletion), and proposed `Sub-Elim` (resp. `Sub-TS`), which has a high-probability regret bound of $\tilde{\mathcal{O}}(d\sqrt{T})$ (resp. $\tilde{\mathcal{O}}(d^{\frac{3}{2}}\sqrt{T})$) and per-round time complexity of $N^{1-\Theta(\frac{1}{T^2\log^2 T})}$ (resp. $N^{1-\Theta(\frac{1}{T})}$). These results are applicable to our setting with $d = K$ and $N = |\mathcal{X}|$. `Q-GLOC` (Jun et al., 2017) applied to our setting has regret bound of $\tilde{O}(K^{\frac{5}{4}}\sqrt{T})$ and per-round time compleixty $\tilde{\mathcal{O}}(K^2|\mathcal{X}|^\rho)$.

[2]A uniformly good algorithm has the expected regret $R(T) = o(T^\alpha)$ hold for any $\alpha > 0$.

`Sub-Elim` (resp. `Sub-TS`) Yang et al. (2022) applied to our setting has regret bound of $\tilde{\mathcal{O}}(K\sqrt{T})$ (resp. $\tilde{\mathcal{O}}(K^{\frac{3}{2}}\sqrt{T})$) and has per-round time complexity of $|\mathcal{X}|^{1-\Theta(\frac{1}{T^2 \log T})}$ (resp. $|\mathcal{X}|^{1-\Theta(\frac{1}{T})}$). These results have worse regret bounds than what we have obtained, and their per-round time complexity can be exponential in K.

C. Membership Oracles for Different Matroids

In this section, we discuss $\mathcal{T}_{\mathrm{member}}$ for the matroids shown in Section 2.

- For uniform matroid, the membership oracle is given $\boldsymbol{x} \in \mathcal{I}$ and $k \in [K] \setminus \operatorname{supp}(\boldsymbol{x})$, and has to check whether $|\operatorname{supp}(\boldsymbol{x} + \boldsymbol{e}_k)| \leq D$. Suppose the number $n = |\operatorname{supp}(\boldsymbol{x})|$ is maintained. Then, it takes $\mathcal{O}(1)$ time to check whether $n + 1 \leq D$, and hence $\mathcal{T}_{\mathrm{member}} = \mathcal{O}(1)$.
- For partition matroids, the membership oracle is given $\boldsymbol{x} \in \mathcal{I}$ and $k \in [K] \setminus \operatorname{supp}(\boldsymbol{x})$, and has to check whether $|\operatorname{supp}(\boldsymbol{x} + \boldsymbol{e}_k) \cap S_i| \leq 1$ for all $i \in [D]$. Suppose there is an integer array A of size K such that $j \in S_{A[j]}$ for each $j \in [K]$, and suppose there is an integer array B of size D such that $B[i] = \sum_{j \in \operatorname{supp}(\boldsymbol{x})} \mathbb{1}\{j \in S_i\}$ for each $i \in [D]$. Then, to decide whether whether $\boldsymbol{x} + \boldsymbol{e}_k \in \mathcal{I}$, it only requires to check whether $B[A[k]] + 1 \leq 1$. This can be implemented in $\mathcal{O}(1)$ time, and thus $\mathcal{T}_{\mathrm{member}} = \mathcal{O}(1)$.
- For graphical matroids, the membership oracle has to detect if there is a cycle. Using the union-find data structure, whether $\operatorname{supp}(\boldsymbol{x}) \cup \{k\}$ has a cycle can be detected in $\mathcal{O}(\log K)$ time, so we have $\mathcal{T}_{\mathrm{member}} = \mathcal{O}(\log K)$. Refer to Section 4.6 in (Kleinberg and Tardos, 2006) for more detailed explanation.
- For transversal matroids, there is little discussion about its membership oracle. Here we present an implementation to answer a query $(\boldsymbol{x}, k)$ about whether $\boldsymbol{x} + \boldsymbol{e}_k \in \mathcal{I}$, where $\boldsymbol{x} \in \mathcal{I}$ and $k \in [K] \setminus \operatorname{supp}(\boldsymbol{x})$. Suppose a maximum matching M on $\operatorname{supp}(\boldsymbol{x}) \cup V$ is maintained. Then, answering whether $\boldsymbol{x} + \boldsymbol{e}_k \in \mathcal{I}$ is equivalent to checking whether an augmenting path on $\operatorname{supp}(\boldsymbol{x} + \boldsymbol{e}_k) \cup V$ from M can be found. Finding a augmentation path can be done by a breadth-first search (BFS) starting from k (see Section 17.2 in (Schrijver, 2003)), and it takes $\mathcal{O}(DK)$ time because there are at most K leaves in the BFS tree and the length of the path from k to each leaf is at most $2D$. Thus, we have $\mathcal{T}_{\mathrm{member}} = \mathcal{O}(DK)$.

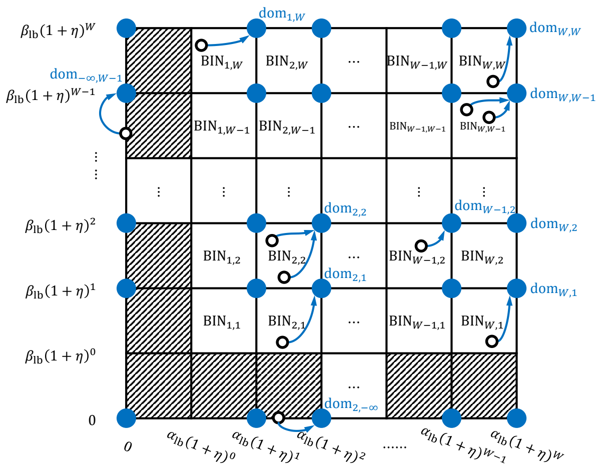


Figure 1. Illustration of feature rounding. There are $|\mathbb{W}|^2$ bins, and features are assumed not to be in (the interior of) the shaded area. Each feature $\boldsymbol{f}_k$ is rounded to its dominating point $\mathrm{dom}(\boldsymbol{f}_k)$, which is specified by a curved arrow.

D. Omitted Proofs in Section 4

Proof of Lemma 4.6. By Eq. (8) and the optimality of $\boldsymbol{x}^*_{\mathrm{dom}}$, for any base $\boldsymbol{x}$, we have

$$\begin{aligned}
&\sum_{k \in \mathrm{supp}(\boldsymbol{x}^*_{\mathrm{dom}})} \langle \boldsymbol{f}_k, \boldsymbol{q} \rangle \\
&> \frac{1}{1+\eta} \cdot \sum_{k \in \mathrm{supp}(\boldsymbol{x}^*_{\mathrm{dom}})} \langle \mathrm{dom}(\boldsymbol{f}_k), \boldsymbol{q} \rangle && \text{(by Eq. (8))} \\
&\geq \frac{1}{(1+\eta)^2} \cdot \sum_{k \in \mathrm{supp}(\boldsymbol{x})} \langle \mathrm{dom}(\boldsymbol{f}_k), \boldsymbol{q} \rangle && \text{(by optimality of } \boldsymbol{x}^*_{\mathrm{dom}}) \\
&\geq \frac{1}{(1+\eta)^2} \cdot \sum_{k \in \mathrm{supp}(\boldsymbol{x})} \langle \boldsymbol{f}_k, \boldsymbol{q} \rangle && \text{(by Eq. (8))} \\
&\geq \frac{1}{1+\epsilon} \cdot \sum_{k \in \mathrm{supp}(\boldsymbol{x})} \langle \boldsymbol{f}_k, \boldsymbol{q} \rangle. && \text{(as } (1+\eta)^2 \leq 1+\epsilon) \quad \square
\end{aligned}$$

Proof of Lemma 4.7. The proof follows from the uniqueness of the maximum-weight base in the case of distinct weights; see, e.g., (Edmonds, 1971). □

Proof of Lemma 4.8. For a query $\boldsymbol{q} \in \mathbb{R}^2$, let $\mathcal{C} \subset \mathbb{R}^2$ be a cell in arrangement of $\mathcal{L}$ whose closure contains $\boldsymbol{q}$ (which may not be uniquely determined). Then, there is a permutation $\pi \in \mathfrak{S}_K$ such that for any vector $\boldsymbol{h} \in \mathcal{H} \cap \mathcal{C}$, we have $\langle \boldsymbol{f}_{\pi(i)}, \boldsymbol{h} \rangle > \langle \boldsymbol{f}_{\pi(j)}, \boldsymbol{h} \rangle$ whenever $i < j$. Since $\boldsymbol{q}$ is in the closure of $\mathcal{C}$, it holds that $\langle \boldsymbol{f}_{\pi(i)}, \boldsymbol{q} \rangle \geq \langle \boldsymbol{f}_{\pi(j)}, \boldsymbol{q} \rangle$ for any $i < j$, implying the proof. □

Proof of Lemma 4.9. Since each cell in arrangement of $\mathcal{L}$ is a polyhedral cone generated by two lines of $\mathcal{L}$ that does not contain any other lines of $\mathcal{L}$, there are only $\mathcal{O}(K^2)$ cells, and each of their internal points can be found by using Algorithm 6, as desired.

□

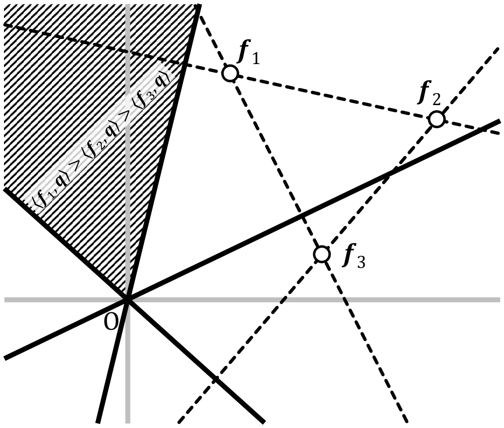


Figure 2. Illustration of characterization of representable permutations. There are three features $\boldsymbol{f}_1, \boldsymbol{f}_2, \boldsymbol{f}_3$ on $\mathbb{R}^2$. Each dashed line denotes $\overleftrightarrow{\boldsymbol{f}_i \boldsymbol{f}_j}$ for some $i \neq j$; each black bold line is orthogonal to some dashed line and intersects the origin. Such black bold lines generate six regions, each corresponding to a distinct permutation. For example, for any query $\boldsymbol{q}$ in the hatched area, it holds that $\langle \boldsymbol{f}_1, \boldsymbol{q} \rangle > \langle \boldsymbol{f}_2, \boldsymbol{q} \rangle > \langle \boldsymbol{f}_3, \boldsymbol{q} \rangle$; i.e., $\boldsymbol{q}$ represents a permutation π such that $(\pi(1), \pi(2), \pi(3)) = (1, 2, 3)$.

Algorithm 6 **GENERATE-HITTING-SET**.

Input: K distinct features $(\boldsymbol{f}_k)_{k \in [K]}$.
let $\Theta \leftarrow \emptyset$;
for *all* $k \neq k'$ **do**
    let L be a unique line that is orthogonal to line $\overleftrightarrow{\boldsymbol{f}_k \boldsymbol{f}_{k'}}$ and intersects $\mathbf{0}$;
    add the angle θ of L and $-\theta$ to Θ;
end
let $\mathcal{H} \leftarrow \emptyset$;
for *all neighboring (but distinct)* θ_1 *and* θ_2 *in* Θ **do**
    let $\boldsymbol{h} \triangleq \left(\cos(\frac{\theta_1+\theta_2}{2}), \sin(\frac{\theta_1+\theta_2}{2})\right)$ be an internal point of a polyhedral cone generated by two half-lines whose angles are θ_1 and θ_2;
    add $\boldsymbol{h}$ to $\mathcal{H}$;
end
return $\mathcal{H}$;

Proof of Theorem 4.4. The correctness of **FIND-BASE** is shown first. Given a query $\boldsymbol{q} \in \mathbb{R}_+^2$, Algorithm 3 finds $\boldsymbol{h} \in \mathcal{H}$ such that $\langle \boldsymbol{f}_k, \boldsymbol{h} \rangle > \langle \boldsymbol{f}_{k'}, \boldsymbol{h} \rangle$ implies $\langle \boldsymbol{f}_k, \boldsymbol{q} \rangle \geq \langle \boldsymbol{f}_{k'}, \boldsymbol{q} \rangle$ due to Lemma 4.8. Calling $\mathcal{A}_{\boldsymbol{h}}$ finds a $(1+\eta)$-approximate maximum-weight base $\boldsymbol{x}^\circ$ of $\mathcal{M}$ with arm k's weight defined as $\langle \mathrm{dom}(\boldsymbol{f}_k), \boldsymbol{h} \rangle$. Since a total order over $[K]$ induced by arm weights $\langle \mathrm{dom}(\boldsymbol{f}_k), \boldsymbol{h} \rangle$ is consistent with that induced by arm weights $\langle \mathrm{dom}(\boldsymbol{f}_k), \boldsymbol{q} \rangle$, by Lemma 4.7, $\boldsymbol{x}^\circ$ is also a $(1+\eta)$-approximate maximum-weight base of $\mathcal{M}$ with arm k's weight defined as $\langle \mathrm{dom}(\boldsymbol{f}_k), \boldsymbol{q} \rangle$. By Lemma 4.6,

$$\sum_{k \in \mathrm{supp}(\boldsymbol{x}^\circ)} \langle \boldsymbol{f}_k, \boldsymbol{q} \rangle \geq \frac{1}{1+\epsilon} \cdot \sum_{k \in \mathrm{supp}(\boldsymbol{x})} \langle \boldsymbol{f}_k, \boldsymbol{q} \rangle, \tag{17}$$

for any base $\boldsymbol{x}$ of $\mathcal{M}$; namely, $\boldsymbol{x}^\circ$ is a $(1+\epsilon)$-approximate maximum-weight base of $\mathcal{M}$ with arm k's weight defined as $\langle \boldsymbol{f}_k, \boldsymbol{q} \rangle$, completing the correctness of **FIND-BASE**.

Subsequently, we bound the time complexity of each subroutine as follows.

INITIALIZE: Construction of $\mathrm{BIN}_{q,r}$ and $\mathrm{dom}_{q,r}$ for all $q, r \in \mathbb{W}$ completes in $\mathcal{O}(K + W^2)$ time. Then, a hitting set $\mathcal{H}$ for $\mathrm{dom}_{q,r}$'s and $\frac{1}{1+\eta} \cdot \mathrm{dom}_{q,r}$'s of size $|\mathcal{H}| = \mathcal{O}(W^4)$ can be constructed in poly (W) time due to Lemma 4.9. There will be $|\mathcal{H}|$ instances of algorithm $\mathcal{A}$ (with different arm weights), creating which takes $\mathcal{O}(W^4 \cdot \mathcal{T}_{\mathrm{init}}(\mathcal{A}; \eta))$ time.

FIND-BASE: Checking whether each $\boldsymbol{h} \in \mathcal{H}$ and query $\boldsymbol{q} \in \mathbb{R}_+^2$ belong to (the closure of) the same cell in arrangement of

$\mathcal{V}$ can be done in $\mathcal{O}(W^2)$ time by comparing the induced total orders. By brute-force search, a desired $\boldsymbol{h}$ can be found in $\mathcal{O}(W^6)$ time. Since calling $\mathcal{A}_{\boldsymbol{h}}$ requires $\mathcal{O}(D)$ time, the entire time complexity is bounded by $\mathcal{O}(\text{ poly }(W) + D)$.

UPDATE-FEATURE: For $|\mathcal{H}|$ instances of $\mathcal{A}$, a single arm's weight would be changed, each of which runs in $\mathcal{T}_{\mathrm{update}}(\mathcal{A}; \eta)$ time.

Observe finally that

$$\begin{aligned} W &= \max\left\{\left\lceil \log_{1+\eta}\left(\frac{\alpha_{\text{ub}}}{\alpha_{\text{lb}}}\right)\right\rceil, \left\lceil \log_{1+\eta}\left(\frac{\beta_{\text{ub}}}{\beta_{\text{lb}}}\right)\right\rceil\right\} \\ &= \mathcal{O}\left(\frac{\log(\frac{\alpha_{\text{ub}}}{\alpha_{\text{lb}}}) + \log(\frac{\beta_{\text{ub}}}{\beta_{\text{lb}}})}{\log(1+\eta)}\right) \\ &= \mathcal{O}\left(\eta^{-1} \cdot \log\left(\frac{\alpha_{\text{ub}}}{\alpha_{\text{lb}}} \cdot \frac{\beta_{\text{ub}}}{\beta_{\text{lb}}}\right)\right) \\ &= \mathcal{O}\left(\epsilon^{-1} \cdot \log\left(\frac{\alpha_{\text{ub}}}{\alpha_{\text{lb}}} \cdot \frac{\beta_{\text{ub}}}{\beta_{\text{lb}}}\right)\right), \end{aligned} \tag{18}$$

where we used the fact that $\frac{1}{\log(1+\eta)} < \frac{1}{\eta}$ when $\eta \in (0, 1)$, completing the proof. $\square$

E. Proofs Related to Regret Analysis

E.1. Proofs Related to the Bijection g_t

Lemma 5.2. There exists a bijection g_t : supp $(\boldsymbol{i}^\star) \to$ supp $(\boldsymbol{x}(t))$ such that (i) $g_t(j) = j$ for $j \in$ supp $(\boldsymbol{i}^\star) \cap$ supp $(\boldsymbol{x}(t))$; (ii) for any $j \in$ supp $(\boldsymbol{i}^\star) \setminus$ supp $(\boldsymbol{x}(t))$,

$$x_{g_t(j)}(t) = 1 \implies \left\langle \mathrm{dom}(\boldsymbol{f}_{g_t(j)}), \boldsymbol{h} \right\rangle \geq \frac{\left\langle \mathrm{dom}(\boldsymbol{f}_j), \boldsymbol{h} \right\rangle}{1 + \frac{1}{3\log^m T}}.$$

Proof: Let $\eta \triangleq \frac{1}{3\log^m T}$. The proof is inspired by Section 4.2 in (Kveton et al., 2014a), and several changes are made to deal with the usage of the dynamic$(1+\eta)$-approximate maximum-weight basis algorithm in the **FIND-BASE** procedure.

Let $\xi_t : [D] \to$ supp $(\boldsymbol{x}(t))$ be the ordering such that $\xi_t(i)$'s arm weight $\left\langle \mathrm{dom}(\boldsymbol{f}_{\xi_t(i)}), \boldsymbol{h} \right\rangle$ is the i-th largest, where $\boldsymbol{h} \in \mathcal{H}$ lies in the same cell as the query $\boldsymbol{q} = (1, \lambda_t)$ when invoking **FIND-BASE** procedure.

Explicit construction of g_t: We define

$$g_t(j) = \xi_t(\pi_t^{-1}(j)), \forall j \in \text{ supp}\,(\boldsymbol{i}^\star),$$

where the function $\pi_t : [D] \to$ supp $(\boldsymbol{i}^\star)$ is a bijection such that the following hold:

(i) $\sum_{i=1}^{k-1} \boldsymbol{e}_{\xi_t(i)} + \boldsymbol{e}_{\pi_t(k)} \in \mathcal{I}$ for all $k \in [D]$

(ii) $\pi_t(k) = \xi_t(k)$ if $\xi_t(k) \in$ supp $(\boldsymbol{i}^\star) \cap$ supp $(\boldsymbol{x}(t))$

The existence of π_t is proved in Lemma E.1 and also by Lemma 1 of (Kveton et al., 2014a).

Show (i) $g_t(j) = j$ for $j \in$ supp $(\boldsymbol{i}^\star) \cap$ supp $(\boldsymbol{x}(t))$: Fix any $j \in$ supp $(\boldsymbol{i}^\star) \cap$ supp $(\boldsymbol{x}(t))$. From the definition of π_t, we have $\pi_t(j) = \xi_t(j)$ and hence $g_t(j) = \xi_t(\pi_t^{-1}(j)) = \xi_t(\xi_t^{-1}(j)) = j$.

Show (ii) $x_{g_t(j)}(t) = 1 \implies \left\langle \mathrm{dom}(\boldsymbol{f}_{g_t(j)}), \boldsymbol{h} \right\rangle \geq \frac{\left\langle \mathrm{dom}(\boldsymbol{f}_j), \boldsymbol{h} \right\rangle}{1+\eta}$: Fix any $j \in$ supp $(\boldsymbol{i}^\star) \setminus$ supp $(\boldsymbol{x})$. Let $k = \pi_t^{-1}(j)$. Observe that the bijection π_t captures the situation that: The algorithm can choose $\pi_t(k) \in$ supp $(\boldsymbol{i}^\star)$ as the k-th element but instead chooses $\xi_t(k) \in$ supp $(\boldsymbol{x}(t))$. By the procedure of Algorithm 3 and Assumption 4.2, this happens when

$$\left\langle \mathrm{dom}(\boldsymbol{f}_{\xi_t(k)}), \boldsymbol{h} \right\rangle \geq \frac{1}{1+\eta} \left\langle \mathrm{dom}(\boldsymbol{f}_{\pi_t(k)}), \boldsymbol{h} \right\rangle,$$

and replacing $k = \pi_t^{-1}(j)$ completes the proof. ☐

Lemma E.1. *Let $\boldsymbol{x}, \boldsymbol{i}^\star \in \mathcal{X}$, and $\xi : [D] \to$ supp $(\boldsymbol{x})$ be an arbitrary bijection. There exists a bijection $\pi : [D] \to$ supp $(\boldsymbol{i}^\star)$ such that $\sum_{i=1}^{k-1} \boldsymbol{e}_{\xi(i)} + \boldsymbol{e}_{\pi(k)} \in \mathcal{I}$ for all $k \in [D]$.*

Proof: This lemma is equivalent to Lemma 1 of (Kveton et al., 2014a). For reader's convenience, we provide a proof here.

For $k = D$, consider $\sum_{i=1}^{D-1} \boldsymbol{e}_{\xi(i)} \in \mathcal{I}$ (due to hereditary property), and $\boldsymbol{i}^\star \in \mathcal{I}$. As the former has $D-1$ element while the latter has D elements, by augmentation property, there exists $\pi(D) \in$ supp $(\boldsymbol{i}^\star)$ such that $\sum_{i=1}^{D-1} \boldsymbol{e}_{\xi(i)} + \boldsymbol{e}_{\pi(D)} \in \mathcal{I}$. For the case when $\xi(D) \in$ supp $(\boldsymbol{i}^\star) \cap$ supp $(\boldsymbol{x})$, we set $\pi(D) = \xi(D)$.

The proof is completed by repeating the following process for $k = D-1, \cdots, 1$. As $\sum_{i=1}^{k-1} \boldsymbol{e}_{\xi(i)} \in \mathcal{I}$ (due to hereditary property) has $k-1$ elements, and $\boldsymbol{i}^\star - \sum_{i=k+1}^{D} \boldsymbol{e}_{\pi(i)} \in \mathcal{I}$ (due to hereditary property) has k elements, by augmentation property, there exists $\pi(k)$ such that $\sum_{i=1}^{k-1} \boldsymbol{e}_{\xi(i)} + \boldsymbol{e}_{\pi(k)} \in \mathcal{I}$. If $\xi(k) \in$ supp $(\boldsymbol{i}^\star) \cap$ supp $(\boldsymbol{x})$, we set $\pi(k) = \xi(k)$. ☐

E.2. Lemmas Related to Regret Analysis of Algorithm 5

In this section, we fix a best action $\boldsymbol{i}^\star \in \mathrm{argmax}_{\boldsymbol{x} \in \mathcal{X}} \langle \boldsymbol{\mu}, \boldsymbol{x} \rangle$ and define $\triangle_{j,k} \triangleq \mu_j - \mu_k$. Let $\{\bar{j}\}_{j=1}^{D}$ be the permutation of supp $(\boldsymbol{i}^\star)$ such that $\mu_{\bar{1}} \geq \cdots \geq \mu_{\bar{D}}$. Define $d_k \triangleq \max\{j \in \text{ supp}\,(\boldsymbol{i}^\star) : \triangle_{\bar{j},k} > 0\}$ and $\triangle_{\min} \triangleq \min_{k \notin \text{ supp}(\boldsymbol{i}^\star)} \triangle_{\overline{d_k},k}$.

Lemma 5.3. Let $\epsilon < \frac{\triangle_{\min}}{b}$. Then, for any $i \in$ supp $(\boldsymbol{i}^\star)$ and any $j \notin$ supp $(\boldsymbol{i}^\star)$,

$$\mu_i - \mu_j > 0 \implies \frac{\mu_i}{1+\epsilon} - \mu_j > 0.$$

Proof: Fix $i \in$ supp $(\boldsymbol{i}^\star)$ and $j \notin$ supp $(\boldsymbol{i}^\star)$ such that $\mu_i - \mu_j > 0$. We want the following to hold:

$$\frac{\mu_i}{1+\epsilon} - \mu_j > 0 \Longleftrightarrow \mu_i - (1+\epsilon)\mu_j > 0 \Longleftrightarrow \mu_i - \mu_j > \epsilon\mu_j.$$

As $\mu_i - \mu_j > \epsilon\mu_j$ must hold for all such i and j, taking the minimum over all possible i and j on the left-hand side, and use the fact that $\mu_j \le b$ for all j on the right-hand side, we derive

$$\frac{\triangle_{\min}}{b} > \epsilon$$

is the condition on ϵ to ensure $\mu_i - \mu_j > 0 \implies \frac{\mu_i}{1+\epsilon} - \mu_j > 0$ holds for all i and j. □

Lemma 5.4. Let $k \notin$ supp $(\boldsymbol{i}^\star)$ and $j \in [d_k]$. For $T > T_0$,

$$\sum_{j=1}^{d_k} \triangle_{\overline{j},k} (I)_{\overline{j},k} \le \sum_{j=1}^{d_k} \triangle_{\overline{j},k} T_0 + \frac{12(b-a)^2 \triangle_{\overline{d_k},k} \log T}{\left(\frac{\mu_{\overline{d_k}}}{1+\log^{-m} T} - \mu_k\right)^2}.$$

Proof: Recall $(I)_{\overline{j},k} = \sum_{t=1}^T \mathbb{E}\left[\mathbb{1}\left\{g_t(\overline{j}) = k, N_k(t) \le n_{\overline{j},k}\right\}\right]$, where $n_{j,k} = \max\left\{\frac{6(b-a)^2 \log T}{\left(\frac{\mu_j}{1+\log^{-m} T} - \mu_k\right)^2}, T_0\right\}$.

First, we claim that: for any $\{a_j\}_{j=1}^{d_k}$ with $a_1 \ge \cdots \ge a_{d_k} \ge 0$,

$$\sum_{j=1}^{d_k} a_j (I)_{\overline{j},k} \le a_1 n_{\overline{1},k} + \sum_{j=2}^{d_k} a_j (n_{\overline{j},k} - n_{\overline{j-1},k}). \tag{19}$$

Show Eq. (19): We show by induction. For the base case, we have

$$a_1 (I)_{\overline{1},k} + a_2 (I)_{\overline{2},k} \le a_1 n_{\overline{1},k} + a_2 (n_{\overline{2},k} - n_{\overline{1},k}). \tag{20}$$

Eq. (20) is derived as follows. Since $a_1, a_2 \ge 0$ and

$$(I)_{\overline{1},k} + (I)_{\overline{2},k} = \sum_{t=1}^T \mathbb{E}\left[\mathbb{1}\{g_t(\overline{1}) = k, N_k(t) \le n_{\overline{1},k}\} + \mathbb{1}\{g_t(\overline{2}) = k, N_k(t) \le n_{\overline{2},k}\}\right] \le \max\{n_{\overline{1},k}, n_{\overline{2},k}\} = n_{\overline{2},k},$$

therefore we can bound $(I)_{\overline{2},k}$ as $(I)_{\overline{2},k} \le n_{\overline{2},k} - (I)_{\overline{1},k}$, yields that:

$$a_1 (I)_{\overline{1},k} + a_2 (I)_{\overline{2},k} \le (a_1 - a_2)(I)_{\overline{1},k} + \triangle_{\overline{2},k} n_{\overline{2},k}.$$

Then, since $a_1 \ge a_2$ and $(I)_{\overline{1},k} \le n_{\overline{1},k}$, we derive

$$a_1 (I)_{\overline{1},k} + a_2 (I)_{\overline{2},k} \le (a_1 - a_2) n_{\overline{1},k} + a_2 n_{\overline{2},k},$$

which shows Eq. (20). Now, assume for any $\{b_j\}_{j=1}^{\ell}$ with $b_1 \ge \cdots \ge b_\ell \ge 0$, the following

$$\sum_{j=1}^{\ell} b_j (I)_{\overline{j},k} \le b_1 n_{\overline{1},k} + \sum_{j=2}^{\ell} b_j (n_{\overline{j},k} - n_{\overline{j-1},k}) \tag{21}$$

holds for $\ell < d_k$.

Fix any $\{a_j\}_{j=1}^{\ell+1}$ with $a_1 \geq \cdots \geq a_{\ell+1} \geq 0$. Consider $\sum_{j=1}^{\ell+1} a_j (I)_{\overline{j},k}$. Since $a_j \geq 0$ for all $j \in [\ell+1]$ and

$$\sum_{j=1}^{\ell+1} (I)_{\overline{j},k} = \sum_{t=1}^{T} \mathbb{E}\left[\sum_{j=1}^{\ell+1} \mathbb{1}\{g_t(\overline{j}) = k, N_k(t) \leq n_{\overline{1},k}\}\right] \leq \max_{j\in[\ell+1]} n_{\overline{j},k} = n_{\overline{\ell+1},k},$$

we can bound $(I)_{\overline{\ell+1},k}$ as $(I)_{\overline{\ell+1},k} \leq n_{\overline{\ell+1},k} - \sum_{j=1}^{\ell} (I)_{\overline{j},k}$, which results in:

$$\sum_{j=1}^{\ell+1} a_j (I)_{\overline{j},k} \leq \sum_{j=1}^{\ell} (a_j - a_{\ell+1})(I)_{\overline{j},k} + a_{\ell+1} n_{\overline{\ell+1},k}.$$

Since $a_1 - a_{\ell+1} \geq \cdots \geq a_\ell - a_{\ell+1} \geq 0$, using inductive hypothesis Eq. (21) with $b_j = a_j - a_{\ell+1}$ for all $j \in [\ell]$, we get

$$\begin{aligned}
\sum_{j=1}^{\ell+1} a_j (I)_{\overline{j},k} &\leq (a_1 - a_{\ell+1}) n_{\overline{1},k} + \sum_{j=2}^{\ell} (a_j - a_{\ell+1})(n_{\overline{j},k} - n_{\overline{j-1},k}) + a_{\ell+1} n_{\overline{\ell+1},k}. \\
&= a_1 n_{\overline{1},k} + \sum_{j=2}^{\ell} a_j (n_{\overline{j},k} - n_{\overline{j-1},k}) - a_{\ell+1}\left(n_{\overline{1},k} + \sum_{j=2}^{\ell}(n_{\overline{j},k} - n_{\overline{j-1},k}) - n_{\overline{\ell+1},k}\right) \\
&= a_1 n_{\overline{1},k} + \sum_{j=2}^{\ell} a_j (n_{\overline{j},k} - n_{\overline{j-1},k}) + a_{\ell+1}(n_{\overline{\ell+1},k} - n_{\overline{\ell},k}) \\
&= a_1 n_{\overline{1},k} + \sum_{j=2}^{\ell+1} a_j (n_{\overline{j},k} - n_{\overline{j-1},k}).
\end{aligned}$$

Thus, Eq. (19) is proved by induction.

Define $\epsilon \triangleq \log^{-m} T$ and $\triangle_{j,k}(\epsilon) \triangleq \frac{\mu_j}{1+\epsilon} - \mu_k$. Using Eq. (19) with $a_j = \triangle_{\overline{j},k}$ for $j \in [d_k]$ and recalling $n_{j,k} \triangleq \max\left\{\frac{6(b-a)^2 \log T}{\triangle_{j,k}(\epsilon)^2}, T_0\right\}$, we have

$$\begin{aligned}
\sum_{j=1}^{d_k} \triangle_{\overline{j},k} (I)_{\overline{j},k} &\leq \triangle_{\overline{1},k} n_{\overline{1},k} + \sum_{j=2}^{d_k} \triangle_{\overline{j},k}(n_{\overline{j},k} - n_{\overline{j-1},k}) \\
&\leq \sum_{j=1}^{d_k} \triangle_{\overline{j},k} T_0 + 6(b-a)^2 \log T \left(\frac{\triangle_{\overline{1},k}}{\triangle_{\overline{1},k}(\epsilon)^2} + \sum_{j=2}^{d_k} \triangle_{\overline{j},k}\left(\frac{1}{\triangle_{\overline{j},k}(\epsilon)^2} - \frac{1}{\triangle_{\overline{j-1},k}(\epsilon)^2}\right)\right).
\end{aligned}$$

We upper bound the last term by:

$$\begin{aligned}
\frac{\triangle_{\overline{1},k}}{\triangle_{\overline{1},k}(\epsilon)^2} + \sum_{j=2}^{d_k} \triangle_{\overline{j},k}\left(\frac{1}{\triangle_{\overline{j},k}(\epsilon)^2} - \frac{1}{\triangle_{\overline{j-1},k}(\epsilon)^2}\right) &= \sum_{j=1}^{d_k-1} \frac{\triangle_{\overline{j},k}(\epsilon) - \triangle_{\overline{j+1},k}(\epsilon)}{(\triangle_{\overline{j},k}(\epsilon))^2} + \frac{\triangle_{\overline{d_k},k}}{\triangle_{\overline{d_k},k}(\epsilon)^2} \\
&\leq \sum_{j=1}^{d_k-1} \frac{\triangle_{\overline{j},k}(\epsilon) - \triangle_{\overline{j+1},k}(\epsilon)}{\triangle_{\overline{j},k}(\epsilon)\triangle_{\overline{j+1},k}(\epsilon)} + \frac{\triangle_{\overline{d_k},k}}{\triangle_{\overline{d_k},k}(\epsilon)^2} \\
&= \sum_{j=1}^{d_k-1}\left(\frac{1}{\triangle_{\overline{j+1},k}(\epsilon)} - \frac{1}{\triangle_{\overline{j},k}(\epsilon)}\right) + \frac{\triangle_{\overline{d_k},k}}{\triangle_{\overline{d_k},k}(\epsilon)^2} \leq \frac{2\triangle_{\overline{d_k},k}}{\triangle_{\overline{d_k},k}(\epsilon)^2},
\end{aligned}$$

where the first inequality is due to $\triangle_{\overline{j},k}(\epsilon) \geq \triangle_{\overline{j+1},k}(\epsilon)$, and the second upperbounds the telescoping series:

$$\sum_{j=1}^{d_k-1}\left(\frac{1}{\triangle_{\overline{j+1},k}(\epsilon)} - \frac{1}{\triangle_{\overline{j},k}(\epsilon)}\right) = \frac{1}{\triangle_{\overline{d_k},k}(\epsilon)} - \frac{1}{\triangle_{\overline{1},k}(\epsilon)} \leq \frac{1}{\triangle_{\overline{d_k},k}(\epsilon)}$$

Hence, we derive an upper bound for the first part relevant to $(I)_{\bar{j},k}$:

$$\sum_{j=1}^{d_k} \triangle_{\bar{j},k}(I)_{\bar{j},k} \leq \sum_{j=1}^{d_k} \triangle_{\bar{j},k} T_0 + \frac{12(b-a)^2 \triangle_{\overline{d_k},k} \log T}{\triangle_{\overline{d_k},k}(\epsilon)^2}.$$

□

Lemma 5.5. Let $k \notin$ supp $(\boldsymbol{i}^\star)$ and $j \in [d_k]$. For $T > T_0$,

$$(II)_{\bar{j},k} \leq \frac{1}{T} + \frac{\pi^2}{6}.$$

Proof of Lemma 5.5: Let $\epsilon = \frac{1}{\log^m T}$. Recall

$$(II)_{\bar{j},k} = \sum_{t=1}^{T} \mathbb{E}\left[\mathbb{1}\left\{g_t(\bar{j}) = k, N_k(t) > n_{\bar{j},k}\right\}\right],$$

where $n_{j,k} = \max\left\{\frac{6(b-a)^2 \log T}{(\frac{\mu_j}{1+\log^{-m} T} - \mu_k)^2}, T_0\right\}$.

First, we claim that

$$g_t(\bar{j}) = k \implies u_k(N_k(t-1), T) \geq \frac{\min_{s<t} u_{\bar{j}}(s,t)}{1+\epsilon}, \tag{14}$$

where $u_k(s,t) = \tilde{\mu}_k(s) + \sqrt{\frac{1.5(b-a)^2 \log t}{s}}$ and $\tilde{\mu}_k(t) = \frac{1}{t}\sum_{s=1}^{t} y_k(s)$.

Show Eq. (14): Observe that $g_t(\bar{j}) = k$ implies

$$\left(1 + \frac{\epsilon}{3}\right)\langle \boldsymbol{f}_k, \boldsymbol{q}\rangle \geq \langle \mathrm{dom}(\boldsymbol{f}_k), \boldsymbol{q}\rangle \geq \frac{\left\langle \mathrm{dom}(\boldsymbol{f}_{\bar{j}}), \boldsymbol{q}\right\rangle}{1+\frac{\epsilon}{3}} \geq \frac{\left\langle \boldsymbol{f}_{\bar{j}}, \boldsymbol{q}\right\rangle}{1+\frac{\epsilon}{3}},$$

where Eq. (8) is used in the first and the last inequality, and the second inequality is due to Lemma 5.2 and Corollary 4.10. By $(1+\frac{\epsilon}{3})^2 \leq 1+\epsilon$ and expanding $\boldsymbol{f}_k = (\hat{\mu}_k(t-1), \frac{1}{\sqrt{N_k(t-1)}})$ and $\boldsymbol{q} = (1, \sqrt{1.5(b-a)^2 \log t})$, we have

$$u_k(N_k(t-1), t) \geq \frac{u_{\bar{j}}(N_{\bar{j}}(t-1), t)}{1+\epsilon}.$$

As $\log T > \log t$ and $N_{\bar{j}}(t-1) \in [t-1]$, we further derive

$$u_k(N_k(t-1), T) \geq \frac{u_{\bar{j}}(N_{\bar{j}}(t-1), t)}{1+\epsilon} \geq \frac{\min_{s<t} u_{\bar{j}}(s,t)}{1+\epsilon},$$

which shows Eq. (14).

Second, let $\mathcal{T}_{\bar{j},k} = \{t \in \{n_{\bar{j},k}+1, \cdots, T\} : g_t(\bar{j}) = k, N_k(t-1) > n_{\bar{j},k}\}$. From Eq. (14), we derive

$$\begin{aligned}
(II)_{\bar{j},k} &= \sum_{t=n_{\bar{j},k}+1}^{T} \mathbb{P}\left[g_t(\bar{j}) = k, N_k(t-1) > n_{\bar{j},k}\right] \\
&\leq \sum_{t=n_{\bar{j},k}+1} \mathbb{P}\left[u_k(N_k(t-1), T) \geq \frac{\min_{s<t} u_{\bar{j}}(s,t)}{1+\epsilon} \text{ and } t \in \mathcal{T}_{\bar{j},k}\right] \\
&\leq \sum_{t=n_{\bar{j},k}+1} \sum_{s<t} \mathbb{P}\left[u_k(N_k(t-1), T) \geq \frac{u_{\bar{j}}(s,t)}{1+\epsilon} \text{ and } t \in \mathcal{T}_{\bar{j},k}\right],
\end{aligned} \tag{22}$$

where the last inequality uses union bound.

Third, we now upper bound each term $\mathbb{P}\left[u_k(N_k(t-1),T) \geq \frac{u_{\bar{j}}(s,t)}{1+\epsilon} \text{ and } t \in \mathcal{T}_{\bar{j},k}\right]$ in Eq. (22). Remind that

$$u_k(N_k(t-1),T) \geq \frac{u_{\bar{j}}(s,t)}{1+\epsilon} \iff \underbrace{\tilde{\mu}_k(N_k(t-1)) + \frac{\lambda_T}{\sqrt{N_k(t-1)}}}_{A_t} \geq \underbrace{\frac{\tilde{\mu}_{\bar{j}}(s) + \frac{\lambda_t}{\sqrt{s}}}{1+\epsilon}}_{B_{t,s}}.$$

Define the event $\mathcal{E}_{t,s} = \{A_t \geq B_{t,s} \text{ and } t \in \mathcal{T}_{\bar{j},k}\}$. We will partition the event $\mathcal{E}_{t,s}$ by comparing A_t to $A'_t = \mu_k + \frac{2\lambda_T}{\sqrt{N_k(t-1)}}$ and comparing $B_{t,s}$ to $B' = \frac{\mu_{\bar{j}}}{1+\epsilon}$ as follows:

- $\mathcal{E}_{t,s} \cap \left\{A_t \geq A'_t \text{ and } t \in \mathcal{T}_{\bar{j},k}\right\} \subseteq \left\{\tilde{\mu}_k(N_k(t-1)) \geq \mu_k + \frac{\lambda_T}{\sqrt{N_k(t-1)}} \text{ and } t \in \mathcal{T}_{\bar{j},k}\right\}$
- $\mathcal{E}_{t,s} \cap \{B_{t,s} \leq B' \text{ and } t \in \mathcal{T}_{\bar{j},k}\} \subseteq \left\{\mu_{\bar{j}} \geq \tilde{\mu}_{\bar{j}}(s) + \frac{\lambda_t}{\sqrt{s}} \text{ and } t \in \mathcal{T}_{\bar{j},k}\right\}$
- $\mathcal{E}_{t,s} \cap \{A_t < A'_t \text{ and } B_{t,s} > B' \text{ and } t \in \mathcal{T}_{\bar{j},k}\} \subseteq \left\{\mu_k + \frac{2\lambda_T}{\sqrt{N_k(t-1)}} > \frac{\mu_{\bar{j}}}{1+\epsilon} \text{ and } t \in \mathcal{T}_{\bar{j},k}\right\}$. The inclusion is because under the event $\mathcal{E}_{t,s} \cap \{A_t < A'_t \text{ and } B_{t,s} > B' \text{ and } t \in \mathcal{T}_{\bar{j},k}\}$, we have

  $$\mu_k + \frac{2\lambda_T}{\sqrt{N_k(t-1)}} = A'_t > A_t \geq B_{t,s} > B' = \frac{\mu_{\bar{j}}}{1+\epsilon},$$

  where the first and last inequalities are due to the event $\{A_t < A'_t \text{ and } B_{t,s} > B' \text{ and } t \in \mathcal{T}_{\bar{j},k}\}$, and the second inequality is due to the event $\mathcal{E}_{t,s} = \{A_t \geq B_{t,s} \text{ and } t \in \mathcal{T}_{\bar{j},k}\}$.

Hence, we have the following inclusion:

$$\begin{aligned}&\{A_t \geq A'_t \text{ and } t \in \mathcal{T}_{\bar{j},k}\} \cup \{B_{t,s} \leq B' \text{ and } t \in \mathcal{T}_{\bar{j},k}\} \cup \{A_t < A'_t \text{ and } B_{t,s} > B' \text{ and } t \in \mathcal{T}_{\bar{j},k}\}\\ &= \{t \in \mathcal{T}_{\bar{j},k}\} \supset \{A_t \geq B_{t,s} \text{ and } t \in \mathcal{T}_{\bar{j},k}\} = \mathcal{E}_{t,s}.\end{aligned}$$

From union bound,

$$\begin{aligned}\mathbb{P}[\mathcal{E}_{t,s}] &\leq \mathbb{P}\left[\{A_t \geq A'_t \text{ and } t \in \mathcal{T}_{\bar{j},k}\} \cap \mathcal{E}_{t,s}\right] + \mathbb{P}\left[\{B_{t,s} \leq B' \text{ and } t \in \mathcal{T}_{\bar{j},k}\} \cap \mathcal{E}_{t,s}\right]\\ &\quad + \mathbb{P}\left[\{A_t < A'_t \text{ and } B_{t,s} > B' \text{ and } t \in \mathcal{T}_{\bar{j},k}\} \cap \mathcal{E}_{t,s}\right]\\ &\leq \mathbb{P}\left[\mu_k + \frac{\lambda_T}{\sqrt{N_k(t-1)}} \leq \tilde{\mu}_k(N_k(t-1)) \text{ and } t \in \mathcal{T}_{\bar{j},k}\right]\\ &\quad + \mathbb{P}\left[\mu_{\bar{j}} \geq \tilde{\mu}_{\bar{j}}(s) + \frac{\lambda_t}{\sqrt{s}} \text{ and } t \in \mathcal{T}_{\bar{j},k}\right] + \mathbb{P}\left[\mu_k + \frac{2\lambda_T}{\sqrt{N_k(t-1)}} > \frac{\mu_{\bar{j}}}{1+\epsilon} \text{ and } t \in \mathcal{T}_{\bar{j},k}\right].\end{aligned} \tag{23}$$

In Eq. (23), recall $\lambda_t = \sqrt{1.5(b-a)^2 \log t}$ and observe that the last term

$$\mathbb{P}\left[\mu_k + 2\sqrt{\frac{1.5(b-a)^2 \log T}{N_k(t-1)}} \geq \frac{\mu_{\bar{j}}}{1+\epsilon} \text{ and } t \in \mathcal{T}_{\bar{j},k}\right] \leq \mathbb{P}\left[\mu_k + 2\sqrt{\frac{1.5(b-a)^2 \log T}{n_{\bar{j},k}+1}} \geq \frac{\mu_{\bar{j}}}{1+\epsilon}\right] = 0,$$

where the inequality is because $t \in \mathcal{T}_{\bar{j},k}$ implies $N_k(t-1) \geq n_{\bar{j},k} + 1$, and the equality is because

$$n_{\bar{j},k} \geq \frac{6(b-a)^2 \log T}{(\frac{\mu_{\bar{j}}}{1+\epsilon} - \mu_k)^2} \implies \frac{6(b-a)^2 \log T}{n_{\bar{j},k}+1} < \left(\frac{\mu_{\bar{j}}}{1+\epsilon} - \mu_k\right)^2$$

and also we have $\frac{\mu_{\bar{j}}}{1+\epsilon} - \mu_k > 0$ which is ensured by Lemma 5.3 as $T > T_0$. Finally, from Eq. (22) and Eq. (23),

$$
\begin{aligned}
(II)_{\bar{j},k} \le & \sum_{t=n_{\bar{j},k}+1}^{T} \sum_{s<t} \mathbb{P}\left[\tilde{\mu}_k(N_k(t-1)) \ge \mu_k + \sqrt{\frac{1.5(b-a)^2 \log T}{N_k(t-1)}} \text{ and } t \in \mathcal{T}_{\bar{j},k}\right] \\
& + \sum_{t=n_{\bar{j},k}+1}^{T} \sum_{s<t} \mathbb{P}\left[\mu_{\bar{j}} \ge \tilde{\mu}_{\bar{j}}(s) + \sqrt{\frac{1.5(b-a)^2 \log t}{s}} \text{ and } t \in \mathcal{T}_{\bar{j},k}\right] \\
\le & \sum_{t=n_{\bar{j},k}+1}^{T} \sum_{s<t} \left(\mathbb{P}\left[\tilde{\mu}_k(t-1) \ge \mu_k + \sqrt{\frac{1.5(b-a)^2 \log T}{t-1}}\right] + \mathbb{P}\left[\mu_{\bar{j}} \ge \tilde{\mu}_{\bar{j}}(s) + \sqrt{\frac{1.5(b-a)^2 \log t}{s}}\right]\right) \\
\le & \sum_{t=n_{\bar{j},k}+1}^{T} \sum_{s<t} \left(e^{-3\log T} + e^{-3\log t}\right),
\end{aligned}
$$

where the second inequality is because $\{N_k(t-1)\}_{t\in\mathcal{T}_{\bar{j},k}}$ is strictly increasing (as $N_k(t) = N_k(t-1)+1$ when $g_t(\bar{j}) = k$) and thus is a subsequence of $\{n_{\bar{j},k}+1, \cdots, T\}$, and the last inequality is due to an application of Hoeffding's inequality (Lemma E.2) with $s = \sqrt{1.5(t-1)(b-a)^2 \log T}$ and $n = t-1$ to bound the first term and with $s = \sqrt{1.5s(b-a)^2 \log t}$ and $n = s$ to bound the second term. The proof is completed by evaluating

$$
\begin{aligned}
\sum_{t=1}^{T}\sum_{s<t} e^{-3\log T} &\le \sum_{t=1}^{T} \frac{t}{T^3} \le \frac{T(T+1)}{2T^3} \le \frac{1}{T}, \\
\sum_{t=1}^{T}\sum_{s<t} e^{-3\log t} &\le \sum_{t=1}^{\infty} \frac{t}{t^3} \le \sum_{t=1}^{\infty} \frac{1}{t^2} \le \frac{\pi^2}{6}.
\end{aligned}
$$

□

Lemma E.2 (Hoeffding's inequality)**.** *Let $X_1, \cdots, X_n$ be independent random variables such that $X_i \in [a,b]$ for all $i \in [n]$. Then, for all $s > 0$,*

$$
\mathbb{P}\left[\sum_{i=1}^{n}(X_i - \mathbb{E}[X_i]) \ge s\right] \le \exp\left(-\frac{2s^2}{n(b-a)^2}\right).
$$